On Premise Computational Infrastructure

APHL Annual Meeting
Infrastructure Solutions for Bioinformatics Implementation in a State Laboratory
June 5 2019
Dr. Kelly F. Oakeson
On Premise Computational Infrastructure
UPHL’s Workflow

- Remote access for troubleshooting & software updates
- Raw data upload to Illumina’s cloud platform for analysis

UPHL Network Connection / No domain access / Web access only
UPHL’s Workflow
UPHL’s Workflow

1. Fastq Files
   - Raw fastq file
   - Transferred to our local Linux box using BaseMount

2. Sequence QC
   - SeqyClean
   - High Quality Sequence Data

3. QC Check
   - FastQC
   - Quick QC check

4. De novo Assembly
   - Shovill
   - Draft Genome Assemblies

5. Assembly QC
   - Quast
   - QC check for Assemblies

6. Genome Annotation
   - Prokka
   - Gene Inventories

7. Phylogenetic Analysis
   - Roary, IQ-Tree & R
   - Phylogenetic Relationships

8. Organism ID
   - MASH
   - Verification of expected organism and QC check

9. Salmonella Serotyping
   - SeqSero
   - Run Salmonella isolates only

10. QC Report
    - MultQC
    - Results end up in one unified QC report

11. E. coli Serotyping
    - SerotypeFinder
    - Using Abricate

12. AMR Prediction
    - ResFinder
    - Using Abricate

13. QC Report
    - MultQC
    - Results end up in one unified QC report

14. C. coli O137:H7 x wr. Sakai
UPHL’s Workflow

A modular tool to aggregate results from bioinformatics analyses across many samples into a single report.


General Statistics

<table>
<thead>
<tr>
<th>Sample Name</th>
<th>N50 (Kbp)</th>
<th>Length (Mbp)</th>
<th>Organism</th>
<th>Contigs</th>
<th>CDS</th>
<th>% Dups</th>
<th>% GC</th>
<th>Length</th>
<th>M Seqs</th>
</tr>
</thead>
<tbody>
<tr>
<td>1279528-UT-M03999-190125</td>
<td></td>
<td></td>
<td>Enterobacter sp.</td>
<td>100.0</td>
<td>4774</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1279528-UT-M03999-190125_S2_L001_R1_001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1279528-UT-M03999-190125_S2_L001_R2_001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1279528-UT-M03999-190125_clean_PE1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1279528-UT-M03999-190125_clean_PE2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1279528-UT-M03999-190125_clean_SE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1279528_UT_M03999_190125_contigs</td>
<td>220.0Kbp</td>
<td>5.1Mbp</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1332073-UT-M03999-190125</td>
<td></td>
<td></td>
<td>Enterobacter cloacae</td>
<td>29.0</td>
<td>4441</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1332073-UT-M03999-190125_S3_L001_R1_001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1332073-UT-M03999-190125_S3_L001_R2_001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1332073-UT-M03999-190125_clean_PE1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1332073-UT-M03999-190125_clean_PE2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
UPHL’s Computational Progression

Laptop

Desktop

Linux Workstation
UPHL’s Computational Progression

Laptop

Desktop

Linux Workstation

aws

The Cloud
It’s Not a Server!
Bioinformatic Analyses of Whole-Genome Sequence Data in a Public Health Laboratory

Kelly F. Oakeson, Jennifer Marie Wagner, Michelle Mendenhall, Andreas Rohrwasser, Robyn Atkinson-Dunn

**Figure.** Steps in the bioinformatics pipeline implemented at Utah Public Health Laboratory.
Thank You

Utah Public Health Laboratory

Robyn Atkinson-Dunn

Andy Rohrwasser

Jenni Wagner

Anna Sangster

Erin Young