University of Antwerp
I GHI I Global Health Institute

Next Generation Sequencing Technologies

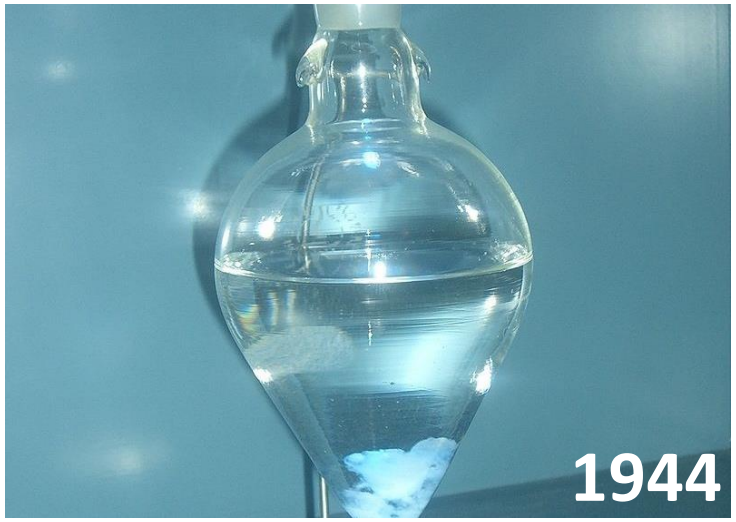
Anzaan Dippenaar

31 August 2024

TORCH
Tuberculosis Omics ResearCH

DNA sequencing

A BRIEF HISTORY



Friedrich Miescher:

- isolated various phosphate-rich chemicals from white blood cells, which he called nuclein

Avery–MacLeod–McCarty:

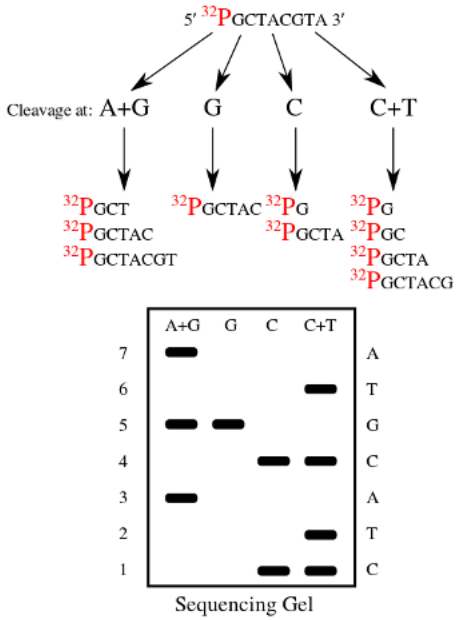
- DNA is the “genetic material”

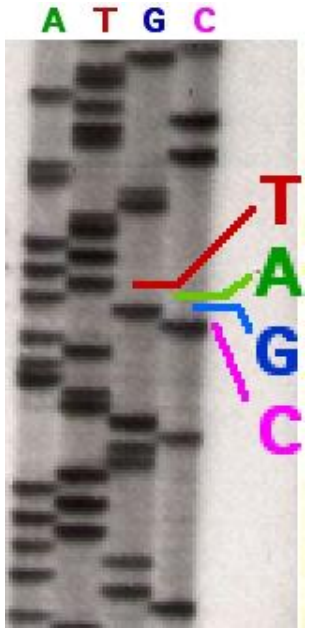
Watson & Crick:

- Double Helix Structure of DNA

History of DNA sequencing

1977

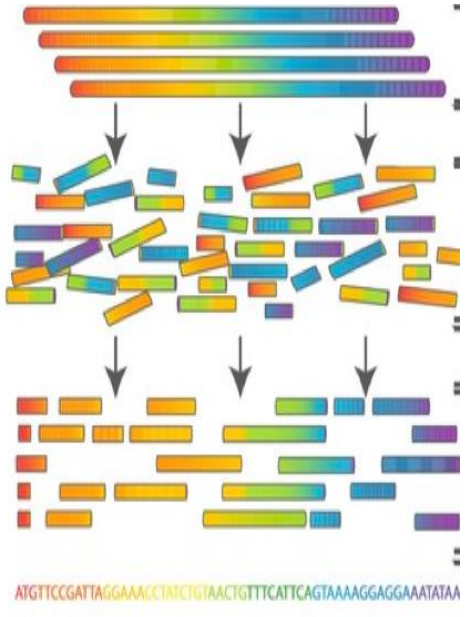




Maxam-Gilbert:
DNA sequencing by chemical degradation

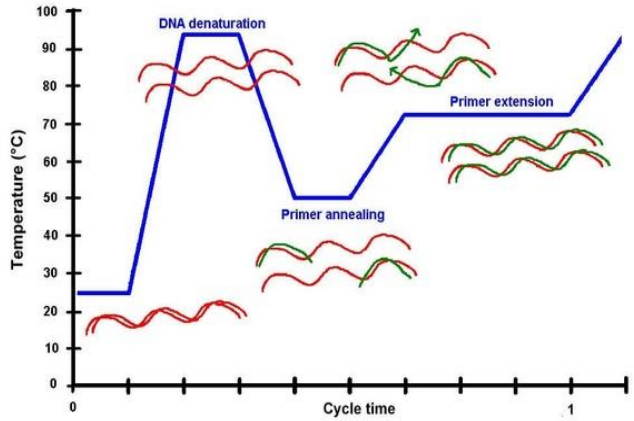
Frederick Sanger *et al*:
DNA sequencing with chain-terminating inhibitors

1981



Messing *et al*:
shotgun sequencing of cauliflower mosaic virus

1983



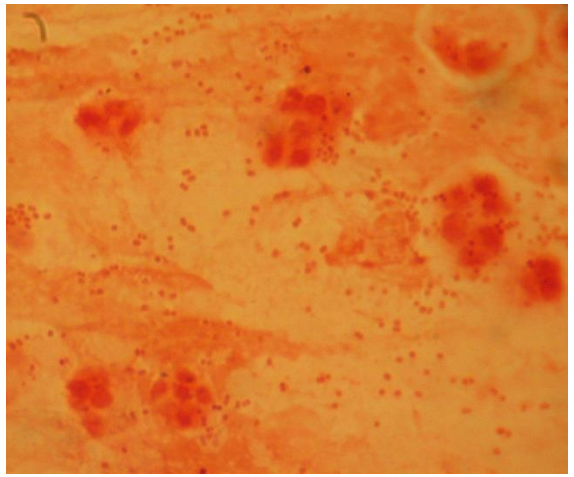
Kary Mullis:
Polymerase Chain reaction (PCR)

The history of sequencing – 1st gen

1987

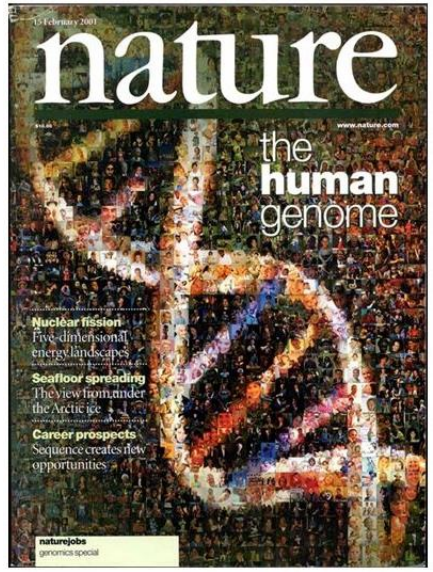
First automated sequencer: ABI Prism 370

1995



Craig Venter and Hamilton Smith:
first complete genome of a free-living organism: *H. influenza*

2001



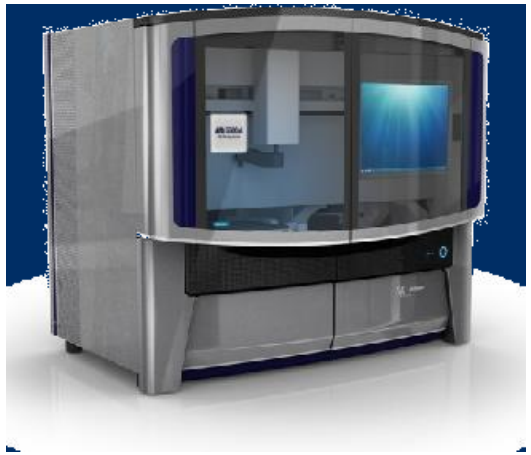
Draft of the human genome published

The history of sequencing: NGS

2005 → 2006 → 2007 → 2011

Roche 454 pyrosequencing

Solexa (Illumina) Next Generation Sequencer



Applied Biosystems Next Generation Sequencer : SOLiD

Ion Torrent: Personal Genome Machine (PGM)

The history of sequencing: 3rd gen

2009

Helicos single molecule sequencer : Helicos Genetic Analyser System

2011

Pacific Biosciences single molecule sequencer : PacBio RS Systems

2012

Oxford Nanopore Technologies – The GridION and MinION

DNA sequencing

THE PRINCIPLES

Next-Generation Sequencing platforms

- **Different approaches are used:**
  1) Sequencing by synthesis - uses DNA polymerase (e.g. Illumina, Pyrosequencing)
  2) Sequencing by ligation - uses DNA ligase (e.g. SOLiD)
  3) Direct molecule sequencing (e.g. ONT)

Choosing the most appropriate platform

Platform	Primary advantages	Primary disadvantages
Sanger Sequencing	Low cost for small studies	High cost for large amounts of data
Illumina	Moderate run cost, low error rate, various read length configurations, standardised data format	High instrument costs, relatively short reads, large amounts of data produced*
PacBio	Longest available read length (single molecule real time sequencing), short run time	Expensive equipment and per-sample costs, high(er) error rate
Ion Torrent	Low-cost instruments and disposable chips, simple machine	High error rate, fewer reads at higher cost per Mb relative to e.g. Illumina MiSeq, smaller user community, data format
Oxford Nanopore Technologies	MinIon – small, portable instrument, low capital (hardware) cost, long reads, error rate does not increase along read length	High error rate, limited library preparation approaches, high amount of input DNA, analysis not standardized

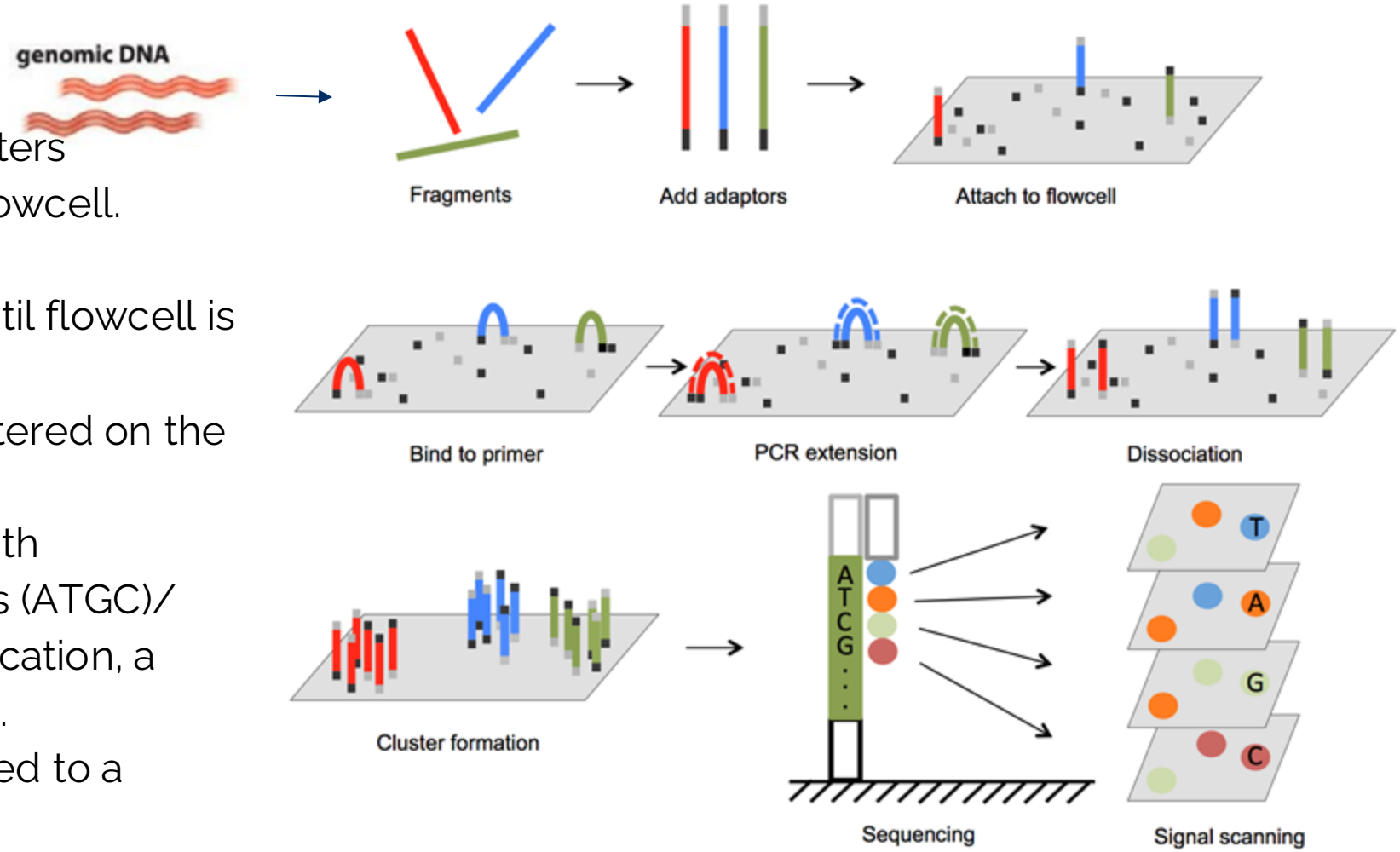
Understanding sequencing – PCR

University of Antwerp
GHI | Global Health Institute

ONLINE VIDEO: PCR SONG - BIORAD

Sequencing by synthesis

1. Fragment genomic DNA
2. Add 2x different synthetic adapters complimentary to that on the flowcell.
3. Hybridize to solid flowcell.
4. Perform bridge amplification until flowcell is covered with DNA strands.
5. Target strands will now be clustered on the flowcell.
6. Add adapter specific primers with fluorescently labeled DNA basis (ATGC)/
7. As a base is used during amplification, a signal is detected and captured.
8. These contigs are then compared to a reference genome.



https://www.youtube.com/watch?v=CZeN-IgjYCo
https://www.illumina.com/science/technology/next-generation-sequencing/beginners/ngs-workflow.html#:~:text=The%20next%2Dgeneration%20sequencing%20workflow,to%20plan%20your%20NGS%20workflow.

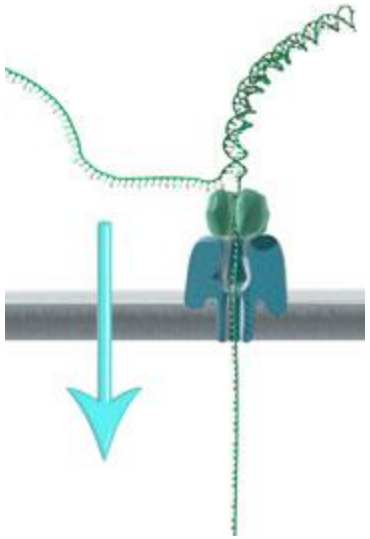
Sequencing by synthesis

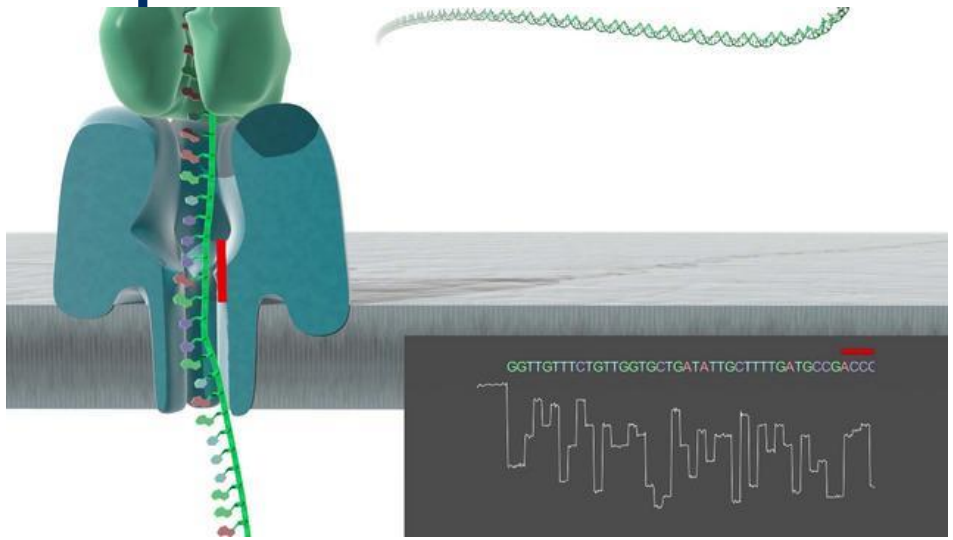
https://www.youtube.com/watch?v=CZeN-IgjYCo
https://www.illumina.com/science/technology/next-generation-sequencing/beginners/ngs-workflow.html#:~:text=The%20next%2Dgeneration%20sequencing%20workflow,to%20plan%20your%20NGS%20workflow.

Nanopore DNA Sequencing

Oxford Nanopore's technology uses nanopores embedded in a lipid membrane.

As a DNA molecule moves through the nanopore, minute changes in the electrical current are recorded, and later interpreted.

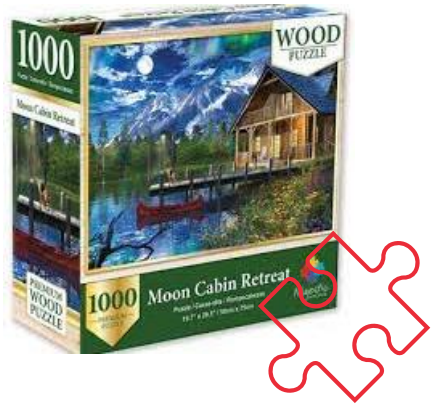




Nanopore DNA Sequencing

ONLINE VIDEO: ONT

NGS analysis approaches



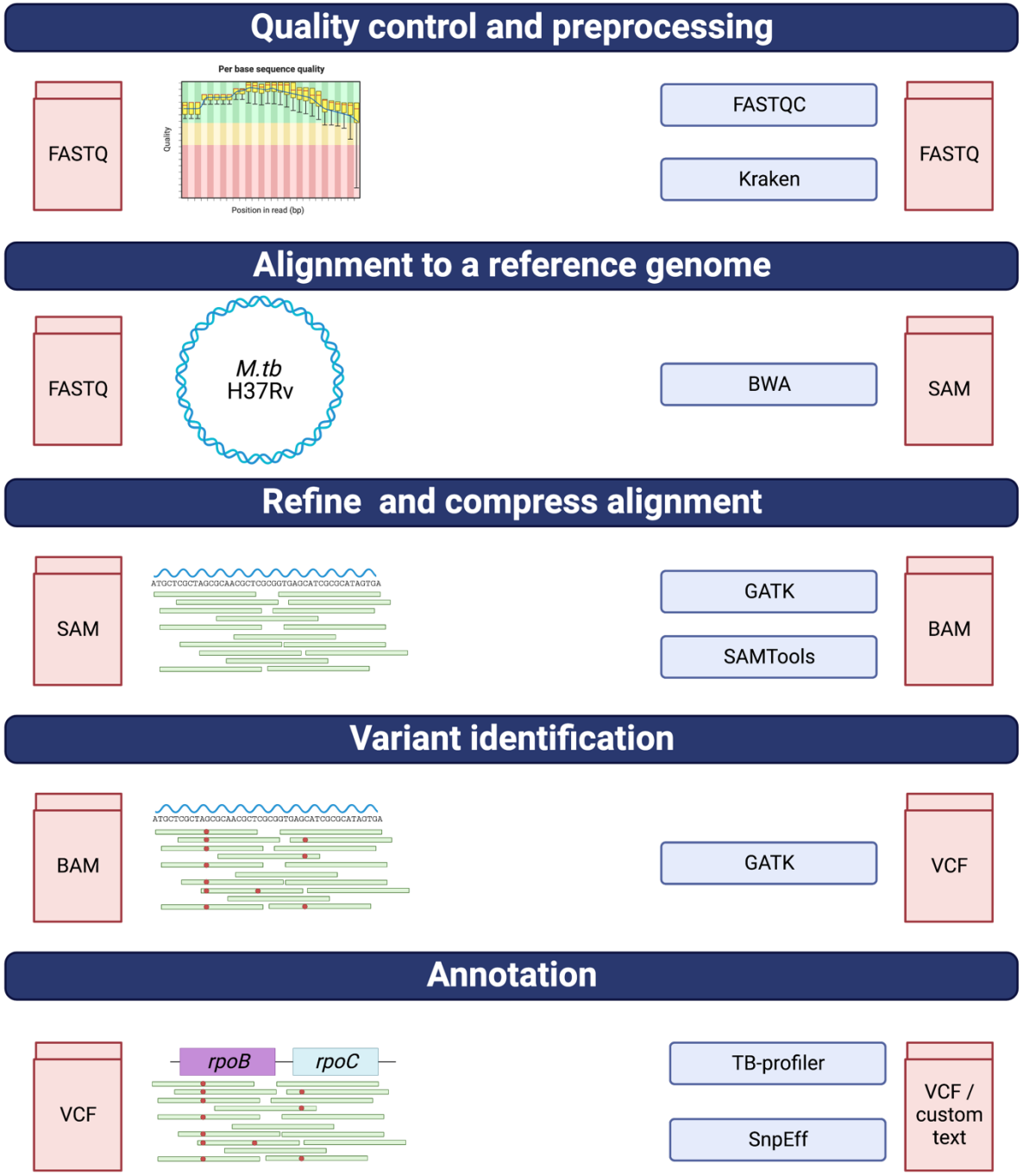
Reference mapping approach

- Known genome sequence of closely related organism
- Analysis is less resource-intensive
- Limited to known genomic regions

De novo assembly

- Limited knowledge of genome sequence or closely related organism
- Analysis is resource-intensive
- Describe novel genomic structures
- Species characterisation

WGS ANALYSIS



•10.1099/mgen.0.000689

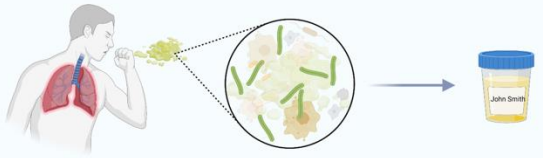
Next generation sequencing

of *M. tuberculosis*

FROM SAMPLE TO SEQUENCE

Microbiology

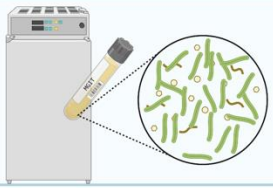
Clinical specimen



Liquefaction

Decontamination

Mtb Enrichment

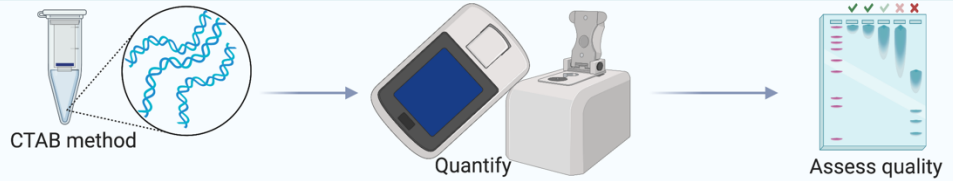


Inactivation

Molecular biology

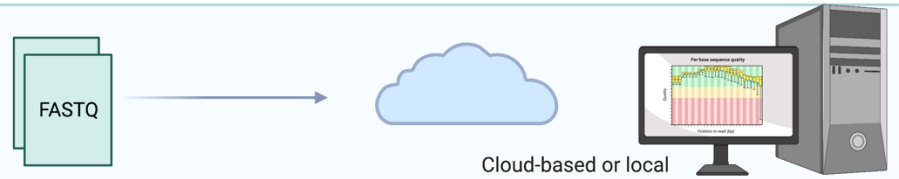
Mtb lysis

DNA extraction and quality control



Library preparation and sequencing

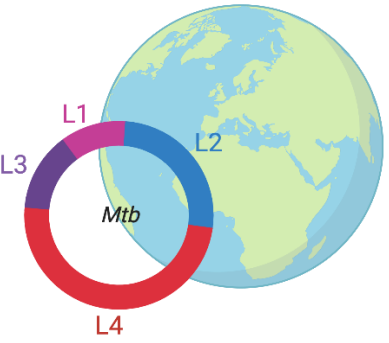
Data analysis



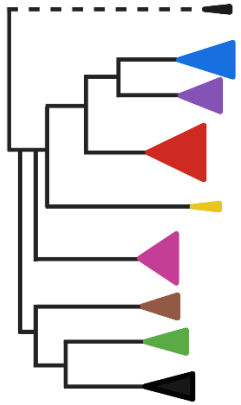
Bioinformatics

NGS APPLICATIONS FOR TB RESEARCH AND CLINICAL CARE

Surveillance

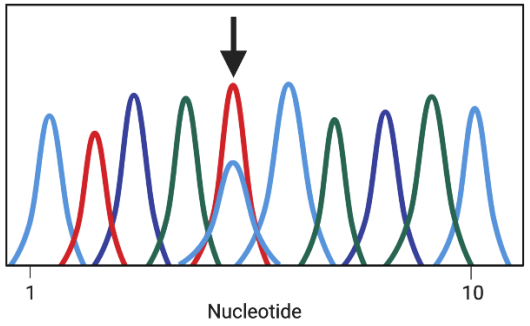


Epidemiology

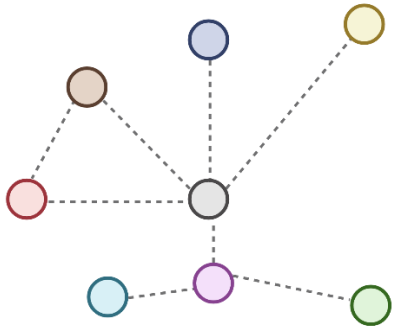


Drug resistance

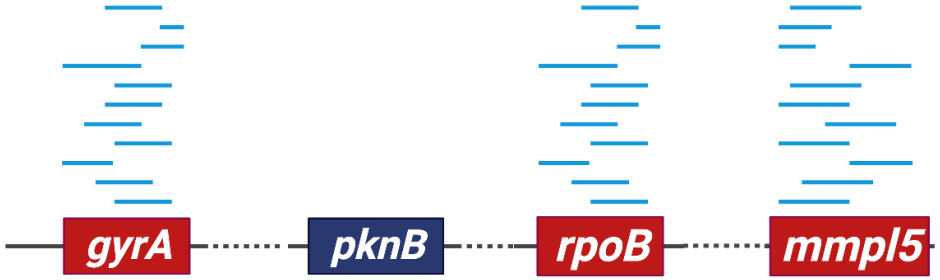
Heteroresistance

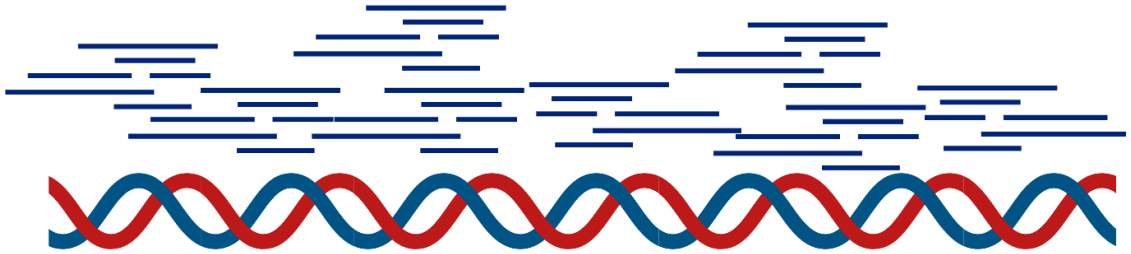


Transmission

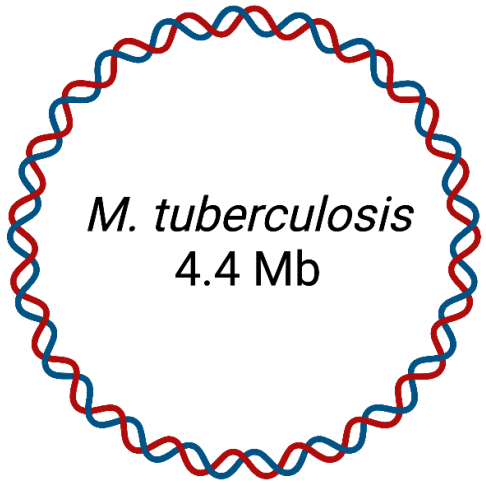


TARGETED vs. WHOLE GENOME SEQUENCING





Isoniazid
Rifampicin
Bedaquiline



Logistics and needs for TDS vs standard WGS vs advanced WGS

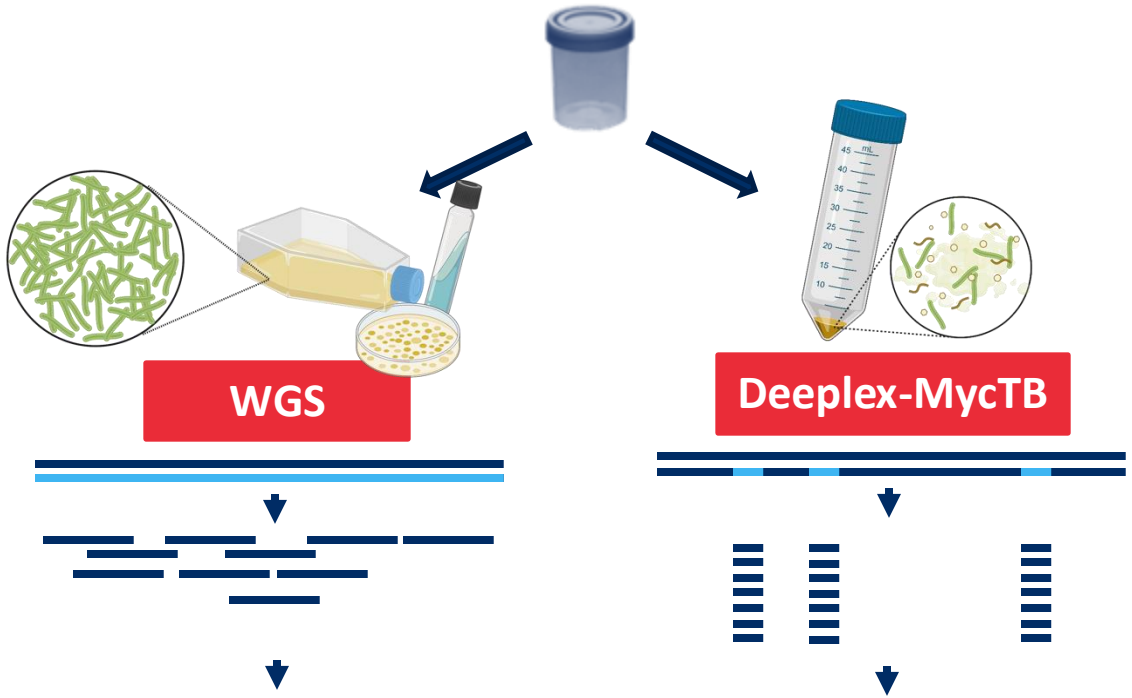
	Targeted Sequencing (Deeplex-MycTB)	Standard WGS	Advanced WGS
Application	Culture-free prediction of resistance to anti-TB drugs	Culture-based prediction of resistance to anti-TB drugs	Culture-based prediction of resistance to anti-TB drugs, transmission dynamics
Input material	DNA extracted directly from specimens (if bacterial load high) or cultured isolates	High quality purified DNA extracted from cultured *Mtb* isolates	High quality purified DNA extracted from cultured *Mtb* isolates
Staff requirements	Sequencing technician	Sequencing technician	Technician and Bioinformatician
Depth of coverage	> 1000 X (up to 100k X)	>10x	>30x (depending on application)
TAT (assuming in-house sequencer)	2 days	3-40 days culture 1 day DNA extraction 2-3 days sequencing 1 day analysis (Need for culture increases TAT)	3-40 days culture 1 day DNA extraction 2-3 days sequencing 2-day analysis (standard) 4+ days analysis (additional) (Need for culture increases TAT)
Data analysis	User friendly cloud-based system	User-friendly graphical interface	Bioinformatics analysis tools. (Graphical interface, command line)

NGS - Applications in *M. tuberculosis*

tNGS vs WGS

PROJECT OVERVIEW: deepMTB

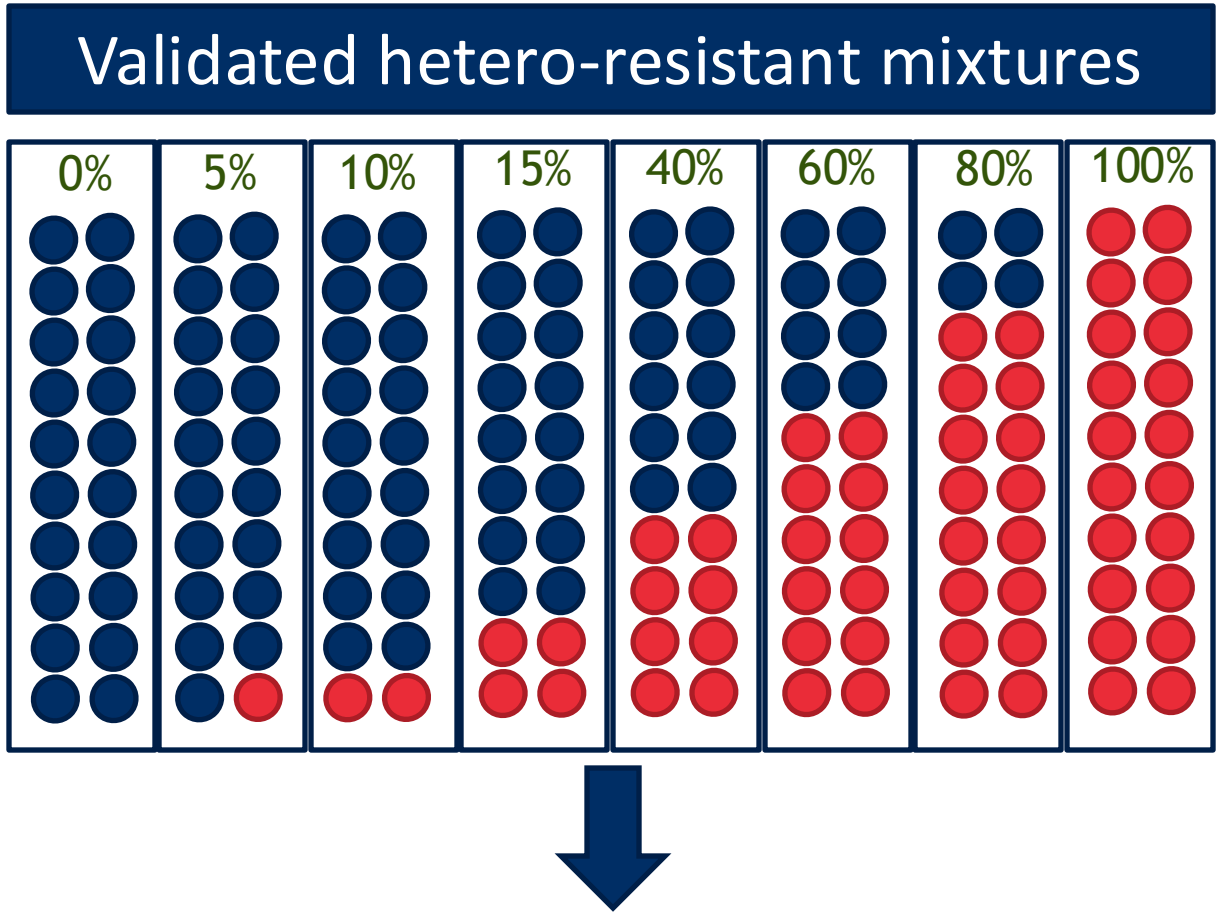
- **Impact of (Hetero)resistance on treatment success of DR-TB**



Comparison of
Hetero-resistance detected by different approaches
+
Quantify prevalence culture bias

TARGETED & WHOLE GENOME SEQUENCING

PROJECT OVERVIEW: deepMTB



- **Lab-made resistant and susceptible *M. tuberculosis* mixtures**
- **Available from BE-based bacterial biobank**
- **Developers of diagnostic tests**

University of Antwerp
GHI | Global Health Institute

BCCM
BELGIAN CO-ORDINATED COLLECTIONS OF MICRO-ORGANISMS

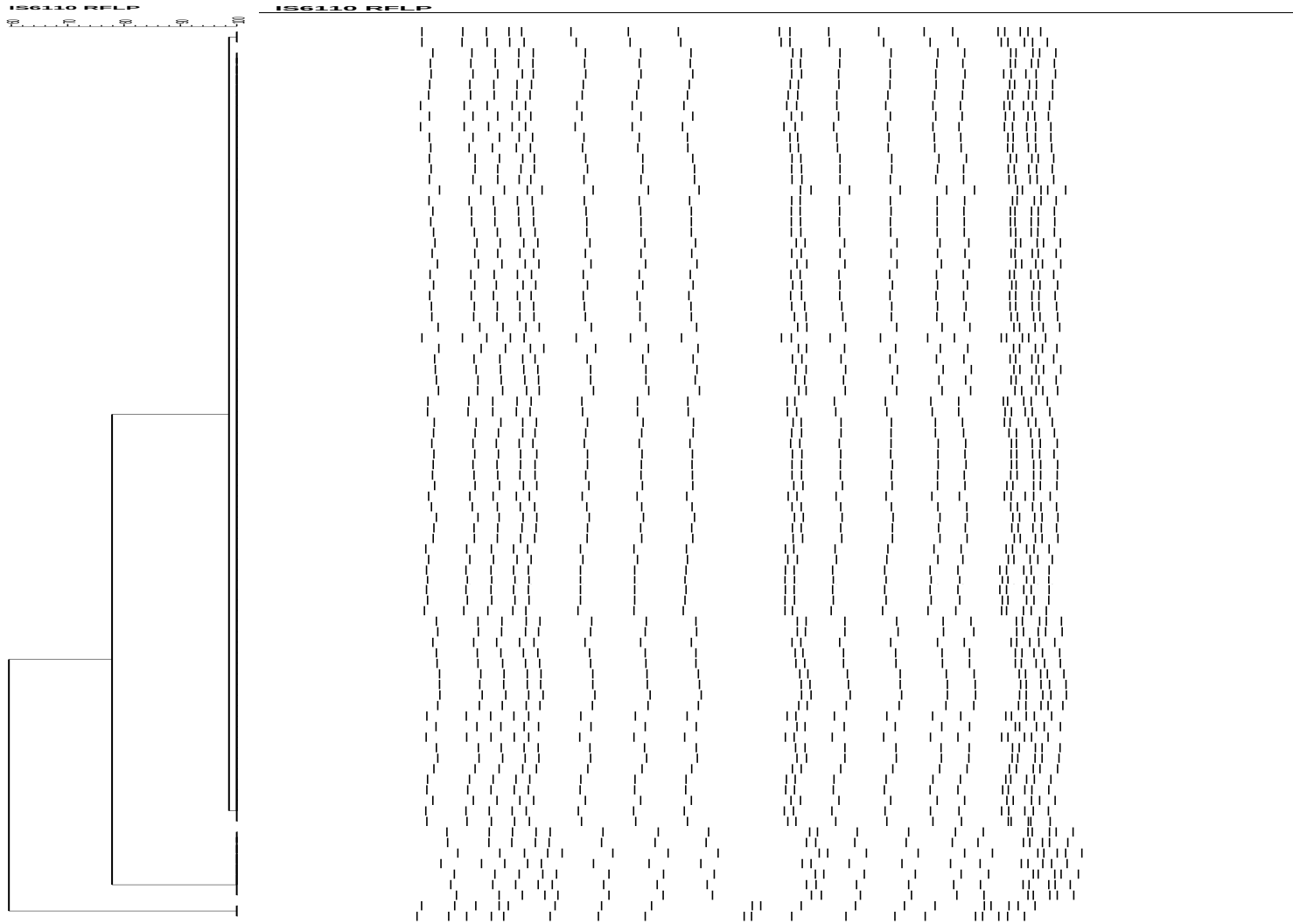
NGS - Applications in *M. tuberculosis*

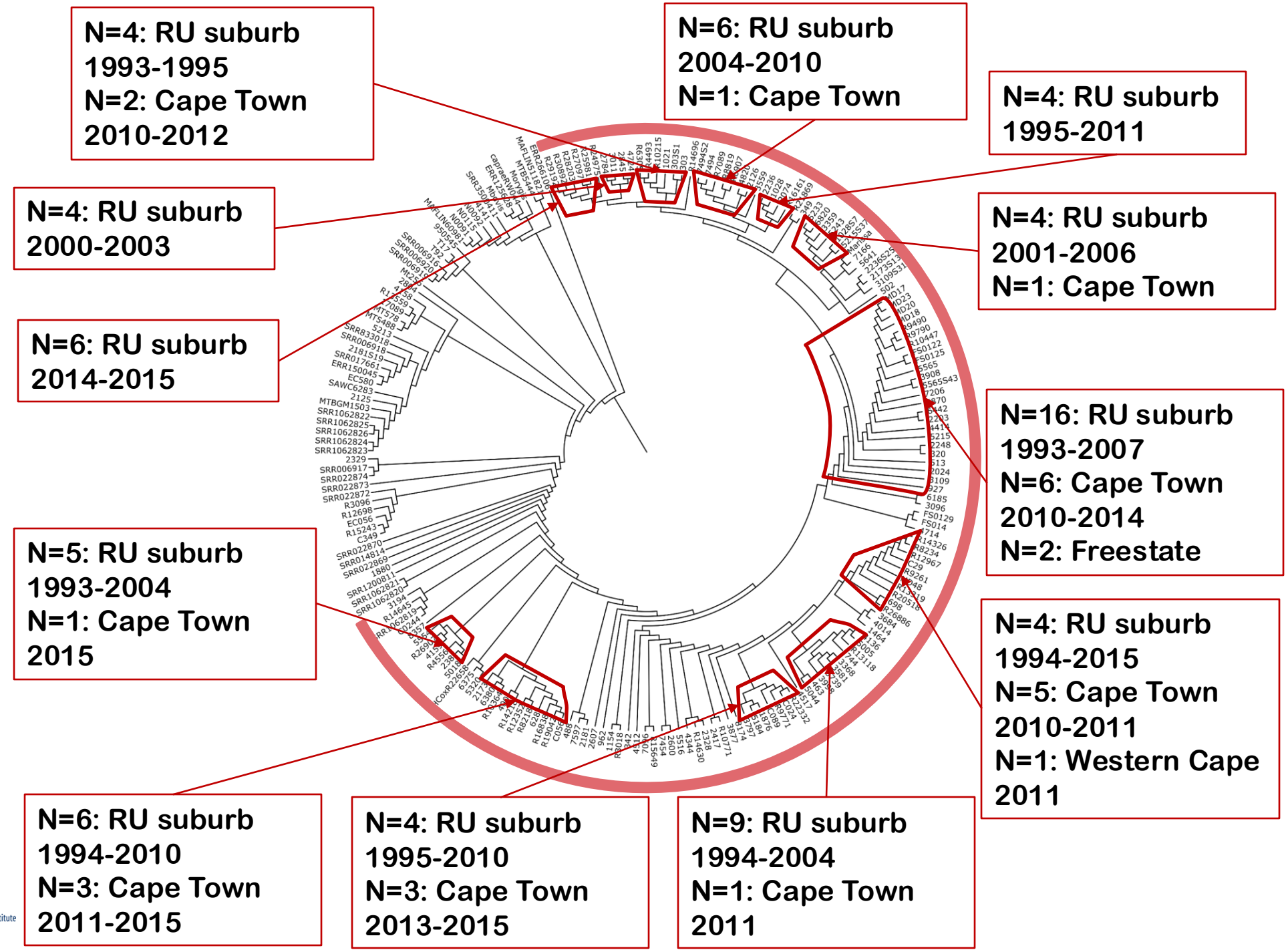
NGS has higher resolution compared to classic genotyping techniques

WGS: MDR Beijing outbreak in CAPE TOWN

- 1993-2015
- 84 patients (127 isolates)
- Identical Beijing IS*6110*-RFLP
- Resistant to 4 frontline drugs

- WGS revealed characteristic: *pncA* 100ATC; 160GCA mutations



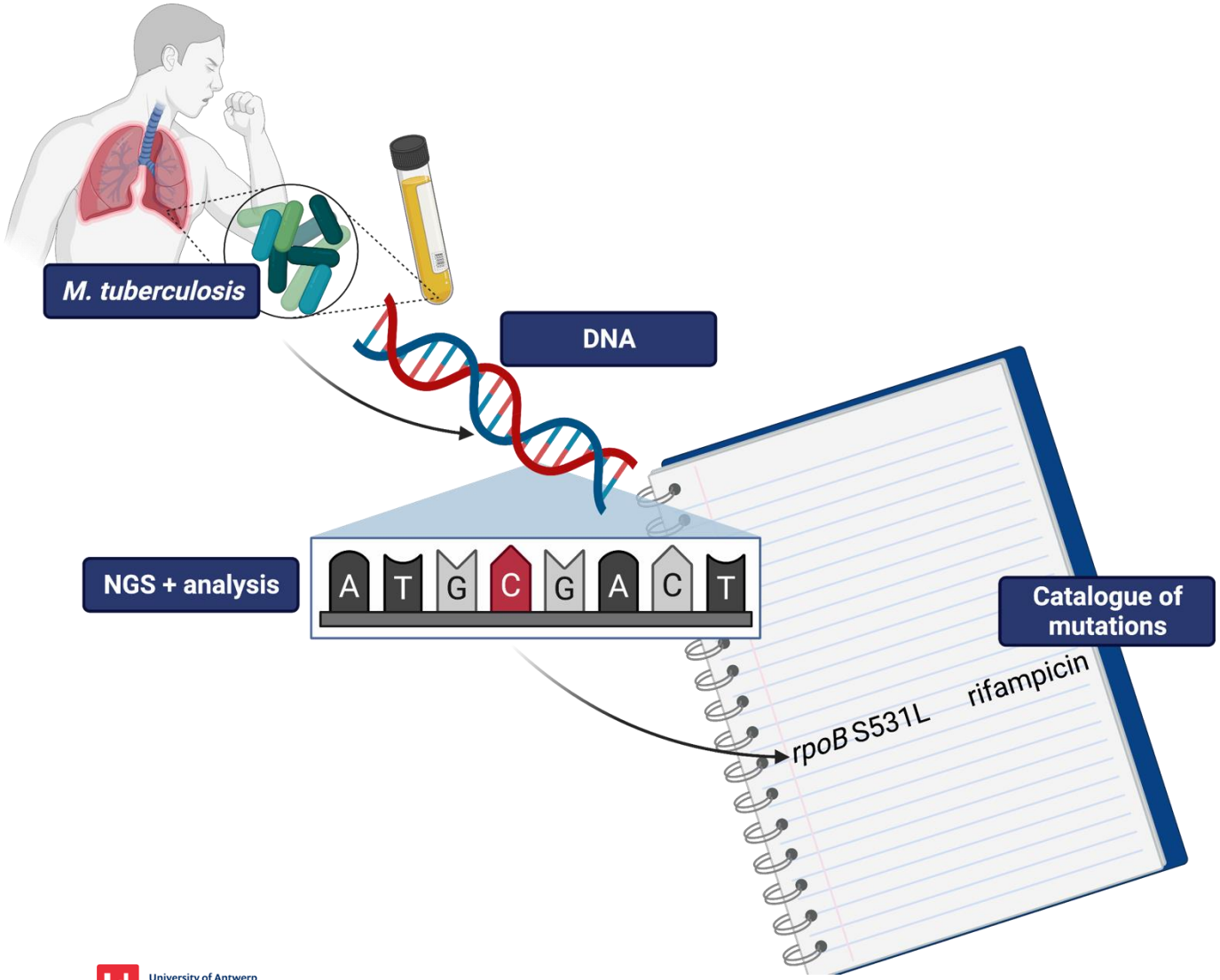


University of Antwerp
GHI | Global Health Institute

WGS of clinical primary *Mtb* liquid cultures for

PRECISION MEDICINE

PROJECT OVERVIEW: SMARTT



- **Pragmatic randomised control trial**
  - WGS and Standard of care
- **WGS-based comprehensive drug-resistance profile to guide DR-TB treatment**
  - Individualised treatment / personalised medicine
- **Low-middle income country**
  - Freestate Province, South Africa
- **Primary outcome**
  - Bacteriological response to treatment

•10.1186/s13063-022-06793-w

WHOLE GENOME SEQUENCING

PROJECT OVERVIEW: SMARTT

MYCOBACTERIUM TUBERCULOSIS GENOME SEQUENCING REPORT

SMARTT

Patient Name		Phone	
Birth date		Sample	LP250
Gender	MALE	Weight (Kg)	55.0
Facility		Facility contact	
Requested by		Sequenced from	MGIT
Sample type	Sputum	Sample date	XXX
Sample number		Report date	2022-11-01

Drug Resistance Profile

This *Mycobacterium tuberculosis* strain is predicted to be:

Resistant to: Ethambutol, Ethionamide, Prothionamide, Rifampicin, Isoniazid, and Isoniazid High Dose

Susceptible to: Levofloxacin, Linezolid, Imipenem, Meropenem, Moxifloxacin, Moxifloxacin High Dose, Bedaquiline, Amikacin, Clofazimine, Terizidone, Streptomycin, Para Aminosalicylic Acid, Rifabutin, Delamanid, and Pyrazinamide

Clinical Information

Hb (g/dL)	10.7	Platelets (*10^9/L)	600
Neutrophils (*10^9/L)	NA	GFR (mL/min)	60
ALT (IU/L)	13	QTc (ms)	391
Hearing loss	No	Painful peripheral neuropathy	No
Psychosis	No	Poorly controlled epilepsy	No
Visual problems	No	Pregnant	No

- **Clinical decision support system**
  - AI trained algorithm
  - Takes DR-profile and clinical information into account
  - Knowledge of drug properties

Treatment recommender:

Proposed Individual Treatment Regimen

Drug	Dosage (mg)
Levofloxacin	1000 mg
Linezolid	600 mg
Bedaquiline	200 mg three times per week
Pyrazinamide	1500 mg

•10.1186/s12911-022-01790-0

Summary

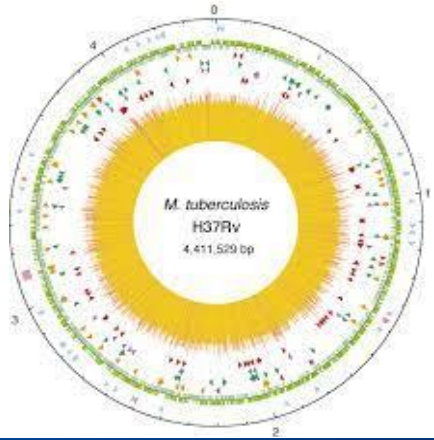
- **NGS important role to accurately track transmission, outbreaks etc.**
- **Importance in WGS for drug resistance in TB:**
  - All known drug resistance conferring variants can be identified
    - Enables individualised treatment
  - Detection of heteroresistance
  - Identify novel mechanisms of resistance

- **Various applications of NGS**
  - Which clinical question are you seeking to address?

Tuberculosis practicals – Monday 2 Sep

- **Full day!**
- **Please prepare – lab manual**
- **Before Thursday:**
  - Create an account on Galaxy: https://usegalaxy.org/
  - Remember your login details…

DEEPLEX



COMPLEXITY

TORCH
Tuberculosis Omics ResearCH

Universiteit Antwerpen

Stellenbosch
UNIVERSITY
IYUNIVESITHI
UNIVERSITEIT

TB Genomics
research group

torch-consortium.com

anzaan.dippenaar@uantwerpen.be

MOLECULAR ASSAYS FOR TUBERCULOSIS: increased levels of complexity

INFORMATION	*Mtb*/no *Mtb* RIF	*Mtb*/no *Mtb* RIF INH	*Mtb*/no *Mtb* 15 drugs Genotype	*Mtb*/no *Mtb* all drugs Genotype Transmission
TARGETS	MTB-specific 1 drug-target	MTB-specific 3 drug-targets	MTB-specific 18 drug-targets	4.4Mb All drug-targets
SAMPLE TYPE	Raw specimen	Treated specimen	Treated Specimen ***	Cultured isolate
MOLECULAR ASSAY	**Simple NAATs**	**Moderate complexity NAATs**	**tNGS**	**WGS**