Perspectives from a Food Regulatory Laboratory: FLORIDA AG’s Experience with WGS Implementation

InFORM 2017
Serena Giovinazzi, Ph.D.

Bureau of Food Laboratories
Division of Food Safety
Florida Department of Agriculture and Consumer Services
Florida Department of Agriculture and Consumer Services (FDACS)
FDACS - Food Laboratory
Accreditation & Collaboration

Accreditation

The Bureau of Food Laboratories gained accreditation to ISO-17025, for the specific tests listed in A2LA certificates 2534.01 and 2534.02 on March 31, 2007.

Collaborations

- **PulseNet** (since 2004)
- FSIS FERN CAP (Microbiology)
- FDA FERN CAP (Microbiology) (shared w/ Chem Residue)
- FDA ISO Accreditation
- **GenomeTrakr** (since 2015)
- LRN
- FL DOH
- APHL
- eLEXNET
Food Testing and Surveillance

Targets Ready-to-Eat Foods for Pathogen Analysis

- Deli Meat
- Cheese
- Sandwiches
- Smoked Fish
- Produce
- Fresh Squeezed Juice
- Ice Cream
Where Do Our Samples Come From?

Within FDACS:
- Division of Food Safety
  - Bureau of Food Inspection
  - Bureau of Dairy Industry
  - Bureau of Chemical Residue
  - Agriculture Environmental Services

Within Florida:
- Department of Health
- Department of Law Enforcement

Federal Agencies:
- FDA: contract (RRT/MFRPS)
- Cooperative Agreements (FSIS/FDA)

Food Inspections
14 Districts
120 Inspectors

Recalls Last Year: 20 (Division of Food Safety)
Molecular Subtyping & Outbreak Detection Workflow

**Foodborne Pathogen**
- Food Commodity
- Microbial Isolation
- *Salmonella, L. monocytogenes, E. Coli O157:H7 and non-O157*

**PFGE**
- PulseNet Submission

**WGS**
- DNA isolation (Qiagen Kit)
- Library prep (Nextera XT)
- Cluster generation
- Sequencing (FastQ Only)
- Analysis (In-House Pipeline)

*Metadata sharing*
In-House Basic WGS Bioinformatics

Data from MiSeq

Raw Sequence Read QC, *de novo* Genome Assembly
In-house pipeline: FastQC -> MetaPhlAn2 -> SPAdes -> QUAST -> BBMap -> Bandage

Metrics Report for Data and Assembly Quality
BBMap, QUAST, MetaPhlAn2

Sequence Comparison VS Database
kSNP3, FastTree2, FigTree

Detect SNPs & Identify SNP Distances
CFSAN SNP pipeline, FastTree2

Produce Phylo Tree
FigTree
## Run Metrics & Quality Control

<table>
<thead>
<tr>
<th>Assembly</th>
<th>#Contigs</th>
<th>Total Length</th>
<th>Average Coverage</th>
<th>Average Insert Size</th>
<th>Pct_Lm</th>
<th>Pct_Slm</th>
<th>Ptc_Ec</th>
<th>Pct_Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample 1</td>
<td>17</td>
<td>3014256</td>
<td>167.42</td>
<td>409.99</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Sample 2</td>
<td>13</td>
<td>3042044</td>
<td>171.09</td>
<td>387.13</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Sample 3</td>
<td>12</td>
<td>3042057</td>
<td>209.09</td>
<td>377.87</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Sample 4</td>
<td>13</td>
<td>3041526</td>
<td>195.65</td>
<td>377.56</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Sample 5</td>
<td>19</td>
<td>3042033</td>
<td>122.33</td>
<td>455.72</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Sample 6</td>
<td>22</td>
<td>3042205</td>
<td>213.15</td>
<td>359.91</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Sample 7</td>
<td>20</td>
<td>3122563</td>
<td>169.19</td>
<td>419.09</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Sample 8</td>
<td>17</td>
<td>2978936</td>
<td>201.67</td>
<td>420.68</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Sample 9</td>
<td>78</td>
<td>5990293</td>
<td>86.16</td>
<td>207.64</td>
<td>55.81541</td>
<td>0</td>
<td>0</td>
<td>44.19459</td>
</tr>
<tr>
<td>Sample 10</td>
<td>17</td>
<td>2990560</td>
<td>211.67</td>
<td>410.56</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Sample 11</td>
<td>23</td>
<td>3112224</td>
<td>108.37</td>
<td>461.13</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Sample 12</td>
<td>17</td>
<td>3041346</td>
<td>142.30</td>
<td>433.72</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Sample 13</td>
<td>15</td>
<td>3041334</td>
<td>186.01</td>
<td>410.94</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Sample 14</td>
<td>18</td>
<td>3047430</td>
<td>171.57</td>
<td>413.28</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Sample 15</td>
<td>21</td>
<td>3045563</td>
<td>181.17</td>
<td>416.88</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Sample 16</td>
<td>48</td>
<td>5097865</td>
<td>113.62</td>
<td>401.26</td>
<td>0</td>
<td>100</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

**Assembly Stats**

**Library Prep Quality**

**Metagenomic Verification**

- OK; Lm
- KO; Data retracted
- OK; Slm
Metagenomic Verification

WGS helps us identify contaminated historical isolates
In-House Basic WGS Bioinformatics

Data from MiSeq or NCBI

Raw Sequence Read QC, *de novo* Genome Assembly
In-house pipeline: FastQC -> MetaPhlAn -> SPAdes -> QUAST -> BBMap -> Bandage

Metrics Report for Data and Assembly Quality
BBMap, QUAST, MetaPhlAn2

Sequence Comparison *VS* Database
kSNP3, FastTree2, FigTree

Detect SNPs & Identify SNP Distances
CFSAN SNP pipeline FastTree2

Produce Phylo Tree
FigTree
Current *L. mono* Phylogenetic Tree

- **kSNP3**
  - (Reference-Free SNP Identification)
- **FastTree**
  - (SNP Phylogenetics)
- **Figtree**
  - (Graphical Viewer)

807 isolates
In-House Basic WGS Bioinformatics

Data from MiSeq or NCBI

Raw Sequence Read QC, *de novo* Genome Assembly
In-house pipeline: FastQC -> MetaPhlAn -> SPAdes -> QUAST -> BBMap -> Bandage

Metrics Report for Data and Assembly Quality
BBMap, QUAST, MetaPhlAn2

Sequence Comparison VS Database
kSNP3, FastTree2, FigTree

Detect SNPs & Identify SNP Distances
CFSAN SNP pipeline
FastTree2

Produce Phylo Tree
FigTree

Brian Caudle
Example Case 1: WGS Analysis in Action

Routine surveillance of Firm X in May 2017
Follow-up samples with matching PFGE pattern
Stop Sales until CLEAR!
WGS analysis: Surveillance and Follow-up samples <10 SNPs
Environmental swabbing on June 1st: Zone 1 & 2 only
Matching environmental were detected in the swabs collected in the meat case by PFGE and WGS
Firm X cleaned and re-opened by the July 4th week-end
Example Case 2: WGS Analysis in Action

Routine surveillance of Firm Y in July 2017: 5 L. mono isolates

1 week later: Follow up samples all NEGATIVE

October 2017: 5 L. mono with matching PFGE pattern

1 week later: 3 Follow-up positives matching by PFGE and WGS

Environmental collected: 13 isolated that matched previous PFGE patterns

Firm Y Deli is closed until clear. No clinical isolates to date.

Less than 3 SNPs apart
**WGS in Outbreak Investigation**

- **May 2014**
  - Routine surveillance of Firm A
  - July 2014: Matching environmentals detected

- **March 2015**
  - Routine surveillance of Firm B
  - Samples matched Firm A isolates
  - Follow-up samples also contained *L. monocytogenes*

- **March 2015**
  - Routine surveillance of Firm C
  - Samples matched Firm A and B isolates
  - Matching environmentals

---

**Total 30 *L. monocytogenes* isolates**

**10 Identical Ascl patterns & very similar Apal patterns**
FLAG was approached by DOH to run WGS analysis on a group of 8 clinical isolates with a potential epi link to firm C.
No WGS matches between food/environmentals and clinical

No concrete evidence to connect all 3 Firms

Firm C temporarily closes in Oct 2015 due to a Food Inspection “stop-processing” order

FLAG maintains routine surveillance (all Firms)
<table>
<thead>
<tr>
<th>Date</th>
<th>Sample Type</th>
<th>Food Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>2017-08-15</td>
<td>Egg Salad</td>
<td>Food</td>
</tr>
<tr>
<td>2017-08-22</td>
<td>HamCheese Sandwich</td>
<td>Food</td>
</tr>
<tr>
<td>2017-08-22</td>
<td>HamCheese Sandwich</td>
<td>Food</td>
</tr>
<tr>
<td>2017-08-15</td>
<td>Egg Salad</td>
<td>Food</td>
</tr>
<tr>
<td>2017-08-22</td>
<td>HamCheese Sandwich</td>
<td>Food</td>
</tr>
<tr>
<td>2017-08-22</td>
<td>HamCheese Sandwich</td>
<td>Food</td>
</tr>
</tbody>
</table>
Conclusions

Concurrently run PFGE and WGS

WGS and the analysis pipeline helped us

- Develop a state-wide local database
- In depth data QC, assess data quality and isolate contaminations
- Local outbreak cluster detection
- To better aid our customers (e.g., direct environmental swabbing)
- During outbreak investigations
- To prove/disprove connection between food, environmental and/or clinical isolates
Future Directions

BioNumerics 7.6

- Ease of access to clinical WGS data

Further training of analytical staff in bioinformatics along with automation efforts

Promote openness to accept isolates or data for regional (or national) support

- Projected to sequence Campy isolates from NC in 2018

Add WGS method under the ISO-17025 accreditation scope
FLORIDA WGS Data

**Listeria monocytogenes**
- FLAG: 882 Food isolates
- FLDOH: 104 Clinical isolates

**Escherichia coli**
- FLAG: 3 Food isolates
- FLDOH: 52 Clinical isolates

**Salmonella**
- FLAG: 5 Food isolates
- FLDOH: 863 Clinical isolates

**GOAL**: Real-time addition of FL-DOH clinical samples for a Florida-centric foodborne pathogen database
Acknowledgements

Thank you for the continued support and collaboration

- Brian Caudle, FLAG
  - For developing the FLAG WGS analysis pipeline

- Rita Johnson/Summer Williams, Florida RRT
  - For funding the MiSeq, computer, and reagents

- Ruth Timme, PhD and Maria Sanchez Leon
  - FDA/CFSAN – GenomeTrakr network

- PulseNet & APHL

- Jason Blanton, PhD, Florida Department of Health
  - For sharing the FL-DOH data
### Contacts

<table>
<thead>
<tr>
<th>Name</th>
<th>Email</th>
<th>Position</th>
<th>Phone</th>
</tr>
</thead>
<tbody>
<tr>
<td>Serena Giovinazzi, PhD</td>
<td><a href="mailto:Serena.Giovinazzi@freshfromflorida.com">Serena.Giovinazzi@freshfromflorida.com</a></td>
<td>Environmental Specialist III</td>
<td>850.488.4226</td>
</tr>
<tr>
<td>Jason Crowe, PhD</td>
<td><a href="mailto:Jason.Crowe@freshfromflorida.com">Jason.Crowe@freshfromflorida.com</a></td>
<td>Biological Administrator II</td>
<td>850.617.7562</td>
</tr>
<tr>
<td>Patty Lewandowski</td>
<td><a href="mailto:Patty.Lewandowski@freshfromflorida.com">Patty.Lewandowski@freshfromflorida.com</a></td>
<td>Bureau Chief</td>
<td>850.617.7559</td>
</tr>
<tr>
<td>Brian Caudle</td>
<td><a href="mailto:Stanley.Caudle@freshfromflorida.com">Stanley.Caudle@freshfromflorida.com</a></td>
<td>Biological Scientist IV</td>
<td>850.617.7560</td>
</tr>
</tbody>
</table>
Thank you!

QUESTIONS?
### Software Requirements/Analytical Tools

<table>
<thead>
<tr>
<th>Analytical Tool</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>De Novo Analyses</strong></td>
<td></td>
</tr>
<tr>
<td>SPAdes</td>
<td><a href="http://cab.spbu.ru/software/spades">http://cab.spbu.ru/software/spades</a></td>
</tr>
<tr>
<td>MetaPhlAn2</td>
<td><a href="http://huttenhower.sph.harvard.edu/metaphlan2">http://huttenhower.sph.harvard.edu/metaphlan2</a></td>
</tr>
<tr>
<td>FastQC</td>
<td><a href="https://www.bioinformatics.babraham.ac.uk/projects/fastqc/">https://www.bioinformatics.babraham.ac.uk/projects/fastqc/</a></td>
</tr>
<tr>
<td>QUAST</td>
<td><a href="http://quast.sourceforge.net/quast.html">http://quast.sourceforge.net/quast.html</a></td>
</tr>
<tr>
<td>BBMap</td>
<td><a href="https://jgi.doe.gov/data-and-tools/bbtools/">https://jgi.doe.gov/data-and-tools/bbtools/</a></td>
</tr>
<tr>
<td>Bandage</td>
<td><a href="https://rrwick.github.io/Bandage/">https://rrwick.github.io/Bandage/</a></td>
</tr>
<tr>
<td><strong>Strain Comparison</strong></td>
<td></td>
</tr>
<tr>
<td>kSNP3</td>
<td><a href="https://sourceforge.net/projects/ksnp/">https://sourceforge.net/projects/ksnp/</a></td>
</tr>
<tr>
<td>FastTree2</td>
<td><a href="http://www.microbesonline.org/fasttree/">http://www.microbesonline.org/fasttree/</a></td>
</tr>
<tr>
<td>CFSAN SNP Pipeline</td>
<td><a href="https://github.com/CFSAN-Biostatistics/snp-pipeline/">https://github.com/CFSAN-Biostatistics/snp-pipeline/</a></td>
</tr>
<tr>
<td>FigTree</td>
<td><a href="http://tree.bio.ed.ac.uk/software/figtree/">http://tree.bio.ed.ac.uk/software/figtree/</a></td>
</tr>
</tbody>
</table>