



EDLB Metagenomics Pilot Projects

Andrew Huang, PhD

Enteric Diseases Laboratory Branch, DFWED/NCEZID/CDC

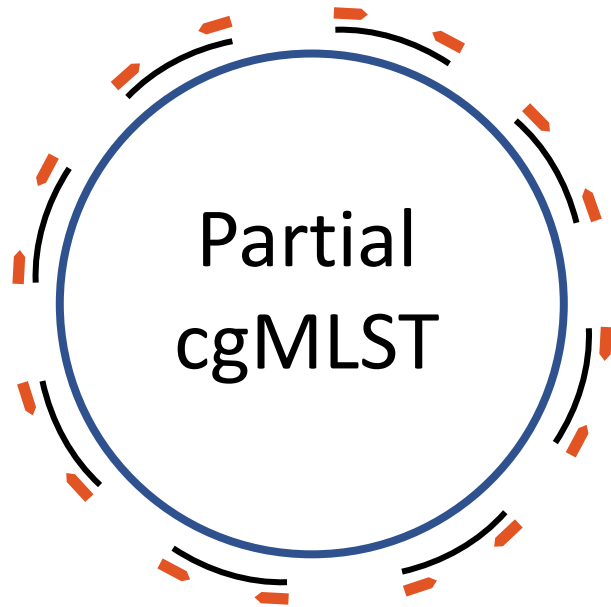
APHL ID Lab Con

13 Mar 2023

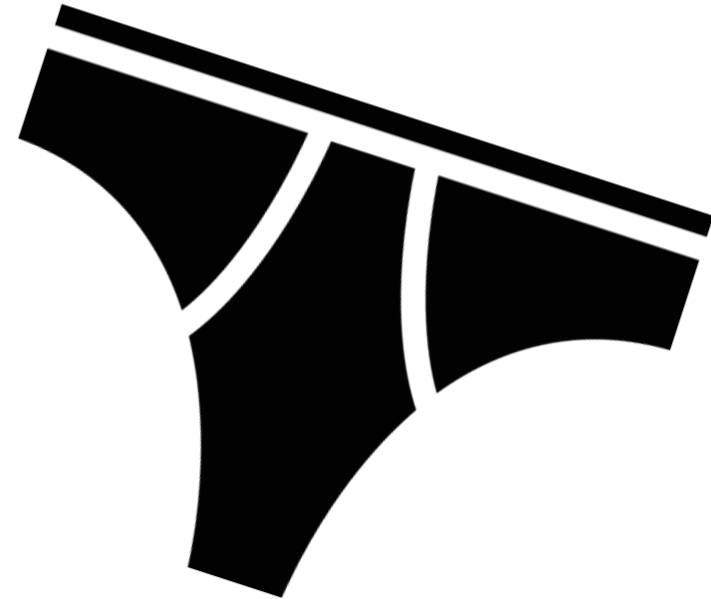
How Do We Detect Clusters Using CIDT + Specimens?



Metagenomics!

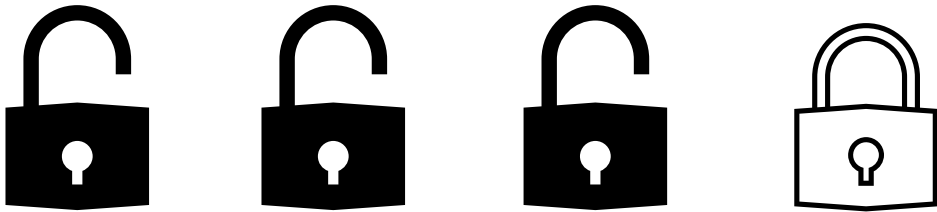


Enteric Surveillance Subtyping
Jo Williams (IGY7@cdc.gov)



Undiagnosed Diarrheal Illnesses
(UNDIs)

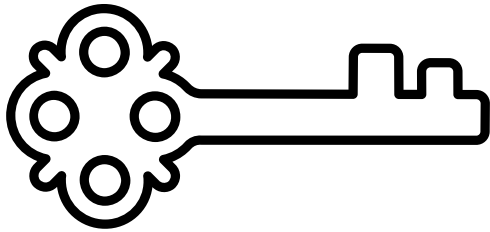
A gap in foodborne outbreak investigation



One in four foodborne outbreaks are caused by unknown agents^{[1](#)}

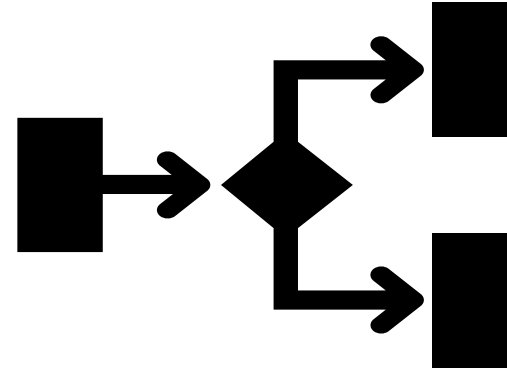
- Current culture dependent methods and culture-independent diagnostic tests do not identify novel pathogens
- CIFOR guidelines offer stool collection guidelines for **outbreaks of unknown etiology**, but no pathway for analyzing samples

The UNDI pipeline the key to unlocking OUEs



UNDIs is an NGS-based **pathogen discovery pipeline** for outbreak samples with strong epidemiological links but negative or inconsistent CIDT results.

It builds on **existing work** with both highly multiplexed amplicon sequencing and shotgun metagenomics.

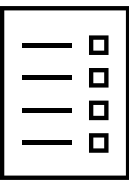


Tier one is a **rapid, inexpensive** HMAS panel, deployable at PHLs

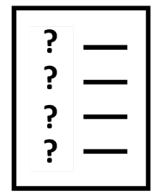
Tier two is a pathogen-agnostic metagenomic shotgun approach designed to **identify shared genomic material** from a small number of outbreak samples



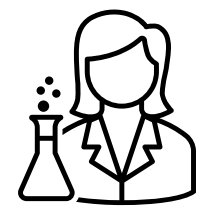
Collection
Collect and extract DNA from multiple outbreak stool samples



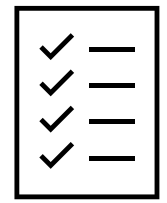
UNDIs Tier 1: HMAS
At local PHLs the UNDIs known pathogen HMAS panel is run on Standard Biotoools Juno or X9



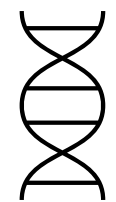
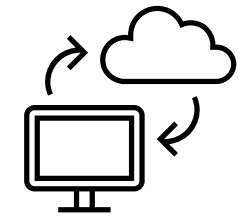
HMAS Negative
UNDIs HMAS panel comes back negative for common diarrheal pathogens



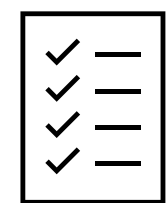
Metagenomic sequencing



HMAS Positive
UNDIs panel is positive for known diarrheal pathogen and directs lab to appropriate workflows



UNDIs Tier 2: metagenomic workflow
UNDIs metagenomic workflow employs a combination of reference-based and reference-free analyses to bin by species



Known pathogen
UNDIs identifies known pathogen that has evaded current CIDTs/UNDIs HMAS



Pathogen discovery
UNDIs identifies an unknown or novel pathogen

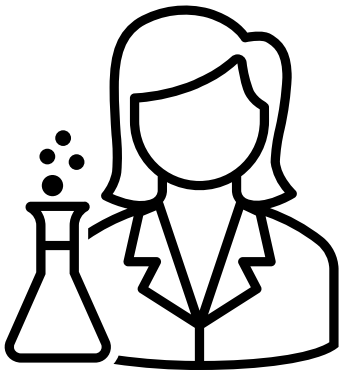
Tier 1: UNDI known pathogen HMAS panel

- *Salmonella*
- *Escherichia/Shigella*
- *Norovirus, Sapovirus, Rotavirus, Adenovirus, Astrovirus*
- *Vibrio*
- *Campylobacter*
- *B. cereus*
- *S. aureus*
- *Plesiomonas shigelloides*
- *Arcobacter butzleri*
- *Aeromonas*
- *Poliovirus*
- *Hepatitis A, Hepatitis E*
- *Saffold virus, Parvovirus*
- *Parasites*



Standard Biotools X9

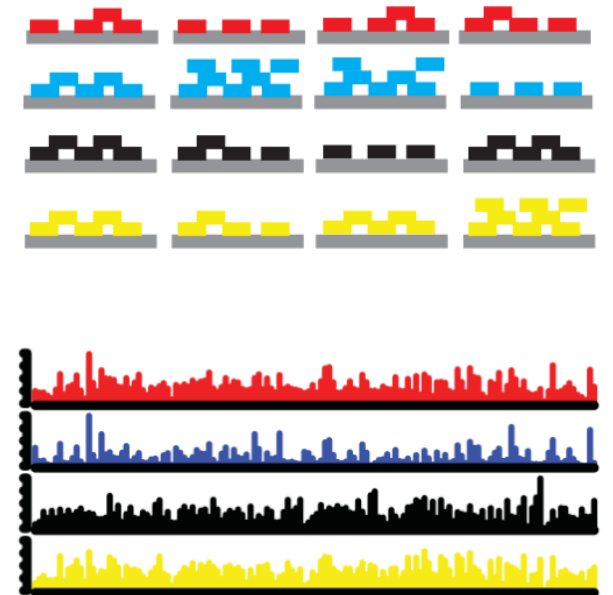
Tier 2: Binning Bioinformatics, Sorting by Species



Metagenomic sequencing



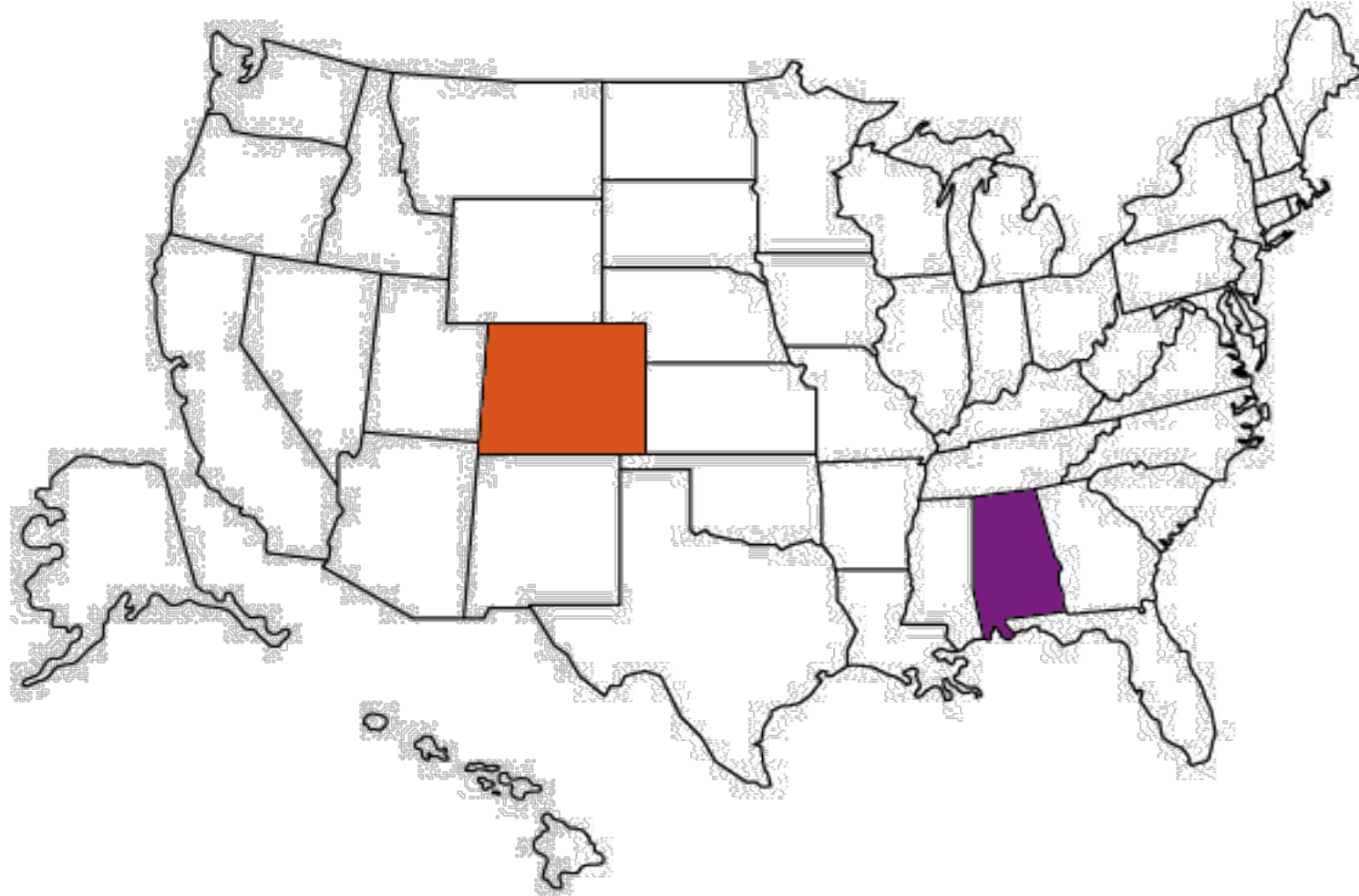
Reference based
MIDAS2



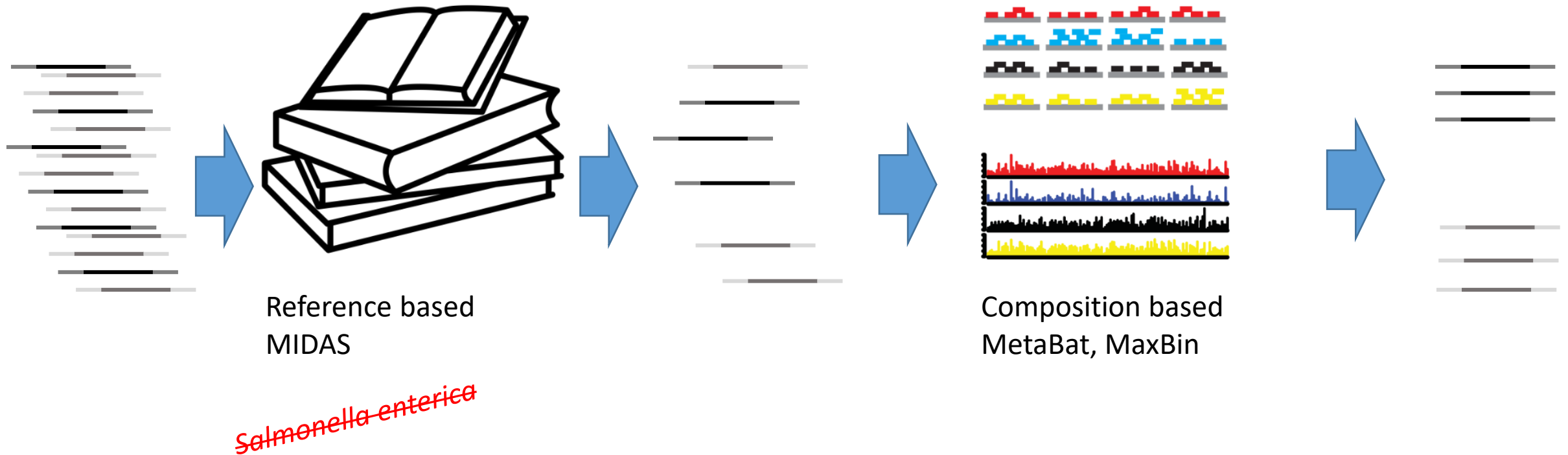
Composition based
MetaBat, MaxBin

Tier 2 simulated: Take two known outbreaks

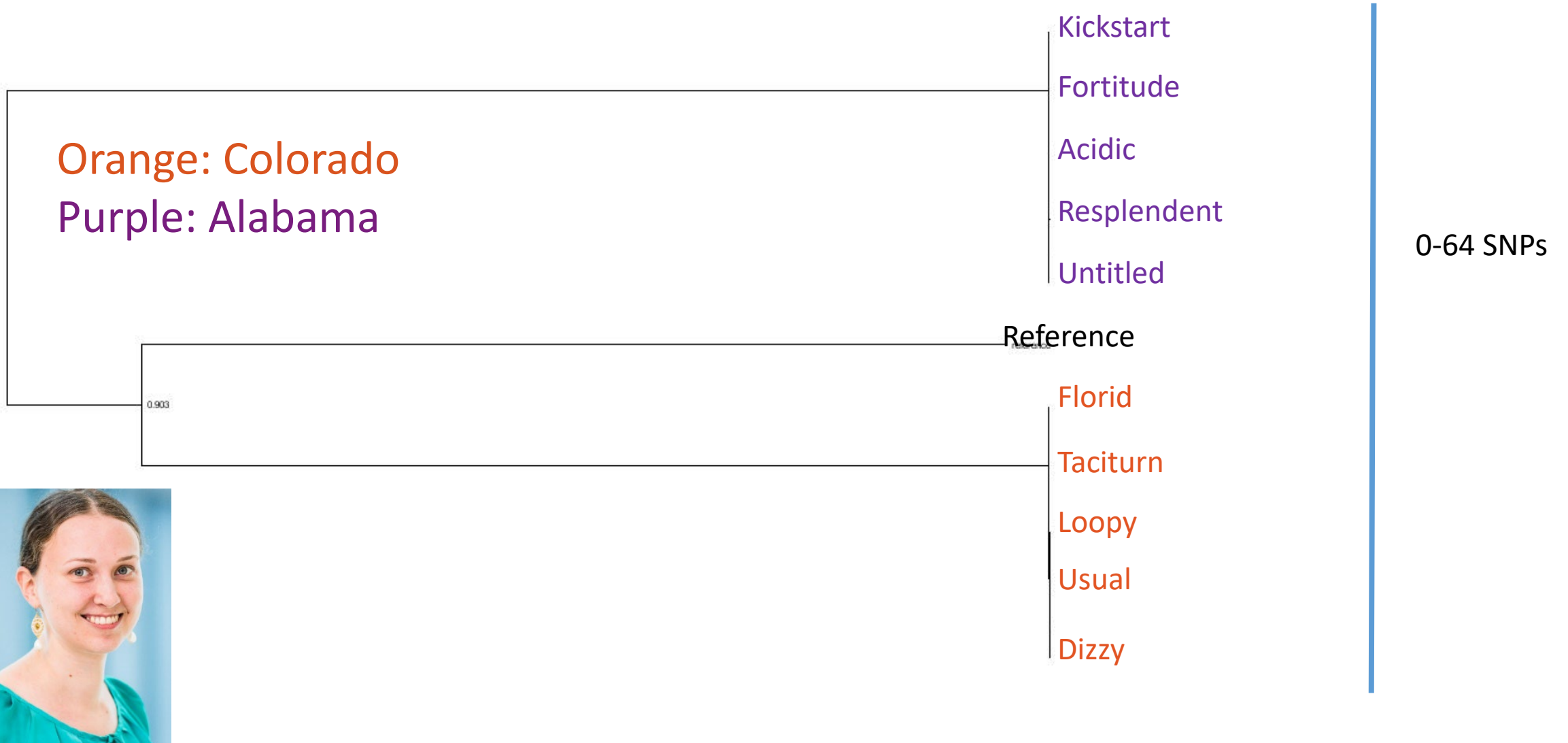
- Two similar 2013 *Salmonella enterica* outbreaks in Colorado and Alabama
- Epi and isolate NGS later showed to be unrelated



Tier 2 simulated: Forget *Salmonella*, bin for discovery

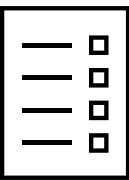


Tier 2 simulated: Discover *Salmonella*, try to make a tree

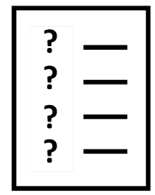




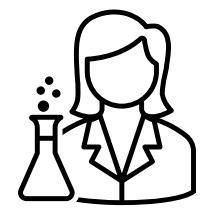
Collection
Collect and extract DNA from multiple outbreak stool samples



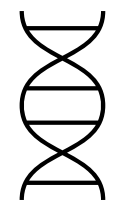
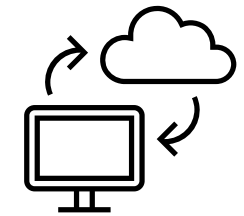
UNDIs Tier 1: HMAS
At local PHLs the UNDIs known pathogen HMAS panel is run on Standard Biotoools Juno



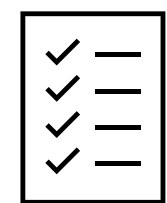
HMAS Negative
UNDIs HMAS panel comes back negative for common diarrheal pathogens



Metagenomic sequencing



UNDIs Tier 2: metagenomic workflow
UNDIs metagenomic workflow employs a combination of reference-based and reference-free analyses to bin by species

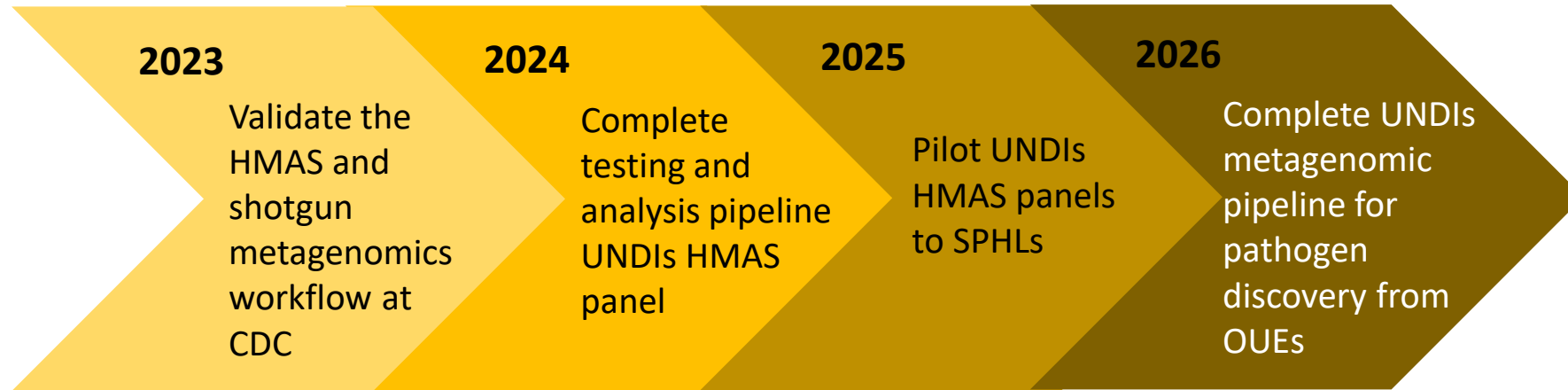


Known pathogen
UNDIs identifies known pathogen that has evaded current CIDs/UNDIs HMAS



Pathogen discovery
UNDIs identifies an unknown or novel pathogen

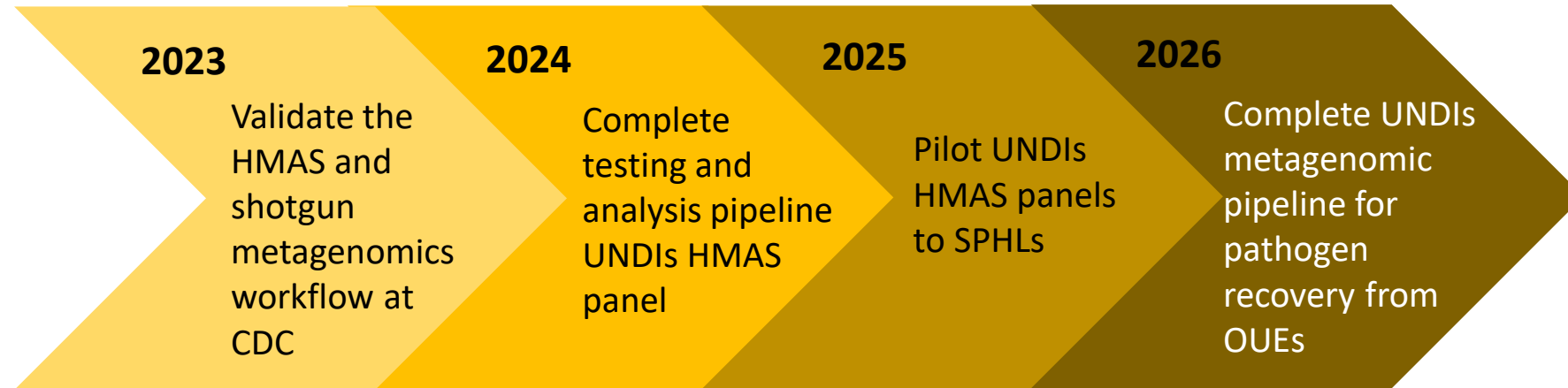
The UNDIS roadmap for pathogen discovery



2023 Roadmap

- Design and validate known pathogen HMAS panel
- Engage epidemiological support
- Collaborate with state public health partners to collect foodborne outbreak stool samples (>2)
- Evaluate shotgun metagenomics at SPHLs

UNDIS project looking for collaborators



Right now: Looking for collaborators to support UNDIS pipeline development by collecting outbreak stools of known and unknown cause.

- Collaborate with us to develop new methods and research unknown etiology outbreaks.
- ELC proposal guidance and APHL fellows available.

Andrew Huang
Ahuang1@cdc.gov

Questions?



Contact

Jo Williams (IGY7@cdc.gov)

Andrew Huang (AHuang1@cdc.gov)

For more information, contact CDC
1-800-CDC-INFO (232-4636)
TTY: 1-888-232-6348 www.cdc.gov

The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

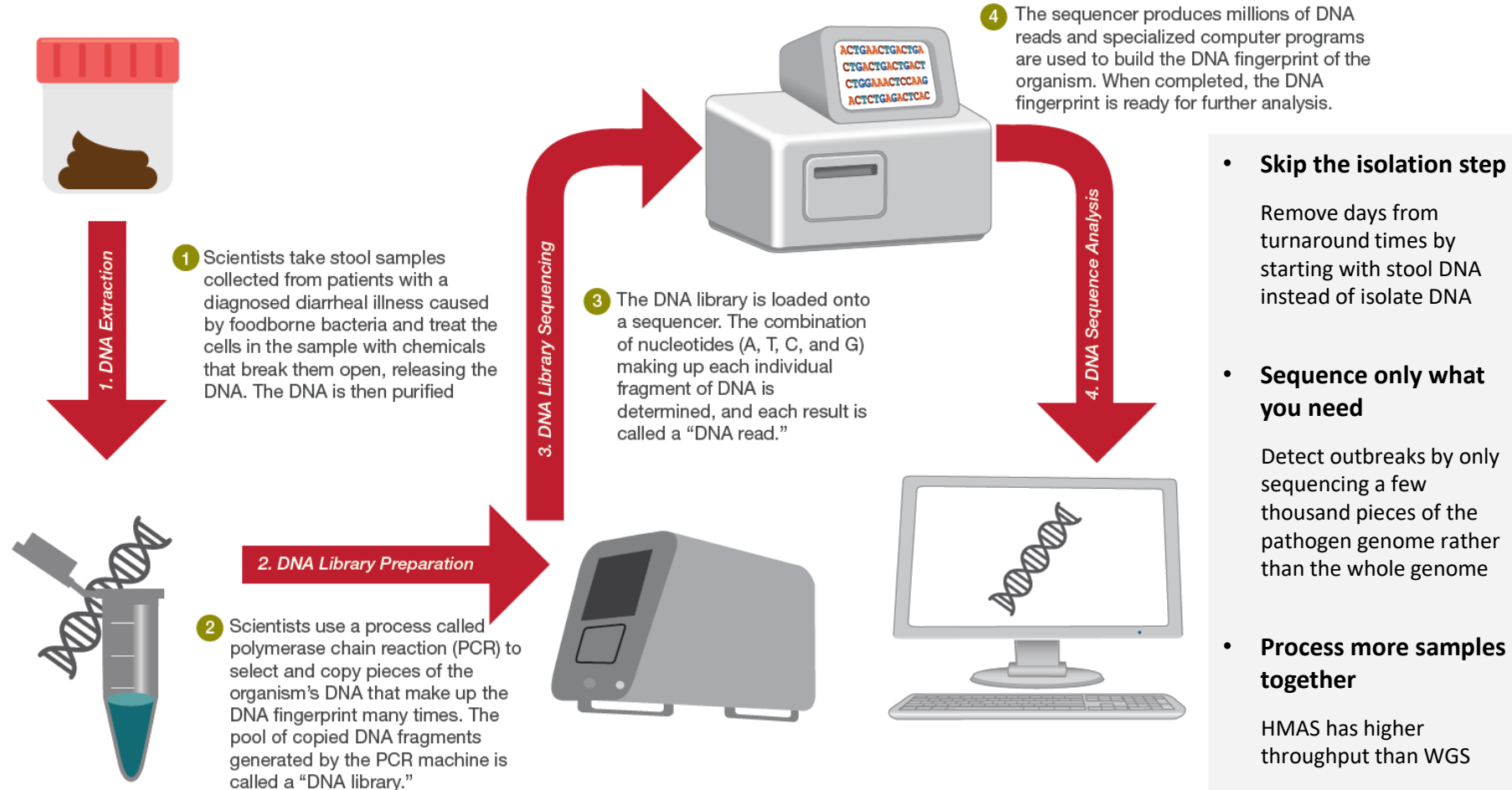


Juno HMAS for Subtyping

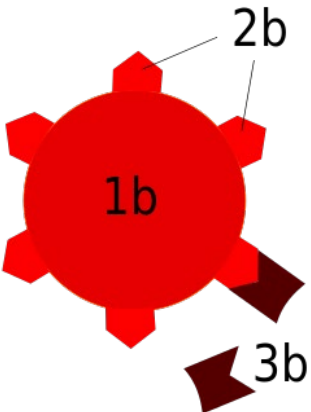
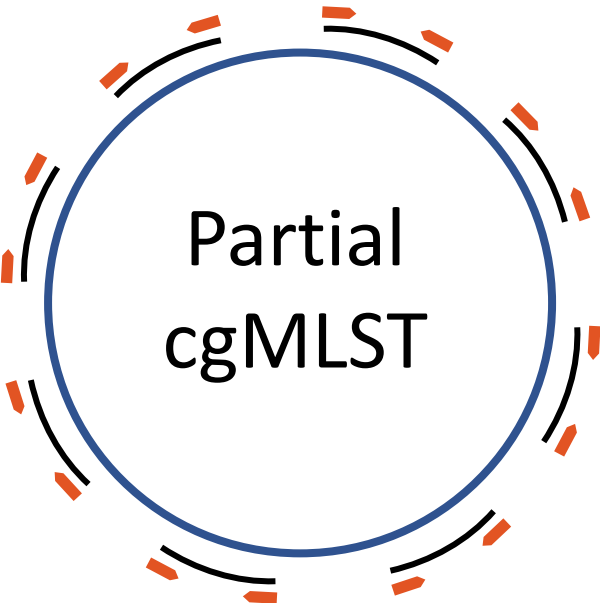
HMAS for Routine Surveillance

The Highly Multiplexed Amplicon Sequencing (HMAS) Process

HMAS is a laboratory procedure that determines the order of bases in the genome of an organism at specific locations that are a part of the organism's DNA fingerprint. These DNA fingerprints can help link cases together allowing an outbreak to be detected and solved sooner. The HMAS procedure is still under development and is expected to become available soon.



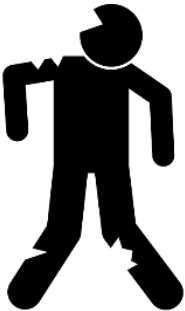
HMAS Subtyping Targets



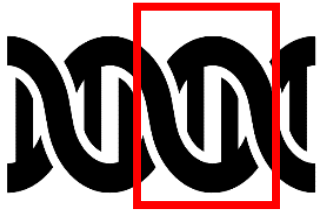
Serotyping



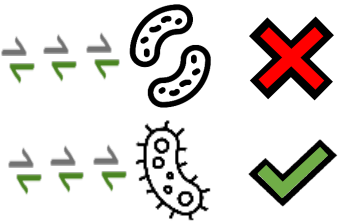
Antimicrobial resistance



Virulence



Short Informative Regions



Pathogen Specific



1000's of Targets

Pilot Sites and Progress

Initial Goals and Participants

- **MN and CO are pilot sites now**
 - Funding through ELC and APHL
 - APHL fellows
- **First assay: *Salmonella* subtyping**
- **Goals**
 - Testing workflow, protocols, compatibility with PHL needs
 - Validating output by comparing isolate WGS, isolate HMAS, and stool HMAS for same sample
 - Cluster detection validation TBA



Standard BioTools Juno Microfluidic Thermocycler

Instrumentation

Juno (current)

- Standard microfluidic thermocycler
 - Function: produce amplicons



X9 (possible future)

- Real time microfluidic thermocycler
 - Function: produce and quantify amplicons



X9 “Fit for Purpose” Testing

- Found a X9 going on loan to a collaborator at Georgia Tech
- Planning 2 initial test runs beginning next week
 - 48.48 IFCs
 - 2461 cgMLST panel for *Salmonella*
 - Bacterial isolates, mixed isolates, isolate-spiked “healthy” stools
 - Replicates of 2 runs planned for Juno later this month for comparison
- Data will help support decisions to upgrade existing pilot sites and start new sites with X9
- Additional applications of real time and library prep functions for enterics under consideration

PulseNet Outbreak Surveillance

HMAS Deployment Timelines

FY23	FY24	FY25
<i>Salmonella</i> assay Support for 2 existing pilots Support for 3 new pilots possible	<i>Salmonella</i> assay (10 SPHLs) STEC assay pilot begins Command line analysis pipeline pilot begins	<i>Salmonella</i> assay (20 SPHLs) STEC assay (20 SPHLs) Analysis pipeline in PulseNet 2.0