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

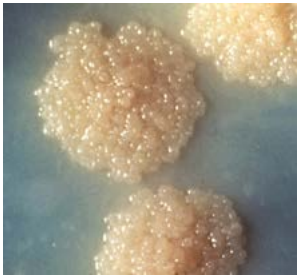
Whole Genome Sequencing for TB Diagnostics at the Wadsworth Center

May 19, 2015

Kimberlee Musser, PhD
Chief, Bacterial Diseases

TB Background

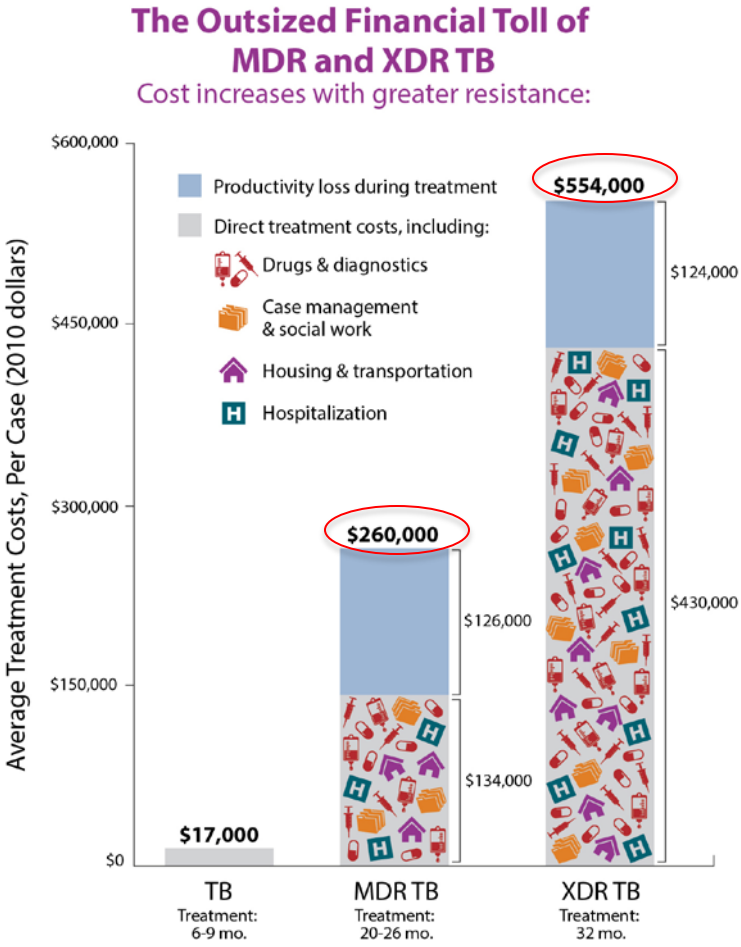
- Tuberculosis is an ancient disease that has been evolving alongside humans
- Caused by *Mycobacterium tuberculosis* and other MTBC species
- Roughly one third of the world's population is infected with TB
- 2013: 9 million new infections, 1.5 million deaths
- Second only to HIV/AIDS as a worldwide killer



Drug Resistant Tuberculosis is a Global Health Concern

- MDR: resistance to at least rifampin and isoniazid
- XDR: also resistant to at least 3 second line drugs

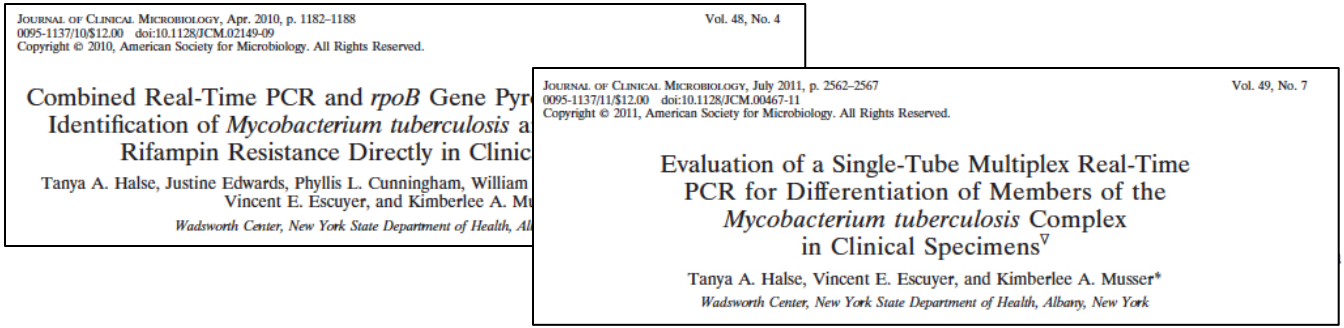
- Standard drug susceptibility testing can take weeks to months to complete
- The tests for drug resistance have certain limitations
  - Number of targets



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

Why perform WGS on *Mycobacterium tuberculosis*?

- **Turn-around time-**Culture and Drug Susceptibility Testing (DST) - weeks to months
- **Early molecular diagnostic testing-**Important to improved TB treatment and TB prevention
- **More comprehensive results-**Detect mixed infections, many predictors of DR, subpopulations of TB may predict success of treatment
- **Cost effective-**Real-time PCR targets (9), Pyrosequencing targets (6), Spoligotyping (Luminex)



JOURNAL OF CLINICAL MICROBIOLOGY, Apr. 2010, p. 1182–1188 Vol. 48, No. 4
0095-1137/10/$12.00 doi:10.1128/JCM.02149-09
Copyright © 2010, American Society for Microbiology. All Rights Reserved.

Combined Real-Time PCR and *rpoB* Gene Pyr
Identification of *Mycobacterium tuberculosis* a
Rifampin Resistance Directly in Clinic

Tanya A. Halse, Justine Edwards, Phyllis L. Cunningham, William
Vincent E. Escuyer, and Kimberlee A. Mu

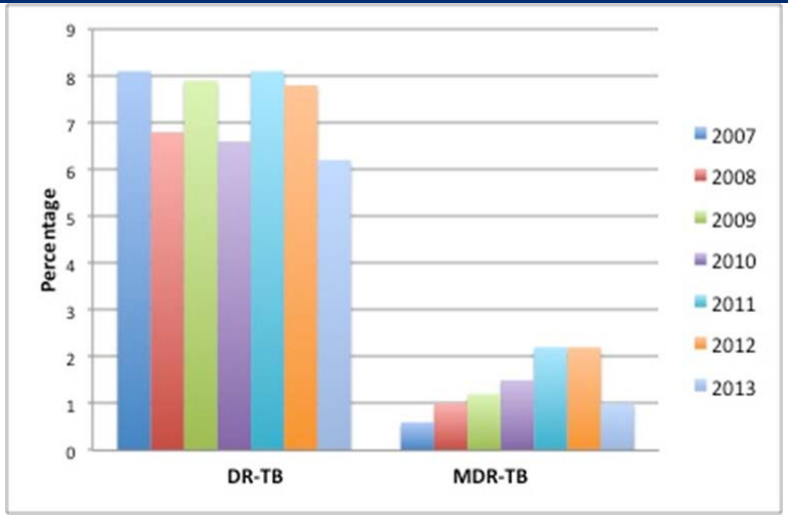
Wadsworth Center, New York State Department of Health, Al

JOURNAL OF CLINICAL MICROBIOLOGY, July 2011, p. 2562–2567 Vol. 49, No. 7
0095-1137/11/$12.00 doi:10.1128/JCM.00467-11
Copyright © 2011, American Society for Microbiology. All Rights Reserved.

Evaluation of a Single-Tube Multiplex Real-Time PCR for Differentiation of Members of the *Mycobacterium tuberculosis* Complex in Clinical Specimens▽

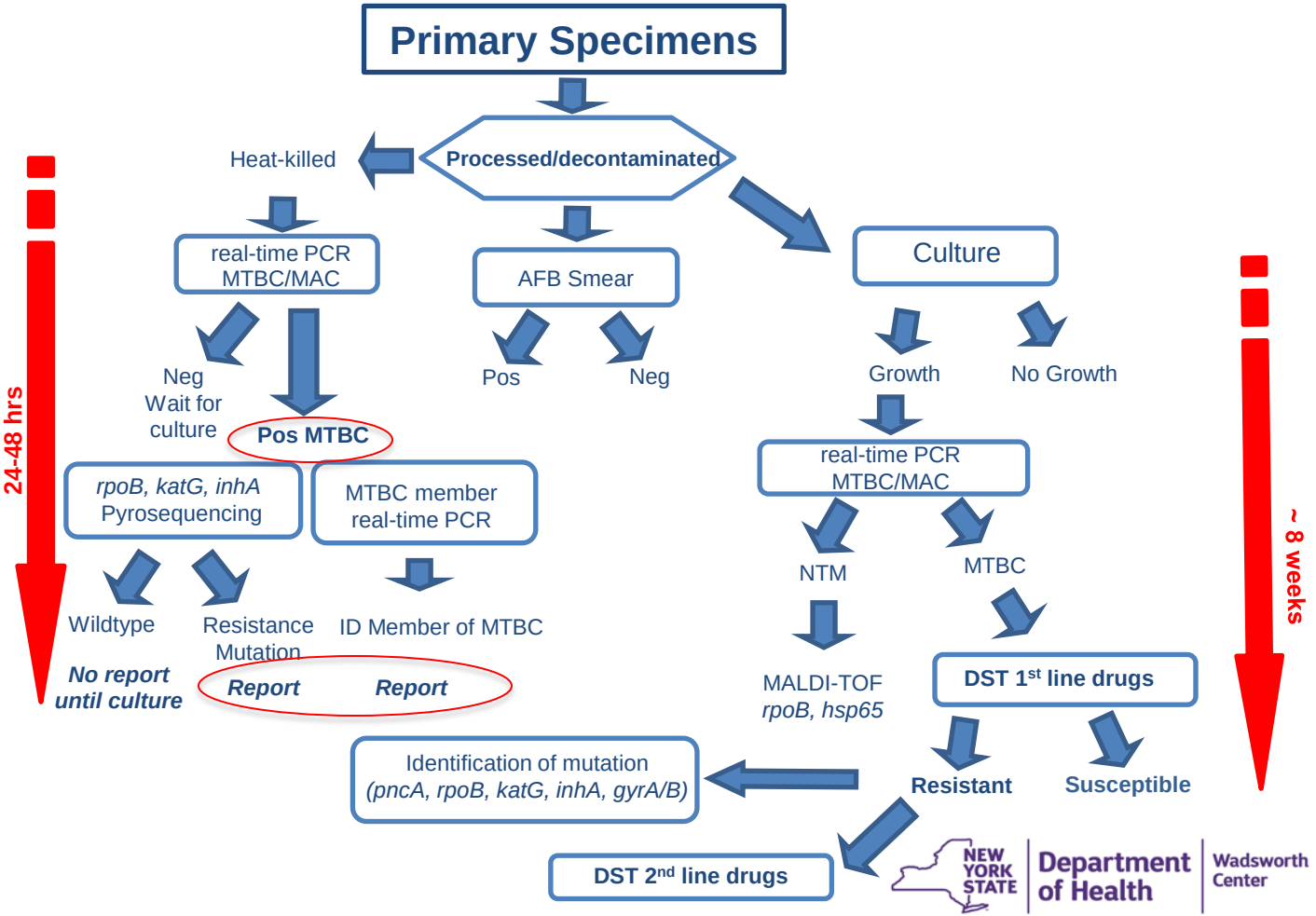
Tanya A. Halse, Vincent E. Escuyer, and Kimberlee A. Musser*

Wadsworth Center, New York State Department of Health, Albany, New York



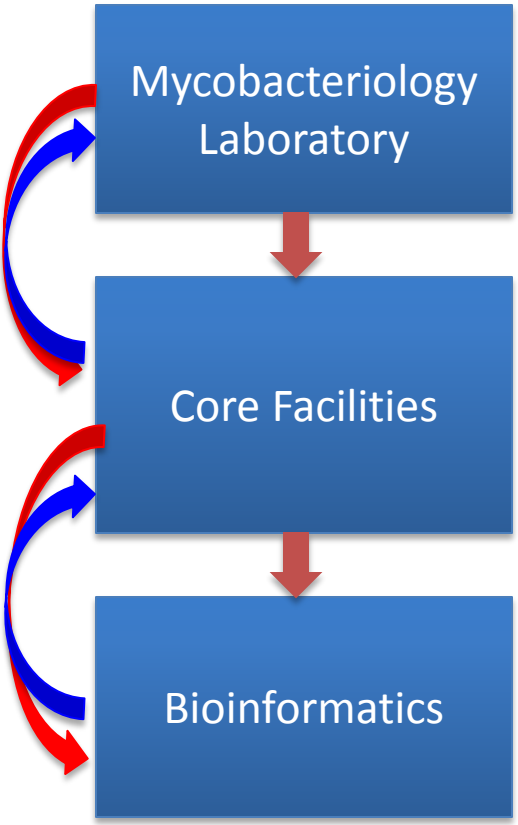
	2007	2008	2009	2010	2011	2012	2013	2014
TB Cases	1175	1200	1007	954	910	864	872	786

Testing Algorithm



NEW YORK STATE Department of Health | Wadsworth Center

Validating a WGS assay for TB



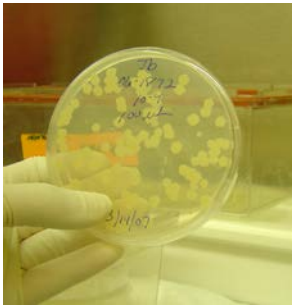
Selecting validation approach, culture, optimization of DNA preparation

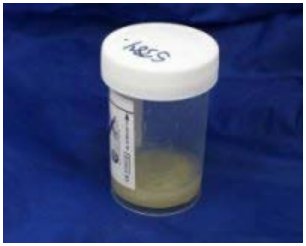
Library preparation and Miseq sequencing, optimizing, planning overall decisions for balancing runs (1 Ion Torrent, 3 Miseq instruments)

Development, testing and continual improvements to pipeline, data storage (Freely available academic software, background in microbiology)

Where to start?

- Isolates
  - Solid **(initial 68)**
  - MGITs
- Primary specimens
  - sputum **(preliminary)**
  - other





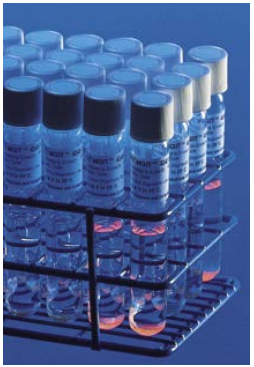
Need to keep in mind available testing volumes, what is needed for other tests, archiving, etc...

How Can We Mimic a Clinical Isolate?

- Grow 2 different strains in MGIT tube
  - *M. tuberculosis* : ATCC strain (ATCC)
  - *M. bovis BCG* : patient strain (BCG)
- Aliquots made and heat killed before leaving BSL-3

- A positive tube contains approximately 10^5 to 10^6 colony-forming units per milliliter (CFU/mL)
- 0,1,2,3 days since flag positive
- Isolates submitted to the lab are usually <2mL

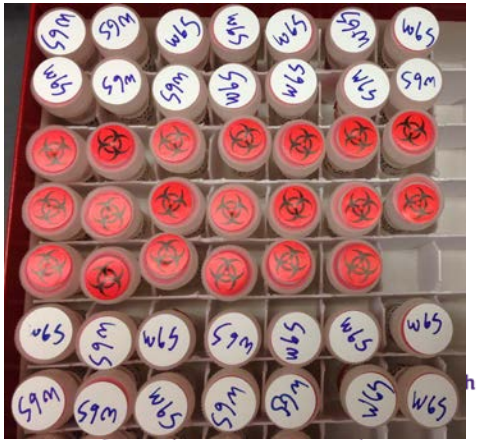
MGIT



HK ATCC x2

HK BCG x2

Live BCG x3



A New Tool box

- Modified an existing real-time PCR assay
- Qubit
  - WGS Core requests ~1ng/uL
- NanoDrop
  - Purity and concentration
- Communication with Sequencing Core, bioinformatician

Breaking TB Open is Critical for DNA Extraction

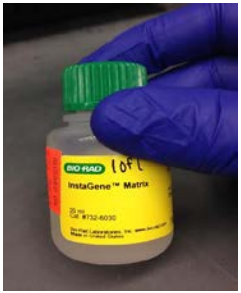
Important TB Characteristics

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

Initial Methods

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

Success- InstaGene matrix and Tissue Homogenizer



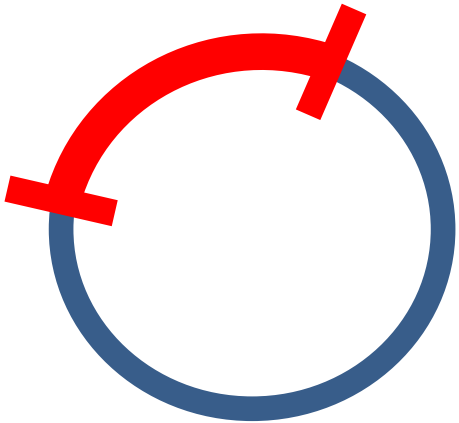
- **InstaGene matrix (Chelex resin)**
  - **When to add beads, when to beat, boil time etc...**

- **Fastprep tissue homogenizer**
  - **Good enough yield to provide reliable WGS data even with 0 day MGIT**

Sample 5 sent for WGS:
Depth: 96.77X
Coverage: 99.8%

Successful WGS

- **Depth: 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**
- **Coverage: A percentage that describes how much of the genome was sequenced**
  - **Best 100%**

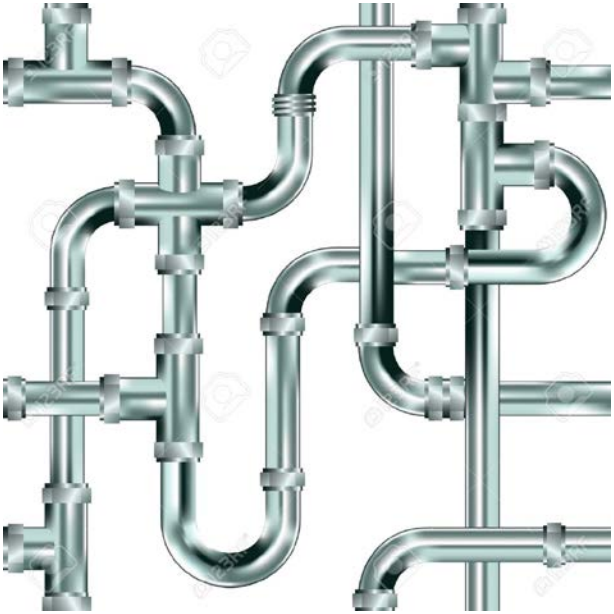


Library Preparation is Another Key Factor

- Votintseva *et al.* suggested using 15 cycle library preparation
  - 2015 paper about WGS of early positive MGIT
- All 4 samples failed using standard 12 cycles
- Successful using 15 cycles
  - Depth: 13-30X
  - Coverage: >96%
- Normalization

Sample	Date	Method	stock ng/uL
BCG (0 day)	2/25	InstaGene	0.268
	2/25	InstaGene	0.344
	2/25	InstaGene	0.346
	2/25	InstaGene	0.38

Bioinformatics Pipeline

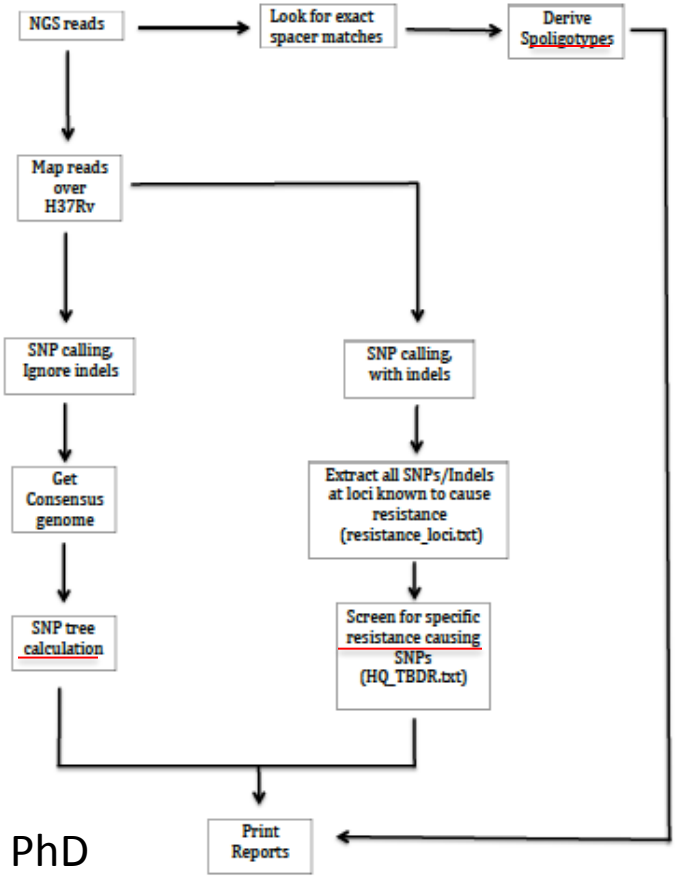


Pascal LaPierre, PhD
Michael Palumbo, PhD

NEW YORK STATE | Department of Health | Wadsworth Center

Pipeline Schematic

Pascal Lapierre, PhD



NEW YORK STATE OF OPPORTUNITY. Department of Health

ANDREW M. CUOMO
Governor

HOWARD A. ZUCKER, M.D., J.D.
Acting Commissioner

SALLY DRESLIN, M.S., R.N.
Executive Deputy Commissioner

May 2015

Validation of Next Generation Sequencing (NGS)-Based Methods for Identification and/or Characterization of Infectious Agents (Isolates only)

The following guidelines are applicable to WGS using NGS-based methods for identification and/or characterization of infectious agent isolates. These guidelines should be used in conjunction with and not in lieu of the existing microbiology molecular guidelines: (http://www.wadsworth.org/labcert/TestApproval/forms/Microbiology_NAAT_Checklist.pdf).

Overall, clinical validation of NGS assays follows the same basic principles for validating most other complex molecular diagnostic procedures. It is anticipated that these guidelines will evolve as the field matures and gains experience. Please make sure you use the most up-to-date version of these guidelines.

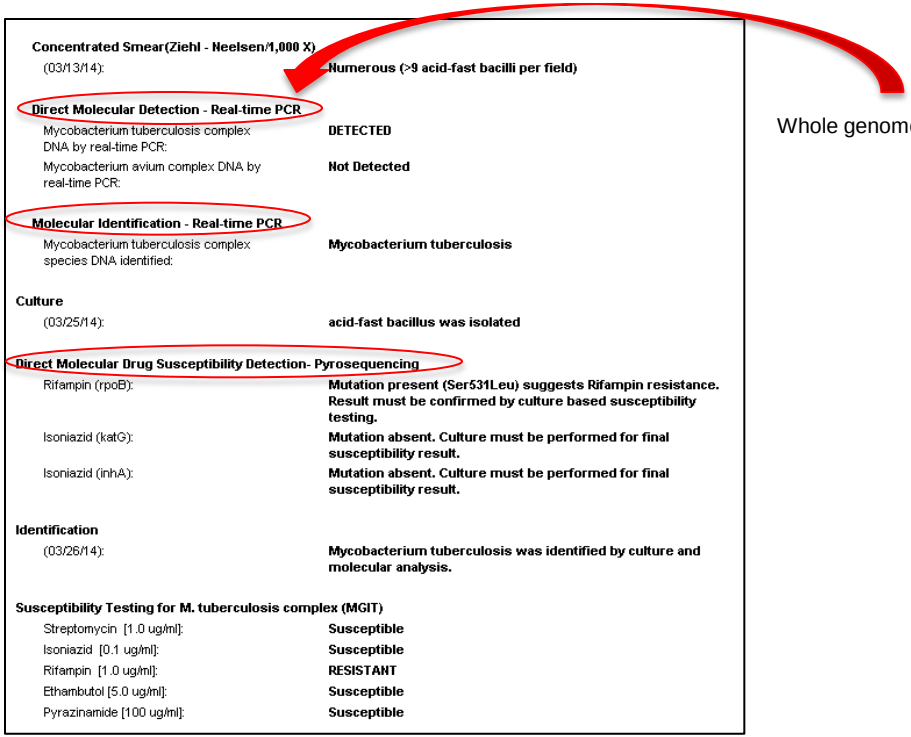
General Requirements:

- The detailed standard operating procedure manual (SOPM) must include all relevant quality assurance and proficiency testing details for this test. The SOPM must include a step-by-step description of all steps involved, from template to library preparation to data analysis and interpretation of results.
- The SOPM must include all expected reporting and reflex testing scenarios. The SOPM must clearly define what will be reported from the NGS/WGS results and what will not be reported. It must include or refer to the procedure(s) for confirmation testing, including clear criteria for when confirmation is required.
- The SOPM must include statements that identify the limitations of the assay.

Reporting

- In addition to the actual results, specimen reports must include an interpretation of the findings. Representative examples of specimen reports must be included with the submission.

TB WGS Reports



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:	**DETECTED**
Mycobacterium avium complex DNA by real-time PCR:	**Not Detected**
Molecular Identification - Real-time PCR	
Mycobacterium tuberculosis complex species DNA identified:	**Mycobacterium tuberculosis**
Culture	
(03/25/14):	**acid-fast bacillus was isolated**
Direct 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]:	**Susceptible**
Isoniazid [0.1 ug/ml]:	**Susceptible**
Rifampin [1.0 ug/ml]:	**RESISTANT**
Ethambutol [5.0 ug/ml]:	**Susceptible**
Pyrazinamide [100 ug/ml]:	**Susceptible**

Whole genome sequencing

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

WGS

Single assay

Identification

Species Differentiation

Genotyping (more accurate)

Drug resistance mutations (more comprehensive)

COST

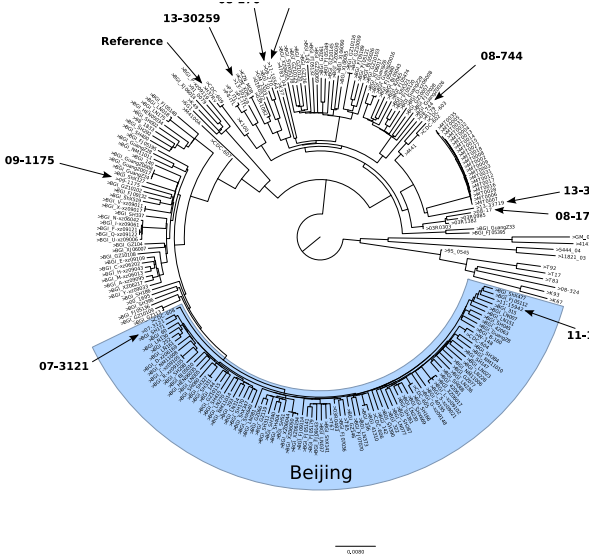
Estimated around $100 per sample

TURNAROUND TIME

DNA preparation (1 days)

WGS result (4-5 days)

WGS Prediction Spoligotypes with increased resolution



Sample #	Extraction	AccId	Phenotype (DST)	Resistance Associated mutations (WGS)	Other mutations noted (WGS)	WGS Spoligotype	CLIMS spoligo	Do results correlate?
21	CTAB	14-7386	SM (low level in 2nd line), INH, RIF, RBT	rpoB aa 450 (RIF) rpsL aa 88 (SM) katG aa 315 (INH)		S00034	Scheduled 12/15/14	Yes
22	CTAB	14-7387	INH (low level), RIF, RBT	rpoB aa 445 (RIF) intergenic hyp-fabG1 aa. -15 (INH)		S00197	S00197	Yes
23	CTAB	14-7388	SM, INH, PZA, EMB (2nd line only), KM, AN,	rpoB aa 450 (RIF) rpsL aa 43 (SM) katG aa 315 (INH) embB aa 306 (EMB/INH)	rrs (AN/SM) pncA aa 119 (PZA)	S00034	S00034	Yes
24	CTAB	14-7391	SM, EMB, CAP, KM, AN	rrs (AN/SM)	embC aa 981 (EMB) gid aa 97 (SM)	S00241	S00241	Yes
25	CTAB	14-7392	SM, RIF, EMB, PZA, KM, RBT, FLQ	gyrA aa 94 (FLQ) rpoB aa 450 (RIF) rpoC aa 452 (RIF) embB aa 306 (EMB/INH) embB aa 497 (EMB/INH)	gid aa 92 (SM) rrs (AN/SM) rpsL aa 43 (SM) pncA TG-TGG (PZA) eis promoter (AN/KM)	S00034	S00034	Yes
26	CTAB	14-7393	INH, RIF, PZA, EMB 5.0 (2nd line only), ETA, RBT,	rpoB aa 450 (RIF) intergenic hyp-fabG1 aa. -15 (INH/ETA) embB aa 306 (EMB/INH)		S00002	S00002	Yes
27	CTAB	14-7395	SM, INH, RIF, EMB, PZA, RBT	rpoB aa 450 (RIF) rpsL aa 43 (SM) katG aa 315 (INH) pncA aa 12 (PZA) embB aa 306 (EMB/INH)		S00034	S00034	Yes
28	CTAB	14-7396	SM, INH, RIF, PZA, EMB (2nd line only), RBT	rpoB aa 450 (RIF) rpsL aa 43 (SM) katG aa 315 (INH) embB aa 306 (EMB/INH)	pncA aa 151 (PZA)	S00034	S00034	Yes

A Glimpse of the One Stop Shop in Action

AccId	Phenotype (DST)	Resistance Associated mutations (WGS)	Other mutations noted (WGS)	WGS Spoligotype	CLIMS spoligo	Do results correlate?
		embB aa 306 (EMB/INH)				
13-19614	INH, RIF, EMB, PZA, RBT 0.5 only, FLQ	gyrA aa 94 (FLQ) rpoB aa 432 (RIF) rpoC aa 483 (RIF) embB aa 306 (EMB/INH)	pncA aa 182 (PZA) intergenic oxyR-ahpC aa. -48 (INH)	S00002 (777777777760771)	S00005 (777777777720771)	Per Linda, spoligotype off by one spacer
13-21110	INH	rpoB aa 452 (RIF)	katG aa 463 (INH) katG aa 121 (INH)	S00034	S00034	Yes/DST RIF susc (Leu511Pro-known to not correlate with DST)
08-532	EMB (SIRE only)	kasA aa 269 (INH)	embB aa 937 (EMB/INH)		S00615	Yes
11-23189	Pan-susceptible; RD12 negative	None			S02110	Yes
12-2066	SM, INH, RIF, EMB, RBT	rpoB aa. 445 (RIF) rpsL aa 88 (SM) katG aa. 315 (INH) embB ss. 306 (EMB)		S01770	S01770	Yes
12-23774	SM (SIRE only), INH, RIF	rpoB aa. 435 (RIF) intergenic hyp-fabG1 aa. -8 (INH) katG aa. 315 (INH)	gid aa 79 (SM)	S00002	S00002	Yes
12-28162	INH (low level)	No HC mutation	embB aa. 937	S00009	S00009	Yes

Evolving Pipeline

- **INH resistant, pyrosequencing failed-** WGS *katG* gene missing
- **Real-time PCR failed to differentiate MTBC-** Regions of Deletion (RD 12, 4, 9) deleted
- **Low level isoniazid, streptomycin-** accumulating data
- **Spoligotyping shows faint bands-**SNPs in spacer
- **Bedaquiline mutations detected**

XDR Case (November 2014)

- **Wadsworth Center-** 13 tests **(0-2 days)**
- **CDC-** 5 additional molecular tests **(7 days)**
- **Wadsworth Center-** first line **(14 days)**
- **Wadsworth Center-** second line **(28 days)**
- **National Jewish Health Advanced Diagnostic Laboratories-** drug susceptibility **(46 days)**

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

XDR 14-36340

Gene	Genome position	Gene Position	SNP	Res. associated	Codon AA change	Known mutation?
rrs	1473246	1400 A -> G	AMI/SM			Putative mutation*
gyrA	7362 61	G -> C	FLQ	21	Glu/Gln	No	GAG -> CAG	
gyrA	7582 281	A -> G	FLQ	94	Asp/Gly	HC mutation	GAC -> GGC	
gyrA	7585 284	G -> C	FLQ	95	Ser/Thr	No	AGC -> ACC	
gyrA	9304 2003	G -> A	FLQ	668	Gly/Asp	No	GGC -> GAC	
rpoB	761155	1349 C -> T	RIF	450	Ser/Leu	HC mutation	TCG -> TTG	
rpoC	764948	1579 T -> G	RIF	527	Leu/Val	No	TTG -> GTG	
rpoC	765150	1781 G -> A	RIF	594	Gly/Glu	No	GGG -> GAG	
tlyA	1917972 CTA -> CTG	33 A -> G	AMI	11	Leu/Leu	No	Silent	
katG	2154678 GCG -> GCC	1434 G -> C	INH	478	Ala/Ala	No	Silent	
katG	2155168	944 G -> C	INH	315	Ser/Thr	HC mutation	AGC -> ACC	
pncA	2289049	193 T -> TA	PZA	Insertion	Frameshift	No		
ahpC	2726409	217 G -> C	INH	73	Asp/His	No	GAC -> CAC	
embC	4242643 CGC -> CGT	2781 C -> T	EMB	927	Arg/Arg	No	Silent	
embC	4242803	2941 G -> C	EMB	981	Val/Leu	No	GTG -> CTG	
embB	4247730	1217 G -> C	EMB/INH	406	Gly/Ala	HC mutation	GGC -> GCC	
embB	4249408 Silent	2895 G -> A CCG -> CCA	EMB/INH	965	Pro/Pro	No		
embB	4249678 Silent	3165 C -> A CGC -> CGA	EMB/INH	1055	Arg/Arg	No		
ethA	4326718	756 CCGCG -> CCGCGCG	ETH	Insertion	Frameshift	No		
gid	4407934	269 T -> G	SM	90	Leu/Arg	No	CTC -> CGC	

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

Spoligotype: S00062 (777740777760771)

Lineage Euro-American
***M. tuberculosis* X1 family**

Future Directions WGS Bacteria

- Finalize validation and implement WGS for TB MGIT testing
- TB Primary specimens
- Foodborne bacteria- FDA GenomeTrakr, WGS AMD- PulseNet
- *Legionella* spp., *E. coli*, *Staphylococcus aureus*, *Streptococcus* spp., antibiotic resistance- (CRE)

Acknowledgements

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

APPLIED GENOMIC TECHNOLOGIES CORE

Matt Schudt
Patrick Van Roey
Pascal Lapierre
Mike Paulmbo

National Jewish Health
Advanced Diagnostic Laboratories

Max Salfinger

FUNDING

Wadsworth Center, NYSDOH
Public Health Genomics Initiative

CDC

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

NIH National Institutes of Health
Office of Extramural Research

R03 NIH- Use of whole genome sequencing for tuberculosis diagnostics

NEW YORK STATE Department of Health | Wadsworth Center