Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

THE MICROBIAL ECOLOGY OF *CAMPYLOBACTER JEJUNI* IN NEW ZEALAND WITHIN A SPATIAL-TEMPORAL FRAMEWORK

A thesis presented in partial fulfilment of the
requirements for the degree of
Doctor of Philosophy
in
Veterinary Science
at Massey University, Manawatu,
New Zealand

Barbara Mary Binney

2015

The journey is the reward.

Chinese proverb

Look deep into nature, and then you will understand everything better.

Albert Einstein

Abstract

Campylobacter jejuni (*C. jejuni*) is an important cause of gastroenteritis internationally; it is a complex bacterium carried by multiple hosts, showing phenotypic and genotypic variation. This thesis systematically examines the molecular ecology and evolution of *C. jejuni* in New Zealand from the levels of population movement, phenotype, genome and metabolism.

First, the demographic history of cattle, sheep and poultry importations into New Zealand (1860-1979) was quantified. Australia was the most common reported source of cattle sheep and poultry, with large numbers of cattle and sheep being imported in the 1860s, and large numbers of poultry imported from the 1960s onwards. This suggests the population structure of cattle and sheep and the microbial organisms they carried may exhibit a founder effect.

The second level investigated the phenotypes of related sequence types (ST) with generalist and specialist lifestyles and compared them at 42℃ and 22℃ on the basis of carbon source utilisation in Biolog phenotypic microarrays. The isolates utilised a total of 29 carbon sources in a pattern that clustered them together on the basis of ST at 42℃ more than lifestyle and host. At 22℃ they utilised a limited palette of carbon sources (9) related to the tricarboxylic acid cycle (TCA).

The third level, used genomic comparisons to identify a putative new species *C.* sp. nov. 4 spp. in the Australian purple swamphen (*Porphyrio porphyrio melanotus*). Overall, the pattern of relationship between isolates associated with the pukeko (*Porphyrio porphyrio melanotus*), takahe (*Porphyrio hochstetteri*) and the Australian swamphen isolates suggested a recent common ancestor and then divergence after separation. Despite high levels of recombination in *C.jejuni*, the genomes grouped by clonal complex and ST, this suggests there are factors restricting regular recombination between more distant *C. jejuni* STs. The draft genomes for the wild-bird and agricultural-related isolates clustered by lineages in a host(s).

The fourth level involved the comparison of *C. jejuni* metabolic pathways (subsystems) to identify host association. Type VI secretion system, Coenzyme A biosynthesis and *Campylobacter* spp. iron metabolism were identified as important pathways in distinguishing between wild-bird and livestock associated isolates.

We stand on the shoulders of giants.

---

used by Sir Isaac Newton

Acknowledgements

My Supervisors: Professor Nigel French, Dr Patrick Biggs, Associate Professor Barbara Holland, Dr Phil Carter, and Dr Jonathan Marshall.

The mEpiLab team: Angie Reynolds, Dr Anne Midwinter, Dr Julie Collins-Emmerson, Anthony Pita, Neville Hack, Sarah Moore, Rukhshana Nahid Akhter and Lynn Rogers.

Helpers from across the Tasman: Dr Mary Valcanis of the Microbiological Diagnostic Unit (MDU) at Melbourne University and Dr Carl Mulnickel of the ASAP Lab. The volunteers who collected samples: Will Steele, Katrina Lee, Christine Hurring, Sandra Martig, Dorothy Dhaeze,Michael Tarburton, Tiffany Miller-Scott, Peter Dann and James O'Connor of Birdlife Australia.

Massey University Support Staff: Jacque Mackenzie, Wendy Maharey, Christine Cunningham and Debbie Hill, and Noelene White.

Software and IT help: Simon Verschaffelt, Xavier Didelot (ClonalFrame), Marco Galardini (Ductape), Bruno Contreras Moreira (Get Homologues), Markus Goker (opm), Korbinian Strimmer (binda), Vladimir Minin (rbrothers), Michael Woodhams (ADD), Sajia Akhter (Phispy) and Maulik Shukla (PATRIC).

Friends and family: I wish to thank the postgrads of the Hopkirk for their help and comradeship especially Anja Friedrich, Hamid Ishad, Fang Fang and Patricia Jaros. The TTP team including Lesley Stringer, Debbie Leader, Jackie Benshop, Chris Jewell, Becky Frost and many more. Edward George Binney for his warmth and companionship that enriched my life.

Funding: This research was made possible by funding from Massey University, the Marsden fund, and IVABs funding.

I have met and been inspired by many people in my life, which has helped me to and through my journey, too many to list here, so I apologise to those I have not directly mentioned.

Contents

List of Figures

List of Tables

Abbreviations

AFLP: Amplified fragment length polymorphism

AMOVA: Analysis of molecular variance

AU: Australia

blood agar: Columbia horse blood agar

bp: base pair

BIGSdb:Bacterial Isolate Genome Sequence Database.

BRENDA: Bacterial restriction endonuclease DNA analysis

contigs: contiguous reads of DNA sequence

CPS: capsular polysaccharides

CDC: Centers for Disease Control, USA

cfu: colony forming unit

CC: Clonal complex

CRISPR: Clustered regularly interspaced short palindromic repeats

DNA: Deoxyribonucleic acid

D.O.C: New Zealand Department of Conservation

ELISA: Enzyme-linked immunosorbent assay

ESCMID: European Society of Clinical Microbiology and Infectious Diseases

ESGEM: Study Group on Epidemiological Markers

ERIC: Enterobacterial repetitive intergenic consensus

ESR: Environmental Science and Research Ltd.

EU: European Union

FAME: Fatty acid methyl ester analysis

FAFLP: fluorescent AFLP

G+C%: guanine–cytosine content as a percent of nitrogenous bases in a DNA molecule

gDNA: genomic Deoxyribonucleic acid.

HGT: Horizontal gene transfer

HL: Heat labile antigens

hr: hour

HS: Heat stable antigens

IF0a: Omnilog inoculating fluid IF-0a GN/GP

kbp: Kilobase pairs (of DNA)

LOS: lipooligosaccharide mCCDA: Modified Cefoperazone Charcoal Desoxycholate agar

MLEE: Multilocus enzyme electrophoresis

MLST: Multilocus sequence typing

mrp: Macrorestriction profiling

MYA: million years ago

NCTC: National Collection of Type Cultures (UK)

nt: Nucleotides

NGS: Next generation sequencing

NZFSA: New Zealand Food Safety Authority

OIE: World Organisation for Animal Health

Omnilog bags: Biolog Gas Bag (P/N 3032)

PBS: Phosphate Buffer solution

PCR: Polymerase Chain Reaction

peg: protein encoding gene, plural pegs

PFGE: Pulsed Field Gel Electrophoresis

PM plates: Biolog Phenotypic Microarray plates.

PM1: Biolog Phenotypic Microarray plate 1.

PM2A: Biolog Phenotypic Microarray plates 2A.

16S rRNA: 16S ribosomal

RAPD: Randomly amplified polymorphic DNA

RAPD-PCR: Random amplified polymorphic DNA Polymerase Chain Reaction

rDNA: Ribosomal DNA

RE: Restriction enzyme

REA: Restriction endonuclease analysis

RFLP: Restriction fragment length polymorphism

rMLST: ribosomal Multilocus Sequence Typing

RNA: Ribonucleic acid

RNAT: RNA thermometers

rps: ribosomal protein subunit

SD: Standard deviation

SNP: single nucleotide polymorphism

ST: Sequence type, plural STs

sp: Species (single)

spp.: Species (multiple)

subsp.: subspecies

SVR: short variable repeats

VBNC: Viable but Nonculturable State

wgMLST: whole genome MLST

WGS: Whole genome sequencing

WHO: World Health Organisation

WTO: World Trade Organisation

This is dedicated to my parents, Ronald Samuel Binney and Noela Alma Binney (nee Jones), who did not live to see this part of my journey, I love you and I am grateful to have been your daughter. It is because of the greatness within you both that I could live this adventure.

This hierarchy of successive integrations, characterised by restrictions and the appearance of new properties at each level, has some consequences. The first is the necessity of analysing complex objects at all levels... The second point concerns predictability. It is possible to make predictions at one level on the basis of what is known at a simpler one? Only to a very limited extent. The properties of a system can be explained by the properties of its components. They cannot be deduced from them.

Francois Jacob 1977

1

Introduction

Understanding the epidemiology and evolution of *Campylobacter* spp. is an important modern public health issue [1]. *Campylobacter jejuni* (*C. jejuni*) is a bacteria that is carried in the gastrointestinal flora of a wide range of warm blooded animals, domesticated and wild. Since the 1970s there has been a growing recognition of the role *C. jejuni* plays in human disease, mainly as a cause of gastroenteritis but also as a contributing factor in complex sequelae like Guillain-Barre syndrome. It is important to understand *C. jejuni* so that measures can be taken to manage the risk to the public. As a multi-host pathogen, that is fastidious to grow in a laboratory, with a tendency to phenotypic and genotypic variation, it has proven to be enigmatic.

New Zealand's geographical isolation has resulted in the development of many endemic species, some of which act as hosts to *C. jejuni*. This segregation was interrupted in the last 1,000 years with the arrival first of the Polynesians, then later European settlers who introduced many new species including pastoral livestock. Modern New Zealand developed often by displacing an ancient fauna and flora with imported ones, the full consequences of these actions and the resulting ecological change for the microbial communities can only be guessed at. From the 1980s to 2006 there were rising levels of reported campylobacteriosis in New Zealand [2]. It was notably associated with a sequence type (ST-474) found in the poultry industry [3] but ST-474 was only reported once

outside of New Zealand. Against this background, questions were raised regarding the evolution and molecular ecology of *C. jejuni* in New Zealand.

Overall, this thesis addresses the two big questions of "Has *C. jejuni* in New Zealand followed own unique trajectory?", and "What is the relationship between phenotype, genotype and host in *C. jejuni*?". This thas been approached as a systematic investigation of the molecular ecology and evolution of *Campylobacter jejuni* (*C. jejuni*) in New Zealand. The approach presented here was to examine *C. jejuni* from the perspective of multiple levels, each of which form a chapter. The first level analyses the demographic history of host population movements in a chapter titled The Past. The second level was an examination of phenotypic variation in a chapter titled The Phenotype. The third level was an investigation of variation at the level of the genome in a chapter titled The Genome. The fourth level looks into variation in the metabolic pathways in a chapter titled The Metabolism. The common thread that was generally running through these levels of investigation was a group of *C. jejuni* sequence types (ST-2381 lineage), frequently found in New Zealand environmental waters, and their recently found hosts, two New Zealand water rails, the pūkeko (*Porphyrio porphyrio melanotus*) and the takahē (*Porphyrio hochstetteri*). This group was compared at the level of phenotype and genome with isolates from the more investigated livestock associated groups (e.g. clonal complexes CC21, CC42, CC45)

The Past: Historical demography of livestock importation into New Zealand New Zealand was an isolated archipelago before the arrival of the Polynesians ∼800 years ago, with birds forming the dominant terrestrial class. The arrival of mankind, especially Europeans ∼200 years ago caused significant changes to the land and the ecology. The Europeans settlers brought pastoral agriculture, as they colonised New Zealand introducing livestock such as cattle, sheep and poultry. These animals act as hosts of *C. jejuni*, and may have introduced sequence types that had not previously been present in New Zealand. New Zealand's isolation, if it was facilitating a unique evolutionary trajectory for *C. jejuni*, could be greatly affected by the arrival of new host species and their associated sequence types. In this chapter the official records were used to quantify the historical demography of live animals imported into New Zealand, in particular cattle, sheep and poultry. An estimation of the quantity and source of livestock introduced into New Zealand has implications not only for *C. jejuni*, it may also provides a better understanding of why many infectious livestock diseases found elsewhere in the world are not present in New Zealand.

The Phenotype: Phenotypic comparison of isolates associated with wild-bird and livestock. *C. jejuni* is known for its phenotypic and genotypic variation [4] [5], as well as its ability to exist in a wide range of hosts [6]. Biolog phenotypic microarrays were used to investigate the question of "Is there a host-associated basis for the phenotypic variation?". Agriculturally-associated sequence types of *C. jejuni* generally colonise a wider range of hosts than the more

host-specific wild-bird sequence types [7] [6]. Often species that are able to inhabit a wider environment range i.e. generalists, can also utilise a wider range of resources and in this case we will test nutrient range in the form of carbon sources.

Eleven New Zealand isolates related by shared MLST alleles were compared for their ability to respire in media containing a wide range of single carbon sources. The sequence types (ST) are associated with different host species; ST-42 isolates are found internationally and associated with agricultural animals, while ST-2381, ST-3655, and ST-3663 have been found in the pūkeko (*Porphyrio porphyrio melanotus*) and the takahē (*Porphyrio hochstetteri*). ST-3673 is a rare sequence type found once in a chicken and in New Zealand water, and ST-3845 has also been found in New Zealand water but not in a live host. These are novel ST that are commonly found in New Zealand environmental water [8] [9]. The ability to utilise a single carbon source was tested at two temperatures 42℃, a temperature associated with host body temperature when it can multiply, and 22℃, a temperature reflective of the environment that *C. jejuni* experiences outside the host and at which it does not multiply. The behaviour of *C. jejuni* at temperatures like 22℃ is an important public health question, because despite being fastidious it survives at this and even lower temperatures as it survives in the environment outside of the host [10]. While the behaviour at 42℃, a temperature similar to host body temperature, allows a comparison of phenotypic behaviour at two different and important temperatures.

The Genome: Characterisation and comparison of *Campylobacter* spp. isolates from Australian purple swamphen (*Porphyrio porphyrio melanotus*) with a New Zealand dataset. This chapter pursued two aims, the first was investigating if the geographical and temporal separation of the Australian purple swamphen (*Porphyrio porphyrio melanotus*) from the New Zealand sub population of the pūkeko (*Porphyrio porphyrio melanotus*) and the takahē (*Porphyrio hochstetteri*) resulted in evolutionary divergence of the *C. jejuni* they host. The second aim, was to use comparative genomics from a range of New Zealand *Campylobacter* spp. isolates to compare agricultural-associated and wild-bird associated *C. jejuni* genomes, thus investigating the genome for a "host signature" that could differentiate those ST from livestock to those found in wild-bird.

First the draft genomes of *Campylobacter* spp. isolates from the Australian purple swamphen (*Porphyrio porphyrio melanotus*) were characterise. Then genomic comparisons were made to identify their relationship to isolates from the pūkeko (*Porphyrio porphyrio melanotus*) and the takahē (*Porphyrio hochstetteri*). *Porphyrio porphyrio melanotus* arrived in Australasia $\sim 600,000$ years ago[11], but the pūkeko sub-population was a recent arrival to New Zealand, ~ 500 years ago [11] [12]. The South island takahē (*Porphyrio hochstetteri*) is closely related to both birds [13] [14]. The pūkeko and the takahē are both associated with a novel lineage of *C. jejuni* water isolates [15] and some newly identified *Campylobacter* spp. [16]. Genomic comparison of the Australian purple

swamphen *Campylobacter* spp. isolates will show if the temporal and geographical separation of the host populations has lead to the New Zealand *Campylobacter* spp. following a unique evolutionary trajectory. Griekspoor et al. 2013 used genotyping of *C. jejuni* from to European and Australian populations of blackbirds (*Turdus merula*) that had been separated for $\sim$ 200 years and found the host association persisted despite the separation [7], this analysis will use a different wild-bird family, a longer separation, and investigate genomic differences.

The genomic comparisons included a range of New Zealand *Campylobacter* spp. isolates associated with both wild-birds and agricultural livestock hosts using the core genome and the pan-genome. This comparison included the Australian purple swamphen, pūkeko and takahē associated isolates, and put their relationship in context with a larger range of *C. jejuni* ST and *Campylobacter* spp.. Most research particularly from the genomic perspective has predominantly used clinical isolates (from people) or from livestock which are the most common cause of clinical cases [17], this analysis provides a wider view of the relationships between *Campylobacter* spp. in New Zealand.

The Metabolism: Genomic comparison of metabolic subsystems present in *Campylobacter* spp. dataset. This chapter is a continuation of the previous genomic comparison chapter and by focusing on finding variation in those *C. jejuni* metabolic pathways found in wild-bird isolates compared to isolates associated with agriculture, it may be possible to identify features related to the adaptation to a given host. In general *C. jejuni* biotyping techniques identify physiological variation in bacteria but suffer from poor reproducibility and discriminatory power due to variation of gene expression caused by environmental factors and lack of biochemical diversity between isolates [18] [19], however by using a genomic approach we can investigate metabolic pathways and their variation without the influence of factors that affect expression. Almost all previous genomic analysis of *C. jejuni* is based on clinical isolates from people or agricultural associated isolates, there is a limited amount of reported genomic analysis into those sequence types associated with wild-birds [17]. By comparing isolates associated with two different wild-bird species, the starling (*Sturnus vulgaris*) and the Australian purple swamphen with a group of isolates from clonal complexes associated with agriculture, significant differences in metabolism between the groups were identified. This will help us better understand the ability of *C. jejuni* to adapt to a host by identifying which metabolic systems are involved and how they change. Examination of the differences between Australian purple swamphen and pūkeko /takahē associated isolates was also performed at the level of the subsystem. A greater understanding of which metabolic features are flexible and which are not, may identify pathways and reactions that may be susceptible to pharmacological manipulation and thus control of *C. jejuni* .

Structure of the thesis Each chapter is set out in the style of the journal to which it has been submitted or to which it is in the process of being submitted i.e. it is written as a standalone

article rather than a thesis chapter. Consequently, there is some repetition, particularly in the Methods section, and there are stylistic differences between chapters. The submitted manuscripts include my supervisors as co-authors. Currently the only article in-print is the basis of the chapter addressing historical demography: Binney, B., Biggs, P. J., Carter, P. E., Holland, B. M., & French, N. P. (2014). Quantification of historical livestock importation into New Zealand 1860–1979. New Zealand Veterinary Journal, 62(6), 309-314. doi:10.1080/00480169.2014.914861

Isolates used in this thesis A large number of isolates were used in this thesis. The isolates used in the phenotypic analysis found in chapter 4 are identified in Table 4.1. The isolates used in the genomic analysis found in chapter 5 are identified in Tables 5.1, 5.3, 5.4, 5.5. The isolates used in the metabolic analysis found in chapter 6 are identified in Table 6.1 and are a subset of the isolates in Table 5.3.

It is a riddle, wrapped in a mystery, inside an enigma

Winston Churchill referring to Russia, 1939.

2

Literature Review

2.1 Taxonomy

The genus *Campylobacter* has been classified as :
Domain: Bacteria
Phylum: Proteobacteria
Class: Epsilonproteobacteria
Order: Campylobacterales
Family: Campylobacteraceae
Genus: *Campylobacter*

At present (23rd March 2015)[1] there are 33 recognised members of the genus *Campylobacter* and 14 subspecies, forming part of the Campylobacteraceae family, which in turn is part of the epsilonproteobacteria class. The family Campylobactereaceae contains the *Campylobacter, Dehalaospirillum, Sulfurospirillum* and *Arcobacter* genera. The Order Campylobacterales contains the Campylobactereaceae and the family Helicobacteraceae which includes the gastric carcinogen *Helicobacter pylori*.

[1] List of prokaryotes with standing nomenclature http://www.bacterio.cict.fr/campylobacter.html

The epsilonproteobacteria is a diverse range of bacteria from some extreme environments: *C. jejuni* inhabits the gastrointestinal tract of birds and mammals while *Hydrogenimonas thermophila* is a chemoautotrophic endosymbiote in the gills of a gastropod (*Alviniconcha aff. hessleri*) that inhabits geothermal vents in the deep sea [20] [21].

Initially *Campylobacter* was not recognised as a separate genus and it was grouped with the bacteria causing cholera as part of the *Vibrio* genus. Being both fastidious and relatively non-reactive biochemically it made the taxonomy of *Campylobacter* difficult to elucidate until more modern molecular based techniques became available. Technical progress in the 1960s enabled reclassification of some members of the Vibrio grouping into the newly formed *Campylobacter* genus based on biochemical, guanine–cytosine content (G+C%) and serological profiles [22]. In 1963, *Vibrio jejuni* was changed to *C. jejuni* but it was not until 1973 that other members such as *Campylobacter fetus* were named as part of the genus [22] [23].

Campylobacteraceae are Gram negative, non-spore forming spiral shaped bacteria. *C. jejuni* are visible under magnification as spirally curved rods, and were eventually named "*Campylobacter*" meaning "curved rod" in Greek. Although under certain conditions such as in response to some stresses, *Campylobacter* are known to change to a coccoid shape [24] usually accompanied by a transformation into a viable but non-culturable state (VBNC) [25]. The genus *Campylobacter* grow under microaerobic conditions (mostly less than 10% oxygen) [26], with many having a characteristic corkscrew motion. While most *Campylobacter* spp. can be grown between $\sim 30 - 37$℃ the optimal temperature differs between species. *C. jejuni*, *Campylobacter coli* (*C. coli*) and *Campylobacter lari*(*C. lari*) are called thermophilic *Campylobacter* as they have an optimum growth at $42 - 43$℃ and do not grow below 28℃ [27]. *C. jejuni* and *C. coli* are similar biochemically except *C. jejuni* is the only *Campylobacter* species that hydrolyses hippurate although *C. jejuni subsp. doylei* varies in its ability to hydrolyse hippurate [28] [29].

2.2 Medical history

A *Campylobacter* spp., was reported by Butzler [30], to have been first observed by Escherich in 1886 [31], however it was not fully recognised as a human pathogen until much later, being first identified as a pathogen in veterinary medicine. Campylobacteriosis in the form of *Campylobacter fetus fetus* (initially called *Vibrio foetus*) was first identified and reported in 1913 as a cause of epizootic abortions in U.K. sheep by McFadyean and Stockman [32] according to Veron and Chatelain [23], although Hindmarsh and Skirrow say McFadyean first reported in 1905 to the British government the association of "spirillae" with abortion of sheep and cattle [33] [34]. Later work by Smith and Taylor in the USA identified a *Campylobacter* spp. causing abortions in cattle in 1919 [35]. Stegenga and Tersptra identified an enzootic sterility in cows caused by a bacteria

they named *Vibrio fetus venerealis* and later renamed *Campylobacter fetus venerealis* [36]. A vibrio-like organism was found in some cattle diarrhoea in 1927 [37], but it was not until 1931 *Vibrio jejuni* (now *Campylobacter jejuni*) was attributed as the cause [38]. In 1944 a vibrio was isolated from the faeces of pigs with diarrhoea, the resulting bacteria was called *Vibrio coli* (now *Campylobacter coli*) [39].

The reporting of veterinary campylobacteriosis occurred later in Australasia, presumably due to lack of diagnosis rather than lack of presence. In 1942 Hindmarsh recorded the first recognised case of "bovine vibrionic abortion" in Australasia [34]. While McFarlane et al. (1952) were the first in New Zealand to report *Campylobacter fetus* subsp. *fetus* after investigating sheep abortions in the 1950 lambing season [40].

C. jejuni did not start being regularly diagnosed as a cause of human gastroenteritis until the 1970s [41] [30], although Levy had earlier reported a vibrio-like organism causing a milk borne outbreak of gastroenteritis in 1946 [42]. Skirrow suggested that the difficulty for most medical laboratories in growing and isolating *Campylobacter* spp. was a major factor in delaying greater recognition as a cause of gastroenteritis in people [43] [27]. In 1977, a selective media made from blood-agar supplemented with the antibiotics vancomycin, polymyxin B, and trimethoprim, was developed to selectively grow *C. jejuni* and *C. coli* [43]. Once microaerophilic conditions and suitable selective media were widely adopted to detect *C. jejuni*, it began emerging as a common cause of human gastroenteritis [27].

2.2.1 Non-*C. jejuni* *Campylobacter* spp.

About 95% of campylobacteriosis is attributed to *C. jejuni* and *C. coli* [44][p123] however there is increasing awareness of the disease potential of other thermophillic *Campylobacter* spp.. *Campylobacter fetus* subsp. *fetus* was recognised in 1913 as a causative agent in sheep fetal abortions [32] and cattle [35]. It can cause diarrhoea, particularly in compromised patients, with a reported annual incidence of 0.1 per 100,000 and it can also cause a bacteraemia [44][p124]. *Campylobacter upsaliensis*(*C. upsaliensis)*, can also cause bacteraemia and diarrhoea but not just in compromised patients. It is thermotolerant and commonly isolated from dogs and cats, although human to human transmission is thought to be a possible source of infection. *Campylobacter hyointestinalis* (*C. hyointestinalis*) was first identified in 1983 as a cause of proliferative enteritis in pigs, but has been isolated from other asymptomatic animals. There are a limited number of human cases reported, usually with diarrhoea [44][p132]. *Campylobacter lari*(*C. lari*) was first found in 1980 in seagulls, since then it has been found in some wild birds and water [44][p133]. Some isolates have the unusual ability for a *Campylobacter* spp. to hydrolyse urea It has been associated with diarrhoeal disease and bacteraemia in both immunocompetent and immunocompromised patients. *Campylobacter sputorum*(*C. sputorum*) forms part of human gingival flora, but has ocassionally been isolated from other parts of the body [44][p133].

There are six known hydrogen requiring *Campylobacter* spp.: *Campylobacter rectus*(*C. rectus*), *Campylobacter curvus*(*C. curvus*), *Campylobacter concisus*(*C. concisus*), *Campylobacter gracilis*(*C. gracilis*), *Campylobacter showae*(*C. showae*), *Campylobacter mucosalis*(*C. mucosalis*) [44][p135]. *C. mucosalis* is associated with proliferative enteritis in pigs. All the rest have been found in human gingival flora associated with diseased gums but have been found in other parts of the body and potentially associated with disease.

2.3 Epidemiology of *C. jejuni*

2.3.1 Clinical Signs in People

Clinically *Campylobacter* enteritis cannot be distinguished from other acute diarrhoeal disease without diagnostic tests, nor is there a clear difference in infections by *C. jejuni* or *C. coli* [45][p99]. A definitive diagnosis requires detection of *Campylobacter* in the faeces, although a recent epidemiological study using multivariate analysis identified a difference between the two *Campylobacter* spp. using risk of infection in France associated with age, travel and season [46]. The mean incubation period is 3.2 days (range eighteen hours to eight days) [45][p101]. The acute inflammatory reaction frequently extends from the small intestine to the colon and rectum. Clinical signs show variation possibly including: acute diarrhoea, fever, chills, myalgia, abdominal cramps with the bowel motions varying from loose and watery to frankly bloody [47][p2].

A late onset complication of *Campylobacter* enteritis is Guillain-Barre Syndrome (GBS) [48]. GBS was first described in 1916 and is an acute autoimmune-mediated disorder of the peripheral nervous system, and believed to be produced by carbohydrate mimicry by a *C. jejuni* lipooligosaccharide ($Gal\beta1-3GalNAc\beta1-4(NeuAc\alpha2-3)Gal\beta1$) of a human GM1 ganglioside [49]. The percentage of cases of campylobacteriosis which result in GBS has been estimated as 0.1% [50]. Miller-Fisher, is a rare variant of GBS affecting the ocular system [51] and is thought to be due to molecular mimicry between GQ1b and the lipooligosaccharides of *C. jejuni* and *Haemophilus influenzae* [52]. Other post infection complications associated with *Campylobacter* enteritis and other bacterial enteritis include a reactive arthritis reported in 1–5% of the cases [53] [50] and possibly a post infection irritable bowel syndrome [54].

2.3.2 Seasonality in Campylobacteriosis

Many countries show a consistent seasonal pattern in the occurrence of campylobacteriosis, although the mechanisms for this are not clearly understood, and it can vary among countries. Temperate zone countries, tend to experience an increase in spring leading to a peak in summer and then a gradual decline [55] [56] [57] [58]. In European countries this effect was more pronounced with increasing latitude [55]. Temperate regions in Australia show a different pattern of incidence to

the subtropical regions in Australia, with subtropical Queensland showing only a modest increase in the warmer months [56] [59].

New Zealand as a whole, shows a pronounced summer peak with *C. jejuni* cases (Figure 2.1) occurring over a longer period than in European countries [55]. In a multinational analysis using population structure data a summer peak was formed by a clade of ST-45 and ST-283 in the Finnish data, but this was absent in the New Zealand and Australian data [60]. However recent work has identified a summer peak formed by the ST-45 clonal complex in New Zealand [61] It has also been observed along the length (North–South) of New Zealand that the degree of seasonal incidence shows variation with geography [62]. All these studies combine to suggest there may be a range of geographical factors affecting the epidemiology of this disease.

Climate related factors such as temperature, sunlight and humidity may affect the seasonality, by affecting the prevalence in reservoirs, survival in the environment and/or human behaviour affecting exposure. In laboratory experiments, natural populations of culturable thermo-tolerant campylobacters, in sea and river water, decreased within less than 10 minutes exposure to simulated summer sunlight, and none could be cultured after 30 minutes [63] [64]. This effect of sunlight would imply less risk from environmental sources like water in the summer when there is a peak of cases. Two peak times were reported in the isolation of *C. jejuni* from the Taieri River, New Zealand, one peak in summer and one in winter [65]. In New Zealand and overseas, there is a seasonal pattern of incidence in rivers and environmental water showing a higher incidence in the winter and declining into the summer [66] [67].

Human behaviour can be affected by season and this could affect the seasonality and the sources of infection: for example overseas travel and vacations can also have an impact with different sequence types of *Campylobacter* seen in the returning travellers [68]. In Minnesota, USA winter spikes of fluoroquinolone-resistant strains have been related to travel to Latin America where fluoroquinolone is used in poultry production [69] and South Korea has also reported unusual antibiotic resistance patterns in isolates that have been related to international travel [68].

2.3.3 Dose and Immunity

Many factors can influence the infective dose, such as prior exposure but in general the infective dose is low, with evidence for as few as 500 bacterial cells resulting in an experimental infection, but exposure does not always result in clinical signs [70] [71]. A large variation in dose response has been reported and was attributed to isolate variation [72] [73]. Variation in response by volunteers in studies suggests short term immunity to homologous strains, but it is not known how wide or how long the resulting immunity lasts [70] [74]. Immunodeficient patients, such as those with hypoglobulinemia and acquired immune deficiency syndrome (AIDs), are associated with more severe clinical symptoms, and conditions such as chronic carriage and possibly bacteraemia [75] [76]. Protective immunity has been shown in people chronically exposed to risky foods such as raw

(unpasteurised) milk [77] [78]. In developing countries where repeated infections during childhood are common, the infection rates and intensity of symptoms decline with age [45][p100]. It has been suggested that after repeated exposure people may become asymptomatic carriers, but there is very limited support for this at the moment [79][p182] [80].

2.3.4 Age and sex

In many developed countries the age distribution has a bimodal peak with higher reported incidence in children under five and a second modest peak between 15–44 years [81] [82]. Other enteric diseases also show a peak in younger children which in part is attributed to behaviours of greater health-seeking and sample taking for this age group [83][p170]. The second peak, people between 15 and 44 years of age, has been attributed to the greater likelihood of young adults to travel [84]. The sex distribution is also similar in developed countries with boys and men experiencing an incidence rate 1.1–1.5 times higher than girls and women [82].

2.3.5 Epidemic patterns

Campylobacteriosis presents most frequently as sporadic infections ($\sim$ 99%) [85][p253], making epidemiological analysis more difficult [86]. It is uncommon to associate *Campylobacter* with common-source outbreaks, unlike some other food pathogens like *Salmonella* because it does not multiply on food at room temperature [87]. The low infection dose, ease of cross-contamination, multiple subtypes contaminating food product, and relatively long incubation can obscure a cluster of related illness [88] [89]. However common-source outbreaks can occur and have been related to food and water, with raw milk and poultry being common sources [90] [91] [92] [15]. In New Zealand *Campylobacter* was identified in approximately 12 - 16% of reported outbreaks compared to 1% in the USA (1980–1996) [93][p16]. Secondary transmission (person to person) is rare in comparison with other enteric infections [94].

2.3.6 Worldwide prevalence

Campylobacter is the most commonly reported cause of bacterial gastroenteritis in developed countries, with *C. jejuni* estimated to infect 1% of the EU population each year; the European Food Safety Authority (EFSA) had 214,268 confirmed human cases (with 31 deaths) in 2012 [95]. 1.3 million cases per year have been estimated in the USA by Centers for Disease Control and Prevention (CDC) [95]. In an Australian retrospective cross sectional study of 1722 patients with gastroenteritis, the commonest cause was *Campylobacter* at 22%. The World Health Organisation say the true incidence of this disease is poorly known, particularly in low to middle income nations but studies in high-income countries have estimated the annual incidence at between 4.4 – 9.3 per 1000 population [96].

2.3.7 An emerging disease in New Zealand

In 1980 campylobacteriosis in people became a legally notifiable disease in New Zealand [97], which resulted in providing the public health authorities with data on the trend of increasing incidence which peaked in 2006 at 383.5 per 100,000 [2]. Molecular epidemiological investigations and source attribution modelling identified the poultry industry as a major source, and the response to new regulations has been a major decline in national levels of campylobacteriosis (Figure 2.1), although compared to the national incidence levels world-wide it is still high [98] [99]. Prior to the regulatory intervention into the New Zealand poultry industry over 70% of human cases were attributable to poultry [97] [100], afterwards (2008–2010) the estimate of source attribution to poultry has declined to less than 50% [101]. Understanding the changing epidemiology requires continued investigation of host source, the strains they carry, and identifying the most informative model so suitable steps can be taken to improve public health [102].

Most developed countries have had some form of surveillance of campylobacteriosis in the human population for the last 25 years which show a trend of increasing reported cases [83][p169]. It is difficult however to compare infection rates of *Campylobacter* enteritis between countries because they are affected by differences in surveillance levels, cultural norms for health-care seeking and diagnosis as well as immune status [82]. Both Iceland (1990s) [103] and Denmark (1992–2001) [104] experienced increasing levels of reported *Campylobacter* enteritis and responded to the situation with multi-pronged public health campaigns involving improved poultry regulation, which resulted in a decline in incidence levels [83][p169]. Both of these countries used techniques like the reduction of *Campylobacter* counts by freezing poultry carcases [105] and knowledge of *Campylobacter* status of farms.

2.3.8 Clinical signs in animals

This review will emphasise disease in animals directly attributable to *C. jejuni* and not include those diseases attributed to the genus *Campylobacter* generally or other pathogenic species, such as *Campylobacter fetus* which causes abortions in cattle and sheep.

Although many animals can carry strains of *C. jejuni* as a commensal in the gastrointestinal tract, it is not always an avirulent part of the microbial flora. Natural infections with *C. jejuni* causing enteritis have been reported in young macaques, young ferrets and minks, cats, and swine while experimentally infected chickens, rodents, ferrets, primates (e.g. *Aotus nancymae* [106]), rabbits and pigs have also shown clinical signs [47][p5] [107] . There is also experimental evidence of abortion *C. jejuni* induced abortions in minks and guinea pigs [107].

Cats and dogs are considered a source of *C. jejuni* for people and it is difficult to attribute *C. jejuni* as the cause of an infection in these animals when it could be normal carriage. However there

[2] ESR Annual Surveillance Summary https://surv.esr.cri.nz/surveillance/annual_surveillance.php

are multiple reports of *C. jejuni* causing diarrhoea, varying in intensity from watery to bloody with mucus, and mainly seen in the young [47][p5]. The reported cases of *C. jejuni* in dogs either experimentally induced or naturally occurring are usually transient and mild, but there are a few reports of death from campylobacter-associated haemorrhagic enteritis [107]. *C. jejuni* has also been isolated from the vaginal swabs of three dog having late pregnancy abortions [108]. In the case of cats there are occasional accounts of *C. jejuni*-associated acute diarrhoea in kittens, while it has proved difficult to establish the presence of clinical signs in cats the few times it has been tried experimentally [107].

Cattle and sheep are also considered to "silently" carry and to be a source of *C. jejuni* for people. There are mixed reports on *C. jejuni*'s ability to cause diarrhoea in cattle and sheep; an early report by Jones et al. (1931) associated experimentally infected calves with diarrhoea [38]. While transient diarrhoea and sometime dysentry has been reported in calves[109], other reports of similar prevalence in cattle with and without diarrhoea weakens the strength of this association [47][p6]. Bovine mastitis due to *C. jejuni* has been experimentally induced with few reports of naturally occurring cases [107]. There are reports of experimentally induced and naturally occurring cases of diarrhoea in lambs due to *C. jejuni* [107]. Abortion due to *C. jejuni* (predominantly ST-8) has been shown in both cattle and sheep [110] [111] [112]. It is possible the ability to cause gastrointestinal disease may depend on the age of the animal, its immune status and the strain of *C. jejuni*.

Swine frequently carry both *C. jejuni* and *C. coli* as normal gastrointestinal flora but there are reports of young immunologically naive animals showing signs of anorexia, fever and diarrhoea after inoculation with *C. jejuni* [47][p6]. It is also reported that co-infections with some viruses [113], bacteria [114] and parasites [115] can increase the disease pathology suggesting that host immunity is a significant determinant of disease outcome [47][p6]. No reports of pig abortions associated with *C. jejuni* were found.

Reports of *C. jejuni* related disease are not common in horses. *C. jejuni* has been isolated from foals with clinical signs of fever, colic and non responsive diarrhoea [107].

While *C. jejuni* is not a commonly attributed cause of goat abortions; experimentally a goat inoculated with *C. jejuni* has aborted and an aborted foetus has culture pure *C. jejuni* [107].

Chickens (*Gallus gallus*) are a significant source of *C. jejuni* infection in people with a high prevalence in many flocks [116][p667]. Normally there are no significant clinical signs to suggest disease, although a period of sub-optimal growth associated with the initial stage of colonisation of a flock with *C. jejuni* has been suggested [107]. Certainly inoculation of very young chicks has produced clinical signs of gastroenteritis [117] and pro-inflammatory response in the intestinal mucosa [118] and even mortality has been reported [107]. It has also been reported that some breeds of chicken are more susceptible to *C. jejuni* resulting in damage to gut mucosa and can even diaharroea [119] [107].

Vibrionic hepatitis is a condition in chickens characterised by focal necrosis in the liver, that shows as a a drop in egg production and an increase in morbidity. After appearing in the 1950s and 1960s it disappeared then reappeared in the early 2000s [107] [120]. While initially attributed to *C. jejuni* alone, more recent investigations do not support the presence in the liver of *C. jejuni* to be sufficient alone to initiate this condition and suggest it is a contributing factor [120].

There are reports of *Campylobacter* spp. being pathogenic in farmed ratites [121]. *C. jejuni* was isolated from ostriches with nondescript clinical signs of poor growth rates, anorexia, lethargy, weakness and ataxia before death with watery urates in the faeces [122]. Amongst the pet birds (parrot, finch, canary) *C. jejuni* has been associated with hepatitis, lethargy, loss of appetite, weight loss, yellow diarrhea and mortality [47][p7].

Although it is frequently stated that *C. jejuni* is avirulent in birds [123], little is known of the pathogenicity of the various strains and their ability to cause disease in the various wild-bird species. A report from Waldenstrom et al (2010) that the inoculation of a strain from a Song Thrush (*Turdus philomelos)* into a European Robin (*Erithacus rubecula*) resulted in a transient weight loss and antibody response [124]. Another report states three-week-old Japanese quails (*Coturnix coturnix japonica*) had diarrhoea for 2 weeks after oral inoculation with *C. jejuni* [107].

Recently *C. jejuni* isolates were reported in grey seals (*Halichoerus grypus*) and attributed to human/poultry sources while there was not sufficient evidence to attribute pathogenicity to the finding, there was concern regarding human to wildlife transmission [125].

2.3.9 Disease control measures

Most food processing hygiene measures are aimed at *E. coli* as a representative model organism, however *C. jejuni* can behave in a different manner than *E. coli* [126]. We are only beginning to understand the mechanisms and behaviours involved in *C. jejuni's* survival; for example the ability to survive inside some amoeba [127], in biofilms [128], the viable but non-culturable state [129] and quorum sensing [130]. It has even been suggested that the presence of other bacteria can aid the survival of *C. jejuni* in the environment [95]. Control methods need to consider these behaviours, along with its improved survival at lower temperatures [131].

In developed countries, the main risk factors for contracting campylobacteriosis are eating undercooked poultry meat, drinking raw milk, or drinking untreated water, and to a lesser degree, living in a household with a cat or dog and in rural New Zealand setting contact with calves [95] [101]. Control measures that have worked for poultry include increased hygiene requirements on-farms to decrease disease incidence in flocks and during processing to reduce contamination [98] [83][p169].

Until recently raw milk was not sold to the public in New Zealand , recently this changed to restricted sales with higher hygiene requirements. Although there is some small evidence that

chronic exposure to raw milk produces a protective immunity, the consumption of raw milk is a significant cause of outbreaks [77]. As treatment of the milk is contrary to the ethos of raw milk consumption, strict controls on its production are required, and monitoring that standards are meet.

In New Zealand, water from the roof or a bore to underground water are common sources of water in isolated rural areas. Options to control the incidence of campylobacteriosis under these conditions exist. Such as small scale treatment methods e.g. boiling drinking water, or increased awareness of good practices for collecting and storing water, e.g. regular cleaning of roof and pipes used in collecting water, or improved monitoring of the collected water e.g. regular checking of the storage tank.

2.4 Molecular Epidemiology and subtyping

This is not a comprehensive review of all possible methods of typing but the aim is to include those that have been regularly applied to and used for epidemiological purposes with *C. jejuni*.

The diversity shown by *C. jejuni* at the phenotypic and genotypic level is well established [132][p59]. A typing system needs to show diversity in a species but stability in a strain [133]. Generally the phenotypic typing systems tended to suffer from an inability to type some isolates leading to ambiguous data [134]. Being a relatively slow growing, fastidious and relatively biochemically inert species *C. jejuni* has been less amenable to some of the more traditional laboratory phenotyping systems [135]. In general the genotypic systems tend to suffer from the high rate of recombination (genetic exchange) observed in isolates distorting relationships [136], although MLST has proven to be a reliable and informative epidemiological tool [137]. The new genomic era offers new typing systems like rMLST and wgMLST that may widen our perspective of *C. jejuni* [138].

2.4.1 *Campylobacter* typing systems

The purpose of bacterial typing methods in molecular epidemiology is to generate isolate specific molecular "fingerprints" for assessment of epidemiological relatedness [139]. Since the discovery of *Campylobacter* not only has its taxonomic name changed but a variety of methods have been developed, and continue to be developed, to aid the identification and relatedness of *Campylobacter* spp. isolates [140]. Many methods have been developed as typing systems for *C. jejuni* and they can be loosely divided into phenotypic and genotypic methods. Each method has its own strengths and weakness which must be considered when applying these methods [141].

In 2007 a new set of criteria for typing systems were proposed by the European Society of Clinical Microbiology and Infectious Diseases (ESCMID) Study Group on Epidemiological Markers (ESGEM) in response to the technological advances in the previous decade [139]. The performance criteria

proposed as an update of the previous guidelines for a typing system in an epidemiological study were: stability of the marker for the period of the study, universal type-ability of all isolates, a useful level of discrimination that is concordant with the epidemiological picture and a high degree of reproducibility [142] [139]. Once the performance criteria is attained, a convenience criteria can be applied which can include: flexibility (range of species), rapidity, accessibility, ease of use, cost and suitability for computerisation and storage of results [139]. Taboada suggested the ideal *Campylobacter* typing/subtyping methods should have the features of: deployability (easy, simple interpretation, high throughput, inexpensive), portability (able to enter into a computer database), accessibility (standard laboratory equipment in front line laboratories), epidemiological specificity (unambiguously linking related and excluding unrelated), versatility (discriminate to the level needed), and stability (not impacted by frequent recombination) [137]. These two sets of criteria for epidemiological typing systems are not the only ones available, and while not being mutually incompatible, they do show that the view on what is "fit for purpose" in an epidemiological typing system has a variety of perspectives and it is from the perspective of the end use (or question being addressed) that a typing system is finally judged.

2.4.2 Phenotyping

Phenotype systems

Phenotypic methods seek to characterise and differentiate isolates based on the products of their gene expression using properties like biochemical profiles, carbon metabolism, antimicrobial susceptibility profiles, bacteriophage types, and antigens present on the cell surface. As gene expression is the basis of these techniques the results can vary due to influences like growth conditions, growth phase of the culture, and phase variation.

Biotyping

Biotyping is the identification of bacterial isolates through the expression of metabolic activities which can include colony morphology, environmental tolerances, and biochemical reactions. Biochemical tests were developed from the 1980s beginning with Skirrow and Benjamin (1980) and then Hebert et al.(1982). These systems were based on using different combinations of hippurate hydrolysis, rapid H_2S production, nalidixic acid resistance, and DNA hydrolysis [143] [144]. Lior proposed a combination of these biotyping tests (hippurate hydrolysis, rapid H_2S production, and DNA hydrolysis) and extending them with serotyping in 1984 [145]. The Preston test utilises 12 tests, consisting of growth at 28℃, hippurate hydrolysis, and 10 resistotyping tests [146]; while the commercial kit API-Campy is also still popular [147]. Resistotyping is a component of biotyping, testing an isolates sensitivity to selected antibiotics, either by agar dilution or disc diffusion and is used in the Lior test (nalidixic acid), API-Campy (four tests), Preston biotyping (10 tests).

In general, biotyping techniques alone suffer from poor reproducibility and discriminatory power,

due to variation of gene expression caused by environmental factors and lack of biochemical diversity between isolates [18] [19], and considered more useful in combination with a more discriminating system. However they do continue to be useful for laboratories that lack access to molecular methods [148] or distinguishing between outbreak and non-outbreak associated isolates [149][p166].

Serotyping

Serotyping is a classical phenotype based tool and has been used since the beginning of the 20th century [150]. It has been applied to many food-borne pathogens for strain differentiation, using antibodies and antisera to detect surface antigens. *Campylobacter* spp. have an array of cell surface structures including lipopolysaccharides, capsular polysaccharides, membrane proteins, and extracellular organelles such as flagella [151].

Serotyping of *Campylobacter* began in 1971 when Berg et. al. identified 3 heat stable antigens that separated *Campylobacter* into three serotypes . This was followed by Penner and Hennessey proposing a scheme in 1980 based on extracted thermostable (HS) antigens located in the outer membrane that were thought to be the lipopolysaccharide (LOS) but are now believed to be predominately related to capsular polysaccharides (CPS) [152] [153] [154]. In 1982 Lior et. al. presented a scheme based on a slide agglutination technique using live bacteria and detecting heat-labile (HL) antigens [155]. The method was developed according to the same principles used to identify flagella antigens in other Gram-negative bacteria, and defined the HL antigens of three species *C. jejuni, C. coli,* and *C. lari*, but flagella may not be the sero-determinant [156] [19]. In a comparison study using 1,405 isolates, 96.1% of the isolates were allocated a type by the Penner method and 92.1% of the isolates were allocated a type by the Lior method [157]. Patton considered the two systems were compatible which has been supported by several studies combining both HS and HL antigens [158] [157] [159].

Both the Penner and Lior typing systems have been widely used but suffer from being time-consuming, have costly reagents, require large panels of antisera, and still result in a high number of untypeable strains [19]. Attempts to minimise these shortcomings have included combining the Penner system with other non-serotyping systems [160]. A modified Penner serotyping scheme (LEP) [161] was also developed to simplify the method and reduce problems with non-specific agglutination and cross reactivity, but McKay found 36% of isolates were untypeable using this method [162]. The high levels of isolates that are un-typeable in the Penner system may be due to some level of specificity to the time and place it was developed [163][p28]. It was developed in the 1980s in Canada so *C. jejuni* antigens that may have arisen since then or in other countries are not accounted for.

Protein profiling

A protein profile is created by extracting proteins from a population of bacteria growing under rigidly defined laboratory conditions. Usually a whole cell lysate is denatured and added to a

vertical polyacrylamide gel for electrophoresis, after which the resulting bands are stained and compared for migration and intensity. This technique can identify *Campylobacter* at the species level, on the basis of species-specific bands [164], but numerous bands are present in this technique making analysis complicated and it is not recommended for use beyond this level of discrimination [165] [166]. The technique is time consuming, laborious, and not considered suitable for large numbers of isolates [167].

Multilocus Enzyme Electrophoresis (MEE)

Multilocus enzyme electrophoresis (MEE) has been in use since the 1980s and applied to population genetics, systematics, and molecular epidemiology [168]. Using the technique, bacterial isolates show variations in the electrophoretic mobility of different constitutive enzymes, usually 9–11 enzymes which separate out under electrophoresis in non-denaturing conditions [168]. Applying this method led Meinersmann to suggest *Campylobacter* may have a clonal framework with other portions of the genome involved in frequent recombination [169]. Although MEE has a high level of discrimination it is not greatly used due to its low reproducibility, difficulty comparing results between laboratories, and it can generate untypeable alleles [149][p167] [19].

Phage typing

Phage typing relies on the pattern of lysis shown by an isolate to various virulent typing bacteriophages [170], and isolates with the same pattern of lysis are related as the same "phage type". This method depends on the *Campylobacter* cell expressing a suitable receptor for the bacteriophage to enter the cell, and instability or lack of expression can effect the outcome. Several schemes for *Campylobacter* phage typing have been proposed using either USA, UK and Canadian bacteriophages [170] [171] [172]. The use of phage-typing in conjunction with other techniques such as serology has also been suggested [173]. The disadvantages of this technique include non-typeable isolates and problems with cross reactivity, it is also labour intensive and expensive, requiring the maintenance of large panels [141].

Fatty acid methyl ester (FAME) analysis

Using a gas-liquid chromatograph and a computer database it is possible to analyse and compare the different fatty acids components of the cell wall (eg, lipid A or lipoteichoic acid) and cell membrane to identify an isolate [174]. If the isolate itself is unknown because the lipid content is phylogenetically conserved within the bacteria, it is possible to establish a degree of relatedness to a known species [174]. This technique has been used to investigate *Campylobacter* physiology [87], and when compared to PGFE showed a lesser degree of discrimination with limited ability to sub-speciate isolates [18]. Once operating, this process is fast simple and all isolates are able to be allocated to a type but standardisation of isolate preparation is necessary for reproducibility [167].

MALDI-TOF

Matrix assisted laser desorption ionisation time of flight mass spectrophotometry (MALDI-TOF) analyses the whole intact cell by laser bombardment of the cell in a matrix aiding the ionisation process and allowing the separation of gas-phase ions on a mass-charge basis [175]. This method has been applied to differentiate *Campylobacter* at the species level [176] and even discriminate specific subtypes within the *C. jejuni* [177]. Although a rapid technique which circumvents the relatively low level of biochemical activity of *C. jejuni*, MALDI-TOF requires rigid standardisation and the resulting data is complex, often requiring computerised cluster analysis with a comparable database [178][p28]. A recent analysis of MALDI-TOF spectra clustering did not correspond with the MLST sequence, biochemical phenotype or phage susceptibility of the *C. jejuni* and *C. coli* isolates [179]. Significant differences in the probability of correct identification at the *Campylobacter* species level can occur due to differences in growth medium, the incubation temperature, time at interpretation (from 24 h to 120 h), and the database applied [180].

Raman spectroscopy

Raman spectroscopy uses laser emitting photons at specific wavelengths to excite specific molecules, then reading the resulting pattern of peaks to identify the phenotype [181]. This approach was used by Read to compare rMLST typed *Campylobacter* showing a 83.95% accuracy at the species level and qualified success comparing clades and clonal complexes, but the results also suggested that the host environment could affect the phenotype and that the phenotypic signal may change as strains moved between hosts [182]. The reported "host effect" suggests, like the FAME and MALDI-TOF analysis, standardisation of isolate preparation is an important factor and can affect the results and their interpretation.

Enzyme profile analysis

The enzymatic profile of 234 *Campylobacter* strains has been evaluated and 40 different enzymes were identified as present at least once, with the mean number of 13 per isolate [183]. Elharrif and Megraud (1986) used a relatively simple technique involving commercially available methods but there is little reference to its usage for recent epidemiological investigations [4]. Although this approach would be able to type isolates to a enzyme profile phase variation could result in mis-allocation of an isolate.

Biolog Phenotypic microarrays

The Biolog phenotypic microarray system uses a series of proprietary plates developed and manufactured by Biolog (Hayward, CA, USA) [184]. The Biolog phenotype assays measures cell respiration with tetrazolium dye reduction, by a bacteria in a minimal nutrient environment containing a defined nutrient. This system can test a bacteria's phenotype using carbon, nitrogen, phosphorus, and sulphur metabolism, sensitivity to various salts and ions and total osmolarity, sensitivity to

pH, and sensitivity to 240 chemicals. The phenotypic microarrays are standardised and consists of 96 wells that can be read in a spectrophotometer or in the purpose built Omnilog plate reader. Using the Biolog phenotypic microarrays in combination with the Omnilog plate reader (Biolog Inc, Heyward, CA) has been called by the manufacturer a "high throughput phenotypic microarray system that produces global phenotypic characterisation".

The Omnilog plate reader acts as an incubator with a robotic arm and CCD camera to spectrophotometrically record the dye colour change in each well on a plate every 15 minutes. The phenotypic microarray plates, PM1 and PM2A each have 95 wells containing a unique carbon source and a single negative control well. It is presumed if the bacteria are able to metabolise the supplied carbon source, the resulting electron flow this will lead to reduction of the tetrazolium dye [185]. When all the cells processes are working normally there is a flow of electrons from the carbon source to NADH then along the electron transport chain of the cell, and finally the electrons are accepted by the tetrazolium dye to produce a purple colour [186]. However the process may not always be this specific, for example molecules other than NADH can donate electrons to the *C. jejuni* electron transport chain or act directly to reduce the tetrazolium dye [187] [188].

Tetrazolium is a water-soluble molecule that is readily reduced to form a formazan in the presence of an electron donor [189]. The colorless tetrazolium molecule can be reduced and depending on its formulation produce a blue, purple, red, or fluorescent formazan [190] in the form of an insoluble crystal or a water soluble dye [191]. The tetrazolium compounds can be formulated to cross the cell membrane and accumulate as an intracellular product or stay extracellular. The Biolog system offers the terazolium dye in several proprietary formulations, recommending a formulation for a given bacteria. There are many examples of non-proprietary formulations and their uses including indicating activity in activated sludge [192] and the presence of bacteria in groundwater [193]. 5-cyano-2,3-ditolyl tetrazolium chloride (CTC) has been suggested as being more indicative of viability than activity in bacteria [194]. A tetrazolium dye, in the form of 5-cyano-2,3-ditolyl tetrazolium chloride-4′,6-diamino-2-phenylindole staining has been used to identify respiration in the viable but non culturable (VBNC) form of *C. jejuni* [195].

The Omnilog plate reader comes with some in-house software to record and provides initial analysis of the colour change of the tetrazolium-based dye in each well recorded over time. The pattern in a well where the bacteria is respiring with the carbon source is called a "respiration curve" as shown in Figure 2.2 [196]. The idealised "respiration curve" can be considered similar to a growth curve and analysed into four parameters [197]. The parameters are shown in Figure 2.2; lag phase (λ), slope of curve (μ), maximum value of curve (A), and the area under the curve (AUC). These parameters can be used to investigate the kinetics of the reactions, for example it has been used to show the effect of starvation on the behaviour of *Escherichia coli* (*E. coli*) O157:H7 [198]. Analysis of this data and parametrisation is facilitated by a purpose built R package called "opm" [199]

[196].

Various uses have been proposed for the application of the Biolog phenotypic microarray technology, particularly in combination with next generation sequencing (NGS) of genomes. Gawand et al. (2013) have suggested its use in developing genome-scale metabolic models [200]. The results from PM1 and PM2A analyses have been used to model the essential genes of *C. jejuni* [201]. Ductape [202] is a programme that combines the data from the phenotypic microarrays with an isolates genome to provide output that includes the metabolic pathways in the form of KEGG maps [203]. The KEGG maps [204] [205] cover a series of metabolic pathways showing the presence of a gene (from the genome) in the form of an enzyme and the level of phenotype activity (from the phenotypic microarray plate) in a reaction.

Potential difficulties with this system include:

- While it has been reported as an irreversible reduction of the terazolium dye there appears to be reversible reduction by some *Enterococcus* spp. (Dr Anne Midwinter, personal communication, June 2012).
- The lack of a standardised threshold makes comparison between investigations difficult. Variation in deciding which parameter to use and a suitable cut off point that ascribes a nutrient as utilised or not has not been standardised.
- False positives: It has been shown that under microaerophilic conditions there is an abiotic reduction of the tetrazolium dye in certain wells on PM1 and PM2A [187]. This results in an inability to determine if the organism being tested or the environment is reducing the tetrazolium dye.

Phenotypic microarray results for *C. jejuni*

The Biolog phenotypic microarray system (PM1 and PM2A) have been previously used to investigate *C. jejuni* and reported in seven articles. The summary of the combined results from four articles: Line et al. (2010) [187], Tang et al. (2010) [206], Gripp et al. (2011) [5], and Wagley et al. (2014) [207] are shown in Table 2.1. Due to the paucity of information provided in the articles by Brandl et al. (2004) [208], Fouts et al. (2005)[209] and Muraoka and Zhang (2011)[210] they were not included in Table 2.1.

Brandl et al. (2004) used Biolog GN and ECO plates and reported utilisation by *C. jejuni* 10 carbon sources, although now some are recognised as likely false positives under these conditions (L-Arabinose, L-Fucose, α-D-Glucose, α-Hydroxybutyric acid, Succinic acid, L-Asparagine, L-Aspartic acid, L-Proline, L-Serine, L-Cysteine). Fouts et al. (2005) reported preliminary Biolog PM plates data showing variation and suggested this may be due to strain variation [209]. The full details of the analysis and results were not reported but the isolate was *C. jejuni* RM1221 (original source chicken [211]) and said to respire using arabinose, fucose, formic acid, acetic acid, fructose, mannose,

hydroxybutyric acid, asparagine, and aspartic acid [209]. Muraoka and Zhang (2011) used the PM1 on three *C. jejuni* strains 11168, 81116, and 81-176 (all originally from clinical sources [211]) to demonstrate phenotypic variation, as some strains were able to utilise L-fucose [210]. From Figure 2 in their article it appears the isolates all utilised 21 single carbon sources (wells) and possibly another 5 single carbon sources (wells) [210].

Line et al. (2010) used PM1 and PM2A plates to compare the phenotype of *C. jejuni* strain 11168 (GS), original source was a clinical sample [211], at 37℃ and 42℃ [187]. The parameter reported was the relative absorbance at 66 hours which is presumably the equivalent (usually) of the maximum height parameter (A). The reported mean values sometimes have a large standard deviation which implies large variation in the replicated tests for the isolate. Line et al. (2010) did report the oxidation of the substrates was generally greater at 37℃ than at 42℃ [187]. They also reported ten substrates (wells) that were showing potential false positive reactions in the microaerobic conditions required by *C. jejuni*: L-Arabinose (PM1 A02), D-Xylose (PM1 B08), D-Ribose (PM1 C04), L-Lyxose (PM1 H06), D-Arabinose (PM2A B05), 2-Deoxy-D-ribose (PM2A B09), D-Glucosamine (PM2A E05), 5-keto-D-gluconate (PM2A E12), Oxalomalate (PM2A F05), Dihydroxyacetone (PM2A H09).

Using PM1 and PM2A single carbon source plates *C. jejuni* ATCC 33560 was phenotyped at 30℃ and 42℃ by Tang et al. (2010) [206]. A total of 15 carbon sources were reported as utilised, the same single carbon sources at both temperatures [206]. The analysis did not explain how they identified utilisation from non-utilisation wells, i.e. the parameter used and the cut-off threshold [206].

Gripp et al. (2011) used PM1 plates to analyse eight ST-21 isolates and one each from ST-51, ST-50, ST-48, ST-45, ST-267, ST-257, ST-607, ST-290, and ST-572 i.e. in total 17 isolates from 10 strains [5]. The eight ST-21 strains were from different sources (human, chicken, food, bovine) and the nine other isolates were all from humans [5]. Gripp et al. (2011) were unable to find phenotypic variation within the ST-21 isolates attributable to source [5]. In this analysis the cut-off point for utilisation was values > 800 for the area under the respiration curve (AUC) [5]. Using this as the cut-off the isolates all utilised the same core 19 single carbon sources and a combined total of 48 different single carbon sources [5]. The plates were incubated in the Omnilog plate reader at 37℃ or 42℃ but in the results it is not identified if this caused a difference in the outcomes [5].

Wagley et al. (2014) tested *C. jejuni* and *C. coli* isolates using the Biolog PM1 and PM2A plates and reported they differed in their ability to utilise propionic acid as a sole carbon source in culture medium [207]. However their suggestion that all *C. jejuni* may lack this ability is contradicted by some earlier work by Gripp et al. (2011) where an ST-50 isolate seemed to utilise propionic acid, however this may be attributed to the two papers using different parameters to identify the cut-off point [5] [207] [210] [206]. The 13 clinical isolates used in the investigation Wagley et al.

(2014) involving PM1 and PM2A plates were from three different countries, but they were not all identified by MLST [207]. Although three isolates from geese were tested for propionic acid utilisation, it is not identified if they were domestic or wild birds [207]. The Biolog PM1 and PM2A plates were incubated at 37℃ but the parameter used or the cut-off point is not stated, however the values in the supplementary data table (Additional file 1. Raw data on carbon utilisation of *C. jejuni* and *C.coli* strains using Biolog analysis) are consistent with the results being the maximum height parameter. It is noticeable that in the three replicates for the each isolate used, there is considerable variation within some wells. Assuming a cut off value of > 100 in any of the replicates, to indicate a positive value for a given isolate in the well, this would result in 18 carbon sources being used by all the isolates and a total of 29 carbon sources were used by at least one isolate once.

Table 2.1 shows 11 carbon sources used by all the isolates in the Biolog phenotypic microarray analysis of the four studies, and a total of 53 different carbon sources were utilised by at least 1 strain as a sole carbon source. These results show there is a significant amount of phenotypic variation between *C. jejuni* isolates [212]

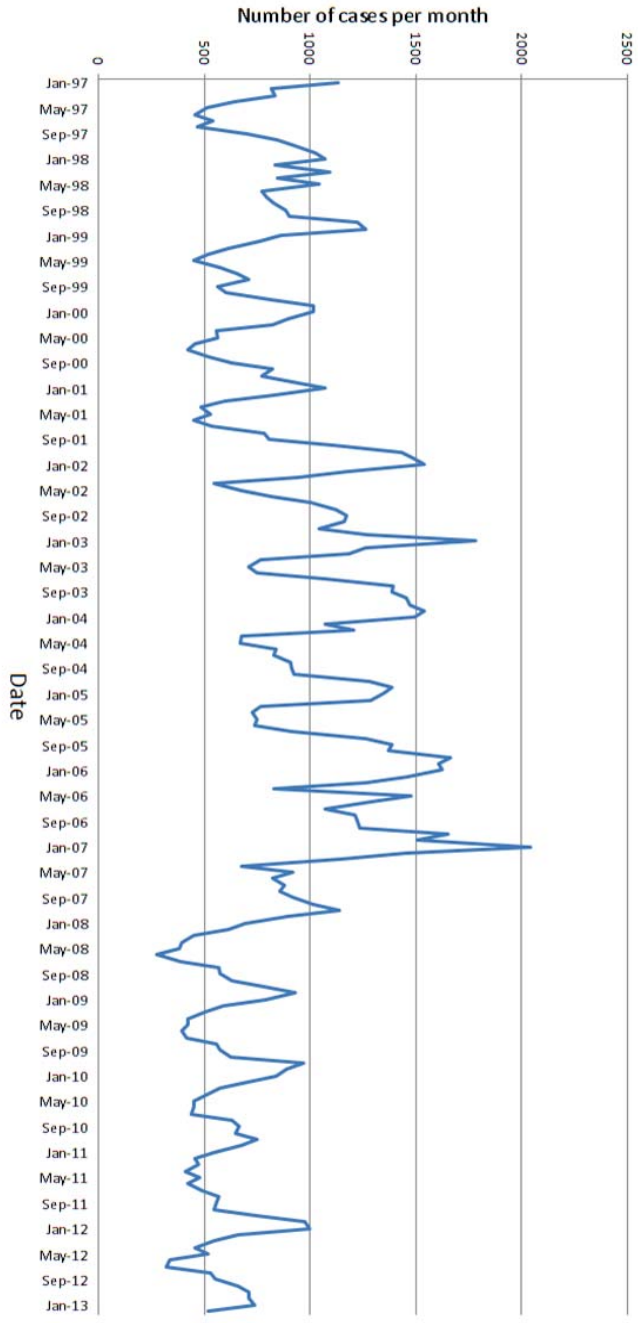


Figure 2.1: **Trends in *Campylobacter* cases in New Zealand January 1997 - May 2012.** The number of notified *Campylobacter* cases by month in New Zealand January 1997 – May 2012 shows two trends. A seasonal pattern and an increasing incidence trend until 2006 when interventions in the poultry industry resulted in a marked decline although the seasonal pattern is still apparent.

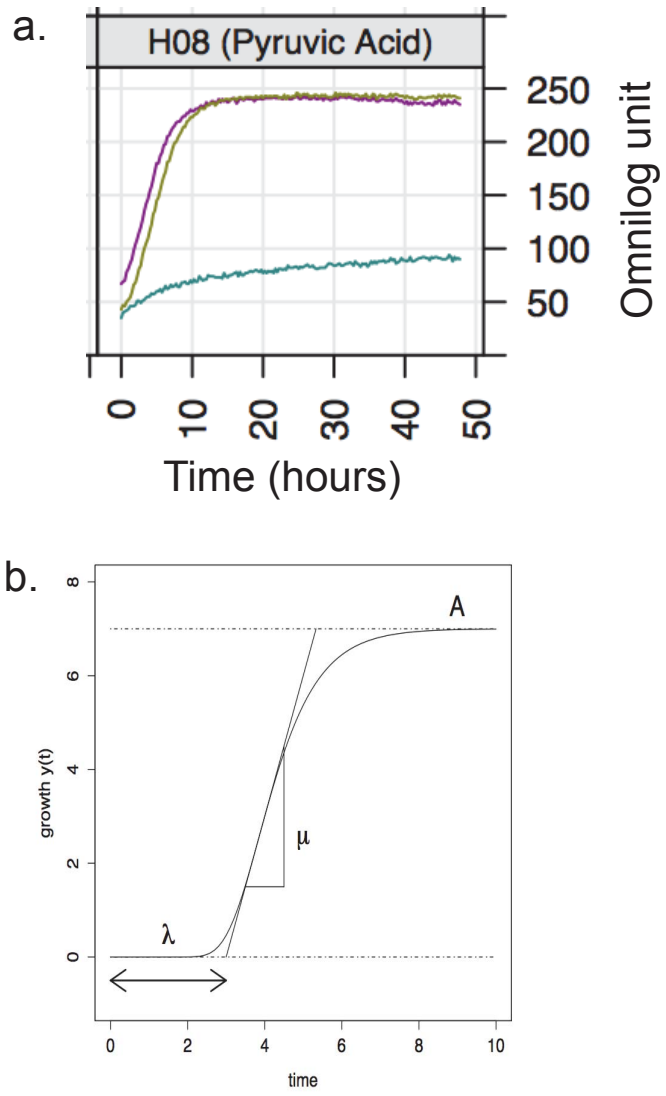


Figure 2.2: **The parametrisation of the Biolog phenotypic microarray respiration curves.** Figure 2.2 (a) shows two respiration curves (red, khaki) produced by a *C. jejuni* isolate in phenotypic microarray plate 1 (PM1) in a well H08 containing pyruvate as the single carbon source in two experiments. The curve is produced by the change in colour of the tetrazolium dye measured in Omnilog units by a spectrophotometer, over a period of 48 hours. The bottom green line is the colour change in the same well (PM1 H08) under the same conditions when no bacteria is present. Figure 2.2 (b) shows an idealised growth curve which can be approximated as a respiration curve seen in Figure 2.2 (a). The parameters are lag phase (λ), slope of curve (μ), maximum value of curve (A), and the area under the curve (AUC). The source of 2.2 (b) is "grofit: Fitting Biological Growth Curves with R" [197].

Table 2.1: **Utilisation of single carbon sources by *C. jejuni* strains using Biolog Phenotypic microarray (PM1 and PM2A).** The name of the carbon source and Chemical Abstracts Service registration number (CAS ID) in a Biolog phenotypic microarray well (PM plate well), are shown with the combined results for four investigations using *C. jejuni* isolates. The results from Line et al. (2010) are in the Line column, and were assessed as a positive (+) if the reported mean value plus 2 standard deviation at either temperature is > 100 or if unreported it is assumed to be (-) The results from Tang et al. (2010) in the Tang column, and were as reported (+) or (-) [206]. The results from Gripp et al. (2011) are in the Gripp column, and a cut-off value of > 800 (AUC) was used as a positive (+) result [5]. The results from Wagley et al. (2014) are in the Wagley column, and a cut-off value of > 100 was applied, when any of the three replicates reached this value it is a positive result (+), else a negative (-) result is assumed [207]. The results in false positive wells are not reported in this table [187]. NT means not tested. A $+/-$ represents inconsistent results for the carbon source being tested.

Carbon source	**CAS ID**	**Line**	**Tang**	**Gripp**	**Wagley**	**PM plate well**
Succinic acid	CAS 6106-21-4	+	+	+	+	PM1 A05
L-Aspartic acid	CAS 3792-50-5	+	+	+/-	+	PM1 A07
L-Proline	CAS 147-85-3	-	+	+/-	+	PM1 A08
D-Mannose	CAS 3458-28-4	-	-	+/-	-	PM1 A11
Dulcitol	CAS 608-66-2	-	-	+/-	-	PM1 A12
L-Fucose	CAS 2438-80-4	-	-	+/-	+/-	PM1 B04
L-Lactic acid	CAS 312-85-6	+	+	+	+	PM1 B09
Formic acid	CAS 141-53-7	-	-	+/-	+/-	PM1 B10
L-Glutamic acid	CAS 6106-04-3	+	+	+	+/-	PM1 B12
D-Glucose-6-Phosphate	CAS 3671-99-6	-	-	+/-	-	PM1 C01
D, L-Malic acid	CAS 6915-15-7	+	+	+/-	+	PM1 C03
Tween 20	CAS 9005-64-5	-	-	+/-	-	PM1 C05

Continued on next page

Combined results of *C. jejuni* isolates in PM1 and PM2A —– *continued from previous page*

Carbon source	**CAS ID**	**Line**	**Tang**	**Gripp**	**Wagley**	**plate well**
Acetic acid	CAS 127-09-3	-	-	+/-	-	PM1 C08
Thymidine	CAS 50-89-5	-	-	+/-	-	PM1 C12
L-Asparagine	CAS 70-47-3	+	+	+/-	+	PM1 D01
a-Ketoglutaric acid	CAS 22202-68-2	+	-	+	+	PM1 D06
a-Ketobutyric acid	CAS 2013-26-5	-	-	+/-	-	PM1 D07
Uridine	CAS 58-96-8	-	-	+/-	-	PM1 D12
L-Glutamine	CAS 56-85-9	-	-	+	+/-	PM1 E01
m-Tartaric acid	CAS147-73-9	-	-	+/-	-	PM1 E02
D-Fructose-6-Phosphate	CAS 26177-86-6	-	-	+/-	-	PM1 E04
a-Hydroxybutyric acid	CAS 19054-57-0	+	+	+	+	PM1 E07
Adenosine	CAS 58-61-7	-	-	+/-	-	PM1 E12
Glycyl-L-Aspartic Acid	CAS 4685-12-5	-	-	+/-	+/-	PM1 F01
Citric acid	CAS 6132-04-3	+	-	+/-	+/-	PM1 F02
Fumaric acid	CAS 17013-01-3	+	+	+/-	+/-	PM1 F05
Bromosuccinic acid	CAS 923-06-8	+	+	+/-	+/-	PM1 F06
Propionic acid	CAS 137-40-6	-	-	+/-	-	PM1 F07
Glycolic acid	CAS 79-14-1	-	-	+/-	+/-	PM1 F09
Glyoxylic acid	CAS 563-96-2	-	-	+/-	+/-	PM1 F10
Inosine	CAS 58-63-9	-	-	+/-	-	PM1 F12
Glycyl-L-Glutamic acid	CAS 7412-78-4	-	-	+/-	+/-	PM1 G01
L-Serine	CAS 56-45-1	+	+	+/-	+	PM1 G03

Continued on next page

Combined results of *C. jejuni* isolates in PM1 and PM2A —– *continued from previous page*

Carbon source	**CAS ID**	**Line**	**Tang**	**Gripp**	**Wagley**	**plate well**
Mono-Methylsuccinate	CAS 3878-55-5	-	+	+	+/-	PM1 G09
Methylpyruvate	CAS 600-22-6	+	+	+	+	PM1 G10
D-Malic acid	CAS 636-61-3	+	+	+	+	PM1 G11
L-Malic acid	CAS 138-09-0	+	+	+	+	PM1 G12
Glycyl-L-Proline	CAS704-15-4	-	-	+/-	+	PM1 H01
p-Hydroxy-Phenylacetic acid	CAS156-38-7	-	-	+/-	-	PM1 H02
m-Hydroxy-Phenylacetic acid	CAS 621-37-4	-	-	+/-	-	PM1 H03
Tyramine	CAS 60-19-5	-	-	+/-	-	PM1 H04
D-Psicose	CAS 551-68-8	-	-	+	-	PM1 H05
Glucuronamide	CAS 3789-97-7	-	-	+/-	-	PM1 H07
Pyruvic acid	CAS 113-24-6	+	-	+	-	PM1 H08
L-Galactonic acid-g-Lactone	CAS 1668-08-2	-	-	+/-	-	PM1 H09
D-Galacturonic acid	CAS 91510-62-2	-	-	+/-	-	PM1 H10
Phenylethylamine	CAS 156-28-5	-	-	+/-	-	PM1 H11
2-Aminoethanol	CAS141-43-5	-	-	+/-	-	PM1 H12
3-O-b-D-Galactopyranosyl-D-Arabinose	CAS 6057-48-3	-	-	NT	+/-	PM2A B12
D-Lactic acid Methyl Ester	CAS 17392-83-5	-	-	NT	+	PM2A F01
Quinic acid	CAS77-95-2	-	-	NT	-	PM2A F06
D-Ribono-1,4-Lactone	CAS 5336-08-3	-	-	NT	+/-	PM2A F07
Succinamic acid	CAS 638-32-4	-	-	NT	+/-	PM2A F10
N-Acetyl-L-Glutamic acid	CAS 1188-37-0	-	-	NT	-	PM2A G03

2.4.3 Genotyping

The entire bacterial genome is the most fundamental "molecule of identity" in the cell and thus represents the best basis for assessing isolate interrelatedness [213] and there are a wide range of techniques that have been developed to utilise this potential. Genotyping involves using the more recently developed molecular techniques usually either PCR, restriction enzymes, or nucleic acid sequencing. The stages of development for genotyping can be divided into first generation plasmid analysis, second generation restriction enzymes and probes, followed by third generation consisting of pulsed field gel electrophoresis (PFGE) and PCR-based methods, and finally fourth-generation DNA sequence-based approaches [213]. In the case of *Campylobacter* spp. found in some marine mammals, genotypic identification has been suggested as more reliable than phenotypic methods [214].

DNA hybridisation was developed in the 1980s and used successfully at the level of genera and species identification [164] [215]. DNA-DNA hybridisation is a time and labour intensive technique and a more simple DNA dot blot hybridisation assay has been developed that has the ability to differentiate atypical *C. jejuni* isolates such as nitrate-negative and nalidixic acid-resistant isolates [216]. Variations of this theme have been developed such as using melting point analysis of ribosomal-DNA (rDNA) probes, but its power of discrimination is only to the species level [217]. The increased availability of genomic analysis has resulted in the production of software like "Jspecies" to undertake *in silico* DNA-DNA hybridisation but this approach has yet to be accepted as a replacement for the laboratory bench technique to establish taxonomic identity [218]. This approach is more for taxonomic analysis than an epidemiological approach.

Plasmid typing

Plasmid typing assesses the presence or absence of a particular sized plasmid (extra-chromosomal, self-replicating double-stranded DNA molecule) using agarose gel electrophoresis. Plasmids can encode genes for antibiotic resistance which can be of epidemiological interest. Although plasmid typing can be used to separate isolates that have been typed to the species level [219] [220], there are problems with many *C. jejuni* lacking plasmids, and are therefore untypeable, and also the lack of consistent association between plasmids and other biological markers e.g. serotyping [221] [222].

Restriction endonuclease analysis (REA) or Bacterial restriction endonuclease DNA analysis (BRENDA)

The DNA content of a cell, genomic (i.e. total DNA including plasmids) or just chromosomal DNA, can be selectively cleaved by a restriction endonuclease at a specific site wherever it is present in the DNA [223] [224]. The choice of cleaving enzyme dictates the level of discrimination and there are a range of restriction endonuclease from which to choose [225] [222] [226]. The different sized fragments are resolved into a linear pattern by conventional agarose gel electrophoresis and

stained with ethidium bromide. In 1988, a study using REA typing on 316 New Zealand isolates which showed a similarity of 49.7% between the human isolates and the poultry isolates, researchers concluded that poultry was a major source of human *C. jejuni* infection [227]. REA has been used to identify the source of outbreaks and identify at the species level [228].

Pulsed Field Gel Electrophoresis (PFGE)

PFGE has emerged as a popular molecular approach for the epidemiological analysis of bacterial pathogens. Since it development in 1984, it has even been called the "gold standard" [149] [213] [229]. PFGE is a variation of the REA method with both methods allowing the examination of the entire genome without the cost and equipment to sequence the entire genome. Digestion of the genomic DNA by rare cutting restriction endonuclease can result in large DNA fragments (>15–20kb) that are not differentiated by standard electrophoresis but can often be separated when voltage is periodically switched among three directions: one that runs through the central axis of the gel and two that run at an angle of 60 degrees either side. The technique has proven valuable in epidemiological examinations of *C. jejuni* [230] [133] [231]. PFGE is reported as being a good predictor of MLST clonal complex, but not MLST sequence type when tested on New Zealand *Campylobacter* isolates [232]. The drawbacks of PFGE include:

- Reproducibility can be difficult but is achievable when well standardised such as on PulseNet [233].
- The time required by the method and its complexity [141].
- There can be difficulties with interpretation of the gel patterns which has been helped by software such as Bionumerics (Applied Maths) but a major drawback still exists in using these macro-restriction profiles to unambiguously identify relationships between isolates although Tenover has proposed a classification system of these relationships [234] [213].

An analysis by Sails reported PFGE was more discriminatory than MLST, but genomic instability (plasticity) due to frequent recombination events makes this technique unsuitable for longitudinal studies [141] [235].

Ribotyping

Ribosomal RNA (rRNA) cistrons are highly conserved but differ widely between bacterial species in both number, location, and primary nucleotide sequence for 5S, 16S, and 23S rRNA [225]. The rRNA DNA cistrons are detected using a labelled probe in a Southern blot upon an isolates genomic DNA after it has been digested with a restriction endonuclease and separated by gel electrophoresis [236]. Sequence variation in the flanking sites to the conserved rRNA results in different banding patterns when the DNA fragments are probed. The process is called ribotyping and the resulting pattern has been called the ribopattern or the ribotype [236] [167]. Ribotyping has been automated using the RiboPrinter™ Microbial characterisation System (Qualicon, Wilmington,

DE) which includes a database. This system has been used to differentiate *C. jejuni* and *C. coli* [237]. Ribotyping has been used for some bacterial epidemiological studies and *Campylobacter* species differentiation [225] [238]. Advantages of this method include the universality of rRNA, which makes prior knowledge of the genome unnecessary, while the lesser number of fragments compared to other techniques like RFLP makes interpretation easier [236]. Disadvantages include less discrimination than PFGE which uses the whole genome rather than being limited to the rRNA genes in *Campylobacter* spp. [19] [160], it is technically demanding, labour intensive and needs to be standardised if done manually [236]. There is an automated system which can avoid these labour problems but this can be an expensive option [236].

Amplified Fragment Length Polymorphism (AFLP) analysis

AFLP combines restriction endonuclease analysis with PCR thus utilising all the genome without requiring prior knowledge of the genome. The isolates genomic DNA is cut, usually with two restriction endonucleases, and the resulting restriction fragments are ligated with end-specific adapters. The adapters provide the basis for the primers to anneal in the PCR reaction but the primer is usually longer than the adapter, often by 1–3 bp [236]. The technique was patented by Zabeau and Vos in 1993 and published in 1995 [239] [240]. The resulting DNA fragments (PCR product) can be detected by conventional gel electrophoresis or automated DNA sequencer. The DNA fragment pattern can be analysed by software such as GeneScan software (PE Applied Biosystems) and provide densitometric values GelCompar v4.1 software (Applied Maths, Kortrijk, Belgium). Lindstedt et al.(2000) found AFLP analysis of *C. jejuni* strains to be a rapid method and it showed more discriminatory power than PFGE or PCR-RFLP [241]. Schouls et al. (2003) found Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR), MLST and AFLP analysis were equally powerful in their comparison study investigating *C. jejuni* outbreaks [242]. The method can be modified to use fluorescently labelled primers for PCR amplification (fluorescent AFLP or FAFLP) and an automated DNA sequencer for fragment detection providing similar discrimination to PFGE [243] [244]. AFLP has good discrimination but this is dependent on the combination of restriction endonucleases and the number of selective nucleotides in the primers, and it has similar reproducibility to other DNA banding patterns [236]. As with other molecular techniques, standardisation of the technique is important and necessary, so comparisons can be made between laboratories. Disadvantages include the equipment which can be expensive and the complex nature of the DNA fragment patterns can make interpretation challenging.

Random amplified polymorphic DNA (RAPD PCR)

RAPD uses arbitrary random primers that are ∼10 bp (10 mer) in length to bind to multiple regions over the chromosome and in the multiplex PCR reaction generate multiple amplicons. Alternatively repetitive extragenic palindromic (REP) primers can be used [245]. This technique does not require previous knowledge of the genome and uses the whole genome. It has been used on

C. lari isolates from shellfish in the Netherlands [246]. Disadvantages include lack of consistency and reproducibility with slight changes in cell numbers and growth conditions altering results, as well as poor correlation with other molecular techniques [167].

Single gene PCR based methods

Using this technique different *Campylobacter* species can be differentiated based on the similarity of the nucleotide sequence at one locus, and this approach has been demonstrated for many loci [149][p166]. Due to the frequency of inter-genomic and intra-genomic recombination in *Campylobacter* spp. relying on a single gene to identify the degree of relatedness between isolates is problematic for epidemiological studies and this review will limit itself to one example of a single gene typing system *fla*A [247] [235] [242]. Discrimination of isolates at the intra-species level has been demonstrated by restriction fragment length polymorphism (RFLP) of the flagellin A (*fla*A) gene or nucleotide sequencing of the short variable region (SVR) of *fla*A [248] [249] but other genes such as *por*A have also been used [250].

Flagella typing

C. jejuni flagella are important for motility and have been associated with pathogenesis [251]. The flagella of *Campylobacter* are complex and their proteins are encoded by a major flagellin gene, *fla*A, and a minor flagellin gene, *fla*B. Both *fla*A and *fla*B are homologous with highly conserved regions. The *fla*A gene has an area of short variable repeats (SVR) that frequently shows variation [252]. Flagella typing has shown to be suitable for the initial grouping of isolates in a surveillance situation but is better used in combination with other methods, especially with data covering long time periods [253] [167]. For example, it has been combined with MLST and with 16S rRNA oligonucleotide probes [254] [255]. Flagella typing is quick and has good discrimination of *C. jejuni* strains but on its own may not differentiate *C. jejuni* and *C. coli* [255]. There is evidence of inter-genomic (between strains) and intra-genomic (between *fla*A and *fla*B) recombination [252] [255], which would result in long term relationships predicted by these genes not reflecting the true clonal relationship when considering the rest of the genome [252]. Flagella typing has been performed using two different methods to type *C. jejuni* at the strain level.

- **Flagellin gene restriction fragment length polymorphism analysis (fla-PCR-RFLP)**
  A PCR-based method, where the *fla*A PCR product is cut using rare restriction enzymes is called restriction fragment length polymorphism (RFLP) or fla-PCR-RFLP. This method has been used to type *Campylobacter fla*A using various restriction endonucleases and primers [256] [257] [258]. The discriminatory power of flagellin gene typing using RFLP is determined by the type of restriction enzyme used [133], therefore standardisation of this technique is required for comparison purposes.

- ***fla*A SVR sequencing**
  The *fla*A can be sequenced and compared to other *fla*A sequences in a database (PubMLST)

[259]. Sequencing of a short variable region (SVR) in the gene encoding flagellin A (*fla*A), called fla-SVR, is a relatively cheap and easy genotyping method suitable for local and short term epidemiological investigations [248] [260]. Analysis of the DNA sequence of a 582-bp segment of *fla*A produced strain groupings similar to that generated by PCR-RFLP but further differentiated two *fla*A PCR-RFLP types (with a 1-bp difference in the 582-bp region) [133].

Clustered regularly interspaced short palindromic repeats (CRISPR)

CRISPR loci comprise two main elements, the CRISPR spacer array and a group of CRISPR-associated (*cas*) genes; combined they are referred to as a "CRISPR-Cas system". CRISPR elements function as a nucleic acid-based adaptive immune system in about ∼85% of archaea and ∼48% of bacteria [261]. Louwen et al. (2013) showed that CRISPR elements of *C. jejuni* are diverse, associated with the lipooligosaccharide (LOS) locus class and some of the spacer sequences were also shown to have high sequence similarities with *C. jejuni* (pro-)phage DNA [262]. Sequencing CRISPR loci has produced effective subtyping of *C. jejuni* for outbreak investigation [263] [242] [262]. However factors such as CRISPR loci not being present in all *C. jejuni* isolates (missing entirely or in part in ∼25% [263]), and variable levels of polymorphism mean its usefulness for strain typing in short term epidemiological studies must be assessed on a case-by-case basis [261]. An approach to increase its epidemiological application used by Kovanen et al. (2014) combined CRISPR dynamic polymorphism with the more stable MLST system which identified variation within the main sequence types [264].

Multilocus sequence typing (MLST)

Multilocus sequence typing (MLST) is a technique that uses comparative DNA sequencing of conserved house-keeping genes (under stabilising selection for the conservation of metabolic function) to characterise haploid organisms by indexing genetic variation under a standardised framework [141]. By using multiple genes distributed around the genome there is some stability despite recombination events (or horizontal gene transfer (HGT)). Several schemes have been described for *C. jejuni*:

- Dingle and colleagues [265] described an MLST scheme using seven housekeeping genes (*asp*A, *gln*A, *glt*A, *gly*A, *pgm* or *glm*M, *tkt*A, *atp*A or *unc*A). 400–600 bp segments are sequenced from each of the loci. Each of the alleles are assigned a unique allele number on the basis of those already present in the database, and the alleles of an isolate combine to form a seven-digit code that is then assigned a sequence type (ST). This grouping of alleles has been referred to as the definitive MLST alleles [266]. This curated database is accessible from a world wide web site for *Campylobacter* PubMLST [3].

- Suerbaum proposed a different set of seven housekeeping genes (*asd*, *atp*A, *ddl*A, *eft*S, *fum*C,

[3]Campylobacter PubMLST http://pubmlst.org/campylobacter/

*nuo*H, *yph*C) [267].

- An extended MLST can include various other genes. *fla*A and *por*A are not conserved housekeeping genes but are under diversifying selection and are also available on the *Campylobacter* PubMLST website. *fla*B has been used in an extended MLST for population analysis [268]. These three genes have been referred to as antigen gene sequence typing (AGST) and can increase the discrimination index in combination with MLST to 0.99 [80] [266].
- Lang et al. (2010) used an expanded 16 gene MLST on *C. coli* [269].
- In the age of genome sequencing the MLST approach can be expanded still further and will be discussed in the genomic typing section.

The advantage of molecular sub-typing methods based on the sequencing of one or more gene loci, are minimal experimental variation, enabling direct inter-laboratory comparisons and the establishment of digital databases [141] [270]. The MLST system has also be used to perform phylogenetic, population genetic analysis and multinational longitudinal epidemiological studies [271]. The MLST system can allocate closely related sequence types into a clonal complex which can help understand the population structure of weakly clonal bacteria like *C. jejuni* and is a relevant unit for short and long term epidemiological analysis [272]. Disadvantages of this system include the time, cost and specialised equipment. Also occasionally a PCR may fail to type a loci, probably due to primers failing to anneal to the correct position. When investigating a short term outbreak there may be insufficient differentiation, i.e. not enough time for the accumulation of mutations in the MLST loci, and the use of more rapidly changing loci e.g. *fla*A is required [141].

Another concern, illustrated by the finding of significant differences between isolates of the same MLST sequence type (ST-474), supports a need for typing of *C. jejuni* genomes [273]. Reports based on other pathogens are already appearing supporting a move from MLST to whole genome typing particularly in low divergent species [274][275]. It is also possible to use genomic data on line to identify an isolates MLST sequence type [276]. The MLST system designed for *C. jejuni* by Dingle et al.(2001) is hosted on PubMLST (http://pubmlst.org/campylobacter/) and was used in this thesis [265]. This system is widely used in epidemiological studies and contains a large amount of data worldwide (376,670 sequences and 7504 MLST profiles reported on 21st January 2015) [277].

2.4.4 Whole Genome typing

Currently genomics and its related fields are driven by emerging technology with huge potential but rather than be speculative this literature review will be limited to those genomic typing methods that have had a demonstrated epidemiological application to *C. jejuni*. Methods using the whole genome include PFGE, RFLP, PCR-RFLP, and RAPD and have been discussed previously in the genotyping section as they do not involve genome sequencing.

Sequencing of the whole genome gives researchers greater opportunity to identify and use loci that will provide suitable stability and variation to answer the question they may be studying. The approaches in this area can be divided into DNA microarrays and next generation sequencing (NGS). There are a range of next generation sequencing technologies [278] emerging to sequence the genome and various methods to assemble the genomes [279] [280] including using the transcriptome with RNA-sequencing (RNA-Seq) [281]. NGS has the potential for faster analysis of isolates than is presently the norm [282] for example bench-top sequencing has enabled genomes to be available for isolates within 5 days of culturing [283]. Hall et al (2010) showed that a simple presence/absence analysis of the pan-genomes from six different species provided better resolution than MLST [284].

DNA Microarrays

DNA microarrays (DNA chip, biochip) are solid surfaces covered in DNA sequences (probes) and when a target DNA sequence hybridises to a probe the combination is detected. DNA microarrays have been shown to detect genetic polymorphisms in a range of bacteria including *C. jejuni* [285][286]. Polymorphism was identified particularly in genomic prophage and hypervariable regions (88% of the variation [287]) [288]. Champion et al. (2005) used this method combined with Bayesian modelling to identify two clades, a livestock and non-livestock split [289]. The drawback of this approach is generally the construction or purchase of DNA microarrays and the performance of strain to strain hybridisation experiments is considered too expensive for routine application. There is also a need for a reference sequence for the probe, meaning sequences present in the isolate being tested but not in the probe DNA sequences will not be identified, making selection of the reference sequences very important [285].

ribosomal Multilocus Sequence Typing (rMLST)

The ribosomal protein subunit (*rps*) genes have the advantages required of a typing system i.e. being universally present but differentially variable [290] [291] and is potentially a universal bacterial identification and typing scheme [138]. The rMLST allelic profile is based on 53 ribosomal genes for most bacteria although only 52 in the case of Campylobacterales as they lack *rpm*D [292]. Like the MLST approach on which it is based, the rMLST employs curated reference sequences to identify gene variants and index them using the BIGSdb[4] platform [293]. This system has been shown a discrimination level equivalent to MLST's clonal complexes [293]. rMLST has also shown to have difficulties in differentiating several closely related species [276] including *Y. pestis* versus *Y. pseudotuberculosis* [294] and *M. tuberculosis* versus *M. bovis* [295] and *B. anthracis* versus *B. thuringiensis* [296]. In cases like *B. anthracis* and *B. thuringiensis* most of the genetic difference is contained in plasmids, making the rMLST approach or any approach relying solely on the core genome unlikely to provide the necessary resolution [296][276]. It has been suggested that as the

[4]http://pubmlst.org/rmlst/

rMLST database grows, its performance will improve [276]. rMLST requires the time, cost and specialised equipment for NGS and internet access, however this technology is improving in ability and decreasing in cost.

Whole Genome Multilocus Sequence Typing (wgMLST)

A complete gene by gene approach to genome comparisons, developed along the lines of MLST, called whole genome MLST (wgMLST) but with the potential to include regions between genes if the adjacent loci are defined, has been developed [297] [298]. PubMLST[5] currently has a Bacterial Isolate Genome Sequence Database (BIGSdb) to which the genomes can be uploaded and processed [297] [277]. The BIGSdb is a Linux based on-line server, that can link gene, genomic and phenotypic data, although the phenotypic data is stored as meta-data [298]. This platform allows for flexibility of approach and scalability of databases, an example of the work-flow is shown in Figure 2.3. DNA can be isolated from a bacterial isolate or community sample and sequenced on an appropriate platform, short-read data can be obtained from the Sequence Read Archive (SRA), and assembled genomes can be downloaded from public databases. The SRA[6] at NCBI stores raw sequence data from NGS technologies including 454, IonTorrent, Illumina, SOLiD, Helicos and Complete Genomics. The DNA sequence data, which is stored as text, can be combined with information on its provenance and phenotypic data in the form of metadata, and can then be grouped within different schemes e.g. MLST or rMLST loci if required [277]. The Genome Comparator module [298] can compare sequences to other data such as reference genomes using algorithms such as BLAST [299]. Loci from a particular scheme (e.g. rMLST, MLST, wgMLST) can be compared to the sequence definition database and alleles indexed if previously identified or the definitions expanded as new variants are identified. Although loci are usually protein encoding genes, this approach is not restricted to these and intergenic regions could be included [298].

Analysis comparing MLST and wgMLST using seasonal data from Finland was applied to a milk borne *C. jejuni* outbreak [300] [292]. wgMLST showed increased sorting ability by demonstrating variation within ST-45 and the results suggested that some "sporadic" outbreaks may be linked by a common source of infection [301].

wgMLST is dependent on NGS technology, bioinformatics skills and internet access which for the immediate future are limitations. These technologies are rapidly improving and decreasing in cost and these features could change from limitations to benefits [137]. As NGS is still a developing technology, and issues of quality control can exist at various stages of the processing;quality control differences between technologies at the sequencing stage [302], variation in assembly or mapping methods and their quality [303] [304], can all affect the data entering the database, which in turn can affect the quality of the final analysis.

[5]PubMLST website http://pubmlst.org/campylobacter

[6]NCBI's Sequence Read Archive is at (http://www.ncbi.nlm.nih.gov/Traces/sra/)

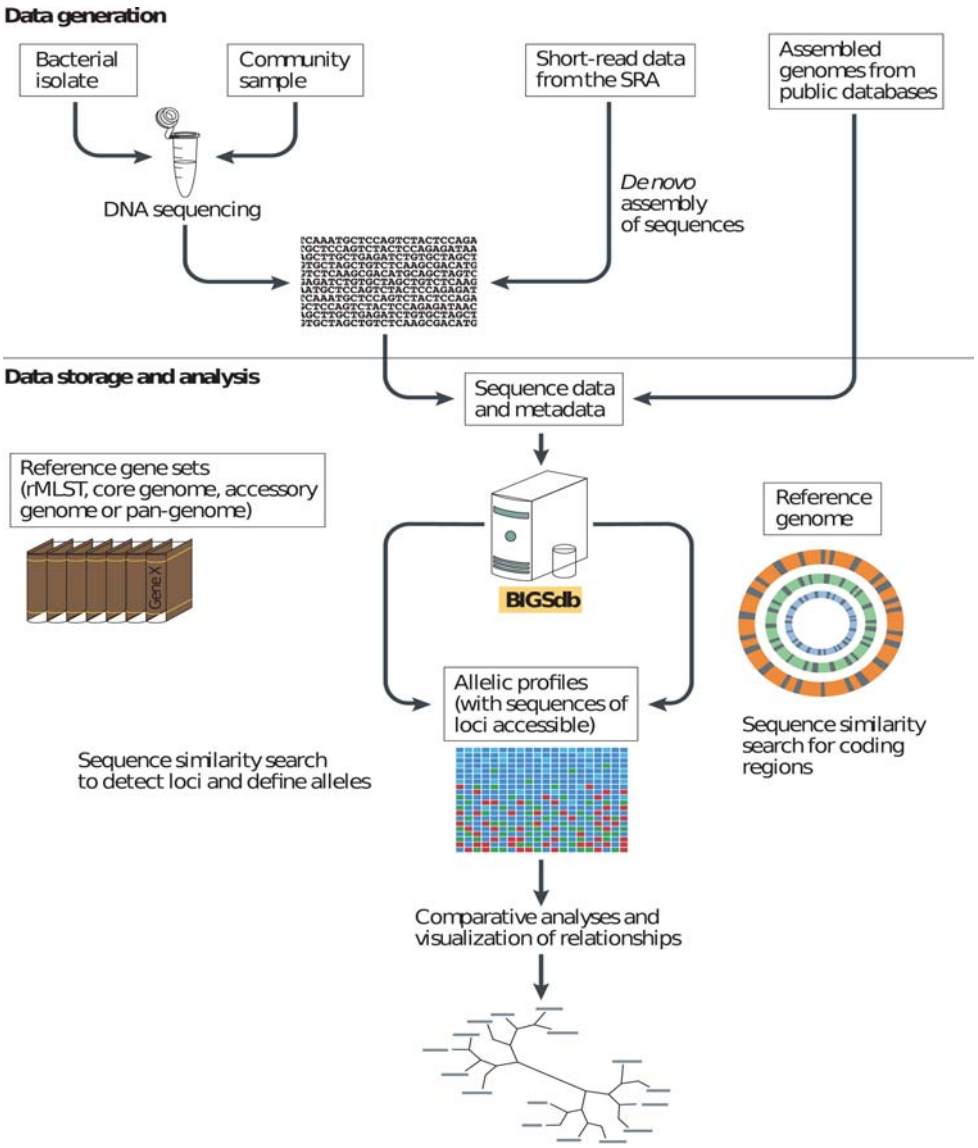


Figure 2.3: **An outline of potential whole genome gene by gene analysis workflows for Bacterial Isolate Genome Sequence Database (BIGSdb).** BIGSdb can input DNA sequences can be derived from bacterial isolates and communities as well as de-novo genomic data from next generation sequencing (NGS) and existing databases. Metadata about a samples origin and phenotype can also be stored in BIGSdb. Loci that are part of a scheme can be identified and alleles are assigned designations to generate an allelic profile. The Genome Comparator module in BIGSdb can be used to produce a distance matrix based on allelic profiles. The source of Figure 2.3 is "MLST revisited: the gene-by-gene approach to bacterial genomics." [298].

2.5 The Host

C. jejuni is a multi-host species isolated from a wide range of warm blooded host species, avian and mammalian. *C. jejuni* has been isolated from water and insects [305] [306], however as a thermophilic bacteria requiring temperatures $\geq$ 30℃ to replicate, these findings suggest the isolates found for example in water are transiting between hosts and while a possible source, it is not a centre of multiplication [307] [308].

Discovery of new *Campylobacter* spp. and hosts

Currently the discovery of new *Campylobacter* species and strains being associated with new hosts is a dynamic and developing field. There are new species and hosts still being reported, for example in the last few years using real-time PCR detection new and existing *Campylobacter* species were reported in 85% of the of free-living Virunga mountain gorillas (*Gorilla beringei beringei*) tested [309] and 16S rRNA PCR showed 11.1% of the sampled Persian fallow deer (*Dama mesopotamica*) had a *Campylobacter* spp. [310]. New putative species have been identified: in captive lion-tailed macaques (*Macaca silenus*) *Campylobacter corcagiensis* sp. nov. [311], *Campylobacter subantarcticus* sp. nov., has been isolated from birds in the sub-Antarctic region [312] and in black-headed gulls (*Larus ridibundus*) *Campylobacter volucris* sp. nov., has been isolated [313]. Campylobacter species are not limited to mammals and birds as they have also been found in reptiles [314]. These results suggest that as we look using modern molecular techniques we will continue to find more hosts and new *Campylobacter* species and strains. There is much to learn about the prevalence of different species and strain types of *Campylobacter* in many wild animals, and by understanding the prevalence and distribution pattern, we will better understand the bacteria and the importance of host association and niche adaptation.

Host attribution of *C. jejuni*

C. jejuni lineages are known to undergo frequent recombination events, but still show a host signature that can be associated either at MLST sequence or clonal complex level [272] [315] [316]. Sequence-based typing such as MLST gives data that is comparable, reproducible and applicable to population genetic analyses [317] [265]. Closely related sequence types can be classified into clonal complex which can help understand the population structure of weakly clonal bacteria like *C. jejuni* and is a relevant unit for short and long term epidemiological analysis [272]. Using the MLST allelic profile, consisting of seven alleles, attribution models have been used to predict the host source [315] [318]. The ability to successfully identify the host from MLST has contributed to the realisation that *C. jejuni* from agricultural species such as cattle, chicken and sheep tend to belong to certain clonal complexes (e.g. CC21, CC42, CC45, CC48) but those from wild-birds tend to be host-specific, and this association trends continues in different geographical areas [318] [7] [319]. The tendency for a division between isolates found in livestock (poultry, cattle, sheep

and to a lesser extent pigs) and wild-birds was also identified by other typing techniques such as BRENDA [227] and PFGE [320] [321]. The serological techniques like Penner was hampered by high levels of unidentified strains particularly in the wild-bird populations, for example, Rosef et al. (1985) could only serotype 65.7% of the isolates from domestic and wild animals [322].

A consequence of the recognition of the host association division into agricultural livestock (cattle, sheep and poultry) and wild birds, is the identification of differing lifestyles between these groups [318] [7] [319]. ST and CC associated with agricultural livestock are spread across mammal (cattle, sheep) and bird (poultry) hosts and referred to as a generalist lifestyle. While ST associated with wild-birds tend to be host specific with a specialist lifestyle.

Wild-birds as hosts

Understanding the role of wild-birds in the epidemiology and the population biology of *C. jejuni* is important to understanding this emerging disease [323]. Wild-birds have been shown to act as a source of campylobacteriosis in people [324], but to a much lesser degree than agricultural animals (sheep, cattle, chickens) [319] [318], based on *C. jejuni* MLST sequence and PFGE profiles [325]. Although not as significant a source of human campylobacteriosis as agricultural associated isolates, wild-bird genotypes could act as a genetic reservoir with the potential to create new food animal strains either by host switching and adapting to the new niche or through genetic transfer by recombination of DNA [319] [6] [17][p266]. If wild-birds do play such a role due to their highly mobile nature, particularly migratory birds, they could act as vectors over long distances [326]. For example they played an important role in the recent spread of West Nile virus and H1N1 avian influenza [327] [328].

Evidence of direct transmission from wild-bird to humans is limited due to minimal contact between the two; however many other bacterial pathogens found in wild-birds follow a more indirect route, such as wild-bird faecal contamination of food and/or water [329]. In a reverse role, Griekspoor et al. (2009) suggested human to bird transmission, when reporting ST-45 in Macaroni penguins (*Eudyptes chrysolophus*) from Antarctica, as ST-45 is normally associated with agricultural animals. [330].

Campylobacter spp. has been found world-wide in wild-birds from different taxa, phylogeny and ecologies [17][p266]. Table 2.2 shows that *Campylobacter* spp. has been reported, so far, in 95 species with varying levels of prevalence within and between species. These surveys were performed from 1983–2014 using a range of approaches making comparisons between them problematic as results are highly dependent on sampling and culturing techniques [17][p266]. For example Kapperud et al. (1983) [331] found *C. jejuni* in 15/35 black-headed gulls (*Larus ridibundus*) but Palmgren et al. (1997) [332] found no *Campylobacter* spp. in the 41 individuals they sampled. Small sample sizes, different isolation methods, and differences in growing and identification techniques could have contributed to the different reported carriage by a species. Such as using only a hippurate

test to identify *C. jejuni* [321] rather than in a PCR based method [333] could have affected the results [323]. The proportion of gases in the microaerobic environment or the addition of hydrogen (H_2), and different incubation temperatures (37℃ vs. 42℃) could also affect the isolation of *Campylobacter* spp. [323] [334]. In some surveys fresh faecal samples were used [335] [318] [336] and it is possible that the same individual was sampled twice, while others sampled direct from the cloaca [337] [338].

There is a large range of variation between survey results for the same species in Table 2.2 for which other explanations are possible aside from differences in sampling and isolation methods. It has been suggested that diet or ecological guild (group of species that exploit the same resources) [323] may affect the occurrence of *Campylobacter* spp. in wild-birds, in particular exposure to human refuse [331] [339]. Other possible explanations for variation in the percentage of individuals identified with *Campylobacter* isolates are seasonal variation [340] and variation of prevalence with age [341].

Table 2.2: **Prevalence of *Campylobacter* spp. and *C. jejuni* isolated from wild-bird species.** The results of multiple surveys involving the same bird species were reported together to produce a range of values for these species. The *Campylobacter* spp. percentage does not include *C. jejuni* isolates where known but some surveys did not identify isolates at the species level.

Bird species	**Common name**	***Campylobacter* spp. (%)**	***C. jejuni* (%)**	**Study**
Family Accipitridae				
Accipiter nisus	Eurasian sparrowhawk	3.1	3.1	[323]
Buteo jamaicensis	Red-tailed hawk		12.5	[342]
Milvus milvus	Red kite	50		[343]
Family Anatidae				
Anas acuta	Northern pintail	100, 2.4	50	[344],[323],[345]
Anas americana	American wigeon		42	[344]
Anas clypeata	Northern shoveler		65.7	[344]
Anas carolinensis	Green-winged teal		16.1, 100	[344],[346]
Anas crecca	Eurasian teal	50		[323]
Anas discors	Blue-winged teal	50		[347]
Anas penelope	Eurasian wigeon	2.3		[333]
Anas platyrhychos	Mallard duck	10	4.8, 33.7, 40, 79	[343],[342],[344],[348][346]
Anas strepera	Gadwall		15.4	[344]
Anser anser	Domestic goose	15	100	[337]
Aythya americana	Redhead		50	[346]
Branta bernicla	Brant goose	25		[323]
Branta canadensis	Canada goose	40	8.2, 11.4,52	[349][350][351][348]
Branta leucopsis	Barnacle goose		10.4, 11.4	[352][353]

Continued on next page

Prevalence of *Campylobacter* spp. in wild-birds —– *continued from previous page*

Bird species	**Common name**	***Campylobacter* spp. (%)**	**C. jejuni (%)**	**Study**
Bucephala clangula	Goldeneye	100		[331]
Cygnus atratus	Black swan	45		[349]
Cygnus olor	Mute swan	3.3		[333]
Somateria mollissima	Common eider	100		[323]
Spatula clypeata	Shoveler		65.7	[344]
Family Alcidae				
Fratercula arctica	Puffin		51.3	[331]
Family Ardeidae				
Bubulcus ibis	Cattle egret	50	50	[347]
Family Cardinalidae				
Pheucticus ludovicianus	Rose-breasted grosbeak		50	[354]
Family Charadriidae				
Charadrius hiaticula	Common ringed plover		33.3	[323]
Pluvialis dominica	Golden plover	33.3		[347]
Pluvialis squatarola	Grey plover	100	100	[323]
Vanellus vanellus	Lapwing	54, 75, 100	46	[355][343][353]
Family Columbidae				
Columba livia domestica	Pigeon		4.3, 12.5	[333][339]
Streptopelia orientalis	Eastern turtledove		1.6	[339]
Family Corvidae				
Corvus brachyrhynchos	American crow		42.8	[342]
Corvus cornix	Hooded crow		89.6, 89.8	[356],[331]

Continued on next page

Prevalence of *Campylobacter* spp. in wild-birds —– *continued from previous page*

Bird species	**Common name**	***Campylobacter* spp. (%)**	**C. jejuni (%)**	**Study**
Corvus corone	Carrion crow	6.25, 39.7	85.4, 50	[331][357]
Corvus frugilegus	Rook	51.2		[343]
Corvus levaillanti, Corvus cornix	Crows		35.5	[339]
Corvus monedula	Jackdraw		17.7, 100	[333] [323]
Corvus ossifragus	Fish crow		40	[342]
Cyanocitta cristata	Blue jay		16.7	[342]
Cyanopica cyanus	Blue magpie		20	[339]
Garrulus glandarius	Eurasian jay		33.3	[333]
Pica pica	Magpie	4.2	2.8	[333]
Family Falconidae				
Falco peregrinus	Peregrine falcon		2.9	[358]
Family Fringillidae				
Carduelis chloris	European greenfinch	3.4		[323]
Emberiza citrinella	Yellowhammer		5.2	[323]
Emberiza hortulana	Ortolan bunting	16		[323]
Emberiza schoeniclus	Reedbunting		20	[331]
Fringilla coelebs	Chaffinch		1.1	[333]
Zonotrichia albicollis	White-throated sparrow	3.6	14.2	[354]
Family Gruidae				
Grus canadensis	Sandhill Crane	48		[335]
Family Haematopodidae				

Continued on next page

Prevalence of *Campylobacter* spp. in wild-birds —– *continued from previous page*

Bird species	**Common name**	***Campylobacter* spp. (%)**	**C. jejuni (%)**	**Study**
Haematotopus ostralegus	Oystercatcher	35.7, 6.3	64.2	[355][333]
Family Icteridae				
Quiscalus quiscula	Common grackle		41.1	[359]
Family Laridae				
Larus argentatus	Herring Gull	30.2, 45	19, 44.2, 100,	[331][341] [360]
Larus atricilla	Laughing gull		33.3	[342]
Larus canus	Common gull	18.9		[331]
Larus delawarensis	Ring-billed gull	15.9		[338]
Larus ridibundus	Black-headed gull	6.4, 62	25, 29.9, 63	[321][341][331][361]
Sterna hirundo	Common tern		5.6	[331]
Family Motacillidae				
Anthus pratensis	Meadow pipit	33.3		[323]
Motacilla alba	White wagtail	31.2	6.2	[323]
Motacilla flava	Western yellow wagtail	33.3		[323]
Family Muscicapidae				
Turdus iliacus	Redwing		50	[323]
Turdus merula	Blackbird	29.5	1.9, 3.2, 15.9	[323][333][332]
Turdus migratorius	American robin		3.3	[342]
Turdus philomelos	Song thrush	4	28	[323]
Turdus pilaris	Fieldfare		14	[323]
Turdus viscivorus	Mistle thrush		100	[323]
Family Paridae				

Continued on next page

Prevalence of *Campylobacter* spp. in wild-birds —– *continued from previous page*

Bird species	**Common name**	***Campylobacter* spp. (%)**	**C. jejuni (%)**	**Study**
Parus major	Great tit	10	0.7	[323],[333]
Family Passeridae				
Passer domesticus	House sparrow	0.9	1.2,1.8	[333][354]
Passer montanus	Eurasian tree sparrow		2.6, 60	[345]
Family Phasianidae				
Coturnix coturnix	Quail	21.4	8.6	[362]
Family Prunellidae				
Prunella modularis	Dunnock		7.7	[323]
Family Pycnonotidae				
Hypsipetes amaurotis	Bulbul		11	[339]
Family Rallidae				
Fulica atra	Eurasian coot	25		[343]
Gallinula chloropus	Common moorhen	100	33.3	[343][333]
Family Regulidae				
Regulus regulus	Goldcrest	1		[323]
Family Scolopacidae				
Actitis hypoleucos	Common sandpiper	100, 100		[323],[353]
Arenaria interpres	Turnstone	3.7, 44.9, 75	55.1	[355][343][323]
Calidris alpina	Dunlin	70.6, 75	9.3, 9, 100	[323][342][353] [7]
Calidris canutus	Red knot	7.1, 66.6		[323][333]
Calidris ferruginea	Curlew sandpiper	88.9	11	[323]
Calidris minuta	Little stint	62.5	25	[323]

Continued on next page

Prevalence of *Campylobacter* spp. in wild-birds —– *continued from previous page*

Bird species	**Common name**	***Campylobacter* spp. (%)**	**C. jejuni (%)**	**Study**
Calidris pusilla	Semi-palmated sandpiper	33.3		[347]
Calidris temminckii	Temminck's stint	50		[323]
Gallinago gallinago	Common snipe	100, 100		[323][347]
Limicola falcinellus	Broad-billed sandpiper	72	36	[323]
Lymnocryptes minimus	Jack snipe	100		[323]
Micropalama himantopus	Stilt sandpiper	75	50	[347]
Philomachus pugnax	Ruff	100	100	[323]
Scolopax rusticola	Eurasian woodcock	100		[323]
Tringa flavipes	Lesser yellowleg	53.8	38.5	[347]
Tringa glareola	Wood sandpiper	16.6	16.6	[323]
Tringa melanoleuca	Greater yellowleg	100		[347]
Tringa nebularia	Common greenshank	100		[323]
Tringa totanus	Redshank	51.9, 86	0.8, 48.1	[355][353]
Family Spheniscidae				
Eudyptes chrysolophus	Macaroni penguin		3	[330]
Family Strigidae				
Asio otus	Long-eared owl		20	[323]
Strix uralensis	Ural owl		100	[331]
Family Sturnidae				
Sturnus cineraceus	Gray starling		14.3	[339]
Sturnus vulgaris	Starling	6.9, 11.1, 12.9, 41	3.2, 11.1, 18, 29.8, 32.3	[323][332][333][340],[318][363]

Continued on next page

Prevalence of *Campylobacter* spp. in wild-birds —– *continued from previous page*

Bird species	**Common name**	***Campylobacter* spp. (%)**	**C. jejuni (%)**	**Study**
Family Troglodytidae				
Troglodytes troglodytes	Eurasian wren	2.7	2.7	[323]
Family Tytonidae				
Tyto alba	Barn owl		4.2	[333]

Table 2.2 is a summary of the available reports on *Campylobacter* spp. or *C. jejuni* isolated from various wild-bird species although some reports of domestic birds were included and some wild-birds excluded. Two surveys were included that contained domestic pigeons [339] or domestic geese [337], however, in both cases these are free range animals exposed to a diverse environment and diet which is more similar to wild-birds than that encountered by intensively farmed animals. A survey by Yogasundram et al. (1989) and another by Petersen et al. (2001) were not included as they are based on birds submitted for necropsy, rather than live healthy birds [364] [320]. Some surveys did not identify by species, but grouped at under a common name such as "gulls" [349] [340] [365] [366] [349] [363]. These were not included in the table as identity to the species level may be important .

Sometimes negative results i.e. no *Campylobacter* spp. isolated in a given population sampled were reported for a wild-bird species in a survey (for example [332] [323] [339]). These results have not been included as Table 2.2 as this review is focusing on where *Campylobacter* spp. in general and *C. jejuni* in particular is found in wild-birds. It is important to acknowledge that *Campylobacter* spp. were not found in all birds surveyed. Waldenstrom and Griekspoor (2014) in their review of studies looking for *Campylobacter* spp. in wild-birds by species showed 58.5% (120/205 identified to the species level not a genus group) with zero *Campylobacter* spp. isolated in any reported study [17][pp267-276]. However 52.5% (63/120) of the results that failed to find *Campylobacter* spp. came from very small sample sizes (≤ 5 birds), indicating the result may be affected by stochastic variation rather than confirmation of zero prevalence.

A potentially important implication of a negative finding is the lack of *C. jejuni* isolated from six sites consisting of eight Antarctic bird species (233 samples in total). This may suggest increased human intrusion poses an increased risk of transmission to potentially naive species/populations [367]. A transmission event may have already occurred from human to Macaroni penguins [330]. Little is known about *C. jejuni*'s potential as a wild-bird pathogen beyond a report of weight loss in a European Robin (*Erithacus rubecula*) artificially inoculated with a strain from the Song Thrush (*Turdus philomelos*) [124].

Genomic analysis of wild-bird isolates

MLST typing has been the gold standard for *C. jejuni* epidemiological studies, and while it has helped with source attribution, it does not deal directly with the accessory genome which could identify features that may contribute to niche adaptation [368]. Some comparative genome analyses using datasets containing wild-bird isolates have been published. Champion et al. (2005) used genomotyping (whole-genome comparisons of microbes using DNA microarrays) to compare 111 *C. jejuni* strains (70 human, 17 chicken, 13 bovine, 5 ovine, 6 environmental) and Bayesian analysis of the phylogeny showed a livestock and non-livestock clade [289]. The analysis correctly placed 88.6% (31/35) of the livestock isolates in the livestock clade [289]. However this approach also

placed 55.7% (39/70) of the human isolates in the non–livestock clade [289], rather than attributing a high percentage of human isolates to the livestock clade (agriculture) as many MLST studies have [315] [318] [7] [319]. Hepworth et al. (2011) compared 80 isolates, (16 birds/duck/water, 15 wild mammals, 49 agricultural/human isolates) using comparative genomic hybridisation (CGH) to show that genomic variation existed between the water/wildlife group and the agricultural group, and within the water/wildlife group [368]. They sequenced two of the wild mammalian isolates, but not a wild-bird isolate [368]. Hepworth et al. (2007) used suppression subtractive hybridisation to compare genomes by identifying sequences not present in NCTC11168 [369], isolates were used from wildlife but limited to common clonal comlexes (CC-21, CC42, CC45 and CC61). In the 2007 analysis 7 of the isolates were identified as having a "bird" host, the results showed a correlation between clonal complex and metabolic gene but not a host preference [369]. Sheppard et al. (2013) performed a genome wide analysis using an association mapping method to compare 192 isolates (including 2 water, 3 starling, 4 goose, 5 ducks) and found a seven-gene region in most cattle isolates but often absent in chickens or wild-birds isolates [370]. Three of the seven genes encoded for vitamin B_5 bio-synthesis [370], and the absence of these genes had previously been observed in environmental isolates [371][p93].

In summary there is limited information on the types and distribution of *Campylobacter* spp. in general and *C. jejuni* in particular in wild-birds. There is also little genomic based analyses investigating wild-bird isolates and the genomic variation they contribute to the *C. jejuni* pan-genome [368] and even less genomic comparison involving Australasian isolates with most analysis involving Northern Hemisphere isolates.

2.6 Temperature and *C. jejuni*

This is a review of the effect of temperature on *C. jejuni* and the range of temperatures it can experience inside different hosts. An intestinal bacteria associated with many different host species *C. jejuni* is exposed to and must survive transmission through a wide range of environments and temperatures. As more is found out about *C. jejuni* it is clear there are major differences between it and *E. coli* the standard organism considered when designing food hygiene standards [372], making an understanding of its response to different temperatures an important public health question.

2.6.1 Host body temperature

C. jejuni has been isolated from a range of species, both mammalian and avian, but poultry is the major source of human infection [2] [91] [92]. The optimal laboratory incubation temperature of 42℃ – 43℃ for *C. jejuni* [373] [89] is often associated with poultry (*Gallus gallus*) which has a reported normal rectal body temperature of 40.6℃ – 43.0℃ [374] [373].

It has been suggested a birds average cloacal reading is between $39.8℃ - 43.3℃$ [375]. An exception is the nocturnal kiwi (*Apteryx spp.*) with daytime levels of $36.4℃ - 37.2℃$ and night time levels of $38.2℃ - 39.9℃$ reported in the North Island Brown (*Apteryx mantelli*) [376]. This variation is more consistent with Prinzinger et. al. (1991) findings that there is a wide range of normal avian body temperatures and the bird's physiological state needs to be considered as an important contributing factor [377]. The results for the mean body temperature in 86 bird species when grouped into three levels of activity show: at rest $38.54℃ \pm 0.96℃$, in an active phase $41.02℃ \pm 1.29℃$, and in a highly active phase $43.85℃ \pm 0.94℃$ [377]. These results suggest that when inside an avian host *C. jejuni* could experience a wide range of temperatures.

Compared to birds, mammals are generally associated with lower core body temperatures although there is variation within species due to different physiological states and between species. During a fever, the mammalian core body temperatures is generally in the range of $38℃ - 41℃$ [378]. A notable mammal is the bat, because when in flight the body temperature is raised and depending on the species normal body temperatures ranges from $35℃ - 42℃$ [379]. Recently a *Campylobacter* spp. was isolated in some European bats [380]. It has been suggested flight in bats and the resulting temperature increase may act as a selective force on some of the pathogenic viruses they can carry (e.g. Ebola, Marburg, Rabies) [379].

C. jejuni's pathogenic behaviour in people rather than poultry has been attributed to a difference in host body temperature [373]. A difference in gene expression profiles has been reported at 37℃, a human associated temperature, compared to 42℃, a poultry associated temperature [373]. Hofreuter et al. (2014) suggested gluconate dehydrogenase GADH, which is more active at 42℃ than 37℃, may be an avian host adaptation [212]. However this review of avian body temperatures would suggest the association of a consistent 42℃ body temperature to an avian host may be an over-simplification. It is possible in parallel with the bat body temperature hypothesis of OShea et al. (2014) that the fever response of a human host may be less effective against a pathogen like *C. jejuni* adapted to a host with a high body temperature rather than the generally lower human body temperature triggering increased pathogenicity [379] [373].

2.6.2 The effect of temperature on growth

C. jejuni is a thermophilic and commonly growing at temperatures ranging from $\sim 30 - 47℃$, with an optimal growth temperature of $42℃ - 43℃$ [373] [89]. Growing thermotolerant *Campylobacter* spp. *in vitro* at 42℃ is both to enhance their growth and repress the growth of other bacteria [334]. However sampling and culturing methods used in an investigation can unwittingly bias the results of an investigation and recent work suggests when sampling from the environment for *Campylobacter* spp. isolates incubating the samples at 37℃ as well 42℃ should be considered, especially if non *C. jejuni* are being investigated [381].

Most other bacteria have a gradual growth rate decline as they near their minimal growth tempera-

ture with the growth rate being linearly proportional to the square root of the growth temperature minus the minimum growth temperature [382]. In contrast *C. jejuni* has an abrupt transition from the growing state (actively dividing cells) to an alive but not growing state [89].

The reported minimal growth temperature, determined by experimentation, for *C. jejuni* is ~ 30℃ with some variation possibly due to strain variation:

- $31 - 32$℃ using a sewerage and a bovine isolate [383]
- $31 - 32$℃ using five human isolates [89]
- $34 - 36$℃ [384]

While the exact minimal growth temperature may vary between isolates, it is agreed that there is no growth at 25℃ by *C. jejuni* [385], or thermophillic *Campylobacter* spp. ≤ 28℃ [27].

Although *C. jejuni* subsp. *doylei* has been reported growing down to 30℃± 0.5℃ [386], no report of growth at less than 30℃ could be found for *C. jejuni* subsp *jejuni* [383] [89]. However a "phoenix" effect has been reported at 30℃ when *C. jejuni* isolates initially growing at 39℃ were inoculated into fresh media and incubated at 30℃ resulting in an initial increase, then, a decrease, followed by a later increase in viable counts [387].

2.6.3 The effect of high temperatures

Exposure to heat stress can induce in bacteria two types of heat shock proteins which are normally conserved; the chaperones and the ATP dependent proteases which repair and prevent damage in the folding of proteins [388]. A range of proteins and genes have been identified that are believed to be fulfilling various roles in protection from increasing temperatures. Homologues for heat shock proteins have been found in the *C. jejuni* genome e.g. *Gro*EL, *Gro*ES, *Clp*P, *Dna*K, *Dna*J, *Grp*E, *Hrc*A, *Htr*A, and *Deg*P [389]. The *C. jejuni* genome lacks the sigma (σ) factor RpoS which induces many changes in other Gram negative bacteria in the stationary phase, including greater heat resistance, but *C. jejuni* has demonstrated an ability to respond to a 46℃ heat shock by inducing 24 proteins [388]. Up regulation was also detected in 10 heat shock proteins (*Gro*EL,*Gro*ES, *Grp*E, *Dna*K, *Dna*J, *Hsp*R, *Cbp*A,*Hrc*A, *Clp*B, *Lon* and a putative *Hsl*U gene), using whole genome array analysis in response to a lesser increase in temperature from $37 - 42$℃ by *C. jejuni* [373]. A potential regulatory role in temperature response by *C. jejuni* has been identified for the *Rac*RS regulon which is differentially expressed at 37℃ compared to 42℃ [390], and are homologues of *Hrc*A and *Hsp*R, which are negative regulators of heat shock in *B. subtilus*.

Heat can cause structural changes in *C. jejuni* with an increase in temperature from 37℃$ - 42$℃ showing a differential expression of some membrane structural proteins [373]. Greater levels of heat stress accelerates the transition from *C. jejuni* normal motile spiral shape to a coccoid one and temperatures > 55℃ rapidly result in cell death which is attributed to the unfolding of the

most thermally labile regions of the ribosome [391] [392] [393].

Some bacteria have mechanisms for sensing environmental changes in temperature using riboswitches and RNA thermometers (RNAT) [394]. Both riboswitches and RNA thermometers use the folding ability of RNA to form complex secondary and tertiary structures as the mechanism to enable regulation [395]. Riboswitches recognise molecules and suppress or activate the genes that control fundamental biosynthesis pathways [396] [397], while RNA thermometers sense temperature to control heat shock response and the expression of some virulence genes [398] [399] [400]. RNAT's control the translation initiation of heat shock or virulence genes by forming a secondary structure that traps the ribosome binding site (RBS) and when there is an increase in temperature, for example, to 37℃ in the case of virulence genes or higher for heat shock genes, the structure is destabilised liberating the RBS, permitting the formation of the translation initiation complex, and the expression of the gene [395]. A report investigating RNATs within the *C. jejuni* genome was not found.

2.6.4 The effect of low temperature

C. jejuni does not respond to lower temperatures like *E. coli* [401] Most of our understanding of bacteria's response to low temperatures is based on the model organisms *E. coli* and *B. subtilis*, for example the production of cold shock proteins (*Csp*) in response to sudden decreases of temperature [402]. *Csp*A regulates at the level of transcription, translation and mRNA stability to control the response to cold, allowing the bacteria to survive and remain metabolically active during long periods of refrigeration [403] [404]. *Csp*A, although wide spread in bacteria is not found in archaea, cyanobacteria or the *C. jejuni* genome [383].

Less is known about the genes involved in *C. jejuni*'s response to cold than to heat. A role has been suggested for *sod*B and *kat*A in freeze thaw survival [405][pp571–590], and polynucleotide phosphorylase (PNPase) has been reported as important to long-term survival at low temperatures [406]. A potential difficulty in measuring metabolic activity at low temperatures could be due to cells changing into the viable but non culturable (VBNC) state when the amount of gene expression is almost undetectable [407].

C. jejuni can be active at lower temperatures, as oxygen consumption, catalase activity, ATP generation, and protein synthesis have been reported down to 4℃ in *C. jejuni* [383]. *C. jejuni* has the ability to take up "naked" DNA from the environment. This natural transformation has been shown to occur in aerobic and microaerobic conditions at temperatures of 20℃, 30℃ and 37℃ , which suggests this process can occur in the environment outside the host [408].

Using DNA microarray and FT-IR spectroscopy on the *C. jejuni* NCTC 11168 isolate at low temperatures, Moen et al. (2005) showed that gene expression was in general downgraded compared to normal growing conditions, and at 25℃ it showed more down regulated genes than at 5℃ [409].

The few up regulated genes included some involved in energy metabolism and the cell envelope [409]. The results lead the authors to suggest this was a survival mechanism, with as many genes as possible down regulated to save energy while genes involved in energy metabolism and modification of the cell wall are up regulated [409].

Membrane adaptation to low temperature by bacteria often results in changes in membrane composition involving (unsaturated) fatty acid and polysaccharide content [410]. In response to low temperatures (6℃), *C. jejuni* did not change its fatty acid composition unlike *E. coli* [411]. In contrast, the cell membrane fatty acid composition varied at 25℃ of *C. jejuni* when the coccoid shape composition was compared to the normal spiral form [87]. The cell lipid composition has also been reported as changing in response to lower temperatures (5℃ and 25℃)[409].

A meta-analysis of *C. jejuni* survival from $0-42$℃ showed greater survival at lower temperatures [10]. *C. jejuni* has shown some variation in its ability to survive at low temperatures in work by Chan et al. (2001) when they tested 19 isolates (10 clinical and 9 poultry) for survivability but they were not MLST typed [403]. The clinical isolates showed better survival ability at 4℃ than non-clinical (poultry) isolates, but the variation of tolerance for freezing at –20℃ was not attributable to either source [403]. Moen et al. (2005) identified that the clinical isolate *C. jejuni* NCTC 11168 isolate survived better at 5℃ than 25℃ [409]. Recently it was observed that *C. jejuni* and *C. coli* could survive cryogenic freezing, indicating a strong ability for low temperature tolerance [412]. Murphy et al. (2006) suggested *C. jejuni*'s ability to survive at very low temperatures maybe due to an uncommon mechanism [126].

There are many gaps in our understanding of *C. jejuni*'s response and survival at various temperatures. Although *Campylobacter* is known to survive and respond to changes in temperature little is known about its behaviour at low temperatures (< 30℃) [413] [387], when it abruptly stops multiplying [383], beyond a greater survivability at lower temperatures [403] and that the mechanism involved may be novel *E. coli* [126]. There is little information in the literature identifying changes in metabolic activity such as carbon metabolism between the active growing state at a host body temperature ($37-42$℃) and at other temperatures such as room temperature ($18-25$℃). Nor is there much information about isolate metabolic variation. Although 42℃ is often reported as the avian body temperature and the optimal growing environment for *C. jejuni*, this review shows there is variation among the different bird species for body temperature, especially in relation to physiological exertion such as flying. Research into how *C. jejuni* adapts to these environmental changes may help us to understand how to better manage *C. jejuni*, particularly in food processing environments.

2.7 Conclusion of the Literature review:

Phenotypic and genotypic variation between *C. jejuni* isolates is a common finding [132][p59] and genetic analysis has shown there are significant differences between the isolates from agricultural animals (sheep, cattle and poultry) and those found in wild-birds [368] [414] [6]. There is little information on isolates associated with wild-birds probably because they are less associated with disease in humans, although this lower level of source attribution to wild-birds could be due to less risk of transmission rather than less pathogenicity [414]. Multiple reports show *C. jejuni* and *Campylobacter* spp. being isolated from wild-birds (Table 2.2)(over 40% of species tested in studies), involving a wide range of different species . Variation in carriage within species and possibly predisposition to carry the bacteria due to diet (guild) and age have been reported [339] [341] [323]. The evidence suggests that wild-bird strains of *C. jejuni* tend to be much more host specific than the more investigated strains related to domestic animals such as cattle, sheep and poultry [7] [319]. This could mean that there are many as yet unidentified strains, with as yet unknown genotypic and phenotypic potential in the worldwide wild-bird reservoir. Investigations of *C. jejuni* show evidence that recombination is more frequent than mutation [415] and that recombination between more similar organisms occurs more frequent than with less similar [416]. Further investigation of these wild-bird associated strains is important to understand the potential they may have to provide genomic sequences for other readily recombining *C. jejuni*, for example could they act as a source of virulence factors for a poultry associated strain which in turn could more readily get into the food chain and be a major public health problem?

C. jejuni is a fastidious intracellular gastrointestinal pathogen that must survive transmission through the environment, which entails exposure to a range of temperatures. It has been shown that *C. jejuni* may not grow at these lower temperatures, but it does survive them [10] [403]. There is limited information on the behaviour of *C. jejuni* at lower temperatures, such as room temperature, beyond its survival and inability to replicate. Addressing questions about *C. jejuni* phenotypic variation at different temperatures especially below ≤ 30℃ have barely been addressed particularly using wild-bird bird isolates rather than those associated with human cases.

The past is a foreign country; they do things differently there.

L.P. Hartley, The Go-Between, 1953.

3

the past:

Historical demography of livestock importation into New Zealand

3.1 Prelude

This chapter investigates the historical importation of livestock into New Zealand, as cattle sheep and poultry act as hosts for *C. jejuni*. New Zealand's long geographical isolation resulted in a predominance of avian fauna prior to the arrival of mankind. Many of these birds were endemic, and the only terrestrial mammals were two bats species. The arrival of the Polynesians ($\sim$ 800 years ago) [417], bringing with them the kiore, a Pacific rat (*Rattus exulans* (Peale)) [418], and the kuri (*Canis lupus familiaris*), a Polynesian dog [419], started ecological and environmental changes that were accelerated by the arrival of Europeans $\sim$ 200 years ago [420][Pg 127]. Into an environment that had been evolving in relative isolation came the introduction of animals like cattle, sheep and poultry as part of the rise to dominance of pastoral agriculture, resulting in momentous

environmental changes.

As a multi-host gastrointestinal bacteria that is not known to reproduce outside the host [307], the successful arrival of *C. jejuni* to a new country can be associated with the arrival of the host animal. Sequence types (ST) of *C. jejuni* have been generally divided into those associated with agriculture and livestock, such as CC21, CC45 and ST-42; and those associated with wildlife, wild-birds in particular [272] [7] [421]. The sequence types associated with wild-birds appear to be specific to a wild-bird species or closely related species [7] [17]; so their arrival and spread into a new country would depend on migration or introduction either by self or aided by humans. However the ST associated with livestock i.e. cattle, sheep and poultry, appear to have a more generalist lifestyle and spread between these hosts, although there may be some degree of host preference within this group [421].

In this chapter, we addressed the big question of how did *C. jejuni* get to the geographically isolated New Zealand? There are many possible routes, including migratory birds and sea birds that travel long distances. However this question, addressed in Chapter 3, was refined to consider livestock only, as livestock were not present in New Zealand prior to the arrival of the Europeans, so their arrival pattern should reflect the arrival pattern of the *C. jejuni* ST associated with cattle, sheep and poultry. *C. jejuni* has a population structure impacted by four evolutionary processes; point mutation, recombination, selection (both positive and negative), and demographic processes such as growth , reduction and admixture [422]. While the exact timing and size of the bacteria's arrival in New Zealand may not be directly estimated, knowing the host association of ST and the arrival pattern of those hosts should reflect it.

This quantification and timing of the livestock introduction into New Zealand in this chapter has been used in to produce coalescent models of demographic growth for *C. jejuni* in New Zealand [423]. Although it is reasonable to assume the pre-existing prolific avian wildlife of New Zealand did not carry the *C. jejuni* ST, over time their were other arrivals that may have contributed to the introduction of these ST. It is not possible to rule out either the early Polynesians, kiore (*Rattus exulans* (Peale)), and the kuri (*Canis lupus familiaris*) as sources of *C. jejuni* nor the introduced 32 mammals (e.g. rats, ferrets, rabbits, deer) that established themselves in New Zealand [424] as sources of the ST we associate with cattle, sheep and poultry. The arrival of pigs (*Sus scrofa*) also has not been documented in this chapter, and while it is more strongly associated with *C. coli* than *C. jejuni*. Some *C. jejuni* sequence types (STs) and clonal complexes (CC) associated with livestock e.g. CC21 and CC42 have been isolated from pigs PubMLST[1].

Australia, as a another country lacking native livestock and recently colonised by Europeans is also likely to have reliable records of the pattern of livestock introduction which could be used in a similar manner. While countries in Africa or Eurasian have had livestock for a long time, both

[1]Campylobacter PubMLST http://pubmlst.org/campylobacter/

domestic and wild, probably the pattern of the arrival of the livestock during their long histories has not been well recorded so the approach used in this chapter is unlikely to be productive. However, there are methods of tracing geographical spread of a host by the phylogenetic relatedness of the micro-organisms they carry which may spread some light on the general patterns of mixing [425] [426].

The basis of this chapter is an article [427] written by the candidate, which has been reproduced here in its entirety. This article forms the basis of this chapter and has been published by Taylor & Francise [2] in the New Zealand Veterinary Journal [427]. The article has been cited [423]. This article contributes to a body of work from the mEpiLab group of Massey University regarding the introduction of livestock to New Zealand and its effect on the subsequent evolution of the microbes that accompanied them.

[2] http://www.tandfonline.com/

Quantification of historical livestock importation into New Zealand 1860 – 1979

3.2 Abstract

AIMS: To quantify the numbers of live cattle, sheep and poultry imported into New Zealand and, where possible, their country of origin from 1860 to 1979.

METHODS: Information on the origin and number of live animal importations into New Zealand was collected for cattle, sheep and poultry for the period 1868 – 1979 from the annual reports compiled by the New Zealand Registrar Generals Office, Government Statisticians Office, Census and Statistics Office, Census and Statistics Department, Customs Department and Department of Statistics. Census data from 1851 to 1871 were also used to estimate the livestock population during this period. The number of animals imported and the mean population for each species in a decade were determined, and the major countries of origin were identified.

RESULTS: A large number of cattle (53,384) and sheep (604,525) were imported in the 1860s, and then there was a marked reduction in importations. Live poultry were imported in relatively small numbers (20,701) from 1880 to 1939, then 1,564,330 live poultry were imported between 1960 and 1979. Australia was the predominant country of origin for sheep between 1868 and 1959 (51,347/60,918; 84.3%) and of cattle between 1868 and 1979 (10,080/15,157; 66.5%). Only 6,712 (11.0%) sheep and 3,909 (25.8%) cattle were imported from the United Kingdom over the same periods, and even fewer from other countries.

CONCLUSIONS: The collated data and historical reports show that from 1860 to 1979 Australia has been the main source of livestock introduced into New Zealand. The pattern of importation showed that large numbers of cattle and sheep were initially imported in the 1860s, probably in response to rapid agricultural expansion. Thereafter importations continued at much reduced numbers. In contrast, relatively small numbers of poultry were introduced until the 1960s when large numbers were imported as part of the development of a modern high-production industry. The overall pattern for both cattle and sheep was of a bottleneck event, as initially a relatively limited number of animals arrived from outside populations, followed by population expansion with ongoing but limited immigration (admixture). Investigation into the genetic population structure of New Zealands cattle and sheep, as well as their host-associated microorganisms, could reflect the impact of these early historical events.

3.3 Introduction

A key step towards a better understanding of the genetic population structure of livestock in New Zealand and their host-associated microorganisms (commensal, pathogen and parasite) is quantification of the pattern of importation of these animals. Previous work on the historical demography suggests that Australia was the main source of early livestock importation [428] [429], although others suggested that there were three main countries of origin of animal introductions to New Zealand, namely the United Kingdom, Australia and Chile [430] [431]. No previously published research quantifying these importation movements from colonial to modern times has been found by the authors. By examining the annual government records ($1840 - 1980$) it should be possible to quantify the size and timing of these important demographic historical events for cattle, sheep and poultry in New Zealand.

The importance of a quantitative analysis of livestock populations and their movements is not limited to the animals themselves, but includes the microbial organisms they carry. Microorganisms that have host-associated strains, such as *Campylobacter jejuni*, may have had their population structure in New Zealand influenced by these historical patterns of live animal importation. The movement of pathogens, parasites and their hosts may affect national biosecurity, global trade and control of emerging infectious diseases [432]. Recent developments in molecular biology, and specifically DNA sequencing, have facilitated detailed investigation into the geographical distribution and biodiversity within populations of microorganisms [433]. Phylogenetic analysis investigates the evolutionary relationships of organisms [434] which, in combination with population genetics modelling, has the potential to show how past events may have affected the development of a particular microorganism, parasite or host population [435] [425] [436].

Modern New Zealand is an isolated archipelago, and one of the last places on earth that mankind colonised with domesticated animals (reviewed in [437]). Polynesian voyagers, who travelled around the Pacific with a variety of domestic animals, arrived about 800 years ago in New Zealand and introduced the kiore (*Rattus exulans*) and the kuri (*Canis familiaris* Linnaeus) [419]. Early European explorers, starting with Captain James Cook, began introducing sheep and other domesticated animals to New Zealand [438] [439]. The Europeans introduced a system of pastoral agriculture and, from 1840 onwards, early settlers accelerated the rate of livestock introduction [440] [441].

Very early in the introduction of sheep to New Zealand sheep scab (*Psoroptes ovis*) was introduced [442] [443] [444]. Sheep scab was reported in New Zealand by 1840 [445] and government mandated control measures were instigated in the Scab Ordinance of 1849 [446] [447]. New Zealand was declared officially free of sheep scab in 1892 [448]. It has been suggested that strict quarantine measures were rapidly developed in New Zealand [431], but others report a more slow and progressive development, from almost no measures to the modern system [443]. A slower development of

national measures may stem from the scientific understanding of microorganisms as a cause of disease (germ theory) only emerging in the late nineteenth century.

The aim of this study was to quantify the numbers of live cattle, sheep and poultry imported into New Zealand from 1860 to 1979 and, where possible, identify their country of origin. A better understanding gained from determining the numbers, timing and origin of these importations could clarify their contribution to the current livestock population structure in New Zealand, and the population structure of their host-associated microorganisms (commensal, pathogenic and parasitic).

3.4 Materials and methods

Data on live animal importation were collected by reviewing and combining information from official New Zealand records, from 1840 onwards, on the importation of sheep, cattle and poultry. From 1840 to 1920 the Blue Book series was the general title given to a series of annual statistical reports produced by various sections of the New Zealand government. The publication of live animal importations data as annual books continued from 1920 to 1979 under various names. The data for $1860 - 1867$ did not include the source of the imports, only the total number imported. Table 3.1 shows the names of the annual official record for a given period and the government body that compiled them. In most years there was an individual report for each year, but occasionally the report was printed as a combined report for two or three or even four consecutive years. A detailed list of all publications can be found in [449].

From 1868 to 1879 poultry importations were not consistently quantified as individual birds in annual official records, therefore these imports were not used in this study. Also, from 1942 to 1948 the records for importing poultry quantified them by total annual value, not by number of individual animals, so these years were also not included in the study.

Census data used to estimate livestock populations in New Zealand between 1851 and 1871 were from reports in Statistics of New Zealand for 1861 [450] and the New Zealand Official Year-book for 1919 [451]. Estimates for population numbers of cattle, sheep and poultry between 1860 and 1979 were derived from reports in Statistics of New Zealand for 1861 [450], Agricultural and Pastoral Statistics of New Zealand 18611954 [452], New Zealand Official Year-book 1919 [451], and Statistics New Zealands Infoshare on-line database [3] [453].

Statistical analysis

The number of live cattle, sheep and poultry reported in the annual official records as being imported each year were summed to provide a total for each decade and combined with data on the origin of

[3] http://www.stats.govt.nz/infoshare/Default.aspx

livestock to estimate the proportions of imports attributable to Australia, the United Kingdom and other countries. The mean population for each decade was calculated from the available estimated population size for a given year using R v2.12 (R Development Core Team 2008, R Foundation for Statistical Computing, Vienna, Austria).

Table 3.1: **The titles of the annual statistical report containing live animal importations from 1853 to 1979 and the compiling office of the New Zealand government.**

Year	Title	Compiled by	Section	Published
1853-1872	Statistics of New Zealand	Registrar Generals Office	3.1	Auckland
1873-1906	Statistics of the Colony of New Zealand	Registrar Generals Office	3.2	Wellington
1907-1908	Statistics of the Dominion of New Zealand for the Year	Registrar Generals Office	3.3	Wellington
1909-1920	Statistics of the Dominion of New Zealand for the Year	Government Statistician's Office	3.3	Wellington
1921	Statistical report on the trade and shipping of the Dominion of New Zealand for the year 1921	Census and Statistics Office	17.1.2	Wellington
1922-1926	Statistical report on the trade and shipping in the Dominion of New Zealand for the Year	Census and Statistics Office	17.1.3	Wellington
1927-1932	Statistical report on the trade and shipping in the Dominion of New Zealand	Census and Statistics Office	17.1.4	Wellington
1933-1935	Statistical report on the trade and shipping of the Dominion of New Zealand	Census and Statistics Office	17.1.5	Wellington
1936-1944	Statistical report on the trade and shipping of the Dominion of New Zealand	Census and Statistics Department	17.1.5	Wellington
1945-1948	Statistical report on the trade and shipping of New Zealand for the Years	Census and Statistics Department	17.1.7	Wellington

Continued on next page

Annual statistical reports – *Continued from previous page*

Year	Title	Compiled by	Section	Published
1949	Statistical report on the external trade of New Zealand for the 1949. Part 1B, Imports	Customs Department	17.1.11	Wellington
1950-1951	Statistical report on the trade and shipping of New Zealand for the Years	Census and Statistics Department	17.1.10	Wellington
1952-1954	Statistical report on the external trade of New Zealand for the years. Part 1B, Imports	Customs Department	17.1.11	Wellington
1955-1961	Statistical report on the external trade of New Zealand for the year. Part 1B, Imports	Customs Department	17.1.12	Wellington
1962	Statistical report on the external trade of New Zealand for the six months January June 1962. Section 2, Imports	Customs Department	17.1.13	Wellington
1962-1972	Imports: Part A Commodity by Country /New Zealand	Department of Statistics	17.3.1	Wellington
1972-1979	External Trade Imports: Commodity by Country and Country by Commodity	Department of Statistics	17.3.3	Wellington

3.5 Results

Table 3.2: **Estimated numbers of domestic livestock in New Zealand** 1851 – 1871 **based on census data.** Data included only stock owned by European settlers.

Year	Sheep	Cattle	Horses	Goats	Pigs	Poultry	Asses and Mules
1851	233,043	34,787	2,890				
1858	1,523,324	137,204	14,912	11,797	40,732		122
1861	2,760,183	193,150	28,270	12,191	43,270	236,098	153
1864	4,937,273	249,760	49,409	12,005	61,276	378,414	339
1867	8,418,579	312,835	65,715	11,964	115,104	676,065	323
1871	9,700,629	436,592	81,028	12,434	151,460	872,174	397

Livestock populations in New Zealand 18511871

Estimated numbers of domestic livestock during early European settlement are shown in Table 3.2. During the 20-year period from 1851 to 1871 there was a rapid increase in both the national cattle (12.5 fold) and sheep (41.6 fold) populations.

Table 3.3: **Estimated number of live sheep and cattle imported into New Zealand and mean populations by decade between** 1860 – 1979**.**

	Cattle		Sheep	
Decade	**Total No. imported**	**Estimated mean population**	**Total No. imported**	**Estimated mean population**
1860	53,384	251,960	604,525	5,372,412
1870	2,030	503,313	13,580	11,491,607
1880	1,176	775,998	6,062	14,774,840
1890	837	1,034,239	4,218	19,247,813
1900	235	1,615,854	4,389	20,556,601
1910	691	2,543,626	3,349	25,031,689
1920	301	3,351,899	5,277	25,122,965
1930	320	4,304,019	4,032	29,965,629
1940	266	4,601,716	2,976	32,571,329
1950	512	5,547,273	4,126	39,288,805
1960	429	7,104,100		
1970	2,516	8,712,775		

Importation of cattle, sheep and poultry 18601979

The estimated numbers of cattle and sheep imported during each decade between 1860 and 1979 are shown in Table 3.3. From 1860 to 1959 an estimated 652,534 sheep were imported into New Zealand; this was larger by a magnitude of 10 than the 62,697 cattle imported from 1860 to 1979. For the decade $1860-1869$, 53,384 cattle, or 85% of all cattle importations for $1860-1979$, occurred, and 604,525, or 93% of all sheep importations for $1860-1959$, occurred (Table 3.3). The proportions of cattle and sheep imported from different countries for each decade between 1868 and 1979 are shown in Table 3.3. During the period 1868 to 1979 a total of 15,157 cattle were imported, with 10,080 (66.5%) from Australia, 3,909 (25.8%) from the United Kingdom and 1,168 (7.7%) from other countries. During the period 1868 to 1959, a total of 60,918 sheep were imported, with 51,347 (84.3%) reported as coming from Australia, 6,712 (11.0%) from the United Kingdom and 2,859 (4.7%) from other countries. Chile was reported as a source of seven cattle and eight sheep in 1891.

In $1960-1979$, when the reporting of sheep and goat numbers was combined, the total number of sheep and goats imported was 8,263, with 8,150 (98.6%) from Australia, 113 (1.4%) from the United Kingdom and none from other countries. Numbers of poultry imported between 1880 and 1979 for each decade, with country of origin, are shown in Table 3.4. The total number of poultry importations recorded between 1880 and 1939 was 20,701, compared to 287,789 in the 1960s and 127,651 in the 1970s, with most importations being from Australia.

Table 3.4: **Estimated total number of live poultry imported into New Zealand, by country of origin, and mean population, by decade between** $1880-1979$. UK is the United Kingdom. Other refers to all countries not including the UK or Australia.

Decade	Total imported	UK	Australia	Other	Estimated mean population
1880	2,866	1,125	1,244	497	1,622,568
1890	5,448	650	1,105	3,693	1,790,070
1900	7,094	1,783	4,751	560	3,191,604
1910	2,252	462	1,548	242	3,580,569
1920	2,518	521	1,789	208	3,886,077
1930	523	156	339	28	4,019,076
1940a	18	16	2	0	4,470,366
1950	54	36	18	0	4,342,581
1960	287,789	2	287,787	0	4,910,947
1970	1,274,741	0	1,274,741	0	7,045,089

a - poultry importations for the 1940s do not include $1942-1948$

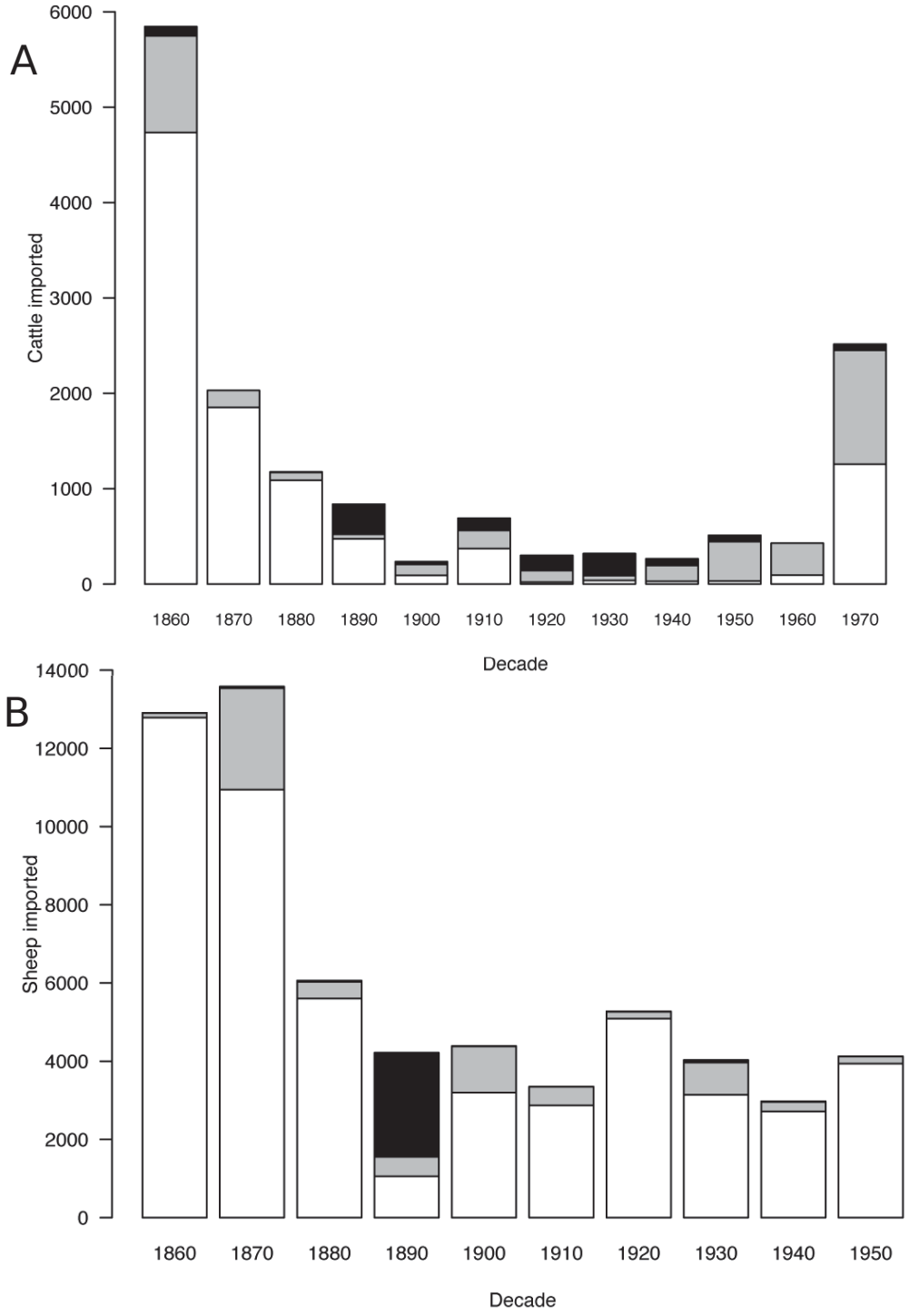


Figure 3.1: **A stacked bar chart of the reported number of (a) cattle and (b) sheep imported into New Zealand for each decade between 1860-1979.** Each bar shows the proportion of each species imported from Australia (white), United Kingdom (grey) or other countries (black). Note, the data for the 1860 decade only include 1868 and 1869.

3.6 Discussion

These results provide strong evidence that historically Australia was the major source of cattle, sheep and poultry into New Zealand, followed by the United Kingdom. A review of the number of live animal importations does not support Chile as being an important source of cattle, sheep or poultry during this period. It could be argued that from the start of the official records in the 1840s until 1867, when the source country was not recorded in the annual official records, a large number of animals could have come from Chile and the United Kingdom. However, this was not supported by reports that merino from Australia was the predominant sheep breed and cattle were a mixture of breeds, mainly from Australia [428] [429].

We propose that the historical events and numbers of sheep and cattle imported reflect a relative genetic bottleneck produced by the New Zealand population being based on a limited subset of animals derived from Australian livestock. The early cattle and sheep breeds introduced to Australia were themselves a selection of certain breeds or from certain regions, suggesting that they were a non-random sample of a non-homogenous world population [454]. In the 1860s 53,384 cattle and 604,525 sheep were imported to New Zealand, equivalent to 1.4% and 2.9% of the Australian cattle (3,846,679) and sheep (20,981,523) populations in 1861 [455]. Although these were relatively large numbers of animals being imported into New Zealand they were neither likely to be randomly selected from the population nor would the population be homogeneously distributed, i.e. the animals on a given farm were more likely to be related to each other than to animals from another farm.

There was a rapid increase in New Zealand livestock numbers during early European settlement $(1851-1871)$. Some of the numbers and origin of importations were not comprehensively reported in the annual official records during the period $1840-1867$, but the effect of these early importations and natural increase can be seen in the expansion in population size shown in Table 3.4 , as pastoral agriculture became established in New Zealand. In terms of population structure this suggests a pattern of a bottleneck event, as initially a comparatively limited number of cattle and sheep arrived from overseas to start the populations, followed by rapid population expansion with ongoing but limited immigration (admixture). Investigation into New Zealands cattle and sheep genetic population structures should reflect these early events. This pattern may be less obvious in the modern dairy cow, unless mitochondrial genes were examined, due to the widespread use of artificial insemination. It is possible that the effect of a population bottleneck resulting initially in reduced genetic diversity may be found in the host related microorganisms carried by the cattle and sheep. There are many other contributing factors, but New Zealand has a limited range of livestock-associated pathogens, for example Q fever and scrapie are absent from the national flock [456] [457]. Further, for other zoonotic pathogens, such as Shiga toxin-producing *Escherichia coli*

and *Campylobacter jejuni*, there is evidence of lower diversity and unique patterns of sequence types in New Zealand compared to other countries [15] [458].

Establishing Australia as the major historical source of sheep to New Zealand helps begin to answer an important biosecurity and international trade question; why is scrapie absent in New Zealand sheep? Our findings support the previous suggestion that the absence of scrapie in New Zealand stems from Australia being the main source of sheep during the early development phase of farming [456]. Other contributing factors may have been the 6-month journey by ship from the United Kingdom acting as a quarantine and eliminating sick animals [456]. Also at the time when most of the importations from the United Kingdom occurred scrapie may have been concentrated in certain breeds which were more confined to specific areas in the United Kingdom, for example Dorset Horn and Hampshire breeds in the county of Wessex [459], and the Suffolk breed that were not imported to New Zealand [456] [460]. The absence of scrapie cannot be attributed to genetic resistance as the susceptible PrP alleles are present in important breeds in both New Zealand and Australia [461] [462] and it has been experimentally induced in New Zealand sheep [463]. Thus the evidence supports the view that at the initial stages of development of the sheep population in New Zealand it depended heavily on sheep from Australia, and due to a combination of factors these sheep were scrapie-free. Both countries have maintained their scrapie freedom in recent years using strict quarantine measures.

An advantage of using the annual official records of the New Zealand government was that they provide combined information on all individual importations from ports all across the country for each year. Furthermore, the fact that they were collected and compiled by a government department suggests that a reasonable degree of rigour was applied to ensure they were correct. However, there were certain drawbacks associated with the use of these records. Because they were produced over such a long period, there were changes in the format and the type of information reported as the emphasis of the reports changed. For example, the term poultry as defined in 1980 by Statistics New Zealand refers to chicken, ducks, geese, turkey, and guinea fowls, and this definition is assumed to be consistent with the earlier records, but over the period under investigation the definition used may have changed (Anonymous 1980). An issue not addressed in this study was the effect of importation of genetic material such as semen, unfertilised eggs, and fertilised eggs. During the nineteenth century, when the bulk of the cattle and sheep importation occurred, the technology for freezing semen and embryos for long-term storage and transport did not exist [464]. However in more recent times importation of genetic material, particularly fertile hatching eggs by the modern poultry industry, became important [465].

The annual official records used in this study do not specify poultry importation numbers prior to 1868, but poultry were known to have been introduced to New Zealand from the time of the European explorer James Cook [447]. The patterns of introduction for cattle and sheep were similar

to each other, with very large numbers imported in the 1860s followed by a marked decline in the numbers imported. The pattern of introduction for poultry differed, with large numbers of importations not starting until the 1960s, as the poultry industry changed from predominantly small-scale poultry production to a modern high-production industry. This was associated with increasing the demand for birds with high genetic production value [466].

The collated data and historical reports show that from the time of European settlement to the 1970s Australia has been the main source of livestock introduced into New Zealand. The pattern of importation showed that large numbers of cattle and sheep were initially imported in the 1860s, probably in response to rapid agricultural expansion, thereafter importations continued with much reduced numbers. In contrast, relatively small numbers of poultry were introduced until the 1960s when large numbers were imported as part of the development of a modern high-production industry. The overall pattern for both cattle and sheep was of a founder event, as initially a relatively limited number of animals arrived from select outside populations, followed by population expansion with ongoing but limited immigration (admixture). Investigation into the genetic population structure of New Zealands cattle and sheep, as well as their host-associated microorganisms, could reflect the impact of these early historical events.

Acknowledgements

The authors wish to thank Emeritus Prof. AN Bruere (ONZM) for initial help and advice. This work was supported by the Marsden Fund (grant number MAU0802) and the Australian Research Council (grant number FT100100031).

Now, my own suspicion is that the universe is not only queerer than we suppose, but queerer than we can suppose..... and must be my excuse for dreaming.

J.B.S. Haldane, Possible Worlds

4

the phenotype:

Phenotypic comparison of isolates associated with wild-birds and livestock.

Prelude

This chapter investigates *C. jejuni* for phenotype, genotype, and host adaptation from a New Zealand perspective. *C. jejuni* is known for its phenotypic and genotypic variation [132], but it has not been established how much phenotypic variation there is between the more investigated sequence types (STs) found in livestock and the less well known STs found in wild-birds. MLST typing has shown a pattern of genetic differences in the seven housekeeping genes that enable sequence types (STs) and clonal complexes(CC) to be attributed to specific hosts or host group [315] [467]. The New Zealand isolates used in this Biolog phenotypic microarray analysis were from sequence types associated with different types of host and lifestyle. The question, "what is the link between phenotype and host?" was refined to compare the ability of different host associated isolates to

utilise carbon sources at two different temperatures. The temperatures represent *C. jejuni* at two different physiological states, 42℃ when actively growing, and 22℃ when not growing. In this chapter the null hypothesis was tested that there is no difference between the number of carbon sources utilised by STs associated with livestock or wild-birds in either physiological state.

The previous chapter on demographic history, established the timing, scale and source of recorded arrivals for cattle, sheep and poultry into New Zealand between 1860–1979. This quantification of the arrival of livestock to New Zealand, an isolated group of islands without these animals prior to European arrival, should reflect the arrival into New Zealand of the livestock associated *C. jejuni* STs e.g. ST-42.

ST-42 has been commonly found internationally and associated with agricultural animals, particularly cattle, so probably arrived in New Zealand with its host as described in the demographic history. Three of the other sequence types have been found in faeces from two New Zealand water rails, the indigenous pūkeko (*Porphyrio porphyrio melanotus*) and the endemic takahē (*Porphyrio hochstetteri*): ST-2381, ST-3663 and ST-3655. Genomic comparisons (Chapter 5) show ST-2381, ST-3663, ST-3655, ST-3845 and ST-3845 are from a shared lineage. The ST-2381 lineage is commonly found in New Zealand environmental waters, but rarely reported outside the country [8]. ST-3845 has been found in New Zealand water but has not yet been found in a host species faeces. ST-3673, is rarely reported, thus far it has been found once in chicken tissue (USA) and once in water (NZ). Because some these isolates are found in two non-migratory native birds and the environmental waters, but rarely if ever reported internationally, identifying their phenotype may identify unique features present in New Zealand *C. jejuni*.

Several factors attributed to the temperature choices in this chapter. In laboratory conditions *C. jejuni* has an optimum growth temperature of $\sim$ 42℃ [373] [89] so it was used as the temperature to test the phenotype of the isolates when actively growing. As shown in the literature review (section 2.6.1), 42℃ has been frequently associated with poultry (*Gallus gallus*) which has a reported normal rectal body temperature of 40.6℃ – 43.0℃ [374] [373], however it has been shown that generally avian body temperature shows a wide range of variation dependent on physiology [377]. The temperature 37℃, which is often associated with average human body temperature, and due to this the behaviour of *C. jejuni* at this temperature has been investigated [373]. However this chapter, is focused on comparing the phenotypic behaviour of *C. jejuni* isolates in their actively growing state (42℃) and in a less active, non growing state (22℃) as experienced in the environment, rather than comparing behaviour at actively growing temperatures with 42℃ vs. 37℃. The lower temperature, when in the non-replicating state was set at 22℃ as this is towards the lower end of operation for the Omnilog plate reader. Options to use lower temperatures like 14℃ the average temperature of New Zealand surface water in summer, or 4℃ a common refrigeration temperature were considered but discarded due to the technical difficulty and issues with comparisons. Biolog Inc. had advised

us a software modification to increase the temperature range achievable in the Biolog Plate reader was to be made available, but no modifications were received during the course of this experiment, nor have we been made aware of any of since its completion.

The basis of the chapter was written as an article to be published in a scientific peer-reviewed journal, so it is written as a stand alone article, rather than a thesis chapter. As a consequence some of the work supporting this investigation is not included in the printed chapter and has been put into Appendix B on the attached DVD. The work in Appendix B supports this investigation of the relationship between phenotype, genotype and host and is part of the thesis chapter. This material may form the basis of future analyses. The work on the DVD can be read with an internet browser as a website.

Different carbon utilisation patterns in *Campylobacter jejuni* associated with host generalist or specialist lifestyle and temperature.

Manuscript prepared for Applied and Environmental Microbiology.

4.0.1 Abstract

Campylobacter jejuni is a principal cause of bacterial gastroenteritis in people around the world. A study was performed using Biolog phenotypic microarrays to identify and compare the phenotypic profiles of 11 phylogenetically related New Zealand *Campylobacter jejuni* isolates at two different temperatures, 42℃ and 22℃. The sequence types used included ST-42 which is found internationally and associated with agricultural animals and has a host generalist lifestyle. In contrast ST-2381, ST-3663, and ST-3655 have been reported in one or both of two closely related indigenous New Zealand birds and are associated with a host specialist lifestyle. ST-3673 has been found in a chicken and ST-3845 has only been found in water. At 42℃, ST-42 used a greater number carbon sources than ST-2381 and ST-3655, which is consistent with a generalist that lives in a wider range of hosts and utilises a wider range of resources, however this pattern did not extend to the other sequence types. The pattern of single carbon source respiration clustered according to sequence type at 42℃ but not at 22℃. In total there was activity in 29 out of 190 phenotypic microarray wells at 42℃ after 48 hours but at 22℃ a subset group of only seven wells showed activity after 96 hours. These results show that the phenotypic profile variation at 42℃ reflects MLST. The carbon sources showing activity at 22℃ tended to be closely related to the tricarboxylic acid cycle and energy production.

4.1 Introduction

Campylobacter jejuni is a major public health problem worldwide, it is the commonest cause of gastroenteritis reported in developed regions like the EU and the USA, with New Zealand reporting a peak of 383.5 cases per 100,000 in 2006 [468] [50] [2]. It seems counter-intuitive that a fastidious bacteria which is difficult to grow in laboratory conditions, preferring relatively high temperatures and with a sensitivity to atmospheric levels of oxygen should emerge as such a major problem in developed countries. Yet despite being a prominent cause of disease in people, many aspects of the biology of *C. jejuni* remain a mystery, such as how a bacteria with a small genome can have the metabolic flexibility to colonise the gastrointestinal tract of multiple hosts species and survive exposure to a variety of external environments encountered during transmission from host to host [469].

C. jejuni is known for both diversity in genotype and phenotype [4][5] and a better understanding of this diversity may lead to better control methods. Analysis of *C. jejuni* population structure and genotypes has lead to the identification of two different lifestyles, a group of generalists that can colonise a wide range of hosts, and a group of specialists each of which tend to colonise a specific species or a few closely related species [6] [7] [5]. The ability to colonise a broad host range may increase the risk of a pathogen becoming an emerging disease in humans and most cases are associated with the host generalist sequence types of *C. jejuni* [470] [471]. A host generalist lifestyle may also indicate the ability to utilise more resources i.e. utilise a wider range of compounds than an isolate with a host specialist lifestyle. Isolates from livestock such as cattle, sheep and poultry are more associated with a host generalist lifestyle, and filling an agricultural food animal niche [7].

Investigations into isolates from wild-birds show they exhibit a host specific lifestyle and although they cannot be ruled out as a cause of human infection they are considered an unlikely source [7] [321] [318]. There is concern that *C. jejuni* isolates associated with wild-birds switch hosts to livestock and/or act as a source of genetic material by horizontal gene transfer [17][p266]. Most investigations involve *C. jejuni* isolates associated with the agricultural niche and human cases, with little known about how much genetic variation and phenotypic diversity exist in the isolates from wild-birds [17][p266].

A high throughput phenotypic microarray system has been developed for global phenotypic characterisation using a series of proprietary plates from Biolog (Hayward, CA, USA) [184]. A range of New Zealand isolates from sequence types associated with agriculture and wild-birds were examined using the Biolog phenotypic microarrays. The phenotypic microarray plates, PM1 and PM2A, each have 96 wells containing a unique carbon source and a single negative control well. It is presumed if the bacteria are able to metabolise the supplied carbon source this will lead to cell respiration

resulting in the reduction of the tetrazolium dye [185]. When all the cell's processes are working normally there is a flow of electrons from the carbon source to NADH, then along the electron transport chain of the cell, until the electrons are accepted by the tetrazolium dye which changes to produce a purple colour [186].

Previous comparisons have been made using Biolog phenotypic microarrays supplying a sole carbon source to *C. jejuni* at 30℃, 37℃ and 42℃ [187] [206] [5] [207]. Tang et al. (2010) found there was no change in the pattern of carbon utilisation between 30℃ and 42℃ using *C. jejuni* ATCC 33560 [206]. Line et al. (2010) reported *C. jejuni* 11168 (GS) in general showed greater levels of oxidation for a given sole carbon source at 42℃ than 37℃ [187]. An examination by Gripp et al. (2011) of eight different ST-21 isolates at 37℃ and 42℃, showed two distinct patterns of carbon metabolism which were neither temperature dependent nor indicative of the isolates source [5]. The same researchers reported that nine other *Campylobacter* sequence types were also tested using the phenotypic microarrays and each produced a distinct metabolic fingerprint [5]. Wagley et al. (2014) compared *C. jejuni* and *C. coli* isolates using the Biolog PM1 and PM2A plates on mainly human case isolates from UK, Pakistan and Thailand, although one was from a goose but it was not identified as domestic or wild [207]. All of the isolates used in these four separate analyses were from chicken, food, bovine, or human cases aside from a single goose, so there is a paucity of information on wild-birds associated isolates and ST.

The biochemical pathways utilised by the family Campylobacteraceae show significant differences to the more studied gastrointestinal bacteria family the Enterobacteriaceae [472] [473]. *C. jejuni* has been shown to utilise a limited range of carbon sources, and unlike *Escherichia coli* (*E coli*) is unable to metabolise any sugars unless a genomic island (Cj0480c–Cj0490) is present that enables the utilisation of L-fucose [474] [210]. Genes coding for aspects of the central carbon metabolism, the Embden–Meyerhof pathway, pentose phosphate pathway and the complete tricarboxylic acid cycle (TCA cycle) have been identified in the sequenced *C. jejuni* genome [389] [44][41–71]. A predominant site for *C. jejuni* in poultry is the lower gastrointestinal tract [475] where anaerobic fermenting bacteria dominate, and generate as by-products small organic acids which along with certain amino acids are utilised both as carbon and energy sources by *C. jejuni* [476] [477] [188].

Understanding *C. jejuni* phenotypic behaviour at different temperatures is important because *C. jejuni*, like all gut bacteria, regularly experience environmental temperatures when in transition from host to host. Potentially contaminated food processing and handling in most household kitchens occurs at room temperature (about $18 - 25$℃) or it is stored in refrigerators (about 4℃), so understanding *C. jejuni* behaviour at these lower temperatures is important. Most microorganisms have a gradual decline in growth rate as they near their minimal growth temperature whereas *C. jejuni* tends to grow in a narrow range of temperatures ($30 - 47$℃) [373] [89]) then

transition abruptly from the growing to a non-growing state [382] [383]. *C. jejuni*, a thermophilic *Campylobacter* spp., is unable to replicate below $\sim$ 30℃ and in the laboratory has an optimum growth temperature of $\sim$ 42℃ [373] [89] which approximates host body temperature. Unlike *E.coli*, *C. jejuni* does not produce the protective cold shock proteins (*Csp*) in response to a marked decrease in temperature, but other protective proteins and some protective mechanisms like polynucleotide phosphorylase (PNPase) have been reported as important to its long-term survival at low temperatures [406] [383]. Despite this, a meta-analysis of *C. jejuni* survival reports for 0℃ to 40℃ produced a mixed effects model with better survival at 0℃ than at 40℃, showing that lower temperatures are more than tolerated [10]. *C. jejuni* is not inert at low temperatures, having been reported as motile and chemotaxic towards formate and malate under microaerobic conditions at 4℃ and 20℃ as well as 40℃ [383]. At 20℃ declining levels of respiration and increased coccoid shape formation have been reported [478]. Indications of cell metabolic activity in the form of oxygen consumption, catalase activity, ATP generation, and protein synthesis have been observed as low as 4℃ [383]. While we are beginning to understand its ability to survive at lower temperatures, there is little information on phenotypic variation of *C. jejuni* at room temperature ($18 - 25$℃).

The genomes of the isolates were sequenced, with Illumina sequencing, to enable comparison between them based on rMLST, and identify if the genomic island associated with fucose utilisation in *C. jejuni.* is present [474] [210].

This study describes the phenotypic profile produced using Biolog phenotypic microarrays on a group of related *C. jejuni* isolates, associated with two different hosts types, at two different temperatures 42℃ and 22℃. The significance of the study is in providing insights into the single source carbon utilisation of a host generalist (agriculture-associated) sequence type in comparison to several host specialist sequence types (wild-bird associated) and a sequence type isolated from water and a chicken. The single source carbon utilisation pattern was compared between the two different host associated groups at 42℃ for 48 hours a temperature associated with being inside a host, and at 22℃ for 96 hours a temperature associated with being in the environment.

4.2 Methods and Materials

Isolates

The 11 isolates used in the phenotype analysis were all from New Zealand and stored in the mEpiLab -80℃ freezers. The isolates were collected as part of other projects and details are on described in [479] and [480]. The isolates were typed using the multilocus sequence typing (MLST) method as described by Dingle et al.(2001), which was based on an index of the alleles present at seven specified housekeeping genes [265] [277]. The isolates were initially chosen based on their close relationship to a common New Zealand water isolate, ST-2381 [8], based on minimum spanning tree

of the MLST allelic profiles from 5092 sequence types downloaded from the PubMLST database[1] see Figure 4.4. Eleven isolates representing six sequence types were used based on their relatedness in shared MLST alleles and more than one isolate represented the same ST. However after this analysis was initiated, work by mEpiLab established the ST-2381 and several related sequence types were associated with the pūkeko (*Porphyrio porphyrio* subsp. *melanotus*) and the takahē (*Porphyrio hochstetteri*).

Table 4.1 shows each isolate, MLST sequence type (ST) and the host associated with the ST. The three ST-42 isolates belong to a sequence type that is considered to be host generalists, and while mainly associated with ruminants particularly cattle, it has been isolated from surface waters, poultry, wild-birds and human cases [481] [482] [8]. ST-42 forms the central genotype of a major clonal complex (CC42) and is found internationally [272] [482].

ST-2381 and ST-3655 have been isolated from New Zealand surface waters and two New Zealand rails, the pūkeko (*Porphyrio porphyrio* subsp. *melanotus*) and the takahē (*Porphyrio hochstetteri*) [8] [16]. ST-3663, ST-3673 and ST-3845 share MLST alleles with both ST-42 and ST-2381 as shown Table 4.1, while the minimum spanning tree used to find this relationship is in Figure 4.4. Our laboratory has isolated ST-3663 in takahē, but it may be able to colonise the pūkeko. We are aware of ST-3673 being reported twice in PubMLST (http://pubmlst.org/campylobacter/), once in New Zealand water and once from chicken meat or offal in the USA although it was classified as *C. coli* in the profile [277]. ST-3845 has been found several times in New Zealand water but has not yet been reported from an identified host.

Whole genome sequencing

The genomes of all the isolates were sequenced by two very similar methods. Seven isolates had genomic DNA extracted and the whole genome sequenced as described in Biggs et al.(2011) [273]: S263a, W83a, N3d, W135a, W63b, W120a, N53. The other four *C. jejuni* isolates (N27, N31, N191, W194b) had genomic DNA extracted using either the Promega wizard genomic DNA purification kit or the Qiagen QIAmp DNA mini kit. Both kits were used as per the manufacturers' instructions. The genomic DNA samples was fragmented by nebulisation for 6 minutes at a pressure of 32 psi, purified, then end repaired, A-tailed, adaptor-ligated, fractionated, purified and enriched according to the manufacturer's instructions, using the TruSeq DNA LT Sample Prep Kit v2-Set A. The prepared libraries were normalized to equal molarity, diluted to 2nM and pooled; 20 libraries per pool. A flow cell was prepared for each of the library pools and sequencing reactions using $9\rho M$ of the pooled libraries were performed on an Illumina MiSeq instrument with the MiSeq Reagent Kit v2, this resulted in approximately 12 to 15 million clusters per run.

An in-house work flow based on Perl scripts was used to assemble, quality control and annotate the genomes. Solexa QA [483] was used for trimming and quality control of the reads (Q30,

[1] 5^{th} October 2011 *Campylobacter* PubMLST http://pubmlst.org/campylobacter/

Table 4.1: **The *C. jejuni* isolates, their sources, host or host group, and MLST allelic profile**. The sequence type (ST) based on the seven loci of the multilocus sequence type (MLST) profile (*asp*A, *gln*A, *glt*A, *gly*A, *pgm*, *tkt*, *unc*A) are reported for each isolate. The Sample refers to the material in which the isolate was found, and this can vary from the host, or host group that the ST has previously been associated. The Host refers to either to a host species that a ST has been associated with or the host group "agriculture" which refers to sequence types associated with livestock such as cattle, sheep and chickens. Isolate W120a (ST-3673) has been reported once in NZ water and once in chicken meat in the USA.

Isolate	ST	*asp*A	*gln*A	*glt*A	*gly*A	*pgm*	*tkt*	*unc*A	Sample	Host	Region	Year
W83a	2381	175	251	216	282	359	293	102	water	pūkeko, takahē	Manawatu, NZ	2006
W194b	2381	175	251	216	282	359	293	102	water	pūkeko, takahē	Manawatu, NZ	2007
N3d	2381	175	251	216	282	359	293	102	water	pūkeko, takahē	Waikato, NZ	2008
W135a	3655	1	6	5	282	261	7	3	water	pūkeko, takahē	Manawatu, NZ	2006
W63b	3663	175	6	216	282	261	7	3	water	takahē	Manawatu, NZ	2006
W120a	3673	175	6	216	4	434	7	3	water	chicken	Manawatu, NZ	2006
N53	3845	1	6	216	282	261	346	3	water	unknown	Waikato, NZ	2009
N27	3845	1	6	216	282	261	346	3	water	unknown	Waikato, NZ	2008
S263a	42	1	2	3	4	5	9	3	cattle faeces	agriculture	Manawatu, NZ	2008
N191	42	1	2	3	4	5	9	3	cattle faeces	agriculture	Waikato, NZ	2010
N31	42	1	2	3	4	5	9	3	water	agriculture	Waikato, NZ	2008

expected probability of incorrect base calling 1:1000). Velvet (ver. 1.2.10) [484] was used for *de novo* assembly of the 2×250 base short reads. A range of kmers ($55 - 250$bp in 10bp steps) were used in the assembly process and the characteristics of the assemblies were compared and ranked using N50, the number of contigs, maximum contig length and assembly length. N50 is a measure of length of the longest contig in an assembly, where at least 50% of all base pairs are contained in contigs of this length or larger. The highest ranking contig assembly for a given genome were then annotated with Prokka (ver. 1.10) [485].

A dataset of 32 *C. jejuni* and 14 *C. coli* genomes (Supplement Table 4.4) was used as a comparison basis for the ribosomal multilocus sequence types (rMLST) with the 11 isolates used in the Biolog phenotypic microarray analysis . The 46 isolates had genomic DNA extracted and their whole genome sequenced as described in Biggs et al [273]. All the genomes used in this analysis are draft genomes.

Genome analysis

All the genomes were uploaded to a local copy of the BIGSdb servers [277]. The Genome Comparator module of BIGSdb used *C. jejuni* NCTC 11168 as an annotated reference genome, and was the source of comparison sequences to BLAST against and generate a whole genome profile for each isolate. The default settings were used; a minimum of 70% identity, a minimum 50% alignment and a BLASTN word size of 15.

The concatenated rMLST sequences of the 11 isolates used in the Biolog phenotypic microarray analysis were aligned in Geneious (ver 7.1.7) using Clustal W [486] and represented as a Neighbor-Net in SplitsTree [487]. The rMLST alleles of all 57 genomes (listed in Table 4.1 and Supplement Table 4.4) were also made into a distance matrix that simply counted of the number of loci that differ between each pair of isolates, missing or incomplete alleles were removed from the pairwise comparison. A Nexus file was created from the distance matrix and made into an unrooted phylogenetic network using the Neighbor-Net in SplitsTree [488]. All of the *Campylobacter* spp. had 52 rMLST alleles and lacked the *rpm*D locus.

A genomic island at (Cj0480c–Cj0490) enables isolates to utilise L-fucose [474] [210]. Using the genome comparator in BIGSdb the 11 isolates used in the Biolog phenotypic microarray analysis were checked for an insertion at: Cj0480c, Cj0484, Cj0485, Cj0486, Cj0487, Cj0488.

Biolog Phenotype Microarray

The isolates were revived from the -80℃ freezer by plating onto Columbia horse blood agar and grown for 48 hours at 42℃ in the MACS VA500 variable atmosphere workstation under microaerobic conditions (85% N_2, 10% CO_2, 5% O_2). A single colony was then inoculated onto Columbia horse blood agar and grown for about 24 hours at 42℃ in the same microaerobic environment. Outside the workstation, in a laminar flow cabinet, the cultures were then harvested from the surface of the

blood agar plates into a mixture of IF0a inoculating fluid (Biolog, Hayward, CA, USA), $NaHCO_3$ (150mM), Bovine serum albumin (BSA 6% solution), sterile water, and tetrazolium violet (Dye D, Biolog, Hayward, CA, USA) as per the manufacturer's instructions (Biolog Inc.). A sample of the inoculum was swabbed onto a blood agar plate and incubated for $24-48$ hours in the MACS VA500 variable atmosphere workstation to check the isolate being tested was viable and a pure culture. The final optical density of the mixture was measured at 52% transmission $\pm$ 2%(Turbidimeter, Biolog). 100 µl aliquots were inoculated into each of the 96 wells in the phenotypic microarray plates, PM1 and PM2A (Biolog). This process took less than 15 minutes then the plates were returned to the MACS VA500 variable atmosphere workstation and sealed in a plastic barrier bag (supplied by Biolog) to maintain the microaerobic environment for rest of the experiment. The bag was double heat sealed with an impulse heat sealer (Impulse sealer TISH 300C) and immediately placed in an Omnilog automatic plate reader (Biolog) for incubation. The bags containing the inoculated phenotypic microarray plates were taped to the tray in the plate reader. Each plate was incubated at 42℃ for 48 hours or at 22℃ for 96 hours. The Biolog phenotypic microarray measures cell respiration by tetrazolium dye reduction [184]. Any visible colour change, showing the reduction of the tetrazolium, was recorded spectrophotometrically in Omnilog units, every 15 minutes using a CCD camera. A cell free negative control phenotypic microarray plate, for both PM1 and PM2A, were tested for abiotic dye reduction to show any false positive reactions at both temperatures. Each plate had a negative control well (PM1 A1, PM2A A1) that lacks a carbon source, in which despite being inoculated with the isolate, no significant colour change of the tetrazolium dye was recorded. The entire process was performed at least twice for each isolate on each plate (PM1, PM2A) at both temperatures. A table of the number of technical and biological replicates is supplied in the Appendix B.

Curve parameter analysis

The value in Omnilog units was recorded over time for each well and when there was a significant level of activity it was plotted to give a respiration curve [199] [196]. The grofit package within R, using the splines option, was use to convert the curve into a set of four parameters (lag phase, slope of growth phase, maximum height, area under the curve) based on an idealised growth curve [197]. This analysis only used the maximum value attained (A) in a well i.e. the maximum height of the idealised growth curve.

Adjustment of data

After the annual service of our Omnilog plate reader there was a noticeable decrease in all the Omnilog values recorded from start to finish, consistent with an adjustment of the baseline value. In order to account for this difference a linear mixed effects model was developed comparing the maximum height value (A parameter or curve maxima) before and after the service values for the same isolate under the same conditions. This resulted in reducing the maximum value for

the pre-service analysis by 35.5 Omnilog units, so the results before and after the service were comparable. Supporting information for this process is in Appendix B.

Identifying carbon source utilisation

A series of steps were taken to identify if the colour change recorded in Omnilog units for each well of the phenotypic microarray plate was due to carbon utilisation by an isolate. First, all the plots of the change in tetrazolium dye colour over time in a given well were manually examined. Previously identified carbon sources (wells) that produced false positive reactions, in both PM1 and PM2A plates [187], were confirmed using an un-inoculated plate which confirmed a significant level of colour change due to abiotic reactions. The false positives (dihydroxyacetone, D-arabinose, L-arabinose, L-lyxose, D-ribose, D-xylose, 2-deoxy-D-ribose, D-glucosamine, oxalomalate, 5-keto-D-gluconate) were removed from the analysis. Density plots of the maximum utilisation value for each well were reviewed (shown in Appendix B), wells where no isolates achieved a value > 100 were considered not to show signs of utilisation i.e. pronounced tetrazolium dye reduction in the presence of the carbon source and the bacteria in the Biolog phenotypic array plate under the conditions being tested. This threshold value was set at a high level to reduce the chances of accepting a false positive, with the results shown in Table 4.2 and Table 4.3. Carbon sources that did not meet this criteria, at either temperature, for carbon utilisation were not included in later analysis.

kmeans

The kmeans were used to cluster the maximum value for all results in a given well into groups of high activity, weak or no activity. The Hartigan and Wong algorithm was used by the kmeans function in R and follows a process of looking for a local optima of the sum of squares for a number of clusters (k), such that moving one point from one cluster to another cluster does not reduce the within sum of squares of the clusters [489]. The kmeans clustered the results for a given well into 2 or 3 groups. The kmeans cluster of higher values were attributed to significant metabolic activity by the isolates in the well (a positive), and the lower scoring kmeans cluster are attributed to no significant activity by the isolates reducing the tetrazolium dye in the well (a negative). In many wells there was a clear distinction between the two clusters ($k = 2$), but in some wells where the distinction was less clear a third cluster ($k = 3$) was used, with the middle cluster representing results that may be weak positives.

Some isolates showed an inconsistent result for a given carbon source (well) by having replicates in both the positive and negative kmeans groups. In these cases all the replicates were manually compared to see if the isolates results varied for a single carbon source or if the whole plate was affected. If an isolate was positive in a given well but not on a replication, and the rest of the results were consistent between plates, an observation with a kmeans scoring a negative result was not included in the calculation of the mean for the isolate with that carbon source. This approach was used in identifying the pattern of carbon source utilisation heatmap but was not needed in the

RE-EM analysis where this within isolate variation was treated as a random effect.

Heatmap

The data used to generate the heatmap was the calculated mean value in Omnilog units for each isolate at 42℃ in each active well. The two-way hierarchical clustering is shown in the two dendograms, both were performed using euclidean distances and complete linkage methods. Figure 4.3 was made using the R package pheatmap [490].

RE-EM tree

A regression tree with random effects was created using the REEMtree package in R which combines the generalisation of a mixed effects model and a tree based partitioning method to account for random effects and longitudinal data (i.e. repeated measures) [491]. The maximum height value of the respiration curve for each observation of an isolate at each temperature with 29 carbon sources was used to create the model; temperature, MLST and carbon source (well) were the fixed effects and isolate the random effect to account for variation between replications. The 29 carbon sources were all identified as showing respiration activity at 42℃. The RE-EM tree partitioned the observations until reaching a terminal leaf, and a terminal leaf was identified by a cross validation process to avoid under or over partitioning [491]. The RE-EM tree approach allowed for interactions between the explanatory variables - temperature, carbon source and isolate.

Statistical analysis

All the statistical analyses were performed using R v2.13 and v3.1.0 (R Development Core Team 2008, R Foundation for Statistical Computing, Vienna, Austria). The Biolog phenotypic microarray data was analysed using the packages opm[199][196], grofit [197], lme [492], kmeans [493], gplots [494], pheatmap [490] and REEMtree [495].

4.3 Results

Genome analysis

The 11 isolates given a phenotypic microarray analysis were compared to each other and to a larger dataset (Table 4.4), using rMLST in Figure 4.1. In Figure 4.1 A the rMLST comparison shows a distinction between *C. coli* and *C. jejuni* as identified by the red line with a 1. The *C. jejuni* then subdivide into their STs, with the 11 isolates used in the phenotypic microarray analysis identified by a yellow star. These isolates separate into a branch with ST-42, and another branch containing ST-2381, ST-3673, ST-3845, ST-3655 and ST-3663. Figure 4.1 B, shows the split between the group of 11 isolates more clearly. The isolate form into 3 main branches, with the top right branch consisting of the ST-42 isolates (S263a, N191, N31) and the lower right branch formed

by the ST-2381 isolates (W83a, W194b, N3d). The middle branch on the left is formed by ST-3655 (W135a), ST3845 (N53,N27), ST3673 (W120a) and ST3663 (W63b).

The genomic island insert (cj0480c-cj0490) containing fucP was found in the ST-3845, ST-3673, ST-3663 and ST-3655 isolates but not in the ST-2381 or ST-42 isolates.

Pattern of carbon source utilisation

In total 29 different carbon sources showed significant levels of respiration by at least one isolate on at least one occasion in Table 4.2. Although no two isolates showed exactly the same pattern, where there is more than one isolate with the same MLST they were more similar to each other, which is shown in the heatmap analysis (Figure 4.3). At 42℃ all the isolates except ST-42 utilise quinic acid. The ability to utilise L-fucose but not D-fucose is found in ST-3845, ST-3673 and ST-3663.

In Table 4.3 there are seven carbon sources that appeared to support respiration by a least one isolate at 22℃; succinic acid, L-lactic acid, formic acid, D,L-malic acid, D-malic acid, pyruvic acid, and succinamic acid. No isolate showed significant activity at 22℃ utilising L-malic acid, although some did at 42℃, so it may be more correct to consider the activity in the D,L-malic acid well to be just D-malic acid respiration. N191 (ST-42) just missed the cut off value for respiration at 22℃ with D-malic acid but did exceed the cut-off (> 100 Omnilog units) for the combined D,L-malic acid. Three isolates W120a (ST-3673), S263a (ST-42) and N27 (ST-3845) did not show respiration activity with formic acid at 22℃ but did at 42℃. Three isolates N191 (ST-42), N53 (ST-3845) and W135a (ST-3655) did not show respiration activity with formic acid at 42℃ but did at 22℃.

RE-EM tree

Figure 4.2 is a simplified representation of the full RE-EM tree shown in the supplementary file Figure S.1 is too big to see on a printed page. The significant first bifurcation in the RE-EM tree (Fig 4.2) is due to temperature. It must be noted that each split partitions the observations/results present in the node into two separate and smaller subgroups. The isolates at the lower temperature (22℃) were in a non-growing state and showed a different phenotypic profile to those in an actively growing state (42℃) as seen in Tables 4.2 and 4.3. At 42℃ after 48 hours, 29 carbon sources (wells) showed significant levels of activity but at 22℃ there were only seven carbon sources (wells) showing significant levels of activity after 96 hours.

At 22℃ the first carbon source node bifurcates into the two carbon sources most commonly used at this temperature by the isolates (L-lactic acid and formic acid), and the remaining 27 carbon sources. Following a node where L-lactic acid and formic acid split, there is a node from which ST-3655 (W135a) splits from the other isolates. The next split for the remaining isolates divides L-lactic acid from formic acid, with L-lactic acid then splitting the remaining isolates into a group with low maximum parameter values (ST-2381, ST-3673, ST-3845) and higher values (ST-42,

Table 4.2: **Utilisation of carbon sources by 11 *C. jejuni* isolates using two Biolog Phenotypic microarrays (PM1, PM2A) at 42℃.** A positive result is (+), a negative (-) and (+/-) shows some replicates were positive and some negative. The identification of the well number for each carbon source is supplied in the supplementary table in Appendix B

Carbon source	ST-42 N191	ST-42 S263a	ST-42 N31	ST-3845 N27	ST-3845 N53	ST-3673 W120a	ST-3663 W63b	ST-3655 W135a	ST-2381 N3d	ST-2381 W194b	ST-2381 W83a
Succinic acid	+	+/-	+/-	+	+	+	+	-	-	-	-
L-Aspartic acid	+	+/-	+/-	-	+	+	+	-	-	-	-
L-Proline	+	-	+/-	+	+	+	+	-	-	-	-
L-Fucose	-	-	-	+	+	+	+/-	-	-	-	-
L-Lactic acid	+	+	+	+	+	+	+	+	+	+/-	-
Formic acid	-	+/-	+/-	+/-	-	+	+	-	+	+/-	+/-
L-Glutamic acid	+	+/-	+	+	+	+	+	-	-	-	-
D,L-Malic acid	+	+	+	+	+	+	+	+/-	+	-	-
L-Asparagine	+	+/-	+/-	-	+	+	+	-	-	-	-
α-Ketoglutaric acid	+	+	+	+	+	+	+	-	-	-	-
L-Glutamine	+	-	+/-	-	+	+	+	-	-	-	-
α-Hydroxybutyric acid	+	+/-	+	-	-	-	+	-	+	-	+
Citric acid	+	-	-	-	+	-	-	-	-	-	-
Fumaric acid	+	+	+/-	-	+	+	+	-	-	-	-
Bromosuccinic acid	+	-	+/-	-	+	-	+/-	-	-	-	-
Glycolic acid	-	-	-	-	-	-	-	-	+	-	-
Glyoxylic Acid	-	-	-	-	-	-	-	-	+/-	-	-
Glycyl-L-Glutamic Acid	-	-	-	-	+	-	-	-	-	-	-
L-serine	+/-	+/-	+/-	-	+	+	+	-	-	-	-
Methylpyruvate	+	+	+/-	+/-	+	+	+/-	-	-	-	-
D-Malic acid	+	+/-	+/-	-	-	-	+	+/-	+/-	-	-
L-Malic acid	+	+	+/-	+	+	+	+	-	-	-	-
Glycyl-L-Proline	-	-	-	-	+/-	-	-	-	-	-	-
Pyruvic acid	+	+/-	+	+	+	+	+	+/-	+/-	+/-	-
D-Lactic acid Methyl Ester	+/-	+	+/-	-	-	-	+/-	-	-	-	+/-
Quinic acid	-	-	-	+	+	+	+	+	+	+	+/-
D-Ribono-1,4-Lactone	-	-	-	-	-	-	-	-	+/-	-	-
Succinamic acid	+	+/-	-	+	+	+	+/-	-	-	-	-
N-Acetyl-L-Glutamic acid	-	-	-	+/-	+	+	+/-	-	-	-	-

Table 4.3: **Utilisation of carbon sources of 11 *C. jejuni* isolates using Biolog Phenotypic microarrays (PM1, PM2A) at** 22℃. A positive result is (+), a negative (-) and (+/-) shows some replicates were positive and some negative. The identification of the well number each carbon source is supplied in the supplementary table in Appendix B

Carbon source	ST-42 N191	ST-42 S263a	ST-42 N31	ST-3845 N27	ST-3845 N53	ST-3673 W120a	ST-3663 W63b	ST-3655 W135a	ST-2381 N3d	ST-2381 W194b	ST-2381 W83a
Succinic acid	+/-	-	-	-	-	-	-	-	-	-	-
L-Lactic acid	+/-	-	-	-	-	-	+/-	+/-	-	-	-
Formic acid	+/-	-	+/-	-	+/-	-	+/-	+	+/-	+	+/-
D,L-Malic acid	+/-	-	-	-	-	-	+/-	-	-	-	-
D-Malic acid	-	-	-	-	-	-	+/-	+/-	-	-	-
Pyruvic acid	+/-	-	-	-	-	-	-	-	-	-	-
Succinamic acid	-	-	-	-	-	+/-	-	-	-	-	-

ST-3663).

The branching pattern of the RE-EM tree on the 42℃ side of the split was more complex. The first split at 42℃ forms two groups, one group (ST-2381, ST-3655) utilised less carbon sources than the other group (ST-3663, ST-3673, ST-3845, ST-42). The less active group (ST-2381, ST-3655) did not later split based on the ST pattern. The more active group (ST-3663, ST-3673, ST-3845, ST-42) showed multiple partitions within the group down several branches. One split showed ST-3673 and ST-3845 splitting from ST-3663 and ST-42 by differences in a combination of carbon sources (citric acid, glycylLglutamic acid, D-lactic acid methyl ester, N-acetyl-L-glutamic acid). ST-3663 and ST-42 tended to have higher maximum values for D-lactic acid methyl ester while ST-3673 and ST-3845 tended to have higher values for the other three carbon sources (citric acid, glycylLglutamic acid, N-acetyl-L-glutamic acid). In another branch ST-3663, ST-3673, ST-3845, and ST-42 were split by L- fucose, formic acid, α-hydroxybutyric acid, bromosuccinic acid, D-malic acid and quinic acid. Differences in the maximum values for these carbon sources split them into a group of ST-3845 and ST-42, and another of ST-3663 and ST-3673. ST-3663 and ST-3673 were not split from each other in this branch. ST-42 showed high maximum values for α-hydroxybutyric acid and D-malic acid but ST-3845 did not which resulted in a split. However ST-3845 showed generally greater values for L-fucose, formic acid, bromosuccinic acid and quinic acid than ST-42 which resulted in a split. In the third branch on the far right of Figure 4.2, it showed a group of ST (ST-3663, ST-3673, ST-3845, ST-42) with generally high maximum values for; succinic acid, L-aspartic acid, L-proline, L-lactic acid, L-glutamic acid, D,L-malic acid, L-asparagine, α-Ketoglutaric acid, L-glutamine, fumaric acid, L-serine, methylpyruvate, L-malic acid, pyruvic acid and succinamic acid. ST-3673 did split off from ST-3663, ST-3845, and ST-42 on two separate lower branches with ST-3673 consistently showing higher values with the 14 carbon sources.

Heatmap

Figure 4.3 shows the pattern of single source carbon utilisation using both PM1 and PM2A plates with the isolates clustering in a pattern consistent with their ST. The resulting pattern shows a clustering of the three host specific ST-2381 isolates (W83a, W194b, and W135a) and the ST-3655 isolate (W135a) with all four showing activity in a more limited range of wells (3–10) compared to the other isolates. The three host generalist ST-42 isolates (S263a, N31, N191) cluster with the ST-3663 isolate (W63b) and show activity in a much wider range of wells (17–20). The third cluster has the ST-3845 isolates (N53, N27) and the ST-3673 (W120a) isolate, which have some similarity with the other clusters by utilizing a wide range of wells (14-22) including three absent in ST-42 (N-Acetyl-L-Glutamic acid, L-Fucose, Quinic acid).

Both Figure 4.3, Table 4.2 and RE-EM tree analysis show the ST-42, ST-2381 and ST-3655 isolates are unable to use L-fucose at 42℃, but ST-3845, ST-3673 and ST-3663 do. In comparison to the reference genome ST-42 and ST-2381 lack the genomic island sequences at Cj0480c–Cj0490 but

ST-3845, ST-3673, ST-3663 and ST-3655 have the sequences present.

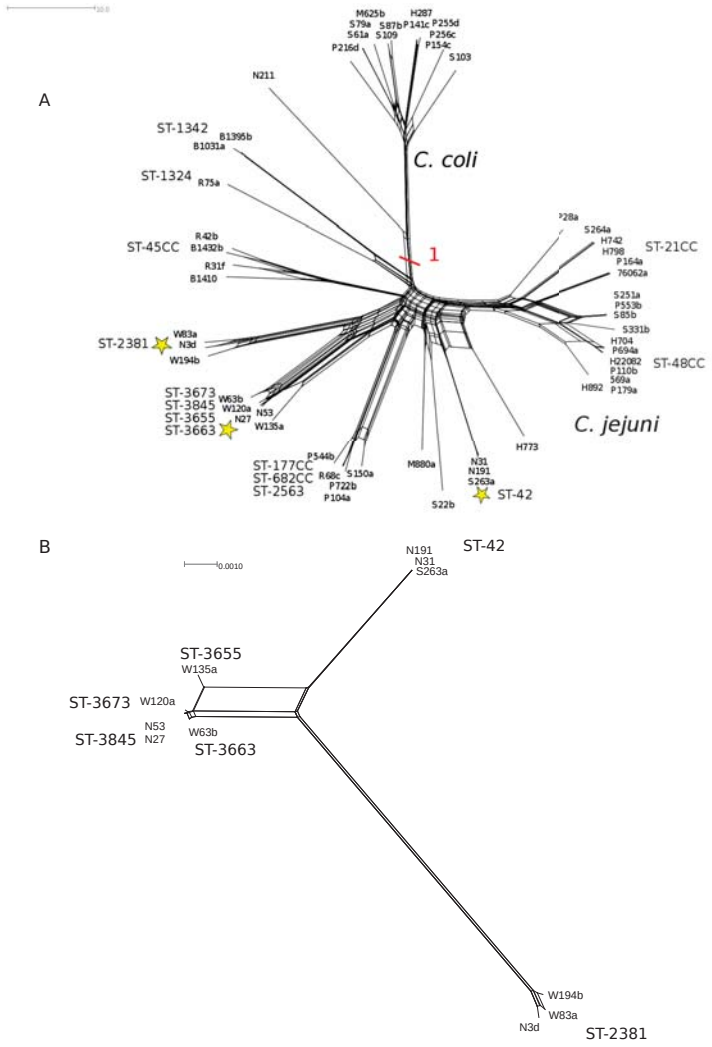


Figure 4.1: **Figure A shows the phylogenetic relationship based on their rMLST profile of 46 *Campylobacter jejuni* and *C. coli* isolates to the 11 isolates analysed with the Biolog Phenotypic microarrays.** The three stars identify the three clusters formed by the 11 isolates analysed with the Biolog Phenotypic microarrays. The red line with a 1 shows the split between the *Campylobacter jejuni* and *C. coli*. The 52 rMLST alleles were compared in a distance matrix then represented as a Neighbor-Net in SplitsTree [487]. **Figure B shows the phylogenetic relationship based on the rMLST sequences of the 11 *Campylobacter jejuni* isolates analysed with the Biolog Phenotypic microarrays.** The concatenated sequences of the 52 rMLST loci forming the allelic profile were aligned in Geneious (ver 7.1.7) using Clustal W [486] and represented as a Neighbor-Net in SplitsTree [487]

Figure 4.2: **RE-EM tree partitioning the effect of temperature, sequence type (ST) and carbon source (well) on the *C. jejuni* isolates**. The nodes of the tree represent the fixed effects (temperature, ST and carbon source) being partitioned into a tree with isolate as the random effect. The blue node represents temperature, the pink node represents carbon source, the orange node is ST, and the green node is a terminal leaf. Under each ST node there is the list of ST and how they were split. The first and most important partition is based on the effect of temperature on the isolates.

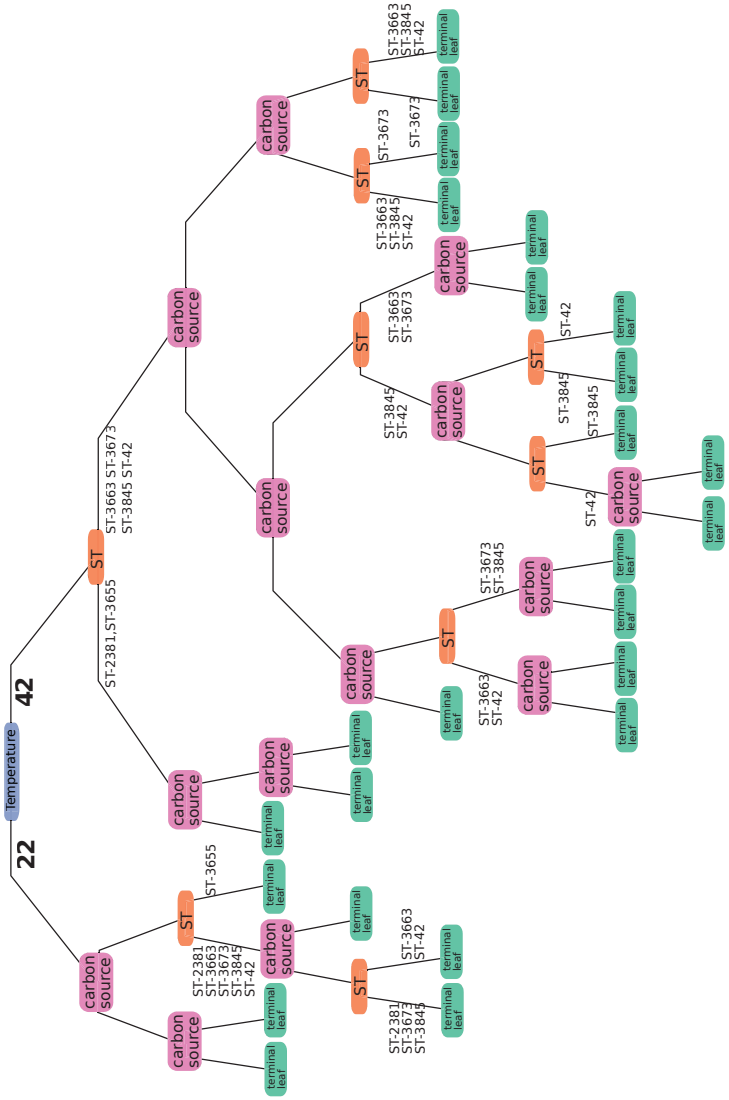
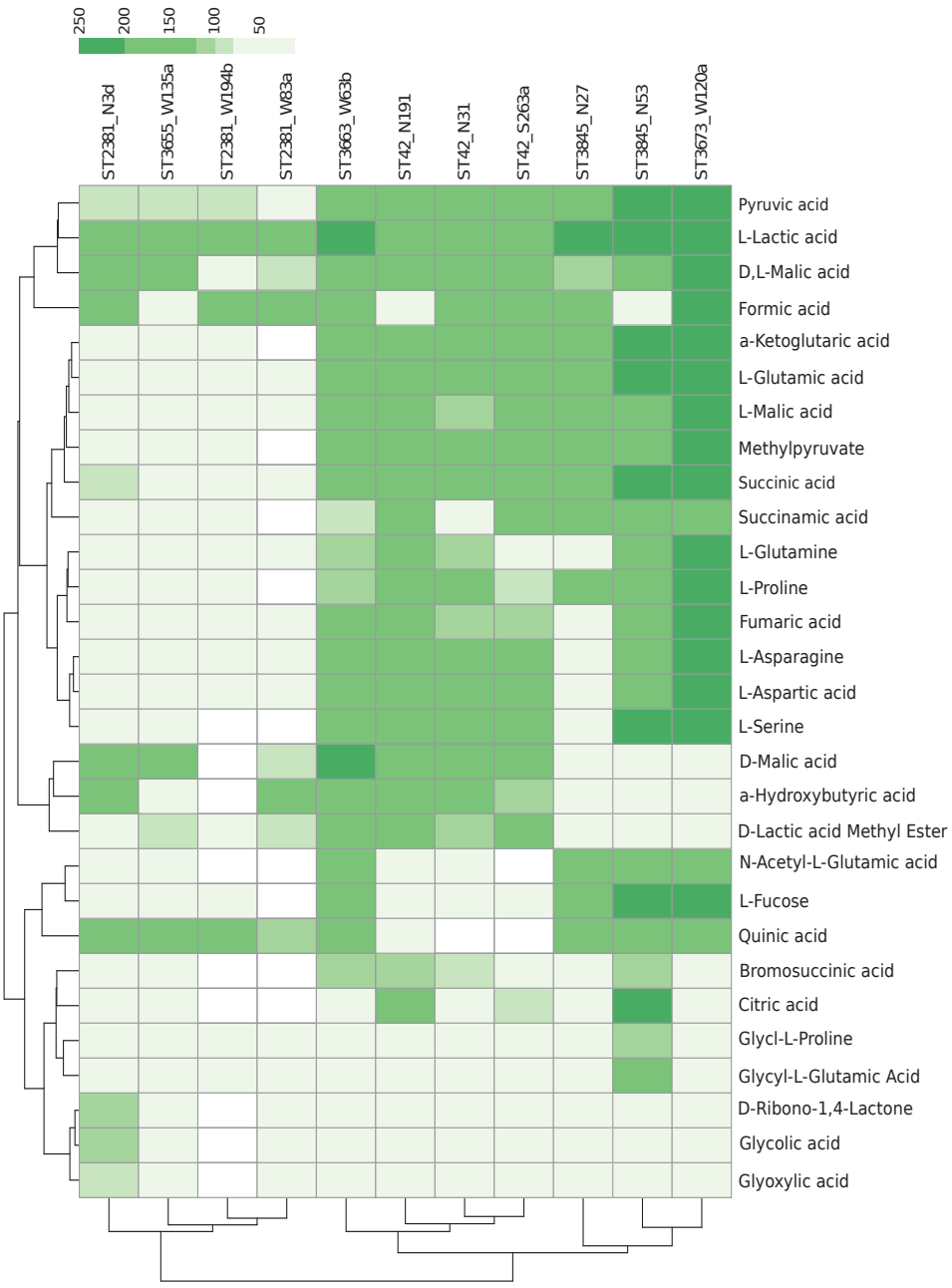


Figure 4.3: **Single source carbon utilisation at** 42℃ **in the phenotypic microarray plates by *C. jejuni* isolates associated with agricultural and wild-bird hosts.** The isolate name and ST are combined to identify each row. The carbon source is identified in each column. The horizontal and vertical dendrograms group the similar rows and columns.



4.4 Discussion

This investigation showed a clear difference based on phenotypic variation at 42℃ between two host specific sequence types (ST-2381, ST-3655) which have been found in wild-birds (pūkeko and takahē) and the host generalist sequence type ST-42 which is found in agricultural livestock. However ST-3663 has also been found in wild-birds (takahē) and has shown a phenotype more similar to ST-42, which does not support the hypothesis that the more host specific wild-bird STs utilise less carbon sources than the host generalist livestock associated STs. ST3673 and ST-3845, which share MLST alleles in their allelic profiles with both ST-42 and ST-2381, based on the number and type of carbons substrates utilised were more similar to ST-42 as seen in Figure 4.3 and Figure 4.3. Phenotypic diversity within *C. jejuni* ST in Biolog phenotypic microarray analysis has been previously report in ST-21, a host generalist ST, and the results did not correlate with isolation source [5]. A possible explanation for isolates from sequence types associated with the same host(s) showing a large amount of phenotypic variation is they may occupy different niches within the same host such as occupying different sections of the gastrointestinal system [421], and the demands of the with-in host niche select for different metabolic abilities. Alternatively it could be another example of phenotypic plasticity in *C. jejuni* [182] or there may be varying degrees of host specificity and host generalisation within *C. jejuni*. For example, ST-42 may not be as wide a generalist in its behaviour as ST-21 .

At 42℃ the isolates results clustered in patterns consistent with their ST and rMLST type, but not consistent with the number of shared alleles in the MLST profile, (Table 4.1) or the closeness of association shown by rMLST typing in Figure 4.1 A and B. ST-2381 and ST-3655 show a similar limited palette of activity in the phenotypic microarray plates at 42℃, although they only share one MLST allele, *gly*A (282). Similarly ST-42 and ST-3663 also show a wide range of activity but only share one MLST allele, *unc*A (3). In contrast both ST-3845 and ST-3673 share two MLST alleles with ST-42 and ST-2381, but they do not cluster with ST-42 or ST-2381 in Figure 4.3. The phylogenetic relationship based on rMLST typing in Figures 4.1A and B shows the isolates form three branches separating into ST-2381, ST-42 and a cluster of ST-3845, ST-3663, ST-3655 and ST-3673. This supports the ST being consistent with rMLST, and while the ST seems to predict, allowing for some variation, the phenotypic profile within a sequence type, it does not fully predict the phenotypic profiles of related sequence types.

The MLST typing system is based on housekeeping genes that are conserved and often part of the core genome, but metabolic phenotypes can be the result of genes in the flexible accessory genome. It has been suggested that in microbial genomes some genes associated with metabolic reactions, like the ability to use simple carbon substrates, are less conserved instead being more easily gained and lost [496] [497]. An example of this flexibility may be the ability to utilise L-fucose in *C. jejuni*

which is associated with a genomic island [474] [210]. Although the ability to utilise L-fucose was found in ST-3845, ST-3673 and ST-3663, it was not found in the ST-42 isolates. It has been reported in the generalist ST-21 clonal complex (CC) but not yet in CC45 or CC61 [421] [498]. The L-fucose related genomic island insert (cj0480c–cj0490) was found in the ST-3845, ST-3673, ST-3663 and ST-3655 isolates but not in the ST-2381 or ST-42 isolates. Three metabolic traits (γ-glutamyl-transpeptidase (GGT), fucose permease (*fuc*P), and secreted L-asparaginase (*ans*B) have previously been demonstrated to be associated with MLST (ST-22, ST-283, CC45, ST-586) [498]. These results suggest there is a relationship between the MLST/rMLST type, which are based on house-keeping genes (core genome), and metabolic phenotypes which are more associated with the flexible accessory genome, in both the agriculture-associated isolates and the wild-bird associated isolates. Further investigation is required to identify the basis of cohesion within sequence types based on carbon source phenotype while at the same time the degree of shared MLST allele profile does not predict phenotypic behaviour.

N-acetyl-L-glutamic acid was utilised by four isolates spread across (ST-3845, ST-3673, ST-3663) at 42℃, using the Biolog phenotypic microarray carbon sources and has not been previously reported by either Tang et al (2010), Line et al (2010), or Wagley et al. (2014) [206] [187] [207]. Quinic acid was utilised at 42℃ by all the pūkeko and takahē associated isolates and ST-3663, but not the ST-42 isolates and has not been previously reported by either Tang et al (2010), Line et al (2010), or Wagley et al. (2014) [206] [187] [207]. The isolates that utilised N-acetyl-L-glutamic acid and quinic acid, except for ST-3663, are associated with the wild-bird hosts, takahē and pūkeko, while the more common agriculture-associated ST-42 isolates did not utilise these carbon sources. This suggests that by investigating isolates from more wild-birds we may find more phenotypic variation within *C. jejuni* than already found [17].

An ability to utilise L-fucose at 42℃ and the presence of the related genomic island insert (cj0480c–cj0490) [474] [210] was shown by ST-3845, ST-3673 and ST-3663. ST-3655 has the genomic island insertion (cj0480c–cj0490) required for L-fucose utilisation but it did not utilise the sugar in this investigation. Some possible reasons for this include a non-functional enzyme sequence or phase variation silencing the gene. A future investigation may consider adding H_2 to the microaerobic environment, the addition could change the phenotype being expressed. However H_2 may also reduce the tetrazolium dye in the Biolog PM plates, so an alternative detection method may be required.

In total 29 of the carbon sources from PM1 and PM2A showed significant activity at 42℃, but only a subset group of seven were identified as similarly active at 22℃. At 22℃ *C. jejuni* is not actively growing so the demands and processes necessary to remain viable change [407]. We suspect, at least in part, the decrease may be due to a down shifting of the level of metabolic activity between 42℃ and 22℃, similar to that reported when *C. jejuni* goes from extracellular to

intracellular [499], or when a culture changes from growth phase to stationary phase [500]. This general down shifting of metabolic activity may in part explain how such a fastidious organism as *C. jejuni* survives outside the host, i.e. a combination of not growing and either reducing or stopping metabolic pathways that could react adversely in the external environment outside of a host [409]. However this is not the only possible explanation as at 22℃ the temperature may have affected the kinetics of some metabolic reactions slowing the process so the 96 hour period may not have been long enough to identify all the activity. A possible solution would be to repeat the process with a higher concentration of bacteria and/or for a longer period so a slower level of activity might be detected or alternatively comparing the transcriptome of different host associated isolates at different temperatures.

At 22℃ with fewer single carbon sources showing activity there is no early clear separation of the ST-2381 isolates phenotype profile from that of ST-42 isolates as seen at 42℃ (Figure 4.2). This lack of differentiation may be due to the phenotypes recorded at 42℃ are those expressed in the host, whereas those expressed at 22℃ are in the external environment, which both share. However ST variation in phenotype at low temperatures cannot be ruled out as isolates from different hosts have shown variation in their ability to tolerate 4℃ [403].

A feature of the seven carbon sources showing activity at 22℃ (Table 4.3) are their close relationship to tricarboxylic acid cycle (TCA or citric acid cycle or Krebs cycle) and energy production. The carbon sources succinic acid, formic acid, and D malic acid are part of the TCA cycle. Succinamic acid can be used by some bacteria in a cyclic-imide-transforming pathway to produce fumarate which is part of the TCA cycle. L-lactic acid converts directly to pyruvic acid and pyruvic acid feeds directly into the TCA cycle. *C. jejuni* has a complex branched electron transport chain that can directly use as electron donors various molecules such as hydrogen, NAD(P)H, formate, malate and succinate, as well as being able to use a range of alternative electron acceptors including fumarate [501] [502][pp275–292].

The continued activity in the metabolic area of energy production at 22℃ is in agreement with a report of up regulation of genes involved in energy metabolism at 5℃ and 25℃ against a general background of down-regulation of most other genes [409]. It also suggests that although other metabolic processes may have stopped or been reduced at 22℃, disinfection compounds like benzalkonium chloride that can inhibit endogenous respiration by the TCA cycle should be effective. However the ability of *C. jejuni* to persist in protective biofilms may change the efficacy of the disinfection agent.

The phenotypic microarray plate method relies on the reduction of tetrazolium dye to indicate the utilisation of a carbon source, but sometimes false positives can occur, for example pentose sugars in a microaerobic environment [187]. At 22℃ when the cells are stressed, increased cell death followed by cell lysis, could have caused reduction of the tetrazolium dye by the released

enzymes acting on the carbon and producing a false positive reaction. It has been demonstrated that periplasmic formate dehydrogenase can reduce formate and produce NADH using membrane vesicles i.e. an intact live cell was not required [503]. This may explain why three isolates that did not utilise formic acid at 42℃ did at 22℃ but not why another three isolates that did utilise it at 42℃ did not at 22℃. It is known molecules other than NADH can donate electrons to the *C. jejuni* electron transport chain [188], or act directly to reduce the tetrazolium dye [187] so it may be best to compliment this low temperature phenotype investigation by repeating it using another detection method.

We experienced some variability between replicates in some wells with the Biolog phenotypic microarray system, variability with this method has been previously reported when characterising both *Pseudomonas aeruginosa* strains [504] and *C. jejuni* [207]. A possible explanation is phase variation which is known to occur in a many bacteria including *C. jejuni* [505]. Phase variation is a reversible, stochastic and high frequency event where clonal bacterial populations can generate a heritable but reversible change in the level of expression in the affected gene. Various underlying mechanisms can cause phase variation, including epigenetic methylation state change, site specific recombination and simple sequence repeats (SSR). An analysis of eight *C. jejuni* genomes showed a high number of phase variable genes with short tracts of homologous repeats (SSR), but there was poor conservation of these phase variation mechanisms at specific loci [506]. Most reports of phase variation in *C. jejuni* refer to surface associated structures such as the capsule, lipo-oligosaccharide and the flagella, which would suggest the phase variation may be affecting transmembrane transport proteins rather than the enzymes involved in the reaction [507] [389]. However in *Haemophilus influenzae* phase variable genes have been identified affecting components of type I and type III restriction modification (RM) systems with the potential to affect the expression of multiple genes, so more investigation would be required to identify if this is phase variation and what is being affected [508] [509] [510].

The reactions reported at 42℃ by this investigation occurred *in vitro*, and may not reflect the *in vivo* metabolism, particularly the intracellular state, as it has been shown that there is a down-shift in metabolic activity when *C. jejuni* transitions from extra-cellular to intracellular [499]. It has been suggested that intracellular pathogenic bacteria shift their metabolism in a host even to the level of being strain specific [511]. This phenotype analysis was *in vitro* with single carbon sources, a situation that does not duplicate the environment in the gastrointestinal system, where there are multiple carbon sources and interactions between competing bacteria and the host, so these results may be more indicative of potential rather than examples of events within the host.

In this analysis we followed the manufacturers protocol which did not include starvation of the isolates prior to inoculation into the minimal single carbon source media of the Biolog phenotypic microarray plates. It is possible some carbon source from the original Columbia horse blood agar

on which they grew could have been transferred with the bacteria but if this had a significant affect on the result it should have affected every well, as the swab from the plate inoculated a liquid media (IF0a) which was then pipetted into all wells. An option not taken in the experiment was starvation of the isolates prior to inoculation into the phenotypic microarray plates. This could change the state of the *C. jejuni* isolates e.g. some could form coccoids and enter the viable but non-culturable state (VBNC) [512]. It is likely such a transformation would affect the resulting carbon source utilisation phenotype expressed and this could reflect future work on the behaviour of *C. jejuni*.

Occasionally a control strain, like NCTC 11168, is used in Biolog phenotypic microarray analysis, for example when identifying a new feature like L-fucose utilisation[210]. In this analysis as the intent was to compare new poorly understood ST on the basis of their host association and phylogenetic relatedness, a control strain was not used as we relied on the highly standardised nature of the Phenotypic microarray plates, their preparation and the Omnilog Reader environment to ensure a level of consistency.

The isolates used were stored at −80℃ from 2006 to 2010 before being restored and grown for the experiments performed in 2012. Each isolate was regrown after removal from the −80℃ storage, and while there was slight variation in colony size between all the isolates, they grew well on the Columbia horse blood agar suggesting that the long storage period had not adversely affected them. However the ST-2381 and ST-3655 isolates, stored between 2006-2008, showed less utilised fewer carbon sources than the other isolates stored between 2006-2010 (Table 4.1). It is possible this storage could have effected their behaviour on the Biolog phenotypic microarray plates, but in contrast ST-3663 and ST-3673 were also in storage from 2006, and utilised many more carbon sources. This would suggest if there was a loss of metabolic ability related to time in storage it did not affect all isolates equally, nor did it affect there ability to be successfully regrown.

Most of these isolates were taken from environmental water, except for two ST-42 isolates (S263a, N191). Originally this analysis was examining the novel ST-2381 and related ST which were commonly found in New Zealand environmental waters, but with no known host [8], however over a time ST-2381, ST-3655 and ST-3663 have isolated from the pūkeko and or the takahē [16]. It has been shown that *C. jejuni* STs can be used for source attribution to a host [67] and they can develop a genomic host signature [315]. This behaviour combined with *C. jejuni* from wild-birds showing host specificity in their ST [318] [7] [319], strongly suggests that although the isolates originate from water their hosts was one of these two rails. ST-3845 and ST-3673, while not found in either bird yet, have a very similar genomic signature with the ST-2381, ST-3655 and ST-3663 suggesting they may have these or a closely related bird as their host. Support for this is shown by ST-42 sharing part of their MLST allelic profile with ST-3845 and ST-3673, but in the genomic comparison, they are quite separate.

It is reasonable to ask if using mainly water isolates has affected the results, rather than an isolate direct from a bird may have on the phenotype. As Gripp et al. (2011) demonstrated with ST-21 from different sources that there was no correlation between phenotypic variation and source using Biolog phenotypic microarrays[5], it suggests that the origin being water rather than animal may not significantly effect the PM results.

The inability of the Biolog phenotypic microarray analysis to differentiate between the generalist and specialist lifestyle of the isolates at either 42℃ or 22℃, does not mean there is no phenotypic difference between these lifestyles in *C. jejuni*. It does suggest the basis of the differences may not be in the number of single source carbons utilised by an isolate. There are many other features that could be involved,such as it may be a matter of differences in specific carbon sources, e.g. the use of quinic acid by some isolates, or it could be related to another non-carbon related metabolic system, or conditions under which a metabolic feature is expressed and optimised.

The Omnilog phenotypic analysis involved a total of 11 isolates, which is a moderate sample number. Within this set variation was identified between repetitions of the same isolates which we suggest may be due to phase variation. Variation was shown between isolates within the same ST, but they still clustered together when compared between STs at 42℃. Such behaviour is consistent with reports from Gripp et al. (2011) of phenotypic flexibility and high genetic micro-diversity in ST-21 when using Biolog phenotypic microarrays [5]. This would suggest that the sample size was large enough to show a level of variation, and that a larger number of isolates may have produced even more variation.

It is important to continue to characterise the metabolic activity of *C. jejuni* at room temperature and below, possibly supplementing biochemical tests with other techniques, like proteome and transcriptome analysis at a series of temperatures below < 30℃, to see how the profile changes at different temperatures and with different sequence types. Continued analysis with a range of isolates from more wild-bird as well as those associated with livestock will identify the extent that host differences may affect response to temperature. *C. jejuni* can transform into a viable but non-culturable (VBNC) state when under stress or at a low temperature and in this state the transcription levels of some virulence associated genes have been shown to be low but detectable [513]. Investigating to what extent there is a down-shifting and switching off of pathways/reactions, and/or if there is more temperature related reduction in reaction kinetics by *C. jejuni* will help explain its behaviour, and hence survival at lower temperatures. Understanding how the metabolic activity changes with temperature may help with kitchen hygiene and cleaning regimes for surfaces and materials that may be contaminated by *C. jejuni*. For example if a metabolic process is important at 22℃, such a metabolic process could be targeted in home kitchen disinfection.

Conclusion

This study identified phenotypic variation between all the isolates at 42℃. The pattern of variation between ST-42 and ST-2381/ST-3655 did suggest the agriculture/generalist lifestyle associated sequence type (ST-42) utilised more carbon sources at 42℃ than the wild-bird/specialist lifestyle associated sequence type (ST-2381), however this simple pattern did not extend to the other isolates of the sequence types ST-3663, ST-3845 and ST-3673. Although the isolates phenotypes did vary they did cluster on the basis of ST, but not on the amount of shared alleles in a ST. The phenotypic profile at 22℃ showed a much reduced number of carbon sources showing activity compared to 42℃. The carbon sources showing activity at 22℃ tended to be closely related to the TCA cycle and energy production.

Acknowledgements

The authors wish to thank Prof. David Kelly and Dr Anne Midwinter for his helpful comments. This work was supported by the Marsden Fund (grant number MAU0802) and the Australian Research Council (grant number FT100100031). This work was part funded by Dr Julie Collins Emersons' Massey University BRSCA fund.

Supplement

Table 4.4: **The 32 *Campylobacter jejuni* and 14 *Campylobacter coli* used in the rMLST comparison.** The Source refers to the material from which the *Campylobacter* were isolated. The multilocus sequence type (MLST) is reported as sequence type (ST) and the clonal complex (CC) if known.

Isolate	Source	Species	ST	Clonal complex
P110b	chicken offal or meat	*C. jejuni*	474	CC48
H22082	human stool	*C. jejuni*	474	CC48
H798	human stool	*C. jejuni*	50	CC21
H892	human stool	*C. jejuni*	48	CC48
P104a	chicken offal or meat	*C. jejuni*	45	CC45
P164a	chicken offal or meat	*C. jejuni*	190	CC21
P694a	chicken offal or meat	*C. jejuni*	474	CC48
S85b	cattle	*C. jejuni*	53	CC21
P216d	chicken offal or meat	*C. coli*	3230	CC828
P256c	chicken offal or meat	*C. coli*	2397	CC828
P255d	chicken offal or meat	*C. coli*	2397	CC828
S109	sheep	*C. coli*	3232	CC828
P722b	chicken offal or meat	*C. jejuni*	25	CC45
P141c	chicken offal or meat	*C. coli*	1581	
P154c	chicken offal or meat	*C. coli*	1581	
S331b	cattle	*C. jejuni*	21	CC21
R68c	wild-bird	*C. jejuni*	45	CC45
R42b	wild-bird	*C. jejuni*	2539	CC71
76062a	human stool	*C. jejuni*	190	CC21
S87b	sheep	*C. coli*	3232	CC828
H742	human stool	*C. jejuni*	50	CC21
B1432b	starling	*C. jejuni*	681	CC682
S22b	cattle	*C. jejuni*	2026	CC403
S79a	cattle	*C. coli*	3072	CC828
B1031a	starling	*C. jejuni*	1342	
S264a	cattle	*C. jejuni*	50	CC21
M625b	pork offal or meat	*C. coli*	1016	CC828
P256c	chicken offal or meat	*C. coli*	2397	CC828
B1395b	starling	*C. jejuni*	1342	

Continued on next page

rMLST comparison isolates —– *continued from previous page*

Isolate	**Source**	**Species**	**ST**	**Clonal complex**
M880a	lamb offal or meat	*C. jejuni*	2341	CC61
569a	human stool	*C. jejuni*	474	CC48
H704	human stool	*C. jejuni*	474	CC48
R31f	wild-bird	*C. jejuni*	2536	
H773	human stool	*C. jejuni*	3711	CC257
P179a	chicken offal or meat	*C. jejuni*	474	CC48
N211	environmental waters	*C. coli*	3302	
R75a	wild-bird	*C. jejuni*	1324	
S61a	cattle	*C. coli*	3072	CC828
S150a	cattle	*C. jejuni*	45	CC45
B1410	starling	*C. jejuni*	177	CC71
P553b	chicken offal or meat	*C. jejuni*	53	CC21
S251a	cattle	*C. jejuni*	53	CC21
P28a	chicken offal or meat	*C. jejuni*	520	CC21
H287	human stool	*C. coli*	1581	
P544b	chicken offal or meat	*C. jejuni*	45	CC45
S103	sheep	*C. coli*	3299	CC828

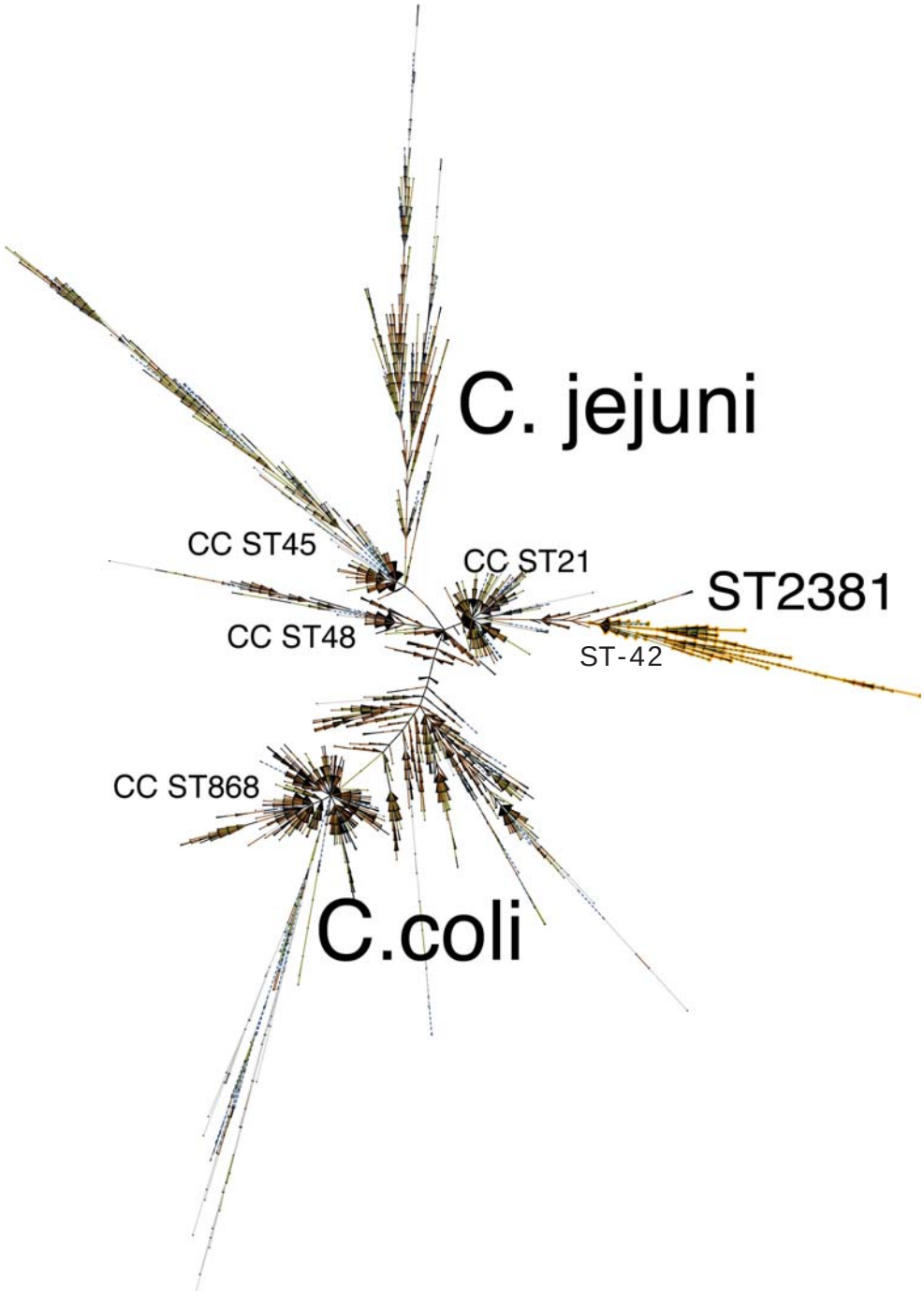


Figure 4.4: **Minimum spanning tree of PubMLST isolates showing a branch with ST-2381 and associated isolates.** On 5^{th} October 2011 the MLST allelic profiles of 5092 sequence types held in the PubMLST database were downloaded and analysed into a minimum spanning tree using the BioNumerics (Applied Maths, Austin, TX, USA) software. The tree separated *C. coli* from *C. jejuni*, which divided into three major branches for CC21 (centred on ST-21), CC45 (centred on ST-45) and CC48 (centred on ST-48). Another branch, highlighted in yellow, contains ST-42, ST-2381, the associated ST used in this chapters analysis as well as ST-179, ST-682 and ST-177.

Appendix B in DVD

Summary of information provided in Appendix B

For completeness, Appendix B provides further details about the data and methods applied in this study. For ease of access Appendix B is on the DVD that is provided with the thesis, and is accessible by opening the file index.html with an internet browser (tested on Safari and Firefox). The file may need to be loaded onto a computer hard drive and unzipped.

1. The supplementary data for the chapter "Different carbon utilisation patterns in *Campylobacter jejuni* associated with host generalist or specialist lifestyle and temperature"

2. R script used in analysis - The R scripts used to generate the results from the Biolog plate reader data.

3. Adjustment factor for the effect of Omnilog plate reader service - The development of the Adjustment factor for the effect of Omnilog plate reader service.

4. Visualisation of adjusted data - The graphical results of the R scripts for the 11 *C. jejuni* isolates tested against PM1 and PM2A plates on an Omnilog plate reader including box-plots, density plots and kmeans. The density plots show the frequency of occurrence of a given value in the data and represents an estimate of the probability density function.The clustering of the maximum height (A) parameter values of each well by k-means. k-means is a method of clustering that partitions n observations into k clusters, so each observation belongs to the cluster with the nearest mean, i.e. the k-mean serves as Visualisation of unadjusted data - XY-plots, level plots, radial plots and heatmaps of the unadjusted results for each phenotypic microarray well at each temperature and for each isolate.

5. Visualisation of unadjusted data - XY-plots, level plots, radial plots and heatmaps of the unadjusted results for each phenotypic microarray well at each temperature and for each isolate.

6. Pilot studies for Chapter 4 - Three pilot studies related to using the Biolog Phenotypic Microarrays and *C. jejuni* isolates. Three trials were undertaken to assess the function of the Omnilog Reader and Biolog Phenotypic Microarrays under conditions used in the analysis. Pilot Study 1, the affect of culture age on PM plate results. Pilot Study 2, the concentration of *C. jejuni* used in the PM plates. Pilot Study 3, the effect of room temperature on *C. jejuni* growth.

7. Genomes mapped to KEGG pathways - The 11 *C. jejuni* isolates have had their genomes sequenced and mapped to KEGG pathways. The maps for each pathway are provided as web pages.

8. Biological and Technical replicates - Table of Biological and Technical replicates used in the Phenotypic microarray analysis.

9. Comparison of five different Biolog phenotypic microarray investigations - Compares the results of this chapter with four other published investigations into different C. jejuni isolates.

10. Original Omnilog output - Screenshots of the Omnilog Reader Software output, comparing two different isolates.

The purpose of computing is insight - not numbers.

R. W. Hamming (1962)

5

the genome:

Characterisation and comparison of *Campylobacter* spp. isolates from Australian purple swamphen (*Porphyrio porphyrio melanotus*) with a New Zealand dataset.

Prelude

This chapter is a wide ranging genomic comparison of *Campylobacter* isolates, performed to address the question of has *C. jejuni* followed a unique evolutionary trajectory in New Zealand? To do this isolates from a bird species in Australia were compared to a lineage associated with a sub-population of the bird species in New Zealand. Genomic comparisons were also made to a dataset of New Zealand *Campylobacter* genomes. This provided both a background for the Australian and New Zealand bird isolates, and gave a wider ranging genomic comparison of agricultural-associated and wild-bird associated *C. jejuni* genomes, thus investigating the genome for a "host signature" that

could differentiate those ST from livestock to those found in wild-bird. In the previous chapter the phenotype was investigated for differences between those ST associated with livestock and those from wild-birds, now this comparison is being carried into the genotype.

In this chapter the genomes from *Campylobacter* spp. isolates collected from the Australian purple swamphen (*Porphyrio porphyrio melanotus*) were compared to STs associated with the New Zealand pūkeko (*Porphyrio porphyrio melanotus*) and takahē (*Porphyrio hochstetteri*) to investigate if the temporal and geographical separation between them has led the *Campylobacter* spp. along a unique evolutionary trajectory, testing a null hypothesis of no divergence between the populations. The host specificity, seen in wild-bird *C. jejuni* STs has been reported to transcend geographical and temporal separation [7]. This suggests a genomic comparison between the isolates found in the Australian purple swamphen and the New Zealand pūkeko will show a shared lineage and identify if there has been any diverging evolution. To provide context for the genomic comparison of *Campylobacter* spp. isolates, a dataset of New Zealand isolates draft genomes was used. The dataset contained the draft genomes of New Zealand isolates that were collected and sequenced by mEpiLab for a range of projects [273] [16]. This dataset was not used as statistically representative group of New Zealand isolates, but as a collection that covers isolates from a range of New Zealand sources, covering livestock, wild-birds and the environmental water.

The ancestors of *Porphyrio porphyrio melanotus* are thought to have arrived in Australia ∼600,000 years ago, while the pūkeko are a geographically separated sub-population that arrived recently in New Zealand ∼500 years ago [11]. The South island takahē (*Porphyrio hochstetteri*) is closely related to both [13] [14], and a much earlier arrival to New Zealand, with fossils dating back ∼38,000 years found in multiple sites around the South Island [514][p38]. It is thought to have arrived in New Zealand ∼2.5 million years ago, and share a volant ancestor with the *Porphyrio porphyrio melanotus* [11].

Genomic characterisation of Australian purple swamphen *(Porphyrio porphyrio melantotus) Campylobacter* spp. and comparison with New Zealand *Campylobacter* spp.

5.0.1 Abstract

Campylobacter jejuni (*C. jejuni*) is an important human pathogen worldwide, while other members of the *Campylobacter* genus are much less frequent causes of human illness. Wild-birds are a source of *Campylobacter* spp., but little is known about the comparative genomics and evolution of *Campylobacter*spp. from wild-bird populations. Samples were collected from the Australian purple swamphen (*Porphyrio porphyrio melanotus*) to compare to isolates associated with the New Zealand pūkeko (*Porphyrio porphyrio melanotus*), which has been a geographically separate sub-population for an estimated 400–600 years. The endemic South Island takahē (*Porphyrio hochstetteri*) is closely related to both, but is a much earlier arrival to New Zealand, and shares some *Campylobacter* with the pūkeko. The genomes of the *Campylobacter* spp. isolates from Australian purple swamphen were characterised and compared to New Zealand pūkeko and takahē isolates, to investigate if the temporal and geographical separation had led to the *Campylobacter* spp. following a unique evolutionary trajectory. To give a context to this comparison a range of New Zealand isolates associated with both wild-bird and agricultural-related hosts were also compared using core genome and pan-genome features. There was a close relationship between *C. jejuni* from the Australian purple swamphen and isolates associated with the New Zealand pūkeko and takahē, in both the core genome and pan-genome analysis that supports a recent common ancestor followed by divergence after separation. A consistent pattern of relationship was found between isolates, both agricultural and wild-bird associated, both in the core genome of *C. jejuni*, as measured by MLST and rMLST, and extending into the patterns of gene presence and absence in the pan-genome. This consistency starts to explain how sequence types (ST) and their clonal complexes (CC) have been successfully attributed to particular hosts, it also suggests there may be some factor restricting recombination between dissimilar ST. A putative new *Campylobacter* spp. was also identified.

5.1 Introduction

Several *Campylobacter* spp. are known to be pathogenic in both people and animals [502][pp5-6], and *C. jejuni* is considered to be the most common cause of gastro-enteritis worldwide [96] [515]. Most investigations have examined *C. jejuni* and *C. coli* which combined account for ∼95% of human disease related to *Campylobacter* spp. [516][p123]. *C. jejuni* exhibits phenotypic and genotypic variation [502][p59] [247] which in combination with being fastidious to grow has made it difficult to investigate [30]. However over the last decade a lot of progress has been made in understanding its epidemiology using multi-locus sequence typing (MLST) [517] [317]. The increased availability of fully sequenced genomes for investigations should greatly improve our understanding and open up new perspectives on the *Campylobacter* genus in many areas [518] even to a re-evaluation of phylogenetic relationships [519].

Epidemiological studies of *C. jejuni* using multilocus sequence typing (MLST) revealed sequence types (ST) associated with wild-birds show a niche specialist's, lifestyle while agriculture-associated ST were able to colonise a wider range of hosts (cattle, sheep and poultry) and had a more generalist lifestyle [6] [7]. Wild-bird STs tend to be associated with a single species, or closely related species, even over large geographical distances [7]. MLST based analysis also showed agricultural-associated ST were the predominant cause of clinical cases [318] [6]. Although this may reflect people being more exposed to agricultural animals, rather than non-pathogenicity in wild-bird associated ST [318], because occasionally ST associated with wild-birds have caused gastro-enteritis [324]

Much of the published genomic information on *C. jejuni* focuses on agriculture-associated isolates and clinical case [368]. Two exceptions were Hepworth et al. (2011) and Champion et al. (2005), who both identified a livestock (agriculture) clade and a non-livestock (wildlife/water) clade [368] [289]. Champion et al. (2005) used DNA microarrays in combination with Bayesian analysis but found the majority of *C. jejuni* isolates from clinical cases were in the non-livestock clade, which was noted as inconsistent with other findings [289]. Hepworth et al. (2011) used comparative genomic hybridisation and genomic sequencing of two non-livestock (wildlife/water) clade members and to show they shared deletions and insertions not present in the reference genome NCTC 11168 [368].

This analysis compares *C. jejuni* isolates from two geographically separated wild bird populations. *Porphyrio porphyrio*, is a member of the rail family and has an extensive range around the world [520] [11]. Based on recent phylogenetic estimates, the subspecies *Porphyrio porphyrio melanotus* arrived in Australasia about 600,000 years ago [11]. The faeces of Australian purple swamphen (*Porphyrio porphyrio melanotus*) in Victoria, Australia, were sampled for *Campylobacter* spp.. Genomic comparison were made to *Campylobacter* spp. isolates associated with the New Zealand pūkeko (*Porphyrio porphyrio melanotus*), a sub-population is thought to have arrived in New

Zealand as recently as 400–600 years ago [12][pp 590-595] [514][pp 38] [11]. The South island takahē (*Porphyrio hochstetteri*) is closely related to both birds [13] [14], and also associated with some of the isolates used in this analysis. The South island takahē are though to have arrived in New Zealand ∼2.5 million years ago [11], and fossil evidence dating back ∼38,000 years [514][p38]. Both the pūkeko and the takahē, are a source of new *Campylobacter* spp. [16][p227], and associated with novel *C. jejuni* found in environmental waters [232].

The two aims of this investigation were first to identify the relationship of the Australian purple swamphen *Campylobacter* spp. isolates, to those of the New Zealand pūkeko and the takahē, by characterising Australian purple swamphen isolates then comparing the genomes. The second, was to compare the relationships between *Campylobacter* spp. isolates from both wild-bird, human and agricultural hosts at two different levels of the genome, the core genome and pan-genome. The pan-genome consists of the core genome, defined as genes present in all isolates, and an accessory genome, which is composed of genes absent in one or more isolates or unique to an isolate (note the term "gene" in this chapter refers to orthologues/protein coding sequences). First the characteristics and identify of the Australian purple swamphen isolates was established. To identify the new *Campylobacter* spp., *in silico* DNA-DNA hybridisation [218] was be used in conjunction with 16S ribosomal RNA (16S rRNA) analysis. Based on next generation sequencing (NGS), the Australian purple swamphen isolates were given MLST and ribosomal multi-locus sequence typing (rMLST) allelic profiles [521] [298]. The mobilome of the Australian purple swamphen isolates was investigated by examining for features that indicate the presence of phages and genomic islands. Next, the genomes of the Australian purple swamphen isolates where put into context by comparing then to a dataset of New Zealand *Campylobacter* spp., which included ST associated to the pūkeko and the takahē as well those associated with wild-birds and agriculture. The core genome was compared on the basis of MLST and rMLST allelic profiles. The pan-genome was compared using a Dollo distance matrix and discriminant analysis of principal components (DAPC) investigating their phylogenetic relationships and population structure.

5.2 Methods and Materials

Biological materials and DNA extraction

Fresh faecal samples from Australian purple swamphen (*Porphyrio porphyrio melantotus*) in the greater Melbourne area of Victoria were swabbed and sent for culturing within 24 hours. The swab was inoculated into Boltons broth and incubated at 42℃ for 2 days, after which a sterile swab inoculated a sample of the broth onto modified charcoal cefoperazone deoxycholate agar (mCCD) which was then grown at 42℃ in a microaerobic environment produced by CampyGen sachet (Oxoid) for two days. The plate was then inspected for colonies that resembled *Campylobacter* and a single colony was taken and inoculated onto a sterile blood agar plate and grown at 42℃ in a microaerobic environment for two days. At this point the plate was checked for contamination. If none was detected the plate was swabbed and the swab transported to the Microbiological Diagnostic Unit Public Health Laboratory at Melbourne University where it was prepared for storage at −70℃. Horse blood agar plate was inoculated with the swab and incubated in a microaerobic environment at 37℃ and/or 42℃ depending on its growth. After 2 days the growth on the plate was inoculated into 20% glycerol and stored at −70℃. When required the isolates were regrown on horse blood agar for two days at 37℃ in a microaerobic environment. From this a loop of growth was taken and stabbed into Wang's medium and incubated at 37℃ in a microaerobic environment for $2-3$ days. The samples were then transported to the mEpilab where a swab was taken and grown on Columbia horse blood agar (Fort Richard, Auckland, New Zealand) in a MAC500 Workstation microaerobic environment (85% N_2, 10% CO_2, 5% O_2) for two days and inoculated into 15% glycerol solution in labelled vials for storage at −80℃. The isolates where regrown from the glycerol solution on horse blood agar in a microaerobic environment for 2 days before DNA extraction.

Whole genome sequencing

The genomic DNA (gDNA) was extracted from 14 Australian purple swamphen isolates (B2121-B2134) and four New Zealand *C. jejuni* isolates (N27, N31, N191, W194b). The gDNA was extracted using either the Wizard genomic DNA purification kit (Promega) or the QIAmp DNA mini kit (Qiagen). Both kits were used as per the manufacturers instructions. The genomic DNA samples were fragmented by nebulisation for 6 minutes at a pressure of 32 psi, purified, then end repaired, A-tailed, adaptor-ligated, fractionated, purified and enriched according to the manufacturer's instructions, using the TruSeq DNA LT Sample Prep Kit v2-Set A. The prepared libraries were normalized to equal molarity, diluted to 2nM and pooled; 20 libraries per pool. A flow cell was prepared for each of the library pools and sequencing reactions using 9 pmoles of the pooled libraries were performed on an Illumina MiSeq instrument with the MiSeq Reagent Kit v2. for a 2×250 base run. This resulted in approximately 12 to 15 million clusters per run.

The remaining 66 New Zealand *Campylobacter* spp. isolates had genomic DNA extracted and

were whole genome sequenced, assembled and mapped as described in Biggs et al [273]. These draft genomes are from a wide range of isolates collected by the mEpiLab, and used as the basis of comparison for the 14 Australian purple swamphen isolate genomes. This data of New Zealand set of *Campylobacter* spp. genomes contains three putative new species (*C.* sp. nov. 1, 2 and 3) that have been found in New Zealand wild-birds [16].

Assembly and annotation of the genomes

An in-house work flow based on Perl scripts was used to assemble, quality control and annotate the genomes. Solexa QA [483] was used for trimming and quality control of the short reads (Q30, expected probability of incorrect base calling 1:1000). Velvet (ver. 1.2.10) [484] was used for *de novo* assembly of the 2×250 base short reads. A range of kmers ($55 - 250$bp in steps of 10bp) were used in the assembly process and the characteristics of the assemblies were compared and ranked using N50, the number of contigs, maximum contig length and assembly length. N50 is a measure of length of the longest contig in an assembly where at least 50% of all base pairs are contained in contigs of this length or larger. The highest ranking contig assembly for a given genome was then annotated with Prokka (ver. 1.10) [485]. The customised Perl script also identified and listed the 16S rRNA genes from each of the 84 draft genomes. The annotated sequences were examined for the presence of Clustered regularly interspaced short palindromic repeats (CRISPR) by looking for CRISPR-associated (cas) sequences. If a Cas-2 sequence was identified it was assumed CRISPR were potentially present as it is a core element in the CRISPR-Cas system [522]. All the genomes used in this analysis, except the reference genomes used in the taxonomy analysis, were draft genomes.

Characterisation of *Campylobacter* spp. genomes from Australian purple swamphen

16S rRNA genes

The 14 Australian purple swamphen *Campylobacter* spp. genomes 16S rRNA sequences were checked in Barnap [523]. All the 16S rRNA sequences belonging to *C.* sp. nov. 1 isolates (4) were fragments and partial sequences so they were removed from this analysis, leaving only full length 16S rRNA sequences. The remaining ten Australian purple swamphen isolates were all 1501bp long and compared to 83 16S rRNA sequences from seven known *Campylobacter* spp. and three previously identified putative new species [16]. The full length 16S rRNA sequences from the 83 sequences were aligned in Geneious (ver.7.1.7) using Clustal W [486] and represented as a Neighbor-Net in SplitsTree [487]. The full identity of the 16 reference sequences are in the supplemental data (Table 5.4). The remaining 57 isolates are from New Zealand and are identified in Table 5.3.

The full length 16S rRNA DNA sequence of ten Australian purple swamphen isolates were also compared to the EzTaxon database [524]. Initially the sequences were compared by the BLASTN [525] program against the EzTaxon database which contains type strains with validly published prokaryotic names and representatives of uncultured phylotypes [524]. The top thirty sequences

with the highest scores were selected for the calculation of pairwise sequence similarity using global alignment algorithm implemented by the EzTaxon server[1] [524]. The final results showed the pairwise similarity of the 16S rRNA for 10 Australian purple swamphen isolates against 25 *Campylobacter* spp. containing nine subspecies. The pairwise similarity comparison was then represented as a heatmap [490].

in silico DNA-DNA hybridisation

A matrix of DNA-DNA hybridisation (DDH) values comparing various genomes was calculated by three approaches with JSpecies [218]. The average nucleotide identity (ANI) was calculated using BLAST ver. 2.2.18(ANIb), and MUMmer ver. 3.23 (ANIm) with a cut-off value of $95-96\%$ [526] [527]. A $95-96\%$ similarity cut-off is thought to reflect a laboratory performed 70% DDH [528]. The tetranucleotide frequency correlation coefficients (Tetra value) is an alignment free analysis of tetranucleotide signatures normally with a cut off of ≥ 0.99, however as draft genomes are used and this can result in more variation [529] [218]. All 14 Australian purple swamphen isolates were compared to 14 *C. jejuni*, one *C. jejuni* subsp. *doylei*, three *C. coli* and six other *Campylobacter* spp. including *Campylobacter curvus*, *Campylobacter lari*, *Campylobacter hominis*, *Campylobacter concisus*, *Campylobacter fetus*. The reference sequences used in the comparison are listed in Table 5.5.

The same three methods of measuring *in silico* DDH were also used to to compare the isolates to 15 New Zealand *Campylobacter* spp. isolates. The isolates represent *C. jejuni*, *C. coli* and three new putative *Campylobacter* spp. (*Campylobacter* sp. nov. 1, *Campylobacter* sp. nov. 2, *Campylobacter* sp. nov. 3) that have been associated with New Zealand wild-birds, particularly water rails. The New Zealand isolates are referred to in Table 5.3.

PathogenFinder

The draft genomes were uploaded to the PathogenFinder [530] website and compared to the ϵ-proteobacteria dataset. The algorithm compared the proteins in the isolates to those found in a dataset of known pathogenic and non-pathogenic ϵ-proteobacteria. Based on sequence similarity and association with pathogenicity, a probability for being a human pathogen was calculated. This process does not depend on identifying known virulence factors, but is based on the similarity of the genome to known pathogens genomes, so as yet unidentified proteins can be accounted for in the comparison.

Mobilome

Phage

The Australian purple swamphen sequences were uploaded as genbank files made in Prokka [485] to the SEED server [531] where they were re-annotated into the RAST system [532]. Using this

[1] http://www.ezbiocloud.net/eztaxon

format the phage-associated sequences were identified with PhiSpy [533] by using a combination of sequence similarity and composition-based methods to identify potential phage sequences. PhiSpy only identifies a prophage sequence when it is continuous for a minimum length equivalent to 40 protein encoding genes (peg) long.

Genomic Islands

The Australian purple swamphen draft genomes were uploaded to the IslandViewer website [534] [535] to computationally predict genomic islands using three distinct methods SIGI-HMM [536], IslandPath-DIMOB [537] and IslandPick [534].

Antibiotic resistance genes

The genomes were analysed for genes known to produce antibiotic resistance to aminoglycosides, β-lacatam, fluoroquinolone, fosfomycin, fusidic acid, macrolide-lincosamide-streptograminB, nitroimidazole, phenicol, rifampicin, sulphonamide, tetracycline, trimethoprim, and glycopeptide using ResFinder (ver. 2.1) [538]; a BLAST based on-line identification system. The minimum length was 60% and the percentage ID threshold was 90% (the percentage of nucleotides that are identical between the best-matching resistance gene in the database and the corresponding sequence in the genome). Although genes associated with antibiotic resistance do not have to be mobile sequences, many are known to be transferred between bacteria by horizontal gene transfer (HGT) [539].

Estimation of core and pan-genome

A core genome and pan-genome was estimated based on comparing 84 *Campylobacter* isolates from a range of sources, including avian, environmental water and agricultural hosts. This collection included the 14 isolates from Australian purple swamphen and 70 New Zealand isolates. The core genome and the pan-genome were estimated using GET HOMOLOGUES (ver. 1.0) [540]. The core genome was estimated from the 84 draft genomes by comparing the results from three different methods of orthologue group prediction, Orthomcl [541] [542], COG [543] and BDBH [540], then using only the orthologue clusters identified by all three methods as present in all 84 isolates. The GET HOMOLOGUES default values were used in the BLAST searches with a minimum E-value (maximum expectation value) of 10^{-5} and a minimum of 75% coverage in pairwise alignment. The pan-genome was estimated using GET HOMOLOGUES [540], with Orthomcl for clustering [541] [542], to make a presence/absence matrix of homologue groups, this matrix was used in further analysis.

Core genome

MLST

The allelic profile for the seven MLST loci was established *in silico* using the genome sequence and MLST (ver1.7) for the 14 Australian purple swamphen isolates (B2121-B2134) [521]. The

validity of this approach with *C. jejuni* was checked by comparing the genomes of 11 isolates that previously had their identity (ST) established using the accepted laboratory protocol [265] and finding 100% agreement.

rMLST

The Australian purple swamphen genomes were uploaded to the BIGSdb servers [277], where the rMLST sequences were identified and compared to 69 New Zealand isolates from both agricultural and wild-bird/water sources. Details on the New Zealand isolates are provided in Table 5.3. The rMLST profiles are based on 52 loci, rather than the full 53 loci as none of the *Campylobacter* spp. have the *rpm*D loci. A distance matrix of the rMLST alleles was calculated from pairwise differences, missing alleles were removed from the pairwise comparison. The distance matrix was made into an unrooted phylogenetic network using the Neighbor-Net in SplitsTree [488].

Pan-genome comparison

Using the discrete character parsimony PARS method from the PHYLIP suite [544] on the pan-genome presence absence matrix created in GET HOMOLOGUES [540] a .phylip file was made. The .phylip file was converted into an Additive Dollo Distance (ADD) matrix, which uses a Dollo process that assumes a trait (gene) can arise once but be lost many times [545]. The ADD matrix was then used to construct an unrooted phylogenetic network using the Neighbor-Net in SplitsTree [488].

Population structure

Discriminant analysis of principal components The discriminant analysis of principal components (DAPC) was performed using the adegenet 1.4–2 package in R [546] [547], as described in Jombart et al. [548]. First the pan-genome presence/absence data was transformed into principal components and then k-means clustering was used to define the groups (8). Thirty components were retained in the principal component analysis, accounting for 82.6% of the variability.

Statistical analysis

The statistical analyses were performed using R v2.13 and v3.1.0 (R Development Core Team 2008, R Foundation for Statistical Computing, Vienna, Austria). The genomic data was analysed using the packages adegenet 1.4–2 package [546] [547] and pheatmap [490].

5.3 Results

Characterisation table of Australian purple swamphen genomes

The draft genome characteristics were summarised in Table 5.1, where in general, the draft genomes share many of the same characteristics. *C.* sp. nov. 1 has a slightly larger genome (20.2–20.3

Mbp) compared to the *C. jejuni* (17.9–18.8 Mbp) and *C.* sp. nov. 4 (17.0–17.2 Mbp). *C.*sp.nov.1 (31.3−31.4%) also has a slightly higher percent of hypothetical proteins than *C. jejuni* (27.3−29.8%) or *C.*sp.nov.4 (27.7 − 28.3%). However the basis for calling these isolates new species was not limited to these genome size differences and is related in the following analysis of both the core and pan-genome.

16S rRNA

All the complete 16S rRNA sequences are ∼1500bp long except *Campylobacter curvus* which was much longer (1703bp) possible due to an insertion sequence [549]. Using full length 16S rRNA sequence (Figure 5.1) there is a clear separation between the isolates of *Campylobacter concisus, Campylobacter curvus, Campylobacter fetus* subsp. *fetus, Campylobacter hominis, Campylobacter lari*, *C.*sp.nov.3, and the grouping of *C. jejuni*, *C. coli* and *C.* sp. nov. 1, 2 and 4. Figure 5.2 provides a closer view of the *C. jejuni*, *C. coli* and *C.* sp. nov. 1, 2 and 4 grouping.

Figure 5.2 shows the full length 16S rRNA sequences of 46 *C. jejuni*, 12 *C. coli*, and four putative *Campylobacter* spp., exhibiting a level of phylogenetic similarity between *C. jejuni* and *C. coli*, and at the same time showing diversity within both species.

All the full length 16S rRNA sequences for the *C. coli* were 1501bp long and grouped together, except ST-3302 which grouped closer to *C.* sp. nov 4. The 16S rRNA sequences differed by a 2bp single nucleotide polymorphism (SNP) between ST-3302 and *C.* sp. nov 4, but ST-3302 differed by 5bp SNP to the rest of the *C. coli*. Nine *C. jejuni* isolates, six ST-474 (clonal complex (CC) CC48) and three isolates from ST-48, ST-2341 and ST-3711 (CC48, CC61 and CC257 respectively), shared the same length (1501bp) and sequence as the group of 11 *C. coli*. The ten international reference *C. jejuni* sequences used as comparisons to the New Zealand isolates were spread across the *C. jejuni/C. coli* grouping in Figure 5.2, suggesting they show a similar amount and type of variation as the New Zealand isolates. There is a noticeable tendency in Figure 5.2 for the isolates from the Australian swamphen and from the two New Zealand rails (pūkeko and takahē Table 5.3) to be in the same general area.

	C. sp. nov. 1				C. *jejuni*								*C.* sp. nov. 4	
	B2121	B2122	B2125	B2131	B2124	B2127	B2128	B2129	B2130	B2132	B2133	B2134	B2123	B2126
Size (bp)	2024631	2021406	2025434	2022724	1868014	1877416	1793035	1863972	1854699	1861956	1863331	1838791	1705599	1723026
N50	56630	59161	56651	48764	192560	166668	165799	289799	151678	290307	142310	166720	100976	128371
% G+C content	30.07	30.08	30.07	30.07	30.01	30.01	30.14	30.04	30.01	30.04	30.05	30.02	30.52	30.56
No contigs	128	118	124	138	154	164	107	165	166	160	178	163	102	93
CDS number	2021	2023	2017	2023	1884	1893	1784	1864	1850	1860	1863	1836	1638	1681
hypothetical proteins (%)	31.4	31.3	31.3	31.4	29.7	29.8	27.3	28.5	29.1	28.7	28.8	28.4	28.3	27.7
Prophage	yes	yes	yes	yes	yes	yes	yes	ND	ND	yes	yes	yes	yes	yes
Genomic island	yes	ND	yes	ND	yes	yes	yes	yes	yes	yes	yes	yes	yes	ND
Antibiotic resistance genes	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
Ribosomal RNA operons	3	5	4	3	6	3	4	4	3	4	4	3	3	3
16S rRNA (bp)	partial	partial	partial	partial	1501	1501	1501	1501	1501	1501	1501	1501	1501	1501
CRISPR	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
tRNA	40	40	41	40	36	36	36	36	36	36	37	36	41	43
tmRNA	1	1	1	1	1	1	1	1	1	1	1	1	1	1
pathogen probability (%)	81.0	81.3	81.0	81.0	82.2	82.5	82.8	82.9	82.6	82.9	82.9	82.6	83.8	80.7

Table 5.1: **General features of the Australian purple swamphen (*Porphyrio porphyrio melantotus*) genomes.** CRISPR stands for Clustered regularly interspaced short palindromic repeats. ND refers to none detected

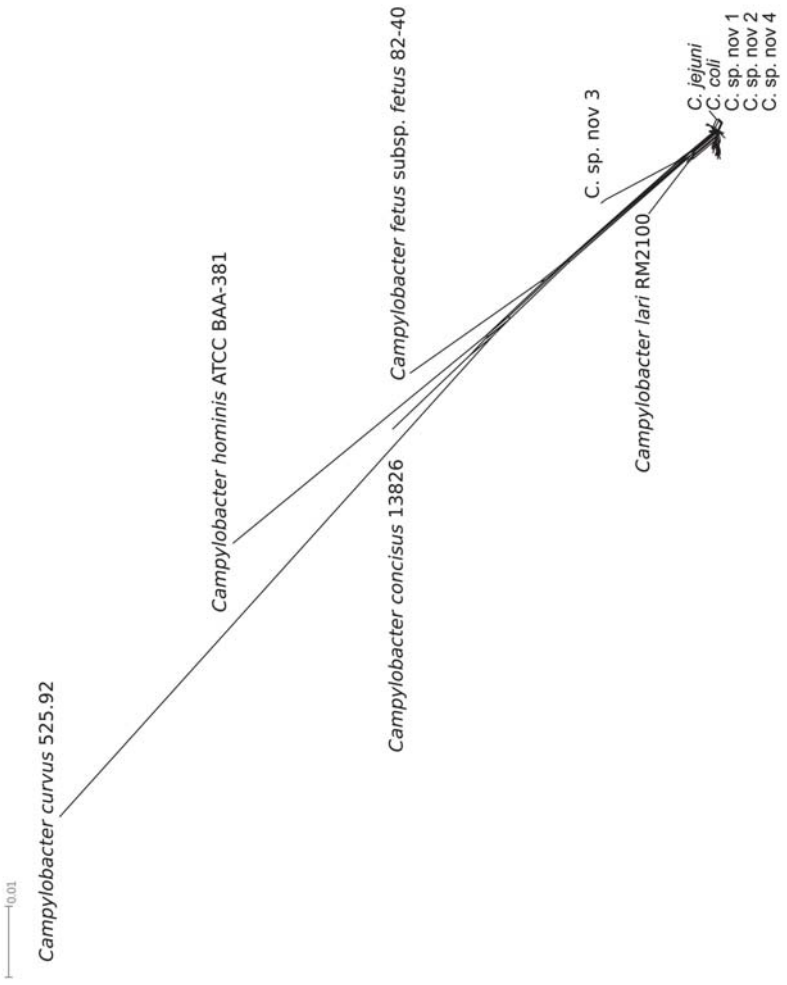


Figure 5.1: **A Neighbor-Net of full length 16S rRNA for 83 *Campylobacter* spp.isolates.** *Campylobacter curvus*,*Campylobacter hominis*, *Campylobacter concisus*, *Campylobacter fetus* subsp. *fetus*, *Campylobacter lari* and *C.* sp. nov. 3 are dispersed and phylogenetically distinct to a tighter grouping of the *C. jejuni, C. coli*, *C.* sp. nov. 1, 2 and 4 isolates. A closer view of this tighter grouping is shown in Figure 5.2. The reference sequences are listed in Table 5.4

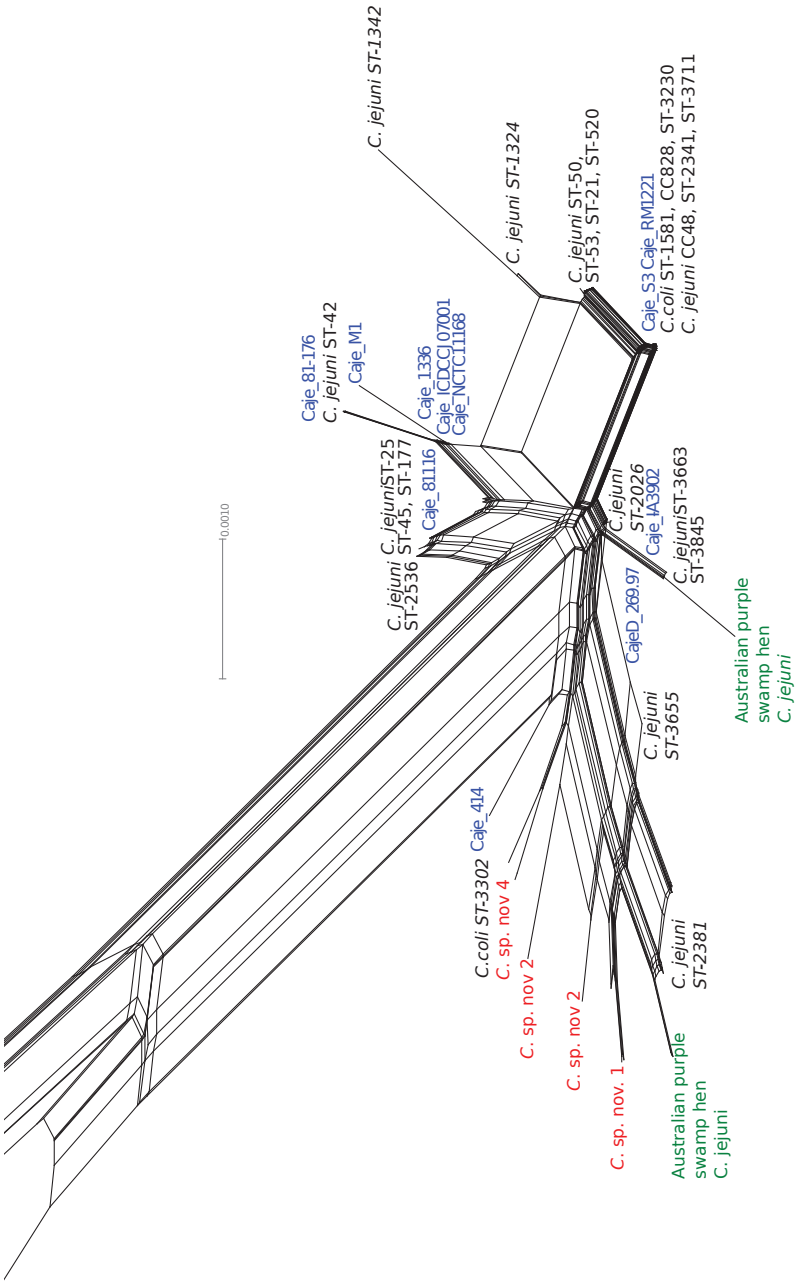


Figure 5.2: **An part-view of the Neighbor-Net of full length 16S rRNA for 83 *Campylobacter* spp..** This is a closer view of the tight grouping of the *C. jejuni/C. coli* complex including *C.* sp. nov. 1, 2 and 4 from Figure 5.1. The red letters are for *C.* sp. nov. 1, 2 and 4. The green letters highlight the Australian purple swamphen *Campylobacter jejuni*. The blue letters show the distribution of the non-New Zealand reference sequences listed in Table 5.4.

The inability of the 16S rRNA sequences alone to clearly distinguish between some *Campylobacter* spp. is also evident in Figure 5.3 where the ten Australian purple swamphen show $\geq 98.65\%$ pairwise similarity to the same four *Campylobacter* species (*Campylobacter jejuni* subsp. *doylei, Campylobacter jejuni* subsp. *jejuni, Campylobacter subantarcticus, Campylobacter lari subsp. concheus, Campylobacter insulaenigrae*, but not *Campylobacter lari* subsp. *lari* or *C. coli*. This application of pairwise similarity [550] shows a similar pattern of relationship for both the *C. jejuni* and the *C.* sp. nov. 4 isolates. However, there is a noticeable splitting of the eight Australian purple swamphen *C. jejuni* isolates into the same two groups, by both the pairwise similarity analysis (Figure 5.3) and in the Neighbor-Net analysis (Figure 5.2). This splitting of the eight Australian purple swamphen *C. jejuni* isolates was the only example in Figure 5.2 of isolates not clustering by ST.

***in silico* DNA-DNA hybridisation**

The *in silico* DNA-DNA hybridisation comparison used three different measures (ANIb, ANIm, Tetra) to compare the genomes of 14 Australian purple swamphen isolates to some international reference sequences and produced results consistent with the rMLST from the core genome. Eight Australian purple swamphen isolates (B2124, B2127, B2128, B2129, B2130, B2132, B2133, B2134) showed similarity to both *C. jejuni* and *C. jejuni* subsp.*doylei*, above the cut-off thresholds for all three measures (Figures 5.4 5.6 5.5). The remaining six Australian purple swamphen isolates did not show similarity, above the cut-off thresholds for the three measures, with any of the other *Campylobacter* spp. used (Figure 5.4 5.5 5.6).

In comparison, when using the dataset of New Zealand isolates, four of the six Australian purple swamphen isolates (B2121, B2122, B2125, B2131) that did not group with any of the international reference genomes, showed a similarity above the cut-off thresholds (Figures 5.8 5.9 5.7) with a previously identified putative species *C.* sp. nov. 1 [16]. The ANIb measure (Figure 5.8) of the four isolates pairwise similarity (94%) is below the $95 - 96\%$ cut off with B1821b, a *C.* sp. nov. 1 takahē isolate. However using the other two measures, ANIm and Tetra, the cut-off values for similarity with B1821b are meet (Figure 5.7 5.9). The same eight of isolates (B2124, B2127, B2128, B2129, B2130, B2132, B2133, B2134) continued to show similarity above the cut-off thresholds with *C. jejuni* by all three measurements (Figures 5.8 5.9 5.7). Two isolates (B2123 and B2126) did not show similarity with any of the *Campylobacter* spp. to which they were compared, and may be another putative new *Campylobacter* spp.. For simplicity these isolates will be referred to as *Campylobacter* sp. nov. 4 (*C.* sp. nov. 4).

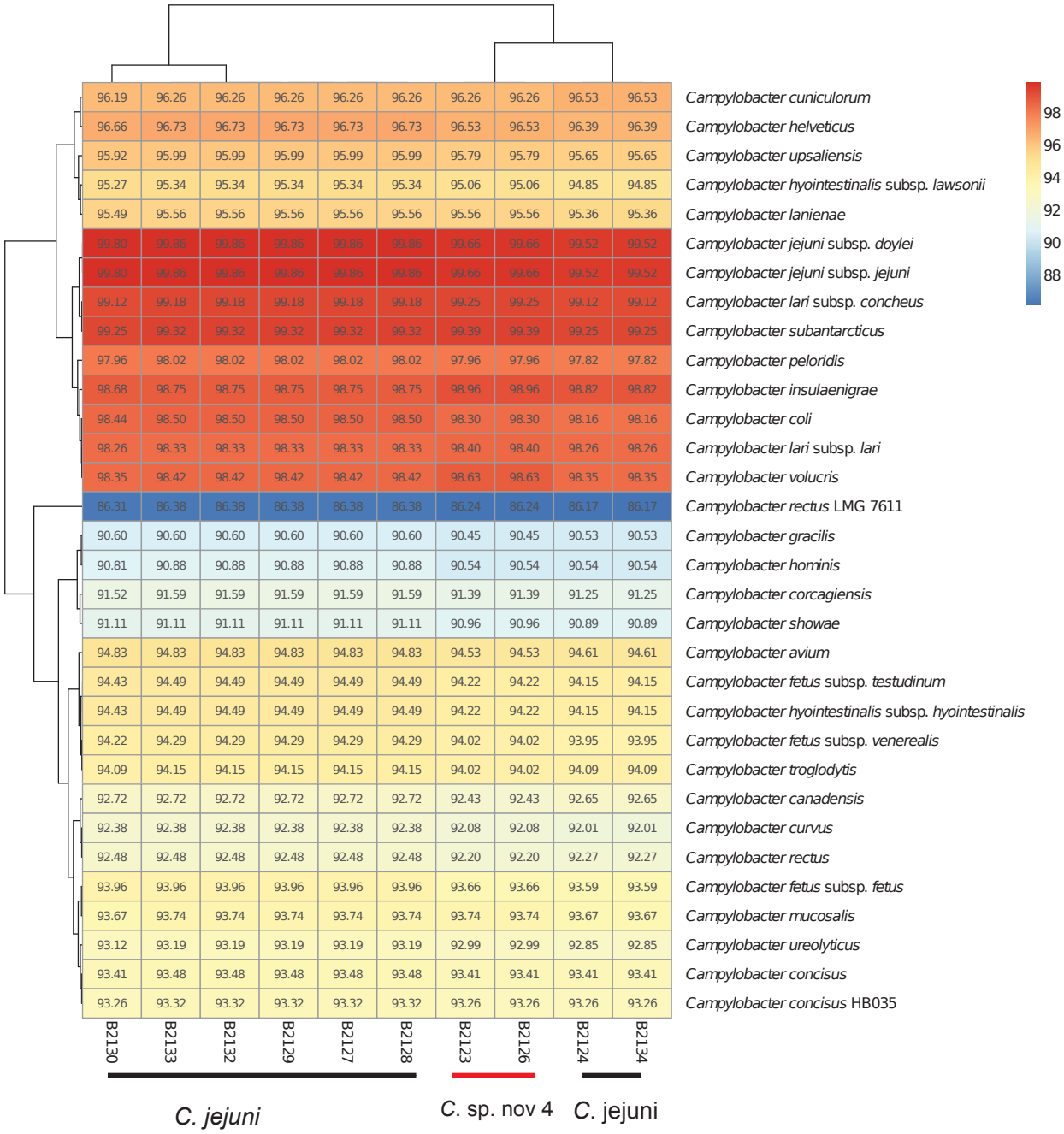


Figure 5.3: **Comparison of ten Australian purple swamphen *Campylobacter* spp. 16S rRNA by pairwise sequence similarity using EzTaxon[524].** All the isolates show $\geq 98.65\%$ similarity to the same group; the two *C. jejuni* subtypes, *C. lari* subsp. *concheus*, *Campylobacter insulaenigrae* and *Campylobacter subantarcticus*.

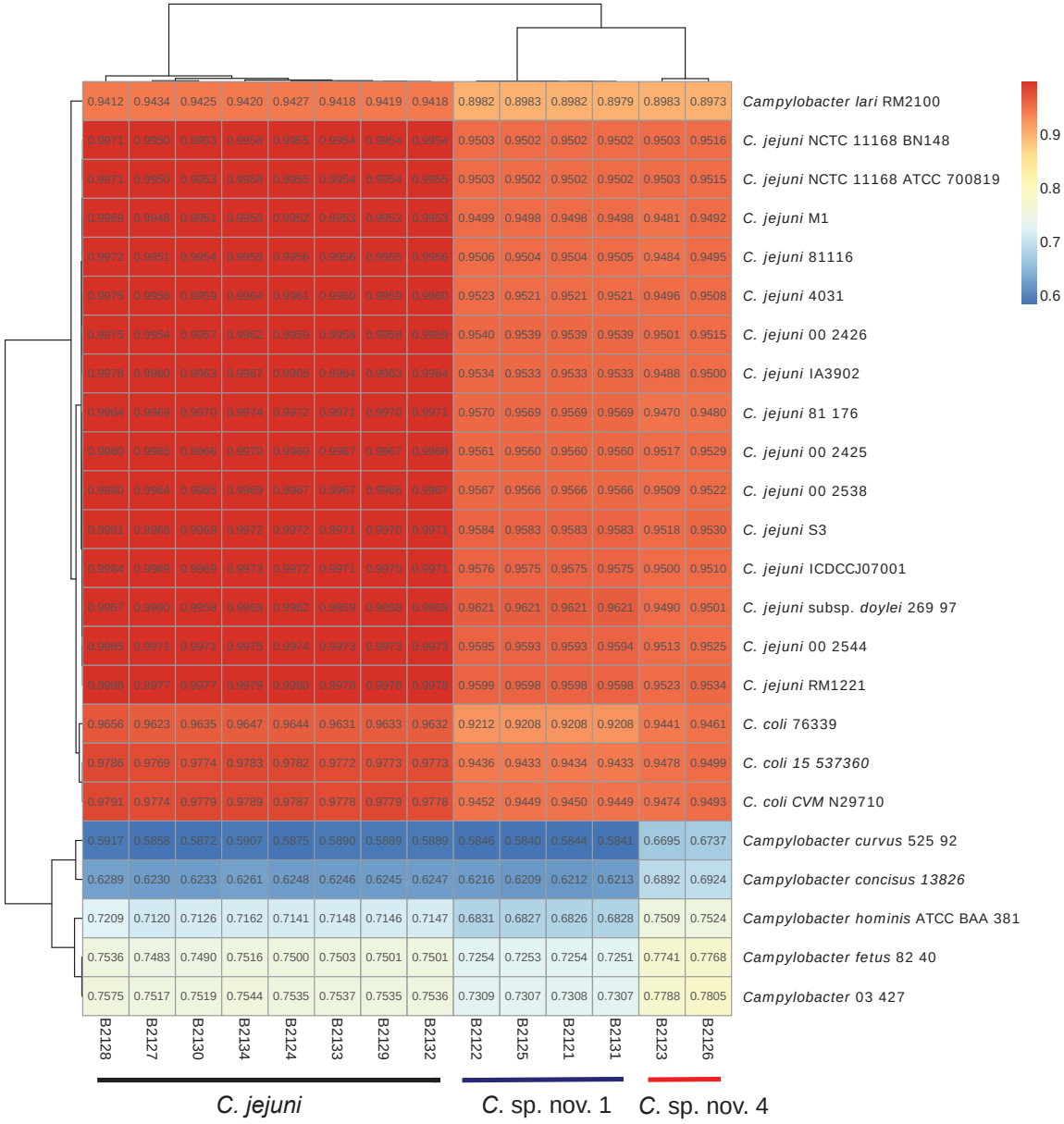


Figure 5.4: **The heatmap of *in silico* DNA-DNA hybridisation of *Campylobacter* spp. genomes compared by tetranucleotide frequency correlation coefficients (Tetra value)**. The 14 draft genomes (B2121-B2134) being compared to known genomes are from the Australian purple swamphen *Campylobacter* spp.. Using a cut off of ≥ 0.99 the genomes divide into *C. jejuni*, *C.* sp. nov. 1 and 4. The identity of the sequences used for the comparison are in Table 5.5.

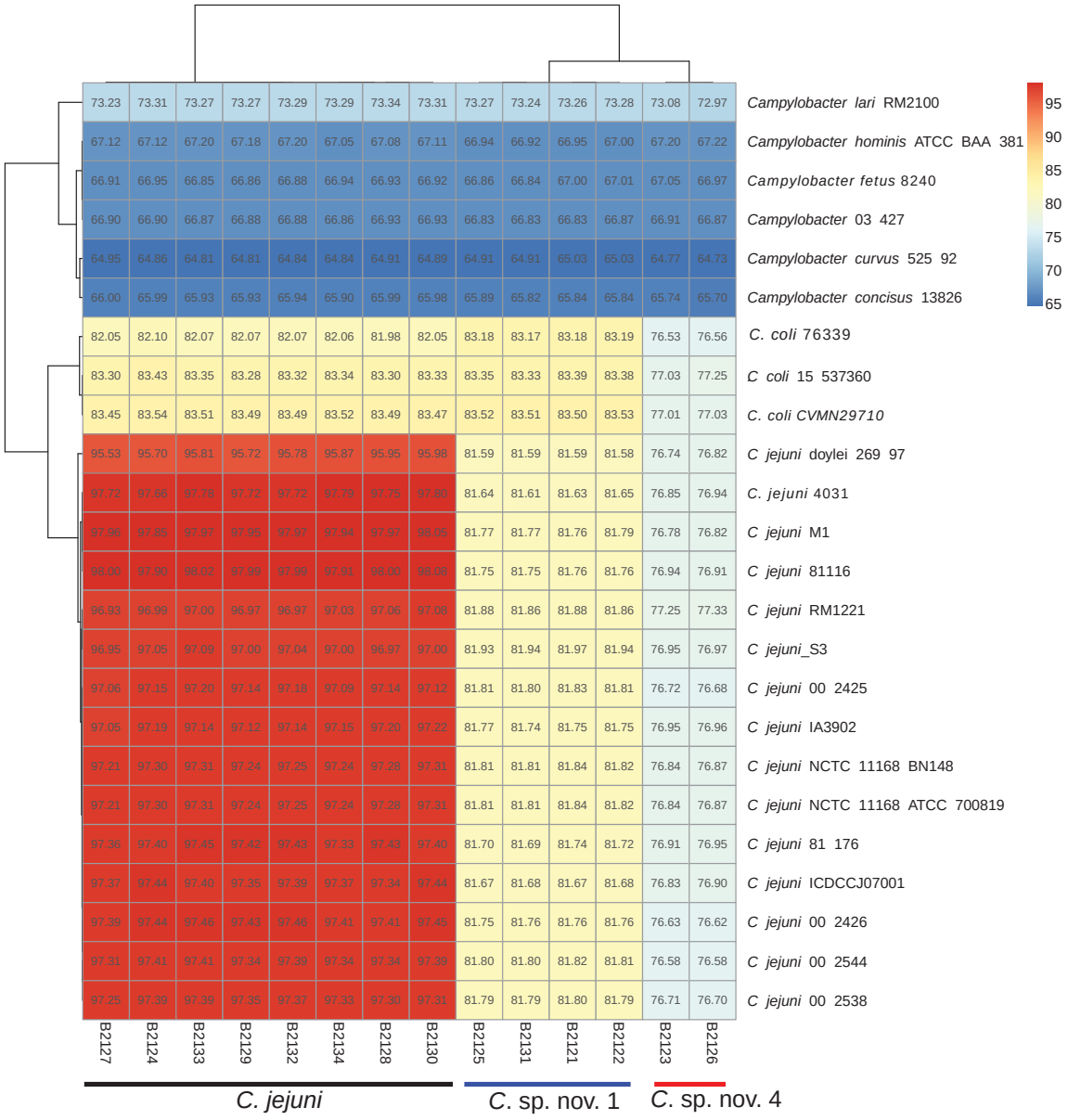


Figure 5.5: **The heatmap of *in silico* DNA-DNA hybridisation of *Campylobacter* spp. genomes compared by average nucleotide identity using BLAST (ANIb)**. The 14 draft genomes (B2121-B2134) being compared to known genomes are from the Australian purple swamphen *Campylobacter* spp.. Using a cut off value of $95 - 96\%$ the genomes divide into *C. jejuni*, *C.* sp. nov. 1 and 4. The identity of the sequences used for the comparison are in Table 5.5.

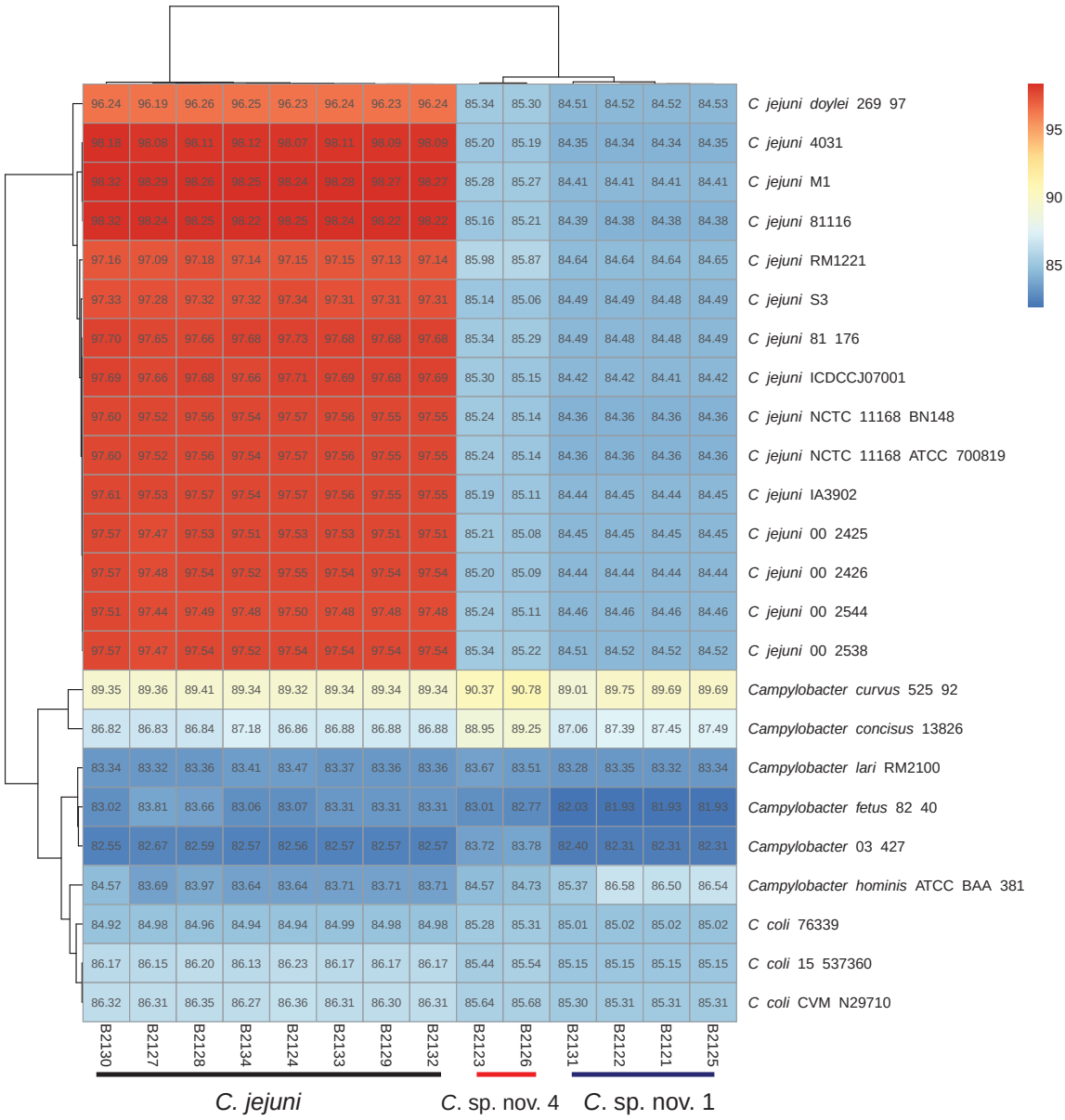


Figure 5.6: **The heatmap of *in silico* DNA-DNA hybridisation of 14 Australian purple swamphen *Campylobacter* spp. genomes measured by average nucleotide identity using MUMmer (ANIm)**. The 14 draft genomes (B2121-B2134) being compared to known genomes are from the Australian purple swamphen *Campylobacter* spp.. Using a cut-off value of $95-96\%$ the genomes divide into *C. jejuni*, *C.* sp. nov. 1 and 4. The identity of the sequences used for the comparison are in Table 5.5.

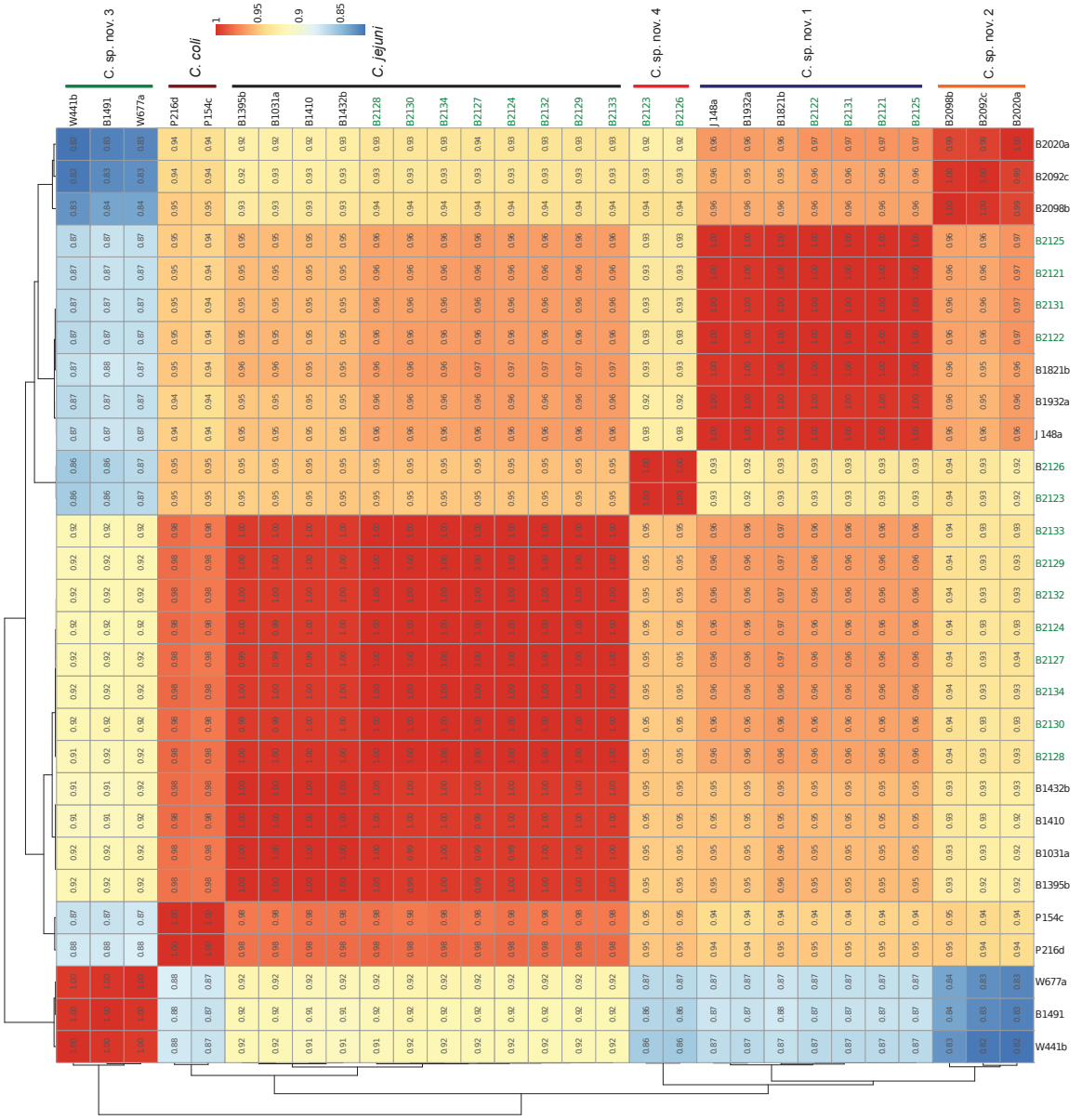


Figure 5.7: **The heatmap of *in silico* DNA-DNA hybridisation of 29 *Campylobacter* spp. genomes measured by tetranucleotide frequency correlation coefficients (Tetra value)**. Using a cut off of ≥ 0.99 the genomes divide into six *Campylobacter* spp. The isolates are shown in Table 5.3 and the Australian purple swamphen isolates are highlighted in green.

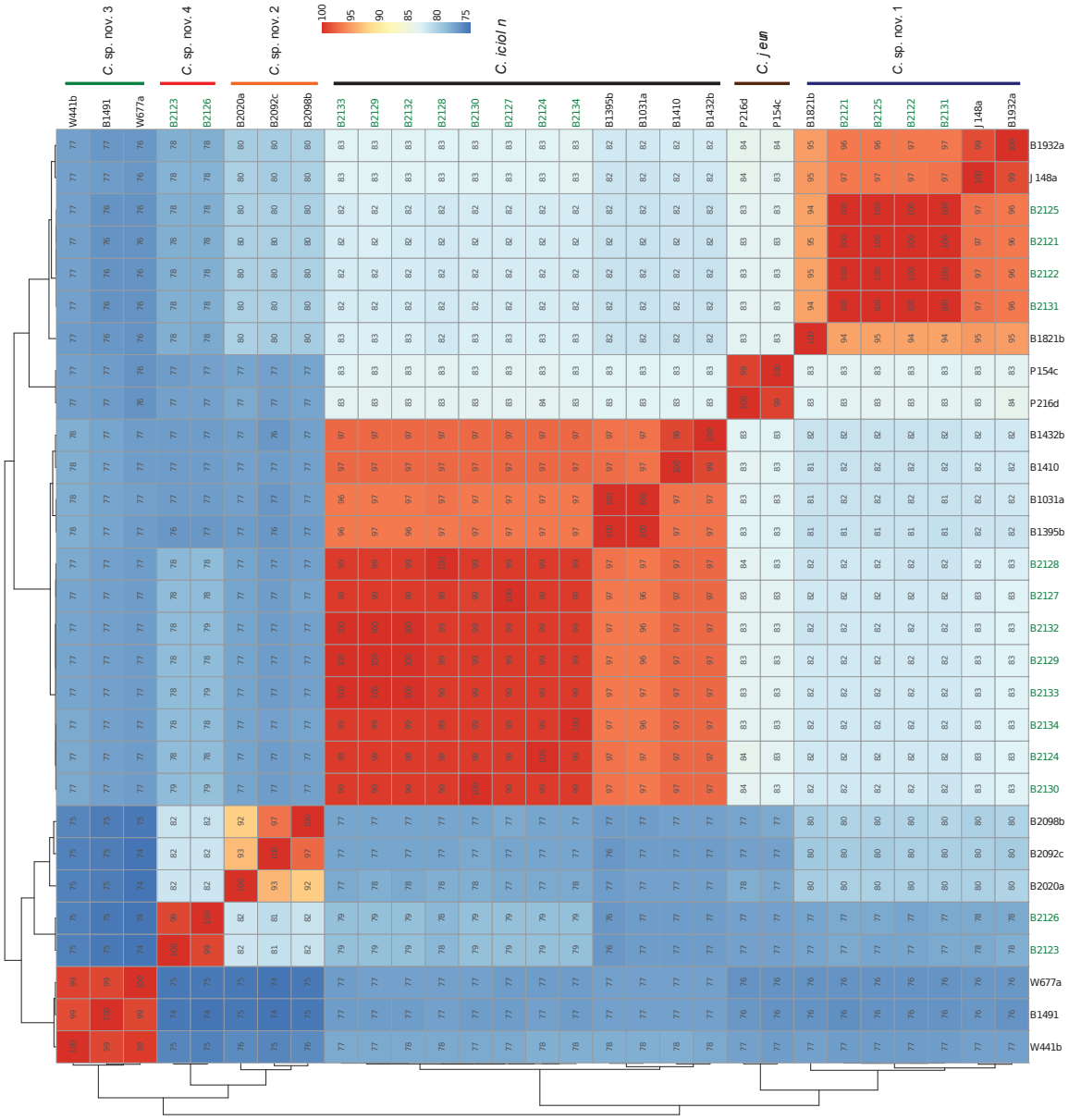


Figure 5.8: **The heatmap of *in silico* DNA-DNA hybridisation of 29 *Campylobacter* spp. genomes measured by average nucleotide identity using BLAST (ANIb).** Using a cut off value of $95 - 96\%$ the genomes divide into six *Campylobacter* spp. The isolates are shown in Table 5.3 and the Australian purple swamphen isolates are highlighted in green.

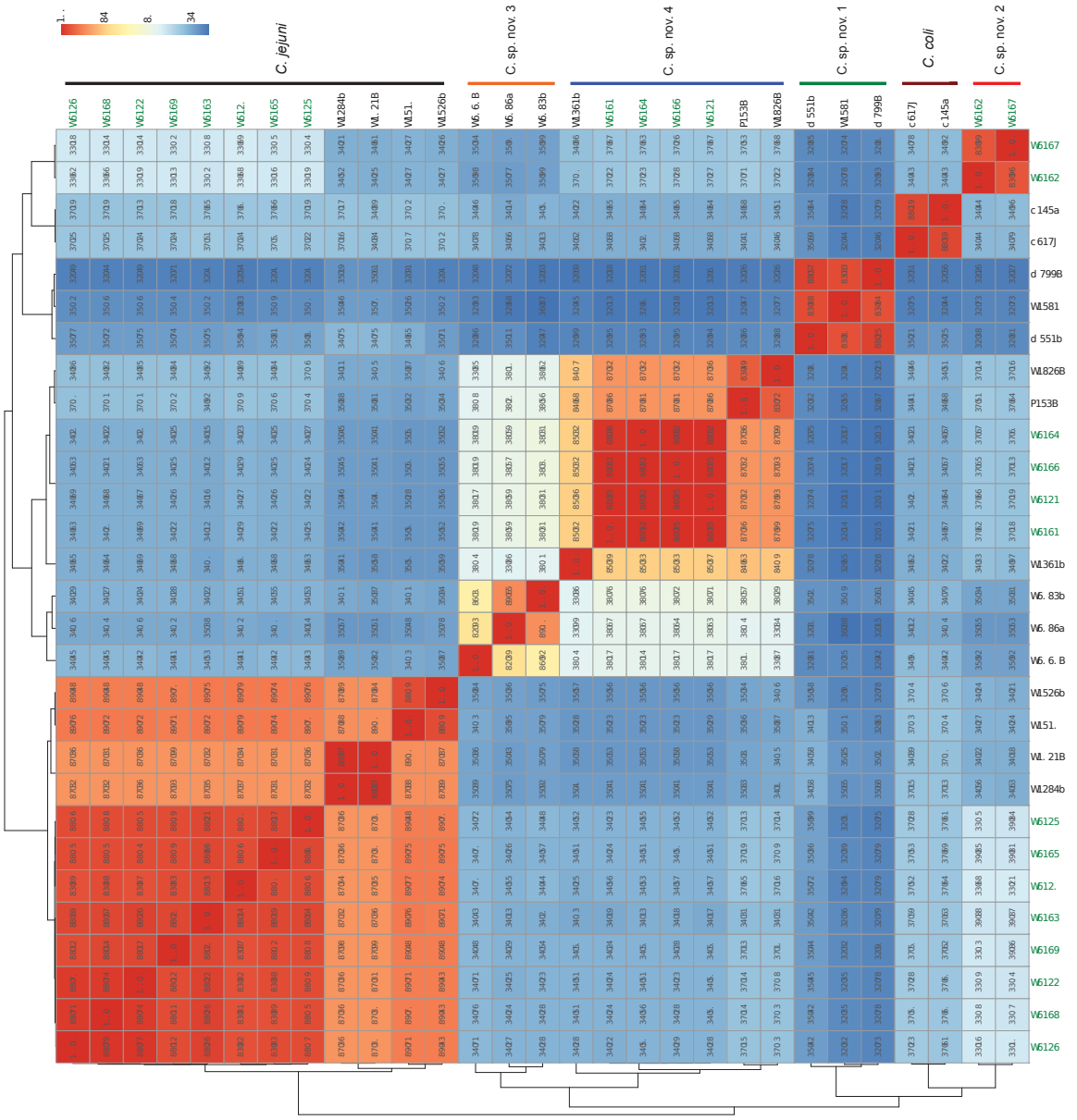


Figure 5.9: **The heatmap of *in silico* DNA-DNA hybridisation of 29 *Campylobacter* spp. genomes measured by average nucleotide identity using MUMmer (ANIm)**. Using a cut off value of $95-96\%$ the genomes divide into six *Campylobacter* spp. The isolates are shown in Table 5.3 and the Australian purple swamphen isolates are highlighted in green.

PathogenFinder

The Australian purple swamphen isolates when analysed in PathogenFinder [530] showed a probability range 80.7% – 83.8% (Table 5.1) of being a human pathogen. As a comparison seven ST-474 isolates (569a, H22082, H704, m73020, P110b, P179a, P694a) were also analysed and all showed an 83% probability of being a human pathogen. ST-474 was a significant contributor to the recent *C. jejuni* epidemic in New Zealand.

Mobilome

All the Australian purple swamphen draft genomes showed features of the mobilome which is consistent with the pan-genome having 43.4% single occurrence or unique homologues. Table 5.1 reveals a tendency towards having sequences indicative of a genomic island or a prophage. There are predicted genomic islands in 77% (11/14) of the draft Australian purple swamphen genomes. Phage-associated sequences were identified in 64% (9/14) of the draft genomes (Table 5.1). No recognised antibiotic resistance genes were identified for 12 different antibiotics in any of the genomes, however this does not rule out a previously unknown antibiotic resistance gene being present.

Estimation of core and pan-genome

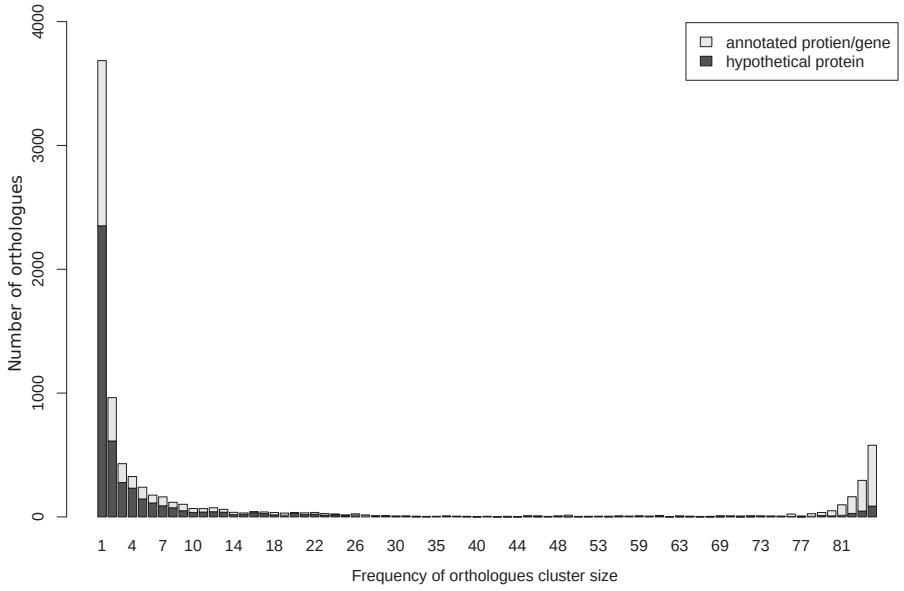


Figure 5.10: **The frequency of occurrence of identified orthologues (genes and hypothetical proteins) shared or unique to 84 *Campylobacter* isolates.** The clustering of orthologues was calculated in Orthomcl [541] [542]

The pan-genome was estimated by GET HOMOLOGUES using Orthomcl. It showed a total of 8495 homologues and orthologues within this group of up to 6 *Campylobacter* spp.. The bar chart in Figure 5.10 shows the number and occurrence of shared and unique orthologues for the pan-

genome of the 84 *Campylobacter* isolates identified in Table 5.3. As seen in other pan-genomes the distribution is bimodal with an increased number of clusters at the core genome end (found in all 84) and the low frequency end (found in only 1 or a few isolates) [551] [552]. 43.4% (3685/8495) of the total number of orthogues were single occurrence homologues, for which there are several possible explanations, one suggestion would be a large "mobilome" i.e. mobile elements of the genome such as transposons, plasmids, bacteriophage, and self-splicing molecular parasites [553][p3]. In the characterisation table (Table 5.1) for the Australian purple swamphen genomes, many of the isolates contained mobile elements, and this may be a common feature in the other genomes. 45.0% (3822/8495) of the orthologue clusters were annotated as hypothetical proteins. The hypothetical proteins were not evenly distributed: 15.7% (87/554) of the core genome was comprised of hypothetical proteins compared to 63.8% (2351/3685) of the single occurring homologues.

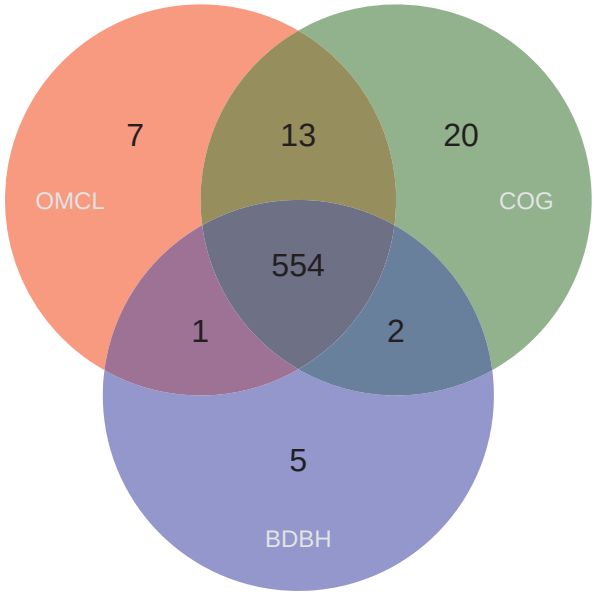


Figure 5.11: **The core genome of 84 isolates across 6 *Campylobacter* spp. estimated by three different methods within GET HOMOLOGUES.** The three methods used to estimate the size of core genome were Orthomcl, COG and BDBH with the Venn diagram showing the differences in estimated core size and shared content between the three results.

This dataset of 84 isolates from 6 putative *Campylobacter* spp. have an estimated core genome size of 554 (Figure 5.11) as defined by orthologues present in all isolates and identified in all three methods (Orthomcl, COG and BDBH) [541] [542] [543] [540]. There were little differences in the estimated core genome size and composition estimated by each algorithm: Orthomcl (575), COG (589) and BDBH (562). These differences could be attributed to differences in the way the methods handled features like in-paralogues [202].

Table 5.2: **The multi-locus sequence type (MLST) allelic profile of 14 Australian purple swamphen (*Porphyrio porphyrio melantotus*) isolates.** The MLST consists of seven loci for each isolate (B2121-B2134) and were derived *in silico* from the genomic data using MLST (ver1.7) [521]. The allelic profiles currently do not have a designated ST, and new-1 and new-2 represent two unique alleles for a given loci.

isolates	**species**	***asp*A**	***gln*A**	***glt*A**	***gly*A**	***pgm***	***tkt***	***unc*A**
B2121	*C.* sp. nov. 1	new-1	new-1	new-1	343	new-1	new-1	new-1
B2122	*C.* sp. nov. 1	new-1	new-1	new-1	343	new-1	new-1	new-1
B2123	*C.* sp. nov. 4	new-2	new-2	new-2	new-2	new-2	new-2	new-2
B2124	*C. jejuni*	1	6	5	92	261	293	3
B2125	*C.* sp. nov. 1	new-1	new-1	new-1	343	new-1	new-1	new-1
B2126	*C.* sp. nov. 4	new-2	new-2	new-2	new-2	new-2	new-2	new-2
B2127	*C. jejuni*	1	6	5	92	261	7	3
B2128	*C. jejuni*	1	6	5	92	261	7	3
B2129	*C. jejuni*	1	6	5	92	261	7	3
B2130	*C. jejuni*	1	6	5	92	261	7	3
B2131	*C.* sp. nov. 1	new-1	new-1	new-1	343	new-1	new-1	new-1
B2132	*C. jejuni*	1	6	5	92	261	7	3
B2133	*C. jejuni*	1	6	5	92	261	7	3
B2134	*C. jejuni*	1	6	5	92	261	7	3

Core genome features

MLST

Seven of the Australian purple swamphen *C. jejuni* isolates were indistinguishable sharing the same allelic profile, and the eighth was a single locus variant to the other isolates (Table 5.2). All the *C. jejuni* isolates had seven existing alleles in the PubMLST database for *C. jejuni* and *C. coli* [266], however the isolates that are believed to be new *Campylobacter* spp. resulted in 6 new MLST alleles for the putative *C.* sp. nov. 1 and seven new MLST alleles for the new putative *C.* sp. nov. 4. The *gly*A allele 343, found in all four *C.* nov. sp. 1 isolates (Table 5.2), has been reported once in PubMLST [2] as a *C. jejuni* chicken isolate (ST-3479) found in the UK in 2013.

rMLST

The rMLST profiles (Figure 5.12) provided a clear division between *Campylobacter* spp.. The *C. jejuni* isolates clustered by ST and CC. Within the *C. coli* split, the Caco-N211 (ST-3302) isolate

[2] 18th Feb 2015 http://pubmlst.org/campylobacter/

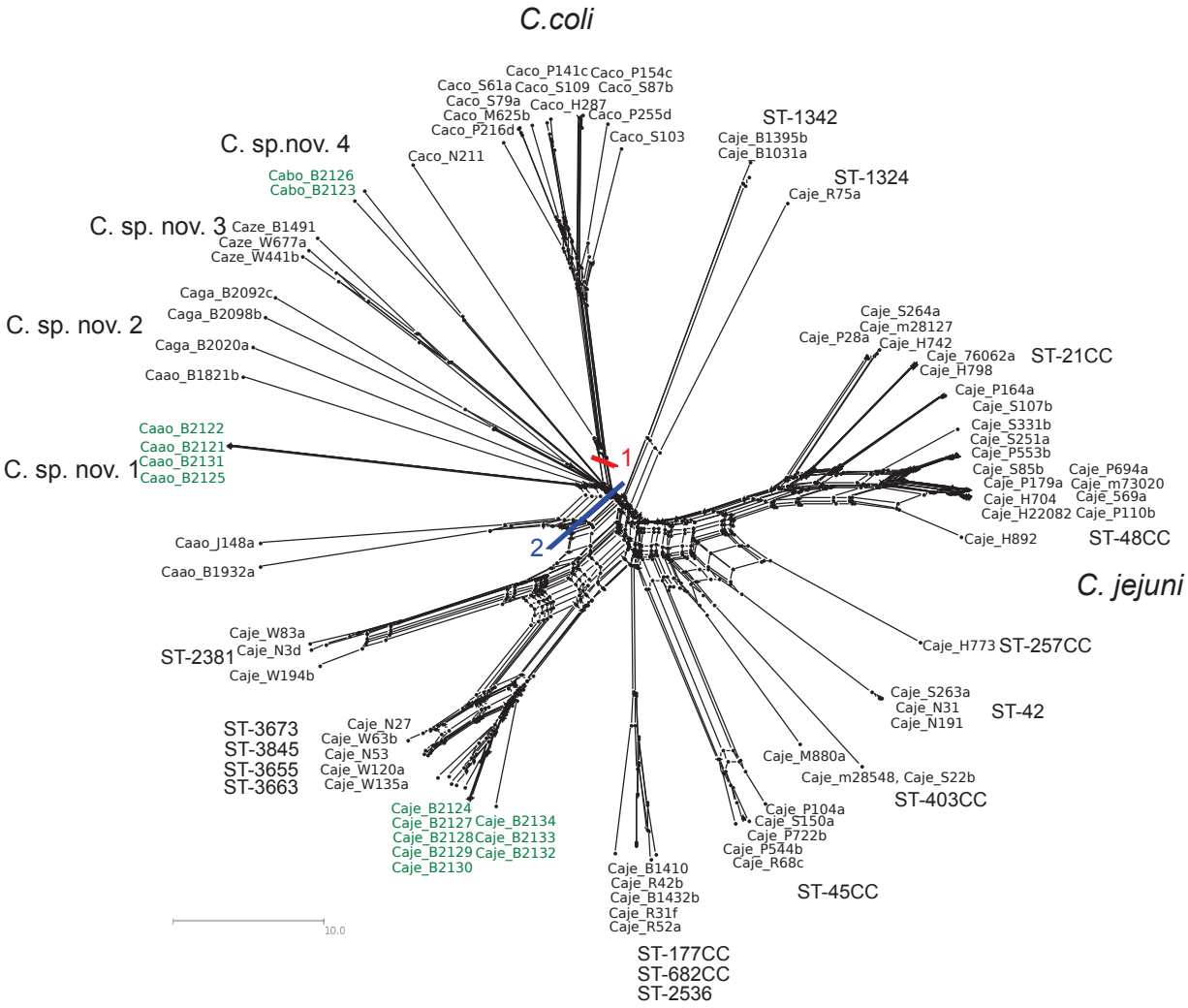


Figure 5.12: **Comparison of 83 *Campylobacter* spp. based on rMLST**. The isolates are labelled by *Campylobacter* spp. with a prefix; Caje for *C. jejuni*, Caco for *C. coli*, Caao for *C.* sp. nov. 1, Caga for *C.* sp. nov. 2, Caze for *C.* sp. nov. 3, Cabo for *C.* sp. nov. 4. Line 1 (red) shows a division for *C. coli* and Line 2 (blue) shows a division for *C. jejuni*. The isolates highlighted in green were found in Australian purple swamphen.

was separated from the other isolates which were predominantly members of CC828, a behaviour consistent with the 16S rRNA analysis 5.2. The seven isolates representing *C.* sp. nov. 1 cluster together, but within the grouping they formed three groups, one associated with Australian purple swamphen isolates and the other two with New Zealand isolates. This suggests *C.* sp. nov. 1 may be a large diverse group with multiple clades similar to *C. coli* and *C. jejuni*.

Pan-genome

Pan-genome comparison

Figure 5.13 shows that the pattern of presence or absence of orthologues in the pan-genome can clearly differentiate the known species *C. jejuni* and *C.coli*, as well as the new proposed species: *C.* nov. 1, *C.* nov. 2, *C.* nov. 3, *C.* nov. 4. The level of differentiation also extends to the *C. jejuni* isolates clustering at the clonal complex level. Another feature is the overall similarity of the pan-genome (Figure 5.13) to the core genome (Figure 5.12). In particular the clustering relationship between the Australian purple swamphen isolates and the takahē/pūkeko associated isolates (ST-2381, ST-3663, ST-3655) was basically the same in both comparisons, core and pan-genome.

Population structure

Discriminant analysis of principal components (DAPC) In the DAPC analysis, the population structure of the *Campylobacter* spp. isolates, as expressed by the pan-genome presence/absence matrix, clustered into 8 groups with a probability of membership $> 99\%$ by each member. A scatter plot based on the first and second principal components shows the distribution of the clusters in Figure 5.14 while Table 5.3 shows the identity of the isolates in each cluster. The isolates generally clustered by *Campylobacter* spp. and ST/CC but not to geographical location. *C.* sp. nov. 1 isolates are all in cluster 1, all the *C .coli* in cluster 2 (mainly clonal complex CC828), and the *C.* sp. nov. 3 are all in cluster 6. Cluster 7 includes both *C.* sp. nov. 2 and 4. *C. jejuni* is divided between 4 clusters, all relatively close to each other but cluster 5 is slightly separate to clusters 3, 4 and 8 in Figure 5.14. The division of *C. jejuni* follows a ST/CC pattern: cluster 3 contains CC45, CC42, CC177, CC682; cluster 4 contains CC48; cluster 8 contains CC21, CC61, CC257, and CC403; cluster 5 contains the Australian purple swamphen *C. jejuni* isolates and some of the ST associated with pūkeko and takahē. Although one ST (ST-3673) in cluster 5 is rare, having only been found once in a chicken and in New Zealand environmental waters, it does share MLST alleles with other New Zealand isolates in the cluster. It is of note that *C.* sp. nov. 2 and *C.* sp. nov. 4 were not placed in separate clusters. Also the ST-2381 isolates which were associated with the pūkeko and the takahē were not in cluster 5 but were in cluster 3. Cluster 3 has a mixture of agriculture-associated and wild-bird associated ST.

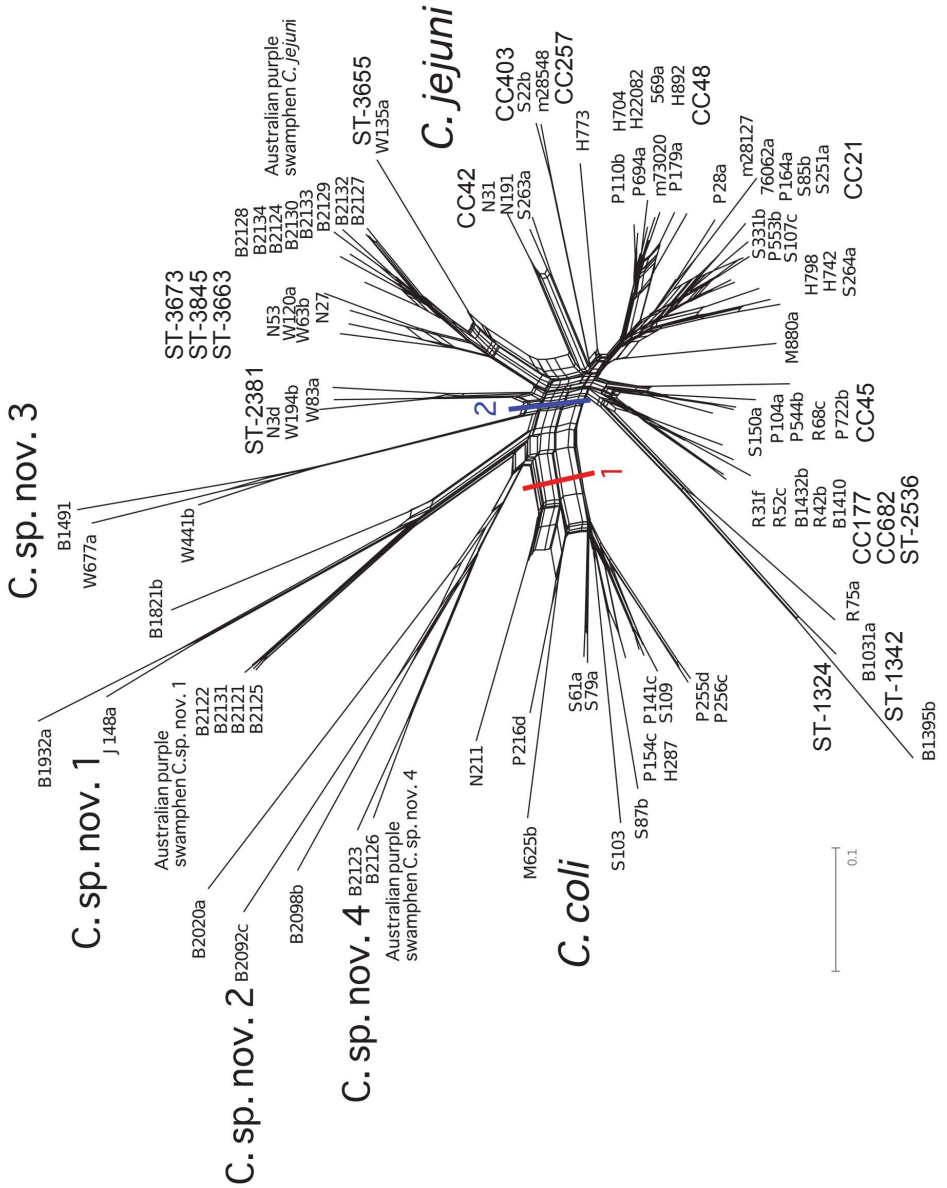


Figure 5.13: **The Neighbor-Net of 84 *Campylobacter* genomes based on the presence or absence of orthologues in the pan-genome**. The red line 1 shows the split dividing off *C.coli* and the blue line 2 shows the split dividing off *C. jejuni* isolates. The branches containing the proposed new species (*C.* nov. 1, *C.* nov. 2, *C.* nov. 3, *C.* nov. 4) and the Australian purple swamphen isolates are identified. The presence/absence matrix was made into an Additive Dollo Distance (ADD) matrix [545] and visualised using SplitsTree [487]

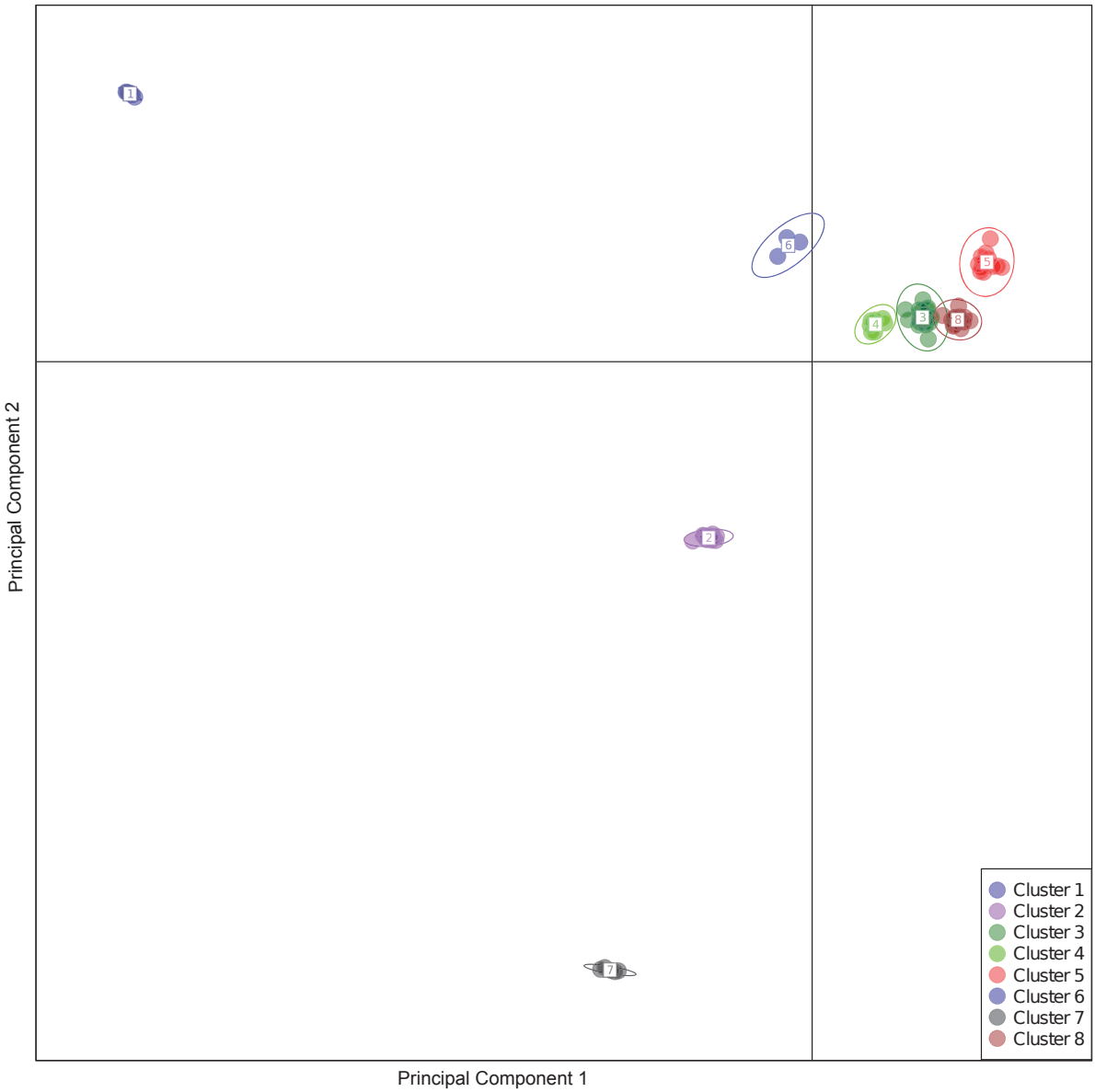


Figure 5.14: **Discriminant analysis of principal components of 84 *Campylobacter* spp. isolates** in a scatter plot based on the first and second principal components. The identify of the isolates in each cluster is shown in Table 5.3. Cluster 1 consists of C. sp.nov. 1; Cluster 2 consists of *C. coli*; Clusters 3,4,5,8 consists of *C. jejuni*; Cluster 6 consists of *C.* sp. nov. 3; and Cluster 7 consists of *C.* sp. nov. 2 and 4.

Table 5.3: **The Characterisation and DAPC clustering of the 84 *Campylobacter* isolates used in the genomic analysis.** The identity of the isolates by species, geographical region and country of origin are reported. Sample is the source from which the isolate was obtained. The Host describes the host species which the ST has been associated, if known, and livestock refers to cattle sheep and chickens. Australia (AU) and New Zealand (NZ) are abbreviated. DAPC cluster refers to the cluster the isolates were grouped in by the discriminant analysis of principal components in Figure 5.14. The multilocus sequence type (ST) and the related clonal complex (CC), if known are also supplied.

Isolate	Country	Region	Source	Host	Species	DAPC cluster	ST	Clonal complex
B1821b	NZ	Wellington	takahē	takahē	*C.* sp. nov. 1	1	-	-
B1932a	NZ	Otago	takahē	takahē	*C.* sp. nov. 1	1	-	-
B2121	AU	Melbourne	swamphen	swamphen	*C.* sp. nov. 1	1	-	-
B2122	AU	Melbourne	swamphen	swamphen	*C.* sp. nov. 1	1	-	-
B2125	AU	Melbourne	swamphen	swamphen	*C.* sp. nov. 1	1	-	-
B2131	AU	Philip Is.	swamphen	swamphen	*C.* sp. nov. 1	1	-	-
J148a	NZ	Manawatu	pūkeko	pūkeko	*C.* sp. nov. 1	1	-	-
H287	NZ	Manawatu	human		*C. coli*	2	1581	-
M625b	NZ	Manawatu	pig		*C. coli*	2	1016	CC828
N211	NZ	Manawatu	water		*C. coli*	2	3302	-
P141c	NZ	Manawatu	chicken	chicken	*C. coli*	2	1581	-
P154c	NZ	Manawatu	chicken	chicken	*C. coli*	2	1581	-
P216d	NZ	Manawatu	chicken	chicken	*C. coli*	2	3230	CC828
P255d	NZ	Manawatu	chicken	chicken	*C. coli*	2	2397	CC828
P256c	NZ	Manawatu	chicken	chicken	*C. coli*	2	2397	CC828

Continued on next page

DAPC —– *continued from previous page*

Isolate	Country	Region	Sample	Host	Species	DAPC cluster	ST	Clonal complex
S103	NZ	Manawatu	sheep	sheep	*C. coli*	2	3299	CC828
S109	NZ	Manawatu	sheep	sheep	*C. coli*	2	3232	CC828
S61a	NZ	Manawatu	cattle	cattle	*C. coli*	2	3072	CC828
S79a	NZ	Manawatu	cattle	cattle	*C. coli*	2	3072	CC828
S87b	NZ	Manawatu	sheep	sheep	*C. coli*	2	3232	CC828
B1031a	NZ	Manawatu	starling	starling	*C. jejuni*	3	1342	-
B1395b	NZ	Manawatu	starling	starling	*C. jejuni*	3	1342	-
R75a	NZ	Manawatu	wild-bird	wild-bird	*C. jejuni*	3	1324	-
B1410	NZ	Manawatu	starling	starling	*C. jejuni*	3	177	CC177
B1432b	NZ	Manawatu	starling	starling	*C. jejuni*	3	681	CC682
N191	NZ	Waikato	cattle	livestock	*C. jejuni*	3	42	CC42
N31	NZ	Waikato	cattle	livestock	*C. jejuni*	3	42	CC42
N3d	NZ	Waikato	water	pūkeko, takahē	*C. jejuni*	3	2381	-
P104a	NZ	Manawatu	chicken	chicken	*C. jejuni*	3	45	CC45
P544b	NZ	Manawatu	chicken	chicken	*C. jejuni*	3	45	CC45
P722b	NZ	Manawatu	chicken	chicken	*C. jejuni*	3	25	CC45
R31f	NZ	Manawatu	wild-bird	wild-bird	*C. jejuni*	3	2536	-
R42b	NZ	Manawatu	wild-bird	wild-bird	*C. jejuni*	3	2539	CC177
R52c	NZ	Manawatu	wild-bird	wild-bird	*C. jejuni*	3	2536	-
R68c	NZ	Manawatu	wild-bird	wild-bird	*C. jejuni*	3	45	CC45
S150a	NZ	Manawatu	cattle	livestock	*C. jejuni*	3	45	CC45

Continued on next page

DAPC —– *continued from previous page*

Isolate	Country	Region	Sample	Host	Species	DAPC cluster	ST	Clonal complex
S263a	NZ	Manawatu	cattle	livestock	*C. jejuni*	3	42	CC42
W194b	NZ	Manawatu	water	pūkeko, takahē	*C. jejuni*	3	2381	-
W83a	NZ	Manawatu	water	pūkeko, takahē	*C. jejuni*	3	2381	-
569a	NZ	Manawatu	human	livestock	*C. jejuni*	4	474	CC48
H22082	NZ	Manawatu	human	livestock	*C. jejuni*	4	474	CC48
H704	NZ	Manawatu	human	livestock	*C. jejuni*	4	474	CC48
H892	NZ	Manawatu	human	livestock	*C. jejuni*	4	48	CC48
m73020	NZ	Manawatu	human	livestock	*C. jejuni*	4	474	CC48
P110b	NZ	Manawatu	chicken	chicken	*C. jejuni*	4	474	CC48
P179a	NZ	Manawatu	chicken	chicken	*C. jejuni*	4	474	CC48
P694a	NZ	Manawatu	chicken	chicken	*C. jejuni*	4	474	CC48
B2124	AU	Melbourne	swamphen	swamphen	*C. jejuni*	5	-	-
B2127	AU	Melbourne	swamphen	swamphen	*C. jejuni*	5	-	-
B2128	AU	Melbourne	swamphen	swamphen	*C. jejuni*	5	-	-
B2129	AU	Philip Is.	swamphen	swamphen	*C. jejuni*	5	-	-
B2130	AU	Philip Is.	swamphen	swamphen	*C. jejuni*	5	-	-
B2132	AU	Philip Is.	swamphen	swamphen	*C. jejuni*	5	-	-
B2133	AU	Philip Is.	swamphen	swamphen	*C. jejuni*	5	-	-
B2134	AU	Philip Is.	swamphen	swamphen	*C. jejuni*	5	-	-
N27	NZ	Waikato	water	unknown	*C. jejuni*	5	3845	-
N53	NZ	Manawatu	water	unknown	*C. jejuni*	5	3845	-

Continued on next page

DAPC —– *continued from previous page*

Isolate	Country	Region	Sample	Host	Species	DAPC cluster	ST	Clonal complex
W120a	NZ	Manawatu	chicken	chicken	*C. jejuni*	5	3673	-
W135a	NZ	Manawatu	water	pūkeko	*C. jejuni*	5	3655	-
W63b	NZ	Manawatu	water	takahē	*C. jejuni*	5	3663	-
B1491	NZ	Manawatu	starling	starling	*C.* sp. nov. 3	6	-	-
W441b	NZ	Manawatu	water	unknown	*C.* sp. nov. 3	6	-	-
W677a	NZ	Canterbury	water	unknown	*C.* sp. nov. 3	6	-	-
B2020a	NZ	Manawatu	weka	weka	*C.* sp. nov. 2	7	-	-
B2092c	NZ	Marlborough	weka	weka	*C.* sp. nov. 2	7	-	-
B2098b	NZ	Golden Bay	banded rail	banded rail	*C.* sp. nov. 2	7	-	-
B2123	AU	Melbourne	swamphen	swamphen	*C.* sp. nov. 4	7	-	-
B2126	AU	Melbourne	swamphen	swamphen	*C.* sp. nov. 4	7	-	-
H773	NZ	Manawatu	human	chicken	*C. jejuni*	8	3711	CC257
m28548	NZ	Manawatu	human	unknown	*C. jejuni*	8	2026	CC403
M880a	NZ	Manawatu	sheep	cattle, sheep	*C. jejuni*	8	2341	CC61
S22b	NZ	Manawatu	cattle	unknown	*C. jejuni*	8	2026	CC403
76062a	NZ	Manawatu	human	chicken, cattle	*C. jejuni*	8	190	CC21
H742	NZ	Manawatu	human	livestock	*C. jejuni*	8	50	CC21
H798	NZ	Manawatu	human	livestock	*C. jejuni*	8	50	CC21
m28127	NZ	Manawatu	human	cattle, sheep	*C. jejuni*	8	520	CC21
P164a	NZ	Manawatu	chicken	livestock	*C. jejuni*	8	190	CC21
P28a	NZ	Manawatu	chicken	cattle, sheep	*C. jejuni*	8	520	CC21

Continued on next page

DAPC —– *continued from previous page*

Isolate	Country	Region	Sample	Host	Species	DAPC cluster	ST	Clonal complex
P553b	NZ	Manawatu	chicken	livestock	*C. jejuni*	8	53	CC21
S107c	NZ	Manawatu	cattle	livestock	*C. jejuni*	8	53	CC21
S251a	NZ	Manawatu	cattle	livestock	*C. jejuni*	8	53	CC21
S264a	NZ	Manawatu	cattle	livestock	*C. jejuni*	8	50	CC21
S331b	NZ	Manawatu	cattle	livestock	*C. jejuni*	8	21	CC21
S85b	NZ	Manawatu	cattle	livestock	*C. jejuni*	8	53	CC21

5.4 Discussion

The overall pattern of clustering between the pūkeko/takahē isolates and Australian purple swamphen is consistent with a recent shared ancestry followed by divergence, suggesting the New Zealand isolates could be following a different evolutionary trajectory over the period of separation and accumulating differences. These results show a close relationship between the genomes of *C. jejuni* from the Australian purple swamphen and the New Zealand isolates associated with pūkeko and takahē, consistent with a strong host-association bond despite the geographical and temporal separation, but allowing for some variation. While the same ST (sequence type) seen in the pūkeko and takahē was not found in the Australian purple swamphen, the ST did cluster together in a consistent manner in both the core and pan-genome analyses.

The phylogenetic pattern, in the core and pan-genome comparisons, generally has the Australian purple swamphen *C. jejuni* clustering closely with ST-3673, ST-3845, ST-3663, and ST-3655 while the ST-2381 isolates are nearby but less closely clustered. A possible cause of this relationship is the ST-3673, ST-3845, ST-3663, and ST-3655 isolates arrived in New Zealand with the pūkeko ancestor but the ST-2381 arrived earlier in the takahē ancestor. Based on this assumption it would suggest that over the longer separation period within the takahē, the ST-2381 diverged more but stayed within some host-association boundary so when pūkeko arrived they were still compatible in both hosts. Alternatively, the ST-2381 could be occupying a similar but slightly different niche within the gastro-intestinal system which has resulted in an accumulation of differences. In the previous phenotypic analysis, ST-2381 and ST-3655 showed a similar pattern of carbon source utilisation at 42℃, which makes ST-2381 being in a slightly different gastro-intestinal niche a less likely explanation. However there are other possible explanations, such as insufficient sample size. The founding *Porphyrio* ancestors that arrived in New Zealand were probably small in number and could have lacked the ST we found in our Australian sample, or our Australian samples may not have been large enough to detect the full range of ST.

Two non-*C. jejuni* species were also identified in the Australian purple swamphen, a new *Campylobacter* spp. *C.* sp nov. 4 and another recently identified *Campylobacter* spp. *C.* sp nov. 1 [16][p227]. More work, such as phenotypic analysis will be required to fully establish this claim as the *in silico* analysis alone, does not meet the minimum standards proposed by Ursing et al. for a new species of Campylobacteraceae [215]. *C.* sp nov. 4, as far as we know has not been previously reported and so far has only been found in the Australian purple swamphen. There are several possible reasons why *C.* sp nov. 4 has not been found in New Zealand birds, particularly rails, they include; being missed in sampling, absent by chance in the birds that colonised New Zealand, absent because the species had not entered the Australian population when New Zealand was colonised, and the species had not evolved when New Zealand was colonised by the pūkeko's ancestor.

Campylobacter sp. nov. 1, along with *Campylobacter* sp. nov. 2 and *Campylobacter* sp. nov. 3 were found in previous investigations by mEpiLab members which were not part of this body of work [16][p227]. *C.* sp nov. 1 had been previously found in the pūkeko and takahē [16][p227] and now it has been isolated from the Australian purple swamphen. The simplest explanation for this distribution would be that *C.* sp nov. 1 arrived in New Zealand with the pūkeko or the takahē, suggesting the species has been around ≥ 500 years. More analysis of these all these new species is required but as *Campylobacter* spp are often sub-divided into thermophilic (e.g. *C. jejuni*, *C. coli*, *C. lari* and *C. upsaliensis*) and non thermophilic (*C. fetus*, *C. curvus*, *C. hominis* etc) due to the isolation method (incubation at 42℃) and demonstrated phylogenetic closeness (see Figure 5.3, 5.12) to *C. coli*/*C. jejuni*, the putative new species (*C.* sp. nov. 1, 2, 3 & 4) are likely to be classified as thermophilic when their taxonomic identification is completed.

Throughout these genomic comparisons there has been a consistency in the relationships between isolates when comparing core genome features (Figure 5.12) or making an across the entire genome comparison(Figure5.13, Table 5.3). Such a consistency suggests the forces acting on the *C. jejuni* core genome are acting on the accessory genome, keeping a connection between the two with an allowance for minor variation. This finding starts to explain how ST and clonal complexes have been successfully used to associate isolates to hosts [7] [6], i.e. the housekeeping genes (core genome and essential genes) are reflecting the type of metabolic capabilities that exist in the accessory genome. Previously an association between ST and metabolic traits had been identified based on specific genes [498][554], and at a genomic level using suppression subtractive hybridisation Hepworth et al (2007) there was a correlation between the common clonal complexes and metabolic genes [369]. However this investigation showed the relatedness exists across the core and accessory genomes, in wild-bird associated and agriculture-associated ST/CC. The strong pattern of association does not appear to extend from a direct connection to a given MLST allele, rather it is reflected in total allelic profiles. This study does not identify the evolutionary mechanism but as *C. jejuni* is known for high levels of recombination it is a likely factor [415] [555]. Barriers to recombination between ST have been reported in agriculture associated lineages: with minimal recombination between CC21 and CC45 [421], and lineage specific restriction-modification systems identified in CC403 [556]. These results suggest at the genome level for both agriculture-associated and wild-bird associated *C. jejuni*, there may be a combination of recombination mainly occurring between more closely related sequence types producing homogeneity, and possibly some purifying selection for host adaptation removing the "less adapted" combinations.

There was a level of demarcation shown between the *Campylobacter* spp. using *in silico* DNA-DNA-hybridisation (DDH) comparisons (Figures 5.4, 5.5, 5.6) that was not apparent in either of the 16S rRNA analysis. The differences between the *Campylobacter* spp. was apparent in the rMLST (Figure 5.12) and the pan-genome comparisons of both the ADD (Figure 5.13) and to a lesser degree in the DAPC clustering (Table 5.3, Figure 5.14). In both the DDH comparison with international

reference genomes (Figures 5.6, 5.5, 5.4), and the New Zealand isolates (Figures 5.8, 5.9, 5.7), eight of the isolates can be identified as *C. jejuni*. By expanding the comparison base to include the New Zealand isolates, it was possible to group the remaining six Australian purple swamphen isolates into two different groups. One group of four (Figures 5.7, 5.8, 5.9) showed a species level similarity to a recently identified new species (*C.* sp. nov. 1) [16]. The remaining two isolates, based on a lack of similarity with other *Campylobacter* spp. and a marked similarity to each other, suggest they represent a new species (*C.* sp. nov. 4). Two of the putative new *Campylobacter* spp. (*C.* sp. nov. 1, *C.* sp. nov. 2) in the New Zealand dataset have previously been isolated from New Zealand water rails (pūkeko and takahē Table 5.3) [16], the third *C.* sp. nov. 4 has now been isolated from Australian purple swamphen, a close relative. The new *Campylobacter* species and new strains found in this investigation and others [312] [330] suggest that as we look at more wild animals we will find more new *Campylobacter* spp. and *C. jejuni* sequence types which will lead to a better understanding of the microbial ecology and evolution of this genus.

The lack of clear differentiation between some of the *Campylobacter* spp. (*C. coli* and *C. jejuni*) in the 16S rRNA comparisons (Figure 5.2, 5.3) is a recognised failing of 16S rRNA analysis [549] [557]. 16S rRNA comparisons have been shown as too conserved to discriminate some closely related species [558]. Despite the 16S rRNA analysis being performed as two different comparisons, a pairwise similarity matrix (Figure 5.3) using international reference sequences, and a Neighbor-Net visual representations of a larger dataset of New Zealand isolates (Figure 5.1, 5.2), the results do not show much distance distinguishing between *C. coli*, *C. jejuni*, *C.* sp. nov. 1, 2 and 4; suggesting they are all closely related. There was 100% similarity between some *C. jejuni* STs, particularly ST-474 (all 6 isolates), and some *C. coli* (mainly CC828 and may be part of clade 1). This level of similarity, seen in Figure 5.2, has been previously reported between other isolates of these species [549]. ST-474 is an important source of poultry associated campylobacteriosis in New Zealand, although rarely reported outside of the country [67] [15]. Based on 16S rRNA the *C. coli* have clustered into two groups, one was *C. coli* alone and the other a mixture of *C. coli* and *C. jejuni* [557]. Sheppard et al. (2011, 2013) suggested there has been introgression between some *C. coli* (clade 1) and *C. jejuni* [559] [560]. It is possible this result reflects such a event with some degree of recombination between ST-474 and *C. coli*.

Two Australian purple swamphen *C. jejuni* isolates (B2124, B2134), did not follow the trend to cluster with the same ST, in both 16S rRNA analyses (Figure 5.2 5.3). Based on ST B2134, should group with B2127, B2128, B2129, B2130, B2132, B2133 (Table 5.2) which share all the same MLST allelic profile. B2124 and B2134 differ from each other at the *pgm* loci (293,7) by 13bp. Both B2124 and B2134 have the same 16S rRNA sequence and this only varies by 1–2bp SNP from the ST-2381 isolates, which are found in the pūkeko and takahē, but vary by 7–8bp SNP to the other Australian purple swamphen *C. jejuni* isolates. This could be an example of variation within ST, genomic variation has been reported between ST-474 isolates [273] as has variation in 16S rRNA within

Campylobacter spp. species [549].

It was first proposed in 1977 [561] to use 16S rRNA to establish phylogenetic relationships, and using a cut off point of $\geq 97\%$ sequence similarity between two isolates has been correlated with a 70% DNA–DNA hybridisation (DDH) value [562]. This analysis used the $\geq 98.65\%$ cut-off suggested for full length 16S rRNA sequences [563]. There was a noticeable difference between the relative closeness of Australian purple swamphen *C. jejuni* to *C. lari*, as shown in the EzTaxon pairwise similarity (Figure 5.3), compared to the marked difference between *C. jejuni* and *C. lari* shown by the Neighbor-Net analysis (Figure 5.1). A possible reason for these variable results is a combination of some *Campylobacter* spp., like *C. jejuni* and *C. coli* are very closely related, and variation in 16S rRNA within a species, for example variation of 16S rRNA within *C. lari* has been reported [549]. This means an approach that tends to choose a single example of a species/phylotype [524] would be vulnerable to missing some similarities unless every cluster/clade is represented.

The Discriminant analysis of principal components (DAPC) analysis in general splits the genomes into groups by *Campylobacter* spp. and the *C. jejuni* population along host-association and ST/CC. An interesting feature of the DAPC analysis is cluster 3 which showed a combination of different ST and CC (CC45, CC42, CC177, CC682). CC45 has been associated with environmental water isolates [564], as well as agricultural livestock. The other clusters of *C. jejuni* isolates did not show this mixing of wild-bird and agriculture related isolates. A possible implication of this grouping is a common ancestry between these ST and CC45. Alternatively this cluster could be an artefact of this DAPC analysis on this dataset either due to the fitting of the data to the optimum number of clusters (8) predicted by kmeans.

The population structure suggested by the DAPC analysis does not as strongly support species differentiation of *C.* sp. nov. 2 and 4, in contrast to the rMLST (Figure 5.12) and the *in silico* DNA-DNA hybridisation results(Figures 5.8, 5.9, 5.7). This may be the effect of the shared environment (host i.e. all the isolates in cluster 7 are from members of the rail family), leading to many shared pathways in the pan-genome, and resulting less differentiation between the species on a simple presence/absence of a gene comparison. In Figure 5.13 which is based on the pan-genome *C.* sp. nov. 2 and 4 are very close. Alternatively the clustering of *C.* sp. nov. 2 and *C.* sp. nov. 4. together could be due the small number of genomes of the putative new species in the dataset and the principal component analysis emphasis is on variability while the features that are predictive of these putative species may be at a lower variance in terms of a pan-genome presence/absence data.

The pan-genome of the 84 *Campylobacter* spp. isolates in our dataset, which is possibly a mixture of 6 species, has a large proportion of homologues (43.4%) that did not cluster with any other sequence i.e. single occurrences. There are various possible reasons for this, for example, some of the single homologues appeared to be truncated versions of genes found in other homologue clusters

e.g. *cdt*A but did not meet the minimum coverage criteria. Another possibility is that there are a small number of several putative new species in this group and if more genomes from each of these species had been included they have found matching homologues. There are a large proportion of hypothetical proteins (55.0%) in the pan-genome and a larger proportion in the lower occurring homologue clusters (63.8% single occurring homologues) [565][p46], which suggests that there is a lot more to be learnt about the genes that make up *Campylobacter*, their functions and what role they play in inter-species and intra-species differences, and are they part of a large mobile element in the pan-genome?

The Dollo process used in the ADD analysis (Figure 5.13) assumes that a trait can only arise once on a phylogenetic tree but may be lost many times [545]. However recombination is a significant factor in *C. jejuni* evolution, as it seems to occur more frequently than mutation [566], so the reintroduction of a "lost" gene may occur more frequently than assumed in the Dollo assumption. This may be countered by ongoing selection removing those that did not aid fitness and it would reduce the probability of such a gene remaining (and being found) i.e. it would be "lost" again. Even if the assumption of the Dollo process is not an exact representation of the situation for the *Campylobacter* species, this analysis (Figure 5.13) showed consistent results with the rMLST comparison (Figure 5.12), which does not make the same assumptions, suggesting it was not a poor fit.

It is difficult to make comparisons of core and pan-genome size to previous estimates of *Campylobacter* spp. as they used alternative methods, parameters and *Campylobacter* spp. datasets. For example, Lefebure et al. (2010) estimated from 96 genomes of two closely related *Campylobacter* spp.(*C. jejuni* and *C. coli*) that each taxon has a pan-genome size of ∼2,600 genes, and the combined pan-genome of both species was ∼3,000 genes [416]. The core genome of *C. jejuni* was estimated as 1369 and 1485 for *C. coli*, with a shared core genome size of 1317 [416]. Zhou et al. (2013) analysed 34 genomes, containing 10 *Campylobacter* spp., then used a genome wide alignment method where all protein coding sequences where placed in the same gene family cluster if the best local alignment between the sequences covered at least 50% of the length of both sequences and contained at least 50% identity. This approach produced a pan-genome of 13,167 gene families and 348 core gene families are at the genus level. The estimated core genome for four species (*C. jejuni*, *C. coli*, *C. lari*, and *C. upsaliensis*) was 1074 gene families. Another estimate of core and pan-genome for *C. jejuni* only comes from Sheppard et al. (2014), using 128 isolates, and a parameter of reciprocal best hits to *C. jejuni* 11168 loci, with at least 70% nucleotide identity and 50% identity in alignment length using the BLAST [421]. This approach produced a core genome with 595 loci and 3,485 accessory genes [421]. For my dataset of 84 genomes containing six putative *Campylobacter* spp. the core genome was estimated as ∼550 orthologues and the pan-genome ∼8,500, based on BLAST searches with a minimum E-value (maximum expectation value) of 10^{-5} and a minimum of 75% coverage in pairwise alignment for inclusion in an orthologue group. Considering the differences in

the methods and *Campylobacter* spp. involved, the results for this analysis are not inconsistent with previous work.

In the PathogenFinder [530] analysis, ST-474, a significant cause of campylobacteriosis in New Zealand [273], showed a 83% probability of being a human pathogen for all seven of the isolates. This suggests the $81.0 - 83.8\%$ estimation of probability for being a human pathogen for the Australian purple swamphen isolates supports the argument they could be pathogenic and/or have the potential to act as a genetic reservoir i.e. a source of virulence factors for agricultural strains [17]. However this needs to be viewed in context as wild-bird associated *C. jejuni* sequence types are currently infrequently identified as the cause of disease in humans [324], and reports of pathogenicity in *Campylobacter* spp. other than *C. coli* or *C. jejuni* are not common although there is growing recognition of such events [516][p123]. More investigation is required into *Campylobacter* virulence genes and genetic flow between these different populations to understand the level of risk that may exist in recombination between wild-bird and agricultural hosts of *Campylobacter* spp..

The mobilome which, by definition, consists of genetic material that is readily exchanged appears to be a common feature in the form of predicted genomic islands and prophages in the Australian purple swamphen isolates, and a similar result has been reported in the *C. lari* group [567]. We have not analysed the frequency of movement and retention of these features, but just as *C. jejuni* has been shown to have high levels of recombination in the core genome [415], it is possible genomic islands and prophages have an important effect on the evolution of this bacteria. No antibiotic resistance genes were detected in this analysis, these genes can be part of mobile genetic elements, and this result may be due to the isolates being unlikely to be exposed to antibiotics, certainly compared to isolates associated with agriculture. The draft genomes were not investigated for the presence of plasmids [568], nor were bench based techniques used in this chapter although such work could be performed in the future as plasmids have been associated with pathogenicity [569].

There are many steps from collection of our samples through to the final culturing that may have affected our results, for example by choosing a single colony that appeared to be *Campylobacter* from each plate, we removed our ability to identify if each bird had more than one species or genotype of a species. In regard to the putative new species it could be important to understand if they occur as co-infections and the degree of interaction and genetic exchange that may occur inside the host. Another element that may have affected our results was using only 42℃ at the initial identification step incubation temperature rather than using both a 37℃ and 42℃ incubation temperature at this stage, as recent work suggests this may affect which isolates are found [381]. Repeated sub-culturing of *C. jejuni* has been associated with both phenotypic and genotypic change in the isolates [570]; we minimised repeated sub-culturing by using long term storage of the isolates at -70℃ and -80℃ instead and reviving the isolates when required.

Conclusion

There is a consistent pattern of close relationship between *C. jejuni* from the Australian purple swamphen (*Porphyrio porphyrio melanotus*) and isolates associated with the New Zealand pūkeko (*Porphyrio porphyrio melanotus*) and takahē (*Porphyrio hochstetteri*), despite an estimated 400–600 years of separation [12][pp 590-595] [514][pp 38]. The phylogeny of both the core genome and pan-genome supports a recent common ancestor and divergence probably after separation. There is also a consistency of relationship in the *C. jejuni* between the core genome as measured by MLST and rMLST, and extending into the patterns of gene presence and absence in the pan-genome for both the wild-bird and agriculture-associated isolates. This relationship between the core and pan-genome starts to explain how particular MLST and clonal complexes have been able to be successfully associated with particular hosts. The allelic profile of the housekeeping genes (core genome) reflected the type of metabolic capabilities that exist in the accessory region (optional section of the pan-genome) linking phenotype genotype and host. Using genomic analysis (Figures 5.5,5.6, 5.4, 5.8, 5.8, 5.7), the presence of a new *Campylobacter* spp. *C.* sp nov. 4, in Australian purple swamphen (*Porphyrio porphyrio melanotus*) has been demonstrated. More work, such as phenotypic analysis will be required to fully establish this claim as by itself the *in silico* analysis alone does not meet the minimum standards proposed by Ursing et al. for a new species of Campylobacteraceae [215]. Also identified was the presence of another recently identified *Campylobacter* spp. *C.* sp nov. 1 [16][p227]. So far *C.* sp nov. 1 has only been reported in two closely related water rail in New Zealand, the pūkeko (*Porphyrio porphyrio melanotus*) and takahē (*Porphyrio hochstetteri*).

Acknowledgements:

Thank you to a group of wonderful volunteers from BirdLife Australia: Will Steele, Katrina Lee, Christine Hurring, Sandra Martig, Dorothy Dhaeze, Michael Tarburton, Tiffany Miller-Scott, and Peter Dann for collecting the samples. Dr Carl Muhlnickel and the staff of ASAP Laboratory of Melbourne Australia, who cultured the samples, then isolated and extracted some of the DNA. Dr Mary Valcanis of the Enteric Group, Reference Section, Microbiological Diagnostic Unit - Public Health Laboratory, Department of Microbiology \& Immunology, The University Of Melbourne, Victoria, Australia helped in storing and transporting of samples to New Zealand. Lynn Rogers and Rukhshana Nahid Akhter performed the MLST PCR analysis of some samples.

Supplement

Table 5.4: **The isolates used as international reference sequences for 16S rRNA.** The full name of species and subspecies or strain is provided along with any abbreviation used in Figures 5.1 and 5.2. Refseq refers to the NCBI (http://www.ncbi.nlm.nih.gov/refseq/) Reference Sequence Database accession number and BioProject (http://www.ncbi.nlm.nih.gov/bioproject) is the NCBI reference to data on the project of origin.

Species and Strain type	**abbreviation**	**Refseq**	**BioProject**
Campylobacter concisus 13826	Cacon_13826	NC_009802.1	PRJNA58667
Campylobacter curvus 525.92	Cacu_525_92	NC_009715.1	PRJNA58669
Campylobacter fetus subsp. *fetus* 82-40	Cafe_82-40	NC_008599.1	PRJNA58545
Campylobacter hominis ATCC BAA-381	Caho_ATCC_BAA-381	NC_009714.1	PRJNA58981
Campylobacter jejuni RM1221	Caje_RM1221	NC_003912.7	PRJNA57899
Campylobacter jejuni subsp. *doylei* 269.97	CajeD_269.97	NC_009707.1	PRJNA58671
Campylobacter jejuni subsp. *jejuni* 1336	Caje_1336	NZ_CM000854.1	PRJNA43391
Campylobacter jejuni subsp. *jejuni* 414	Caje_414	NZ_CM000855.1	PRJNA43389
Campylobacter jejuni subsp. *jejuni* NCTC 81116	Caje_81116	NC_009839.1	PRJNA58771
Campylobacter jejuni subsp. *jejuni* 81-176	Caje_81-176	NC_008787.1	PRJNA58503
Campylobacter jejuni subsp. *jejuni* IA3902	Caje_IA3902	NC_017279.1	PRJNA159531
Campylobacter jejuni subsp. *jejuni* ICDCCJ07001	Caje_ICDCCJ07001	NC_014802.1	PRJNA61249
Campylobacter jejuni subsp. *jejuni* M1	Caje_M1	NC_017280.1	PRJNA159535
Campylobacter jejuni subsp. *jejuni* NCTC 11168	Caje_NCTC11168	NC_002163.1	PRJNA57587
Campylobacter jejuni subsp. *jejuni* S3	Caje_S3	NC_017281.1	PRJNA159533
Campylobacter lari RM2100	Cala_RM2100	NC_012040.1	PRJNA58115

Table 5.5: **The reference sequences used in the *in silico* DNA-DNA hybridisation comparison of Australian purple swamphen (*Porphyrio porphyrio melantotus*) isolates** . These isolates are used in in Figures 5.4, 5.5, 5.6. RefSeq refers to the NCBI (http://www.ncbi.nlm.nih.gov/refseq/) Reference Sequence Database accession number and uid is the NCBI unique identity number for the sequence.

Species and Strain	abbreviation	RefSeq	uid
Campylobacter jejuni subsp. jejuni 00-2544	C_ jejuni_00_2544	NC_022353.1	uid219326
Campylobacter jejuni subsp. jejuni 00-2425	C_ jejuni00_2425	NC_022362.1	uid219359
Campylobacter jejuni RM1221	C_ jejuni_RM1221	NC_003912.7	uid57899
Campylobacter jejuni subsp. jejuni 00-2538	C_ jejuni_00_2538	NC_022351.1	uid219325
Campylobacter jejuni subsp. jejuni 00-2426	C_ jejuni_00_2426	NC_022352.1	uid219324
Campylobacter jejuni subsp. jejuni M1	C_ jejuni_M1	NC_017280.1	uid159535
Campylobacter jejuni subsp. jejuni 81116	C_ jejuni_81116	NC_009839.1	uid58771
Campylobacter jejuni subsp. jejuni IA3902	C_ jejuni_IA3902	NC_017279.1	uid159531
Campylobacter jejuni subsp. jejuni S3	C_ jejuni_S3	NC_017281.1	uid159533
Campylobacter jejuni subsp. jejuni 81-176	C_ jejuni_81_176	NC_008787.1	uid58503
Campylobacter jejuni subsp. jejuni ICDCCJ07001	C_ jejuni_ICDCCJ07001	NC_014802.1	uid61249
Campylobacter jejuni subsp. jejuni NCTC 11168-BN148	C_ jejuni_NCTC_11168_BN148	NC_018521.1	uid174152
Campylobacter jejuni subsp. jejuni NCTC 11168 = ATCC 700819	C_ jejuni_NCTC_11168__ATCC_700819	NC_002163.1	uid57587
Campylobacter jejuni subsp. jejuni 4031	C_ jejuni_4031	GCA_ 000493495.1	
Campylobacter jejuni subsp. doylei	C_ jejuni_ doylei_269_97	NC_009707.1	uid58671
Campylobacter coli 76339	C_ coli_76339	NC_022132.1	uid217050
Campylobacter coli 15-537360	C_ coli_15_537360	NC_022660.1	uid226113
Campylobacter coli CVM N29710	C_ coli_CVM_N29710	NC_022347.1	uid219322
Campylobacter fetus subsp. fetus 82-40	Cafe_82_40	NC_008599.1	uid58545
Campylobacter sp. 03-427	Campylobacter_03_427	NC_022759.1	uid226993
Campylobacter curvus 525.92	Cacu_525_92	NC_009715.1	uid58669
Campylobacter lari RM2100	Cala_RM2100	NC_012040.1	uid58115
Campylobacter hominis ATCC BAA-381	Caho_ATCC_BAA_381	NC_009714.1	uid58981
Campylobacter concisus 13826	Cacon_13826	NC_009802.1	uid58667

Appendix C in DVD

Summary of information provided in Appendix C

Appendix C on the DVD provided contains:

1. Protocol for collection of Australian purple swamphen samples.
2. Australian purple swamphen KEGG maps - The 14 Australian purple swamphen isolates have had their genomes sequenced and mapped to KEGG pathways. The maps for each pathway are provided as web pages.
3. *Campylobacter* spp. isolates - A list of *Campylobacter* spp. isolates from New Zealand pūkeko in 2009 and 2012, the isolates used in chapter 4, and all the Australian purple swamphen isolates.
4. *Porphyrio* spp. - provides background information on the *Porphyrio* spp. bird and their distribution.

The file may need to be loaded onto a computer hard drive and unzipped.

Science may be described as the art of systemic over-simplification - the art of discerning what may with advantage be omitted

Karl Popper

6

the metabolism:

Genomic comparison of metabolic subsystems present in *Campylobacter* spp. dataset.

Prelude

This is the final investigation into the link between phenotype, genotype, and host adaptation in *C. jejuni*, and asks if at the metabolic level there is an identifiable "host signature" i.e. showing metabolic features that suggest they are adapted to a given host. This tests a null hypothesis that there is no metabolic host adaptation.

A dataset of predominantly New Zealand *C. jejuni* draft genomes was annotated into curated metabolic pathways, which grouped them by functional categories and into a series of subsystems or metabolic pathways. The isolates belonged to three host groups: starlings (*Sturnus vulgaris*), Australian purple swamphen, and common clonal complexes for livestock isolates (CC21, CC42,

CC45, CC48). The three groups of genomes were used as a training set to identify features that separated them from each other. Three subsystems were identified as being an important difference between these groups and were compared to a larger set of *Campylobacter* spp. genomes, to identify if the three subsystems were host-associated differences. This approach also enables a comparison, using the three subsystems, between the Australian purple swamphen and the pūkeko/takahē associated isolates to identify changes that may have evolved between them.

This chapter contains the metabolic analysis, based on genomic composition, of the mEpiLab genomes dataset. The mEpiLab genomes dataset contained 70 draft *Campylobacter* genomes from New Zealand and 14 Australian purple swamphen *Campylobacter* genomes from Australia. The results of this analysis were not verified using Biolog phenotypic microarray plates.

This chapter was written as an article to be published in the Journal of Basic Microbiology

Comparison of *Campylobacter jejuni* genomes: host associated metabolic differences.

6.0.1 Abstract

Internationally *C. jejuni* is an important cause of gastroenteritis. Epidemiological analysis has shown most isolates associated with human disease originate from agricultural sources while wild-bird hosts are a less likely source of infection. In this analysis we used a unique genomic dataset containing *C. jejuni* from three host groups: starlings (*Sturnus vulgaris*), Australian purple swamphen (*Porphyrio porphyrio melanotus*), and a group of clonal complexes (CC21, CC42, CC45, CC48) associated with agriculture. The genomes were annotated by Rapid Annotations using Subsystems Technology (RAST) that allocated the protein sequences into metabolic pathways called subsystems. Three important subsystems were identified as showing marked variation between the host groups; Coenzyme A biosynthesis, Type VI secretion system, and *Campylobacter* spp. iron metabolism. The three subsystems were investigated using genomes from an expanded dataset to identify their distribution pattern. Vitamin B_5 (pantothenate) biosynthesis, a part of the Coenzyme A biosynthesis subsystem was present in cattle isolates but not in the wild-bird isolates. The Type VI secretion system subsystem was only found in a small group containing isolates from the Australian purple swamphen, CC403, ST-2381, ST-3655 including a new *Campylobacter* spp.. There was variation in the *Campylobacter* spp. iron metabolism subsystem between the isolates. These results show there are important metabolic differences between isolates based on host association, suggesting host adaptation, in both wild-bird and agriculture associated *C. jejuni*.

6.1 Introduction

Campylobacteriosis is considered to be the most common cause of gastro-enteritis worldwide with over 400 million cases estimated annually [96] [571] and most developed countries showing a upward trend in reported cases [83]. As an emerging disease associated with the consumption of contaminated food or water it is a major area of interest: however difficulties in growing isolates in the laboratory and low biochemical reactivity, combined with marked phenotypic and genotypic variation, made early progress slow [30]. In developed countries food producing animals are the main source of *C. jejuni* associated with human cases, with chicken usually being the commonest source [1]. Occasionally human cases have been attributed to sequence types associated with wild-birds [324], however it is not known if the low level of infection ascribed to wild-bird isolates compared to livestock isolates is due to lack of pathogenicity or lack of opportunity for transmission [414]. Epidemiological analysis using multilocus sequence typing (MLST) of isolates from livestock and human cases has shown that many of the sequence types (ST) can be grouped into closely related clonal complexes such as clonal complex ST-21 (CC21), clonal complex ST-42 (CC42), clonal complex ST-45 (CC45), and clonal complex ST-48 (CC48) [498] [272]. These agriculture-associated ST often have the ability to colonise multiple livestock species in particular sheep, cattle and chickens [6]. Meanwhile investigations into *C. jejuni* isolates from wild-birds has shown a different pattern, with ST showing host specificity, that persists even over large geographical distances [7] [319].

In 2000, the first complete genome (*C. jejuni* NCTC 11168) was reported [389], and continued investigation into the genome is opening up the ability of a wide range of disciplines, including taxonomy, epidemiology and physiology to contribute to our understand of this pathogen [572] [300] [573]. While most *C. jejuni* genomic research has focused on human and agricultural related isolates, some genomic analysis has differentiated between agriculture-associated and wildlife-associated isolates [368], for example, the presence of a flagellin glycosylation locus (Cj1321-Cj1326) in livestock associated isolates [289]. On the basis of ST, metabolic genes have been associated to several mainly agriculture-associated CC [554] [498]. Further investigation into differences in metabolism-associated genome should identify metabolic features that represent host adaptation.

In the previous chapter we have shown by comparing genomes from wild-birds and agricultural sources a significant and persistent pattern of differences in both core and pan-genomes. These differences also appear to be between wild-birds associated ST, if they are not closely related e.g. starling and pūkeko. However further investigation is required to identify where these differences appear at a metabolic level i.e. the different metabolic pathways that may be involved in host adaptation by *C. jejuni*. To do this, the genomes were annotated using the RAST and SEED servers

[532] [531] and the protein encoding genes (peg single, pegs plural) were identified and allocated into metabolic pathways called subsystems. In this analysis we started with a training dataset containing three host associated groups, then once notable metabolic features that differentiated the three groups were identified, they were scrutinised against a larger group of *Campylobacter* spp. genomes to see if they remained informative about host association. Three potentially important subsystems were identified and investigated: Type VI secretion system, Coenzyme A biosynthesis and *Campylobacter* spp. iron metabolism.

Type VI secretion system (T6SS) has been identified in a range of bacterial species and has been associated with pathogenicity [574], with cytotoxic effects that may have originated as a mechanism against flagellates, amoeba and ciliates [575] [576]. Relatively recently, T6SS has been identified in *C. jejuni* found in several countries [577] [578]. T6SS's role in pathogenicity may involve host cell adherence and invasion [579] [580]. A complex core of 13 subunits have been identified, which may act as a single functional unit [581]. Horizontal gene transfer (HGT) is believed to be the major transmission pathway for T6SS between bacteria [581]. The structure of T6SS resembles a bacteriophage sheath tube and tail spike proteins and it is hypothesised that this is a delivery mechanism for molecules from the bacterial host across the cell membrane into another cell [582].

Coenzyme A (CoA) is a universal and essential cofactor, vital for the regulation of key metabolic enzymes and involved in numerous metabolic and energy-yielding reactions [583]. The Coenzyme A biosynthesis subsystem, used in this analysis has 9 steps, including three steps involving vitamin B_5 (pantothenate) biosynthesis (Figure 6.4). If the enzymes involved in these reactions are absent, the cell will need to acquire vitamin B_5 from the environment in the host. Vitamin B_5 biosynthesis has been previously reported as a possible factor in host adaptation in *C. jejuni* with a tendency to be absent in non-agriculture-associated isolates [502][p93] and present in cattle isolates [370].

Iron is essential to the *Campylobacter* metabolism [469], as a cofactor of many enzymes and involved in the electron transport system and redox reactions [584]. The iron metabolism also plays an important role not only in growth of *C. jejuni* but in survival from oxidative stress in aerobic conditions [585], and it has also been associated with pathogenicity [586]. While iron acquisition is essential for *C. jejuni*, genomic microarray analysis has previously indicated strain variation in this metabolic pathway [288]. The *Campylobacter* spp. iron metabolism subsystem used in this analysis consists of 26 pegs involved in the uptake and regulation of iron.

This investigation aims to identify metabolic features that may be associated with host adaptation by comparing the genomes associated with three different host groups. To do this, *C. jejuni* genomes, associated with either agricultural livestock, starlings, or Australian purple swamphen were annotated as protein encoding genes (pegs) and classified by function into metabolic pathways called subsystems. Three subsystems were identified as important differences between these groups

and compared to a larger set of *Campylobacter* spp. genomes to identify if these three subsystems still exhibited host associated differences.

6.2 Methods and Materials

Biological materials and DNA extraction

Fresh faecal samples from Australian purple swamphen (*Porphyrio porphyrio melantotus*) in the Melbourne region of Victoria were swabbed and sent for culturing within 24 hours. The swabs were inoculated into Boltons broth and incubated at 42℃ for two days, after which a sterile swab inoculated a sample of the broth onto modified charcoal cefoperazone deoxycholate agar (mCCD) which was then grown at 42℃ in a microaerobic environment produced by CampyGen sachet (Oxoid) for two days. The plate was then inspected for colonies that resembled *Campylobacter* spp. and a single colony was taken and inoculated onto a sterile blood agar plate and grown at 42℃ in a microaerobic environment for two days. At this point the plate was checked for contamination; if none was detected the plate was swabbed and the swab transported to the Microbiological Diagnostic Unit Public Health Laboratory at Melbourne University where it was prepared for storage at −70℃. A horse blood agar plate was inoculated with the swab and incubated in a microaerobic environment, produced by CampyGen sachet (Oxoid), at 37℃ and/or 42℃ depending on its growth; after two days the growth on the plate was inoculated into 20% glycerol and stored at −70℃. When required the isolates were regrown on horse blood agar for two days at 37℃ in a microaerobic environment, from this a loop of growth was taken and stabbed into Wang's medium and incubated at 37℃ in a microaerobic environment for $2-3$ days. The samples were transported to the mEpilab where a swab was taken and grown on Fort Richard Columbia horse blood agar in a MAC500 Workstation microaerobic environment (85% N_2, 10% CO_2, 5% O_2) for two days and inoculated into 15% glycerol solution in labelled vials for storage at −80℃. The isolates where regrown from the glycerol solution on horse blood agar in a microaerobic environment for two days when required for DNA extraction.

Whole genome sequencing

Fourteen Australian purple swamphen isolates (B2121-B2134), consisting of eight *C. jejuni*, *C.* sp. nov. 1 and two *C.* sp. nov. 1, along with four New Zealand *C. jejuni* isolates (N27, N31, N191, W194b) had genomic DNA extracted for next generation sequencing of the genome. The genomic DNA was extracted using either the Promega wizard genomic DNA purification kit or the Qiagen QIAmp DNA mini kit. Both kits were used as per the manufacturers instructions. The genomic DNA samples was fragmented by nebulisation for 6 minutes at a pressure of 32 psi, purified, then end repaired, A-tailed, adaptor-ligated, fractionated, purified and enriched according to the manufacturers instructions, using the TruSeq DNA LT Sample Prep Kit v2-Set A. The prepared

libraries were normalized to equal molarity, diluted to 2nM and pooled: 20 libraries per pool. A flow cell was prepared for each of the library pools and sequencing reactions using 9 pmol of the pooled libraries were performed on an Illumina MiSeq instrument with the MiSeq Reagent Kit v2 for a 2×250 base run, this resulted in approximately 12 to 15 million clusters per run.

The remaining 47 New Zealand *C. jejuni* isolates had genomic DNA extracted and were whole genome sequenced as described in Biggs et al [273]. All 61 isolates used in this chapter are listed in Table 6.1.

Assembly and annotation of the genomes

An in-house work flow based on Perl scripts was used to assemble, quality control and annotate the genomes. Solexa QA [483] was used for trimming and quality control of the reads (Q30, expected probability of incorrect base calling 1:1000). Velvet (ver. 1.2.10) [484] was used for *de novo* assembly of the 2×250 base short reads. A range of kmers were used in the assembly process and the characteristics of the assemblies were compared and ranked using N50, the number of contigs, maximum contig length and assembly length. N50 is a measure of length of the longest contig in an assembly where at least 50% of all base pairs are contained in contigs of this length or larger. The highest ranking contig assembly for a given genome was then annotated with Prokka (ver. 1.10) [485]. All the genomes used in this analysis are draft genomes.

Analysis by RAST Subsystem

The genomes were uploaded to the RAST[1] server [532] [531], in .fna and .fasta format, and using the default setting the sequences were annotated as protein encoding genes (pegs) which were allocated a role in a subsystem [532] [531]. A group of related subsystems form a subcategory, and a group of related subcategories combine into one of 26 metabolic categories. The subsystems represent functional units, like metabolic pathways, and a peg can have roles in more than one subsystem.

A training set of 39 *C. jejuni* genomes, associated either with starlings, Australian purple swamphen or agriculture-associated clonal complexes, was made into a presence/absence matrix based on pegs each with a role in a subsystem. The affinity propagation based agglomerative clustering analysis was undertaken to show that when the genomes were annotated in the RAST as pegs [532] [531], it was possible to distinguish between the three host-associated groups i.e. they clustered into the host-based groups. Next, a binary multi-group discriminant analysis with binary predictors [587] was performed on the presence/absence matrix from 39 *C. jejuni* isolates, to rank each peg in order of importance in distinguishing between the three host-associated groups. To identify a cut-off point, for the number of pegs needed to distinguish between the three host-associated groups, a hierarchical cluster was performed on a presence/absence matrix of the chosen number of pegs. If

[1] http://rast.nmpdr.org

the hierarchical clustering could cluster the genomes back into their original groups, it was assumed enough important pegs had been included. By examining the list of pegs and their associated subsystems it was possible to identify the important subsystems, that showed significance difference between the host-based groups, based on the presence or absence of pegs. Once the subsystems were identified, the pattern of presence/absence was evaluated in a larger dataset of *C. jejuni* genomes, including two new *Campylobacter* spp. found in Australian purple swamphen, to see if the pattern of host association continued.

Affinity propagation based agglomerative clustering

This analysis was limited to *C. jejuni* genomes from two species of wild-bird or clonal complexes associated with agricultural isolates to identify subsystems that differed between the three groups [588]. A total of 39 *C. jejuni* genomes were used: eight Australian purple swamphen isolates, four starling, 12 isolates from CC21, eight from CC48, four CC45, and three from CC42. A presence/ absence matrix was made based on the RAST annotated and categorised pegs [532] [531]. A single peg could be allocated to more one subsystems and this resulted in a total of 1157 subsystem associated pegs across 39 genomes i.e. a peg was linked to a role to a subsystem in the matrix. Using the ApCluster package [588] in R, an affinity propagation based agglomerative clustering process was used employing a binary similarity matrix.

Binary multi-group discriminant analysis

A presence/absence matrix, made from the 39 *C. jejuni* genomes, and based on each peg within a subsystem, was analysed by binary multi-group discriminant analysis with binary predictors [587]. The genomes were placed in one of three host groups: starling, Australian purple swamphen, and the agriculture-associated clonal complexes (CC21, CC48, CC45, CC42) A shrinkage factor was used, due to the small sample size, it was not specified but estimated from the data. A score was calculated for each peg by computing corresponding t-scores between the group means and the pooled mean, which were then ranked. A positive t-score relates to the presence of a peg within a subsystem in the host group while a negative t-score indicates its absence from the host group. The t-scores were ranked in order of highest score, and the top ranked pegs were placed in a presence/absence matrix made from the 39 *C. jejuni* genomes to generate a hierarchical cluster. Using the top 150 ranked pegs produced a clustering pattern, that differentiated between the two wild-bird host-groups and the agriculture-associated group divided into clonal complexes. Then manually reviewing the top 150 pegs, three subsystems were identified as potentially involved in host adaptation: Coenzyme A Biosynthesis, Type VI secretion system (T6SS), and *Campylobacter* spp. Iron Metabolism.

a. Coenzyme A Biosynthesis

The coenzyme A biosynthesis subsystem, in the initial comparison of the three host groups, consisted of either six or nine pegs. A larger dataset was used to investigate if the same pattern of the

coenzyme A biosynthesis subsystem having six or nine pegs was host dependent. A presence/absence matrix of peg in the coenzyme A biosynthesis subsystem was made from 61 genomes, consisting of 55 *C. jejuni* and six putative *Campylobacter* spp. (*C.* sp. nov. 1 and 4) from Australian purple swamphen. A $\tilde{\chi}^2$ test was performed to see if the differences between host groups in the larger dataset were significant.

b. Type VI secretion system (T6SS)

In the initial comparison of the three host groups, the 13 core pegs were either all present or absent in an isolate. In our full dataset, only 20 of the draft genomes showed the presence of T6SS sequences. The sequences for each of the conserved 13 loci [577] were aligned with Muscle [589] and compared in SplitsTree [487] with an unrooted phylogenetic network using a Neighbor-Net [488]. The nomenclature of the RAST/SEED annotation system [532] [531] was used to name the loci: *Imp*B, *Imp*C, *Pvc*109, *Imp*G, *Imp*H, *Vas*D, *Imp*J, *Imp*K, *Clp*B, *Icm*F, *Vgr*G, and *Hcp*. All of the loci except *Vgr*G were concatenated, aligned with Muscle [589] and compared by an unrooted phylogenetic network using Neighbor-Net [488] in SplitsTree [487]. *Vgr*G was not included due to multiple duplications of the loci in some genomes, and this resulted in different phylogenetic relationships between the alleles suggesting they may be in-paralogues under different selection forces (see supplement Figure 6.17).

c. *Campylobacter* Iron Metabolism

The 25 of the 26 pegs in the *Campylobacter* Iron Metabolism were present in the list of 150 top ranked pegs. The relationship between the 61 isolates was investigated using the *Campylobacter* iron metabolism subsystem. The 26 pegs in the *Campylobacter* iron metabolism subsystem (listed in supplementary data 6.5) were analysed as a presence/absence matrix, converted into a distance matrix by the binary method and clustered into four groups using kmeans. A kmeans size of four was chosen based on the hierarchical clustering of the data.

6.3 Results

Affinity propagation based agglomerative clustering

The heatmap for the affinity propagation based agglomerative clustering in Figure 6.1 shows the 39 *C. jejuni* cluster into seven groups based on the pattern of presence/absence of pegs. Each of the 150 top ranked pegs are listed in the supplemental data (Table 6.2). The seven clusters generally correspond with the agricultural associated clonal complexes and the two wild-bird hosts. Cluster 1 contains the Australian purple swamphen isolates (which consists of the same ST and a single loci variant) . Cluster 2 contains only CC21. Cluster 3 contains CC48 and two ST-50 isolates from

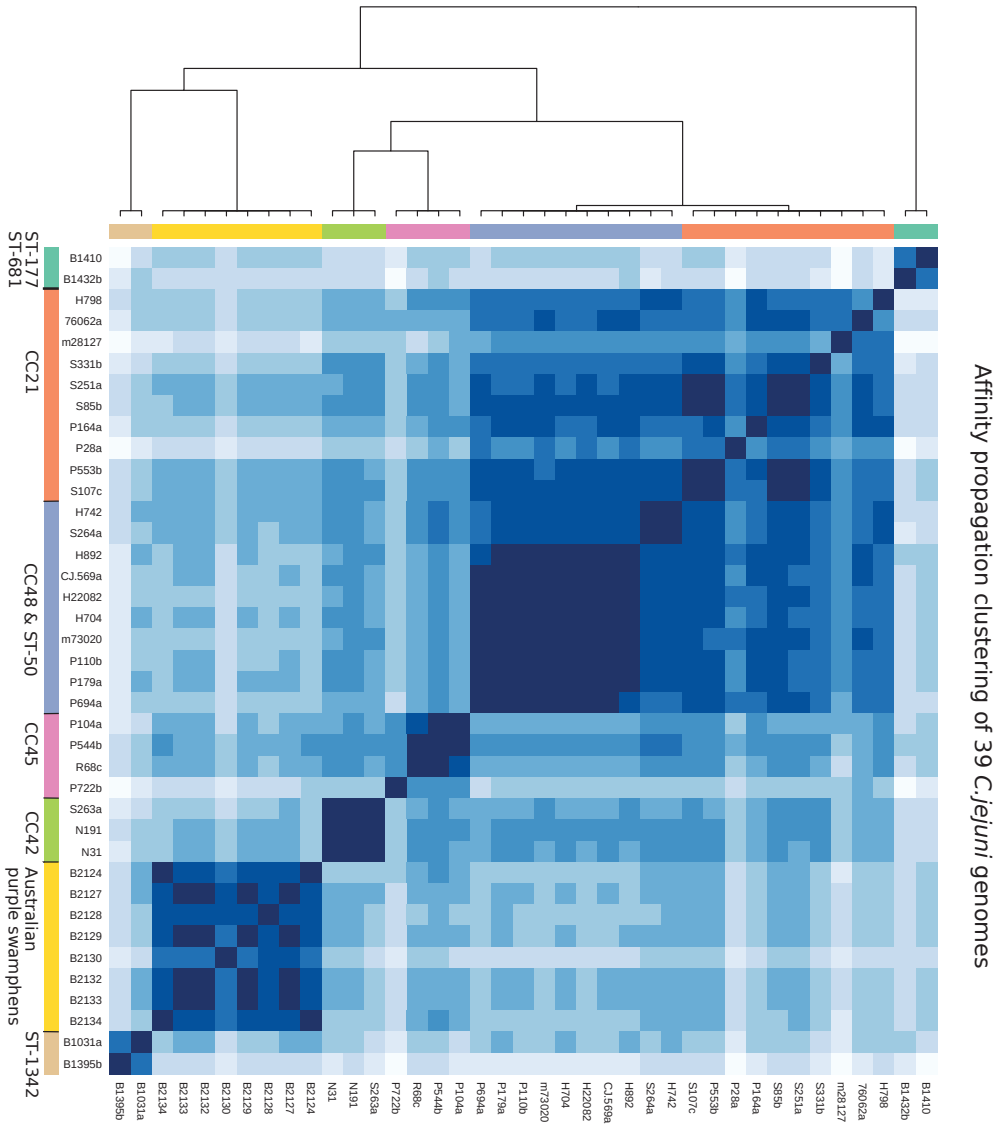


Figure 6.1: **Affinity propagation based agglomerative clustering of 39 *C. jejuni* genomes annotated by RAST functional subsystems.** The isolates cluster into seven groups that generally align with sequence type/clonal complex or wild-bird host. Each cluster is identified by a colour in the sidebar. The heatmap has nine shades of blue, the darker the shade of blue the more similar the isolates are based on the presence or absence of 1157 RAST [531][532] defined pegs.

CC21. Cluster 4 contains ST-45. Cluster 5 contains ST-42. Cluster 6 contains ST-177 and ST-681, both from starlings. Cluster 7 contains ST-1342 from starlings. This result supports there being enough difference between the three host-based groups, using draft genomes annotated into RAST curated subsystems, to differentiate between them.

Binary multi-group discriminant analysis (binda)

A binary multi-group discriminant analysis, i.e. using binary predictors [587] was performed, on a presence absence matrix of 1157 pegs from 39 *C. jejuni* isolates, and each peg was ranked on its t-score, from largest to smallest.

The list of the top 150 ranked pegs and their RAST Category, i.e. group of related subsystems, is displayed in Table 6.2 (Supplementary data). Using only the top 150 ranked pegs for hierarchical clustering in Figure 6.2 rather than needing all 1157 subsystem associated pegs, the isolates clustered in a pattern consistent with Australian purple swamphen, starling and agricultural-associated clonal complex groupings. If only the top 50 pegs are used in the analysis, the divisions between agricultural isolate clonal complexes are lost, but the division of the three groupings (Australian purple swamphen, starlings and agricultural associated clonal complexes) maintained (not shown here).

The list of the frequency of a category's occurrence in the top 150 scoring pegs for differentiation of starling, Australian purple swamphen and major agricultural clonal complexes shows:

Iron acquisition and metabolism 30 (20.0%) includes *Campylobacter* iron metabolism subsystem
Cell Wall and Capsule 18 (12.0%)
Protein Metabolism 17 (11.3%)
Cofactors, Vitamins, Prosthetic Groups, Pigment 14 (9.3%) includes Coenzyme A Biosynthesis subsystem
Respiration 14 (9.3%)
Membrane Transport 14 (9.3%) includes T6SS subsystem
Carbohydrates 11 (7.3%)
Amino Acids and Derivatives 10 (6.6%)
Nucleosides and Nucleotides 5 (3.3%) Virulence, Disease and Defence 5 (3.3%)
Fatty Acids, Lipids, and Isoprenoid 3 (2.0%)
DNA Metabolism 2 (1.3%)
Phosphorus Metabolism 2 (1.3%)
Motility and Chemotaxis 1 (0.7%) Nitrogen Metabolism 1 (0.7%)
Regulation and Cell signalling 1 (0.7%)
RNA Metabolism 1 (0.7%)
Stress Response 1 (0.7%)

a. Coenzyme A Biosynthesis

The presence or absence of pegs involved in coenzyme A biosynthesis subsystem divided the 61 genomes into two groups. The results are shown in Table 6.1. Group "A" consists of isolates

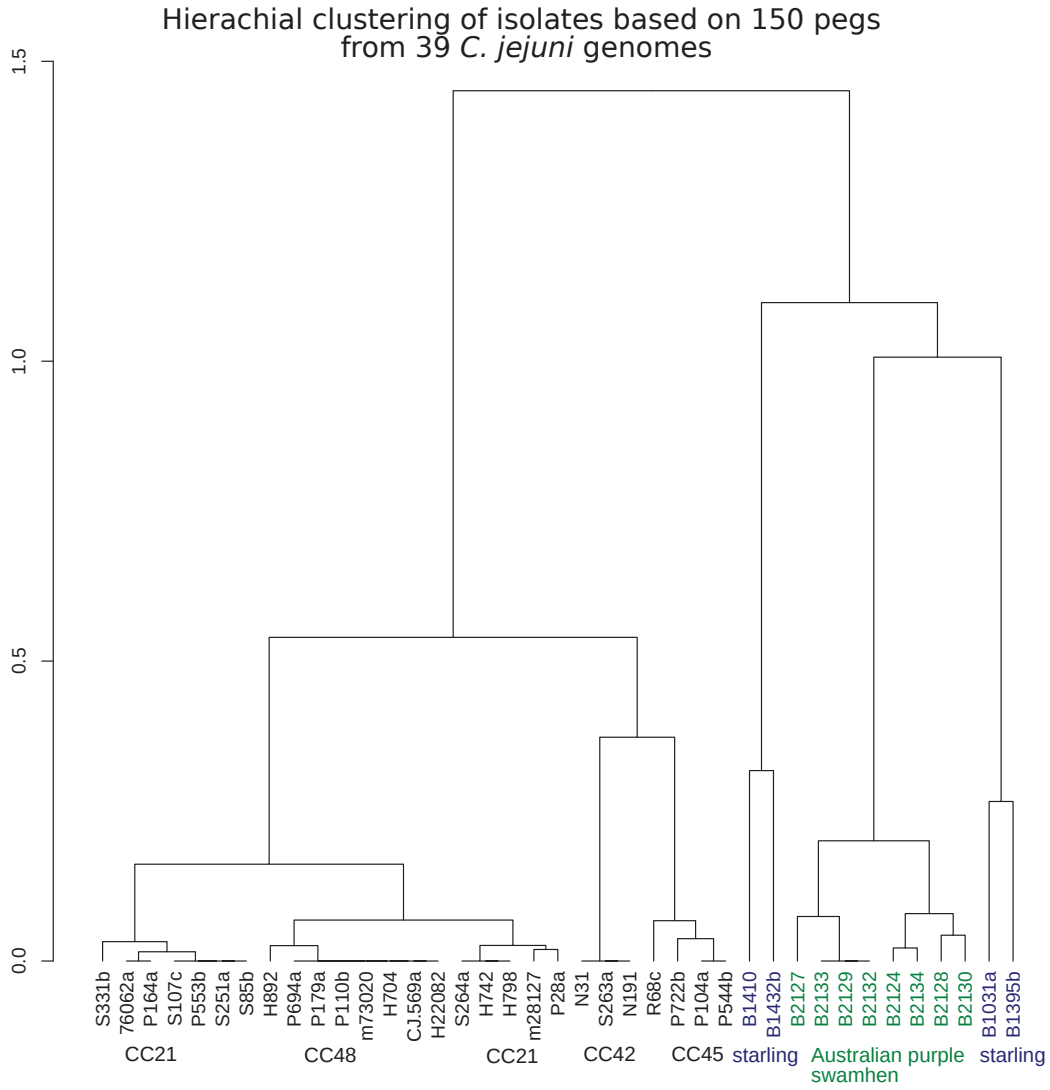


Figure 6.2: **A hierarchical clustering based on the presence or absence of 150 protein encoding genes (peg) from 39 *C. jejuni* genomes**. The top 150 ranked pegs were identified using a binary multi-group discriminant analysis. Based on 150 pegs it was possible to differentiate the genomes into the agriculture-associated clonal complex (CC21, CC48, CC42, CC45) and the two wild-bird species; starling and Australian purple swamphen.

from wild-birds (species unknown), starlings, Australian purple swamphen (*C. jejuni* and putative new species *C.* sp. nov. 1), chickens and water. Member of group A have six pegs containing : dephospho-CoA kinase (EC 2.7.1.24), pantothenate kinase type III coaX-like (EC 2.7.1.33), ketol-acid reductoisomerase (EC 1.1.1.86), Phosphopantetheine adenylyltransferase (EC 2.7.7.3), Phosphopantothenoylcysteine synthetase (EC 6.3.2.5) and phosphopantothenoylcysteine decarboxylase (EC 4.1.1.36). Member of group B have nine pegs, the same six as group A and three more involved in *de novo* pantothenate (vitamin B_5) biosynthesis: 3-methyl–2-oxobutanoate hydrox-

ymethyltransferase (EC 2.1.2.11), aspartate 1-decarboxylase (EC 4.1.1.11), pantoate-beta-alanine ligase (EC 6.3.2.1). Group B includes human, cattle, chicken and water *C. jejuni* isolates and two Australian purple swamphen isolates from a putative new species *C.* sp. nov. 4. The relationship of the genes as enzymes to each other for Coenzyme A biosynthesis is shown in Figure 6.4 (Supplementary data). There are no *C. jejuni* wild-bird isolates with a nine peg Coenzyme A biosynthesis subsystem and no isolates from human cases or cattle with only a six peg Coenzyme A biosynthesis subsystem. The chicken isolates were split with 4/9 (44%) of isolates not having the pantothenate (vitamin B_5) biosynthesis pathway (Table 6.1). Based on $\tilde{\chi}^2$ the presence of the pantothenate (vitamin B_5) biosynthesis pathway in the cattle *C. jejuni* isolates is significant ($p< 0.05$) and the absence in *C. jejuni* from wild-birds is significant ($p< 0.05$).

b. Type VI secretion system (T6SS)

The 13 conserved T6SS genes [577] were only found present in 20 of the 61 genomes: all 14 Australian purple swamphen isolates (eight *C. jejuni* and six non-*C. jejuni*), three ST-2381 isolates, a ST-3655 isolate and two ST-2026 isolates (CC403). Figure 6.3 shows a Neighbor-Net of the concatenated sequences of 12/13 genes, *Vgr*G is not included. The network splits *C. jejuni* and non-*C. jejuni*(*C.* sp. nov. 1& 4) (Figure 6.3). There is also a pattern of splitting within the *C. jejuni* forming three groups: the two ST-2026 isolates (S22b, m28548) which belong to CC403, the three ST-2381 isolates (W83a, W194ba, N3da) and the 8 Australian purple swamphen isolates (B2124, B2127, B2128, B2129, B2130, B2132, B2133, B2134) grouping with ST-3655 (W135a).

The supplemental data has phylogenetic trees for individual gene: *Clp*B Figure 6.5, *Hcp* Figure 6.6, *Icm*F Figure 6.7, *Imp*K Figure 6.8, *Imp*A Figure 6.9, *Imp*B Figure 6.10, *Imp*C Figure 6.11, *Imp*G Figure 6.12, *Imp*H Figure 6.13, *Imp*J Figure 6.14, *Pvc*109 Figure 6.15, *Vas*D Figure 6.16, and *Vgr*G Figure 6.17. The Neighbor-Net of each individual gene (*Imp*A, *Imp*B, *Imp*C, *Pvc*109, *Imp*G, *Imp*H, *Vas*D, *Imp*J, *Imp*K, *Clp*B, *Icm*F, *Vgr*G, and *Hcp*) show a similar relationship between the isolates as seen in Figure 6.3, except for *Vgr*G gene. The phylogenetic relationship between the isolates shown by the *Vgr*G loci is not consistent with the one shown in the other 12 genes (see supplement data). The *Vgr*G gene was duplicated in 80% (16/20) of the draft genomes and the sequences when compared showed variation in both size and sequence.

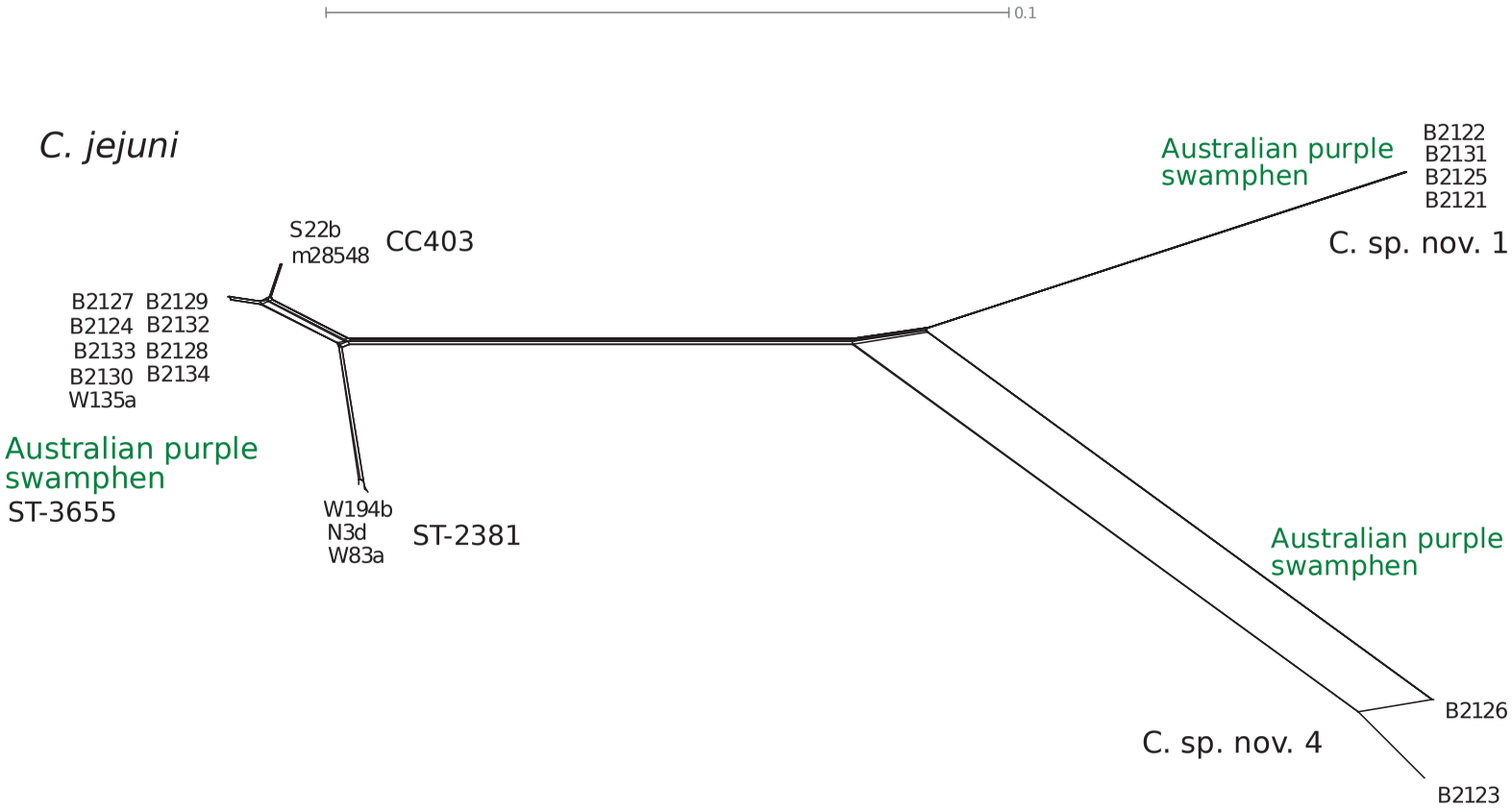


Figure 6.3: **A Neighbor-Net of the concatenated sequences of 12 core genes of the Type VI secretion system found in 20 *Campylobacter* spp. isolates.** The 12 core genes included in this Neighbor-Net are: *Imp*A, *Imp*B, *Imp*C, *Pvc*109, *Imp*G, *Imp*H, *Vas*D, *Imp*J, *Imp*K, *Clp*B, *Icm*F, and *Hcp*

c. *Campylobacter* iron metabolism

The list of top 150 ranked pegs (Table 6.2) contains 30 pegs in the category Iron acquisition and metabolism but the *Campylobacter* iron metabolism subsystem investigated here is a subset of this and represented by 26 pegs.

Using kmeans the isolates clustered into four groups based on the presence or absence of pegs in the *Campylobacter* spp. iron metabolism subsystems, and the results were placed in Table 6.1. The genomes clustered into two small and two large. The two small groups were: group A four isolates including two from CC403, and group D three isolates from wild-birds, two from starlings (CC177, CC682). The two large groups included a range of hosts both wild-bird and agriculture-associated: group B contained 32 isolates including the Australian purple swamphen *C. jejuni* isolates and CC21, CC42, CC45 but not CC48, group C contained 16 isolates including CC48, CC21, CC61 and CC257. The difference between group B and C was group C lacked a peg for a high-affinity Fe_2^+/Pb_2^+ permease precursor.

Table 6.1: **Characteristics of 61 *Campylobacter* spp. isolates in the three subsystems.** The clonal complex (CC) and sequence type (ST) were reported if known."$VitB_5$" shows the division for an isolate in the Coenzyme A biosynthesis subsystem either six protein encoding genes (pegs) in group A or nine pegs in group B. "T6SS" is the Type VI secretion system and isolates lacking this subsystem are in group A, while it is present in group B. "Iron" represents the *Campylobacter* iron metabolism subsystem, isolates cluster into group A, B, C, or D.

isolate	source	species	ST	clonal complex	VitB5	T6SS	iron
N191	cattle	*C. jejuni*	42	CC42	B	A	B
S107c	cattle	*C. jejuni*	53	CC21	B	A	B
S22b	cattle	*C. jejuni*	2026	CC403	B	B	A
S251a	cattle	*C. jejuni*	53	CC21	B	A	B
S263a	cattle	*C. jejuni*	42	CC42	B	A	B
S264a	cattle	*C. jejuni*	50	CC21	B	A	B
S331b	cattle	*C. jejuni*	21	CC21	B	A	B
S85b	cattle	*C. jejuni*	53	CC21	B	A	B
P110b	chicken	*C. jejuni*	474	CC48	B	A	C
P179a	chicken	*C. jejuni*	474	CC48	B	A	C
P28a	chicken	*C. jejuni*	520	CC21	B	A	B
P553b	chicken	*C. jejuni*	53	CC21	B	A	B
P694a	chicken	*C. jejuni*	474	CC48	B	A	C
m28127	human	*C. jejuni*	520	CC21	B	A	B
m28548	human	*C. jejuni*	2026	CC403	B	B	A
m73020	human	*C. jejuni*	474	CC48	B	A	C
569a	human	*C. jejuni*	474	CC48	B	A	C
76062a	human	*C. jejuni*	190	CC21	B	A	C
H22082	human	*C. jejuni*	474	CC48	B	A	C
H704	human	*C. jejuni*	474	CC48	B	A	C
H742	human	*C. jejuni*	50	CC21	B	A	B
H773	human	*C. jejuni*	3711	CC257	B	A	C
H798	human	*C. jejuni*	50	CC21	B	A	B
H892	human	*C. jejuni*	48	CC48	B	A	C
M880a	sheep	*C. jejuni*	2341	CC61	B	A	C
B2123	swamphen	*C.* sp. nov. 4	-	-	B	B	-
B2126	swamphen	*C.* sp. nov. 4	-	-	B	B	-
N31	water	*C. jejuni*	42	CC42	B	A	B

Continued on next page

Characteristics of Isolates —– *continued from previous page*

isolate	source	species	ST	clonal complex	VitB5	T6SS	iron
P104a	chicken	*C. jejuni*	45	CC45	A	A	B
P164a	chicken	*C. jejuni*	190	CC21	A	A	C
P544b	chicken	*C. jejuni*	45	CC45	A	A	B
P722b	chicken	*C. jejuni*	25	CC45	A	A	B
B1031a	starling	*C. jejuni*	1342	-	A	A	C
B1395b	starling	*C. jejuni*	1342	-	A	A	B
B1410	starling	*C. jejuni*	177	CC177	A	A	D
B1432b	starling	*C. jejuni*	681	CC682	A	A	D
B2121	swamphen	*C.* sp. nov. 1	-	-	A	B	-
B2122	swamphen	*C.* sp. nov. 1	-	-	A	B	-
B2124	swamphen	*C. jejuni*	-	-	A	B	B
B2125	swamphen	*C.* sp. nov. 1	-	-	A	B	-
B2127	swamphen	*C. jejuni*	-	-	A	B	B
B2128	swamphen	*C. jejuni*	-	-	A	B	B
B2129	swamphen	*C. jejuni*	-	-	A	B	B
B2130	swamphen	*C. jejuni*	-	-	A	B	B
B2131	swamphen	*C.* sp. nov. 1	-	-	A	B	-
B2132	swamphen	*C. jejuni*	-	-	A	B	B
B2133	swamphen	*C. jejuni*	-	-	A	B	B
B2134	swamphen	*C. jejuni*	-	-	A	B	B
N27	water	*C. jejuni*	3845	-	A	A	C
N3d	water	*C. jejuni*	2381	-	A	B	B
N53	water	*C. jejuni*	3845	-	A	A	C
W120a	water	*C. jejuni*	3673	-	A	A	A
W135a	water	*C. jejuni*	3655	-	A	B	B
W194b	water	*C. jejuni*	2381	-	A	B	B
W63b	water	*C. jejuni*	3663	-	A	A	A
W83a	water	*C. jejuni*	2381	-	A	B	B
R31f	wild bird	*C. jejuni*	2536	-	A	A	B
R42b	wild bird	*C. jejuni*	2539	CC177	A	A	D
R68c	wild bird	*C. jejuni*	45	CC45	A	A	B
R75a	wild bird	*C. jejuni*	1324	-	A	A	C
R52c	wild bird	*C. jejuni*	2536	-	A	A	B

6.4 Discussion

These results demonstrate that there are identifiable genomic differences, at a metabolic pathway level, between *C.jejuni* found in the starling (*Sturnus vulgaris*), the Australian purple swamphen (*Porphyrio porphyrio melanotus*), and isolates from the agriculture-associated clonal complexes (CC21, CC42, CC45, CC48). There was a tendency for agricultural-associated *C.jejuni* clonal complexes, particularly isolates from cattle, to vary from the wild-bird associated isolates based on the ability to synthesis vitamin B_5, which is a part of the Coenzyme A biosynthesis subsystem. The Australian purple swamphen isolates tended to have present the 13 central genes for the Type VI secretion system (T6SS), while the other isolates did not. While the *Campylobacter* iron metabolism did cluster all the Australian purple swamphen *C.jejuni* isolates together, the starling and the agriculture-associated clonal complexes were spread across the four groups.

There are several interesting traits suggested by the comparative phylogeny of the T6SS sequences in this dataset (Figure 6.3). The first trait is the grouping at the species level (*Campylobacter* spp.) and the clustering of isolates within *C. jejuni* into the same groups. Figures 6.5, 6.6, 6.7, 6.8, 6.9, 6.10, 6.11, 6.12, 6.13, 6.14, 6.15, 6.16 show this same pattern persists in 12 of the 13 genes (not *Vgr*G). The 13 core T6SS genes are thought to act as a single functional unit [590]. Horizontal gene transfer (HGT) is believed to be the main method of transmission for T6SS [581], and this pattern is consistent with 12 of the loci being transmitted as a single unit rather than as individual genes. This pattern of relationship could be the consequence of the T6SS sequences having been inserted in a *Campylobacter* spp. ancestor prior to the speciation event that split *C. jejuni*, *C.* sp. nov. 1 and 4. It is intriguing that *Vgr*G is the only loci showing variation from this pattern, and it does so within each *Campylobacter* spp. group. A possible reason is the *Vgr*G loci is not under the same purifying selection as the 12 other core subunits, producing loci duplication and polymorphism.

The second trait of note is the consistent pattern of splitting by 12 of the core T6SS genes (not including *Vgr*G) within *C. jejuni*, as the New Zealand ST-3655 isolate (W135a) grouped with the Australian purple swamphen isolates, while the three other New Zealand ST-2381 isolates (W194b, W63b, W83a) grouped separately. Four of the isolates ST-2381 (W194b, W63b, W83a) and ST-3655 (W135a) were collected from New Zealand water samples. Both ST-2381 and ST-3655 have been isolated from pūkeko (*Porphyrio porphyrio melantotus*)[16][p228], while ST-2381 has also been isolated from takahē *Porphyrio hochstetteri* (Dr Jonathan Marshall, personal communication, January 2015) The pūkeko is a sub-population of *Porphyrio porphyrio melantotus*, and closely related to the Australian purple swamphen, with the ancestor of the pūkeko believed to have arrived in New Zealand between 400–600 years ago [12][pp 590–595] [514][pp 38] [11]. The South Island takahē (*Porphyrio hochstetteri*) is closely related to both birds [13] [14], and a much earlier arrival

to New Zealand, with fossils dating back ∼38,000 years found in multiple sites around the South Island [514][p38]. The close relationship, in the 12 core T6SS genes, between the ST-3655 isolate (W135a) and the eight Australian purple swamphen isolates is consistent with the pūkeko recent geographical separation from the Australian population. The marked and consistent difference in the ST-2381 isolates, to the eight Australian purple swamphen isolates, may have a more engaging explanation. It is feasible the ST-2381 isolates originated from another New Zealand rail, the takahē which has been separated from the common *Porphyrio* spp. ancestor for a long period, during which time the T6SS genes in ST-2381 accumulated more differences. The CC403 isolates are part of agriculture associated clonal complex, and in the 12 core T6SS genes they tend to be closer to the ST-3655 and Australian purple swamphen isolates than the ST-2381 isolates. As CC403 has a markedly different host, it would suggest the amount of separation seen in the ST-2381 isolates is important. Alternatively, it is possible the ST-2381 isolates may occupy a different niche within the bird gastrointestinal system and this may be a barrier to genetic exchange and create the pattern of difference [421].

The presence of a functional T6SS in *C. jejuni* has been associated with pathogenicity and niche-adaptation [579]. T6SS is often found in the Proteobacteria, but is less reported in the Epsilonproteobacteria [581]. Previously reported prevalences in *C. jejuni* were between $10 - 14\%$ [577][580], which is lower than the 25.4% (14/55) found in this dataset. 12 of the 14 *C. jejuni* isolates found carrying T6SS were associated with wild-bird hosts, suggesting other wild-bird isolates should be investigated for T6SS. It has been suggested that even if wild-bird isolates are not pathogenic themselves, they could act as a source of genetic material (genetic reservoirs) for the agriculture-associated isolates [17], therefore it is important to identify if wild-bird *C. jejuni* do carry potential pathogenic features like T6SS or antibiotic resistant genes, and if there is gene flow to agricultural associated *C. jejuni*.

Coenzyme A (CoA) is an essential cofactor in numerous metabolic and energy-yielding reactions and is involved in the regulation of key metabolic enzymes. The pantothenate biosynthesis pathway, a part of the Coenzyme A biosynthesis subsystem (Figure 6.4) and has been previously reported as a possible factor in host adaptation in *C. jejuni* due to its tendency to be absent in non-agriculture-associated isolates [502][p93] and present in cattle isolates [370]. The lack of *de-novo* pantothenate (vitamin B_5) biosynthesis in pathogenic bacteria has been previously identified in some pathogenic bacteria resulting in the suggestion that they rely on scavenging exogenous pantothenate [583]. Our results, showed as significant, that all the *C. jejuni* isolates from cattle had the pantothenate biosynthesis pathway (9 pegs in the Coenzyme A subsystem) but none of the wild bird *C. jejuni* isolates (6 pegs in the Coenzyme A subsystem). It is curious that the extra three pegs required for pantothenate biosynthesis were found in both *C.* sp. nov. 4 isolates from the Australian purple swamphen, but not in their *C. jejuni*. A reason suggested for the pattern of Coenzyme A biosynthesis subsystem presence/absence has been the diet of the host [370], but this indicates

that host diet may not be the sole reason for pantothenate biosynthesis.

Iron acquisition is essential for *C. jejuni* and associated with host colonisation [469]. It is interesting that a metabolic feature so integral to survival can show marked variation based on the presence or absence of genes from a subsystem, between sequence types, rather than variation limited to differences in alleles. However strain variation, within *C. jejuni*, related to iron metabolism has been previously reported [288], suggesting it is an important area of variation. A possible reason that *Campylobacter* spp. iron metabolism subsystem was identified as important in the differentiating between the host groups is two of the four starling isolates (B1432b, B1410) were annotated as inactive subsystems due to insufficient components. However this identification by the RAST annotation as having insufficient components also occurred to several isolates from the expanded dataset. As the iron metabolism is essential for *Campylobacter* spp. [584] rather than being inactive, it is possible there was an error in the sequencing, assembly or annotation and the relevant genes were missed. Investigation into the *Campylobacter* iron metabolism of the affected starling isolates, both at the genomic and cell physiology levels may explain these events.

In the initial comparison of the three host associated groups: starlings, Australian purple swamphen and agriculture-associated clonal complexes, there were enough differences within their RAST classifications (Figure 6.1) to show a pattern of clustering based on metabolic features. The multilocus sequence typing system uses an allelic profile based on loci that are part of the core genome and this has been used for source attribution [99] [268]. The pattern of variation in the subsystems in this analysis (Figure 6.1), which represent metabolic pathways, can also be related to the ST/CC and to host. An exception in this analysis of ST clustering to clonal complex is shown by two ST-50 isolates (Figure 6.1) which clustered with CC48 and not CC21 (Table 6.1). A possible explanation is CC21 is a large and diverse generalist group and they may share many metabolic pathway genes with other non-CC21 ST like ST-50 [482]. In the previous chapter both these clonal complexes clustered near each other in the pan-genome and rMLST comparison.

It is likely that the identification of T6SS and *Campylobacter* spp. iron metabolism subsystems, in particular, were dependent on the wild-bird hosts used, and a dataset based on some other unrelated wild-bird species would possibly not produce the same result. Any answer to host adaptation in *C. jejuni* will be heavily influenced by the dataset and the type of analysis undertaken. A novel dataset and a functional subsystem approach was taken which to the best of our knowledge has not been used in an investigation of host adaptation in *C. jejuni*. In the future the investigation should be expanded to isolates from other wild-bird species and include genomic and experimental confirmation of differences that may be host adaptations.

This analysis has shown that based on the pattern of presence and absence of pegs, it is possible to differentiate between agricultural associated isolates and wild-bird isolates, even if the number of peg is reduced e.g. Figure 6.2 uses only 150 pegs. This supports the hypothesis that *C.jejuni* does

show a level of host adaptation in the accessory genome. We identified three RAST subsystems that appeared to have a significant affect in differentiating the isolates in our dataset: Coenzyme A biosynthesis, Type VI secretion system (T6SS) and the *Campylobacter* iron metabolism subsystem. The phylogenetic relationship between 12 of the core genes in T6SS showed ST-3655, a pūkeko associated isolate, is closer to Australian swamphen *C.jejuni* than ST-2381 raising the possibility that ST-2381 originated in the takahē rather than the pūkeko. (T6SS) and the *Campylobacter* iron metabolism have been reported as related to pathogenicity [586], and Coenzyme A biosynthesis has been associated with host adaptation.

Acknowledgements

Maulik Shukla, Senior Computer Scientist, PATRIC Production Lead, Cyberinfrastructure Division,Virginia Bioinformatics Institute, Virginia Tech helped with using the RAST/SEED/PATRIC system for analysis of the genomes.

6.5 Supplementary data:

Top 150 ranked protein encoding genes(peg)

Table 6.2: **The top 150 ranked protein encoding genes (pegs) showing variation between the three *C. jejuni* host groups.** The peg are listed by their role and Category within the Rapid Annotations using Subsystems Technology (RAST) and SEED servers [532] [531]. EC is the Enzyme Commission number.

Category	Role
Membrane Transport	*Clp*B protein
Membrane Transport	*Icm*F-related protein
Membrane Transport	Outer membrane protein *Imp*K/*Vas*F, *Omp*A/*Mot*B domain
Membrane Transport	Protein *Imp*G/*Vas*A
Membrane Transport	Type VI secretion lipoprotein/*Vas*D
Membrane Transport	Uncharacterised protein *Imp*A
Membrane Transport	Uncharacterised protein *Imp*B
Membrane Transport	Uncharacterised protein *Imp*C
Membrane Transport	Uncharacterised protein *Imp*H/*Vas*B
Membrane Transport	Uncharacterised protein *Imp*J/*Vas*E
Membrane Transport	Uncharacterised protein similar to VCA0109
Membrane Transport	*Vgr*G protein
Nitrogen Metabolism	Nitrate/nitrite transporter
Stress Response	transcriptional regulator, *Crp*/*Fnr* family
Respiration	Arsenate reductase (EC 1.20.4.1)
Virulence, Disease and Defence	Arsenate reductase (EC 1.20.4.1)
Virulence, Disease and Defence	Arsenic efflux pump protein
Virulence, Disease and Defence	Beta-lactamase (EC 3.5.2.6)
Carbohydrates	L-lactate dehydrogenase (EC 1.1.1.27)
Cell Wall and Capsule	UDP-N-acetylglucosamine 2-epimerase (EC 5.1.3.14)
Regulation and Cell signaling	Death on curing protein, Doc toxin
Carbohydrates	*Tcu*A
Carbohydrates	*Tcu*B
Carbohydrates	*Tcu*C
Virulence, Disease and Defence	Arsenical resistance operon repressor

Continued on next page

top 150 peg —– *continued from previous page*

Category	**Role**
Cofactors, Vitamins, Prosthetic Groups, Pigments	Thiamin ABC transporter, transmembrane component
Iron acquisition and metabolism	Ferric iron ABC transporter, ATP-binding protein
Carbohydrates	Gluconate 2-dehydrogenase (EC 1.1.99.3), membrane-bound, gamma subunit
Respiration	ATP synthase alpha chain (EC 3.6.3.14)
Respiration	ATP synthase beta chain (EC 3.6.3.14)
Respiration	ATP synthase delta chain (EC 3.6.3.14)
Respiration	ATP synthase epsilon chain (EC 3.6.3.14)
Respiration	ATP synthase F0 sector subunit a (EC 3.6.3.14)
Respiration	ATP synthase F0 sector subunit b (EC 3.6.3.14)
Respiration	ATP synthase F0 sector subunit b' (EC 3.6.3.14)
Respiration	ATP synthase gamma chain (EC 3.6.3.14)
Cell Wall and Capsule	Bacillosamine/Legionaminic acid biosynthesis aminotransferase *Pgl*E
Cell Wall and Capsule	Legionaminic acid biosynthesis protein *Ptm*A
Cell Wall and Capsule	Legionaminic acid biosynthesis protein *Ptm*F
Cell Wall and Capsule	Legionaminic acid cytidylyltransferase (EC 2.7.7.43)
Cell Wall and Capsule	Legionaminic acid synthase (EC 2.5.1.56)
Cell Wall and Capsule	UDP-N-acetylglucosamine 4,6-dehydratase (EC 4.2.1.-)
Protein Metabolism	4-keto-6-deoxy-N-Acetyl-D-hexosaminyl-(Lipid carrier) aminotransferase
Amino Acids and Derivatives	Homoserine O-succinyltransferase (EC 2.3.1.46)
Amino Acids and Derivatives	O-acetylhomoserine sulfhydrylase (EC 2.5.1.49)
Amino Acids and Derivatives	O-succinylhomoserine sulfhydrylase (EC 2.5.1.48)
Cofactors, Vitamins, Prosthetic Groups, Pigments	3-methyl-2-oxobutanoate hydroxymethyltransferase (EC 2.1.2.11)
Cofactors, Vitamins, Prosthetic Groups, Pigments	Aspartate 1-decarboxylase (EC 4.1.1.11)
Cofactors, Vitamins, Prosthetic Groups, Pigments	Pantoate–beta-alanine ligase (EC 6.3.2.1)
Cell Wall and Capsule	N-Acetylneuraminate cytidylyltransferase (EC 2.7.7.43)
Cell Wall and Capsule	Legionaminic acid biosynthesis protein *Ptm*G

Continued on next page

top 150 peg —– *continued from previous page*

Category	**Role**
Cell Wall and Capsule	N-acetylneuraminate synthase (EC 2.5.1.56)
Amino Acids and Derivatives	Delta-1-pyrroline-5-carboxylate dehydrogenase (EC 1.2.1.88)
Carbohydrates	Predicted D-lactate dehydrogenase, Fe-S protein, FAD/FMN-containing
Cell Wall and Capsule	Beta-1,4-N-acetylgalactosaminyltransferase (EC 2.4.1.-)
Amino Acids and Derivatives	Glycerate kinase (EC 2.7.1.31)
Amino Acids and Derivatives	L-serine dehydratase, alpha subunit (EC 4.3.1.17)
Amino Acids and Derivatives	L-serine dehydratase, beta subunit (EC 4.3.1.17)
Carbohydrates	Glycerate kinase (EC 2.7.1.31)
Carbohydrates	L-serine dehydratase, alpha subunit (EC 4.3.1.17)
Carbohydrates	L-serine dehydratase, beta subunit (EC 4.3.1.17)
Fatty Acids, Lipids, and Isoprenoids	Glycerate kinase (EC 2.7.1.31)
Iron acquisition and metabolism	Enterochelin uptake ATP-binding protein
Iron acquisition and metabolism	Enterochelin uptake periplasmic binding protein
Iron acquisition and metabolism	Enterochelin uptake permease *Ceu*B
Iron acquisition and metabolism	Enterochelin uptake permease *Ceu*C
Iron acquisition and metabolism	Ferric iron ABC transporter, iron-binding protein
Iron acquisition and metabolism	Ferric iron ABC transporter, permease protein
Iron acquisition and metabolism	Ferric receptor *Cfr*A
Iron acquisition and metabolism	Ferric siderophore transport system, biopolymer transport protein *Exb*B
Iron acquisition and metabolism	Ferric siderophore transport system, periplasmic binding protein *Ton*B
Iron acquisition and metabolism	Ferric uptake regulation protein
Iron acquisition and metabolism	Ferrous iron transport protein A, putative
Iron acquisition and metabolism	Ferrous iron transport protein B
Iron acquisition and metabolism	Haemin uptake system ATP-binding protein
Iron acquisition and metabolism	Haemin uptake system outer membrane receptor
Iron acquisition and metabolism	Haemin uptake system periplasmic haemin-binding protein
Iron acquisition and metabolism	Haemin uptake system permease protein
Iron acquisition and metabolism	Hemerythrin-like iron-binding protein

Continued on next page

top 150 peg —– *continued from previous page*

Category	**Role**
Iron acquisition and metabolism	Magnesium and cobalt transport protein *Cor*A
Iron acquisition and metabolism	Nonheme iron-containing ferritin
Iron acquisition and metabolism	Periplasmic protein p19 involved in high-affinity Fe_2^+ transport
Iron acquisition and metabolism	Peroxide stress regulator
Iron acquisition and metabolism	Possible bacterioferritin
Iron acquisition and metabolism	Putative heme oxygenase
Iron acquisition and metabolism	Putative iron-uptake ABC transport system ATP-binding protein
Iron acquisition and metabolism	Putative lipoprotein of ferric iron transporter system
Virulence, Disease and Defence	Filamentous haemagglutinin domain protein
Amino Acids and Derivatives	Carboxynorspermidine dehydrogenase, putative (EC 1.1.1.-)
Amino Acids and Derivatives	L-serine dehydratase (EC 4.3.1.17)
Carbohydrates	L-serine dehydratase (EC 4.3.1.17)
Cofactors, Vitamins, Prosthetic Groups, Pigments	Biotin-protein ligase (EC 6.3.4.15)
Cofactors, Vitamins, Prosthetic Groups, Pigments	Flavodoxin
Cofactors, Vitamins, Prosthetic Groups, Pigments	Molybdopterin binding motif, CinA N-terminal domain
Cofactors, Vitamins, Prosthetic Groups, Pigments	tRNA (cytidine(34)-2'-O)-methyltransferase (EC 2.1.1.207)
Nucleosides and Nucleotides	Cytosine/purine/uracil/thiamine/allantoin permease family protein
Respiration	Type *cbb*3 cytochrome oxidase biogenesis protein CcoG, involved in Cu oxidation
Fatty Acids, Lipids, and Isoprenoids	3-oxoacyl-[acyl-carrier-protein] synthase, KASIII (EC 2.3.1.41)
Membrane Transport	Oligopeptide transport ATP-binding protein *Opp*D (TC 3.A.1.5.1)
Cofactors, Vitamins, Prosthetic Groups, Pigments	Long-chain-fatty-acid–CoA ligase (EC 6.2.1.3)
Carbohydrates	Aldehyde dehydrogenase A (EC 1.2.1.22)

Continued on next page

top 150 peg —– *continued from previous page*

Category	**Role**
Iron acquisition and metabolism	Ferric siderophore transport system, biopolymer transport protein ExbD
Iron acquisition and metabolism	Putative lipoprotein, similar to CjrA of Escherichia coli O164
Iron acquisition and metabolism	Putative outer membrane siderophore receptor
Membrane Transport	Ferric siderophore transport system, biopolymer transport protein *Exb*D
Cell Wall and Capsule	Adenylylsulfate kinase (EC 2.7.1.25)
Cell Wall and Capsule	Putative amidotransferase (Type 1 glutamine amidotransferase - GATase1)
Cell Wall and Capsule	Putative sugar nucleotidyltransferase
Cell Wall and Capsule	Pyruvate phosphate dikinase
DNA Metabolism	Type III restriction-modification system methylation subunit (EC 2.1.1.72)
Cofactors, Vitamins, Prosthetic Groups, Pigments	Menaquinone via futalosine step 1
Cofactors, Vitamins, Prosthetic Groups, Pigments	Ubiquinone/menaquinone biosynthesis methyltransferase UbiE (EC 2.1.1.-)
Amino Acids and Derivatives	N-acetyl-L,L-diaminopimelate deacetylase (EC 3.5.1.47)
Cell Wall and Capsule	D-glycero-D-manno-heptose 1-phosphate guanosyltransferase
DNA Metabolism	Type III restriction-modification system DNA endonuclease res (EC 3.1.21.5)
Cell Wall and Capsule	dTDP-glucose 4,6-dehydratase (EC 4.2.1.46)
Cell Wall and Capsule	UDP-glucose 4-epimerase (EC 5.1.3.2)
Cofactors, Vitamins, Prosthetic Groups, Pigments	Gene SCO4494, often clustered with other genes in menaquinone via futalosine pathway
Cofactors, Vitamins, Prosthetic Groups, Pigments	Menaquinone via futalosine step 3
Cofactors, Vitamins, Prosthetic Groups, Pigments	Menaquinone via futalosine step 4
Iron acquisition and metabolism	Fe_2^+ ABC transporter, permease protein 2
RNA Metabolism	Permease of the drug/metabolite transporter (DMT) superfamily

Continued on next page

top 150 peg —– *continued from previous page*

Category	**Role**
Phosphorus Metabolism	Phosphoenolpyruvate phosphomutase (EC 5.4.2.9)
Phosphorus Metabolism	Phosphonopyruvate decarboxylase (EC 4.1.1.82)
Fatty Acids, Lipids, and Isoprenoids	Alcohol dehydrogenase (EC 1.1.1.1)
Motility and Chemotaxis	Flagellar hook protein FlgE
Nucleosides and Nucleotides	CTP synthase (EC 6.3.4.2)
Nucleosides and Nucleotides	Deoxycytidine triphosphate deaminase (EC 3.5.4.13)
Nucleosides and Nucleotides	Thioredoxin reductase (EC 1.8.1.9)
Nucleosides and Nucleotides	Thymidylate kinase (EC 2.7.4.9)
Protein Metabolism	Apolipoprotein N-acyltransferase (EC 2.3.1.-)
Protein Metabolism	GTP-binding and nucleic acid-binding protein *Ych*F
Protein Metabolism	GTP-binding protein *Eng*A
Protein Metabolism	GTP-binding protein *Eng*B
Protein Metabolism	GTP-binding protein *Era*
Protein Metabolism	GTP-binding protein *Obg*
Protein Metabolism	GTP-binding protein *Typ*A/*Bip*A
Protein Metabolism	GTPase and tRNA-U34 5-formylation enzyme *Trm*E
Protein Metabolism	Methionyl-tRNA formyltransferase (EC 2.1.2.9)
Protein Metabolism	Periplasmic thiol
Protein Metabolism	Ribosome-binding factor A
Protein Metabolism	Signal recognition particle receptor protein *Fts*Y (=alpha subunit) (TC 3.A.5.1.1)
Protein Metabolism	Signal recognition particle, subunit *Ffh* SRP54 (TC 3.A.5.1.1)
Protein Metabolism	Translation initiation factor 1
Protein Metabolism	Translation initiation factor 2
Protein Metabolism	Translation initiation factor 3
Respiration	hydrogenase, (NiFe)/(NiFeSe) small subunit family
Respiration	Protein *hyd*E
Respiration	Quinone-reactive Ni/Fe hydrogenase, cytochrome b subunit
Respiration	Quinone-reactive Ni/Fe-hydrogenase large chain (EC 1.12.5.1)

Coenzyme A Biosynthesis pathway

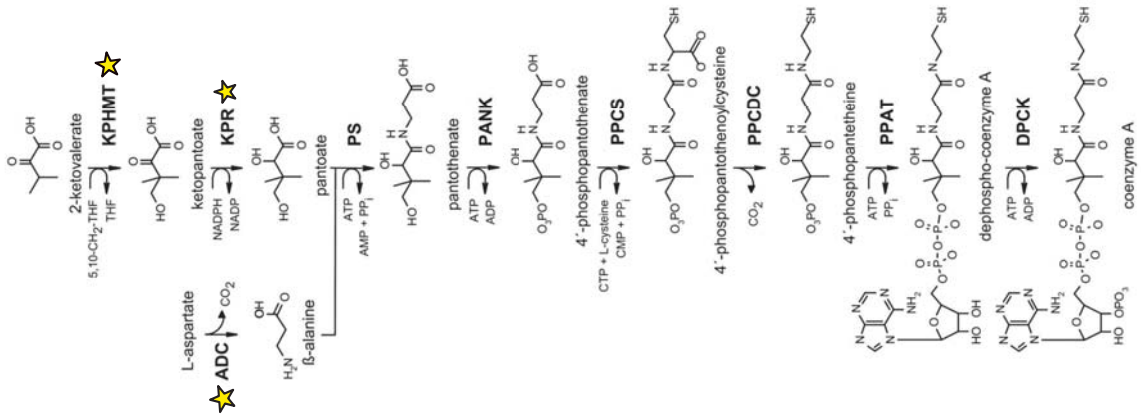


Figure 6.4: **The bacterial biosynthesis of Coenzyme A pathway**. The yellow stars denote the three *de-novo* steps that some of the *C.jejuni* lack. In bacteria, *de novo* biosynthesis of CoA comprises of nine steps in total. β-alanine is produced from aspartate by the action of aspartate a-decarboxylase (EC 4.1.1.11; ADC). Keto-pantoate hydroxymethyltransferase (EC 2.1.2.11; KPHMT) converts a-ketoisovalerate to ketopantoate, which is reduced to pantoate by ketopantoate reductase (EC 1.1.1.169; KPR). Pantothenate synthetase (EC 6.3.2.1; PS) catalyses the condensation of pantoate and β-alanine to pantothenate. Phosphorylation of pantothenate is catalysed by pantothenate kinase (EC 2.7.1.33; PANK) which gives 49- phosphopantothenate. 49- phosphopantothenate is condensed with cysteine by phosphopantothenoylcysteine synthetase (EC 6.3.2.5; PPCS) to yield 49-phospho-N-pantothenoylcysteine. Next, 49-phosphopantetheine is produced by the action of phosphopantothenoylcysteine decarboxylase (EC 4.1.1.36; PPCDC). An adenylyl-group is transfered to 49- phosphopantetheine and catalysed by phosphopantetheine adenylyltransferase (EC 2.7.7.3; PPAT). Finally, dephospho-CoA is phosphorylated by dephospho-coenzyme A kinase (EC 2.7.1.24; DPCK) resulting in Coenzyme A. The Figure is attributted to Genschel (2004)[583].

Type VI secretion system (T6SS)

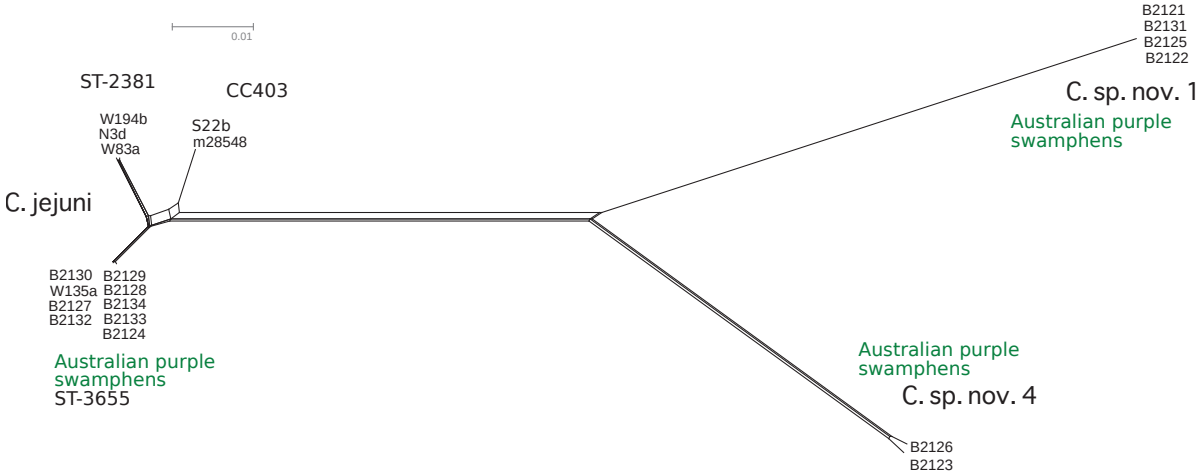


Figure 6.5: A Neighbor-Net [591] of the *Clp*B gene of the Type VI secretion system found in 20 *Campylobacter* spp. isolates visualised in SplitsTree [487].

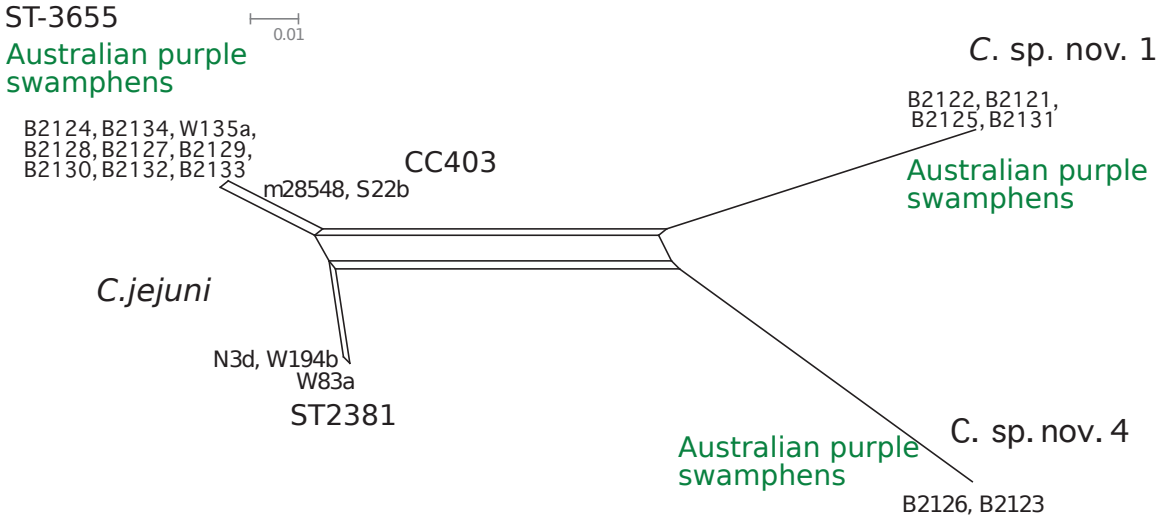


Figure 6.6: A Neighbor-Net[591] of the *hcp*A gene of the Type VI secretion system found in 20 *Campylobacter* spp. isolates visualised in SplitsTree [487].

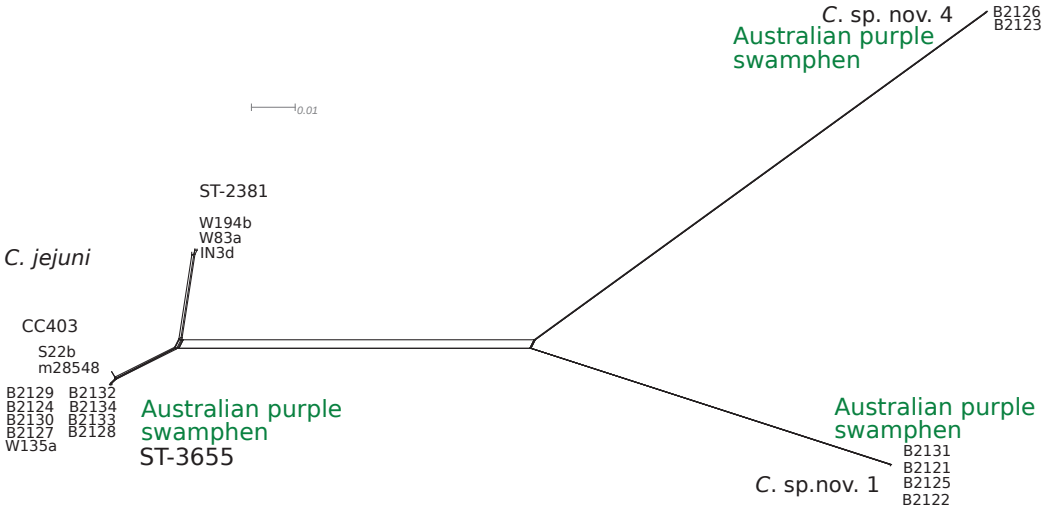


Figure 6.7: A Neighbor-Net[591] of the *Icm*F gene in the Type VI secretion system found in 20 *Campylobacter* spp. isolates visualised in SplitsTree [487].

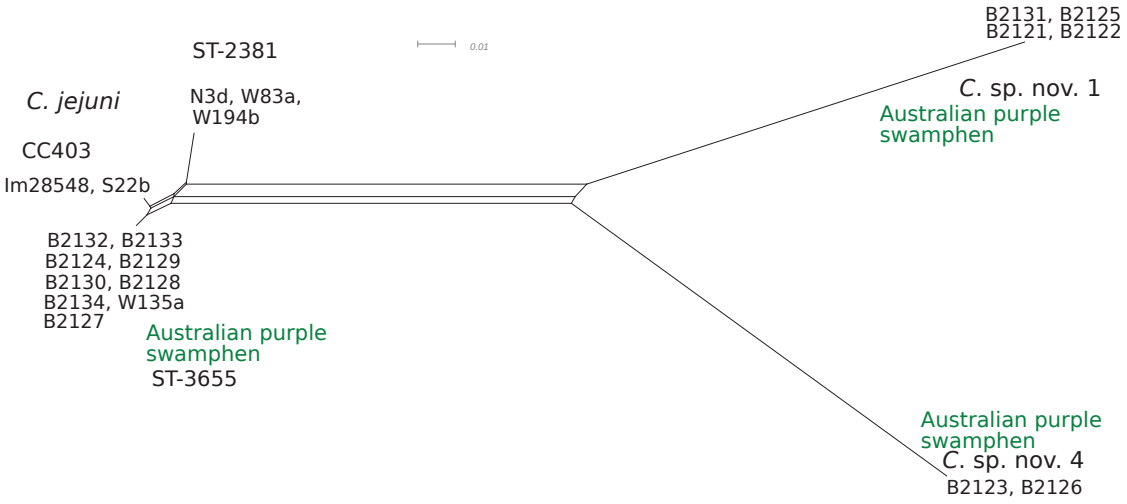


Figure 6.8: A Neighbor-Net[591] of the *ImK* gene in the Type VI secretion system found in 20 *Campylobacter* spp. isolates visualised in SplitsTree [487].

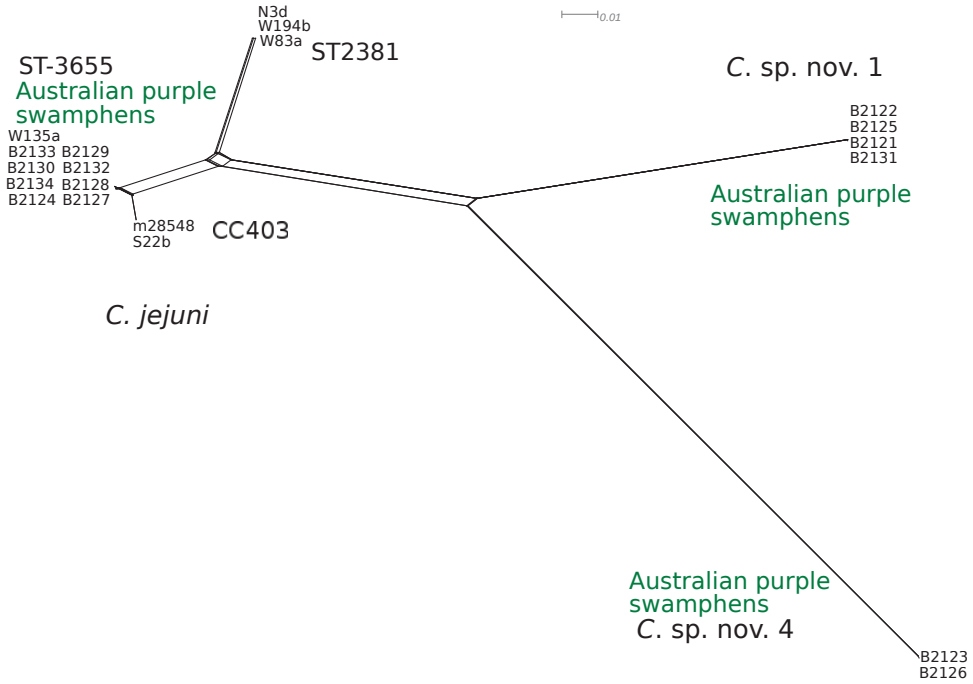


Figure 6.9: A Neighbor-Net [591] of the *Imp*A gene in the Type VI secretion system found in 20 *Campylobacter* spp. isolates visualised in SplitsTree [487].

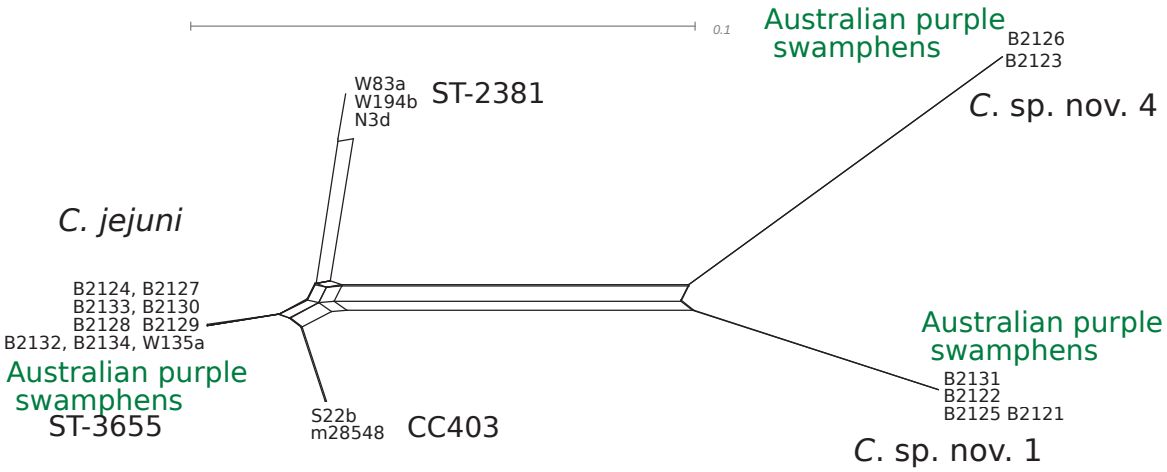


Figure 6.10: A Neighbor-Net[591] of the *Imp*B gene of the Type VI secretion system found in 20 *Campylobacter* spp. isolates visualised in SplitsTree [487].

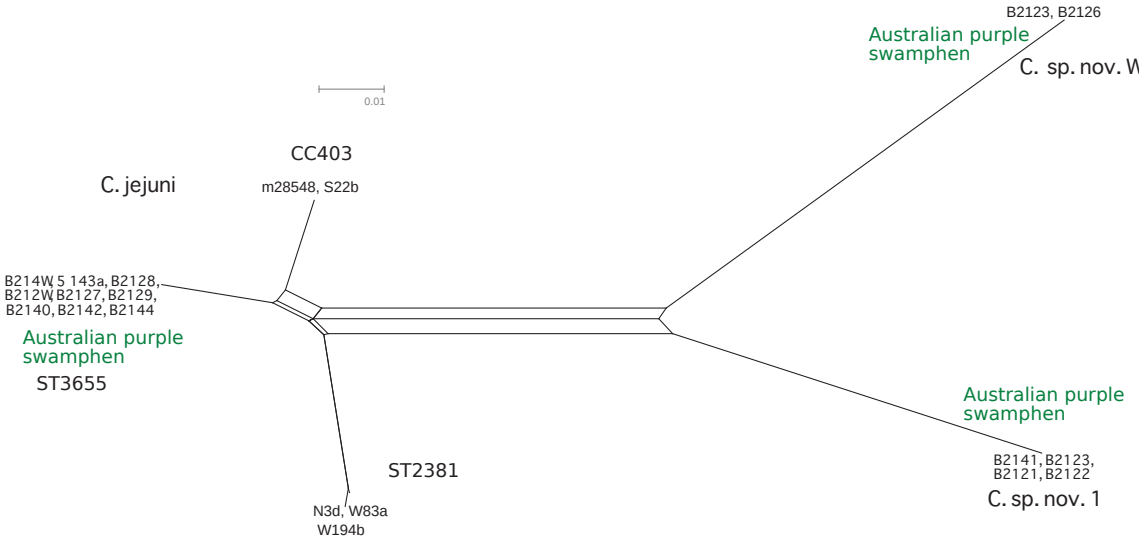


Figure 6.11: A Neighbor-Net[591] of the *Imp*C gene of the Type VI secretion system found in 20 *Campylobacter* spp. isolates visualised in SplitsTree [487].

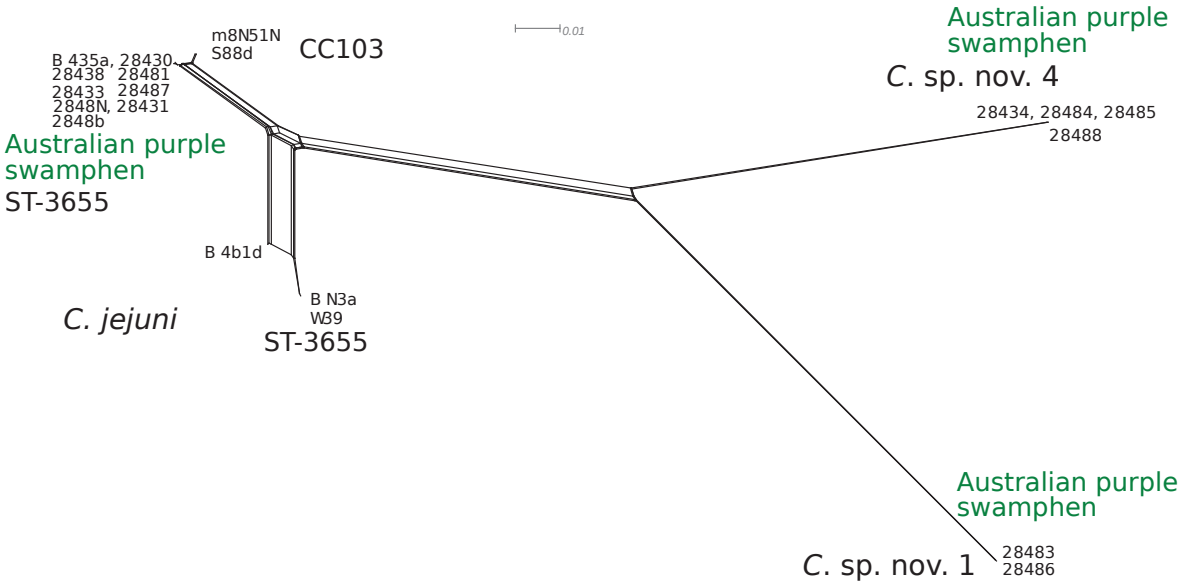


Figure 6.12: A Neighbor-Net[591] of the *Imp*G gene in the Type VI secretion system found in 20 *Campylobacter* spp. isolates visualised in SplitsTree [487].

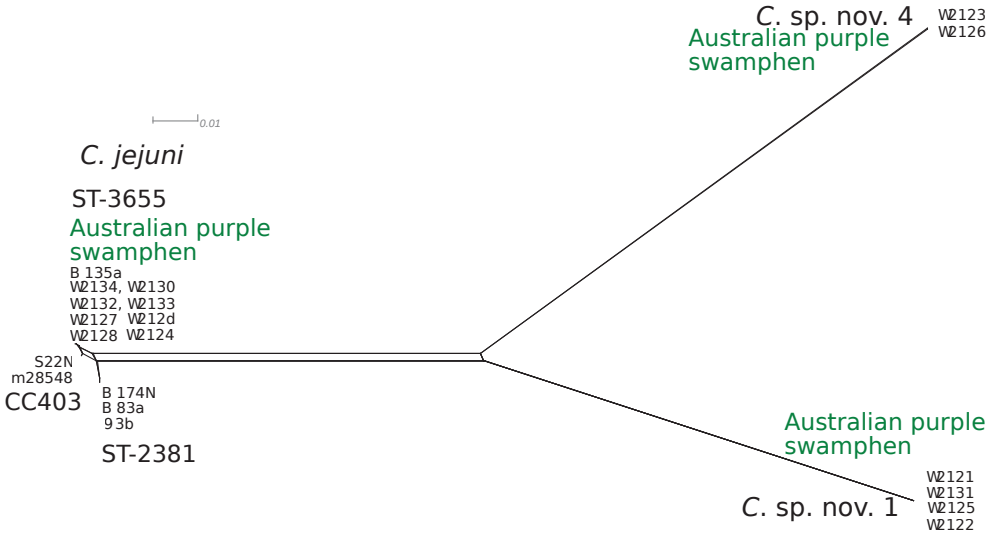


Figure 6.13: A Neighbor-Net[591] of the *Imp*H gene in the Type VI secretion system found in 20 *Campylobacter* spp. isolates visualised in SplitsTree [487].

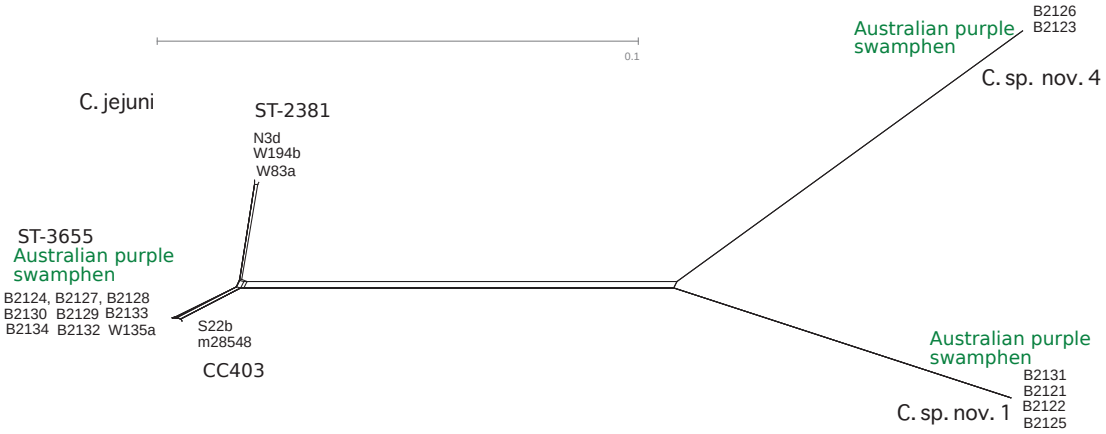


Figure 6.14: A Neighbor-Net[591] of the *Imp*J gene in the Type VI secretion system found in 20 *Campylobacter* spp. isolates visualised in SplitsTree [487].

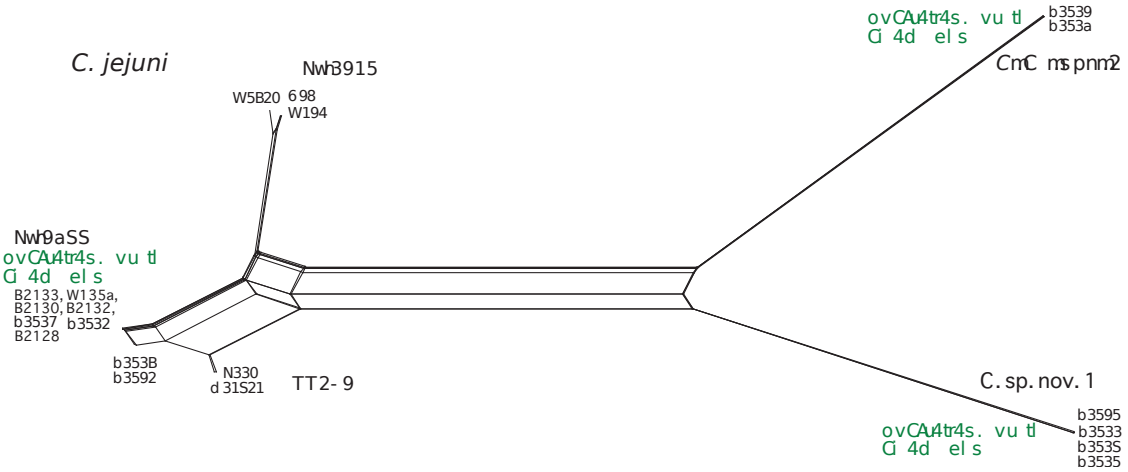


Figure 6.15: A Neighbor-Net[591] of the *Pvc*109 gene in the Type VI secretion system found in 20 *Campylobacter* spp. isolates visualised in SplitsTree [487].

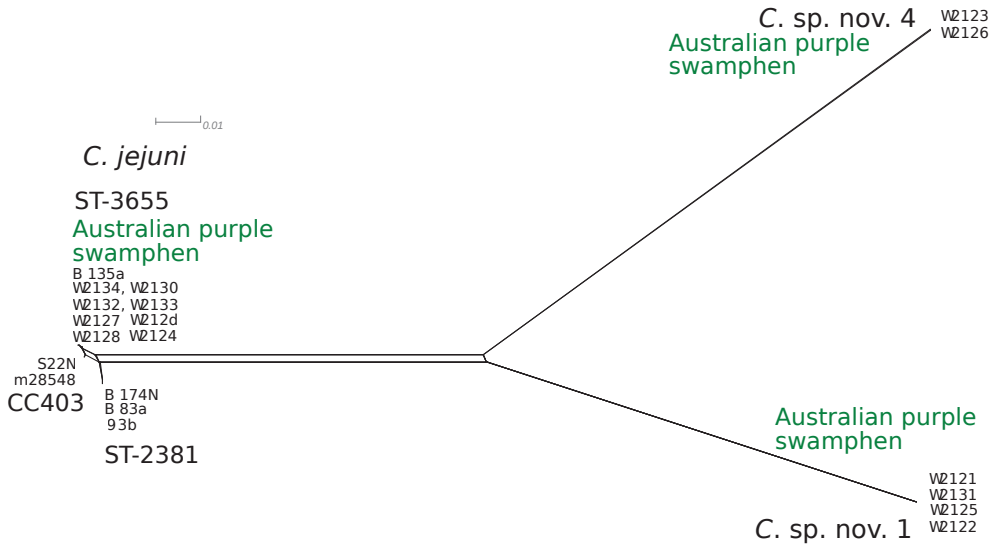


Figure 6.16: A Neighbor-Net[591] of the *Vas*D gene of the Type VI secretion system found in 20 *Campylobacter* spp. isolates visualised in SplitsTree [487].

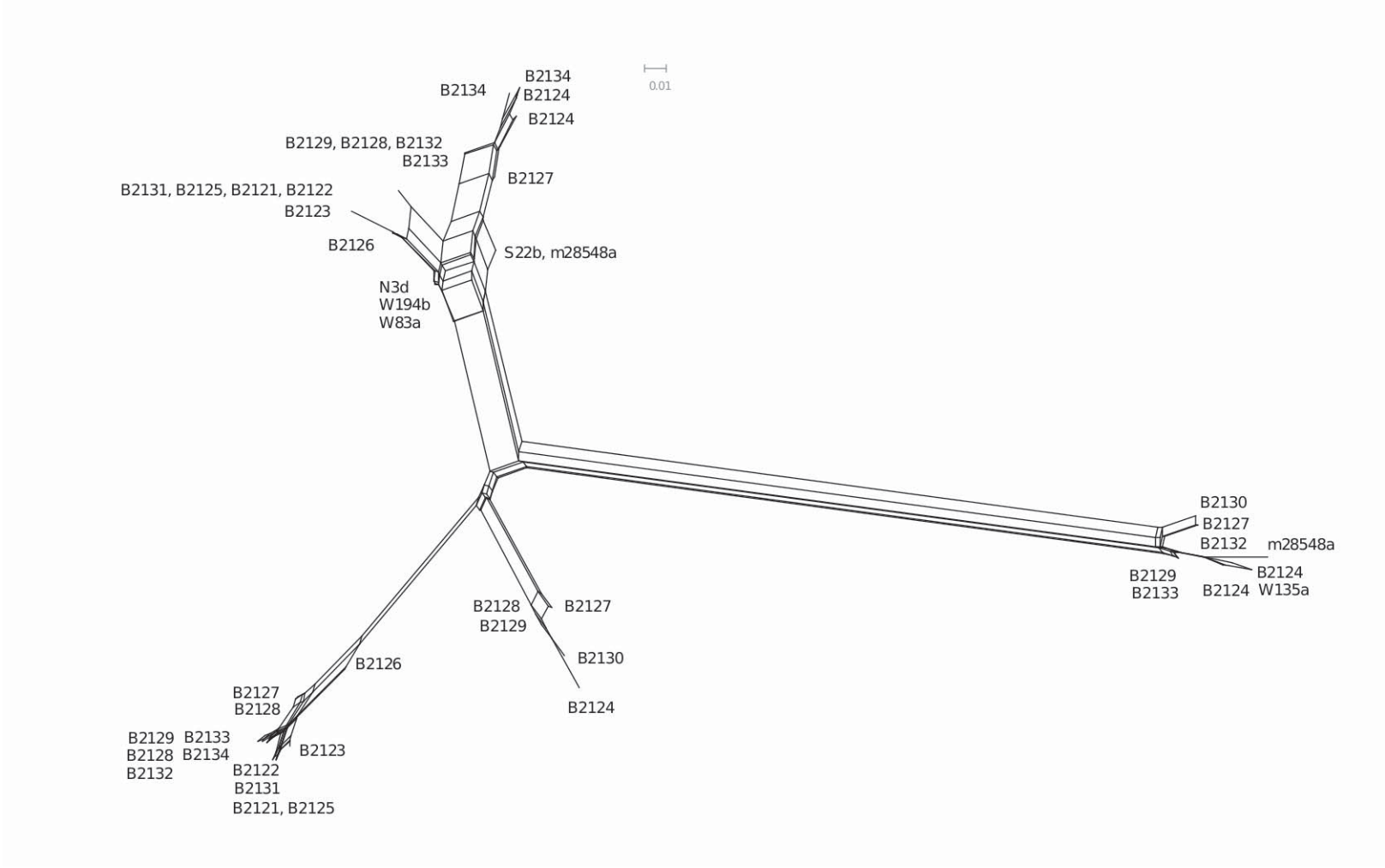


Figure 6.17: A Neighbor-Net[591] of the *Vgr* gene of the Type VI secretion system found in 20 *Campylobacter* spp. isolates visualised in SplitsTree [487]. Some of the isolates had multiple versions so they appear multiple times.

The *Campylobacter* Iron Metabolism subsystem

The roles of the 26 pegs in the *Campylobacter* Iron Metabolism subsystem:

Enterochelin uptake ATP-binding protein

Enterochelin uptake periplasmic binding protein

Enterochelin uptake permease *Ceu*B

Enterochelin uptake permease *Ceu*C

Ferric iron ABC transporter, iron-binding protein

Ferric iron ABC transporter, permease protein

Ferric receptor *Cfr*A

Ferric siderophore transport system, biopolymer transport protein *Exb*B

Ferric siderophore transport system, periplasmic binding protein *Ton*B

Ferric uptake regulation protein

Ferrous iron transport protein A, putative

Ferrous iron transport protein B

Haemin uptake system ATP-binding protein

Haemin uptake system outer membrane receptor

Haemin uptake system periplasmic haemin-binding protein

Haemin uptake system permease protein

Hemerythrin-like iron-binding protein

High-affinity Fe_2^+/Pb_2^+ permease precursor

Magnesium and cobalt transport protein *Cor*A

Nonheme iron-containing ferritin

Periplasmic protein p19 involved in high-affinity Fe_2^+ transport

Peroxide stress regulator

Possible bacterioferritin

Putative heme oxygenase

Putative iron-uptake ABC transport system ATP-binding protein

Putative lipoprotein of ferric iron transporter system

Appendix D in DVD

Summary of information provided in Appendix D

In Appendix D on the DVD provided there are 61 pie-charts showing the gene distribution by subsystem category for each isolate. The file may need to be loaded onto a computer hard drive and unzipped.

Every object that biology studies is a system of systems.

Francois Jacob (1974)

7

Conclusion

This thesis addresses the complex and enigmatic bacteria *C. jejuni*, from a multifaceted viewpoint, an approach consistent with a One Health perspective where human, animal and environmental health are all considered to be interconnected. Several themes have been running through these chapters, one has been the investigation of different levels going from population movements recorded as historical demography (migration), through phenotype behaviour, to the genome sequences and bacterial cell metabolism. By combining four levels of analysis, four different snapshots of its behaviour in response to the forces that shape this pathogen can seen. This approach does not provide the sort of in-depth perspective obtained from only investigating a single system e.g. the effects of a single gene, but does provide a broad strokes picture of *C. jejuni* in New Zealand.

Another theme in this thesis has been the comparison of agricultural-associated and wild-bird associated *C jejuni* sequence types in a New Zealand context: as they both have markedly different host range and behaviour. The historical movements of cattle, sheep and poultry into New Zealand reflects the entry pattern of the agricultural-associated sequence types. Analysis of the phenotypes compared New Zealand isolates, an agricultural-associated sequence type and several associated with the wild-birds (indigenous pūkeko and the endemic takahē). The genome and metabolism comparisons highlighted similarities and differences between agricultural-associated

and the wild-bird associated sequence types.

The third theme, has been the examination of isolates associated with the pūkeko and the takahē, two closely related New Zealand water rails. The pūkeko form a spatial and temporal separated sub-population of *Porphyrio porphyrio melanotus*. The genomes of isolates from the Australian purple swamphen represent a population that the pūkeko has been separated from for $\sim$ 500 years [11]. Comparisons were made between *Campylobacter* spp. from these two long separated populations of *Porphyrio porphyrio melanotus*, one in New Zealand and one in Australia.

Within this thesis, each investigation has addressed facets of the two larger questions. Has *C. jejuni* followed a unique evolutionary trajectory in New Zealand? And what is the link between phenotype, genotype and host adaptation? In chapter 3 (the past), it quantifies the importation of livestock to New Zealand as its isolation resulted in the only terrestrial mammals, prior to the arrival of man, being two species of bat. The livestock would act as hosts to some microbes, previously not present in New Zealand, such as the *C. jejuni* ST associated with livestock. While this does not directly address if *C. jejuni* has followed a unique evolutionary trajectory in New Zealand, it does show some of the contributing history. Chapter 4 (the phenotype) examines the link between phenotype, genotype and host adaptation by investigating carbon utilisation patterns using Biolog phenotypic microarrays and comparing a ST associated with livestock (ST-42) with the ST-2381 lineage that is associated with wild-birds (pūkeko and takahē). Chapter 5 (the genome) uses comparative genomics to address both questions of this thesis. Isolates from an Australian wild bird (*Porphyrio porphyrio melanotus*) are compared to a lineage associated with two New Zealand wild bird (*Porphyrio porphyrio melanotus*, *Porphyrio hochstetteri*) to see how the geographical and temporal separation has affected the evolution of *C. jejuni* between the two groups. And a larger data set of New Zealand draft genomes associated with wild birds and livestock were compared to investigate genome and host adaptation. Finally in Chapter 6 (the metabolism) the genomic comparison of draft genomes associated with wild birds and livestock was made based on metabolic pathways to investigate phenotype, genotype and host adaptation.

The Past The published article "Quantification of historical livestock importation into New Zealand 1860-1979" demonstrated that Australia was a significant source of livestock to New Zealand, followed by the UK. The pattern of importation showed a large number of cattle and sheep arrived in NZ during early colonial times as shown in the rapid population expansion from 1851 to 1871, with the largest recorded numbers arriving during the 1860s. The pattern of arrival for poultry was quite different to cattle and sheep, as relatively small numbers were imported until the $\sim 1960s$ at which point large numbers started to be imported. This work was based on official records, however as the colony of New Zealand was only established in 1840 some of the data for the early years is missing . The methods, purpose for collecting it, and actual recording of the data changed over the decades which resulted in making some comparisons over such long periods

of time difficult. This analysis supports earlier reports that Australia was a significant source of livestock to New Zealand, although these earlier reports lacked quantification of the size and timing of demographic movement. The significance of this result for *C. jejuni* evolution is that it indicates the timing, size and source of the arrival of agriculture-associated STs (i.e. hosted by cattle, sheep, poultry) into New Zealand. It also provides an insight into New Zealand's biosecurity status, as the arriving livestock were potentially hosts for other pathogenic microbes, i.e. the demographic history relates to the host species and a range of microbes associated with cattle, sheep and poultry, not just *C. jejuni*.

The Phenotype Two Biolog phenotypic microarrays (PM1, PM2A) were used to examine the behaviours at two different temperatures (42℃, 22℃) of a group of eleven related New Zealand isolates. The isolates were from a ST associated with livestock and a generalist lifestyle (ST-42), and a group of isolates associated with two water rails, the pūkeko and the takahē representing a host specific lifestyle (ST-2381, ST-3655, ST-3663). While ST-3845 and ST-3673 are both rare isolates they have not yet been found in either bird; however both show a shared lineage with other STs (ST-2381, ST-3655, ST-3663). The pattern of variation between ST-42 and both ST-2381 and ST-3655 was consistent with the agriculture/generalist lifestyle associated sequence type (ST-42) utilising more carbon sources (resources) at 42℃ than the wild-bird/specialist lifestyle associated sequence types (ST-2381, ST-3655), however this simple pattern did not extend to the other isolates in ST-3663, ST-3845 and ST-3673. There was phenotypic variation between all the isolates at 42℃. Isolates from the same ST clustered together with similar patterns of carbon source utilisation, but clustering between different STs were not based on the degree of shared allelic profiles. Using the number of carbon sources utilised in the Biolog phenotypic microarrays at 42℃ did not support differentiating host association between livestock and wild-bird. However there may be a basis for differentiating between these groups on the ability to utilise specific carbon sources, as the only the wild-bird associated isolates were able to utilised Quinic acid at 42℃. In total at 42℃ the *C. jejuni* utilised 29 carbon sources, while at 22℃ only a subset group of seven was utilised. The carbon sources showing activity at 22℃ tended to be closely related to the TCA cycle and energy production. 22℃ is the type of temperature *C. jejuni* would experience when in the environment between hosts, or in a kitchen. These results suggest a control mechanism aimed at the TCA cycle in *C. jejuni* would be effective at room temperature, whereas mechanisms aimed at a growth-related metabolism would not.

The Genome The pūkeko, is a close relative of the takahē, and a geographically isolated sub-population to the Australian swamphen. The pūkeko is the host of some novel *C. jejuni* commonly found in New Zealand waters [8]. A genomic comparison was made between isolates from Australian swamphen isolates obtained in Victoria, and a New Zealand dataset including pūkeko and takahē

associated STs. This identified several *Campylobacter* spp. in the Australian swamphen and put them in context to *C. jejuni* ST associated with the New Zealand pūkeko and takahē, as well as the overall relationship of New Zealand wild-bird and agriculture-associated isolates. There was a close relationship between *C. jejuni* from the Australian purple swamphen and isolates associated with the New Zealand pūkeko and takahē, in both the core genome and pan-genome analysis that supports a recent common ancestor followed by divergence, probably after separation. However ST-2381 appears to be less close to the Australian purple swamphen isolates than other pūkeko and takahē associated isolates. A putative new *Campylobacter* spp. *C.* sp. nov. 4 was identified in the Australian swamphen isolates by genomic analysis, although more bench-based analysis is needed to confirm this finding and establish a taxonomy. *C.* sp nov 1 which has previously only been identified in the pūkeko and takahē was also present [16]. The comparison of full length 16S rRNA sequences showed ST-474, an important cause of campylobacteriosis in New Zealand, has the same sequence as some *C. coli* isolates. This could be the result of recombination between some *C. coli*, probably clade 1, and some *C. jejuni* in this case ST-474. The core genome as represented by MLST and rMLST was able to separate the species that were not *C. jejuni* and within *C. jejuni* identify the isolates into MLST based clonal complexes. The pan-genome formed by our dataset when represented by a presence/absence matrix of homologues, displayed very similar relationships as the core genome with the isolates grouping by species and MLST clonal complex groups, in two different analyses; the Discriminant analysis of principal components (DAPC) of the population structure and the Additive Dollo Distance (ADD) matrix of the pan-genome. These results suggest at the genome level for both agriculture-associated and wild-bird associated *C. jejuni*, there may be a combination of recombination mainly occurring between more closely related sequence types producing homogeneity, and possibly purifying selection for host adaptation removing the "less adapted" combinations.

The Metabolism The *C. jejuni* genomes from starlings, Australian purple swamphen and agriculture-associated clonal complexes were put into three groups and annotated using RAST [531][532]. Then using the genomes of the three groups, a comparison was made based on the presence or absence of a protein encoding gene (peg), with each peg having a role in a previously defined metabolic pathway called a subsystem. The three subsystems identified as being important differences between the three host groups were: type VI secretion system (T6SS), coenzyme A biosynthesis and the *Campylobacter* iron metabolism. Comparing the three subsystems to a wider group of genomes showed some interesting patterns. The T6SS subsystem was found in a limited number of *C. jejuni* genomes; the ST-430 clonal complex, ST-2381, ST-3655 and isolates from the Australian purple swamphen. The T6SS subsystem was also present in the new *Campylobacter* spp. (*C.* sp. nov. 1 and 4) found in the Australian purple swamphen. T6SS has previously been reported in *C. jejuni* and is associated with pathogenicity, suggesting if wild-bird isolates do act as genetic

reservoirs for the livestock associated ST, they could transfer pathogenic features. Three extra enzymes in coenzyme A biosynthesis, have previously been found in cattle isolates [370] and absent in environmental isolates, presumably wild-birds [502][p93], a distribution pattern which these results support. There was variation in the *Campylobacter* spp. iron metabolism subsystem found in the dataset of isolates but the pattern of differences between host groups was not clear.

Summary This work has contributed to the understanding of the molecular ecology and evolution of *C. jejuni* by examining its behaviour at four levels across a temporal-spatial framework:

- Confirming Australia as an important source of cattle and sheep importations, notably during the 1860s when significant numbers were imported. By quantifying the historical size, source and distribution entry pattern of livestock (cattle, sheep, poultry) to New Zealand, a country that historically had no livestock, we can develop a greater understanding of the past and current population structures of livestock and many micro-organisms that they host. The historical demography of livestock arrival to New Zealand also helps to understand the basis for the absence of many infectious diseases of livestock which resulted in a high national biosecurity status in terms of freedom from disease.

- Phenotypic variation was tested for host association using Biolog phenotypic microarray single carbon source utilisation patterns at 22℃ and 42℃. The pattern of *C. jejuni* phenotypic variation was closely related to ST and in total 29 carbons sources were utilised. At 22℃ it was a more restricted range of seven carbon sources in total, and all were closely related to the TCA cycle and energy production. In terms of total carbons sources used the phenotypic variation was not associated with host at 22℃ or 42℃, but at 42℃ the ST-2381 lineage did use Quinic acid.

- A putative new *Campylobacter* spp. *C.* sp. nov. 4 was identified in the Australian purple swamphen isolates by genomic analysis, more bench-based analysis is needed to confirm this finding. *C.* sp. nov. 1, previously only identified in the pūkeko and takahē, was also present.

- Combining the results showed *C. jejuni* having a similar relationship with each other when comparing the core genome, as typed by MLST and rMLST, and the accessory genome as expressed by the presence/absence of genes, metabolic subsystems and phenotypes. The relationship does not appear to be based on the individual alleles of the MLST, but the combination reflected by a ST (an allelic profile). This provides a basis to understand why the MLST was so successful in host association. These results suggest that a combination of recombination predominantly between closely related STs produces a degree of homogeneity, and with purifying selection drive host adaptation in *C. jejuni* lineages.

- There was a close relationship between *C. jejuni* from the Australian purple swamphen and isolates associated with the New Zealand pūkeko and takahē, in both the core genome and

pan-genome analysis that supports a recent common ancestor followed by divergence, probably after separation. However ST-2381 appears to be slightly less close to the Australian purple swamphen isolates and the other pūkeko and takahē associated isolates in the phenotype, core genome, pan-genome and T6SS subsystem comparisons.

- Three metabolic pathways (type VI secretion system, coenzyme A biosynthesis and *Campylobacter* iron metabolism) were identified as being important differences between the *C. jejuni* genomes from three host groups (starlings, Australian swamphen and several agricultural-associated clonal complexes).

Future work At each of the four levels of investigation, despite new findings, new questions were also raised.

The demographic history of the introduction of livestock species to New Zealand, was well recorded in official documents, but demographic records for the annual arrival of the migratory birds and sea birds that travel long distances is less well known. In terms of national biosecurity, understanding the arrival and departure of highly mobile bird populations, as well as the pathogens they may vector is important. There is limited information on any *C. jejuni* or *Campylobacter* spp. they may carry or other pathogens.

The association between *Porphyrio porphyrio* and *Campylobacter* spp. should be investigated in more depth. The finding of several putative new *Campylobacter* spp. in pūkeko, takahē and Australian purple swamphen suggests the purple swamphen family as a whole should be investigated world-wide to see if they are associated with more new *Campylobacter* spp.. If the swamphen, in general, are found to be associated with a high prevalence of new *Campylobacter* spp. then it would be important to understand their interactions with *Campylobacter* spp. and identify what is causing the speciation events? Is the nature of *Porphyrio porphyrio* guild (resources they exploit) important to the prevalence [323] and type of associated *Campylobacter* spp. or are other factors more important, like gastrointestinal niche within the bird?

A better understanding of the association between *Porphyrio porphyrio* and *Campylobacter* spp. may help understand the relationship to wild-birds in general. Although this thesis has found the differences between wild-bird associated ST and the agriculture-associated ST permeates the whole genome, the amount of interaction particularly, to what extent can/do wild-bird associated ST act as genetic reservoirs for the agricultural associated sequence types has yet to be established. What is the potential amount of genetic exchange and what is the actual amount occurring between the different host associated groups with an emphasis on wild-birds? Are some ST more likely to interact with others, and could some species like the chicken be acting as mixing point, even on a temporary basis and facilitating genetic exchange? By investigating the role of wild-bird associated ST in campylobacteriosis, we will better understand the past and future potential for evolution of

C. jejuni, and its ability to act as a pathogen.

An important area for public health is a better understanding of the behaviour *C. jejuni* at lower temperatures i.e. < 30℃, already work has shown it has a diverse array of survival behaviours when in the environment (biofilms [592], carriage inside amoeba [127], the viable but non-culturable (VBNC) state [593]). Despite having a small genome *C. jejuni* has shown great phenotypic and genotypic variation, including phase variation [506]. However at lower temperatures there is a general down regulation of many of the expressed genes [409], the factors controlling this decrease in gene expression and identification of essential pathways should identify mechanisms that could be used in the food processing industry and kitchens to control *C. jejuni* .

In this thesis, shows a relationship between the core and accessory genome was demonstrated, which needs more investigation. De Haan et al.(2012) and Zautner et al. (2011) identified the link between MLST and metabolic traits based on specific genes, but this investigation showed the relatedness exists across the core and accessory genomes, in wild-bird associate and agriculture-associated ST/CC [498][554]. What is required is a better understanding of the nature of the relatedness between the core as expressed by ST/CC and the accessory genome, as seen in phenotypic behaviour, metabolic pathways and the pattern of presence or absence of genes. Is this the result of a dynamic homogenisation process due to a combination of recombination predominantly with closely related ST and selection? How is a bacteria with a high rate of recombination and genotypic variation maintaining this level of relationship between the core and accessory genome? Barriers to recombination between more different ST have been reported [421] [556]. By understanding these evolutionary forces and the factors they operate through, we may better understand *C. jejuni* and pathogens in general. A better understanding of the forces operating on a pathogen will enable better strategies to be developed that limit the evolution of features like antibiotic resistance, increased pathogenicity, or evasion of immunological recognition.

Bibliography

[1] S. V. R. Epps, R. B. Harvey, M. E. Hume, T. D. Phillips, R. C. Anderson, and D. J. Nisbet, "Foodborne *Campylobacter*: infections, metabolism, pathogenesis and reservoirs.," *International Journal of Environmental Research and Public Health*, vol. 10, pp. 6292–6304, Dec. 2013.

[2] A. Sears, M. G. Baker, N. Wilson, J. Marshall, P. Muellner, D. M. Campbell, R. J. Lake, and N. P. French, "Marked Campylobacteriosis Decline after Interventions Aimed at Poultry, New Zealand," *Emerging Infectious Diseases*, vol. 17, pp. 1007–1015, June 2011.

[3] P. Muellner, P. E. Carter, A. C. Midwinter, J. M. Collins-Emerson, R. Pattison, and N. P. French, "Molecular epidemiology of *Campylobacter jejuni* : comparing Multilocus Sequence Types from humans and poultry .," in *Proceedings of the 11th International Symposium on Veterinary Epidemiology and Economics,*, pp. 1–4, Nov. 2011.

[4] Z. Elharrif and D. F. Mégraud, "Characterization of thermophilic *Campylobacter*: I. Carbon-substrate utilization tests," *Current Microbiology*, vol. 13, pp. 117–122, May 1986.

[5] E. Gripp, D. Hlahla, X. Didelot, F. Kops, S. Maurischat, K. Tedin, L. Ellerbroek, D. Schomburg, T. Janssen, P. Bartholomäus, S. Woltemate, M. Uhr, B. Brenneke, P. Gruening, G. Gerlach, L. Wieler, S. Suerbaum, and C. Josenhans, "Closely related *Campylobacter jejuni* strains from different sources reveal a generalist rather than a specialist lifestyle," *BMC Genomics*, vol. 12, p. 584, Nov. 2011.

[6] S. K. Sheppard, F. M. Colles, N. M. McCarthy, N. J. C. Strachan, I. D. Ogden, K. J. Forbes, J. F. Dallas, and M. C. J. Maiden, "Niche segregation and genetic structure of *Campylobacter jejuni* populations from wild and agricultural host species.," *Molecular Ecology*, vol. 20, pp. 3484–3490, Aug. 2011.

[7] P. Griekspoor, F. M. Colles, N. M. McCarthy, P. M. Hansbro, C. Ashhurst-Smith, B. Olsen, D. Hasselquist, M. C. J. Maiden, and J. Waldenstrom, "Marked host specificity and lack of phylogeographic population structure of *Campylobacter jejuni* in wild birds," *Molecular Ecology*, vol. 22, pp. 1463–1472, Jan. 2013.

[8] P. E. Carter, S. M. McTavish, H. J. L. Brooks, D. Campbell, J. M. Collins-Emerson, A. C. Midwinter, and N. P. French, "Novel clonal complexes with an unknown animal reservoir dominate *Campylobacter jejuni* isolates from river water in New Zealand.," *Applied and Environmental Microbiology*, vol. 75, pp. 6038–6046, Oct. 2009.

[9] N. P. French, J. Marshall, V. Moran, and mEpiLab, "New and emerging data on typing of *Campylobacter* spp. strains in animals, environmental matrices and humans," Tech. Rep. Final report: 07-10436, Ministry of Agriculture and Forestry, 2011.

[10] J.-M. Membré, M. Laroche, and C. Magras, "Meta-analysis of *Campylobacter* spp. Survival Data within a Temperature Range of 0 to 42°C," *Journal of Food Protection*, vol. 76, pp. 1726–1732, Oct. 2013.

[11] J. C. García-R and S. A. Trewick, "Dispersal and speciation in purple swamphens (Rallidae: *Porphyrio*)," *The Auk*, vol. 132, pp. 140–155, Nov. 2014.

[12] P. R. Millener, *The Quaternary avifauna of the North Island, New Zealand*. PhD thesis, University of Auckland, 1981.

[13] S. L. Olson, "Classification of Rallidae," *Wilson bulletin*, vol. 85, Jan. 1973.

[14] C. G. Sibley and J. E. Alquist, *Phylogeny and Classification of Birds*. A Study in Molecular Evolution, Yale University Press, 1990.

[15] S. M. McTavish, C. E. Pope, C. Nicol, K. Sexton, N. P. French, and P. E. Carter, "Wide geographical distribution of internationally rare *Campylobacter* clones within New Zealand.," *Epidemiology and Infection*, vol. 136, pp. 1244–1252, Sept. 2008.

[16] N. P. French, S. Yu, P. J. Biggs, B. Holland, P. Fearnhead, B. Binney, A. J. Fox, D. Grove-White, J. Leigh, W. G. Miller, P. Muellner, and P. E. Carter, "Evolution of *Campylobacter* species in New Zealand," in *Campylobacter Ecology and Evolution* (S. K. Sheppard and G. Meric, eds.), pp. 221–240, Norfolk, UK: Caister Academic Press, Jan. 2014.

[17] J. Waldenstrom and P. Griekspoor, "Ecology and Host association of *Campylobacter* in wild birds ," in *Campylobacter Ecology and Evolution* (G. Meric, ed.), pp. 265–285, Caister Academic Press, Nov. 2014.

[18] M. Steele, B. McNab, L. Fruhner, S. DeGrandis, D. Woodward, and J. A. Odumeru, "Epidemiological typing of *Campylobacter* isolates from meat processing plants by pulsed-field gel electrophoresis, fatty acid profile typing, serotyping, and biotyping.," *Applied and Environmental Microbiology*, vol. 64, pp. 2346–2349, July 1998.

[19] K. N. Eberle and A. S. Kiess, "Phenotypic and genotypic methods for typing *Campylobacter jejuni* and *Campylobacter coli* in poultry," *Poultry Science*, vol. 91, pp. 255–264, Dec. 2011.

[20] K. Takai, K. H. Nealson, and K. Horikoshi, "*Hydrogenimonas thermophila* gen. nov., sp. nov., a novel thermophilic, hydrogen-oxidizing chemolithoautotroph within the ε-Proteobacteria, isolated from a black smoker in a Central Indian Ridge hydrothermal field," *International Journal Of Systematic And Evolutionary Microbiology*, vol. 54, no. 1, pp. 25–32, 2004.

[21] Y. Suzuki, T. Sasaki, M. Suzuki, Y. Nogi, T. Miwa, K. H. Nealson, and K. Horikoshi, "Novel Chemoautotrophic Endosymbiosis between a Member of the Epsilonproteobacteria and the Hydrothermal-Vent

Gastropod *Alviniconcha aff. hessleri (Gastropoda: Provannidae)* from the Indian Ocean," *Applied and Environmental Microbiology*, vol. 71, pp. 5440–5450, Sept. 2005.

[22] M. Sebald and M. Vernon, "Base DNA content and classification of *Vibrios*," *Annales de l'Institut Pasteur*, vol. 105, pp. 897–910, 1963.

[23] M. Veron and R. Chatelain, "Taxonomic study of the genus *Campylobacter* Sebald and Veron and designation of the neotype strain for the type species, *Campylobacter fetus* (Smith and Taylor) Sebald and Veron," *International Journal of Systematic Bacteriology*, vol. 23, no. 2, pp. 122–134, 1973.

[24] A. P. Moran and M. E. Upton, "A comparative study of the rod and coccoid forms of *Campylobacter jejuni* ATCC 29428.," *The Journal of Applied Bacteriology*, vol. 60, pp. 103–110, Feb. 1986.

[25] D. M. Rollins and R. R. Colwell, "Viable but nonculturable stage of *Campylobacter jejuni* and its role in survival in the natural aquatic environment.," *Applied and Environmental Microbiology*, vol. 52, pp. 531–538, Sept. 1986.

[26] C. A. Woodall, M. A. Jones, P. A. Barrow, J. Hinds, G. L. Marsden, D. J. Kelly, N. Dorrell, B. W. Wren, and D. J. Maskell, "*Campylobacter jejuni* gene expression in the chick cecum: evidence for adaptation to a low-oxygen environment.," *Infection and Immunity*, vol. 73, pp. 5278–5285, Aug. 2005.

[27] M. B. Skirrow, "*CAMPYLOBACTER* — Properties and Occurrence," in *Encyclopedia of Food Sciences and Nutrition*, pp. 779–786, Elsevier, 2003.

[28] P. Vandamme and H. Goossens, "Taxonomy of *Campylobacter, Arcobacter,* and *Helicobacter*: a review.," *International journal of medical microbiology*, vol. 276esher, no. 4, pp. 447–472, 1992.

[29] P. Vandamme, L. Debruyne, E. De Brandt, and E. Falsen, "Reclassification of *Bacteroides ureolyticus* as *Campylobacter ureolyticus* comb. nov., and emended description of the genus Campylobacter.," *International Journal Of Systematic And Evolutionary Microbiology*, vol. 60, pp. 2016–2022, Sept. 2010.

[30] J. P. Butzler, "*Campylobacter,* from obscurity to celebrity.," *Clinical Microbiology and Infection*, vol. 10, no. 10, pp. 868–876, 2004.

[31] T. Escherich, "Beitrage zur Kenntniss der Darmbacterien," *Muenchener Medizinische Wochenschrift*, vol. 33, pp. 815–817, 1886.

[32] J. McFadyean and S. Stockman, "Report of the Departmental Committee appointed by the Board of Agriculture and Fisheries to inquire into epizootic abortion," tech. rep., Great Britain. Board of Agriculture and Fisheries. Committee to Inquire into Epizootic Abortion.

[33] M. B. Skirrow, "John McFadyean and the centenary of the first isolation of *Campylobacter* species.," *Clinical Infectious Diseases*, vol. 43, pp. 1213–1217, Oct. 2006.

[34] W. L. Hindmarsh, "A case of bovine vibrionic abortion in New South Wales," *Australian Veterinary Journal*, vol. 18, no. 4, pp. 175–176, 1942.

[35] T. Smith and M. S. Taylor, "Some morphological and biological characters of the Spirilla (*Vibrio fetus*, n. sp.) associated with disease of the fetal membranes in cattle.," *Journal of Experimental Medicine*, vol. 30, pp. 299–311, Oct. 1919.

[36] T. Stegenga and J. I. Terpstra, "Over *Vibrio foetus* infecties bij het rund en enzootische steriliteit," *Tijdschr Diergeneesk*, vol. 74, pp. 293–296, 1949.

[37] T. Smith and M. L. Orcutt, "Vibrios From Calves and Their Serological Relation To Vibrio Fetus," *Journal of Experimental Medicine*, vol. 45, pp. 391–397, Feb. 1927.

[38] F. S. Jones, M. Orcutt, and R. B. Little, "Vibrios (Vibrio jejuni, n. sp.) associated with intestinal disorders of cows and calves," *Journal of Experimental Medicine*, vol. 53, no. 6, pp. 853–864, 1931.

[39] L. P. Doyle, "A vibrio associated with swine dysentery," *American Journal of Veterinary Research*, 1944.

[40] D. McFarlane, R. M. Salisbury, H. G. Osborne, and J. L. Jebson, "Investigation into sheep abortion in New Zealand during the 1950 lambing season," *Australian Veterinary Journal*, vol. 28, no. 9, pp. 221–226, 1952.

[41] J. P. Butzler, P. Dekeyser, M. Detrain, and F. Dehaen, "Related vibrio in stools.," *The Journal of pediatrics*, vol. 82, no. 3, pp. 493–495, 1973.

[42] A. J. Levy, "A gastro-enteritis outbreak probably due to a bovine strain of vibrio," *The Yale Journal of Biology and Medicine*, 1946.

[43] M. B. Skirrow, "Campylobacter enteritis: a "new" disease.," *British Medical Journal*, vol. 2, pp. 9–11, July 1977.

[44] I. Nachamkin, C. M. Szymanski, and M. J. Blaser, eds., Campylobacter*, Third Edition*. New York: American Society of Microbiology, 3 ed., Jan. 2008.

[45] M. J. Blaser and J. Engberg, "Clinical aspects of *Campylobacter jejuni* and *Campylobacter coli* infections," in *Campylobacter, Third Edition* (I. Nachamkin, C. M. Szymanski, and M. J. Blaser, eds.), pp. 99–122, Washington, DC: American Society of Microbiology, Jan. 2008.

[46] E. Bessède, P. Lehours, L. Labadi, S. Bakiri, and F. Megraud, "Comparison of Characteristics of Patients Infected by *Campylobacter jejuni, Campylobacter coli*, and *Campylobacter fetus*," *Journal of Clinical Microbiology*, vol. 52, pp. 328–330, Dec. 2013.

[47] V. B. Young and L. S. Mansfield, "*Campylobacter* Infection - Clinical context," in *Campylobacter* (J. M. Ketley and M. E. Konkel, eds.), pp. 1–12, Horizon Scientific Press, 2005.

[48] P. C. R. Godschalk, A. P. Heikema, M. Gilbert, T. Komagamine, C. W. Ang, J. Glerum, D. Brochu, J. Li, N. Yuki, B. C. Jacobs, A. van Belkum, and H. P. Endtz, "The crucial role of *Campylobacter jejuni* genes in anti-ganglioside antibody induction in Guillain-Barre syndrome.," *The Journal of clinical investigation*, vol. 114, pp. 1659–1665, Dec. 2004.

[49] N. Yuki, "Carbohydrate mimicry between human ganglioside GM1 and *Campylobacter jejuni*

lipooligosaccharide causes Guillain – Barre syndrome," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, no. 31, pp. 11404–11409, 2004.

[50] S. F. Altekruse, N. J. Stern, P. I. Fields, and D. L. Swerdlow, "*Campylobacter jejuni*–an emerging foodborne pathogen.," *Emerging Infectious Diseases*, vol. 5, no. 1, pp. 28–35, 1999.

[51] M. Fisher, "An Unusual Variant of Acute Idiopathic Polyneuritis (Syndrome of Ophthalmoplegia, Ataxia and Areflexia)," *New England Journal of Medicine*, vol. 255, pp. 57–65, July 1956.

[52] M. Mori, S. Kuwabara, and N. Yuki, "Fisher syndrome: clinical features, immunopathogenesis and management," *Expert Review of Neurotherapeutics*, vol. 12, no. 1, pp. 39–51, 2012.

[53] J. E. Pope, A. Krizova, A. X. Garg, H. Thiessen-Philbrook, and J. M. Ouimet, "Campylobacter reactive arthritis: a systematic review.," *Seminars in Arthritis and Rheumatism*, vol. 37, no. 1, pp. 48–55, 2007.

[54] J. P. Thornley, D. Jenkins, K. R. Neal, T. Wright, J. Brough, and R. C. Spiller, "Relationship of *Campylobacter* toxigenicity in vitro to the development of postinfectious irritable bowel syndrome.," *The Journal of Infectious Diseases*, vol. 184, no. 5, pp. 606–609, 2001.

[55] G. Nylen, F. Dunstan, S. R. Palmer, Y. Andersson, F. Bager, J. Cowden, G. Feierl, Y. Galloway, G. Kapperud, F. Megraud, K. Molbak, L. R. Petersen, and P. Ruutu, "The seasonal distribution of *campylobacter* infection in nine European countries and New Zealand.," *Epidemiology and Infection*, vol. 128, pp. 383–390, June 2002.

[56] R. S. Kovats, S. J. Edwards, D. Charron, J. Cowden, R. M. D'Souza, K. L. Ebi, C. Gauci, P. Gerner-Smidt, S. Hajat, S. Hales, G. Hernández Pezzi, B. Kriz, K. Kutsar, P. McKeown, K. Mellou, B. Menne, S. O'Brien, and H. Schmid, "Climate variability and *campylobacter* infection: an international study.," *International journal of Biometeorology*, vol. 49, pp. 207–214, Mar. 2005.

[57] G. Miller, G. M. Dunn, and A. Smith-Palmer, "Human campylobacteriosis in Scotland: seasonality, regional trends and bursts of infection.," *Epidemiology and Infection*, vol. 132, pp. 585–593, Aug. 2004.

[58] M. B. Skirrow, "A demographic survey of *Campylobacter, Salmonella* and *Shigella* infections in England. A Public Health Laboratory Service Survey.," *Epidemiology and Infection*, vol. 99, pp. 647–657, Dec. 1987.

[59] T. Tenkate, R. Stafford, and B. McCall, "A five year review of *Campylobacter* infection in Queensland," *Communicable Diseases Intelligence*, vol. 20, p. 478, Oct. 1996.

[60] N. M. McCarthy, I. A. Gillespie, A. J. Lawson, J. Richardson, K. R. Neal, P. R. Hawtin, M. C. J. Maiden, and S. J. O'Brien, "Molecular epidemiology of human *Campylobacter jejuni* shows association between seasonal and international patterns of disease," *Epidemiology and Infection*, vol. 140, no. 12, pp. 1–9, 2012.

[61] A. Friedrich, J. C. Marshall, P. J. Biggs, A. C. Midwinter, and N. P. French, "Seasonality of

Campylobacter jejuni isolates associated with human campylobacteriosis in the Manawatu region, New Zealand.," *Journal of Hygiene*, pp. 1–9, Sept. 2015.

[62] M. Hearnden, C. Skelly, R. Eyles, and P. Weinstein, "The regionality of campylobacteriosis seasonality in New Zealand.," *International journal of environmental health research*, vol. 13, pp. 337–348, Dec. 2003.

[63] K. Obiri-Danso, N. Paul, and K. Jones, "The effects of UVB and temperature on the survival of natural populations and pure cultures of *Campylobacter jejuni, Camp. coli, Camp. lari* and urease-positive thermophilic *campylobacters* (UPTC) in surface waters," *Journal of Applied Microbiology*, vol. 90, pp. 256–267, Feb. 2001.

[64] K. Jones, "*Campylobacters* in water, sewage and the environment.," *Journal Applied Microbiology*, no. 30, pp. 68S–79S, 2001.

[65] R. Eyles, D. Niyogi, C. Townsend, G. Benwell, and P. Weinstein, "Spatial and temporal patterns of *Campylobacter* contamination underlying public health risk in the Taieri River, New Zealand.," *Journal of Environmental Quality*, vol. 32, pp. 1820–1828, Sept. 2003.

[66] H. H. Abulreesh, T. A. Paget, and R. Goulder, "*Campylobacter* in Waterfowl and Aquatic Environments: Incidence and Methods of Detection," *Environmental Science & Technology*, vol. 40, pp. 7122–7131, Dec. 2006.

[67] P. Muellner, E. Pleydell, R. Pirie, M. G. Baker, D. Campbell, P. E. Carter, and N. P. French, "Molecular-based surveillance of campylobacteriosis in New Zealand–from source attribution to genomic epidemiology.," *Eurosurveillance*, vol. 18, no. 3, 2013.

[68] I. Cha, N. Kim, J.-G. Nam, E.-S. Choi, G. T. Chung, Y.-H. Kang, and S. Hong, "Genetic Diversity of *Campylobacter jejuni* Isolates from Korea and Travel-Associated Cases from East and Southeast Asian Countries," *Japanese Journal of Infectious Diseases*, vol. 67, no. 6, pp. 490–494, 2014.

[69] K. E. Smith, J. M. Besser, C. W. Hedberg, F. T. Leano, J. B. Bender, J. H. Wicklund, B. P. Johnson, K. A. Moore, and M. T. Osterholm, "Quinolone-resistant *Campylobacter jejuni* infections in Minnesota, 1992–1998," *New England Journal of Medicine*, vol. 340, no. 20, pp. 1525–1532, 1999.

[70] R. E. Black, M. M. Levine, L. M. Clements, T. P. Hughes, and M. J. Blaser, "Experimental *Campylobacter jejuni* Infection in Humans," *The Journal of Infectious Diseases*, vol. 157, no. 3, pp. 472–479, 1988.

[71] D. A. Robinson, "Campylobacter Infection," *The Journal of the Royal Society for the Promotion of Health*, vol. 101, no. 4, pp. 138–140, 1981.

[72] G. J. Medema, P. F. Teunis, A. H. Havelaar, and C. N. Haas, "Assessment of the dose-response relationship of *Campylobacter jejuni.*," *International Journal of Food Microbiology*, vol. 30, pp. 101–111, June 1996.

[73] P. F. Teunis and A. H. Havelaar, "The Beta Poisson dose-response model is not a single-hit model.," *Risk analysis : an official publication of the Society for Risk Analysis*, vol. 20, pp. 513–520, Aug. 2000.

[74] R. E. Black, M. M. Levine, D. Perlman, M. L. Clements, and M. J. Blaser, "Human volunteer studies with *Campylobacter jejuni*," in *Campylobacter jejuni. Current status and future trends.* (U. A. M. Research and D. Command, eds.), 1993.

[75] F. J. Sorvillo, L. E. Lieb, and S. H. Waterman, "Incidence of campylobacteriosis among patients with AIDS in Los Angeles County.," *Journal of Acquired Immune Deficiency Syndromes*, vol. 4, no. 6, pp. 598–602, 1991.

[76] M. R. Nelson, D. C. Shanson, D. A. Hawkins, and B. G. Gazzard, "Salmonella, Campylobacter and Shigella in HIV-seropositive patients," *AIDs*, vol. 6, no. 12, pp. 1495–1498, 1992.

[77] M. J. Blaser, E. Sazie, and Williams, L. P., "The Influence of Immunity on Raw Milk– Associated *Campylobacter* Infection," *The Journal of the American Medical Association*, vol. 257, no. 1, pp. 43–46, 1987.

[78] D. M. Jones, D. A. Robinson, and J. Eldridge, "Serological studies in two outbreaks of *Campylobacter jejuni* infection," *Journal of Hygiene*, vol. 87, no. 2, pp. 163–170, 1981.

[79] S. K. Sheppard, M. C. J. Maiden, and D. Falush, "Population genetics of *Campylobacter*," in *Bacterial Population Genetics in Infectious Disease* (D. A. Robinson, D. Falush, and E. J. Feil, eds.), pp. 181–194, Hoboken, NJ, USA: John Wiley & Sons, Inc., Apr. 2010.

[80] K. E. Dingle, N. M. McCarthy, A. J. Cody, T. E. Peto, and M. C. J. Maiden, "Extended sequence typing of *Campylobacter* spp., United Kingdom.," *Emerging Infectious Diseases*, vol. 14, pp. 1620–1622, Oct. 2008.

[81] R. V. Tauxe, "Incidence, trends and sources of campylobacteriosis in developed countries: an overview.," in *The increasing incidence of human campylobacteriosis. Report and proceedings of a WHO consultation of experts*, (Copenhagen, Denmark), pp. 42–43, Nov. 2000.

[82] A. Amato-Gauci, "ECDC to launch first report on communicable diseases epidemiology in the European Union," *Eurosurveillance*, vol. 12, no. 23, 2007.

[83] C. K. Olson, S. Ethelberg, and R. V. Tauxe, "Epidemiology of *Campylobacter* infections in industrialized nations ," in *Campylobacter, Third Edition* (I. Nachamkin, C. M. Szymanski, and M. J. Blaser, eds.), pp. 163–190, Washington, DC: American Society of Microbiology, Jan. 2008.

[84] G. Kapperud and S. Aasen, "Descriptive epidemiology of infections due to thermotolerant *Campylobacter* spp. in Norway, 1979-1988," *APMIS*, vol. 100, no. 7-12, pp. 883–890, 1992.

[85] N. J. C. Strachan and K. J. Forbes, "Extensive spatial and temporal clustering of Campylobacter infections," in *Campylobacter Ecology and Evolution* (G. Meric, ed.), University Librarian, Nov. 2014.

[86] R. G. Pebody, M. J. Ryan, and P. G. Wall, "Outbreaks of campylobacter infection: rare events for a common pathogen.," *Communicable disease report. CDR review*, vol. 7, pp. R33–7, Mar. 1997.

[87] W. C. Hazeleger, J. D. Janse, P. M. Koenraad, R. R. Beumer, F. M. Rombouts, and T. Abee, "Temperature-dependent membrane fatty acid and cell physiology changes in coccoid forms of *Campylobacter jejuni.*," *Applied and Environmental Microbiology*, vol. 61, pp. 2713–2719, July 1995.

[88] I. A. Gillespie, S. J. O'Brien, G. K. Adak, C. C. Tam, J. A. Frost, F. J. Bolton, D. S. Tompkins, and Campylobacter Sentinel Surveillance Scheme Collaborators, "Point source outbreaks of Campylobacter jejuni infection–are they more common than we think and what might cause them?," *Epidemiology and Infection*, vol. 130, pp. 367–375, June 2003.

[89] M. P. Doyle and D. J. Roman, "Growth and survival of *Campylobacter fetus* subsp. *jejuni* as a function of temperature and pH .," *Journal of Food Protection*, vol. 44, pp. 596–601, Jan. 1981.

[90] R. C. Wood, K. L. MacDonald, and M. T. Osterholm, "Campylobacter enteritis outbreaks associated with drinking raw milk during youth activities. A 10-year review of outbreaks in the United States.," *The Journal of the American Medical Association*, vol. 268, pp. 3228–3230, Dec. 1992.

[91] K. J. Forbes, F. J. Gormley, J. F. Dallas, O. Labovitiadi, M. MacRae, R. J. Owen, J. Richardson, N. J. C. Strachan, J. M. Cowden, I. D. Ogden, and C. C. McGuigan, "*Campylobacter* immunity and coinfection following a large outbreak in a farming community.," *Journal of Clinical Microbiology*, vol. 47, pp. 111–116, Jan. 2009.

[92] A. P. Black, M. D. Kirk, and G. Millard, "Campylobacter outbreak due to chicken consumption at an Australian Capital Territory restaurant.," *Communicable diseases intelligence quarterly report*, vol. 30, no. 3, pp. 373–377, 2006.

[93] R. Lake, "Transmission Routes for Campylobacteriosis in New Zealand," Tech. Rep. FW0424, New Zealand Food Safety Authority, May 2006.

[94] M. J. Blaser, P. Checko, C. Bopp, A. Bruce, and J. M. Hughes, "Campylobacter enteritis associated with foodborne transmission.," *American Journal of Epidemiology*, vol. 116, pp. 886–894, Dec. 1982.

[95] I. Indikova, T. J. Humphrey, and F. Hilbert, "Survival with a Helping Hand: Campylobacter and Microbiota," *Frontiers in Microbiology*, vol. 6, p. 1055, Nov. 2015.

[96] World Health Organization, "The global view of Campylobacteriosis : report of an expert consultation.," tech. rep., World Health Organization, Utrecht, The Netherlands., 2012.

[97] J. Eberhart-Phillips, N. Walker, D. Bell, D. Sinclair, W. Rainger, and M. Bates, "Campylobacteriosis in New Zealand: results of a case-control study.," *Journal of Epidemiology & Community Health*, vol. 51, no. 6, pp. 686–691, 1997.

[98] G. B. McBride, A. Ball, S. Harper, B. Horn, R. Lake, S. Elliott, J. Marshall, and P. van der Logt, "*Campylobacter* in Food and the Environment :," tech. rep., Ministry of Agriculture and Forestry, July 2011.

[99] P. Muellner and S. Hathaway, "Source attribution of food-borne zoonoses in New Zealand: a modified Hald model.," *Risk analysis : an official publication of the Society for Risk Analysis*, vol. 29, no. 7, pp. 970–984, 2009.

[100] R. Ikram, S. Chambers, P. Mitchell, M. A. Brieseman, and O. H. Ikam, "A case control study to determine risk factors for *campylobacter* infection in Christchurch in the summer of 1992-3.," *The New Zealand Medical Journal*, vol. 107, pp. 430–432, Oct. 1994.

[101] N. P. French and J. Marshall, "Source Attribution Studies for Campylobacteriosis in New Zealand," Tech. Rep. 2014/11, Ministry of Primary Industry, Mar. 2014.

[102] N. P. French, P. Muellner, and E. Pleydell, "Attribution of potentially foodborne enteric diseases: human campylobacteriosis - ruminant attribution," Tech. Rep. Final report: 11777, New Zealand Food Safety Authority, Jan. 2011.

[103] N. J. Stern and K. L. Hiett, "*Campylobacter* spp. in Icelandic poultry operations and human disease.," *Epidemiology and Infection*, vol. 130, no. 1, pp. 23–32, 2003.

[104] B. Borck, C. Galliano, S. Nordentoft, S. Ethelberg, L. Boysen, B. Helwigh, and H. Rosenquist, "Effects of the Danish intervention strategies aimed at reducing Campylobacter in broiler meat," in *International Workshop on Campylobacter, Helicobacter and related organisms 14*, p. 1742, 2007.

[105] F. Georgsson, A. E. Thornorkelsson, M. Geirsdóttir, J. Reiersen, and N. J. Stern, "The influence of freezing and duration of storage on *Campylobacter* and indicator bacteria in broiler carcasses.," *Food Microbiology*, vol. 23, pp. 677–683, Oct. 2006.

[106] F. R. Jones, S. Baqar, A. Gozalo, G. Nunez, N. Espinoza, S. M. Reyes, M. Salazar, R. Meza, C. K. Porter, and S. E. Walz, "New World monkey Aotus nancymae as a model for Campylobacter jejuni infection and immunity.," *Infection and Immunity*, vol. 74, pp. 790–793, Jan. 2006.

[107] M. B. Skirrow, "Diseases due to Campylobacter, Helicobacter and related bacteria," *Journal of Comparative Pathology*, vol. 111, pp. 113–149, Aug. 1994.

[108] M. W. Odendaal, K. G. de Cramer, M. L. van der Walt, A. D. Botha, and P. M. Pieterson, "First isolation of *Campylobacter jejuni* from the vaginal discharge of three bitches after abortion in South Africa.," *The Onderstepoort journal of veterinary research*, vol. 61, pp. 193–195, June 1994.

[109] T. G. Hungerford, *Diseases of Livestock*. McGraw-Hill, 9th ed., 1990.

[110] O. R. Hedstrom, R. J. Sonn, E. D. Lassen, B. D. Hultgren, R. O. Crisman, B. B. Smith, and S. P. Snyder, "Pathology of *Campylobacter jejuni* abortion in sheep.," *Veterinary Pathology*, vol. 24, pp. 419–426, Sept. 1987.

[111] E. R. Burrough, O. Sahin, P. J. Plummer, Q. Zhang, and M. J. Yaeger, "Pathogenicity of an emergent, ovine abortifacient *Campylobacter jejuni* clone orally inoculated into pregnant guinea pigs.," *American Journal of Veterinary Research*, vol. 70, pp. 1269–1276, Oct. 2009.

[112] Z. Wu, R. Sippy, O. Sahin, P. Plummer, A. Vidal, D. G. Newell, and Q. Zhang, "Genetic diversity and antimicrobial susceptibility of Campylobacter jejuni isolates associated with sheep abortion in the United States and Great Britain.," *Journal of Clinical Microbiology*, vol. 52, pp. 1853–1861, June 2014.

[113] M. E. Konkel and L. A. Joens, "Effect of enteroviruses on adherence to and invasion of HEp-2 cells by *Campylobacter* isolates.," *Infection and Immunity*, vol. 58, pp. 1101–1105, Apr. 1990.

[114] G. Bukholm and G. Kapperud, "Expression of *Campylobacter jejuni* invasiveness in cell cultures coinfected with other bacteria.," *Infection and Immunity*, vol. 55, pp. 2816–2821, Nov. 1987.

[115] L. S. Mansfield, D. T. Gauthier, S. R. Abner, K. M. Jones, S. R. Wilder, and J. F. Urban, "Enhancement of disease and pathology by synergy of *Trichuris suis* and *Campylobacter jejuni* in the colon of immunologically naive swine.," *The American journal of tropical medicine and hygiene*, vol. 68, pp. 70–80, Mar. 2003.

[116] J. A. Wagenaar, W. F. Jacobs-Reitsma, M. Hofshagen, and D. G. Newell, "Poultry colonization with *Campylobacter* and its control at the primary production level," in *Campylobacter, Third Edition* (I. Nachamkin, C. M. Szymanski, and M. J. Blaser, eds.), pp. 667–678, Washington, DC: American Society of Microbiology, Jan. 2008.

[117] S. L. Welkos, "Experimental gastroenteritis in newly-hatched chicks infected with *Campylobacter jejuni*," *Journal of Medical Microbiology*, vol. 18, no. 2, pp. 233–248, 1984.

[118] C. K. Smith, M. Abu-Oun, S. A. Cawthraw, T. J. Humphrey, L. Rothwell, P. Kaiser, P. A. Barrow, and M. A. Jones, "*Campylobacter* colonization of the chicken induces a proinflammatory response in mucosal tissues," *FEMS Immunology & Medical Microbiology*, vol. 54, pp. 114–121, Oct. 2008.

[119] S. Humphrey, G. Chaloner, K. Kemmett, N. Davidson, N. Williams, A. Kipar, T. J. Humphrey, and P. Wigley, "*Campylobacter jejuni* Is Not Merely a Commensal in Commercial Broiler Chickens and Affects Bird Welfare," *mBio*, vol. 5, pp. e01364–14–e01364–14, July 2014.

[120] J. L. Jennings, L. C. Sait, C. A. Perrett, C. Foster, L. K. Williams, T. J. Humphrey, and T. A. Cogan, "Campylobacter jejuni is associated with, but not sufficient to cause vibrionic hepatitis in chickens.," *Veterinary Microbiology*, vol. 149, pp. 193–199, Apr. 2011.

[121] M. Cooke, "Disease entities of farmed ratites in New Zealand," *Surveillance*, 1998.

[122] C. P. Stephens, S. L. W. On, and J. A. Gibson, "An outbreak of infectious hepatitis in commercially reared ostriches associated with *Campylobacter coli* and *Campylobacter jejuni*," *Veterinary Microbiology*, vol. 61, pp. 183–190, Mar. 1998.

[123] M. R. Khanna, S. P. Bhavsar, and B. P. Kapadnis, "Effect of temperature on growth and chemotactic behaviour of *Campylobacter jejuni*.," *Letters in Applied Microbiology*, vol. 43, no. 1, pp. 84–90, 2006.

[124] J. Waldenstrom, D. Axelsson-Olsson, B. Olsen, D. Hasselquist, P. Griekspoor, L. Jasson, S. Teneberg, L. Svensson, and P. Ellström, "*Campylobacter jejuni* colonization in wild birds: results from an infection experiment," *PLoS ONE*, 2010.

[125] J. L. Baily, G. Meric, S. Bayliss, G. Foster, S. E. Moss, E. Watson, B. Pascoe, J. Mikhail, R. Pizzi, R. J. Goldstone, D. G. E. Smith, K. Willoughby, A. J. Hall, S. K. Sheppard, and M. P. Dagleish, "Evidence of land-sea transfer of the zoonotic pathogen *Campylobacter* to a wildlife marine sentinel species.," *Molecular Ecology*, vol. 24, pp. 208–221, Nov. 2014.

[126] C. Murphy, C. Carroll, and K. N. Jordan, "Environmental survival mechanisms of the foodborne pathogen *Campylobacter jejuni*.," *Journal of Applied Microbiology*, vol. 100, no. 4, pp. 623–632, 2006.

[127] D. Axelsson-Olsson, J. Olofsson, L. Svensson, P. Griekspoor, J. Waldenstrom, P. Ellström, and

B. Olsen, "Amoebae and algae can prolong the survival of *Campylobacter* species in co-culture.," *Experimental parasitology*, vol. 126, pp. 59–64, Sept. 2010.

[128] A. Culotti and A. I. Packman, "Pseudomonas aeruginosa facilitates Campylobacter jejuni growth in biofilms under oxic flow conditions.," *FEMS Microbiology Ecology*, vol. 91, p. fiv136, Dec. 2015.

[129] C. Bronowski, C. E. James, and C. Winstanley, "Role of environmental survival in transmission of *Campylobacter jejuni.*," *FEMS Microbiology Letters*, vol. 356, pp. 8–19, July 2014.

[130] F. Boyen, V. Eeckhaut, F. Van Immerseel, F. Pasmans, R. Ducatelle, and F. Haesebrouck, "Quorum sensing in veterinary pathogens: mechanisms, clinical importance and future perspectives.," *Veterinary Microbiology*, vol. 135, no. 3-4, pp. 187–195, 2009.

[131] N. Y. Park, E. Y. Ro, H. J. Jo, K. S. Park, and K. S. Yoon, "Effect of Packaging and Temperature on Survival Kinetics of Campylobacter jejunion Precooked Chicken Breast," *Journal of Food Safety*, vol. 34, pp. n/a–n/a, Sept. 2014.

[132] C. Fitzgerald, A. D. Sails, and P. I. Fields, "*Campylobacter jejuni* strain variation," in *Campylobacter* (J. M. Ketley and M. E. Konkel, eds.), pp. 59–77, Horizon Scientific Press, 2005.

[133] C. Fitzgerald, L. O. Helsel, M. a. Nicholson, S. J. Olsen, D. L. Swerdlow, R. Flahart, J. Sexton, and P. I. Fields, "Evaluation of methods for subtyping *Campylobacter jejuni* during an outbreak involving a food handler.," *Journal of Clinical Microbiology*, vol. 39, pp. 2386–2390, July 2001.

[134] B. L. Pike and P. Guerry, "Global Distribution of *Campylobacter jejuni* Penner Serotypes: A Systematic Review," *PLoS ONE*, vol. 8, no. 6, p. e67375, 2013.

[135] I. Nachamkin, "Microbiologic approaches for studying *Campylobacter* species in patients with Guillain-Barré syndrome.," *The Journal of Infectious Diseases*, vol. 176 Suppl 2, pp. S106–14, Dec. 1997.

[136] S. K. Sheppard, N. M. McCarthy, D. Falush, and M. C. J. Maiden, "Convergence of *Campylobacter* species: implications for bacterial evolution.," *Science*, vol. 320, pp. 237–239, Apr. 2008.

[137] E. N. Taboada, C. G. Clark, E. L. Sproston, and C. D. Carrillo, "Current methods for molecular typing of *Campylobacter* species.," *Journal of Microbiological Methods*, vol. 95, pp. 24–31, Oct. 2013.

[138] K. A. Jolley and M. C. J. Maiden, "Using MLST to study bacterial variation: prospects in the genomic era.," *Future Microbiology*, vol. 9, no. 5, pp. 623–630, 2014.

[139] A. van Belkum, P. T. Tassios, L. Dijkshoorn, S. Haeggman, B. Cookson, N. K. Fry, V. Fussing, J. Green, E. J. Feil, P. Gerner-Smidt, and M. Struelens, "Guidelines for the validation and application of typing methods for use in bacterial epidemiology.," *Clinical Microbiology and Infection*, vol. 13 Suppl 3, pp. 1–46, Aug. 2007.

[140] S. L. W. On, "Identification methods for *Campylobacters, Helicobacters*, and related organisms," *Clinical Microbiology Reviews*, vol. 9, no. 3, pp. 405–239, 1996.

[141] A. D. Sails, B. Swaminathan, and P. I. Fields, "Clonal Complexes of *Campylobacter jejuni* Identified by Multilocus Sequence Typing Correlate with Strain Associations Identified by Multilocus Enzyme Electrophoresis," *Journal of Clinical Microbiology*, vol. 41, pp. 4058–4067, Sept. 2003.

[142] M. J. Struelens and Members of the European Study Group on Epidemiological Markers, "Consensus guidelines for appropriate use and evaluation of microbial epidemiologic typing systems.," *Clinical Microbiology and Infection*, vol. 2, pp. 2–11, Aug. 1996.

[143] M. B. Skirrow and J. Benjamin, "Differentiation of enteropathogenic *Campylobacter*.," *Journal of Clinical Pathology*, vol. 33, pp. 1122–1122, Nov. 1980.

[144] G. A. Hebert, D. G. Hollis, R. E. Weaver, M. A. Lambert, M. J. Blaser, and C. W. Moss, "30 years of campylobacters: biochemical characteristics and a biotyping proposal for *Campylobacter* jejuni.," *Journal of Clinical Microbiology*, vol. 15, pp. 1065–1073, June 1982.

[145] H. Lior, "New, extended biotyping scheme for *Campylobacter jejuni, Campylobacter coli*, and "*Campylobacter laridis*"," *Journal of Clinical Microbiology*, vol. 20, pp. 636–640, Oct. 1984.

[146] F. J. Bolton, A. V. Holt, and D. N. Hutchinson, "*Campylobacter* biotyping scheme of epidemiological value.," *Journal of Clinical Pathology*, vol. 37, pp. 677–681, June 1984.

[147] M. B. Huysmans, J. D. Turnidge, and J. H. Williams, "Evaluation of API Campy in comparison with conventional methods for identification of thermophilic *campylobacters*.," *Journal of Clinical Microbiology*, vol. 33, pp. 3345–3346, Dec. 1995.

[148] J. E. Moore and R. H. Madden, "Comparison of eight phenotypic methods for subspecies characterization of thermophilic *Campylobacter* spp. isolated from pig liver.," *Journal of Food Protection*, vol. 66, pp. 1079–1084, June 2003.

[149] M. E. Konkel, J. E. Christensen, A. M. Keech, M. R. Monteville, J. D. Klena, and S. G. Garvis, "Identification of a fibronectin-binding domain within the *Campylobacter jejuni Cad*F protein.," *Molecular Microbiology*, vol. 57, pp. 1022–1035, Aug. 2005.

[150] G. H. Nuttall, "Progress report upon the biological test for blood as applied to over 500 bloods from various sources, together with a preliminary note upon a method for measuring the degree of reaction.," *British Medical Journal*, vol. 1, pp. 825–827, Apr. 1902.

[151] D. Linton, A. V. Karlyshev, and B. W. Wren, "Deciphering *Campylobacter jejuni* cell surface interactions from the genome sequence," *Current Opinion in Microbiology*, vol. 4, pp. 35–40, Jan. 2001.

[152] J. L. Penner, "The genus *Campylobacter*: a decade of progress.," *Clinical Microbiology Reviews*, vol. 1, pp. 157–172, Apr. 1988.

[153] J. L. Penner and J. N. Hennessy, "Passive hemagglutination technique for serotyping *Campylobacter fetus* subsp. *jejuni* on the basis of soluble heat-stable antigens.," *Journal of Clinical Microbiology*, vol. 12, pp. 732–737, Dec. 1980.

[154] A. V. Karlyshev, D. Linton, N. A. Gregson, and A. J. Lastovica, "Genetic and biochemical evidence of a *Campylobacter jejuni* capsular polysaccharide that accounts for Penner serotype specificity.," *Molecular Microbiology*, vol. 35, pp. 529–541, Feb. 2000.

[155] H. Lior, D. L. Woodward, J. A. Edgar, L. J. Laroche, and P. Gill, "Serotyping of *Campylobacter*

jejuni by slide agglutination based on heat-labile antigenic factors.," *Journal of Clinical Microbiology*, vol. 15, pp. 761–768, May 1982.

[156] R. A. Alm, P. Guerry, M. E. Power, H. Lior, and T. J. Trust, "Analysis of the role of flagella in the heat-labile Lior serotyping scheme of thermophilic *Campylobacters* by mutant allele exchange.," *Journal of Clinical Microbiology*, vol. 29, pp. 2438–2445, Nov. 1991.

[157] C. M. Patton, T. J. Barrett, and G. K. Morris, "Comparison of the Penner and Lior methods for serotyping *Campylobacter* spp.," *Journal of Clinical Microbiology*, vol. 22, pp. 558–565, Oct. 1985.

[158] D. L. Woodward and F. G. Rodgers, "Identification of *Campylobacter* heat-stable and heat-labile antigens by combining the Penner and Lior serotyping schemes," *Journal of Clinical Microbiology*, vol. 40, no. 3, pp. 741–745, 2002.

[159] C. J. Jackson, D. M. Jones, D. R. Wareing, and D. N. Hutchinson, "Associations between heat-stable (O) and heat-labile (HL) serogroup antigens of *Campylobacter jejuni*: evidence for interstrain relationships within three O/HL serovars.," *Journal of Clinical Microbiology*, vol. 36, pp. 2223–2228, Aug. 1998.

[160] E. M. Nielsen, J. Engberg, V. Fussing, L. Petersen, C. H. Brogren, and S. L. W. On, "Evaluation of phenotypic and genotypic methods for subtyping *Campylobacter jejuni* isolates from humans, poultry, and cattle.," *Journal of Clinical Microbiology*, vol. 38, pp. 3800–3810, Oct. 2000.

[161] J. A. Frost, A. N. Oza, R. T. Thwaites, and B. Rowe, "Serotyping scheme for *Campylobacter jejuni* and *Campylobacter coli* based on direct agglutination of heat-stable antigens.," *Journal of Clinical Microbiology*, vol. 36, pp. 335–339, Feb. 1998.

[162] D. McKay, J. Fletcher, P. Cooper, and F. M. Thomson-Carter, "Comparison of two methods for serotyping *Campylobacter* spp.," *Journal of Clinical Microbiology*, vol. 39, pp. 1917–1921, May 2001.

[163] D. G. Newell, J. A. Frost, B. Duim, J. A. Wagenaar, R. H. Madden, J. van der Plas, and S. L. W. On, "New developments in the subtyping of *Campylobacter* species," in *Campylobacter, Third Edition* (I. Nachamkin and M. J. Blaser, eds.), New York: American Society of Microbiology, Jan. 2000.

[164] S. L. W. On, H. I. Atabay, J. E. Corry, C. S. Harrington, and P. Vandamme, "Emended description of *Campylobacter sputorum* and revision of its infrasubspecific (biovar) divisions, including *C. sputorum* biovar paraureolyticus, a urease-producing variant from cattle and humans.," *International Journal of Systematic Bacteriology*, vol. 48 Pt 1, pp. 195–206, Jan. 1998.

[165] K. Kersters and J. De Ley, "Identification and grouping of bacteria by numerical analysis of their electrophoretic protein patterns.," *Journal of General Microbiology*, vol. 87, pp. 333–342, Apr. 1975.

[166] D. A. Ferguson and D. W. Lambe, "Differentiation of *Campylobacter* species by protein banding patterns in polyacrylamide slab gels.," *Journal of Clinical Microbiology*, vol. 20, pp. 453–460, Sept. 1984.

[167] J. D. Klena, New Zealand Ministry of Health, and New Zealand Ministry of Agriculture and Forestry,

"A Survey of Phenotypic and Genetic Methods Used to Identify and Differentiate Thermotolerant *Campylobacter* Spp. Strains," Ministry of Health, 2001.

[168] P. Boerlin, "Applications of multilocus enzyme electrophoresis in medical microbiology," *Journal of Microbiological Methods*, vol. 28, no. 3, pp. 221–231, 1997.

[169] R. J. Meinersmann, C. M. Patton, G. M. Evins, I. K. Wachsmuth, and P. I. Fields, "Genetic diversity and relationships of *Campylobacter* species and subspecies," *International Journal Of Systematic And Evolutionary Microbiology*, vol. 52, no. 5, pp. 1789–1797, 2002.

[170] B. A. Grajewski, J. W. Kusek, and H. M. Gelfand, "Development of a bacteriophage typing system for *Campylobacter jejun*i and *Campylobacter coli*.," *Journal of Clinical Microbiology*, vol. 22, pp. 13–18, July 1985.

[171] S. M. Salama and D. N. Hutchinson, "Application of a new phagetyping scheme to *Campylobacters* isolated during outbreaks.," *Epidemiology and Infection*, vol. 104, pp. 405–411, June 1990.

[172] R. Khakhria and H. Lior, "Extended phage-typing scheme for *Campylobacter jejuni* and *Campylobacter coli*.," *Epidemiology and Infection*, vol. 108, pp. 403–414, June 1992.

[173] J. A. Frost, J. M. Kramer, and S. A. Gillanders, "Phage typing of *Campylobacter jejuni* and *Campylobacter coli* and its use as an adjunct to serotyping.," *Epidemiology and Infection*, vol. 123, pp. 47–55, Aug. 1999.

[174] D. F. Welch, "Microbial Identification by GLC," *Clinical Microbiology Newsletter*, vol. 17, no. 14, pp. 105–108, 1995.

[175] J. J. Bright, M. A. Claydon, M. Soufian, and D. B. Gordon, "Rapid typing of bacteria using matrix-assisted laser desorption ionisation time-of-flight mass spectrometry and pattern recognition software.," *Journal of Microbiological Methods*, vol. 48, pp. 127–138, Feb. 2002.

[176] E. Bessède, O. Solecki, E. Sifré, L. Labadi, and F. Megraud, "Identification of *Campylobacter* species and related organisms by matrix assisted laser desorption ionization-time of flight (MALDI-TOF) mass spectrometry," *Clinical Microbiology and Infection*, vol. 17, pp. 1735–1739, Mar. 2011.

[177] A. E. Zautner, W. O. Masanta, M. Weig, U. Gro, and O. Bader, "Discrimination of multilocus sequence typing-based *Campylobacter jejuni* subgroups by MALDI-TOF mass spectrometry," *BMC Microbiology*, vol. 13, pp. 1–1, Nov. 2013.

[178] S. L. W. On, "Taxonomy, Phylogeny, and methods for the identification of *Campylobacter* species," in *Campylobacter* (J. M. Ketley and M. E. Konkel, eds.), pp. 13–42, Horizon Scientific Press, 2005.

[179] S. Kittler, S. Fischer, A. Abdulmawjood, G. Glunder, and G. Klein, "Colonisation of a Phage Susceptible *Campylobacter jejuni* Population in Two Phage Positive Broiler Flocks," *PLoS ONE*, vol. 9, p. e94782, Apr. 2014.

[180] D. Martiny, A. Visscher, B. Catry, S. Chatellier, and O. Vandenberg, "Optimization of *Campylobacter* growth conditions for further identification by matrix-assisted laser desorption/ionization time-of-flight

mass spectrometry (MALDI-TOF MS)," *Journal of Microbiological Methods*, vol. 94, pp. 221–223, Sept. 2013.

[181] R. J. Swain and M. M. Stevens, "Raman microspectroscopy for non-invasive biochemical analysis of single cells.," *Biochemical Society transactions*, vol. 35, pp. 544–549, June 2007.

[182] D. S. Read, D. J. Woodcock, N. J. C. Strachan, K. J. Forbes, F. M. Colles, M. C. J. Maiden, F. Clifton-Hadley, A. Ridley, A. Vidal, J. Rodgers, A. S. Whiteley, and S. K. Sheppard, "Evidence for phenotypic plasticity among multihost *Campylobacter jejuni* and *C. coli* lineages, obtained using ribosomal multilocus sequence typing and Raman spectroscopy.," *Applied and Environmental Microbiology*, vol. 79, pp. 965–973, Feb. 2013.

[183] Z. Elharrif and F. Mégraud, "Characterization of thermophilic *Campylobacter*. II. Enzymatic profiles," *Current Microbiology*, vol. 13, pp. 317–322, Nov. 1986.

[184] B. R. Bochner, "Global phenotypic characterization of bacteria," *FEMS Microbiology Reviews*, vol. 33, no. 1, pp. 191–205, 2009.

[185] B. R. Bochner and M. A. Savageau, "Generalized indicator plate for genetic, metabolic, and taxonomic studies with microorganisms.," *Applied and Environmental Microbiology*, vol. 33, pp. 434–444, Feb. 1977.

[186] B. R. Bochner, P. Gadzinski, and E. Panomitros, "Phenotype microarrays for high-throughput phenotypic testing and assay of gene function.," *Genome Research*, vol. 11, pp. 1246–1255, July 2001.

[187] J. E. Line, K. L. Hiett, J. Guard-Bouldin, and B. S. Seal, "Differential carbon source utilization by *Campylobacter jejuni* 11168 in response to growth temperature variation.," *Journal of Microbiological Methods*, vol. 80, pp. 198–202, Feb. 2010.

[188] D. R. Weerakoon and J. W. Olson, "The *Campylobacter jejuni* NADH:ubiquinone oxidoreductase (complex I) utilizes flavodoxin rather than NADH.," *Journal of Bacteriology*, vol. 190, no. 3, pp. 915–925, 2008.

[189] F. Denizot and R. Lang, "Rapid colorimetric assay for cell growth and survival: modifications to the tetrazolium dye procedure giving improved sensitivity and reliability," *Journal of Immunological Methods*, vol. 89, no. 2, pp. 271–277, 1986.

[190] G. G. Rodriguez, D. Phipps, K. Ishiguro, and H. F. Ridgway, "Use of a fluorescent redox probe for direct visualization of actively respiring bacteria.," *Applied and Environmental Microbiology*, vol. 58, pp. 1801–1808, June 1992.

[191] A. Shea, M. Wolcott, S. Daefler, and D. A. Rozak, "Biolog phenotype microarrays," in *Microbial Systems Biology* (A. Navid, ed.), Totowa, NJ: Humana Press, 2012.

[192] C. McCluskey, J. P. Quinn, and J. W. McGrath, "An evaluation of three new-generation tetrazolium salts for the measurement of respiratory activity in activated sludge microorganisms.," *Microbial Ecology*, vol. 49, no. 3, pp. 379–387, 2005.

[193] P. B. Hatzinger, P. Palmer, R. L. Smith, C. T. Peñarrieta, and T. Yoshinari, "Applicability of

tetrazolium salts for the measurement of respiratory activity and viability of groundwater bacteria.," *Journal of Microbiological Methods*, vol. 52, pp. 47–58, Jan. 2003.

[194] V. Créach, A. C. Baudoux, G. Bertru, and B. L. Rouzic, "Direct estimate of active bacteria: CTC use and limitations.," *Journal of Microbiological Methods*, vol. 52, pp. 19–28, Jan. 2003.

[195] J. L. Tholozan, J.-M. Cappelier, J. P. Tissier, G. Delattre, and M. Federighi, "Physiological characterization of viable-but-nonculturable *Campylobacter jejuni* cells.," *Applied and Environmental Microbiology*, vol. 65, pp. 1110–1116, Mar. 1999.

[196] L. A. I. Vaas, J. Sikorski, V. Michael, H.-P. Klenk, and M. Göker, "Visualization and curve-parameter estimation strategies for efficient exploration of phenotype microarray kinetics.," *PLoS ONE*, vol. 7, no. 4, p. e34846, 2012.

[197] M. Kahm, G. Hasenbrink, H. Lichtenberg-Frate, J. Ludwig, and M. Kschischo, "Grofit: Fitting biological growth curves," *Journal Of Statistical Software*, vol. 33, Feb. 2010.

[198] R. N. Xavier, H. W. Morgan, I. R. McDonald, and H. Withers, "Effect of long-term starvation on the survival, recovery, and carbon utilization profiles of a bovine *Escherichia coli* O157:H7 isolate from New Zealand.," *Applied and Environmental Microbiology*, vol. 80, pp. 4383–4390, July 2014.

[199] L. A. I. Vaas, J. Sikorski, B. Hofner, A. Fiebig, N. Buddruhs, H.-P. Klenk, and M. Göker, "opm: an R package for analysing OmniLog(R) phenotype microarray data.," *Bioinformatics*, vol. 29, pp. 1823–1824, July 2013.

[200] P. Gawand, L. Yang, W. R. Cluett, and R. Mahadevan, "Metabolic Model Refinement Using Phenotypic Microarray Data," in *Systems Metabolic Engineering* (H. S. Alper, ed.), pp. 47–59, Totowa, NJ: Humana Press, Jan. 2013.

[201] A. Metris, M. Reuter, D. J. Gaskin, J. Baranyi, and A. H. van Vliet, "In vivo and in silico determination of essential genes of *Campylobacter jejuni*," *BMC Genomics*, vol. 12, p. 535, Jan. 2014.

[202] M. Galardini, A. Mengoni, E. G. Biondi, R. Semeraro, A. Florio, M. Bazzicalupo, A. Benedetti, and S. Mocali, "DuctApe: A suite for the analysis and correlation of genomic and OmniLog™ Phenotype Microarray data.," *BMC Genomics*, vol. 103, pp. 1–10, Jan. 2014.

[203] M. Kanehisa, S. Goto, Y. Sato, M. Kawashima, M. Furumichi, and M. Tanabe, "Data, information, knowledge and principle: back to metabolism in KEGG.," *Nucleic Acids Research*, vol. 42, pp. D199–205, Jan. 2014.

[204] M. Kanehisa, S. Goto, S. Kawashima, Y. Okuno, and M. Hattori, "The KEGG resource for deciphering the genome.," *Nucleic Acids Research*, vol. 32, no. Database issue, pp. D277–80, 2004.

[205] S. Okuda, T. Yamada, M. Itoh, T. Katayama, P. Bork, S. Goto, and M. Kanehisa, "KEGG Atlas mapping for global analysis of metabolic pathways.," *Nucleic Acids Research*, vol. 36, no. Web Server issue, pp. W423–6, 2008.

[206] J. Y. H. Tang, J. Carlson, F. Mohamad Ghazali, A. A. Saleha, M. Nishibuchi, Y. Nakaguchi, and S. Radu, "Phenotypic MicroArray (PM) profiles (carbon sources and sensitivity to osmolytes and pH)

of *Campylobacter jejuni* ATCC 33560 in response to temperature," *International Journal of Food Research*, vol. 17, pp. 837–844, 2010.

[207] S. Wagley, J. Newcombe, E. Laing, E. Yusuf, C. M. Sambles, D. J. Studholme, R. M. La Ragione, R. W. Titball, and O. L. Champion, "Differences in carbon source utilisation distinguish *Campylobacter jejuni* from *Campylobacter coli*.," *BMC Microbiology*, vol. 14, p. 262, Oct. 2014.

[208] M. T. Brandl, A. F. Haxo, A. H. Bates, and R. E. Mandrell, "Comparison of Survival of Campylobacter jejuni in the Phyllosphere with That in the Rhizosphere of Spinach and Radish Plants," *Applied and Environmental Microbiology*, vol. 70, pp. 1182–1189, Feb. 2004.

[209] D. E. Fouts, E. F. Mongodin, R. E. Mandrell, W. G. Miller, D. A. Rasko, J. Ravel, L. M. Brinkac, R. T. DeBoy, C. T. Parker, S. C. Daugherty, R. J. Dodson, A. S. Durkin, R. Madupu, S. A. Sullivan, J. U. Shetty, M. A. Ayodeji, A. Shvartsbeyn, M. C. Schatz, J. H. Badger, C. M. Fraser, and K. E. Nelson, "Major Structural Differences and Novel Potential Virulence Mechanisms from the Genomes of Multiple *Campylobacter* Species," *PLoS Biology*, vol. 3, no. 1, p. e15, 2005.

[210] W. T. Muraoka and Q. Zhang, "Phenotypic and genotypic evidence for L-fucose utilization by *Campylobacter jejuni*.," *Journal of Bacteriology*, vol. 193, pp. 1065–1075, Feb. 2011.

[211] T. Duong and M. E. Konkel, "Comparative studies of *Campylobacter jejuni* genomic diversity reveal the importance of core and dispensable genes in the biology of this enigmatic food-borne pathogen.," *Current Opinion in Biotechnology*, vol. 20, pp. 158–165, Apr. 2009.

[212] D. Hofreuter, "Defining the metabolic requirements for the growth and colonization capacity of *Campylobacter jejuni*.," *Frontiers in Cellular and Infection Microbiology*, vol. 4, p. 137, 2014.

[213] R. V. Goering, "Pulsed field gel electrophoresis: A review of application and interpretation in the molecular epidemiology of infectious disease," *Infection*, vol. 10, no. 7, pp. 866–875, 2010.

[214] R. A. Stoddard, W. G. Miller, J. E. Foley, J. Lawrence, F. M. D. Gulland, P. A. Conrad, and B. A. Byrne, "*Campylobacter insulaenigrae* isolates from northern elephant seals (*Mirounga angustirostris*) in California.," *Applied and Environmental Microbiology*, vol. 73, no. 6, pp. 1729–1735, 2007.

[215] J. B. Ursing, H. Lior, and R. J. Owen, "Proposal of minimal standards for describing new species of the family Campylobacteraceae.," *International Journal of Systematic Bacteriology*, vol. 44, pp. 842–845, Oct. 1994.

[216] T. Ezaki, N. Takeuchi, S. L. Liu, A. Kai, H. Yamamoto, and E. Yabuuchi, "Small-scale DNA preparation for rapid genetic identification of *Campylobacter* species without radioisotope.," *Microbiology and immunology*, vol. 32, no. 2, pp. 141–150, 1988.

[217] J. M. Logan, K. J. Edwards, N. A. Saunders, and J. Stanley, "Rapid identification of *Campylobacter* spp. by melting peak analysis of biprobes in real-time PCR.," *Journal of Clinical Microbiology*, vol. 39, pp. 2227–2232, June 2001.

[218] M. Richter and R. Rossello-Mora, "Shifting the genomic gold standard for the prokaryotic species

definition," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, pp. 19126–19131, Nov. 2009.

[219] W. C. Bradbury, M. A. Marko, J. N. Hennessy, and J. L. Penner, "Occurrence of plasmid DNA in serologically defined strains of *Campylobacter jejuni* and *Campylobacter coli.*," *Infection and Immunity*, vol. 40, pp. 460–463, May 1983.

[220] F. C. Tenover, S. Williams, K. P. Gordon, N. Harris, C. Nolan, and J. J. Pwrde, "Utility of Plasmid Fingerprinting for Epidemiological Studies of *Campylobacter jejuni* Infections," *Journal of Infectious Diseases*, vol. 149, pp. 279–279, Feb. 1984.

[221] A. Fayos, R. J. Owen, J. Hernandez, C. Jones, and A. Lastovica, "Molecular subtyping by genome and plasmid analysis of *Campylobacter jejuni* serogroups O1 and O2 (Penner) from sporadic and outbreak cases of human diarrhoea," *Epidemiology and Infection*, vol. 111, no. 03, p. 415, 1993.

[222] C. M. Patton, I. K. Wachsmuth, G. M. Evins, J. a. Kiehlbauch, B. D. Plikaytis, N. Troup, L. Tompkins, and H. Lior, "Evaluation of 10 methods to distinguish epidemic-associated *Campylobacter* strains.," *Journal of Clinical Microbiology*, vol. 29, pp. 680–688, Apr. 1991.

[223] H. O. Smith and K. W. Wilcox, "A restriction enzyme from *Hemophilus influenzae*. I. Purification and general properties.," *Journal of Molecular Biology*, vol. 51, pp. 379–391, July 1970.

[224] K. Danna and D. Nathans, "Specific cleavage of simian virus 40 DNA by restriction endonuclease of *Hemophilus influenzae*," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 68, pp. 2913–2917, Dec. 1971.

[225] R. J. Owen, M. Costas, and C. Dawson, "Application of different chromosomal DNA restriction digest fingerprints to specific and subspecific identification of Campylobacter isolates.," *Journal of Clinical Microbiology*, vol. 27, pp. 2338–2343, Oct. 1989.

[226] V. Korolik, P. J. Coloe, and V. Krishnapillai, "A specific DNA probe for the identification of *Campylobacter jejuni.*," *Journal of General Microbiology*, vol. 134, pp. 521–529, Feb. 1988.

[227] C. K. Kakoyiannis, P. J. Winter, and R. B. Marshall, "The relationship between intestinal *Campylobacter* species isolated from animals and humans as determined by BRENDA.," *Epidemiology and Infection*, vol. 100, pp. 379–387, June 1988.

[228] C. K. Kakoyiannis, P. J. Winter, and R. B. Marshall, "Identification of *Campylobacter coli* isolates from animals and humans by bacterial restriction endonuclease DNA analysis.," *Applied and Environmental Microbiology*, vol. 48, pp. 545–549, Sept. 1984.

[229] D. C. Schwartz and C. R. Cantor, "Separation of yeast chromosome-sized DNAs by pulsed field gradient gel electrophoresis," *Cell*, 1984.

[230] Y. Suzuki, M. Ishihara, M. Saito, N. Ishikawa, and T. Yokochi, "Discrimination by means of pulsed-field gel electrophoresis between strains of *Campylobacter jejuni* Lior type 4 derived from sporadic cases and from outbreaks of infection," *The Journal of infection*, vol. 29, pp. 183–187, Sept. 1994.

[231] R. Karenlampi, H. Rautelin, and M.-L. Hänninen, "Evaluation of genetic markers and molecular

typing methods for prediction of sources of *Campylobacter jejuni* and *C. coli* infections.," *Applied and Environmental Microbiology*, vol. 73, pp. 1683–1685, Mar. 2007.

[232] S. M. McTavish and D. M. Campbell, "Multilocus sequence typing of *Campylobacter jejuni,* and the correlation between clonal complex and pulsed-field gel electrophoresis macrorestriction profile.," *FEMS Microbiology Letters*, vol. 298, no. 2, pp. 149–156, 2009.

[233] D. Wasyl, A. ElSedawy, and S. Lukinmaa, "PulseNet Europe - international molecular subtyping network for food-borne disease surveillance.," *Medycyna Weterynaryjna*, vol. 64, no. 2, pp. 123–126, 2008.

[234] F. C. Tenover, R. D. Arbeit, R. V. Goering, P. A. Mickelsen, B. E. Murray, and D. H. Persing, "Interpreting chromosomal DNA restriction patterns produced by pulsed-field gel electrophoresis: criteria for bacterial strain typing.," *Journal of Clinical Microbiology*, vol. 33, pp. 2233–2239, Sept. 1995.

[235] P. de Boer, J. A. Wagenaar, R. P. Achterberg, and B. Duim, "Generation of *Campylobacter jejuni* genetic diversity *in vivo*.," *Molecular Microbiology*, vol. 44, no. 2, pp. 351–359, 2002.

[236] W. Li, D. Raoult, and P.-E. Fournier, "Bacterial strain typing in the genomic era.," *FEMS Microbiology Reviews*, vol. 33, pp. 892–916, Sept. 2009.

[237] G. Manfreda, A. De Cesare, and V. Bondioli, "Ribotyping characterisation of *Campylobacter* isolates randomly collected from different sources in Italy," *Diagnostic Microbiology and Infectious Disease*, vol. 47, no. 2, pp. 385–392, 2003.

[238] E. Bingen, E. Denamur, and J. Elion, "Use of ribotyping in epidemiological surveillance of nosocomial outbreaks.," *Clinical Microbiology Reviews*, vol. 7, pp. 311–327, July 1994.

[239] M. Zabeau and P. Vos, "Selective restriction fragment amplification: a general method for dna fingerprinting," Apr. 1993.

[240] P. Vos, R. Hogers, M. Bleeker, M. Reijans, T. van de Lee, M. Hornes, A. Frijters, J. Pot, J. Peleman, and M. Kuiper, "AFLP: a new technique for DNA fingerprinting.," *Nucleic Acids Research*, vol. 23, pp. 4407–4414, Nov. 1995.

[241] B. A. Lindstedt, E. Heir, T. Vardund, K. K. Melby, and G. Kapperud, "Comparative fingerprinting analysis of *Campylobacter jejuni* subsp. jejuni strains by amplified-fragment length polymorphism genotyping.," *Journal of Clinical Microbiology*, vol. 38, pp. 3379–3387, Sept. 2000.

[242] L. M. Schouls, S. Reulen, B. Duim, J. A. Wagenaar, R. J. L. Willems, K. E. Dingle, F. M. Colles, and J. D. A. Van Embden, "Comparative Genotyping of *Campylobacter jejuni* by Amplified Fragment Length Polymorphism, Multilocus Sequence Typing, and Short Repeat Sequencing: Strain Diversity, Host Range, and Recombination," *Journal of Clinical Microbiology*, vol. 41, pp. 15–26, Jan. 2003.

[243] B. A. Lindstedt, E. Heir, T. Vardund, and G. Kapperud, "Fluorescent amplified-fragment length polymorphism genotyping of *Salmonella enterica* subsp. *enterica* serovars and comparison with

pulsed-field gel electrophoresis typing.," *Journal of Clinical Microbiology*, vol. 38, pp. 1623–1627, Apr. 2000.

[244] S. Zhao, S. E. Mitchell, J. Meng, and S. Kresovich, "Genomic typing of *Escherichia coli* O157: H7 by semi-automated fluorescent AFLP analysis," *Microbes and Infection*, 2000.

[245] J. Versalovic, T. Koeuth, and J. R. Lupski, "Distribution of repetitive DNA sequences in eubacteria and application to fingerprinting of bacterial genomes.," *Nucleic Acids Research*, vol. 19, pp. 6823–6831, Dec. 1991.

[246] H. P. Endtz, J. S. Vliegenthart, P. Vandamme, H. W. Weverink, N. P. van den Braak, H. A. Verbrugh, and A. van Belkum, "Genotypic diversity of *Campylobacter lari* isolated from mussels and oysters in The Netherlands.," *International Journal of Food Microbiology*, vol. 34, pp. 79–88, Jan. 1997.

[247] T. M. Wassenaar and B. Geilhausen, "Evidence of genomic instability in *Campylobacter jejuni* isolated from poultry.," *Applied and Environmental Microbiology*, vol. 64, pp. 1816–1821, May 1998.

[248] R. J. Meinersmann, L. O. Helsel, P. I. Fields, and K. L. Hiett, "Discrimination of *Campylobacter jejuni* isolates by *fla* gene sequencing.," *Journal of Clinical Microbiology*, vol. 35, pp. 2810–2814, Nov. 1997.

[249] I. Nachamkin, K. Bohachick, and C. M. Patton, "Flagellin gene typing of *Campylobacter jejuni* by restriction fragment length polymorphism analysis.," *Journal of Clinical Microbiology*, vol. 31, pp. 1531–1536, June 1993.

[250] A. J. Cody, M. C. J. Maiden, and K. E. Dingle, "Genetic diversity and stability of the porA allele as a genetic marker in human *Campylobacter* infection," *Microbiology*, vol. 155, no. 12, pp. 4145–4154, 2009.

[251] C. C. Grant, M. E. Konkel, W. Cieplak, and L. S. Tompkins, "Role of flagella in adherence, internalization, and translocation of *Campylobacter jejuni* in nonpolarized and polarized epithelial cell cultures.," *Infection and Immunity*, vol. 61, pp. 1764–1771, May 1993.

[252] C. S. Harrington, F. M. Thomson-Carter, and P. E. Carter, "Evidence for recombination in the flagellin locus of *Campylobacter jejuni:* implications for the flagellin gene typing scheme.," *Journal of Clinical Microbiology*, vol. 35, pp. 2386–2392, Sept. 1997.

[253] L. Petersen and S. L. W. On, "Efficacy of flagellin gene typing for epidemiological studies of *Campylobacter jejuni* in poultry estimated by comparison with macrorestriction profiling," *Letters in Applied Microbiology*, vol. 31, pp. 14–19, Dec. 2001.

[254] K. T. Elvers, C. R. Helps, T. M. Wassenaar, V. M. Allen, and D. G. Newell, "Development of a strain-specific molecular method for quantitating individual campylobacter strains in mixed populations.," *Applied and Environmental Microbiology*, vol. 74, pp. 2321–2331, Apr. 2008.

[255] K. E. Dingle, F. M. Colles, D. Falush, and M. C. J. Maiden, "Sequence typing and comparison of population biology of *Campylobacter coli* and *Campylobacter jejuni*.," *Journal of Clinical Microbiology*, vol. 43, pp. 340–347, Jan. 2005.

[256] S. Fujimoto, B. M. Allos, N. Misawa, C. M. Patton, and M. J. Blaser, "Restriction fragment length polymorphism analysis and random amplified polymorphic DNA analysis of *Campylobacter jejuni* strains isolated from patients with Guillain-Barré syndrome.," *The Journal of Infectious Diseases*, vol. 176, pp. 1105–1108, Oct. 1997.

[257] T. Zorman, M. Heyndrickx, S. Uzunović-Kamberović, and S. Smole Mozina, "Genotyping of Campylobacter coli and *C. jejuni* from retail chicken meat and humans with campylobacteriosis in Slovenia and Bosnia and Herzegovina.," *International Journal of Food Microbiology*, vol. 110, pp. 24–33, July 2006.

[258] B. Duim, T. M. Wassenaar, a. Rigter, and J. A. Wagenaar, "High-resolution genotyping of *Campylobacter* strains isolated from poultry and humans with amplified fragment length polymorphism fingerprinting.," *Applied and Environmental Microbiology*, vol. 65, pp. 2369–2375, June 1999.

[259] B. M. Korczak, M. Zurfluh, S. Emler, J. Kuhn-Oertli, and P. Kuhnert, "Multiplex strategy for multilocus sequence typing, *fla* typing, and genetic determination of antimicrobial resistance of *Campylobacter jejuni* and *Campylobacter coli* isolates collected in Switzerland.," *Journal of Clinical Microbiology*, vol. 47, pp. 1996–2007, July 2009.

[260] T. M. Wassenaar and R. Alonso, "Comparison of *Campylobacter* fla-SVR genotypes isolated from humans and poultry in three European regions.," *Letters in Applied Microbiology*, vol. 49, no. 3, pp. 388–395, 2009.

[261] H. Deveau, J. E. Garneau, and S. Moineau, "CRISPR/Cas system and its role in phage-bacteria interactions.," *Annual Review of Microbiology*, vol. 64, no. 1, pp. 475–493, 2010.

[262] R. Louwen, D. Horst-Kreft, A. G. de Boer, L. van der Graaf, G. de Knegt, M. Hamersma, A. P. Heikema, A. R. Timms, B. C. Jacobs, J. A. Wagenaar, H. P. Endtz, J. van der Oost, J. M. Wells, E. E. S. Nieuwenhuis, A. H. M. van Vliet, P. T. J. Willemsen, P. van Baarlen, and A. van Belkum, "A novel link between *Campylobacter jejuni* bacteriophage defence, virulence and Guillain-Barré syndrome.," *European Journal of Clinical Microbiology & Infectious Diseases*, vol. 32, pp. 207–226, Feb. 2013.

[263] E. P. Price, H. Smith, F. Huygens, and P. M. Giffard, "High-resolution DNA melt curve analysis of the clustered, regularly interspaced short-palindromic-repeat locus of *Campylobacter jejuni.*," *Applied and Environmental Microbiology*, vol. 73, no. 10, pp. 3431–3436, 2007.

[264] S. M. Kovanen, R. I. Kivisto, M. Rossi, and M.-L. Hänninen, "A combination of MLST and CRISPR typing reveals dominant *Campylobacter jejuni* types in organically farmed laying hens.," *Journal of Applied Microbiology*, vol. 117, pp. 249–257, July 2014.

[265] K. E. Dingle, F. M. Colles, D. R. Wareing, R. Ure, A. J. Fox, F. E. Bolton, H. J. Bootsma, R. J. Willems, R. Urwin, and M. C. J. Maiden, "Multilocus sequence typing system for *Campylobacter jejuni.*," *Journal of Clinical Microbiology*, vol. 39, pp. 14–23, Jan. 2001.

[266] F. M. Colles and M. C. J. Maiden, "*Campylobacter* sequence typing databases: applications and future prospects.," *Microbiology*, vol. 158, pp. 2695–2709, Nov. 2012.

[267] S. Suerbaum, M. Lohrengel, A. Sonnevend, F. Ruberg, and M. Kist, "Allelic Diversity and Recombination in *Campylobacter jejuni*," *Journal of Bacteriology*, vol. 183, no. 8, pp. 2553–2559, 2001.

[268] S. Kittl, G. Heckel, B. M. Korczak, and P. Kuhnert, "Source attribution of human *Campylobacter* isolates by MLST and fla-typing and association of genotypes with quinolone resistance," *PLoS ONE*, 2013.

[269] P. Lang, T. Lefébure, W. Wang, P. Pavinski Bitar, R. J. Meinersmann, K. Kaya, and M. J. Stanhope, "Expanded multilocus sequence typing and comparative genomic hybridization of *Campylobacter coli* isolates from multiple hosts.," *Applied and Environmental Microbiology*, vol. 76, pp. 1913–1925, Mar. 2010.

[270] M. C. J. Maiden, J. a. Bygraves, E. Feil, G. Morelli, J. E. Russell, R. Urwin, Q. Zhang, J. Zhou, K. Zurth, D. A. Caugant, I. M. Feavers, M. Achtman, and B. G. Spratt, "Multilocus sequence typing: A portable approach to the identification of clones within populations of pathogenic microorganisms," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 95, pp. 3140–3145, Mar. 1998.

[271] M. Pérez-Losada, E. B. Browne, A. Madsen, T. Wirth, R. P. Viscidi, and K. A. Crandall, "Population genetics of microbial pathogens estimated from multilocus sequence typing (MLST) data.," *Infection, genetics and evolution*, vol. 6, pp. 97–112, Mar. 2006.

[272] K. E. Dingle, F. M. Colles, R. Ure, J. A. Wagenaar, B. Duim, F. J. Bolton, A. J. Fox, D. R. A. Wareing, and M. C. J. Maiden, "Molecular characterization of *Campylobacter jejuni* clones: a basis for epidemiologic investigation.," *Emerging Infectious Diseases*, vol. 8, pp. 949–955, Sept. 2002.

[273] P. J. Biggs, P. Fearnhead, G. Hotter, V. Mohan, J. Collins-Emerson, E. Kwan, T. E. Besser, A. Cookson, P. E. Carter, and N. P. French, "Whole-genome comparison of two *Campylobacter jejuni* isolates of the same sequence type reveals multiple loci of different ancestral lineage.," *PLoS ONE*, vol. 6, no. 11, p. e27121, 2011.

[274] S. J. Salipante, D. J. SenGupta, L. A. Cummings, T. A. Land, D. R. Hoogestraat, and B. T. Cookson, "Application of whole genome sequencing for bacterial strain typing in molecular epidemiology.," *Journal of Clinical Microbiology*, pp. JCM.03385–14, Jan. 2015.

[275] E. Castro-Nallar, N. A. Hasan, T. A. Cebula, R. R. Colwell, R. A. Robison, W. E. Johnson, and K. A. Crandall, "Concordance and discordance of sequence survey methods for molecular epidemiology," *PeerJ*, vol. 3, no. 8, pp. e761–24, 2015.

[276] M. V. Larsen, S. Cosentino, O. Lukjancenko, D. Saputra, S. Rasmussen, H. Hasman, T. Sicheritz-Pontén, F. M. Aarestrup, D. W. Ussery, and O. Lund, "Benchmarking of Methods for Genomic Taxonomy," *Journal of Clinical Microbiology*, vol. 52, pp. 1529–1539, Apr. 2014.

[277] K. A. Jolley and M. C. J. Maiden, "BIGSdb: Scalable analysis of bacterial genome variation at the population level," *BMC Bioinformatics*, vol. 11, no. 1, p. 595, 2010.

[278] M. L. Metzker, "Sequencing technologies - the next generation.," *Nature Reviews Genetics*, vol. 11, no. 1, pp. 31–46, 2010.

[279] J. R. Miller, S. Koren, and G. Sutton, "Assembly algorithms for next-generation sequencing data," *BMC Genomics*, vol. 95, pp. 315–327, June 2010.

[280] M. Hunt, C. Newbold, M. Berriman, and T. D. Otto, "A comprehensive evaluation of assembly scaffolding tools," *Genome Biology*, vol. 15, no. 3, p. R42, 2014.

[281] J. A. Martin and Z. Wang, "Next-generation transcriptome assembly," *Nature Reviews Genetics*, 2011.

[282] X. Didelot, R. Bowden, D. J. Wilson, T. E. Peto, and D. W. Crook, "Transforming clinical microbiology with bacterial genome sequencing.," *Nature Reviews Genetics*, vol. 13, pp. 601–612, Sept. 2012.

[283] D. W. Eyre, T. Golubchik, N. C. Gordon, R. Bowden, P. Piazza, E. M. Batty, C. L. C. Ip, D. J. Wilson, X. Didelot, L. O'Connor, R. Lay, D. Buck, A. M. Kearns, A. Shaw, J. Paul, M. H. Wilcox, P. J. Donnelly, T. E. Peto, A. S. Walker, and D. W. Crook, "A pilot study of rapid benchtop sequencing of Staphylococcus aureus and Clostridium difficile for outbreak detection and surveillance," *BMJ Open*, vol. 2, pp. e001124–e001124, Jan. 2012.

[284] B. G. Hall, G. D. Ehrlich, and F. Z. Hu, "Pan-genome analysis provides much higher strain typing resolution than multi-locus sequence typing," *Microbiology*, vol. 156, pp. 1060–1068, Mar. 2010.

[285] J. Garaizar, A. Rementeria, and S. Porwollik, "DNA microarray technology: a new tool for the epidemiological typing of bacterial pathogens?," *FEMS Immunology & Medical Microbiology*, vol. 47, pp. 178–189, July 2006.

[286] B. M. Pearson, C. Pin, J. Wright, K. I'Anson, T. J. Humphrey, and J. M. Wells, "Comparative genome analysis of *Campylobacter jejuni* using whole genome DNA microarrays," *FEBS Letters*, vol. 554, no. 1-2, pp. 224–230, 2003.

[287] L. G. Pittenger, J. G. Frye, V. McNerney, J. Reeves, J. Haro, P. J. Fedorka-Cray, M. A. Harrison, and M. D. Englen, "Analysis of *Campylobacter jejuni* whole-genome DNA microarrays: significance of prophage and hypervariable regions for discriminating isolates.," *Foodborne Pathogens and Disease*, vol. 9, pp. 473–479, May 2012.

[288] N. Dorrell, J. A. Mangan, K. G. Laing, J. Hinds, D. Linton, H. Al-Ghusein, B. G. Barrell, J. Parkhill, N. G. Stoker, A. V. Karlyshev, P. D. Butcher, and B. W. Wren, "Whole genome comparison of *Campylobacter jejuni* human isolates using a low-cost microarray reveals extensive genetic diversity.," *Genome Research*, vol. 11, pp. 1706–1715, Oct. 2001.

[289] O. L. Champion, M. W. Gaunt, O. Gundogdu, A. Elmi, A. A. Witney, J. Hinds, N. Dorrell, and B. W. Wren, "Comparative phylogenomics of the food-borne pathogen *Campylobacter jejuni* reveals genetic markers predictive of infection source.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, pp. 16043–16048, Nov. 2005.

[290] N. Yutin, P. Puigbò, E. V. Koonin, and Y. I. Wolf, "Phylogenomics of prokaryotic ribosomal proteins.," *PLoS ONE*, vol. 7, no. 5, p. e36972, 2012.

[291] M. Wu and J. a. Eisen, "A simple, fast, and accurate method of phylogenomic inference.," *Genome Biology*, vol. 9, no. 10, p. R151, 2008.

[292] A. J. Cody, N. M. McCarthy, M. Jansen van Rensburg, T. Isinkaye, S. D. Bentley, J. Parkhill, K. E. Dingle, I. C. J. W. Bowler, K. A. Jolley, and M. C. J. Maiden, "Real-time genomic epidemiological evaluation of human *Campylobacter* isolates by use of whole-genome multilocus sequence typing.," *Journal of Clinical Microbiology*, vol. 51, pp. 2526–2534, Aug. 2013.

[293] K. A. Jolley, C. M. Bliss, J. S. Bennett, H. B. Bratcher, C. Brehony, F. M. Colles, H. Wimalarathna, O. B. Harrison, S. K. Sheppard, A. J. Cody, and M. C. J. Maiden, "Ribosomal multilocus sequence typing: universal characterization of bacteria from domain to strain.," *Microbiology*, vol. 158, pp. 1005–1015, Apr. 2012.

[294] M. Achtman, K. Zurth, G. Morelli, G. Torrea, A. Guiyoule, and E. Carniel, "*Yersinia pestis*, the cause of plague, is a recently emerged clone of *Yersinia pseudotuberculosis*.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 96, pp. 14043–14048, Nov. 1999.

[295] S. Sreevatsan, X. Pan, K. E. Stockbauer, N. D. Connell, B. N. Kreiswirth, T. S. Whittam, and J. M. Musser, "Restricted structural gene polymorphism in the *Mycobacterium tuberculosis* complex indicates evolutionarily recent global dissemination.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 94, pp. 9869–9874, Sept. 1997.

[296] E. Helgason, O. A. Okstad, D. A. Caugant, H. A. Johansen, A. Fouet, M. Mock, I. Hegna, and A. B. Kolstø, "Bacillus anthracis, Bacillus cereus, and Bacillus thuringiensis–one species on the basis of genetic evidence.," *Applied and Environmental Microbiology*, vol. 66, pp. 2627–2630, June 2000.

[297] S. K. Sheppard, K. A. Jolley, and M. C. J. Maiden, "A Gene-By-Gene Approach to Bacterial Population Genomics: Whole Genome MLST of *Campylobacter*.," *Genes*, vol. 3, no. 2, pp. 261–277, 2012.

[298] M. C. J. Maiden, M. J. Jansen van Rensburg, J. E. Bray, S. G. Earle, S. A. Ford, K. A. Jolley, and N. M. McCarthy, "MLST revisited: the gene-by-gene approach to bacterial genomics.," *Nature Reviews Microbiology*, vol. 11, pp. 728–736, Oct. 2013.

[299] C. Camacho, G. Coulouris, V. Avagyan, N. Ma, J. Papadopoulos, K. Bealer, and T. L. Madden, "BLAST+: architecture and applications.," *BMC Bioinformatics*, vol. 10, no. 1, p. 421, 2009.

[300] J. Revez, A.-K. Llarena, T. Schott, M. Kuusi, M. Hakkinen, R. Kivistö, M.-L. Hänninen, and M. Rossi, "Genome analysis of *Campylobacter jejuni* strains isolated from a waterborne outbreak.," *BMC Genomics*, vol. 15, p. 768, Sept. 2014.

[301] S. M. Kovanen, R. I. Kivistö, M. Rossi, T. Schott, U.-M. Kärkkäinen, T. Tuuminen, J. Uksila, H. Rautelin, and M.-L. Hänninen, "Multilocus sequence typing (MLST) and whole-genome MLST of *Campylobacter jejuni* isolates from human infections in three districts during a seasonal peak in Finland.," *Journal of Clinical Microbiology*, vol. 52, pp. 4147–4154, Dec. 2014.

[302] M. D. Cao, S. Balasubramanian, and M. Bodén, "Sequencing technologies and tools for short tandem repeat variation detection," *bib.oxfordjournals.org*, Feb. 2014.

[303] C. Alkan, S. Sajjadian, and E. E. Eichler, "Limitations of next-generation genome sequence assembly," *Nature Methods*, vol. 8, pp. 61–65, Nov. 2010.

[304] S. Bao, R. Jiang, W. Kwan, B. Wang, X. Ma, and Y.-Q. Song, "Evaluation of next-generation sequencing software in mapping and assembly," *Journal of Human Genetics*, vol. 56, pp. 406–414, Apr. 2011.

[305] C. Bates, K. L. Hiett, and N. J. Stern, "Relationship of *Campylobacter* isolated from poultry and from darkling beetles in New Zealand," *Avian Diseases*, vol. 48, no. 1, pp. 138–147, 2004.

[306] E. L. Sproston, I. D. Ogden, M. MacRae, K. J. Forbes, J. F. Dallas, S. K. Sheppard, A. J. Cody, F. M. Colles, M. J. Wilson, and N. J. C. Strachan, "Multi-locus sequence types of *Campylobacter* carried by flies and slugs acquired from local ruminant faeces.," *Journal of Applied Microbiology*, vol. 109, no. 3, pp. 829–838, 2010.

[307] C. Thomas, H. Gibson, D. J. Hill, and M. Mabey, "Campylobacter epidemiology: an aquatic perspective.," *Journal of Applied Microbiology*, vol. 85 Suppl 1, pp. 168S–177S, Dec. 1998.

[308] G. L. Nichols, "Fly transmission of Campylobacter.," *Emerging Infectious Diseases*, vol. 11, pp. 361–364, Mar. 2005.

[309] C. A. Whittier, M. R. Cranfield, and M. K. Stoskopf, "Real-time PCR detection of *Campylobacter* spp. In free-ranging mountain gorillas (*Gorilla beringei beringei*).," *Journal of Wildlife diseases*, vol. 46, pp. 791–802, July 2010.

[310] R. Khoshbakht, M. Tabatabaei, H. Shirzad Aski, and H. Shayegh, "Distribution of *Salmonella, Arcobacter*, and thermophilic *Campylobacter* spp. among Persian fallow deer (*Dama mesopotamica*) population in Dasht-e-Arzhan Wildlife refuge, southern Iran," *Comparative Clinical Pathology*, Aug. 2014.

[311] M. Koziel, P. O'Doherty, P. Vandamme, G. D. Corcoran, R. D. Sleator, and B. Lucey, "*Campylobacter corcagiensis* sp. nov., isolated from faeces of captive lion-tailed macaques (*Macaca silenus*)," *International Journal Of Systematic And Evolutionary Microbiology*, vol. 64, pp. 2878–2883, Aug. 2014.

[312] L. Debruyne, T. Broman, S. Bergström, B. Olsen, S. L. W. On, and P. Vandamme, "*Campylobacter subantarcticus* sp. nov., isolated from birds in the sub-Antarctic region," *International Journal Of Systematic And Evolutionary Microbiology*, vol. 60, pp. 815–819, Apr. 2010.

[313] L. Debruyne, T. Broman, S. Bergström, B. Olsen, S. L. W. On, and P. Vandamme, "*Campylobacter volucris* sp. nov., isolated from black-headed gulls (*Larus ridibundus*)," *International Journal Of Systematic And Evolutionary Microbiology*, vol. 60, pp. 1870–1875, Aug. 2010.

[314] M. J. Gilbert, M. Kik, W. G. Miller, B. Duim, and J. A. Wagenaar, "*Campylobacter iguaniorum* sp.

nov., isolated from reptiles.," *International Journal Of Systematic And Evolutionary Microbiology*, Jan. 2015.

[315] N. M. McCarthy, F. M. Colles, K. E. Dingle, M. C. Bagnall, G. Manning, M. C. J. Maiden, and D. Falush, "Host-associated genetic import in *Campylobacter jejuni*.," *Emerging Infectious Diseases*, vol. 13, pp. 267–272, Feb. 2007.

[316] W. G. Miller and I. V. Wesley, "Identification of host-associated alleles by multilocus sequence typing of Campylobacter coli strains from food animals Printed in Great Britain," *Microbiology*, vol. 152, no. 1, pp. 245–255, 2006.

[317] M. C. J. Maiden, "Multilocus sequence typing of bacteria.," *Annual Review of Microbiology*, vol. 60, no. 1, pp. 561–588, 2006.

[318] F. M. Colles, K. E. Dingle, A. J. Cody, and M. C. J. Maiden, "Comparison of *Campylobacter* populations in wild geese with those in starlings and free-range poultry on the same farm.," *Applied and Environmental Microbiology*, vol. 74, pp. 3583–3590, June 2008.

[319] S. K. Sheppard, F. M. Colles, J. Richardson, A. J. Cody, R. Elson, A. Lawson, G. Brick, R. Meldrum, C. L. Little, R. J. Owen, M. C. J. Maiden, and N. M. McCarthy, "Host association of *Campylobacter* genotypes transcends geographic variation.," *Applied and Environmental Microbiology*, vol. 76, pp. 5269–5277, Aug. 2010.

[320] L. Petersen, E. M. Nielsen, J. Engberg, S. L. W. On, and H. H. Dietz, "Comparison of genotypes and serotypes of *Campylobacter jejuni* isolated from Danish wild mammals and birds and from broiler flocks and humans.," *Applied and Environmental Microbiology*, vol. 67, pp. 3115–3121, July 2001.

[321] T. Broman, H. Palmgren, S. Bergström, M. Sellin, J. Waldenstrom, M. L. Danielsson-Tham, and B. Olsen, "*Campylobacter jejuni* in Black-Headed Gulls (*Larus ridibundus*): Prevalence, Genotypes, and Influence on *C. jejuni* Epidemiology," *Journal of Clinical Microbiology*, vol. 40, pp. 4594–4602, Dec. 2002.

[322] O. Rosef, G. Kapperud, S. Lauwers, and B. Gondrosen, "Serotyping of *Campylobacter jejuni, Campylobacter coli, and Campylobacter laridis* from domestic and wild animals," *Applied and Environmental Microbiology*, vol. 49, no. 6, pp. 1507–1510, 1985.

[323] J. Waldenstrom, T. Broman, I. Carlsson, D. Hasselquist, R. P. Achterberg, J. A. Wagenaar, and B. Olsen, "Prevalence of *Campylobacter jejuni, Campylobacter lari*, and *Campylobacter coli* in Different Ecological Guilds and Taxa of Migrating Birds," *Applied and Environmental Microbiology*, vol. 68, no. 12, pp. 5911–5917, 2002.

[324] P. S. L. Kwan, C. Xavier, M. Santovenia, J. Pruckler, S. Stroika, K. Joyce, T. Gardner, P. I. Fields, J. McLaughlin, R. V. Tauxe, and C. Fitzgerald, "Multilocus Sequence Typing Confirms Wild Birds as the Source of a *Campylobacter* Outbreak Associated with the Consumption of Raw Peas," *Applied and Environmental Microbiology*, vol. 80, no. 15, pp. 4540–4546, 2014.

[325] T. Broman, J. Waldenstrom, D. Dahlgren, I. Carlsson, I. Eliasson, and B. Olsen, "Diversities and

similarities in PFGE profiles of *Campylobacter jejuni* isolated from migrating birds and humans," *Journal of Applied Microbiology*, vol. 96, pp. 834–843, Apr. 2004.

[326] Z. Hubálek, "An annotated checklist of pathogenic microorganisms associated with migratory birds.," *Journal of Wildlife diseases*, vol. 40, pp. 639–659, Oct. 2004.

[327] D. W. Crowder, E. A. Dykstra, J. M. Brauner, A. Duffy, C. Reed, E. Martin, W. Peterson, Y. Carrière, P. Dutilleul, and J. P. Owen, "West nile virus prevalence across landscapes is mediated by local effects of agriculture on vector and host communities.," *PLoS ONE*, vol. 8, no. 1, p. e55006, 2013.

[328] H. Chen, G. J. D. Smith, S. Y. Zhang, K. Qin, J. Wang, K. S. Li, R. G. Webster, J. S. M. Peiris, and Y. Guan, "Avian flu: H5N1 virus outbreak in migratory waterfowl," *Nature*, vol. 436, pp. 191–192, July 2005.

[329] L. Clark, "Disease risk posed by wild birds associated with agricultural landscapes," in *The Produce Contamination Problem* (G. M. Sapers, E. B. Solomon, and K. R. Mathews, eds.), pp. 139–165, Academic Press, Sept. 2014.

[330] P. Griekspoor, B. Olsen, and J. Waldenstrom, "*Campylobacter jejuni* in penguins, Antarctica," *Emerging Infectious Diseases*, vol. 15, no. 5, pp. 847–847, 2009.

[331] G. Kapperud and O. Rosef, "Avian wildlife reservoir of *Campylobacter fetus* subsp. *jejuni, Yersinia* spp., and *Salmonella* spp. in Norway.," *Applied and Environmental Microbiology*, vol. 45, pp. 375–380, Feb. 1983.

[332] H. Palmgren, M. Sellin, S. Bergström, and B. Olsen, "Enteropathogenic bacteria in migrating birds arriving in Sweden.," *Scandinavian journal of infectious diseases*, vol. 29, no. 6, pp. 565–568, 1997.

[333] L. A. Hughes, M. Bennett, P. Coffey, J. Elliott, T. R. Jones, R. C. Jones, A. Lahuerta-Marin, K. McNiffe, D. Norman, N. J. Williams, and J. Chantrey, "Molecular epidemiology and characterization of *Campylobacter* spp. isolated from wild bird populations in northern England.," *Applied and Environmental Microbiology*, vol. 75, pp. 6292–6298, Sept. 2009.

[334] S. L. W. On, "Isolation, identification and subtyping of *Campylobacter:* Where to from here? ," *Journal of Microbiological Methods*, vol. 95, pp. 3–7, Oct. 2013.

[335] J. Lu, H. Ryu, J. Vogel, J. Santo Domingo, and N. J. Ashbolt, "Molecular detection of *Campylobacter* spp. and fecal indicator bacteria during the northern migration of sandhill cranes (*Grus canadensis)* at the central Platte River.," *Applied and Environmental Microbiology*, vol. 79, pp. 3762–3769, June 2013.

[336] N. P. French, A. Midwinter, B. Holland, J. Collins-Emerson, R. Pattison, F. M. Colles, and P. E. Carter, "Molecular epidemiology of *Campylobacter jejuni* isolates from wild-bird fecal material in children's playgrounds.," *Applied and Environmental Microbiology*, vol. 75, pp. 779–783, Feb. 2009.

[337] F. Aydin, H. I. Atabay, and M. Akan, "The isolation and characterization of *Campylobacter jejuni* subsp. *jejuni* from domestic geese (*Anser anser*).," *Journal of Applied Microbiology*, vol. 90, pp. 637–642, Apr. 2001.

[338] S. Quessy and S. Messier, "Prevalence of *Salmonella* spp., *Campylobacter* spp. and *Listeria* spp. in ring-billed gulls (*Larus delawarensis*).," *Journal of Wildlife diseases*, vol. 28, pp. 526–531, Oct. 1992.

[339] K. Ito, Y. Kubokura, K.-i. Kaneko, Y. Totake, and M. Ogawa, "Occurence of *Campylobacter jejuni* in free-living wild birds from Japan," *Journal of Wildlife diseases*, vol. 24, pp. 467–470, Jan. 1988.

[340] V. Mohan, M. A. Stevenson, J. Marshall, P. Fearnhead, B. R. Holland, G. Hotter, and N. P. French, "*Campylobacter jejuni* colonization and population structure in urban populations of ducks and starlings in New Zealand.," *Microbiology Open*, vol. 2, pp. 659–673, Aug. 2013.

[341] G. Glunder, U. Neumann, S. Braune, J. Prüter, S. Petersen, and G. Vauk, "The occurrence of *Campylobacter* spp. and *Salmonella* spp. in gulls in northern Germany.," *DTW. Deutsche tierärztliche Wochenschrift*, vol. 98, pp. 152–155, Apr. 1991.

[342] J. I. Keller, W. G. Shriver, J. Waldenstrom, P. Griekspoor, and B. Olsen, "Prevalence of *Campylobacter* in wild birds of the mid-Atlantic region, USA.," *Journal of Wildlife diseases*, vol. 47, pp. 750–754, July 2011.

[343] G. Glunder, "Characterization of *Campylobacter* spp. from wild birds.," *Berliner und Munchener tierarztliche Wochenschrift*, vol. 102, pp. 49–52, Feb. 1989.

[344] N. A. Luechtefeld, M. J. Blaser, L. B. Reller, and W. L. Wang, "Isolation of *Campylobacter fetus* subsp. *jejuni* from migratory waterfowl.," *Journal of Clinical Microbiology*, vol. 12, pp. 406–408, Sept. 1980.

[345] M. Fukuyama, T. Kamimura, T. Itoh, K. Saito, M. Takahashi, S. Sakai, M. Murata, K. Kohzaki, M. Hara, and T. Shimizu, "Distribution of *Campylobacter jejuni* in wild birds and serogroup of isolates by slide agglutination technique.," *Nihon juigaku zasshi. The Japanese journal of veterinary science*, vol. 48, pp. 487–493, June 1986.

[346] R. E. Pacha, G. W. Clark, E. A. Williams, and A. M. Carter, "Migratory birds of central Washington as reservoirs of *Campylobacter jejuni*," *Canadian Journal of Microbiology*, vol. 34, pp. 80–82, Jan. 1988.

[347] S. N. Workman, G. E. Mathison, and M. C. Lavoie, "Pet dogs and chicken meat as reservoirs of Campylobacter spp. in Barbados.," *Journal of Clinical Microbiology*, vol. 43, pp. 2642–2650, June 2005.

[348] D. M. Fallacara, C. M. Monahan, T. Y. Morishita, and R. F. Wack, "Fecal Shedding and Antimicrobial Susceptibility of Selected Bacterial Pathogens and a Survey of Intestinal Parasites in Free-Living Waterfowl," *Avian Diseases*, vol. 45, p. 128, Jan. 2001.

[349] E. M. Moriarty, N. Karki, M. Mackenzie, L. W. Sinton, D. R. Wood, and B. J. Gilpin, "Faecal indicators and pathogens in selected New Zealand waterfowl," *New Zealand Journal of Marine and Freshwater Research*, vol. 45, pp. 679–688, Dec. 2011.

[350] M. E. Rutledge, R. M. Siletzky, and W. Gu, "Characterization of *Campylobacter* from resident Canada geese in an urban environment," *Journal of Wildlife diseases*, vol. 49, no. 1, pp. 1–9, 2013.

[351] H. Wahlström, E. Tysén, E. O. Engvall, B. Brändström, E. Eriksson, T. Morner, and I. Vågsholm, "Survey of *Campylobacter* species, VTEC O157 and *Salmonella* species in Swedish wildlife.," *The Veterinary Record*, vol. 153, pp. 74–80, July 2003.

[352] C. P. A. de Haan, K. Lampén, J. Corander, and M.-L. Hänninen, "Multilocus sequence types of environmental *Campylobacter jejuni* isolates and their similarities to those of human, poultry and bovine *C. jejuni* isolates," *Zoonoses and Public Health*, vol. 60, pp. 125–133, July 2012.

[353] J. Waldenstrom, S. L. W. On, and B. Olsen, "Species diversity of *Campylobacteria* in a wild bird community in Sweden.," *Journal of Applied Microbiology*, vol. 102, no. 2, pp. 424–432, 2007.

[354] R. Sippy, O. Sahin, P. Plummer, and Q. Zhang, "Occurrence and molecular analysis of *Campylobacter* in wildlife on livestock farms.," *Veterinary Microbiology*, vol. 157, no. 3-4, pp. 369–375, 2012.

[355] C. R. Fricker and N. Metcalfe, "*Campylobacters* in wading birds (Charadrii): incidence, biotypes and isolation techniques," *Zentralblatt für Bakteriologie, Mikrobiologie und Hygiene. 1. Abt. Originale B, Hygiene*, vol. 179, pp. 469–475, Oct. 1984.

[356] O. Olsvik and B. P. Berdal, "Survey of antibodies against *Francisella tularensis* and some enteropathogenic bacteria in the hooded crow *Corvus corone* cornix," *FAUNA NORVEGICA*, vol. 4, pp. 29–30, Jan. 1983.

[357] P. Robino, L. Tomassone, C. Tramuta, M. Rodo, M. Giammarino, G. Vaschetti, and P. Nebbia, "Prevalence of *Campylobacter jejuni, Campylobacter coli* and enteric *Helicobacter* in domestic and free living birds in North-Western Italy.," *Schweizer Archiv für Tierheilkunde*, vol. 152, pp. 425–431, Sept. 2010.

[358] H. Palmgren, T. Broman, J. Waldenstrom, P. Lindberg, A. Aspán, and B. Olsen, "*Salmonella Amager*, *Campylobacter jejuni*, and urease-positive thermophilic *Campylobacter* found in free-flying peregrine falcons (Falco peregrinus) in Sweden.," *Journal of Wildlife diseases*, vol. 40, pp. 583–587, July 2004.

[359] J. J. SACKS, S. Lieb, L. M. Baldy, S. Berta, C. M. Patton, M. C. White, W. J. Bigler, and J. J. Witte, "Epidemic campylobacteriosis associated with a community water supply.," *American Journal of Public Health*, vol. 76, pp. 424–428, Apr. 1986.

[360] C. D. Whelan, P. Monaghan, R. W. A. Girdwood, and C. R. Fricker, "The significance of wild birds (*Larus sp.*) in the epidemiology of *campylobacter* infections in humans," *Epidemiology and Infection*, vol. 101, p. 259, Apr. 1988.

[361] W. Sixl, R. Karpísková, Z. Hubálek, J. Halouzka, M. Mikulásková, and J. Salava, "*Campylobacter* spp. and *Salmonella* spp. in black-headed gulls (*Larus ridibundus*).," *Central European journal of public health*, vol. 5, pp. 24–26, Mar. 1997.

[362] L. Dipineto, T. P. Russo, A. Gargiulo, L. Borrelli, L. M. De Luca Bossa, A. Santaniello, P. Buonocore, L. F. Menna, and A. Fioretti, "Prevalence of enteropathogenic bacteria in common quail (*Coturnix coturnix*).," *Avian Pathology*, vol. 43, pp. 1–10, Sept. 2014.

[363] V. Mohan, "Faeco-prevalence of *Campylobacter jejuni* in urban wild birds and pets in New Zealand.," *BMC Research Notes*, vol. 8, p. 1, Feb. 2015.

[364] K. Yogasundram, S. M. Shane, and K. S. Harrington, "Prevalence of Campylobacter jejuni in Selected Domestic and Wild Birds in Louisiana," *Avian Diseases*, vol. 33, p. 664, Oct. 1989.

[365] L. Casanovas, M. de Simón, M. D. Ferrer, J. Arqués, and G. Monzón, "Intestinal carriage of campylobacters, salmonellas, yersinias and listerias in pigeons in the city of Barcelona.," *The Journal of Applied Bacteriology*, vol. 78, pp. 11–13, Jan. 1995.

[366] J. Cabrita, J. Rodrigues, F. Braganca, C. Morgado, I. Pires, and A. P. Gonçalves, "Prevalence, biotypes, plasmid profile and antimicrobial resistance of *Campylobacter* isolated from wild and domestic animals from Northeast Portugal," *Journal of Applied Bacteriology*, vol. 73, pp. 279–285, Mar. 1992.

[367] J. Bonnedahl, T. Broman, J. Waldenstrom, H. Palmgren, T. Niskanen, and B. Olsen, "In Search of Human-associated Bacterial Pathogens in Antarctic Wildlife: Report from Six Penguin Colonies Regularly Visited by Tourists," *Ambio*, vol. 34, pp. 430–432, Aug. 2005.

[368] P. J. Hepworth, K. E. Ashelford, J. Hinds, K. A. Gould, A. A. Witney, N. J. Williams, H. Leatherbarrow, N. P. French, R. J. Birtles, C. Mendonca, N. Dorrell, B. W. Wren, P. Wigley, N. Hall, and C. Winstanley, "Genomic variations define divergence of water/wildlife-associated *Campylobacter jejuni* niche specialists from common clonal complexes.," *Environmental Microbiology*, vol. 13, pp. 1549–1560, June 2011.

[369] P. J. Hepworth, H. Leatherbarrow, C. A. Hart, and C. Winstanley, "Use of suppression subtractive hybridisation to extend our knowledge of genome diversity in Campylobacter jejuni.," *BMC Genomics*, vol. 8, no. 1, p. 110, 2007.

[370] S. K. Sheppard, X. Didelot, G. Meric, A. Torralbo, K. A. Jolley, D. J. Kelly, S. D. Bentley, M. C. J. Maiden, J. Parkhill, and D. Falush, "Genome-wide association study identifies vitamin B5 biosynthesis as a host specificity factor in *Campylobacter*.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 110, pp. 11923–11927, July 2013.

[371] N. Dorrell, O. L. Champion, and B. W. Wren, "Advances in *Campylobacter jejuni* comparative genomics through whole genome DNA microarrays," in *Campylobacter* (J. M. Ketley and M. E. Konkel, eds.), pp. 79–99, Horizon Scientific Press, 2005.

[372] S. F. Park, "The physiology of *Campylobacter* species and its relevance to their role as foodborne pathogens.," *International Journal of Food Microbiology*, vol. 74, no. 3, pp. 177–188, 2002.

[373] A. Stintzi, "Gene Expression Profile of *Campylobacter jejuni* in Response to Growth Temperature Variation," *Journal of Bacteriology*, vol. 185, pp. 2009–2016, Mar. 2003.

[374] D. Robertshaw, "Temperature Regulation and Thermal Environment ," in *Dukes' physiology of domestic animals.* (O. W. Reece, ed.), New Delhi: Panima Publishing Corporation., Jan. 2004.

[375] W. R. Branson, J. H. Greg, and R. H. Linda, *Avian medicine (principles and application).* Lake Worth (Flo): Wingers Publishing, 1994.

[376] D. S. Farner, N. Chivers, and T. Riney, "The body temperatures of the North Island Kiwis.," *EMU*, vol. 56, no. 3, pp. 199–206, 1956.

[377] R. Prinzinger, A. Pressmar, and E. Schleucher, "Body temperature in birds," *Comparative Biochemistry and Physiology Part A: Physiology*, vol. 99, pp. 499–506, Jan. 1991.

[378] C. M. Blatteis, "Fever: pathological or physiological, injurious or beneficial?," *Journal of Thermal Biology*, vol. 28, no. 1, pp. 1–13, 2003.

[379] T. J. O'Shea, P. M. Cryan, A. A. Cunningham, A. R. Fooks, D. T. S. Hayman, A. D. Luis, A. J. Peel, R. K. Plowright, and J. L. N. Wood, "Bat flight and zoonotic viruses.," *Emerging Infectious Diseases*, vol. 20, pp. 741–745, May 2014.

[380] W. C. Hazeleger, W. F. Jacobs-Reitsma, P. H. C. Lina, M. A. P. van Bergen, T. Bosch, and R. R. Beumer, "Bugs in bats 1: ," in *11th conference of the Dutch Society for Wildlife*, (Utrecht), May 2014.

[381] I. U. H. Khan, S. Hill, E. Nowak, and T. A. Edge, "Effect of Incubation Temperature on the Detection of Thermophilic *Campylobacter* Species from Freshwater Beaches, Nearby Wastewater Effluents, and Bird Fecal Droppings," *Applied and Environmental Microbiology*, vol. 79, pp. 7639–7645, Nov. 2013.

[382] M. H. Zwietering, J. T. de Koos, B. E. Hasenack, J. C. de Witt, and K. van't Riet, "Modeling of bacterial growth as a function of temperature.," *Applied and Environmental Microbiology*, vol. 57, pp. 1094–1101, Apr. 1991.

[383] W. C. Hazeleger, J. A. Wouters, F. M. Rombouts, and T. Abee, "Physiological activity of *Campylobacter jejuni* far below the minimal growth temperature," *Journal of Clinical Microbiology*, vol. 64, pp. 3917–3922, Sept. 1998.

[384] C. O. Gill and L. M. Harris, "Limiting conditions of temperature and pH for growth of "thermophilic" *Campylobacters* on solid media," *Journal of Food Protection*, 1983.

[385] C. O. Gill and L. M. Harris, "Contamination of red-meat carcasses by *Campylobacter fetus* subsp. *jejuni*.," *Applied and Environmental Microbiology*, vol. 43, pp. 977–980, May 1982.

[386] T. W. Steele and R. J. Owen, "*Campylobacter jejnui* subsp. *doylei* sunbsp. nov., a sunspecies of Nitrate-Negative *Campylobacters* Isolated from Human Clinical Specimens," *International Journal of Systematic Bacteriology*, vol. 38, pp. 316–318, July 1988.

[387] A. F. Kelly, A. Martínez-Rodriguez, R. A. Bovill, and B. M. Mackey, "Description of a "Phoenix" Phenomenon in the Growth of *Campylobacter jejuni* at Temperatures Close to the Minimum for Growth," *Applied and Environmental Microbiology*, vol. 69, pp. 4975–4978, Aug. 2003.

[388] M. E. Konkel, B. J. Kim, J. D. Klena, C. R. Young, and R. Ziprin, "Characterization of the thermal stress response of *Campylobacter jejuni*.," *Infection and Immunity*, vol. 66, pp. 3666–3672, Aug. 1998.

[389] J. Parkhill, B. W. Wren, K. Mungall, J. M. Ketley, C. Churcher, D. Basham, T. Chillingworth, R. M. Davies, T. Feltwell, S. Holroyd, K. Jagels, A. V. Karlyshev, S. Moule, M. J. Pallen, C. W. Penn, M. A. Quail, M.-A. Rajandream, K. M. Rutherford, A. H. M. van Vliet, S. Whitehead, and B. G. Barrell, "The genome sequence of the food-borne pathogen *Campylobacter jejuni* reveals hypervariable sequences.," *Nature*, vol. 403, pp. 665–668, Feb. 2000.

[390] A. M. Brás and S. Chatterjee, "A novel *Campylobacter jejuni* two-component regulatory system important for temperature-dependent growth and colonization.," *Journal of Bacteriology*, vol. 181, pp. 3298–3302, May 1999.

[391] A. Klančnik, N. Botteldoorn, L. Herman, and S. S. Možina, "Survival and stress induced expression of *gro*EL and *rpo*D of *Campylobacter jejuni* from different growth phases.," *International Journal of Food Microbiology*, vol. 112, no. 3, pp. 200–207, 2006.

[392] P. Tangwatcharin, S. Chanthachum, P. Khopaibool, and M. W. Griffiths, "Morphological and physiological responses of Campylobacter jejuni to stress.," *Journal of Food Protection*, vol. 69, pp. 2747–2753, Jan. 2006.

[393] H. T. T. Nguyen, J. E. L. Corry, and C. A. Miles, "Heat resistance and mechanism of heat inactivation in thermophilic *Campylobacters*.," *Applied and Environmental Microbiology*, vol. 72, pp. 908–913, Jan. 2006.

[394] K. Papenfort and J. Vogel, "Regulatory RNA in Bacterial Pathogens," *Cell Host and Microbe*, vol. 8, pp. 116–127, July 2010.

[395] F. Narberhaus, "Translational control of bacterial heat shock and virulence genes by temperature-sensing mRNAs.," *RNA biology*, vol. 7, pp. 84–89, Jan. 2010.

[396] E. Nudler and A. S. Mironov, "The riboswitch control of bacterial metabolism," *Trends in Biochemical Sciences*, vol. 29, no. 1, pp. 11–17, 2004.

[397] M. Mandal, B. Boese, J. E. Barrick, W. C. Winkler, and R. R. Breaker, "Riboswitches control fundamental biochemical pathways in *Bacillus subtilis* and other bacteria.," *Cell*, vol. 113, pp. 577–586, May 2003.

[398] J. Kortmann and F. Narberhaus, "Bacterial RNA thermometers: molecular zippers and switches.," *Nature Reviews Microbiology*, vol. 10, pp. 255–265, Apr. 2012.

[399] F. Narberhaus, T. Waldminghaus, and S. Chowdhury, "RNA thermometers.," *FEMS Microbiology Reviews*, vol. 30, pp. 3–16, Jan. 2006.

[400] E. Loh, E. Kugelberg, A. Tracy, Q. Zhang, B. Gollan, H. Ewles, R. Chalmers, V. Pelicic, and C. M. Tang, "Temperature triggers immune evasion by *Neisseria meningitidis*," *Nature*, vol. 502, pp. 237–240, Sept. 2013.

[401] T. J. Humphrey, S. O'Brien, and M. Madsen, "Campylobacters as zoonotic pathogens: A food production perspective," *International Journal of Food Microbiology*, vol. 117, pp. 237–257, July 2007.

[402] S. Phadtare, "Recent developments in bacterial cold-shock response.," *Current issues in molecular biology*, vol. 6, pp. 125–136, July 2004.

[403] K. F. Chan, H. Le Tran, R. Y. Kanenaka, and S. Kathariou, "Survival of Clinical and Poultry-Derived Isolates of Campylobacter jejuni at a Low Temperature (4 C)," *Applied and Environmental Microbiology*, vol. 67, pp. 4186–4191, Sept. 2001.

[404] S. Phadtare, J. Alsina, and M. Inouye, "Cold-shock response and cold-shock proteins," *Current Opinion in Microbiology*, vol. 2, no. 2, pp. 175–180, 1999.

[405] S. L. Svensson and E. Friedrich, "Survival strategies of *Campylobacter jejuni:* Responses, the viable but non culturable state and biofilms," in *Campylobacter, Third Edition* (I. Nachamkin, C. M. Szymanski, and M. J. Blaser, eds.), pp. 571–590, Washington, DC: American Society of Microbiology, Jan. 2008.

[406] M. W. Qoronfleh, C. Debouck, and J. Keller, "Identification and characterization of novel low-temperature-inducible promoters of *Escherichia coli.*," *Journal of Bacteriology*, vol. 174, no. 24, pp. 7902–7909, 1992.

[407] J. T. Trevors, J. D. van Elsas, and L. Van Overbeek, "Bacterial gene expression at low temperatures.," *Extremophiles*, vol. 16, no. 2, pp. 167–176, 2012.

[408] C. S. Vegge, L. Brondsted, M. Ligowska-Marzketa, and H. Ingmer, "Natural Transformation of *Campylobacter jejuni* Occurs Beyond Limits of Growth," *PLoS ONE*, vol. 7, p. e45467, Sept. 2012.

[409] B. Moen, A. Oust, Ø. Langsrud, N. Dorrell, G. L. Marsden, J. Hinds, A. Kohler, B. W. Wren, and K. Rudi, "Explorative multifactor approach for investigating global survival mechanisms of *Campylobacter jejuni* under environmental conditions.," *Applied and Environmental Microbiology*, vol. 71, pp. 2086–2094, Apr. 2005.

[410] T. M. Fuchs, K. Neuhaus, and S. Scherer, "Life at Low Temperatures," in *The Prokaryotes*, pp. 375–420, Berlin, Heidelberg: Springer Berlin Heidelberg, 2013.

[411] R.-A. Hughes, K. Hallett, T. Cogan, M. Enser, and T. J. Humphrey, "The response of *Campylobacter jejuni* to low temperature differs from that of *Escherichia coli.*," *Applied and Environmental Microbiology*, vol. 75, pp. 6292–6298, Oct. 2009.

[412] N. W. Gunther IV, K. T. Rajkowski, and C. Sommers, "Survival after Cryogenic Freezing of *Campylobacter* Species in Ground Turkey Patties Treated with Polyphosphates," *Journal of Food Protection*, vol. 78, no. 2, pp. 419–423, 2015.

[413] W. J. Snelling, M. Matsuda, J. E. Moore, and J. S. G. Dooley, "Under the microscope *Campylobacter jejuni.*," *Letters in Applied Microbiology*, vol. 41, no. 4, pp. 297–302, 2005.

[414] F. M. Colles, K. Jones, R. M. Harding, and M. C. J. Maiden, "Genetic diversity of *Campylobacter jejuni* isolates from farm animals and the farm environment," *Applied and Environmental Microbiology*, 2003.

[415] P. Fearnhead, N. G. C. Smith, M. Barrigas, A. Fox, and N. P. French, "Analysis of recombination

in *Campylobacter jejuni* from MLST population data.," *Journal of Molecular Evolution*, vol. 61, pp. 333–340, Sept. 2005.

[416] T. Lefébure, P. D. P. Bitar, H. Suzuki, and M. J. Stanhope, "Evolutionary dynamics of complete *Campylobacter* pan-genomes and the bacterial species concept.," *Genome Biology and Evolution*, vol. 2, no. 0, pp. 646–655, 2010.

[417] J. M. Wilmshurst, A. J. Anderson, T. F. G. Higham, and T. H. Worthy, "Dating the late prehistoric dispersal of Polynesians to New Zealand using the commensal Pacific rat.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105, pp. 7676–7680, June 2008.

[418] R. H. Taylor, "What limits kiore (*Rattus exulans*) distribution in New Zealand?," *New Zealand Journal of Zoology*, vol. 2, no. 4, pp. 473–477, 1975.

[419] A. J. Anderson, "Advances in New Zealand mammalogy 1990-2000 : Polynesian dog or kuri," *Journal of the Royal Society of New Zealand*, vol. 31, no. 1, pp. 161–163, 2001.

[420] K.-J. Wilson, *Flight of the huia : ecology and conservation of New Zealand's frogs, reptiles, birds and mammals*. Christchurch, N.Z.: Canterbury University Press, 2003.

[421] S. K. Sheppard, L. Cheng, G. Meric, C. P. A. de Haan, A.-K. Llarena, P. Marttinen, A. Vidal, A. Ridley, F. Clifton-Hadley, T. R. Connor, N. J. C. Strachan, K. Forbes, F. M. Colles, K. A. Jolley, S. D. Bentley, M. C. J. Maiden, M.-L. Hänninen, J. Parkhill, W. P. Hanage, and J. Corander, "Cryptic ecology among host generalist *Campylobacter jejuni* in domestic animals.," *Molecular Ecology*, vol. 23, pp. 2442–2451, May 2014.

[422] K. E. Dingle and M. C. J. Maiden, "Population Genetics of *Campylobacter jejuni*," in *Campylobacter* (J. M. Ketley and M. E. Konkel, eds.), pp. 43–57, Horizon Scientific Press, 2005.

[423] D. Prangle, P. Fearnhead, M. P. Cox, P. J. Biggs, and N. P. French, "Semi-automatic selection of summary statistics for ABC model choice.," *Statistical applications in genetics and molecular biology*, vol. 13, pp. 67–82, Feb. 2014.

[424] K. Wodzicki and S. Wright, "Introduced Birds and Mammals in New Zealand and Their Effect on the Environment," *Tuatara : Journal of the Biological Society*, vol. Volume 27,, no. Issue 2,, 1984.

[425] P. Lemey, A. Rambaut, J. J. Welch, and M. A. Suchard, "Phylogeography takes a relaxed random walk in continuous space and time.," *Molecular Biology and Evolution*, vol. 27, no. 8, pp. 1877–1885, 2010.

[426] Y. Yamaoka, "Helicobacter pylori typing as a tool for tracking human migration.," *Clinical Microbiology and Infection*, vol. 15, pp. 829–834, Sept. 2009.

[427] B. Binney, P. J. Biggs, P. E. Carter, B. R. Holland, and N. P. French, "Quantification of historical livestock importation into New Zealand 1860-1979.," *New Zealand Veterinary Journal*, vol. 62, pp. 309–314, Nov. 2014.

[428] B. L. Evans, *A history of agricultural production and marketing in New Zealand*. Palmerston North,: Keeling & Mundy, 1st ed., 1969.

[429] R. L. Peden, *Sheep farming practice in colonial Canterbury 1843 to 1882*. PhD thesis, University of Canterbury, 2002.

[430] K. L. Hughes, "History of veterinary public health in Australasia.," *Revue scientifique et technique*, vol. 10, no. 4, pp. 1019–1040, 1991.

[431] J. A. Crump, D. R. Murdoch, and M. G. Baker, "Emerging infectious diseases in an island ecosystem: the New Zealand perspective.," *Emerging Infectious Diseases*, vol. 7, no. 5, pp. 767–772, 2001.

[432] C. Brown, "Emerging zoonoses and pathogens of public health significance–an overview.," *Revue scientifique et technique*, vol. 23, pp. 435–442, Aug. 2004.

[433] J. B. H. Martiny, B. J. M. Bohannan, J. H. Brown, R. K. Colwell, J. A. Fuhrman, J. L. Green, M. C. Horner-Devine, M. Kane, J. A. Krumins, C. R. Kuske, P. J. Morin, S. Naeem, L. Ovreas, A.-L. Reysenbach, V. H. Smith, and J. T. Staley, "Microbial biogeography: putting microorganisms on the map.," *Nature Reviews Microbiology*, vol. 4, pp. 102–112, Feb. 2006.

[434] R. D. Sleator, "Phylogenetics.," *Archives of Microbiology*, vol. 193, no. 4, pp. 235–239, 2011.

[435] B. T. Grenfell, O. G. Pybus, J. R. Gog, J. L. N. Wood, J. M. Daly, J. a. Mumford, and E. C. Holmes, "Unifying the epidemiological and evolutionary dynamics of pathogens.," *Science*, vol. 303, no. 5656, pp. 327–332, 2004.

[436] M. Stoneking and J. Krause, "Learning about human population history from ancient and modern genomes," *Nature Reviews Genetics*, vol. 12, pp. 603–614, Sept. 2011.

[437] G. P. Wallis and S. A. Trewick, "New Zealand phylogeography: evolution on a small continent.," *Molecular Ecology*, vol. 18, no. 17, pp. 3548–3580, 2009.

[438] C. Diez-Tascon, R. P. Littlejohn, P. A. Almeida, and A. M. Crawford, "Genetic variation within the Merino sheep breed: analysis of closely related populations using microsatellites.," *Animal Genetics*, vol. 31, pp. 243–251, Oct. 2000.

[439] J. Parkes, "Advances in New Zealand mammalogy 1990–2000: feral livestock," *Journal of the Royal Society of New Zealand*, vol. 31, no. 1, pp. 233–241, 2001.

[440] J. R. Fairweather, "White-Settler Colonial Development: Early New Zealand Pastoralism and the Formation of Estates1," *Journal of Sociology*, vol. 21, pp. 237–257, Jan. 1985.

[441] B. Patterson, "Laagers in the Wilderness ," *Stout Centre Review*, vol. 1, pp. 3–14, Apr. 1991.

[442] A. H. Clark, *The Invasion of New Zealand by People Plants and Animals*. The South Island, Rutgers University Press, 2nd printi ed., 1949.

[443] A. E. Pierce, "An historical review disease and quarantine of animal movement in New Zealand and Australia," *New Zealand Veterinary Journal*, vol. 23, no. 7, pp. 125–136, 1975.

[444] J. R. Fisher, "The growth of the livestock industries and the veterinary profession in Australia to 1850.," *Australian Veterinary Journal*, vol. 71, no. 8, pp. 248–253, 1994.

[445] R. P. Hargreaves, "Waimate: Pioneer New Zealand Farm," *Agricultural History*, vol. 36, no. 1, pp. 38–45, 1962.

[446] New Zealand Government, "No. IV. Sheep scab ordinance," pp. 397–400, Aug. 1849.

[447] G. M. Thomson, *The naturaliasation of animals and plants in New Zealand*. The Naturalisation of Animals and Plants in New Zealand, Cambridge, England.: Cambridge University Press, 1922.

[448] R. Wolfe, *A short history of sheep in New Zealand*. Auckland [N.Z.]: Random House New Zealand, 2006.

[449] Statistics New Zealand, *Statistical publications, 1840 - 2000*. Wellington, N.Z.: Statistics New Zealand, 2004.

[450] New Zealands Registrar Generals Office, *Statistics of New Zealand for 1862*. W.C. Wilson for the New Zealand Government, 1863.

[451] New Zealand Government Statistician's Office, "The New Zealand official year-book 1919," tech. rep., New Zealand Government, 1919.

[452] B. L. Evans, "Agricultural and pastoral statistics of New Zealand 1861-1954." R.E. Owen Govt. Printer, Wellington, N.Z., 1956.

[453] Statistics New Zealand, "Infoshare," tech. rep., Statistics New Zealand, 2011.

[454] M. Pearson and J. Lennon, *Pastoral Australia : Fortunes, Failures & Hard Yakka: A Historical Overview 1788-1967*. CSIRO PUBLISHING, Apr. 2010.

[455] T. A. Coghlan, *A Statistical account of the seven colonies of Australasia (1894)*. Statistician, Sydney: Charles Potter, Government Printer, 1894.

[456] A. N. Bruere, "Scrapie in New Zealand - Its history and what it could mean," in *Proc 15th Annual Seminar, Society of Sheep & Beef Cattle Veterinarians NZVA*, pp. 9–22, The society of Sheep and Beef Cattle Veterinarians of the New Zealand Veterinary Association, 1985.

[457] F. Hilbink, M. Penrose, E. Kovacova, and J. Kazar, "Q fever is absent from New Zealand," *International Journal of Epidemiology*, vol. 22, no. 5, pp. 945–949, 1993.

[458] P. Jaros, A. L. Cookson, D. M. Campbell, G. E. Duncan, D. Prattley, P. E. Carter, T. E. Besser, S. Shringi, S. Hathaway, J. C. Marshall, and N. P. French, "Geographic Divergence of Bovine and Human Shiga Toxin–Producing *Escherichia coli* O157:H7 Genotypes, New Zealand1," *Emerging Infectious Diseases*, vol. 20, pp. 1980–1989, Dec. 2012.

[459] N. Hunter, "Scrapie—Uncertainties, biology and molecular approaches," *Biochimica et Biophysica Acta*, vol. 1772, pp. 619–628, June 2007.

[460] A. N. Bruere, "Scrapie freedom - the New Zealand story," *Surveillance*, vol. 30, no. 4, pp. 3–7, 2003.

[461] N. Hunter and D. Cairns, "Scrapie-free Merino and Poll Dorset sheep from Australia and New Zealand have normal frequencies of scrapie-susceptible PrP genotypes.," *The Journal of General Virology*, vol. 79, no. 8, pp. 2079–2082, 1998.

[462] A. Bossers, F. L. Harders, and M. A. Smits, "PrP genotype frequencies of the most dominant sheep breed in a country free from scrapie," *Archives of Virology*, vol. 144, no. 4, pp. 829–834, 1999.

[463] E. F. Houston, S. I. Halliday, M. Jeffrey, W. Goldmann, and N. Hunter, "New Zealand sheep with scrapie-susceptible PrP genotypes succumb to experimental challenge with a sheep-passaged scrapie isolate (SSBP/1).," *The Journal of General Virology*, vol. 83, pp. 1247–1250, May 2002.

[464] R. H. Foote, "The history of artificial insemination: Selected notes and notables," *Journal of Animal Science*, vol. 80, no. 2, pp. 1–10, 2002.

[465] B. Cooper-Blanks, "The New Zealand poultry meat industry: An education and industry resource." Poultry Industry Association of New Zealand, Auckland, Jan. 1999.

[466] W. R. B. Stone, "The poultry industry in New Zealand," *New Zealand Veterinary Journal*, vol. 20, pp. 142–145, Aug. 1972.

[467] A. J. Cody, N. M. McCarthy, J. E. Bray, H. M. L. Wimalarathna, F. M. Colles, M. J. Rensburg, K. E. Dingle, J. Waldenstrom, and M. C. J. Maiden, "Wild bird associated Campylobacter jejuni isolates are a consistent source of human disease, in Oxfordshire, United Kingdom," *Environmental microbiology reports*, pp. n/a–n/a, June 2015.

[468] EFSA Panel on Biological Hazards., "Scientific Opinion on a review on the European Union Summary Reports on trends and sources zoonoses , zoonotic agents and food-borne outbreaks in 2009 and 2010 – specifically for the data on *Salmonella, Campylobacter* , verotoxigenic *Escherichia coli , Listeria monocytogenes* and foodborne outbreaks," *EFSA Journal*, vol. 10, no. 6, p. 2726, 2012.

[469] M. Stahl, J. Butcher, and A. Stintzi, "Nutrient acquisition and metabolism by *Campylobacter jejuni*," *Frontiers in Cellular and Infection Microbiology*, vol. 2, pp. 1–10, 2012.

[470] S. Cleaveland, M. K. Laurenson, and L. H. Taylor, "Diseases of humans and their domestic mammals: pathogen characteristics, host range and the risk of emergence.," *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, vol. 356, pp. 991–999, July 2001.

[471] S. K. Sheppard, J. F. Dallas, N. J. C. Strachan, M. MacRae, N. M. McCarthy, D. J. Wilson, F. J. Gormley, D. Falush, I. D. Ogden, M. C. J. Maiden, and K. J. Forbes, "*Campylobacter* genotyping to determine the source of human infection.," *Clinical Infectious Diseases*, vol. 48, pp. 1072–1078, Apr. 2009.

[472] S. L. W. On and B. Holmes, "Classification and Identification of *Campylobacters, Helicobacters* and Allied Taxa by Numerical Analysis of Phenotypic Characters," *Systematic and Applied Microbiology*, vol. 18, pp. 374–390, Jan. 1995.

[473] P. Vandamme and J. De Ley, "Proposal for a new family, Campylobacteraceae," *International Journal of Systematic Bacteriology*, vol. 41, no. 3, pp. 451–455, 1991.

[474] M. Stahl, X. Liu, J. Li, C. M. Szymanski, and A. Stintzi, "L-fucose utilization provides *Campylobacter jejuni* with a competitive advantage," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 108, no. 17, pp. 7194–7199, 2011.

[475] J. T. Beery, M. B. Hugdahl, and M. P. Doyle, "Colonization of gastrointestinal tracts of chicks by *Campylobacter jejuni.*," *Applied and Environmental Microbiology*, vol. 54, pp. 2365–2370, Oct. 1988.

[476] J. Lu, U. Idris, B. Harmon, C. Hofacre, J. J. Maurer, and M. D. Lee, "Diversity and succession of the intestinal bacterial community of the maturing broiler chicken.," *Applied and Environmental Microbiology*, vol. 69, pp. 6816–6824, Nov. 2003.

[477] D. R. Weerakoon and J. W. Olson, "The role of respiratory donor enzymes in *Campylobacter jejuni* host colonization and physiology," *Microbial Pathogenesis*, vol. 47, pp. 8–15, July 2009.

[478] B. Lázaro, J. Cárcamo, A. Audícana, I. Perales, and A. Fernández-Astorga, "Viability and DNA maintenance in nonculturable spiral *Campylobacter jejuni* cells after long-term exposure to low temperatures.," *Applied and Environmental Microbiology*, vol. 65, pp. 4677–4681, Oct. 1999.

[479] H. Irshad, A. L. Cookson, C. M. Ross, P. JAROS, D. J. PRATTLEY, A. Donnison, G. McBRIDE, J. Marshall, and N. P. French, "Diversity and relatedness of Shiga toxin-producing Escherichia coli and Campylobacter jejuni between farms in a dairy catchment.," *Journal of Hygiene*, pp. 1–12, Nov. 2015.

[480] P. Müllner, S. E. F. Spencer, D. J. Wilson, G. Jones, A. D. Noble, A. C. Midwinter, J. M. Collins-Emerson, P. E. Carter, S. Hathaway, and N. P. French, "Assigning the source of human campylobacteriosis in New Zealand: a comparative genetic and epidemiological approach.," *Infection, genetics and evolution*, vol. 9, pp. 1311–1319, Dec. 2009.

[481] S. K. Sheppard, J. F. Dallas, M. MacRae, N. M. McCarthy, E. L. Sproston, F. J. Gormley, N. J. C. Strachan, I. D. Ogden, M. C. J. Maiden, and K. J. Forbes, "*Campylobacter* genotypes from food animals, environmental sources and clinical disease in Scotland 2005/6.," *International Journal of Food Microbiology*, vol. 134, pp. 96–103, Aug. 2009.

[482] D. H. Grove-White, A. J. H. Leatherbarrow, P. J. Cripps, P. J. Diggle, and N. P. French, "Molecular epidemiology and genetic diversity of *Campylobacter jejuni* in ruminants," *Epidemiology and Infection*, vol. 1, no. 1, pp. 1–11, 2010.

[483] M. P. Cox, D. A. Peterson, and P. J. Biggs, "SolexaQA: At-a-glance quality assessment of Illumina second-generation sequencing data.," *BMC Bioinformatics*, vol. 11, no. 1, p. 485, 2010.

[484] D. R. Zerbino and E. Birney, "Velvet: algorithms for de novo short read assembly using de Bruijn graphs.," *Genome Research*, vol. 18, pp. 821–829, May 2008.

[485] T. Seemann, "Prokka: rapid prokaryotic genome annotation.," *Bioinformatics*, vol. 30, pp. 2068–2069, Apr. 2014.

[486] M. A. Larkin, G. Blackshields, N. P. Brown, R. Chenna, P. A. McGettigan, H. McWilliam, F. Valentin, I. M. Wallace, A. Wilm, R. Lopez, J. D. Thompson, T. J. Gibson, and D. G. Higgins, "Clustal W and Clustal X version 2.0.," *Bioinformatics*, vol. 23, pp. 2947–2948, Nov. 2007.

[487] D. H. Huson, "SplitsTree: analyzing and visualizing evolutionary data.," *Bioinformatics*, vol. 14, no. 1, pp. 68–73, 1998.

[488] D. H. Huson and D. Bryant, "Application of phylogenetic networks in evolutionary studies.," *Molecular Biology and Evolution*, vol. 23, pp. 254–267, Feb. 2006.

[489] J. A. Hartigan and M. A. Wong, "Algorithm AS 136: A k-means clustering algorithm," *Journal of the Royal Statistical Society: Series C*, vol. 28, no. 1, pp. 100–108, 1979.

[490] R. Kolde, "pheatmap: pretty heatmaps," *R package version*, Sept. 2013.

[491] R. J. Sela and J. S. Simonoff, "RE-EM trees: a data mining approach for longitudinal and clustered data," *Machine Learning*, vol. 86, no. 2, pp. 169–207, 2011.

[492] D. M. Bates, M. Maechler, B. Bolker, and S. Walker, "lme4: Linear mixed-effects models using Eigen and S4," tech. rep., CRAN, Jan. 2014.

[493] R Core Team 2014, "R: A language and environment for statistical computing.," Jan. 2014.

[494] H. Wickham, *ggplot2: Elegant graphics for data analysis*. Springer New York, 1 ed., Jan. 2009.

[495] R. J. Sela and J. Simonoff, *Package ' RE-EM tree '*, 2013.

[496] O. X. Cordero and M. F. Polz, "Explaining microbial genomic diversity in light of evolutionary ecology," *Nature Reviews Microbiology*, vol. 12, pp. 263–273, Mar. 2014.

[497] A. C. Martiny, K. Treseder, and G. Pusch, "Phylogenetic conservatism of functional traits in microorganisms," *The ISME Journal*, vol. 7, pp. 1–9, Dec. 2012.

[498] C. P. A. de Haan, A.-K. Llarena, J. Revez, and M.-L. Hänninen, "Association of *Campylobacter jejuni* metabolic traits with multilocus sequence types.," *Applied and Environmental Microbiology*, vol. 78, pp. 5550–5554, Aug. 2012.

[499] X. Liu, B. Gao, V. Novik, and J. E. Galan, "Quantitative proteomics of intracellular *Campylobacter jejuni* reveals metabolic reprogramming," *PLoS Pathogens*, 2012.

[500] J. A. Wright, A. J. Grant, D. Hurd, M. Harrison, E. J. Guccione, D. J. Kelly, and D. J. Maskell, "Metabolite and transcriptome analysis of *Campylobacter jejuni in vitro* growth reveals a stationary-phase physiological switch.," *Microbiology*, vol. 155, pp. 80–94, Jan. 2009.

[501] M. J. Sellars, S. J. Hall, and D. J. Kelly, "Growth of *Campylobacter jejuni* Supported by Respiration of Fumarate, Nitrate, Nitrite, Trimethylamine-N-Oxide, or Dimethyl Sulfoxide Requires Oxygen," *Journal of Bacteriology*, vol. 184, pp. 4187–4196, Aug. 2002.

[502] J. M. Ketley and M. E. Konkel, *Campylobacter*. Molecular and Cellular Biology, Horizon Scientific Press, 2005.

[503] P. S. Hoffman and T. G. Goodman, "Respiratory physiology and energy conservation efficiency of *Campylobacter jejuni*.," *Journal of Bacteriology*, vol. 150, pp. 319–326, Apr. 1982.

[504] D. A. Johnson, S. G. Tetu, K. Phillippy, J. Chen, Q. Ren, and I. T. Paulsen, "High-throughput phenotypic characterization of *Pseudomonas aeruginosa* membrane transport genes.," *PLoS Genetics*, vol. 4, no. 10, p. e1000211, 2008.

[505] C. D. Bayliss, "Evolution of simple sequence repeat-mediated phase variation in bacterial genomes.," *Annals of the New York Academy of Sciences*, vol. 1267, no. 1, pp. 39–44, 2012.

[506] C. D. Bayliss, J. M. Ketley, P. A. Barrow, and M. A. Jones, "Phase variable genes of *Campylobacter jejuni* exhibit high mutation rates and specific mutational patterns but mutability is not the major determinant of population structure during host colonization.," *Nucleic Acids Research*, vol. 40, no. 13, pp. 5876–5889, 2012.

[507] A. V. Karlyshev, D. Linton, N. A. Gregson, and B. W. Wren, "A novel paralogous gene family involved in phase-variable flagella-mediated motility in *Campylobacter jejuni*.," *Microbiology*, vol. 148, pp. 473–480, Feb. 2002.

[508] R. Jayaraman, "Phase variation and adaptation in bacteria: A "Red Queen's Race"," *Current Science*, vol. 100, pp. 1163–1171, 2011.

[509] K. L. Fox, S. J. Dowideit, A. L. Erwin, Y. N. Srikhanta, A. L. Smith, and M. P. Jennings, "*Haemophilus influenzae* phasevarions have evolved from type III DNA restriction systems into epigenetic regulators of gene expression," *Nucleic Acids Research*, vol. 35, pp. 5242–5252, July 2007.

[510] C. D. Bayliss, M. J. Callaghan, and E. R. Moxon, "High allelic diversity in the methyltransferase gene of a phase variable type III restriction-modification system has implications for the fitness of *Haemophilus influenzae*.," *Nucleic Acids Research*, vol. 34, no. 14, pp. 4046–4059, 2006.

[511] W. Eisenreich, T. Dandekar, J. Heesemann, and W. Goebel, "Carbon metabolism of intracellular bacterial pathogens and possible links to virulence," *Nature Reviews Microbiology*, vol. 8, pp. 401–412, May 2010.

[512] J.-M. Cappelier, A. Rossero, and M. Federighi, "Demonstration of a protein synthesis in starved Campylobacter jejuni cells.," *International Journal of Food Microbiology*, vol. 55, pp. 63–67, Apr. 2000.

[513] W. Chaisowwong, A. Kusumoto, M. Hashimoto, T. Harada, K. Maklon, and K. Kawamoto, "Physiological Characterization of *Campylobacter jejuni* under Cold Stresses Conditions: Its Potential for Public Threat," *The Journal of Veterinary Medical Science*, vol. 74, no. 1, pp. 43–50, 2012.

[514] W. G. Lee and I. G. Jamieson, *The Tahake Fifty Years of Conservation Management and Research* . Dunedin, New Zealand: University of Otago Press, 1st ed., Jan. 2001.

[515] J. A. Platts-Mills and M. Kosek, "Update on the burden of Campylobacter in developing countries," *Current Opinion in Infectious Diseases*, vol. 27, pp. 444–450, Oct. 2014.

[516] A. J. Lastovica and B. M. Allos, "Clinical significance of *Campylobacter* and related species other than *Campylobacter jejuni* and *Campylobacter coli*," in *Campylobacter, Third Edition* (M. J. Blaser, I. Nachamkin, and C. M. Szymanski, eds.), pp. 123–150, Washington, DC: American Society of Microbiology, Jan. 2008.

[517] N. J. C. Strachan, F. J. Gormley, O. Rotariu, I. D. Ogden, G. Miller, G. M. Dunn, S. K. Sheppard, J. F. Dallas, T. M. S. Reid, H. Howie, M. C. J. Maiden, and K. J. Forbes, "Attribution of *Campylobacter*

infections in northeast Scotland to specific sources by use of multilocus sequence typing.," *The Journal of Infectious Diseases*, vol. 199, pp. 1205–1208, Apr. 2009.

[518] N. J. Croucher and X. Didelot, "The application of genomics to tracing bacterial pathogen transmission," *Current Opinion in Microbiology*, vol. 23, pp. 62–67, Feb. 2015.

[519] L. van der Graaf van Bloois, W. G. Miller, E. Yee, M. Rijnsburger, J. A. Wagenaar, and B. Duim, "Inconsistency of phenotypic and genomic characteristics of *Campylobacter fetus* subspecies requires re-evaluation of current diagnostics," *Journal of Clinical Microbiology*, vol. 52, pp. 4183–4188, Sept. 2014.

[520] S. A. Trewick, "Flightlessness and phylogeny amongst endemic rails (Aves:Rallidae) of the New Zealand region.," *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, vol. 352, pp. 429–446, Apr. 1997.

[521] M. V. Larsen, S. Cosentino, S. Rasmussen, C. Friis, H. Hasman, R. L. Marvig, L. Jelsbak, T. S. Ponten, D. W. Ussery, F. M. Aarestrup, and O. Lund, "Multilocus Sequence Typing of Total-Genome-Sequenced Bacteria," *Journal of Clinical Microbiology*, vol. 50, no. 4, pp. JCM.06094–11–, 2012.

[522] R. Barrangou, "CRISPR-Cas systems and RNA-guided interference.," *Wiley interdisciplinary reviews. RNA*, vol. 4, pp. 267–278, May 2013.

[523] T. Seemann, "barrnap 0.3 : Bacterial/Archaeal Ribosomal RNA Predictor ,"

[524] O.-S. Kim, Y.-J. Cho, K. Lee, S.-H. Yoon, M. Kim, H. Na, S.-C. Park, Y.-S. Jeon, J.-H. Lee, H. Yi, S. Won, and J. Chun, "Introducing EzTaxon-e: a prokaryotic 16S rRNA gene sequence database with phylotypes that represent uncultured species.," *International Journal Of Systematic And Evolutionary Microbiology*, vol. 62, pp. 716–721, Mar. 2012.

[525] S. F. Altschul, T. L. Madden, A. A. Schäffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman, "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.," *Nucleic Acids Research*, vol. 25, pp. 3389–3402, Sept. 1997.

[526] J. Goris, K. T. Konstantinidis, J. A. Klappenbach, T. Coenye, P. Vandamme, and J. M. Tiedje, "DNA-DNA hybridization values and their relationship to whole-genome sequence similarities.," *International Journal Of Systematic And Evolutionary Microbiology*, vol. 57, pp. 81–91, Jan. 2007.

[527] S. Kurtz, A. Phillippy, A. L. Delcher, M. Smoot, M. Shumway, C. Antonescu, and S. L. Salzberg, "Versatile and open software for comparing large genomes.," *Genome Biology*, vol. 5, no. 2, p. R12, 2004.

[528] B. J. Tindall, R. Rossello-Mora, H.-J. Busse, W. Ludwig, and P. Kämpfer, "Notes on the characterization of prokaryote strains for taxonomic purposes.," *International Journal Of Systematic And Evolutionary Microbiology*, vol. 60, pp. 249–266, Jan. 2010.

[529] H. Teeling, A. Meyerdierks, M. Bauer, R. Amann, and F. O. Glöckner, "Application of tetranucleotide

frequencies for the assignment of genomic fragments.," *Environmental Microbiology*, vol. 6, pp. 938–947, Sept. 2004.

[530] S. Cosentino, M. V. Larsen, F. M. Aarestrup, and O. Lund, "PathogenFinder - Distinguishing Friend from Foe Using Bacterial Whole Genome Sequence Data," *PLoS ONE*, vol. 8, p. e77302, Oct. 2013.

[531] R. K. Aziz, S. Devoid, T. Disz, R. a. Edwards, C. S. Henry, G. J. Olsen, R. Olson, R. A. Overbeek, B. Parrello, G. D. Pusch, R. L. Stevens, V. Vonstein, and F. Xia, "SEED servers: high-performance access to the SEED genomes, annotations, and metabolic models.," *PLoS ONE*, vol. 7, no. 10, p. e48053, 2012.

[532] R. K. Aziz, D. Bartels, A. A. Best, M. DeJongh, T. Disz, R. a. Edwards, K. Formsma, S. Gerdes, E. M. Glass, M. Kubal, F. Meyer, G. J. Olsen, R. Olson, A. L. Osterman, R. A. Overbeek, L. K. McNeil, D. Paarmann, T. Paczian, B. Parrello, G. D. Pusch, C. Reich, R. Stevens, O. Vassieva, V. Vonstein, A. Wilke, and O. Zagnitko, "The RAST Server: rapid annotations using subsystems technology.," *BMC Genomics*, vol. 9, no. 1, p. 75, 2008.

[533] S. Akhter, R. K. Aziz, and R. A. Edwards, "PhiSpy: a novel algorithm for finding prophages in bacterial genomes that combines similarity- and composition-based strategies," *Nucleic Acids Research*, vol. 40, pp. e126–e126, Sept. 2012.

[534] M. G. I. Langille, W. W. L. Hsiao, and F. S. L. Brinkman, "Evaluation of genomic island predictors using a comparative genomics approach.," *BMC Bioinformatics*, vol. 9, no. 1, p. 329, 2008.

[535] B. K. Dhillon, T. A. Chiu, M. R. Laird, M. G. I. Langille, and F. S. L. Brinkman, "IslandViewer update: Improved genomic island discovery and visualization.," *Nucleic Acids Research*, vol. 41, pp. W129–32, July 2013.

[536] S. Waack, O. Keller, R. Asper, T. Brodag, C. Damm, W. Fricke, K. Surovcik, P. Meinicke, and R. Merkl, "Score-based prediction of genomic islands in prokaryotic genomes using hidden Markov models," *BMC Bioinformatics*, vol. 7, no. 1, p. 142, 2006.

[537] W. Hsiao, I. Wan, S. J. Jones, and F. S. L. Brinkman, "IslandPath: aiding detection of genomic islands in prokaryotes," *Bioinformatics*, vol. 19, pp. 418–420, Feb. 2003.

[538] E. Zankari, H. Hasman, S. Cosentino, M. Vestergaard, S. Rasmussen, O. Lund, F. M. Aarestrup, and M. V. Larsen, "Identification of acquired antimicrobial resistance genes.," *Journal of Antimicrobial Chemotherapy*, vol. 67, pp. 2640–2644, Nov. 2012.

[539] M. R. Gillings, "Evolutionary consequences of antibiotic use for the resistome, mobilome and microbial pangenome.," *Frontiers in Microbiology*, vol. 4, p. 4, 2013.

[540] B. Contreras-Moreira and P. Vinuesa, "GET HOMOLOGUES, a versatile software package for scalable and robust microbial pangenome analysis.," *Applied and Environmental Microbiology*, vol. 79, pp. 7696–7701, Nov. 2013.

[541] L. Li, C. J. Stoeckert, and D. S. Roos, "OrthoMCL: identification of ortholog groups for eukaryotic genomes.," *Genome Research*, vol. 13, pp. 2178–2189, Sept. 2003.

[542] F. Chen, A. J. Mackey, C. J. Stoeckert, and D. S. Roos, "OrthoMCL-DB: querying a comprehensive multi-species collection of ortholog groups.," *Nucleic Acids Research*, vol. 34, pp. D363–8, Jan. 2006.

[543] D. M. Kristensen, L. Kannan, M. K. Coleman, Y. I. Wolf, A. Sorokin, E. V. Koonin, and A. Mushegian, "A low-polynomial algorithm for assembling clusters of orthologous groups from intergenomic symmetric best matches.," *Bioinformatics*, vol. 26, pp. 1481–1487, June 2010.

[544] J. Felsenstein, "PHYLIP - Phylogeny Inference Package (Version 3.2)," *Cladistics*, pp. 164–166, Jan. 1989.

[545] M. Woodhams, D. A. Steane, R. C. Jones, D. Nicolle, and V. Moulton, "Novel distances for dollo data," *Syst Biol*, vol. 62, pp. 62–77, June 2012.

[546] T. Jombart, "adegenet: a R package for the multivariate analysis of genetic markers.," *Bioinformatics*, vol. 24, pp. 1403–1405, June 2008.

[547] T. Jombart and I. Ahmed, "adegenet 1.3-1: new tools for the analysis of genome-wide SNP data," *Bioinformatics*, vol. 27, pp. 3070–3071, Oct. 2011.

[548] T. Jombart, S. Devillard, and F. Balloux, "Discriminant analysis of principal components: a new method for the analysis of genetically structured populations.," *BMC Genetics*, vol. 11, no. 1, p. 94, 2010.

[549] S. L. W. On, "Taxonomy of *Campylobacter, Arcobacter, Helicobacter* and related bacteria: current status, future prospects and immediate concerns.," *Journal Applied Microbiology*, vol. 90, no. 30, pp. 1S–15S, 2001.

[550] J. Chun, J. H. Lee, Y. Jung, M. Kim, S. Kim, B. K. Kim, and Y. W. Lim, "EzTaxon: a web-based tool for the identification of prokaryotes based on 16S ribosomal RNA gene sequences," *International Journal Of Systematic And Evolutionary Microbiology*, vol. 57, no. 10, pp. 2259–2261, 2007.

[551] M. Touchon, C. Hoede, O. Tenaillon, V. Barbe, S. Baeriswyl, P. Bidet, E. Bingen, S. Bonacorsi, C. Bouchier, O. Bouvet, A. Calteau, H. Chiapello, O. Clermont, S. Cruvellier, A. Danchin, M. Diard, C. Dossat, M. E. Karoui, E. Frapy, L. Garry, J. M. Ghigo, A. M. Gilles, J. Johnson, C. Le Bouguénec, M. Lescat, S. Mangenot, V. Martinez-Jéhanne, I. Matic, X. Nassif, S. Oztas, M. A. Petit, C. Pichon, Z. Rouy, C. S. Ruf, D. Schneider, J. Tourret, B. Vacherie, D. Vallenet, C. Médigue, E. P. C. Rocha, and E. Denamur, "Organised genome dynamics in the *Escherichia coli* species results in highly diverse adaptive paths.," *PLoS Genetics*, vol. 5, p. e1000344, Jan. 2009.

[552] J. Bohlin, L. Snipen, A. Cloeckaert, K. Lagesen, D. Ussery, A. B. Kristoffersen, and J. Godfroid, "Genomic comparisons of *Brucella* spp. and closely related bacteria using base compositional and proteome based methods," *BMC Evolutionary Biology*, vol. 10, no. 1, p. 249, 2010.

[553] J. Siefert, "Defining the Mobilome," in *Methods in Molecular Biology* (M. Gogarten, J. Gogarten, and L. Olendzenski, eds.), pp. 13–27–27, Totowa, NJ: Humana Press, 2009.

[554] A. E. Zautner, S. Herrmann, J. Corso, A. M. Tareen, T. Alter, and U. Gross, "Epidemiological

association of different *Campylobacter jejuni* groups with metabolism-associated genetic markers.," *Applied and Environmental Microbiology*, vol. 77, pp. 2359–2365, Apr. 2011.

[555] D. J. Wilson, E. Gabriel, A. J. H. Leatherbarrow, J. Cheesbrough, S. Gee, E. Bolton, A. Fox, C. A. Hart, P. J. Diggle, and P. Fearnhead, "Rapid evolution and the importance of recombination to the gastroenteric pathogen *Campylobacter jejuni*.," *Molecular Biology and Evolution*, vol. 26, pp. 385–397, Feb. 2009.

[556] L. Morley, A. McNally, K. Paszkiewicz, J. Corander, G. Meric, S. K. Sheppard, J. Blom, and G. Manning, "Gene loss and lineage specific restriction-modification systems associated with niche differentiation in the *Campylobacter jejuni* Sequence Type 403 clonal complex.," *Applied and Environmental Microbiology*, pp. AEM.00546–15, Mar. 2015.

[557] I. Hansson, M. Persson, L. Svensson, E. O. Engvall, and K.-E. Johansson, "Identification of nine sequence types of the 16S rRNA genes of *Campylobacter jejuni* subsp. *jejuni* isolated from broilers.," *Acta veterinaria Scandinavica*, vol. 50, no. 1, p. 10, 2008.

[558] G. E. Fox, J. D. Wisotzkey, and P. Jurtshuk, "How Close Is Close: 16S rRNA Sequence Identity May Not Be Sufficient To Guarantee Species Identity," *International Journal of Systematic Bacteriology*, vol. 42, pp. 166–170, Jan. 1992.

[559] S. K. Sheppard, X. Didelot, K. A. Jolley, A. E. Darling, B. Pascoe, G. Meric, D. J. Kelly, A. Cody, F. M. Colles, N. J. C. Strachan, I. D. Ogden, K. Forbes, N. P. French, P. E. Carter, W. G. Miller, N. M. McCarthy, R. Owen, E. Litrup, M. Egholm, J. P. Affourtit, S. D. Bentley, J. Parkhill, M. C. J. Maiden, and D. Falush, "Progressive genome-wide introgression in agricultural *Campylobacter coli*.," *Molecular Ecology*, vol. 22, pp. 1051–1064, Feb. 2013.

[560] S. K. Sheppard, N. M. McCarthy, M. C. J. Maiden, and K. A. Jolley, "Introgression in the genus *Campylobacter*: generation and spread of mosaic alleles.," *Microbiology*, vol. 157, no. 4, pp. 1066–1074, 2011.

[561] C. R. Woese and G. E. Fox, "Phylogenetic structure of the prokaryotic domain: the primary kingdoms.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 74, pp. 5088–5090, Nov. 1977.

[562] E. Stackebrandt and B. M. Goebel, "Taxonomic Note: A Place for DNA-DNA Reassociation and 16S rRNA Sequence Analysis in the Present Species Definition in Bacteriology," *International Journal of Systematic Bacteriology*, vol. 44, pp. 846–849, Oct. 1994.

[563] M. Kim, H. S. Oh, S. C. Park, and J. Chun, "Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes," *International Journal Of Systematic And Evolutionary Microbiology*, vol. 64, pp. 346–351, Feb. 2014.

[564] N. P. French, M. Barrigas, P. Brown, P. Ribiero, N. Williams, H. Leatherbarrow, R. Birtles, E. Bolton, P. Fearnhead, and A. Fox, "Spatial epidemiology and natural population structure of *Campylobacter jejuni* colonizing a farmland ecosystem.," *Environmental Microbiology*, vol. 7, pp. 1116–1126, Aug. 2005.

[565] M. T. Le and A. H. van Vliet, "Genome and Transcriptome Evolution in the Genus *Campylobacter*," in *Campylobacter Ecology and Evolution* (G. Meric, ed.), pp. 43–55, Caister Academic Press, 2014.

[566] P. Fearnhead, P. J. Biggs, and N. P. French, "Learning about recombination in *Campylobacter*," in *Campylobacter Ecology and Evolution* (G. Meric, ed.), pp. 9–23, Norfolk, U.K.: Caister Academic Press, 2014.

[567] W. G. Miller, E. Yee, M. H. Chapman, T. P. L. Smith, J. L. Bono, S. Huynh, C. T. Parker, P. Vandamme, K. Luong, and J. Korlach, "Comparative genomics of the *Campylobacter lari* group.," *Genome Biology and Evolution*, Nov. 2014.

[568] F. Zhou and Y. Xu, "cBar: a computer program to distinguish plasmid-derived from chromosome-derived sequence fragments in metagenomics data," *Bioinformatics*, vol. 26, pp. 2051–2052, Aug. 2010.

[569] L. M. Friis, C. Pin, D. E. Taylor, B. M. Pearson, and J. M. Wells, "A role for the tet(O) plasmid in maintaining Campylobacter plasticity," *Plasmid*, vol. 57, pp. 18–28, Jan. 2007.

[570] E. C. Gaynor, S. Cawthraw, G. Manning, J. K. MacKichan, S. Falkow, and D. G. Newell, "The Genome-Sequenced Variant of Campylobacter jejuni NCTC 11168 and the Original Clonal Clinical Isolate Differ Markedly in Colonization, Gene Expression, and Virulence-Associated Phenotypes," *Journal of Bacteriology*, vol. 186, pp. 503–517, Dec. 2003.

[571] G. M. Ruiz-Palacios, "The health burden of *Campylobacter* infection and the impact of antimicrobial resistance: playing chicken.," *Clinical Infectious Diseases*, vol. 44, pp. 701–703, Mar. 2007.

[572] P. Vandamme, D. Gevers, and L. Debruyne, "Taxonomy of the Family Campylobacteraceae," in *Campylobacter, Third Edition*, pp. 3–25, New York: American Society of Microbiology, Jan. 2008.

[573] R. M. Howlett, M. P. Davey, W. Paul Quick, and D. J. Kelly, "Metabolomic analysis of the food-borne pathogen *Campylobacter jejuni*: application of direct injection mass spectrometry for mutant characterisation," *Metabolomics*, vol. 10, pp. 887–896, Mar. 2014.

[574] S. Shrivastava and S. S. Mande, "Identification and Functional Characterization of Gene Components of Type VI Secretion System in Bacterial Genomes," *PLoS ONE*, vol. 3, p. e2955, Aug. 2008.

[575] S. B. Neogi, S. Yamasaki, M. Alam, and R. J. Lara, "The role of wetland microinvertebrates in spreading human diseases," *Wetlands Ecology and Management*, vol. 22, pp. 469–491, Aug. 2014.

[576] S. J. Coulthurst, "The Type VI secretion system - a widespread and versatile cell targeting system.," *Research in microbiology*, vol. 164, pp. 640–654, July 2013.

[577] M. Ugarte-Ruiz, R. A. Stabler, L. Domínguez, M. C. Porrero, B. W. Wren, N. Dorrell, and O. Gundogdu, "Prevalence of Type VI Secretion System in Spanish *Campylobacter jejuni* Isolates," *Zoonoses and Public Health*, pp. n–a–n–a, Dec. 2014.

[578] J. W. Harrison, T. T. N. Dung, F. Siddiqui, S. Korbrisate, H. Bukhari, M. P. V. Tra, N. V. M. Hoang, J. Carrique-Mas, J. Bryant, J. I. Campbell, D. J. Studholme, B. W. Wren, S. Baker, R. W. Titball,

and O. L. Champion, "Identification of Possible Virulence Marker from *Campylobacter jejuni* isolates," *Emerging Infectious Diseases*, vol. 20, pp. 1026–1029, June 2014.

[579] K. Lertpiriyapong, E. R. Gamazon, Y. Feng, D. S. Park, J. Pang, G. Botka, M. E. Graffam, Z. Ge, and J. G. Fox, "*Campylobacter jejuni* type VI secretion system: roles in adaptation to deoxycholic acid, host cell adherence, invasion, and in vivo colonization.," *PLoS ONE*, vol. 7, no. 8, p. e42842, 2012.

[580] N. M. C. Bleumink-Pluym, L. B. van Alphen, L. I. Bouwman, and M. M. S. M. Wosten, "Identification of a functional type VI secretion system in *Campylobacter jejuni* conferring capsule polysaccharide sensitive cytotoxicity.," *PLoS Pathogens*, vol. 9, p. e1003393, May 2013.

[581] F. Boyer, G. Fichant, J. Berthod, Y. Vandenbrouck, and I. Attree, "Dissecting the bacterial type VI secretion system by a genome wide in silico analysis: what can be learned from available microbial genomic resources?," *BMC Genomics*, vol. 10, no. 1, p. 104, 2009.

[582] A. B. Russell, S. B. Peterson, and J. D. Mougous, "Type VI secretion system effectors: poisons with a purpose," *Nature Reviews Microbiology*, vol. 12, no. 2, pp. 137–148, 2014.

[583] U. Genschel, "Coenzyme A biosynthesis: reconstruction of the pathway in archaea and an evolutionary scenario based on comparative genomics.," *Molecular Biology and Evolution*, vol. 21, pp. 1242–1251, July 2004.

[584] A. H. M. van Vliet, J. M. Ketley, S. F. Park, and C. W. Penn, "The role of iron in *Campylobacter* gene regulation, metabolism and oxidative stress defense.," *FEMS Microbiology Reviews*, vol. 26, no. 2, pp. 173–186, 2002.

[585] A. Stintzi, A. H. M. van Vliet, and J. M. Ketley, "Iron metabolism, transport and regulation.," in *Campylobacter, Third Edition* (I. Nachamkin, C. M. Szymanski, and M. J. Blaser, eds.), pp. 591–610, New York: American Society of Microbiology, Jan. 2008.

[586] A. E. Zautner, C. Ohk, A. M. Tareen, R. Lugert, and U. Gross, "Epidemiological association of *Campylobacter jejuni* groups with pathogenicity-associated genetic markers.," *BMC Microbiology*, vol. 12, no. 1, p. 171, 2012.

[587] S. Gibb and K. Strimmer, *binda: Multi-Class Discriminant Analysis using Binary Predictors. R package version 1.0.0.*, Jan. 2014.

[588] U. Bodenhofer, A. Kothmeier, and S. Hochreiter, "APCluster: an R package for affinity propagation clustering.," *Bioinformatics*, vol. 27, pp. 2463–2464, Sept. 2011.

[589] R. C. Edgar, "MUSCLE: multiple sequence alignment with high accuracy and high throughput.," *Nucleic Acids Research*, vol. 32, no. 5, pp. 1792–1797, 2004.

[590] J. M. Silverman, Y. R. Brunet, E. Cascales, and J. D. Mougous, "Structure and Regulation of the Type VI Secretion System," *Annual Review of Microbiology*, vol. 66, pp. 453–472, Oct. 2012.

[591] D. Bryant and V. Moulton, "NeighborNet: An agglomerative method for the construction of planar phylogenetic networks," *Algorithms in Bioinformatics*, vol. 2452, no. Chapter 28, pp. 375–391, 2002.

[592] C. M. Buswell, Y. M. Herlihy, L. M. Lawrence, J. T. McGuiggan, P. D. Marsh, C. W. Keevil, and S. A. Leach, "Extended survival and persistence of Campylobacter spp. in water and aquatic biofilms and their detection by immunofluorescent-antibody and -rRNA staining.," *Applied and Environmental Microbiology*, vol. 64, no. 2, pp. 733–741, 1998.

[593] W. Baffone, A. Casaroli, B. Citterio, L. Pierfelici, R. Campana, E. Vittoria, E. Guaglianone, and G. Donelli, "*Campylobacter jejuni* loss of culturability in aqueous microcosms and ability to resuscitate in a mouse model," *International Journal of Infectious Diseases*, 2006.