Application of Next Generation Sequencing to a Raw Milk Associated *Campylobacter* Outbreak

APHL Annual Meeting 2017
Providence, RI
Dr. Kelly F. Oakeson Ph.D.
PFGE to Whole Genome Sequencing

PFGE

WGS

PFGE

WGS
Whole Genome Sequencing

Raw Data
Maximum Likelihood Tree of 993 Concatenated Gene Sequences

- PNUSAC000775
- PNUSAC000374
- PNUSAC000770
- PNUSAC000360
- PNUSAC000443
- PNUSAC000881
- PNUSAC000883
- PNUSAC000598
- PNUSAC000689
- PNUSAC000370

Campylobacter jejuni
subsp. jejuni
NCTC 11168 = ATCC 70081

Molecular Evolutionary Analysis

hqSNPs
UPHL Bioinformatic Workflow

Computational Requirements & Throughput

Bioinformatic Pipeline Steps:

- Read Quality Control
- Reference Strain Determination
- Read Mapping to Reference Strain
- SNP & Indel Determination
- De novo Genome Assembly
- Genome Annotation
- Phylogenetic Tree Construction
- Phylogenetic Analysis

Typical Laptop / Desktop Computer:
- Throughput: 1-10 Samples

High-end Desktop Computer:
- Throughput: 1-10 Samples

High-end Linux Workstation:
- Throughput: 100s of Samples

Oakeson, K. F., et al. (In Press) Implementation of Bioinformatic Analyses of Whole Genome Sequence Data in a Public Health Laboratory: Emerging Infectious Diseases
**Campylobacter jejuni**

- May 2014, three confirmed cases of *C. jejuni* infections
- Identical PFGE patterns
- All patients reported raw milk consumption from dairy “A”
- Additional cases identified during May and June
- Outbreak investigation initiated June 10, 2014
- Total of 99 cases
Campylobacter jejuni

Campylobacter jejuni PFGE

- PFGE performed on 79 isolates
- 61 patient derived isolates
- 18 isolates derived from bulk milk storage tanks
- 76 of 79 isolates have indistinguishable SmaI PFGE patterns
UPHL Bioinformatic Workflow

Computational Requirements & Throughput

Bioinformatic Pipeline Steps:

Read Quality Control
Reference Strain Determination
Read Mapping to Reference Strain
SNP & Indel Determination
De novo Genome Assembly
Genome Annotation
Phylogenetic Tree Construction
Phylogenetic Analysis

Typical Laptop / Desktop Computer:
Throughput: 1-10 Samples

High-end Desktop Computer:
Throughput: 1-10 Samples

High-end Linux Workstation:
Throughput: 100s of Samples

Oakeson, K. F., et al. (In Press) Implementation of Bioinformatic Analyses of Whole Genome Sequence Data in a Public Health Laboratory: Emerging Infectious Diseases
**Salmonella enterica**

- Complex multi-state outbreak
- Associated with rotisserie chicken
- Five distinct PFGE patterns
- 88 isolates in total
- 80 patient derived isolates
- 8 environmental isolates
- Sequence data obtained from SRA
UPHL Bioinformatic Workflow

Computational Requirements & Throughput

Bioinformatic Pipeline Steps:

Read Quality Control  Reference Strain Determination  Read Mapping to Reference Strain  SNP & Indel Determination  De novo Genome Assembly  Genome Annotation  Phylogenetic Tree Construction  Phylogenetic Analysis

Typical Laptop / Desktop Computer: Throughput: 1-10 Samples

High-end Desktop Computer: Throughput: 1-10 Samples

High-end Linux Workstation: Throughput: 100s of Samples

Oakeson, K. F., et al. (In Press) Implementation of Bioinformatic Analyses of Whole Genome Sequence Data in a Public Health Laboratory: Emerging Infectious Diseases
Thank You

Utah Public Health Laboratory

Robyn Atkinson-Dunn
Andy Rohrwasser
Jenni Wagner

University of Utah

Jon Seger
Patrice Showers Corneli