Core Genome Multilocus Sequence Typing for *Campylobacter coli*

Sampa Mukherjee[1], Lucas Harrison[1], Chih-Hao Hsu[1], Errol Strain[1], Qijing Zhang[2], Shaohua Zhao[1]

[1]- United States Food and Drug Administration, Center for Veterinary Medicine

[2] - Iowa State University, College of Veterinary Medicine

Abstract

Campylobacter species are among the leading foodborne bacterial agents of human diarrheal illness. The majority of campylobacteriosis has been attributed to *Campylobacter jejuni* (85% or more), followed by *Campylobacter coli* (5 to 10%). The distribution of *C. jejuni* and *C. coli* varies by host, suggesting the contribution to human infection may differ between sources. To address this, core genome multilocus sequence type with a 200-allele scheme (cgMLST200) was used to determine cgMLST type for 3,432 *C. coli* isolated from food animals (n=2,613), retail poultry meats (n=389), clinical settings (n=285), and environmental sources (n=145). Source attribution was determined by analyzing the core genome using the STRUCTURE program and a minimal multilocus distance methodology. Source attribution revealed the clinical isolates were divided into groups with the greatest probability of originating from environmental, swine, poultry and cattle sources. A higher proportion of the clinical *C. coli* population was attributed to retail poultry (23.7%) and cecal chicken (20.3%) sources compared with those from swine (8.7%) or wild birds (3.2%). Within the population of *C. coli* clinical isolates, 70% of the isolates that were attributed to retail poultry, dairy cattle, beef cattle and environmental waters came from two $cgMLST_{200}$ groups from each source. The most prevalent antibiotic resistance genes among all *C. coli* were *tetO* (65.6%), $bla_{OXA-193}$ (54.2%), *aph(3')-IIIa* (23.5%) and *aadE-Cc* (20.1%). Within $cgMLST_{200}$ groups, 17/17 $cgMLST_{200}$-435 and 89/92 $cgMLST_{200}$-707 isolates encoded for *aph(3')-VIIa* and 16/16 $cgMLST_{200}$-319 harbored *aph(2')-If* genes. Distribution of bla_{OXA} alleles showed 49/50 $cgMLST_{200}$-5 isolates contained $bla_{OXA-498}$ while $bla_{OXA-460}$ was present in 37/38 $cgMLST_{200}$-650 isolates. The $cgMLST_{200}$-514 group revealed both *ant(6)-Ia* and *sat4* resistance genes in 23/23 and 22/23 isolates, respectively. Isolation source also correlated with specific $cgMLST_{200}$ groups as 637/676 $cgMLST_{200}$-287 were isolated from cattle sources and 10 $cgMLST_{200}$ groups with population sizes 5 - 61 were isolated exclusively from retail poultry. These findings illustrate the contribution of known sources of C. coli to the human burden of campylobacteriosis and the use of cgMLST as an indicator of antimicrobial resistance in *C. coli*.

Introduction

Campylobacter species are the leading cause of human bacterial enteritis worldwide. While most cases of campylobacteriosis are caused by *C. jejuni*, *C. coli* populations are more commonly associated with resistance to multiple classes of antibiotics. *C. coli* populations can reside in food animal populations without any apparent effect on the host. This trait of asymptomatic host infection complicates efforts to identify the source of human disease. The prevalence and distribution of *C. jejuni* and *C. coli* varies by host, suggesting the contribution to human infection may differ between sources. Identifying the contribution of food animal sources of *C. coli* is important for the food industry and healthcare professionals to develop strategies to prevent the spread of antibiotic resistant organisms.

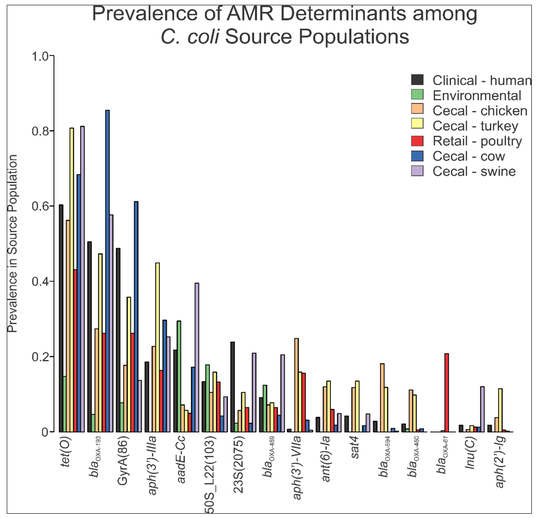
Our previous study evaluating *C. jejuni* populations using $cgMLST_{200}$ showed an excellent correlation with traditional MLST typing and had a greater utility for source attribution and correlation with AMR profiles. Since then, a minimal multilocus distance (MMD) analysis of *Campylobacter* genomes was demonstrated to be an effective tool for source attribution. The objective of this study is to use both cgMLST and MMD methods to characterize the relative contributions of *C. coli* sources to act as reservoirs for antibiotic resistance genes and the relative contribution of each source to human campylobacteriosis.

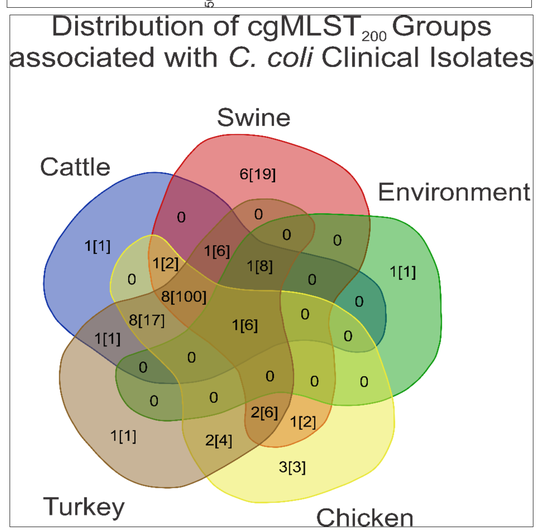
Disclaimer

The views expressed in this poster are those of the author(s) and may not reflect the official policy of the Department of Health and Human Services, the U.S. Food and Drug Administration, or the U.S. Government.

Materials and Methods

The sequences and metadata of 3,432 *C. coli* isolates originating from human, food animal, retail meat and environmental sources were collected and assigned core genome groups based on the 1,343 loci of the *Campylobacter* core genome. Strains were clustered into groups that allow for a 200-allele difference between strains, or $cgMLST_{200}$ groups. A minimum spanning tree of core genomes was generated using GrapeTree v1.5.0 and antimicrobial resistance genotypes were determined by the AMRFinder tool. Finally, source attribution of individual strains was performed using a minimal multilocus distance (MMD) algorithm developed by Pérez-Reche et al. Data visualizations were generated in R.





Results and Discussion

Analysis of the distribution of AMR determinants in *C. coli* populations among isolation source revealed that only one AMR gene, bla_{OXA-61} was preferentially isolated from a single source (**Figure 1**). Evaluating AMR determinants by $cgMLST_{200}$ groups shows a much higher correlation with specific $cgMLST_{200}$ groups and AMR determinants (**Tables 1 & 2**). Comparing $cgMLST_{200}$ groups to isolation source showed that the majority of groups that contain strains isolated from humans are able to infect multiple other hosts (**Figure 2**). When evaluating the proportion of isolates from each source that belong to human pathogenic $cgMLST_{200}$ groups, swine populations appear to have the least risk of contributing to human infections (**Table 3**). While the data from table 3 show that the majority of cattle isolates belong to $cgMLST_{200}$ groups capable of causing disease in humans, our MMD model shows that only 18.6% of clinical isolates were attributed to a cattle source (**Figure 3**). This discrepancy between the two methods may be due to processing intervention methods that limit the transmission of *C. coli* strains present in the caeca of food animals to the food products that humans are exposed to.

Table 1. Prevalence of AMR genes in $cgMLST_{200}$ groups

$cgMLST_{200}$ group	AMR gene	Prevalence
$cgMLST_{200}$-435	*aph(3)-VIIa*	17/17 (100%)
$cgMLST_{200}$-654	*aph(3)-VIIa*	89/92 (96.7%)
$cgMLST_{200}$-266	*aph(2)-If*	16/16 (100%)
$cgMLST_{200}$-5	$bla_{OXA-489}$	49/50 (98%)
$cgMLST_{200}$-597	$bla_{OXA-460}$	37/38 (97.4%)
$cgMLST_{200}$-461	*ant(6)-Ia*	23/23 (100%)
$cgMLST_{200}$-461	*sat4*	22/23 (95.7%)

Table 2. Prevalence of AMR genes in $cgMLST_{200}$ groups

$cgMLST_{200}$ Group	AMR Substitution	Prevalence
$cgMLST_{200}$-266	GyrAT86I	16/16 (100%)
$cgMLST_{200}$-84	GyrAT86I	14/15 (93.3%)
$cgMLST_{200}$-248	GyrAT86I	146/184 (79.3%)
$cgMLST_{200}$-221	GyrAT86I	49/65 (75.4%)
$cgMLST_{200}$-234	GyrAT86I	452/676 (66.9%)
$cgMLST_{200}$-5	L22A103V	50/50 (100%)
$cgMLST_{200}$-558	L22A103V	23/24 (95.8%)
$cgMLST_{200}$-597	L22A103V	30/38 (78.9%)
$cgMLST_{200}$-266	L23A2075G	16/16 (100%)

Table 3. Distribution of human-pathogenic $cgMLST_{200}$ groups by source

Source	# of isolates	Total # of $cgMLST_{200}$ groups	$cgMLST_{200}$ groups shared with Human	Source population within human-pathogenic $cgMLST_{200}$ groups
Swine	898	617	21	69 (7.7%)
Environment	129	83	3	26 (20.1%)
Cattle	964	117	22	845 (87.7%)
Chicken	849	233	26	223 (26.3%)
Turkey	307	123	25	103 (33.6%)

Figure 1 (Upper left): Prevalence of AMR determinants in *C. coli* populations among different isolation sources. Comparison of determinants across sources reveals that bla_{OXA-61} was the only resistance determinant that was selective for a single isolation source, retail poultry.

Figure 2 (Lower left): Distribution of $cgMLST_{200}$ groups that contain human pathogenic strains of *C. coli*. The majority of human isolates belong to $cgMLST_{200}$ groups containing isolates from a variety of sources. The first number listed in each source shows how many $cgMLST_{200}$ groups contain human-pathogenic *C. coli*. The second number, in brackets, is the number of human-pathogenic strains in the $cgMLST_{200}$ groups.

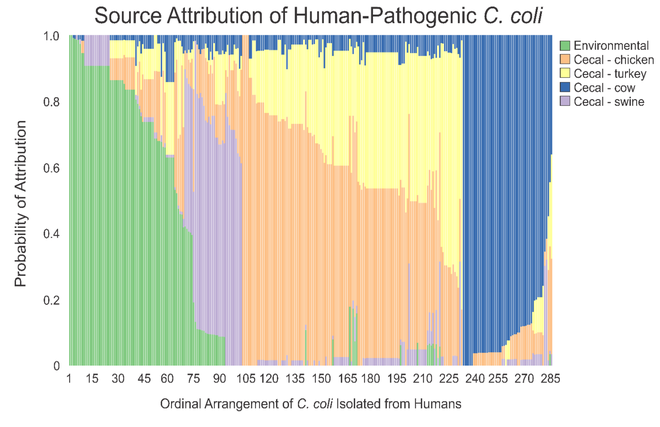


Figure 3. Minimal multilocus distance (MMD) source attribution of human-pathogenic *C. coli* to environmental and food animal cecal sources. Each bar represents a single *C. coli* isolate obtained from a human source and the color composition of the bar shows the likelihood of the strain as originating from the evaluated source. Only 4.2% of the clinical isolates could be attributed to a single source with 100% likelihood and 11.9% of the population could be attributed at >95% likelihood. In order, *C. coli* were attributed to chicken (36.5%), environmental (23.9), cattle (18.6%), swine (11.9%) and turkey (9.1%) sources. The strong co-occurrence of chicken and turkey attribution to the same clinical isolates indicates genomic similarity of the two poultry populations.

Conclusions

1. Source attribution analysis indicate that a higher proportion of *C. coli* clinical isolates are attributed to poultry and environmental sources than from cattle and swine.
2. Within the population of *C. coli* clinical isolates, 70% of the isolates that were attributed to retail poultry, dairy cattle, beef cattle and environmental waters came from two $cgMLST_{200}$ groups from each source. This indicates that some, but not all, $cgMLST_{200}$ groups contain strains that are adapted to survival in specific hosts.
3. With the exception of bla_{OXA-61} in retail poultry, antimicrobial resistance determinants were not preferentially identified in any of the isolation sources.
4. Using both $cgMLST_{200}$ and MMD analyses, we have demonstrated the utility of large-scale genomic analysis of *C. coli* for identifying risk factors of campylobacteriosis, including the source of antimicrobial resistant *C. coli* infections.

Acknowledgements

We would like to thank Glenn E. Tillman, Jovita Haro and Cesar Morales of the United Stated Department of Agriculture Food Safety and Inspection System for their expertise and contributions to this project.