Transitioning Public Health Microbiology to Whole Genome Sequencing: Experiences and Plans for Bacterial Foodborne Pathogens

Peter Gerner-Smidt, MD ScD, Chief
Enteric Diseases Laboratory Branch

2015 APHL Annual Meeting
Indianapolis, IN, May 19, 2015

National Center for Emerging and Zoonotic Infectious Diseases
Division of Foodborne, Waterborne, and Environmental Diseases

Would You Keep Doing This?

CDC Specimen ID: **2014003970** CDC Unique ID: **N8K7DBC1** CDC Local ID: **2014C-3008**

GENUS/SPECIES: *Escherichia coli*

SEROTYPE: O104:H4*

PATHOTYPE: Shiga toxin producing and Enteroaggregative *E. coli* (STEC % Eagg Ec)

VIRULENCE PROFILE: *stx2* –pos (subtype stx2a), *eae-negative, EhxA-negative, Eagg plasmid-positive*

ANTIMICROBIAL RESISTANCE: Ampicillin, Cefoxitin, Ceftriaxone, Streptomycin, Tetracycline, Sulfamethoxazole/Trimethoprim

Comments:

***Disclaimer** - This test has not been cleared or approved by the FDA. The performance characteristics have not been fully established. The results of this test should NOT be used for the diagnosis, treatment, or assessment of patient health or management.

Explanantion of Virulence Markers

stx1, *stx2*, *eae*, and *ehxA* are virulence markers present in E.coli O157 and other Shiga toxin-producing *E.coli*. These organisms are important causes of hemolytic uremic syndrome (HUS), hemorrhagic colitis and diarrhea worldwide.

EAgg plasmid is a genetic marker for E.coli that has a distinctive aggregating pattern of adherence on HEp-2 cells in vitro. Enteroaggregative *E.coli* (EAgg Ec) have been associated with persistent childhood diarrhea and have been isolated from travelers and persons with AIDS and diarrhea.

Approved by: Nancy Strockbine, Ph.D.
Ph: 404-639-4186
Fax: 404-639-3333
E-mail: nas6@cdc.gov

Biochemical 'panel'

O and H agglutination

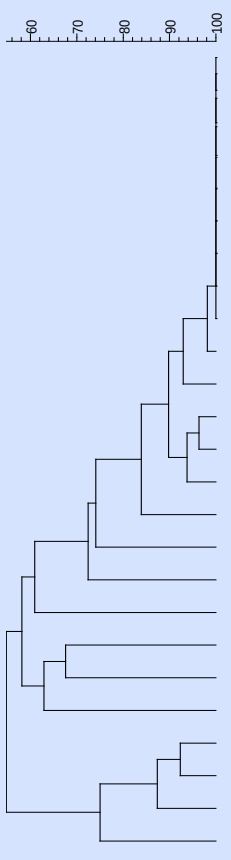
Min. 3 PCRs + RFLP

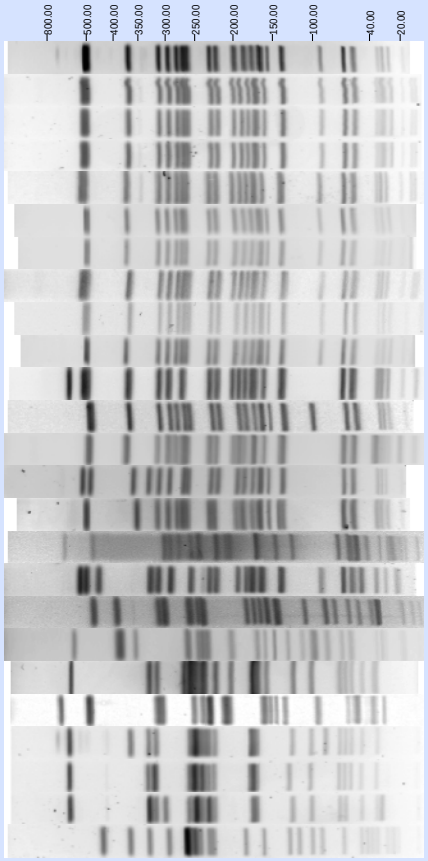
Disc diffusion OR broth micro dilution

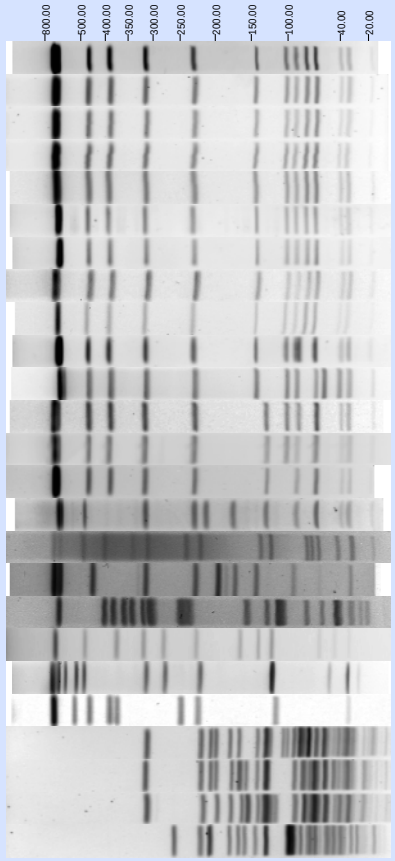
TAT: 1- 2 weeks

And This?

PFGE is high-discriminatory but NOT phylogenetically relevant







PFGE-XbaI-pattern	PFGE-BlnI-pattern	Country
EXAX01.0003	EXAA26.0003	
EXAX01.0003	EXAA26.0003	
EXAX01.0003	EXAA26.0003	
EXAX01.0003	EXAA26.0003	
EXAX01.0003	EXAA26.0003	Germany
EXAX01.0003	EXAA26.0003	Germany
EXAX01.0003	EXAA26.0003	Denmark
EXAX01.0003	EXAA26.0003	France
EXAX01.0003	EXAA26.0003	
EXAX01.0003	EXAA26.0004	
EXAX01.0018	EXAA26.0019	
EXAX01.0002	EXAA26.0002	Georgia
EXAX01.0009	EXAA26.0002	
EXAX01.0001	EXAA26.0001	Georgia
EXAX01.0019	EXAA26.0018	
EXAX01.0010	EXAA26.0010	
EXAX01.0020	EXAA26.0020	
EXAX01.0014	EXAA26.0011	
EXAX01.0015	EXAA26.0015	
EXAX01.0016	EXAA26.0016	
EXAX01.0017	EXAA26.0017	
EXAX01.0006	EXAA26.0009	
EXAX01.0007	EXAA26.0008	
EXAX01.0008	EXAA26.0007	
EXAX01.0005	EXAA26.0006	

If You Could Do It All This In One Shot?

By whole genome sequencing (WGS)

TAT: 3-4 days

GENUS/SPECIES: *Escherichia coli*

SEROTYPE: O104:H4*

PATHOTYPE: Shiga toxin producing and Enteroaggregative *E. coli* (STEC & EAEC)

VIRULENCE PROFILE: *stx2a, aggR, aggA, sigA, sepA, pic, aatA, aaiC, aap*

SEQUENCE TYPE: ST678

ANTIMICROBIAL RESISTANCE GENES: bla_{TEM-1} , $bla_{CTX-M-15,}$ *strAB*, *sul2*, *tet(A)*A, *dfrA7*

All characteristics have been determined by whole genome sequencing (WGS)

Comments:

***Disclaimer** - This test has not been cleared or approved by the FDA. The performance characteristics have not been fully established. The results of this test should NOT be used for the diagnosis, treatment, or assessment of patient health or management.

Explanantion of Virulence Markers

The strain contains Shiga toxin subtype 2a typically associated with virulent STEC
It does <u>not</u> contain adherence and virulence factors (*eae, ehxA*) typically associated with virulent STEC
It contains adherence and virulence factors typically associated with virulent EAEC (*aggR, aggA, sigA, sepA, pic, aatA, aaiC, aap*)
It is likely resistant to penicillins, cephalosporins, streptomycin, sulphonamides, trimethoprim and tetracyclines
This genotype is associated with extremely high (>10%) rates of hemolytic uremic syndrome (HUS)

Approved by: Nancy Strockbine, Ph.D.
Ph: 404-639-4186
Fax: 404-639-3333
E-mail: nas6@cdc.gov

Species

Serotype

All known virulence genes

Sequence type

Resistance genes

Including Phylogenetically Relevant High-Precision, High-Discriminatory Subtyping?

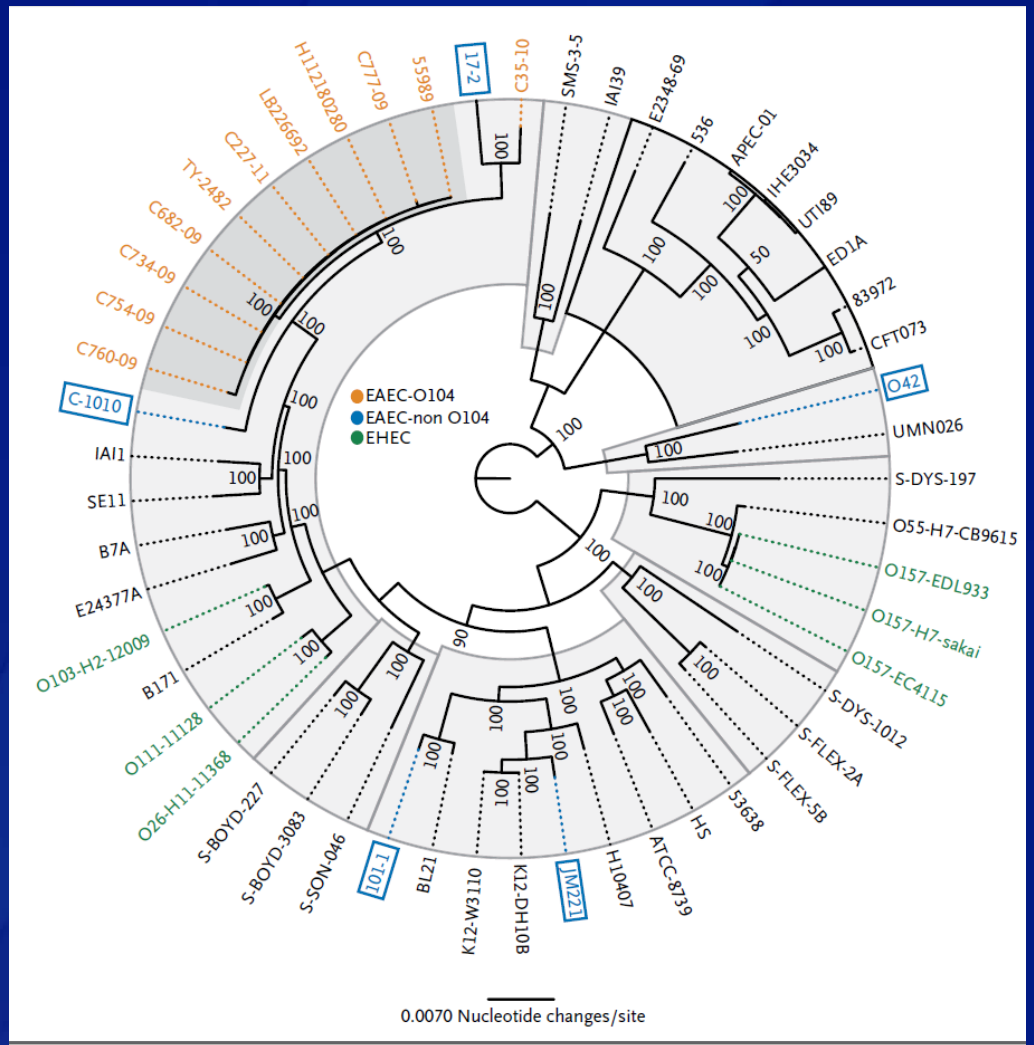


Figure 2. Phylogenetic Comparisons of 53 *Escherichia coli* and Shigella Isolates.

At the Same Or Lower Cost

From: Rasko DA et al. N Engl J Med. 2011 Aug 25;365(8):709-17.

"Transforming Public Health Microbiology – PulseNet and Beyond"

- **Replacing traditional microbiology with WGS**
  - For Enteric Reference Labs:
    - Consolidation of multiple workflows: Identification – serotyping – virulence profiling – antimicrobial resistance characterization – subtyping
    - Fast: Decrease TAT to (2)-3-4 days
    - Cheaper: Supplies for traditional STEC and *Campylobacter* reference testing cost ~ 2 x the cost of sequencing supplies

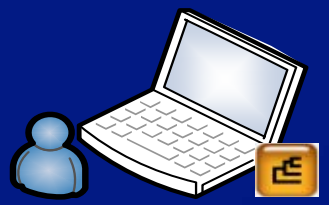
CDC's AMD Initiative

- **5 year budget initiative that started in fiscal year 2014 - initial investment of $30 million; level funding requested for each of the remaining years**

Role of State and Local Public Health Labs in the Future WGS-based Surveillance

Existing personnel will:

Isolate, sequence and perform routine analysis of WGS data from their own jurisdiction for the use in local and national laboratory surveillance (& diagnostics)

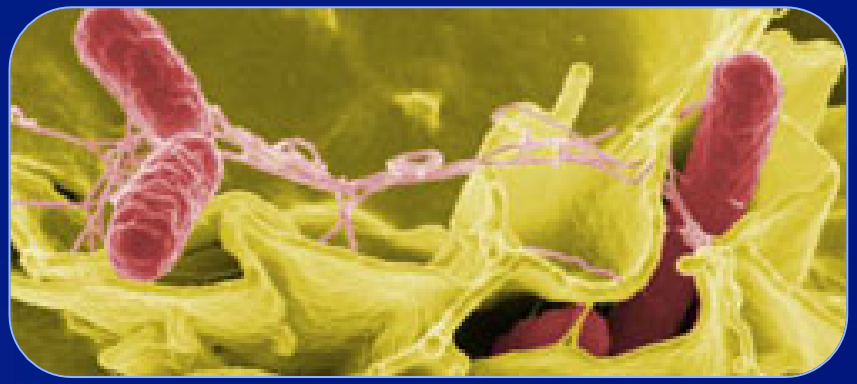


Path to Implementation of WGS in Public Health

- **Analysis on the BioNumerics platform**
  - Already in all public health labs
  - WGS scripts will be made available to partners in the states once developed
- **Training in sequencing and analysis provided for state partners**
  - Three training session completed

➡ **Less isolates to be sent to CDC for characterization**

- Faster and more efficient surveillance



Role of CDC Laboratories In The World of WGS

- **Data management & data analysis**
- **Surge capacity for WGS**
- **WGS Training & Troubleshooting**
- **National organism specific subject matter expertise**
- **‘Center for Classical Microbiology’**
  - When WGS fails or new strains emerge
  - Sentinel surveillance using classical methods
- **More integration of laboratory and epidemiology**
  - Laboratory expertise is needed to use and interpret the data in epidemiological contexts

The WGS- Based National Laboratory Surveillance System

- Consolidated subtyping and reference laboratory workflows
  - Backwards compatible (?)
  - Stable nomenclature
  - High epidemiological concordance
- Fast & economical
- Compatible with epidemiology and regulatory tracking systems
- Harmonized comparison and communication of results locally, nationally, globally
- Local control
  - By non-bioinformaticians using standard Windows software on standard PCs

WGS-Based Surveillance Of Foodborne Infections

- A click button analysis tools that pull out data for reference and surveillance needs in a single workflow

- Currently in all PulseNet laboratories
- May be implemented with minimal training of existing personnel

Gene-Gene Approach
Multi Locus Sequence Typing (MLST)

- Multi-locus sequence typing (MLST) and/or identification of individual genes for reference characterization
- Assess variations (‘alleles’) within each gene:
  - SNP(s), indels, rearrangements

‘Locus’ (gene)	Strain 1	Strain 2	Strain3	Strain 4
A	ACTAGAGGGAA **allele 1**	ACTAGAGG**C**AA **allele 2**	ACT**-**GAGGG**T**AA **allele 3**	AC**GGGAGAT**AA **allele 4**
B	TAGCCAGGGTC **allele 1**	TAGC**A**AGGGTC **allele 2**	TAGC**---**GGTC **allele 3**	TAGGCAGGGTC **allele 1**
C, D, E, etc.…	**alleles 5,2,8…**	**alleles 1,4,7…**	**alleles 1,3,9…**	**alleles 6,2,9…**

- The gene- gene approach may provide you all the information you need in a reference laboratory

Gene – Gene Approach
Multi Locus Sequence Typing (MLST)

- **Tiered approach to strain characterization**

Genus/Species
Serotype
AR, virulotype
7- gene MLST

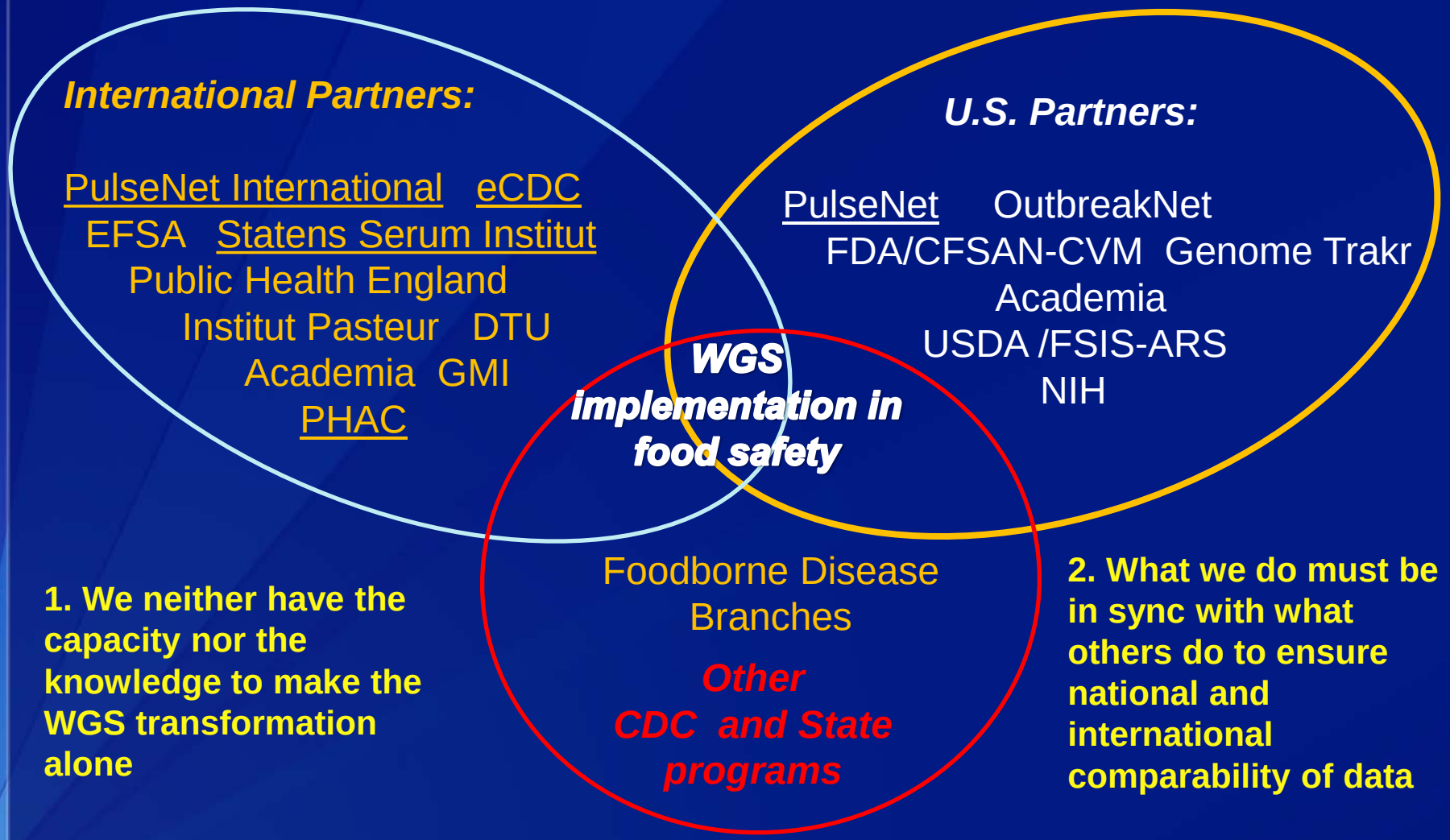
- **Definitive subtyping**
  - Leads to nomenclature
  - *Requires curation*
- **Hardly any isolates will be identical by wgMLST**
  - **…. but surely at lower tier levels**

MLST vs SNP

	SNP	MLST
Epidemiological concordance	High	High
Stable nomenclature	No	**Yes**
Reference characterization: identification, serotyping, virulence & resistance markers	No	**Yes**
Speed	Slow SNP calling, slow analysis	Slow allele calling, fast analysis
Local computing requirements	Medium-High	**Low**
Local bioinformatics expertise	Yes	**No**
Reference used to perform analysis	Sequence of closely related annotated strain	Allele database
Requires curation	No	Yes

MLST is the primary approach for public health surveillance; SNP is used if more detail is needed or MLST fails

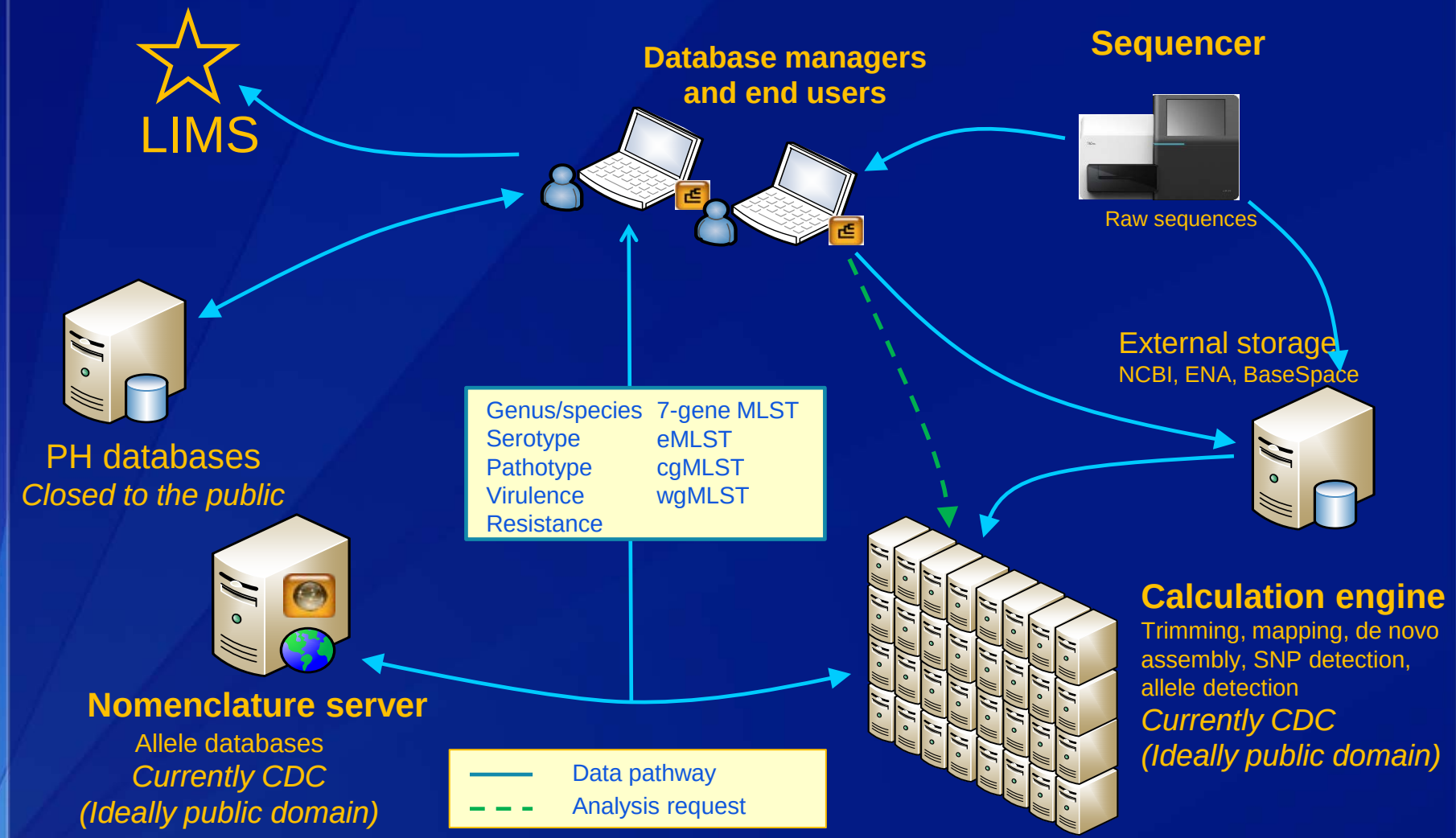
Partners In System Development



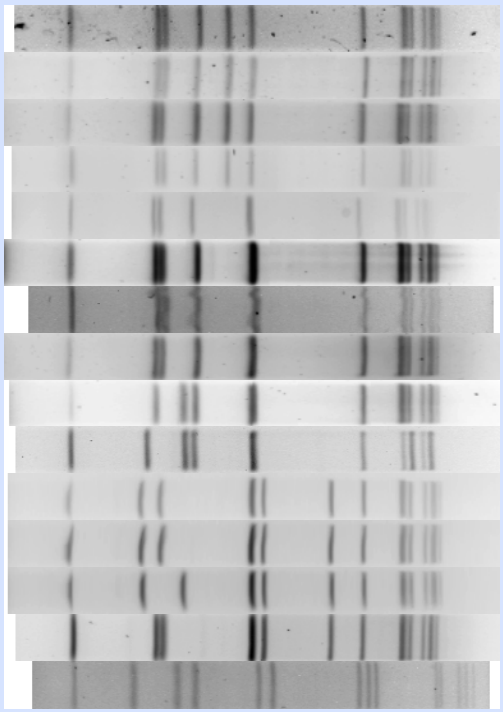
1. We neither have the capacity nor the knowledge to make the WGS transformation alone

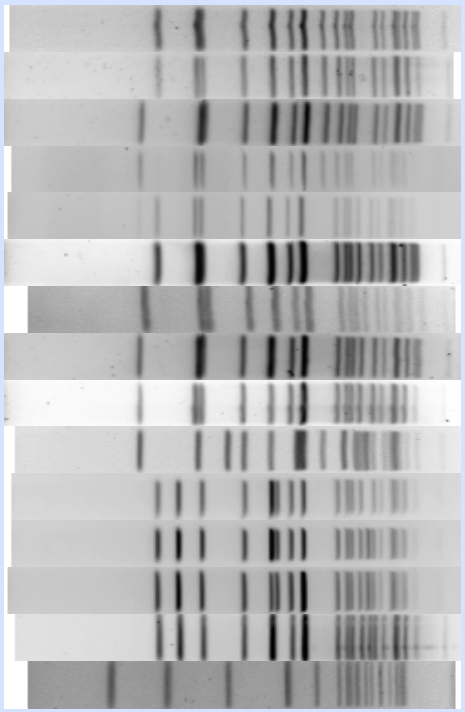
2. What we do must be in sync with what others do to ensure national and international comparability of data

Public Health WGS Workflow



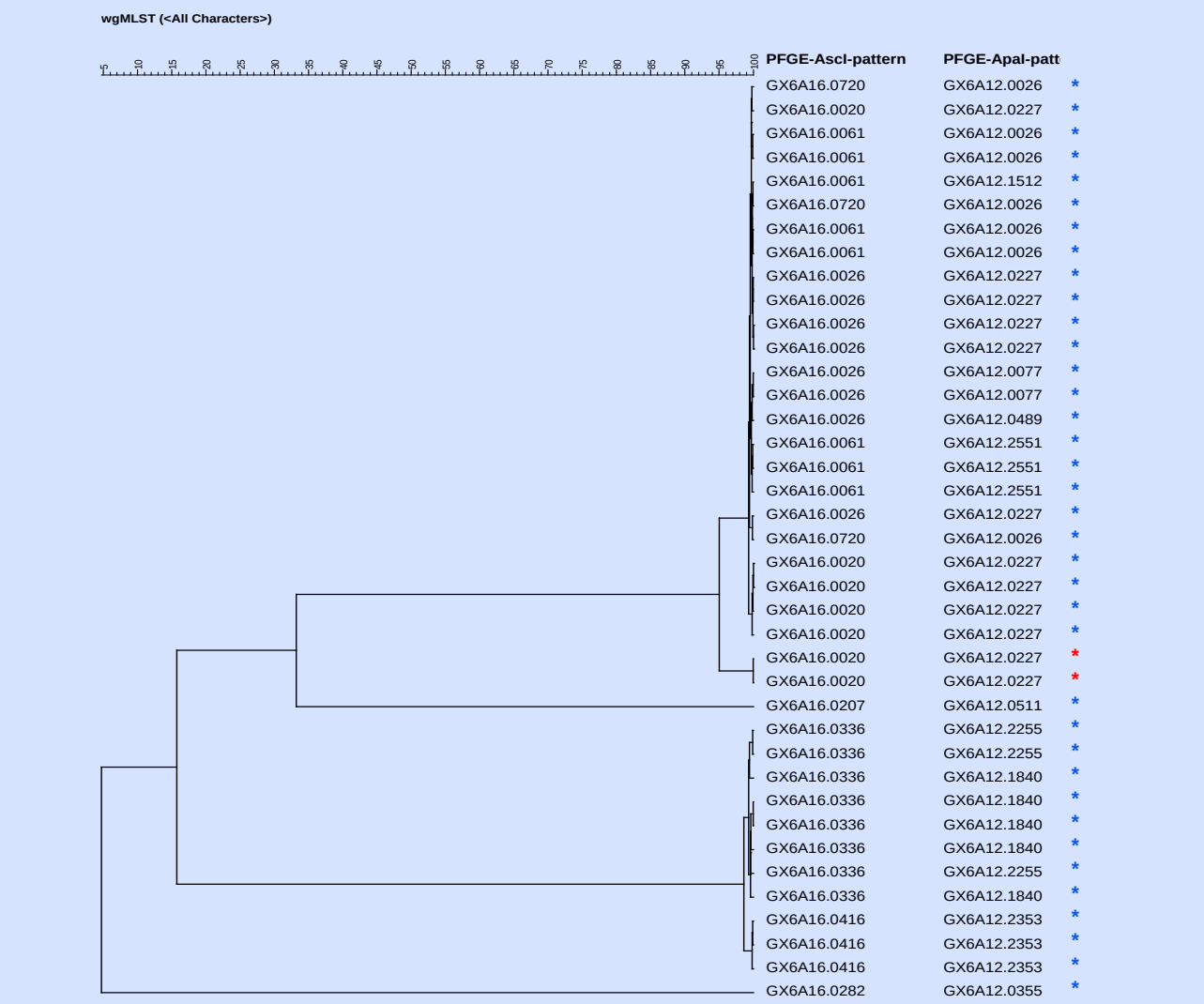
PFGE Patterns Seen In The Blue Bell (*) and Jeni's (*) Ice Cream *Listeria* Recalls





N		Serotype
1	*	1/2b
1	*	1/2b
3	*	1/2b
28	*	1/2b
1	*	1/2b
3	*	1/2b
15	*	1/2b
6	*	1/2b
13	**	3b
1	*	Not typed
18	*	3b
36	*	3b
2	*	3b
4	*	Not typed
1	*	1/2a

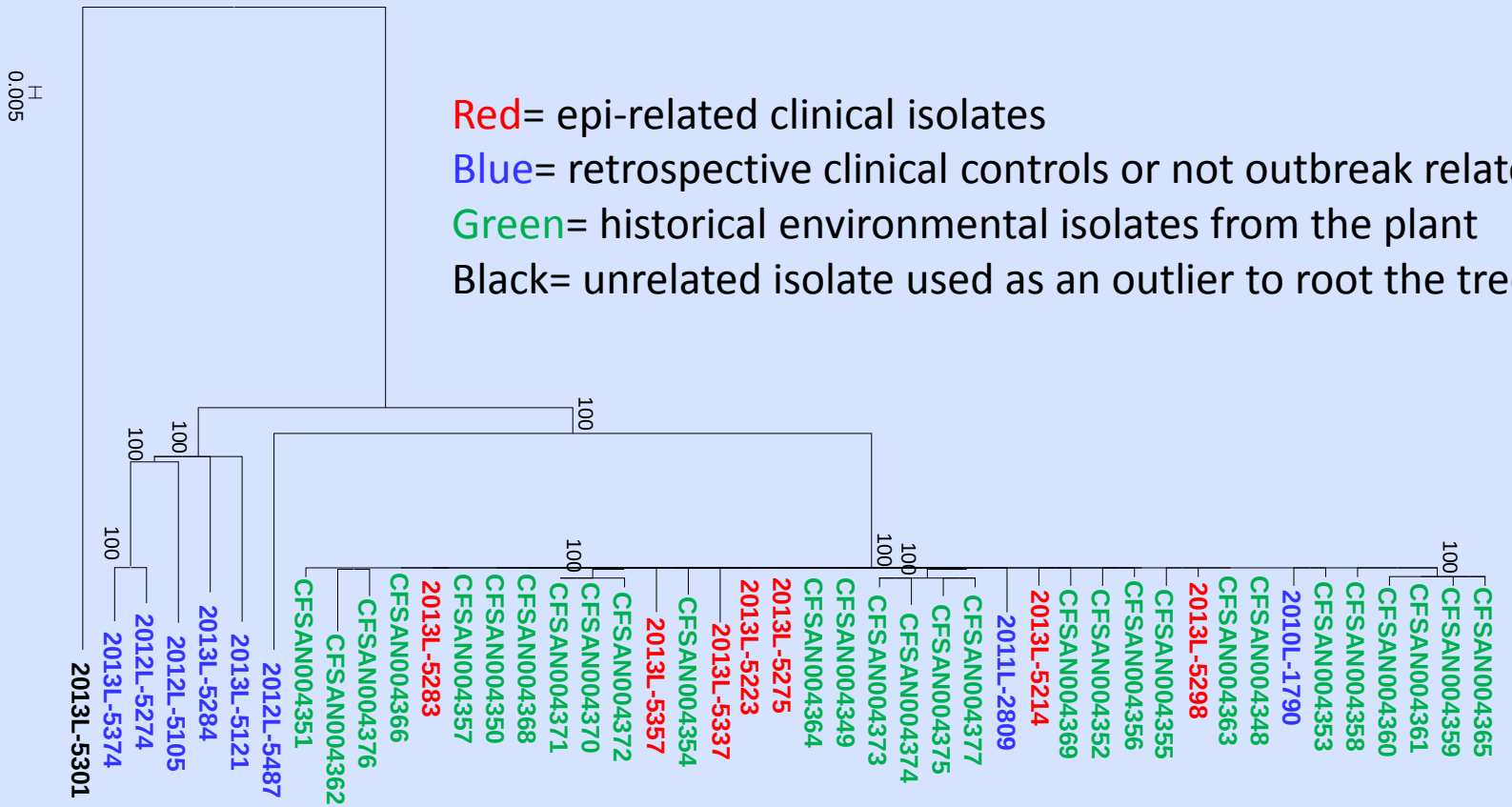
wgMLST Tree Of Isolates From The Blue Bell (*) and Jeni's (*) Ice Cream *Listeria* Recalls



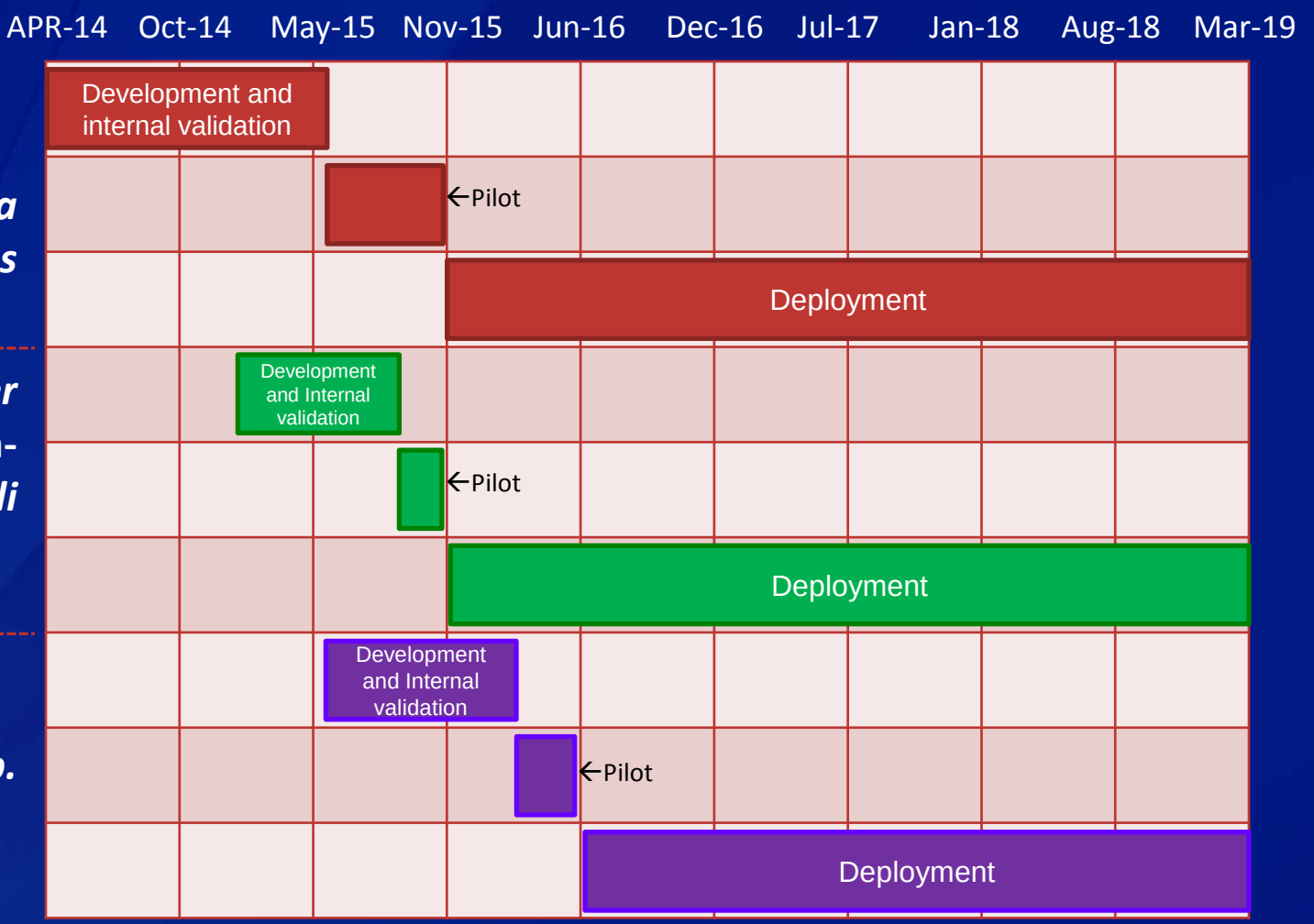
WGS Of *Listeria monocytogenes* From The Crave Brothers Cheese outbreak (2013)

hqSNP

Historical isolates from the plant environment added to the comparison (courtesy FDA/CFSAN)



Projected wgMLST Database Implementation Timeline



Acknowledgements


Public Health Agency of Canada


Colleagues in EDLB
University of Oxford: Martin Maiden

Disclaimers:

"The findings and conclusions in this presentation are those of the author and do not necessarily represent the official position of the Centers for Disease Control and Prevention"

"Use of trade names is for identification only and does not imply endorsement by the Centers for Disease Control and Prevention or by the U.S. Department of Health and Human Services."