Integrating Whole Genome Sequencing of Salmonella enterica Serovar Enteritidis into the Public Health Laboratory for Surveillance and Outbreak Investigations


INTRODUCTION

Salmonella enterica serovar Enteritidis is a leading cause of foodborne illness in the United States. The low genetic diversity of S. Enteritidis limits how well pulse-field gel electrophoresis (PFGE) can detect outbreaks of enteric pathogens. To improve discrimination between sporadic and outbreak-associated isolates, the Wadsworth Center performs whole genome sequencing (WGS) single nucleotide polymorphism (SNP) based phylogenetic typing on all S. Enteritidis isolates in addition to PFGE typing.

Study Objectives:
- Identify the benefits and challenges of incorporating WGS-based typing into routine surveillance.
- Establish an efficient means for reporting that is useful for both laboratorians and epidemiologists.

Methods

Whole genome sequencing subdivides PFGE types

- PFGE Type
- Number of GCs Detected
- Number of Isolates in GC (% of total)
- Total number of Isolates in PFGE Type

<table>
<thead>
<tr>
<th>PFGE Type</th>
<th>Number of GCs Detected</th>
<th>Number of Isolates in GC (% of total)</th>
<th>Total number of Isolates in PFGE Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>JEGX01.0002</td>
<td>8</td>
<td>25 (52.1)</td>
<td>48</td>
</tr>
<tr>
<td>JEGX01.0004</td>
<td>45</td>
<td>121 (56.0)</td>
<td>216</td>
</tr>
<tr>
<td>JEGX01.0005</td>
<td>13</td>
<td>64 (66.7)</td>
<td>96</td>
</tr>
<tr>
<td>JEGX01.0021</td>
<td>12</td>
<td>20 (54.1)</td>
<td>37</td>
</tr>
<tr>
<td>JEGX01.0034</td>
<td>5</td>
<td>10 (40.0)</td>
<td>25</td>
</tr>
<tr>
<td>JEGX01.0065</td>
<td>5</td>
<td>10 (58.8)</td>
<td>17</td>
</tr>
<tr>
<td>JEGX01.0001</td>
<td>1</td>
<td>14 (83.3)</td>
<td>15</td>
</tr>
<tr>
<td>JEGX01.0019</td>
<td>1</td>
<td>4 (80.0)</td>
<td>5</td>
</tr>
<tr>
<td>JEGX01.0023</td>
<td>6</td>
<td>8 (80.0)</td>
<td>10</td>
</tr>
<tr>
<td>JEGX01.0030</td>
<td>2</td>
<td>10 (100)</td>
<td>5</td>
</tr>
<tr>
<td>JEGX01.0199</td>
<td>0</td>
<td>0 (0)</td>
<td>2</td>
</tr>
<tr>
<td>JEGX01.0616</td>
<td>1</td>
<td>3 (100)</td>
<td>3</td>
</tr>
</tbody>
</table>

For endemic PFGE types (highlighted in gray) WGS analysis reveals many covert clusters
- For non-endemic rare PFGE types WGS confirms associations

17% of Genomic Clusters Contain Two PFGE Patterns

For the most common pairs shown below, this occurs from loss of a 59kb plasmid (SLA5)

- 53 genomic clusters contain JEGX01.0004 and/or JEGX01.0034
- 13 genomic clusters contain JEGX01.0005 and/or JEGX01.0030
- 13 genomic clusters contain JEGX01.0021 and/or JEGX01.0023

Frequency Distribution of Isolates in Genomic Clusters (GC)

Most genomic clusters were comprised of 4 or fewer isolates (n=86 genomic clusters, mean=3.37, SD=2.87)

Study Description

502 isolates sequenced in real time

Collected over 598 days (8/27/2013 to 4/16/2015)
32 PFGE patterns represented
- 438/502 (87%) isolates were part of endemic patterns
86 Genomic Clusters (GC) identified in dataset
- GC defined as 54 SNPs diversity among isolates
- 1-5 (17%) GC’s contained 2 different PFGE patterns

Cluster resolution is improved by WGS for 502 prospectively collected clinical samples

Prioritizing Genomic Clusters of Interest (GCOI) for rapid epidemiological follow-up:

The LLWW Plot

Takes into account the frequency of isolate acquisition into a cluster:
- Variables can be changed depending on need

SUMMARY

WGS can subdivide endemic PFGE patterns into genomic clusters and can better discriminate between genomic clusters and sporadic.
- Created a tool to prioritize genomic clusters that can be utilized in real time by both epidemiologists and laboratorians.

ACKNOWLEDGMENTS

We gratefully acknowledge support of this research by the Wadsworth Center Public Health Genomics Initiative, the Wadsworth Center Applied Genomic Technologies and Bioinformatics Cores, and the FDA Genome Tracker Initiative.

REFERENCES
