Retrospective whole-genome sequencing analysis distinguished PFGE and drug resistance matched retail meat and clinical *Salmonella* isolates

Andrea B. Keefer[a], Lingzi Xiaoli[a], Nkuchia M. M'ikanatha[b], Kuan Yao[c], Maria Hoffmann[c], and Edward G. Dudley[a,#]

[a]Department of Food Science, The Pennsylvania State University, University Park, Pennsylvania, USA

[b]Pennsylvania Department of Health, Harrisburg, Pennsylvania, USA

[c]Center for Food Safety and Applied Nutrition (CFSAN), Food and Drug Administration (FDA), College Park, Maryland, USA

Running Title: WGS distinguished historic *Salmonella* isolates

[#]Address correspondence to Edward Dudley, egd100@psu.edu

Keywords: Whole-genome sequencing, retail meat, *Salmonella*, antimicrobial resistance

Abstract

Non-typhoidal *Salmonella* are a leading cause of outbreak and sporadic-associated foodborne illnesses in the U.S. These infections have been associated with a range of foods, including retail meats. Traditionally, pulsed-field gel electrophoresis (PFGE) and antibiotic susceptibility testing (AST) have been used to facilitate public health investigations of *Salmonella* infections. However, whole-genome sequencing (WGS) has emerged as an alternative tool that can be routinely implemented. To assess its potential in enhancing integrated surveillance in Pennsylvania, WGS was used to directly compare the genetic characteristics of 7 retail meat and 43 clinical historic *Salmonella* isolates, subdivided into three subsets based on PFGE and AST results, to retrospectively resolve their genetic relatedness and identify antimicrobial resistance (AMR) determinants. Single nucleotide polymorphism (SNP) analyses revealed the retail meat isolates within *S*. Heidelberg, *S*. Typhimurium var. O5- subset 1, and *S*. Typhimurium var. O5- subset 2 were separated from each primary PFGE pattern-matched clinical isolate by 6-12, 41-96, and 21-81 SNPs, respectively. Fifteen resistance genes were identified across all isolates, including *fosA7*, a gene only recently found in a limited number of *Salmonella* and a ≥ 95% phenotype to genotype correlation was observed for all tested antimicrobials. Moreover, AMR was primarily plasmid-mediated in *S*. Heidelberg and *S*. Typhimurium var. O5- subset 2; whereas, AMR was chromosomally-carried in *S*. Typhimurium var. O5- subset 1. Similar plasmids were identified in both the retail meat and clinical isolates. Collectively, these data highlight the utility of WGS in retrospective analyses and enhancing integrated surveillance of *Salmonella* from multiple sources.

Importance

Due to its enhanced resolution, whole-genome sequencing has emerged as a public health tool that can be utilized for pathogen monitoring, outbreak investigations, and surveillance for antimicrobial resistance. This study demonstrated that historical isolates that are indistinguishable by pulsed-field gel electrophoresis, a conventional genotyping method, and antibiotic susceptibility testing, could in fact be different strains, further highlighting the power of whole-genome sequencing. Moreover, we evaluated the role of whole-genome sequencing in integrated surveillance for drug-resistant *Salmonella* from retail meat and clinical sources in Pennsylvania and found a high correlation between antimicrobial resistance phenotype, as determined by antibiotic susceptibility testing, and genotype. Furthermore, the genomic context of each resistance gene was elucidated, which is critical to understanding how resistance is spreading within *Salmonella* in Pennsylvania. Taken together, these results demonstrate the utility and validity of whole-genome sequencing in characterizing human and food-derived *Salmonella*.

Introduction

Non-typhoidal *Salmonella enterica* subsp. *enterica* are the leading bacterial etiological agent of foodborne illness, hospitalization, and death in the U.S. (1). The Centers for Disease Control and Prevention (CDC) estimates that non-typhoidal *Salmonella* cause 1.2 million infections, 23,000 hospitalizations, and 450 deaths annually (2). Furthermore, compared to other bacterial pathogens, non-typhoidal *Salmonella* account for the majority of foodborne outbreaks that occur in the U.S., with associated food commodities including eggs, vegetables, fruits, and retail meats (3, 4). Two important serovars of *Salmonella* are *S*. Typhimurium, including its variant, *S*. Typhimurium var. O5-, and *S*. Heidelberg; these serovars are consistently ranked within the top ten most commonly isolated from humans and retail meats in the U.S. (5, 6).

Although *Salmonella* infections are typically self-limiting, antimicrobial treatment can be necessary in some cases (7); accordingly, drug-resistant non-typhoidal *Salmonella* are categorized by the CDC as a serious public health threat (2). Indeed, an estimated 100,000 drug resistant non-typhoidal *Salmonella* infections occur annually in the U.S. (2). Specifically in Pennsylvania, to contribute to the One Health approach outlined by the White House for combating antimicrobial resistance (AMR), the Pennsylvania Department of Health (PADOH) conducts integrated AMR surveillance in enteric bacteria, including *Salmonella*, isolated from clinical samples and retail meats, as part of the National Antimicrobial Resistance Monitoring System (NARMS) (8, 9).

Moreover, the current standard methods employed to conduct integrated surveillance and foodborne outbreak investigations are antibiotic susceptibility testing (AST) and pulsed-field gel electrophoresis (PFGE). Even though PFGE has been considered the gold standard molecular epidemiological tool for decades, it has numerous documented limitations (10), including the

93 inability to differentiate between clonal or low genetic diversity isolates, such as *S*. Heidelberg
94 (11). Similarly, despite its utility, AST has several shortcomings, including MIC breakpoint
95 inconsistencies (12), inability to efficiently test all known drugs, and only providing phenotype-
96 level resolution (13), which is not adequate, if AMR gene alleles and/or transmission
97 mechanisms need to be discerned.

98 Due to increases in affordability and ease of performance, whole-genome sequencing
99 (WGS) has emerged as an attractive tool that can be utilized for foodborne outbreak
100 investigations and pathogen-specific surveillance. Compared to conventional subtyping methods,
101 WGS yields increased discriminatory power, which results from its nucleotide-level resolution,
102 enabling the differentiation between clonal and closely related bacterial isolates (11, 14–17).
103 Accordingly, prior studies have assessed and established the utility of WGS in *Salmonella*
104 outbreak investigations, primarily by performing WGS retroactively on outbreak-associated
105 isolates from various serovars (11, 14–22). WGS is also effective at source-tracking (20), as it
106 provides the resolution needed to establish a genetic link between clinical and food isolates,
107 which traditionally, can be difficult to attain (17, 18).

108 In addition, WGS has significant potential for use as an AMR surveillance tool that can
109 be used to monitor and track resistance in human, animal, and food isolates. Indeed, NARMS has
110 recently incorporated WGS into its AMR monitoring efforts (23). WGS allows for the
111 elucidation of a bacterium's full resistome; this information can be used to associate resistance
112 phenotype with genotype and reveal possible transmission mechanisms, by characterizing the
113 genomic context of each AMR gene (13, 24, 25).

114 Due to the enhanced resolution conferred by WGS, there is now motivation to re-examine
115 historic collections of isolates to further resolve their relatedness and genetic AMR profiles.

116 Therefore, this study aimed to 1) use WGS to retrospectively resolve the genetic relatedness of a
117 historic collection of PFGE-matched and multi-drug resistant (MDR) retail meat and clinical *S.*
118 Heidelberg and *S.* Typhimurium var. O5- isolates, respectively and 2) to ascertain the genetic
119 AMR profile of each isolate in that collection, in an effort to assess the role of WGS in integrated
120 surveillance for drug-resistant *Salmonella* from clinical and retail meat sources in Pennsylvania.

121 **Materials and Methods**

122 ***Bacterial isolates***

123 All bacterial isolates (n = 50) sequenced in this study, along with their associated
124 metadata, are listed in Table 1. Henceforth, isolates SH-01 through SH-12 will collectively be
125 referred to as the *S.* Heidelberg subset; isolates SC-01 through SC-28 will be referred to as *S.*
126 Typhimurium var. O5- subset 1; and isolates SC-29 through SC-38 will be referred to as *S.*
127 Typhimurium var. O5- subset 2.

128 As part of NARMS surveillance, isolates SH-01, SH-02, SC-01, SC-29, and SC-35
129 through SC-37 were recovered and identified from retail meats (ground turkey, pork chop, and
130 chicken breast), purchased throughout the state of Pennsylvania between 2009-2014, following
131 standard protocols by the PADOH (26). Of note, the *S.* Heidelberg retail meat isolates were
132 derived from meats processed at different facilities; however, the *S.* Typhimurium var. O5-
133 subset 2 retail meat isolates SC-29, SC-35, and SC-36 were all derived from meats that were
134 originally processed at the same facility. All clinical *Salmonella* were from a collection of human
135 isolates that had indistinguishable PFGE patterns with retail meat isolates collected as part of
136 NARMS in Pennsylvania; the clinical isolates were submitted to the PADOH Bureau of
137 Laboratories by clinical laboratories in compliance with public health reporting requirements
138 (27).

Pulsed-field gel electrophoresis (PFGE) and antibiotic susceptibility testing (AST)

All retail meat and clinical isolates previously underwent PFGE and AST characterization through the PADOH. Each retail meat and clinical isolate was subjected to PFGE using the XbaI restriction enzyme, following the CDC's PulseNet protocol for *Salmonella* subtyping (28). For all isolates, except SH-09, SH-10, SC-19, and SC-24 through SC-28, the BlnI restriction enzyme was also used to obtain a secondary PFGE pattern. All retail meat and matched clinical isolates were also subjected to AST as previously described (29) for susceptibility to the antimicrobial agents: gentamicin, streptomycin, ampicillin, amoxicillin-clavulanic acid, cefoxitin, ceftiofur, ceftriaxone, azithromycin, chloramphenicol, nalidixic acid, ciprofloxacin, sulfisoxazole, trimethoprim-sulfamethoxazole, and tetracycline. Resistance was defined using CLSI criteria if available, otherwise NARMS interpretative criteria were used (30).

Genomic DNA extraction, library construction, and genome sequencing

A 3 mL Lysogeny Broth (LB) overnight culture, derived from a single colony, was prepared for each isolate. Total genomic DNA was extracted using the Wizard© Genomic DNA Purification Kit (Promega, Madison, WI, USA) following the manufacturer's instructions. Genomic DNA purity was confirmed via an A_{260}/A_{280} measurement (target ≥ 1.8) and the concentration was determined using the Qubit™ dsDNA Broad-Range quantification kit (Thermo Fisher Scientific, Waltham, MA, USA). Following quantification, genomic DNA was diluted to 0.2 ng/µL. A paired-end DNA library was prepared and normalized using the Nextera XT DNA Library Prep Kit (Illumina, Inc., San Diego, CA, USA). The resulting library was sequenced on an Illumina MiSeq sequencer (Illumina, Inc., San Diego, CA, USA) using a MiSeq reagent v2, 500-cycle kit, with 250 bp read length. Additionally, a representative isolate from each subset (SH-04, SC-09, and SC-31, respectively) was sequenced on a PacBio *RS* II (Pacific

Biosciences, Menlo Park, CA, USA) as previously described (31). Specifically, we prepared the library using 10 μg genomic DNA that was sheared to a size of 20-kb fragments by g-tubes (Covaris, Inc., Woburn, MA, USA) according to the manufacturer's instruction. The SMRTbell 20-kb template library was constructed using DNA Template Prep Kit 1.0 with the 20-kb insert library protocol (Pacific Biosciences, Menlo Park, CA, USA). Size selection was performed with BluePippin (Sage Science, Beverly, MA, USA). The library was sequenced using the P6/C4 chemistry on 2 single-molecule real-time (SMRT) cells with a 240-min collection protocol along with stage start. Analysis of the sequence reads was implemented using SMRT Analysis 2.3.0. The best *de novo* assembly was established with the PacBio Hierarchical Genome Assembly Process (HGAP3.0) program, which resulted in the closed chromosome of each isolate (Table S1). Each closed chromosome was annotated using Rapid Annotation using Subsystem Technology (RAST) (32).

Sequencing quality control

Following sequencing, Illumina read quality was confirmed using FastQC v0.11.5 (33). Raw Illumina reads for each isolate within each subset were aligned to the closed chromosome of SH-04, SC-09, and SC-31, respectively using Burrows-Wheeler Aligner v0.7.15 (BWA-MEM) (34). Subsequently, average genome coverage was calculated using the SAMtools v1.4 depth command (35). Additionally, Illumina reads from each isolate were *de novo* assembled using SPAdes v3.9 (36). QUAST v4.5 (37) was used to assess assembled/draft genome quality. All isolates sequenced in this study by Illumina technology had > 30X coverage, < 200 contigs, an N50 score > 200,000, and total assembly length between 4.4-5.l megabases (Mb) (Fig. S1A, B). The previously determined agglutination derived serotype of each isolate was also confirmed using SeqSero (38).

Comparative SNP and phylogenetic analyses

The validated single nucleotide polymorphism/variant (SNP or SNV) calling pipeline SNVPhyl v1.0.1 (39) was utilized to identify variants. Default parameters were used with the exception of the following: minimum coverage was set to 10X, minimum mean mapping was set to 30, and the SNV abundance ratio was set to 0.75. Preliminarily, the assembled genomes (from Illumina sequencing) of SH-04, SC-09, and SC-38 were used as the reference genomes for SNP and phylogenetic analyses in each subset (data not shown). Subsequently, the complete chromosome sequences (from PacBio sequencing) of isolates SH-04, SC-09, and SC-31 were used as the reference genomes for the *S.* Heidelberg subset, *S.* Typhimurium var. O5- subset 1, and *S.* Typhimurium var. O5- subset 2, respectively. The SNVPhyl pipeline outputted a concatenated SNP alignment file, a pairwise SNP distance matrix, and a maximum likelihood phylogenetic tree generated using PhyML v3.1.1 for each subset. The SNP alignment file was also used to construct a custom maximum likelihood phylogenetic tree using PhyML v3.1.1 (40). This tree was constructed using the GTR + gamma model with 1,000 bootstrap replicates.

SNP annotation

Concatenated base call files from SNVPhyl were downloaded and converted to VCF files using BCFtools v1.3.1 view (35); vcf-subset v0.1.13 (41) was used to filter out all non-variant positions. The filtered VCF files for each isolate were combined into a single VCF file using BCFtools merge for each subset. A custom SnpEff database was built for annotation of each subset using the RAST-generated GenBank file for either SH-04, SC-09, or SC-31. SNPs were then annotated using SnpEff v4.3 (42). SNP annotations were subsequently filtered to only include valid SNPs, as determined by SNVPhyl. Gene names for each valid SNP were extracted

from the RAST-generated GFF file of either SH-04, SC-09, or SC-31. SNP annotation script can be found at https://github.com/DudleyLabPSU/SNP-Annotation.

Genetic AMR profile determination

Genetic AMR determinants were identified in all genomes, using the Bacterial Antimicrobial Resistance Reference Gene Database (BARRGD) (Accession number PRJNA313047; accessed April 2018) and BLAST+ (43). To confirm BARRGD results and determine if any AMR-associated chromosomal point mutations were present, ResFinder 3.0 (44) was used. AMR genes that were detected by either method were only considered to be present if they had ≥ 90% nucleotide identity and ≥ 60% coverage (these parameters align with ResFinder's default search settings).

Plasmid identification and characterization

Known plasmid replicon sequences were identified using the PlasmidFinder (45) database and BLAST+. BLAST alignments were filtered to only include those present at ≥ 95% nucleotide identity and ≥ 60% coverage (these parameters align with PlasmidFinder's default search settings), with the exception of IncX1 in the *S*. Heidelberg subset, which was included due to its presence at 94.9% nucleotide identity. PacBio sequencing and subsequent processing (as described above) were also used to attain complete plasmid sequences within SH-04, SC-09, and SC-31 (Table S1). Each closed plasmid sequence was annotated using RAST (32). BLAST Ring Image Generator (BRIG) (46) was used to visualize plasmid comparisons in all subsets. Complete plasmid sequences were also compared to plasmids deposited in NCBI's GenBank database (Table S2) using BLAST (47) and BRIG, to assess novelty.

Accession numbers

Raw Illumina WGS data were submitted to NCBI and subsequently, assigned BioSample and SRA accession numbers (Table S3); all data can also be found under BioProject PRJNA357723. The closed chromosome and plasmid sequences of isolates SH-04, SC-09, and SC-31 were submitted to GenBank and their accession numbers are also found in Table S3.

Results

Comparison of retail meat and human isolates based on PFGE and AST results

In total, the PFGE patterns of 86 retail meat isolates were indistinguishable from the PFGE patterns of 1,665 clinical isolates located in the Pennsylvania surveillance database. From that larger collection, three subsets were chosen to whole-genome sequence and are referred to as the *S*. Heidelberg subset, *S*. Typhimurium var. O5- subset 1, and *S*. Typhimurium var. O5- subset 2. These subsets were selected based on quantity (≥ one retail meat matching multiple clinical isolates), AMR (primarily identical MDR patterns within each subset), and serovar relevance.

All isolates within the *S*. Heidelberg subset (two retail meat and ten clinical) matched by primary PFGE pattern (JF6X01.0058) and all isolates, except SH-09 and SH-10 (secondary patterns unknown), shared the same secondary PFGE pattern (JF6A26.0076) (Table 1). Isolates SH-01 through SH-10 displayed resistance to ampicillin, gentamicin, streptomycin, and tetracycline antimicrobials; whereas, isolates SH-11 and SH-12 were only resistant to ampicillin, gentamicin, and tetracycline (Table 1). All *S.* Typhimurium var. O5- subset 1 isolates (one retail meat and 27 clinical) shared the same primary PFGE pattern (JPXX01.0018) and isolates SC-01 through SC-17 also shared the same secondary PFGE pattern (JPXA26.0156) (Table 1). Additionally, all isolates within this subset were phenotypically resistant to ampicillin, chloramphenicol, streptomycin, sulfisoxazole, and tetracycline (ACSSuT pattern) (Table 1). All

253 *S.* Typhimurium var. O5- subset 2 isolates (four retail meat and six clinical) matched by primary
254 PFGE pattern (JPXX01.1283) and all isolates except SC-33 and SC-34 matched by secondary
255 PFGE pattern (JPXA26.0397) (Table 1). However, this subset was defined by two distinct AMR
256 profiles: isolates SC-29 through SC-34 displayed resistance to ampicillin, amoxicillin-clavulanic
257 acid, ceftiofur, ceftriaxone, cefoxitin, sulfisoxazole, and tetracycline; whereas, isolates SC-35
258 through SC-38 were resistant to only sulfisoxazole and tetracycline antimicrobials (Table 1).

259 ***Comparative SNP analysis resolved phylogenetic relationships in each subset***

260 To retrospectively resolve the genetic relatedness of each PFGE-matched retail meat and
261 clinical isolate within all three subsets, the validated SNP-calling pipeline, SNVPhyl was
262 utilized. In total, the *S.* Heidelberg subset was defined by 35 SNPs. These SNPs were generally
263 distributed over the length of the reference chromosome (Fig. S2A, top) and were predominantly
264 non-synonymous (Fig. S2B); notably, SNP annotation revealed that missense mutations were
265 located in flagellar-associated genes (*fliC* and *fliE*) in some isolates. In terms of SNP distances,
266 the retail meat isolates (SH-01 and SH-02) were separated from their ten PFGE-matched clinical
267 isolates by 6 to 12 SNPs (Table 2). Interestingly, despite being derived from two different
268 processing facilities, the two retail meat isolates were separated from one another by 1 unique
269 SNP (Table 2); a missense variant (Ala138Thr) in *rapA*. Furthermore, within this subset, six
270 clinical isolates were previously determined to be outbreak-associated (outbreak A). These
271 isolates (SH-03 through SH-06, SH-11, and SH-12) were separated from each other by 3 to 9
272 SNPs (Table 2). Additionally, clinical isolates SH-09 and SH-10 were previously determined to
273 be part of a separate outbreak (outbreak B). Indeed, under these experimental conditions, they
274 were separated by 0 SNPs (Table 2). Moreover, to visualize the phylogenetic relationships of the
275 *S.* Heidelberg isolates, a maximum likelihood phylogenetic tree was constructed. The retail meat

276 isolates branched distinctly away from the *S.* Heidelberg clinical isolates (Fig. 1). The clinical
277 isolates all clustered on the same primary branch; notably, sporadic clinical isolate SH-08
278 clustered with the confirmed outbreak isolates (Fig. 1). Furthermore, the use of the draft genome
279 of SH-04 as the reference genome resulted in nearly identical SNP distances and phylogenetic
280 tree topologies; these similarities were also observed in each *S.* Typhimurium var. O5- subset,
281 when the SC-09 or SC-38 assembled genome was utilized as the reference (data not shown).
282 Furthermore, *S.* Typhimurium var. O5- subset 1 was defined by 482 SNPs that were
283 uniformly distributed across the reference chromosome (Fig. S2A, middle). The majority of the
284 SNPs were non-synonymous (Fig. S2C). Moreover, in some *S.* Typhimurium var. O5- subset 1
285 isolates, missense mutations were located in multiple fimbrial-associated genes (*fimD*, *stdB,*
286 *fimF*), type III secretion system 1 (T3SS1)-associated genes (*spaN*, *sirC*), and flagellar-
287 associated genes (*fliC*, *motA*). The overall pairwise SNP distances within this collection ranged
288 from 0 to 105 (Table 3). Even though the retail meat isolate (SC-01) shared the same primary
289 and secondary PFGE patterns and drug resistance profile as clinical isolates SC-02 through SC-
290 17, on the genome level, it was surprisingly separated from each of those isolates by 41 to 75
291 SNPs (Table 3). Furthermore, SC-01 was separated from clinical isolates SC-18 through SC-28
292 by 46 to 96 SNPs, despite matching by primary PFGE pattern and drug resistance (Table 3).
293 Nonetheless, some clinical isolates were genetically indistinguishable from one another (SC-05
294 and SC-06 and SC-14 and SC-19) or otherwise very closely related (shown in light grey shading
295 in Table 3). A maximum likelihood tree was also constructed for visualization of phylogenetic
296 relationships. Retail meat isolate SC-01 generally clustered with clinical isolates SC-02 through
297 SC-06 and SC-17, but importantly, SC-01 was still separated from those isolates by 41 to 44
298 SNPs (Fig. 2 and Table 3). Interestingly, these isolates all originated from the southern portion of

299 Pennsylvania. SC-01 and SC-06 were classified as southeast; SC-04 and SC-05 were classified
300 as southcentral; SC-02, SC-03, and SC-17 were classified as southwest. The remainder of the
301 clinical isolates branched separately, with some forming distinct clusters consistent with the
302 pairwise SNP distances (Fig. 2).

303 Lastly, *S.* Typhimurium var. O5- subset 2 was defined by 225 SNPs that were evenly
304 dispersed across the reference chromosome (Fig. S2A, bottom). Similar to the previous two
305 subsets, the majority of the SNPs were non-synonymous (Fig. S2D). Interestingly, in some *S.*
306 Typhimurium var. O5- subset 2 isolates missense or nonsense mutations were located in genes
307 that encode the T3SS1 effector *sipA* and the type III secretion system 2 (T3SS2) structural
308 component and effector protein, *ssaV* and *sseF*. In terms of relatedness, the first retail meat
309 isolate, SC-29, was separated from its primary PFGE, secondary PFGE, and AST-matched
310 clinical isolates (SC-30 through SC-32) by 58, 61, and 27 SNPs, respectively (Table 4). The
311 remaining three retail meat isolates, SC-35, SC-36, and SC-37, and clinical isolate SC-38
312 matched by both PFGE patterns and drug resistance. However, SC-38 was separated from SC-
313 35, SC-36, and SC-37 by 24, 27, and 31 SNPs, respectively (Table 4). Of note, the smallest
314 pairwise SNP distance between a clinical and retail meat isolate was 21, which occurred between
315 clinical isolate SC-32 and retail meat isolate SC-35 (Table 4); these isolates matched by both
316 PFGE patterns, but did not share the same resistance phenotype. In total, all four retail meat
317 isolates were separated from the six clinical isolates in this subset by 21 to 81 SNPs (Table 4).
318 When visualized on a phylogenetic tree, two distinct clusters of isolates were observed (Fig. 3).
319 Clinical isolates SC-30, SC-31, SC-33, and SC-34 all distinctly branched away from the four
320 retail meat isolates and furthermore, were fairly dissimilar from one another as well (Fig. 3).
321 Conversely, clinical isolates SC-32 and SC-38 clustered with all four retail meat isolates, but

importantly, these isolates were still separated from the retail meat isolates by 21 to 31 SNPs (Fig. 3 and Table 4). In general, the trend of relatedness was the same in both *S.* Typhimurium var. O5- subsets, in that, isolates that matched by conventional methods, were all separated by a relatively large number of SNPs, especially when compared to the *S.* Heidelberg collection.

Genetic AMR profile generally aligned with observed AMR phenotype in each subset

To further characterize these isolates, assess whether they carry the same AMR genes, and determine AMR phenotype/genotype correlations, BARRGD and ResFinder 3.0 were used. Within the *S.* Heidelberg subset, six AMR genes were identified: *aadA1* (STR^r), *aac(3)-IId* (GEN^r), *aph(3'')-Ib/strA* (STR^r), *bla*$_{\text{TEM-1B}}$ (broad-spectrum β-lactam^r), *tet(A)* (TET^r), and *fosA7* (fosfomycin; FOF^r) (Fig. 4A). All six of these genes, except for *aph(3'')-Ib/strA,* were identified in the assembled genomes of all isolates; only isolates SH-01 through SH-03, SH-06, SH-07, SH-09, and SH-10 carried *aph(3'')-Ib/strA* (Fig. 4A). Interestingly, the genome of SH-04, when sequenced by PacBio, housed an additional copy of *bla*$_{\text{TEM-1B}}$ (Fig. 4A). Indeed, *bla*$_{\text{TEM-1B}}$ was located in a relatively short contig in the draft genomes of all isolates, likely indicating it did not assemble well, providing a possible explanation for why it was detected once in the MiSeq-based assembly, but twice in the PacBio-based genome. No known AMR-conferring chromosomal mutations were identified in any isolate within this subset or in either *S.* Typhimurium var. O5- subset.

Within *S.* Typhimurium var. O5- subset 1, a total of six different AMR genes were identified: *aadA2* (STR^r), *bla*$_{\text{PSE}}$/*bla*$_{\text{CARB-2}}$ (broad-spectrum β-lactam^r), *tet(G)* (TET^r), *floR* (CHL^r), *sul1* (FIS^r), and *catA1* (CHL^r) (Fig. 4B). In addition, *qacEΔ*1, a gene that confers resistance to quaternary ammonium compounds was identified in all isolates (Fig. 4B). All AMR genes, but *catA1*, were identified in each isolate; *catA1* was only present in clinical isolate SC-23

(Fig. 4B). Within the PacBio genome of SC-09, an additional copy of *qacEΔ1* was identified, as well as a partial and complete copy of *sul1*, which was only partially present once (~66% query coverage) in the assembled genome of each isolate (Fig. 4B).

Within the last subset, *S.* Typhimurium var. O5- subset 2, three AMR genes were identified: $bla_{\text{CMY-2}}$ (broad and extended-spectrum β-lactam^{r}), *tet(A)* (TET^{r}), and *sul2* (FIS^{r}), (Fig. 4C). Each AMR gene, except $bla_{\text{CMY-2}}$, was identified in each assembled genome; $bla_{\text{CMY-2}}$ was only identified in isolates SC-29 through SC-34 (Fig. 4C). Moreover, within the PacBio genome of SC-31, two $bla_{\text{CMY-2}}$ genes were identified (Fig. 4C).

Accordingly, across all three subsets, there was a 100% correlation between AMR phenotype and genotype for β-lactam, gentamicin (an aminoglycoside), chloramphenicol, sulfonamide, and tetracycline antimicrobials; whereas, there was a 95% correlation for streptomycin (an aminoglycoside) (Table 5). Additionally, no correlation for fosfomycin could be calculated, as resistance to this drug was not phenotypically tested for.

Identification of plasmid and chromosomal AMR determinants in each subset

Lastly, to further elucidate the genetic AMR profile of each isolate, plasmids were identified in an effort to determine and compare AMR gene location in each isolate. PacBio sequencing of SC-09, from *S.* Typhimurium var. O5- subset 1, revealed one plasmid with the replicon sequences, IncFIB(S), IncFII(S). This plasmid (pSC-09-1) was conserved in all isolates within this subset and was found to house multiple virulence-associated genes, but no AMR genes were identified (Fig. S3). This observation indicated that AMR was chromosomally carried in this subset. Indeed, all AMR genes identified in this subset were located on the SC-09 chromosome. Furthermore, following alignment, it was found that all isolates carried their AMR genes within an approximate 12-kb chromosomal region (Fig. 5). Moreover, annotation

368 demonstrated that approximately 3-kb upstream of *sul1* is an integrase and immediately
369 downstream of *aadA2* is an additional integrase, suggesting that all AMR genes are associated
370 with a similar mobile genetic element in this subset.

371 Within the *S.* Heidelberg subset, PacBio sequencing of SH-04 resulted in two plasmids,
372 pSH-04-1 (IncI1-alpha) and pSH-04-2 (IncX1). pSH-04-1 was present in both retail meat and all
373 ten clinical isolates and housed four (*aac(3)-IId*, $bla_{\text{TEM-1B}}$, *aadA1*, and *tet(A)*) of the six
374 identified AMR genes (Fig. 6A). pSH-04-2 was also conserved in all twelve isolates and carried
375 an additional copy of $bla_{\text{TEM-1B}}$ (Fig. 6B). Interestingly, a BLAST search indicated that pSH-04-2
376 was different than other known *Salmonella* plasmids, with the top three hits sharing 98-99%
377 nucleotide identity and 60-95% query coverage (pFDAARGOS_312_2, pSE95-0621-1, and
378 pSTY3-1898) (Fig. S4A; Table S2); the main differences between the plasmids in Fig. S4A were
379 the presence/absence of various mobile element and hypothetical proteins. Additionally, through
380 annotation of the SH-04 chromosome, it was determined that *fosA7* was chromosomally-carried.

381 Finally, sequencing of SC-31, from *S.* Typhimurium var. O5- subset 2, resulted in three
382 complete plasmid sequences. pSC-31-1 (IncI1-alpha) housed the beta-lactamase, $bla_{\text{CMY-2}}$.
383 Moreover, comparative sequence analysis demonstrated that pSC-31-1 was only present in
384 isolates SC-29 through SC-33 (Fig. 7A). This result aligned clearly with phenotype as SC-29
385 through SC-33 were resistant to β-lactams, but SC-35 through SC-38 were sensitive. However,
386 SC-34 also displayed resistance to β-lactam antibiotics, but did not appear to have this plasmid,
387 despite carrying the $bla_{\text{CMY-2}}$ gene (Fig. 4C). Therefore, it is postulated that SC-34 carries this
388 gene on its chromosome. Supporting this hypothesis, SC-31 also carried a chromosomal copy of
389 $bla_{\text{CMY-2}}$. Conversely, pSC-31-2 (IncA/C2) was present in all isolates within this subset (Fig. 7B)
390 and was found to house the remaining two AMR genes, *sul2* and *tet(A)*. Furthermore, a BLAST

search revealed that only two other plasmids have been sequenced that show high similarity to this plasmid: pCFSAN001921 (99% nucleotide identity/100% query coverage) and pFDAARGOS_312_3 (99% nucleotide identity/92% query coverage) (Fig. S4B). Interestingly, both of those plasmids were also isolated from *S.* Typhimurium var. O5-. Lastly, pSC-31-3 (ColpVC) was only fully present in SC-29, SC-31, and SC-34, and housed no known AMR determinants (Fig. 7C).

Discussion

An important source of *Salmonella* is retail meat; in Pennsylvania, previous studies have demonstrated that non-typhoidal *Salmonella*, including isolates that are drug-resistant, are prevalent on meat products sold at grocery stores and farmers' markets (48, 49). Here, we exploited the strength of WGS to compare the genetic characteristics of three historic subsets of drug-resistant and PFGE-matched retail meat and clinical *S.* Heidelberg and *S.* Typhimurium var. O5- isolates, to reassess their relatedness and identify their resistome.

Two previous studies have used WGS to compare *S.* Heidelberg isolated from multiple sources, including humans and retail meats. Hoffmann *et al.* (14) utilized 454 sequencing and SNP analysis to distinguish outbreak-associated Heidelberg isolates from non-outbreak isolates, collected from several sources between 1982 and 2011, with similar PFGE patterns. Edirmanasinghe *et al.* (24) used WGS to characterize *S.* Heidelberg from various sources, collected through routine surveillance in Canada. However, in this study, we focused on directly comparing the genomic characteristics of 50 historic retail meat and human *S.* Heidelberg and *S.* Typhimurium var. O5- isolates that matched, importantly, by multiple conventional methods (PFGE and AST) and were collected through routine surveillance in Pennsylvania between 2009 and 2014. The narrow geographical region and temporal distribution of this collection is

reflective of what other state public health laboratories would examine during their own retrospective and comparative analyses using WGS.

Comparative SNP analysis in *S.* Heidelberg subset reveals isolates are closely related

Within the *S.* Heidelberg subset, comparative SNP analysis revealed that the retail meat isolates, SH-01 and SH-02, were separated by 6 to 12 SNPs from the ten matched clinical isolates (Table 2). Prior studies have reported SNP differences of 0 to 4, an average of 17, and 4 to 19 for previous confirmed *S.* Heidelberg outbreaks (11, 14, 19). Thus, the observed SNP distances between the retail meat and clinical isolates within this subset align with outbreak-associated values reported in the literature for the Heidelberg serovar, indicating that these isolates are closely related and may share a common source. However, due to the historic nature of these isolates, an epidemiological link was not investigated between the food and human isolates in this case. Moreover, the retail meat isolates were collected in 2013, whereas, the majority of the clinical isolates were isolated from 2010 and 2011. Accordingly, these data underscore the importance of interpreting genomic results in the context of epidemiological data, which is particularly crucial when analyzing historic isolate collections.

Comparative SNP analysis in *S.* Typhimurium var. O5- subsets reveals genetic differences

Conversely, comparative SNP analysis revealed a range of phylogenetic relationships within each *S.* Typhimurium var. O5- subset (Tables 3, 4). Importantly, the retail meat isolate within subset 1 was separated from all clinical isolates by 41 to 96 SNPs (Table 3); similarly, the four retail meat isolates within subset 2 were separated from the six clinical isolates by 21 to 81 SNPs (Table 4). Previous reports have determined that *S.* Typhimurium outbreak-associated isolates have been separated by 2 to 12, 3 to 30, 0 to 12, 0 to 7, and a maximum of 3 SNPs (19, 22, 50, 51). Furthermore, others have proposed specific SNP cutoffs to classify an isolate as

437 outbreak-associated (22); however, a clearly defined consensus in the field on a maximum SNP
438 distance threshold for outbreak analysis has not been established. Nonetheless, our data suggest
439 that there is not a genetic link between the retail meat and clinical isolates within the two *S.*
440 Typhimurium var. O5- subsets, despite them matching by conventional methods. Indeed,
441 previous studies have found similar genetic distances between PFGE-matched sporadic and
442 outbreak isolates; for example, one sporadic *S.* Bareilly isolate that shared the same primary
443 PFGE pattern as an *S.* Bareilly outbreak was separated by 117 SNPs from those strains (intra-
444 outbreak SNP distance was 1 to 6) (20).

445 **AMR phenotype and resistome were well correlated in each subset**

446 Within this study, we also determined the resistome of each isolate, which included the
447 identification of specific AMR genes and elucidation of each gene's genomic context, in an
448 effort to assess the role of WGS in enhancing integrated surveillance for drug-resistant
449 *Salmonella* in Pennsylvania. Overall, the AMR profiles, as determined by conventional AST,
450 correlated well with AMR genotype. Across all subsets, a 100% correlation was observed for all
451 tested antimicrobials, except streptomycin, where a 95% correlation was observed (Table 5).
452 Similarly, in a comprehensive *Salmonella* study encompassing 640 isolates, the AMR
453 phenotype/genotype correlation was 99% (13); these results are corroborated by a smaller study
454 that observed a 100% correlation between AMR phenotype and genotype for 49 *Salmonella*
455 isolates from swine (52). Furthermore, the slightly lower streptomycin correlation that we
456 observed has been noted previously. McDermott *et al*. (13) postulate that this discordance
457 between phenotype and genotype is likely the result of a MIC breakpoint value that is too high.
458 Consequently, another study found that by lowering the breakpoint MIC value for streptomycin
459 resistance from ≥ 64 μg/L to ≥ 32 μg/L, a higher correlation was observed (53). Thus, it is

460 plausible that the lower correlation observed in this study is also the result of not enough isolates
461 being considered resistant based on standard MIC testing; however, silent genes or a non-
462 functional protein product could also be responsible.

463 **AMR is primarily plasmid-mediated in the *S.* Heidelberg subset**

464 Within the *S.* Heidelberg subset, six AMR genes were identified (Fig. 4A). These genes
465 or their close variants have been identified in *S.* Heidelberg previously (14, 54–58). Specifically,
466 *fosA7* was identified in *Salmonella* for the first time in 2017; this particular gene sequence was
467 only found in 35 of the approximately 40,000 *Salmonella* draft and complete genomes in NCBI,
468 of which 75% were *S.* Heidelberg (58). Notably, *fosA7* was located in both retail meat and all ten
469 clinical *S.* Heidelberg isolates in this study. Moreover, PacBio sequencing revealed that this gene
470 was located on the chromosome of SH-04, consistent with prior data suggesting that *fosA7* is
471 exclusively chromosomal in *Salmonella* (58). PacBio sequencing also elucidated the presence of
472 multi-copy AMR genes, which were not detected in the draft genomes within this subset (i.e.
473 $bla_{\text{TEM-1B}}$) and in each *S.* Typhimurium var. O5- subset as well; this observation highlights a
474 potential advantage of incorporating long-read sequencing into routine surveillance methods.

475 In addition, comparative plasmid analysis identified two plasmids that were present in
476 each retail meat and clinical isolate. pSH-04-1, an IncI1-alpha plasmid, carried four of the six
477 AMR genes (*aadA1, aac(3)-IId*, $bla_{\text{TEM-1B}}$, *tet(A)*) (Fig. 6A). Similar resistance genes have been
478 reported on an IncI1 plasmid in *S.* Heidelberg previously (14). This replicon type plasmid has
479 also been found to house sulfonamide resistance and importantly, the bla_{CMY} beta-lactamase,
480 which encodes resistance to extended-spectrum cephalosporins, in *S.* Heidelberg (24, 57, 59).
481 pSH-04-2, an IncX1 plasmid, housed an additional copy of $bla_{\text{TEM-1B}}$ (Fig. 6B); IncX1 plasmids
482 have been identified in *S.* Heidelberg isolates previously (14).

AMR is chromosomally-carried in *S.* Typhimurium var. O5- subset 1

Within *S.* Typhimurium var. O5- subset 1, seven different resistance genes were identified (Fig. 4B). Annotation and subsequent alignment of each isolate's chromosome determined that all of the identified genes, with the exception of *catA1* in isolate SC-23, were located within a 12-kb chromosomal region (Fig. 5). This AMR gene topology is typical of *Salmonella* Typhimurium strains that also display the ACSSuT penta-resistance pattern. Previous work has demonstrated that a region of the chromosome, termed *Salmonella* genomic island 1 (SGI1), houses the AMR gene cluster that is responsible for this phenotype; within this region, the AMR genes, *floR* and *tet(G)*, are flanked by integrons carrying the AMR genes, *aadA2* and bla_{PSE}/$bla_{\text{CARB-2}}$ (60, 61). Consistent with this observation, the single plasmid identified within this subset did not carry AMR genes.

AMR is plasmid-mediated in *S.* Typhimurium var. O5- subset 2

Lastly, within *S.* Typhimurium var. O5- subset 2, three AMR genes were identified (Fig. 4C). These genes have previously been identified in *S.* Typhimurium and *S.* Typhimurium var. O5- (62–64). Comparative plasmid analysis revealed that pSC-31-1 (IncI1-alpha), which carried $bla_{\text{CMY-2}}$, was present only in isolates SC-29 through SC-33 (Fig. 7A). Indeed, IncI1 plasmids are frequently associated with this gene in *Salmonella*, including *S.* Typhimurium var. O5- (62). Moreover, this genetic observation is consistent with the AST results, as isolates SC-29 through SC-33 were resistant to β-lactams, whereas, SC-35 through SC-38 were sensitive. Intriguingly, isolate SC-34 was resistant to β-lactams and was found to carry the $bla_{\text{CMY-2}}$ gene; however, the IncI1-alpha plasmid was not present in this isolate. Indeed, a previous *S.* Heidelberg study identified $bla_{\text{CMY-2}}$ on the chromosome, which is suggestive of plasmid integration being possible (24). Conversely, the second plasmid, pSC-31-2, was found in all retail meat and clinical isolates

within this subset (Fig. 7B). This 188 kb IncA/C2 plasmid housed the AMR genes, *sul2* and *tet(A)*. Notably, a BLAST query indicated that this plasmid was only similar to two other plasmids, both of which were from other *S.* Typhimurium var. O5- isolates (Fig. S4B). When comparing these plasmids to pSC-31-2, the main differences were the presence or absence of multiple conjugal transfer-associated and hypothetical proteins. Accordingly, these data suggest that pSC-31-2 is a novel version of an IncA/C2, AMR-encoding plasmid that appears to generally be restricted to the serovar Typhimurium var. O5-.

In summary, this study demonstrated that historic retail meat and human *Salmonella* isolates, collected through routine monitoring, that are indistinguishable by the conventional methods PFGE and AST, could be different strains, further underscoring the power of WGS. These data also highlight the importance and necessity of interpreting WGS data in the context of epidemiological findings—a point that is particularly crucial, when analyzing historic isolate collections. In addition, we evaluated the role of WGS in enhancing integrated surveillance of drug-resistant *Salmonella* from retail meat and clinical sources in Pennsylvania, by identifying resistance genes and characterizing their genomic environment, as this information is vital to understanding how resistance is disseminating. We observed that resistance phenotype and genotype correlated well in each isolate and that the same AMR-encoding plasmids were found in both the retail meat and clinical isolates. As one of the first studies to directly compare historic retail meat and clinical *Salmonella* isolates using WGS, these results demonstrate the usefulness and value of WGS to public health laboratories performing retrospective comparisons of bacterial isolates from multiple sources.

Acknowledgements

We thank Lisa Dettinger, James Tait, Carina Davis, Melinda Johnston, and Barry Perry for assistance with preparation of *Salmonella* isolates. A.B.K. was supported by The Black Endowed Graduate Fellowship through the Penn State College of Agricultural Sciences. This work was supported by the U.S. Food and Drug Administration grant number 1U18FD006222-01 to E.G.D. for support of GenomeTrakr in Pennsylvania, the U.S. Food and Drug Administration grant number 1U01FD006253-01 to PADOH for collaboration in the National Antimicrobial Resistance Monitoring System, Centers for Disease Control and Prevention Epidemiology and Laboratory Capacity (ELC) for collaboration in National Antimicrobial Resistance Monitoring (CDC-RFA-CI10-101204PPHF13), and the USDA National Institute of Food and Agriculture Federal Appropriations under Project PEN04522 and Accession number 0233376.

References

1. Scallan E, Hoekstra RM, Angulo FJ, Tauxe RV, Widdowson M-A, Roy SL, Jones JL, Griffin PM. 2011. Foodborne illness acquired in the United States—major pathogens. Emerg Infect Dis 17:7–15.
2. CDC. 2013. Antibiotic resistance threats in the United States, 2013. Atlanta, Georgia. https://www.cdc.gov/drugresistance/threat-report-2013/pdf/ar-threats-2013-508.pdf
3. Jackson BR, Griffin PM, Cole D, Walsh KA, Chai SJ. 2013. Outbreak-associated *Salmonella enterica* serotypes and food commodities, United States, 1998-2008. Emerg Infect Dis 19:1239–1244.
4. CDC. 2017. Surveillance for foodborne disease outbreaks, United States, 2015, annual report. Atlanta, Georgia.

https://www.cdc.gov/foodsafety/pdfs/2015FoodBorneOutbreaks_508.pdf

5. Boore AL, Hoekstra RM, Iwamoto M, Fields PI, Bishop RD, Swerdlow DL. 2015. *Salmonella enterica* infections in the United States and assessment of coefficients of variation: a novel approach to identify epidemiologic characteristics of individual serotypes, 1996-2011. PLoS One 10:e0145416.

6. FDA. 2015. 2014-2015 Retail meat interim report. Rockville, MD. https://www.fda.gov/downloads/AnimalVeterinary/SafetyHealth/AntimicrobialResistance/NationalAntimicrobialResistanceMonitoringSystem/UCM498134.pdf

7. Acheson D, Hohmann EL. 2001. Nontyphoidal salmonellosis. Clin Infect Dis 32:263–269.

8. White House. 2014. National Strategy For Combating Antibiotic-Resistant Bacteria. Washington DC. https://obamawhitehouse.archives.gov/sites/default/files/docs/carb_national_strategy.pdf

9. Gilbert JM, White DG, McDermott PF. 2007. The US National Antimicrobial Resistance Monitoring System. Future Microbiol 2:493–500.

10. Salipante SJ, SenGupta DJ, Cummings LA, Land TA, Hoogestraat DR, Cookson BT. 2015. Application of whole-genome sequencing for bacterial strain typing in molecular epidemiology. J Clin Microbiol 53:1072–1079.

11. Bekal S, Berry C, Reimer AR, Van Domselaar G, Beaudry G, Fournier E, Doualla-Bell F, Levac E, Gaulin C, Ramsay D, Huot C, Walker M, Sieffert C, Tremblay C. 2016. Usefulness of high-quality core genome single-nucleotide variant analysis for subtyping the highly clonal and the most prevalent *Salmonella enterica* serovar Heidelberg clone in the context of outbreak investigations. J Clin Microbiol 54:289–295.

12. Kahlmeter G. 2014. Defining antibiotic resistance-towards international harmonization.

575 Ups J Med Sci 119:78–86.

576 13. McDermott PF, Tyson GH, Kabera C, Chen Y, Li C, Folster JP, Ayers SL, Lam C, Tate
577 HP, Zhao S. 2016. Whole-genome sequencing for detecting antimicrobial resistance in
578 Nontyphoidal *Salmonella*. Antimicrob Agents Chemother 60:5515–5520.

579 14. Hoffmann M, Zhao S, Pettengill J, Luo Y, Monday SR, Abbott J, Ayers SL, Cinar HN,
580 Muruvanda T, Li C, Allard MW, Whichard J, Meng J, Brown EW, McDermott PF. 2014.
581 Comparative genomic analysis and virulence differences in closely related *Salmonella*
582 *enterica* serotype Heidelberg isolates from humans, retail meats, and animals. Genome
583 Biol Evol 6:1046–1068.

584 15. Phillips A, Sotomayor C, Wang Q, Holmes N, Furlong C, Ward K, Howard P, Octavia S,
585 Lan R, Sintchenko V. 2016. Whole genome sequencing of *Salmonella* Typhimurium
586 illuminates distinct outbreaks caused by an endemic multi-locus variable number tandem
587 repeat analysis type in Australia, 2014. BMC Microbiol 16.

588 16. den Bakker HC, Allard MW, Bopp D, Brown EW, Fontana J, Iqbal Z, Kinney A,
589 Limberger R, Musser KA, Shudt M, Strain E, Wiedmann M, Wolfgang WJ. 2014. Rapid
590 whole-genome sequencing for surveillance of *Salmonella enterica* serovar Enteritidis.
591 Emerg Infect Dis 20:1306–1314.

592 17. Allard MW, Luo Y, Strain E, Li C, Keys CE, Son I, Stones R, Musser SM, Brown EW.
593 2012. High resolution clustering of *Salmonella enterica* serovar Montevideo strains using
594 a next-generation sequencing approach. BMC Genomics 13:32.

595 18. Wilson MR, Brown E, Keys C, Strain E, Luo Y, Muruvanda T, Grim C, Jean-Gilles
596 Beaubrun J, Jarvis K, Ewing L, Gopinath G, Hanes D, Allard MW, Musser S. 2016.
597 Whole genome DNA sequence analysis of *Salmonella* subspecies *enterica* serotype

598 Tennessee obtained from related peanut butter foodborne outbreaks. PLoS One
599 11:e0146929.
600 19. Leekitcharoenphon P, Nielsen EM, Kaas RS, Lund O, Aarestrup FM. 2014. Evaluation of
601 whole genome sequencing for outbreak detection of *Salmonella enterica*. PLoS One
602 9:e87991.
603 20. Hoffmann M, Luo Y, Monday SR, Gonzalez-Escalona N, Ottesen AR, Muruvanda T,
604 Wang C, Kastanis G, Keys C, Janies D, Senturk IF, Catalyurek U V., Wang H, Hammack
605 TS, Wolfgang WJ, Schoonmaker-Bopp D, Chu A, Myers R, Haendiges J, Evans PS, Meng
606 J, Strain EA, Allard MW, Brown EW. 2016. Tracing origins of the *Salmonella* Bareilly
607 strain causing a food-borne outbreak in the United States. J Infect Dis 213:502–508.
608 21. Taylor AJ, Lappi V, Wolfgang WJ, Lapierre P, Palumbo MJ, Medus C, Boxrud D. 2015.
609 Characterization of foodborne outbreaks of *Salmonella enterica* serovar Enteritidis with
610 whole-genome sequencing single nucleotide polymorphism-based analysis for
611 surveillance and outbreak detection. J Clin Microbiol 53:3334–3340.
612 22. Octavia S, Wang Q, Tanaka MM, Kaur S, Sintchenko V, Lan R. 2015. Delineating
613 community outbreaks of *Salmonella enterica* serovar Typhimurium by use of whole-
614 genome sequencing: insights into genomic variability within an outbreak. J Clin Microbiol
615 53:1063–1071.
616 23. CDC. 2014. NARMS 2014 human isolates surveillance report. Atlanta, Georgia.
617 https://www.cdc.gov/narms/pdf/2014-annual-report-narms-508c.pdf
618 24. Edirmanasinghe R, Finley R, Parmley EJ, Avery BP, Carson C, Bekal S, Golding G,
619 Mulvey MR. 2017. A whole genome sequencing approach to study cefoxitin-resistant
620 *Salmonella enterica* serovar Heidelberg from various aources. Antimicrob Agents

621 Chemother 61:e01919-e16.

622 25. Zhao S, Tyson GH, Chen Y, Li C, Mukherjee S, Young S, Lam C, Folster JP, Whichard
623 JM, McDermott PF. 2015. Whole-genome sequencing analysis accurately predicts
624 antimicrobial resistance phenotypes in *Campylobacter* spp. Appl Environ Microbiol
625 82:459–466.

626 26. FDA. 2012. 2012 Retail meat report National Antimicrobial Resistance Monitoring
627 System. Laurel, Maryland.
628 https://www.fda.gov/downloads/animalveterinary/safetyhealth/antimicrobialresistance/nati
629 onalantimicrobialresistancemonitoringsystem/ucm442212.pdf

630 27. Pennsylvania Department of Health. List of reportable diseases.
631 http://www.health.pa.gov/Your-Department-of-
632 Health/Offices%20and%20Bureaus/epidemiology/Pages/Reportable-
633 Diseases.aspx#.Wy6PXBJKho6

634 28. CDC. 2017. Standard Operating Procedure for PulseNet PFGE of *Escherichia coli*
635 O157:H7, *Escherichia coli* non-O157 (STEC), *Salmonella* serotypes, *Shigella sonnei* and
636 *Shigella flexneri*. Atlanta, GA.
637 https://www.cdc.gov/pulsenet/pdf/ecoli-shigella-salmonella-pfge-protocol-508c.pdf

638 29. Sandt CH, Fedorka-Cray PJ, Tewari D, Ostroff S, Joyce K, Mikanatha NM. 2013. A
639 comparison of non-typhoidal *Salmonella* from humans and food animals using pulsed-
640 field gel electrophoresis and antimicrobial susceptibility patterns. PLoS One 8:e77836.

641 30. FDA. Breakpoints Table 1. interpretive criteria used for susceptibility testing of
642 *Salmonella* and *E. coli*.
643 https://www.fda.gov/downloads/AnimalVeterinary/SafetyHealth/AntimicrobialResistance/

644 NationalAntimicrobialResistanceMonitoringSystem/UCM581395.pdf

645 31. Yao K, Muruvanda T, Roberts RJ, Payne J, Allard MW, Hoffmann M. 2016. Complete
646 genome and methylome sequences of *Salmonella enterica* subsp. *enterica* serovar Panama
647 (ATCC 7378) and *Salmonella enterica* subsp. *enterica* serovar Sloterdijk (ATCC 15791).
648 Genome Announc 4:e00133-16.

649 32. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S,
650 Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil
651 LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O,
652 Vonstein V, Wilke A, Zagnitko O. 2008. The RAST Server: Rapid Annotations using
653 Subsystems Technology. BMC Genomics 9:75.

654 33. Babraham Bioinformatics. 2016. FastQC v0.11.5.

655 34. Li H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-
656 MEM. arXiv 0:1–3.

657 35. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G,
658 Durbin R, 1000 Genome Project Data Processing Subgroup 1000 Genome Project Data
659 Processing. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics
660 25:2078–2079.

661 36. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM,
662 Nikolenko SI, Pham S, Prjibelski AD, Pyshkin A V, Sirotkin A V, Vyahhi N, Tesler G,
663 Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its
664 applications to single-cell sequencing. J Comput Biol 19:455–477.

665 37. Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: quality assessment tool for
666 genome assemblies. Bioinformatics 29:1072–1075.

667 38. Zhang S, Yin Y, Jones MB, Zhang Z, Deatherage Kaiser BL, Dinsmore BA, Fitzgerald C,
668 Fields PI, Deng X. 2015. *Salmonella* serotype determination utilizing high-throughput
669 genome sequencing data. J Clin Microbiol 53:1685–1692.
670 39. Petkau A, Mabon P, Sieffert C, Knox NC, Cabral J, Iskander M, Iskander M, Weedmark
671 K, Zaheer R, Katz LS, Nadon C, Reimer A, Taboada E, Beiko RG, Hsiao W, Brinkman F,
672 Graham M, Van Domselaar G. 2017. SNVPhyl: a single nucleotide variant phylogenomics
673 pipeline for microbial genomic epidemiology. Microb Genomics 3.
674 40. Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New
675 algorithms and methods to estimate maximum-likelihood phylogenies: assessing the
676 performance of PhyML 3.0. Syst Biol 59:307–321.
677 41. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, Depristo MA, Handsaker RE,
678 Lunter G, Marth GT, Sherry ST, Mcvean G, Durbin R, Project G, Group A. 2011. The
679 variant call format and VCFtools. Bioinforma Appl Note 27:2156–2158.
680 42. Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden
681 DM. 2012. A program for annotating and predicting the effects of single nucleotide
682 polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118;
683 iso-2; iso-3. Fly (Austin) 6:80–92.
684 43. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL.
685 2009. BLAST+: architecture and applications. BMC Bioinformatics 10:421.
686 44. Zankari E, Hasman H, Cosentino S, Vestergaard M, Rasmussen S, Lund O, Aarestrup FM,
687 Larsen MV. 2012. Identification of acquired antimicrobial resistance genes. J Antimicrob
688 Chemother 67:2640–2644.
689 45. Carattoli A, Zankari E, Garcìa-Fernandez A, Larsen MV, Lund O, Villa L, Aarestrup FM,

690 Hasman H. 2014. PlasmidFinder and pMLST: in silico detection and typing of plasmids.
691 Antimicrob Agents Chemother 58:3895–3903.
692 46. Alikhan N-F, Petty NK, Ben Zakour NL, Beatson SA. 2011. BLAST Ring Image
693 Generator (BRIG): simple prokaryote genome comparisons. BMC Genomics 12:402.
694 47. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment
695 search tool. J Mol Biol 215:403–410.
696 48. M 'ikanatha NM, Sandt CH, Localio AR, Tewari D, Rankin SC, Whichard JM, Altekruse
697 SF, Lautenbach E, Folster JP, Russo A, Chiller TM, Reynolds SM, Mcdermott PF. 2010.
698 Multidrug-resistant *Salmonella* isolates from retail chicken meat compared with human
699 clinical isolates. Foodborne Pathog Dis 7:929–934.
700 49. Scheinberg J, Doores S, Cutter CN. 2013. A microbiological comparison of poultry
701 products obtained from farmers' markets and supermarkets in Pennsylvania. J Food Saf
702 33:259–264.
703 50. Gymoese P, Sørensen G, Litrup E, Olsen JE, Nielsen EM, Torpdahl M. 2017.
704 Investigation of outbreaks of *Salmonella enterica* serovar Typhimurium and its
705 monophasic variants using whole-genome sequencing, Denmark. Emerg Infect Dis
706 23:1631–1639.
707 51. Ashton PM, Peters T, Ameh L, McAleer R, Petrie S, Nair S, Muscat I, de Pinna E,
708 Dallmand T. 2015. Whole genome sequencing for the retrospective investigation of an
709 outbreak of *Salmonella* Typhimurium DT 8. PLoS Curr 7.
710 52. Zankari E, Hasman H, Kaas RS, Seyfarth AM, Agerso Y, Lund O, Larsen M V.,
711 Aarestrup FM. 2013. Genotyping using whole-genome sequencing is a realistic alternative
712 to surveillance based on phenotypic antimicrobial susceptibility testing. J Antimicrob

713 Chemother 68:771–777.

714 53. Tyson GH, Li C, Ayers S, McDermott PF, Zhao S. 2016. Using whole-genome
715 sequencing to determine appropriate streptomycin epidemiological cutoffs for *Salmonella*
716 and *Escherichia coli*. FEMS Microbiol Lett 363.

717 54. Chen S, Zhao S, White DG, Schroeder CM, Lu R, Yang H, McDermott PF, Ayers S,
718 Meng J. 2004. Characterization of multiple-antimicrobial-resistant *Salmonella* serovars
719 isolated from retail meats. Appl Environ Microbiol 70:1–7.

720 55. Folster JP, Pecic G, Rickert R, Taylor J, Zhao S, Fedorka-Cray PJ, Whichard J,
721 McDermott P. 2012. Characterization of multidrug-resistant *Salmonella enterica* serovar
722 Heidelberg from a ground turkey-associated outbreak in the United States in 2011.
723 Antimicrob Agents Chemother 56:3465–3466.

724 56. Patchanee P, Zewde BM, Tadesse DA, Hoet A, Gebreyes WA. 2008. Characterization of
725 multidrug-resistant *Salmonella enterica* serovar Heidelberg isolated from humans and
726 animals. Foodborne Pathog Dis 5:839–851.

727 57. Han J, Lynne AM, David DE, Tang H, Xu J, Nayak R, Kaldhone P, Logue CM, Foley SL.
728 2012. DNA sequence analysis of plasmids from multidrug resistant *Salmonella enterica*
729 serotype Heidelberg isolates. PLoS One 7:e51160.

730 58. Rehman MA, Yin X, Persaud-Lachhman MG, Diarra MS. 2017. First detection of a
731 fosfomycin resistance gene, *fosA7*, in *Salmonella enterica* serovar Heidelberg isolated
732 from broiler chickens. Antimicrob Agents Chemother 61:e00410–e00417.

733 59. Folster JP, Pecic G, Singh A, Duval B, Rickert R, Ayers S, Abbott J, Mcglinchey B,
734 Bauer-Turpin J, Haro J, Hise K, Zhao S, Fedorka-Cray PJ, Whichard J, Mcdermott PF.
735 2012. Characterization of extended-spectrum cephalosporin–resistant *Salmonella enterica*

736 serovar Heidelberg isolated from food animals, retail meat, and humans in the United
737 States 2009. Foodborne Pathog Dis 9:638–645.
738 60. Boyd DA, Peters GA, Ng L-K, Mulvey MR. 2000. Partial characterization of a genomic
739 island associated with the multidrug resistance region of *Salmonella enterica*
740 Typhymurium DT104. FEMS Microbiol Lett 189:285–291.
741 61. Boyd D, Peters GA, Cloeckaert A, Boumedine KS, Chaslus-Dancla E, Imberechts H,
742 Mulvey MR. 2001. Complete nucleotide sequence of a 43-kilobase genomic island
743 associated with the multidrug resistance region of *Salmonella enterica* serovar
744 Typhimurium DT104 and its identification in phage type DT120 and serovar Agona. J
745 Bacteriol 183:5725–5732.
746 62. Gray JT, Hungerford LL, Fedorka-Cray PJ, Headrick ML. 2004. Extended-spectrum-
747 cephalosporin resistance in *Salmonella enterica* isolates of animal origin. Antimicrob
748 Agents Chemother 48:3179–3181.
749 63. Frech G, Kehrenberg C, Schwarz S. 2003. Resistance phenotypes and genotypes of
750 multiresistant *Salmonella enterica* subsp. *enterica* serovar Typhimurium var. Copenhagen
751 isolates from animal sources. J Antimicrob Chemother 51:180–182.
752 64. Gebreyes WA, Altier C. 2002. Molecular characterization of multidrug-resistant
753 *Salmonella enterica* subsp. *enterica* serovar Typhimurium isolates from swine. J Clin
754 Microbiol 40:2813–2822.
755
756
757
758

Figure Legends

Figure 1. Phylogenetic relationships of *S.* Heidelberg isolates. Maximum likelihood phylogenetic tree of *S.* Heidelberg retail meat (diamonds) and clinical (circles) isolates generated by PhyML v3.1.1 (40) using the GTR + gamma model and 1,000 bootstrap replicates. Isolates SH-03 through SH-06, SH-11, and SH-12 were previously determined to be part of an outbreak (outbreak A); isolates SH-09 and SH-10 were previously determined to be part of a separate outbreak (outbreak B). The SNP alignment file produced by SNVPhyl v1.0.1 (39) contained 35 SNPs and served as the input for PhyML. The closed chromosome of clinical isolate SH-04 was used as the reference genome (grey arrow). Bootstrap values above 50 are included on the tree.

Figure 2. Phylogenetic relationships of *S.* Typhimurium var. O5- subset 1 isolates. Maximum likelihood phylogenetic tree of *S.* Typhimurium var. O5- subset 1 retail meat (diamond) and clinical (circles) isolates constructed by PhyML v3.1.1 (40) using the GTR + gamma model and 1,000 bootstrap replicates. The SNP alignment file produced by SNVPhyl v1.0.1 (39) served as the input for PhyML and contained 482 SNPs. The closed chromosome of clinical isolate SC-09 was used as the reference genome (grey arrow). Bootstrap values above 65 are included on the tree.

Figure 3. Phylogenetic relationships of *S.* Typhimurium var. O5- subset 2 isolates. Maximum likelihood phylogenetic tree of *S.* Typhimurium var. O5- subset 2 retail meat (diamonds) and clinical (circles) isolates constructed by PhyML v3.1.1 (40) using the GTR + gamma model and 1,000 bootstrap replicates. The SNP alignment file produced by SNVPhyl v1.0.1 (39) served as the input for PhyML and contained 225 SNPs. The closed chromosome of

clinical isolate SC-31 was used as the reference genome (grey arrow). Bootstrap values above 70 are included on the tree.

Figure 4. Identification of genetic resistance determinants in each subset. (A) *S.* Heidelberg subset. (B) *S.* Typhimurium var. O5- subset 1. (C) *S.* Typhimurium var. O5- subset 2. Each column represents the resistance genes identified in each isolate's assembled genome, with the exception of the last column in each graph, which represents the genes identified in the chromosome and plasmid(s) of each isolate sequenced by PacBio: SH-04 in (A), SC-09 in (B), and SC-31 in (C). A combination of BARRGD (Accession number PRJNA313047; accessed April 2018) and ResFinder 3.0 (44) databases were used to identify genes. Solid-colored genes were present at > 99.3% nucleotide identity and 100% query coverage. Pattern-colored genes were present at > 99.3% nucleotide identity and > 63% query coverage. Each isolate's phenotypic drug resistance profile is noted along the bottom. AMP=ampicillin; AMC=amoxicillin-clavulanic acid; TIO=ceftiofur; AXO=ceftriaxone; FOX=cefoxitin; CHL=chloramphenicol; GEN=gentamicin; STR=streptomycin; FIS=sulfisoxazole; TET=tetracycline.

Figure 5. Chromosomally carried AMR in *S.* Typhimurium var. O5- subset 1. Alignment of the SC-09 chromosome, closed by PacBio sequencing, against the assembled genomes of SC-01 through SC-08 and SC-10 through SC-28 using BRIG (46). Each colored concentric ring represents one assembled genome aligning to the reference genome. The outermost ring represents open reading frames (ORFs) within the SC-09 chromosome and AMR gene locations are indicated.

Figure 6. Comparative plasmid analysis in the *S.* Heidelberg subset. (A) Alignment of pSH-04-1 (IncI1-alpha), closed by PacBio sequencing, against the assembled genomes of SH-01 through SH-03 and SH-05 through SH-12 using BRIG (46). (B) Alignment of pSH-04-2 (IncX1), closed by PacBio sequencing, against the assembled genomes of SH-01 through SH-03 and SH-05 through SH-12. In (A) and (B), each colored concentric ring represents one assembled genome aligning to pSH-04-1 or pSH-04-2. The outermost ring in each panel represents ORFs in pSH-04-1 or pSH-04-2; the AMR annotations are included.

Figure 7. Comparative plasmid analysis in *S.* Typhimurium var. O5- subset 2. (A) Alignment of pSC-31-1 (IncI1-alpha), closed by PacBio sequencing, against the assembled genomes of SC-29, SC-30, and SC-32 through SC-38 using BRIG (46). The assembled genomes of SC-34 through SC-38 did not align to the reference plasmid. (B) Alignment of pSC-31-2 (IncA/C2), closed by PacBio sequencing, against the assembled genomes of SC-29, SC-30, and SC-32 through SC-38. (C) Alignment of pSC-31-3 (ColpVC), closed by Illumina sequencing, against the assembled genomes of SC-29, SC-30, and SC-32 through SC-38. In (A-C), each colored concentric ring represents one assembled genome aligning to either pSC-31-1, pSC-31-2, or pSC-31-3. The outermost ring in each panel represents ORFs in each plasmid; the AMR annotations are included.

828 **Tables**

829 **Table 1. Metadata of *S.* Heidelberg and *S.* Typhimurium var. O5- isolates sequenced in this**

830 **study.**

PSU Isolate Identifier[a]	State	Date	Source	PFGE Primary Pattern[b]	PFGE Secondary Pattern[c]	Phenotypic AMR Profile[d]
***S.* Heidelberg subset**						
SH-01	PA	2013	Ground Turkey	JF6X01.0058	JF6A26.0076	AMP, GEN, STR, TET
SH-02	PA	2013	Ground Turkey	JF6X01.0058	JF6A26.0076	AMP, GEN, STR, TET
SH-03[e]	PA	2011	Human	JF6X01.0058	JF6A26.0076	AMP, GEN, STR, TET
SH-04[e]	PA	2011	Human	JF6X01.0058	JF6A26.0076	AMP, GEN, STR, TET
SH-05[e]	PA	2011	Human	JF6X01.0058	JF6A26.0076	AMP, GEN, STR, TET
SH-06[e]	PA	2011	Human	JF6X01.0058	JF6A26.0076	AMP, GEN, STR, TET
SH-07	PA	2012	Human	JF6X01.0058	JF6A26.0076	AMP, GEN, STR, TET
SH-08	PA	2013	Human	JF6X01.0058	JF6A26.0076	AMP, GEN, STR, TET
SH-09[f]	PA	2010	Human	JF6X01.0058	-	AMP, GEN, STR, TET
SH-10[f]	PA	2010	Human	JF6X01.0058	-	AMP, GEN, STR, TET
SH-11[e]	PA	2011	Human	JF6X01.0058	JF6A26.0076	AMP, GEN, TET
SH-12[e]	PA	2011	Human	JF6X01.0058	JF6A26.0076	AMP, GEN, TET
***S.* Typhimurium var. O5- subset 1**						
SC-01	PA	2013	Pork Chop	JPXX01.0018	JPXA26.0156	AMP, CHL, STR, FIS, TET
SC-02	PA	2009	Human	JPXX01.0018	JPXA26.0156	AMP, CHL, STR, FIS, TET
SC-03	PA	2009	Human	JPXX01.0018	JPXA26.0156	AMP, CHL, STR, FIS, TET

SC-04	PA	2009	Human	JPXX01.0018	JPXA26.0156	AMP, CHL, STR, FIS, TET
SC-05	PA	2010	Human	JPXX01.0018	JPXA26.0156	AMP, CHL, STR, FIS, TET
SC-06	PA	2010	Human	JPXX01.0018	JPXA26.0156	AMP, CHL, STR, FIS, TET
SC-07	PA	2010	Human	JPXX01.0018	JPXA26.0156	AMP, CHL, STR, FIS, TET
SC-08	PA	2010	Human	JPXX01.0018	JPXA26.0156	AMP, CHL, STR, FIS, TET
SC-09	PA	2010	Human	JPXX01.0018	JPXA26.0156	AMP, CHL, STR, FIS, TET
SC-10	PA	2010	Human	JPXX01.0018	JPXA26.0156	AMP, CHL, STR, FIS, TET
SC-11	PA	2010	Human	JPXX01.0018	JPXA26.0156	AMP, CHL, STR, FIS, TET
SC-12	PA	2010	Human	JPXX01.0018	JPXA26.0156	AMP, CHL, STR, FIS, TET
SC-13	OH	2010	Human	JPXX01.0018	JPXA26.0156	AMP, CHL, STR, FIS, TET
SC-14	PA	2010	Human	JPXX01.0018	JPXA26.0156	AMP, CHL, STR, FIS, TET
SC-15	PA	2011	Human	JPXX01.0018	JPXA26.0156	AMP, CHL, STR, FIS, TET
SC-16	PA	2012	Human	JPXX01.0018	JPXA26.0156	AMP, CHL, STR, FIS, TET
SC-17	PA	2013	Human	JPXX01.0018	JPXA26.0156	AMP, CHL, STR, FIS, TET
SC-18	NJ	2010	Human	JPXX01.0018	JPXA26.0634	AMP, CHL, STR, FIS, TET
SC-19	PA	2011	Human	JPXX01.0018	-	AMP, CHL, STR, FIS, TET
SC-20	PA	2011	Human	JPXX01.0018	JPXA26.0003	AMP, CHL, STR, FIS, TET
SC-21	PA	2011	Human	JPXX01.0018	JPXA26.0155	AMP, CHL, STR, FIS, TET
SC-22	PA	2012	Human	JPXX01.0018	JPXA26.0490	AMP, CHL, STR, FIS, TET
SC-23	PA	2012	Human	JPXX01.0018	JPXA26.0490	AMP, CHL, STR, FIS, TET
SC-24	PA	2013	Human	JPXX01.0018	-	AMP, CHL, STR, FIS, TET
SC-25	PA	2013	Human	JPXX01.0018	-	AMP, CHL, STR, FIS, TET

SC-26	PA	2013	Human	JPXX01.0018	-	AMP, CHL, STR, FIS, TET
SC-27	PA	2014	Human	JPXX01.0018	-	AMP, CHL, STR, FIS, TET
SC-28[g]	PA	2014	Human	JPXX01.0018	-	AMP, CHL, STR, FIS, TET
***S.* Typhimurium var. O5- subset 2**						
SC-29	PA	2012	Chicken Breast	JPXX01.1283	JPXA26.0397	AMP, AMC, TIO, AXO, FOX, FIS, TET
SC-30	PA	2010	Human	JPXX01.1283	JPXA26.0397	AMP, AMC, TIO, AXO, FOX, FIS, TET
SC-31	PA	2010	Human	JPXX01.1283	JPXA26.0397	AMP, AMC, TIO, AXO, FOX, FIS, TET
SC-32	PA	2011	Human	JPXX01.1283	JPXA26.0397	AMP, AMC, TIO, AXO, FOX, FIS, TET
SC-33	PA	2012	Human	JPXX01.1283	JPXA26.0786	AMP, AMC, TIO, AXO, FOX, FIS, TET
SC-34	PA	2012	Human	JPXX01.1283	JPXA26.0667	AMP, AMC, TIO, AXO, FOX, FIS, TET
SC-35	PA	2011	Chicken Breast	JPXX01.1283	JPXA26.0397	FIS, TET
SC-36	PA	2012	Chicken Breast	JPXX01.1283	JPXA26.0397	FIS, TET
SC-37	PA	2013	Chicken Breast	JPXX01.1283	JPXA26.0397	FIS, TET
SC-38	PA	2010	Human	JPXX01.1283	JPXA26.0397	FIS, TET

[a]Strain identifier; "SH" indicates *S.* Heidelberg; "SC" indicates *S.* Copenhagen (more commonly referred to as *S.* Typhimurium var. O5-)

[b]XbaI enzyme restriction pattern

[c]BlnI enzyme restriction pattern

[d]AMP=Ampicillin; AMC=Amoxicillin-Clavulanic Acid; TIO=Ceftiofur; AXO=Ceftriaxone; FOX=Cefoxitin; CHL=Chloramphenicol; GEN=Gentamicin; STR=Streptomycin; FIS=Sulfisoxazole; TET=Tetracycline

[e]Isolate belongs to outbreak A

[f]Isolate belongs to outbreak B

[g]Serotyped as *S.* Typhimurium by agglutination and *S.* Typhimurium var. O5- by SeqSero

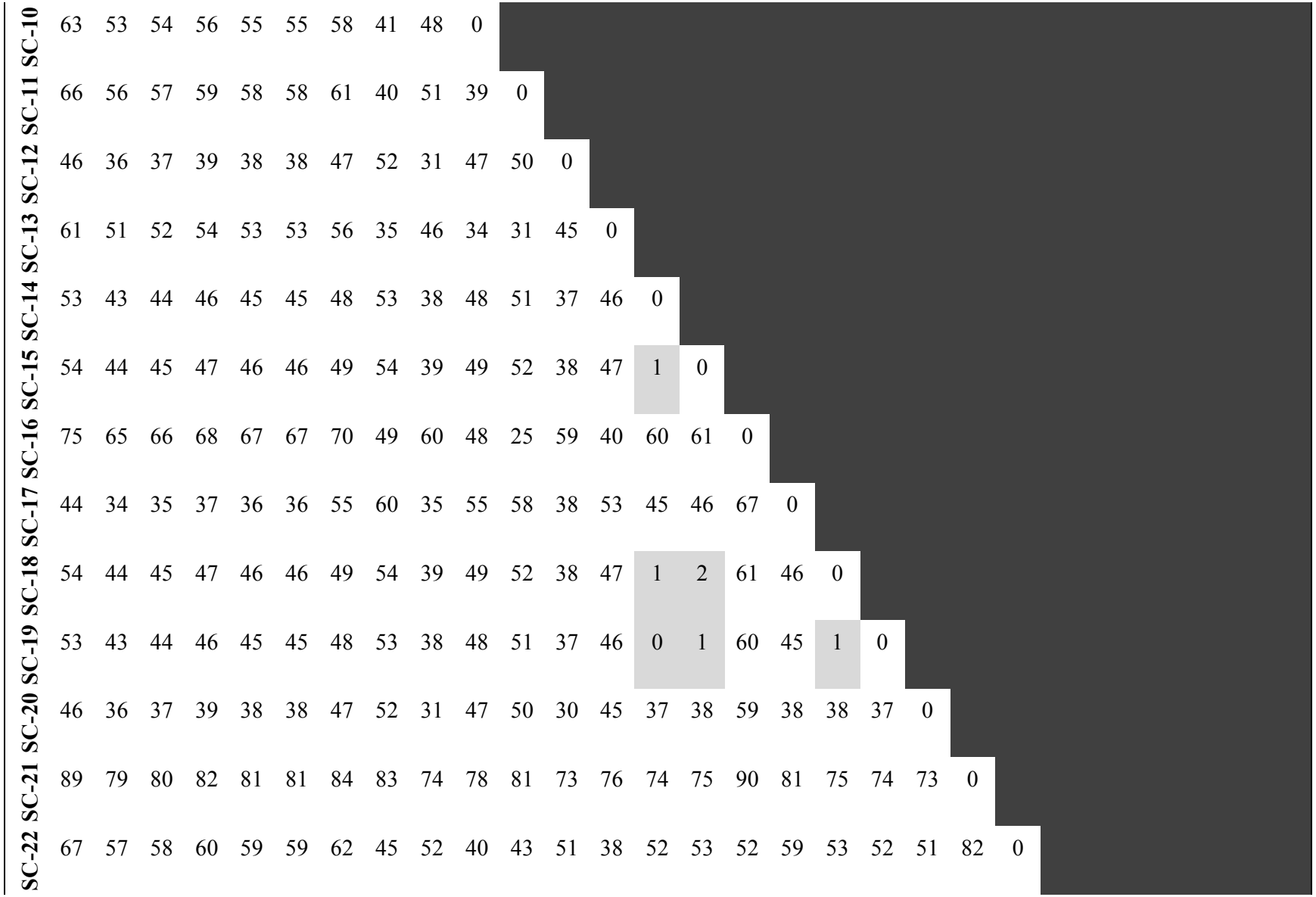
Table 2. Pairwise SNP distances between *S.* Heidelberg isolates.

	Retail Meat		Outbreak A Clinical				Clinical		Outbreak B Clinical		Outbreak A Clinical	
	SH-01	**SH-02**	**SH-03**	**SH-04**	**SH-05**	**SH-06**	**SH-07**	**SH-08**	**SH-09**	**SH-10**	**SH-11**	**SH-12**
SH-01	0											
SH-02	1	0										
SH-03	7	6	0									
SH-04	9	8	4	0								
SH-05	8	7	3	5	0							
SH-06	12	11	7	9	8	0						
SH-07	10	9	7	9	8	12	0					

SH-08	10	9	5	7	6	10	10	0				
SH-09	10	9	5	7	6	10	10	8	0			
SH-10	10	9	5	7	6	10	10	8	0	0		
SH-11	8	7	3	5	4	8	8	6	6	6	0	
SH-12	8	7	3	5	4	8	8	6	6	6	4	0

858 **Table 3. Pairwise SNP distances between *S.* Typhimurium var. O5- subset 1 isolates.**

		Retail Meat	Clinical																								
		SC-01	SC-02	SC-03	SC-04	SC-05	SC-06	SC-07	SC-08	SC-09	SC-10	SC-11	SC-12	SC-13	SC-14	SC-15	SC-16	SC-17	SC-18	SC-19	SC-20	SC-21	SC-22	SC-23	SC-24	SC-25	SC-26	SC-27	SC-28
Retail Meat	SC-01	0																											
Clinical	SC-02	42	0																										
Clinical	SC-03	41	15	0																									
Clinical	SC-04	43	17	2[a]	0																								
Clinical	SC-05	44	34	35	37	0																							
Clinical	SC-06	44	34	35	37	0	0																						
Clinical	SC-07	63	53	54	56	55	55	0																					
Clinical	SC-08	68	58	59	61	60	60	63	0																				
Clinical	SC-09	43	33	34	36	35	35	48	53	0																			



	SC-10	SC-11	SC-12	SC-13	SC-14	SC-15	SC-16	SC-17	SC-18	SC-19	SC-20	SC-21	SC-22
SC-10	0												
SC-11	39	0											
SC-12	47	50	0										
SC-13	34	31	45	0									
SC-14	48	51	37	46	0								
SC-15	49	52	38	47	1	0							
SC-16	48	25	59	40	60	61	0						
SC-17	55	58	38	53	45	46	67	0					
SC-18	49	52	38	47	1	2	61	46	0				
SC-19	48	51	37	46	0	1	60	45	1	0			
SC-20	47	50	30	45	37	38	59	38	38	37	0		
SC-21	78	81	73	76	74	75	90	81	75	74	73	0	
SC-22	40	43	51	38	52	53	52	59	53	52	51	82	0

SC-23	66	56	57	59	58	58	61	44	51	39	42	50	37	51	52	51	58	52	51	50	81	5	0						
SC-24	68	58	59	61	60	60	63	10	53	41	40	52	35	53	54	49	60	54	53	52	83	45	44	0					
SC-25	69	59	60	62	61	61	64	63	54	58	61	53	56	54	55	70	61	55	54	53	80	62	61	63	0				
SC-26	69	59	60	62	61	61	64	39	54	42	41	53	36	54	55	50	61	55	54	53	84	46	45	39	64	0			
SC-27	69	59	60	62	61	61	64	47	54	42	45	53	40	54	55	54	61	55	54	53	84	8	7	47	64	48	0		
SC-28	96	86	87	89	88	88	91	90	81	85	88	80	83	81	82	97	88	82	81	80	[b]	89	88	90	87	91	91	0	

[a]Light grey shading denotes pairwise SNPs distances of < 20 between clinical isolates

[b]Pairwise SNP distance of 105

861 **Table 4. Pairwise SNP distances between *S.* Typhimurium var. O5- subset 2 isolates.**

	Retail Meat	Clinical					Retail Meat			Clinical
	SC-29	**SC-30**	**SC-31**	**SC-32**	**SC-33**	**SC-34**	**SC-35**	**SC-36**	**SC-37**	**SC-38**
SC-29	0									
SC-30	58	0								
SC-31	61	57	0							
SC-32	27	57	60	0						
SC-33	72	66	73	73	0					
SC-34	71	67	70	70	83	0				
SC-35	28	58	61	21	74	71	0			
SC-36	31	61	64	24	77	74	17	0		
SC-37	35	65	68	28	81	78	21	24	0	
SC-38	24	52	55	23	68	65	24	27	31	0

862

863

864 **Table 5. Overall correlation between AMR phenotype and genotype across all subsets.**

Antibiotic Class	Phenotype Resistant	Genotype Resistant	Phenotype Sensitive	Genotype Sensitive	Correlation
Aminoglycosides					
GEN	12	12	38	38	100%
STR	38	40	12	10	95%
β-lactam	46	46	4	4	100%
Chloramphenicol	28	28	22	22	100%
Fosfomycin	[a]	12	[a]	38	[a]
Sulfonamide	38	38	12	12	100%
Tetracycline	50	50	0	0	100%

865 [a]Isolates were phenotypically not tested for resistance to the drug fosfomycin

866

867

868

869

870

871

Figures and Figure Legends

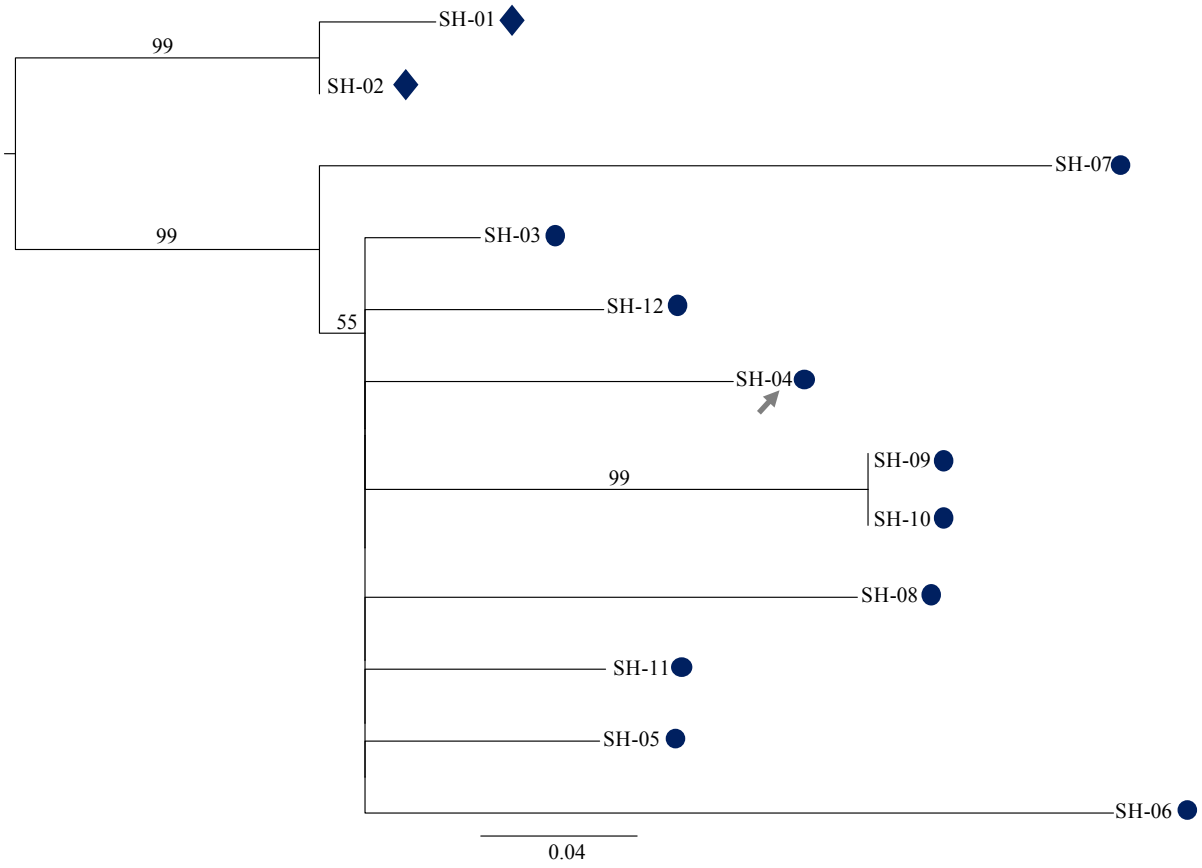


Figure 1. Phylogenetic relationships of *S.* Heidelberg isolates. Maximum likelihood phylogenetic tree of *S.* Heidelberg retail meat (diamonds) and clinical (circles) isolates generated by PhyML v3.1.1 (40) using the GTR + gamma model and 1,000 bootstrap replicates. Isolates SH-03 through SH-06, SH-11, and SH-12 were previously determined to be part of an outbreak (outbreak A); isolates SH-09 and SH-10 were previously determined to be part of a separate outbreak (outbreak B). The SNP alignment file produced by SNVPhyl v1.0.1 (39) contained 35 SNPs and served as the input for PhyML. The closed chromosome of clinical isolate SH-04 was used as the reference genome (grey arrow). Bootstrap values above 50 are included on the tree.

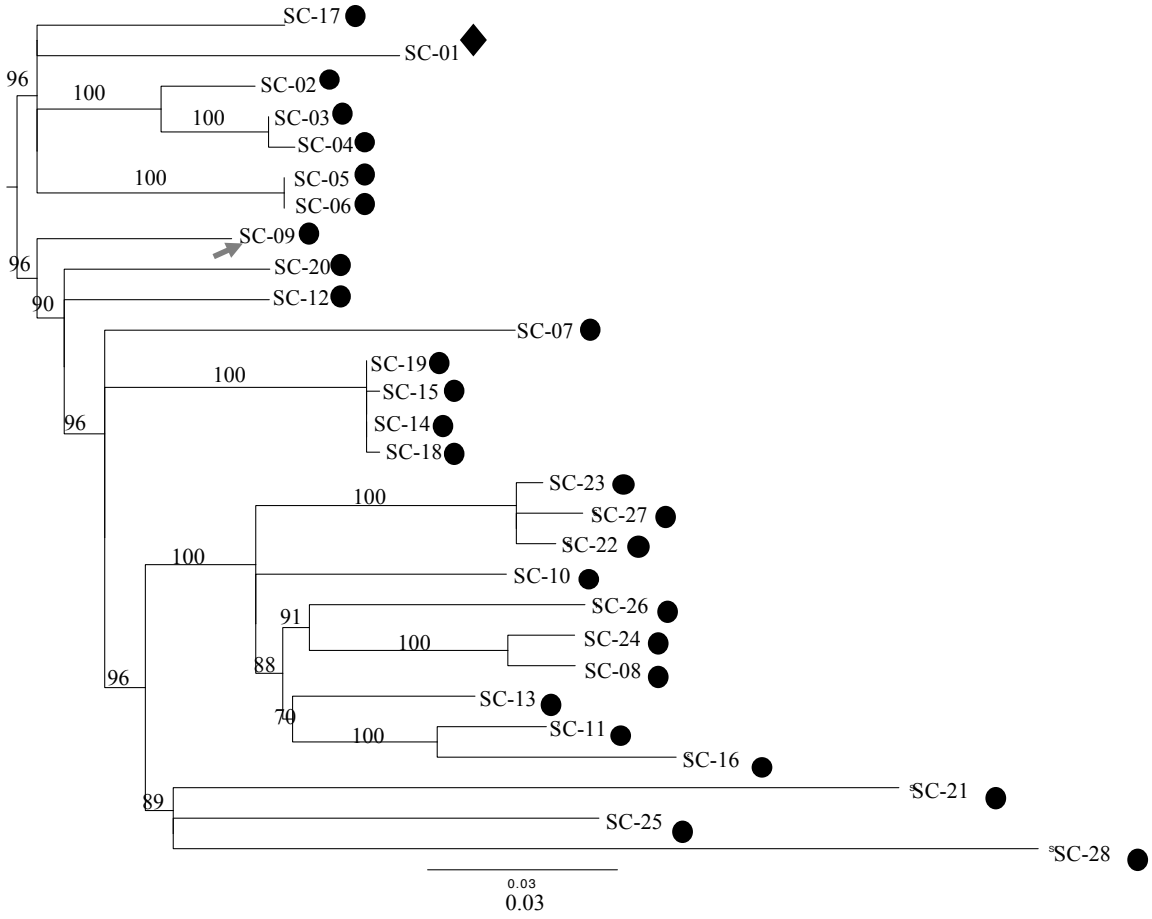


Figure 2. Phylogenetic relationships of *S.* Typhimurium var. O5- subset 1 isolates. Maximum likelihood phylogenetic tree of *S.* Typhimurium var. O5- subset 1 retail meat (diamond) and clinical (circles) isolates constructed by PhyML v3.1.1 (40) using the GTR + gamma model and 1,000 bootstrap replicates. The SNP alignment file produced by SNVPhyl v1.0.1 (39) served as the input for PhyML and contained 482 SNPs. The closed chromosome of clinical isolate SC-09 was used as the reference genome (grey arrow). Bootstrap values above 65 are included on the tree.

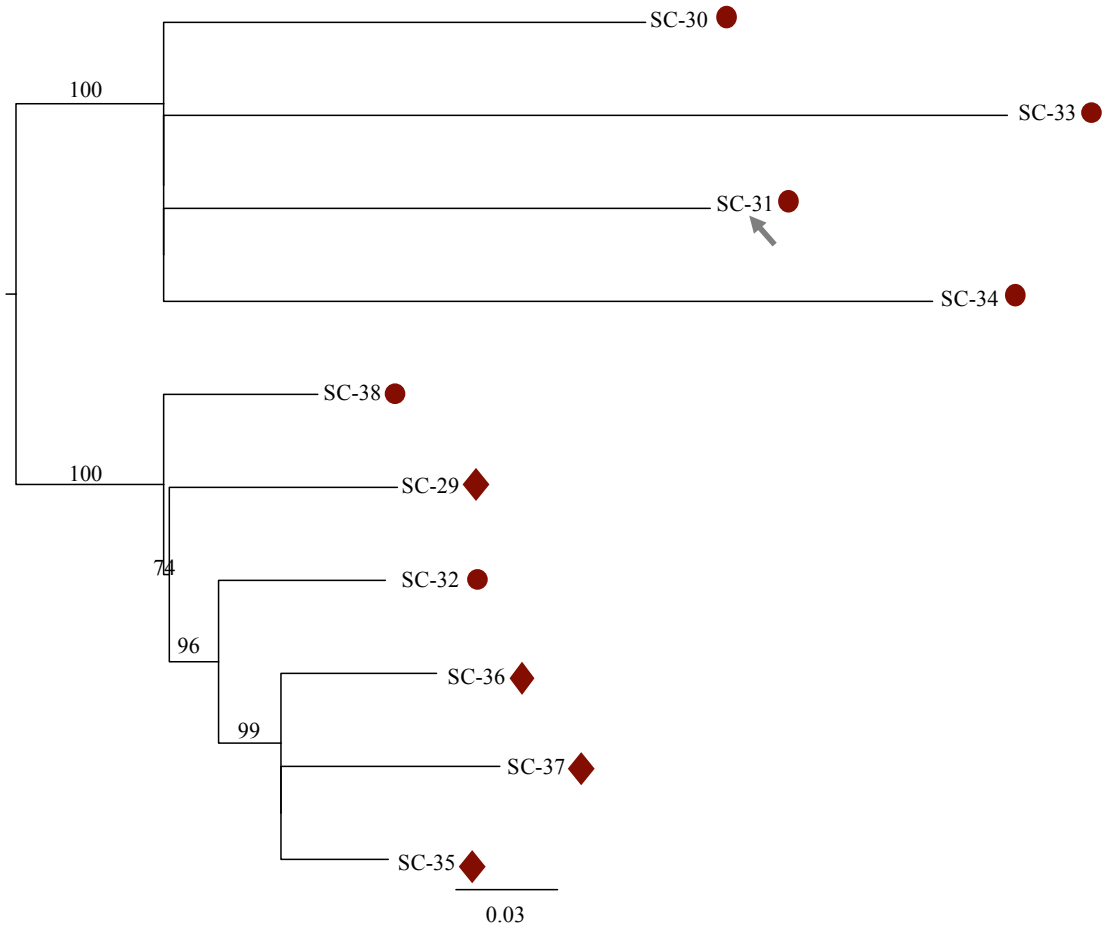


Figure 3. Phylogenetic relationships of *S.* Typhimurium var. O5- subset 2 isolates. Maximum likelihood phylogenetic tree of *S.* Typhimurium var. O5- subset 2 retail meat (diamonds) and clinical (circles) isolates constructed by PhyML v3.1.1 (40) using the GTR + gamma model and 1,000 bootstrap replicates. The SNP alignment file produced by SNVPhyl v1.0.1 (39) served as the input for PhyML and contained 225 SNPs. The closed chromosome of clinical isolate SC-31 was used as the reference genome (grey arrow). Bootstrap values above 70 are included on the tree.

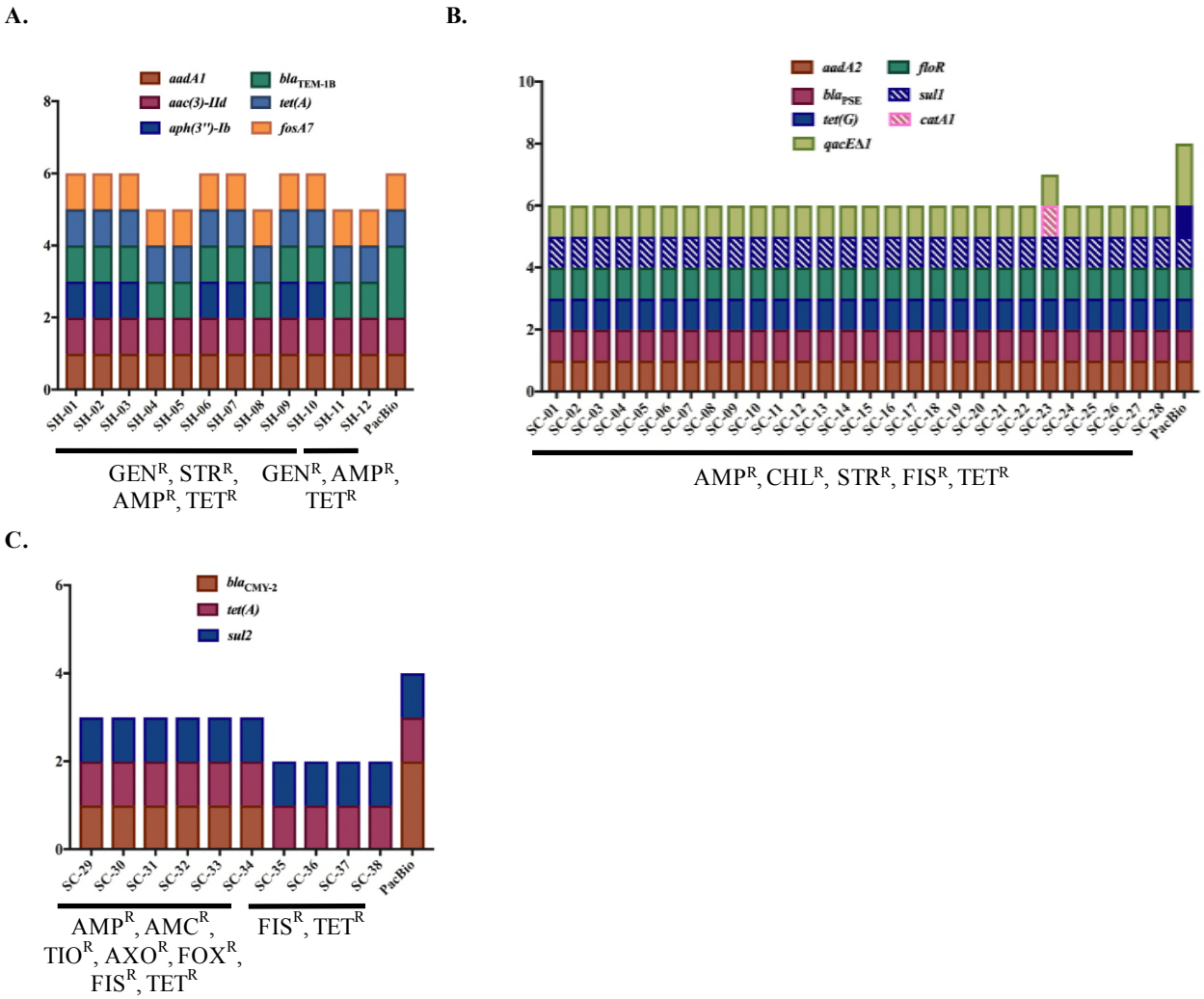


Figure 4. Identification of genetic resistance determinants in each subset. (A) *S.* Heidelberg subset. (B) *S.* Typhimurium var. O5- subset 1. (C) *S.* Typhimurium var. O5- subset 2. Each column represents the resistance genes identified in each isolate's assembled genome, with the exception of the last column in each graph, which represents the genes identified in the chromosome and plasmid(s) of each isolate sequenced by PacBio: SH-04 in (A), SC-09 in (B), and SC-31 in (C). A combination of BARRGD (Accession number PRJNA313047; accessed April 2018) and ResFinder 3.0 (44) databases were used to identify genes. Solid-colored genes were present at > 99.3% nucleotide identity and 100% query coverage. Pattern-colored genes

were present at > 99.3% nucleotide identity and > 63% query coverage. Each isolate's phenotypic drug resistance profile is noted along the bottom. AMP=ampicillin; AMC=amoxicillin-clavulanic acid; TIO=ceftiofur; AXO=ceftriaxone; FOX=cefoxitin; CHL=chloramphenicol; GEN=gentamicin; STR=streptomycin; FIS=sulfisoxazole; TET=tetracycline.

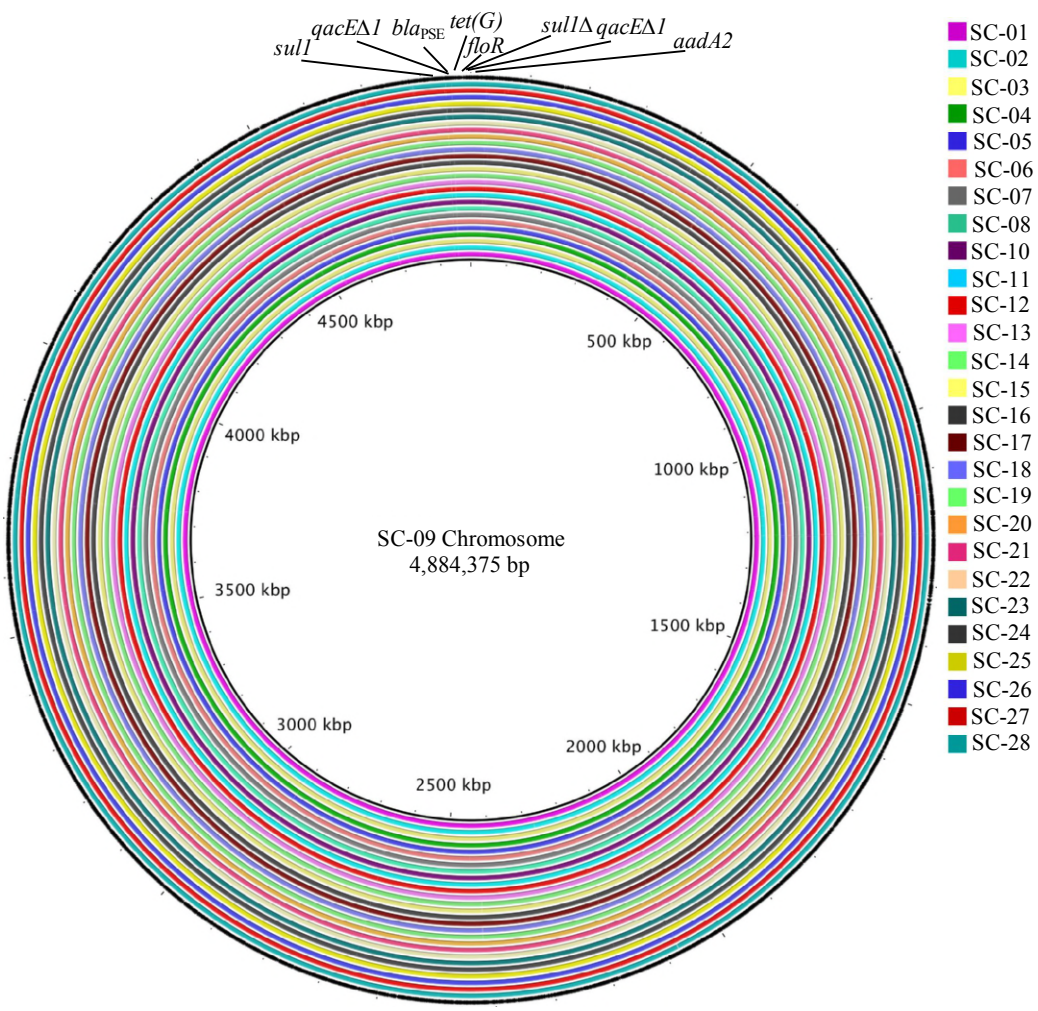


Figure 5. Chromosomally carried AMR in *S.* Typhimurium var. O5- subset 1. Alignment of the SC-09 chromosome, closed by PacBio sequencing, against the assembled genomes of SC-01 through SC-08 and SC-10 through SC-28 using BRIG (46). Each colored concentric ring represents one assembled genome aligning to the reference genome. The outermost ring represents open reading frames (ORFs) within the SC-09 chromosome and AMR gene locations are indicated.

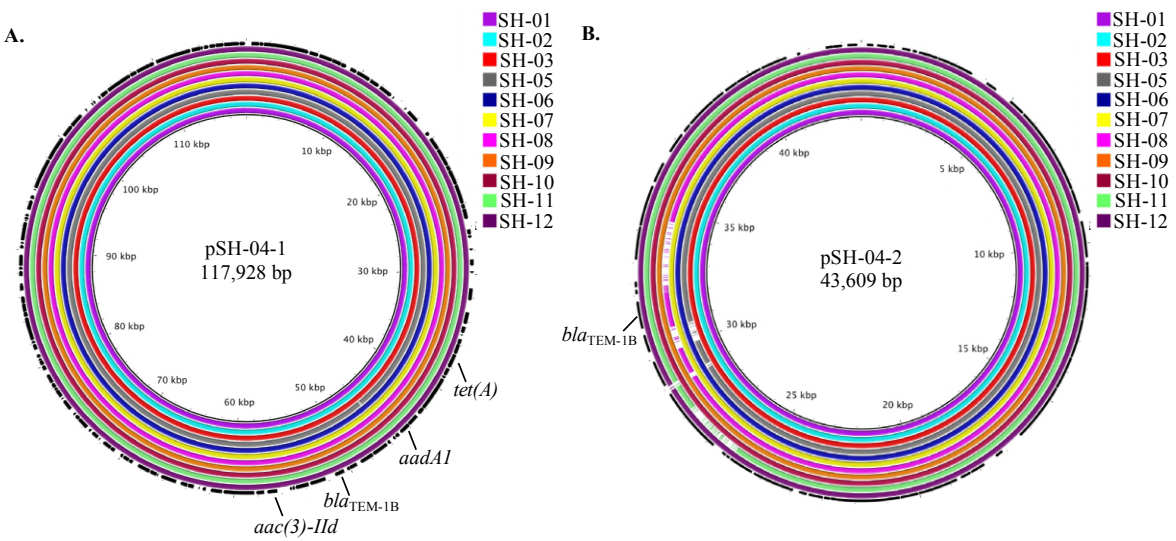


Figure 6. Comparative plasmid analysis in the *S.* Heidelberg subset. (A) Alignment of pSH-04-1 (IncI1-alpha), closed by PacBio sequencing, against the assembled genomes of SH-01 through SH-03 and SH-05 through SH-12 using BRIG (46). (B) Alignment of pSH-04-2 (IncX1), closed by PacBio sequencing, against the assembled genomes of SH-01 through SH-03 and SH-05 through SH-12. In (A) and (B), each colored concentric ring represents one assembled genome aligning to pSH-04-1 or pSH-04-2. The outermost ring in each panel represents ORFs in pSH-04-1 or pSH-04-2; the AMR annotations are included.

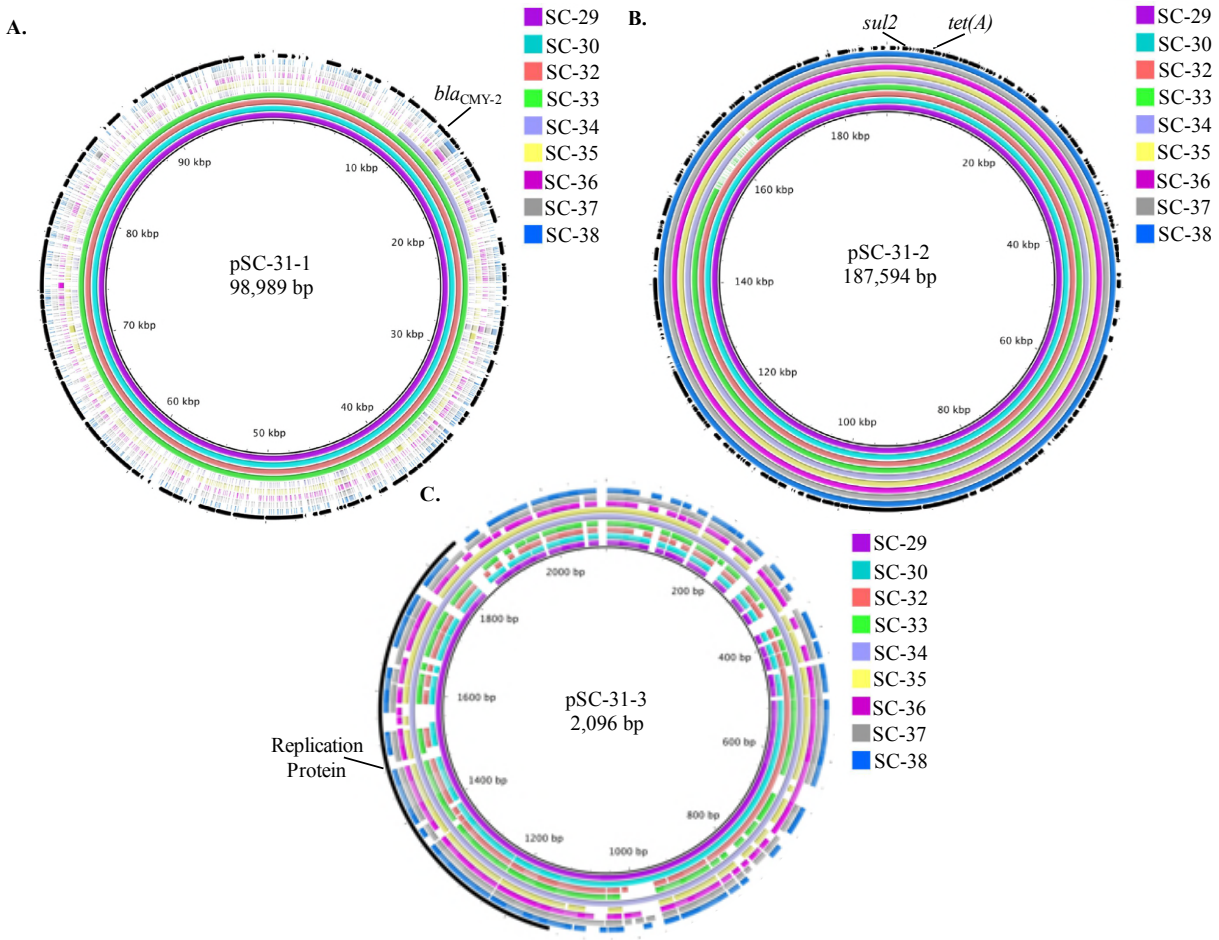


Figure 7. Comparative plasmid analysis in *S.* Typhimurium var. O5- subset 2. (A) Alignment of pSC-31-1 (IncI1-alpha), closed by PacBio sequencing, against the assembled genomes of SC-29, SC-30, and SC-32 through SC-38 using BRIG (46). The assembled genomes of SC-34 through SC-38 did not align to the reference plasmid. (B) Alignment of pSC-31-2 (IncA/C2), closed by PacBio sequencing, against the assembled genomes of SC-29, SC-30, and SC-32 through SC-38. (C) Alignment of pSC-31-3 (ColpVC), closed by Illumina sequencing, against the assembled genomes of SC-29, SC-30, and SC-32 through SC-38. In (A-C), each colored concentric ring represents one assembled genome aligning to either pSC-31-1, pSC-31-2, or pSC-31-3. The outermost ring in each panel represents ORFs in each plasmid; the AMR annotations are included.