The *Listeria* Initiative in a New Era of Whole Genome Sequencing

Benjamin Silk, PhD, MPH
John Besser, PhD
Rajal Mody, MD, MPH

Division of Foodborne, Waterborne, and Environmental Diseases

InFORM, San Antonio, Texas
November 21, 2013
Outline

1. Overview of *Listeria* Initiative: Ben Silk, CDC
2. Overview of the new WGS Project: John Besser, CDC
3. Examples of how WGS has been of value in the first 3 months of the project: Ben Silk, CDC
4. Discussion of what can be done to make WGS even more effective: Raj Mody, CDC
5. Q&A
# Listeria Infection (Listeriosis) by Risk Group

<table>
<thead>
<tr>
<th>Group</th>
<th>Typical illness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elderly</td>
<td>Bloodstream infection or meningitis</td>
</tr>
<tr>
<td>Persons with immunocompromising conditions</td>
<td>Bloodstream infection or meningitis</td>
</tr>
</tbody>
</table>
| Pregnant women and their newborn infants   | **Pregnant women:** Asymptomatic infection or “flu-like” illness with fever, then sometimes miscarriage or stillbirth  
                                          | **Infants:** Bloodstream infection or meningitis                                |
| Healthy persons                            | Diarrhea with fever                                                              |
## Listeria Infection (Listeriosis) by Risk Group

<table>
<thead>
<tr>
<th>Group</th>
<th>Typical illness</th>
<th>Surveillance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elderly</td>
<td>Bloodstream infection or meningitis</td>
<td>Yes</td>
</tr>
<tr>
<td>Persons with immunocompromising conditions</td>
<td>Bloodstream infection or meningitis</td>
<td>Yes</td>
</tr>
<tr>
<td>Pregnant women and their newborn infants</td>
<td>Pregnant women: Asymptomatic infection or “flu-like” illness with fever, then sometimes miscarriage or stillbirth</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Infants: Bloodstream infection or meningitis</td>
<td></td>
</tr>
<tr>
<td>Healthy persons</td>
<td>Diarrhea with fever</td>
<td>No</td>
</tr>
</tbody>
</table>
Surveillance Case Definition for Invasive Listeriosis

Isolation of *Listeria monocytogenes* from

- a normally sterile site
  (usually blood or cerebrospinal fluid)

or

- products of conception
PulseNet Data Analysis Involves Searching for Clusters

- PulseNet teams at CDC and in states search for similar PFGE patterns.
- When a cluster is identified, they report it to epidemiologists.

Each row is PFGE pattern from one isolate. Box shows cluster of 4 rows with same pattern.
Multistate Outbreak of Listeriosis from Cantaloupe, July – October, 2011
Rapid Investigation of Cantaloupe Outbreak

- Largest known U.S. outbreak of listeriosis (147 cases)
  - deadliest U.S. foodborne outbreak in >90 years (33 deaths)
- Rapid, coordinated identification of food source using PulseNet and the *Listeria* Initiative questionnaires prevented an estimated
  - prevented ~36 illnesses and ~7 deaths
Is it better to be lucky or good?

(Actually both are best.)
Status of Listeriosis Control in 2004

- **Incidence**
  - Marked decline in 1990s, stable from 2001 through 2003

- **Outbreaks**
  - Marked increase in number of outbreaks detected
  - Marked decrease in size of outbreaks
  - Detecting clusters of patients with similar PFGE patterns for which could not find source
Challenges of Listeriosis Outbreaks

- **Cases are often geographically dispersed**
  - so detecting outbreaks is difficult

- **Getting food histories from ill persons is difficult**
  - incubation period may be weeks, so memories have faded
  - elderly and ill patients may have poor food recall
  - ill persons may die before interview

- **Controls are needed, but are difficult to find**
  - need to ask controls who are similar to cases about foods consumed

- **Many clusters of cases are small**
  - which makes determining the source difficult
New Approach: *Listeria* Initiative, 2004

- **State and local health depts**
  - Promptly interview all patients (whether or not part of an outbreak) with a standard form that asks about foods
  - Rapidly perform PFGE for all *L. monocytogenes* isolates upload patterns to PulseNet

- **CDC**
  - Scans for clusters of similar isolates
  - When cluster detected, epidemiologists analyze questionnaires
### III. Food Consumption History

**Instructions for Interviewer:** Please read all options to case-patient in each category. For the names of purchase sites, it is preferable to use codes: A1 for first grocery store, A2 for second grocery store, A3 for third grocery store, A4 for fourth grocery store, A5 for fifth grocery store. A DELI COUNTER serves portions or helpings of salads, cheeses, and meats sliced ON-SITE. A specified counter within a grocery store, food market, or delicatessen. Foods sliced and packaged AT THE FACTORY and sold as pre-packaged containers in self-serve refrigerated display cases are NOT considered to be from a deli counter.

**Instructions to Read to Case-Patient (or Surrogate):**

Now I’d like to ask you about the foods that you ate between _/__/__ (date 4 weeks before) through _/__/__ (specimen collection/delivery date). For each food, please give me your best guess as to whether you ATE the food, you’re not sure but you LIKELY ATE the food, you’re not sure but you LIKELY DID NOT EAT the food, or you know you DID NOT EAT the food.

#### MEATS

In the 4 week period, did you eat any of the following COLD CUT, DELI MEAT, OR LUNCHEON MEAT items?

<table>
<thead>
<tr>
<th>Item</th>
<th>Ate</th>
<th>Likely Ate</th>
<th>Did Not</th>
<th>Did Not Eat</th>
<th>How Often?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ham</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>~1–2 x/month</td>
</tr>
<tr>
<td>Bologna</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>~1–2 x/month</td>
</tr>
<tr>
<td>Turkey Breast</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>~1–2 x/month</td>
</tr>
<tr>
<td>Other turkey deli meat (e.g., turkey ham)</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>~1–2 x/month</td>
</tr>
</tbody>
</table>
**Listeria Initiative:**
**Case-case Analysis**

- **What:** Statistical comparison of food exposure frequencies
- **How:** Use of *Listeria* Initiative database
  - Outbreak cases: patients with outbreak-related listeriosis
  - Sporadic cases: patients with listeriosis unrelated to any cluster
- **Why:** Generate hypothesis about outbreak source based on epidemiological measure of association (e.g., odds ratios and confidence intervals)
**Listeria Initiative (LI): Case-Case Analyses**

<table>
<thead>
<tr>
<th>Date</th>
<th>Food</th>
<th>Outbreak Cases</th>
<th>LI Sporadic Cases*</th>
<th>OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>September 9</td>
<td>Ham</td>
<td>7/11 (64)</td>
<td>360/774 (47)</td>
<td>2.0 (0.5–9.4)</td>
</tr>
<tr>
<td></td>
<td>Watermelon</td>
<td>5/10 (50)</td>
<td>49/82 (60)</td>
<td>0.7 (0.1–3.2)</td>
</tr>
<tr>
<td></td>
<td>Cantaloupe</td>
<td>11/11 (100)</td>
<td>54/85 (64)</td>
<td>8.5 (1.3–∞)</td>
</tr>
<tr>
<td>September 12</td>
<td>Cantaloupe</td>
<td>13/13 (100)</td>
<td>54/85 (64)</td>
<td>10.1 (1.6–∞)</td>
</tr>
<tr>
<td>September 14</td>
<td>Cantaloupe</td>
<td>19/19 (100)</td>
<td>54/85 (64)</td>
<td>14.9 (2.4–∞)</td>
</tr>
</tbody>
</table>

*In sporadic cases of listeriosis, 2004–2010, limited to adults ≥60 years of age and isolation dates in August (melons only)
1. Overview of *Listeria* Initiative: Ben Silk, CDC

2. Overview of the new WGS Project: John Besser, CDC

3. Examples of how WGS has been of value in the first 3 months of the project: Ben Silk, CDC

4. Discussion of what can be done to make WGS even more effective: Raj Mody, CDC

5. Q&A
Frederick Sanger, 95, Two-Time Winner of Nobel and Pioneer in Genetics, Dies

By DENISE GELLENE
Published: November 20, 2013

Frederick Sanger, a British biochemist whose discoveries about the chemistry of life led to the decoding of the human genome and to the development of new drugs like human growth hormone, earning him two Nobel Prizes, a distinction held by only three other scientists, died on Tuesday in Cambridge, England. He was 95.

His death was confirmed by Adrian Penrose, communications manager at the Medical Research Council in Cambridge. Dr. Sanger lived in a nearby village called Swaffham Bulbeck.

Dr. Sanger won his first Nobel Prize, in chemistry, in 1958 for showing how amino acids link together to form insulin. The discovery
U.S. Nationwide Real-Time Listeriosis Surveillance using Whole Genome Sequencing: Proof-of-Concept Study

U.S. collaborators
- State health agencies
- NCBI
- FDA
- USDA
- CDC

Collaborating countries
- France
- UK
- Denmark
- Australia
- Canada
Core Question

Will the use of whole genome sequencing as part of real-time surveillance lead to a public health benefit?
Goals

- “Real-time” whole genome sequencing (WGS) and analysis of isolates from all U.S. cases of *Listeria monocytogenes* and food/environmental isolates
- Assess technological and interpretive challenges
- Develop a path to a WGS-based PulseNet infrastructure
- Improve epidemiological follow-up of cases and clusters
- Develop new epidemiological tools
Budget

$139,000 (one year)

Timeline

- **July, 2013**: Started planning,
- **September 1, 2013**: official start date
Technical Questions

- What clustering markers have the strongest epidemiological concordance? (e.g. k-mer SNPS; hq SNPS; wg MLST)
- What sort of infrastructure is needed? (sequencing capacity, platform compatibility, quality control, standardization, bioinformatic pipelines, bandwidth, storage, etc)
Interpretive Challenges

- Continuous versus categorical data
  - Strain naming
  - Defining “closeness”
  - Defining clusters (in a real-time context)

- Interpretive challenges
  - How much epidemiological inference can we make from (agent) phylogenetic information?
Parallel Workflow

State and Local Public health laboratories

State Departments of Agriculture

PFGE Case Definition

WGS Case Definition

?
Real-Time WGS Surveillance

- Shipment of isolate to CDC: 4 days
- WGS and analysis: 1 week
Listeria Project Data Flow

- CDC Labs
  - isolates
  - sequences
  - Patient-level metadata (PulseNet database)
  - Reduced metadata

- State Health Laboratories
  - isolates
  - sequences

- FDA Genome Trackr Labs
  - isolates
  - sequences

- NCBI
  - sequences
  - cluster analyses

- Public Release
<table>
<thead>
<tr>
<th>Strains Sequenced; 11/15/2013</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number</td>
</tr>
<tr>
<td>Real-time WGS program</td>
</tr>
<tr>
<td>100K Pathogen study</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>
What Do We Hope To See: Outbreaks Solved

- Faster, with fewer cases (e.g. 2-3)
- Strong hypotheses from food/environmental findings
- Better determination of scope
- WGS findings lead to alternate hypotheses
Interpretive Challenges

- Continuous versus categorical data
  - Strain naming
  - Defining “closeness”
  - Defining clusters (in a real-time context)

- Interpretive challenges
  - How much epidemiological inference can we make from (agent) phylogenetic information?
PFGE: Categorical Data (strains indistinguishable or not)
Whole Genome Sequences: Continuous Data

Listeria Cluster September 19, 2013

Whole genome sequence differences
## Classification of Cases (by Agent)

<table>
<thead>
<tr>
<th></th>
<th>Case</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exposed</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>Not Exposed</td>
<td>C</td>
<td>D</td>
</tr>
</tbody>
</table>

\[ \text{OR} = \frac{A \times D}{C \times B} \]

**PFGE Case Definition**
Whole Genome Sequence Case Definition (exact match)
Proper Classification of Cases Improves Measures of Association

Cluster definitions
Whole Genome Sequencing of *Listeria monocytogenes* during the Crave Brothers Cheese outbreak

High confidence core SNP
Whole Genome Sequencing of *Listeria monocytogenes* during the Crave Brothers Cheese outbreak

High confidence core SNP

excluded case
retrospective controls

Epi-related
PFGE match- from outbreak period
excluded case

retrospective controls

Most inclusive case definition

High confidence core SNP

Epi-related

PFGE match - from outbreak period

Most specific case definition
Options for Defining Clusters

- Use multiple case definitions
- Develop / test measures of clade significance
Sensitivity

Specificity

Inclusive Case Definition

Specific Case Definition
Key Steps in an Outbreak Investigation

- **Detect Cluster**
- **Develop Hypothesis**
- **Test Hypothesis**
- **Determine Scope**

**Clinical Diagnoses**

**Use more specific case definition**

**Food Monitoring**

**Targeted Food/Env Testing**

**Trace-backs**

**Use more inclusive case definition to find additional cases**
Options for Defining Clusters

- Use multiple case definitions
- Develop / test measures of clade significance
Statistical Options for Cluster Recognition?

Example: Fixation Index (measures strain consolidation above background rate)

The fixation index ($F_{ST}$) is a measure of population differentiation due to genetic structure. It is frequently estimated from genetic polymorphism data, such as single-nucleotide polymorphisms (SNP) or microsatellites. Developed as a special case of Wright's $F$-statistics, it is one of the most commonly used statistics in population genetics.

$$F_{ST} = \frac{\sigma_S^2}{\sigma_T^2} = \frac{\sigma_S^2}{\bar{p}(1 - \bar{p})}$$

From Hannes Poussele; Applied Maths
Continuous versus categorical data
- Strain naming
- Defining “closeness”
- Defining clusters (in a real-time context)

Interpretive challenges
- How much epidemiological inference can we make from (agent) phylogenetic information?
Central Assumption

Phylogenetic similarity (and distance) between disease-causing organisms can be used to infer epidemiological associations (between people, animals, food, and the environment)
Complex Ecology of Outbreaks
Meaning of a “close match”: High likelihood of some connection between isolates (from cases, food, environment)
Outline

1. Overview of *Listeria* Initiative: Ben Silk, CDC

2. Overview of the new WGS Project: John Besser, CDC

3. Examples of how WGS has been of value in the first 3 months of the project: Ben Silk, CDC

4. Discussion of what can be done to make WGS even more effective: Raj Mody, CDC

5. Q&A
**Listeria Outbreaks, 1983-2012**

- **Before PulseNet**
  - 1983-1997 (15 years)
  - 5 outbreaks
  - Median 69 cases/outbreak

- **PulseNet’s first years**
  - 1998-2003 (6 years)
  - 14 outbreaks
  - Median 11 cases/outbreak

- **Listeria Initiative & PulseNet**
  - 2004-2012 (9 years)
  - 28 outbreaks
  - Median 5.5 cases/outbreak
Listeria Outbreaks and Incidence, 1983-2012

Before PulseNet
1983-1997 (15 years)
5 outbreaks
Median 69 cases/outbreak

PulseNet's first years
1998-2003 (6 years)
14 outbreaks
Median 11 cases/outbreak

Listeria Initiative & PulseNet
2004-2012 (9 years)
28 outbreaks
Median 5.5 cases/outbreak
You are here
An Epidemiologist's Wish List for WGS

- Epidemiological concordance
Whole genome sequencing data on isolates with 2-enzyme PFGE match

- Green: Isolates from cheese factory, 2010-2011
- Red: Isolates from patients in current outbreak
- Blue: Isolates from patients in past, exposure is not known
Whole Genome Sequencing of *Listeria monocytogenes* during the Crave Brothers Cheese outbreak

High confidence core SNP

NY – no soft cheese exposure

excluded CO case
retrospective controls

Epi-related
PFGE match- from outbreak period

TX - no exposure info
Cluster in Midwest State A

<table>
<thead>
<tr>
<th>PFGE-ApaI+PFGE-Ascl</th>
<th>PFGE-Ascl</th>
<th>PFGE-Apal</th>
<th>Ascl_Apal</th>
<th>SNP</th>
</tr>
</thead>
<tbody>
<tr>
<td>State A</td>
<td>GX6A16.0005</td>
<td>GX6A12.0088</td>
<td>Ref</td>
<td></td>
</tr>
<tr>
<td>State B</td>
<td>GX6A16.0164</td>
<td>GX6A12.0088</td>
<td>145</td>
<td></td>
</tr>
<tr>
<td>State A</td>
<td>GX6A16.0164</td>
<td>GX6A12.0197</td>
<td>62</td>
<td></td>
</tr>
<tr>
<td>State A</td>
<td>GX6A16.0005</td>
<td>GX6A12.0088</td>
<td>31</td>
<td></td>
</tr>
<tr>
<td>State A</td>
<td>GX6A16.0005</td>
<td>GX6A12.0088</td>
<td>53</td>
<td></td>
</tr>
<tr>
<td>State A</td>
<td>GX6A16.0005</td>
<td>GX6A12.0088</td>
<td>46</td>
<td></td>
</tr>
<tr>
<td>State C</td>
<td>GX6A16.1244</td>
<td>GX6A12.0021</td>
<td>229</td>
<td></td>
</tr>
<tr>
<td>State B</td>
<td>GX6A16.1244</td>
<td>GX6A12.0021</td>
<td>182</td>
<td></td>
</tr>
</tbody>
</table>

*PA*=pregnancy-associated

- **PA**, Mexican-style cheese (State A)
- **PA**, Mexican-style cheese (State A)
- **PA**, No exposure info (State A)
- **PA**, Mexican-style cheese (State A)
- **PA**, Mexican-style cheese (State A)
- **PA**, Mexican-style cheese (State A)
- **PA**, Mexican-style cheese (State A)
- **PA**, Mexican-style cheese (State A)
- **PA**, Mexican-style cheese (State A)
- **PA**, Mexican-style cheese (State A)
- **PA**, Mexican-style cheese (State A)
- **PA**, Mexican-style cheese (State A)
- **PA**, Mexican-style cheese (State A)
- **PA**, Mexican-style cheese (State A)
- **PA**, Mexican-style cheese (State A)
An Epidemiologist's Wish List for WGS

- Epidemiological concordance
- Subtyping value added (beyond PFGE)
Number of Patients by Date of Illness Onset and PFGE Pattern — United States, 2008–2011

Date of Illness Onset

Number of Isolates

- Pattern 4
- Pattern 3
- Pattern 2
- Pattern 1

2008 | 2009 | 2010 | 2011
Company A

<table>
<thead>
<tr>
<th>Date</th>
<th>State</th>
<th>Pattern</th>
<th>GX6A 16.0018</th>
<th>GX6A 12.1394</th>
</tr>
</thead>
<tbody>
<tr>
<td>7/13/2013</td>
<td>State B</td>
<td>1/2a</td>
<td>GX6A 16.0018</td>
<td>GX6A 12.0022</td>
</tr>
<tr>
<td>8/12/2013</td>
<td>State C</td>
<td>1/2a</td>
<td>GX6A 16.0018</td>
<td>GX6A 12.0222</td>
</tr>
</tbody>
</table>

**Ascl_Apal**

**PFGE-Ascl**

**PFGE-Apal**

GX6A16.0018  GX6A12.0022
GX6A16.0018  GX6A12.1394
<table>
<thead>
<tr>
<th>Clinical isolate for patient with exposure to Company A product</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinical isolate for patient with exposure to Company B product</td>
</tr>
</tbody>
</table>

**Node A**

<table>
<thead>
<tr>
<th>Clinical isolate for patient with exposure to Company A product</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinical isolate for patient with exposure to Company B product</td>
</tr>
</tbody>
</table>

**Node B**

<table>
<thead>
<tr>
<th>Clinical isolate for patient with exposure to Company A product</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinical isolate for patient with exposure to Company B product</td>
</tr>
</tbody>
</table>
Large, Multistate Cluster
```
<table>
<thead>
<tr>
<th>Date</th>
<th>State</th>
<th>Location</th>
<th>Sample Type</th>
<th>Sample Code</th>
<th>Organism</th>
<th>Pathogen Code</th>
<th>Matrix A</th>
<th>Matrix B</th>
<th>Id</th>
</tr>
</thead>
<tbody>
<tr>
<td>8/9/2013</td>
<td>OH</td>
<td></td>
<td>blood</td>
<td>2013L-5439</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9/13/2013</td>
<td>MI</td>
<td></td>
<td>blood</td>
<td>2013L-5527</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6/13/2013</td>
<td>MD</td>
<td></td>
<td>blood</td>
<td>2013L-5473</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8/15/2013</td>
<td>NY</td>
<td></td>
<td>blood</td>
<td>2013L-5521</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6/16/2013</td>
<td>MI</td>
<td></td>
<td>blood</td>
<td>2013L-5548</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9/25/2013</td>
<td>OH</td>
<td></td>
<td>blood</td>
<td>2013L-5560</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5/26/2013</td>
<td>NY</td>
<td></td>
<td>blood</td>
<td>2013L-5345</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9/10/2013</td>
<td>MI</td>
<td></td>
<td>blood</td>
<td>2013L-5540</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9/12/2013</td>
<td>CA</td>
<td></td>
<td>blood</td>
<td>2013L-5545</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9/19/2013</td>
<td>WA</td>
<td></td>
<td>blood</td>
<td>2013L-5556</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5/17/2013</td>
<td>HI</td>
<td></td>
<td>blood</td>
<td>2013L-5308</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6/25/2013</td>
<td>SC</td>
<td></td>
<td>blood</td>
<td>2013L-5314</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10/11/2013</td>
<td>WA</td>
<td></td>
<td>blood</td>
<td>2013L-5587</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6/13/2013</td>
<td>SC</td>
<td></td>
<td>blood</td>
<td>2013L-5333</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Subcluster**
```
An Epidemiologist's Wish List for WGS

✓ Epidemiological concordance
✓ Subtyping value added (beyond PFGE)

☐ Ease of interpretation?
  ☐ How close is close (phylogenetically)?
  ☐ What node on a tree is most meaningful?
MAXIMIZING THE VALUE OF WGS TO IMPROVE IDENTIFICATION OF FOOD SOURCES OF LISTERIA OUTBREAKS
**Ways to Improve WGS-Epi Data Integration to Better Detect and Solve Listeria Outbreaks**

1. **Now:** CDC following-up with states
   - To assure, on a regular basis, that Listeria Initiative (LI) data and isolate data are linked as best as possible

2. **Proposed:** collect more exposure information
   - Produce, specialty deli items, foods eaten at restaurants, recalled products, others?
   - Greater use of open-ended interviewing

3. **Proposed:** build better IT systems
   - Web-based *Listeria* Initiative data entry
   - Integrated Epi-WGS database that allows for point-and-click assessment of exposures at different nodes in a phylogenetic tree
Denominator is *Listeria* cases reported to National Notifiable Disease Surveillance System
Isolate and Questionnaire Submission, 2004-2012

Denominator is *Listeria* cases reported to National Notifiable Disease Surveillance System.
Isolate and Questionnaire Submission, 2004-2012

Denominator is *Listeria* cases reported to National Notifiable Disease Surveillance System
Novel Data Closeout in 2013: Improved Linkage

Asked states to

- Review existing linkages between LI data and isolates for any errors
- If possible, link any un-linked LI forms and isolates
- If any *Listeria* Initiative forms or isolates remain missing or unlinked, explain why

Findings

- 45 jurisdictions respond
- Increased cases with linked LI data & isolate by 7%
- Isolates are often not sent to state laboratories
- Patients are often unable to be interviewed
Ways to Improve WGS-Epi Data Integration to Better Detect and Solve Listeria Outbreaks

1. **Now: CDC following-up with states**
   - To assure, on a regular basis, that Listeria Initiative (LI) data and isolate data are linked as best as possible

2. **Proposed: collect more exposure information**
   - Produce, specialty deli items, foods eaten at restaurants, recalled products, others?
   - Greater use of open-ended interviewing

3. **Proposed: build better IT systems**
   - Web-based *Listeria* Initiative data entry
   - Integrated Epi-WGS database that allows for point-and-click assessment of exposures at different nodes in a phylogenetic tree
Questions?

For thoughts on strengthening the Listeria Initiative, please contact Raj Mody, rmody@cdc.gov

The findings and conclusions in this presentation are those of the author and do not necessarily represent the views of the Centers for Disease Control and Prevention.
Causal Pathways in Outbreaks: Polyclonal Outbreaks

No molecular linkage (when an epidemiological linkage exists)
Molecular Markers for Outbreak Detection: Interpretation Issues

- Does a WGS “mis-match” rule out?
- Does a WGS match prove association?
Causal Pathways in Outbreaks: Direct

Molecular linkage = likely epidemiological linkage
Causal Pathways in Outbreaks: Indirect (or cross-contamination)

Molecular linkage...possibly misleading or not useful
Causal Pathways in Outbreaks: Multiple Contamination Events; One leads to Outbreak

Molecular linkage...possibly incorrect conclusion

NO ISOLATE
Causal Pathways in Outbreaks: Secondary and Tertiary Transmission

Molecular linkage....incorrect conclusion, or no conclusion
Meaning of a WGS “Match”

High likelihood of *some* connection between isolates (from cases, food, environment)*

* Everything else requires epidemiology
Getting the Word Out

When Food Bites Back
Protecting people from deadly Listeria food poisoning

48M
About 48 million people in the US (1 in 6 people) get sick from contaminated food each year.

3rd
Listeria is the 3rd leading cause of death from food poisoning.

90%+
At least 90% of people who get Listeria infections are pregnant women and their newborns, people 65 or older, or people with weakened immune systems.

Many germs can be spread through food. Some, like Listeria, can be deadly. Listeria strikes hard at pregnant women and their newborns, older adults, and people with weakened immune systems. Listeria can cause miscarriage and meningitis. Most people found to have Listeria infection require hospital care and about 1 in 5 people with the infection die. Outbreak investigations tell us what foods make people sick and what needs to change to identify steps to make food safer and save lives. We have made some progress against Listeria, which is the third leading cause of death from food poisoning. However, we can do more to protect people at higher risk for food poisoning and make food safer for everyone.

If you, or someone you make food for, are pregnant, 65 or older, or have a weakened immune system, you must be especially careful when selecting, preparing, and storing foods.

- Know your risk of food poisoning.
- Select, prepare, and store food safely.
- Follow the safe food guidelines – Clean, Separate, Cook, Chill – at food.safety.gov.

Learn more about how to prevent food poisoning and outbreaks.

Who has a higher risk of getting Listeria food poisoning?

Lessons from Listeria outbreaks: Food poisoning can happen to anyone. Each year, about 48 million people in the US (1 in 6) get sick from eating contaminated food. It can be especially dangerous for pregnant women and their newborns; older adults; and people with immune systems weakened by cancer, cancer treatments, or other serious conditions (like diabetes, kidney failure, liver disease, and HIV/AIDS). Listeria is a prime example of how germs that contaminate food can cause sickness and death in these groups.

Pregnant women, fetuses, and newborn infants

- Listeria can pass from pregnant women to their fetuses and newborns.
- It can cause miscarriages, stillbirths, and newborn deaths.

People with weakened immune systems

- Listeria can spread through the bloodstream to cause meningitis, and often kills. The weaker your immune system, the greater the risk.

Contaminated celery
LISTERIA OUTBREAK: Pre-cut celery in chicken salad served at hospitals sickened 10 people who had other serious health problems. Five of them died as a result.

Adults 65 or older

- Listeria can spread through the bloodstream to cause meningitis, and often kills. The older you are, the greater the risk.

Tainted cantaloupes
LISTERIA OUTBREAK: Contaminated whole cantaloupes sickened 147 people in 29 states and caused one of the deadliest foodborne outbreaks in the US. There were 33 deaths, mostly in adults over 65, reported during the outbreak.

What foods are risky?

When it comes to Listeria, some foods are more risky than others. Most of the other foods where Listeria is known to hide:

- Raw Sprouts
- Raw Milk
- Soft Cheeses
- Deli Meats and Hot Dogs (cold, not heated)
- Smoked Seafood

SOURCE: CDC, 2013
Listeriosis is much less common, but much more deadly, than other major bacterial pathogens

<table>
<thead>
<tr>
<th>Pathogen</th>
<th>Illnesses</th>
<th>Deaths</th>
<th>Case-fatality rate</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Listeria</em></td>
<td>1,600</td>
<td>260</td>
<td>16%</td>
</tr>
<tr>
<td><em>Campylobacter</em></td>
<td>1,300,000</td>
<td>120</td>
<td>0.1%</td>
</tr>
<tr>
<td><em>Salmonella</em></td>
<td>1,230,000</td>
<td>450</td>
<td>0.5%</td>
</tr>
</tbody>
</table>
# Incidence of Listeriosis by Risk Group, United States, 2004-2009

(overall 2.7 per million persons)

<table>
<thead>
<tr>
<th>Risk group</th>
<th>Incidence (per million)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Persons ≥65 years old</td>
<td>12</td>
</tr>
<tr>
<td>Persons &lt;65 years old (most immunocompromised; not pregnant)</td>
<td>1</td>
</tr>
<tr>
<td>Pregnant women and their newborns, non-Hispanic</td>
<td>23</td>
</tr>
<tr>
<td>Pregnant women and their newborns, Hispanic</td>
<td>89</td>
</tr>
</tbody>
</table>

*Silk BJ et al, CID 2012*
## Incidence of Listeriosis by Risk Group, United States, 2004-2009 (overall 2.7 per million persons)

<table>
<thead>
<tr>
<th>Risk group</th>
<th>Incidence (per million)</th>
<th>Proportion of Cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Persons $\geq 65$ years old</td>
<td>12</td>
<td>53%</td>
</tr>
<tr>
<td>Persons $&lt;65$ years old (most immunocompromised; not pregnant)</td>
<td>1</td>
<td>31%</td>
</tr>
<tr>
<td>Pregnant women and their newborns, non-Hispanic</td>
<td>23</td>
<td>9%</td>
</tr>
<tr>
<td>Pregnant women and their newborns, Hispanic</td>
<td>89</td>
<td>6%</td>
</tr>
</tbody>
</table>
Captures outbreak data on agents, foods, and settings responsible for illness
### Food Vehicles for *Listeria* Outbreaks, 1998-2003

(11 outbreaks with known vehicle)

<table>
<thead>
<tr>
<th>Food Vehicle</th>
<th># of Outbreaks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delicatessen meat (usually turkey)</td>
<td>5</td>
</tr>
<tr>
<td>Hot dogs</td>
<td>3</td>
</tr>
<tr>
<td>Cheese</td>
<td>2</td>
</tr>
<tr>
<td>Other</td>
<td>1</td>
</tr>
</tbody>
</table>

Data from Foodborne Disease Outbreak Surveillance System
Food Vehicles for *Listeria* Outbreaks, 2004-2012 (9 years, 20 outbreaks with known vehicle)

<table>
<thead>
<tr>
<th>Food Vehicle</th>
<th># of Outbreaks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cheese</td>
<td>10</td>
</tr>
<tr>
<td>Produce (sprouts, celery, cantaloupe)</td>
<td>3</td>
</tr>
<tr>
<td>Deli meat</td>
<td>2</td>
</tr>
<tr>
<td>Other*</td>
<td>5</td>
</tr>
</tbody>
</table>

*Data from Foodborne Disease Outbreak Surveillance System; *milk, chicken, sushi rolls, taco salad, tuna salad*
Listeria Wake-up Call, 1985

- **Outbreak in California**
  - 142 ill
    - 93 pregnancy-related
    - 40 not pregnancy-related
  - 28 deaths, 20 miscarriages
  - traced to Mexican-style soft cheese (*queso fresco*)

- **Began surveillance for individual cases**
Epidemiologic Notes and Reports Listeriosis Associated with Consumption of Turkey Franks

Microbiologic implication of meat as a source of clinical listeriosis has not previously been documented. In December 1988, a woman was found to have listeria monocytogenes (LM) in her throat. LM was isolated from an open package of Plantation Brand turkey franks from the patient’s refrigerator. LM was also isolated from two unopened packages of Plantation Brand turkey franks from a local store. Cultures, however, showed negative results for LM. LM isolates from the patient and from the opened and unopened packages had the same electrophoretic enzyme type. On April 14, 1989, the company voluntarily recalled the franks, and the U.S. Department of Agriculture confirmed that multiple products from the same producer were involved. The company confirmed that LM of the same serotype and enzyme type, and had the same origin.

Reference


Disclaimer

All MMWR HTML documents published before January 1993 are electronic conversions from ASCII text into HTML. This conversion may not be perfect. Users should not rely on this HTML document, but are referred to the original MMWR paper copy for the official text, figures, and tables. Each week, Public Health Reports prints the complete paper in the "Excerpts from the Weekly Surveillance Summaries" section. For more information, please visit www.cdc.gov/mmwr.

Questions or messages regarding errors in formatting should be addressed to mmwrq@cdc.gov.
Hot Dog Outbreak, 1998

- 108 ill
  - 13 pregnancy-related
  - 95 not pregnancy-related
- 14 deaths, 4 miscarriages
- Patients in 24 states
  - isolates linked by PFGE pattern
- Hot dogs contaminated at one processing plant after cooking
- Outcome: industry efforts to make hot dogs safer

Mead et al, Epidemiol Infect 2006;134:744-751
Hot Dog Outbreaks:
A Regulatory and Industry Success

- Hot dogs probably an important source of illness before the 1998 outbreak
  - PulseNet made hidden multistate outbreaks detectable
- Following the 1998 outbreak, regulatory agencies and industry improved safety

No hot dog outbreak detected since 1999

Delicatessen Turkey Meat Outbreak, 2002

- 54 ill
  - 12 pregnancy-related
  - 42 not pregnancy-related
- 8 deaths, 3 miscarriages
- Patients in 9 states
  - isolates linked by PFGE pattern

Deli Meat Outbreaks: A Regulatory and Industry Success

- Contamination occurs in the processing plant, after meat cooked
- After the 2002 outbreak
  - US Department of Agriculture tightened regulations
  - Industry added growth inhibitors

Only 2 (small) deli meat outbreaks since 2002

**Listeria Incidence and Percentage of Ready-to-Eat Meats Positive for *L. monocytogenes*, 1983-2012**

- **% Processed Meat**
- **Incidence (per million pop)**

**Before PulseNet**
- PulseNet’s first years
- *Listeria* Initiative & PulseNet
**Listeria Incidence and Percentage of Ready-to-Eat Meats Positive for L. monocytogenes, 1983-2012**

- **% Processed Meat**
- **Incidence (per million pop)**

- **Before PulseNet**
- **PulseNet’s first years**
- **Listeria Initiative & PulseNet**

The graph shows a decrease in the incidence of L. monocytogenes in processed meats from 1983 to 2012, with notable improvements during the PulseNet’s first years and the Listeria Initiative & PulseNet periods.
Listeriosis Outbreak from Cantaloupe, July – October, 2011

- Detected by Colorado state health department
  - notified CDC September 2
- Rapid nationwide investigation using *Listeria* Initiative
  - cantaloupe from a single farm recalled September 14
- 147 ill
  - 7 pregnancy-related
  - 139 not pregnancy-related
- 33 deaths, 1 miscarriage
- Patients in 28 states
  - isolates linked by PFGE pattern
    - outbreak caused by strains with 5 patterns
Environmental Assessment at Farm

- Processing facility most likely source of contamination
- *Listeria* outbreak strains were found on cantaloupes and surfaces at the facility
- Deficiencies
  - cantaloupes not cooled after harvest
  - equipment designed and used improperly
  - sanitation inadequate
Characteristics of Cheese Outbreaks, United States, 2004-2012 (n = 10 outbreaks)

- Cheese type
  - 60% Mexican-style, soft
  - 40% other, usually soft, often specialty

- 80% made with pasteurized milk
  - suggests environmental contamination during or after cheese-making
Outbreak from Imported Cheese that Cross-Contaminated Other Cheeses, 2012

- **Index patient**
  - 68 year-old person in Pennsylvania with *Listeria* meningitis
  - consumed 2 cheeses purchased from a store (“Grocery Store A”)

- **Health department cultured the 2 cheeses in his refrigerator**
  - both had the patient’s strain of *Listeria*
Investigation of Index Patient, July 2012

- Health department went to Grocery Store A
  - isolated outbreak strain from a wedge of l’Édel de Cléron cheese that had been cut from a wheel of cheese and repackaged
  - did not isolate *Listeria* from Blue cheese

- CDC checked PulseNet database
  - a few isolates with same rare PFGE pattern from patients in other states in 2012
Investigation of Index Patient, July 2012

- Health department went to Grocery Store A
  - isolated outbreak strain from a wedge of l’Édel de Cléron cheese that had been cut from a wheel of cheese and repackaged
  - did not isolate *Listeria* from Blue cheese

- CDC checked PulseNet database
  - a few isolates with same rare PFGE pattern from patients in other states in 2012

Was l’Édel de Cléron cheese the source of their illnesses?
Initial Analysis of *Listeria* Initiative Data

- Compared food exposures of
  - 10 patients with rare PFGE pattern
  - 204 other patients with listeriosis during 2008-2012

- Result
  - 9 (90%) of patients whose isolates had the rare pattern, but only 37% of control patients had recently consumed soft cheese (odds ratio 17, p-value <0.01)
Epidemiologic Investigation

- Designed a supplemental questionnaire to collect more information about cheese
  - types and brands, including l’Édel de Cléron
  - packaging information
  - where purchased
Initial Analysis of Supplemental Questionnaire

- Re-interviewed 9 patients with supplemental questionnaire
- Patients consumed a variety of cheeses
  - none reported consuming l’Édel de Cléron!
  - no common brand of cheese named
- 6 had information about packaging
  - all 6 consumed cheese from a wedge that had been cut from a wheel and repackaged
Meanwhile, in California....

- On a routine inspection at a distributor (“Distributor Z”), regulators cultured samples of cheese
  - isolated outbreak strain from 2 other types of cheese
- Distributor Z did not supply cheese to Store A (index patient’s store)
  - suggested that the cheeses at Distributor Z and Store A were contaminated by another cheese (“Cheese X”) when they were cut and repackaged
Finding Cheese X

- CDC Identified 17 cheeses that were cut and repackaged at both
  - Distributor Z in California
  - Store A in Pennsylvania
- FDA obtained intact wheels of cheese and began culturing
- On September 5, isolated outbreak strain from ricotta salata
  (imported from Italy, made from pasteurized sheep milk)
Summary of Ricotta Salata Outbreak, March-October 2012

- 22 ill
  - 9 pregnancy-related
  - 13 not pregnancy-related
- 4 deaths, 1 miscarriage
- Patients in 13 states
  - isolates linked by PFGE
- All 16 patients with complete food histories consumed cheese
  - 7 ricotta salata
  - 9 another cheese
    - probably cross-contaminated by ricotta salata when cut and re-packaged
- Ricotta salata was recalled and WHO notified
Reads are chopped up into sections of length=k

We hope/assume that k is long enough to be unique in the genome. However, the assumption will be wrong in repeat regions, like in 16S.
Each kmer is a sliding window. Not mutually exclusive

Many overlapping kmers
The purpose of kmers in this project?

To pattern-match quickly across the millions of base pairs in hundreds of genomes. To find differences between genomes.

K=25
(in this example)

Genome 1

`AAAAAAAAAAAAAAAAATAAAAAAAAAAAAAAAAAA`

Genome 2

`AAAAAAAAAAAAAAAAACAAAAAAAAAAAAAA`

[Genome 3]

[Genome 4]

..... [and so on]

Check for variation
The purpose of kmers in this project?

Ultimately, find relatedness across all genomes. Find the common ancestor. Find support for which genomes might be in a cluster.