Journal of Microbiological Methods 95 (2013) 24–31

Contents lists available at ScienceDirect

Journal of Microbiological Methods

journal homepage: www.elsevier.com/locate/jmicmeth

Current methods for molecular typing of *Campylobacter* species

Eduardo N. Taboada [a], Clifford G. Clark [b], Emma L. Sproston [c], Catherine D. Carrillo [d,*]

[a] Laboratory for Foodborne Zoonoses, Public Health Agency of Canada, PO Box 640, Township Rd. 9-1, Lethbridge, AB T1J 3Z4, Canada
[b] National Microbiology Laboratory, Public Health Agency of Canada, 1015 Arlington St., Winnipeg, Manitoba R3E 3R2, Canada
[c] Bureau of Microbial Hazards, Health Canada, 251 Sir Frederick Banting Driveway, Ottawa, ON K1A 0C6, Canada
[d] Canadian Food Inspection Agency, 960 Carling Avenue, Ottawa, ON K1A 0C6, Canada

ARTICLE INFO

Article history:
Received 1 May 2013
Received in revised form 1 July 2013
Accepted 2 July 2013
Available online 17 July 2013

Keywords:
Campylobacter
Molecular typing
flaA short variable region sequence typing
Multi-locus sequence typing
Comparative genomic fingerprinting
Whole genome sequencing

ABSTRACT

Campylobacter remains one of the most common bacterial causes of gastroenteritis worldwide. Tracking sources of this organism is challenging due to the large numbers of human cases, and the prevalence of this organism throughout the environment due to growth in a wide range of animal species. Many molecular subtyping methods have been developed to characterize *Campylobacter* species, but only a few are commonly used in molecular epidemiology studies. This review examines the applicability of these methods, as well as the role that emerging whole genome sequencing technologies will play in tracking sources of *Campylobacter* spp. infection.

1. Introduction

Campylobacter spp. are among the most prevalent bacterial enteric pathogens in both developed and developing nations (ECDC, 2013; Kirkpatrick and Tribble, 2011). Most human disease has been attributed to *Campylobacter jejuni*, with *C. coli*, *C. lari*, and *C. upsaliensis* causing the majority of the remaining human cases of infection.

While incidence of campylobacteriosis is high, most illnesses occur sporadically, and cases are rarely (typically less than 1%) associated with outbreaks (ECDC, 2013; Silva et al., 2011). However, it is likely that outbreaks or smaller case clusters occur far more frequently than currently detected due to a number of confounding factors (Fussing et al., 2007; Miller et al., 2004; Pebody et al., 1997). To begin with, current and past methods used to detect clusters of cases may not detect any but the most obvious clusters from large point-source outbreaks of extremely limited duration and geographical extent. The use of improved molecular typing methods for determining the relatedness of isolates has already been shown to result in better detection of outbreaks (Fussing et al., 2007; Taylor et al., 2013). However, due to the large number of cases of campylobacteriosis, there may not be sufficient resources to perform molecular typing on the isolated organisms in order to identify common genotypes responsible for infections. In Europe, for example, characterization of most of the clinical *Campylobacter* isolates does not extend beyond genus level identification (ECDC, 2013). Finally, to further complicate matters, studies have found that several genotypes of *Campylobacter* can be isolated from a single clinical sample (Smith et al., 1999; Gilpin et al., 2012), and point source outbreaks may comprise more than one genotype of this pathogen (Hedberg et al., 2001).

Despite the fact that only a small proportion of cases appear to be outbreak-related, *Campylobacter* was one of the three most commonly reported causes of foodborne outbreaks in Europe in 2010 (ECDC, 2013). Outbreaks due to *Campylobacter* spp. are most commonly associated with dairy products, poultry products and untreated water (Greig and Ravel, 2009; Ravel et al., 2009; Taylor et al., 2013). However, common sources of infection cannot be determined from outbreaks alone, as routes of infection may differ between outbreaks and sporadic cases (Taylor et al., 2013).

Molecular typing methodologies have been instrumental in enhancing epidemiological investigations aimed at tracking sources of sporadic infections with *Campylobacter* spp. by providing information on the genetic subtypes in circulation. Poultry products are frequently contaminated with *Campylobacter* spp. and molecular typing data has linked *Campylobacter* spp. on these products to human infections (Batz et al., 2012; Müllner et al., 2009; Nadeau et al., 2002), particularly in urban areas (Müllner et al., 2010b). Though *Campylobacter* is frequently found in fresh water and in other food animals, such as cows and pigs,

Abbreviations: PFGE, pulsed field gel electrophoresis; flaA-SVR, *flaA* short variable region sequence typing; *flaA*-RFLP, *flaA* restriction fragment length polymorphism analysis; MLST, multi-locus sequence typing; eMLST, extended MLST; RAPD, random amplification of polymorphic DNA; WGS, whole genome sequencing; NGS, next-generation sequencing; MCGH, microarray comparative genomic hybridization; CGF, comparative genomic fingerprinting; SNPs, single nucleotide polymorphisms; HRM, high-resolution melting analysis.

* Corresponding author. Tel.: +1 613 759 1255.
E-mail addresses: Eduardo.Taboada@phac-aspc.gc.ca (E.N. Taboada), Clifford.Clark@phac-aspc.gc.ca (C.G. Clark), Emma.Sproston@hc-sc.gc.ca (E.L. Sproston), Catherine.Carrillo@inspection.gc.ca (C.D. Carrillo).

0167-7012/$ – see front matter.
http://dx.doi.org/10.1016/j.mimet.2013.07.007

and in wild animals, the role that these sources play in the epidemiology of campylobacteriosis, or how transmission to human hosts would commonly occur is not yet clear (Friesema et al., 2012; Taylor et al., 2013). A new generation of molecular typing methodologies will help to further characterize sources of *Campylobacter* infection.

The aim of this review is to provide an overview of current methods used for molecular typing of *Campylobacter jejuni* and *C. coli*. These methods will be discussed in the context of their contribution to *Campylobacter* epidemiology, both in the clinical setting, and in the food production continuum. The role of whole genome sequencing (WGS) in elucidating the population structure and epidemiology of *Campylobacter* spp. as well as in the development and the assessment of novel methods will also be examined.

2. Detecting clusters of human disease and sporadic cases

Routine surveillance, either to detect clusters that may be outbreaks or to quantify the burden of disease through detection of sporadic cases, which appear to be unlinked based on epidemiologic or molecular evidence, is a vital function of basic public health (Tauxe et al., 2010). Outbreak investigations can provide extremely valuable information about the organisms causing disease in human populations and data about the route the pathogen traveled from its origin in food animals, plants, or the environment (Batz et al., 2012). Such information can then be used to set up targeted investigations designed to evaluate the risk factors that promote survival of the organism at each stage and suggest methods to mitigate the risk of the organism in the food chain.

National or regional passive surveillance captures only a fraction of the available information about human infections, but can provide very valuable data on the incidence and prevalence of circulating subtypes as well as population dynamics represented as temporal or spatial changes. Sentinel-site active surveillance (e.g. C-Enternet in Canada (http://www.phac-aspc.gc.ca/c-enternet/) or FoodNet in the US (http://www.cdc.gov/foodnet/)), provides continuously collected information within a small geographic area, usually with an enriched data set representing epidemiological information and case information about each patient. These data sets should provide the most representative information about what subtypes are circulating within the boundaries of the sentinel site and can also provide insights into the association of particular bacterial subtypes with specific symptoms or disease presentations.

The inclusion of animals or environmental reservoirs in such surveillance schemes can extend the range of knowledge about how particular subtypes of a bacterial pathogen behave; direct comparisons can be made if the same subtyping methodology is used. Point-prevalence studies can frequently supplement this information by providing a snapshot of the organisms (types and subtypes) that can be detected within a specified geographic area within a specified span of time.

2.1. The PulseNet paradigm: a system for active molecular epidemiologic surveillance

The implementation of pulsed-field gel electrophoresis (PFGE) for bacterial typing (Barrett et al., 1994) and development of PulseNet in the U.S. (Swaminathan et al., 2001) demonstrated the value of "real-time" surveillance for the detection of public health events. Rather than the traditional "church supper" type of event that mainly identified larger point-source clusters or outbreaks of short duration, clusters and outbreaks that were distributed over larger areas have been identified (Tauxe et al., 2010). These outbreaks could not have otherwise been detected within the high background of cases that appear to be sporadic based on the lack of epidemiologic or molecular linkage.

Enhanced case cluster identification is made possible through the analysis of patterns of incidence of specific molecular subtypes among cases identified to public health authorities. Even when the numbers of each unique subtype within a single jurisdiction may not provoke an alert or a public health response, sharing data from many jurisdictions has been found useful for identification of more dispersed outbreaks (Swaminathan et al., 2001), especially when data are shared widely and in real-time. The use of molecular subtyping in conjunction with real-time data-sharing among large networks allows the identification and characterization of outbreaks that are geographically widespread, of longer duration, that occur due to extensive distribution of contaminated foods or other vehicles of infection.

A key role of PulseNet laboratories is to rapidly notify epidemiologists of newly-identified case-clusters (Boxrud et al., 2010). Laboratories frequently include data on the historical prevalence of the cluster pattern in their jurisdiction, obtained from comprehensive, continuously updated, and curated databases, in order to aid interpretation of the data. Epidemiological investigations are required to verify that nonrandom clusters of cases may indicate an outbreak. However, identification of clusters through the detection of a particular molecular type – or types – provides additional complexity for both laboratory and epidemiological investigations. The pathogen subtype then becomes part of the epidemiological case definition used to include other cases caused by organisms with the same type and exclude organisms of the same genus and species but with different molecular types (Tauxe, 2006). Early in the course of an outbreak this can represent an assumption that may not be valid. For example, multiple genotypes may be present in a source of infection (Gilpin et al., 2012). This may limit the number of cases meeting the case definition to an inadequate number of cases required for statistical analysis to identify likely vehicles of infection.

While the PulseNet model has been useful for a number of priority pathogens, currently clinical *Campylobacter* isolates are not routinely typed. This may be largely due to limitations in the resources that would be required for the analysis of the large number of clinical isolates routinely identified. Consequently, few outbreaks are detected through this system.

2.2. Subtyping rationale and strategies

Surrogate measures for characterizing a bacterium and for identifying distinct lineages and sub-lineages within a bacterial population are generally termed as "type" or "subtype". The term "molecular fingerprinting" has historically denoted methodologies aimed at differentiating organisms without a phylogenetic context. However, currently the two terms are used interchangeably.

The requirements for molecular typing/subtyping methods will vary based on the organism and the ultimate use of the information (van Belkum et al., 2007). For example, for surveillance and epidemiology of *Campylobacter* spp., cost and ease of use are primary considerations due to the very large numbers of isolates obtained from human cases (Frost, 2001). Molecular or phenotypic markers selected for epidemiological investigation should provide enough differentiation to separate, for instance, outbreak and non-outbreak isolates. Source tracking initiatives have also found it useful to group larger populations into clones or lineages that are specific for particular animals or environmental niches (Muellner et al., 2013; Sheppard et al., 2010).

For many pathogens, molecular typing is performed in the context of outbreak investigation and the rapidity of the response becomes the most important consideration to minimize the ultimate impact of an emerging outbreak (Boxrud et al., 2010; Hedberg et al., 2008). Methods used for molecular typing must therefore have a very rapid turnaround time from the time an organism is isolated to completion of typing. The method must have a high throughput especially when the number of cases is rapidly increasing, such as in the early stages of an outbreak. Implementation should be in front-line reference laboratories to reduce delays associated with shipping of strains and laboratory capacity issues (Swaminathan et al., 2001).

Molecular typing analyses have not frequently been applied to sporadic cases of campylobacteriosis, partly due to the resources that would