Public Health Ontario | Santé publique Ontario

To view an archived recording of this presentation please click the following link:

https://youtu.be/3XE2aDxfHyE

Please scroll down this file to view a copy of the slides from the session.

Disclaimer

This document was created by its author and/or external organization. It has been published on the Public Health Ontario (PHO) website for public use as outlined in our Website Terms of Use. PHO is not the owner of this content. Any application or use of the information in this document is the responsibility of the user. PHO assumes no liability resulting from any such application or use.

Progress and Challenges in Implementing a Tuberculosis Genomic Epidemiology Program

Jennifer L. Guthrie, MSc, PhD
Assistant Professor, Western University
Dept of Microbiology & Immunology
Dept of Epidemiology & Biostatistics

Adjunct Scientist
Public Health Ontario

PHO Microbiology Rounds
November 17, 2022

Schulich MEDICINE & DENTISTRY

Western

Disclosures

I have no financial or other conflicts of interest

Learning Objectives

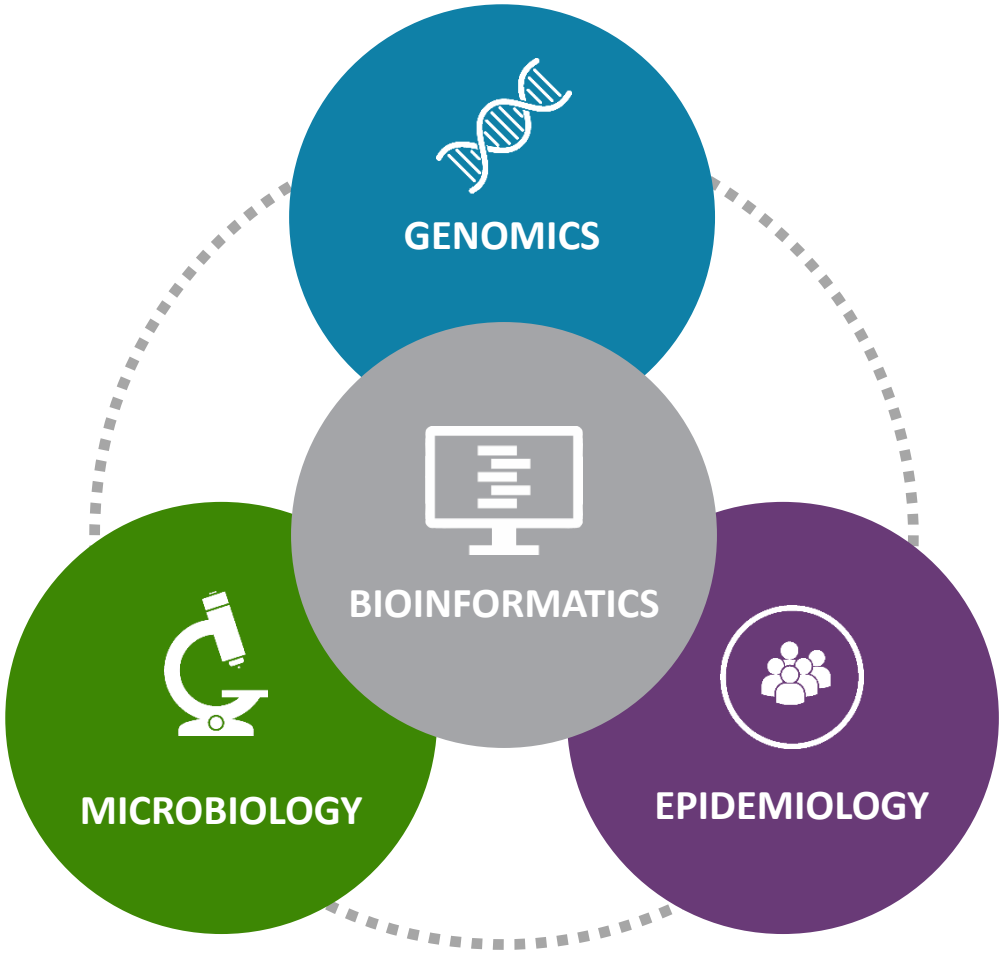
- Explain genomic epidemiology in plain language
- Describe how genomic epidemiology is applied to public health using a real-world example
- Discuss the challenges in implementing a tuberculosis genomic epidemiology program

epidemiology (noun, \ ˌe-pə-ˌdē-mē-ˈä-lə-jē \

: the study and analysis of the distribution, patterns and determinants of health and disease conditions in a defined population

GENOMIC EPIDEMIOLOGY

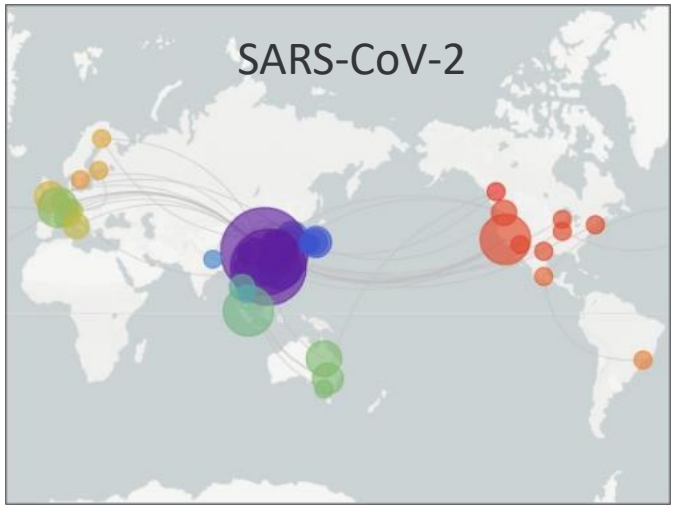
Use of **pathogen genomic data** to determine the distribution and spread of an infectious disease in a specified population and its application to public health decision making



GENOMIC EPIDEMIOLOGY

Applications

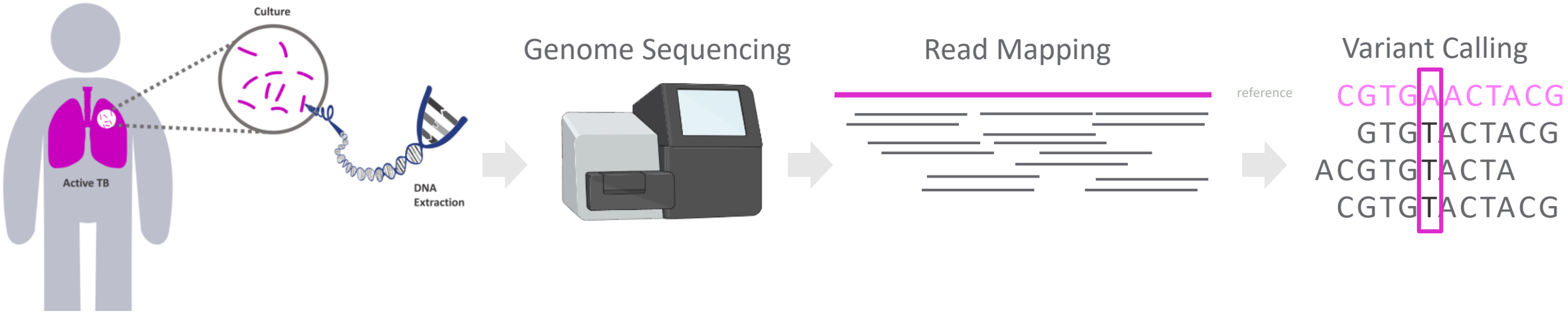
- **Population-level surveillance**
  - e.g., monitor variants to inform vaccine development
- **Outbreak detection**
  - e.g., link hospital cases with no known connections
- **Outbreak investigation**
  - e.g., identify source of food-borne outbreak
- **Infection control**
  - e.g., pinpoint contaminated hospital equipment
- **Track spread of a pathogen, locally to globally**
  - e.g., SARS-CoV-2



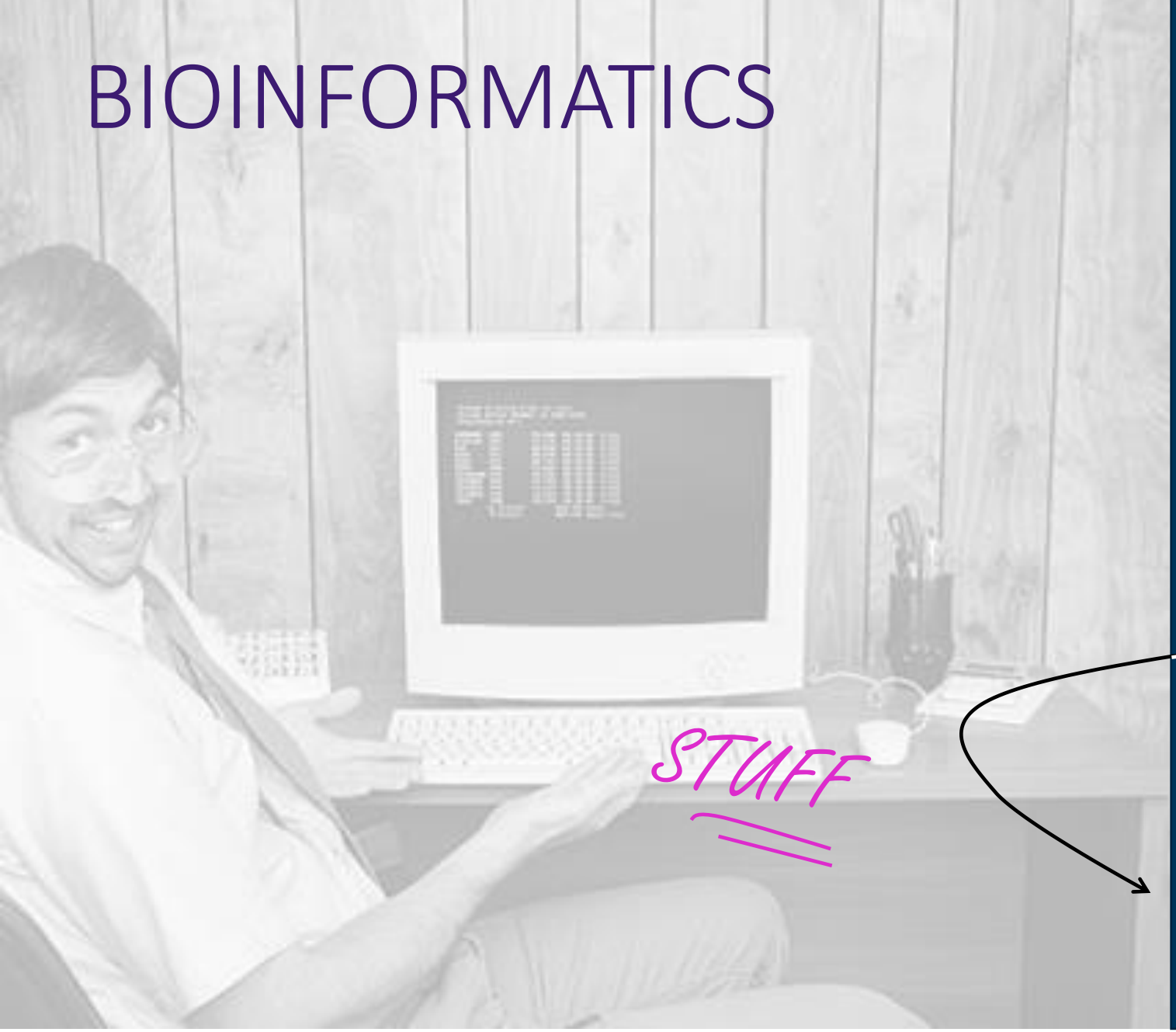
WHOLE GENOME SEQUENCING

WHOLE GENOME SEQUENCING

Methods



BIOINFORMATICS



DNA extraction

Library prep

Genome sequencing

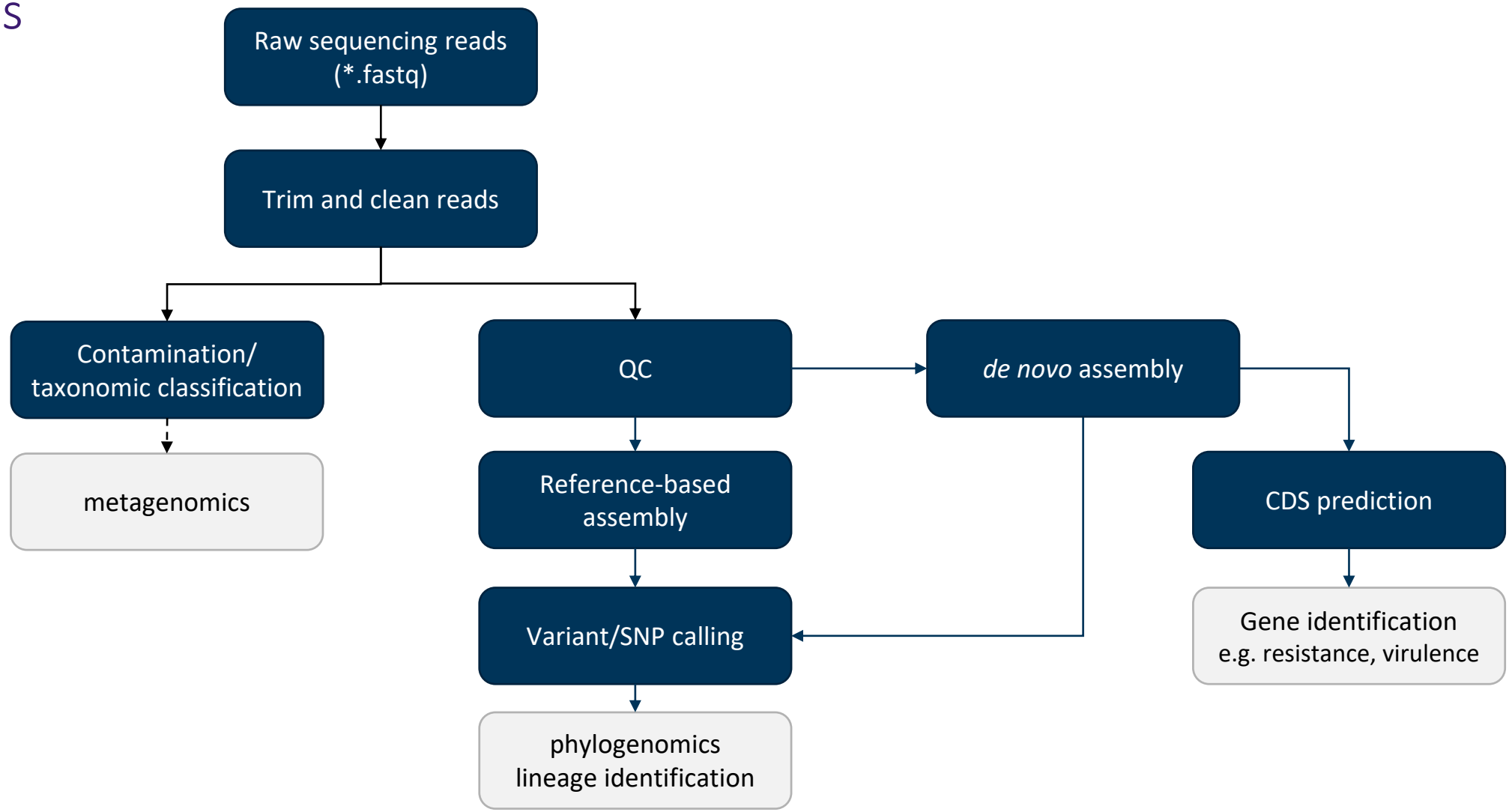
Raw data

SNP identification

Phylogenomics

BIOINFORMATICS

Pipelines



BIOINFORMATICS

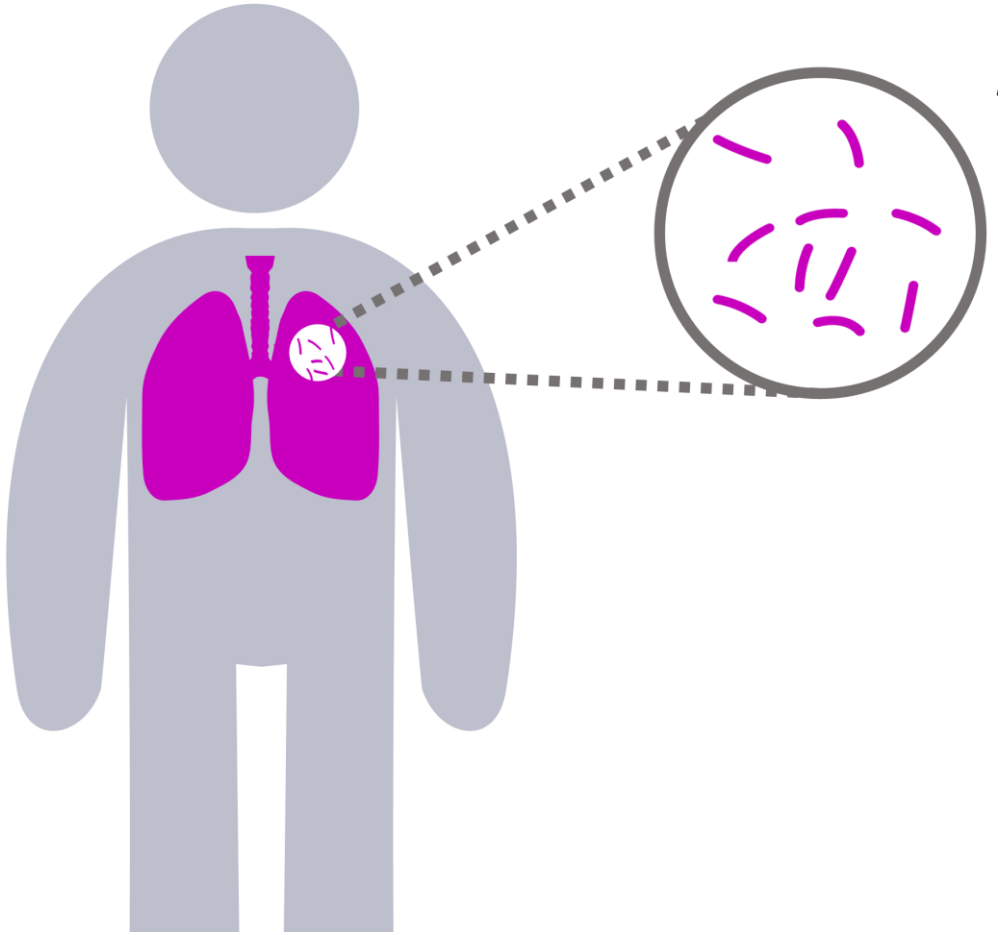
Pipelines

```
$ perl ~/dev/git/cortex/scripts/analyse_variants/make_branch_fasta.pl --callfile cortex_called_annotated.fa

$ bwa index CP000717.fa
[bwa_index] Pack FASTA... 0.05 sec
[bwa_index] Construct BWT for the packed sequence...
[bwa_index] 4.30 seconds elapse.
[bwa_index] Update BWT... 0.06 sec
[bwa_index] Pack forward-only FASTA... 0.82 sec
[bwa_index] Construct SA from BWT and Occ... 1.48 sec
[main] Version: 0.7.5a-r405 [main] CMD: bwa index CP000717.fa
[main] Real time: 9.992 sec; CPU: 6.721 sec bwa mem CP000717.fa cortex_called_annotated.fa.branches.fasta >
cortex_called_annotated.fa.branches.fasta.map_to_F11.sam
[M::main_mem] read 1504 sequences (96567 bp)...
[main] Version: 0.7.5a-r405 [main] CMD: bwa mem CP000717.fa cortex_called_annotated.fa.branches.fasta
[main] Real time: 2.134 sec
$
```

TUBERCULOSIS

Mycobacterium tuberculosis



- Respiratory infection
- Airborne transmission
- 5% – 10% develop active TB

ortality

1.6 million deaths annually

anada

Low-incidence (4.7/100,000)

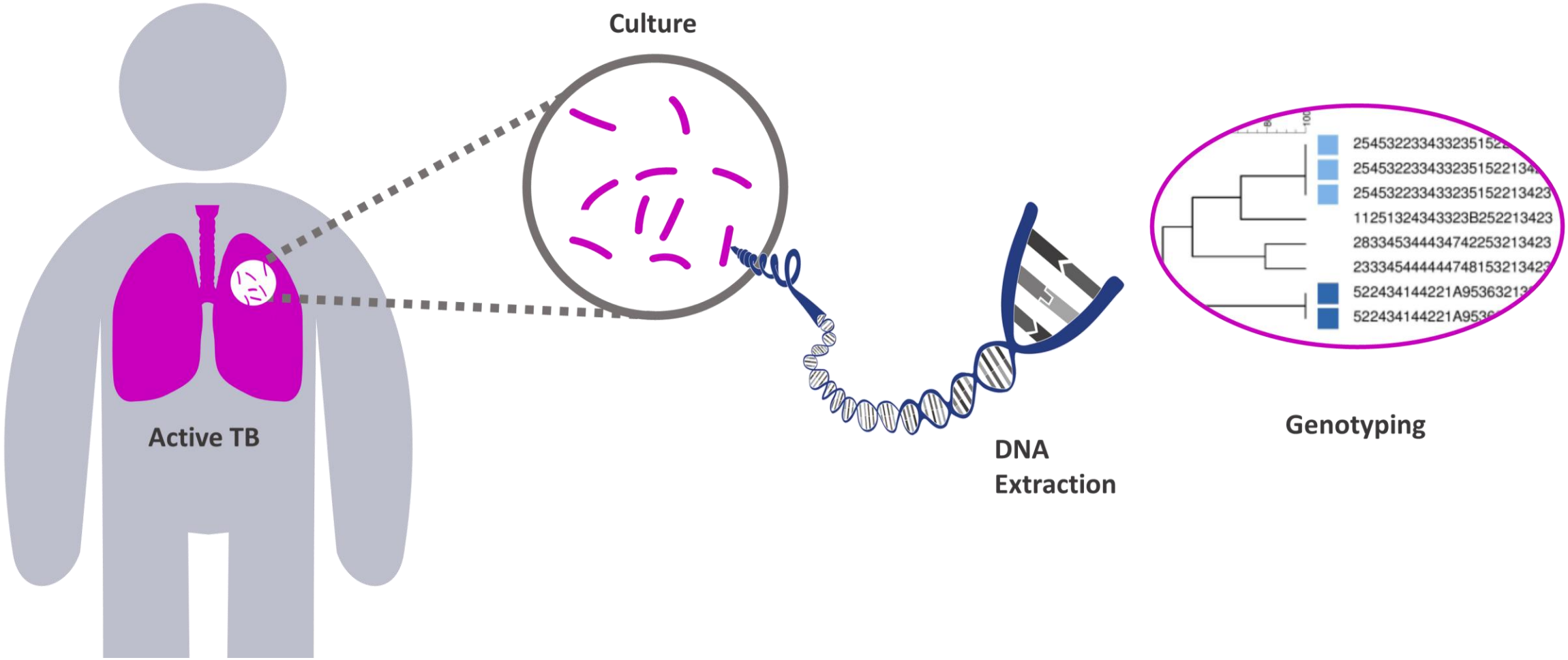
ukon erritory

Above the national rate incidence (13.9/100,000)

TB GENOMIC EPIDEMIOLOGY

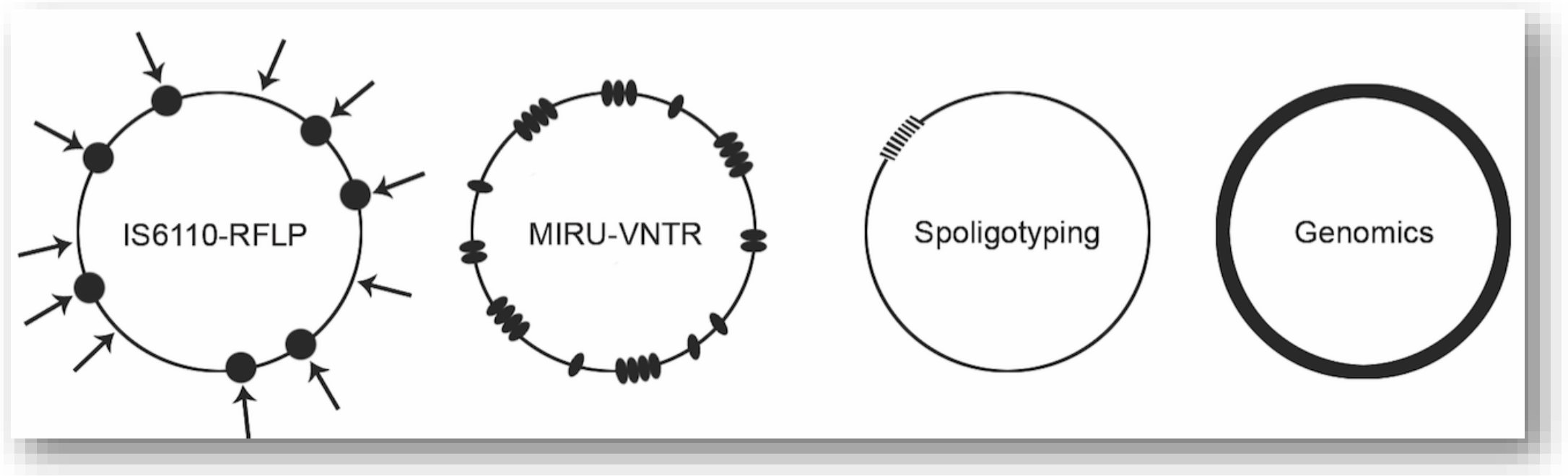
MOLECULAR METHODS

Genotyping



COMMON METHODS

Distinguishing Between *Mtb* Strains



POLLING QUESTION 1

What proportion of the tuberculosis genome do genotyping methods use to differentiate between strains?

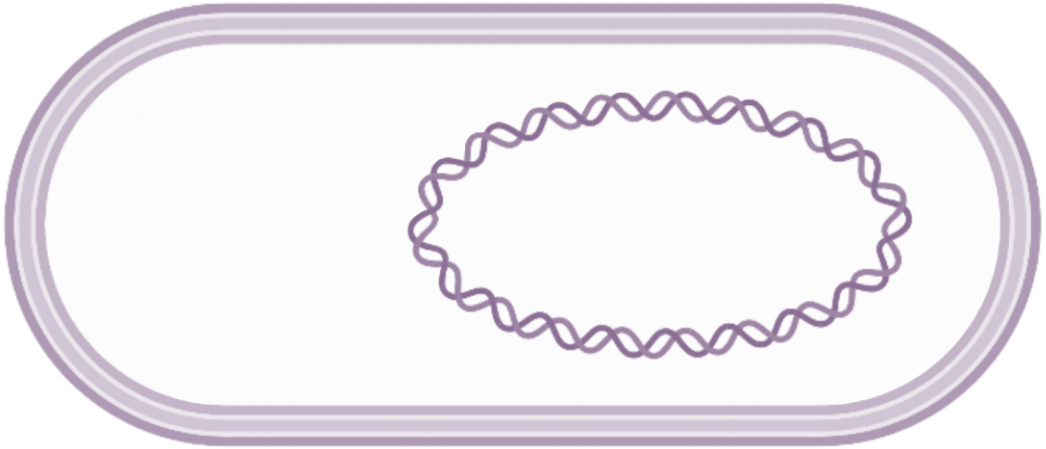
A. <1%

B. 1 - 25%

C. 26 - 50%

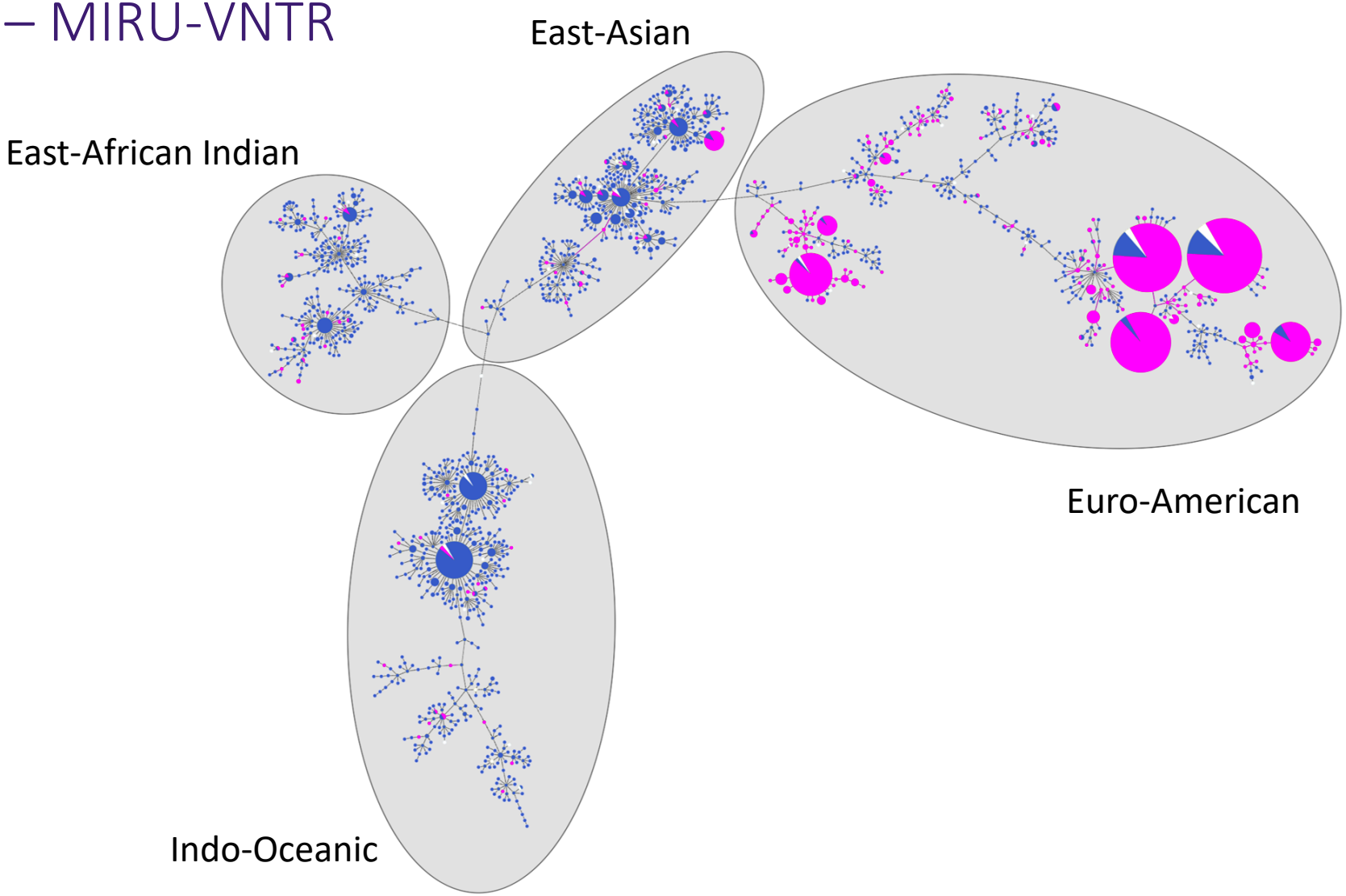
D. 51 -75%

E. 76 - 100%



MOLECULAR EPIDEMIOLOGY of TB

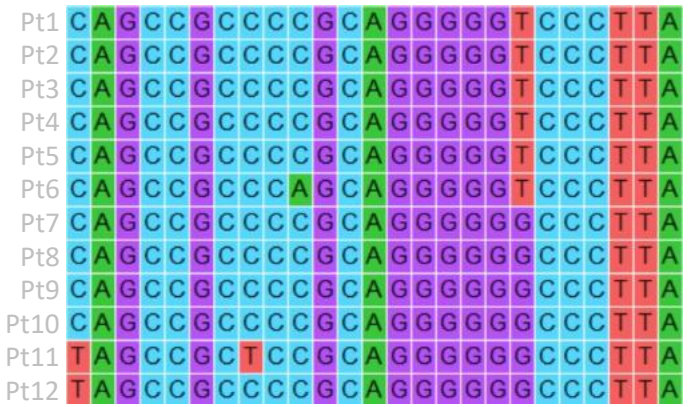
Minimum Spanning Tree – MIRU-VNTR

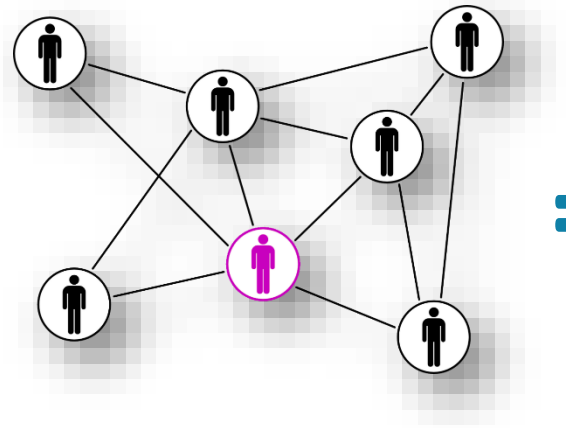


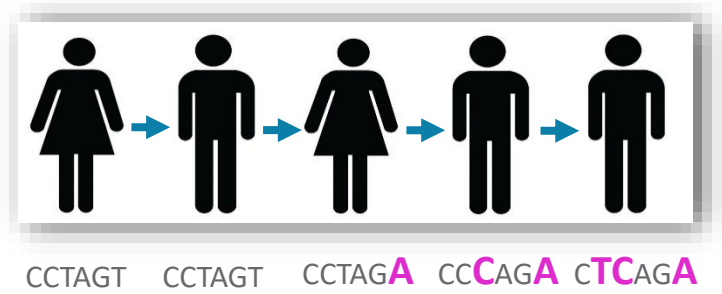
Guthrie JL et al. (2018) Clin. Infect. Dis. 66(6): 849–856.

GENOMIC EPIDEMIOLOGY

Tracking Transmission







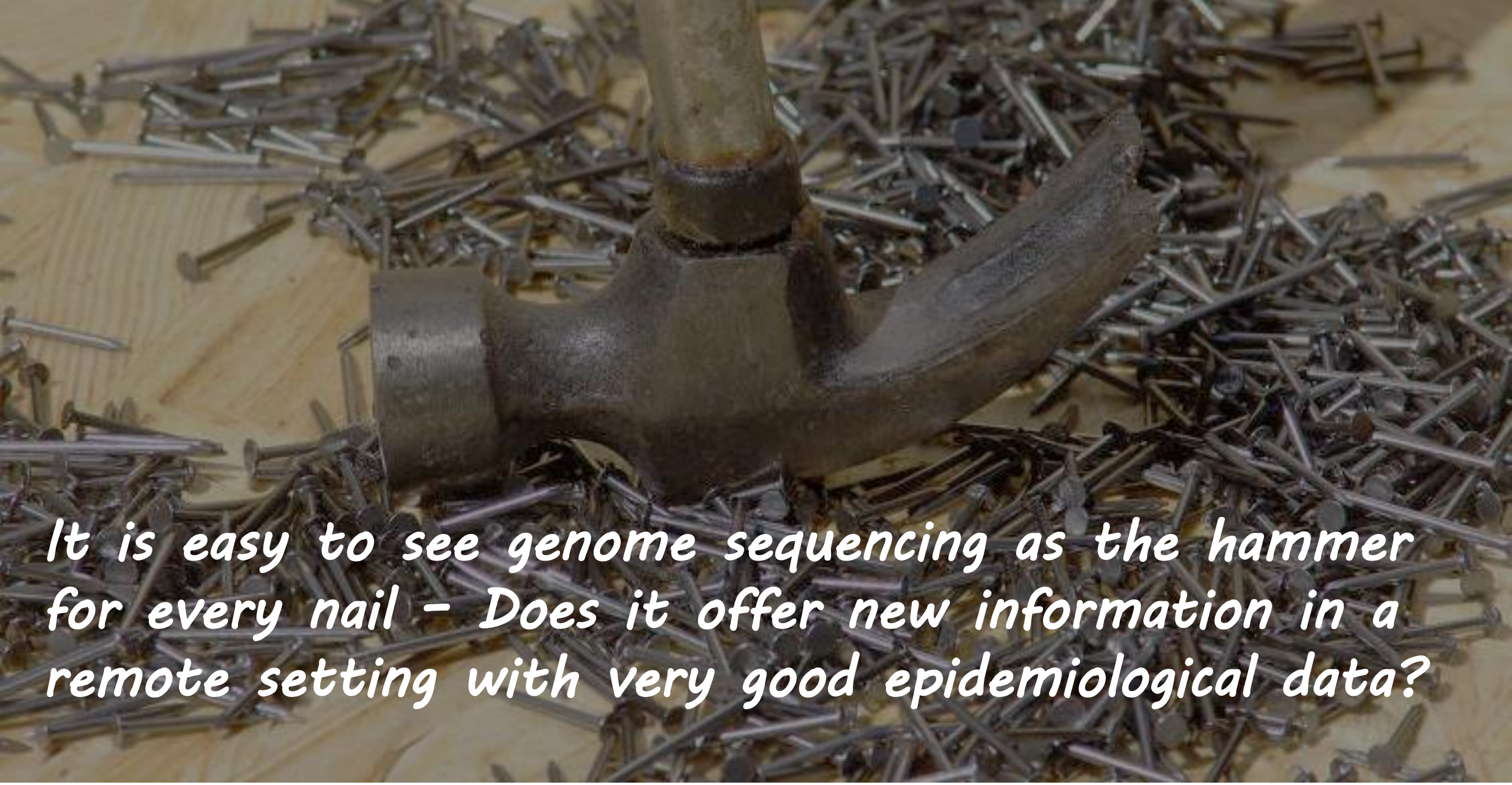
Whole genome sequencing

Contact investigation data
Key clinical and demographic data

Identify chains of transmission

Tuberculosis Transmission in a Remote Circumpolar Region

Area: 482,443 km²
Population: 38,400


It is easy to see genome sequencing as the hammer for every nail - Does it offer new information in a remote setting with very good epidemiological data?

PROJECT GOAL

Use small well-defined populations to examine molecular and genomic epidemiology and reconstruct transmission events

ulture

Resurrected **>2,300** *Mtb* isolates from frozen stocks

Extracted DNA from each culture

Performed 24-locus MIRU-VNTR genotyping

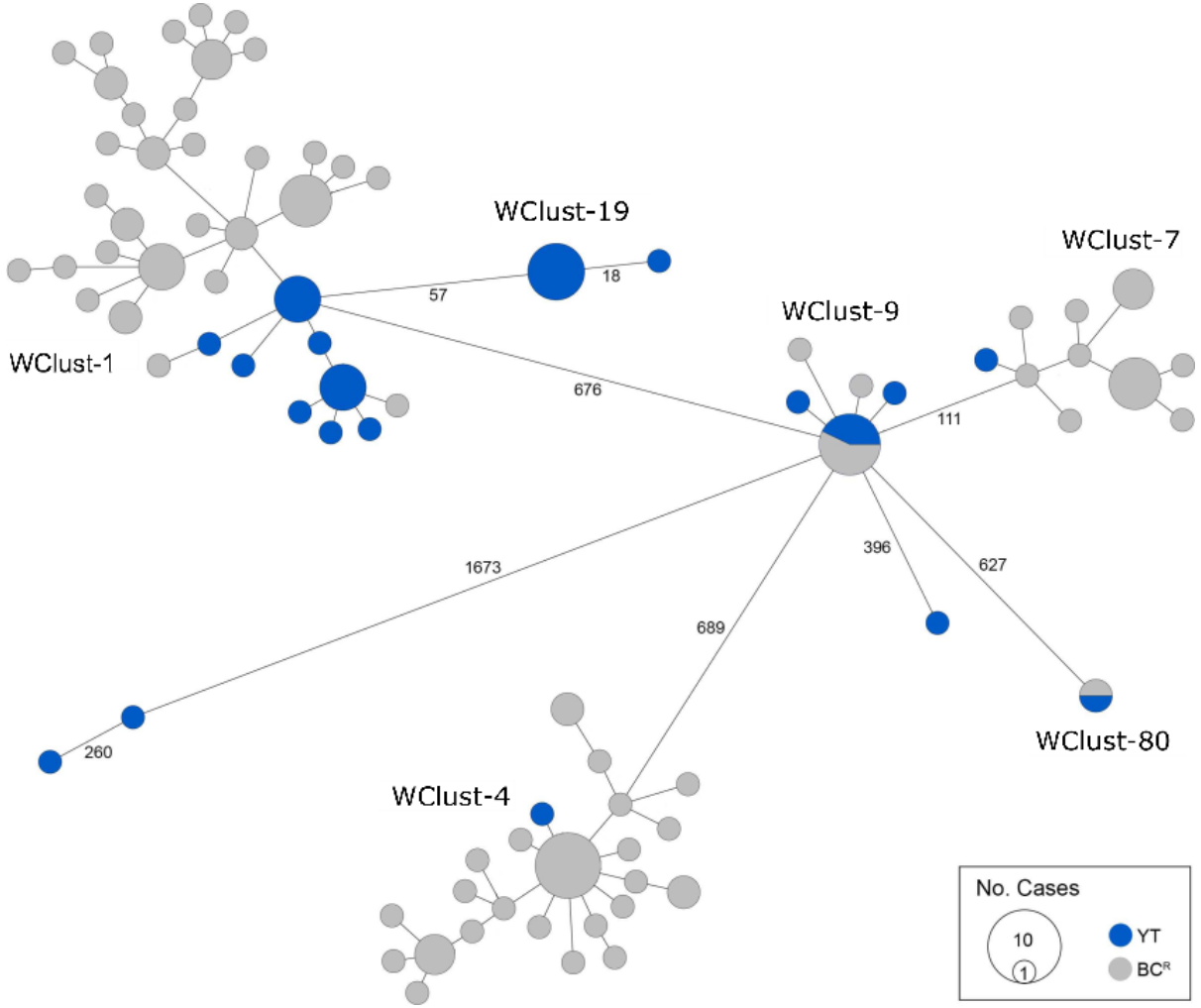
Sequencing of **>1,200** isolates and bioinformatics analyses

nalysis

Predict transmission networks and compare to epi based transmission networks

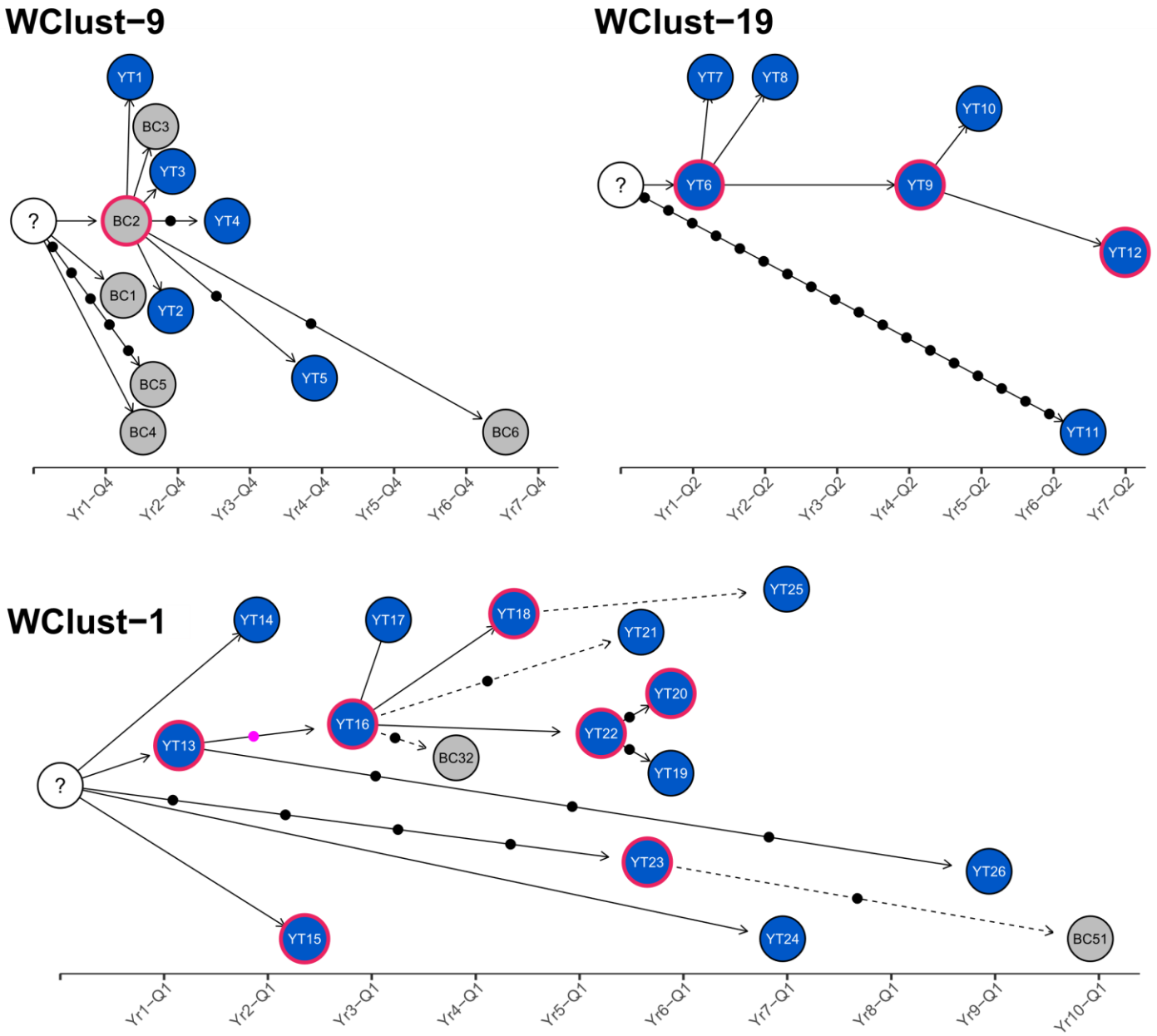
GENOMIC CLUSTERS

- Integrate genomic results of BC and YT
- Generate a minimum spanning tree to visualize genomic clusters



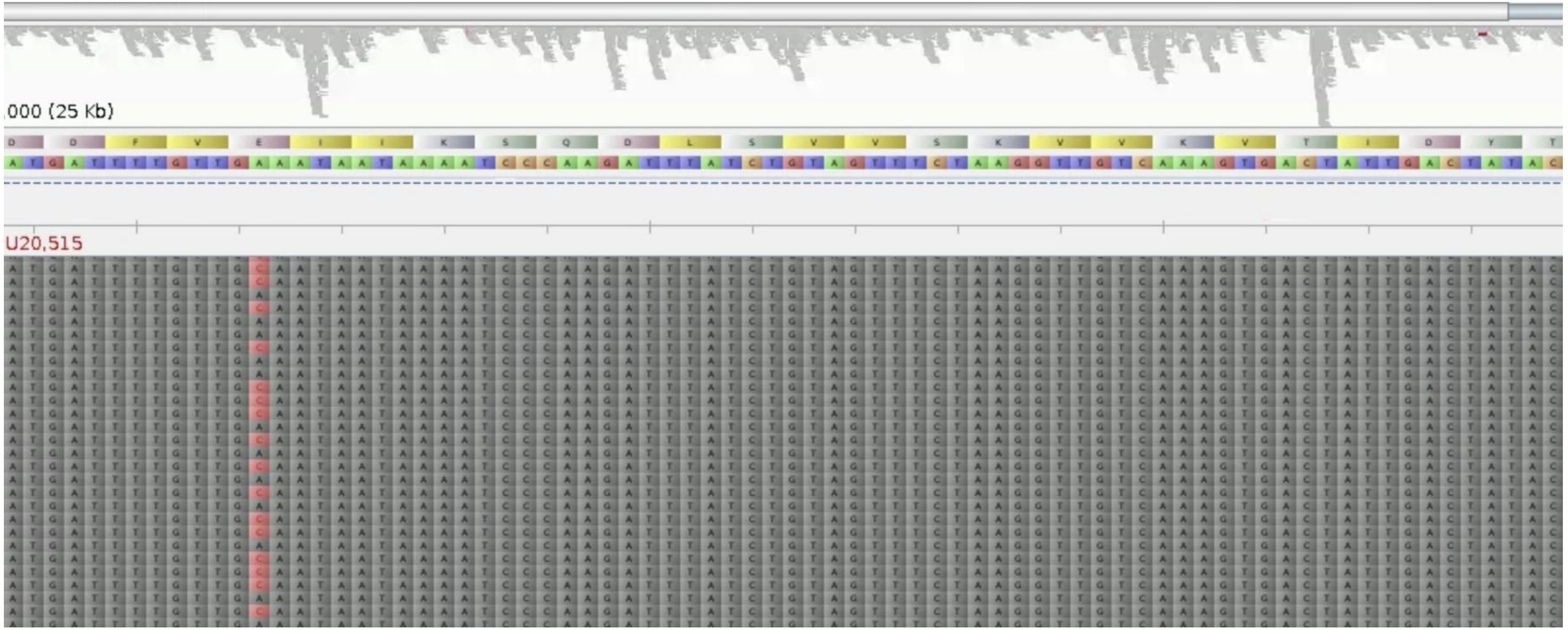
Guthrie JL et al. (2019) Epidemiol. Infect. 147, E188.

THREE GENOMICALLY DISTINCT CLUSTERS REPRESENTING TRANSMISSION

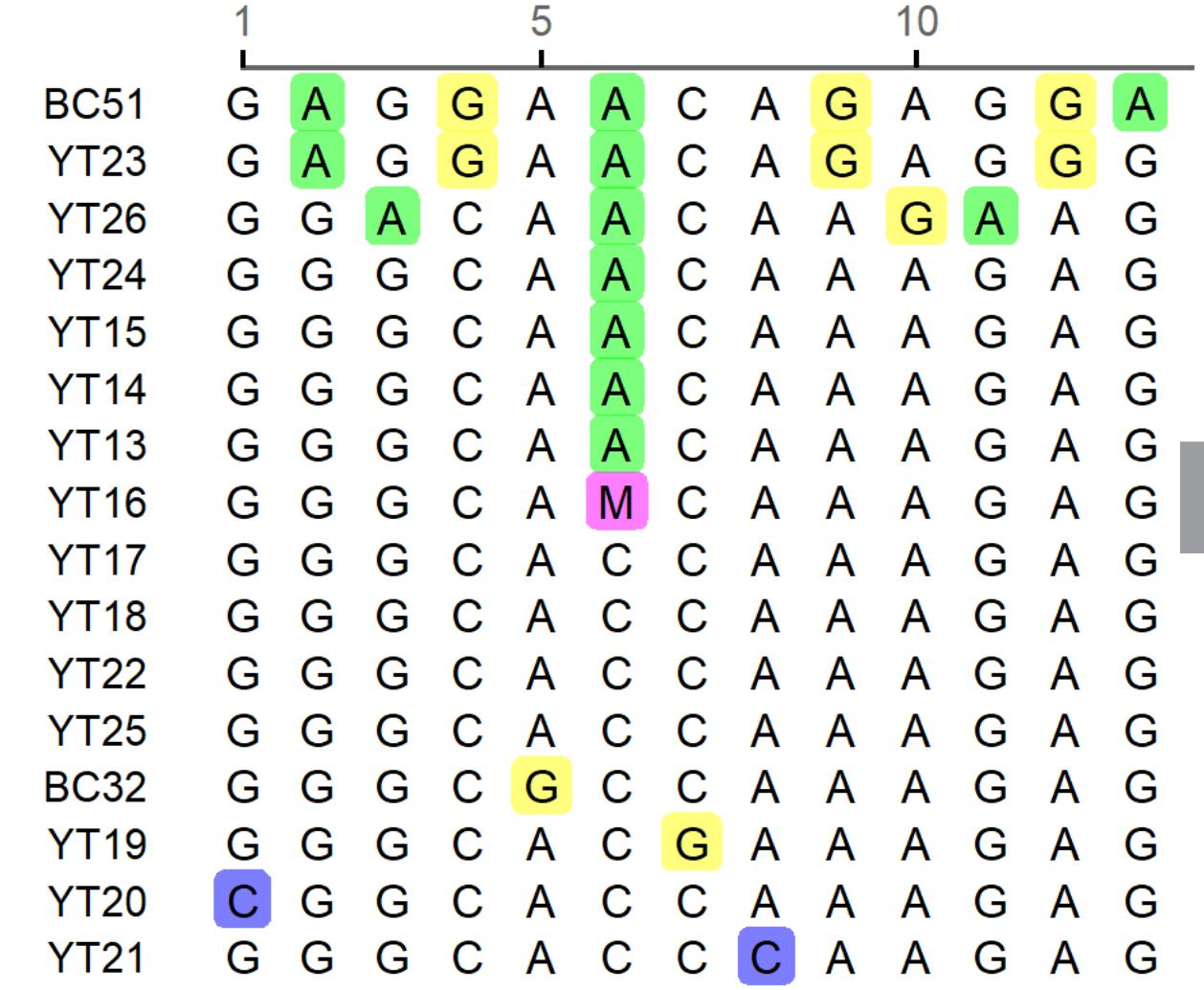


Guthrie JL et al. (2019) Epidemiol. Infect. 147, E188.

MIXED POSITION



WClust-1



MINOR

15% of reads A

85% of reads C

82% OF WGS MATCHED EPIDEMIOLOGY IN LOCATION OF ACQUISITION

SOURCE CASE IDENTIFICATION MATCHING VARIED BY CLUSTER

Guthrie JL et al. (2020) Epidemiol. Infect. 148, E15.

"To have had the WGS data, would have saved many hours of discussion – would have helped to focus the discussion by narrowing the list of potential sources"

ACCURACY
MIRU-VNTR conflicted with known epi connections, genomics more closely aligned and provided some new insights

PROGRAM ASSESSMENT
Assess the effectiveness of treatment and prevention programs

"Many of these confirmed our suspicions"

"It was nice to know this was a reactivation and not a contact of a missed source"

82% OF WGS MATCHED EPIDEMIOLOGY IN LOCATION OF ACQUISITION
SOURCE CASE IDENTIFICATION MATCHING VARIED BY CLUSTER

Guthrie JL et al. (2020) Epidemiol. Infect. 148, E15.

"To have had the WGS data, would have saved many hours of discussion – would have helped to focus the discussion by narrowing the list of potential sources"

ACCURACY
MIRU-VNTR conflicted with known epi connections, genomics more closely aligned and provided some new insights

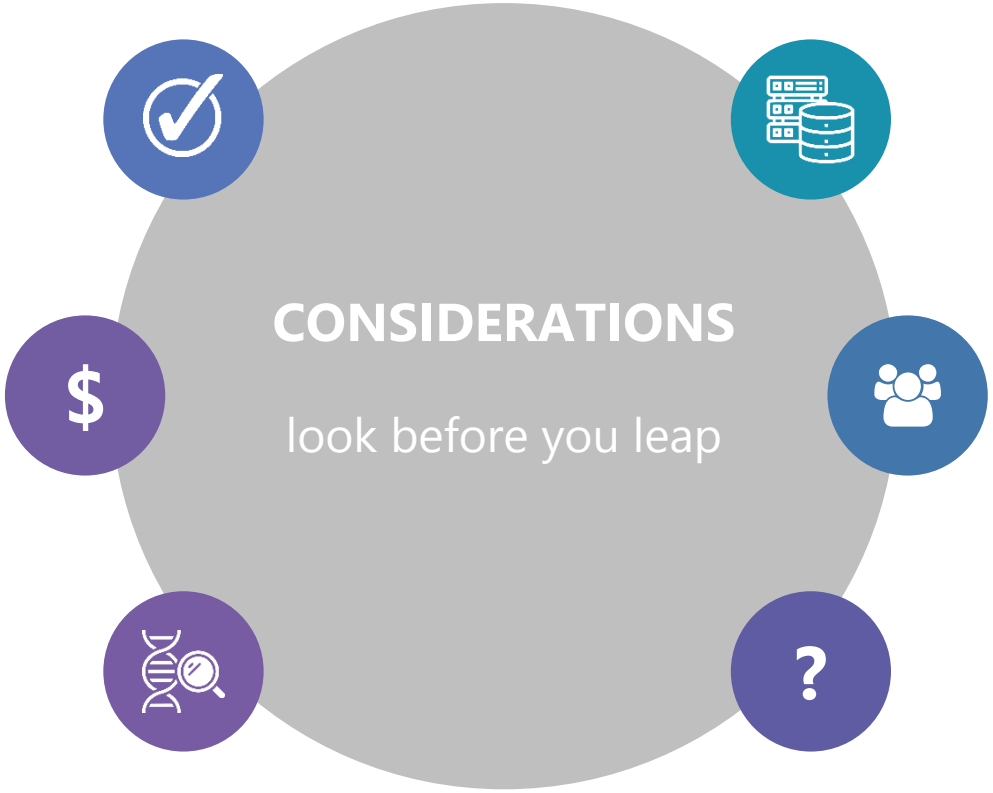
PROGRAM ASSESSMENT
Assess the effectiveness of treatment and prevention programs

"Many of these confirmed our suspicions"

"It was nice to know this was a reactivation and not a contact of a missed source"

GENOMIC EPIDEMIOLOGY PROGRAM

GENOMIC EPIDEMIOLOGY PROGRAM



POSSIBLE

Is sequencing even possible?

DATA

Access to epidemiological data?

COST

Can you afford it?

EXPERTISE

Do you have genomic epidemiology expertise?

INTERPRETABILITY

Is the genome itself readily interpretable in an epidemiological context?

JUSTIFICATION

Why use genomic epidemiology?

Guthrie JL and Gardy JL. (2017) Ann. N.Y. Acad. Sci. 1388:59–77.

GENOMIC EPIDEMIOLOGY PROGRAM

Challenges

WGS Data

- Standardized nomenclature and automated processes for data, including sample and file naming
- Genomics database that tracks samples and captures sequencing related data
- Sufficient physical storage capacity for large genome sequencing files
- Organized hierarchical file structure to efficiently store WGS files allowing for automated downstream processes

GENOMIC EPIDEMIOLOGY PROGRAM

Challenges

Epidemiological Data

- Collection and entry of data relevant to surveillance and outbreaks in standardized fields within CCM/iPHIS
- Routine and automated data extraction from CCM/iPHIS and linkage to WGS outputs

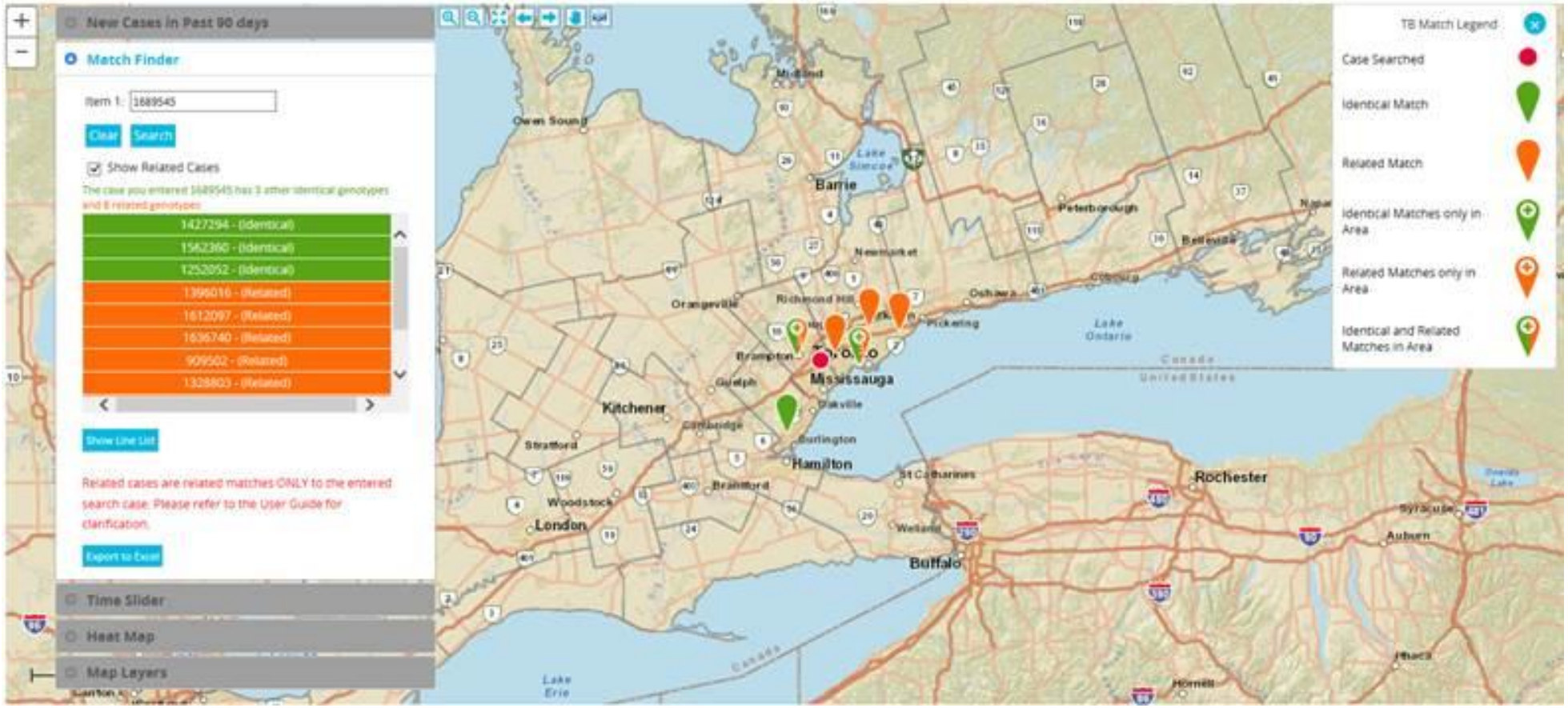
GENOMIC EPIDEMIOLOGY PROGRAM

Challenges

Data Analysis & Communication

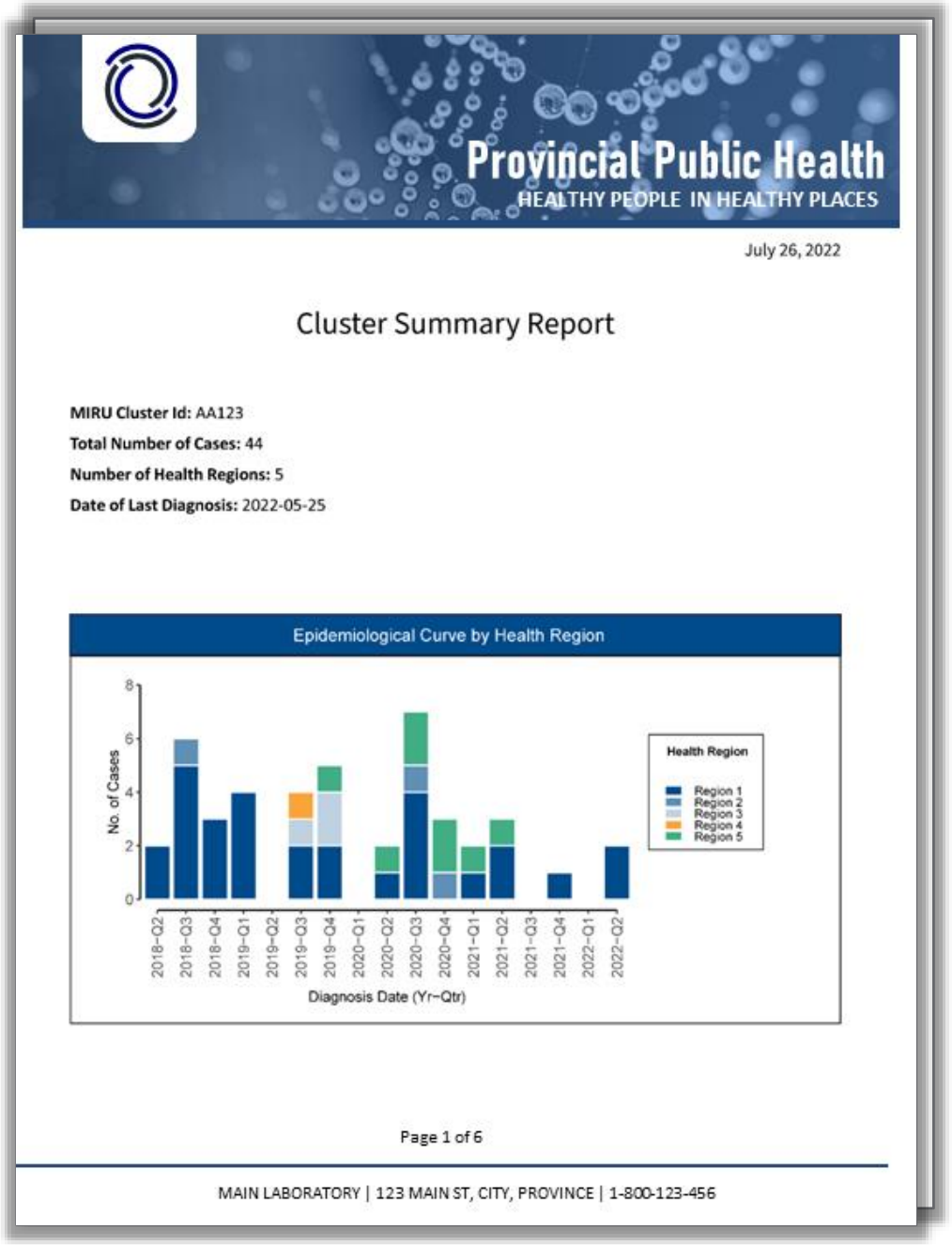
- Personnel with a combination of epidemiological, bioinformatics, and evolutionary biology expertise and disease and pathogen specific knowledge
- Effective means to communicate TB genomic epidemiology to stakeholders
- Timeliness of alerts to possible transmission and outbreak investigations such that it is actionable

OUT-TB WEB


New Cases in Past 90 days
Match Finder
Item 1: 1689545
Clear
Search
Show Related Cases
1427294 - (Identical)
1562360 - (Identical)
1252052 - (Identical)
1396016 - (Related)
1612097 - (Related)
1636740 - (Related)
909502 - (Related)
1328803 - (Related)
Show Line List
Related cases are related matches ONLY to the entered search case. Please refer to the User Guide for clarification.
Export to Excel
Time Slider
Heat Map
Map Layers
TB Match Legend
Case Searched
Identical Match
Related Match
Identical Matches only in Area
Related Matches only in Area
Identical and Related Matches in Area
Barrie
Peterborough
Oshawa
Mississauga
Kitchener
Hamilton
London
Buffalo
Rochester
Lake Ontario
Lake Erie

CLUSTER REPORT

Prototype



Provincial Public Health
HEALTHY PEOPLE IN HEALTHY PLACES

July 26, 2022

Cluster Summary Report

MIRU Cluster Id: AA123
Total Number of Cases: 44
Number of Health Regions: 5
Date of Last Diagnosis: 2022-05-25

Epidemiological Curve by Health Region

Page 1 of 6

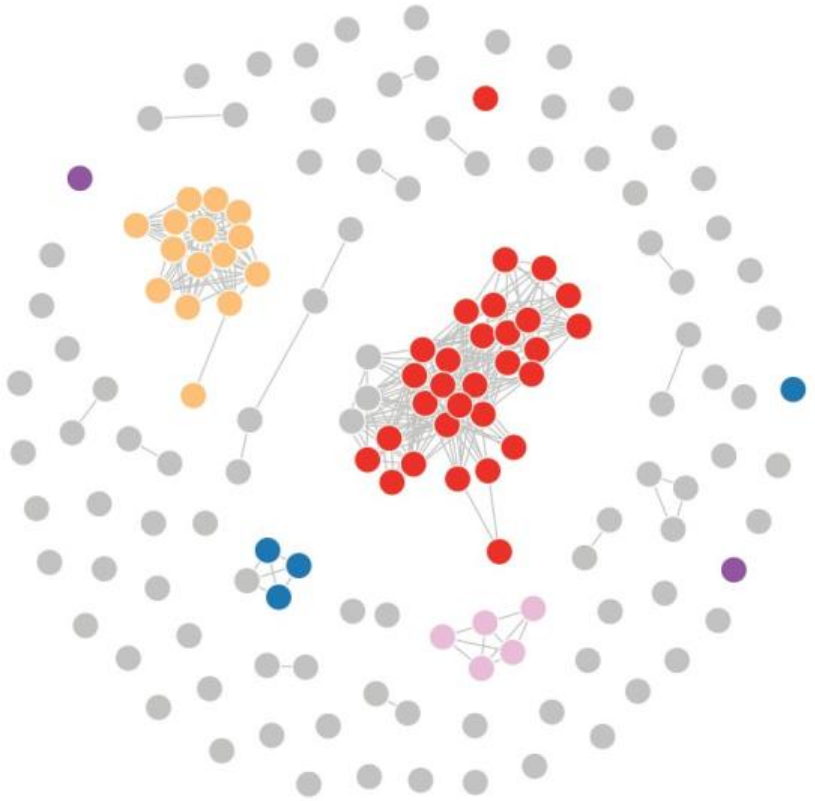
MAIN LABORATORY | 123 MAIN ST, CITY, PROVINCE | 1-800-123-456

POLLING QUESTION 2

Genomic epidemiology can resolve all tuberculosis outbreaks?

A. True

B. False



CONCLUSIONS

Genomics offers a **deeper understanding** of infectious disease epidemiology. Paired with active surveillance programs can resolve more outbreaks, suggest new modes of transmission, and reveal new reservoirs of disease

…however, implementation of a public health genomic epidemiology program remains challenging and **requires significant investments** in infrastructure and personnel with appropriate expertise.

ACKNOWLEDGEMENTS

British Columbia

Dr. Jennifer Gardy

Dr. James Johnston

Dr. Victoria Cook

Dr. David Roth

Dr. Inna Sekirov

Dr. Mabel Rodrigues

Danielle Jorgensen

BCCDC PHL

Yukon

Dr. Brendan Hanley

Lori Strudwick

YCDC

Ontario

Dr. Melissa Richard-Greenblatt

Dr. Fran Jamieson

Alex Marchand-Austin

Carla Duncan

Karen Lam

Daria Pyskir

PHO TB Laboratory

Oxford

Dr. Derrick Crook

Dr. Tim Walker

Suppor t

THE UNIVERSITY OF BRITISH COLUMBIA

Lauréats KILLAM Laureates

BC Centre for Disease Control
An agency of the Provincial Health Services Authority

CIHR IRSC
Canadian Institutes of Health Research | Instituts de recherche en santé du Canada

Canada Research Chairs | Chaires de recherche du Canada

Questions?