Long-Read Sequencing Reveals Evolution and Acquisition of Antimicrobial Resistance and Virulence Genes in Salmonella enterica

Cong Li1*, Gregory H. Tyson1†, Chih-Hao Hsu1, Lucas Harrison1, Errol Strain1, Thu-Thuy Tran1, Glenn E. Tillman2, Uday Dessai3, Patrick F. McDermott1 and Shaohua Zhao1*

1 Center for Veterinary Medicine, United States Food and Drug Administration, Laurel, MD, United States, 2 Food Safety and Inspection Service, United States Department of Agriculture, Athens, GA, United States, 3 Food Safety and Inspection Service, United States Department of Agriculture, Washington, DC, United States

Salmonella enterica is a significant and phylogenetically diverse zoonotic pathogen. To understand its genomic heterogeneity and antimicrobial resistance, we performed long-read sequencing on Salmonella isolated from retail meats and food animals. A collection of 134 multidrug-resistant isolates belonging to 33 serotypes were subjected to PacBio sequencing. One major locus of diversity among these isolates was the presence and orientation of Salmonella pathogenic islands (SPI), which varied across different serotypes but were largely conserved within individual serotypes. We also identified insertion of an IncQ resistance plasmid into the chromosome of fourteen strains of serotype I4,[5],12:i:– and the Salmonella genomic island 1 (SGI-1) in five serotypes. The presence of various SPIs, SGI-1 and integrated plasmids contributed significantly to the genomic variability and resulted in chromosomal resistance in 55.2% (74/134) of the study isolates. A total of 93.3% (125/134) of isolates carried at least one plasmid, with isolates carrying up to seven plasmids. We closed 233 plasmid sequences of thirteen replicon types, along with twelve hybrid plasmids. Some associations between Salmonella isolate source, serotype, and plasmid type were seen. For instance, IncX plasmids were more common in serotype Kentucky from retail chicken. Plasmids IncC and IncHI had on average more than five antimicrobial resistance genes, whereas in IncX, it was less than one per plasmid. Overall, 60% of multidrug resistance (MDR) strains that carried >3 AMR genes also carried >3 heavy metal resistance genes, raising the possibility of co-selection of antimicrobial resistance in the presence of heavy metals. We also found nine isolates representing four serotypes that carried virulence plasmids with the spv operon. Together, these data demonstrate the power of long-read sequencing to reveal genomic arrangements and integrated plasmids with a high level of resolution for tracking and comparing...
resistant strains from different sources. Additionally, the findings from this study will help expand the reference set of closed Salmonella genomes that can be used to improve genome assembly from short-read data commonly used in One Health antimicrobial resistance surveillance.

**Keywords:** Salmonella, multidrug resistance (MDR), plasmid, Salmonella genomic island (SGI), Salmonella pathogenicity island (SPI)

---

**INTRODUCTION**

Salmonella enterica is an important zoonotic pathogen that causes over 125 million illnesses in the United States each year (Scallan et al., 2011). The term enteric has classically subdivided into serotypes and over 2,600 serotypes have been identified thus far. While many serotypes may be capable of causing infections in humans and animals, a limited number of serotypes cause most human infections in the United States (Scallan et al., 2011). Recent advancements in whole-genome sequencing with WGS offer a unique opportunity to dissect and investigate Salmonella serotypes at the nucleotide level and to further our understanding of evolutionary changes. The main features associated with Salmonella evolution include acquisition and recombination of mobile genetic elements such as plasmids, transposons, integrons, and plasmids. Among others (Partridge et al., 2018), plasmids and plasmids among other factors provide important information to understand resistance, host and environmental adaptations, and sources of resistant Salmonella infections.

While most Salmonella infections are self-limiting, serious infections can require antimicrobial therapy (Tack et al., 2020). Antimicrobial resistance (AMR) is compromising clinical care and increasing healthcare costs and antibiotic resistance in Salmonella. AMR is typically attained through horizontal acquisition of antimicrobial resistance genes (ARGs). Although chromosomal mutations also contribute to resistance, ARGs are the most important source of resistance in Salmonella infections.

One way that Salmonella strains acquire ARGs is through plasmid acquisition (Emond-Rheault et al., 2020). Plasmids carry not only ARGs but also heavy metal and disinfectant resistance genes, which may contribute to co-selection for other ARGs (Vijayakumar and Hundle, 2019). The types of ARGs that Salmonella carry can vary considerably as they may include species-specific and conjugative plasmids, and conjugative plasmids found widely among Enterobacterales (Redondo-Salvo et al., 2021). Some plasmid types are highly associated with specific serotypes and sources of Salmonella (Zhou et al., 2020). These plasmid types provide important information for outbreak investigations and AMR source attribution. Traditionally, incompatibility plasmid types have been used to assign plasmid into different groups based on plasmid replication machinery (Carattoli, 2013). This approach has not been particularly useful for Salmonella plasmids and PlasmidSpades2.PLACNet software has helped to expand analyses of plasmid genomes derived from short-read sequencing data (de Curraize et al., 2020).

Characterization of plasmids and other resistance elements in Salmonella has been studied extensively by WGS. The use of short-read sequencing in conjunction with programs such as PlasmidSpades2.PLACNet software has helped to expand analyses of genomes derived from short-read sequencing data (de Curraize et al., 2020). There have been relatively few large-scale, long-read sequencing studies which yield more complete genomic information with higher resolution.

Aside from plasmids, ARGs also are commonly carried by chromosomally encoded SGI s and plasmids (SGIs). SGI-1 was first reported in Salmonella Typhimurium DT104 in 2001. It contained a 27 kb backbone plus a 15 kb complex with 25 heavy metal resistant genes, an integron with ARGs conferring resistance to five antimicrobial classes (Boyd et al., 2000). Different variant SGI-1 have been described with diversity in ARG profile and multidrug resistance (MDR) regions (Hall, 2010). Additional SGIs including SGI-0, SGI-2, SGI-3, and SGI-4 have been identified based on genomic structure and resistance gene content. Both SGI-0 and SGI-2 refer to the same location in S. Typhimurium DT104 and S. Enteritidis, while SGI-1 and SGI-3 are different (Branchu et al., 2019). SGIs are initially described as distinct SGI s, but they may share the same chromosomal location. SGI-3 and SGI-4 are considered the same SGI (Arai et al., 2019; Branchu et al., 2019). SGI-4 is known to carry ARGs, whereas SGI-3, which is otherwise similar, does not carry ARGs (Arai et al., 2019; Branchu et al., 2019). Together, these features of SGI-3 and SGI-4 are considered important for AMR determinants of mobile genetic elements and contribute to the genomic diversity found in Salmonella.

Salmonella pathogenicity islands (SPIs) play a pivotal role in Salmonella virulence (Hensel, 2004; Ryckewaert et al., 2009). Here, we describe the diversity of plasmids and SGIs among Salmonella. SPIs are described, with a diversity of ARG alleles in multidrug resistance classes (Boyd et al., 2000). Different variants of SGI-1 have been described with diversity in ARG profile and multidrug resistance (MDR) regions (Hall, 2010). Additional SGIs including SGI-0, SGI-2, SGI-3, and SGI-4 have been identified based on genomic structure and resistance gene content. Both SGI-0 and SGI-2 refer to the same location in S. Typhimurium DT104 and S. Enteritidis, while SGI-1 and SGI-3 are different (Branchu et al., 2019). SGI-4 is known to carry ARGs, whereas SGI-3, which is otherwise similar, does not carry ARGs (Arai et al., 2019; Branchu et al., 2019). Together, these features of SGI-3 and SGI-4 are considered important for AMR determinants of mobile genetic elements and contribute to the genomic diversity found in Salmonella.

The history of Salmonella epidemiology has relied on various features of categorized strains. Following biochemical profiling, serotyping has long been the basis of Salmonella strain typing and tracking. Later, plasmid profiling by electrophoresis and pulsed field gel electrophoresis (PFGE) was used. For AMR monitoring, in national programs such as the United States National Antimicrobial Resistance Monitoring System (NARMS), minimum inhibitory concentration (MIC) testing followed by multiplex PCR and conjugation assays. These commonly used methods to track resistance and genotype have been characterized by routine using short-read DNA sequencing chemistries. While this provides a comprehensive picture of strain relatedness and resistance carriage, a closed genomes
are needed to reveal the detailed gene arrangements and structural changes. In this report, we describe the use of PacBio long-read sequencing to characterize 33 isolates representing 33 Salmonella serotypes, isolated from a variety of food animals. This study helped us elucidate the genomic structure and location of virulence and resistance genes in their colocation on mobile DNA elements and now these traits are studied in Salmonella evolution. We also propose a new approach to simplify naming of GIs based on their genomic position.

MATERIALS AND METHODS

Isolate Sources and PacBio Sequencing

One hundred thirty-four isolates representing 33 serotypes were collected at slaughter from swine, turkey, beef, and pork products as well as cecal/gut samples collected at slaughter from swine, turkey, cattle, and chicken from 2016–2018 across 31 different states. Isolates were selected for Pacific Biosciences (PacBio) long-read sequencing to represent diverse resistance patterns including three pan-susceptible isolates of diverse serotypes and different NARMs sources.

For long-read sequencing, DNA libraries were prepared using 11 kb of template preparation protocol with SMRTbell template prep kit v1.0. Sequencing was performed using PacBio Bioscience technology on the Sequel platform with sequencing kit 3.0, as described previously (Tate et al., 2021).

Sequencing data are available in BioProject PRJNA292661. Isolate-level accession numbers are listed in Supplementary Table 1.

Resistance Gene and Plasmid Identification

Antimicrobial resistance genes, biocide resistance genes, and HMRGs were identified with the AMRFinder Plus version 3.8 (Feldgarden et al., 2019) and AMRFinder Plus virulence genes and ARGs outside the AMRFinder core genes were not reported due to their limited elevation in this Salmonella study.

To identify plasmid replicons, sequence files were used for PlasmidFinder with cutoffs of 50% identity and 60% length (Carattoli et al., 2014). The sequences of theavrRBCD operon was extracted from the plasmid pOU1115 carried by S. dublin (Strain Accession DQ115388) and local last analysis was performed to identify the presence of theavrRBCD operon.

Integrons were identified by a similar approach with PlasmidFinder being used to identify replicons in some cases. Last analysis was conducted to identify similarity with known plasmids (Table 2).

Salmonella Pathogenic Islands and Salmonella Genomic Island Identification

Sequences of 24 SPIs were downloaded from GenBank local database (Fookes et al., 2011; Hayward et al., 2014; Cheng et al., 2019) and the size of the SPIs ranged from 0.7 to 33.3 kb. the coding sequences (CDS) of thevirulence genes were used to construct a phylogeny with six of them (I 4, 5, 12, I 4, 5, 12) and two of them (I 4, 5, 12) and other of them (I 4, 5, 12) and others of them (I 4, 5, 12). The clade including S. Typhimurium DT104 (Boyd et al., 2000) was further analyzed to identify additional SGI variants identified by the existence of SGI-4 as discovered in the 341 strain (Accession DQ115388). A local last analysis with the same cutoffs was performed to identify the presence of this SGI-4.

Phylogenetic Tree

The program KSNP3.0 was used to generate single nucleotide polymorphisms (SNPs) from all subset of 44 complete chromosomes to represent all 33 serotypes and cases of chromosomal heterogeneity within serotypes (Gardner et al., 2015). Prior to KSNP generation, the sequence size was chosen by Kchooser included in KSNP3.0 (Gardner et al., 2015). The maximum likelihood phylogenetic tree was constructed by MEGA7.0 with 2500 bootstrap repetitions. The clade including S. Schwarzengrund was placed at the root based on established literature (Worley et al., 2018).

RESULTS

Presence of Salmonella Pathogenic Islands and Arrangement in the Chromosome

Assembly of long-read sequences produced circular closed chromosomes of 4,923,868 bp for S. Dublin and 4,923,868 bp for S. Typhimurium DT104. The maximum likelihood phylogenetic tree was constructed by MEGA7.0 with 2500 bootstrap repetitions. The clade including S. Schwarzengrund was placed at the root based on established literature (Worley et al., 2018).

The size of the SPIs ranged from 0.7 to 33.3 kb. The coding sequences of virulence genes were used to construct a phylogeny with six of them (I 4, 5, 12) and others of them (I 4, 5, 12). The clade including S. Typhimurium DT104 (Boyd et al., 2000) was further analyzed to identify additional SGI variants identified by the existence of SGI-4 as discovered in the 341 strain (Accession DQ115388). A local last analysis with the same cutoffs was performed to identify the presence of this SGI-4.
Salmonella pathogenic islands contain a variety of genes that contribute to Salmonella virulence and part of Salmonella serotype evolution (Marcus et al., 2000). To assess the complement of SPIs that contribute to the diversity of chromosomal sequences from different serotypes, we constructed a phylogenetic tree using SNPs across 44 chromosomes from 34 serotypes (Figure 1A). This tree reflects the phylogenetic relationships among different Salmonella serotypes and is reflective of their entire genomic content.

The SPIs 1-6, SPI-9, SPI-11, and SPI-24 were present in all isolates (Figure 1B and Supplementary Table 1). SPI-24 was previously called CS54 (Sabbagh et al., 2010) but later renamed as SPI-24 (Cheng et al., 2019). Each of these conserved SPIs were largely in conserved locations as well. Except for large inversions encompassing multiple SPIs in serotypes Infantis, Muenchen, Typhimurium, and Enteritidis (Figures 1B, C). Three SPIs, including SPI-20, SPI-21, and SPI-22, were not identified in our isolate collection.

**FIGURE 1 |** Presence of SPIs in different Salmonella serotypes. (A) A phylogenetic tree was constructed based on chromosomal SNPs of each of the isolates. (B) SPI arrangements are listed with representative sequences of each serotype. Red indicates that SPIS are conserved in all serotypes, yellow indicates insertions of SGIs or plasmids. Other colors represent SPIS whose presence are variable among different serotypes/isolates. (C) The conserved structure across all serotypes is shown.

---

**Supplementary Table 1**

| Serotype      | Size (kb) | Location |
|---------------|-----------|----------|
| Infantis      | 6.3       | 693500-1093749 |
| Muenchen      | 6.1       | 683500-1083749 |
| Typhimurium   | 6.5       | 693500-1093749 |
| Enteritidis   | 6.3       | 693500-1093749 |

---

**Supplementary Figure 1**

A phylogenetic tree was constructed using SNPs across 44 chromosomes from 34 serotypes. The tree reflects the phylogenetic relationships among different Salmonella serotypes and is reflective of their entire genomic content.
Salmonella pathogenic islands were largely conserved within serotypes but showed varying degrees of diversity between serotypes, consistent with previous reports (Hsu et al., 2019; Hao et al., 2020). As expected, from the close relationship between SPI-5 and Typhimurium and S. monophagic variant strain N179016, the SGI-1 flanked on the Typhimurium chromosome and was located between SPI-4 and SPI-6 (Figure 1A). As previously reported, the large fragment missing in SPI-6 belongs to SPI-1 and inverted (Figure 1B).

**Acquisition of Genomic Islands and Associated Antimicrobial Resistance Genes and Heavy Metal Resistant Genes**

In this study, SGI-1 and SGI-1 variants were found among S. Typhimurium strains, including serotypes of Typhimurium (n=8), S. enterica serotype Derby (n=2), S. saintpauli (n=1), and S. galavani (n=1) (Table 1). The location of insertion was unique within SGI-1 variants among S. Typhimurium strains. Interestingly, none of the SPI-1 genomic islands were serotype-specific. As a result, we identified SPIs were distributed among multiple serotypes. Large inversions were observed within S. Typhimurium. For example, S. Typhimurium strain N18S0476 and S. Enteritidis strain N18S1634 each carried similar backbone structure and ARGs, whereas N18S0476 contained only sul1 and blaCARB-2. Therefore, the S. infantis strain S. Typhimurium strain N18S0476 contains a larger SGI-1 sequenced, and the S. infantis strain contains a SGI-1 located between SPI-2 and SPI-6 (Figure 1A).

Integration of Plasmids Into the Chromosome and Associated Antimicrobial Resistance Genes and Heavy Metal Resistant Genes

In this study, SGI-1 and SGI-1 variants were found among S. Typhimurium strains, including serotypes of Typhimurium and S. monophagic variant strain N179016, except that in N17S016, S. Typhimurium strain, and all S. Typhimurium SGI-1 were found among S. Typhimurium strains from retail chicken. For example, S. Typhimurium strain N18S0597 (N18S1595 and N18S2170) was found among S. Typhimurium strains from retail chicken. For example, S. Typhimurium strain N18S0597 (N18S1595 and N18S2170) was found among S. Typhimurium strains from retail chicken.

There are additional examples of ARGs or HMRGs in the chromosome that may have resulted from plasmid integrations. The most common serotypes with chromosomal ARGs were S. enterica serotype Typhimurium (20/24 isolates), I 4,

Integration of Plasmids Into the Chromosome and Associated Antimicrobial Resistance Genes and Heavy Metal Resistant Genes

In this study, SGI-1 and SGI-1 variants were found among S. Typhimurium strains, including serotypes of Typhimurium and S. monophagic variant strain N179016, except that in N17S016, S. Typhimurium strain, and all S. Typhimurium SGI-1 were found among S. Typhimurium strains from retail chicken. For example, S. Typhimurium strain N18S0597 (N18S1595 and N18S2170) was found among S. Typhimurium strains from retail chicken.
### TABLE 1 | Putative Salmonella genomic islands in the genomes of Salmonella isolates.

| Strain ID | Serotype | SGI | Size | Accession | Position | Top non-Salmonella hit of organism and coverage, identity | Accession# | Resistance genes |
|-----------|----------|-----|------|-----------|----------|----------------------------------------------------------|-----------|------------------|
| F18S012   | Typhimurium | SGI-1 | 47,722 | OK209931 | SPI-3, SPI-4 | Proteus mirabilis (89%, 100%) | KJ186152.1 | qacEdelta1 bla<sub>CARB-2</sub> tet(G) floR sul1 qacEdelta1 aadA2 |
| F18S028   | Typhimurium | SGI-1 | 38,450 | OK209935 | SPI-3, SPI-4 | Proteus mirabilis (89%, 100%) | KJ186152.1 | sul1 qacEdelta1 bla<sub>CARB-2</sub> |
| N16S132   | Typhimurium | SGI-1 | 43,948 | OK209937 | SPI-3, SPI-4 | Proteus mirabilis (89%, 100%) | MK422178.1 | tet(A) merR merT merP merC sul1 qacEdelta1 |
| N17S016   | Typhimurium | SGI-1 | 44,852 | OK209939 | SPI-3, SPI-4 | Proteus mirabilis (64%, 100%) | MK422178.1 | tet(A) merR merT merP merC sul1 qacEdelta1 aadA2 dfrA12 |
| N17S1441  | Derby     | SGI-1 | 63,305 | OK209933 | SPI-3, SPI-4 | Proteus mirabilis (93%, 100%) | KJ439039.1 | sul1 qacEdelta1 tet(A) aadA1 tet(A) aph(6)-Id aph(3<sup>00</sup>)-Ib |
| N16S319   | Alachua   | SGI-1 | 47,273 | OK209932 | SPI-3, SPI-4 | Proteus mirabilis (86%, 100%) | KJ439039.1 | sul1 qacEdelta1 tet(A) aadA1 tet(A) aph(6)-id aph(3<sup>00</sup>)-Ib |
| N17S834   | Senftenberg | SGI-1 | 117,891 | OK209936 | SPI-3, SPI-4 | Citrobacter koseri (66%, 89%) | CP026697.1 | tet(A) merR merT merR merP merC sul1 qacEdelta1 aadA1 tet(A) |
| N16S0175  | Saintpaul | SGI-1 | 29,745 | OK209938 | SPI-3, SPI-4 | Proteus mirabilis (97%, 100%) | CP056647.1 | sul1 qacEdelta1 tet(A) aadA1 tet(A) aph(6)-id aph(3<sup>00</sup>)-Ib |
| F18S010   | I 4,[5],12:i:- | SGI-4 | 81,780 | MN730129.1 | SPI-4, SPI-6 | Citrobacter sp. (45%, 97%) | CP056647.1 | pcoS pcoR pcoD pcoC pcoA silP silA silB silC silD silS silE silC arsBarsA arsD arsR |
| F18S014   |           |       |      |           |          |                                                          |           |                  |
| F18S032   |           |       |      |           |          |                                                          |           |                  |
| F18S040   |           |       |      |           |          |                                                          |           |                  |
| F18S043   |           |       |      |           |          |                                                          |           |                  |
| F16S144   |           |       |      |           |          |                                                          |           |                  |
| N17S056   |           |       |      |           |          |                                                          |           |                  |
| N17S1466  |           |       |      |           |          |                                                          |           |                  |
| F18S030   |           |       |      |           |          |                                                          |           |                  |
| N18S0173  |           |       |      |           |          |                                                          |           |                  |
| N16S319   | Alachua   | SGI-4 | ~85,000 | OK209934 | SPI-4, SPI-6 | Enterobacter hormaechei (97%, 100%) | CP042551.1 | silE silS silR silC silF silB silA silP pcoA pcoB pcoD pcoR pcoS pcoE |
| ID          | Serotype       | Estimated size of plasmid on chromosome | Location of insertion | AMR and HMR on inserted plasmid | Match with reference plasmid (accession No) | Plasmid types |
|-------------|----------------|----------------------------------------|-----------------------|----------------------------------|---------------------------------------------|---------------|
| F18S002     | I 4,[5],12:i:- | 16 kb                                  | SPI-1, SPI-9          | *tet(B)* merR merT merP merC sul2 aph(3’)-Id aph(6)-Id bl*TEM-1 | pHCM1 (CP029645.1)                            | IncQ          |
| F18S004     | Muenster       | 12 kb                                  | SPI-1, SPI-9          | su2 merC merP merT merR tet(B)   | pHCM1 (CP029645.1)                            | IncQ          |
| F18S007     | Johannesburg  | 31.6 kb                                | SPI-6, SPI-11         | *sul2 aph(3’)-Id*                | pF18S044-1 (ready for submission)            | IncHI2 IncHI2A|
| F18S023     |                   | 31.6 kb                                | SPI-3, SPI-8          | *sul2 merC merP merT merR tet(B) | pHCM1 (CP029645.1)                            | IncQ          |
| N18S009     |                   | 31.6 kb                                | SPI-3, SPI-8          | *sul2 merC merP merT merR tet(B) | pHCM1 (CP029645.1)                            | IncQ          |
| F18S033     |                   | 31.6 kb                                | SPI-5, SPI-11         | *sul2 merC merP merT merR tet(B) | pHCM1 (CP029645.1)                            | IncQ          |
| N18S034     |                   | 31.6 kb                                | SPI-3, SPI-13         | *sul2 merC merP merT merR tet(B) | pHCM1 (CP029645.1)                            | IncQ          |
| F18S036     |                   | 77.3 kb                                | SPI-1, SPI-3          | *sul2 merC merP merT merR tet(B) | pHCM1 (CP029645.1)                            | IncQ          |
| F18S039     |                   | 71 kb                                  | SPI-5, SPI-11         | *sul2 merC merP merT merR tet(B) | pHCM1 (CP029645.1)                            | IncQ          |
| N18S042     |                   | 71 kb                                  | SPI-6, SPI-16         | *sul2 merC merP merT merR tet(B) | pHCM1 (CP029645.1)                            | IncQ          |
| F18S045     |                   | 54 kb                                  | SPI-4, SPI-6          | *sul2 merC merP merT merR tet(B) | pHCM1 (CP029645.1)                            | IncQ          |
| N18S059     |                   | 85 kb                                  | SPI-1, SPI-9          | *sul2 merC merP merT merR tet(B) | pHCM1 (CP029645.1)                            | IncQ          |
| N18S078     |                   | 76.4 kb                                | SPI-1, SPI-9          | *sul2 merC merP merT merR tet(B) | pHCM1 (CP029645.1)                            | IncQ          |
| N18S097     |                   | 107 kb                                 | SPI-6, SPI-16         | *sul2 merC merP merT merR tet(B) | pHCM1 (CP029645.1)                            | IncQ          |
| (Continued) |

Frontiers in Microbiology | www.frontiersin.org

November 2021 | Volume 12 | Article 777817
TABLE 2 | (Continued)

| ID     | Serotype | Estimated size of plasmid on chromosome | Location of insertion | AMR and HMR on inserted plasmid | Match with reference plasmid (accession No) | Plasmid types     |
|--------|----------|------------------------------------------|-----------------------|---------------------------------|--------------------------------------------|-------------------|
| N18S1595 | Typhimurium | 56 kb                                  | SPI-6, SPI-16          | merR tet(A) sul2                  | pF18S004                                      | Unknown - not closed |
| N17S0166 | I 4,[5],12:i:- | 158 kb                                | SPI-1, SPI-9           | pccoS mcr-9.1 resC aadA1          | pN53063 (CP049311.1)                          | IncHI2 IncHI2A     |
| N18S0736 | I 4,[5],12:i:- | 38 kb                                  | SPI-1, SPI-9           | sul1, qacEdelta1, aac(3)-Vla aadA1 resR | pN53063 (CP049311.1)                          | IncHI2 IncHI2A     |
| F18S004 | Typhimurium | 21 kb                                  | SPI-6, SPI-9           | tetW tetZ tetD                   | pF18S044-1                                    | IncHI2 IncHI2A     |
| F18S013 | Typhimurium | 3.8 kb                                  | SPI-5, SPI-14          | bla<sub>CMY</sub>-2              | pF18S003-1/pF18S007-1/pN16S065-2              | IncC/IncI/IncB/O/K/Z |
| N16S021 | N16S070   | 3.8 kb                                  | SPI-9, SPI-12          | bla<sub>CMY</sub>-2              | pF18S003-1/pF18S007-1/pN16S065-3              | IncC/IncI/IncB/O/K/Z |
| N16S009 | N16S098   | 3.8 kb                                  | SPI-3, SPI-13          | bla<sub>CMY</sub>-2              | pF18S009-1/pF18S007-1/pN16S065-10             | IncC/IncI/IncB/O/K/Z |
| N16S189 | N16S214   | 3.8 kb                                  | SPI-2, SPI-12          | aph(3')-Ib aph(6)-Ib tet(A)       | pH1038-142(KJ448434.1)                        | IncN IncFil        |
| N18S0645 | N18S2188  | 3.8 kb                                  | SPI-2, SPI-12          | tet(A) merR merP merC             | pN16S024 (CP052840.1)                         | IncHI2            |

FIGURE 2 | IncQ plasmid integrated into Salmonella I 4,[5],12:i:- chromosome. The circle on the left represents a chromosome structure of a S. Typhimurium strain F18S013 with SPIs (in red) distributed on the chromosome. The green area was the sequence homologous to a Salmonella I 4,[5],12:i:- strain F18S010. The grey area shows differences between the two strains. The detailed comparison of the genomic structure of the two strains on the right is from the area with black box on the chromosome.

N18S0173 and N18S2170 carried multiple chromosomal ARGs (Supplementary Table S1). These isolates encompassed seven different serotypes.

Chromosomal HMRGs were found in 51 of 134 isolates, including genes encoding resistance to copper, silver, mercury, and arsenic. These comprised 29 turkey, 12 pig, 1 chicken, and 1 cattle isolate. Thirty-nine of the 51 HMRG isolates with more than three chromosomal HMRGs also had chromosomal ARGs. In many cases, HMRGs and ARGs were physically linked (Table 2).

Together these findings among 134 Salmonella genomes showed that the maintenance and spread of chromosomal ARGs and HMRGs in Salmonella is accomplished through a complex interplay of genomic islands and integrated plasmids. Further work will be needed to understand whether the acquisition of these genes is specifically selected for by exposure to heavy metals and/or antimicrobials or connected to other survival and fitness challenges faced by Salmonella.

Plasmid Types and Association With Resistance Genes, Sources, and Serotypes

From the 134 Salmonella isolates, we developed high-quality sequence for 285 plasmids, 245 of which were fully resolved into circular structures. On average, there were two plasmids per isolate, although seven isolates had no plasmids and one isolate...
FIGURE 3 | Plasmid types present across different food animal sources. Replicon information was based on analysis from PlasmidFinder. Only the most common plasmid types are represented in the figure.

TABLE 3 | Association of plasmid types with antimicrobial resistance genes (ARGs) and heavy metal resistant genes (HMRGs).

| Plasmid types | Total number | Avg. size in kb (range) | Most common serotype (number) | Avg. ARGs (range) | Avg. HMRGs (range) |
|---------------|--------------|-------------------------|-------------------------------|-------------------|---------------------|
| IncA          | 2            | 133 (90–176)            | Heidelberg, I, 4,[5],12:i:- (1) | 5.5 (5–6)         | 2.5 (0–6)           |
| IncB/O/K/Z    | 1            | 115                     | Kentucky                       | 4                 | 4                   |
| IncC          | 36           | 142 (52–232)            | Typhimurium (11)               | 6.6 (0–13)        | 3.8 (0–8)           |
| IncF          | 23           | 112 (16–164)            | Kentucky (9)                   | 2.1 (0–6)         | 0.8 (0–7)           |
| IncH          | 15           | 270 (145–354)           | Kentucky (3)                   | 5.4 (1–14)        | 13.9 (3–22)         |
| IncI          | 39           | 99 (53–125)             | Kentucky, I, 4,[5],12:i:- (7)  | 1.6 (0–4)         | 0.7 (0–15)          |
| IncN          | 3            | 57 (43–71)              | Enteritidis, Heidelberg, I, 4,[5],12:i:- | 3 (1–5) | 0           |
| IncP          | 1            | 18                      |                                | 2                 | 0                   |
| IncQ          | 11           | 11 (8–12)               | Reading                        | 3.3 (1–4)         | 0                   |
| IncR          | 1            | 70                      | Muenser                        | 10                | 0                   |
| IncX          | 22           | 42 (31–53)              | Kentucky (9)                   | 0.4 (0–8)         | 0                   |
| IncY          | 1            | 92                      | I, 4,[5],12:i:-               | 0                 | 0                   |
| Col           | 69           | 5 (2–15)                | Typhimurium (11)               | 0.3 (0–2)         | 0                   |
| Phage-like     | 1            | 91                      | Typhimurium                    | 0                 | 0                   |
| Combination   | 12           | 237 (75–389)            | Infantis (4)                   | 6.3 (0–12)        | 6.3 (0–18)          |
| Unknown (no replicon) | 28 | 18 (1–186) | Typhimurium (5) | 0.6 (0–6) | 0.5 (0–14) |
| Chromosome    | 134          | 4.8 Mb (4.5–5.1 Mb)     | N/A                            | 1.7 (0–7)         | 7.3 (0–25)          |

As many as seven plasmids were analyzed for closed, circular sequences (Supplementary Table 1). We sought to identify the AMR and HMRGs on these plasmids and evaluate if any plasmids had animal source or serotype-specificity. Although some plasmid types were most common in certain sources, most were widely distributed across all animal sources (Figure 3). Some association between plasmid type and serotype was also noticed (Table 3).

**IncC**

Sizes of the 36 IncC plasmids varied considerably, from 52 to 232 kb, and were found among all animal sources (Table 3). These plasmids were found among 11 different serotypes, with S. Typhimurium most prevalent. Some common resistance patterns emerged. IncC-2 plasmids harbored tetA and tetB, and IncC-5 harbored aph(6)-Id/aph(3’)-Ib. Twelve of the isolates also had at least one qacE gene and 29 had at least one mercury resistance gene.

**IncHI**

There were 15 IncHI plasmids identified among twelve different serotypes, present in all food animal sources, with swine being most common (7/15, 47%). All were large plasmids of 145–354 kb and possessed from one to fourteen ARGs. Seventeen also had the qacE biocide resistance gene, and all had at least one metal resistance operon, including those conferring resistance to silver, copper, mercury, and arsenic.

**IncI**

There were 39 IncI plasmids in our collection of 134 genomes, from 19 different serotypes, with at least one ARG found in 33 isolates.
other serotypes. Plasmid type and resistance genes was not observed in any two of the plasmids had MRGs and the qacE genes were seen in eight of the plasmids. Most common among the sources of IncI plasmids was turkey, with 22 isolates, and of these were Serotype A. [5], 12:i:-.

Other Plasmid Types

Several other plasmid types were identified, including IncX, IncN, IncC, IncHI, IncR, IncF, IncP, IncQ, and IncX replicons. The most strongly associated with Serotype Agona had the genes integrated into the chromosome. Some interesting Serotype-specific findings. Serotype Agona had blaCMY-2 and tetA, which were found in eight of the plasmids. The most common animal source for Serotypes, including Typhimurium, one I 4,[5],12:i:-, two Salomonella enterica, and all five Serotypes, including Montevideo and all five Salmonella enterica.

DISCUSSION

Here we present the results from long-read sequencing of Salmonella genomes, including over two hundred plasmids with a diverse range of ARGs. The distribution of ARGs was not uniform across different genes. Some genes that were plasmid-specific included atA and ancr-90 only on IncHI plasmids [Tyson et al., 2020] and aprB only on Col plasmids. Other genes such as tet-M were more widely distributed across different plasmid types (Fig. 1).

Overall, there was an association between the presence of MRGs and the resistance genes on the same plasmid (Fig. 2). Some genes that were plasmid-specific included atA and ancr-90 only on IncHI plasmids. Other genes such as tet-M and tetB, were more widely distributed across different plasmid types (Fig. 1). Overall, there was an association between the presence of MRGs and the resistance genes on the same plasmid (Fig. 2).
We identified twelve isolates with SGI-1 sequences, either by homology to the reference SGI-1 sequence or from insertions in the same region. All SGI-1s were similar to SGI-1 in Proteus mirabilis, but their closest relatives also included sequences from Citrobacter and Enterobacter. This finding further helped us understand how these genomic islands were horizontally acquired.

We had other novel findings, including an SGI-1 sequence in serotype Alachua and large SGI-1s of different origin in serotypes Senftenberg and Saintpaul. These SGI-1s and a Salmonella SGI-1 variant were horizontally acquired. The great diversity made it impossible to name the variants alphabetically. A typical approach in this study was to resolve these issues associated with the naming of SGI sequences by proposing a new approach based on their relative position in the genome. For instance, SGI-0 and SGI-2 (Levings et al., 2008; de Curraize et al., 2020) as previously named and named as SGI-1 based on their consistent position with other SGI-1 variants. An additional example is SGI-IV, which is shown by a novel SGI island containing HMRGs in S. Alachua. Even though it has limited homology with a previously reported SPI-1 (95% identity and 50% length), it was named as SGI-4 variant because of its genomic location. By naming SGIs based on location, we developed a streamlined process for SGI nomenclature in future work to identify diverse SGI sequences identified and help identify the potential new variants. It would be interesting to further investigate the prevalence and distribution of GI in Salmonella from other sources, including from humans and sick animals. In this study, we also found ARGs and HMRGs on many chromosomal sequences with 100% homology to plasmids, indicating fragment of plasmid integration into the chromosome. Table 2 shows fourteen isolates of S. Typhi, S. Paratyphi A and S. Paratyphi B with MDR IncQ plasmids were inserted into the same location in the chromosome. (Figure 2). This monophasic serotype of S. Typhi is the most common cause of typhoid fever in the world.
results from different insertions, deletions, or other disruptions of genes.

Supplementary Material

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

AUTHOR CONTRIBUTIONS

CL, GH, TZ, and ZK conceptualized the ideas, validated the formal analysis, investigated the data handling, and contributed to the original draft preparation. ES contributed to the software. ES, GET, UD, and PM supervised the project. CL, GHT, and C-HH managed the resources. CL and T-TT performed the visualization. CL, GHT, C-HH, and SZ developed the methodology. CL, C-HH, and ES performed the analysis, investigation, and writing—original draft preparation. ES, GET, UD, and PM were involved in the writing—review and editing. CL, GHT, C-HH, and SZ were involved in the visualization. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at https://www.frontiersin.org/articles/10.3389/fmicb.2021.777817/full#supplementary-material.

REFERENCES

Abebe, T., Kuga, J., and Okayama, M. (2020). Review of major food-borne non-tuberculous bacteria pathogen. S. Afr. J. Food Sci. Tech. 2020. 47:235.

Amavisit, P., Lightfoot, D., Browning, G. F., and Markham, P. F. (2003). Variation between Salmonella pathogenicity islands. J. Bacteriol. 185, 3624–3635. doi:10.1128/JB.185.12.3624-3635.2003

Blondel, B. R., Contreras, I., Santiviago, C. A., and Bawn, M. (2009). Comparative genomic analysis uncovers novel mobile genetic elements in the Salmonella enterica serovar Typhimurium DT104. Genome Biol. 10:354. doi:10.1186/1471-2164-10-354

Boyd, D. P., Peters, J. R., and Carter, A. M. (2000). Partial characterization of an inner membrane protein associated with the multidrug resistance region of Salmonella enterica serovar Typhimurium. Antimicrob. Agents Chemother. 44, 2180–2187.

Bearson, B. L., Bearson, B. L., and Shippy, D. C. (2019). Sgi-4 in monophasic Salmonella enterica serovar i 4,[5],12:i:- pork outbreak isolate usda15wa-1. FEMS Microbiol. Lett. 189, 285–291. doi:10.1098/rstb.2019.0072

Kim, J., et al. (2016). Plasmids in Salmonella enterica: Assembly of plasmids from whole-genome sequencing data. Bioinformatic. 32, 380–387.

Krause, J., Branchuch, C. J., Bearson, B. L., and Ovington, C. M. (2020). Linear plasmids of Salmonella enterica serovars and their metal tolerance. Antimicrob. Agents Chemother. 64, e00429–19. doi:10.1128/AAC.00429-19

Antipov, D., Hartwick, N. J., and Pevzner, P. A. (2021). Clusters of plasmids conferring resistance to three or more antimicrobial agents (3 AMRGs) are widespread among multidrug-resistant Salmonella enterica serotypes. Front. Microbiol. 12:777817. doi:10.3389/fmicb.2021.777817/full#supplementary-material
enhance resistance to copper. Front. Microbiol. 11, 01118. doi: 10.3389/fmicb.2019.01118

Carroll, A. R. (2013). Plasmids and the spread of resistance. Int. J. Med. Microbiol. 303, 98–99. doi: 10.1016/j.ijmm.2013.02.000

Carroll, A. R., Zancki, A., Garcia-Fernandez, J. A., Yohannes, L. D., Mitha, M., and Best, S. (2014). Antibiotic resistance of human salmonellas, plasmidfinders and plasmid multilocus sequence typing. Antimicrob. Agents Chemother. 58, 3895–3903. doi: 10.1128/aac.02412-14

Cheng, S. R., S. Fadde, N. D., and Wiedmann, M. M. (2019). Embracing diversity through discovery and development of plasmid mechanisms. J. Biotechnol. 320, 176–182. doi: 10.1016/j.jbiotec.2018.10.012

Comparative genomic analysis of 450 strains of Salmonella enterica serovars typhimurium and typhi. Antimicrob. Agents Chemother. 50(5). doi: 10.1128/aac.01030-16

Partridge, R. J., K. Wong, S. R., Fireth, N. D., and ensen, R. E. (2018). Mobile genetic elements associated with antimicrobial resistance in Salmonella enterica. Microb. Ecol. 61, 203–216. doi: 10.1007/s00248-017-0857-1
Conflicts of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s Note: All claims made in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.