Adaptation via Optimal Control of Mutation in Artificial and Natural Systems

Roman V. Belavkin

Faculty of Science and Technology
Middlesex University, London NW4 4BT, UK

December 5, 2018
Bioinfomatics Educational Seminar
Tokyo University of Science

Evolution and Adaptation

Mutation

Optimal Control of the Muation Rate

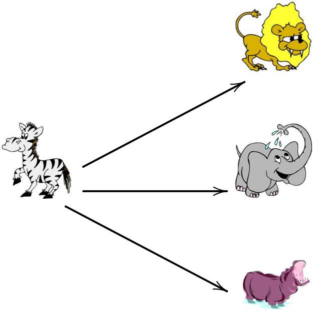
Recombination

Natural selection and adaptation

Definition (Fitness)

A real-valued function $f : \Omega \to \mathbb{R}$ such that

- $f(a) \leq f(b)$ means b has a better reproductive success than a.
- $f(a) \leq f(b)$ means b is better adapted to the environment than a.
- Fitness can be absolute or relative (e.g. may depend on other individuals or environment).



Beneficial	if $f(a) < f(b)$
Neutral	if $f(a) = f(b)$
Deleterious	if $f(a) > f(b)$

Mechanisms of adaptation

Mutaion $a \mapsto b$

$a =$ATAGGACTCA

↓

$b =$ATGGGATTCA

- Mutation rate μ
- Fitness

Recombination $a \mapsto ab$

$a =$ATAGGACTCA
$b =$GTAGGCTAGT

↓

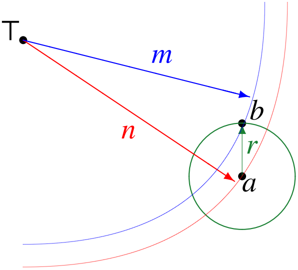
$ab =$ATAGGCTAGA
$ba =$GTAGGACTCT

- Recombination rate μ
- Fitness of parents
- Similarity of parents

Question

Is there an evolutionary strategy to control of parameters of mutation or recombination to maximize their adaptation rates?

Geometric model of adaptation by Fisher (1930)



- Considered Euclidean space $\mathbb{R}^l$ of l traits
- There exists an optimal individual $\top \in \mathbb{R}^l$
- Fitness = negative Euclidean distance to ⊤:

$$f(a) = -d(\top, a)$$

- Probability of adaptation:

$$P_+ := P(m < n \mid n, r) = \sum_{m=0}^{n-1} P(m \mid n, r)$$

(Fisher, 1930):

- How does P_+ depend on mutation radius r?
- P_+ decreases exponentially with r for all $n = d(\top, a) \in [0, \infty)$
- Evolution is more likely to occur via small mutations

Anti-microbial resistance (AMR)

- In autumn 2016, all members of the United Nations agreed to sign a declaration to fight antimicrobial resistance, the secretary general declaring it a 'fundamental threat' to global health and safety.
- New antibiotics can be developed, but they are more expensive than their predecessors (Gelband et al., 2015).
- Bacteria can adapt in just a few hours after introduction of antibiotics (Zhang et al., 2011).

Objective

Understanding how can microbes achieve such an astonishing rate of adaptation.

Evolution and Adaptation

Mutation

Optimal Control of the Muation Rate

Recombination

Geometry of a Hamming space $\mathcal{H}_\alpha^l$

- Instead of Euclidean space consider Hamming space of DNA strings:

$$\mathcal{H}_\alpha^l := \{0, \dots, \alpha - 1\}^l$$

- $\{0, \dots, \alpha - 1\}$ finite alphabet of size α:

$$\{0, 1, 2, 3\} \qquad \{A, C, T, G\} \qquad \alpha = 4$$

- $\{0, \dots, \alpha - 1\}^l$ set of all strings of length l
- Hamming distance between $a, b \in \mathcal{H}_\alpha^l$:

$$d(a, b) = l - \sum_{i=1}^{l} \delta_{a_i b_i}\,, \qquad \delta_{a_i b_i} = \begin{cases} 1 & \text{if } a_i = b_i \\ 0 & \text{otherwise} \end{cases}$$

- Hamming space $\mathcal{H}_\alpha^l$ has finite diameter l.

Evolutionary dynamics

- Probability space $(\Omega \equiv \mathcal{H}_\alpha^l, 2^\Omega, P_t)$ at time $t \geq 0$
- $P_t(\omega)$ is a probability distribution of genotypes at time $t \geq 0$
- Distribution of fitness levels

$$P_t(n) := P_t\{\omega : f(\omega) = n\}$$

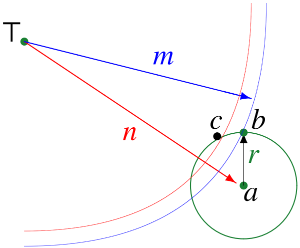
- Markov evolutionary dynamics:

$$P_{t+s}(m) = \left[\sum_{n=0}^{l} P_\mu(m \mid n)\,(\cdot)\right]^s P_t(n)$$

Optimal evolutionary dynamics

Find optimal $T(\cdot) = \sum_{n=0}^{l} P(m \mid n)(\cdot)$

Mutation



$\top$	=	A	A	A	A	A
a	=	C	A	C	A	C
b	=	A	C	T	A	C

$$d(a,b) = 3$$

$$d(\top, a) = 3$$

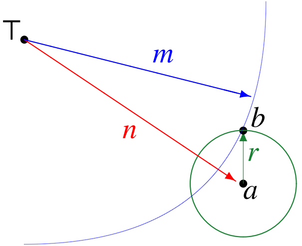
$$d(\top, b) = 3$$

Problem

Find conditional probabilty $P(d(\top, b) = m \mid d(\top, a) = n, d(a,b) = r)$

$$P_{t+1}(m) = \sum_{n=0}^{l} \left(\sum_{r=0}^{l} P(m \mid n, r) \underbrace{P(r \mid n)}_{\textit{Mutation}} \right) P_t(n)$$

Mutation



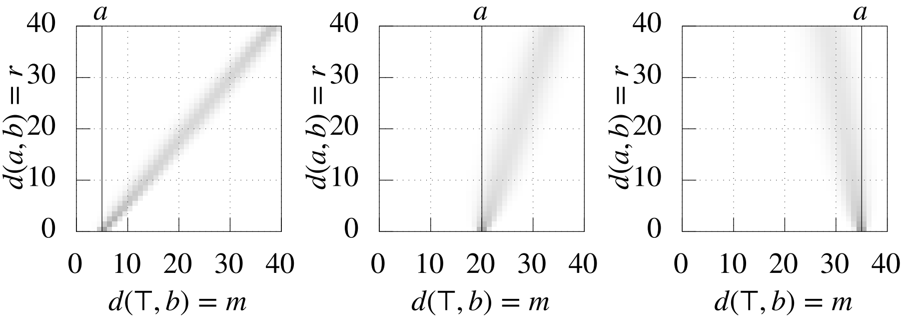
T	$=$	A	A	A	A	A
a	$=$	C	A	C	A	C
b	$=$	A	C	T	A	C
		r_+	r_-	r_0		

- $r_+ + r_- + r_0 = r$
- $r_+ - r_- = n - m$

Theorem (Belavkin (2011))

$$P(m \mid n, r) = \frac{\sum_{r_+=0}^{r} (\alpha - 2)^{r_0} \binom{n - r_+}{r_0} (\alpha - 1)^{r_-} \binom{l - n}{r_-} \binom{n}{r_+}}{(\alpha - 1)^r \binom{l}{r}}$$

Effect of the mutation radius r



$$\mathbb{E}_P\{m \mid n, r\} = n + \left(1 - \frac{n}{l(1 - 1/\alpha)}\right) r$$

Mutation radius and mutation rate

Assumptions

Independence : mutations at different sites are independent:

$$P(b \mid a) = P(b_1, \dots, b_l \mid a_1, \dots, a_l) = \prod_{i=1}^{l} P(b_i \mid a_i)$$

Neutrality : selection does not act on sites $P(b_i \mid a_i) = P(b_j \mid a_j)$.

One parameter : $P(b_i \mid a_i) = \mu$ for all $i \in \{0, \dots, \alpha - 1\}$, called the mutation rate (Jukes & Cantor, 1969).

Random mutation radius

In this case mutation radius r is binomial random variable:

$$P_\mu(r) = \binom{l}{r} \mu^r (1 - \mu)^{l-r}$$

Evolution and Adaptation

Mutation

Optimal Control of the Muation Rate

Recombination

Optimal control of mutation rate

The fact that P_+ varies with mutation radius and distance to optimum means that we can find optimal mutation rate control functions $\mu(n)$, which are solutions to the following optimization problems (Belavkin, 2010, 2012):

- Maximum adaptation in no more than λ generations

$$\text{maximize}_{\mu(x)} \quad \mathbb{E}\{f_{s+t}\} \quad \text{subject to} \quad t \leq \lambda$$

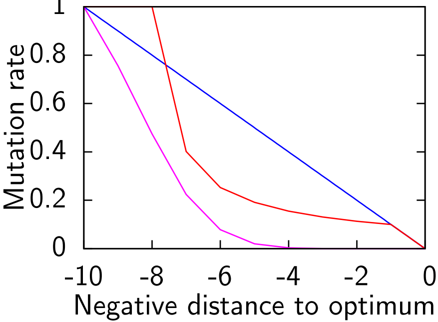
- Maximum adaptation in no more than λ bits between p_s and p_{s+t}:

$$\text{maximize}_{\mu(x)} \quad \mathbb{E}\{f_{s+t}\} \quad \text{subject to} \quad I(p_{s+t}, p_s) \leq \lambda$$

- Cumulative criterion:

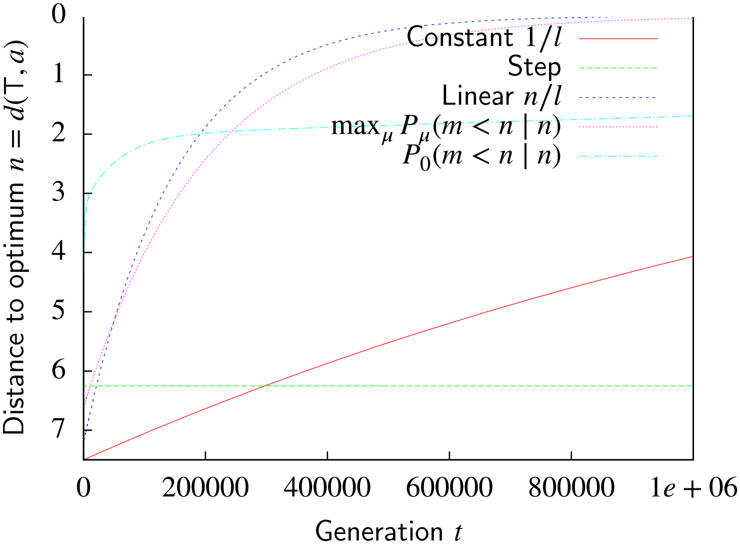
$$\sup_{\mu(x)} \sum_{t=0}^{\lambda} \mathbb{E}_{\mu(x)}\{f_{s+t}\} \leq \sum_{t=0}^{\lambda} \sup_{\mu(x)} \mathbb{E}_{\mu(x)}\{f_{s+t}\}$$

Optimal mutation rate control functions in $\mathcal{H}_4^{10}$



Belavkin, R. V., Channon, A., Aston, E., Aston, J., Krasovec, R., Knight, C. G. (2016). Monotonicity of Fitness Landscapes and Mutation Rate Control. *Journal of Mathematical Biology*, Vol. 73:6, pp 1491–1524, Springer.

Expected fitness in time



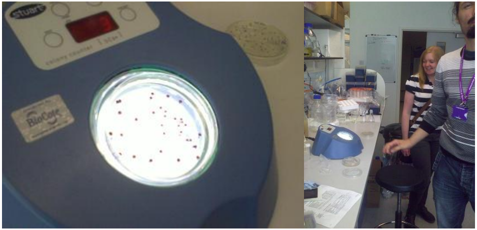
Evolution as an information dynamic system

- EPSRC Sandpit '*Math of Life*' (July, 2009):

- Three year project (2010–13)
- Followed by two BBSRC project.

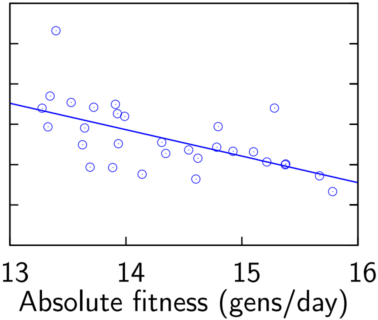
Middlesex University : Roman Belavkin
University of Warwick : John Aston
University of Keele : Alastair Channon & Elizabeth Aston
University of Manchester : Chris Knight, Rok Krašovec & Danna Gifford

Mutation rate control in *E. coli*



- Used strains of *Escherichia coli* K-12 MG1665
- Fluctuation test using media 50μg/ml of Rifamipicin
- Estimated mutation rates μ in *E.coli* strains grown in Davis minimal medium with different amount of glucose.

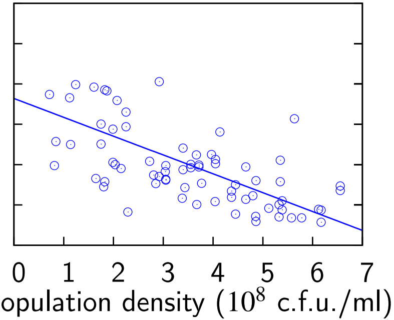
Experimental results (Krašovec et al., 2014)



- Strong relationship between μ and density of cells ($p < .0001$).
- No such relationship in the *luxS* quorum sensing mutant ($p = .0234$).

Krašovec, R., Belavkin, R., Aston, J., Channon, A., Aston, E., Rash, B., Kadirvel, M., Forbes. S., Knight, C. G. (2014, April). Mutation-rate-plasticity in rifampicin resistance depends on Escherichia coli cell-cell interactions. *Nature Communications*, Vol. 5 (3742).

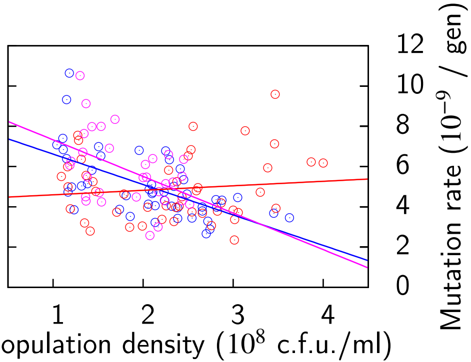
Experimental results (Krašovec et al., 2014)



- Strong relationship between μ and density of cells ($p < .0001$).
- No such relationship in the *luxS* quorum sensing mutant ($p = .0234$).

Krašovec, R., Belavkin, R., Aston, J., Channon, A., Aston, E., Rash, B., Kadirvel, M., Forbes. S., Knight, C. G. (2014, April). Mutation-rate-plasticity in rifampicin resistance depends on Escherichia coli cell-cell interactions. *Nature Communications*, Vol. 5 (3742).

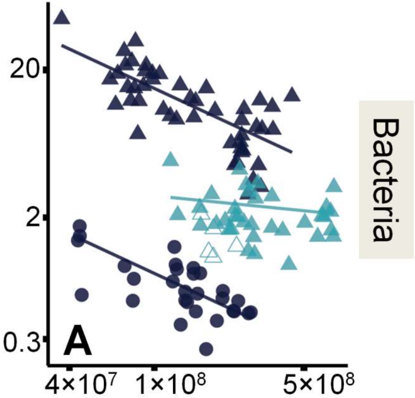
Experimental results (Krašovec et al., 2014)



- Strong relationship between μ and density of cells ($p < .0001$).
- No such relationship in the *luxS* quorum sensing mutant ($p = .0234$).

Krašovec, R., Belavkin, R., Aston, J., Channon, A., Aston, E., Rash, B., Kadirvel, M., Forbes. S., Knight, C. G. (2014, April). Mutation-rate-plasticity in rifampicin resistance depends on Escherichia coli cell-cell interactions. *Nature Communications*, Vol. 5 (3742).

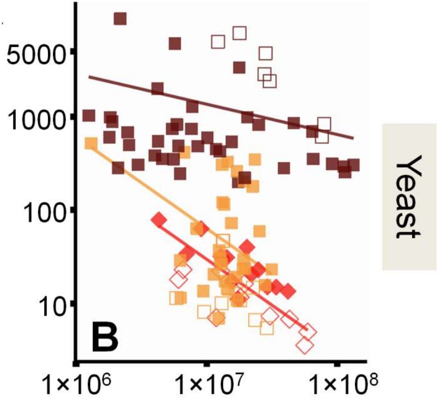
Plastic mutation rates in bacteria (Krašovec et al., 2017)



- *rifampicin* (triangles)
- *nalidixic acid* in *E.coli* (dark circles) and in *P. aeruginosa* (light circles)

Krašovec, R., Richards, H., Gifford, D. R., Hatcher, C., Faulkner, K. J., Belavkin, R. V., Channon, A., Aston, E., McBain, A. J., Knight, C. G. (2017). Spontaneous mutation rate is a plastic trait associated with population density across domains of life. *PLoS Biology*, 15:8.

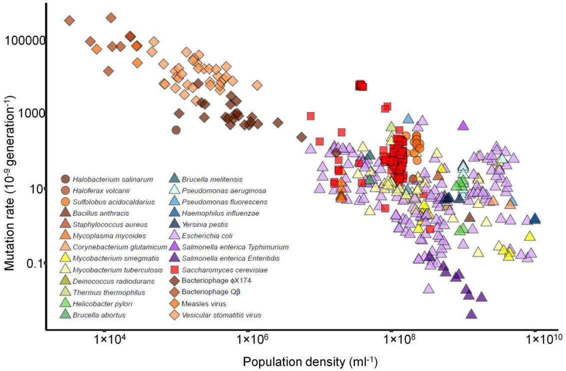
Plastic mutation rates in yeast (Krašovec et al., 2017)



- *hygromycin B* (squares) in *S. cerevisiae*
- *5-FOA* (diamonds)

Krašovec, R., Richards, H., Gifford, D. R., Hatcher, C., Faulkner, K. J., Belavkin, R. V., Channon, A., Aston, E., McBain, A. J., Knight, C. G. (2017). Spontaneous mutation rate is a plastic trait associated with population density across domains of life. *PLoS Biology*, 15:8.

Plastic rates in all domains of life (Krašovec et al., 2017)



>70 years of published data (1943–2016), 67 studies, 26 species.

Krašovec, R., Richards, H., Gifford, D. R., Hatcher, C., Faulkner, K. J., Belavkin, R. V., Channon, A., Aston, E., McBain, A. J., Knight, C. G. (2017). Spontaneous mutation rate is a plastic trait associated with population density across domains of life. *PLoS Biology*, 15:8.

Evolution and Adaptation

Mutation

Optimal Control of the Muation Rate

Recombination

Mechanisms of adaptation

Mutaion $a \mapsto b$

$a =$ATAGGACTCA
↓
$b =$ATGGGATTCA

- Mutation rate μ
- Fitness

Recombination $a \mapsto ab$

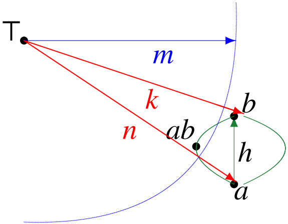
$a =$ATAGGACTCA
$b =$GTAGGCTAGT
↓
$ab =$ATAGGCTAGA
$ba =$GTAGGACTCT

- Recombination rate μ
- Fitness of parents
- Similarity of parents

Horizontal gene transfer (HGT)

A form of recombination — part of a DNA string is picked up from another cell or from extracellular environment.
May occur via different mechanisms (e.g. conjugation, transduction and natural transformation).

Recombination



$\top$	=	A	A	A	A	A
a	=	C	A	C	A	C
b	=	A	C	T	A	C
ab	=	A	C	C	A	C

- $d(\top, a) = d(\top, b) = d(a, b) = 3$
- $d(\top, ab) = 3$

Problem

Find $P(d(\top, ab) = m, d(\top, ba) = m' \mid d(\top, a) = n, d(\top, b) = k, d(a, b) = h, r)$

$$P(m, m') = \sum_{n=0}^{l} \sum_{k=0}^{l} \left(\sum_{h=0}^{l} \sum_{r=0}^{l} P(m, m' \mid n, k, h, r) \underbrace{P(r, h \mid n, k)}_{\textit{Recombination}} \right) P(n, k)$$

Probability of recombination onto $S(\mathrm{T}, m)$

Theorem

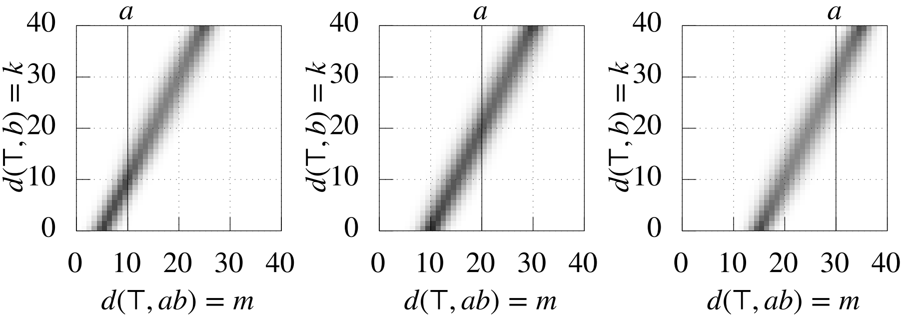
$$P(m \mid n, k, h, r) = \frac{\sum_{n_+} (\alpha - 2)^{n_0} \binom{n-n_+}{n_0} (\alpha - 1)^{n_-} \binom{l-n}{n_-} \binom{n}{n_+} \sum_{r_+} \binom{l-n_+-n_-}{r-r_+-r_-} \binom{n_-}{r_-} \binom{n_+}{r_+}}{\binom{l}{r} \sum_{n_+} (\alpha - 2)^{n_0} \binom{n-n_+}{n_0} (\alpha - 1)^{n_-} \binom{l-n}{n_-} \binom{n}{n_+}}$$

where

$$\begin{aligned} n_+ - n_- = n - k &\Rightarrow & n_- = n_+ - (n - k) \geq 0 \\ n_+ + n_- + n_0 = h &\Rightarrow & n_0 = h - 2n_+ + (n - k) \geq 0 \\ r_+ - r_- = n - m &\Rightarrow & r_- = r_+ - (n - m) \geq 0 \, , \; r - r_+ - r_- \geq 0 \end{aligned}$$

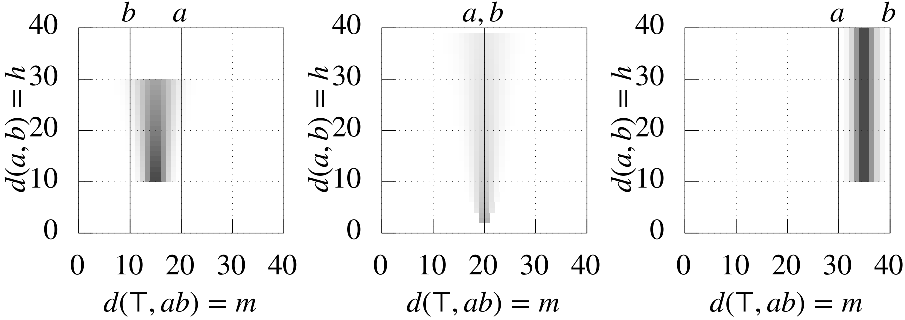
Triangle inequalities $|n - h| \leq k \leq n + h$, $|n - \min\{r, h\}| \leq m \leq n + \min\{r, h\}$

Effect of the second parent's distance k



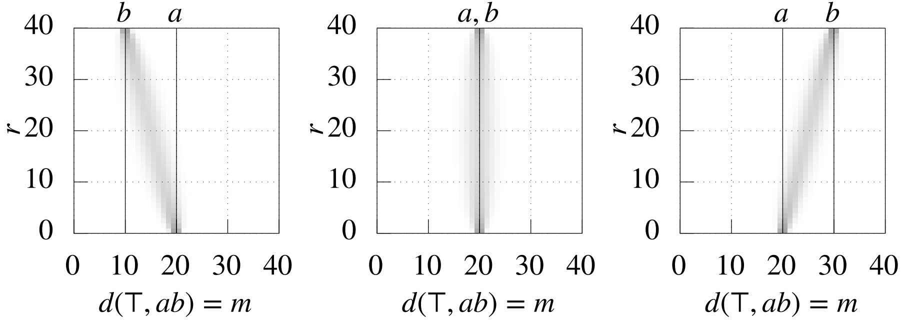
$$\mathbb{E}_P\{m \mid n, k, h, r\} = n + \frac{(k - n)}{l} r$$

Effect of the recombination height h



$$\mathbb{E}_P\{m \mid n, k, h, r\} = n + \frac{(k - n)}{l} r$$

Effect of the recombination radius r



$$\mathbb{E}_P\{m \mid n, k, h, r\} = n + \frac{(k-n)}{l} r$$

Acknowledgements

U. of Manchester : Christopher Knight, Rok Krašovec, Danna Gifford

U. of Keele : Alastair Channon & Elizabeth Aston

U. of Warwick John Aston (now in Cambridge)

EPSRC *SANDPIT: Evolution as an Information Dynamic System* (EP/H031936/1)

BBSRC *The theory and practice of evolvability: Effects and mechanisms of mutation rate* (BB/L009579/1)

BBSRC *Adaptive landscapes of antibiotic resistance: population size and 'survival-of-the-flattest'* (BB/M021106/1)

Thank you!

Evolution and Adaptation

Mutation

Optimal Control of the Muation Rate

Recombination

Belavkin, R. V. (2010). Information trajectory of optimal learning. In M. J. Hirsch, P. M. Pardalos, & R. Murphey (Eds.), *Dynamics of information systems: Theory and applications* (Vol. 40, pp. 29–44). Springer.

Belavkin, R. V. (2011). Mutation and optimal search of sequences in nested Hamming spaces. In *IEEE Information Theory Workshop.* IEEE.

Belavkin, R. V. (2012). On evolution of an information dynamic system and its generating operator. *Optimization Letters*, *6*(5), 827–840.

Fisher, R. A. (1930). *The genetical theory of natural selection*. Oxford: Oxford University Press.

Gelband, H., Miller-Petrie, M., Pant, S., Gandra, S., Levinson, J., Barter, D., et al. (2015). *The state of the World's antibiotics.* Center for Disease Dynamics, Economics & Policy.

Jukes, T. H., & Cantor, C. R. (1969). *Evolution of protein molecules*. New York: Academic Press.

Krašovec, R., Belavkin, R. V., Aston, J. A. D., Channon, A., Aston, E., Rash, B. M., et al. (2014, April). Mutation rate plasticity in rifampicin resistance depends on escherichia coli cell-cell interactions. *Nature Communications*, *5*(3742).

Krašovec, R., Richards, H., Gifford, D. R., Hatcher, C., Faulkner, K. J., Belavkin, R. V., et al. (2017). Spontaneous mutation rate is a plastic trait associated with population density across domains of life. *PLoS Biology*, *15*(8).

Zhang, Q., Lambert, G., Liao, D., Kim, H., Robin, K., Tung, C. kuan, et al. (2011). Acceleration of emergence of bacterial antibiotic resistance in connected microenvironments. *Science*, *333*, 1764–1767.