Establishing a Whole Genome Sequence-Based national network for the detection and traceback of Foodborne Pathogens

Marc W. Allard Ph.D.
Microbiologist, Division of Microbiology, ORS,
Center for Food Safety and Applied Nutrition, (FDA) USA

June. 16th, 2014, efsa Scientific Colloguium Series:
Use of Whole Genome Sequencing (WGS) of food-borne pathogens for public health protection

S. Bareilly Outbreak (April-June 2012)

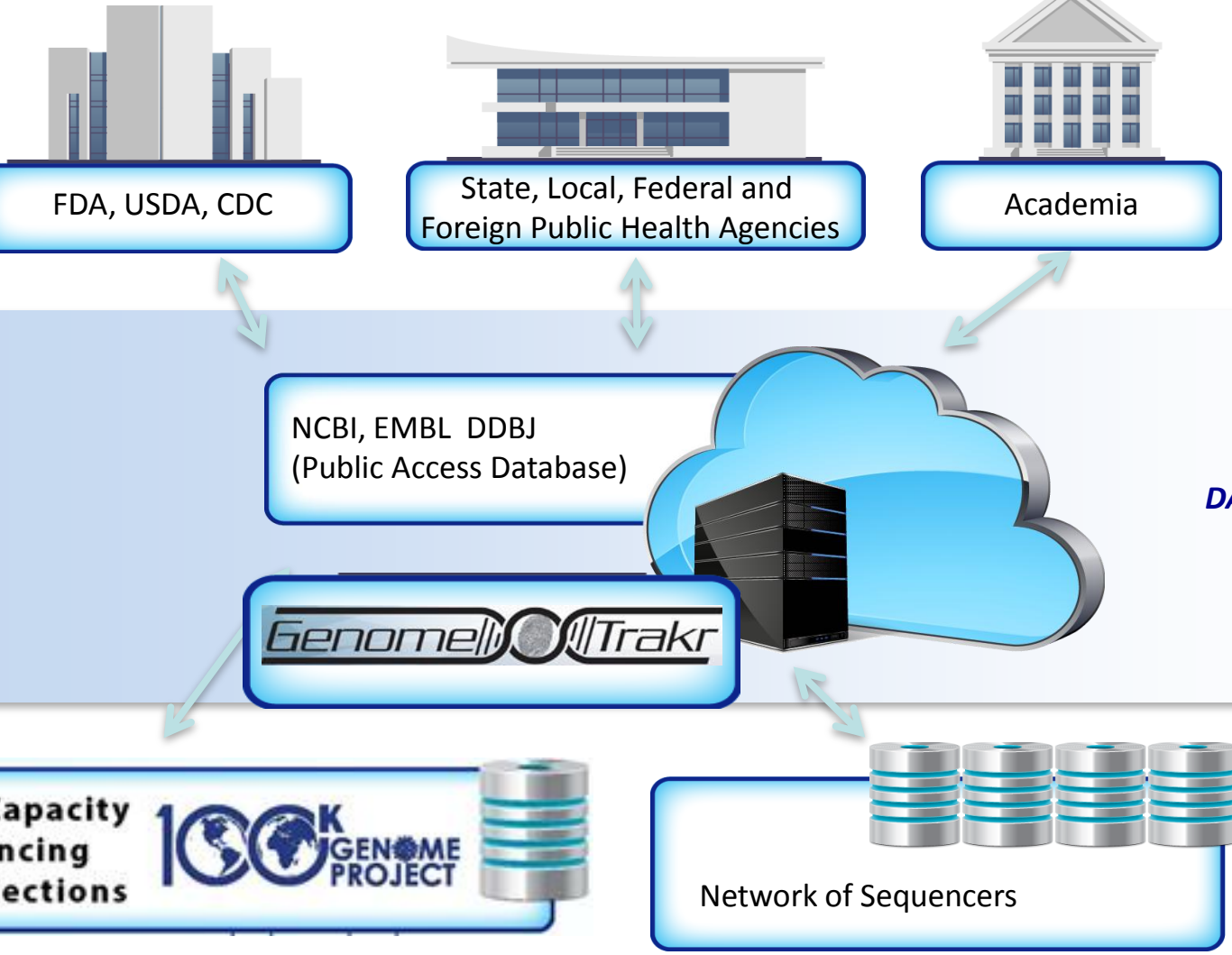
PFGE Match

110-130 SNPs

Roles for regulatory use of whole genome sequencing

1. Source information
    1. Country or state/region of origin
    2. Infer a type of food or environment
    3. Better target investigational resources
2. Surveillance
3. Risk assessment and modeling
4. Replace traditional bacteriological typing procedures
    1. Speed
    2. Accuracy
    3. Cost

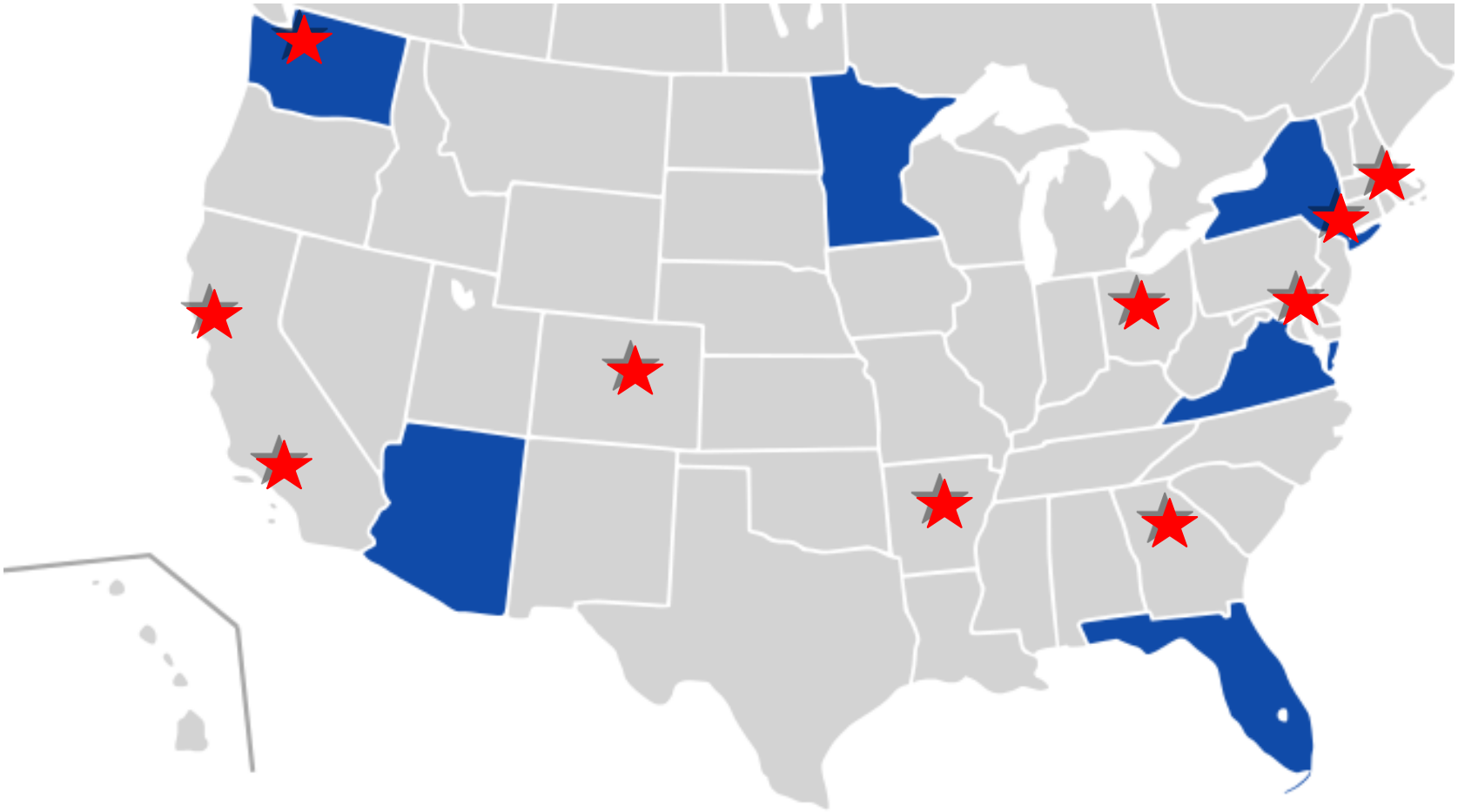
Our Current Model – Public



Network of Sequencers

7 state health depts.

+ 10 FDA-ORA



Inputs

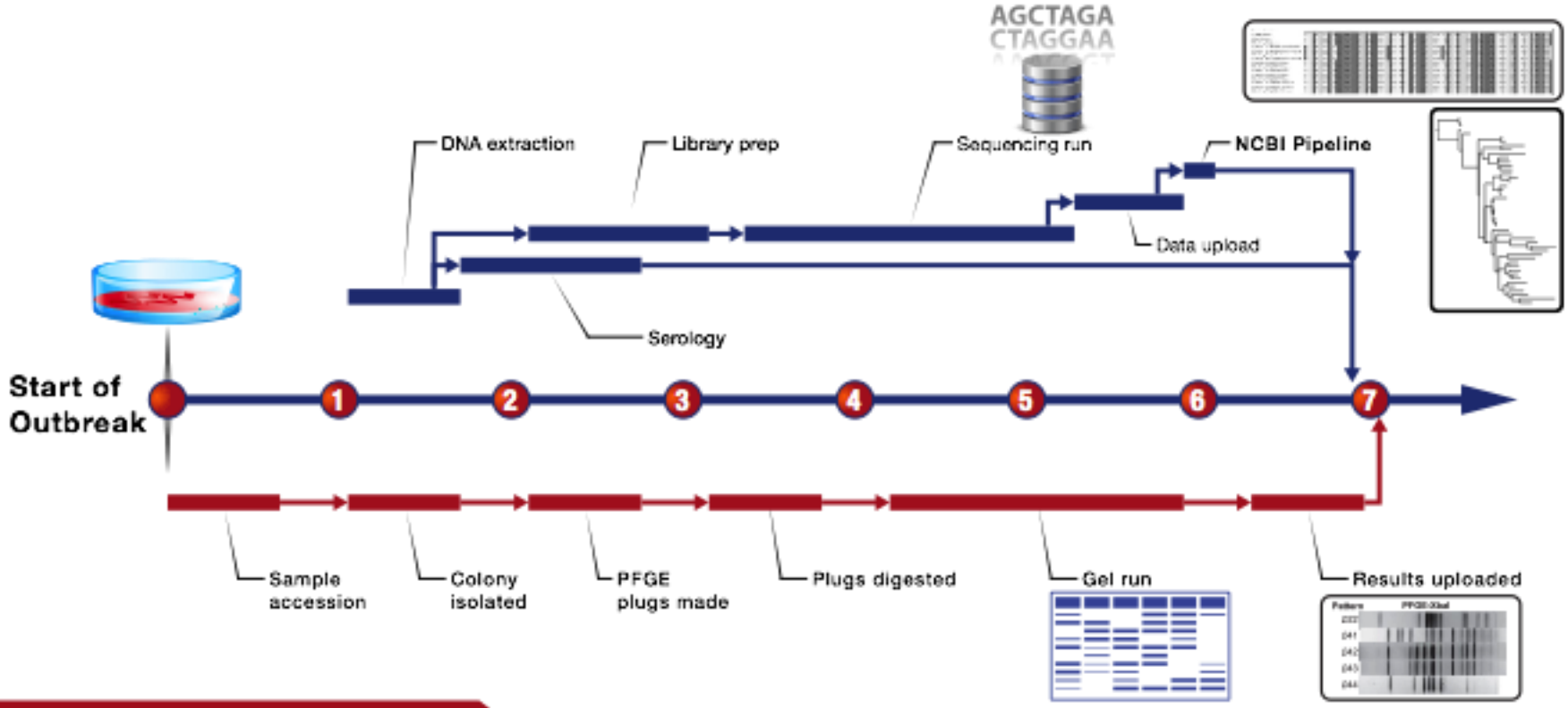
- 1 Miseq system
- Sufficient reagents to sequence > 300 genomes per year
- Dedicated scientific staff (bioinformatics and/or laboratory support) through Oak Ridge Institute for Science and Education (ORISE)
- Bioinformatics and laboratory support, analysis pipeline

Deliverables

- Minimum ~300 genomes with metadata uploaded to NCBI per annum, minimum 20X coverage
- food and environmental related bacterial (prefer *Salmonella*) isolates

Next-generation Lab Response vs. Conventional Lab Response

Next-Generation Lab Response



Conventional Lab Response

1 = Day 1

11/27/2013

MINIMAL PATHOGEN METADATA (FOODBORNE OUTBREAKS)

What

sample_name
organism
strain/isolate

Category (attribute_package)
1a) **Clinical/Host-associated**
- 1a1) **specific_host**
- 1a2) **isolation_source**
- 1a3) host-disease

OR

1b) **Environmental/Food/Other**
- 1b1) **isolation_source**

When

collection_date

Where

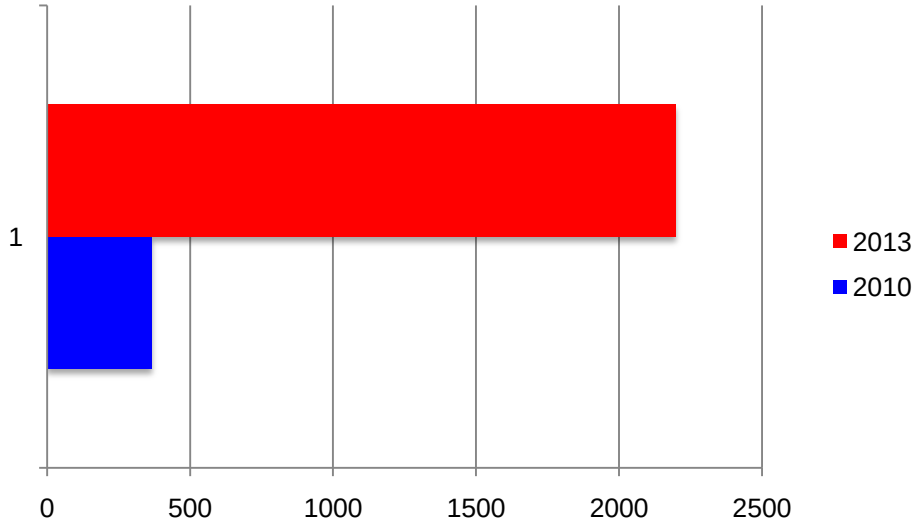
Geographic location

6a) **geo_loc_name**
OR
6b) **lat_lon**

Who

collected by

Sequences generated by Genometrakr and FDA



FDA-State Desktop Pilot called GenomeTrakr - Public Face at NCBI

http://www.ncbi.nlm.nih.gov/bioproject/183844

Food and Drug Administration, Center for Food Safety and Applied Nutrition: GenomeTrakr Project

Accession: PRJNA183844 ID: 183844

Currently encompasses whole genome sequencing of cultured pathogens as part of a surveillance project for the rapid detection of outbreaks of foodborne illnesses

Project Type: Umbrella project (***Subtype:*** Disease)

Relevance: Agricultural, Medical, Industrial, Environmental

Project Data:

Resource Name	Number of Links
SEQUENCE DATA	
Nucleotide (total)	31172
WGS master	194
Genomic DNA	30978
SRA Experiments	437
Protein Sequences	901056
PUBLICATIONS	
PubMed	12
PMC	9
OTHER DATASETS	
BioSample	1104
GEO DataSets	1

Project Type	Number of Projects
Genome sequencing	
Highest level of assembly :	
SRA or Trace	1
No data links	4
Total	5

BioProject accession	Assembly level	Name	Title
PRJNA183847	-	GenomeTrakr Project: Arizona State Public Health Laboratory	GenomeTrakr Project: Arizona State Public Health Laboratory (Arizona State Public Health...)
PRJNA183848	-	GenomeTrakr Project: Florida Department of Health	GenomeTrakr Project: Florida Department of Health (Florida Department of Health)
PRJNA186035	SRA or Trace	GenomeTrakr Project: Food and Drug Administration, Center for Food Safety and Applied Nutrition	GenomeTrakr Project: Food and Drug Administration, Center for Food Safety and Applied Nutrition (Center for Food Safety and...)
PRJNA183850	-	GenomeTrakr Project: New York State Dept. of Health, Wadsworth Center	GenomeTrakr Project: New York State Dept. of Health, Wadsworth Center (New York State Dept. of Health...)
PRJNA183851	-	GenomeTrakr Project: Washington State Department of Health Public Health Laboratory	GenomeTrakr Project: Washington State Department of Health Public Health Laboratory (Washington State Department...)

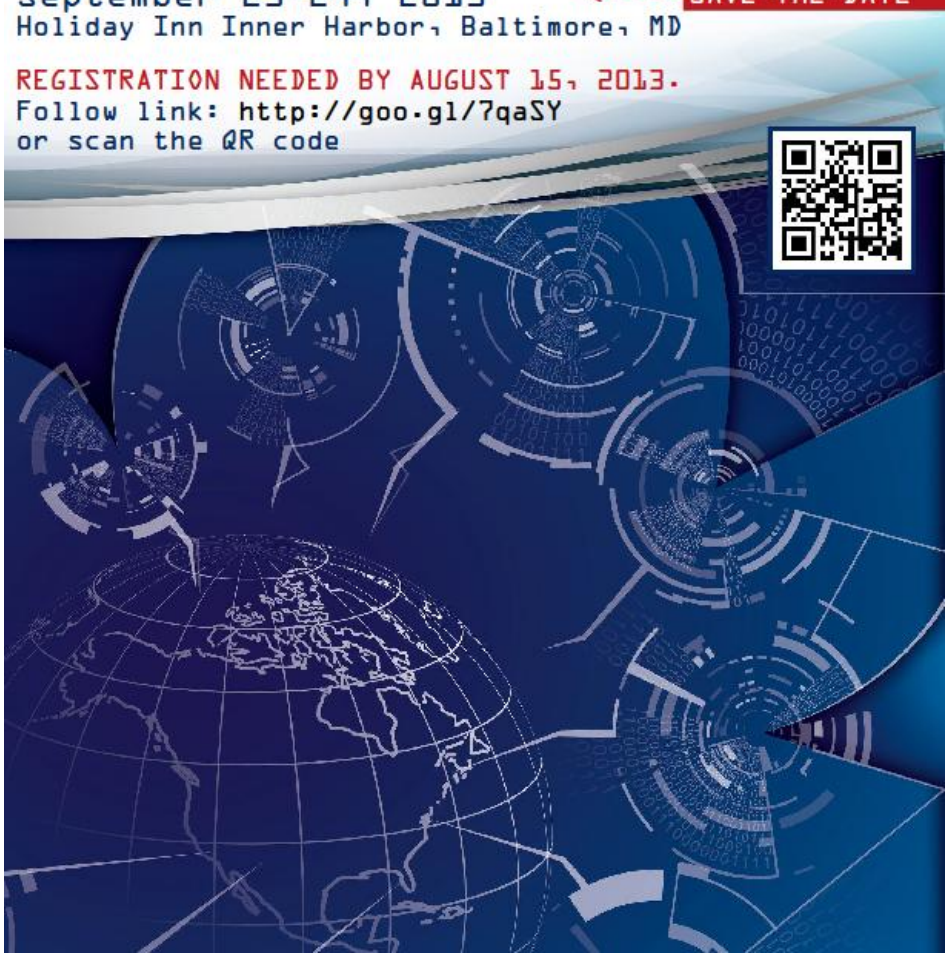
MN and VA most recent partners. Mexico Sinaloa 1st international partner.
SRA completed experiments are ~4,000 records to date. ~500/mo.

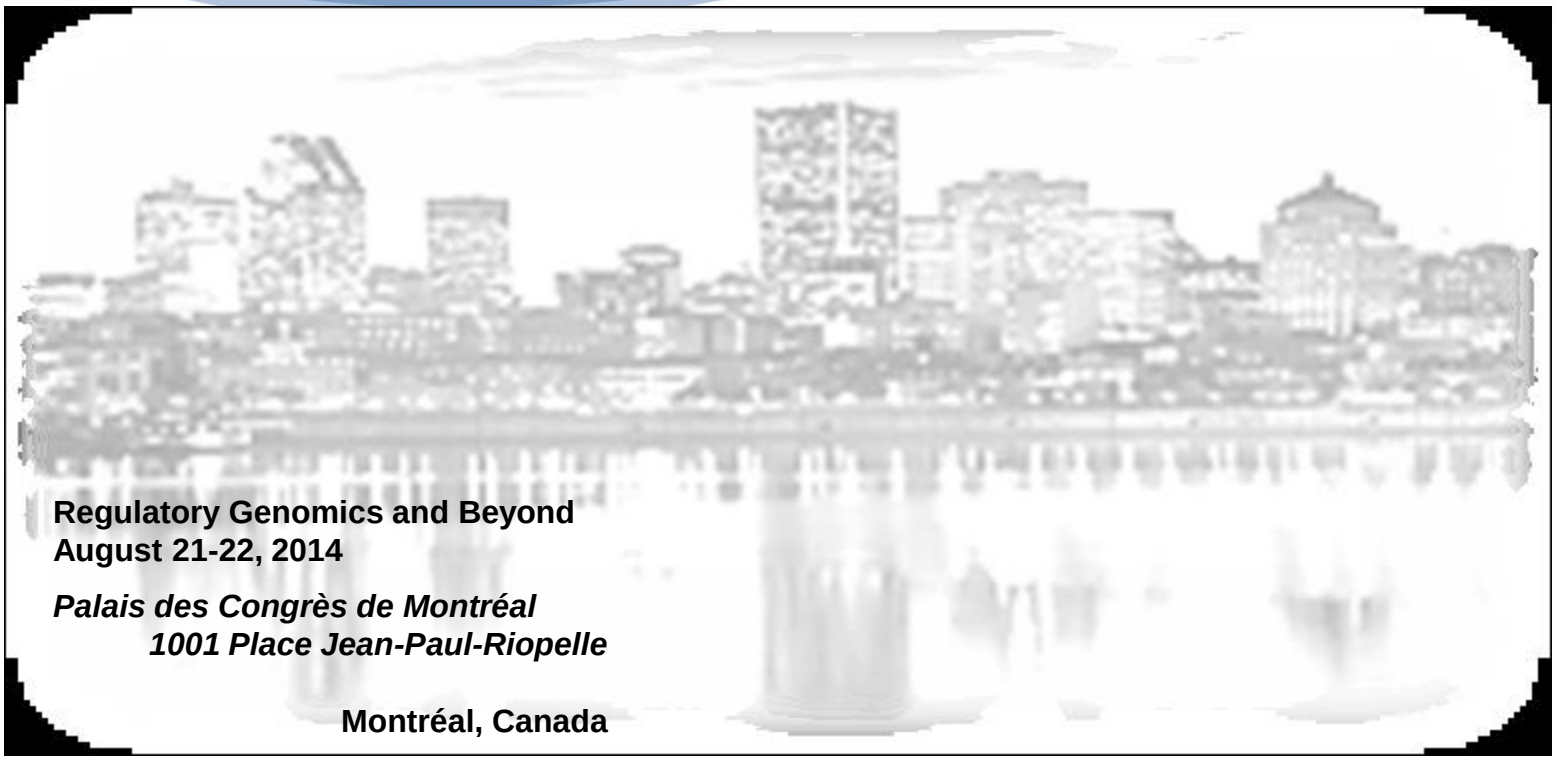
NGN-FPT

Next Generation Sequencing Network for Food Pathogen Traceability Project

September 23-24, 2013 <-----SAVE-THE-DATE
Holiday Inn Inner Harbor, Baltimore, MD

REGISTRATION NEEDED BY AUGUST 15, 2013.
Follow link: http://goo.gl/7qaSY
or scan the QR code




Regulatory Genomics and Beyond
August 21-22, 2014
Palais des Congrès de Montréal
1001 Place Jean-Paul-Riopelle
Montréal, Canada


Global Microbial Identifier
7th Meeting
11 - 12 September 2014, York, UK

Example of how the network functions – Collaboration with NCBI and CDC on real-time *Listeria* outbreak detection – 8 months into project

The goal of this project is to type every clinical and food/environmental isolate of *Listeria monocytogenes* collected in the US.

Listeria monocytogenes

Accession: PRJNA211

Project Type: Umbrella project (***Subtype:*** Disease)

Relevance: Medical, Environmental

Keyword: GMI

▼ SRA Data Details

Parameter	Value
Data volume, Gbases	202
Data volume, Tbytes	0.11

Listeria monocytogenes encompasses the following 16 sub-projects:

Project Type

Genome sequencing

Highest level of assembly :
SRA or Trace

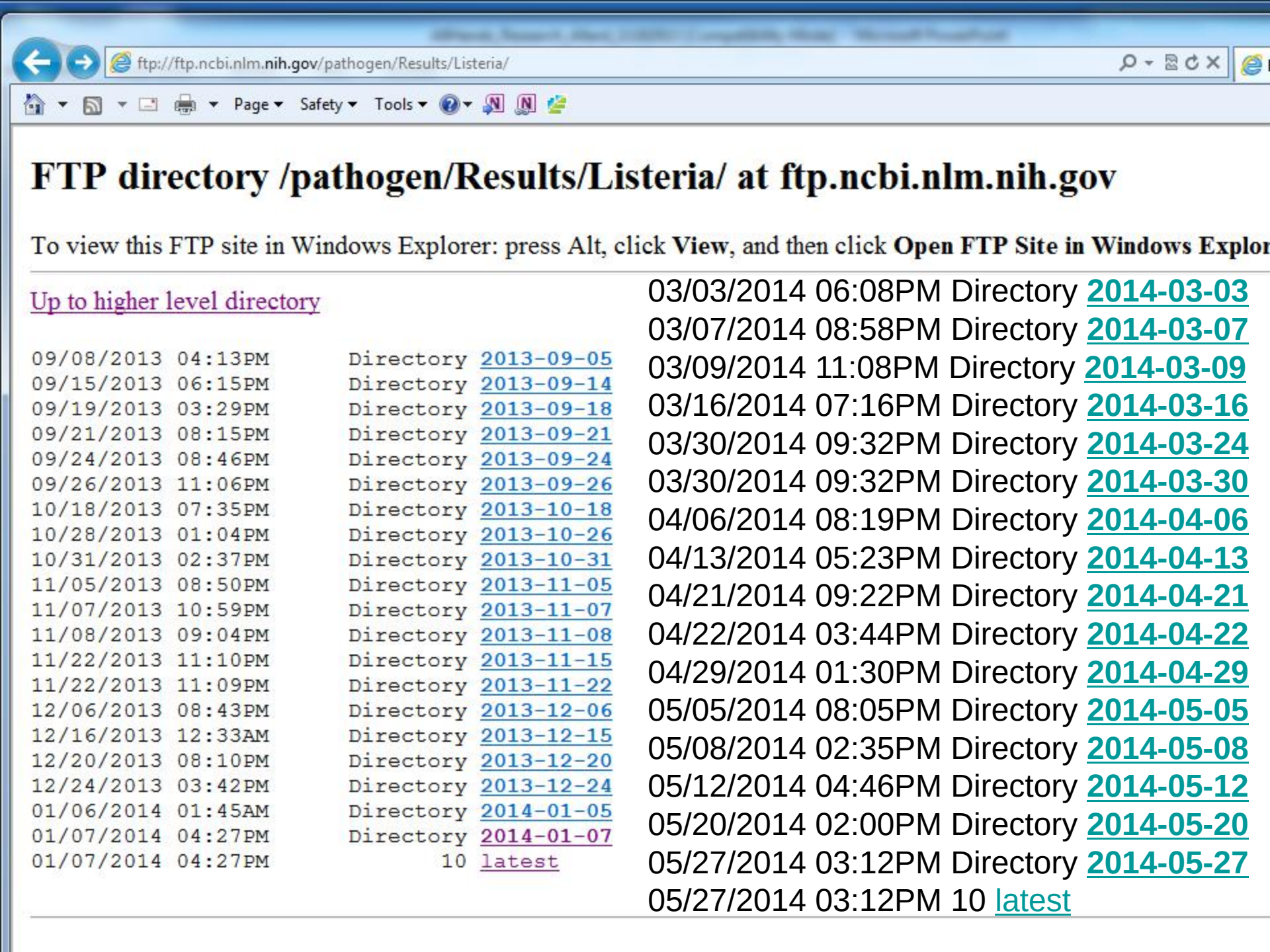
BioProject accession	Assembly level	Organism	Title
PRJNA218587	SRA or Trace	Listeria	Listeria
PRJNA212117	SRA or Trace	Listeria monocytogenes	Listeria
PRJNA215355	SRA or Trace	Listeria monocytogenes	Listeria
PRJNA218633	SRA or Trace	Listeria monocytogenes CFSAN003725	Listeria
PRJNA218634	SRA or Trace	Listeria monocytogenes CFSAN003726	Listeria
PRJNA218635	SRA or Trace	Listeria monocytogenes CFSAN003727	Listeria
PRJNA218636	SRA or Trace	Listeria monocytogenes CFSAN003728	Listeria
PRJNA218637	SRA or Trace	Listeria monocytogenes CFSAN003729	Listeria
PRJNA218638	SRA or Trace	Listeria monocytogenes CFSAN003806	Listeria

FTP directory /pathogen/Results/ at ftp.ncbi.nlm.nih.gov

To view this FTP site in Windows Explorer: press Alt, click **View**, and then click **Open FTP Site in Windows Explor**

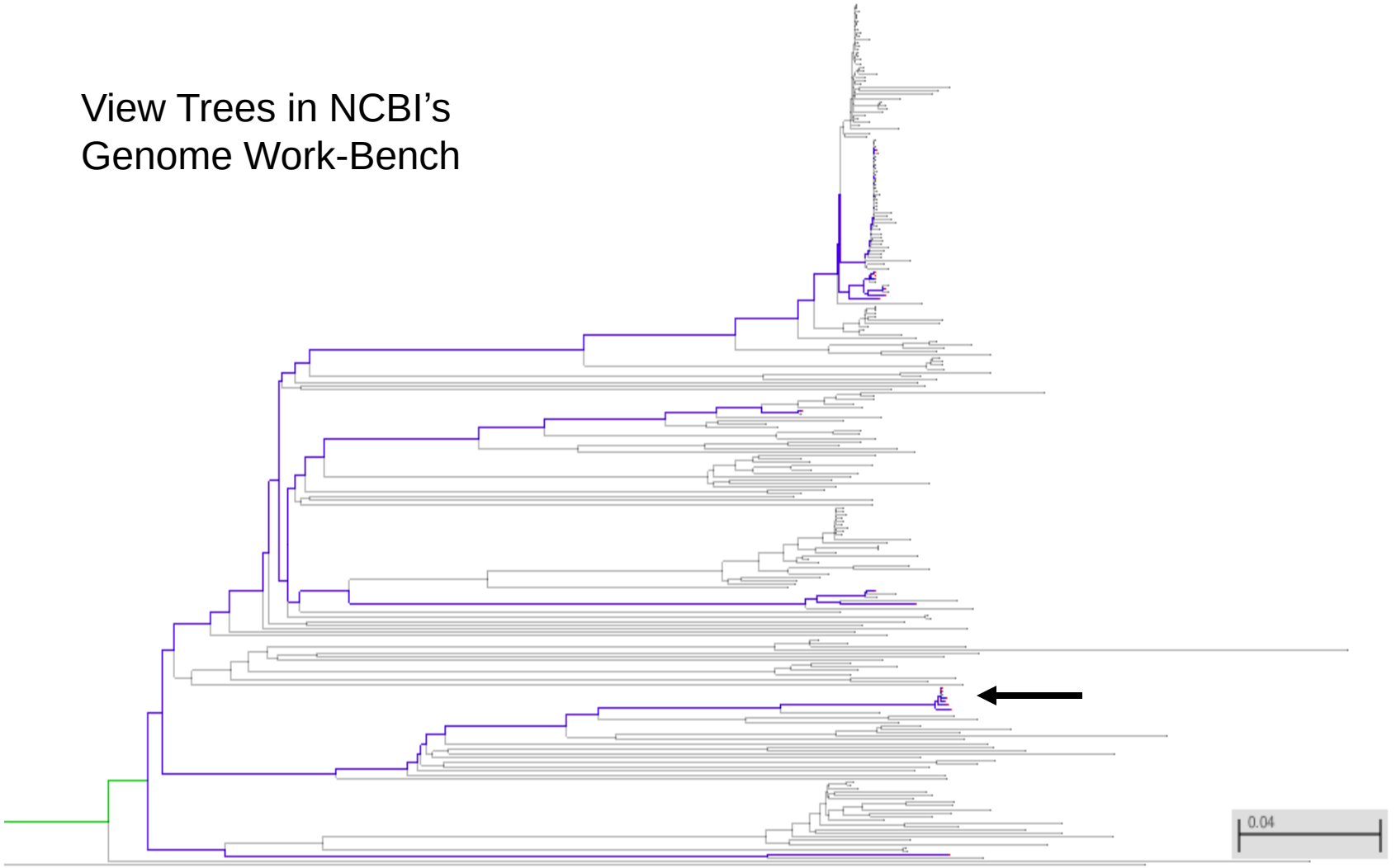
Up to higher level directory

09/23/2013 12:50AM Directory Campylobacter
01/07/2014 03:41PM Directory Escherichia_coli_Shigella
01/07/2014 04:27PM Directory Listeria
01/06/2014 08:56PM Directory Salmonella

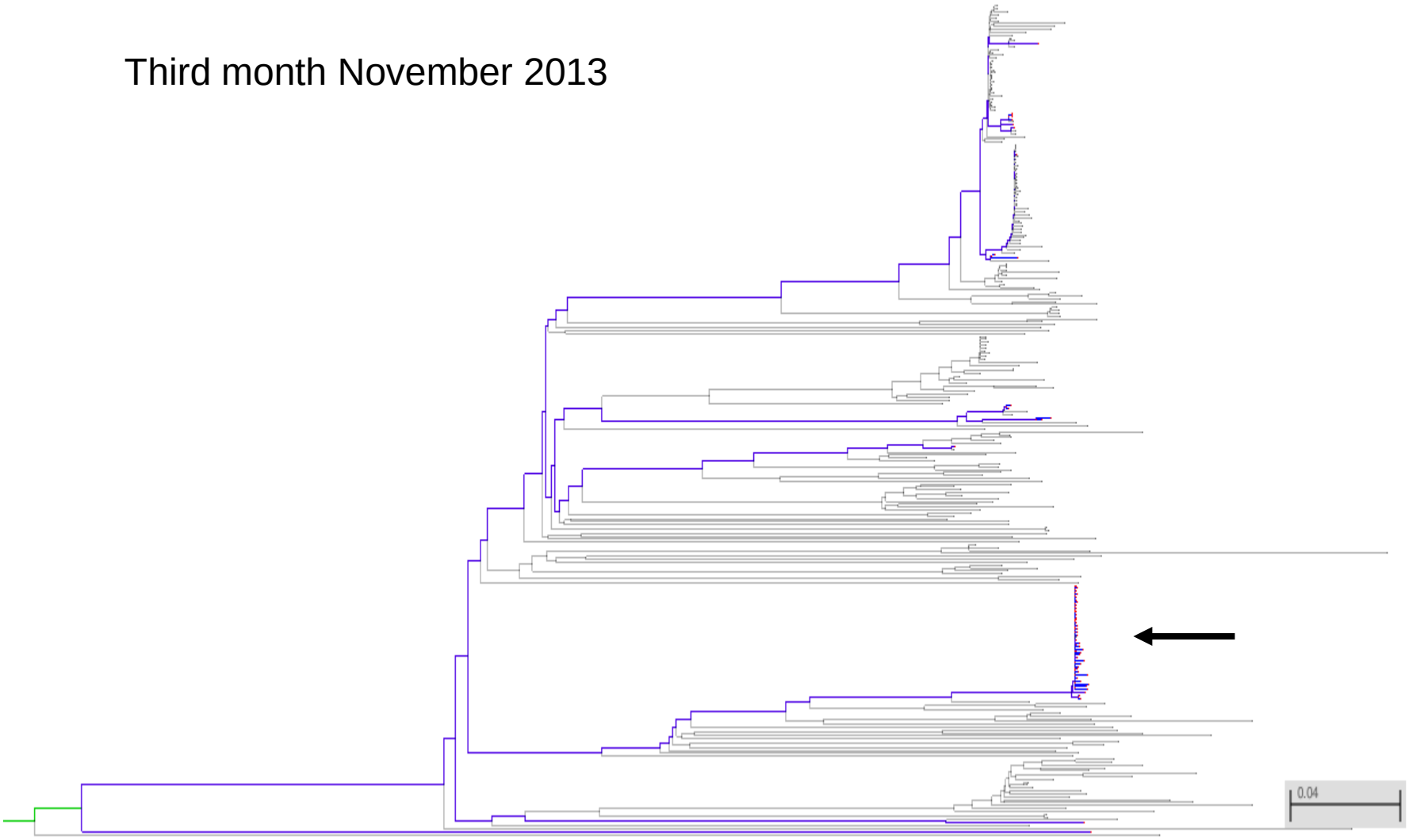

ftp://ftp.ncbi.nlm.nih.gov/pathogen/Results/Listeria/
Page
Safety
Tools
FTP directory /pathogen/Results/Listeria/ at ftp.ncbi.nlm.nih.gov
To view this FTP site in Windows Explorer: press Alt, click View, and then click Open FTP Site in Windows Explor
Up to higher level directory
09/08/2013 04:13PM Directory 2013-09-05
09/15/2013 06:15PM Directory 2013-09-14
09/19/2013 03:29PM Directory 2013-09-18
09/21/2013 08:15PM Directory 2013-09-21
09/24/2013 08:46PM Directory 2013-09-24
09/26/2013 11:06PM Directory 2013-09-26
10/18/2013 07:35PM Directory 2013-10-18
10/28/2013 01:04PM Directory 2013-10-26
10/31/2013 02:37PM Directory 2013-10-31
11/05/2013 08:50PM Directory 2013-11-05
11/07/2013 10:59PM Directory 2013-11-07
11/08/2013 09:04PM Directory 2013-11-08
11/22/2013 11:10PM Directory 2013-11-15
11/22/2013 11:09PM Directory 2013-11-22
12/06/2013 08:43PM Directory 2013-12-06
12/16/2013 12:33AM Directory 2013-12-15
12/20/2013 08:10PM Directory 2013-12-20
12/24/2013 03:42PM Directory 2013-12-24
01/06/2014 01:45AM Directory 2014-01-05
01/07/2014 04:27PM Directory 2014-01-07
01/07/2014 04:27PM 10 latest
03/03/2014 06:08PM Directory 2014-03-03
03/07/2014 08:58PM Directory 2014-03-07
03/09/2014 11:08PM Directory 2014-03-09
03/16/2014 07:16PM Directory 2014-03-16
03/30/2014 09:32PM Directory 2014-03-24
03/30/2014 09:32PM Directory 2014-03-30
04/06/2014 08:19PM Directory 2014-04-06
04/13/2014 05:23PM Directory 2014-04-13
04/21/2014 09:22PM Directory 2014-04-21
04/22/2014 03:44PM Directory 2014-04-22
04/29/2014 01:30PM Directory 2014-04-29
05/05/2014 08:05PM Directory 2014-05-05
05/08/2014 02:35PM Directory 2014-05-08
05/12/2014 04:46PM Directory 2014-05-12
05/20/2014 02:00PM Directory 2014-05-20
05/27/2014 03:12PM Directory 2014-05-27
05/27/2014 03:12PM 10 latest

After second month October 2013

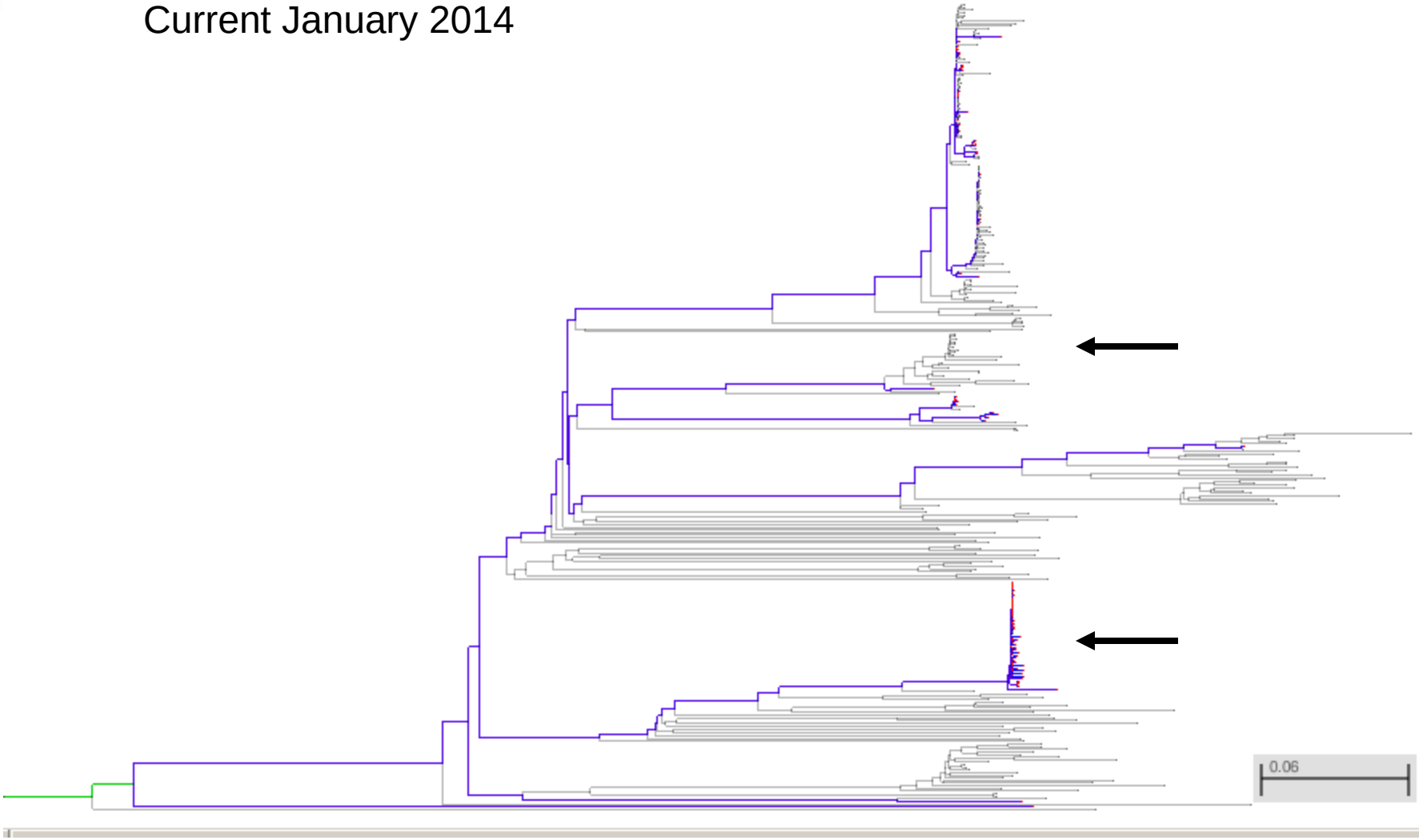
View Trees in NCBI's
Genome Work-Bench



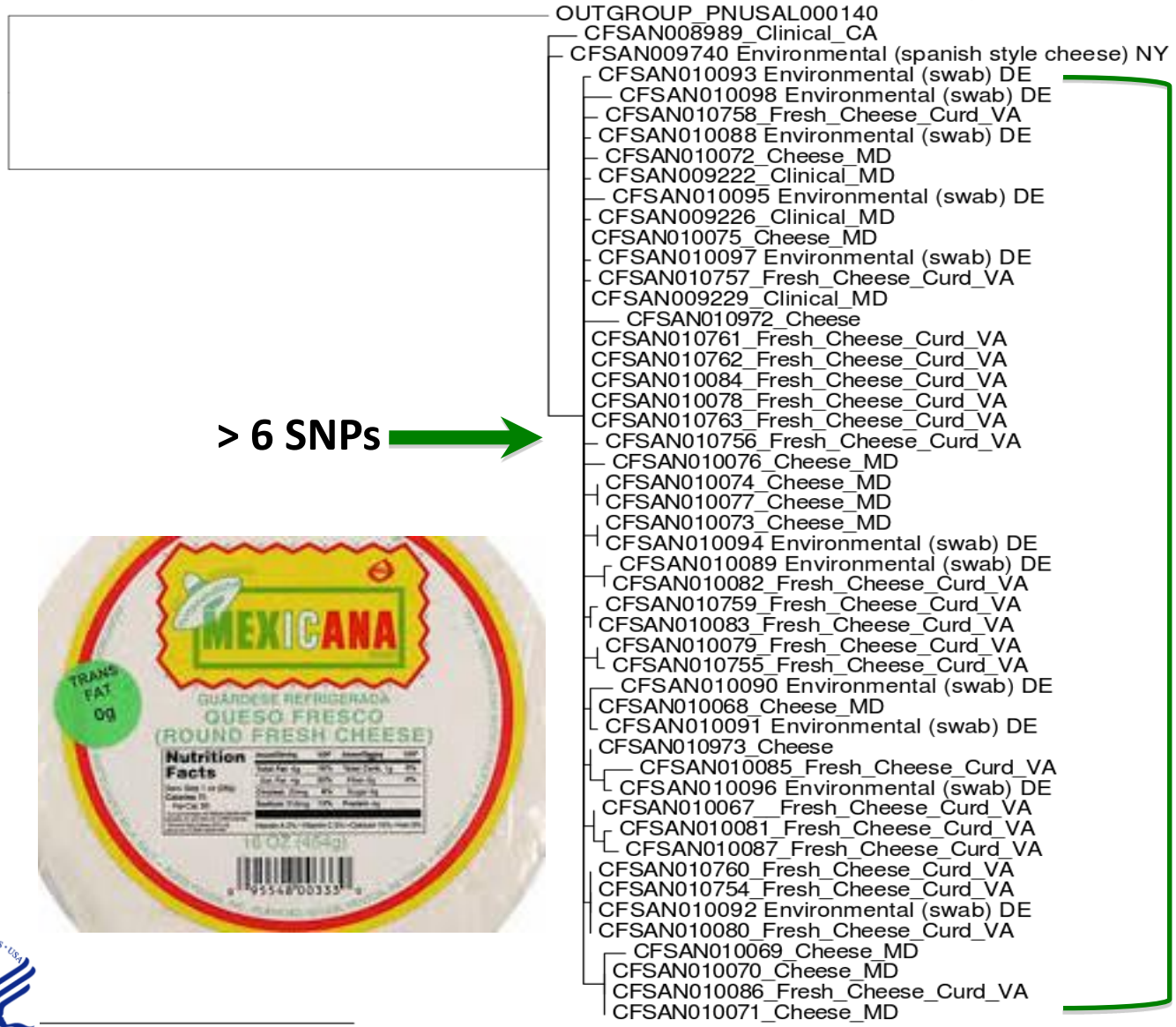
Third month November 2013



Current January 2014



Listeria monocytogenes and Support of FDA Compliance Standards



> 6 SNPs

Isolates from Roos facility, distributed product, and patients who consumed product

"Whole-genome sequencing (WGS) of *Listeria monocytogenes* strains isolated from Roos Foods cheese products has been performed by the FDA and Virginia's Division of Consolidated Laboratory Services. These strains were found to be highly related by WGS to the *Listeria* strains isolated from patients in this outbreak, adding further confidence that cheese products produced by Roos Foods were a likely source of the outbreak. WGS provides genetic information that allows investigators to rapidly identify differences among isolates. Compared with pulsed-field gel electrophoresis (PFGE), WGS provides clearer distinction between cases and foods that are likely part of a given outbreak and those that are not."

FDA
Center for Food
Safety and Applied Nutrition

Questions

Errol Strain: Errol.Strain@fda.hhs.gov

Marc Allard: Marc.Allard@fda.hhs.gov