Population Biology and Comparative Genomics of *Campylobacter* Species

Lennard Epping, Esther-Maria Antão, and Torsten Semmler

Contents

Abstract The zoonotic pathogen *Campylobacter* is the leading cause for bacterial foodborne infections in humans. *Campylobacters* are most commonly transmitted via the consumption of undercooked poultry meat or raw milk products. The decreasing costs of whole genome sequencing enabled large genome-based analyses of the evolution and population structure of this pathogen, as well as the development of novel high-throughput molecular typing methods. Here, we review the evolutionary development and the population diversity of the two most clinically relevant *Campylobacter* species; *C. jejuni* and *C. coli.* The state-of-the-art phylogenetic studies showed clustering of *C. jejuni* lineages into host specialists and generalists with

The original version of this chapter was revised: This chapter has been changed to open access under the terms of the Creative Commons Attribution 4.0 International License. The correction to this chapter is available at https://doi.org/10.1007/978-3-030-65481-8_12

L. Epping · T. Semmler (✉)
Microbial Genomics, Robert Koch Institute, Nordufer 20, 13353 Berlin, Germany
e-mail: SemmlerT@rki.de

L. Epping
e-mail: EppingL@rki.de

E.-M. Antão
Robert Koch Institute, Nordufer 20, 13353 Berlin, Germany
e-mail: AntaoE@rki.de

S. Backert (ed.), *Fighting Campylobacter Infections*, Current Topics in Microbiology and Immunology 431, https://doi.org/10.1007/978-3-030-65481-8_3

coexisting lifestyles in chicken and livestock-associated hosts, as well as the separation of *C. coli* isolates of riparian origin (waterfowl, water) from *C. coli* isolated from clinical and farm-related samples. We will give an overview of recombination between both species and the potential impact of horizontal gene transfer on host adaptation in *Campylobacter*. Additionally, this review briefly places the current knowledge of the population structure of other *Campylobacter* species such as *C. lari*, *C. concisus* and *C. upsaliensis* into perspective. We also provide an overview of how molecular typing methods such as multilocus sequence typing (MLST) and whole genome MLST have been used to detect and trace *Campylobacter* outbreaks along the food chain.

1 Introduction

Campylobacter is one of the most common causes of foodborne infections worldwide (Kaakoush et al. 2015). To date, the genus *Campylobacter* includes 32 formally described species and 9 subspecies (Costa and Iraola 2019) and is part of the natural microbiota in the intestines of farm and wild animals (Altekruse et al. 1999). The most commonly known species are *Campylobacter jejuni* and *Campylobacter coli* that are mainly associated with campylobacteriosis in humans (Møller Nielsen 1997; Gillespie et al. 2002). *Campylobacter lari*, *Campylobacter concisus* and *Campylobacter upsaliensis* are less important for human gastrointestinal infections, but still can be frequently isolated from clinically relevant samples (Man 2011). Most notably, their multi-host lifestyles and ability for adaptation make *C. jejuni* and *C. coli* dangerous pathogens that are typically transmitted through the food chain (Oyarzabal and Backert 2012). Mainly spread through undercooked chicken meet or raw milk, these bacteria infect around 550 million people annually as reported by the World Health Organization (WHO), resulting in worldwide healthcare costs and economy loss of billions of dollars (Kaakoush et al. 2015).

Since the first complete genome sequence of the *Campylobacter* species *C. jejuni* published in 2000 (Parkhill et al. 2000), the functionality of whole genome sequencing (WGS) such as next-generation sequencing (NGS) or long read sequencing technology, namely Oxford Nanopore Technology (ONT) and Pacific Bioscience (PacBio), has massively improved. Time-consuming and low-resolution methods like pulsed-field gel electrophoresis (PFGE) (Yan et al. 1991; Potturi-Venkata et al. 2007) and *flaA* typing (Nachamkin et al. 1993) have been replaced by multilocus sequence typing (MLST) or whole/core genome MLST (wgMLST, cgMLST) that have since been frequently used for epidemiological studies (Tagini and Greub 2017). Instead of only analyzing a small part of the genome, e.g., a single gene (*flaA* typing) or MLST, which accounts for only 0.2% of the genome (Sheppard and Maiden 2015), wgMLST differentiates isolates by using all coding regions of the genomes incorporating hundreds of genes. This high discriminatory power even allows to link transmission events in epidemiological studies. Thus, high-throughput sequencing has become a time- and cost-effective method for typing, transmission-tracing, evolutionary analyses and surveillance of *Campylobacter*.

Besides comprehensive typing methods, NGS provides a broad range of possibilities to study genetic variations with respect to phenotypic difference. Powerful tools such as pan-genomic studies (Medini et al. 2005) or genome-wide association studies (GWAS), which were recently applied to microbial genomics (Falush 2016), allow very detailed correlation of the presence/absence and the allelic variants of all genes within a bacterial species population with specific phenotypes (see Sect. 4.1 below). These WGS-driven approaches enable researchers to effectively study the important aspects of host-specificity and adaptation of *Campylobacter* and help to understand the transmission and emergence of *Campylobacter* infections.

In this review, we give a broad overview of the historical evolution of *Campylobacter* and how the current population structure has been formed by niche adaptation together with inter- and intra-*Campylobacter* species recombination. Furthermore, we describe the huge potential of high-throughput and computational methods used to study relationships of *Campylobacter* strains in an agricultural and clinical environment that have provided new evidence regarding host and niche segregation.

2 Evolution Theory and Concepts for the Genus *Campylobacter*

In order to understand evolutionary and ecological processes within bacterial evolution, it is important to measure the molecular rate of mutations per replication event, also known as a molecular clock (Duchêne et al. 2016). The mutation rate of bacteria can be influenced by several different evolutionary processes such as selection pressure, genetic drift or the bottleneck effect that might play an important role in a host-adapted species like *C. jejuni* (Toft and Andersson 2010). The general approach of Ochman and Wilson (1987) to analyze the molecular clock is based on ancestral diversification calculated by 1% divergence in 16S rRNA nucleotides per 50 million years. Using this method, the divergence time of the genus *Campylobacter* was estimated to have started around 10 million years ago and clade formation of *C. coli* around 2.5 million years ago (Sheppard and Maiden 2015). However, *Campylobacter* was identified to evolve more rapidly than *Escherichia coli* and *Salmonella Typhimurium*, which have been used by Ochman and Wilson. *Campylobacter* has an unusually high rate of recombination, as horizontal gene transfer was estimated to generate two times more genetic diversity than de novo mutations (Wilson et al. 2009). Furthermore, bacterial lineages accumulate genetic substitutions more rapidly while they undergo adaptive evolution (Eyre-Walker and Keightley 2007). For all these reasons, Wilson et al. (2009) proposed a novel approach to estimate divergence in *Campylobacter* population by applying a more rapid rate of the molecular clock. They estimated the divergence of *C. coli* and *C. jejuni* to 6,580 years ago, with 95% confidence intervals (CI) of 3,580–12,400. This estimate fits within the time frame of the first domestication of wild animals during the agricultural revolution (Neolithic Revolution). The Neolithic Revolution started around 10,000–12,000 BC

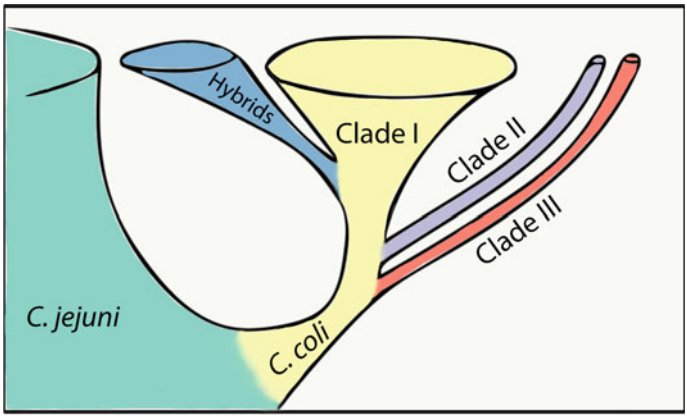

Fig. 1 Schematic representation of an evolutionary scenario of *C. coli* and *C. jejuni* (adapted from Sheppard et al. 2013a). *C. coli* and *C. jejuni* separated into two species. Due to different ecological niches *C. coli* differentiated into three clades (I–III) (Sheppard et al. 2008, 2013a). Recent recombination between strains from *C. coli* clade I and *C. jejuni* lead to the development of *C. coli* hybrid strains with substantial genomic introgression from *C. jejuni* (Sheppard et al. 2008; Golz et al. 2020)

in the Middle East and spread to central Europe 3,000–5,000 BC, providing novel niches and possibilities to emerge for commensal and pathogenic bacteria (Mira et al. 2006). The divergence of *C. coli* into three distinct clades was estimated to 1,000–1,700 years ago, and clonal complexes of *C. jejuni* started to evolve 400 years ago (Fig. 1). This timeline indicates that the emergence of *C. jejuni* and *C. coli* as individual species is a very recent event compared to *E. coli* where the main population without members of related genera has been formed around five million years ago (Wirth et al. 2006).

Independent of the model used, it is clear that the clonal complexes and clade forming lineages separated after the ancestral split of the genus into these major species that currently play a significant role in clinical and foodborne diseases. However, the development of two distinct species did not force a strict recombination barrier between them (Sheppard et al. 2013a). While the speciation within the genus *Campylobacter* was probably triggered by the agricultural revolution thousands of years ago, methods of the modern food industry, globalization or environmental changes form novel evolutionary niches and selection pressure for bacteria in general (de Mazancourt et al. 2008; Van Alfen 2015; Caniça et al. 2019). In case of *Campylobacter*, there is evidence that *C. coli* started to converge toward *C. jejuni* due to a change in their ecology, e.g., by colonizing the same niche or host (Sheppard et al. 2008), which has been facilitating recombination between these species, as will be discussed below.

3 Population Structure

Since *Campylobacter* spp. have become more and more relevant for public health, high-throughput molecular typing plays an important role in surveillance programs and outbreak control. Most importantly, MLST and NGS provide a generic approach and, additionally, have a massive impact on understanding the population structure of *Campylobacter*. MLST is a generic scheme based on allelic variants from seven housekeeping genes used to classify bacteria into related or distant lineages (Maiden et al. 1998). *C. jejuni* and *C. coli* are characterized by the same MLST scheme which analyzes allelic variants of the same orthologous loci in both species, enabling the possibility of directly comparing the species with each other (Dingle et al. 2001; Miller et al. 2005). With the advent of high-throughput NGS, epidemiological studies made use of more detailed and complex schemes and methods developed for comparative genomics, which generated in-depth knowledge about the population structure of microbes. In this section, we will describe the population structure of both *C. jejuni* and *C. coli* that have an average nucleotide identity (ANI) of 85% (Fig. 2) (Dingle et al. 2005). Furthermore, we will give an overview of recombination events between these species, which resulted in the emergence of "hybrid" strains.

3.1 Diversity and Population Structure of C. Jejuni *and* C. Coli

C. jejuni is a natural part of the gut microbiota in a wide range of hosts such as chicken, cattle, pigs or wild birds and can also be found in environmental reservoirs such as water (Altekruse et al. 1999). This multi-host lifestyle is reflected by its broad diversity, which can even be detected by a low-resolution method like MLST, representing less than 1% of the genomic DNA in *Campylobacter*. Based on phylogenetic analyses (Fig. 3), resulting from a concatenated alignment of the genes used for cgMLST, *C. jejuni* forms a weak clonal complex structure (Dingle et al. 2001; Suerbaum et al. 2001). The clonal complexes CC-45 and CC-21 harbor the most relevant clinical and outbreak strains and are among the most prevalent isolates at PubMLST database (https://pubmlst.org/), with 24% and 9% of the entries, respectively, emphasizing their importance. Isolates belonging to these complexes are known to be "host-generalist" that can colonize cattle, chicken or human hosts (Manning et al. 2003; Dearlove et al. 2016). Their ability to switch rapidly between hosts makes them a dangerous threat for human health through consumption of contaminated milk and of undercooked chicken products. Geographical signatures in *Campylobacter* are relatively weak as "identical" host associated lineages emerge all over the world (Pascoe et al. 2017). However, the frequency of specific STs can vary between countries. For example, ST-22 has been identified in Finland (Revez et al. 2011), ST-4526 in Japan (Asakura et al. 2012), and ST-190 and ST-474 were observed to emerge rapidly in New Zealand (McTAVISH et al. 2008; Mohan et al.

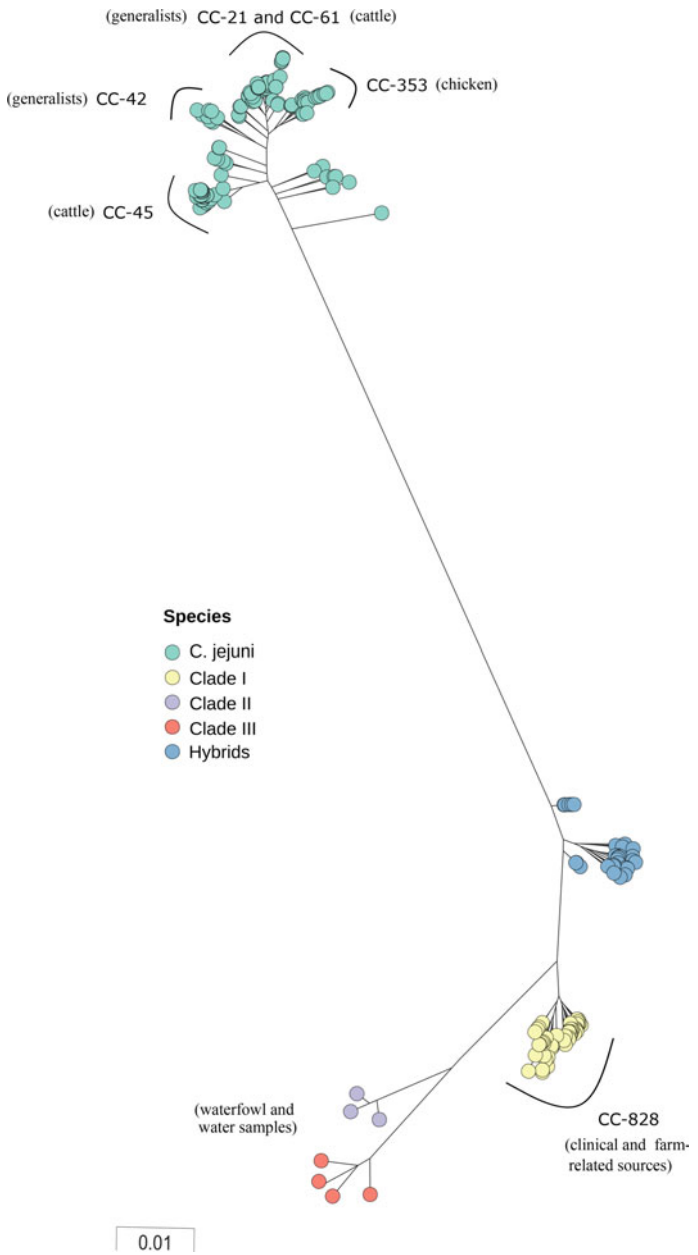

Fig. 2 Graphical visualization of pairwise ANI values of *C. coli* and *C. jejuni* genomes. *C. coli* Clade I (yellow), Clade II (red) and Clade III (purple) are clearly separated based on ANI. *C. jejuni* (turquoise) and *C. coli* are distinct species, with approximately 85% ANI. Hybrid strains formed a separate cluster but were classified as *C. coli* based on 97% ANI, in contrast to 88% ANI between the hybrid strains and *C. jejuni*. Data were taken from (Sheppard et al. 2013a, b; Golz et al. 2020). ANI was calculated using FastANI (Jain et al. 2018) and visualized with pheatmap (Kolde 2015)

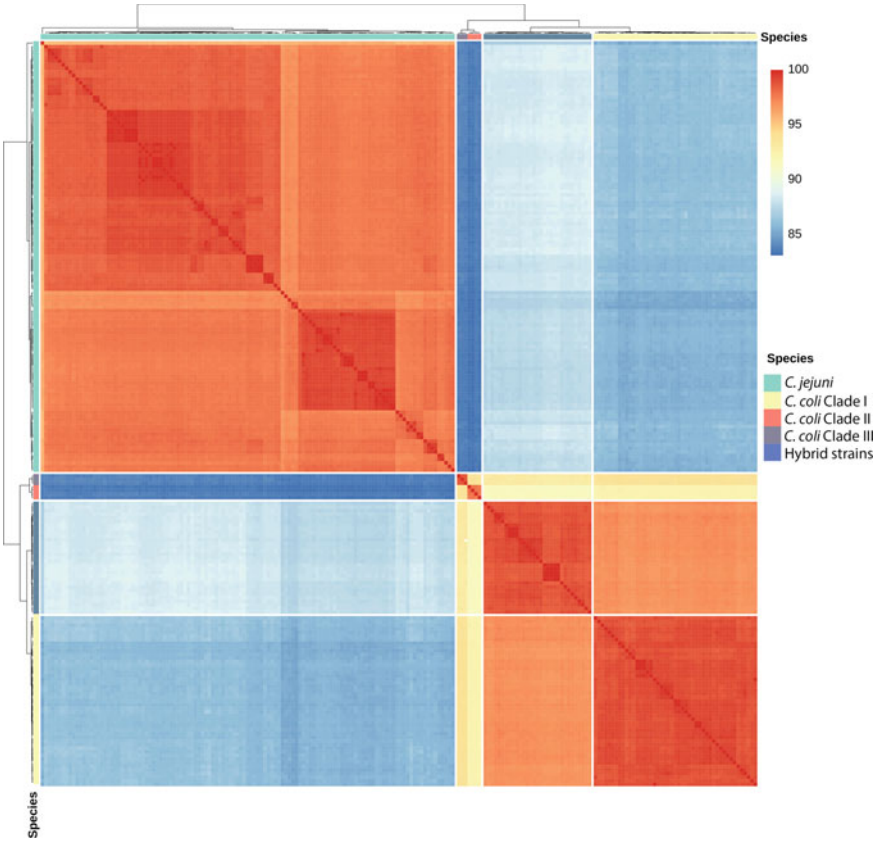

Fig. 3 Core genome-based phylogeny of *C. coli, C. jejuni* and hybrid strains. *C. jejuni* (turquoise) shows a diverse lineage-specific population structure with CC-21 and CC-45 (both host generalists), CC-42 and CC-61 (predominantly isolated from cattle), and CC-353 (from chicken). *C. coli* shows a three-clade structure with Clade I (yellow: from clinical- and farm-related sources), Clade II (purple) and Clade III (red: both from waterfowl and water samples). Clade I mainly consists of CC-828. Hybrid genomes with high DNA introgression from *C. jejuni* are colored in blue. Data were taken from (Sheppard et al. 2013a, b; Golz et al. 2020), and the phylogenetic tree was created with FastTree v2.1 (Price et al. 2010) based on 874 core genes including 123,227 variable sites

2013). Besides host generalists, repetivitve some lineages of host specialists can also cause human infections through food products. Those include CC-42 and CC-61 that are associated with cattle and sheep (Colles et al. 2003), and several different STs and CCs associated with chicken, including CC-257, CC-353 or CC-443 (Sheppard et al. 2011a, 2014). Other lineages such as CC-177 and CC-682 can be isolated from wild birds and water, causing the so-called water-born *Campylobacter* infections (Colles et al. 2009; Mohan et al. 2013). *C. jejuni* also shows a high level of diversity within the same barn or herd—e.g., isolates belonging to more than 10 distinct CCs have been found within a single chicken flock (Colles et al. 2008; Vidal et al. 2016).

However, *C. jejuni* CCs may be subject to a strong recombination barrier even if they colonize the same host (Sheppard et al. 2014). This might be forced by a niche separation within the same host, due to subsequent colonization events at different time points, which limit the horizontal gene transfer (Sheppard and Maiden 2015).

Even isolates assigned to the same ST based on the seven housekeeping genes can vary to great extent in their genetic diversity. For example, 16 strains assigned to ST-45 that were isolated during an outbreak in Finland formed three distinct strain clusters in wgMLST. Out of approximately 1200 shared loci, these clusters differed from each other by alleles in 293, in 414, and in 453 loci, respectively, indicating the presence of clearly different strains. In contrast, within the individual strain clusters the genomes differed by between zero and eighteen loci, suggesting clonal descent of those isolates (Kovanen et al. 2014). The other frequently isolated STs from this outbreak, including ST-230, ST-267 and ST-677, showed a maximum of 40 different alleles among genome clusters within each ST (Kovanen et al. 2014).

In contrast to *C. jejuni*, *C. coli* forms three distinct clades (I-III) (Figs. 1, 2 and 3), colonizing different ecological niches. Isolates from clade I are generally associated with an agricultural origin, whereas isolates belonging to clade II or clade III can most likely be found in environmental sources like water (Sheppard et al. 2008, 2013a; Skarp-de Haan et al. 2014). To date, around 81% of the genotyped isolates included in the PubMLST database belong to clonal complex CC-828 of clade I, reflecting the clinical relevance and industrial importance of this lineage (Miller et al. 2006; Thakur et al. 2006; Cody et al. 2012; Nohra et al. 2016). The second-most predominant clonal complex, also part of clade I, is CC-1150, comprising around 5% of *C. coli* isolates submitted to the PubMLST database. Clade I has a lower rate of diversity compared to *C. coli* clade II, to *C. coli* clade III, or to the general population structure of *C. jejuni* (Duim et al. 1999; Dingle et al. 2005; Sheppard et al. 2010b). The relatively low variation within the housekeeping genes as well as the lack of a proper lineage separation, especially in clade I, indicate the effect of a recent bottleneck and thus an early phase of lineage separation in the *C. coli* population (Sheppard et al. 2010b). Due to the distinct ecological niches, an ecological recombination barrier might have led to the development of three clades in *C. coli* (Sheppard et al. 2010b). However, recombination between *C. coli* clade I and *C. jejuni* resulted in hybrid strains (Figs. 1 and 3), as has been shown in several studies (Sheppard et al. 2008, 2013a; Sheppard and Maiden 2015; Golz et al. 2020).

3.2 Inter Species Recombination and Hybrid Species

Bacterial evolution is highly influenced by horizontal or lateral gene transfer (HGT or LGT) through transformation, transduction or conjugation. For recombination events, one has to distinguish between DNA introgression of complete genes or gene loci and intragenic recombination between loci leading to new mosaic allelic variants. Mosaic alleles consist of sequence content derived from different evolutionary and ancestral backgrounds (Smith 1992). As previously mentioned, early

inter-species recombination, especially between *C. jejuni* and *C. coli*, plays a major role in the evolution of the genus *Campylobacter*, which might compensate for the small genome size of this genus (Suerbaum et al. 2001). Indeed, about 18.6% of the allelic variants of the seven MLST genes in *C. coli* exhibit *C. jejuni* ancestry, whereas just 2.3% of *C. jejuni* alleles were acquired from *C. coli*, indicating asymmetric gene flow between the two species (Sheppard et al. 2008). A more detailed analysis of the mosaic ancestry patterns among the seven housekeeping genes revealed an average inter-species gene flow of around 8.3% from *C. jejuni* to *C. coli* clade I, but less than 0.5% from *C. coli* clade I to *C. jejuni* (Sheppard et al. 2011b). Even in *C. coli* clade I, the genome-wide DNA introgression rate differs substantially among the predominant clonal complexes. CC-828 showed an overall introgression of approximately 10% whereas CC-1150 was found to contain up to 23% of its genome acquired from *C. jejuni* in agriculture-associated samples. Recombination mainly happened in agriculturally relevant isolates rather than in non-agricultural *C. coli* isolates and thus might be an important adaptation and niche aggregation factor. In *C. coli* clade II and clade III, genome-wide recombination with *C. jejuni* played a minor role as those isolated had only 0.2–1.2% inferred *C. jejuni* ancestry (Sheppard et al. 2013a).

Apart from the single allele exchanges, it is possible that multiple loci in the genomes have been exchanged between *C. jejuni* and *C. coli*. This would lead to the appearance of several hybrid strains (Fig. 1) that cannot clearly be identified by routine polymerase chain reaction (PCR) typing with single species differentiation marker genes and need to be investigated further by WGS. Several of such untypeable *Campylobacter* strains were isolated from egg shells of chickens in Germany (Golz et al. 2020). These isolates showed a DNA introgression of up to 15% from *C. jejuni*. However, they were still identified as *C. coli* as they exhibited 97% average nucleotide identity with *C. coli* clade I, but only 88% ANI with *C. jejuni* (Fig. 2). Furthermore, detailed genome analysis provided evidence that these recombination events are not distributed randomly across the chromosome. Instead, they particularly affect genes that are involved in general stress response, in DNA repair and in cell wall synthesis mechanisms and thus might enhance the fitness of *C. coli* for survival under harsh environmental conditions.

3.3 Additional Species

C. jejuni and *C. coli* are the most prevalent species concerning food contamination and clinical *Campylobacter* infections. Besides these, 13 additional *Campylobacter* species, sporadically causing clinically relevant symptoms, have been summarized (Costa and Iraola 2019). In the following subsection, we exemplarily describe the population structure of *C. lari*, *C. upsaliensis* and *C. concisus* that are frequently found in gastroenteritis patients (Man 2011).

C. lari is usually found in coastal regions and marine environments. It is mainly associated with shorebirds, like gulls, albatrosses, redshanks, to name a few, but also

with marine mammals and shellfish, and occasionally causes gastroenteritis infections (Costa and Iraola 2019). However, the species definition of *C. lari* is an ongoing process, and several *C. lari*-like species have been described, including *Campylobacter insulaenigrae*, *Campylobacter peloridis*, *Campylobacter subantarcticus* and *Campylobacter volucris*. In 2009, *C. lari* was divided into two subspecies, namely *C. lari* subsp. *lari* and *C. lari* subsp. *concheus* (Debruyne et al. 2009). All *C. lari* and *C. lari*-like species are summarized as *Campylobacter lari* group (Miller et al. 2014).

C. concisus colonizes the human oral cavity and consists of two genetically distinct genomospecies (GS1 and GS2) that cannot be distinguished on the phenotypic level despite DNA binding values of only 42–50% in DNA-DNA hybridization experiments (Vandamme et al. 1989; Aabenhus et al. 2005). However, both genomospecies include multiple strains that have been isolated from healthy as well as diarrheic patients, which makes it difficult to make a general assumption on its pathogenicity (Chung et al. 2016). In particular, *C. concisus* GS2 seems to be more pathogenic as it is more often isolated from clinical patients with bloody diarrhea (Kalischuk and Inglis 2011). In addition, a recent study discovered novel genomic markers and a specific plasmid which are associated with *C. concisus* GS2 from patients suffering from Crohn's Disease (Liu et al. 2018).

C. upsaliensis is commonly found in domestic animals like cats and dogs (Goossens et al. 1990), but has also been isolated all over the world from clinical cases of bloody diarrhea (Bourke et al. 1998). This *Campylobacter* species is closely related to *C. coli* and *C. jejuni* based on 16S rRNA comparison (Vandamme et al. 1991). In contrast to *C. concisus*, *C. upsaliensis* shows a homogenous population structure with 80–96% DNA-DNA hybridization between strains (Sandstedt et al. 1983), even though it possesses a high degree of diversity on a genotypic level (Lentzsch et al. 2004). Besides this, little is known about the emergence of *C. upsaliensis*, which needs to be investigated in further studies.

4 Host Association of *Campylobacter*

Comparative genomic methods not only had a major influence on our understanding of population structures, but also advanced our knowledge and understanding of host adaptive mechanisms of *Campylobacter*. Besides the MLST and cgMLST schemes, (pan-genome) approaches and genome-wide association studies have opened the door for large-scale genome analyses of these traits.

4.1 Impact of Genomic High-Throughput Methods

Pan-genomic analyses have become powerful tools to study a variety of bacterial species (Rouli et al. 2015). The term "pan-genome" describes the entire set of genes

composed of core and accessory genes within a bacterial population. Genes that occur in at least 99% of the population are marked as core genes whereas accessory genes only have to occur at least once in the population. Core genes mostly encode proteins that are involved in housekeeping functions of the organisms. Accessory genes on the other hand can have an adaptive function toward a specific environment or selection pressure and are usually acquired by HGT. Therefore, it is highly probable that these parts of the genome are involved in niche or host adaptation of *Campylobacter.* CgMLST and wgMLST make use of the concept of pan-genomes and establish a novel typing scheme for bacterial strains that, in contrast to MLST, includes all core genes of a species and thereby provides a high resolution by comprising the whole genetic diversity (Sheppard et al. 2013b). Similar to the MLST scheme for *C. jejuni* and *C. coli,* the cgMLST scheme combines *C. jejuni* and *C. coli* and utilizes 1343 gene loci to describe the genetic variation among the strains (Cody et al. 2017).

Due to decreasing costs in WGS and a subsequent increase in bacterial genome sequencing, the concept of GWAS has emerged in the field of microbial genomics (Chen and Shapiro 2015; Lees and Bentley 2016). GWAS is a statistical concept to compare two different phenotypes in order to identify trait-associated genomic compounds. This can be generally used to analyze epidemiology-, resistance- or, in case of *C. jejuni* host-related determinants based on WGS data. Different methods have been developed to apply this method either on entire genes, *k-mer* (word of length *k*), or single nucleotide polymorphism (SNP) level to bacterial populations. In comparison to GWAS tools that are made for human genetic research, these take into account the clonal and lineage-related phylogenetic structure of bacterial populations (Brynildsrud et al. 2016; Power et al. 2017). In order to investigate the host association of *C. jejuni*, a couple of GWAS have been applied in this field of research, mainly for the clinically relevant lineages CC-21 and CC-45 (Sheppard et al. 2013b; Yahara et al. 2017; Thépault et al. 2017; Buchanan et al. 2017). These complexes contain isolates from different hosts of predominantly avian and ruminant origin. Thus, these strains need to adapt frequently to varying environments. For example, chicken and cattle hosts substantially differ in their body temperature, pH level or in the microbiome of their digestive tract. In addition, bacterial cells are exposed to oxidative stress outside the host gut (Kim et al. 2015) during transmission to a new host. Intentionally, many of these studies used a gene-by-gene approach (Yahara et al. 2017; Buchanan et al. 2017), whereas others also keep in mind that core genome adaptation might play a role in host adaptation, especially in host-adapted lineages. Therefore, a *k-mer* approach (Sheppard et al. 2013b; Lees et al. 2018) can not only be applied in order to detect the presence of entire genes but also to identify specific alleles of core genes that may be involved in host adaptation.

4.2 Source Attribution in Clinical and Agricultural Setting

Host-adapted clonal lineages can be observed in several different bacterial pathogens, such as *C. jejuni*, *Staphylococcus aureus* or *Salmonella enterica* on different genetic

levels (2010a, 2011a; Weinert et al. 2012; Hayward et al. 2016; Sheppard et al. 2018). Gene sets of these lineages are affected by several factors, including the host, the composition of food, and by antibiotics and interactions with the host microbiome that can either lead to a temporary or to a permanent adaptation. Genetic mechanisms like DNA replication errors that lead to point mutations, insertions, deletions or recombination events may result in rapid adaptation and the formation of host-specific lineages. In general, *Campylobacter* species are distributed differentially among livestock animals; *C. coli* is dominant in pig-associated samples (Thakur et al. 2006) whereas *C. jejuni* is more abundant in cattle and chicken hosts. Additionally, there might also exist a geographic factor. For example, *Campylobacter* cases in France are more likely to be caused by isolates from ruminant hosts than in other countries (Thépault et al. 2017).

Several studies investigated host adaptation, especially from *C. jejuni*, as this species shows a well-defined lineage separation based on MLST data that distinguish the population into host-specialist and host-generalist clonal complexes (Sheppard et al. 2014). Several colonization studies revealed that modification and differential transcription of motility genes in *C. jejuni* play a key role in adaptation and transmission (Hermans et al. 2011; de Vries et al. 2017; Ren et al. 2018). These data were supported by in vitro experiments as well as by genomic data and by RNA sequencing. Apart from traditional WGS analysis, the novel concepts of GWAS provided great in-depth knowledge about host adaptation, colonization and clinically relevant factors of *C. jejuni*. The group of Sheppard and co-workers discovered multiple genes involved in vitamin B5 biosynthesis and iron uptake within cattle-related strains of the CC-45 complex by applying a *k-mer*-based GWAS (Sheppard et al. 2013b). These genes might be related to different nutrition of cattle host in comparison to poultry. Independently, the same genes have also been detected within a set of 25 diagnostic marker genes by a pan-genome approach leading to the identification of clinically relevant *C. jejuni* isolates with up to 90% accuracy (Buchanan et al. 2017). However, even strains of the clinically relevant complexes CC-21 and CC-45, isolated from poultry processing chains, show substantially different genotypes and carry different genes involved in lipooligosaccharide synthesis (*kpsC, kpsD)*, metabolic processes (*glmS*), oxidative stress response (*nuoK* and *fumC)* as well as genes involved in nucleotide salvage (cj1377c*)* and antimicrobial resistance like efflux proteins (cj1375*)* (Yahara et al. 2017). A pan-genomic approach by Thèpault and co-workers identified 15 additional host-segregation markers in *C. jejuni* isolates from France that might aid to determine the source of clinical cases. Those genes are mainly involved in metabolic processes and nucleotide metabolism. These markers had been utilized to trace back the source of *C. jejuni* infections with an average accuracy of 80.7% for chicken-induced cases and of 68.2% for ruminant-caused cases (Thépault et al. 2017). While numerous studies have focused on the source of *Campylobacter* infections, less work has been dedicated to understanding the genetic mechanisms behind livestock- and environment-specific STs in chickens, cattle or water sources. However, this might generate valuable insights into the evolution and relevant host-specific factors of *Campylobacter* in order to deal with the spread and contamination in livestock environment and further understand the process of adaptation toward clinically relevant pathogens.

4.3 Relevance for Public Health (Applications)

WGS has not only improved our general understanding and knowledge of bacterial populations, adaptations and recombination to date, but is also an important part of routine high-throughput diagnostics for hospitals, for animal husbandry and for surveillance programs of foodborne diseases (Gerner-Smidt et al. 2019). Due to these programs, it is possible to detect a sudden increase in case numbers within a specific time interval. When applying WGS-driven approaches to outbreak detection and source tracking, it is important to distinguish between geographically restricted point-source outbreaks and clusters of cases that are not necessarily related to each other geographically (Llarena et al. 2017). Most outbreaks are diffuse and show a spatial and time-dependent clustering of *Campylobacter* genotypes or subtypes within livestock and clinical cases (Llarena et al. 2017). These outbreaks can spread across several countries, but can be linked to contaminated food products with a low level of contamination. The difficulty in detecting these outbreaks is to be able to distinguish them from sporadic *Campylobacter* cases and to handle the high rate of genetic exchange and recombination within the species (Llarena et al. 2017). This might be achieved by WGS-based molecular characterization in combination with wgMLST or cgMLST that provide the necessary resolution for the genomic comparisons of closely related strains (Deurenberg et al. 2017). For example, a recent wgMLST-based study on the genomic diversity of *C. jejuni* isolates from Israel detected 29 diffuse clusters of genetically related strains that have shown a low variance in allelic differences (Rokney et al. 2018). Importantly, this study further identified adapted clones that kept causing infections over the span of several years. Another study from Finland showed that *C. jejuni* infections, which increased during the summer, were mainly related to three STs with 16 to 37 allelic differences between the cluster, and thus, due to the short period of time, probably belonged to the same source (Kovanen et al. 2014).

In addition to such diffuse outbreaks, point-source outbreaks can also occur; however, those are less frequent and are usually locally restricted. They are mostly related to restaurant meals (Glashower et al. 2017), canteen food (Moffatt et al. 2016) or farming communities (Forbes et al. 2009) and are associated with a high level of contamination within the food products. An appropriate methodology to identify these outbreaks is based on single nucleotide variants (SNVs), because diversity in general should be low and resulting in only a small amount of allelic variants. This approach has been successfully applied in several studies. Moffatt et al. showed a high level of identity in a chicken-related outbreak in Australia with two different genotypes with a SNV difference of only 3–8 SNPs and 30 SNPs, respectively. Additionally, several studies conducted by Revez and colleagues (Revez et al. 2014) demonstrated how wgMLST can be applied in outbreak investigations and source tracing. Patient isolates from milk-born outbreaks shared 1432 loci with isolates from a milk source and only showed three SNPs difference between the strains. Just like for many other bacterial pathogens, WGS-based methods provide a great benefit for *Campylobacter*-related public health applications. However, in contrast to other

bacteria, the high species diversity of *Campylobacter* within the same host often requires an adapted approach.

5 Concluding Remarks

The species of genus *Campylobacter* show a very individual population structure ranging from less clonal diversity to strictly separated clonal lineages. Horizontal gene transfer and recombination events may occur at various levels within the individual population but also between the *Campylobacter* species. Even "hybrid" strains exist that contain large proportion of genomic elements from two species. Modern next-generation sequencing-based methods paved the way for high-resolution molecular typing of outbreak and disease-related strains by applying a standardized typing scheme based on the whole core genome and, additionally, the pangenome. Further, they also allowed the identification of genomic factors that contribute to host adaptation of individual lineages on the gene and allele level and to trace the source of several *Campylobacter* lineages. This contribution to tracing and unraveling transmission and infection chains results in important public health applications to contain this important zoonotic pathogen.

Acknowledgements This work is supported by the German Federal Ministry of Education and Research (BMBF) in frame of the zoonoses research consortium PAC-Campylobacter to T. S. (project IP10B/01KI1725F)

References

Aabenhus R, On SLW, Siemer BL, Permin H, Andersen LP (2005) Delineation of *Campylobacter concisus* genomospecies by amplified fragment length polymorphism analysis and correlation of results with clinical data. J Clin Microbiol 43:5091–5096. https://doi.org/10.1128/JCM.43.10.5091-5096.2005

Altekruse SF, Stern NJ, Fields PI, Swerdlow DL (1999) *Campylobacter jejuni*—an emerging foodborne pathogen. Emerg Infect Dis 5:28–35. https://doi.org/10.3201/eid0501.990104

Asakura H, Brüggemann H, Sheppard SK, Ekawa T, Meyer TF, Yamamoto S, Igimi S (2012) Molecular evidence for the thriving of *Campylobacter jejuni* ST-4526 in Japan. PLoS ONE 7:e48394. https://doi.org/10.1371/journal.pone.0048394

Bourke B, Chan VL, Sherman P (1998) *Campylobacter upsaliensis*: waiting in the wings. Clin Microbiol Rev 11:440–449. https://doi.org/10.1128/CMR.11.3.440

Brynildsrud O, Bohlin J, Scheffer L, Eldholm V (2016) Erratum to: rapid scoring of genes in microbial pan-genome-wide association studies with Scoary. Genome Biol 17:262. https://doi.org/10.1186/s13059-016-1132-8

Buchanan CJ, Webb AL, Mutschall SK, Kruczkiewicz P, Barker DORR, Hetman BM, Gannon VPJJ, Abbott DW, Thomas JE, Inglis GD, Taboada EN (2017) A genome-wide association study to identify diagnostic markers for human pathogenic *Campylobacter jejuni* strains. Front Microbiol 8:1224. https://doi.org/10.3389/fmicb.2017.01224

Caniça M, Manageiro V, Abriouel H, Moran-Gilad J, Franz CMAP (2019) Antibiotic resistance in foodborne bacteria. Trends Food Sci Technol 84:41–44. https://doi.org/10.1016/j.tifs.2018.08.001

Chen PE, Shapiro BJ (2015) The advent of genome-wide association studies for bacteria. Curr Opin Microbiol 25:17–24. https://doi.org/10.1016/j.mib.2015.03.002

Chung HKL, Tay A, Octavia S, Chen J, Liu F, Ma R, Lan R, Riordan SM, Grimm MC, Zhang L (2016) Genome analysis of *Campylobacter concisus* strains from patients with inflammatory bowel disease and gastroenteritis provides new insights into pathogenicity. Sci Rep 6:38442. https://doi.org/10.1038/srep38442

Cody AJ, Bray JE, Jolley KA, McCarthy ND, Maiden MCJ (2017) Core genome multilocus sequence typing scheme for stable, comparative analyses of *Campylobacter jejuni* and *C. coli* human disease isolates. J Clin Microbiol 55:2086–2097. https://doi.org/10.1128/JCM.00080-17

Cody AJ, McCarthy NM, Wimalarathna HL, Colles FM, Clark L, Bowler ICJW, Maiden MCJ, Dingle KE (2012) A Longitudinal 6-year study of the molecular epidemiology of clinical *Campylobacter* isolates in Oxfordshire, United Kingdom. J Clin Microbiol 50:3193–3201. https://doi.org/10.1128/JCM.01086-12

Colles FM, Jones K, Harding RM, Maiden MCJ (2003) Genetic diversity of *Campylobacter jejuni* isolates from farm animals and the farm environment. Appl Environ Microbiol 69:7409–7413. https://doi.org/10.1128/AEM.69.12.7409-7413.2003

Colles FM, Jones TA, McCarthy ND, Sheppard SK, Cody AJ, Dingle KE, Dawkins MS, Maiden MCJJ (2008) *Campylobacter* infection of broiler chickens in a free-range environment. Environ Microbiol 10:2042–2050. https://doi.org/10.1111/j.1462-2920.2008.01623.x

Colles FM, McCarthy ND, Howe JC, Devereux CL, Gosler AG, Maiden MCJ (2009) Dynamics of *Campylobacter* colonization of a natural host, *Sturnus vulgaris* (European Starling). Environ Microbiol 11:258–267. https://doi.org/10.1111/j.1462-2920.2008.01773.x

Costa D, Iraola G (2019) Pathogenomics of emerging *Campylobacter* species. Clin Microbiol Rev 32:0001. https://doi.org/10.1128/CMR.00072-18

de Mazancourt C, Johnson E, Barraclough TG (2008) Biodiversity inhibits species' evolutionary responses to changing environments. Ecol Lett 11:380–388. https://doi.org/10.1111/j.1461-0248.2008.01152.x

de Vries SPW, Gupta S, Baig A, Wright E, Wedley A, Jensen AN, Lora LL, Humphrey S, Skovgard H, MacLeod K, Pont E, Wolanska DP, L'Heureux J, Mobegi FM, Smith DGEE, Everest P, Zomer A, Williams N, Wigley P, Humphrey T, Maskel DJ, Grant AJ, Skovgård H, MacLeod K, Pont E, Wolanska DP, L'Heureux J, Mobegi FM, Smith DGEE, Everest P, Zomer A, Williams N, Wigley P, Humphrey T, Maskell DJ, Grant AJ, Skovgard H, MacLeod K, Pont E, Wolanska DP, L'Heureux J, Mobegi FM, Smith DGEE, Everest P, Zomer A, Williams N, Wigley P, Humphrey T, Maskell DJ, Grant AJ, Skovgård H, MacLeod K, Pont E, Wolanska DP, L'Heureux J, Mobegi FM, Smith DGEE, Everest P, Zomer A, Williams N, Wigley P, Humphrey T, Maskell DJ, Grant AJ (2017) Genome-wide fitness analyses of the foodborne pathogen *Campylobacter jejuni* in in vitro and in vivo models. Sci Rep 7:1251. https://doi.org/10.1038/s41598-017-01133-4

Dearlove BL, Cody AJ, Pascoe B, Méric G, Wilson DJ, Sheppard SK (2016) Rapid host switching in generalist *Campylobacter* strains erodes the signal for tracing human infections. ISME J 10:721–729. https://doi.org/10.1038/ismej.2015.149

Debruyne L, On SLW, De Brandt E, Vandamme P (2009) Novel *Campylobacter lari*-like bacteria from humans and molluscs: description of *Campylobacter peloridis* sp. nov., *Campylobacter lari* subsp. *concheus* subsp. nov. and *Campylobacter lari* subsp. *lari* subsp. nov. Int J Syst Evol Microbiol 59:1126–1132. https://doi.org/10.1099/ijs.0.000851-0

Deurenberg RH, Bathoorn E, Chlebowicz MA, Couto N, Ferdous M, García-Cobos S, Kooistra-Smid AMD, Raangs EC, Rosema S, Veloo ACM, Zhou K, Friedrich AW, Rossen JWA (2017) Application of next generation sequencing in clinical microbiology and infection prevention. J Biotechnol 243:16–24. https://doi.org/10.1016/j.jbiotec.2016.12.022

Dingle KE, Colles FM, Falush D, Maiden MCJ (2005) Sequence typing and comparison of population biology of *Campylobacter coli* and *Campylobacter jejuni*. J Clin Microbiol 43:340–347. https://doi.org/10.1128/JCM.43.1.340-347.2005

Dingle KE, Colles FM, Wareing DRAA, Ure R, Fox AJ, Bolton FE, Bootsma HJ, Willems RJL, Urwin R, Maiden MCJ (2001) Multilocus sequence typing system for *Campylobacter jejuni*. J Clin Microbiol 39:14–23. https://doi.org/10.1128/JCM.39.1.14-23.2001

Duchêne S, Holt KE, Weill F-X, Le Hello S, Hawkey J, Edwards DJ, Fourment M, Holmes EC (2016) Genome-scale rates of evolutionary change in bacteria. Microbial Genomics 2:e000094. https://doi.org/10.1099/mgen.0.000094

Duim B, Wassenaar TM, Rigter A, Wagenaar J (1999) High-resolution genotyping of *Campylobacter* strains isolated from poultry and humans with amplified fragment length polymorphism fingerprinting. Appl Environ Microbiol 65:2369–2375. https://doi.org/10.1128/AEM.65.6.2369-2375.1999

Eyre-Walker A, Keightley PD (2007) The distribution of fitness effects of new mutations. Nat Rev Genet 8:610–618. https://doi.org/10.1038/nrg2146

Falush D (2016) Bacterial genomics: microbial GWAS coming of age. Nat Microbiol 1:16059. https://doi.org/10.1038/nmicrobiol.2016.59

Forbes KJ, Gormley FJ, Dallas JF, Labovitiadi O, MacRae M, Owen RJ, Richardson J, Strachan NJC, Cowden JM, Ogden ID, McGuigan CC (2009) *Campylobacter* immunity and coinfection following a large outbreak in a farming community. J Clin Microbiol 47:111–116. https://doi.org/10.1128/JCM.01731-08

Gerner-Smidt P, Besser J, Concepción-Acevedo J, Folster JP, Huffman J, Joseph LA, Kucerova Z, Nichols MC, Schwensohn CA, Tolar B (2019) Whole genome sequencing: bridging one-health surveillance of foodborne diseases. Front Public Health 7:0001. https://doi.org/10.3389/fpubh.2019.00172

Gillespie IA, O'Brien SJ, Frost JA, Adak GK, Horby P, Swan AV, Painter MJ, Neal KR (2002) A case-case Comparison of *Campylobacter coli* and *Campylobacter jejuni* infection: a tool for generating hypotheses. Emerg Infect Dis 8:937–942. https://doi.org/10.3201/eid0809.010817

Glashower D, Snyder J, Welch D, McCarthy S (2017) Notes from the field: outbreak of *Campylobacter jejuni* associated with consuming undercooked chicken liver Mousse—Clark County, Washington, 2016. MMWR Morbidity and Mortality Weekly Report 66:1027. https://doi.org/10.15585/mmwr.mm6638a4

Golz JC, Epping L, Knüver M-T, Borowiak M, Hartkopf F, Deneke C, Malorny B, Semmler T, Stingl K (2020) Whole genome sequencing reveals extended natural transformation in *Campylobacter* impacting diagnostics and the pathogens adaptive potential. Sci Rep 10:3686. https://doi.org/10.1038/s41598-020-60320-y

Goossens H, Vlaes L, De Boeck M, Levy J, De Mol P, Butzler J-P, Kersters K, Pot B, Vandamme P (1990) Is "*Campylobacter upsaliensis*" an unrecognised cause of human diarrhoea? Lancet 335:584–586. https://doi.org/10.1016/0140-6736(90)90359-D

Hayward MR, Petrovska L, Jansen VAA, Woodward MJ (2016) Population structure and associated phenotypes of *Salmonella enterica* serovars Derby and Mbandaka overlap with host range. BMC Microbiol 16:15. https://doi.org/10.1186/s12866-016-0628-4

Hermans D, Van Deun K, Martel A, Van Immerseel F, Messens W, Heyndrickx M, Haesebrouck F, Pasmans F (2011) Colonization factors of *Campylobacter jejuni* in the chicken gut. Vet Res 42:82. https://doi.org/10.1186/1297-9716-42-82

Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S (2018) High throughput ANI analysis of 90 K prokaryotic genomes reveals clear species boundaries. Nat Commun 9:5114. https://doi.org/10.1038/s41467-018-07641-9

Kaakoush NO, Castaño-Rodríguez N, Mitchell HM, Man SM (2015) Global epidemiology of *Campylobacter* infection. Clin Microbiol Rev 28:687–720. https://doi.org/10.1128/CMR.00006-15

Kalischuk LD, Inglis GD (2011) Comparative genotypic and pathogenic examination of *Campylobacter concisus* isolates from diarrheic and non-diarrheic humans. BMC Microbiol 11:53. https://doi.org/10.1186/1471-2180-11-53

Kim J-C, Oh E, Kim J, Jeon B (2015) Regulation of oxidative stress resistance in *Campylobacter jejuni*, a microaerophilic foodborne pathogen. Front Microbiol 6:0001. https://doi.org/10.3389/fmicb.2015.00751

Kolde R (2015) Pheatmap : pretty heatmaps. R package version 1.0.8

Kovanen SM, Kivisto RI, Rossi M, Schott T, Karkkainen U-M, Tuuminen T, Uksila J, Rautelin H, Hanninen M-L (2014) Multilocus sequence typing (MLST) and whole-genome MLST of *Campylobacter jejuni* isolates from human infections in three districts during a seasonal peak in Finland. J Clin Microbiol 52:4147–4154. https://doi.org/10.1128/JCM.01959-14

Lees JA, Bentley SD (2016) Bacterial GWAS: not just gilding the lily. Nat Rev Microbiol 14:406. https://doi.org/10.1038/nrmicro.2016.82

Lees JA, Galardini M, Bentley SD, Weiser JN, Corander J (2018) pyseer: a comprehensive tool for microbial pangenome-wide association studies. Bioinformatics 34:4310–4312. https://doi.org/10.1093/bioinformatics/bty539

Lentzsch P, Rieksneuwohner B, Wieler LH, Hotzel H, Moser I (2004) High-resolution genotyping of *Campylobacter upsaliensis* strains originating from three continents. J Clin Microbiol 42:3441–3448. https://doi.org/10.1128/JCM.42.8.3441-3448.2004

Liu F, Ma R, Tay CYA, Octavia S, Lan R, Chung HKL, Riordan SM, Grimm MC, Leong RW, Tanaka MM, Connor S, Zhang L (2018) Genomic analysis of oral *Campylobacter concisus* strains identified a potential bacterial molecular marker associated with active Crohn's disease. Emer Micro Inf 7:1–14. https://doi.org/10.1038/s41426-018-0065-6

Llarena A-K, Taboada E, Rossi M (2017) Whole-genome sequencing in epidemiology of *Campylobacter jejuni* infections. J Clin Microbiol 55:1269–1275. https://doi.org/10.1128/JCM.00017-17

Maiden MCJ, Bygraves JA, Feil E, Morelli G, Russell JE, Urwin R, Zhang Q, Zhou J, Zurth K, Caugant DA, Feavers IM, Achtman M, Spratt BG (1998) Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. Proc Natl Acad Sci 95:3140–3145. https://doi.org/10.1073/pnas.95.6.3140

Man SM (2011) The clinical importance of emerging *Campylobacter* species. Nat Rev Gastroenterol Hepatol 8:669–685. https://doi.org/10.1038/nrgastro.2011.191

Manning G, Dowson CG, Bagnall MC, Ahmed IH, West M, Newell DG (2003) Multilocus sequence typing for comparison of veterinary and human isolates of *Campylobacter jejuni*. Appl Environ Microbiol 69:6370–6379. https://doi.org/10.1128/AEM.69.11.6370-6379.2003

McTavish SM, Pope CE, Nicol C, Sexton K, French N, Carter PE (2008) Wide geographical distribution of internationally rare *Campylobacter* clones within New Zealand. Epidemiol Infect 136:1244–1252. https://doi.org/10.1017/S0950268807009892

Medini D, Donati C, Tettelin H, Masignani V, Rappuoli R (2005) The microbial pan-genome. Curr Opin Genet Dev 15:589–594. https://doi.org/10.1016/j.gde.2005.09.006

Miller WG, Englen MD, Kathariou S, Wesley IV, Wang G, Pittenger-Alley L, Siletz RM, Muraoka W, Fedorka-Cray PJ, Mandrell RE (2006) Identification of host-associated alleles by multilocus sequence typing of *Campylobacter coli* strains from food animals. Microbiology 152:245–255. https://doi.org/10.1099/mic.0.28348-0

Miller WG, On SLW, Wang G, Fontanoz S, Lastovica AJ, Mandrell RE (2005) Extended multilocus sequence typing system for *Campylobacter coli*, *C. lari*, *C. upsaliensis*, and *C. helveticus*. J Clin Microbiol 43:2315–2329. https://doi.org/10.1128/JCM.43.5.2315-2329.2005

Miller WG, Yee E, Chapman MH, Smith TPL, Bono JL, Huynh S, Parker CT, Vandamme P, Luong K, Korlach J (2014) Comparative genomics of the *Campylobacter lari* Group. Genome Biol Evol 6:3252–3266. https://doi.org/10.1093/gbe/evu249

Mira A, Pushker R, Rodríguez-Valera F (2006) The Neolithic revolution of bacterial genomes. Trends Microbiol 14:200–206. https://doi.org/10.1016/j.tim.2006.03.001

Moffatt CRM, Greig A, Valcanis M, Gao W, Seemann T, Howden BP, Kirk MD (2016) A large outbreak of *Campylobacter jejuni* infection in a university college caused by chicken liver pâté, Australia, 2013. Epidemiol Infect 144:2971–2978. https://doi.org/10.1017/S0950268816001187

Mohan V, Stevenson M, Marshall J, Fearnhead P, Holland BR, Hotter G, French NP (2013) *Campylobacter jejuni* colonization and population structure in urban populations of ducks and starlings in New Zealand. MicrobiologyOpen 2:659–673. https://doi.org/10.1002/mbo3.102

Møller Nielsen E (1997) Distribution of serotypes of *Campylobacter jejuni* and *C. coli* from Danish patients, poultry, cattle and swine. FEMS Immunol Med Microbiol 19:47–56. https://doi.org/10.1016/S0928-8244(97)00049-7

Nachamkin I, Bohachick K, Patton CM (1993) Flagellin gene typing of *Campylobacter jejuni* by restriction fragment length polymorphism analysis. J Clin Microbiol 31:1531–1536. https://doi.org/10.1128/JCM.31.6.1531-1536.1993

Nohra A, Grinberg A, Midwinter AC, Marshall JC, Collins-Emerson JM, French NP (2016) Molecular epidemiology of *Campylobacter coli* strains isolated from different sources in New Zealand between 2005 and 2014. Appl Environ Microbiol 82:4363–4370. https://doi.org/10.1128/AEM.00934-16

Ochman H, Wilson AC (1987) Evolution in bacteria: evidence for a universal substitution rate in cellular genomes. J Mol Evol 26:74–86. https://doi.org/10.1007/BF02111283

Oyarzabal OA, Backert S (2012) Microbial food safety: an introduction (food science text series). Springer, Heidelberg

Parkhill J, Wren BW, Mungall K, Ketley JM, Churcher C, Basham D, Chillingworth T, Davies RM, Feltwell T, Holroyd S, Jagels K, Karlyshev AV, Moule S, Pallen MJ, Penn CW, Quail MA, Rajandream M-A, Rutherford KM, van Vliet AHM, Whitehead S, Barrell BG (2000) The genome sequence of the food-borne pathogen *Campylobacter jejuni* reveals hypervariable sequences. Nature 403:665–668. https://doi.org/10.1038/35001088

Pascoe B, Méric G, Yahara K, Wimalarathna H, Murray S, Hitchings MD, Sproston EL, Carrillo CD, Taboada EN, Cooper KK, Huynh S, Cody AJ, Jolley KA, Maiden MCJ, McCarthy ND, Didelot X, Parker CT, Sheppard SK (2017) Local genes for local bacteria: evidence of allopatry in the genomes of transatlantic *Campylobacter* populations. Mol Ecol. https://doi.org/10.1111/mec.14176

Potturi-Venkata LP, Backert S, Lastovica AJ, Vieira SL, Norton RA, Miller RS, Pierce S, Oyarzabal OA (2007) Evaluation of different plate media for direct cultivation of *Campylobacter* species from live broilers. Poult Sci 86:1304–1311. https://doi.org/10.1093/ps/86.7.1304

Power RA, Parkhill J, de Oliveira T (2017) Microbial genome-wide association studies: lessons from human GWAS. Nat Rev Genet 18:41–50. https://doi.org/10.1038/nrg.2016.132

Price MN, Dehal PS, Arkin AP (2010) FastTree 2—approximately maximum-likelihood trees for large alignments. PLoS ONE 5:e9490. https://doi.org/10.1371/journal.pone.0009490

Ren F, Li X, Tang H, Jiang Q, Yun X, Fang L, Huang P, Tang Y, Li Q, Huang J, Jiao X (2018) Insights into the impact of flhF inactivation on *Campylobacter jejuni* colonization of chick and mice gut. BMC Microbiol 18:149. https://doi.org/10.1186/s12866-018-1318-1

Revez J, Rossi M, Ellström P, de Haan C, Rautelin H, Hänninen M-L (2011) Finnish *Campylobacter jejuni* strains of multilocus sequence type ST-22 complex have two lineages with different characteristics. PLoS ONE 6:e26880. https://doi.org/10.1371/journal.pone.0026880

Revez J, Zhang J, Schott T, Kivistö R, Rossi M, Hänninen ML (2014) Genomic variation between *Campylobacter jejuni* isolates associated with milk-borne-disease outbreaks. J Clin Microbiol 52(8):2782–6. https://doi.org/10.1128/JCM.00931-14.

Rokney A, Valinsky L, Moran-Gilad J, Vranckx K, Agmon V, Weinberger M (2018) Genomic epidemiology of *Campylobacter jejuni* transmission in Israel. Front Microbiol 9:0001. https://doi.org/10.3389/fmicb.2018.02432

Rouli L, Merhej V, Fournier P-E, Raoult D (2015) The bacterial pangenome as a new tool for analysing pathogenic bacteria. New Microbes New Infections 7:72–85. https://doi.org/10.1016/j.nmni.2015.06.005

Sandstedt K, Ursing J, Walder M (1983) Thermotolerant *Campylobacter* with no or weak catalase activity isolated from dogs. Curr Microbiol 8:209–213. https://doi.org/10.1007/BF01579548

Sheppard SK, Cheng L, Méric G, De Haan CPA, Llarena AK, Marttinen P, Vidal A, Ridley A, Clifton-Hadley F, Connor TR, Strachan NJC, Forbes K, Colles FM, Jolley KA, Bentley SD, Maiden MCJ, Hänninen ML, Parkhill J, Hanage WP, Corander J (2014) Cryptic ecology among host generalist *Campylobacter jejuni* in domestic animals. Mol Ecol 23:2442–2451. https://doi.org/10.1111/mec.12742

Sheppard SK, Colles F, Richardson J, Cody AJ, Elson R, Lawson A, Brick G, Meldrum R, Little CL, Owen RJ, Maiden MCJ, McCarthy ND (2010a) Host association of *Campylobacter* genotypes transcends geographic variation. Appl Environ Microbiol 76:5269–5277. https://doi.org/10.1128/AEM.00124-10

Sheppard SK, Colles FM, McCarthy ND, Strachan NJC, Ogden ID, Forbes KJ, Dallas JF, Maiden MCJ (2011a) Niche segregation and genetic structure of *Campylobacter jejuni* populations from wild and agricultural host species. Mol Ecol 20:3484–3490. https://doi.org/10.1111/j.1365-294X.2011.05179.x

Sheppard SK, Dallas JF, Wilson DJ, Strachan NJCC, McCarthy ND, Jolley KA, Colles FM, Rotariu O, Ogden ID, Forbes KJ, Maiden MCJ (2010b) Evolution of an agriculture-associated disease causing *Campylobacter coli* clade: evidence from national surveillance data in Scotland. PLoS ONE 5:e15708. https://doi.org/10.1371/journal.pone.0015708

Sheppard SK, Didelot X, Jolley KA, Darling AE, Pascoe B, Meric G, Kelly DJ, Cody A, Colles FM, Strachan NJC, Ogden ID, Forbes K, French NP, Carter P, Miller WG, McCarthy ND, Owen R, Litrup E, Egholm M, Affourtit JP, Bentley SD, Parkhill J, Maiden MCJ, Falush D (2013a) Progressive genome-wide introgression in agricultural *Campylobacter coli*. Mol Ecol 22:1051–1064. https://doi.org/10.1111/mec.12162

Sheppard SK, Didelot X, Meric G, Torralbo A, Jolley KA, Kelly DJ, Bentley SD, Maiden MCJ, Parkhill J, Falush D (2013b) Genome-wide association study identifies vitamin B5 biosynthesis as a host specificity factor in *Campylobacter*. Proc Natl Acad Sci USA 110:11923–11927. https://doi.org/10.1073/pnas.1305559110

Sheppard SK, Guttman DS, Fitzgerald JR (2018) Population genomics of bacterial host adaptation. Nat Rev Genet 19:549–565. https://doi.org/10.1038/s41576-018-0032-z

Sheppard SK, Maiden MCJ (2015) The Evolution of *Campylobacter jejuni* and *Campylobacter coli*. Cold Spring Harb Perspect Biol 7:a018119. https://doi.org/10.1101/cshperspect.a018119

Sheppard SK, McCarthy ND, Falush D, Maiden MCJ (2008) Convergence of *Campylobacter* species: implications for bacterial evolution. Science 320:237–239. https://doi.org/10.1126/science.1155532

Sheppard SK, McCarthy ND, Jolley KA, Maiden MCJ (2011b) Introgression in the genus *Campylobacter*: generation and spread of mosaic alleles. Microbiology 157:1066–1074. https://doi.org/10.1099/mic.0.045153-0

Skarp-de Haan CPA, Culebro A, Schott T, Revez J, Schweda EKH, Hänninen M-L, Rossi M (2014) Comparative genomics of unintrogressed *Campylobacter coli* clades 2 and 3. BMC Genom 15:129. https://doi.org/10.1186/1471-2164-15-129

Smith J (1992) Analyzing the mosaic structure of genes. J Mol Evol 34:0001. https://doi.org/10.1007/BF00182389

Suerbaum S, Lohrengel M, Sonnevend A, Ruberg F, Kist M (2001) Allelic Diversity and recombination in *Campylobacter jejuni*. J Bacteriol 183:2553–2559. https://doi.org/10.1128/JB.183.8.2553-2559.2001

Tagini F, Greub G (2017) Bacterial genome sequencing in clinical microbiology: a pathogen-oriented review. Eur J Clin Microbiol Infect Dis 36:2007–2020. https://doi.org/10.1007/s10096-017-3024-6

Thakur S, Morrow WEM, Funk JA, Bahnson PB, Gebreyes WA (2006) Molecular Epidemiologic investigation of *Campylobacter coli* in swine production systems, using multilocus sequence typing. Appl Environ Microbiol 72:5666–5669. https://doi.org/10.1128/AEM.00658-06

Thépault A, Méric G, Rivoal K, Pascoe B, Mageiros L, Touzain F, Rose V, Béven V, Chemaly M, Sheppard SK (2017) Genome-wide identification of host-segregating epidemiological markers for source attribution in *Campylobacter jejuni*. Appl Environ Microbiol 83:e03085–e03116. https://doi.org/10.1128/AEM.03085-16

Toft C, Andersson SGE (2010) Evolutionary microbial genomics: insights into bacterial host adaptation. Nat Rev Genet 11:465–475. https://doi.org/10.1038/nrg2798

Van Alfen NK (2015) Encyclopedia of agriculture and food systems. Choice Rev Online 52:52-2875–52-2875. https://doi.org/10.5860/CHOICE.188216

Vandamme P, Falsen E, Pot B, Hoste B, Kersters K, De Ley J (1989) Identification of EF group 22 campylobacters from gastroenteritis cases as *Campylobacter concisus*. J Clin Microbiol 27:1775–1781. https://doi.org/10.1128/JCM.27.8.1775-1781.1989

Vandamme P, Falsen E, Rossau R, Hoste B, Segers P, Tytgat R, De Ley J (1991) Revision of *Campylobacter*, *Helicobacter*, and *Wolinella* taxonomy: emendation of generic descriptions and proposal of Arcobacter gen. nov. Int J Syst Bacteriol 41:88–103. https://doi.org/10.1099/00207713-41-1-88

Vidal AB, Colles FM, Rodgers JD, McCarthy ND, Davies RH, Maiden MCJ, Clifton-Hadley FA (2016) Genetic diversity of *Campylobacter jejuni* and *Campylobacter coli* isolates from conventional broiler flocks and the impacts of sampling strategy and laboratory method. Appl Environ Microbiol 82:2347–2355. https://doi.org/10.1128/AEM.03693-15

Weinert LA, Welch JJ, Suchard MA, Lemey P, Rambaut A, Fitzgerald JR (2012) Molecular dating of human-to-bovid host jumps by *Staphylococcus aureus* reveals an association with the spread of domestication. Biol Let 8:829–832. https://doi.org/10.1098/rsbl.2012.0290

Wilson DJ, Gabriel E, Leatherbarrow AJH, Cheesbrough J, Gee S, Bolton E, Fox A, Hart CA, Diggle PJ, Fearnhead P (2009) Rapid evolution and the importance of recombination to the gastroenteric pathogen *Campylobacter jejuni*. Mol Biol Evol 26:385–397. https://doi.org/10.1093/molbev/msn264

Wirth T, Falush D, Lan R, Colles F, Mensa P, Wieler LH, Karch H, Reeves PR, Maiden MCJ, Ochman H, Achtman M (2006) Sex and virulence in *Escherichia coli*: an evolutionary perspective. Mol Microbiol 60:1136–1151. https://doi.org/10.1111/j.1365-2958.2006.05172.x

Yahara K, Méric G, Taylor AJ, de Vries SPW, Murray S, Pascoe B, Mageiros L, Torralbo A, Vidal A, Ridley A, Komukai S, Wimalarathna H, Cody AJ, Colles FM, McCarthy N, Harris D, Bray JE, Jolley KA, Maiden MCJ, Bentley SD, Parkhill J, Bayliss CD, Grant A, Maskell D, Didelot X, Kelly DJ, Sheppard SK (2017) Genome-wide association of functional traits linked with *Campylobacter jejuni* survival from farm to fork. Environ Microbiol 19:361–380. https://doi.org/10.1111/1462-2920.13628

Yan W, Chang N, Taylor DE (1991) Pulsed-field gel electrophoresis of *Campylobacter jejuni* and *Campylobacter coli* genomic DNA and its epidemiologic application. J Infect Dis 163:1068–1072. https://doi.org/10.1093/infdis/163.5.1068