

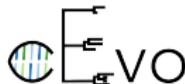
# Computational Biology

Lecturers:  
Tanja Stadler, Carsten Magnus & Tim Vaughan

Teaching Assistants:  
Jūlija Pečerska, Jérémie Sciré,  
Sarah Nadeau & Marc Manceau

Computational Evolution  
Department of Biosystems Science and Engineering

HS 2019



- Structured Populations
  - Population structure
  - Spatial structure
  - Non-spatial structure
  - Structure in phylodynamic analyses
  - Structured birth-death models
  - Structured coalescent models
  - Structural dynamics
- Phyldynamics in Action
- References

# Bayesian Inference: Questions

- ② Does a Bayesian phylogenetic analysis of the kind described here allow one to directly infer ancestral sequences?  
Why/Why not?

## Structured Populations

Population structure

Spatial structure

Non-spatial structure

Structure in phylodynamic analyses

Structured birth-death models

Structured coalescent models

Structural dynamics

## Phylodynamics in Action

## References

# Bayesian Inference: Questions

- ② Does a Bayesian phylogenetic analysis of the kind described here allow one to directly infer ancestral sequences?  
Why/Why not?
- ② How might we test to see whether a Bayesian MCMC analysis has explored the full state space supported by the posterior?

## Structured Populations

- Population structure
- Spatial structure
- Non-spatial structure
- Structure in phylodynamic analyses
- Structured birth-death models
- Structured coalescent models
- Structural dynamics

## Phyldynamics in Action

## References

# Bayesian Inference: Questions

- ② Does a Bayesian phylogenetic analysis of the kind described here allow one to directly infer ancestral sequences?  
Why/Why not?
- ② How might we test to see whether a Bayesian MCMC analysis has explored the full state space supported by the posterior?
- ② Suppose you have conducted a Bayesian phylodynamic analysis and recovered a 95% HPD interval for the birth rate parameter. If you take this result and use it to construct a new prior for this parameter and use this prior to analyze the *same data*, would the resulting second posterior be valid?

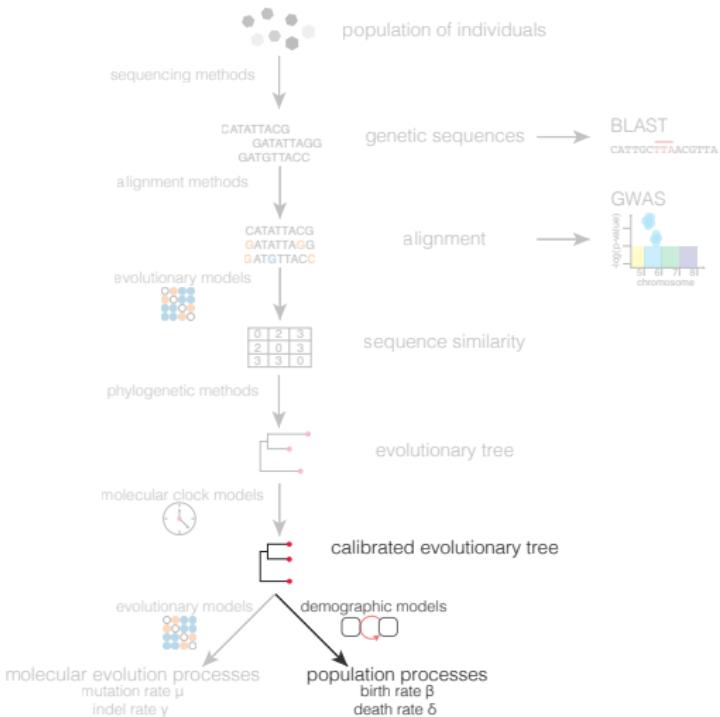
## Structured Populations

- Population structure
- Spatial structure
- Non-spatial structure
- Structure in phylodynamic analyses
- Structured birth-death models
- Structured coalescent models
- Structural dynamics

## Phylodynamics in Action

## References

# Overview



## Structured Populations

- Population structure
- Spatial structure
- Non-spatial structure
- Structure in phylodynamic analyses
- Structured birth-death models
- Structured coalescent models
- Structural dynamics

## Phylodynamics in Action

### References

# Structured Populations

Biological populations often have some internal “structure”:

## Structured Populations

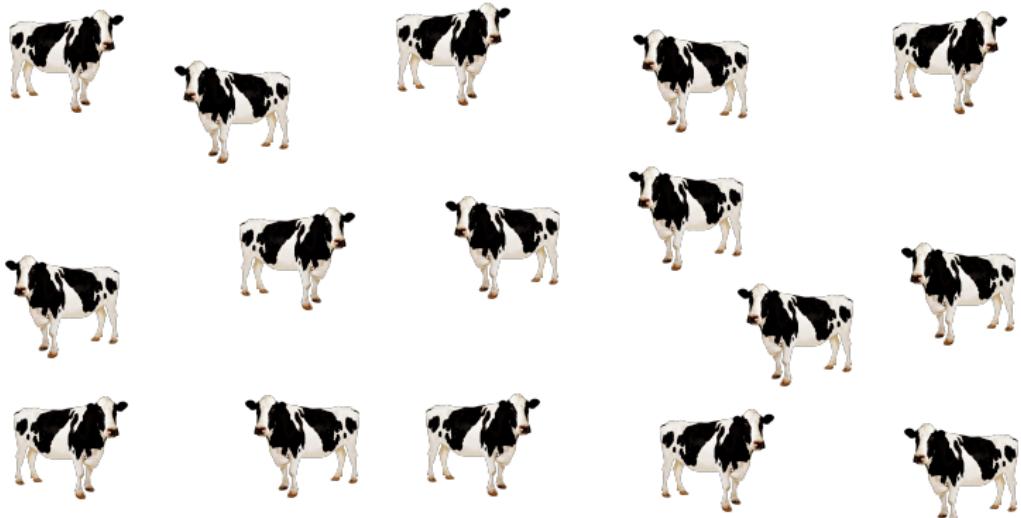
- Population structure
- Spatial structure
- Non-spatial structure
- Structure in phylodynamic analyses
- Structured birth-death models
- Structured coalescent models
- Structural dynamics

## Phylodynamics in Action

## References

# Structured Populations

Biological populations often have some internal “structure”:



## Structured Populations

Population structure

Spatial structure

Non-spatial structure

Structure in phylodynamic analyses

Structured birth-death models

Structured coalescent models

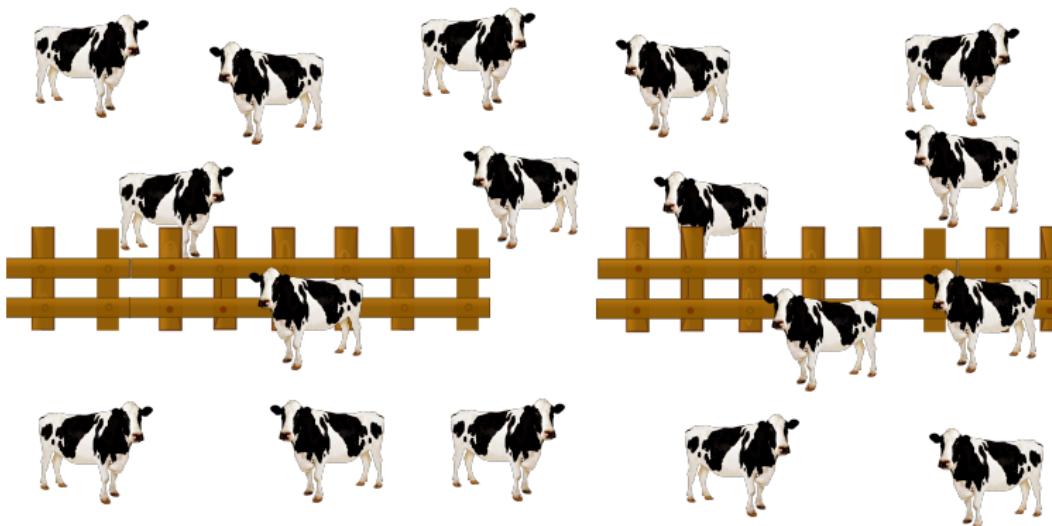
Structural dynamics

## Phyldynamics in Action

## References

# Structured Populations

Biological populations often have some internal “structure”:



## Structured Populations

Population structure

Spatial structure

Non-spatial structure

Structure in phylodynamic analyses

Structured birth-death models

Structured coalescent models

Structural dynamics

## Phyldynamics in Action

## References

# Structured Populations

What do we really mean by a “structured population”?

## Our Definition

A population is structured if its members possess one or more traits (e.g. location, group membership, ...) that affects their phylodynamic parameters (e.g. birth rate, death rate, sampling rate, coalescence rate).

### Structured Populations

Population structure

Spatial structure

Non-spatial structure

Structure in phylodynamic analyses

Structured birth-death models

Structured coalescent models

Structural dynamics

### Phylodynamics in Action

### References

# Geographic/Spatial structuring



## Structured Populations

- Population structure
- Spatial structure
- Non-spatial structure
- Structure in phylodynamic analyses
- Structured birth-death models
- Structured coalescent models
- Structural dynamics

## Phylodynamics in Action

## References

# Geographic/Spatial structuring

- ▶ Gene flow limited by spatial/geographic segregation of subpopulations.



## Structured Populations

- Population structure
- Spatial structure
- Non-spatial structure
- Structure in phylodynamic analyses
- Structured birth-death models
- Structured coalescent models
- Structural dynamics

## Phyldynamics in Action

## References

# Geographic/Spatial structuring

- ▶ Gene flow limited by spatial/geographic segregation of subpopulations.
- ▶ Impact of structure depends on the rate of migration across boundaries relative to the local birth-rate.



Google

## Structured Populations

- Population structure
- Spatial structure
- Non-spatial structure
- Structure in phylodynamic analyses
- Structured birth-death models
- Structured coalescent models
- Structural dynamics

## Phyldynamics in Action

## References

# Geographic/Spatial structuring

- ▶ Gene flow limited by spatial/geographic segregation of subpopulations.
- ▶ Impact of structure depends on the rate of migration across boundaries relative to the local birth-rate.
- ▶ A population spanning an archipelago is a classic example of a spatially structured population.



Google

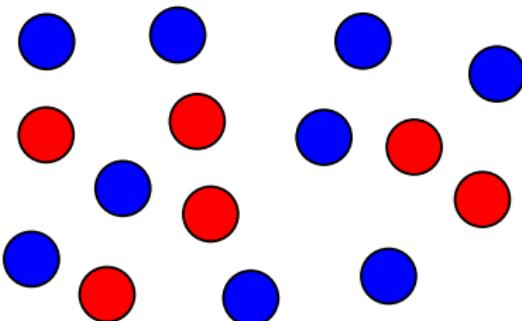
## Structured Populations

- Population structure
- Spatial structure
- Non-spatial structure
- Structure in phylodynamic analyses
- Structured birth-death models
- Structured coalescent models
- Structural dynamics

## Phyldynamics in Action

## References

# Non-spatial structuring



Even populations that are spatially mixed can be structured in other ways:

## Structured Populations

Population structure

Spatial structure

Non-spatial structure

Structure in phylodynamic analyses

Structured birth-death models

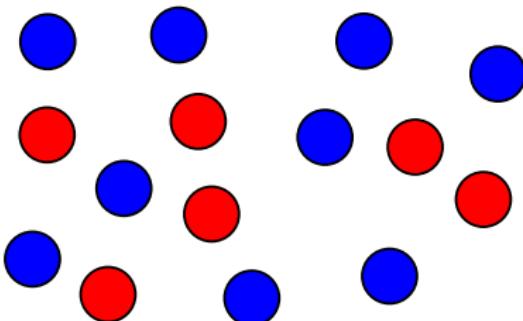
Structured coalescent models

Structural dynamics

## Phylodynamics in Action

## References

# Non-spatial structuring



Even populations that are spatially mixed can be structured in other ways:

- ▶ pathogen populations are generally composed of many within-host sub-populations,

## Structured Populations

Population structure

Spatial structure

Non-spatial structure

Structure in phylodynamic analyses

Structured birth-death models

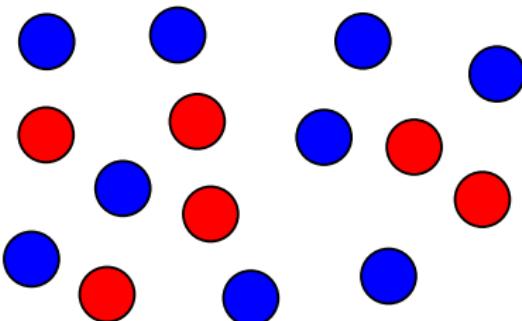
Structured coalescent models

Structural dynamics

## Phylodynamics in Action

## References

# Non-spatial structuring



Even populations that are spatially mixed can be structured in other ways:

- ▶ pathogen populations are generally composed of many within-host sub-populations,
- ▶ pathogen sub-populations may possess traits (e.g. drug resistance/susceptibility) affecting reproductive success,

## Structured Populations

Population structure

Spatial structure

Non-spatial structure

Structure in phylodynamic analyses

Structured birth-death models

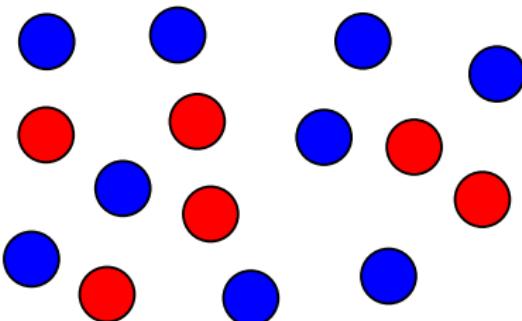
Structured coalescent models

Structural dynamics

## Phylodynamics in Action

## References

# Non-spatial structuring



Even populations that are spatially mixed can be structured in other ways:

- ▶ pathogen populations are generally composed of many within-host sub-populations,
- ▶ pathogen sub-populations may possess traits (e.g. drug resistance/susceptibility) affecting reproductive success,
- ▶ infected individuals may be in different epidemiological states (e.g. exposed vs infectious),

## Structured Populations

Population structure

Spatial structure

Non-spatial structure

Structure in phylodynamic analyses

Structured birth-death models

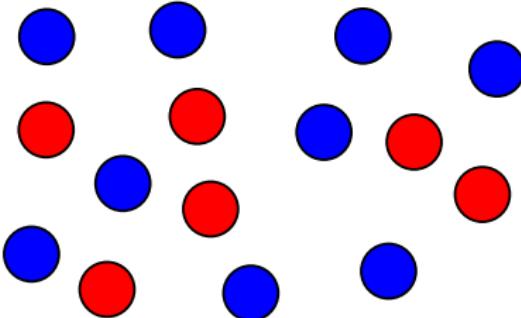
Structured coalescent models

Structural dynamics

## Phylodynamics in Action

## References

# Non-spatial structuring



Even populations that are spatially mixed can be structured in other ways:

- ▶ pathogen populations are generally composed of many within-host sub-populations,
- ▶ pathogen sub-populations may possess traits (e.g. drug resistance/susceptibility) affecting reproductive success,
- ▶ infected individuals may be in different epidemiological states (e.g. exposed vs infectious),
- ▶ sampled animals may be members of different species, between which there may be (extremely rare) horizontal gene transfer.

## Structured Populations

Population structure

Spatial structure

Non-spatial structure

Structure in phylodynamic analyses

Structured birth-death models

Structured coalescent models

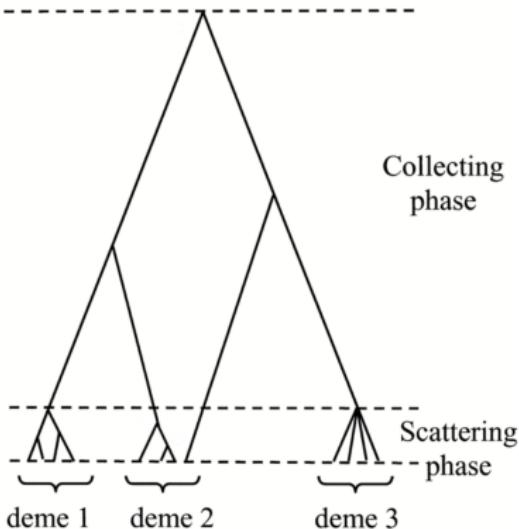
Structural dynamics

## Phylodynamics in Action

## References

# Importance to phylodynamic analyses

- ▶ Population structure can play an important role in shaping the phylogenetic relationships between samples.



[Pannell, 2003]

## Structured Populations

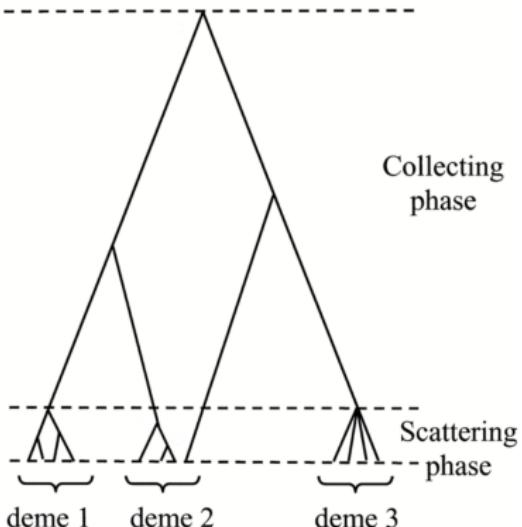
- Population structure
- Spatial structure
- Non-spatial structure
- Structure in phylodynamic analyses
- Structured birth-death models
- Structured coalescent models
- Structural dynamics

## Phylodynamics in Action

## References

# Importance to phylodynamic analyses

- ▶ Population structure can play an important role in shaping the phylogenetic relationships between samples.
- ▶ Failing to account for existing structure in phylodynamic analyses can bias results.



[Pannell, 2003]

## Structured Populations

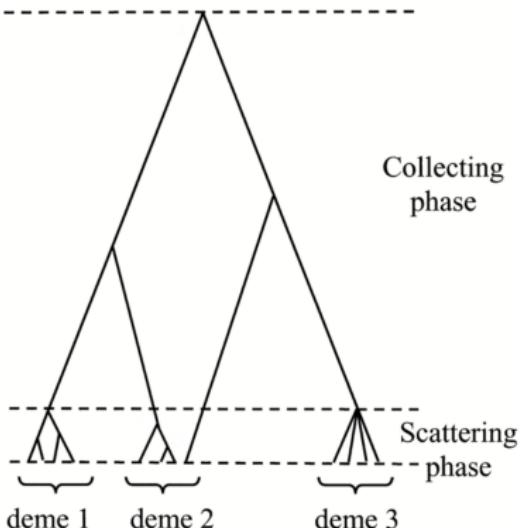
- Population structure
- Spatial structure
- Non-spatial structure
- Structure in phylodynamic analyses
- Structured birth-death models
- Structured coalescent models
- Structural dynamics

## Phylodynamics in Action

## References

# Importance to phylodynamic analyses

- ▶ Population structure can play an important role in shaping the phylogenetic relationships between samples.
- ▶ Failing to account for existing structure in phylodynamic analyses can bias results.
- ▶ We can also learn about parameters of structured models (e.g. local birth/death rates and sub-population sizes) using structure-aware phylodynamic models.



[Pannell, 2003]

## Structured Populations

- Population structure
- Spatial structure
- Non-spatial structure
- Structure in phylodynamic analyses
- Structured birth-death models
- Structured coalescent models
- Structural dynamics

## Phylodynamics in Action

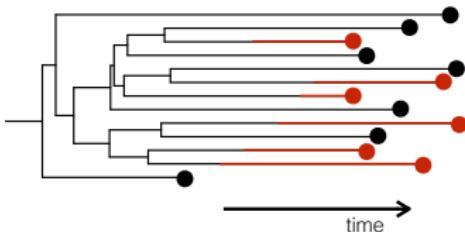
## References

# Phylogenetic trees contain information about population structure

Transmitted drug resistance:



*De novo* drug resistance:



- ▶ Each tip corresponds to one patient. The tree corresponds to the transmission chain.

## Structured Populations

Population structure

Spatial structure

Non-spatial structure

Structure in phylodynamic analyses

Structured birth-death models

Structured coalescent models

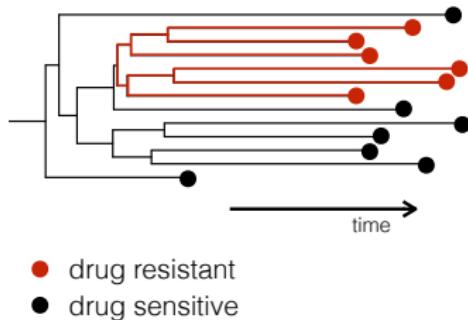
Structural dynamics

## Phylodynamics in Action

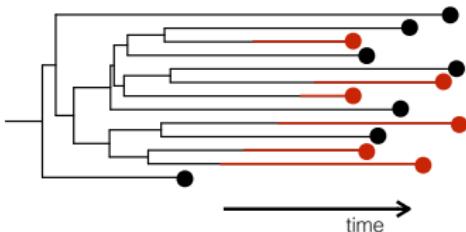
## References

# Phylogenetic trees contain information about population structure

Transmitted drug resistance:



*De novo* drug resistance:



- ▶ Each tip corresponds to one patient. The tree corresponds to the transmission chain.
- ▶ Counting number of sensitive and number of resistant samples through time does not provide insight into which scenario happened (transmitted or *de novo* drug resistance).

## Structured Populations

- Population structure
- Spatial structure
- Non-spatial structure
- Structure in phylodynamic analyses
- Structured birth-death models
- Structured coalescent models
- Structural dynamics

## Phylodynamics in Action

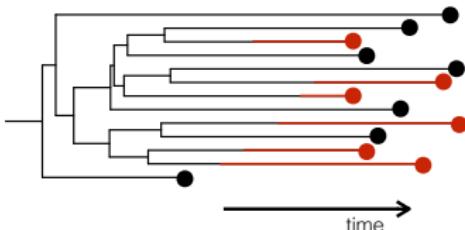
## References

# Phylogenetic trees contain information about population structure

Transmitted drug resistance:



*De novo* drug resistance:



## Structured Populations

- Population structure
- Spatial structure
- Non-spatial structure
- Structure in phylodynamic analyses
- Structured birth-death models
- Structured coalescent models
- Structural dynamics

## Phylodynamics in Action

## References

- ▶ Each tip corresponds to one patient. The tree corresponds to the transmission chain.
- ▶ Counting number of sensitive and number of resistant samples through time does not provide insight into which scenario happened (transmitted or *de novo* drug resistance).
- ▶ However, the phylogeny with tip labels contains information about the scenarios...

# Phylogenetic trees contain information about population structure

## Structured Populations

Population structure

Spatial structure

Non-spatial structure

Structure in phylodynamic analyses

Structured birth-death models

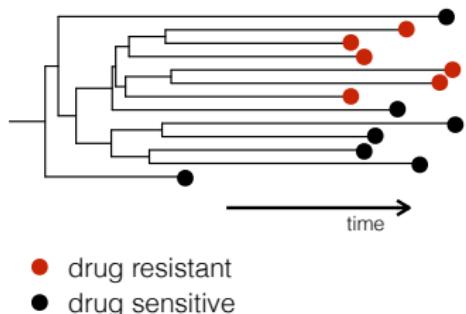
Structured coalescent models

Structural dynamics

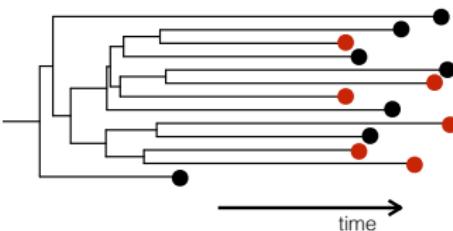
## Phylodynamics in Action

## References

Transmitted drug resistance:

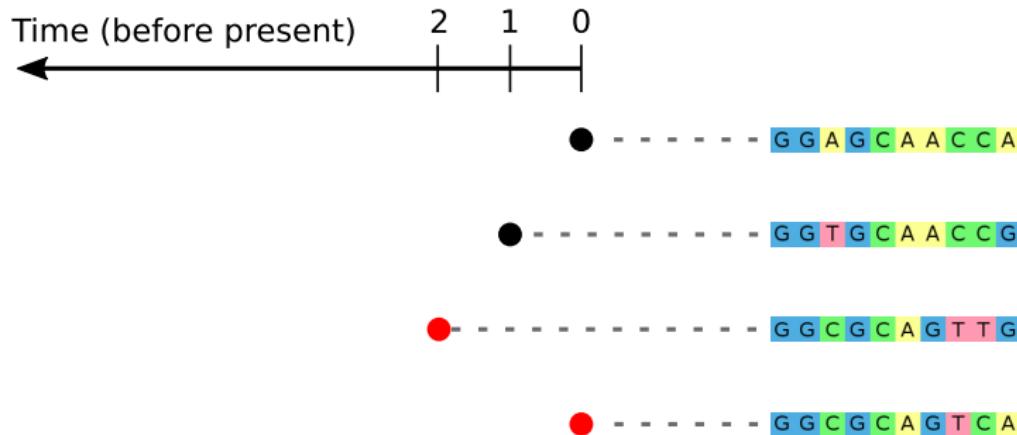


*De novo* drug resistance:



(Note that this information is hidden from traditional epidemiological data such as hospital case records.)

# Basic structured pop. inference problem



## Structured Populations

- Population structure
- Spatial structure
- Non-spatial structure
- Structure in phylodynamic analyses

Structured birth-death models

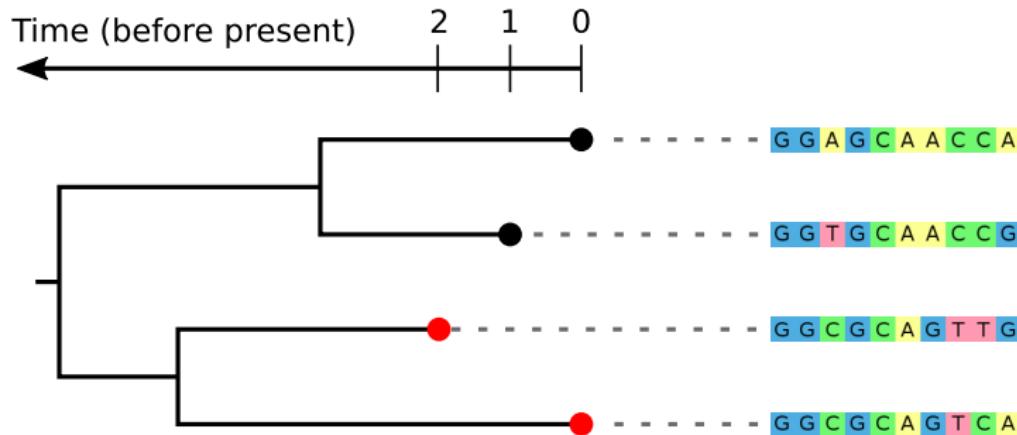
Structured coalescent models

Structural dynamics

## Phylodynamics in Action

## References

# Basic structured pop. inference problem



Structured Populations

Population structure

Spatial structure

Non-spatial structure

Structure in phylodynamic analyses

Structured birth-death models

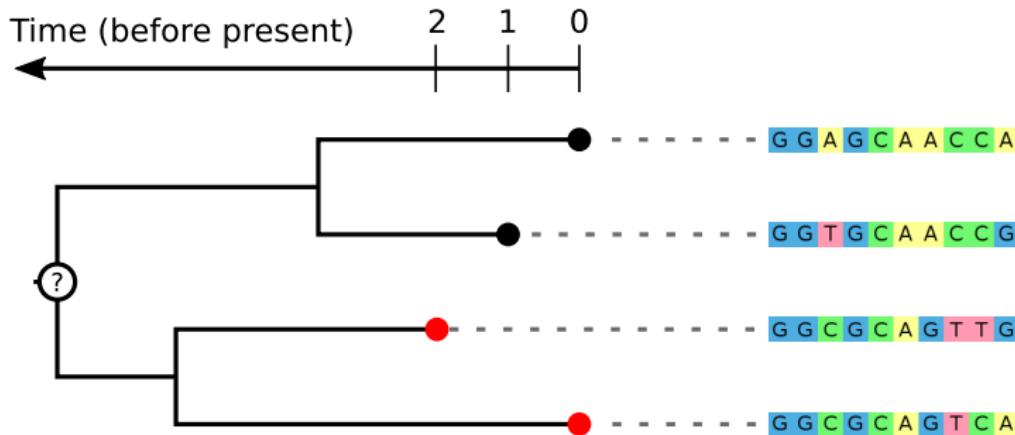
Structured coalescent models

Structural dynamics

Phylodynamics in Action

References

# Basic structured pop. inference problem



Structured  
Populations

Population structure

Spatial structure

Non-spatial structure

Structure in phylodynamic analyses

Structured birth-death models

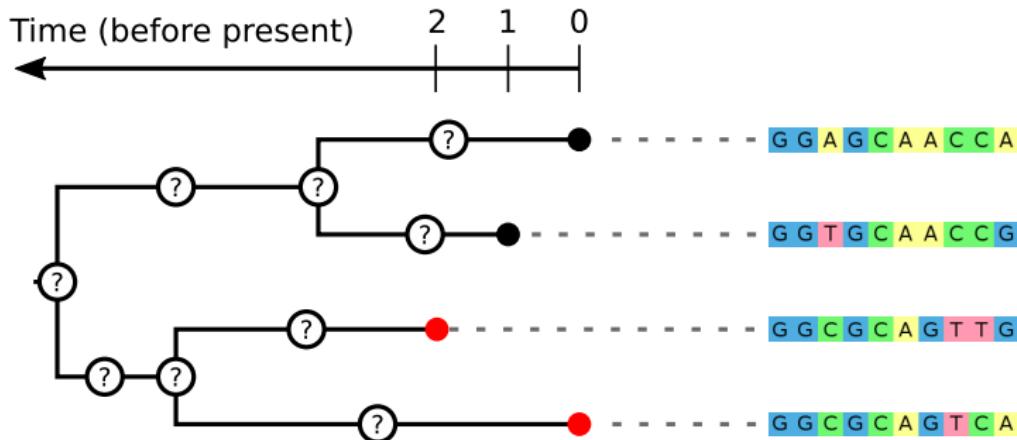
Structured coalescent models

Structural dynamics

Phylodynamics in Action

References

# Basic structured pop. inference problem



Structured  
Populations

Population structure

Spatial structure

Non-spatial structure

Structure in phylodynamic analyses

Structured birth-death models

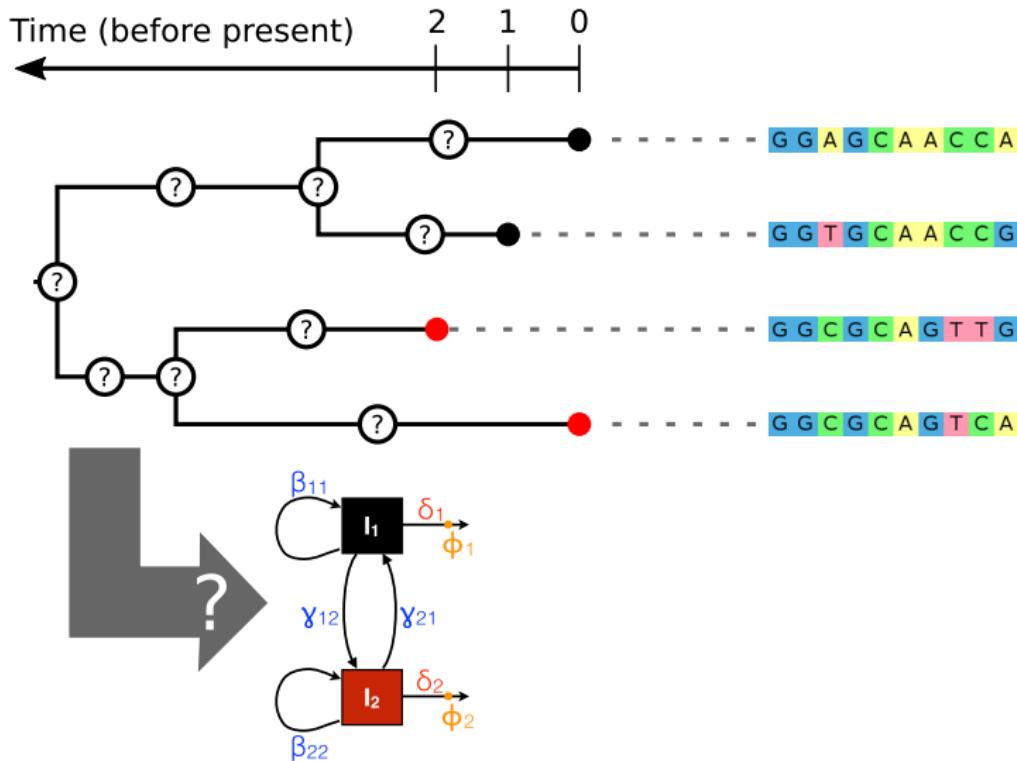
Structured coalescent models

Structural dynamics

Phylodynamics in Action

References

# Basic structured pop. inference problem



Structured  
Populations

Population structure

Spatial structure

Non-spatial structure

Structure in phylodynamic analyses

Structured birth-death models

Structured coalescent models

Structural dynamics

Phylodynamics in Action

References

# Structured birth-death models

Several models augment the existing birth-death models to include individuals of different types, including:

## Binary State Speciation and Extinction (BiSSE)

Allows for 2 states affecting birth and death rates,  
assumes complete present-day sampling.

[Maddison et al., 2007]

## Multi-State Speciation and Extinction (MuSSE)

Allows for > 2 states, allows for incomplete  
present-day sampling. [FitzJohn et al., 2009]

## Multi-type Birth Death (MTBD)

Allows for > 2 states, allows for both incomplete  
present-day and through-time sampling.  
[Kühnert et al., 2016, Scire et al., ]

### Structured Populations

- Population structure
- Spatial structure
- Non-spatial structure
- Structure in phylodynamic analyses

- Structured birth-death models

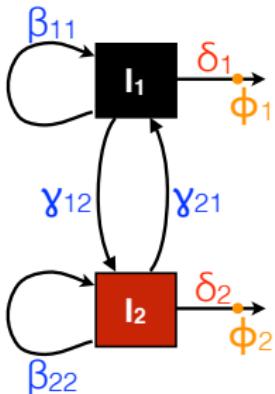
- Structured coalescent models

- Structural dynamics

### Phyldynamics in Action

### References

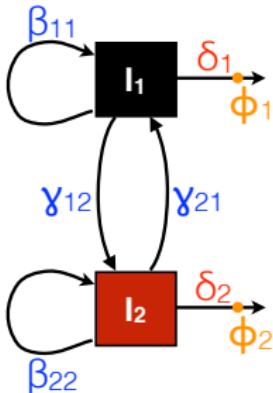
# Simple structured birth-death model



- Structured Populations
  - Population structure
  - Spatial structure
  - Non-spatial structure
  - Structure in phylodynamic analyses
  - Structured birth-death models
  - Structured coalescent models
  - Structural dynamics
- Phylodynamics in Action
- References

- ▶ In this case,  $\beta_{11}$  is transmission (birth) rate of drug sensitive strains,  $\beta_{22}$  is transmission (birth) rate of drug resistant strains,  $\gamma_{12}$  is the rate of resistance evolution.

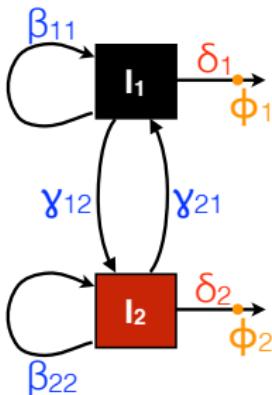
# Simple structured birth-death model



<b>Structured Populations</b>
Population structure
Spatial structure
Non-spatial structure
Structure in phylodynamic analyses
<b>Structured birth-death models</b>
Structured coalescent models
Structural dynamics
<b>Phyldynamics in Action</b>
References

- ▶ In this case,  $\beta_{11}$  is transmission (birth) rate of drug sensitive strains,  $\beta_{22}$  is transmission (birth) rate of drug resistant strains,  $\gamma_{12}$  is the rate of resistance evolution.
- ▶ In general:
  - ▶ Different compartments may represent pathogen strains, geographic locations, host risk groups, etc.

# Simple structured birth-death model



<b>Structured Populations</b>
Population structure
Spatial structure
Non-spatial structure
Structure in phylodynamic analyses
<b>Structured birth-death models</b>
Structured coalescent models
Structural dynamics
<b>Phylodynamics in Action</b>
References

- ▶ In this case,  $\beta_{11}$  is transmission (birth) rate of drug sensitive strains,  $\beta_{22}$  is transmission (birth) rate of drug resistant strains,  $\gamma_{12}$  is the rate of resistance evolution.
- ▶ In general:
  - ▶ Different compartments may represent pathogen strains, geographic locations, host risk groups, etc.
  - ▶ The  $\beta_{ij}$  represent the rate at which individuals of type  $i$  **produce** individuals of type  $j$ , while  $\gamma_{ij}$  is the rate at which individuals of type  $i$  **become** type  $j$ .

# Multi-type birth-death phylodynamic likelihood

Very similar to the unstructured case:

## Structured Populations

- Population structure
- Spatial structure
- Non-spatial structure
- Structure in phylodynamic analyses

## Structured birth-death models

- Structured coalescent models

## Structural dynamics

## Phylodynamics in Action

## References

# Multi-type birth-death phylodynamic likelihood

Very similar to the unstructured case:

- ▶ Let  $p_i(t)$  be the probability that an individual of type  $i$ , alive at time  $t$  in the past, gives rise to *no* sampled descendants.

## Structured Populations

- Population structure
- Spatial structure
- Non-spatial structure
- Structure in phylodynamic analyses

## Structured birth-death models

- Structured coalescent models

## Structural dynamics

## Phylodynamics in Action

## References

# Multi-type birth-death phylodynamic likelihood

Very similar to the unstructured case:

- ▶ Let  $p_i(t)$  be the probability that an individual of type  $i$ , alive at time  $t$  in the past, gives rise to *no* sampled descendants.
- ▶ Satisfies the following ODE:

$$\begin{aligned} \frac{d}{dt} p_i(t) = & - \left( \sum_{j=1}^d (\beta_{ij} + \gamma_{ij}) + \delta_i \right) p_i(t) \\ & + \sum_{j=1}^d \beta_{ij} p_i(t) p_j(t) + \sum_{j=1}^d \gamma_{ij} p_j(t) + \delta_i \end{aligned}$$

## Structured Populations

- Population structure
- Spatial structure
- Non-spatial structure
- Structure in phylodynamic analyses

## Structured birth-death models

## Structured coalescent models

## Structural dynamics

## Phylodynamics in Action

## References

# Multi-type birth-death phylodynamic likelihood

Very similar to the unstructured case:

- ▶ Let  $p_i(t)$  be the probability that an individual of type  $i$ , alive at time  $t$  in the past, gives rise to *no* sampled descendants.
- ▶ Satisfies the following ODE:

$$\begin{aligned} \frac{d}{dt} p_i(t) = & - \left( \sum_{j=1}^d (\beta_{ij} + \gamma_{ij}) + \delta_i \right) p_i(t) \\ & + \sum_{j=1}^d \beta_{ij} p_i(t) p_j(t) + \sum_{j=1}^d \gamma_{ij} p_j(t) + \delta_i \end{aligned}$$

- ▶ Let  $g_i^e(t)$  be the probability that an individual of type  $i$ , alive at time  $t$  in the past and belonging to sampled tree edge  $e$ , gives rise to the sampled phylogeny below that edge. (This obeys a similar ODE.)

## Structured Populations

Population structure

Spatial structure

Non-spatial structure

Structure in phylodynamic analyses

Structured birth-death models

Structured coalescent models

Structural dynamics

## Phylodynamics in Action

## References

# Multi-type birth-death phylodynamic likelihood

Very similar to the unstructured case:

- ▶ Let  $p_i(t)$  be the probability that an individual of type  $i$ , alive at time  $t$  in the past, gives rise to *no* sampled descendants.
- ▶ Satisfies the following ODE:

$$\frac{d}{dt} p_i(t) = - \left( \sum_{j=1}^d (\beta_{ij} + \gamma_{ij}) + \delta_i \right) p_i(t) \\ + \sum_{j=1}^d \beta_{ij} p_i(t) p_j(t) + \sum_{j=1}^d \gamma_{ij} p_j(t) + \delta_i$$

- ▶ Let  $g_i^e(t)$  be the probability that an individual of type  $i$ , alive at time  $t$  in the past and belonging to sampled tree edge  $e$ , gives rise to the sampled phylogeny below that edge. (This obeys a similar ODE.)

Unlike unstructured case, no known analytical solutions to these equations exist: must be solved numerically to compute structured phylodynamic likelihood.

## Structured Populations

Population structure

Spatial structure

Non-spatial structure

Structure in phylodynamic analyses

Structured birth-death models

Structured coalescent models

Structural dynamics

## Phylodynamics in Action

## References

# Bayesian inference for multi-type birth-death models

The phylogenetic posterior for a multi-type birth-death analysis becomes

$$P(\mathcal{T}, Q, \eta | A, L) = \frac{P(A|Q, \mathcal{T})P(\mathcal{T}, L|\eta)P(Q)P(\eta)}{P(A, L)}$$

- ▶  $\mathcal{T}$  is a phylogenetic tree.
- ▶  $L$  are the locations/types associated with the sequences.
- ▶  $\eta$  are the parameters of the multi-type birth-death model.
- ▶  $P(\mathcal{T}, L|\eta)$  is the structured phylodynamic likelihood.

As before,  $A$  represents the sequence alignment,  $Q$  the substitution rate matrix.

## Structured Populations

- Population structure
- Spatial structure
- Non-spatial structure
- Structure in phylodynamic analyses
- Structured birth-death models

## Structured coalescent models

- Structural dynamics

## Phyldynamics in Action

## References

# Bayesian inference for multi-type birth-death models

The phylogenetic posterior for a multi-type birth-death analysis becomes

$$P(\mathcal{T}, Q, \eta | A, L) = \frac{P(A|Q, \mathcal{T})P(\mathcal{T}, L|\eta)P(Q)P(\eta)}{P(A, L)}$$

- ▶  $\mathcal{T}$  is a phylogenetic tree.
- ▶  $L$  are the locations/types associated with the sequences.
- ▶  $\eta$  are the parameters of the multi-type birth-death model.
- ▶  $P(\mathcal{T}, L|\eta)$  is the structured phylodynamic likelihood.

As before,  $A$  represents the sequence alignment,  $Q$  the substitution rate matrix.

- ▶ Likelihood computed using equations on previous page integrates out (i.e. averages over) ancestral states.

## Structured Populations

Population structure

Spatial structure

Non-spatial structure

Structure in phylodynamic analyses

Structured birth-death models

Structured coalescent models

Structural dynamics

## Phylodynamics in Action

## References

# Bayesian inference for multi-type birth-death models

The phylogenetic posterior for a multi-type birth-death analysis becomes

$$P(\mathcal{T}, Q, \eta | A, L) = \frac{P(A|Q, \mathcal{T})P(\mathcal{T}, L|\eta)P(Q)P(\eta)}{P(A, L)}$$

- ▶  $\mathcal{T}$  is a phylogenetic tree.
- ▶  $L$  are the locations/types associated with the sequences.
- ▶  $\eta$  are the parameters of the multi-type birth-death model.
- ▶  $P(\mathcal{T}, L|\eta)$  is the structured phylodynamic likelihood.

As before,  $A$  represents the sequence alignment,  $Q$  the substitution rate matrix.

- ▶ Likelihood computed using equations on previous page integrates out (i.e. averages over) ancestral states.
- ▶ Can also derive probability of “coloured” tree with ancestral states marked. In that case, MCMC must be performed on this expanded state space of coloured trees.

Structured Populations
Population structure
Spatial structure
Non-spatial structure
Structure in phylodynamic analyses
Structured birth-death models
Structured coalescent models
Structural dynamics
Phyldynamics in Action
References

# Bayesian inference for multi-type birth-death models

The phylogenetic posterior for a multi-type birth-death analysis becomes

$$P(\mathcal{T}, Q, \eta | A, L) = \frac{P(A|Q, \mathcal{T})P(\mathcal{T}, L|\eta)P(Q)P(\eta)}{P(A, L)}$$

- ▶  $\mathcal{T}$  is a phylogenetic tree.
- ▶  $L$  are the locations/types associated with the sequences.
- ▶  $\eta$  are the parameters of the multi-type birth-death model.
- ▶  $P(\mathcal{T}, L|\eta)$  is the structured phylodynamic likelihood.

As before,  $A$  represents the sequence alignment,  $Q$  the substitution rate matrix.

- ▶ Likelihood computed using equations on previous page integrates out (i.e. averages over) ancestral states.
- ▶ Can also derive probability of “coloured” tree with ancestral states marked. In that case, MCMC must be performed on this expanded state space of coloured trees.
  - + Posterior distribution for ancestral types directly available.

## Structured Populations

Population structure

Spatial structure

Non-spatial structure

Structure in phylodynamic analyses

Structured birth-death models

Structured coalescent models

Structural dynamics

## Phyldynamics in Action

## References

# Bayesian inference for multi-type birth-death models

The phylogenetic posterior for a multi-type birth-death analysis becomes

$$P(\mathcal{T}, Q, \eta | A, L) = \frac{P(A|Q, \mathcal{T})P(\mathcal{T}, L|\eta)P(Q)P(\eta)}{P(A, L)}$$

- ▶  $\mathcal{T}$  is a phylogenetic tree.
- ▶  $L$  are the locations/types associated with the sequences.
- ▶  $\eta$  are the parameters of the multi-type birth-death model.
- ▶  $P(\mathcal{T}, L|\eta)$  is the structured phylodynamic likelihood.

As before,  $A$  represents the sequence alignment,  $Q$  the substitution rate matrix.

- ▶ Likelihood computed using equations on previous page integrates out (i.e. averages over) ancestral states.
- ▶ Can also derive probability of “coloured” tree with ancestral states marked. In that case, MCMC must be performed on this expanded state space of coloured trees.
  - + Posterior distribution for ancestral types directly available.
  - Analyses much more computationally demanding.

Structured Populations
Population structure
Spatial structure
Non-spatial structure
Structure in phylodynamic analyses
Structured birth-death models
Structured coalescent models
Structural dynamics
Phyldynamics in Action
References

# Example: Seasonal influenza (H3N2)

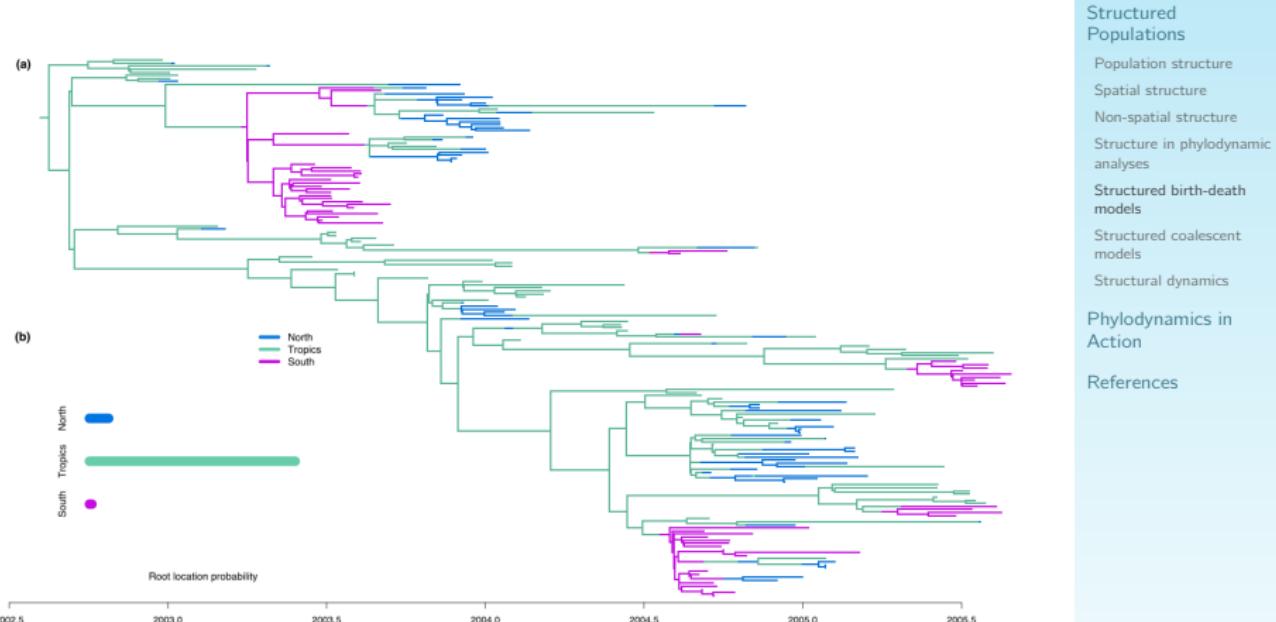


Figure adapted from [Kühnert et al., 2016]

Phylogenetic tree of geographic spread of seasonal influenza inferred under a multi-type birth-death model.

## Structured Populations

- Population structure
- Spatial structure
- Non-spatial structure
- Structure in phylodynamic analyses
- Structured birth-death models
- Structured coalescent models
- Structural dynamics

## Phylodynamics in Action

## References

# Example: Seasonal influenza (H3N2)

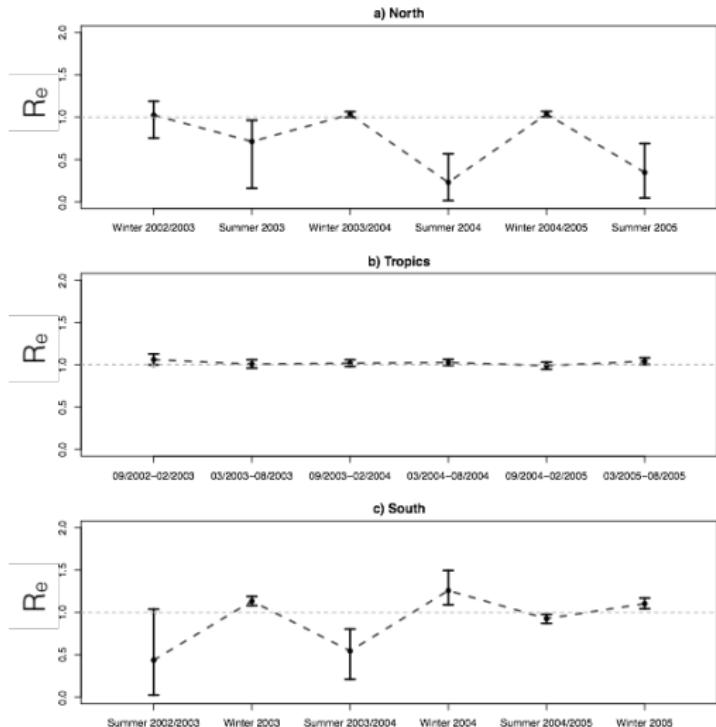


Figure adapted from [Kühnert et al., 2016]

Effective reproductive number of seasonal influenza inferred using the same multi-type birth-death model.

## Structured Populations

- Population structure
- Spatial structure
- Non-spatial structure
- Structure in phylodynamic analyses
- Structured birth-death models
- Structured coalescent models
- Structural dynamics

## Phylodynamics in Action

### References

# Unknown Leaf Types

- Newer methods are available that address the problem of accounting for unknown leaf types/locations.

## Structured Populations

Population structure

Spatial structure

Non-spatial structure

Structure in phylodynamic analyses

Structured birth-death models

Structured coalescent models

Structural dynamics

## Phylogenetics in Action

## References

# Unknown Leaf Types

- ▶ Newer methods are available that address the problem of accounting for unknown leaf types/locations.
- ▶ This problem is particularly challenging when the **number** of types/locations is also unknown.

Structured Populations

Population structure

Spatial structure

Non-spatial structure

Structure in phylodynamic analyses

Structured birth-death models

Structured coalescent models

Structural dynamics

Phylodynamics in Action

References

# Unknown Leaf Types

- ▶ Newer methods are available that address the problem of accounting for unknown leaf types/locations.
- ▶ This problem is particularly challenging when the **number** of types/locations is also unknown.

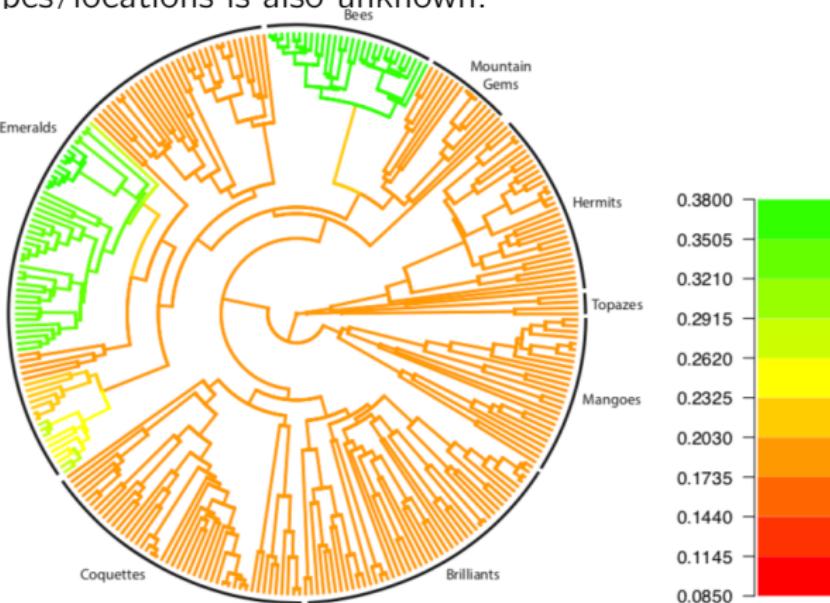


Figure adapted from [Barido-Sottani et al., 2018]  
Speciation rate classes across phylogeny of hummingbird species.

Structured Populations  
Population structure  
Spatial structure  
Non-spatial structure  
Structure in phydodynamic analyses  
Structured birth-death models  
Structured coalescent models  
Structural dynamics  
  
Phyldynamics in Action  
  
References

# Structured coalescent models

- ▶ Just as for the birth-death models, we can augment coalescent models to include multiple types.

## Structured Populations

- Population structure
- Spatial structure
- Non-spatial structure
- Structure in phylodynamic analyses
- Structured birth-death models
- Structured coalescent models
- Structural dynamics

## Phylodynamics in Action

## References

# Structured coalescent models

- ▶ Just as for the birth-death models, we can augment coalescent models to include multiple types.
- ▶ Instead of type-specific birth, death and sampling rates however, this yields models with type-specific population sizes.

## Structured Populations

- Population structure
- Spatial structure
- Non-spatial structure
- Structure in phylodynamic analyses
- Structured birth-death models
- Structured coalescent models
- Structural dynamics

## Phyldynamics in Action

## References

# Structured coalescent models

- ▶ Just as for the birth-death models, we can augment coalescent models to include multiple types.
- ▶ Instead of type-specific birth, death and sampling rates however, this yields models with type-specific population sizes.
- ▶ In a similar way to the unstructured coalescent process, the structured coalescent process can be derived from a variety of distinct forward-time models of population dynamics.

## Structured Populations

- Population structure
- Spatial structure
- Non-spatial structure
- Structure in phylodynamic analyses
- Structured birth-death models
- Structured coalescent models
- Structural dynamics

## Phyldynamics in Action

## References

# Structured coalescent models

- ▶ Just as for the birth-death models, we can augment coalescent models to include multiple types.
- ▶ Instead of type-specific birth, death and sampling rates however, this yields models with type-specific population sizes.
- ▶ In a similar way to the unstructured coalescent process, the structured coalescent process can be derived from a variety of distinct forward-time models of population dynamics.
- ▶ We focus here on an extension to the Wright-Fisher model.

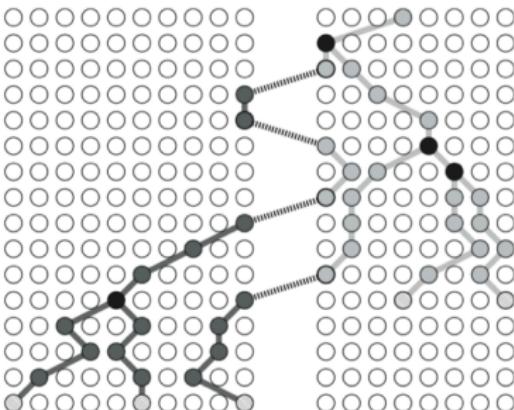
## Structured Populations

Population structure  
Spatial structure  
Non-spatial structure  
Structure in phylodynamic analyses  
Structured birth-death models  
Structured coalescent models  
Structural dynamics

## Phyldynamics in Action

## References

# Structured Wright-Fisher model



## Structured Populations

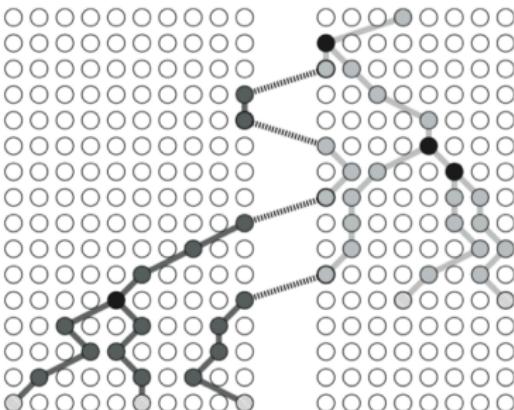
- Population structure
- Spatial structure
- Non-spatial structure
- Structure in phylodynamic analyses
- Structured birth-death models
- Structured coalescent models
- Structural dynamics

## Phylodynamics in Action

## References

- ▶ Introduced by [Notohara, 1990] (and probably others).

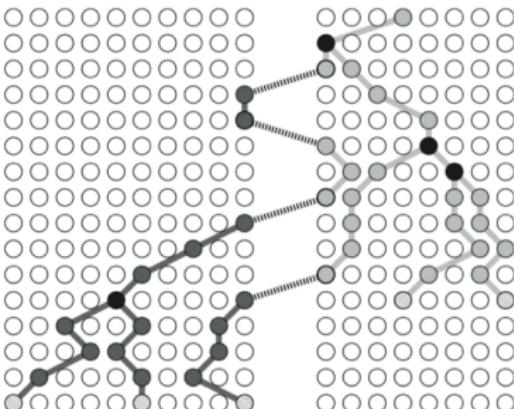
# Structured Wright-Fisher model



- Structured Populations
  - Population structure
  - Spatial structure
  - Non-spatial structure
  - Structure in phylodynamic analyses
  - Structured birth-death models
  - Structured coalescent models
  - Structural dynamics
- Phyldynamics in Action
- References

- ▶ Introduced by [Notohara, 1990] (and probably others).
- ▶ Assumes a single population is divided into sub-populations (demes) of size  $N_i$  for  $i \in [1, d]$ .

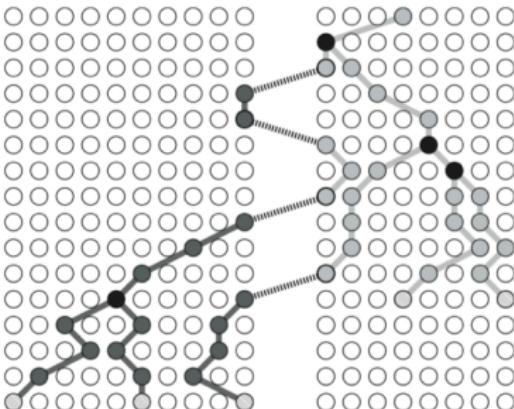
# Structured Wright-Fisher model



- Structured Populations
  - Population structure
  - Spatial structure
  - Non-spatial structure
  - Structure in phylodynamic analyses
  - Structured birth-death models
  - Structured coalescent models
  - Structural dynamics
- Phyldynamics in Action
- References

- ▶ Introduced by [Notohara, 1990] (and probably others).
- ▶ Assumes a single population is divided into sub-populations (demes) of size  $N_i$  for  $i \in [1, d]$ .
- ▶ Allows for migration between demes at rate  $q_{ij}$ .

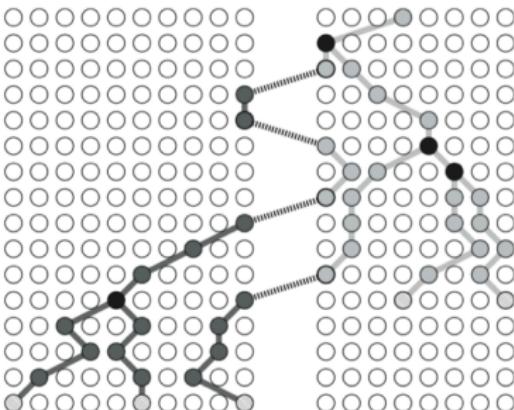
# Structured Wright-Fisher model



- Structured Populations
  - Population structure
  - Spatial structure
  - Non-spatial structure
  - Structure in phylodynamic analyses
  - Structured birth-death models
  - Structured coalescent models
  - Structural dynamics
- Phyldynamics in Action
- References

- ▶ Introduced by [Notohara, 1990] (and probably others).
- ▶ Assumes a single population is divided into sub-populations (demes) of size  $N_i$  for  $i \in [1, d]$ .
- ▶ Allows for migration between demes at rate  $q_{ij}$ .
- ▶ As in unstructured case, assume a fixed time interval  $g$  between successive generations.

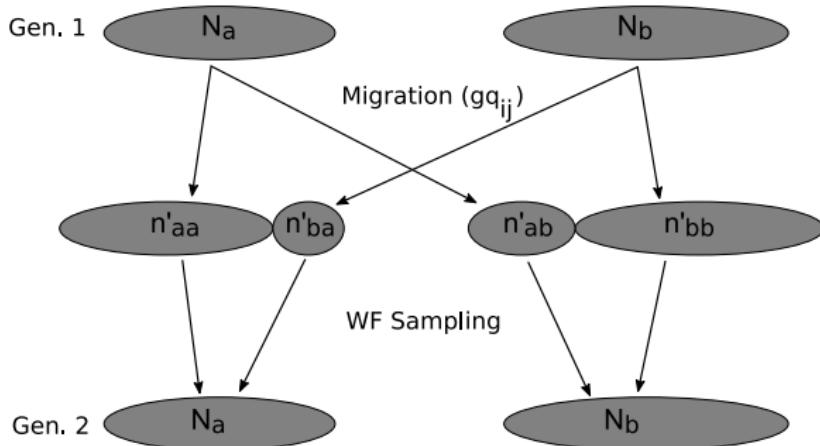
# Structured Wright-Fisher model



- Structured Populations
  - Population structure
  - Spatial structure
  - Non-spatial structure
  - Structure in phylodynamic analyses
  - Structured birth-death models
  - Structured coalescent models
  - Structural dynamics
- Phyldynamics in Action
- References

- ▶ Introduced by [Notohara, 1990] (and probably others).
- ▶ Assumes a single population is divided into sub-populations (demes) of size  $N_i$  for  $i \in [1, d]$ .
- ▶ Allows for migration between demes at rate  $q_{ij}$ .
- ▶ As in unstructured case, assume a fixed time interval  $g$  between successive generations.
- ▶ Assumes that sub-population sizes are unaffected by migration in the long term.

# Structured Wright-Fisher model



- ▶ Each generation in the structured WF model includes a migration phase and a sampling phase.

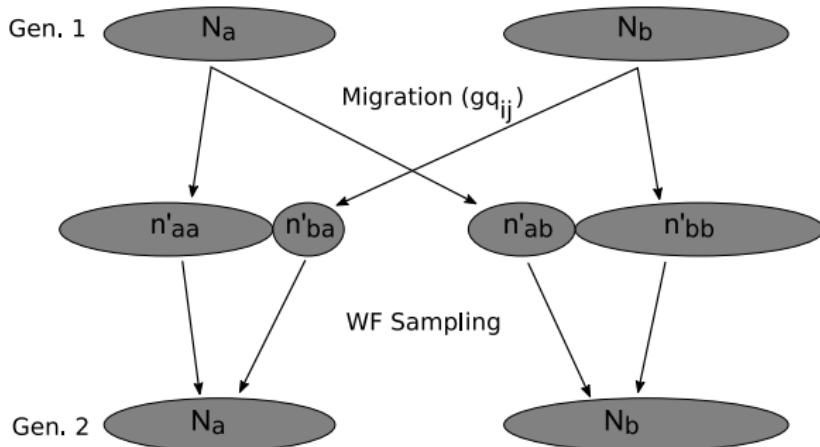
## Structured Populations

- Population structure
- Spatial structure
- Non-spatial structure
- Structure in phylodynamic analyses
- Structured birth-death models
- Structured coalescent models
- Structural dynamics

## Phyldynamics in Action

## References

# Structured Wright-Fisher model



- ▶ Each generation in the structured WF model includes a migration phase and a sampling phase.
  - ▶ Migration: Individuals migrate according to probabilities  $gq_{ij}$  to form intermediate generation.

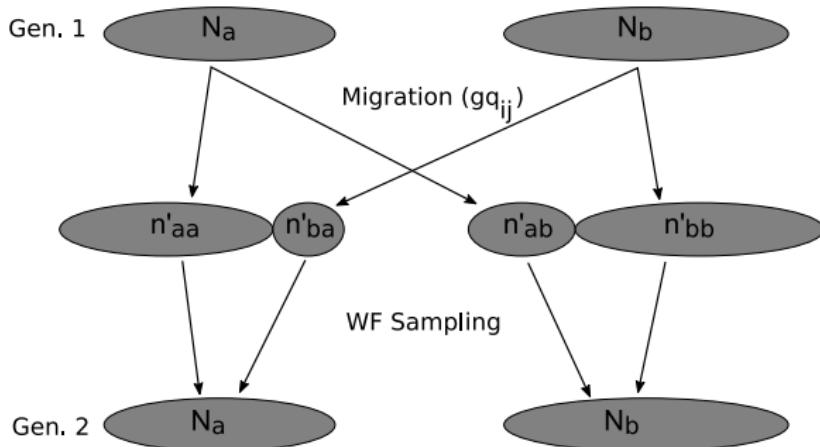
## Structured Populations

- Population structure
- Spatial structure
- Non-spatial structure
- Structure in phydodynamic analyses
- Structured birth-death models
- Structured coalescent models
- Structural dynamics

## Phyldynamics in Action

### References

# Structured Wright-Fisher model



- ▶ Each generation in the structured WF model includes a migration phase and a sampling phase.
  - ▶ Migration: Individuals migrate according to probabilities  $gq_{ij}$  to form intermediate generation.
  - ▶ Sampling: Children in next generation sample parents uniformly from intermediate population.

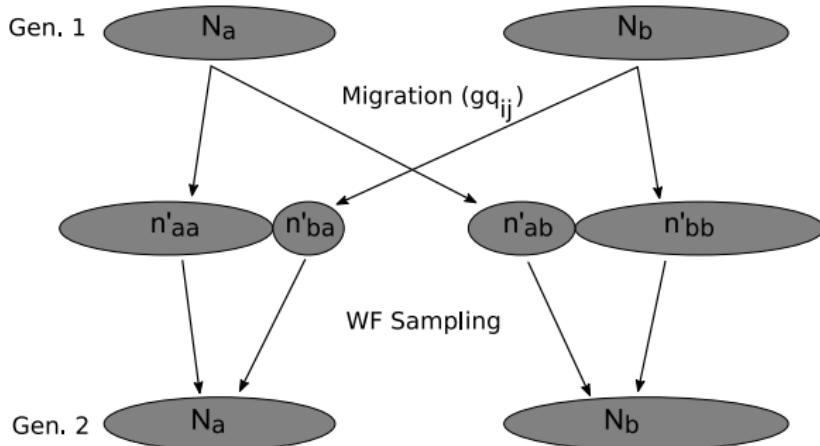
## Structured Populations

- Population structure
- Spatial structure
- Non-spatial structure
- Structure in phydodynamic analyses
- Structured birth-death models
- Structured coalescent models
- Structural dynamics

## Phyldynamics in Action

## References

# Structured Wright-Fisher model



- ▶ Each generation in the structured WF model includes a migration phase and a sampling phase.
  - ▶ Migration: Individuals migrate according to probabilities  $gq_{ij}$  to form intermediate generation.
  - ▶ Sampling: Children in next generation sample parents uniformly from intermediate population.
- ▶ Ensures subpopulation sizes remain fixed, even with asymmetric migration rates between demes.

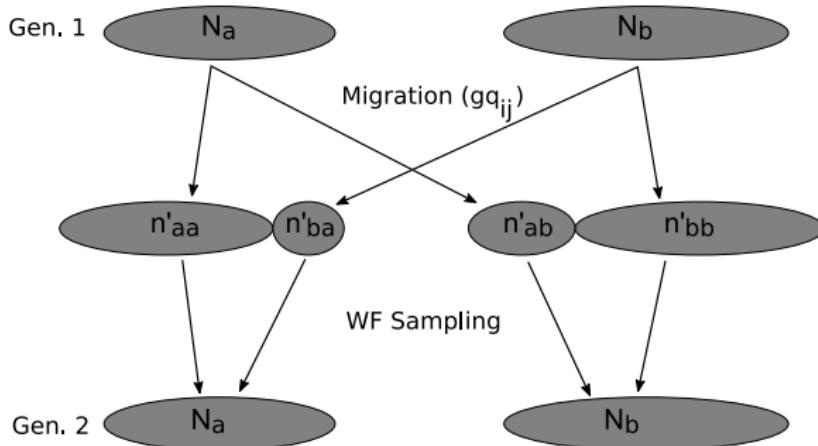
## Structured Populations

- Population structure
- Spatial structure
- Non-spatial structure
- Structure in phylodynamic analyses
- Structured birth-death models
- Structured coalescent models
- Structural dynamics

## Phyldynamics in Action

## References

# Structured Wright-Fisher model



- ▶ Each generation in the structured WF model includes a migration phase and a sampling phase.
  - ▶ Migration: Individuals migrate according to probabilities  $gq_{ij}$  to form intermediate generation.
  - ▶ Sampling: Children in next generation sample parents uniformly from intermediate population.
- ▶ Ensures subpopulation sizes remain fixed, even with asymmetric migration rates between demes.
- ▶ (Assumption: migration rate is much slower than time necessary for sub-populations to return to equilibrium population size.)

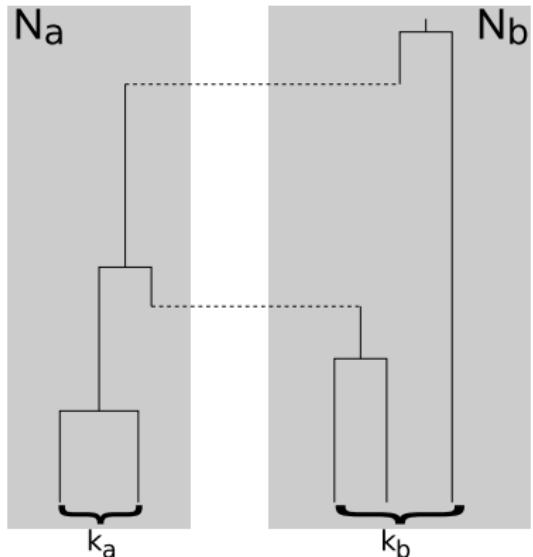
## Structured Populations

- Population structure
- Spatial structure
- Non-spatial structure
- Structure in phylodynamic analyses
- Structured birth-death models
- Structured coalescent models
- Structural dynamics

## Phyldynamics in Action

## References

# The structured coalescent



This is a backward in time tree generation process corresponding to the coalescent limit of the structured WF model.

Coalescence rate in deme i:

$$\binom{k_i}{2} \frac{1}{gN_i}$$

Migration rate (backward)  $i \rightarrow j$ :

$$k_i m_{ij}$$

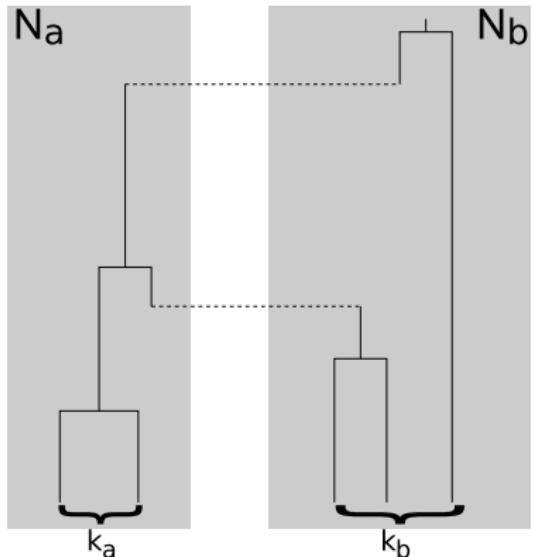
## Structured Populations

- Population structure
- Spatial structure
- Non-spatial structure
- Structure in phylodynamic analyses
- Structured birth-death models
- Structured coalescent models
- Structural dynamics

## Phyldynamics in Action

## References

# The structured coalescent



This is a backward in time tree generation process corresponding to the coalescent limit of the structured WF model.

Coalescence rate in deme  $i$ :

$$\binom{k_i}{2} \frac{1}{gN_i}$$

Migration rate (backward)  $i \rightarrow j$ :

$$k_i m_{ij}$$

- The backward-time migration rate  $m_{ij}$  (also called the immigration rate) is related to the forward time rate  $q_{ji}$  from the structured WF model by

$$m_{ij} = q_{ji} \frac{N_j}{N_i}$$

## Structured Populations

- Population structure
- Spatial structure
- Non-spatial structure
- Structure in phylodynamic analyses
- Structured birth-death models
- Structured coalescent models
- Structural dynamics

## Phyldynamics in Action

## References

# Expected coalescence time

For a symmetric 2 deme model ( $N_1 = N_2 = N$  and  $m_{12} = m_{21} = m$ ) we can derive the *expected* time for two lineages to coalesce:

## Structured Populations

- Population structure
- Spatial structure
- Non-spatial structure
- Structure in phylodynamic analyses

- Structured birth-death models

- Structured coalescent models

- Structural dynamics

## Phyldynamics in Action

## References

# Expected coalescence time

For a symmetric 2 deme model ( $N_1 = N_2 = N$  and  $m_{12} = m_{21} = m$ ) we can derive the *expected* time for two lineages to coalesce:

- ▶ Define  $T_d$  and  $T_s$  as the expected coalescence times when lineages are in different/same demes respectively.

## Structured Populations

Population structure

Spatial structure

Non-spatial structure

Structure in phylodynamic analyses

Structured birth-death models

Structured coalescent models

Structural dynamics

## Phyldynamics in Action

## References

# Expected coalescence time

For a symmetric 2 deme model ( $N_1 = N_2 = N$  and  $m_{12} = m_{21} = m$ ) we can derive the *expected* time for two lineages to coalesce:

- ▶ Define  $T_d$  and  $T_s$  as the expected coalescence times when lineages are in different/same demes respectively.
- ▶ Lineages in distinct demes cannot coalesce, so we have:

$$T_d = \frac{1}{2m} + T_s$$

## Structured Populations

Population structure

Spatial structure

Non-spatial structure

Structure in phylodynamic analyses

Structured birth-death models

Structured coalescent models

Structural dynamics

## Phyldynamics in Action

## References

# Expected coalescence time

For a symmetric 2 deme model ( $N_1 = N_2 = N$  and  $m_{12} = m_{21} = m$ ) we can derive the *expected* time for two lineages to coalesce:

- ▶ Define  $T_d$  and  $T_s$  as the expected coalescence times when lineages are in different/same demes respectively.
- ▶ Lineages in distinct demes cannot coalesce, so we have:

$$T_d = \frac{1}{2m} + T_s$$

- ▶ Lineages in the same deme wait for average time  $1/(2m + 1/Ng)$  before either coalescing or migrating:

$$T_s = \frac{1}{2m + 1/Ng} + \frac{1/Ng}{2m + 1/Ng} 0 + \frac{2m}{2m + 1/Ng} T_d$$

<b>Structured Populations</b>
Population structure
Spatial structure
Non-spatial structure
Structure in phylodynamic analyses
Structured birth-death models
Structured coalescent models
Structural dynamics
<b>Phyldynamics in Action</b>
References

# Expected coalescence time

For a symmetric 2 deme model ( $N_1 = N_2 = N$  and  $m_{12} = m_{21} = m$ ) we can derive the *expected* time for two lineages to coalesce:

- ▶ Define  $T_d$  and  $T_s$  as the expected coalescence times when lineages are in different/same demes respectively.
- ▶ Lineages in distinct demes cannot coalesce, so we have:

$$T_d = \frac{1}{2m} + T_s$$

- ▶ Lineages in the same deme wait for average time  $1/(2m + 1/Ng)$  before either coalescing or migrating:

$$T_s = \frac{1}{2m + 1/Ng} + \frac{1/Ng}{2m + 1/Ng} 0 + \frac{2m}{2m + 1/Ng} T_d$$

- ▶ Solving this pair of simultaneous equations yields:  
 $T_s = 2Ng$  and  $T_d = \frac{1}{2m} + 2Ng$ .

## Structured Populations

Population structure

Spatial structure

Non-spatial structure

Structure in phylodynamic analyses

Structured birth-death models

Structured coalescent models

Structural dynamics

## Phyldynamics in Action

## References

# Expected coalescence time

For a symmetric 2 deme model ( $N_1 = N_2 = N$  and  $m_{12} = m_{21} = m$ ) we can derive the *expected* time for two lineages to coalesce:

- ▶ Define  $T_d$  and  $T_s$  as the expected coalescence times when lineages are in different/same demes respectively.
- ▶ Lineages in distinct demes cannot coalesce, so we have:

$$T_d = \frac{1}{2m} + T_s$$

- ▶ Lineages in the same deme wait for average time  $1/(2m + 1/Ng)$  before either coalescing or migrating:

$$T_s = \frac{1}{2m + 1/Ng} + \frac{1/Ng}{2m + 1/Ng} 0 + \frac{2m}{2m + 1/Ng} T_d$$

- ▶ Solving this pair of simultaneous equations yields:

$$T_s = 2Ng \text{ and } T_d = \frac{1}{2m} + 2Ng.$$

Interesting: Expected time to coalesce from the same deme is independent of migration rate!

<b>Structured Populations</b>
Population structure
Spatial structure
Non-spatial structure
Structure in phylodynamic analyses
Structured birth-death models
Structured coalescent models
Structural dynamics
<b>Phyldynamics in Action</b>
References

# Bayesian inference for structured coalescent models

The phylogenetic posterior for a structured coalescent analysis becomes

$$P(\mathcal{T}_{\text{col}}, Q, \theta | A, L) = \frac{P(A|Q, \mathcal{T}_{\text{col}})P(\mathcal{T}_{\text{col}}|\theta, L)P(Q)P(\theta)}{P(A, L)}$$

- ▶  $\mathcal{T}_{\text{col}}$  is a phylogenetic tree with ancestral locations marked.
- ▶  $L$  are the locations/types associated with the sequences.
- ▶  $\theta$  are the parameters of the structured coalescent model.
- ▶  $P(\mathcal{T}_{\text{col}}|\theta, L)$  is the structured coalescent likelihood.

## Structured Populations

- Population structure
- Spatial structure
- Non-spatial structure
- Structure in phylodynamic analyses
- Structured birth-death models
- Structured coalescent models

## Structural dynamics

## Phyldynamics in Action

## References

# Bayesian inference for structured coalescent models

The phylogenetic posterior for a structured coalescent analysis becomes

$$P(\mathcal{T}_{\text{col}}, Q, \theta | A, L) = \frac{P(A|Q, \mathcal{T}_{\text{col}})P(\mathcal{T}_{\text{col}}|\theta, L)P(Q)P(\theta)}{P(A, L)}$$

- ▶  $\mathcal{T}_{\text{col}}$  is a phylogenetic tree with ancestral locations marked.
- ▶  $L$  are the locations/types associated with the sequences.
- ▶  $\theta$  are the parameters of the structured coalescent model.
- ▶  $P(\mathcal{T}_{\text{col}}|\theta, L)$  is the structured coalescent likelihood.

As before,  $A$  represents the sequence alignment,  $Q$  the substitution rate matrix.

## Structured Populations

- Population structure
- Spatial structure
- Non-spatial structure
- Structure in phylodynamic analyses
- Structured birth-death models
- Structured coalescent models

## Structural dynamics

## Phyldynamics in Action

## References

# Bayesian inference for structured coalescent models

The phylogenetic posterior for a structured coalescent analysis becomes

$$P(\mathcal{T}_{\text{col}}, Q, \theta | A, L) = \frac{P(A|Q, \mathcal{T}_{\text{col}})P(\mathcal{T}_{\text{col}}|\theta, L)P(Q)P(\theta)}{P(A, L)}$$

- ▶  $\mathcal{T}_{\text{col}}$  is a phylogenetic tree with ancestral locations marked.
- ▶  $L$  are the locations/types associated with the sequences.
- ▶  $\theta$  are the parameters of the structured coalescent model.
- ▶  $P(\mathcal{T}_{\text{col}}|\theta, L)$  is the structured coalescent likelihood.

As before,  $A$  represents the sequence alignment,  $Q$  the substitution rate matrix.

- ▶  $P(\mathcal{T}_{\text{col}}|\theta, L)$  is a simple extension of the unstructured coalescent expression.

## Structured Populations

- Population structure
- Spatial structure
- Non-spatial structure
- Structure in phylodynamic analyses
- Structured birth-death models
- Structured coalescent models

## Structural dynamics

## Phyldynamics in Action

## References

# Bayesian inference for structured coalescent models

The phylogenetic posterior for a structured coalescent analysis becomes

$$P(\mathcal{T}_{\text{col}}, Q, \theta | A, L) = \frac{P(A|Q, \mathcal{T}_{\text{col}})P(\mathcal{T}_{\text{col}}|\theta, L)P(Q)P(\theta)}{P(A, L)}$$

- ▶  $\mathcal{T}_{\text{col}}$  is a phylogenetic tree with ancestral locations marked.
- ▶  $L$  are the locations/types associated with the sequences.
- ▶  $\theta$  are the parameters of the structured coalescent model.
- ▶  $P(\mathcal{T}_{\text{col}}|\theta, L)$  is the structured coalescent likelihood.

As before,  $A$  represents the sequence alignment,  $Q$  the substitution rate matrix.

- ▶  $P(\mathcal{T}_{\text{col}}|\theta, L)$  is a simple extension of the unstructured coalescent expression.
- ▶ Unlike the multi-type birth-death case, it is difficult to integrate over ancestral locations.

## Structured Populations

- Population structure
- Spatial structure
- Non-spatial structure
- Structure in phylodynamic analyses
- Structured birth-death models
- Structured coalescent models

## Structural dynamics

## Phyldynamics in Action

## References

# Bayesian inference for structured coalescent models

The phylogenetic posterior for a structured coalescent analysis becomes

$$P(\mathcal{T}_{\text{col}}, Q, \theta | A, L) = \frac{P(A|Q, \mathcal{T}_{\text{col}})P(\mathcal{T}_{\text{col}}|\theta, L)P(Q)P(\theta)}{P(A, L)}$$

- ▶  $\mathcal{T}_{\text{col}}$  is a phylogenetic tree with ancestral locations marked.
- ▶  $L$  are the locations/types associated with the sequences.
- ▶  $\theta$  are the parameters of the structured coalescent model.
- ▶  $P(\mathcal{T}_{\text{col}}|\theta, L)$  is the structured coalescent likelihood.

As before,  $A$  represents the sequence alignment,  $Q$  the substitution rate matrix.

- ▶  $P(\mathcal{T}_{\text{col}}|\theta, L)$  is a simple extension of the unstructured coalescent expression.
- ▶ Unlike the multi-type birth-death case, it is difficult to integrate over ancestral locations.
- ▶ Has been done approximately  
[De Maio et al., 2015, Müller et al., 2017].

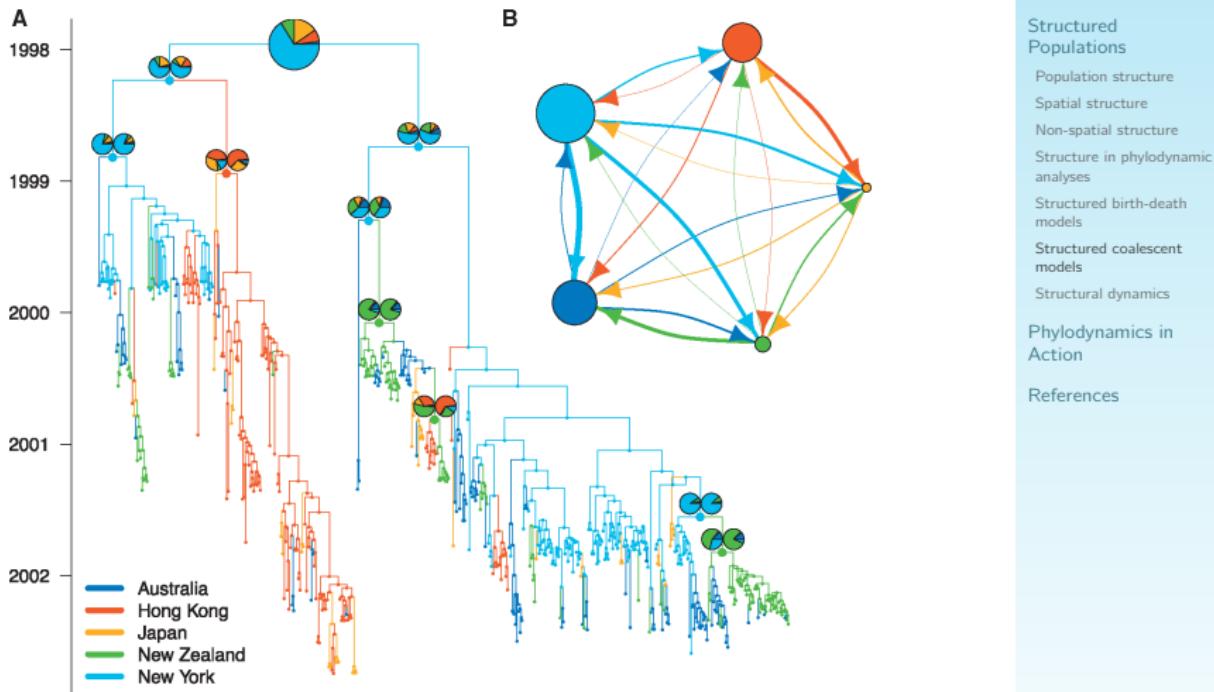
## Structured Populations

- Population structure
- Spatial structure
- Non-spatial structure
- Structure in phylodynamic analyses
- Structured birth-death models
- Structured coalescent models
- Structural dynamics

## Phyldynamics in Action

## References

# Example: Seasonal influenza (H3N2)



Tree with ancestral locations inferred from influenza sequences sampled from five locations. Effective sub-population sizes and migration rates are also inferred. [Müller et al., 2018]

# Unsampled Types (Ghost Demes)

- ▶ It is often the case (particularly for spatial structure) that not all types/locations are sampled.

## Structured Populations

Population structure

Spatial structure

Non-spatial structure

Structure in phylodynamic analyses

Structured birth-death models

Structured coalescent models

Structural dynamics

## Phyldynamics in Action

## References

# Unsampled Types (Ghost Demes)

- ▶ It is often the case (particularly for spatial structure) that not all types/locations are sampled.
- ▶ Unsampled demes are known as “ghost demes” and ignoring them can lead to overestimates of sub-population size.

## Structured Populations

Population structure

Spatial structure

Non-spatial structure

Structure in phylodynamic analyses

Structured birth-death models

Structured coalescent models

Structural dynamics

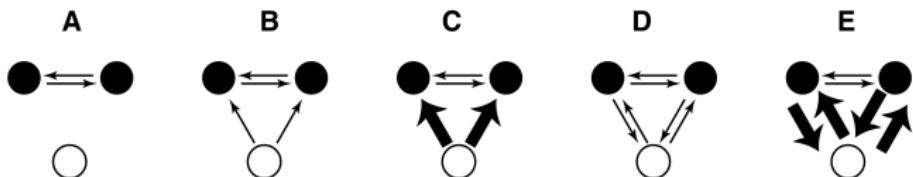
## Phyldynamics in Action

## References

# Unsampled Types (Ghost Demes)

<b>Structured Populations</b>
Population structure
Spatial structure
Non-spatial structure
Structure in phylodynamic analyses
Structured birth-death models
Structured coalescent models
Structural dynamics
<b>Phyldynamics in Action</b>
References

- ▶ It is often the case (particularly for spatial structure) that not all types/locations are sampled.
- ▶ Unsampled demes are known as “ghost demes” and ignoring them can lead to overestimates of sub-population size.



Example models with an explicit ghost deme. [Beerli, 2004]

# Populations with dynamic structure

Coalescent models have also been extended to account for **structure** which changes through time.

## Structured Populations

- Population structure
- Spatial structure
- Non-spatial structure
- Structure in phylodynamic analyses
- Structured birth-death models
- Structured coalescent models
- Structural dynamics

## Phyldynamics in Action

## References

# Populations with dynamic structure

Coalescent models have also been extended to account for **structure** which changes through time.

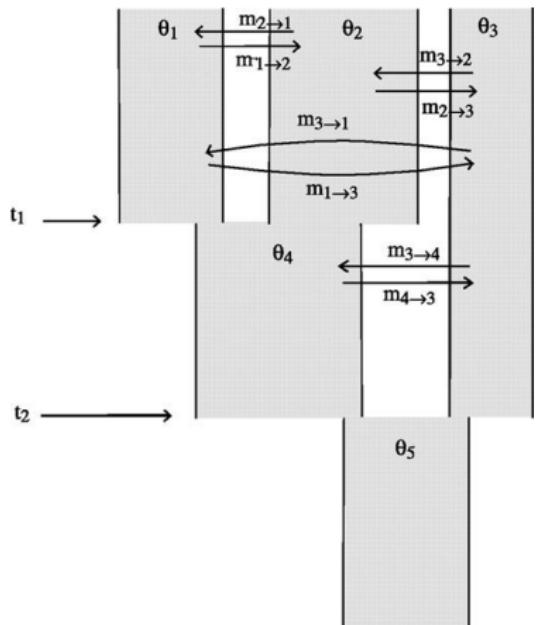


Figure adapted from [Hey, 2009]

Isolation-with-migration model.

## Structured Populations

- Population structure
- Spatial structure
- Non-spatial structure
- Structure in phydodynamic analyses
- Structured birth-death models
- Structured coalescent models
- Structural dynamics

## Phyldynamics in Action

## References

# Populations with dynamic structure

Coalescent models have also been extended to account for **structure** which changes through time.

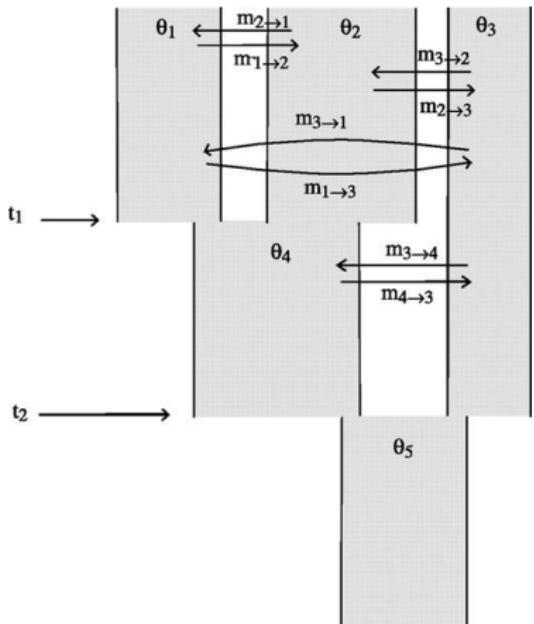


Figure adapted from [Hey, 2009]

Isolation-with-migration model.

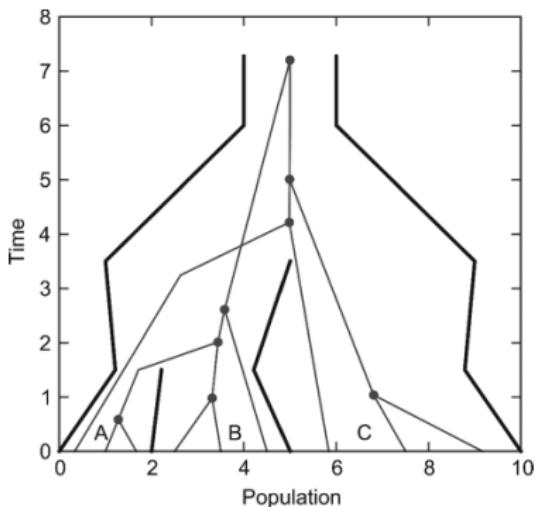


Figure adapted from  
[Heled and Drummond, 2010]

Multi-species coalescent model.

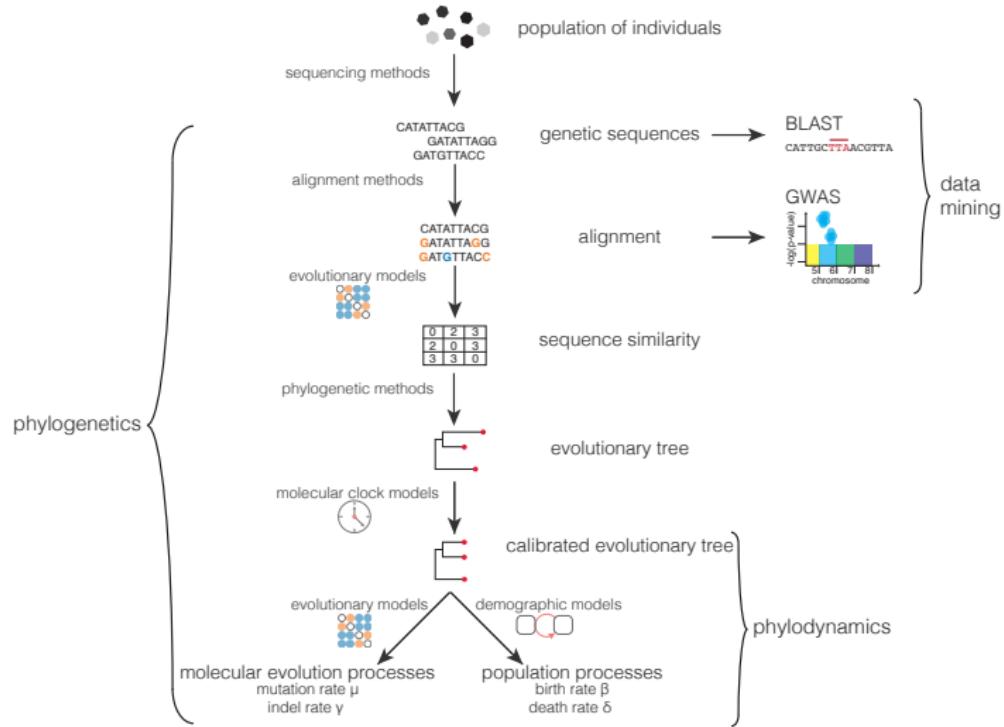
## Structured Populations

- Population structure
- Spatial structure
- Non-spatial structure
- Structure in phydodynamic analyses
- Structured birth-death models
- Structured coalescent models
- Structural dynamics

## Phyldynamics in Action

## References

# Phylodynamics and Phylogenetics in Action



## Structured Populations

- Population structure
- Spatial structure
- Non-spatial structure
- Structure in phylodynamic analyses
- Structured birth-death models
- Structured coalescent models
- Structural dynamics

## Phylodynamics in Action

### References

# Macroevolution

individuals = species

## ► **(Molecular) Evolution**

Genetic information and morphology of species changes through time.

## ► **Phylogenetics**

Phylogeny displays species relationship.

## ► **Phylogenetics**

Population dynamics is the speciation and extinction process.

Examples: Dinosaurs; penguins.

### Structured Populations

Population structure

Spatial structure

Non-spatial structure

Structure in phylodynamic analyses

Structured birth-death models

Structured coalescent models

Structural dynamics

### Phylodynamics in Action

### References

# Epidemiology

individuals = infected hosts

## ► **Molecular Evolution**

Genetic information of pathogens changes through time.

## ► **Phylogenetics**

Phylogeny displays transmission history.

## ► **Phylodynamics**

Population dynamics is the transmission and recovery process.

Examples: Ebola, HCV, HIV, Zika.

### Structured Populations

Population structure

Spatial structure

Non-spatial structure

Structure in phylodynamic analyses

Structured birth-death models

Structured coalescent models

Structural dynamics

### Phylodynamics in Action

### References

# Immunology: antibody response

individuals = B cells

## ► Molecular Evolution

B cells change through time due to recombination and somatic hypermutation, as response to pathogen exposure.

## ► Phylogenetics

Phylogeny displays B cell evolution.

## ► Phylodynamics

Population dynamics is the B cell generation and loss process.

### Structured Populations

Population structure

Spatial structure

Non-spatial structure

Structure in phylodynamic analyses

Structured birth-death models

Structured coalescent models

Structural dynamics

### Phylodynamics in Action

### References

# Immunology

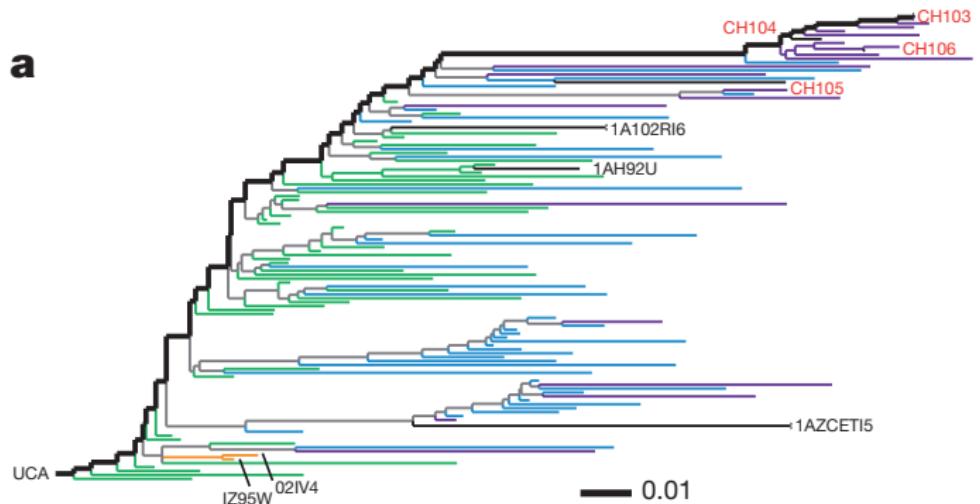
**a**

Figure adapted from [Liao et al., 2013]

Phylogeny of B-cells of an HIV infected individual. Future challenge is to investigate co-evolution of HIV and B-cell response.

## Structured Populations

- Population structure
- Spatial structure
- Non-spatial structure
- Structure in phylodynamic analyses
- Structured birth-death models
- Structured coalescent models
- Structural dynamics

## Phylodynamics in Action

### References

# Developmental Biology

individuals = cells of a multicellular organism

## ► “Evolution”

Cell types change from stem cells to highly specialized cells.

## ► Phylogenetics

Phylogeny displays differentiation of cells through time.

## ► Phydynamics

Population dynamics is the gain and loss of cell types.

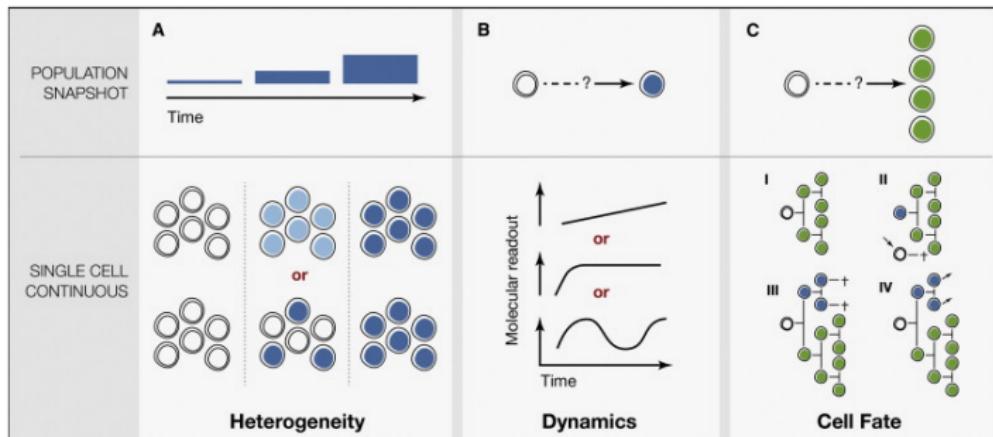


Figure adapted from [Etzrodt et al., 2014]

## Structured Populations

- Population structure
- Spatial structure
- Non-spatial structure
- Structure in phydodynamic analyses
- Structured birth-death models
- Structured coalescent models
- Structural dynamics

## Phydynamics in Action

## References

# Human migration

individuals = human populations

- ▶ **Evolution:** Human genomes evolve slowly and recombination makes analysis very hard!

*Solution:* Study portions of the genome that do not recombine (Y chromosome + mtDNA)

- ▶ **Phylogenetics**

Phylogeny displays genetic relationships between human populations.

- ▶ **Phylodynamics**

Population dynamics is the migration process out of Africa.

## Structured Populations

Population structure

Spatial structure

Non-spatial structure

Structure in phylodynamic analyses

Structured birth-death models

Structured coalescent models

Structural dynamics

## Phylodynamics in Action

## References

# Human migration

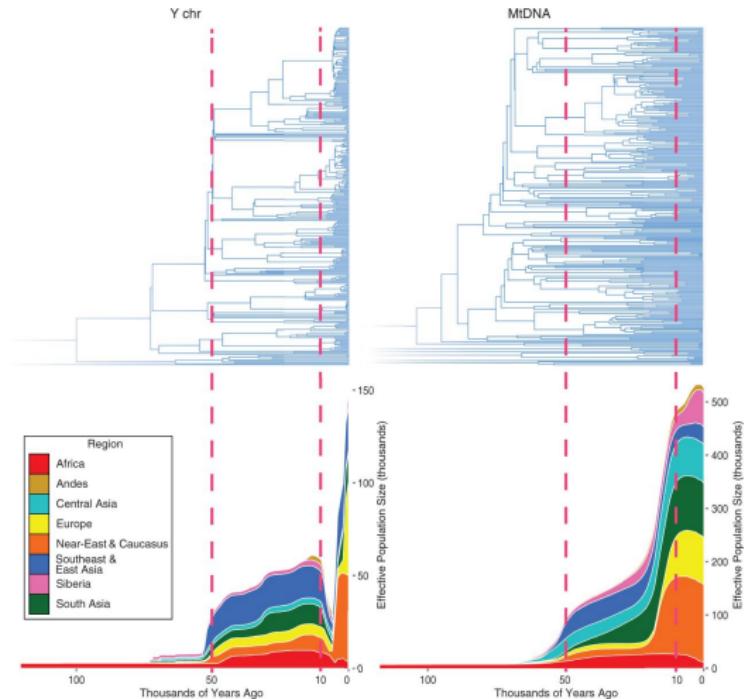


Figure adapted from [Karmin et al., 2015]

Bayesian skyline plot population dynamic inferences from Y chromosomes and mtDNA. (GTR substitution model.)

## Structured Populations

- Population structure
- Spatial structure
- Non-spatial structure
- Structure in phylodynamic analyses
- Structured birth-death models
- Structured coalescent models
- Structural dynamics

## Phylodynamics in Action

### References

# Language evolution

individuals = languages

## ► Evolution

Words and letters change through time.

## ► Phylogenetics

Phylogeny displays language history.

## ► Phylodynamics

Population dynamics is the gain and loss of languages.

### Structured Populations

Population structure

Spatial structure

Non-spatial structure

Structure in phylodynamic analyses

Structured birth-death models

Structured coalescent models

Structural dynamics

### Phylodynamics in Action

### References

<b>English</b>	<i>all</i>	<i>sea</i>	<i>water</i>	<i>when</i>
<b>German</b>	<i>alle</i>	<i>see, meer</i>	<i>wasser</i>	<i>wann</i>
<b>French</b>	<i>tout</i>	<i>mer</i>	<i>eau</i>	<i>quand</i>
<b>Italian</b>	<i>tutto</i>	<i>mare</i>	<i>acqua</i>	<i>quando</i>
<b>Greek</b>	<i>pant</i>	<i>thalasa</i>	<i>nero</i>	<i>pote</i>
<b>Hittite</b>	<i>dapiya</i>	<i>aruna-</i>	<i>watar</i>	<i>kuwapi</i>

Figure adapted from <http://language.cs.auckland.ac.nz/what-we-did>

# Language evolution

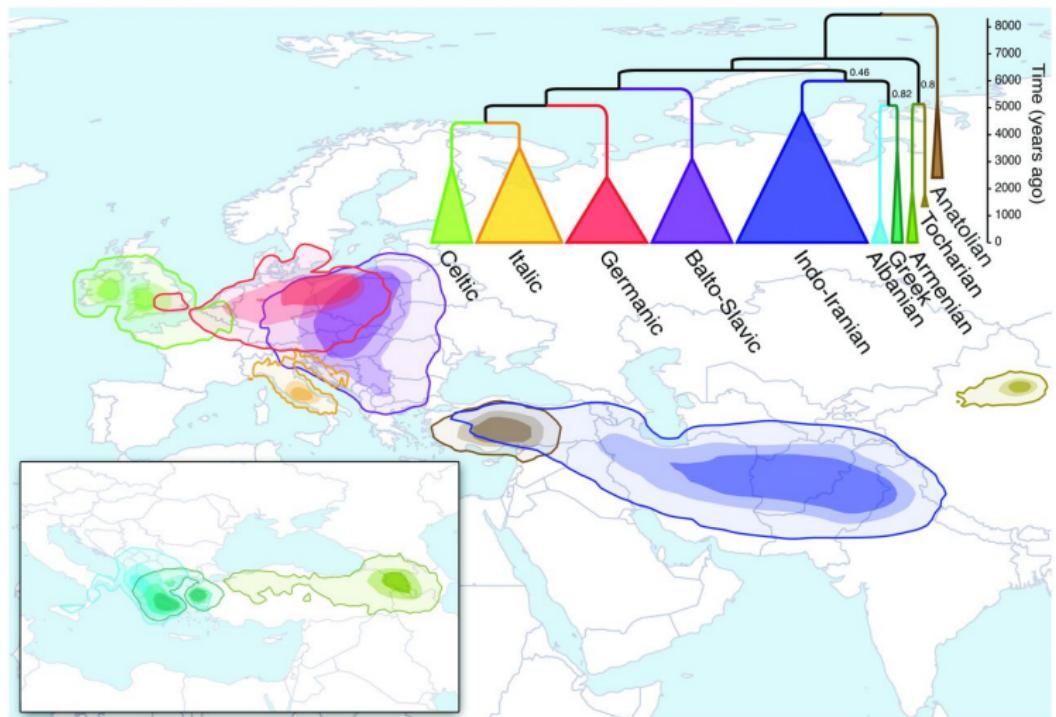


Figure adapted from [Bouckaert et al., 2012]

Mapping the origins and expansion of the Indo-European language family using a structured phylodynamic model.

- Structured Populations
- Population structure
- Spatial structure
- Non-spatial structure
- Structure in phylodynamic analyses
- Structured birth-death models
- Structured coalescent models
- Structural dynamics
- Phylodynamics in Action
- References

# Cultural evolution

individuals = human populations

- ▶ **Political systems:** How does the complexity of systems change over time [Currie et al., 2010]?
- ▶ **Religion:** Are social structures correlated with certain religious practices, such as ritual human sacrifice? [Watts et al., 2016]?

## Structured Populations

Population structure  
Spatial structure  
Non-spatial structure  
Structure in phylodynamic analyses  
Structured birth-death models

Structured coalescent models  
Structural dynamics

## Phylodynamics in Action

## References

# Cultural evolution: political systems

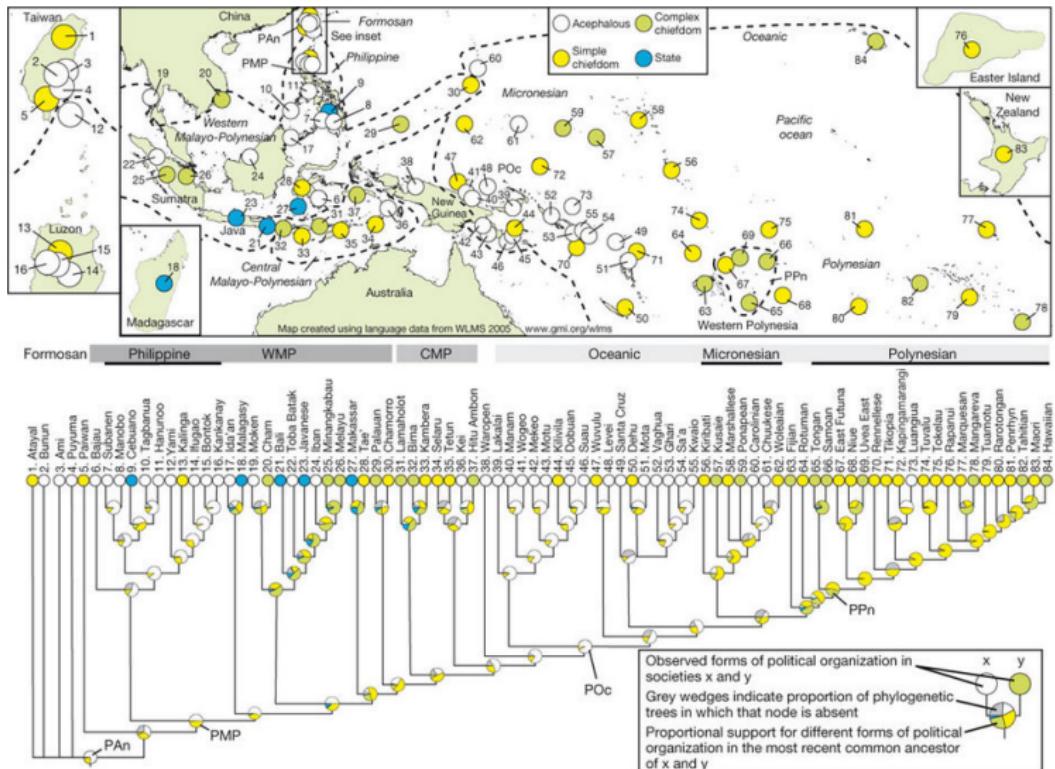


Figure adapted from [Currie et al., 2010]

## Structured Populations

- Population structure
- Spatial structure
- Non-spatial structure
- Structure in phylodynamic analyses
- Structured birth-death models
- Structured coalescent models
- Structural dynamics

## Phylodynamics in Action

### References

# Cultural evolution: political systems

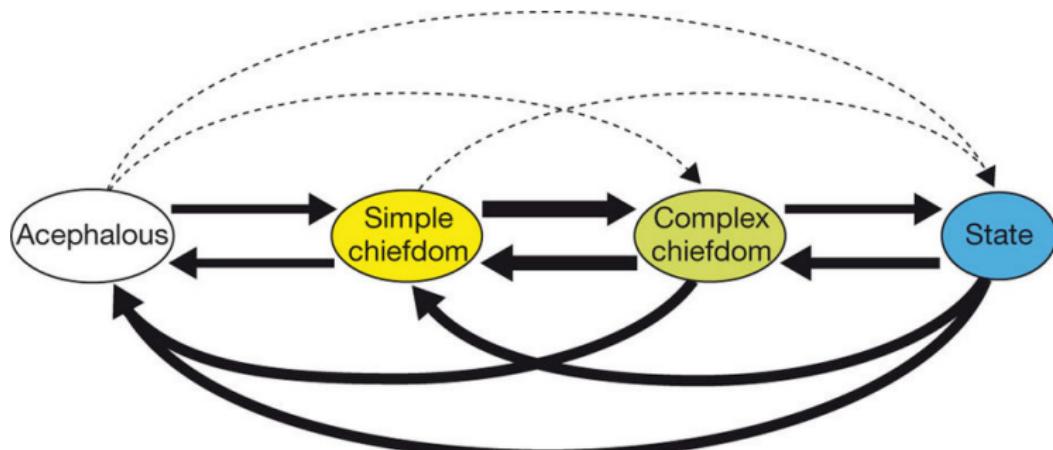


Figure adapted from [Currie et al., 2010]

Inferred rate of transitions between systems (thicker lines represent faster rates).

## Structured Populations

Population structure

Spatial structure

Non-spatial structure

Structure in phylodynamic analyses

Structured birth-death models

Structured coalescent models

Structural dynamics

## Phylodynamics in Action

## References

## Cultural evolution: religious practice

CB

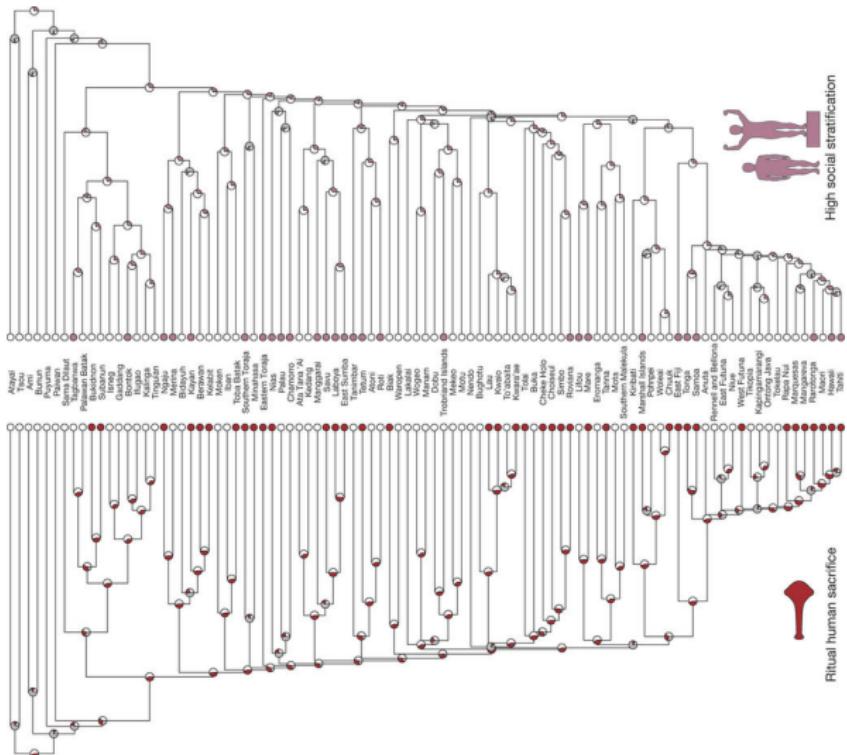


Figure adapted from [Watts et al., 2016]

Evolution of ritual human sacrifice (bottom) is correlated with evolution of social stratification (top).

## Structured Populations

## Population structure

## Spatial structure

## Non-spatial structure

## Structure in phylodynamic analyses

## Structured birth-death models

## Structured coalescent models

## Phylogenetics in Action

### References

# Questions

- ?
- Under a structured birth-death model, how do the sub-population sizes vary (if at all) through time?
- ?
- Suppose you perform a structured coalescent analysis on sequences collected from a relatively *unstructured* population. Would you expect the posterior migration rate to be very low or very high? Why?
- ?
- How might the evolution of languages violate the assumptions of a substitution+birth/death phylodynamic model?

## Structured Populations

- Population structure
- Spatial structure
- Non-spatial structure
- Structure in phylodynamic analyses
- Structured birth-death models

## Structured coalescent models

## Structural dynamics

## Phyldynamics in Action

## References

# References I

- Barido-Sottani, J., Vaughan, T. G., and Stadler, T. (2018). A multi-state birth-death model for Bayesian inference of lineage-specific birth and death rates. *bioRxiv preprint*.
- Beerli, P. (2004). Effect of unsampled populations on the estimation of population sizes and migration rates between sampled populations. *Molecular Ecology*, 13(4):827–836.
- Bouckaert, R., Lemey, P., Dunn, M., Greenhill, S. J., Alekseyenko, A. V., Drummond, A. J., Gray, R. D., Suchard, M. A., and Atkinson, Q. D. (2012). Mapping the origins and expansion of the indo-european language family. *Science*, 337(6097):957–960.
- Currie, T. E., Greenhill, S. J., Gray, R. D., Hasegawa, T., and Mace, R. (2010). Rise and fall of political complexity in island south-east asia and the pacific. *Nature*, 467(7317):801–804.
- De Maio, N., Wu, C.-H., O'Reilly, K. M., and Wilson, D. (2015). New routes to phylogeography: A Bayesian structured coalescent approximation. *PLoS Genetics*, 11(8):e1005421.
- Etzrodt, M., Endele, M., and Schroeder, T. (2014). Quantitative single-cell approaches to stem cell research. *Cell stem cell*, 15(5):546–558.
- FitzJohn, R. G., Maddison, W. P., and Otto, S. P. (2009). Estimating trait-dependent speciation and extinction rates from incompletely resolved phylogenies. *Systematic Biology*, 58(6):595–611.
- Heled, J. and Drummond, A. J. (2010). Bayesian inference of species trees from multilocus data. *Mol Biol Evol*, 27(3):570–80.
- Hey, J. (2009). Isolation with migration models for more than two populations. *Molecular biology and evolution*, 27(4):905–920.
- Karmin, M., Saag, L., Vicente, M., Sayres, M. A. W., Järve, M., Talas, U. G., Roots, S., Ilumäe, A.-M., Mägi, R., Mitt, M., et al. (2015). A recent bottleneck of y chromosome diversity coincides with a global change in culture. *Genome research*, 25(4):459–466.
- Kühnert, D., Stadler, T., Vaughan, T. G., and Drummond, A. J. (2016). Phylodynamics with migration: A computational framework to quantify population structure from genomic data. *Molecular biology and evolution*, page msw064.
- Liao, H.-X., Lynch, R., Zhou, T., Gao, F., Alam, S. M., Boyd, S. D., Fire, A. Z., Roskin, K. M., Schramm, C. A., Zhang, Z., et al. (2013). Co-evolution of a broadly neutralizing hiv-1 antibody and founder virus. *Nature*, 496(7446):469–476.
- Maddison, W. P., Midford, P. E., and Otto, S. P. (2007). Estimating a binary character's effect on speciation and extinction. *Syst Biol*, 56(5):701–710.

## Structured Populations

- Population structure
- Spatial structure
- Non-spatial structure
- Structure in phylodynamic analyses
- Structured birth-death models
- Structured coalescent models
- Structural dynamics

## Phyldynamics in Action

## References

# References II

- Müller, N. F., Rasmussen, D., and Stadler, T. (2018). MASCOT: parameter and state inference under the marginal structured coalescent approximation. *Bioinformatics*.
- Müller, N. F., Rasmussen, D. A., and Stadler, T. (2017). The structured coalescent and its approximations. *Molecular Biology and Evolution*, 34(11):2970–2981.
- Notohara, M. (1990). The coalescent and the genealogical process in geographically structured population. *J Math Biol*, 29(1):59–75.
- Pannell, J. R. (2003). Coalescence in a metapopulation with recurrent local extinction and recolonization. *Evolution*, 57(5):949–961.
- Scire, J., Barido-Sottani, J., Kühnert, D., Vaughan, T. G., and Stadler, T. Improved multi-type birth-death phylodynamic inference in BEAST 2. In preparation.
- Watts, J., Sheehan, O., Atkinson, Q. D., Bulbulia, J., and Gray, R. D. (2016). Ritual human sacrifice promoted and sustained the evolution of stratified societies. *Nature*, 532(7598):228–231.

## Structured Populations

Population structure

Spatial structure

Non-spatial structure

Structure in phylodynamic analyses

Structured birth-death models

Structured coalescent models

Structural dynamics

## Phyldynamics in Action

## References