AMERICAN SOCIETY FOR MICROBIOLOGY | Journal of Clinical Microbiology®

MINIREVIEW

Whole-Genome Sequencing in Epidemiology of *Campylobacter jejuni* Infections

Ann-Katrin Llarena,[a] Eduardo Taboada,[b] Mirko Rossi[a]

Department of Food Hygiene and Environmental Health, Faculty of Veterinary Medicine, University of Helsinki, Helsinki, Finland[a]; National Microbiology Laboratory at Lethbridge, Animal Diseases Research Institute, Public Health Agency of Canada, Lethbridge, Alberta, Canada[b]

ABSTRACT This review describes the current state of knowledge regarding the application of whole-genome sequencing (WGS) in the epidemiology of *Campylobacter jejuni*, the leading cause of bacterial gastroenteritis worldwide. We describe how WGS has increased our understanding of the evolutionary and epidemiological dynamics of this pathogen and how WGS has the potential to improve surveillance and outbreak detection. We have identified hurdles to the full implementation of WGS in public health settings. Despite these challenges, we think that ample evidence is available to support the benefits of integrating WGS into the routine monitoring of *C. jejuni* infections and outbreak investigations.

KEYWORDS foodborne pathogens, genome analysis, molecular subtyping, surveillance studies

Accepted manuscript posted online 1 March 2017

Citation 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.

Editor Colleen Suzanne Kraft, Emory University

Address correspondence to Mirko Rossi, mirko.rossi@helsinki.fi.

The ability to conduct epidemiological investigations and to intervene to control and to prevent foodborne and environmentally transmitted diseases is a major task of public health authorities. The identification, over a short period of time, of a sudden increase in the number of expected cases of a disease in a population in a limited geographical area (point-source outbreak) or clusters of cases with a presumptive common source not necessarily clustered geographically (diffuse outbreak) depends on knowledge of the baseline infectious state of the population regarding that disease, which is usually acquired through surveillance. In this regard, identifying epidemiologically linked cases and differentiating them from concurrent sporadic incidences are essential for risk assessment, outbreak investigations, and source attribution for foodborne pathogens. These processes have relied increasingly on traditional epidemiological investigations supplemented with molecular subtyping of the etiological agent, and no method offers a higher degree of resolution than whole-genome sequencing (WGS). The recent development of high-throughput sequencing technologies for WGS (i.e., next-generation sequencing [NGS]) has resulted in large-scale sequencing of various pathogens. Allowing the detection of all possible epidemiologically significant variations between strains (1), WGS is progressively replacing traditional typing methods (serotyping, phenotyping, pulsed-field gel electrophoresis [PFGE], and amplified fragment length polymorphism [AFLP] analysis) and sequence-based investigations (PCR-based methods). Moreover, the declining costs of NGS and the availability of benchtop analyzers are increasingly facilitating the application of WGS for routine surveillance and outbreak investigations of bacterial and viral infectious diseases by public health authorities (2). As a result, WGS analysis is currently being used in several countries for real-time surveillance of *Listeria monocytogenes* and *Salmonella enterica* (http://www.fda.gov/Food/FoodScienceResearch/WholeGenomeSequencingProgramWGS) (3), and similar approaches for other foodborne pathogens are expected to come into use shortly. *Campylobacter jejuni* is one of the most frequent causes of bacterial gastroenteritis globally; the epidemiology

of this pathogen is only partially understood, however, and shedding new light on this area is difficult, as most cases appear to be sporadic and remain unreported (4). Additionally, the occurrence of immune individuals in the population, coupled with many possible reservoirs, transmission pathways, and regional epidemiological differences, make source tracing and risk assessment challenging. To date, the two most common molecular typing methods, namely, PFGE and the traditional seven-locus multilocus sequence typing (MLST) scheme, have been essential research tools for studying the epidemiology of *C. jejuni* infections and have greatly contributed to current knowledge regarding the nature of *C. jejuni* infections in human patients and potential reservoirs (1). However, these methods have been difficult to implement in the context of routine surveillance; moreover, limitations of these methods have been uncovered, suggesting that these methods may be unsuitable as the sole subtyping methods for assessing epidemiological links among *C. jejuni* isolates (5–7). Therefore, developing a new generation of efficient tools for epidemiological surveillance of *C. jejuni* is clearly necessary and a high priority in order to overcome the limitations of available typing methods (1). WGS is set to emerge as the typing method of choice for *C. jejuni* outbreak investigations, and it has the potential for implementation in routine surveillance. However, several issues regarding epidemiology and genomic diversity need to be addressed before WGS can become a useful and reliable working tool in the public health sector response to campylobacteriosis. In particular, understanding the baseline ecological diversity and population structure of *C. jejuni* and the genomic diversity within a single human infection is necessary for defining interpretation criteria for determining whether isolates are likely to be clonal, as would be expected in a cluster of cases or an outbreak with a common source. In this review, we show how WGS has improved, or has the capacity to improve, surveillance and outbreak detection for *C. jejuni* infections and we discuss how WGS has increased our understanding of both the evolutionary and epidemiological dynamics of campylobacteriosis caused by *C. jejuni*.

DEFINING THE BASELINE GENOMIC DIVERSITY OF *C. JEJUNI*

Assessing the genomic diversity present in a *C. jejuni* population is critical for deducing whether genomic differences between any two *C. jejuni* isolates are sufficiently small for assuming a high likelihood of epidemiological relatedness, i.e., whether the isolates share a source. An important consideration that has rarely been discussed in the literature is the potential for heterogeneity in the levels of genetic diversity observed within various lineages in the *C. jejuni* population. However, defining such diversity is necessary for accurate interpretation of the observed genetic relatedness among isolates; if the lineage diversity is generally low, then the observed genetic distance between two isolates might fall below the presumed cutoff value for cluster definition even in the absence of epidemiological linkage. Evidence for variations in genetic diversity among *C. jejuni* lineages is mounting. For instance, Kovanen and colleagues found limited genomic diversity within three common sequence types (STs), namely, ST-230, ST-267, and ST-677, while ST-45 was considerably more diverse and was separated into three main sublineages (8). In a follow-up publication, Llarena et al. investigated the nature of the population structure of ST-45 clonal complex (CC), focusing on identifying a possible spatial-temporal evolutionary signal. The authors found that the occurrence and strength of the geographical signal varied between sublineages of ST-45 CC strains, but they could not find evidence of a temporal signal (6). In addition, the authors unexpectedly identified certain sublineages of ST-45 with extremely similar genomes regardless of the time and location of sampling. These successful monomorphic clones were persistently isolated from animal hosts and human patients over a decade, from several countries around the world (6). There is no reason to think that clonal dispersal is limited to certain sublineages of ST-45. Indeed, in a recent publication, the clonal expansion during the 1970s of the highly virulent ST-8 lineage, which is now causing vast numbers of ovine abortions in the United States, was reported (9). The presence of clonal populations makes genomic distinction

between epidemiologically associated and nonrelated isolates difficult and can adversely affect the use of WGS methodology in surveillance and outbreak investigations for public health purposes; more studies exploring the genomic diversity and possible occurrence of clonal subpopulations in other *C. jejuni* lineages are needed to resolve this predicament.

GENOMIC DIVERSITY DURING *C. JEJUNI* INFECTION AND COLONIZATION

The genomic variants that accumulate during human passage should be considered when the level of genetic similarity between epidemiologically linked isolates is being established. The generation of diversity in this population is dependent on the mutation rate and mutational patterns of *C. jejuni*. The fixation of such mutations is determined by evolutionary processes such as genetic drift, the bottleneck effect, and selection and can be observed due to host adaption (10). WGS has been applied to study the genomic diversity and changes that can accumulate during human infection (11, 12) and animal colonization (13, 14). The genomic changes that occurred during host passage (either human or animal) were small and, except for rare single nucleotide variants (SNVs) at one or two loci, limited to indels in homopolymeric tracts in contingency loci. These frameshift mutations were generally found in genes regulating phase variations of surface structures and therefore are likely to play an important role in host adaptation (11–14). Furthermore, these mutations accumulated rapidly in the *C. jejuni* population infecting a single individual, even in the absence of selective pressure, demonstrating mutation rates between 10 and 100 times faster than mutation rates in other parts of the genome (10). Due to their intrinsic genetic instability, variations observed in these genomic regions cannot be used to infer epidemiological relationships between isolates. Therefore, to ensure that the genetic signals used for tracing and source attribution of isolates are independent of genomic changes introduced by the host, such homopolymeric tracts should be excluded from genome comparisons in the context of public health investigations (7, 8, 15, 16).

APPLICATION OF WGS OF *C. JEJUNI* IN POINT-SOURCE OUTBREAK INVESTIGATIONS

Several studies have aimed to investigate the applicability of WGS analysis in identifying clonal *C. jejuni* isolates in point-source outbreaks, and the retrospective examination of foodborne and waterborne outbreaks of campylobacteriosis has gained significant research attention. This work has clearly shown that, regardless of the methodology used, the genomic diversity among epidemiologically linked isolates within an outbreak investigation is limited to a few changes, which mainly affect homopolymeric runs. Revez et al., analyzing a milk-borne outbreak, observed up to 3 SNVs and phase variations in 12 loci between the milk- and patient-derived isolates, using a PubMLST-hosted whole-genome multilocus sequence typing (wgMLST) schema featuring 1,738 loci (15). A similar number of genomic differences (3 SNVs) was observed between one human isolate and the assumed source isolate in a waterborne outbreak (7). Further studies have confirmed the initial findings of Revez and colleagues (7, 15). Zhang et al. identified up to seven allele differences between clonal isolates collected from three waterborne outbreaks, using a reference-based gene-by-gene approach (defined as *ad hoc* analysis) with the program Genome Profiler (https://sourceforge.net/projects/genomeprofiler) (17). Lahti and colleagues, investigating an Australian outbreak associated with chicken paté by using a reference-based core-genome MLST analysis implemented in SeqSphere+ (Ridom GmbH), with a total of 1,271 shared loci, described a maximum of 1 allele difference between the clinical isolates (18). Furthermore, four isolates associated with the Walkerton outbreak in Canada differed from one another by a total of 15 SNVs, corresponding to an average of ~4 allele differences, using a 732-core-gene schema (19). In all of those studies, epidemiologically linked strains were clearly separated from nonrelated control strains. Due to variations in the numbers of genes analyzed and the overall genetic diversity of the studied strains, the allelic differences between control strains and outbreak-

associated strains varied between 10 to 15, as observed in the Australian and the Canadian studies (18, 19), and 454, as described in the Finnish milk outbreak (15). In real-life settings, however, the possibility that the variation observed among isolates within an outbreak might represent the genomic variation present in the initial inoculum, due to mixed-strain coinfection, cannot be excluded. The circulation of more than one strain might complicate the correct clustering of cases during an investigation and thus should be taken into consideration during specification of the case definition for the outbreak. For example, the number of observed genomic differences between the assumed source isolate and a second human outbreak isolate analyzed by Revez et al. exceeded what might be expected to accumulate over the course of the outbreak, being similar (69 SNVs) to what was observed between the water isolate and a chicken strain obtained 12 years later (7). In addition, an Australian study that investigated a foodborne outbreak by extrapolating SNVs using Snippy 3.0 (https://github.com/tseemann/snippy) reported the circulation of at least two different strains among patients, one showing SNVs in the range of 3 to 8 and the second up to 30 (20).

APPLICATION OF WGS IN SURVEILLANCE OF *C. JEJUNI* INFECTIONS AND DETECTION OF DIFFUSE OUTBREAKS

In the absence of comprehensive and systematic surveillance, the detection of diffuse outbreaks is difficult, as these cases are masked against the background of sporadic cases. Earlier studies verified the existence of diffuse outbreaks of campylobacteriosis by demonstrating the presence of temporal, spatial, and genotypic clusters among apparently sporadic cases. Such clusters were several times more common than point-source outbreaks (21–23). High-resolution genotyping can improve the detection of clusters with diffuse epidemiological signals. Since WGS offers the highest possible discriminatory power, the integration of genome sequencing in surveillance would be expected to facilitate the identification of possible case clusters, which should allow for the implementation of more effective intervention strategies targeted at the prevention and control of cases of campylobacteriosis. Some studies have already examined the use of genomics in *Campylobacter* surveillance, and all applied a wgMLST methodology (24). Cody and colleagues performed the first real-time genomic epidemiological investigation of *C. jejuni* strains collected through a surveillance program, using a hierarchical wgMLST approach (25). In the first all-against-all analysis, BIGSdb Genome Comparator (26), implemented in PubMLST (http://pubmlst.org), was used to extract the 1,026 loci shared among the 379 investigated isolates obtained from the Oxfordshire monitoring program from a total of 1,643 loci listed in the Gundogdu schema (http://pubmlst.org/campylobacter) (25). Based on the level of genomic diversity observed among isolates collected from a single patient, the cutoff value for investigating the presence of possible clusters was set to 20 allele differences among the 1,026 shared loci (25). A follow-up wgMLST analysis, again using the Gundogdu schema, was used to define diversity within the clusters identified. With this approach, the authors were able to identify temporally associated clusters of strains from cases demonstrating no apparent epidemiological links but showing similar levels of genetic diversity, as observed for same-patient isolates (differences of 3 to 14 loci among 1,478 to 1,586 shared loci) (25). As observed in point-source outbreaks (7, 15, 17), the loci that recurrently showed variations between same-patient isolates and epidemiologically linked isolates were homopolymeric tracts in contingency genes. Similar results were reported by Kovanen et al., who also used a hierarchical wgMLST approach to investigate the genomic relationships of apparently sporadic cases collected during a seasonal peak in Finland (8). The authors clustered the isolates based on the 7-locus MLST scheme and subsequently performed an *ad hoc* wgMLST analysis, using BIGSdb (26) to extract allele information for all 1,738 loci for each ST group. This approach allowed the authors to identify subclusters of isolates (which varied in only a few loci) obtained from cases spread over a large geographical area, despite lacking an apparent epidemiological link (8). In their subsequent study, Kovanen and colleagues attempted to identify the possible sources of these predicted diffuse outbreaks, again applying a

hierarchical wgMLST approach for each ST group and using the reference-based gene-by-gene method implemented in the Genome Profiler software (16, 17). Thereafter, the authors manually screened for allele differences using a 5-SNV cutoff value, while excluding indels in homopolymeric tracts, to define clusters of genetically indistinguishable isolates (16). When accounting for temporal clustering, the authors were able to link up to 24% of the human cases to specific chicken slaughter batches (16). To identify the possible existence of a diffuse outbreak, Fernandes et al. applied the method introduced by Cody and colleagues (25) to compare *C. jejuni* isolates collected through a surveillance program with a reference population of nonrelated isolates from the same lineage (ST-21) (27). The authors showed that 20 of the 23 apparently sporadic cases were indeed part of a cluster, with less than 8 (mean of 4) allele differences across 1,577 shared loci. Simultaneously, the authors demonstrated that the genetic distance between the cluster isolates and the reference population was at least 20 alleles (27). The results from those studies support the idea that putative diffuse outbreaks have a clear impact on the epidemiology of campylobacteriosis and that, by integrating WGS in the surveillance of *C. jejuni* infections, we can potentially distinguish between clustered and sporadic cases. The common belief is that the gene-by-gene method is more suitable for this task than an SNV-calling approach, since the gene-by-gene approach is efficient, easy to automate, computationally less intensive, and less likely to be affected by multi-SNV import resulting from allelic replacement due to homologous gene transfer (24). Although a consensus regarding the analytical (experimental and bioinformatic) pipeline to use in the gene-by-gene approach and the interpretation criteria are still under evaluation, it seems clear that hierarchical wgMLST methodology (8, 16, 25), in which a first-level genomic clustering based on a core set of loci is followed by an *ad hoc* analysis with an increased number of loci for added discriminatory power, will be the basic approach for these types of investigations.

APPLICATION OF WGS IN *C. JEJUNI* SOURCE ATTRIBUTION AND ANALYSIS OF HOST ADAPTATION

Identification of the most frequent transmission routes and foodborne sources of campylobacteriosis is of utmost importance for prioritizing food safety interventions and setting public health goals. *C. jejuni* has complex epidemiology, and transmission can occur in numerous ways, including contaminated food, water, and raw milk and direct animal and environmental contact. To overcome this problem, methods to quantify the relationships between human patient data and possible infection reservoirs, transmission routes, and risk factors have been developed (28). Approaches based on genotypic methods compare the proportions of *C. jejuni* subtypes for different sources and reservoirs with those for isolates collected from human patients. These methods rely on different subtypes having different abilities to colonize different hosts or ecological restrictions preventing an equal subtype distribution between reservoirs, which is often referred to as host adaptation (29). Adaptation to a specific host might result in the selection of reservoir-associated traits, such as the presence of specific genes or clusters of genes (29). Although source attribution models using WGS data as input have yet to be applied to campylobacteriosis cases, phylogenic and population structure analyses of *C. jejuni* isolates collected from different studies have been used to trace the sources of sporadic cases, as described above (16). Furthermore, major efforts have been devoted to detecting and elucidating the mechanisms of host adaptation in *C. jejuni* with the use of WGS, especially identifying host-specific traits (29, 30). The discovery of such loci gained or lost makes these genetic elements possible targets for source tracking; the genetic make-up of a human isolate could theoretically help determine the source of that isolate. We are still far from that possibility, however, as only a few host-associated loci have been found in *C. jejuni*. Furthermore, the high prevalence and wide distribution of so-called "generalist lineages" complicate source attribution, due to their lacking genetic responses to host colonization (31). For instance, Dearlove and colleagues found that ST-21 CC and ST-45 CC undergo rapid

host switching, which appears to erode the development of host signals; therefore, WGS appears no more informative than 7-gene MLST in defining the source of infections for generalist lineages (31).

CONCLUSION

A significant challenge in *Campylobacter* research has been the lack of the systematic surveillance required to monitor emerging trends in the epidemiology of campylobacteriosis. In large part, this is due to the significant resources required to deploy subtyping methods for routine analysis of strains, particularly given the limited success of these approaches in epidemiological investigations. In this context, the various studies highlighted in this review suggest the benefits of integrating WGS in the routine surveillance of *C. jejuni* infections. Furthermore, international standardization is needed, due to the global and regional nature of campylobacteriosis, and WGS allows full comparability of molecular data for strains collected from different areas, as facilitated by public databases. However, several key issues will need to be addressed before the full potential of these approaches can be realized. A fundamental hurdle to the implementation of WGS is the limited knowledge regarding the genetic diversity within *C. jejuni* populations. Undeniably, the successful development of criteria for case cluster definition will require species-specific context based on a detailed understanding of the *C. jejuni* population structure, including genetic variations within and between lineages. This, in turn, will require a critical mass of WGS data from isolates with high-quality metadata from surveillance efforts worldwide, including data on isolates from both point-source and diffuse outbreaks. Moreover, studies to determine the optimal resolution depth of WGS data interpretation, to balance the power to resolve isolates from the main *C. jejuni* lineages while also allowing the detection of clusters of epidemiologically linked strains from outbreaks and source-to-patient transmission events, will be required. The latter might include creating and agreeing on a suitable nomenclature system for tracking and communicating *C. jejuni* trends worldwide. Finally, the development of approaches for the standardization of data analysis tools is warranted, since this would facilitate comparisons of isolates at the national and international levels, simplifying global sharing of data. Regardless of the existing hurdles, current evidence suggests that routine analysis of *C. jejuni* isolates using WGS is likely to provide fundamental insights regarding various aspects of the biology and epidemiology of this important pathogen, underscoring the benefits of implementing WGS in the service of public health.

ACKNOWLEDGMENTS

We thank João André Carriço (University of Lisbon, Lisbon, Portugal) and Alejandra Culebro (University of Helsinki, Helsinki, Finland) for their suggestions for improving the manuscript.

A.-K.L. has been supported by the INNUENDO project, cofunded by the European Food Safety Authority (grant agreement GP/EFSA/AFSCO/2015/01/CT2).

The conclusions, findings, and opinions expressed in this review article reflect only the view of the authors and not the official position of the European Food Safety Authority.

REFERENCES

1. Carrillo CD, Kruczkiewicz P, Mutschall S, Tudor A, Clark C, Taboada EN. 2012. A framework for assessing the concordance of molecular typing methods and the true strain phylogeny of *Campylobacter jejuni* and *C. coli* using draft genome sequence data. Front Cell Infect Microbiol 2:57.
2. Köser CU, Ellington MJ, Cartwright EJP, Gillespie SH, Brown NM, Farrington M, Holden MTG, Dougan G, Bentley SD, Parkhill J, Peacock SJ. 2012. Routine use of microbial whole genome sequencing in diagnostic and public health microbiology. PLoS Pathog 8:e1002824. https://doi.org/10.1371/journal.ppat.1002824.
3. Inns T, Ashton PM, Herrera-Leon S, Lighthill J, Foulkes S, Jombart T, Rehman Y, Fox A, Dallman T, Pinna ED, Browning L, Coia JE, Edeghere O, Vivancos R. 2017. Prospective use of whole genome sequencing (WGS) detected a multi-country outbreak of *Salmonella* Enteritidis. Epidemiol Infect 145:289–298. https://doi.org/10.1017/S0950268816001941.
4. 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.
5. Biggs PJ, Fearnhead P, Hotter G, Mohan V, Collins-Emerson J, Kwan E, Besser TE, Cookson A, Carter PE, French NP. 2011. Whole-genome comparison of two *Campylobacter jejuni* isolates of the same sequence type reveals multiple loci of different ancestral lineage. PLoS One 6:e27121. https://doi.org/10.1371/journal.pone.0027121.

6. Llarena AK, Zhang J, Vehkala M, Välimäki N, Hakkinen M, Hänninen ML, Roasto M, Mäesaar M, Taboada EN, Barker D, Garfolo G, Cammà C, Di Giannatale ED, Corander J, Rossi M. 2016. Monomorphic genotypes within a generalist lineage of *Campylobacter jejuni* show signs of global dispersion. Microb Genomics 10:2.
7. Revez J, Llarena AK, Schott T, Kuusi M, Hakkinen M, Kivistö R, Hänninen ML, Rossi M. 2014. Genome analysis of *Campylobacter jejuni* strains isolated from a waterborne outbreak. BMC Genomics 15:768. https://doi.org/10.1186/1471-2164-15-768.
8. Kovanen S, Kivistö R, Rossi M, Schott T, Kärkkäinen UM, Tuuminen T, Uksila J, Rautelin H, Hänninen ML. 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.
9. Wu Z, Periaswamy B, Sahin O, Yaeger M, Plummer P, Zhai W, Shen Z, Dai L, Chen SL, Zhang Q. 2016. Point mutations in the major outer membrane protein drive hypervirulence of a rapidly expanding clone of *Campylobacter jejuni*. Proc Natl Acad Sci U S A 113:10690–10695. https://doi.org/10.1073/pnas.1605869113.
10. Bayliss CD, Bidmos FA, Anjum A, Manchev VT, Richards RL, Grossier J-P, Wooldridge KG, Ketley JM, Barrow PA, Jones MA, Tretyakov MV. 2012. 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 Res 40:5876–5889. https://doi.org/10.1093/nar/gks246.
11. Revez J, Schott T, Llarena AK, Rossi M, Hänninen ML. 2013. Genetic heterogeneity of *Campylobacter jejuni* NCTC 11168 upon human infection. Infect Genet Evol 16:305–309. https://doi.org/10.1016/j.meegid.2013.03.009.
12. Thomas DK, Lone AG, Selinger LB, Taboada EN, Uwiera RRE, Abbott DW, Inglis GD. 2014. Comparative variation within the genome of *Campylobacter jejuni* NCTC 11168 in human and murine hosts. PLoS One 9:e88229. https://doi.org/10.1371/journal.pone.0088229.
13. Jerome JP, Bell JA, Plovanich-Jones AE, Barrick JE, Brown CT, Mansfield LS. 2011. Standing genetic variation in contingency loci drives the rapid adaptation of *Campylobacter jejuni* to a novel host. PLoS One 6:e16399. https://doi.org/10.1371/journal.pone.0016399.
14. Kim JS, Artymovich KA, Hall DF, Smith EJ, Fulton R, Bell J, Dybas L, Mansfield LS, Tempelman R, Wilson DL, Linz JE. 2012. Passage of *Campylobacter jejuni* through the chicken reservoir or mice promotes phase variation in contingency genes *Cj0045* and *Cj0170* that strongly associates with colonization and disease in a mouse model. Microbiology 158:1304–1316. https://doi.org/10.1099/mic.0.057158-0.
15. 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 52:2782–2786. https://doi.org/10.1128/JCM.00931-14.
16. Kovanen S, Kivistö R, Llarena AK, Zhang J, Kärkkäinen UM, Tuuminen T, Uksila J, Hakkinen M, Rossi M, Hänninen ML. 2016. Tracing isolates from domestic human *Campylobacter jejuni* infections to chicken slaughter batches and swimming water using whole-genome multilocus sequence typing. Int J Food Microbiol 226:53–60. https://doi.org/10.1016/j.ijfoodmicro.2016.03.009.
17. Zhang J, Halkilahti J, Hänninen ML, Rossi M. 2015. Refinement of whole-genome multilocus sequence typing analysis by addressing gene paralogy. J Clin Microbiol 53:1765–1767. https://doi.org/10.1128/JCM.00051-15.
18. Lahti E, Löfdahl M, Ågren J, Hansson I, Olsson Engvall E. 2017. Confirmation of a campylobacteriosis outbreak associated with chicken liver pâté using PFGE and WGS. Zoonoses Public Health 64:14–20. https://doi.org/10.1111/zph.12272.
19. Clark CG, Berry C, Walker M, Petkau A, Barker DOR, Guan C, Reimer A, Taboada EN. 2016. Genomic insights from whole genome sequencing of four clonal outbreak *Campylobacter jejuni* assessed within the global *C. jejuni* population. BMC Genomics 17:990. https://doi.org/10.1186/s12864-016-3340-8.
20. 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.
21. Jonsson ME, Heier BT, Norström M, Hofshagen M. 2010. Analysis of simultaneous space-time clusters of *Campylobacter* spp. in humans and in broiler flocks using a multiple dataset approach. Int J Health Geogr 9:48. https://doi.org/10.1186/1476-072X-9-48.
22. Strachan NJC, Forbes KJ. 2014. Extensive spatial and temporal clustering of *Campylobacter* infections evident in high-resolution genotypes, p 253–263. *In* Sheppard SK, Meric G (ed), *Campylobacter* ecology and evolution, 1st ed. Caister Academic Press, Norfolk, UK.
23. Castrodale LJ, Provo GM, Xavier CM, McLaughlin JB. 2016. Calling all Campy: how routine investigation and molecular characterization impacts the understanding of campylobacteriosis epidemiology: Alaska, United States, 2004–2013. Epidemiol Infect 144:265–267. https://doi.org/10.1017/S0950268815001466.
24. Maiden MCJ, Jansen van Rensburg MJ, Bray JE, Earle SG, Ford SA, Jolley KA, McCarthy ND. 2013. MLST revisited: the gene-by-gene approach to bacterial genomics. Nat Rev Microbiol 11:728–736. https://doi.org/10.1038/nrmicro3093.
25. Cody AJ, McCarthy ND, van Rensburg MJ, Isinkaye T, Bentley SD, Parkhill J, Dingle KE, Bowler ICJW, Jolley KA, Maiden MCJ. 2013. Real-time genomic epidemiological evaluation of human *Campylobacter* isolates by use of whole-genome multilocus sequence typing. J Clin Microbiol 51:2526–2534. https://doi.org/10.1128/JCM.00066-13.
26. Jolley KA, Maiden MC. 2010. BIGSdb: scalable analysis of bacterial genome variation at the population level. BMC Bioinformatics 11:595. https://doi.org/10.1186/1471-2105-11-595.
27. Fernandes AM, Balasegaram S, Willis C, Wimalarathna HML, Maiden MC, McCarthy ND. 2015. Partial failure of milk pasteurization as a risk for the transmission of *Campylobacter* from cattle to humans. Clin Infect Dis 61:903–909. https://doi.org/10.1093/cid/civ431.
28. Mullner P, Spencer SEF, Wilson DJ, Jones G, Noble AD, Midwinter AC, Collins-Emerson JM, Carter P, Hathaway S, French NP. 2009. Assigning the source of human campylobacteriosis in New Zealand: a comparative genetic and epidemiological approach. Infect Genet Evol 9:1311–1319. https://doi.org/10.1016/j.meegid.2009.09.003.
29. Sheppard SK, Didelot X, Meric G, Torralbo A, Jolley KA, Kelly DJ, Bentley SD, Maiden MCJ, Parkhill J, Falush D. 2013. Genome-wide association study identifies vitamin B_5 biosynthesis as a host specificity factor in *Campylobacter*. Proc Natl Acad Sci U S A 110:11923–11927. https://doi.org/10.1073/pnas.1305559110.
30. Morley L, McNally A, Paszkiewicz K, Corander J, Méric G, Sheppard SK, Blom J, Manning G. 2015. Gene loss and lineage-specific restriction-modification systems associated with niche differentiation in the *Campylobacter jejuni* sequence type 403 clonal complex. Appl Environ Microbiol 81:3641–3647. https://doi.org/10.1128/AEM.00546-15.
31. 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.