Salmonella enterica in Swine Production: Assessing the Association between Amplified Fragment Length Polymorphism and Epidemiological Units of Concern\textsuperscript{\dag}\textsuperscript{†}

Bing Wang,\textsuperscript{1,*} Chong Wang,\textsuperscript{1,2} James D. McKean,\textsuperscript{1} Catherine M. Logue,\textsuperscript{3} Wondwossen A. Gebreyes,\textsuperscript{4} Kelly A. Tivendale,\textsuperscript{5} and Annette M. O’Connor\textsuperscript{1}

Department of Veterinary Diagnostic and Production Animal Medicine, College of Veterinary Medicine, Iowa State University, Ames, Iowa 50011; Department of Statistics, College of Liberal Arts and Sciences, Iowa State University, Ames, Iowa 50011; Department of Veterinary Microbiology and Preventive Medicine, College of Veterinary Medicine, Iowa State University, Ames, Iowa 50011; Department of Veterinary Preventive Medicine, College of Veterinary Medicine, Ohio State University, Columbus, Ohio 43210; and Department of Veterinary Science, The University of Melbourne, Parkville, Victoria 3010, Australia

Received 12 January 2011/Accepted 13 September 2011

The aims of this study were to determine the ability of amplified fragment length polymorphism (AFLP) to differentiate Salmonella isolates from different units of swine production and to demonstrate the relatedness of Salmonella between farms and abattoirs by AFLP. Twenty-four farms in the midwestern United States were visited four times from 2006 to 2009. At each farm or abattoir visit, 30 fecal samples or 30 mesenteric lymph nodes were collected, respectively. A total of 220 Salmonella isolates were obtained, serotyped, and genotyped by multilocus sequence typing (MLST) and AFLP. These 220 isolates clustered into 21 serotypes, 18 MLST types, and 14 predominant AFLP clusters based on a genetic similarity threshold level of 60%. To assess genetic differentiation between farms, harvest cohorts, and pigs, analysis of molecular variance was conducted using AFLP data. The results showed 65.62% of overall genetic variation was attributed to variance among pigs, 27.21% to farms, and 7.17% to harvest cohorts. Variance components at the farm (P = 0.003) and pig (P = 0.001) levels were significant, but not at the harvest cohort level (P = 0.079). A second analysis, a permutation test using AFLP data, indicated that on-farm and at-abattoir Salmonella from pigs of the same farms were more related than from different farms. Therefore, among the three subtyping methods, serotyping, MLST, and AFLP, AFLP was the method that was able to differentiate among Salmonella isolates from different farms and link contamination at the abattoir to the farm of origin.

It has been estimated that approximately 1% of Salmonella infections in humans are caused by the consumption of contaminated pork or processed foods derived from pork in the United States (13). Salmonella contamination of pork can be related to preharvest infection and postharvest cross-contamination (1, 4). A recent quantitative risk assessment model aided by meta-analysis suggested that Salmonella-infected pigs entering the abattoir imparted a source of contamination during processing that explained 75% to 80% of the total contamination associated with carcasses (2). These proportions were estimated by construction of a linear regression model between the proportion of Salmonella carrier pigs entering the harvest lines and the proportion of contaminated eviscerated carcasses based on bootstrap simulation. Data used in this study were extracted from primary research studies that used Salmonella spp. as the outcome, reported as presence or absence. Therefore, the link between the epidemiological unit of swine production, i.e., the pigs entering the harvest chain and the pig carcasses, was established based on the regression association rather than evidence of clonal organisms on the carcasses traced from the pigs entering the abattoir. Such an approach to attribution is associated with a high level of uncertainty. However, the approach cannot currently be avoided, because the applied research in pork food safety has focused on the presence or absence of Salmonella spp. or serotyping, tools which do not allow effective attribution.

Although great progress has been made in recent years to understand the ecology of Salmonella in pork, accelerating our understanding of the influence of Salmonella contamination at the preharvest stage on public health will require the use of molecular methods that have a clear link to an epidemiological unit of concern. To purposefully design and conduct risk factor, intervention, or attribution studies for Salmonella in pork, it will be necessary to understand the variation in molecular subtyping methods within the swine production structure, i.e., the epidemiological unit the method is able to differentiate. Such information would inform the unit of differentiation (farm, harvest cohort, pig) at which interventions and risk factors can be assessed using the subtyping methods.

Currently, the most commonly used methods for subtyping Salmonella are serotyping and pulsed-field gel electrophoresis (PFGE). However, these methods have limitations. For example, although serotyping is commonly used for many studies, the method is not useful for estimating risk factors for, or attribution to, levels of swine production because the epidemi-
Salmonella diversity in swine production

MATERIALS AND METHODS

Study population. This study was a post hoc analysis using Salmonella isolates collected for a project with different purposes, for which the epidemiological unit of concern was known. Details of the project have been published elsewhere (23).

For each farm, 1 to 3 days before harvest, approximately 10 g of fresh feces was collected for a project with different purposes, for which the epidemiological units of concern in swine production and how the method should or can be used to further our understanding of the epidemiology of Salmonella in pork.

Given this current gap in knowledge, the aims of this study were to determine the ability of AFLP to differentiate Salmonella isolates from different units of swine production and to demonstrate the relatedness of Salmonella between farms and abattoirs by using AFLP.

AFLP. The AFLP fingerprinting was conducted using the AFLP-based microbial genome-mapping kits (EcoRI microbial AFLP ligation/amplification module and MseI microbial AFLP ligation/amplification module; Life Technologies Corporation, Carlsbad, CA) following the manufacturer’s protocol with some modifications. Briefly, genomic DNA extractions (prepared as for MLST above) were adjusted to a concentration of 50 ng ml⁻¹ in 10 μl water. One hundred nanograms of genomic DNA (2 μl) was digested with the restriction enzymes EcoRI (5 units) and MseI (1 unit), along with 3 μl of 2× restriction buffer 4 (New England BioLabs Inc., Ipswich, MA), 0.5 μl of 1.0-mg ml⁻¹ bovine serum albumin (BSA) at 37°C in a thermal cycler with a heated cover, followed by incubation at 70°C for 15 min. Adaptor oligosequences for EcoRI and MseI from the commercial kit were ligated to the restriction fragments by using 0.08 μl T4 DNA ligase (New England BioLabs Inc., Ipswich, MA), with 1 μl of 1× T4 DNA ligation buffer, 0.13 μl of bovine serum albumin, and then incubated with shaking overnight at 37°C. Dried DNA fragments (4 μl) prepared by restriction and ligation were then initially amplified using EcoRI and MseI core sequence (0.5 μl each) complementary to the EcoRI and MseI end, with 15 μl AFLP amplification core mix (Life Technologies Corporation, Carlsbad, CA). The conditions for preselective amplification were 72°C for 2 min, followed by 20 cycles of 94°C for 20 s, 56°C for 30 s, and 72°C for 2 min. The amplification products were diluted in 0.1× TAE-Tris-EDTA buffer (1× TAE-Tris-EDTA buffer=139 mM Tris-HCl, 5 mM sodium EDTA, pH 8.0) and then incubated with a selective amplification using a carboxyfluorescein (FAM) dye-labeled EcoRI primer (1 μM; 0.5 μl) with an additional adenine base at the 3’ end (EcoRI+A), MseI-selective primer (5 μM; 0.5 μl) with an additional cytosine base at the 3’ end (MseI+C), and 7.5 μl AFLP amplification core mix. The conditions for selective amplification were 94°C for 2 min; 94°C for 20 s, 65°C for 30 s, and decreasing at 1°C/ cycle for 11 cycles to 56°C; then 72°C for 2 min; 19 cycles of 94°C for 20 s, 56°C for 30 s, and 72°C for 2 min; and finally 60°C for 30 min.

The FAM-labeled PCR products were sent to the DNA Facility at Iowa State University for sequencing. The amplified fragments were separated by capillary electrophoresis using an 3730 XL DNA analyzer (Applied Biosystems, Foster City, CA) at a temperature of 4°C for 120 s, injection for 5 s at 20,000 V, and separation at 15.0 kV for 1,600 s. Amplified bands located between 50-bp and 1,000-bp fragment lengths were scored and analyzed using GeneMapper version 4.0 software.

Subtyping methods. Three methods of subtyping were used for comparison, i.e., serotyping, MLST, and AFLP.

Serotyping. All putative Salmonella isolates were submitted to the National Veterinary Services Laboratories for serotyping (USDA NVSL, Ames, IA). A sample was classified as Salmonella positive if its serotype was determined.

MLST. A single colony was inoculated in 1 ml sterile Luria-Bertani broth (LB broth; BD Diagnostic System Inc., Sparks, MD) in a 1.5-ml microcentrifuge tube and then incubated with shaking overnight at 37°C. The culture was centrifuged at 10,000 × g for 2 min, and the supernatant was removed. The pellet was resuspended in 200 μl sterile double-distilled H₂O, boiled at 100°C in a heating block (Eppendorf AG, Hamburg, Germany) for 10 min, and then centrifuged at 10,000 × g for 2 min. The supernatant containing the extracted DNA was aspirated and stored in a sterile microcentrifuge tube at −20°C. MLST of the seven housekeeping genes (aroC, dnaN, hemD, hisD, purE, susC, and thrA) for Salmonella enterica was performed using amplification and sequencing primers according to the MLST scheme found at http://mlst.ucc.ie/mlst/dbs/Senterica/documents/primersEnterica.html.

PFGE is an expensive, low-throughput, labor- and time-intensive method that requires specialized training. Consequently, few large production-level studies have employed it.

An alternative method of molecular subtyping is amplified fragment length polymorphism (AFLP), a PCR-based, high-throughput, relatively inexpensive method. Its use has been reported in studies of Salmonella in swine (11, 12). However, despite the use of AFLP in studies of Salmonella in swine, it is unclear what association the method has with epidemiological units of concern in swine production and how the method should or can be used to further our understanding of the epidemiology of Salmonella in pork.

Given this current gap in knowledge, the aims of this study were to determine the ability of AFLP to differentiate Salmonella isolates from different units of swine production and to demonstrate the relatedness of Salmonella between farms and abattoirs by using AFLP.
were sized with the MapMarker 1000 size standard using the size-calling of the local Southern method. The sum of signal in the range of 50 to 1,000 bp for all samples within the same run was used to perform the signal normalization. The sized peaks were labeled as 1 if the peak amplitude was greater than 100; otherwise they were labeled as 0. The output of genotypes from the GeneMapper software was imported into the BioNumerics version 5.1 software package (Applied Maths, Inc., Austin, TX) for cluster analysis.

The reproducibility of the AFLP protocol was analyzed. One Salmonella strain that was confirmed by culture and serotyping was randomly selected. The AFLP protocol was independently conducted 8 times for this strain, each time starting from the same DNA template extracted as previously described. The resulting AFLP fingerprints of the 8 replicates were then compared using the BioNumerics software. The similarity of the 8 fingerprints of the selected strain ranged from 92.5 to 97.3%, indicating a high level of reproducibility of the AFLP protocol employed here. Based on this result, isolates sharing genetic similarity higher than 92.5% were considered the same AFLP fingerprint.

Statistical analysis. (i) Descriptive analysis. For fecal and mesenteric lymph node samples, the percentage of positive farms or harvest cohorts was calculated with a 95% confidence interval (CI), where the positive farms or harvest cohorts were defined as having at least one Salmonella-positive sample. The median prevalence and range of the within-farm and within-harvest cohort samples were determined.

Further descriptive analyses included comparison of the isolate-level discriminatory abilities of serotyping, MLST, and AFLP using Simpson’s index of diversity (DI) with a 95% CI (17). Simpson’s index of diversity calculated the probability that two unrelated strains sampled from the test population would be placed in different genotypes. Higher values of DI indicated greater discrimination. This study also reported the resolution value for each method, i.e., the number of isolates divided by the type frequency.

To describe the agreement between MLST STs and serotypes, a frequency distribution was carried out to compare matches. Typing system concordance between MLST and serotyping was assessed by the adjusted Rand index (16) using EpiCompare software version 1.0 (Ridom GmbH, Würzburg, Germany).

For this analysis, all possible pairs of isolates were determined by cross-classifying on the basis of matched and mismatched MLST STs and serotypes, and the resulting output was the percentage concordance.

AFLP patterns were used to generate an unweighted pair group method with arithmetic (UPGMA) clustering tree using the Dice similarity coefficient (BioNumerics version 5.1; Applied Maths, Inc., Austin, TX). An arbitrary threshold breakpoint of genetic similarity of 60% was chosen to cluster isolates into distinct AFLP clonal types. This cutoff was selected based on its plausibility using various criteria, particularly phenotypic characteristics (6); therefore, the cutoff was determined by observing the highest concordance shared between AFLP and serotyping/MLST based on the adjusted Rand index. The agreement levels of predominant AFLP clusters with MLST and serotypes were then presented by frequency matrices.

(ii) Evaluation of the epidemiological unit of differentiation: analysis of molecular variance. Analysis of molecular variance (AMOVA) was utilized to reflect the molecular variation at multilevel epidemiological units of concern, i.e., farms, harvest cohorts, and pigs (7). From the AMOVA, the variance components of farm, harvest cohort, and pig level were calculated. Total variance was estimated as the sum of variance components at all levels. The ratio of variance component of farm, harvest cohort, and pig to the total variance was interpreted as the percentage of the total variance that could be attributed to the among-farms difference. The among-harvest cohorts within farms difference, and the among-pigs within harvest cohorts difference. AMOVA was conducted based on the molecular typing method employed with the highest discriminatory ability in this study. The null hypothesis that variance components of all levels were equal to zero was tested using a permutation test with 1,000 iterations. AMOVA and tests for the significance of variance components were conducted for both farm fecal and mesenteric lymph node samples. Data analysis was performed using commands in package ade4 of R (version 2.9.2). We used maximum likelihood estimates of variances to prevent negative variance component estimates by fixing the initially generated negative variance components at zero (21). If the variance component of the harvest cohort did not test significantly different from zero, the inference was that within the same farm, isolates from harvest cohorts were not significantly different from each other in genotypic diversity. Alternatively, if the molecular method was able to differentiate among farms, the farm was considered the epidemiological unit of differentiation.

(iii) Epidemiological relatedness between isolates from farms and abattoirs: permutation test. A permutation test was used to assess the ability of the AFLP method to associate Salmonella contamination at the abattoir to Salmonella infection on farm (R version 2.9.2). The genetic distance between farm fecal isolates (on-farm isolates) and mesenteric lymph node isolates (at-abattoir isolates) from the level of the epidemiological unit was calculated as the test statistic. The results of the aforementioned AMOVA were used to determine the epidemiological unit used for this permutation test.

The null hypothesis was that the average genetic distances between the on-farm and at-abattoir isolates from matched epidemiological units were the same as from unmatched epidemiological units. The genetic distance of any two isolates was defined as the Euclidean distance between their genotypic patterns based on the highest discriminatory method in this study. For matched epidemiological units, the average genetic distance was calculated by summing the genetic distances of all the possible pairs in that unit and then dividing by the number of pairs. The test statistic was the summation of the average genetic distances of isolates from all the epidemiological units. The observed test statistic was calculated using the genetic distances that were from each matched epidemiological unit. The isolates were randomly permuted, and the random test statistic was calculated for each permutation following the same formula as the observed test statistic. The significance of the permutation test was evaluated by comparing the observed test statistic to the distribution of 1,000 random test statistics. Statistical significance was considered at the P < 0.05 level. For example, a one-sided P-value less than 0.05 on the left tail could indicate a significantly strong genetic relationship between the on-farm and at-abattoir isolates from each farm and at-abattoir isolates of pigs.

**RESULTS**

Study population descriptive analysis. Of 27 farms that agreed to participate in the study, 3 subsequently declined, as they were unable to provide pigs at the time required. Twenty-four farms with 50 harvest cohorts (1 to 4 cohorts per farm) were visited to collect samples. Within farms, the visits of adjacent cohorts were 2.5 months apart on average, ranging from 1 to 8 months. The final data set included fecal samples from 21 farms and mesenteric lymph nodes from 16 farms.

Results from fecal samples were missing because poor weather in winter restricted abattoir visits, or the finishing pigs were shipped to the abattoir at a time when Iowa State University staff were unable to make the visit, or due to processing errors that resulted in lost identification. The missing collection of mesenteric lymph nodes occurred due to the original study design, in which we did not propose to collect mesenteric lymph nodes from all cohorts, or due to poor weather.

The overall prevalence of Salmonella was 3.4% (n = 1,490; 95% CI, 2.6 to 4.5%) in farm feces and 20.1% (n = 839; 95% CI, 17.4 to 22.9%) in mesenteric lymph nodes. One or more Salmonella-positive isolates from feces were identified from pigs on 15 of 21 farms (71.4%; 95% CI, 49.8 to 86.1%) and 20 of 50 harvest cohorts (40%; 95% CI, 27.4 to 53.9%). Salmonella-positive isolates were isolated from one or more mesenteric lymph nodes on 12 of 16 farms (75%; 95% CI, 50.1 to 91.5%) and 21 of 28 harvest cohorts (75%; 95% CI, 56.7 to 88.4%). The median prevalence and range of the within-farm and within-cohort positive pig prevalence are shown in Table 1.

A total of 220 Salmonella isolates (51 from farm feces, 169 from mesenteric lymph nodes) were identified. Twenty-one different serotypes and 18 MLST STs were detected among 220 Salmonella isolates. ST684 was newly identified from 2 isolates of S. enterica serovar Uganda and 4 isolates of S. Uganda var. 15+ and was added to the MLST database. A clustering tree generated from AFLP patterns is shown in Fig. S1 in the supplemental material along with the serotypes and MLST STs for each isolate.

Agreement between MLST sequence types and serotypes. A total of 220 Salmonella isolates were included in this analysis.
Most isolates with the same serotype had the same MLST type, except where a single isolate accounted for an additional ST. For example, 66 isolates were serotyped as *S.* Typhimurium (Copenhagen), and all were assigned to ST13, while only 1 isolate was serotyped as *S.* Enteritidis, which was also the only isolate assigned to ST11. However, there were some exceptions. Sixty-one *Salmonella* Derby isolates were assigned to two different STs (17 isolates to ST13 and 44 isolates to ST40). Fifty-four isolates serotyped as *S.* Agona (37; 68.5%) and *S.* Derby (17; 31.5%) were both assigned to ST13 (Fig. 1). The remaining frequencies are shown in Fig. 1. The MLST approach shared 72.2% concordance with serotyping based on a genetic similarity index, at which AFLP shared the highest concordance of 36.9% with serotyping and MLST based on an adjusted Rand index. Thirteen isolates did not cluster with any groups and were considered to be sporadic genotypes. The frequency matrices presenting the agreement of AFLP types with serotypes and MLST types are shown in Fig. 1.

**Comparison of isolate-level discriminatory ability.** A total of 220 *Salmonella* isolates were used in this analysis. Considering each pattern as a distinct subtype, AFLP had the highest DI, 1, followed by serotyping, with a DI of 0.80, and MLST, with a DI of 0.77 (Table 2). On average, 10.47 isolates clustered in a single serotype, a similar resolution to MLST (12.22 isolates per MLST ST). Because it demonstrated the highest resolution for differentiating *Salmonella* isolates from swine origin, only genotypic results of AFLP were selected for the epidemiological unit of differentiation and epidemiological relatedness analyses.

**Epidemiological unit of differentiation: AMOVA.** Fifty-one isolates from farm feces were used for the on-farm epidemiological unit of differentiation analysis. For the farm fecal samples, the percentage of the among-pigs-within-harvest cohorts variance component was 65.62%, followed by among-farm (27.21%) and among-harvest cohorts within farms (7.17%). The variance components of among farms and among pigs within harvest cohorts were significantly different from zero (\( P = 0.003 \) for among farms; \( P = 0.001 \) for among pigs among harvest cohorts) (Table 3).

A total 169 isolates from mesenteric lymph nodes were used for the at-abattoir epidemiological unit of differentiation analysis. For the mesenteric lymph node samples, most of the AFLP type diversity was found among pigs within harvest cohorts (75.19%), followed by data among harvest cohorts within farms (24.81%) and among farms (0%); the variance components of among harvest cohorts within farms and among pigs within harvest cohorts were statistically significant (for both, \( P = 0.001 \)) (Table 3).

The AMOVA of on-farm and at-abattoir isolates inferred (i) pigs and carcasses from a specific farm carried *Salmonella* isolates with substantial genotypic variation; (ii) the genotypic distribution of the *Salmonella* population in live pigs on a specific farm did not change significantly across harvest cohorts of pigs from the same farm; (iii) AFLP could differentiate the on-farm *Salmonella* isolates originating from the different farms.

**Epidemiological relatedness between isolates from farms and abattoirs: permutation test.** AMOVA (above) indicated that AFLP could differentiate *Salmonella* isolates across farms; therefore, the epidemiological unit of differentiation of AFLP was the farm. The permutation procedure was used to detect the relatedness of on-farm and at-abattoir isolates from the same farms. A total of 108 isolates (18 from farm feces and 90 from mesenteric lymph nodes) from the matched farms were used in this analysis (Table 4).

The permutation test results showed \( P = 0.036 \) on the left tail (Fig. 2), that is, a 3.6% or lower probability that the sum of the average genetic distance of each farm was less than the observed average genetic distance for matched farms. The shorter the genetic distance, the stronger the genetic relationship. Therefore, the AFLP subtyping method was capable of demonstrating significant epidemiological relatedness between *Salmonella* isolates of pigs and carcasses from the same farms.

**DISCUSSION**

The aims of this study were to determine the ability of AFLP to differentiate *Salmonella* isolates from different units of swine production and to demonstrate the relatedness of *Salmonella* between farms and abattoirs by AFLP. Many methods have been developed to differentiate isolates of a particular bacterial species, and many publications have catalogued the ability of these methods to differentiate these organisms. However, molecular epidemiology requires not just the ability to differentiate but also the ability to associate “different isolates” with meaningful epidemiological outcomes. Such knowledge is critical to our understanding of the ecology of microorganisms and the epidemiology of disease. An example of this idea is the use of PFGE to identify food source contamination in salmonellosis outbreaks. The ability of PFGE to identify pathogens in contaminated food sources with those from human cases within the same outbreak and to separate distinct outbreaks is...
Interestingly, the employment of PFGE in outbreak investigations is currently being scrutinized. In the multistate *Salmonella Enteritidis* outbreak related to shell eggs in 2010, PFGE was not sufficiently discriminatory for the outbreak because *Salmonella* isolates that appeared to have no epidemiological link to the outbreak had similar PFGE patterns (5). To determine the preharvest source of *Salmonella* contamination and assess the impact of

![FIG. 1. The frequency matrix of isolates typed by serotyping, MLST, and AFLP types.](http://aem.asm.org/)

- (A) Agreement between serotypes and MLST STs.
- (B) Agreement between serotypes and AFLP types.
- (C) Agreement between MLST STs and AFLP types.
interventions in the context of swine production, the application of an inappropriate genotyping method could produce a similar issue. Previous studies have shown that some methods of differentiating *Salmonella* are unable to associate *Salmonella* isolates with the epidemiological units of concern in swine production (8). For example, serotyping and MLST were not useful in attributing *S.* Typhimurium or a sequence type found at the abattoir to a particular farm, because the serotype or sequence type could be found in most of the farms (8). Therefore, it is necessary to attribute *Salmonella* contamination at the abattoir to the farm origin by adopting an appropriate genotyping method that is able to associate the unit of differentiation with the unit of production.

The AMOVA for isolates from farm feces used in this study indicated that AFLP was able to differentiate *Salmonella* populations originating from multiple swine farms. The ability to differentiate between farms might make it possible to attribute *Salmonella* contamination at the abattoir back to a particular farm, because the serotype or sequence type could be found in most of the farms (8). Therefore, it is necessary to attribute *Salmonella* contamination at the abattoir to the farm origin by adopting an appropriate genotyping method that is able to associate the unit of differentiation with the unit of production.

The AMOVA for isolates from farm feces used in this study indicated that AFLP was able to differentiate *Salmonella* populations originating from multiple swine farms. The ability to differentiate between farms might make it possible to attribute *Salmonella* contamination at the abattoir back to a particular farm, because the serotype or sequence type could be found in most of the farms (8). Therefore, it is necessary to attribute *Salmonella* contamination at the abattoir to the farm origin by adopting an appropriate genotyping method that is able to associate the unit of differentiation with the unit of production.

The AMOVA for isolates from farm feces used in this study indicated that AFLP was able to differentiate *Salmonella* populations originating from multiple swine farms. The ability to differentiate between farms might make it possible to attribute *Salmonella* contamination at the abattoir back to a particular farm, because the serotype or sequence type could be found in most of the farms (8). Therefore, it is necessary to attribute *Salmonella* contamination at the abattoir to the farm origin by adopting an appropriate genotyping method that is able to associate the unit of differentiation with the unit of production.

Table 2 shows a comparison of abilities of serotyping, MLST, and AFLP to discriminate between 220 *Salmonella enterica* isolates.

**Table 2. Comparison of abilities of serotyping, MLST, and AFLP to discriminate between 220 Salmonella enterica isolates**

<table>
<thead>
<tr>
<th>Method</th>
<th>Type frequency</th>
<th>DI (95% CI)</th>
<th>Resolution value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Serotyping</td>
<td>21</td>
<td>0.80 (0.77–0.83)</td>
<td>10.47</td>
</tr>
<tr>
<td>MLST</td>
<td>18</td>
<td>0.77 (0.74–0.80)</td>
<td>12.22</td>
</tr>
<tr>
<td>AFLP</td>
<td>220</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

* The resolution value was calculated by dividing the number of isolates by the type frequency.

* Considering each pattern as a distinct subtype, each isolate had a different AFLP type.

**Table 3. AMOVA results of AFLP data for Salmonella enterica isolates categorized by farm and by harvest cohort of pigs**

<table>
<thead>
<tr>
<th>Sample source and analysis group</th>
<th>Degrees of freedom</th>
<th>Sum of squares</th>
<th>Variance component</th>
<th>% of total variance</th>
<th>P value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Farm feces</td>
<td>14</td>
<td>808.72</td>
<td>9.68</td>
<td>27.21</td>
<td>0.003</td>
</tr>
<tr>
<td>Among harvest cohorts within farms</td>
<td>5</td>
<td>156.62</td>
<td>2.55</td>
<td>7.17</td>
<td>0.079</td>
</tr>
<tr>
<td>Among pigs within harvest cohorts</td>
<td>33</td>
<td>770.21</td>
<td>23.34</td>
<td>65.62</td>
<td>0.001</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mesenteric lymph nodes</th>
<th>Degrees of freedom</th>
<th>Sum of squares</th>
<th>Variance component</th>
<th>% of total variance</th>
<th>P value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Among farms</td>
<td>11</td>
<td>988</td>
<td>0</td>
<td>0</td>
<td>0.82</td>
</tr>
<tr>
<td>Among harvest cohorts within farms</td>
<td>9</td>
<td>699.34</td>
<td>7.88</td>
<td>24.81</td>
<td>0.001</td>
</tr>
<tr>
<td>Among pigs within harvest cohorts</td>
<td>146</td>
<td>3,486.03</td>
<td>23.88</td>
<td>75.19</td>
<td>0.001</td>
</tr>
</tbody>
</table>

* P values are for testing the null hypothesis that variance components of different levels are equal to zero.
Salmonella required to identify enough be needed for a particular farm to account for between-pig substantial pig-level noise, and many pig-level samples would be used to differentiate between farms, but there would be the time period of the present study. AFLP could potentially Salmonella method, these results suggest that the genetic diversity of a significance level of 0.05. Using AFLP as a differentiation ance "among harvest cohorts within farms" did not attain the significant factor in explaining genetic diversity, while the vari-ance "among pigs within harvest cohorts" and "among farms" was a three times greater than "among farms." Both the variance attributable to "among pigs within harvest cohorts" was three times greater than "among farms." Both the variance "among pigs within harvest cohorts" and "among farms" was a significant factor in explaining genetic diversity, while the vari-ance "among harvest cohorts within farms" did not attain the significance level of 0.05. Using AFLP as a differentiation method, these results suggest that the genetic diversity of a Salmonella population on a farm is relatively small across co-horts harvested within a relatively short period, such as within the time period of the present study. AFLP could potentially be used to differentiate between farms, but there would be substantial pig-level noise, and many pig-level samples would be needed for a particular farm to account for between-pig variation. Therefore, a large number of fecal samples would be required to identify enough Salmonella isolates to capture the genotypic distribution of the Salmonella population on a par-ticular swine farm.

The permutation test for epidemiological relatedness indicated that AFLP could be used to identify the flowthrough contamination from farms to abattoirs. The contaminated car-casses at the abattoir can be attributed to infected pigs (flow-through contamination) as well as healthy but later cross-in-fected pigs prior to harvest. Previous studies have documented that transportation via trucks and holding in lairage pens were major sources for Salmonella contamination after pigs leave swine farms (19, 20). Therefore, Salmonella isolated at the abattoir can arise partly from the original pigs on the farms but also from other pigs from different cohorts or farms. Following this theory, the observed test statistic should be in the middle part of the permutation test statistic distribution, which suggests no genetic relationship between the on-farm and at-abattoir isolates originating from pigs of the same farm. However, the permutation test results indicated that Salmonella isolates from farms and abattoirs originating from the same farm were genetically correlated. The genetic relatedness im-plies that the introduction of external source isolates from truck or lairage transit infection was not sufficient to block the genetic link from farms to abattoirs, which makes it possible to trace the postharvest Salmonella contamination back to a par-ticular farm by using AFLP.

MLST and serotyping methods were not capable of discrim-inating between epidemiological units of concern. Compared to serotyping and PFGE, MLST has the advantage of being reproducible and easily exchanged between laboratories, but MLST is not clearly associated with an epidemiological unit of concern. Others have reported the use of MLST in Salmonella from swine, for example, a total of 110 S. enterica isolates were typed using the seven-gene MLST scheme (as used in this study) and 43 STs were identified (22). However, the epidemi-ological origins of the human and veterinary isolates were not presented, and the level of epidemiological unit that MLST could differentiate was not explored. For example, it was not reported whether the MLST method could differentiate among human and veterinary isolates from different outbreaks (22). Similarly, Fakhr et al. (8) conducted MLST (4-gene scheme: manB, pduF, glxA, and spaM) for the genetic discrimination of 85 S. Typhimurium isolates and found no genetic diversity among the isolates tested, with 100% identity to the sequence reported in GenBank for the S. Typhimurium LT 2 strain, and obviously no link to an epidemiological unit was reported. The limited discrimination of MLST between closely related isolates may be due to a relatively small part of the genome being used in an MLST investigation, as well as to a moderate-to-slow mutation rate within the targeted housekeeping genes (15).

In this study, one of our interests was the apportionment of total variation to different sources. However, a negative esti-mate of the variance component of the farm level was observed from the R output of AMOVA when applied to the isolates from mesenteric lymph nodes. One possible reason for the negative variance component of the farm level might be that the variability among cohorts was too large. The expected variation among farms was calculated as the sum of the ex-pected variation among cohorts and the product of the vari-ance component of farm level and appropriate degrees of freedom. Therefore, a negative variance component of farm level could occur when the observed variation of “among har-vest cohorts within farms” is greater than “among farms.” One explanation for the greater “among harvest cohorts within farms” variation when applied to the mesenteric lymph node isolates is that Salmonella contamination of a harvest cohort at the abattoir has influences from both farms and abattoir lair-age pens. As discussed previously, the lairage pens potentially represent a sampling of all farms for that day/week. Salmonella can be isolated from pig mesenteric lymph nodes 2 h after the

![Histogram of test statistic of average genetic distance between farm and abattoir samples](http://aem.asm.org/)

**FIG. 2.** Histogram of test statistic of average genetic distance between farm and abattoir samples, when the epidemiological unit of differentiation of AFLP was the farm. The histogram is based on 1,000 permutations. The vertical line is the test statistic and is based on observed data from truly matched farms. Of 1,000 simulated farm-level test statistics, only 3.6% simulations reported a test statistic smaller than the observed one.

of Salmonella at the pig level; i.e., AFLP identified 220 types (based on a 100% similarity threshold level) among 220 Salmonella isolates. Compared to serotyping and MLST, the issue is “too much” differentiation rather than too little. The AMOVA partitioned the overall diversity of Salmonella iso-lates into multiple epidemiological units of concern. The variance attributable to “among pigs within harvest cohorts” was three times greater than “among farms.” Both the variance “among pigs within harvest cohorts” and “among farms” was a significant factor in explaining genetic diversity, while the vari-ance “among harvest cohorts within farms” did not attain the significance level of 0.05. Using AFLP as a differentiation method, these results suggest that the genetic diversity of a Salmonella population on a farm is relatively small across co-horts harvested within a relatively short period, such as within the time period of the present study. AFLP could potentially be used to differentiate between farms, but there would be substantial pig-level noise, and many pig-level samples would be needed for a particular farm to account for between-pig variation. Therefore, a large number of fecal samples would be required to identify enough Salmonella isolates to capture the genotypic distribution of the Salmonella population on a par-ticular swine farm.

The permutation test for epidemiological relatedness indicated that AFLP could be used to identify the flowthrough contamination from farms to abattoirs. The contaminated car-casses at the abattoir can be attributed to infected pigs (flow-through contamination) as well as healthy but later cross-in-fected pigs prior to harvest. Previous studies have documented that transportation via trucks and holding in lairage pens were major sources for Salmonella contamination after pigs leave swine farms (19, 20). Therefore, Salmonella isolated at the abattoir can arise partly from the original pigs on the farms but also from other pigs from different cohorts or farms. Following this theory, the observed test statistic should be in the middle part of the permutation test statistic distribution, which suggests no genetic relationship between the on-farm and at-abattoir isolates originating from pigs of the same farm. However, the permutation test results indicated that Salmonella isolates from farms and abattoirs originating from the same farm were genetically correlated. The genetic relatedness im-plies that the introduction of external source isolates from truck or lairage transit infection was not sufficient to block the genetic link from farms to abattoirs, which makes it possible to trace the postharvest Salmonella contamination back to a par-ticular farm by using AFLP.

MLST and serotyping methods were not capable of discrim-inating between epidemiological units of concern. Compared to serotyping and PFGE, MLST has the advantage of being reproducible and easily exchanged between laboratories, but MLST is not clearly associated with an epidemiological unit of concern. Others have reported the use of MLST in Salmonella from swine, for example, a total of 110 S. enterica isolates were typed using the seven-gene MLST scheme (as used in this study) and 43 STs were identified (22). However, the epidemi-ological origins of the human and veterinary isolates were not presented, and the level of epidemiological unit that MLST could differentiate was not explored. For example, it was not reported whether the MLST method could differentiate among human and veterinary isolates from different outbreaks (22). Similarly, Fakhr et al. (8) conducted MLST (4-gene scheme: manB, pduF, glxA, and spaM) for the genetic discrimination of 85 S. Typhimurium isolates and found no genetic diversity among the isolates tested, with 100% identity to the sequence reported in GenBank for the S. Typhimurium LT 2 strain, and obviously no link to an epidemiological unit was reported. The limited discrimination of MLST between closely related isolates may be due to a relatively small part of the genome being used in an MLST investigation, as well as to a moderate-to-slow mutation rate within the targeted housekeeping genes (15).

In this study, one of our interests was the apportionment of total variation to different sources. However, a negative esti-mate of the variance component of the farm level was observed from the R output of AMOVA when applied to the isolates from mesenteric lymph nodes. One possible reason for the negative variance component of the farm level might be that the variability among cohorts was too large. The expected variation among farms was calculated as the sum of the ex-pected variation among cohorts and the product of the vari-ance component of farm level and appropriate degrees of freedom. Therefore, a negative variance component of farm level could occur when the observed variation of “among har-vest cohorts within farms” is greater than “among farms.” One explanation for the greater “among harvest cohorts within farms” variation when applied to the mesenteric lymph node isolates is that Salmonella contamination of a harvest cohort at the abattoir has influences from both farms and abattoir lair-age pens. As discussed previously, the lairage pens potentially represent a sampling of all farms for that day/week. Salmonella can be isolated from pig mesenteric lymph nodes 2 h after the
animals become exposed to a Salmonella-contaminated environment (18). Therefore, Salmonella isolates from the pigs of a harvest cohort at the abattoir can potentially originate from different farms. In addition, the lack of enough numbers of cohorts per farm might be the reason for the large variance estimation for “among cohorts.” We conducted an ad hoc solution to fix the variance component of farm level as zero. The restriction was reasonable, because the P value (0.82) showed the farm-level variance component was not significantly different from zero.

In this study, the error of the AMOVA model was variance between pigs. We identified one Salmonella isolate per pig, which ignored the variation among isolates within pigs. It is possible that different genotypes from multiple Salmonella isolates occur in a single pig. However, because our aim was to find a tool useful for Salmonella preharvest intervention and such an intervention at the farm or harvest cohort level is more relevant, it might not be necessary to be concerned about the isolate-isolate variation below the pig level.

Overall, the findings of this study have important implications. First, AFLP could differentiate Salmonella isolates between the epidemiological unit farms, but not the harvest cohorts within farms. Second, AFLP could link the abattoir contamination to the farm origin. Finally, there is substantial pig-to-pig genotypic variation. Although using AFLP to subtype Salmonella isolates on a swine farm might be able to capture the genetic character of the Salmonella population on that swine farm, practical application of AFLP for molecular epidemiology in the market chain requires large sample sizes per source farm.

ACKNOWLEDGMENTS

We thank all the swine farms and producers who participated. We thank the swine product safety laboratories at Iowa State University for collecting and determining Salmonella enterica isolates. We thank Daniel Tadesse (Wondwosen Gebreyes’ laboratory) for offering the AFLP protocol. We thank Robert C. Foster (Chong Wang’s student) for writing the function for the permutation test using R.

Funding for this work was provided by USDA-CSREES grant 2005-35212-15929.

REFERENCES

5. Centers for Disease Control and Prevention. 2 December 2010, final posting date. Investigation update: multistate outbreak of human Salmonella Enteritidis infections associated with shell eggs. Centers for Diseases Control and Prevention, Atlanta, GA.