A two-locus model with interacting mutation rates and fluctuating selection

Peter Pfaffelhuber

University of Freiburg
Abteilung für Mathematische Stochastik
joint work with Franz Baumdicker, Elisabeth Huss

June 25th, 2018

Second order selection

- An individual is **first-order-fit** if it produces on average more offspring than others.
- An individual is **second-order-fit** if its offspring produce on average more offspring than others.

- Example (from Mao et al (1997)):
  In *E. coli*, the fraction of cells lacking a certain DNA repair mechanism, is $\sim 10^{-5}$.
  $\rightarrow$ they carry a mutator allele
  Treat a population with antibiotics. After four rounds of treatments, only mutators have survived.

- Conclusion: **Mutators can be second-order fit.**

Res. Microbiol. 152 (2001) 11–16

S0923-2508(00)01163-3/REV

Mini-review

Second-order selection in bacterial evolution: selection acting on mutation and recombination rates in the course of adaptation

Olivier Tenaillon*, François Taddei, Miroslav Radman, Ivan Matic

Inserm E9916, Faculté de médecine Necker Enfants Malades, Université Paris V, 156, rue de Vaugirard, 75730 Paris cedex 15, France

Received 24 October 2000; accepted 26 October 2000

Abstract – The increase in genetic variability of a population can be selected during adaptation, as demonstrated by the selection of mutator alleles. The dynamics of this phenomenon, named second-order selection, can result in an improved adaptability of bacteria through regulation of all facets of mutation and recombination processes. © 2001 Éditions scientifiques et médicales Elsevier SAS

PROCEEDINGS OF THE ROYAL SOCIETY B

rspb.royalsocietypublishing.org

Adaptation through genetic time travel? Fluctuating selection can drive the evolution of bacterial transformation

Jan Engelstädter[1] and Danesh Moradigaravand[2]

[1]School of Biological Sciences, The University of Queensland, Brisbane, Queensland 4072, Australia
[2]Institute of Biogeochemistry and Pollutant Dynamics, ETH Zurich, Zurich 8092, Switzerland

The model

There is an A-locus (types: ℓ, h) and a B-locus (types: 0,1)
The A-allele determines the mutation rate at the B-locus

		B-locus		
		0		1
A-locus	ℓ	$X_{\ell 0}$	$\longleftrightarrow$	$X_{\ell 1}$
	h	X_{h0}	$\longleftrightarrow$	X_{h1}

$$
\begin{aligned}
dX_{\ell 0} &= \theta_\ell (X_{\ell 1} - X_{\ell 0})dt + \sqrt{X_{\ell 0}(1 - X_{\ell 0})}dW_{\ell 0}, \\
dX_{\ell 1} &= \theta_\ell (X_{\ell 0} - X_{\ell 1})dt + \sqrt{X_{\ell 1}(1 - X_{\ell 1})}dW_{\ell 1}, \\
dX_{h0} &= \theta_h (X_{h1} - X_{h0})dt + \sqrt{X_{h0}(1 - X_{h0})}dW_{h0}, \\
dX_{h1} &= \theta_h (X_{h1} - X_{h0})dt + \sqrt{X_{h1}(1 - X_{h1})}dW_{h1}
\end{aligned}
$$

The model

There is an A-locus (types: ℓ, h) and a B-locus (types: 0,1)
The A-allele determines the mutation rate at the B-locus

A-locus \ B-locus	0		1
ℓ	$X_{\ell 0}$	$\longleftrightarrow$	$X_{\ell 1}$
h	X_{h0}	$\longleftrightarrow$	X_{h1}

Environment favors 1 over 0 with $\sigma n Z$ for $Z \in \{-1, 1\}$

$Z \to -Z$ at rate $\gamma n^2/2$

$$
\begin{aligned}
dX_{\ell 0} &= -\sigma n Z X_{\ell 0} X_1 dt + \theta_\ell (X_{\ell 1} - X_{\ell 0}) dt + \sqrt{X_{\ell 0}(1 - X_{\ell 0})} dW_{\ell 0}, \\
dX_{\ell 1} &= \sigma n Z X_{\ell 1} X_0 dt + \theta_\ell (X_{\ell 0} - X_{\ell 1}) dt + \sqrt{X_{\ell 1}(1 - X_{\ell 1})} dW_{\ell 1}, \\
dX_{h0} &= -\sigma n Z X_{h0} X_1 dt + \theta_h (X_{h1} - X_{h0}) dt + \sqrt{X_{h0}(1 - X_{h0})} dW_{h0}, \\
dX_{h1} &= \sigma n Z X_{h1} X_0 dt + \theta_h (X_{h1} - X_{h0}) dt + \sqrt{X_{h1}(1 - X_{h1})} dW_{h1}
\end{aligned}
$$

Stochastic averaging

- The genertor of (X, Z) is

$$Gf(x,z) = \underbrace{G_0 f(x,z)}_{\substack{\text{mutation,}\\ \text{resampling}}} + \underbrace{nG_1 f(x,z)}_{\text{selection}} + \underbrace{n^2 G_2 f(x,z)}_{\text{environment}}$$

- Any Markov process can be characterized via a martingale problem. Here, this means that

$$\begin{aligned} f(X_t, Z_t) &- \int_0^t Gf(X_s, Z_s) ds \\ &= f(X_t, Z_t) - \int_0^t (G_0 f + nG_1 f + n^2 G_2 f)(X_s, Z_s) ds \end{aligned}$$

  is a martingale for all smooth, bounded f.
- Is there a limit of X as $n \to \infty$?

Stochastic averaging

- Goal: Averaging out a fast variable (environemnt).
- Dates back at least to Khashminskii (1966)

  A limit theorem for the solutions of differential equations with random right-hand sides. *Theor. Probability Appl.*, 11(11):390–406, 1966.

- General reference is Kurtz (1992)

  Averaging for martingale problems and stochastic approximation. In *Applied stochastic analysis (New Brunswick, NJ, 1991), volume 177 of Lecture Notes in Control and Inform. Sci.*, 186–209. Springer, Berlin, 1992.

- For processes on three time-scales:
  - Semigroup approach: See Theorem 1.7.6 of Ethier, Kurtz (1986)

    Markov processes: Characterization and conver- gence. Wiley Series in Probability and Mathematical Statistics: Probability and Mathe- matical Statistics. John Wiley & Sons Inc., New York, 1986.
  - Special case of diffusion operators E. Pardoux and A. Yu. Veretennikov. On Poisson equation and diffusion approximation 1 and 2. *Ann. Probab.*, 29:1061–1085, 2001 and 31:1166–1192, 2003.
  - Martingale-problem approach Hutzentaler, Pfaffelhuber, Printz (2018) Stochastic averaging for multiscale Markov processes with an application to a Wright-Fisher model with fluctuating selection. *Submitted.*

Stochastic averaging

- The generator is

$$G = G_0 + nG_1 + n^2 G_2$$

with $G_2 f = 0$ if f only depends on x.

- For $n \to \infty$,

$$f \qquad (X_t \quad) - \int_0^t G \; f \qquad (X_s, Z_s) ds$$

$$\approx f(X_t) - \int_0^t G_0 f(X_s) + nG_1 f(X_s, Z_s)$$

$$ds$$

is a martingale.

Stochastic averaging

- The generator is

$$G = G_0 + nG_1 + n^2 G_2$$

with $G_2 f = 0$ if f only depends on x.
Let f only depend on x and find h such that $G_2 h = -G_1 f$.

- For $n \to \infty$, Z is a fast process with equilibrium π_x, so

$$\begin{aligned}
&(f+\tfrac{1}{n}h)\,(X_t, Z_t) - \int_0^t G(f+\tfrac{1}{n}h)\,(X_s, Z_s)ds \\
&\approx f(X_t) - \int_0^t G_0 f(X_s) + nG_1 f(X_s, Z_s) \\
&\qquad + G_1 h(X_s, Z_s) + nG_2 h(X_s, Z_s) ds \\
&\approx f(X_t) - \int_0^t \underbrace{G_0 f(X_s) + \overbrace{\mathbb{E}_{\pi_{X_s}}[G_1 h(X_s, Z)]}^{=:\bar{G}_1 h(X_s)}}_{\text{potential generator of limit process}} ds
\end{aligned}$$

is a martingale.

Stochastic averaging

UNI FREIBURG

- The generator is

$$Gf(x,z) = \underbrace{G_0 f(x,z)}_{\text{mutation, resampling}} + \underbrace{nG_1 f(x,z)}_{\text{selection}} + \underbrace{n^2 G_2 f(x,z)}_{\text{environment}}$$

- Let f only depend on x and find h with $G_2 h = -G_1 f$
- For the limit process X,

$$f(X_t) - \int_0^t G_0 f(X_s) + \bar{G}_1 h(X_s) ds$$

is a martingale.

Stochastic averaging

- The generator is

$$Gf(x,z) = \underbrace{G_0 f(x,z)}_{\text{mutation, resampling}} + \underbrace{nG_1 f(x,z)}_{\text{selection}} + \underbrace{n^2 G_2 f(x,z)}_{\text{environment}}$$

$$G_2 f(x,z) = \mathbb{E}_\pi[f(x,Z)] - f(x,z)$$

- Let f only depend on x and find h with $G_2 h = -G_1 f$
- For the limit process X,

$$f(X_t) - \int_0^t G_0 f(X_s) + \bar{G}_1 h(X_s) ds$$

  is a martingale.
- For $h = G_1 f$,

$$G_2 h(x,z) = \mathbb{E}_\pi[Z] g(x) - G_1 f(x,z) = -G_1 f(x,z)$$

- Limit has generator, for smooth f,

$$Gf = G_0 f + \mathbb{E}_\pi[G_1 G_1 f(x,Z)]$$

Main result

UNI FREIBURG

Theorem (Baumdicker, Huss, P, 2018)

As $n \to \infty$, $X = X^n$ converges weakly to the unique solution of

$$\begin{aligned}
dX_{\ell 0} &= \tfrac{\sigma^2}{\gamma} X_{\ell 0} X_1 (X_1 - X_0)dt + \theta_\ell (X_{\ell 1} - X_{\ell 0})dt \\
&\qquad + \sqrt{X_{\ell 0}(1 - X_{\ell 0})}dW_{\ell 0} + \sqrt{\tfrac{2\sigma^2}{\gamma}} X_{\ell 0} X_1 dW, \\
dX_{\ell 1} &= \tfrac{\sigma^2}{\gamma} X_{\ell 1} X_0 (X_0 - X_1)dt + \theta_\ell (X_{\ell 0} - X_{\ell 1})dt \\
&\qquad + \sqrt{X_{\ell 1}(1 - X_{\ell 1})}dW_{\ell 1} - \sqrt{\tfrac{2\sigma^2}{\gamma}} X_{\ell 1} X_0 dW, \\
dX_{h0} &= \tfrac{\sigma^2}{\gamma} X_{h0} X_1 (X_1 - X_0)dt + \theta_h (X_{h1} - X_{h0})dt \\
&\qquad + \sqrt{X_{h0}(1 - X_{h0})}dW_{h0} + \sqrt{\tfrac{2\sigma^2}{\gamma}} X_{h0} X_1 dW, \\
dX_{h1} &= \tfrac{\sigma^2}{\gamma} X_{h1} X_0 (X_0 - X_1)dt + \theta_h (X_{h0} - X_{h1})dt \\
&\qquad + \sqrt{X_{h1}(1 - X_{h1})}dW_{h1} - \sqrt{\tfrac{2\sigma^2}{\gamma}} X_{h1} X_0 dW,
\end{aligned}$$

where W is an independent Brownian motion.

Main result

UNI FREIBURG

Corollary 1 (Karlin-Levikson model)

If $\theta = \theta_h = \theta_\ell$, $X_0 = X$ satisfies

$$dX = \frac{2\sigma^2}{\gamma} X(1-X)(\tfrac{1}{2} - X)dt + \theta(1-2X)dt + \sqrt{X(1-X)}dW + \sqrt{\frac{2\sigma^2}{\gamma}} X(1-X)dW'$$

with independent Brownian motions W, W'.

Corollary 2 (High versus low mutators)

$$dX_h = \frac{\sigma^2}{\gamma}(X_{h0}X_{\ell 1} - X_{h1}X_{\ell 0})(X_1 - X_0)dt + \sqrt{X_h X_\ell}dW + \sqrt{\frac{2\sigma^2}{\gamma}}(X_{h0}X_{\ell 1} - X_{h1}X_{\ell 0})dW'$$

with independent Brownian motions W, W'.

When is a high mutation rate beneficial?

UNI FREIBURG

Theorem (Baumdicker, Huss, P; 2018)

If $X_h(0) = x, X_{h0}(0) = px, X_{\ell 0} = q(1-x)$,

$$\mathbf{P}_x(X_h(\infty) = 1) = x + \frac{\sigma^2}{8\gamma} x(1-x) f + o(\sigma^2/\gamma)$$

with

$$\begin{aligned} f &= (1-2x)\left(\frac{1}{3} - \frac{4}{(3+\theta_h+\theta_\ell)}(1-2p)(1-2q)\right) \\ &\quad + \frac{3}{(3+2\theta_\ell)}\left((1-x)(1-2q)^2 - \frac{1}{1+\theta_\ell}\right) \\ &\quad - \frac{3}{(3+2\theta_h)}\left(x(1-2p)^2 - \frac{1}{1+\theta_h}\right) \\ &\quad + 3\frac{\theta_h - \theta_\ell}{(1+\theta_\ell)(1+\theta_h)}. \end{aligned}$$

When is a high mutation rate beneficial?

- Assume $p = q = \frac{1}{2}$, i.e.

$$\mathbf{P}_x(X_h(\infty) = 1) \approx x + \frac{\sigma^2}{8\gamma} x(1-x) \cdot$$
$$\Big(\tfrac{1}{3}(1-2x) - \frac{3}{(3+2\theta_\ell)}\frac{1}{1+\theta_\ell} + \frac{3}{(3+2\theta_h)}\frac{1}{1+\theta_h} + 3\frac{\theta_h - \theta_\ell}{(1+\theta_\ell)(1+\theta_h)} \Big)$$

- For a maximum, solve

$$-4(\theta_\ell - 1)\theta_h^2 + 8(2-\theta_\ell)\theta_h + 2(7-2\theta_\ell) = 0,$$
$$\theta_h = \frac{4 - 2\theta_\ell + \sqrt{2(\theta_\ell+1)}}{2(\theta_\ell - 1)}.$$

$\Rightarrow$ there is a fixed point at $\theta \approx 1.78$.

When is a high mutation rate beneficial?

- For small σ^2/γ,

$$\mathbf{P}_x^{\sigma^2/\gamma}[X_h(\infty)=1] = \mathbf{E}_x^{\sigma^2/\gamma}[X_h(\infty)] = x + \int_0^\infty \mathbf{E}_x^{\sigma^2/\gamma}[\bar{G}X_h(s)]ds$$

$$= x + \frac{\sigma^2}{\gamma}\int_0^\infty \mathbf{E}_x^{\sigma^2/\gamma}[(X_{h0}(t)X_{\ell 1}(t) - X_{h1}(t)X_{\ell 0}(t))(X_1(t)-X_0(t))]dt$$

When is a high mutation rate beneficial?

UNI FREIBURG

- For small σ^2/γ,

$$\mathbf{P}_x^{\sigma^2/\gamma}[X_h(\infty)=1] = \mathbf{E}_x^{\sigma^2/\gamma}[X_h(\infty)] = x + \int_0^\infty \mathbf{E}_x^{\sigma^2/\gamma}[\bar{G}X_h(s)]ds$$

$$= x + \frac{\sigma^2}{\gamma}\int_0^\infty \mathbf{E}_x^{\sigma^2/\gamma}[(X_{h0}(t)X_{\ell 1}(t) - X_{h1}(t)X_{\ell 0}(t))(X_1(t)-X_0(t))]dt$$

$$= x + \frac{\sigma^2}{\gamma}\int_0^\infty \underbrace{\mathbf{E}_x^{0}[(X_{h0}(t)X_{\ell 1}(t) - X_{h1}(t)X_{\ell 0}(t))(X_1(t)-X_0(t))]}_{\text{can be computed using Kingman's coalescent}} dt$$

$$+o(\sigma^2/\gamma)$$

Summary

- Second-order selection favors types which have fit offspring (rather than being fit themselves).
- Fluctuating selection can be trated with stochastic averaging.
- For small selection strength or fast fluctuating environments, we computed the optimal mutation rate.