



United States Department of Agriculture

One Team, One Purpose



Food Safety and Inspection Service

Protecting Public Health and Preventing Foodborne Illness





Food Safety and Inspection Service:

Application of WGS Data in Examining *Listeria monocytogenes* Harborage in Food Processing Environments

Glenn Tillman, Ph.D.

USDA-FSIS Office of Public Health

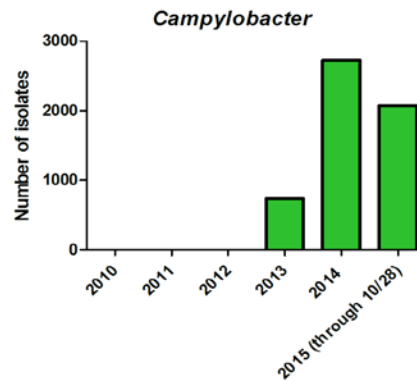
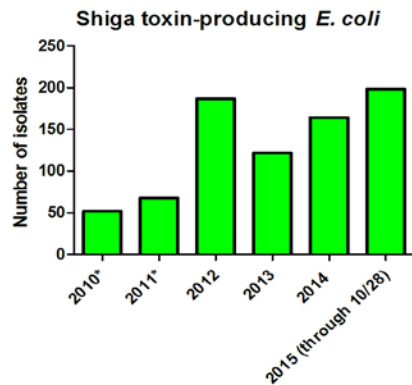
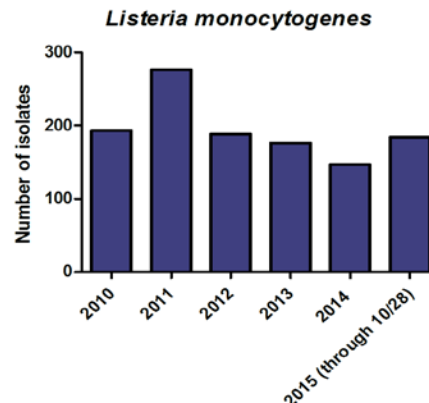
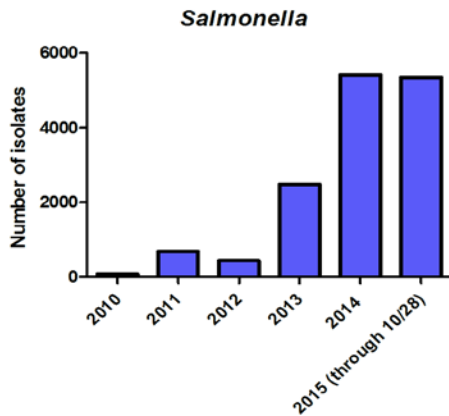
Athens, GA



Food Safety and Inspection Service: FSIS Mission and Sampling Programs

- FSIS is the public health agency in the U.S. Department of Agriculture responsible for ensuring that the nation's commercial supply of meat, poultry, and processed egg products is safe, wholesome, and correctly labeled and packaged
- Regulates more than 6,000 slaughter and processing establishments nationwide
- Verifies safety of approximately 100 billion pounds of product annually

Food Safety and Inspection Service: FSIS Mission and Sampling Programs



- FSIS Field Service Laboratories, State laboratory partners, and contract laboratories screen regulatory, baseline, and investigation samples
- **Isolates are submitted to Outbreaks Section Eastern Laboratory (OSEL) for further characterization**



Food Safety and Inspection Service:

Why is FSIS becoming involved in WGS?

- **Supports FSIS mission goals**
 - Effectively use science to understand foodborne illness and emerging trends
- **Improved resolution for epidemiology investigation and foodborne illness investigations**
 - Improved strain discrimination, cluster detection, and case classification
- **Method alignment with our Public Health partners**
 - Adherence to a CDC-PulseNet model



Food Safety and Inspection Service: FSIS WGS status and progress

- ***Listeria monocytogenes* real-time surveillance project**
 - 249 *Lm* isolates sent to FDA CFSAN from August 2013- March 2015
 - 126 *Lm* isolates sequenced by FSIS since April 2015
- ***Salmonella* isolates sequenced**
 - 393 *Salmonella* Egg Baseline isolates in 2015 (FDA CFSAN)
 - 348 isolates uploaded as part of investigations
 - 94 NARMS cecal isolates
- **Shiga toxin-producing *Escherichia coli***
 - 158 STEC from surveillance or baseline projects
- ***Campylobacter* spp.**
 - 38 from surveillance or baseline projects
 - 108 NARMS cecal isolates

Food Safety and Inspection Service: *Lm* Harborage Project – Purpose

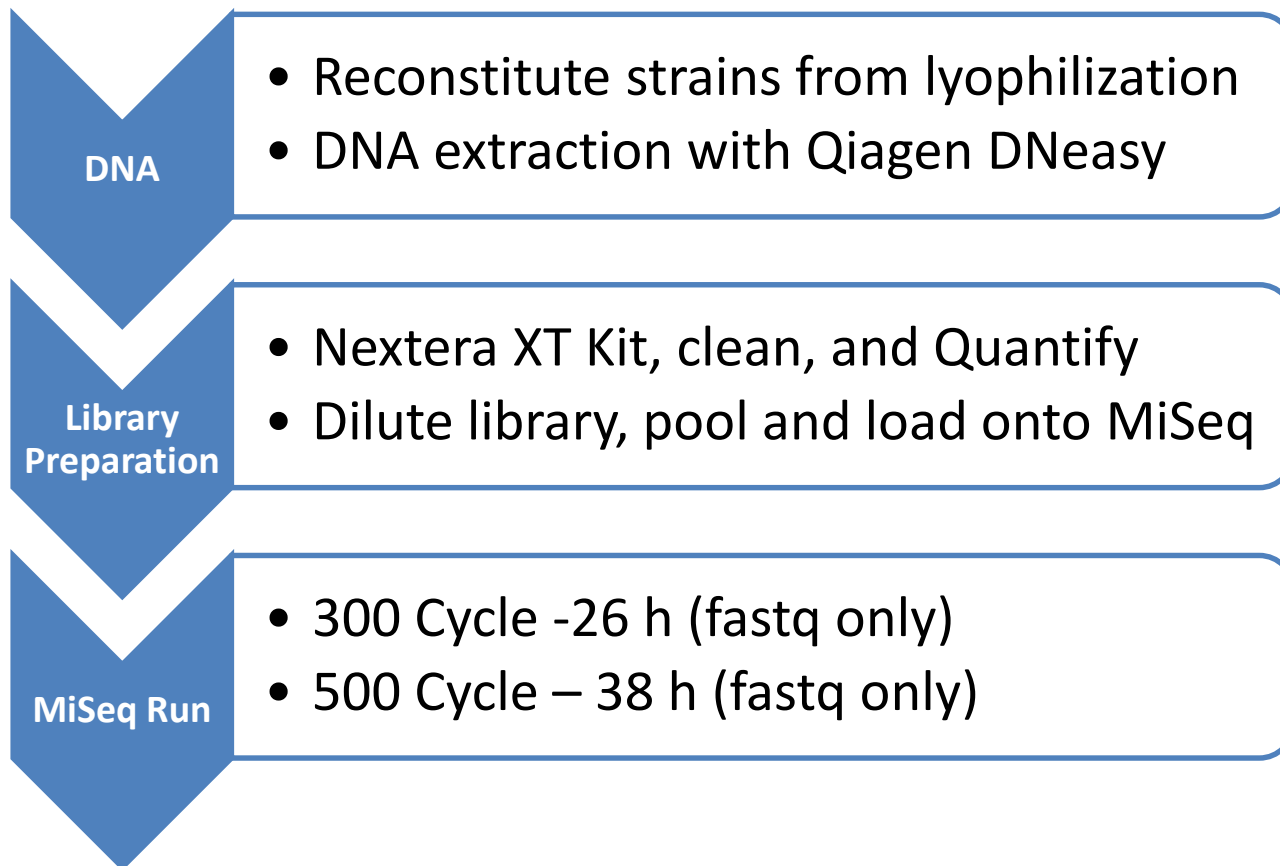
- Examine questions of importance to FSIS
 - New tools to look at an old problem (*Lm* strain harborage)
- Advance FSIS knowledge of comparing clusters using PFGE data vs. WGS data
- Advance FSIS knowledge in evaluating harborage of *Lm* strains in a food processing environment
- Integrate existing bioinformatics tools to answer questions about inferred phylogeny

Food Safety and Inspection Service: *Lm* Harborage Project – Isolate Selection

- Datasets consisted of groups of *Lm* isolates from same establishment
- Establishments chosen had:
 - At least 10 *Lm* isolates (both food and environmental)
 - Isolates collected over multiple inspection/collection events
 - Isolates with PFGE pattern information available

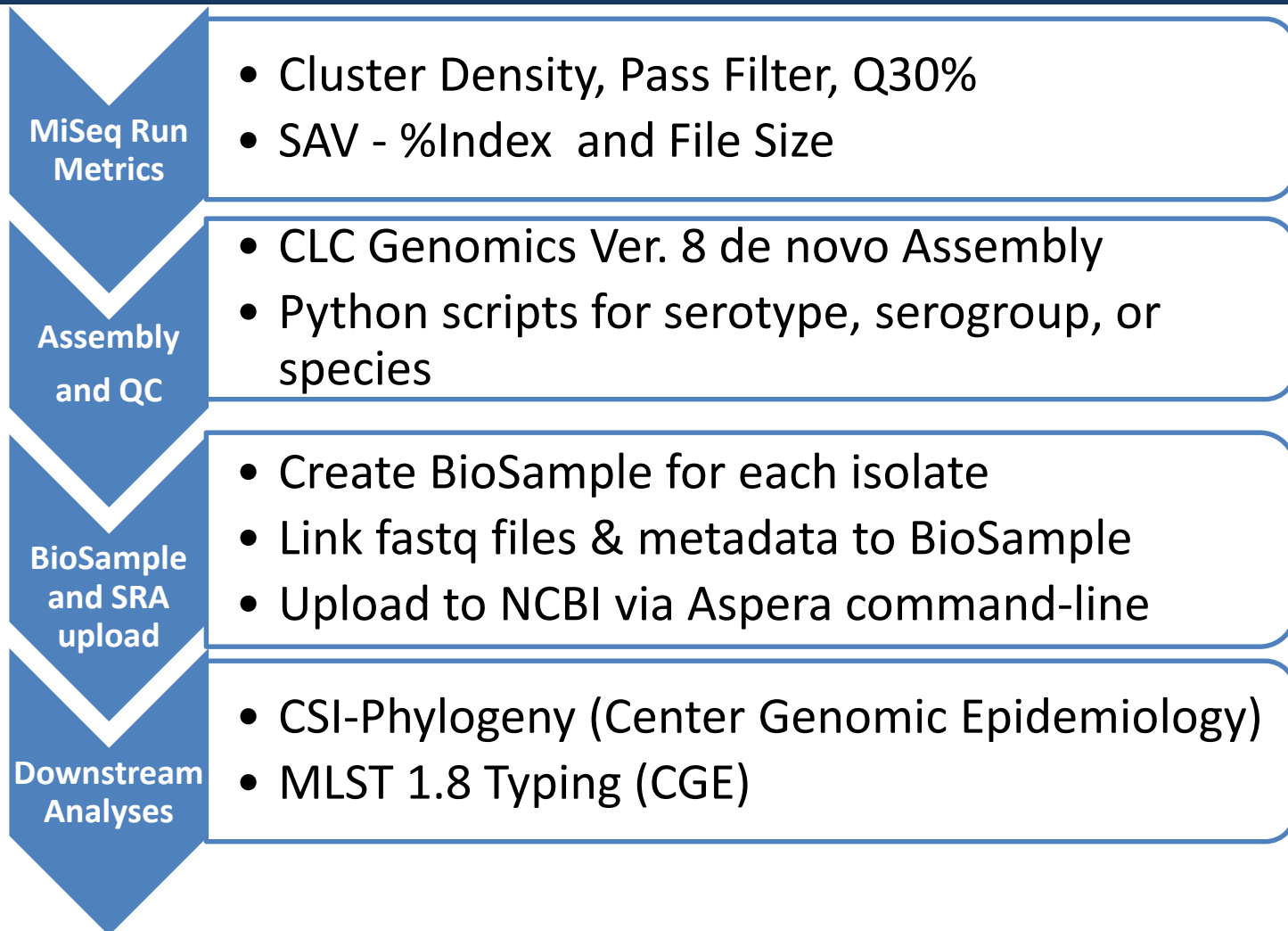
Food Safety and Inspection Service:

Lm Harborage Project – WGS Wet-Lab Workflow



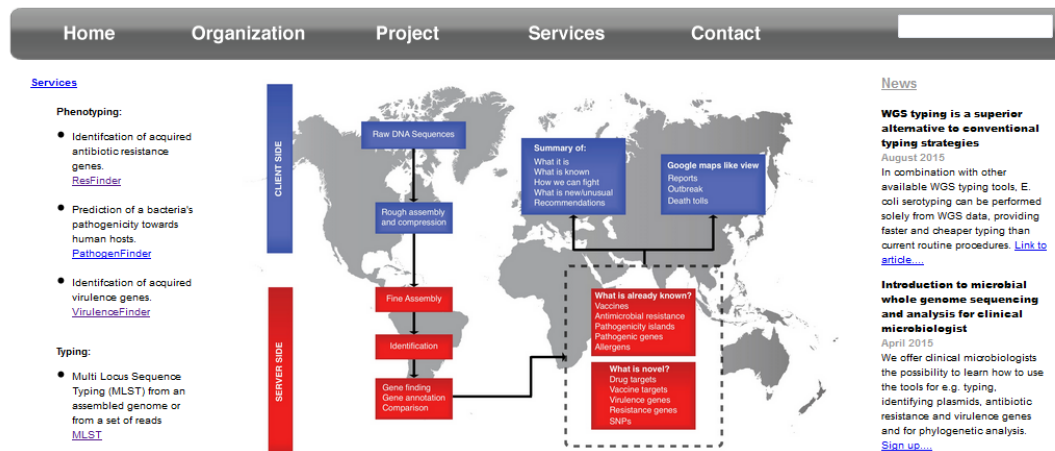
Food Safety and Inspection Service:

Lm Harborage Project - WGS Dry-Lab Workflow



Food Safety and Inspection Service: *Lm* Harborage Project - WGS Dry-Lab Workflow

Center for Genomic Epidemiology



- [CSIPhylogeny](#) inferred phylogeny based on the concatenated alignment of the high quality SNPs
- [MLST](#) Multi Locus Sequence Typing (MLST) from an assembled genome or from a set of reads
- [KmerFinder](#) Prediction of bacterial species using a fast K-mer algorithm
- [ResFinder](#) Identification of acquired antibiotic resistance genes.

Food Safety and Inspection Service: *Lm* Harborage Project - CSIPhylogeny

CSIPhylogeny Pipeline

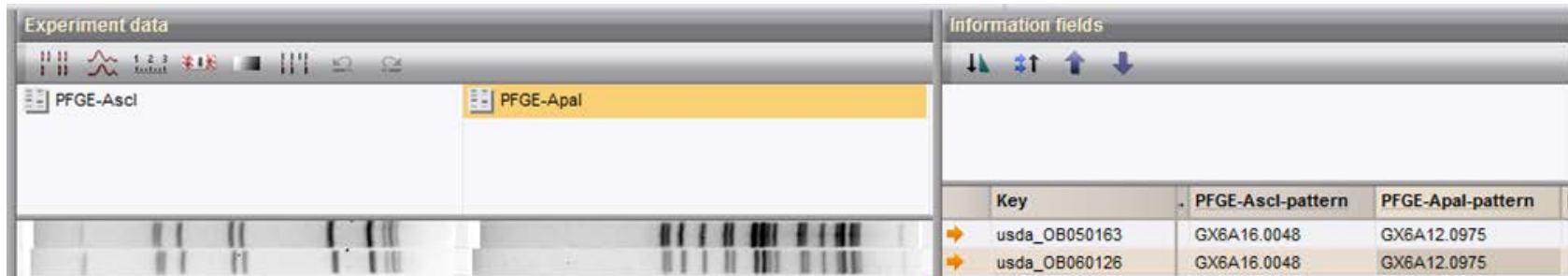
- Kaas *et al.* (2014) Solving the Problem of Comparing Whole Bacterial Genomes across Different Sequencing Platforms. PloS ONE 9(8) e104984.
- BWA - Mapping of reads to reference
- genomeCoverageBed in BEDtools – Calculate depth at each mapped position
- SAMTools – mpileup to call SNPs
- SNP Filtering (user input) – desired coverage, support, read mapping
- FastTree (optional) - Maximum Likelihood (ML) Tree
- Other tree program (Garli, FigTree, Mega 6) to build ML tree
 - Multiple Sequence Alignment (MSA) based on concatenated SNPs

Food Safety and Inspection Service: *Lm* Harborage Project – Establishment A

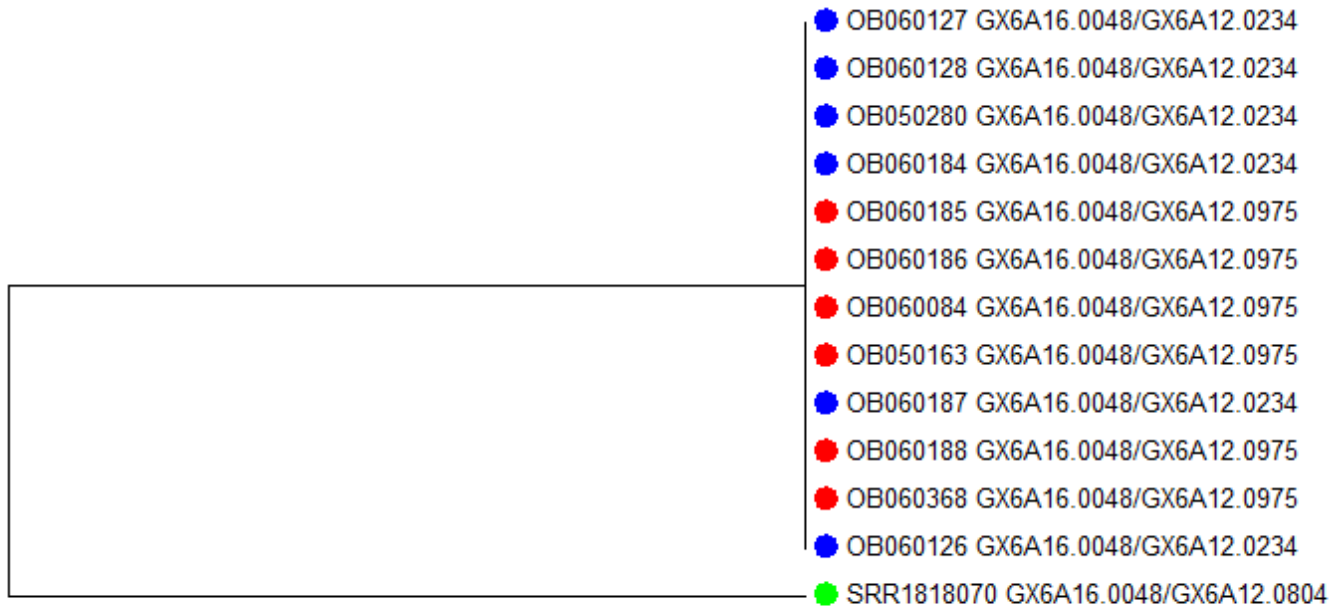
- 12 isolates from Establishment A chosen for WGS
 - 6 with pattern combination .0048/.0234
 - 6 with pattern combination .0048/.0975
 - 9 Environmental and 3 Food isolates
 - Collected over a 2 year period (2005-2006)

GX6A16.0048(0.92%)/ GX6A12.0234 (0.09%) Combined – (0.09%)

GX6A16.0048 (0.92%)/ GX6A12.0975 (0.05%) combined – (0.03%)



Food Safety and Inspection Service: *Lm* Harborage Project – Establishment A



—
2

- Reads mapped against OB050163.fa (56 contigs) assembly
- 20X coverage at called SNP with 90% relative depth at SNP
- 5bp between SNP and Q30 Minimum SNP Quality
- Maximum Likelihood Tree constructed in Mega 6.0 using 500 bootstrap reps

Food Safety and Inspection Service: Lm Harborage Project – Establishment A

Establishment A – hqSNP matrix

	OB050163	OB050280	OB060084	OB060126	OB060127	OB060128	OB060184	OB060185	OB060186	OB060187	OB060188	OB060368	SRR181807
OB050163	0	1	1	1	2	1	1	1	1	0	0	0	524
OB050280	1	0	2	2	1	0	0	0	0	1	1	1	525
OB060084	1	2	0	2	3	2	2	2	2	1	1	1	525
OB060126	1	2	2	0	3	2	2	2	2	1	1	1	525
OB060127	2	1	3	3	0	1	1	1	1	2	2	2	526
OB060128	1	0	2	2	1	0	0	0	0	1	1	1	525
OB060184	1	0	2	2	1	0	0	0	0	1	1	1	525
OB060185	1	0	2	2	1	0	0	0	0	1	1	1	525
OB060186	1	0	2	2	1	0	0	0	0	1	1	1	525
OB060187	0	1	1	1	2	1	1	1	1	0	0	0	524
OB060188	0	1	1	1	2	1	1	1	1	0	0	0	524
OB060368	0	1	1	1	2	1	1	1	1	0	0	0	524
SRR1818070	524	525	525	525	526	525	525	525	525	524	524	524	0

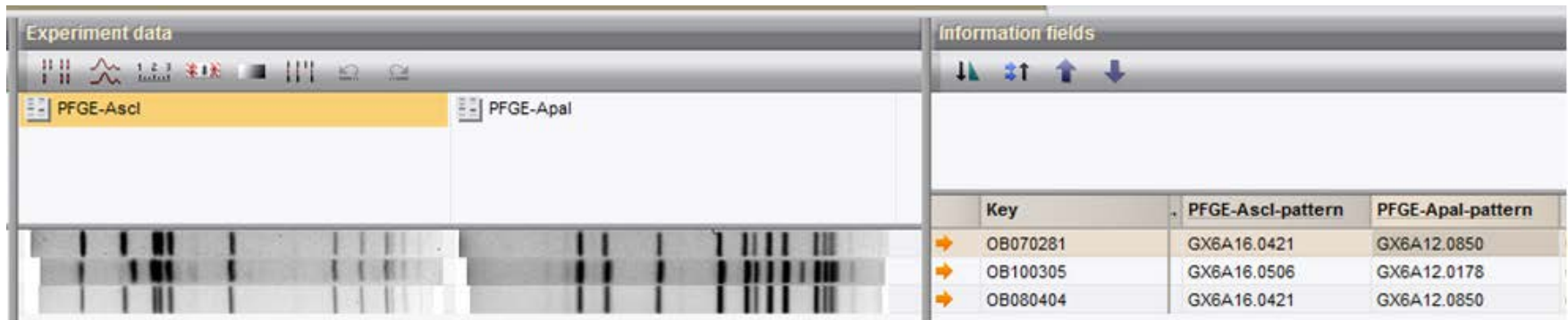
Only 1 MLST type based on WGS data for 7 housekeeping genes

- ST-88

Food Safety and Inspection Service: *Lm* Harborage Project – Establishment B

- 14 isolates from Establishment B chosen for WGS
 - 11 with pattern combination .0421/.0850
 - 1 with pattern combination .0506/.0178
 - 11 Environmental and 3 Food isolates
 - Collected over a 5 year span

GX6A16.0421(0.16%)/ GX6A12.0850(0.46%) Combined – (0.09%)



Food Safety and Inspection Service: *Lm* Harborage Project – Establishment B



—
2

- Reads mapped against OB070281.fa (103 contigs) assembly
- 20X coverage at called SNP with 90% relative depth at SNP
- 5bp between SNP and Q30 Minimum SNP Quality
- Maximum Likelihood Tree constructed in Mega 6.0 using 500 bootstrap reps

Food Safety and Inspection Service: Lm Harborage Project – Establishment B

Establishment B – hqSNP matrix

	100038091	100038095	OB070281	OB070306	OB080404	OB080499	OB100305	OB100401	OB100402	OB100403	OB100404	OB100405	OB100406	OB100407	SRR1449234
100038091	0	0	2	2	2	1	1	0	0	0	2	2	2	1	271
100038095	0	0	2	2	2	1	1	0	0	0	2	2	2	1	271
OB070281	2	2	0	2	2	1	1	2	2	2	2	2	2	1	271
OB070306	2	2	2	0	2	1	1	2	2	2	2	2	2	1	271
OB080404	2	2	2	2	0	1	1	2	2	2	2	2	2	1	271
OB080499	1	1	1	1	1	0	0	1	1	1	1	1	1	0	270
OB100305	1	1	1	1	1	0	0	1	1	1	1	1	1	0	270
OB100401	0	0	2	2	2	1	1	0	0	0	2	2	2	1	271
OB100402	0	0	2	2	2	1	1	0	0	0	2	2	2	1	271
OB100403	0	0	2	2	2	1	1	0	0	0	2	2	2	1	271
OB100404	2	2	2	2	2	1	1	2	2	2	0	0	0	1	271
OB100405	2	2	2	2	2	1	1	2	2	2	0	0	0	1	271
OB100406	2	2	2	2	2	1	1	2	2	2	0	0	0	1	271
OB100407	1	1	1	1	1	0	0	1	1	1	1	1	1	0	270
SRR1449234	271	271	271	271	271	270	270	271	271	271	271	271	271	270	0

3 different MLST types based on WGS data for 7 housekeeping genes

- ST-9
- ST-34
- ST-510

Food Safety and Inspection Service: *Lm* Harborage Project – Establishment B

MLST-1.8 Server - Typing Results

Sequence Type: ST-34

Locus	% Identity	HSP Length	Allele Length	Gaps	Allele
abcz	100.00	537	537	0	abcz_6
bgla	100.00	399	399	0	bgla_5
cat	100.00	496	496	0	cat_6
clape	100.00	462	462	0	clape_4
claf	100.00	471	471	0	claf_1
ldh	100.00	453	453	0	ldh_15
lhka	100.00	480	480	0	lhka_1

[extended output](#)

MLST Profile: *Imonocytogenes*

Organism: *Listeria monocytogenes*

MLST-1.8 Server - Typing Results

Sequence Type: ST-9

Locus	% Identity	HSP Length	Allele Length	Gaps	Allele
abcz	100.00	537	537	0	abcz_5
bgla	100.00	399	399	0	bgla_5
cat	100.00	496	496	0	cat_6
clape	100.00	462	462	0	clape_4
claf	100.00	471	471	0	claf_1
ldh	100.00	453	453	0	ldh_4
lhka	100.00	480	480	0	lhka_1

[extended output](#)

MLST Profile: *Imonocytogenes*

Organism: *Listeria monocytogenes*

MLST-1.8 Server - Typing Results

Sequence Type: ST-510

Locus	% Identity	HSP Length	Allele Length	Gaps	Allele
abcz	100.00	537	537	0	abcz_6
bgla	100.00	399	399	0	bgla_5
cat	100.00	496	496	0	cat_6
clape	100.00	462	462	0	clape_4
claf	100.00	471	471	0	claf_1
ldh	100.00	453	453	0	ldh_200
lhka	100.00	480	480	0	lhka_1

[extended output](#)

MLST Profile: *Imonocytogenes*

Organism: *Listeria monocytogenes*

ST-34

SNP within *ldh* = 3
different MLST types

Identity	550												560											
1. OB070306	T	G	G	A	G	A	C	A	C	A	C	A	A	C	A	A	A							
2. OB070281	T	G	G	A	G	C	C	A	C	A	C	A	A	C	A	A	A							
3. OB080404	T	G	G	A	G	C	C	A	C	A	C	A	A	C	A	A	A							

Identity	630												
1. OB070306	T	A	T	T	T	T	T	C	G	T	A	A	G
2. OB070281	T	A	T	T	T	T	T	C	G	T	A	A	G
3. OB080404	T	A	T	T	T	T	C	C	G	T	A	A	G

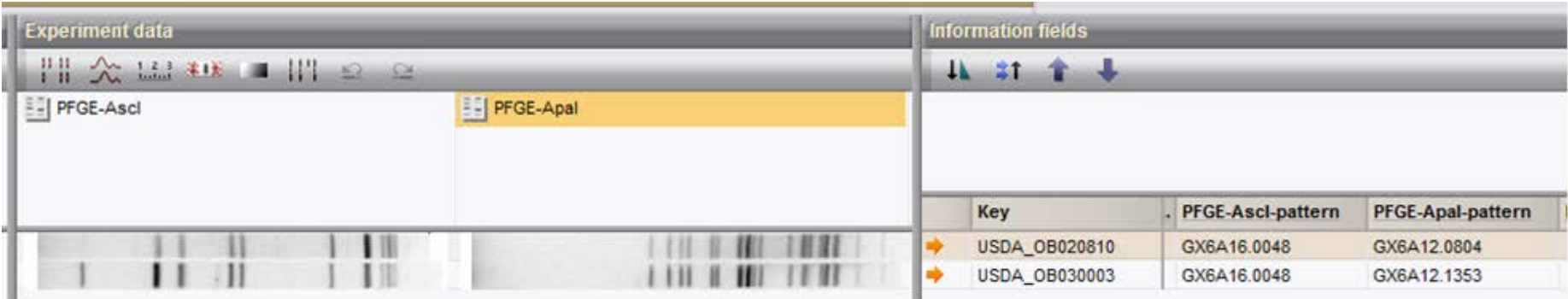
ST-510

Food Safety and Inspection Service: *Lm* Harborage Project – Establishment C

- 11 isolates from Establishment C chosen for WGS
 - 10 with pattern combination .1048/.0804
 - 1 with pattern combination .1048/.1353
 - 6 Environmental and 5 Food isolates
 - Collected over a 6 year span (2002-2009)

GX6A16.0048 (0.92%)/GX6A12.0804 (0.2%) combined - 0.2%

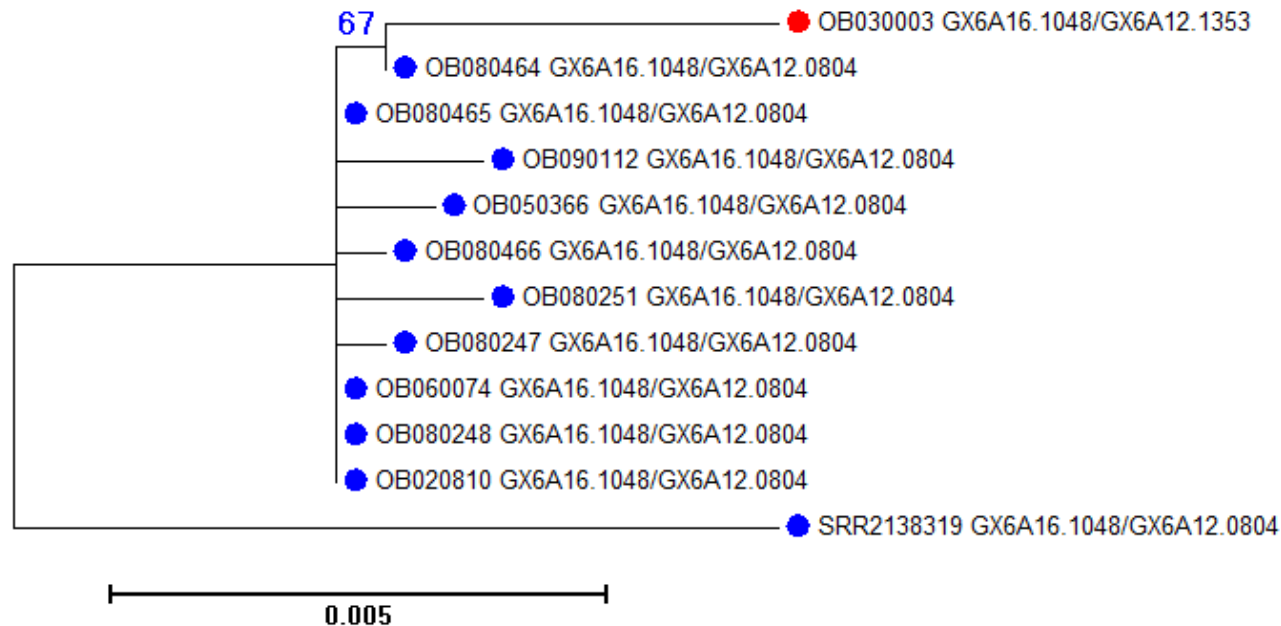
GX6A16.0048(0.92%)/GX6A12.1353 (0.07%) combined – 0.06%



The screenshot displays a software interface for PFGE analysis. On the left, under 'Experiment data', there are two gel images labeled 'PFGE-Ascl' and 'PFGE-Apal'. On the right, under 'information fields', there is a table with the following data:

Key	PFGE-Ascl-pattern	PFGE-Apal-pattern
→ USDA_OB020810	GX6A16.0048	GX6A12.0804
→ USDA_OB030003	GX6A16.0048	GX6A12.1353

Food Safety and Inspection Service: *Lm* Harborage Project – Establishment C



- Reads mapped against OB020810.fa (34 contigs) assembly
- 20X coverage at called SNP with 90% relative depth at SNP
- 5bp between SNP; Q30 Minimum SNP Quality
- Maximum Likelihood Tree constructed in Mega 6.0 using 500 bootstrap reps

Food Safety and Inspection Service: Lm Harborage Project – Establishment C

Establishment C – hqSNP matrix

	OB080465	OB080466	OB090112	OB030003	OB050366	OB060074	OB080247	OB080248	OB080251	OB080464	SRR2138319
OB080465	0	1	3	9	2	0	1	0	3	1	22
OB080466	1	0	4	10	3	1	2	1	4	2	23
OB090112	3	4	0	12	5	3	4	3	6	4	25
OB030003	9	10	12	0	11	9	10	9	12	8	31
OB050366	2	3	5	11	0	2	3	2	5	3	24
OB060074	0	1	3	9	2	0	1	0	3	1	22
OB080247	1	2	4	10	3	1	0	1	4	2	23
OB080248	0	1	3	9	2	0	1	0	3	1	22
OB080251	3	4	6	12	5	3	4	3	0	4	25
OB080464	1	2	4	8	3	1	2	1	4	0	23
SRR2138319	22	23	25	31	24	22	23	22	25	23	0

Only 1 MLST type based on WGS data for 7 housekeeping genes

- ST-296



Food Safety and Inspection Service: *Lm Harborage Project – Summary*

- Three establishments chosen with Lm isolates collected over multiple sampling events
- WGS data showed good concordance with PFGE data
- Existing bioinformatics tools can assist with questions
- Next Steps
 - Integrate other pipelines for data analyses
 - Examine more establishments (more common PFGE patterns)



Food Safety and Inspection Service: Acknowledgements

- **FSIS:**
OPHS Whole Genome Sequencing Work Group
- **CDC:**
Heather Carleton, Peter Gerner-Smidt, Lee Katz, Ashley Sabol, Steven Stroika, Eija Trees
- **NIH:**
Bill Klimke, David Lipman, Jim Ostell, Martin Shumway
- **FDA:**
Peter Evans, Patrick McDermott, Justin Payne, Ruth Timme, Charlie Wang