

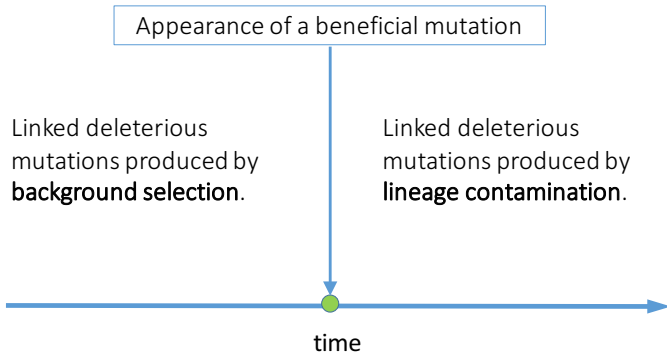
Can excessive mutation prevent adaptive evolution? A multitype branching process model.

Sophie Péni^on
Universit^oe Paris-Est Cr^oteil

joint work with Philip GERRISH and Paul SNI^oEGOWSKI
Georgia Institute of Technology and University of Pennsylvania

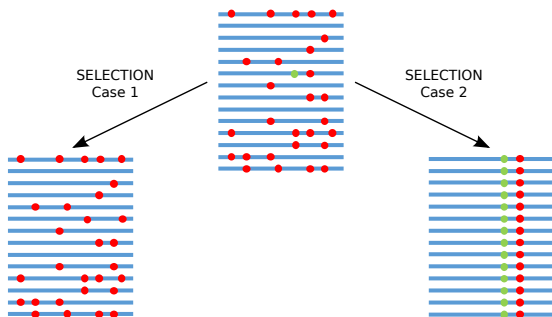
IV Workshop on Branching Processes and their Applications
University of Extremadura
10th-13th April 2018

- 1 Biological context
 - Background selection
 - Lineage contamination
- 2 Stochastic model
- 3 Results
 - Mean evolution of the population size
 - Extinction probability
 - Relative fitness
 - Mutational meltdown
 - Effect on adaptive dynamics



Background selection

- s_b, s_d = selective advantage, disadvantage
- u = genomic mutation rate
- number of deleterious mutations on each genome $Z \sim \mathcal{P}(u/s_d)$
average fitness of the population = $\mathbb{E} \left[(1 - s_d)^Z \right] = e^{-u}$

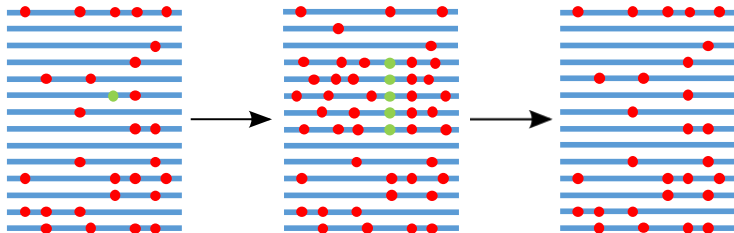


$(1 + s_b)(1 - s_d) < e^{-u}$
loss of the beneficial mutation

$(1 + s_b)(1 - s_d) > e^{-u}$
hitchhiking of deleterious mutations

Background selection

What happens at high mutation rate?



$$(1 + s_b)(1 - s_d) > e^{-u}$$

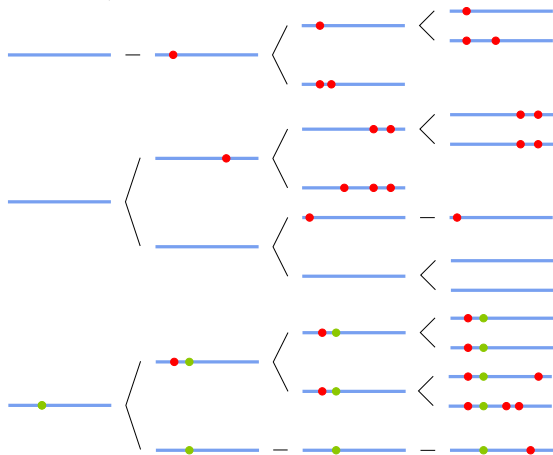
beneficial mutation is lost despite initially having a net selective advantage.

This is due to what we call **lineage contamination**.

Lineage contamination

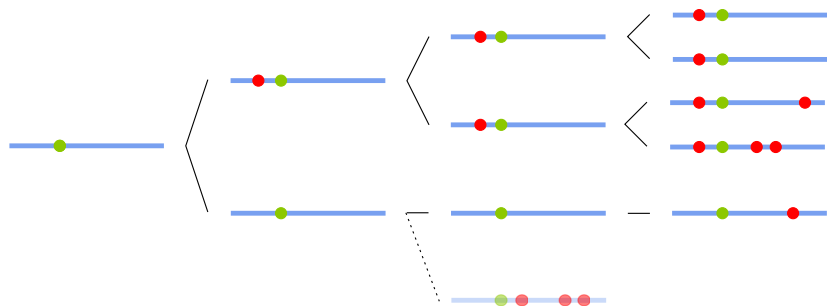
Random **accumulation of deleterious mutations** in a growing lineage founded by the occurrence of a **single beneficial mutation** in an otherwise homogeneous population (wild-type population starting with N lineages).

Initial population size $N + 1$



- 1 Biological context
 - Background selection
 - Lineage contamination
- 2 Stochastic model
- 3 Results
 - Mean evolution of the population size
 - Extinction probability
 - Relative fitness
 - Mutational meltdown
 - Effect on adaptive dynamics

Lineage contamination model



- **binary reproduction:** individuals produce 2 descendants
- **accumulation of deleterious mutations:** a descendant of an individual carrying i deleterious mutations may accumulate k additional mutations with probability $e^{-u} u^k / k!$
- **fitness:** such a descendant is then selected with probability proportional to its fitness $w_{i+k} = (1 + s_b)(1 - s_d)^{i+k}$

Multitype branching process

- a **type** $\{0, 1, \dots\}$ corresponds to the number of accumulated deleterious mutations
- each individual of type i produces a total number of 0, 1 or 2 descendants of types $\{i, i + 1, \dots\}$

Offspring generating function for the type i : for all $\mathbf{r} \in [0, 1]^{\mathbb{N}}$

$$f_i(\mathbf{r}) = \mathbb{E}_{\mathbf{e}_i} [\mathbf{r}^{\mathbf{X}_1}] = \left(1 - \sum_{k \geq 0} \frac{e^{-u} u^k}{k!} \frac{w_{i+k}}{2} (1 - r_{i+k}) \right)^2$$

Mean matrix: upper triangular

$$m_{ij} = \mathbb{E}_{\mathbf{e}_i} [X_{1,j}] = \frac{w_j e^{-u} u^{j-i}}{(j-i)!}$$

- wild-type population $(\mathbf{X}_t)_{t \in \mathbb{N}}$
 - ▶ $\mathbf{X}_0 = (N, 0, \dots)$
 - ▶ no selective advantage, $s_b = 0$
- beneficial lineage $(\mathbf{X}_t^b)_{t \in \mathbb{N}}$
 - ▶ $\mathbf{X}_0^b = (1, 0, \dots)$
 - ▶ selective advantage, $s_b > 0$

Goals

- **Extinction probability** of the beneficial lineage, as a function of mutation rate
- **Relative fitness dynamics** of the beneficial lineage within the wild-type population
- **Mutational meltdown** of the beneficial lineage (timing of sequential extinctions of fittest classes)

Mathematical challenges

- branching process with **infinite** set of types \mathbb{N}
- no back mutations \implies **reducible** process
- studied quantities not classical in the literature

- 1 Biological context
 - Background selection
 - Lineage contamination
- 2 Stochastic model
- 3 Results
 - Mean evolution of the population size
 - Extinction probability
 - Relative fitness
 - Mutational meltdown
 - Effect on adaptive dynamics

Mean evolution of the population size

Mean population size at time $t \in \mathbb{N}$:

$$\mathbb{E}(|\mathbf{X}_t^b|) = (1 + s_b)^t e^{-ut} e^{\frac{u(1-(1-s_d)^t)(1-s_d)}{s_d}}$$

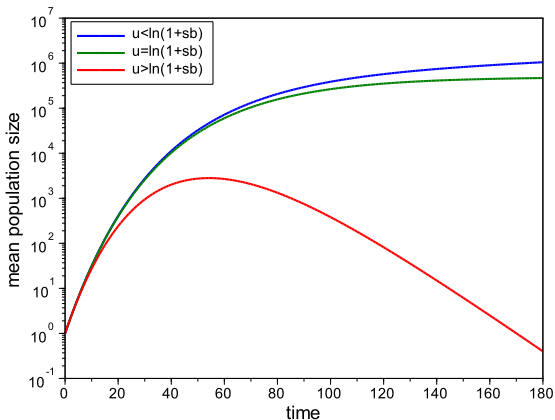


Figure. Evolution over time of the mean size of a beneficial lineage, with $s_d = 0.03$ and $s_b = 0.5$.

Extinction probability

Generally, if $(\mathbf{X}_t)_t$ is reducible with an infinite set of types,

$$p_{\text{ext}}^{\text{glob}} = \mathbb{P}\left(\lim_{t \rightarrow \infty} |\mathbf{X}_t| = 0\right) \leq \mathbb{P}\left(\forall i \in \mathbb{N}, \lim_{t \rightarrow \infty} X_{t,i} = 0\right) = p_{\text{ext}}^{\text{part}}.$$

Proposition

Partial vs. global extinction

$$p_{\text{ext}}^{\text{part}} = p_{\text{ext}}^{\text{glob}} = p_{\text{ext}}.$$

Proposition

Extinction threshold

$$p_{\text{ext}} = 1 \iff u \geq \ln(1 + s_b).$$

Proof. For branching random walks¹

$$p_{\text{ext}}^{\text{part}} < 1 \iff \limsup_{t \in \mathbb{N}} \left(m_{0,0}^{(t)}\right)^{\frac{1}{t}} > 1 \iff e^{-\mu} (1 + s_b) > 1.$$

¹Zucca (2011)

Extinction probability

Obtained² as the limit of the extinction probability of a branching process $(\mathbf{X}_t^{(D)})_t$ with a **finite** number of types D

- = $(\mathbf{X}_t)_t$ in which none of the individuals with more than D deleterious mutations are counted
- = branching process with set of types $\{0, \dots, D\}$ and offspring generating functions $f_i(r_0, \dots, r_D, 1, 1, \dots)$

$$p_{\text{ext}} = \lim_{D \rightarrow \infty} \mathbb{P} \left(\lim_{t \rightarrow \infty} |\mathbf{X}_t^{(D)}| = 0 \right)$$

Let $D_0 = \max\{i \in \mathbb{N} : e^{-u} w_i > 1\}$.

Proposition

$$p_{\text{ext}} = \mathbb{P} \left(\lim_{t \rightarrow \infty} |\mathbf{X}_t^{(D_0)}| = 0 \right),$$

and only $D_0 + 1$ equations are required to compute^a p_{ext} .

^aFirst coordinate of the smallest non-negative fixed point of the offspring generating function with $D_0 + 1$ types

²Hautphenne *et al* (2013)

Survival probability

Proposition

The survival probability $p_{svl} = 1 - p_{ext}$ is a monotonically decreasing function of s_d which reaches its minimum at $s_d = s_c = 1 - e^u / (1 + s_b)$.

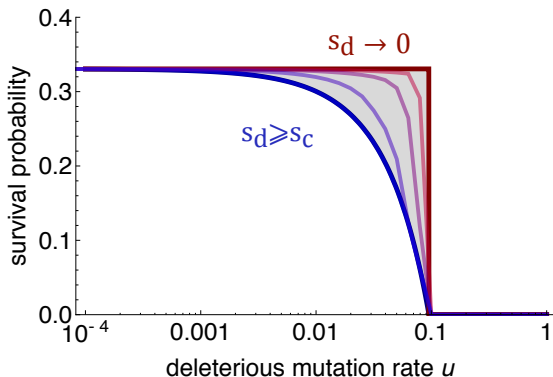


Figure. Survival probability p_{svl} of a beneficial lineage, with $s_b = 0.1$.

Relative fitness

Absolute fitnesses

$$\mathcal{W}(\mathbf{x}_t^b) = \sum_{i \in \mathbb{N}} w_i \frac{X_{t,i}^b}{|\mathbf{x}_t^b|}, \quad t < T_{ext}^b,$$

Relative fitness

$$\frac{\mathcal{W}(\mathbf{x}_t^b)}{\mathcal{W}(\mathbf{x}_t^b + \mathbf{x}_t)}.$$

How do the fitnesses of the adaptive sub-population and of the large wild-type population compare?

Intuition

Since both populations are accumulating deleterious mutations at the same rate, the *relative* fitness should not depend on this deleterious mutation rate.

Relative Fitness

Large population approximation

Proposition

$$\lim_{N \rightarrow \infty} \frac{\mathcal{W}(\mathbf{X}_t^b)}{\mathcal{W}(\mathbf{X}_t^b + \mathbf{X}_t)} \stackrel{\text{a.s.}}{=} e^{u(1-s_d)(1-(1-s_d)^t)} \mathcal{W}(\mathbf{X}_t^b).$$

Proof. Branching property for $(\mathbf{X}_t)_t$ (N independent copies of a process starting with one individual) and strong law of large numbers:

$$\begin{aligned} \lim_{N \rightarrow \infty} \frac{1}{\mathcal{W}(\mathbf{X}_t^b + \mathbf{X}_t)} &= \lim_{N \rightarrow \infty} \frac{|\mathbf{X}_t^b| + |\mathbf{X}_t|}{\sum_{i \in \mathbb{N}} (1-s_d)^i \left((1+s_b) X_{t,i}^b + X_{t,i} \right)} \\ &\stackrel{\text{a.s.}}{=} \frac{\mathbb{E}(|\mathbf{X}_t|)}{\sum_{i \in \mathbb{N}} (1-s_d)^i \mathbb{E}(X_{t,i})}. \end{aligned}$$

□

Relative Fitness

Long-time limit of the mean relative fitness

Proposition

$$\lim_{t \rightarrow +\infty} \mathbb{E} \left(\frac{\mathcal{W}(\mathbf{X}_t^b)}{\mathcal{W}(\mathbf{X}_t^b + \mathbf{X}_t)} \mathbf{1}_{t < T_{\text{ext}}^b} \right) = (1 + s_b) p_{\text{svl}}.$$

Proof. Property for (finite-type) reducible branching processes³ with largest eigenvalue > 1 :

$$\lim_{t \rightarrow \infty} e^{tu} (1 + s_b)^{-t} \mathbf{X}_t^b \stackrel{\text{a.s.}}{=} W \mathbf{v},$$

where

- $\mathbf{v} = \left(1, \frac{u(1-s_d)}{s_d}, \frac{1}{2!} \left(\frac{u(1-s_d)}{s_d} \right)^2, \dots \right)$ is the left eigenvector of the mean matrix \mathbf{M} for its maximal eigenvalue $e^{-u} (1 + s_b)$,
- W is a one-dimensional random variable with $\mathbb{P}(W > 0) = p_{\text{svl}}$.

□

³Kesten and Stigum, 1967

Relative Fitness

We obtain upper and lower bounds for the mean relative fitness which only involve computable quantities of the processes $(|\mathbf{X}_t^b|)_{t \in \mathbb{N}}$ and $(X_{t,i}^b)_{t \in \mathbb{N}}$.

- $\mathbb{E}(r^{|\mathbf{X}_t^b|})$ obtained by t iterations of $f_0(r, r, \dots)$
- $\mathbb{E}(r^{X_{t,i}^b})$ obtained by t iterations of $f_0(1, \dots, 1, r, 1, \dots)$

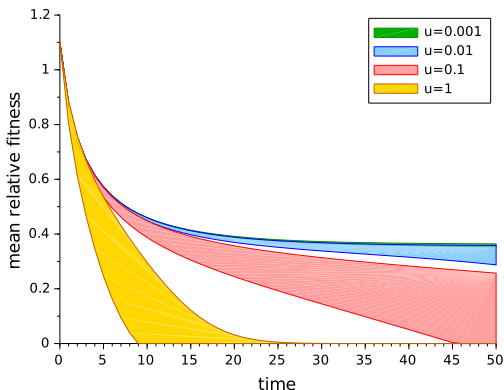
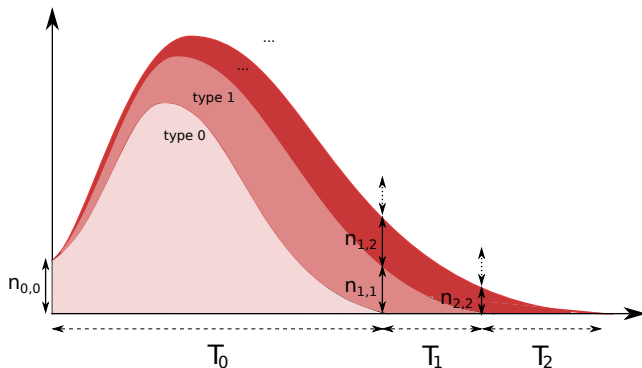


Figure. Time evolution of the mean relative fitness ($s_b = 0.1, s_d = 0.03$).

Mutational meltdown

Extinction time of the fittest class

$T_i := \inf\{t \geq 0 : X_{t,i} = 0\}$ conditionally on $\mathbf{X}_0 = \mathbf{n}_i$; the (rounded) average composition of the population at T_{i-1} .



Timing of sequential extinction of fittest classes

$$\mathbb{E}_{\mathbf{n}_0}(T_0), \mathbb{E}_{\mathbf{n}_1}(T_1), \dots$$

Mutational meltdown

- $T_i \stackrel{\mathcal{L}}{\sim}$ extinction time of the monotype branching process $(Y_t)_{t \geq 0}$ with offspring generating function $g(r) = f_i(1, \dots, 1, r, 1, \dots)$ and $Y_0 = n_{i,j}$
- $\phi_t(r) = \mathbb{E}_1(r^{Y_t})$ solution of the Ricatti equation

$$\frac{\partial}{\partial t} \phi_t(r) = g(\phi_t(r)) - \phi_t(r), \quad \phi_t(0) = r.$$

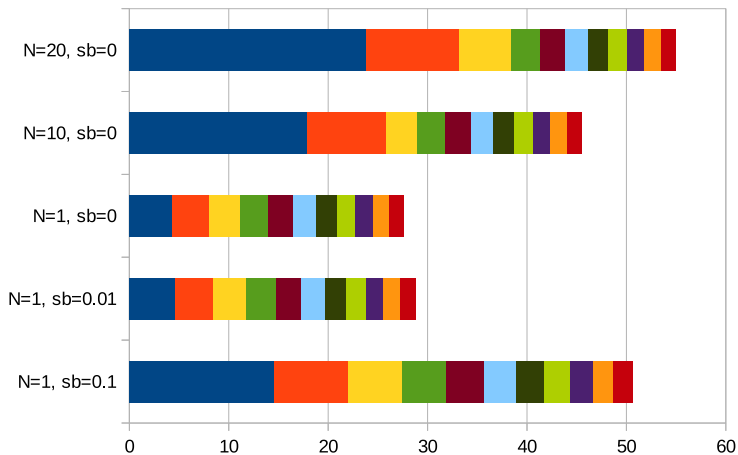
Proposition

For each $i \in \mathbb{N}$ and $\mathbf{x} \in \mathbb{N}^{\mathbb{N}}$ with $x_j = 0$ for $j > i$,

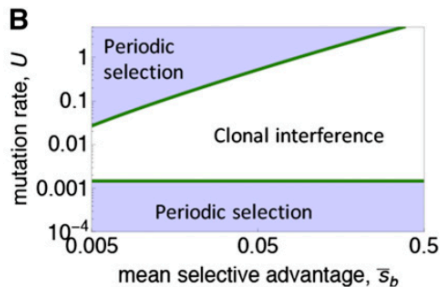
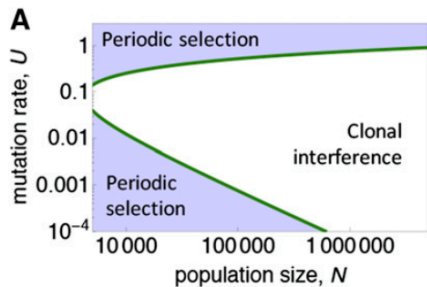
- the extinction time cumulative distribution $\mathbb{P}_{\mathbf{x}}(T_i \leq t)$,
 - the mean extinction time $\mathbb{E}_{\mathbf{x}}(T_i)$,
 - the mean population composition at the extinction time $\mathbb{E}_{\mathbf{x}}(X_{T_i,j})$,
- can be computed explicitly.

Mutational meltdown

Sequence of the mean extinction times of the fittest classes ($u = 0.1$, $s_d = 0.03$)



Effect on adaptive dynamics



Thank you for your attention!