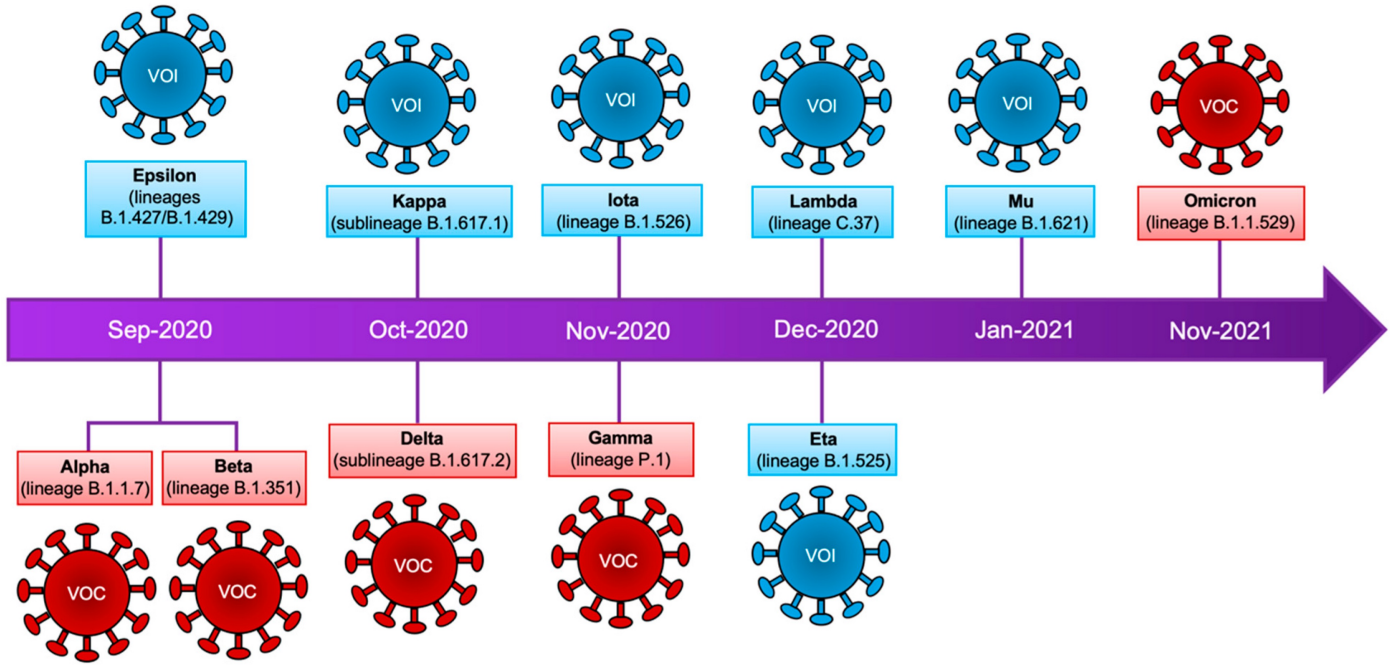
University of Glasgow
Virus evolution and molecular clocks
WORLD CHANGING GLASGOW
A WORLD TOP 100 UNIVERSITY

# Virus evolution

- Changes in the genetic sequences of viruses over time
- Can occur through various mechanisms (e.g. mutation & recombination), leading to the emergence of new virus variants
- Why should we care?



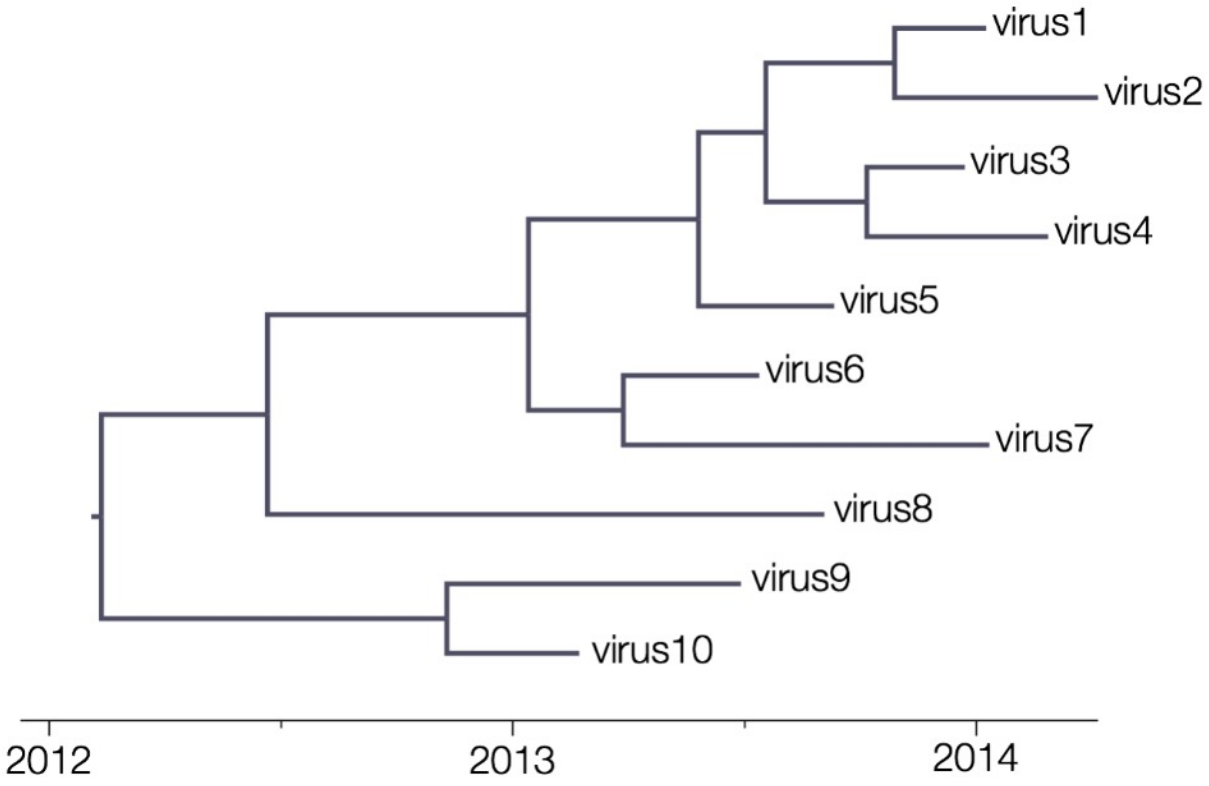

*Flores-Vega et al., 2022*

# Molecular clocks

- Molecular sequence data from rapidly evolving organisms are often sampled at different points in time
- Sampling times can then be used for molecular clock calibration
- The root-to-tip (RTT) regression is an essential tool to assess the degree to which the data behave in a clock-like fashion
- Types of molecular clocks:
  - ✓Strict vs relaxed
  - ✓Local vs global

# The strict molecular clock

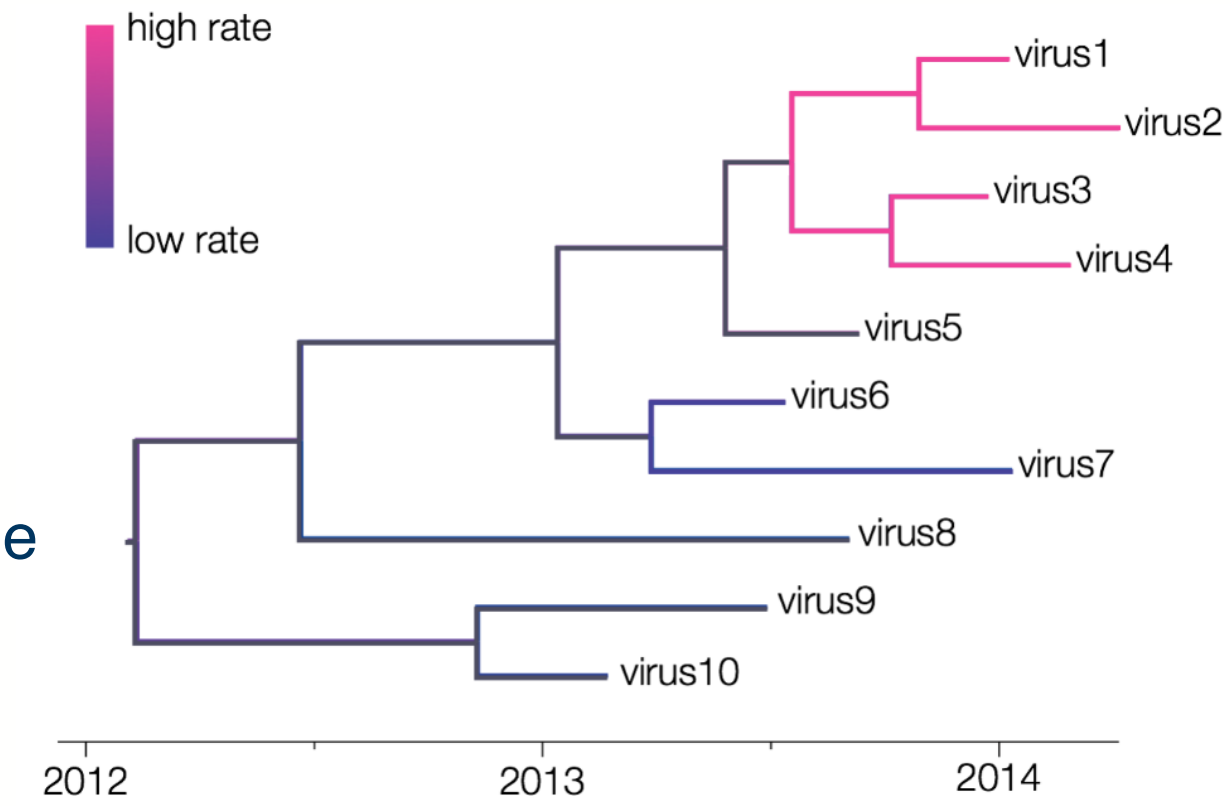
- A strict molecular clock
  - ✓The simplest molecular clock model
  - ✓Assumes a constant evolutionary rate (rate of substitution per unit time)
- When the evolutionary rate is constant throughout a phylogenetic tree, the term ***global molecular clock*** is used




*Under the strict clock model, evolution occurs at the same rate at every branch.*

# The strict molecular clock

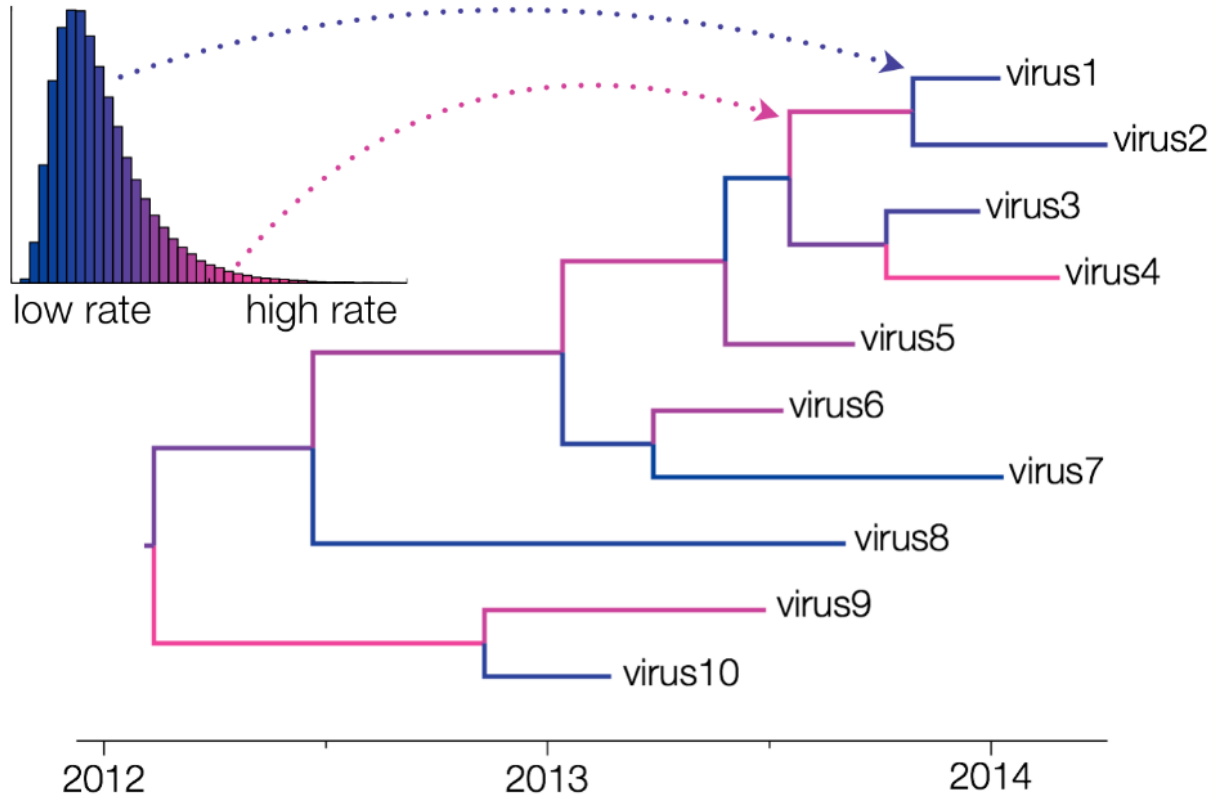
- ***A strict local clock*** - different substitution rates apply to different monophyletic groups within a tree
  - ✓sampling from different host populations, host species, or pathogen lineages
- Larger datasets are more likely to sample from distinct populations, making local clocks increasingly important




*Under the fixed local clock model, a shift in evolutionary rate occurs at the TMRCA of each defined taxon set.*

# The relaxed molecular clock

- Multivariate statistical approach i.e. a different rate for each branch
- Computationally demanding

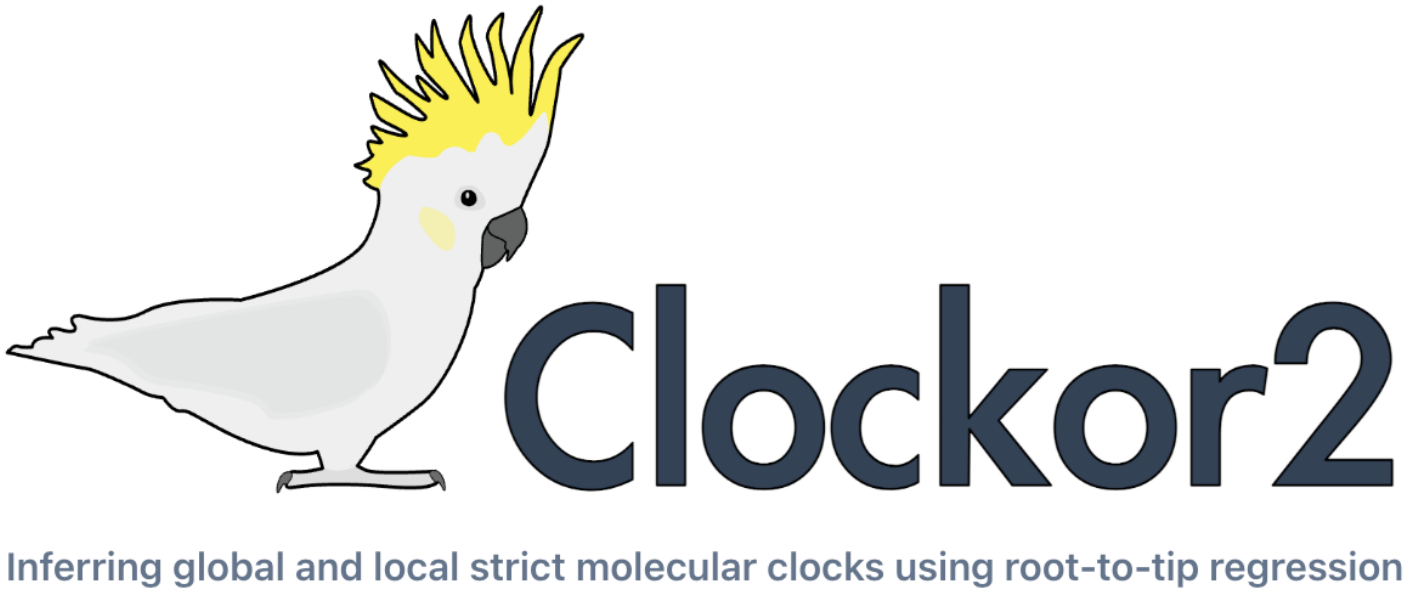



*Under the uncorrelated relaxed clock model, the evolutionary rate at each branch is drawn from an underlying distribution.*

University of Glasgow

# Clockor2

- A scalable and accessible client-side web application for exploring the fit of local clocks, with results available in seconds to minutes
- Enables rapid inference of global and local strict molecular clocks from phylogenetic trees using root-to-tip regression (RTT)




*clockor2.github.io*

University of Glasgow

# Input

- A newick tree
- Sampling dates & group identifiers (parsed from tip labels or separate files on input)

University of Glasgow

# The Best Fitting Root

- Based on the $R^2$ or residual mean square (RMS) of the RTT regression of a global clock model for the input tree
- The root position along the branch leading to the highest $R^2$ or RMS is selected
- The best-fitting root is inferred using a single global clock because it presents the most parsimonious model of the evolutionary rate for a given tree
- Clockor2 does not find the best-fitting root for local clock models because the search space of best-fitting roots and local clock configurations quickly becomes prohibitive and is possibly unidentifiable

# Testing the fit of a clock

- Local clock and or global clock configurations can also be compared using an information criterion
  - ✓Bayesian Information Criterion (BIC)
  - ✓Akaike Information Criterion (AIC)
  - ✓corrected Akaike Information Criterion (AICc)

**Local Clock** AICc **-585.68** AIC **-592.32** BIC **-567.18**

| CLOCK | #TIPS | RATE | $R^2$ | RESIDUAL MEAN SQUARED | INTERCEPT |
|---|---|---|---|---|---|
| Bauchi | 14 | 5.003 x 10^-4 | 0.913 | 8.234 x 10^-7 | 1945.074 |

University of Glasgow
Let’s put Clockor2 to work!
#UofGWorldChangers
@UofGlasgow