

Whole Genome Sequencing for TB Diagnostics

Kimberlee Musser, PhD Chief, Bacterial Diseases Wadsworth Center

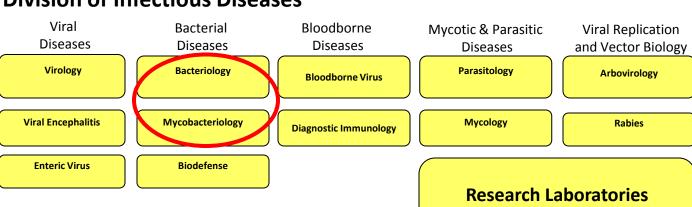


- 900,000 sq. ft. state-of-the-art-facilities- 5 locations
- ~700 staff, >150 doctoral level scientists
- \$25 million in external grant funding
- Laboratories in four scientific divisions:
 - Environmental Health, Infectious Disease, Genetics, Translational Medicine



PI- grant funded programs

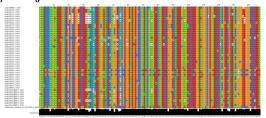
Division of Infectious Diseases



Roles of the Wadsworth Center Bacteriology and Mycobacteriology Laboratories

- Reference services
- Outbreak and hospital investigations
- Specialized testing
- Support of disease surveillance and epidemiology investigations
- Preparedness and response
- Applied research (NIH, CDC, contracts)







TB Background



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

A Typical TB Case Requires:



PLUS

- X-rays
- Lab tests
- Follow-up & testing of contacts

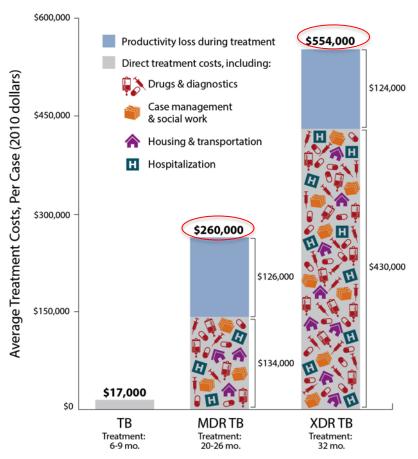


Total cost to U.S for TB cases in 2014.



The Outsized Financial Toll of MDR and XDR TB

Cost increases with greater resistance:



Source: U.S. Centers for Disease Control and Prevention

Preventing and Controlling MDR and XDR TB in the U.S. Requires:

BETTER TREATMENT OPTIONS

RAPID DIAGNOSIS

EXPERT
TREATMENT
OF EVERY
TB CASE

IMPROVING GLOBAL TB DIAGNOSIS AND TREATMENT

http://www.cdc.gov/nchhstp/newsroom/2014/WorldTBDay-graphics.html

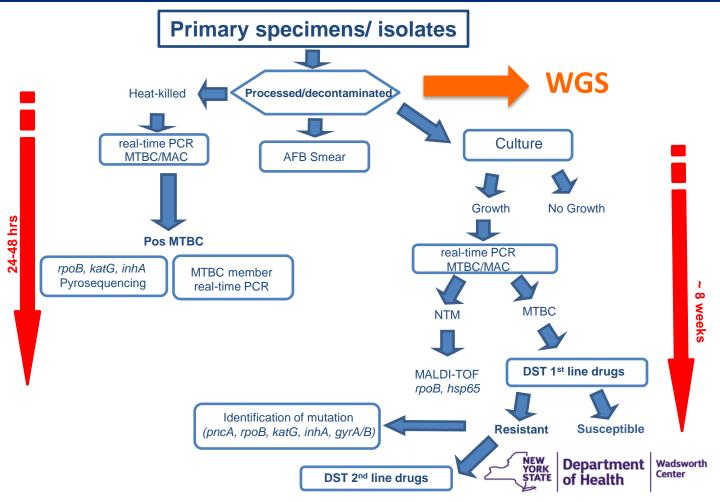


Rapid Diagnosis of Mycobacterium tuberculosis with WGS

- Faster turn-around time
- More comprehensive results
 - Detect mixed infections
 - Many predictors of drug resistance
 - Emerging resistance
- Cost effective
 - Replace existing assays (real-time PCR, pyrosequencing, spoligotyping)
 - Staff time savings

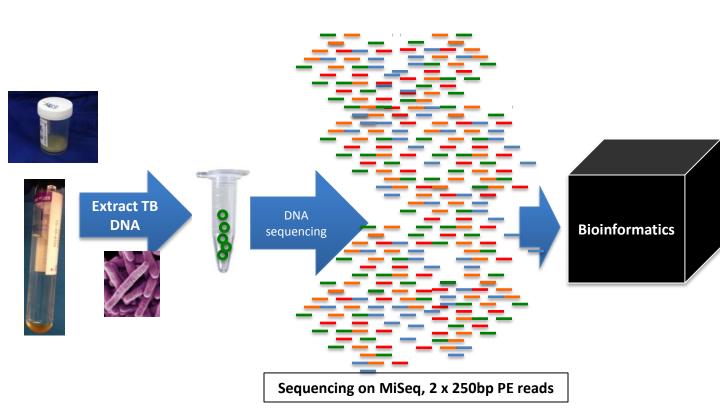


Testing Algorithm



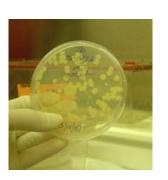
Whole Genome Sequencing

Next Generation Sequencing



Where to start?

- Isolates
 - Solid
 - MGITs
- Primary specimens
 - sputum
 - other







Optimizing TB isolate preparation for WGS

- Assess methods used in lab
- Research TB WGS methods
- Assess worse case scenario
 - 1- 2 ml MGIT
 - early MGIT positive (Day 0-3 flagged positive)
- Ease of use, cost
- DNA concentration
- Ultimately- WGS 40X depth and close to 100% coverage



Breaking TB Open is Critical for DNA Extraction

Important TB Characteristics

- ~24 hour doubling time
- TB clumps together
- Unique cell wall
 - Rich in lipids (>60%)
 - Mycolic acids



Initial Methods Tested

- Typical bacterial extraction
- Zymo Research Kit
 - Meant for tough to lyse fungi/ bacteria
- CTAB method
 - Ideal for plant cell nucleic acid extraction/ MTB



DNA yield too low, labor intensive, WGS variable results



November 15, 2015

InstaGene Matrix and Tissue Homogenizer

InstaGene matrix (Chelex resin)

 The Chelex matrix binds to PCR inhibitors rather than DNA, preventing DNA loss due to irreversible DNA binding.



Fastprep tissue homogenizer

Good enough yield to provide reliable WGS data even with 0 day MGIT

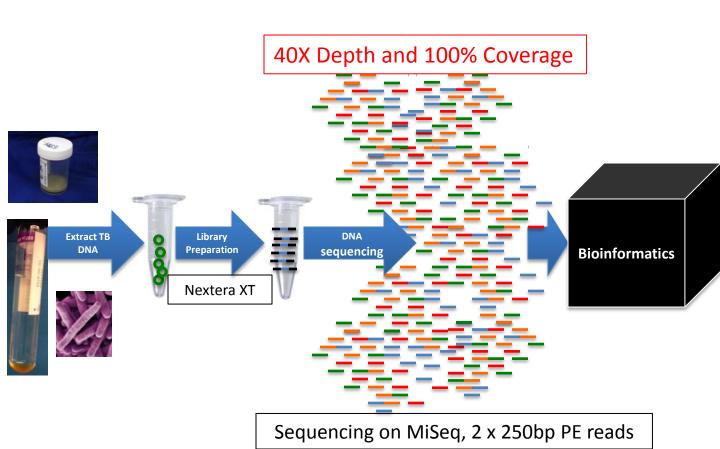






Whole Genome Sequencing

Next Generation Sequencing

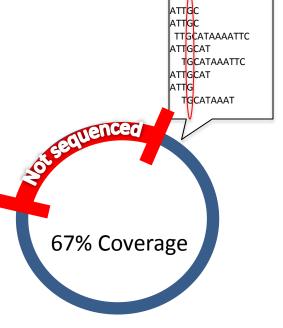


Successful WGS

 <u>Depth</u>: Essentially the number of times the base was read; measure of confidence in correct call

- Can be given as a genome average
- We are aiming for 40X

- <u>Coverage</u>: A percentage that describes how much of the genome was sequenced
 - Best 100%



8X Depth

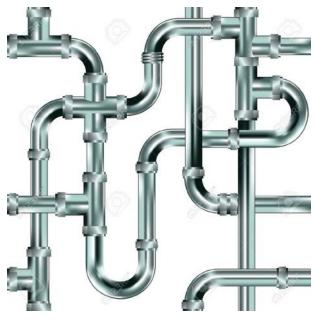


Library Preparation is Another Key Factor

- Votintseva et al. suggested using 15 cycle library preparation
 - 2015 paper about WGS of early positive MGIT

			12 cycle lil	brary prep	1 <u>5 cycle</u>	library prep
Sample	Method	stock ng/ul	Avg depth	coverage %	Avg depth	coverage %
	InstaGene	0.268	FAIL	FAIL	27.66	97.23
M. bovis BCG (Oday)	InstaGene	0.344	FAIL	FAIL	19.4	97.07
(oddy)	InstaGene	0.346	FAIL	FAIL	14.22	96.78

Bioinformatics Pipeline

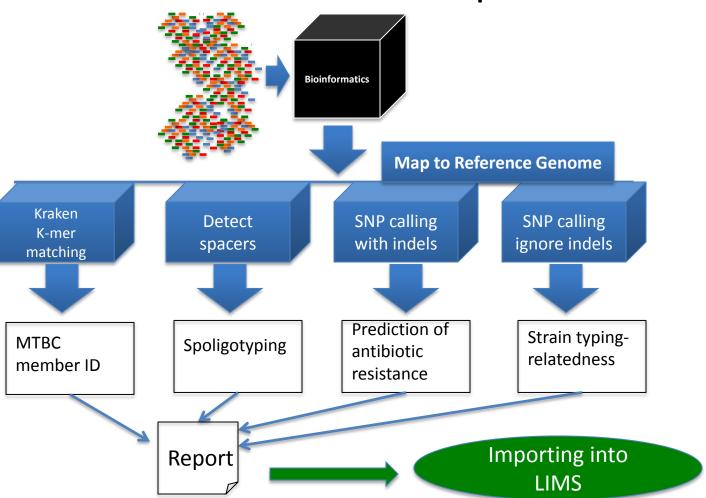


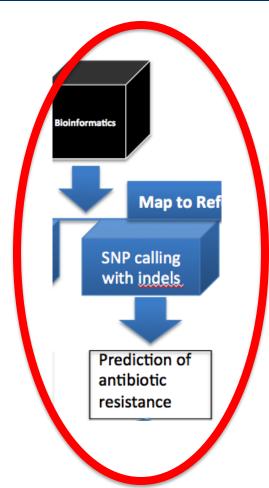
A PRINCIPATION AND AND THE PROPERTY CONTROL OF THE CONTROL OF THE

Pascal Lapierre, PhD Michael Palumbo, PhD



TB Bioinformatics Pipeline





Evaluating 8 Drug Classes

- 12 loci
- Hundreds of potential SNPs
- Additional loci throughout the genome
 - frameshifts
 - insertions
 - deletions

Validation of TB WGS for isolates

- SOP, reports, interpretation, QC, assay controls, metrics
- Specificity, intra-assay and inter-assay reproducibility
- Retrospective testing
- Prospective testing
- Evaluate each drug



Sequence Confirmation- Using another molecular method (used to predict resistance)

Drug	Gene	# Mutations found	# Confirmed	# Not Confirmed
Rifampin	гроВ	46	44	2 ¹
Isoniazid	katG	40	36	4 ¹
Isoniazid &	inhA/	22	17	5 ¹
Ethionamide	mabA			
Fluoroquinolones	gyrA	13	12	1 ²
Streptomycin	rrs	11	11	0
	rpsL	28	26	2 ^{2,3}
Pyrazinamide	pncA	31	19	12 ^{1,2,3}
Ethambutol	embB	28	26	2 ^{1,3}
Kanamycin	rrs	8	8	0
	eis	1	0	1 ¹

¹No assay/ Mutation found outside range of confirmatory assay/ deletion caused assay failure

²pending

³No remaining DNA stock

Fluoroquinolone comparison

		DST Ph	enotype
	Fluoroquinolones	R	S
WGS	R	13	0
Genotype	S	0	61

Resistance Predictive Value= 100% Susceptible Predictive value= 100%

Isoniazid comparison

		DST Ph	enotype
	Isoniazid	R	S
WGS	R	55	11
Genotype	S	62	32

¹This SNP is known to be a good but not perfect predictor of INH resistance (14/15 resistant)

Resistance Predictive Value= 98% Susceptible Predictive value= 84%



² Each of the 6 has a different mutation that could potentially account for the missed resistance

Prospective Testing



First Batch of Prospective specimens

Results as of 8/3/15:

	?	Current@method@results@			WGS@esults 					
	Sample	DST₪	Pyro? results?	spoligo?	Genome2 coverage2	Ave2 Depth2	ID②	High? confidence? mutations? detected?	Frameshift2 and/ordarge2 deletions2 detected2	spoligo®
	IDR15- 510872	invalid;② PZA IS ②	No mutations	ND?	98.362	91.052	Mtb2	None⊡	none⊡	S000342
	IDR15- 520242	pending;2 PZA IS 2	ND®	ND®	98.462	75.327	Mtb₪	gyrA? Ala90Val;? gyrA? Asp94Gly? (FLQ)?	none⊡	Unknown② (new)②
-	IDR15- 52248®	pending®	ND®	ND®	98.821	79.082	Mtb⊡	embB2 Met305Val2 (EMB);2 gyrA2 Ser91Pro2 (FLQ);2 katG2 Ser315Thr2 (INH)2	rpoB@(+TTC)@ in-frame@ insertion@ (RIF)@	S00034®

No Results

WGS complete



First Batch of Prospective specimens

Results as of 8/18/15:

	?	Current@method@esults@			WGS⊞esults⊡					
	Sample⊡	DST⊡	Pyro2 results2	spoligo?	Genome® coverage®		IDī	High2 confidence2 mutations2 detected2	Frameshift2 and/ordarge2 deletions2 detected2	spoligo⊡
	IDR15- 510872	Pan- Susceptible	No? mutations?	ND2	98.362	91.052	Mtb2	None?	none⊡	S000342
	IDR15- 520242	pending;? PZA :5 ?	gyrA:2 Asp94Gly2	pending?	98.462	75.322	Mtb⊡	gyrA? Ala90Val;? gyrA? Asp94Gly? (FLQ)?	none⊡	Unknown2 (new)2
-	IDR15- 52248 [®]	pending?	rpoB:回 insertion回 ofเ密動ases回 at配odon回 514;國atG:回 Ser315Thr ;隱yrA:回 Ser91Pro回	ND®	98.827	79.082	Mtb2	embB2 Met305Val2 (EMB);2 gyrA2 Ser91Pro2 (FLQ);1katG2 Ser315Thr2 (INH)2	rpoBI[+TTC)] in-frame] insertion[] (RIF)]	S00034®

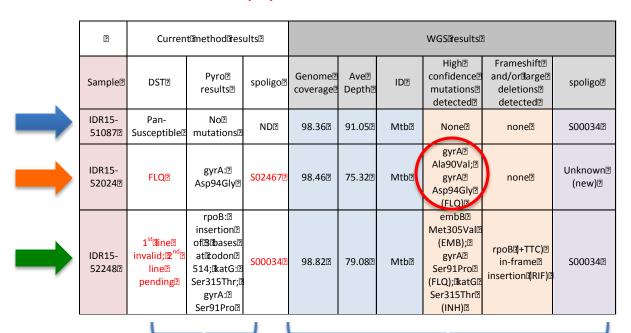
Some Results

WGS complete



First Batch of Prospective specimens

Results as of 10/8/15:



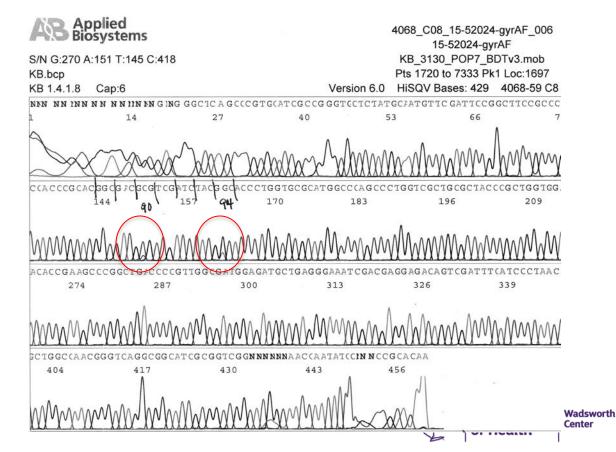
More Results

WGS complete



Wadsworth Center

Heteroresistance



Turn-Around Time

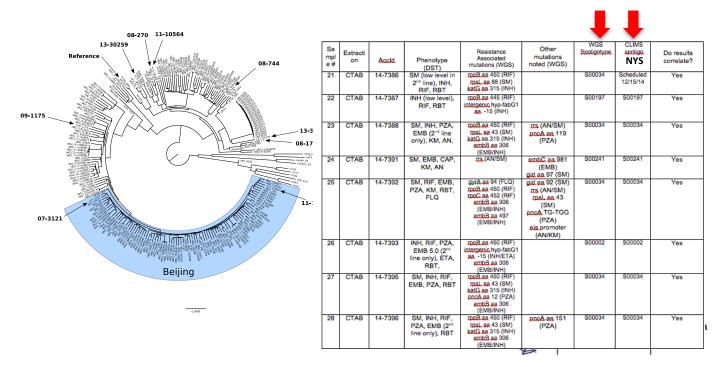


Sample	Date Received	Date processed and sent to core	Date of DST results	Date of WGS report	DST TAT from receipt of specimen to DST results (days)	WGS TAT from receipt of specimen to report (days)
IDR15- 51087	7/3/15	7/17/15	8/13/15	7/30/15	41	27
IDR15- 52024	7/10/15	7/17/15	8/24/15 (1 st line); 9/1/15 (2 nd line)	7/30/15	53	20
IDR15- 52248	7/13/15	7/17/15	1 st line invalid; 2 nd line still pending	7/30/15	>60	17

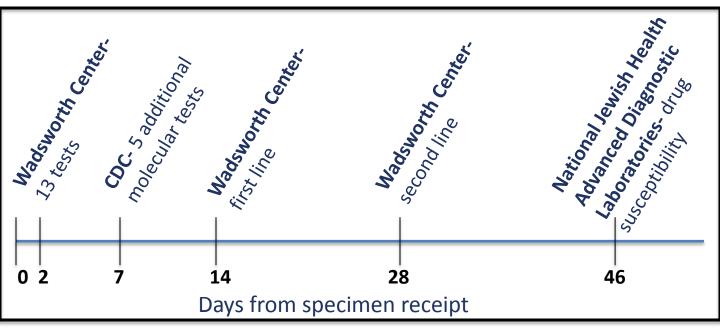




WGS prediction spoligotypes and genotyping with increased resolution



XDR Case (November 2014)



Can we develop <u>one</u> assay capable of generating the same results...and more? Can we do it in <1 week?

XDR Case (November 2014)

Spoligotype: S00062 (777740777760771)

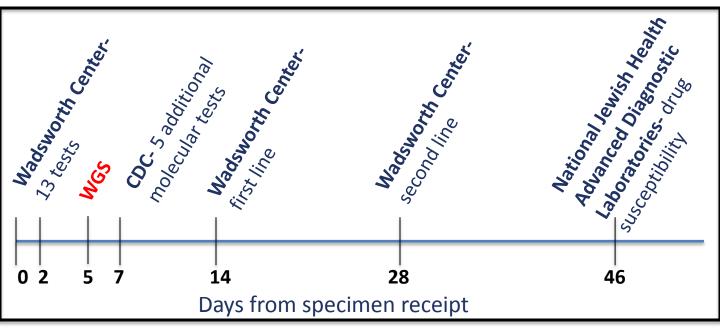
							Pos	ition	SNP	Res.	assoc	ciated	Codo	n AA
chang		Known												
rrs	14732	246	14	00	A	->	G	AMI/S	M			Puta	tive	
mutat	ion*													
gyrA	7362	61	G ·	->	C		FL(21	Glu/	Gln	No	GAG	-> CAG	
gyrA	7582	281	A ·	->	G		FL(94	Asp/	Gly	HC m	utatio	n GAC	-> GGC
gyrA	7585	284	G ·	->	C		FL(95	Ser/	Thr	No	AGC	-> ACC	
gyrA	9304	2003	G ·	->	A		FL(668	Gly/	Asp	No	GGC	-> GAC	
rpoB	76115	55	13	49	C	->	T	RIF	450	Ser/	Leu	HC m	utatio	n TCG
-> TI	G													
rpoC	76494	18	15	79	T	->	G	RIF	527	Leu/	Val	No	TTG .	-> GTG
rpoC	76515	50	17	81	G	->	A	RIF	594	Gly/	Glu	No	GGG -	-> GAG
tlyA	19179 CTA -	972 -> CTG	33		A	->	G	IMA	11	Leu/	Leu	No	Sile	nt
katG	000000000000000000000000000000000000000	578 -> GCC	14	34	G	->	С	INH	478	Ala/	Ala	No	Sile	nt
katG	21551		94	4	G	->	C	INH	315	Ser/	Thr	HC m	utatio	n AGC
-> AC	C													
pncA	22890	149	19	3	T	->	TA	PZA	Inse	rtion	Fram	eshift	: No	
ahpC	27264	109	21	7	G	->	C	INH	73	Asp/	His	No	GAC -	-> CAC
embC	42426 CGC -	543 -> CGT	27	81	C	->	T	EMB	927	Arg/	Arg	No	Sile	nt
embC	42428	303	29	41	G	->	C	EMB	981	Val/	Leu	No	GTG -	-> CTG
embB	42477	730	12	17	G	->	C	EMB/I	NH	406	Gly/	Ala	HC	
mutat	ion	GGC -	> G	CC							(Section # Co.			
embB	42494	108	28	95	G	->	A	EMB/I	NH	965	Pro/	Pro	No	
	Siler	ıt	CCI	G -	->	CCA	3							
embB	42496	578	31	65	C	->	A	EMB/I	NH	1055	Arg/	Ara	No	
Company of	Siler	nt	CG	C -	->	CGA	9000	143600000	155634	STATE TO		Second .	55557	
ethA	43267							CCGCGCG	ETH	Inse	rtion	Fram	eshift	No
gid	44079		26			->		SM	90		Arg	No	CTC .	-> CGC

Lineage Euro-American *M. tuberculosis* X1 family

Drug Resistant phenotype:

- ✓ FLQ (OFL,LVX, MX)
- ✓ RIF
- ✓ INH
- ✓ SM
- **✓** EMB
- ✓ PZA
- ✓ RBT
 - **KAN**
- ✓ AMI CAP (11%)

XDR Case (November 2014)



TB WGS Predicted Drug Resistant phenotype: FLQ (OFL,LVX, MX)
RIF, INH, SM, EMB
PZA, RBT, KAN, AMI

CAP (11%)

Exciting Anecdotal Findings

- 1- MDR identification in ~ 2 weeks on not even known as TB case months before DST available.
- 2- Resolution/ early identification of mixed samples (NTM/TB)
- 3- Resolves inconclusive identifications MTB complex due to missing RD regions
- 4- Resolves issues where pyrosequencing or Sanger sequencing will FAIL due to deletion in target genes
- 5- Finds mutations outside of pyrosequencing region of target genes
- 6- Finds mutations in 2nd line drugs which would never have been found when 1st line drugs are susceptible
- 8- Clears up spoligotyping issues
- 9- Can identify heteroresistance
- 10- Predicts resistance when DST is invalid. WGS is even more valuable because the normal time to susceptibility results is pushed back. In some cases these specimens turn out to have contamination so that DST can never be completed and is canceled.



What does it really cost?

Existing Testing Methods

	Cost per specimen (\$)					
	Reagents	Labor ^{a,b}				
Real-time PCR-	3.98	10.02				
detect MTBC	3.96	10.02				
Real-time PCR-						
detect MTBC	4.62	10.02				
members						
Molecular DST	16.78	42.06				
(rpoB, katG, inhA)	10.78	42.00				
Spoligotyping	7.91	21.52				
Total costs	\$33.29	\$83.62				

\$116.91

Whole Genome Sequencing*

[Cost per sp	ecimen (\$)
	Reagents	Labor ^{a,b}
DNA Extraction	3.49	11.02
MiSeq® library preparation	45.00	67.00
MiSeq® Sequencing	77.00	13.00
Total costs	\$125.49	\$91.02

\$216.51



TB WGS Reports

DETECTED

Concentrated Smear(Ziehl - Neelsen/1,000 X)

(03/13/14): Numerous (>9 acid-fast bacilli per field)

Direct Molecular Detection - Real-time PCR

Mycobacterium tuberculosis complex

DNA by real-time PCR:

Mycobacterium avium complex DNA Not Detected

real-time PCR1

Molecular Identification - Real-time PCR

Mycobacterium tuberculosis complex Mycobacterium tuberculosis

species DNA identified:

Culture

(03/25/14): acid-fast bacillus was isolated

Birect Molecular Drug Susceptibility Detection- Pyrosequencing

Rifampin (rpoB): Mutation present (Ser531Leu) suggests Rifampin resistance.

Result must be confirmed by culture based susceptibility

testing.

Isoniazid (katG): Mutation absent. Culture must be performed for final

susceptibility result.

Isoniazid (inhA): Mutation absent. Culture must be performed for final

susceptibility result.

Identification

(03/26/14): Mycobacterium tuberculosis was identified by culture and

molecular analysis.

Susceptibility Testing for M. tuberculosis complex (MGIT)

Streptomycin [1.0 ug/ml]:

Isoniazid [0.1 ug/ml]:

Rifampin [1.0 ug/ml]:

Ethambutol [5.0 ug/ml]:

Pyrazinamide [100 ug/ml]:

Susceptible

Susceptible

Whole genome sequencing

Molecular Drug Susceptibility Prediction- Whole Genome Sequencing

Rifampin

rpoB: Mutation present: Ser531Trp suggests resistance

Note: XX% of isolates in our in-house evaluation of XX clinical isolates with this mutation are resistant.

Isoniazid

Mutation present: Ser315Thr suggests resistance katG:

Note: XX% of isolates in our in-house evaluation of XX clinical isolates with

this mutation are resistant. inhA:

Mutation present: C-15T suggests resistance Note: XX% of isolates in our in-house evaluation of XX clinical isolates with this mutation are resistant.

Pyrazinamide

Mutation present: Trp68Arg suggests resistance pncA:

Ethambutol

embB: Mutation present: Met306lle suggests resistance

Note: XX% of isolates in our in-house evaluation of XX clinical isolates with this mutation are resistant.

Streptomycin rrs:

Mutation present A>C at 513 suggests resistance Note: XX% of isolates in our in-house evaluation of XX clinical isolates with this mutation are resistant.

gyrA No Mutation

gyrB No mutation

Kanamycin eis:

Ofloxacin

No mutation rrs:

No mutation

Ethionamid

Mutation: C-15T suggests resistance inhA: Note: XX% of isolates in our in-house evaluation of XX clinical isolates with this mutation are resistant.

Disclaimer: A negative result (e.g. no mutation) does not rule out contributory mutations present elsewhere in the genome. CULTURE MUST BE PERFORMED FOR FINAL SUSCEPTIBILITY RESULT.

Whole Genome Sequencing of TB: A "One Stop Shop"

WGS

Single assay

Species identification

Genotyping (more accurate)

Drug resistance mutations

(more comprehensive)

COST

Estimated around \$100-\$200 per sample

TURNAROUND TIME

DNA preparation (1 days)
WGS result (4-5 days)

Reality for TB Cultures!

Next challenge TB specimens



Future Directions WGS TB

- Finalize validation and implement WGS for TB culture testing
- Evaluate TAT, sample numbers
- Data interpretation/ notes
- LIMS importing
- NCBI
- Data Storage/ assessing data over time
- TB Primary specimens



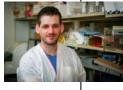
Wadsworth Center NEW YORK STATE DEPARTMENT OF HEALTH

MYCOBACTERIOLOGY LAB

Vincent Escuyer
Donna Kohlerschmidt
Michelle Isabelle
Susan Wolfe
Dennis Biggins

BACTERIOLOGY LAB

Joe Shea
Tanya Halse
Tammy Quinlan
Justine Edwards
Linda Gebhardt





APPLIED GENOMIC TECHNOLOGIES CORE

Matt Schudt
Patrick VanRoey
Pascal Lapierre
Mike Palumbo



FUNDING

Wadsworth Center, NYSDOH
Public Health Genomics Initiative



National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention



R03 NIH- Use of whole genome sequencing for tuberculosis diagnostics





Master of Science in Laboratory Sciences



Imagine a career:

- Protecting and improving the health of our communities
- Diagnosing diseases of public health importance
- Developing cutting edge testing methods





Master of Science in Laboratory Sciences

Department

of Health

Wadsworth Center



24 Month Full-time intensive program:

- 27 didactic credits
- 8 eight-week rotations through different scientific focus areas
- 1 Capstone research project

Tuition and Scholarships:

- No tuition fees
- Maximum of 4 scholarships of \$10,000 will be awarded each year based on academic merit.
 - * Applications due: March 2016

Questions?

