

Next Generation Sequencing Workflow and Applications for *Mycobacterium tuberculosis*

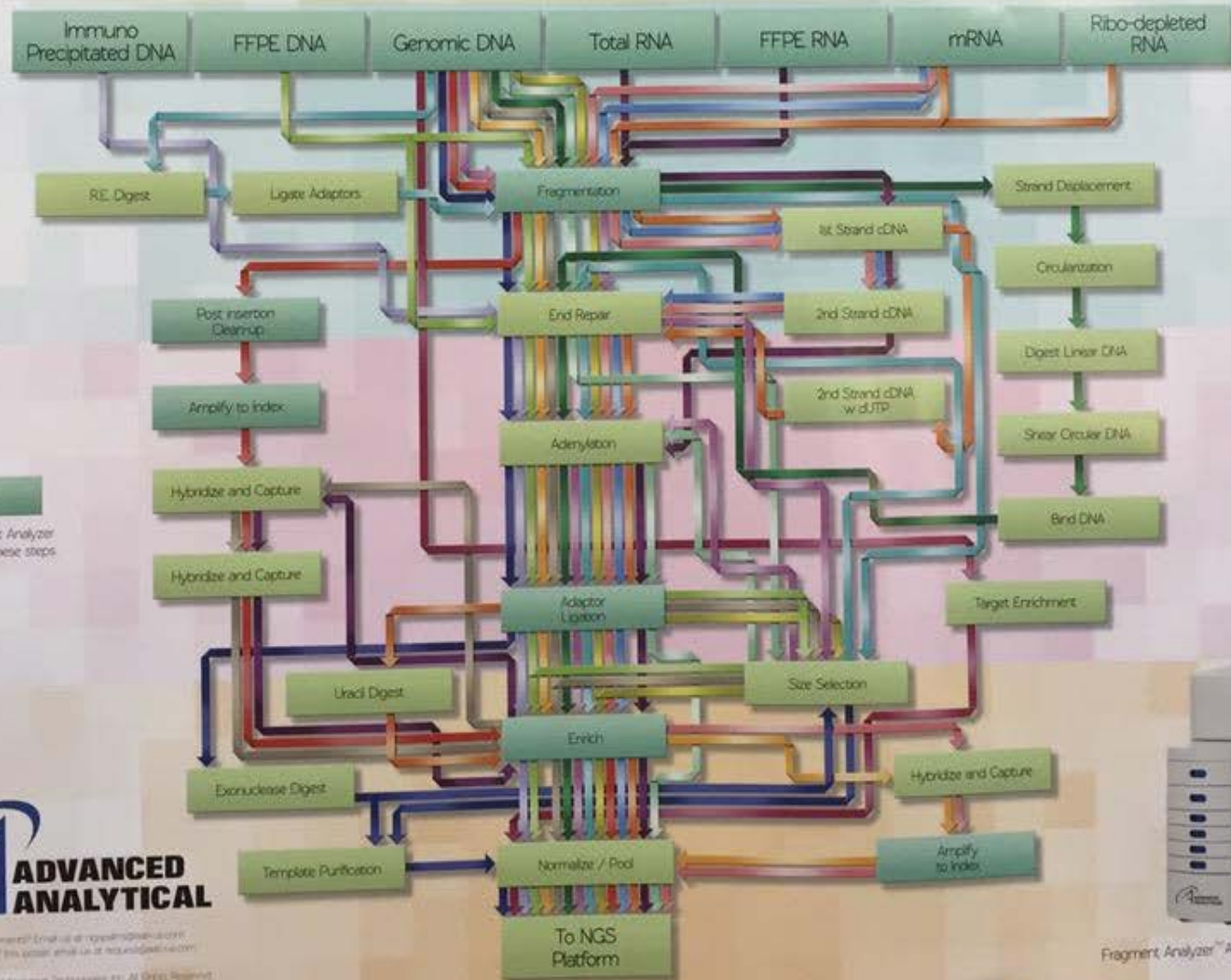
Jamie Posey, PhD
Applied Research Team Lead

June 8, 2015

The Paths to NGS

Workflow Types

- | | | | |
|--------------------------|----------------------------------|----------------------------------|---------------------------------|
| Immuno Precipitated gDNA | Single Molecule gDNA | Double Hybridize Enrichment gDNA | Total RNA / mRNA |
| FFPE gDNA | Transposon-Based gDNA | Mate Paired gDNA | miRNA |
| Target Enrichment | Low Input gDNA | PCR-Free gDNA | Small Hybridize Enrichment mRNA |
| Restriction Digated gDNA | Single Hybridize Enrichment gDNA | Standard gDNA | miRNA - Ribo-Depleted RNA |



Fragment Analyzer used in these steps



Questions? Comment or Email us at info@advancedanalytical.com or support@advancedanalytical.com

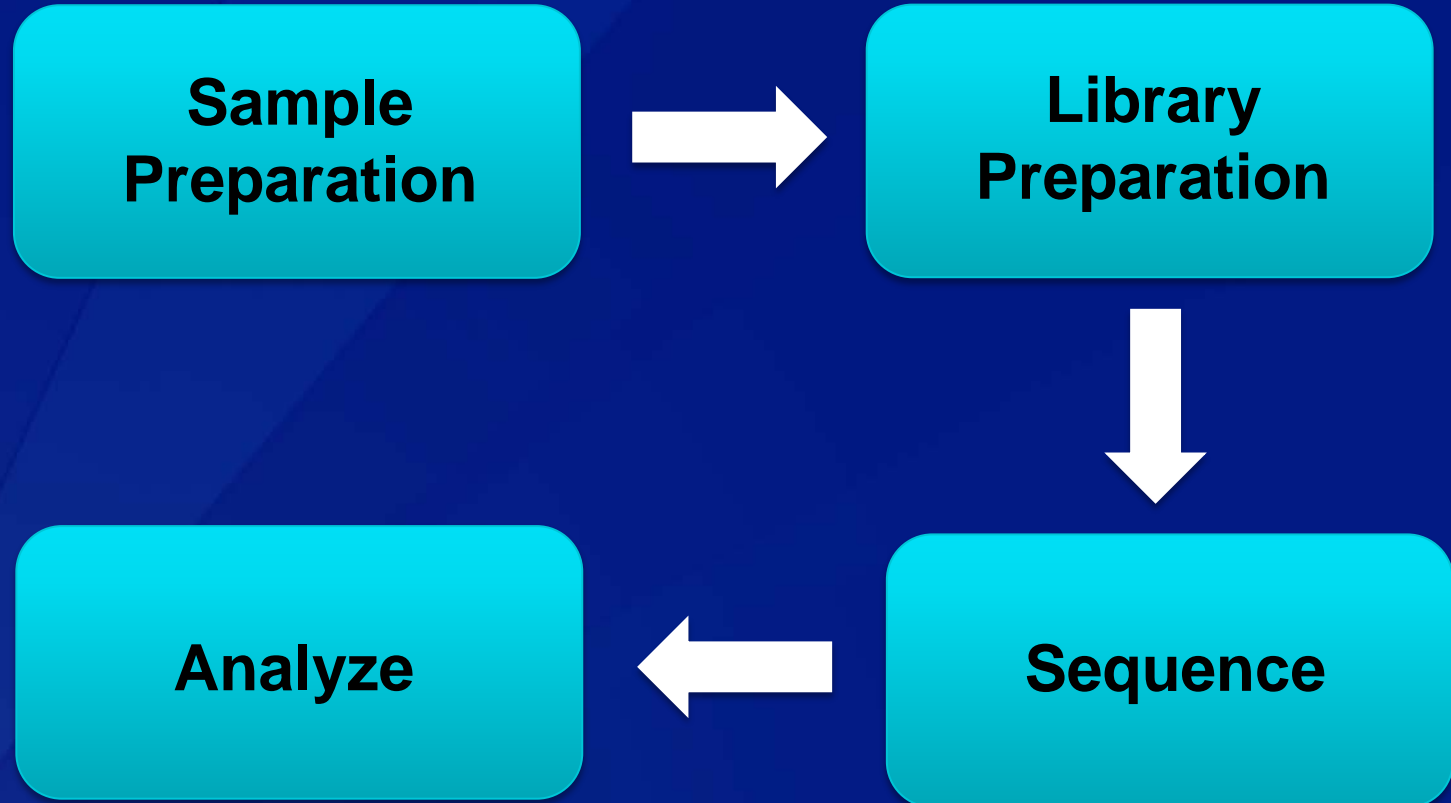
Copyright 2015 Advanced Analytical Technologies, Inc. All Rights Reserved



Fragment Analyzer™ Automated CE Instrument



Next Generation Sequencing (NGS) Workflow



Sample Preparation for Whole Genome Sequencing (WGS)

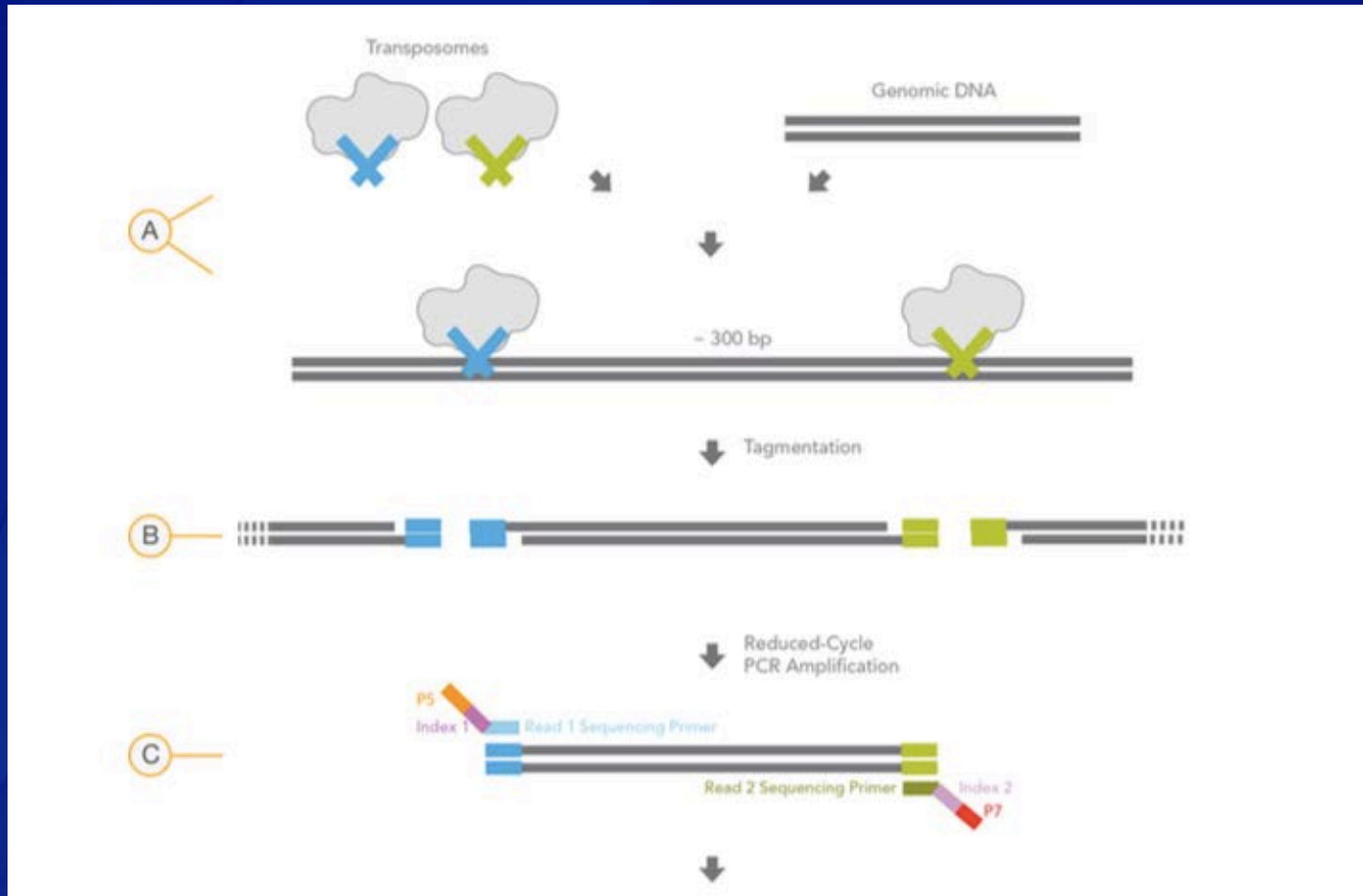
❑ Isolation of DNA

- Chemical lysis (CTAB)
- Mechanical lysis (FastPrep-24)
- Purify DNA

❑ Shear genomic DNA

- Physical
- Enzymatic

Library Preparation for Illumina Platforms



Sequence Libraries



Illumina

<https://www.youtube.com/embed/HMyCqWhwB8E?iframe&rel=0&autoplay=1>

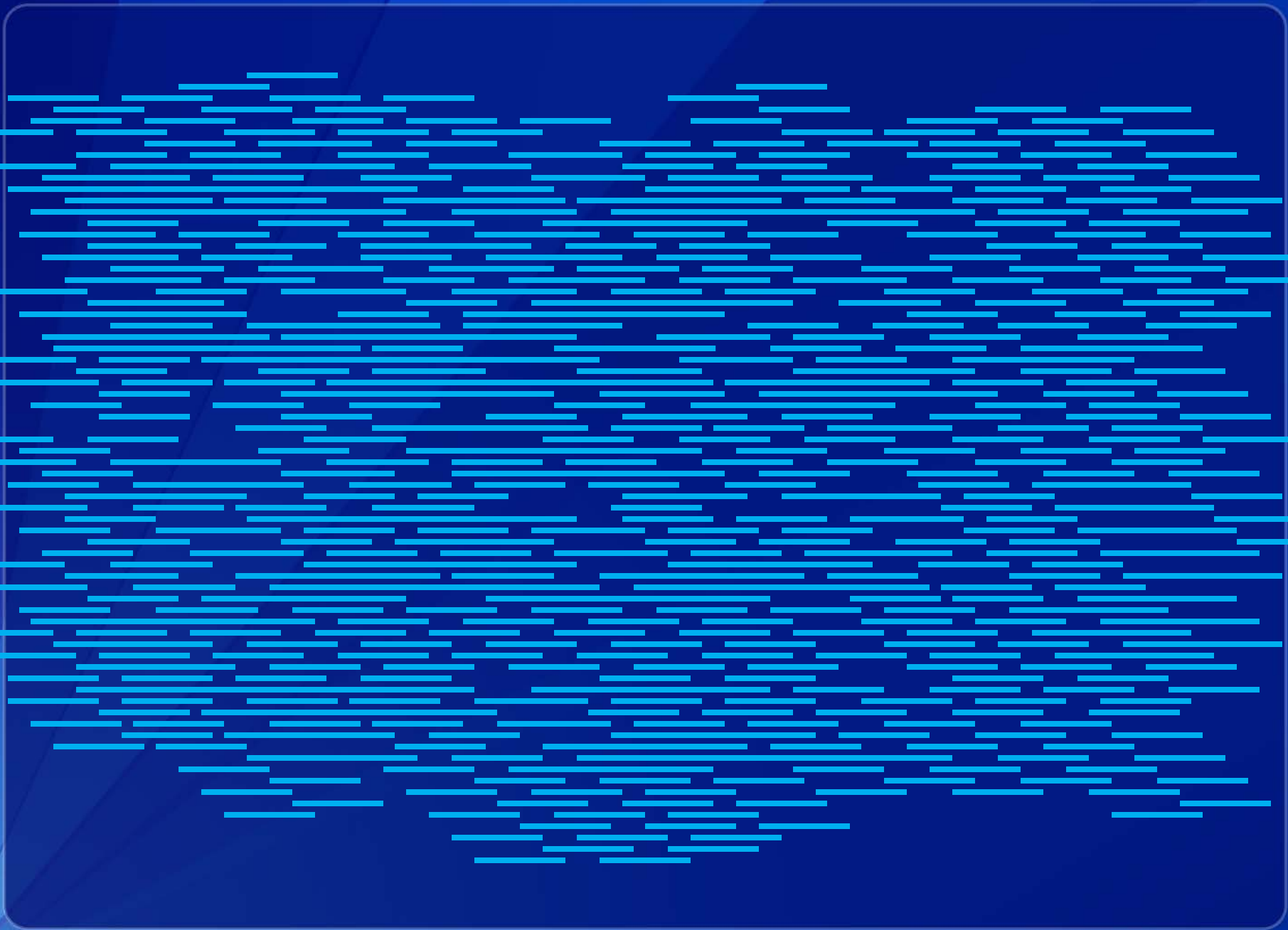
https://www.youtube.com/watch?t=10&v=L_jAtDSB8kA

PacBio SMRT

<https://www.youtube.com/watch?v=v8p4ph2MAvI>

Ion Torrent PGM by Life Technologies

<https://www.youtube.com/watch?v=WYBzbxIfuKs>



Genome Assembly

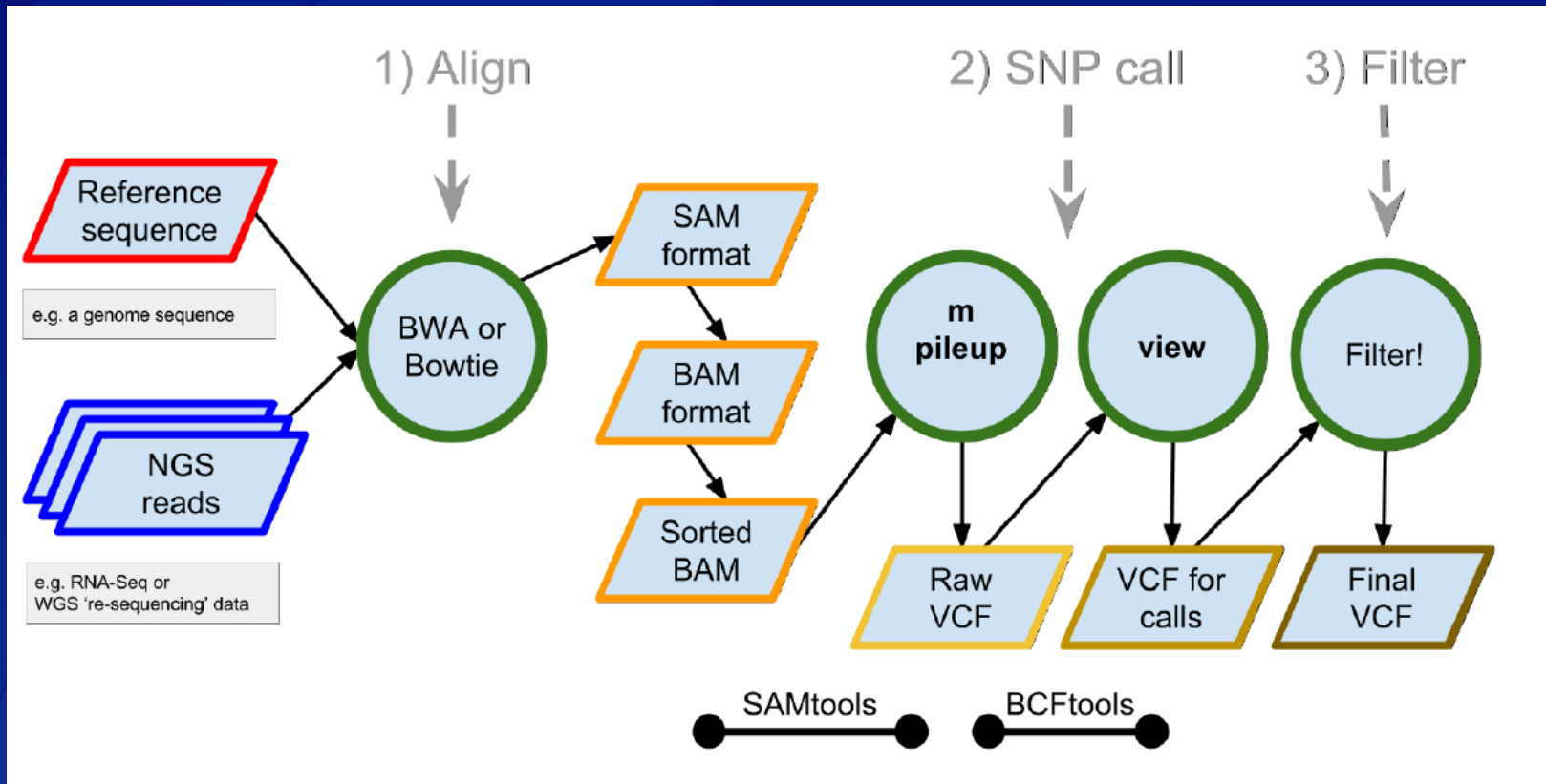


<http://www.jigsawplanet.com/?rc=createpuzzle>

Bioinformatic Tools



Simple Variant Call Pipeline



Examples of Commercial Products



DNASTAR®



BaseSpace®
Genomics Cloud Computing

The companies and products depicted here are not endorsed by CDC.

Reference-Guided (Mapped) Assembly



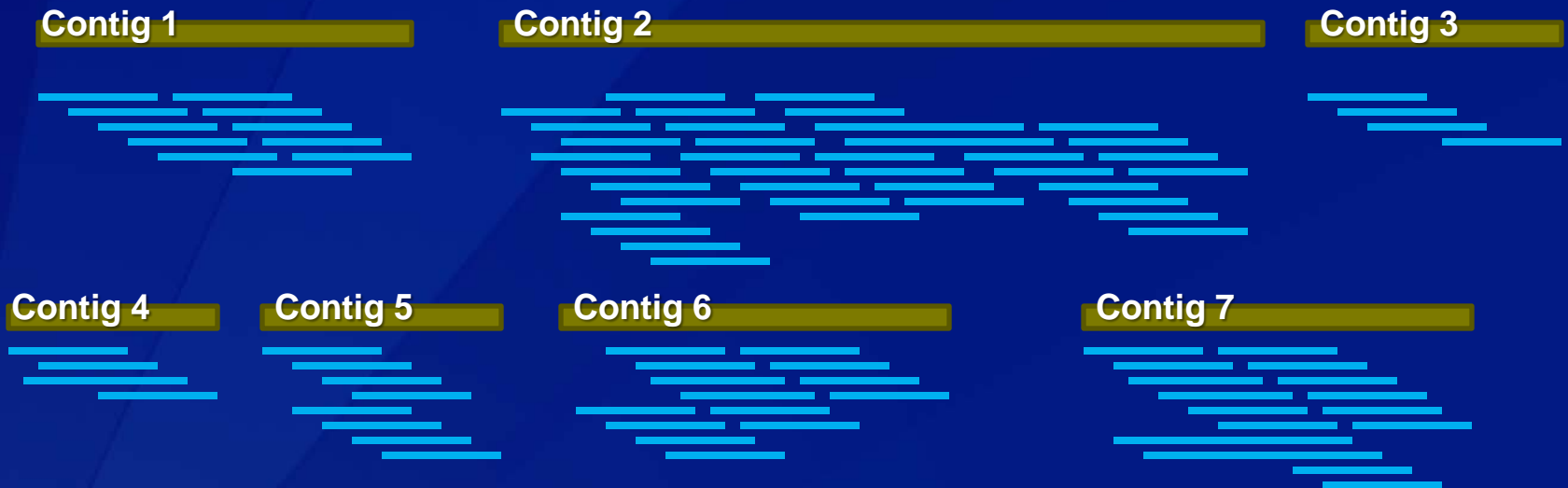
UNMAPPED READS

1. Sequences not present in the reference.
2. Plasmids or other extrachromosomal.
3. DNA Structural Variation/Rearrangement

ADVANTAGES: Relatively fast, well-suited to highly-conserved genomes.

DISADVANTAGES: Issues with high diversity, mobile elements

De-Novo Assembly

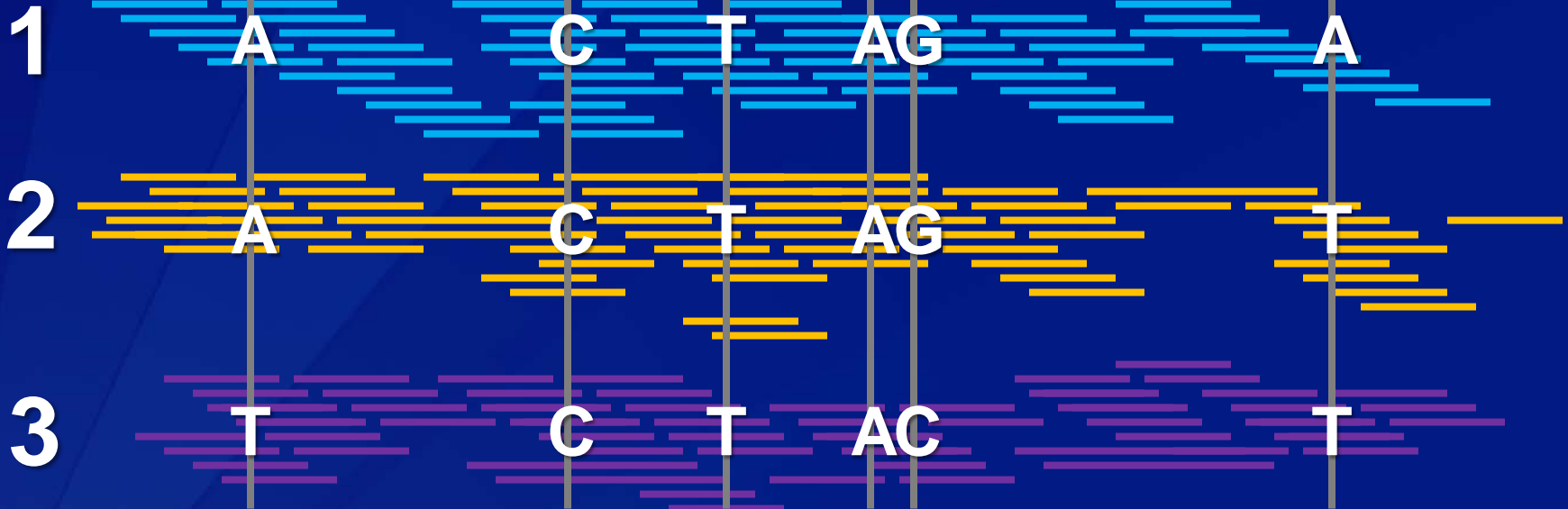


ADVANTAGES: Reference agnostic: assembles all the reads it can into contigs.

DISADVANTAGES: Doesn't always get things right. Repeat sequences etc.

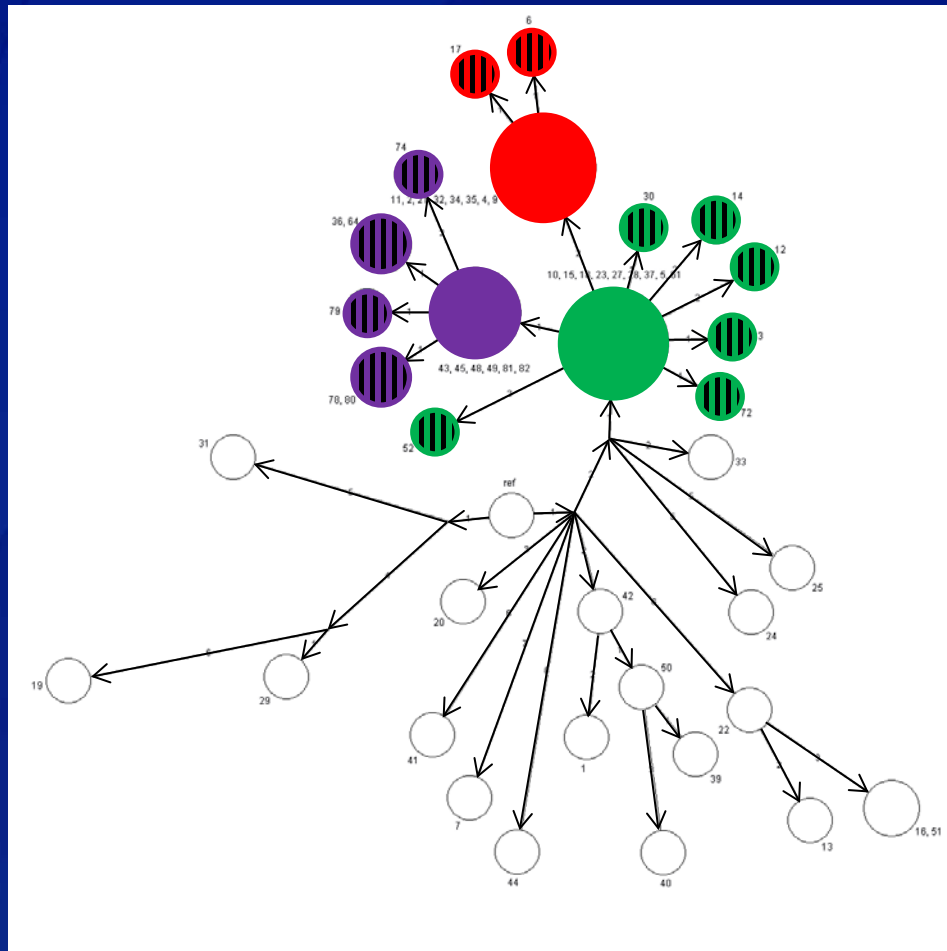
Whole Genome SNP Typing

Reference Sequence/Core Genome

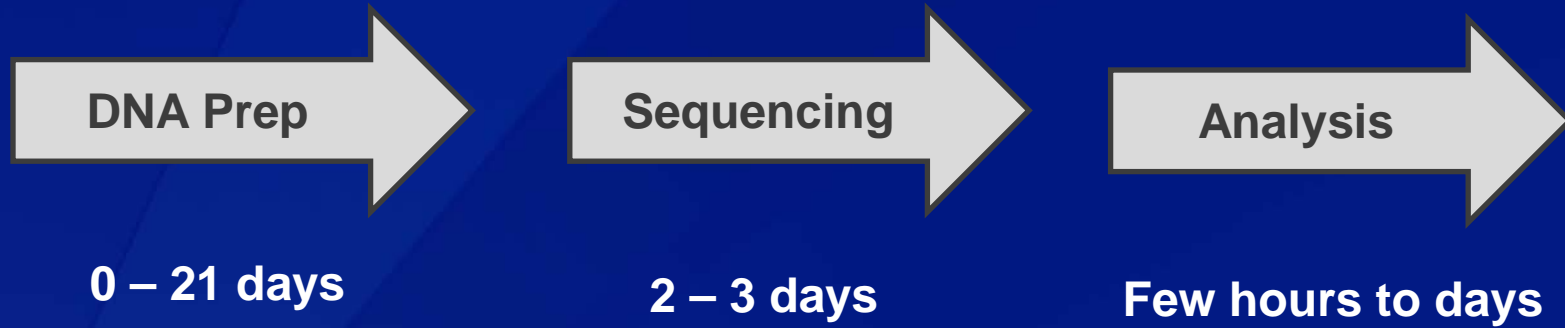


ACTAGA
ACTAGT
TCTACT

NJ tree visualization

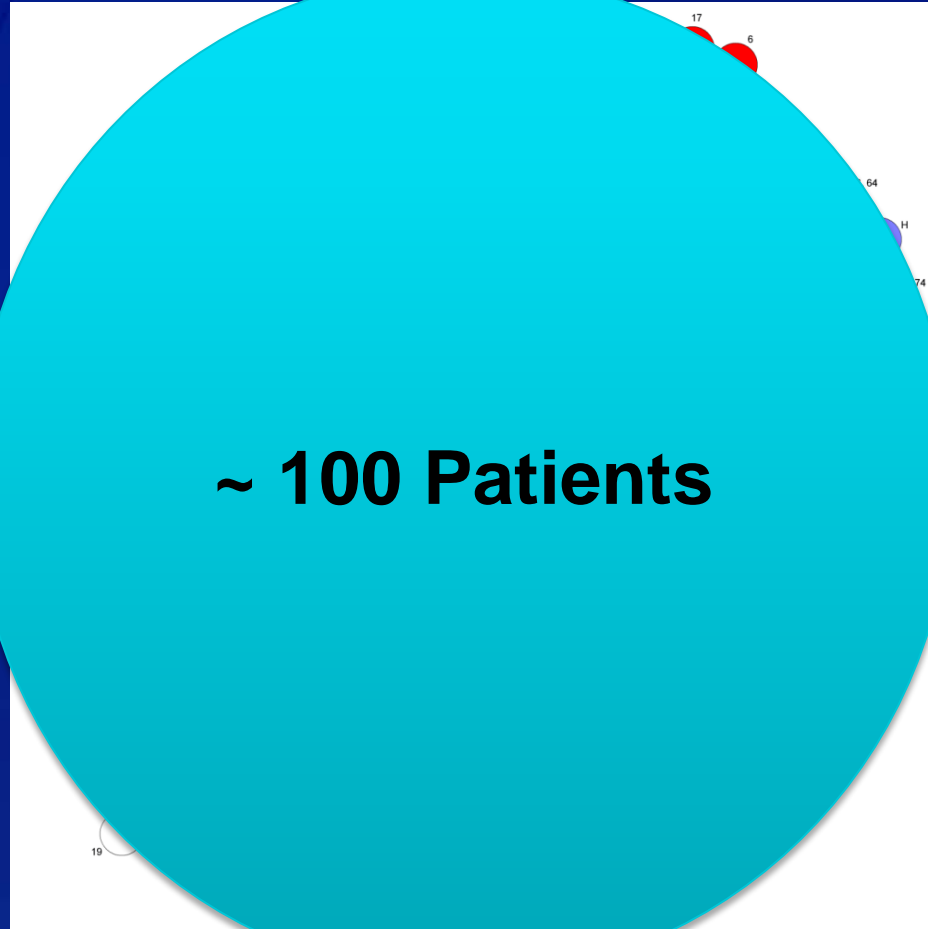


Turnaround Time



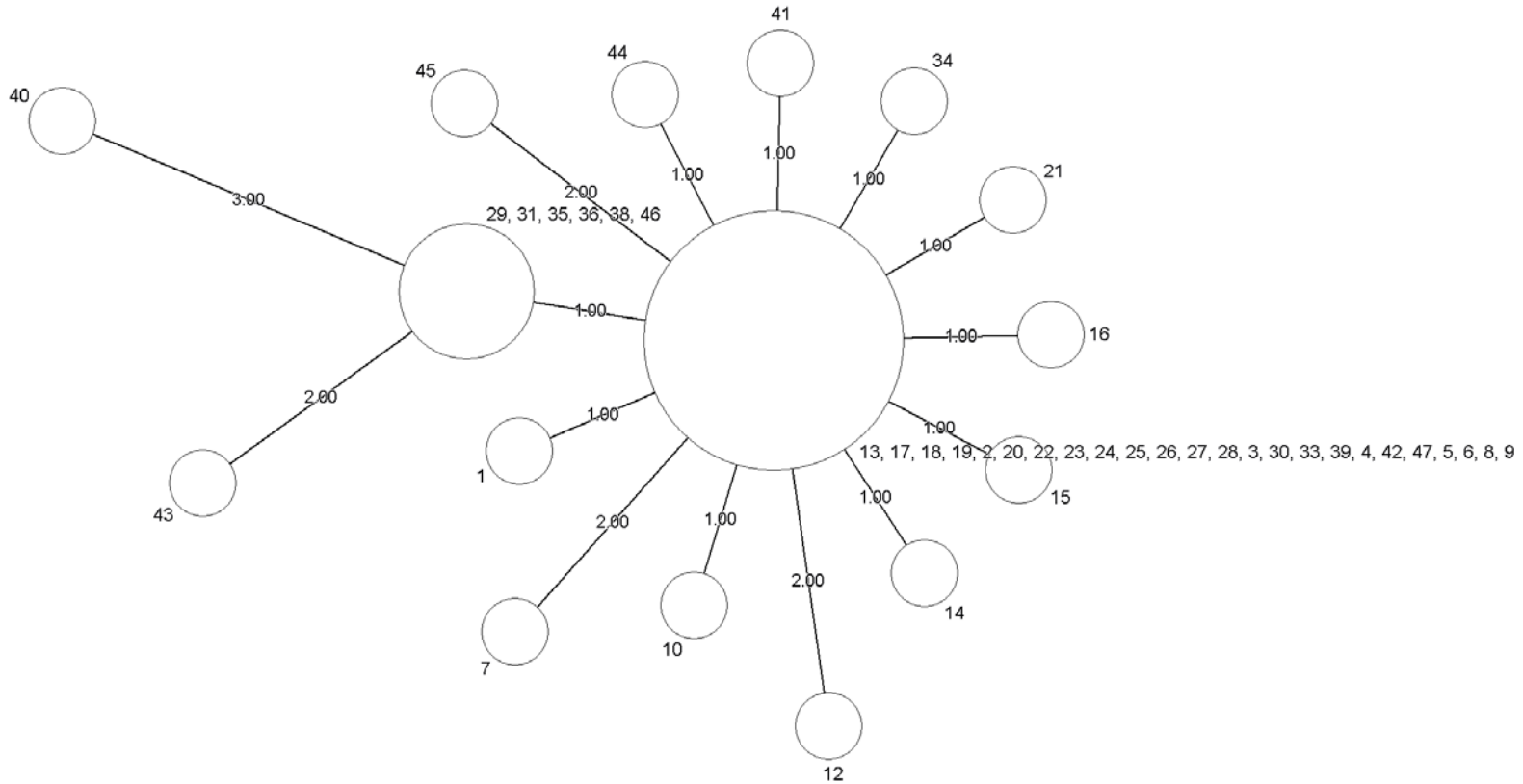
MOLECULAR EPIDEMIOLOGY

Cluster 1

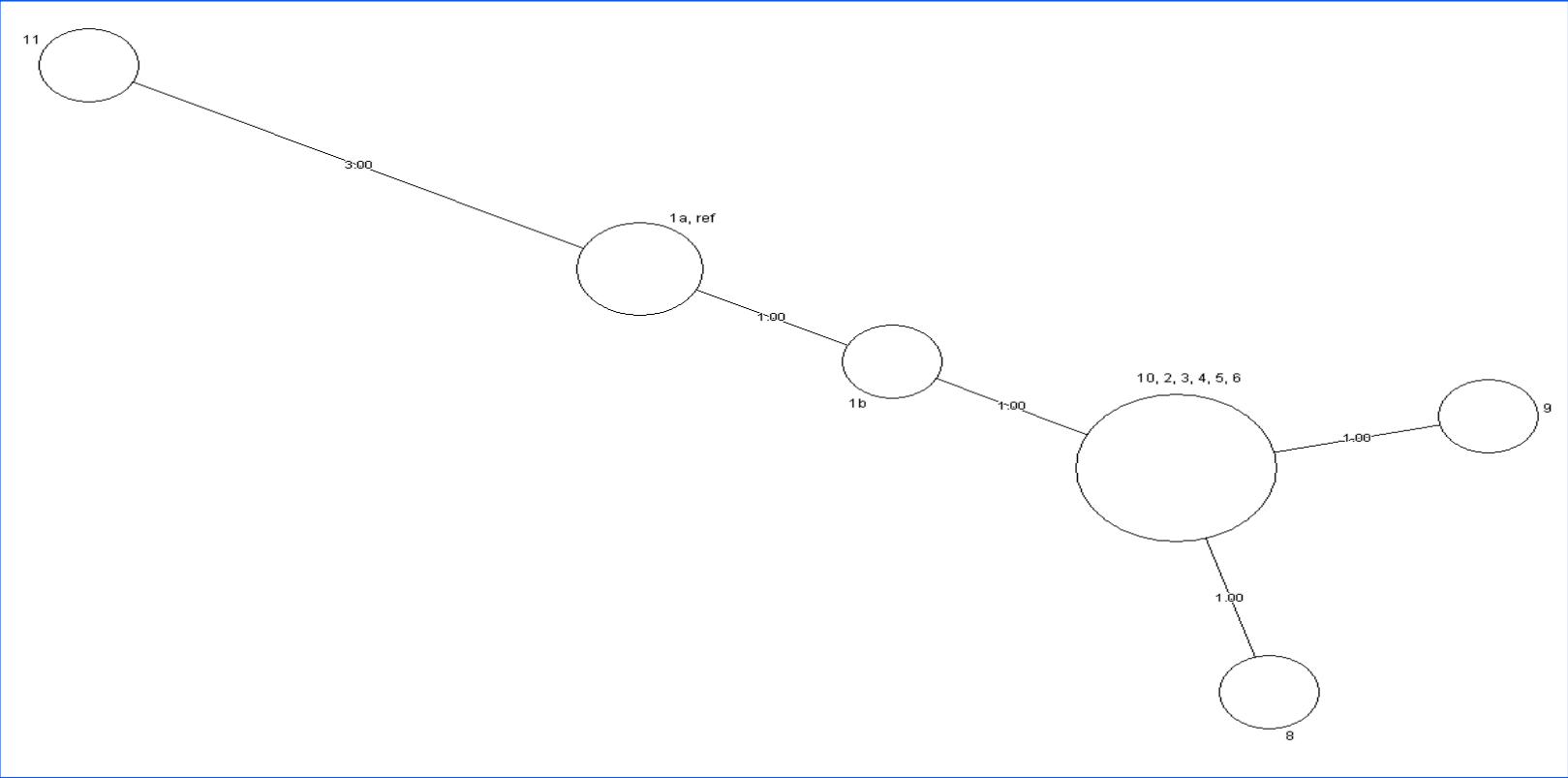


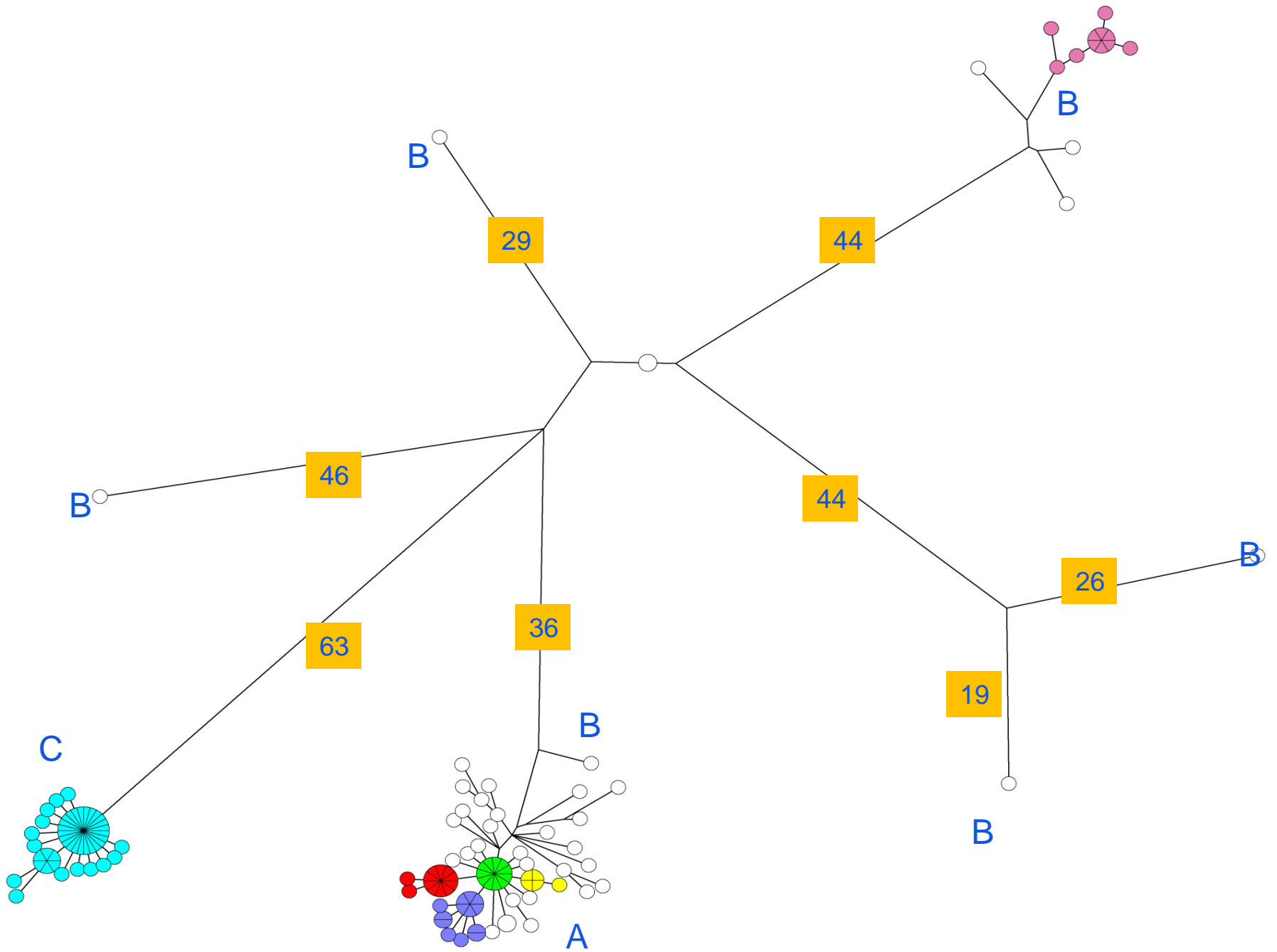
~ 100 Patients

Cluster 2



Cluster 3

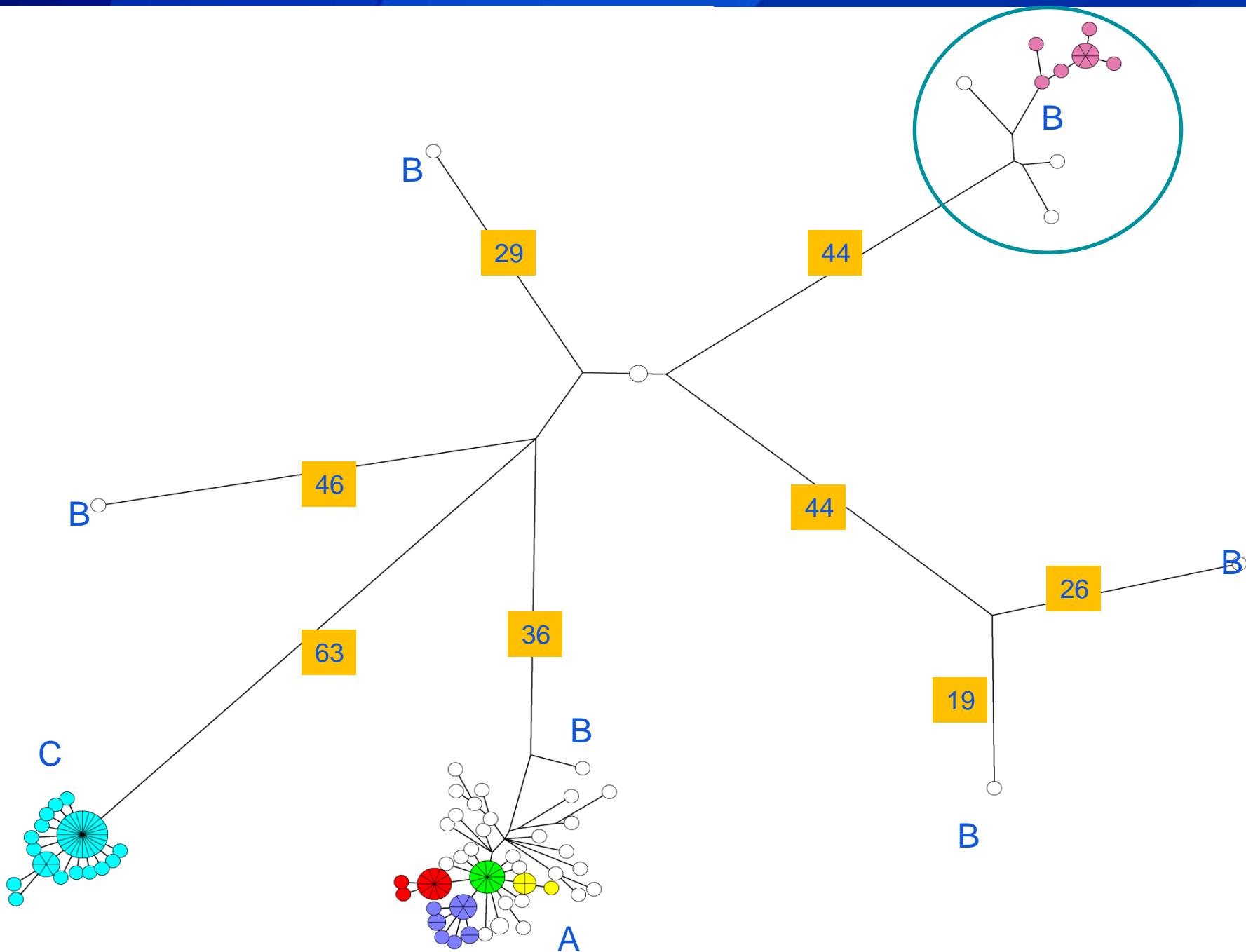




DRUG RESISTANCE

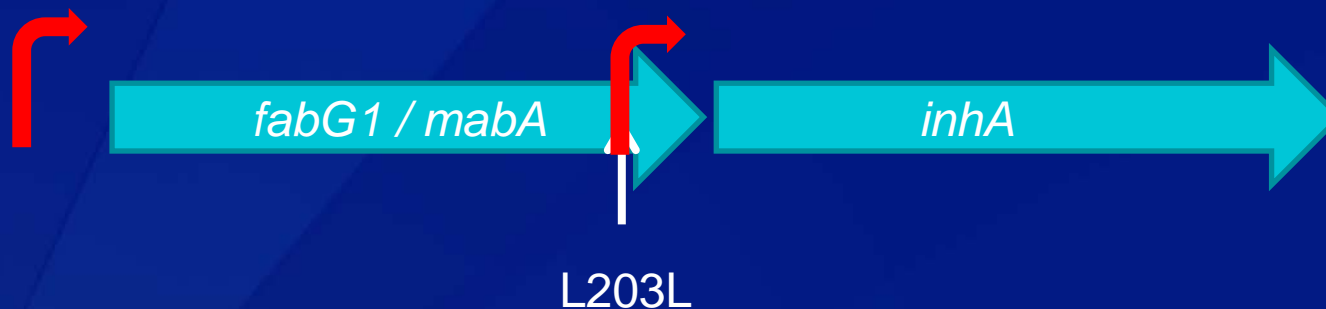
Applications of WGS for Drug Resistance

- ❑ Surveillance
- ❑ Clinical management
- ❑ Identify new mechanisms



C	G	C31998G	58 bp upstream Rv0029		C	T	C1663856 T	acn
C	T	C51403T	Rv0047c		G	A	G1674048 A	fabG1
C	T	C118832T	Rv0102		T	C	T1877958 C	pks7
C	T	C247984T	Rv0207c		C	G	C1888075 G	pks9
T	C	T362962C	PE_PGRS 5		T	C	T2087076 C	171 bp upstream of Rv1838c
C	G	C477188G	Rv0398c		C	T	C2372126 T	Rv2112c
C	A	C480678A	mmpL1		G	C	G2402463 C	Rv2142c
A	G	A649974 G	ubiE		T	C	T2614547 C	46 bp upstream Rv2339
C	T	C761147T	rpoB		G	A	G2751471 A	Rv2449c
G	C	G765719C	rpoC		G	A	G2958534 A	Rv2631
C	G	C799139G	Rv0698		G	A	G3126489 A	Rv2819c
G	A	G905686 A	Rv0811c		G	A	G3137406 A	echA16
C	A	C926861A	PE_PGRS 13		C	T	C3213150 T	lepB
C	G	C1023436 G	betP		A	C	A3377940 C	PPE46
T	G	T1093459 G	PE_PGRS 17		A	C	A3380380 C	PPE47
G	A	G1114491 A	Rv0997		A	G	A3416480 G	Rv3055
C	T	C1208858 T	Rv1084		C	T	C3455434 T	Rv3088
C	T	C1231660 T	Rv1104		G	C	G3608047 C	Rv3230c
A	C	A1246730 C	bpoB		C	T	C3764285 T	PPE56
C	T	C1266797 T	Rv1139c		G	A	G3765280 A	PPE56
C	T	C1309314 T	fdxC		C	A	C3777772 A	spoU
A	C	A1320356 C	papA3		A	G	A4026439 G	5 bp upstream Rv3585
G	A	G1353888 A	tagA		C	A	C4037284 A	PE_PGRS 59
G	A	G1421085 A	Rv1272c		T	A	T4072484 A	Rv3633
					G	C	G4084482 C	topA
					A	G	A4314271 G	bfrB

Identify New Mechanisms of Resistance

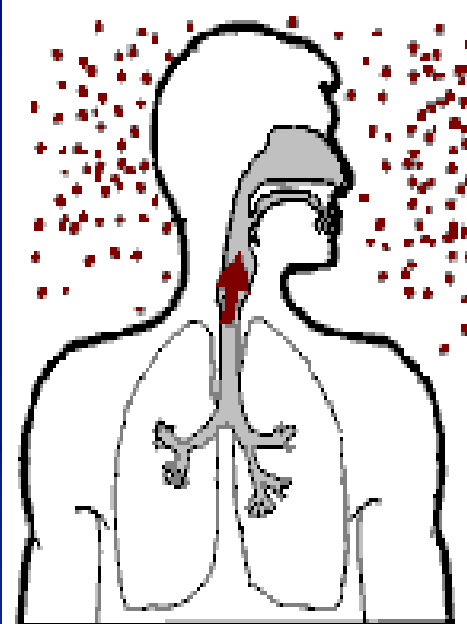


A silent mutation in *mabA* confers isoniazid resistance on *Mycobacterium tuberculosis*.

[Ando H](#), [Miyoshi-Akiyama T](#), [Watanabe S](#),
[Kirikae T](#).

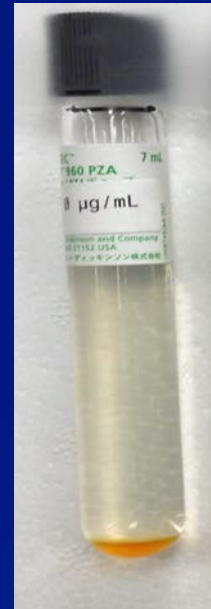
METAGENOMICS

Starting Material



Sputum

?



Dx Culture

✓



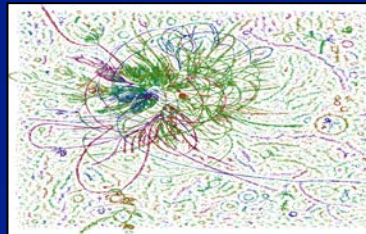
Subculture

✓

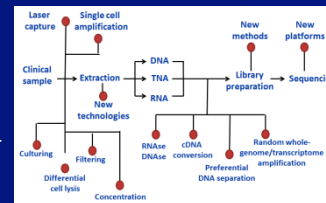
AMD Metagenomics Project Overview



Clinical specimen



Sample extraction/
processing



Clutter Mitigation



High throughput sequencing

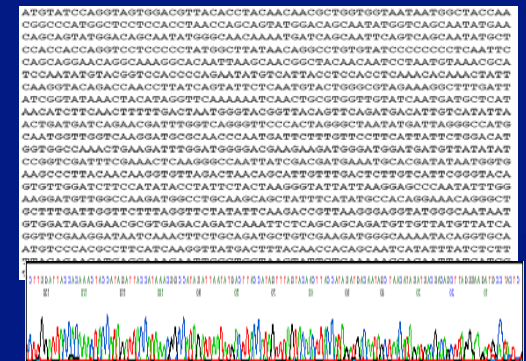
Identification &
characterization



Comparative
pathogen
sequence
database



Subtractive
host
sequence
database



RNA/DNA content sequence

Clinical Specimen Sets



Human whole blood (EDTA), 2 liters

- Acquired through normal channels



Human sputum, 2 liters

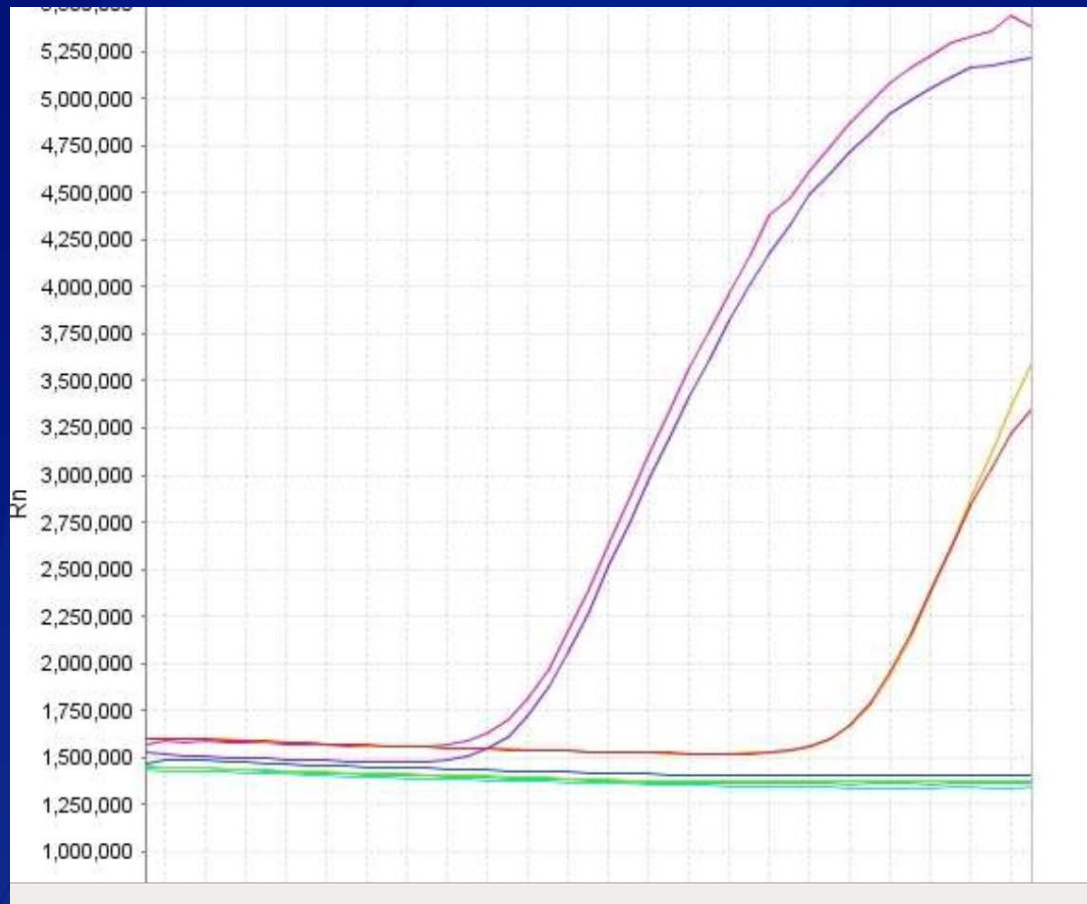
- Multiple donors
- Took over 6 months to accumulate



Human stool, 2 liters

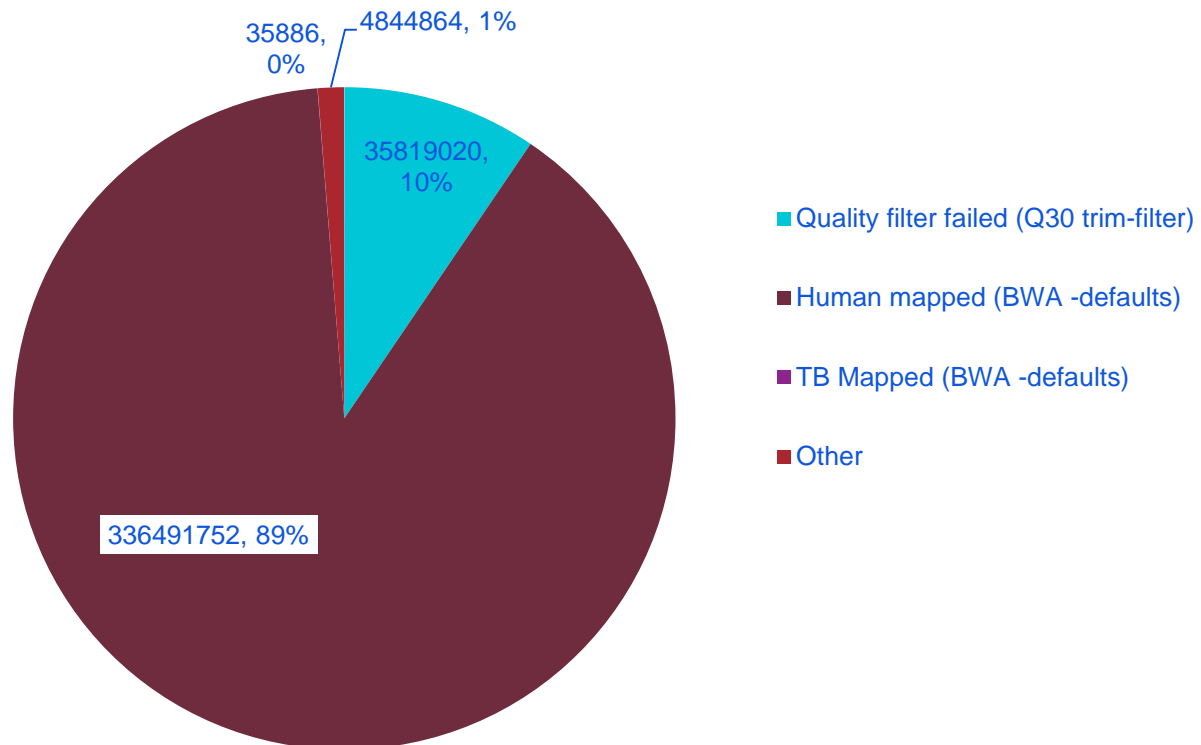
- Obtained from single donor
- “Genome diet”

Test Sputum for Mtb



Metagenomics

Sputum Background Sample Raw Read processing



Project Workflow

MTb gDNA + hu gDNA at different ratios

MTb gDNA + Sputum gDNA at different ratios

MTb:huDNA Ratio (%)

All mixtures @ 25 ng μL^{-1}

MTb:Sputum Ratio (%)



Transposon-Based Library Prep

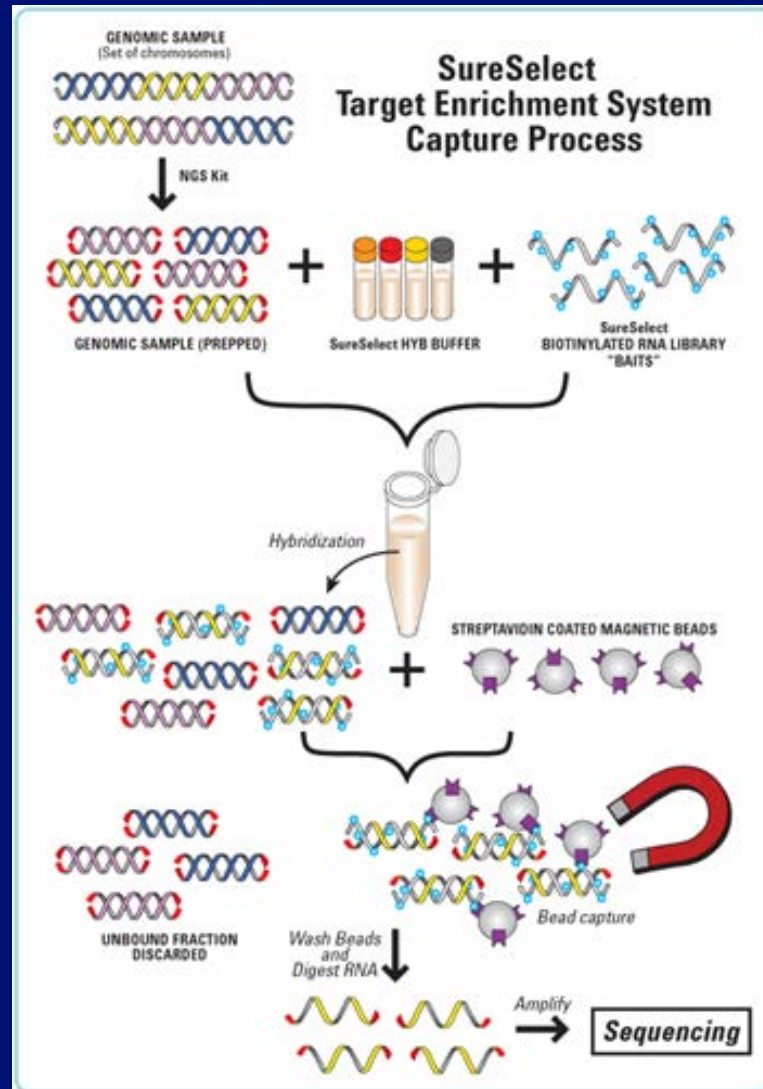


Sequence Capture

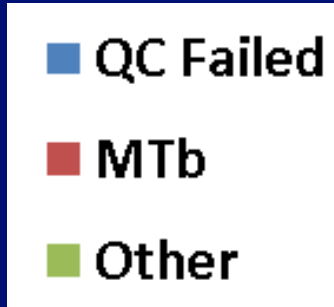


MiSeq Sequencing and Analysis

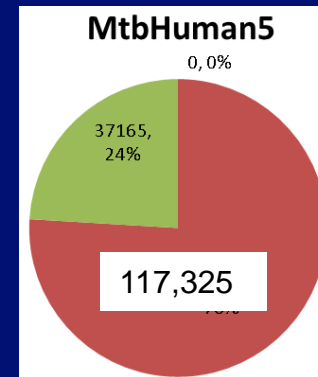
Agilent SureSelect



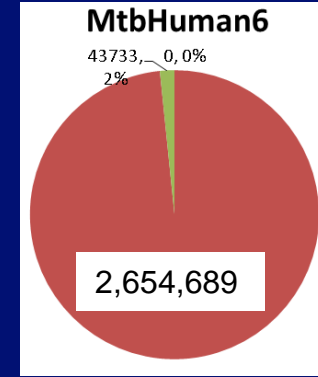
Mtb SureSelect Enrichment



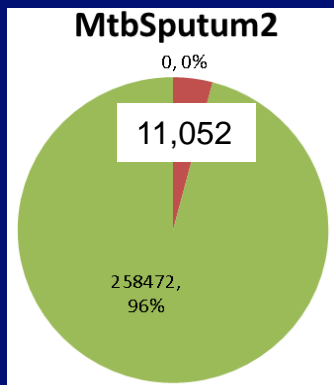
0.1%



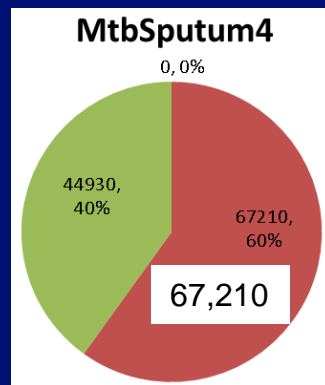
1 %
(82.5% , 4.5x)



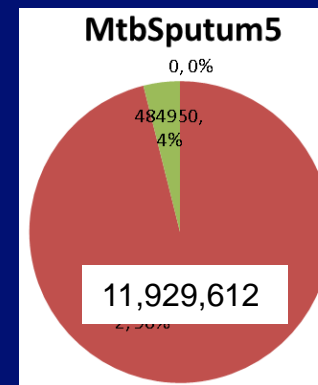
20 %
(99.4% , 77x)



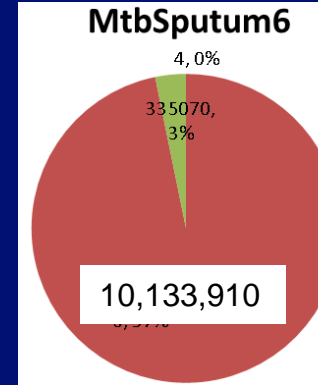
0.01 %



1 %
(72% , 3x)



10 %
(98.9% , 270x)



20 %
(98.9% , 251x)

Recent Publication

Accepted Manuscript Posted Online

JCM Accepted Manuscript Posted Online 13 May 2015
J. Clin. Microbiol. doi:10.1128/JCM.00486-15
Copyright © 2015, American Society for Microbiology. All Rights Reserved.

1 **Rapid Whole Genome Sequencing of *M. tuberculosis* directly from**
2 **clinical samples**

3 Amanda C. Brown^{1,2*}, Josephine M. Bryant^{3*}, Katja Einjer-Jensen⁴, Jolyon

4 Holdstock¹, Darren T Houniet¹, Jacqueline Z.M. Chan¹, Daniel P. Depledge³,

5 Vladyslav Nikolayevskyy⁵, Agnieszka Broda⁵, Madeline J. Stone⁶, Mette T.

6 Christiansen³, Rachel Williams³, Michael B. McAndrew¹, Helena Tutill³, Julianne

7 Brown³, Mark Melzer⁷, Caryn Rosmarin⁷, Timothy D. McHugh⁸, Robert J. Shorten^{8,9},

8 Francis Drobniowski⁵, Graham Speight¹, Judith Breuer³

9 Study Group: PATHSEEK consortium

Conclusions

- ❑ **Some infrastructure is required for NGS**
 - Laboratory work
 - Bioinformatics
 - IT and data management

- ❑ **Developed SOPs and optimized analysis**
 - Surveillance
 - Drug resistance

- ❑ **Work in progress**
 - Universal genotyping using WGS
 - Metagenomics
 - Fund at least two state laboratories for WGS

Acknowledgements

❑ **Laboratory Branch / Applied Research Team**

- Lauren Cowan
- Melisa Willby
- Paige Gupton
- Kartee Johnson

❑ **Core Facility**

- Mike Frace
- Mili Sheth
- Jamie Davis

❑ **AMD Metagenomics Team**

- Chris Hopkins
- Eishita Tyagi
- Scott Burns