# Genomic Evidence Reveals Numerous *Salmonella enterica* Serovar Newport Reintroduction Events in Suwannee Watershed Irrigation Ponds

Baoguang Li,[a] Scott A. Jackson,[a*] Jayanthi Gangiredla,[a] Weimin Wang,[a] Huanli Liu,[a*] Ben D. Tall,[b] Junia Jean-Gilles Beaubrun,[b] Michele Jay-Russell,[c] George Vellidis,[d] Christopher A. Elkins[a]

Division of Molecular Biology[a] and Division of Virulence Assessment,[b] Center for Food Safety and Applied Nutrition, U.S. FDA, Laurel, Maryland, USA; University of California, Davis Western Center for Food Safety, Davis, California, USA[c]; University of Georgia, Tifton, Georgia, USA[d]

**Our previous work indicated a predominance (56.8%) of *Salmonella enterica* serovar Newport among isolates recovered from irrigation ponds used in produce farms over a 2-year period (B. Li et al., Appl Environ Microbiol 80:6355–6365, http://dx.doi.org /10.1128/AEM.02063-14). This observation provided a valuable set of metrics to explore an underaddressed issue of environmental survival of *Salmonella* by DNA microarray. Microarray analysis correctly identified all the isolates ($n = 53$) and differentiated the *S.* Newport isolates into two phylogenetic lineages (*S.* Newport II and *S.* Newport III). Serovar distribution analysis showed no instances where the same serovar was recovered from a pond for more than a month. Furthermore, during the study, numerous isolates with an indistinguishable genotype were recovered from different ponds as far as 180 km apart for time intervals as long as 2 years. Although isolates within either lineage were phylogenetically related as determined by microarray analysis, subtle genotypic differences were detected within the lineages, suggesting that isolates in either lineage could have come from several unique hosts. For example, strains in four different subgroups (A, B, C, and D) possessed an indistinguishable genotype within their subgroups as measured by gene differences, suggesting that strains in each subgroup shared a common host. Based on this comparative genomic evidence and the spatial and temporal factors, we speculated that the presence of *Salmonella* in the ponds was likely due to numerous punctuated reintroduction events associated with several different but common hosts in the environment. These findings may have implications for the development of strategies for efficient and safe irrigation to minimize the risk of *Salmonella* outbreaks associated with fresh produce.**

*S*almonella spp. are ubiquitous inhabitants of the environment and intestinal tracts of animals; annually, they cause approximately 1.3 billion cases of nontyphoidal salmonellosis worldwide (2). In the United States, approximately 1.2 million cases of salmonellosis occur each year (3), and its economic burden is more than a billion dollars (4). Moreover, despite enormous efforts and advances in the detection and monitoring of foodborne pathogens during the past decade, human infections caused by *Salmonella* still remain a challenging health problem globally (5).

The genus *Salmonella* comprises a diverse group of pathogens that have evolved to survive in a wide range of environments and hosts (6). There are more than 2,600 serovars (7), and the majority (over 1,500) of them belong to *Salmonella enterica* subsp. *enterica*, which encompasses most of the serovars that are of greatest clinical relevance (8). *Salmonella* spp. traditionally have been regarded as a foodborne pathogen of animal origin; however, water has increasingly been recognized as a source and vector of *Salmonella* (9). Surface water obtained from irrigation ponds is a higher-risk source than water obtained from wells and other sources (10). However, farmers continue to use it as a water source because irrigation is an essential component of fruit and vegetable production, and surface water is still the most feasible and economic irrigation choice (10). On 16 January 2013, the Food and Drug Administration (FDA) issued a proposed rule to establish science-based standards for the growing, harvesting, packing, and holding by domestic and foreign farms of produce consumed in the United States (11) as directed by the 2011 Food Safety Modernization Act (FSMA). On 29 September 2014, FDA issued a supplemental notice of proposed rulemaking, reopening the rulemaking

for a limited number of key issues and taking into account public input on the 2013 proposed rule (12). The proposed rule focuses on identifying recognized routes of microbial contamination of produce on the farm, and one of its components includes agricultural water used during production, harvest, and postharvest activities. Agricultural water, which is defined, in part, "as any water that is intended to, or likely to, contact, or cover produce or food-contact surfaces" (11). Under the proposed Produce Safety Rule of

the FSMA, farmers would be required to monitor and assess the microbial quality of agricultural water such as irrigation water and take corrective actions, if needed, to ensure the safety of fresh produce (12).

The facts that *Salmonella* spp. are ubiquitous in the environment (13) and that most *Salmonella* serovars are zoonotic pathogens (14) underscore the complex nature of their presence in the environment and transmission to humans. In spite of the growing evidence indicating contaminated irrigation water as a source of enteric pathogens during numerous produce-related outbreaks (15), the genomic diversity and distribution of *Salmonella enterica* in irrigation surface water are not well known. A better understanding of such topics would be beneficial for the development of effective strategies to mitigate the risk of produce contamination by *Salmonella*, and insight into the persistence of *Salmonella* spp. in irrigation water would facilitate the development of strategies for efficient and safe irrigation procedures to minimize preharvest *Salmonella* contamination of produce. To achieve this objective, the use of robust and high-resolution subtyping methodologies is imperative.

Currently, conventional methods for identification and subtyping of *Salmonella* are based on cultural, serological, and biochemical properties of these foodborne pathogens. These methods work but are not without drawbacks, e.g., being time-consuming and labor-intensive and lacking in discriminatory power. Such drawbacks of these conventional methods severely limit the ability of public health officials to attribute disease to food sources in epidemiological and foodborne outbreak investigations. In the last few decades, numerous DNA-based subtyping methods, such as pulsed-field gel electrophoresis (PFGE) (16), multilocus sequence typing (MLST) (17), repetitive element PCR (Rep-PCR) (18), multilocus variable-number tandem-repeat analysis (MLVA) (19), and the use of clustered regularly interspaced short palindromic repeats (CRISPR) (20), have been used to characterize the genetic relatedness and diversity of outbreak pathogens. Each of these subtyping methods has its own advantages and drawbacks in terms of sensitivity, robustness, specificity, cost, and speed. In general, PFGE is the most commonly used method, and it is widely considered the gold standard for subtyping *Salmonella enterica*. However, limits with PFGE include difficulties in serovar determination of polyphyletic patterns (21).

Furthermore, most of these above-mentioned subtyping methods provide little information about genetic traits, such as those associated with pathogenicity, virulence, and antimicrobial resistance, possessed by pathogenic strains to confirm the serovar predication for closely related strains associated with outbreaks. For instance, *S.* Newport and *S.* Enteritidis are notoriously enigmatic in regard to attribution of contamination sources within outbreaks (22, 23), and these two serovars are among the most common serovars associated with human and animal salmonellosis worldwide (24). Rapid and accurate serovar prediction in epidemiological investigations will help in the development of strategies against these foodborne pathogens (21). In such cases, high-resolution DNA microarray genotyping becomes a particularly useful tool to attribute suspected pathogens from outbreaks. Moreover, microarray analysis can also provide a quick snapshot of the genomic content of the pathogens; such information, in turn, will augment the detection, identification, and attribution of strains associated with outbreaks. Therefore, the aims of this study were 2-fold: (i) to develop and validate a DNA microarray for
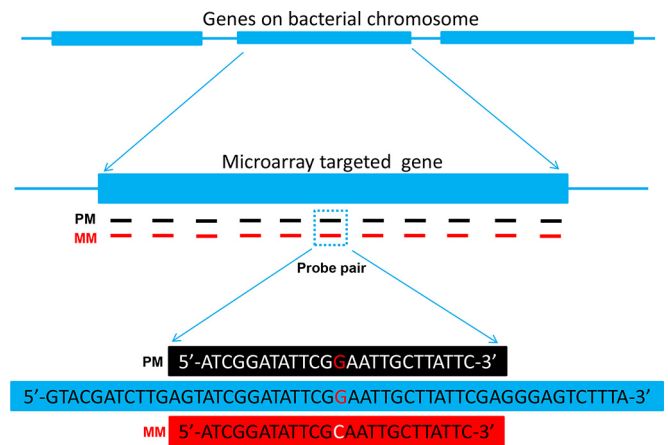


FIG 1 Probe set design used in the microarray.

subtyping *Salmonella enterica* and (ii) to take advantage of the high discriminatory power that DNA microarray offers to investigate the 2-year predominance of *Salmonella* Newport in surface irrigation pond environments in the southeastern United States and to gain insight into a underaddressed issue in the field of microbiology and food safety, i.e., survival of *Salmonella enterica* in environmental water (25, 26), in particular, in irrigation pond water used for produce production.

## MATERIALS AND METHODS

**Sampling locations and schedules.** This is a follow-up study of previous work which was carried out from July 2011 to September 2013 within the Suwannee River basin area, as described by Li et al. (1). In the present study, the same set of *Salmonella* isolates was used, and for context, the background of the *Salmonella* isolates is given here. Briefly, the isolates were recovered from water samples ($n = 170$) which were collected monthly from 10 selected irrigation ponds associated with Suwannee watershed produce farms over a 27-month period of time. The ponds were coded BB, VH, RT1, RT, CC1, SC, NP, LV, CC2, and MD1. Specifically, five ponds, SC, NP, LV, CC2, and MD1, were sampled from July 2011 to July 2013, and the other five ponds, BB, VH, RT1, RT, and CC1, were sampled from July 2011 to February 2012; the geographic distances between the ponds ranged from 1.8 to 180 km (1).

**Microarray design.** The DNA microarray is a custom Affymetrix (Santa Clara, CA) array designed by FDA to determine the gene content of individual strains of foodborne pathogens and was developed using genomes of *Salmonella enterica* ($n = 38$), *Escherichia coli* ($n = 27$), *Shigella* spp. ($n = 10$), and *Vibrio cholerae* ($n = 10$) (see Table S1 in the supplemental material). In total, the DNA microarray covers over 80,000 unique genes representing the pan-genomes of these four foodborne pathogens, including known antibiotic resistance and virulence genes. The genomes and genes represented on the array are specifically described in Table S1 in the supplemental material. Each gene target is represented on the array by 22 individual oligonucleotide probes (25-mer). The 22 probes together make up 11 probe pair sets, and each probe pair set includes 11 perfect-match (PM) probes and 11 mismatch (MM) probes per gene. A PM probe matches the reference sequence perfectly, while an MM probe contains a single nucleotide mismatch in the central (13th) position of the probe (Fig. 1). Incorporation of MM probes in the probe pair set was to allow for measurement of, and thereby correction for, nonspecific hybridization.

**Microarray assay.** *Salmonella* isolates were grown overnight in 3 ml of Luria broth at 37°C in a shaking incubator set at 150 rpm. Genomic DNA was isolated from 2 ml of overnight culture using the chemistries described for the Qiagen DNeasy tissue kit (Qiagen, Valencia, CA) in conjunction with the automation of the QiaCube instrument (Qiagen), fol-

lowing the manufacturer's instructions. Amounts of 10 to 20 μg of genomic DNA were recovered in a final volume of 200 μl. The DNA was concentrated and purified using Microcon YM-30 microcentrifuge filters (Merck Millipore, Darmstadt, Germany) to a final volume of approximately 30 μl. Fragmentation of 10 μg of the purified DNA was performed by incubating at 37°C for 1 min in a 20-μl reaction mixture containing 1× One-Phor-All Plus buffer (GE Healthcare, Piscataway, NJ) and 0.01 U of DNase I (Amersham Biosciences, NJ) and was terminated by heating at 99°C for 15 min. The fragmented DNA was labeled by adding 10 μl of 5× terminal transferase buffer (Promega, Madison, WI), 2 μl of 1 mM biotin-11-ddATP (PerkinElmer, Waltham, MA), and 2 μl (60 U) of terminal transferase enzyme (Promega), incubating at 37°C for 3 h, and inactivating at 95°C for 10 min.

Hybridization was done following the Affymetrix GeneChip expression analysis technical manual for the 49-format array (27). Briefly, 200 μl of hybridization solution containing 10 μg of labeled DNA, 100 mM morpholineethanesulfonic acid (MES), 1 M Na$^+$, 20 mM EDTA, 0.01% Tween 20, 50 pM control oligonucleotide B2, 0.1 mg/ml salmon sperm DNA (Sigma, St. Louis, MO), 7.8% dimethyl sulfoxide (DMSO) (Sigma), and 0.5 mg/ml bovine serum albumin (BSA) (Sigma) was added to the Affymetrix arrays and incubated at 45°C with rotation (60 rpm) for 16 h in a hybridization oven. After hybridization, the arrays were washed and stained on an Affymetrix FS-450 fluidics station using the mini_prok2v1_450 fluidics script (27). Reagents for washing and staining were prepared according to the GeneChip expression analysis technical manual (27). The following exception was made to the wash-and-stain procedure: streptavidin solution mix (vial 1) was replaced with SAPE solution mix (Life Technologies, Grand Island, NY).

Arrays were scanned using a GeneChip scanner 3000 (Affymetrix) running AGCC software. For each gene represented on the array, probe set intensities were summarized using the robust multiarray averaging (RMA) function in the Affymetrix package of R-Bioconductor. Hierarchical clustering was performed with the RMA-summarized probe set intensities using the MADE4 package of R-Bioconductor. Scatter plots were generated with the RMA-summarized probe set intensities using Spotfire.

**Microarray data analysis.** All microarray data analysis was carried out using R-Bioconductor; summarized probe set intensities were calculated for each strain by using the RMA method (28) as implemented in the Affymetrix package of R-Bioconductor or Affymetrix Power Tools (APT). Briefly, RMA summarization of probe level data was done by performing three individual treatments on all of the experimental CEL data in succession. First, probe-specific correction of the PM probes was done using a model based on the observed intensities being the sum of signal and noise. Second, quantile normalization was performed on the corrected PM probe intensities. Finally, a median polishing algorithm was used to summarize the background-corrected and the normalized probe intensities to generate a final probe set value.

**Calculating gene differences and generating dendrograms.** RMA-summarized probe set intensities were compared across all strains assayed for each gene. If the same gene in two (different) strains had an RMA intensity difference of greater than 8-fold ($\log_2 = 3$), then this gene was considered to be "different" between those two isolates. Using this criterion, a strain-versus-strain gene difference matrix was generated, where the difference matrix represents the number of genes/alleles that differs between any two isolates. Gene difference matrices were converted to dendrograms using the hclust function in the base package as well as the phylo function in the ape package of R-Bioconductor.

## RESULTS

**Sensitivity, specificity, and discriminatory power of microarray analysis to characterize the diversity of pond-associated *Salmonella* strains.** During the study, two presumptive *Salmonella* colonies (double picks) were picked from six randomly selected water samples for microarray analysis to determine whether multiple picks from a water sample would improve the recovery of *Salmo-*

*nella* from irrigation water by microarray analysis. As a result, on three occasions, the double-picks from water samples C143, C188, and C234 recovered two different serovars; i.e., water sample C143 yielded C143 (*S.* Thompson) and C143A (*S.* Montevideo), water sample C188 yielded C188 (*S.* Drypool) and C188C (*S.* Montevideo), and water sample C234 yielded C234 (*S.* Newport) and 234A (*S.* Saintpaul). On three occasions, double picks from water samples C110, C151, and C180 recovered the same serovar; i.e., water sample C110 yielded isolates C110 and C110A (*S.* Newport), water sample C151 yielded isolates C151 and C151A (*S.* Newport), and water sample C180 yielded isolates C180 and C180A (*S.* Enteritidis).

This double-pick strategy provided three pairs of strains possessing the same serovar (C110-C110A [*S.* Newport], C151-C151A [*S.* Newport], and C180-C180A [*S.* Enteritidis]). Thus, we tested whether these three pairs of isolates could be used as three sets of replicates of identical genotypes (reference strains for quality control) for the analysis. As a result, each isolate pair showed an indistinguishable microarray genotype (0 gene differences), suggesting that each isolate pair was a clone of the strain from the same source or origin (Fig. 2A, B, and C). Furthermore, the six *S.* Enteritidis isolates were recovered from different ponds and at different sampling time points; indistinguishable genomic profiles were demonstrated by microarray analysis (see Fig. S1 in the supplemental material).

To further test the microarray's discriminatory power and specificity, we used the microarray to understand the genomic diversity associated with other closely related strains. For example, the *S.* Newport isolate from water sample C75 was compared with numerous *S.* Newport isolates from other water samples, including C177, strain SL1511 (an *S.* Newport isolate from a cilantro sample from 2013) and strain SARB36 (*S.* Newport). In contrast, this *S.* Newport isolate (C75) showed various degrees of genomic divergence from these strains, showing 0 to 240 gene differences (Fig. 2D, E, and F; see Fig. S1 in the supplemental material). Moreover, all the *Salmonella* isolates from the irrigation ponds, strains from recent outbreaks and strains from the Division of Molecular Biology (DMB), FDA, strain collection were correctly identified, with clear results without any ambiguity in determining a strain identity encountered during the analysis. Thus, the sensitivity, specificity, and discriminatory power of the microarray analysis were demonstrated to be satisfactory for foodborne pathogen identification and source attribution.

**Subtyping *Salmonella enterica* by microarray analysis.** The microarray data on the examined isolates were compared with our in-house FDA DMB microarray database. Serovars of the 56 *Salmonella* isolates recovered from the irrigation ponds were determined by microarray analysis, and all the isolates were successfully subtyped, resulting in 53 strains (three pairs of isolates were identified as the same strain) (Table 1); a cluster dendrogram among the 56 isolates analyzed was then generated by microarray analysis (Fig. 3). The majority of the isolates were *S.* Newport, accounting for 54.7% (29/53), and the second most common serovar was *S.* Enteritidis, accounting for 11.3% (6/53); these were followed by *S.* Muenchen and *S.* Javiana, each accounting for 7.5% (4/53). There were two isolates identified as *S.* Thompson, two isolates identified as *S. enterica* subsp. *diarizonae*, and two isolates identified as *S.* Saintpaul, each accounting for 3.8% (2/53). Isolates of other serovars included *S.* Bareilly, *S.* Montevideo, *S.* Inverness, and *S.* Drypool (Table 1).
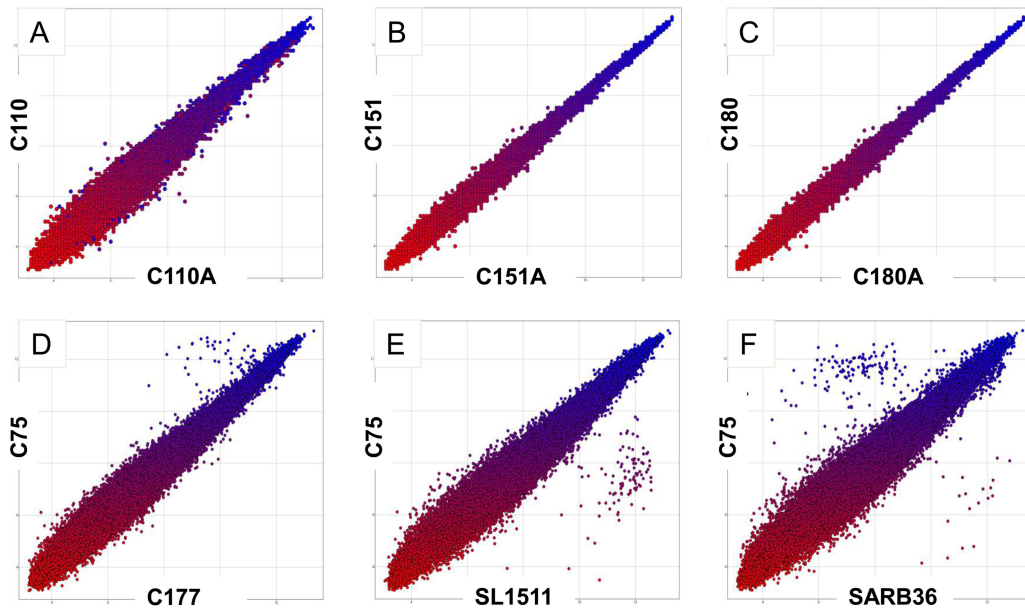
**FIG 2** Microarray scatter plots of pairwise comparisons demonstrating gene-level differences between strains analyzed in this study. (A to C) Three pairs of isolates C110 (*S.* Newport) (A), C151 (*S.* Newport) (B), and C180 (*S.* Enteritidis) (C), each of which was recovered from the same irrigation pond, were compared. (D to F) *S.* Newport isolate C75 was compared with *S.* Newport isolate C177 (recovered from a different pond) (D), *S.* Newport isolate SL1511 (recovered from recent outbreaks) (E), and strain SARB36 (F).

Microarray analysis further differentiated the 29 *S.* Newport isolates into two lineages, *S.* Newport II and *S.* Newport III. *S.* Newport II included 16 isolates, and *S.* Newport III covered 13 isolates (Fig. 3). It is noteworthy that these microarray results not only well complemented the PFGE analysis described in our previous report (1) but also identified two isolates which were not typeable by PFGE, i.e., C188 and CC227 (Table 2). Additionally, microarray analysis confirmed the strain-level synteny among *S.* Newport clusters and revealed subtle genomic differences within the clusters, showing that

there were 0 to 30 gene differences among the *S.* Newport II isolates and 0 to 16 gene differences among the *S.* Newport III isolates (see Fig. S2 in the supplemental material).

**Tracking the subtle genomic differences among isolates of *Salmonella* Newport by microarray analysis.** With the high discriminatory power of microarray analysis, not only could we subtype *S.* Newport to two different lineages, but we also were able to detect subtle gene differences within each lineage and further differentiate the two lineages into four subgroups, A, B, C, and D. For

**TABLE 1** Distribution of *Salmonella* serovars of the isolates from irrigation ponds as determined by microarray analysis

| Pond code | 2011 | | | | | | 2012 | | | | | | | | | | | | 2013 | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Jul. | Aug. | Sep. | Oct. | Nov. | Dec. | Jan. | Feb. | Mar. | Apr. | May | Jun. | Jul. | Aug. | Sep. | Oct. | Nov. | Dec. | Jan. | Feb. | Mar. | Apr. | May | Jun. | Jul. | Aug. | Sep. |
| MD1 | | | | C109[a] | | | | C149 | C152 | | C162 | | | C177 | C182 | | | C197 | | | | | | C227 | | | C242 |
| CC2 | | C90 | | C110 | | | | C150 | C153 | | C163 | | | | C183 | C188[b] | C198 | | | | | | | | | | C243 |
| LV | | C82 | | | | C122[f] | | C142 | C154 | | | | C174 | C179 | | | | | | | | | | | C234[e] | | C244 |
| NP | | C83 | | | | C123 | C143[c] | | | | | | | C180 | | | | | | | | | | | | C235 | C240 | C245 |
| SC | | C84 | | | | C124 | | | | | | C166 | | C181[d] | C186 | | | | | | | | | | | | C241 | C246 |
| BB | C75 | | | | | C125 | | | | | | | | | | | | | | | | | | | | | |
| VH | | | | | | C126 | No sampling | | | | | | | | | | | | | | | | | | | | |
| RT1 | | | | C107 | | | | C147 | | | | | | | | | | | | | | | | | | | |
| RT2 | | C88 | | C108 | | | | C148 | | | | | | | | | | | | | | | | | | | |
| CC1 | | | | C111 | | | | C151 | | | | | | | | | | | | | | | | | | | |

[a] *Salmonella* isolates from the samples are highlighted by color as follows: yellow, *S.* Newport II; red, *S.* Newport III; green, *S.* Enteritidis; light blue, *S.* Muenchen; gray, *S.* Javiana; dark blue, *S.* Thompson; orange, *S. enterica* subsp. *diarizonae*.

[b] Samples that are not highlighted were as follows: C186, *S.* Inverness; C188, *S.* Drypool; C198, *S.* Saintpaul; and C245, *S.* Bareilly.

[c] Two different strains were isolated from the sample, i.e., C143 (*S.* Thompson) and C143A (*S.* Montevideo).

[d] Two different strains were isolated from the sample, i.e., C181 (*S.* Enteritidis) and C181C (*S.* Javiana).

[e] Two different strains were isolated from the sample, i.e., C234 (*S.* Newport) and C234A (*S.* Saintpaul).

[f] Samples in brackets were collected on 5 December 2011, and samples in boxes and brackets were collected on 21 February 2012.
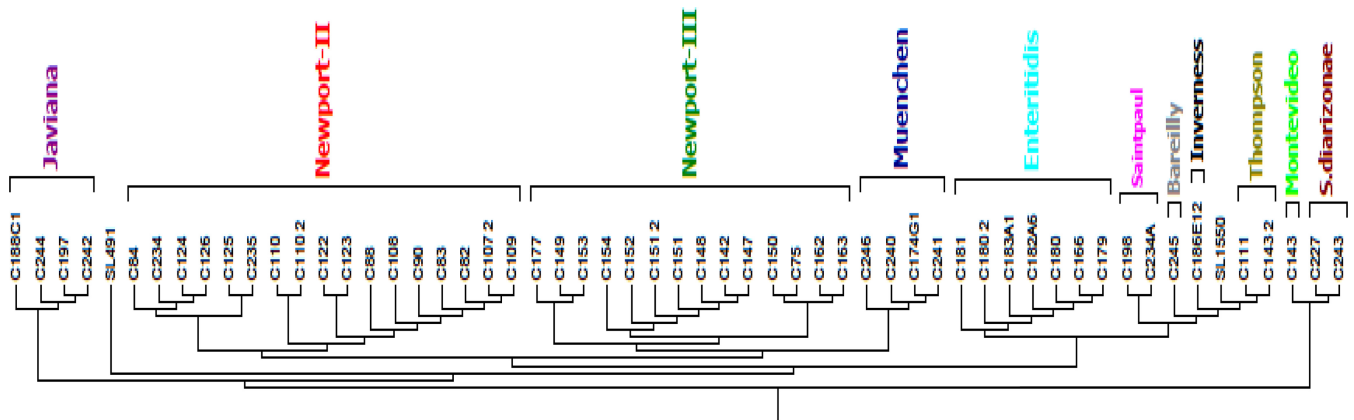
**FIG 3** Hierarchical clustering of RMA-summarized microarray data employing a database of over 760 *Salmonella* reference strains and outbreak strains from DMB, FDA, collections integrated with 56 isolates analyzed in this study. The resulting dendrogram represents a large comprehensive phylogenic tree which can be viewed in Fig. S2 in the supplemental material. Using the same strategy, a streamlined cluster dendrogram was generated among the 56 isolates analyzed by microarray analysis.

instance, in lineage II, strains in subgroups A (C84, C124, C126, and C234) and B (C122, C123, C124, C125, and C126) showed an identical genotype within their subgroup as measured by gene differences ($n = 0$), suggesting that these strains in this subgroup may share a host; in lineage III, strains in subgroups C (C142, C147, C148, C150, C151, C152, C153, and C154) and D (C75, C162, and C163) had an identical genotype within the subgroup, suggesting that these strains may also share a host (Table 3; Fig. 4). Although no gene differences were found within these subgroups, a small number of gene differences were detected outside these subgroups (Table 3; Fig. 4). These microarray data can be used to understand the relationship among these closely related *S*. Newport isolates from each irrigation pond.

**Antimicrobial susceptibility profiles determined by microarray analysis.** Previously, 16 *S*. Newport isolates were found to be resistant to amoxicillin-clavulanic acid, ampicillin-sulbactam, ceftazidime, ceftriaxone, cephalothin, chloramphenicol, streptomycin, and tetracycline by the Kirby-Bauer method (1). In this study, the microarray data confirmed the antimicrobial resistance phenotypes of the isolates by the detection of the presence of the five genes that were responsible for the antimicrobial resistance genotypes among the 16 *S*. Newport isolates, as shown in Table 4. Hence, confirmation of antimicrobial resistance genotypes of *S*. Newport isolates by microarray analysis demonstrated the microarray's ability to rapidly discern the antimicrobial susceptibility profiles of foodborne pathogens such as *Salmonella* in outbreaks, demonstrating the versatility of this technology for use in microbiological and epidemiological studies.

## DISCUSSION

Molecular typing methods, such as PFGE, MLST, and Rep-PCR, have several advantages over traditional methods, such as increased discriminatory power, standardization, and reproducibility (29). The most common molecular subtyping method is PFGE. In the present study, we used a custom-designed microarray to subtype the *Salmonella enterica* isolates and to assess the microarray's capacity as an alternative method to PFGE analysis in subtyping of *Salmonella enterica*.

Microarray identified all of the 53 *Salmonella* isolates, compared with 96.1% (49/51) identification of the isolates by PFGE

(Table 2). Overall, these two subtyping technologies complemented each other. Similar to our previously reported PFGE analysis for these strains (1), microarray analysis was able to further differentiate the 29 *S*. Newport isolates into two lineages (Fig. 3). PFGE analysis was able to provide specific PFGE (XbaI) patterns (PulseNet patterns) with most of the assayed isolates. By querying PulseNet with these patterns, we were able to find the relationship between the isolates and recent *Salmonella* outbreaks or clinical isolates (1). On the other hand, two isolates (C188 and C227) were not identified (1). This was not surprising, because these isolates were recovered from environmental water samples, and PulseNet is more focused toward food-related and clinical isolates. However, here, isolates C188 and C227 were successfully identified as *S*. Drypool and *S*. *enterica* subsp. *diarizonae*, respectively, by microarray analysis (Table 2). Additionally, isolate C243 was identified as *S*. Virchow by PFGE but was identified as *S*. *enterica diarizonae* by microarray analysis (Table 2). Subsequently, these identification results of the three strains (C188, C227, and C234) by microarray analysis were confirmed by whole-genome sequencing (WGS) analysis (data not shown).

The predominance (29/53) of *S*. Newport among *Salmonella* isolates in surface water was intriguing and puzzling. Many of these isolates were recovered from different water samples and from different sampling events (pond locations and sampling dates). However, these peculiar serovar distribution patterns blended with their spatial and temporal factors (ten selected irrigation ponds surveyed for 2 years) and provided a set of valuable metrics to study the underaddressed issue of survival of *Salmonella* in environmental water, such as irrigation pond water.

Previous studies have demonstrated that *S*. Newport is polyphylogenetic and is further differentiated into three different genomic clusters (23, 30). Based on the epidemiological study of *S*. Newport, isolates in the *S*. Newport cluster with multidrug resistance (MDR) genes are more likely to be in the *S*. Newport II lineage, whereas the *S*. Newport pan-susceptible isolates predominate in the *S*. Newport III lineage (23, 30). In our previous study, despite the highly polyphyletic serovar nature of *S*. Newport, all strains within each lineage were indistinguishable by PFGE (1). However, owing to the high resolution of microarray analysis, we

TABLE 2 Comparison of subtyping of *Salmonella* isolates from the irrigation ponds and sampling day observations at the ponds

| Isolate | Pond code | Sample date (mo/day/yr) | Subtype determined by: | | Observations for signs of possible hosts/reservoirs of *Salmonella* and other environmental factors |
|---|---|---|---|---|---|
| | | | PFGE | Microarray analysis | |
| C75 | BB | 07/05/11 | *S.* Newport | *S.* Newport | Pond almost full |
| C82 | LV | 08/15/11 | *S.* Newport | *S.* Newport | Recent rain event |
| C83 | NP | 08/15/11 | *S.* Newport | *S.* Newport | Wading birds present, recent crop harvested |
| C84 | SC | 08/15/11 | *S.* Newport | *S.* Newport | Recent crop harvested |
| C88 | RT2 | 08/29/11 | *S.* Newport | *S.* Newport | Dry land with cotton growing, turtles |
| C90 | CC2 | 08/29/11 | *S.* Newport | *S.* Newport | Low water level and much algae in water |
| C107 | RT1 | 10/24/11 | *S.* Newport | *S.* Newport | Geese present, turtle and deer tracks |
| C108 | RT2 | 10/24/11 | *S.* Newport | *S.* Newport | Beaver, turtle, egret present |
| C109 | RD1 | 10/24/11 | *S.* Newport | *S.* Newport | Turtle present |
| C110 | CC2 | 10/24/11 | *S.* Newport | *S.* Newport | Deer tracks, beaver activity |
| C111 | CC1 | 10/24/11 | *S.* Thompson | *S.* Thompson | No information |
| C122 | LV | 12/05/11 | *S.* Newport | *S.* Newport | Bobcat and canine tracks |
| C123 | NP | 12/05/11 | *S.* Newport | *S.* Newport | Tracks of deer, canine, coyote, turkey, porcine, and armadillo |
| C124 | SC | 12/05/11 | *S.* Newport | *S.* Newport | Raccoon tracks |
| C125 | BB | 12/05/11 | *S.* Newport | *S.* Newport | No information |
| C126 | VH | 12/05/11 | *S.* Newport | *S.* Newport | Crop soil prepn |
| C142 | LV | 02/06/12 | *S.* Newport | *S.* Newport | Feline tracks |
| C143 | NP | 02/06/12 | *S.* Thompson | *S.* Thompson | Deer, skunk, and canine tracks |
| C143A | NP | 02/06/12 | Not done | *S.* Montevideo | Deer, skunk, and canine tracks |
| C147 | RT1 | 02/21/12 | *S.* Newport | *S.* Newport | Crop soil prepn |
| C148 | RT2 | 02/21/12 | *S.* Newport | *S.* Newport | Crop soil prepn |
| C149 | MD1 | 02/21/12 | *S.* Newport | *S.* Newport | Crop soil prepn |
| C150 | CC2 | 02/21/12 | *S.* Newport | *S.* Newport | Recent rain event |
| C151 | CC1 | 02/21/12 | *S.* Newport | *S.* Newport | Recent rain event, wading birds present |
| C152 | MD1 | 03/19/12 | *S.* Newport | *S.* Newport | Deer tracks, beaver activity |
| C153 | CC2 | 03/19/12 | *S.* Newport | *S.* Newport | Deer tracks, beaver activity |
| C154 | LV | 03/26/12 | *S.* Newport | *S.* Newport | No information |
| C162 | MD1 | 05/07/12 | *S.* Newport | *S.* Newport | No information |
| C163 | CC2 | 05/07/12 | *S.* Newport | *S.* Newport | Two armadillos, beaver activity, and deer tracks |
| C166 | SC | 05/21/12 | *S.* Enteritidis | *S.* Enteritidis | Pumping in and out of pond |
| C174 | LV | 07/23/12 | *S.* Muenchen | *S.* Muenchen | Crop spoil prepn |
| C177 | MD1 | 08/13/12 | *S.* Newport | *S.* Newport | No information |
| C179 | LV | 08/27/12 | *S.* Enteritidis | *S.* Enteritidis | No information |
| C180 | NP | 08/27/12 | *S.* Enteritidis | *S.* Enteritidis | Deer tracks |
| C181 | SC | 08/27/12 | *S.* Enteritidis | *S.* Enteritidis | Coyote tracks |
| C182 | MD1 | 09/17/12 | *S.* Enteritidis | *S.* Enteritidis | Crop soil prepn |
| C183 | CC2 | 09/17/12 | *S.* Enteritidis | *S.* Enteritidis | Multiple egret and wood storks present, beaver tracks |
| C186 | SC | 09/17/12 | *S.* Inverness | *S.* Inverness | Pumping in and out of pond |
| C188 | CC2 | 10/01/12 | Not typeable | *S.* Drypool | Beaver tracks |
| C188C | CC2 | 10/01/12 | Not done | *S.* Montevideo | Beaver tracks |
| C197 | MD1 | 12/03/12 | *S.* Javiana | *S.* Javiana | Beaver activity, wading birds |
| C198 | CC2 | 12/03/12 | *S.* Saintpaul | *S.* Saintpaul | Beaver activity, wading birds, raccoon tracks |
| C227 | MD1 | 06/13/13 | Not typeable | *S. enterica* subsp. *diarizonae* | Turtle present |
| C234 | LV | 07/22/13 | *S.* Newport | *S.* Newport | Wading birds and crows present |
| C234A | LV | 07/22/13 | *S.* Saintpaul | *S.* Saintpaul | Wading birds and crows present |
| C235 | NP | 07/22/13 | *S.* Newport | *S.* Newport | Wading birds present |
| C240 | NP | 08/19/13 | *S.* Muenchen | *S.* Muenchen | No information |
| C241 | SC | 08/19/13 | *S.* Muenchen | *S.* Muenchen | House flies, armadillo tracks |
| C242 | MD1 | 09/16/13 | *S.* Javiana | *S.* Javiana | No information |
| C243 | CC2 | 09/16/13 | *S.* Virchow | *S. enterica* subsp. *diarizonae* | No information |
| C244 | LV | 09/23/13 | *S.* Javiana | *S.* Javiana | Crows and geese present |
| C245 | SC | 09/23/13 | *S.* Bareilly | *S.* Bareilly | No information |
| C246 | SC | 09/23/13 | *S.* Muenchen | *S.* Muenchen | Area up slope cleared of trees and debris |
| Total | | | 51 | 53 | |

**TABLE 3** Gene differences among isolates of two lineages of *Salmonella* Newport as detected by microarray analysis[a]

| Lineage | Subgroup | Isolate name | Pond code | Sampling date | C125 | C235 | C84 | C124 | C126 | C234 | C122 | C123 | C110 | C88 | C108 | C82 | C83 | C90 | C107 | C109 | C149 | C177 | C142 | C151 | C153 | C147 | C148 | C150 | C152 | C154 | C75 | C162 | C163 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| II | | C125 | BB | 12/05/11 | 0 | 2 | 3 | 3 | 3 | 3 | 4 | 4 | 29 | 13 | 15 | 16 | 8 | 13 | 8 | 17 | | | | | | | | | | | | | |
| | | C235 | NP | 07/22/13 | 2 | 0 | 2 | 1 | 1 | 0 | 4 | 6 | 20 | 10 | 12 | 12 | 7 | 10 | 7 | 12 | | | | | | | | | | | | | |
| | A | **C84** | **SC** | **08/15/11** | 3 | 2 | 0 | 0 | 0 | 0 | 1 | 3 | 24 | 10 | 13 | 12 | 6 | 10 | 6 | 14 | | | | | | | | | | | | | |
| | | C124 | SC | 12/05/11 | 3 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 19 | 3 | 4 | 6 | 1 | 3 | 1 | 10 | | | | | | | | | | | | | |
| | | C126 | VH | 12/05/11 | 3 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 18 | 8 | 9 | 9 | 1 | 7 | 1 | 10 | | | | | | | | | | | | | |
| | | **C234** | **LV** | **07/22/13** | 3 | 0 | 0 | 0 | 0 | 0 | 3 | 0 | 12 | 3 | 4 | 5 | 0 | 0 | 0 | 4 | | | | | | | | | | | | | |
| | | C122 | LV | 12/05/11 | 4 | 4 | 1 | 1 | 1 | 3 | 0 | 0 | 1 | 1 | 2 | 0 | 0 | 1 | 2 | 1 | | | | | | | | | | | | | |
| | | C123 | NP | 12/05/11 | 4 | 6 | 3 | 1 | 0 | 0 | 0 | 0 | 6 | 0 | 1 | 1 | 0 | 0 | 0 | 1 | | | | | | | | | | | | | |
| | | C110 | CC2 | 10/24/11 | 29 | 20 | 24 | 21 | 18 | 14 | 1 | 6 | 0 | 7 | 7 | 8 | 6 | 7 | 7 | 7 | | | | | | | | | | | | | |
| | | C88 | RT2 | 08/29/11 | 13 | 10 | 10 | 3 | 8 | 3 | 1 | 0 | 7 | 0 | 1 | 1 | 0 | 0 | 0 | 1 | | | | | | | | | | | | | |
| | | C108 | RT2 | 10/24/11 | 15 | 12 | 13 | 4 | 9 | 4 | 2 | 1 | 7 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | | | | | | | | | | | | | |
| | B | C82 | LV | 08/15/11 | 16 | 12 | 12 | 6 | 9 | 5 | 0 | 1 | 7 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | | | | | | | | | | | | | |
| | | C83 | NP | 08/15/11 | 8 | 7 | 6 | 1 | 1 | 0 | 0 | 0 | 6 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | | | | | | | | | | | | | |
| | | C90 | CC2 | 08/29/11 | 13 | 10 | 10 | 3 | 7 | 0 | 1 | 0 | 7 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | | | | | | | | | | | | | |
| | | C107 | RT1 | 10/24/11 | 8 | 7 | 6 | 1 | 1 | 0 | 2 | 0 | 7 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | | | | | | | | | | | | | |
| | | C109 | MD1 | 10/24/11 | 17 | 12 | 14 | 10 | 10 | 4 | 1 | 1 | 7 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | | | | | | | | | | | | | |
| III | | C149 | MD1 | 02/21/12 | | | | | | | | | | | | | | | | | 0 | 0 | 0 | 0 | 0 | 0 | 4 | 7 | 7 | 5 | 11 | 15 | 10 |
| | | C177 | MD1 | 08/13/12 | | | | | | | | | | | | | | | | | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 4 | 5 | 4 | 8 | 11 | 7 |
| | C | C142 | LV | 02/06/12 | | | | | | | | | | | | | | | | | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | | C151 | CC1 | 02/21/12 | | | | | | | | | | | | | | | | | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | | C153 | CC2 | 03/19/12 | | | | | | | | | | | | | | | | | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | | C147 | RT1 | 02/21/12 | | | | | | | | | | | | | | | | | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 4 | 1 |
| | | C148 | RT2 | 02/21/12 | | | | | | | | | | | | | | | | | 4 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | | C150 | CC2 | 02/21/12 | | | | | | | | | | | | | | | | | 7 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | | C152 | MD1 | 03/19/12 | | | | | | | | | | | | | | | | | 7 | 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | | C154 | LV | 03/26/12 | | | | | | | | | | | | | | | | | 5 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | D | C75 | BB | 07/05/11 | | | | | | | | | | | | | | | | | 11 | 8 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | | C162 | MD1 | 05/07/12 | | | | | | | | | | | | | | | | | 15 | 11 | 0 | 0 | 0 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | | C163 | CC2 | 05/07/12 | | | | | | | | | | | | | | | | | 10 | 7 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

[a] Strains that demonstrated no gene differences (*n* = 0) as determined by microarray analysis are highlighted with the same color (green, yellow, red, or blue).

were able to trace the minor changes in genomic content within these strains associated with each cluster (Table 4). For instance, microarray analysis clearly pointed out the subtle genomic differences within the two *S.* Newport lineages, showing 0 to 29 gene differences within the *S.* Newport II lineage and 0 to 16 gene differences within the *S.* Newport III lineage (Table 4). Therefore, the nearly identical genomic profiles of these *S.* Newport isolates manifested an important scenario: the two genotypes of *S.* Newport circulating in the irrigation ponds in the Suwannee River watershed were clonal and were likely from two independent hosts or sources.

There are over 2,600 known *Salmonella* serovars, and their ubiquity in the environment highlights the complex nature of *Salmonella* transmission. Numerous studies have reported that *Salmonella* can be transmitted through water (31, 32); however, few if any reports have addressed how *Salmonella* disseminates by water or have provided phylogenetic evidence of survival of bacteria in water. Interestingly, the two lineages of *S.* Newport (II and III) observed in this study have been circulating in the 10 irrigation ponds for 2 years (Table 1). Specifically, the *S.* Newport II genotype was found during the period between August 2011 and July, 2013, whereas the *S.* Newport III genotype was present during the period between July 2011 and August 2012. Serovar distribution analysis, however, showed there were no instances where the same serovar was recovered from a pond for more than a month, except on one occasion, where three *S.* Newport III isolates were recov-

ered from three ponds (MD1, CC2, and LV) during two consecutive months (February to March 2012). However, both microarray and whole-genome sequencing (WGS) (data not shown) indicated that more gene differences were detected vertically (same pond at different time points) than horizontally (different ponds at the same time point) among the six isolates (Table 3; see Fig. S2 in the supplemental material), suggesting that none of the three isolates recovered in March 2012 were clones of the same strain that was isolated in February 2012; *Salmonella* in the pond water may not last for more than a month. However, this punctuated but lingering presence of *S.* Newport in the irrigation ponds lasted for 2 years. For instance, *S.* Newport II (C83) was isolated from irrigation pond NP in August 2011, but no *Salmonella* was recovered from NP in the next 3 months; again, *S.* Newport II was isolated in December 2011(C123) and July 2013 (C234), respectively.

Additionally, during the 2-year survey, five more different serovars of *Salmonella* were isolated from irrigation pond NP, i.e., in February 2012 (C143 and C143A), August 2012 (C180), August 2013 (C240), and September 2013 (C245) (Tables 1 and 2). This particular irrigation pond circulation pattern of *S.* Newport and the microarray evidence of subtle genomic differences among the isolates within the two different lineages of *S.* Newport (II and III) indicated that some source(s) or host(s) that carried *S.* Newport continuously shed the pathogen into the environment, causing numerous reintroduction events in the ponds. These syntenic *S.*
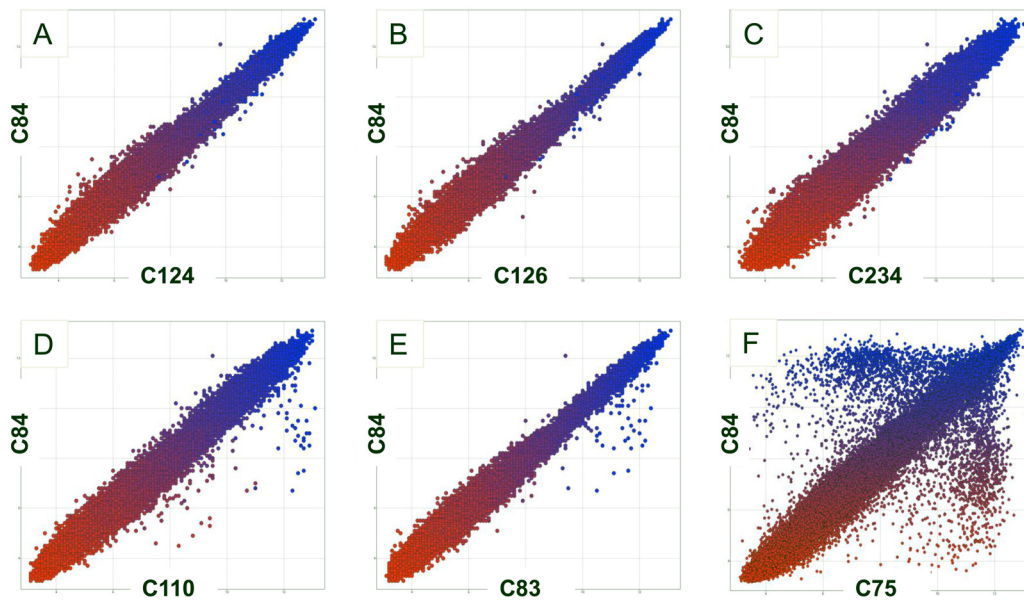
FIG 4 Microarray scatter plots of pairwise comparisons demonstrating gene-level differences between strains analyzed in this study. Four *S.* Newport II isolates C84, C124, C126, and C234, were recovered from different ponds at different sample dates. (A to C) The genotype of C84 was compared with those of C124 (A), C126 (B), and C234 (C). (D to F) The genotype of C84 was also compared with those of two *S.* Newport II isolates, C110 (D) and C83 (E), and *S.* Newport II strain C75 (F).

Newport strains were present in a relatively large surveyed region (over 180 km long), and some of these strains were isolated on the same sampling dates but at different locations. Furthermore, on two different sampling dates (samples taken on 5 December 2011 and 21 February 2012 and indicated by brackets in Table 1), an *S.* Newport isolate was recovered from each of the five samples on either sampling date, and the isolates on the same sampling dates shared a common S. Newport genotype lineage (II or III, respectively) (Table 1; Fig. 3). In light of the average recovery (30%) of *Salmonella* from the ponds as reported previously (1), the 100% recovery on two different sampling dates from the ponds with two lineages of *S.* Newport (II and III) clearly revealed a strong synchronous nature of the occurrence of *Salmonella* in the irrigation ponds despite the spatial factors associated with the ponds (as far as 180 km apart). Hence, one plausible scenario might be that multiple mobile hosts, such as wildlife, which might have been colonized with *Salmonella* either transiently or through an active zoonotic infection came in contact with the pond environment. These hosts not only inhabited these areas, but also continuously disseminated *Salmonella* into the ponds directly or indirectly. This notion gained some credence from the observation records at the ponds on the sampling day, which showed signs of a variety of

wildlife, such as tracks of deer, coyote, etc., in the proximity of the ponds (Table 2). Previously, Gorski et al. (14) reported a similar circulating pattern and suggested that one or more sources maintained the pathogen population (e.g., by colonization) and continuously disseminated the bacteria into the environment. *Salmonella* has been reported to persist in the environment for several months to more than a year (14, 33–35), and *Salmonella* in irrigation ponds could come from residential or industrial water, runoff water, rainfall, soil, crops, livestock, wildlife, or husbandry activities (35). In 1996, Beuchat proposed a broad and comprehensive model to encompass all foodborne pathogens, including viruses, and the likely transmission routes, such as feces, soil, irrigation water, manure, air, etc., which are involved in contamination of produce (36). However, how the bacteria from these routes or sources were introduced into irrigation water to result in this particular distribution pattern remains unclear. In this study, we attempted to summarize the possible factors that may contribute to the survival of *Salmonella* in the surface water transmission routes through the produce production and consumption to establish a conceptual model to visualize how *Salmonella* is transmitted from the environment to irrigation ponds and vice versa, as shown in

TABLE 4 Genotype for antimicrobial susceptibility of *Salmonella* isolates detected by microarray analysis

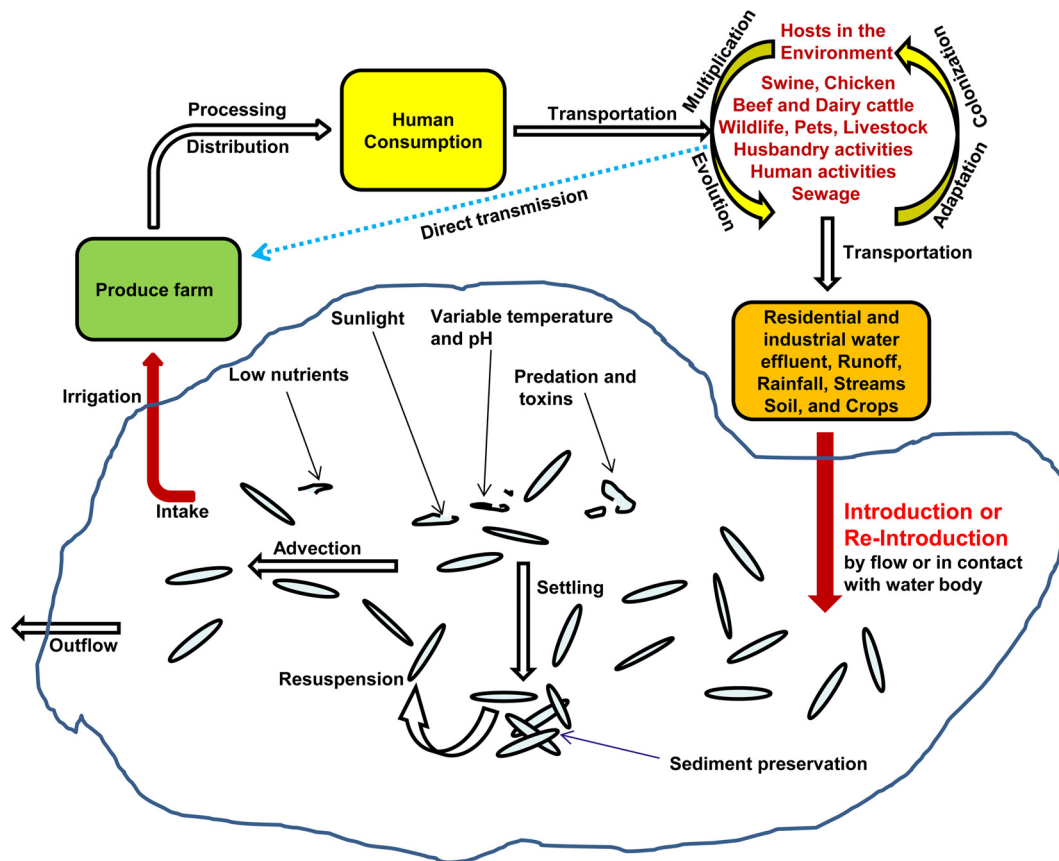| Resistance genotype | Predicted phenotype | Presence in isolate: | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | C75 | C82 | C83 | C84 | C88 | C90 | C107 | C108 | C109 | C110 | C122 | C123 | C124 | C125 | C126 | C240 | C234 | C235 |
| *strB* | Streptomycin | − | + | + | + | + | + | + | + | + | + | + | + | + | + | + | − | + | + |
| *strA* | Streptomycin | − | + | + | + | + | + | + | + | + | + | + | + | + | + | + | − | + | + |
| *bla*$_{CMY-59}$ | Beta-lactam | − | + | + | + | + | + | + | + | + | + | + | + | + | + | + | − | + | + |
| *floR* | Phenicol | − | + | + | + | + | + | + | + | + | + | + | + | + | + | + | − | + | + |
| *sul2* | Sulfonamide | − | + | + | + | + | + | + | + | + | + | + | + | + | + | + | − | + | + |
| *tet*(A) | Tetracycline | − | + | + | + | + | + | + | + | + | + | + | + | + | + | + | − | + | + |

**FIG 5** Conceptual model for *Salmonella* transmission through surface water, showing microbial interactions and intersections between the environment, animals, and humans that influence the dynamic nature of the *Salmonella* population within an irrigation pond ecology through which the consumption of produce contaminated with irrigation water leads to human illness.

Fig. 5. It may be helpful in the development of efficient and safe irrigation strategies for produce farmers.

This hypothesis of reintroduction still needs to be further tested and validated by independent studies in various geographic and ecological regions. Survival of *Salmonella* in the environment has been a longstanding issue (37, 38), but survival of *Salmonella* in surface water has been underaddressed (25, 26). In fact, most of the reports describing the persistence of *Salmonella* were related to poultry houses (39, 40), swine and dairy farms (41–43), produce and fruit farms (44–46), and the process of adaptation and/or colonization of various hosts by *Salmonella* that might be involved (35, 45), where numerous ecological factors affect the *Salmonella* survival niches in the environment. Determining the fate of *Salmonella* in such complex ecosystems was a daunting task because it is almost impossible to differentiate reintroduction or regrowth of *Salmonella* from the environment with *Salmonella* persisting in the environment without having access to high-resolution pathogen or source attribution technologies and constant monitoring/real-time monitoring/24-h surveillance of the ecosystems.

To address these challenges, other studies have used inoculation-based approaches, i.e., spiking water or dairy lagoons, which were contained in microcosms with high concentrations (up to $10^{10}$ CFU/g) of *Salmonella*. The microcosms were subjected to changes in various parameters such as temperature, pH, nutrients, time, etc., and survival of *Salmonella* changed drastically under

these various *in vitro* conditions (25, 26, 47). Ravva and Sarreal reported that *Salmonella* populations decreased significantly in aerated microcosms held under summer conditions, with a decimal reduction time of 3.4 days (47), while Cevallos-Cevallos et al. found that *Salmonella* survived for at least 63 days, depending on the microcosm eutrophication level and the temperature (26).

However, these data were derived from samples under certain experimental circumstances (microcosms) rather than directly from natural environmental samples, and the microcosms used were isolated containers that were set at specific constant parameters or conditions, mimicking natural conditions except that they lacked the dynamic influences from runoff, rainfall, sunlight, wildlife, human activities, predation, toxins, etc. (35). On the other hand, these data seemed to more accurately reflect the *Salmonella* population dynamics under such specific conditions by excluding the factors of reintroduction and/or regrowth of *Salmonella* from the environment. Such findings may also help explain why, in the present study, strains of the same serovar isolated from a pond were not recovered in the following month. As an example, first, in the summer, the *Salmonella* decimal reduction time could be as short as 3.4 days, and second (and more importantly), these studies reminded us that the *Salmonella* populations in the ponds were much lower than the concentrations (up to $10^{10}$ CFU/g) used in the microcosm experiments by these groups (25, 26, 47), which were deemed to be much higher than the possible concentrations

in the real-world scenarios in the environment, as noted by Moore et al. (25). Therefore, one plausible explanation for the lack of *Salmonella* persistence in the ponds was low initial *Salmonella* populations in the ponds and the faster-than-previously-assumed *Salmonella* decimal reduction time. In fact, the low levels of *Salmonella* in the ponds were reflected by the high threshold cycle ($C_T$) values of most of the samples in the quantitative PCR (qPCR) screening step even after the samples were concentrated and enriched in our previous work (1), and this was also corroborated by the assessment of the same sets of ponds for a similar period of time by a different group (48), who reported that the *Salmonella* levels remained consistently low (mean of <1 most probable number [MPN] per liter) during the study period (48). In retrospect, the predominance of *S.* Newport among isolates from the irrigation ponds on produce farms in our previous work (1) provided us an ideal ecosystem that reflected the spatial and temporal factors needed to address the issue of survival of *Salmonella* in surface water, while the DNA microarray analysis was a robust and efficient tool that allowed us to obtain adequate phylogenic evidence to attribute the dynamic distribution pattern of *Salmonella* in the irrigation ponds; i.e., this peculiar distribution pattern of *Salmonella* in the irrigation ponds might result from numerous reintroduction events from different hosts in the environment.

The present study is not without limits; for instance, even if this hypothesis holds, it would be more meaningful to specifically identify the hosts or sources that contaminated the irrigation ponds. This might be residential water, industrial water, runoff water, livestock, wildlife, husbandry activities, etc., and fast and accurate identification on the hosts or sources of the contamination of the ponds would be more relevant and helpful for farmers to take the appropriate measurements to minimize the risk of *Salmonella* contamination associated with irrigation water. Efforts regarding such aspects may be worth making for food safety and public health professionals.

**Conclusions.** Microarray analysis offers high discriminatory power to identify genomic diversity of the *Salmonella* serovars isolated from irrigation water. The application of genome-wide comparative analysis indicated that the two lineages of *S.* Newport were clonal; however, subtle genotypic differences were found within the lineages. Moreover, these findings suggested that the isolates in either lineage could have come from several unique hosts, and *Salmonella* in the ponds may not last for more than a month. Based on this phylogenic evidence, we established a conceptual model for *Salmonella* transmission in surface water; i.e., the presence of *Salmonella* in the ponds on the produce farms likely resulted from numerous reintroduction events from some hosts or reservoirs, such as cattle, horses, swine, husbandry activities, wildlife, etc., associated with the pond environment. These findings not only have theoretical implications in epidemiology but also may be useful for the development of efficient and safe irrigation practices for produce farmers.

## ACKNOWLEDGMENTS

## REFERENCES

1. **Li B, Vellidis G, Liu H, Jay-Russell M, Zhao S, Hu Z, Wright A, Elkins CA.** 2014. Diversity and antimicrobial resistance of Salmonella enterica isolates from surface water in southeastern United States. Appl Environ Microbiol **80:**6355–6365. http://dx.doi.org/10.1128/AEM.02063-14.
2. **Chimalizeni Y, Kawaza K, Molyneux E.** 2010. The epidemiology and management of non typhoidal salmonella infections. Adv Exp Med Biol **659:**33–46. http://dx.doi.org/10.1007/978-1-4419-0981-7_3.
3. **Scallan E, Griffin PM, Angulo FJ, Tauxe RV, Hoekstra RM.** 2011. Foodborne illness acquired in the United States—unspecified agents. Emerg Infect Dis **17:**16–22. http://dx.doi.org/10.3201/eid1701.P21101.
4. **Todd EC.** 1989. Costs of acute bacterial foodborne disease in Canada and the United States. Int J Food Microbiol **9:**313–326. http://dx.doi.org/10.1016/0168-1605(89)90099-8.
5. **CDC.** 2010. Preliminary FoodNet data on the incidence of infection with pathogens transmitted commonly through food—10 states, 2009. MMWR Morb Mortal Wkly Rep **59:**418–422.
6. **Foley SL, Johnson TJ, Ricke SC, Nayak R, Danzeisen J.** 2013. Salmonella pathogenicity and host adaptation in chicken-associated serovars. Microbiol Mol Biol Rev **77:**582–607. http://dx.doi.org/10.1128/MMBR.00015-13.
7. **Guibourdenche M, Roggentin P, Mikoleit M, Fields PI, Bockemuhl J, Grimont PA, Weill FX.** 2010. Supplement 2003-2007 (no. 47) to the White-Kauffmann-Le Minor scheme. Res Microbiol **161:**26–29. http://dx.doi.org/10.1016/j.resmic.2009.10.002.
8. **CDC.** 2014. *Salmonella* surveillance: annual summary, US. Centers for Disease Control and Prevention, Atlanta, GA.
9. **Jacobsen CS, Bech TB.** 2012. Soil survival of *Salmonella* and transfer to freshwater and fresh produce. Food Res Int **45:**557–566. http://dx.doi.org/10.1016/j.foodres.2011.07.026.
10. **Jones LA, Worobo RW, Smart CD.** 2014. UV light inactivation of human and plant pathogens in unfiltered surface irrigation water. Appl Environ Microbiol **80:**849–854. http://dx.doi.org/10.1128/AEM.02964-13.
11. **US Food and Drug Administration.** 2013. Food Safety Modernization Act: standards for the growing, harvesting, packing, and holding of produce for human consumption; proposed rule, 2013 (docket no. FDA-2011-N-0921). US Food and Drug Administration, Washington, DC.
12. **FDA.** 19 September 2014, posting date. Food Safety Modernization Act (FSMA): produce safety standards. US Food and Drug Administration, Silver Spring, MD. http://www.fda.gov/Food/GuidanceRegulation/FSMA/ucm304045.htm.
13. **Fedorka-Cray PJ, Dargatz DA, Thomas LA, Gray JT.** 1998. Survey of Salmonella serotypes in feedlot cattle. J Food Prot **61:**525–530.
14. **Gorski L, Parker CT, Liang A, Cooley MB, Jay-Russell MT, Gordus AG, Atwill ER, Mandrell RE.** 2011. Prevalence, distribution, and diversity of Salmonella enterica in a major produce region of California. Appl Environ Microbiol **77:**2734–2748. http://dx.doi.org/10.1128/AEM.02321-10.
15. **Kozak GK, MacDonald D, Landry L, Farber JM.** 2013. Foodborne outbreaks in Canada linked to produce: 2001 through 2009. J Food Prot **76:**173–183. http://dx.doi.org/10.4315/0362-028X.JFP-12-126.
16. **Schwartz DC, Cantor CR.** 1984. Separation of yeast chromosome-sized DNAs by pulsed field gradient gel electrophoresis. Cell **37:**67–75. http://dx.doi.org/10.1016/0092-8674(84)90301-5.
17. **Enright MC, Spratt BG.** 1999. Multilocus sequence typing. Trends Microbiol **7:**482–487. http://dx.doi.org/10.1016/S0966-842X(99)01609-1.
18. **Versalovic J, Koeuth T, Lupski JR.** 1991. Distribution of repetitive DNA sequences in eubacteria and application to fingerprinting of bacterial genomes. Nucleic Acids Res **19:**6823–6831. http://dx.doi.org/10.1093/nar/19.24.6823.
19. **van Oppen MJ, Rico C, Turner GF, Hewitt GM.** 2000. Extensive homoplasy, nonstepwise mutations, and shared ancestral polymorphism at a complex microsatellite locus in Lake Malawi cichlids. Mol Biol Evol **17:**489–498. http://dx.doi.org/10.1093/oxfordjournals.molbev.a026329.
20. **Shariat N, Dudley EG.** 2013. Where are we heading with Salmonella molecular subtyping? Future Microbiol **8:**1231–1233. http://dx.doi.org/10.2217/fmb.13.107.
21. **Shi C, Singh P, Ranieri ML, Wiedmann M, Moreno Switt AI.** 2013. Molecular methods for serovar determination of Salmonella. Crit Rev Microbiol http://dx.doi.org/10.3109/1040841X.2013.837862.
22. **Zheng J, Keys CE, Zhao S, Meng J, Brown EW.** 2007. Enhanced subtyping scheme for Salmonella enteritidis. Emerg Infect Dis **13:**1932–1935. http://dx.doi.org/10.3201/eid1312.070185.
23. **Sangal V, Harbottle H, Mazzoni CJ, Helmuth R, Guerra B, Didelot X,**

Paglietti B, Rabsch W, Brisse S, Weill FX, Roumagnac P, Achtman M. 2010. Evolution and population structure of Salmonella enterica serovar Newport. J Bacteriol 192:6465–6476. http://dx.doi.org/10.1128/JB.00969-10.

24. Galanis E, Lo Fo Wong DM, Patrick ME, Binsztein N, Cieslik A, Chalermchikit T, Aidara-Kane A, Ellis A, Angulo FJ, Wegener HC. 2006. Web-based surveillance and global Salmonella distribution, 2000–2002. Emerg Infect Dis 12:381–388. http://dx.doi.org/10.3201/eid1203.050854.

25. Moore BC, Martinez E, Gay JM, Rice DH. 2003. Survival of Salmonella enterica in freshwater and sediments and transmission by the aquatic midge Chironomus tentans (Chironomidae: Diptera). Appl Environ Microbiol 69:4556–4560. http://dx.doi.org/10.1128/AEM.69.8.4556-4560.2003.

26. Cevallos-Cevallos JM, Gu G, Richardson SM, Hu J, van Bruggen AH. 2014. Survival of Salmonella enterica Typhimurium in water amended with manure. J Food Prot 77:2035–2042. http://dx.doi.org/10.4315/0362-028X.JFP-13-472.

27. Affymetrix. 2014. Expression analysis technical manual with specific protocols for use with the hybridization wash and stain kit. http://media.affymetrix.com/support/downloads/manuals/expression_analysis_technical_manual.pdf.

28. Bolstad BM, Irizarry RA, Astrand M, Speed TP. 2003. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics 19:185–193. http://dx.doi.org/10.1093/bioinformatics/19.2.185.

29. Herrera-Leon S, Ramiro R, Arroyo M, Diez R, Usera MA, Echeita MA. 2007. Blind comparison of traditional serotyping with three multiplex PCRs for the identification of Salmonella serotypes. Res Microbiol 158:122–127. http://dx.doi.org/10.1016/j.resmic.2006.09.009.

30. Harbottle H, White DG, McDermott PF, Walker RD, Zhao S. 2006. Comparison of multilocus sequence typing, pulsed-field gel electrophoresis, and antimicrobial susceptibility typing for characterization of Salmonella enterica serotype Newport isolates. J Clin Microbiol 44:2449–2457. http://dx.doi.org/10.1128/JCM.00019-06.

31. Haley BJ, Cole DJ, Lipp EK. 2009. Distribution, diversity, and seasonality of waterborne salmonellae in a rural watershed. Appl Environ Microbiol 75:1248–1255. http://dx.doi.org/10.1128/AEM.01648-08.

32. Walters SP, Thebo AL, Boehm AB. 2011. Impact of urbanization and agriculture on the occurrence of bacterial pathogens and stx genes in coastal waterbodies of central California. Water Res 45:1752–1762. http://dx.doi.org/10.1016/j.watres.2010.11.032.

33. Thomason BM, Biddle JW, Cherry WB. 1975. Dection of salmonellae in the environment. Appl Microbiol 30:764–767.

34. Parker CT, Huynh S, Quinones B, Harris LJ, Mandrell RE. 2010. Comparison of genotypes of Salmonella enterica serovar Enteritidis phage type 30 and 9c strains isolated during three outbreaks associated with raw almonds. Appl Environ Microbiol 76:3723–3731. http://dx.doi.org/10.1128/AEM.03053-09.

35. Winfield MD, Groisman EA. 2003. Role of nonhost environments in the lifestyles of Salmonella and Escherichia coli. Appl Environ Microbiol 69:3687–3694. http://dx.doi.org/10.1128/AEM.69.7.3687-3694.2003.

36. Beuchat LR. 1996. Pathogenic microorganisms associated with fresh produce. J Food Prot 59:204–216.

37. Wilkoff LJ, Westbrook L, Dixon GJ. 1969. Factors affecting the persistence of Staphylococcus aureus on fabrics. Appl Microbiol 17:268–274.

38. Davies RH, Wray C. 1995. Observations on disinfection regimens used on Salmonella enteritidis infected poultry units. Poult Sci 74:638–647. http://dx.doi.org/10.3382/ps.0740638.

39. Pedersen TB, Olsen JE, Bisgaard M. 2008. Persistence of Salmonella Senftenberg in poultry production environments and investigation of its resistance to desiccation. Avian Pathol 37:421–427. http://dx.doi.org/10.1080/03079450802216561.

40. Bailey JS, Stern NJ, Fedorka-Cray P, Craven SE, Cox NA, Cosby DE, Ladely S, Musgrove MT. 2001. Sources and movement of Salmonella through integrated poultry operations: a multistate epidemiological investigation. J Food Prot 64:1690–1697.

41. Baloda SB, Christensen L, Trajcevska S. 2001. Persistence of a Salmonella enterica serovar Typhimurium DT12 clone in a piggery and in agricultural soil amended with Salmonella-contaminated slurry. Appl Environ Microbiol 67:2859–2862. http://dx.doi.org/10.1128/AEM.67.6.2859-2862.2001.

42. Cobbold RN, Rice DH, Davis MA, Besser TE, Hancock DD. 2006. Long-term persistence of multi-drug-resistant Salmonella enterica serovar Newport in two dairy herds. J Am Vet Med Assoc 228:585–591. http://dx.doi.org/10.2460/javma.228.4.585.

43. Fossler CP, Wells SJ, Kaneene JB, Ruegg PL, Warnick LD, Bender JB, Godden SM, Halbert LW, Campbell AM, Zwald AM. 2004. Prevalence of Salmonella spp on conventional and organic dairy farms. J Am Vet Med Assoc 225:567–573. http://dx.doi.org/10.2460/javma.2004.225.567.

44. Shi X, Namvar A, Kostrzynska M, Hora R, Warriner K. 2007. Persistence and growth of different Salmonella serovars on pre- and postharvest tomatoes. J Food Prot 70:2725–2731.

45. van Overbeek LS, van Doorn J, Wichers JH, van Amerongen A, van Roermund HJW, Willemsen PTJ. 2014. The arable ecosystem as battleground for emergence of new human pathogens. Front Microbiol 5:104. http://dx.doi.org/10.3389/fmicb.2014.00104.

46. Uesugi AR, Danyluk MD, Mandrell RE, Harris LJ. 2007. Isolation of Salmonella Enteritidis phage type 30 from a single almond orchard over a 5-year period. J Food Prot 70:1784–1789.

47. Ravva SV, Sarreal CZ. 2014. Survival of Salmonella enterica in aerated and nonaerated wastewaters from dairy lagoons. Int J Environ Res Public Health 11:11249–11260. http://dx.doi.org/10.3390/ijerph111111249.

48. Luo Z, Gu G, Ginn A, Giurcanu MC, Adams P, Vellidis G, van Bruggen AH, Danyluk MD, Wright AC. 2015. Distribution and characterization of Salmonella enterica isolates from irrigation ponds in the southeastern United States. Appl Environ Microbiol 81:4376–4387. http://dx.doi.org/10.1128/AEM.04086-14.