

State Hygienic Laboratory

Whole genome sequencing - past, present, and future of foodborne outbreak investigations

Valérie Reeb, Ph.D.

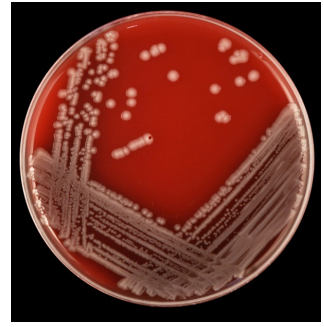
Assistant Research Scientist, SHL Molecular division

Iowa Food Safety Task Force Meeting

April 11, 2023

Overview

- Outbreak source tracking
- Whole genome sequencing basic
- Public sequence database
- Iowa outbreak investigations
- Antibiotic resistance
- Culture independent subtyping
- SHL capabilities

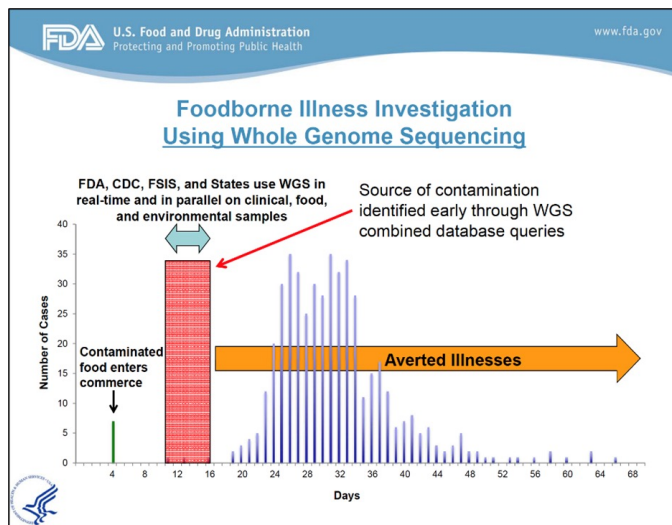


Outbreak Source Tracking



→ Traditional foodborne Illness Investigation

- Identify pathogen in clinical samples
- Case investigation by epidemiologist
- Track food source between cases
- Identify pathogen in food samples



→ Whole Genome Sequencing Outbreak Investigation

- FDA GenomeTrakr Network
- CDC PulseNet

Outbreak Source Tracking

Step 1. Collect pathogen samples from

- Sick people (medical professionals)
- Food (FDA, USDA, states and local agencies)
- Productions facilities, restaurant, farms, etc. (Federal, State and Local investigators)

Step 2. Sequence pathogens whole genome

Step 3. Compare Genomic Sequences

Action

- Partners work to prevent more illnesses
- Determine how the pathogen got into the food supply chain

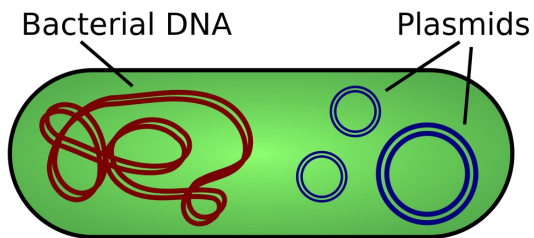


<https://www.fda.gov/media/111130/download>

Genomic data

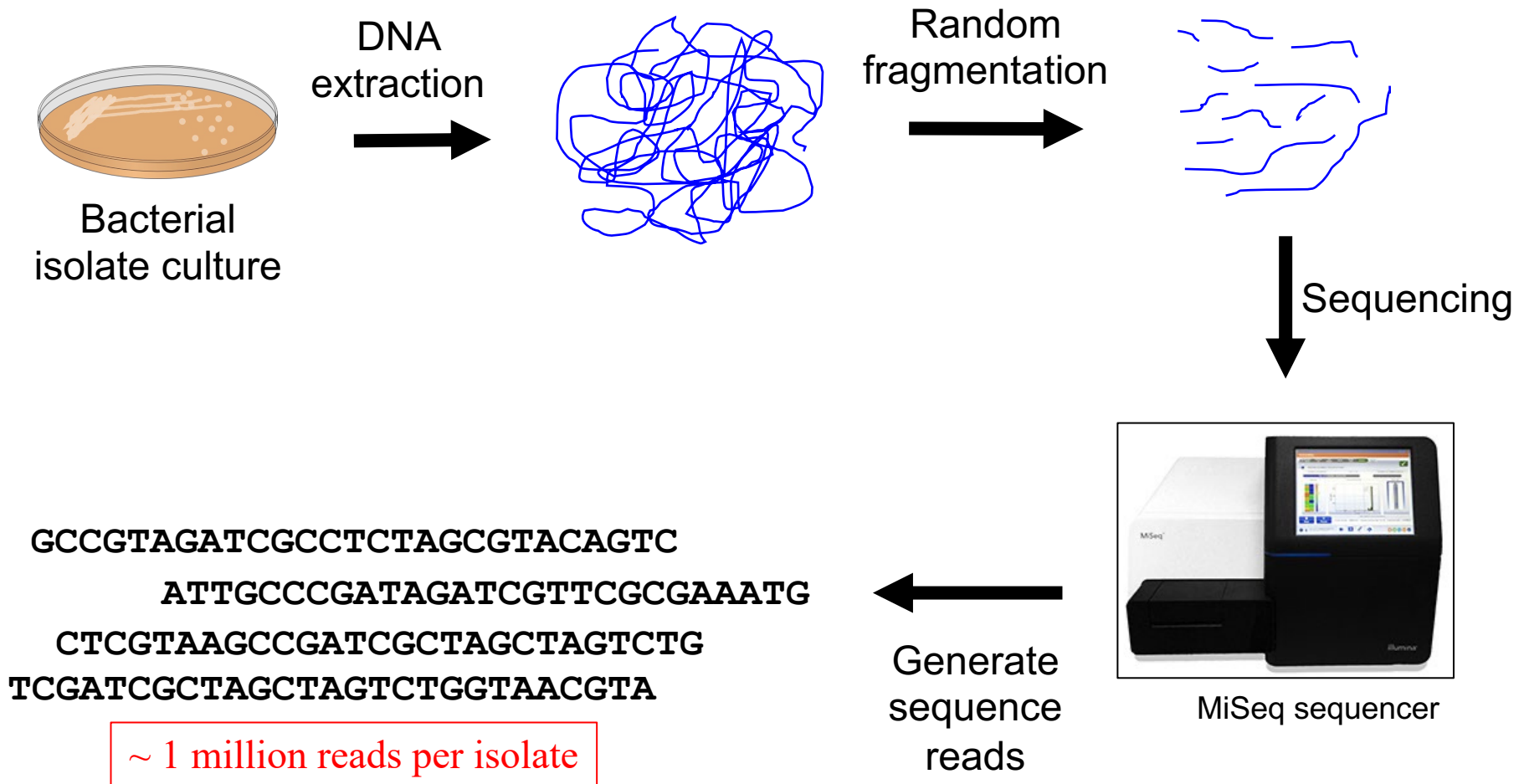


Each bacteria can be identified by its unique DNA sequence



One long sequence of 130,000 – 15,000,000 nucleotides (A, T, C, G)

Whole Genome Sequencing (WGS)

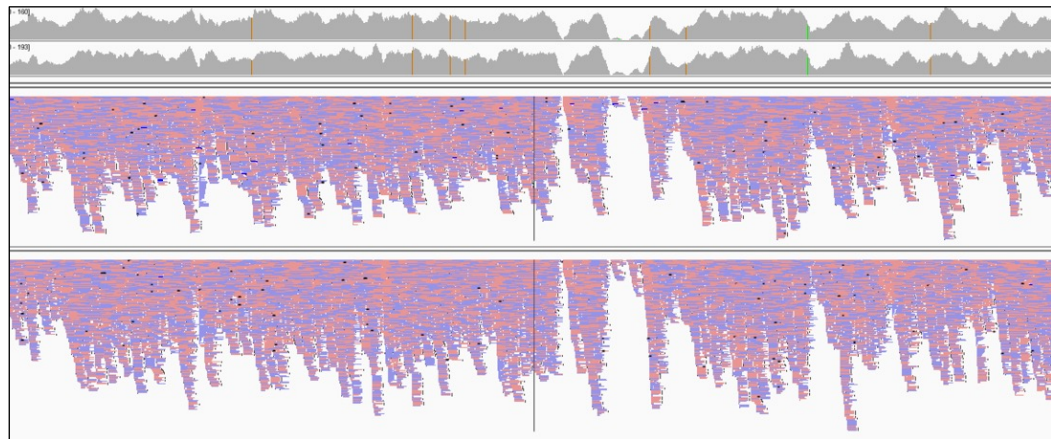


Reads assembly

Re-create the whole bacterial DNA sequence

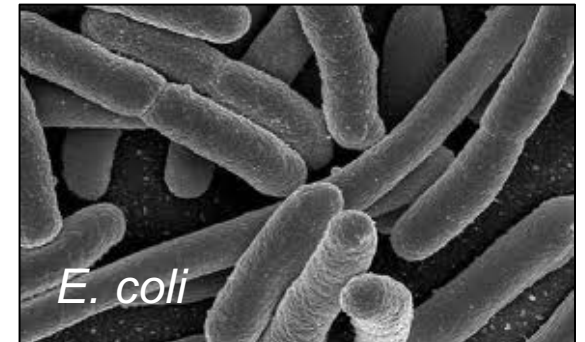
```
AGGTTTCGATGCTAGTGCCTTA
  TTCGATGCTAGTGCCTTAGAT
    CTAGTGCCTTAGATCGATCG
      TAGTGCCTTAGATCGATCGG
        CGATCGGTAGCCAAGTCGCT
          CCAAGTCGCTAGATTCGGAT
            AGTCGCTAGATTCGGATAAC
              CGCTAGATTCGGATAACCTG

AGGTTTCGATGCTAGTGCCTTAGATCGATCGGTAGCCAAGTCGCTAGATTCGGATAACCTG
```



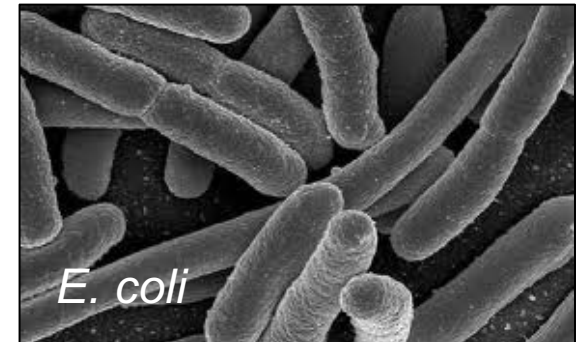
Genome comparison

Listeria : AGGT**T**CGATGCTAAGTGC**A**GTACCGTATAACG...
Salmonella: AGGTCCG**G**TGCTAAGTGC**A**GTACCGTATAACG...
E. Coli : AGGTCCGATGCTA**T**GTGCCGTACC**C**TATAACG...
Patient : AGGT**T**CGATGCTAAGTGC**A**GTACCGTATAACG...



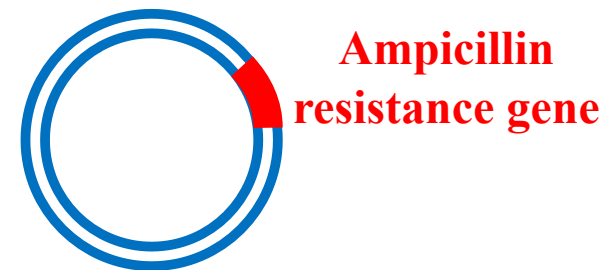
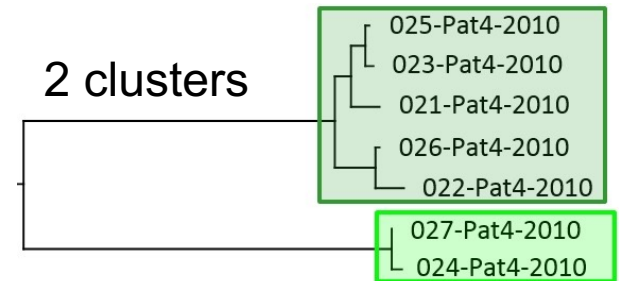
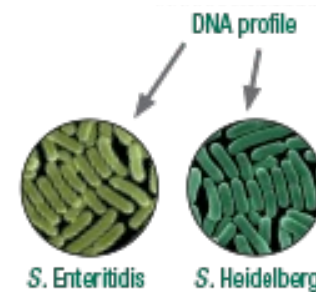
Genome comparison

Listeria	AGGT T CGATGCTAAGTGC A GTACCGTATAACG...
Salmonella	AGGTCCG G TGCTAAGTGC A GTACCGTATAACG...
E. Coli	AGGTCCGATGCTA T GTGCCGTACC C TATAACG...
Patient bacteria	AGGT T CGATGCTAAGTGC A GTACCGTATAACG...



Advantages of WGS

- Improved identification of bacteria
 - e.g., Salmonella Serotyping
- Better discrimination between isolates
- Improved outbreak detection
 - Link clinical cases
 - Find source earlier
 - Solve more outbreaks while still small
- Identification of antimicrobial resistance genes and virulence factors



WGS limitation



What is the
source of
contamination
?

NCBI Pathogen Detection database

- Sequences repository for foodborne pathogens
 - 1.4 Millions sequences (182,000)
 - 78 species
- Associated metadata
 - Geographical location
 - Source
 - Environmental or clinical
- Link isolates from food or the environment to human illness
- Larger database = faster source detection

Species	Isolate sequenced
<i>Salmonella enterica</i>	526,819
<i>E.coli and Shigella</i>	302,223
<i>Staphylococcus aureus</i>	95,579
<i>Campylobacter jejuni</i>	85,184
<i>Klebsiella pneumoniae</i>	77,527
<i>Listeria monocytogenes</i>	56,413
<i>Streptococcus pneumoniae</i>	38,139
<i>Enterococcus faecium</i>	28,526
<i>Neisseria gonorrhoeae</i>	26,978
<i>Pseudomonas aeruginosa</i>	25,139
<i>Clostridioides difficile</i>	23,909
<i>Acinetobacter baumannii</i>	23,737
<i>Streptococcus pyogenes</i>	20,471
<i>Vibrio parahaemolyticus</i>	10,082

As of April 6, 2023

NCBI Pathogen detection database

Cluster results within 1 week of submission



PDT000002130.2 | SAM... [icons]

dist: 6
target_acc: PDT000002130.2
target_creation_date: 2015-02-11
Run: SRR1068291
sra_release_date: 2014-01-03
LibraryLayout: PAIRED
Platform: ILLUMINA
sra_center: CFSAN
fullasm_id: 0
complete_fl: 0
asm_stats_length_bp: 5113245
asm_stats_contig_n50: 767722
asm_stats_n_contig: 51
bioproject_acc: PRJNA186035
bioproject_title: GenomeTrakr
Project: US Food and Drug Administration
bioproject_center: Center for Food Safety and Applied Nutrition
biosample_acc: SAMN02345006
attribute_package: Pathogen: environmental/food/other
collected_by: FSW
collection_date: 2010-11-08
geo_loc_name: Indonesia
isolation_source: frz raw black tiger shrimp
lat_lon: missing
sample_name: FSW0076
serovar: Weltevreden
strain: FSW0076
sub_species: enterica
HHS_region: 0
tax-id: 59201
species_taxid: 28901
scientific_name: Salmonella enterica subsp. enterica

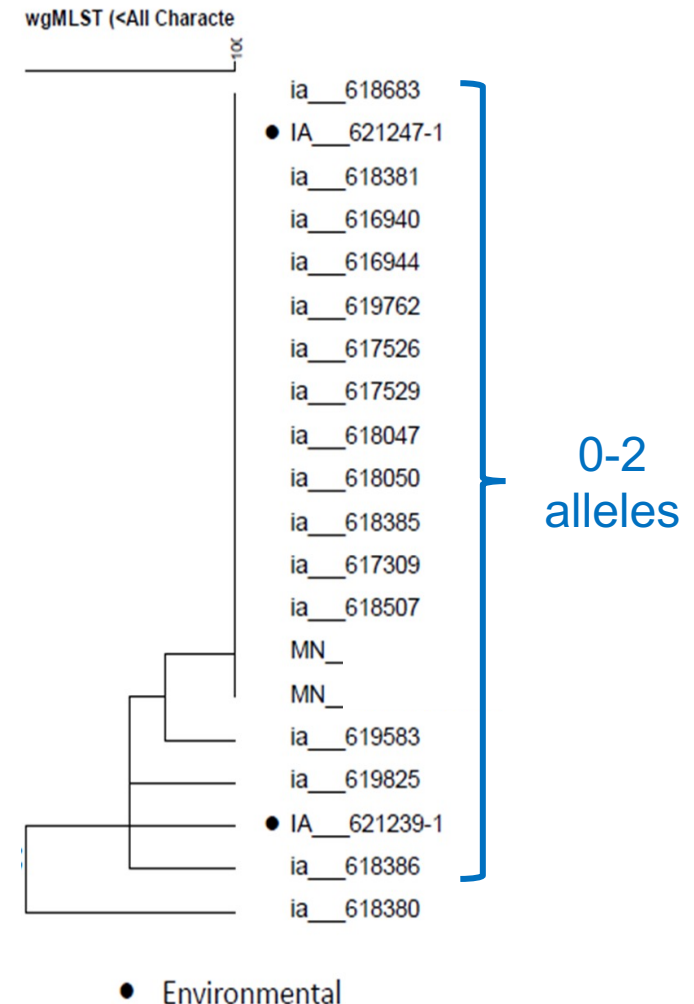
2018 Multistate outbreak of *Salmonella* Typhimurium in Chicken Salad

→ Outbreak summary

- 264 cases 8 states (240 in IA)
- 87% reported eating chicken salad

Timeline

- Increased # of Salmonellosis cases
- 4 patients reported eating chicken salad from the same grocery store
- Voluntary recall of the chicken salad
- Manufacture producing chicken salad was identified
- Chicken salad from 7 stores tested at SHL
- 2 presumptive positives IDed by PCR
- 60 samples collected at the manufacture (only FDA areas) – no positives
- Both positives from the chicken salad and clinical isolates closely related by WGS



2019 *E. coli* O103 traced to Sprout

→ Outbreak summary

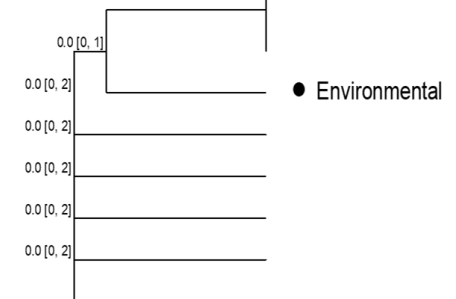
- 22 cases from 10 Iowa counties



→ Timeline

- Outbreak identify through routine surveillance testing on 12/18/2019
- Sub-sandwich franchise owner and sprout owner are notified
- Spent irrigation water and 16 clover sprout samples were collected
- All clinical and environmental samples were identified as *E. coli* O103 by PCR
- WGS shows close relationship of two environmental and clinical samples
- Related to an *E. coli* O103 multistate outbreak in Feb 2020

0-2 alleles

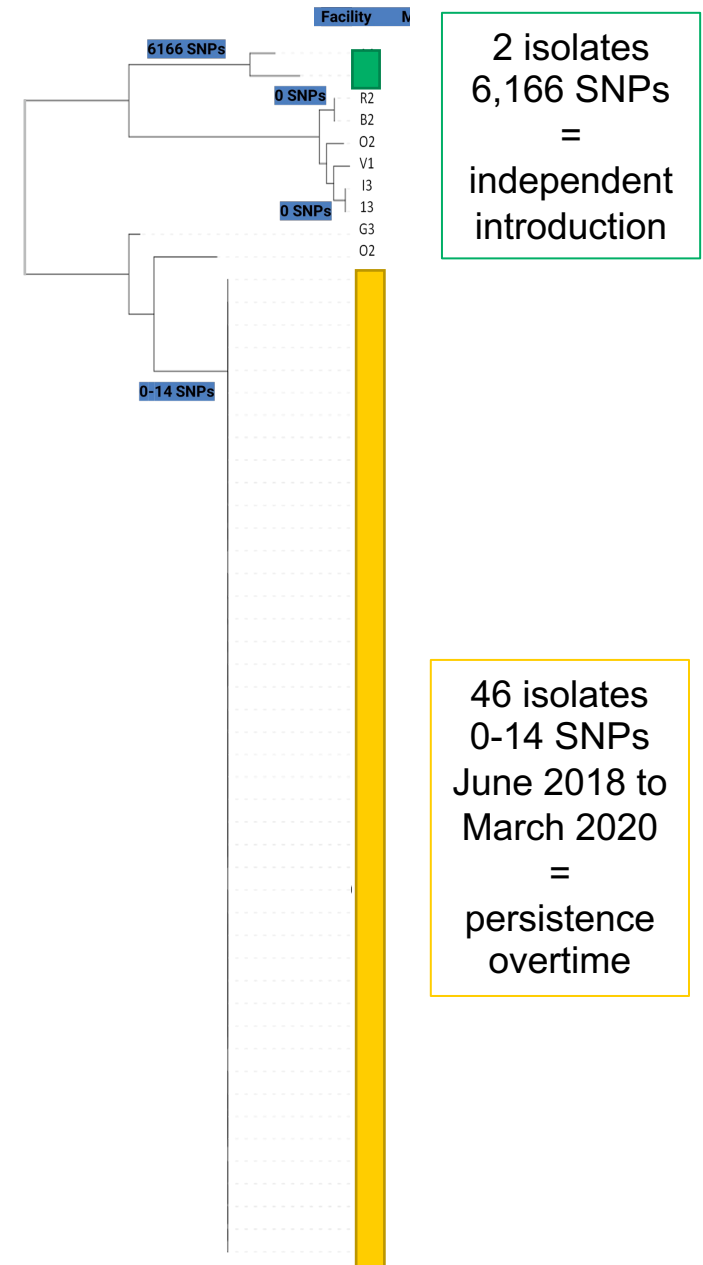


Persistence overtime

→ Environmental *Listeria* in Iowa

- 3,394 environmental samples collected between 2013-2020
- 18 facilities
- 15.8% positivity (136 samples)
- WGS for 56 isolates
- 9 sequence types identified

WGS help us distinguish between repeated introductions versus persistent overtime



Antibiotic resistance pathogens in food

→ Statistics

- 2.8 millions AR infections per year
- 1 in 4 infections caused by food and animals
- 35,000 deaths

→ Spread of AR bacteria in food

- Consumption of animal products
- Animal feces in surrounding environment
- Fruits and vegetables in contact with soil, water or fertilizer containing animal feces

→ Strategic approach to reduce infection

- Develop and enhance tools for conducting surveillance
- Expand tracking of environmental sources contributing to spread (One Health)
- Promoting responsible use of antimicrobials
- Development of new products (new antimicrobials, vaccines, etc)



Antibiotic resistance pathogens in food

→ Automatic detection in clinical and environmental samples

- Pathogen Detection
- AMRFinderPlus tool
- Identify antimicrobial genes from pathogen genome sequence

→ Wastewater surveillance

- Target specific AR genes

 National Library of Medicine
National Center for Biotechnology Information

Pathogen Detection BETA

Complete (3)

acrF

blaEC

mdtM

Point (2)

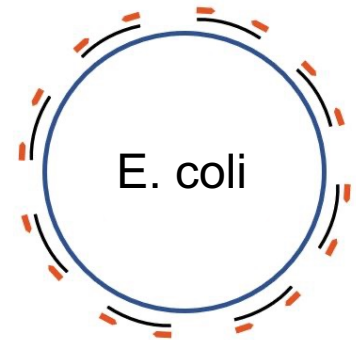
glpT_E448K

pmrB_Y358N

Culture Independent Subtyping

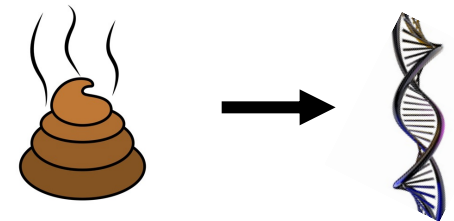
→ Identify pathogens directly from samples (fast TAT)

- Amplicon sequencing (target specific)
 - Targeted pathogens identification
 - Targeted pathogen serotype
 - Targeted AR and virulence genes
 - Low cost
- Shotgun metagenomics (Sequence everything)
 - All pathogens sequenced
 - All AR and virulence genes sequenced (but origin unknown)
 - Advantageous for discovery
 - Expensive



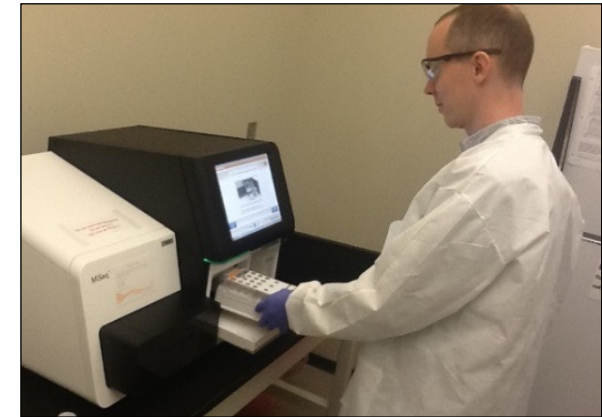
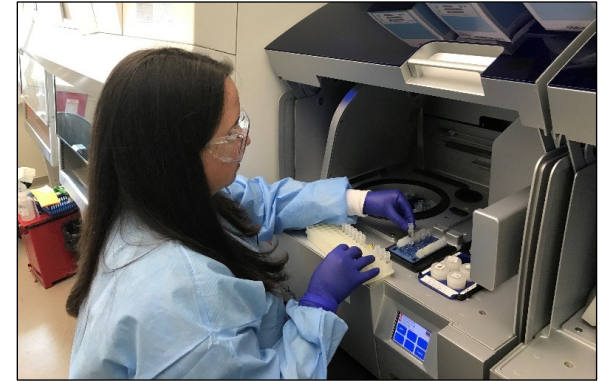
→ CDC is piloting Salmonella subtyping from stool

- Highly Multiplex Amplicon Sequencing (HMAS)



SHL Capabilities

- Over 9,000 bacteria and 20,000 SARS-CoV-2 viruses since 2015
- Bacteria sequenced: Foodborne pathogen, *Legionella*, Carbapenem-resistant bacteria, *Staphylococcus*, *Streptococcus*, *Mycobacterium*, *Neisseria*, etc.
- Eight technicians trained for genome sequencing
- Dedicated spaces for sequencing, different sequencing instruments (Illumina MiSeq, Nextseq, Oxford Nanopore MinIon) and Liquid handler (EpMotion, Clear DX)
- Bioinformatics pipelines installed on the UI HPC



Acknowledgment

SHL genomic team

- Jeff Benfer
- Kris Eveland
- Wesley Hottel
- Murphy Keller
- Gina Kline
- Erik Twait
- Anna Yakos
- Arielle Watson
- Kristen Zanon

SHL Environmental Micro

- Ryan Jepson
- Cathy Lord

Epidemiology support

- Johnson County Public Health Dept.
- Iowa Dept. of Public Health.

Environmental sampling

- Johnson County Public Health Dept.
- Iowa Dept. of Inspection and Appeals
- Iowa Dept. of Agriculture

Funding and support

- CDC (ELC)
- APHL
- FDA Genome Trakr
- University of Iowa CHEEC

IOWA

State Hygienic Laboratory

Thank you