



United States Department of Agriculture

One Team, One Purpose



Food Safety and Inspection Service

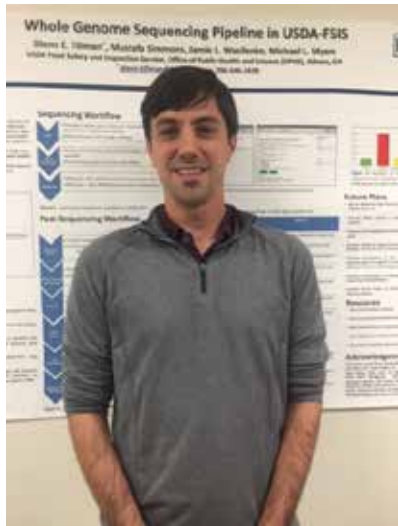
Protecting Public Health and Preventing Foodborne Illness



Food Safety and Inspection Service:

FSIS: Analytical Tools for WGS

October 26, 2017



Glenn E. Tillman, PhD
Chief, Microbiology Characterization Branch
Eastern Laboratory (Athens, GA)
Office of Public Health Science
Food Safety Inspection Service, USDA

Food Safety and Inspection Service:

FSIS: Update on WGS

- Background
- What are we doing (or planning to do) with our data?
- Concluding remarks

Food Safety and Inspection Service:

Why WGS?

- **Improved resolution for foodborne illness investigations**
 - Improved strain discrimination, illness cluster detection, and case classification
- **Supports FSIS mission goals**
 - Effectively use science to understand foodborne illness and emerging microbiological trends
 - Identification of environmental harborage or recurrences of pathogens in FSIS-regulated establishments/products to further support the inspection and verification process
- **Alignment of pathogen surveillance with our domestic public health and regulatory partners**
 - Collaborative efforts with US Food and Drug Administration Center for Food Safety and Applied Nutrition (FDA-CFSAN), the US Centers for Disease Control and Prevention (CDC), the US National Institutes of Health National Center for Biotechnology Information (NCBI), and also state/local health partners/laboratories

Food Safety and Inspection Service:

WGS at FSIS: Current Status

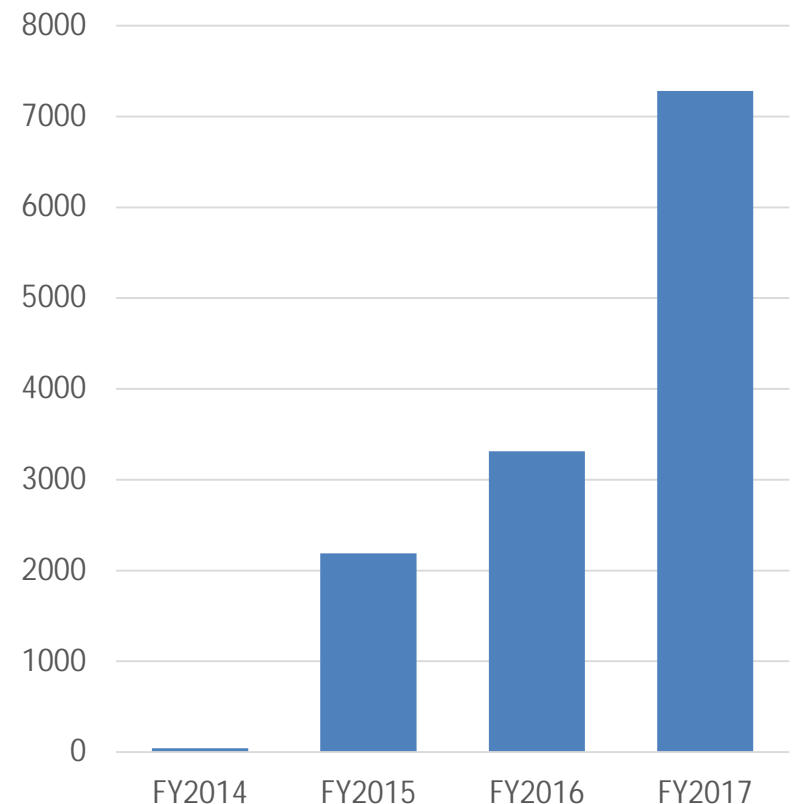
- FSIS has built capacity for conducting WGS on all isolates obtained from FSIS sampling programs
 - Currently 12 sequencers in FSIS Field Service Laboratories, and expect all to be operational in early FY18
 - In FY17, FSIS sequenced 7,282 isolates
- In collaboration with our public health and regulatory partners, FSIS currently considers available WGS analyses in addition to PFGE, epidemiological and traceback information to further understand the relationship between clinical and food isolates
- FSIS works with National Antimicrobial Resistance Monitoring System (NARMS) partners (FDA, CDC) to understand the occurrence or introduction of antimicrobial resistance genes in pathogens of interest

Food Safety and Inspection Service: FSIS Sequencing Efforts and Data Sharing

FSIS Submissions to NCBI Bioprojects

- PRJNA242847
 - GenomeTrakr Project: USDA-FSIS (*Salmonella*)
- PRJNA215355
 - GenomeTrakr Project: FDA (*Listeria monocytogenes*)
- PRJNA287430
 - USDA-FSIS: *Campylobacter*
- PRJNA268206
 - GenomeTrakr Project: USDA-FSIS (STEC)
- PRJNA292666
 - FSIS NARMS *Salmonella*
- PRJNA292667
 - FSIS NARMS *E. coli*
- PRJNA292668
 - FSIS NARMS *Campylobacter*
- PRJNA292669
 - FSIS NARMS *Enterococcus*

Number of Isolates Sequenced by Fiscal Year



Food Safety and Inspection Service: WGS Data Sharing – Metadata

The screenshot shows the NCBI BioSample web interface for the sample FSIS1606985. The browser address bar shows the URL: <https://www.ncbi.nlm.nih.gov/biosample/?term=FSIS1606985>. The page header includes the NCBI logo, navigation links (Resources, How To), and user options (usda fsis, My NCBI, Sign Out). The main content area displays the following information:

- Pathogen:** environmental/food/other sample from *Listeria monocytogenes*
- Identifiers:** BioSample: SAMN05366775, Sample name: FSIS1606985, SRA: SRS1548159
- Organism:** [Listeria monocytogenes](#)
cellular organisms; Bacteria; Terrabacteria group; Firmicutes; Bacilli; Bacillales;>Listenaceae; Listeria
- Package:** [Pathogen: environmental/food/other, version 1.0](#)
- Attributes:**
 - strain:** FSIS1606985
 - collected by:** USDA-FSIS
 - collection date:** 2016
 - geographic location:** [USA_WV](#)
 - isolation source:** Product-Swab-Non meat
 - latitude and longitude:** missing
- BioProject:** [PR_INA215355](#) Listeria monocytogenes
Retrieve [all samples](#) from this project
- Submission:** USDA FSIS, Glenn Tillman, 2016-07-08
- Accession:** SAMN05366775 ID: 5366775
[BioProject](#) [SRA](#)

On the right side, there are sections for 'Send to' (with options for Related information, BioProject, SRA, Taxonomy), 'Search details' (with a search box and 'Search' button), and 'Recent activity' (with a 'Turn Off' button and a list of recent searches).

- Description about an isolate that has an experiment assigned
- Metadata, such as source type, organism, serotype etc.
- BioSample number is specific to a bacterial isolate for FSIS

Food Safety and Inspection Service:

FSIS: WGS Analytical Tools Overview

- Quality Control Assessment
- Antibiotic Resistance gene detection
- Salmonella and STEC serotype determination
- STEC Virulence Gene characterization
- Phylogenetic Comparisons

Food Safety and Inspection Service:

Further Characterization of Isolates Using WGS

- ***Campylobacter* speciation**
- **Serotyping/serogrouping**
 - *Salmonella*
 - Adulterant STEC
- **Antimicrobial Resistance (Phenotype prediction)**
 - *Salmonella*
 - *Campylobacter*
 - *E. coli*
 - *Enterococcus*
- **Identify characterized genes of interest**
 - Resistance to environmental factors (heat, acid, etc)
 - Virulence factors
 - *stx/ae* sub-types (STEC)
- **Alternative to PFGE for comparison of genotypes**
 - wgMLST analyses
 - SNP analyses

A single workflow for many characterization approaches *via* informatics

Food Safety and Inspection Service: FSIS: WGS Data Analyses Work Flow Overview



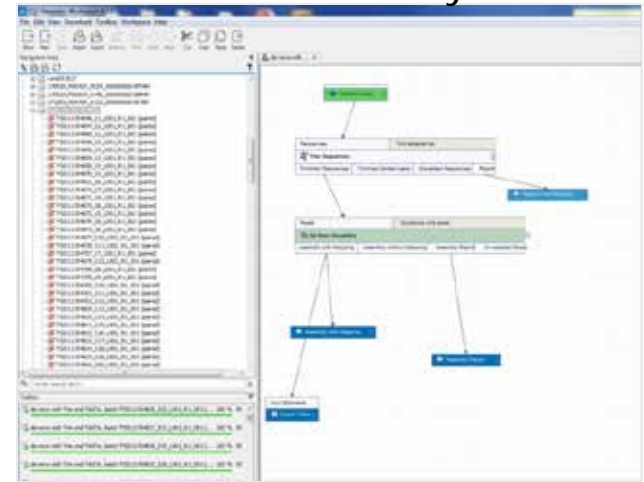
Input: FASTQ

- QC Pipeline
 - Coverage
 - Average Quality
 - Nucleotide balance

```
@M02848:54:000000000-AJ347:1:1101:14587:1926 1:N:0:4
TCCGTGCTCAGTTACACGGACAAAATACCGGCGAAAAACCTGGGCCCTCCCTGCCGACATGGGATT
+
AA>1A1AAD31DF3F3B111A0000B1BB00A/////0AAGG11/BFFGFFG1/>//E?10E0FGDGF1
```

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

De novo Assembly



Output: FASTA

- QC Pipeline
 - File Size
 - N50 & No. contigs
 - Correct organism

Input: FASTA

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

```
>FSIS1609314_s10_L001_R1_001_(paired)_trimmed_(paired)_contig_1
TGCGGTTTTGTACGTTCAAATTTTTCTTTAGACACGGCTATATTCCTTACTATAGCGCTC
TCCCCTTCAGGAGAGAGCACGGGATTTGGTTTTTAACCCCTCGGGCTATTTACCACGGG
CTTCGATTACGGCCTGAGCAACGTTGTTTCGGCGCATCATCATACTTCAGGAATCCATAG
TGTACGATCGCGACCTTTGGTCAGAGAACCGACGCTGAGTTGCATATCCGAACATTTAG
ACAGCGGTACTTCAGCGGGAATCTTAACGCCGTAACTTCAGATTCCTGACCTTTGAGCA
TACCACGACGGCGGCTAAGGTCCGGGATAACGTCACCGGTAITCTCTTCAGGTGTTTCTA
CTTCAACCTTCATGATCGGCTCAAGCAGAAGCTGTTTTGCTTTCTTAAAGCCTCTTTTAA
AGCGGATAGACGCAGCCAGTTTAAACGCCAGCTCAGAGGAGTCAACGTCATGGTAAGAAC
```

Food Safety and Inspection Service:

Genotypic Screening for Antimicrobial Resistance Using WGS Data

- Ability to rapidly identify new genes of concern in sequenced isolates
- FSIS works with NARMS and other public health and regulatory partners to identify presence of genes or genetic markers of interest
 - When possible, FSIS currently compares genotype prediction to phenotype using the NARMS panel
- Examples:
 - Extended spectrum beta-lactamases (*bla*_{CTX-M-65})
 - Colistin
 - Quinolone
 - Linezolid
 - Daptomycin



2015

Salmonella and
Campylobacter
from **cecal** with
 $R \geq 1$
antimicrobial



2016

All *Salmonella*
and
Campylobacter
from **cecal**



2017

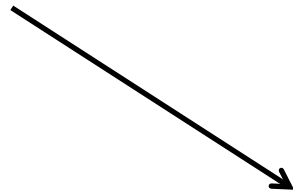
All *Salmonella*
and
Campylobacter
from both **cecal**
and **HACCP**

AMR *E. coli* and
Enterococcus

Food Safety and Inspection Service: WGS usage: Antibiotic Resistance Gene Detection



FASTQ files



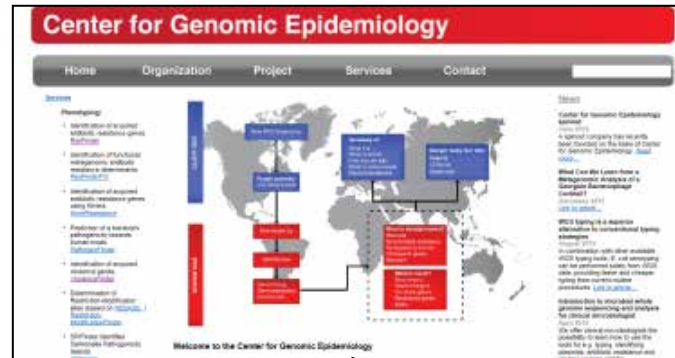
QC &
Assembly
Pipeline



Resfinder
(BLAST)



Resistance gene profile

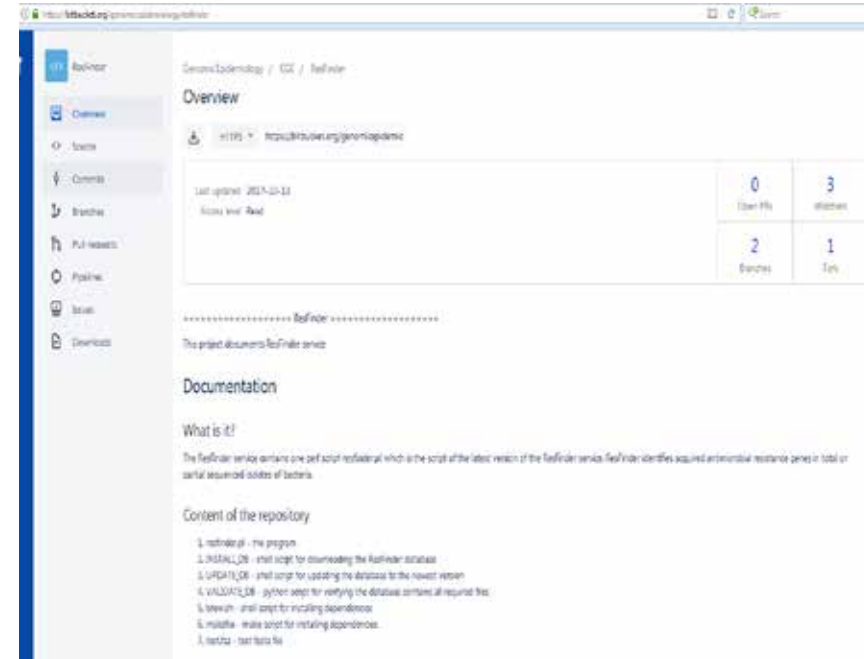


Food Safety and Inspection Service: WGS usage: Antibiotic Resistance Gene Detection

```

limsadmin@ASIFEDD447N3X12: /media/limsadmin/Data1/170929_M04790_0132_000000000-BDP3M/clc_assemblies/for_slide
P3M/clc_assemblies/for_slide$ class_resfinder170825.sh
Processing: FSIS11704693.fa
Found 2 aminoglycoside genes in FSIS11704693.fa
Processing: FSIS11704693.fa
Found 1 beta-lactam genes in FSIS11704693.fa
Processing: FSIS11704693.fa
Found 0 colistin genes in FSIS11704693.fa
Processing: FSIS11704693.fa
Found 0 fosfomycin genes in FSIS11704693.fa
Processing: FSIS11704693.fa
Found 0 fusidicacid genes in FSIS11704693.fa
Processing: FSIS11704693.fa
Found 0 macrolide genes in FSIS11704693.fa
Processing: FSIS11704693.fa
Found 0 nitroimidazole genes in FSIS11704693.fa
Processing: FSIS11704693.fa
Found 0 oxazolidinone genes in FSIS11704693.fa
Processing: FSIS11704693.fa
Found 2 phenicol genes in FSIS11704693.fa
Processing: FSIS11704693.fa
Found 0 quinolone genes in FSIS11704693.fa
Processing: FSIS11704693.fa
Found 1 rifampicin genes in FSIS11704693.fa
Processing: FSIS11704693.fa
Found 1 sulphonamide genes in FSIS11704693.fa
Processing: FSIS11704693.fa
Found 1 tetracycline genes in FSIS11704693.fa
Processing: FSIS11704693.fa
Found 0 trimethoprim genes in FSIS11704693.fa
Processing: FSIS11704693.fa
Found 0 vancomycin genes in FSIS11704693.fa

```



<https://bitbucket.org/genomicepidemiology/resfinder>

GENE	%COVERAGE	%IDENTITY	#FILE	CATEGORY	SEQUENCE
blaCTX-M-65	100	100	FSIS11704693	Beta-lactam resistance	FSIS11704693_S10_L001_R1_001_(paired)_trimmed_(paired)_contig_19
aadA1	95.37	99.89	FSIS11704693	Aminoglycoside resistance	FSIS11704693_S10_L001_R1_001_(paired)_trimmed_(paired)_contig_45
tet(A)	100	100	FSIS11704693	Tetracycline resistance	FSIS11704693_S10_L001_R1_001_(paired)_trimmed_(paired)_contig_45
sul1	100	100	FSIS11704693	Sulphonamide resistance	FSIS11704693_S10_L001_R1_001_(paired)_trimmed_(paired)_contig_45
floR	99.92	98.19	FSIS11704693	Phenicol resistance	FSIS11704693_S10_L001_R1_001_(paired)_trimmed_(paired)_contig_70

Food Safety and Inspection Service:

Antimicrobial Resistance (Phenotype Prediction) Using WGS Analyses

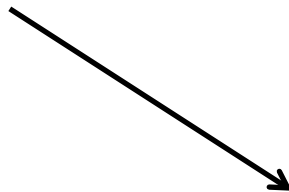
- Analyzed ~1190 *Salmonella* isolates from FY2016 NARMS cecal sampling
 - Includes isolates from all animals tested
- Compared genotypic prediction for resistance phenotype to reported phenotype using NARMS panel
 - Overall concordance for 1130 isolates (95.0%)

Phenotype	Number of Isolates with Phenotype Using NARMS Panel	Percent Concordance of Genotype with Phenotype
Pan-susceptible	769	99.1%
Beta-lactam resistance	110	99.1%
Aminoglycoside resistance	167	98.8%
Phenicol resistance	42	97.6%
Tetracycline resistance	330	99.7%

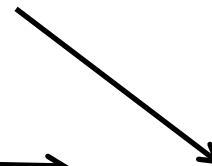
Food Safety and Inspection Service: WGS usage: Salmonella Serotype Determination



FASTQ files



Assembly
Pipeline



SeqSero
(BLASTdb)



Serotype Determination

Food Safety and Inspection Service:

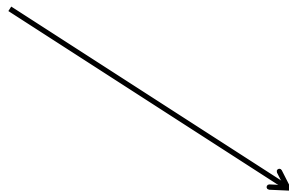
Using WGS Analyses for Serotype Determination

- FSIS sequenced 4205 *Salmonella* isolates from various sampling programs from Jan 2015 through Jun 2017
- Compared serotype reported by routine methods (molecular serotyping or traditional serology) with serotype determined using a custom query to SeqSero database
 - For 4045/4205 (96.2%) of isolates, WGS that matched reported serology result
 - For 160/4205 (3.8%) of isolates, WGS did not match reported serology result
 - Includes isolates with incomplete genetic factor set (cannot call/identify serotype)

Food Safety and Inspection Service: WGS usage: MLST and virulence typing in Shiga toxin-producing *Escherichia coli*



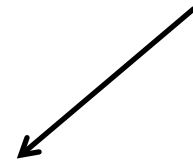
FASTQ files



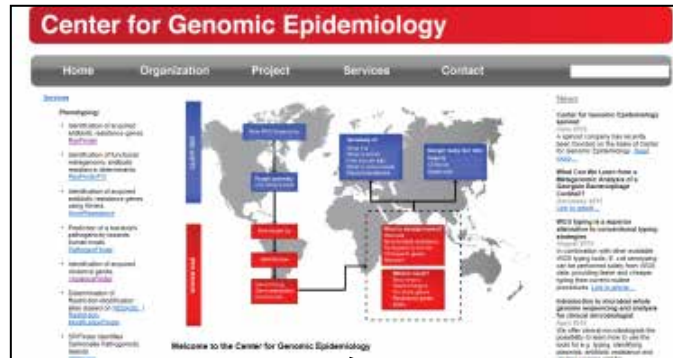
Assembly
Pipeline



VirulenceFinder
SerotypeFinder
(BLAST)



Virulence genes and
serotype



Food Safety and Inspection Service: WGS usage: MLST and virulence typing in Shiga toxin-producing *Escherichia coli*

```
limsadmin@ASIFEDD447N3X12: /media/limsadmin/Data3/ecoli_isolates/for_slide
limsadmin@ASIFEDD447N3X12: /media/limsadmin/Data3/ecoli_isolates/for_slide$ stec_characterization
Processing: FSIS1608854.fa
Found 2 O antigen genes in FSIS1608854.fa
Processing: FSIS1608854.fa
Found 1 H antigen genes in FSIS1608854.fa
Processing: FSIS1608854.fa
Found 7 MLST genes in FSIS1608854.fa
Processing: FSIS1608854.fa
Found 1 shiga toxin genes in FSIS1608854.fa
Processing: FSIS1608854.fa
Found 1 intimin genes in FSIS1608854.fa
limsadmin@ASIFEDD447N3X12: /media/limsadmin/Data3/ecoli_isolates/for_slide$ █
```

FSIS_Number	Serotype	Sequence_type	stx_type	eae_allele
FSIS1608854	O26:H11	ST-21	stx1a	Beta1

Food Safety and Inspection Service:

Using WGS Analyses for MLST and virulence typing in Shiga toxin-producing *Escherichia coli*

Serogroup (no. sequenced)	Top 7 gene MLST sequence types	Top <i>stx</i> types	<i>eae</i> types	Top Serotype
O26 (50)	ST-21 (82%)	<i>stx1a</i> (80%) <i>stx2a</i> (16%)	Beta1 (100%)	O26:H11 (100%)
O45 (20)	ST-17 (95%)	<i>stx1a</i> (100%)	Epsilon (95%)	O111:H8 (100%)
O103 (147)	ST-17 (67.8%)	<i>stx1a</i> (98.0%)	Epsilon (84.9%)	O103:H2 (84.9%)
O111 (44)	ST-16 (93.0%)	<i>stx1a</i> (86.0%)	Theta (100%)	O111:H8 (100%)
O121 (15)	ST-655 (93.3%)	<i>stx2a</i> (100%)	Epsilon (100%)	O121:H19 (100%)
O145 (18)	ST-32 (100%)	<i>stx1a stx2d</i> (38.9%) <i>stx2a</i> (27.7%) <i>stx2c</i> (11.1%)	Gamma-1 (100%)	O145:H28 (100%)
O157 (149)	ST-11 (96.00%)	<i>stx1a stx2a</i> (28.2%) <i>stx2c</i> (22.8%) <i>stx2a</i> (20.1%)	Gamma-1 (100%)	O157:H7 (100%)

- *E. coli* serotypes, O157:H11 & O157:H29 identified to be lacking *eae/stx* through WGS
- *stx*-negative strains can be identified as serotype *E. coli* O157:H7 based on WGS

Food Safety and Inspection Service:

WGS Analyses for Phylogenetic Context: wgMLST and hgSNP Analyses

FSIS uses pipelines developed by public health partners

Lyve-SET

<https://github.com/lskatz/lyve-SET/blob/master/>

NCBI Pathogen Detection Isolates Browser

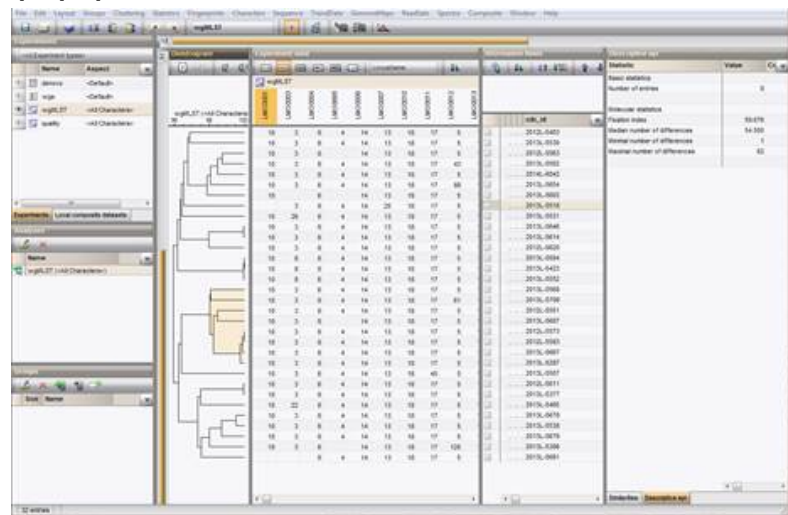
<http://www.ncbi.nlm.nih.gov/pathogens>

FDA SNP Pipeline

<https://github.com/CFSAN-Biostatistics/snp-pipeline>

wgMLST

BioNumerics 7.6 CDC-PulseNet



Food Safety and Inspection Service

WGS usage: SNP tree from NCBI Pathogen Pipeline

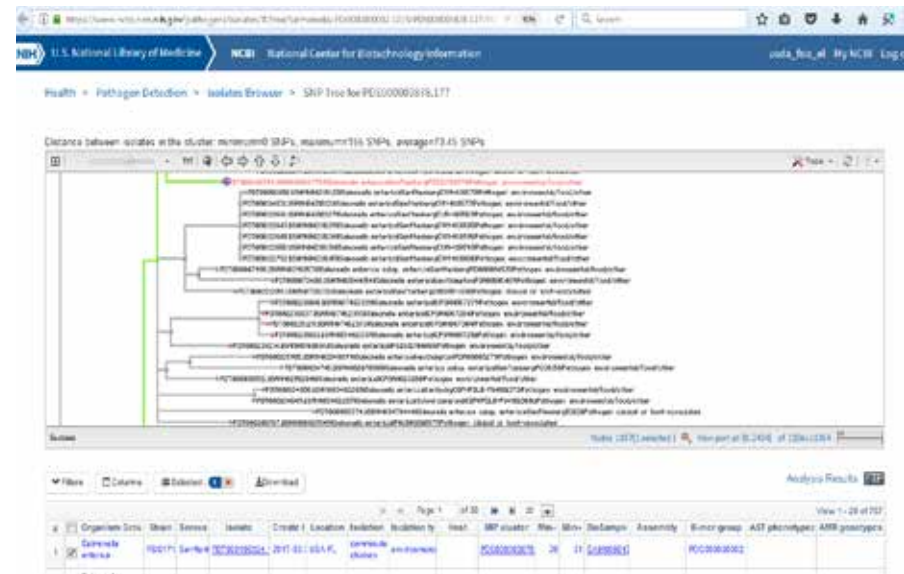
<https://www.ncbi.nlm.nih.gov/pathogens/isolates/>

Health > Pathogen Detection > Isolates Browser

Search:

Organism Group:

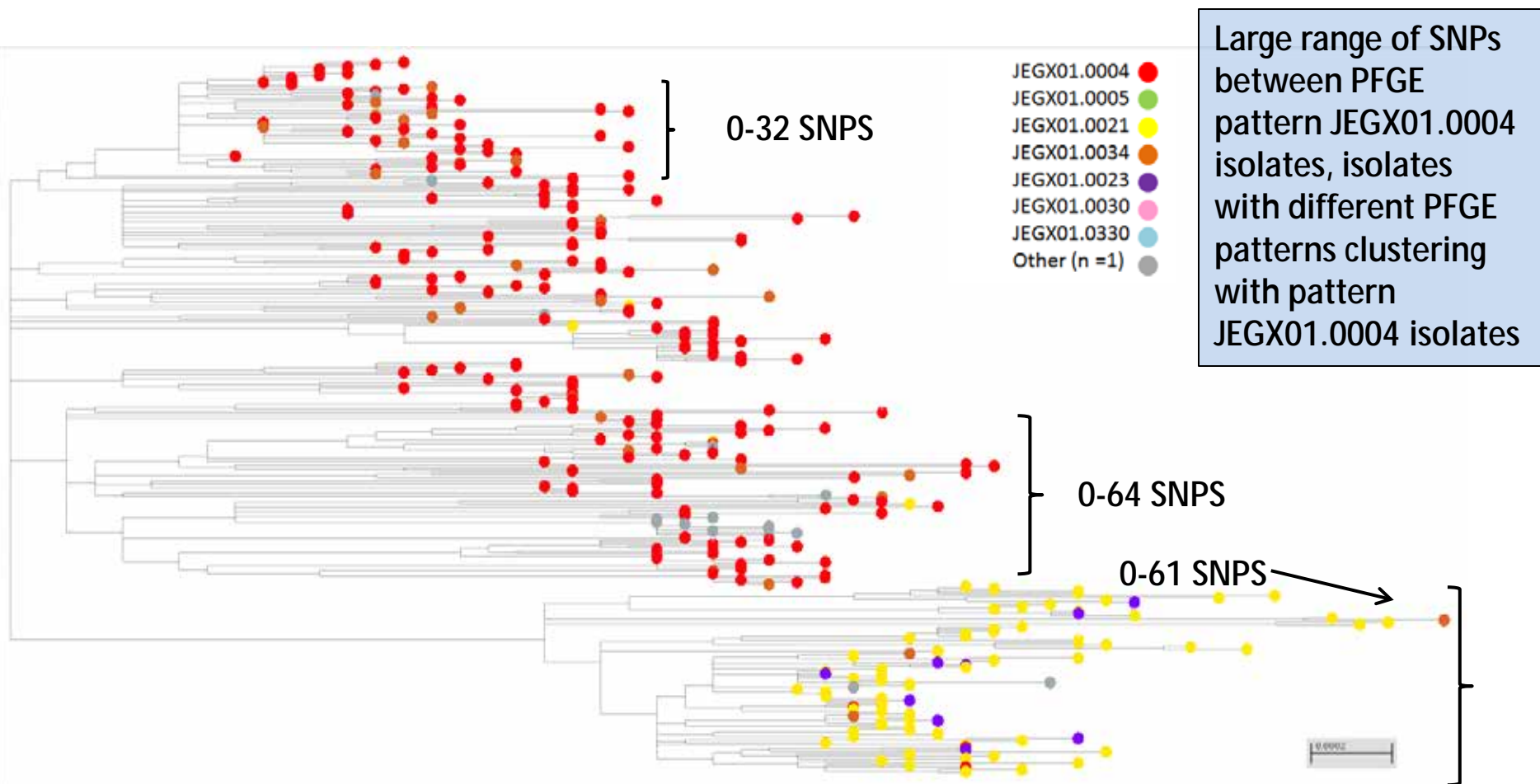
#	Organism Group	Strain	Accession	Isolate	Country	Location	Isolation Date	Isolation Type	Host	SNP cluster	Min. Size	Max. Size	Seq. Length	Assembly	Bin group	AST pt	AST group(s)
1	Campylobacter jejuni	F30111006	FC1388191006	F	USA	MA	2017-01-21	Animal - Swine - Market	environmental		44	44	2489620001	FC1388191006	FC1388191006		NA20A:101
2	Campylobacter jejuni	F30111007	FC1388191007	F	USA	TX	2017-01-21	Animal - Cattle - Feed	environmental		44	44	2489620008	FC1388191007	FC1388191007		NA20A:401
3	Campylobacter jejuni	F30111008	FC1388191008	F	USA	VA	2017-01-21	Animal - Cattle - Feed	environmental	FC0000010412	6	16	2489620002	FC1388191008	FC1388191008		NA20A:101
4	Campylobacter jejuni	F30111009	FC1388191009	F	USA	VA	2017-01-21	Animal - Swine - Market	environmental		44	44	2489620010	FC1388191009	FC1388191009		NA20A:101
5	Campylobacter jejuni	F30111010	FC1388191010	F	USA	MA	2017-01-21	Animal - Cattle - Feed	environmental	FC0000010412	11	9	2489620019	FC1388191010	FC1388191010		WFO
6	Campylobacter jejuni	F30111011	FC1388191011	F	USA	TX	2017-01-21	Animal - Cattle - Feed	environmental	FC0000010412	8	11	2489620005	FC1388191011	FC1388191011		NA20A:101
7	Campylobacter jejuni	F30111012	FC1388191012	F	USA	TX	2017-01-21	Animal - Cattle - Feed	environmental	FC0000010412	41	44	2489620001	FC1388191012	FC1388191012		NA20A:101



Product	Serotype	Total #	% in SNP cluster (#)	% w/in 20 SNPs of clinical	% w/in 10 SNPs of clinical
Total Serotypes	Enteritidis	447	99.8% (446)	98.9% (442)	90.6% (405)
	Kentucky	607	99.8% (606)	0% (0)	0% (0)
	Heidelberg	74	98.6% (73)	98.6% (73)	86.4% (64)
	Schwarzengrund	122	100% (122)	47.5% (58)	7.3% (9)
	Infantis	218	94.5% (206)	92.2% (201)	50.9% (111)
	Typhimurium	155	96.8% (150)	30.3% (47)	10.3% (16)

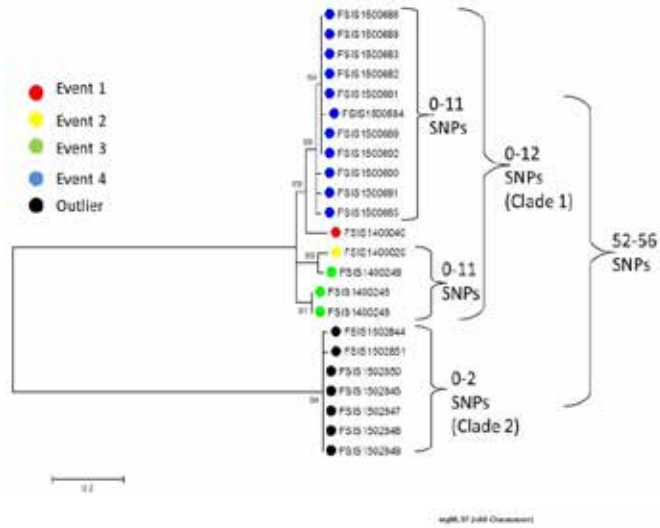
Food Safety and Inspection Service

WGS usage: SNP tree from NCBI Pathogen Pipeline

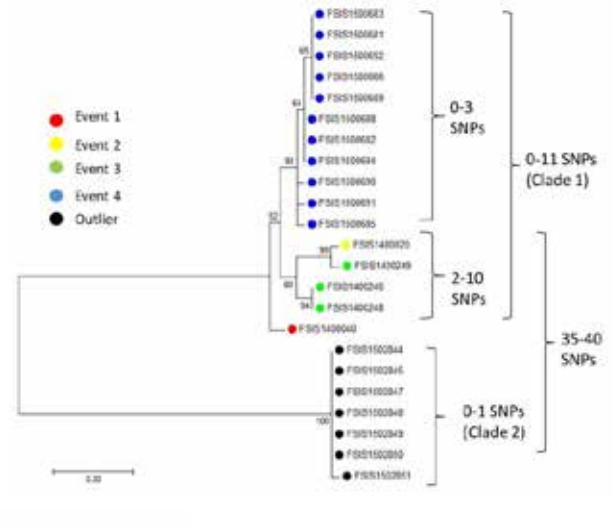


Food Safety and Inspection Service

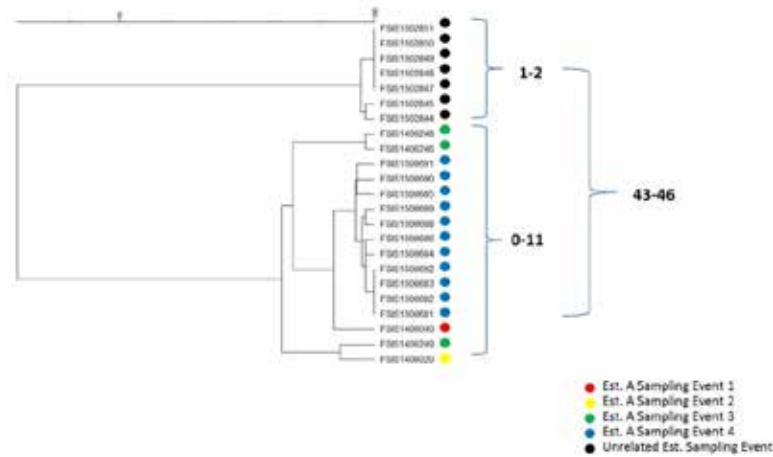
WGS usage: *hq*SNP analyses and *wg*MLST with *Listeria monocytogenes*



Snp Pipeline



Lyve-SET

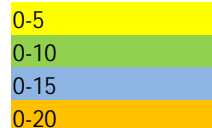


wgMLST

Food Safety and Inspection Service

WGS usage: hqSNP analyses and wgMLST with *Listeria monocytogenes*

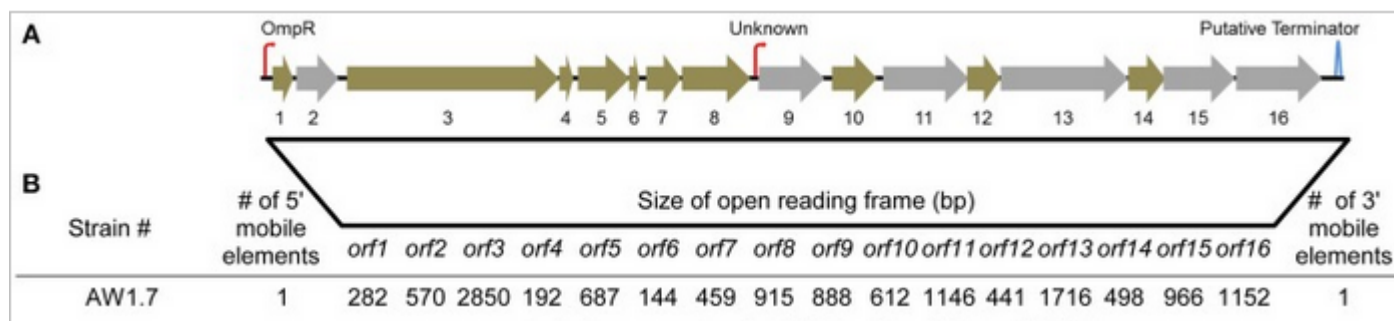
Lab Number	Primary_PN_Pattern	Sec_PN_Pattern	Lyve-SET SNP Clade	CFSAN Snp Pipeline Clade	wgMLST Clade
FSIS1400040	GX6A16.0018	GX6A12.0022	1	1	1
FSIS1400020	GX6A16.0018	GX6A12.0022	1	1	1
FSIS1400249	GX6A16.0018	GX6A12.0022	1	1	1
FSIS1400248	GX6A16.0018	GX6A12.0022	1	1	1
FSIS1400246	GX6A16.0018	GX6A12.0022	1	1	1
FSIS1500681	GX6A16.0018	GX6A12.0022	1	1	1
FSIS1500682	GX6A16.0018	GX6A12.0022	1	1	1
FSIS1500683	GX6A16.0018	GX6A12.0022	1	1	1
FSIS1500684	GX6A16.0018	GX6A12.0022	1	1	1
FSIS1500685	GX6A16.0018	GX6A12.0022	1	1	1
FSIS1500686	GX6A16.0018	GX6A12.0022	1	1	1
FSIS1500688	GX6A16.0018	GX6A12.0022	1	1	1
FSIS1500689	GX6A16.0018	GX6A12.0022	1	1	1
FSIS1500690	GX6A16.0018	GX6A12.0022	1	1	1
FSIS1500691	GX6A16.0018	GX6A12.0022	1	1	1
FSIS1500692	GX6A16.0018	GX6A12.0022	1	1	1
FSIS1502844	GX6A16.0018	GX6A12.0022	2	2	2
FSIS1502845	GX6A16.0018	GX6A12.0022	2	2	2
FSIS1502847	GX6A16.0018	GX6A12.0022	2	2	2
FSIS1502848	GX6A16.0018	GX6A12.0022	2	2	2
FSIS1502849	GX6A16.0018	GX6A12.0022	2	2	2
FSIS1502850	GX6A16.0018	GX6A12.0022	2	2	2
FSIS1502851	GX6A16.0018	GX6A12.0022	2	2	2



Food Safety and Inspection Service:

WGS usage: Usage of genotypic data to predict phenotypic characteristics

- FSIS is beginning to explore using genotypic data to predict phenotypic characteristics
 - Previous examples of antimicrobial resistance markers
 - Resistance to other environmental factors (i.e. heat, acid, certain chemicals, etc)
- Example: Locus of Heat Resistance (LHR) - 14-kb genomic island
 - Present in a diverse group of Enterobacteriaceae, including *Cronobacter sakazakii*, *Klebsiella pneumoniae*, *Enterobacter cloacae*, *E. coli*, and *Salmonella*
 - A BLAST database including genes contained in the LHR was built based on the published sequence of *E. coli* AW1.7
 - *Salmonella* isolates from FSIS-regulated products were sequenced, assembled and queried against the LHR BLAST database
 - 11 FSIS *Salmonella* isolates contained LHR genes of interest
 - Additional phenotypic testing required to determine if isolates exhibit resistance to heat
- Heat resistance can be linked to different biological pathways/genes
 - Advantage of WGS: As new pathways/genes are identified, data can be queried again



Food Safety and Inspection Service:

Concluding Remarks

- FSIS' new strategic plan is focused on the use of new technology to prevent foodborne illnesses and protect public health
 - Includes two objectives specific to incorporating WGS into FSIS surveillance and regulatory efforts
- FSIS has built sufficient capacity for conducting WGS on all FSIS pathogen isolates
 - Sequenced 7282 isolates in FY17 with the future target of ~10,000 isolates/year
- FSIS is exploring how we can use WGS data beyond outbreak investigations, including understanding the link between genotypes and phenotypes of interest
- FSIS continues to engage with national and international partners
- FSIS continues to use WGS analyses in conjunction with other metadata, including epidemiological and traceback information, to further understand the relationship between clinical, food and environmental isolates

Food Safety and Inspection Service: **Acknowledgements**

- USDA FSIS Offices
- USDA ARS
- CDC PulseNet and NARMS
- FDA CFSAN
- FDA CVM
- NCBI
- State Laboratories

