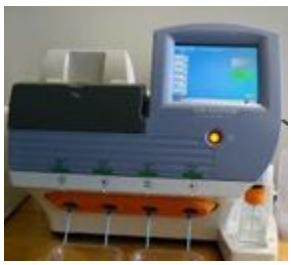
NEW YORK STATE OF OPPORTUNITY. Department of Health | Wadsworth Center

Whole Genome Sequencing - How It's Changing Clinical Practice

Division of Disease Control
Annual Infectious Disease Conference
November 15, 2017

William Wolfgang, PhD
Wadsworth Center, NYS DOH
william.wolfgang@health.ny.gov



ION Torrent

Illumina MiSeq
Next Generation
DNA sequencer

NEW YORK STATE OF OPPORTUNITY. Department of Health

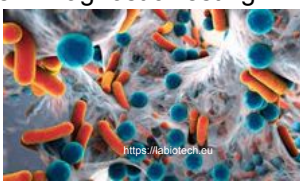
A lot is expected from whole genome sequencing

- Improve cluster detection to aid in source identification.
- Provide comprehensive & rapid detection of drug resistance and virulence factors.
- Rapid pathogen identification.
  - Serotype and genotype identification.
- Characterization of complex microbial communities.

- Improve turn around times.
- Improve benefit to cost ratio.

NEW YORK STATE OF OPPORTUNITY. Department of Health

Three stories

Surveillance	Virulence & Drug resistance	Culture Independent Diagnostic Testing
		
• *Salmonella* Enteritidis story	• Tb drug resistance story	• *E. coli* surveillance from primary stool samples
• Tools are in place – State PHL, CDC and FDA	• Tools are being developed – Wadsworth	• Tools are being developed
• Close to full national implementation in PulseNet	• National standards in development	• National standards?

NEW YORK STATE OF OPPORTUNITY. Department of Health

Food poisoning in America is very **very** common

Each year

- 1 in 6 (48 million) get sick.
- 128,000 are hospitalized
- 3,000 die
- For about 60% the cause is unknown.

CDC statistics - 2011.

- ***Listeria*** 1,500 cases
  - 1,455 hospitalizations and 255 deaths
- ***Salmonella*** 1 million cases.
  - 19,000 hospitalizations and 378 deaths.

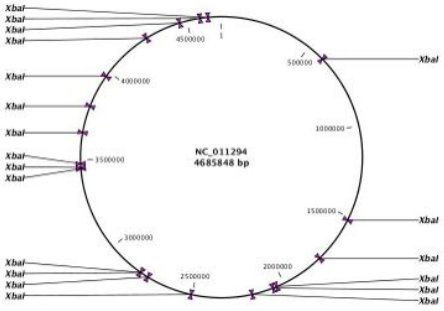
NEW YORK STATE OF OPPORTUNITY. Department of Health

Salmonella is everywhere

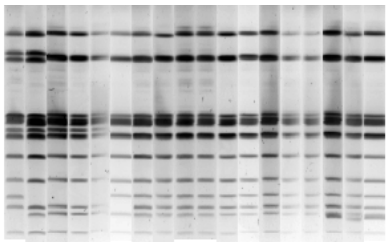
- Peanut Butter – *Salmonella* Bredeney
- Hedgehogs – *Salmonella* Typhimurium
- Mangoes – *Salmonella* Braenderup
- Cantaloupe – *Salmonella* Typhimurium and *Salmonella* Newport
- Ground Beef - *Salmonella* Enteritidis
- Live Poultry – *Salmonella* Hadar
- Dry Dog Food - *Salmonella* Infantis
- Raw Scraped Ground Tuna Product - *Salmonella* Bareilly and *Salmonella* Nchanga
- Small Turtles - *Salmonella* Sandiego, *Salmonella* Pomona, and *Salmonella* Poona
- Restaurant Chain A - *Salmonella* Enteritidis

NEW YORK STATE OF OPPORTUNITY. Department of Health

To aid epidemiologists a DNA fingerprint is created by Pulsed Field Gel Electrophoresis (PFGE)



DNA fingerprints from *Salmonella*



-1Mbp

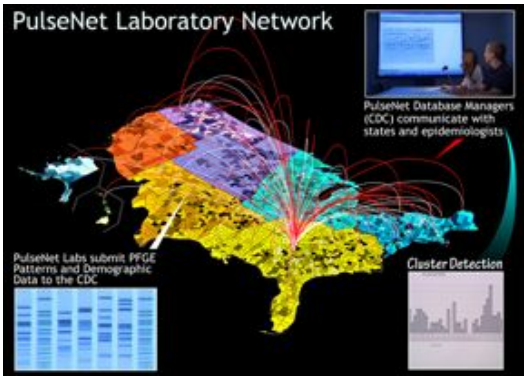
5'…TCTAGA…3'
3'…AGATCT…5'

- Chromosome is cut at specific sites by a restriction enzyme.
- The fragments are separated by size on a gel.
- If **patterns match** then the bugs **may be** closely related.

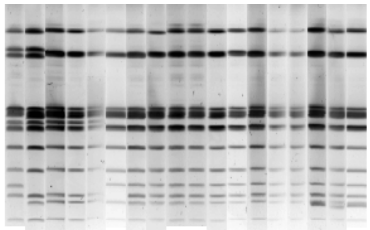
NEW YORK STATE OF OPPORTUNITY. Department of Health

Each year the Wadsworth Center performs PFGE on about 2,000 bacterial

- These PFGE patterns are uploaded to a CDC PulseNet database
- To detect outbreaks in the patient population.
- To find source of the outbreak.



PulseNet USA
The National Molecular Subtyping Network for Foodborne Disease Surveillance
CDC



NEW YORK STATE OF OPPORTUNITY. Department of Health

Societal and economic benefits of PulseNet are great

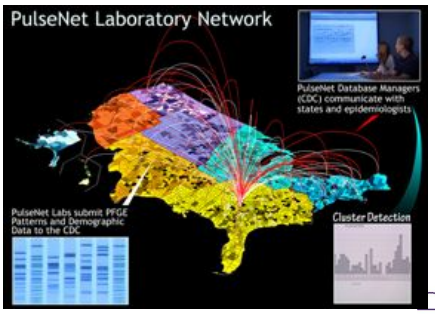
Illnesses averted annually

- 266,522 *Salmonella*,
- 9,489 *Escherichia coli*
- 56 *Listeria monocytogenes*

$507 million are saved annually in medical and productivity costs.

- returns **$70** for every dollar spent.

Scharff et al. 2016 Am. J. of Preventive Medicine
http://dx.doi.org/10.1016/j.amepre.2015.09.018



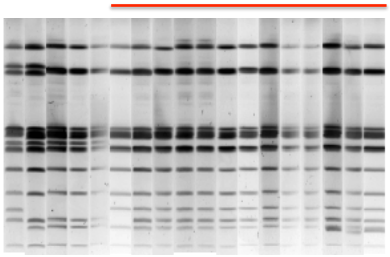
But the PFGE fingerprints can be misleading

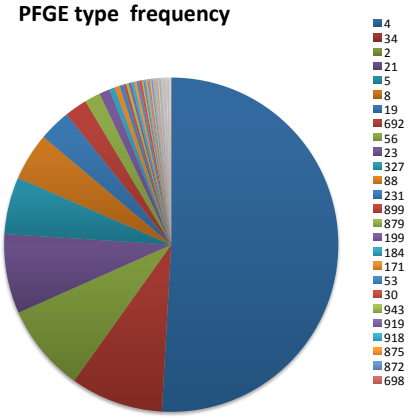
i. Low resolution
ii. Identify false clusters
iii. Instability

NEW YORK STATE OF OPPORTUNITY. Department of Health

For *Salmonella* Enteritidis PFGE has low resolution

- 1/2 have the same PFGE DNA fingerprint.
- And 2/3 have a very common PFGE DNA fingerprint.
- These common patterns are of limited use to our epidemiologists.





52 PFGE types

NEW YORK STATE OF OPPORTUNITY. Department of Health

Greater discrimination can be achieved by Whole Genome Sequencing

Identification of a Salmonellosis Outbreak by Means of Molecular Sequencing

E. Kurt Lienau, Ph.D.
Errol Strain, Ph.D.
Charles Wang, B.S.
Jie Zheng, D.V.M., Ph.D.
Andrea R. Ottesen, Ph.D.
Christine E. Keys, M.S.
Thomas S. Hammack, M.S.
Steven M. Musser, Ph.D.
Eric W. Brown, Ph.D.
Marc W. Allard, Ph.D.
Food and Drug Administration
College Park, MD
marc.allard@fda.hhs.gov
Guojie Cao, M.S.
Jianghong Meng, D.V.M., Ph.D.
University of Maryland
College Park, MD
Robert Stones, M.S.
Food and Environment Research Agency
York, United Kingdom

n engl j med 364;10

Applied and Environmental Microbiology, Dec. 2011, p. 8648–8655 Vol. 77, No. 24
0099-2240/11/$12.00 doi:10.1128/AEM.06538-11
Copyright © 2011, American Society for Microbiology. All Rights Reserved.

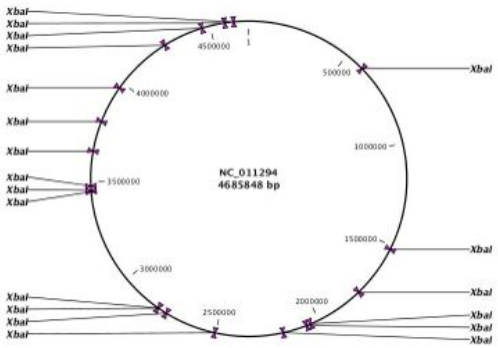
A Whole-Genome Single Nucleotide Polymorphism-Based Approach To Trace and Identify Outbreaks Linked to a Common *Salmonella enterica* subsp. *enterica* Serovar Montevideo Pulsed-Field Gel Electrophoresis Type †

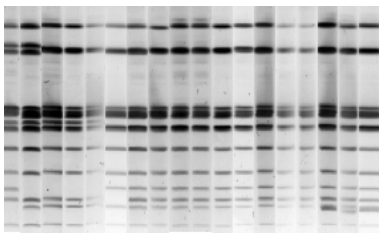
Henk C. den Bakker,[1]* Andrea I. Moreno Switt,[1] Craig A. Cummings,[2] Karin Hoelzer,[1] Lovorka Degoricija,[2] Lorraine D. Rodriguez-Rivera,[1] Emily M. Wright,[1] Rixun Fang,[2] Margaret Davis,[3] Tim Root,[4] Dianna Schoonmaker-Bopp,[4] Kimberlee A. Musser,[4] Elizabeth Villamil,[4] HaeNa Waechter,[5] Laura Kornstein,[5] Manohar R. Furtado,[2] and Martin Wiedmann[1]

Department of Health

WGS provides a better way to fingerprint

- PFGE detects changes at only about 12 sites on the genome.
- Whole Genome Sequencing detects 4.5 million bases in a *Salmonella* genome.

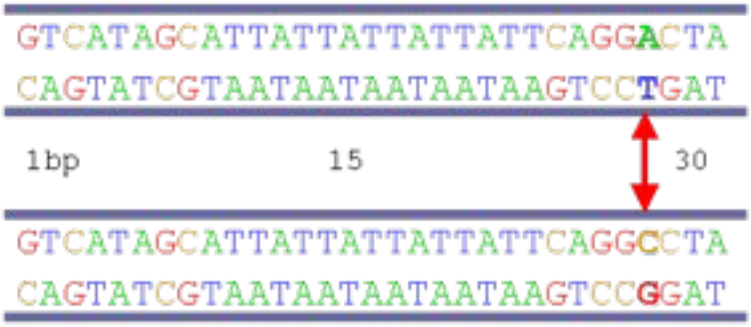




NEW YORK STATE OF OPPORTUNITY. Department of Health

Bioinformatic software that can analyze massive amounts of sequence data

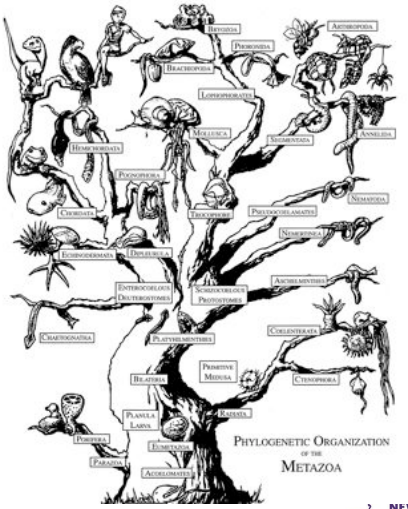
Computer programs can compare these whole genomes to detect Single Nucleotide Polymorphisms (SNP).



NEW YORK STATE OF OPPORTUNITY. Department of Health

Other programs can then infer evolutionary relationships from the pattern of SNPs.

- Critically this reveals lineage as well as closeness.



NEW YORK STATE OF OPPORTUNITY. Department of Health

A 2013 Retrospective study WGS revealed a cluster where PFGE had failed

- Sept. 2010 Connecticut Dept. of Health identifies a *Salmonella* outbreak in a long term care facility (LTCF).
- Outbreak was linked to cannoli from a Westchester bakery.
- Both NY and CT cases consumed cannoli's.

- Isolates had the most common PFGE pattern.



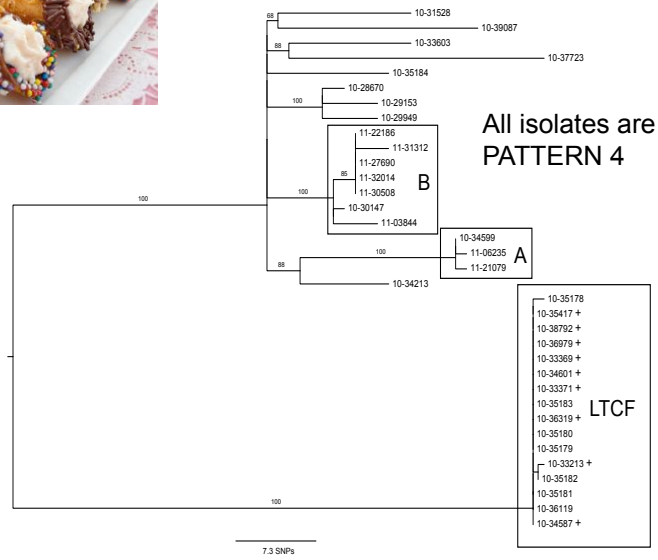

NEW YORK STATE OF OPPORTUNITY. Department of Health

Cannoli cohort

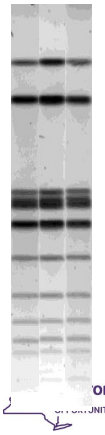
Key	County	Date	PFGE
IDR1000029153	Cattaraugus	8/10/10	JEGX01.0004
IDR1000031528	Rockland	8/26/10	JEGX01.0004
IDR1000033213	Putnam	9/10/10	JEGX01.0004
IDR1000033369	Putnam	9/10/10	JEGX01.0004
IDR1000033371	Putnam	9/11/10	JEGX01.0004
IDR1000034601	Washington	9/13/10	JEGX01.0004
IDR1000034587	Westchester	9/20/10	JEGX01.0004
IDR1000035417	Putnam	9/22/10	JEGX01.0004
IDR1000035178	Westchester	9/13/10	JEGX01.0004
IDR1000035179	Greenwich CT	9/12/10	JEGX01.0004
IDR1000035180	Westchester	9/12/10	JEGX01.0004
IDR1000035181	Westchester	9/13/10	JEGX01.0004
IDR1000035182	Westchester	9/12/10	JEGX01.0004
IDR1000035183	Greenwich CT	9/16/10	JEGX01.0004
IDR1000036119		9/17/10	JEGX01.0004
IDR1100035184	Westchester	9/16/10	JEGX01.0004
IDR1000036319	Putnam	9/28/10	JEGX01.0004
IDR1000036979	Putnam	10/8/10	JEGX01.0004
IDR1000038792	Nassau	10/29/10	JEGX01.0004
IDR1000034599	Orange	9/15/10	JEGX01.0004
IDR1100006235	Westchester	2/21/11	JEGX01.0004
IDR1100021079	Rockland	7/13/11	JEGX01.0004
IDR1000030147	Out-Of-State	8/22/10	JEGX01.0004
IDR1100003844	Onondaga	2/1/11	JEGX01.0004
IDR1100022186	Yates	7/22/11	JEGX01.0004
IDR1100027690	Erie	9/6/11	JEGX01.0004
IDR1100030508	Madison	10/9/11	JEGX01.0004
IDR1100031312	Suffolk	10/5/11	JEGX01.0004
IDR1100032014	Onondaga	10/22/11	JEGX01.0004
IDR1000028670	Nassau	8/8/10	JEGX01.0004
IDR1000029949	Suffolk	8/16/10	JEGX01.0004
IDR1000033603	Erie	9/14/10	JEGX01.0004
IDR1000034213	Erie	9/13/10	JEGX01.0004
IDR1000037723	Westchester	10/4/10	JEGX01.0004
IDR1000039087	Westchester	10/27/10	JEGX01.0004

NEW YORK STATE OF OPPORTUNITY. Department of Health

WGS can identify an outbreak cluster not detected by PFGE



All isolates are PFGE PATTERN 4



NEW YORK STATE OF OPPORTUNITY. Department of Health

Cannoli cohort expanded by WGS

Epi alone

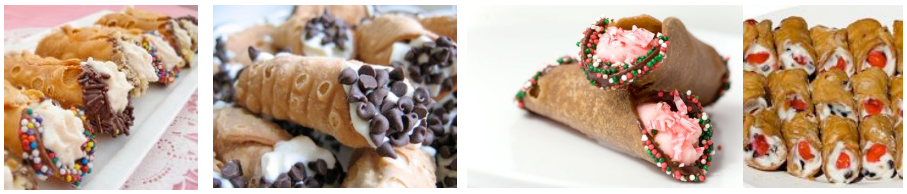
Key	County	Date	PFGE
IDR1000029153	Cattaraugus	8/10/10	JEGX01.0004
IDR1000031528	Rockland	8/26/10	JEGX01.0004
IDR1000033213	Putnam	9/10/10	JEGX01.0004
IDR1000033369	Putnam	9/10/10	JEGX01.0004
IDR1000033371	Putnam	9/11/10	JEGX01.0004
IDR1000034601	Washington	9/13/10	JEGX01.0004
IDR1000034587	Westchester	9/20/10	JEGX01.0004
IDR1000035417	Putnam	9/22/10	JEGX01.0004
IDR1000035178	Westchester	9/13/10	JEGX01.0004
IDR1000035179	Greenwich CT	9/12/10	JEGX01.0004
IDR1000035180	Westchester	9/12/10	JEGX01.0004
IDR1000035181	Westchester	9/13/10	JEGX01.0004
IDR1000035182	Westchester	9/12/10	JEGX01.0004
IDR1000035183	Greenwich CT	9/16/10	JEGX01.0004
IDR1000036119		9/17/10	JEGX01.0004
IDR1000035184	Westchester	9/16/10	JEGX01.0004
IDR1000036319	Putnam	9/28/10	JEGX01.0004
IDR1000036979	Putnam	10/8/10	JEGX01.0004
IDR1000038792	Nassau	10/29/10	JEGX01.0004
IDR1000034599	Orange	9/15/10	JEGX01.0004
IDR1100006235	Westchester	2/21/11	JEGX01.0004
IDR1100021079	Rockland	7/13/11	JEGX01.0004
IDR1000030147	Out-Of-State	8/22/10	JEGX01.0004
IDR1100003844	Onondaga	2/1/11	JEGX01.0004
IDR1100022186	Yates	7/22/11	JEGX01.0004
IDR1100027690	Erie	9/6/11	JEGX01.0004
IDR1100030508	Madison	10/9/11	JEGX01.0004
IDR1100031312	Suffolk	10/5/11	JEGX01.0004
IDR1100032014	Onondaga	10/22/11	JEGX01.0004
IDR1000028670	Nassau	8/8/10	JEGX01.0004
IDR1000029949	Suffolk	8/16/10	JEGX01.0004
IDR1000033603	Erie	9/14/10	JEGX01.0004
IDR1000034213	Erie	9/13/10	JEGX01.0004
IDR1000037723	Westchester	10/4/10	JEGX01.0004
IDR1000039087	Westchester	10/27/10	JEGX01.0004

Epi and WGS

Key	County	Date	PFGE
IDR1000029153	Cattaraugus	8/10/10	JEGX01.0004
IDR1000031528	Rockland	8/26/10	JEGX01.0004
IDR1000033213	Putnam	9/10/10	JEGX01.0004
IDR1000033369	Putnam	9/10/10	JEGX01.0004
IDR1000033371	Putnam	9/11/10	JEGX01.0004
IDR1000034601	Washington	9/13/10	JEGX01.0004
IDR1000034587	Westchester	9/20/10	JEGX01.0004
IDR1000035417	Putnam	9/22/10	JEGX01.0004
IDR1000035178	Westchester	9/13/10	JEGX01.0004
IDR1000035179	Greenwich CT	9/12/10	JEGX01.0004
IDR1000035180	Westchester	9/12/10	JEGX01.0004
IDR1000035181	Westchester	9/13/10	JEGX01.0004
IDR1000035182	Westchester	9/12/10	JEGX01.0004
IDR1000035183	Greenwich CT	9/16/10	JEGX01.0004
IDR1000036119		9/17/10	JEGX01.0004
IDR1000035184	Westchester	9/16/10	JEGX01.0004
IDR1000036319	Putnam	9/28/10	JEGX01.0004
IDR1000036979	Putnam	10/8/10	JEGX01.0004
IDR1000038792	Nassau	10/29/10	JEGX01.0004
IDR1000034599	Orange	9/15/10	JEGX01.0004
IDR1100006235	Westchester	2/21/11	JEGX01.0004
IDR1100021079	Rockland	7/13/11	JEGX01.0004
IDR1000030147	Out-Of-State	8/22/10	JEGX01.0004
IDR1100003844	Onondaga	2/1/11	JEGX01.0004
IDR1100022186	Yates	7/22/11	JEGX01.0004
IDR1100027690	Erie	9/6/11	JEGX01.0004
IDR1100030508	Madison	10/9/11	JEGX01.0004
IDR1100031312	Suffolk	10/5/11	JEGX01.0004
IDR1100032014	Onondaga	10/22/11	JEGX01.0004
IDR1000028670	Nassau	8/8/10	JEGX01.0004
IDR1000029949	Suffolk	8/16/10	JEGX01.0004
IDR1000033603	Erie	9/14/10	JEGX01.0004
IDR1000034213	Erie	9/13/10	JEGX01.0004
IDR1000037723	Westchester	10/4/10	JEGX01.0004
IDR1000039087	Westchester	10/27/10	JEGX01.0004

NEW YORK STATE OF OPPORTUNITY. Department of Health

In this case WGS was better than PFGE



1) Revealed clusters that were not detected by PFGE.
2) Identified isolates in the cluster that were missed by epidemiologists.

Isolates associated with the outbreak were 0 to 4 SNPs apart.

- This range of SNP diversity has been seen in many studies.

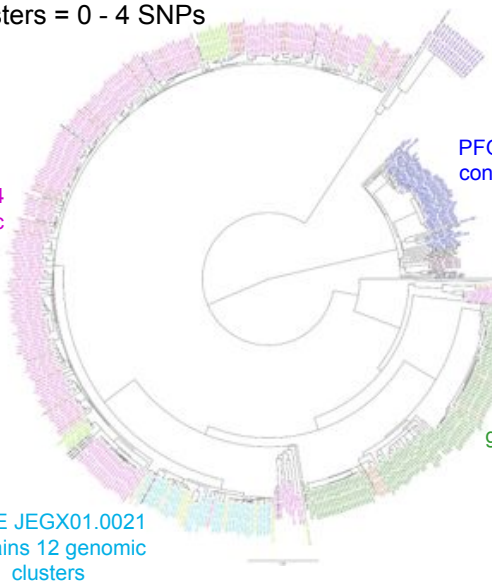
NEW YORK STATE OF OPPORTUNITY. Department of Health

For *Salmonella* Enteritidis how many clusters are we missing using PFGE?

NEW YORK STATE OF OPPORTUNITY. Department of Health

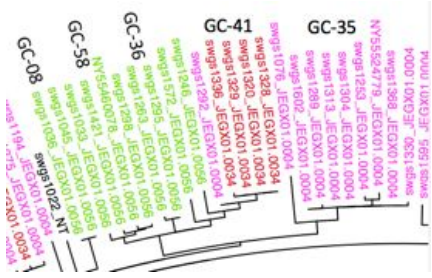
In a 2 year study we sequenced all *S* Enteritidis

- 684 isolates.
- Genomic clusters = 0 - 4 SNPs



NEW YORK STATE OF OPPORTUNITY. Department of Health

2 year study



- **Identified 118 Genomic Clusters (4 SNPs or less)**
  - 42% of isolates reside in clusters
- **Only 7** genomic clusters contained non-endemic PFGE types.
- Most clusters contain 4 or fewer isolates
- Largest cluster contained over 17
- Can persist for almost 2 years
- **31 genomic** (26%) clusters contained multiple PFGE types.

NEW YORK STATE OF OPPORTUNITY. Department of Health

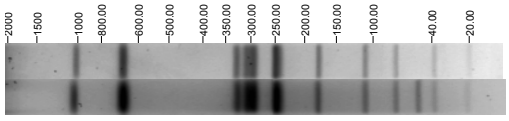
NYC restaurant outbreak

NYCDOHMH requested WGS.

- 6 cases two separate reservations
  PFGE JEGX01.0034
- 1 case dined at around the same time.
  PFGE JEGX01.0004

Q: Was the 7th patient related?

A: Yes; There was 0 SNPs difference.



. Enteritidis JEGX01.0034

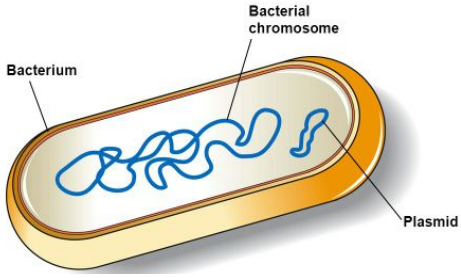
. Enteritidis JEGX01.0004

SLA 5 plasmid

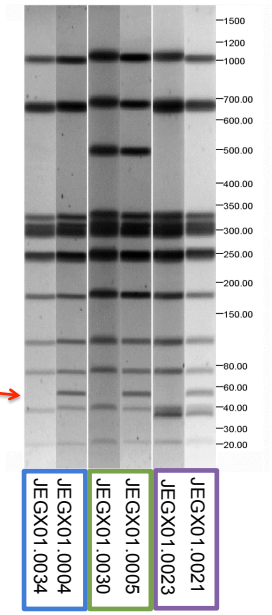
NEW YORK STATE OF OPPORTUNITY. Department of Health

PFGE patterns are unstable because of mobile elements

- Pattern 4 becomes a 34 by loss of a single plasmid (SLA 5).
- Seen in other PFGE types that occur in a single genomic cluster.



SLA5 plasmid



NEW YORK STATE OF OPPORTUNITY. Department of Health

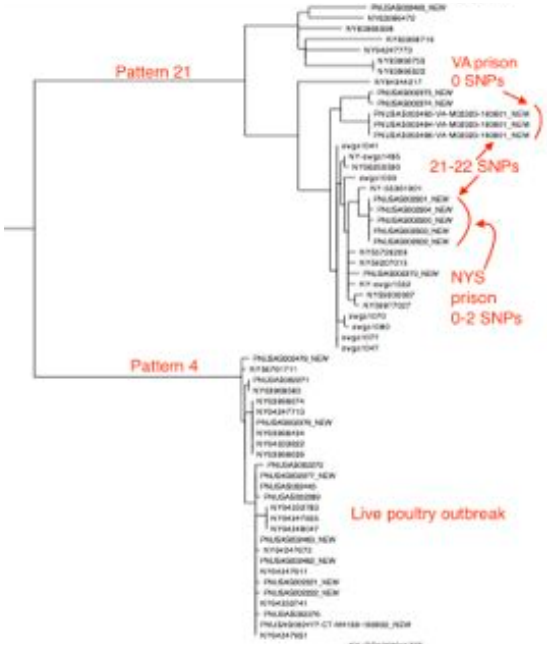
Were two prison outbreaks in NY and VA from a common source?

- **5/16**-**Virginia** reports an SE cluster associated with chicken consumption at a correctional facility.
- At the same time **NYS** is investigating a SE cluster also associated with chicken at a correctional facility.
- Both have the same PFGE pattern: JEGX01.0021

Q. Could they be from a common source?

NEW YORK STATE OF OPPORTUNITY. Department of Health

Prison associated cluster



- Clusters were 21 SNPS apart
- Indicating a different source for each outbreak.

NEW YORK STATE OF OPPORTUNITY. Department of Health

The National Genomic Enteric Surveillance Machine of the future

- CDC PulseNet Network of 39 state labs
  - mostly clinical isolates
- FDA GenomeTrakr Network of 60 federal, state and international labs
  - mostly environmental isolates
- NCBI provides a centralized public repository for all sequences and metadata.
  - Publicly accessible
- CDC and NCBI provide centralized analysis tools to states.
  - CDC uses Bionumerics 7.6 and is private
  - NCBI has developed a Pathogen Browser and is public

PulseNet USA The National Molecular Subtyping Network for Foodborne Disease Surveillance CDC

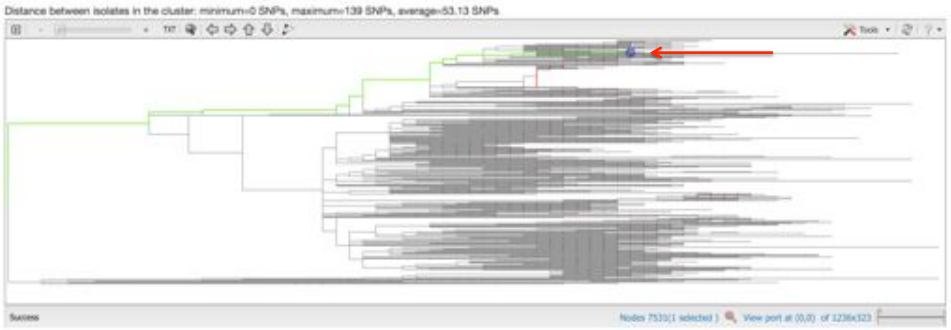
FDA

NCBI

NEW YORK STATE OF OPPORTUNITY. Department of Health

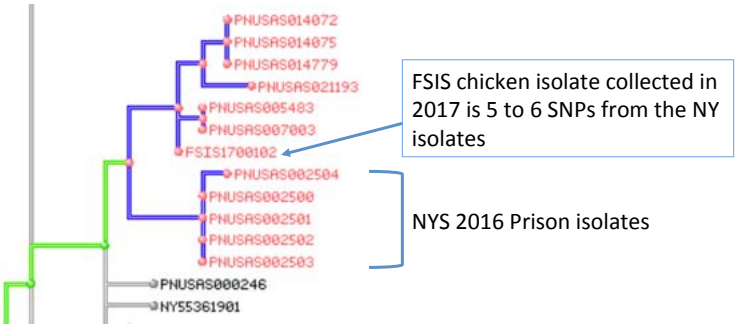
There are 92,839 *Salmonella* genomes at NCBI

- This part of the tree has 5,407 genomes.
- Arrow indicate NYS prison isolate.



NEW YORK STATE OF OPPORTUNITY. Department of Health

Zoom in



- FSIS collected from chicken in 2017 in Virginia.

NEW YORK STATE OF OPPORTUNITY. Department of Health

WGS is great for surveillance!!!

Provides ultimate resolution.

- Subdivide Endemic PFGE types.
- It is more stable.
- Clusters are detected sooner.

It works well to confirm or refute epi. findings.
Not so well to inform investigations up front.

Challenges still exist.

- Standardization of analysis
- Communication with partners
- Increased number of clusters detected:
  - Prioritization
- When do we stop doing PFGE?

NEW YORK STATE OF OPPORTUNITY. Department of Health

Outcomes for WGS surveillance

- Laboratory
  - Clusters will be detected more rapidly and from fewer isolates.
- Epidemiology
  - Allow identification of clusters within **endemic** patterns.
  - Identify more sources of contamination.
- Public Health
  - More efficient identification and removal of pathogen sources.

NEW YORK STATE OF OPPORTUNITY. Department of Health

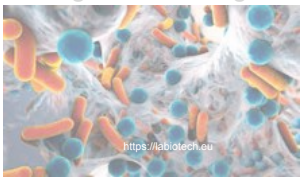
Surveillance

- *Salmonella* Enteritidis story
- Tools are in place
  - CDC and FDA
- Close to full national implementation

Virulence & Drug resistance

- Tb drug resistance story
- Tools are being developed
  - Wadsworth
- National standards in development

Culture Independent Diagnostic Testing

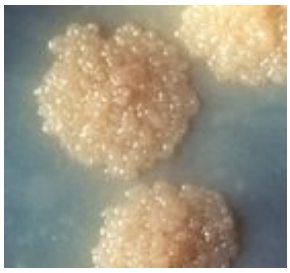


- *E. coli* surveillance from primary stool samples
- Tools are being developed
- National standards?

NEW YORK STATE OF OPPORTUNITY. Department of Health

TB Background

- Caused by *Mycobacterium tuberculosis* and other MTBC species
- Roughly one third of the world's population is infected with TB
- In 2013 there were 9 million new infections, 1.5 million deaths
- Second only to HIV/AIDS as a worldwide killer



NEW YORK STATE OF OPPORTUNITY. Department of Health

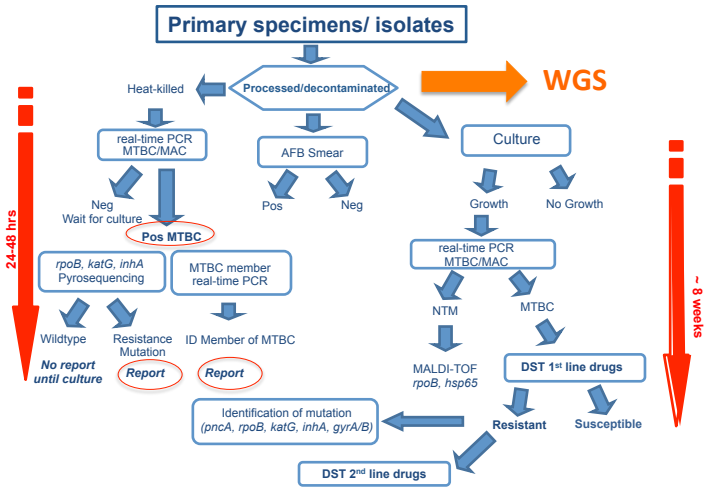
New York State has a high incidence of Tb

	2009	2010	2011	2012	2013	2014	2015	2016
TB Cases*	1007	954	910	864	872	784	766	768
DR-TB	80	63	74	67	54	109	115	
MDR-TB	12	14	20	19	8	12	7	11
XDR-TB	0	0	2	2	0	2	0	0

* New York ranks third or fourth nationally in cases each year

NEW YORK STATE OF OPPORTUNITY. Department of Health

The current work flow to detect drug resistance is slow and complex



NEW YORK STATE OF OPPORTUNITY. Department of Health

Why perform WGS on *Mycobacterium tuberculosis*?

- One stop shop – Simplified workflow
  - Replace 7 tests with 1
    - 1 real-time PCR, 5 pyrosequencing, 1 luminex spoligotyping
- Faster turn-around time
- More comprehensive results
  - More Drug resistance mutations are detected
  - Detect mixed infections and Emerging resistance
- Save money and staff time

NEW YORK STATE OF OPPORTUNITY. Department of Health

A prospective study using WGS was conducted from January 2016 to August 2017

Species identification

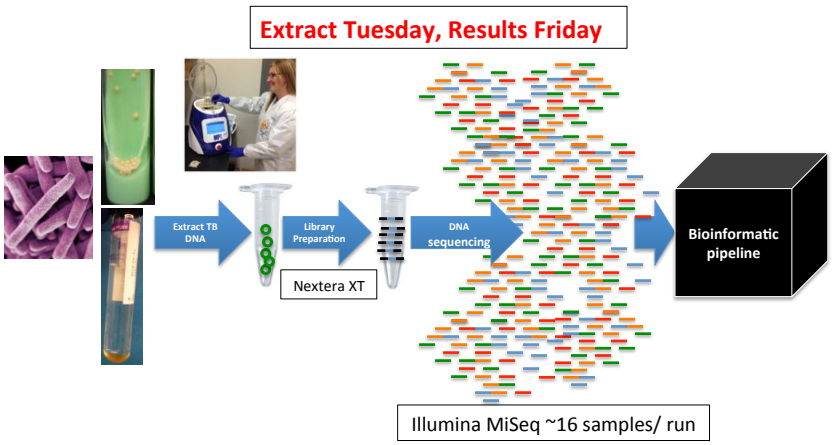
- 1005 *M. tuberculosis*
- 16 *M. bovis*
- 16 *M. bovis*-BCG
- 10 *M. africanum*
- 3 *M. orygis*

Resistance profiling

– 832 pan-susceptible (81%)
- 197 resistant (19%) including:
  – 27 MDR, 1 XDR

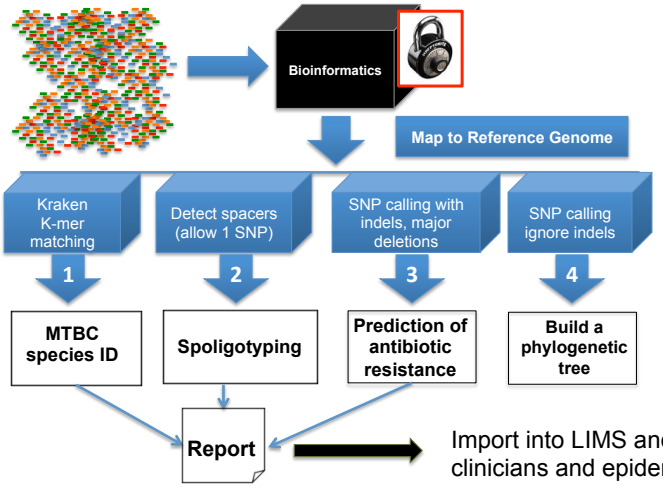
NEW YORK STATE OF OPPORTUNITY. Department of Health

Whole Genome Sequencing Workflow: TAT of less than 1 week



NEW YORK STATE OF OPPORTUNITY. Department of Health

The output from the TB Whole Genome Sequencing Pipeline includes surveillance as well as drug resistance data



Import into LIMS and share with clinicians and epidemiologists

NEW YORK STATE OF OPPORTUNITY. Department of Health

Criteria for assigning high confidence mutations associated with the drug resistance.

- **Supported in literature** by 2 or more independent publications
- **At least one strain received by the lab with the candidate mutation exhibits resistance in culture.** The association must be robust (i.e. no confounding second mutations).
- **A curated database** links drug susceptibility results with a specific mutation at a defined drug concentrations.
- Try to be conservative.

NEW YORK STATE OF OPPORTUNITY. Department of Health

Mutations that predict with High-Confidence Drug Resistance

- Only report **well-characterized** mutations to limit false predictions of resistance
- Mutations of unknown significance **<u>are not</u>** currently reported
- Critically-detect and track all SNPs
- WGS predictions will become more robust

Drug	Locus	Codon/NT position
Rifampin (RIF)	*rpoB*	251, 511, 513, 516, 522, 526, 531, 533, 572
Isoniazid (INH)	*katG* *oxyR-ahpC* promoter region *mabA* promoter region *mabA* *inhA*	279, 315, 394, 525 -81 -17, -15, -8 203 94
Pyrazinamide (PZA)	*pncA/pncA* promoter region	Any nonsynonymous change
Fluroquinolones (FLQ)	*gyrA* *gyrB*	74, 90, 91, 94 510
Ethionamide (ETH)	*mabA* promoter region *mabA* *ethA*	-17, -15, -8 203 Frameshift/STOP

Red = not detected by previous molecular assays

NEW YORK STATE OF OPPORTUNITY. Department of Health

A single comprehensive report is generated by the Bioinformatic pipeline

Gene of Interest	Result
rpoB	Ser531Leu
katG	Ser315Thr
oxyR-ahpC promoter region	No high-confidence mutation
inhA	No high-confidence mutation
mabA-inhA promoter region	No high-confidence mutation
mabA	No high-confidence mutation
pncA	Leu27Pro
pncA promoter region	No high-confidence mutation
embB	Met306Val
rpsL	Lys43Arg
rrs 512, 513, 516, 906	No high-confidence mutation
rrs 1400	No high-confidence mutation
eis promoter region	No high-confidence mutation
gyrA	Asp94Gly
gyrB	No high-confidence mutation
ethA	DELETION

Drug	Result
Rifampin	RESISTANT (predicted)
Isoniazid	RESISTANT (predicted)
Pyrazinamide	RESISTANT (predicted)
Ethambutol	RESISTANT (predicted)
Streptomycin	RESISTANT (predicted)
Kanamycin/Amikacin	Susceptible (suggested)
Kanamycin	Susceptible (suggested)
Fluoroquinolones	RESISTANT (predicted)
Ethionamid	RESISTANT (predicted)

Αλσο σπεχιεσ ιδεντιφιχατιον ανδ σπολιγογτψπε

Species	*Mycobacterium tuberculosis*
WGS_SPOLIGO	000000000003771 (S00034)

NEW YORK STATE OF OPPORTUNITY. Department of Health

Turn around time for reporting of drug susceptibility is shortened by WGS

WGS reported sooner than culture based testing	259	82%
Same Date	30	9.5%
WGS reported later than culture based testing	26	8.2%

- Average TAT is 7 days faster
- Numbers are improving over time
- WGS is reported later for samples for which WGS fails and needs re-culturing

NEW YORK STATE OF OPPORTUNITY. Department of Health

Summary

- WGS improved turn around time compared to conventional drug resistance testing.
- Provides a more comprehensive assessment of drug resistance.
- Will inform the genetic basis for drug resistance going forward.
- Saves time and money.

NEW YORK STATE OF OPPORTUNITY. Department of Health

Future

- Adding additional drug resistance targets
- Determine drug resistance in **primary** samples by an amplicon based approach.
  - Significant decrease in turn around time
  - No new mutations detected
  - No high resolution genotyping

NEW YORK STATE OF OPPORTUNITY. Department of Health

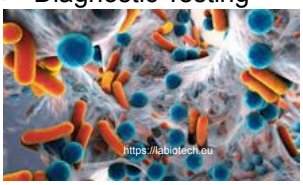
Surveillance

- *Salmonella* Enteritidis story
- Tools are in place
  - CDC and FDA
- Close to full national implementation

Virulence & Drug resistance

- Tb drug resistance story
- Tools are being developed
  - Wadsworth
- National standards in development

Culture Independent Diagnostic Testing

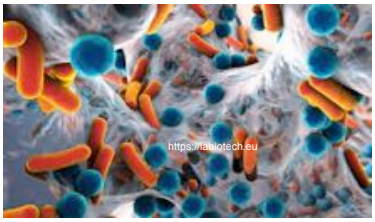


- *E. coli* surveillance from primary stool samples
- Tools are being developed
- National standards?

NEW YORK STATE OF OPPORTUNITY. Department of Health

WGS directly from primary stool samples for STEC

- Currently almost all our STEC come in as stools or broths.
- Sequencing pathogens directly from stool would substantially shorten the time to diagnosis for public health reporting.
- Significant challenges because of the mixture of multiple sequences from different organisms present in the specimen.

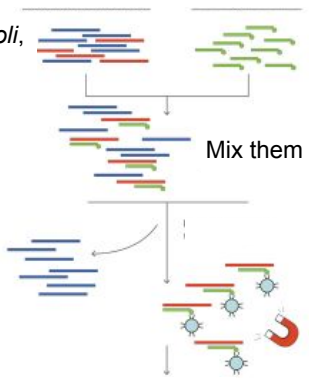


NEW YORK STATE OF OPPORTUNITY. Department of Health

A combined lab / bioinfomatic approach was used to attach this challenge

RNA baiting

DNA in stool; Red = *E.coli*, Blue = species



RNA bait made to *E. coli* (green) is magnetic

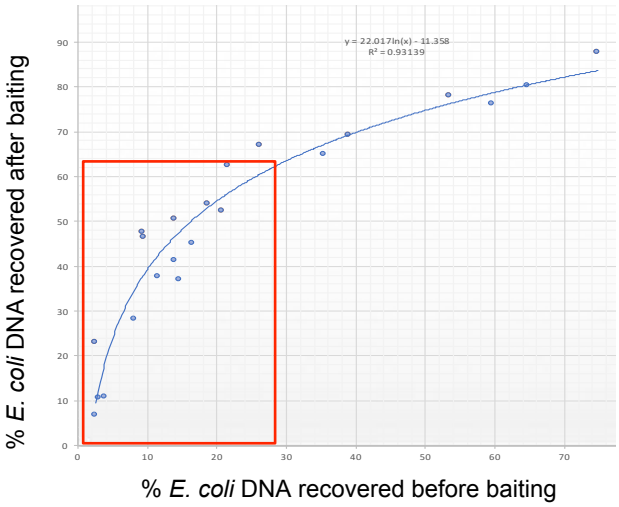
Mix them up

Use a magentic to pull out DNA from *E. coli* stuck to the bait

Sequence target DNA

NEW YORK STATE OF OPPORTUNITY. Department of Health

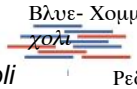
RNA baiting enriches *E. coli* DNA up to 5X



NEW YORK STATE OF OPPORTUNITY. Department of Health

Next Bioinformatic magic happens

- *E. coli* and *shigella* sequences are identified with Kraken.
- All other sequences- human, bacterial, viral and fungal are discarded.
- More magic
  - Align the reads to a reference
  - Determine if there are two populations of *E. coli*
  - Sort the Red reads from the Blue reads.



Βλυε- Χομμενσαλ *E. coli*

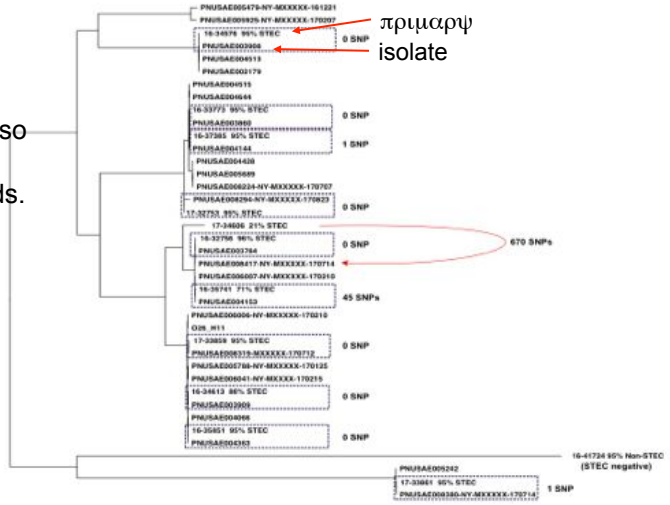
Ρεδ - ΣΤΕΧ

- Assemble red reads and blue reads separately to a reference.
  - Call SNPs

NEW YORK STATE OF OPPORTUNITY. Department of Health

The complete STEC genome is recovered

- Sequenced STEC genomes from 11 stools.
- From same stools also isolated STEC by conventional methods.
- Compared SNP differences in a tree.
- 9 of 11 virtually 0-1 SNP difference.



NEW YORK STATE OF OPPORTUNITY. Department of Health

Summary

- A combination of laboratory techniques and bioinformatic tools can accurately assemble a pure genome from a mixed primary sample.
- The bioinformatics pipeline continues to be optimized to improve the accuracy, particularly when the STEC population is low.
- Eventually will be able to identify virulence factors and serotype.
- Save time and money because there is no need to isolate the pathogen from the stool.
- **Will no longer need isolates.**

NEW YORK STATE OF OPPORTUNITY. Department of Health

Acknowledgments

Wadsworth Bacteriology
- Kim Musser
- Joe Shea
- Tanya Halse
- Lisa Mingle
- Michelle Dickinson
- Samantha Wirth

Wadsworth Center Genomics Core
- Patrick VanRoey
- Matt Shudt
- Zhen Zhang
- Melissa Leisner
- Helen Lee
- Josh Williams
- Nathalie Boucher

Bioinformatics Core
- Mike Palumbo
- Pascal LaPierre
- Erica Lasek-Nesselquist
- Navjot Singh

Wadsworth PulseNet Lab
- Deb Baker
- Lisa Thompson

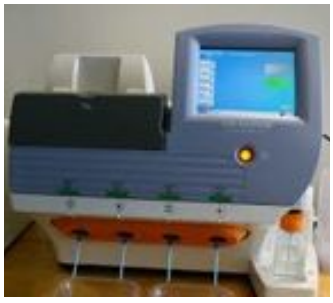
Cornell
- Martin Wiedmann
- Henk den Bakker

CDC
- Eija Trees
- Heather Carlton

FDA
- Eric Brown
- Marc Allard
- Errol Strain
- Ruth Timme

NCBI
- Bill Klimke
- Martin Shumway

NEW YORK STATE OF OPPORTUNITY. Department of Health



NEW YORK STATE OF OPPORTUNITY. Department of Health