# WGS Program for Food Safety and Public Health

Patrick McDermott MS, PhD, D(AAM)
US Food and Drug Administration
Center for Veterinary Medicine
Laurel, MD

# Background

- In the US, each year about 48 million persons become ill, 128,000 are hospitalized, and 3,000 die as a result of consuming foods contaminated with known microbial pathogens.
- The annual cost of medical care and lost wages caused by infections of the 14 principal foodborne pathogens was estimated to be $1,950,000,000 in 2009.
- Identifying and controlling food safety problems focuses public health resources on those illnesses that are most likely to be related as part of an outbreak.
  - Approximately 800 outbreaks are investigated and reported each year in the U.S.
- Foodborne illnesses are reported to clinics for follow-up. Usually, the pathogen is sent to a public health laboratory for further characterization.

# Role of FDA in Foodborne Outbreak Investigations

- FDA's role during outbreak investigations includes:
  - Traceback of suspected foods to their source
  - Food and environmental testing
  - Product and regulatory actions
  - Trace-forward (if needed)
  - Environmental assessments of farm or production facilities
  - Public communications
- Food Safety Modernization Act (FSMA) 2011
  - New regulatory tools
  - Rules emphasize prevention of foodborne illnesses

# Challenges

1. The landscape of food safety is constantly evolving:
   - production techniques,
   - trade patterns/changing supply chains
   - changing consumer taste
2. Detecting outbreaks amidst the background of sporadic cases
3. Inability to identify the food vehicles and causes of food contamination responsible for those sporadic cases not associated with known outbreaks, which account for the vast majority of estimated foodborne illnesses.

# The Complex and Global Etiology of Foods

**Salad**

Shrimp – India
Cilantro – Mexico
Romaine – Salinas, CA
Cheddar – Wisconsin
Carrots – Idaho
Gruyere – Switzerland
Pecans – Georgia
Sprouts – Chicago
Red Cabbage - NY

**Sushi**

Shrimp – Indonesia
Imitation Crab – Alaska
Tuna Scrape – India
Fish Roe – Seychelles
Salmon – Puget Sound
Soy Sauce – China
Rice – Thailand
Seaweed Wrap – CA
Avocado – Mexico
Cucumber – Maryland
Wasabi – Japan
Pepper – Vietnam

**Fruit platter**

Watermelon – Delaware
Blackberries – Guatemala
Blueberries – New Jersey
Pineapple – Guam
Grapes – California
Kiwi – New Zealand
Apples – New York
Pears – Oregon
Cantaloupe – Costa Rica
Honeydew – Arizona
Papaya – Mexico
Banana – Costa Rica

# Benefits of WGS in Relation to Food Safety

- Performance: greater discriminatory power leads to more targeted response
  - Greater certainty when matching clinical, environmental, and product sample isolates
  - Links between illnesses and the potential source of contamination can be made with fewer isolates
  - Clues to geographic origin of pathogens
- Cost – a single method (e.g., no more need for typing sera).
- Speed - Faster identification of the food involved in the outbreak
- Universality
- Ease of learning and use
- Ease of sharing (common language)
- Flexible and amenable to re-analysis

# Benefits of WGS in Relation to Food Safety

- Investigators can be deployed in a more targeted manner, saving resources
- Potential to help reduce the number of foodborne illnesses and deaths over time
- Understand foodborne illness and emerging microbiological trends, including AMR
- Recurrences of pathogens in regulated food establishments/ products to further support the inspection and verification process

**Greater confidence in food safety actions**

# Potential Drawbacks of WGS in Relation to Food Safety

- Cost – in particular developing countries! Priorities?
- Perception of cost
- Data storage (global data sharing mechanisms)
- Infra-structure (internet connection/speed)
- Data handling (national capacity, int. networks, partnerships)
- Interpretation of WGS data (especially in combination with epi)
- Trust (ownership, privacy, ultimate use of data)
- **Need for basic epi, surveillance and food monitoring/testing infrastructure**

# Basic Data Flow for Global WGS Public Access Databases

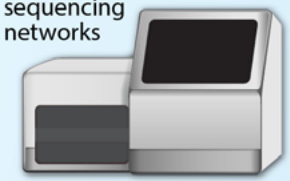
FDA

**DATA ACQUISITION**

*Sequence and upload genomic and geographic data*

GENOME TRAKR

Other distributed sequencing networks

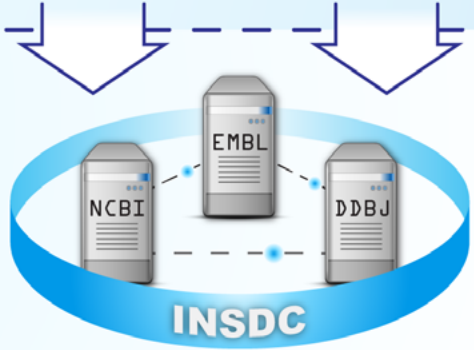


**DATA ASSEMBLY, ANALYSIS, AND STORAGE**

*International Nucleotide Sequence Database Collaboration (INSDC)*

*Shared Public Access Databases*

- *NCBI – National Center for Biotechnology Information*
- *EMBL – European Molecular Biology Laboratory*
- *DDBJ – DNA Databank of Japan*



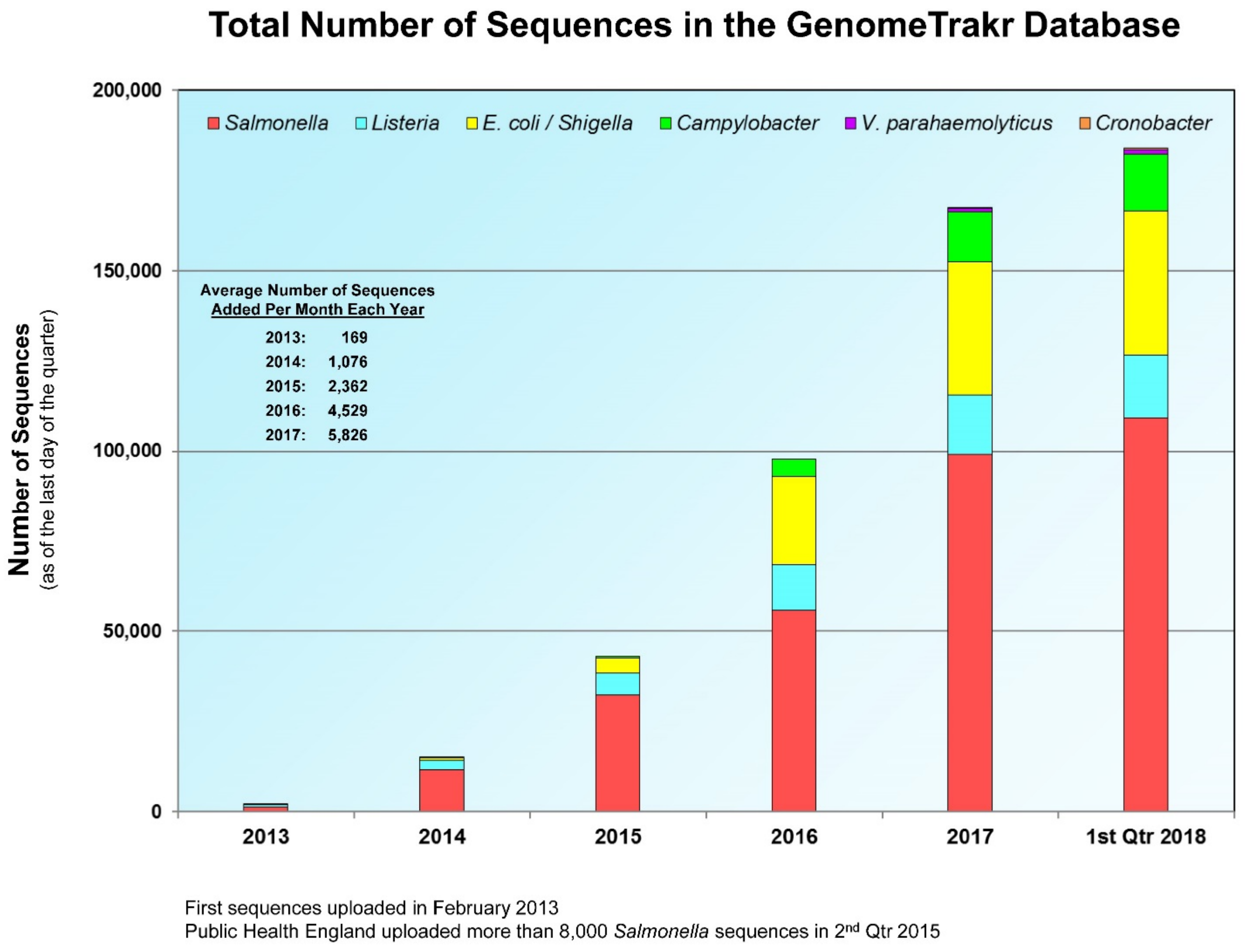

**PUBLIC HEALTH APPLICATION AND INTERPRETATION OF DATA**

- *Find clinical links*
- *Identify clusters*
- *Conduct traceback*
- *Develop rapid methods*
- *Develop culture independent tests*
- *Develop new analytical software*

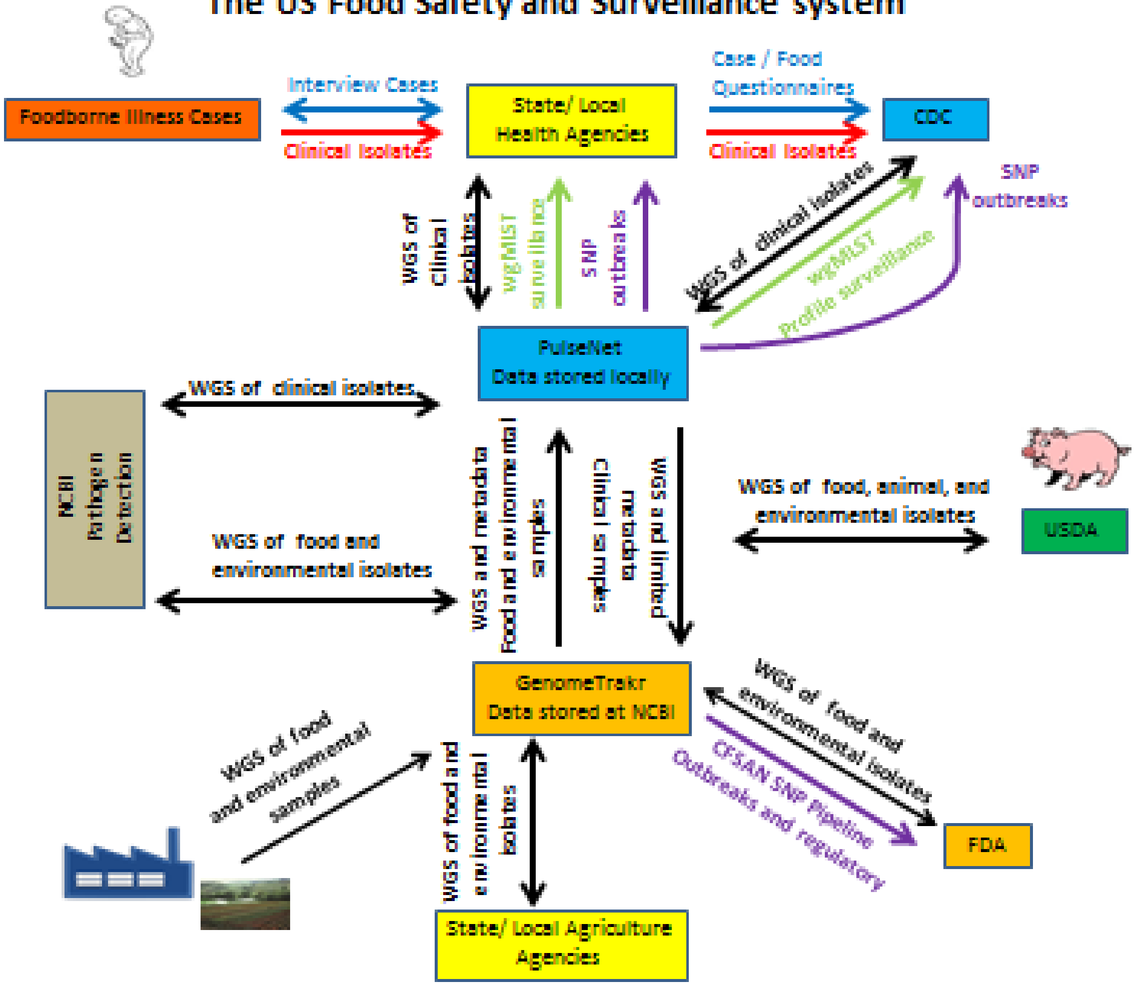
# Current Scope of GenomeTrakr Network

- Network includes labs at FDA, CDC, FSIS, 17 state health and university labs, 1 U.S. hospital lab, and 11 labs located outside the U.S.
  - Contributing labs are on 4 continents and in 10 countries
- The network provides high resolution genomic sequences of food pathogens, ex. *Salmonella, Listeria*, STEC's, others. Greater than 130,000 sequences in the database
- New GenomeTrakr labs are coming on-line
- Partnered with CDC in 2013 to study all clinical and environmental isolates of *Listeria monocytogenes, now E. coli, (Salmonella coming)*

FDA GenomeTrakr website
http://www.fda.gov/Food/FoodScienceResearch/WholeGenomeSequencingProgramWGS/ucm363134.htm


Total Number of Sequences in the GenomeTrakr Database
Salmonella
Listeria
E. coli / Shigella
Campylobacter
V. parahaemolyticus
Cronobacter
Average Number of Sequences Added Per Month Each Year
2013: 169
2014: 1,076
2015: 2,362
2016: 4,529
2017: 5,826
Number of Sequences
(as of the last day of the quarter)
200,000
150,000
100,000
50,000
0
2013
2014
2015
2016
2017
1st Qtr 2018
First sequences uploaded in February 2013
Public Health England uploaded more than 8,000 Salmonella sequences in 2nd Qtr 2015

# The US Food Safety and Surveillance system



https://www.ncbi.nlm.nih.gov/pathogens/

# Pathogen Detection BETA

View the recent webinar: 'Introducing the Pathogen Detection Isolates Browser'.

NCBI Pathogen Detection integrates bacterial pathogen genomic sequences originating in food, environmental sources, and patients. It quickly clusters and identifies related sequences to uncover potential food contamination sources, helping public health scientists investigate foodborne disease outbreaks.

Find isolates now!

Examples:
1. Search for isolates encoding a mobile colistin resistance gene and a KPC beta-lactamase
search: AMR_genotypes:mcr* AND AMR_genotypes:blaKPC*
2. Search for Salmonella isolates from the USA
search: geo_loc_name: USA AND taxgroup_name:"Salmonella enterica"

## Explore the Data

| Species | New Isolates | Total Isolates |
|---|---|---|
| **Salmonella enterica** | 108 | 153,255 |
| **E.coli and Shigella** | 73 | 54,679 |
| **Campylobacter jejuni** | 63 | 21,838 |
| **Listeria monocytogenes** | 5 | 20,879 |

**See more organisms...**

## Learn More

About
FAQ
Browser Factsheet
Antimicrobial Resistance Factsheet
Antimicrobial Resistance
Contributors

## Data Resources

Isolates Browser
Antimicrobial resistance reference gene database
Isolates with antibiotic resistant phenotypes
Beta-lactamase resources
Download analysis results (FTP)

## Submit

How to submit data
How to submit antibiotic resistance phenotypes
How to submit beta-lactamases
NCBI Submission Portal

FDA

# General guidelines for establishing that isolates arose from a common source

| | Supports | Neutral | Refutes |
|---|---|---|---|
| SNP distance | < 20 | 20 – 100 | > 100 |
| Bootstrap support | > 0.90 | 0.80 – 0.90 | < 0.80 |
| Tree topology | Monophyletic | Paraphyletic | Polyphyletic |

1. Supporting epidemiology or traceback information are required to justify decisions.

2. Isolates with any combination of supporting or neutral evidence may ultimately be determined to match (but see point 1).

3. A finding that any WGS evidence refutes a match is sufficient to eliminate the possibility of an overall match between two isolates (at least until more data are collected).

This approach reduces the chance that minor variations in a category of evidence will lead to significant changes in the interpretation of WGS analyses.

# *E. coli* O121 in Flour

Food Safety News

*Breaking news for everyone's consumption*

Google Custom Sea

Home | Outbreaks | Food Recalls | Food Politics | Calendar | Subscribe | Directory | Media Kit | About Us | Contact Us

## General Mills expands flour recall after more illnesses are reported

CDC update: 42 E. coli O121 cases in 21 states

BY CATHY SIEGNER | JULY 1, 2016

General Mills announced Friday that it was expanding its recall of Gold Medal flour, Wondra flour, and Signature Kitchens flour to include flour made earlier in the fall that may still be in consumers' pantries.

The company, based in Minneapolis, stated that the recall is being expanded due to a newly reported illness which "appears to have stemmed from the consumption of raw dough or batter linked to flour produced last fall." (The expanded list of recalled flour products is toward the end of this story.)

The recalled flour has been linked to four new cases of infection with Shiga toxin-producing (STEC) E. coli O121 since June 1, according to an update posted Friday by the U.S. Centers for Disease Control and Prevention (CDC).


There are now 42 cases in 21 states, with 11 related hospitalizations. One new state, Indiana, has been added to the list of states with ill people, the agency added.

No one has developed hemolytic uremic syndrome, a dangerous type of kidney failure associated with E. coli

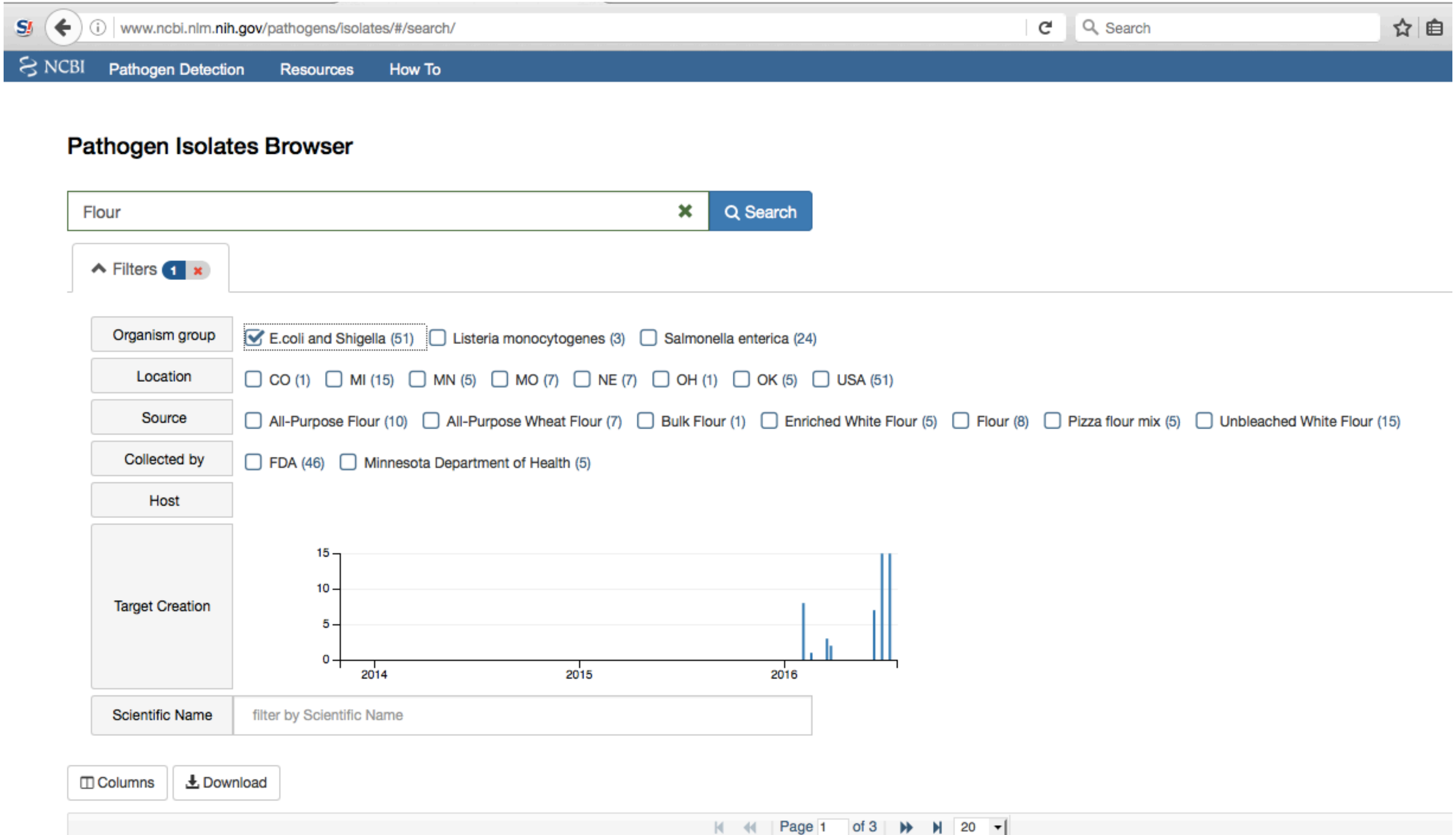
### FOOD RECALLS

- 3,000 cases of fruit-flavored ice pops recalled for Listeria risk
- Wegmans recalls detox tea because of Salmonella risk
- Another company recalls kratom after illness reported
- FDA Salmonella finding prompts recall of three raw dog foods

### HOT FOOD BLOGS

- Connecticut Department of Health weighs in on E. coli Outbreak Linked to Yuma Romaine
- Frozen Pops Recall over Listeria Concerns
- New Jersey Reports 7 with E. coli Linked to Romaine
- Letter From The Editor: Saluting barfblog and eFoodAlert
- FDA Flexes Muscles To Achieve

# *E. coli* 0121 in Flour



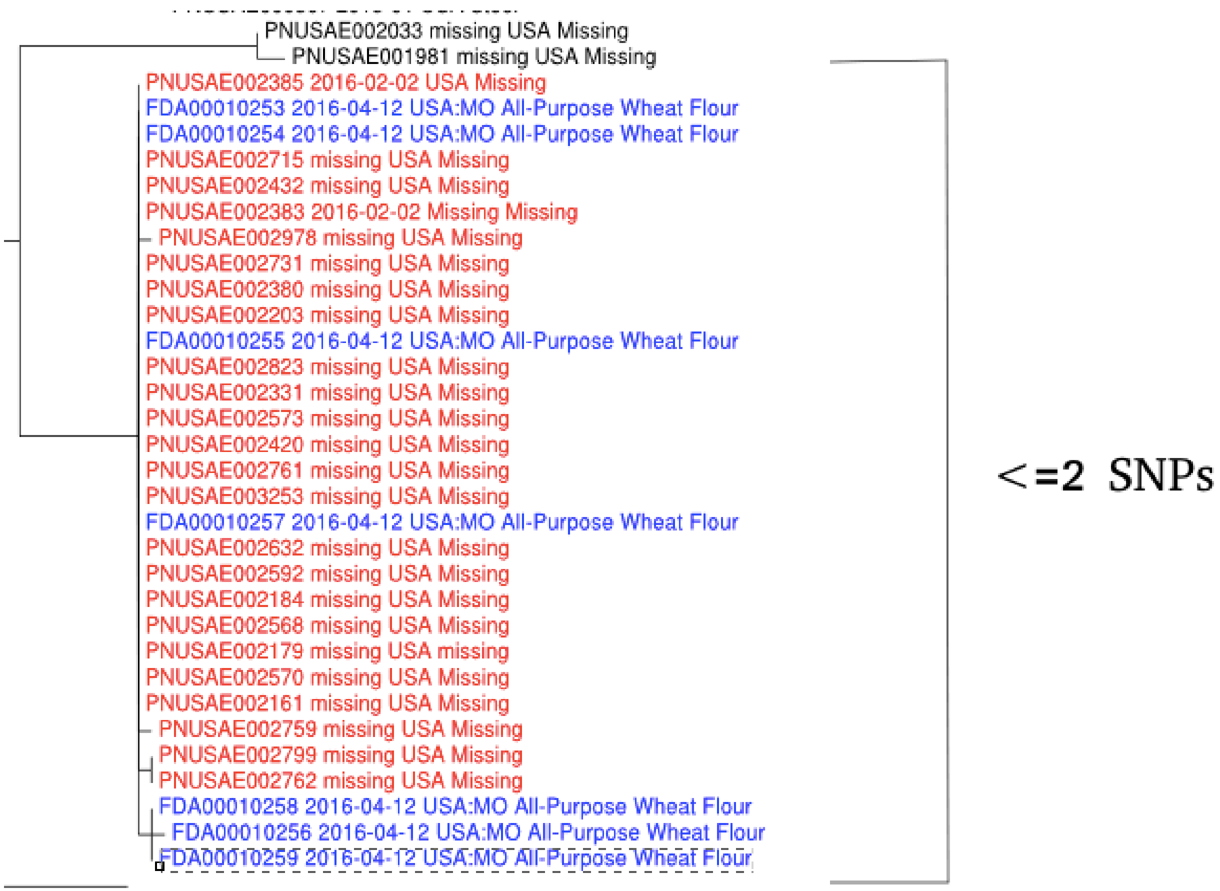
0-3 SNPs to clinical isolates

| Target | Cluster | Min-same | Min-opp | attribute_package | biosample_acc | isolation_sourc |
|---|---|---|---|---|---|---|
| PDT000133980.1 | PDS000003441.33 | 0 | 0 | Pathogen: environmental/food/other | SAMN05215988 | All-Purpose Wheat Flour |
| PDT000133981.1 | PDS000003441.33 | 0 | 0 | Pathogen: environmental/food/other | SAMN05215989 | All-Purpose Wheat Flour |
| PDT000133982.1 | PDS000003441.33 | 0 | 0 | Pathogen: environmental/food/other | SAMN05215990 | All-Purpose Wheat Flour |
| PDT000133983.1 | PDS000003441.33 | 3 | 3 | Pathogen: environmental/food/other | SAMN05215991 | All-Purpose Wheat Flour |
| PDT000133984.1 | PDS000003441.33 | 0 | 0 | Pathogen: environmental/food/other | SAMN05215992 | All-Purpose Wheat Flour |
| PDT000133985.1 | PDS000003441.33 | 0 | 0 | Pathogen: environmental/food/other | SAMN05215993 | All-Purpose Wheat Flour |
| PDT000133986.1 | PDS000003441.33 | 1 | 1 | Pathogen: environmental/food/other | SAMN05215994 | All-Purpose Wheat Flour |

0-3 SNPs to other food/env isolates


17

# CFSAN SNP Pipeline



# Inspections of High-Risk Facilities

Home | Foodborne Illness Outbreaks | Food Recalls | Food Politics | Events | Subscribe | About Us

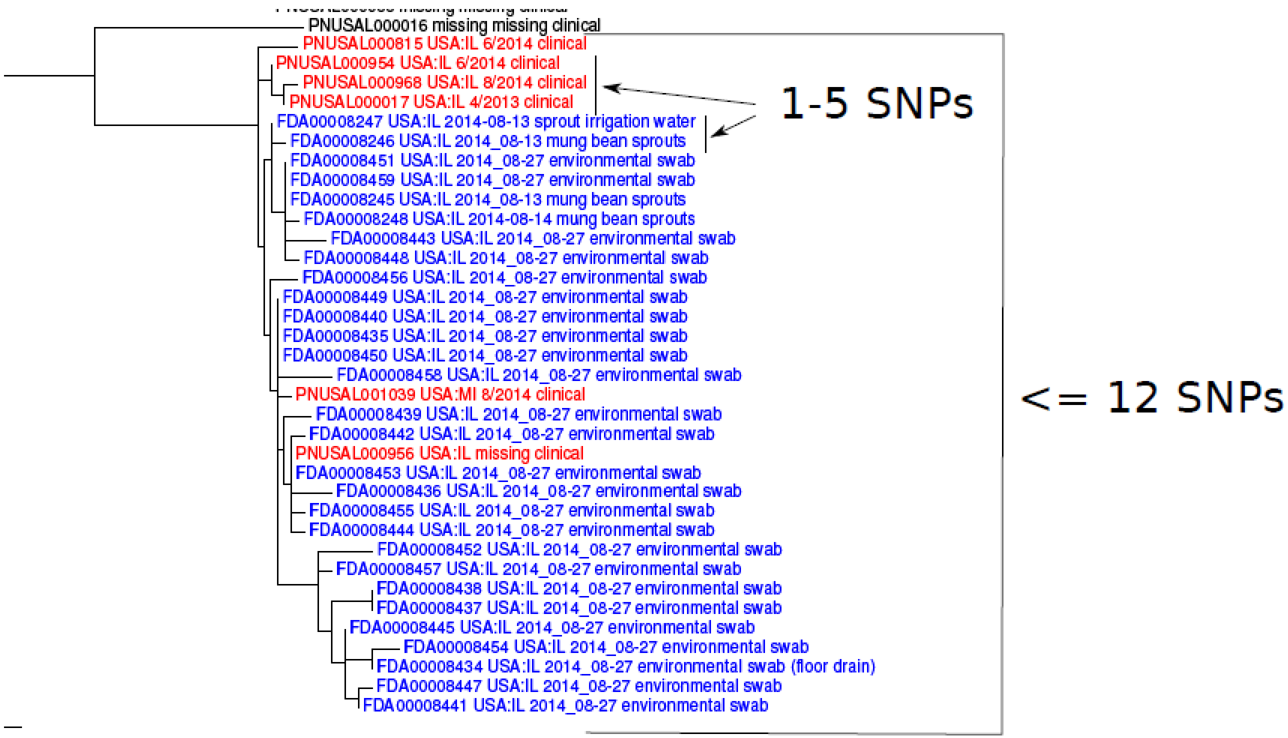
## Recalled Bean Sprouts Linked to 2 Listeria Deaths, 3 Hospitalizations

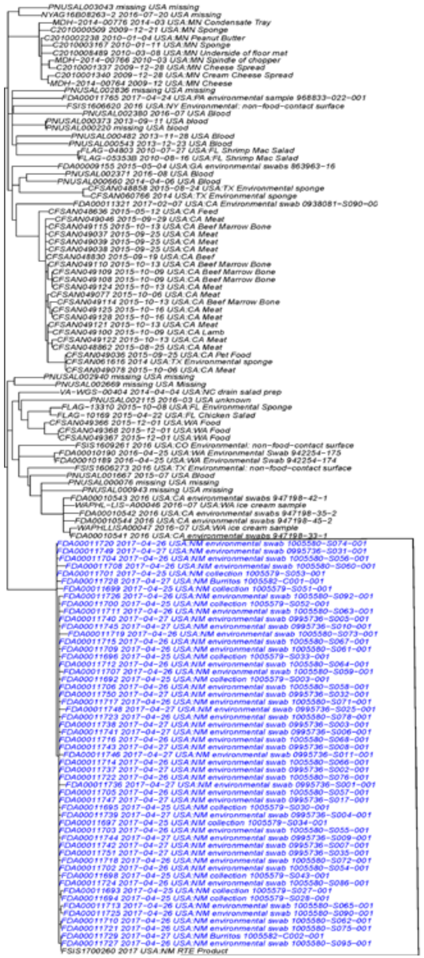
BY **NEWS DESK** | NOVEMBER 8, 2014

Two people have died and three others have been hospitalized after eating Listeria-contaminated bean sprouts produced by Wholesome Soy Products of Chicago, according to the U.S. Centers for Disease Control and Prevention.

The victims became ill between June and August 2014, but this is the first announcement of the outbreak. It was detected retroactively using whole-genome sequencing, a new technology for detecting outbreaks which utilizes DNA sequencing of bacteria.

# L. mono in Sprouts





<= 9 SNPs

# *Observation*:

Isolates with small genetic distances are often from the same facility.

# Fitness Traits of Interest

(1) Thermal tolerance
(2) Dessication resistance
(3) Osmotic/Ionic tolerance
(4) Quat resistance
(5) Chlorine resistance
(6) Biofilm persistence
(7) Surface adherence
(8) Antibiotic resistance
(9) Antimicrobial resistance
(10) Ecological fitness
(11) Heavy metal resistance
(12) Metabolic persistence
(13) Enhanced hydrophobic fitness
(14) Produce invasiveness
(15) Flower invasiveness
(16) Root system invasiveness
(17) Acid resistance
(18) Surface water fitness
(19) In vivo plant migratory fitness
(20) Soil fitness
(21) Capsaicin resistance
(22) Swarming
(23) Trans-ovarian poultry colonization
(24) Fecal persistence (poultry)
(25) Yolk content invasion
(26) Multidrug resistance
(27) External amoeba harborage
(28) Internal amoeba harborage
(29) Acyl-homoserine lactone (AHL)
(30) KatE stationary-phase catalase
(31) In vivo migratory fitness
(32) RDAR phenotype
(33) The 'Weltevreden' type
(34) Peristence within the tomato**

# Summary

- **WGS is now routine** in US outbreak response and compliance surveillance. In collaboration with other public health agencies (FSIS, CDC), WGS has been used in numerous foodborne contamination events, including those cause by antibiotic resistant bacteria.
- We **expect many more small outbreaks** to be identified that were previously categorized as sporadic infections by linking them to specific food or environmental sources.
- Numerous additional applications exist for using WGS including supply chain management, quality assurance, process evaluation, etc.
- Genome sequences are **portable, instantly cross-compatible and highly scalable**. One technology approach regardless of organism.
- Have to balance the need for increased number of well characterized **environmental** (food, water, facility, etc.) sequences with the need for extensive clinical isolates
- WGS, unlike PFGE, is more than a surveillance tool. It provides **comprehensive information** on traits of medical and food safety importance, including historical reference to pathogen emergence.
- Instead of multiple food safety programs using different technologies, all will exploit the same data set for different purposes. **Sample is king.**

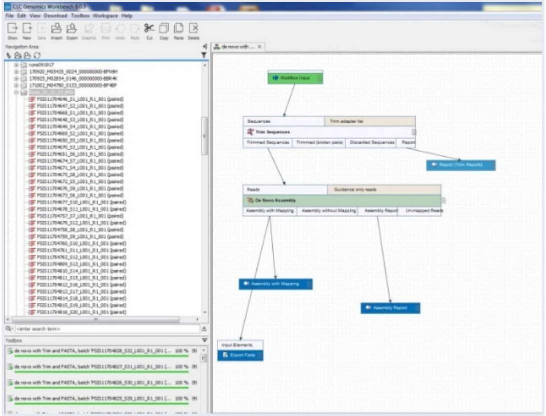
U.S. FOOD & DRUG ADMINISTRATION

# WGS Data Analyses Work Flow Overview


Input: FASTQ

QC Pipeline
- Coverage
- Average Quality
- Nucleotide balance

De novo Assembly



- wgMLST BioNumerics 7.6
- Lyve-SET, SNP Pipeline
- NCBI Pathogen Isolate Browser

QC Pipeline
- File Size
- N50 & Contigs
- Correct organism

Output: FASTA

```
>FSIS1609314_S10_L001_R1_001_(paired)_trimmed_(paired)_contig_1
TGCGGTTTTGTACGTTCAAATTTTTCTTTAGACACGGCTATATTCCTTACTATAGCGCTC
TCCCCTTCAGGAGAGAGCACGGGATTTTGGTTTTTAACCCTGCGGCTTATTTACCACGGG
CTTCGATTACGGCCTGAGCAACGTTGTTCGGCGCATCATCATACTTCAGGAATTCCATAG
TGTACGATGCGCGACCTTTGGTCAGAGAACGCAGCTGAGTTGCATATCCGAACATTTCAG
ACAGCGGTACTTCAGCGTGGATCTTAACGCCTGTAACTTCAGATTCCTGACCTTTGAGCA
TACCACGACGGCGGCTAAGGTCGCCGATAACGTCACCGGTATTCTCTTCAGGTGTTTCTA
CTTCAACCTTCATGATCGGCTCAAGCAGAACTGGTTTTGCTTTCTTAAAGCCTTCTTTAA
AGGCGATAGACGCAGCCAGTTTAAACGCCAGCTCAGAGGAGTCAACGTCATGGTAAGAAC
```

Input: FASTA

- MLST Sequence Type
- Antibiotic Resistance genes
- Virulence Profile
- Salmonella and STEC serotype
- MASH Tree comparison