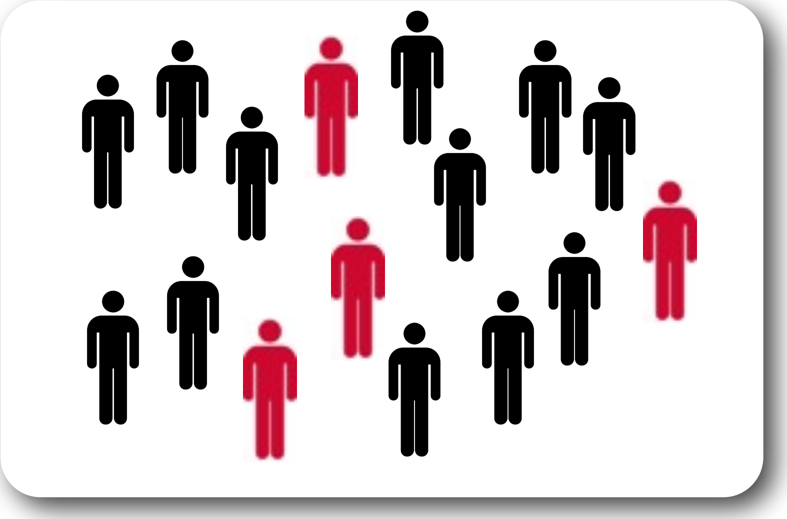
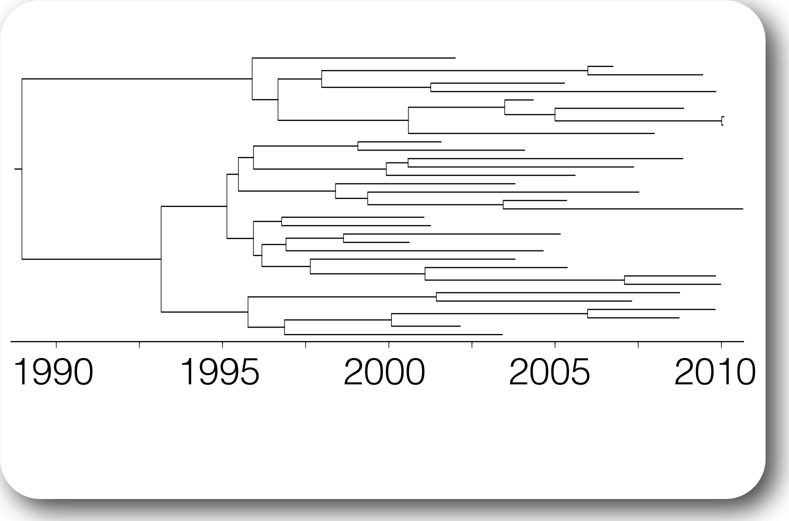
Epidemiology meets phylogenetics

Inferring epidemiological dynamics based on genetic sequence data

Gabriel Leventhal, Sebastian Bonhoeffer

Denise Kühnert, Alexei Drummond





Tanja Stadler
Theoretical Biology group — Department of Environmental Systems Science
tanja.stadler@env.ethz.ch

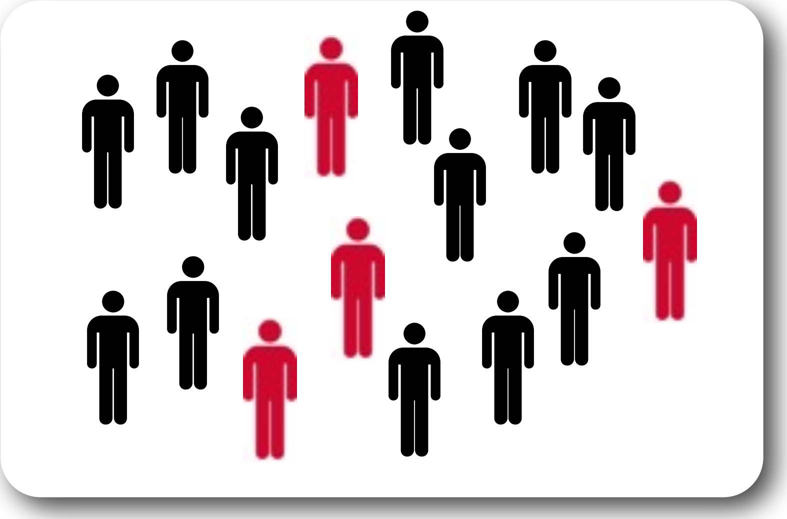
ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

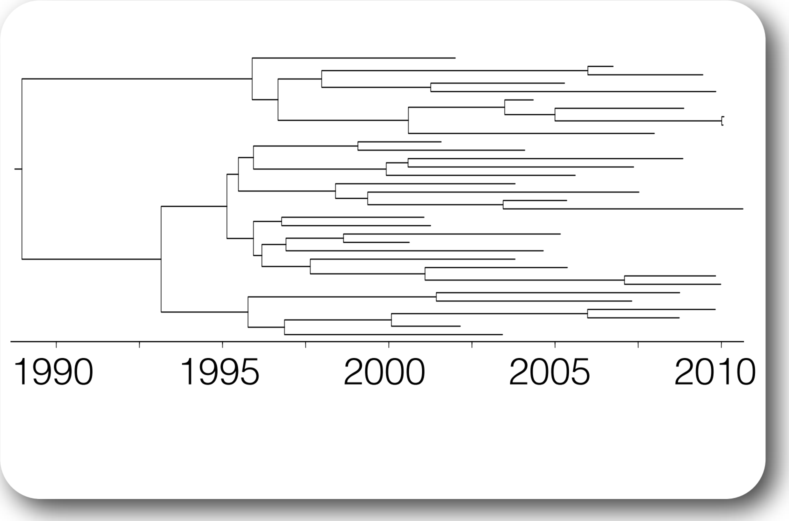
Epidemiology meets phylogenetics

Inferring epidemiological dynamics based on genetic sequence data

Gabriel Leventhal, Sebastian Bonhoeffer

Denise Kühnert, Alexei Drummond

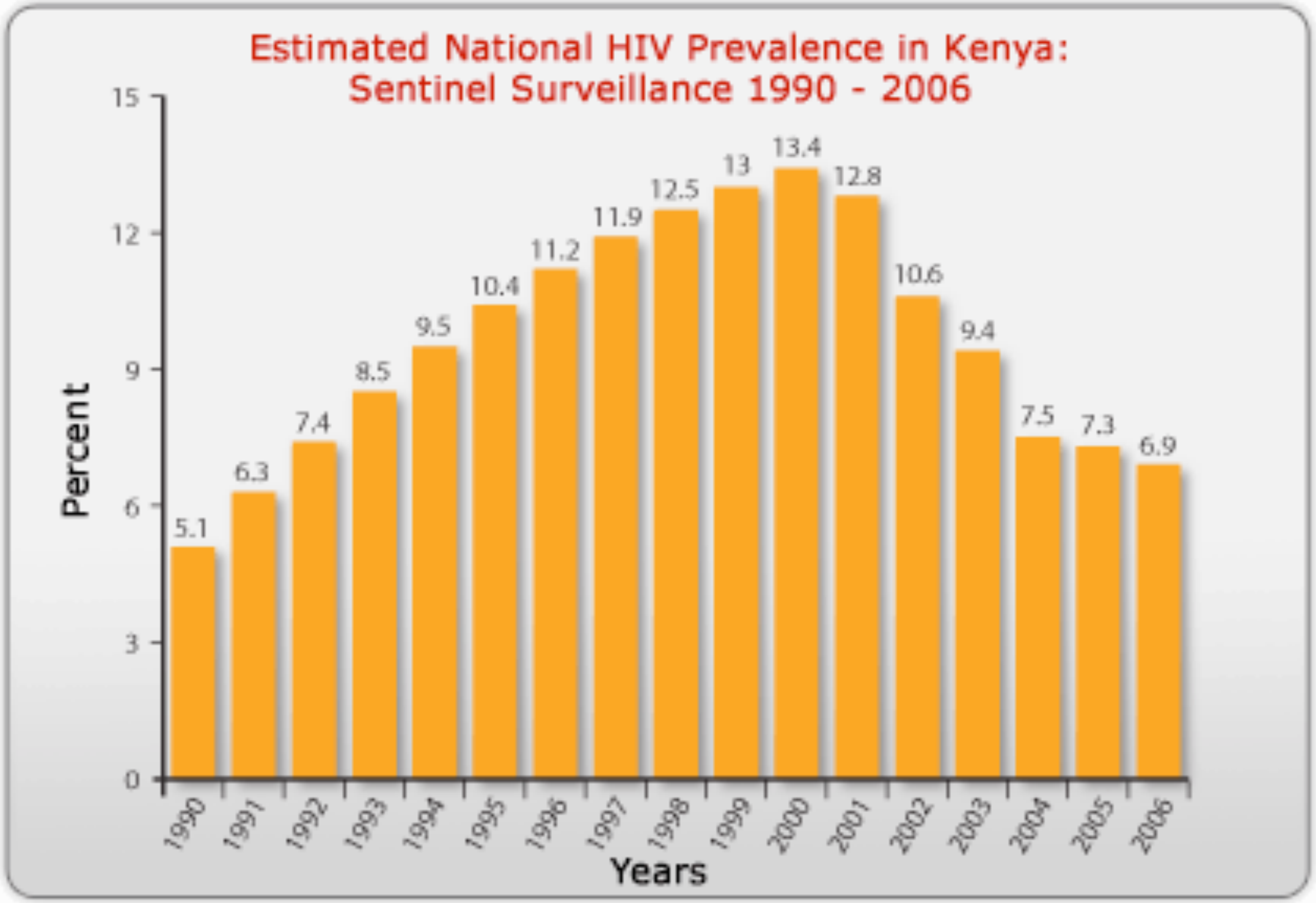




Tanja Stadler
Theoretical Biology group — Department of Environmental Systems Science
tanja.stadler@env.ethz.ch

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

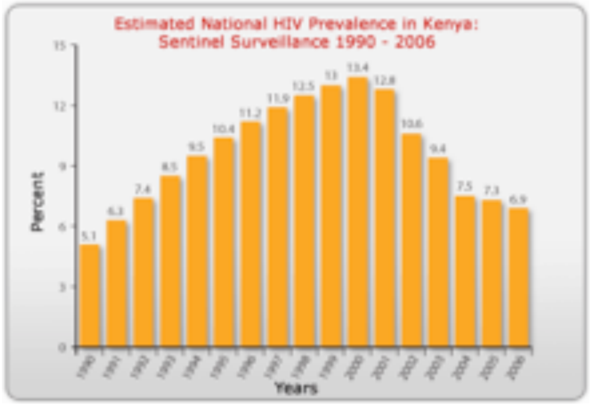
Epidemiology



Epidemiology

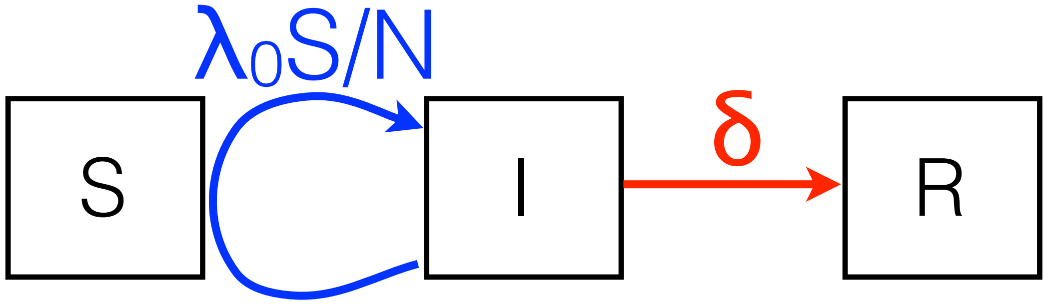
Data

# of infected individuals through time



Models

Population dynamics described by SIR models:

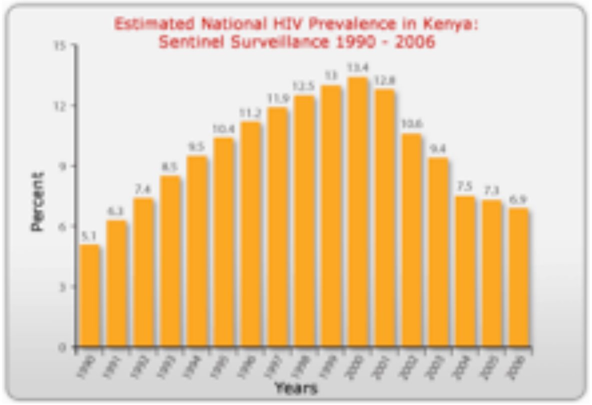


$$
\begin{aligned}
\frac{dS}{dt} &= -\lambda_0 IS/N \\
\frac{dI}{dt} &= \lambda_0 IS/N - \delta I \\
\frac{dR}{dt} &= \delta I
\end{aligned}
$$

Epidemiology

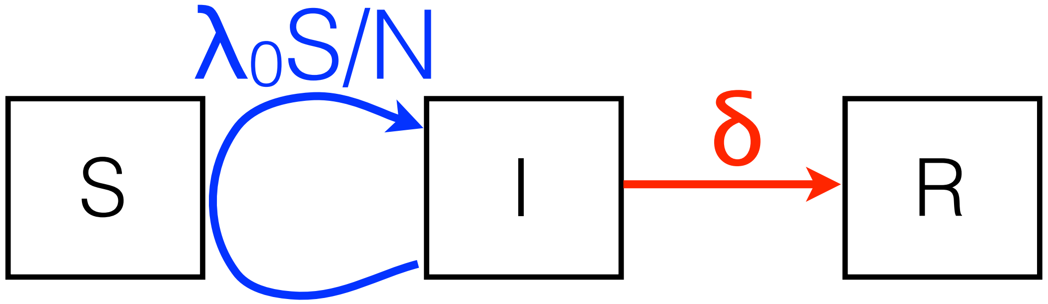
Data

# of infected individuals through time



Models

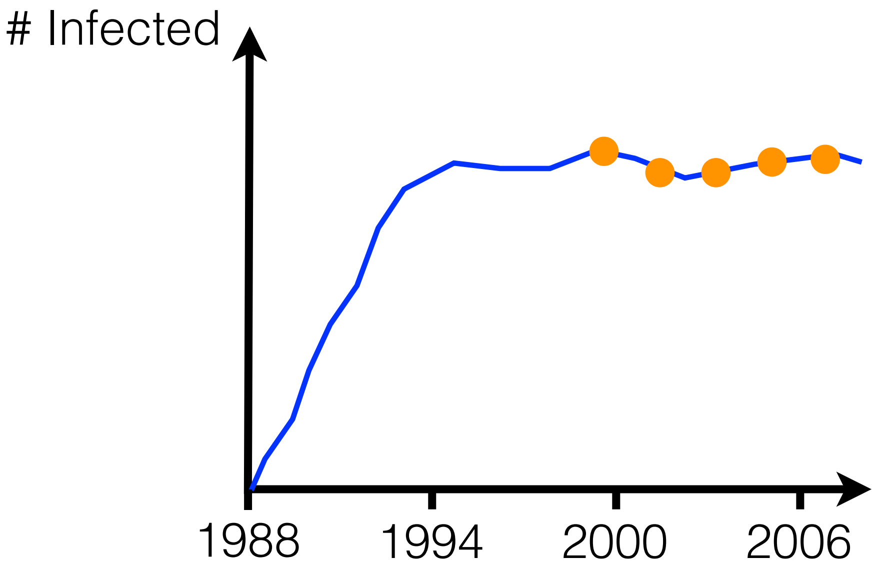
Population dynamics described by SIR models:



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

Limitations

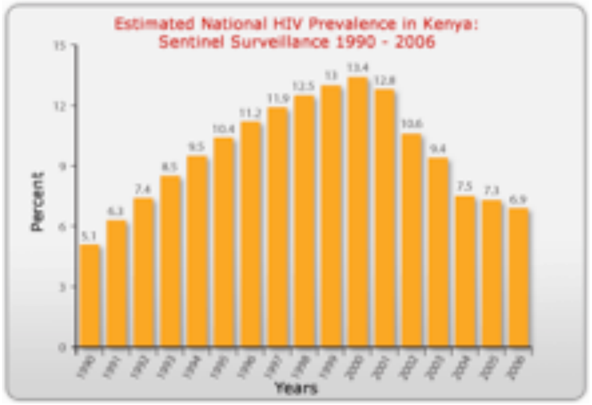
If sampling in early epidemic was missed:



Epidemiology

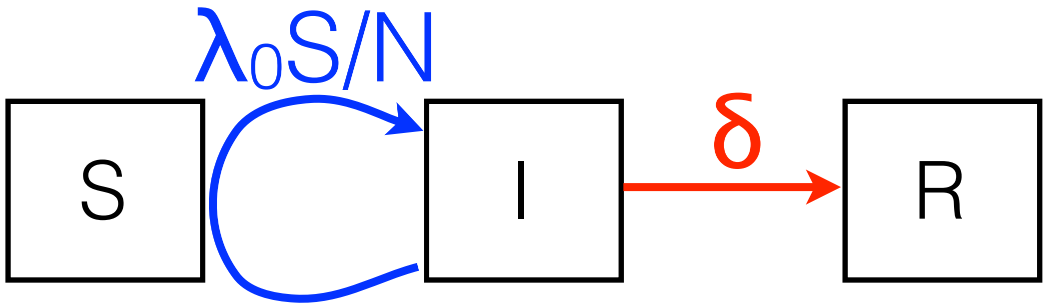
Data

# of infected individuals through time



Models

Population dynamics described by SIR models:

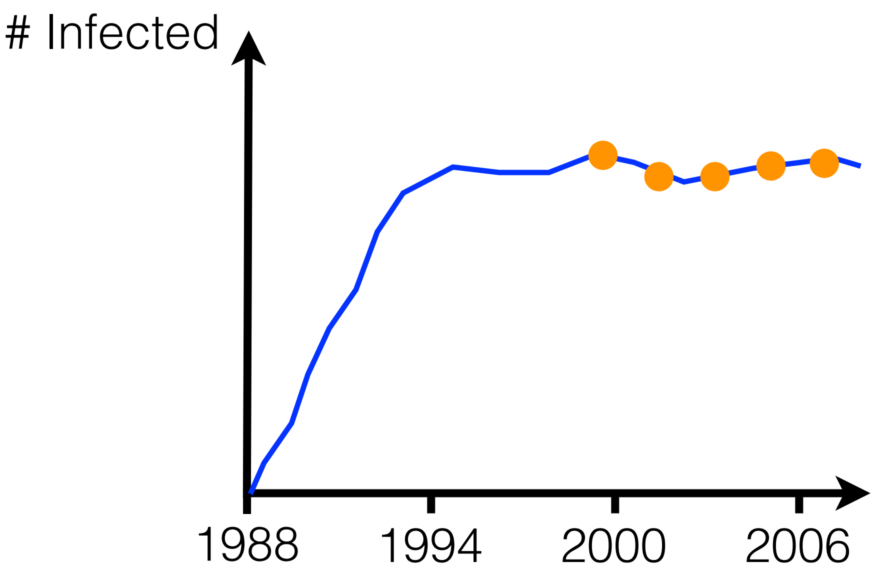


$$
\begin{aligned}
\frac{dS}{dt} &= -\lambda_0 IS/N \\
\frac{dI}{dt} &= \lambda_0 IS/N - \delta I \\
\frac{dR}{dt} &= \delta I
\end{aligned}
$$

Limitations

If sampling in early epidemic was missed:

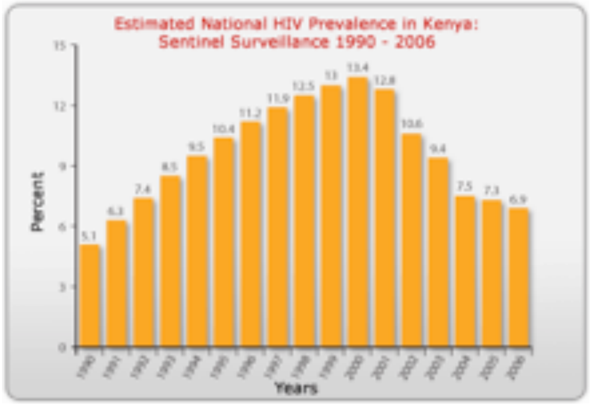
- **Time of epidemic outbreak?**
- **Basic reproductive number R_0?**



Epidemiology

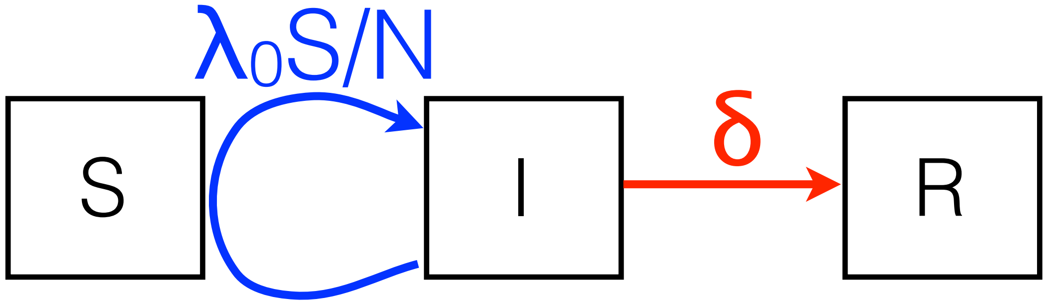
Data

# of infected individuals through time



Models

Population dynamics described by SIR models:



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

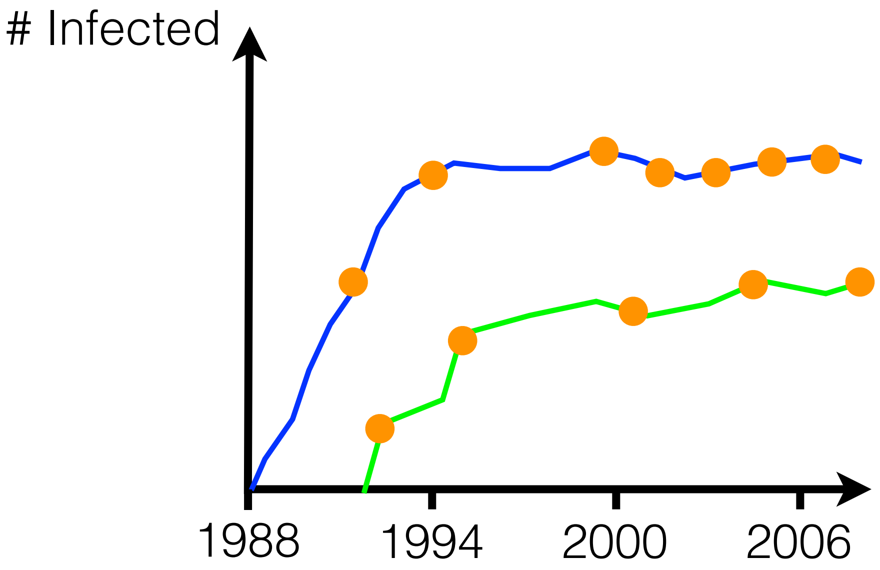
Limitations

If sampling in early epidemic was missed:

- **Time of epidemic outbreak?**
- **Basic reproductive number R_0?**

Data does not tell who infected whom:

- **Population structure?**



Genetic data entering epidemiology

Phylo-
genetics

Inf1	`ACACACCC`
Inf2	`TCACACCT`
Inf3	`AAAGACTT`
Inf4	`ACAGACTT`

Genetic data entering epidemiology

Phylo-genetics

Input:

Pathogen genetic data from different hosts

Inf1	`ACACACCC`
Inf2	`TCACACCT`
Inf3	`AAAGACTT`
Inf4	`ACAGACTT`

Genetic data entering epidemiology

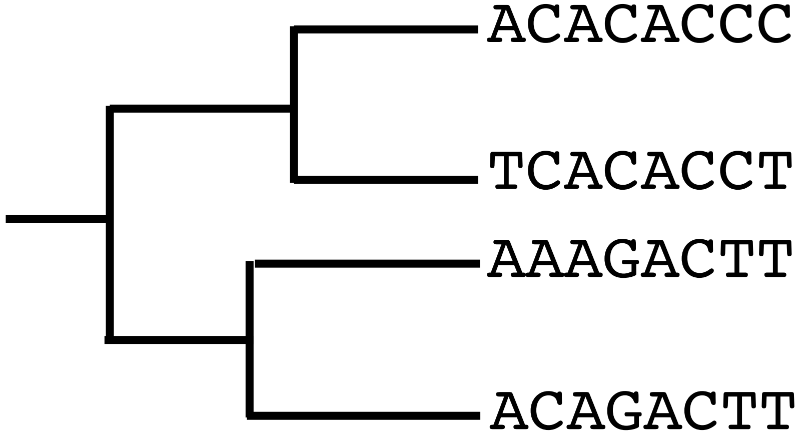
Phylo-
genetics

Input:

Pathogen genetic data from different hosts

Output:

Transmission chain (who infected whom)



Genetic data entering epidemiology

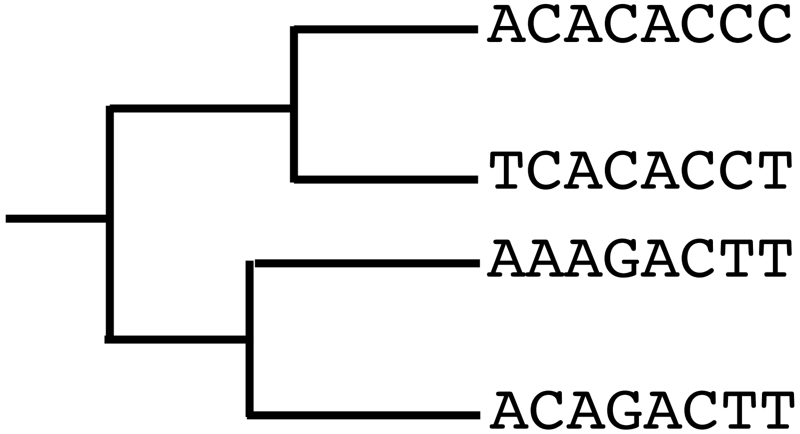
Phylo-
genetics

Input:

Pathogen genetic data from different hosts

Output:

Transmission chain (who infected whom)



Example
key result

Emergence of HIV

Epidemic identified in 1980s

Genetic data entering epidemiology

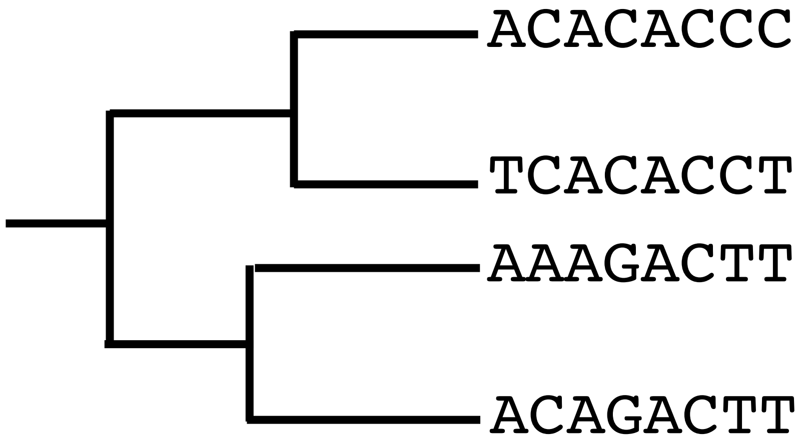
Phylo-genetics

Input:

Pathogen genetic data from different hosts

Output:

Transmission chain (who infected whom)

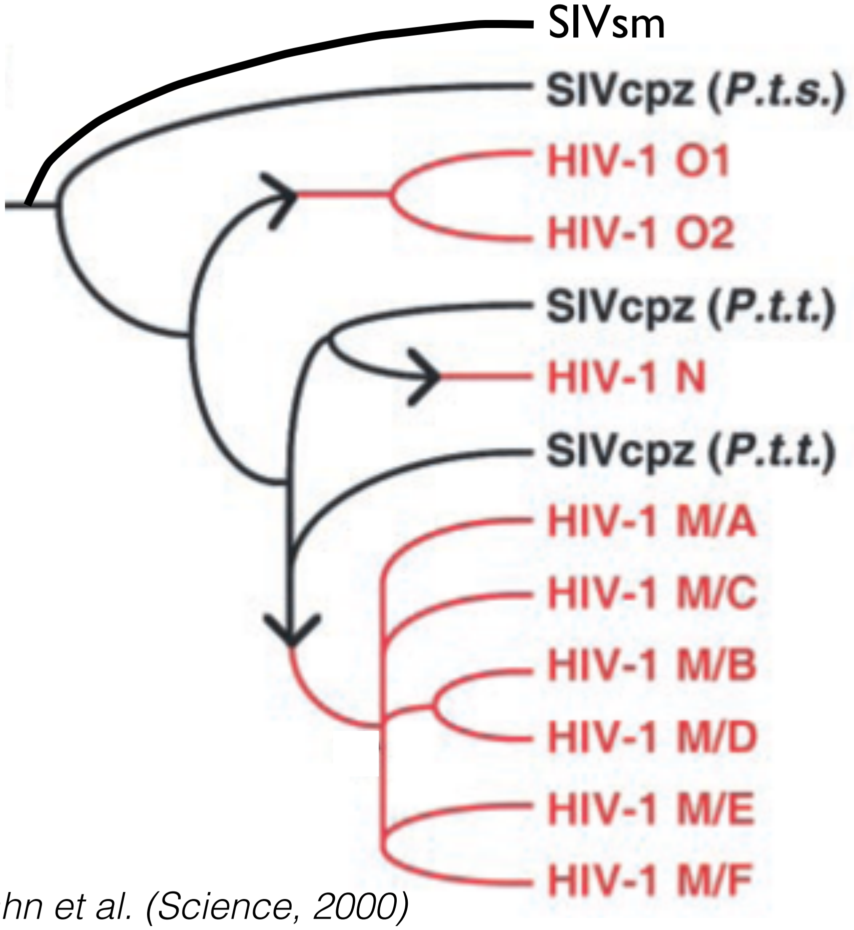


Example key result

Emergence of HIV

Epidemic identified in 1980s

Sequence data tells us about pre-1980



adopted from Hahn et al. (Science, 2000)

Genetic data entering epidemiology

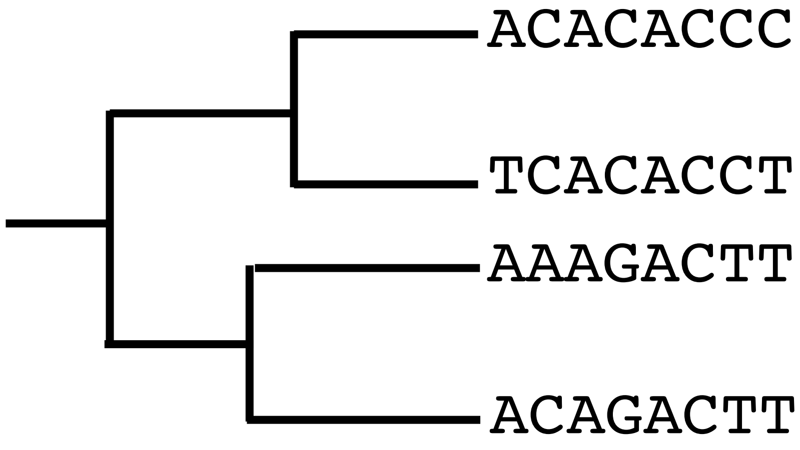
Phylo-genetics

Input:

Pathogen genetic data from different hosts

Output:

Transmission chain (who infected whom)



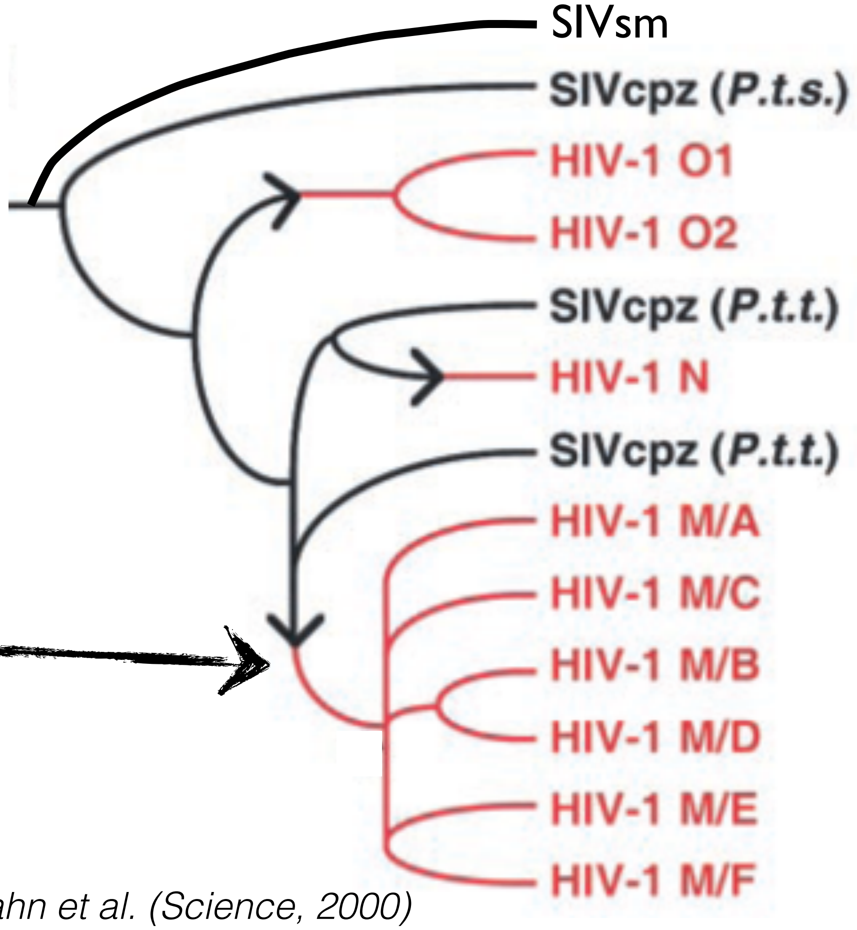
Example key result

Emergence of HIV

Epidemic identified in 1980s

Sequence data tells us about pre-1980

- Zoonosis from chimps
- In Africa
- Early 20th century



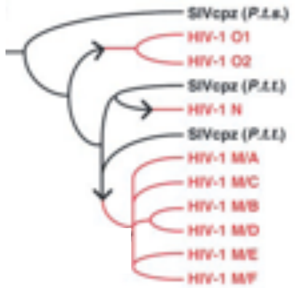
adopted from Hahn et al. (Science, 2000)

Genetic data contains information about the epidemiological dynamics

Phylo-
genetics

State of process

- Time and geographic location of outbreak



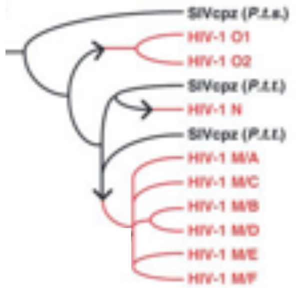
Emergence of a
pathogen
(here HIV)
Hahn et al. (Science, 2000)

Genetic data contains information about the epidemiological dynamics

Phylo-genetics

State of process

- Time and geographic location of outbreak



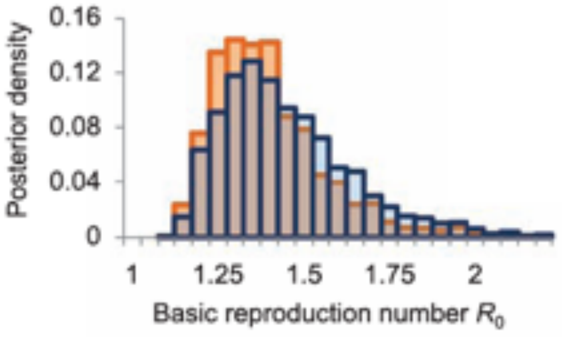
Emergence of a pathogen (here HIV)

Hahn et al. (Science, 2000)

Phylo-dynamics

Dynamics of process

- Transmission process



Basic reproductive number of an emerging pathogen (here H1N1)

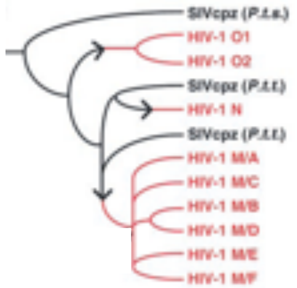
Fraser et al. (Science, 2009)

Genetic data contains information about the epidemiological dynamics

Phylo-genetics

State of process

- Time and geographic location of outbreak



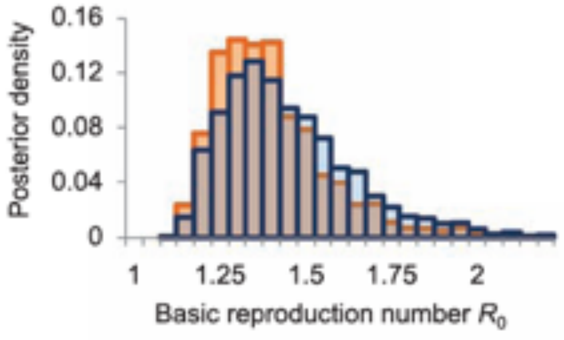
Emergence of a pathogen (here HIV)

Hahn et al. (Science, 2000)

Phylo-dynamics

Dynamics of process

- Transmission process



Basic reproductive number of an emerging pathogen (here H1N1)

Fraser et al. (Science, 2009)

Epidemio-logical models

Ideally: assume epidemiological model in phylogenetics (birth-death model)

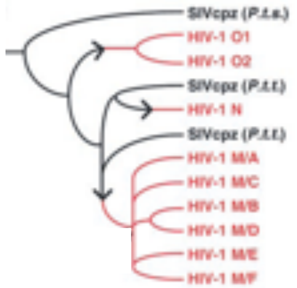
- Quantification of all model parameters!

Genetic data contains information about the epidemiological dynamics

Phylogenetics

State of process

- Time and geographic location of outbreak

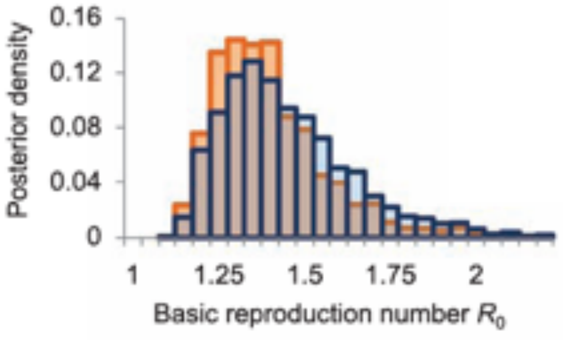


Emergence of a pathogen (here HIV)
Hahn et al. (Science, 2000)

Phylodynamics

Dynamics of process

- Transmission process



Basic reproductive number of an emerging pathogen (here H1N1)
Fraser et al. (Science, 2009)

Epidemiological models

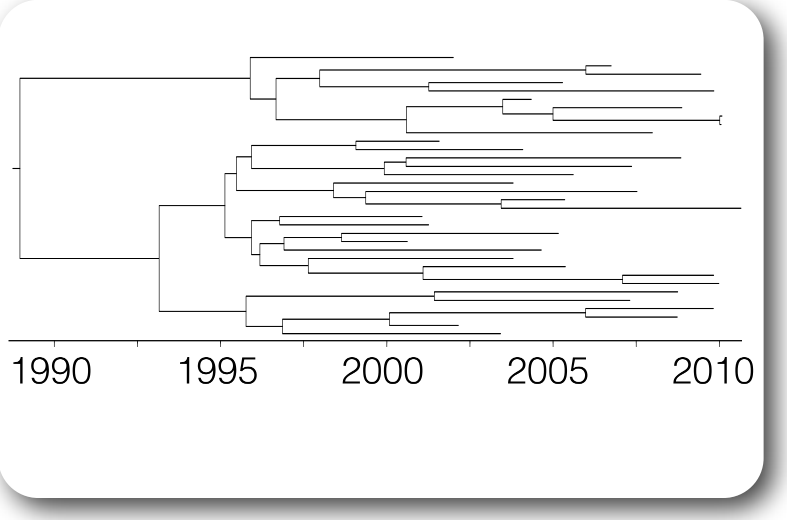
<u>Ideally:</u> assume epidemiological model in phylogenetics (birth-death model)

- Quantification of all model parameters!

<u>Approximate:</u> by population genetic model (coalescent)

- Only deterministic population size is parameterized!

Merge epidemiology and phylogenetics

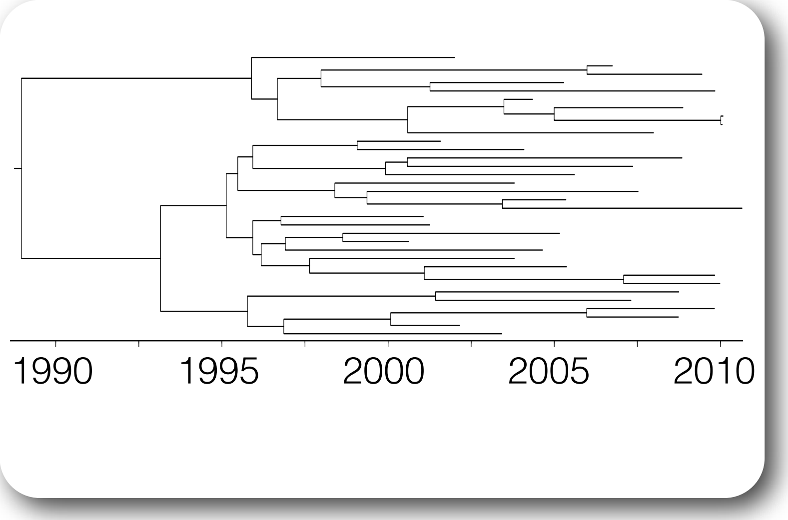


Methodology to:

1) Reconstruct phylogenetic trees

2) Fit epidemiological models to phylogenetic trees

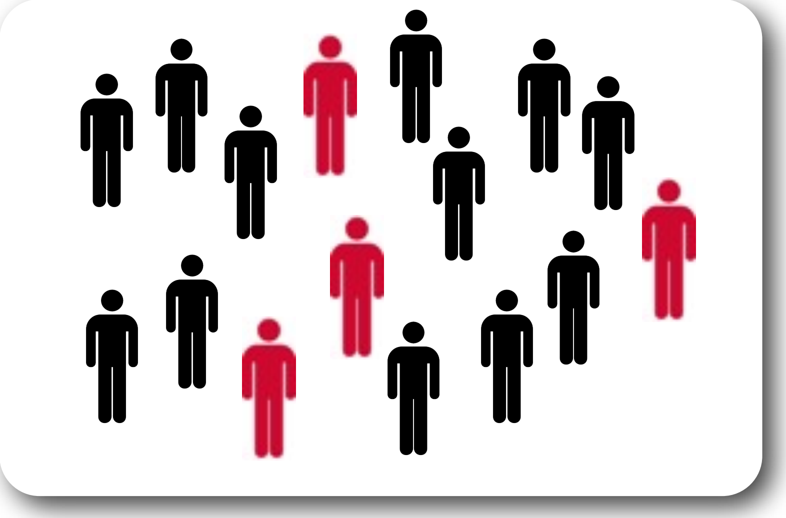
Merge epidemiology and phylogenetics



Methodology to:

1) Reconstruct phylogenetic trees

2) Fit epidemiological models to phylogenetic trees



Quantification of spread of infectious diseases?

Factors governing transmission dynamics?
(e.g. population structure, environment, vaccines)

General modeling framework

Epidemiological model

Growth of tree

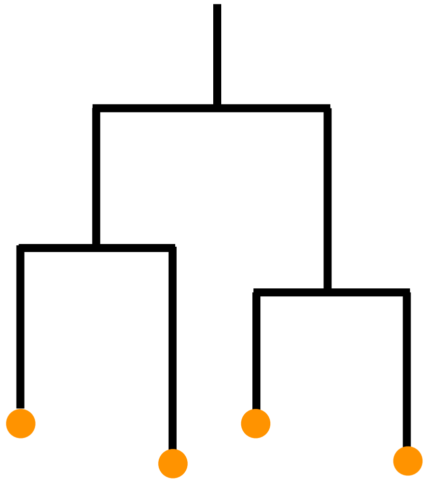
- parameters **η**

General modeling framework

Epidemiological model

Growth of tree

- parameters η



General modeling framework

Epidemiological model

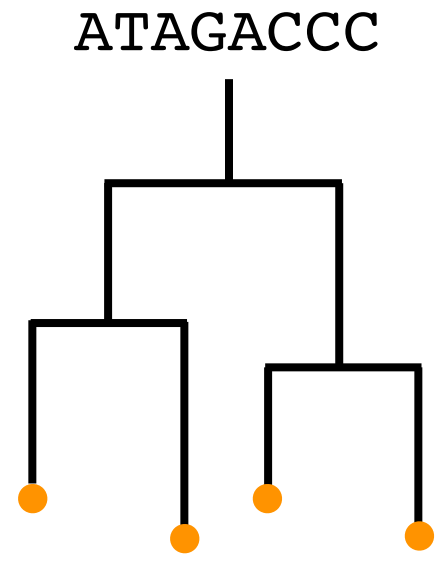
Growth of tree

- parameters η

Evolution-ary model

Evolution of sequences along tree

- parameters θ



General modeling framework

Epidemiological model

Growth of tree

- parameters η

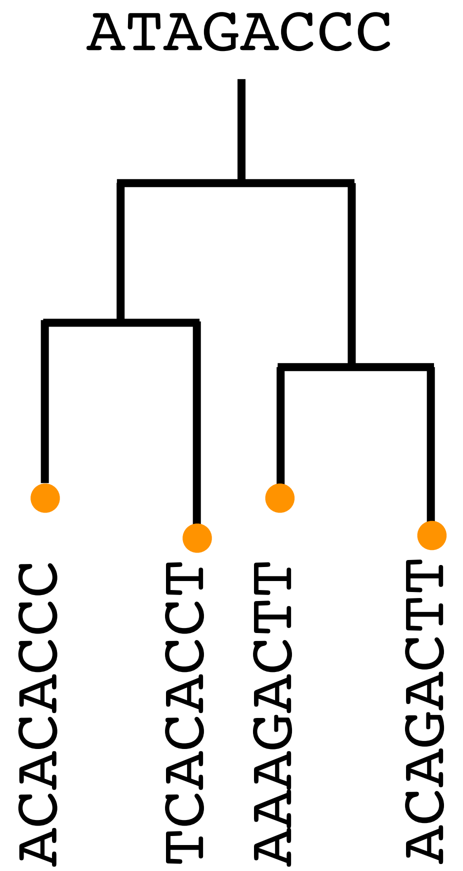
Evolutionary model

Evolution of sequences along tree

- parameters θ

Tree distribution

Defined through η and θ



General modeling framework

Epidemiological model

Growth of tree

- parameters η

Evolutionary model

Evolution of sequences along tree

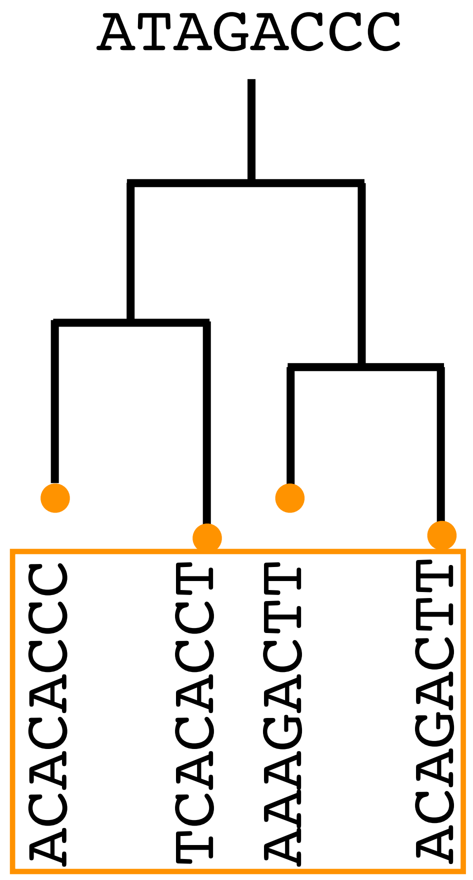
- parameters θ

Tree distribution

Defined through η and θ

Data

Sequence alignment



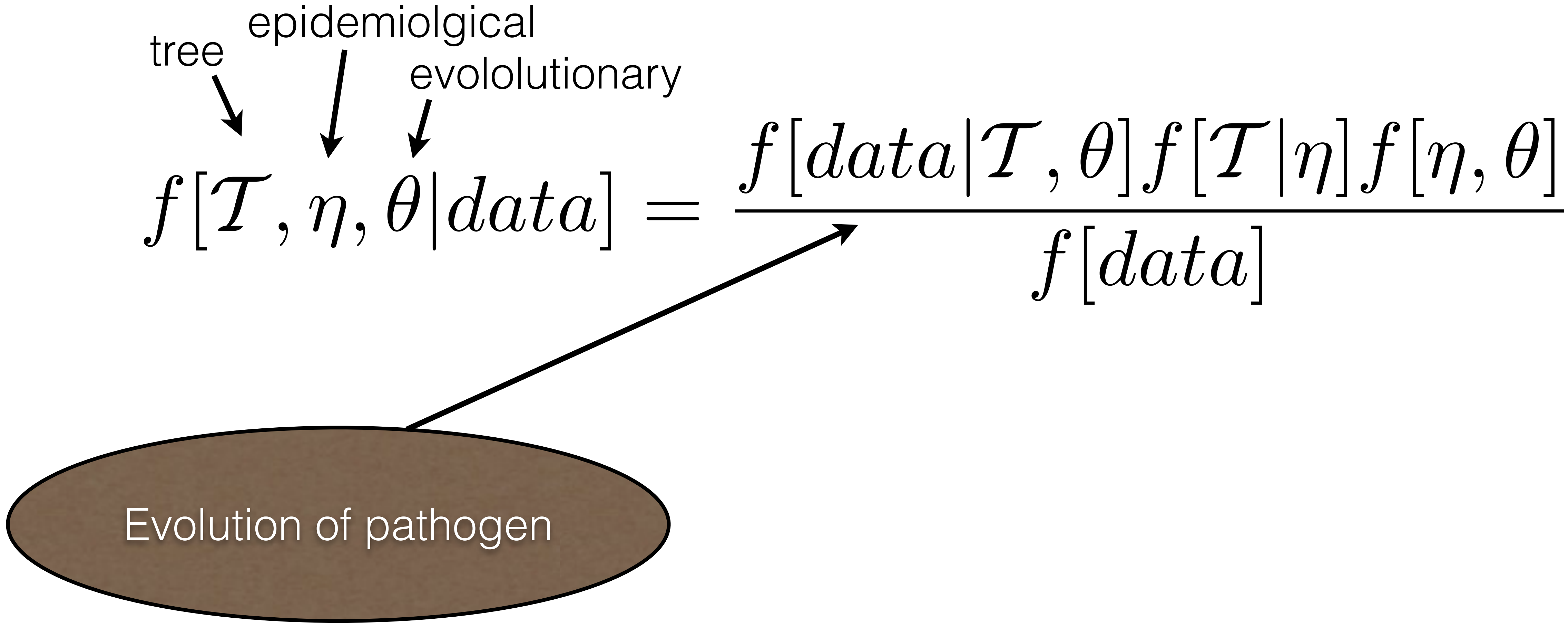
Bayesian approach for estimating epidemiological parameters

$$f[\overset{\text{tree}}{\overset{\searrow}{\mathcal{T}}}, \overset{\text{epidemiolgical}}{\overset{\downarrow}{\eta}}, \overset{\text{evololutionary}}{\overset{\downarrow}{\theta}}|data]$$

Bayesian approach for estimating epidemiological parameters

$$f[\overset{\substack{\text{tree}\\ \searrow}}{\mathcal{T}}, \overset{\substack{\text{epidemiolgical}\\ \downarrow}}{\eta}, \overset{\substack{\text{evolololutionary}\\ \swarrow}}{\theta}|data] = \frac{f[data|\mathcal{T}, \theta]f[\mathcal{T}|\eta]f[\eta, \theta]}{f[data]}$$

Bayesian approach for estimating epidemiological parameters



tree

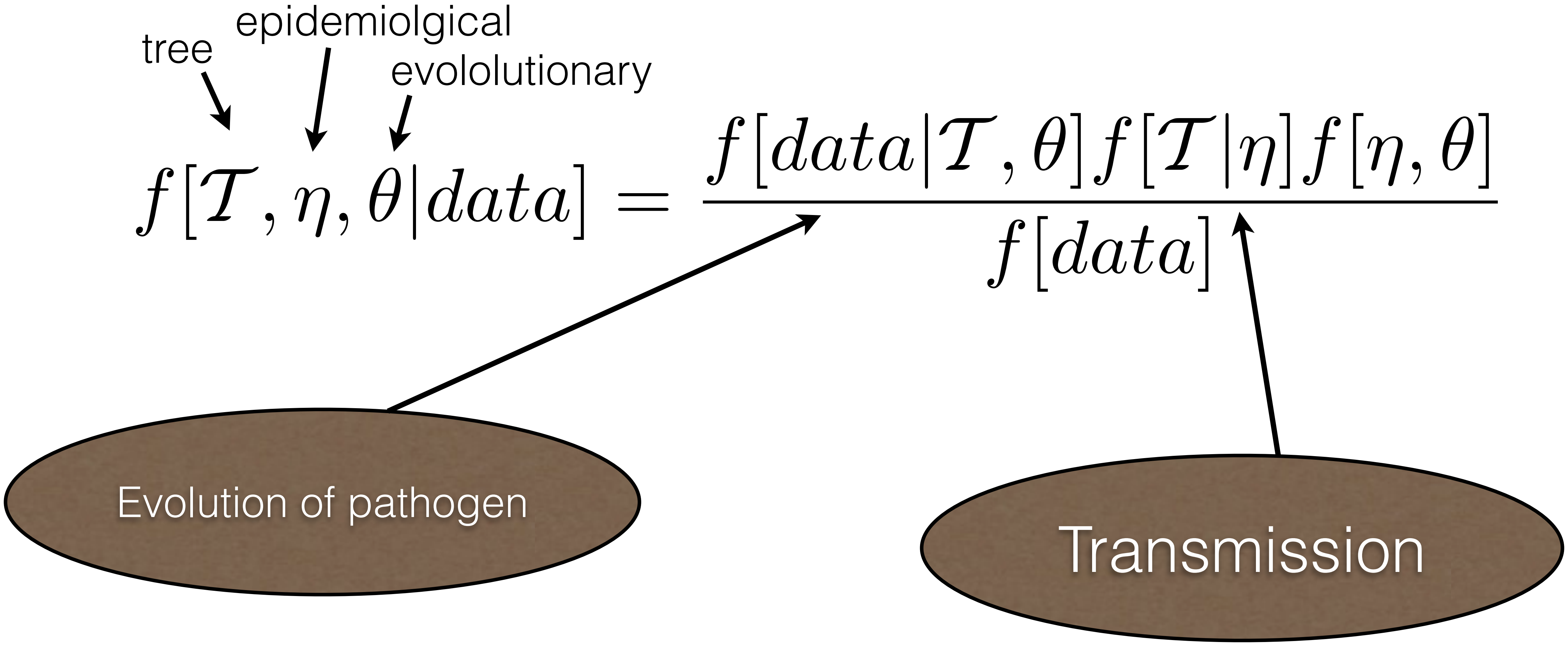
epidemiolgical

evolololutionary

$$f[\mathcal{T}, \eta, \theta | data] = \frac{f[data|\mathcal{T}, \theta] f[\mathcal{T}|\eta] f[\eta, \theta]}{f[data]}$$

Evolution of pathogen

Bayesian approach for estimating epidemiological parameters



tree

epidemiolgical

evololutionary

$$f[\mathcal{T}, \eta, \theta | data] = \frac{f[data|\mathcal{T}, \theta] f[\mathcal{T}|\eta] f[\eta, \theta]}{f[data]}$$

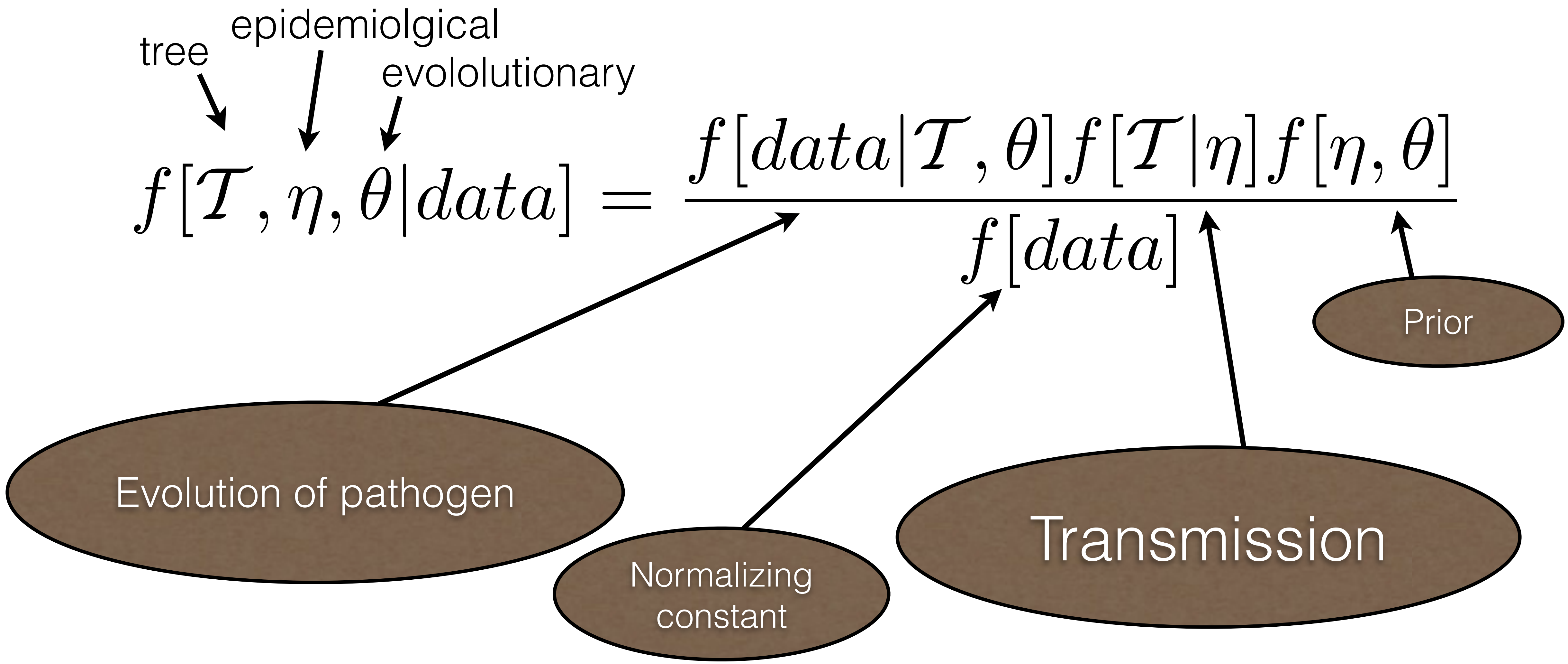
Evolution of pathogen

Transmission

Bayesian approach for estimating epidemiological parameters

$$f[\overset{\text{tree}}{\mathcal{T}}, \overset{\text{epidemiolgical}}{\eta}, \overset{\text{evololutionary}}{\theta}|data] = \frac{\underbrace{f[data|\mathcal{T},\theta]}_{\text{Evolution of pathogen}}\underbrace{f[\mathcal{T}|\eta]}_{\text{Transmission}}\underbrace{f[\eta,\theta]}_{\text{Prior}}}{f[data]}$$

Bayesian approach for estimating epidemiological parameters



tree

epidemiolgical

evolololutionary

$$f[\mathcal{T},\eta,\theta|data] = \frac{f[data|\mathcal{T},\theta]f[\mathcal{T}|\eta]f[\eta,\theta]}{f[data]}$$

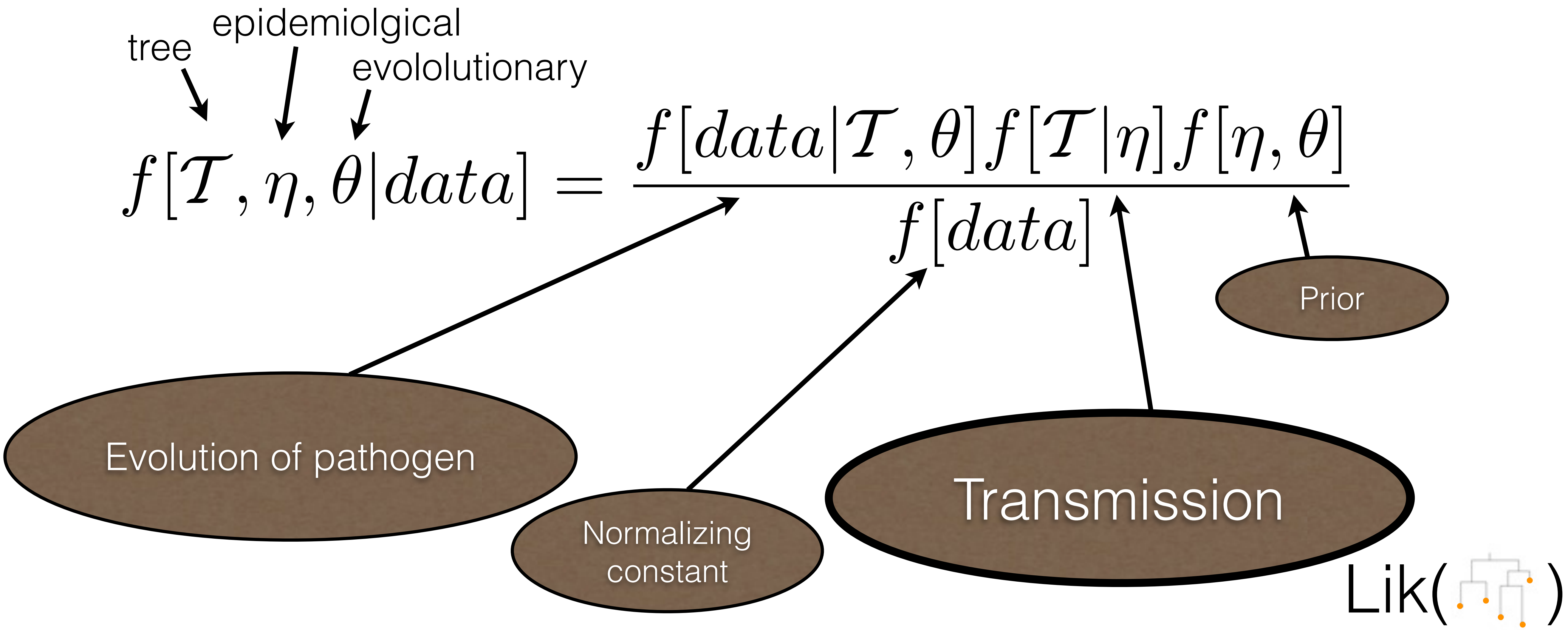
Evolution of pathogen

Normalizing constant

Transmission

Prior

Bayesian approach for estimating epidemiological parameters



tree

epidemiolgical

evolololutionary

$$f[\mathcal{T},\eta,\theta|data] = \frac{f[data|\mathcal{T},\theta]f[\mathcal{T}|\eta]f[\eta,\theta]}{f[data]}$$

Evolution of pathogen

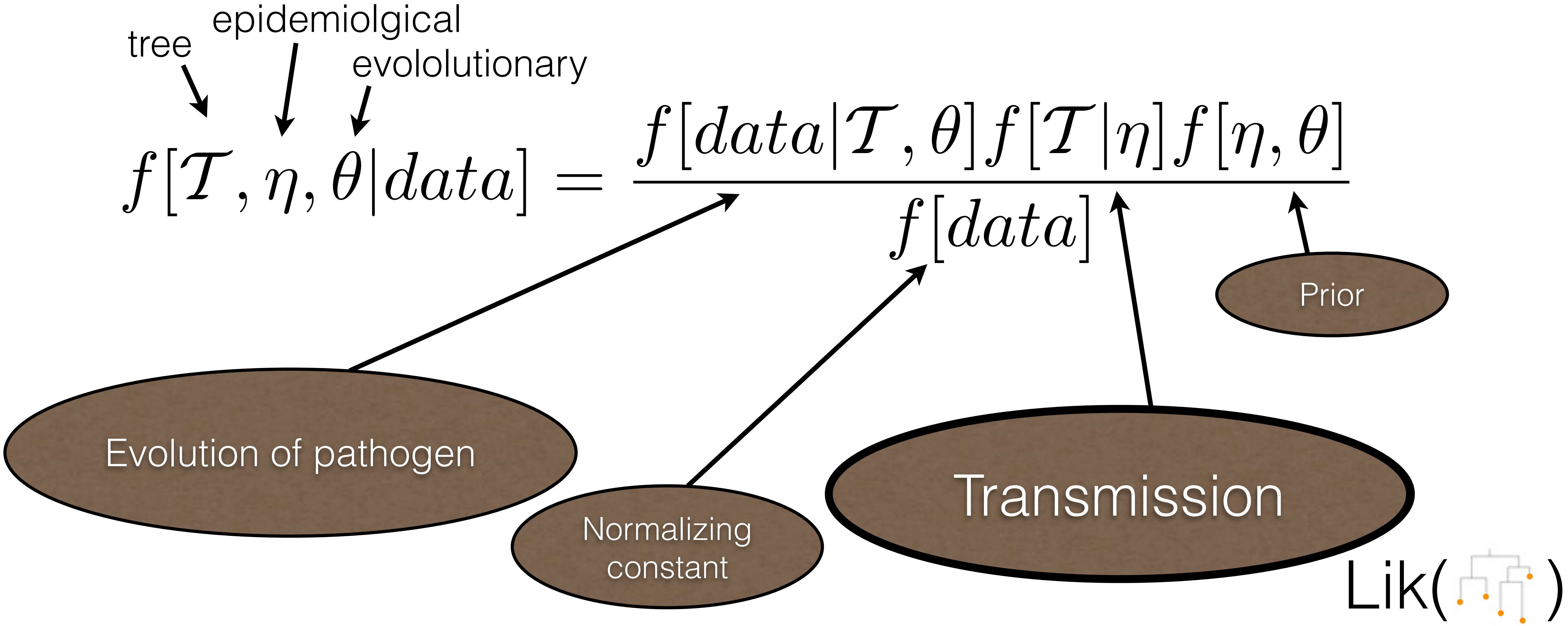
Normalizing constant

Transmission

Prior

Lik( )

Bayesian approach for estimating epidemiological parameters



tree

epidemiolgical

evololutionary

$$f[\mathcal{T}, \eta, \theta | data] = \frac{f[data|\mathcal{T}, \theta] f[\mathcal{T}|\eta] f[\eta, \theta]}{f[data]}$$

Evolution of pathogen

Normalizing constant

Transmission

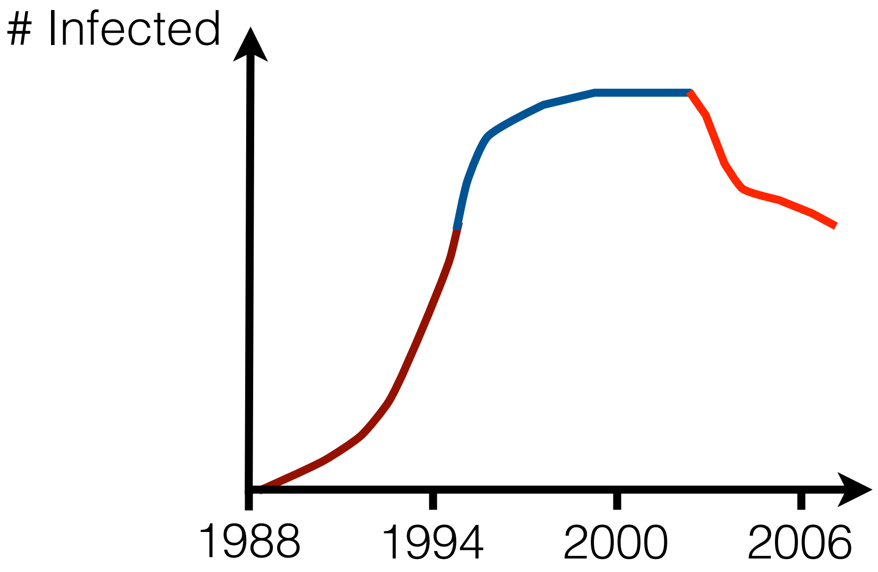
Prior

Lik( )

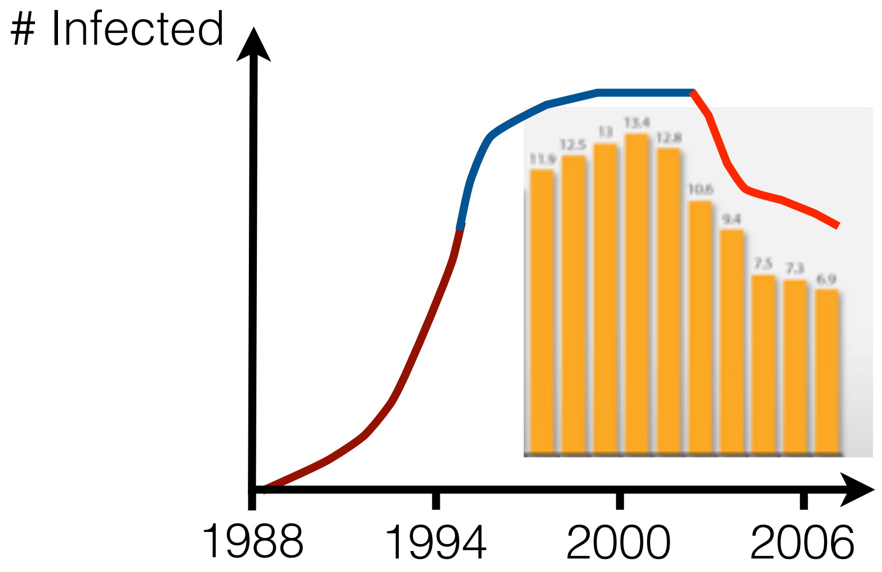
Implemented in Beast

Transmission model: coalescent or birth-death model

Phylogenetic methods for different epidemic dynamics

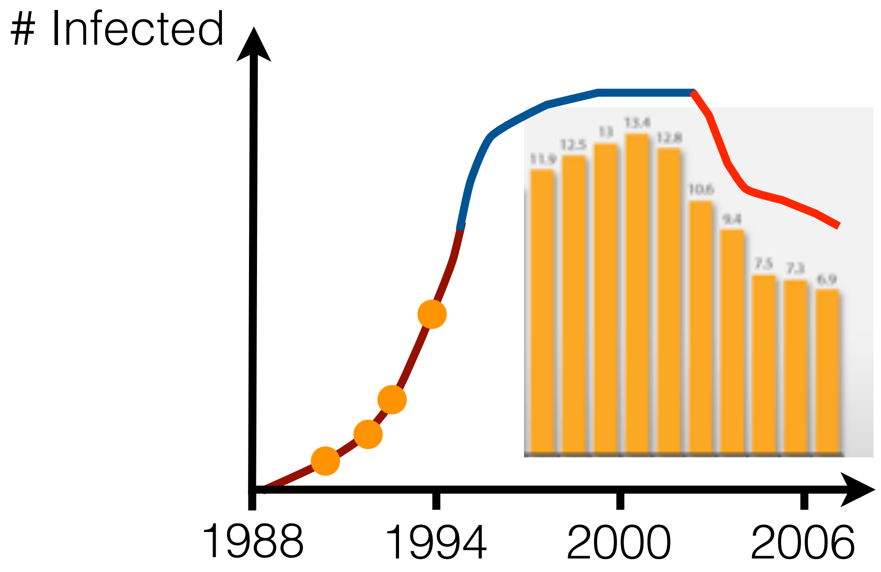


Phylogenetic methods for different epidemic dynamics

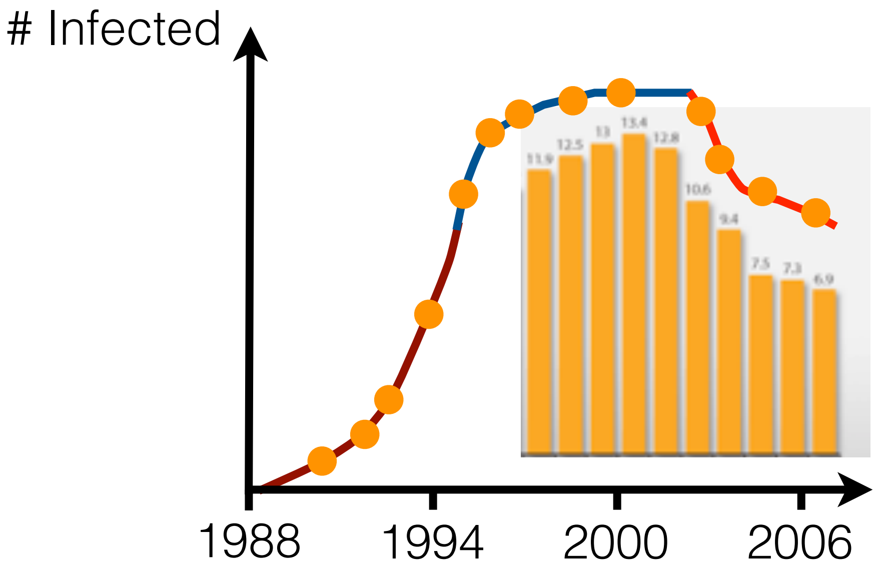


Phylogenetic methods for different epidemic dynamics

1. ...for analyzing epidemic outbreaks
    - *C: Drummond et al. (Genetics, 2002)*
    - *BD: Stadler et al. (MBE, 2012)*



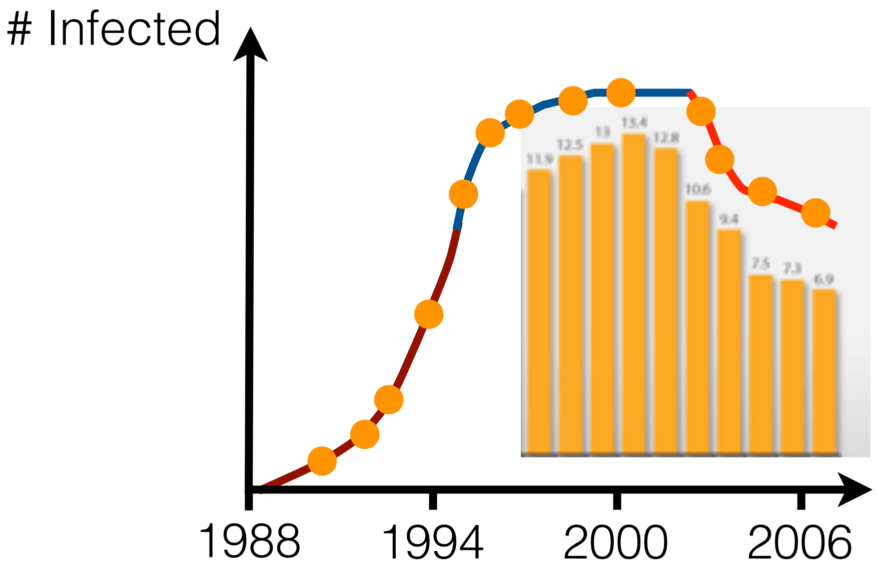
Phylogenetic methods for different epidemic dynamics



1. ...for analyzing epidemic outbreaks
    - *C: Drummond et al. (Genetics, 2002)*
    - *BD: Stadler et al. (MBE, 2012)*

2. ...for identifying epidemiological changes over time
    - *C: Drummond et al. (MBE, 2005)*
    - *BD: Stadler, Kühnert et al. (PNAS, 2013)*

Phylogenetic methods for different epidemic dynamics

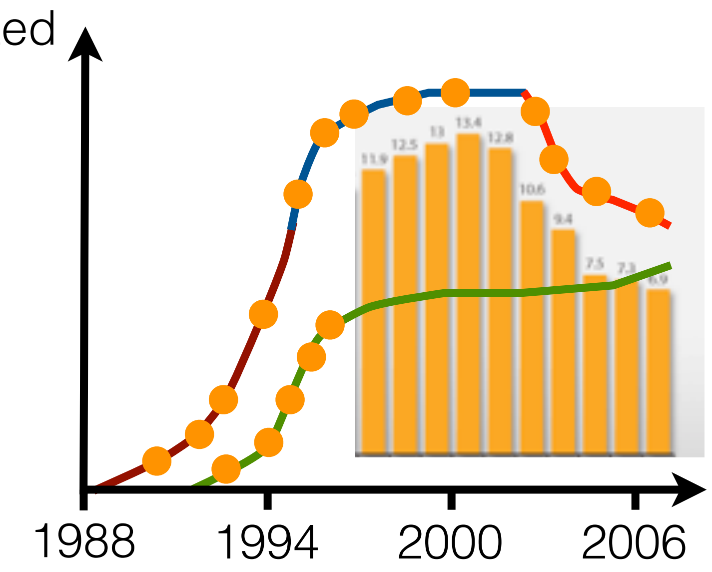


1. ...for analyzing epidemic outbreaks
    - *C: Drummond et al. (Genetics, 2002)*
    - *BD: Stadler et al. (MBE, 2012)*

2. ...for identifying epidemiological changes over time
    - *C: Drummond et al. (MBE, 2005)*
    - *BD: Stadler, Kühnert et al. (PNAS, 2013)*

3. ...accounting for SIR model dynamics
    - *C: Volz et al. (Genetics, 2009)*
    - *BD: Kühnert et al., Leventhal et al. (both submitted)*

Phylogenetic methods for different epidemic dynamics



1. ...for analyzing epidemic outbreaks
   - *C: Drummond et al. (Genetics, 2002)*
   - *BD: Stadler et al. (MBE, 2012)*

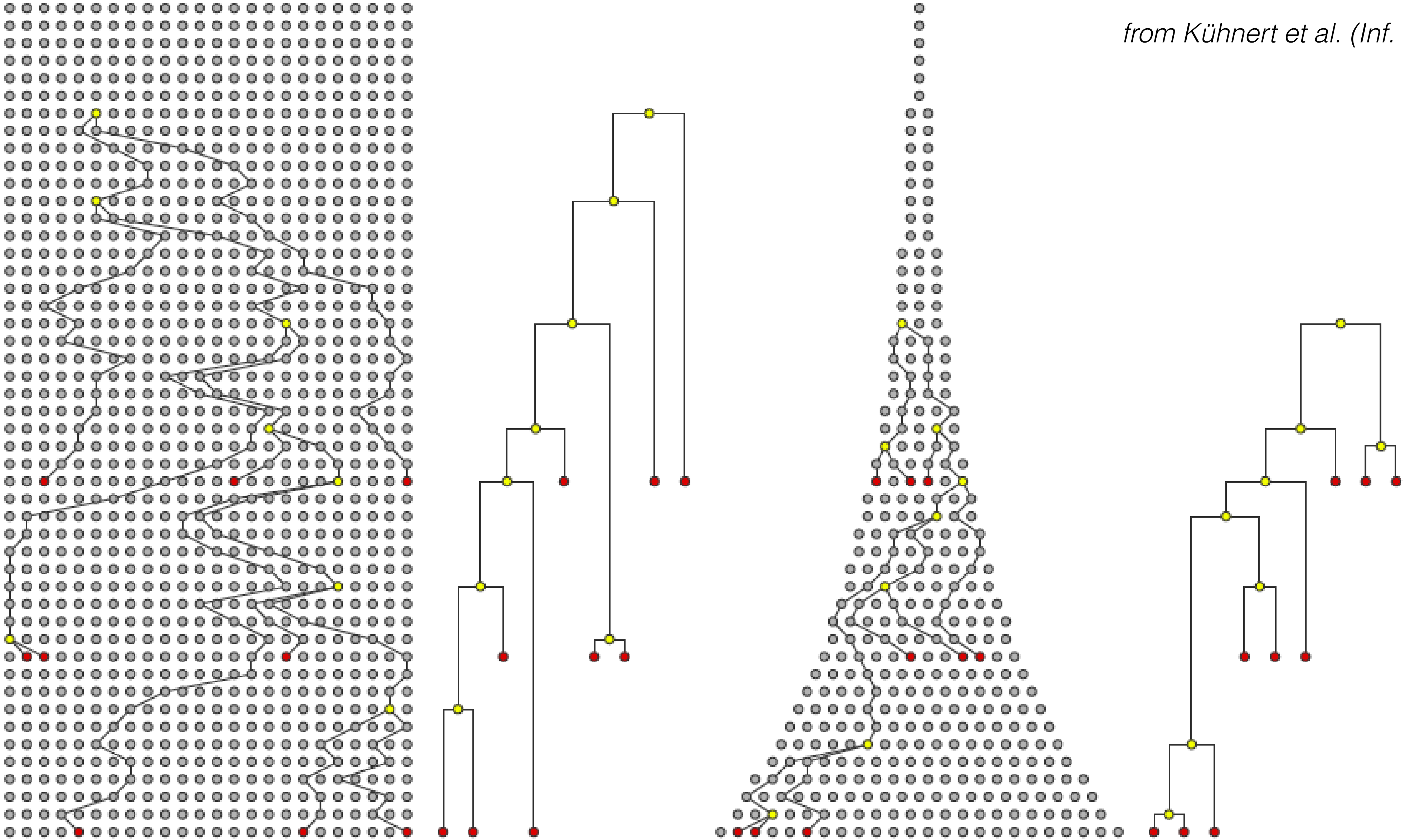
2. ...for identifying epidemiological changes over time
   - *C: Drummond et al. (MBE, 2005)*
   - *BD: Stadler, Kühnert et al. (PNAS, 2013)*

3. ...accounting for SIR model dynamics
   - *C: Volz et al. (Genetics, 2009)*
   - *BD: Kühnert et al., Leventhal et al. (both submitted)*

4. ...accounting for host population structure
   - *C: Volz (Genetics, 2012)*
   - *BD: Stadler & Bonhoeffer (Phil. Trans. Roy. Soc. B, 2013)*

Part A

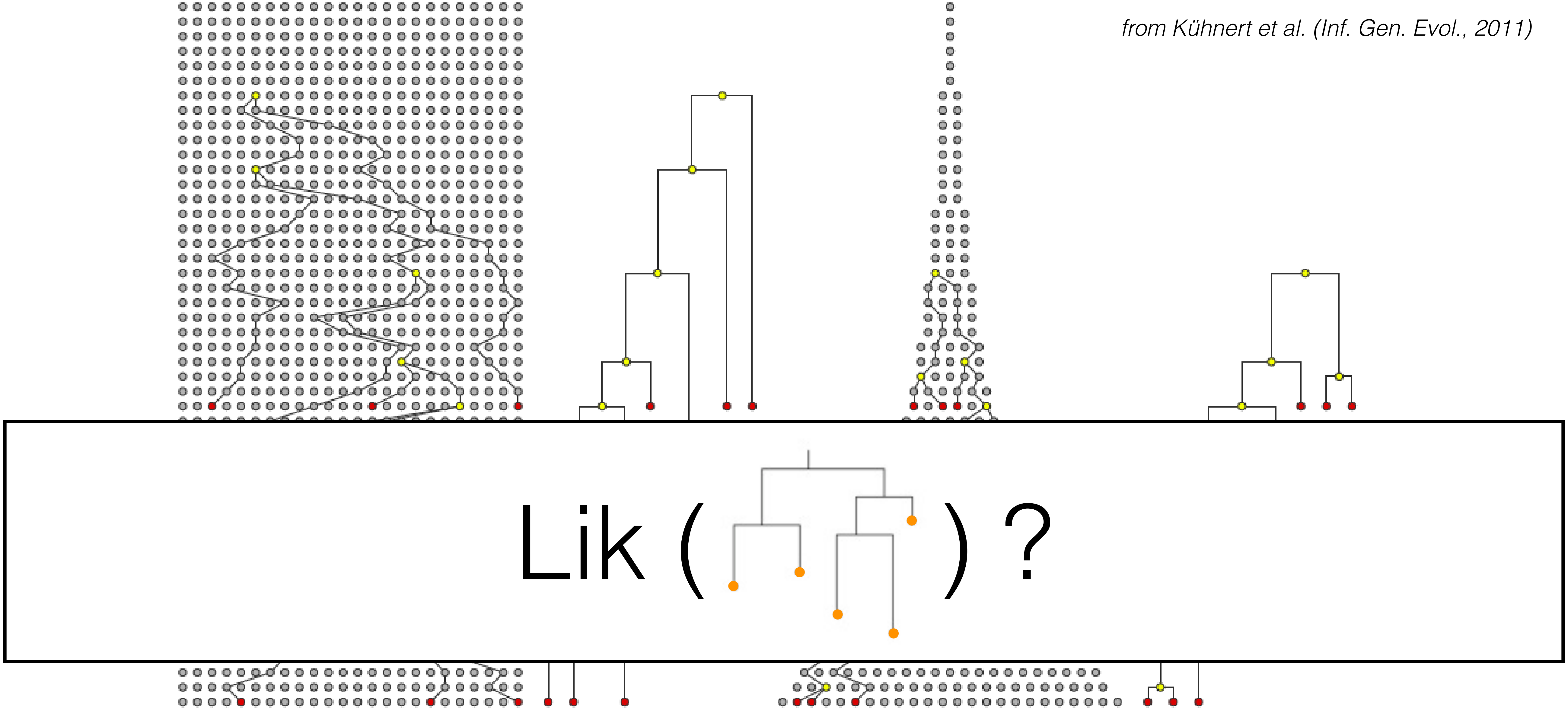
Coalescent as a model for transmission



from Kühnert et al. (Inf. Gen. Evol., 2011)

Part A

Coalescent as a model for transmission

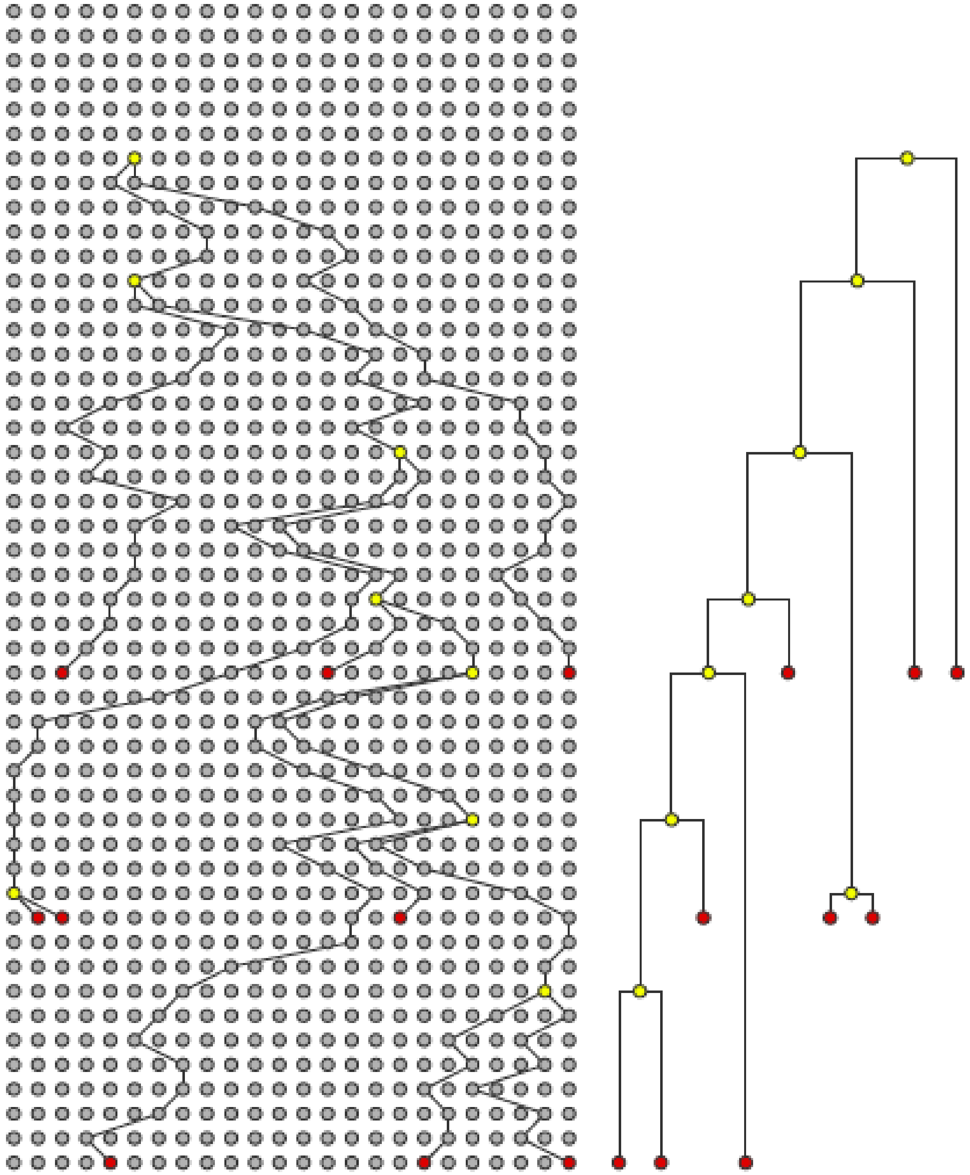


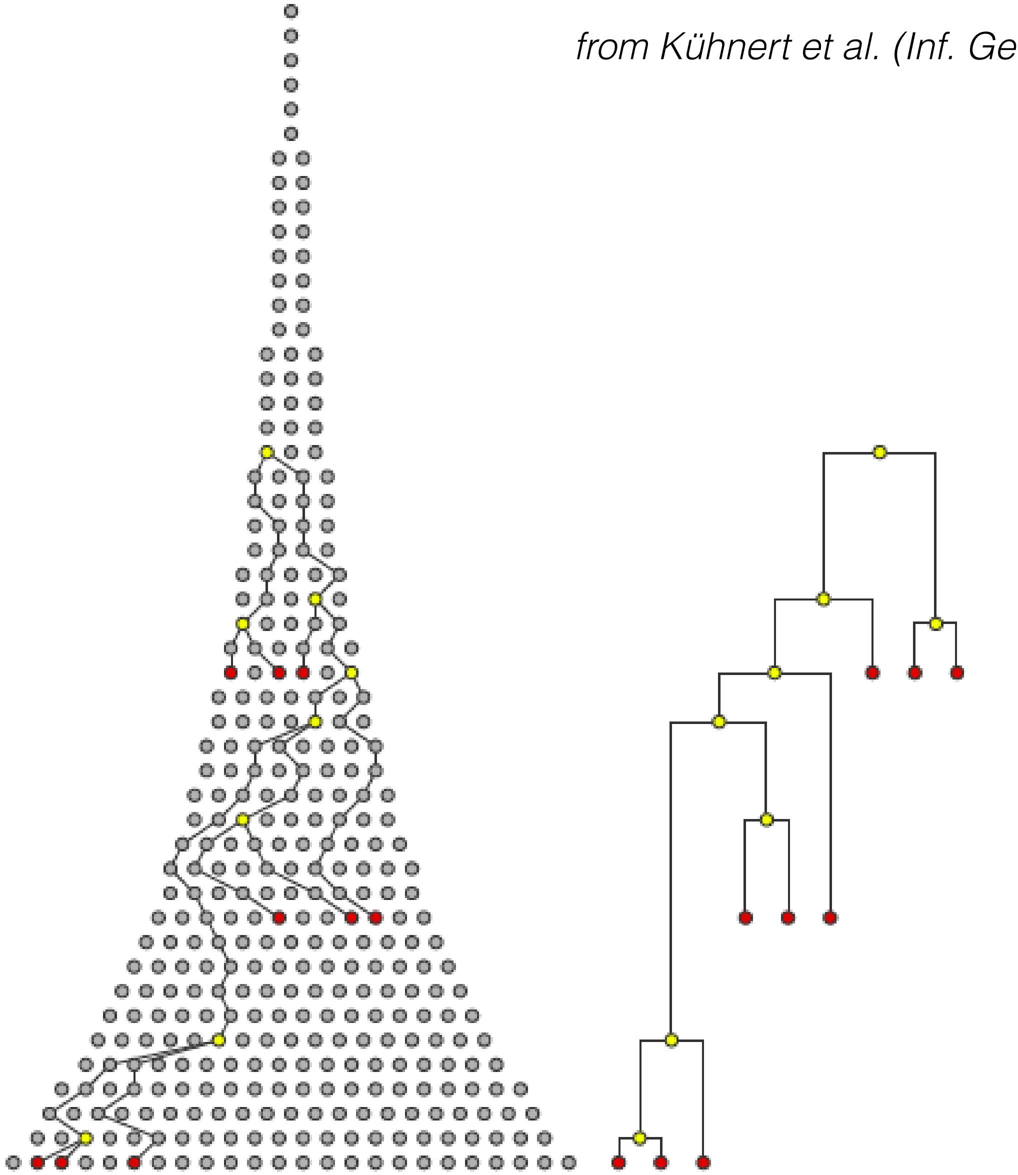
from Kühnert et al. (Inf. Gen. Evol., 2011)

Lik ( ) ?

Part A

Coalescent as a model for transmission





from Kühnert et al. (Inf. Gen. Evol., 2011)

$$f_G(g|\theta) = \frac{1}{\theta^{n-1}} \cdot \prod_{i=2}^{2n-1} e^{(-k_i(k_i-1)/2\theta)(t_i - t_{i-1})}$$

$$f_G(g|\theta,\, r) = \frac{1}{\theta^{n-1}} \cdot \prod_{i=2}^{2n-1} e^{rt_i}\, e^{(-k_i(k_i-1)/2\theta r)(e^{rt_i} - e^{rt_{i-1}})}$$

from Drummend et al. (Genetics, 2002)

Estimating R_0 from sequencing data

Hepatitis C

Little known before first isolation, ie. pre-1989.

How can we determine basic reproductive number R_0?

from Pybus et al. (Science, 2001)

Estimating R_0 from sequencing data

Hepatitis C

Little known before first isolation, ie. pre-1989.

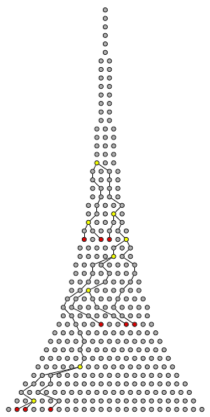
How can we determine basic reproductive number R_0?

Inference

40-100 sequences per subtype

Fit exponential-growth coalescent to sequences and obtain growth rate r (i.e. $\eta=r$)

$R_0 = rD+1$ (where D is expected time of infectiousness)



from Pybus et al. (Science, 2001)

Estimating R_0 from sequencing data

Hepatitis C

Little known before first isolation, ie. pre-1989.

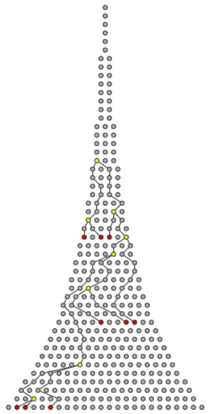
How can we determine basic reproductive number R_0?

Inference

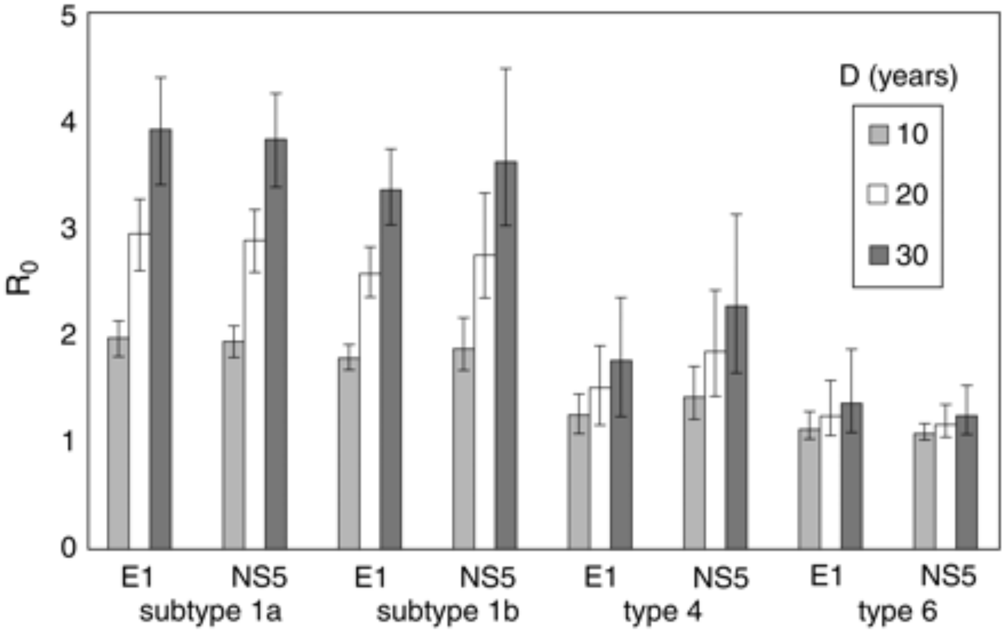
40-100 sequences per subtype

Fit exponential-growth coalescent to sequences and obtain growth rate r (i.e. $\eta=r$)

$R_0 = rD+1$ (where D is expected time of infectiousness)

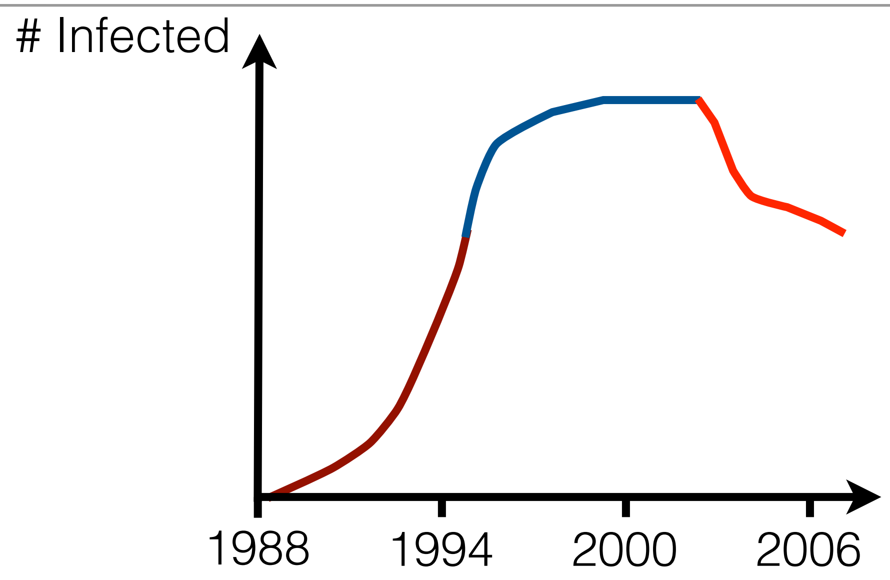


Results



from Pybus et al. (Science, 2001)

Bayesian skyline plot relaxes constant growth rate

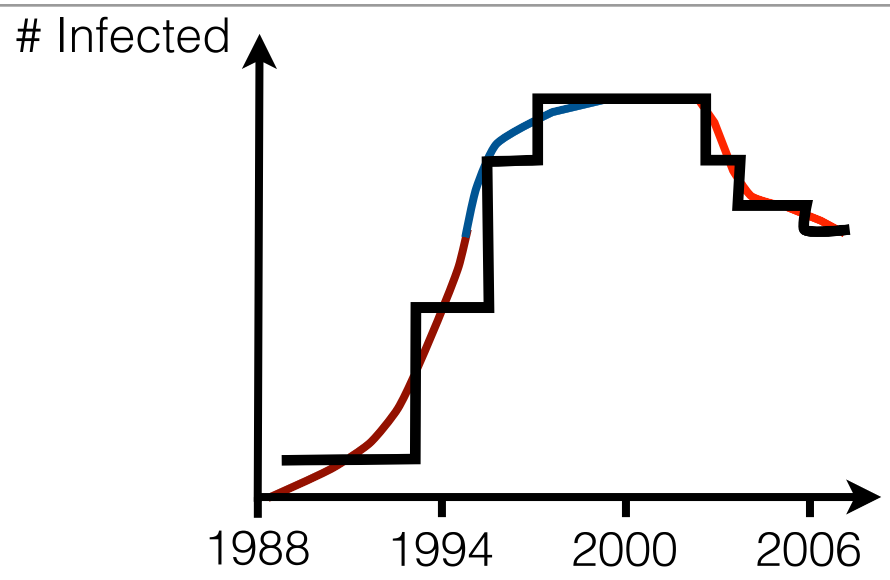


from Drummond et al. (MBE, 2005)

Bayesian skyline plot relaxes constant growth rate

Assump-
tion

Piecewise constant population size

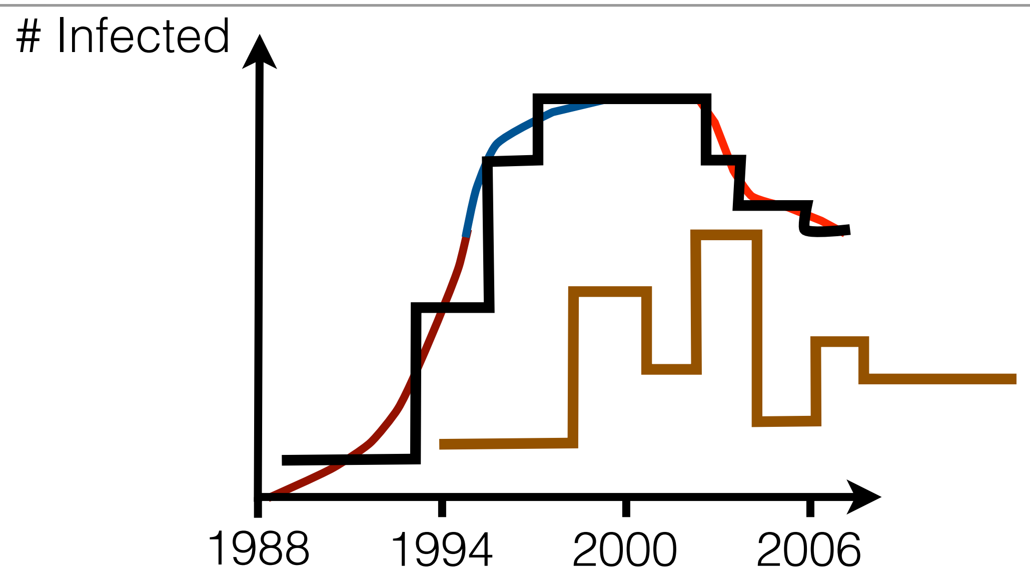


from Drummond et al. (MBE, 2005)

Bayesian skyline plot relaxes constant growth rate

Assump-
tion

Piecewise constant population size

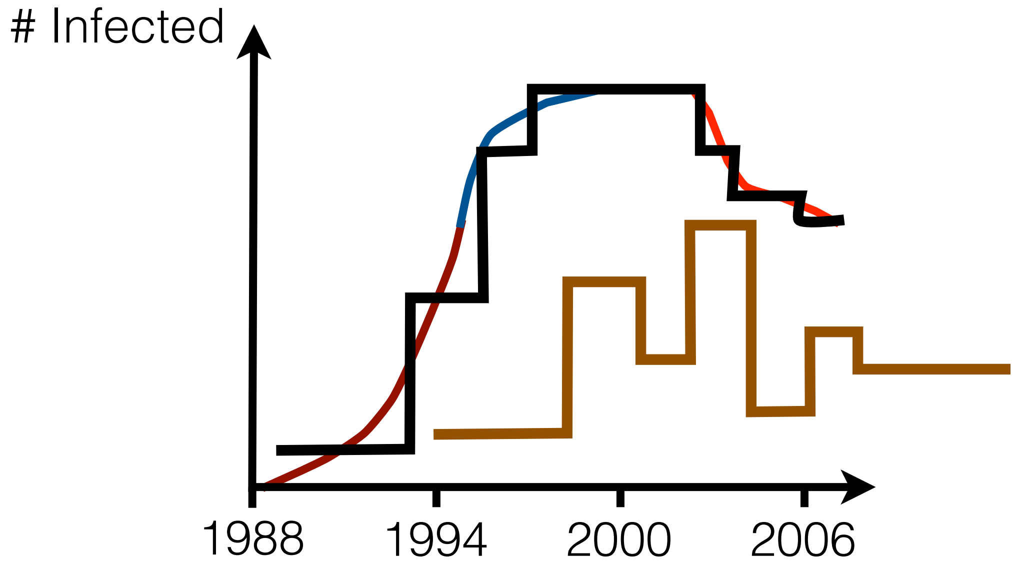


from Drummond et al. (MBE, 2005)

Bayesian skyline plot relaxes constant growth rate

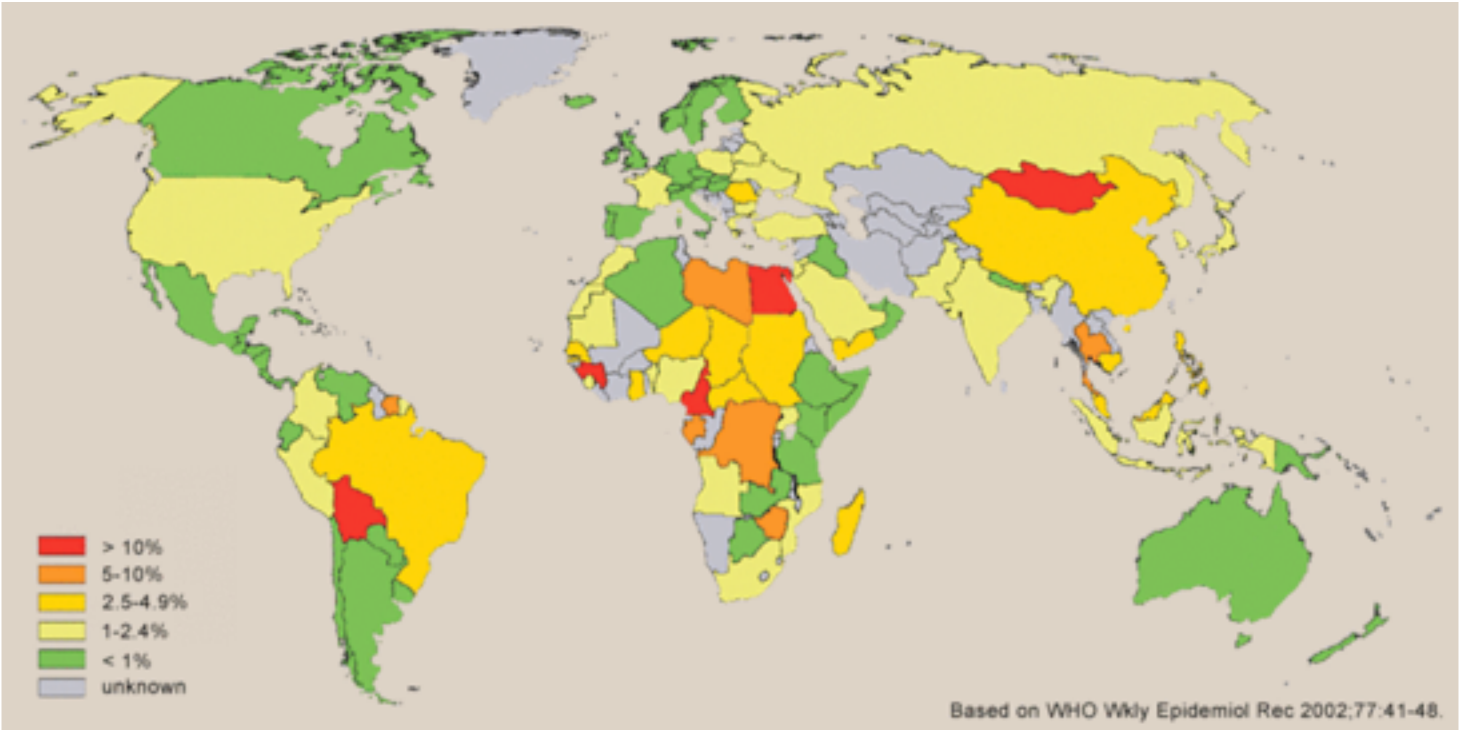
Assumption

Piecewise constant population size



HCV in Egypt

- Egypt highest HCV prevalence worldwide (about 13 %)
- Neighboring countries low prevalence
- When did the high prevalence emerge?
- Why?



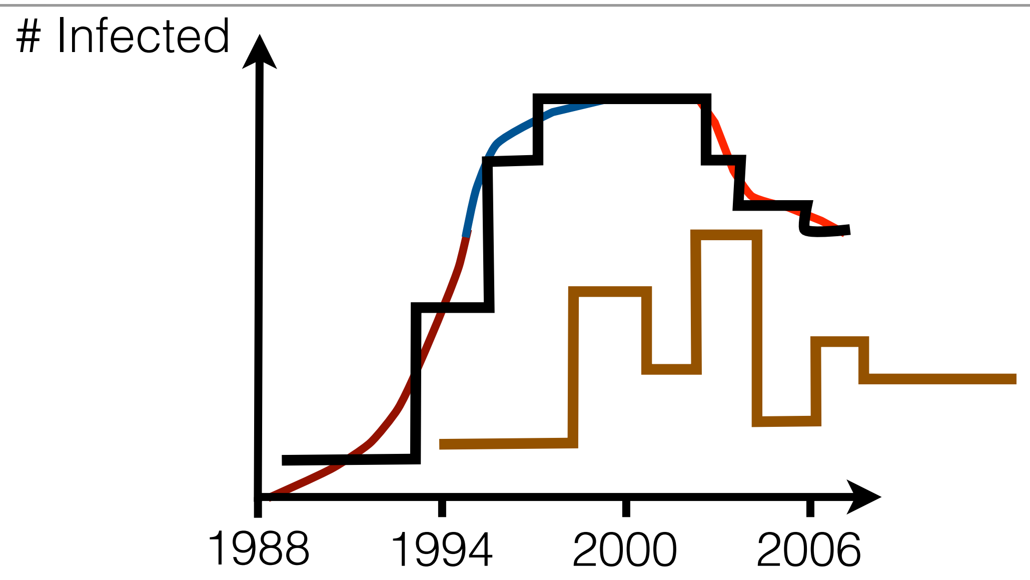
Based on WHO Wkly Epidemiol Rec 2002;77:41-48.

from Drummond et al. (MBE, 2005)

Bayesian skyline plot relaxes constant growth rate

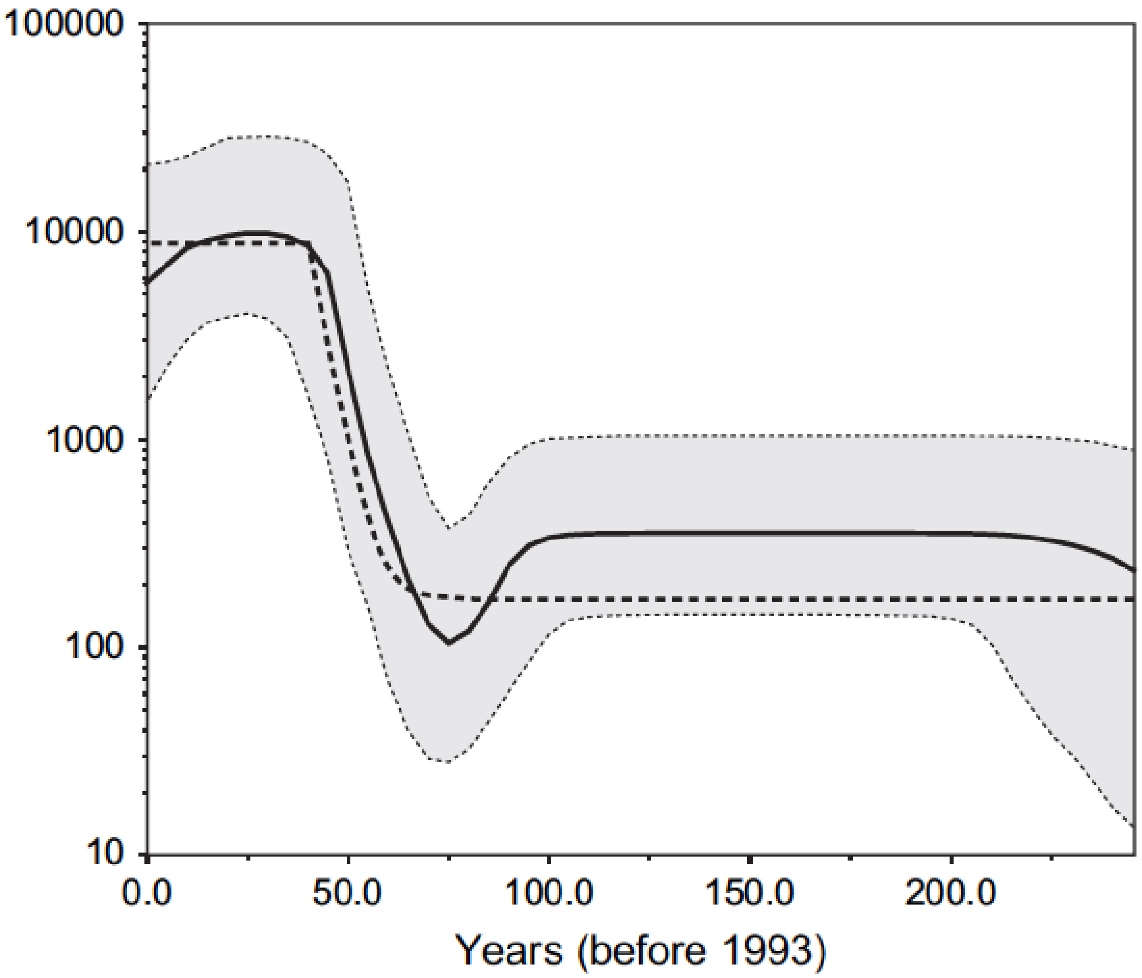
Assump-tion

Piecewise constant population size



HCV in Egypt

- Egypt highest HCV prevalence worldwide (about 13 %)
- Neighboring countries low prevalence
- When did the high prevalence emerge?
- Why?



from Drummond et al. (MBE, 2005)

Limitations of classic coalescent

Coalescent

Coalescent rates only depend on # infected, i.e. $\lambda_0 I$!

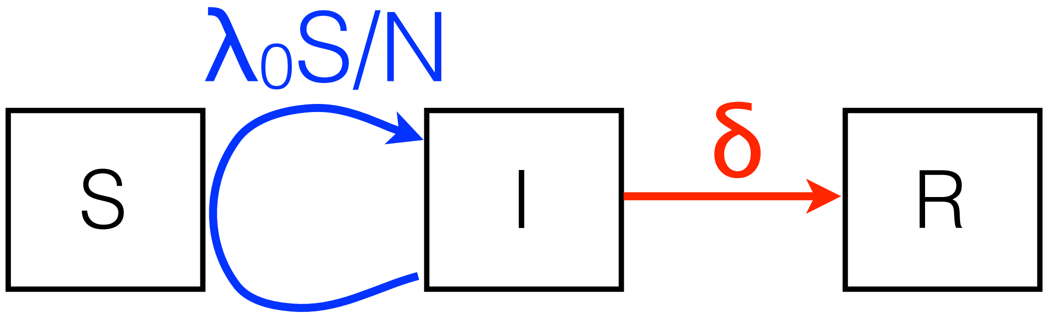
from Volz et al. (Genetics, 2009)

Limitations of classic coalescent

Coales-cent

Coalescent rates only depend on # infected, i.e. $\lambda_0 I$!

SIR extension



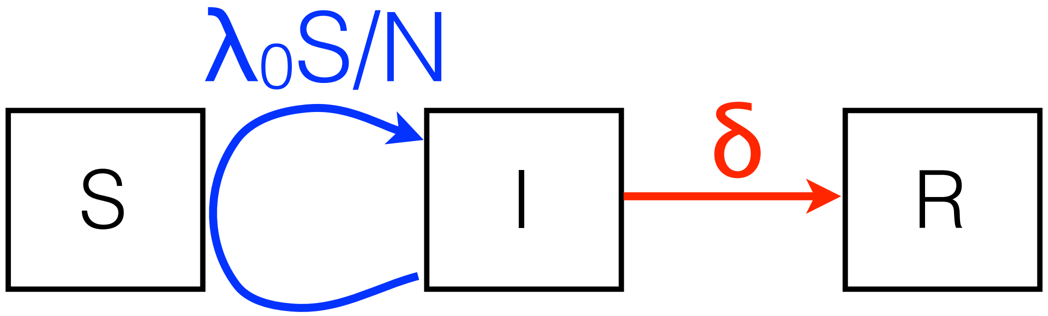
from Volz et al. (Genetics, 2009)

Limitations of classic coalescent

Coalescent

Coalescent rates only depend on # infected, i.e. $\lambda_0 I$!

SIR extension



overall coalescent rate: $\lambda_0 IS/N$

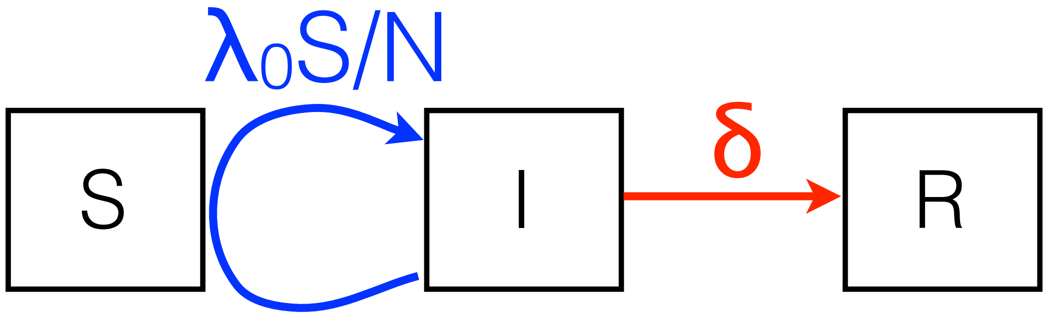
from Volz et al. (Genetics, 2009)

Limitations of classic coalescent

Coalescent

Coalescent rates only depend on # infected, i.e. $\lambda_0 I$!

SIR extension



$\lambda_0 S/N$

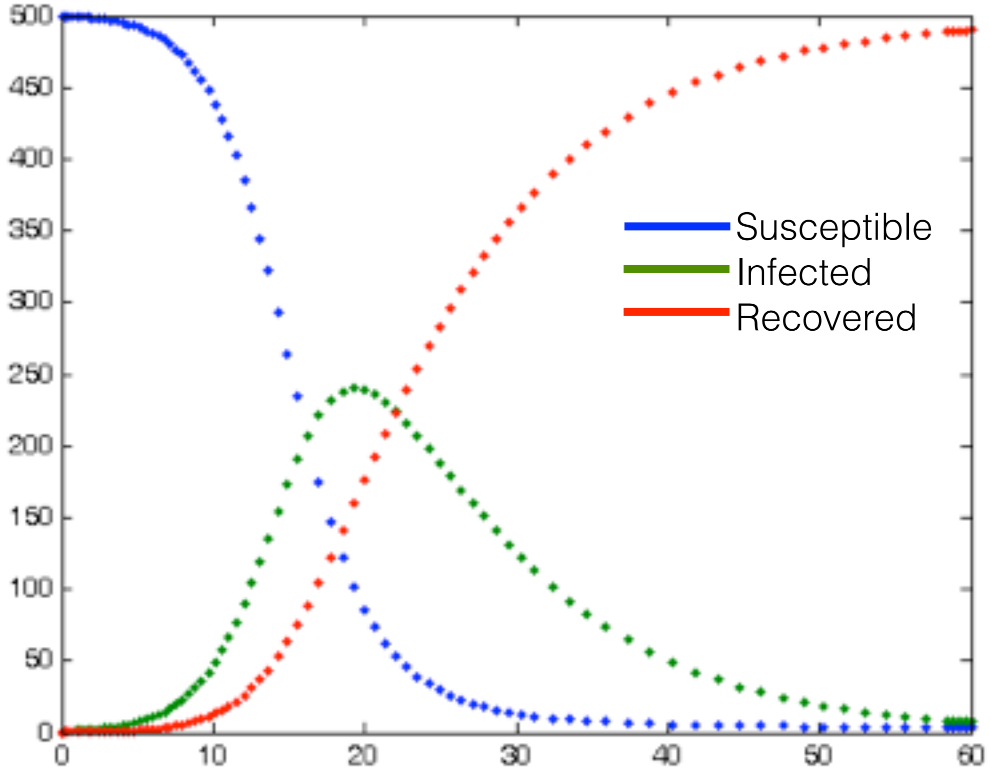
S

I

δ

R

overall coalescent rate: $\lambda_0 IS/N$



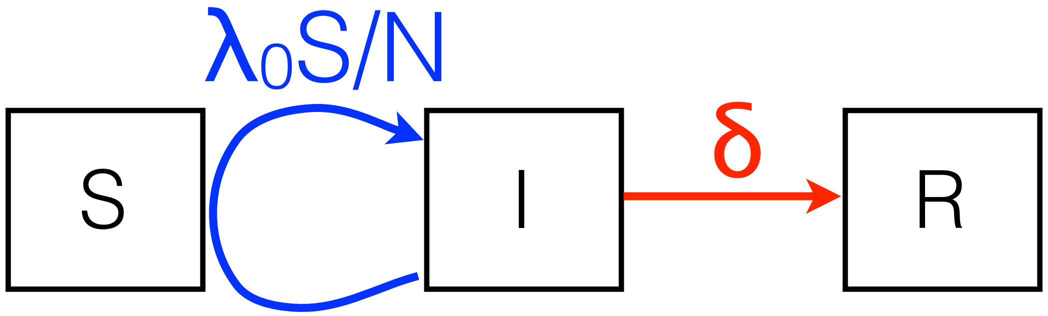
from Volz et al. (Genetics, 2009)

Limitations of classic coalescent

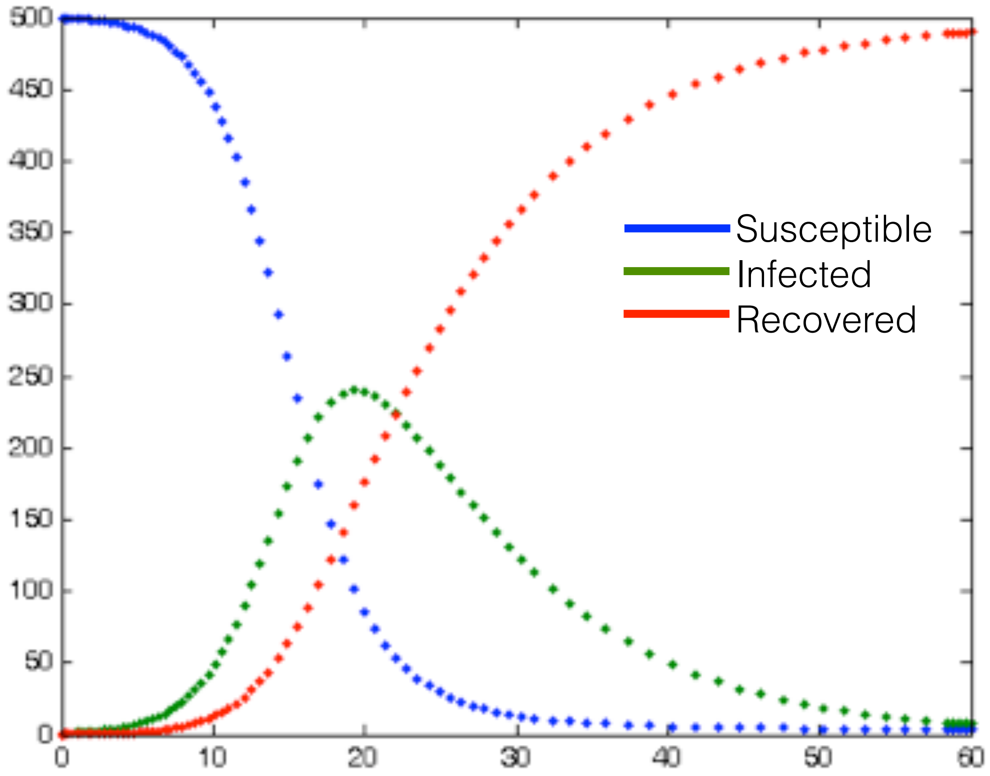
Coalescent

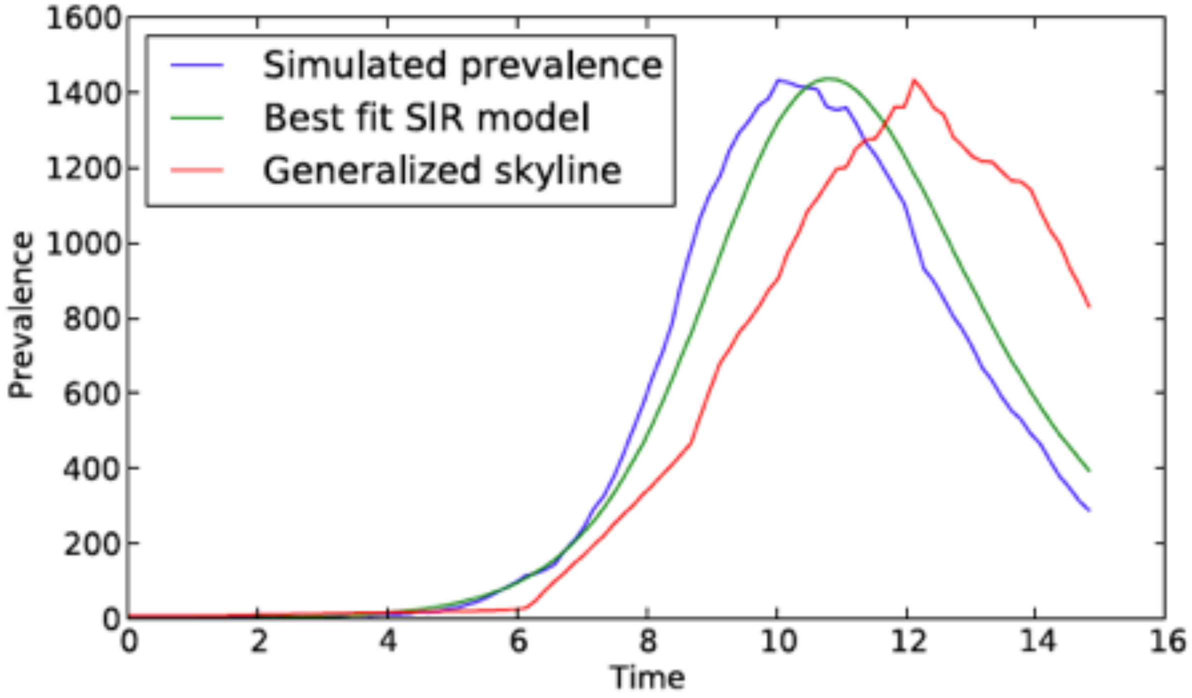
Coalescent rates only depend on # infected, i.e. $\lambda_0 I$!

SIR extension



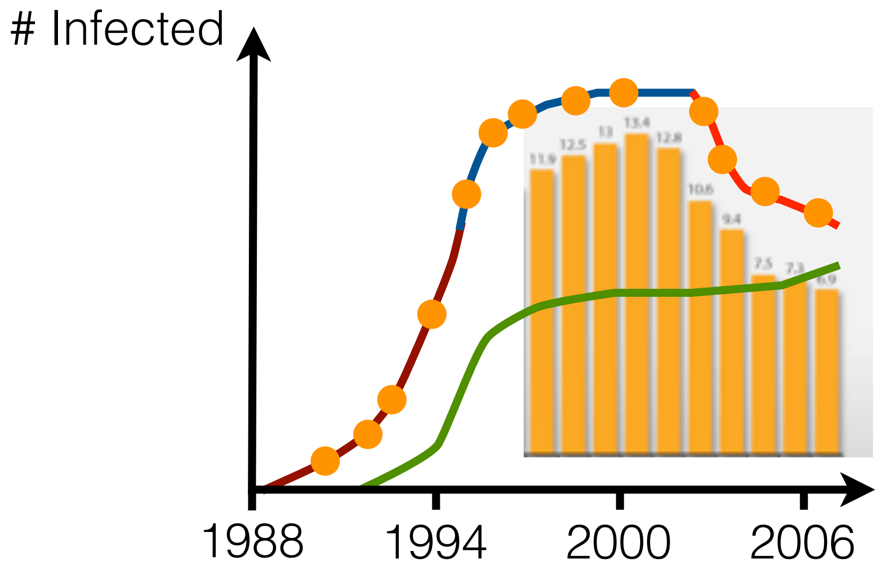
overall coalescent rate: $\lambda_0 IS/N$





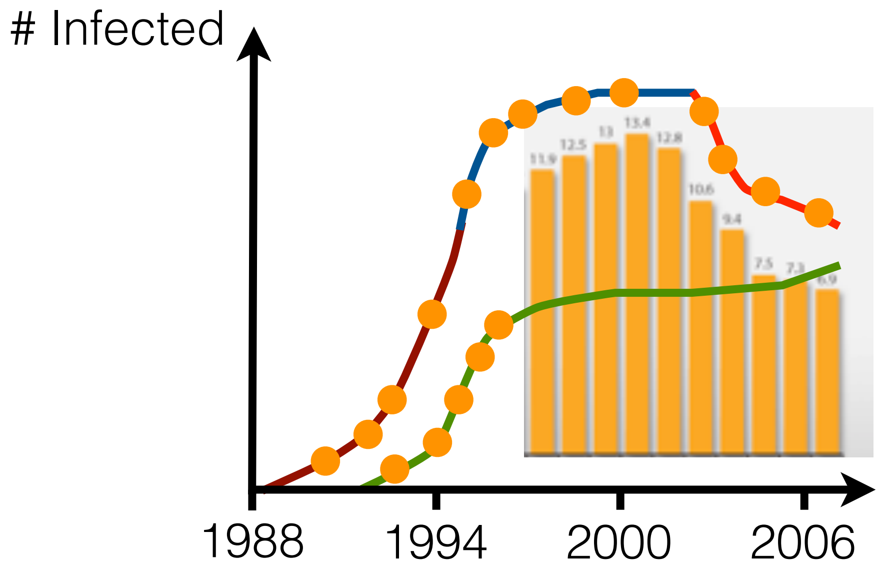
from Volz et al. (Genetics, 2009)

Phylogenetic methods for different epidemic dynamics



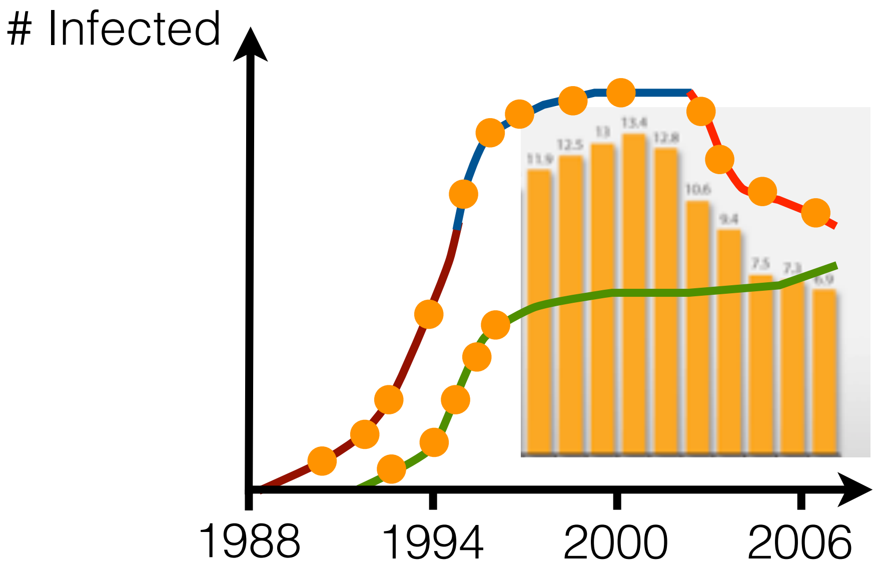
1. ...for analyzing epidemic outbreaks
    - *C: Drummond et al. (Genetics, 2002)*
2. ...for identifying epidemiological changes over time
    - *C: Drummond et al. (MBE, 2005)*
3. ...accounting for SIR model dynamics
    - *C: Volz et al. (Genetics, 2009)*
4. ...accounting for host population structure
    - *C: Volz (Genetics, 2012)*

Phylogenetic methods for different epidemic dynamics



1. ...for analyzing epidemic outbreaks
    - *C: Drummond et al. (Genetics, 2002)*
2. ...for identifying epidemiological changes over time
    - *C: Drummond et al. (MBE, 2005)*
3. ...accounting for SIR model dynamics
    - *C: Volz et al. (Genetics, 2009)*
4. ...accounting for host population structure
    - *C: Volz (Genetics, 2012)*

Phylogenetic methods for different epidemic dynamics

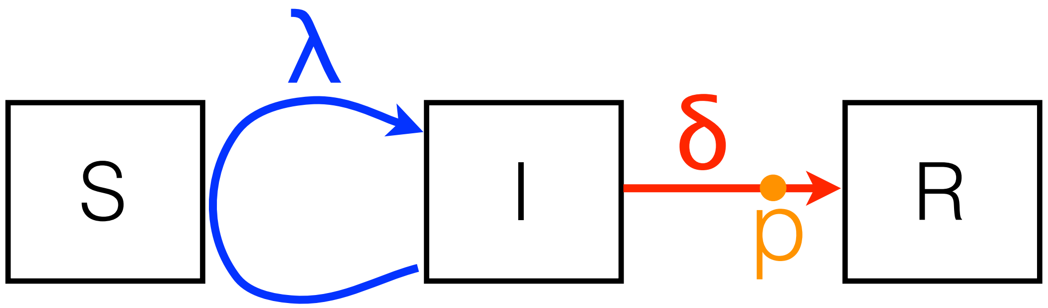


1. ...for analyzing epidemic outbreaks
    - *C: Drummond et al. (Genetics, 2002)*
    - *BD: Stadler et al. (MBE, 2012)*
2. ...for identifying epidemiological changes over time
    - *C: Drummond et al. (MBE, 2005)*
    - *BD: Stadler, Kühnert et al. (PNAS, 2013)*
3. ...accounting for SIR model dynamics
    - *C: Volz et al. (Genetics, 2009)*
    - *BD: Kühnert et al., Leventhal et al. (both submitted)*
4. ...accounting for host population structure
    - *C: Volz (Genetics, 2012)*
    - *BD: Stadler & Bonhoeffer (Phil. Trans. Roy. Soc. B, 2013)*

Part 2

Birth-death model for transmission

Birth-death model



- transmission rate λ
- “becoming-non-infectious” rate δ
- sampling probability p

Parameters may depend on:

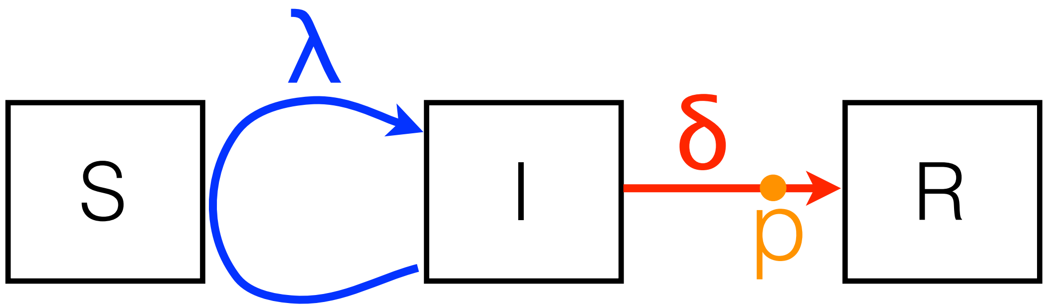
1) constant (epidemic outbreak)
2) time (environmental effect)
3) # of susceptibles (SIR model)
4) type of infected individual

Stadler (JTB, 2010)

Part 2

Birth-death model for transmission

Birth-death model

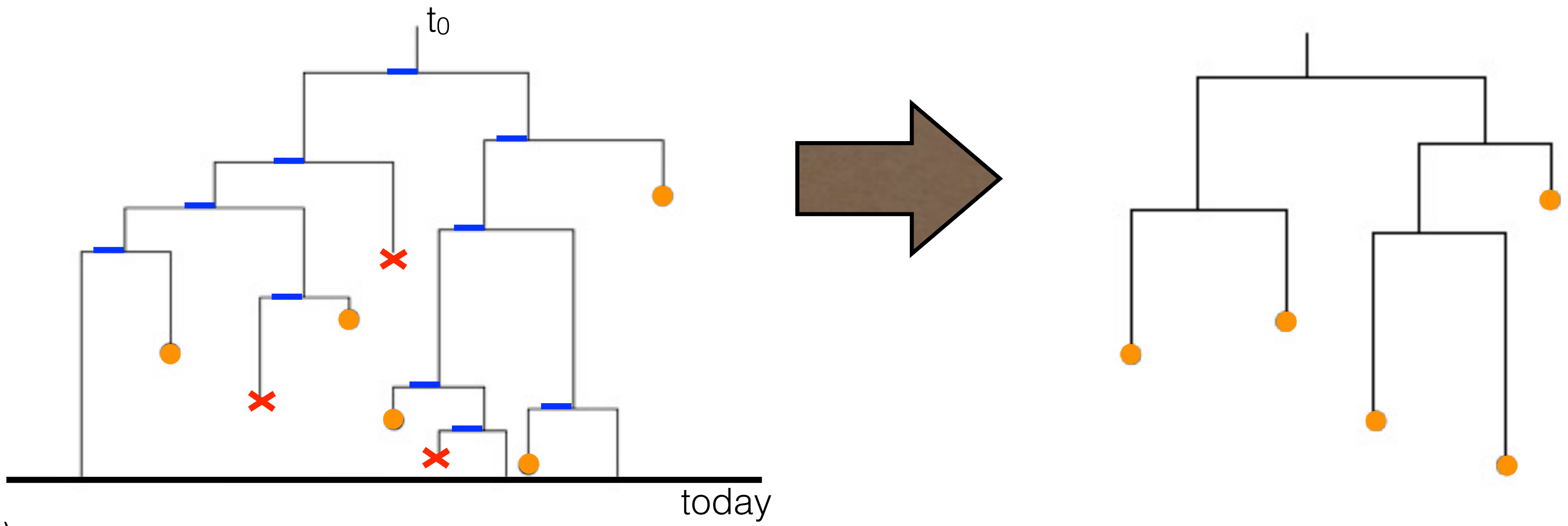


- transmission rate λ
- "becoming-non-infectious" rate δ
- sampling probability p

Parameters may depend on:

1) constant (epidemic outbreak)
2) time (environmental effect)
3) # of susceptibles (SIR model)
4) type of infected individual

Phylo-genetic trees

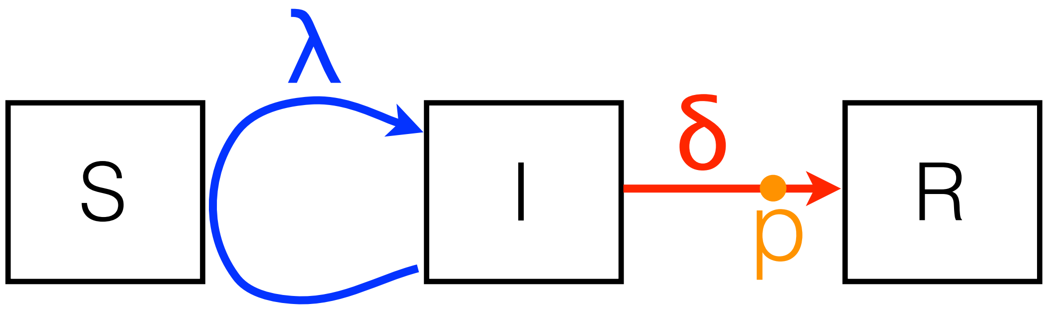


Stadler (JTB, 2010)

Part 2

Birth-death model for transmission

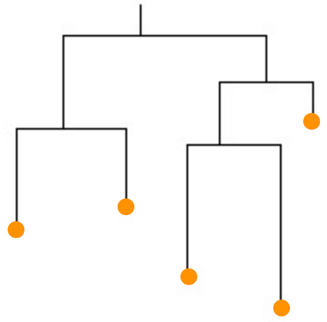
Birth-death model



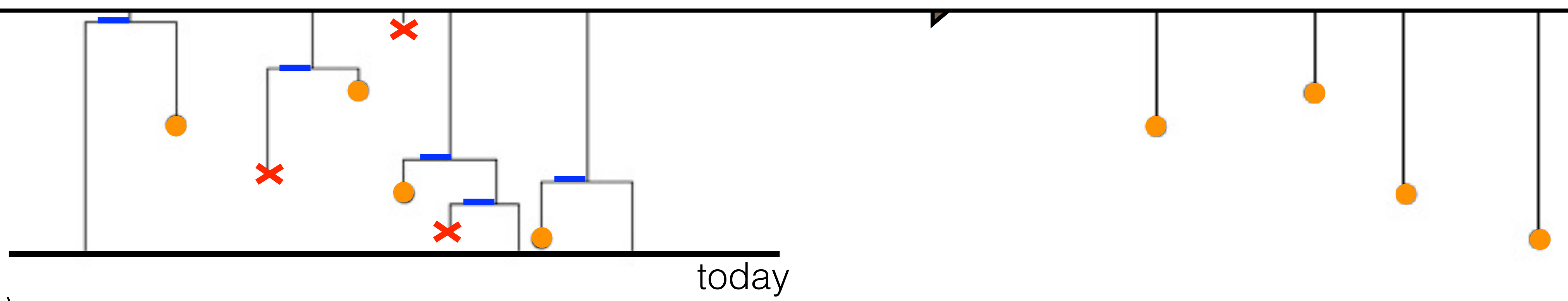
- transmission rate λ
- “becoming-non-infectious” rate δ
- sampling probability p

Parameters may depend on:

1) constant (epidemic outbreak)
2) time (environmental effect)
3) # of susceptibles (SIR model)
4) type of infected individual

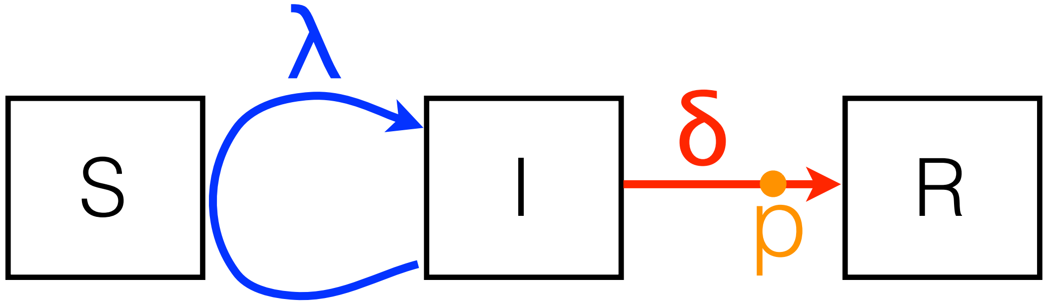
Lik () ?

genetic trees



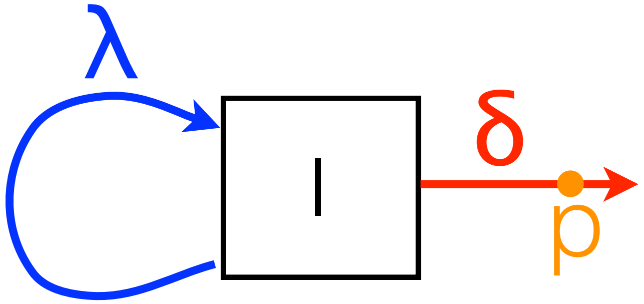
Stadler (JTB, 2010)

1) Epidemic outbreaks



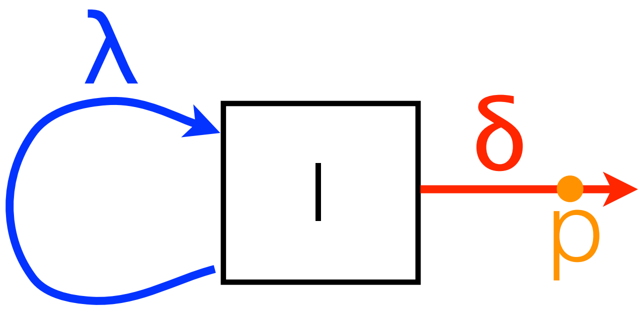
Stadler (JTB, 2010)

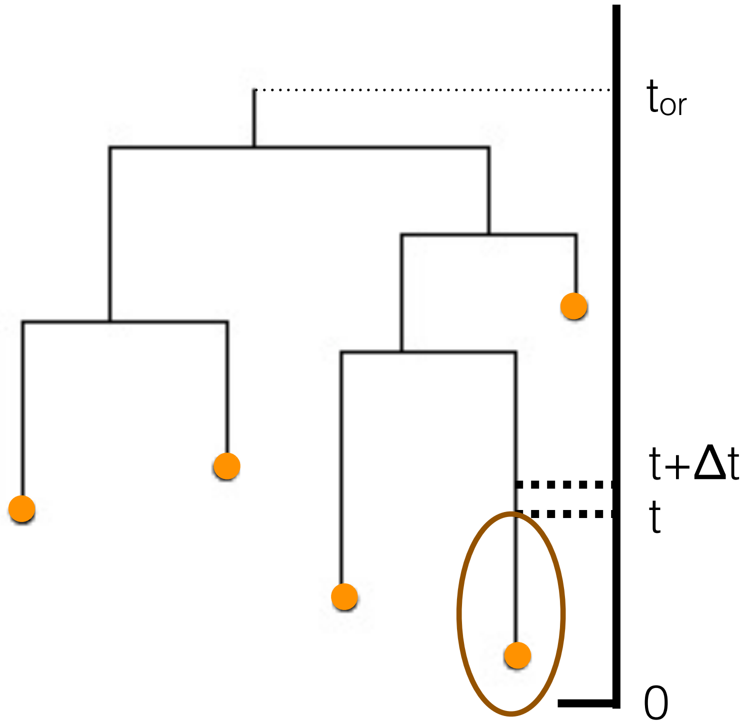
1) Epidemic outbreaks



Stadler (JTB, 2010)

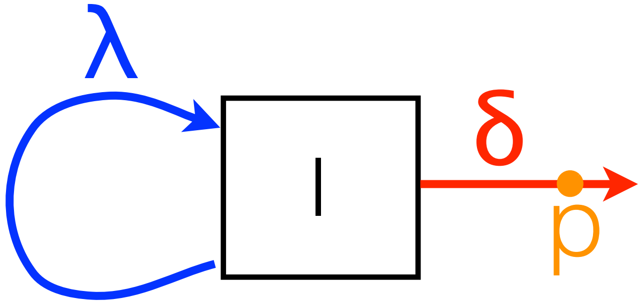
1) Epidemic outbreaks

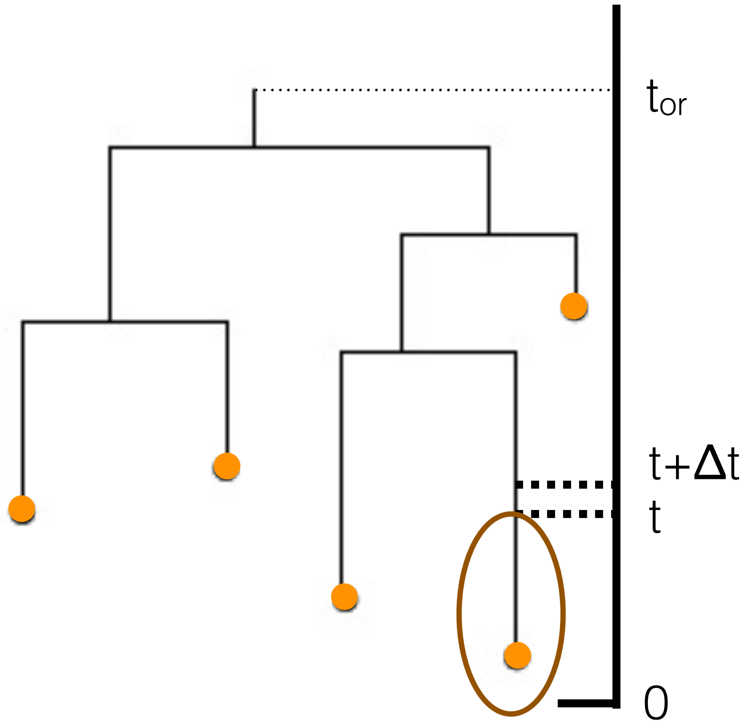




Stadler (JTB, 2010)

1) Epidemic outbreaks

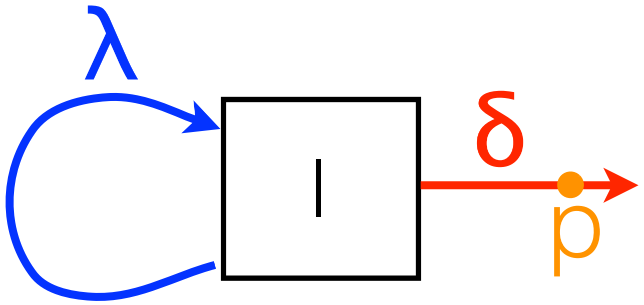


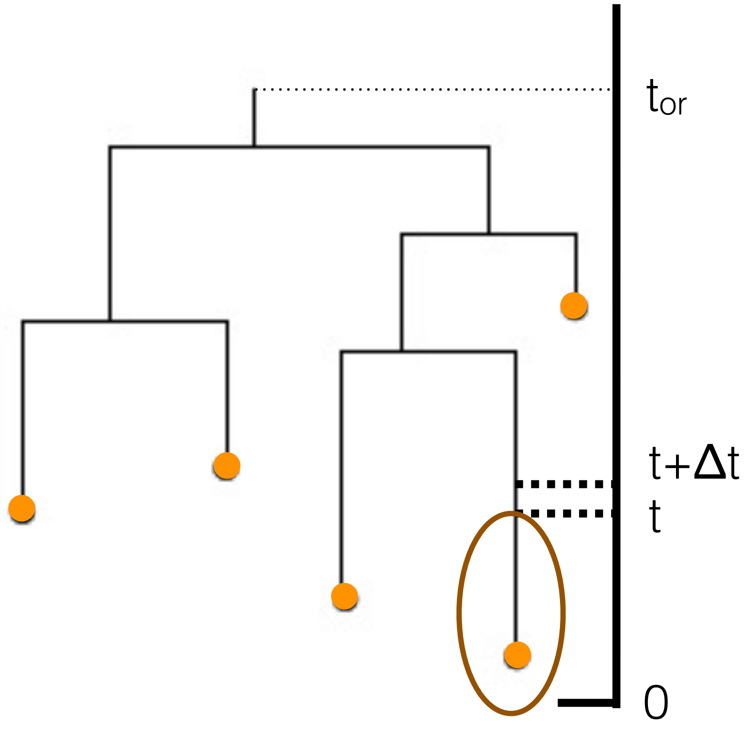


<u>p(t):</u> probability density that a given individual at time t produces an descending tree as observed.

Stadler (JTB, 2010)

1) Epidemic outbreaks



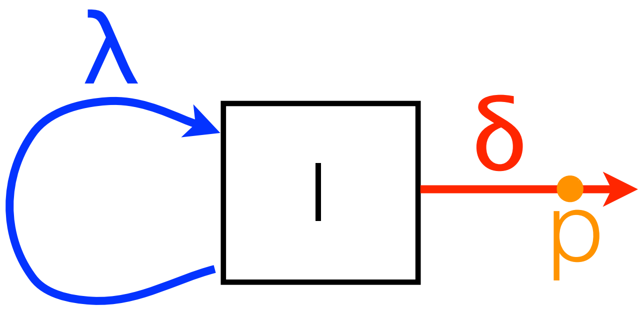


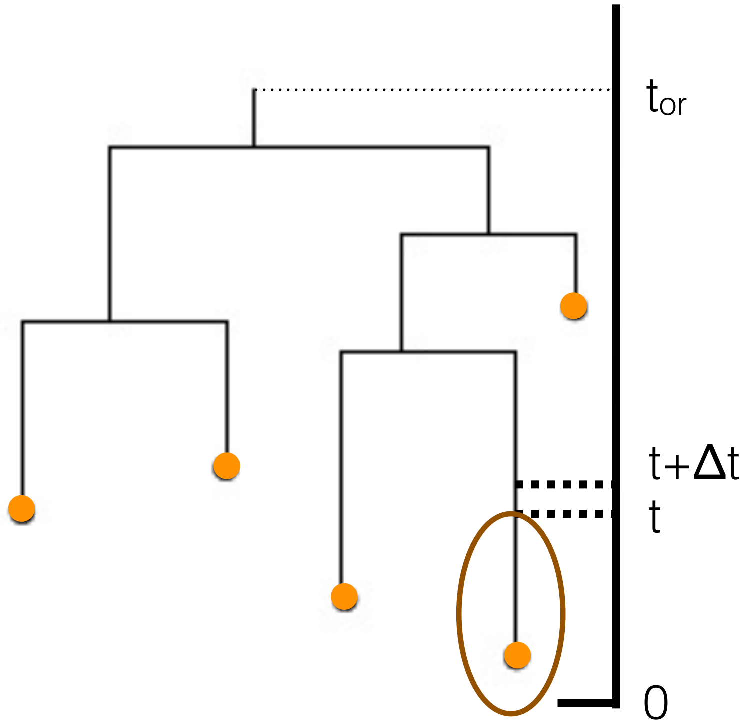
p(t): probability density that a given individual at time t produces an descending tree as observed.

p(t_{or}) = Lik(tree): probability density of sampled tree with origin at time t_{or}.

Stadler (JTB, 2010)

1) Epidemic outbreaks





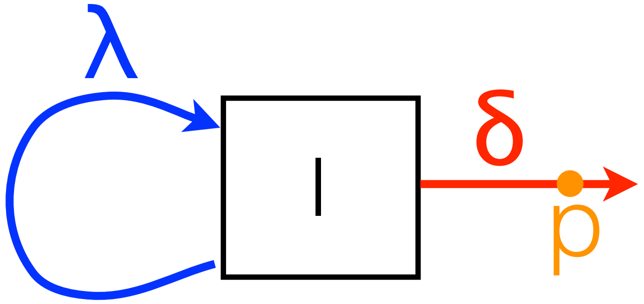
<u>p(t):</u> probability density that a given individual at time t produces an descending tree as observed.

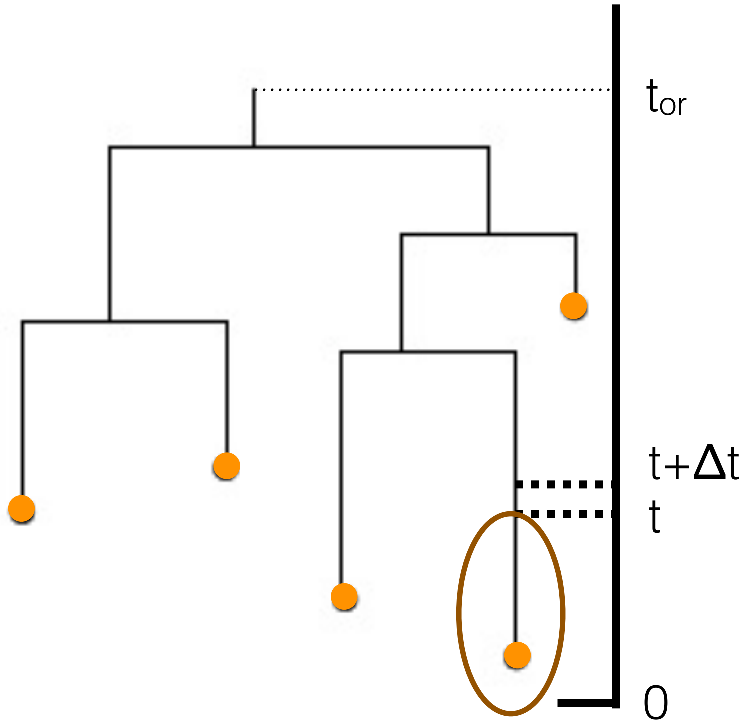
<u>$p(t_{or})$ = Lik(tree):</u> probability density of sampled tree with origin at time t_{or}.

<u>$p_0(t)$:</u> probability density that a given individual at time t has no sampled descendants.

Stadler (JTB, 2010)

1) Epidemic outbreaks





p(t): probability density that a given individual at time t produces an descending tree as observed.

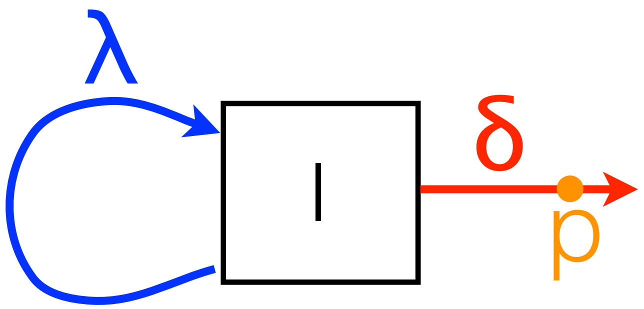
p(t_{or}) = Lik(tree): probability density of sampled tree with origin at time t_{or}.

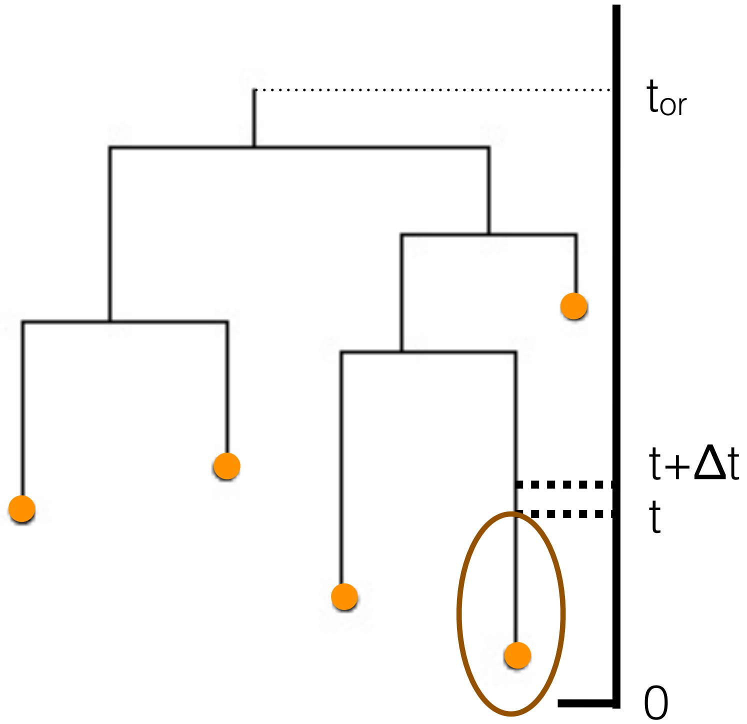
p_0(t): probability density that a given individual at time t has no sampled descendants.

$$p(t+\Delta t) = (1-(\lambda+\delta)\Delta t - O(\Delta t^2))p(t) + \lambda \Delta t 2 p_0(t) p(t) + O(\Delta t^2)$$

Stadler (JTB, 2010)

1) Epidemic outbreaks





p(t): probability density that a given individual at time t produces an descending tree as observed.

p(t_{or}) = Lik(tree): probability density of sampled tree with origin at time t_{or}.

p_0(t): probability density that a given individual at time t has no sampled descendants.

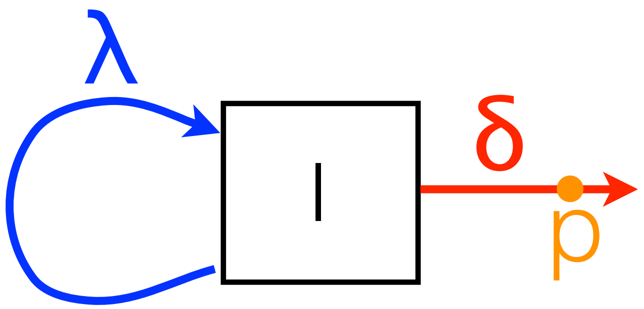
$$p(t+\Delta t) = (1-(\lambda+\delta)\Delta t - O(\Delta t^2))p(t) + \lambda \Delta t 2 p_0(t) p(t) + O(\Delta t^2)$$

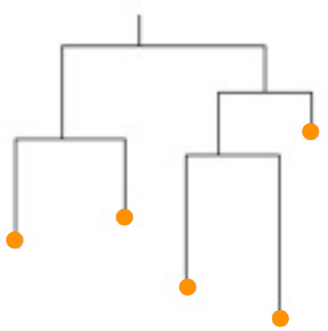
Δt → 0

$$\frac{d}{dt}p(t) = -(\lambda+\delta)p(t) + 2\lambda p_0(t)p(t)$$

Stadler (JTB, 2010)

Tree likelihood



Lik ()

$$= \lambda^{m-1} \prod_{i=0}^{m-1} \frac{1}{q(x_i)} \prod_{i=1}^{m} p\delta q(y_i)$$

where

$$q(t) = 2(1 - c_2^2) + e^{-c_1 t}(1 - c_2)^2 + e^{c_1 t}(1 + c_2)^2$$

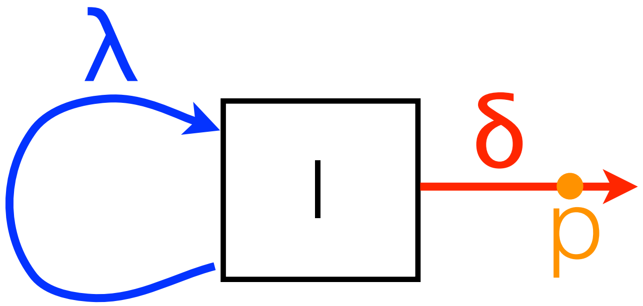
$$\begin{aligned} c_1 &= |\sqrt{(\lambda - \delta)^2) + 4p\lambda\delta}| \\ c_2 &= -\frac{\lambda - \delta}{c_1} \end{aligned}$$

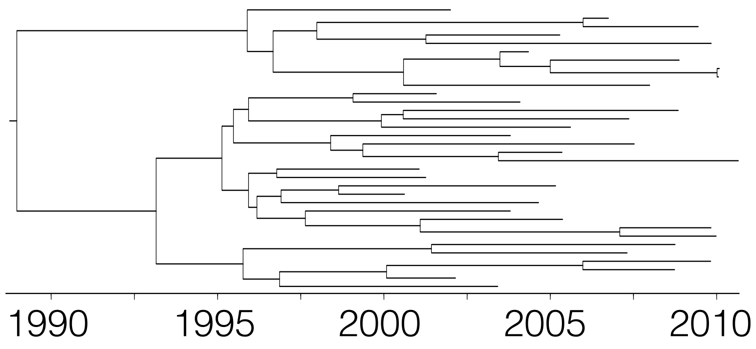
and x_i branching times, y_i sampling times.

Stadler (JTB, 2010)

Simulation study: Comparing exact method to coalescent

Simulating epidemic outbreak

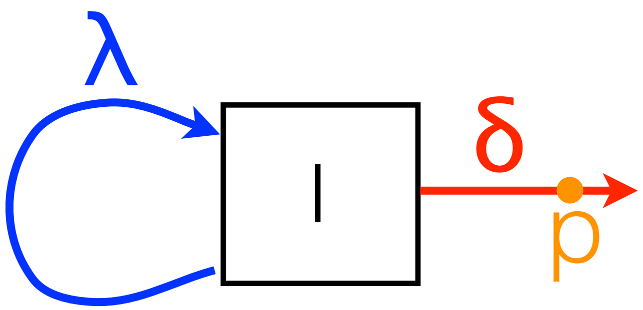


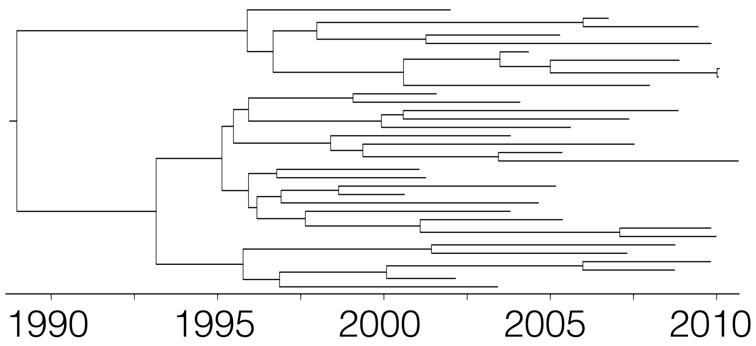


Stadler et al. (Mol. Biol. Evol., 2012)

Simulation study:
Comparing exact method to coalescent

Simulating epidemic outbreak





Exact method

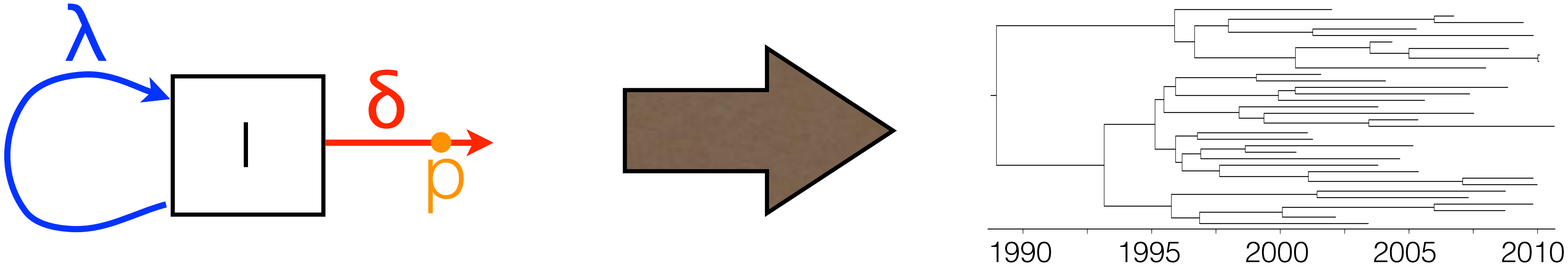
λ-δ is population size growth rate

- 95% *confidence* interval contains true value in **97%** of the simulated trees

Stadler et al. (Mol. Biol. Evol., 2012)

Simulation study:
Comparing exact method to coalescent

Simulating epidemic outbreak



Exact method

λ-δ is population size growth rate

- 95% *confidence* interval contains true value in **97%** of the simulated trees

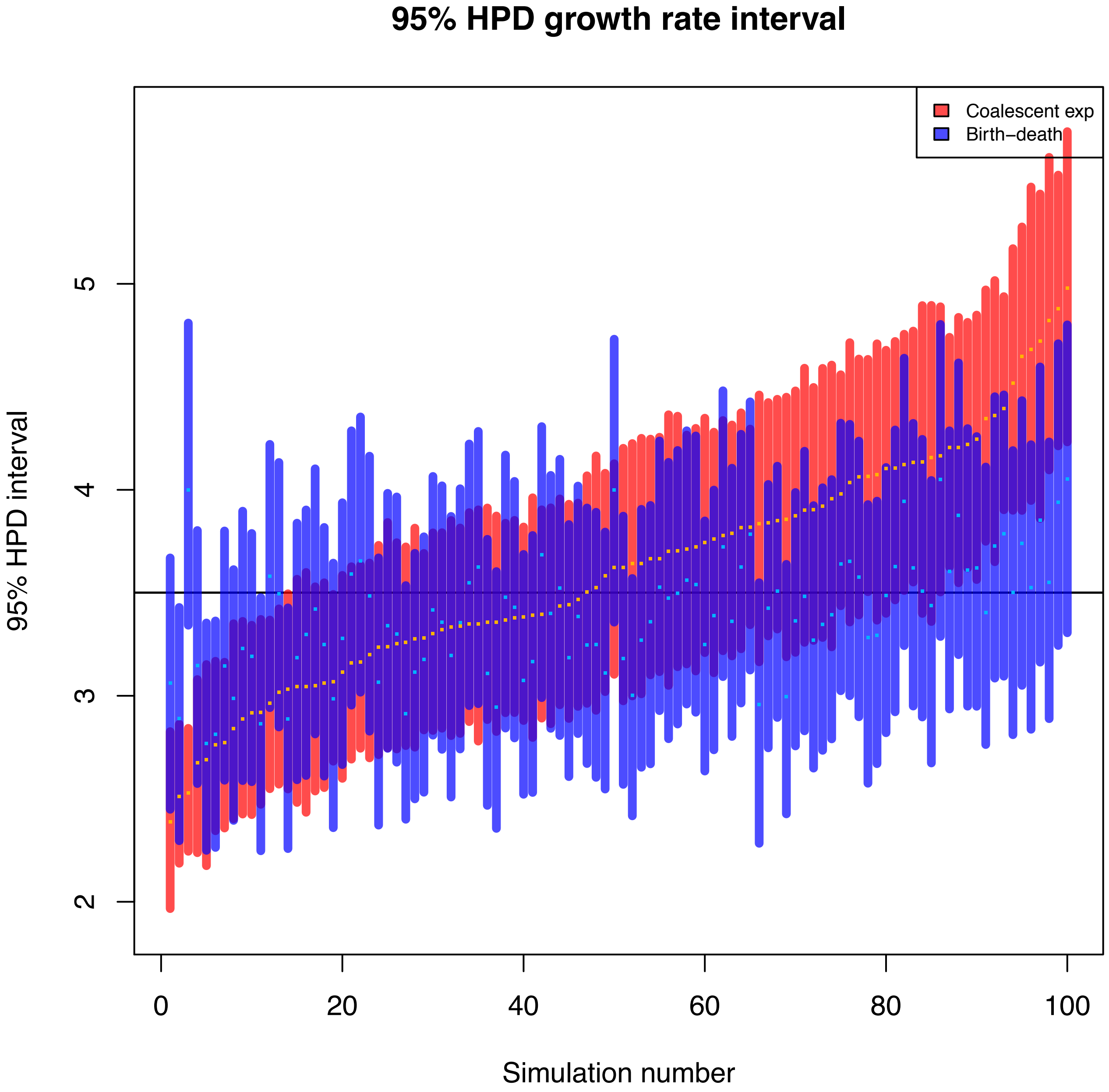
Coalescent approximation

Parameterizes deterministic population size growth

- *95% confidence* interval contains true value in **55%** of the simulated trees

Stadler et al. (Mol. Biol. Evol., 2012)

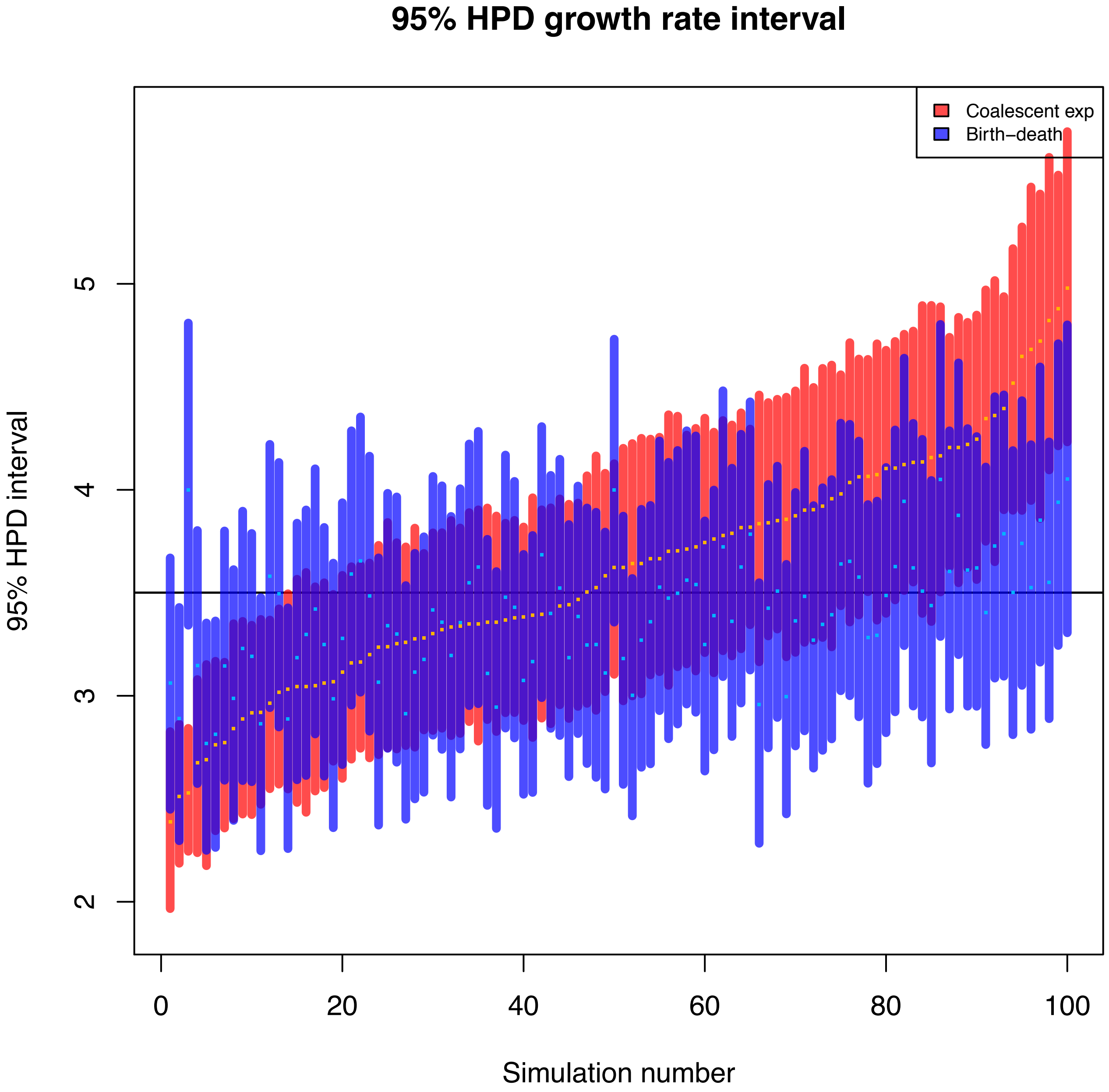
Simulating birth-death trees



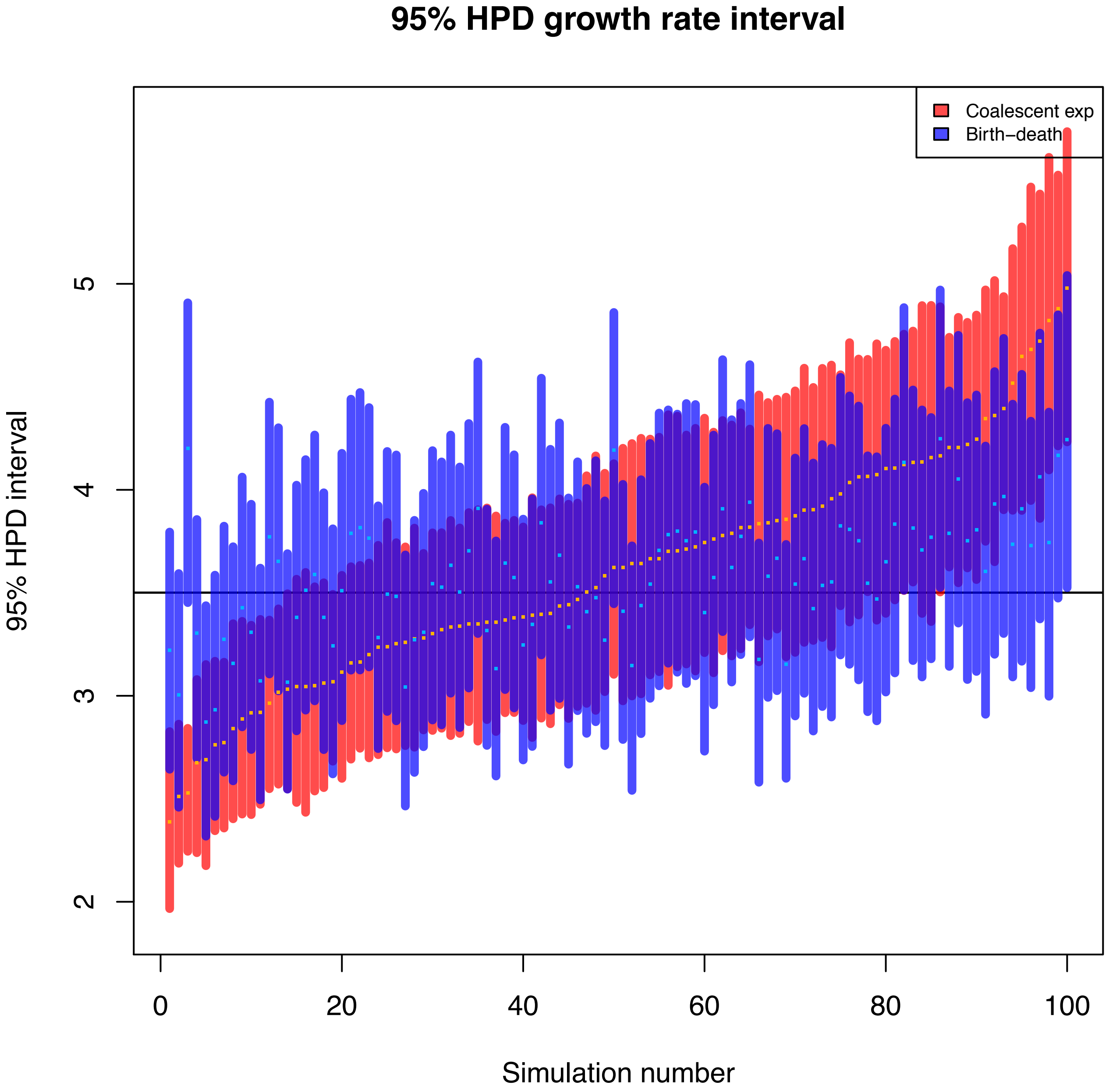
p=0.5

Boskova et al., in preparation

Simulating birth-death trees



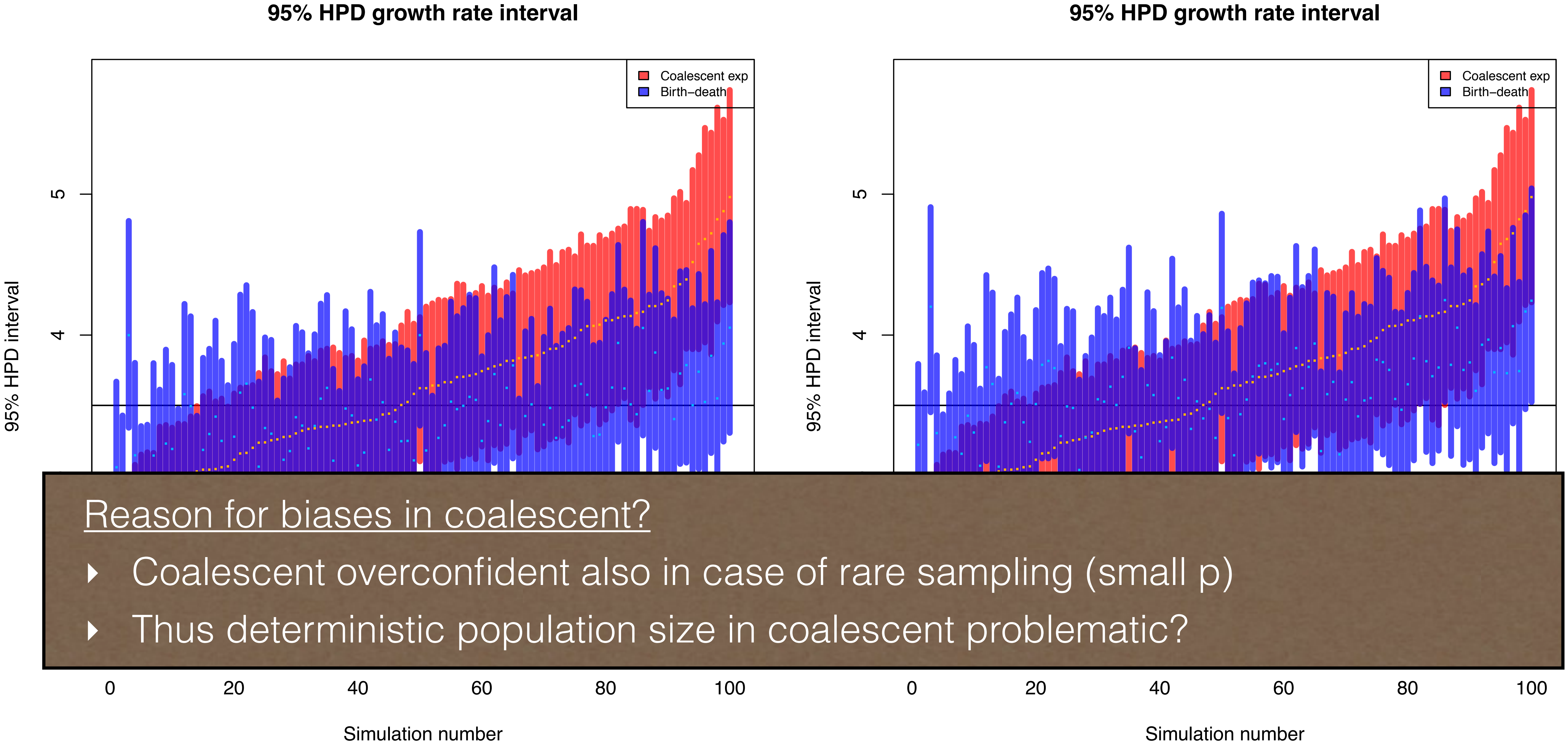
p=0.5



p=0.01

Boskova et al., in preparation

Simulating birth-death trees



Reason for biases in coalescent?

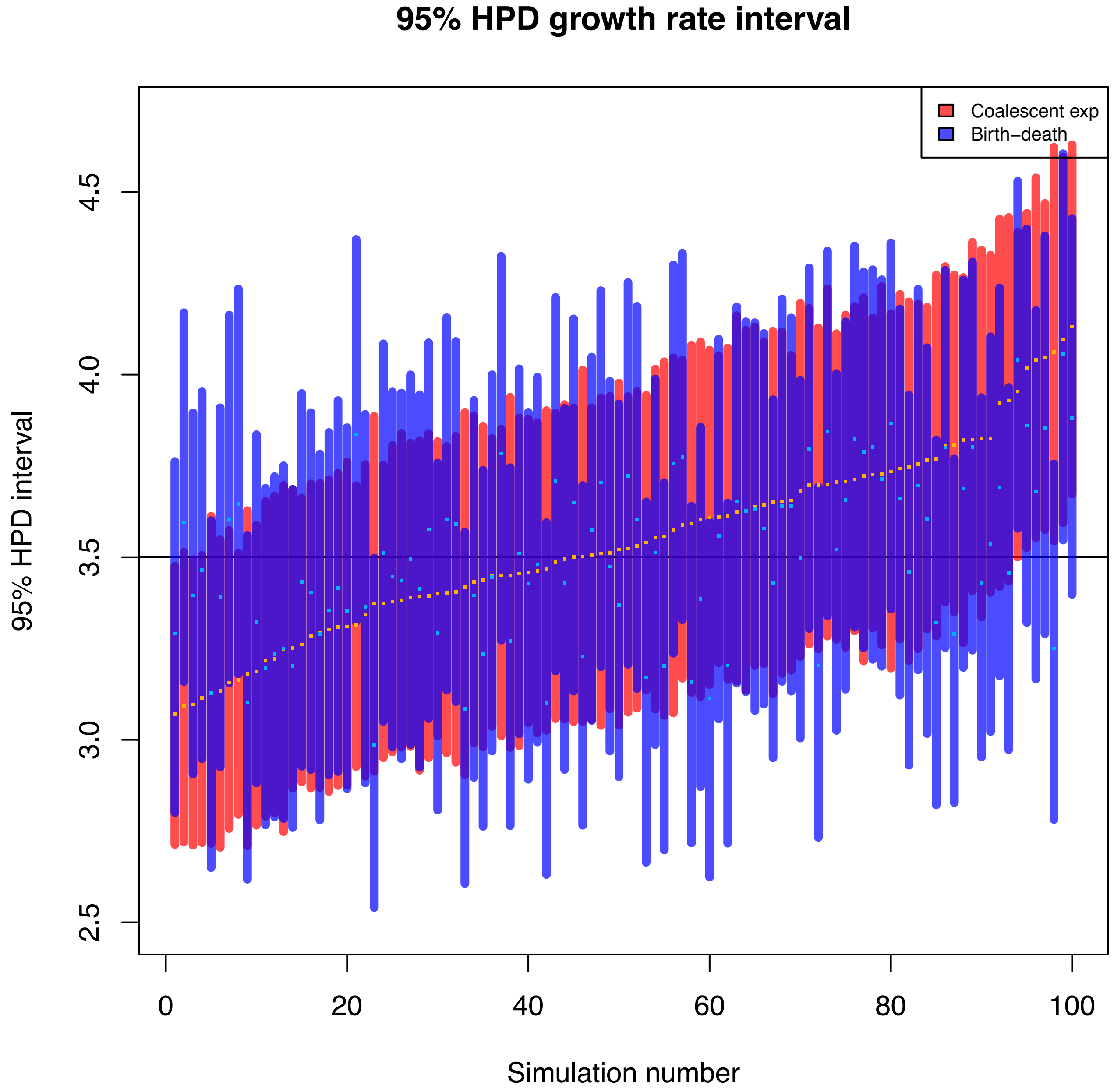
- Coalescent overconfident also in case of rare sampling (small p)
- Thus deterministic population size in coalescent problematic?

p=0.5

p=0.01

Boskova et al., in preparation

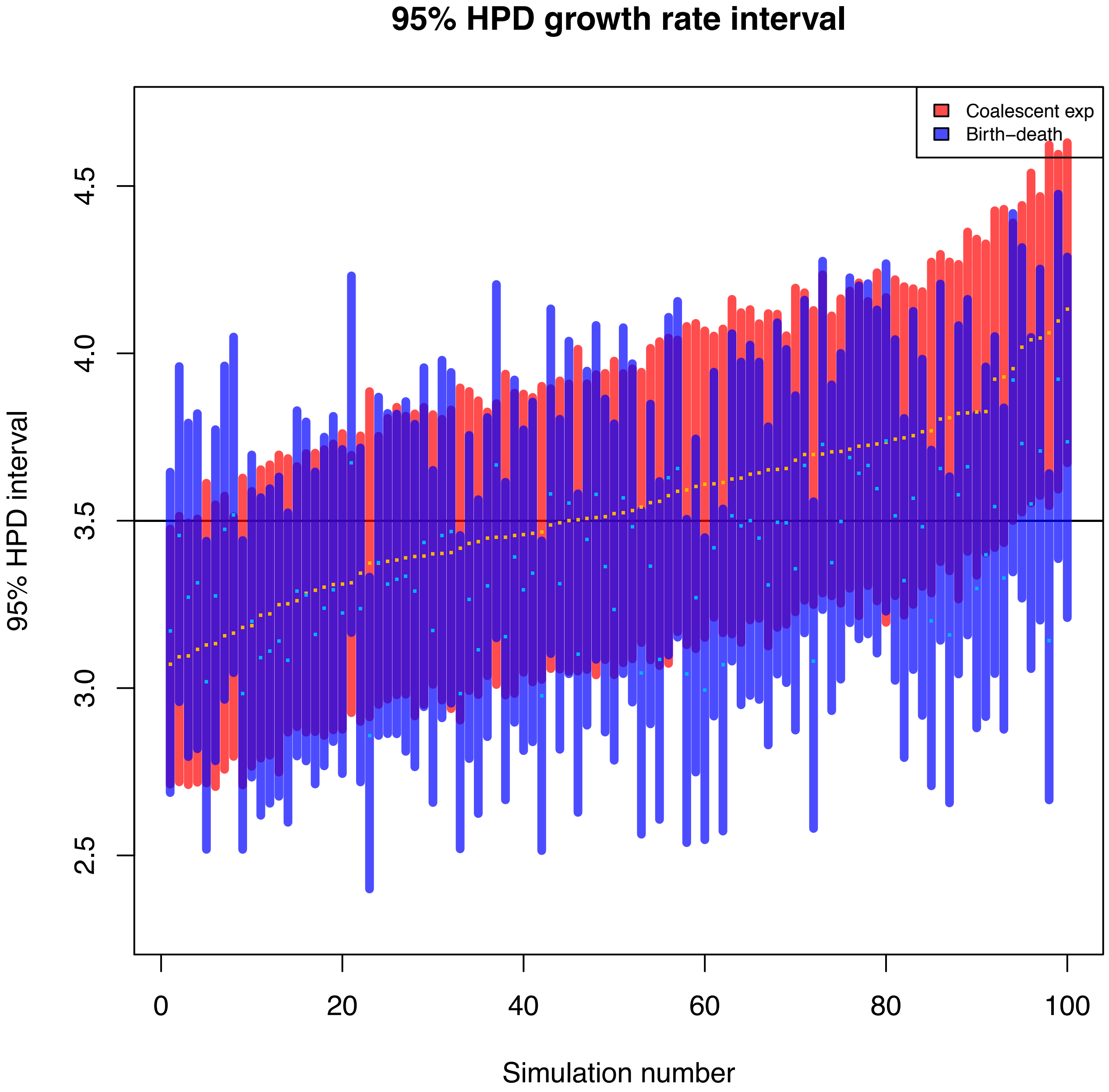
Simulating coalescent trees



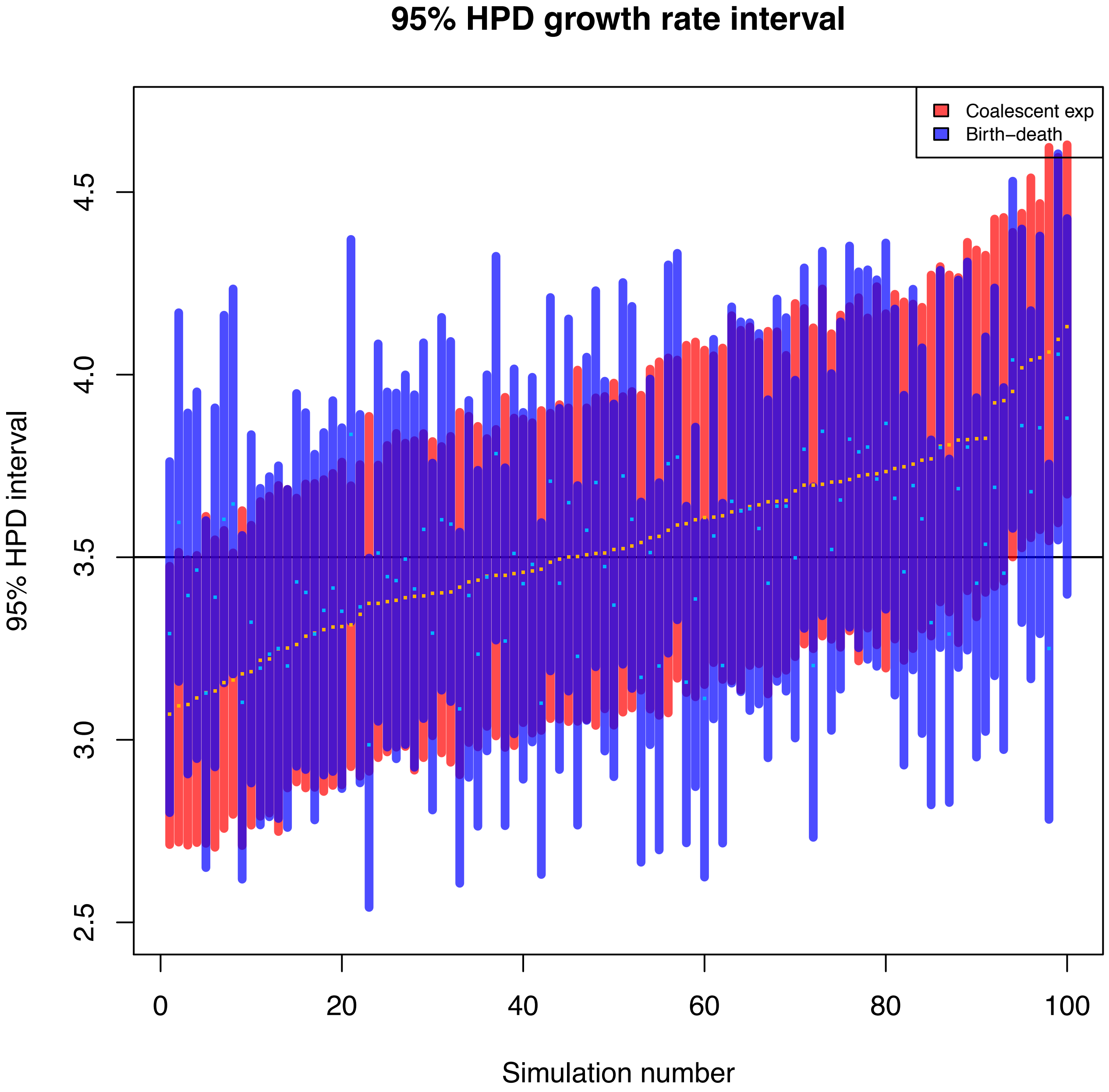
p=0.01

Boskova et al., in preparation

Simulating coalescent trees



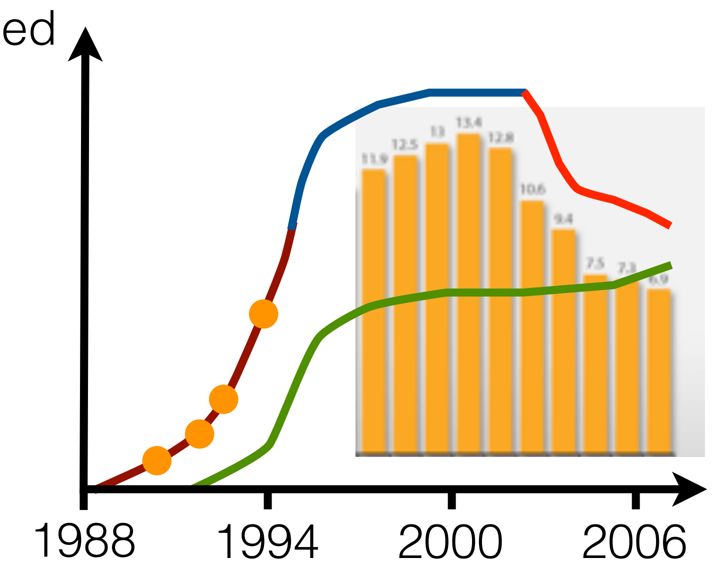
p=0.5



p=0.01

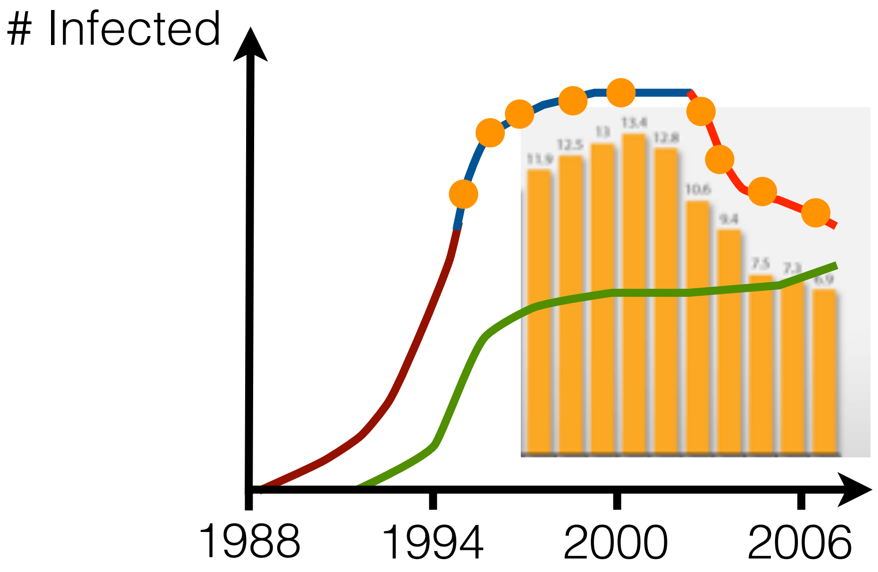
Boskova et al., in preparation

Phylogenetic methods...



1. ...for analyzing epidemic outbreaks
    - *C: Drummond et al. (Genetics, 2002)*
    - *BD: Stadler et al. (MBE, 2012)*
2. ...for identifying epidemiological changes over time
    - *C: Drummond et al. (MBE, 2005)*
    - *BD: Stadler, Kühnert et al. (PNAS, 2013)*
3. ...accounting for SIR model dynamics
    - *C: Volz et al. (Genetics, 2009)*
    - *BD: Kühnert et al., Leventhal et al. (both submitted)*
4. ...accounting for host population structure
    - *C: Volz (Genetics, 2012)*
    - *BD: Stadler & Bonhoeffer (Phil. Trans. Roy. Soc. B, 2013)*

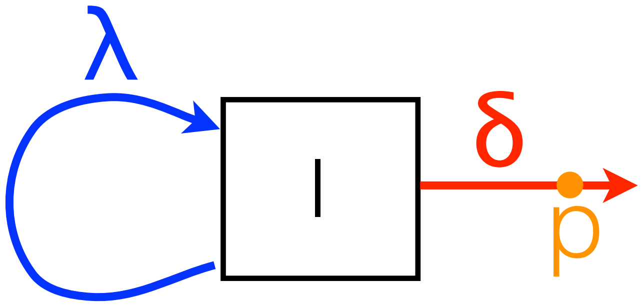
Phylogenetic methods...



1. ...for analyzing epidemic outbreaks
    - *C: Drummond et al. (Genetics, 2002)*
    - *BD: Stadler et al. (MBE, 2012)*
2. ...for identifying epidemiological changes over time
    - *C: Drummond et al. (MBE, 2005)*
    - *BD: Stadler, Kühnert et al. (PNAS, 2013)*
3. ...accounting for SIR model dynamics
    - *C: Volz et al. (Genetics, 2009)*
    - *BD: Kühnert et al., Leventhal et al. (both submitted)*
4. ...accounting for host population structure
    - *C: Volz (Genetics, 2012)*
    - *BD: Stadler & Bonhoeffer (Phil. Trans. Roy. Soc. B, 2013)*

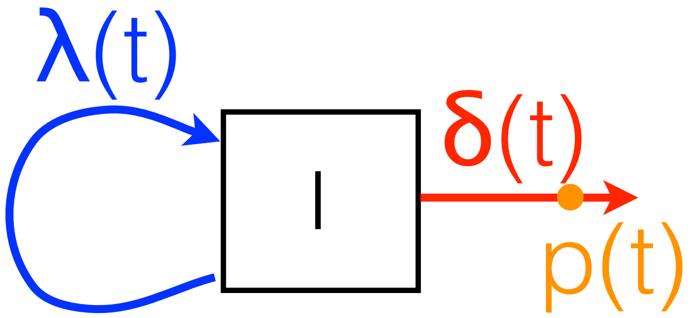
Birth-death-skyline plot

Epidemiological rates may change through time:



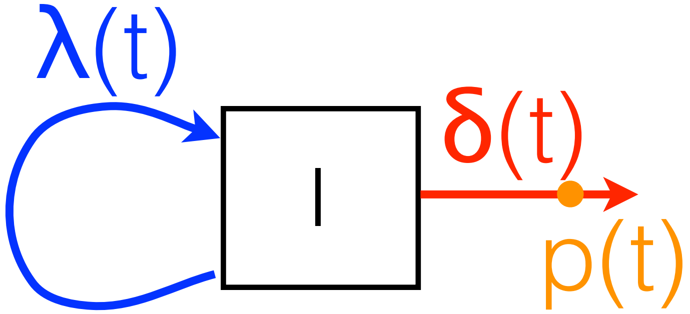
Birth-death-skyline plot

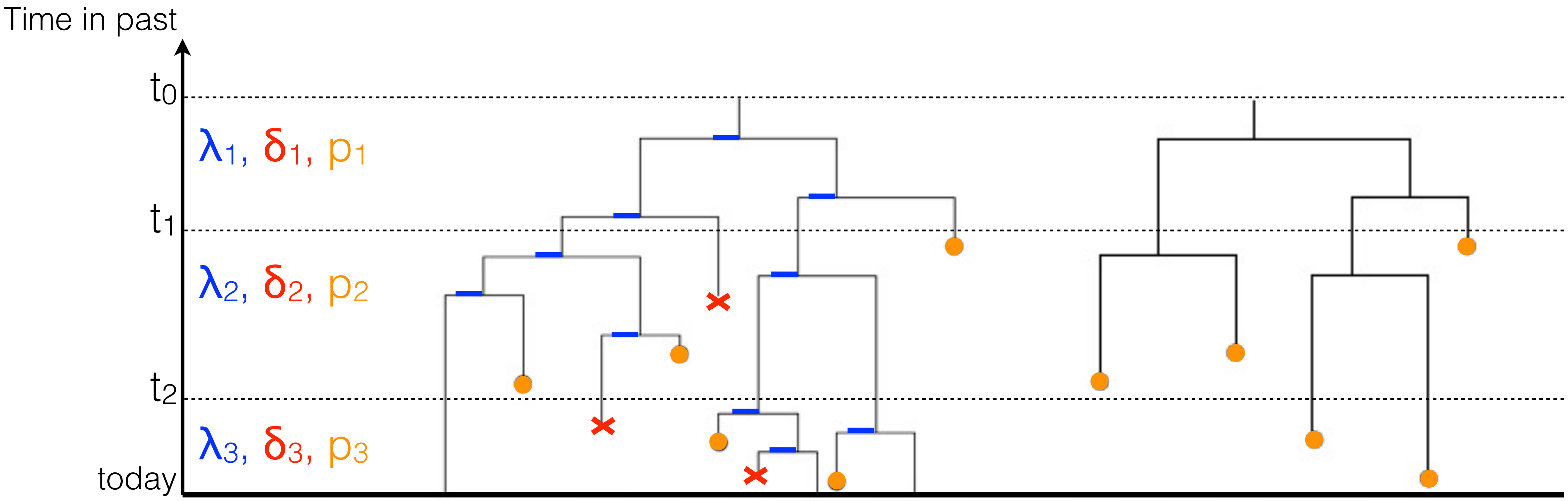
Epidemiological rates may change through time:



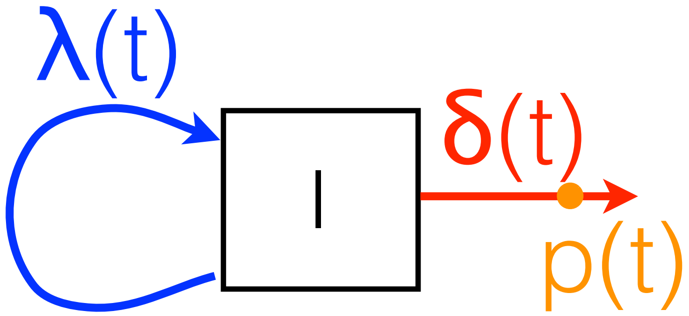
Birth-death-skyline plot

Epidemiological rates may change through time:





Hepatitis C virus in Egypt



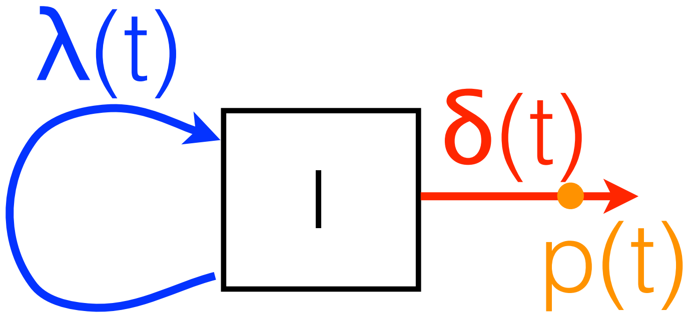
Effective reproductive number $\mathbf{R_e(t)}$

generalizes

Basic reproductive number $\mathbf{R_0 = R_e(0)}$

Stadler, Kühnert et al. (PNAS, 2013)

Hepatitis C virus in Egypt



Effective reproductive number **$R_e(t)$**

generalizes

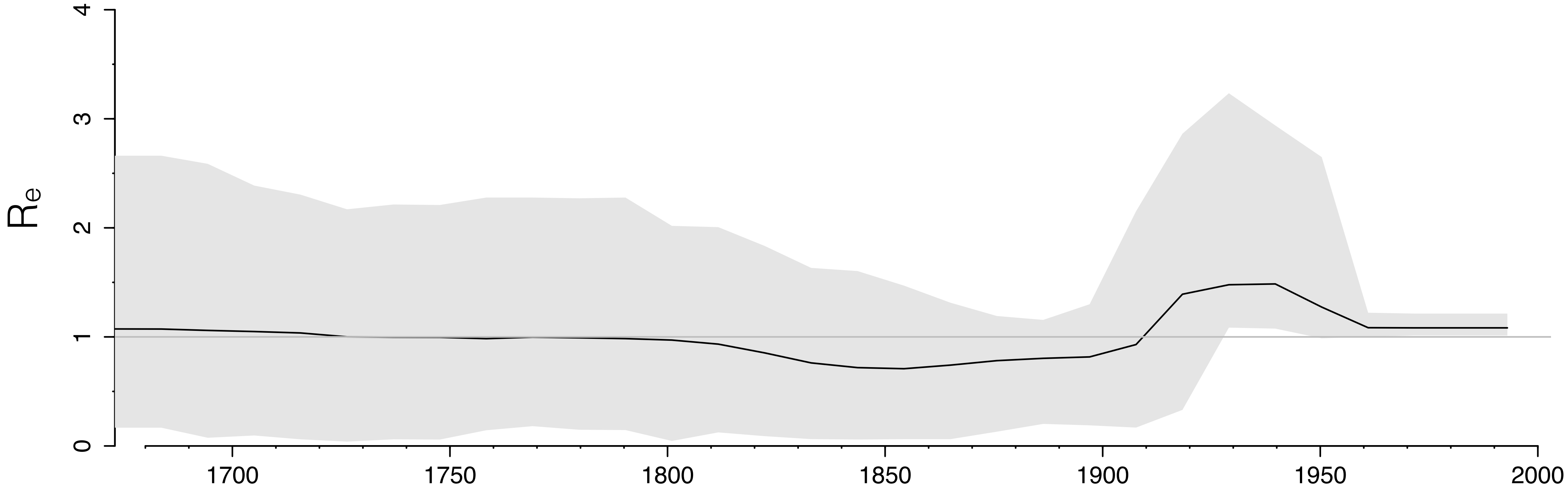
Basic reproductive number **$R_0 = R_e(0)$**

Analysis of 63 seq:

$R_e(t) = \lambda(t)/\delta(t)$

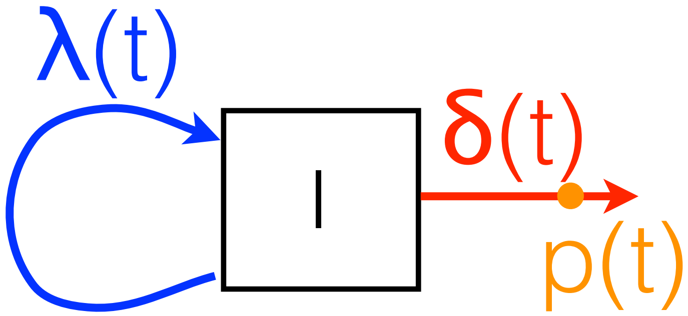
— median

▨ 95% HPD interval



Stadler, Kühnert et al. (PNAS, 2013)

Hepatitis C virus in Egypt



Effective reproductive number **$R_e(t)$**

generalizes

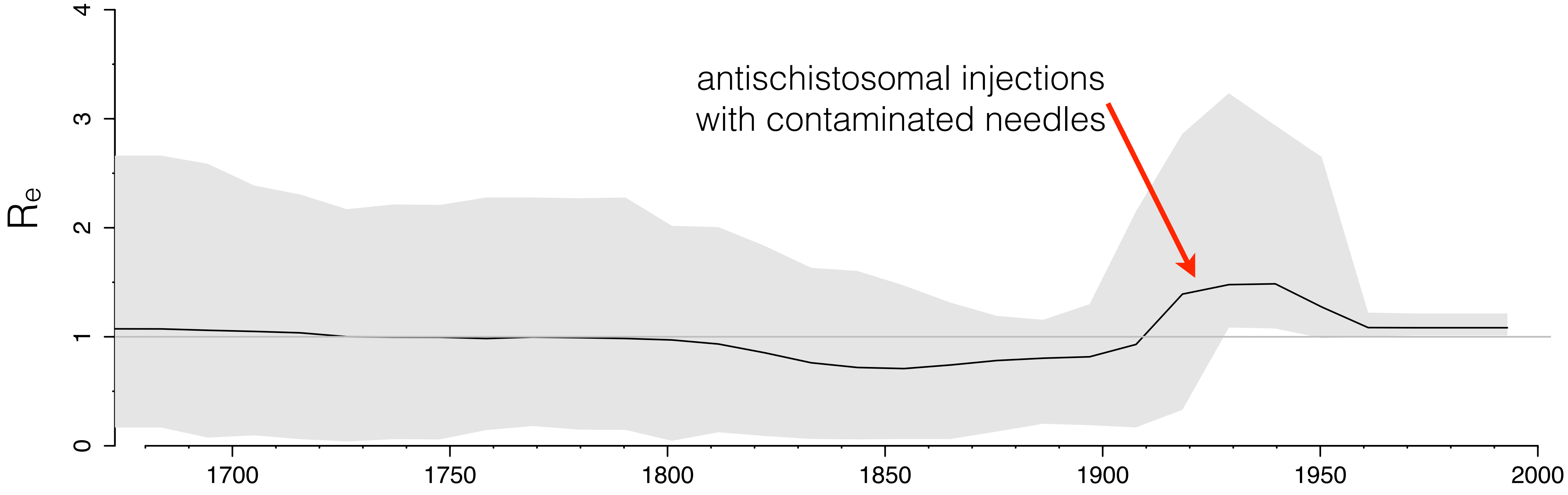
Basic reproductive number **$R_0 = R_e(0)$**

Analysis of 63 seq:

$R_e(t)$ = $\lambda(t)/\delta(t)$

— median

▨ 95% HPD interval

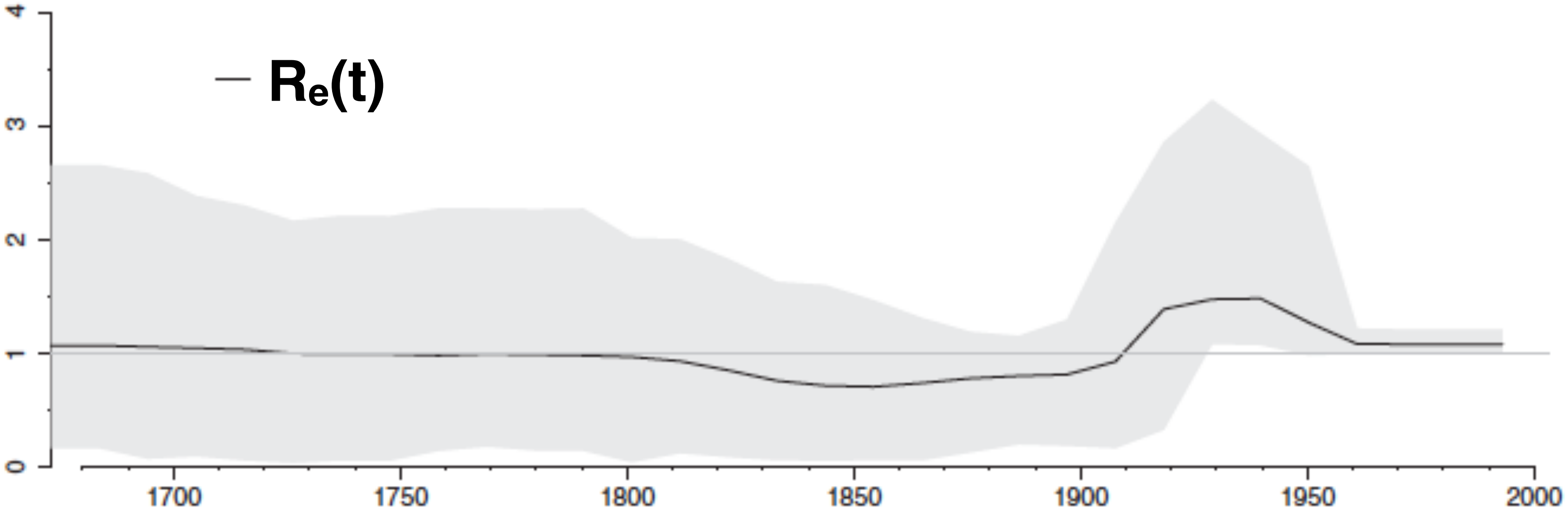


antischistosomal injections with contaminated needles

Stadler, Kühnert et al. (PNAS, 2013)

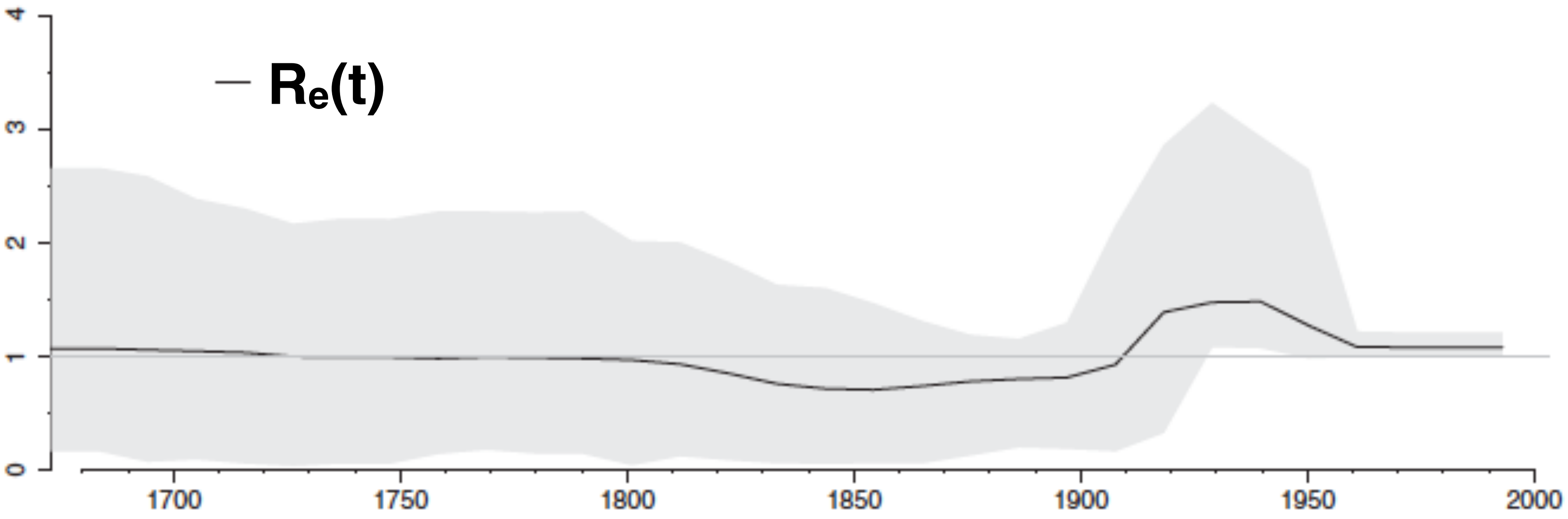
The two skyline models

Birth-death skyline plot: effective reproductive number

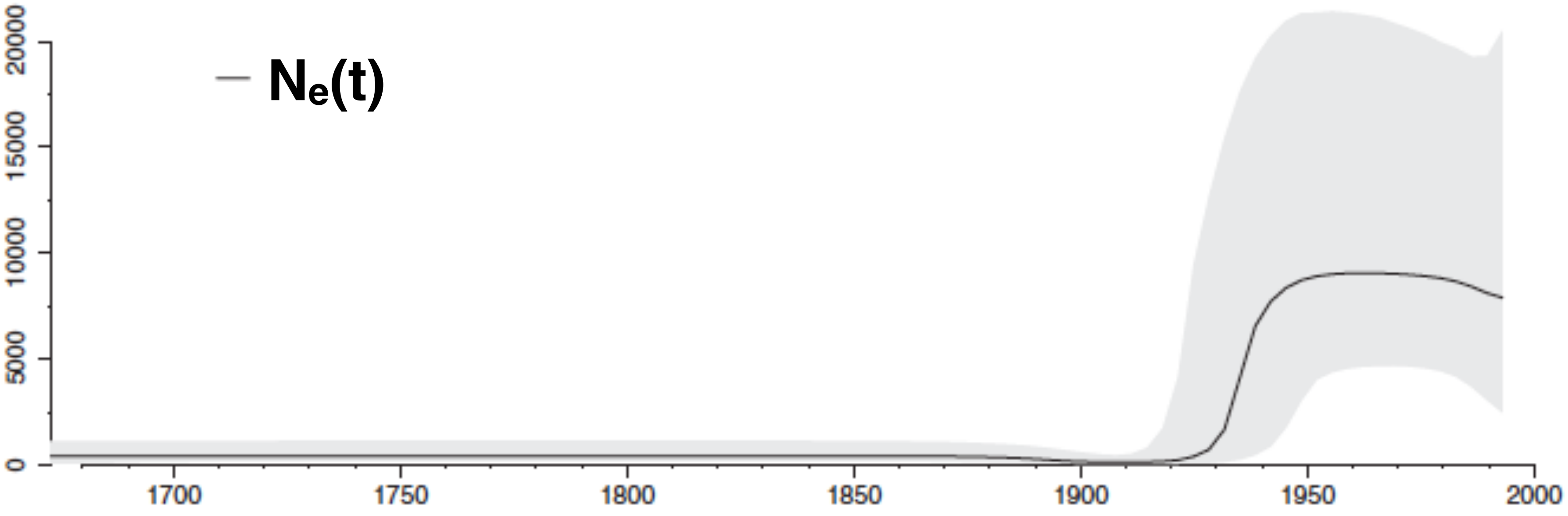


The two skyline models

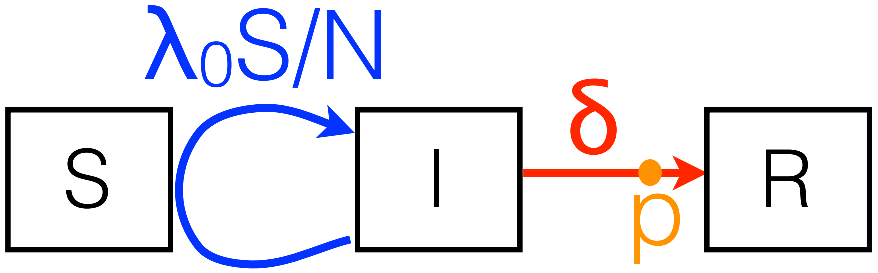
Birth-death skyline plot: effective reproductive number



Coalescent skyline plot: effective population size



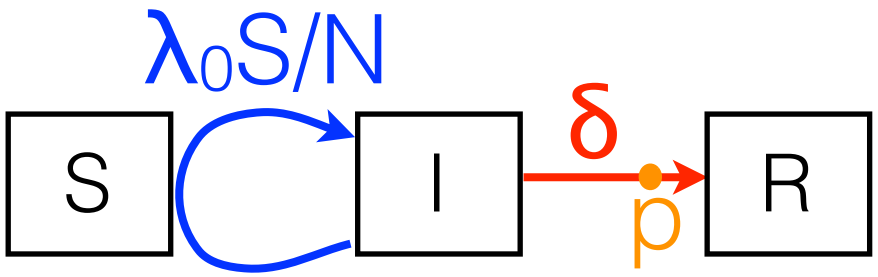
3) SIR models



SIR simulation (trees with 200 tips sampled):
population size: N=5000, start with one infected, N=S+I+R
$\lambda(S) = \lambda_0 S/N$
$R_0 = \lambda_0/\delta = 4$
host recovery rate: $\delta = 0.5$
sampling probability: $p = 0.05$

Kühnert et al. (submitted)

3) SIR models

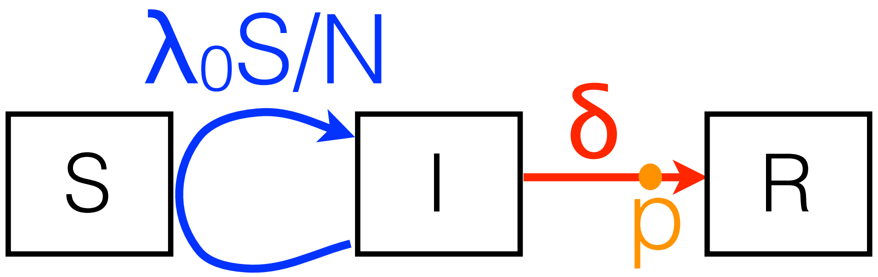


SIR simulation (trees with 200 tips sampled):
population size: $N=5000$, start with one infected, $N=S+I+R$
$\lambda(S) = \lambda_0 S/N$
$R_0 = \lambda_0/\delta = 4$
host recovery rate: $\delta = 0.5$
sampling probability: $p = 0.05$

Inference: 10 intervals for λ, δ and p constant across interval
meaning $\lambda(S)=\lambda_0 S/N$ is approximated by $\lambda(t)$

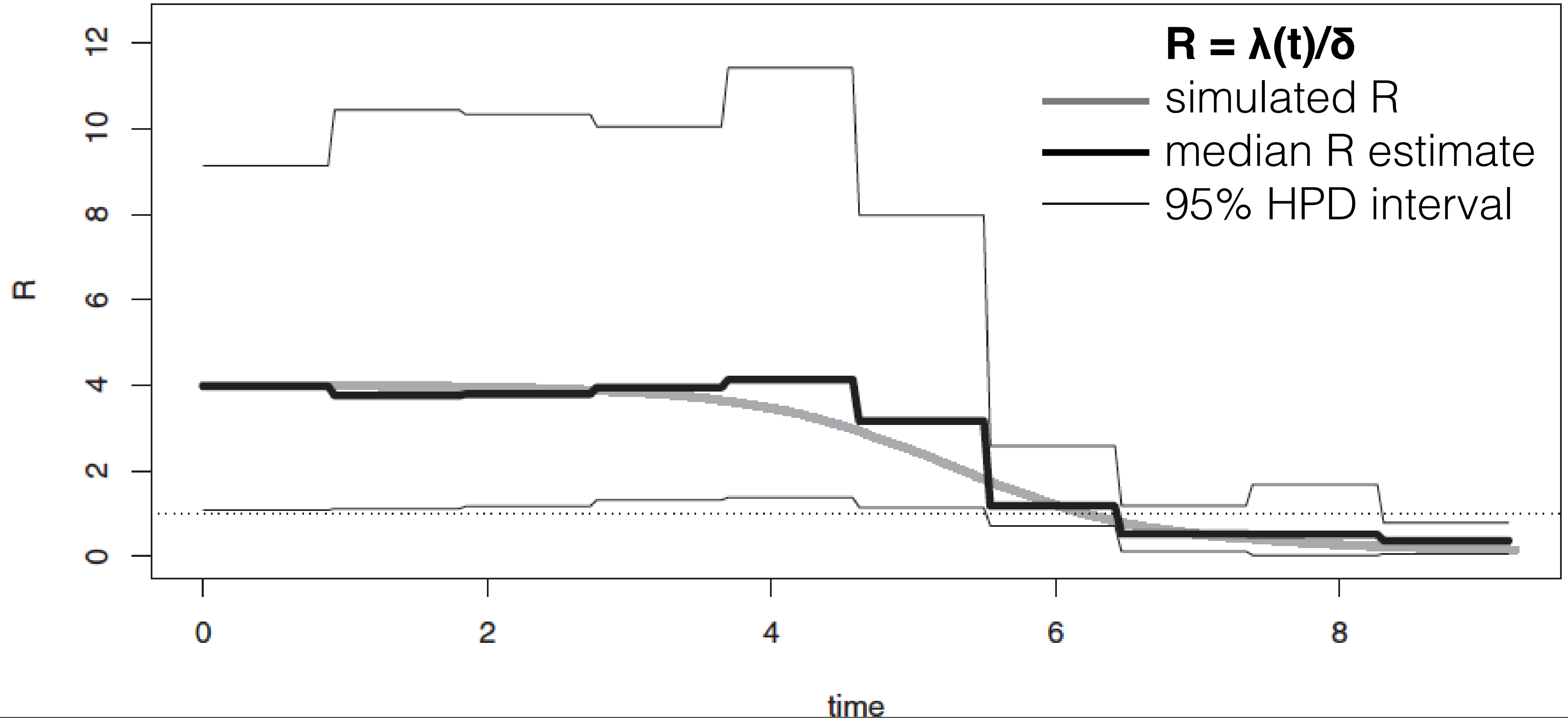
Kühnert et al. (submitted)

3) SIR models



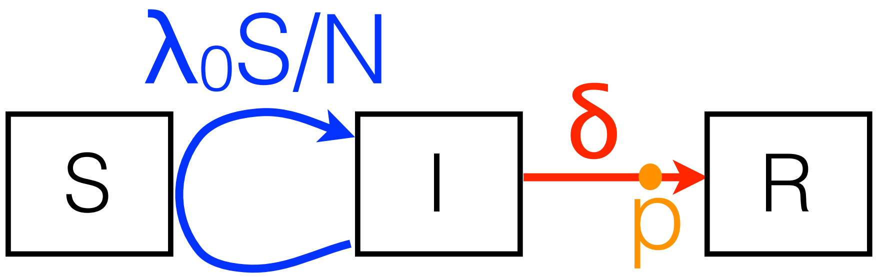
SIR simulation (trees with 200 tips sampled):
population size: N=5000, start with one infected, N=S+I+R
$\lambda(S) = \lambda_0 S/N$
$R_0 = \lambda_0/\delta = 4$
host recovery rate: $\delta = 0.5$
sampling probability: $p = 0.05$

Inference: 10 intervals for λ, δ and p constant across interval
meaning $\lambda(S)=\lambda_0 S/N$ is approximated by $\lambda(t)$



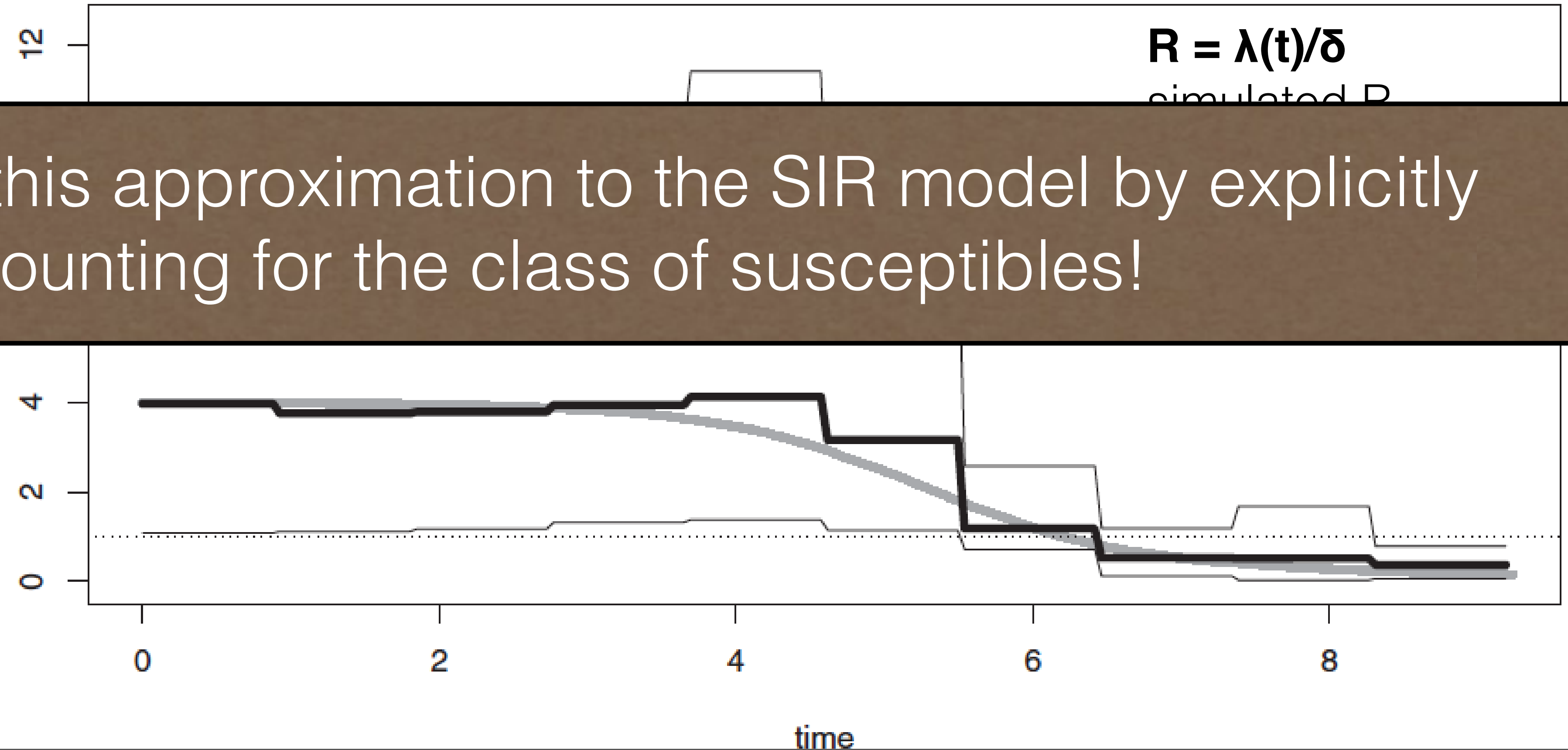
Kühnert et al. (submitted)

3) SIR models



SIR simulation (trees with 200 tips sampled):
population size: N=5000, start with one infected, N=S+I+R
$\lambda(S) = \lambda_0 S/N$
$R_0 = \lambda_0/\delta = 4$
host recovery rate: $\delta = 0.5$
sampling probability: $p = 0.05$

Inference: 10 intervals for λ, δ and p constant across interval
meaning $\lambda(S)=\lambda_0 S/N$ is approximated by $\lambda(t)$

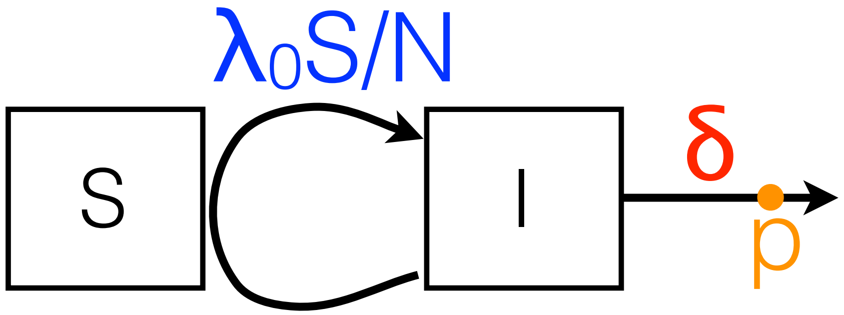


We improved this approximation to the SIR model by explicitly accounting for the class of susceptibles!

Kühnert et al. (submitted)

SI dynamics - exact method

Para-
meters

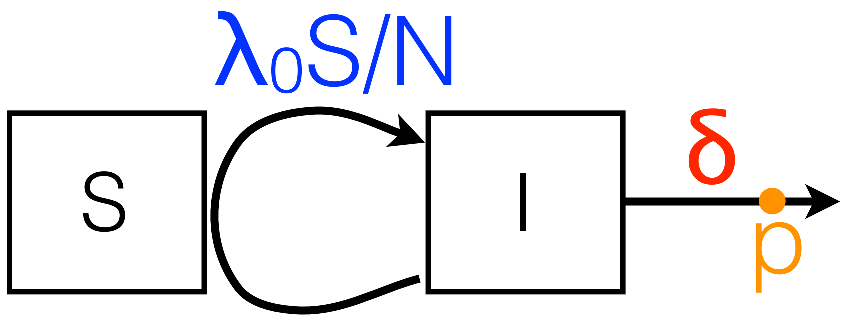


N=S+I (constant)

Leventhal et al. (submitted)

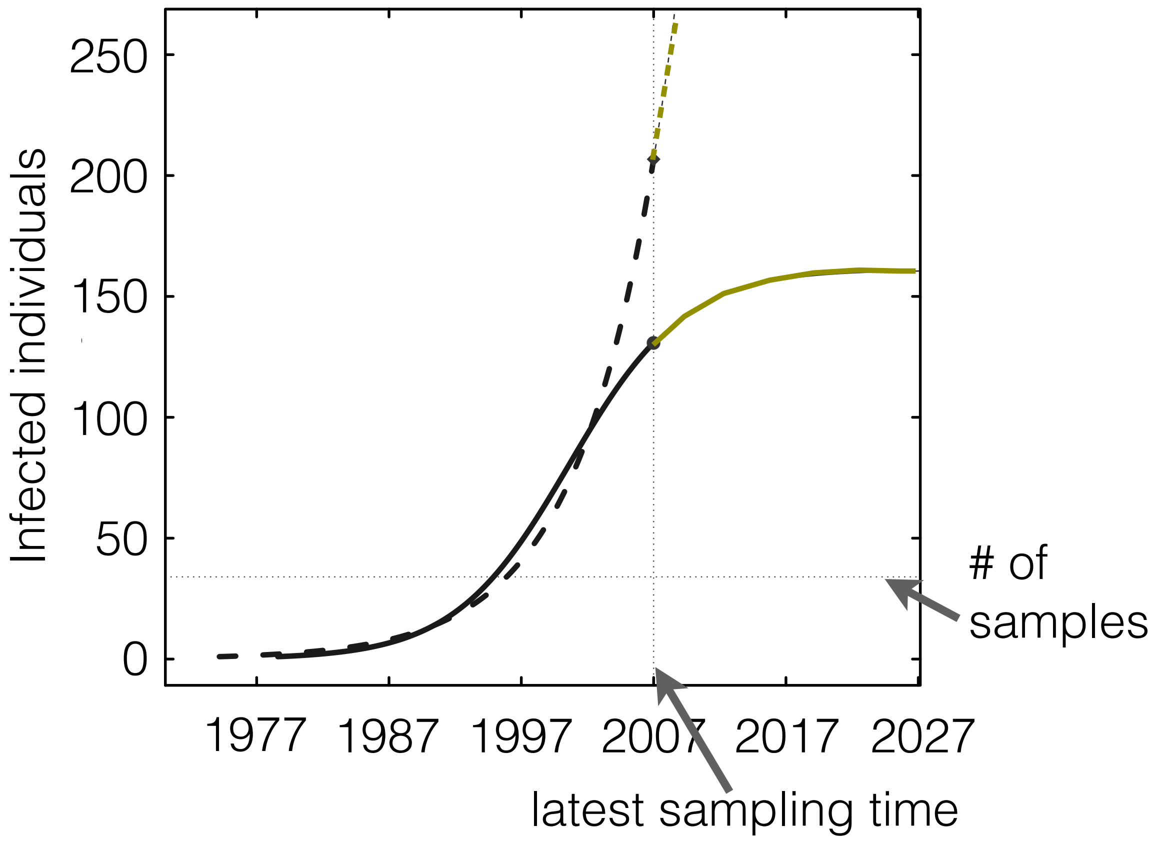
SI dynamics - exact method

Para-meters



N=S+I (constant)

HIV transmission cluster in Switzerland

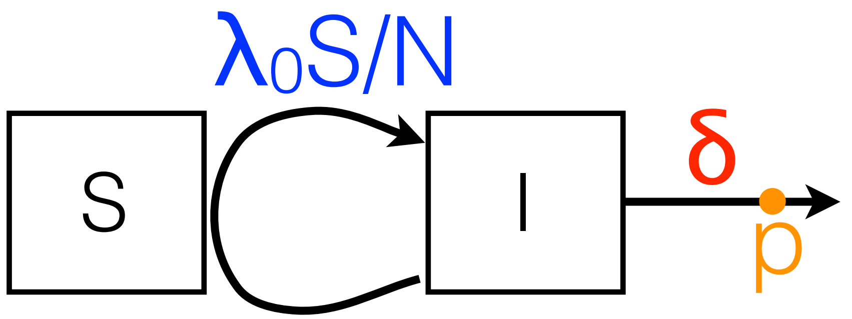


— SI model
- - - - epidemic outbreak model

Leventhal et al. (submitted)

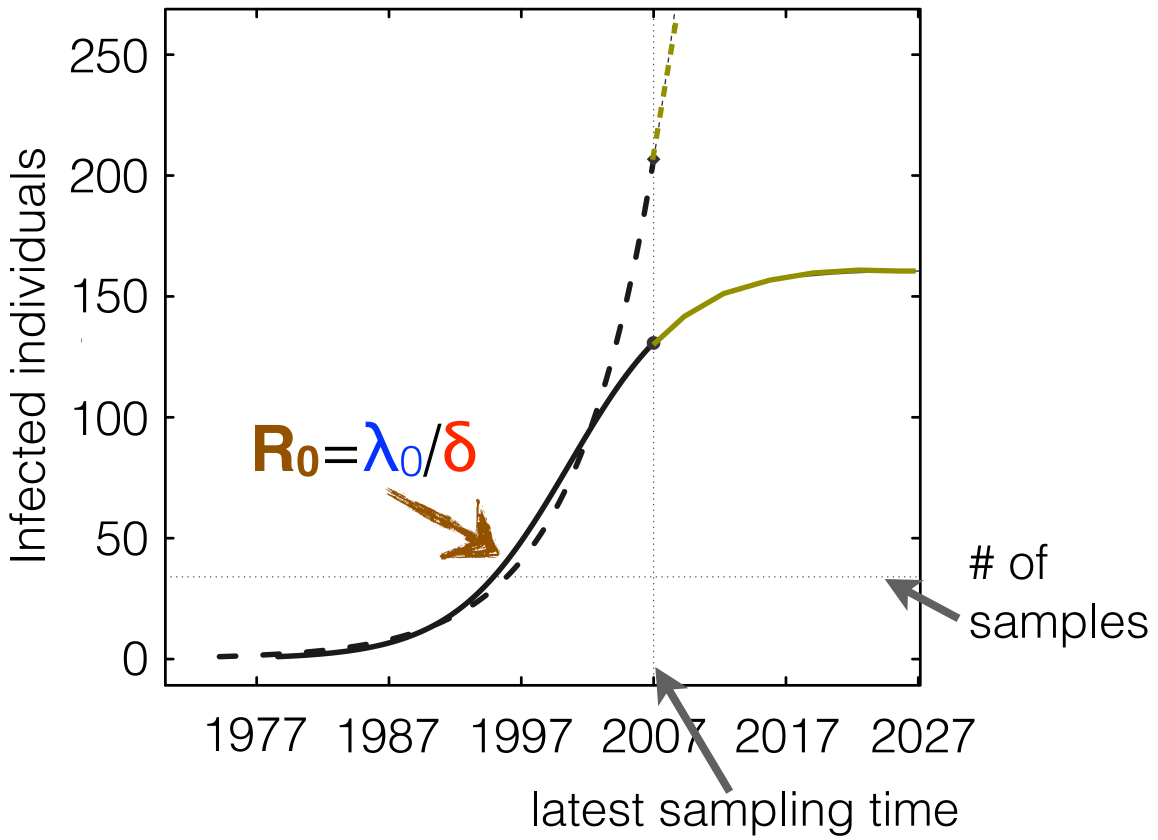
SI dynamics - exact method

Para-meters



N=S+I (constant)

HIV transmission cluster in Switzerland



R_0=λ_0/δ

# of samples

latest sampling time

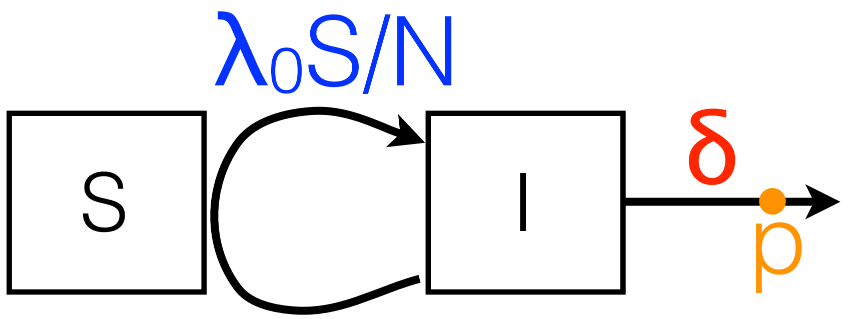
— SI model

- - - - epidemic outbreak model

Leventhal et al. (submitted)

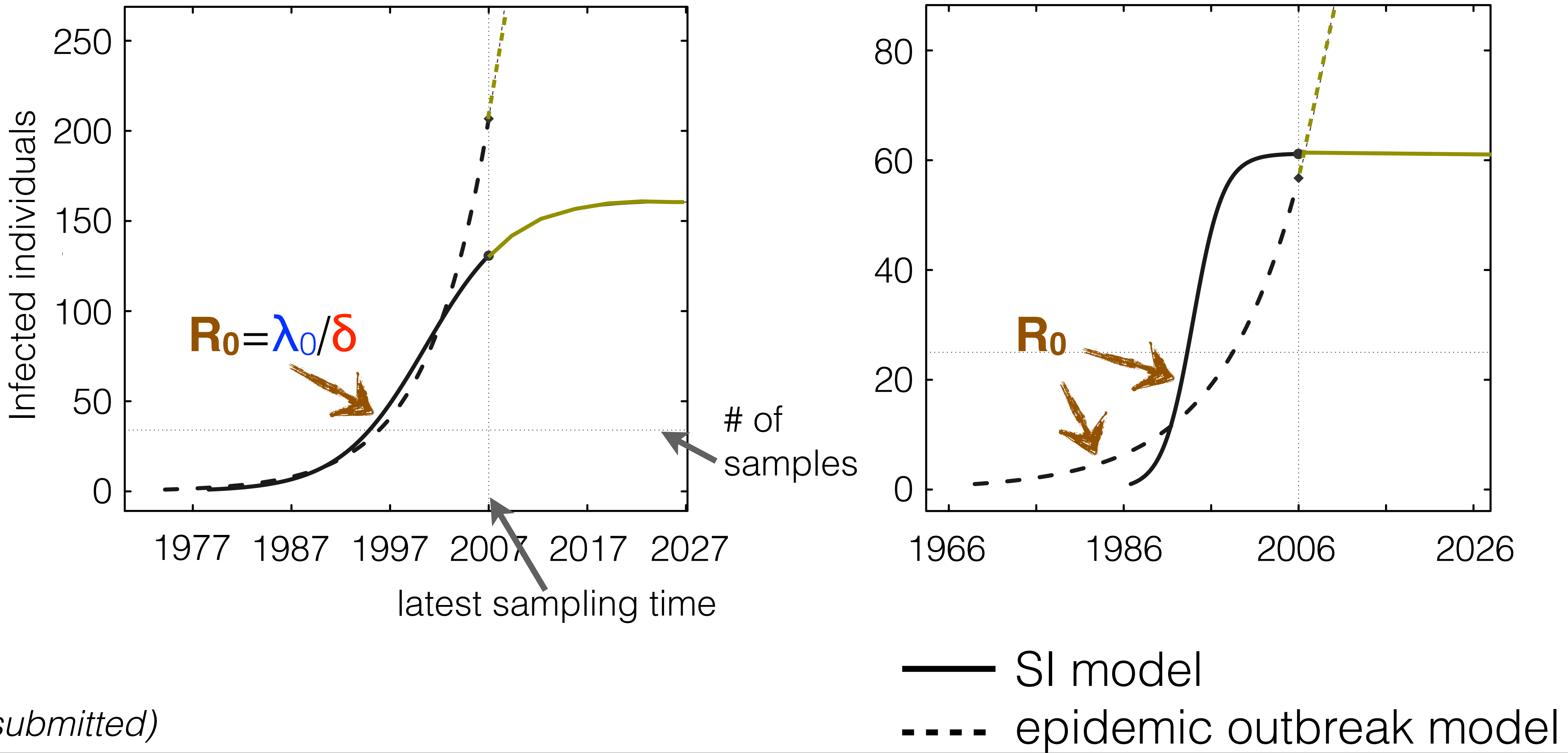
SI dynamics - exact method

Para-meters



N=S+I (constant)

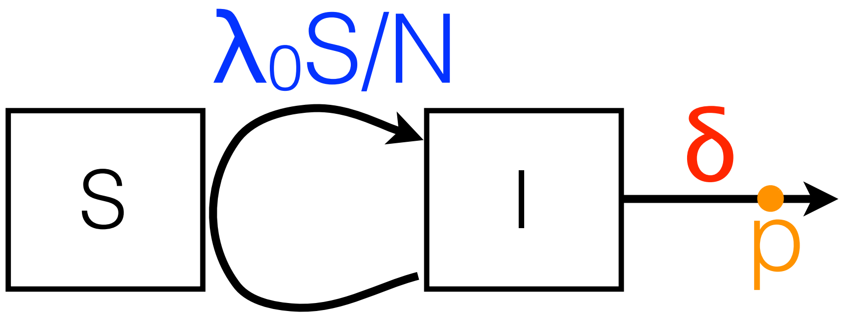
HIV transmission cluster in Switzerland



Leventhal et al. (submitted)

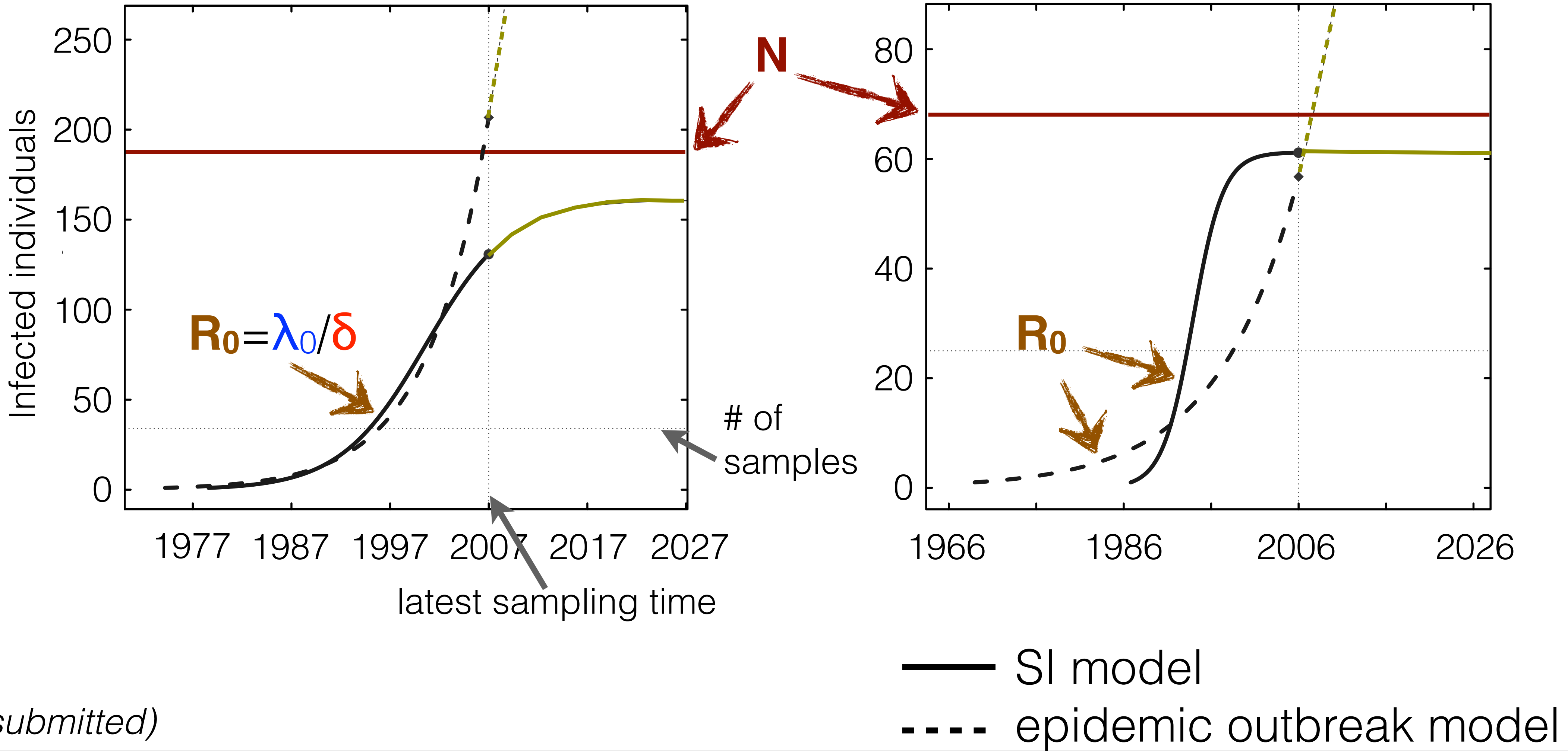
SI dynamics - exact method

Parameters



N=S+I (constant)

HIV transmission cluster in Switzerland



N

R_0=λ_0/δ

R_0

# of samples

latest sampling time

Leventhal et al. (submitted)

SIR models: Birth-death or coalescent?

Coalescent approximation

Volz et al., Genetics, 2009

So far sampling at one time point, not in Beast

Tim Vaughan implements it in Beast for comparison

SIR models: Birth-death or coalescent?

Coales-cent approxi-mation

Volz et al., Genetics, 2009

So far sampling at one time point, not in Beast

Tim Vaughan implements it in Beast for comparison

Birth-death approxi-mation

BDSIR add-on in Beast v2.0

Simulations show good performance

Flexible framework (SI,SIS,SIR,SEIR, etc)

Poster: Denise Kühnert

SIR models: Birth-death or coalescent?

Coales-cent approxi-mation

Volz et al., Genetics, 2009

So far sampling at one time point, not in Beast

Tim Vaughan implements it in Beast for comparison

Birth-death approxi-mation

BDSIR add-on in Beast v2.0

Simulations show good performance

Flexible framework (SI,SIS,SIR,SEIR, etc)

Poster: Denise Kühnert

Birth-death exact

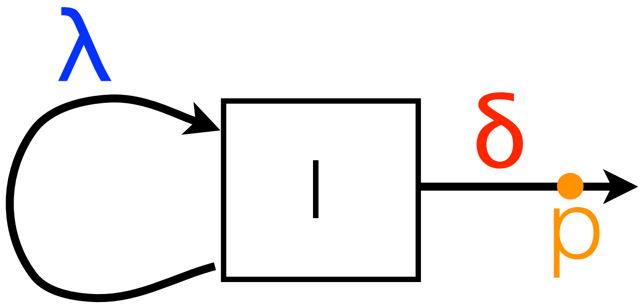
expoTree (R package)

Promising to estimate overall population size N=S+I

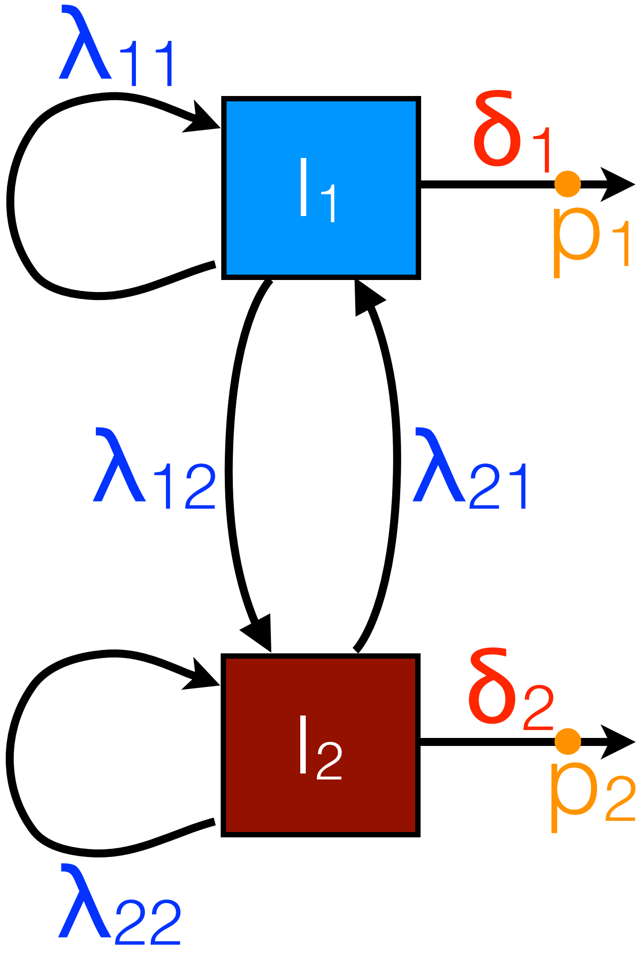
Future study will identify potential limitations of approximate method

Poster: Gabriel Leventhal

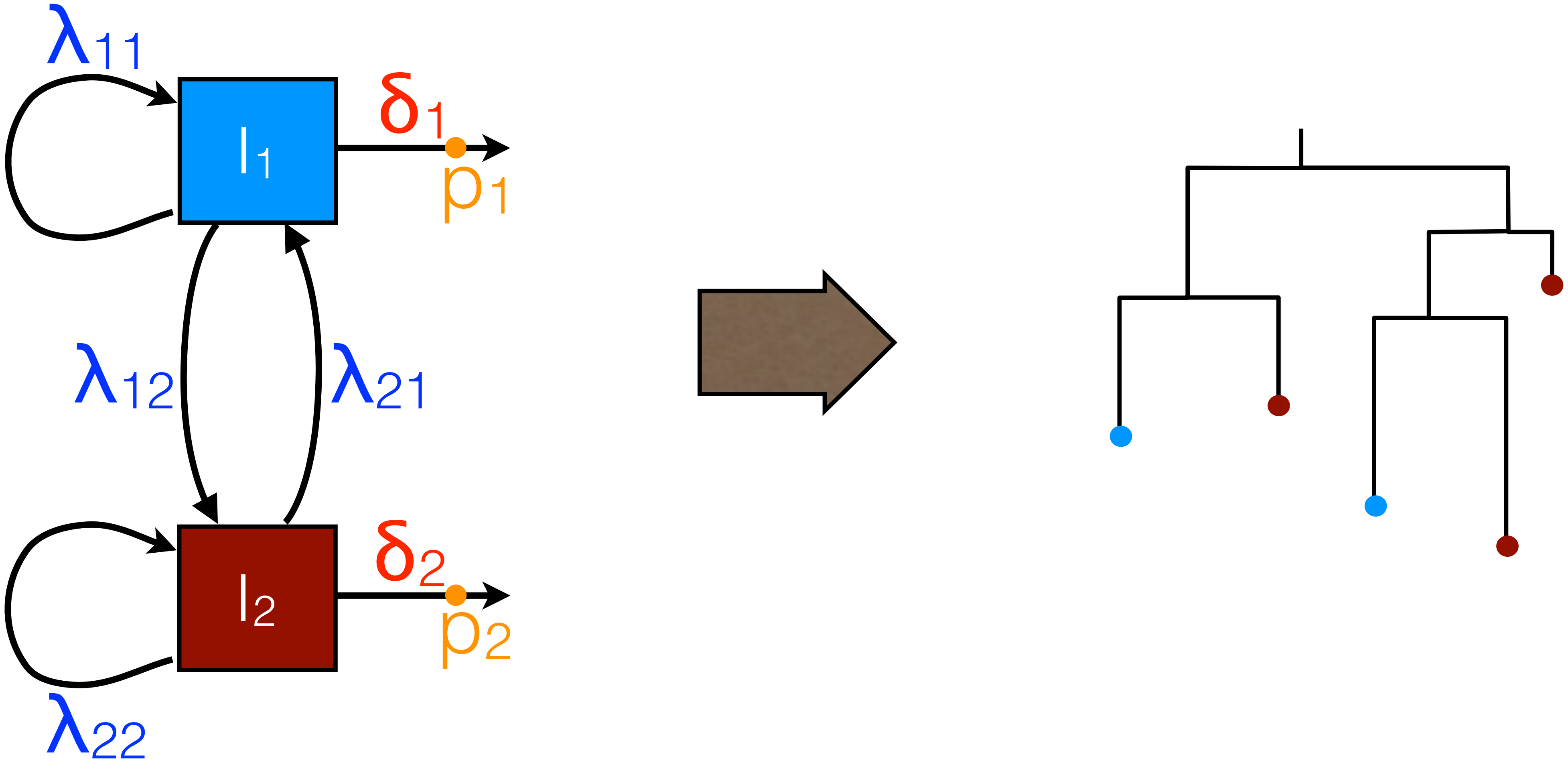
4) Host population structure



4) Host population structure



4) Host population structure



Quantification of risk-group dependent spread of HIV

Two states

Heterosexual

Intravenous drug user

Stadler et al. (Phil. Trans. Roy. Soc. B, 2013)

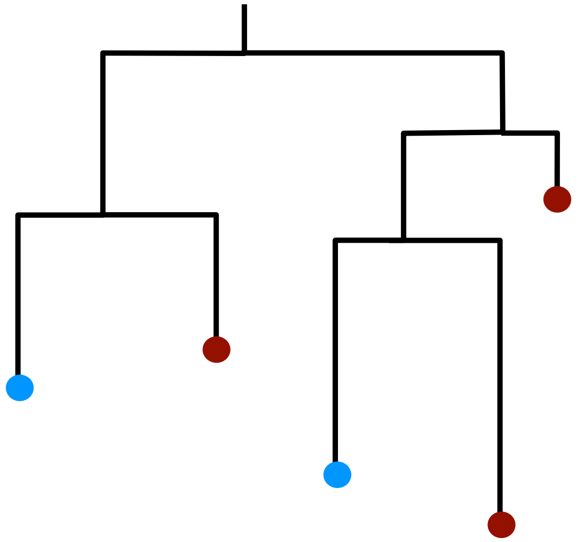
Quantification of risk-group dependent spread of HIV

Two states

Heterosexual

Intravenous drug user

Data



220 Latvian sequences from
Balode et al. (2012)

Stadler et al. (Phil. Trans. Roy. Soc. B, 2013)

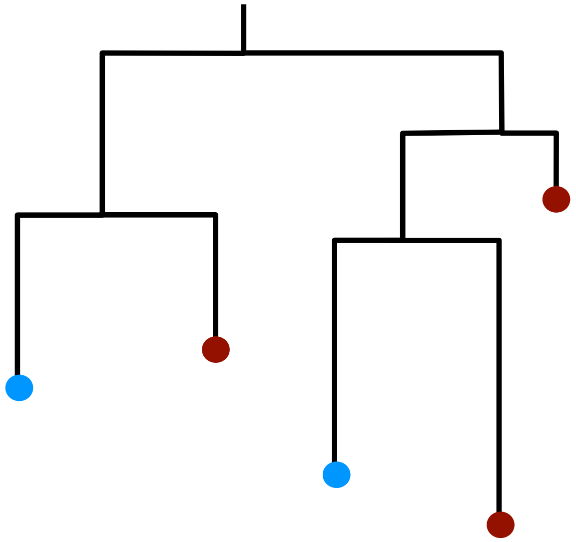
Quantification of risk-group dependent spread of HIV

Two states

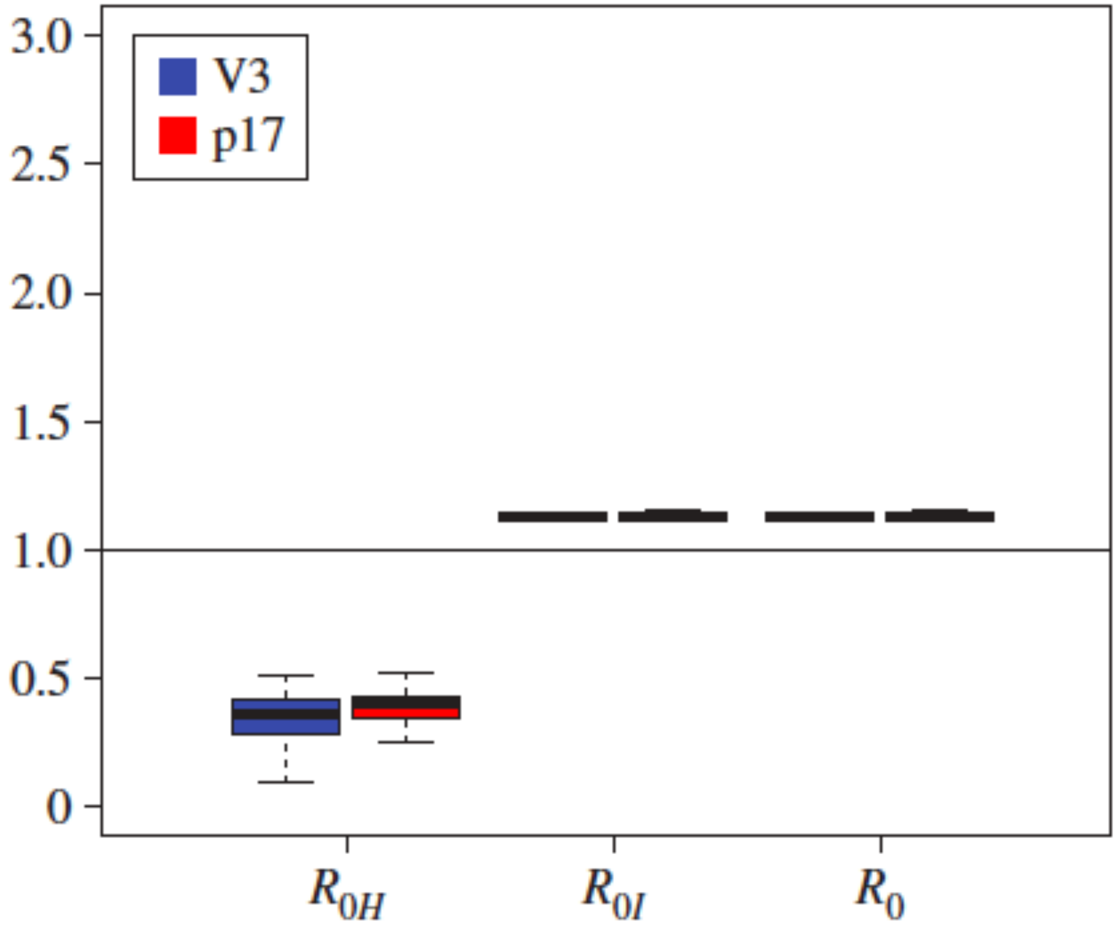
Heterosexual

Intravenous drug user

Data



220 Latvian sequences from Balode et al. (2012)



Stadler et al. (Phil. Trans. Roy. Soc. B, 2013)

Quantification of super-spreading in HIV

Two states

Super-spreader

Normal-spreader

Stadler et al. (Phil. Trans. Roy. Soc. B, 2013)

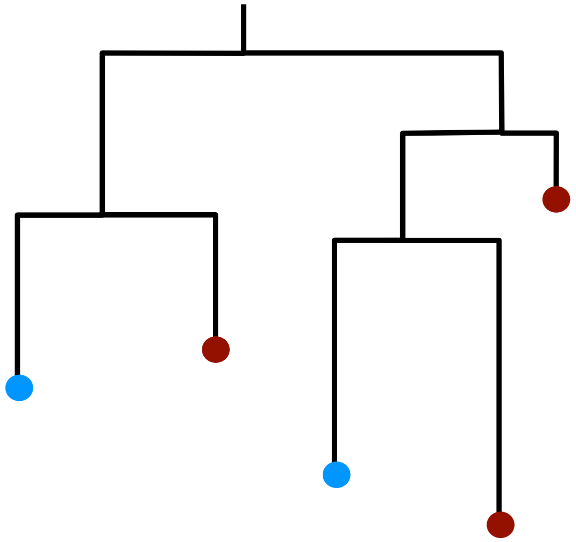
Quantification of super-spreading in HIV

Two states

Super-spreader

Normal-spreader

Data



40 Latvian sequences from
Balode et al. (2012)

Stadler et al. (Phil. Trans. Roy. Soc. B, 2013)

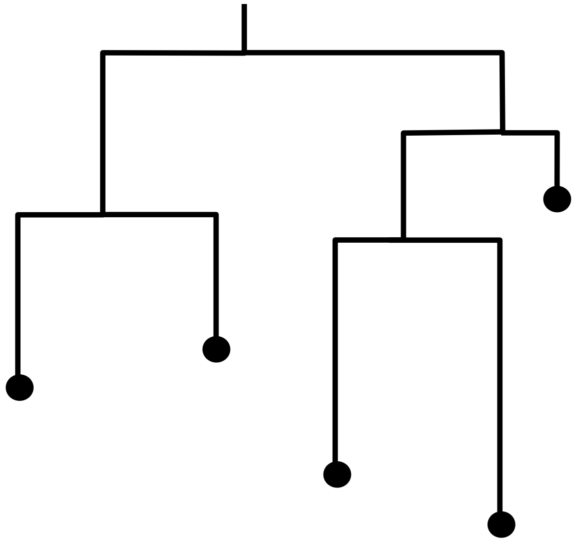
Quantification of super-spreading in HIV

Two states

Super-spreader

Normal-spreader

Data



40 Latvian sequences from
Balode et al. (2012)

Stadler et al. (Phil. Trans. Roy. Soc. B, 2013)

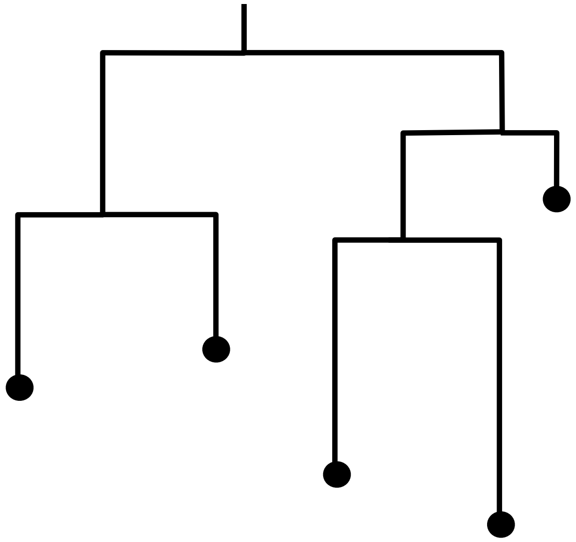
Quantification of super-spreading in HIV

Two states

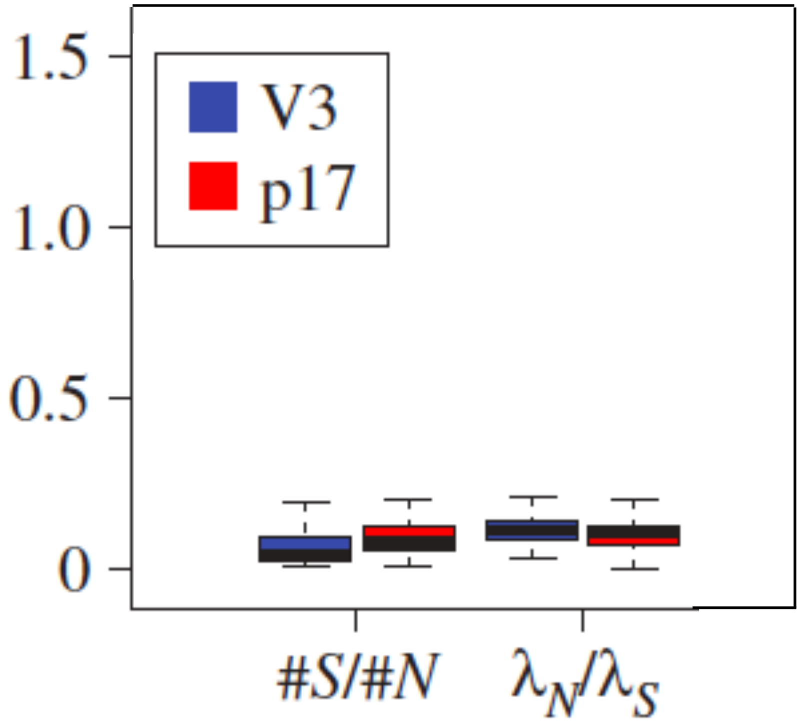
Super-spreader

Normal-spreader

Data



40 Latvian sequences from
Balode et al. (2012)



Stadler et al. (Phil. Trans. Roy. Soc. B, 2013)

Key questions in infectious disease ecology approached through phylogenetics

Human pathogens

PUBLIC HEALTH

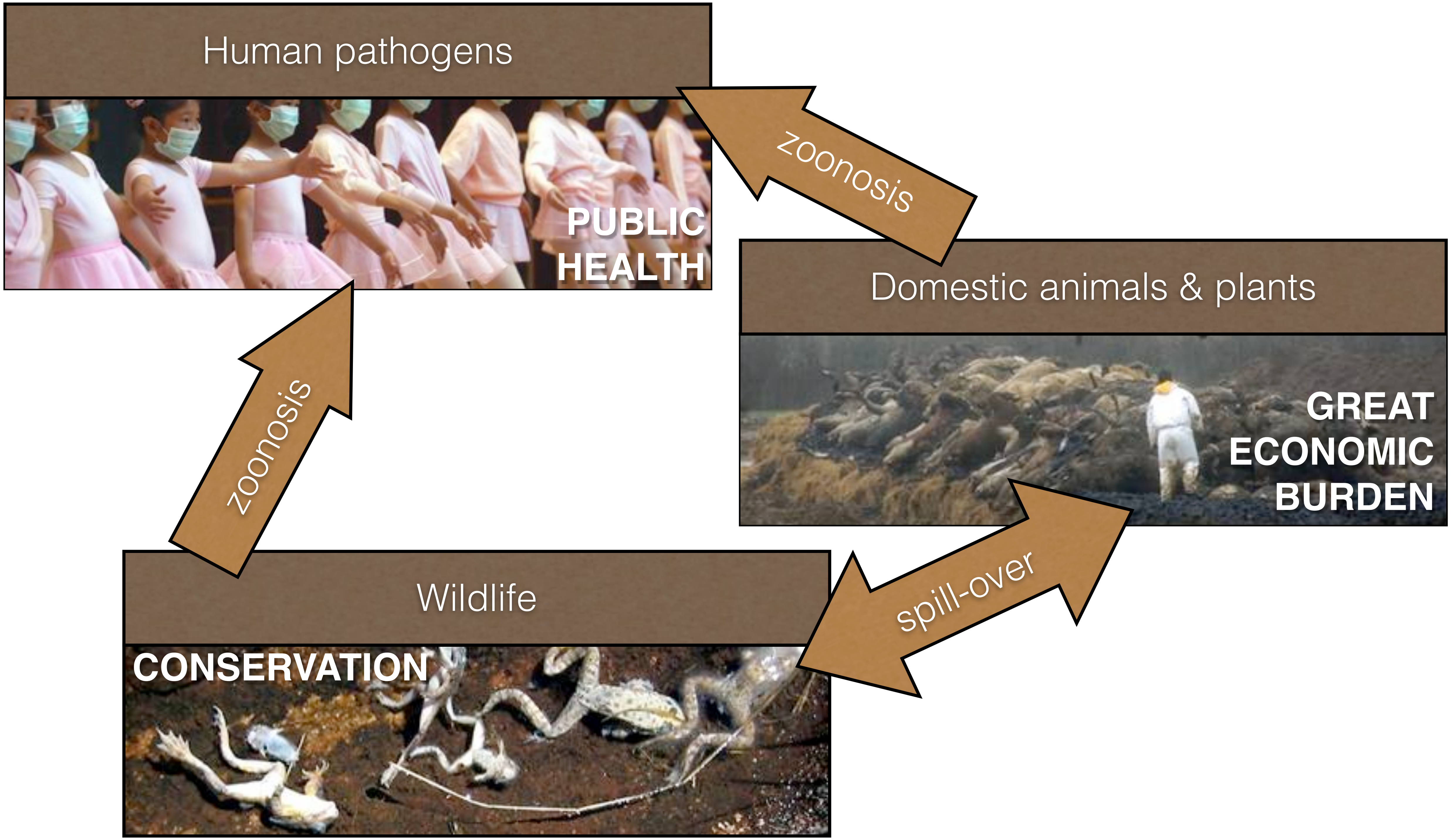
Domestic animals & plants

GREAT ECONOMIC BURDEN

Wildlife

CONSERVATION

Key questions in infectious disease ecology approached through phylogenetics



Human pathogens

PUBLIC HEALTH

zoonosis

Domestic animals & plants

GREAT ECONOMIC BURDEN

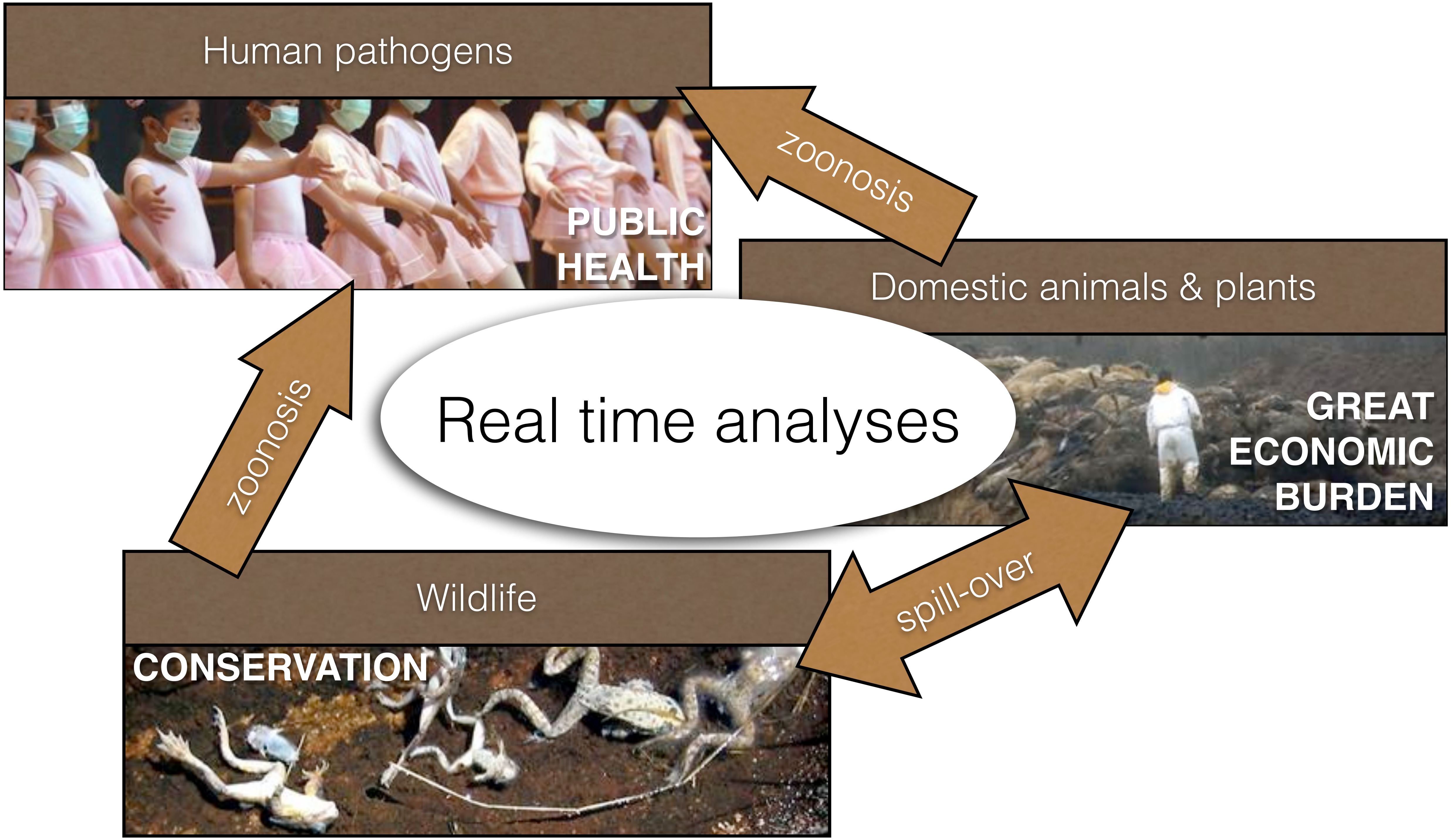
zoonosis

spill-over

Wildlife

CONSERVATION

Key questions in infectious disease ecology approached through phylogenetics



Human pathogens

PUBLIC HEALTH

zoonosis

Domestic animals & plants

GREAT ECONOMIC BURDEN

Real time analyses

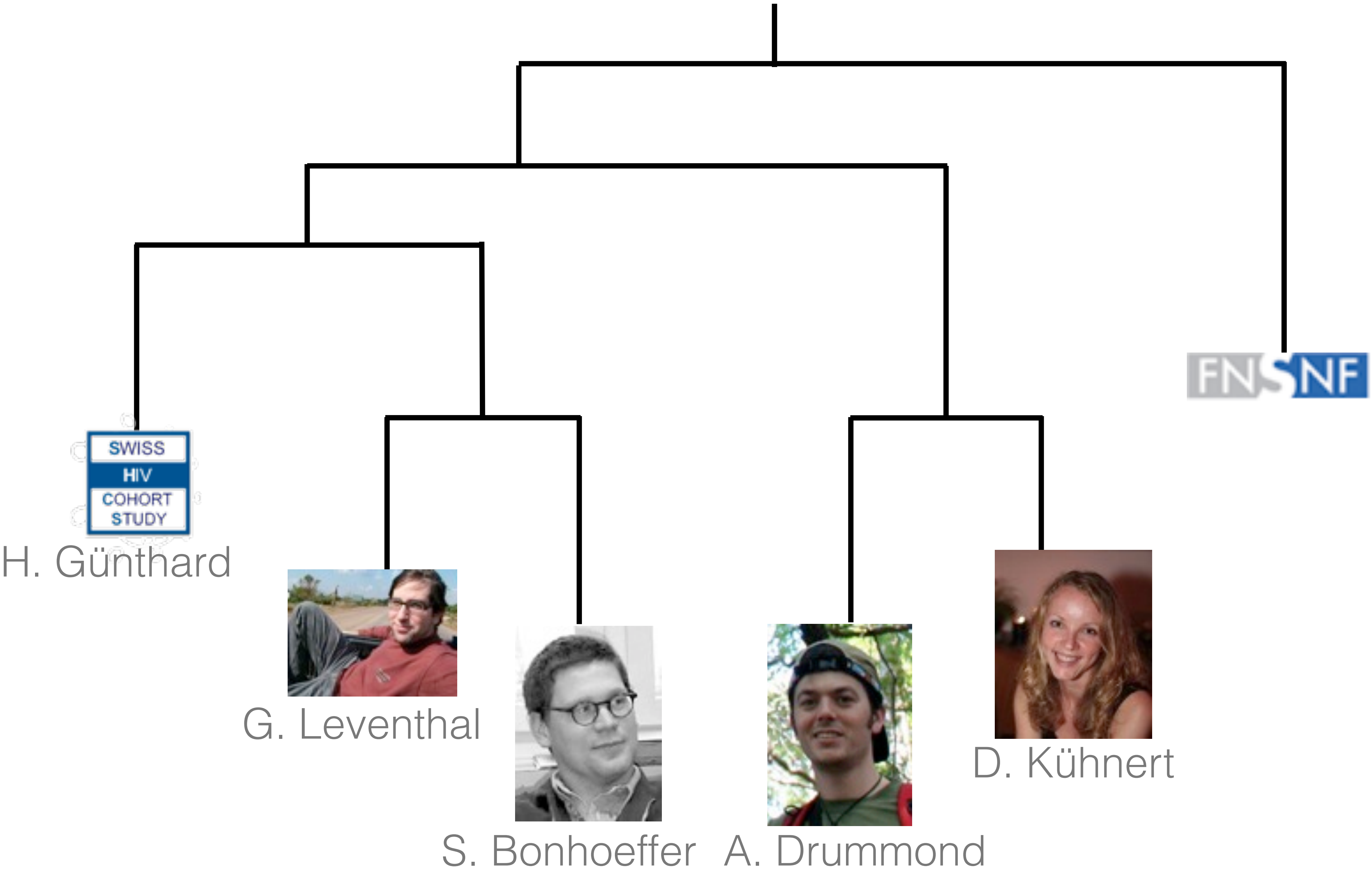
zoonosis

Wildlife

CONSERVATION

spill-over

Phylogeny of Acknowledgements



SWISS HIV COHORT STUDY

H. Günthard

G. Leventhal

S. Bonhoeffer

A. Drummond

D. Kühnert

FNSNF