GeoHealth

RESEARCH ARTICLE

10.1029/2023GH000877

Key Points:

- *Salmonella enterica* isolated from water samples collected near swine farms after Hurricane Florence were from an environmental reservoir
- In relation to the chromosome, we also found that one mobile genetic element is more promiscuous than the other
- Environmental *S. enterica* were less likely than swine-associated *S. enterica* to have antibiotic resistance genes present in plasmids

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

Y. Mao,
yuqingm2@illinois.edu

Citation:

Mao, Y., Zeineldin, M., Usmani, M., Jutla, A., Shisler, J. L., Whitaker, R. J., & Nguyen, T. H. (2023). Local and environmental reservoirs of *Salmonella enterica* after Hurricane Florence flooding. *GeoHealth*, *7*, e2023GH000877. https://doi.org/10.1029/2023GH000877

Received 8 JUN 2023
Accepted 13 OCT 2023

Author Contributions:

Conceptualization: Yuqing Mao, Mohamed Zeineldin, Rachel J. Whitaker, Thanh H. Nguyen
Data curation: Yuqing Mao
Formal analysis: Yuqing Mao, Joanna L. Shisler, Thanh H. Nguyen
Funding acquisition: Antarpreet Jutla, Joanna L. Shisler, Rachel J. Whitaker, Thanh H. Nguyen
Investigation: Yuqing Mao, Mohamed Zeineldin

© 2023 The Authors. GeoHealth published by Wiley Periodicals LLC on behalf of American Geophysical Union. This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

Local and Environmental Reservoirs of *Salmonella enterica* After Hurricane Florence Flooding

Yuqing Mao[1,2], Mohamed Zeineldin[2], Moiz Usmani[3], Antarpreet Jutla[3], Joanna L. Shisler[2,4], Rachel J. Whitaker[2,4], and Thanh H. Nguyen[1,2,5]

[1]Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign, Urbana, IL, USA, [2]Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, IL, USA, [3]Engineering School of Sustainable Infrastructure & Environment, University of Florida, Gainesville, FL, USA, [4]Department of Microbiology, University of Illinois at Urbana-Champaign, Urbana, IL, USA, [5]Carle Illinois College of Medicine, University of Illinois at Urbana-Champaign, Urbana, IL, USA

Abstract In many regions of the world, including the United States, human and animal fecal genetic markers have been found in flood waters. In this study, we use high-resolution whole genomic sequencing to examine the origin and distribution of *Salmonella enterica* after the 2018 Hurricane Florence flooding. We specifically asked whether *S. enterica* isolated from water samples collected near swine farms in North Carolina shortly after Hurricane Florence had evidence of swine origin. To investigate this, we isolated and fully sequenced 18 independent *S. enterica* strains from 10 locations (five flooded and five unflooded). We found that all strains have extremely similar chromosomes with only five single nucleotide polymorphisms (SNPs) and possessed two plasmids assigned bioinformatically to the incompatibility groups IncFIB and IncFII. The chromosomal core genome and the IncFIB plasmid are most closely related to environmental *Salmonella* strains isolated previously from the southeastern US. In contrast, the IncFII plasmid was found in environmental *S. enterica* strains whose genomes were more divergent, suggesting the IncFII plasmid is more promiscuous than the IncFIB type. We identified 65 antibiotic resistance genes (ARGs) in each of our 18 *S. enterica* isolates. All ARGs were located on the *Salmonella* chromosome, similar to other previously characterized environmental isolates. All isolates with different SNPs were resistant to a panel of commonly used antibiotics. These results highlight the importance of environmental sources of antibiotic-resistant *S. enterica* after extreme flood events.

Plain Language Summary Coastal flooding is increasing in frequency due to climate change. It cripples civil infrastructure serving humans and damages many livestock facilities. It is usually assumed that flooding of wastewater, septic systems, and/or livestock manure transmits antibiotic-resistant bacteria and antibiotic resistance genes to the surrounding environment. However, conclusive identification of the source of the microbial contaminants has not been reported. Moreover, the above assumption rules out environmental reservoirs as a potential source for spread. Here, we report that antibiotic-resistant *Salmonella enterica*, isolated from water samples collected near swine farms after Hurricane Florence, were not from animals or manure. Instead, they were from an environmental reservoir. Our findings were based on analyzing chromosomes and plasmids independently and collectively using long-read high-fidelity whole-genome sequences. Knowledge of potential pathogen sources can help mitigate the spread of pathogenic bacteria after hurricanes to minimize the impact of floods on human health.

1. Introduction

Climate change is increasing the frequency and magnitude of extreme precipitation and flooding events around the world (Seneviratne et al., 2021; Tabari, 2020). Floods can be billion-dollar disasters due to disease transmission, infrastructure damage, water resource contamination, and others (Ohl & Tapsell, 2000). The climate change crisis is developing in tandem with increased resistance to antibiotics, a public health crisis that could result in 10 million deaths globally by 2050 (de Kraker et al., 2016; Murray et al., 2022). It is usually assumed that flooding of wastewater, septic systems, and/or livestock manure transmits antibiotic-resistant bacteria and antibiotic resistance genes (ARGs) to the surrounding environment. However, conclusive identification of the source of the microbial contaminants has not been reported. If we could identify the reservoirs of antibiotic-resistant pathogens after massive flooding, then this information is the first step for controlling the spread of these microbial contaminants, leading to building resilience to extreme flooding due to climate change.

Methodology: Yuqing Mao, Mohamed Zeineldin, Moiz Usmani, Antarpreet Jutla, Joanna L. Shisler, Rachel J. Whitaker, Thanh H. Nguyen
Project Administration: Rachel J. Whitaker, Thanh H. Nguyen
Resources: Rachel J. Whitaker, Thanh H. Nguyen
Software: Yuqing Mao, Mohamed Zeineldin, Moiz Usmani, Antarpreet Jutla
Supervision: Rachel J. Whitaker, Thanh H. Nguyen
Validation: Yuqing Mao
Visualization: Yuqing Mao
Writing – original draft: Yuqing Mao
Writing – review & editing: Yuqing Mao, Mohamed Zeineldin, Joanna L. Shisler, Rachel J. Whitaker, Thanh H. Nguyen

There are multiple human bacterial pathogens (e.g., *Salmonella enterica*, *Escherichia coli*, and *Vibrio cholerae*) that can replicate in the environment and in multiple host animals. The most-used microbial source tracking tool for fecal indicators, quantitative polymerase chain reaction (qPCR), can only detect positive correlations between the queried host-source specific genetic markers and pathogens (Harris et al., 2021; Harwood et al., 2014; Jang et al., 2017). qPCR for conserved chromosomal genes has been used to identify elevated levels of pathogen indicators and fecal contaminants in water samples after flooding events caused by Hurricanes Harvey (Texas) (Moghadam et al., 2022; P. Yu et al., 2018) and Florence (North Carolina (NC)) (Harris et al., 2021; Niedermeyer et al., 2020). While these culture-independent and genomic approaches can identify the abundance of a pathogen, they do not have the genetic resolution to identify the source. Next-generation sequencing of 16S rRNA genes from environmental samples can estimate the degree of fecal contamination in a natural water body. However, this approach, too, does not resolve the differences between operational taxonomic units and the species level of a pathogen present. Whole-genome sequencing is a more promising approach to tracking the source of a pathogen in a complex environment (Amirsoleimani et al., 2019; Flach et al., 2021; Henriot et al., 2023; Raza et al., 2021). Short-read whole-genome sequencing allows accurate reconstruction of the phylogenetic relationship of the sampled bacteria. However, the bioinformatic pipelines for short-read whole-genome sequencing used to identify key indicators of transmission such as single nucleotide polymorphisms (SNPs) and mobile genetic element (MGE) sequences or chromosome sequences are laborious or impossible due to the complexity and number of repeated sequences in most bacterial chromosomes. It is important to assemble chromosome and MGE sequences separately but analyze them as coexisting in the same cell, to determine the different transfer routes of different genome components. In addition, previous studies found strong evidence that ARGs and virulence factors could be disseminated by MGEs (Balbuena-Alonso et al., 2022; de Been et al., 2014). These findings emphasized the importance of the identification of the evolutionary history of MGEs. Third-generation long-read high-fidelity sequencing is the most advanced whole-genome sequencing technology, and it can separate the sequences of bacterial chromosomes and extra-chromosomal MGEs to reveal the source of the target bacteria (Hon et al., 2020). With long-read PacBio HiFi sequencing, the phylogenetic relationship of the chromosomes can be directly reconstructed to show the origin of the *Salmonella* isolates, and the MGEs can provide information on the interaction of the *Salmonella* with bacteria from other potential contamination sources. Here, for the first time, we use this technology for the *S. enterica* strains isolated from a flood region to investigate the source of bacteria and MGEs. With this knowledge, pathogen control strategies can be applied more accurately to minimize flooding impacts on risks of infectious disease spread to humans.

Hurricane Florence brought extensive flooding over a large region in the southeastern US, including NC in 2018 (Aly et al., 2021; Callaghan, 2020). This flood had a significant public health impact. For example, there was an 85% increase in emergency room admissions due to bacterial acute gastrointestinal illness after flooding from Hurricane Florence in 2018 (Quist et al., 2022a). Several studies on source tracking of pathogens for water samples collected right after the hurricane have been conducted. Human and swine fecal indicators and pathogens including *Arcobacter butzleri*, *Listeria* spp., and *Salmonella* spp. were detected in water samples collected in NC after Hurricane Florence (Harris et al., 2021; Niedermeyer et al., 2020). Yet, there was no conclusive identification of the source of the contamination of water (human, swine, or environmental) (Harris et al., 2021; Niedermeyer et al., 2020). *Salmonella* spp. are endemic in swine (Bearson, 2021), and our initial hypothesis was that the *S. enterica* collected in this study would be from swine sources. To test this hypothesis, we isolated and sequenced *S. enterica* from water samples collected after this flood. To balance the sequencing quality and the cost, we combined results from Illumina NGS and PacBio HiFi sequencing with the analysis of available whole genome sequences from NCBI. This approach allowed us to identify with high resolution the source of the *S. enterica* isolates. Our work also contributes to a better understanding of the distribution and the spread of pathogens after flood events.

2. Methods

2.1. *Salmonella enterica* Isolation From Flood Samples

Three weeks after Hurricane Florence, on 7 October 2018, we collected water samples from water bodies downstream of the swine farms in an agricultural production area in NC. The description on the selection of sampling locations and sampling protocols has been described previously (Mao et al., 2021). Briefly, we collected water samples from 25 sampling sites, including 16 flooded sites and nine unflooded sites (Figure S1 in Supporting

Information S1). The initial stage in precisely determining sampling sites involved describing the inundated areas following hurricanes. In this study, we employed a relatively novel approach to map inundation, relying on the geomorphometric principles of the landform (Bolch et al., 2011; Clubb et al., 2017; Sofia et al., 2014), though historically hydrological and hydrodynamic models were used to simulate inundation caused by intense rainfall. Here, the fundamental idea is to allow the topography to determine the flow of water across a specific terrain. This approach enables a rapid yet stationary calculation, accurately pinpointing areas prone to flooding following intense rainfall (Usmani et al., 2023). The inundated sites were determined as "flooded," and the sites that were not inundated were determined as "unflooded." To strengthen these outcomes, we conducted independent observations and documented our own records, along with engaging local individuals on-site to verify the occurrence of flooding. Because the swine farms are private properties, the sampling sites were on public lands which were hundreds of meters downstream of the swine farms.

The unflooded site U4 was selected as the sampling negative control because it had the highest elevation among all sampling sites and there was no swine farm around. All other 24 sampling sites had swine farms upstream. Approximately 2 L of the water samples and two to four replicates of water samples were collected from each site. Each replicate of a water sample was processed individually in the lab. Right before pre-processing, $MgCl_2$ solution was added to the sample bags to the final concentration of 25 mM to aid flocculation. After flocculation, the water samples were vacuum filtered in a biosafety cabinet through 1.6 μm glass fiber filters (Millipore) to collect solid particles. When a filter clogged, it was replaced with a new filter until all the water sample in the bag was filtered. After filtering, the filters were cut into quarters for different analyses. Two of these quarters were stored at 4°C in Whirl-Pak standard sample bags (Nasco) sealed by packing tape until being used for isolation of *Salmonella*.

S. enterica was isolated from the filter pieces as follows. First, a filter piece from each replicate of a water sample was inserted into an individual 15 mL centrifuge tube containing 10 mL antibiotic-free LB broth. These tubes were incubated by shaking at 37°C overnight. A tube of LB broth without filters inserted was used as the blank control for *S. enterica* enrichment. A portion of the broth was streaked to xylose lysine tergitol 4 (XLT4) (Thermo Scientific) agar plates for *S. enterica* isolation and selection. Black-colored colonies were picked, re-streaked to new XLT4 agar plates, and incubated at 37°C overnight until single, black-colored colonies were visible ("Xylose lysine tergitol 4 (XLT4) agar," 2003). XLT4 plates streaked by clean loops were used as the blank control for *S. enterica* isolation. Briefly, DNA was extracted from each isolate using the GenFind V2 extraction kit (Beckman Coulter Life Sciences, USA) according to the manufacturer's protocol. The extracted DNA was used for PCR amplification using universal 16S rRNA primers (Universal 1492r; GGY TAC CTT GTT ACG ACT T and Bacterial 27F; AGR GTT TGA TYM TGG CTC AG) (Ishii et al., 2017). The blank control for PCR was made by replacing the DNA sample with molecular biology grade water. The isolated PCR products were purified and sequenced by Sanger sequencing to determine the species of the isolates at the Roy J. Carver Biotechnology Center at the University of Illinois at Urbana-Champaign, Illinois, USA. The concentrations of PCR amplicons were quantified by NanoDrop before sending for sequencing. Then the sequences were aligned using the Clustal W program and compared with known 16S rRNA gene sequences in the GenBank database using nucleotide BLAST (Sayers et al., 2019; Thompson et al., 1994). *S. enterica* was successfully isolated from 10 sampling sites, including five flooded sites (F1, F3, F6, F7, and F12) and five unflooded sites (U1, U2, U5, U6, and U8), as shown in Figure 1a. For each site, four independent *S. enterica* isolates were picked from each agar plate and re-streaked for isolation, resulting in a total of 40 isolates. A portion of each colony was then frozen as a stock source by mixing a part of the colony in LB broth. Then, a 50% glycerol solution was added to the inoculated LB culture in a 1:1 volume ratio for storage −80°C until sequencing. Although samples were collected outside the flood regions (sample U4), no *S. enterica* was isolated from this sample (Figure S1 in Supporting Information S1).

Because the sampling sites were downstream of swine farms, selected *S. enterica* isolates were tested for susceptibility to antimicrobials at Diagnostic Lab, College of Veterinary Medicine, the University of Illinois at Urbana Champaign using bovine/porcine (with erythromycin) minimum inhibitory concentration (MIC) format plate (Sensititer BOPO6F, Thermo Fisher Scientific, Oakwood Village, OH) in accordance with the manufacturer's instructions. The antibiotics used were ampicillin, clindamycin, danofloxacin, enrofloxacin, florfenicol, gamithromycin, gentamicin, neomycin, tetracycline, penicillin, sulphadimethoxine, spectinomycin, trimethoprim-sulfamethoxazole, tiamulin, tilmicosin, tildipirosin, tulathromycin, tylosin tartrate, and erythromycin. The positive control was set by default at the G12, H11, and H12 wells on the plate. The results of the antimicrobial susceptibility test included the numerical value corresponding to the MIC and the categorical criterion of

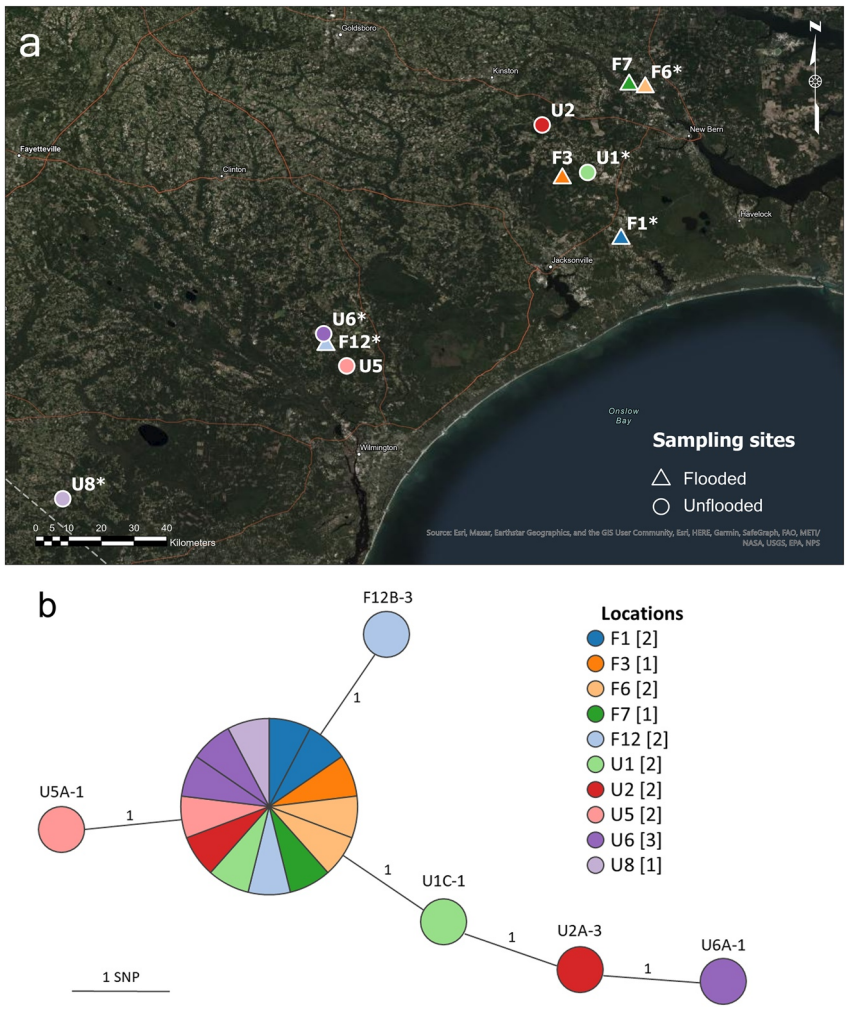

Figure 1. The geographic locations of *S. enterica* isolates collected after Hurricane Florence flooding in North Carolina (NC) are not dependent on proximity to flooded swine farms (a) *S. enterica* was isolated from water from 10 sampling sites in NC, including five flooded (F1, F3, F6, F7, and F12) and five unflooded sites (U1, U2, U5, U6, and U8). The circular dots represent unflooded sites, and the triangle dots represent flooded sites. Sites with PacBio-long-read-sequenced isolates are marked with a "*" in their labels. (b) single nucleotide polymorphism (SNP)-based phylogenetic tree for the *S. enterica* isolates sequenced in this study, constructed by using the five SNPs identified in the chromosomes. Each node includes a group of isolates that share identical whole-genome sequences. Each segment within a node shows the number of isolates. The colors of the pies show the sample collection site of the isolates. The colors in the tree match the colors in the map. The numbers on the branches that connect the nodes show the number of SNPs between the two connected nodes.

resistant, or susceptible, were interpreted on the basis of the Clinical and Laboratory Standards Institute (CLSI) standards (Patel et al., 2016). See Figure S2 in Supporting Information S1 for the overall methodology of sample processing and analysis.

2.2. Whole-Genome Sequencing and SNP Identification

We conducted whole-genome sequencing of all 40 *S. enterica* isolates by short-read Illumina technology. Because the short-read sequencing analysis revealed nearly identical genomes, we selected six isolates for long-read PacBio HiFi technology. These six isolates included three from flooded sites (F1C-1, F6B-2, and F12B-3), and three from unflooded sites (U2A-1, U6A-1, and U8A-3). The short-read sequencing was then mapped onto long-read sequencing for SNP identification. The long-read sequences were also used for genome assembling into chromosome and chromosome-independent MGEs. The sequence data sets were submitted to NCBI under BioProject accession number PRJNA1006749. See Figure S3 in Supporting Information S1 for the overall methodology of whole-genome sequencing and SNP identification.

For short-read sequencing, total DNA was extracted from a 1.0 mL overnight LB culture inoculated from the frozen stock using the GenFind V2 extraction kit (Beckman Coulter Life Sciences) according to the manufacturer's protocol. Extracted DNA quality was analyzed and quantified using Qubit 4.0 fluorometric quantification (Thermo Fisher Scientific, MA, USA), and agarose gel electrophoresis (Bio-Rad Laboratories, Inc, Hercules, CA, USA). All DNA samples were stored at −80°C pending further analysis. Sequencing libraries for the 40 *S. enterica* DNA samples were prepared from extracted genomic DNA using Nextera XT Kit (Illumina, San Diego, CA, USA) and quantified using Qubit High-Sensitivity DNA (Life Technologies, Grand Island, NY, United States). The libraries were normalized using a bead-based procedure and pooled together at equal volumes. The pooled library was denatured and sequenced using Illumina HiSeq sequencing (Illumina, San Diego, CA, USA) at the W. M. Keck Center for Comparative and Functional Genomics, the University of Illinois at Urbana-Champaign. The raw sequence files were de-multiplexed and converted to fastq files using Casava v.1.8.2 (Illumina, Inc, San Diego, CA). Illumina sequencing adapters and low-quality bases were removed using Trimmomatic version 0.36 (Bolger et al., 2014). Draft assemblies for all sequenced genomes were carried out using SPAdes (Bankevich et al., 2012).

For long-read sequencing, the frozen stock was re-streaked to XLT4 agar plates three times for purification. Then, one colony from each of the six lineages was picked and enriched in 900 μL of LB broth with 37°C overnight incubation. A tube with 900 μL of LB broth without bacteria inoculated was used as a blank control for enrichment. After enrichment, the *Salmonella* DNA was extracted by Nanobind CBB big DNA kit (Pacific Biosciences) following the instructions provided by the manufacturer. The quality and quantity of the DNA extracts were analyzed by NanoDrop and gel imaging. The DNA extracts were sheared into an average of 13 kb fragments by Megarupter 3 (Diagenode, NJ, USA). The sheared DNA was converted to barcoded libraries by SMRTBell Express Template Prep kit 2.0 (PacBio, Menlo Park, CA, USA). Then the pooled library was sequenced on one SMRTcell 8M on PacBio Sequel IIe (PacBio) using the CCS sequencing mode and 30 hr of movie time. An internal control was included in the sequencing run to ensure the read length and the read depth of the libraries.

The long-read sequencing reads were assembled by Trycycler (Wick et al., 2021) by dividing the raw read file into 24 read subsets with Raven (Vaser & Šikić, 2021) as the assembler. The contigs assembled from each subset were clustered, reconciled, aligned, partitioned, and combined to get polished assemblies for the whole genomes. The 24 assembled contigs for each genome were clustered by mash distance (Ondov et al., 2016). The chromosome was identified as a cluster with tightly correlated contigs with lengths of approximately 5 million base pairs. Chromosome-independent MGEs were identified as clusters that were separated from the chromosome cluster, and the contigs within those clusters were tightly correlated. The contigs within one cluster that were not highly similar to the other contigs in the same cluster would be identified as poorly assembled contigs. The poorly assembled contigs were discarded for downstream analyses. After identifying all clustered contigs and removing the poorly assembled contigs, the selected contigs went through reconciling, multi-sequence aligning, and read partitioning steps for correction. Each cluster would have one consensus fasta file generated. The combination of the consensus fasta files of all clusters was the final assembly output. DNA sequences from each of the six selected *S. enterica* genomes identified four contigs, including one chromosome and three chromosome-independent MGEs.

Both short-read and long-read whole-genome sequences of the *S. enterica* isolates were submitted to the Bacterial and Viral Bioinformatics Resource Center (BV-BRC) for genome annotation after assembly (Brettin et al., 2015; Olson et al., 2023). Specialty genes including ARGs and virulence factors were determined by the annotation algorithm of BV-BRC. ARGs were searched by BLAT against the sequences recorded in Comprehensive Antibiotic Resistance Database (CARD) (Alcock et al., 2020; Kent, 2002). Then, SeqSero was used to identify the serotypes of the *S. enterica* isolates (Zhang et al., 2015). MOB-suite was used for classifying the incompatibility group of the MGEs (Robertson & Nash, 2018). The six long-read sequences were compared to each other by aligning all six PacBio HiFi raw reads to six assembled genomes respectively by minimap2 (H. Li, 2018). SNP identification was achieved by aligning the 40 short-read sequencing data of *S. enterica* with the long-read sequence of F6B-2 using breseq (Deatherage & Barrick, 2014). After SNP identification, the *S. enterica* isolates isolated from the same water filter sample with identical genomes were identified as clones. These clones were excluded from downstream analysis. The SNP identification results were plotted as a grape tree by GrapeTree (Zhou et al., 2018), using the authors' online tool on 11 March 2023. At the time of revision (August 2023), the backend server of GrapeTree was not responding. We, therefore, re-plotted the grape tree manually using Microsoft PowerPoint following the same proportion and scale of the original GrapeTree graph.

2.3. In Silico Analysis for Potential Linkages Between the *S. enterica* Isolates Obtained in This Study and Publicly Available *S. enterica*

We found samples with similar whole-genome sequences as this study's isolates by using the long-read PacBio HiFi sequence of F6B-2 as the query to search against all bacterial genomes in the BV-BRC database (Olson et al., 2023). The BV-BRC search engine Similar Genome Finder calculated the mash distances between the query sequence and the online sequences (Ondov et al., 2016). Then, the results were sorted from the lowest mash distance to the highest. The lower the mash distance is, the more similar the two genomes are. Similar Genome Finder was run to collect a maximum of 50 hits targeting all publicly available bacterial and archeal genomes. Thirty-three genomes were returned in Similar Genome Finder. We then collected their Sequence Read Archive (SRA) from NCBI.

For these 33 genomes, a pipeline called "Spine-Nucmer-SNPs" (https://github.com/Alan-Collins/Spine-Nucmer-SNPs) written by Dr. Alan J. Collins was used to extract the core genomes and the SNPs in the core genomes. This pipeline combined Spine (Ozer et al., 2014), Nucmer (Marçais et al., 2018), and customized codes to extract SNPs from the core genomes shared in a certain group of genomes. To better reflect the similarity in chromosomes rather than the whole genomes, in the core genome identification step with Spine, only the chromosome part of the F6B-2 whole genome sequence was used as the reference genome. Then, Nucmer was used to align the 34 chromosomal core genomes (i.e., F6B-2 and 33 other *S. enterica* from Similar Genome Finder) to find the location of the SNPs. The identified SNPs were summarized by customized codes into an aligned fasta file. A distance tree of the SNPs extracted from this group of core genomes was plotted by ggtree based on the aligned fasta (G. Yu et al., 2017).

To find previously recorded bacterial genomes that share similar MGEs with the *S. enterica* collected in this study, we used the sequences of the three MGEs (i.e., the IncFIB plasmid, the IncFII plasmid, and the phage) assembled in the long-read PacBio HiFi sequence of F6B-2 as the queries to run nucleotide BLAST on NCBI against the nucleotide collection database on 5 September 2022. The max target sequences were set as 5,000, which was the highest max target number that can be reached in the web edition of nucleotide BLAST, to include as many sequences as possible. Only the top BLAST hits with identity and query coverage greater than 99% were collected for the downstream analyses. No hits met these criteria in the nucleotide BLAST for the suspected phage sequence. With the top BLAST hits returned in the two plasmid nucleotide BLAST runs (i.e., the IncFIB plasmid as the query and the IncFII plasmid as the query), we traced back to the bacterial whole genome SRA records that contain those BLAST hits. All top hits were from *S. enterica* genomes. These publicly available *S. enterica* genomes were divided into two groups by their BLAST queries: IncFIB plasmid hosts (eight genomes) and IncFII plasmid hosts (10 genomes). No *S. enterica* genome overlapped in these two groups.

The matrices of pairwise average nucleotide identity (ANI) were generated by fastANI for plasmids and chromosomes. For each plasmid ANI matrix, we included the assembly of the plasmid in F6B-2 and the plasmid sequences of the other *S. enterica* which were collected in its corresponding nucleotide BLAST. For example, the ANI matrix for the IncFIB plasmid included the IncFIB plasmid from F6B-2 and the eight plasmid sequences in the other IncFIB plasmid hosts (i.e., the eight *S. enterica* in the top hits of IncFIB plasmid BLAST). The plasmid sequences of those online records were directly downloaded from GenBank.

As for the chromosome ANI matrix, we included not only F6B-2 and the two groups of plasmid hosts but also the 33 *S. enterica* genomes obtained by Similar Genome Finder that shared highly similar whole-genome sequences with F6B-2. Four *S. enterica* genomes obtained by Similar Genome Finder overlapped with the IncFIB plasmid hosts. The total number of this big group of *S. enterica* was 48. Because most online publicly available *S. enterica* were assembled by Illumina, their chromosomes were not directly identified. We used the chromosome sequence of F6B-2 as the reference in Spine to extract core genomes of all these 48 *S. enterica* genomes to represent their chromosome sequences to the maximum extent. Then, fastANI was run on the chromosomal core genomes extracted by Spine to generate the chromosome ANI matrices for IncFIB plasmid hosts and IncFII plasmid hosts. The chromosomal core genome of F6B-2 was also included in both ANI matrices. The ANI values of the plasmids and the ANI values of the chromosomal core genomes were compared statistically by two-sample Kolmogorov-Smirnov tests in OriginPro 2021b, because their distributions could not pass the normality test. A p-value of 0.05 was set as the threshold of significant difference. See Figure S4 in Supporting Information S1 for the overall methodology of Similar Genome Finder and BLAST search.

2.4. In Silico Analysis for Potential Linkages Between the Swine Farms and the *S. enterica* Isolates Obtained in This Study

Swine-source *S. enterica* that were isolated in NC and Georgia (GA) in the NCBI BioSample database were searched and collected by "*Salmonella enterica*" [Organism] as input keywords. Then "(NC OR GA) AND (pig OR swine)" was applied to the search results to yield 870 accession numbers on 21 September 2022. These accession numbers were used to search for metadata by NCBI Entrez (Sayers et al., 2022). A second screening was conducted for the metadata to remove 21 samples with wrong matches in geographical sampling locations. The number of NC and GA swine source *S. enterica* BioSamples that passed the metadata screening was 849. Genome sequences of these samples were retrieved by pysradb (Choudhary et al., 2019), which gives the sequence read achieve identification numbers (SRA run IDs). For those genomes with multiple SRA run IDs, we selected the ones with the highest quality (i.e., long-read run if available, or the run with the highest number of spots and bases). *S. enterica* genome assembling was achieved by the BV-BRC platform, which includes a read-trimming step to remove adapters and low-quality reads from raw sequencing data (Bankevich et al., 2012; Gurevich et al., 2013; Koren et al., 2017; Wick et al., 2015, 2017). After this trimming step, short-read and long-read sequencing data were assembled by Unicycler (Wick et al., 2017) and Canu (Koren et al., 2017), respectively. After genome assembly, the assemblies with N50 lower than 50,000 bp were excluded from the downstream analyses based on the QC report generated by QUAST (Gurevich et al., 2013). The serotypes of the *S. enterica* genomes were determined by SeqSero (Zhang et al., 2015). In total, 808 swine source *S. enterica* assemblies passed the quality screening.

A pilot study was performed to determine the phylogenetic relationship by analyzing the genome sequence of one isolate F6B-2 collected in this study, 808 swine source *S. enterica*, and the *S. enterica* reference chromosome genome NC_003197.2. We used Spine (Ozer et al., 2014) to identify the core genomes, which are the DNA sequences shared among all genomes within this group. Because we found only 230 kbp of the 5.2 Mbp (4.6%) whole-genome length as the core genome for F6B-2 by Spine, we applied another method proposed by Worley et al. (2018) to extract core genomes with higher precision. A nucleotide BLAST database was built using 857 *S. enterica* genomes, including F6B-2 collected in this study, 808 swine source *S. enterica* collected in NC and GA between 1999 and 2022, and 48 *S. enterica* collected in both Similar Genome Finder and nucleotide BLAST using MGEs of F6B-2 as query (described in detail in the above sections). The coding sequences (CDS) in the *S. enterica* reference chromosome genome NC_003197.2 were extracted (Worley et al., 2018). Then tBLASTn was performed for all CDS against the database made by all assemblies, with an e-value threshold of 0.001. The tBLASTn results were then screened by a custom Python script to select CDS that had hits in all genomes. A few genomes that shared an extremely low number of CDS with other genomes (e.g., hundreds vs. thousands) were identified in this step and excluded to ensure at least 3,000 of the 4,446 CDS in NC_003197.2 could be included in the second screening step. Then, a second screening by a custom Python script was conducted to select the hits with coverage and identity greater than 50%. If a CDS could not hit against all genomes after the second screening, the CDS would also be excluded from the core genome. After the two CDS screening steps, 710 *S. enterica* genomes with 2,996 shared CDS were selected for building the phylogenetic tree (Table S1). All CDS that passed the screening were aligned separately by MUSCLE (Edgar, 2004). The aligned CDS were then cascaded to form the aligned core genomes of all *S. enterica* genomes. The SNPs in the core genomes were extracted by snp-sites (Page et al., 2016). A distance tree of the SNPs in the core genomes was plotted by FastTree 2 using the extracted SNPs using the Le-Gascuel 2008 model and rescaled the branch lengths to optimize the gamma20 likelihood after optimizing the tree under the CAT approximation (Price et al., 2010). The distance tree with the metadata of the *S. enterica* genomes was plotted by iTOL (Letunic & Bork, 2021).

A machine learning tool called gplas was used to identify the MGEs in Illumina short-read sequences of the swine source genomes (Arredondo-Alonso et al., 2020). PlasFlow was the classifier in the gplas run (Krawczyk et al., 2018). The potential MGEs were output by gplas as DNA sequence bins. After obtaining all potential MGEs in swine source *S. enterica*, mash distances were calculated between every two bins identified by gplas and the three MGEs in our *S. enterica* genomes, to determine their similarity (Ondov et al., 2016). A network was plotted by Cytoscape using the mash distance threshold of 0.0026, which means more than 900/1,000 k-mer counts were shared between the two sequences (Killcoyne et al., 2009). A similar threshold of mash distance was used in a previous study to reconstruct the plasmidome of *Listeria monocytogenes* (Palma et al., 2020). See Figure S5 in Supporting Information S1 for the overall methodology of in silico analysis for swine source *S. enterica*.

2.5. In Silico Analysis for ARG Locations of Swine and Environmental *S. enterica* That Were Isolated in NC

To have a better understanding of the location of ARGs in the genomes of *S. enterica* isolated from swine and environmental sources, we searched for *S. enterica* genomes in the NCBI database with available PacBio sequencing reads. For NC environmental *S. enterica*, we first used the keywords "*Salmonella enterica*" [Organism] to search in the NCBI BioSample database. Then a filter of keywords "(USA NC)AND(water OR river OR creek OR pond OR soil OR forest)" was applied to the search results. There were 622 BioSamples that passed the filter. For NC swine *S. enterica*, the 808 *S. enterica* genomes that passed all quality screening steps were collected. The sequencing platforms were searched among these 1430 *S. enterica* BioSamples using pysradb (Choudhary et al., 2019). The PacBio-sequenced genomes were then re-assembled using trycycler with the same method described previously (Wick et al., 2021). The genome assemblies were submitted to BV-BRC for annotation, to determine the location of ARGs (Brettin et al., 2015; Olson et al., 2023).

3. Results

3.1. There Is a Low Genetic Diversity Amongst *S. enterica* Isolates, Suggesting They Were From a Stable and Persistent Source

S. enterica was isolated successfully from five flooded sites (F1, F3, F6, F7, and F12) and five unflooded sites (U1, U2, U5, U6, and U8) (Figure 1a). No *S. enterica* was isolated from the negative control site U4. Four *S. enterica* isolates were collected from each site, for a total of 40 isolates for analysis. We selected six of our 40 isolates as references for PacBio HiFi whole genome sequencing. Of these, the F6B-2 genome was the only long-read genome whose chromosome could be circularized during assembly, while other genomes may not be fully sequenced without errors. F6B-2 sequencing revealed the presence of one chromosome (5.02 Mbp), two plasmid-like elements (IncFIB; 123.79 kbp and IncFII; 26.03 kbp), and one phage-like element (52.21 kbp) (Figure S6 in Supporting Information S1). When mapping the PacBio sequencing reads of the other five isolates onto F6B-2, no insertions or deletions were identified.

Using this F6B-2 as a reference, we mapped short-read Illumina sequences from the 40 isolates to identify SNPs because SNPs analysis can often identify the phylogenetic relatedness of isolates. We identified a relatively low number of SNPs: five were found in six of the 40 isolates (Table S2 in Supporting Information S1). According to the SNP analysis, we excluded the potential clones for downstream analyses which were identified as the isolates with identical genomes that were isolated from the same water filter sample. One to three independent *S. enterica* genomes were kept for each site, resulting in 18 total independent *S. enterica* genomes. We used a grape tree graph to visualize this finding (Figure 1b), following the method developed previously (Zhou et al., 2018). This finding implies that these 18 isolates are highly related and stable. The connections of the nodes of U1C-1, U2A-3, and U6A-1 in the grape tree plot show the changes in SNPs among them, suggesting their phylogenetic relationship as a lineage. There was no obvious correlation between the presence of the SNPs and the geographical location of the sample or the presence of flooding. It should be noted that the culturing method used to grow and isolate *S. enterica* in this study will select dominant *S. enterica* strains, while rare variants might be missed. This suggests we have isolated the dominant strain shared in each location but may have under-sampled rare variants.

3.2. Publicly Available *S. enterica* Genomes That Shared the Highest Overall Similarity to Our Isolates Were Collected Mainly From Environmental Sources

We queried the F6B-2 whole-genome sequence (including both the chromosome and the chromosome-independent MGEs) against all publicly available bacterial and archeal genomes to find genomes with the highest overall similarity to F6B-2, using the Similar Genome Finder tool of BV-BRC. No records of bacterial genomes identical to F6B-2 were found in this search. Thirty-three *S. enterica* whole genomes with the lowest mash distances (i.e., the highest overall similarity) were returned in the search and were accessible through SRA (Table S3). Within these 33 *S. enterica* genomes, the mash distances ranged from 0.0018 to 0.0021, corresponding to 936/1,000 to 919/1,000 k-mer counts. Using the SRA sequencing reads, we assembled the genomes of 33 strains and determined the backbone of their chromosomal core genomes shared by F6B-2 *S. enterica* 4.49 Mbp in length (89.4% of the 5.02 Mbp sequence) (Table S4). We then estimated a phylogeny shown in Figure 2 using the aligned chromosomal core genomes. This phylogenetic tree showed that the 34 core genomes were divided into three well-resolved clades. Clade I includes the F6B-2 isolate from this study, three isolates from NC (SRR10740290,

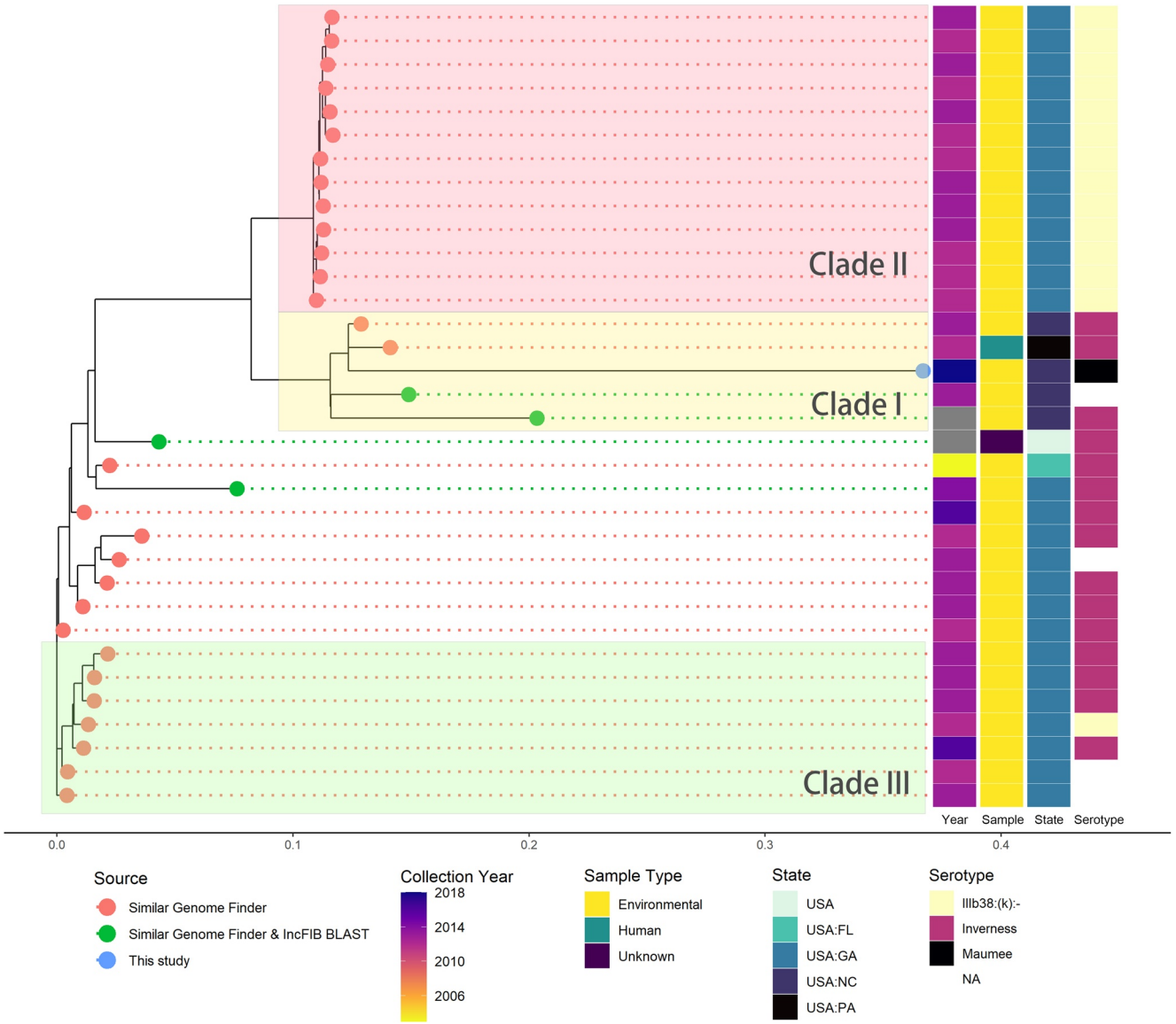

Figure 2. Comparison of genomes of isolates from this study to other published sequences, focusing on single nucleotide polymorphisms in chromosomal core genes. Isolates from this study were in the same clade (Clade I) as all three other North Carolina environmental isolates collected by Similar Genome Finder, as well as a PA human clinical isolate. The GA and FL isolates were divided into other clades. Clade II contained 13 GA isolates; Clade III contained seven GA isolates. The colors of the nodes show the different collection sources of the genomes (blue: in this study, red: found only by Similar Genome Finder, green: found by both Similar Genome Finder and IncFIB plasmid BLAST). The heatmaps on the right show the year of sample collection, the sample type, the geographical location of the sample collection, and the serotype of the isolates. Serotypes were determined based on the White-Kauffmann-Le scheme.

SRR11317976, and SRR5366674), and one isolate from Pennsylvania (PA) which is a human clinical sample (SRR1509573). Clade II members consisted of 13 isolates from GA that are all *S. enterica* serovar IIIb38:(k):- and were all from environmental sources. Clade III members were also from GA and isolated from environmental samples. In this case, four isolates were serovar Inverness. Nine genomes did not fall into a clade and included a sample from Florida (FL). With the exception of one unknown sampling location, this group of isolates were collected from environmental sources. It was striking that the strains that are closest to the F6B-2 strain were predominantly collected from the environment, suggesting that our isolate origins are not from animals, but from the environment.

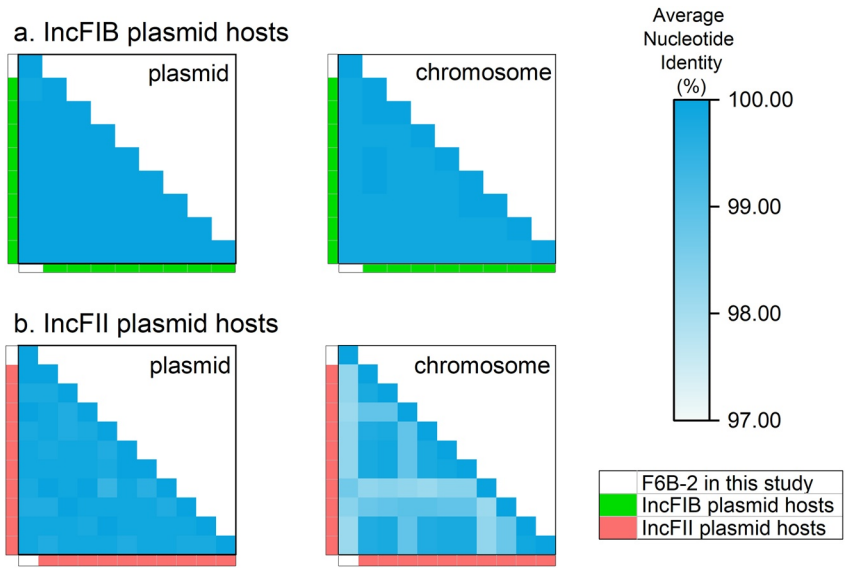

Figure 3. Evidence of horizontal gene transfer of the IncFII plasmid but not the IncFIB plasmid. (a) The average nucleotide identity (ANI) matrices of the IncFIB plasmids and the chromosomal core genomes among publicly available IncFIB plasmid hosts and F6B-2 collected in this study. (b) The ANI matrices of the IncFII plasmids and the chromosomal core genomes among publicly available IncFIB plasmid hosts and F6B-2 collected in this study. The length of the chromosomal core genome backbone of this large group was 3.47 Mbp of the 5.02 Mbp query F6B-2 chromosome sequence.

3.3. Differing Relationships Between Plasmid and Chromosome Compartments Suggest Plasmid Mobility

All 18 *S. enterica* isolates from this study possessed two plasmids (IncFIB and IncFII) and one phage-like element. The two plasmids and the phage-like element do not share similarities with each other. We were interested in determining if the IncFIB and IncFII plasmids were transferred either horizontally or vertically among *S. enterica* strains. We performed a web-based version of the nucleotide BLAST search against each plasmid on 5 September 2022. We identified eight matches to IncFIB and 10 matches to IncFII with greater than 99% identity and coverage. None of the matches were above 90% coverage of the phage-like element, suggesting it is a novel phage. Based on BLAST accession IDs from each of the 18 plasmid hits with high identity, we downloaded the associated host whole genomes to better understand the origins and method of spread for each plasmid. All 18 host genomes were *S. enterica.* The length of the chromosomal core genome backbone of this large group was 3.47 Mbp of the 5.02 Mbp query F6B-2 chromosome sequence (Table S5). The chromosomal core genome ANIs ranged from 99.83% to 99.88% for IncFIB plasmid hosts (Figure 3a, right), and from 98.14% to 98.66% for IncFII plasmid hosts (Figure 3b, right). As a comparison, we also calculated the ANIs among the IncFIB and IncFII plasmids. The ANIs among IncFIB plasmids were from 99.89% to 99.94% (Figure 3a, left) and those among IncFII plasmids were from 99.69% to 99.95% (Figure 3b, left). The ANI values between F6B-2 and IncFIB plasmid hosts' chromosomal core genomes were significantly higher than those for IncFII plasmid hosts' core genomes ($p < 0.001$). This finding suggests that the chromosomes of the hosts of the IncFIB plasmid were more related to F6B-2 (collected in this study) than to all other IncFII plasmid hosts and suggests that the IncFII plasmids are more promiscuous among environmental *S. enterica* strains. In addition, the ANI values among the IncFII plasmids were significantly higher than the ANI values among the chromosomal core genomes of the IncFII plasmid hosts ($p < 0.001$) while those among the IncFIB plasmid hosts were not significantly different. These findings suggested that the IncFII plasmid may be horizontally transferred among *S. enterica*, while the IncFIB plasmid was more likely to be vertically transferred.

The 18 *S. enterica* IncFIB and IncFII plasmid hosts collected in plasmid nucleotide BLAST have similar locations and dates of collection, and most of them, except for one unknown and one lab strain of IncFIB plasmid hosts, were from environmental sources, as shown in Figure S7 in Supporting Information S1. It is also worth noting that four of the eight IncFIB plasmid hosts were the same strains as those *S. enterica* obtained in Similar Genome Finder (Figure 2). In contrast, none of the hosts for the IncFII plasmids are from hosts that are similar

to the strains we isolated. The *S. enterica* serovar hosts of the IncFII plasmid versus the IncFIB plasmid did not overlap. These findings further supported the theory that the IncFII plasmid was horizontally transferred while the IncFIB plasmid was vertically transferred.

3.4. *S. enterica* Genomes Isolated From Our Study Did Not Share Similar Properties With *S. enterica* Isolated Previously From Swine Sources in NC or GA

Data in Figure 2 suggested that the 18 *S. enterica* isolates may not be from animal sources. To explore this further, we compared the core genome of this study's F6B-2 with the *S. enterica* core genomes collected from swine or pork samples in NC and GA from the Sequence Read Achieve. We also included the *S. enterica* genomes from Similar Genome Finder and nucleotide BLAST for comparison. There were in total 710 *S. enterica* genomes in this set, including the F6B-2 collected in this study, that passed all screening steps for core genome searching mentioned in Section 2.4 (Table S1). Figure 4 shows that F6B-2 *S. enterica* did not closely cluster with any other swine-source *S. enterica*. In contrast, swine-source *S. enterica* were clustered into a few different clades. Such a separation supports the finding that our *S. enterica* isolates were a distinct lineage from the swine isolates and are likely environmental. Figure 4 does identify three environmental *S. enterica* that clustered with 94 swine *S. enterica* were IncFII plasmid hosts, but the IncFII plasmid was not found in any of the 94 swine *S. enterica* genomes. This supports the finding of the potential environmental source of the IncFII plasmid, which could be horizontally transferred among environmental *S. enterica*.

Additional data that supported the lack of linkage between this study's isolates and swine source isolates came from analyzing the MGEs between this study's isolates and swine source isolates. A network was built for the mash distance results of the MGEs (Figure S8 in Supporting Information S1). Each node represented an MGE from a *S. enterica* genome, and an edge would connect two nodes if the two MGEs had a mash distance below 0.0026. Thirteen clusters of nodes connected by edges were formed for MGEs from swine sources. None of the three MGEs in this study's isolates were included in the 13 clusters, suggesting a low chance of horizontal gene transfer of the MGEs occurring between our isolates and those from swine sources.

3.5. Analysis of ARG Location in Chromosomes Versus Plasmids Also Suggests a Reservoir Other Than Swine Farms for This Study's *S. enterica*

We used CARD to search for ARGs in the 18 genomes of this study's *S. enterica* isolates (Alcock et al., 2020; Kent, 2002). We identified 65 ARGs and all were located in the chromosomal genomes.

ARGs can be in different locations in *S. enterica* genomes (in the bacterial chromosome or in MGEs). High mobility of ARGs has been reported in farm-related environments to suggest that ARGs from human or animal sources of *S. enterica* are spread via horizontal gene transfer of MGEs (Checcucci et al., 2020; Mazhar et al., 2021; Yang et al., 2020; Zhu et al., 2013). To examine the locations of ARGs in the genomes of swine and environmental isolates, we analyzed *S. enterica* genomes from 13 swine and 18 environmental sources in NC. We chose these genomes because they were sequenced by PacBio technology, which would provide resolution of the location of ARGs. We identified chromosome-independent MGEs from these 35 samples using the tree generated for contig clustering in Trycycler. Table 1 shows these results. Among the 13 swine-source *S. enterica*, 12 contained at least one (and as many as four) MGEs, and ARGs were present in both the chromosome and the plasmids (Table 1). These ARGs include genes resistant to aminoglycoside, tetracycline, sulfonamide, beta-lactam, and chloramphenicol (Table S6). We next examined the sequences of the environmental *S. enterica* strains. Among the 22 environmental *S. enterica* whole genomes available for downloading, four genomes were excluded from further analysis because of low sequencing depths or poor assembly results, leaving 18 remaining. Two of the 18 strains were free of MGEs, based on their assembly results. In contrast to the swine samples, only one of the 16 isolates possessed a susceptible ARG (*vgaC*) with 90% identity on an MGE. In summary, ARGs were present in the chromosomes and plasmids of swine-collected *S. enterica*, but only on the chromosomes of *S. enterica* collected from the environment. In our 18 non-clonal *S. enterica* isolates, we only observed ARGs on the bacterial chromosome, and this is another indication that these 18 isolates are from environmental sources. In summary, the difference in ARG location is consistent with the environmental origin of *S. enterica* collected near the swine farms after Hurricane Florence.

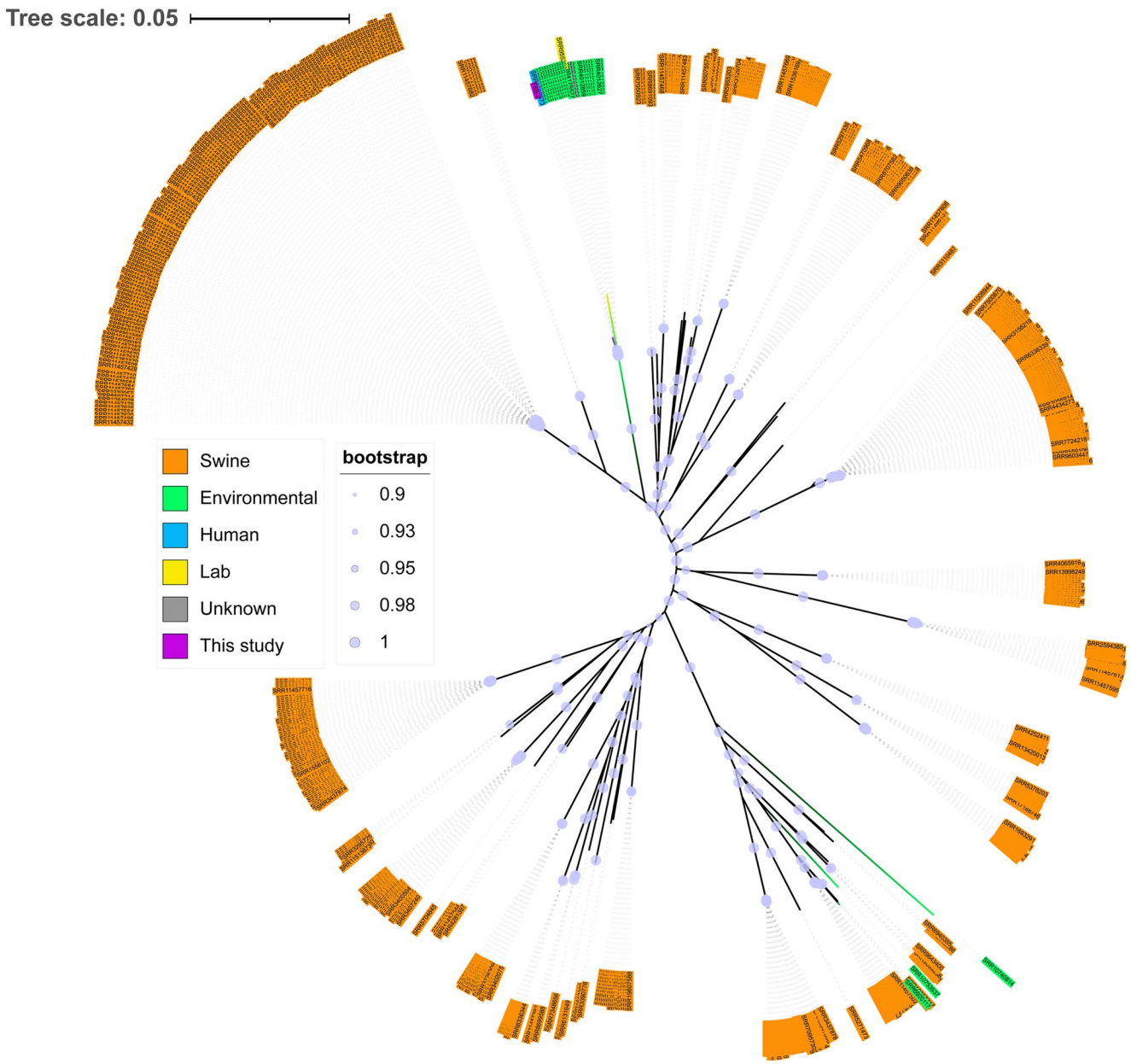

Figure 4. *S. enterica* isolates collected in this study do not cluster with known *S. enterica* swine isolates collected in North Carolina (NC) or GA. The phylogenetic tree is plotted by the single nucleotide polymorphisms in the chromosomal core genomes from *S. enterica* isolated from NC or GA swine sources, the *S. enterica* F6B-2 isolated in this study, and the strains isolated from the environment, humans, labs, or unknown sources that shared either similar whole genomes or similar MGEs with F6B-2. The branches of the environmental isolates are colored in green. The swine-source *S. enterica* included in this tree were collected from 1999 to 2022. The branch lengths are represented by Jukes-Cantor distance. The longer the branch lengths are, the less similar the two core genomes are. The bootstrap values greater than 0.9 are shown as the light purple dots on the tree branches. Bigger dots mean higher bootstrap values. For clarity, only F6B-2 is included in this tree because all 40 isolates collected in this study were nearly identical.

3.6. Antibiotic Resistance Phenotypes of *S. enterica* Isolates From This Study

We examined the susceptibility of six *S. enterica* isolates with different SNPs (F6B-1, F12B-3, U5A-1, U1C-1, U2A-3, and U6A-1) to 20 antibiotics used most frequently in livestock (cows and pigs), using a standard clinical method to determine the MIC (Tables S7 and S8) in which antibiotics can inhibit growth.

Importantly, all six isolates had nearly identical resistance profiles. For example, all of them were resistant to florfenicol, penicillin, gentamycin, neomycin, tiamulin, and erythromycin. The resistance to erythromycin and

Table 1
Swine S. enterica Had More Antibiotic Resistance Genes on Plasmids Than Environmental S. enterica

# Of samples with good sequencing quality	Sample source	Total # of ARGs identified per cell	# Of ARGs on plasmids with >95% identity and coverage
18	This study	65	0
13	Swine (NCBI BioSample)	70–97 (7 assemblies with good-quality chromosomes)	1–11 (12 assemblies with MGEs)
18	Environmental (NCBI BioSample)	65–68 (5 assemblies with good-quality chromosomes)	0 (16 assemblies with MGEs)

penicillin was as expected because the natural resistance to these two drugs has been reported for decades in most gram-negative bacteria (Soares et al., 2012). Gentamycin and neomycin are antibiotics that are usually effective against *Salmonella*, even in cases where a resistant test result is obtained (Mathew et al., 2002). In addition, the antibiotic-resistance phenotypes of this study's *S. enterica* did not fully overlap with the swine-source *Salmonella* collected in the same year. For example, this study's *S. enterica* were susceptible to tilmicosin and clindamycin, but the same susceptibility was not found in any of the 1,052 porcine source *Salmonella* isolates collected in 2018 by the veterinary diagnostic laboratory of Iowa State University ("Bacterial Susceptibility Profiles|Iowa State University," 2023). These findings also suggested a low correlation between this study's *S. enterica* isolates and the swine-source *S. enterica* isolates.

4. Discussion

Precipitation events are occurring at higher frequencies and intensities as a result of climate change. Because precipitation is the main driver of floods, an increase in precipitation results in extensive flooding (Breinl et al., 2021; G. Sofia & Nikolopoulos, 2020). Extreme precipitation events correlate with increased risks of enteric infectious diseases in the US (Jiang et al., 2015; Lee et al., 2019; Saingam et al., 2021; Soneja et al., 2016). For example, Haley et al. (2009) have reported a positive correlation between precipitation and *Salmonella* densities in rural watersheds in Georgia. In NC, there is a positive association between living near concentrated animal feeding operations and emergency room visits due to gastrointestinal illness, and these positive associations became stronger in the weeks after heavy rainfall (Quist et al., 2022b). Flooding usually poses a more serious threat than regular heavy rainfall, because floods can produce standing water and submerge infrastructures allowing a favorable environment for pathogens to thrive. However, the source and spread of harmful bacteria after flooding events remain unknown.

We initially hypothesized that such pathogens would be from flooded swine farms. We tested this hypothesis by analyzing the chromosomes and the MGE sequences of the *S. enteric*a isolates from the flood water samples collected after Hurricane Florence from surface water bodies near swine farms in the flooded region of NC. We used a high-resolution whole-genome sequencing method that can separate the MGE sequences from the chromosome sequences. Analyzing genome sequences of MGEs, in particular the plasmids, has been used to find the sources of foodborne pathogens (Balbuena-Alonso et al., 2022; de Been et al., 2014) because MGEs can have a different evolutionary history and dynamics from their host cells (S. Li et al., 2019).

Our results disproved the initial hypothesis. By assembling chromosomes and MGEs separately but analyzing them collectively, we were able to identify and align core genomes, extract, and compare MGEs, showing that our 18 *S. enterica* isolates were not closely related to swine isolates but rather to environmental isolates. The two plasmids identified in this study, though they were transferred through different routes, were both closely linked to other environmental *S. enterica*, suggesting the exchange of genetic patterns of this study's isolates with the other environmental strains. We also examined the presence of ARGs in chromosomes and plasmids as another means of source tracking. In our isolates, ARGs were found only on the chromosome. We also found our isolates were resistant to multiple antibiotics. The use of antibiotics in the swine industry has caused widespread concerns related to the potential transmission of antibiotic-resistant pathogens from swine to humans (Van Boeckel et al., 2015). A previous study found a higher prevalence of antibiotic resistance in *Staphylococcus aureus* isolates from freshly slaughtered hogs from industrial operations than in isolates from antibiotic-free hog farms (Rhodes et al., 2021). Higher concentrations of swine fecal indicator and more antibiotic-resistant *E. coli* are found in NC from watersheds with commercial hog operations than those without commercial hog operations (Christenson et al., 2022). Although the US swine industry reserves antibiotic use for disease treatment instead

of growth promotion (Muurinen et al., 2021; Rahman et al., 2022), swine fecal indicator Pig2Bac was detectable in some channel water samples in this region from 1 to 5 weeks post Hurricane Florence, suggesting consistent swine-associated contamination during this period (Harris et al., 2021). Our finding that phenotypically and genotypically antibiotic-resistant bacteria spread by Hurricane Florence flooding were from environmental reservoirs is not in contradiction with these previous findings. Our isolation method selects the most dominant and persistent bacteria, while the indicator method could identify rare markers. Nevertheless, our findings suggest a favorable condition for environmental *S. enterica* to persist in surface water bodies surrounding these swine farms. High nutrient loads have been identified in the water near large swine facilities that may allow *S. enterica* in the natural environment to survive and grow (Miralha et al., 2022). The significant role of natural reservoirs for pathogens is likely to increase over time in certain locations due to the warming climate. In addition, higher temperatures and water contaminated with nutrients could be favorable conditions for pathogens to survive and grow (Brown et al., 2020; Miralha et al., 2021; Morgado et al., 2021).

We propose that environmental reservoirs of bacteria (including but may not be limited to *Salmonella* spp.), may be an important source of disease-causing pathogens during and after flooding events. Interestingly, our *S. enterica* strains, their closest relatives, and the other strains containing the IncFII plasmid in databases were predominantly from the natural environment of the southeastern US, suggesting a local genomic signature for environmental *S. enterica*. Very few studies already imply the environmental origin of enteric diseases outbreak in humans. For example, *Salmonella* serovars from environmental origins were detected in patients admitted to hospitals in GA (Lee et al., 2019). Both clinical and environmental non-O1/O139 populations of *V. cholera* were reported to be related to a cholera outbreak in 2010 in Haiti, based on whole-genome sequencing data (Hasan et al., 2012). In contrast, low phylogenetic similarity was observed between Canadian human and animal extended-spectrum β-lactamases-producing *Salmonella* (Bharat et al., 2022). In the future, using the methods described here, the source of disease outbreaks could be identified.

Several limitations of this study are noted. Due to safety and accessibility reasons, the water samples were taken 3 weeks after Hurricane Florence. Thus, we could not collect samples to identify the immediate impact of the flood on the distribution of *S. enterica* in this region. Nearly no precipitation was recorded in the flooded area from the end of Hurricane Florence until our sample dates, as reported by Weather Underground ("Wilmington, NC Weather History|Weather Underground," 2023). Therefore, we do not expect that the impact of precipitation events between the flood and the sampling was significant. Because *Salmonella* can survive in the soil for up to 216 days (Alegbeleye & Sant'Ana, 2023), this delay in sampling is not likely to reduce *S. enterica* concentrations in the water bodies near the swine farms. Harris et al. (2021) examined water samples collected one and 5 weeks after Hurricane Florence in the same region. No significant differences were found for *Salmonella* and *Arcobacter* prevalence between their 1- and 5-week sampling points. Meanwhile, the concentrations of *E. coli* and *Listeria* were significantly higher 5 weeks after the hurricane than the 1-week samples. These authors concluded that the water samples might be more diluted at 1 week versus 5 weeks post-hurricane. When examining pathogen levels in bayou water after Hurricane Harvey, *E. coli* concentrations were slightly lower at 12 days post-hurricane versus 3 days, but both concentrations were above the 7-year average value (P. Yu et al., 2018). Based on these two studies, we expect that *S. enterica* levels detected 21 days post-hurricane to be higher than levels occurring soon after Hurricane Florence.

Another limitation of this study was that all *S. enterica* were isolated after an enrichment step by culturing in LB broth, a step that could potentially select for bacteria that are more adapted to nutrient-rich environments, leading to a lower *S. enterica* diversity. In the future, directly streaking concentrated water samples onto LB plates or loading the water filters to solid selective media to preserve the diversity of the target bacteria can avoid this weakness. Our analysis relied on the online records of bacterial genomes and the findings depend on previous genome sequencing data. Reduced accuracy may also happen when the metadata of the genomes is not completely recorded. We recommend that the sample collection date, the sample collection geographical location, and the sample type be included when new whole genomes are being uploaded. It is critical to have updated, properly curated, and publicly available genome sequencing data, as previously suggested (Hasan et al., 2012).

In this study, we found no simple correlation between the presence of *S. enterica* and proximity to flooded swine farms, or between culturable *Salmonella* and the quantity of the *Salmonella ttrC* genetic marker (data shown previously in Mao et al. (2021)). In the places we sampled, we found *S. enterica* to be from environmental sources. Of course, this does not rule out the possibility that Hurricane Florence caused flooding of septic or

sewer systems, and, in these cases, human waste was the source of *S. enterica*. In addition, we were not able to obtain samples from emergency room visits after the flood. Thus, we are unable to assess if these same environmental *S. enterica* strains caused human disease. One possible direction of future study is to obtain those clinical samples and verify whether the increased gastroenteritis cases are caused by *S. enterica* and whether the disease-causing *S. enterica* can be from environmental sources.

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Data Availability Statement

The data and the customized codes used for building phylogenetic relationships have been uploaded to Zenodo (Nguyen205, 2023) for free downloading. All raw sequencing reads generated in this study can be accessed by NCBI BioProject PRJNA1006749 (https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1006749).

Acknowledgments
We thank Dr. Alan J. Collins for his Spine-Nucmer-SNPs workflow (https://github.com/Alan-Collins/Spine-Nucmer-SNPs) for analyzing SNPs in core genomes. We thank Roy J. Carver Biotechnology Center for conducting Illumina and PacBio sequencing. We acknowledge funding support from the Carl R. Woese Institute for Genomic Biology, the Grainger College of Engineering, the Distinguished Investigator Grant from the Allen Foundation, and EPA Grant R840487. This study has not been formally reviewed by EPA. The views expressed in this document are solely those of the authors and do not necessarily reflect those of the agency. EPA does not endorse any products or commercial services mentioned in this publication.

References

Alcock, B. P., Raphenya, A. R., Lau, T. T. Y., Tsang, K. K., Bouchard, M., Edalatmand, A., et al. (2020). CARD 2020: Antibiotic resistome surveillance with the comprehensive antibiotic resistance database. *Nucleic Acids Research*, *48*(D1), D517–D525. https://doi.org/10.1093/NAR/GKZ935

Alegbeleye, O., & Sant'Ana, A. S. (2023). Survival of *Salmonella* spp. under varying temperature and soil conditions. *The Science of the Total Environment*, *884*, 163744. https://doi.org/10.1016/J.SCITOTENV.2023.163744

Aly, N. A., Casillas, G., Luo, Y. S., McDonald, T. J., Wade, T. L., Zhu, R., et al. (2021). Environmental impacts of Hurricane Florence flooding in Eastern North Carolina: Temporal analysis of contaminant distribution and potential human health risks. *Journal of Exposure Science and Environmental Epidemiology*, *31*(5), 810–822. https://doi.org/10.1038/S41370-021-00325-5

Amirsoleimani, A., Brion, G. M., Diene, S. M., François, P., & Richard, E. M. (2019). Prevalence and characterization of *Staphylococcus aureus* in wastewater treatment plants by whole genomic sequencing. *Water Research*, *158*, 193–202. https://doi.org/10.1016/j.watres.2019.04.035

Arredondo-Alonso, S., Bootsma, M., Hein, Y., Rogers, M. R. C., Corander, J., Willems, R. J. L., & Schürch, A. C. (2020). gplas: A comprehensive tool for plasmid analysis using short-read graphs. *Bioinformatics*, *36*(12), 3874–3876. https://doi.org/10.1093/BIOINFORMATICS/BTAA233

Bacterial Susceptibility Profiles|Iowa State University. (2023). Bacterial susceptibility profiles. Retrieved from https://vetmed.iastate.edu/vdpam/research/disease-topics/bacterial-susceptibility-profiles

Balbuena-Alonso, M. G., Cortés-Cortés, G., Kim, J. W., Lozano-Zarain, P., Camps, M., & del Carmen Rocha-Gracia, R. (2022). Genomic analysis of plasmid content in food isolates of *E. coli* strongly supports its role as a reservoir for the horizontal transfer of virulence and antibiotic resistance genes. *Plasmid*, *123–124*, 102650. https://doi.org/10.1016/J.PLASMID.2022.102650

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., et al. (2012). SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology: A Journal of Computational Molecular Cell Biology*, *19*(5), 455–477. https://doi.org/10.1089/CMB.2012.0021

Bearson, S. M. D. (2021). Annual review of animal Biosciences *Salmonella* in swine: Prevalence, multidrug resistance, and vaccination strategies. https://doi.org/10.1146/annurev-animal-013120

Bharat, A., Mataseje, L., Parmley, E. J., Avery, B. P., Cox, G., Carson, C. A., et al. (2022). One health genomic analysis of extended-spectrum β-lactamase-producing *Salmonella enterica*, Canada, 2012–2016. *Emerging Infectious Diseases*, *28*(7), 1410–1420. https://doi.org/10.3201/EID2807.211528

Bolch, T., Peters, J., Yegorov, A., Pradhan, B., Buchroithner, M., & Blagoveshchensky, V. (2011). Identification of potentially dangerous glacial lakes in the northern Tien Shan. *Natural Hazards*, *59*(3), 1691–1714. https://doi.org/10.1007/s11069-011-9860-2

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120. https://doi.org/10.1093/BIOINFORMATICS/BTU170

Breinl, K., Lun, D., Müller-Thomy, H., & Blöschl, G. (2021). Understanding the relationship between rainfall and flood probabilities through combined intensity-duration-frequency analysis. *Journal of Hydrology*, *602*, 126759. https://doi.org/10.1016/J.JHYDROL.2021.126759

Brettin, T., Davis, J. J., Disz, T., Edwards, R. A., Gerdes, S., Olsen, G. J., et al. (2015). RASTtk: A modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Scientific Reports*, *5*(1), 1–6. https://doi.org/10.1038/srep08365

Brown, C. N., Mallin, M. A., & Loh, A. N. (2020). Tracing nutrient pollution from industrialized animal production in a large coastal watershed. *Environmental Monitoring and Assessment*, *192*(8), 1–17. https://doi.org/10.1007/S10661-020-08433-9

Callaghan, J. (2020). Extreme rainfall and flooding from hurricane florence. *Tropical Cyclone Research and Review*, *9*(3), 172–177. https://doi.org/10.1016/J.TCRR.2020.07.002

Checcucci, A., Trevisi, P., Luise, D., Modesto, M., Blasioli, S., Braschi, I., & Mattarelli, P. (2020). Exploring the animal waste resistome: The spread of antimicrobial resistance genes through the use of livestock manure. *Frontiers in Microbiology*, *11*, 1416. https://doi.org/10.3389/FMICB.2020.01416

Choudhary, S., Andrews, S., & Dale, R. K. (2019). pysradb: A Python package to query next-generation sequencing metadata and data from NCBI Sequence Read Archive. *F1000Research*, *8*, 532. https://doi.org/10.12688/f1000research.18676.1

Christenson, E., Wickersham, L., Jacob, M., & Stewart, J. (2022). A watershed study assessing effects of commercial hog operations on microbial water quality in North Carolina, USA. *Science of the Total Environment*, *838*, 156085. https://doi.org/10.1016/J.SCITOTENV.2022.156085

Clubb, F. J., Mudd, S. M., Milodowski, D. T., Valters, D. A., Slater, L. J., Hurst, M. D., & Limaye, A. B. (2017). Geomorphometric delineation of floodplains and terraces from objectively defined topographic thresholds. *Earth Surface Dynamics*, *5*(3), 369–385. https://doi.org/10.5194/esurf-5-369-2017

Deatherage, D. E., & Barrick, J. E. (2014). Identification of mutations in laboratory-evolved microbes from next-generation sequencing data using *breseq*. *Methods in Molecular Biology*, *1151*, 165–188. https://doi.org/10.1007/978-1-4939-0554-6_12

de Been, M., Lanza, V. F., de Toro, M., Scharringa, J., Dohmen, W., Du, Y., et al. (2014). Dissemination of cephalosporin resistance genes between *Escherichia coli* strains from farm animals and humans by specific plasmid lineages. *PLoS Genetics*, *10*(12), e1004776. https://doi.org/10.1371/JOURNAL.PGEN.1004776

de Kraker, M. E. A., Stewardson, A. J., & Harbarth, S. (2016). Will 10 million people die a year due to antimicrobial resistance by 2050? *PLoS Medicine*, *13*(11), 1002184. https://doi.org/10.1371/JOURNAL.PMED.1002184

Edgar, R. C. (2004). MUSCLE: A multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics*, *5*(1), 1–19. https://doi.org/10.1186/1471-2105-5-113

Flach, C. F., Hutinel, M., Razavi, M., Åhrén, C., & Larsson, D. G. J. (2021). Monitoring of hospital sewage shows both promise and limitations as an early-warning system for carbapenemase-producing Enterobacterales in a low-prevalence setting. *Water Research*, *200*, 117261. https://doi.org/10.1016/J.WATRES.2021.117261

Gurevich, A., Saveliev, V., Vyahhi, N., & Tesler, G. (2013). QUAST: Quality assessment tool for genome assemblies. *Bioinformatics*, *29*(8), 1072–1075. https://doi.org/10.1093/BIOINFORMATICS/BTT086

Haley, B. J., Cole, D. J., & Lipp, E. K. (2009). Distribution, diversity, and seasonality of waterborne salmonellae in a rural watershed. *Applied and Environmental Microbiology*, *75*(5), 1248–1255. https://doi.org/10.1128/AEM.01648-08

Harris, A. R., Fidan, E. N., Nelson, N. G., Emanuel, R. E., Jass, T., Kathariou, S., et al. (2021). Microbial contamination in environmental waters of rural and agriculturally-dominated landscapes following hurricane florence. *ACS ES and T Water*, *1*(9), 2012–2019. https://doi.org/10.1021/ACSESTWATER.1C00103

Harwood, V. J., Staley, C., Badgley, B. D., Borges, K., & Korajkic, A. (2014). Microbial source tracking markers for detection of fecal contamination in environmental waters: Relationships between pathogens and human health outcomes. *FEMS Microbiology Reviews*, *38*(1), 1–40. https://doi.org/10.1111/1574-6976.12031

Hasan, N. A., Choi, S. Y., Eppinger, M., Clark, P. W., Chen, A., Alam, M., et al. (2012). Genomic diversity of 2010 Haitian cholera outbreak strains. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(29), E2010–E2017. https://doi.org/10.1073/PNAS.1207359109

Henriot, C. P., Celle, H., Klaba, V., Biguenet, A., Miège, C., Daval, A., et al. (2023). Effect of a karst system (France) on extended spectrum beta-lactamase (ESBL)-producing *Escherichia coli*. *Water Research*, *230*, 119582. https://doi.org/10.1016/J.WATRES.2023.119582

Hon, T., Mars, K., Young, G., Tsai, Y. C., Karalius, J. W., Landolin, J. M., et al. (2020). Highly accurate long-read HiFi sequencing data for five complex genomes. *Scientific Data*, *7*(1), 1–11. https://doi.org/10.1038/s41597-020-00743-4

Ishii, K., Fujitani, H., Soh, K., Nakagawa, T., Takahashi, R., & Tsuneda, S. (2017). Enrichment and physiological characterization of a cold-adapted nitrite-oxidizing *Nitrotoga* sp. from an eelgrass sediment. *Applied and Environmental Microbiology*, *83*(14), e00549-17. https://doi.org/10.1128/AEM.00549-17

Jang, J., Hur, H. G., Sadowsky, M. J., Byappanahalli, M. N., Yan, T., & Ishii, S. (2017). Environmental *Escherichia coli*: Ecology and public health implications—A review. *Journal of Applied Microbiology*, *123*(3), 570–581. https://doi.org/10.1111/JAM.13468

Jiang, C., Shaw, K. S., Upperman, C. R., Blythe, D., Mitchell, C., Murtugudde, R., et al. (2015). Climate change, extreme events and increased risk of salmonellosis in Maryland, USA: Evidence for coastal vulnerability. *Environment International*, *83*, 58–62. https://doi.org/10.1016/J.ENVINT.2015.06.006

Kent, W. J. (2002). BLAT—The BLAST-like alignment tool. *Genome Research*, *12*(4), 656–664. https://doi.org/10.1101/GR.229202

Killcoyne, S., Carter, G. W., Smith, J., & Boyle, J. (2009). Cytoscape: A community-based framework for network modeling. *Methods in Molecular Biology*, *563*, 219–239. https://doi.org/10.1007/978-1-60761-175-2_12

Koren, S., Walenz, B. P., Berlin, K., Miller, J. R., Bergman, N. H., & Phillippy, A. M. (2017). Canu: Scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Research*, *27*(5), 722–736. https://doi.org/10.1101/GR.215087.116

Krawczyk, P. S., Lipinski, L., & Dziembowski, A. (2018). PlasFlow: Predicting plasmid sequences in metagenomic data using genome signatures. *Nucleic Acids Research*, *46*(6), E35. https://doi.org/10.1093/NAR/GKX1321

Lee, D., Chang, H. H., Sarnat, S. E., & Levy, K. (2019). Precipitation and salmonellosis incidence in Georgia, USA: Interactions between extreme rainfall events and antecedent rainfall conditions. *Environmental Health Perspectives*, *127*(9), 097005. https://doi.org/10.1289/EHP4621

Letunic, I., & Bork, P. (2021). Interactive tree of life (iTOL) v5: An online tool for phylogenetic tree display and annotation. *Nucleic Acids Research*, *49*(W1), W293–W296. https://doi.org/10.1093/nar/gkab301

Li, H. (2018). Minimap2: Pairwise alignment for nucleotide sequences. *Bioinformatics*, *34*(18), 3094–3100. https://doi.org/10.1093/bioinformatics/bty191

Li, S., Zhang, S., Baert, L., Jagadeesan, B., Ngom-Bru, C., Griswold, T., et al. (2019). Implications of mobile genetic elements for *Salmonella enterica* single-nucleotide polymorphism subtyping and source tracking investigations. *Applied and Environmental Microbiology*, *85*(24), e01985-19. https://doi.org/10.1128/AEM.01985-19

Mao, Y., Zeineldin, M., Usmani, M., Uprety, S., Shisler, J. L., Jutla, A., et al. (2021). Distribution and antibiotic resistance profiles of *Salmonella enterica* in rural areas of North Carolina after hurricane florence in 2018. *GeoHealth*, *5*(2), e2020GH000294. https://doi.org/10.1029/2020GH000294

Marçais, G., Delcher, A. L., Phillippy, A. M., Coston, R., Salzberg, S. L., & Zimin, A. (2018). MUMmer4: A fast and versatile genome alignment system. *PLOS Computational Biology*, *14*(1), e1005944. https://doi.org/10.1371/journal.pcbi.1005944

Mathew, A. G., Jackson, F., & Saxton, A. M. (2002). Effects of antibiotic regimens on resistance of *Escherichia coli* and *Salmonella serovar* Typhimurium in swine. *Journal of Swine Health and Production*, *10*(1), 7–13. Retrieved from http://www.aasv.org/shap.html

Mazhar, S. H., Li, X., Rashid, A., Su, J. M., Xu, J., Brejnrod, A. D., et al. (2021). Co-selection of antibiotic resistance genes, and mobile genetic elements in the presence of heavy metals in poultry farm environments. *The Science of the Total Environment*, *755*(Pt 2), 142702. https://doi.org/10.1016/J.SCITOTENV.2020.142702

Miralha, L., Muenich, R. L., Schaffer-Smith, D., & Myint, S. W. (2021). Spatiotemporal land use change and environmental degradation surrounding CAFOs in Michigan and North Carolina. *Science of the Total Environment*, *800*, 149391. https://doi.org/10.1016/J.SCITOTENV.2021.149391

Miralha, L., Sidique, S., & Logsdon Muenich, R. (2022). The spatial organization of CAFOs and its relationship to water quality in the United States. *Journal of Hydrology*, *613*, 128301. https://doi.org/10.1016/J.JHYDROL.2022.128301

Moghadam, S. V., Vadde, K. K., Phan, D. C., Jafarzadeh, A., & Kapoor, V. (2022). Assessing the impact of flooding on bacterial community structure and occurrence of potentially pathogenic bacteria in Texas Rivers after Hurricane Harvey. *Journal of Hazardous Materials Letters*, *3*, 100058. https://doi.org/10.1016/J.HAZL.2022.100058

Morgado, M. E., Jiang, C., Zambrana, J., Upperman, C. R., Mitchell, C., Boyle, M., et al. (2021). Climate change, extreme events, and increased risk of salmonellosis: Foodborne diseases active surveillance network (FoodNet), 2004–2014. *Environmental Health: A Global Access Science Source*, *20*(1), 1–11. https://doi.org/10.1186/S12940-021-00787-Y

Murray, C. J., Ikuta, K. S., Sharara, F., Swetschinski, L., Robles Aguilar, G., Gray, A., et al. (2022). Global burden of bacterial antimicrobial resistance in 2019: A systematic analysis. *The Lancet*, *399*(10325), 629–655. https://doi.org/10.1016/S0140-6736(21)02724-0

Muurinen, J., Richert, J., Wickware, C. L., Richert, B., & Johnson, T. A. (2021). Swine growth promotion with antibiotics or alternatives can increase antibiotic resistance gene mobility potential. *Scientific Reports*, *11*(1), 1–13. https://doi.org/10.1038/s41598-021-84759-9

Nguyen205. (2023). Nguyen205/GeoHealth_Salmonella: GeoHealth (v1.0) [Dataset]. Zenodo. https://doi.org/10.5281/ZENODO.8279911

Niedermeyer, J. A., Miller, W. G., Yee, E., Harris, A., Emanuel, R. E., Jass, T., et al. (2020). Search for *Campylobacter* spp. reveals high prevalence and pronounced genetic diversity of arcobacter butzleri in floodwater samples associated with hurricane florence in North Carolina, USA. *Applied and Environmental Microbiology*, *86*(20), 1–14. https://doi.org/10.1128/AEM.01118-20

Ohl, C. A., & Tapsell, S. (2000). Flooding and human health: The dangers posed are not always obvious. *BMJ British Medical Journal*, *321*(7270), 1167–1168. https://doi.org/10.1136/BMJ.321.7270.1167

Olson, R. D., Assaf, R., Brettin, T., Conrad, N., Cucinell, C., Davis, J. J., et al. (2023). Introducing the bacterial and viral bioinformatics resource center (BV-BRC): A resource combining PATRIC, IRD and ViPR. *Nucleic Acids Research*, *51*(D1), D678–D689. https://doi.org/10.1093/NAR/GKAC1003

Ondov, B. D., Treangen, T. J., Melsted, P., Mallonee, A. B., Bergman, N. H., Koren, S., & Phillippy, A. M. (2016). Mash: Fast genome and metagenome distance estimation using MinHash. *Genome Biology*, *17*(1), 1–14. https://doi.org/10.1186/S13059-016-0997-X

Ozer, E. A., Allen, J. P., & Hauser, A. R. (2014). Characterization of the core and accessory genomes of Pseudomonas aeruginosa using bioinformatic tools Spine and AGEnt. *BMC Genomics*, *15*(1), 1–17. https://doi.org/10.1186/1471-2164-15-737

Page, A. J., Taylor, B., Delaney, A. J., Soares, J., Seemann, T., Keane, J. A., & Harris, S. R. (2016). SNP-Sites: Rapid efficient extraction of SNPs from multi-FASTA alignments. *Microbial Genomics*, *2*(4), e000056. https://doi.org/10.1099/MGEN.0.000056

Palma, F., Brauge, T., Radomski, N., Mallet, L., Felten, A., Mistou, M. Y., et al. (2020). Dynamics of mobile genetic elements of Listeria monocytogenes persisting in ready-to-eat seafood processing plants in France. *BMC Genomics*, *21*(1), 1–20. https://doi.org/10.1186/S12864-020-6544-X

Patel, J. B., Eliopoulos, G. M., Jenkins, S. G., James Lewis, F. S., II, Brandi Limbago, P., Nicolau, D. P., et al. (2016). Performance standards for antimicrobial susceptibility testing performance standards for antimicrobial susceptibility testing suggested citation (pp. 100–125).

Price, M. N., Dehal, P. S., & Arkin, A. P. (2010). FastTree 2—Approximately maximum-likelihood trees for large alignments. *PLoS One*, *5*(3), e9490. https://doi.org/10.1371/JOURNAL.PONE.0009490

Quist, A. J. L., Fliss, M. D., Wade, T. J., Delamater, P. L., Richardson, D. B., & Engel, L. S. (2022a). Hurricane flooding and acute gastrointestinal illness in North Carolina. *The Science of the Total Environment*, *809*, 151108. https://doi.org/10.1016/J.SCITOTENV.2021.151108

Quist, A. J. L., Holcomb, D. A., Fliss, M. D., Delamater, P. L., Richardson, D. B., & Engel, L. S. (2022b). Exposure to industrial hog operations and gastrointestinal illness in North Carolina, USA. *Science of the Total Environment*, *830*, 154823. https://doi.org/10.1016/j.scitotenv.2022.154823

Rahman, M. R. T., Fliss, I., & Biron, E. (2022). Insights in the development and uses of alternatives to antibiotic growth promoters in poultry and swine production. *Antibiotics*, *11*(6), 766. https://doi.org/10.3390/ANTIBIOTICS11060766/S1

Raza, S., Kim, J., Sadowsky, M. J., & Unno, T. (2021). Microbial source tracking using metagenomics and other new technologies. *Journal of Microbiology*, *59*(3), 259–269. https://doi.org/10.1007/S12275-021-0668-9

Rhodes, S., Christenson, E., Nguyen, A., Larsen, J., Price, L. B., & Stewart, J. (2021). Getting ahead of antibiotic-resistant *Staphylococcus aureus* in U.S. hogs. *Environmental Research*, *196*, 110954. https://doi.org/10.1016/J.ENVRES.2021.110954

Robertson, J., & Nash, J. H. E. (2018). MOB-Suite: Software tools for clustering, reconstruction and typing of plasmids from draft assemblies. *Microbial Genomics*, *4*(8), e000206. https://doi.org/10.1099/MGEN.0.000206

Saingam, P., Li, B., Sung, S., & Yan, T. (2021). Immediate impact of Hurricane lane on microbiological quality of coastal water in Hilo Bay, Hawaii. *Environmental Science and Technology*, *55*(5), 2960–2967. https://doi.org/10.1021/ACS.EST.0C07082

Sayers, E. W., Bolton, E. E., Brister, J. R., Canese, K., Chan, J., Comeau, D. C., et al. (2022). Database resources of the national center for biotechnology information. *Nucleic Acids Research*, *50*(D1), D20–D26. https://doi.org/10.1093/NAR/GKAB1112

Sayers, E. W., Cavanaugh, M., Clark, K., Ostell, J., Pruitt, K. D., & Karsch-Mizrachi, I. (2019). GenBank. *Nucleic Acids Research*, *47*(Database issue), D94–D99. https://doi.org/10.1093/NAR/GKY989

Seneviratne, S. I., Zhang, X., Adnan, M., Badi, W., Dereczynski, C., Di Luca, A., et al. (2021). Weather and climate extreme events in a changing climate. In *Climate change 2021: The physical science basis. Contribution of working group I to the sixth assessment report of the intergovernmental panel on climate change* (pp. 1513–1766). https://doi.org/10.1017/9781009157896.013

Soares, G. M. S., Figueiredo, L. C., Faveri, M., Cortelli, S. C., Duarte, P. M., & Feres, M. (2012). Mechanisms of action of systemic antibiotics used in periodontal treatment and mechanisms of bacterial resistance to these drugs. *Journal of Applied Oral Science*, *20*(3), 295–309. https://doi.org/10.1590/S1678-77572012000300002

Sofia, G., Fontana, G. D., & Tarolli, P. (2014). High-resolution topography and anthropogenic feature extraction: Testing geomorphometric parameters in floodplains. *Hydrological Processes*, *28*(4), 2046–2061. https://doi.org/10.1002/hyp.9727

Sofia, G., & Nikolopoulos, E. I. (2020). Floods and rivers: A circular causality perspective. *Scientific Reports*, *10*(1), 5175. https://doi.org/10.1038/S41598-020-61533-X

Soneja, S., Jiang, C., Romeo Upperman, C., Murtugudde, R., Mitchell, C. S., Blythe, D., et al. (2016). Extreme precipitation events and increased risk of campylobacteriosis in Maryland, U.S.A. *Environmental Research*, *149*, 216–221. https://doi.org/10.1016/J.ENVRES.2016.05.021

Tabari, H. (2020). Climate change impact on flood and extreme precipitation increases with water availability. *Scientific Reports*, *10*(1), 1–10. https://doi.org/10.1038/s41598-020-70816-2

Thompson, J. D., Higgins, D. G., & Gibson, T. J. (1994). CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, *22*(22), 4673–4680. https://doi.org/10.1093/NAR/22.22.4673

Usmani, M., Uprety, S., Bonham, N., Jamal, Y., Mao, Y., Sano, D., et al. (2023). Assessment of pathogens in flood waters in coastal rural regions: Case study after Hurricane Michael and Florence. *PLoS One*, *18*(8), e0273757. https://doi.org/10.1371/JOURNAL.PONE.0273757

Van Boeckel, T. P., Brower, C., Gilbert, M., Grenfell, B. T., Levin, S. A., Robinson, T. P., et al. (2015). Global trends in antimicrobial use in food animals. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(18), 5649–5654. https://doi.org/10.1073/PNAS.1503141112

Vaser, R., & Šikić, M. (2021). Time- and memory-efficient genome assembly with Raven. *Nature Computational Science*, *1*(5), 332–336. https://doi.org/10.1038/s43588-021-00073-4

Wick, R. R., Judd, L. M., Cerdeira, L. T., Hawkey, J., Méric, G., Vezina, B., et al. (2021). Trycycler: Consensus long-read assemblies for bacterial genomes. *Genome Biology*, *22*(1), 1–17. https://doi.org/10.1186/S13059-021-02483-Z

Wick, R. R., Judd, L. M., Gorrie, C. L., & Holt, K. E. (2017). Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Computational Biology*, *13*(6), e1005595. https://doi.org/10.1371/JOURNAL.PCBI.1005595

Wick, R. R., Schultz, M. B., Zobel, J., & Holt, K. E. (2015). Bandage: Interactive visualization of de novo genome assemblies. *Bioinformatics*, *31*(20), 3350–3352. https://doi.org/10.1093/BIOINFORMATICS/BTV383

Wilmington, NC Weather History|Weather Underground. (2023). Wilmington, NC weather history. Retrieved from https://www.wunderground.com/history/weekly/us/nc/wilmington/KILM/date/2018-10-8

Worley, J., Meng, J., Allard, M. W., Brown, E. W., & Timme, R. E. (2018). *Salmonella enterica* phylogeny based on whole-genome sequencing reveals two new clades and novel patterns of horizontally acquired genetic elements. *mBio*, *9*(6), 10-1128. https://doi.org/10.1128/MBIO.02303-18

Xylose lysine tergitol 4 (XLT4) agar. (2003). Progress in Industrial Microbiology, 37(C) (pp. 635–637). https://doi.org/10.1016/S0079-6352(03)80106-6

Yang, F., Han, B., Gu, Y., & Zhang, K. (2020). Swine liquid manure: A hotspot of mobile genetic elements and antibiotic resistance genes. *Scientific Reports*, *10*(1), 1–10. https://doi.org/10.1038/s41598-020-72149-6

Yu, G., Smith, D. K., Zhu, H., Guan, Y., & Lam, T. T. Y. (2017). ggtree: An r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods in Ecology and Evolution*, *8*(1), 28–36. https://doi.org/10.1111/2041-210X.12628

Yu, P., Zaleski, A., Li, Q., He, Y., Mapili, K., Pruden, A., et al. (2018). Elevated levels of pathogenic indicator bacteria and antibiotic resistance genes after Hurricane Harvey's flooding in Houston. *Environmental Science and Technology Letters*, *5*(8), 481–486. https://doi.org/10.1021/acs.estlett.8b00329

Zhang, S., Yin, Y., Jones, M. B., Zhang, Z., Kaiser, B. L. D., Dinsmore, B. A., et al. (2015). Salmonella serotype determination utilizing high-throughput genome sequencing data. *Journal of Clinical Microbiology*, *53*(5), 1685–1692. https://doi.org/10.1128/JCM.00323-15/SUPPL_FILE/ZJM999094259SO9.XLSX

Zhou, Z., Alikhan, N. F., Sergeant, M. J., Luhmann, N., Vaz, C., Francisco, A. P., et al. (2018). GrapeTree: Visualization of core genomic relationships among 100,000 bacterial pathogens. *Genome Research*, *28*(9), 1395–1404. https://doi.org/10.1101/GR.232397.117

Zhu, Y. G., Johnson, T. A., Su, J. Q., Qiao, M., Guo, G. X., Stedtfeld, R. D., et al. (2013). Diverse and abundant antibiotic resistance genes in Chinese swine farms. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(9), 3435–3440. https://doi.org/10.1073/PNAS.1222743110