

# Stochastic dynamics in adaptive biology

Sylvie Méléard, Ecole Polytechnique

BUENOS AIRES, 37th Conference on Stochastic Processes and  
their Applications - July 2014

S. Billiard, N. Champagnat, R. Ferrière, C.V. Tran



ANR **MANEGE**



# Darwinian or Adaptive Evolution

The population has the propensity to generate as well to select individual diversity.

The ability of an individual to survive and reproduce depends on some genetic characteristics, called **traits**.

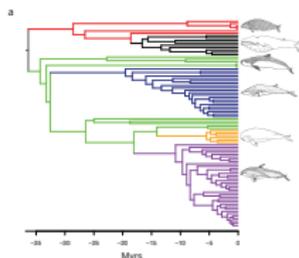
The evolution of the trait distribution results from three basic mechanisms.

- **Heredity**. Transmission of the ancestral trait to the offsprings.
- **Mutations**. Generate variability in the traits.
- **Natural selection**. Individuals with characteristics which increase their probability of survival or their reproduction ability will spread through the population over time.

Evolution consists in successive invasions of successful mutants.

- Much of the molecular diversity measured by population geneticists involve DNA sequences which are selectively neutral (not submitted to the natural selection), called **(neutral) markers**.

Example: Molecular phylogeny = genealogy of neutral markers.



- It is now widely accepted in phylogenies models that evolution rates among and along the branches can change.
- Question: How to explain these changes by the natural selection (without the use of external factors)?

## (1) Adaptive Dynamics: Successive invasions of successful mutants.

- Hofbauer-Sigmund 1990, Marrow-Law-Cannings 1992, Metz-Geritz-Meszéna et al. 1992, 1996, Dieckmann-Law 1996.
- Focus on the interplay between ecology and evolution and emphasizes the ecological interactions.
- The trait is a phenotypic parameter, with values in a continuum.
- The selection comes from a trade-off between reproduction ability and competition for resources.
- Large population and rare mutations.
- **Individual-based model** (birth and death process with mutation and selection): (Bolker-Pacala 97, Kisdi 99, Dieckmann-Law 00, Fournier-M. 04, Ferrière-Champagnat-M. 06, Champagnat 06, Champagnat-M. 10)

## (2) Population genetics approach: how neutral diversity is affected by selection?

- The population size and the selection parameter are constant (and fixed a priori) and independent.
- The traits and markers are alleles on different loci.
- Abundant literature.

Barton, Etheridge-Pfaffelhüber-Wakolbinger,  
Durrett-Schweinsberg.

**Our goal: To construct a stochastic dynamics of adaptive trait and neutral marker driven by eco-evolutionary feedbacks**

# Individual-based model

- Phenotypic trait under selection  $x$  in a subset  $\mathcal{X}$  of  $\mathbb{R}^d$  (rate of nutrient intake, body size at maturity, age at maturity ...).
- Neutral marker  $u$  in a subset  $\mathcal{U}$  of  $\mathbb{R}^\ell$ .
- The type of an individual  $i$ :  $(x_i, u_i)$ .
- $K$  scales the size of the population.
- $p_K$  scales the mutation probability of the traits under selection.
- $q_K$  scales the mutation probability of the neutral markers.
- Population of  $N^K(t)$  individuals weighted by  $\frac{1}{K}$ .

It is represented by the point measure

$$\nu_t^K = \frac{1}{K} \sum_{i=1}^{N^K(t)} \delta_{(x_i, u_i)} \quad ; \quad N^K(t) = K \langle \nu_t^K, \mathbf{1} \rangle.$$

# Transitions

## BIRTHS:

Each individual with characteristics  $(x, u)$  gives birth to a single individual at (inhomogeneous) rate  $b(x)$ ;  $0 \leq b(x) \leq \bar{b}$ .

At each birth time:

- with probability  $(1 - p_K)(1 - q_K)$ , the offspring inherits of  $(x, u)$ . (Clonal reproduction)
- Otherwise mutations on trait and marker occur independently with probability  $p_K$  and  $q_K$ .

**Main assumption: the markers mutate much faster than the traits under selection.**

$$p_K \sim \frac{1}{K^2}; \quad q_K = r_K p_K; \quad q_K \rightarrow 0; \quad r_K \rightarrow +\infty.$$

- Trait mutation: the new trait is  $x + k$  chosen according to  $m(x, k)dk$ .

- Marker mutation: the new marker is  $u + h$  chosen according to  $G_K(u, dh)$ .
- $\exists (A, \mathcal{D}(A))$  the infinitesimal generator of a Markovian semigroup such that  $\forall g \in \mathcal{D}(A)$

$$\lim_K \sup_u \left| \frac{r_K}{K} \int (g(u+h) - g(u)) G_K(u, dh) - Ag \right| = 0.$$

**Three timescales: ecological births and deaths, marker mutations, trait mutations.**

## Examples

- **Example 1** :  $\mathcal{U} = [u_1, u_2]$  and  $G_K \sim N(\mu_K, \sigma_K^2)$  with  $\mu_K \rightarrow 0$  and

$\sigma_K \rightarrow 0$  and  $\lim_K \frac{r_K \mu_K}{K} = \mu > 0$  ;  $\lim_K \frac{r_K \sigma_K^2}{K} = \sigma^2$ . Then

$$Af = \mu f' + \frac{\sigma^2}{2} f'' \quad \text{for } f \in C^2 \text{ and } f'(u_1) = f'(u_2) = 0.$$

- **Example 2** :  $\mathcal{U} = \mathbb{R}$  and  $G_K$  Pareto law with index  $\alpha \in (1, 2)$  renormalized by  $K^{\frac{\eta}{\alpha}}$  such that  $\eta < 1$ ,  $\lim_K \frac{r_K}{K^{1+\eta}} = \bar{r}$  and

$$Af = \bar{r} \int (f(u+h) - f(u) - hf'(u)1_{|h|\leq 1}) \frac{dh}{|h|^{1+\alpha}}.$$

- **Example 3** :  $\mathcal{U} = \{0, 1\}$  and

$G_K(u, dh) = 1_{u=0} q_0 \delta_1(dh) + 1_{u=1} q_1 \delta_{-1}(dh)$  ;  $\lim_K \frac{r_K}{K} = \bar{r}$ .  
Then

$$Af(u) = \bar{r} \left( 1_{u=0} q_0 (f(1) - f(0)) + 1_{u=1} q_1 (f(0) - f(1)) \right).$$

## DEATHS:

- Each individual with characteristics  $(x, u)$  dies at rate

$$d(x) + \frac{1}{K} \sum_{i=1}^{N^K(t)} C(x - x_i) = d(x) + C * \nu_t^K(x).$$

- The term  $C * \nu_t^K(x)$  describes the competition pressure for external resources.

## Assumptions:

(i)

$$b(x) - d(x) > 0; \quad C(x - y) \geq \underline{c} > 0.$$

(ii) The ecological functions  $b, d, C$  are bounded continuous and the mutation measures  $m(\cdot, k)dk$  and  $G_K(\cdot, h)dh$  are continuous.

*Moment conditions propagate and imply the existence and uniqueness of the process.*

## Marker and trait population process

$$\nu_t^K = \frac{1}{K} \sum_{i=1}^{N^K(t)} \delta_{(x_i, u_i)} \in \left\{ \frac{1}{K} \sum_{i=1}^n \delta_{(x_i, u_i)} ; n \geq 0, (x_1, u_1), \dots, (x_n, u_n) \in \mathcal{X} \times \mathcal{U} \right\}.$$

Trait marginal measure (on  $\mathcal{X}$ ):

$$X_t^K(dx) = \frac{1}{K} \sum_{i=1}^{N^K(t)} \delta_{x_i} = \int_{\mathcal{U}} \nu_t^K(dx, du).$$

Marker distribution (probability on  $\mathcal{U}$ ) for a given trait value  $x$ :

$$\pi_t^K(x, du) = \frac{\sum_{i=1}^{N^K(t)} \mathbf{1}_{x_i=x} \delta_{u_i}}{\sum_{i=1}^{N^K(t)} \mathbf{1}_{x_i=x}}.$$

We get

$$\nu_t^K(dx, du) = X_t^K(dx) \pi_t^K(x, du).$$

**Behavior when  $K$  tends to infinity?**

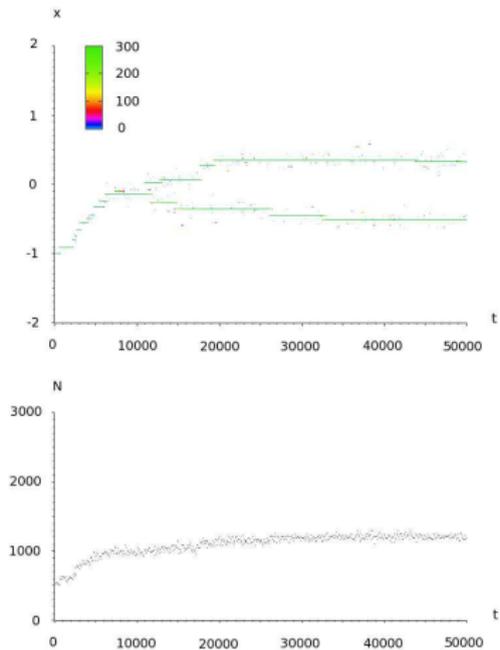
# Numerical Example

Inspired by a model of beak's size (Dieckmann-Doebeli 1999).

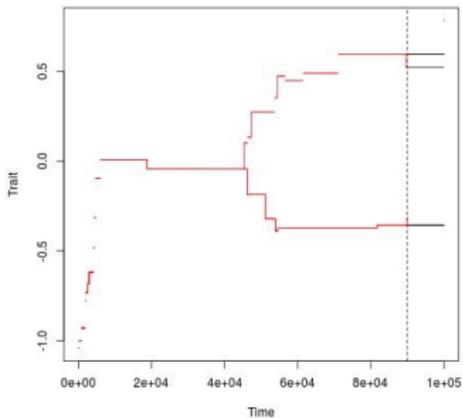
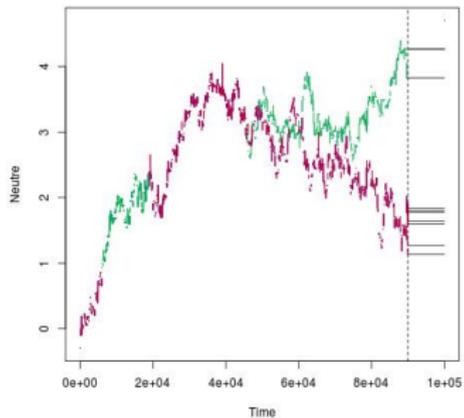
- $\mathcal{X} = [-2, 2]$  ;  $\mathcal{U} = [-2, 2]$  .
- $p_K = \frac{1}{K^2}$  ;  $r_K = K^{3/2}$  ;  $q_K = \frac{1}{\sqrt{K}}$  .
- $m \sim N(0, 10^{-1})$  and  $G_K \sim N(0, \frac{1}{\sqrt{K}})$  are two Gaussian laws conditioned to  $[-2, 2]$ .
- $b(x) = \exp(-x^2/2\sigma_b^2)$  ;  $\sigma_b = 0.9$ , maximum at 0.
- Symmetric competition for resources.

$$C(x - y) = \exp(-(x - y)^2/2\sigma_C^2) , \sigma_C = 0.8.$$

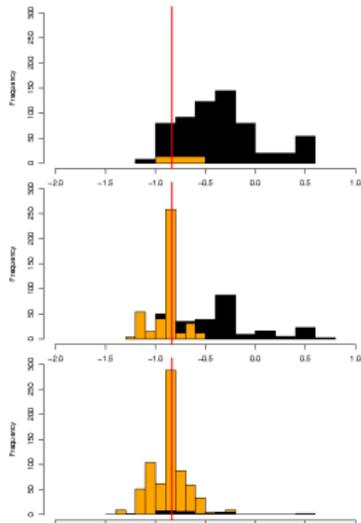
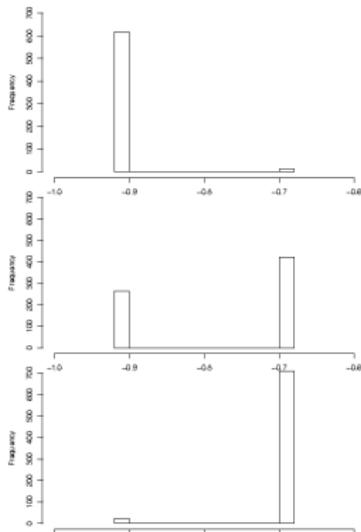
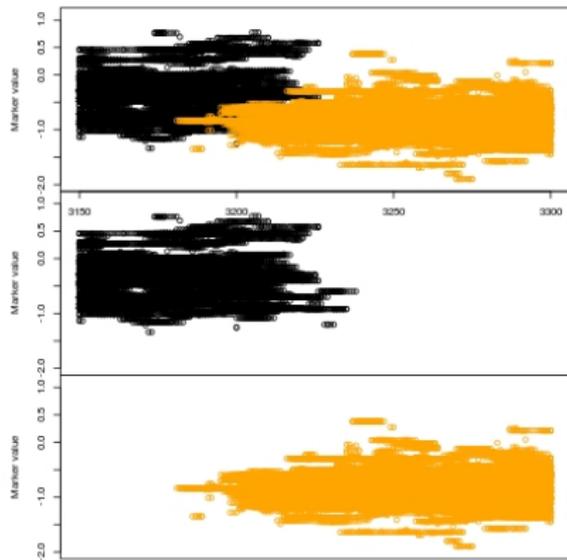
# Simulation of the trait process



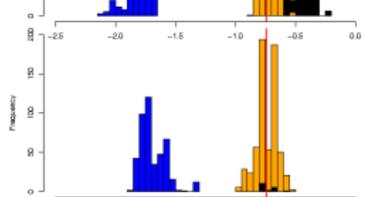
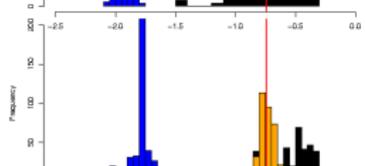
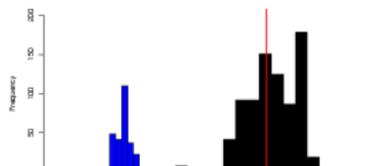
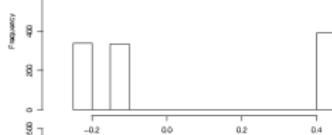
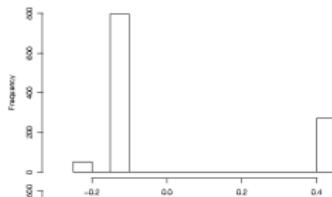
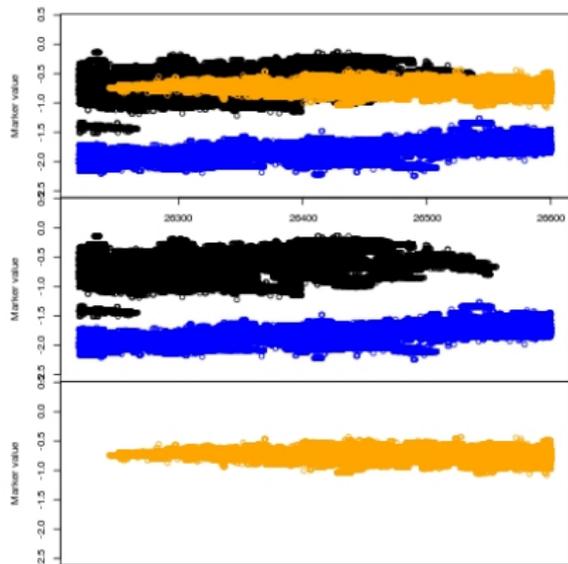
# Simulations of genealogies in a long time scale



# The hitchhiking effect



## Coexistence case.



# Asymptotics of a monomorphic population size

All individuals have the same trait  $x$ .

## Theorem

Assume that  $X_0^K(dy) = \frac{N_0^K}{K} \delta_x(dy)$  and  $\frac{N_0^K}{K} \rightarrow n_0$ .

Then the process  $(\frac{N_t^K}{K}, t \geq 0)$  converges to  $(n_t, t \geq 0)$ , where  $n_t$  is the solution of the (deterministic) logistic equation

$$\dot{n}_t = (b(x) - d(x) - C(0)n_t) n_t,$$

which converges when  $t$  tends to infinity to  $\hat{n}_x = \frac{b(x) - d(x)}{C(0)}$ .

**No trait mutations at this scale.**

**To see something, we consider the process at the much longer trait mutation scale:  $Kt$ .**

# Asymptotic behavior of the trait process ( $X_{Kt}^K$ )

Remark that  $\forall V > 0, \ln K \ll K t \ll \exp(VK),$  for large  $K$ .

**Trait Substitution Sequence** (Metz et al. 1996; Champagnat 06)

## Theorem

Assume that  $X_0^K = n_0^K \delta_{x_0}$  with  $\lim_{K \rightarrow \infty} n_0^K = \hat{n}_{x_0}$ .

Until the first trait coexistence time, the trait population process ( $X_{Kt}^K, t \geq 0$ ) converges to the pure jump process ( $\hat{n}_{Y_t} \delta_{Y_t}; t \geq 0$ ) on monomorphic states, issued from  $\hat{n}_{x_0} \delta_{x_0}$ .

The process ( $\hat{n}_{Y_t} \delta_{Y_t}; t \geq 0$ ) jumps from  $\hat{n}_x \delta_x$  to  $\hat{n}_{x+k} \delta_{x+k}$  at rate

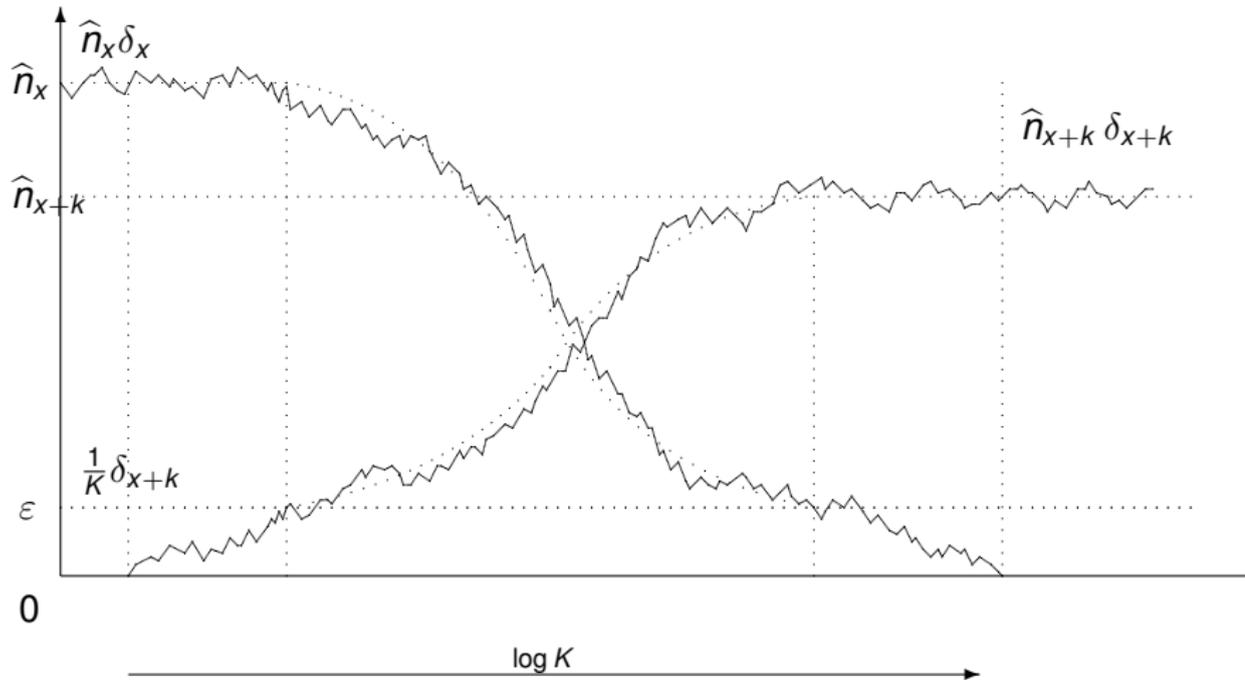
$$b(x) \hat{n}_x \frac{[f(x+k; x)]_+}{b(x+k)} m(x, k) dk.$$

- The invasion fitness function is given by the growth rate

$$f(x + k; x) = b(x + k) - d(x + k) - C(k)\hat{n}_x.$$

- Each jump corresponds to the **successful invasion of a new mutant trait**.
- The selection process has sufficient time between two (sufficiently slow) trait mutations to eliminate disadvantaged traits.
- Succession of phases of **trait mutant invasion**, and phases of **competition** between traits.

population size



# Genetical bottleneck

Let us consider a mutant  $(x + k, v)$  appearing in a monomorphic population with trait  $x$  and marker distribution  $\pi^K(x, du)$ . Assume that  $f(x + k; x) > 0$ .

## Proposition

There exists  $\varepsilon > 0$  such that for  $(t_K; K \in \mathbb{N}^*)$  with  $\log K \ll t_K \ll K$ ,

$$\begin{aligned}\lim_{K \rightarrow +\infty} \mathbb{P}(\langle \nu_{t_K}^K, \mathbf{1}_{x+k} \rangle > \varepsilon) &= \frac{f(x + k; x)}{b(x + k)}; \\ \lim_{K \rightarrow +\infty} \mathbb{P}(\langle \nu_{t_K}^K, \mathbf{1}_{x+k} \rangle = 0) &= 1 - \frac{f(x + k; x)}{b(x + k)}; \\ \lim_{K \rightarrow +\infty} \mathbb{P}(\pi_{t_K}^K(x + k, du) = \delta_v(du)) &= \frac{f(x + k; x)}{b(x + k)}.\end{aligned}$$

The fixation of the mutant creates a genetical bottleneck.

# The marker distribution

$\nu_0^K(dy, dv) = n_0^K \delta_{(x,u)}(dy, dv)$  and  $n_0^K \rightarrow \hat{n}_x$ .

## Proposition

Before the first trait mutation, the process  $(\pi_{Kt}^K(x, dv), t \geq 0)$  converges in law to the Fleming-Viot process  $(F_t^u(x, dv), t \geq 0)$  started at  $\delta_u$  and parametrized by  $x$ .

The Fleming-Viot process  $(F_t^u(x, \cdot), t \geq 0)$  is a probability measure-valued process on  $\mathcal{U}$  starting at  $\delta_u$  such that for  $g \in \mathcal{D}(A)$ ,

$$\int_{\mathcal{U}} g(v) F_t^u(x, dv) = g(u) + b(x) \int_0^t \langle F_s^u(x, \cdot), Ag \rangle ds + M_t^{(x,u)}(g).$$

$M^{(x,u)}(g)$  is a continuous square integrable martingale with quadratic variation

$$\frac{2b(x)}{\hat{n}_x} \int_0^t \left( \langle F_s^u(x, \cdot), g^2 \rangle - \langle F_s^u(x, \cdot), g \rangle^2 \right) ds.$$

# Slow-fast dynamics of traits and markers

## Substitution Fleming-Viot Process

### Theorem

Assume  $\nu_0^K(dy, dv) = n_0^K \delta_{(x_0, u_0)}(dy, dv)$  with  $\lim_{K \rightarrow \infty} n_0^K = \hat{n}_{x_0}$ .

Until the first trait-coexistence, the population process  $(\nu_{Kt}^K, t \geq 0)$  converges on  $M_F(\mathcal{X} \times \mathcal{U})$  to the process  $(V_t, t \geq 0)$  defined by

$$V_t(dy, dv) = \hat{n}_{Y_t} \delta_{Y_t}(dy) F_t^{U_t}(Y_t, dv).$$

The process  $(Y_t, U_t)$  started at  $(x_0, u_0)$  jumps from  $(x, u)$  to  $(x + k, v)$  with the jump measure

$$b(x) \hat{n}_x \frac{[f(x + k; x)]_+}{b(x + k)} F_t^u(x, dv) m(x, k) dk.$$

The convergence holds in the sense of finite dimensional distributions on  $M_F(\mathcal{X} \times \mathcal{U})$  and in the sense of occupation measures.

# Evolutionary Wright-Fisher process

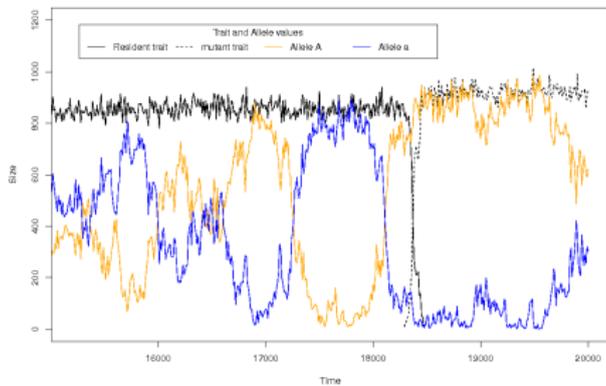
Discrete marker space  $\mathcal{U} = \{A, a\}$ , probabilities  $q_0$  and  $q_1$  to mutate from  $A$  to  $a$  and from  $a$  to  $A$ . The process  $(\nu_{kt}^K)_t$  converges to

$$\hat{n}_{Y_t} \left( W_t^a \delta_{(Y_t, a)}(dy, du) + (1 - W_t^a) \delta_{(Y_t, A)}(dy, du) \right),$$

where

$$dW_t^a = \bar{r} b(Y_t) (q_0(1 - W_t^a) - q_1 W_t^a) dt + \sqrt{\frac{2b(Y_t)}{\hat{n}_{Y_t}} W_t^a (1 - W_t^a)} dB_t.$$

The limiting process jumps with the TSS  $(Y_t)_t$  and at jump time  $t$ , the process  $(W_t^a, 1 - W_t^a)$  goes to  $(1, 0)$  with probability  $W_t^a$  and to  $(0, 1)$  with probability  $1 - W_t^a$ .



# Extension to co-existing traits

(Polymorphic Evolution Sequence, Champagnat-M. PTRF 2011).

Between the jumps of the PES, the marker distribution is the sum of independent Fleming-Viot processes parametrized by the coexisting traits and the total size at equilibrium.