American Public Health Laboratories
May 31 2014
General Meeting, Little Rock, AR

The Emerging Role of Whole-Genome Sequencing in FDA's Food Safety and Regulatory Science Program

Eric W. Brown, Ph.D.
Director
Division of Microbiology
Center for Food Safety & Applied Nutrition
U.S. Food & Drug Administration
College Park, Maryland 20740

Center for Food Safety and Applied Nutrition

Office of Regulatory Science

"The Smokin' Hot Pepper"

Los Angeles Times

FOOD SAFETY

Salmonella 'smoking gun' located

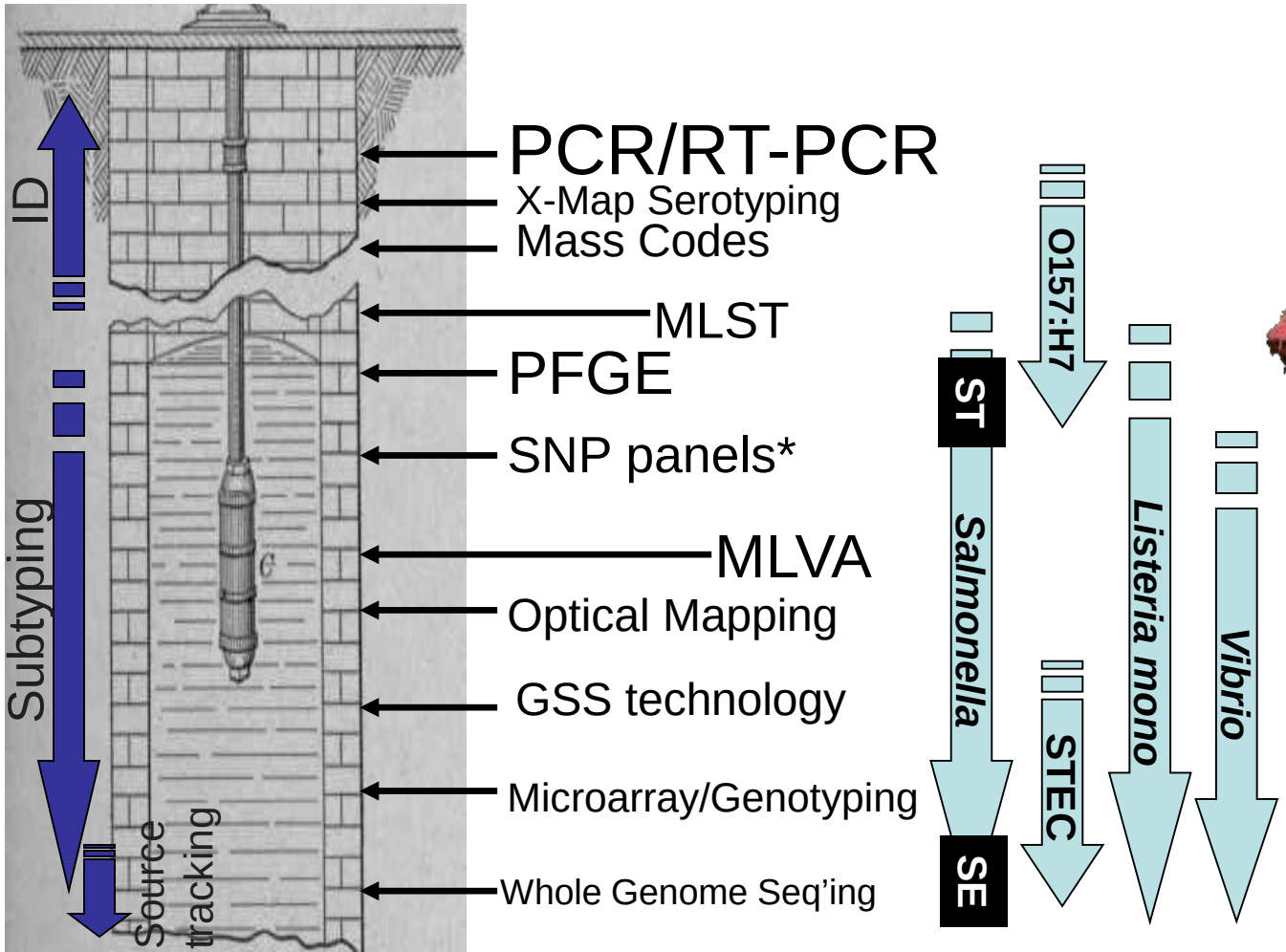
July 31, 2008 | Tiffany Hsu | Times Staff Writer

A strain of the salmonella bacteria that sickened more than 1,300 people has been found in a serrano pepper and a sample of irrigation water at a farm in Mexico, U.S. health officials said Wednesday.

They called the discovery a "breakthrough" but cautioned that tomatoes may still be a culprit in the nearly four-month outbreak that has alarmed consumers and cost the domestic produce industry hundreds of millions of dollars.

(Salmonella Saintpaul outbreak – Summer 2008)

Genomic triaging is now possible for surveillance, testing, and traceability of foodborne contamination



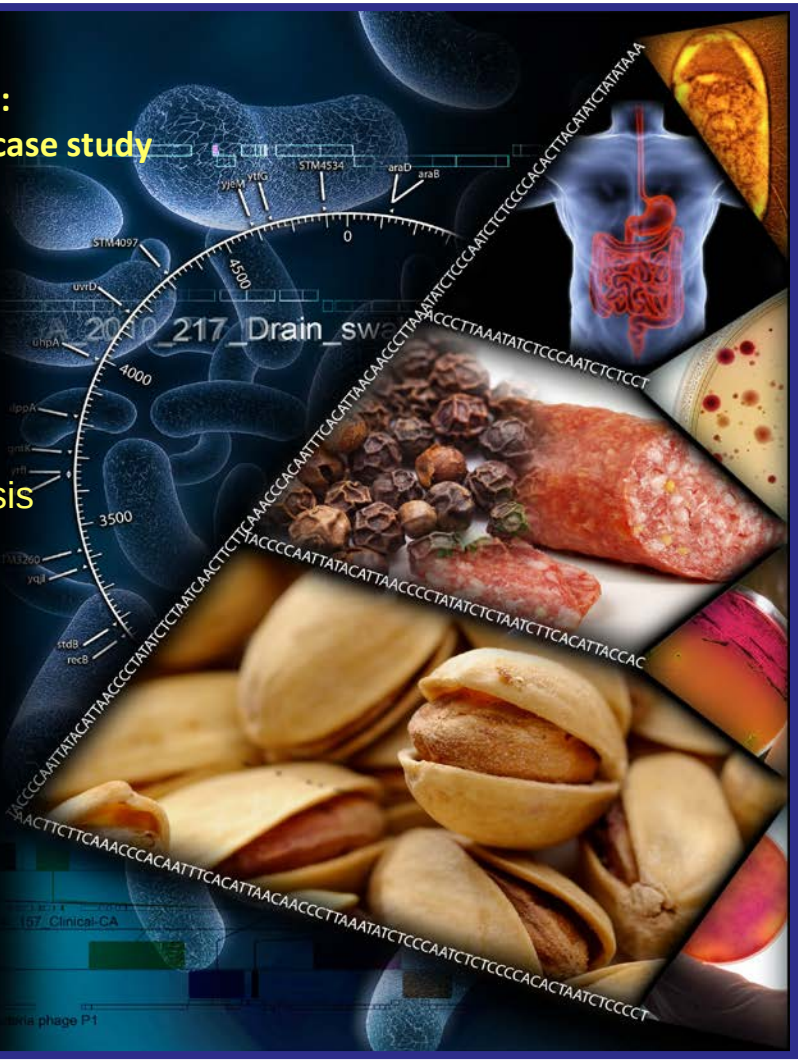
THE EARLY YEARS…

FOODBORNE OUTBREAK INVESTIGATION:
WGS analysis of foodborne salmonellae case study

This investigation focused on *Salmonella* Montevideo samples associated with red and black pepper used in the production of Italian-style spiced meats in a New England processing facility. This manufacturer was implicated in a major salmonellosis outbreak that affected more than 272 people in 44 states and the District of Columbia.

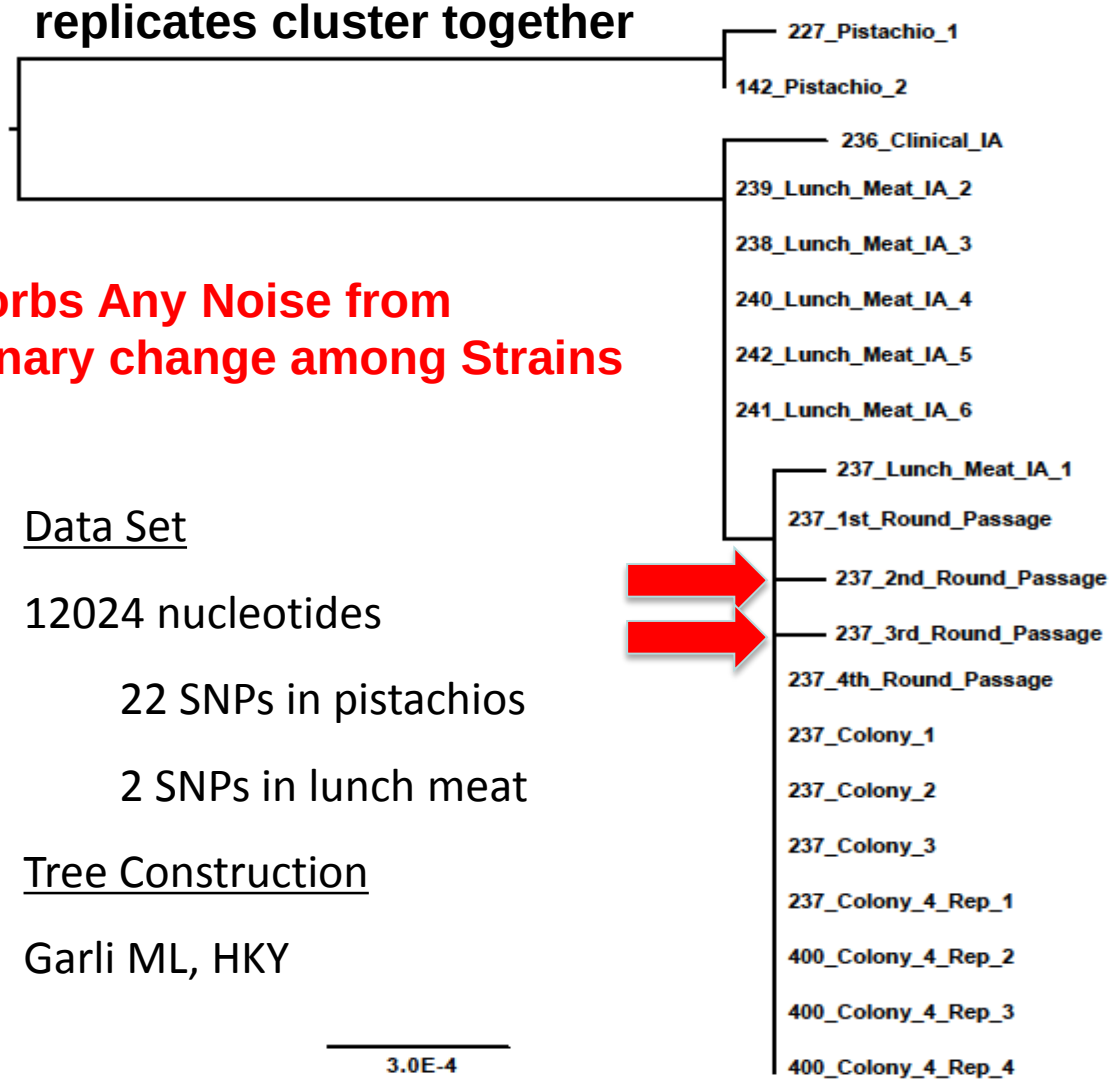
15-20x shot gun sequencing
35 pure culture isolates
from patients, foods and
Environmental samples.

Concatenate 40 variable genes for
Phylogenetic analysis



Allard et al. 2012
BMC Genomics

Sequence variation arising from lab/bioinformatics sources did not affect phylogenetic inference IA_2010008282 replicates cluster together



The Tree Absorbs Any Noise from Microevolutionary change among Strains

Data Set

12024 nucleotides

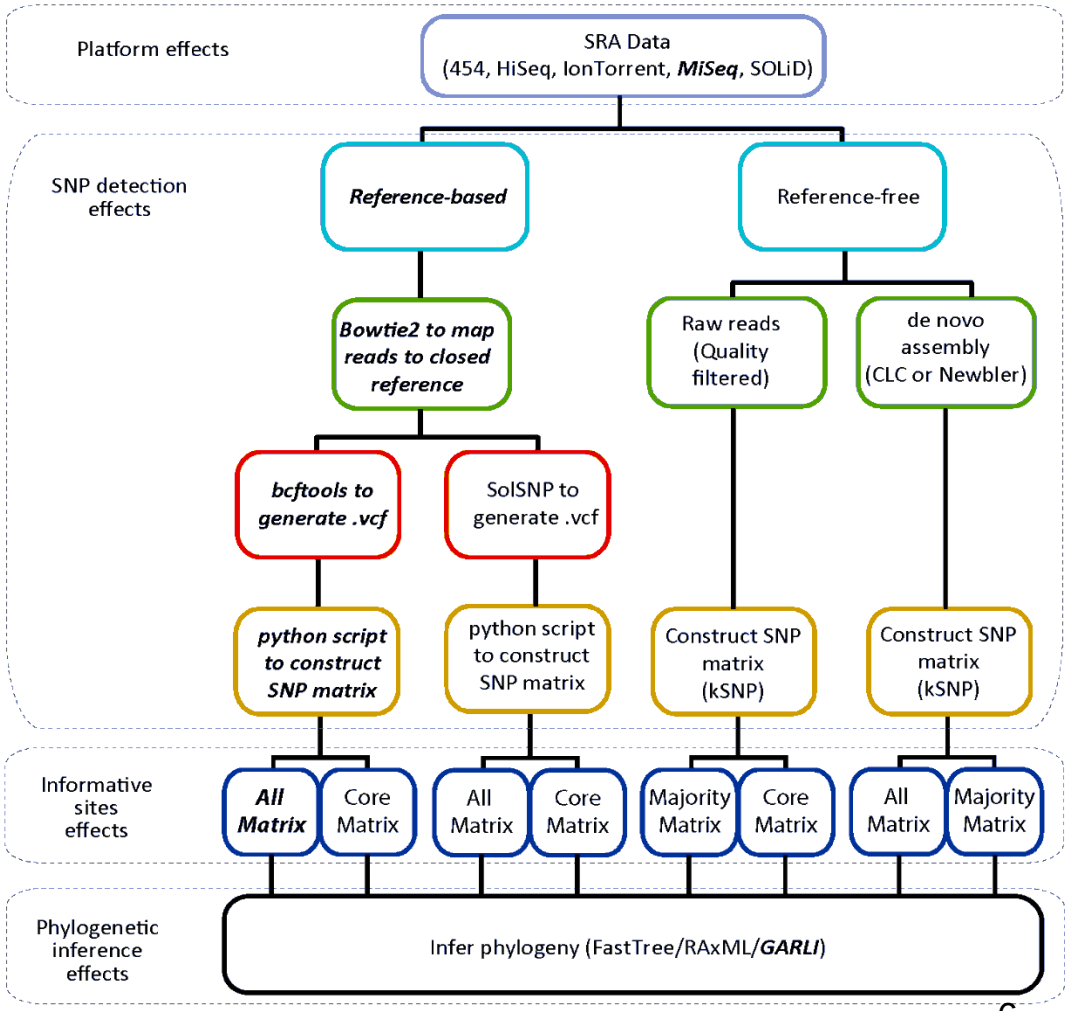
22 SNPs in pistachios

2 SNPs in lunch meat

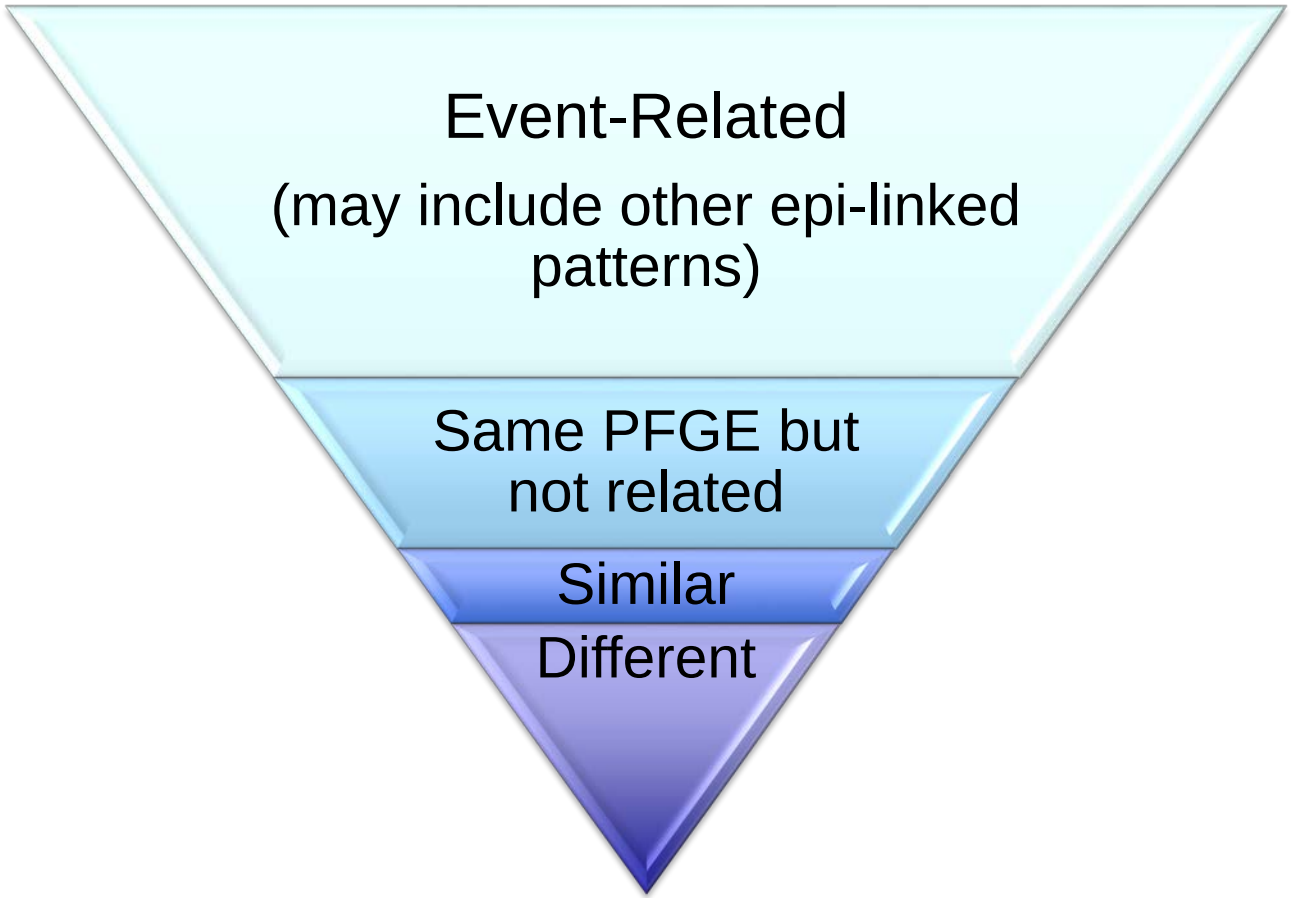
Tree Construction

Garli ML, HKY

Salmonella Montevideo



How we choose the isolates to include



NGS Analysis Strategy

FDA WGS Application to Food Contamination Events

Montevideo black and red pepper
Senftenberg black and red pepper
Enteritidis shell/liquid eggs
Heidelberg ground turkey
Heidelberg chicken broilers
Heidelberg chicken livers
Enteritidis custard
Bareilly tuna scrape
Tennessee peanut butter/peanut butter paste
Typhimurium peanut butter
Braenderup peanut butter/nut butter
Tennessee cilantro
Agona dry cereal
Agona papaya
Newport tomatoes
Newport environmental
Kentucky - Cerro dairy/dairy farms
Anatum spices/pepper flakes
Javiana cantaloupes
Saintpaul hot peppers
4,5,12: i –

Lmono cantaloupes
Lmono queso cheese
Lmono potato salad
Lmono artisanal cheeses
Lmono avocados
Lmono ricotta
Lmono celery/chix salad
Lmono smoked fish
Lmono other herbs

Cronobacter infant formula

V para oysters

EcO157:H7 lettuce

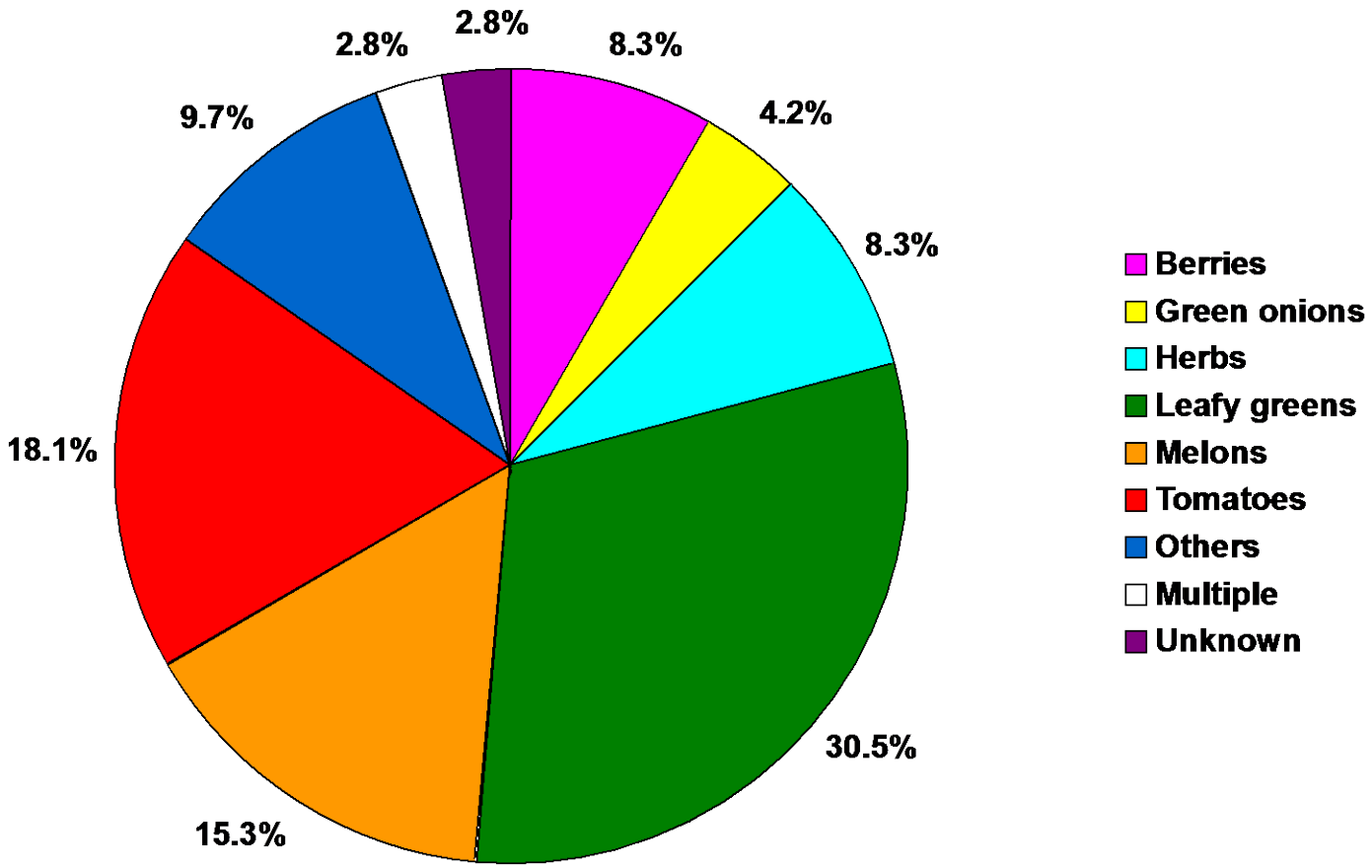
STEC beef

…Numerous other taxa

Applications of WGS in the Food Safety Environment

- Delimiting scope and traceback of food contamination events (Track-N-Trace)
- Quality control for FDA testing and surveillance (Confidence in Regulatory Actions)
- Preventive control monitoring for compliance standards (the “repeat offender” project)
- ID, geno/pheno typing schemes (AST,Serotyping, VP) (CVM,CDRH,CFSAN) – risk assessment and adaptive change in *Salmonella* and *Listeria*

The Impact of WGS in the Traceability of Food Contamination Events



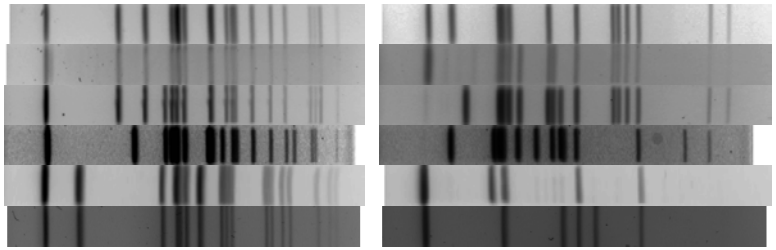
Prevalence of *Salmonella* in all Produce

- **Recent Outbreaks**

– 1998	*S.* Baildon	86 cases
– 2000	*S.* Thompson	29 cases
– 2002	*S.* Newport	512 cases
	S. Newport	12 cases
	S. Javiana	90 cases
– 2004	*S.* Javiana	471 cases
	S. Braenderup	123 cases
– 2005	*S.* Newport	71 cases
	S. Enteriditis	77 cases
	S. Braenderup	76 cases
– 2006	*S.* Newport	107 cases
	S. Typhimurium	186 cases
– 2008	*S.* Saintpaul	1,442 cases

- Widely dispersed, individual patient-cases in many states
- Low attack rates, epidemiology is tedious
- Tracebacks are difficult due to complexity of the supply chain
- Intermittent, low-level contamination
- Implicated produce is rarely still available, the crop is no longer in the field

Newport Outbreak: Tomatoes

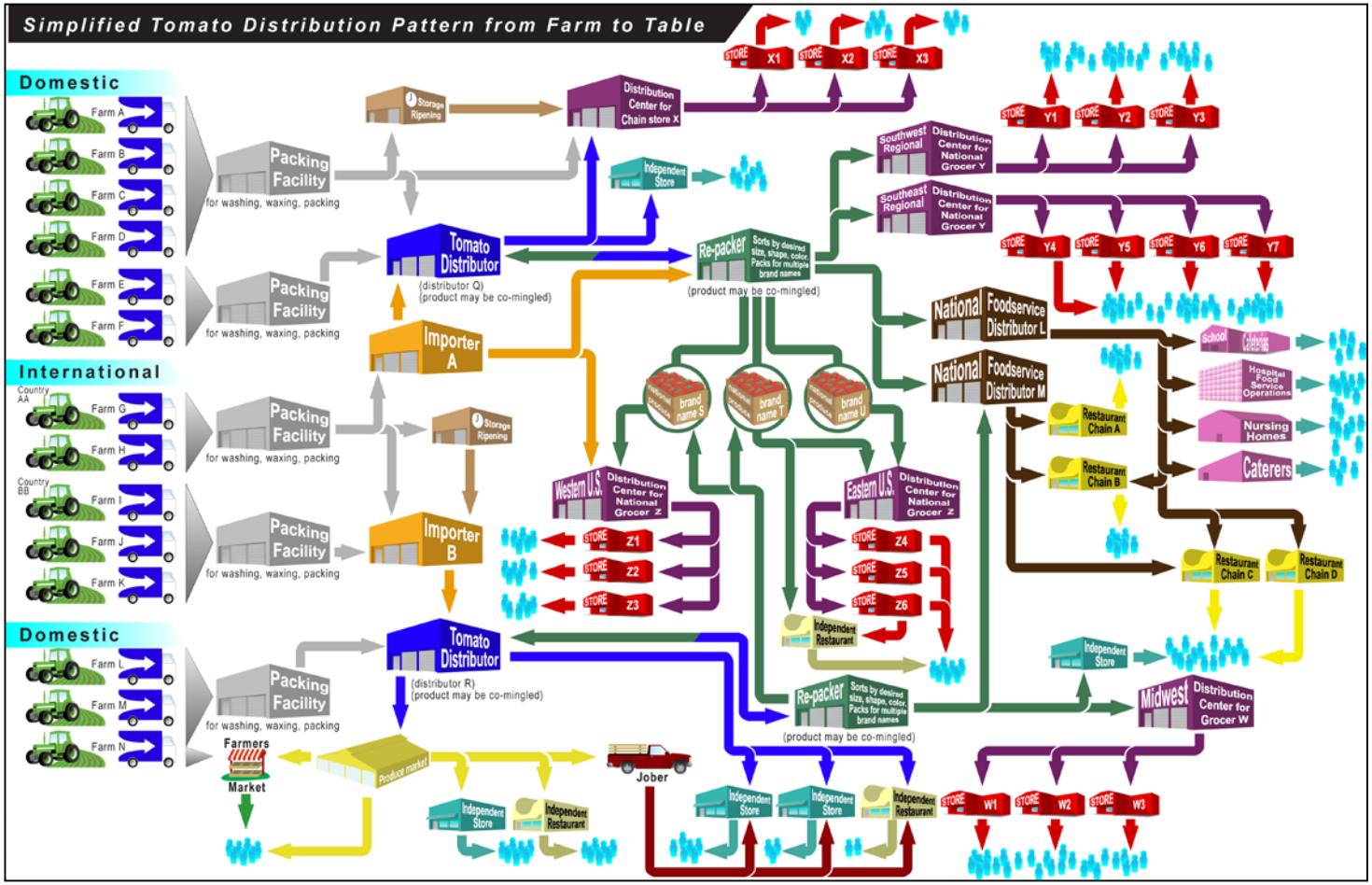


PFGE-XbaI-pattern	PFGE-BlnI-pattern
JJPX01.0011	JJPA26.0075
JJPX01.0011	JJPA26.0044
JJPX01.0030	JJPA26.0045
JJPX01.0061	JJPA26.0021
JJPX01.3101	JJPA26.0056
JJPX01.3101	JJPA26.0273

Note: There is some variation in 2nd enzyme pattern during clusters involving XbaI pattern .0011

The 2010 "DC Cluster"
(actually, part of 3 separate clusters)

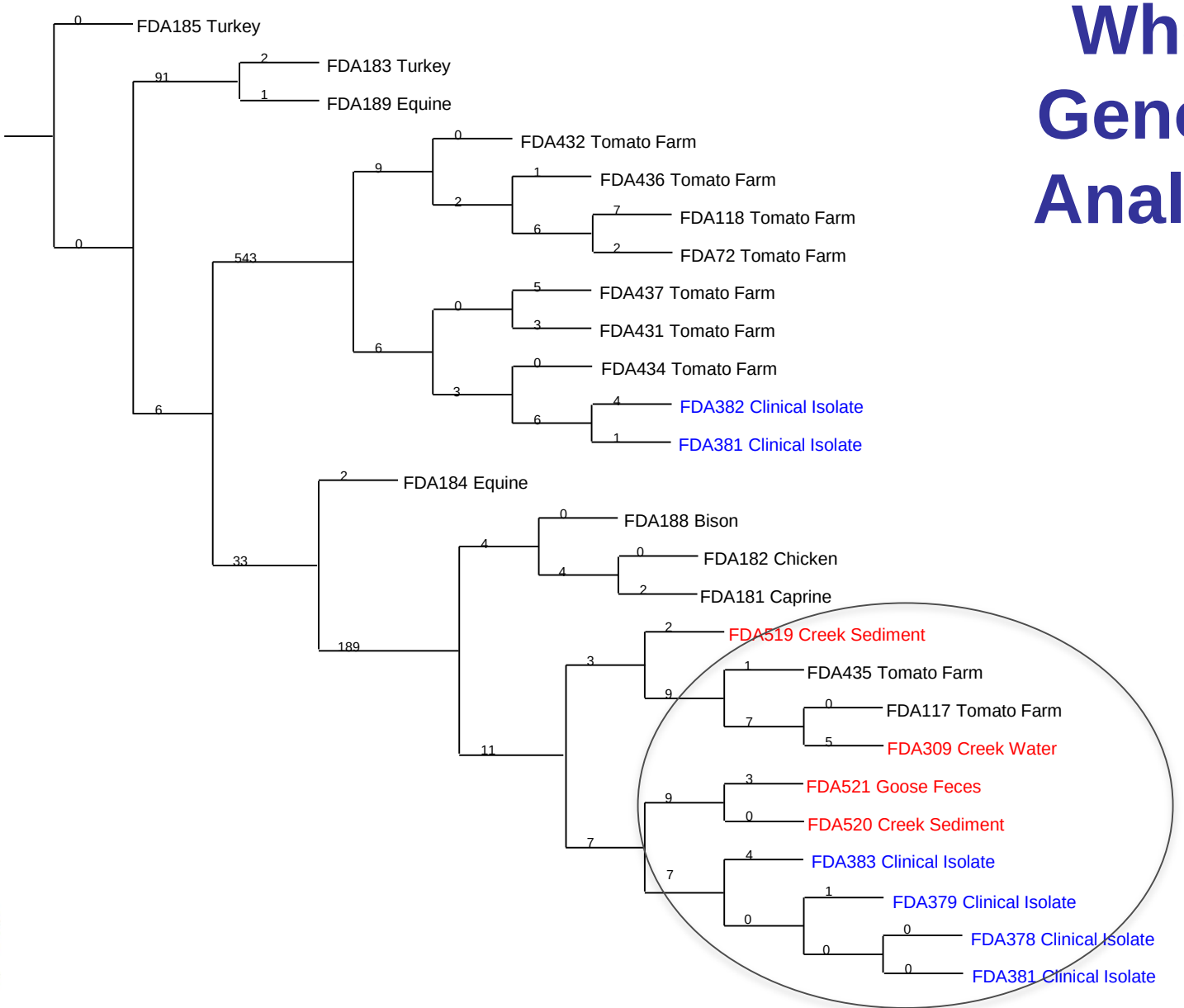
Once tomatoes reach the supply chain, things really "*simplify*".



The Fresh-cut Tomato Supply Chain

FDA's Team Tomato

Whole Genome Analysis



2010 Traceback from a Washington, DC Restaurant with 3 *S.* Newport Patterns

Packer E
CA
Invoice
#9812382

Re-packer D
NJ

PO# ab1203981239
PO# ab1209813487
PO# ab1298129812

Packer F
VA
Invoice #0230293

Farm M
VA

Common Farm: Farm M

Distributor B
MD

7/27/10 Invoice# 0404-3
7/29/10 Invoice# 02302

Restaurant A

11 illnesses

Patterns:
JJPX01.0061 (3 cases)
JJPX01.0011 (3 cases)
JJPX01.0030 (5 cases)

7/30/10 exposure

Packer F
VA

Lot #939 7/15/10
Lot# 0787 7/17/10
Lot# 232 7/22/10

Lot #4545 F #256934
Lot #5221 F #256975
Lot #1111 F #257182

Farm M
VA

Farm N
Farm O
VA

Matching research sample collected from creek adjacent to Farm M.

Pattern JJPX01.0011

Distributor/Re-packer C

Washington, DC

7/23/10 Lot# 4548
7/24/10 Lot# 4548, **2356**
7/26/10 Lot# 1132, 1191
7/27/10 **Lot# 1132**, 1169
7/28/10 Lot# 1555, **1132**
7/30/10 Lot# 1555

Packer G
TN
Supplied to C
Lot #4548 7/12/10

Packer K
AL
Supplied to C
Lot #1189
7/23/10

Packer L
AL
Supplied to C
Lot #1555 7/16/10

Packer H
CA
Supplied to C
Lot #1132 7/20/10

Salmonella Tennessee & Peanut Butter

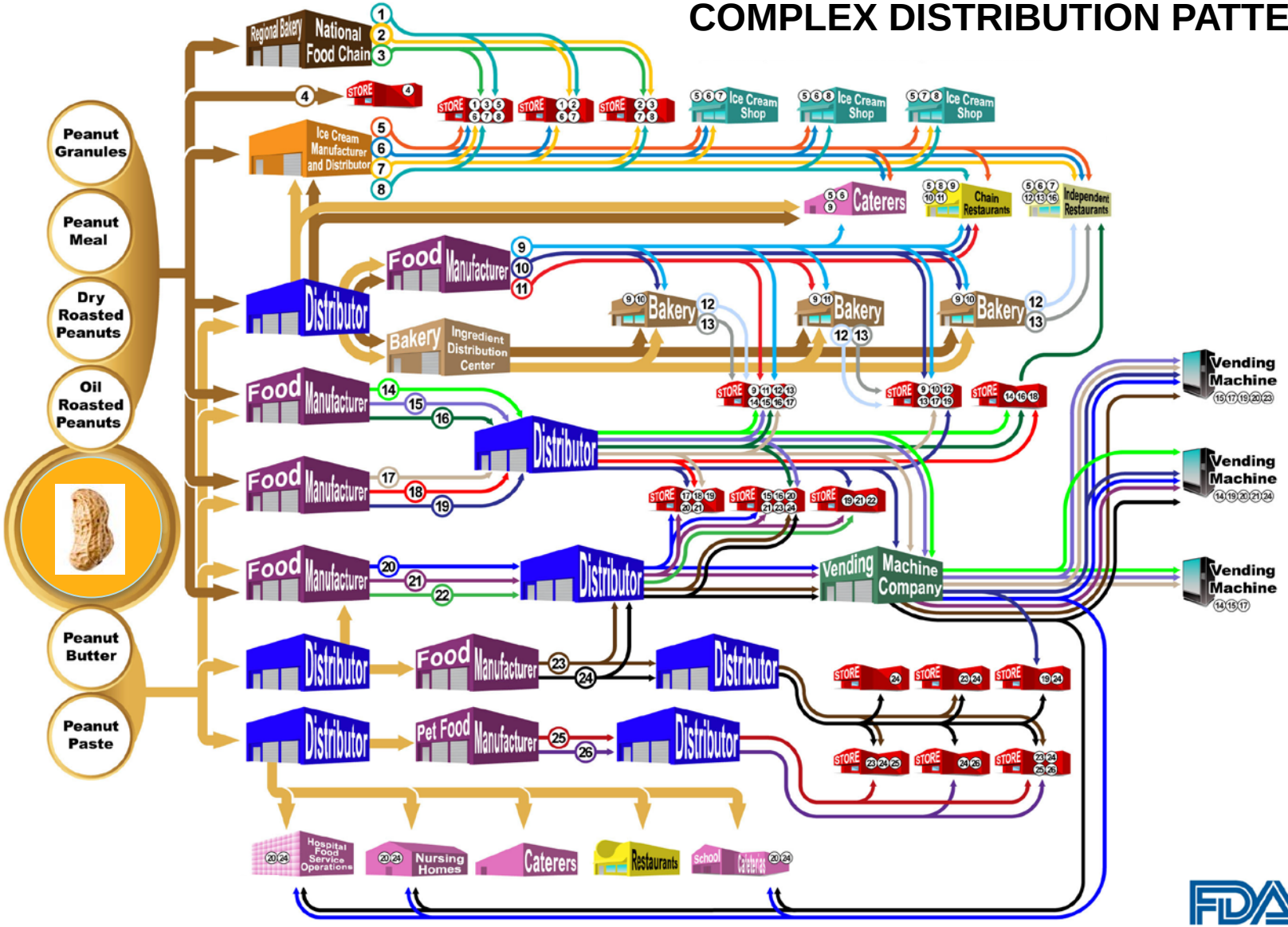
Comparative Genomic Analysis Of Peanut Butter-associated *Salmonella* Tennessee Provides Evidence For Assigning Contamination Sources Among Multiple And Contiguous Foodborne Outbreaks

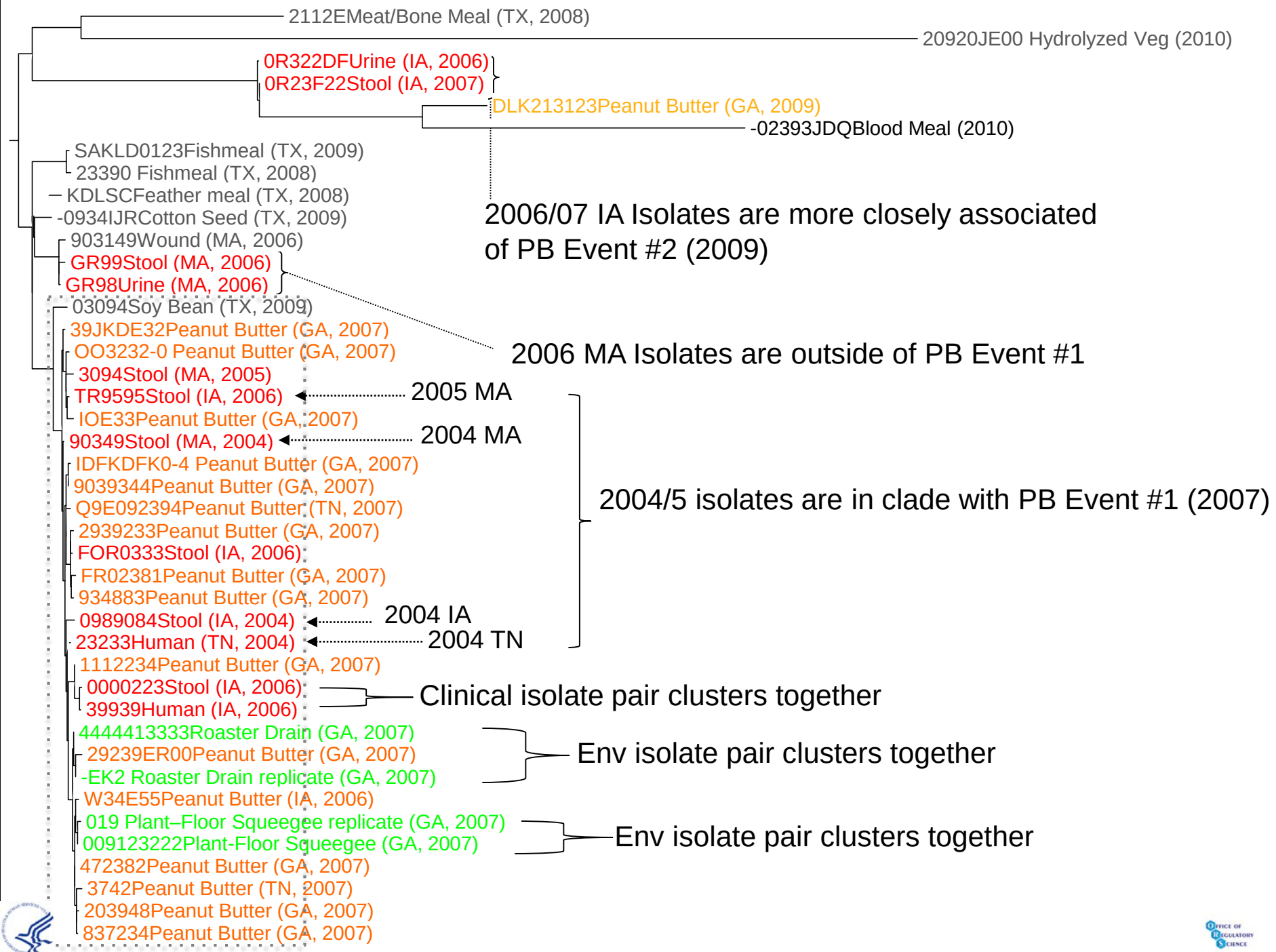
Authors: E. W. Brown[1], C. Keys[1], E. Strain[1], Y. Luo[1], C. Wang[1], T. Muruvanda[1], C. Pirone[1], S. Musser[1], M. Wilson[2], M. Allard[1];
[1]U.S. FDA, College Park, MD, [2]Forensic Sci. Program Western Carolina Univ., Cullowhee, NC

Poster Presentation Day/Date/Time: Monday May 19, 2014 10:45 AM - 12:00 PM
Poster Board Number: 1422

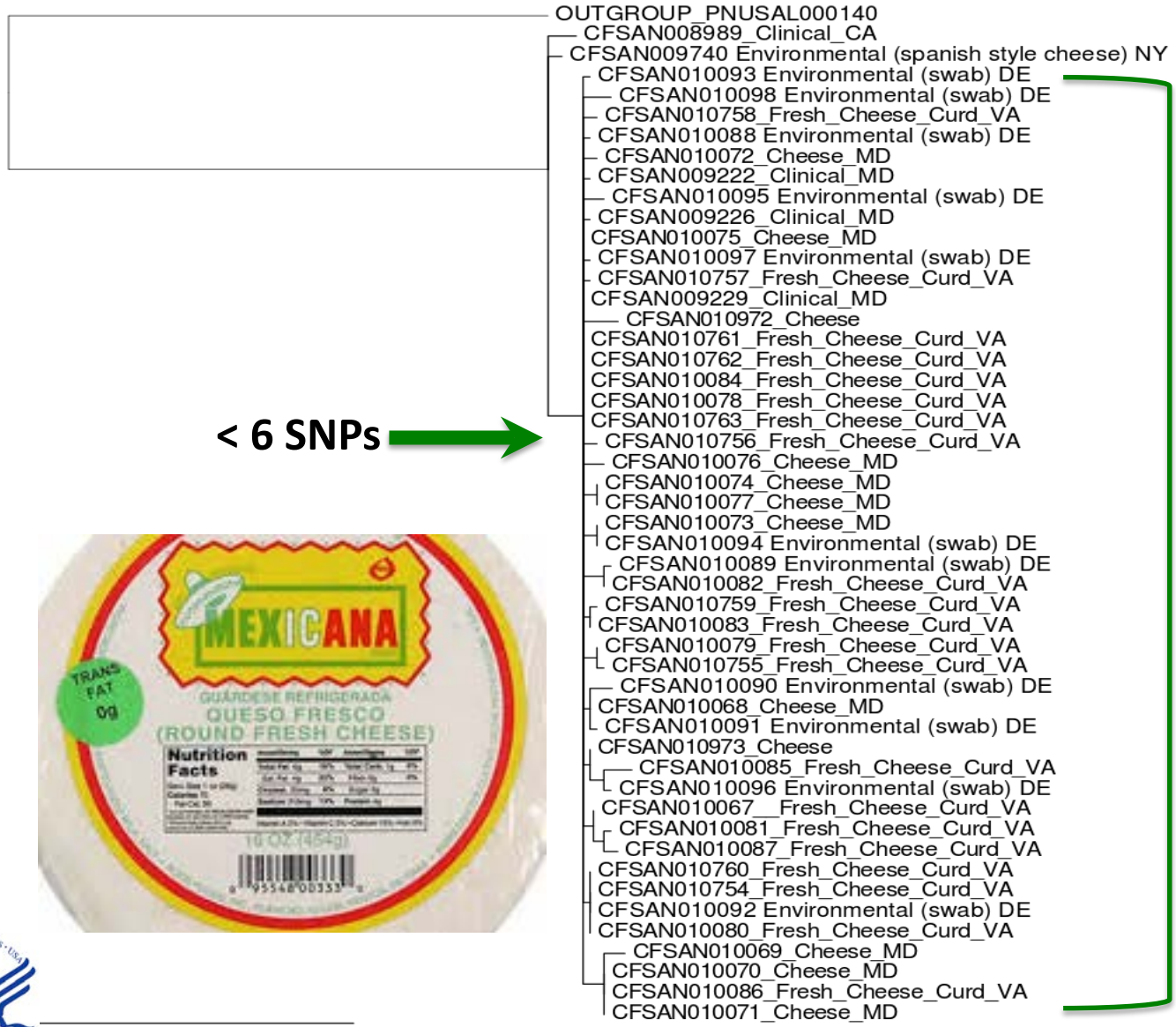
- **In 2007, nearly 630 persons infected with an outbreak strain of *Salmonella* Tennessee from 47 states.**
- **Strongly associated with consumption of peanut butter from single facility.**
- **Outbreak strain of *Salmonella* Tennessee isolated from several opened and unopened jars of peanut butter and from environmental samples obtained from the plant.**
- **In 2008-2009, a second outbreak was reported and also associated with peanut butter from an unrelated facility.**
- **Investigation determined that an *S*. Tennessee isolate detected during this second outbreak had a pulse-field gel electrophoresis (PFGE) pattern that was indistinguishable from the earlier 2007 *S*. Tennessee outbreak strains.**

COMPLEX DISTRIBUTION PATTERNS




2112EMeat/Bone Meal (TX, 2008)
20920JE00 Hydrolyzed Veg (2010)
0R322DFUrine (IA, 2006)
0R23F22Stool (IA, 2007)
DLK213123Peanut Butter (GA, 2009)
-02393JDQBlood Meal (2010)
SAKLD0123Fishmeal (TX, 2009)
23390 Fishmeal (TX, 2008)
KDLSCFeather meal (TX, 2008)
-0934IJRCotton Seed (TX, 2009)
903149Wound (MA, 2006)
GR99Stool (MA, 2006)
GR98Urine (MA, 2006)
03094Soy Bean (TX, 2009)
39JKDE32Peanut Butter (GA, 2007)
OO3232-0 Peanut Butter (GA, 2007)
3094Stool (MA, 2005)
TR9595Stool (IA, 2006)
IOE33Peanut Butter (GA, 2007)
90349Stool (MA, 2004)
IDFKDFK0-4 Peanut Butter (GA, 2007)
9039344Peanut Butter (GA, 2007)
Q9E092394Peanut Butter (TN, 2007)
2939233Peanut Butter (GA, 2007)
FOR0333Stool (IA, 2006)
FR02381Peanut Butter (GA, 2007)
934883Peanut Butter (GA, 2007)
0989084Stool (IA, 2004)
23233Human (TN, 2004)
1112234Peanut Butter (GA, 2007)
0000223Stool (IA, 2006)
39939Human (IA, 2006)
4444413333Roaster Drain (GA, 2007)
29239ER00Peanut Butter (GA, 2007)
-EK2 Roaster Drain replicate (GA, 2007)
W34E55Peanut Butter (IA, 2006)
019 Plant–Floor Squeegee replicate (GA, 2007)
009123222Plant-Floor Squeegee (GA, 2007)
472382Peanut Butter (GA, 2007)
3742Peanut Butter (TN, 2007)
203948Peanut Butter (GA, 2007)
837234Peanut Butter (GA, 2007)
2006/07 IA Isolates are more closely associated of PB Event #2 (2009)
2006 MA Isolates are outside of PB Event #1
2005 MA
2004 MA
2004/5 isolates are in clade with PB Event #1 (2007)
2004 IA
2004 TN
Clinical isolate pair clusters together
Env isolate pair clusters together
Env isolate pair clusters together

Listeria monocytogenes and Support of FDA Compliance Standards



< 6 SNPs

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

A parting perspective…

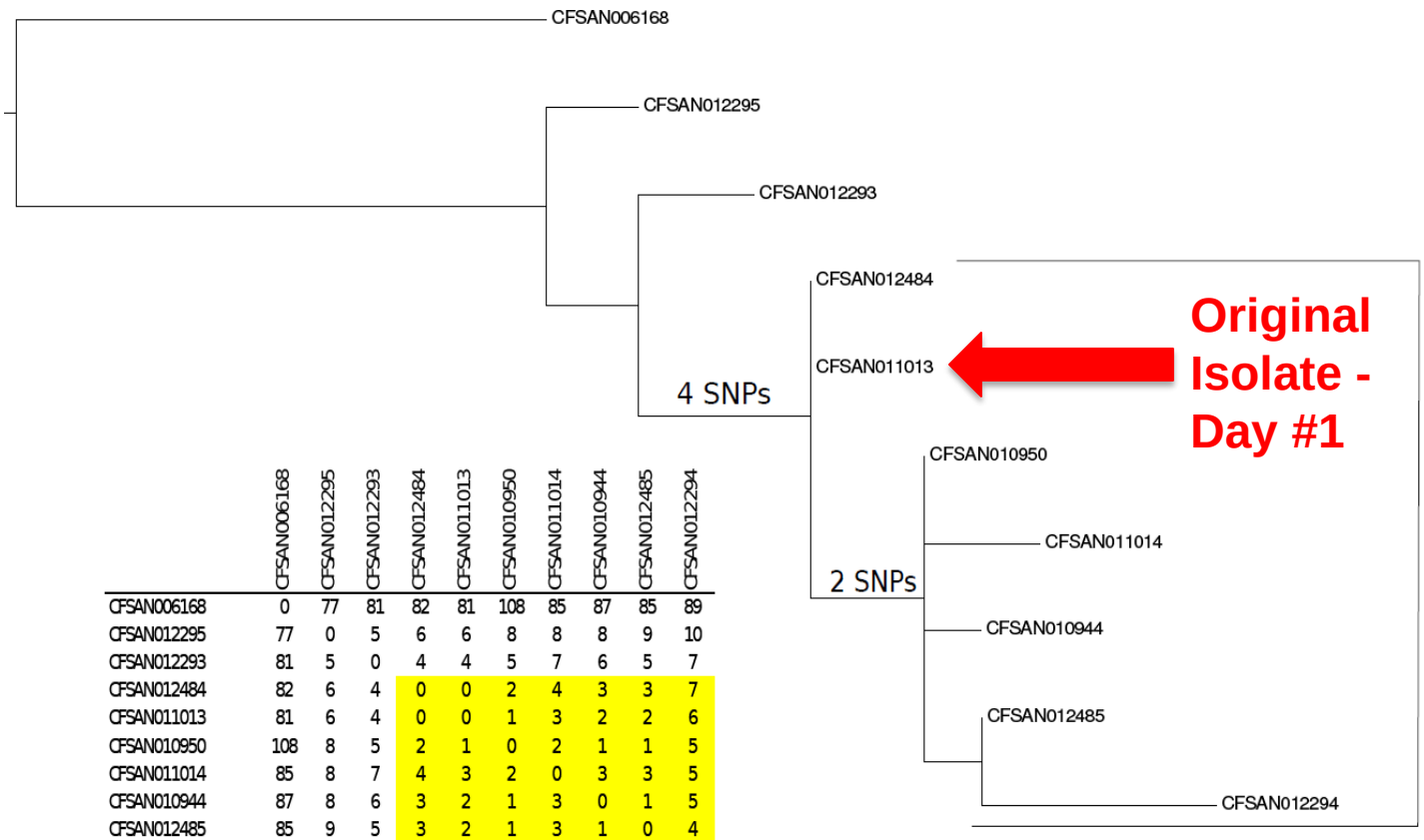
"The foodborne disease surveillance system is to the food industry what radar is to automobile drivers – It is the "threat" of being caught that helps drive compliance with best safety practices."

from Mike Doyle, Ph.D;
CDC Public Health Grand Rounds November 2009

Whole Genome Sequencing = one very powerful radar gun

The Role of WGS in Quality Assurance of Microbiological Sampling

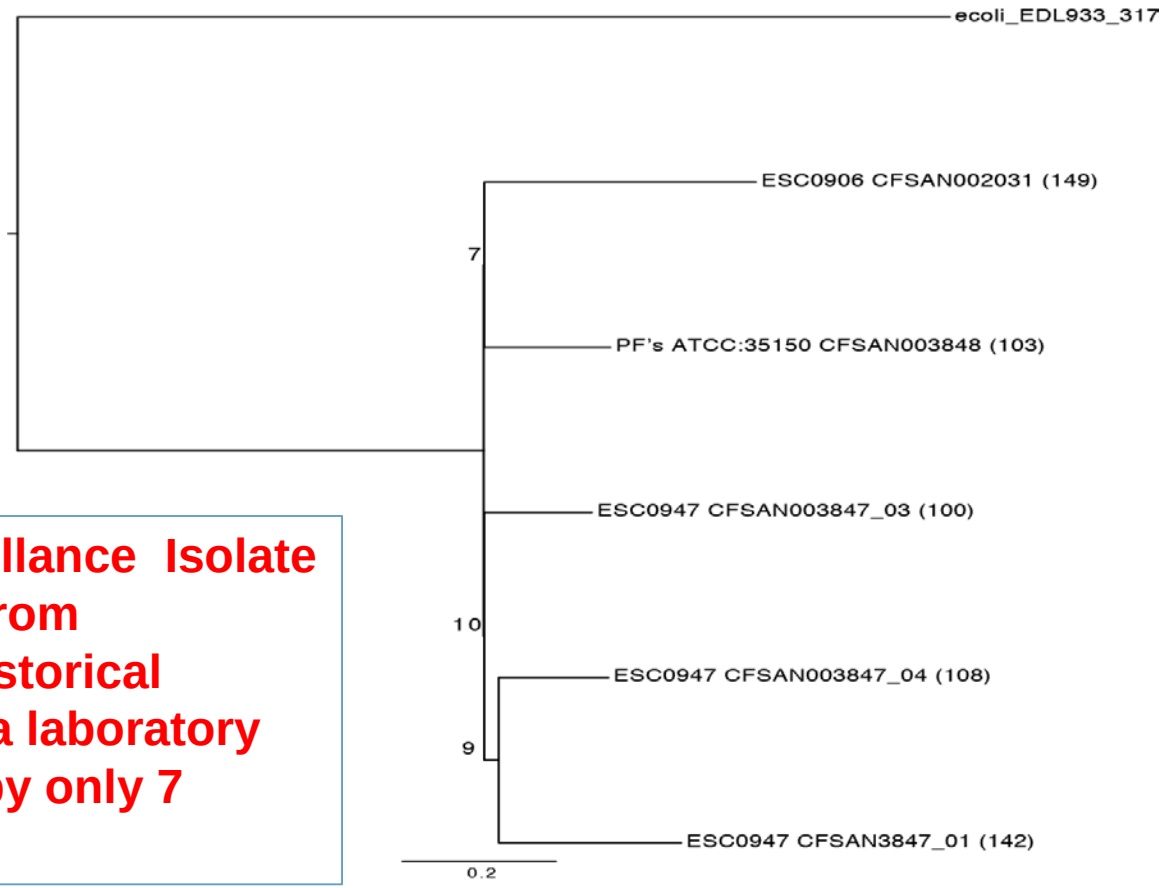
Quality Checks for Multiple Isolations from a Single Laboratory – *Salmonella enterica*



	CFSAN006168	CFSAN012295	CFSAN012293	CFSAN012484	CFSAN011013	CFSAN010950	CFSAN011014	CFSAN010944	CFSAN012485	CFSAN012294
CFSAN006168	0	77	81	82	81	108	85	87	85	89
CFSAN012295	77	0	5	6	6	8	8	8	9	10
CFSAN012293	81	5	0	4	4	5	7	6	5	7
CFSAN012484	82	6	4	0	0	2	4	3	3	7
CFSAN011013	81	6	4	0	0	1	3	2	2	6
CFSAN010950	108	8	5	2	1	0	2	1	1	5
CFSAN011014	85	8	7	4	3	2	0	3	3	5
CFSAN010944	87	8	6	3	2	1	3	0	1	5
CFSAN012485	85	9	5	3	2	1	3	1	0	4
CFSAN012294	89	10	7	7	6	5	5	5	4	0

Multiple Isolates From Different Food Sources Over a 60-day Period diverge by only 0 to 6 SNPs

Quality Checks for Isolations Matching Laboratory Control Standards – *E. coli* O157:H7



Test Isolate

C

C

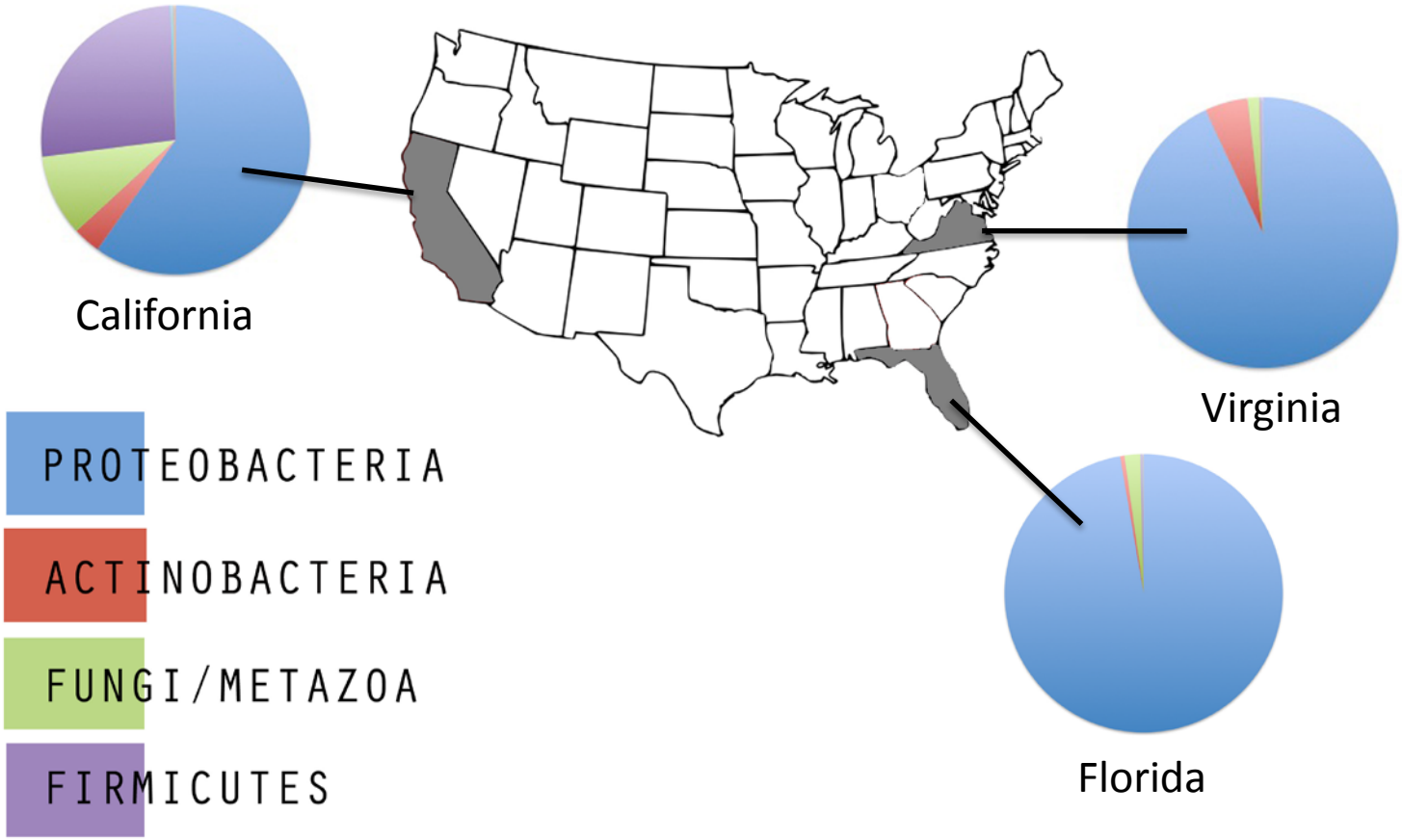
C

C

One surveillance Isolate Diverged from Multiple historical Copies of a laboratory Standard by only 7 SNPs

Ecology and Adaptation: The emergence and persistence of Salmonella across the food supply

Biogeographical Metagenomic Surveys of Tomato Phyllosphere Microflora



Summary of statistically significant abundance of bacterial genera in VA and CA

CA

- *Weissella*
- *Lactococcus*
- *Leuconostoc*
- *Bacillus*
- *Enterobacteriaceae, other*

VA

- *Pseudomonas*
- *Sphingomonas*
- *Xanthomonas*
- *Agrobacterium*
- *Hymenobacter*

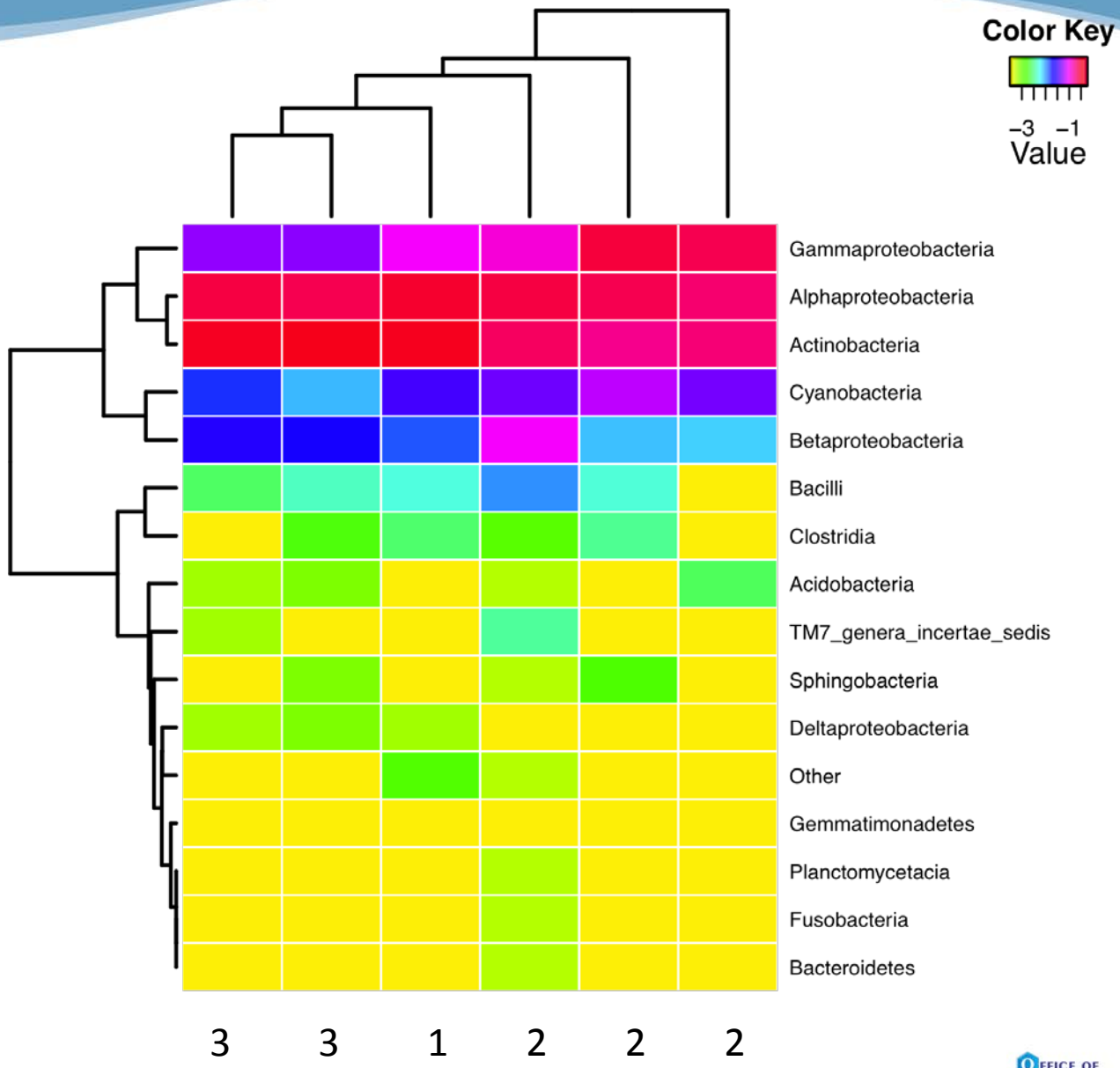
- Strong antibiotic activity associated with *Weissella, Lactococcus, Leuconostoc*
- Niche competition by Enterobacteriaceae, other ?
- Habitat preference created by *Pseudomonas, Sphingomonas*, *Xanthomonas* etc.

Agricultural Practices

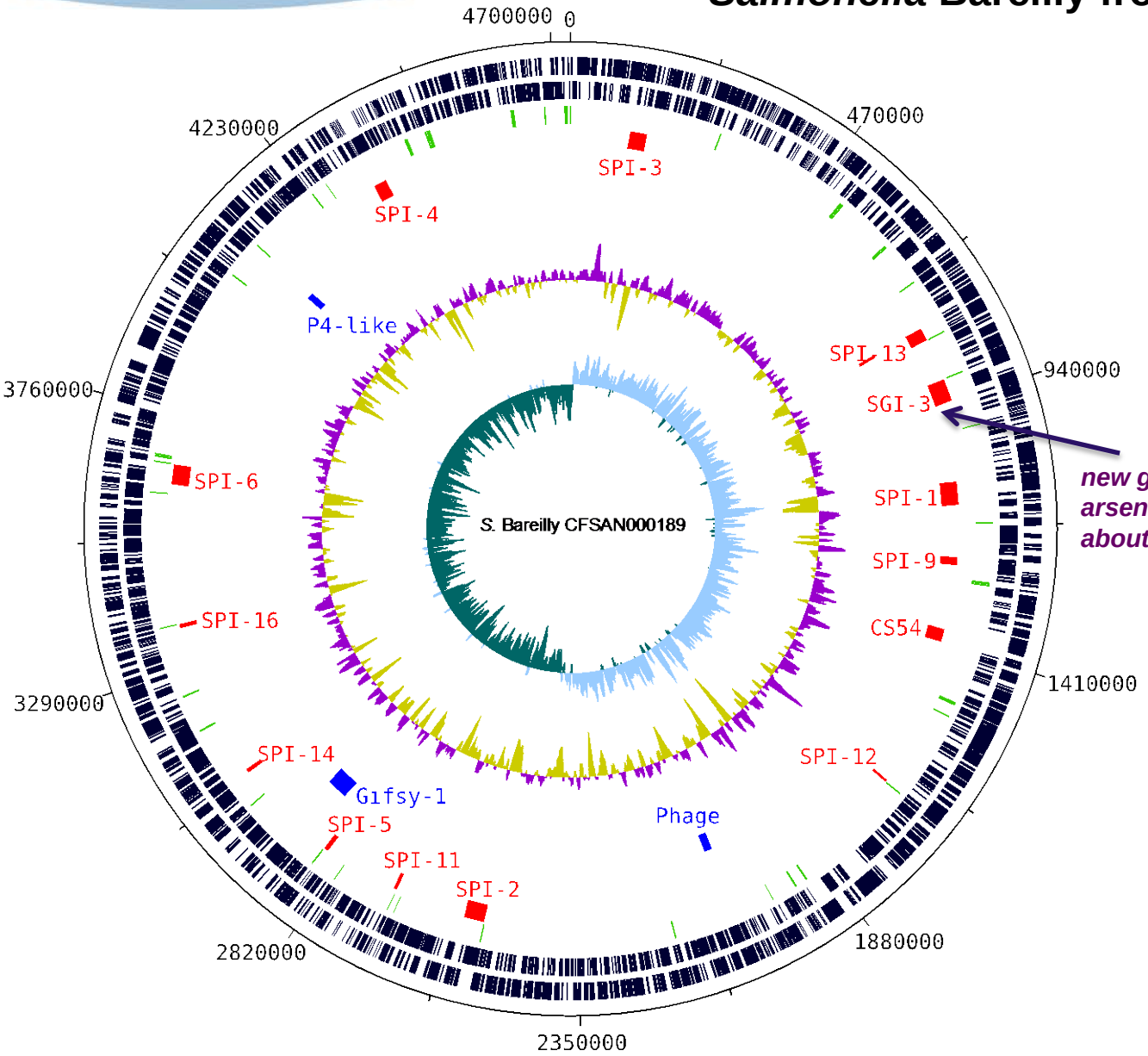
Fungicide Treatment – Tomatoes
Methyl Bromide Treatments – Tomatoes

Water Sources used for Irrigation and Pesticide Applications – Tomatoes

Fungicide Trials
- 1 Control
- 2 Actigard
- 3 Kocide



Salmonella Bareilly from Tuna



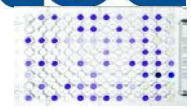
new genomic island arsenic resistance operon about 40 kb

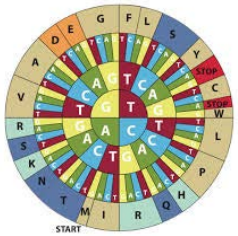
New Genomic Island

Name	Minimum	Maximum	Length	Direction	product
toxin CDS	43,465	43,788	324	forward	toxin
antitoxin CDS	43,124	43,444	321	forward	antitoxin
transcriptional regulator CDS	38,480	39,184	705	forward	transcriptional regulator
fatty acid transporter CDS	36,027	37,214	1,188	forward	fatty acid transporter
mechanosensitive ion channel protein MscS CDS	34,410	35,996	1,587	forward	mechanosensitive ion channel protein MscS
alkyl sulfatase CDS	32,360	34,267	1,908	forward	alkyl sulfatase
histidine phosphatase CDS	30,833	31,447	615	reverse	histidine phosphatase
transcriptional regulator CDS	28,266	28,532	267	reverse	transcriptional regulator
magnesium transporter CDS	24,110	24,766	657	forward	magnesium transporter
fimbrial usher protein CDS	20,181	22,691	2,511	reverse	fimbrial usher protein
fimbrial chaperone protein StdC CDS	19,404	20,129	726	reverse	fimbrial chaperone protein StdC
Positive regulator GrlA CDS	17,713	18,189	477	reverse	Positive regulator GrlA
transcriptional regulator CDS	16,860	17,720	861	reverse	transcriptional regulator
membrane protein CDS	14,535	15,275	741	reverse	membrane protein
ArsR family transcriptional regulator CDS	13,772	14,101	330	reverse	ArsR family transcriptional regulator
NADPH-dependent FMN reductase CDS	13,057	13,770	714	reverse	NADPH-dependent FMN reductase
RNA polymerase sigma 70 CDS	12,503	13,048	546	reverse	RNA polymerase sigma 70
arsenic resistance operon repressor CDS	12,065	12,427	363	reverse	arsenic resistance operon repressor
arsenic transporter ATPase CDS	10,287	12,044	1,758	reverse	arsenic transporter ATPase
ModE family transcriptional regulator CDS	9,845	10,213	369	forward	ModE family transcriptional regulator
arsenate reductase CDS	8,388	8,819	432	forward	arsenate reductase
arylsulfatase CDS	7,086	8,375	1,290	forward	arylsulfatase
arsenic transporter ATPase CDS	5,287	7,038	1,752	forward	arsenic transporter ATPase
arsenic resistance operon repressor CDS	4,907	5,269	363	forward	arsenic resistance operon repressor
arsenic resistance operon repressor CDS	4,506	4,859	354	forward	arsenic resistance operon repressor
nucleotidyltransferase CDS	2,538	4,313	1,776	forward	nucleotidyltransferase
integrase CDS	1	1,194	1,194	forward	integrase

Functional Assays for SNPs

BIOLOG





Better understanding of adaptive change in *Salmonella* and *Lm* may provide more accurate risk assessment as well as enhanced preventive control measures on the farm and in the processing plant.

Summary…

- WGS is revolutionizing the laboratory contribution to public health microbiology, outbreak investigation - food safety included.
- WGS can be used to mitigate tracebacks and delimit the scope of food contamination events unlike ever before – numerous offshoot applications exist (i.e., compliance, quality assurance, risk assessment)
- The development of international open source databases will empower WGS for sentinel surveillance work on a global scale
- We are continuing to deploy the technology for applied food safety investigations and are continuing to develop metagenomic and next next gen applications for identification of foodborne pathogens.
- Genome sequences are agnostic, portable, and instantly cross-compatible. One biological assay could work on all pathogen species. To be immediately useful all we need is the genome and a little metadata.

Acknowledgments

CFSAN Steven Musser

Division of Microbiology-FDA

Marc Allard Charles Wang Peter Evans
Jie Zheng George Kastanis Yan Luo
Chris Keys Tim Muravunda Ruth Timme

Becky Bell Justin Payne Thomas Hammack Dave Melka

Division of Biostatistics-FDA

Errol Strain Yan Luo James Pettengill

FERA-UK

Robert Stones Alisdair Wotherspoon

OARSA-FDA

Mark Mammel Darcy Hanes Chris Grim Karen Jarvis Laura Ewing

Joint Institute for Food Safety & Applied Nutrition

Guojie Cao Jianghing Meng

National Institutes of health

David Lipman
Jim Ostell
William Klimke

Center for Veterinary Medicine-FDA

Shaohua Zhao Pat McDermott Daniel Tadesse

Office of Regulatory Science-FDA

Kelly Bunning John Callahan

FDA Office of Foods Dave White Palmer Orlandi

CDC and FSIS

John Besser Eija Trees Patti Fields Stephanie Defibaugh & oth

International Association for Food Protection Meeting, August 3-6, Indianapolis, IN

New Professional Development Group

Advanced Molecular Detection Analytics Organizational Meeting

August 3rd, 1-3:00 PM, Room 116

To provide a forum for the exchange and sharing of information related to the development and use of advanced molecular approaches for the detection and identification of microbial contaminants of food and related commodities