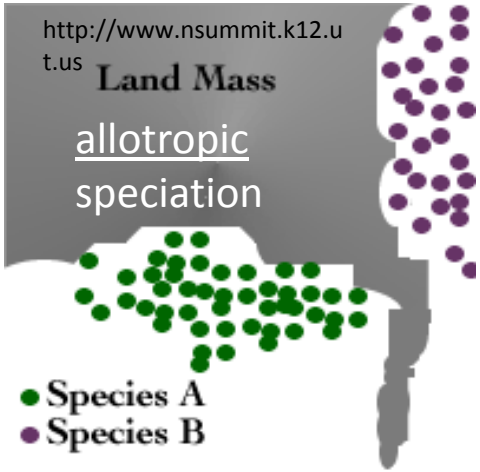
Structured environments, spatial gene drives and pushed genetic waves

1. **Range Expansions in Structured Environments (pulled genetic waves)**

Wolfram Moebius *Andrew Murray*

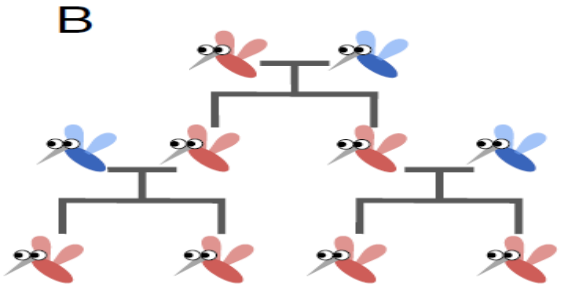
Moebius, Murray and drn, PLoS Comp. Biol. 11: e1004615 (2015)



2. **Spatial gene drives and pushed genetic waves**

Hidenori Tanaka *Howard Stone*

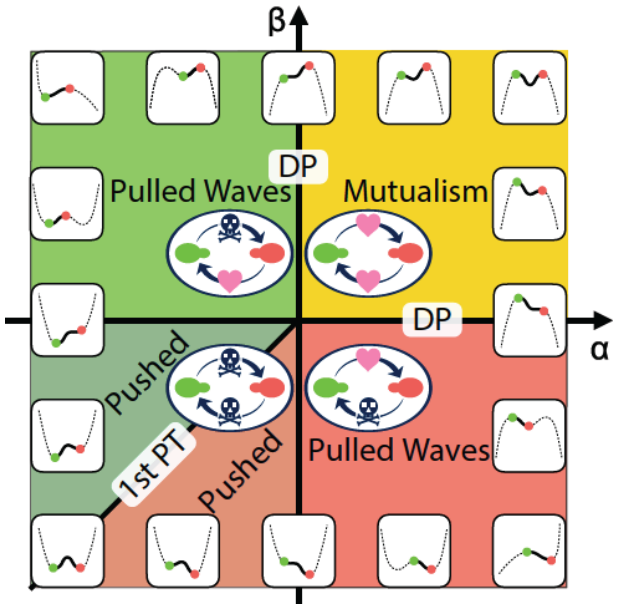
Tanaka, Stone and drn, PNAS 114, 452 (2017)



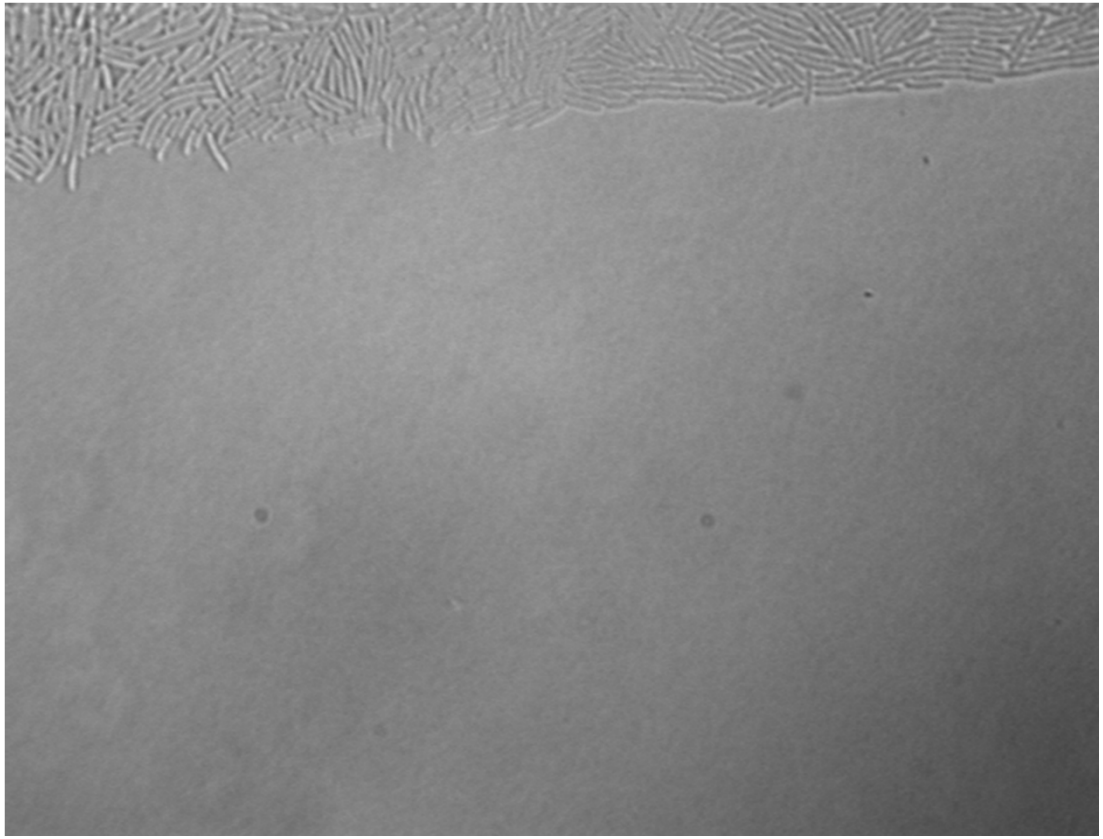
pushed waves and <u>sympatric</u> speciation

3. **Pushed waves, pulled waves and inter-species interactions**

(K. Korolev, M. Lavrentovich & drn)

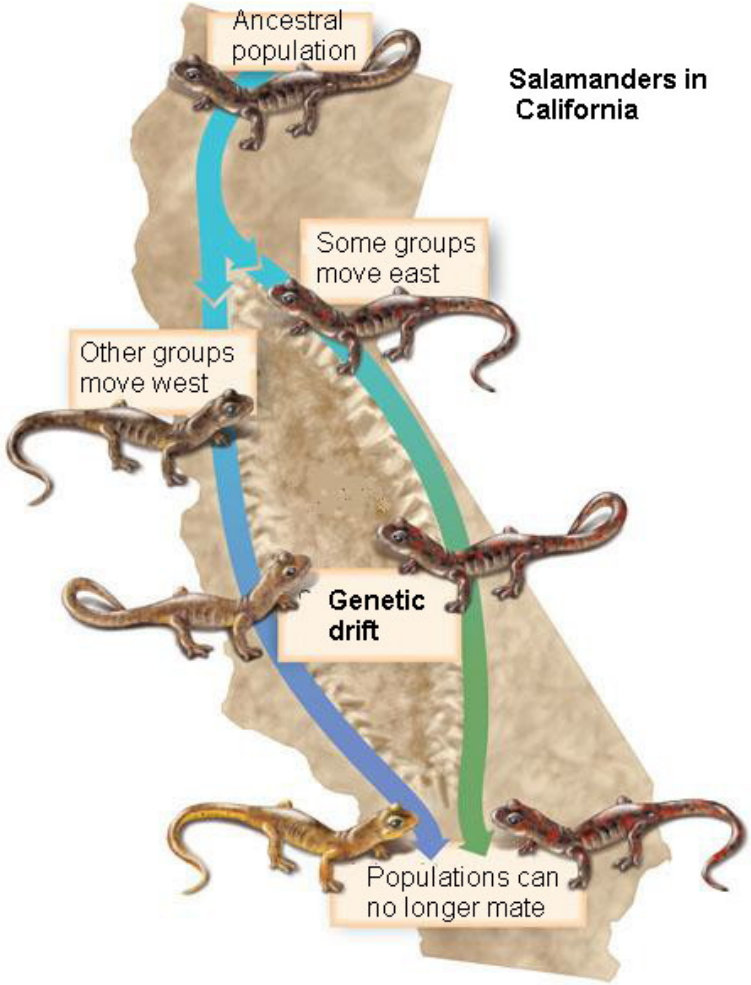


Range expansions can be unstructured or structured



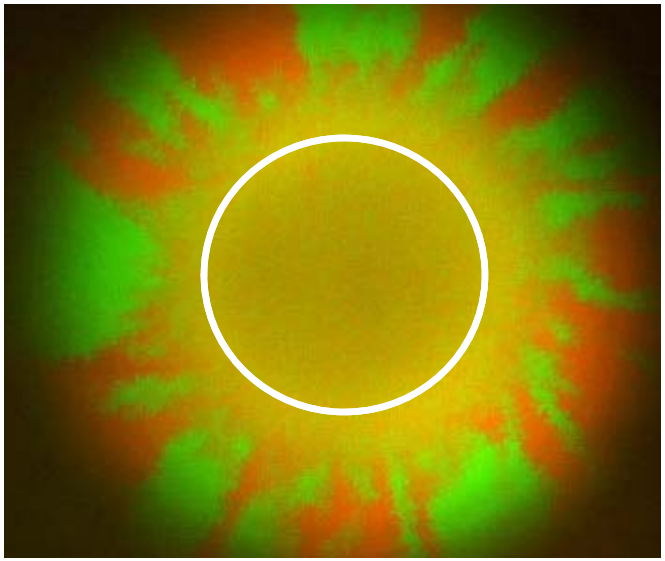
unstructured range expansion
(E. coli on a Petri dish, O. Hallatschek & drn Ramanathan lab)

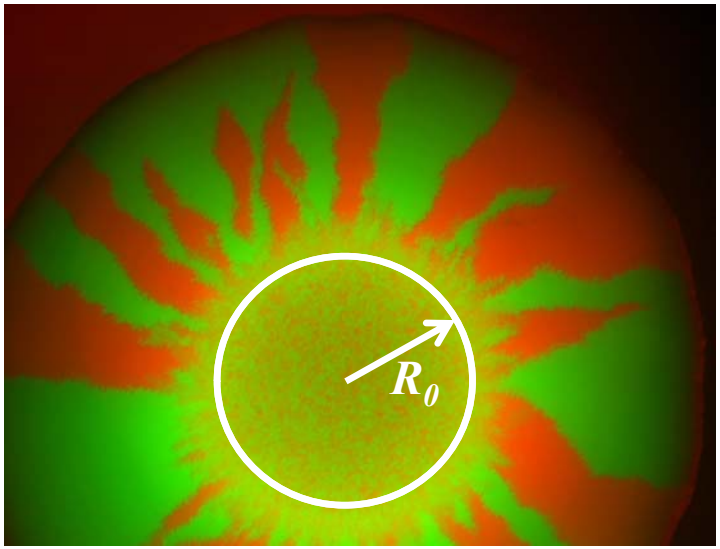
http://legacy.hopkinsville.kctcs.edu

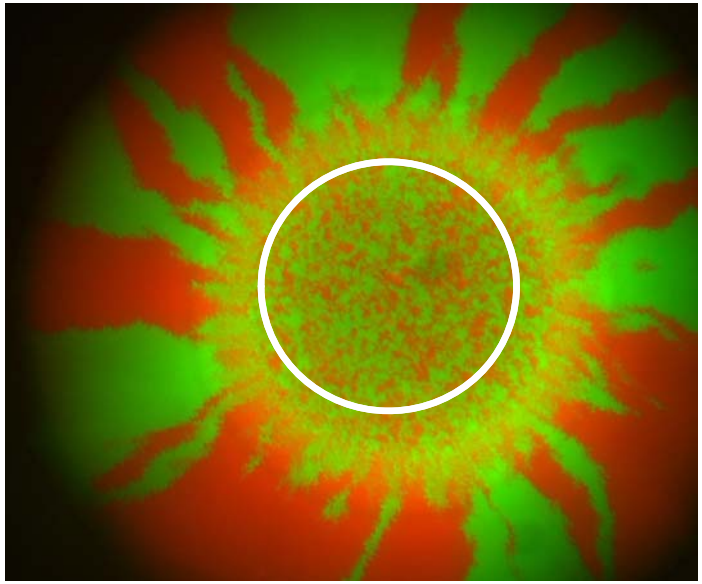


structured range expansion
(leads to allopatric speciation)

Reproducible structured environments from range expansions?







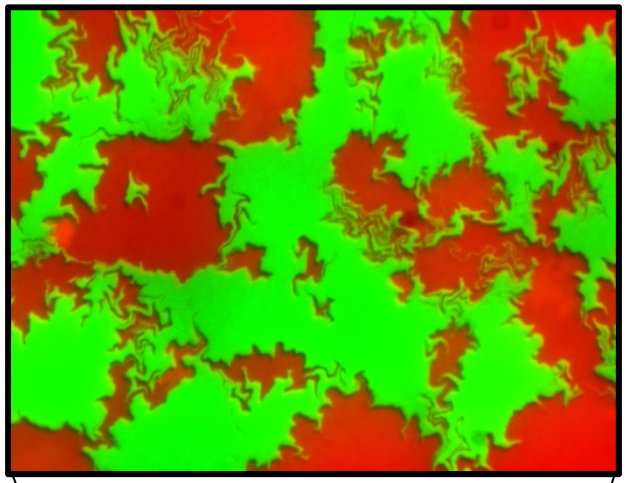
(can infer the radius R_0 of the homeland from data at the boundary:

$N_{\sec} = \sqrt{\pi R_0 v / 2D_W}$

Hallatschek & drn)

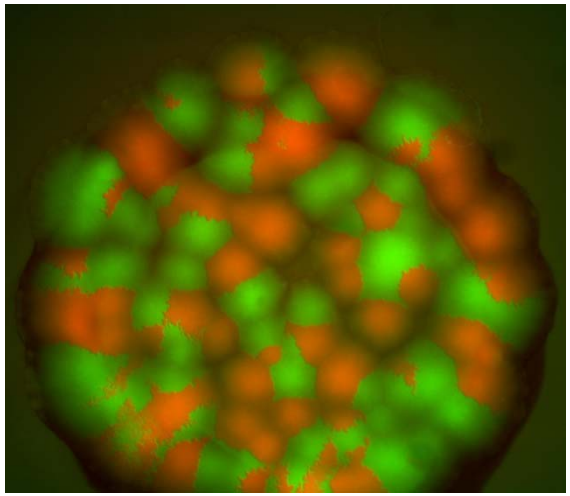
Will focus here on the structured environment <u>inside</u> the homeland…

Can the “homeland” be regarded as an “ecological landscape”?

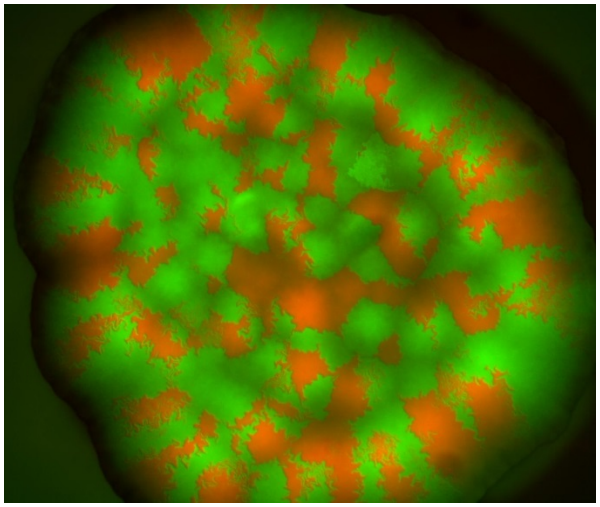


50-50 circular inoculants
24 hours after inoculation;
(~ 1mm spot size)

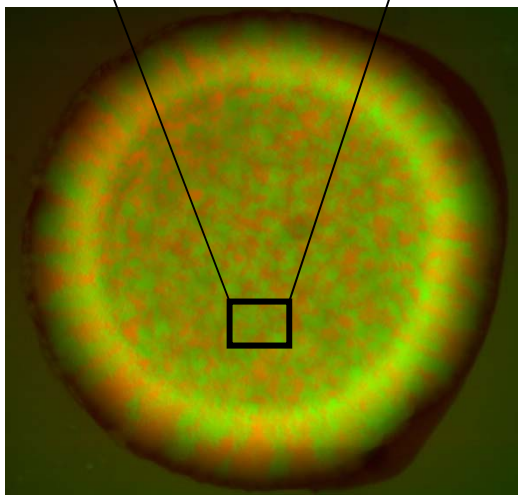
~25 green &
25 “red” viable
founder bacteria



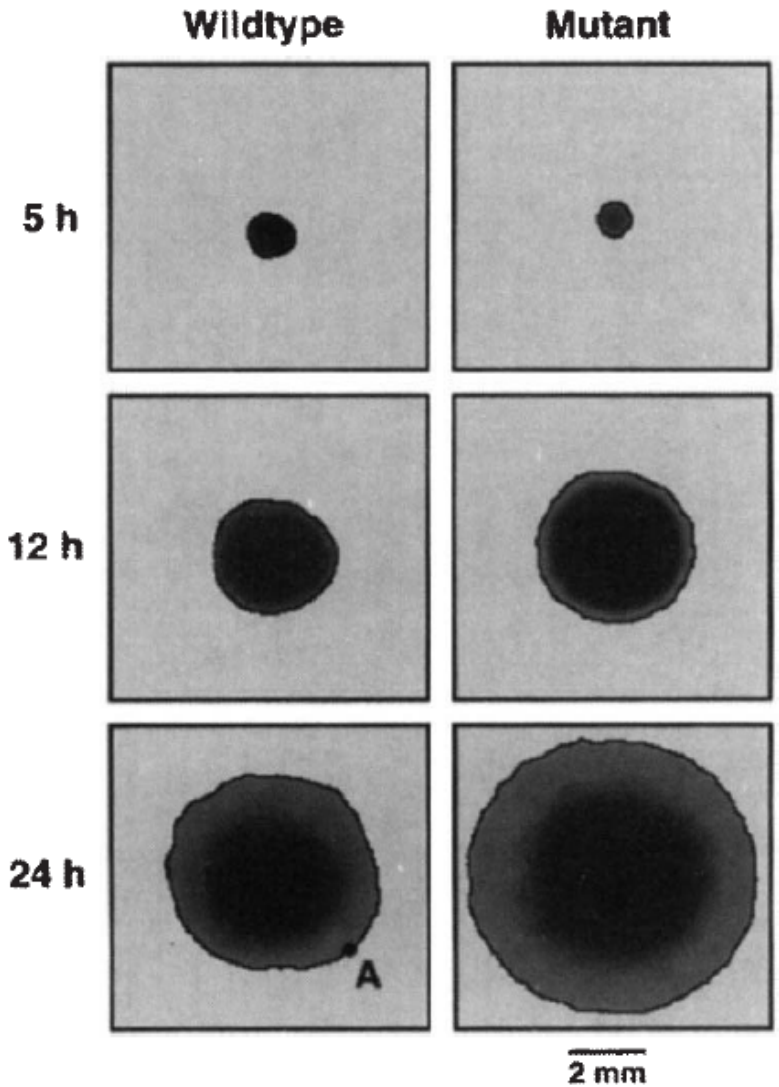
~250 green &
250 “red” viable
founder bacteria



~2500 green &
2500 “red” viable
founder bacteria

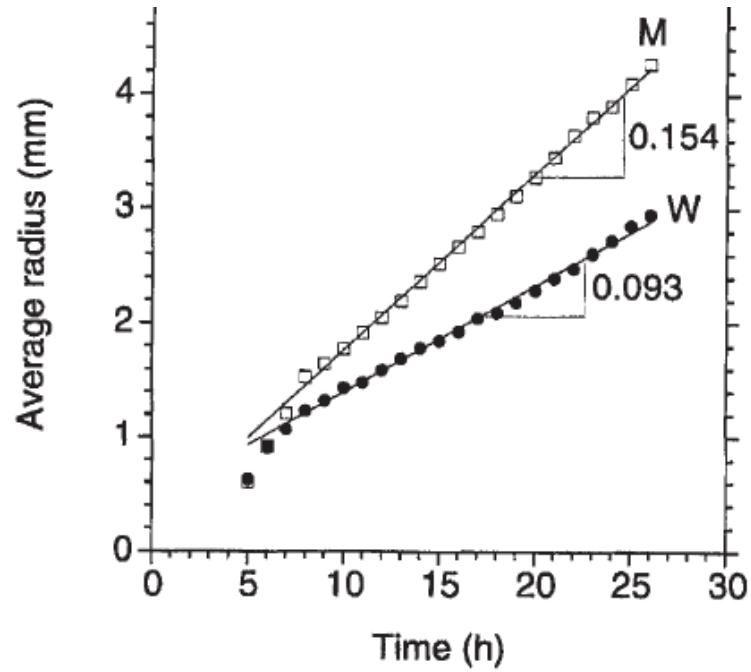


Front velocity of an T7 plaque invading a bacterial lawn



Mutant isolated from stab taken at point A of the wild type descendants

Y. Lee and J. Yin, Nat. Biotech. ***14,*** *491 (1996).*



J. Yin and J. S. McCaskill, Biophys. J. ***61,*** *1540 (1992).*

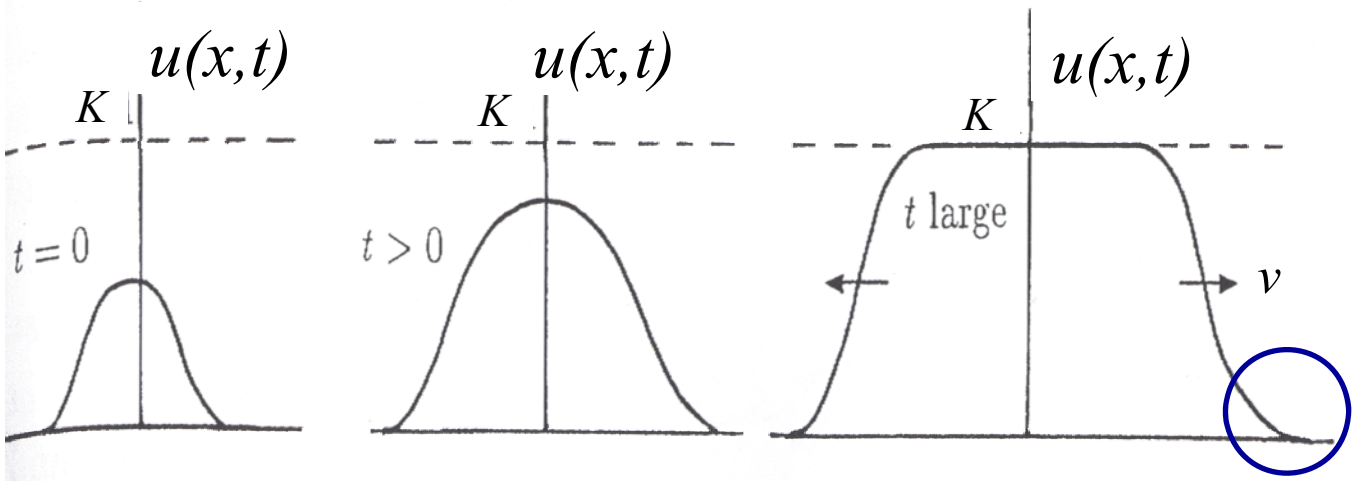
$$V + B \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} I \xrightarrow{k_2} Y \cdot V$$

Infection resembles a Fisher population wave; T7 susceptible and T7 bacteria exist…

Fisher wave of plaque growth model in d = 1

$u(x,t)$ = viral concentration in space and time; τ_g = gen. time

K = carrying capacity; $\partial_t u(x,t) = D\partial_x^2 u(x,t) + u(x,t)[1 - u(x,t)/K]/\tau_g$



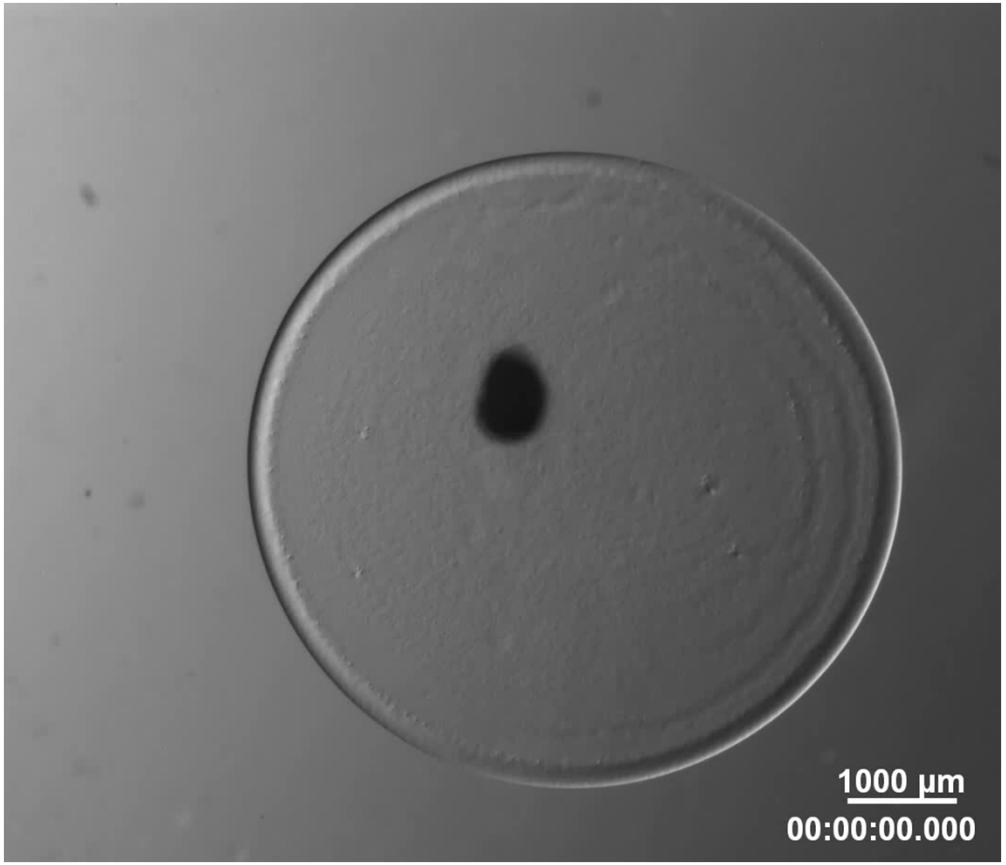
Schematic time development of a wavefront solution of Fisher's equation on the infinite line. (J.D. Murray, Mathematical Biology)

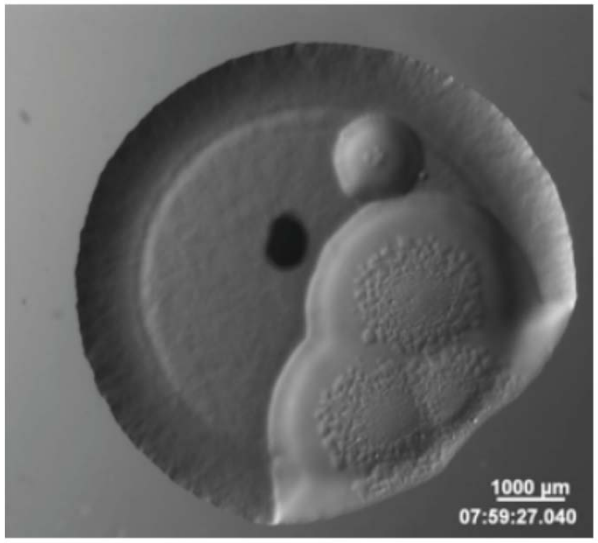
Interface velocity v is $2\sqrt{D/\tau_g}$ *Interface width is* $\sqrt{D\tau_g}$

Classic example of a *pulled* wave dominated by growth at the "foot": $v \sim \sqrt{D\tau_g}/\tau_g$

(However, gene drives will provide examples of both pulled and *pushed* genetic waves)

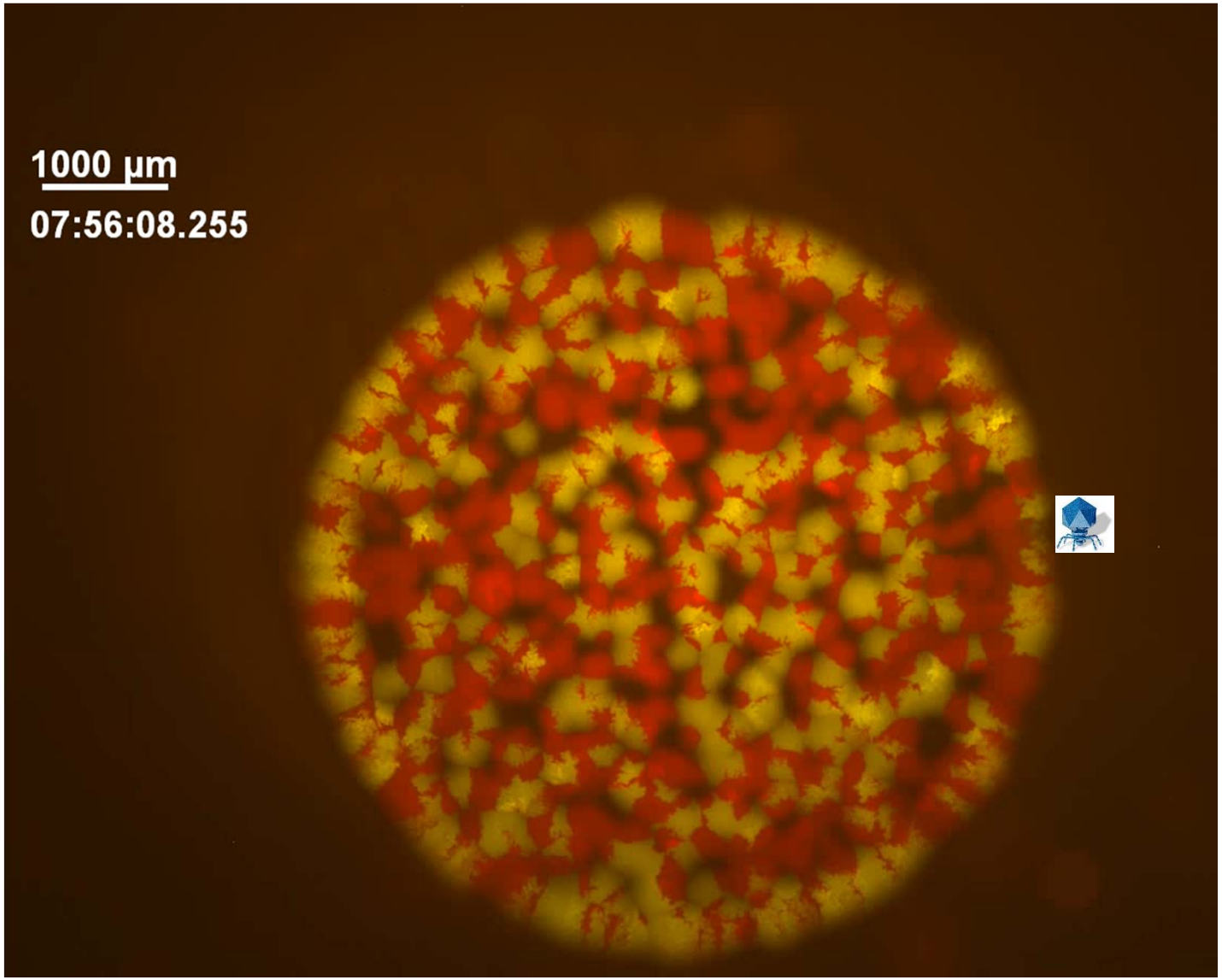
“Refraction” of a T7 population wave….



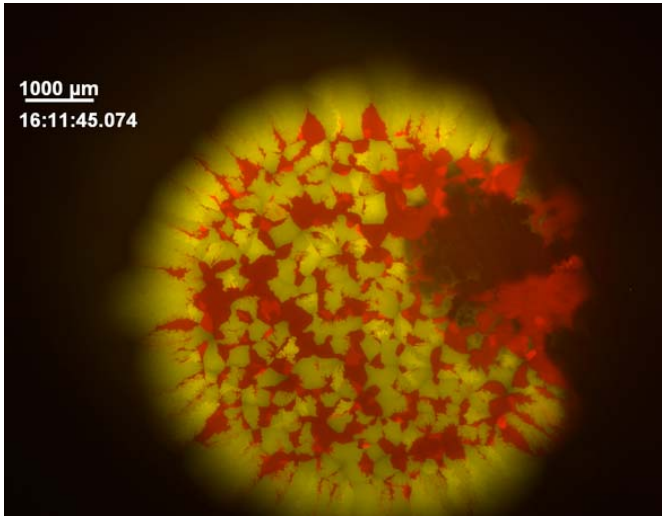


Spread of a T7 epidemic through the homeland

= T7-susceptible E. coli = T7-resistant E. coli

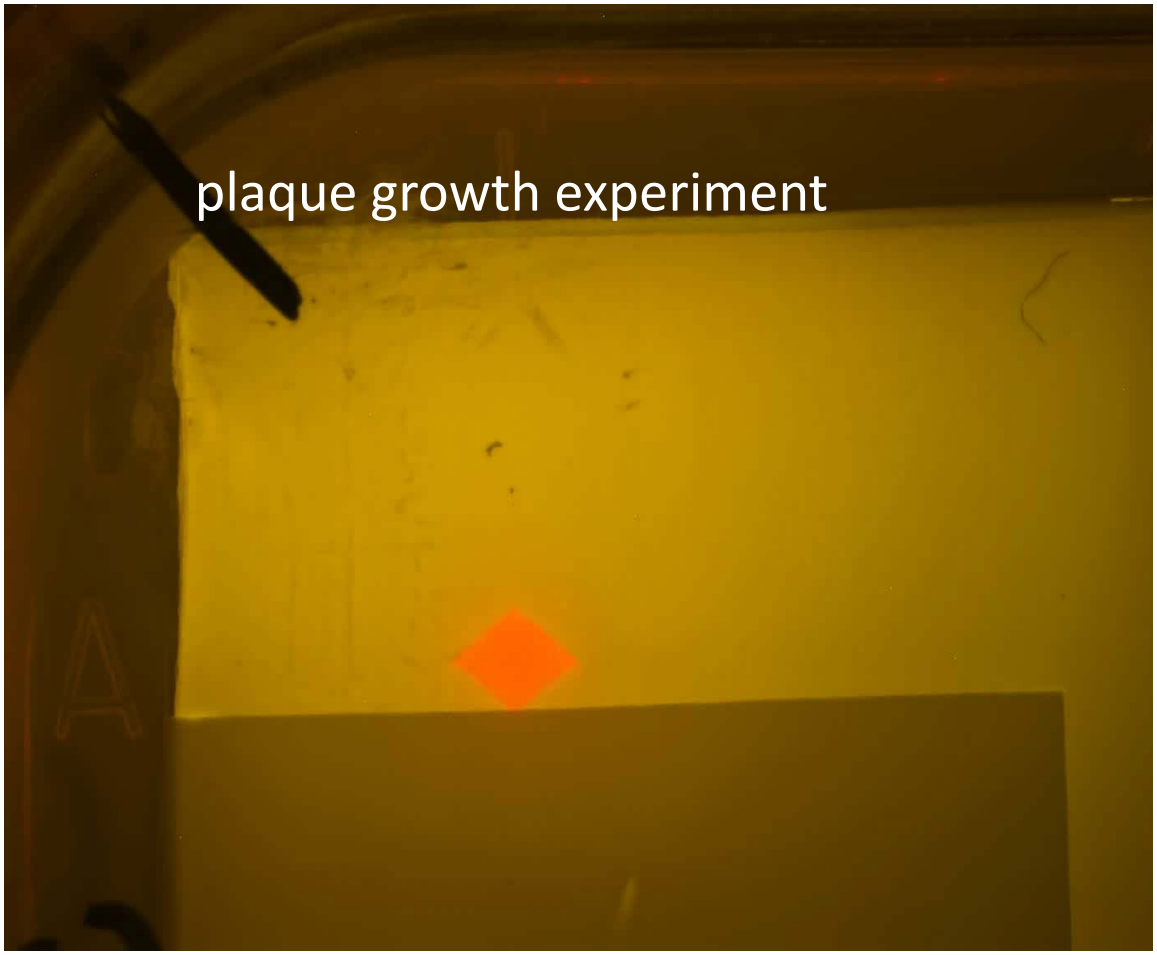


Simplify by focusing on a few well defined obstacles…



Wolfram's ink jet printer: (see also: Leibler Lab, PLoS 1, (2007)

Simplified problem: population dynamics around well-defined obstacles = "lakes"

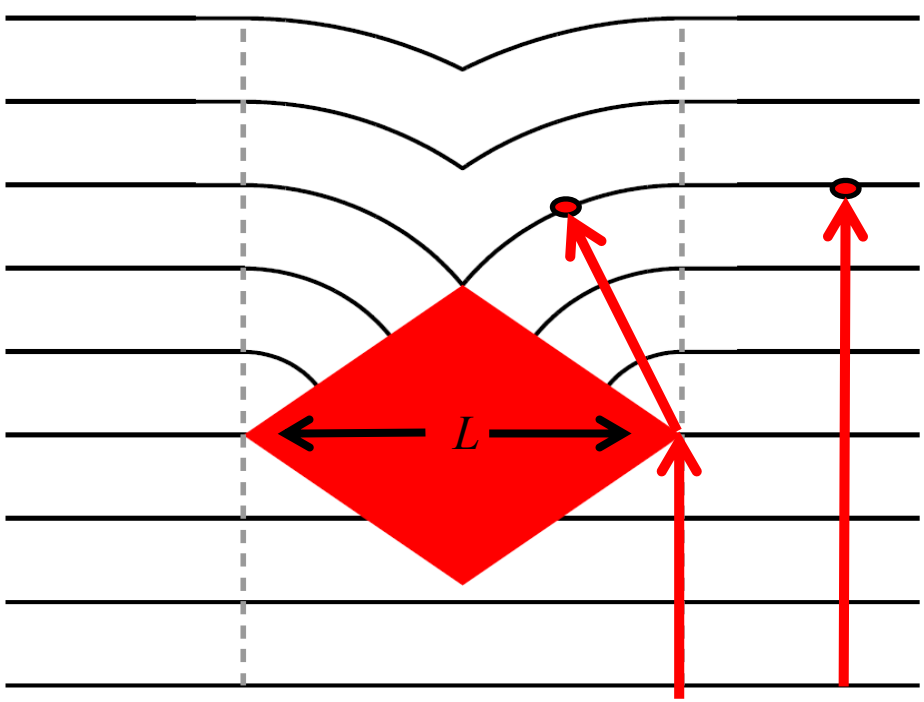


underwater caustics formed by refraction of sunlight

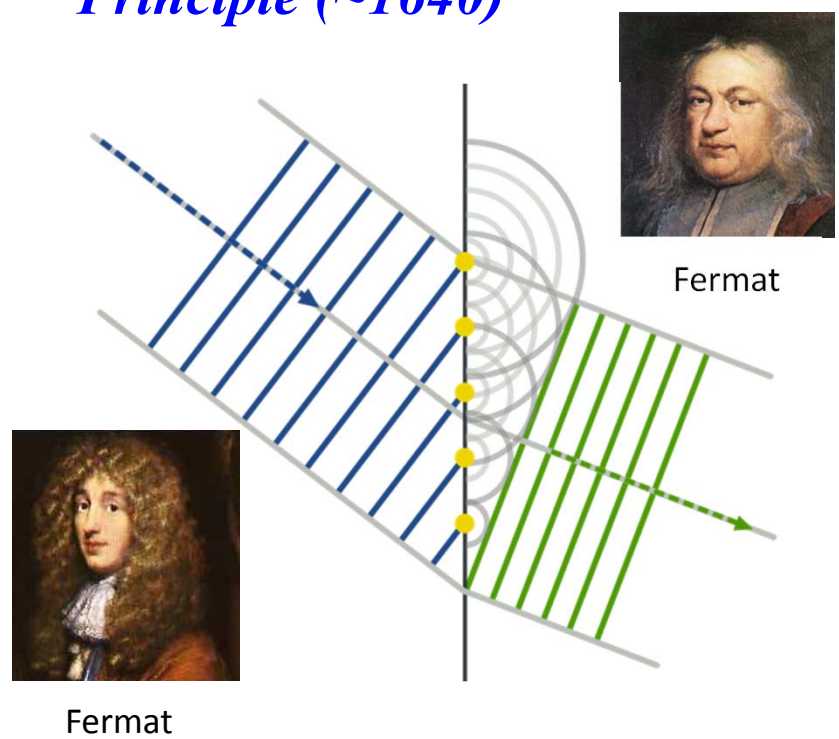
Huygens-Fermat's principle for viral plaques

- ***At any given point on the frontier of the advancing wave of viral destruction, ask "where did you ancestors come from?"***

- ***Huygens-Fermat "principle of least time" reformulated in a Darwinian way… "survival of the fastest"***

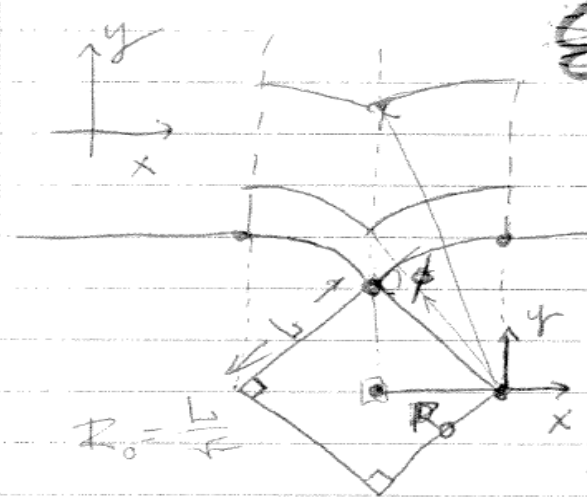


Huygens/Fermat Refraction Principle (~1640)



ENVELOPEMENT OF OBSTACLES IN RANGE EXPANSIONS

DIAMOND-SHAPED OBSTACLE



$R_0 = \frac{L}{\sqrt{2}}$

$$(x - R_0)^2 + y^2 = (\sqrt{2}R_0 + vt)^2$$

$$2\,dx\,(x - R_0) + 2\,dy\,y = 0$$

$$\Rightarrow \frac{dy}{dx} = -\frac{(x - R_0)}{y} = \tan\phi$$

$$\tan\phi = \left.\frac{dy}{dx}\right|_{x=0} = \frac{R_0}{y}, \text{ but } y = \sqrt{(\sqrt{2}R_0 + vt)^2 - R_0^2}$$

$$= \sqrt{R_0^2 + 2\sqrt{2}R_0 vt + v^2t^2}$$

$$\tan\phi = \frac{L/\sqrt{2}}{\sqrt{\frac{L^2}{2} + 2Lvt + v^2t^2}}$$

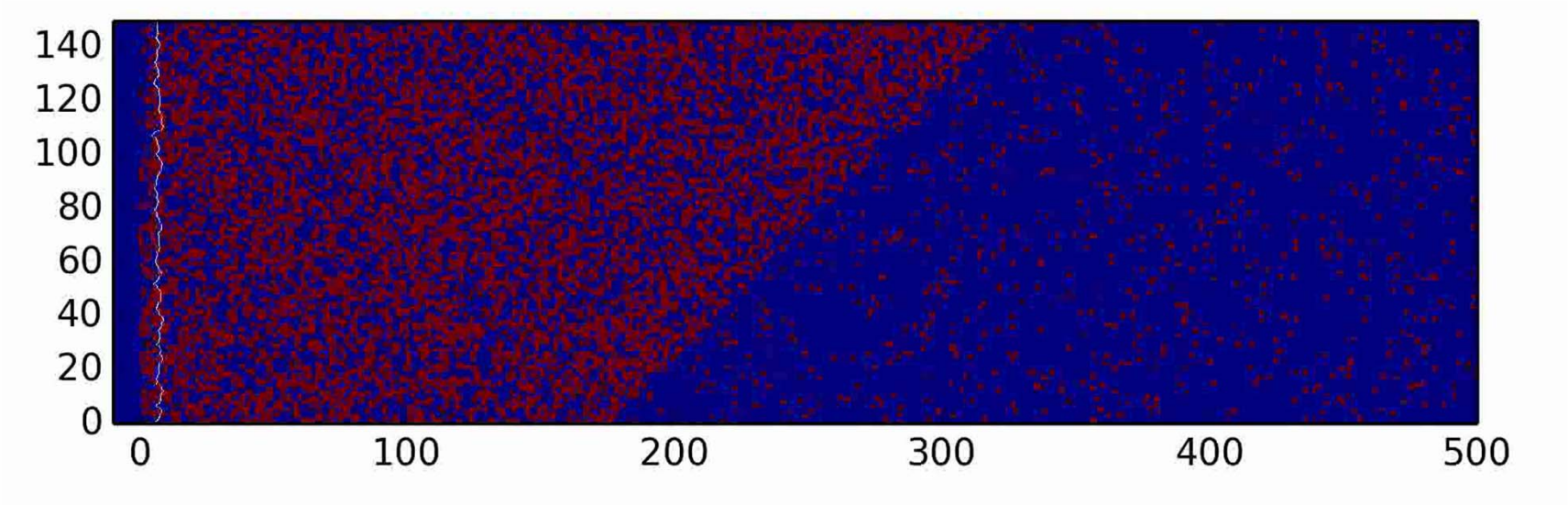
OR

$$\boxed{\tan\phi = \frac{L}{\sqrt{L^2 + 4Lvt + 2v^2t^2}}}$$

Thus $\boxed{\text{as } t \to \infty \quad \phi(t) \sim 1/2vt}$

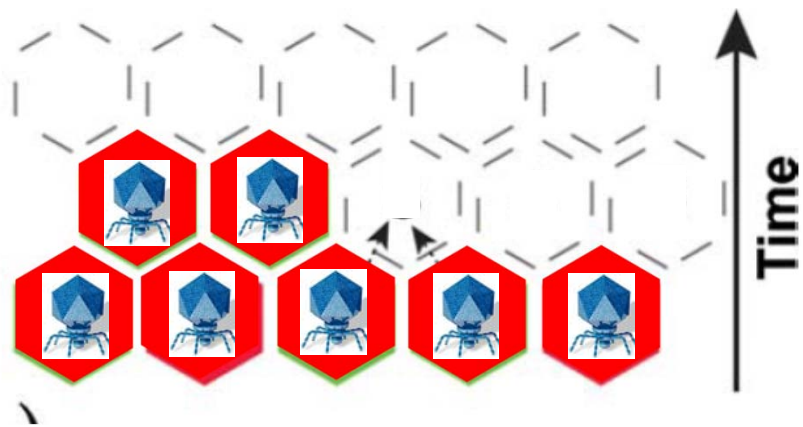
Complex arrays are characterized by an "index of refraction" = reduced population wave velocity

A dense array of obstacles leads to a reduced propagation velocity -- population waves can be bent, like light by a prism

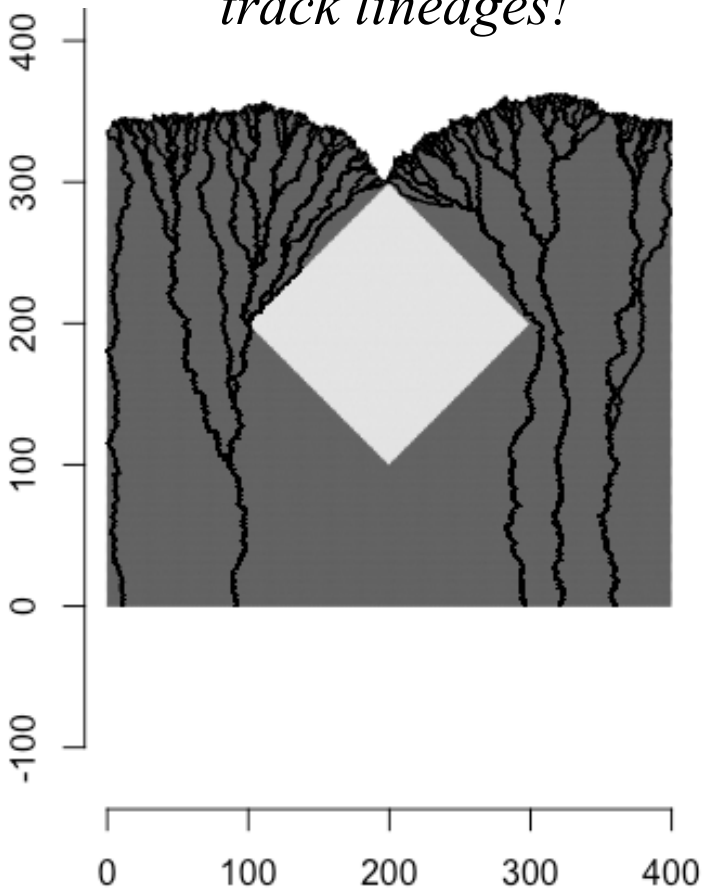


A bent wave-front emerges from the front of prism of disordered media (K. Alards, M. Moebius, F. Tesser & F. Toschi, paper in preparation)

However: the Huygens-Fermat principle neglects discreteness of viruses and cells...

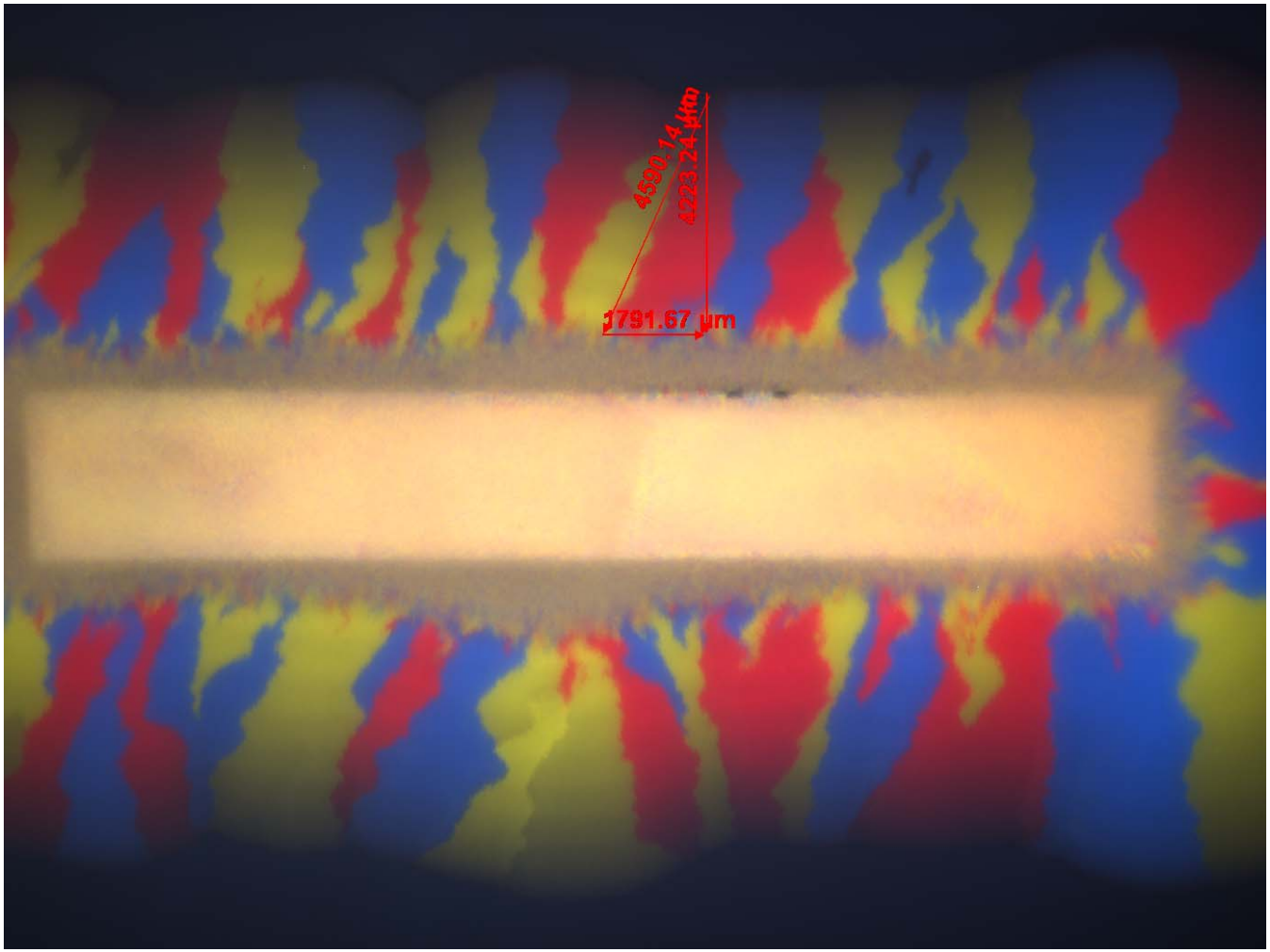


track lineages!



It's only a good first approximation, like ray optics, which neglects photons and the wave nature of light!

Population genetics and range expansions (note genetic drift!)

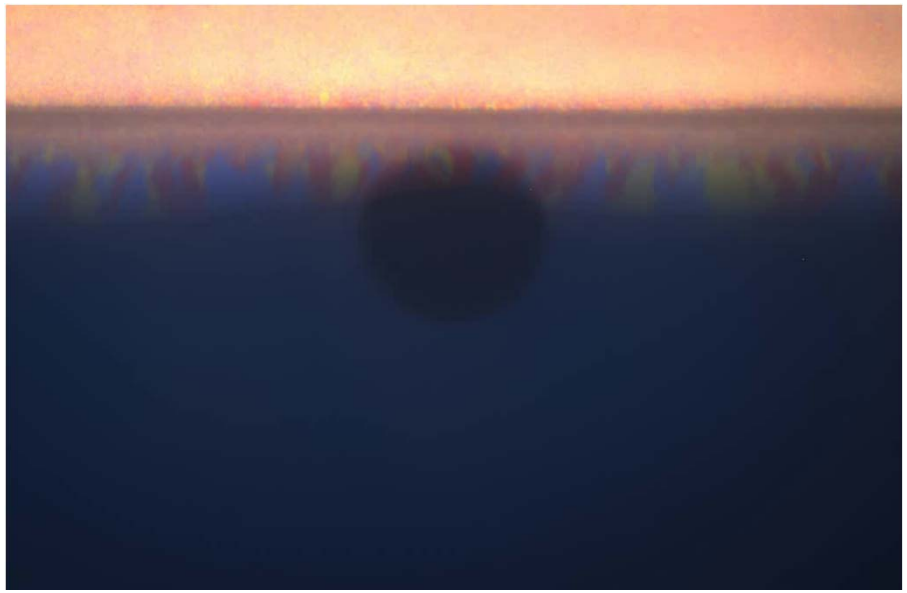


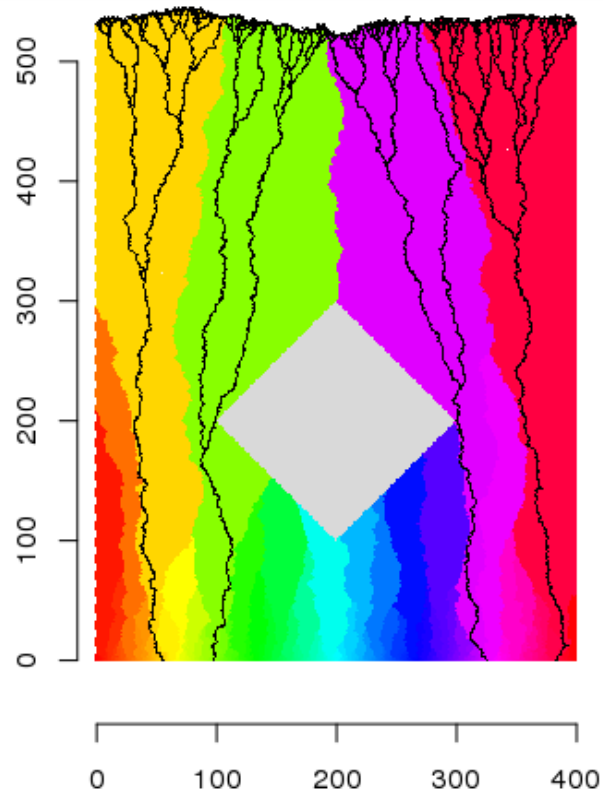
Wandering of genetic boundaries during range expansions given by a "wall diffusion constant".

$D_w \approx a^2/\tau_g$
a = cell size
τ_g = division time

What happens if we add an obstacle????

Coalescent lineages for range expansions around a "lake"





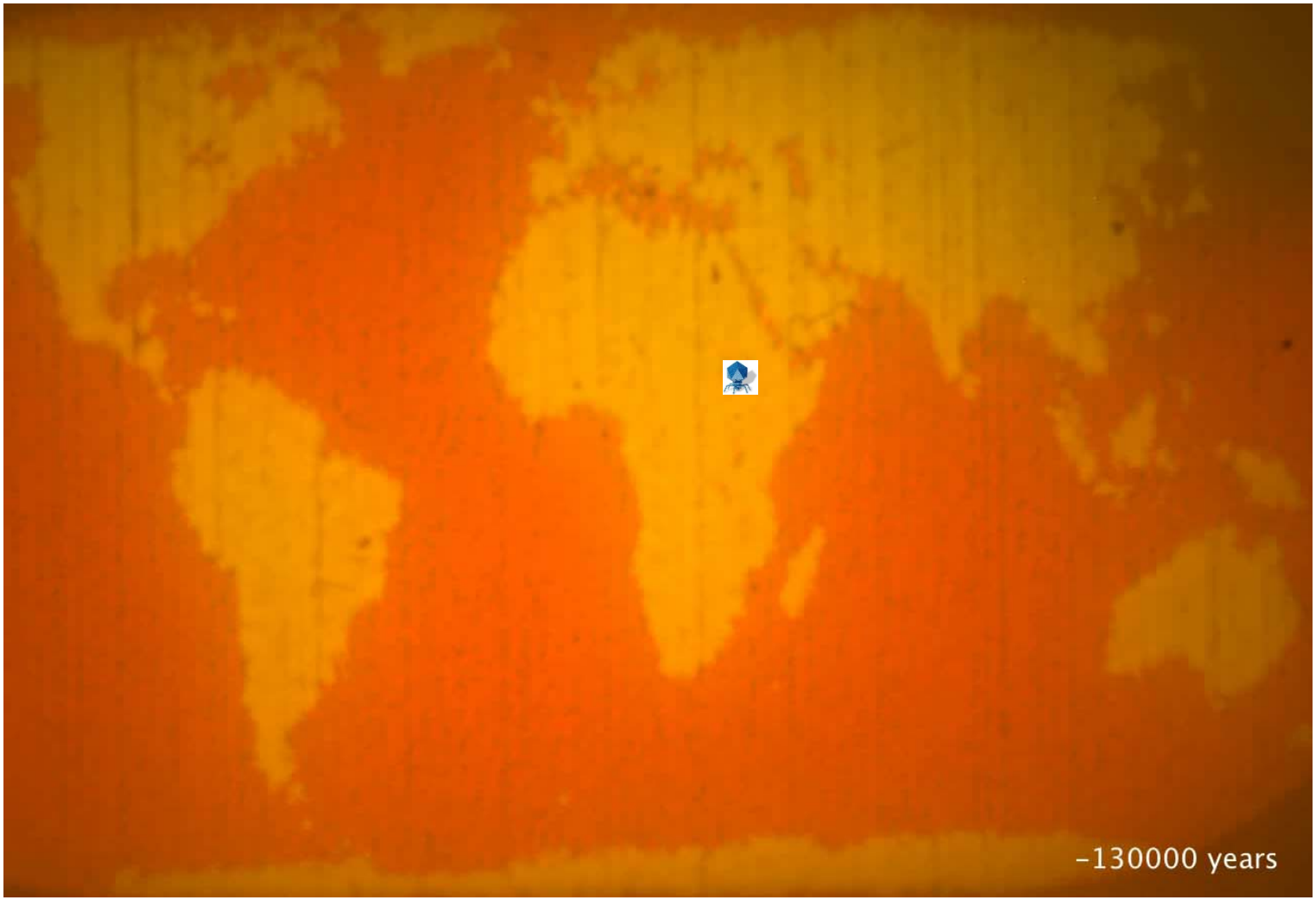
With Dan Beller: by studying lineages <u>at the front</u>, can we determine the location and width of the obstacle?

simulations ⟷ least time argument⟷ bacterial growth experiment

Genetically structured vs. spatially structured populations

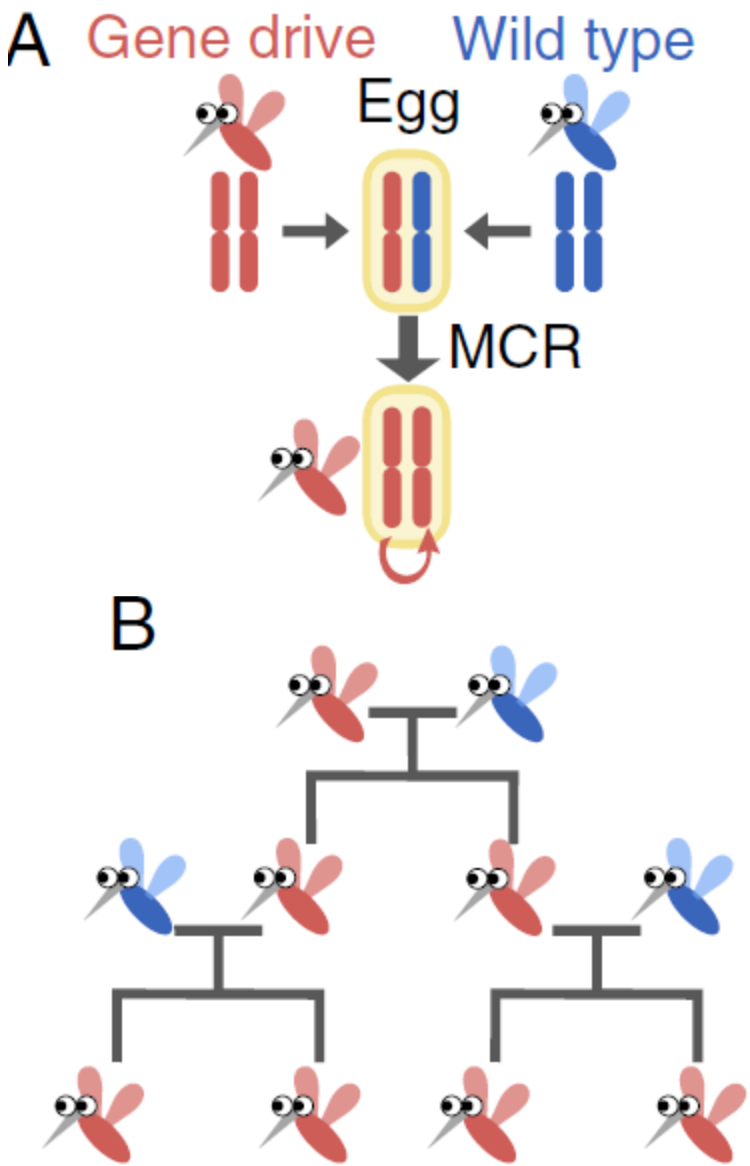
T7 "out of Africa": can print bacterial lawns for T7 in arbitrary patterns…. Wolfram Moebius, unpublished

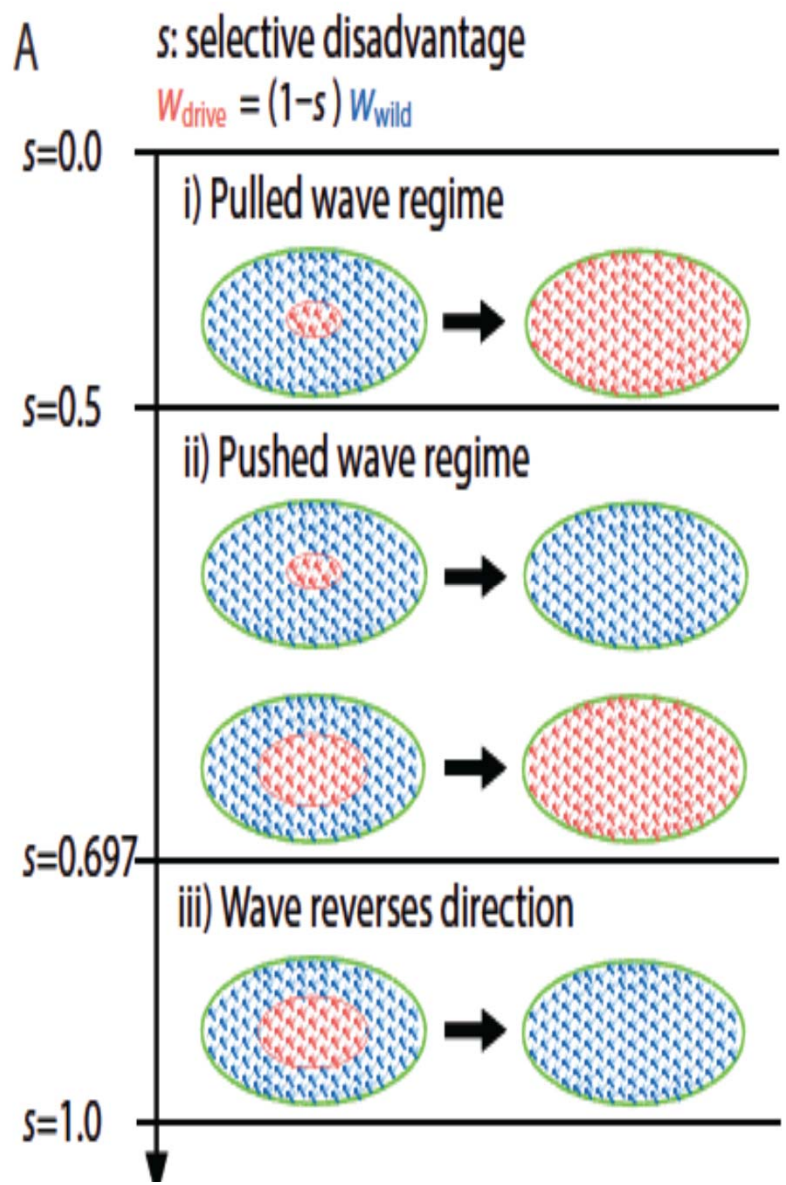
= T7-susceptible E. coli = T7-resistant E. coli



Wolfram Moebius

What about the spatial spread of gene drives?





(Neglect evolution of resistance due to NHEJ, c. f. Philipp Messer)

Non-Mendelian population genetics of drives with 100% conversion

- Consider N_{aa} homozygous wild type individuals, N_{AA} individuals homozygous for the gene drive, and N_{aA} heterozygous individuals in a region that supports a population of size $N=N_{aa}+N_{AA}+N_{aA}$. The fractions p and q of wild type and gene drive gametes are given by

$$p=\frac{2N_{aa}+N_{aA}}{2N_{aa}+2N_{aA}+2N_{AA}},\quad q=\frac{2N_{AA}+N_{aA}}{2N_{aa}+2N_{aA}+2N_{AA}}=1-p$$

- Random mating produces fertilized eggs in the ratio $p^2 : 2pq : q^2$ for diploids *(a,a), (a,A) and (A,A)* respectively. A gene drive with *c= 1 =100%* efficiency leads to only homozygotes *(a,a)* and *(A,A)* in the ratio $p^2 : q^2+2pq$.

- Let *w* describe the reproductive success in a generation time τ_g of the wild type homozygote *(a,a)* and $w'=(1-s)w$ that of the gene drive homozygote , assumed to have a selective disadvantage or fitness cost *s*. The genotype numbers in the next generation are $N_{aa}'=Nwp^2$ and $N_{AA}'=Nw(1-s)(q^2+2pq)$, which leads to

$$q'=\frac{(1-s)(2q-q^2)}{(1-s)(2q-q^2)+(1-q)^2},\quad p'=1-q' \qquad \text{when } s=0,\ \tilde{q}_{n+1}=2\tilde{q}_n(1-\tilde{q}_n),\ q_n=2\tilde{q}_n$$

the time evolution of the gene drive gamete fraction q(t) is give by

$$\frac{dq}{dt}\approx\frac{q'-q}{\tau_g}=\frac{1}{\tau_g}\frac{sq(1-q)(q-q^*)}{1-sq(2-q)},\quad q^*=(2s-1)/s$$

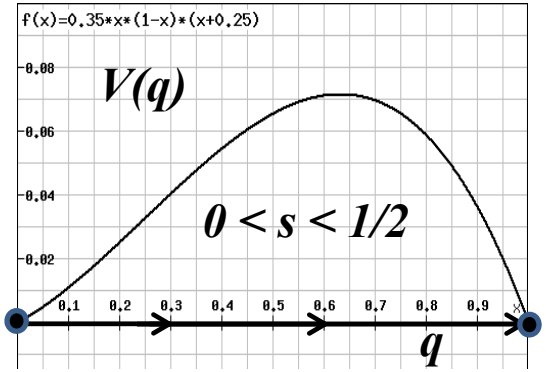
$$\tau_g\frac{dq(t)}{dt}\approx sq(1-q)(q-q^*),\quad q^*=(2s-1)/s$$

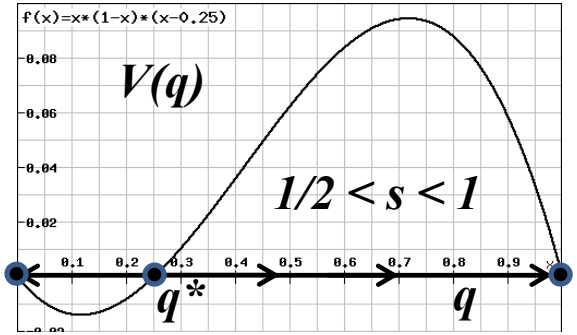
neglecting s in the denominator

Population genetics of gene drives depends on the size of the selective disadvantage s

$$\tau_g \frac{dq(t)}{dt} = sq(1-q)(q-q^*) \equiv V(q), \quad q^* = (2s-1)/s$$

- *Consider, e.g., the malaria plasmodium, carried by mosquitoes and injected with its saliva into humans.*
- *A gene drive could alter the function of a protein manufactured in the salivary gland of females from, say, type a, anesthetizing nerve cells when it bites humans, to instead type A, clogging up essential chemo-receptors in plasmodium and thus killing the plamodium.*
- *Normally, there would be a <u>selective disadvantage s</u>, i.e., a fitness cost to losing this protein, and the new type of mosquitoes would not take over.*
- *However this trait could easily spread if linked to a gene drive with conversion efficiency c close to 1.*
- *High conversion efficiencies c ≈ 1 have already been achieved with CRISPR-Cas 9 in the budding yeast S. Cereviseae and in fruit flies.*



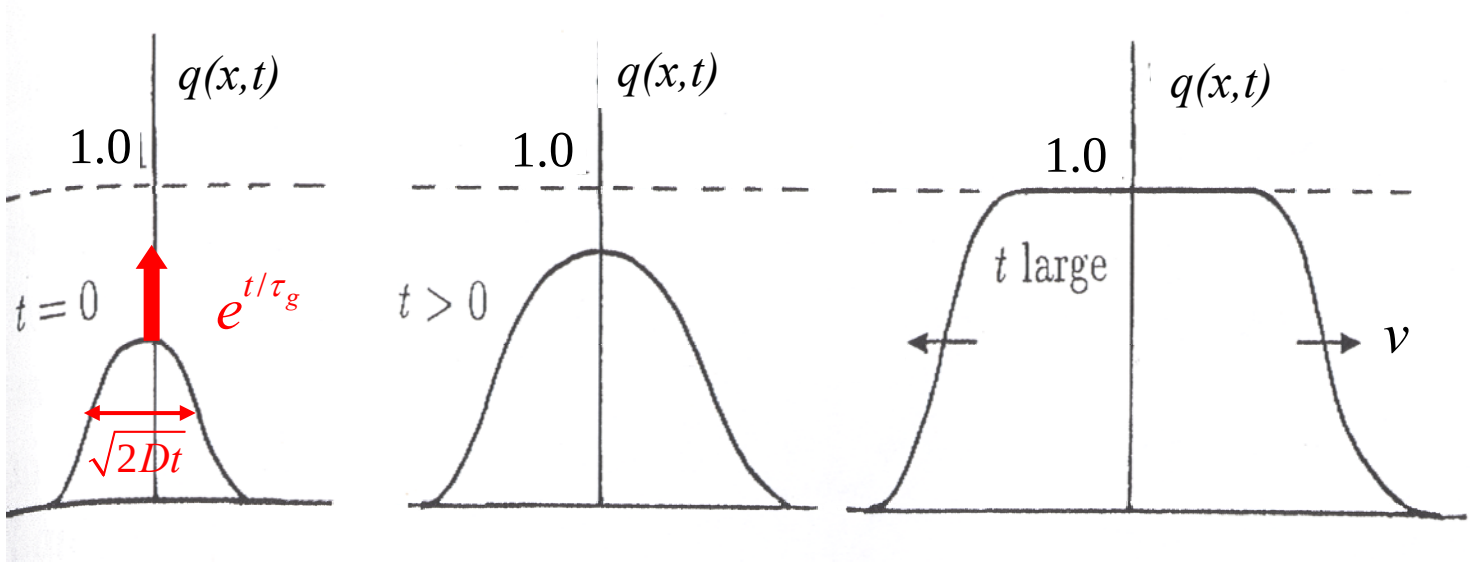


Socially responsible gene drives should be engineered such that *s* > ½, so accidentally releases from a laboratory will not spread: Unckless, Messer et al. Genetics 201, 425 (2015)

Add spatial diffusion: A pulled Fisher wave appears for s <0.5

$$\frac{\partial q(x,t)}{\partial t} \approx D\nabla^2 q + \left(s/\tau_g\right) q(1-q)(q-q^*); \quad q^* = (2s-1)/s,$$

$D =$ diffusion constant. e.g., for $s << 1$, $\dfrac{\partial q(x,t)}{\partial t} \approx D\nabla^2 q + q(1-q)/\tau_g$



Schematic time development of a wavefront solution of Fisher's equation on the infinite line. (J.D. Murray, Mathematical Biology)

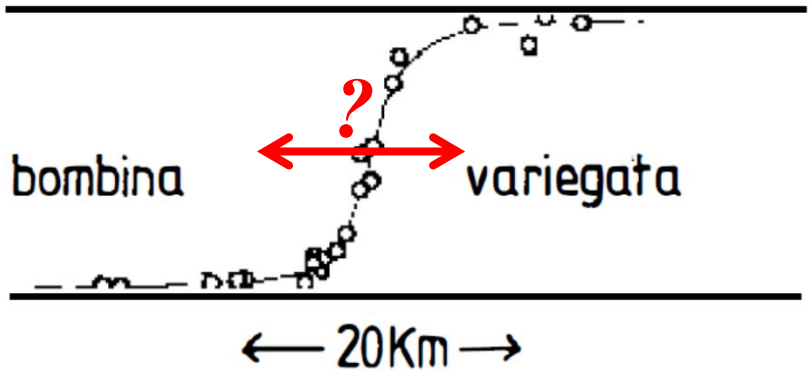
Front velocity = $v_F = 2\sqrt{D/\tau_g}$, Interface width = $\sqrt{D\tau_g}$

However, qualitatively new behavior (a pushed wave!) appears when s > 1/2…

<u>Pushed</u> genetic waves were originally studied in "hybrid zones"

N. H. Barton & G. M. Hewitt Ann. Rev. Ecol. & Sys. **16**, 113 (1985)

- Hybrid zones are narrow regions in which genetically distinct populations meet, mate and produce hybrids. Hundreds of examples known. (e.g., the grasshopper *Podisma pedestris*, the butterfly *Heliconis.)* Hybrid zones can be a kilometers thick and hundreds of kilometers long…



Inferred profile from electrophoretic variations across the hybrid zone of the toads *Bombina bombina* and *Bombina variegata* near Cracow, Poland

- *Which way the interface moves depends on more than just the selective advantage – for example, recombination near the interface can break up favorable clusters of genes.*
- *In some cases, boundaries can exhibit a kind of line tension, as well as a pressure to advance in a particular direction. Promotes sympatric speciation!*

Pushed Fisher waves ($s > ½$)

$$\frac{\partial q(x,t)}{\partial t} \approx D\nabla^2 q + \left(s/\tau_g\right) q(1-q)(q-q^*);\ q^* = (2s-1)/s,$$

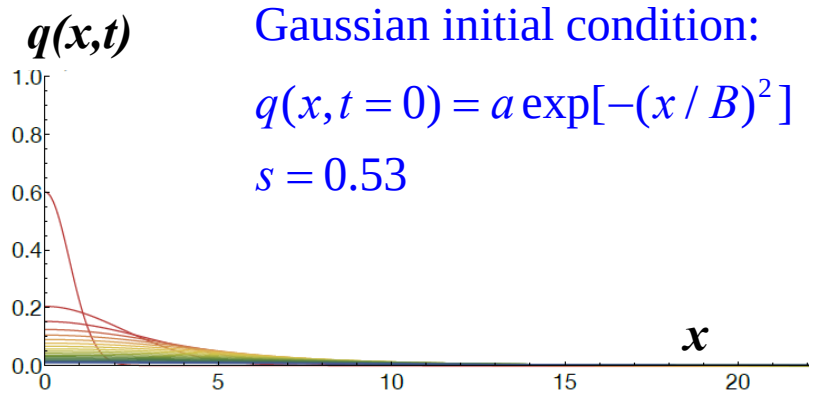
an exact traveling wave solution exists!

$$q(x,t) = \frac{1}{1+\exp[(x-v_E t)\sqrt{s/2\tau_g D}]};\quad v_E = \sqrt{sD/2\tau_g}\,(2-3s)$$

N. H. Barton & M. Turelli, The American Naturalist **178**, E48 (2011)

*A "critical propagule" size for the gene drive is required to nucleate the wave

*Wave reverses direction if s exceeds a maximum selective disadvantage $s_m = 2/3$

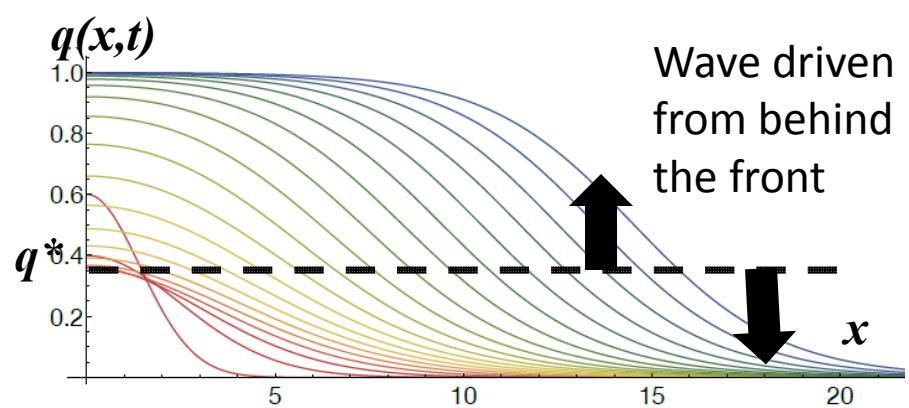


Gaussian initial condition:

$q(x,t=0) = a\exp[-(x/B)^2]$

$s = 0.53$

(a) Too small to nucleate: $a = 0.6, B = 1.0$



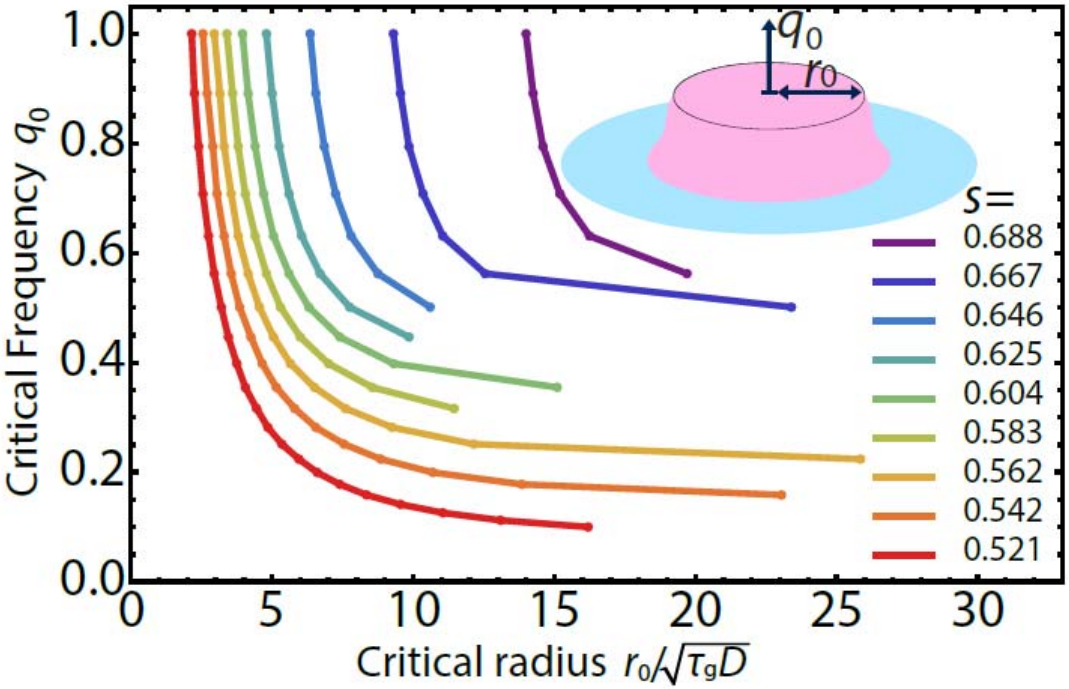
Wave driven from behind the front

(b) Successful nucleation: $a = 0.6, B = 2.0$

Engineering the critical nucleus for pushed excitable waves in two dimensions

$$\tau_g \frac{\partial q}{\partial t} = \tau_g D\left(\frac{\partial^2 q}{\partial r^2} + \frac{1}{r}\frac{\partial q}{\partial r}\right) + \frac{sq(1-q)(q-q^*)}{1-sq(2-q)}.$$

$$q(r, t=0) = \frac{q_0}{1+e^{10(r-r_0)/\sqrt{\tau_g D}}},$$



Double well potential, nucleation theory and Maxwell construction

Recast dynamics in terms of a functional derivative:

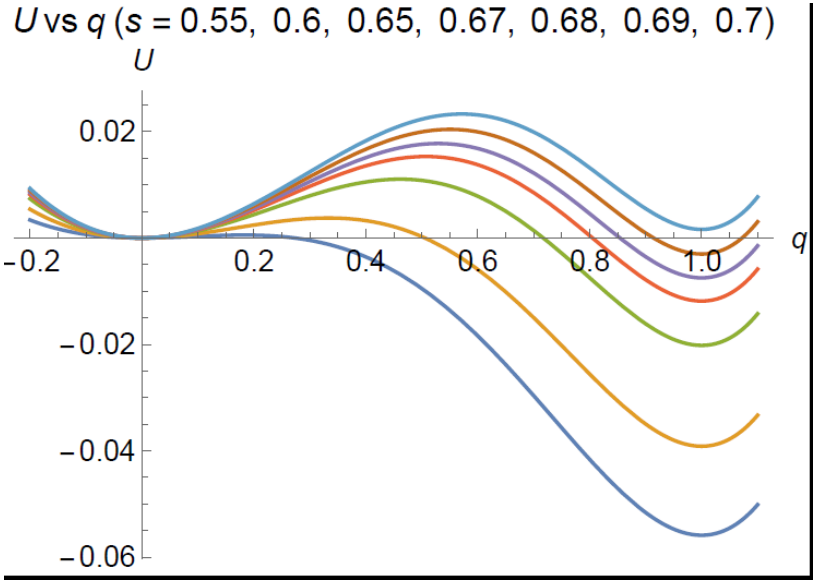
$$\frac{\partial q(x,t)}{\partial t} = -\frac{\delta F}{\delta q(x,t)}, \quad F(t) = \int_{-\infty}^{\infty}\left[\frac{1}{2}\tau_g D\left(\frac{\partial q(y,t)}{\partial y}\right)^2 + U[q(y,t)]\right]dy$$

$$U(q) = -\frac{1}{\tau_g}\int_0^q \frac{sq(1-q)(q-q^*)}{1-sq(2-q)}dq$$

Can show that the dynamics is such that $F(t)$ always decreases:

$$\frac{dF(t)}{dt} = -\int_{-\infty}^{\infty}\left(\frac{\partial q(x,t)}{\partial t}\right)^2 dx < 0$$

[$U(q)$ has a double well structure...]



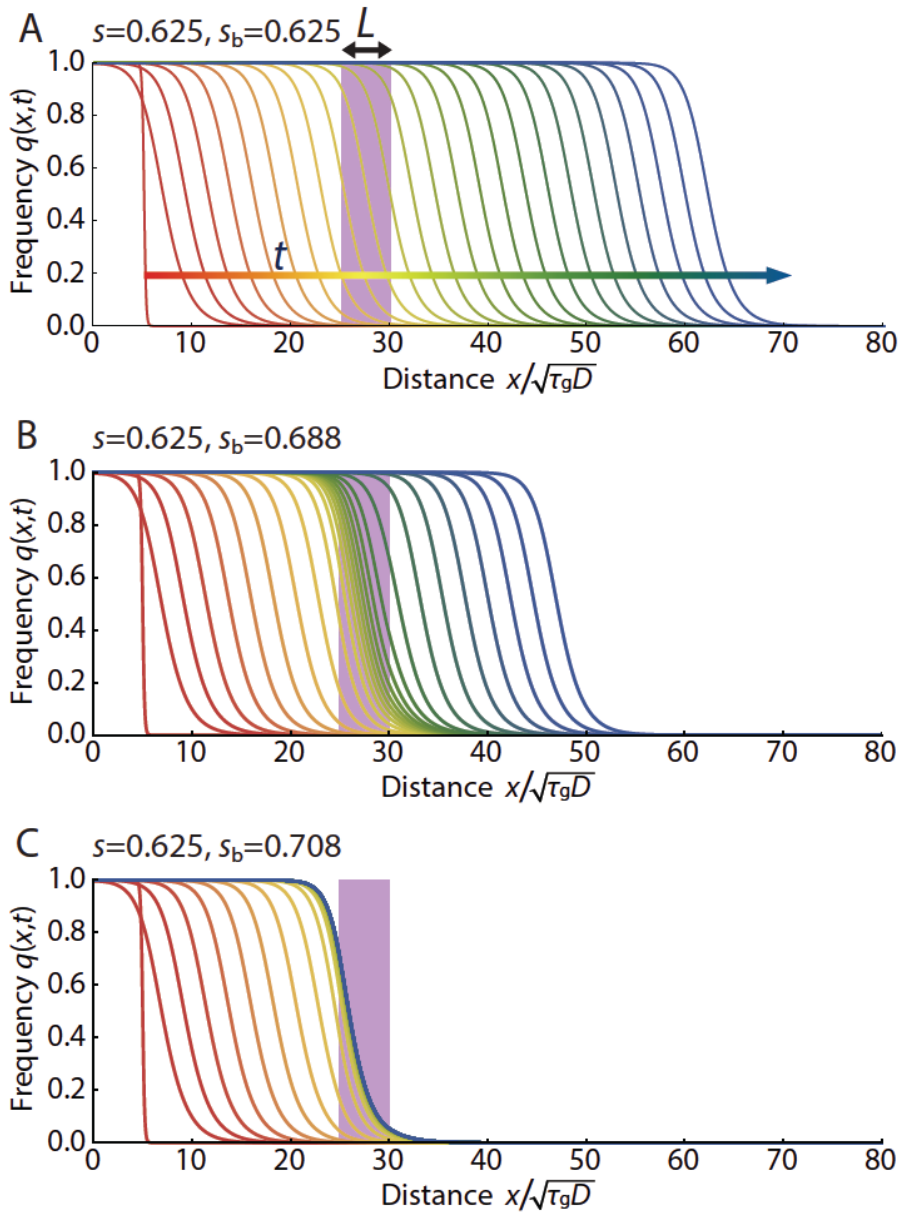
Critical value of $s = s_m$ determined by the condition that the two minima have the same depth, i.e.,

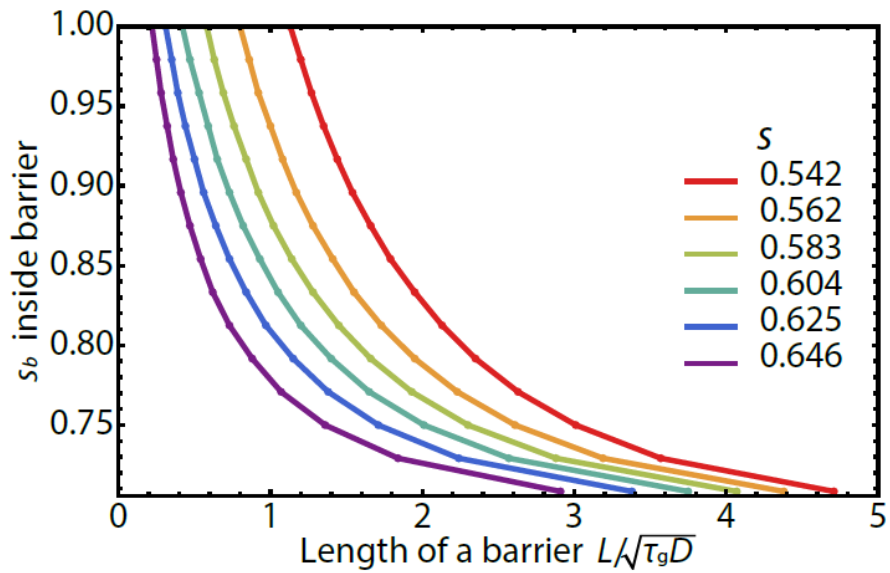
$$U(1) = \int_0^1 \frac{dU(q)}{dq}dq = -\int_0^1 \frac{sq(1-q)(q-q^*)}{1-sq(2-q)}dq = U(0)$$

(Maxwell construction)

$$\Rightarrow \quad s_m = 0.697 > 2/3$$

Stopping power of a selective disadvantage barrier in one dimension





Pushed gene drive wave can be stopped even by a selective disadvantage barrier perforated by gaps

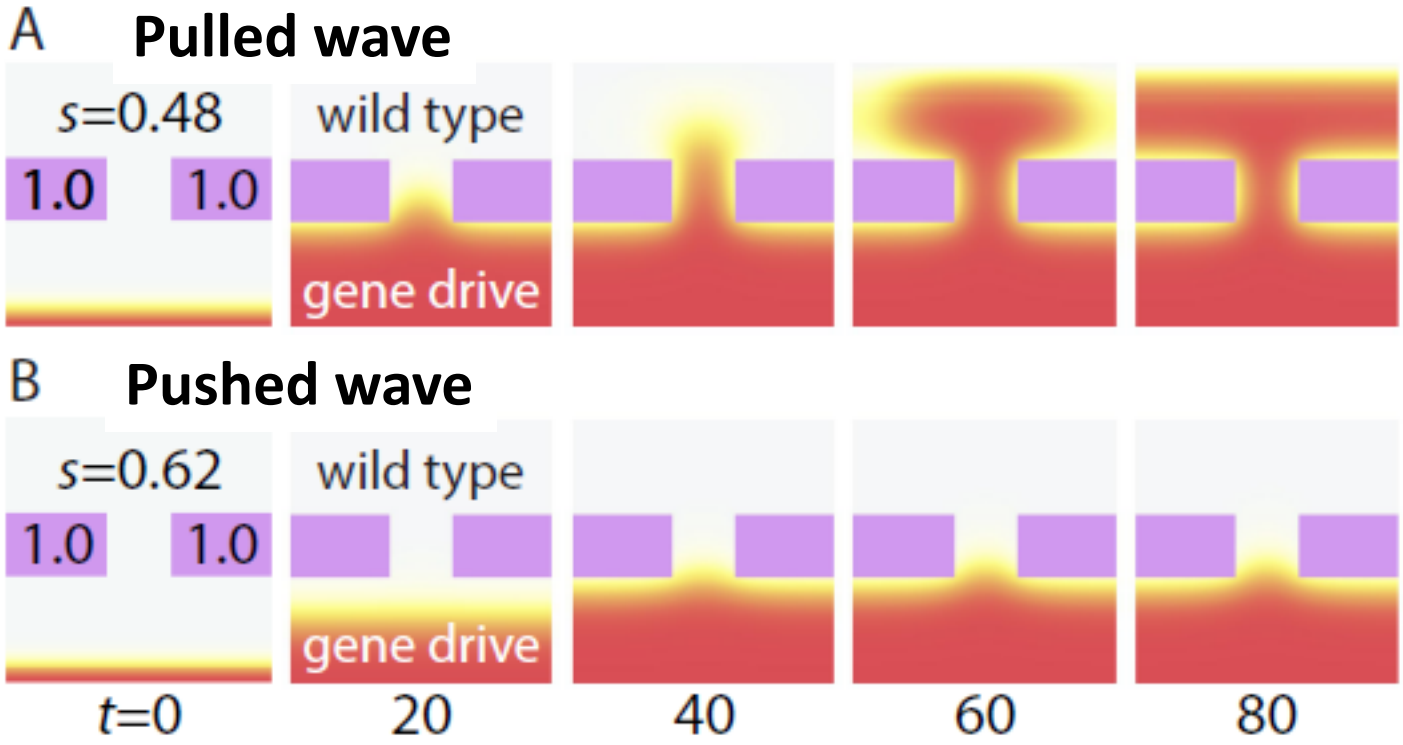
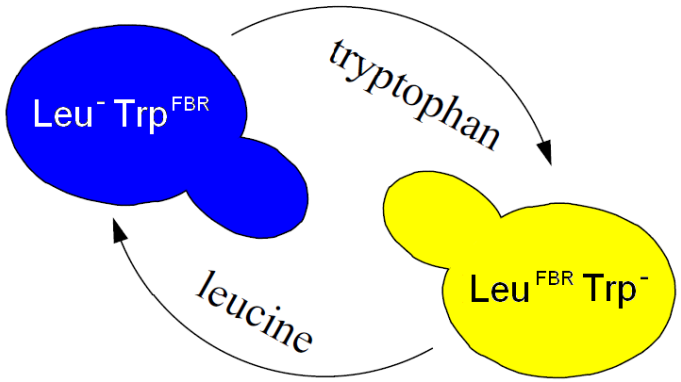
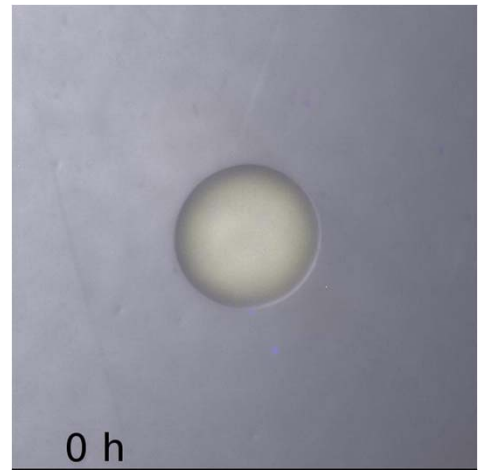


Fig. 7. Population waves impeded by a selective disadvantage barrier of strength $s_b = 1.0$ (colored purple) with a gap. This imperfect barrier has a region without insecticide in the middle of width $6\sqrt{\tau_g D}$. (A) The pulled Fisher wave with $s = 0.48 < 0.5$ always leaks through the gap and reestablishes the gene drive wave (colored red and yellow). (B) The pushed wave that arises when $s = 0.62 > 0.5$ is deexcited by a gapped barrier, provided the gap width is comparable to or smaller than the width of the gene drive wave.

Pulled waves, pushed waves and inter-species interactions

e.g., cross-feeding yeast populations
(M. Mueller, A. Murray & drn)





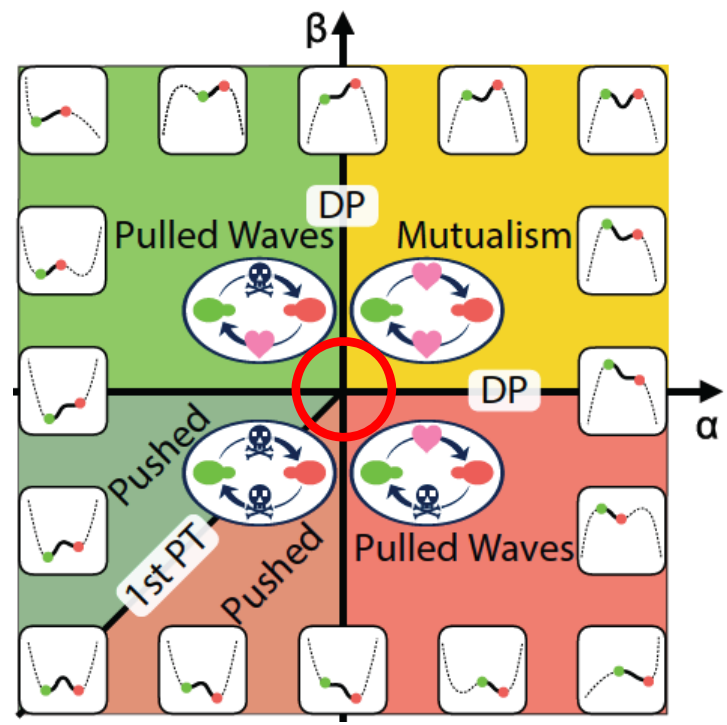
Describe interactions by...

$w_{blue}(x,t) = g + \beta(1 - f(x,t))$, f=blue fraction

$w_{yellow}(x,t) = g + \alpha f(x,t)$, $|\alpha|, |\beta| << g$

$$\frac{\partial f(x,t)}{\partial t} = D\frac{\partial^2 f(x,t)}{\partial x^2} + s_0 f(1-f)(f^* - f) + \sqrt{f(1-f)/2N}\,\Gamma(x,t)$$

$$s_0 = (\alpha+\beta)/g, \quad f^* = \alpha/(\alpha+\beta)$$



neutral model is a set of measure zero!

Thank You!
1000 µm