Physical and evolutionary constraints *at the population scale*

Anne-Florence Bitbol

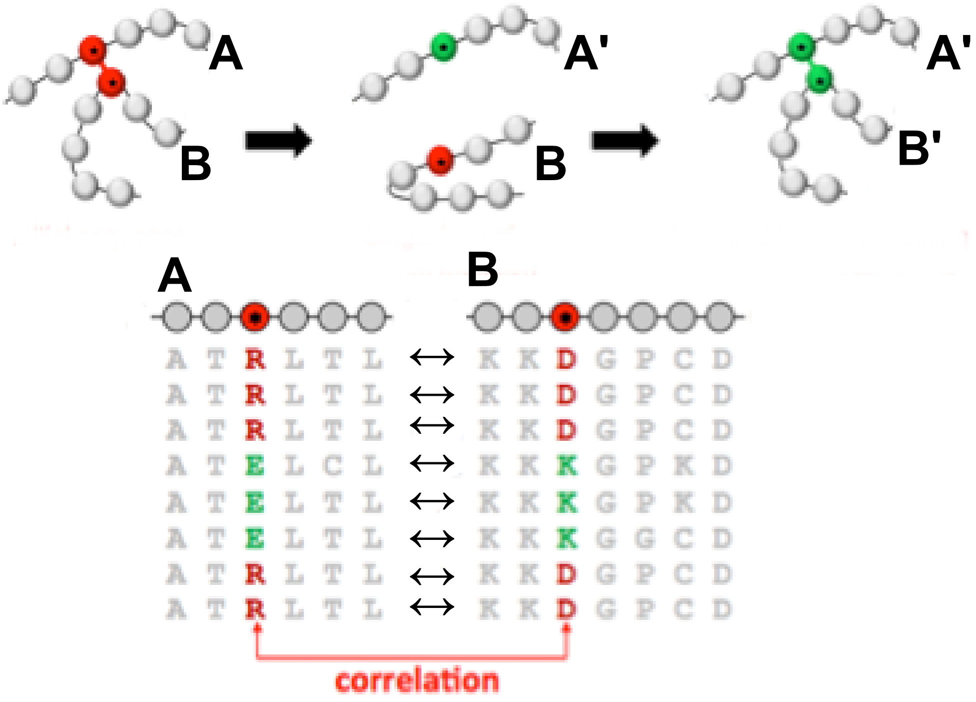
LABORATOIRE Jean Perrin

SORBONNE UNIVERSITÉ

ICTS Program on “Living Matter” - Bangalore, India
April 2018

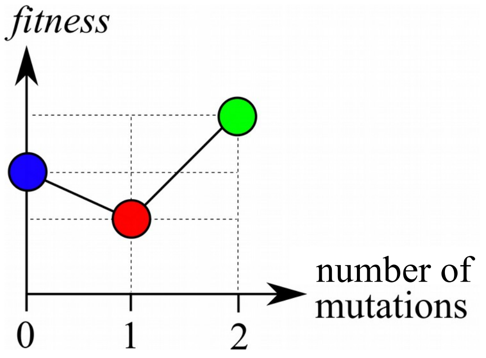
Introduction

▪ Evolutionary coupling and rugged fitness landscapes



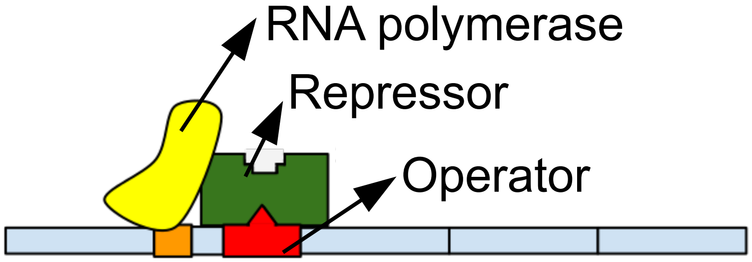
Cf. previous talk (molecular scale)

Typical associated fitness landscape:

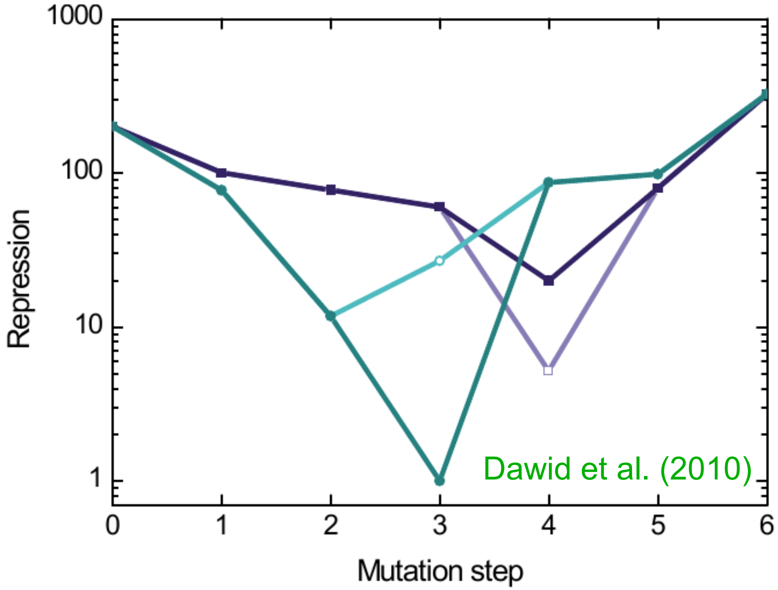


Co-evolving systems → fitness valleys

The *lac* operon:



Wikimedia

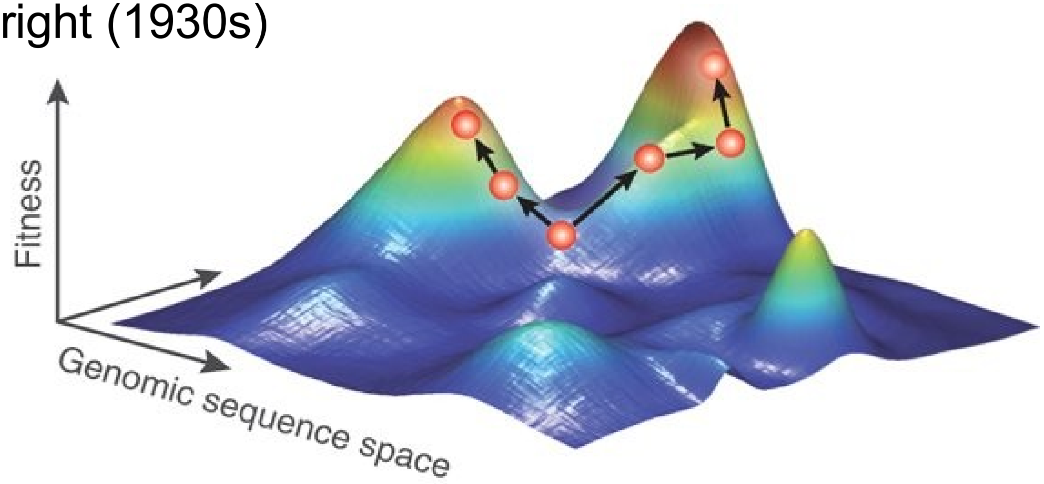


Dawid et al. (2010)

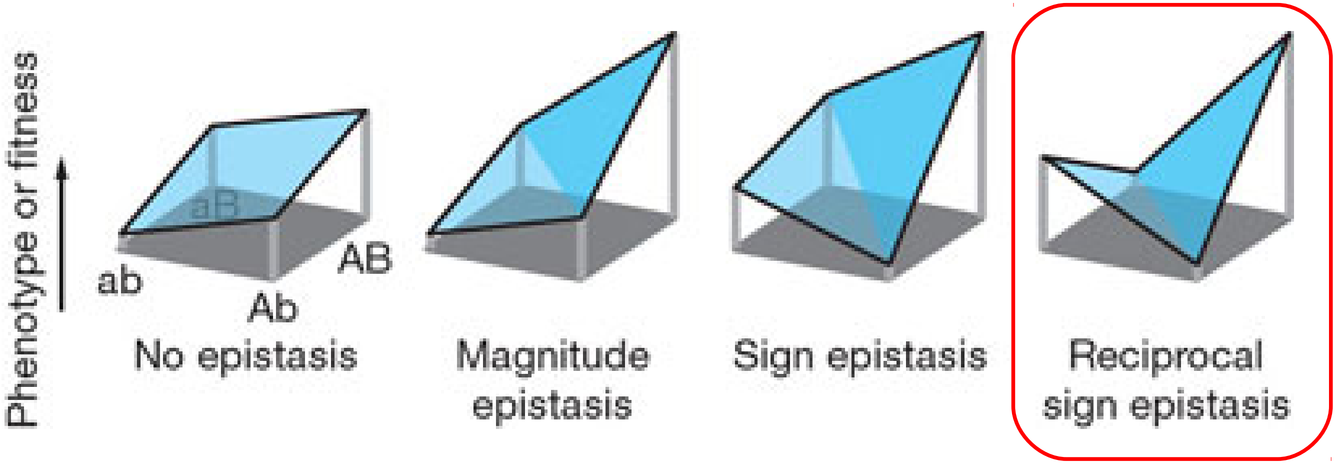
Introduction

- **Fitness landscape**

Wright (1930s)



- **Origin of fitness valleys: epistasis**



Can give rise to multiple peaks

Poelwijk, Kiviet, Weinreich and Tans (2007)

Outline

I. Population subdivision and evolution on rugged fitness landscapes

II. Evolution of antimicrobial resistance

Population subdivision and evolution on rugged fitness landscapes

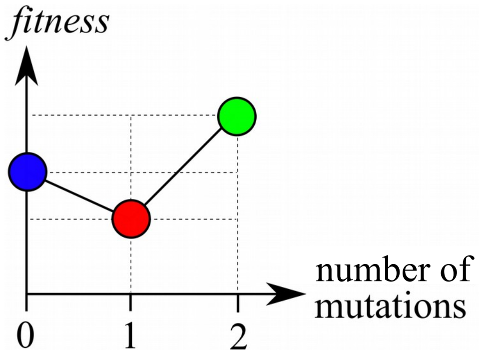
Anne-Florence Bitbol and David J. Schwab

Theoretical Biophysics Group
PIs: Ned Wingreen, William Bialek and Curtis Callan

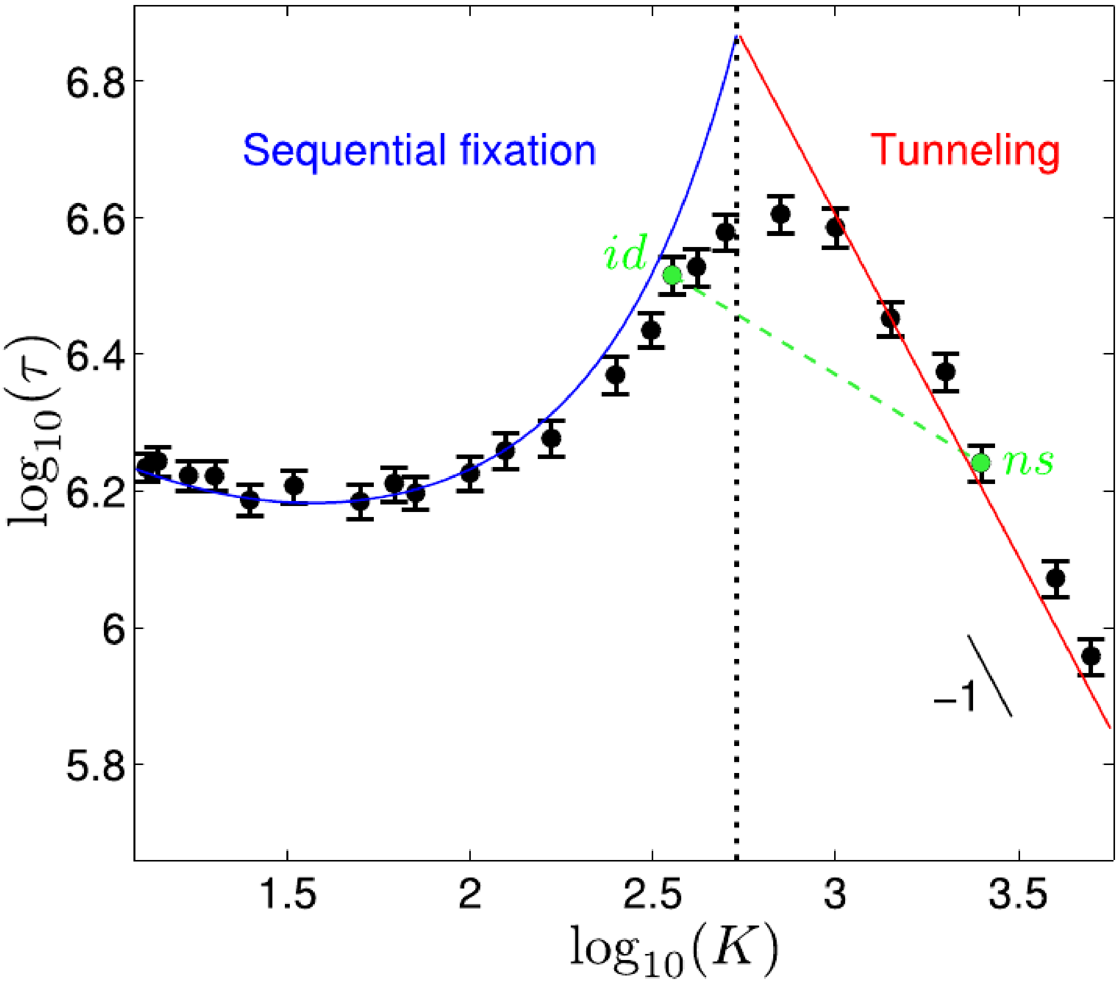
Introduction

- **Effect of population size on fitness valley crossing**

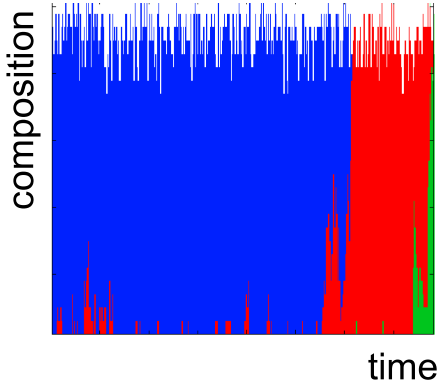
Smaller population → stochasticity is more important
Deleterious / neutral mutations can drift to fixation



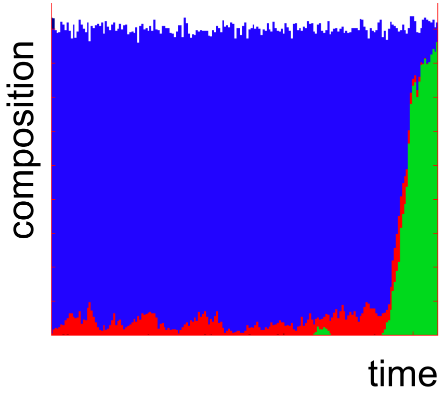
- **Valley crossing time vs. population size: two regimes**



Weinreich and Chao (2005)
Weissman, Desai, Fisher and Feldman (2009)



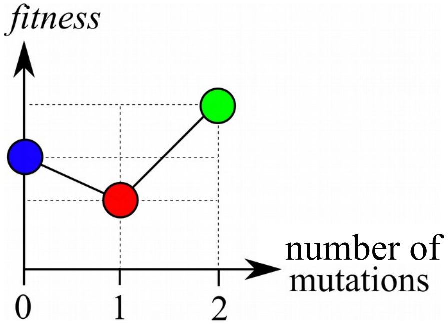
Sequential fixation



Tunneling

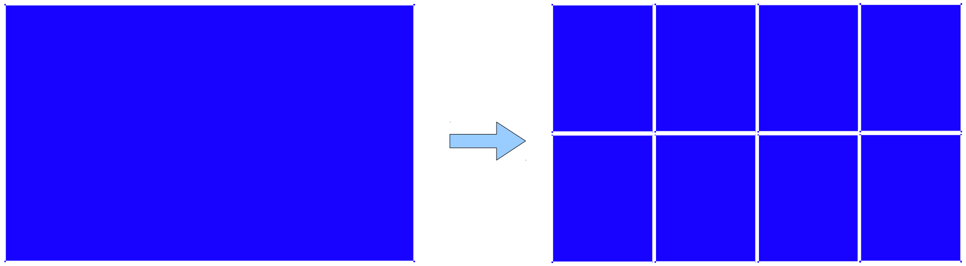
Model & Question

- **Population subdivision: a minimal model**

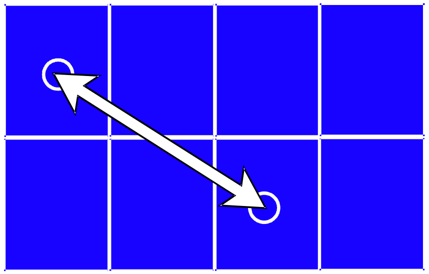


Asexual population
Fixed size

Demes with
identical size



Migration



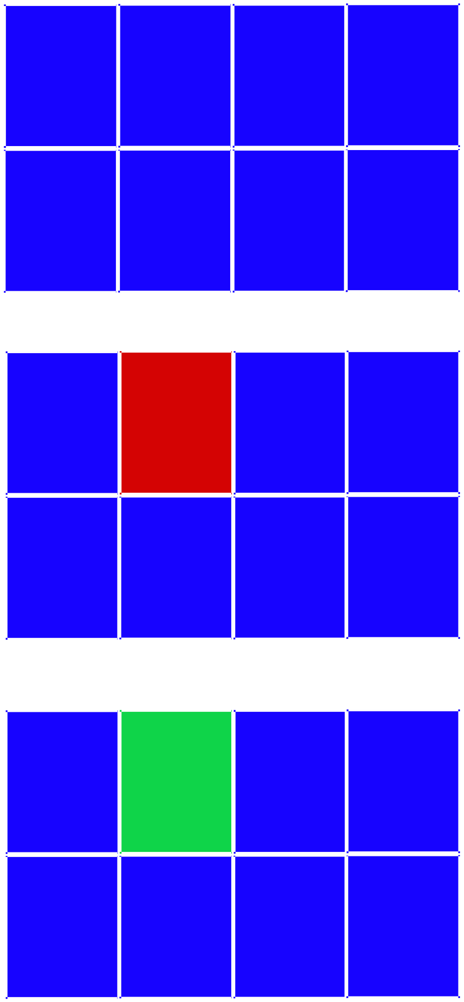
→ Can subdivision with migration (*alone*) accelerate fitness valley crossing?

N.B.: Wright's shifting balance theory (1930s)

Here: Minimal geographic structure
No extinction / founding
No environment heterogeneity
Constant migration rate
Single fitness valley

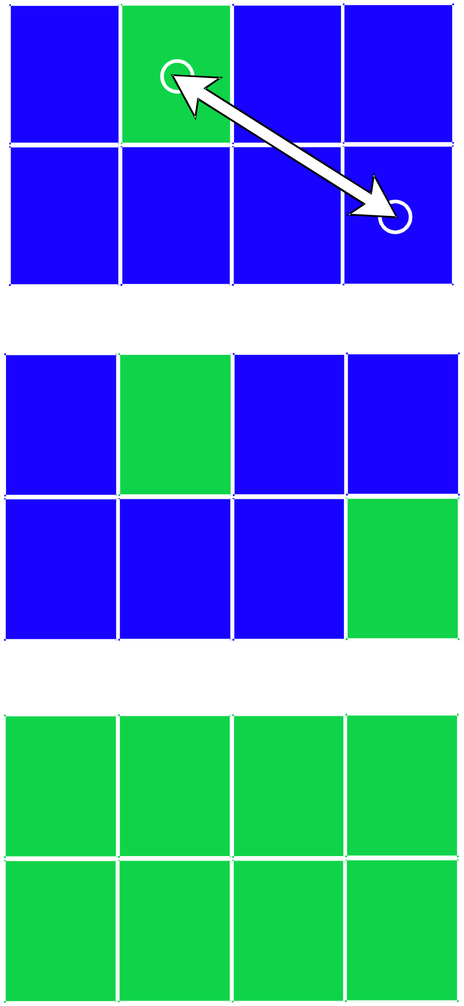
Best scenario

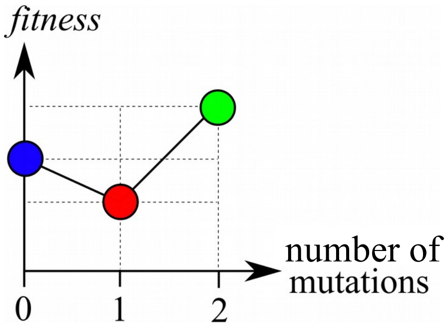
1. Valley crossing by the champion deme



← if demes are in the sequential fixation regime

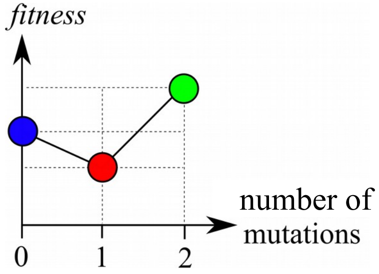
2. Spreading by migration





At best: valley crossing time dominated by that of the *champion* (fastest) deme
→ Speedup in this best scenario?
→ Conditions?

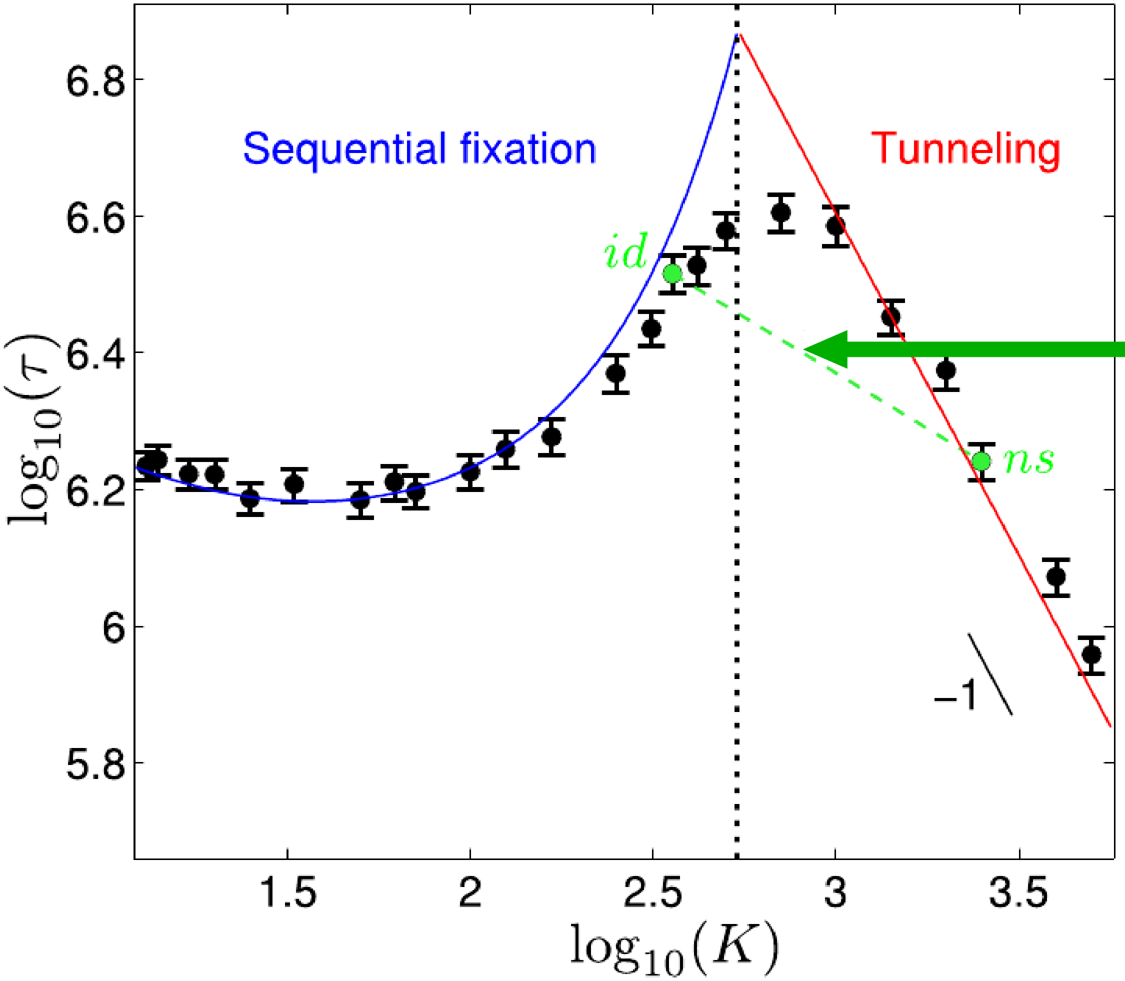
Best scenario



- **Necessary conditions to obtain speedups**

Best scenario → $\tau_m \approx \tau_c$ with $\frac{\tau_c}{\tau_{id}} \approx \frac{1}{D}$
(*m*: metapopulation)

Hence, to have a speedup by subdivision ($\tau_m < \tau_{ns}$), we need $\frac{\tau_{id}}{\tau_{ns}} < D$



Slope needs to be larger (less negative) than -1

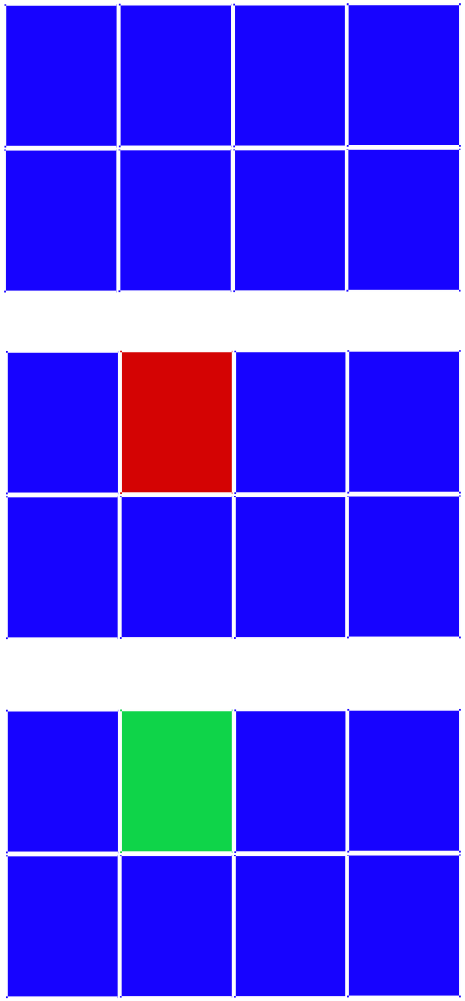
Consequence: Sequential fixation in individual demes is necessary in order to get speedups

Reciprocally: Demes in the sequential fixation regime → speedups *in the best scenario*

→ **Conditions under which the best scenario is attained?**

Conditions

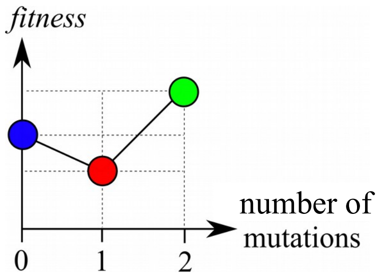
- **Condition 1**

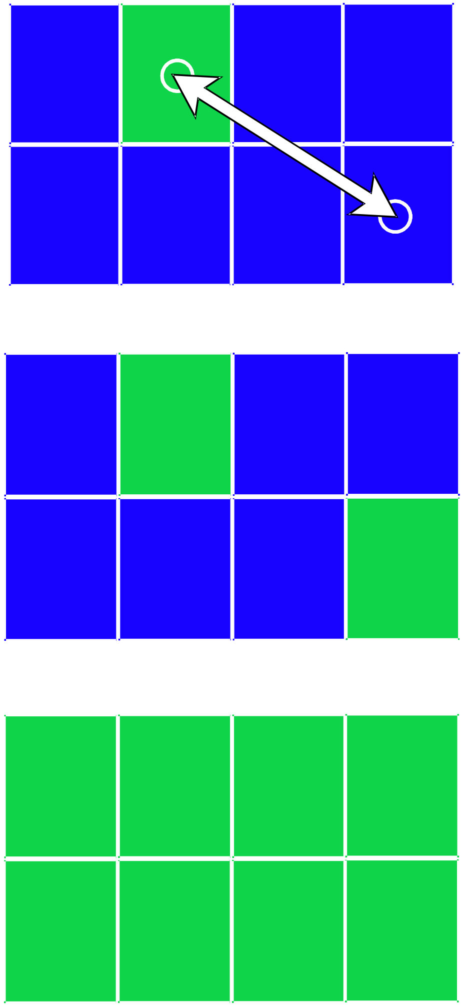


The champion deme must be shielded from migration while in the deleterious state

→ **Upper bound on the migration rate**

- **Condition 2**



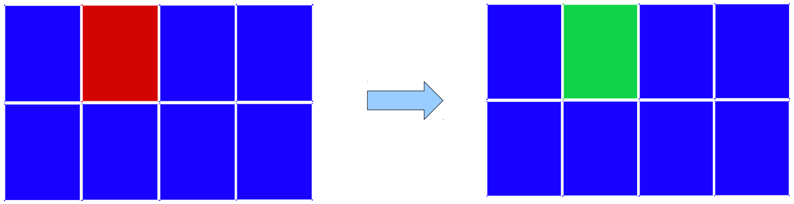


Spreading of the beneficial mutation must be faster than valley crossing by the champion deme

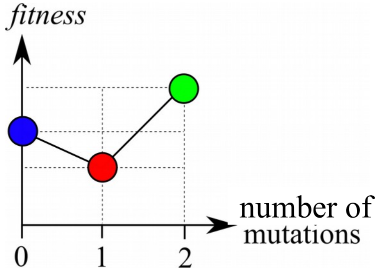
→ **Lower bound on the migration rate**

Expression of the conditions

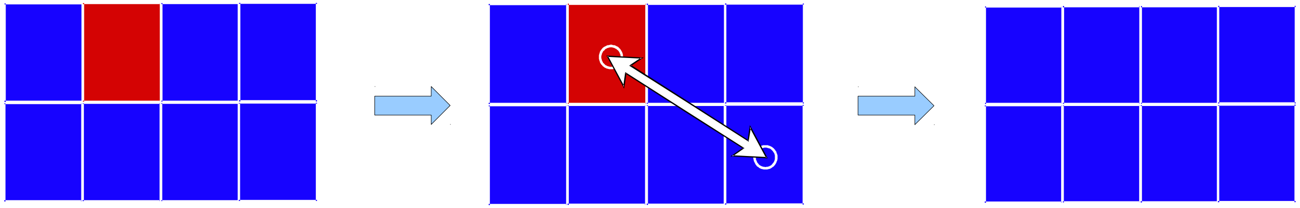
Condition 1 – quasi-independence



Timescale: $\tau_{12} = \dfrac{1}{N\mu d\, p_{12}}$



must occur faster than



Timescale: $t_e = \dfrac{n_e}{DNm}$

n_e = average number of migrations for extinction of '1'

→**Condition 1:** $\tau_{12} < t_e \;\rightarrow\; \dfrac{m}{\mu d} < \dfrac{n_e p_{12}}{D}$: upper bound

Condition 2 – fast spreading

Similarly: $t_s < \tau_c \;\rightarrow\; \dfrac{n_s}{DNm} < \dfrac{1}{DN\mu d\, p_{01}} \;\rightarrow\; p_{01} n_s < \dfrac{m}{\mu d}$: lower bound

Expressions for n_s and n_e: obtained using finite Markov chain theory

(+ intuitive calculation for $\delta \ll 1,\; s \ll 1,\; N\delta \gg 1,\; Ns \gg 1$)

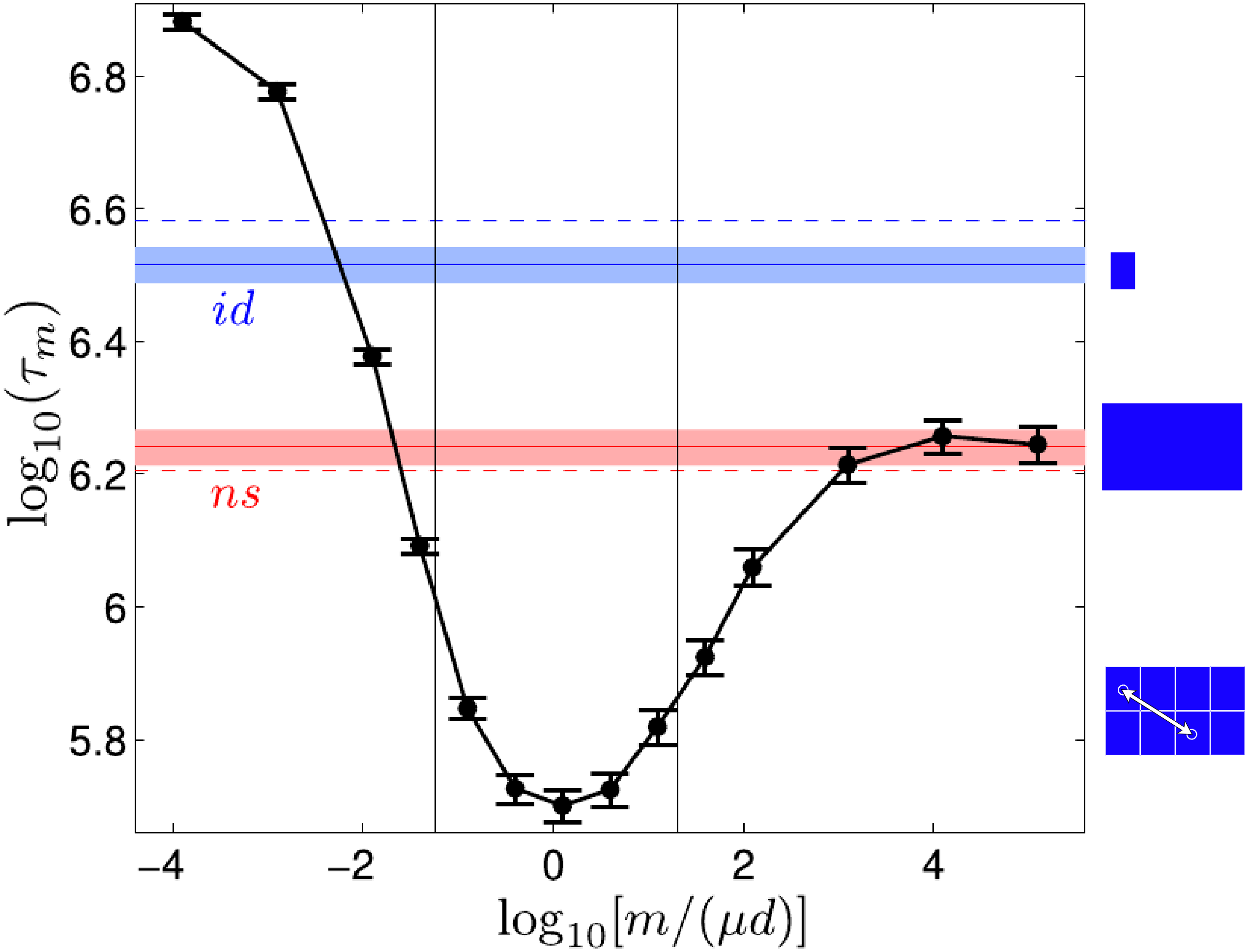
Prediction:

$$\boxed{n_s p_{01} \ll \frac{m}{\mu d} \ll \frac{n_e p_{12}}{D}}$$

→optimal scenario, and $\dfrac{\tau_m}{\tau_{id}} \approx \dfrac{1}{D}$

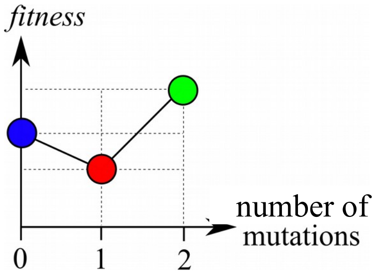
Test: stochastic simulation

- **Simulation (Gillespie algorithm) → crossing time vs. migration rate**



Parameter values:

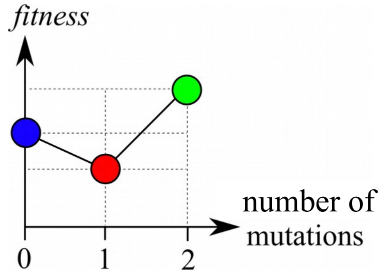
s = 0.3
δ = 0.006
K = 357
D = 7
μ = 8 x 10^{-6}
d = 0.1



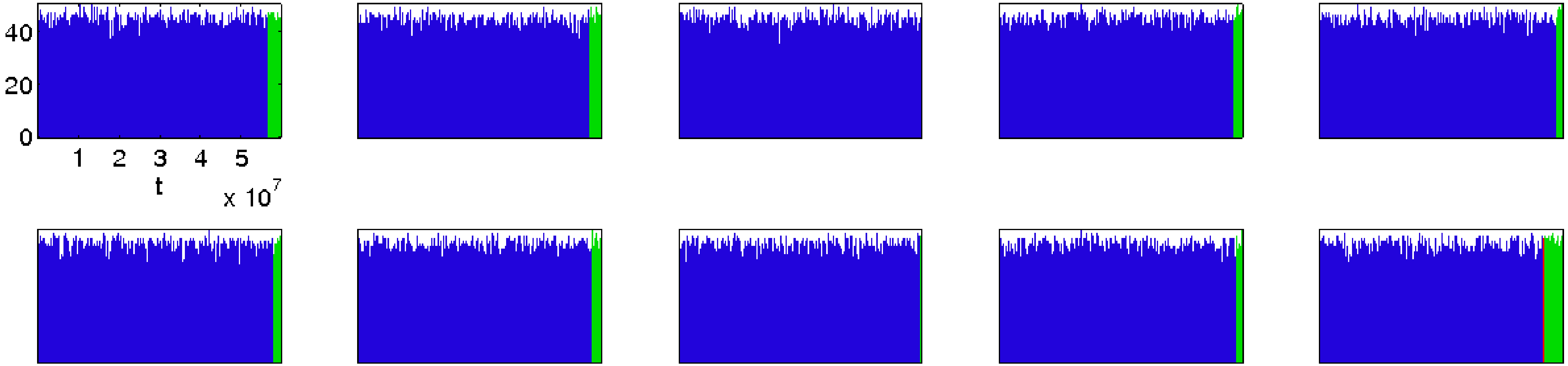
Minimum → $$\left.\begin{aligned}\tau_m &= (5.02 \pm 0.14) \times 10^5 \\ \tau_{id} &= (3.28 \pm 0.10) \times 10^6\end{aligned}\right\}$$ → factor of 6.54, close to *D* = 7

Test: stochastic simulation

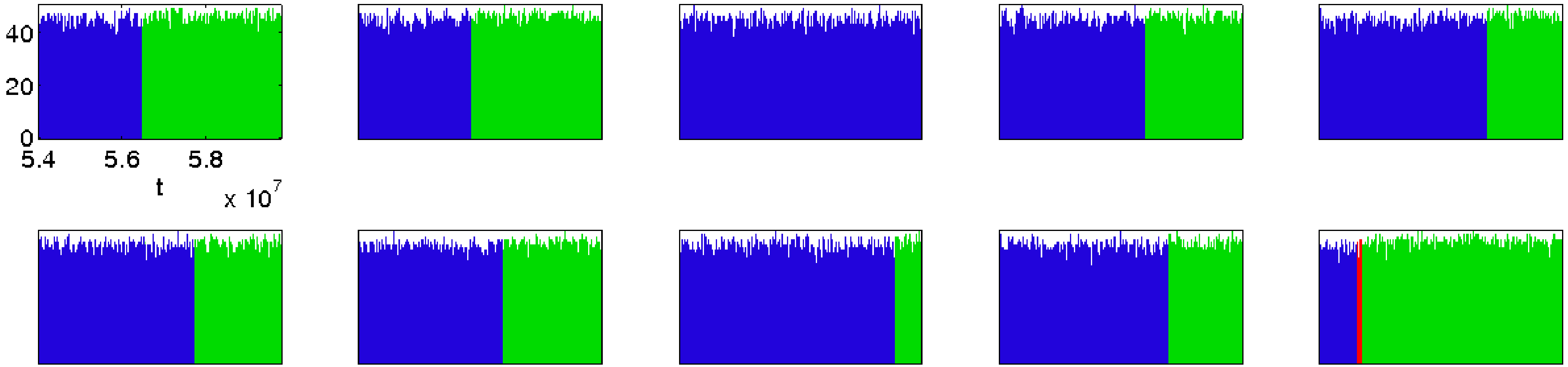
- **Valley crossing at the optimum**



One realization:

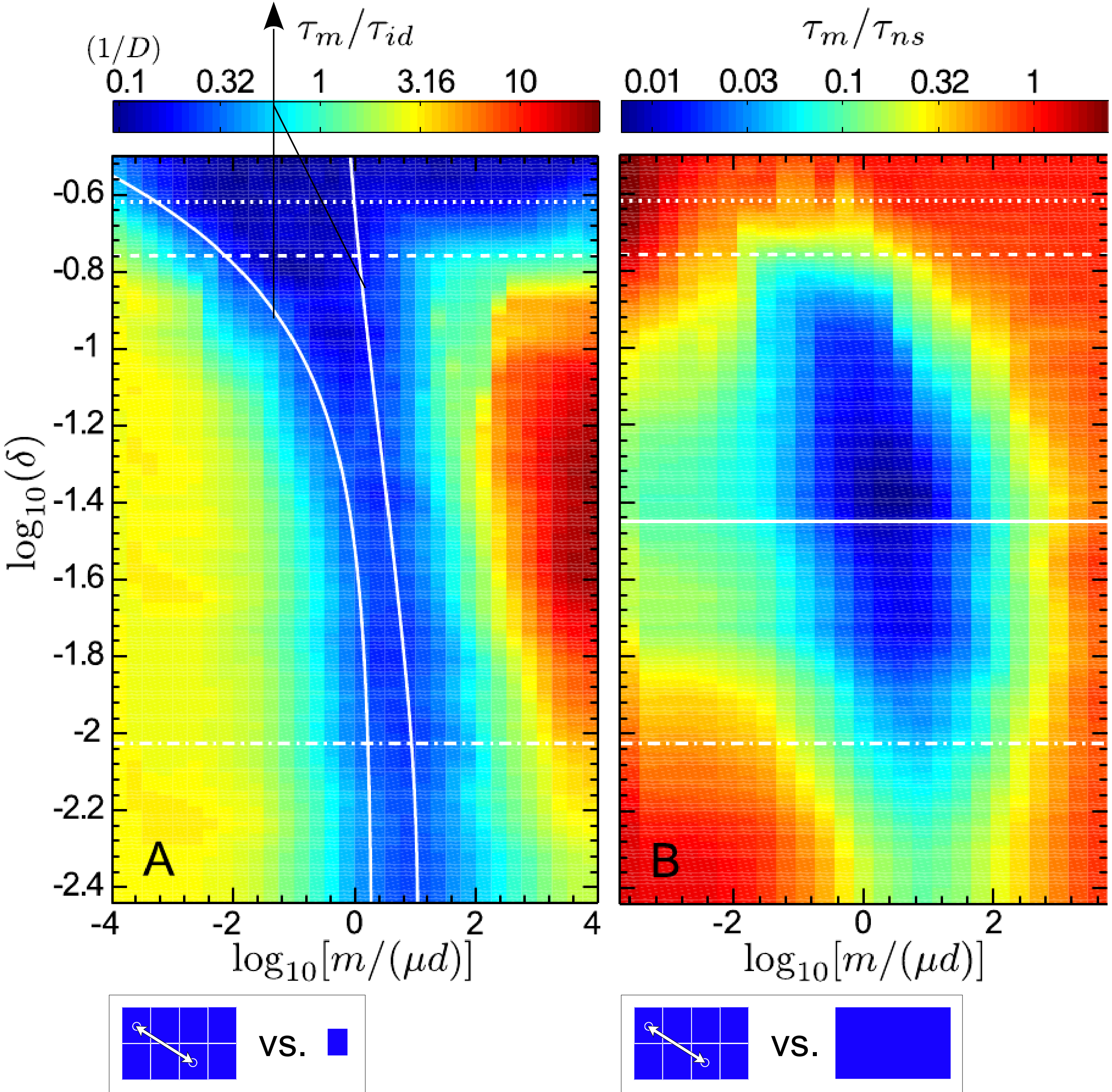


End of the process:



Heatmaps

Parameters:
s = 0.3
K = 50
D = 10
μ = 5 10^{-6}
d = 0.1



Predicted bounds for optimal region

Tunneling starts dominating for isolated demes

Valley depth predicted to yield the highest speedup by subdivision

Tunneling starts dominating for the nonsubdivided population

Highest speedup & trade-off

Highest possible speedup by subdivision

Optimal case → speedup gained by subdividing a population: $\frac{\tau_m}{\tau_{ns}} = \frac{\tau_c}{\tau_{ns}}$

Assume:
- isolated deme in the sequential fixation regime
- nonsubdivided population in the tunneling regime

$$2\sqrt{\mu s} \ll \delta \ll 1 \quad \rightarrow \quad \frac{\tau_m}{\tau_{ns}} = \mu s \frac{e^{N\delta} - 1}{\delta^2}$$

At fixed *N*, this ratio is minimal for $\delta \approx \frac{1.594}{N}$ (→ importance of general calculations)

Its minimal value is $\frac{\tau_m}{\tau_{ns}} \approx 1.544\, N^2 \mu s$

Heatmaps → optimal valley depth: $\delta \approx 0.035 \approx 10^{-1.45}$

A trade-off in the choice of *D*

Fixed $\mathcal{N} = ND$ → highest speedup: $\frac{\tau_m}{\tau_{ns}} \approx 1.544 \frac{\mathcal{N}^2 \mu s}{D^2}$

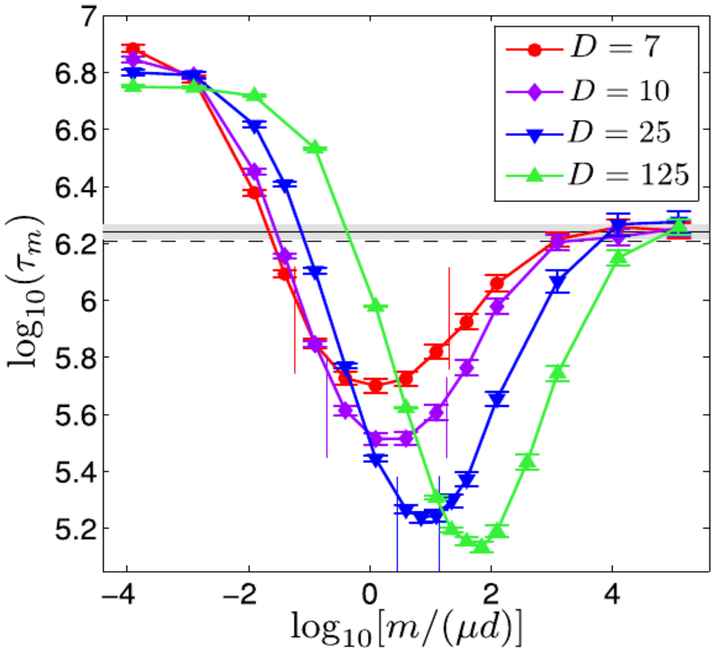
Increase *D* → gain more speedup

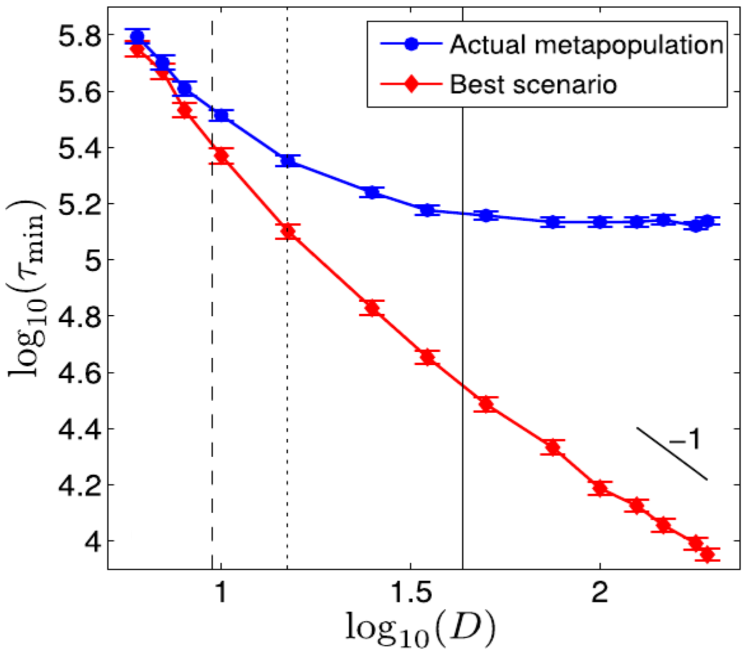
But $\frac{\delta e^{-N\delta}}{s} D \log D \ll \frac{m}{\mu d} \ll \frac{1}{2}\left(1 + \frac{s}{\delta}\right)$

Increase *D* → narrower optimal parameter range

Generalizing

Varying the degree of subdivision

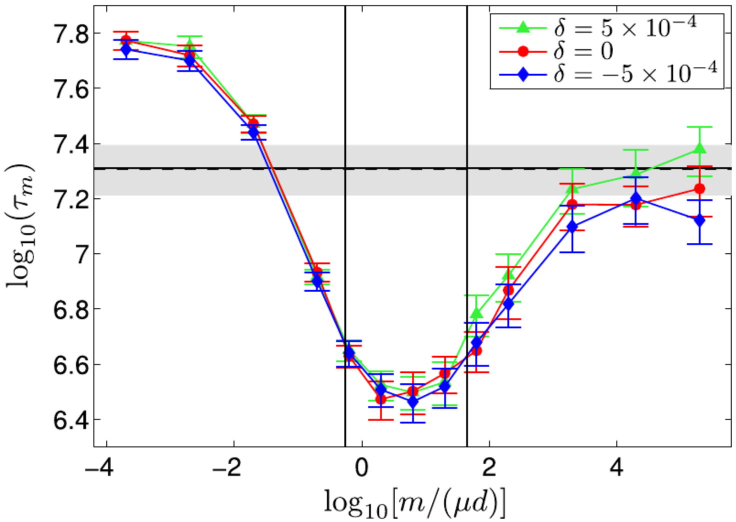




Parameter values:

s = 0.3
δ = 0.006
D K = 2500
μ = 8 x 10^{-6}
d = 0.1

Effectively neutral intermediates



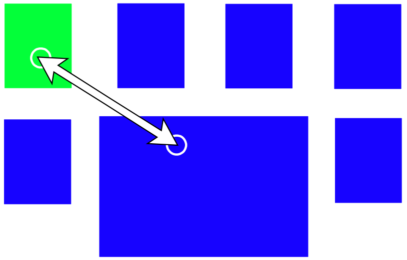
$$|\delta| < \max(\sqrt{\mu s}, 1/N)$$

→ effectively neutral

Parameter values:

s = 0.5
N = 130
D = 10
μ = 5 x 10^{-7}
d = 0.1

Mainland-island(s)



Champion island can drive valley crossing

Application

- **An example**

E. coli → $\mu \approx 8.9\times 10^{-11}$ Wielgoss et al. (2011)

Take $N = 5 \times 10^4$ (small but realistic) Rozen et al. (2008)
$D = 100$ (96-well plates)

Plateau → sequential fixation below $N_\times = 1/\sqrt{\mu s}$

$s = 10^{-2}$ → isolated demes in the sequential fixation regime
for $0 \leq \delta \lesssim 2.2 \times 10^{-4}$

The optimal range of migration rates spans 2 to 4 orders of magnitude depending on δ

Speedup factor from 18 to 2.7×10^2

- **More generally**

For given *N* and *D*, we can predict:
- for which valleys subdivision speeds up crossing
- the highest speedups obtained
- the range of migration rates for which they are reached

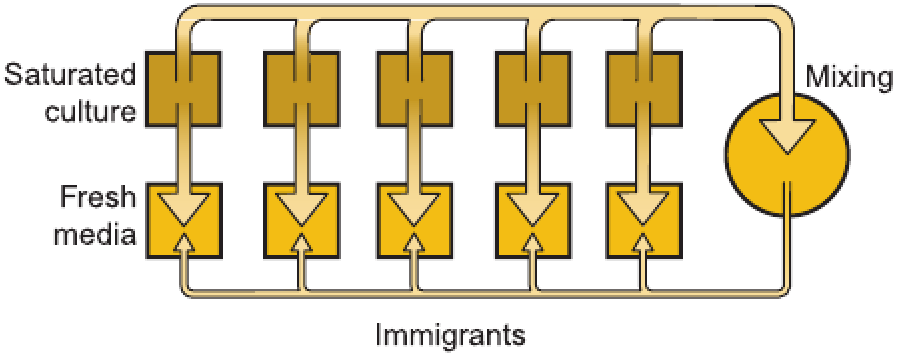
Conclusion

Summary

- Subdivision with migration (alone) can significantly accelerate fitness valley & plateau crossing
- Sufficiently small demes (performing sequential fixation) are necessary

Some related experimental studies

- Kryazhimskiy, Rice & Desai (Evolution 2012) → evolution of subdivided populations of yeast



→ no evidence of any advantage of subdivided populations

- Nahum, Godfrey-Smith, Harding, Marcus, CarlsonStevermer & Kerr (PNAS 2015)
  → evolution of subdivided populations of bacteria
  → some advantage of subdivision

→ Importance of understanding quantitatively the conditions under which subdivision is beneficial

Conclusion

Perspectives

- More complex population structures
- Case of sexual populations (recombination)
- Spatial structure (expanding front)
- Effect of population subdivision on the evolution of antibiotic resistance

Acknowledgements

David J. Schwab

Ned S. Wingreen

The Princeton Biophysics Journal Club

Reference

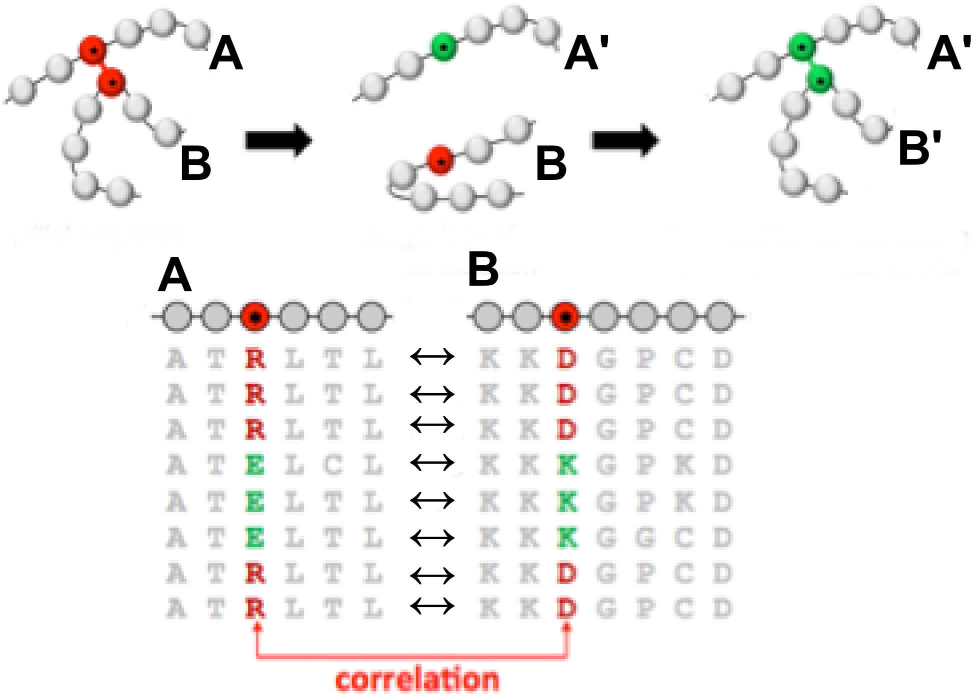
A.-F. Bitbol and D.J. Schwab, *Quantifying the role of population subdivision in evolution on rugged fitness landscapes*, PLOS Computational Biology, 10 (8): 003778 (2014)

Evolution of antimicrobial resistance

Loïc Marrec and Anne-Florence Bitbol

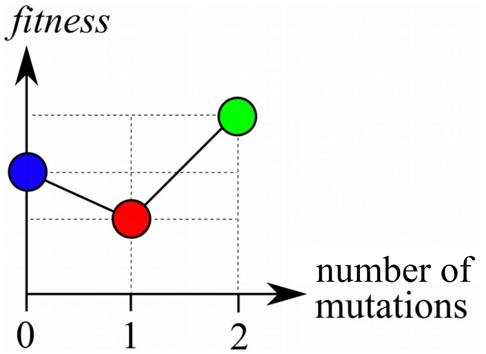
Introduction

Evolutionary coupling and rugged fitness landscapes

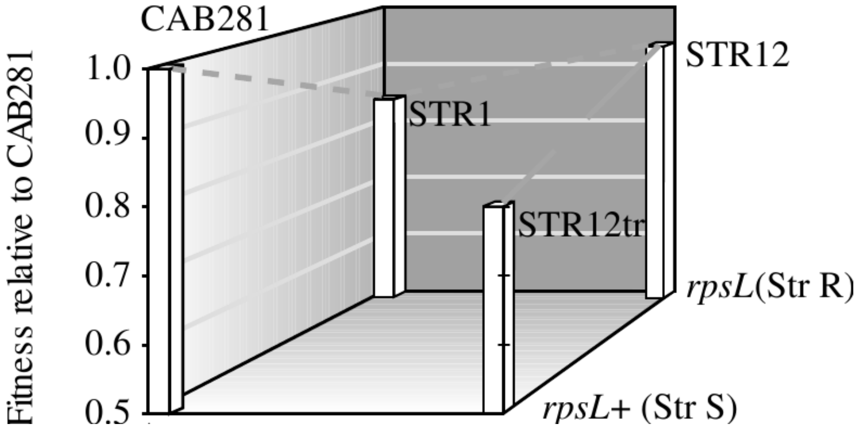


Cf. previous talk (molecular scale)

Typical associated fitness landscape:



Fitness costs in the evolution of antibiotic resistance



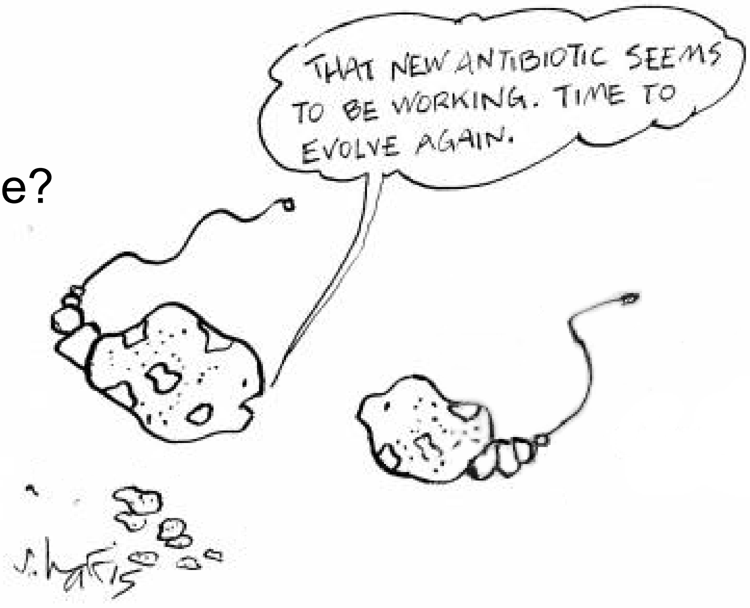
Evolution of streptomycin resistance in *E. coli*

Schrag, Perrot and Levin (1997)

Introduction

A major public health issue

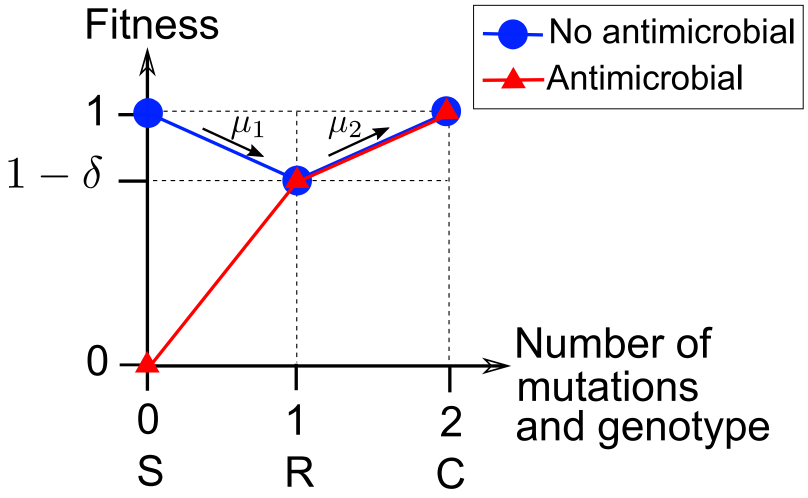
- What conditions favor / challenge the evolution of resistance?
- How to optimize antimicrobial dosage?
- Impact of antimicrobials in the environment?

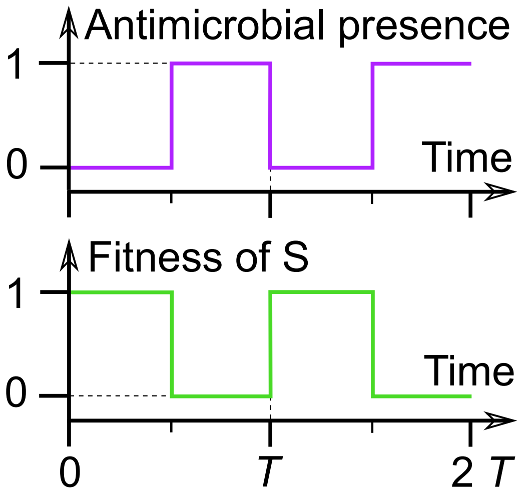


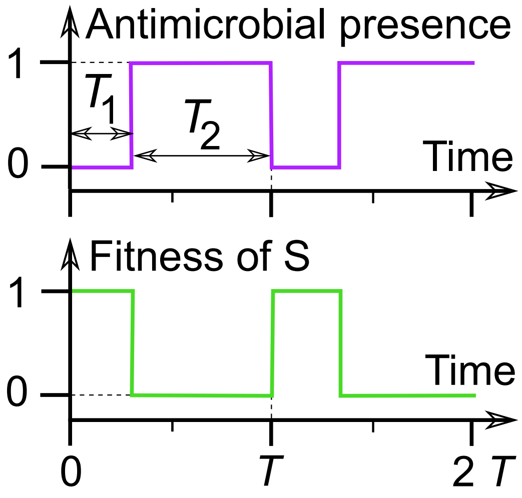
A variable fitness landscape

- Most mutations giving antimicrobial resistance come with a fitness cost
- This cost can be compensated by a second mutation

→ fitness valley ***in the absence of antimicrobial***







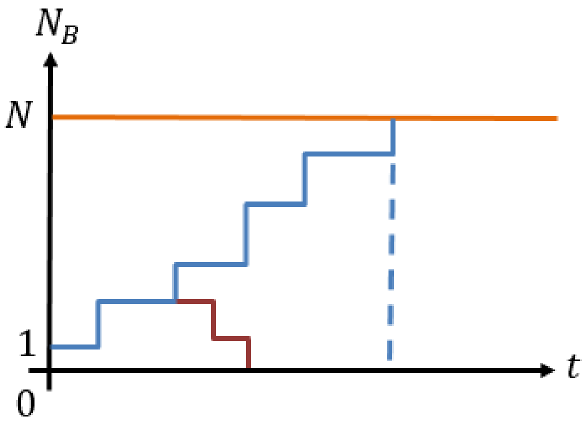
The fitness landscape depends on whether antimicrobial is present or not

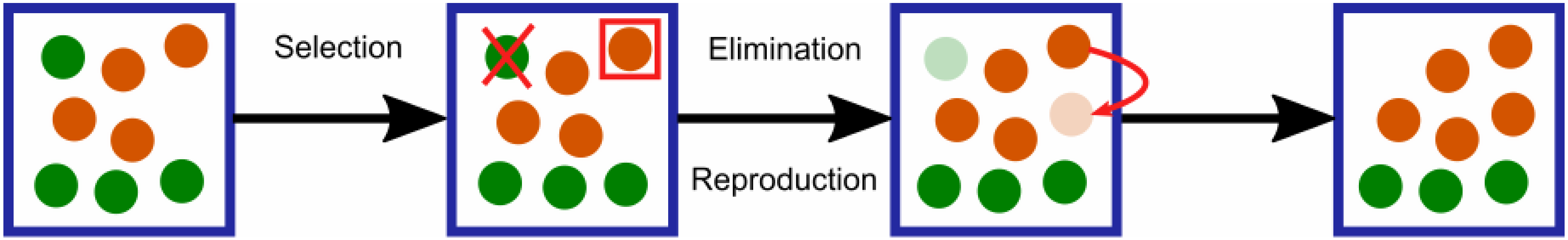
→ **Impact of variable antimicrobial concentrations on the evolution of resistance?**

Stochastic model

Moran (1958)
Ewens (1979)

Moran Process
(constant size)





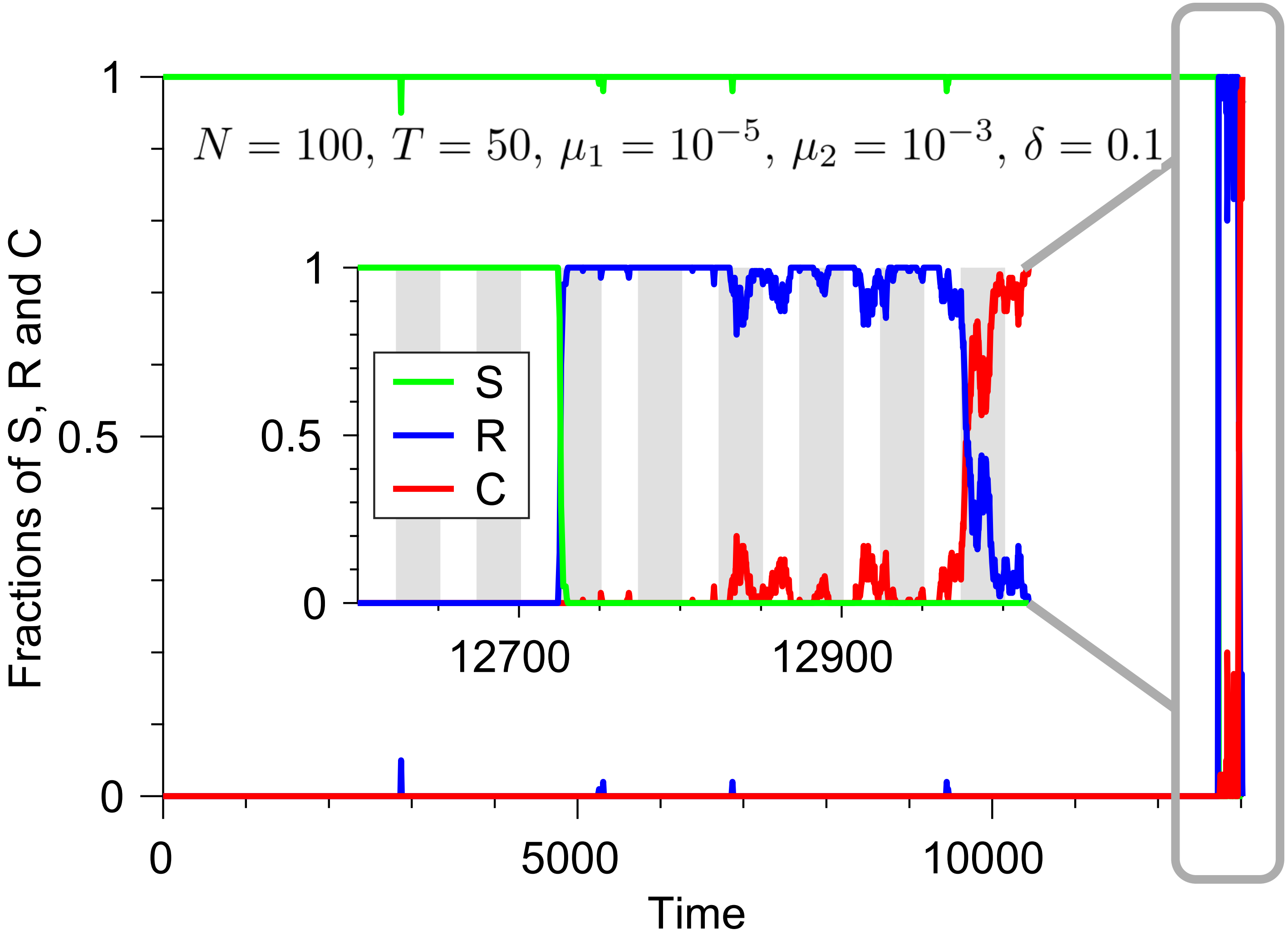
Fixation probabilities; fixation times of each genotype

Markov chain

Absorption probabilities; unconditional and conditional first-passage times (to absorbing states)

Process studied

- **Stochastic model (well-mixed population, fixed size → Moran process)**



Example of a simulation run, starting from a population of S individuals

- Without antimicrobial: valley crossing → 10^7 generations
- Continuous presence of antimicrobial → resistance can't evolve

→ **A periodic antimicrobial presence greatly accelerates the evolution of resistance**

Analytical predictions

Acquisition of resistance

The population can evolve resistance:
- By valley crossing (cf. earlier) Weissman et al (2009)
- Due to “switches” of the environment, i.e. introduction of antimicrobial

What ultimately matters is the fastest process
Focus on the 2nd one (and then compare)

- R mutants can only appear without antimicrobial
- R mutants have a fitness advantage only when antimicrobial is added

→ **Key question: When do R mutants first exist in the presence of antimicrobial?**

t_R^a : average time when R mutants start growing in the presence of antimicrobial

Three key timescales:

- Average time of appearance of a resistant (R) mutant: $1/(N\mu_1)$
- Average time of disappearance of the lineage of a R mutant: τ_R^d
- Period of the alternation of antimicrobial absence and presence: T

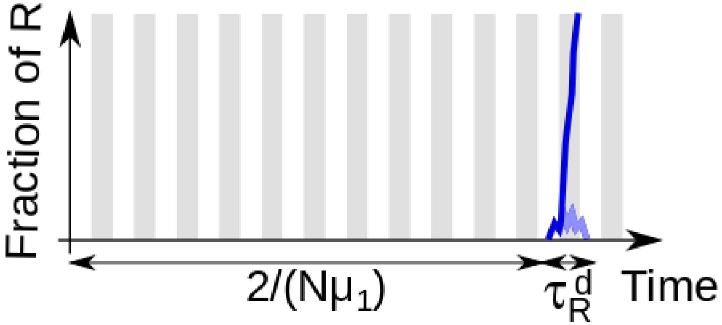
$$\left.\begin{array}{l} 1/(N\mu_1) \\ \tau_R^d \end{array}\right\} \longrightarrow \text{evolution}$$

$$T \longrightarrow \text{environment}$$

Analytical predictions

- **Time until resistants start growing: different regimes**

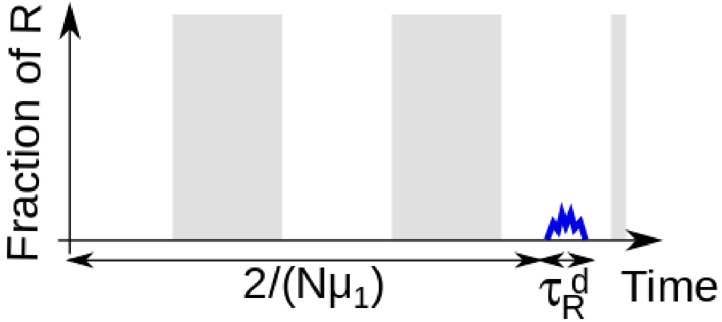
$$\frac{T}{2} \ll \tau_R^d$$



The lineage of the first R mutant lives until antimicrobial is added

$$\rightarrow t_R^a = \frac{2}{N\mu_1}$$

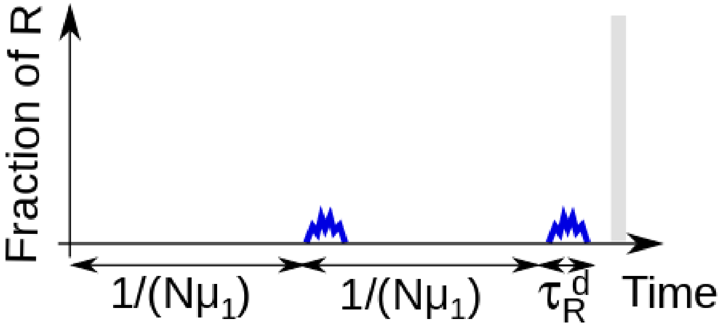
$$\tau_R^d \ll \frac{T}{2} \ll \frac{1}{N\mu_1}$$



Probability that R appeared within τ_R^d before adding antimicrobial:

$$p = \frac{2\tau_R^d}{T} \rightarrow t_R^a = \frac{2}{N\mu_1} \times \frac{1}{p} = \frac{T}{N\mu_1\tau_R^d}$$

$$\frac{T}{2} \gg \frac{1}{N\mu_1}$$



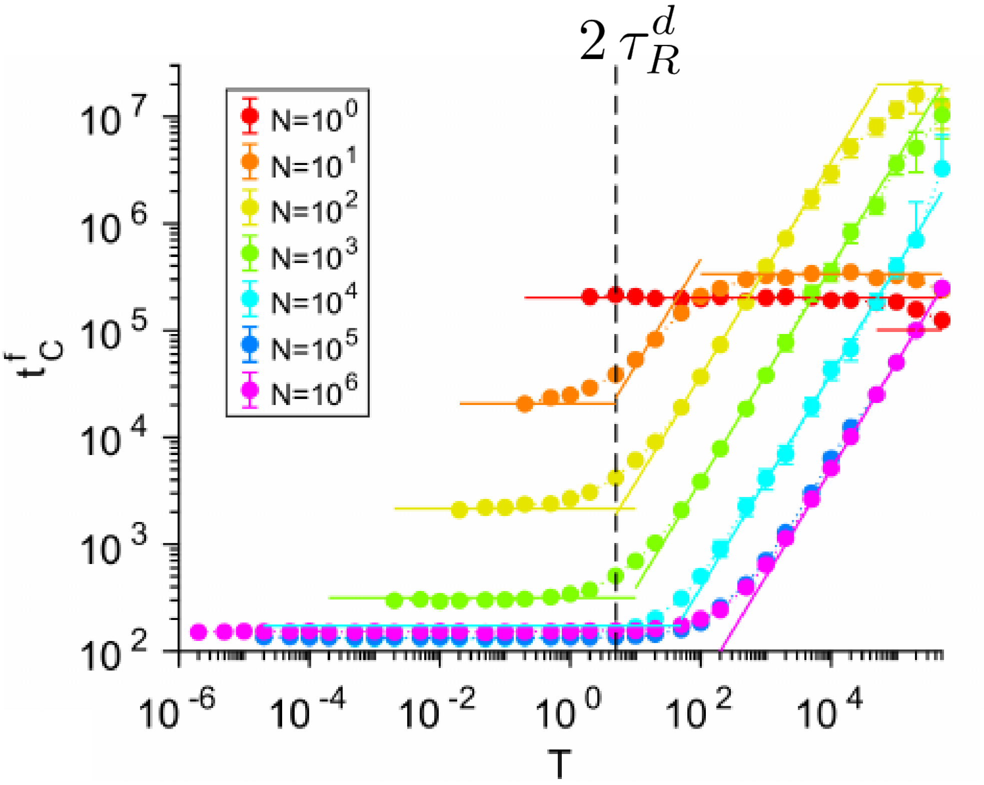
Probability that a lineage of R exists upon a given addition of antimicrobial:

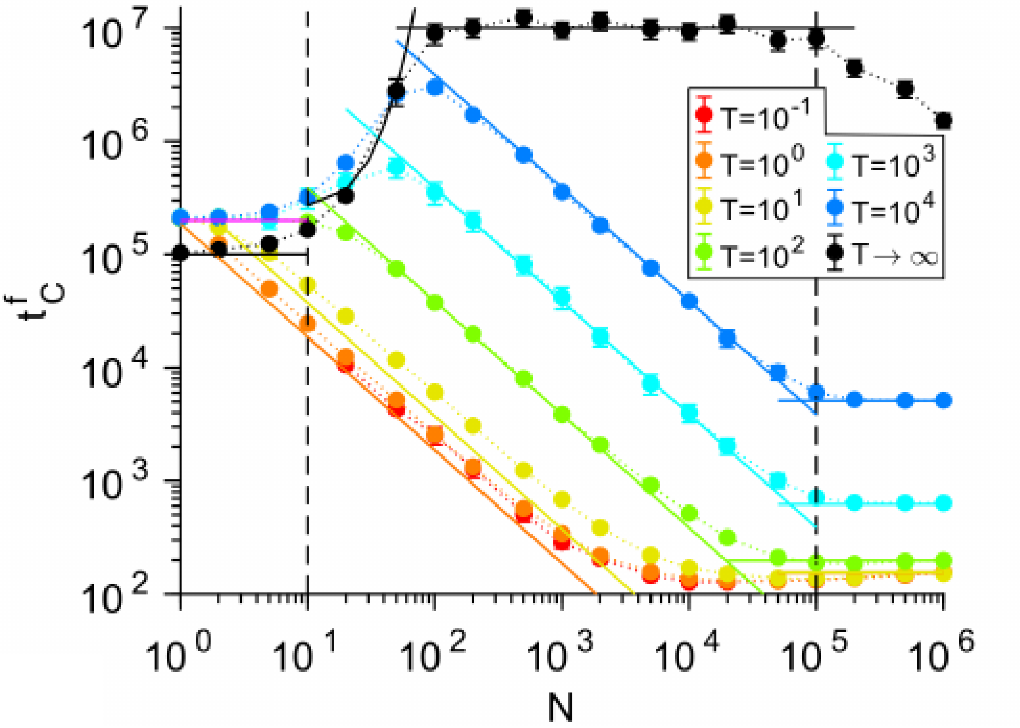
$$p' = N\mu_1\tau_R^d \rightarrow t_R^a = T \times \frac{1}{p'} = \frac{T}{N\mu_1\tau_R^d}$$

$$\rightarrow \boxed{t_R^a = \frac{T}{N\mu_1 \min(\tau_R^d, T/2)}}$$

Results

- **Simulations (Moran process)** $t_C^f \approx t_R^a + \tau_R^f + t_C^a + \tau_C^f$





Solid lines: analytical predictions

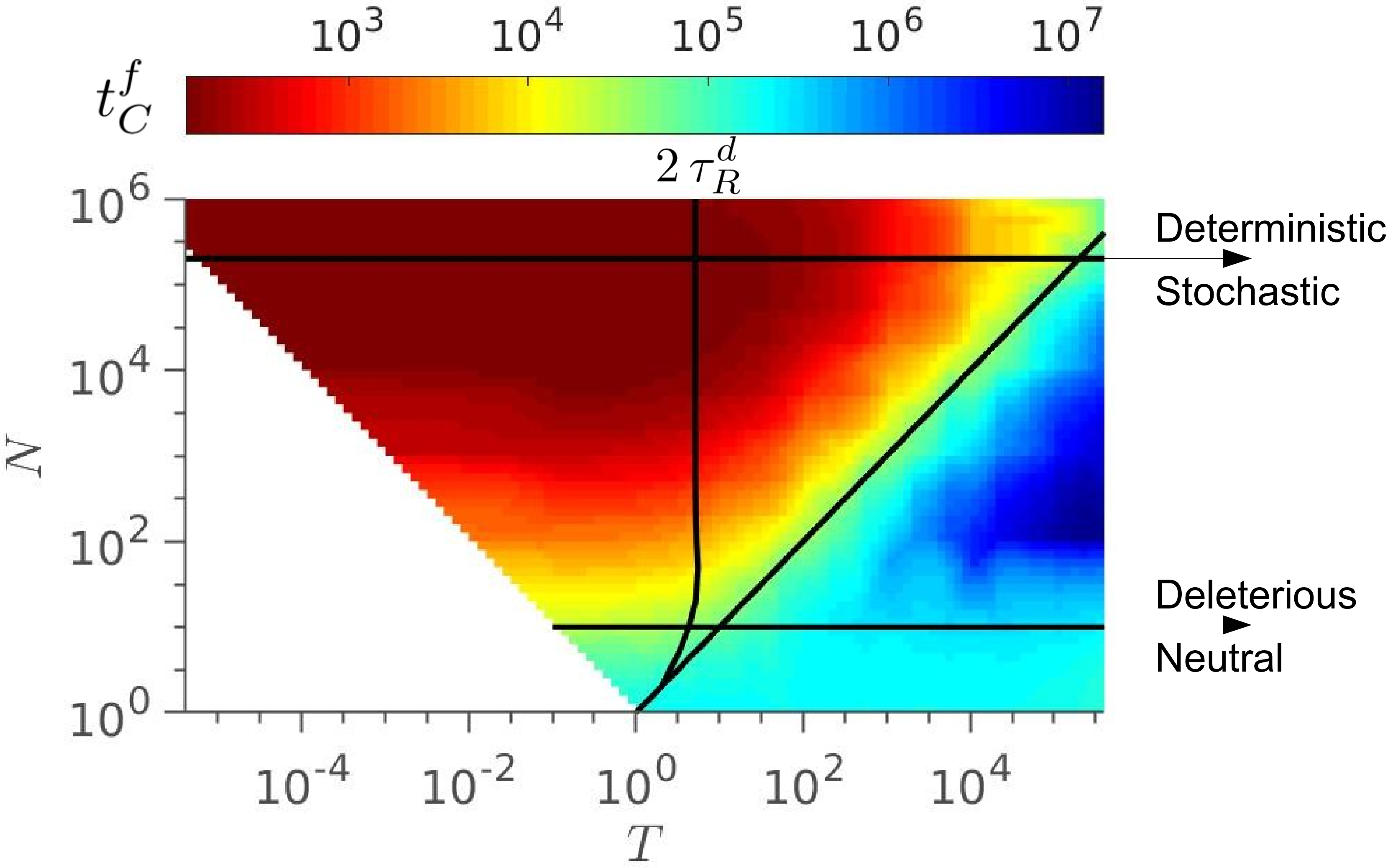
$\mu_1 = 10^{-5},\ \mu_2 = 10^{-3},\ \delta = 0.1$

- **In summary**

- Fast alternations accelerate the evolution of resistance, especially for large populations
- For short enough periods, the first R mutant that appears yields resistance evolution
- What matters is the shortest timescale between the valley-crossing process and the switch-driven process

Results

- **Heatmap**



Deterministic

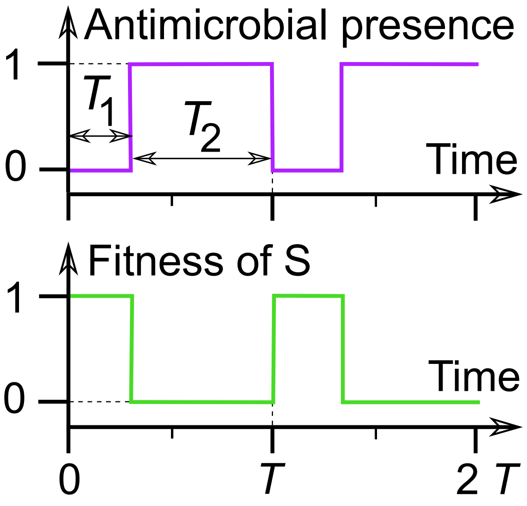
Stochastic

Deleterious

Neutral

Asymmetric alternations

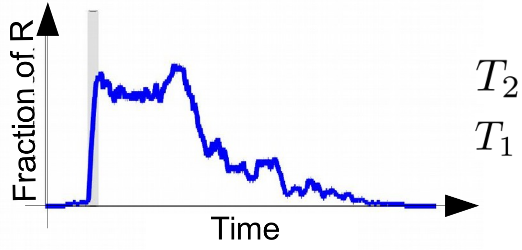
Prediction and results



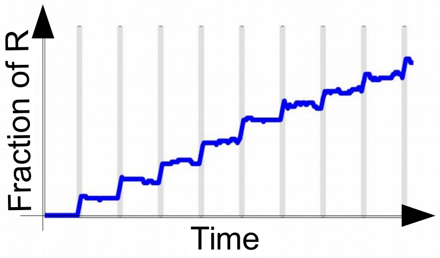
$$t_R^a = \frac{T_1 + T_2}{N\mu_1 \min(\tau_R^d, T_1)}$$

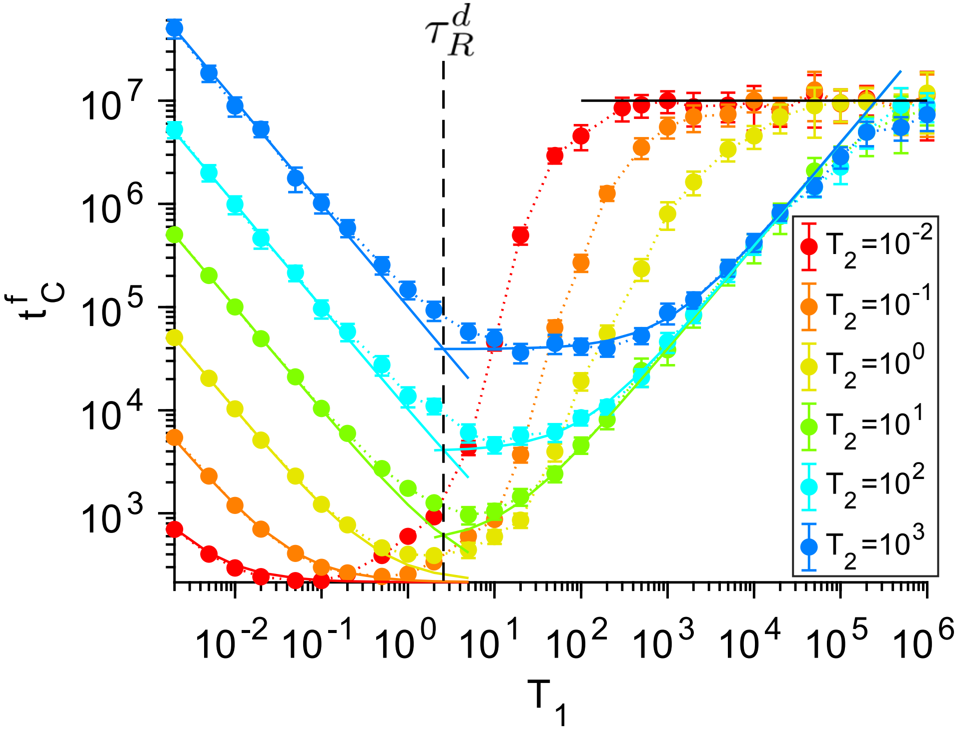
Particular regimes when T_2 is short:

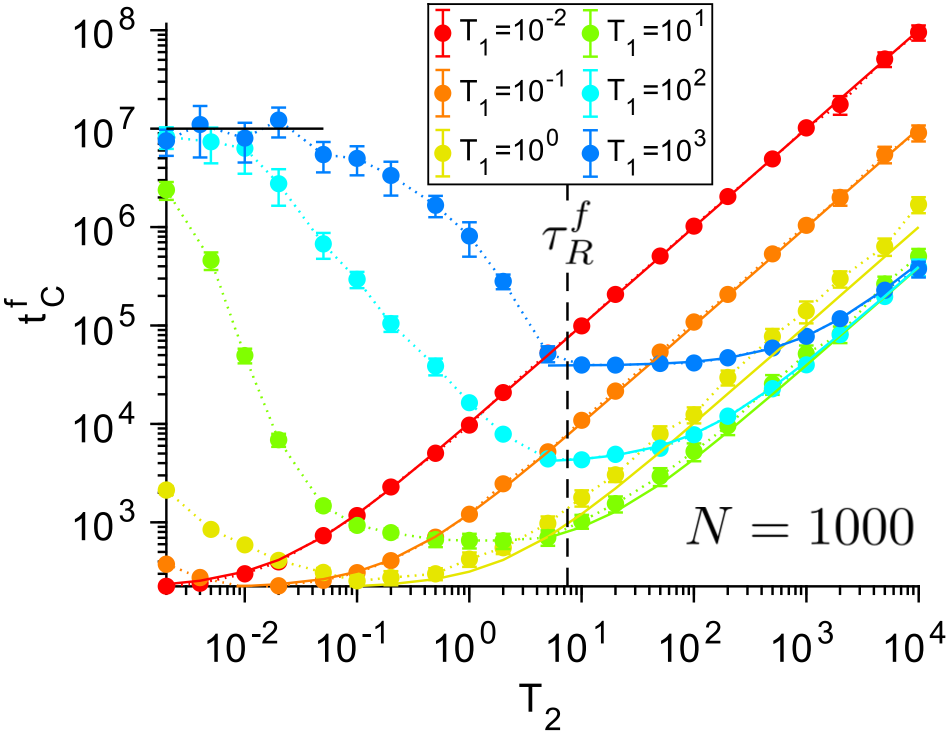
$T_2 \ll \tau_R^f$
$T_1 \gg \tau_R^d$



$T_2 \ll \tau_R^f$
$T_1 \ll \tau_R^d$



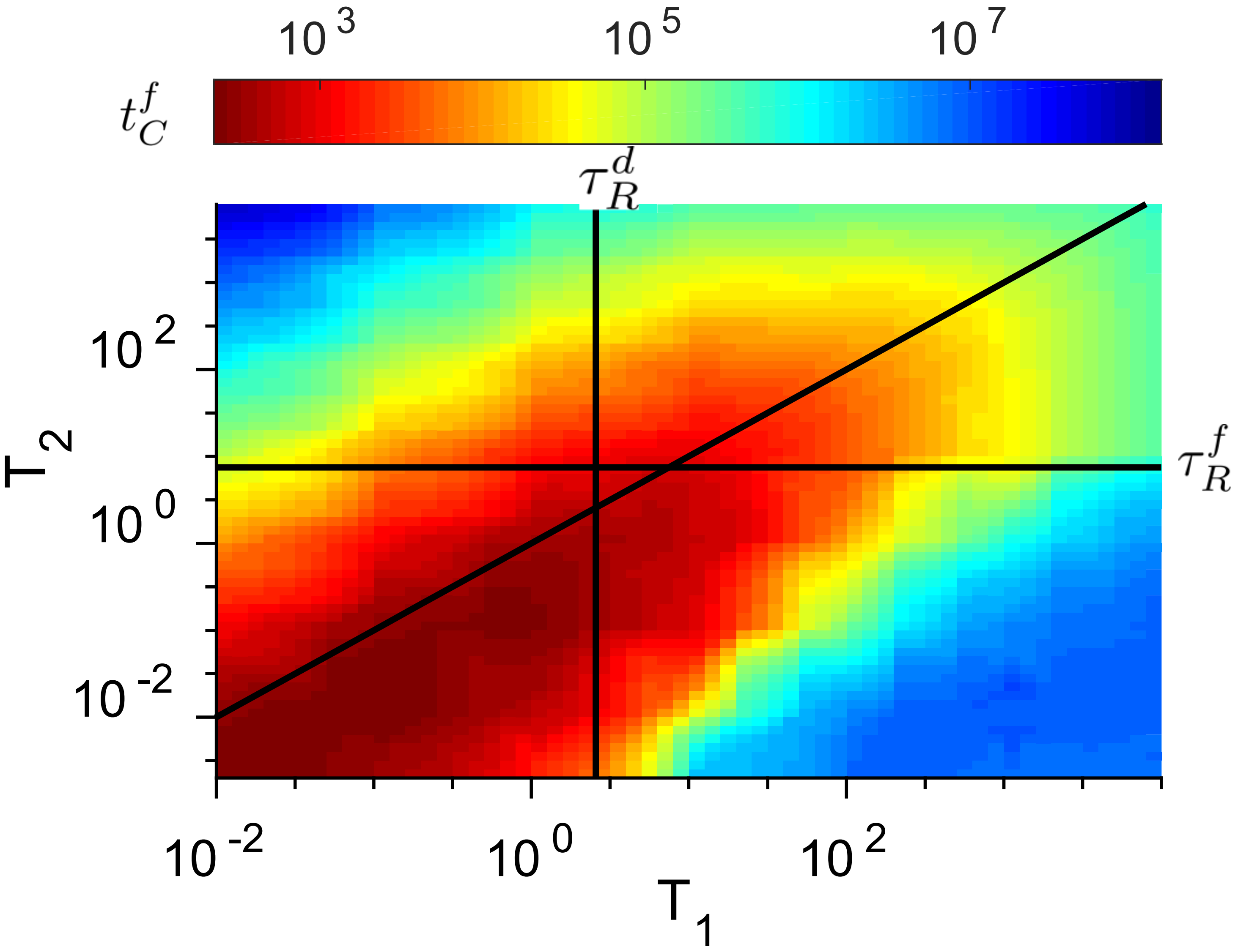




→ **Minimum of the time needed to evolve resistance**

Results

- **Heatmap**



Conclusion

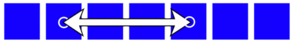
Results for a homogeneous population with fixed size

- Fast alternations accelerate the evolution of resistance, especially for large populations
- Asymmetric alternations → minimum of the time needed to evolve resistance

Discussion

- The worst-case scenario ($T_1 \approx T_2$, lasting a few generations or less) is quite realistic:
  - consider a drug taken every 8 to 12 hours and bacteria dividing every few hours
  - a goal in treatment design is for antimicrobial serum concentration to exceed the MIC for at least 40 to 50% of the time Jacobs (2001)
- Resistance acquisition can then occur in about a day (one S to R mutation; large N)
- Clinically, it is thus important to control for such incursions below the MIC (argument in favor of extended-release antimicrobial formulations)
- The parameters most accelerating resistance can be harnessed in evolution experiments (e.g. in chemostats or morbidostats)

Next steps

- Population with variable size (logistic growth), allowing extinction
- Structured population (realistic – organs, patients) 

Reference

Preprint: http://biorxiv.org/cgi/content/short/279091v1

Acknowledgements

Loïc Marrec Claude Loverdo, David J. Schwab, Raphaël Voituriez

Thanks!