KEMENTERIAN PERTANIAN
DAN KETERJAMINAN MAKANAN

WEBINAR

Sesi Perkongsian Latihan Luar Negara
16 Februari 2026

CHAN ZUCKERBERG BIOHUB

Agency for Science, Technology and Research SINGAPORE

DukeNUS Medical School

ASIA PATHOGEN GENOMICS Initiative

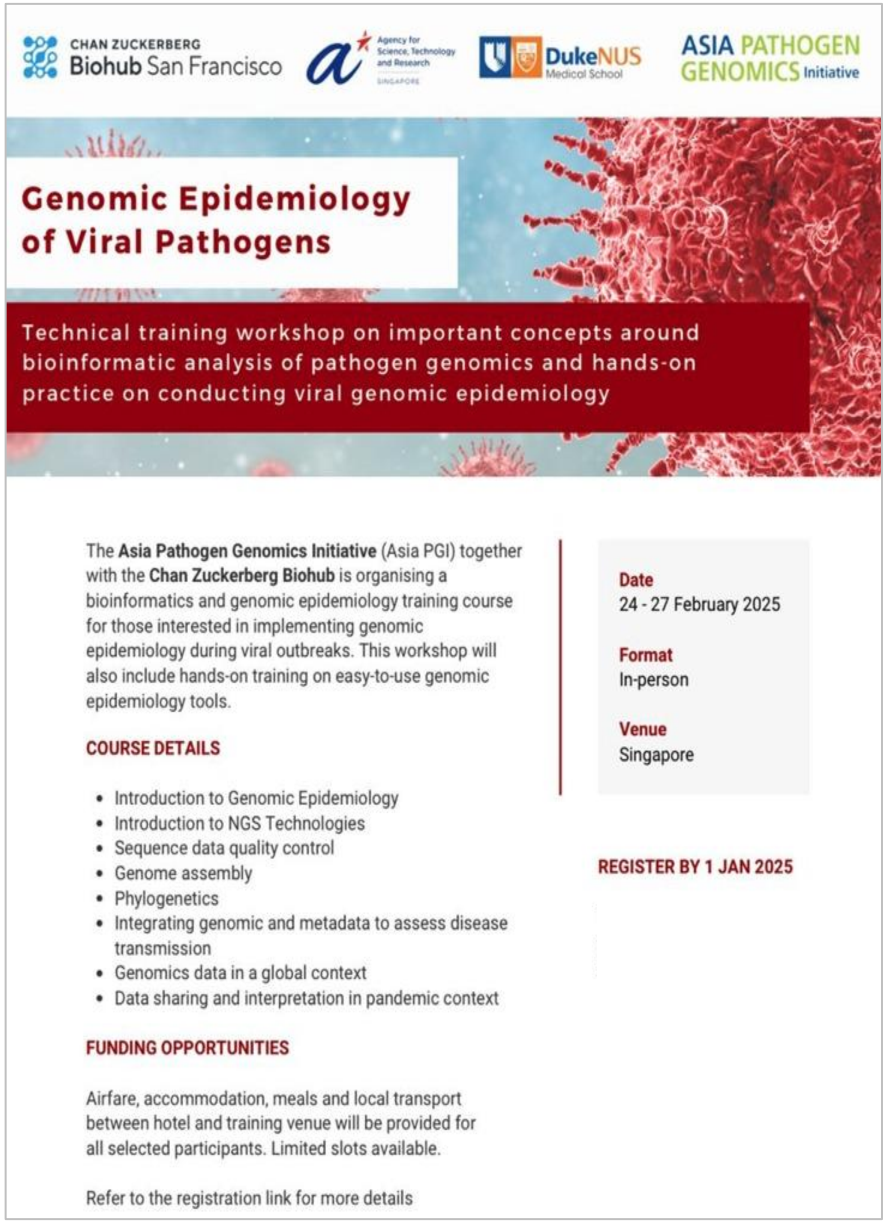
GENOMIC EPIDEMIOLOGY OF VIRAL PATHOGENS

Duke NUS Medical School, Singapore
24-27 February 2025

MOHAMMAD MASRIN BIN AZAMI
Pegawai Penyelidik
Institut Penyelidikan Veterinar, Ipoh

PENGENALAN KURSUS

- Genomic Epidemiology Of Viral Pathogens dianjurkan oleh **Asia Pathogen Genomics Initiative (Asia PGI)** bersama **Chan Zuckerberg Biohub.**
- **Lokasi:** Shaw Foundation Alumni House, NUS.
- **Peserta**: 28 peserta dari 16 negara
  - Malaysia – 3 orang: VRI, MKAK, UKM
- Peserta dipilih berdasarkan latarbelakang teknikal dan kuota mengikut negara.



PENGANJUR

Chan Zuckerberg Biohub

- Nonprofit biomedical research organization established in 2016, funded by Mark Zuckerberg and Priscilla Chan
- To accelerate fundamental biomedical research with a collaborative and technology-forward approach

A*STAR

- Agency for Science, Technology and Research
- Singapore government agency under the Ministry of Trade and Industry

Duke NUS Medical School

- Collaboration between Duke University (USA) and the National University of Singapore (NUS)

Asia Pathogen Genomics Iniative (ASIA-PGI)

- Flagship collaboration under the Duke–NUS Centre for Outbreak Preparedness (COP). Established in late 2021 with support from the Bill & Melinda Gates Foundation,
- Goal is to **boost regional health security by expanding access to pathogen genomic sequencing across South and Southeast Asia**

KEMUDAHAN KURSUS

Shaw Foundation Alumni House, NUS

12 KM

Mercure Singapore Bugis

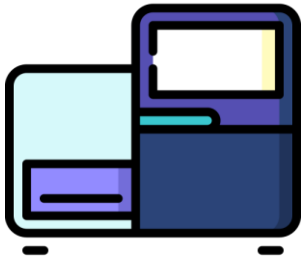
- ✓ **Logistik** – Airfare, airport transfer, daily ground transport
- ✓ **Penginapan** – Mercure Singapore Bugis
- ✓ **Makan minum** semasa kursus

JADUAL KURSUS

23 FEB	24 FEB	25 FEB	26 FEB	27 FEB	28 FEB	1 MAC
IPH - SIN	**AM** • Registration • Welcome introduction • Ice breaking • Pre-Survey **PM** • Sequencing data QC • Genome assembly • Genomic epidemiology Welcome dinner @ Botanic Garden	**AM (Lecture)** • Genomic Metadata & integration • Genomic epidemiology tools **PM (Practical)** • Metagenomic • Practical data analysis	**AM (Lecture)** • Phylogenetics analysis • CZ Gen EPI for genomic epidemiology analysis **PM (Practical)** • Genomic and phylogenetics interpretation • Practical data analysis	**AM (Lecture)** • Genomic tools and databases • Sequence submission **PM** • Networking • Wrap up & certificate	• Trip around Singapore • Free and easy	SIN - IPH

MATLAMAT KURSUS

Pendedahan **aplikasi dan teknik analisa bioinformatic asas dan lanjutan**

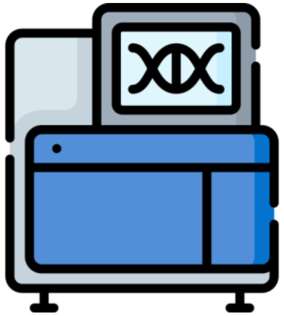


Perkongsian **pengalaman** dan **teknologi terkini** berkaitan **Next Generation Sequencing (NGS)**

Pengenalan konsep **genomik epidemiologi** dalam penyiasatan wabak

Mewujudkan **jaringan kolaborasi professional dalam bidang genomik** dan **kesihatan**

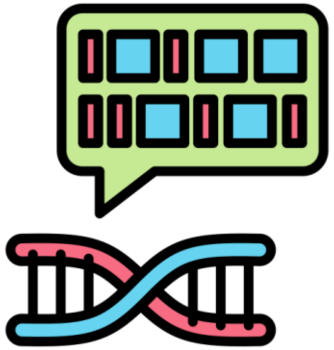
AKTIVITI / TOPIK UTAMA



1

Teknologi Next Generation Sequencing (NGS)

- Jenis sequencing
- Teknologi terkini NGS
- Sequencing



2

Analisa data sequencing

- Raw data
- Quality control
- Assembly
- Downstream analysis

3

Genomic Epidemiology

- Tools & application
- Phylogenetics
- Databases

1 TEKNOLOGI NGS

Jenis / Aplikasi Sequencing

Sanger

- 1 gen / fragmen
- Saiz DNA pendek (<1,000 bp)

One fragment → one sequence → chromatogram output

WGS

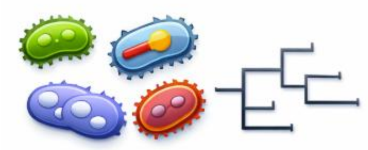
- Keseluruhan genom
- Saiz data besar dan analisa yang kompleks

Entire genome → fragmentation → assembly

Amplicon

- Bahagian spesifik dalam genom
- Target gen / Patogen spesifik

Targeted region → PCR amplify → deep sequencing



16s

- Merujuk kepada gen 16s bakteria
- Menentukan komposisi komuniti bakteria

Bacteria identification → phylogeny → microbiome profiling

Metagenomic

- Kesemua DNA/RNA material di dalam sampel
- Tidak perlu target / primer spesifik

All DNA in sample → shotgun sequencing → taxonomic profiling

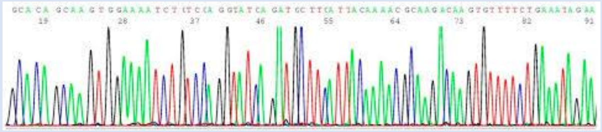
1 TEKNOLOGI NGS

Teknologi sequencing

1st Generation

Low throughput, short read sequencing

Sanger sequencing



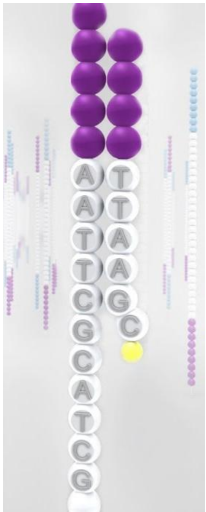
1 x 500bp sequence per well

1 PCR product → 1 read

2nd Generation

High throughput, short read sequencing

Illumina sequencing

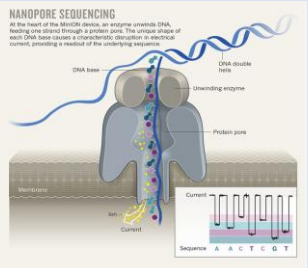


Millions of 300-600bp sequences

3rd Generation

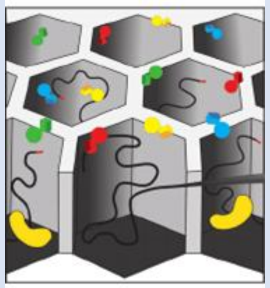
High throughput, long read sequencing

Nanopore Sequencing



Thousands to millions of 1-20kb sequences

PacBio sequencing



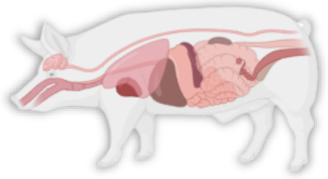
Thousands of 1 - 30kb sequences

2 ANALISA DATA SEQUENCING

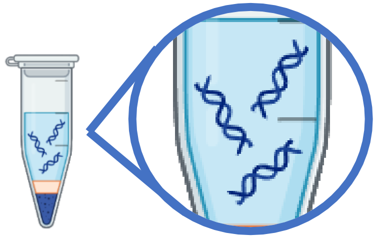
WORKFLOW / PIPELINE

- Rangkaian langkah kerja yang tersusun dan berurutan
- Output satu langkah menjadi input kepada langkah seterusnya.

Sequencing Workflow

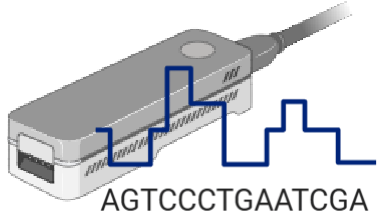


Sample Selection



DNA/RNA extraction

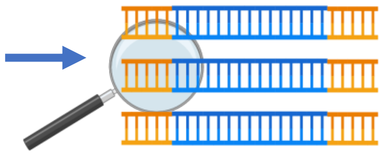
Library Preparation



Sequencing

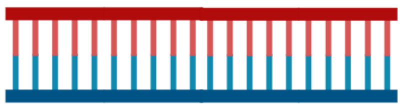
Bioinformatics Analysis Pipeline

Raw data

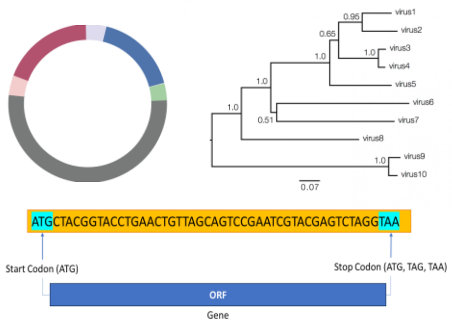


Quality check / control

Assembly



Consensus



Downstream analysis

3 GENOMIC EPIDEMIOLOGY

- **Asas Epidemiologi :**
  - Siapa (**Who**) : Maklumat hos / jantina / status klinikal / vaksinasi / bil. kematian
  - Bila (**When**) : Maklumat masa, tarikh mula symptom, tarikh sampling, tempoh wabak
  - Di mana (**Where**) : Maklumat Lokasi – metadata geografi / Koordinat GPS
  - Apa (**What**) : Jenis penyakit / agen penyebab / strain / variant
  - Bagaimana (**How**) : Cara penularan / sumber jangkitan / laluan jangkitan

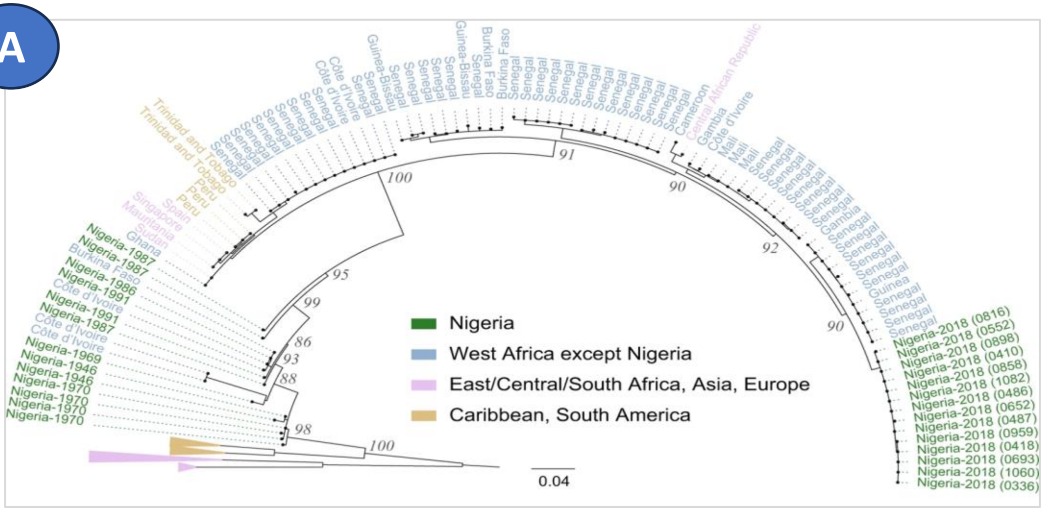
- **Epidemiologi genomik** merupakan cabang epidemiologi **moden** yang meng**integrasi**kan **analisis genom patogen** bagi menilai taburan, *transmission dynamic*, dan faktor penentu /penyebab kejadian penyakit berjangkit dalam populasi.

3 GENOMIC EPIDEMIOLOGY

Pathogen Identity

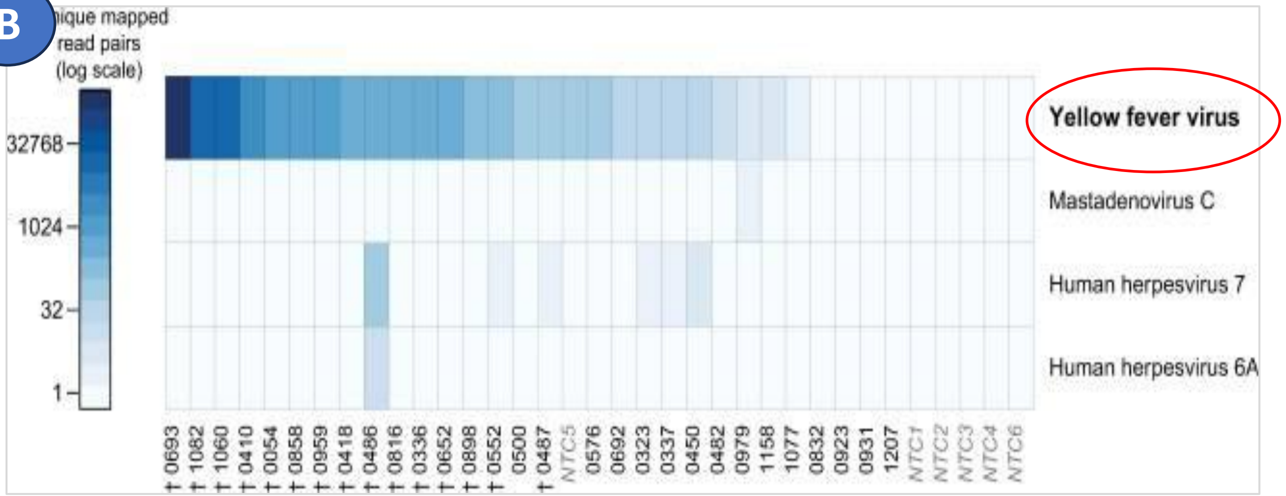
Ajogbasile *et al.*, 2020. Scientific Reports

A



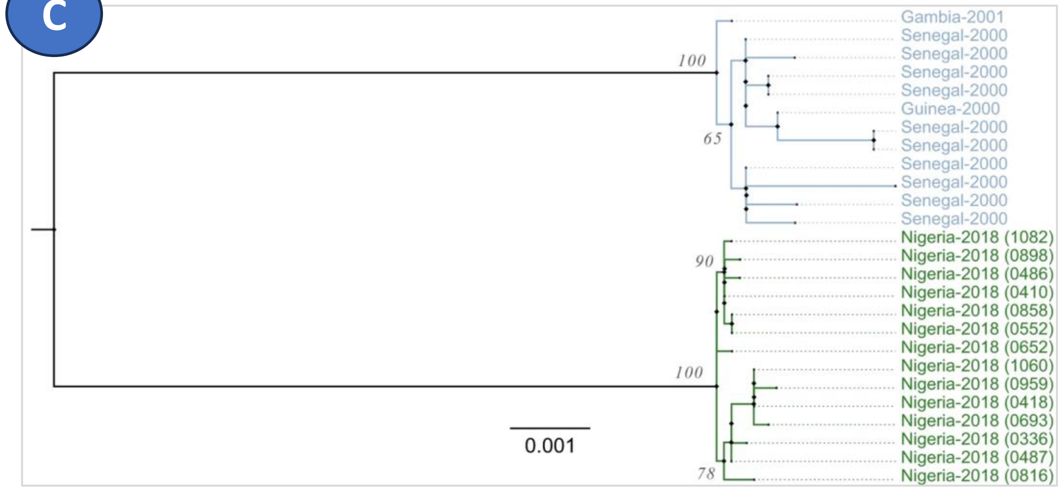
Analisa hubungan genetik mengikut Rantau – mengenalpasti hubungan wabak tempatan dengan wabak global

B



Pengenalpastian jenis virus melalui perbandingan kadar bilangan sequence yang berjaya dipetakan kepada reference sequence virus yang disyaki – **pengesahan diagnosis berasaskan data genomik**

C

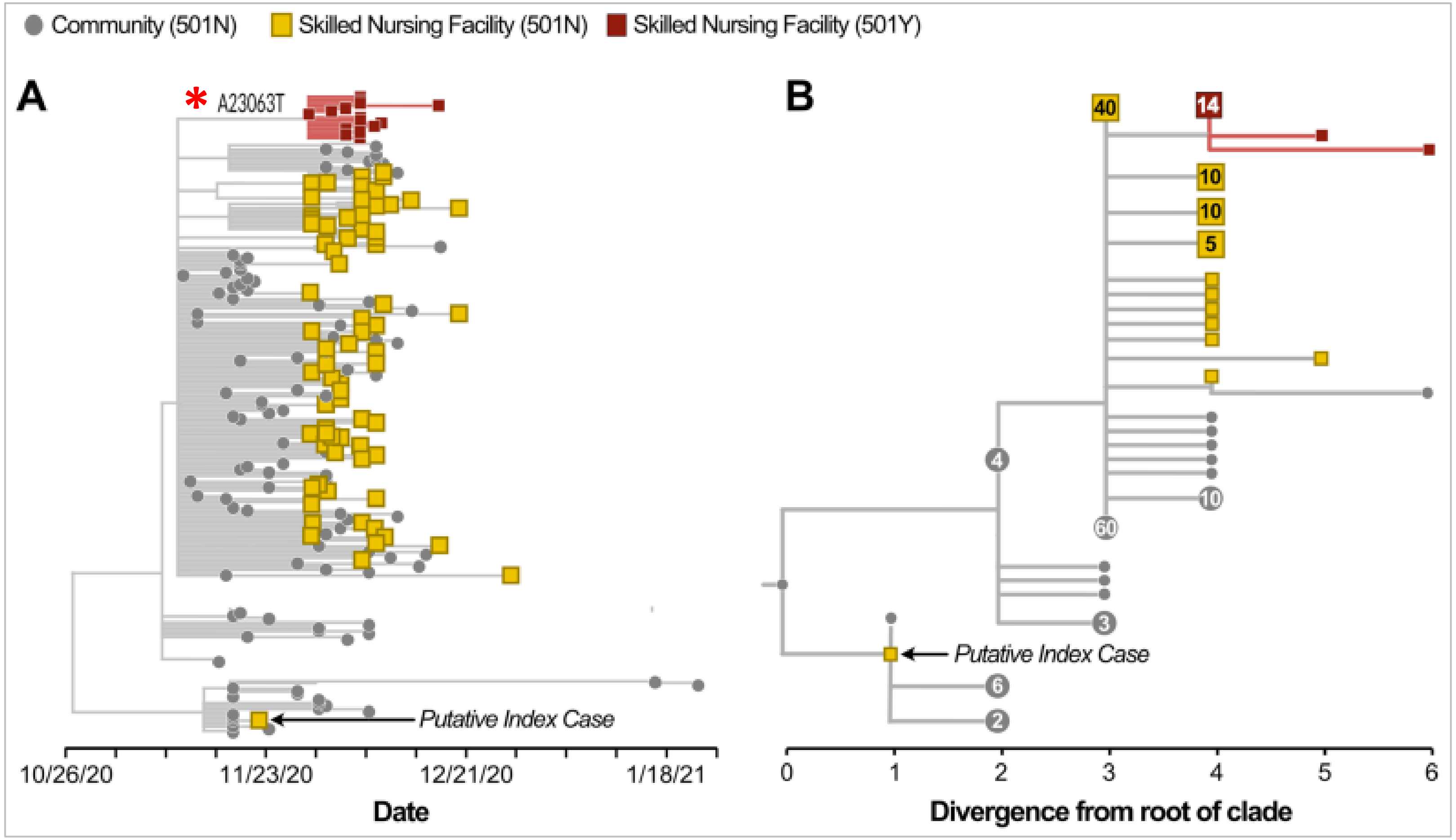


Analisa hubungan genetik untuk mengenalpasti tahap persamaan dan perbezaan, hubungan genetik dan asal usul virus (parents – child).

- Pengenalpastian jenis patogen
- Pengesahan diagnosis penyakit
- Pengenalpastian hubungan wabak tempatan dengan wabak global
- Pengenalpastian tahap persamaan / perbezaan
- Asal-usul / hubungan genetik

Relatedness to other outbreaks / organisms

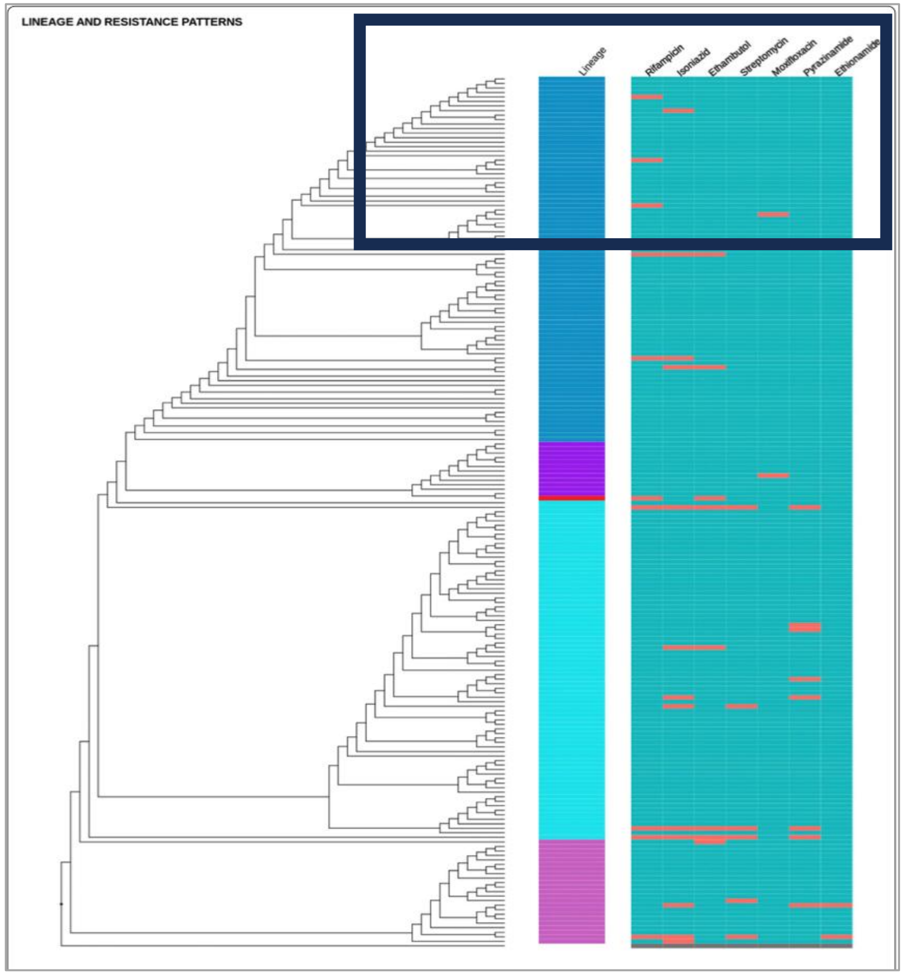
*Terdapat perbezaan pada kedudukan 501 aa (spike protein). 501N (Wildtype) 501Y (mutasi N→ Y)

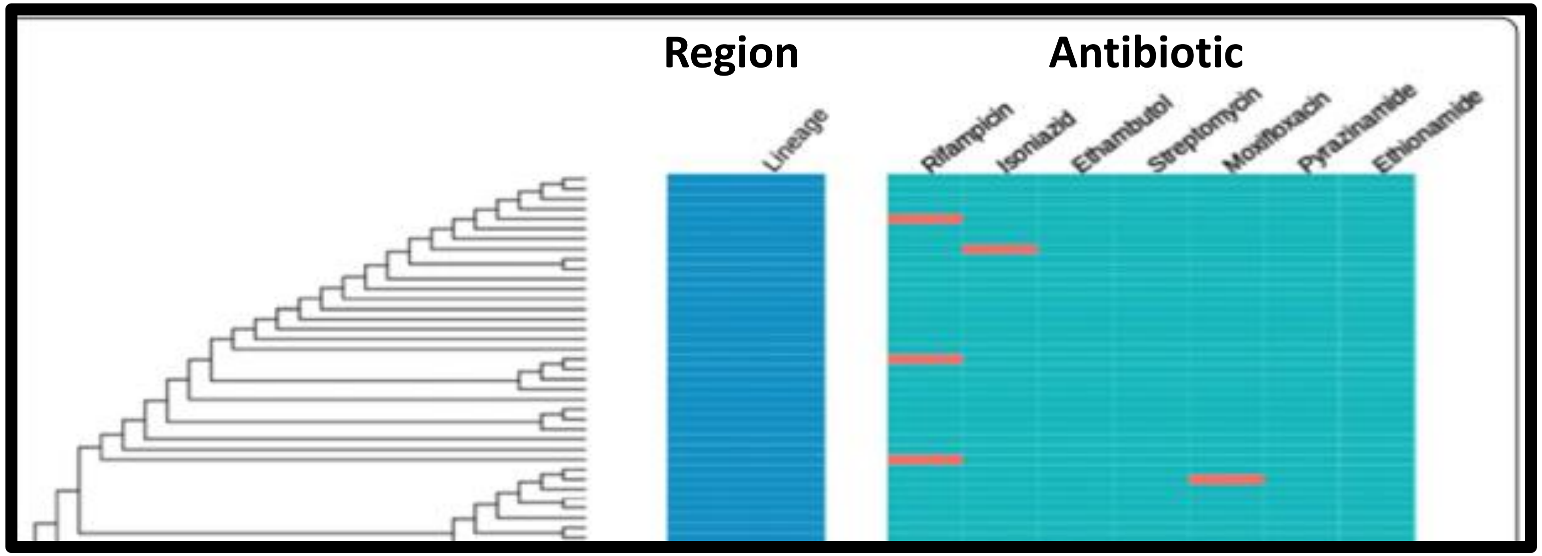


- Perbandingan antara kejadian wabak - Patogen daripada wabak yang sama akan mempunyai **genom yang sangat serupa,** manakala patogen daripada wabak berbeza akan menunjukkan **perbezaan genetik yang lebih besar**.
- Mengenalpasti sama ada jangkitan/wabak adalah dari **rantaian penularan yang sama atau kemasukan baru berdasarkan hubungan genetik**

Stoddard, Black et al, 2022. BMC Public Health

Sensitivity / Resistance to Drugs



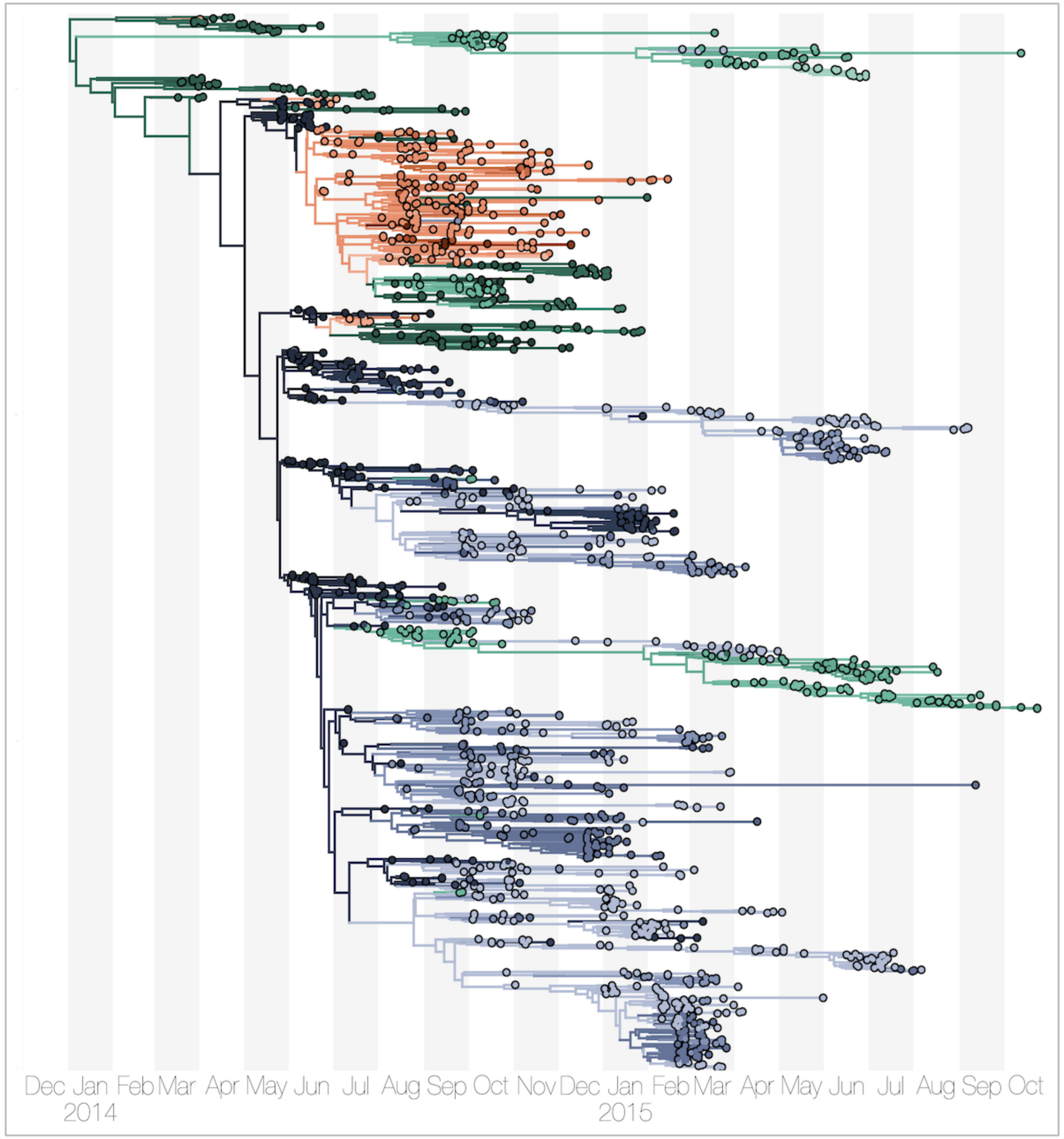


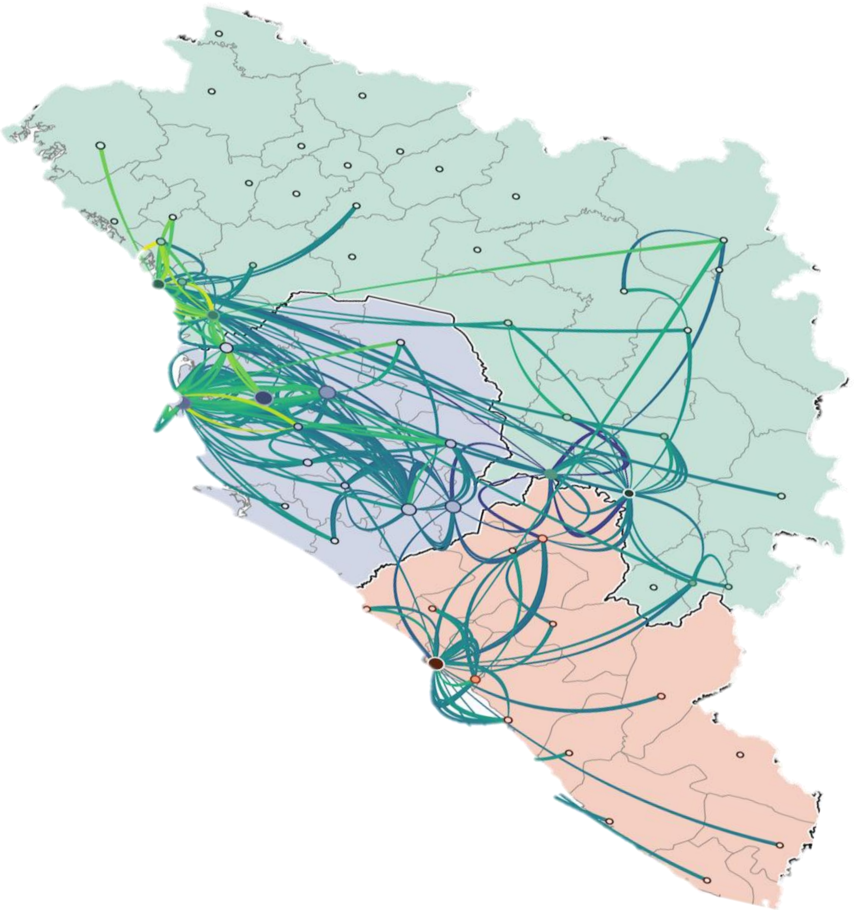
Sensitive Resistance

Mengenalpasti **sensitiviti atau rintangan terhadap antibiotik**, serta hubungannya dengan **lineage atau kluster genetik** melalui pemetaan data genomik.

Mutayoba et al, 2022. BMC Genomics.

Identifying and Tracking Emerging or New Infectious Threats

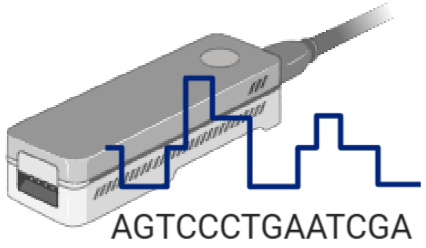




Data genom patogen digunakan untuk menjejak **kemunculan, evolusi, dan pergerakan penyakit berjangkit dari semasa ke semasa dan merentas lokasi.**

Dudas et al., 2017. Nature.

Computational Biology



Sequencing

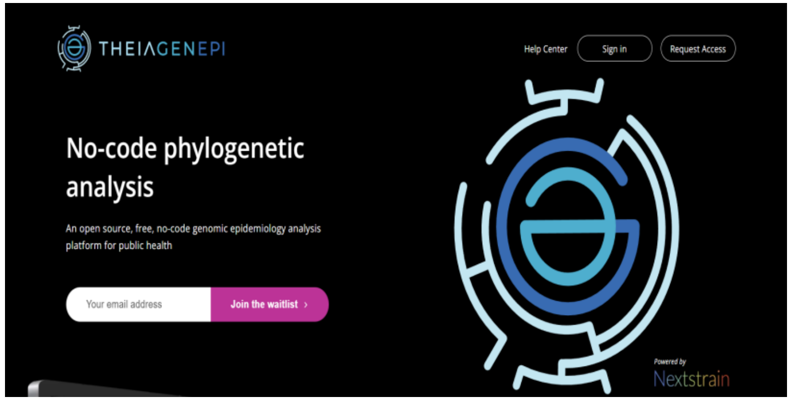
Genomic Data

In silico analysis

All the interesting findings/insights were made possible through the use of several **genomic epidemiology tools**

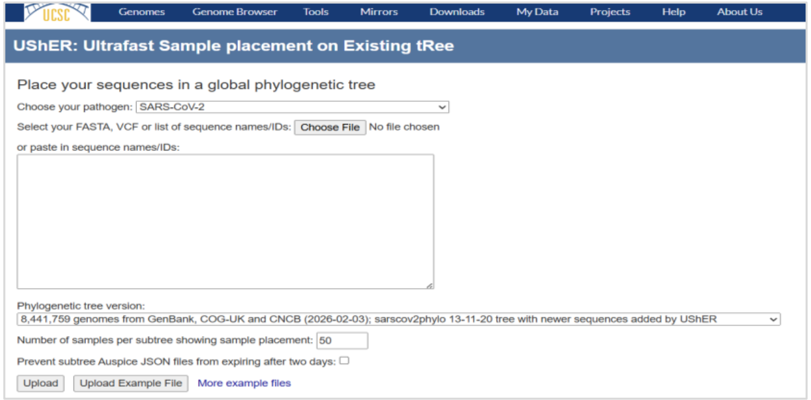
1. By sequencing platform - Oxford Nanopore Technology, illumina, hybrid
2. By licensing - open source, commercial
3. By computational resources - low demand (local machine) to high demand (cloud/server)
4. By operating system - Linux, windows, cross-platform
5. By usage mode - CLI, GUI, cloud

Tools



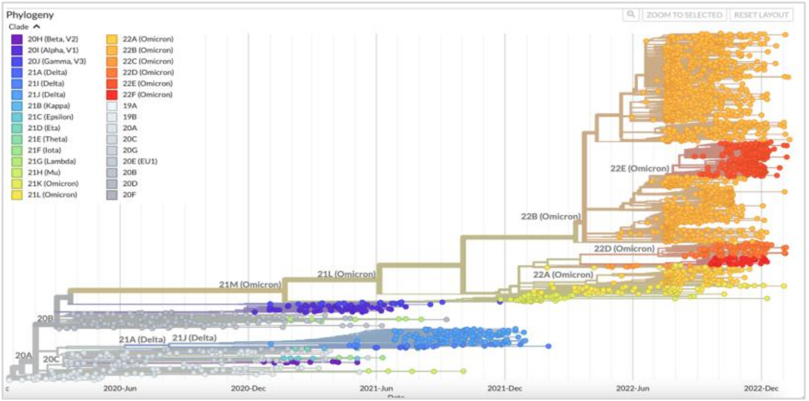
THEIAGENEPI

- Theiagen Genomic Epidemiology Platform
- Cloud-based genomic epidemiology
- Genome Assembly & Mapping
- Typing & Characterization
- AMR & Virulence Analysis
- Phylogenetics



UShER

- Ultrafast Sample placement on Existing tRees
- Rapidly placing new genomes onto an existing large phylogenetic tree (SARS-Cov-2)
- Mutation Mapping
- Outbreak & Transmission Investigation

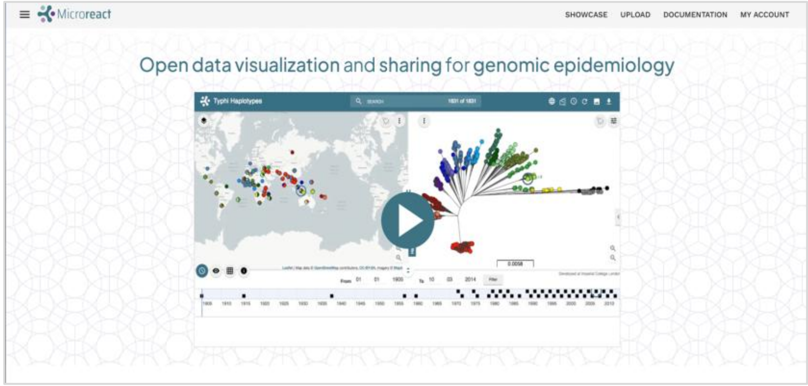


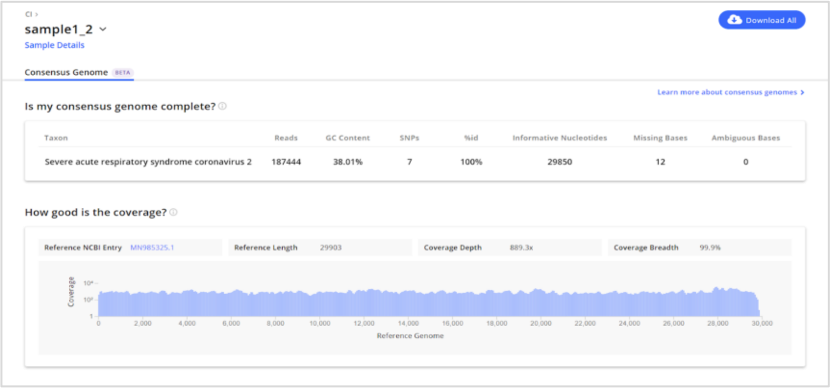
Nextstrain

- Powerful analytics and interactive visualizations.
- Molecular clock estimation
- Evolution rate
- Time to Most Recent Common Ancestor (tMRCA)
- Mutation & Clade Analysis

3 GENOMIC EPIDEMIOLOGY

Tools





Pathogenwatch

- **Genome-based Surveillance & Detailed Characterization**
- Species & taxonomy assignment
- MLST (Multi-Locus Sequence Typing) and cgMLST clustering for strain classification.
- Antimicrobial resistance (AMR) prediction

Microreact

- **Interactive Visualization of Genomic & Epidemiological Data**
- Visualize and explore combined genomic and epidemiological data — especially trees, maps, timelines, and tables
- Map and geographic plots
- **Timeline views**
- Metadata linkage

CZID

- **Metagenomic Detection & Profiling**
- Taxonomic classification
- Viral and microbial detection
- Heatmaps, phylogenetic tree generation, sample comparison

Databases

1. General sequence databases
2. Specialized databases
   - Pathogen
   - Bacteria, AMR
3. Metagenomic and microbiome databases
4. Protein and functional databases

NCBI GenBank

Ensembl ASIA

GISAID

Nextstrain
Real-time tracking of pathogen evolution

GSA Genome Sequence Archive

EMBL European Molecular Biology Laboratory

BV-BRC

UNIVERSITY OF CALIFORNIA SANTA CRUZ Genomics Institute UCSC Genome Browser

DDBJ DNA Data Bank of Japan

GENOMIC EPIDEMIOLOGY OF VIRAL PATHOGENS
24-27 February 2025

Sandy Mak
Senior Scientist
Bioinformatics Institute, A*STAR

Paul Oluniyi
Scientist
Chan Zuckerberg Biohub, USA

Feng Zhu
Senior Researcher
Duke-NUS

Carlo Lipid
Bioinformatics Specialist
Philippine Genome Centre (PGC)

Dr. Lee Soo Cheng
Pakar Perubatan Kesihatan Awam
Makmal Kesihatan Awam Kebangsaan (MKAK)

Dr. Nor Azila Muhammad Azami
Research Fellow
UKM Medical Molecular Biology Institute (UMBI)

PENGETAHUAN DAN KEMAHIRAN YANG BOLEH DIAPLIKASIKAN

Teknik analisa genomik / bioinformatik

- **Pemilihan** tools yang bersesuaian
- Cara **penggunaan tools** dengan betul
- **Pemilihan database** yang bersesuaian
- Mambangunkan **pipeline dan protocol** analisa bioinformatic mengikut keperluan semasa

CABARAN

Modul Latihan yang **padat** dan sesi ceramah yang ringkas

Kepelbagaian bidang – animal / human health, plants, environment, fisheries, wild life

Keupayaan computer untuk menjalankan analisis

Ketersediaan data untuk melengkapkan analisis

CADANGAN PENAMBAHBAIKAN KEPADA DVS

- **Memperkukuh kapasiti kepakaran** bidang bioinformatic
  - Latihan dan pendedahan berterusan kepada kakitangan DVS

- Mewujudkan **pengkalan data asas genomic** bagi isolate tempatan.
  - Pembangunan baseline genomic database bagi pathogen penting
  - Integrasi data genomic dengan data epidemiologi (masa, lokasi, hos dll.)

- **Pemusatan** data genomik
  - Sistem penyimpanan dan pengurusan data yang teratur
  - Perkongsian data genomik untuk penyelidikan genomik lanjutan

- **Mengaktifkan dan memperluaskan** aktiviti sequencing
  - Pelaksanaan WGS dan Analisa lanjutan bagi kes terpilih dan wabak

KESIMPULAN

- Penerapan **teknik analisis sequencing** membolehkan pengenalpastian dan penyiasatan patogen penyakit dilaksanakan secara ***in silico***, sekaligus **menyokong keputusan makmal berasaskan data molekular**.

- Pembangunan **kepakaran tempatan** dalam **analisis sequencing** dapat mengurangkan **kebergantungan kepada sumber dan kepakaran luar**.

- Hasil analisis yang diperoleh dapat meningkatkan kesiapsagaan wabak dan membantu pembuat dasar membuat keputusan yang **berasaskan bukti saintifik**.


thank you