MODELLING POPULATIONS UNDER FLUCTUATING SELECTION

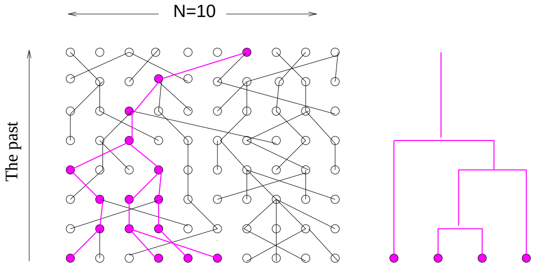
Alison Etheridge

With Aleksander Klimek (Oxford) and Niloy Biswas (Harvard)

The simplest imaginable model of inheritance

A population of fixed size, N, evolving in discrete generations.

Each individual inherits its genetic type from a parent chosen uniformly at random from the previous generation.

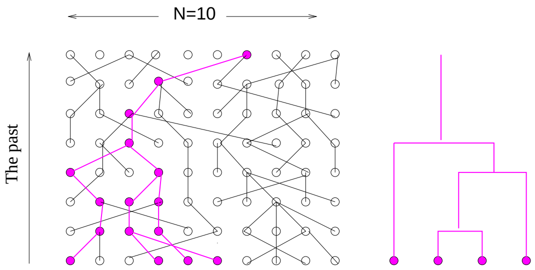


The simplest imaginable model of inheritance

A population of fixed size, N, evolving in discrete generations.

Each individual inherits its genetic type from a parent chosen uniformly at random from the previous generation.

$t \mapsto t/N$, $N \mapsto \infty$

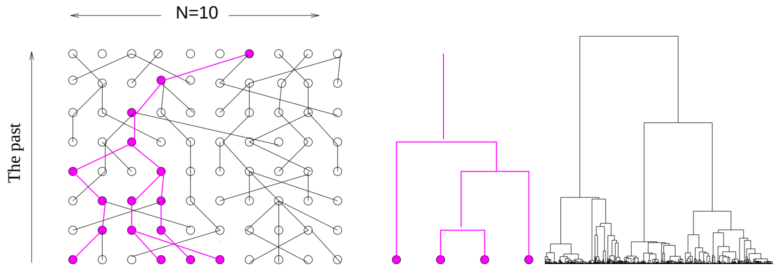


The simplest imaginable model of inheritance

A population of fixed size, N, evolving in discrete generations.

Each individual inherits its genetic type from a parent chosen uniformly at random from the previous generation.

$t \mapsto t/N$, $N \mapsto \infty$



Kingman 1982

Forwards in time? $t \mapsto t/N$, $N \mapsto \infty$

Each individual inherits its genetic type from a parent chosen uniformly at random from the previous generation. Two types a and A. $p(t) =$ proportion of type a. $\delta t = 1/N$.

Forwards in time? $t \mapsto t/N,\ N \mapsto \infty$

Each individual inherits its genetic type from a parent chosen uniformly at random from the previous generation. Two types a and A. $p(t) =$ proportion of type a. $\delta t = 1/N$.

Forwards in time

- $\mathbb{E}[\Delta p] = 0$ (neutrality)
- $\mathbb{E}[(\Delta p)^2] = \delta t p(1-p)$
- $\mathbb{E}[(\Delta p)^3] = O(\delta t)^2$

$$\mathrm{d}p_t = \sqrt{p_t(1-p_t)}\mathrm{d}W_t$$

Genetic drift

Forwards in time? $t \mapsto t/N$, $N \mapsto \infty$

Each individual inherits its genetic type from a parent chosen uniformly at random from the previous generation. Two types a and A. $p(t) =$ proportion of type a. $\delta t = 1/N$.

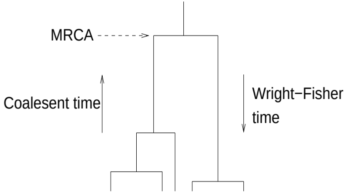
Forwards in time

- $\mathbb{E}[\Delta p] = 0$ (neutrality)
- $\mathbb{E}[(\Delta p)^2] = \delta t p(1-p)$
- $\mathbb{E}[(\Delta p)^3] = O(\delta t)^2$

$$\mathrm{d}p_t = \sqrt{p_t(1-p_t)}\mathrm{d}W_t$$

Genetic drift

Backwards in time



Kingman's coalescent

Forwards in time? $t \mapsto t/N,\ N \mapsto \infty$

Each individual inherits its genetic type from a parent chosen uniformly at random from the previous generation. Two types a and A. $p(t) =$ proportion of type a. $\delta t = 1/N$.

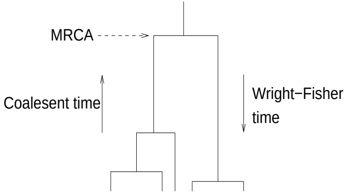
Forwards in time

- $\mathbb{E}[\Delta p] = 0$ (neutrality)
- $\mathbb{E}[(\Delta p)^2] = \delta t p(1-p)$
- $\mathbb{E}[(\Delta p)^3] = O(\delta t)^2$

$$\mathrm{d}p_t = \sqrt{p_t(1-p_t)}\mathrm{d}W_t$$

Genetic drift

Backwards in time

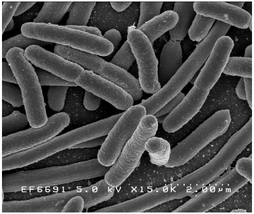


Kingman's coalescent

$$\mathrm{d}p_\tau = \sqrt{\frac{1}{N_e} p_\tau(1-p_\tau)}\mathrm{d}W_\tau,$$

Coalescence rate $\frac{1}{N_e}\binom{k}{2}$

Universality



$N_e = 25,000,000$

$N_e = 2,000,000$

$N_e = 20,000$ (Europe),
$< 50,000$ (Whole world)

$N_e < 100$

Images: Escherichia coli by Rocky Mountain Laboratories NIAID; and "Drosophila melanogaster -side (aka)" by André Karwath, Flickr-moses namkung-The Crowd for DMB 1, "SpottedSalamander" by Camazine at en.wikipedia, all via Wikimedia Commons.

Selection

Relative fitness types a and A are $1+s:1$.
If proportion of type a parents is p, each offspring (independently) type a with probability

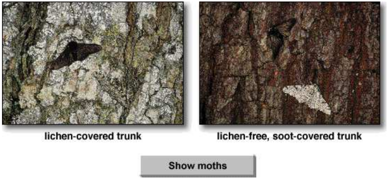
$$\frac{(1+s)p}{1+sp} = (1+s)p\{1-sp\} + \mathcal{O}(s^2) = p + sp(1-p) + \mathcal{O}(s^2).$$

- $\mathbb{E}[\Delta p] = \delta t N sp(1-p)$ (selection)
- $\mathbb{E}[(\Delta p)^2] = \delta t p(1-p) + O(s\delta t)$
- $\mathbb{E}[(\Delta p)^3] = O(\delta t)^2$

If $Ns \to \mathbf{s}$,

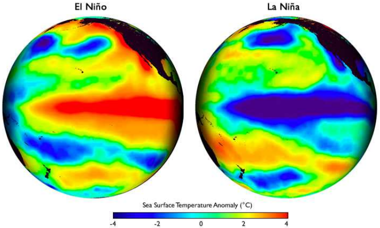
$$\mathrm{d}p_t = \mathbf{s}p(1-p)\mathrm{d}t + \sqrt{p_t(1-p_t)}\mathrm{d}W_t$$

Rapidly fluctuating environments



Gillespie:

"If fitnesses do depend on the state of the environment, as they surely must, then they must just as assuredly change in both time and space, driven by temporal and spatial fluctuations in the environment."



Fluctuating selection

Environment state $Z \in \{-1, +1\}$, $Z \mapsto -Z$ at Poisson rate $1/2$.
Relative fitness types a and A are $1 + Zs : 1$.

If proportion of type a parents is p, each offspring (independently) type a with probability

$$\frac{(1+Zs)p}{1+Zsp} = (1+Zs)p\{1-Zsp\} + \mathcal{O}(s^2) = p + Zsp(1-p) + \mathcal{O}(s^2).$$

Assume $s\sqrt{N}$ is $\mathcal{O}(1)$, so selection *much* stronger than before.

$$\text{“}\mathrm{d}p = Z\sqrt{N}sp(1-p)\mathrm{d}t + \sqrt{p(1-p)}\mathrm{d}W\text{”}$$

The “Kurtz trick”

$$\text{“}\mathrm{d}p = Z\sqrt{N}\mathbf{s}p(1-p)\mathrm{d}t + \sqrt{p(1-p)}\mathrm{d}W\text{”}$$

Frequency type a characterised by a martingale problem.

$$f(p_t, Z_t) - \int_0^t \mathcal{L}f(p_s, Z_s)\mathrm{d}s = f(p_t, Z_t)$$

$$-\int_0^t \Big(\underbrace{\mathcal{L}^{\texttt{neu}}}_{\frac{1}{2}p(1-p)f_{pp}} + \sqrt{N} \underbrace{\mathcal{L}^{\texttt{sel}}}_{\mathbf{s}Zp(1-p)f_p} + N \underbrace{\mathcal{L}^{\texttt{env}}}_{\mathbb{E}_\pi[f(Z)]-f(z)} \Big) f(p_s, Z_s)\mathrm{d}s$$

is a martingale.

The “Kurtz trick” (cont.)

Suppose f depends only on p.

$$(f+\frac{1}{\sqrt{N}}\mathcal{L}^{\texttt{sel}}f)(p_t,Z_t)-\int_0^t(\mathcal{L}^{\texttt{neu}}f(p_s,Z_s)+\frac{1}{\sqrt{N}}\mathcal{L}^{\texttt{neu}}\mathcal{L}^{\texttt{sel}}f(p_s,Z_s)$$

$$+\sqrt{N}\mathcal{L}^{\texttt{sel}}f(p_s,Z_s)+\mathcal{L}^{\texttt{sel}}\mathcal{L}^{\texttt{sel}}f(p_s,Z_s)+\underbrace{\sqrt{N}\mathcal{L}^{\texttt{env}}\mathcal{L}^{\texttt{sel}}f(p_s,Z_s)}_{\sqrt{N}\mathbb{E}_\pi[Zp(1-p)f_p]-\sqrt{N}\mathcal{L}^{\texttt{sel}}f}\ \mathrm{d}s$$

$$\approx f(p_t)-\int_0^t\Big(\mathcal{L}^{\texttt{neu}}f(p_s,Z_s)+\mathcal{L}^{\texttt{sel}}\mathcal{L}^{\texttt{sel}}f(p_s,Z_s)\Big)\mathrm{d}s$$

is a martingale.

Limiting diffusion

$$f(p_t) - \int_0^t \big(\mathcal{L}^{\texttt{neu}} f(p_s) + \mathcal{L}^{\texttt{sel}}\mathcal{L}^{\texttt{sel}} f(p_s, Z_s)\big)\mathrm{d}s$$

is a martingale (for all nice f).

$$\begin{aligned}\mathcal{L}^{\texttt{sel}}\mathcal{L}^{\texttt{sel}} f(p, Z) &= \mathcal{L}^{\texttt{sel}}(Z\mathbf{s}p(1-p)f_p) \\ &= Z^2\mathbf{s}^2\big(p^2(1-p)^2 f_{pp} + p(1-p)(1-2p)f_p\big).\end{aligned}$$

Limiting diffusion satisfies

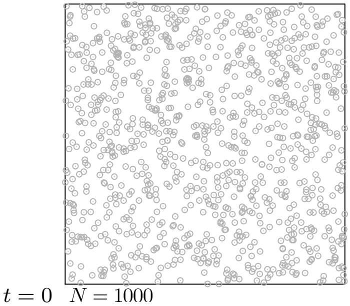
$$\mathrm{d}p = \underbrace{\mathbf{s}^2 p(1-p)(1-2p)\mathrm{d}t}_{\text{balancing selection}} + \underbrace{\sqrt{p(1-p)}\mathrm{d}B_t^1}_{\text{genetic drift}} + \underbrace{\sqrt{2}\mathbf{s}p(1-p)\mathrm{d}B_t^2}_{\text{fluctuating environment}},$$

where B^1, B^2 independent Brownian motions.

Space: the Wright-Malécot model

How do correlations in genetic types decay with distance?

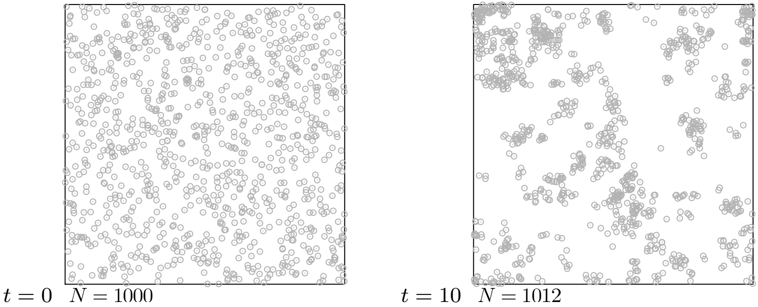
Average one offspring per individual; location of each offspring independent Gaussian pick around position of parent



Space: the Wright-Malécot model

How do correlations in genetic types decay with distance?

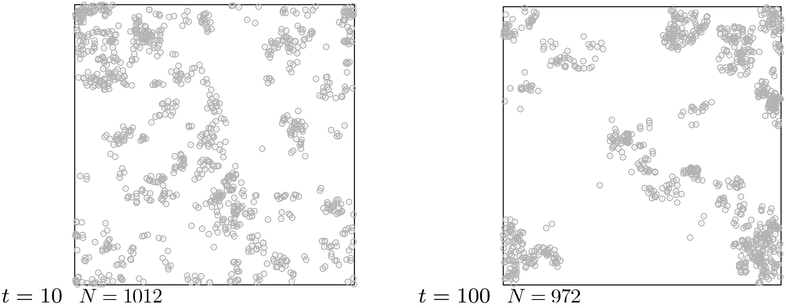
Average one offspring per individual; location of each offspring independent Gaussian pick around position of parent



Space: the Wright-Malécot model

How do correlations in genetic types decay with distance?

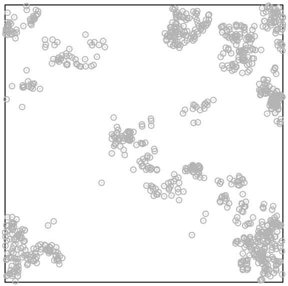
Average one offspring per individual; location of each offspring independent Gaussian pick around position of parent



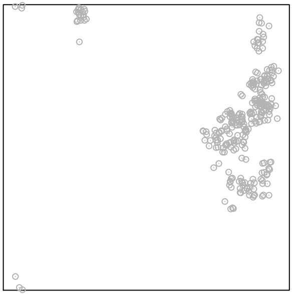
Space: the Wright-Malécot model

How do correlations in genetic types decay with distance?

Average one offspring per individual; location of each offspring independent Gaussian pick around position of parent



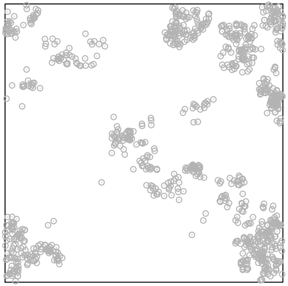
$t = 100 \quad N = 972$



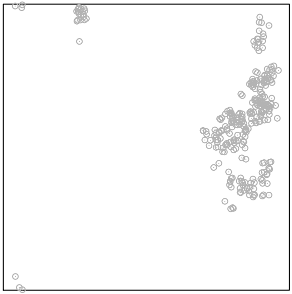
$t = 1000 \quad N = 293$

Space: the Wright-Malécot model

"The pain in the torus" Felsenstein (1975)



$t = 100 \quad N = 972$



$t = 1000 \quad N = 293$

The Λ-Fleming-Viot process

Before writing down a spatial model, need a model in which significant proportion population can be replaced in each reproduction event.

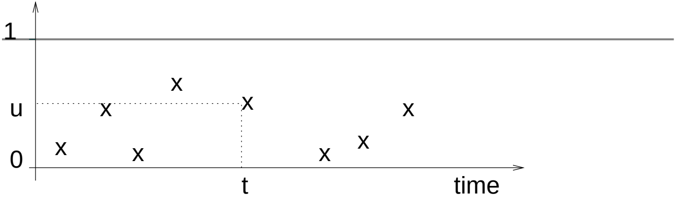
Donnelly & Kurtz (1999)
Bertoin & Le Gall (2003)

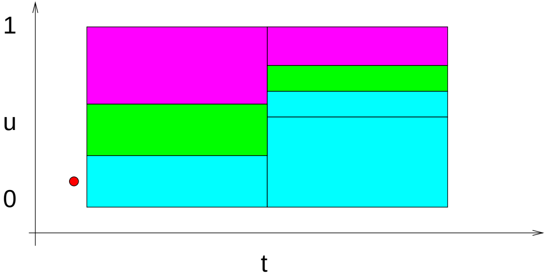
State $\{\rho(t,\cdot) \in \mathcal{M}_1(K), t \geq 0\}$.

- Poisson point process intensity $dt \otimes u^{-2}\Lambda(du)$
- individual sampled at random from population
- proportion u of population replaced by offspring of chosen individual

$$\rho(t,\cdot) = (1-u)\rho(t-,\cdot) + u\delta_k.$$

The Λ-Fleming-Viot process

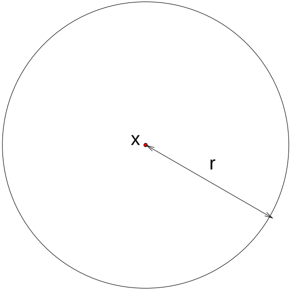




The (neutral) spatial Λ-Fleming-Viot process

State $\{\rho(t,x,\cdot) \in \mathcal{M}_1(K), x \in \mathbb{R}^d, t \geq 0\}$.
Π Poisson point process rate $dt \otimes dx \otimes \xi(dr, du)$ on $[0,\infty) \times \mathbb{R}^d \times [0,\infty) \times [0,1]$.

Dynamics: for each $(t,x,r,u) \in \Pi$,

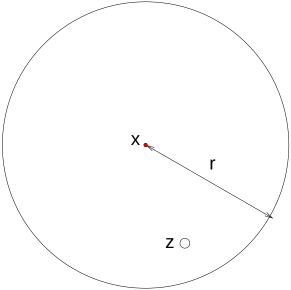


The (neutral) spatial Λ-Fleming-Viot process

State $\{\rho(t,x,\cdot) \in \mathcal{M}_1(K), x \in \mathbb{R}^d, t \geq 0\}$.
Π Poisson point process rate $dt \otimes dx \otimes \xi(dr, du)$ on $[0,\infty) \times \mathbb{R}^d \times [0,\infty) \times [0,1]$.

Dynamics: for each $(t,x,r,u) \in \Pi$,

- $z \sim U(\mathcal{B}_r(x))$

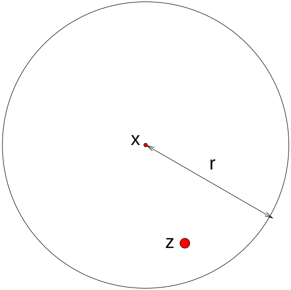


The (neutral) spatial Λ-Fleming-Viot process

State $\{\rho(t,x,\cdot) \in \mathcal{M}_1(K), x \in \mathbb{R}^d, t \geq 0\}$.
Π Poisson point process rate $dt \otimes dx \otimes \xi(dr, du)$ on $[0,\infty) \times \mathbb{R}^d \times [0,\infty) \times [0,1]$.

Dynamics: for each $(t,x,r,u) \in \Pi$,

- $z \sim U(\mathcal{B}_r(x))$
- $k \sim \rho(t-, z, \cdot)$.



The (neutral) spatial Λ-Fleming-Viot process

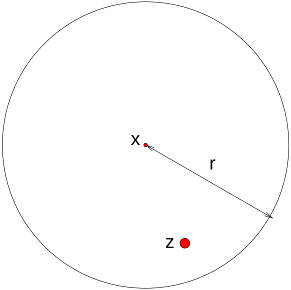
State $\{\rho(t,x,\cdot) \in \mathcal{M}_1(K), x \in \mathbb{R}^d, t \geq 0\}$.
Π Poisson point process rate $dt \otimes dx \otimes \xi(dr,du)$ on $[0,\infty) \times \mathbb{R}^d \times [0,\infty) \times [0,1]$.

Dynamics: for each $(t,x,r,u) \in \Pi$,

- $z \sim U(\mathcal{B}_r(x))$
- $k \sim \rho(t-,z,\cdot)$.

For all $y \in \mathcal{B}_r(x)$,

$$\rho(t,y,\cdot) = (1-u)\rho(t-,y,\cdot) + u\delta_k.$$



Incorporating (fluctuating) selection

Specialise to $K = \{a, A\}$.

$w(t,x)$ proportion of type a at the point x at time t.

Environment:
$Z(t,\cdot)$ random field with

$$\mathbb{P}\left[Z(t,x) = 1\right] = \mathbb{P}\left[Z(t,x) = -1\right] = \frac{1}{2},$$
$$\mathbb{E}\left[Z(t,x)Z(t,y)\right] = g_0(x,y).$$

At times t of Poisson Process $\Pi^{\mathtt{env}}$, sample independent copy of Z.

Dynamics of allele frequencies

For each $(t, x, r, u) \in \Pi$, if

$$\overline{w}(t-, x) = \frac{1}{|\mathcal{B}_r(x)|} \int_{\mathcal{B}_r(x)} w(t-, y) dy,$$

- type K of parent of event is type a with probability

$$\frac{(1 + sZ(t, x))\overline{w}(t-, x)}{1 + sZ(t, x)\overline{w}(t-, x)}$$

  otherwise it is type A.

For all $y \in \mathcal{B}_r(x)$,

$$w(t, y) = (1 - u)w(t-, y) + u\delta_{k=a}.$$

Zooming out

Interested in large scale phenomena, so e.g. shape of events is not important. For simplicity, radius and impact of events deterministic.

Set

$$u_n = \frac{u}{n^{1/3}}, \qquad s_n = \frac{\mathbf{s}}{n^{2/3-\alpha}}.$$

$$w_n(t,x) = w\big(nt, n^{1/3}x\big), \qquad Z_n(t,x) = Z\big(n^{2\alpha}t, n^{1/3}x\big).$$

Correlations in environment such that

$$\lim_{n\to\infty} g_n\big(n^{1/3}x, n^{1/3}y\big) = g(x,y).$$

- Space-time scaling for w is diffusive;
- if selection didn't fluctuate, would need $\alpha = 0$;
- we assume long-range correlations in environment.

Theorem: scaling limit

Suppose $\overline{w}_n(0,x) \Rightarrow w(0,x)$. Then $\overline{w}_n \Rightarrow w$ where

$$\begin{aligned} \mathrm{d}w = \big(\kappa_r u \Delta w + u^2 \mathbf{s}^2 w(1-w)(1-2w)\big)\,\mathrm{d}t \\ + \sqrt{2} u \mathbf{s} V_r w(1-w) W(\mathrm{d}t, \mathrm{d}x) \\ + \mathbf{1}_{d=1} u V_r \sqrt{w(1-w)} \mathcal{W}(\mathrm{d}t, \mathrm{d}x), \end{aligned}$$

where W is a coloured noise with quadratic variation given by

$$\langle W(\phi) \rangle_t = t \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} g(x,y)\phi(x)\phi(y)\mathrm{d}x\mathrm{d}y,$$

and $\mathcal{W}$ is a space-time white noise.
c.f. non-spatial case

$$\mathrm{d}p = \mathbf{s}^2 p(1-p)(1-2p)\mathrm{d}t + \sqrt{p(1-p)}\mathrm{d}B_t^1 + \sqrt{2}\mathbf{s}p(1-p)\mathrm{d}B_t^2.$$

Some remarks

What we'd really like is a way to model genealogical trees relating individuals in the populations.

Under quite general conditions, there should be a stationary distribution in two dimensions:

$$\mathrm{d}p = \mathbf{s}^2 p(1-p)(1-2p)\mathrm{d}t + \sqrt{2}\alpha \mathbf{s} p(1-p)\mathrm{d}B_t^2$$

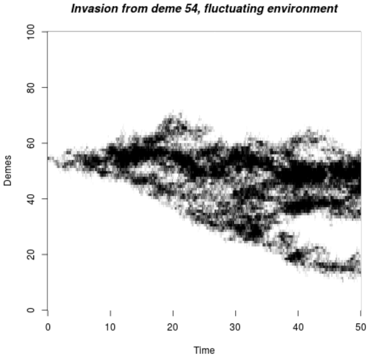
has a non-trivial stationary distribution as soon as $\alpha < 1$.

Without space, there is a branching and annihilating dual...

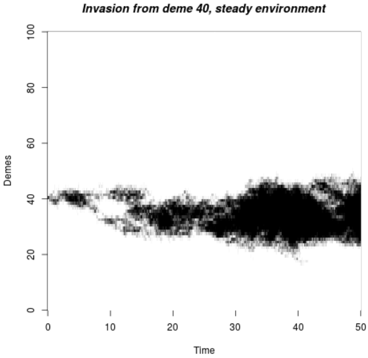
- Genetic drift disappears in $d \geq 2$;
- if follow a rare mutant, recover 'superprocess in random environment' of Mytnik 1996;
- more generally, can write down evolution of subset of a-individuals ('tracer dynamics').

Experiments with an individual based model

Population in discrete demes on a torus on $\mathbb{Z}$.
Two regions with completely anticorrelated environments.
Two scenarios: (i) environment resampled on timescale on the order of generations, (ii) environment fixed.



Invasion from deme 54, fluctuating environment



Invasion from deme 40, steady environment