Phylodynamics of infectious diseases

Recent advances

Luiz Max Carvalho

THE UNIVERSITY
of EDINBURGH

Acknowledgements

Andrew Rambaut
University of
Edinburgh

Luiz Ricardo
UFRJ

Outline

What is phylodynamics?

Concepts and tools.

Outline

What is phylodynamics?

Concepts and tools.

Examples

Influenza, West Nile Virus, Ebola. In space* and time.

Outline

What is phylodynamics?

Concepts and tools.

Examples

Influenza, West Nile Virus, Ebola. In space* and time.

Methodological advances

Lab in a suitcase, GPUs and loads of maths.

Outline

What is phylodynamics?

Concepts and tools.

Examples

Influenza, West Nile Virus, Ebola. In space* and time.

Methodological advances

Lab in a suitcase, GPUs and loads of maths.

Where are we headed?

We should prepare for an era of plenty.

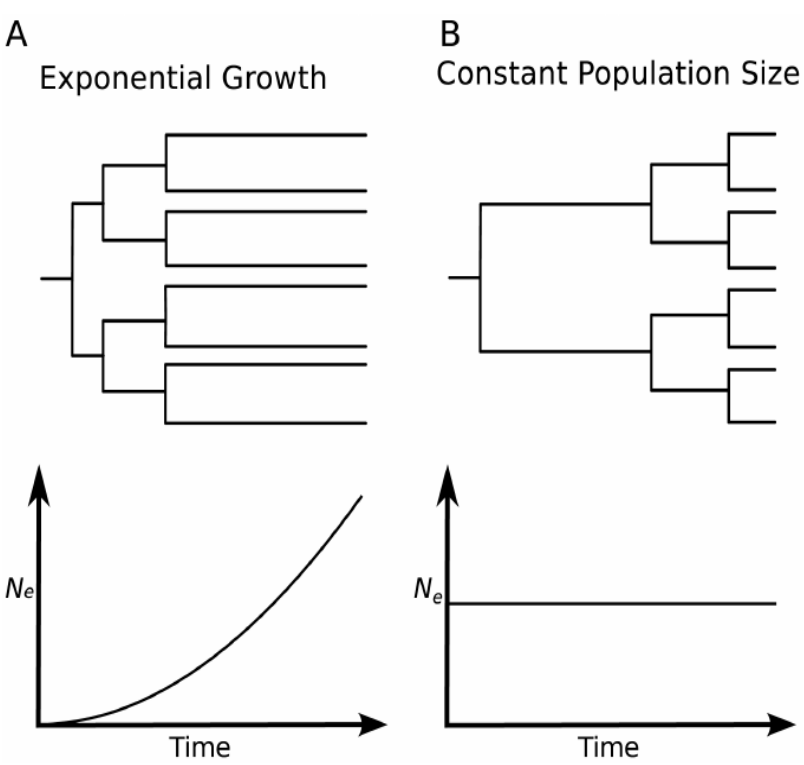
Motivation

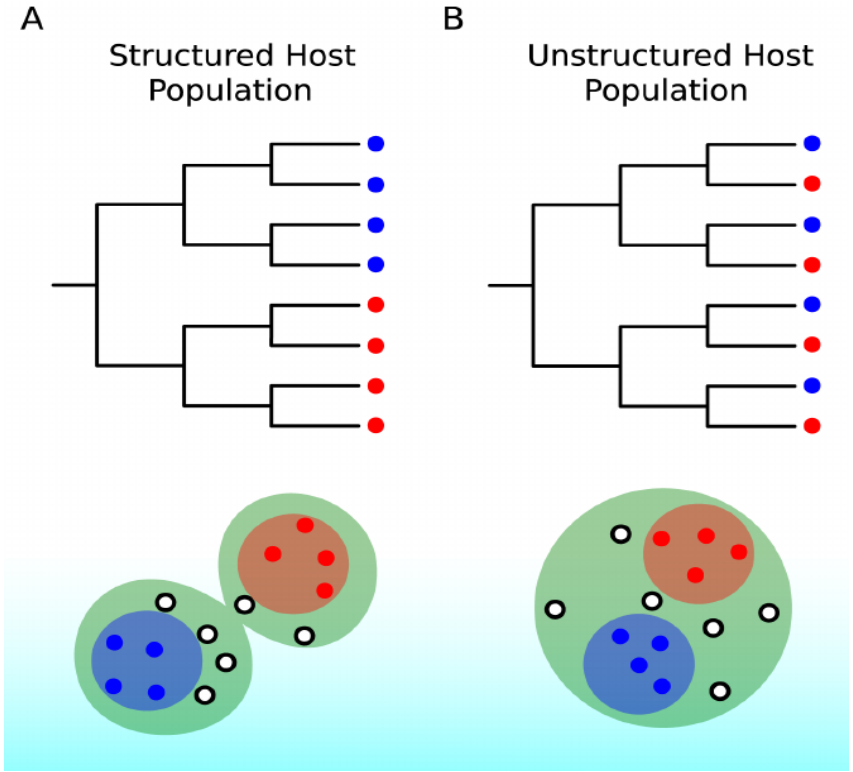
Phylodynamics of fast-evolving pathogens

Inferring spatial and temporal dynamics from genomic data:

Phylogenies*!

* plus complicated models





Trees and the coalescent

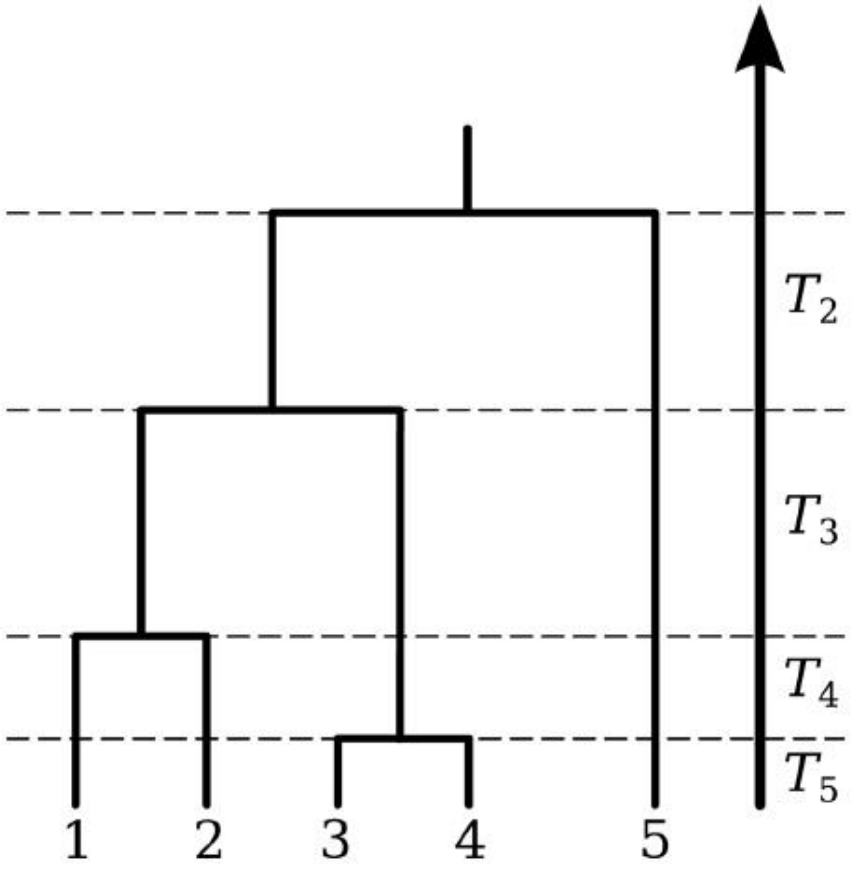


Figure: Figure 4 from Volz et al. (2013).

Let T_n denote the time for n lineages to *coalesce*, i.e., merge into one ancestral lineage, in a population of size N_e. Then:

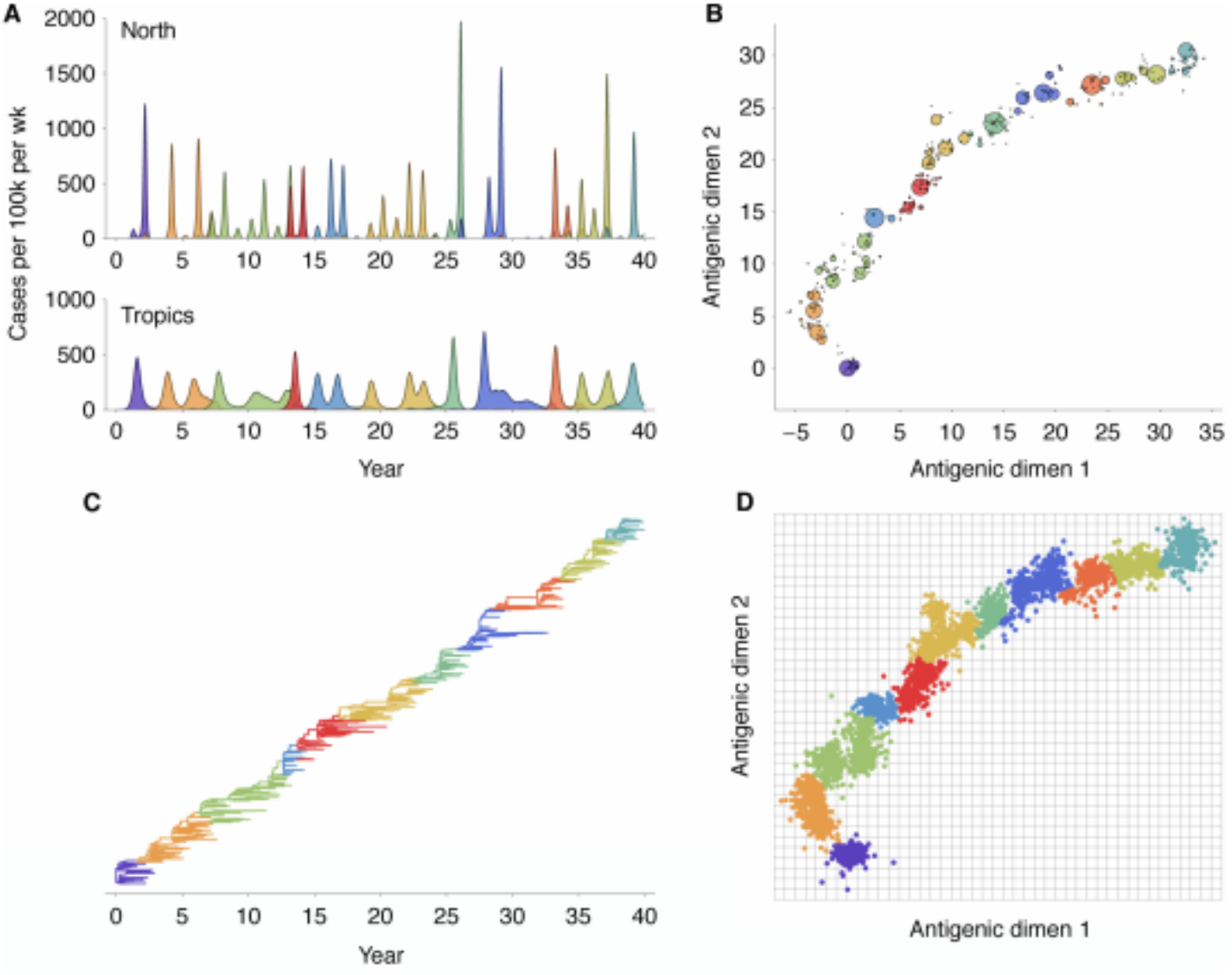
$$Pr(T_n = t) = \lambda_n e^{-\lambda_n t}$$

$$\lambda_n = \binom{n}{2} \frac{1}{N_e} = \binom{n}{2} \frac{1}{N_e \tau}$$

where N_e is the effective population size and τ is the generation time. Let T_{mrca} denote the age of the most recent common ancestor:

$$\begin{aligned} \mathbb{E}[T_{\text{mrca}}] &= \mathbb{E}[T_n] + \mathbb{E}[T_{n-1}] + \ldots + \mathbb{E}[T_2] \\ &= 1/\lambda_n + 1/\lambda_{n-1} + \ldots + 1/\lambda_2 \\ &= 2N_e \left(1 - \frac{1}{n}\right) \end{aligned}$$

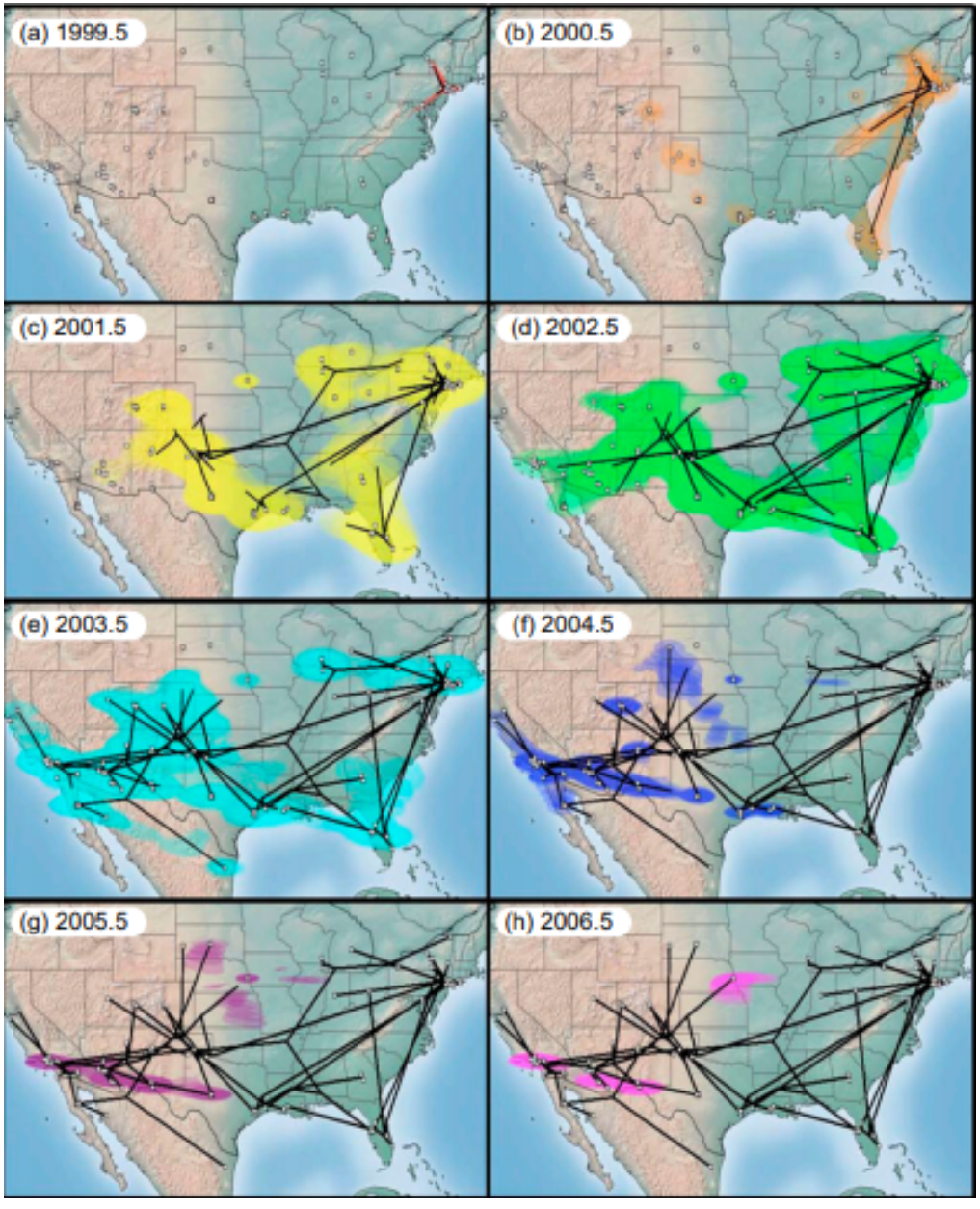
Example I: Antigenic evolution in Influenza H3N2



Antigenic evolution in Influenza H3N2: details

Antigenic evolution in Influenza H3N2: findings

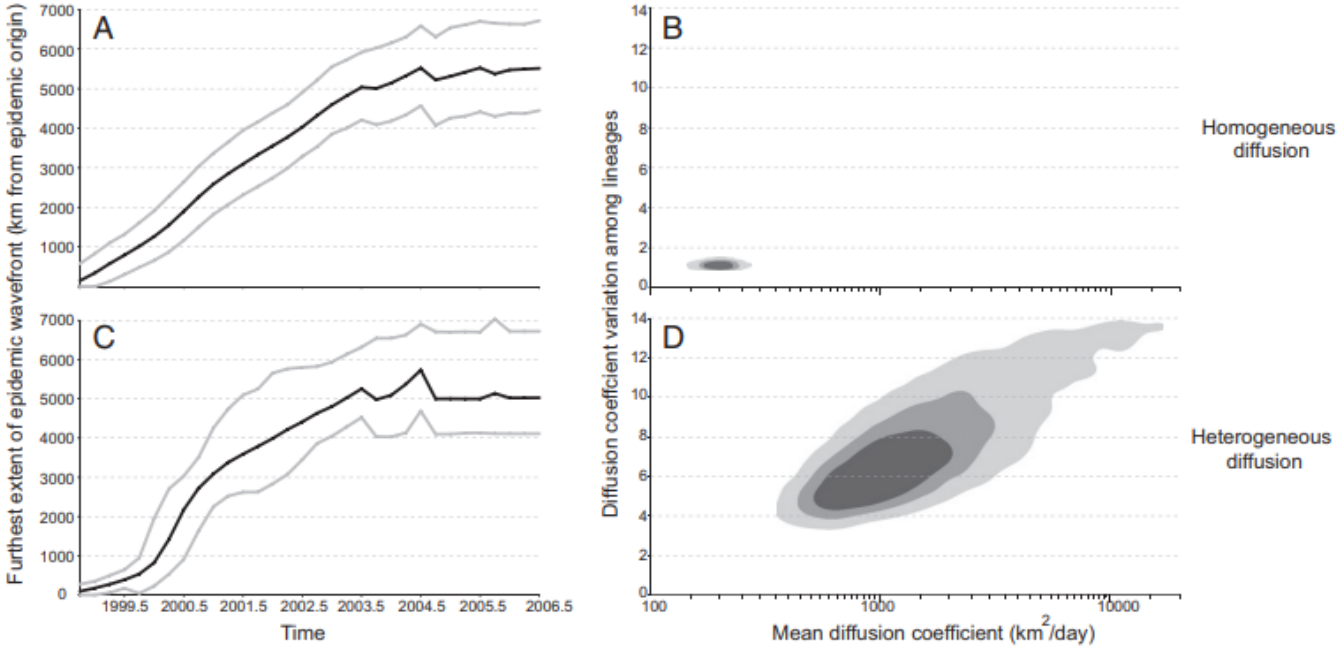
Example II: West Nile Virus in the United States of America



West Nile Virus in the USA: details

- Study from 2012, used 104 complete WNV genomes (11,029 nt);
- Employed a random walk diffusion model.

$$D \approx \frac{1}{n} \sum_{i=1}^{n} \left(\frac{d_i}{2t_i} \right)^2 \tag{1}$$



West Nile Virus in the USA: findings

- Epidemic acceleration rate estimated from genomic data is almost **identical** to what was estimated from large-scale patterns of spatiotemporal WNV incidence;

West Nile Virus in the USA: findings

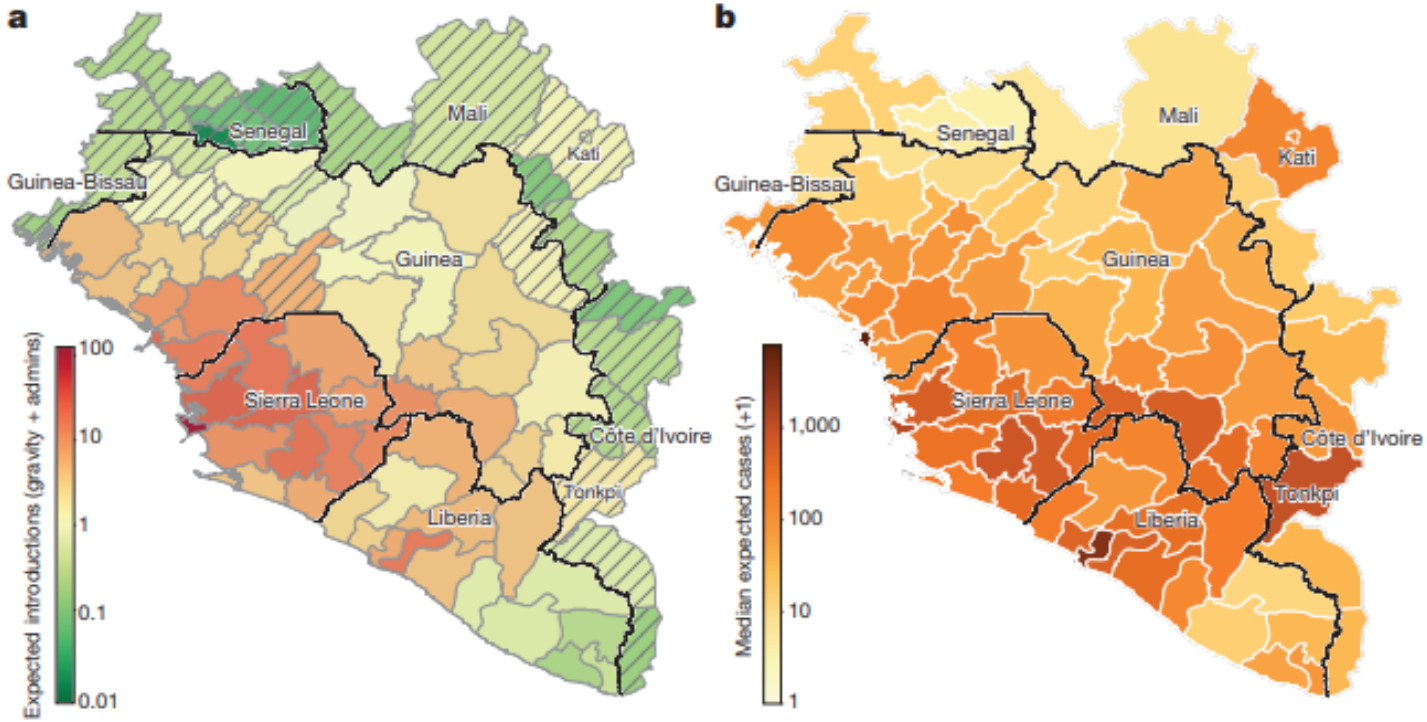
- Epidemic acceleration rate estimated from genomic data is almost **identical** to what was estimated from large-scale patterns of spatiotemporal WNV incidence;
- Rare, long range movements **unlikely to be detected in the field**;

West Nile Virus in the USA: findings

- Epidemic acceleration rate estimated from genomic data is almost **identical** to what was estimated from large-scale patterns of spatiotemporal WNV incidence;
- Rare, long range movements **unlikely to be detected in the field**;
- Traditional approaches overestimate R_0 by ignoring heterogeneity between lineages.

Example IIIa: Ebola epidemics in West Africa

[animation]



Ebola epidemics in West Africa: details

- Study from 2017, used 1610 (!) full EBOV genomes (18,992 nt);

Ebola epidemics in West Africa: details

- Study from 2017, used 1610 (!) full EBOV genomes (18,992 nt);
- Employed a phylogeographic generalised linear model:

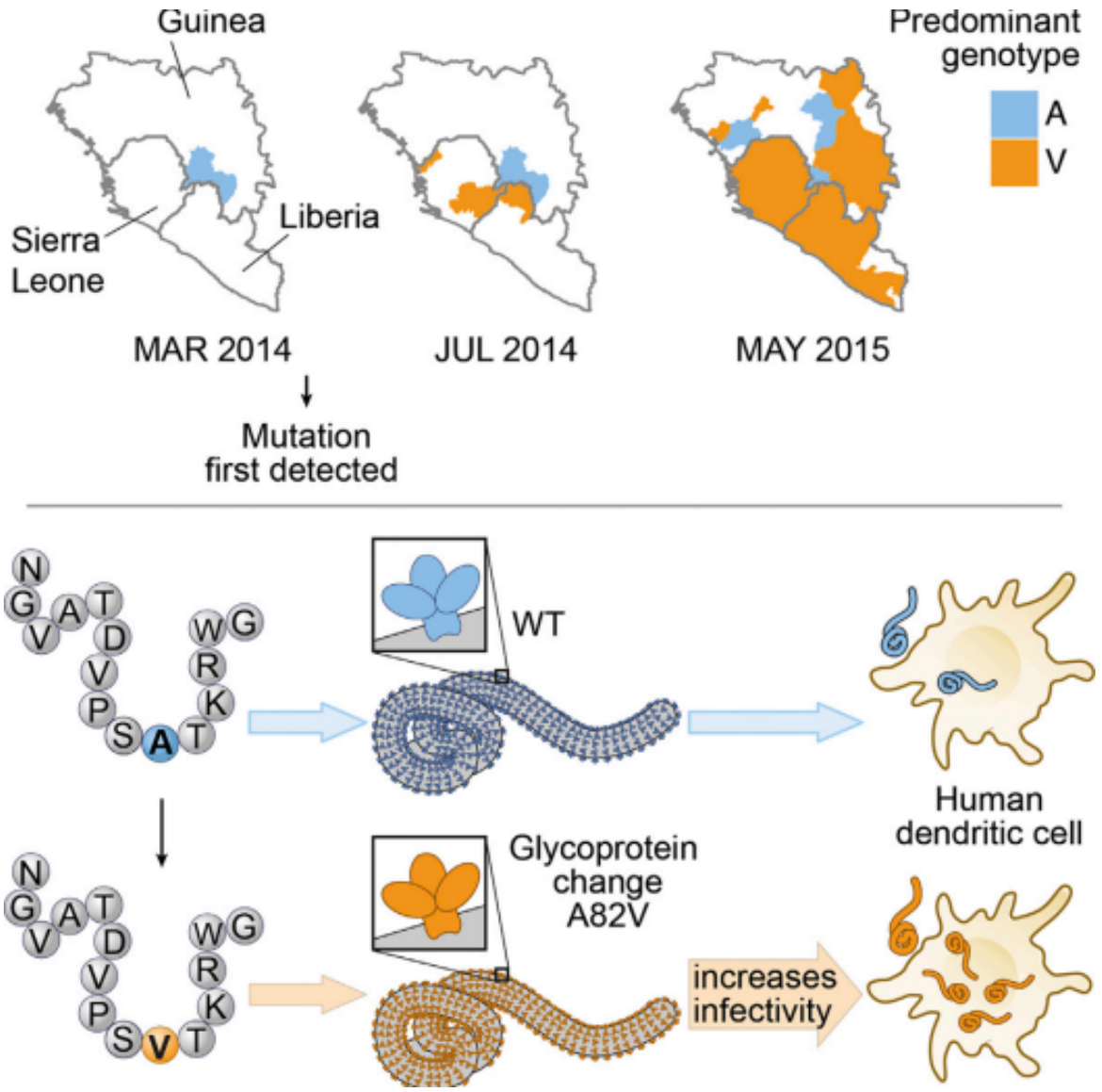
$$\log \Lambda_{ij} = X_{ij}^T \delta\boldsymbol{\beta} + \epsilon_i + \epsilon_j,$$
$$\epsilon_k \sim \text{Normal}(0, \sigma^2) \text{ for } k = 1, \ldots, K, \text{ with}$$
$$\sigma^2 \sim \text{Inverse-Gamma}(0.001, 0.001),$$
$$\beta_j \sim \text{Normal}(0, 16).$$

Table 2 | Summary of generalized linear model results with case counts as the response variable

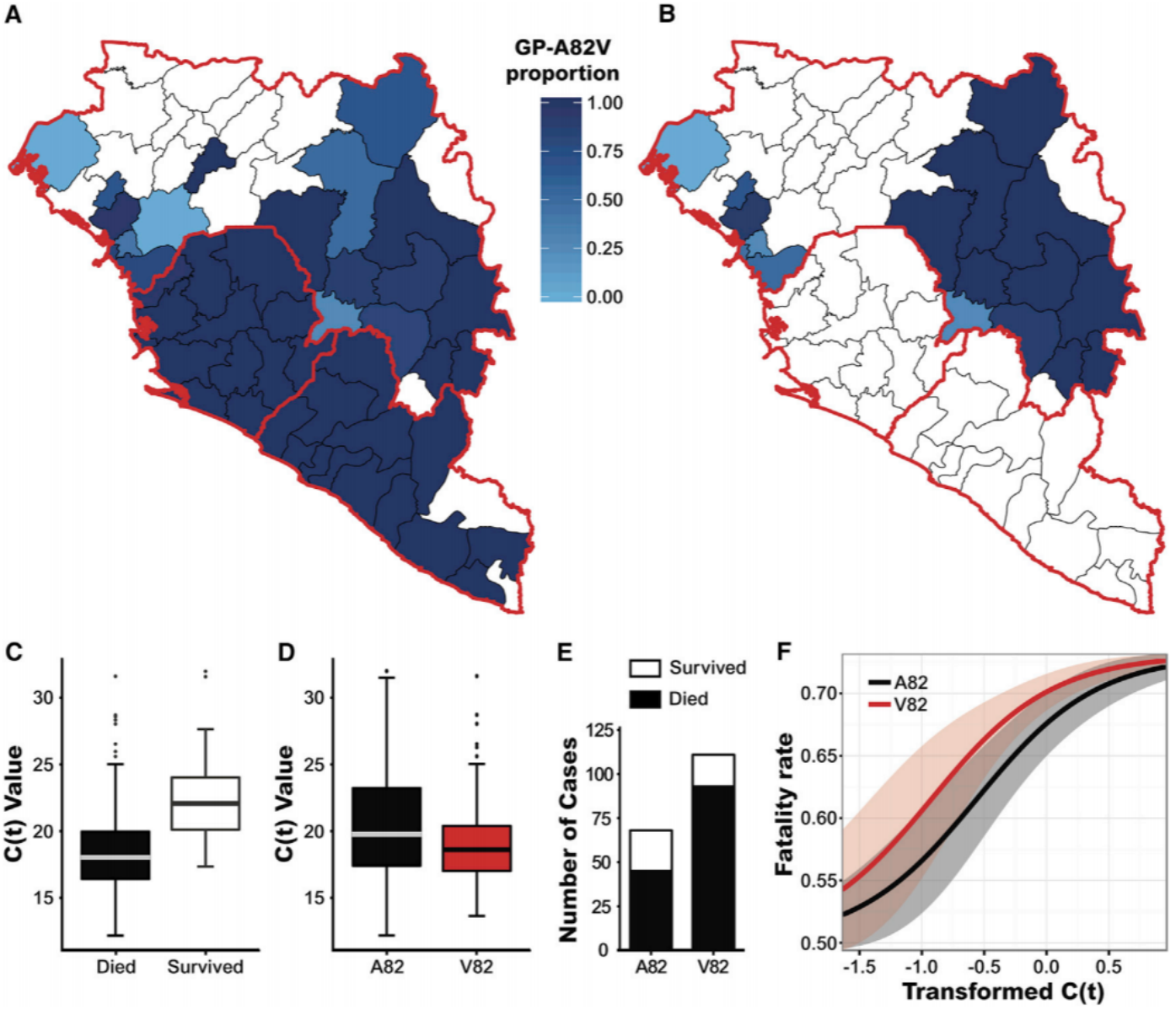
Predictor*	Description	Coefficient†	95% CI‡	Inclusion§	BF‖
TempSS	Temperature seasonality	−1.1	−1.6, −0.5	0.83	>50
TT50K	Time to travel to a population centre of 50,000 people	−0.9	−1.4, −0.4	0.62	32.4
PopSize	Population size	0.9	0.3, 1.6	0.60	29.6
Precip	Precipitation	0.8	0.2, 1.3	0.18	4.4
TT100K	Time to travel to a population centre of 0.1 million people	−0.8	−1.7, −0.1	0.16	3.8

*Predictors included in the model with Bayes factor >3.
†Mean coefficient.
‡95% highest posterior density credible interval (CI).
§Probability that the predictor was included in the model.
‖BF, Bayes factor.

Example IIIb: GPA82V mutation and mortality



GPA82V mutation and mortality



Lessons learned so far

Phylodynamics can help us

- Integrate genetic and antigenic data to understand immune escape;

Lessons learned so far

Phylodynamics can help us

- Integrate genetic and antigenic data to understand immune escape;
- Uncover previously undetectable routes of pathogen spatial spread;

Lessons learned so far

Phylodynamics can help us

- Integrate genetic and antigenic data to understand immune escape;
- Uncover previously undetectable routes of pathogen spatial spread;
- Estimate the rate of spatial spread of an epidemic, its R_0 and drivers of spread;

Lessons learned so far

Phylodynamics can help us

- Integrate genetic and antigenic data to understand immune escape;
- Uncover previously undetectable routes of pathogen spatial spread;
- Estimate the rate of spatial spread of an epidemic, its R_0 and drivers of spread;
- Study clinical outcomes while controlling for underlying dependencies.

New developments: ARTIC network

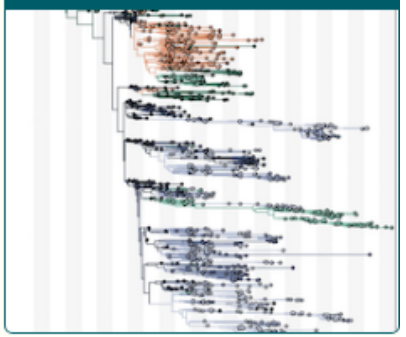
http://artic.network/

The field laboratory

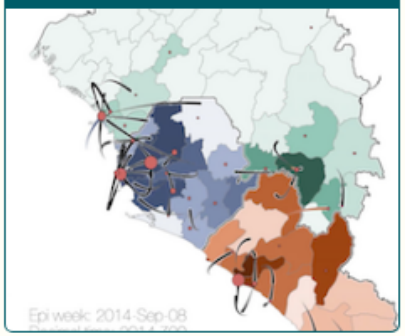
Portable sequencing

Deployable bioinformatics

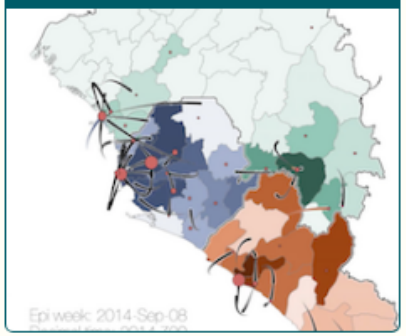
Real-time phylodynamics



Visualisation and interpretation



Tools and resources



Funded by the Wellcome Trust

Collaborators Award 206298/Z/17/Z --- ARTIC network

Lab in a suitcase

- Develop portable sample inactivation and reagent/sample preparation;

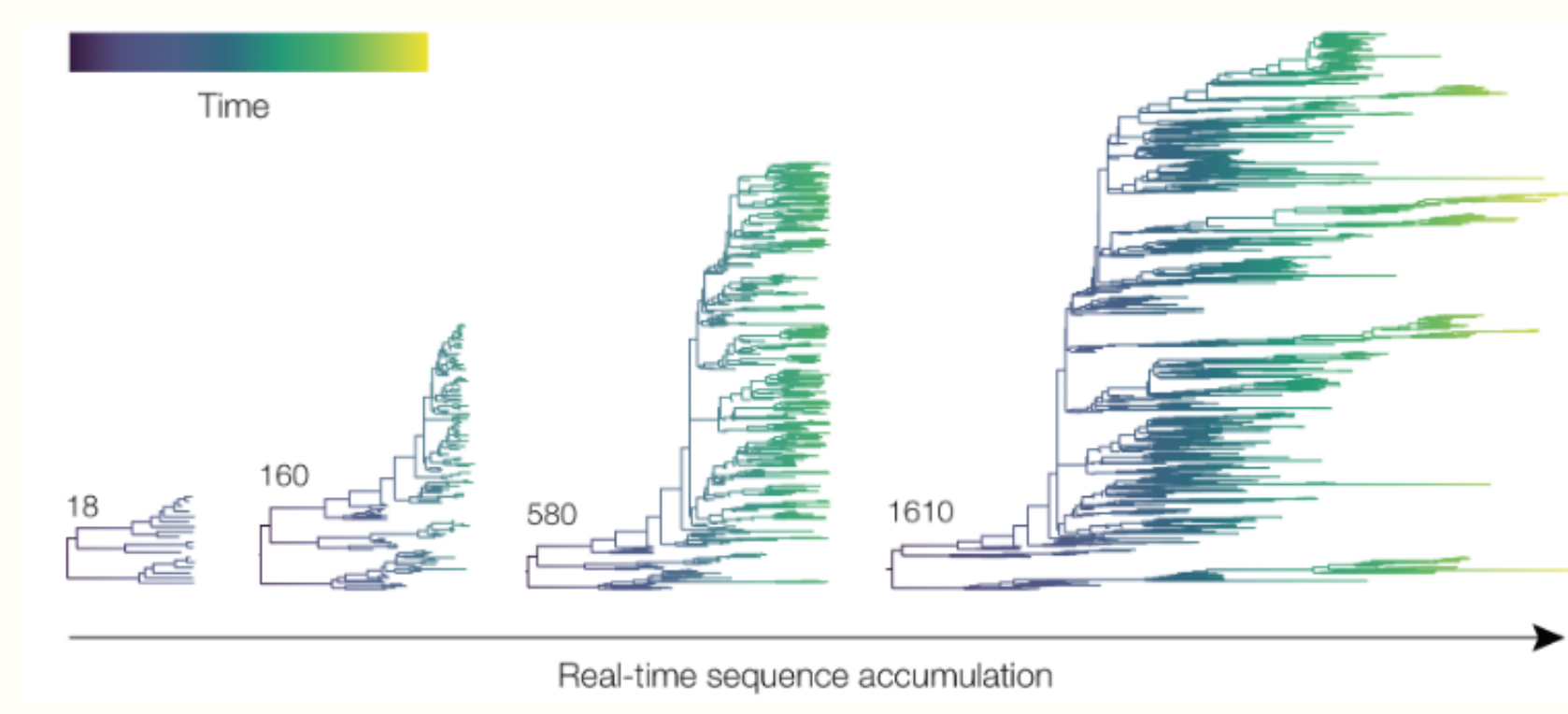
Lab in a suitcase

- Develop portable sample inactivation and reagent/sample preparation;
- Integration of portable lab in a suitcase;

Lab in a suitcase

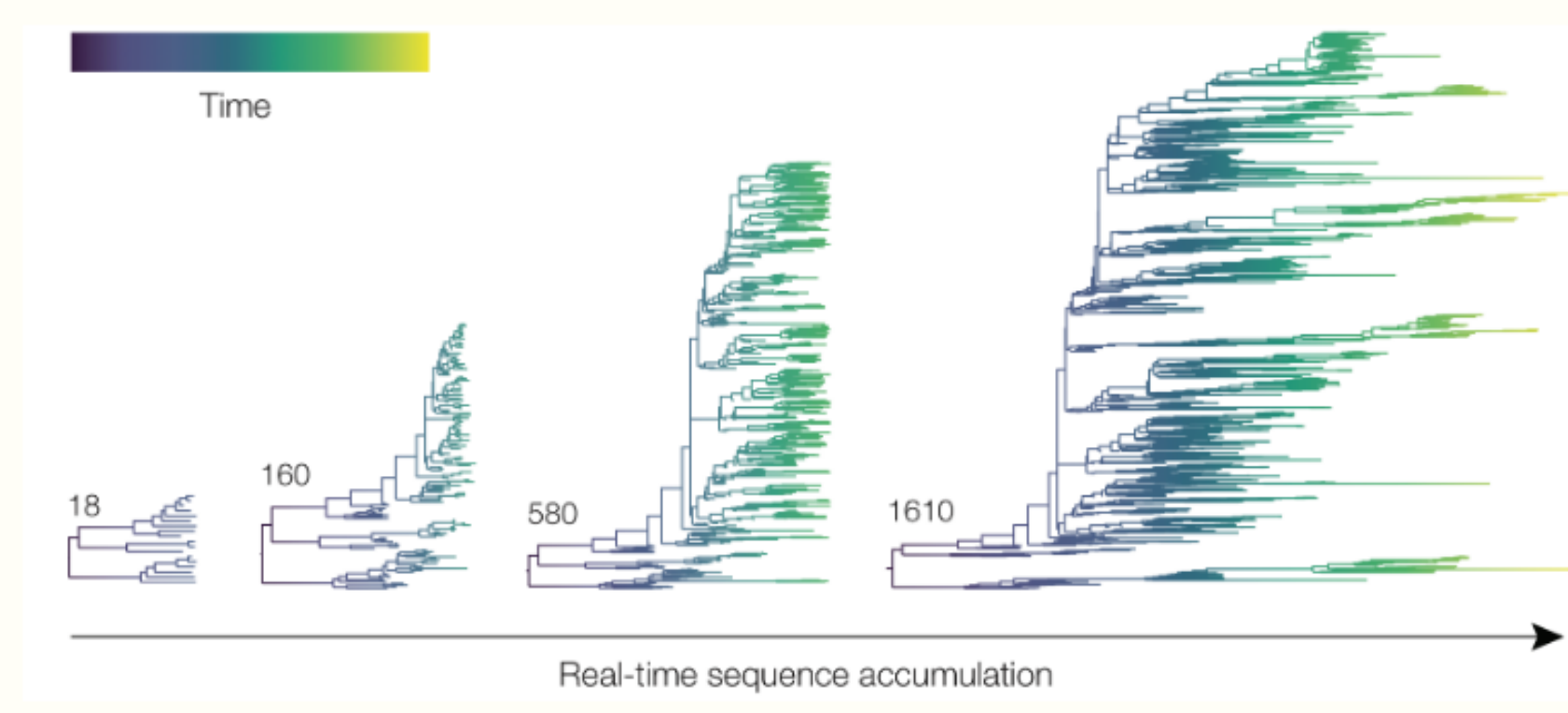
- Develop portable sample inactivation and reagent/sample preparation;
- Integration of portable lab in a suitcase;
- Reducing contamination risk.

Real time phylodynamics



- Build trees and make inferences as sequences arrive;

Real time phylodynamics



- Build trees and make inferences as sequences arrive;
- Sequential Monte Carlo (SMC).

Take home

Phylodynamics is a powerful tool

DNA sequences from pathogens +
environmental/socio-economic data can give us insight

Take home

Phylodynamics is a powerful tool

DNA sequences from pathogens + environmental/socio-economic data can give us insight

Searching trees is **hard**

Developing better statistical models and computational tools is crucial.

Take home

Phylodynamics is a powerful tool

DNA sequences from pathogens + environmental/socio-economic data can give us insight

Searching trees is **hard**

Developing better statistical models and computational tools is crucial.

Data integration is crucial

We need more and better data.

Take home

Phylodynamics is a powerful tool

DNA sequences from pathogens + environmental/socio-economic data can give us insight

Searching trees is **hard**

Developing better statistical models and computational tools is crucial.

Data integration is crucial

We need more and better data.

Nature is complicated

We need better models to go along.

THE END