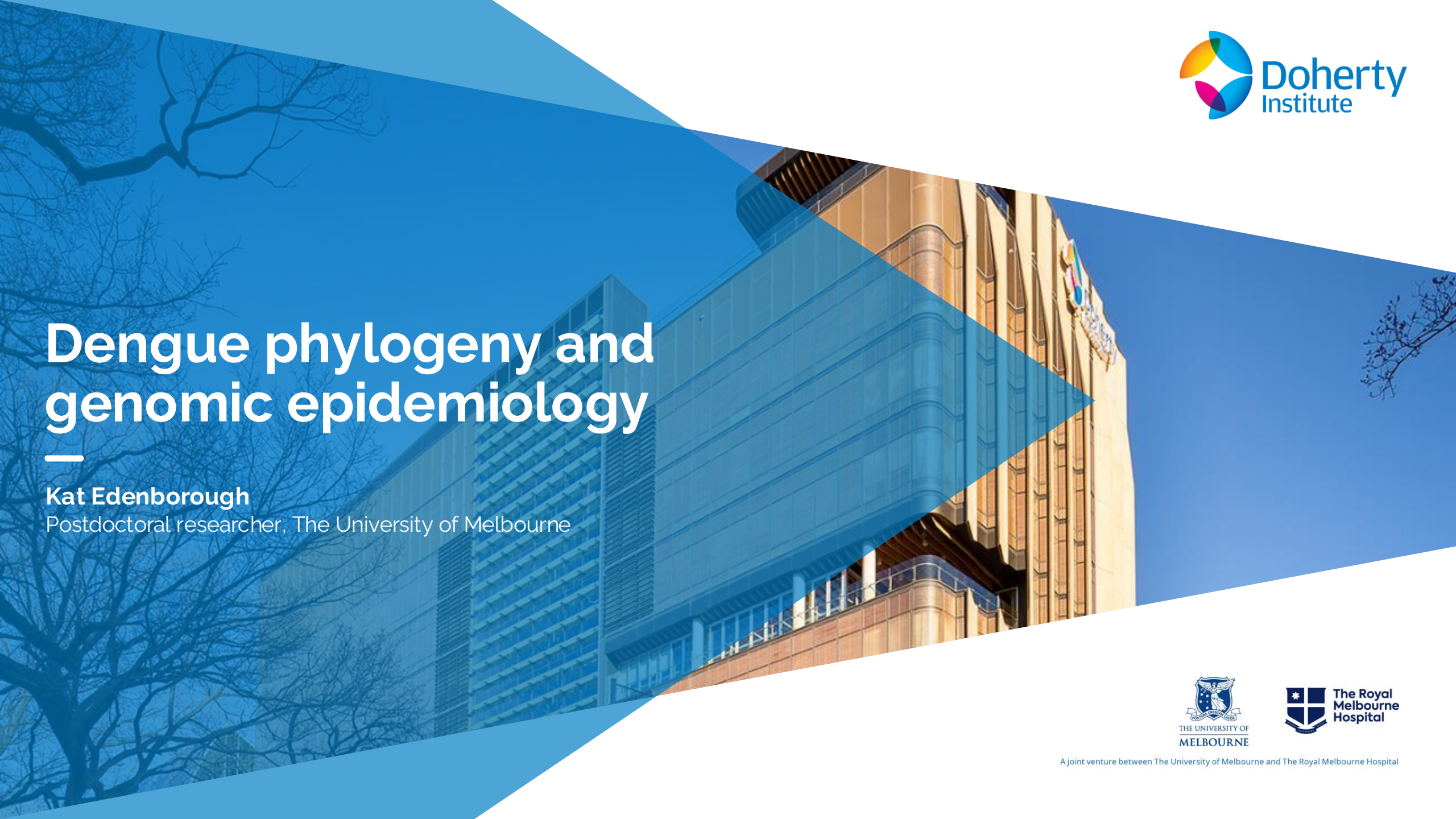


Dengue phylogeny and genomic epidemiology

Kat Edenborough
Postdoctoral researcher, The University of Melbourne

Doherty Institute

THE UNIVERSITY OF MELBOURNE

The Royal Melbourne Hospital

A joint venture between The University of Melbourne and The Royal Melbourne Hospital

Overview

Phylogeny

- Flaviviridae
- Dengue serotypes 1-4
- Genotypes - prior and current classification systems
- Phylogenetic outputs
- Case studies
  - Endemic dengue – Yogyakarta, Indonesia
  - Imported dengue – Queensland, Australia

Genomic epidemiology

- Distinction from phylogenomics
- Application to dengue scenarios
- Hands-on example

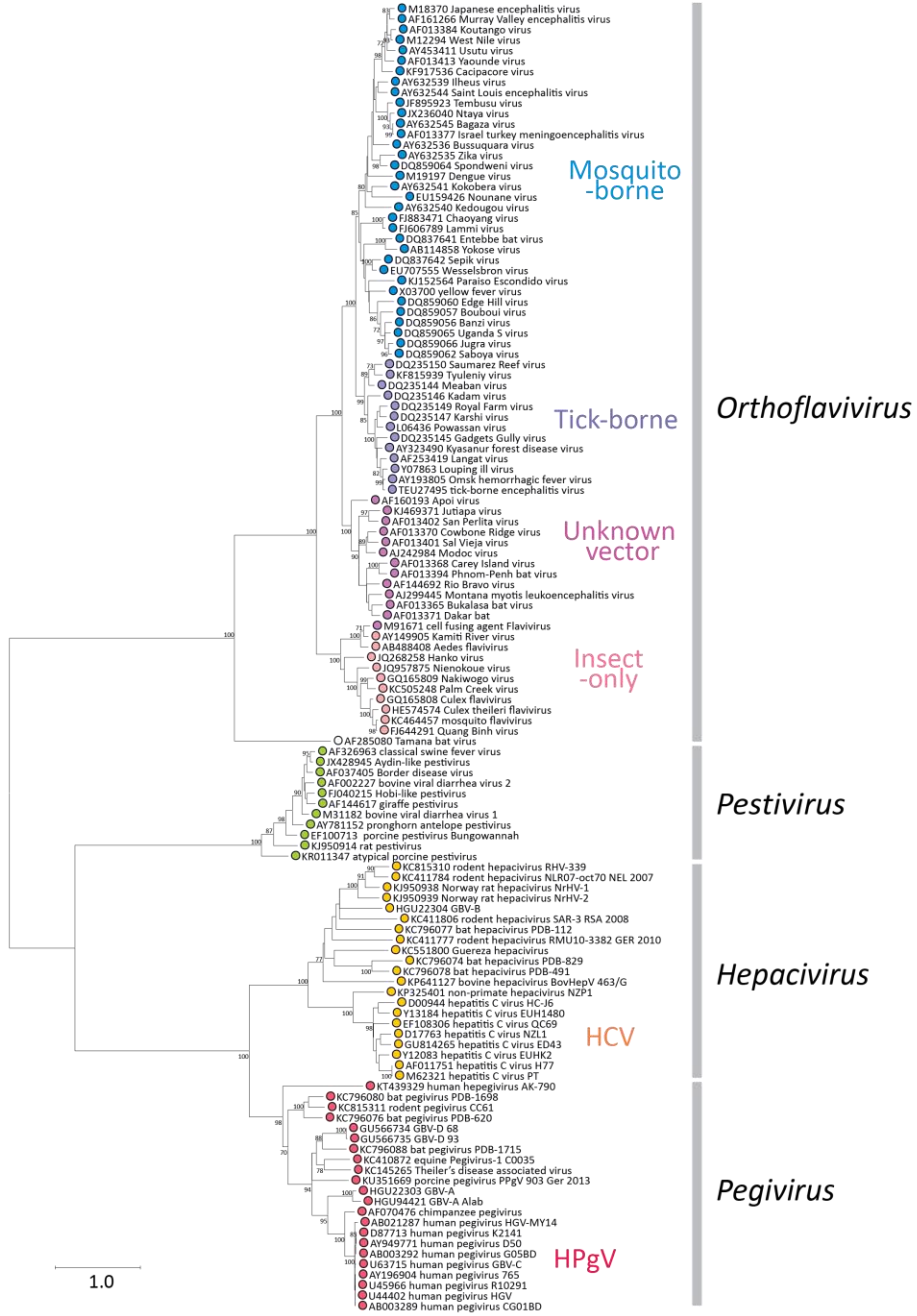
Phylogeny | *Flaviviridae*

Family, genus, species

Dengue viruses situated within the mosquito-borne group of the orthoflavivirus genus

Dengue viruses are close relatives of other medically-important arboviruses

- YFV, JEV, WNV, MVEV



https://ictv.global/report/chapter/flaviviridae/flaviviridae

Phylogeny | DENV-1-4

Phylogenetics has been used to infer the ancestry of DENV-1-4 (E sequences)

- Provided evidence that the four serotypes of dengue emerged in humans independently from the sylvatic cycle
- One tree per serotype is common practice

For DENV, WGS is becoming more commonplace. Small numbers of complete genomes is still a limitation in many analyses.

Whether lack of global representation will affect your analysis will depend on the questions being asked?

- Global trends – genotypes?
- Local trends – major/minor lineages?

One tree might not be able to achieve all your aims

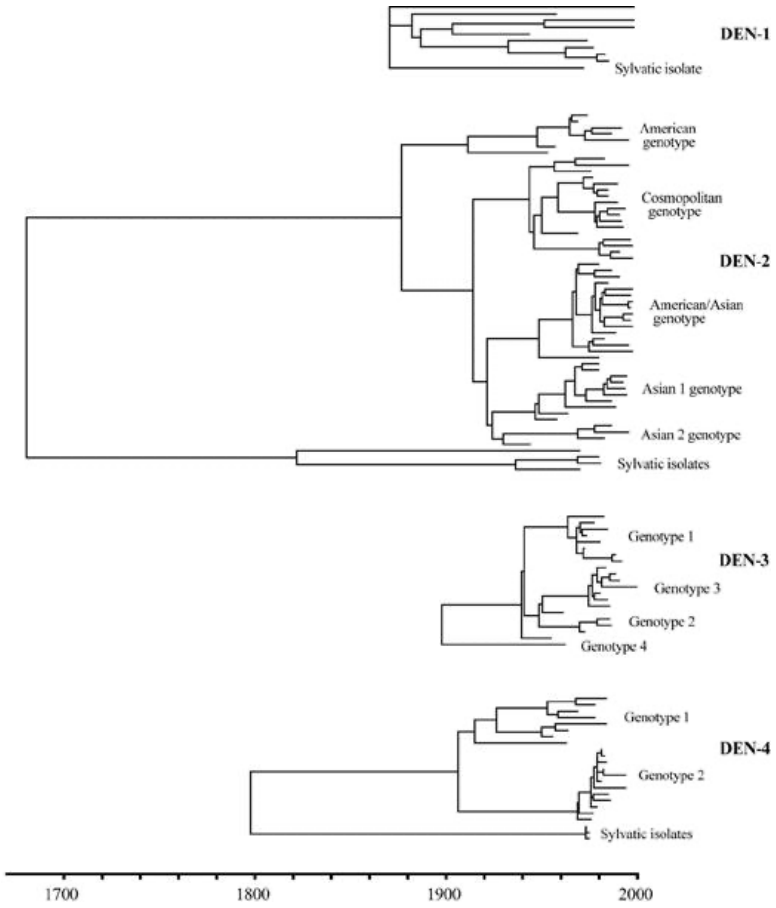
What to include:

Ancestral sequence to root the tree (sylvatic sequence and mid-point root)

Reference that you mapped your reads against (QC)

Your consensus sequence(s) of interest

To provide global/local context, sequences obtained from a repository (NCBI, viruses). Filter to include only whole genomes and exclude E/NS1 or other partial sequences.



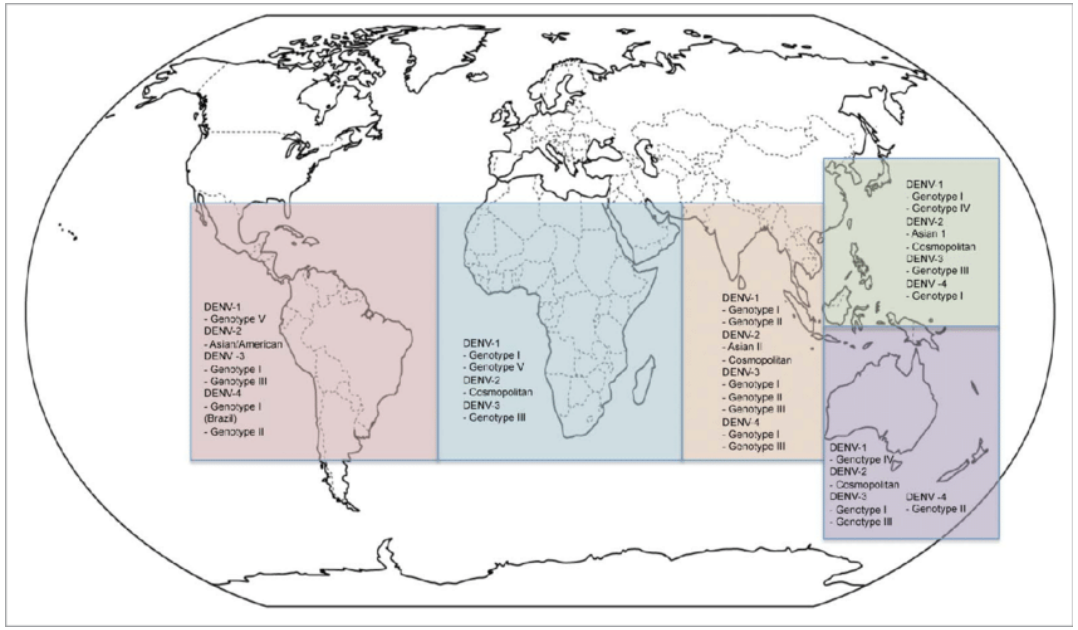
The origin, emergence and evolutionary dynamics of dengue virus. Holmes and Twiddy. Infection genetics and evolution. 2003

Phylogeny | Previous classification - genotype

Dengue genotypes were named in the late 1990s and early 2000s based on partial sequences

Later based on genetic distance in E sequence (>6% sequence divergence)

Named by initial geographical distribution, which has since changed for some genotypes



The relevance of dengue virus genotype surveillance at country level before vaccine approval. Usme-Ciro et al.. Human vaccines and immunotherapies. 2014

Phylogeny | New system

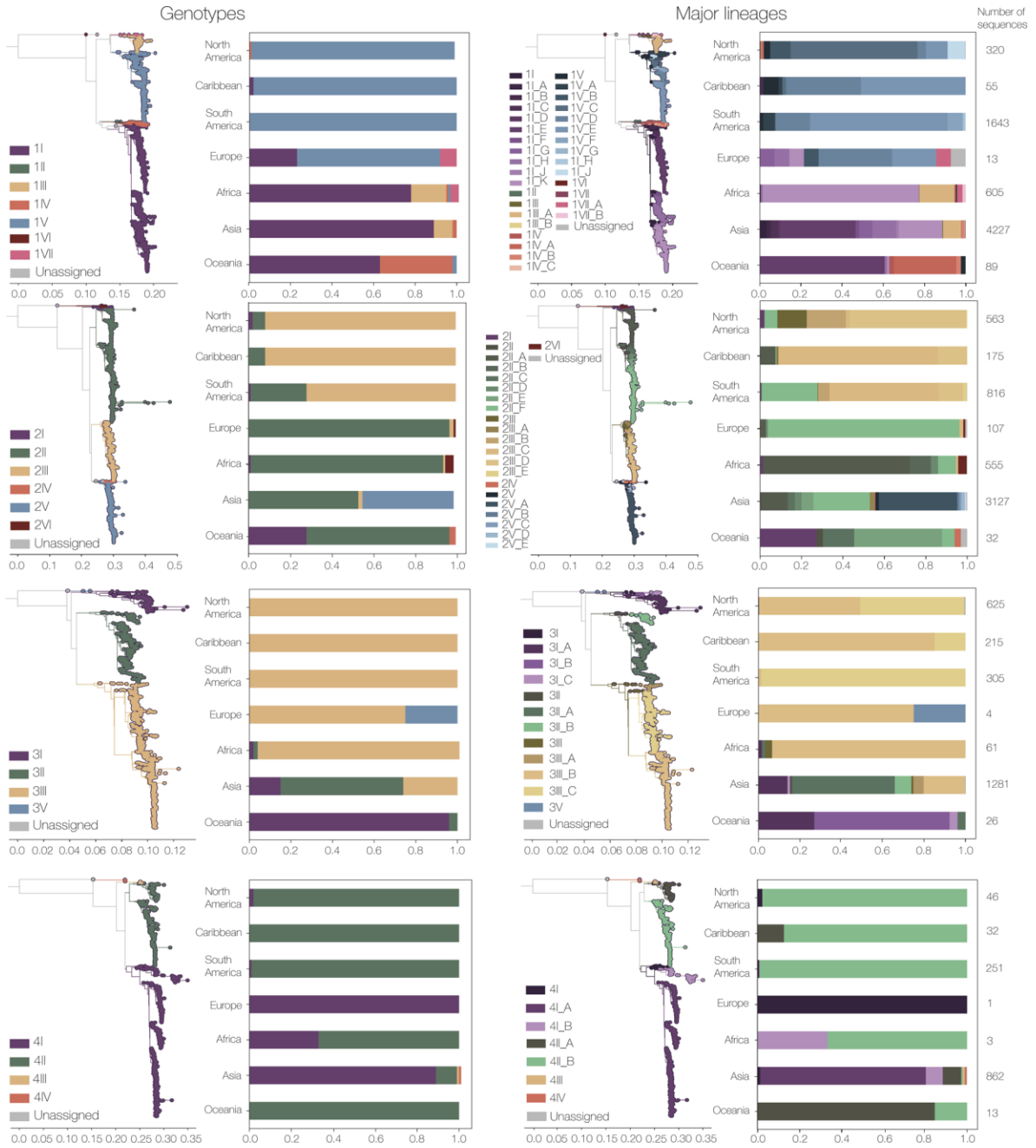
Enables diversity to be captured at finer resolution – genotype and lineage

Genotypes that are regionally distributed

- i.e. **1V** predominantly sampled in the Americas and not elsewhere
- 1V has diverged into 7 major lineages in the Americas which have distinct local circulation patterns

Genotypes that are globally distributed

- i.e. **2II** across America, Europe, Africa and Asia
- 2II has diverged into 6 major lineages and those circulating in Europe are different from those in Africa



A new lineage nomenclature to aid genomic surveillance of dengue virus. Hill et al. PlosBio, 2024

Phylogeny | Outputs for DENV-1-4

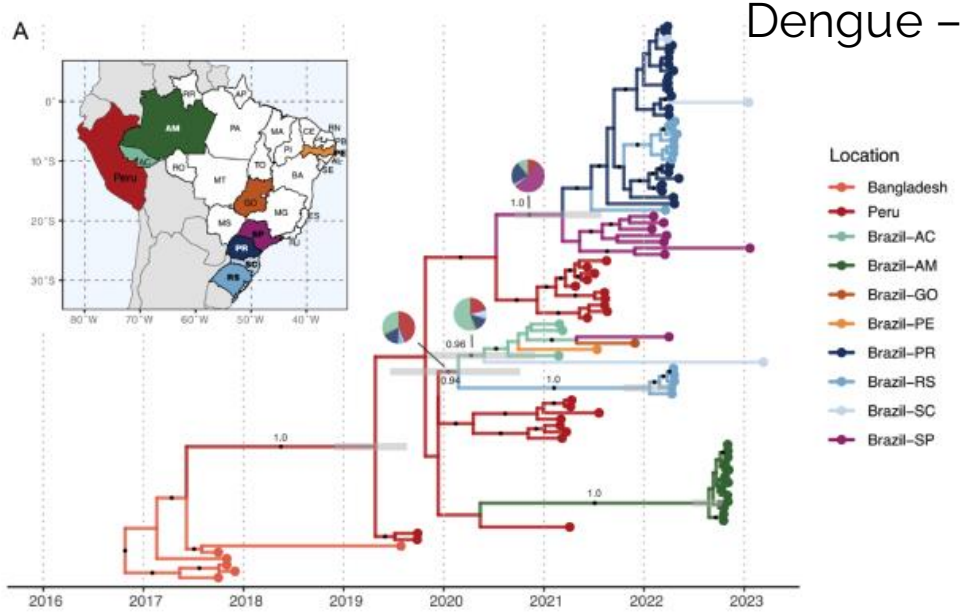
Descriptive outputs:

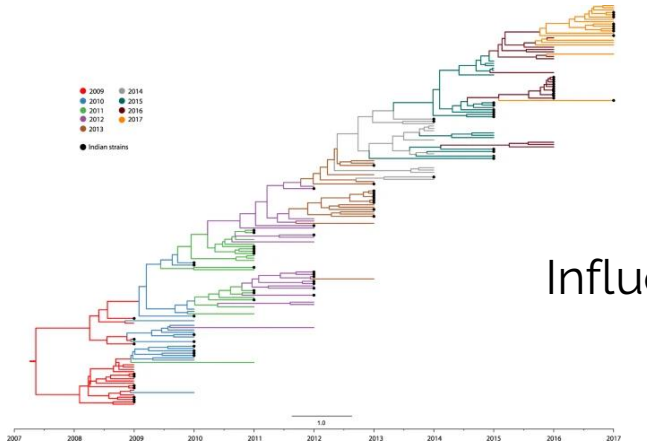
- Define tree structure/topology:
  - DENV has an irregular spatiotemporal structure
  - Influenza has a regular spatiotemporal structure
- Identify key events: Lineage replacement, introduction of genotypes/serotypes

Numerical outputs:

- Clock rates
  - DENV: ~4 – 9x10^{-4} substitutions/site/year
  - Clock rates are slower relative to respiratory viruses due to the presence of population bottlenecks and fitness tradeoffs in the transmission cycle
- Sequence distribution among cases to identify predominant genotypes and lineages

Dengue – 2II





Influenza- H1 HA

Multiple introductions and country-wide spread of DENV-2 genotype II (Cosmopolitan) in Brazil.. Graf et al. Virus Evo, 2023

Evolutionary, genetic, structural characterization and its functional implications for the influenza A (H1N1) infection outbreak in India from 2009 to 2017. Scientific reports. 2019

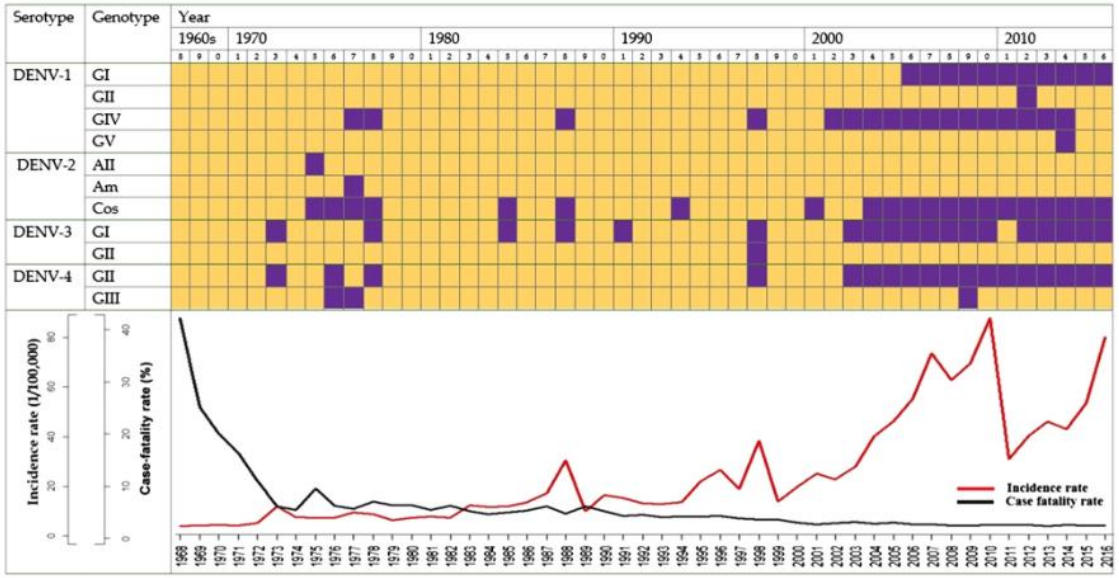
Country case studies - Indonesia

Significant dengue burden

- At least 100,000 cases/year (variable)
- Increasing incidence over the past 50 years
- Peaks in rainy season (December-April) due to mosquito abundance
- Most individuals infected in childhood
  - Seroconversion 50% of 5 yo
  - Seroconversion 90% of 18 yo

Hyperendemicity

- Multiple serotypes in circulation since the 1970s
- Prior to 2000, the predominant virus in circulation was genotype 3I
- In the early 2000s, DENV-3 was replaced with DENV-1, DENV-2 and DENV-4.
- Replacement was associated with significant increase in incidence and number of severe dengue cases.



Dengue viruses circulating in Indonesia: A systemic review and phylogenetic analysis of data from five decades. Harapan et al. Rev Med Virol, 2019

Yogyakarta, Indonesia

Typical trends in endemic regions

- Sequences cluster into clades
- Clades form major/minor lineages
- Multiple lineages might be detected within a genotype (2II – Cosmopolitan, lineage F, D and C)
- Multiple genotypes might be detected within a serotype
- Usually, a predominance of one or two serotypes (DENV-2 and DENV-4)
- Low frequency of coinfection with multiple serotypes (~2%)

ML phylogenies of dengue genomes circulating in Yogyakarta, 2018 – 2020, shown in red and global sequences shown in black

Dengue virus genomic surveillance in the AWED trial reveals genotypic efficacy and disruption of focal transmission. Edenborough et al. Scientific Reports. 2024

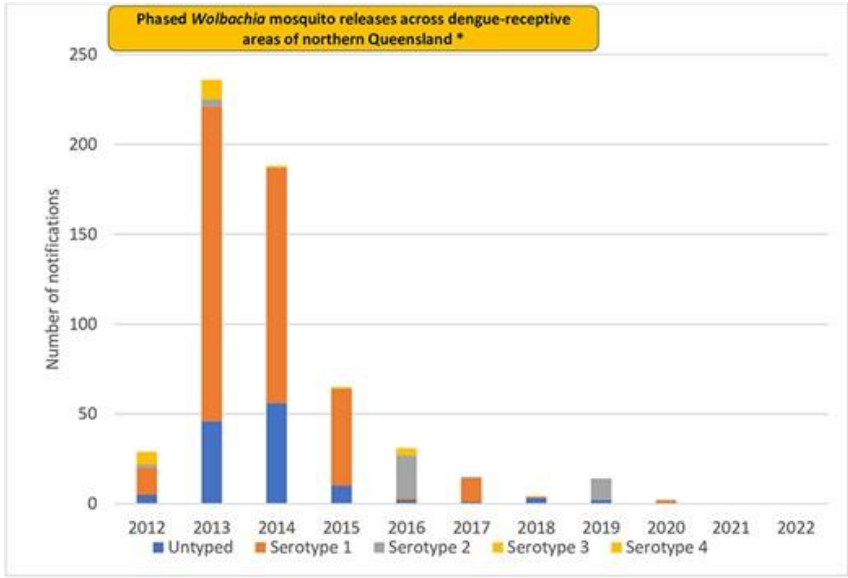
Country case studies - Australia

Relatively low dengue burden

- ~ 1000 cases/year are mostly imported
- Young adults are the most frequently affected
- Before 2020, importations led to local outbreaks in areas with competent vectors (Qld)
- Local transmission has ceased due to control efforts – biocontrol

Not endemic, outbreaks and cases caused by importation from travel

- All four serotypes detected – DENV-1 and DENV-2 predominant
- Serotype dominance mirrors that in neighbouring countries in Asia and Oceania – Indonesia, Thailand and India
- Clustering indicates local transmission and might trigger additional interventions

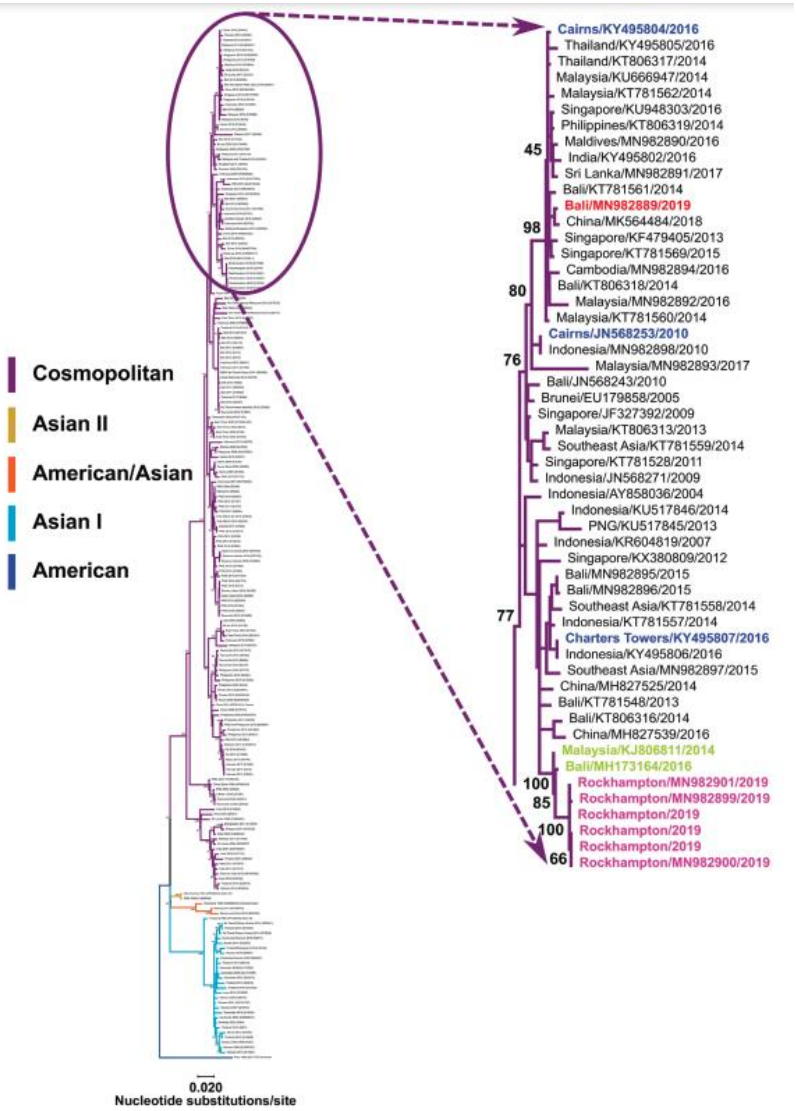


The epidemiology of imported and locally acquired dengue in Australia, 2012-2022, Sohail et al. J Travel Med, 2024

Queensland (Qld), Australia

Typical trends

- Sequences appear as singletons if imported
- Clustering would indicate local transmission
- Lineages/genotypes/serotypes detected will mirror what is predominant and circulating in the region



ML phylogeny of DENV-2 E sequences

Blue – prior outbreaks in QLD

Pink sequences – QLD 2019 outbreak

Re-emergence of dengue virus in regional Queensland: 2019 dengue virus outbreak in Rockhampton, Central Queensland, Australia. Walker et al. Commun Dis Intell. 2018.

Phylogenomics and genomic epidemiology

Phylogenomics

- characterises genomic diversity among cases
- describes evolutionary trends and relationships between sequences from different cases
- commonly used method in academic research

Includes epidemiological data

Genomic epidemiology

- is usually quantitative
- interprets the sequences given an epidemiological context
- conclusions speak to public health questions and usually contain caveats
- requires metadata, which is highly variable in detail between studies
- accuracy is important due to health and legal implications

DENV molecular resolution should align with epidemiological investigation purpose

Classification level	Method	Line of investigation	Epidemiological metadata required
Serotype	Typing PCR	• Investigate seasonal changes in serotype	Yearly case numbers stratified by type
Genotype	WGS or E sequencing	• Investigate seasonal changes in genotype and identify associations with clinical severity, demographic factors and regional trends • Assess if genotypes are regionally novel or being replaced	Clinical information, sex, age, region of residence, and yearly case numbers stratified by genotype
Major and minor lineage	WGS	• Investigate lineage circulation patterns in the regional/global context • Differentiate between imported and endemic lineages (context dependent) • Map fine-scale transmission dynamics • Assess if lineages are regionally novel or being replaced	Clinical information, sex, age, region of residence, travel history, date of symptom onset
Cluster level	WGS	• Link cases based on spatiotemporal proximity and genomic relatedness • Map fine-scale transmission dynamics and phylogenetic clusters across time and place	Clinical information, sex, age, GPS coordinates of residential address, travel history, date of symptom onset

Genomic epidemiology | Hands-on example

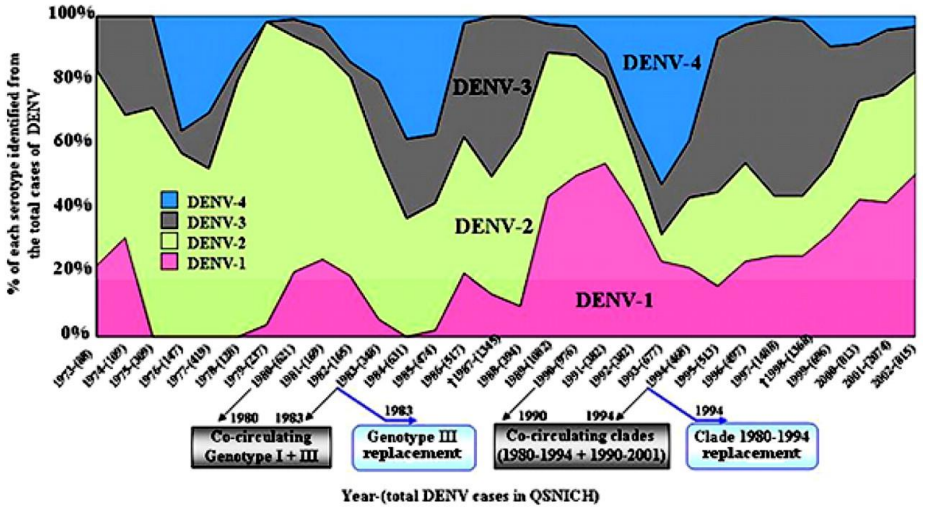
Serotype, genotype and lineage replacement events are commonplace in countries with endemic dengue

Problematic because replacement:

- Increases secondary infection risk. Heterotypic, secondary infection is the greatest predictor for severe dengue outcomes
  - Antibody dependent enhancement
- Specific genotypes might be be better adapted/transmissible in the local *Aedes* vector
- Less serotype-specific immunity existing in the population and a larger susceptible population

Overall, is associated with increased incidence of severe dengue

Dengue cases in Bangkok, Thailand



Clade Replacements in Dengue Virus Serotypes 1 and 3 are associated with changing serotype prevalence. Journal of Virology, 2005

Genomic epidemiology | Hands-on example

Epidemiological context: A regional public health department has noticed increasing trends in severe dengue cases requiring hospitalisation in 2025. The majority of cases admitted with severe dengue so far in 2025 have been diagnosed with DENV-4. Historically, dengue cases have been detected in the region most years within the rainy season and DENV-4 genotype II has been predominantly detected through surveillance in the past 4 years.

Sampling: Blood samples were collected from individuals with febrile illness through hospital and clinic-based arbovirus surveillance programs. Samples collected from laboratory-confirmed dengue cases were subjected to whole genome sequencing. DENV-4 genome sequences were obtained from 88 dengue cases in 2025. The 88 sequences were aligned with an additional 9 genome sequences, sampled from dengue cases in the same region in 2024, and a maximum-likelihood phylogenomic tree was constructed.

Question: You are in a team of genomic epidemiologists and want to understand if the increase in severe dengue is associated with introduction of a novel DENV-4 genotype or lineage circulating in the region. Work through the following steps and questions to assist you with the investigation.

Genomic epidemiology | Hands-on example

Step 1. Drag and drop the phylogenetic tree (DV4.newick) into https://auspice.us/

Step 2. Examine the tree without parsing the metadata and orientate yourself to the different parts.

Q1. What do the tips represent?

Q2. What do the nodes represent?

Q3. What unit is divergence expressed in and what does it mean?

Step 3. Drag in the csv file containing the metadata. Colour the tree by assigned lineages.

Q4. How many genotypes were detected?

Step 4. Colour the tree by year.

Q5. Determine which genotypes were predominant and in which year(s)?

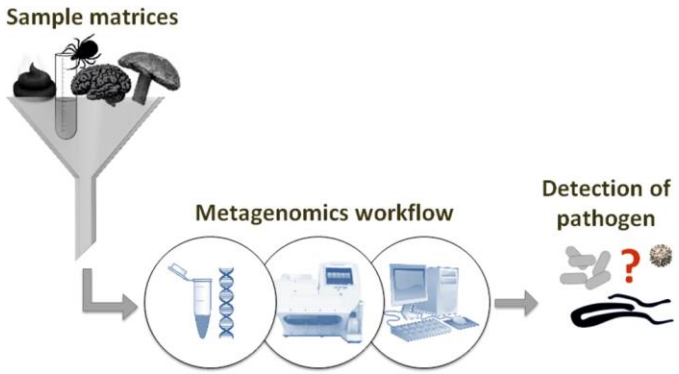
Q6. Could any of the genotypes observed be newly introduced? If yes, what caveats might be required when presenting this information?

Q7. What additional epidemiological data would be useful to investigate associations between genotype and severe disease outcomes?

The future of public health genomics: from swabs to air samplers

Viral metagenomics: pathogen discovery

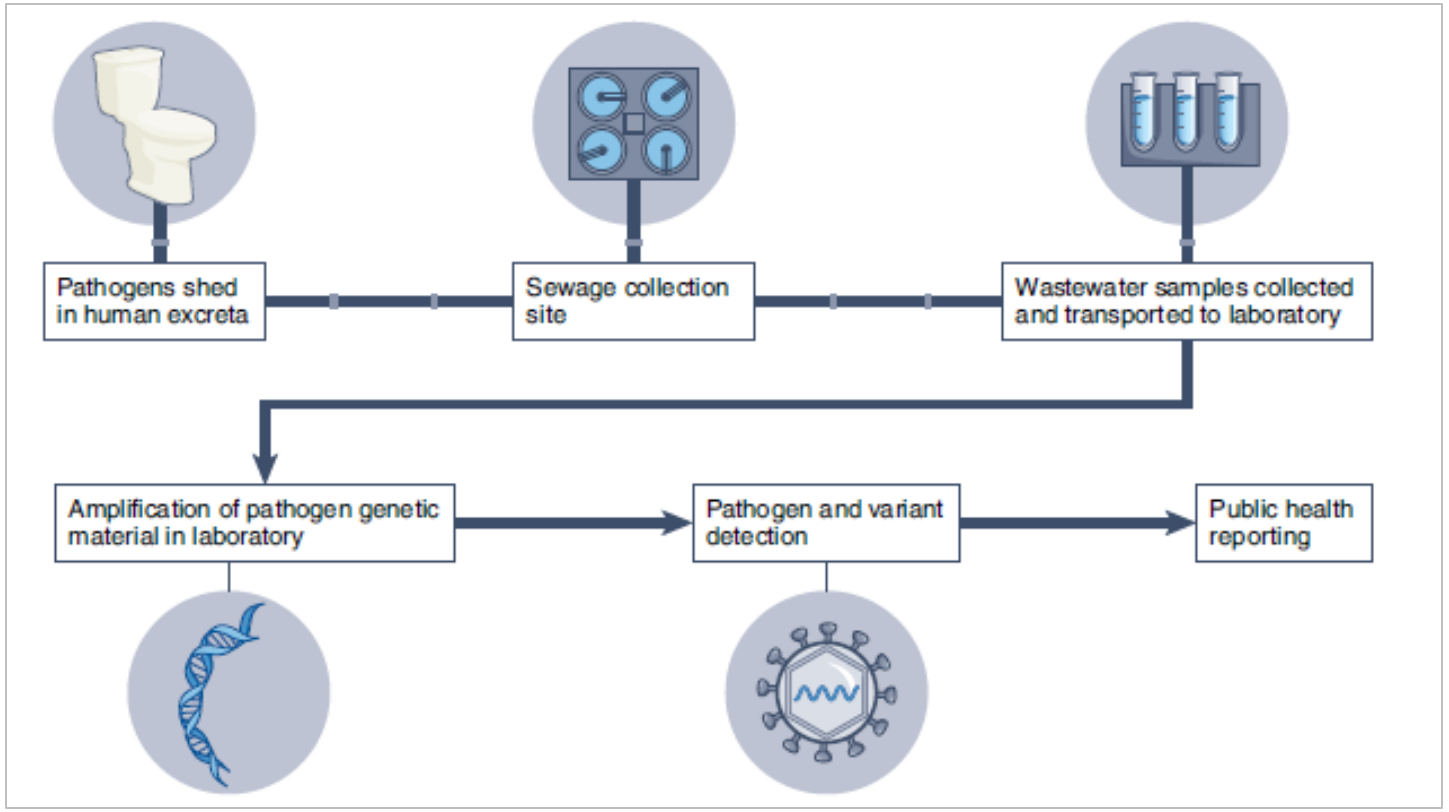
- Conventional public health genomics has traditionally been used to generate a complete genome of a specific infecting viral pathogen
- Metagenomics enables pathogen discovery by enabling assessment of an entire virus community within a given patient sample
  - Identification of an emergent or distantly related viral pathogen
  - Clinical management and patient care: precision medicine



https://www.nature.com/articles/s41598-018-31496-1

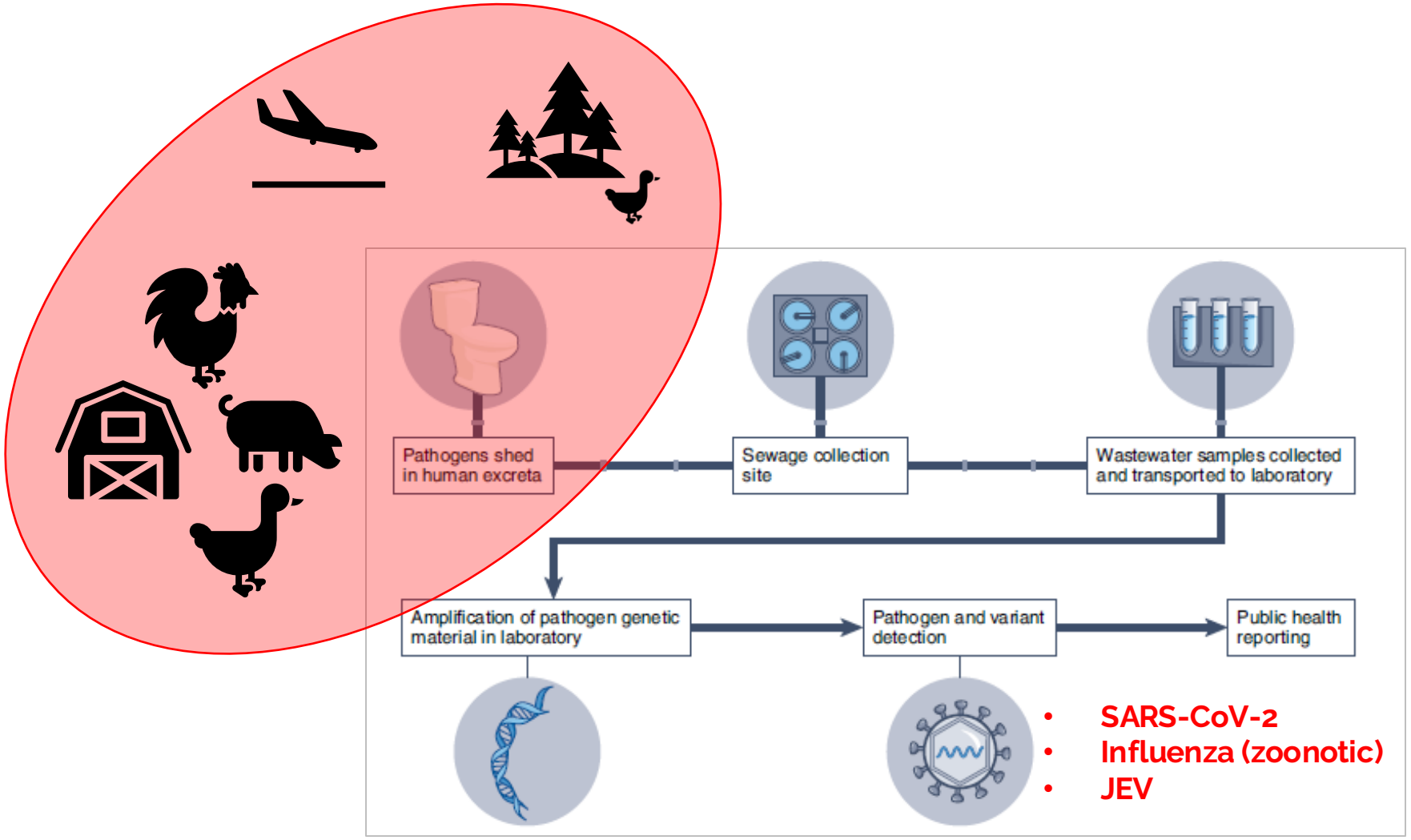
Wastewater genomics: pathogen spread

- Prior to the pandemic, primarily used only to monitor spread of polio virus
- Very important surveillance tool for the eradication of wild polio virus



https://www.nature.com/articles/s41591-022-01940-x

Wastewater genomics: pathogen spread



https://www.nature.com/articles/s41591-022-01940-x

Air samplers: pathogen diversity

- Emergent and highly sensitive form of environmental sampling
- Difficult to accurately capture complexity of virus populations in complex settings:
  - Wet markets
  - Bat caves (pictured)
  - Food production sites: farms/abbotoirs
- Directly sample breadth and complexity of virus populations at hot spot sites to identify emergent viral pathogens

https://www.science.org/content/article/armed-air-samplers-rope-tricks-and-yes-ants-virus-hunters-spot-threats-new-ways

Air samplers: pathogen diversity

- Novel application in translational public health genomics
- Samplers can be attached to remote controlled vehicles (drones/cars) to enable access to remote/inaccessible sites
- Airborne virus particles are captured within the device which can be taken back to a lab for processing
- Replaces need to constantly collect and analyse samples from multiple animals/environmental sites to scan for emergent pathogens
- New tool in pandemic toolkit

Conclusions

- Technologies and applications continue to evolve to increase the timeliness and utility of genomic data obtained
- Epidemiological context is essential to enable rapid risk assessment of genomic signals identified
- Move away from testing individual samples to applying novel methods including wastewater and air sampling for representative sampling and early warning/detection