



Visualization and Communication

**Whole Genome Sequencing
Live Learning Series - Session 3**

July 2, 2025

Introduction



Mugdha Golwalkar, MPH

- Epidemiologist/Outbreak Coordinator
- Foodborne and Enteric Diseases Program
- Tennessee Department of Health

Mugdha.Golwalkar@tn.gov

Learning Objectives

Data Viz	Evaluate impact vs. investment of routine use of data visualization tools
	Recognize how data visualization can support cluster/outbreak investigations
Communicating WGS Results	Use and interpret standard phrasing for communicating WGS results
	Summarize complex laboratory testing and genetic comparisons during outbreak investigations for a general audience
PulseNet 2.0	List the requirements to obtain access to PulseNet 2.0



Cluster/WGS Data Visualization

WGS/Cluster Data Visualization Value

Easily identify
similarities and
differences

Track processes
efficiently

Describe trends
and gaps

Impact vs. Investment

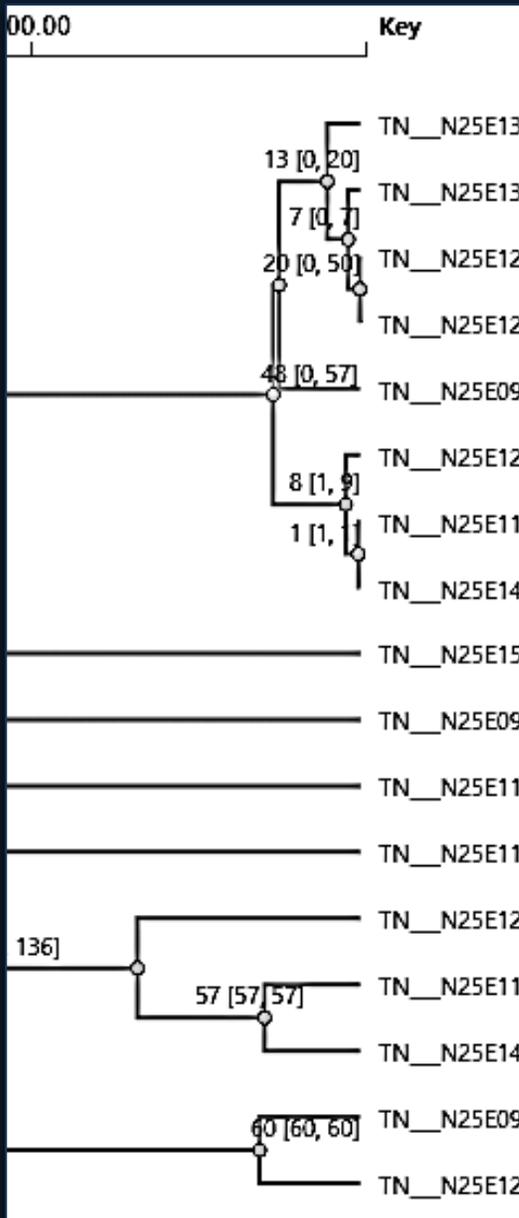
Impact

- Clear communication with partners
- Reduced time and manual processes
- Simplifies monitoring and evaluation

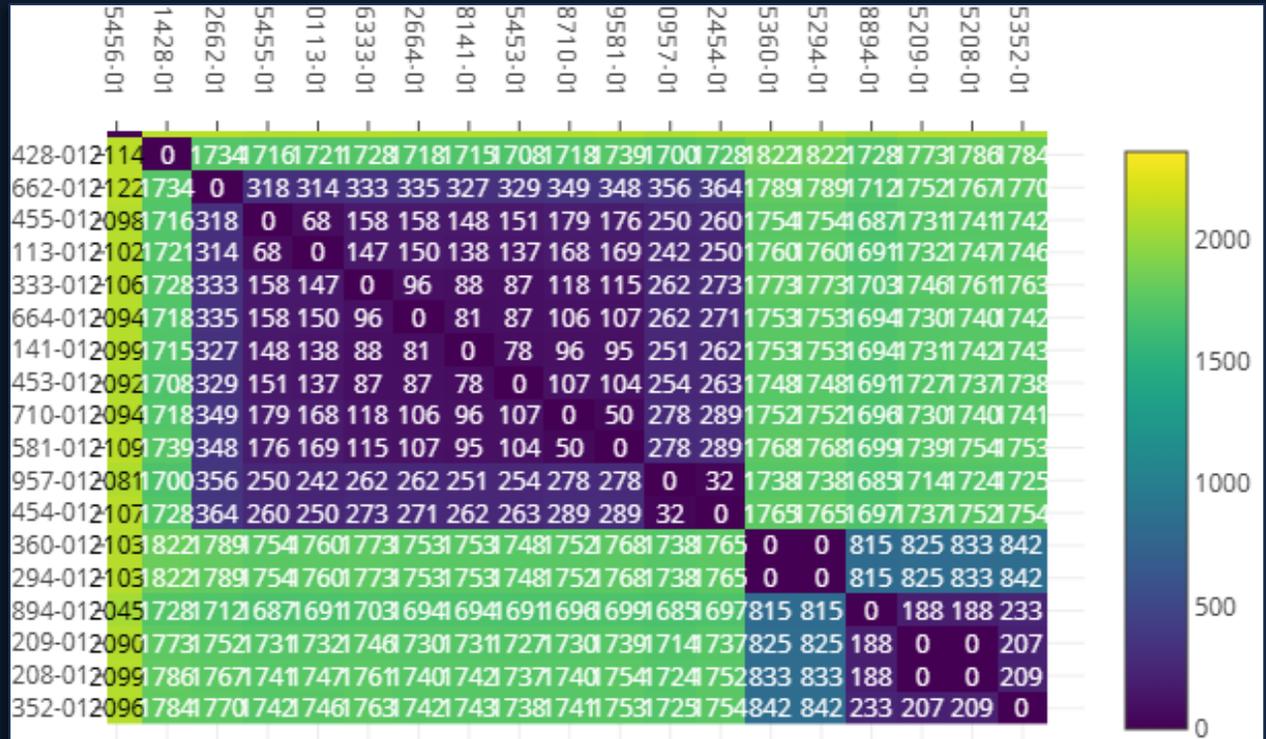
Investment

- Agency buy-in and license cost
- Time to build sufficient expertise
- Building end user data literacy and technical skill

Example #1: Dendrograms and Matrices



Created from PulseNet 2.0



Example #2: Allele Code Dashboard

Salmonella Allele Codes as of 2024

Top 10 Salmonella Allele Code Counts in Tennessee, 2022-2024
(Bionumerics Data)

SALM1.0 - 6743.2.4x	Enteritidis	27
SALM1.1-6.6.6.6x.8x	Enteritidis	21
SALM1.1-6.37.1036.1524.1892.2374	Enteritidis	7
SALM1.1-309.511.669.883.1087.1376	Africana	7
SALM1.0 - 6785.1.2x	I4:i:-	7
SALM1.1-76	Bareilly	6
SALM1.1-56.65.70.76.81.100	Sundsvall	6
SALM1.1-6.19.294.433.526.1110	Enteritidis	5
SALM1.1-22.140.166.192.821.15862	Typhimurium	5
SALM1.1-1672.4033.6559.10210.12477.17932	Newport	5

Most common Salmonella allele code, 2024
SALM1.1-6.6.6.6x.8x
9 cases

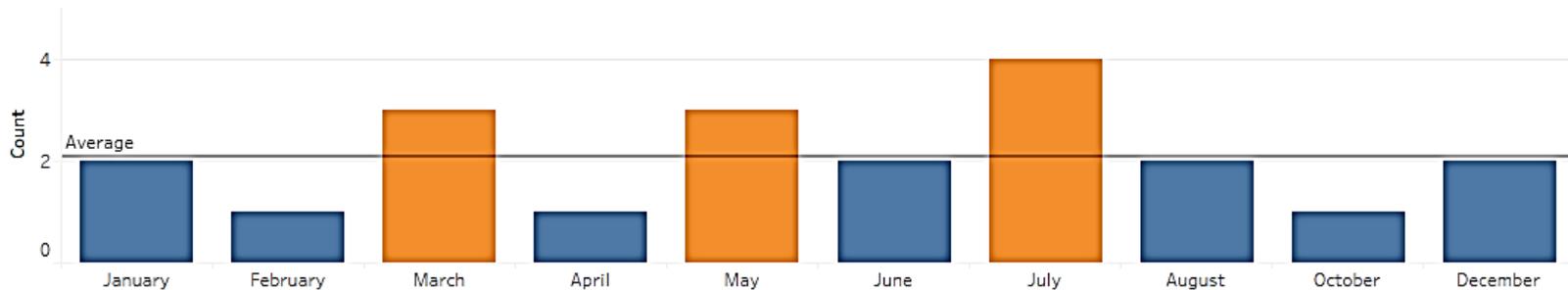
Top Allele Code (SALM1.0 - 6743.2.4x) by Year



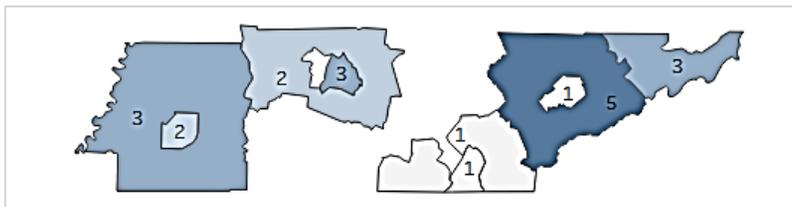
Select an allele code for the below charts:

SALM1.1-6.6.6.6x.8x

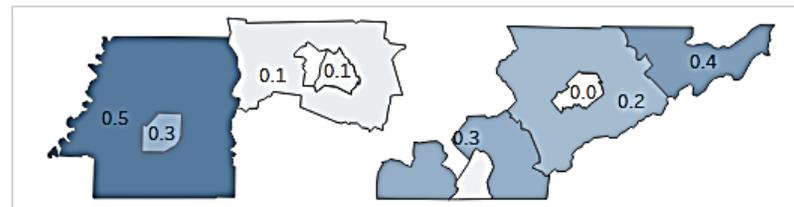
Salmonella (Allele Code SALM1.1-6.6.6.6x.8x) by Month of Received Date, 2024



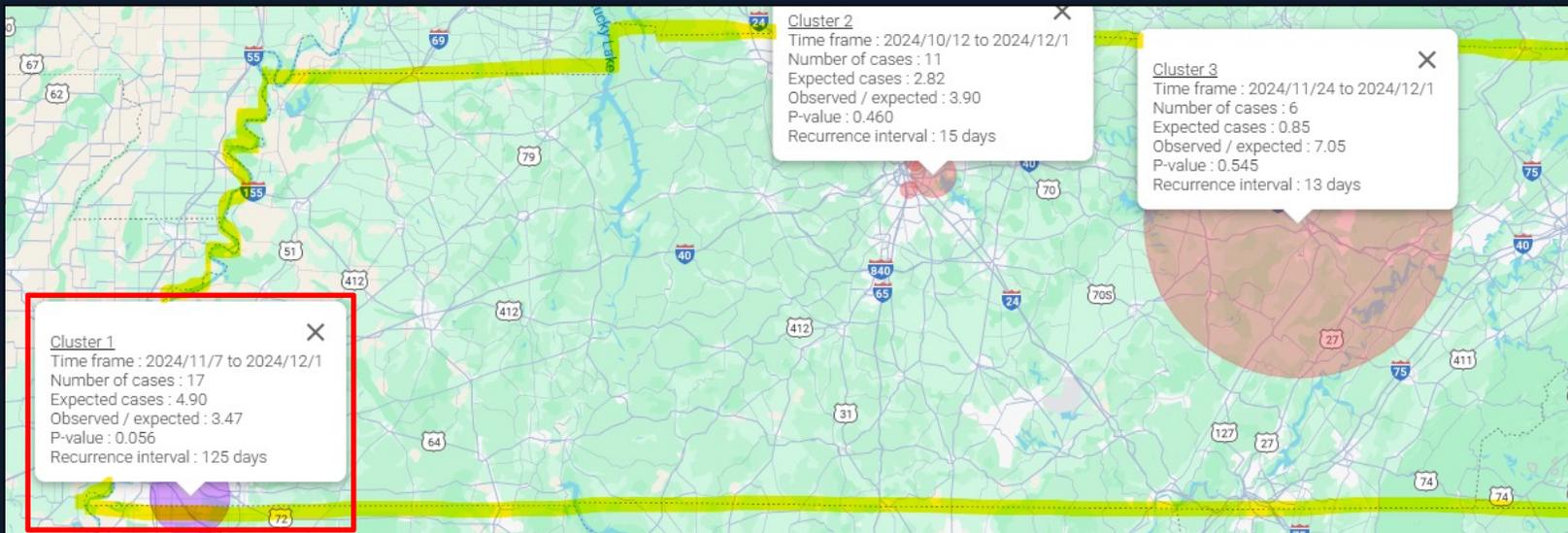
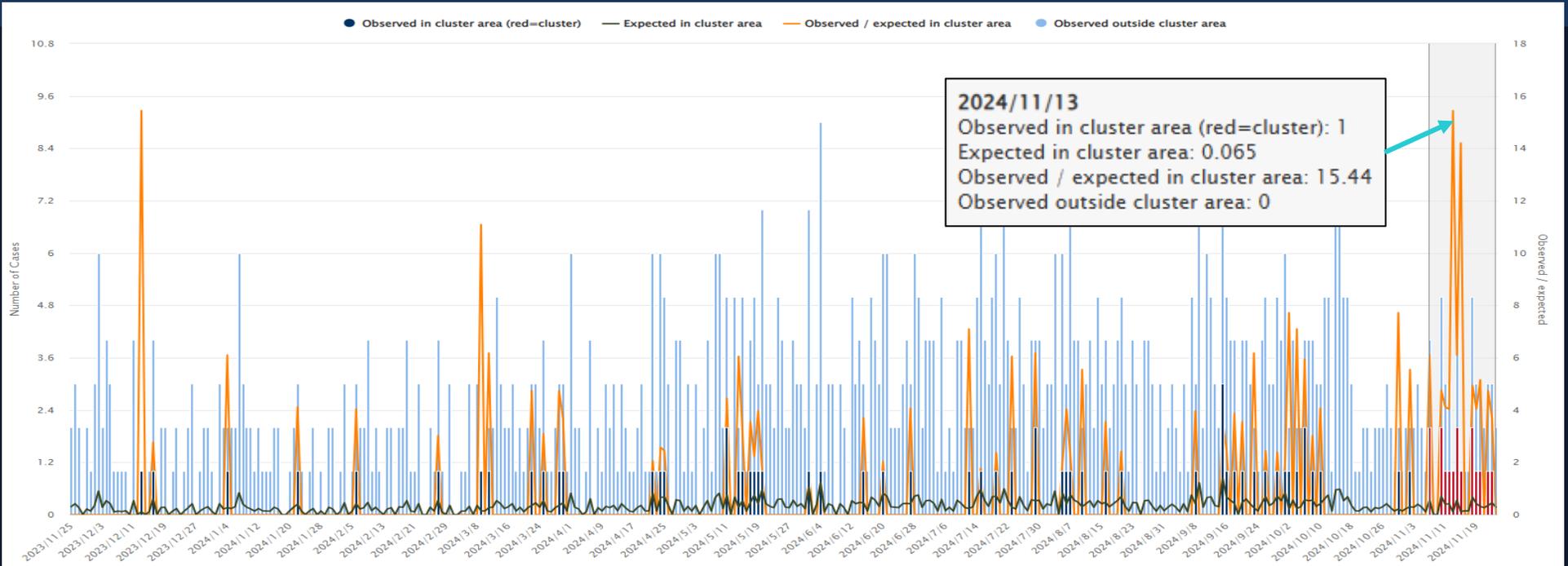
Salmonella Counts (Allele Code SALM1.1-6.6.6.6x.8x) by PHR, 2022, 2023, 2024



Salmonella Rates (Allele Code SALM1.1-6.6.6.6x.8x) by PHR, 2022, 2023, 2024



Example #3: SatScan



Example #4: Closeout Dashboard

Cluster Investigations with Incomplete Individual Case Records

CTDB_RedcapID	FC_ClusterID	Investig..	FC - Cluster To Outbre..	FC - Supp IX?	IX by TN	IX by CDC	IX Outcome
1082	1082	CAS1896..	Not answered	Null	0	0	Null
		CAS1898..	Not answered	Null	0	0	Null
		CAS1899..	Not answered	Null	0	0	Null
		CAS1903..	Not answered	Null	0	0	Null
1083	1083	CAS1899..	No	No	0	0	Null
		CAS1900..	No	No	0	0	Null
1085	1085	CAS1901..	No	No	0	0	Null
		CAS1901..	No	No	0	0	Null
		CAS1901..	No	No	0	0	Null
1086	Null	CAS1900..	No	Yes	0	0	Null
	1086	CAS1900..	No	Yes	0	0	Null
		CAS1901..	No	Yes	0	0	Null
1088	1088	CAS1901..	No	Yes	0	0	Completed
		CAS1903..	No	Yes	0	0	Null

Assign To	Investig..	Miss_Demographics	Miss_Symptoms	Miss_Hospitalizatio..	Miss_Date of Death	Miss_Daycare Detail
B.	CAS166..	1	0	0	0	0
	CAS167..	1	0	0	0	0
	CAS167..	1	0	0	0	0
M	CAS165..	1	0	0	0	0
	CAS165..	1	0	0	0	0
	CAS167..	1	0	0	0	0
P	CAS165..	0	0	0	0	1
	CAS166..	1	0	0	0	0
	CAS168..	1	0	0	0	0
W	CAS165..	1	0	0	0	0
	CAS166..	0	0	0	0	1
	CAS166..	1	0	0	0	0
	CAS166..	0	0	0	0	1
	CAS168..	1	0	0	0	0
	CAS168..	0	0	0	0	1

Interviews
with Missing
Data



Knowledge Check

Which of the following is NOT a challenge in Cluster/WGS Data Visualization?

- A. Ability for users to access visualizations
- B. Rapidly detecting commonalities
- C. Software security and cost
- D. Interpreting the data accurately

Salmonella I 4:i- Outbreak Example

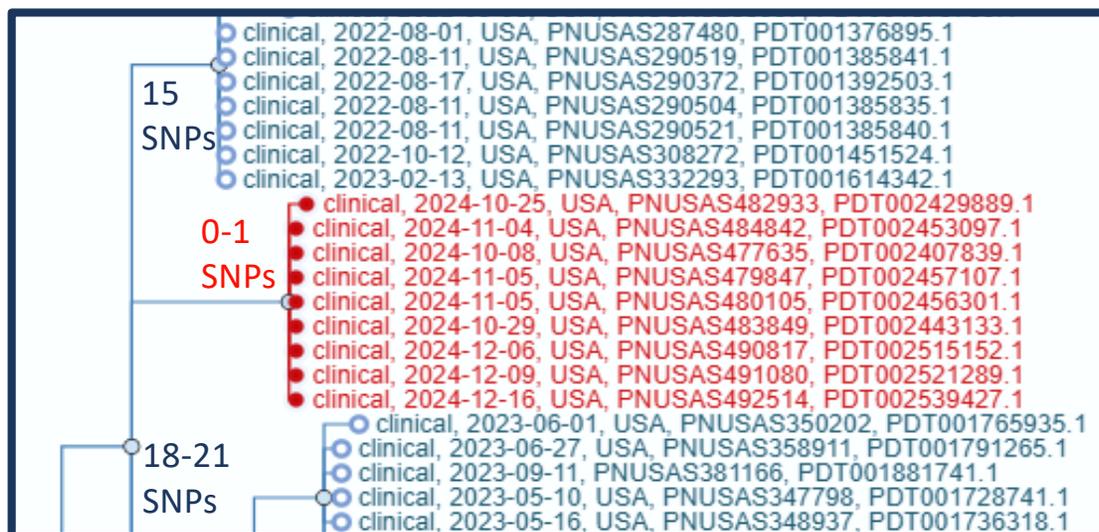
October 15, 2024 Cluster Dendrogram – First Identification

N24E30	PNUSAS479847	I 4:i:-	9/25/2024	10/2/2024	Stool	Human	Shelby	SALM1.1-13.57.232x.844.26767.36545
N24E30	PNUSAS480105	I 4:i:-	9/18/2024	10/2/2024	Stool	Human	Shelby	SALM1.1-13.57.232x.844.26767.36545

Initially lost to follow up; cluster closed

December 10, 2024 Cluster Dendrogram – Reopened with new cases; checked NCBI/SEDRIC

TN__N24E	PNUSAS490817	I 4:i:-	11/21/2024	11/26/2024	Stool	Human	Shelby	MALE	SALM1.1-13.57.232x.844.26767.36545
TN__N24E	PNUSAS483849	I 4:i:-	10/8/2024	10/18/2024	Stool	Human	Shelby	MALE	SALM1.1-13.57.232x.844.26767.36545
TN__N24E	PNUSAS491080	I 4:i:-	11/19/2024	11/27/2024	Stool	Human	Shelby	FEMALE	SALM1.1-13.57.232x.844.26767.36545



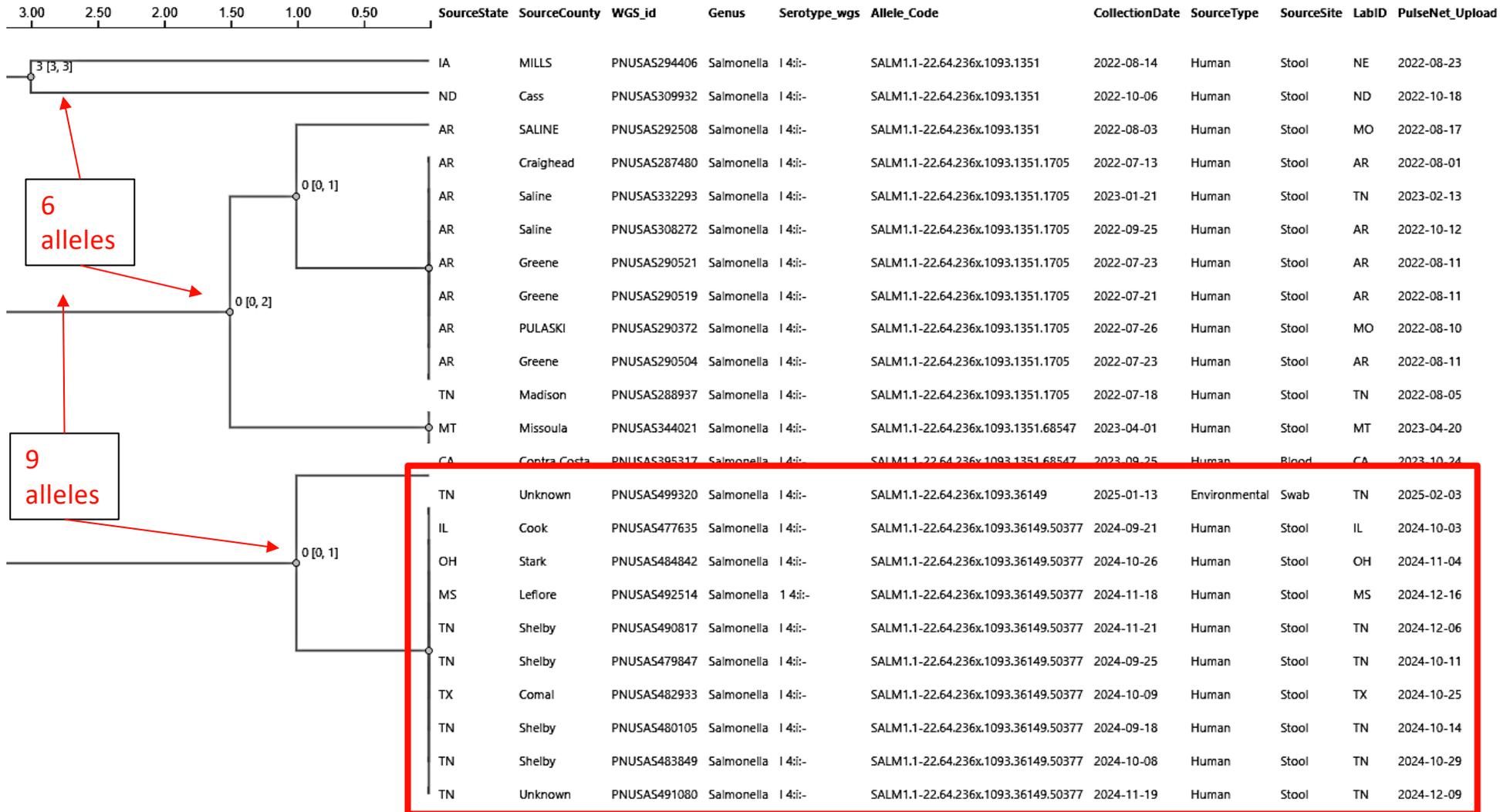
Exposure Review – Tableau Dashboard

Outcomes		
Deceased: 0/6	Hospitalized: 2/6	
Meat Exposures		
Beef exposure: 4/6	Egg exposure: 5/6	Chicken exposure: 4/6
Eat Pork: 3/6	Eat Turkey: 3/6	Eat Seafood: 0/6
Eat Raw Seafood: 0/6	Eat Fish: 1/6	Eat Raw Fish: 0/6
Animal Exposures		
Cow, Goat, or Sheep exposure: 0/6	Chicken or Turkey exposure: 0/6	Reptile or Amphibian exp: 0/6
Dog exposure: 4/6	Cat exposure: 1/6	Any exposure: 4/6
Travel Exposures		
Any travel: 0/6	International Travel: 0/6	Domestic Travel: 0/6

Water Exposures		
Drink Natural Water: 0/6	Well Water: 0/6	Swim in Natural Water: 0/6
Dairy Exposures		
Drink Raw Milk: 0/6	Eat Soft Cheese: 1/6	Eat Other Raw Dairy: 0/6
Fruit Exposures		
Eat Berries: 2/6	Eat Tropical Fruit: 2/6	Eat Cantaloupe: 1/6
Vegetable Exposures		
Eat Spinach: 1/6	Eat Onion: 1/6	Eat Pepper: 0/6
Eat Lettuce: 0/6	Eat Tomato: 2/6	
Other Exposures		
Daycare: 1/6	Foodhandler: 0/6	

Daycare As..	Foodhandle..	Eat Out Cha..	Daycare De..	Daycare Na..	Rest Detail
No or unknown	No or unknown	Yes	Null	Null	M, P&S, BBB , S
No or unknown	No or unknown	Yes	Null	Null	HB, F, CC, BBB
Yes	No or unknown	Yes	five days/week	Null	H
No or unknown	No or unknown	No or unknown	Null	Null	
No or unknown	No or unknown	Yes	Null	Null	BBB , EM, CFA, JM,HH
No or unknown	No or unknown	Yes	Null	Null	SFT, BBB

Final Dendrogram





PulseNet 2.0

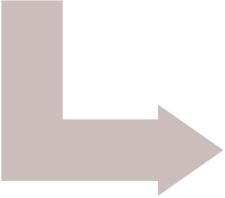
Benefits of Using PN2.0

- When compared to SEDRIC, PN2.0 can:
 - Create dendrograms and matrices to visualize clusters more clearly
 - Show all isolates and clusters at once
 - Find *Shigella* clusters and *Campylobacter* clusters other than *jejuni*
 - Immediately show when a new isolate is added to a cluster
- Epi access can allow for workload-sharing and streamline synthesis with epi/clinical data
- Internally consistent review by Epis can help:
 - Determine when epi data supports expanding the time range
 - Weed out household clusters
 - Track clusters that may close and reopen frequently

Steps to Access PN2.0 for Epidemiologists

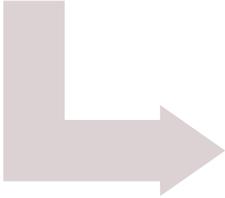
Request Access

- Ask a laboratorian to email pulsenet@cdc.gov
- Request Read Only Certification



Complete Agreements

- Need SAMS access
- Need PulseNet Sharepoint Access
- Data Use Agreement



Pass Quiz

- Review user guide and practice in test environment
- Complete open book quiz

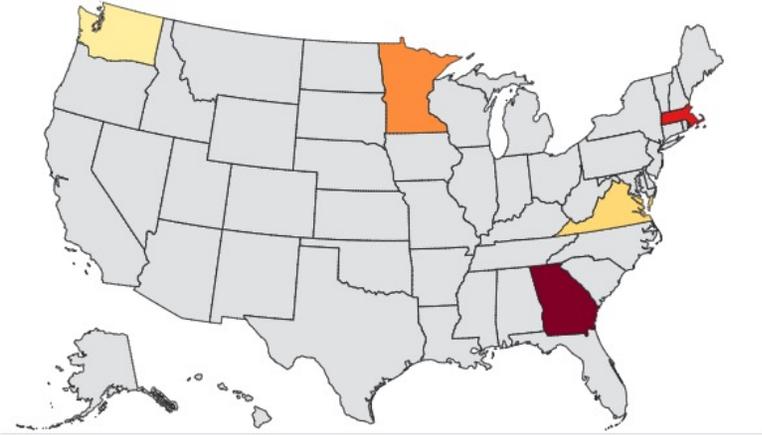


Consult SPHL

- Laboratorians can train on current process
- Can create new or share existing views

Other AMD Resources

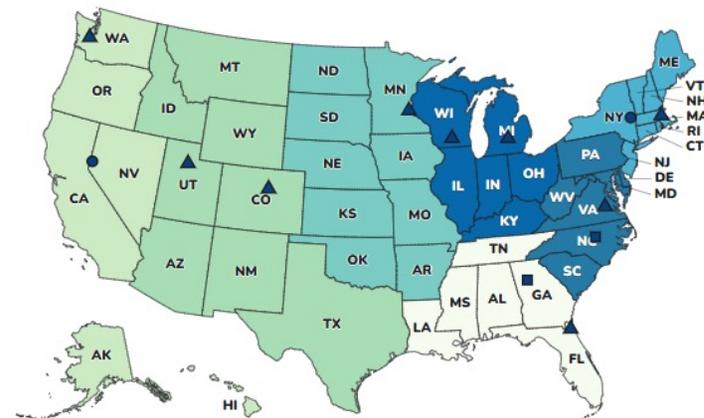
Pathogen Genomics Centers of Excellence



Principal Investigators

- Georgia Department of Public Health
- Massachusetts Department of Health
- Minnesota Department of Health
- Virginia Division of Consolidated Laboratory
- Washington State Department of Health

Bioinformatics Regional Resource Centers (BRRs)



U.S. territories

- AS
- GU
- MP
- PR
- VI

Freely associated states

- FM
- MH
- PW



Knowledge Check

Epis must have SAMS access and pay for a software license to access PulseNet 2.0

A. TRUE

B. FALSE



Communicating Results

Communicating WGS Results

Discussion:

What communication challenges have you faced around genetic results during outbreak investigations?

Common Challenges with Communication

Scientific
jargon

Delay in or
lack of lab
results

Different
access to
data/PHI

Interpreting
Uncertainty

Potential Solutions

Scientific
jargon

Delay in or
lack of lab
results

Different
access to
data/PHI

Interpreting
Uncertainty

Challenge Scenario #1

Sequence data suggests relatedness, but Epi/EH data is missing or does not align

- Cases in animal contact outbreaks with clinical isolates matching environmental isolates, but cases do not report exposures to these sources
- Stealth ingredient implicated but not identified
- Complaint EH investigation finds violations unrelated to likely transmission of the implicated pathogen

Knowledge Check

WGS Links with No/Unsupportive Epi/EH Data

Which of the following would best help address this challenge?

- A. Plain language description of why highly related sequencing data alone is a reliable link
- B. Describing all the reasons epi and/or EH data could be biased or unreliable
- C. Communicating next steps that could address these challenges
- D. A and C
- E. All of the above

Challenge Scenario #2

No/unclear genetic relatedness with clinical or environmental samples, but linked to a source due to epidemiological information

- Cyclosporiasis cases linked to a common ingredient or restaurant
- *Salmonella* outbreak cases reporting the same Mexican-styled restaurant but different serotypes, no environmental samples to link clinical samples
- Wide allele range due to length of outbreak or setting/transmission mode
- Environmental/food sample testing is negative

Knowledge Check

No/Unclear WGS Linkage

Which of the following could help address this communication challenge?

- A. Describing in concise, plain language how EH and Epi data supports the hypothesis and can “rule out” other sources
- B. Explaining why specimens may cluster more loosely or not at all for some serotypes or transmission modes than others
- C. Blame lack of cooperation from ill people or facility for lack of good lab data

Developing Standard Language/Comms

- Plain Language Interpretation of WGS
 - Less detail:
 - *“Through lab testing known as whole genome sequencing, we compared the type of Salmonella bacteria that each case had and found they were highly genetically related to each other. This means it is very likely these infections came from the same source.”*
 - For the more curious:
 - *“Whole genome sequencing lets us compare all of the Salmonella DNA between cases. As Salmonella bacteria replicate, they can acquire mutations to their DNA. If we have two Salmonella DNA sequences that are exactly identical to each other, the bacteria likely had less time to replicate and mutate as they infected people and made them sick. If they had less time to mutate, it’s very likely the bacteria that made people sick came from the same source.”*
- Use available resources for standard language
 - CDC/FDA Outbreak Notices
 - Previous press releases (or work with Comms team to develop language)
- Expectation Setting
- Know Your Audience

Salmonella I 4:i- Outbreak Example Wrap-Up

- Weekly written situation reports with updated dendrograms/interpretations shared with investigators
- Shared understanding and language established within the outbreak team on outbreak calls/communications
- Communication with restaurant ownership done with local points of contact, including their routine inspector
- Developed a final written report including clear interpretation of results in the discussion

Final
Report
Language

“Ten clinical specimens and sixteen environmental samples were collected during this outbreak. All clinical specimens were culture positive for Salmonella I 4:i- and genetically indistinguishable by whole genome sequencing. Of the sixteen environmental samples collected from the restaurant, one was positive for Salmonella I 4:i- and genetically matched the clinical samples, i.e., the outbreak strain causing human illnesses.”

Questions?