Genomic Diversity of *Salmonella* Kentucky Isolated from Humans, Food, and Animal Sources in the United States

Heather Tate[1], Chih-Hao Hsu[1], Jing Han[2], Steven Foley[2], Jason Folster[3], Epiphanie Nyirabahizi[1], Shaohua Zhao[1]
[1] FDA CVM- Laurel MD, [2] FDA NCTR- Jefferson AK, [3] CDC NCEZID- Atlanta GA

Abstract

In the U.S., *Salmonella* serovar Kentucky is frequently isolated from domestic chickens and dairy cattle, but incidence of *S.* Kentucky in humans is low. We aimed to better describe the genetic diversity, antimicrobial resistance, and virulence determinants, of *S.* Kentucky isolates from humans, domestic food animals and meat products, and imported foods. We downloaded and analyzed the genomes of 773 *S.* Kentucky isolates from NCBI. Isolates were collected from humans, chicken, turkey, cattle, swine, and imported foods between 2002 and 2020. Sequence types (ST) were assigned to 752 isolates. Approximately 63% (54/86) of human isolates were ST198, 33% (29/86) were ST152, and 3 isolates were ST314. The majority (90.6%, 570/629) of animal and animal food isolates were ST152 or ST152-like (one allele difference), and 9.2% (58/629) were ST198. Isolates from imported food were mostly ST198 (60%, 22/37) and ST314 (29.7%, 11/37). ST152 isolates from dairy cattle appeared to have a lineage (Clade 2) distinct from ST152 isolates from chicken (Clade 4), and half of ST152 human isolates appeared to cluster within 2 other clades (Clades 1 and 3), largely distinct from Clades 2 and 4. Isolates in Clades 1, 2, and 3 had significantly less antimicrobial resistance determinants, compared to isolates in Clade 4. ST198 isolates clustered into two main lineages: one comprised mostly of isolates from humans and imported foods (Clade ST198.2) that contained triple quinolone resistance determining region (QRDR) mutations, conferring resistance to fluoroquinolones; and one that was largely absent of QRDR mutations but with greater source diversity (Clade ST198.1). Sublineages in Clade ST198.1 revealed source specific clusters and a small cluster containing three isolates with one QRDR mutation in *gyrA* (Clade ST198.1.1). Using a maximum parsimony tree, we found unique differences in the composition of fimbrial gene clusters and other virulence genes among the four STs. ST198 appears to be a major contributor to human infections in the United States but has low prevalence in domestic food animals and meats. The selective presence of certain virulence genes in some STs may provide clues to the apparent host specificity and pathogenicity of *S.* Kentucky lineages.

Introduction

Salmonella Kentucky (*S.* Kentucky) is a polyphyletic serovar, meaning it consists of multiple sequence types (ST) that do not share a common ancestor. Some research points to these STs having different ecologies, host characteristics, and geographic dispersions (1,2). In the U.S., not much is known about the ST distribution among human cases, however food sources have been studied. Most *S.* Kentucky isolated from domestic agricultural sources are ST152 (2). ST198 is also found in mostly dairy cattle sources (2). ST198 can be resistant to multiple antibiotics including ciprofloxacin, a first line therapy for treatment of complicated salmonellosis infection. In countries where ciprofloxacin- resistant (Cip^R) ST198 circulates, poultry is thought to be a major source of infection (1). Cip^R ST198 can be found in high levels among poultry isolates in countries where there is uncontrolled administration of fluoroquinolones in poultry production. Extensive use of fluoroquinolones may provide the selective pressure necessary for distribution and persistence of this strain. Although U.S. regulations prohibit the use of fluoroquinolones in poultry production, these strains may still reach consumers through other food sources, including imported foods. Travel has also been shown to be a risk factor for Cip^R ST198 infection (3). Here, we sought to examine the phylogenetic profiles of a large number of *S.* Kentucky isolates from ill persons in the U.S., domestic agricultural sources, and imported foods. Using these data, we can begin to understand the relatedness of *S.* Kentucky derived from various sources and make inferences about potential sources of foodborne *S.* Kentucky infection.

Materials and Methods

Bacterial strains and sequencing. As part of routine National Antimicrobial Resistance Monitoring System (NARMS) surveillance, *S.* Kentucky were isolated from retail meat, animal ceca and human sampling that occurred at various timepoints between 1996-2020. Imported foods were collected for various FDA sampling assignments occurring between 2005 and 2019. NARMS isolates were sequenced by short-read sequencing on the Illumina MiSeq using v2 or v3 reagent kits (Illumina, San Diego, CA, USA) with 2x300 bp paired-end reads. The libraries were prepared with Nextera XT kit by Illumina. Illumina sequences were submitted to the sequence read archive (SRA) in the National Center for Biotechnology Information (NCBI) under Bioprojects PRJNA292661, PRJNA292666, and PRJNA590944.

Data mining. Genomic sequences were downloaded from NCBI for all human (93), turkey ceca (24), retail ground turkey meat (5), dairy cattle ceca (42), beef cattle ceca (27) retail ground beef (8), sow ceca (2), market swine ceca (10) and retail pork (1) *S.* Kentucky isolates recovered during the timeframe of the study. We downloaded sequences of only 149 isolates from chicken ceca and 375 isolates from retail chicken meat samples because of the relatively large number of *S.* Kentucky collected from these sources during the timeframe of the study. Isolates were selected to ensure that at least 1 isolate from each site, month, and year were included. We also downloaded sequences of 37 *S.* Kentucky recovered from FDA imported product sampling. These included fish, spices, animal feed and pet food.

Bioinformatic analysis. The analysis of classical multilocus sequence typing (MLST) was performed based on the whole-genome sequencing (WGS) data. *S.* Kentucky MLST allelic profiles and sequences were downloaded from the PubMLST database (https://pubmlst.org/). A total of 6,901 profiles for 7 different loci were used for the MLST analysis. The SRST2 pipeline (Inouye et al. 2014) was used to determine the MLST type for our *S.* isolates. *Salmonella enterica* Kentucky isolates with the same MLST sequence type (ST) were included into a single-nucleotide polymorphism (SNP) analysis. The SNP analysis was done by CFSAN-SNP-Pipeline (http://snp-pipeline.readthedocs.io/en/latest/). VarScan (Koboldt et al. 2012) was used to detect SNPs. The plasmid sequences were excluded in the analysis since they evolved through horizontal gene transfer (HGT) which will misrepresent the evolution path of the strains. Phylogenetic tree was constructed using the GARLI pipeline (Maddison et al. 2020). Antimicrobial resistance genes were identified using the Resfinder database (Bortolaia et al. 2020) downloaded from (https://cge.cbs.dtu.dk/services/ResFinder/) in October 2020, containing 3,123 antimicrobial resistance genes. Resistance mutations were also determined using an in-house tool to determine any mutations in the *gyrA* gene that have been shown to cause resistance in *Salmonella*. An NCTR-developed *Salmonella* virulence factor database (4) was queried to identify putative virulence genes. Virulence genes were extracted from the NCTR-database, transformed to binary data and imported into BioNumerics for phylogenetic analysis using Dice Coefficients and UPGMA (Applied Maths, Austin, TX, USA).

References

1. Le Hello S, et al. Front Microbiol. 2013 Dec 18;4:395
2. Haley BJ, et al. PLoS One. 2016 Ocr 3;11(10):e0161225
3. Rickert-Hartmen, R and Folster, JP. Emerg Infect Dis. 2014 May; 20(5): 910-911
4. Aljahdali, NH, et al. Genes. 2020 Nov; 11(11):1307

Results and Discussion

Overall, 752 of 774 isolates sequenced fell among 4 predominate sequence types: ST152 (N=400), an ST152-like clade (similar to ST152 but with one allele difference; N=203), ST198 (N=134), and ST314 (N=15). ST152 isolates separated into 4 major clades (**Fig 1**) . Clades 1 and 2 contained only human isolates. Clade 2 contained isolates from cattle. Clade 4, the largest, contained isolates from varied source types, but predominately chicken sources. Only 5% of isolates in Clades 1, 2, and 3 had a functional resistance gene. However, 75% of isolates in Clade 4 had predicted resistance to at least one antibiotic (not shown).

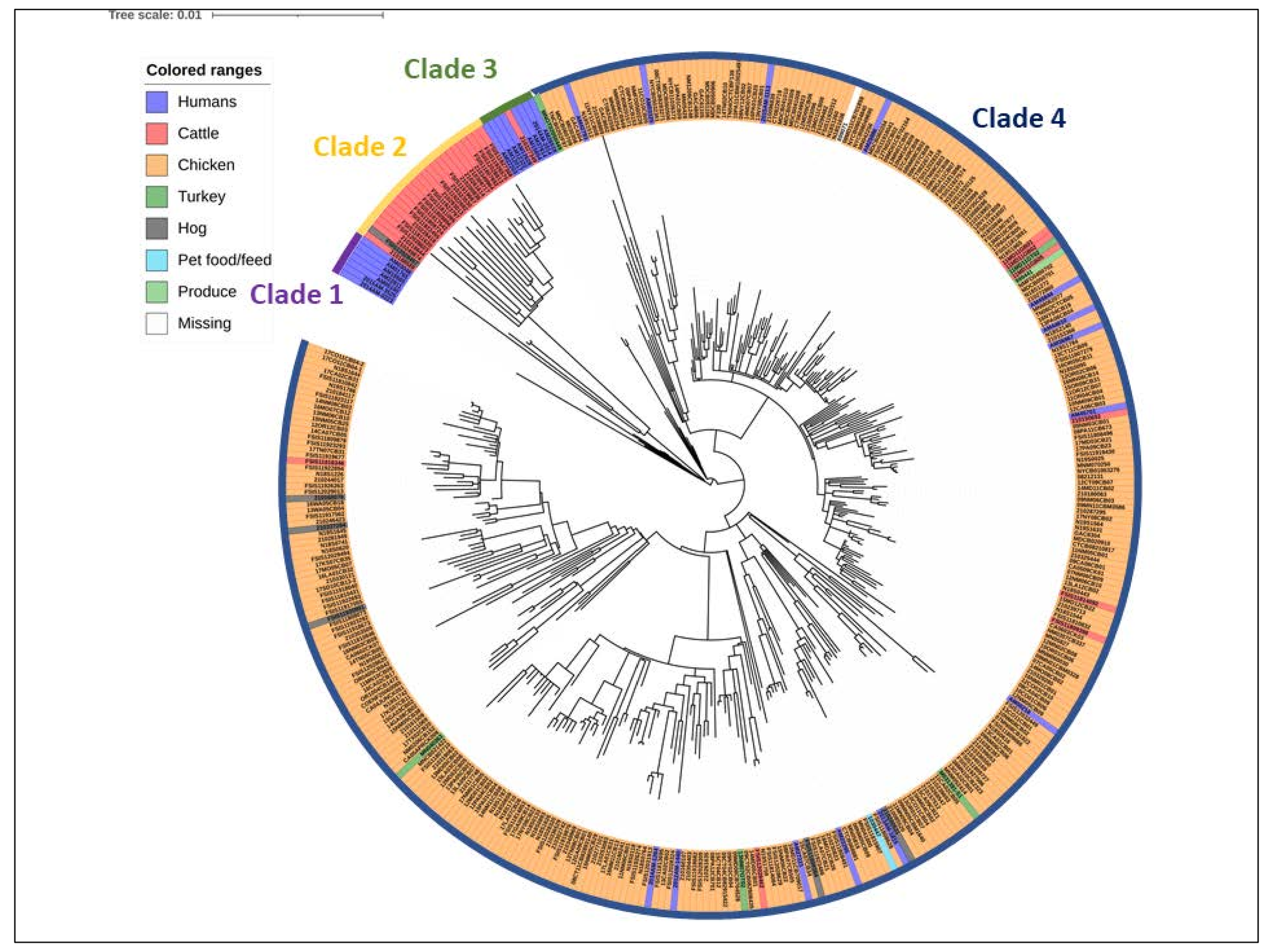


Figure 1. Maximum likelihood phylogeny of 400 ST152 isolates

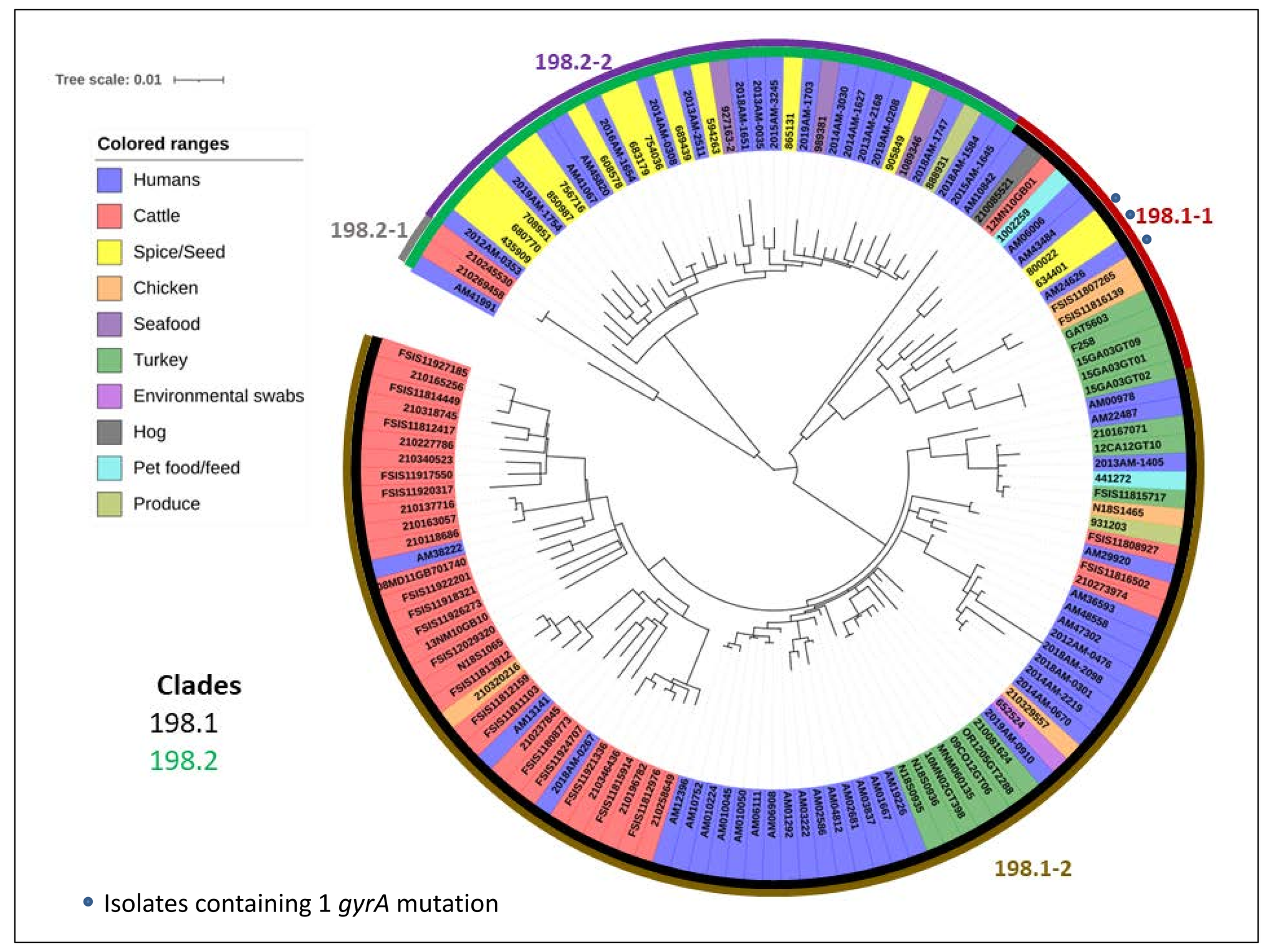


Figure 2. Maximum likelihood phylogeny of 134 ST198 isolates

ST198 isolates were divided into a mostly Cip^S clade (198.1) comprised predominately of isolates from humans and domestic agricultural sources and a Cip^R clade (198.2) comprised of isolates from humans and imported foods and containing triple QRDR mutations (**Fig 2**). Sublineages of these clades demonstrated source specificity. Additionally, we found that 198.1-1 contained three isolates with only one QRDR mutation. ST134 isolates appeared to have geographic interdependency, with one lineage containing isolates derived only from foods sampled for import product testing and the other containing isolates from the US (**Fig 3**).

Figure 3. SNP analysis of 15 ST314 isolates

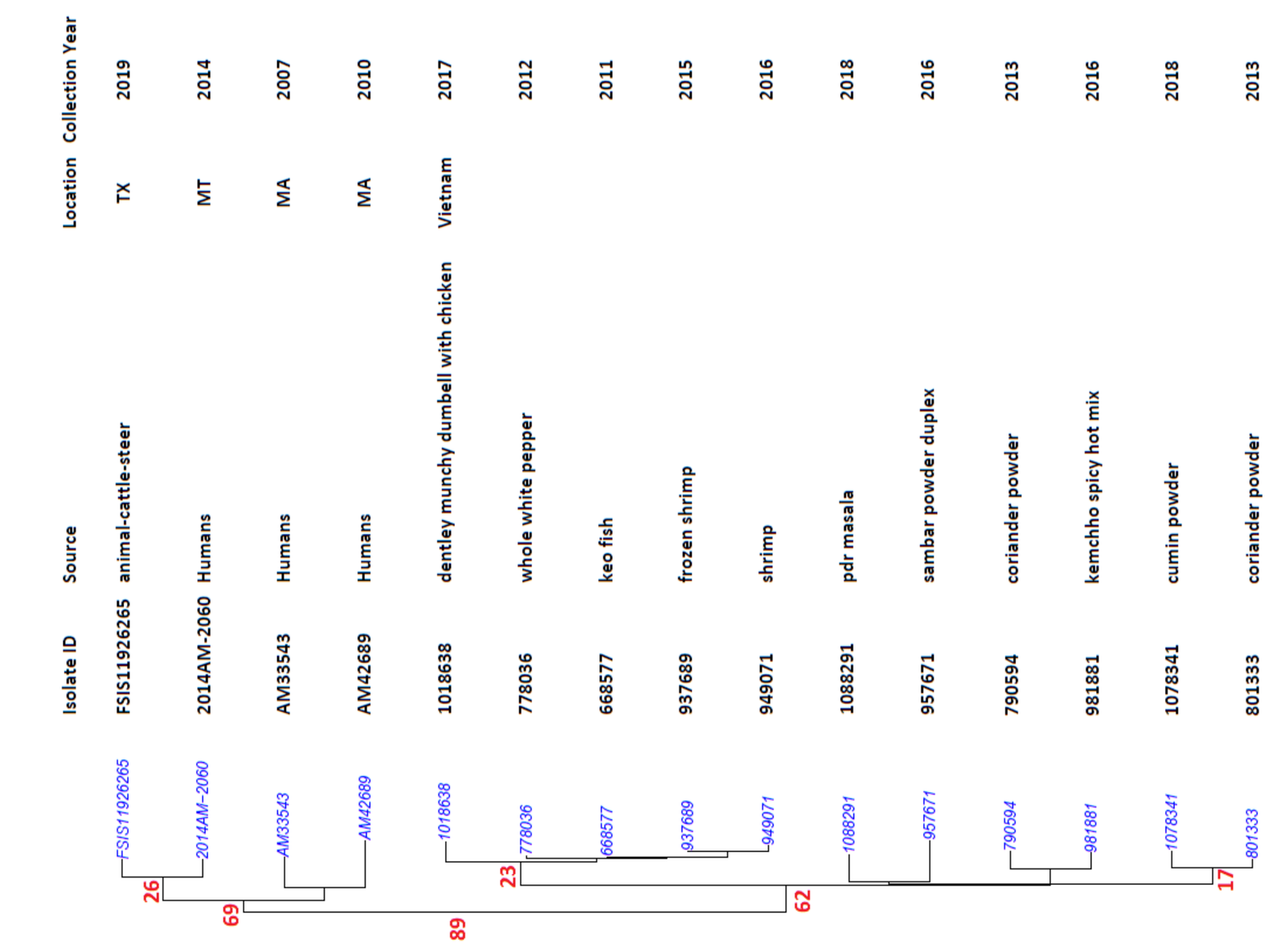
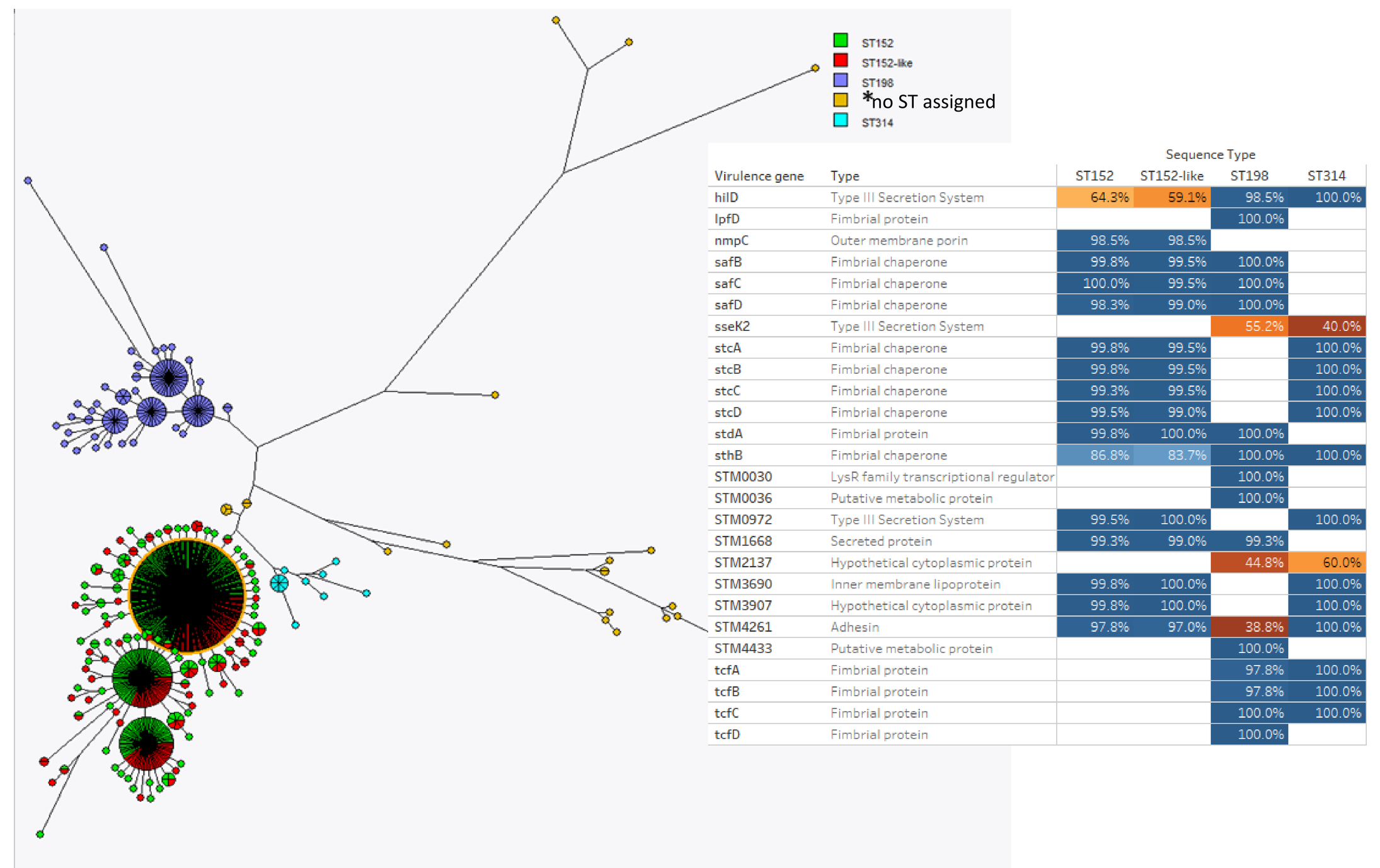


Figure 4. Maximum parsimony tree of all 774 *S.* Kentucky isolates and analysis of virulence genes



A maximum parsimony tree based on virulence gene profiles shows that STs are genomically different (**Fig 4**). These differences were based on the 26 virulence genes shown. For example, *lpfD* was present in all ST198, but were absent from all ST152/152-like and ST314 isolates.

Conclusion

- In our study, the majority (63%) of clinical infections of *S.* Kentucky in humans were ST198.
- Cip^R ST198 appears to be a major contributor to human infections in the U.S., but has low prevalence among *S.* Kentucky isolates from domestic food animals and meats.
- Cip^R ST198 infections may be linked to the consumption of imported (or foreign, if consumed while traveling) food product.
- 90% of ST152 isolates from humans clustered separately from domestic food animals and products, suggesting the majority of sporadic Cip^S ST152 *S.* Kentucky infections may be attributed to a source other than food products derived from domestic cattle, chicken, turkey, or pig.
- There are unique differences in the composition of fimbrial gene clusters and other virulence genes among the ST groups.
- The different composition of virulence genes among the ST groups and AMR genes within the ST groups may provide clues to the host specificity and pathogenicity of *S.* Kentucky isolates