

# Inference with selection, varying population size and emerging population structure

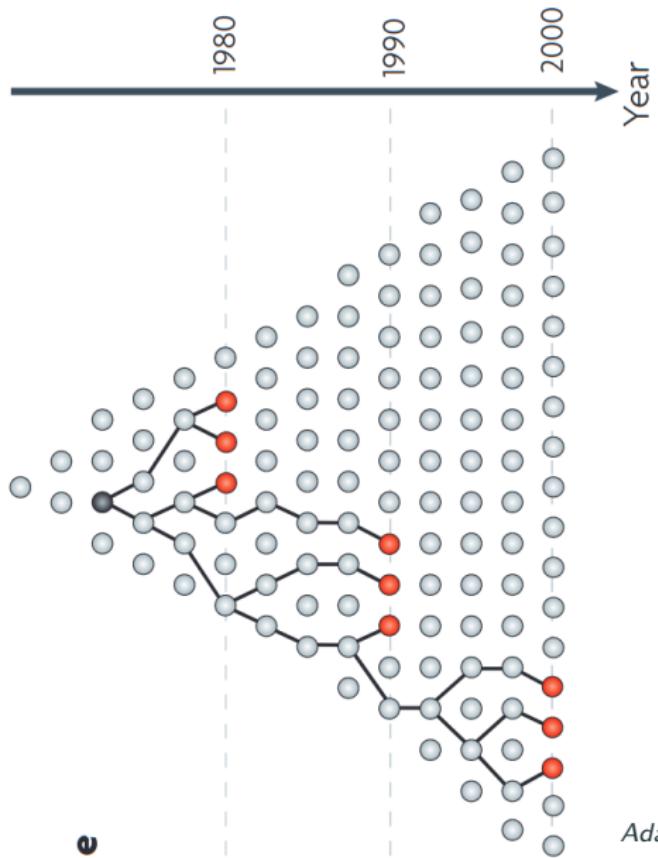
*ABC on a forward-backward coalescent with interactions*

---

Clotilde Lepers, Sylvain Billiard\*, Régis Ferrière, Matthieu Porte, Sylvie Méléard, Viet Chi Tran

\* Université de Lille - France - Lab. Evo-Eco-Paléo

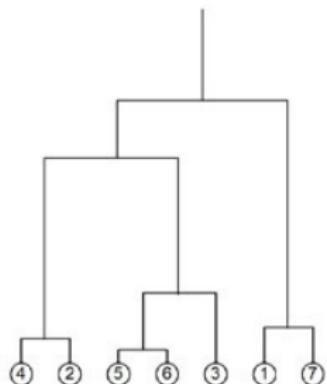
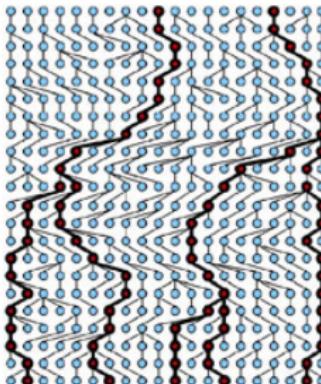
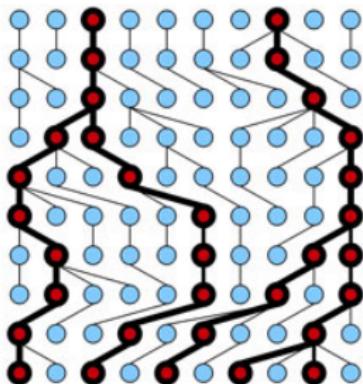
# An inverse problem



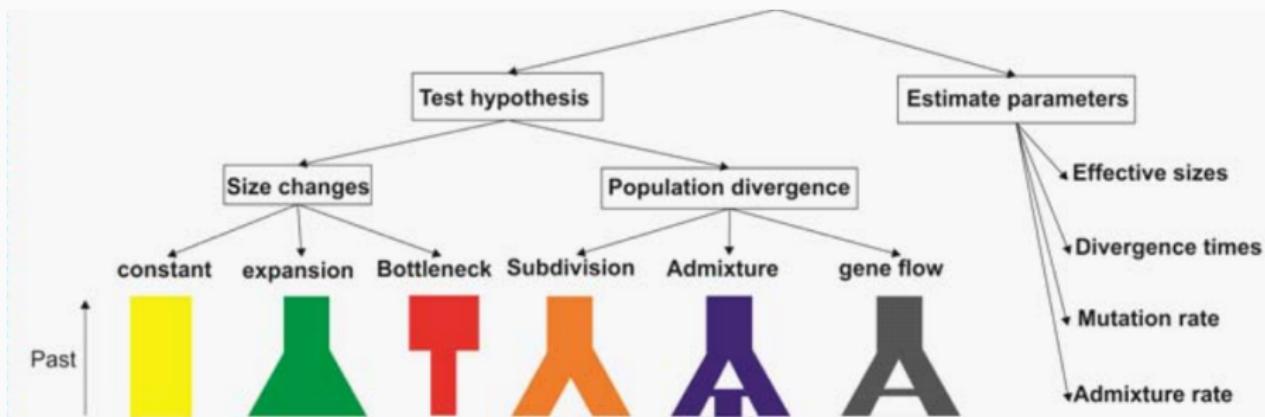
Adapted from Pybus and Rambaut-2009

# Kingman's coalescent and derivations

- Wright Fisher, Kingman-1982



# Applications: Kingman's coalescent and derivations

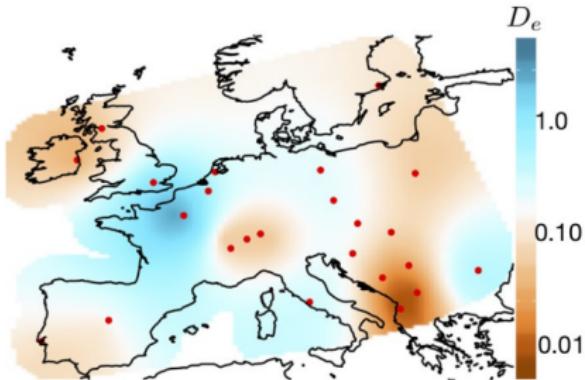
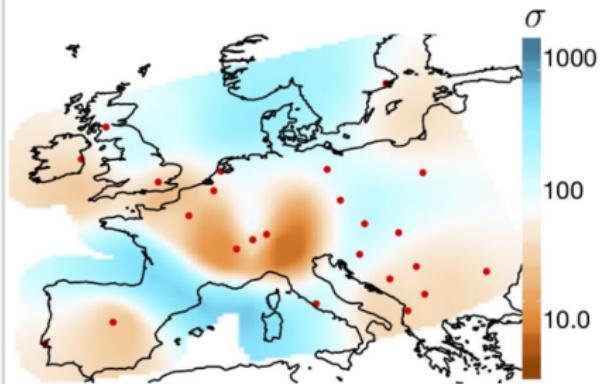


Adapted from Barrandeguy et al.-2017

## Example: Inferring dispersal and population densities

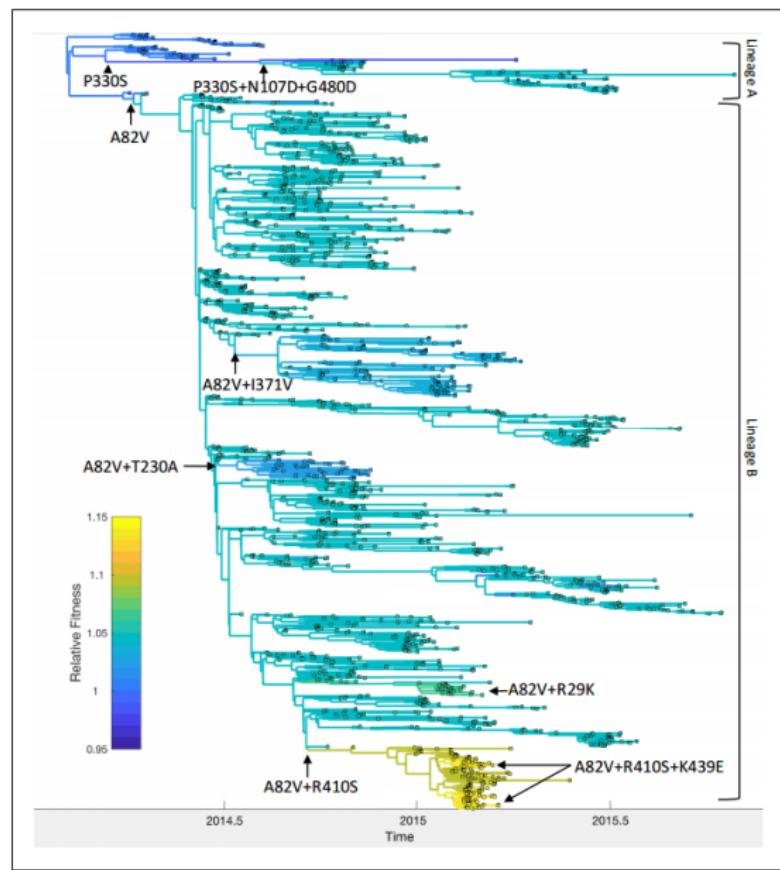
b)

5-10cm (~675 years)



Al-Asadi et al.-2019

# An inverse problem: multitype Birth-Death models



Rasmussen and Stadler-2019

## Limitations and drawbacks

1. Population size variation: exogenous and *a priori* known
2. Population structure: *idem*
3. No Interactions *within* classes, No interactions *between* classes

## Our goals: Inference from a model where...

1. Population size variation: *emerge* from individuals interactions
2. Population structure: *idem*
3. Interactions within *and* between classes

# **1. A stochastic individual-based model**

---

# Assumptions

- Continuous time, competition, asexual, mutations at birth
- Population structured in trait  $x$  and marker  $u$ 
  - Birth rate  $b(x)$
  - Death rate  $d(x) + \frac{\eta(x)}{K} \sum_{i=1}^{N_t^K} C(x - x_i)$
- Trait (rare) mutations probability  $p_K \sim 1/K^2$ .
- Marker (rare) mutations probability  $q_K = a_K p_K$ .
- Marker space  $u \in \mathcal{U}$  e.g.  $\mathcal{U} = \{A, a\}, \mathcal{U} = \mathbb{R}, \dots$

## Three time-scales

---

- Evolution of trait  $x$ : the slowest.
- Evolution of the marker  $u$ : intermediate.
- Ecological processes (births, deaths, interactions): the fastest.

## Evolution of the marker in a structured population with trait $x$

- $F_t^u$ : Distribution of the 'neutral' marker
- Roughly speaking (large population)

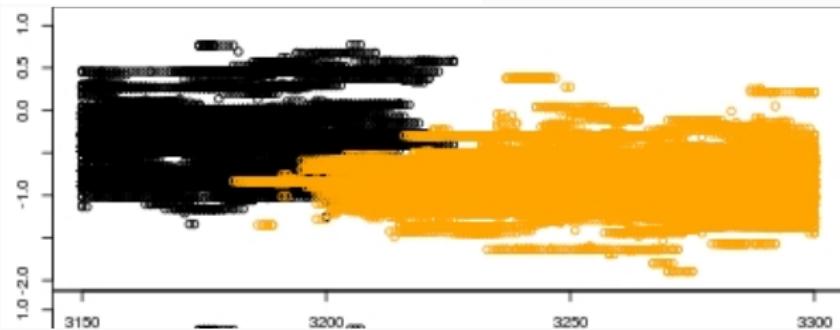
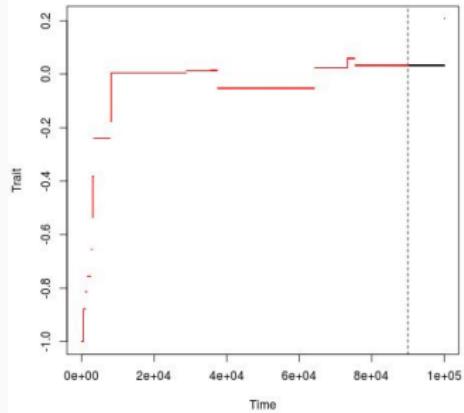
$$F_t^u = b(x) \int_0^t \text{Mutation } ds + M_t^x$$

where  $M_t^x$ : demographic fluctuations

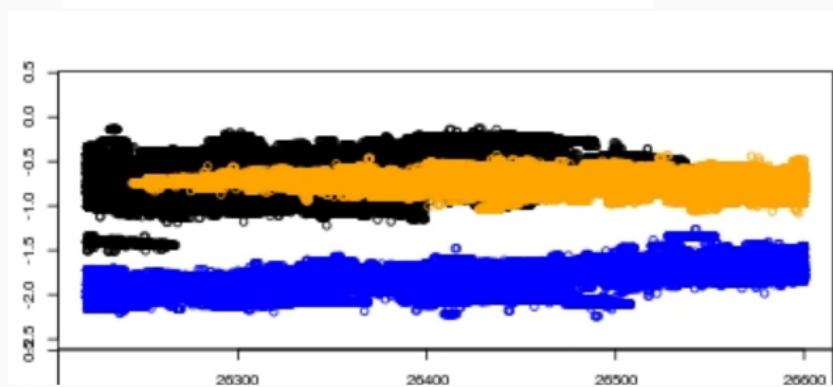
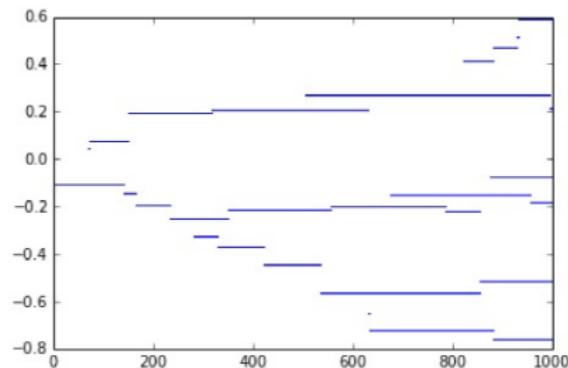
$$\text{Var}(M_t^x) = \frac{2b(x)}{\hat{n}(x; x_1, \dots, x_p)} \int_0^t \text{Variance of } F_s^u \ ds.$$

# Evolution of the marker in a structured population with trait $x$

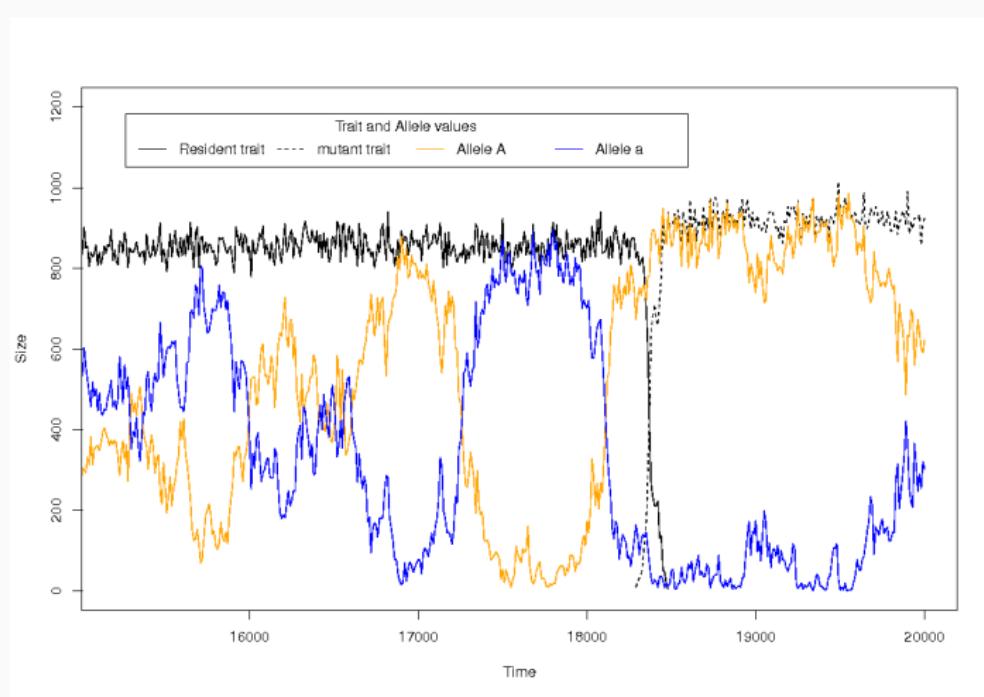
Neutral Marker = continuous trait



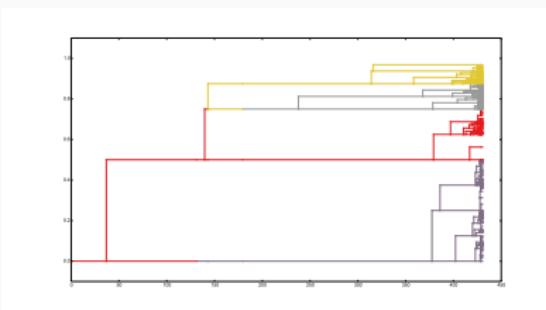
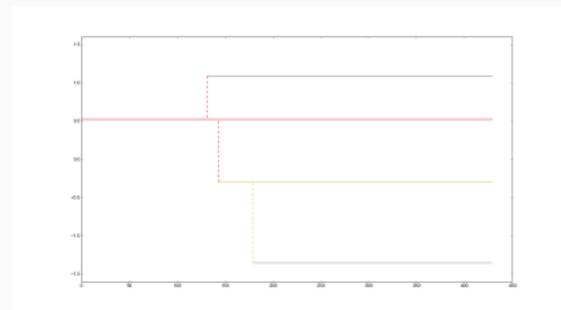
# Evolution of the marker in a structured population with trait $x$



# Evolution of the marker in a structured population with trait $\times$ Neutral Marker = 2 alleles



# Genealogy of a sample of $n$ individuals? A Forward-Backward Coalescent

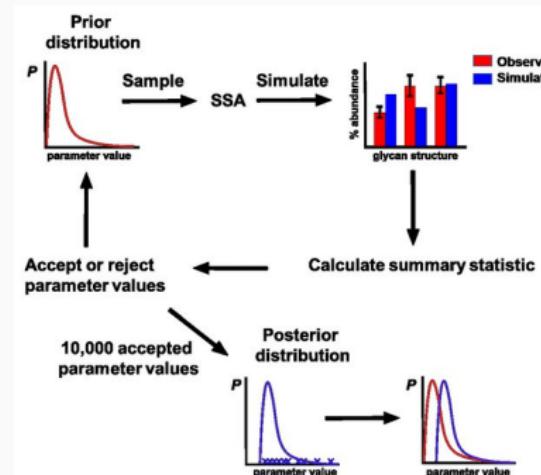


- 1. Forward: Adaptative jumps → trait genealogy
- 2. Backward: Coalescent within classes → 'neutral' marker genealogy.

## **2: Inference**

---

# Likelihood not explicit: Approximate Bayesian Computations



Beaumont et al.-2002, Blum-2010, Marin Pudlo et al.-2012.

- Draw parameters  $\theta_i$  in a prior ( $N$  times)
- $\theta_i \rightarrow$  simulation of neutral phylogenies in the model,
- Descriptive statistics  $S_i$  compared with data  $s_{obs}$
- Weight  $w_i$  for  $\theta_i$  depending on the distance  $d(S_i, s_{obs})$
- Weighted sample  $(\theta_i, w_i) \rightarrow$  posterior distribution.

# Descriptive statistics

## Pop structure from the total pop

- nb of coexisting sub-pop
- mean abundance of the coexisting sub-pop
- var of the abundances
- mean trait  $x$
- trait variance

## Pop structure from the sampled ind

- nb of sampled sub-pop
- relative abundance variance between sampled sub-pop
- mean trait of sample
- trait variance

## Pop structure from neutral markers

- Allele nbr

- M index
- Gene Diversity
- Unbiased Gene Diversity

## Statistics between sub-pop

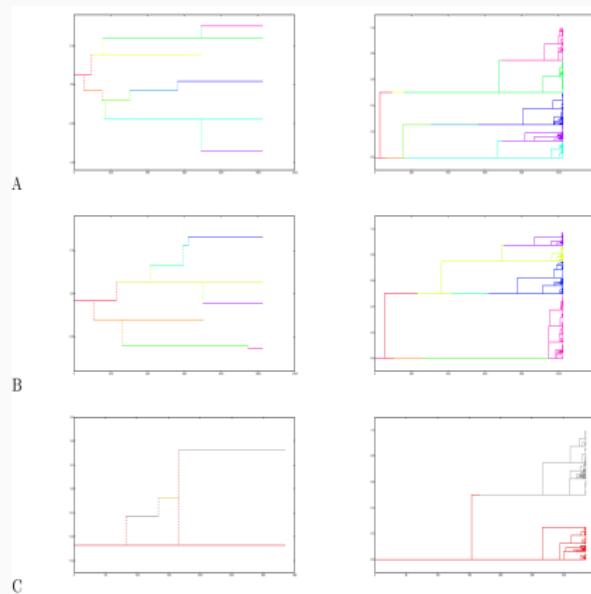
- $F_{ST}$
- $\delta\mu^2$
- Nei's  $D_A$  dist
- Nei's standard genetic dist  $D_S$

## Statistics measured on the whole PES

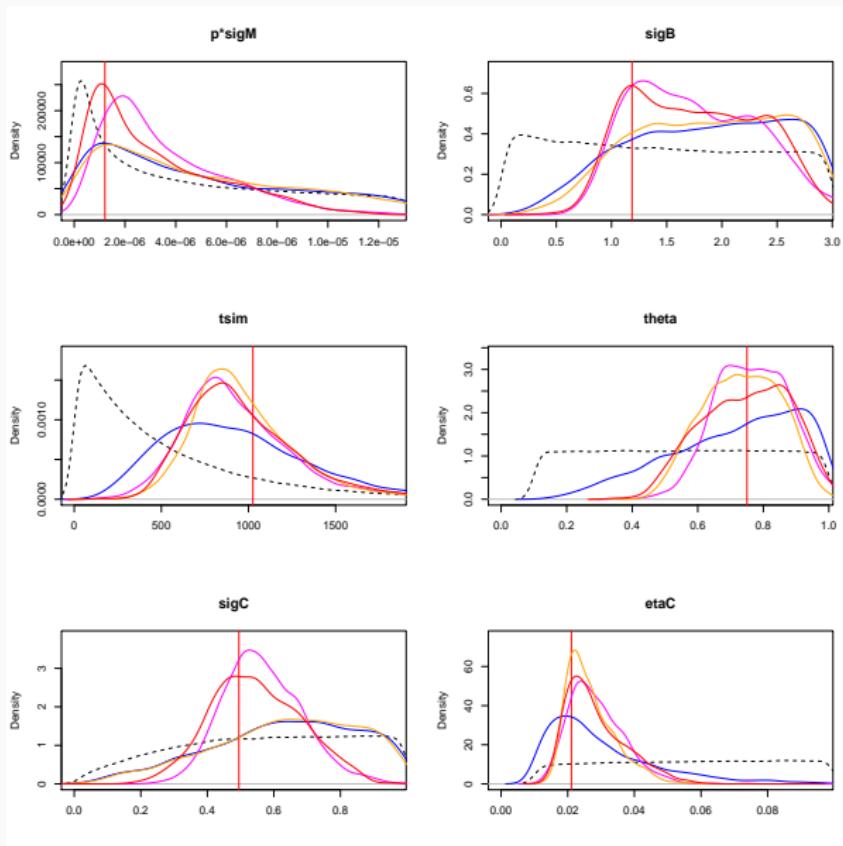
- nb of trait mutations
- Sackin's index.
- sum of branch lengths
- sum of external branch lengths
- nb of cherries
- time of MRCA

# Applications 1. Inference of the parameters in a toy model

- Birth rate:  $b(x) = \exp(-x^2/2\sigma_b^2)$
- Death rate:  $d(x) = d_C$
- Competition:  $C(x, y) = \eta_c \exp(-(x - y)^2/2\sigma_c^2)$
- Trait:  $x \in [-1, 1]$ , Marker:  $u \in [-2, 2]$
- Mutation kernel: truncated Gaussian

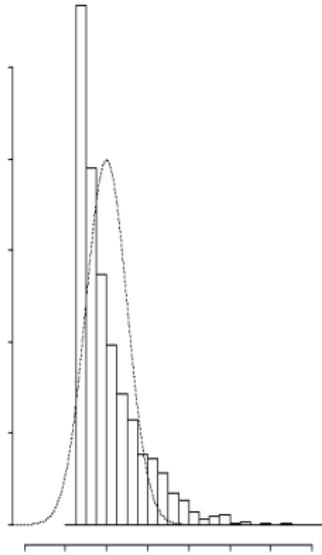


# Applications 1. Prior and posterior distributions

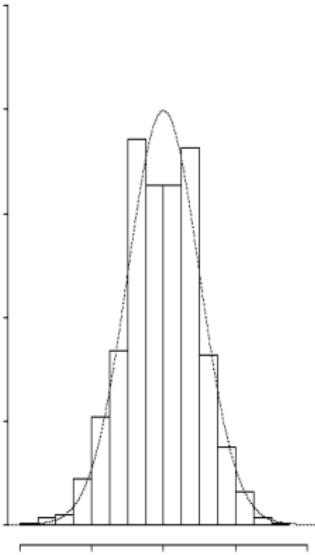


## Applications 1. Discrepancy from a Kingman's coalescent?

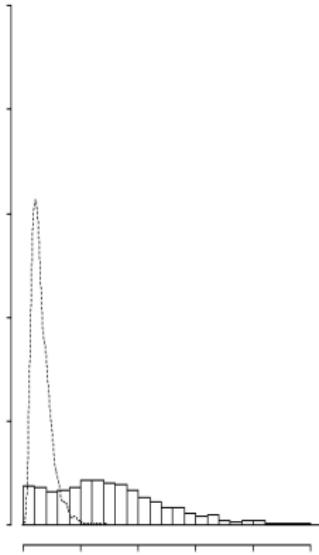
#### • External Branch Length



### • Number of cherries

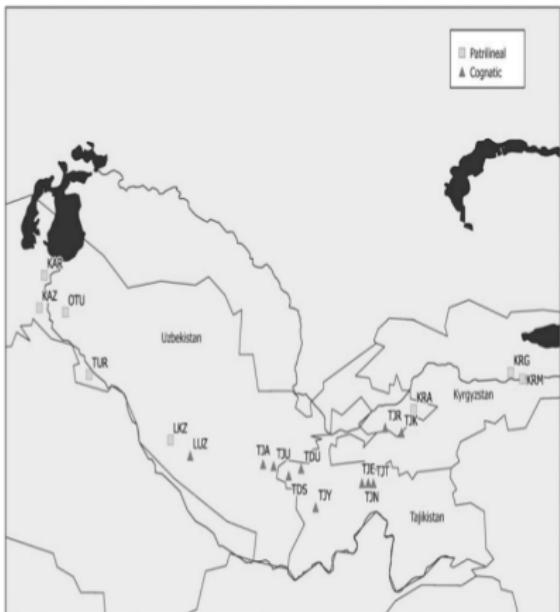


## • Time to MRCA



## Applications 2. Social and Genetic structures in Central Asia

- Different fertility rates between socio-cultural organizations?



- Two socio-cultural organizations
  - Pastoralists with descent groups, exogamous marriages
  - Farmers with nuclear families, endogamous marriages
- 19 populations, 631 individuals
- 10 microsatellites, ch. Y

Data from Chaix et al.-2007, Heyer et al.-2015

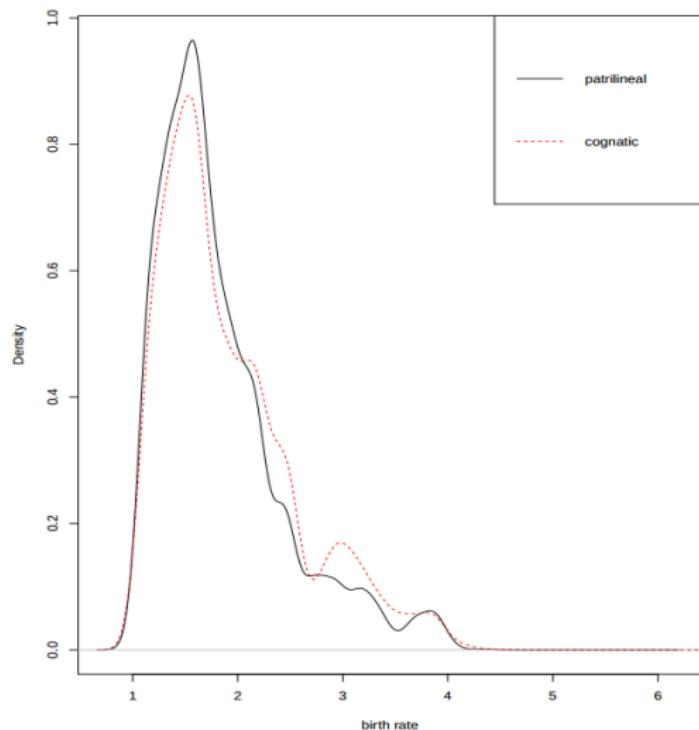
## Applications 2. A socio-demographic model

---

- Trait:  $x$ : demographic processes,  $y$ : geographical location
- Birth rate: Pastoralists  $b_0$ , Farmers  $b_1$
- Competition:
  - $C(y, y') = \eta_0 \exp(-(y - y')^2/2\sigma_c^2)$  (Pastoralists)
  - $C(y, y') = \eta_1 \exp(-(y - y')^2/2\sigma_c^2)$  (Farmers)
- Trait mutation:
  - socio-cultural switch (Pastoralist → Farmer)
  - dispersal and population foundations
- Neutral marker mutations

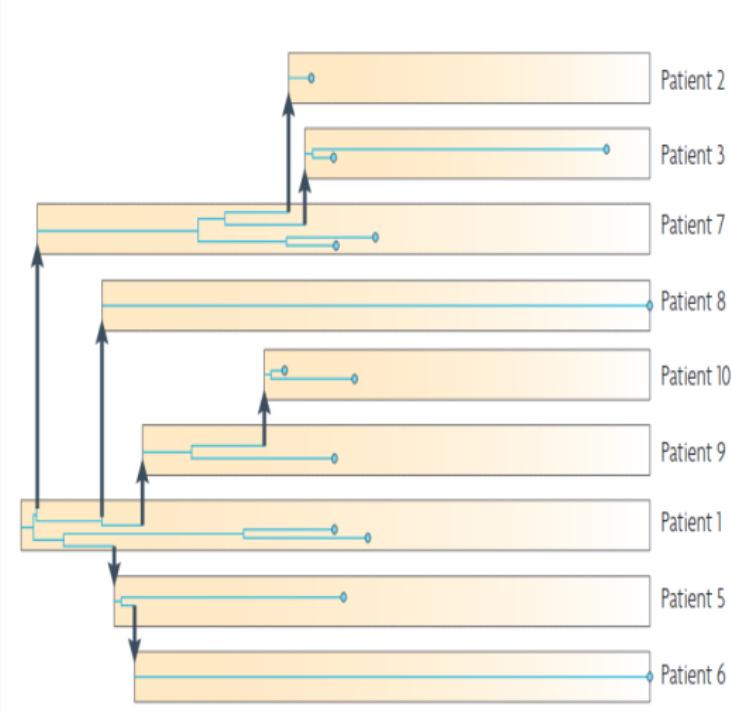
Question:  $b_0 < b_1$ ?

## Applications 2. Posterior distributions: No fertility differences



# Why bothering? 1. Inference in explicit ecological models

- e.g. HIV transmission and within-host evolutionary dynamics



# Why bothering? 2. Pursueing endogenization

Synthese

<https://doi.org/10.1007/s11229-018-1832-6>



CrossMark

S.I.: FIRST PRINCIPLES IN SCIENCE

## The strategy of endogenization in evolutionary biology

Samir Okasha<sup>1</sup> 

- Features cease to be part of the background and are brought within the fold of evolutionary theory
- How evolutionary biology has increased its explanatory power over time



Université  
de Lille

