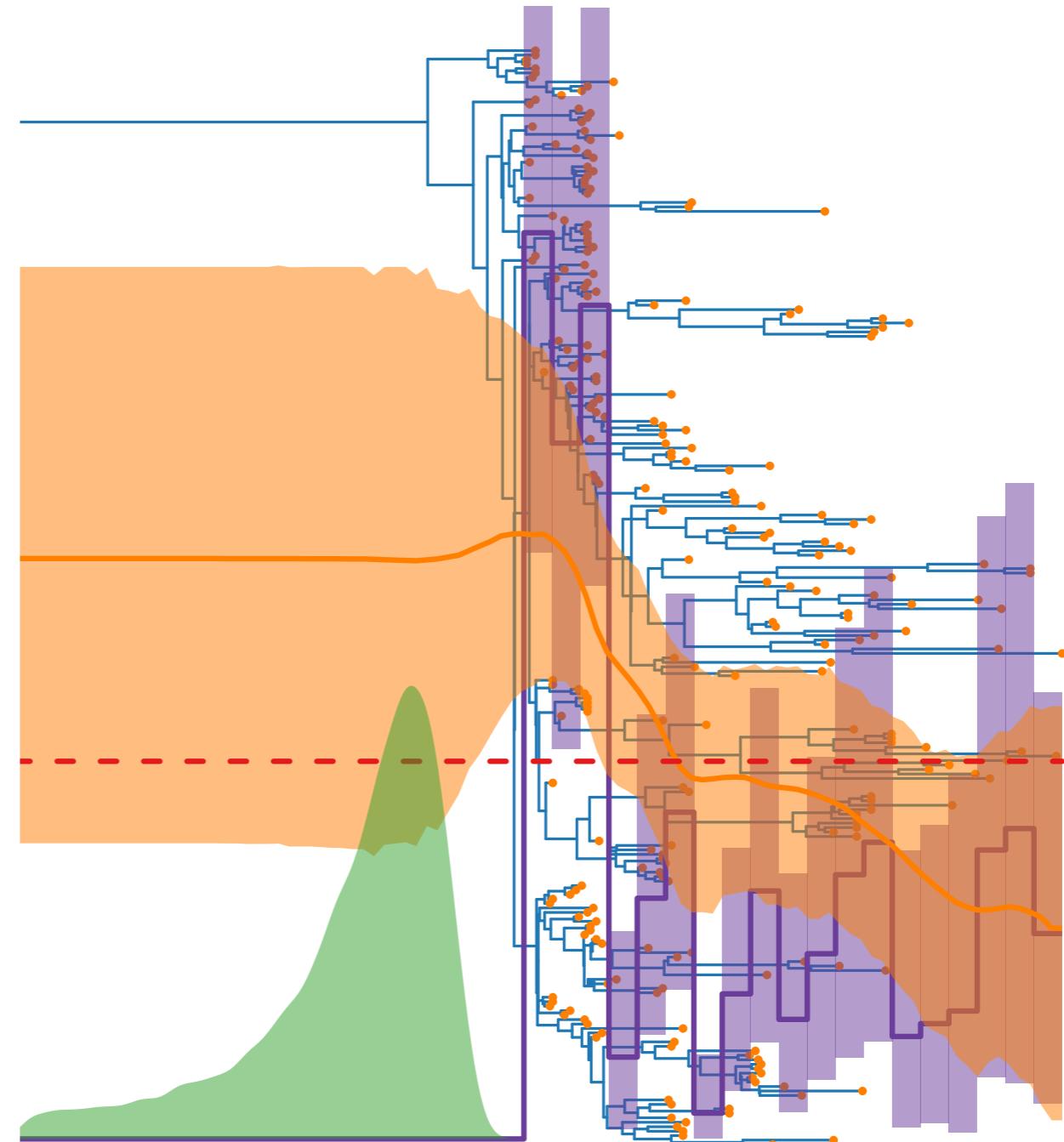


# POPULATION DYNAMICS OF INFECTIOUS DISEASES WITH THE BIRTH-DEATH SKYLINE



LOUIS DU PLESSIS

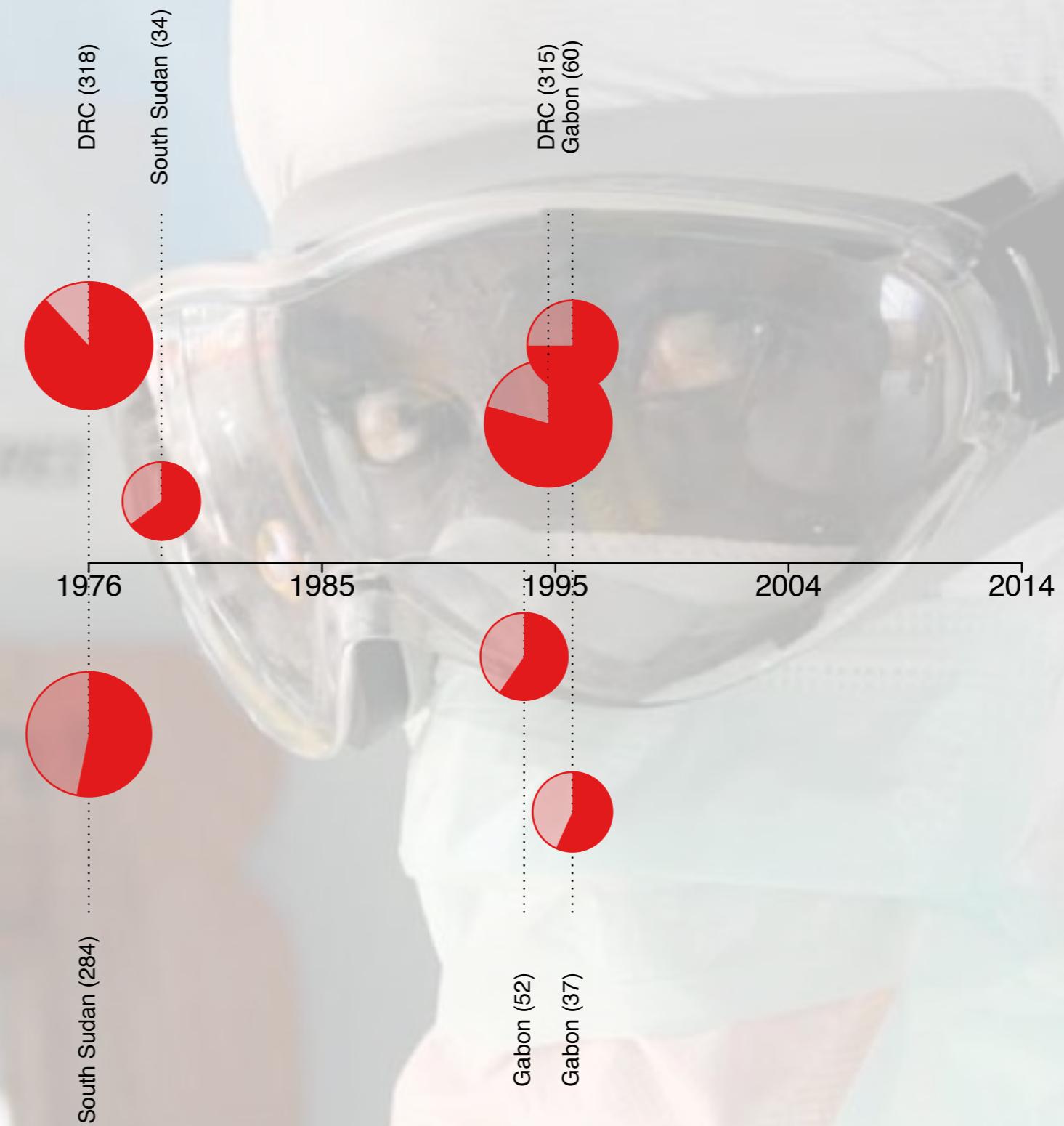
# Emerging and re-emerging infections

---

# Emerging and re-emerging infections

**1976 - 2000**

7 Ebola outbreaks



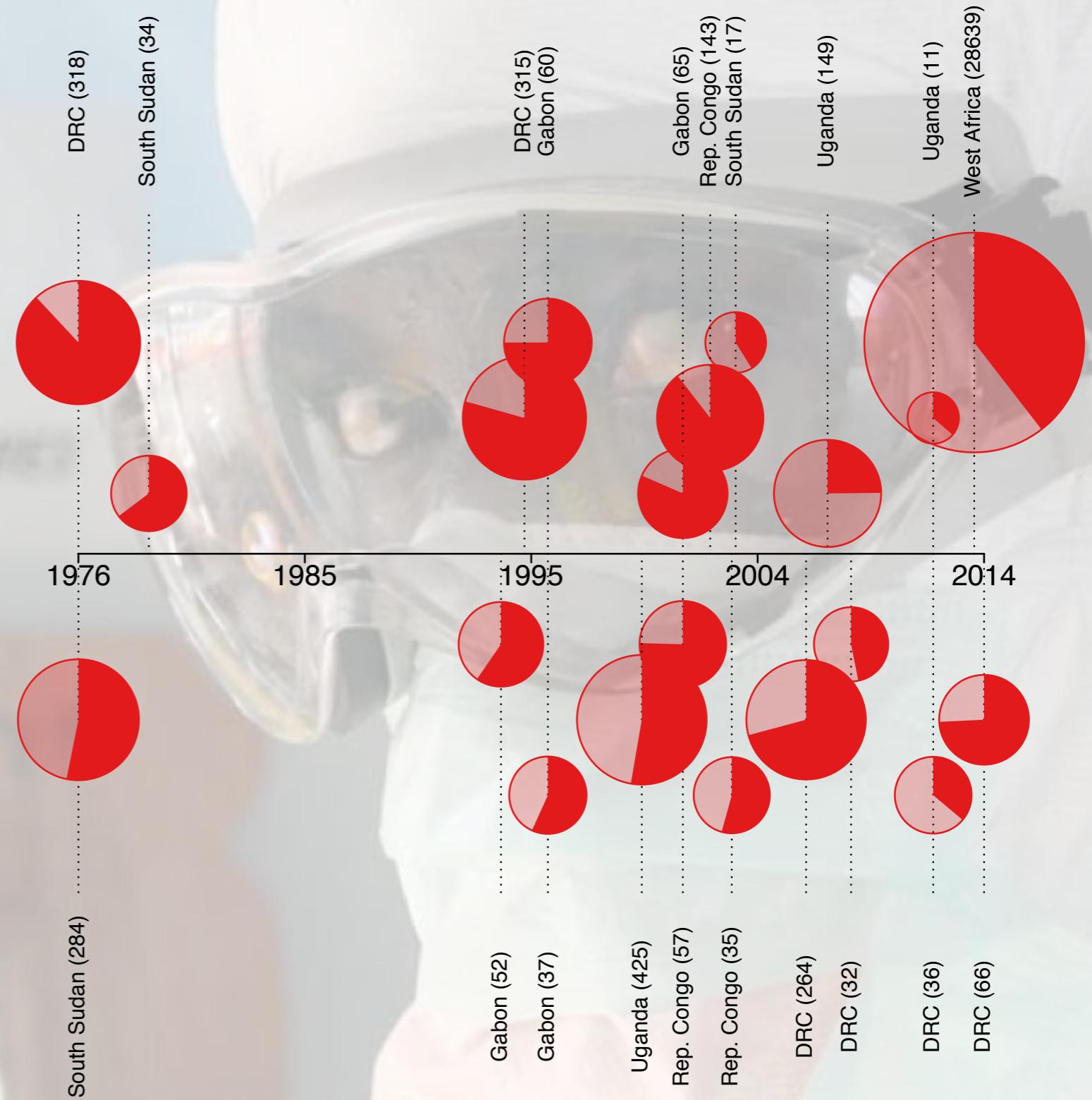
# Emerging and re-emerging infections

**1976 - 2000**

7 Ebola outbreaks

**2000 - 2014**

13 outbreaks



# Emerging and re-emerging infections

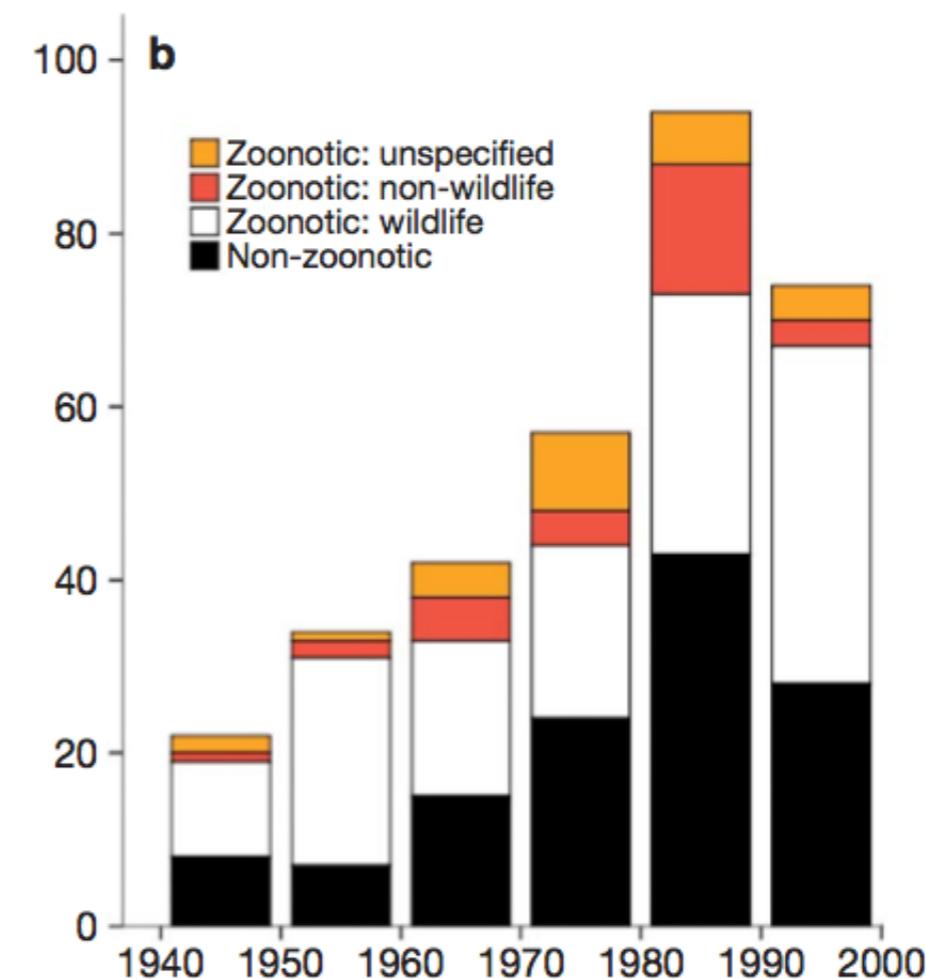
1976 - 2000

7 Ebola outbreaks

2000 - 2014

13 outbreaks

Increase in emerging infectious disease events



Jones et al. Nature 2008

# Emerging and re-emerging infections

**1976 - 2000**

7 Ebola outbreaks

**2000 - 2014**

13 outbreaks

**Increase in frequency**

Higher population density



More contact with animal reservoirs

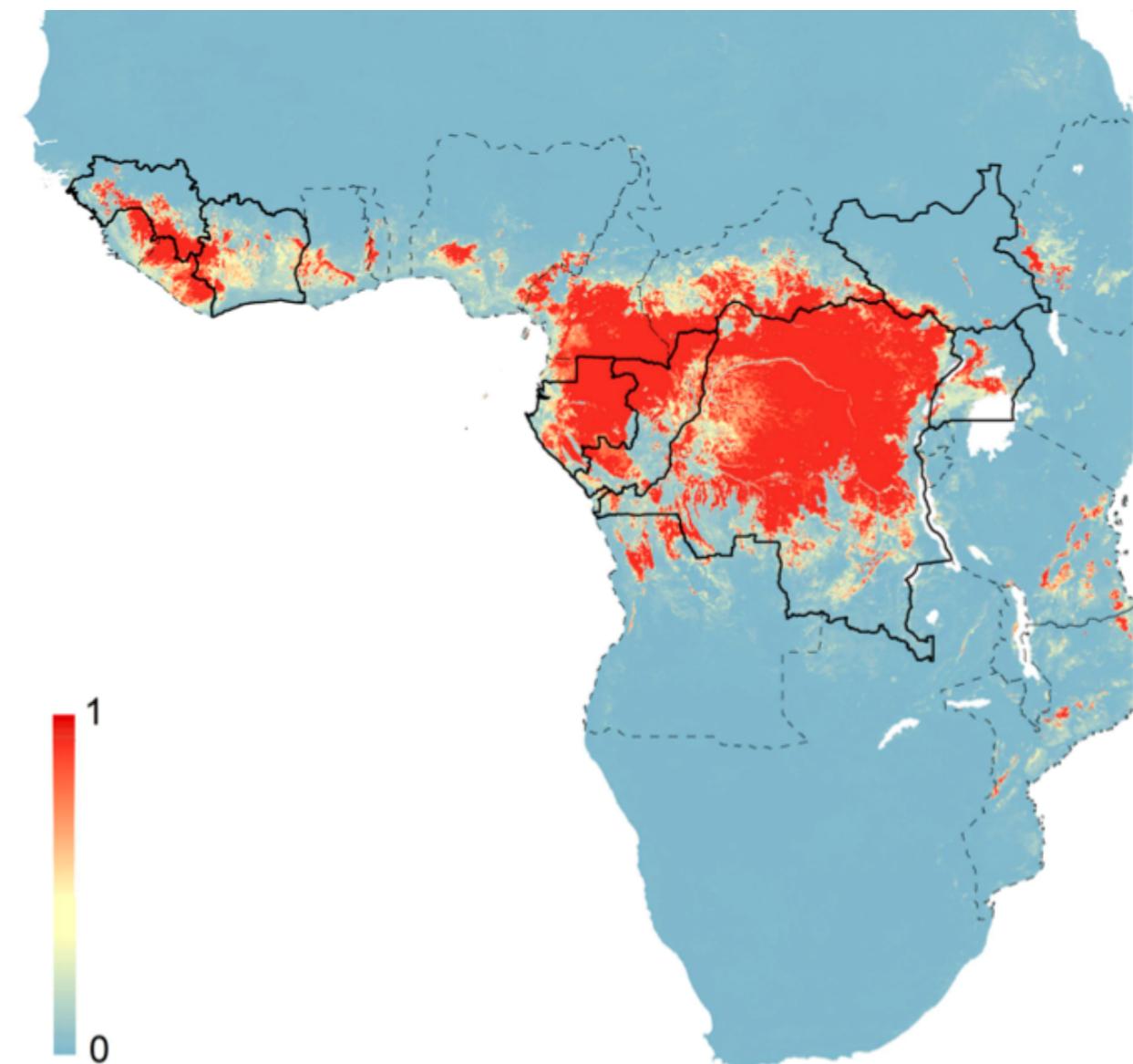
**Increased epidemic potential**

Better infrastructure



More chance to spread

**Ebola outbreak risk**



Pigott et al. eLife 2014



## During future outbreaks:

- A much larger proportion of genomes will be sequenced
- Sequencing will be done faster and directly in the field



Quick et al. Nature 2016

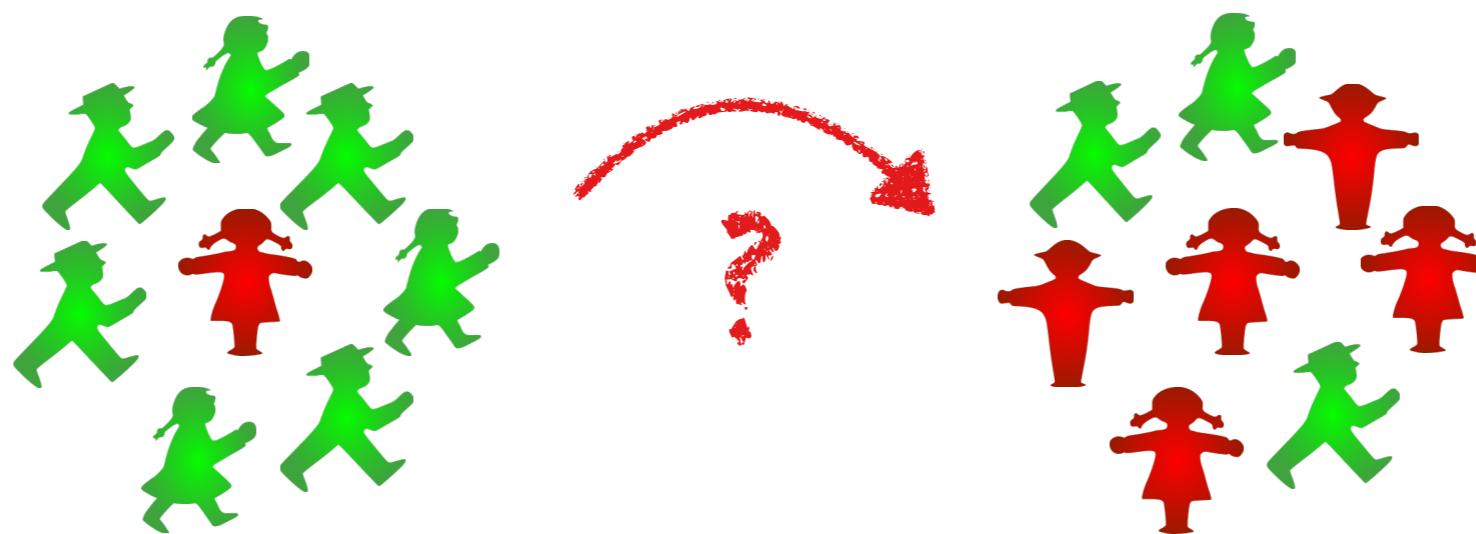


## During future outbreaks:

- A much larger proportion of genomes will be sequenced
- Sequencing will be done faster and directly in the field

# The effective reproduction number ( $R_e$ )

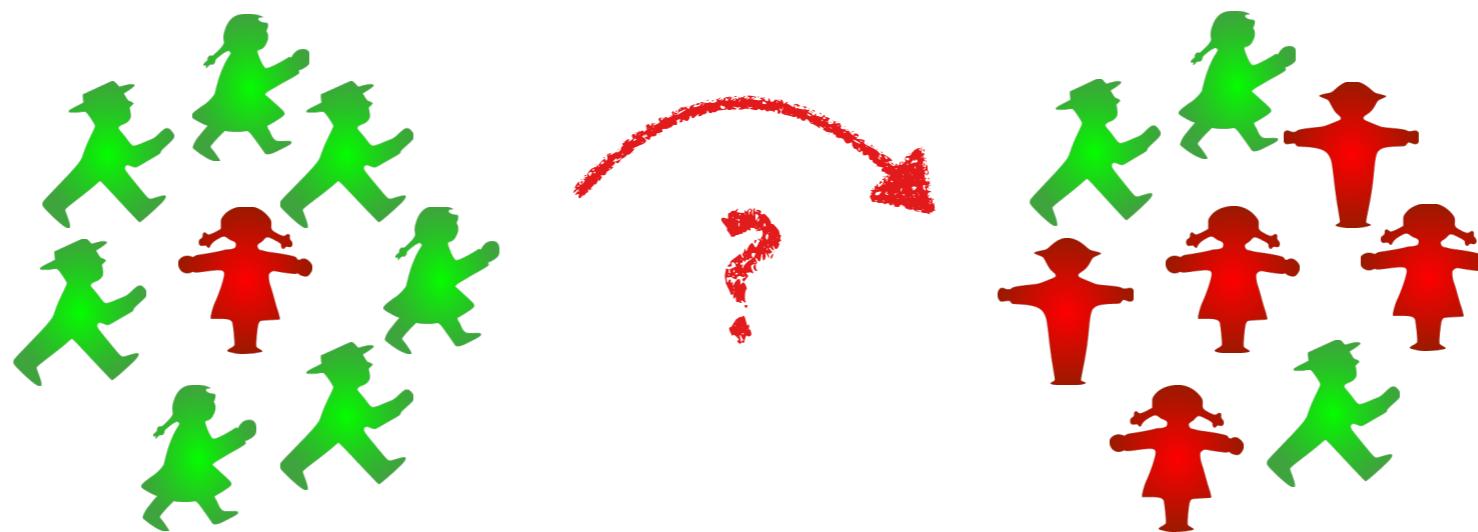
---



- How many secondary infections does the average infected person cause?
- Are public health interventions effective?

# The effective reproduction number ( $R_e$ )

---



- How many secondary infections does the average infected person cause?
- Are public health interventions effective?

$R_e < 1$  — Winning 😊

$R_e > 1$  — Losing 😥

# Transmission dynamics from sequencing data

---

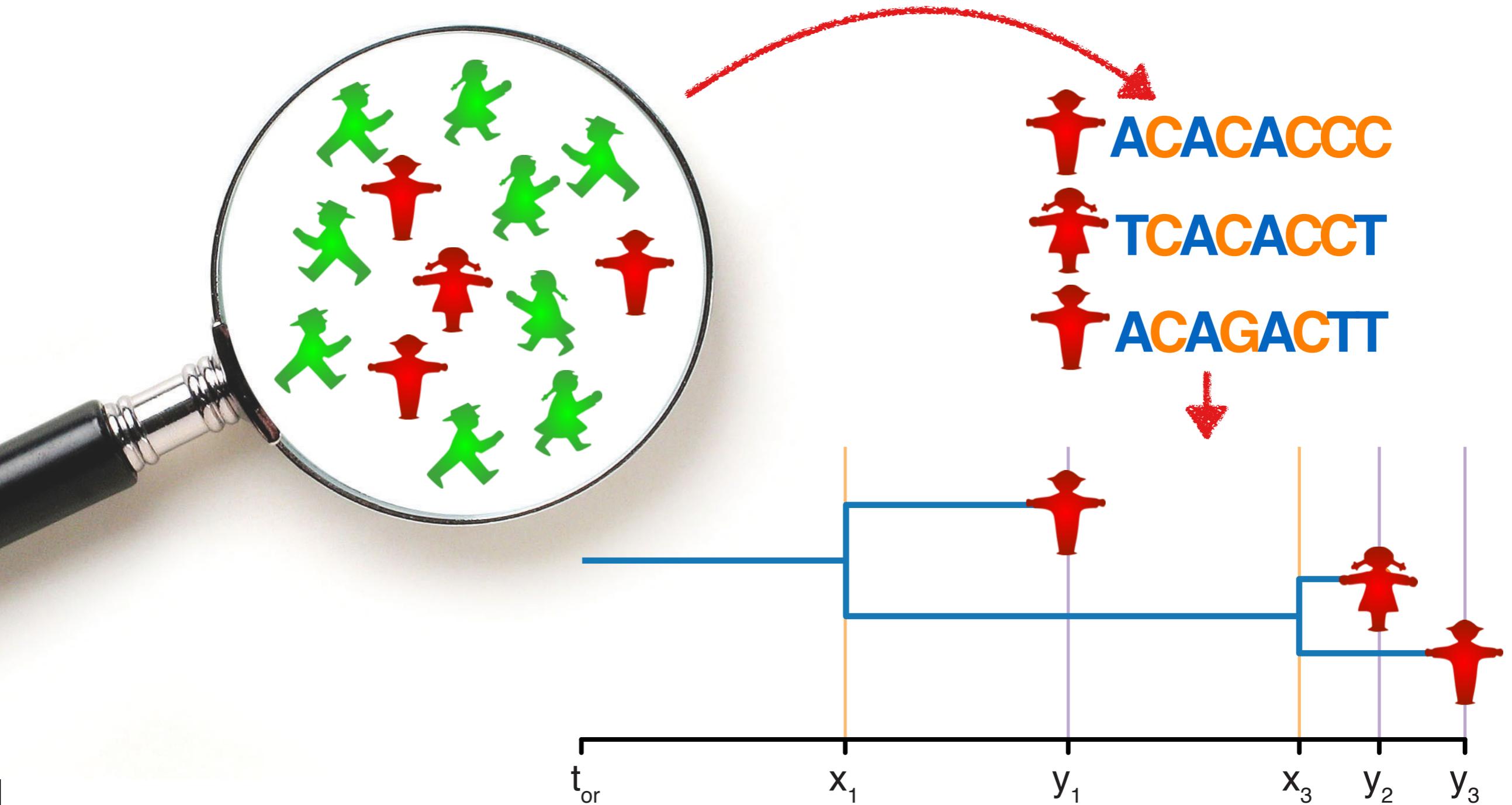


# Transmission dynamics from sequencing data

---



# Transmission dynamics from sequencing data

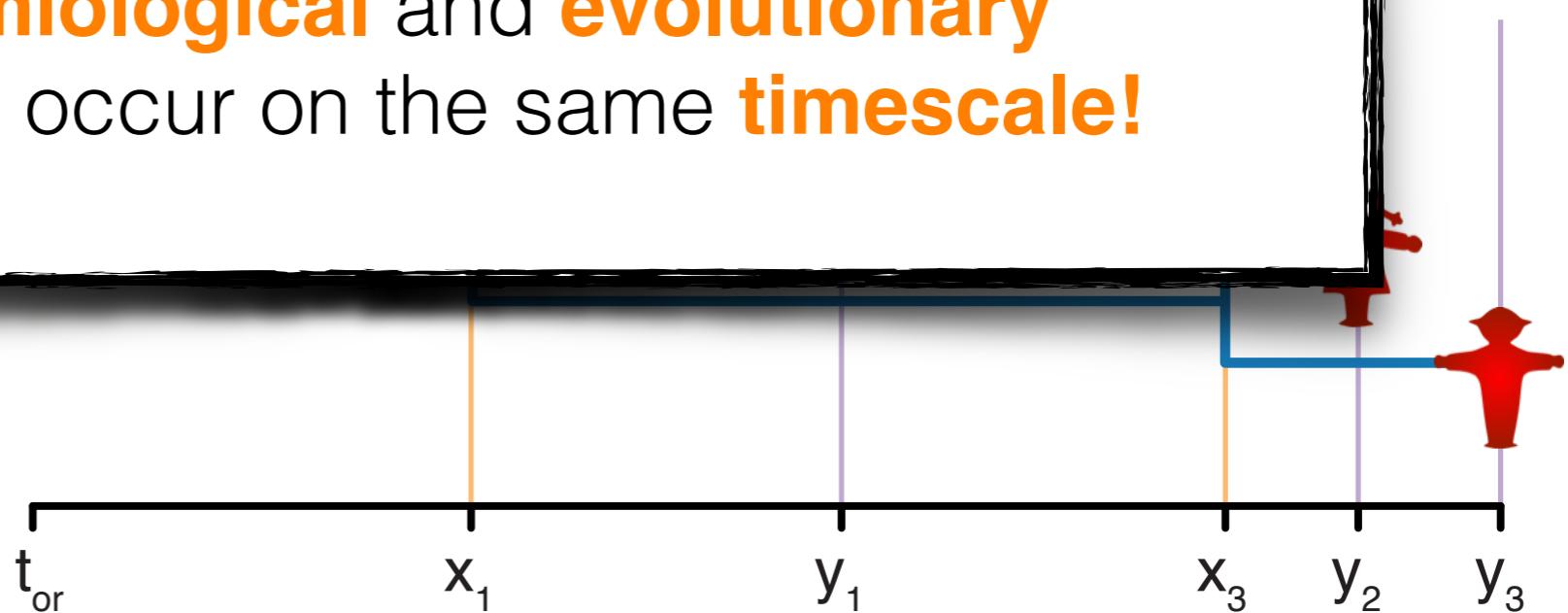


# Transmission dynamics from sequencing data

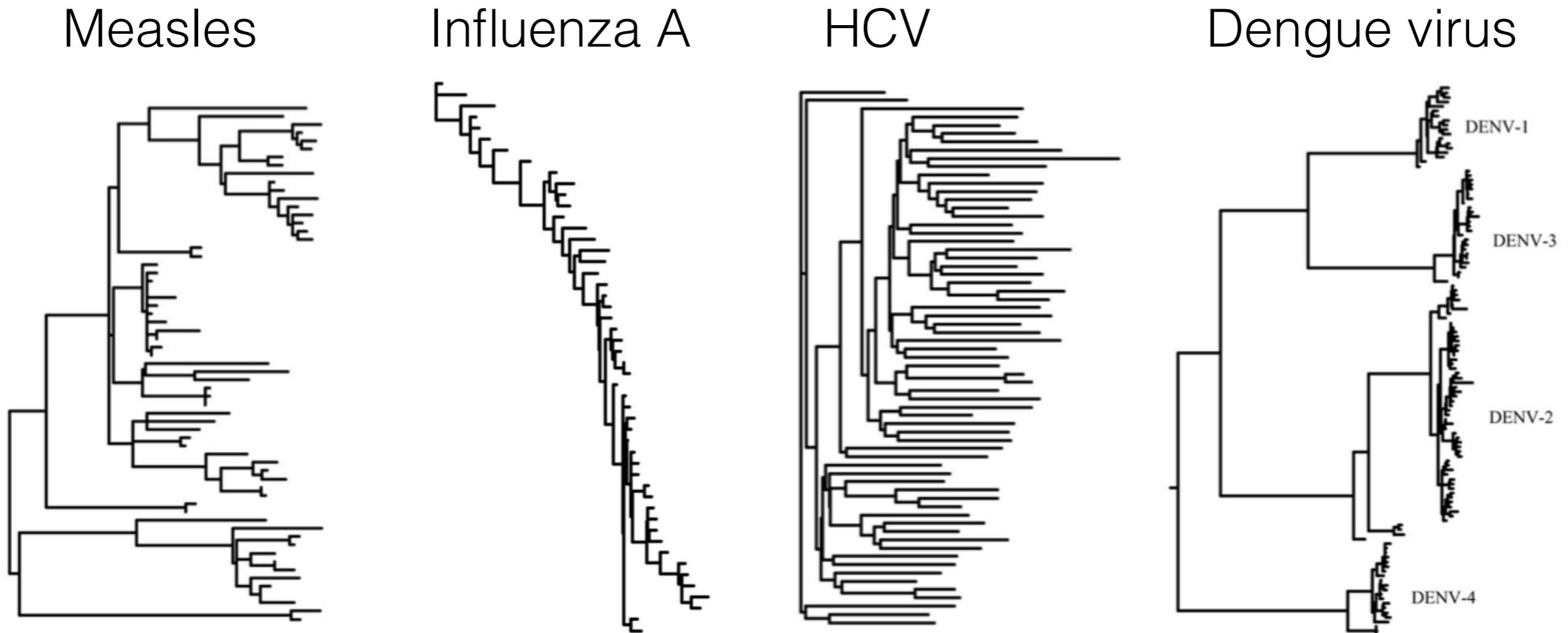
**Need a measurably evolving population!**

- Large population size
- High mutation rate
- Short generation times

**Epidemiological** and **evolutionary** dynamics occur on the same **timescale!**

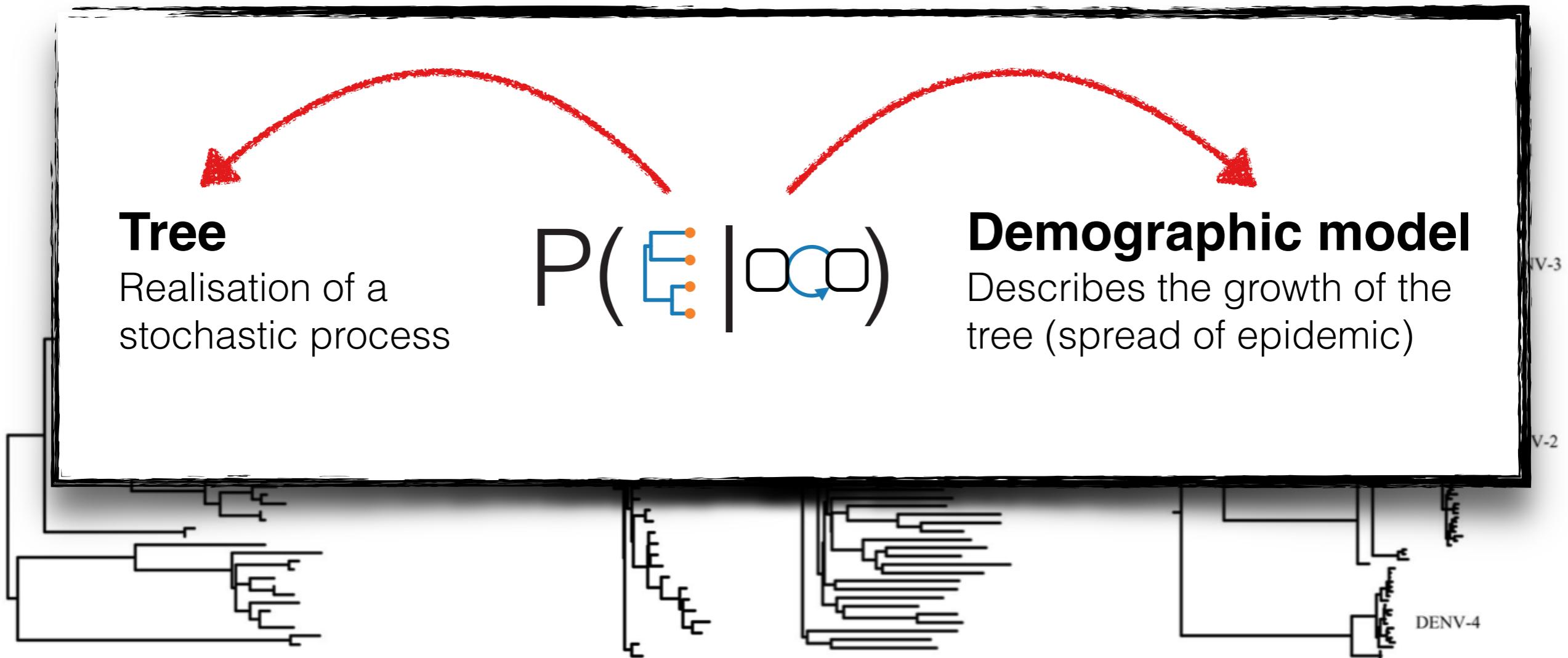


# Genomes contain a signature of the epidemiological dynamics



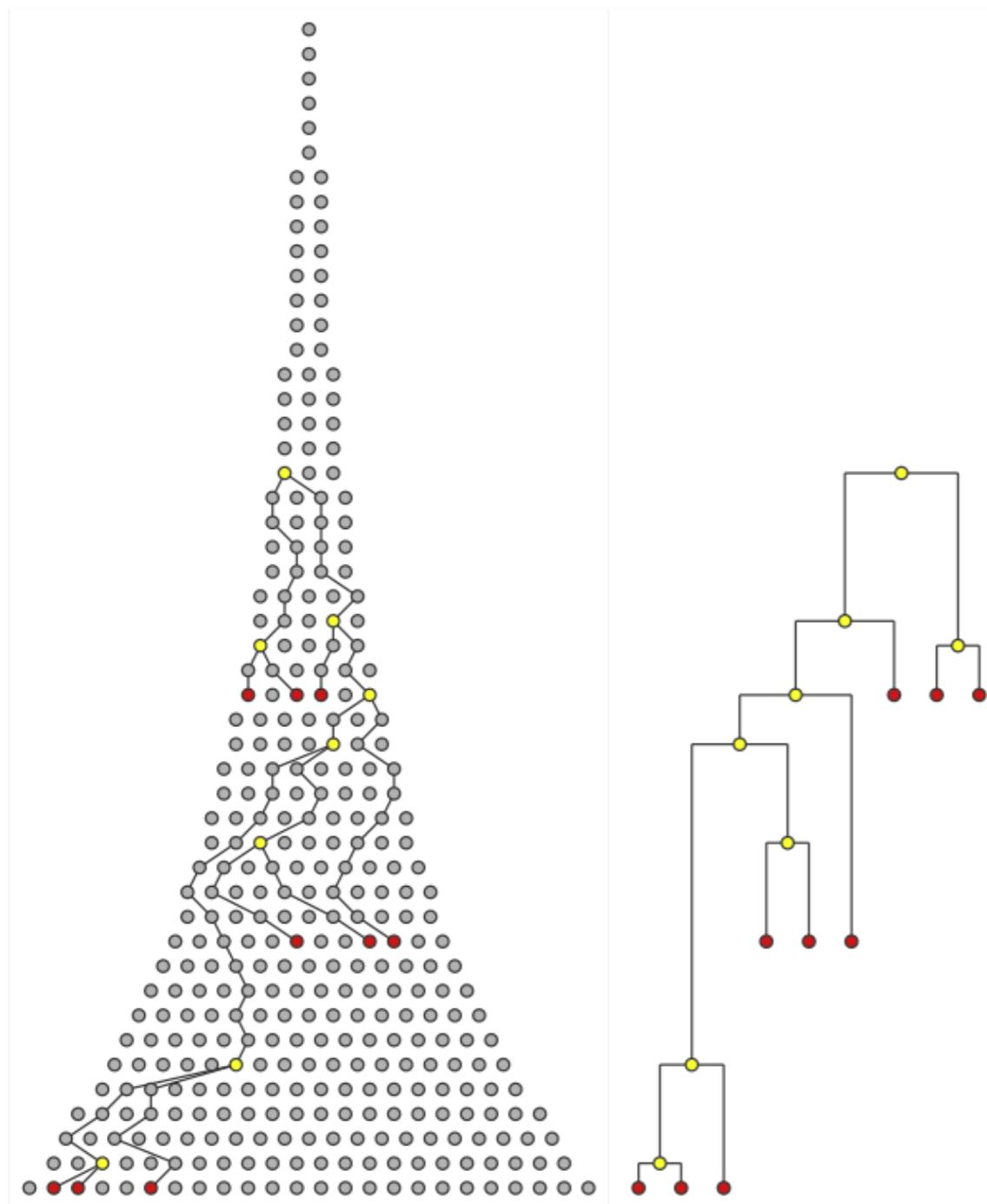
**Phyldynamics** aims to infer the **dynamics** responsible for the observed **phylogenies**

# Genomes contain a signature of the epidemiological dynamics



**Phylodynamics** aims to infer the **dynamics** responsible for the observed **phylogenies**

# The coalescent approximation



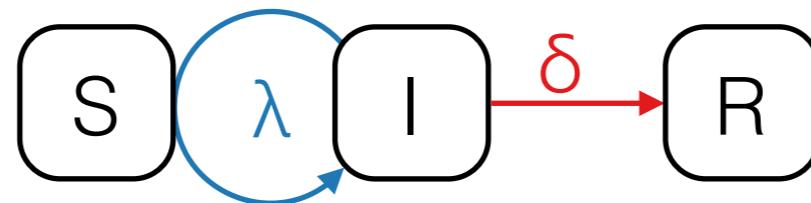
- Given an  $N_e$  (effective population size)
- Calculate the time it takes for 2 nodes in the tree to coalesce

$$P(E | \text{---})$$

A red arrow points down to the letter E in the equation above.

# Compartmental epidemiological models

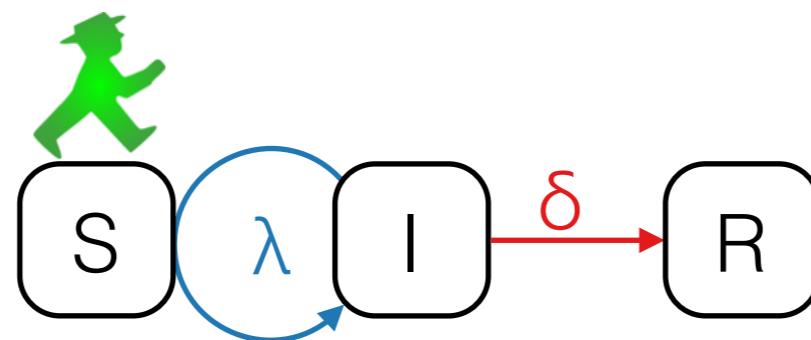
---



- $\lambda$  — infection rate ( $\lambda_0^{IS/N}$ )
- $\delta$  — becoming-noninfectious rate

# Compartmental epidemiological models

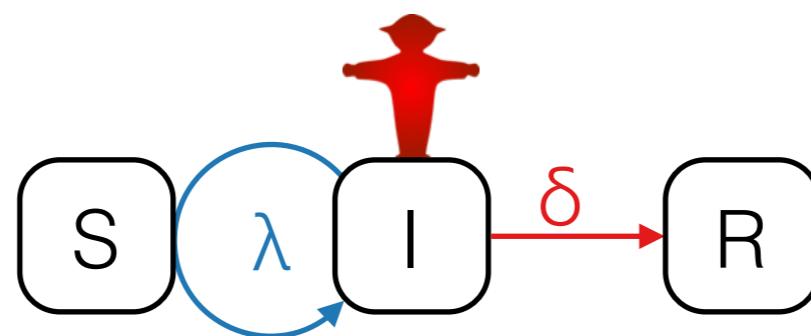
---



- $\lambda$  — infection rate ( $\lambda_0^{IS/N}$ )
- $\delta$  — becoming-noninfectious rate

# Compartmental epidemiological models

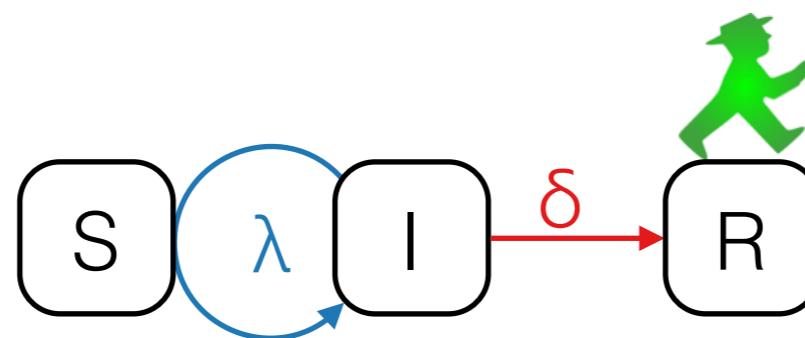
---



- $\lambda$  — infection rate ( $\lambda_0^{IS/N}$ )
- $\delta$  — becoming-noninfectious rate

# Compartmental epidemiological models

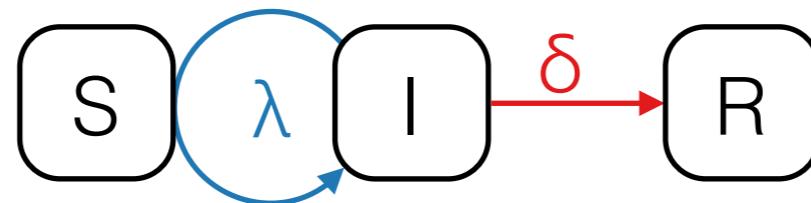
---



- $\lambda$  — infection rate ( $\lambda_0^{IS/N}$ )
- $\delta$  — becoming-noninfectious rate

# Compartmental epidemiological models

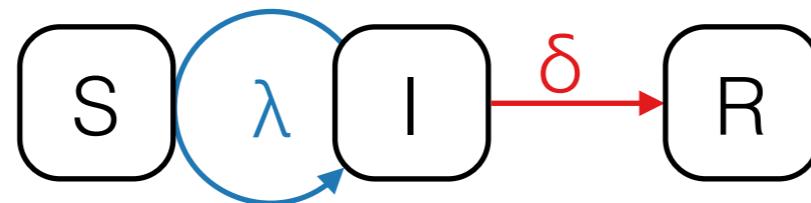
---



- $\lambda$  — infection rate ( $\lambda_0^{IS/N}$ )
- $\delta$  — becoming-noninfectious rate

# Compartmental epidemiological models

---



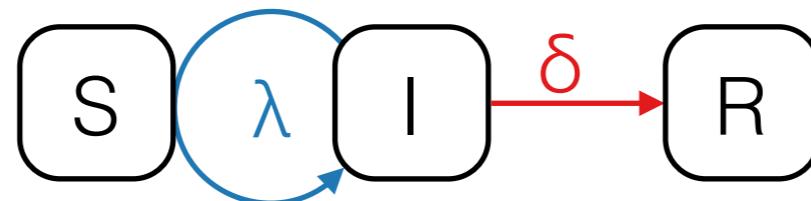
- $\lambda$  — infection rate ( $\lambda_0^{IS/N}$ )
- $\delta$  — becoming-noninfectious rate

$$\frac{dS}{dt} = -\lambda_0^{IS/N}$$

$$\frac{dI}{dt} = \lambda_0^{IS/N} - \delta I$$

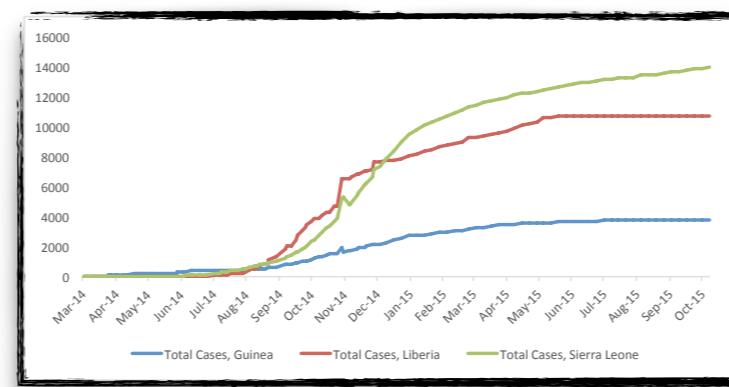
$$\frac{dR}{dt} = \delta I$$

# Compartmental epidemiological models

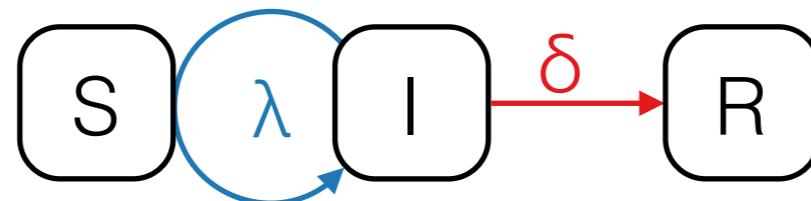


- $\lambda$  — infection rate ( $\lambda_0^{IS/N}$ )
- $\delta$  — becoming-noninfectious rate

$$\begin{aligned} \frac{dS}{dt} &= -\lambda_0^{IS/N} \\ \frac{dI}{dt} &= \lambda_0^{IS/N} - \delta I \\ \frac{dR}{dt} &= \delta I \end{aligned}$$

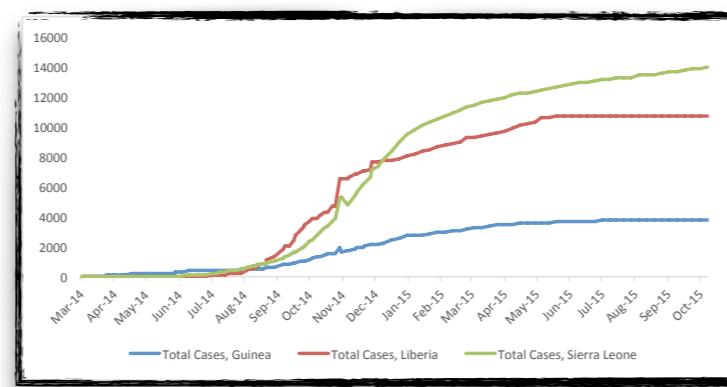


# Compartmental epidemiological models



- $\lambda$  — infection rate ( $\lambda_0^{IS/N}$ )
- $\delta$  — becoming-noninfectious rate

$$\begin{aligned} \frac{dS}{dt} &= -\lambda_0^{IS/N} \\ \frac{dI}{dt} &= \lambda_0^{IS/N} - \delta I \\ \frac{dR}{dt} &= \delta I \end{aligned}$$

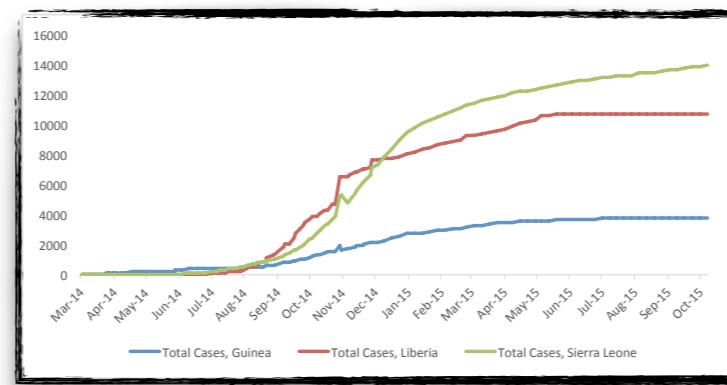


$$R_e = \lambda / \delta$$

# Compartmental epidemiological models

**Completely neglects the transmission tree!  
(who infected whom)**

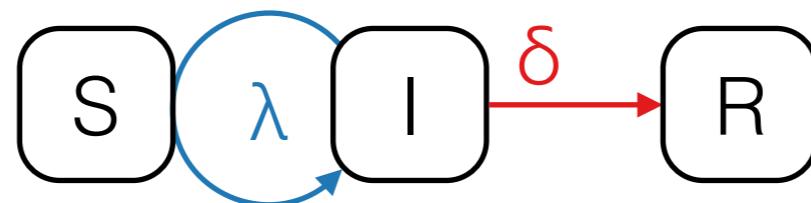
$$\begin{aligned} \frac{dS}{dt} &= -\lambda_0 S/N \\ \frac{dI}{dt} &= \lambda_0 S/N - \delta I \\ \frac{dR}{dt} &= \delta I \end{aligned}$$



$$R_e = \lambda / \delta$$

# The birth-death model

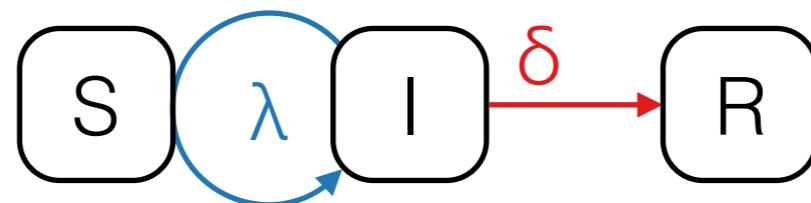
---



- $\lambda$  — infection rate
  - $\delta$  — becoming-noninfectious rate
- 
- Forward in time branching process that generates a tree
  - Assume  $\lambda$  and  $\delta$  are exponentially distributed rates

# The birth-death model

---



- $\lambda$  — infection rate
- $\delta$  — becoming-noninfectious rate

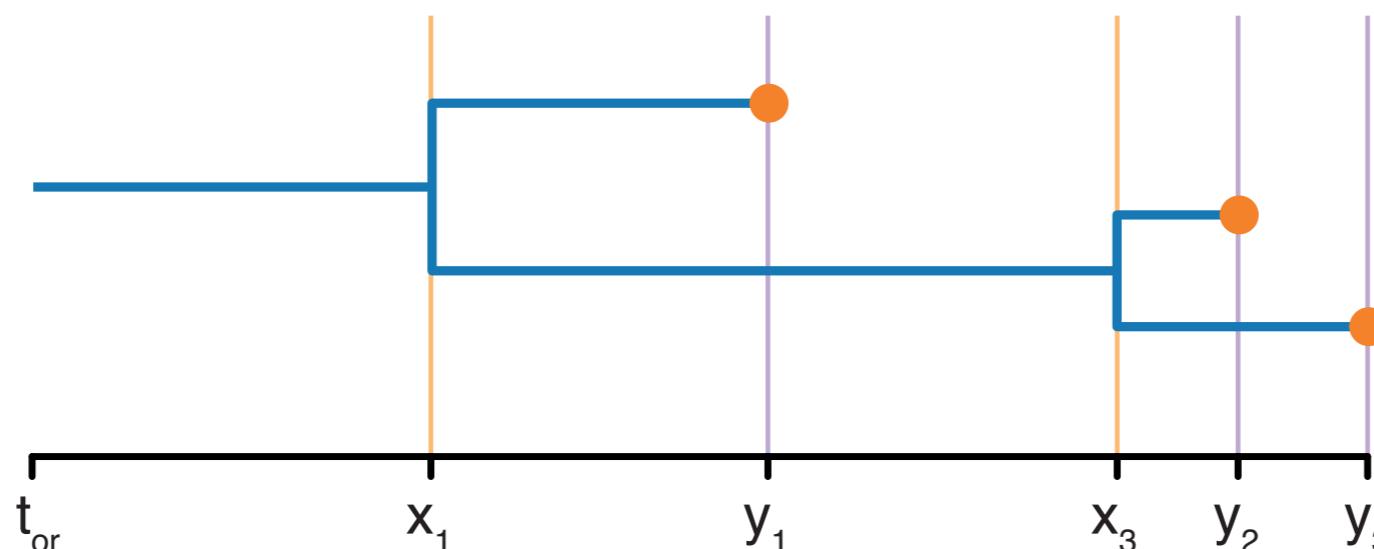
- Forward in time branching process that generates a tree
- Assume  $\lambda$  and  $\delta$  are exponentially distributed rates

$\downarrow$   
P( | )?



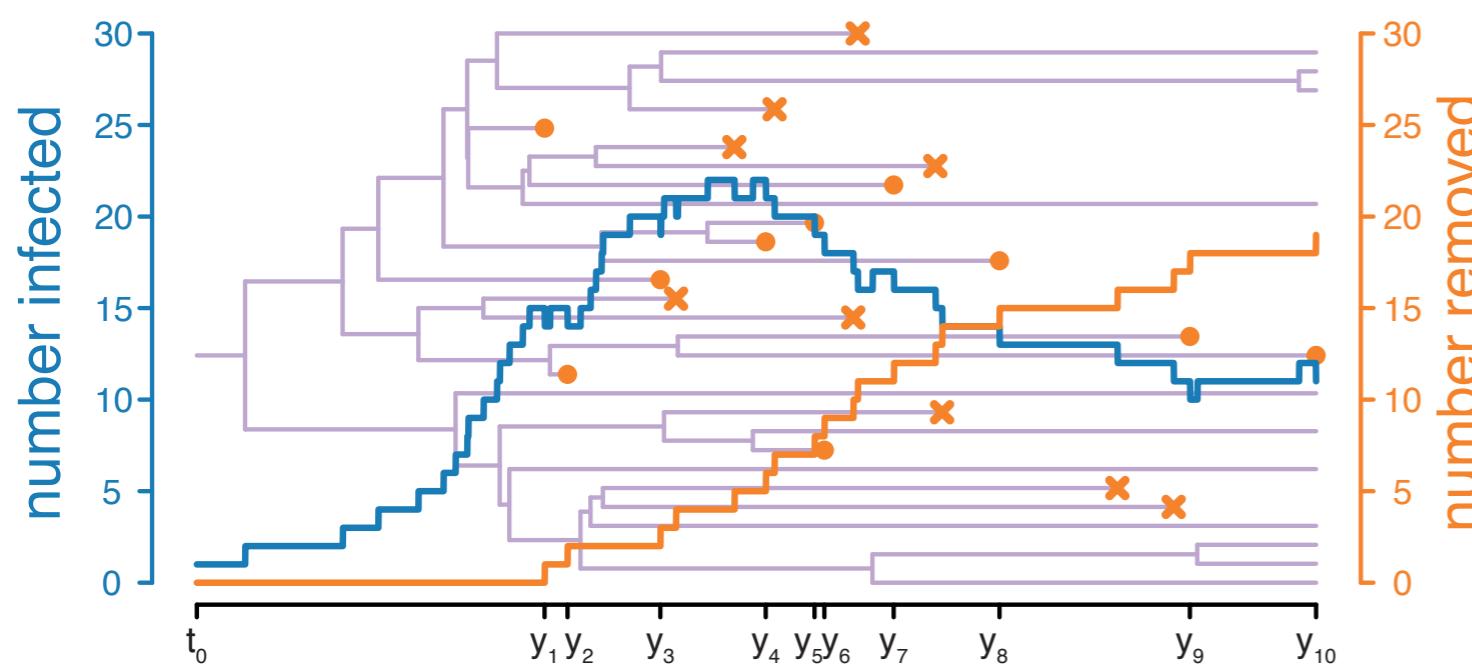
# Types of tree-generating models

**Birth-death →**

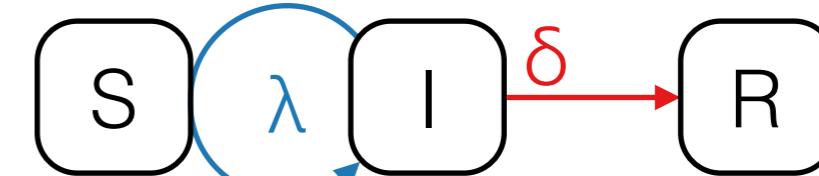


← **Coalescent**

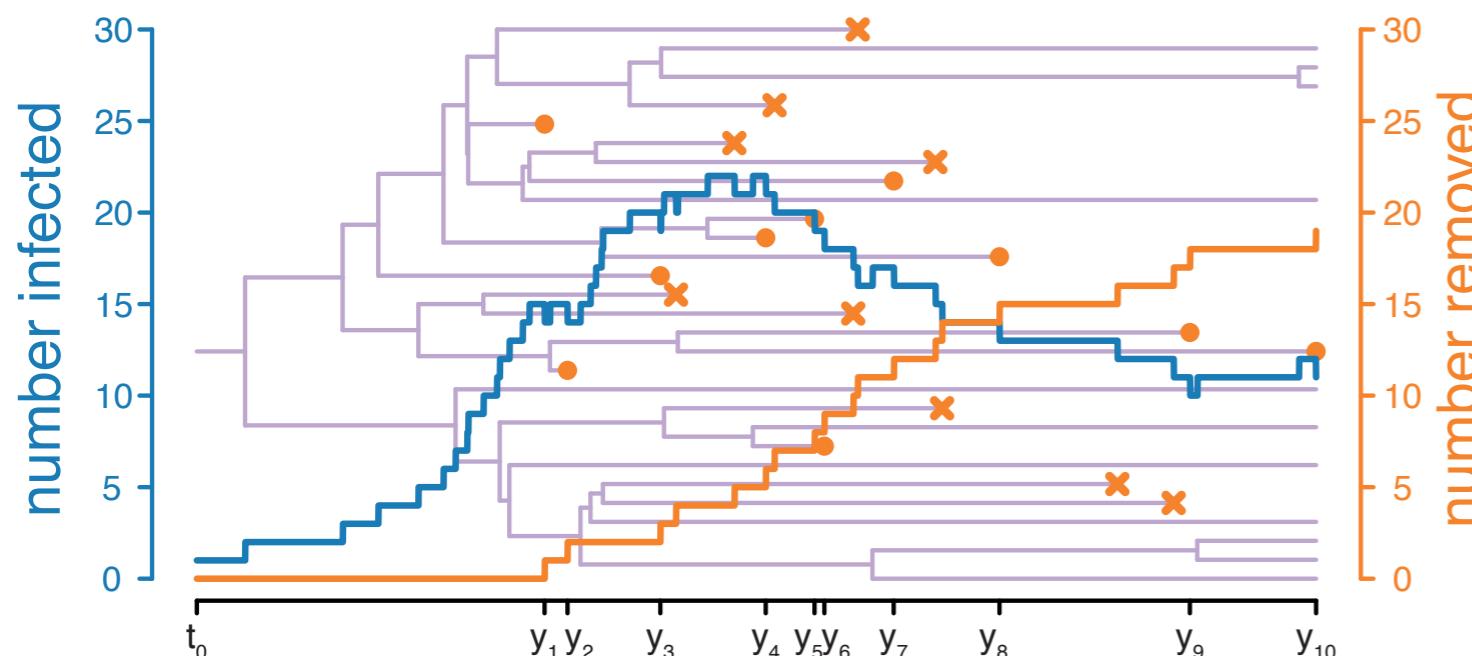
# Full vs. sampled tree



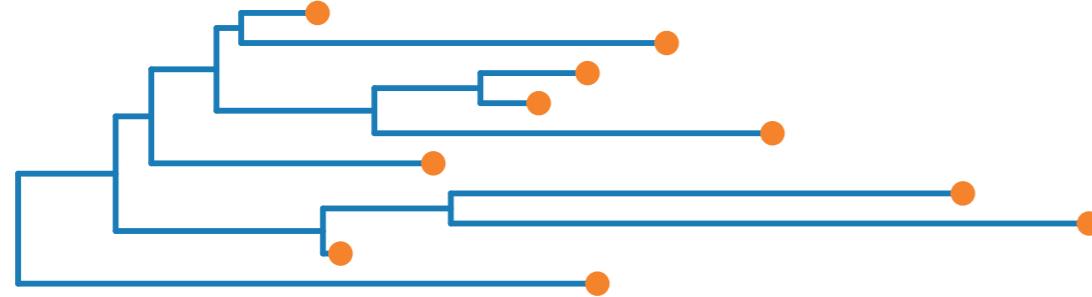
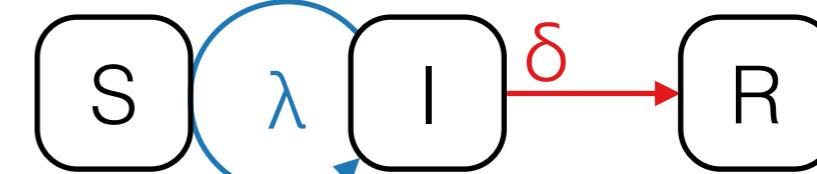
**Full tree  
generated  
by an SIR model**



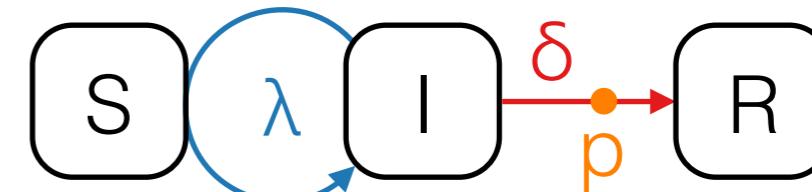
# Full vs. sampled tree



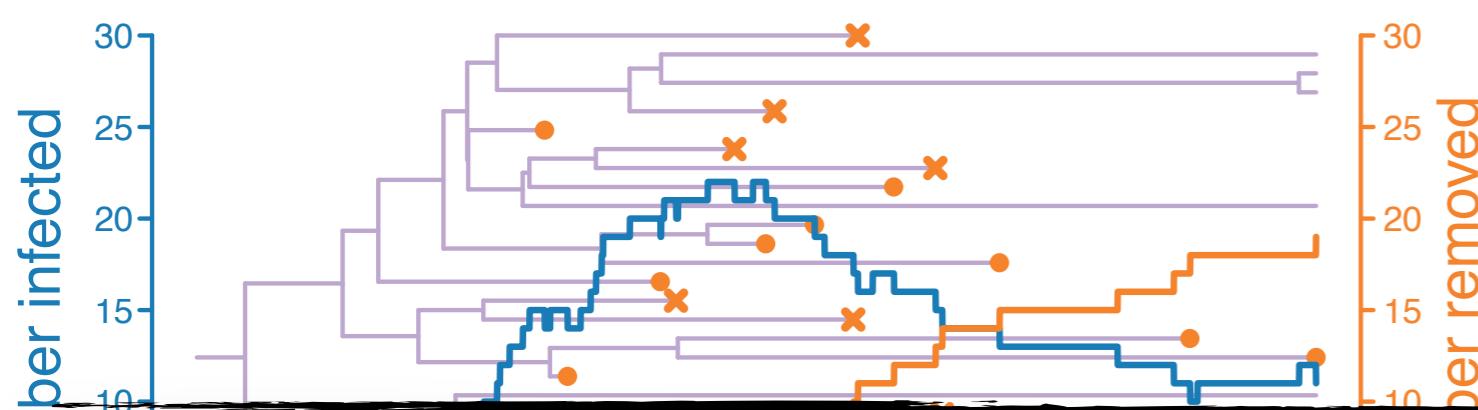
**Full tree  
generated  
by an SIR model**



**Sampled tree**

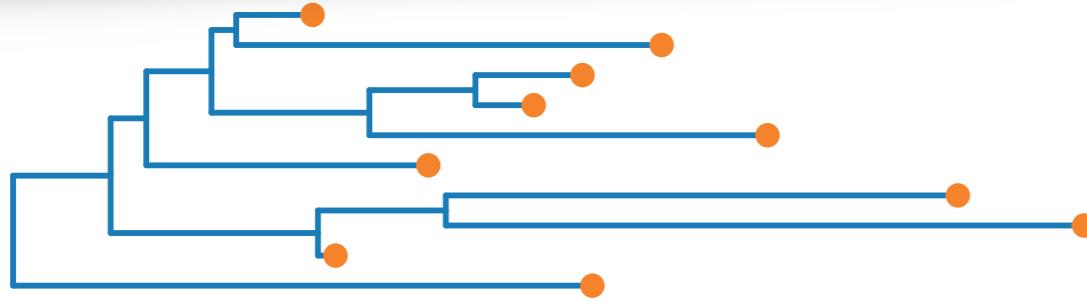


# Full vs. sampled tree

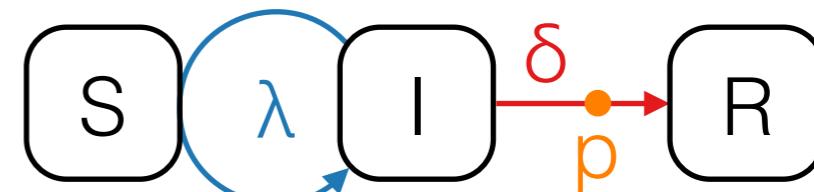


**Full tree  
generated  
by an SIR model**

**Introduce an explicit sampling process**

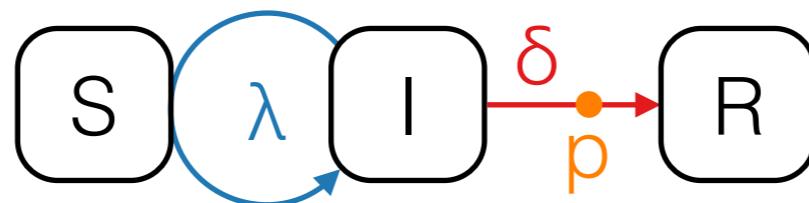


**Sampled tree**



# The birth-death **serial sampling** model

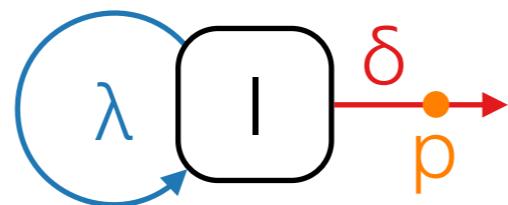
---



- $\lambda$  — infection rate
- $\delta$  — becoming-noninfectious rate
- $p$  — sampling probability
- Forward in time branching process that generates a tree
- Assume  $\lambda$  and  $\delta$  are exponentially distributed rates

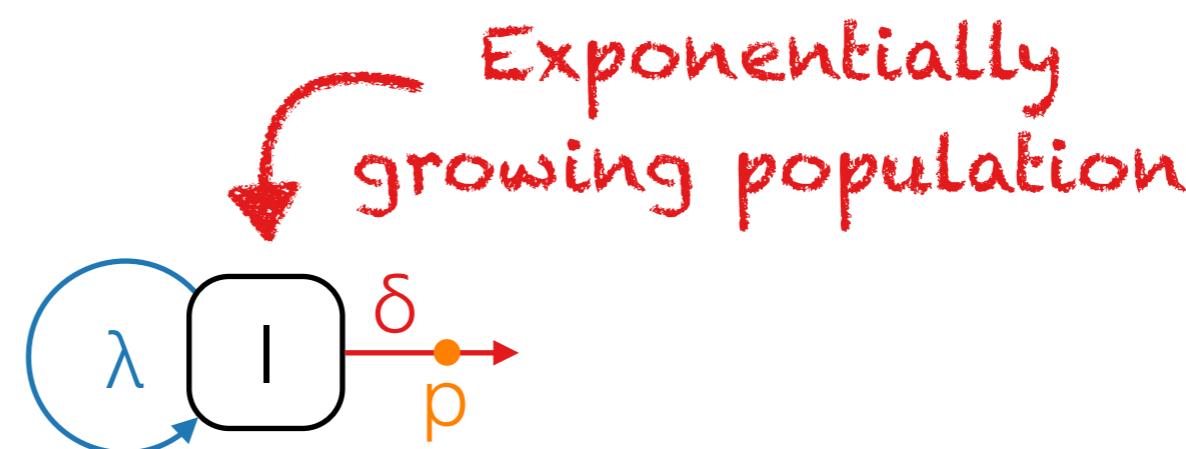
# The birth-death **serial sampling** model

---



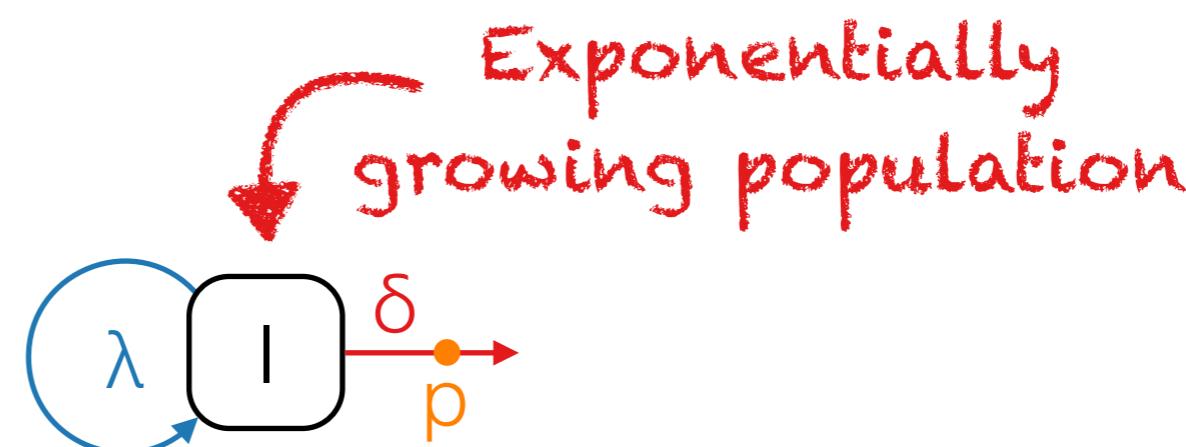
- $\lambda$  — infection rate
- $\delta$  — becoming-noninfectious rate
- $p$  — sampling probability
- Forward in time branching process that generates a tree
- Assume  $\lambda$  and  $\delta$  are exponentially distributed rates

# The birth-death **serial sampling** model

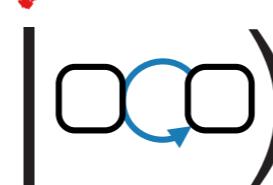


- $\lambda$  — infection rate
- $\delta$  — becoming-noninfectious rate
- $p$  — sampling probability
- Forward in time branching process that generates a tree
- Assume  $\lambda$  and  $\delta$  are exponentially distributed rates

# The birth-death **serial sampling** model

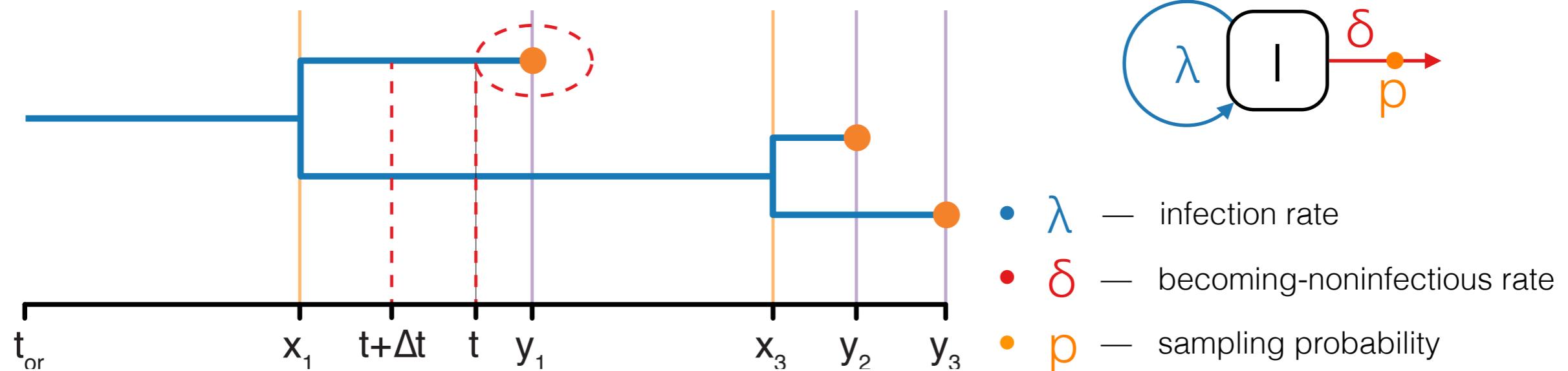


- $\lambda$  — infection rate
  - $\delta$  — becoming-noninfectious rate
  - p — sampling probability
- Forward in time branching process that generates a tree
  - Assume  $\lambda$  and  $\delta$  are exponentially distributed rates

P( | ) ?

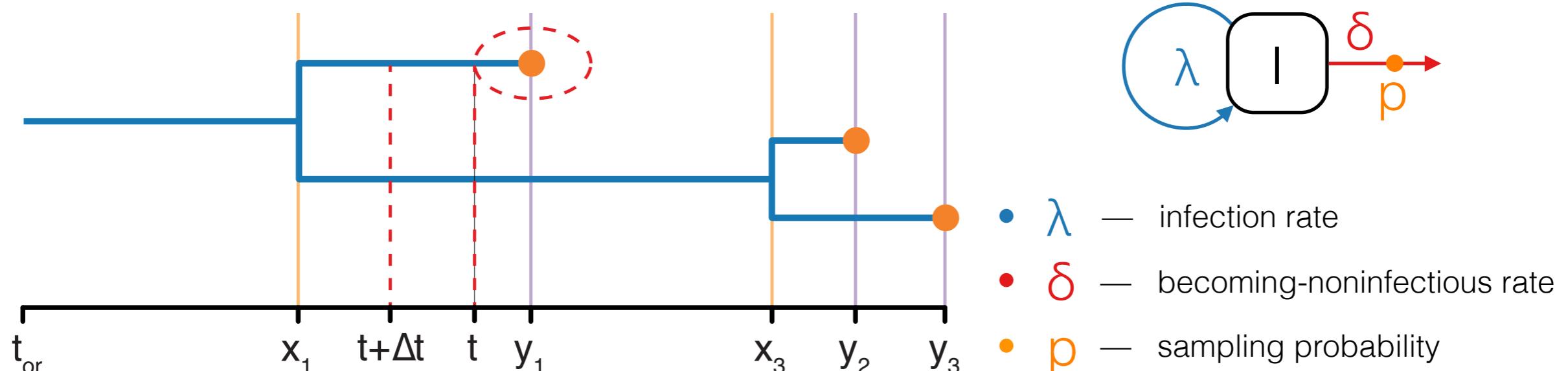
# The birth-death likelihood

P( $E | \infty$ ) ?



# The birth-death likelihood

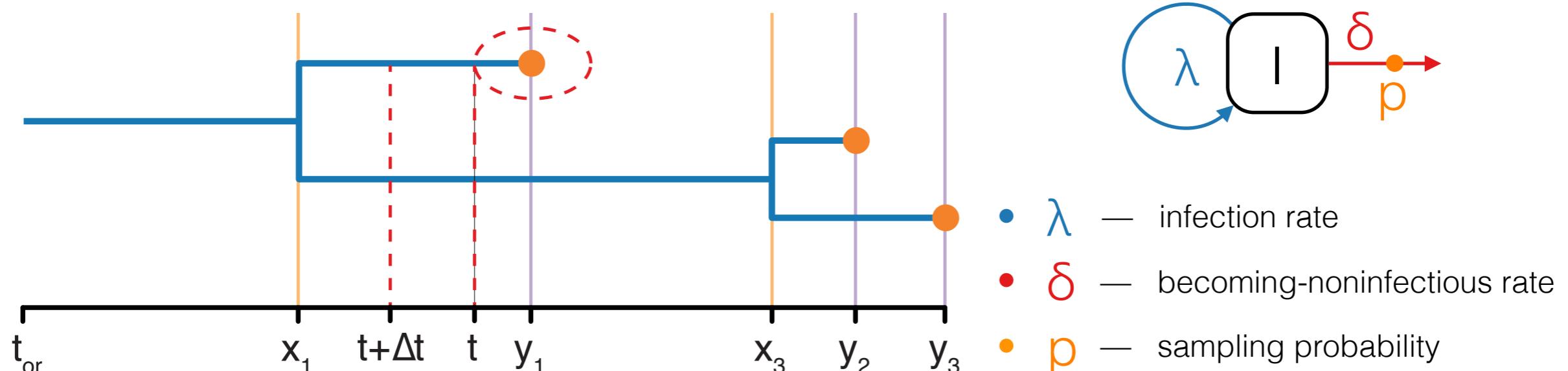
P( $E | \infty$ ) ?



- $p(t)$ : probability density that a given individual at time  $t$  produces the descending tree as observed

# The birth-death likelihood

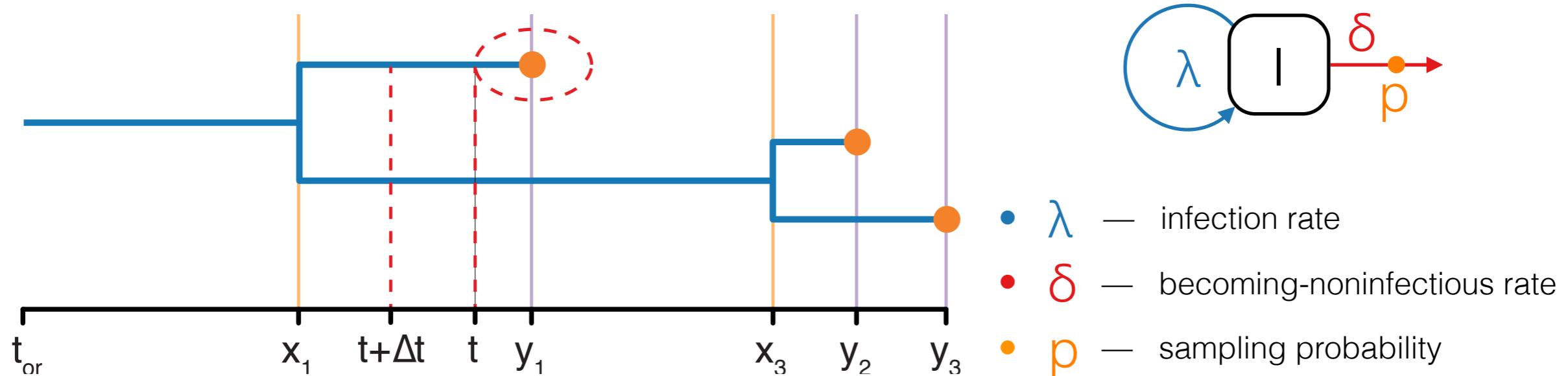
P( $E | \infty$ ) ?



- $p(t)$ : probability density that a given individual at time  $t$  produces the descending tree as observed
- $p_0(t)$ : probability density that a given individual at time  $t$  has no sampled descendants

# The birth-death likelihood

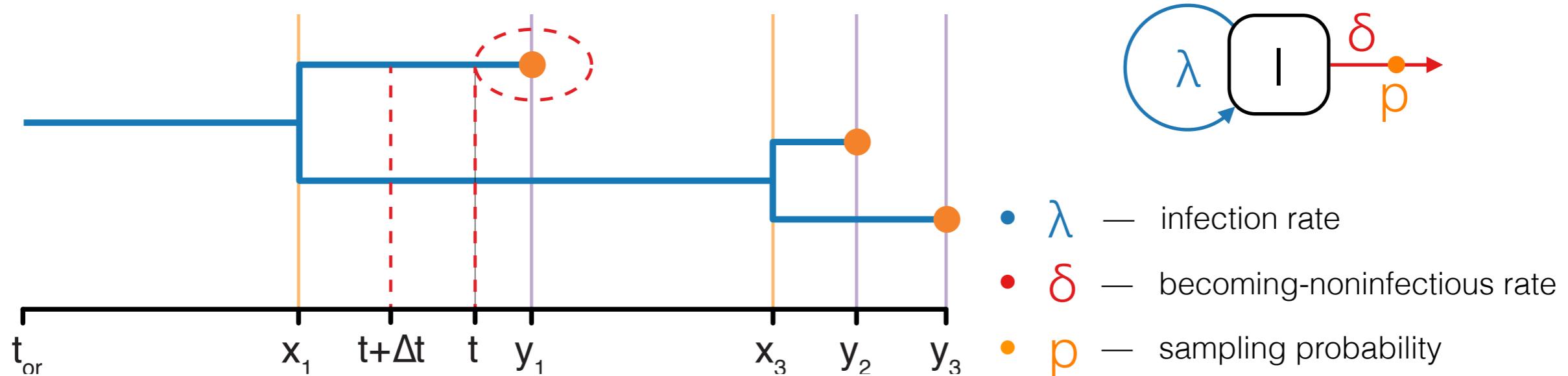
P( $E | \text{tree}$ ) ?



- $p(t)$ : probability density that a given individual at time  $t$  produces the descending tree as observed
- $p_0(t)$ : probability density that a given individual at time  $t$  has no sampled descendants
- $p(t_{\text{or}}) = P(\text{tree})$ : probability density of sampled tree with origin at time  $t_{\text{or}}$

# The birth-death likelihood

P( $E | \text{tree}$ ) ?

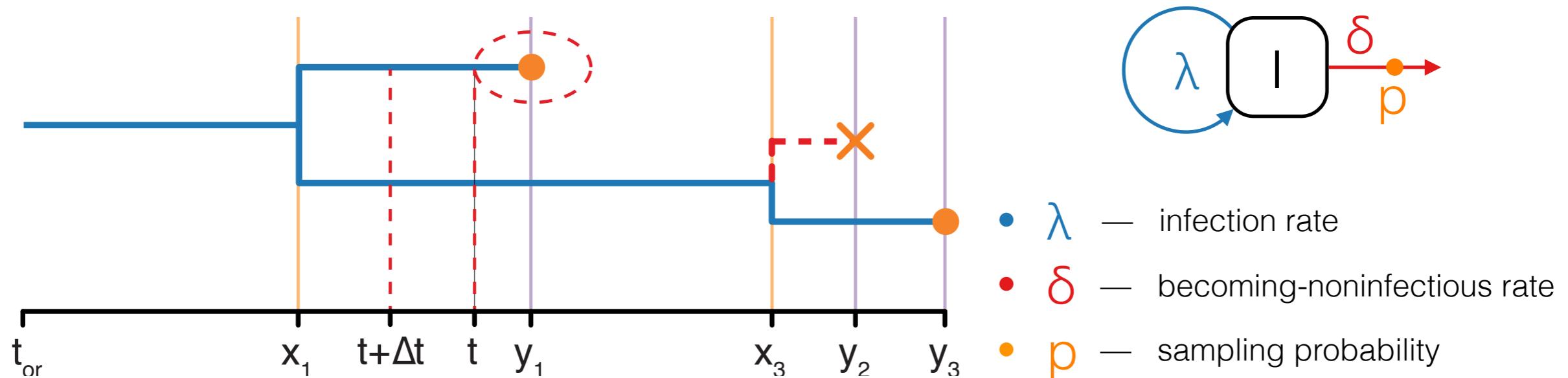


- $p(t)$ : probability density that a given individual at time  $t$  produces the descending tree as observed
- $p_0(t)$ : probability density that a given individual at time  $t$  has no sampled descendants
- $p(t_{\text{or}}) = P(\text{tree})$ : probability density of sampled tree with origin at time  $t_{\text{or}}$

Goal!

# Sampled branches — $p(t) \rightarrow p(t+\Delta t)$

---

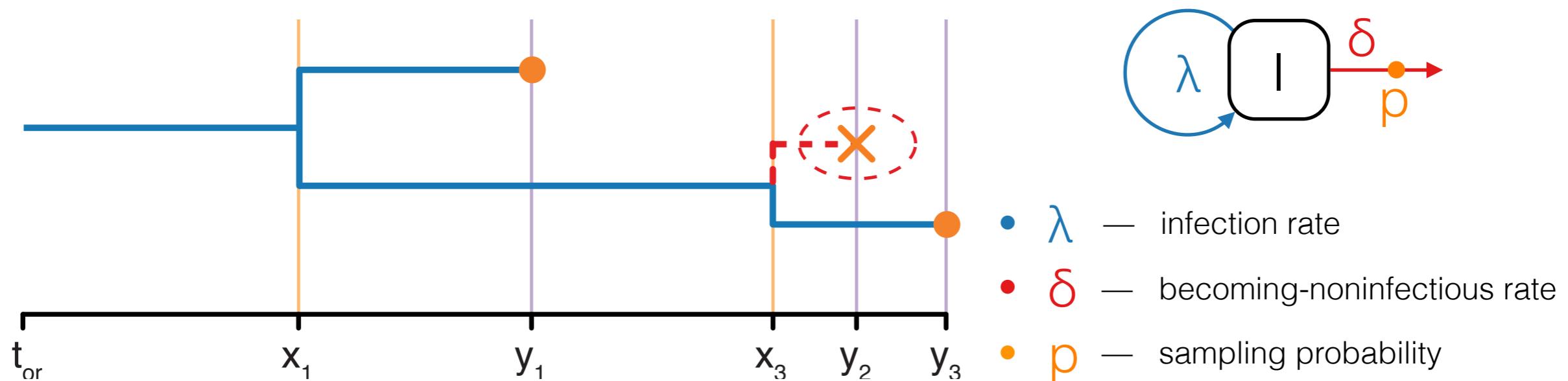


## Possible events in $\Delta t$ :

- Nothing happens:  $(1 - (\lambda + \delta)\Delta t)p(t)$
- A birth happens and is not observed:  $2\lambda p(t)p_0(t)^2\Delta t$

# Unsampled branches — $p_0(t) \rightarrow p_0(t+\Delta t)$

---



## Possible events in $\Delta t$ :

- Nothing happens:  $(1 - (\lambda + \delta)\Delta t)p_0(t)$
- A birth happens and is not observed:  $\lambda p_0(t)^2 \Delta t$
- A death happens and is not observed:  $\delta(1 - p)\Delta t$

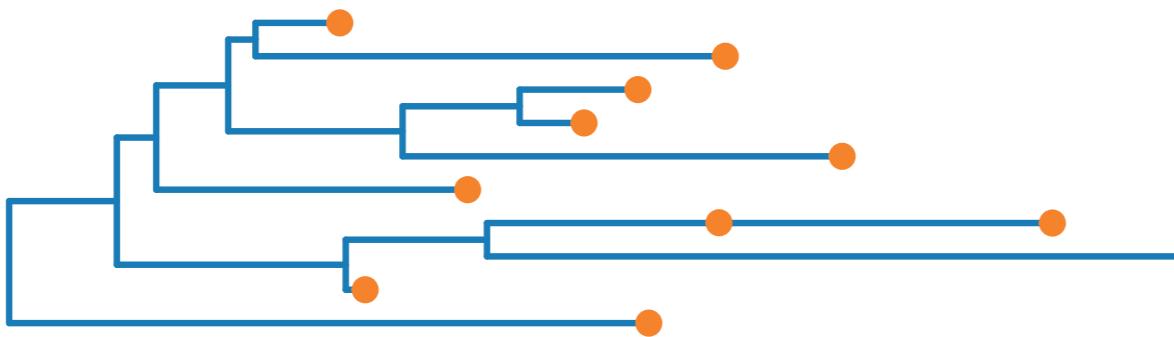
# Unsampled branches — $p_0(t) \rightarrow p_0(t+\Delta t)$

## Probability density of the tree

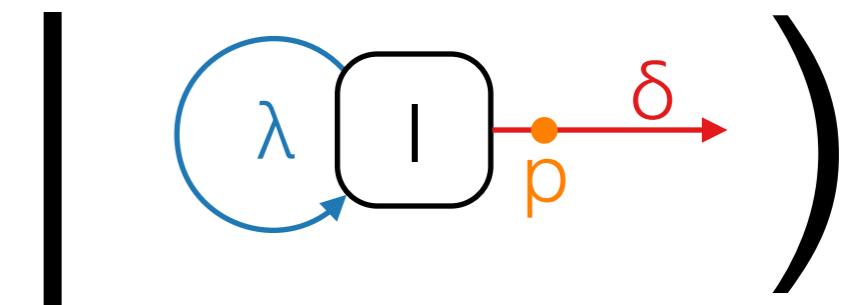
- Start with probability of the tree in the present,  $\mathbf{p(0)}$
  - $\mathbf{p(t)} \rightarrow \mathbf{p(t+\Delta t)}$
  - Adjust  $\mathbf{p(t)}$  at sampling or branching times
  - $\mathbf{p(t_{or}) = P(tree)}$
- 
- Nothing happens:  $(1 - (\lambda + \delta) \Delta t) p_0(t)$
  - A birth happens and is not observed:  $\lambda p_0(t)^2 \Delta t$
  - A death happens and is not observed:  $\delta(1 - p) \Delta t$

# Estimation of parameters in BEAST2

P(



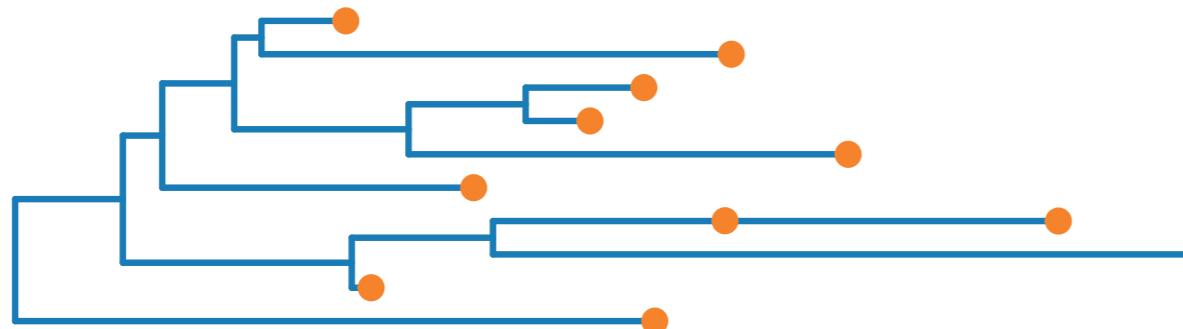
Describes growth of  
the transmission tree



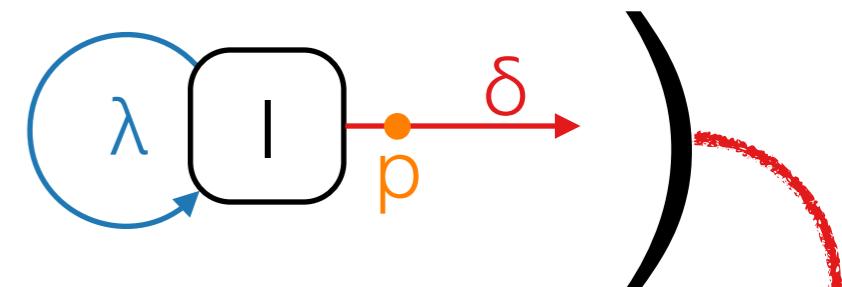
- $\lambda$  — infection rate
- $\delta$  — becoming-noninfectious rate
- $p$  — sampling probability
- $t_0$  — Origin of the process

# Estimation of parameters in BEAST2

P)



Describes growth of  
the transmission tree



$$P(E \cap O \cap C \cap B \cap A | ACAC, TCAC, ACAG) = \frac{P(ACAC, TCAC, ACAG) P(E | ACAC) P(O | TCAC) P(C | ACAG) P(B | E, O, C) P(A | E, O, C, B)}{P(ACAC, TCAC, ACAG)}$$

ACAC . . .  
TCAC . . .  
ACAG . . .  
genetic  
sequences

 genealogy

# demographic model

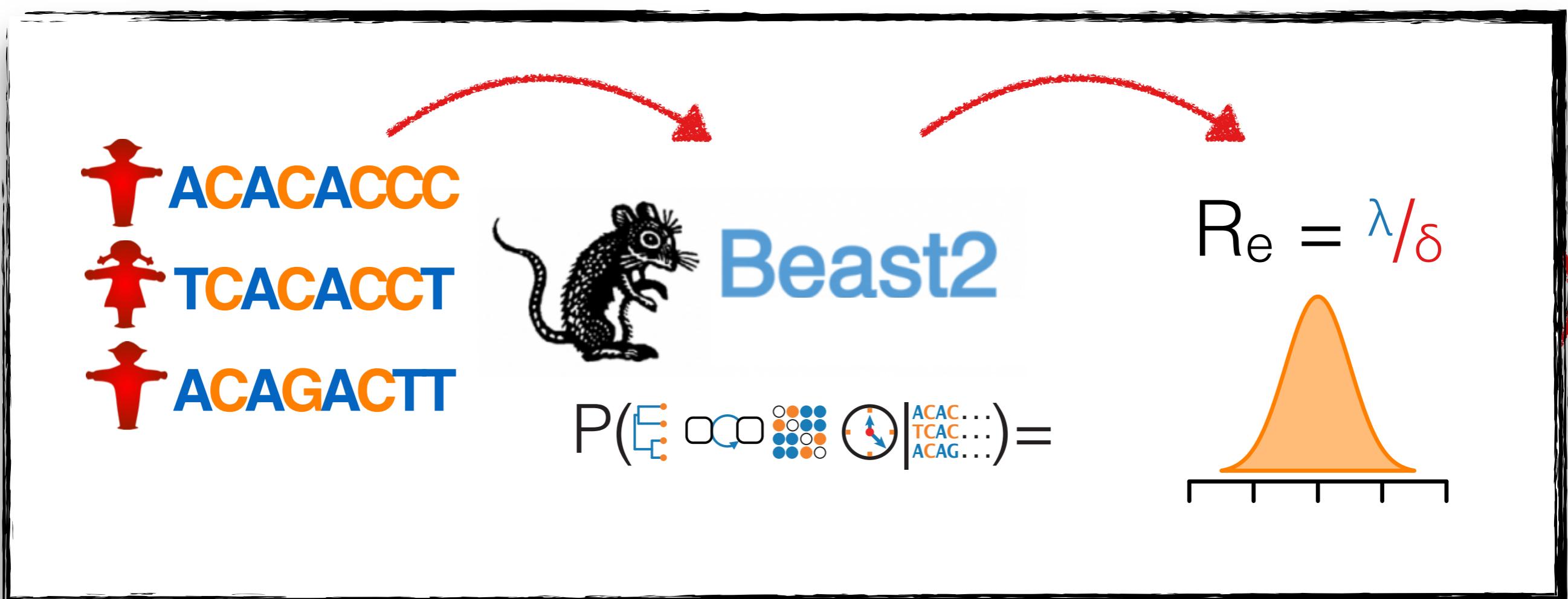


# substitution model



# molecular clock model

# Estimation of parameters in BEAST2



**ACAC...**  
**TCAC...**  
**ACAG...**

genetic sequences

  
genealogy

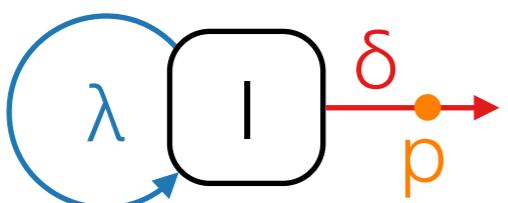
  
demographic model

  
substitution model

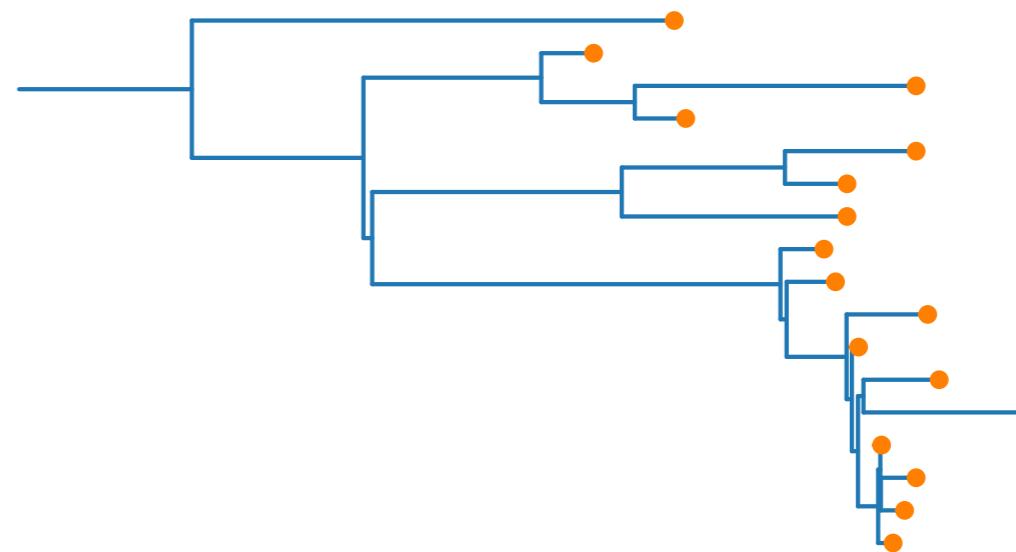
  
molecular clock model

# Birth-death skyline

---

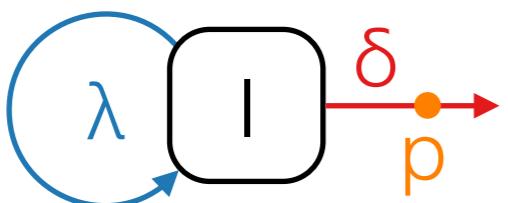


- $\lambda$  — infection rate
- $\delta$  — becoming-noninfectious rate
- $p$  — sampling probability

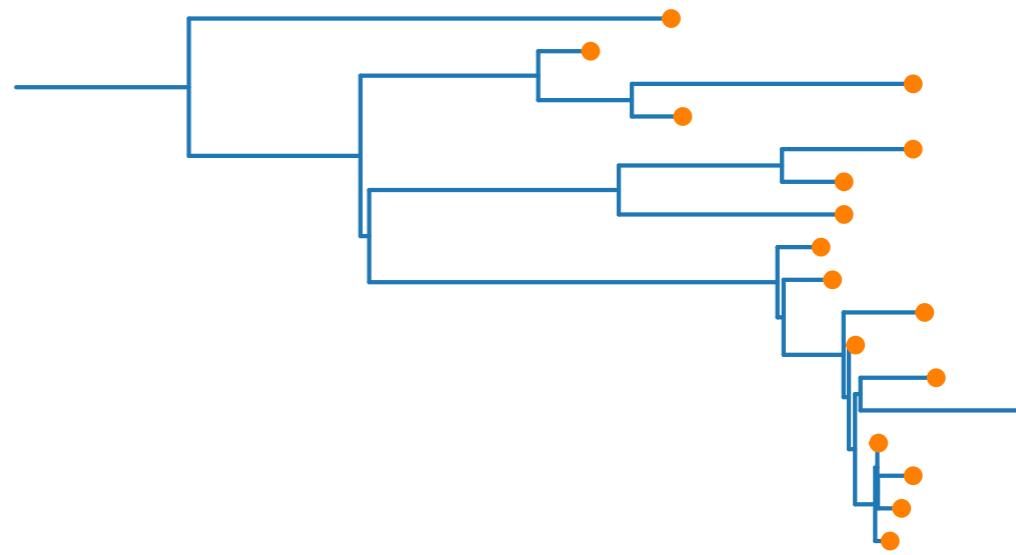


# Birth-death skyline

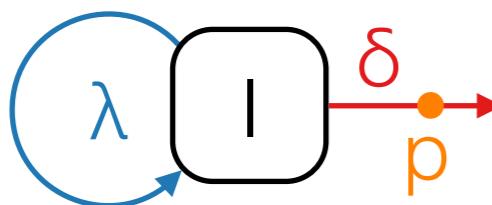
---



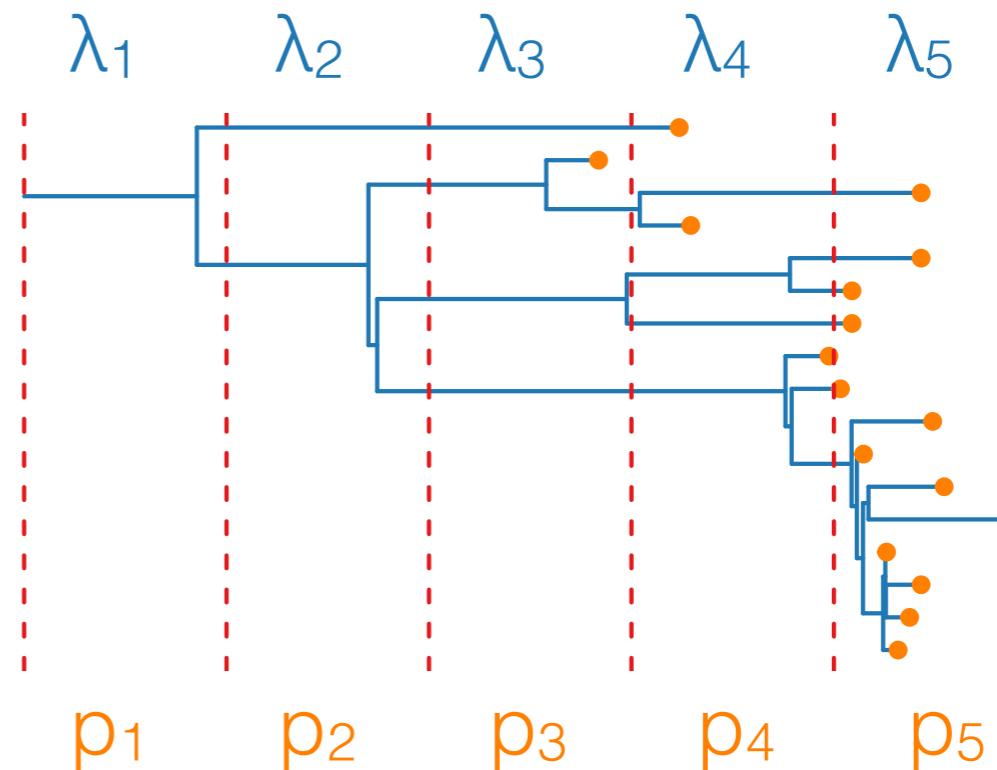
- $\lambda$  — infection rate
  - $\delta$  — becoming-noninfectious rate
  - $p$  — sampling probability
- 
- Model parameters are rates that describe the growth of the tree



# Birth-death skyline



- $\lambda$  — infection rate
- $\delta$  — becoming-noninfectious rate
- $p$  — sampling probability

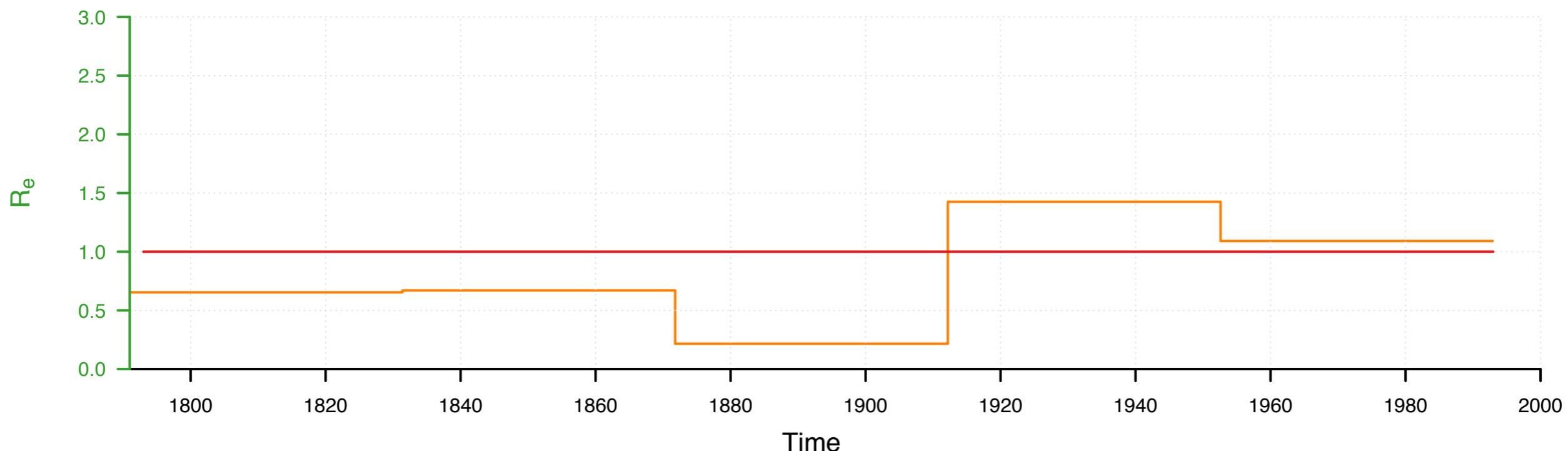


- Model parameters are rates that describe the growth of the tree
- Allow parameters to change through time
- Shifts in rates can be anywhere  
(usually evenly distributed between origin and present)
- Shift times for different parameters can be different!
- Very flexible model, but not all options are available in BEAUTi

# Birth-death skyline plot

---

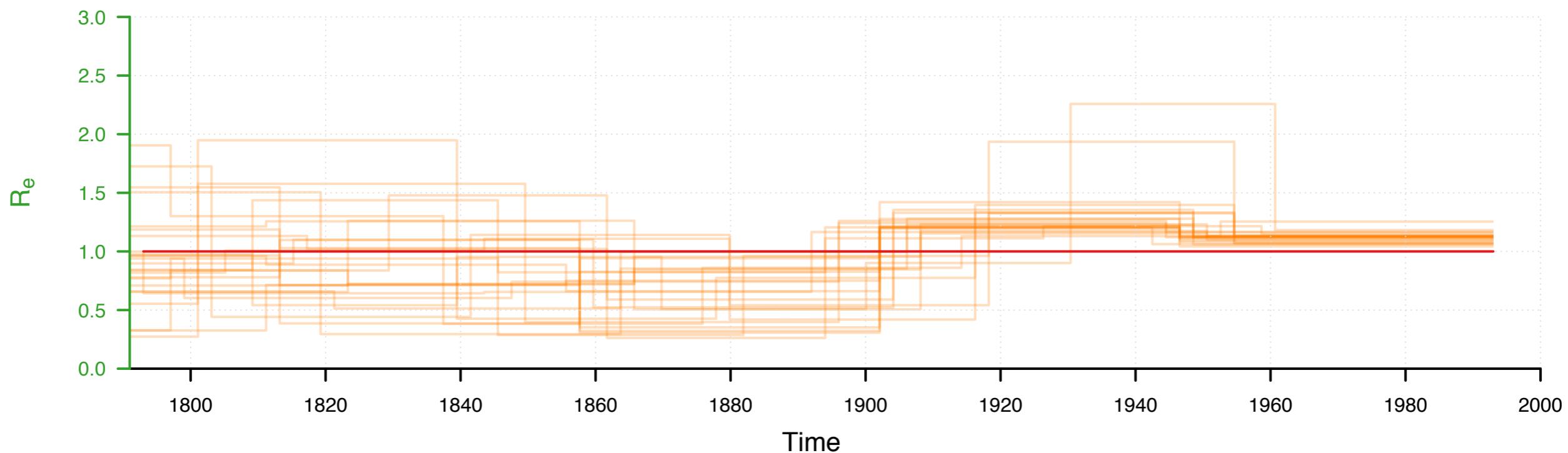
- 63 Egyptian HCV sequences from 1993
- Viral contamination by antischistosomiases injections from 1920s



# Birth-death skyline plot

---

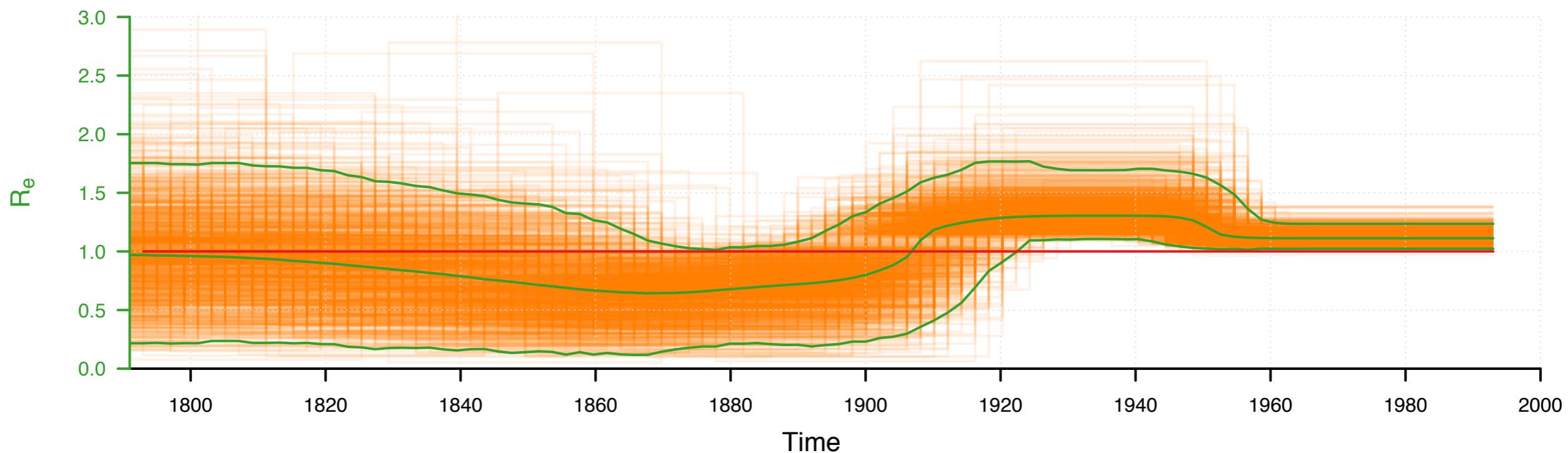
- 63 Egyptian HCV sequences from 1993
- Viral contamination by antischistosomiases injections from 1920s



# Birth-death skyline plot

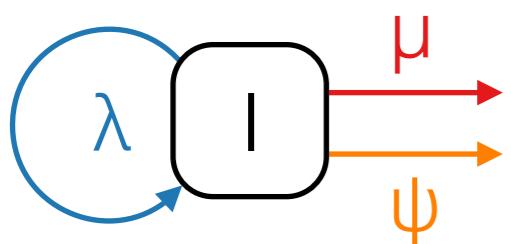
---

- 63 Egyptian HCV sequences from 1993
- Viral contamination by antischistosomiases injections from 1920s



# Model Parameterization: Birth-death process

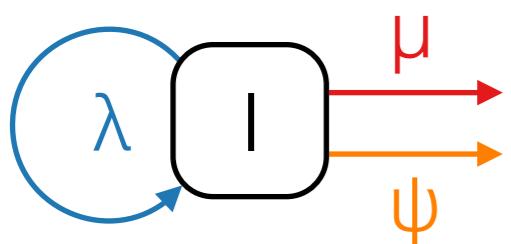
---



- $\lambda$  — Birth-rate (speciation / transmission)
- $\mu$  — Death-rate (extinction / death)
- $\psi$  — Sampling rate

# Model Parameterization: Birth-death process

---

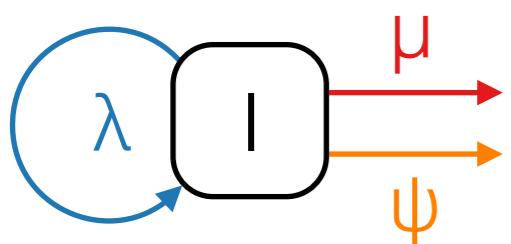


- $\lambda$  — Birth-rate (speciation / transmission)  $\lambda \in [0, \infty)$
- $\mu$  — Death-rate (extinction / death)  $\mu \in [0, \infty)$
- $\psi$  — Sampling rate  $\psi \in [0, \infty)$

Priors?

# Model Parameterization: Birth-death process

---



- $\lambda$  — Birth-rate (speciation / transmission)  $\lambda \in [0, \infty)$
- $\mu$  — Death-rate (extinction / death)  $\mu \in [0, \infty)$
- $\psi$  — Sampling rate  $\psi \in [0, \infty)$

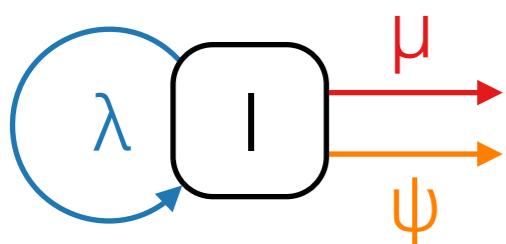
Priors?

## Infectious diseases

- $R = \lambda/(\mu + \psi)$  — Reproduction number  
(is it spreading or not?)
- $\delta = \mu + \psi$  — Becoming noninfectious rate  
( $1/\delta$  = infectious period)
- $p = \psi/(\mu + \psi)$  — Sampling proportion

# Model Parameterization: Birth-death process

---



- $\lambda$  — Birth-rate (speciation / transmission)  $\lambda \in [0, \infty)$
- $\mu$  — Death-rate (extinction / death)  $\mu \in [0, \infty)$
- $\psi$  — Sampling rate  $\psi \in [0, \infty)$

Priors?

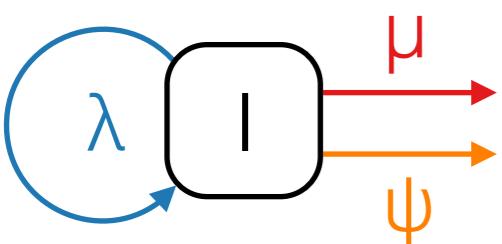
## Infectious diseases

- $R = \lambda/(\mu + \psi)$  — Reproduction number  
(is it spreading or not?)
- $\delta = \mu + \psi$  — Becoming noninfectious rate  
( $1/\delta$  = infectious period)
- $p = \psi/(\mu + \psi)$  — Sampling proportion

$$R \in [0, \infty) \quad \delta \in [0, \infty) \quad p \in [0, 1]$$

More natural priors!

# Model Parameterization: Birth-death process



- $\lambda$  — Birth-rate (speciation / transmission)  $\lambda \in [0, \infty)$
  - $\mu$  — Death-rate (extinction / death)  $\mu \in [0, \infty)$
  - $\psi$  — Sampling rate  $\psi \in [0, \infty)$

## Priors?

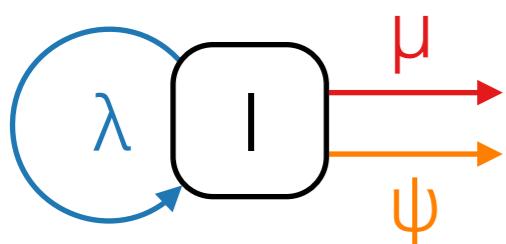
# Infectious diseases

- $R = \lambda / (\mu + \psi)$  — Reproduction number  
(is it greater than one or not?)
  - $\delta = \mu + \psi - R$  — Gross instantaneous rate of change of the infected proportion (per infectious period)
  - $R = \frac{\delta}{\mu + \psi}$  — Proportion of the population that is infected

BDSKY package

# More natural priors!

# Model Parameterization: Birth-death process



- $\lambda$  — Birth-rate (speciation / transmission)  $\lambda \in [0, \infty)$
- $\mu$  — Death-rate (extinction / death)  $\mu \in [0, \infty)$
- $\psi$  — Sampling rate  $\psi \in [0, \infty)$

Priors?

## Infectious diseases

- $R = \lambda/(\mu + \psi)$  — Reproductive number  
(is it a growth rate or not?)
- $\delta = \mu + \psi$  — Infectious rate  
(per unit time period)
- $p$  — Sampling proportion

BDSKY package

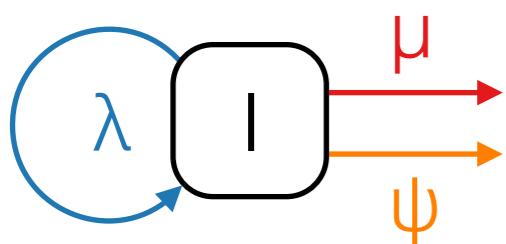
$$R \quad \delta \in [0, \infty) \quad p \in [0, 1]$$

## Speciation

- $\lambda - \mu$  — Growth rate
- $\mu/\lambda$  — Relative death rate
- $p = \psi/(\mu + \psi)$  — Sampling proportion

More natural priors!

# Model Parameterization: Birth-death process



- $\lambda$  — Birth-rate (speciation / transmission)  $\lambda \in [0, \infty)$
- $\mu$  — Death-rate (extinction / death)  $\mu \in [0, \infty)$
- $\psi$  — Sampling rate  $\psi \in [0, \infty)$

Priors?

## Infectious diseases

- $R = \lambda/(\mu + \psi)$  — Reproductive number  
(is it a growth rate or not?)
- $\delta = \mu + \psi - \lambda$  — Infectious rate  
(is it a growth period)
- $p$  — Sampling proportion

BDSKY package

$$R \quad \delta \in [0, \infty) \quad p \in [0, 1]$$

More natural priors!

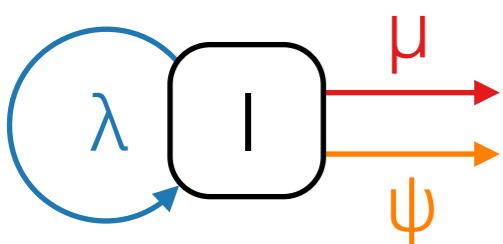
## Speciation

- $\lambda - \mu$  — Growth rate
- $\mu/\lambda$  — Relative death rate
- $p = \psi/(\mu + \psi)$  — Sampling proportion

$$(\lambda - \mu) \in [0, \infty) \quad \mu/\lambda \in [0, 1] \quad p \in [0, 1]$$

Smaller parameter space!

# Model Parameterization: Birth-death process



- $\lambda$  — Birth-rate (speciation / transmission)  $\lambda \in [0, \infty)$
- $\mu$  — Death-rate (extinction / death)  $\mu \in [0, \infty)$
- $\psi$  — Sampling rate  $\psi \in [0, \infty)$

Priors?

## Infectious diseases

- $R = \lambda/(\mu+\psi)$  — Reproductive number (is it a growth period?)
- $\delta = \mu + \psi - \lambda$  — Infectious rate
- $r = \psi/\lambda$  — Sampling proportion

BDSKY package

$$R \quad \delta \in [0, \infty) \quad p \in [0, 1]$$

More natural priors!

## Speciation

- $\lambda - \mu$  — Growth rate
- $\mu/\lambda$  — Relative extinction rate
- $p = \psi/\lambda$  — Sampling proportion

FBD package

$$(\lambda - \mu) \quad \mu/\lambda \in [0, 1] \quad p \in [0, 1]$$

Smaller parameter space!



news    tutorials    contribute

# Skyline plots

Inference of past population dynamics using Bayesian Coalescent Skyline and Birth-Death Skyline plots.  
by Nicola F. Müller and Louis du Plessis

**Tutorial**

- [Skyline-plots.pdf](#)
- [Github repository](#)
- [License](#)
- [Statistics](#)

**Data**

- [hcv.nexus](#)

**XML**

- [hcv\\_bdsky.xml](#)
- [hcv\\_coal.xml](#)

**Scripts**

- [Figure\\_Utils.R](#)
- [LogFile\\_Utils.R](#)
- [Skyline\\_Example.R](#)
- [SkylinePlot.R](#)

---

## Background

Population dynamics influence the shape of the tree and consequently, the shape of the tree contains some information about past population dynamics. The so-called Skyline methods allow to extract this information from phylogenetic trees in a non-parametric manner. It is non-parametric since there is no underlying system of differential equations governing the inference of these dynamics. In this tutorial we will look at two different methods to infer these dynamics from sequence data. The first one is the Bayesian Coalescent Skyline plot (Drummond, Rambaut, Shapiro, & Pybus, 2005), which is based on the coalescent model, and the second one is the Birth-Death skyline (Stadler, Kuhnert, Bonhoeffer, & Drummond, 2013) plot based on the birth-death model. The conceptual difference between coalescent and birth-death approaches lies in the direction of the flow of time. In the coalescent, the time is modeled to go backwards, from present to past, while in the birth-death approach it is modeled to go forwards. Two other fundamental differences are the parameters that are inferred and the way sampling is treated.

## Programs used in this Exercise

[about](#)    [workshops](#)    [contact](#)    [license](#)

powered by GitHub and Jekyll





news tutorials contribute

# Skyline plots

Inference of past population dynamics using Bayesian Coalescent Skyline plots.  
by Nicola F. Müller and Louis du Plessis

**Tutorial**

- [Skyline-plots.pdf](#)
- [Github repository](#)
- [License](#)
- [Statistics](#)

**Data**

- [hcv.nexus](#)

**XML**

- [hcv\\_bdsky.xml](#)
- [hcv\\_coal.xml](#)

**Scripts**

- [Figure\\_Utils.R](#)
- [LogFile\\_Utils.R](#)
- [Skyline\\_Example.R](#)
- [SkylinePlot.R](#)

## Background

Population dynamics inferred from a phylogenetic tree. The shape of the tree contains some information about the evolution of the population. These methods allow to extract this information from the tree. This is non-parametric since there is no underlying system of differential equations. In this tutorial we will look at two different methods to infer population dynamics. One is the Bayesian Coalescent Skyline plot (Drummond, Rambaut, & Suchard, 2013) based on the coalescent model, and the second one is the Birth-Death skyline plot (Drummond, 2013) based on the birth-death model. The conceptual difference between the two approaches lies in the direction of the flow of time. In the coalescent, the time is modeled to go backwards, from present to past, while in the birth-death approach it is modeled to go forwards. Two other important differences are the parameters that are inferred and the way sampling is treated.

## Programs used in this Exercise

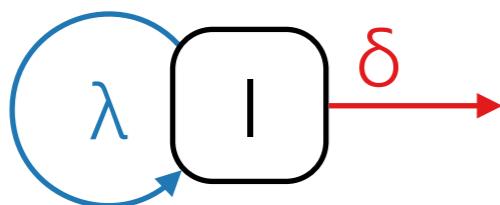
[about](#) [workshops](#) [contact](#) [license](#)

powered by GitHub and Jekyll 

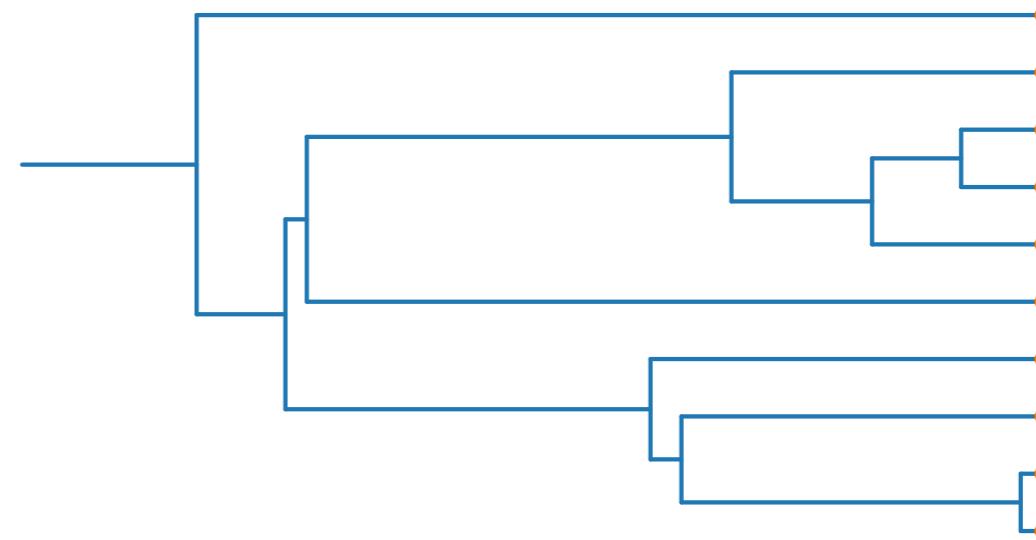
*Birth-death skyline contemporary*

# Birth-death skyline contemporary

---



- $\lambda$  — infection rate
- $\delta$  — becoming-noninfectious rate
- $\rho$  — present day sampling proportion



- No sampling through time
- Sample some proportion of the infected population at one time point
- Can also be applied as a speciation tree-prior (speciation/extinction rate)
- Need to calibrate the clock somehow



news    tutorials    contribute

# Divergence Time Estimation using BEAST v2.x

## Dating Species Divergences with the Fossilized Birth-Death Process

by Tracy A. Heath

**Tutorial**

-  [FBD-tutorial.pdf](#)
-  [Github repository](#)
-  [License](#)
-  [Statistics](#)

Central among the questions explored in biology are those that seek to understand the timing and rates of evolutionary processes. Accurate estimates of species divergence times are vital to understanding historical biogeography, estimating diversification rates, and identifying the causes of variation in rates of molecular evolution.

This tutorial will provide a general overview of divergence time estimation and fossil calibration using a stochastic branching process and relaxed-clock model in a Bayesian framework. The exercise will guide you through the steps necessary for estimating phylogenetic relationships and dating species divergences using the program BEAST v2.x.

Adapted from <http://treethinkers.org/tutorials/divergence-time-estimation-using-beast/>

**Data**

-  [bears\\_cytb\\_fossils.nex](#)
-  [bears\\_irbp\\_fossils.nex](#)

**XML**

-  [bearsDivtime\\_FBD.1.xml](#)
-  [bearsDivtime\\_FBD.2.xml](#)

**Scripts**

[about](#)    [workshops](#)    [contact](#)    [license](#)

powered by GitHub and Jekyll



taming-the-beast.github.io/tutorials/FBD-tutorial/

Taming the BEAST news tutorials contribute

# Divergence Time Estimation v2.x

## Dating Species Divergences with the Fossilized Birth-Death Process

by Tracy A. Heath

**Tutorial**

- [FBD-tutorial.pdf](#)
- [Github repository](#)
- [License](#)
- [Statistics](#)

Central among the questions in evolutionary biology is understanding the timing and rates of evolutionary processes. These times are vital to understanding historical biogeography and the causes of variation in rates of molecular evolution.

This tutorial will guide you through divergence time estimation and fossil calibration using a stochastic birth-death process in a Bayesian framework. The exercise will guide you through the steps of calibrating phylogenetic relationships and dating species divergences using the program BEAST v2.x.

[View tutorial](#)

**Data**

- [bears\\_cytb\\_fossils.nex](#)
- [bears\\_irbp\\_fossils.nex](#)

**XML**

- [bearsDivtime\\_FBD.1.xml](#)
- [bearsDivtime\\_FBD.2.xml](#)

**Scripts**

about workshops contact license

powered by GitHub and Jekyll



*Fossilised birth-death process*



news    tutorials    contribute

# Prior selection

Prior selection and clock calibration using Influenza A data  
by Veronika Bošková and Venelin Mitov

**Tutorial**

- [Prior-selection.pdf](#)
- [Github repository](#)
- [License](#)
- [Statistics](#)

**Data**

- [InfluenzaAH3N2\\_HAgene\\_2009....](#)
- [InfluenzaAH3N2\\_HAgene\\_2009....](#)

**XML**

- [Heterochronous.xml](#)
- [Homochronous.xml](#)
- [Homochronous\\_tMRCA.xml](#)

**Output**

- [InfluenzaAH3N2\\_HAgene\\_2009....](#)
- [InfluenzaAH3N2\\_HAgene\\_2009....](#)

## Background

In the Bayesian analysis of sequence data, priors play an important role. When wrongly selected, the runs may take very long to converge, not converge at all or cause a bias in the inferred trees and model parameters. Selection of proper priors and starting values is crucial and can be a difficult exercise at the start. It is not always easy to pick a proper model of tree generation (tree prior), substitution model, molecular clock model or the prior distribution for an unknown parameter.

The molecular clock model aims to estimate the substitution rate of the data. It is important to understand under which circumstances to use which model and when molecular calibration works. This will help the investigator determine which estimates of parameters can be trusted and which can not.

In this tutorial, we will explore how priors are selected and how molecular clock calibration works using the H3N2 influenza A data from the flu virus spreading in the USA in 2009.

## Programs used in this Exercise

[about](#)    [workshops](#)    [contact](#)    [license](#)

powered by GitHub and Jekyll



news tutorials contribute

# Prior selection

Prior selection and clock calibration using Influenza A data  
by Veronika Bošková and Venelin Mitov

**Tutorial**

- [Prior-selection.pdf](#)
- [Github repository](#)
- [License](#)
- [Statistics](#)

**Data**

- [InfluenzaAH3N2\\_HAgene\\_2009...](#)
- [InfluenzaAH3N2\\_HAgene\\_2009...](#)

**XML**

- [Heterochronous.xml](#)
- [Homochronous.xml](#)
- [Homochronous\\_tMRCA.xml](#)

**Output**

- [InfluenzaAH3N2\\_HAgene...](#)
- [InfluenzaAH3N2\\_HAgene...](#)

## Background

In the Bayesian analysis, prior selection is a critical step. When wrongly selected, the runs may take very long to converge. It is important to understand which priors are appropriate for the data and which are not. Selection of proper priors requires a good understanding of the data and the system being analyzed. This is a difficult exercise at the start. It is not always easy to pick a proper prior for a specific parameter, such as the substitution rate or the molecular clock model or the prior distribution for an unknown parameter.

## Programs used in this Exercise

*Birth-death skyline serial*

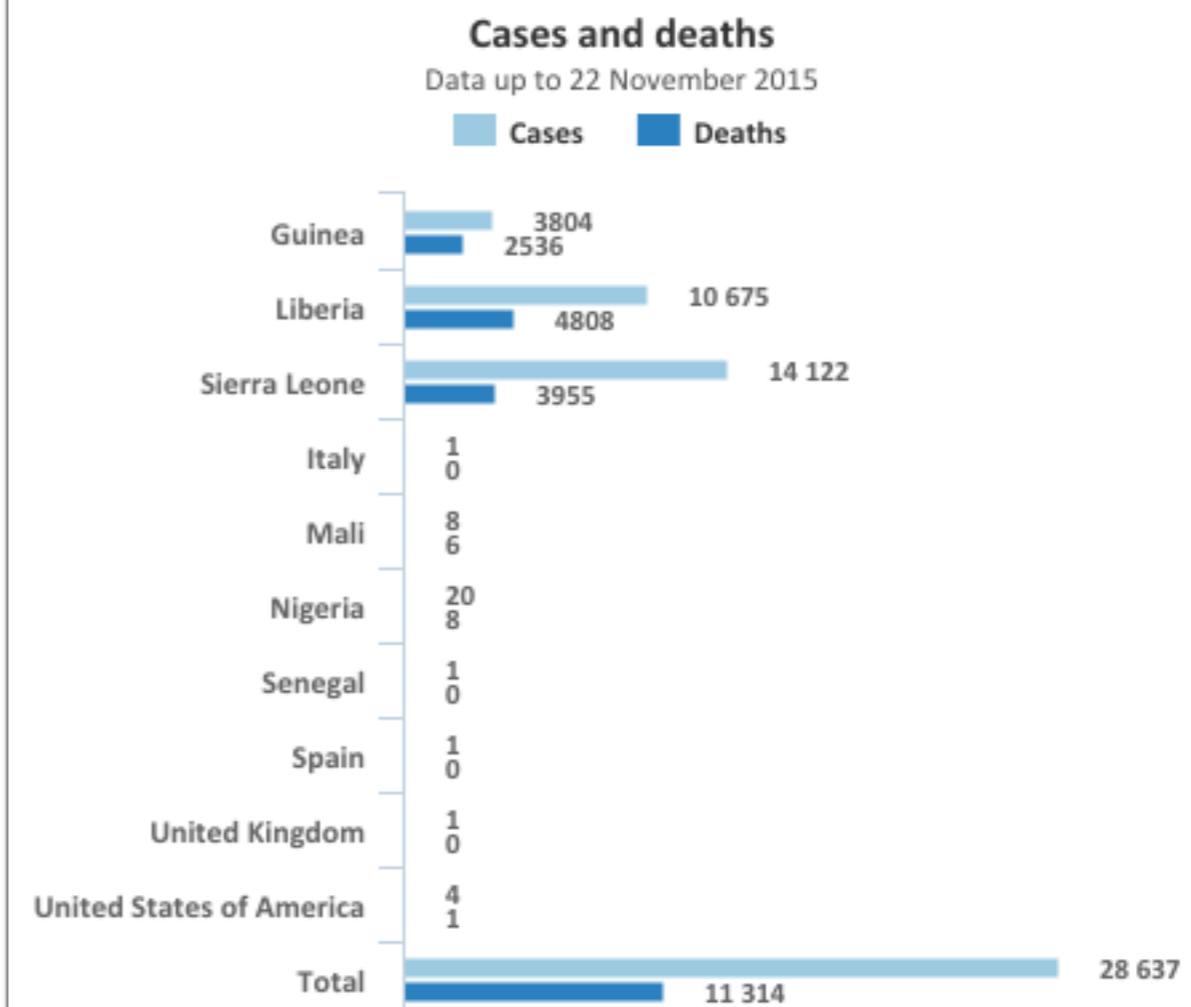
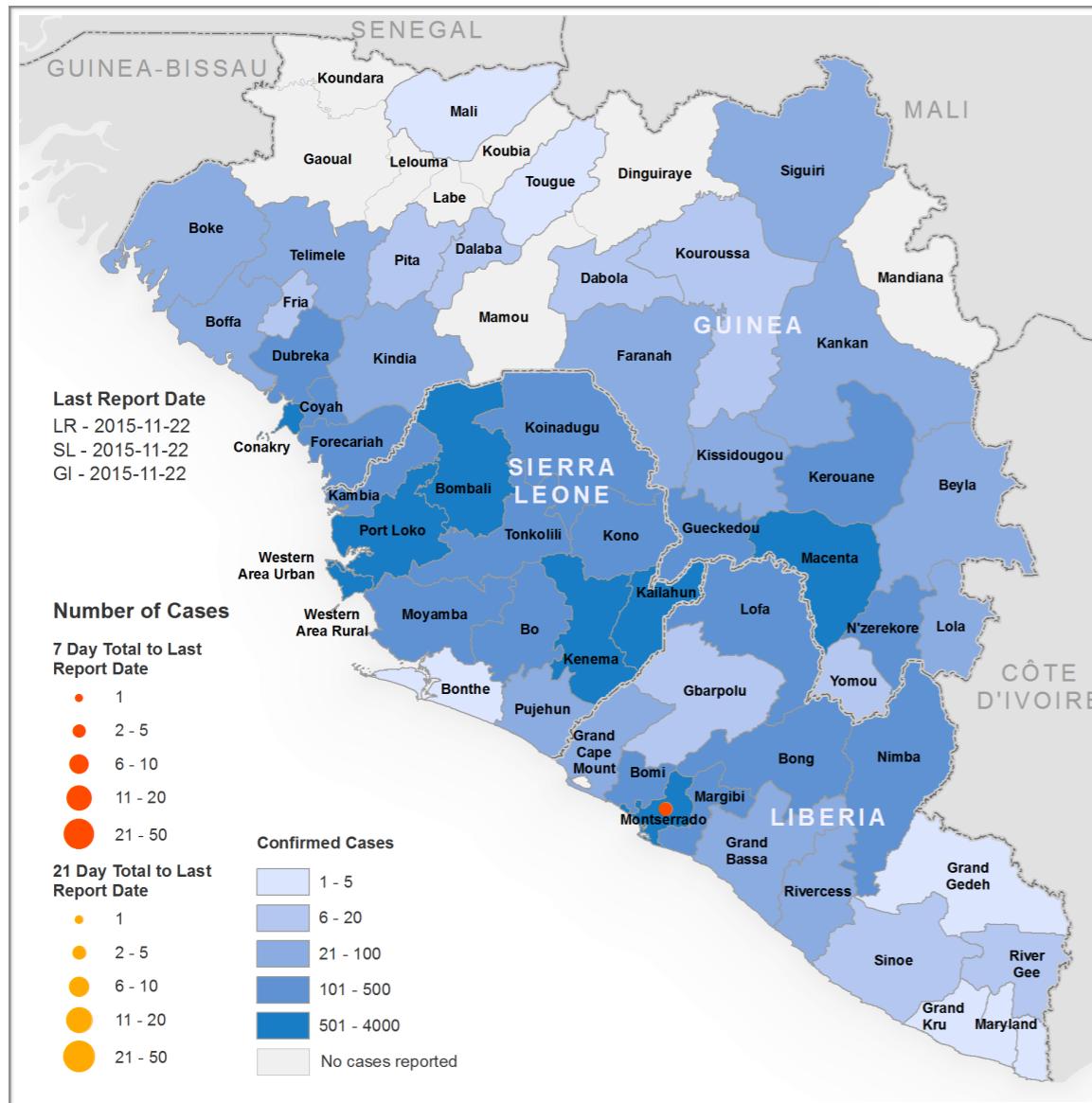
about workshops contact license powered by GitHub and Jekyll 

# PHYLOGENETIC DYNAMICS OF AN EBOLA EPIDEMIC WITH BIRTH- DEATH MODELS



Nick Brandt

# 2014-2015 West African **Ebola** epidemic

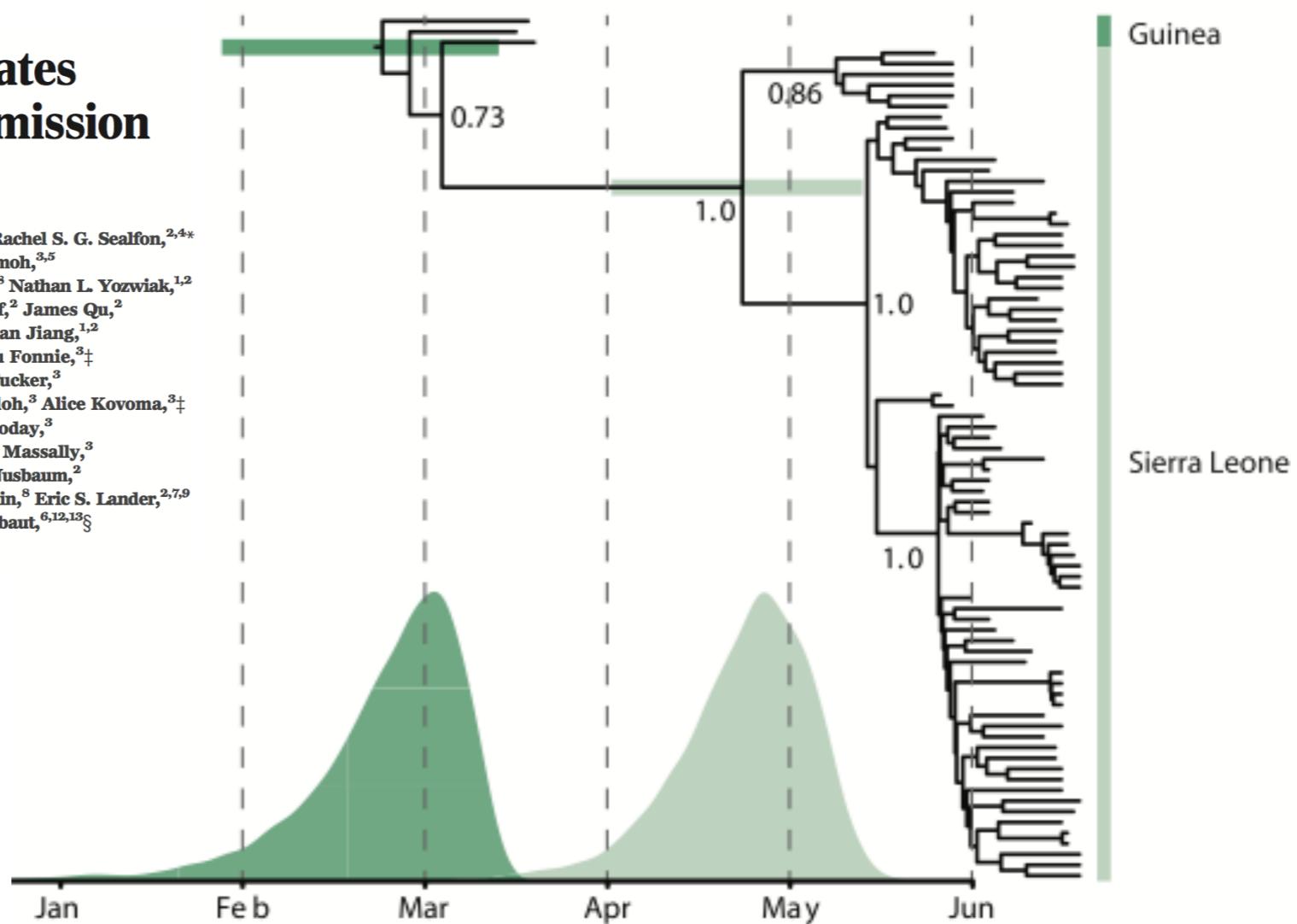


# 2014-2015 West African **Ebola** epidemic

## VIRAL EVOLUTION

### Genomic surveillance elucidates Ebola virus origin and transmission during the 2014 outbreak

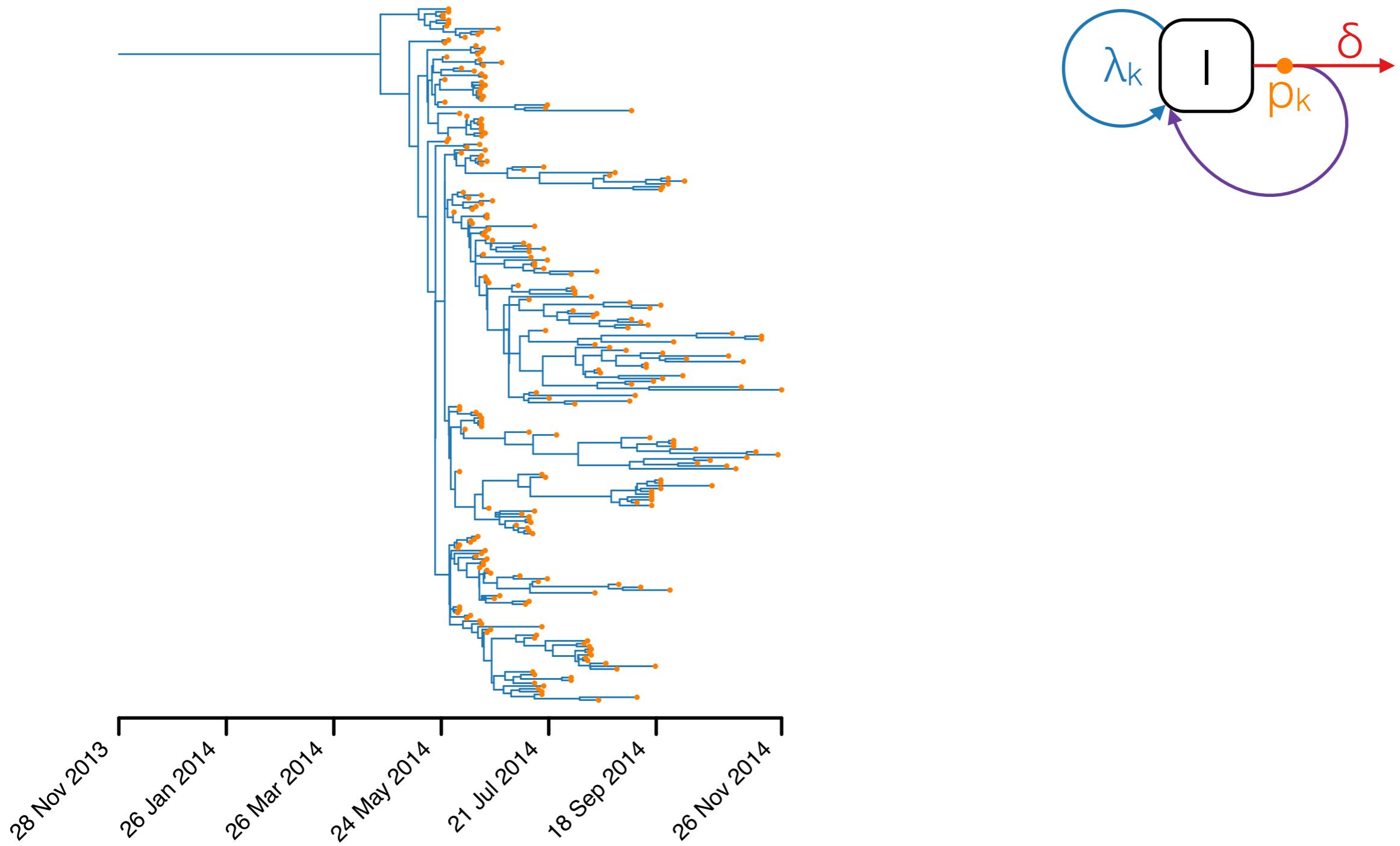
Stephen K. Gire,<sup>1,2\*</sup> Augustine Goba,<sup>3\*†</sup> Kristian G. Andersen,<sup>1,2\*†</sup> Rachel S. G. Sealson,<sup>2,4\*</sup> Daniel J. Park,<sup>2\*</sup> Lansana Kanneh,<sup>3</sup> Simbirie Jalloh,<sup>3</sup> Mambu Momoh,<sup>3,5</sup> Mohamed Fullah,<sup>3,5†</sup> Gytis Dudas,<sup>6</sup> Shirlee Wohl,<sup>1,2,7</sup> Lina M. Moses,<sup>8</sup> Nathan L. Yozwiak,<sup>1,2</sup> Sarah Winnicki,<sup>1,2</sup> Christian B. Matranga,<sup>2</sup> Christine M. Malboeuf,<sup>2</sup> James Qu,<sup>2</sup> Adrienne D. Gladden,<sup>2</sup> Stephen F. Schaffner,<sup>1,2</sup> Xiao Yang,<sup>2</sup> Pan-Pan Jiang,<sup>1,2</sup> Mahan Nekoui,<sup>1,2</sup> Andres Colubri,<sup>1</sup> Moinya Ruth Coomber,<sup>3</sup> Mbalu Fonnie,<sup>3†</sup> Alex Moigboi,<sup>3†</sup> Michael Gbakie,<sup>3</sup> Fatima K. Kamara,<sup>3</sup> Veronica Tucker,<sup>3</sup> Edwin Konuwa,<sup>3</sup> Sidiki Saffa,<sup>3†</sup> Josephine Sellu,<sup>3</sup> Abdul Azziz Jalloh,<sup>3</sup> Alice Kovoma,<sup>3†</sup> James Koninga,<sup>3</sup> Ibrahim Mustapha,<sup>3</sup> Kandeh Kargbo,<sup>3</sup> Momoh Foday,<sup>3</sup> Mohamed Yillah,<sup>3</sup> Franklyn Kanneh,<sup>3</sup> Willie Robert,<sup>3</sup> James L. B. Massally,<sup>3</sup> Sinéad B. Chapman,<sup>2</sup> James Bochicchio,<sup>2</sup> Cheryl Murphy,<sup>2</sup> Chad Nusbaum,<sup>2</sup> Sarah Young,<sup>2</sup> Bruce W. Birren,<sup>2</sup> Donald S. Grant,<sup>3</sup> John S. Scheiffelin,<sup>8</sup> Eric S. Lander,<sup>2,7,9</sup> Christian Happi,<sup>10</sup> Sahr M. Gevao,<sup>11</sup> Andreas Gmirke,<sup>2§</sup> Andrew Rambaut,<sup>6,12,13§</sup> Robert F. Garry,<sup>8§</sup> S. Humarr Khan,<sup>3†§</sup> Pardis C. Sabeti<sup>1,2†§</sup>



Gire et al. Science 2014

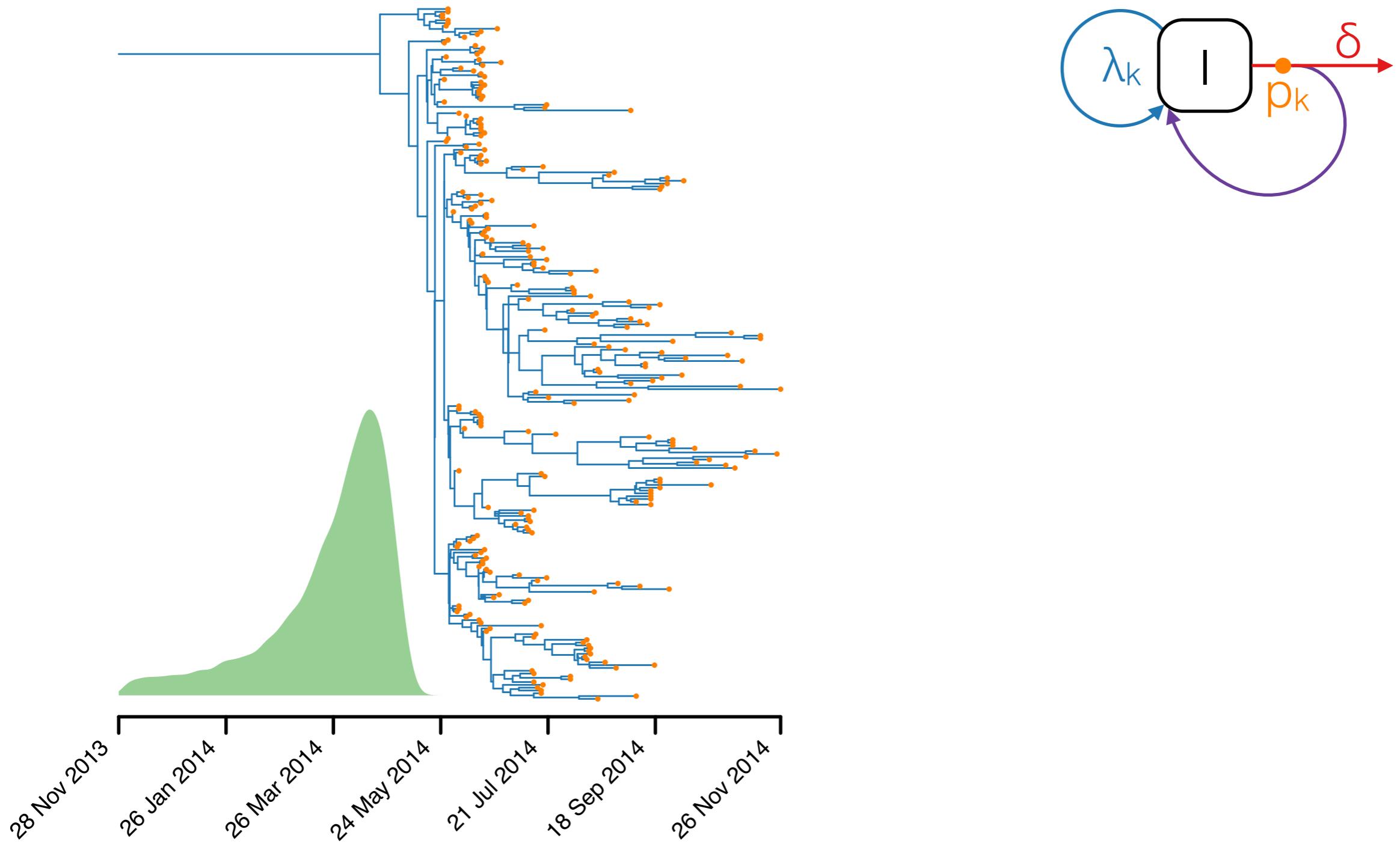
# Results in **Southeast Sierra Leone**

(Based on **246** genomes sampled between **25 May** and **26 Nov 2014**)



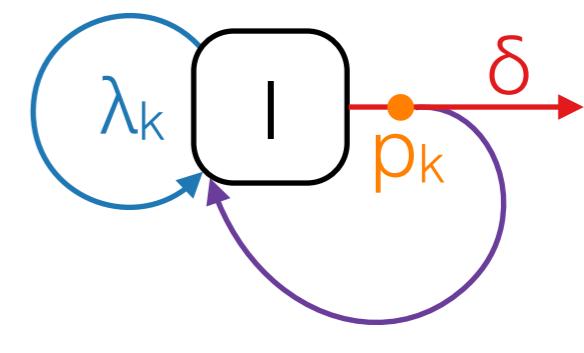
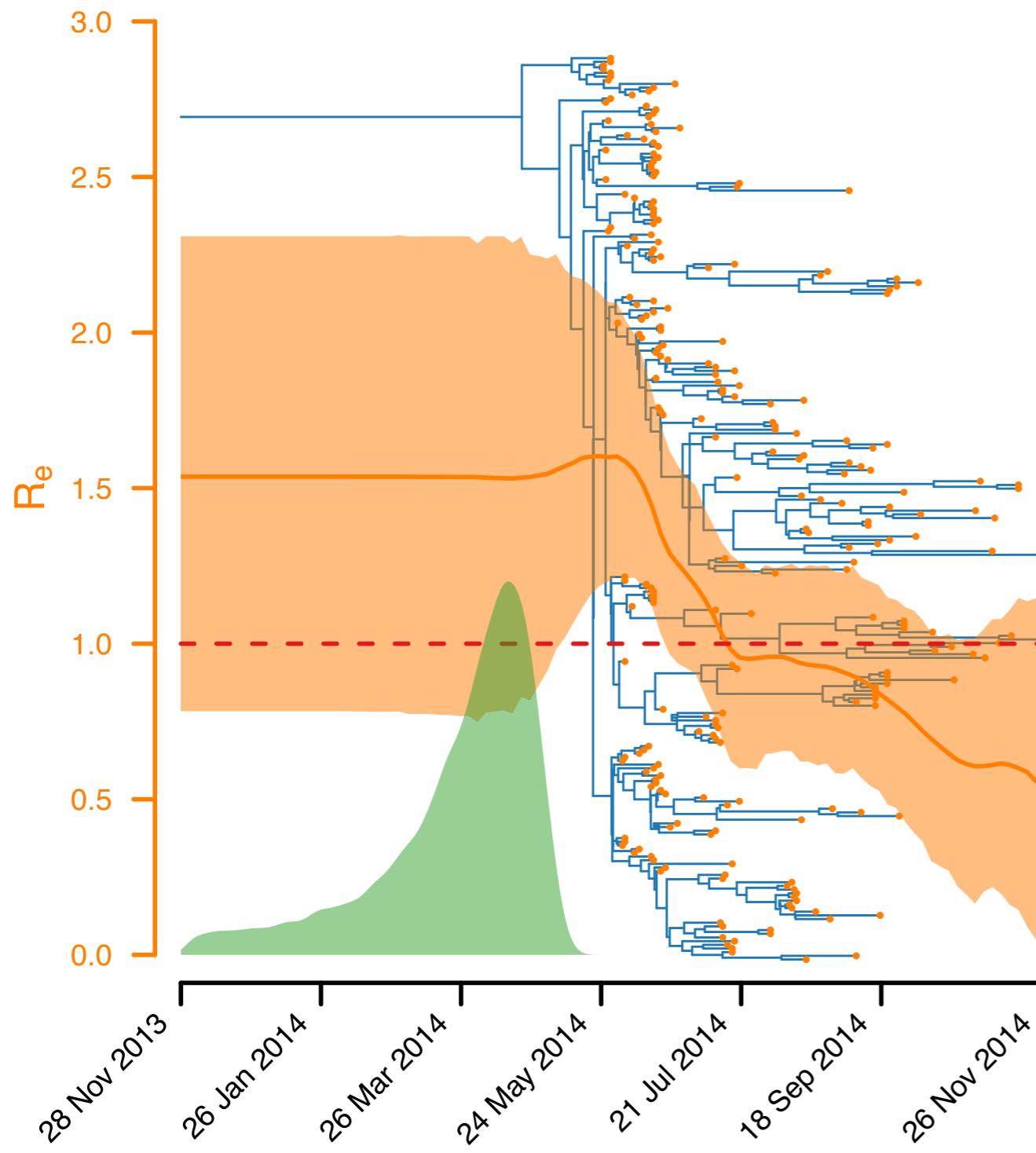
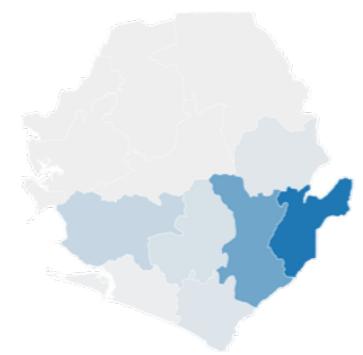
# Results in **Southeast Sierra Leone**

(Based on **246** genomes sampled between **25 May** and **26 Nov 2014**)



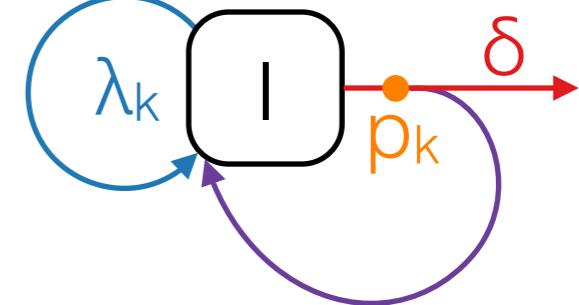
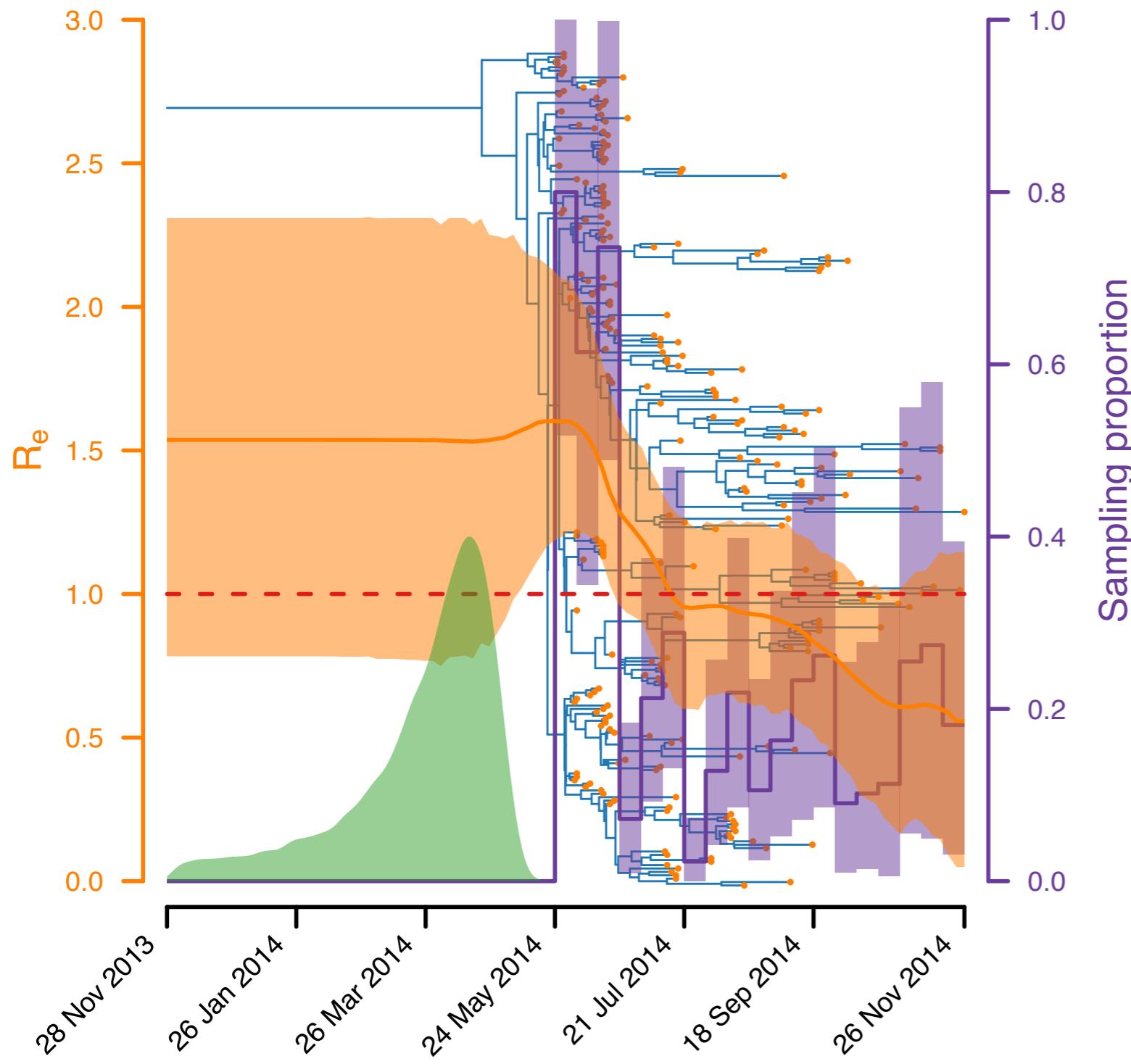
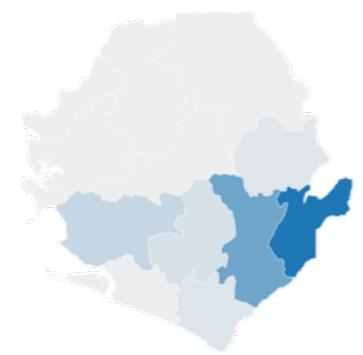
# Results in **Southeast Sierra Leone**

(Based on **246** genomes sampled between **25 May** and **26 Nov 2014**)



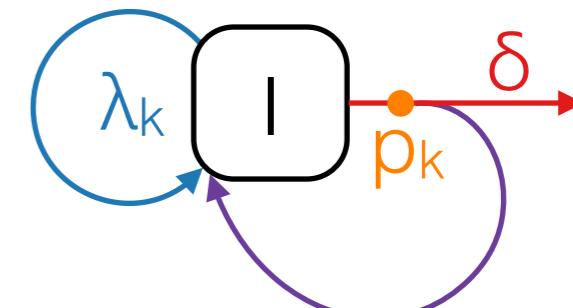
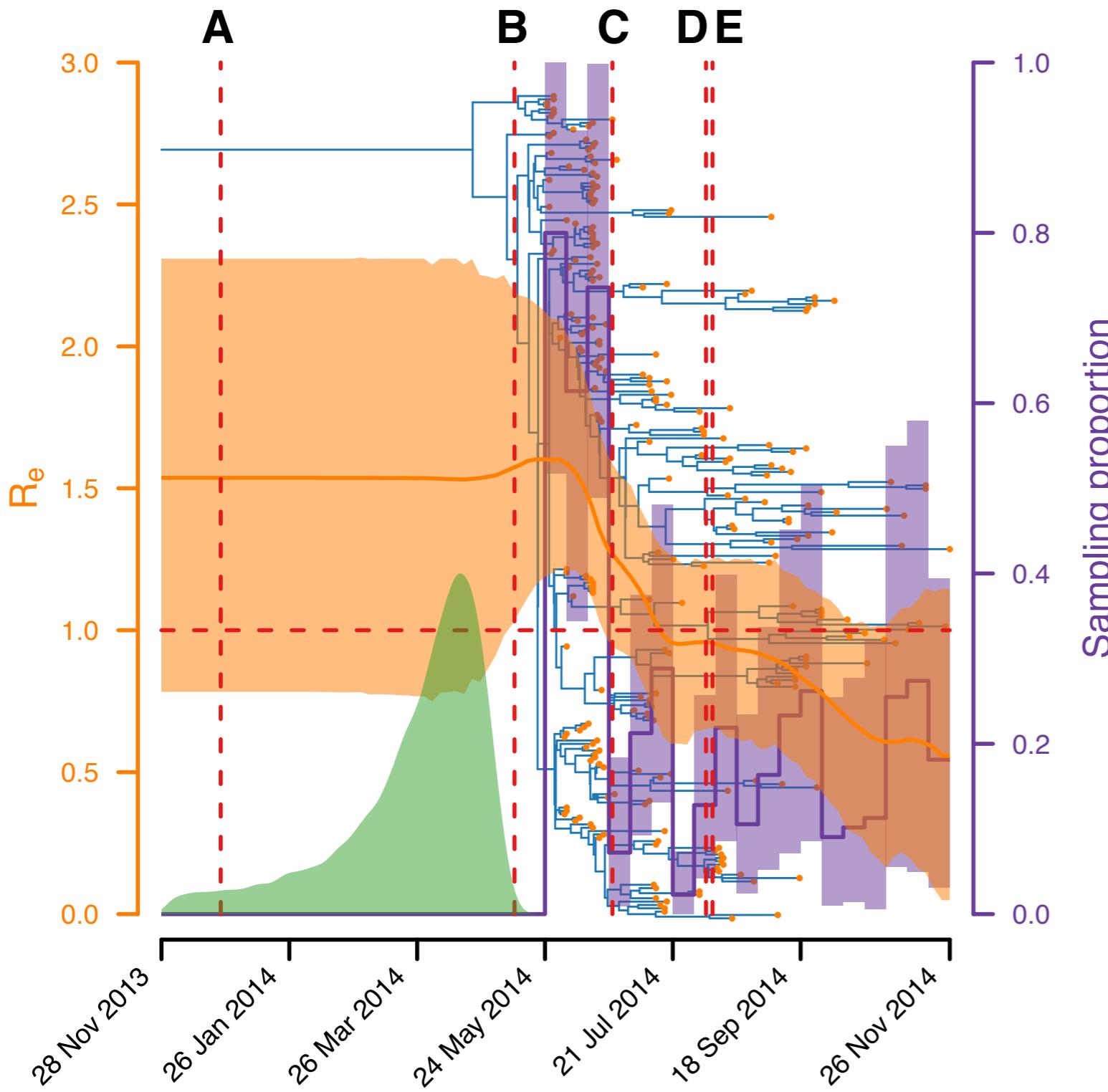
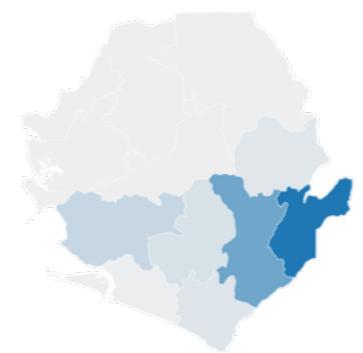
# Results in **Southeast Sierra Leone**

(Based on **246** genomes sampled between **25 May** and **26 Nov 2014**)



# Results in **Southeast Sierra Leone**

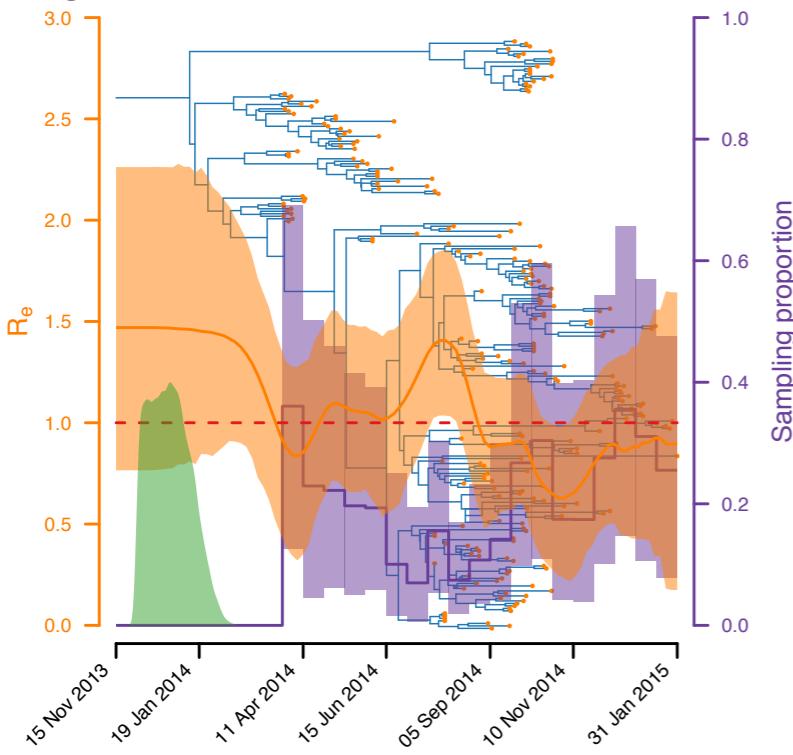
(Based on **246** genomes sampled between **25 May** and **26 Nov 2014**)



- A. Suspected index case
- B. Funeral of traditional healer
- C. MSF opens a treatment centre in Kailahun
- D. National state of emergency
- E. Global public health emergency declared

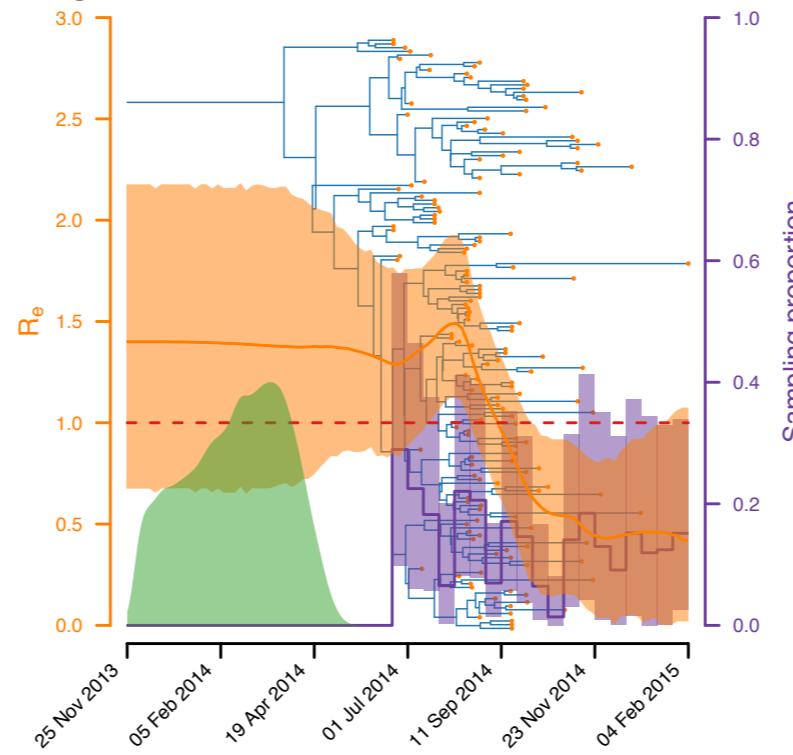
# Guinea

(236 genomes)



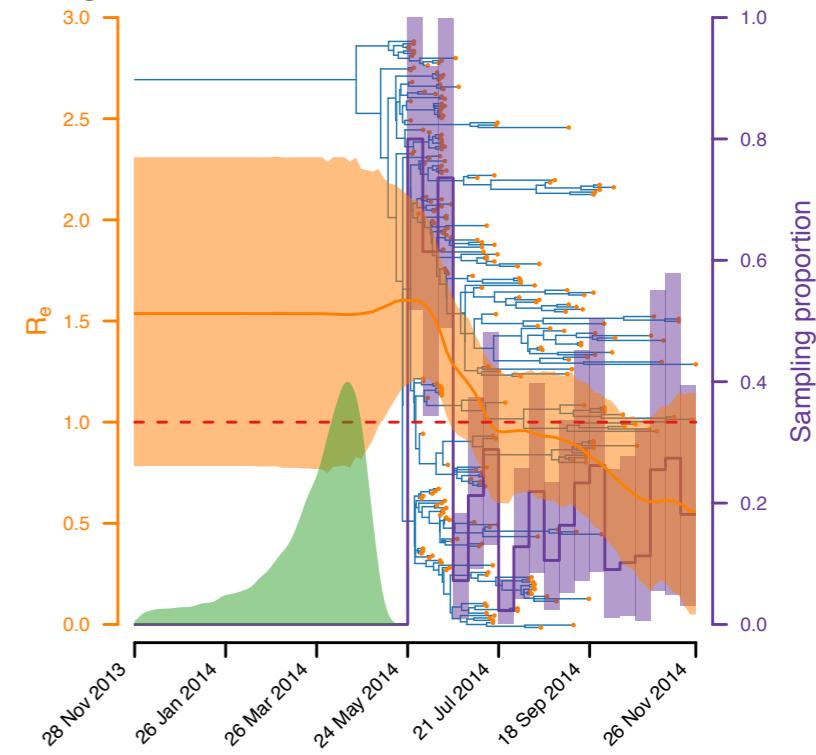
# Liberia

(159 genomes)

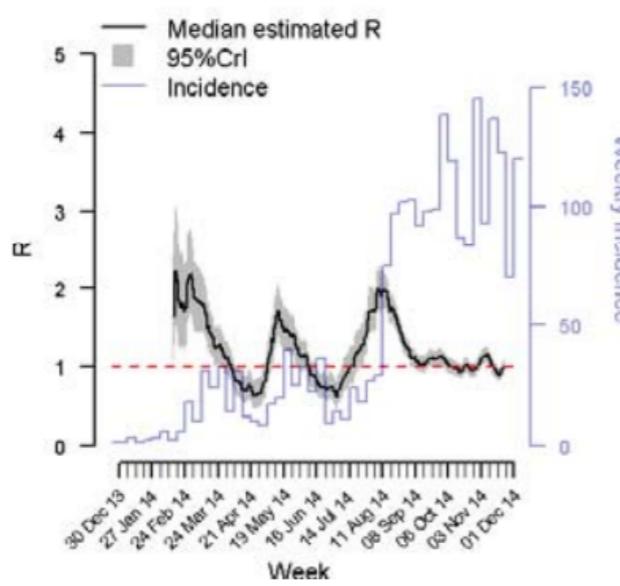


# Southeast Sierra Leone

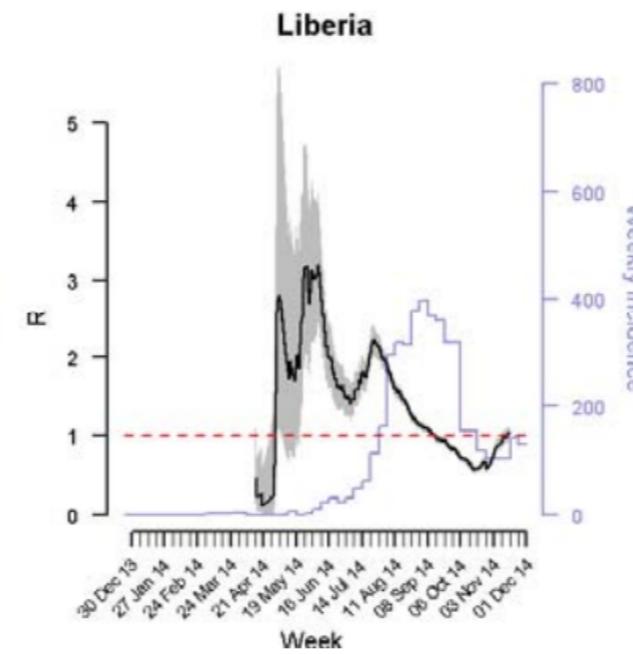
(246 genomes)



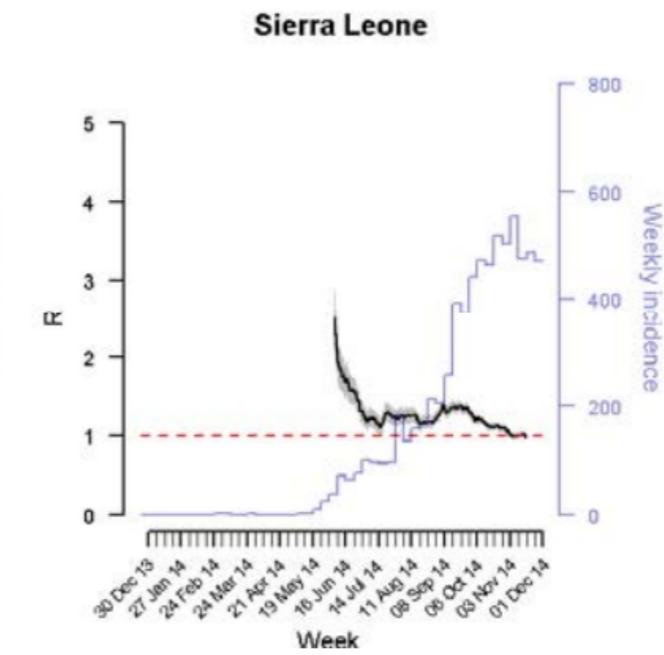
## Guinea



## Liberia



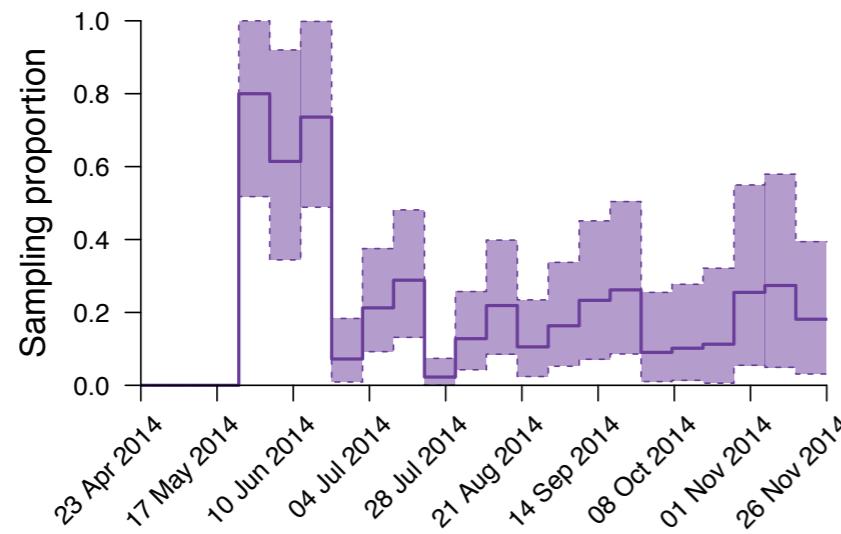
## Sierra Leone



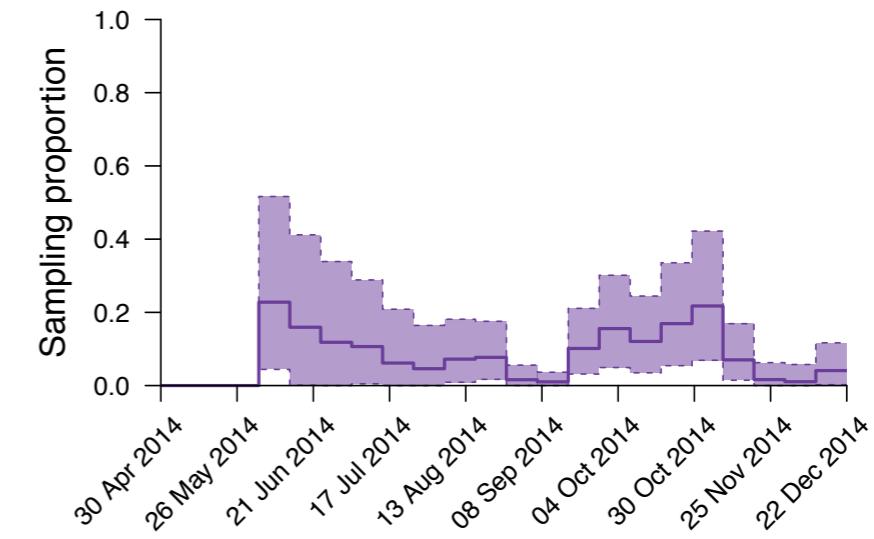
# Changes in the **sampling proportion**

---

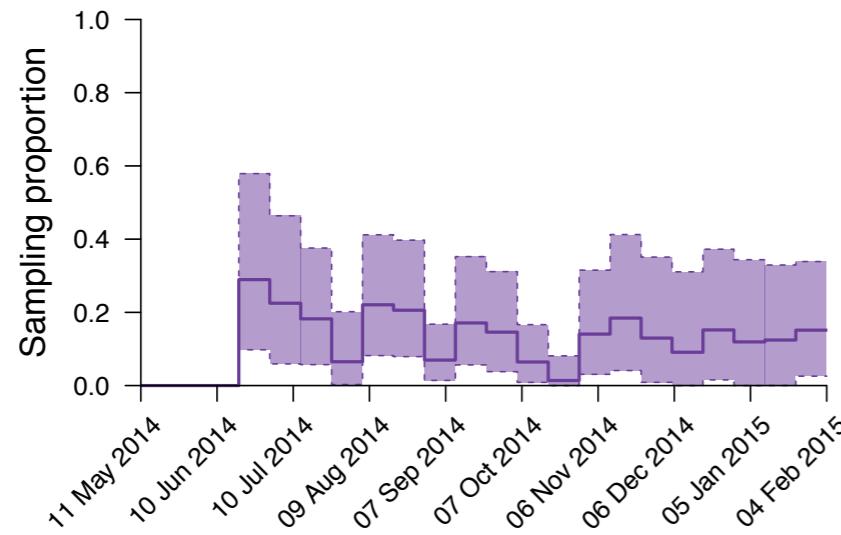
**Southeast Sierra Leone**



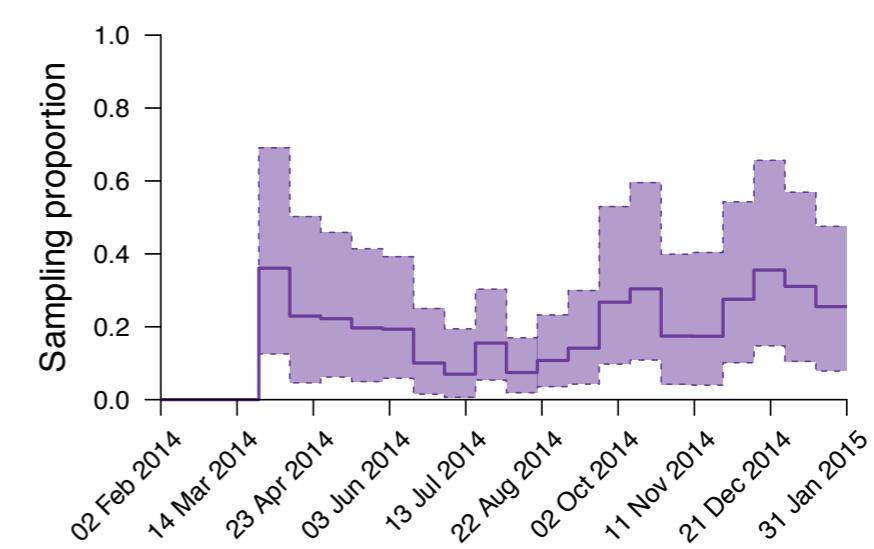
**Northwest Sierra Leone**



**Liberia**



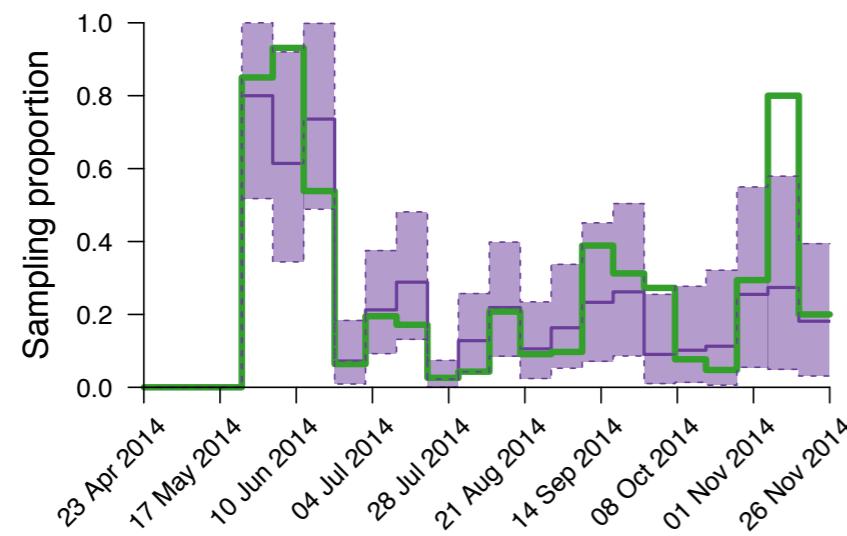
**Guinea**



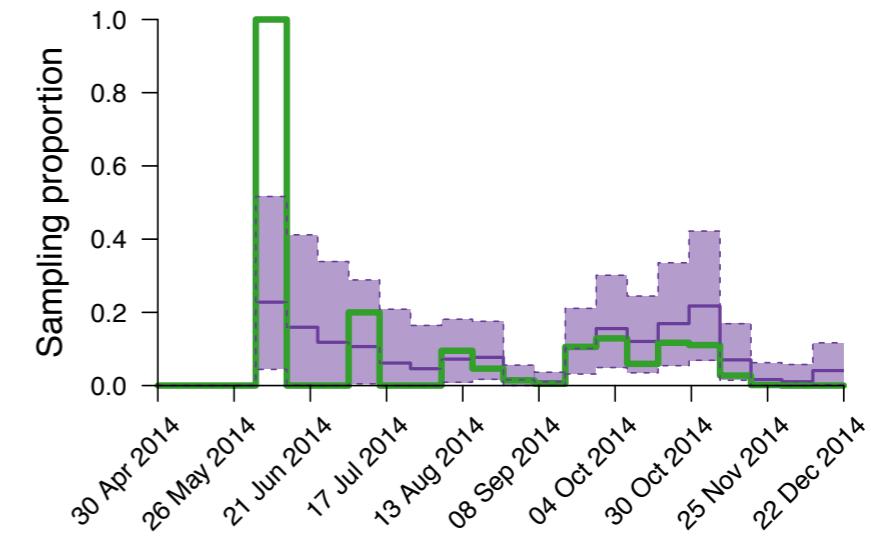
# Changes in the **sampling proportion**

(Compared to the number of **known cases in highly sampled areas**)

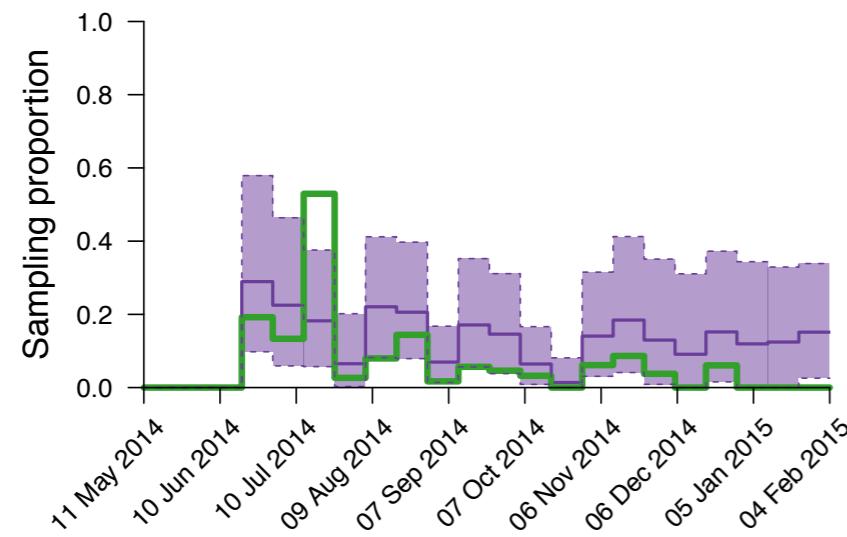
**Southeast Sierra Leone**



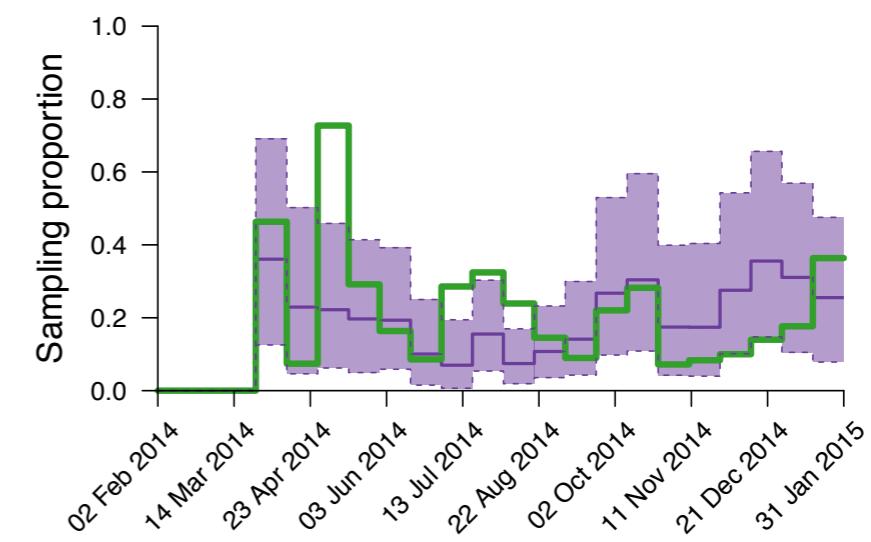
**Northwest Sierra Leone**



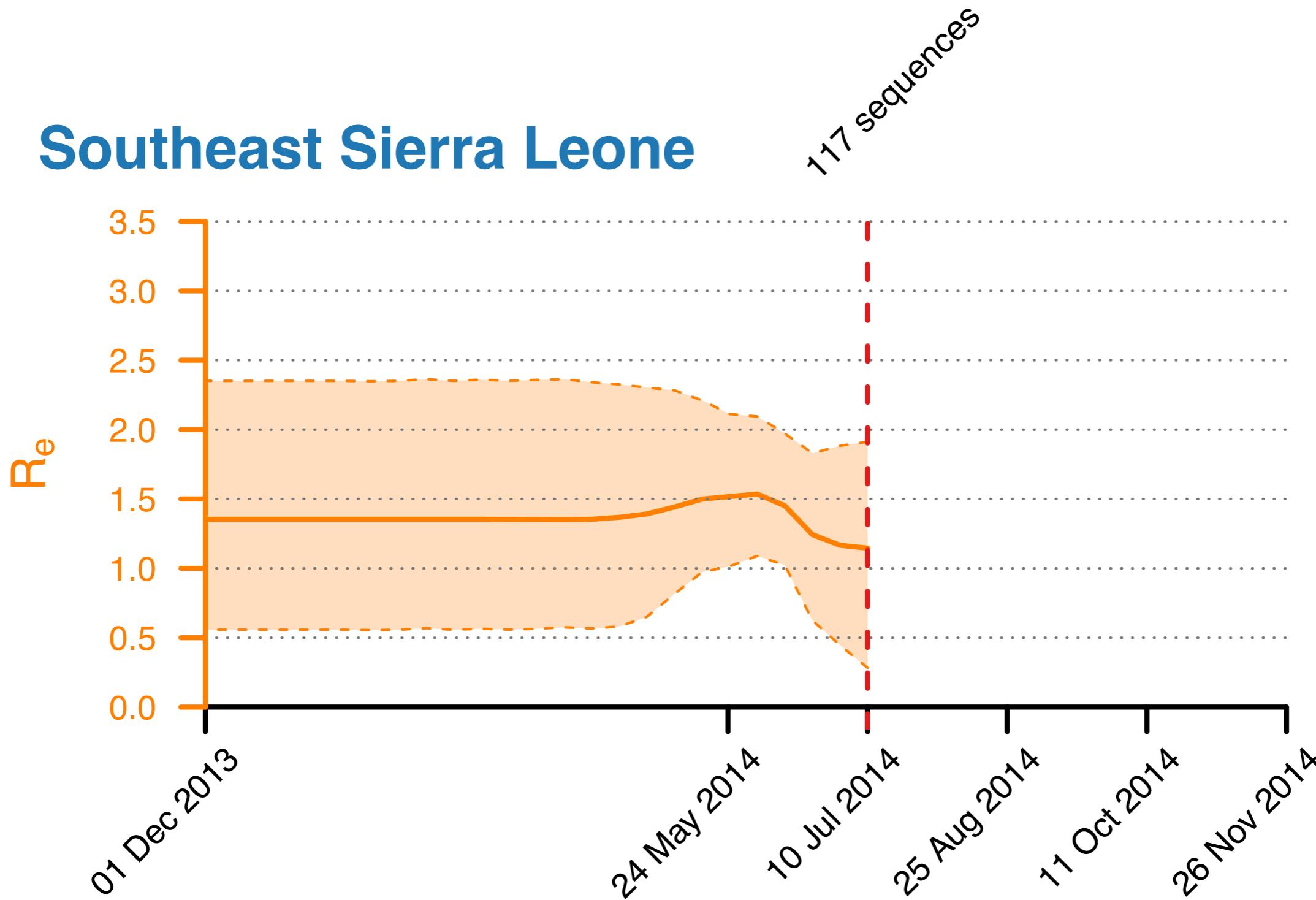
**Liberia**



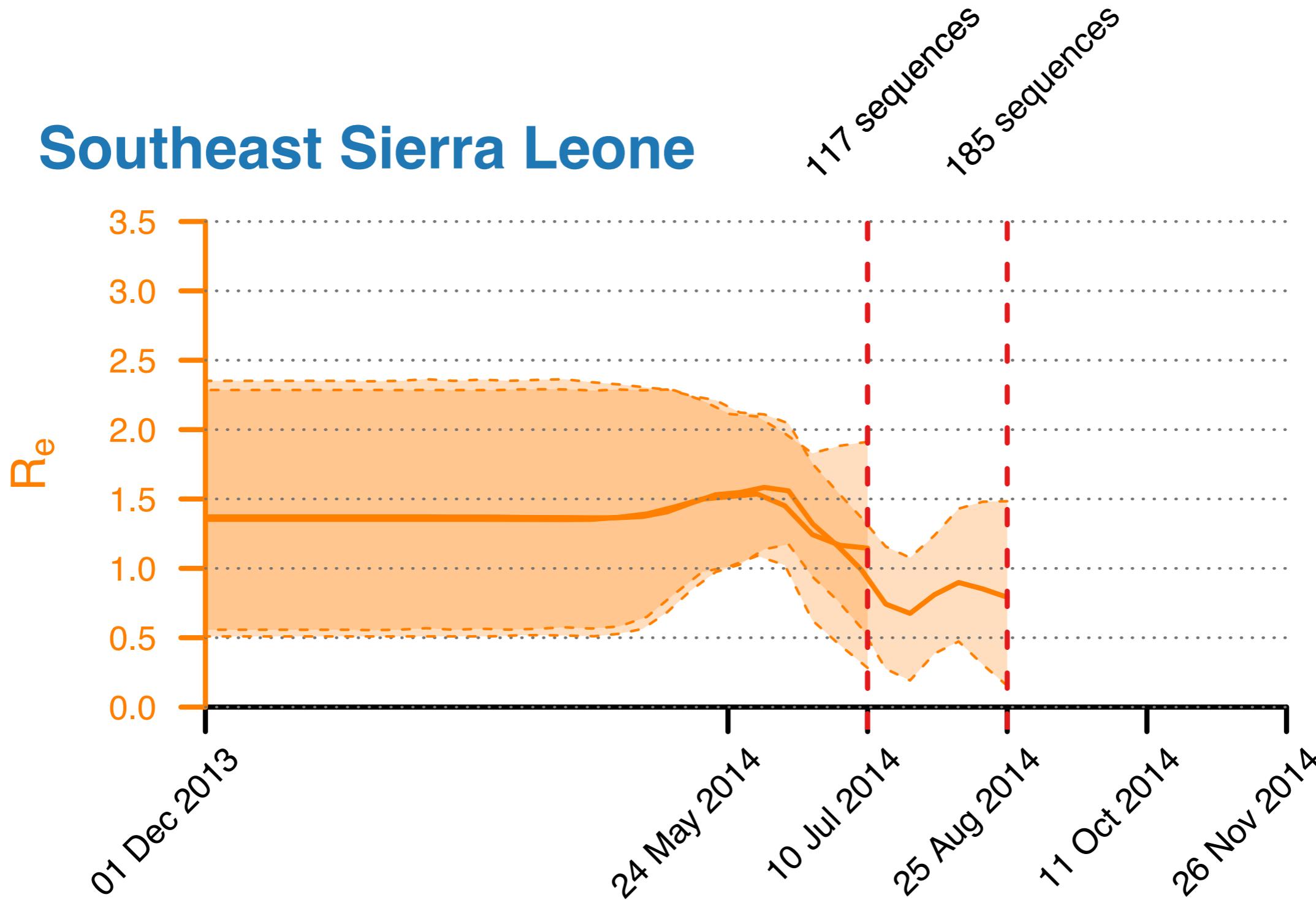
**Guinea**



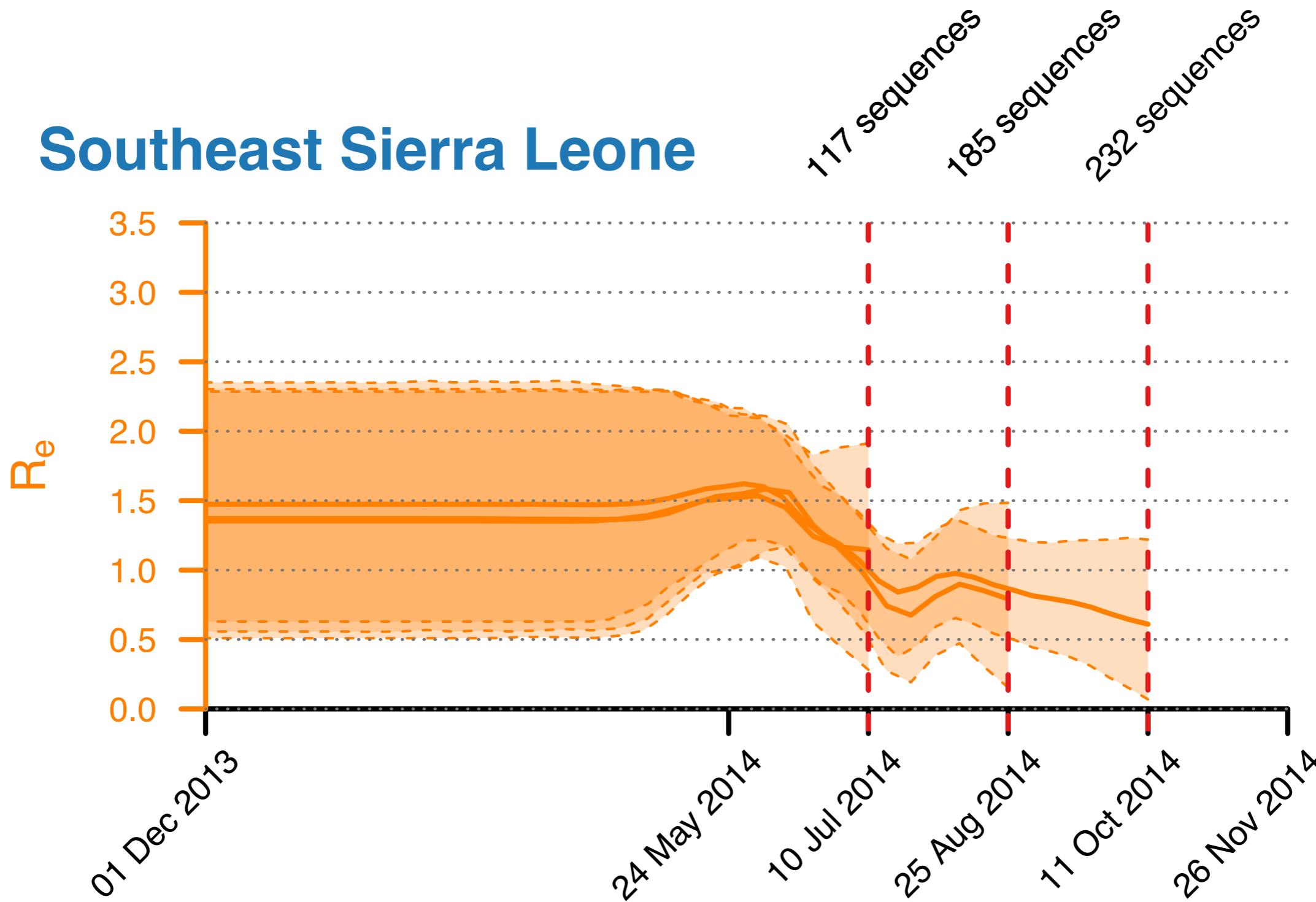
# Time dependence of $R_e$ estimates



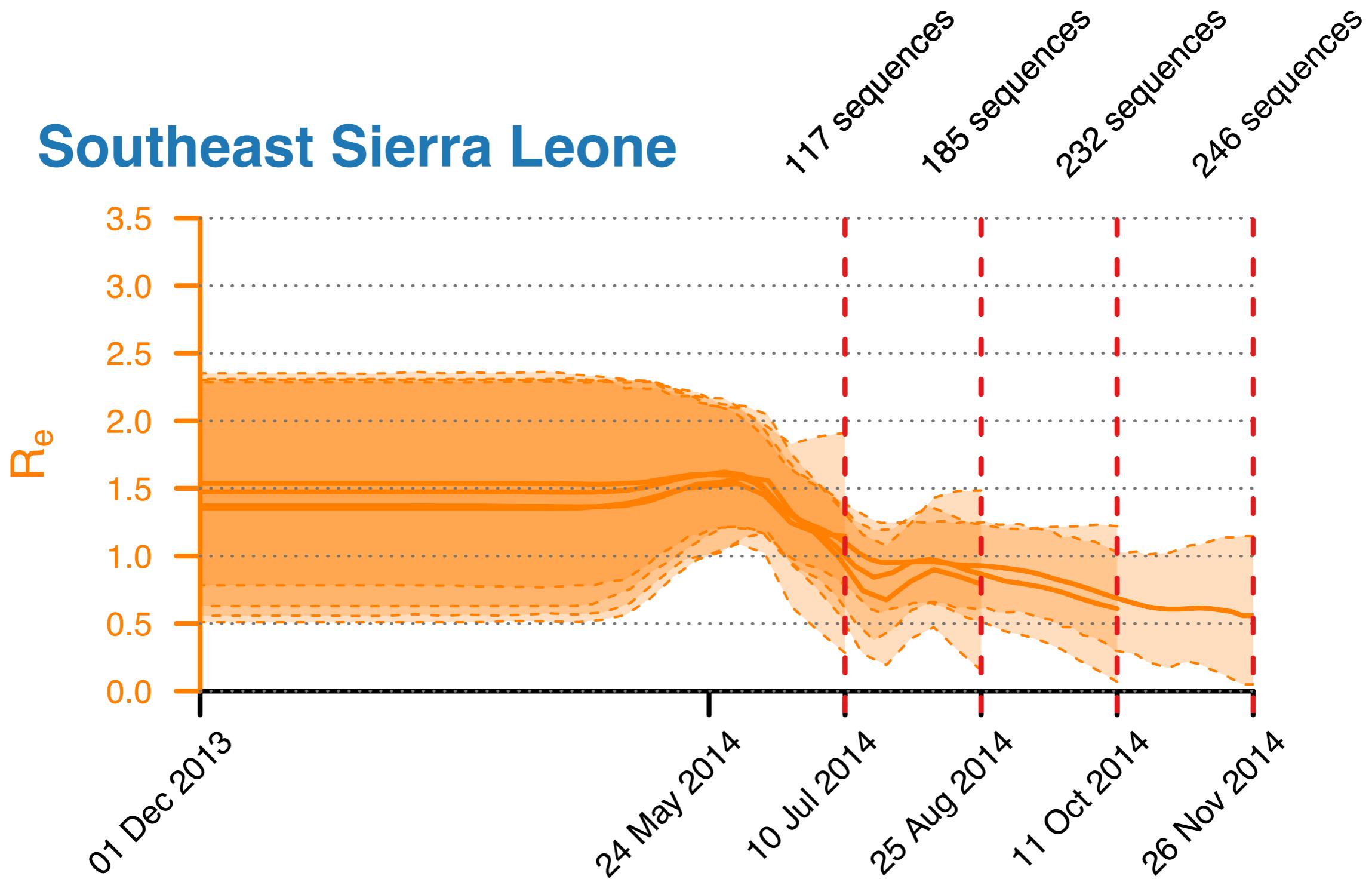
# Time dependence of $R_e$ estimates



# Time dependence of $R_e$ estimates

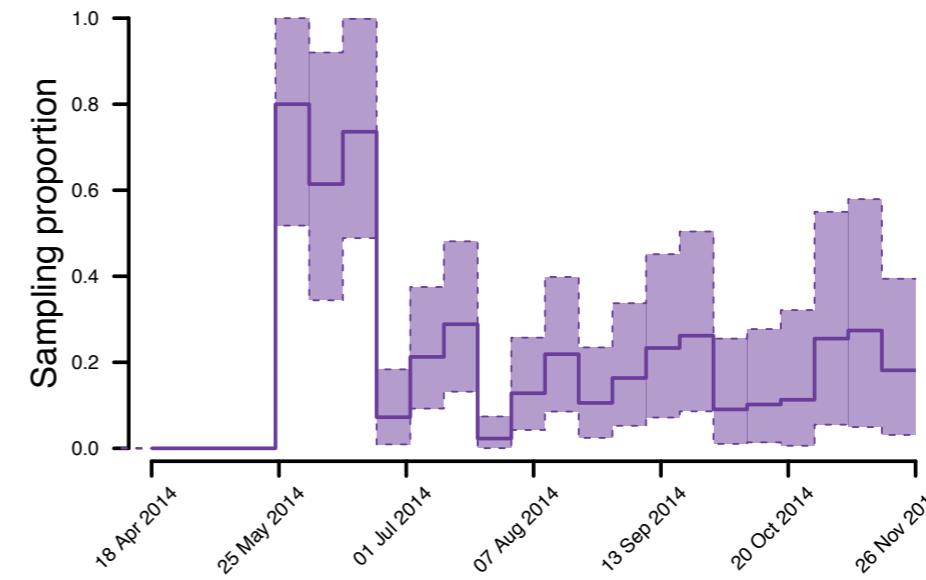
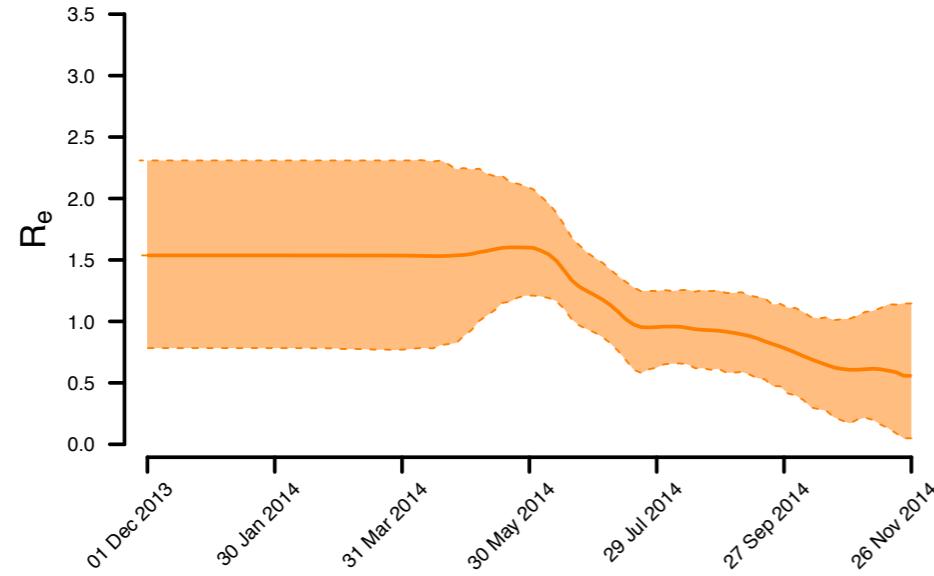


# Time dependence of $R_e$ estimates

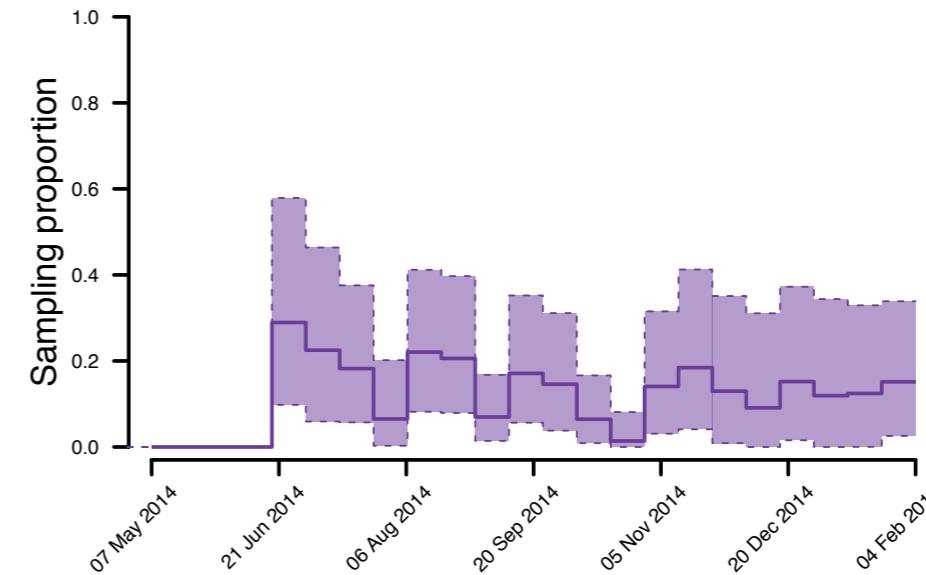
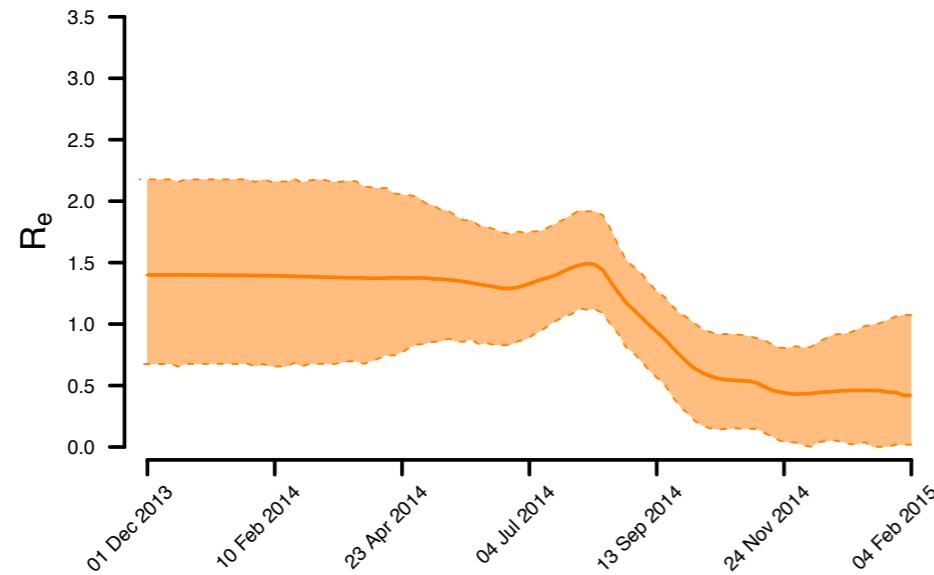


# Sensitivity of skylines

## Southeast Sierra Leone



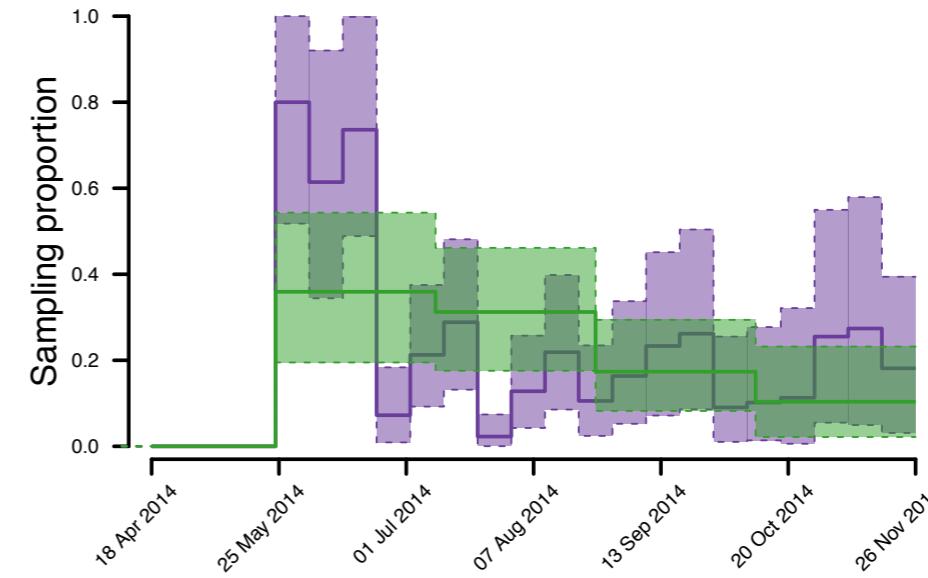
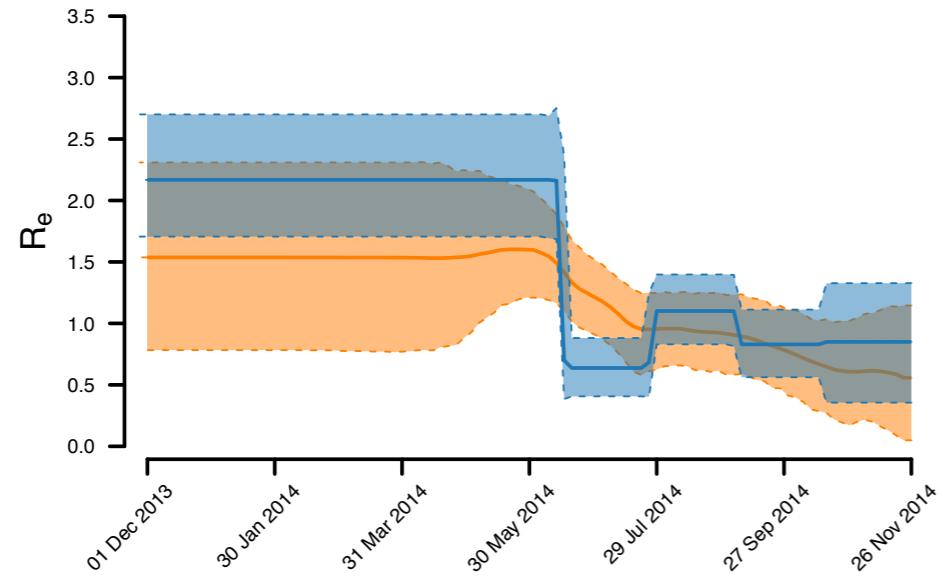
## Liberia



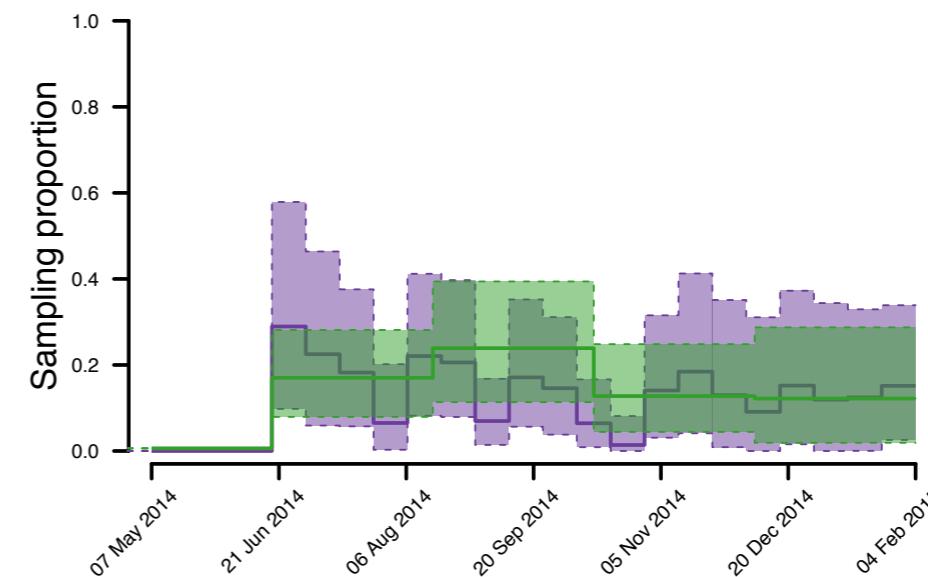
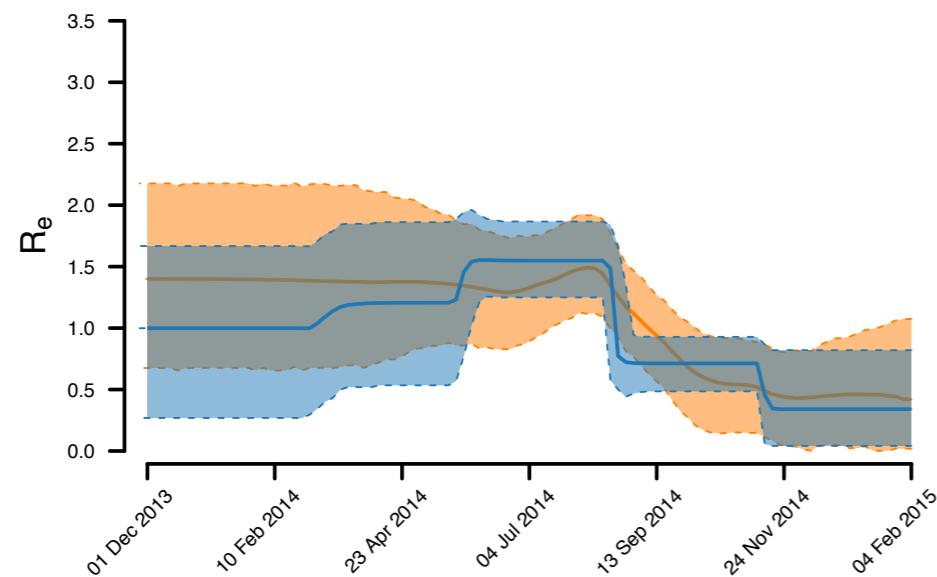
# Sensitivity of skylines

(40 intervals + smoothing prior vs. 5 intervals)

## Southeast Sierra Leone



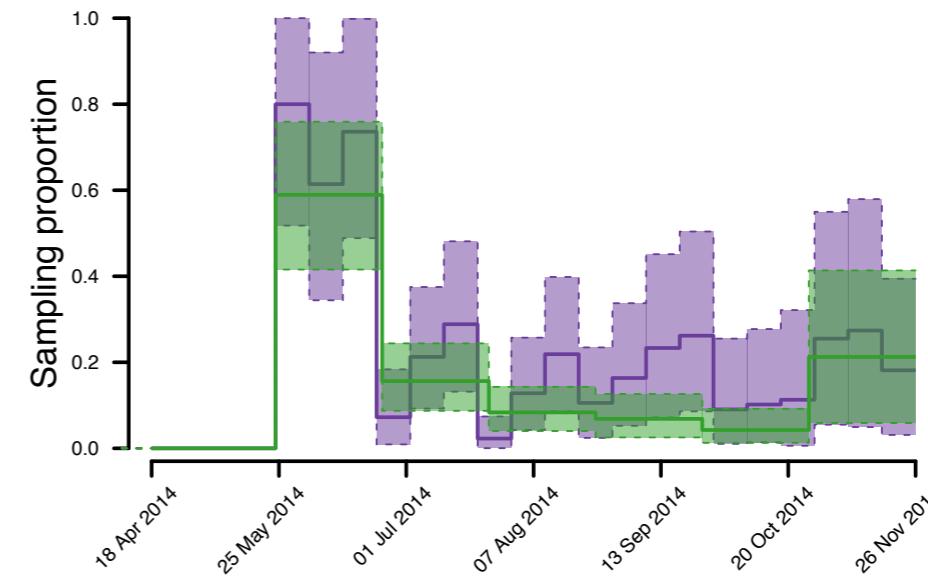
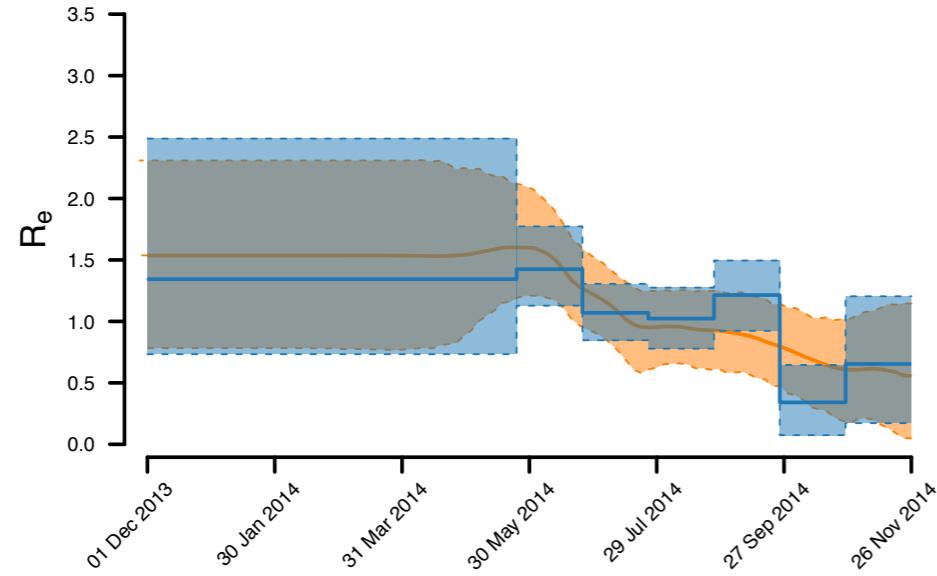
## Liberia



# Sensitivity of skylines

(40 intervals + smoothing prior vs. 7 fixed intervals)

## Southeast Sierra Leone



## Liberia

