Infection, Genetics and Evolution 95 (2021) 105018

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

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

Campylobacter coli isolated in Brazil typed by core genome Multilocus Sequence Typing shows high genomic diversity in a global context

Carolina Nogueira Gomes [a], Dillon Oliver Reese Barker [b], Sheila da Silva Duque [c], Emily Victoria Che [b], Vasena Jayamanna [b], Eduardo Napoleon Taboada [b,*], Juliana Pfrimer Falcão [a,*]

[a] Departamento de Análises Clínicas, Toxicológicas e Bromatológicas, Faculdade de Ciências Farmacêuticas de Ribeirão Preto-Universidade de São Paulo, Ribeirão Preto, São Paulo, Brazil
[b] National Microbiology Laboratory, Public Health Agency of Canada, Winnipeg, Manitoba, Canada
[c] Fundação Oswaldo Cruz-FIOCRUZ, Instituto Oswaldo Cruz-IOC, Rio de Janeiro, Brazil

ARTICLE INFO

Keywords:
Campylobacter coli
MLST
cgMLST
Global context
Brazil

ABSTRACT

Campylobacter has been one of the most common causative agent of bacterial food-borne gastroenteritis in humans worldwide. However, in Brazil the campylobacteriosis has been a neglected disease and there is insufficient data to estimate the incidence of this pathogen in the country.

Aims: The current study aimed to determine the phylogenetic relationships among *Campylobacter coli* strains isolated in Brazil and to compare them with international *Campylobacter* isolates available in some public databases.

Methods and results: A total of 63*C. coli* strains isolated in Brazil were studied. The MLST analysis showed 18 different STs including three STs not yet described in the PubMLST database. The cgMLST allocated the Brazilian strains studied into five main clusters and each cluster comprised groups of strains with nearly identical cgMLST profiles and with significant genetic distance observed among the distinct clusters. The comparison of the Brazilian strains with 3401 isolates from different countries showed a wide distribution of these strains isolated in this country.

Conclusions: The results showed a high similarity among some strains studied and a wide distribution of the Brazilian strains when compared to isolates from different countries, which is an interesting data set since it showed a high genetic diversity of these strains from Brazil in a global context. This study contributed for a better genomic characterization of *C. coli* strains isolated in Brazil and provided important information about the diversity of this clinically-relevant pathogen.

1. Introduction

Campylobacter has been one of the most common causative agent of bacterial food-borne gastroenteritis in humans worldwide. In an analysis performed by the Foodborne Disease Burden Epidemiology Reference Group (FERG) of the World Health Organization (WHO), *Campylobacter* spp. were rated second among the most frequent causes of foodborne infections after norovirus, and responsible for 96 million cases globally each year (Kirk et al., 2015; WHO, 2015; Rokney et al., 2018).

According to the Centers for Disease Control and Prevention (CDC), *Campylobacter* causes an estimated 1.5 million illnesses each year in the United States (CDC, 2020). In 2018, the number of confirmed cases of human campylobacteriosis reported was 246,571 in the European Union, while the number of salmonellosis cases was 91,857 (EFSA and ECDC, 2019).

The clinical symptoms of campylobacteriosis include diarrhea, abdominal cramps, fever, vomiting and can lead to serious medical sequelae such as Guillain-Barré Syndrome, a debilitating and sometimes fatal paralysis (Nachamkin, 2002; Cameron et al., 2012; Xia et al., 2013). The most important potential transmission route of *Campylobacter* spp., specifically *Campylobacter coli* and *Campylobacter jejuni*, to humans is the consumption and handling of improperly prepared poultry meat (Skarp et al., 2016; EFSA and ECDC, 2019). However, the consumption of unpasteurized milk, raw red meat, fruits and vegetables,

* Corresponding authors.
E-mail addresses: eduardo.taboada@canada.ca (E.N. Taboada), jufalcao@fcfrp.usp.br (J.P. Falcão).

https://doi.org/10.1016/j.meegid.2021.105018
Received 1 June 2021; Received in revised form 21 July 2021; Accepted 25 July 2021
Available online 28 July 2021

activities related to recreational waters and contact with farm animals or pets have also been identified as possible causes of transmission (Man, 2011; Silva et al., 2011; Facciolà et al., 2017).

Although Brazil is currently the largest exporter of poultry meat in the world (ABPA, 2019), campylobacteriosis has been a neglected disease. Brazil has no surveillance program, which makes it difficult to estimate the number of cases and the incidence rate for gastroenteritis caused by *Campylobacter* (Coker et al., 2002; Epps et al., 2013; Gomes et al., 2016; Gomes et al., 2018; Silva et al., 2018; Gomes et al., 2019).

Molecular typing methodologies have been of utmost importance in epidemiological investigations for determining routes and sources of infections, identifying outbreaks and cross-contamination of health-related pathogens, identifying virulent strains and assessing the effectiveness of microbiological control measures (Nadeau et al., 2002;

Table 1
Sequence Read Archive (SRA) numbers, year, source, State of origin, Sequence Typing (ST), Clonal Complex (CC) and core genome Multilocus Sequence Typing (cgMLST) profile of the 63 *Campylobacter coli* strains studied.

SRA Number	Strains	Year	Source	State	ST	CC	cgMLST cluster
SRS7405071	CCAMP 771	1995	Sewage	RJ	8158	828	D
SRS7405106	CCAMP 773	1995	Sewage	RJ	8157	828	C
SRS7405094	CCAMP 820	1995	Monkey	RJ	1581	–	C
SRS7405083	CCAMP 821	1995	Monkey	RJ	8158	828	D
SRS7405072	CCAMP 840	1995	Sewage	RJ	860	828	E
SRS7405077	CCAMP 625	1996	Monkey	RJ	1166	828	B
SRS7405103	CCAMP 761	1996	Sewage	RJ	10,881	828	D
SRS7405073	CCAMP 764	1996	Sewage	RJ	8157	828	C
SRS7405133	CCAMP 765	1996	Sewage	RJ	8157	828	C
SRS7405134	CCAMP 767	1996	Sewage	RJ	8161	–	E
SRS7405076	CCAMP 768	1996	Monkey	RJ	1166	828	B
SRS7405074	CCAMP 774	1996	Sewage	RJ	8157	828	C
SRS7405117	CCAMP 787	1996	Sewage	RJ	8157	828	C
SRS7405128	CCAMP 819	1996	Sewage	RJ	1581	–	C
SRS7405075	CCAMP 825	1996	Monkey	RJ	8157	828	C
SRS7405080	CCAMP 775	1997	Sewage	RJ	829	828	E
SRS7405078	CCAMP 791	1997	Monkey	RJ	832	828	B
SRS7405079	CCAMP 818	1997	Sewage	RJ	1595	828	E
SRS7405084	CCAMP 490	1998	Human	RJ	830	828	E
SRS7405082	CCAMP 494	1998	Human	RJ	830	828	D
SRS7405081	CCAMP 495	1998	Human	RJ	8159	828	B
SRS7405085	CCAMP 841	1998	Monkey	RJ	829	828	E
SRS7405087	CCAMP 498	1999	Human	RJ	1166	828	B
SRS7405086	CCAMP 502	1999	Human	RJ	10,888	828	A
SRS7405088	CCAMP 975	1999	Monkey	RJ	860	828	E
SRS7405089	CCAMP 988	1999	Monkey	RJ	8157	828	C
SRS7405091	CCAMP 503	2000	Human	RJ	1166	828	B
SRS7405090	CCAMP 726	2000	Monkey	RJ	8157	828	C
SRS7405092	CCAMP 834	2000	Environmental Water	RJ	860	828	E
SRS7405093	CCAMP 595	2001	Human	RJ	8159	828	B
SRS7405095	CCAMP 667	2002	Monkey	RJ	1112	828	A
SRS7405096	Cc 01	2002	Human	SP	1581	–	C
SRS7405101	Cc 03	2003	Human	SP	830	828	E
SRS7405098	Cc 04	2003	Human	SP	1581	–	C
SRS7405102	Cc 05	2003	Human	SP	830	828	E
SRS7405097	Cc 10	2003	Human	SP	8158	828	D
SRS7405100	CCAMP 165	2003	Monkey	RJ	3858	828	D
SRS7405105	CCAMP 170	2003	Monkey	RJ	860	828	E
SRS7405099	CCAMP 182	2003	Monkey	RJ	3858	828	D
SRS7405104	CCAMP 73	2003	Monkey	RJ	8158	828	D
SRS7405111	CCAMP 394	2004	Monkey	RJ	825	828	E
SRS7405107	CCAMP 446	2004	Monkey	RJ	3858	828	E
SRS7405108	CCAMP 463	2004	Potable Water	MG	860	828	E
SRS7405110	CCAMP 464	2004	Potable Water	MG	2869	828	A
SRS7405113	CCAMP 465	2004	Environmental Water	MG	860	828	E
SRS7405115	CCAMP 466	2004	Environmental Water	MG	860	828	E
SRS7405114	CCAMP 467	2004	Environmental Water	MG	860	828	E
SRS7405109	CCAMP 469	2004	Potable Water	MG	2869	828	A
SRS7405112	CCAMP 769	2004	Sewage	MG	8157	828	C
SRS7405116	CCAMP 392	2007	Monkey	RJ	1096	828	A
SRS7405119	CCAMP 1000	2007	Monkey	RJ	4309	828	–
SRS7405118	CCAMP 1010	2007	Monkey	RJ	854	828	A
SRS7405120	CCAMP 1117	2009	Monkey	RJ	1096	828	A
SRS7405121	CCAMP 1062	2010	Chicken wing	RJ	10,885	–	E
SRS7405122	CCAMP 1067	2010	Chicken liver	RJ	1581	–	C
SRS7405123	CCAMP 1063	2010	Chicken gizzard	RJ	1166	828	B
SRS7405124	CCAMP 1064	2010	Chicken wing	RJ	830	828	E
SRS7405125	CCAMP 1066	2010	Chicken liver	RJ	830	828	E
SRS7405126	CCAMP 1068	2010	Chicken wing	RJ	830	828	E
SRS7405127	CCAMP 1071	2011	Chicken wing	RJ	1166	828	B
SRS7405130	CCAMP 1073	2011	Chicken wing	RJ	1166	828	B
SRS7405131	CCAMP 1074	2011	Chicken gizzard	RJ	1166	828	B
SRS7405129	CCAMP 1075	2011	Chicken liver	RJ	1166	828	B

Monkey: not diarrheal faeces; Human diarrheal faeces; MG: Minas Gerais; RJ: Rio de Janeiro; RS: Rio Grande do Sul; SP: São Paulo.

Müllner et al., 2009; Batz et al., 2012; Taboada et al., 2013). In this way, bacterial typing has contributed substantially for improving the effectiveness of surveillance systems and control strategies in public health (Tenover et al., 1997; Maccannell, 2013; Ranjbar et al., 2014).

Several molecular subtyping schemes have been developed and used successfully in characterization, epidemiological and phylogenetic studies of *C. coli* and *C. jejuni* isolates (Meinersmann et al., 1997; Dingle et al., 2001; Clark et al., 2005; Price et al., 2007; Sheppard et al., 2009; Ahmed et al., 2012; Gomes et al., 2016; Cody et al., 2017; Llarena et al., 2017). Of these methods, Multilocus Sequence Typing (MLST), which measures the allelic diversity of seven conserved housekeeping genes, has been a widely used method in part due to the facility of comparison of nucleotide sequence based typing among different laboratories worldwide (Dingle et al., 2001).

The recent development of high-throughput sequencing technologies for genome sequencing, specifically using next-generation sequencing (NGS) has greatly increased the amount of genetic information available for the characterization of bacterial isolates (Cody et al., 2017; Llarena et al., 2017; Medini et al., 2008).

Core genome Multilocus Sequence Typing (cgMLST), which extends MLST from seven conserved genes found in all members of the population to all such conserved genes within a genus, has been widely used to study the molecular epidemiological characteristics of several pathogens owing to the higher resolution compared to MLST (de Been et al., 2015; Janowicz et al., 2018; Guglielmini et al., 2019; Fida et al., 2020; Lan et al., 2020). However, for *Campylobacter* there is a paucity of studies using this methodology (Clark et al., 2016; Cody et al., 2017; Marotta et al., 2020). It is important to highlight that in Brazil there is a lack of data on this pathogen of worldwide importance, which makes the characterization of *Campylobacter* strains isolated in Brazil extremely important.

The current study analyzed *C. coli* strains isolated from human, animal, food and environmental sources isolated during a 16-year period in Brazil. These strains were typed by MLST and cgMLST and compared with 3401 international *Campylobacter* isolates that had been deposited in Bacterial Isolate Genome Sequence Database (BIGSdb), Sequence Read Archive (SRA) and National Center for Biotechnology Information (NCBI) sequence databases. We aimed to achieve a better understanding of the genotypic diversity of these strains.

2. Materials and methods

2.1. Bacterial strains

A total of 63*C. coli* strains were studied. Those strains were isolated from human faeces (12 strains), animals (21 strains), the environment (20 strains) and food (10 strains) from Rio de Janeiro, Sao Paulo and Minas Gerais States in the Southeast region of Brazil between 1995 and 2011. These strains were selected from the collection of the *Campylobacter* Reference Laboratory of the Oswald Cruz Institute of Rio de Janeiro (FIOCRUZ) as described by Gomes et al. (2018). Table 1 lists the year, source, State of origin and accession numbers of the 63*C. coli* strains studied.

2.2. Genome sequencing

The DNA of each strain was extracted according to Campioni and Falcão (2014). Libraries were prepared using 0.05 ng/ μL of genomic DNA with the Nextera XT Sample Prep Kit (Illumina, San Diego, CA).

The genomes were sequenced using the NextSeq 500 desktop sequencer using the NextSeq 500/550 Mid Output Kit V2 (300 cycles) (Illumina, San Diego, CA) according to the manufacturer's instructions in the Public Health Agency of Canada, National Microbiology Laboratory, Winnipeg, Manitoba, Canada.

2.3. Genome assembly

The assemblies were generated from raw reads using Shovill 0.9.044, which uses SPAdes 3.1245 as the assembler (available at https://github.com/tseemann/shovill). The contigs for each isolate (draft genomes) were annotated using Rapid Prokaryotic Genome Annotation (PROKKA) (Seemann, 2014).

2.4. RefSeqMasher Pipeline

All the assemblies were submitted to RefSeqMasher Pipeline, available in the Integrated Rapid Infectious Disease Analysis platform (IRIDA) (Matthews et al., 2018). RefSeqMasher compares the Jaccard distances between MinHash k-mers of the query genome, and those of the subject genomes in RefSeq. This allows the identification of the RefSeq species most similar to the submitted query sequence (Ondov et al., 2016).

2.5. Multilocus Sequence Typing (MLST) and Core genome Multilocus Sequence Typing (cgMLST)

The MLST profile of the strains studied was obtained using the MLST prediction software developed by Seemann (available at https://github.com/tseemann/mlst), which identifies the allelic sequence of each of the seven housekeeping genes comprising the *Campylobacter* MLST schema in submitted assemblies and compares them with sequences deposited in the PubMLST database (pubmlst.org). Thus, allele numbers and allelic profiles were obtained, which were then used to identify the ST (Sequence Type) and CC (Clonal Complex) of each strain studied.

The assemblies were analyzed for genetic similarity based on a comparison using cgMLST. Initially the assemblies were submitted to the chewBBACA Pipeline (available at https://github.com/B-UMMI/chewBBACA) to define a highly conserved gene set that was considered the core genome for *C. coli* and the alleles for each core gene were identified in the 63*C. coli* strains studied. The phylogenetic analysis of these *C. coli* strains was performed using a cgMLST scheme comprising 646 genes (Table S2). The result of this analysis was subjected to clustering using the program GrapeTree (Zhou et al., 2018). The resulting phylogenetic tree was subsequently visualized using the Interactive Tree of Life (iTOL) v5 program (available at https://itol.embl.de) annotated to indicate the number of allele differences.

2.6. Genomes for comparison

The genetic diversity of the 63*C. coli* strains isolated in Brazil was compared against an international cohort of 3401 sequenced *Campylobacter* strains isolated from different countries between 1992 and 2019 (Table S1) using cgMLST analysis as described above. Initially the whole genome sequences of 2400 *Campylobacter* isolates from Europe, 367*C. coli* isolated from unidentified countries, 606*C. coli* isolated from North America, 14*C. coli* isolated from Oceania, 11*C. coli* isolated from Asia and 3*C. coli* isolated from Africa, were downloaded. These genomes were available in BIGSdb (https://pubmlst.org/organisms/campylobacter-jejunicoli) and SRA; NCBI (https://www.ncbi.nlm.nih.gov/sra) databases (Table S1).

All assemblies were analyzed using the chewBBACA Pipeline, which was used to define a cgMLST scheme for *C. coli* and to identify alleles in the submitted sequences. The result of this analysis was visualized using a Minimum Spanning Tree (MST) generated using GrapeTree version 1.5.0 (Zhou et al., 2018).

3. Results

3.1. Species confirmation and metadata

All the 63 sequences submitted to RefSeqMasher Pipeline were

confirmed to belong to the *Campylobacter coli* species (Table 1). The metadata of the 3401 sequenced *Campylobacter* strains isolated from different countries is available at Supplementary Table 1 (Table S1).

3.2. Multilocus Sequence Typing (MLST) and Core Genome Multilocus Sequence Typing (cgMLST)

The MLST analysis showed 18 different STs among the 63*C. coli* strains studied, including three STs (10,881, 10,885 and 10,888) not yet described in the database among the strains CCAMP 502, CCAMP 761 and CCAMP 1062 (Table 1). The ST 1166 was represented by nine *C. coli* strains isolated from humans, animals, and food, as well as ST 8157 that was represented by 9*C. coli* strains isolated from animals and the environment. The ST 860 and 830 were represented by eight *C. coli* strains isolates from animals and the environment, and seven *C. coli* strains isolated from humans and food. The STs 1581, 8158 and 3858 were represented by five *C. coli* strains isolated from humans, animals, food and the environment, four strains isolated from humans, animals and the environment and three strains isolated from animals, respectively. The STs 829, 1096, 2869 and 8159 were represented by two *C. coli* strains each one and the STs 825, 832, 854, 1112, 1595, 4309 and 8161 were represented by only one strain each. Fifteen out of the 18 STs found belong to the Clonal Complex (CC) 828, which is the main CC of *C. coli* species. The STs 1581, 8161 and 10,885 were not related to any Clonal Complex (Table 1).

The 63*C. coli* strains in this study were allocated into 5 main clusters denoted A, B, C, D and E by cgMLST (Fig. 1). Each cluster comprised groups of strains with nearly identical cgMLST profiles and with significant genetic distance observed among the clusters. Cluster A comprised seven strains and included a pair of strains with very high genetic similarity that were strains CCAMP464 and CCAMP 469. These two strains were isolated from the environment in 2004 in the State of Minas Gerais and differed in only one allele among the 646 core genome genes studied. Cluster B grouped 12 strains, including a main sub-cluster that included 11 strains with a high similarity among them and one additional strain (CCAP 791) which differed in 365 alleles from the remaining strains in the cluster. Cluster C comprised 14 strains that could be further divided into two sub-clusters of nine strains (CCAMP 726, CCAMP 774, CCAMP 988, CCAMP 764, CCAMP 765, CCAMP 769, CCAMP 787, CCAMP 825 and CCAMP 773) and five strains (CCAMP 820, CCAMP 819, CCAMP 1067, Cc 04 and Cc 01) respectively. Each sub-cluster shared a large number of alleles. Cluster D grouped nine

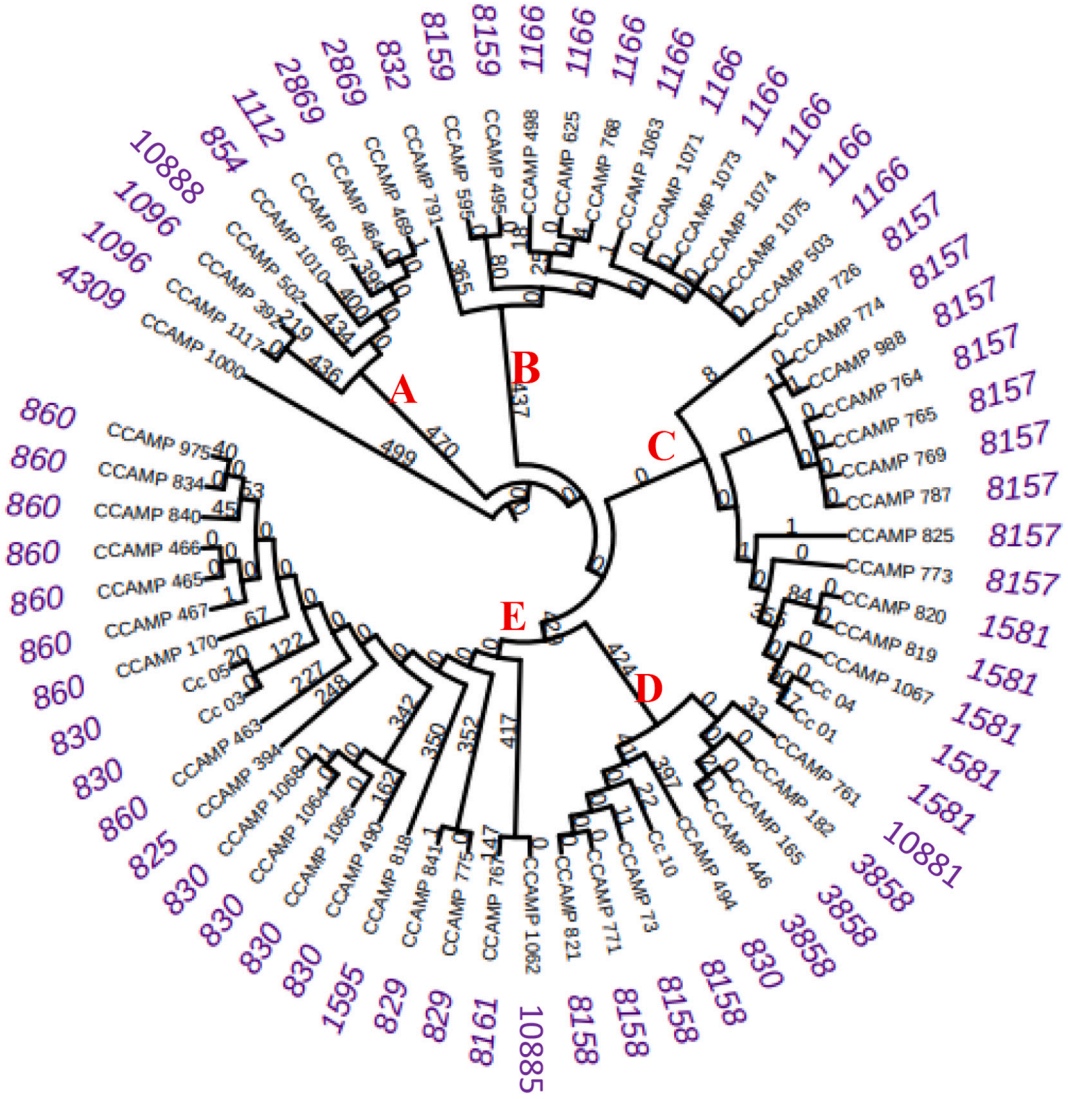

Fig. 1. Phylogenetic tree of the core genome Multilocus Sequence Typing (cgMLST), of the 63 *Campylobacter coli* strains studied, generated by the Interactive Tree Of Life (iTOL) v5 program. The STs of the strains studied are represented in purple and the number shown on the branches of the phylogenetic tree indicated the number of alleles that the strains differed among them.

strains, with four strains (CCAMP 761, CCAMP 182, CCAMP 165 and CCAMP 446) that shared high genetic similarity and differed from the remaining strains in the cluster by 411 alleles. In the same way, a high genetic similarity could be observed among the strains Cc 10, CCAMP 73, CCAMP 771 and CCAMP 821, which differed from the strain CCAMP 494 by 397 alleles. Cluster E grouped 20 strains of which seven (CCAMP 975, CCAMP 834, CCAMP 840, CCAMP 466, CCAMP 465, CCAMP 467, CCAMP 170) shared a large number of alleles. Finally, the strain CCAMP 1000 did not group with any other strain in the dataset and differed by a minimum of 499 alleles when compared to the other strains of this study (Fig. 1).

3.3. Genomes for comparison

The phylogenetic analysis by cgMLST of the 63*C. coli* strains isolated from different sources in Brazil compared to 3401 isolates from Europe, North America, Oceania, Asia, Africa and from countries not identified in the databases was performed based on cgMLST scheme of 484 genes described for the first time in this present study. Fig. 2 shows a wide distribution across the minimum spanning tree of the 63*C. coli* strains studied. However, some Brazilian strains clustered closely to one another, indicating a high genomic similarity among this specific strains (Fig. 2).

4. Discussion

The genome sequencing analyses has been successfully used in *Campylobacter* strains characterization, especially *C. jejuni* species, both in epidemiological and phylogenetic studies (Biggs et al., 2011; Revez et al., 2014; Clark et al., 2016; Cody et al., 2017; Joensen et al., 2018). However, it is important to note that there is a relative paucity of genomic studies on *C. coli* and that there are no published studies using NGS methodologies comparing *C. coli* strains circulating in Brazil among them and with other *C. coli* strains isolated in other countries (Cody et al., 2017; Liu et al., 2016; Ghatak et al., 2017; Cantero et al., 2018).

The molecular epidemiology of *C. coli* remains a challenge due to the nature of genome evolution of this microorganism and the extensive genomic and phenotypic diversity of this bacterial species. The genomic evolution of *C. coli* and *C. jejuni* has been strongly influenced by recombination, including frequent genomic rearrangements and genetic exchanges (De Boer et al., 2002; Ridley et al., 2008; Wilson et al., 2009); however, there is evidence of stability for some clones (Nielsen et al., 2001). Thus, some molecular subtyping methods has been developed for characterization and for epidemiological investigation of *C. jejuni* and *C. coli* isolates.

The MLST methodology is highly reproducible, and the results can be made available and shared in specific databases that make it possible to compare the genetic diversity of bacteria isolated from different countries (Maiden et al., 1998; Dingle et al., 2001; Schouls et al., 2003; Sheppard et al., 2009; Oporto et al., 2011; Llarena et al., 2015; Collado et al., 2018; Levican et al., 2019).

The MLST analysis showed 18 different STs among the 63*C. coli* strains studied, with three strains presenting STs not yet described in the database (Table 1). The MLST analysis allowed us to observe a high genetic similarity among some strains studied suggesting that the environment and food represent a possible source of contamination and/or transmission to humans and animals in Brazil and no predominance of specific STs was observed when the strains were compared based on their year of isolation. Moreover, 15 out of the 18 STs found belong to the Clonal Complex (CC) 828 (Table 1), which suggested the presence of a 3-clade *C. coli* population structure. In this genetic structure, horizontal gene transfer within each clade would be more common than among members of different clades resulting in a low diversity of CCs (Sheppard and Maiden, 2015).

In addition, the CC-828 is represented by approximately 70% of the *C. coli* isolates submitted to PubMLST / *campylobacter* database; the majority of the other isolates shared alleles with this Clonal Complex and are therefore phylogenetically related (Table 1). The ST 1166 and ST 8157 were commonly found among the strains of the present study and it is important to note that these STs have not been frequently

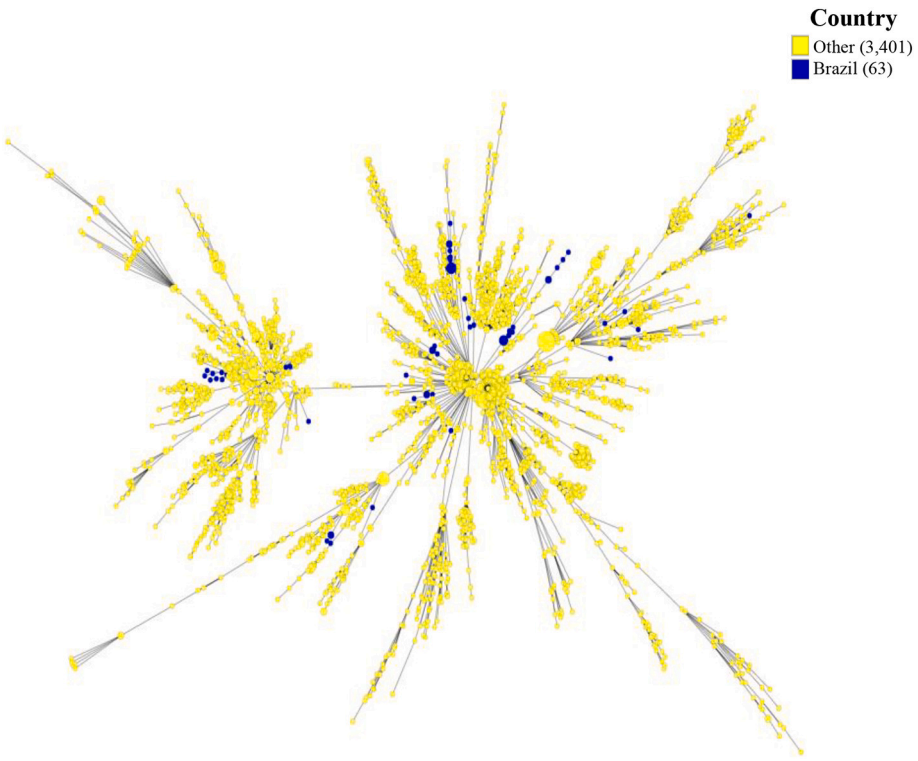

Fig. 2. Minimum Spanning Tree (MST) generated by GrapeTree version 1.5.0 program for the genomic phylogenetic analysis based on the core genome Multilocus Sequence Typing (cgMLST) of 3401 *Campylobacter* strains isolated from different countries (yellow circles) and 63 strains isolated in Brazil (blue circles).

reported in *C. coli* studies in other countries. In China, for example, the most commonly STs found have been ST 9227 (Zhang et al., 2020) while in the United States the predominant ST has been ST 1413 (Thakur et al., 2006). Thus, data reported by different researchers highlight the week clonality and high genomic diversity of this bacterial species.

A cgMLST scheme defines a comprehensive set of loci present in most members of a bacterial group, balancing very high resolution with comparability across the diversity of the group (Cody et al., 2017). In this study, the phylogenetic analysis of these *C. coli* strains was performed using a cgMLST scheme comprising 646 genes while the MLST measures the allelic diversity based on seven conserved housekeeping genes. The cgMLST analysis of the 63 sequenced *C. coli* strains clustered the strains into 5 main clusters which presented a high genomic diversity among them (Fig. 1). Nevertheless, a more detailed analysis of each of these clusters allowed us to observe a high genetic similarity among some of the strains studied.

The high similarity among some strains isolated in Brazil studied in this present work differed from previous studies in which the *C. coli* strains showed a high genetic diversity (Giacomelli et al., 2012; Zhang et al., 2014) and corroborate with the data from Pulse-field gel electrophoresis (PFGE) published by our research group (Gomes et al., 2016). It is important to mention that the whole genome sequencing provided more detailed and precise data than PFGE. Instead of only having the ability to compare bacterial genomes using 15–30 bands that appear in a PFGE pattern, the WGS have millions of bases to compare (CDC, 2021).

In addition, a cluster analysis based on the source of isolation allows us to observe that this high genetic similarity occurred among some strains from different sources suggesting that the environment and food have been a possible source of contamination and/or transmission to humans and animals, over 16 years in Brazil. No correlation based on the year of isolation was observed (Table 1).

Moreover, strains with the same ST were grouped in the same cluster except for strains belonging to ST 830, which could be found in clusters D and E (Fig. 1). It was also observed that some strains belonging to the same cluster differed in a large number of alleles that demonstrate a genetic diversity among strains belonging to the same ST and the high discriminatory power of the cgMLST methodology, which uses a much larger number of core or housekeeping genes when compared to the MLST.

The phylogenetic analysis by cgMLST of the 63*C. coli* strains isolated from different sources in Brazil compared to 3401 isolates from different countries showed a wide distribution of the 63*C. coli* strains, suggesting that strains isolated in Brazil have a high genetic diversity and represent many of the different lineages that have been found globally. This is interesting data since Brazil is the largest poultry meat exporter in the world and its main importers have been countries in Africa, America, Asia, Europe, Oceania and the Middle East of which strains were used for comparisons in this study (ABPA, 2019). Despite the high prevalence of *Campylobacter* in several countries worldwide, until now the research on this pathogen has not been included in the legislation that establishes the microbiological standards of food produced in Brazil (Agência Nacional de Vigilância Sanitária, 2020).

Efforts to reduce the *Campylobacter* infections have been directly related to knowledge about the genetics, the pathogen-host relationships, and possible contamination and transmission routes of this pathogen. Thus, the results obtained in this study are extremely important since they highlighted the need for a broader analysis of the circulating strains in Brazil.

In conclusion, the MLST and cgMLST analyses showed a high similarity among some *C. coli* strains studied suggesting that the environment and food have been a possible source of contamination and/or transmission to humans and animals in Brazil over the years studied. Besides that, a wide distribution of the 63*C. coli* strains isolated in Brazil when compared to isolates from different countries is an interesting data since it shows a high genetic diversity of Brazilian strains.

This study contributed for a better genomic characterization of *C. coli* strains isolated from various sources in Brazil and provided important information about the diversity of this clinically relevant pathogen. Moreover, this study is unique about the fact that genome sequencing data has never been published for *C. coli* strains isolated in Brazil, contributing to the limited amount of genomic data of this species in the literature.

Author contributions

Gomes, C. N- Wrote the manuscript and did the analysis.
Barker, D. O. R – Bioinformatics analysis and manuscript reviewer.
Duque, S. S. – Strain isolation and provider.
Che, E. V. – Sequencing of the samples.
Jayamanna, V. – Bioinformatics analysis.
Taboada, E. N – Bioinformatics analysis and was the supervisor at Public Health Agency of Canada.
Falcão, J. P. – Did the manuscript scientific revision, writing and editing and was the supervisor at University of São Paulo.

Declaration of competing interest

No conflict of interest is declared.

Acknowledgements

The study was supported by the Public Health Agency of Canada research under the supervision of Dr. Eduardo N. Taboada and by São Paulo Research Foundation - FAPESP (Proc. 2016/24716-3 and Proc. 2019/19338-8) under Dr. Juliana P. Falcão supervision. During the course of this work, Dr. Carolina N. Gomes was supported by São Paulo Research Foundation -FAPESP (Proc. 2015/23408-0 and Proc. 2018/26043-1). Falcão, J.P. received a productive fellowship from Council for Scientific and Technological Development (CNPq) grant 304399/2018-3.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.meegid.2021.105018.

References

Agência Nacional de Vigilância Sanitária, 2020. Alimentos: normas estabelecem padrões microbiológicos. In: Resolução da Diretoria Colegiada 331/2019 e Instrução Normativa 60/2019 sobre padrões microbiológicos para alimentos. Brasil. http://portal.anvisa.gov.br/noticias/-/asset_publisher/FXrpx9qY7FbU/content/alimentos-normas-estabelecem-padroes-microbiologicos/219201?inheritRedirect=false [accessed 26 May 2020].

Ahmed, M.U., Dunn, L., Ivanova, E.P., 2012. Evaluation of current molecular approaches for genotyping of *Campylobacter jejuni* strains. Foodborne Pathog. Dis. 9 (5), 375–385.

Associação Brasileira de Proteína Animal. Annual report (ABPA), 2019. http://abpa-br.org/mercados/ [accessed 17 April 2020].

Batz, M.B., Hoffmann, S., Morris Jr., J.G., 2012. Ranking the disease burden of 14 pathogens in food sources in the United States using attribution data from outbreak investigations and expert elicitation. J. Food Prot. 75, 1278–1291.

Biggs, P.J., Fearnhead, P., Hotter, G., Mohan, V., Collins-Emerson, J., Kwan, E., Besser, T. E., Cookson, A., Carter, P.E., French, N.P., 2011. Whole-genome comparison of two *Campylobacter jejuni* isolates of the same sequence type reveals multiple loci of different ancestral lineage. PLoS One 6 (11), e27121.

Cameron, A., Frirdich, E., Huynh, S., Parker, C.T., Gaynor, E.C., 2012. Hyperosmotic stress response of *Campylobacter jejuni*. J. Bacteriol. 194, 6116–6130.

Campioni, F., Falcão, J.P., 2014. Genotypic diversity and virulence markers of *Yersinia enterocolitica* biotype 1A strains isolated from clinical and non-clinical origins. APMIS 122, 215–222.

Cantero, G., Correa-Fiz, F., Ronco, T., Strube, M., Cerdà-Cuéllar, M., Pedersen, K., 2018. Characterization of *Campylobacter jejuni* and *Campylobacter coli* broiler isolates by whole-genome sequencing. Foodborne Pathog. Dis. 15 (3), 145–152.

Centers for Disease Control and Prevention (CDC), 2020. Foodborne Diseases Active Surveillance Network (FoodNet): FoodNet Surveillance Report for 2018. U.S. Department of Health and Human Services, Atlanta, GA.

Centers for Disease Control and Prevention (CDC), 2021. PulseNet Methods & Protocols.

Clark, C.G., Bryden, L., Cuff, W.R., Johnson, P.L., Jamieson, F., Ciebin, B., Wang, G., 2005. Use of the Oxford multilocus sequence typing protocol and sequencing of the flagellin short variable region to characterize isolates from a large outbreak of waterborne *Campylobacter* sp. strains in Walkerton, Ontario, Canada. J. Clin. Microbiol. 43 (5), 2080–2091.

Clark, C.G., Berry, C., Walker, M., Petkau, A., Barker, D.O.R., Guan, C., Reimer, A., Taboada, E.N., 2016. Genomic insights from whole genome sequencing of four clonal outbreak *Campylobacter jejuni* assessed within the global *C. jejuni* population. BMC Genomics 17 (990), 2080–2091.

Cody, A.J., Bray, J.E., Jolley, K.A., Mccarthy, N.D., Maiden, M.C.J., 2017. Core genome multilocus sequence typing scheme for stable, comparative analyses of *Campylobacter jejuni* and *C. coli* human disease isolates. J. Clin. Microbiol. 55, 2086–2097.

Coker, A.O., Isokpehi, R.D., Thomas, B.N., Amisu, K.O., Obi, C.L., 2002. Human campylobacteriosis in developing countries. Emerg. Infect. Dis. 8, 237–243.

Collado, L., Muñoz, N., Porte, L., Ochoa, S., Varela, C., Muñoz, I., 2018. Genetic diversity and clonal characteristics of ciprofloxacin-resistant *Campylobacter jejuni* isolated from Chilean patients with gastroenteritis. Infect. Genet. Evol. 58, 290–293.

de Been, M., Pinholt, M., Top, J., Bletz, S., Mellmann, A., van Schaik, W., Brouwer, E., Rogers, M., Kraat, Y., Bonten, C., Corander, J., Westh, H., Harmsen, D., Willems, R.J. L., 2015. Core genome multilocus sequence typing scheme for high-resolution typing of *Enterococcus faecium*. J. Clin. Microbiol. 53, 3788–3797.

De Boer, P., Wagenaar, J.A., Achterberg, R.P., Van Putten, J.P., Schouls, L.M., Duim, B., 2002. Generation of *Campylobacter jejuni* genetic diversity in vivo. Mol. Microbiol. 44, 351–359.

Dingle, K.E., Van Den Braak, N., Colles, F.M., Price, L.J., Woodward, D.L., Rodgers, F.G., Endtz, H.P., Van Belkum, A., Maiden, M.C., 2001. Sequence typing confirms that *Campylobacter jejuni* strains associated with Guillain-Barre and Miller-Fisher syndromes are of diverse genetic lineage, serotype, and flagella type. J. Clin. Microbiol. 39 (9), 3346–3349.

Epps, S.V.R., Harvey, R.B., Hume, M.E., Phillips, T.D., Anderson, R.C., Nisbet, D.J., 2013. Foodborne *Campylobacter*: infections, metabolism, pathogenesis and reservoirs. Int. J. Env. Res. Pub. He 10, 6292–6304.

European Food Safety Authority and European Centre for Disease Prevention and Control (EFSA and ECDC), 2019. The European Union One Health 2018 Zoonoses report. EFSA J. 17 (12), 5926.

Facciolà, A., Riso, R., Avventuroso, E., Visalli, G., Delia, S.A., Laganà, P., 2017. *Campylobacter*: from microbiology to prevention. J. Prev. Med. Hyg. 58, 79–92.

Fida, M., Cunningham, S.A., Murphy, M.P., Bonomo, R.A., Hujer, K.M., Hujer, A.M., Kreiswirth, B.N., Nicholas, C., Patricio, R.J., Nelson, H., Zinsmaster, N.M., Toraskar, N., Chang, W., Patel, R., Antibacterial Resistance Leadership Group, 2020. Core genome MLST and Resistome analysis of *Klebsiella pneumoniae* using a clinically-amenable workflow. Diagn. Microbiol. Infect. Dis. 114996.

Ghatak, S., He, Y., Reed, S. Jr T. S, Irwin, P., 2017. Whole genome sequencing and analysis of *Campylobacter coli* YH502 from retail chicken reveals a plasmid-borne type VI secretion system. Genomics Data 11, 128–131.

Giacomelli, M., Andrighetto, C., Rossi, F., Lombardi, A., Rizzotti, L., Martini, M., Piccirillo, A., 2012. Molecular characterization and genotypic antimicrobial resistance analysis of *Campylobacter jejuni* and *Campylobacter coli* isolated from broiler flocks in northern Italy. Avian Pathol. 41 (6), 579–588.

Gomes, C.N., Souza, A., Passaglia, J., Medeiros, M.C.I., Duque, S.S., Falcão, J.P., 2016. Genotyping of *Campylobacter coli* strains isolated in Brazil suggests possible contamination among environmental, human, animal and food sources. J. Med. Microbiol. 65, 80–90.

Gomes, C.N., Passaglia, J., Vilela, F.P., Pereira Da Silva, F.M.H.S., Duque, S.S., Falcão, J. P., 2018. High survival rates of *Campylobacter coli* under different stress conditions suggest that more rigorous food control measures might be needed in Brazil. Food Microbiol. 73, 327–333.

Gomes, C.N., Frazão, M.R., Passaglia, J., Duque, S.S., Medeiros, M.I.C., Falcão, J.P., 2019. Molecular epidemiology and resistance profile of *Campylobacter jejuni* and *Campylobacter coli* strains isolated from different sources in Brazil. Microb. Drug Resist. 26 (12), 1–10.

Guglielmini, J., Bourhy, P., Schiettekatte, O., Zinini, F., Brisse, S., Picardeau, M., 2019. Genus-wide *Leptospira* core genome multilocus sequence typing for strain taxonomy and global surveillance. PLoS Negl. Trop. Dis. 13 (4), e0007374.

Janowicz, A., De Massis, F., Ancora, M., Cammà, C., Patavino, C., Battisti, A., Prior, K., Harmsen, D., Scholz, H., Zilli, K., Sacchini, L., Di Giannatale, E., Garofolo, G., 2018. Core genome multilocus sequence typing and single nucleotide polymorphism analysis in the epidemiology of *Brucella melitensis* infections. J. Clin. Microbiol. 56, e00517–e00518.

Joensen, K.G., Kuhn, K.J., Müller, L., Björkman, J.T., Torpdahl, M., Engberg, J., Holt, H. M., Nielsen, H.L., Petersen, A.M., Ethelberg, S., Nielsen, E.M., 2018. Whole-genome sequencing of *Campylobacter jejuni* isolated from Danish routine human stool samples reveals surprising degree of clustering. Clin. Microbiol. Infect. 24, 5–8.

Kirk, M.D., Pires, S.M., Black, R.E., Caipo, M., Crump, J.A., Devleesschauwer, B., Döpfer, D., Fazil, A., Fischer-Walker, C.L., Hald, T., Hall, A.J., Keddy, K.H., Lake, R. J., Lanata, C.F., Torgerson, P.R., Havelaar, A.H., Ángulo, F.J., 2015. World Health Organization estimates of the global and regional disease burden of 22 foodborne bacterial, protozoal, and viral diseases, 2010: a data synthesis. PLoS Med. 12, e1001921.

Lan, P., Shi, Q., Zhang, P., Chen, Y., Yan, R., Hua, X., Jiang, Y., Zhou, J., Yu, Y., 2020. Core genome allelic profiles of clinical *Klebsiella pneumoniae* strains using a random forest algorithm based on multilocus sequence typing scheme for hypervirulence analysis. J. Inf. Secur. 221.

Levican, A., Ramos-Tapia, I., Briceño, I., Guerra, F., Mena, B., Varela, C., Porte, L., 2019. Genomic analysis of Chilean strains of *Campylobacter jejuni* from human faeces. Biomed. Res. Int. 1–12.

Liu, K.C., Jinneman, K.C., Neal-Mckinney, J., Wu, W.H., Rice, D.H., 2016. Genome sequencing and annotation of a *Campylobacter coli* strain isolated from milk with multidrug resistance. Genomics Data 8, 123–125.

Llarena, A.K., Huneau, A., Hakkinen, M., Hänninen, M.L., 2015. Predominant *Campylobacter jejuni* sequence types persist in Finnish chicken production. PLoS One 10 (2), 18.

Llarena, A.K., Taboada, E., Rossia, M., 2017. Whole-genome sequencing in epidemiology of *Campylobacter jejuni* infections. J. Clin. Microbiol. 55, 1269–1275.

Maccannell, D., 2013. Bacterial strain typing. Clin. Lab. Med. 33, 629–650.

Maiden, M.C.J., Bygraves, J.A., Feil, E., Morelli, G., Russell, J.E., Urwin, R., Zhang, Q., Zhou, J., Zurth, K., Caugant, D.A., Feavers, I.M., Achtman, M., Spratt, B.G., 1998. Multilocus sequence typing: A portable approach to the identification of clones within populations of pathogenic microorganisms. Proc. Natl. Acad. Sci. U. S. A. 95 (6), 3140–3145.

Man, S.M., 2011. The clinical importance of emerging *Campylobacter* species. Nat. Rev. Gastroenterol. Hepatol. 8, 669–685.

Marotta, F., Janowicz, A., Di Marcantonio, L., Ercole, C., Di Donato, G., Garofolo, G., Di Giannatale, E., 2020. Molecular characterization and antimicrobial susceptibility of *C. jejuni* isolates from Italian wild bird populations. Pathogens 9 (4), 304.

Matthews, T.C., Bristow, F.R., Griffiths, E.J., Petkau, A., Adam, J., Dooley, D., Kruczkiewicz, P., Curatcha, J., Cabral, J., Fornika, D., Winsor, G.L., Courtot, M., Bertelli, C., Roudgar, A., Feijao, P., Mabon, P., Enns, E., Thiessen, J., Keddy, A., Isaac-Renton, J., Gardy, J.L., Tang, P., The IRIDA Consortium João A Carriço, Chindelevitch, L., Chauve, C., Graham, M.R., McArthur, A.G., Taboada, E.N., Beiko, R.G., Brinkman, F.S.L., Hsiao, W.W.L., Domselaar, G.V., 2018. The integrated rapid infectious disease analysis (IRIDA) platform. bioRxiv 1–34.

Medini, D., Serruto, D., Parkhill, J., Relman, D.A., Donati, C., Moxon, R., Falkow, S., Rappuoli, R., 2008. Microbiology in the post-genomic era. Nat. Rev. Microbiol. 6, 419–430.

Meinersmann, R., Helsel, L.O., Fields, P.I., Hiett, K.L., 1997. Discrimination of *Campylobacter jejuni* isolates by *fla* gene sequencing. J. Clin. Microbiol. 35 (11), 2810–2814.

Müllner, P., Spencer, S.E.F., Wilson, D.J., Jones, G., Noble, A.D., Midwinter, A.C., Collins-Emerson, J.M., Carter, P., Hathaway, S., French, N.P., 2009. Assigning the source of human campylobacteriosis in New Zealand: a comparative genetic and epidemiological approach. Infect. Genet. Evol. 9, 1311–1319.

Nachamkin, I., 2002. Chronic effects of *Campylobacter* infection. Microbes Infect. 4, 399e403.

Nadeau, E., Messier, S., Quessy, S., 2002. Prevalence and comparison of genetic profiles of *Campylobacter* strains isolated from poultry and sporadic cases of campylobacteriosis in humans. J. Food Prot. 65, 73–78.

Nielsen, E.M., Engberg, J., Fussing, V., 2001. Genotypic and serotypic stability of *Campylobacter jejuni* strains during in vitro and in vivo passage. Int. J. Med. Microbiol. 291, 379–385.

Ondov, B.D., Treangen, T.J., Melsted, P., Mallonee, A.B., Bergman, N.H., Koren, S., Phillippy, A.M., 2016. Mash: fast genome and metagenome distance estimation using MinHash. Genome Biol. 17 (1).

Oporto, B., Juste, R.A., López-Portoléz, J.A., Hurtado, A., 2011. Genetic diversity among *Campylobacter jejuni* isolates from healthy livestock and their links to humans isolates in Spain. Zoonoses Public Health 58 (5), 365–375.

Price, E.P., Smith, H., Huygens, F., Giffard, P.M., 2007. High-resolution DNA melt curve analysis of the clustered, regularly interspaced short-palindromic-repeat locus of *Campylobacter jejuni*. Appl. Environ. Microbiol. 73 (10), 3431–3436.

Ranjbar, R., Karami, A., Farshad, S., Giammanco, G.M., Mammina, C., 2014. Typing methods used in the molecular epidemiology of microbial pathogens: a how-to guide. New Microbiol. 37, 1–15.

Revez, J., Llarena, A., Schott, T., Kuusi, M., Hakkinen, M., Kivistö, R., Hänninen, M., Rossi, M., 2014. Genome analysis of *Campylobacter jejuni* strains isolated from a waterborne outbreak. BMC Genomics 15, 1–8.

Ridley, A.M., Toszeghy, M.J., Cawthraw, S.A., Wassenaar, T.M., Newell, D.G., 2008. Genetic instability is associated with changes in the colonization potential of *Campylobacter jejuni* in the avian intestine. J. Appl. Microbiol. 105, 95–104.

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, 2432.

Schouls, L.M., Reulen, S., Duim, B., Wagenaar, J.A., Willems, R.J., Dingle, K.E., Colles, F. M., Van Embden, J.D., 2003. Comparative genotyping of *Campylobacter jejuni* by amplified fragment length polymorphism, multilocus sequence typing, and short repeat sequencing: strain diversity, host range, and recombination. J. Clin. Microbiol. 41 (1), 15–26.

Seemann, T., 2014. Prokka: rapid prokaryotic genome annotation. Bioinform. Appl. Note 30, 2068–2069.

Sheppard, S.K., Maiden, M.C.J., 2015. The evolution of *Campylobacter jejuni* and *Campylobacter coli*. Cold Spring Harb. Perspect. Biol. 7 (8), a018119.

Sheppard, S.K., Dallas, J.F., MacRae, M., McCarthy, N.D., Sproston, E.L., Gormley, F.J., Strachan, N.J.C., Ogden, I.D., Maiden, M.C.J., Forbes, K.J., 2009. *Campylobacter* genotypes from food animals, environmental sources and clinical disease in Scotland 2005/6. Int. J. Food Microbiol. 134, 96–103.

Silva, J., Leite, D., Fernandes, M., Mena, C., Gibbs, P.A., 2011. *Campylobacter* spp. as a foodborne pathogen: a review. Front. Microbiol. 2, 1–12.

Silva, W.C., Targino, B.N., Mendonça, R.S., Sant'Ana, A. S., Hungaro, H. M., 2018. *Campylobacter:* Na overview of cases, occurence in food, contamination sources, and antimicrobial resistance in Brazil. Food Rev. Int. 34, 364–389.

Skarp, C.P., Hänninen, M.L., Rautelin, H.I., 2016. *Campylobacteriosis*: the role of poultry meat. Clin. Microbiol. Infect. 22, 103–109.
Taboada, E.N., Clark, C.G., Sproston, E.L., Carrilo, C.D., 2013. Current methods for molecular typing of *Campylobacter* species. J. Microbiol. Methods 95 (1), 24–31.
Tenover, F.C., Arbeit, R.D., Goering, R.V., 1997. How to select and interpret molecular strain typing methods for epidemiological studies for bacterial infections: a review for healthcare epidemiologists. Infect. Control Hosp. Epidemiol. 5, 426–439.
Thakur, S., Morrow, W.E., Funk, J.A., Bahnson, P.B., Gebreyes, W.A., 2006. Molecular epidemiologic investigation of *Campylobacter coli* in swine production systems, using multilocus sequence typing. Appl. Environ. Microbiol. 72, 5666–5669.
Wilson, D.J., Gabriel, E., Leatherbarrow, A.J.H., Cheesbrough, J., Gee, S., Bolton, E., Fox, A., Hart, C.A., Diggle, P.J., Fearnhead, P., 2009. Rapid evolution and the importance of recombination to the gastroenteric pathogen *Campylobacter jejuni*. Mol. Biol. Evol. 26, 385–397.
World Health Organization (WHO), 2015. Foodborne Disease Burden Epidemiology Reference Group 2007–2015. World Health Organization, Amsterdam, The Netherland. Available from: https://www.who.int/activities/estimating-the-burden-of-foodborne-diseases.
Xia, Q., Muraoka, W.T., Shen, Z., Sahin, O., Wang, H., Wu, Z., Liu, P., Zhang, Q., 2013. Adaptive mechanisms of *Campylobacter jejuni* to erythromycin treatment. BMC 13, 1–16.
Zhang, M.J., Liu, X., Xu, X., Gu, Y., Tao, X., Yang, X., Yan, G., Zhang, J., 2014. Molecular subtyping and antimicrobial susceptibilities of *Campylobacter coli* isolates from diarrheal patients and food-producing animals in China. Foodborne Pathog. Dis. 11 (8), 610–619.
Zhang, P., Zhang, X., Liu, Y., Jiang, J., Shen, Z., Chen, Q., 2020. Multilocus sequence types and antimicrobial resistance of *Campylobacter jejuni* and *C. coli* isolates of human patients from Beijing, China, 2017–2018. Front. Microbiol. 11, 554784.
Zhou, Z., Alikhan, N., Sergeant, M.J., Luhmann, N., Vaz, C., Francisco, A.P., Carriço, J.A., Achtman, M., 2018. GrapeTree: visualization of core genomic relationships among 100,000 bacterial pathogens. Genome Res. 28, 1395–1404.