

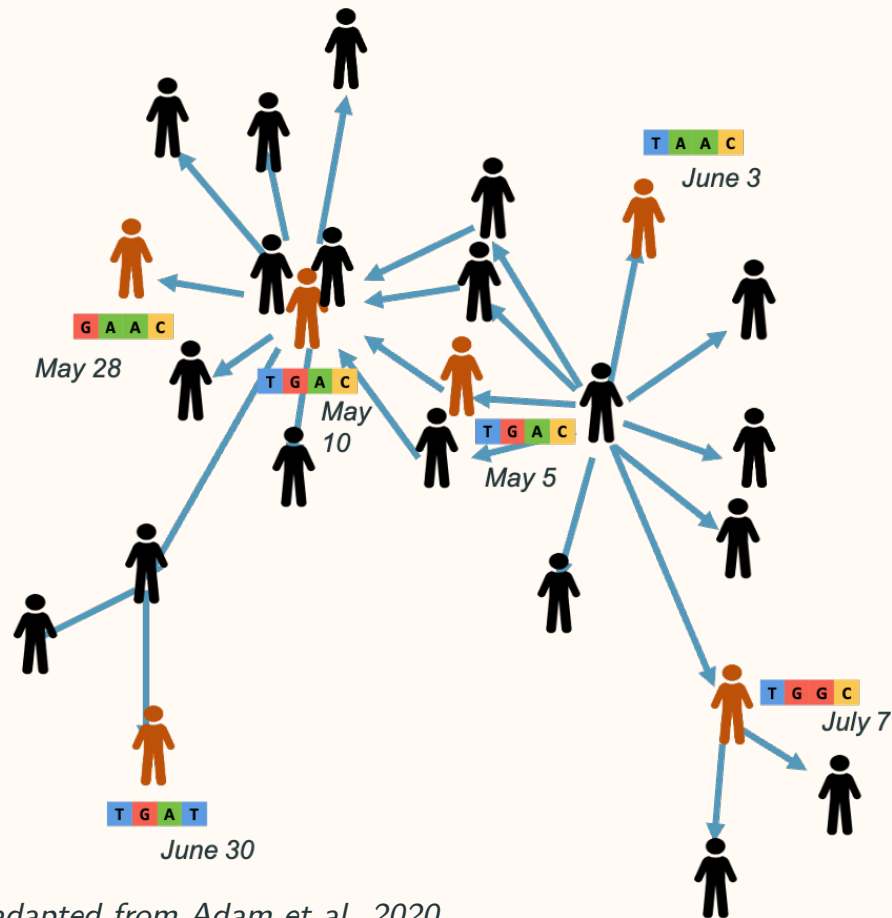
# Phylogenetic Inference across Micro and Macroevolutionary Scales

---

Taming the BEAST: August 16th, 2023

Simon Fraser University

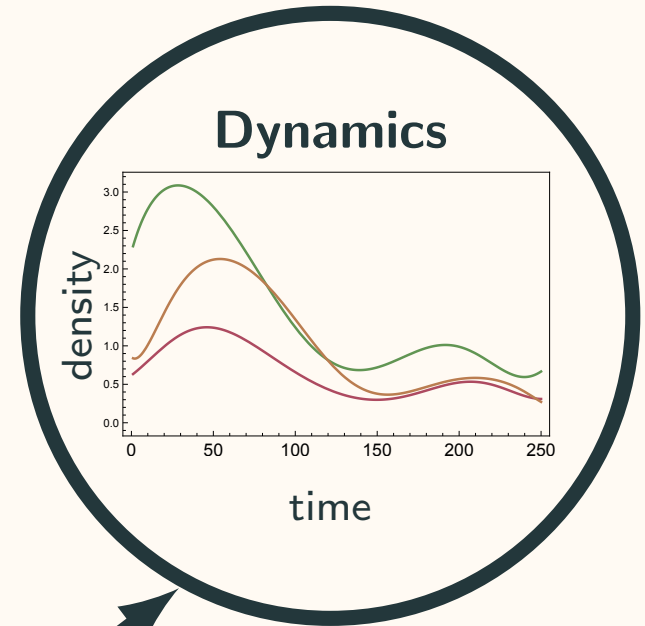
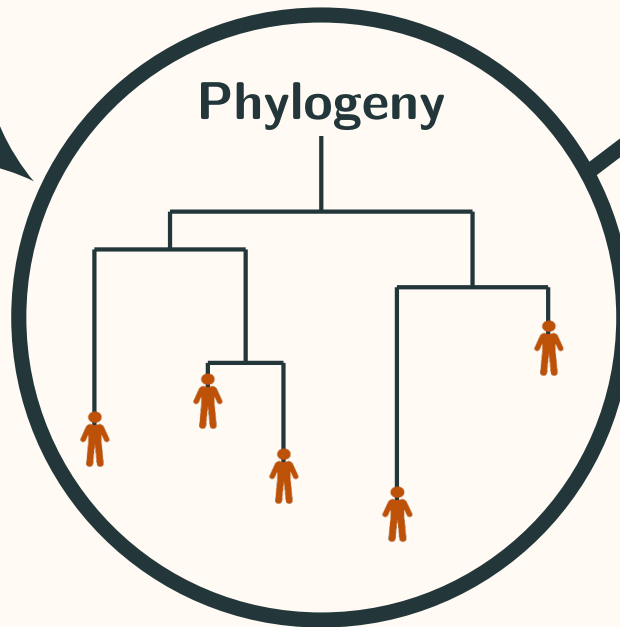
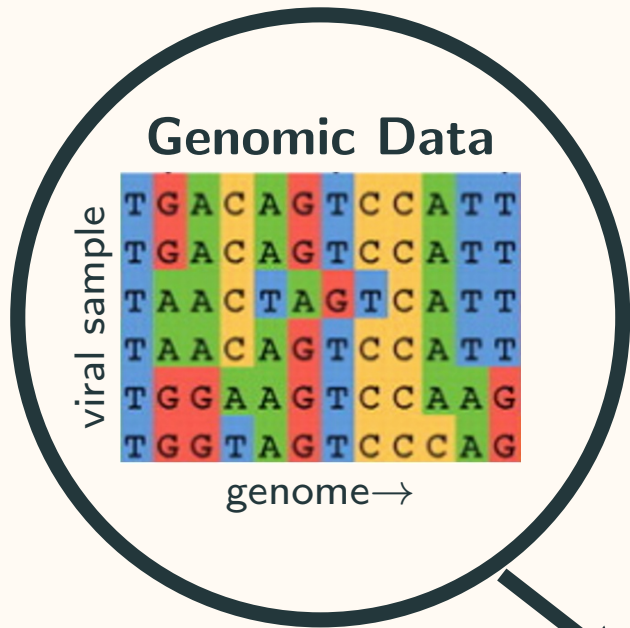
# Viral Evolution



—adapted from Adam et al. 2020

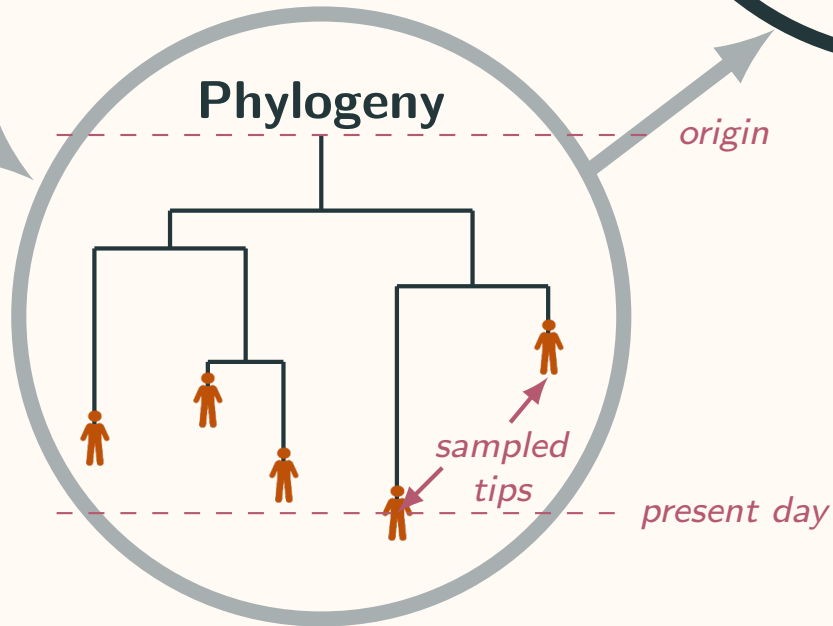
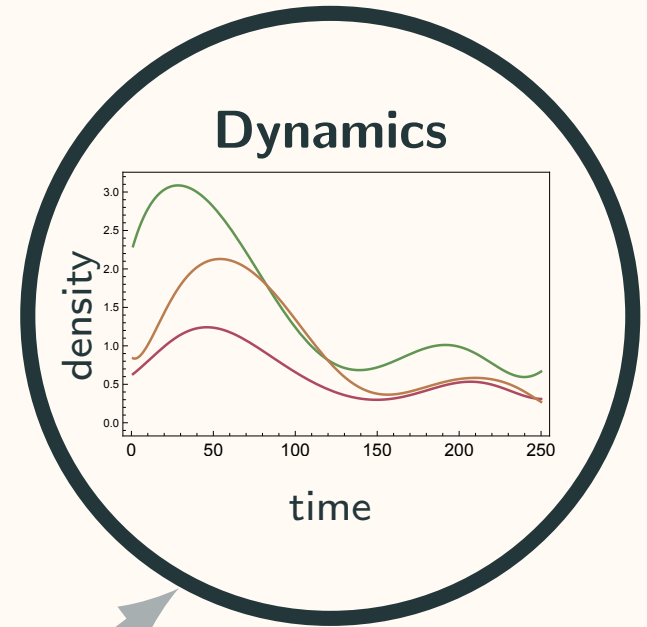
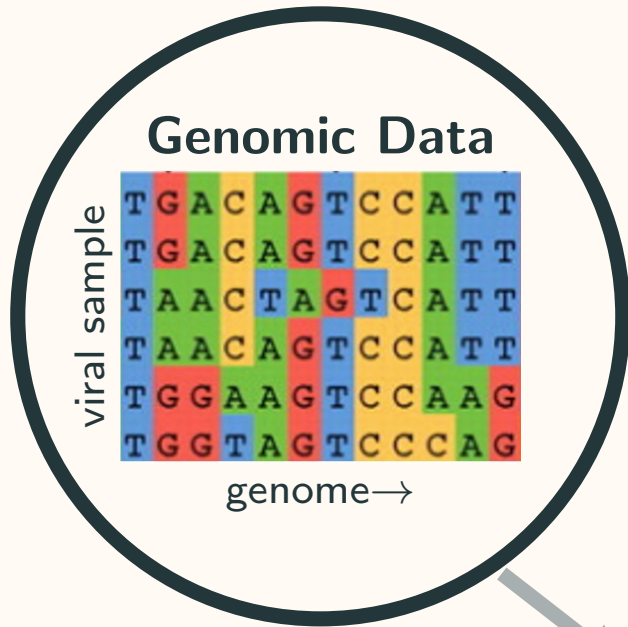
# Viral Phylodynamics

## “Phylodynamics”



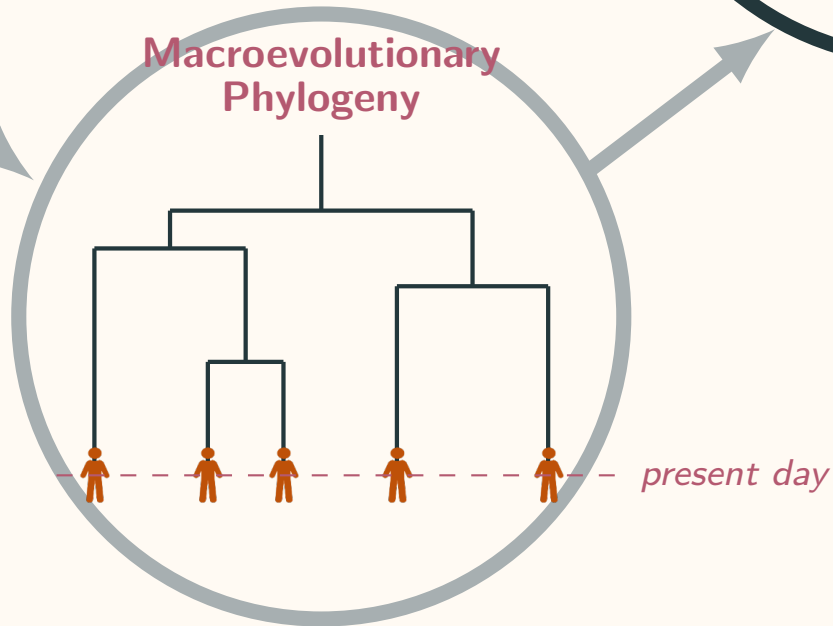
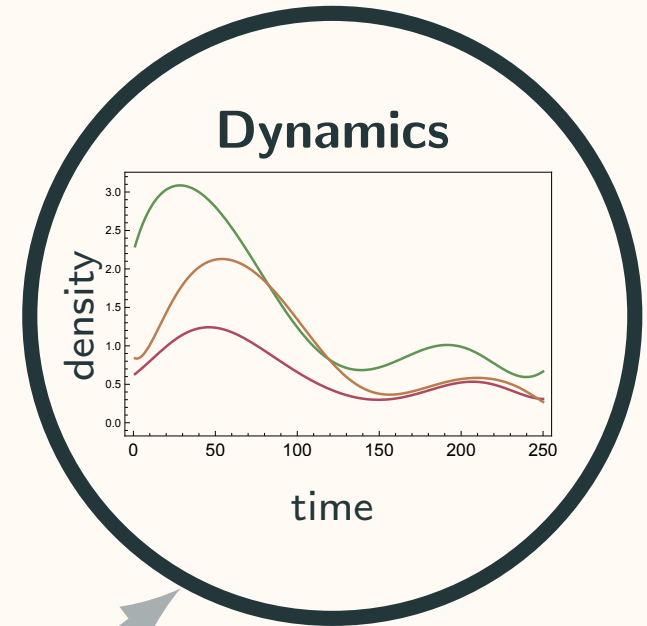
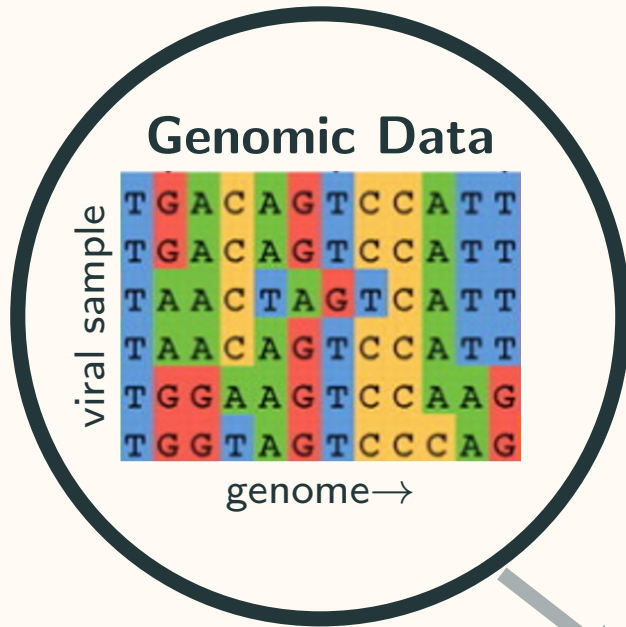
# Viral Phylodynamics

## “Phylodynamics”



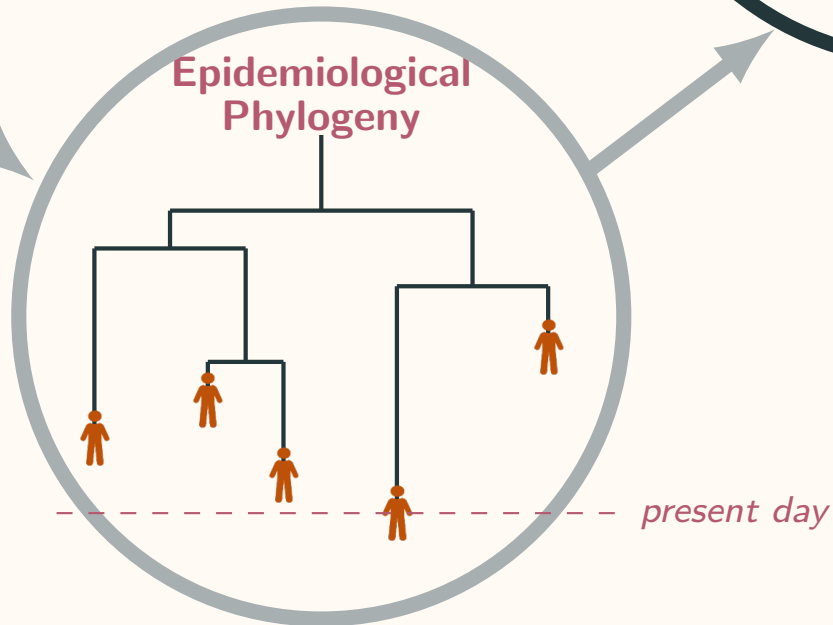
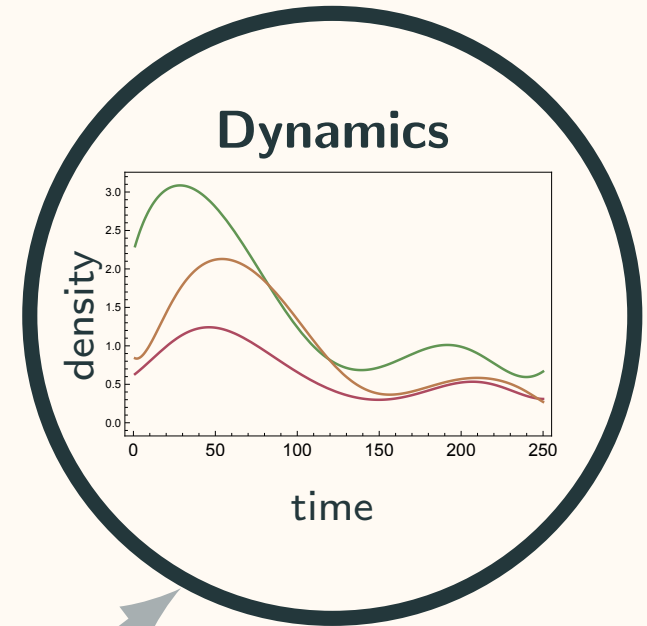
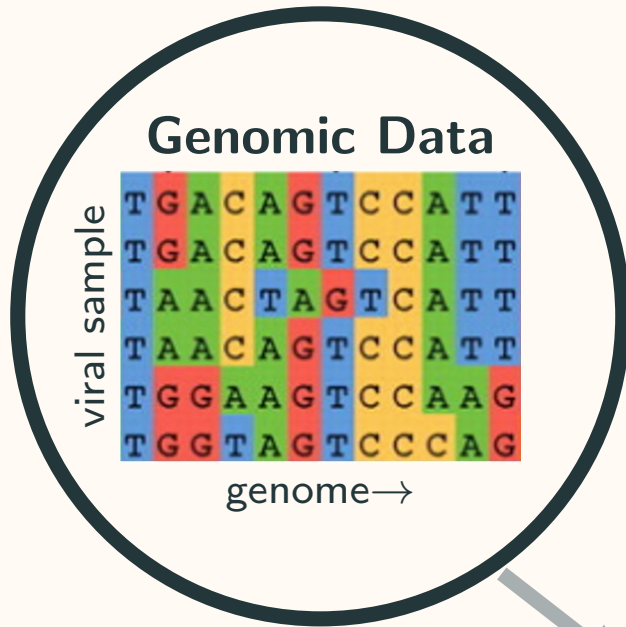
# Viral Phylodynamics

## “Phylodynamics”



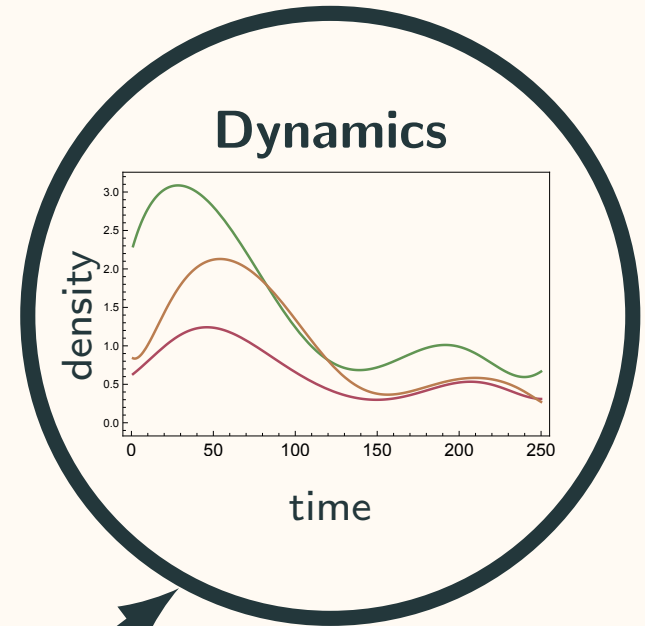
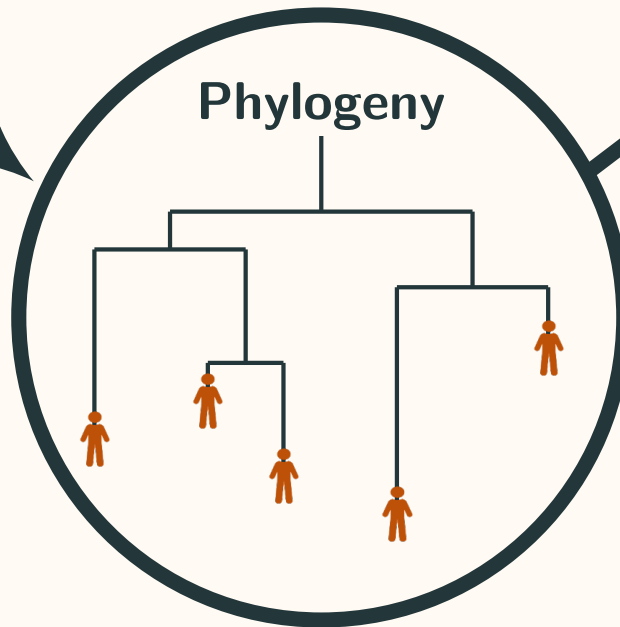
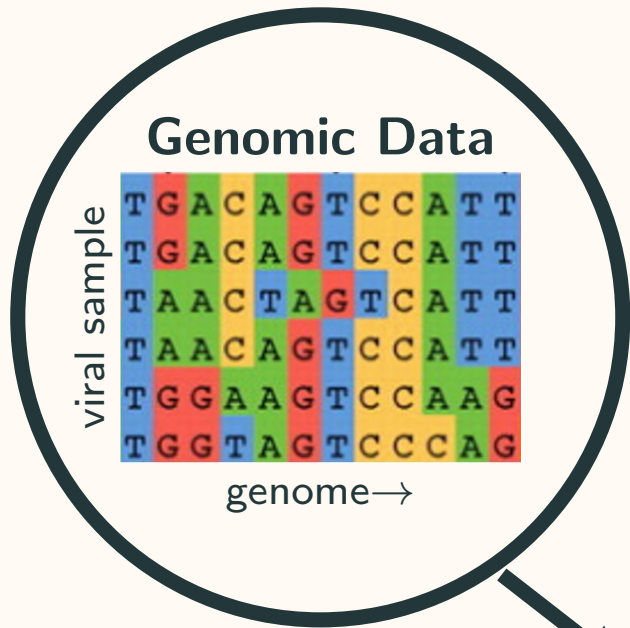
# Viral Phylodynamics

## “Phylodynamics”



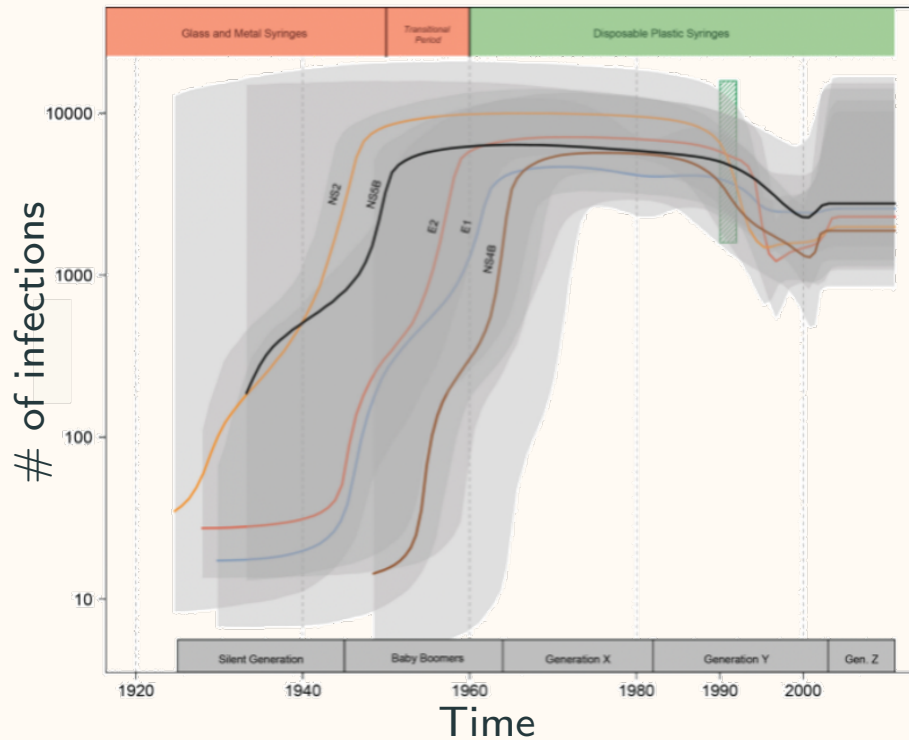
# Viral Phylodynamics

## “Phylodynamics”



# Phylodynamic Example

## HCV in the Americas



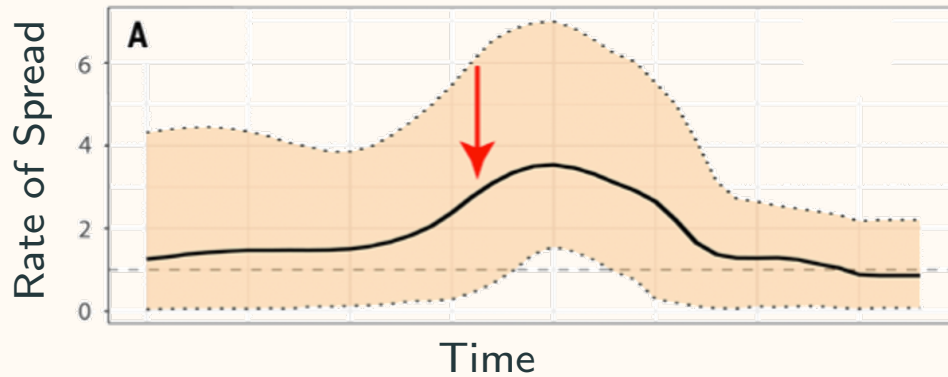
—Joy et al. 2016

Did hospital glassware contribute to the spread of Hepatitis C Virus?



# Phylodynamic Example

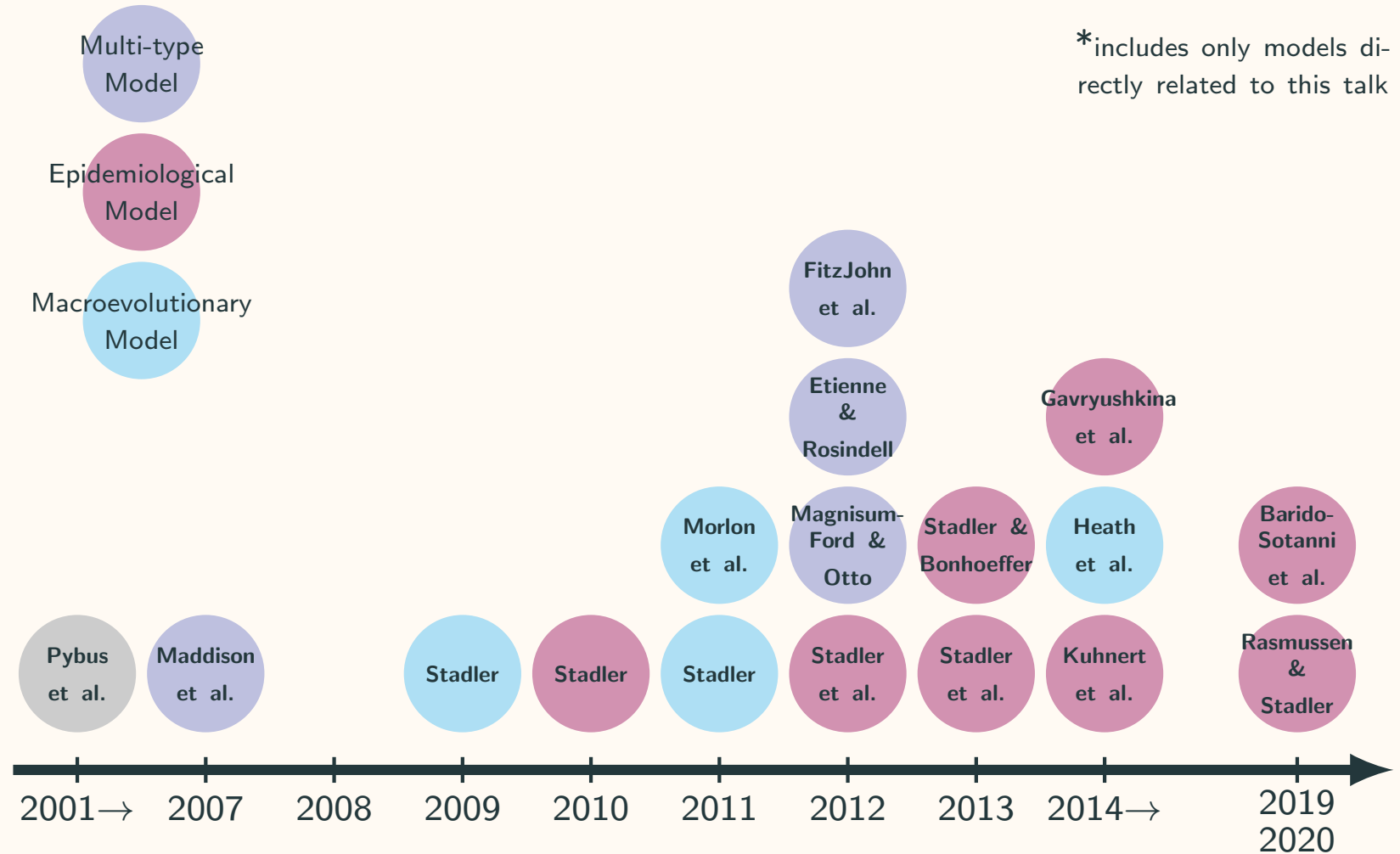
## Cancer in Tasmanian Devils



– Patton et al. 2020

Will facial tumors drive Tasmanian devils extinct?

# Historical Phylodynamic Model Developments



# A General Birth-Death-Sampling Model

---

**Aim:** Develop a general birth-death sampling model that allows rates to vary through time.

**Aim:** Develop a general birth-death sampling model that allows rates to vary through time.

*Why develop a general model?*

- I. Clarifies assumptions of and connections between existing models
- II. Allows us to explicitly fit an SIR epidemiological model
- III. Enables mathematical analysis of the general model class

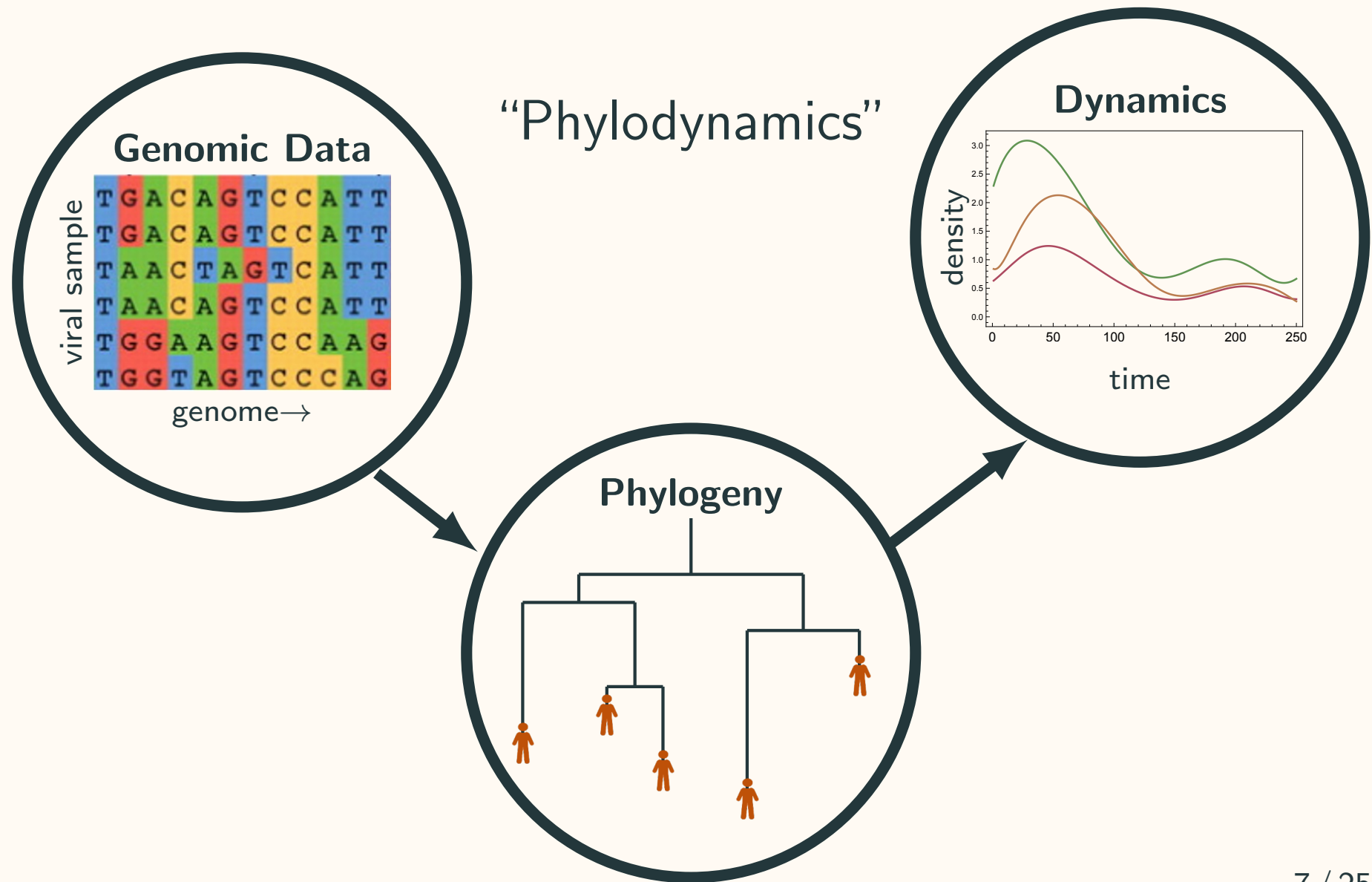
**Aim:** Develop a general birth-death sampling model that allows rates to vary through time.

*Why develop a general model?*

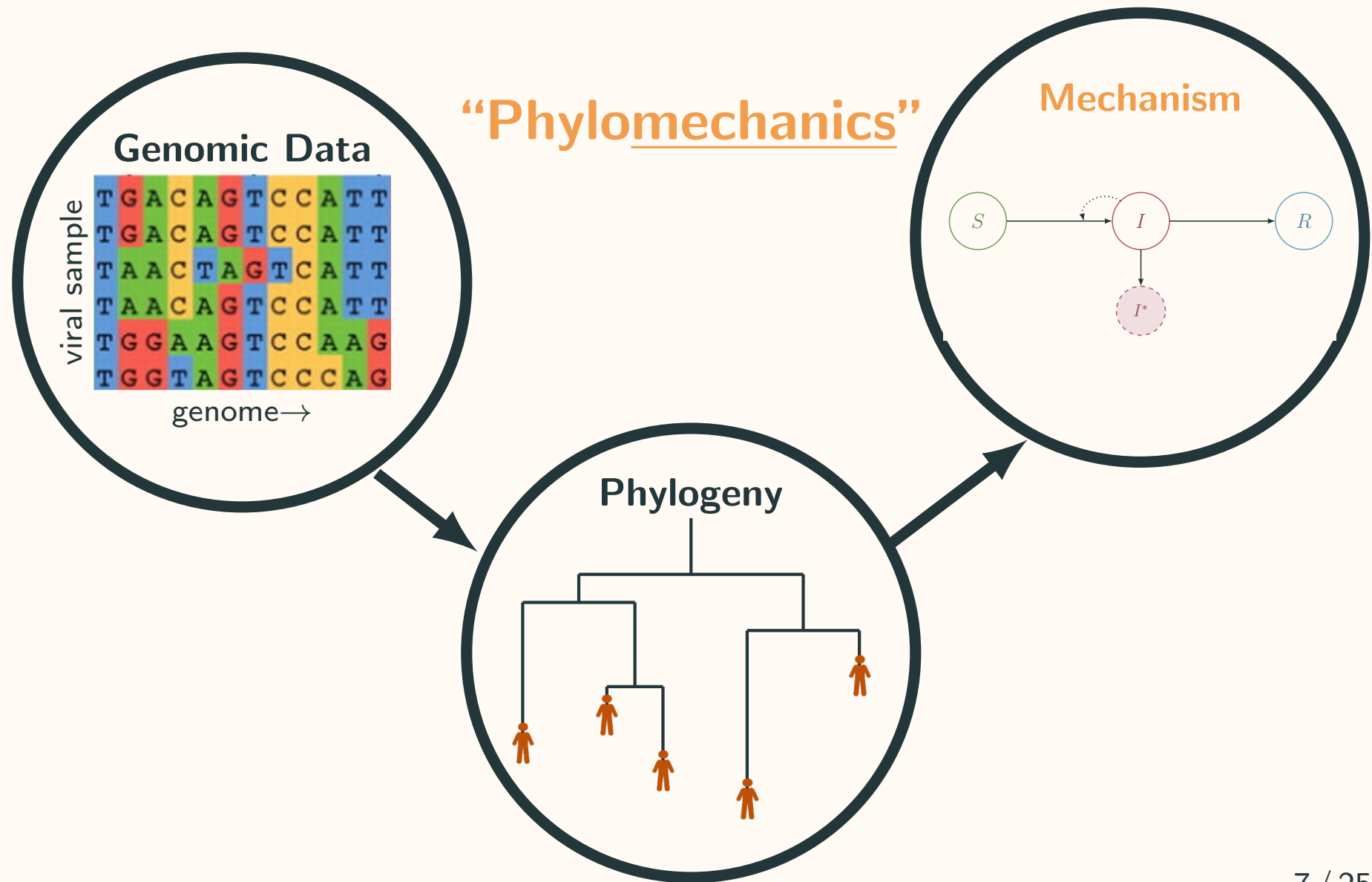
**Formulate general model**

- I. Clarifies assumptions of and connections between existing models
- II. Allows us to explicitly fit an SIR epidemiological model
- III. Enables mathematical analysis of the general model class

# Viral Phylogenetic Reconstruction

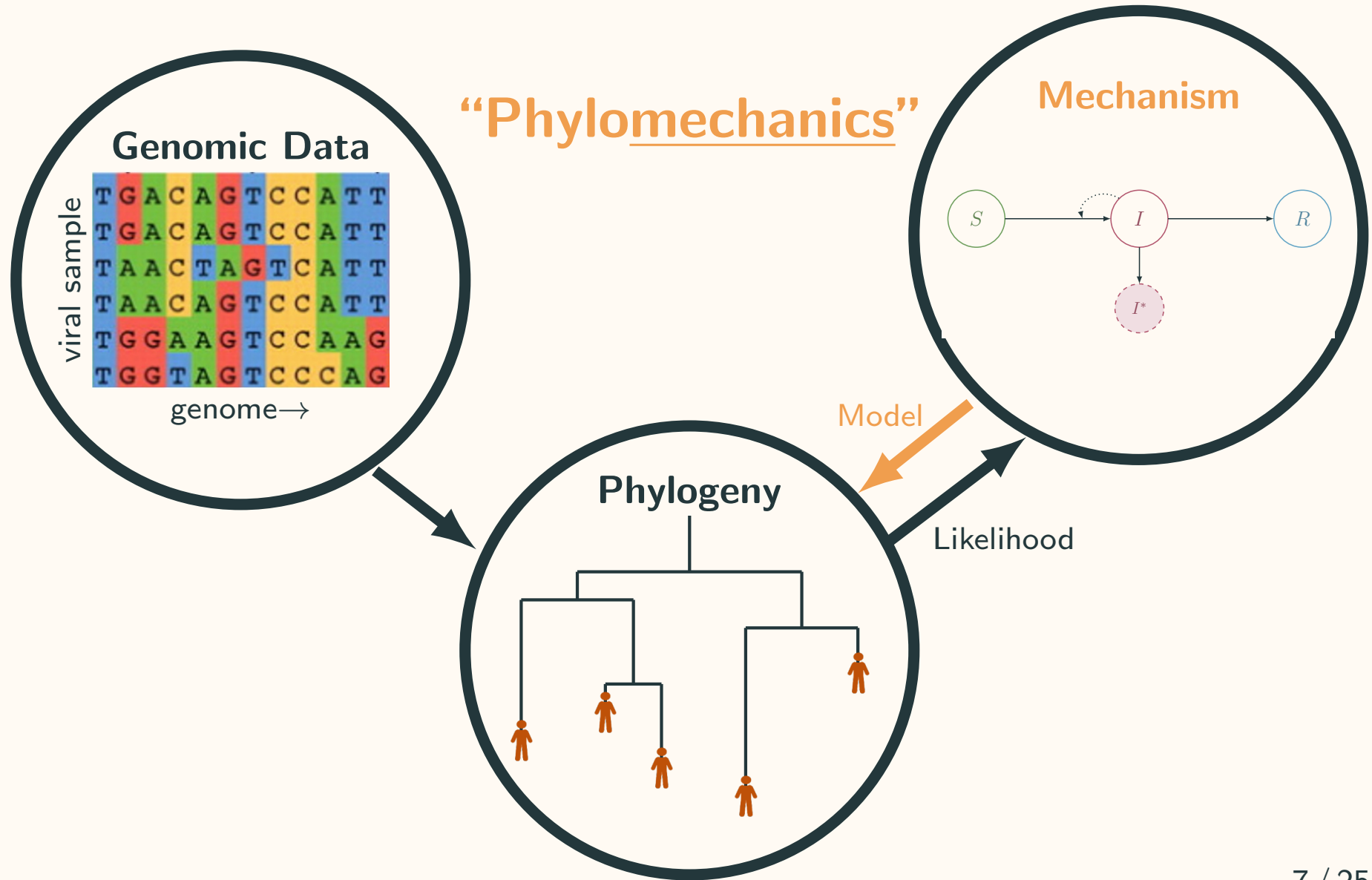


# Viral Phylogenetic Reconstruction



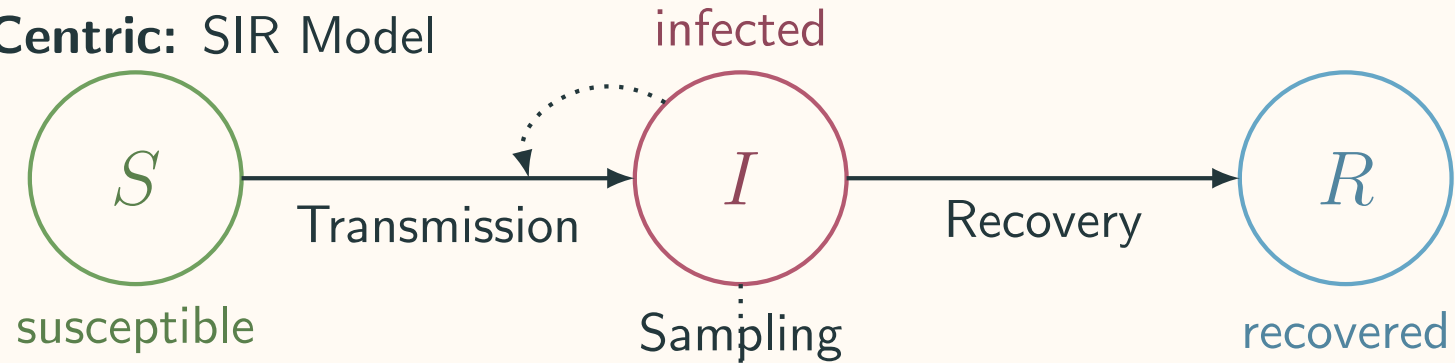


# Viral Phylogenetic Reconstruction

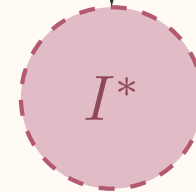


# The SIR Model → BDS Model

**Host Centric: SIR Model**

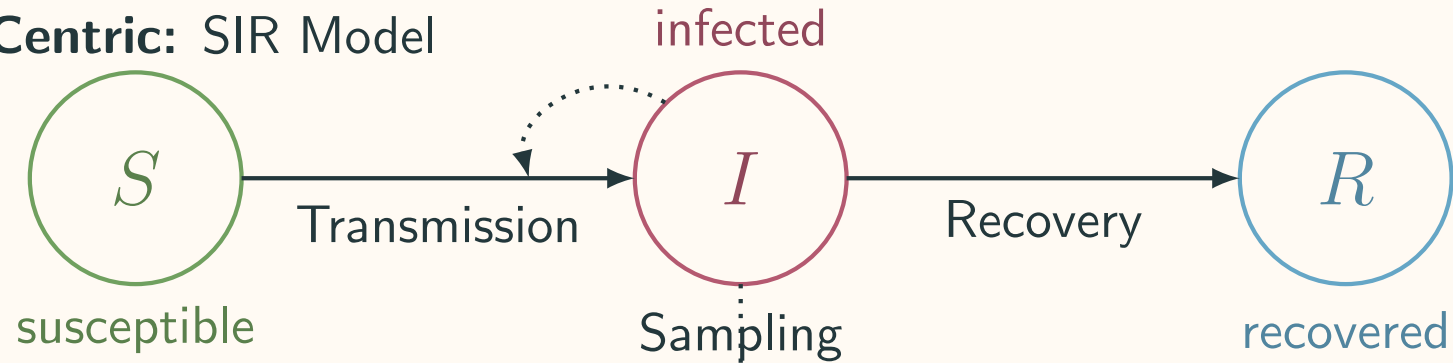


**Pathogen Centric: BDS Model**

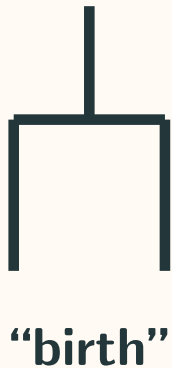


# The SIR Model → BDS Model

**Host Centric: SIR Model**

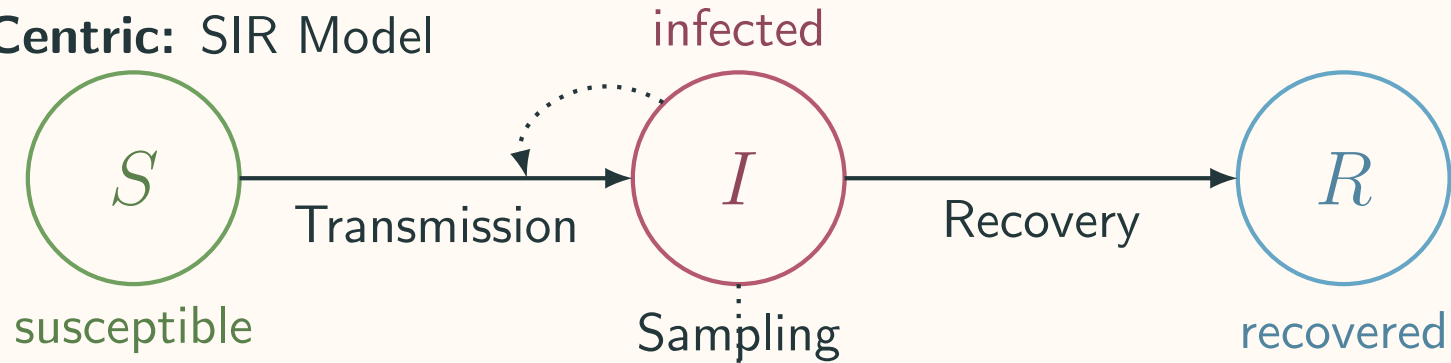


**Pathogen Centric: BDS Model**

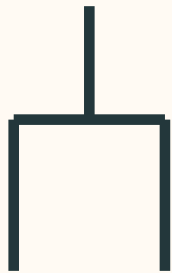


# The SIR Model → BDS Model

**Host Centric: SIR Model**



**Pathogen Centric: BDS Model**



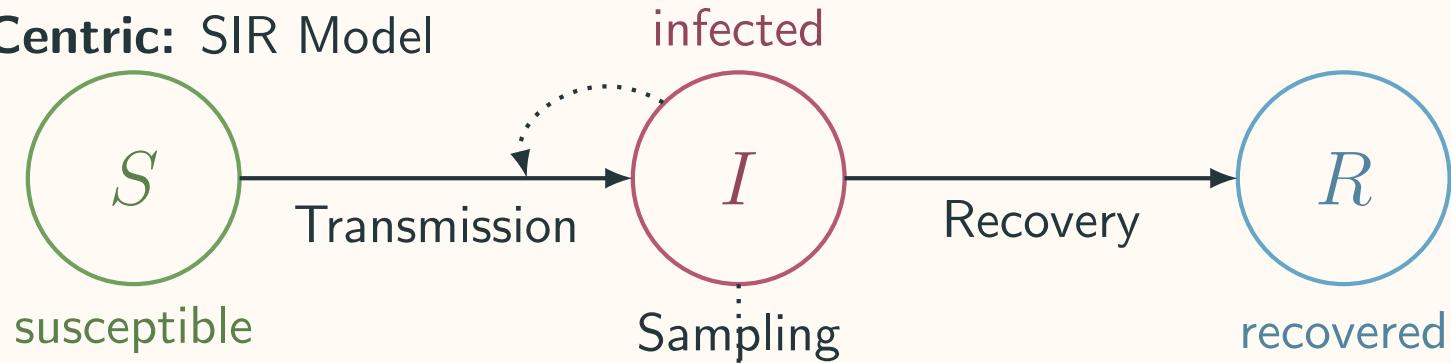
“birth”



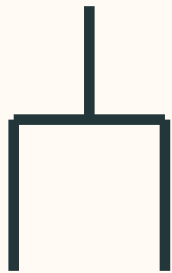
“death”

# The SIR Model → BDS Model

**Host Centric: SIR Model**



**Pathogen Centric: BDS Model**



“birth”



“sampling”



“death”

# A General Birth-Death-Sampling Framework

## Six-Step Framework

### Step 1: Specify the Model

time of origin  $T$

rates  $\lambda(\tau), \mu(\tau), \psi(\tau)$

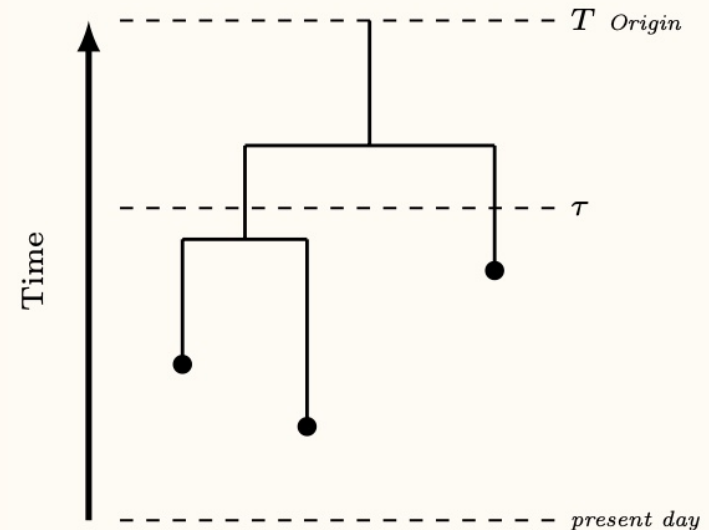
sampling  $\rho_i, r$

mass extinction  $\nu_i$

conditioning  $\mathcal{S}$

### Model Parameters:

$$\Theta = \{\lambda, \mu, \psi, \rho, r, \nu, \mathcal{S}\}$$



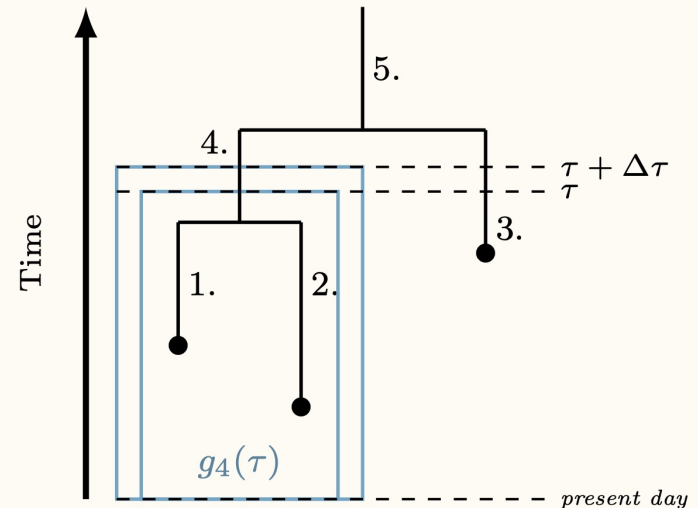
# A General Birth-Death-Sampling Framework

## Six-Step Framework

**Step 1:** Specify the Model

**Step 2:** Derive IVP for  $g_e(\tau)$

$g_e(\tau)$  is the probability that an edge  $e$  alive at time  $\tau$  gives rise to the observed phylogeny between  $\tau$  and the present day.



$$\frac{dg_e}{d\tau} = \underbrace{-(\lambda + \mu + \psi)g_e(\tau)}_{\text{nothing happens}} + \underbrace{2\lambda g_e(\tau)E(\tau)}_{\text{birth}}$$

$$g_e(s_e) = \begin{cases} \lambda g_{e1}(s_e)g_{e2}(s_e) & \text{birth} \\ \psi & \text{sample} \end{cases}$$

# A General Birth-Death-Sampling Framework

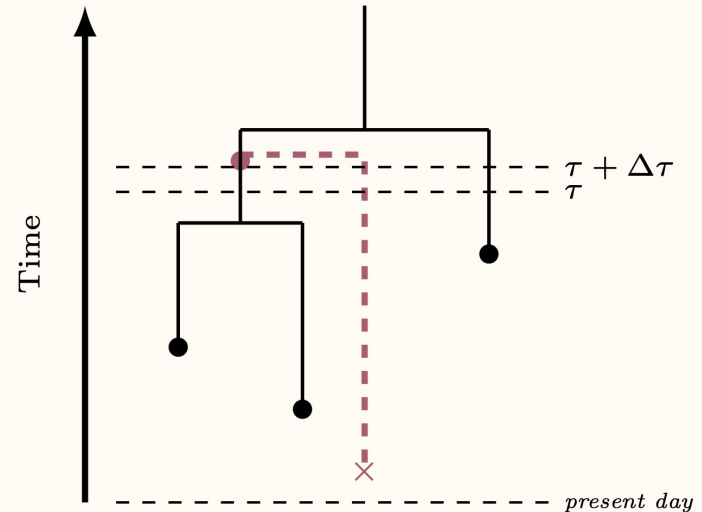
## Six-Step Framework

**Step 1:** Specify the Model

**Step 2:** Derive IVP for  $g_e(\tau)$

**Step 3:** Derive IVP for  $E(\tau)$

$E(\tau)$  is the probability that a lineage alive at time  $\tau$  leaves no sampled descendants between  $\tau$  and the present day.



$$\frac{dE}{d\tau} = \underbrace{-(\lambda + \mu + \psi)E(\tau)}_{\text{nothing happens}} + \underbrace{\lambda E(\tau)^2}_{\text{birth}} + \underbrace{\mu}_{\text{death}}$$

$$E(0) = 1 - \rho_0$$



# A General Birth-Death-Sampling Framework

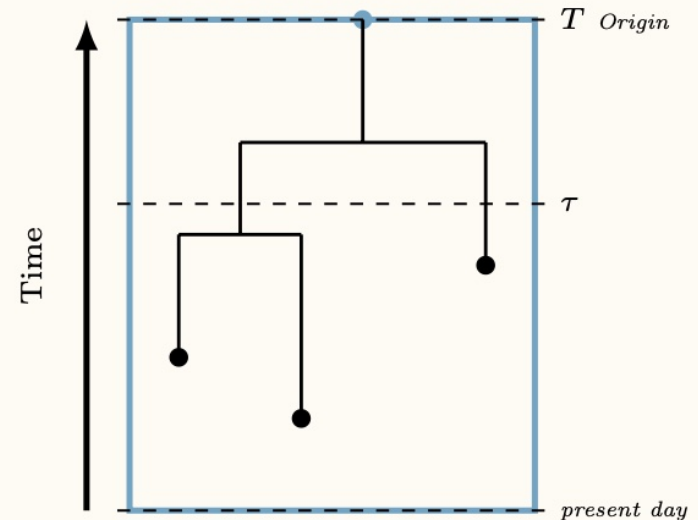
## Six-Step Framework

**Step 1:** Specify the Model

**Step 2:** Derive IVP for  $g_e(\tau)$

**Step 3:** Derive IVP for  $E(\tau)$

**Step 4:** Obtain likelihood ( $g_e(T)$ )



$$g_e(T) = \underbrace{\prod_i \lambda}_{\text{births}} \times \underbrace{\rho_0^J \prod_j \psi}_\text{sampling} \times \underbrace{\prod_e \Psi(s_e, t_e)}_\text{edges}$$

# A General Birth-Death-Sampling Framework

## Six-Step Framework

**Step 1:** Specify the Model

**Step 2:** Derive IVP for  $g_e(\tau)$

**Step 3:** Derive IVP for  $E(\tau)$

**Step 4:** Obtain likelihood ( $g_e(T)$ )

**Step 5:** Alternative Representation

ODE for  $g_e(\tau)$

$$\frac{dg_e(\tau)}{d\tau} = -(\lambda + \mu + \psi)g_e(\tau) + 2\lambda E(\tau)g_e(\tau)$$

# A General Birth-Death-Sampling Framework

## Six-Step Framework

**Step 1:** Specify the Model

**Step 2:** Derive IVP for  $g_e(\tau)$

**Step 3:** Derive IVP for  $E(\tau)$

**Step 4:** Obtain likelihood ( $g_e(T)$ )

**Step 5:** Alternative Representation

ODE for  $g_e(\tau)$

$$\frac{dg_e(\tau)}{d\tau} = -(\lambda + \mu + \psi)g_e(\tau) + 2\lambda E(\tau)g_e(\tau)$$

General Solution: Probability flow

$$g_e(\tau) = \underbrace{g_e(s_e)}_{\text{initial}} \exp \left[ \underbrace{\int_{s_e}^{\tau} -(\lambda + \mu + \psi) + 2\lambda E(x) dx}_{\text{Prob. flow } \Psi(s_e, \tau)} \right]$$

# A General Birth-Death-Sampling Framework

## Six-Step Framework

**Step 1:** Specify the Model

**Step 2:** Derive IVP for  $g_e(\tau)$

**Step 3:** Derive IVP for  $E(\tau)$

**Step 4:** Obtain likelihood ( $g_e(T)$ )

**Step 5:** Alternative Representation

ODE for  $g_e(\tau)$

$$\frac{dg_e(\tau)}{d\tau} = -(\lambda + \mu + \psi)g_e(\tau) + 2\lambda E(\tau)g_e(\tau)$$

General Solution: Probability flow

$$g_e(\tau) = \underbrace{g_e(s_e)}_{\text{initial}} \exp \left[ \underbrace{\int_{s_e}^{\tau} -(\lambda + \mu + \psi) + 2\lambda E(x) dx}_{\text{Prob. flow } \Psi(s_e, \tau)} \right]$$

Define  $\Phi$

$$\Psi(s_e, t_e) = \frac{\Phi(t_e)}{\Phi(s_e)}$$

# A General Birth-Death-Sampling Framework

## Six-Step Framework

**Step 1:** Specify the Model

**Step 2:** Derive IVP for  $g_e(\tau)$

**Step 3:** Derive IVP for  $E(\tau)$

**Step 4:** Obtain likelihood ( $g_e(T)$ )

**Step 5:** Alternative Representation

ODE for  $g_e(\tau)$

$$\frac{dg_e(\tau)}{d\tau} = -(\lambda + \mu + \psi)g_e(\tau) + 2\lambda E(\tau)g_e(\tau)$$

General Solution: Probability flow

$$g_e(\tau) = \underbrace{g_e(s_e)}_{\text{initial}} \underbrace{\exp \left[ \int_{s_e}^{\tau} -(\lambda + \mu + \psi) + 2\lambda E(x) dx \right]}_{\text{Prob. flow } \Psi(s_e, \tau)}$$

Define  $\Phi$

$$\Psi(s_e, t_e) = \frac{\Phi(t_e)}{\Phi(s_e)}$$

Reparameterize

$$g_e(T) = \underbrace{\Phi(T)}_{\text{origin}} \times \underbrace{\prod_i \lambda \Phi(x_i)}_{\text{births}} \times \underbrace{\rho_0^J \prod_j \frac{\psi}{\Phi(y_j)}}_{\text{sampling}}$$

# A General Birth-Death-Sampling Framework

## Six-Step Framework

**Step 1:** Specify the Model

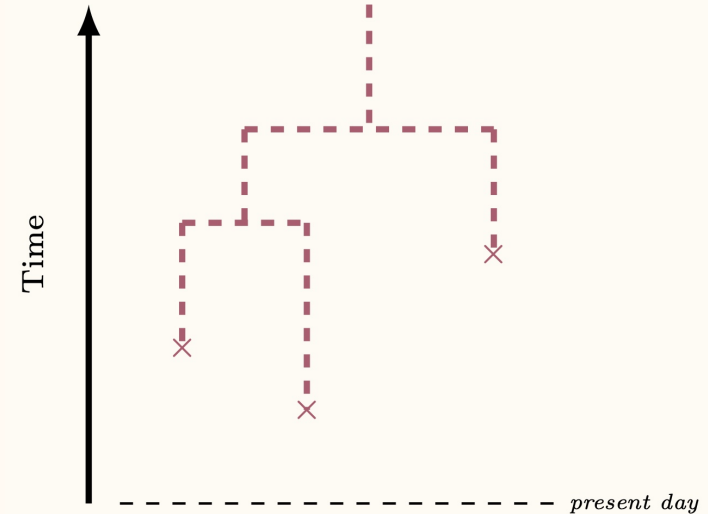
**Step 2:** Derive IVP for  $g_e(\tau)$

**Step 3:** Derive IVP for  $E(\tau)$

**Step 4:** Obtain likelihood ( $g_e(T)$ )

**Step 5:** Alternative Representation

**Step 6:** Condition the likelihood



$$\mathcal{L} = \underbrace{\mathcal{S}}_{\text{cond.}} \times \underbrace{\Phi(T)}_{\text{origin}} \times \underbrace{\prod_i \lambda \Phi(x_i)}_{\text{births}} \times \underbrace{\rho_0^J \prod_j \frac{\psi}{\Phi(y_j)}}_{\text{sampling}}$$

**Aim:** Develop a general birth-death sampling model that allows rates to vary through time.

*Why develop a general model?*

- I. **Clarifies assumptions of and connections between existing models**
- II. Allows us to explicitly fit an SIR epidemiological model
- III. Enables mathematical analysis of the general model class

# Model Connections

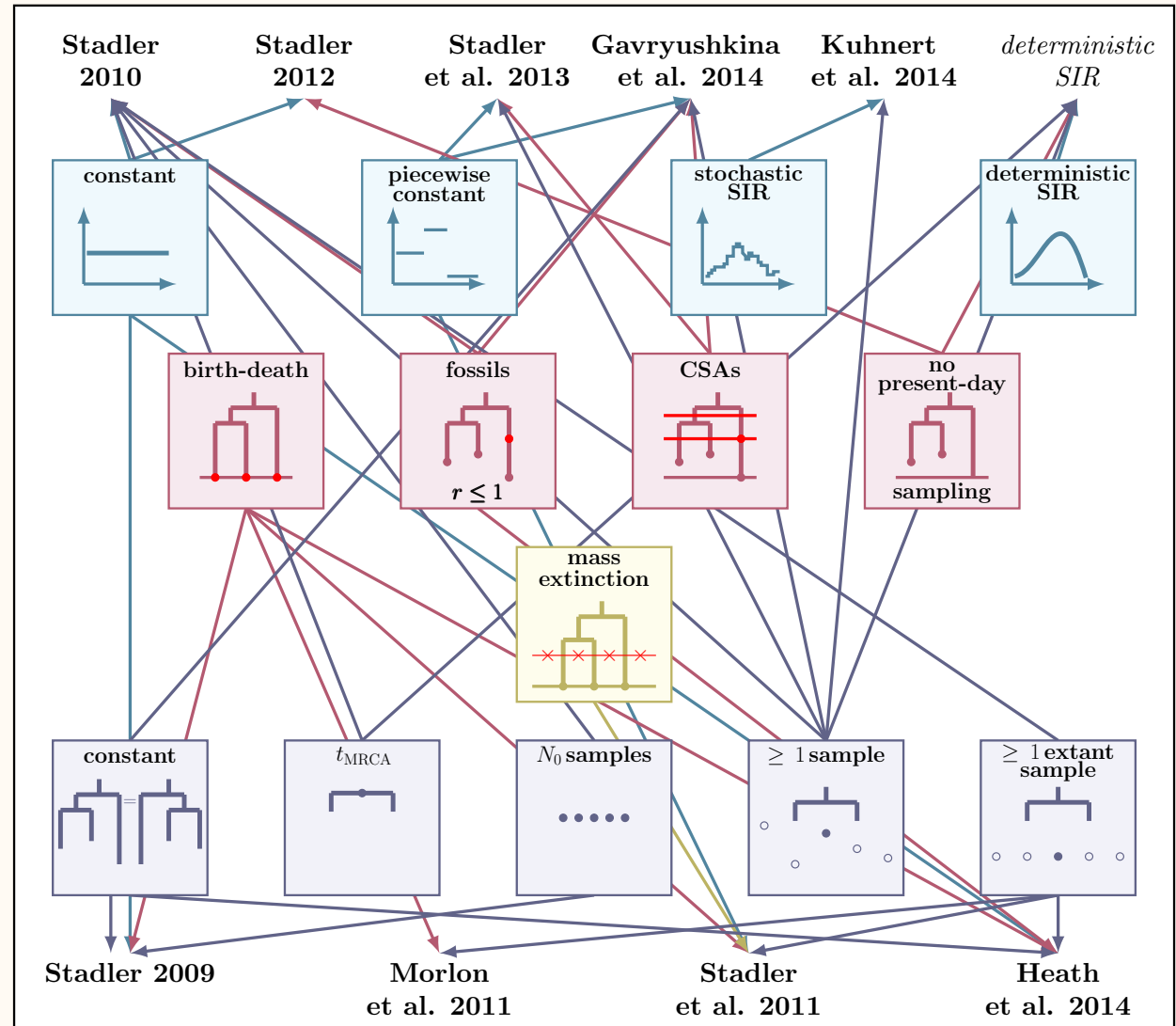
## Assumptions:

rates

sampling

mass extinction

conditioning





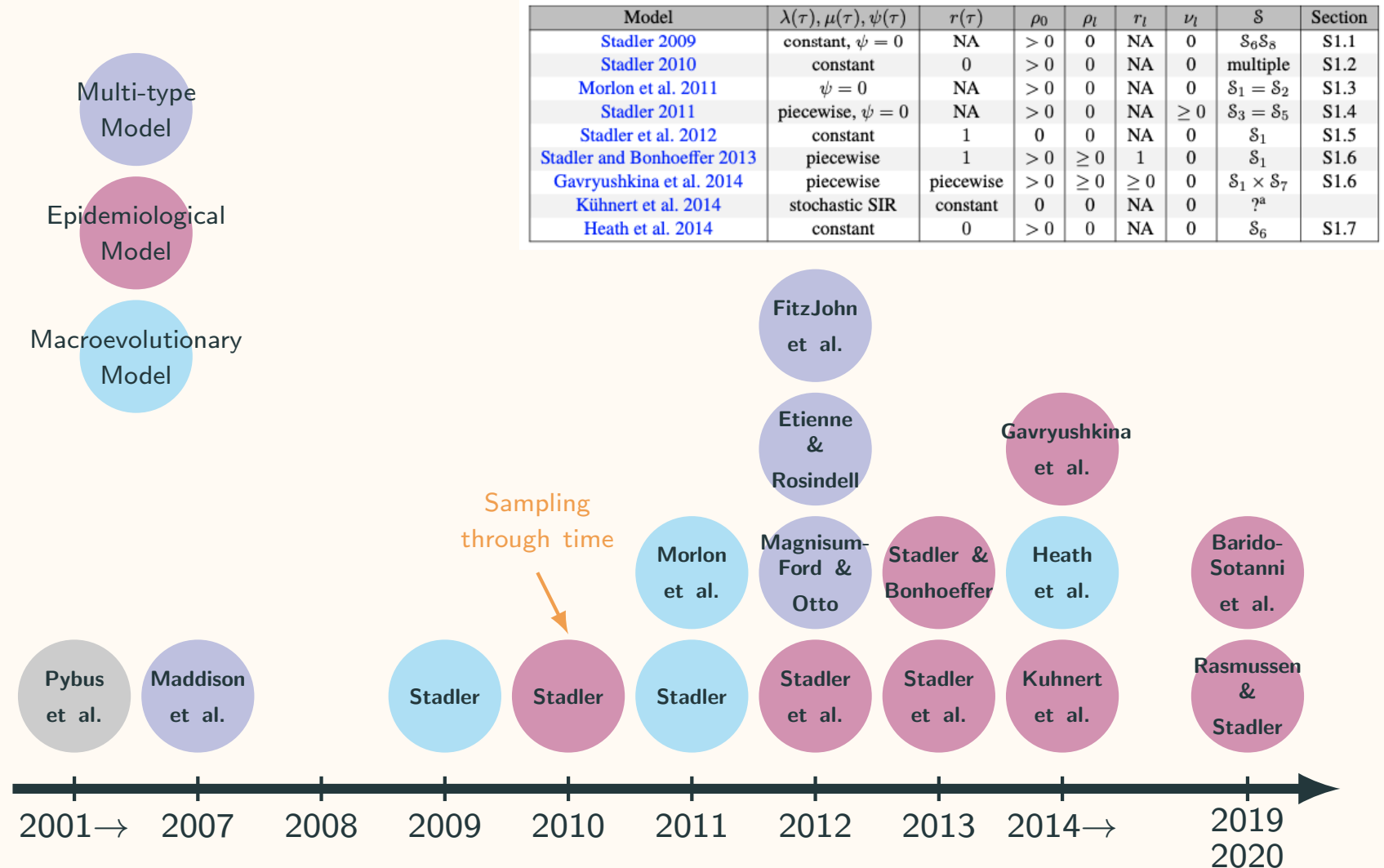
# A Note about conditioning

## Existing Conditioning Regimes

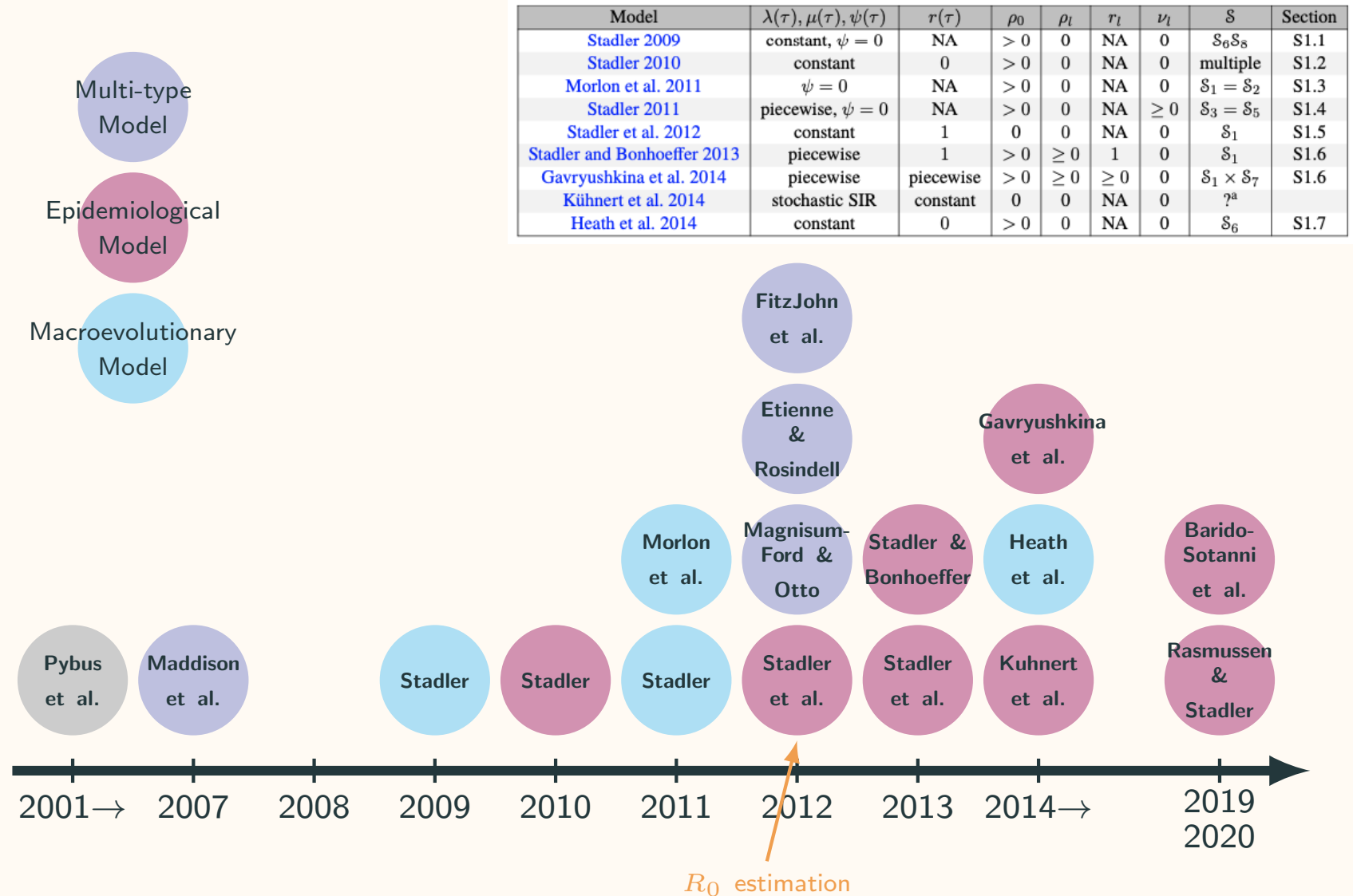
| Condition                                                                                                                          | Description                                                                                   | Examples                                                                                                                                                    |
|------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------|
| $\mathcal{S}_0 = 1$                                                                                                                | No conditioning                                                                               | <b>Eq.3</b> ( <a href="#">Stadler, 2010</a> ),<br><b>Eq.1</b> ( <a href="#">Stadler et al., 2012</a> ) & <b>Thrm. 2.6</b> ( <a href="#">Stadler, 2011</a> ) |
| $\mathcal{S}_1 = \frac{\Phi(x_1)}{\Phi(T)}$                                                                                        | Likelihood given the $t_{\text{MRCA}}$ rather than on the time of origin.                     |                                                                                                                                                             |
| $\mathcal{S}_2 = \frac{1}{1-E(T)}$                                                                                                 | At least one sampled descendent either at or before the present day (given $T$ ).             | <b>Eq.2</b> <a href="#">Morlon et al. (2011)</a> ,<br><b>Thrm.1</b> <a href="#">Stadler and Bonhoeffer (2013)</a>                                           |
| $\mathcal{S}_3 = \frac{\Phi(x_1)}{\Phi(T)(1-E(x_1))^2}$                                                                            | At least one sampled descendent at or before the present day given $\tau_{\text{MRCA}} = x_1$ | ( <a href="#">Heath et al., 2014</a> )                                                                                                                      |
| $\mathcal{S}_4 = \frac{1}{1-\hat{E}(T)}$                                                                                           | At least one <i>extant</i> sampled lineage (given $T$ ).                                      | <b>Cor.3.7</b> ( <a href="#">Stadler, 2010</a> ) & <b>Thrm.2.7</b> ( <a href="#">Stadler, 2011</a> )                                                        |
| $\mathcal{S}_5 = \frac{\Phi(x_1)}{\Phi(T)2(1-\hat{E}(x_1))(1-E(x_1))}$                                                             | At least one <i>extant</i> sampled lineage given $\tau_{\text{MRCA}} = x_1$                   |                                                                                                                                                             |
| $\mathcal{S}'_5 = \frac{\Phi(x_1)}{\Phi(T)(1-\hat{E}(x_1))^2}$                                                                     | Both daughters of MRCA have at least one <i>extant</i> sampled lineage                        | <b>Eq. 5</b> <a href="#">Stadler (2010)</a>                                                                                                                 |
| $\mathcal{S}_6 = \frac{1}{\hat{E}_{N_0}(T)}$                                                                                       | Exactly $N_0$ <i>extant</i> sampled lineages (given $T$ ).                                    | <b>Eq.2</b> <a href="#">Stadler (2009)</a> ,<br><b>Eq.4</b> ( <a href="#">Stadler, 2010</a> )<br>& <b>Cor.3.6</b> <a href="#">Stadler (2010)</a>            |
| $\mathcal{S}_7 = \frac{\Phi(x_1)}{\Phi(t_{\text{or}})} \left( \sum_{i=1}^{N_0-1} \hat{E}_i(x_1) \hat{E}_{N_0-i}(x_1) \right)^{-1}$ | Exactly $N_0$ <i>extant</i> sampled lineages given $t_{\text{MRCA}}$ .                        | <b>Eq.6</b> <a href="#">Stadler (2010)</a>                                                                                                                  |
| $\mathcal{S}_8 = \text{constant}$                                                                                                  | Multiply by a constant                                                                        | <a href="#">Stadler (2009)</a> & <a href="#">Gavryushkina et al. (2014)</a> <sup>a</sup> & <a href="#">Heath et al. (2014)</a>                              |

There does not exist a natural conditioning for  $\geq n_{\min}$  lineages

# Historical Developments Revisited

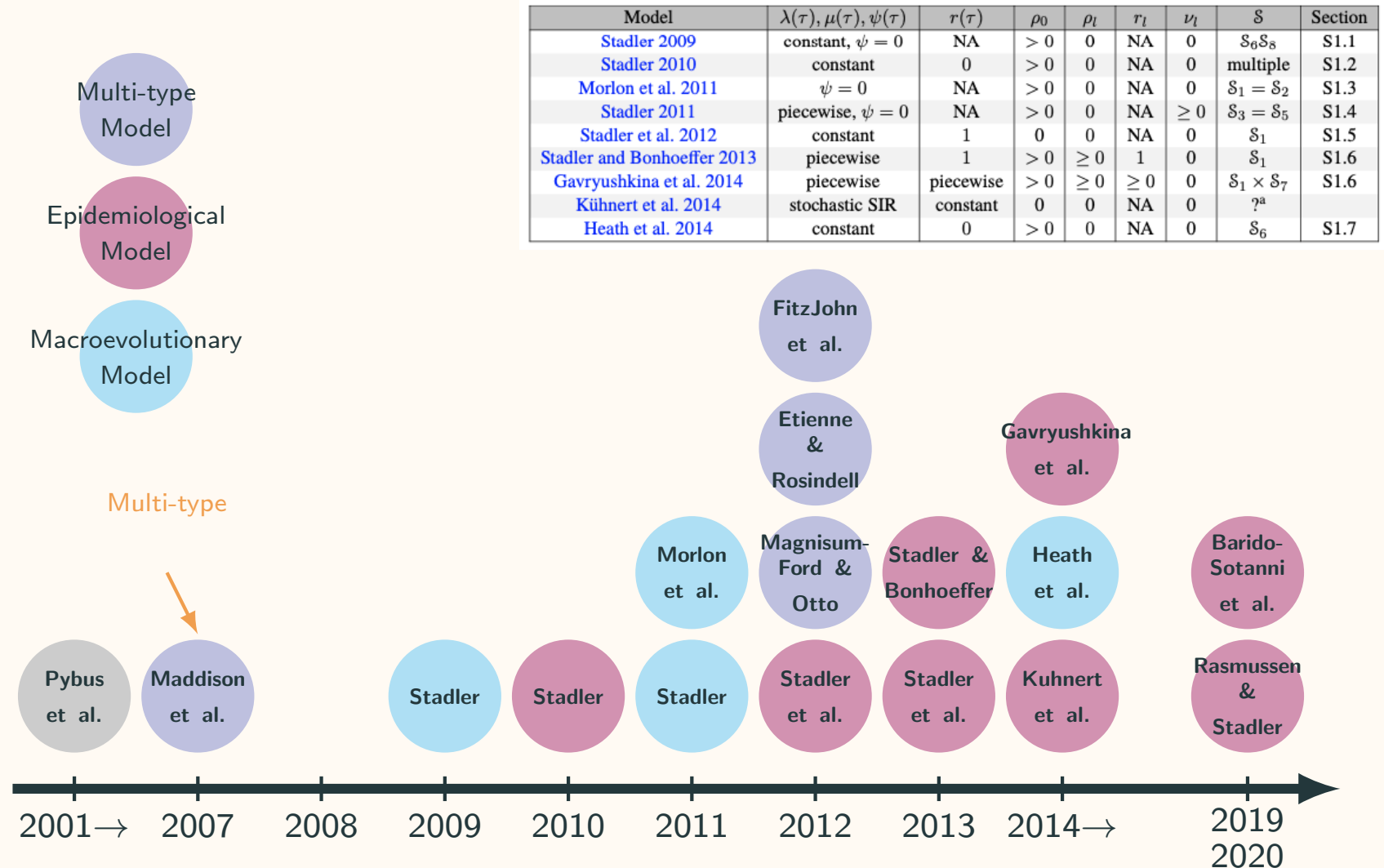


# Historical Developments Revisited

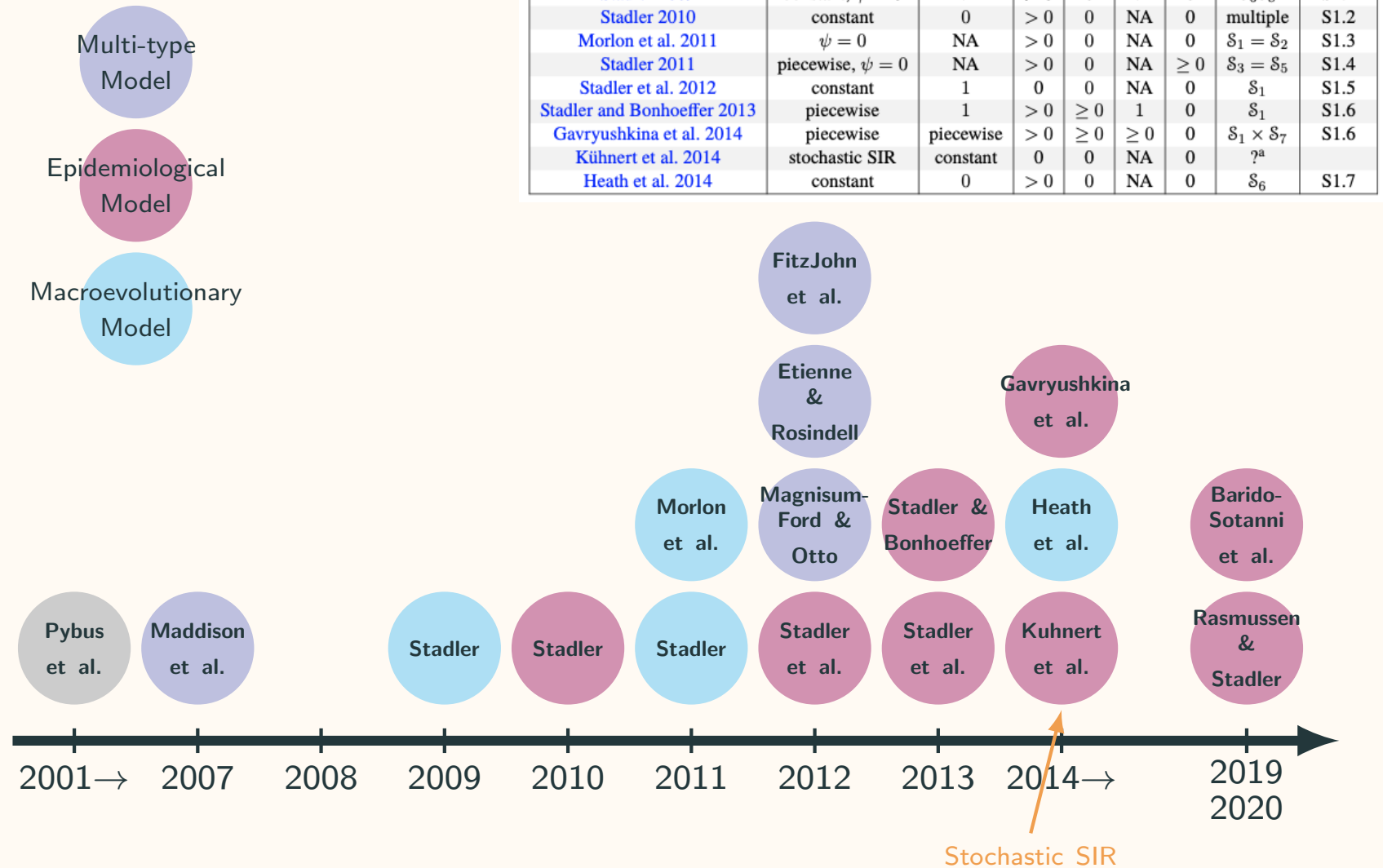


| Model                       | $\lambda(\tau), \mu(\tau), \psi(\tau)$ | $r(\tau)$ | $\rho_0$ | $\rho_l$ | $r_l$    | $\nu_l$  | $\mathcal{S}$                        | Section |
|-----------------------------|----------------------------------------|-----------|----------|----------|----------|----------|--------------------------------------|---------|
| Stadler 2009                | constant, $\psi = 0$                   | NA        | $> 0$    | 0        | NA       | 0        | $\mathcal{S}_6 \mathcal{S}_8$        | S1.1    |
| Stadler 2010                | constant                               | 0         | $> 0$    | 0        | NA       | 0        | multiple                             | S1.2    |
| Morlon et al. 2011          | $\psi = 0$                             | NA        | $> 0$    | 0        | NA       | 0        | $\mathcal{S}_1 = \mathcal{S}_2$      | S1.3    |
| Stadler 2011                | piecewise, $\psi = 0$                  | NA        | $> 0$    | 0        | NA       | $\geq 0$ | $\mathcal{S}_3 = \mathcal{S}_5$      | S1.4    |
| Stadler et al. 2012         | constant                               | 1         | 0        | 0        | NA       | 0        | $\mathcal{S}_1$                      | S1.5    |
| Stadler and Bonhoeffer 2013 | piecewise                              | 1         | $> 0$    | $\geq 0$ | 1        | 0        | $\mathcal{S}_1$                      | S1.6    |
| Gavryushkina et al. 2014    | piecewise                              | piecewise | $> 0$    | $\geq 0$ | $\geq 0$ | 0        | $\mathcal{S}_1 \times \mathcal{S}_7$ | S1.6    |
| Kühnert et al. 2014         | stochastic SIR                         | constant  | 0        | 0        | NA       | 0        | ? <sup>a</sup>                       |         |
| Heath et al. 2014           | constant                               | 0         | $> 0$    | 0        | NA       | 0        | $\mathcal{S}_6$                      | S1.7    |

# Historical Developments Revisited



# Historical Developments Revisited

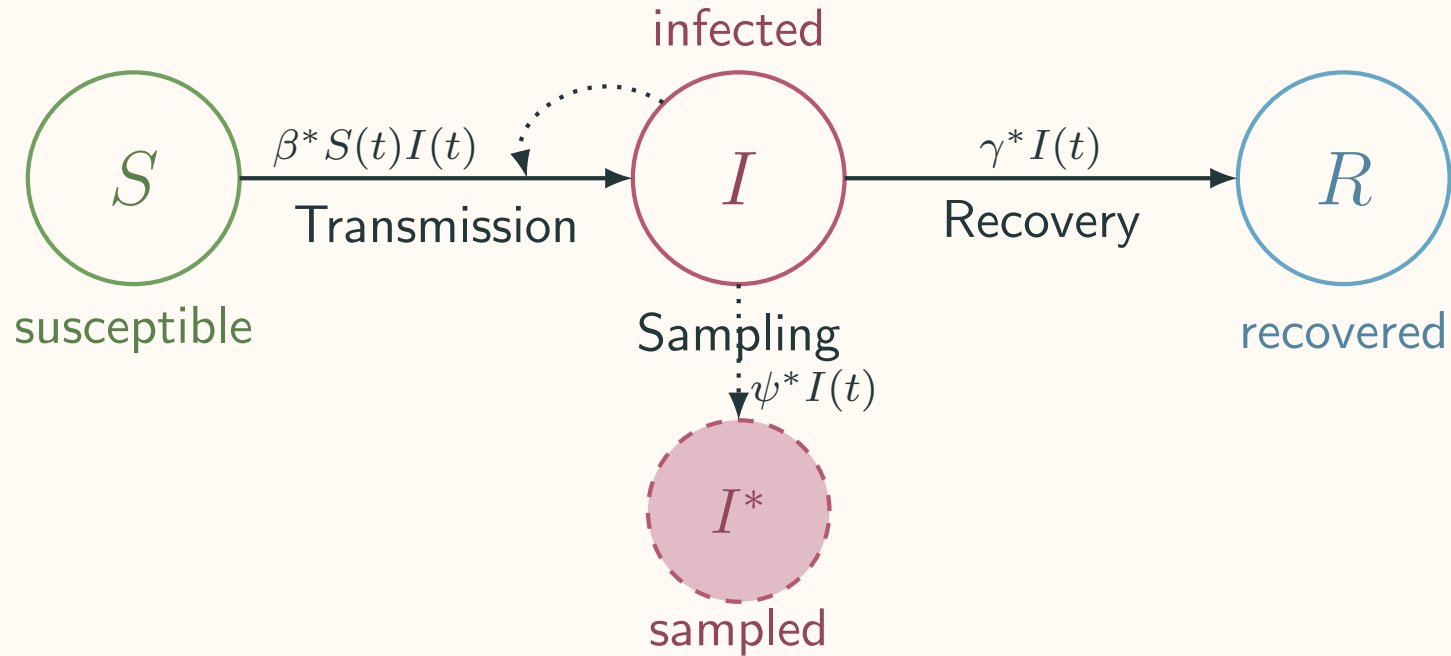


**Aim:** Develop a general birth-death sampling model that allows rates to vary through time.

*Why develop a general model?*

- I. Clarifies assumptions of and connections between existing models
- II. Allows us to explicitly fit an SIR epidemiological model**
- III. Enables mathematical analysis of the general model class

# The SIR Model Revisited



$$\begin{aligned}
 \mathcal{L}(\vec{x}, \vec{y}, \vec{z}, N_0 | \Theta_{SIR}, \mathcal{S}) &= \mathcal{S} \rho_0^{N_0} \Phi(T) \prod_{i=1}^I \lambda(x_i) \Phi(x_i) \\
 &\times \prod_{j=1}^n \frac{\psi(y_j)}{\Phi(y_j)} [(1 - r(y_j))E(y_j) + r(y_j)] \prod_{k=1}^m \psi(z_k) (1 - r(z_k)) \\
 \lambda(\tau) &= \beta^* S(T - \tau), \quad \mu(\tau) = \gamma^*, \quad \psi(\tau) = \psi^*
 \end{aligned}$$

**Aim:** Develop a general birth-death sampling model that allows rates to vary through time.

*Why develop a general model?*

- I. Clarifies assumptions of and connections between existing models
- II. Allows us to explicitly fit an SIR epidemiological model
- III. Enables mathematical analysis of the general model class**



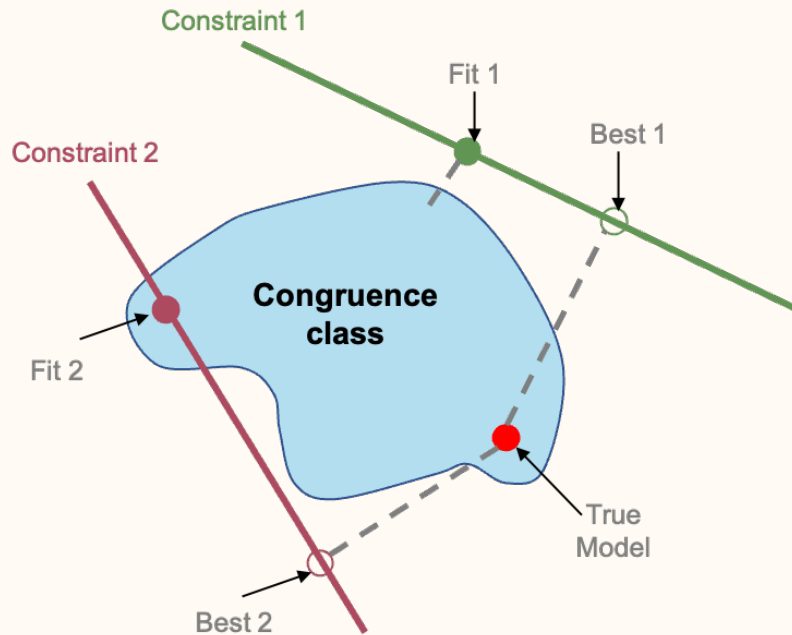
# Parameter Identifiability

---

# Parameter Identifiability in Macroevolution

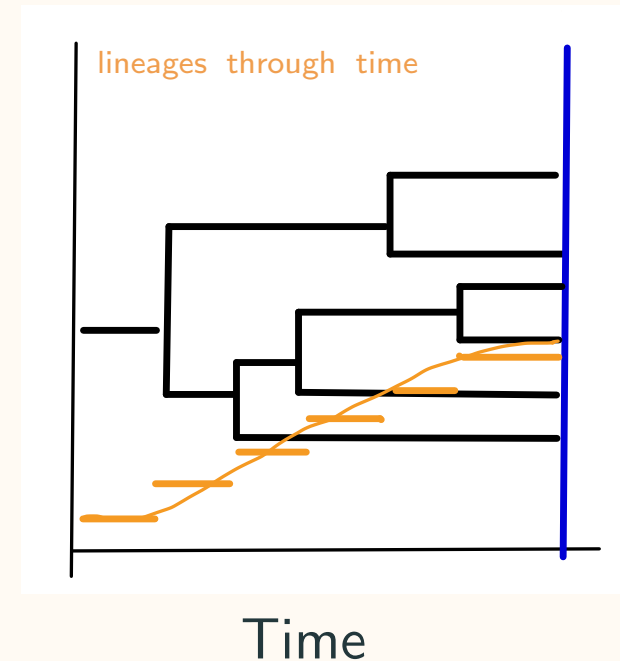
Original Likelihood:

$$\mathcal{L}(\mathcal{T}|\lambda, \mu)$$



Reparameterization:

$$\mathcal{L}(\mathcal{T}|\tilde{\lambda})$$



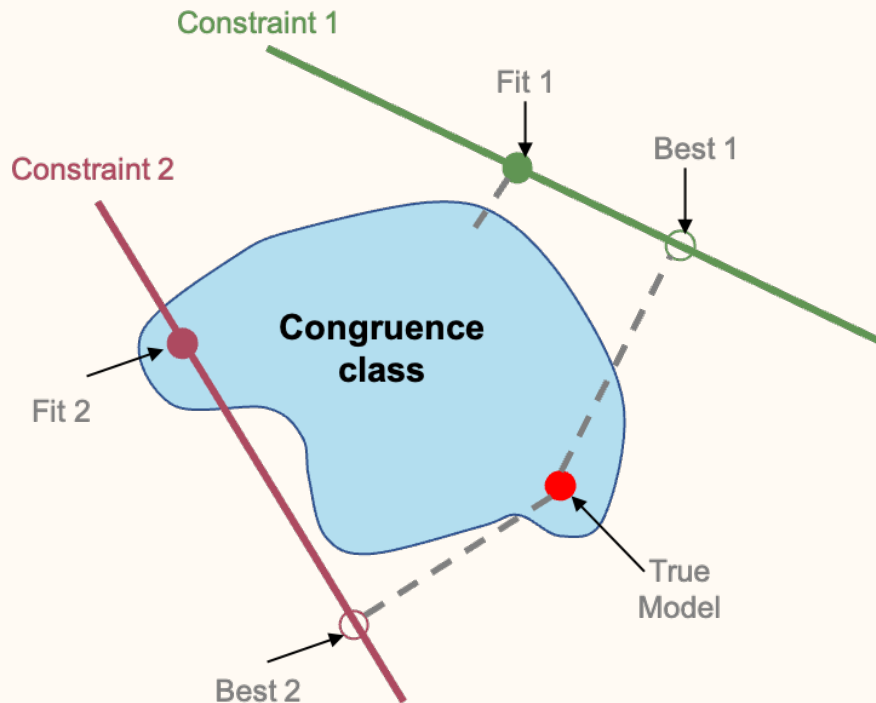
## Macroevolutionary Congruence Class:

Models  $\{\lambda(\tau), \mu(\tau)\}$  that have identical likelihoods and hence are indistinguishable given the data.

# The Epidemiological Congruence Classes

Original Likelihood:

$$\mathcal{L}(\mathcal{T}|\lambda, \mu, \psi)$$

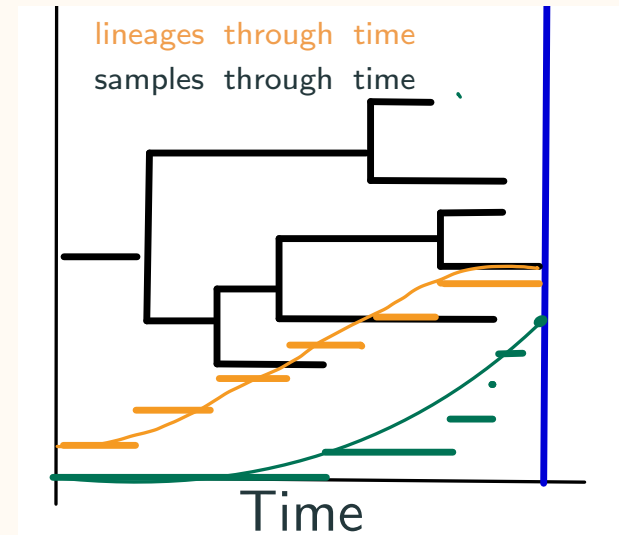


Reparameterization:

$$\mathcal{L}(\mathcal{T}|\tilde{\lambda}, \tilde{\psi})$$

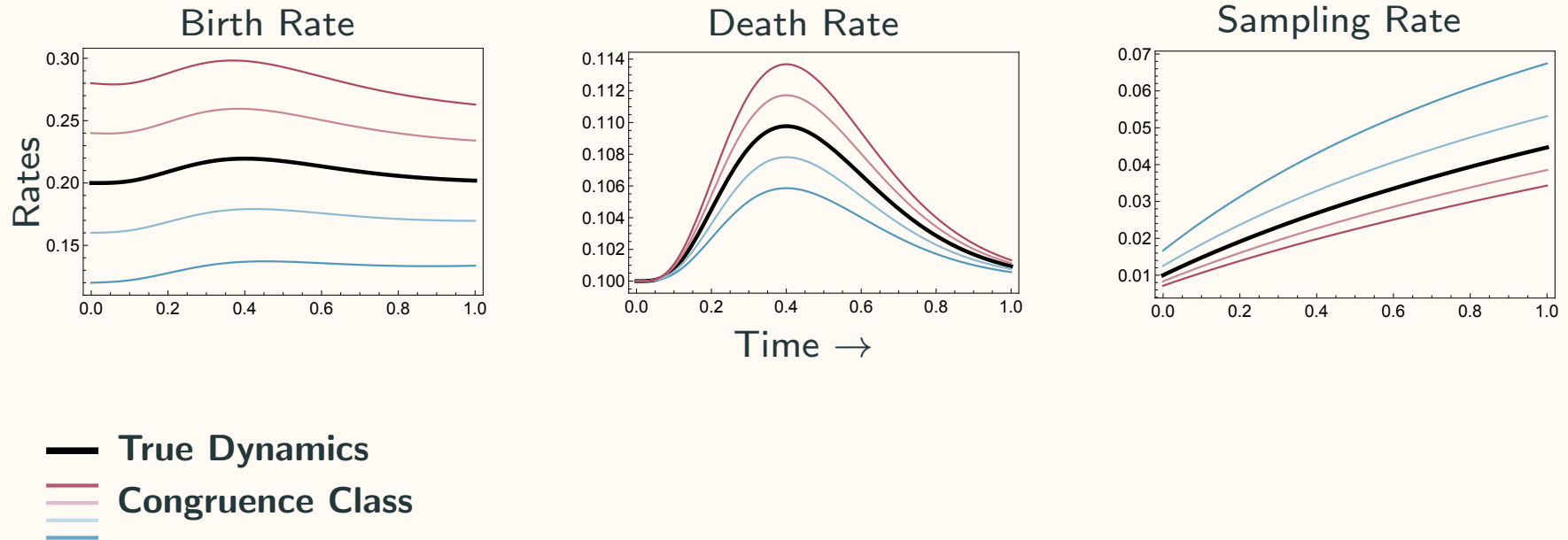
$$\tilde{\lambda}(t) = \lambda(t)(1 - E(t))$$

$$\tilde{\psi}(t) = \frac{\lambda(t)}{(1 - E(t))}$$



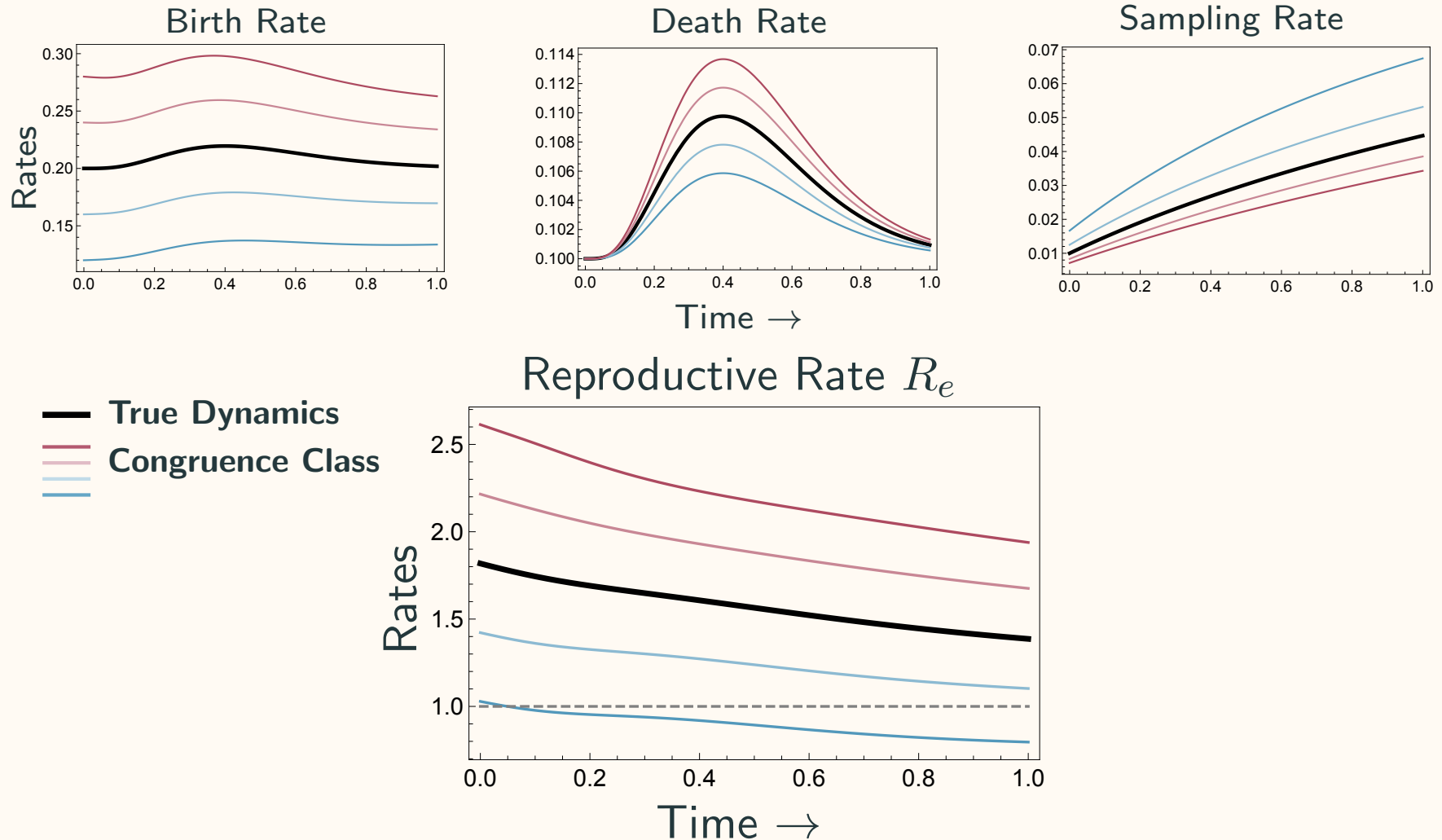
Louca, McLaughlin, MacPherson, Joy, and Pennell (Molecular Biology and Evolution)

# Impacts on Inference 1: The effective reproductive number



*Louca, McLaughlin, MacPherson, Joy, and Pennell (Molecular Biology and Evolution)*

# Impacts on Inference 1: The effective reproductive number



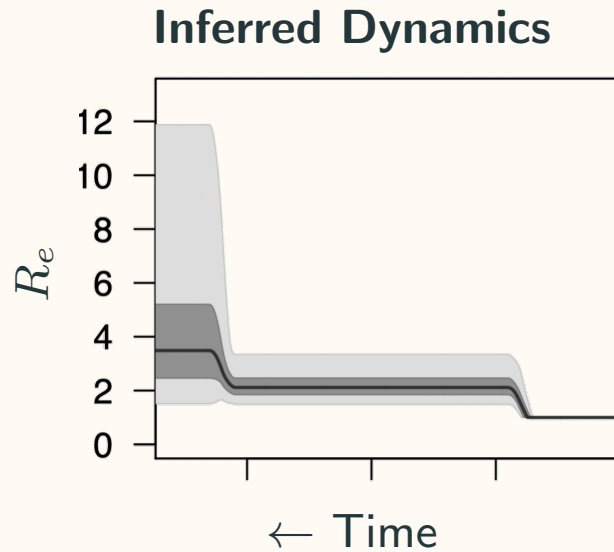
Effective Reproductive Number  $R_e(t)$ : The rate of disease spread at time  $t$

# Impacts on Inference 3: HIV in Alberta Canada

563 HIV Sequences

Bayesian birth-death skyline (in BEAST)

Constrains functions to be piecewise constant



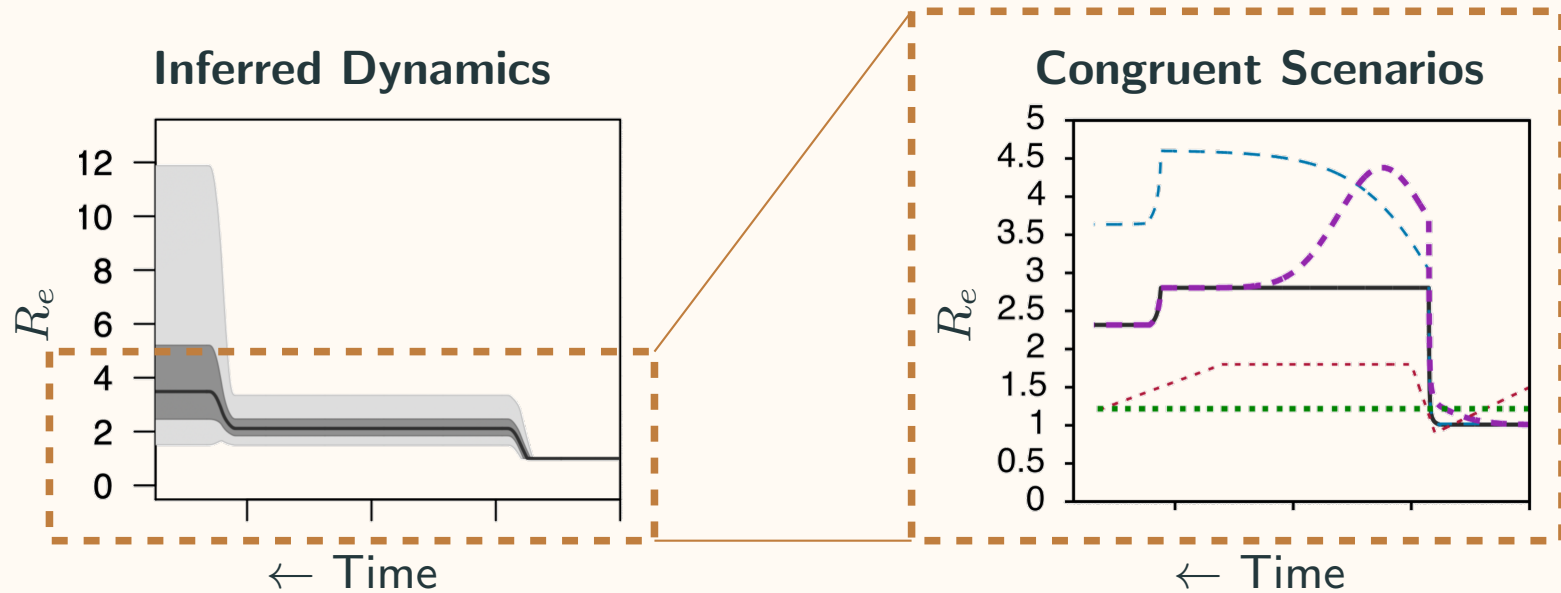
*Louca, McLaughlin, MacPherson, Joy, and Pennell (Molecular Biology and Evolution)*

# Impacts on Inference 3: HIV in Alberta Canada

563 HIV Sequences

Bayesian birth-death skyline (in BEAST)

Constrains functions to be piecewise constant



*Louca, McLaughlin, MacPherson, Joy, and Pennell (Molecular Biology and Evolution)*

**Aim:** Develop a general birth-death sampling model that allows rates to vary through time.  
General model derivation with 6-step framework

*Why develop a general model?*

- I. Clarifies assumptions of and connections between existing models  
General model identifies for key model assumptions
- II. Allows us to explicitly fit an SIR epidemiological model  
General model provides straightforward parameterization of the general model
- III. Enables mathematical analysis of the general model class  
Phylogenetic inference is in general limited by parameter identifiability

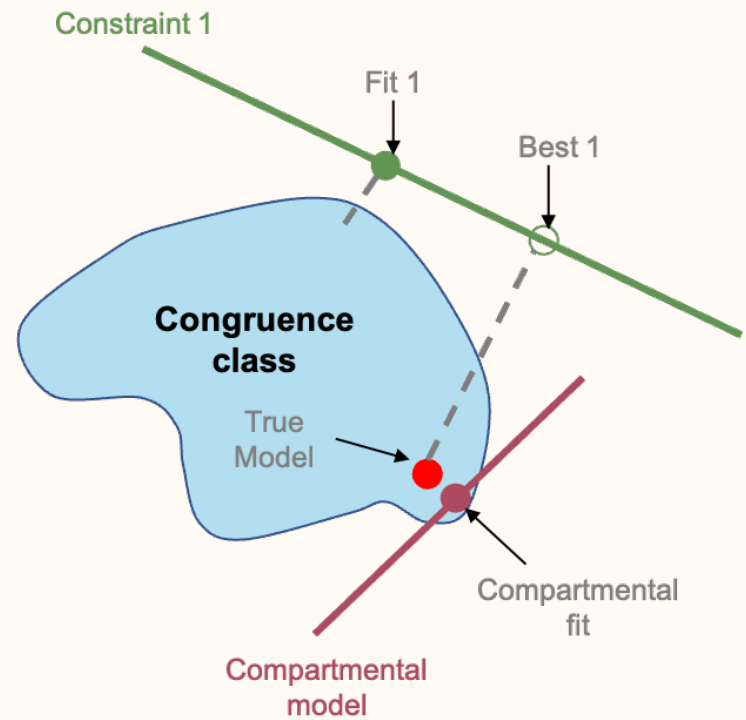


# Ways forward

---

# Ways Forward

- I. Smoothing
- II. Very strict priors
- III. Concerted Sampling Attempts
- IV. **Mechanistic models**

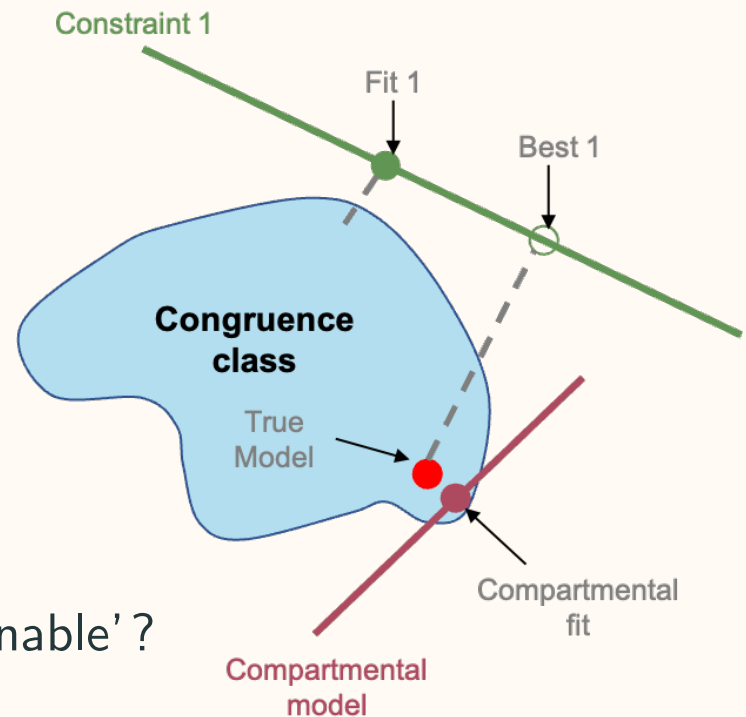


# Ways Forward

- I. Smoothing
- II. Very strict priors
- III. Concerted Sampling Attempts
- IV. **Mechanistic models**

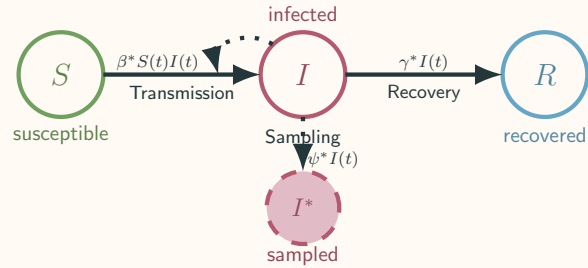
What if we constrain the rates to be 'reasonable'?

The subset with a plausible **mechanism**.



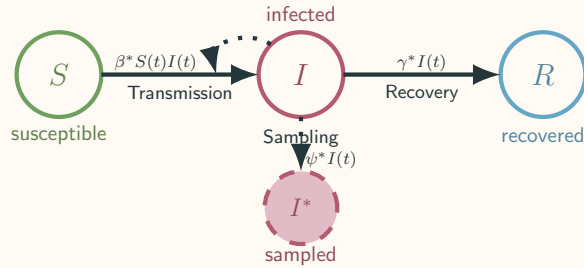
# Bayesian Inference in an SIR Model (Ongoing)

## Epidemiological Mechanism

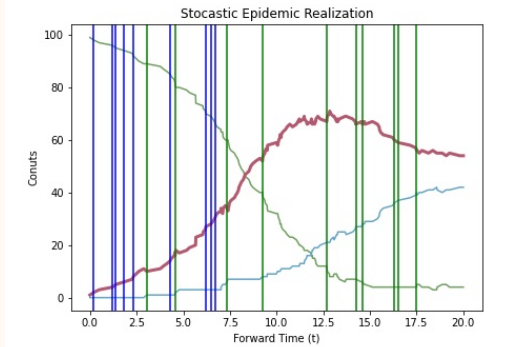


# Bayesian Inference in an SIR Model (Ongoing)

## Epidemiological Mechanism



## Epidemic with Sampling

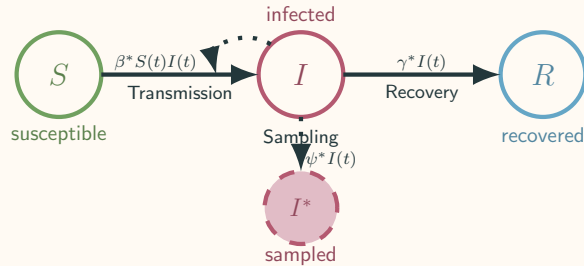


Birth:  $x_i$

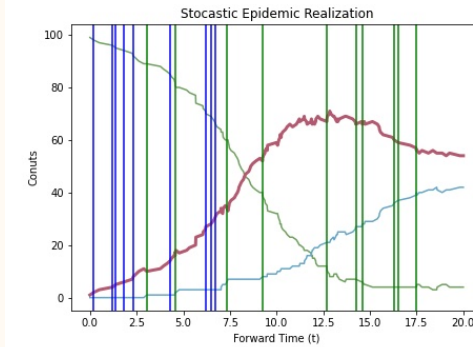
Sampling:  $y_j$

# Bayesian Inference in an SIR Model (Ongoing)

## Epidemiological Mechanism



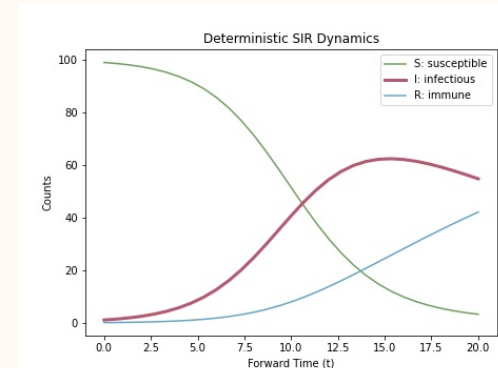
## Epidemic with Sampling



Birth:  $x_i$

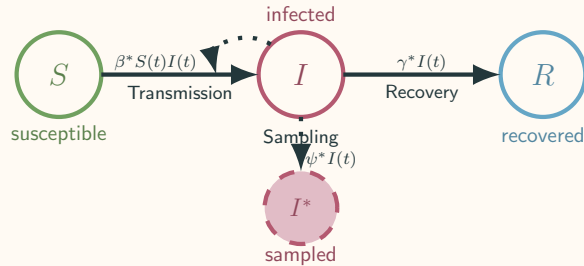
Sampling:  $y_j$

**Deterministic Dynamics**

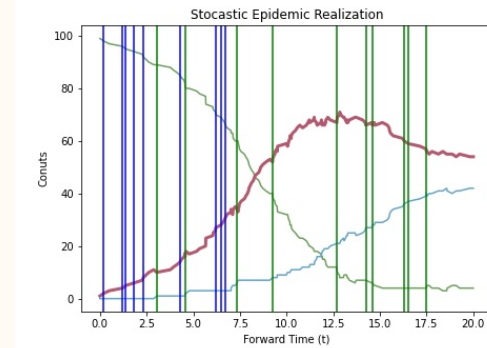


# Bayesian Inference in an SIR Model (Ongoing)

## Epidemiological Mechanism



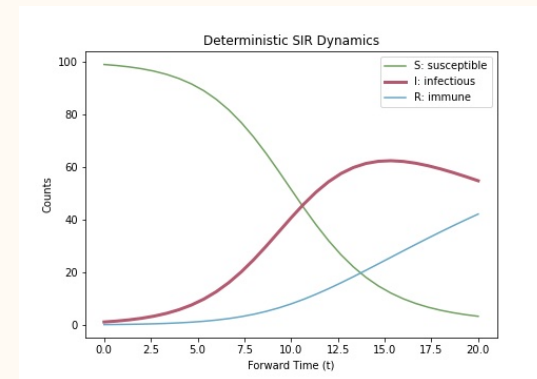
## Epidemic with Sampling



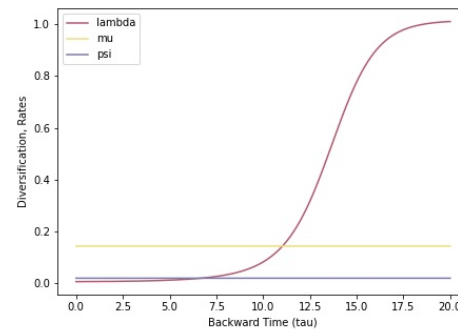
Birth:  $x_i$

Sampling:  $y_j$

**Deterministic Dynamics**

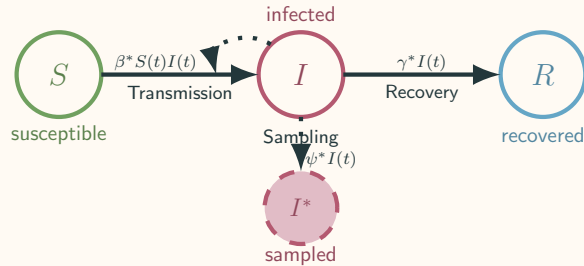


## Diversification Model Likelihood

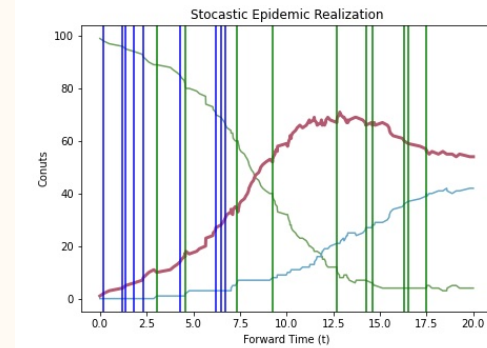


# Bayesian Inference in an SIR Model (Ongoing)

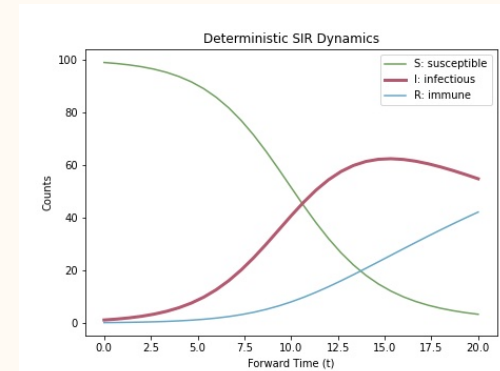
## Epidemiological Mechanism



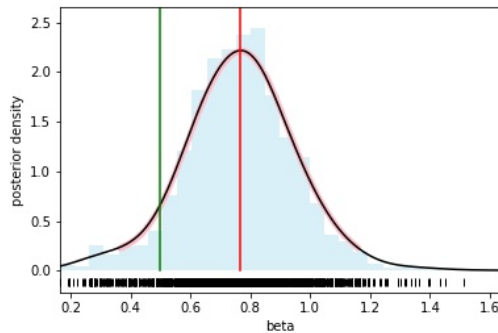
## Epidemic with Sampling

Birth:  $x_i$ Sampling:  $y_j$ 

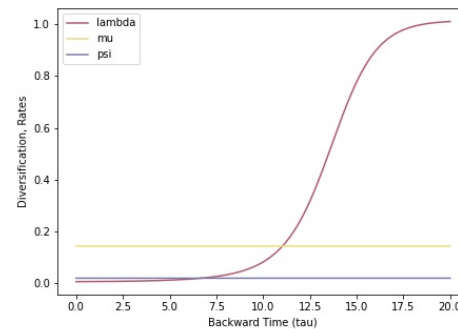
## Deterministic Dynamics



## Posterior Estimation

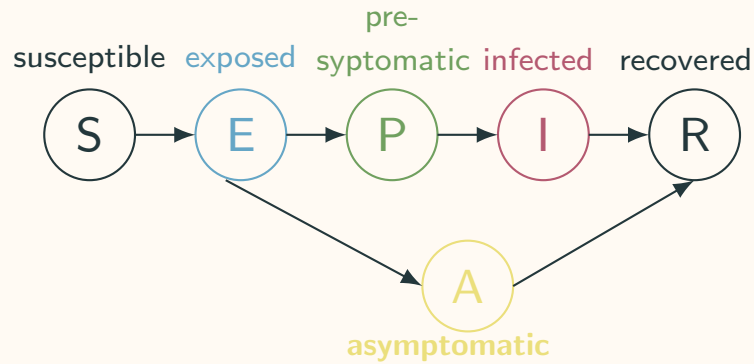


## Diversification Model Likelihood



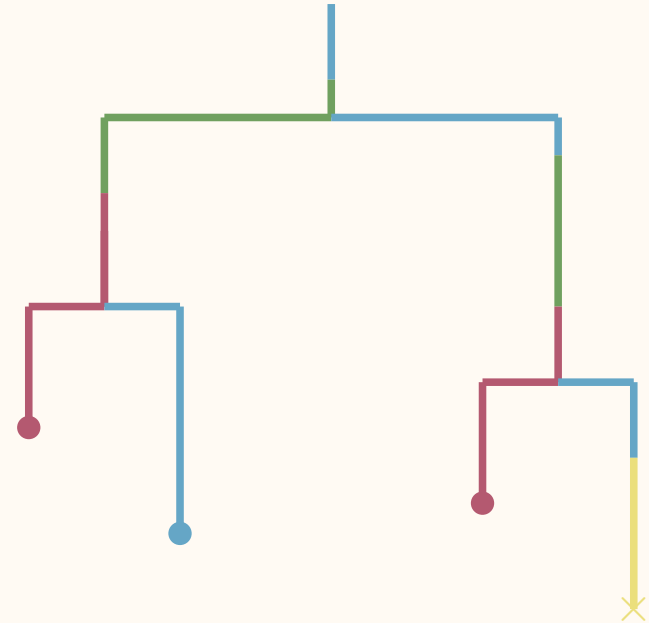


# Future Directions: Fitting complex epidemiological models



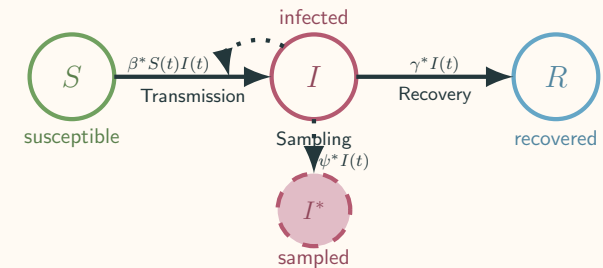
—adapted from Day et al. 2020

## Multi-type Tree

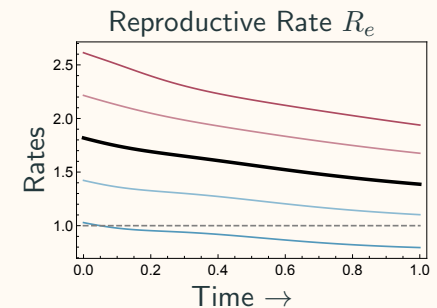


# Conclusions

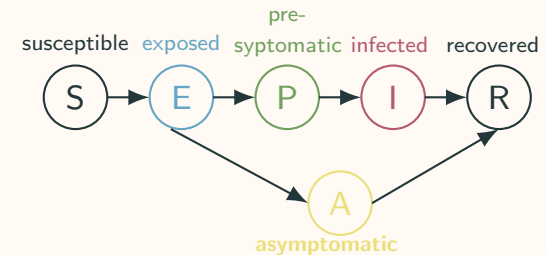
I. We derived a general phylodynamic model



II. We characterized parameter (un)-identifiability



III. Can epidemiological mechanism resolve unidentifiability?



—adapted from Day et al. 2020

Many Thanks To:

**Coauthors:** Matt Pennell (USC), Angela Mclaughlin and Jeff Joy (UBC), Stilian Louca (UofO)

**Funding:** Univ. of Toronto, NSERC, CRC