Recent developments in biological evolution

- Chin-Kun Hu (胡進錕)

  Institute of Physics, Academia Sinica, Taipei 11529, Taiwan and NCTS, Hsinchu
- E-mail: **huck@phys.sinica.edu.tw**
- http://proj1.sinica.edu.tw/~statphys/

Collaborators:
A. E. Allahverdyan, Yerevan Physics Institute, Armenia
Guan-Rong Huang, National Center for Theoretical Sciences, Hsinchu, Taiwan
David B. Saakian, Institute of Physics, AS, Taipei and Yerevan Physics Institute, Armenia
Makar H. Ghazaryan, Yerevan Engineering University, Armenia
Tatiana Yakushkina, National Research University Higher School of Economics, Moscow


China
Myanmar
Laos
Cambodia
Philippines
Indonesia
Papua New Guinea


CHINA
Chi-lung
TAIPEI
Su-ao
TAIWAN STRAIT
Chang-hua
Pescadores
Ma-kong
T'ai-nan
T'ai-tung
Kao-hsiung
PACIFIC
OCEAN
LUZON STRAIT


中華民國
REPUBLIC OF CHINA
TAIWAN
護照
PASSPORT

Homepage of Professor Chin-Kun Hu - Mozilla Firefox

檔案(F) 編輯(E) 檢視(V) 歷史(S) 書籤(B) 工具(T) 說明(H)

http://proj1.sinica.edu.tw/~statphys/people/huck/index.html

最常瀏覽 新手上路 即時新聞

Homepage of Professor Chin-Kun Hu

Professor Chin-Kun Hu 胡進錕

Laboratory of Statistical and Computational Physics

Address: Institute of Physics, Academia Sinica, Nankang, Taipei 11529, Taiwan

Tel: +886-2-27896720 (Office)

+886-2-27335033 (Home)

FAX: +886-2-27834187

Email: huck@phys.sinica.edu.tw

URL: http://www.sinica.edu.tw/~statphys/

Biographically - *Position Held*

1991.6.-1993.6.	Visiting Scholar, Department of Physics, Harvard Univ., Cambridge, MA 02138, USA
1984.9.-Present	Research Fellow, Institute of Physics, Academia Sinica, Taipei, Taiwan
1983.9.-1984.8.	Visiting Expert, Institute of Physics, Academia Sinica, Taipei, Taiwan
1981.9.-1983.8.	Research Associate, Dept. of Chemistry, University of Toronto, Toronto, Ontario, Canada
1979.7.-1981.8.	Research Associate, Dept. of Physics and Astronomy, University of Maine, Orono, Maine, USA
1978.11.-1979.7.	Postgraduate Research Physicist,Inst. for Pure and Applied Physical Sciences, Univ. of California at San Diego, USA
1978.6.-1978.10.	Associate Research Fellow, Precision Instrument Development Center of National Science Council in Taiwan
1977.1.-1978.5.	Associate Professor, Chung Cheng Institute of Technology in Taiwan

檔案(F) 編輯(E) 檢視(V) 我的最愛(A) 工具(T) 說明(H)

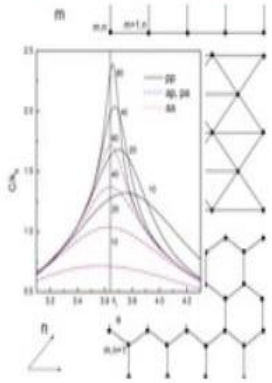
網頁快訊圖庫 建議的網站 新聞 外交部未宣布與甘國斷交...

Laboratory of Statistical and Computational Physics

The LSCP, Academia Sinica

Welcome to Laboratory of Statistical and Computational Physics (LSCP) at the Institute of Physics (IOP) of Academia Sinica, located in Nankang, Taipei, Taiwan. LSCP is devoted to frontier research in statistical and computational physics (SCP), applications of SCP to problems in physical, biological, and social sciences, sponsoring meetings in SCP, and promoting education and research of SCP in developing countries.

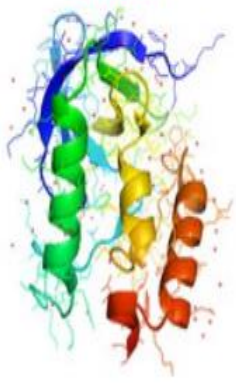
Equilibrium Statistical Physics



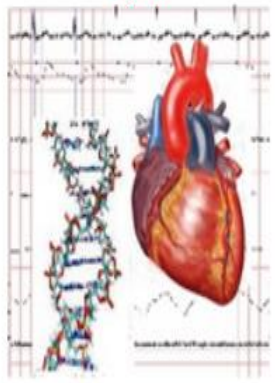
Non-equilibrium Statistical Physics

Nonlinear Dynamics & Avalanche Process

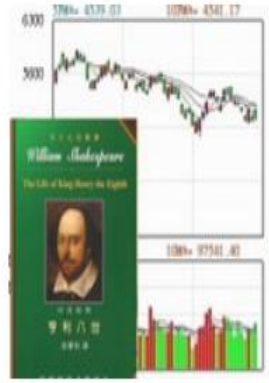
Parallel Computing & Protein Simulation



Theoretical Biophysics & Medical Physics



Economical & Social Model



上午 10:52
2014/7/23

Liquid-gas, Ising & L-J models have same exponents

Jan V. Sengers · Joseph G. Shanks

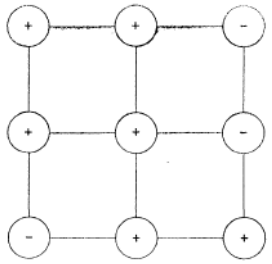
Experimental Critical-Exponent Values for Fluids

J Stat Phys (2009) 137: 857–877

$\gamma = 1.238 \pm 0.012,$

$\nu = 0.629 \pm 0.003$

Beta=0.3245



H. W. J. Blote, E. Luijten, J. R. Heringa.
3D Ising model: Monte Carlo study
J. Phys. A 28, 6289 (1995)

Exponent	Expressed in RG exp.	Value
α	$2 - d/y_t$	0.110 (2)
β	$(d - y_h)/y_t$	0.3267 (10)
γ	$(2y_h - d)/y_t$	1.237 (2)
δ	$y_h/(d - y_h)$	4.786 (14)
η	$2 - 2y_h + d$	0.037 (3)
ν	$1/y_t$	0.6301 (8)
θ	$-y_i/y_t$	0.52 (4)

H. Watanabe, N. Ito, and C.-K. Hu, Phase diagram and universality of the Lennard-Jones gas-liquid system, J. Chem. Phys.,136, 204102 (2012).

$\beta = 0.3285(7).$ $\nu = 0.63(4).$

Yang and Lee, Phys. Rev. 87, 404 (1952) and Lee and Yang, Phys. Rev. 87, 410 (1952):
Lattice gas model and hence Ising model can represent gas-liquid systems.

Biological Polymorphism 生物多形現象

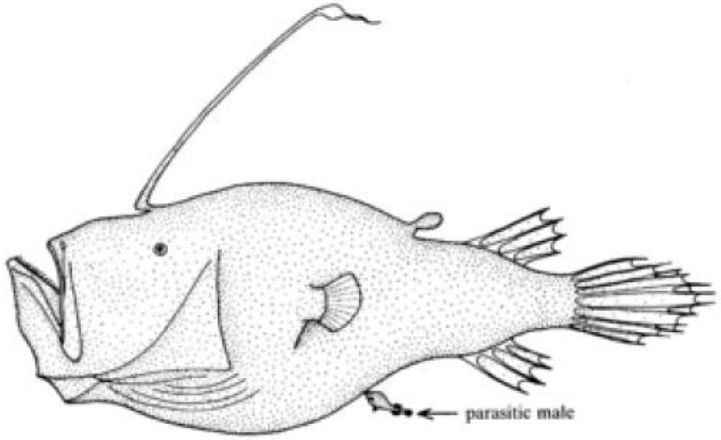
Two or more clearly different phenotypes exist in the same population of a species — in other words, the occurrence of more than one *form* or *morph*.

A. E. Allahverdyan and **C.-K. Hu: PRL** 102, 058102 (2009).

Examples:sexual dimorphism (兩性異形)

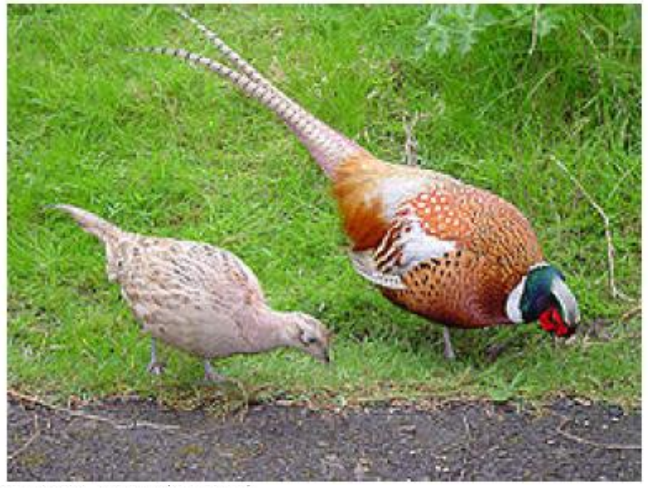
http://en.wikipedia.org/wiki/Sexual_dimorphism

Spider (蜘蛛)



Female Triplewart seadevil, an anglerfish, with male attach

Common Pheasant(雉雞)



Definition of polymorphism: two or more different phenotypes co-exist in one interbreeding population **(雜種繁殖群體)**

-- May or may not have a genetic basis

-- Frequently related with varying environment

Example: land snail grove snail (小樹林蝸牛)*Cepaea Nemoralis*.
Two morphs brown (**B**) and yellow (**Y**) compete for resources.

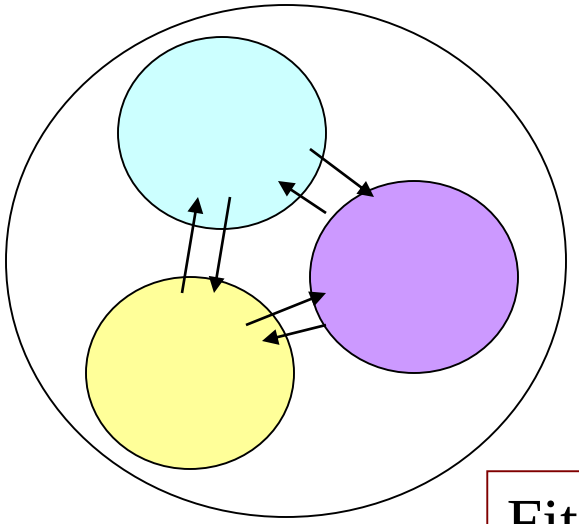
B (**Y**) is less visible for predators at spring (summer and autumn).

Y is resistant to high and low temperatures: advantage at summer and winter.

B and **Y** have different advantages under different environmental conditions.

LM Cook, 1998

Replicator approach: Population consists of several interacting groups



$p_k = N_k / N$ frequency of each group

$$\dot{N}_k = N_k f_k(p)$$

Births-deaths process

Fitness = births – deaths + inter-group interactions.

Probability of meeting for two individuals depend on *p*'s

$$\dot{p}_k = p_k [\, f_k(a,p) - \sum_{l=1}^{n} p_l f_l \,]$$

Replicator equation

$$f_k(a,p) = \sum_{l=1}^{n} a_{kl} p_l$$

Simplest assumption about interactions

Interaction parameters: payoffs

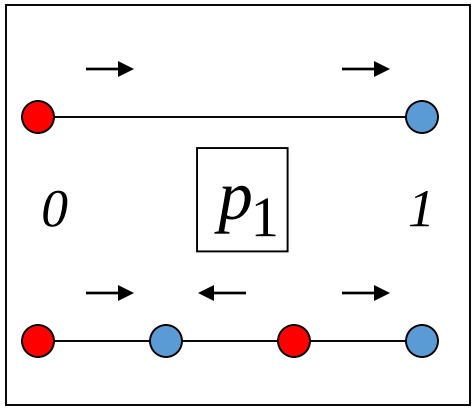
Two morphs.

$$\dot{\bar{p}}_1 = \bar{p}_1(1-\bar{p}_1)[\bar{A} - \bar{B}\bar{p}_1 - C\bar{p}_1(1-\bar{p}_1)]$$

$$a_{12}(\tau) = a_{21}(\tau)$$

$$C \equiv \frac{1}{\omega}[\overline{\hat{a}_{12}(\tilde{a}_{22} - \tilde{a}_{11})} + \overline{\hat{a}_{22}\tilde{a}_{11}}]$$

$\bar{a}_{11} > \bar{a}_{12} > \bar{a}_{22}$ $\quad a_{kl}(\tau) = \bar{a}_{kl} + \tilde{a}_{kl}(\tau).$ $\quad \partial_\tau \hat{a}_{kl}(\tau) = \tilde{a}_{kl}(\tau), \quad \overline{\hat{a}_{kl}} = 0.$



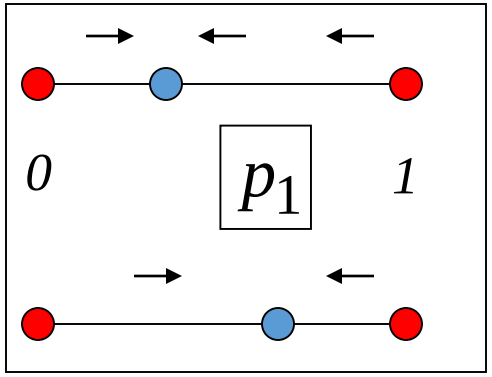
=> Morph # 1 (generalist) dominates for *C=0*

Two new rest-points: Polymorphism for sufficiently positive *C*

$$A(t) \equiv a_{12}(t) - a_{22}(t),$$

$$B(t) \equiv 2a_{12}(t) - a_{11}(t) - a_{22}(t),$$

$\bar{a}_{12} > \bar{a}_{11}, \bar{a}_{22}$

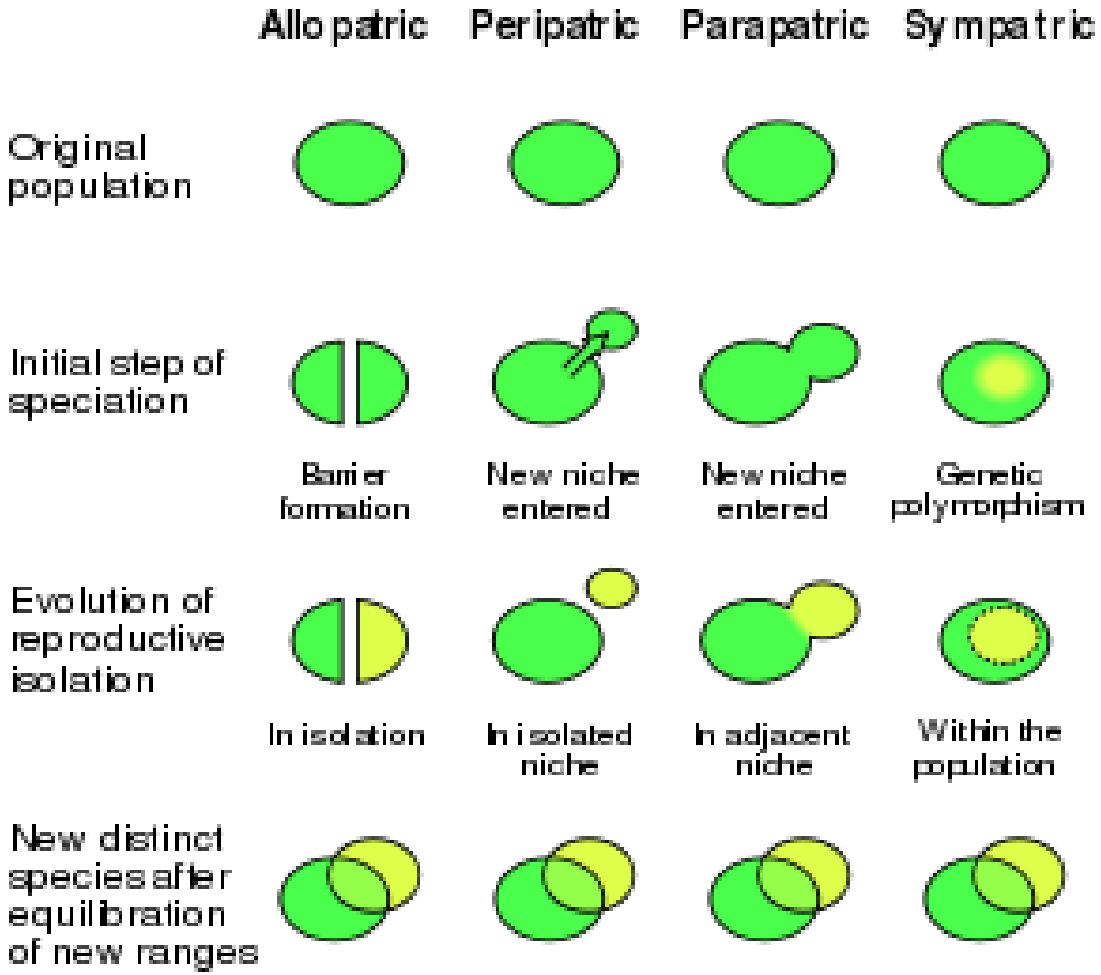


=> Coexistence for *C=0 (heterozygote advantage)*

No new rest-points: Polymorphism is kept for any *C*

Allopatric speciation(異域物種形成) vs. Sympatric speciation (同域物種形成)

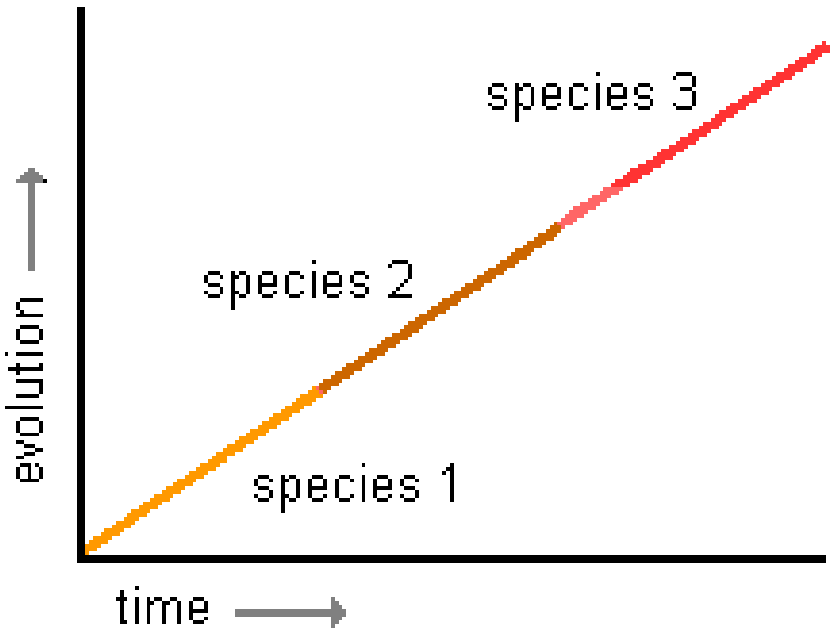
- Allopatric speciation, also known as **geographic speciation**, is the phenomenon whereby biological populations are physically isolated by an extrinsic barrier and evolve intrinsic (genetic) reproductive isolation, such that if the barrier breaks down, individuals of the populations can no longer interbreed.
- Sympatric speciation is the genetic divergence of various populations (from a single parent species) inhabiting the same geographic region, such that those populations become different species.
- **Our theory gives positive impact on sympatric speciation.**

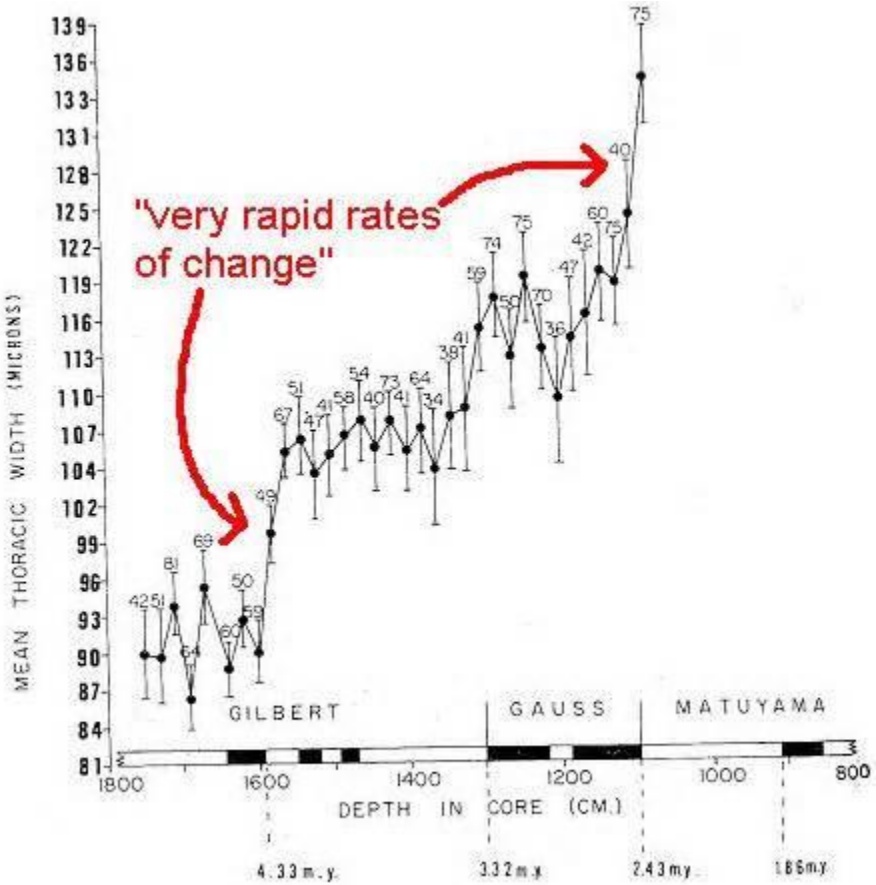


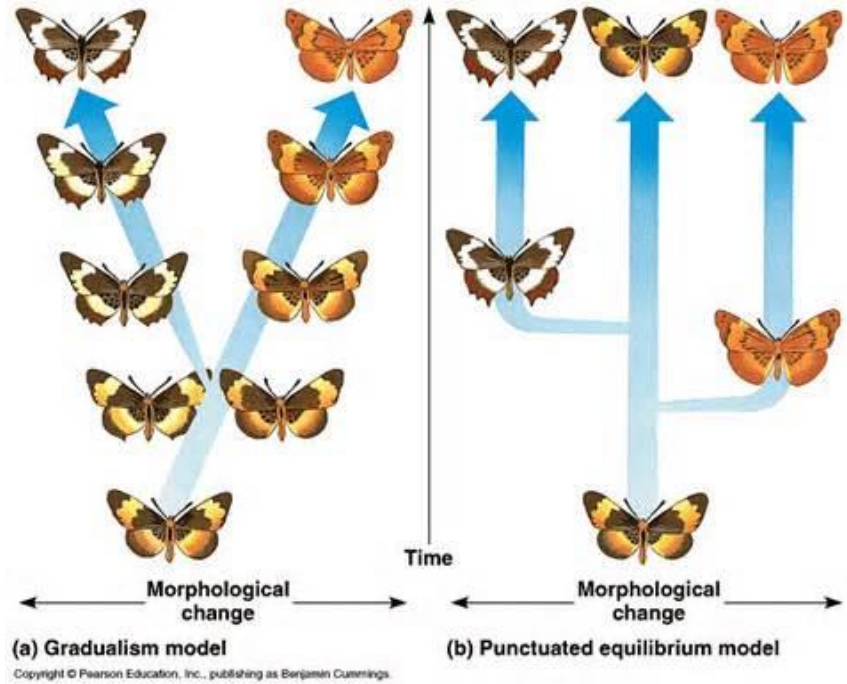
Micro and Macro Evolution

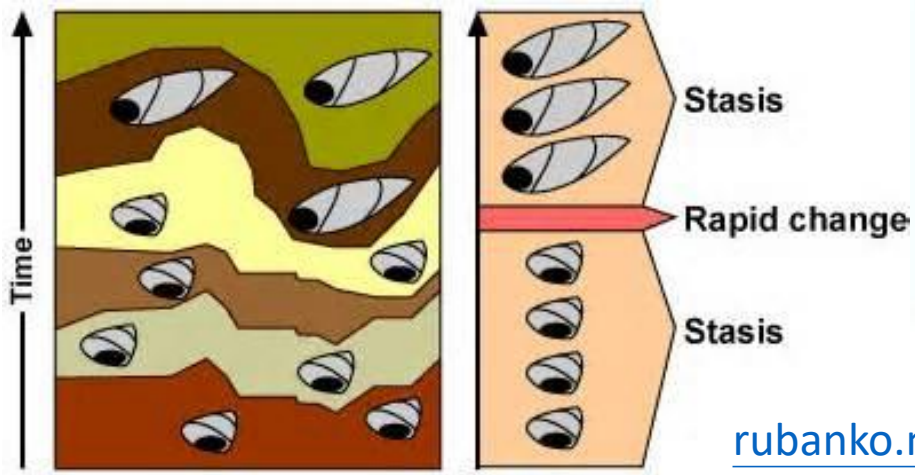
http://anthro.palomar.edu/synthetic/synth_9.htm

Throughout most of the 20th century, researchers developing the synthetic theory of evolution primarily focused on **microevolution** , which is slight genetic change over a few generations in a population. Until the 1970's, it was generally thought that these changes from generation to generation indicated that past species evolved gradually into other species over millions of years. This model of long term gradual change is usually referred to as **gradualism** or phyletic gradualism . It is essentially the 19th century Darwinian idea that species evolve slowly at a more or less steady rate. A natural consequence of this sort of **macroevolution** would be the slow progressive change of one species into the next in a line, as shown by the graph on the right









EOLSS - PATTERNS AND RATES OF SPECIES EVOLUTION

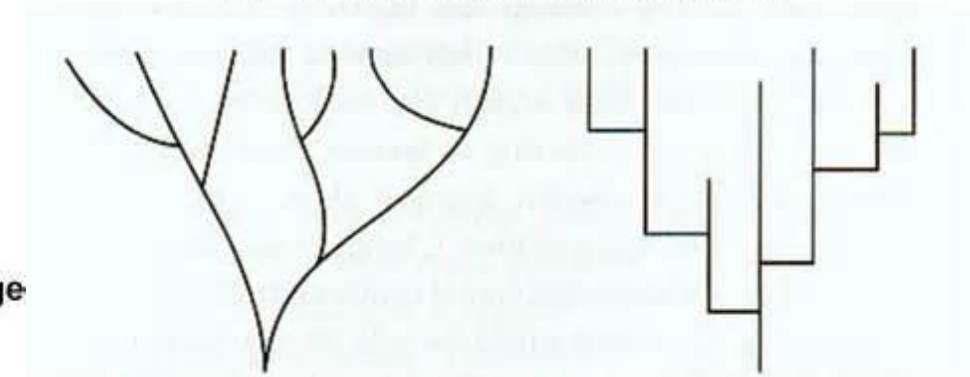
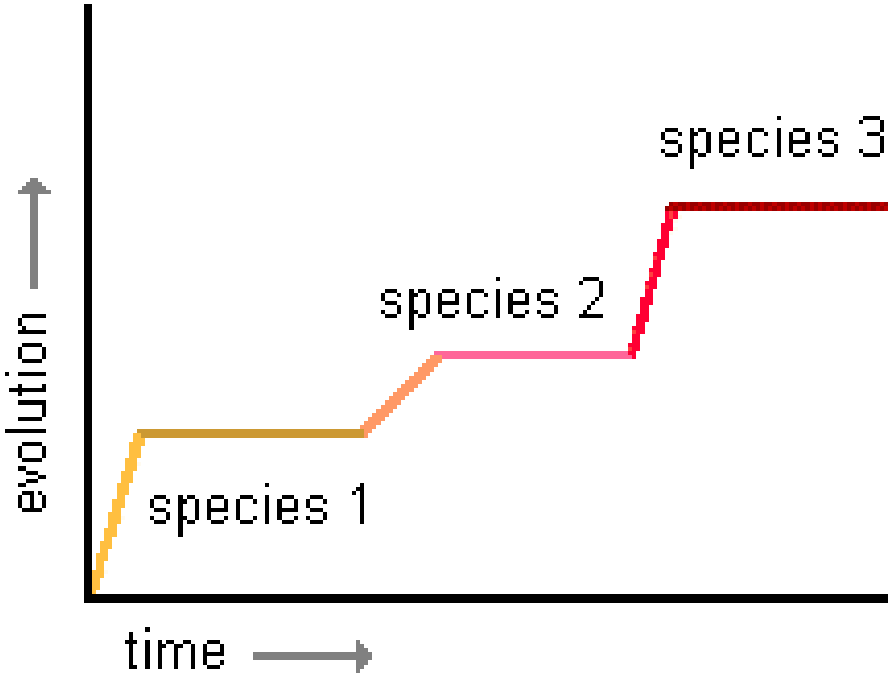
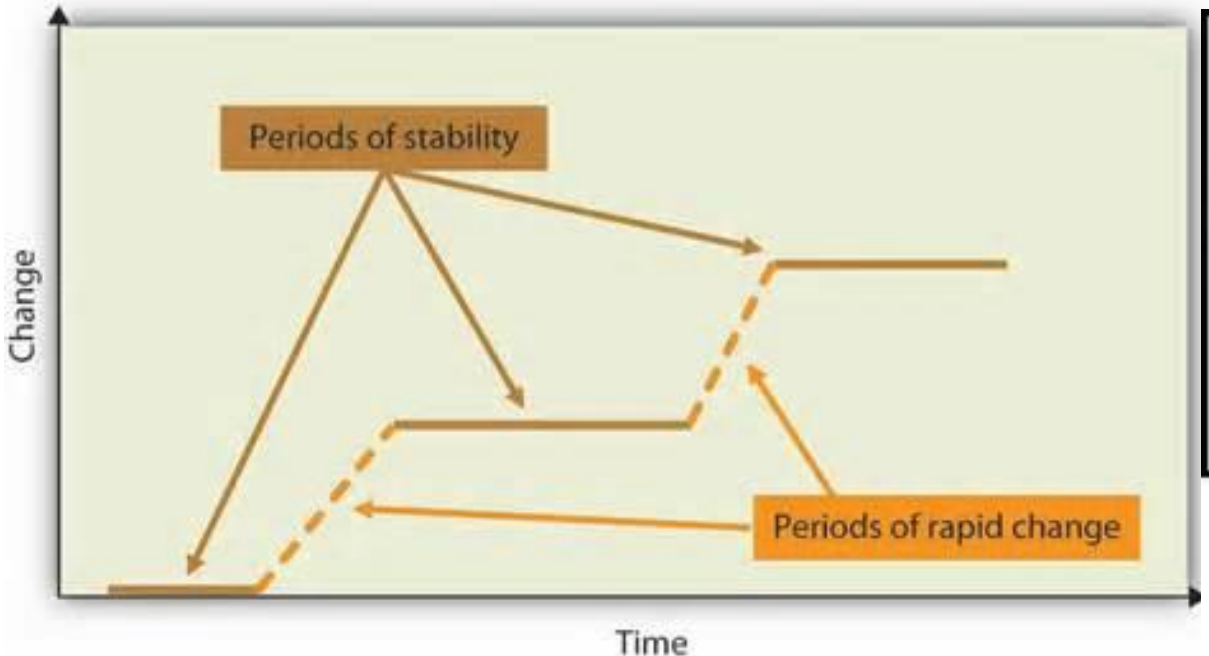


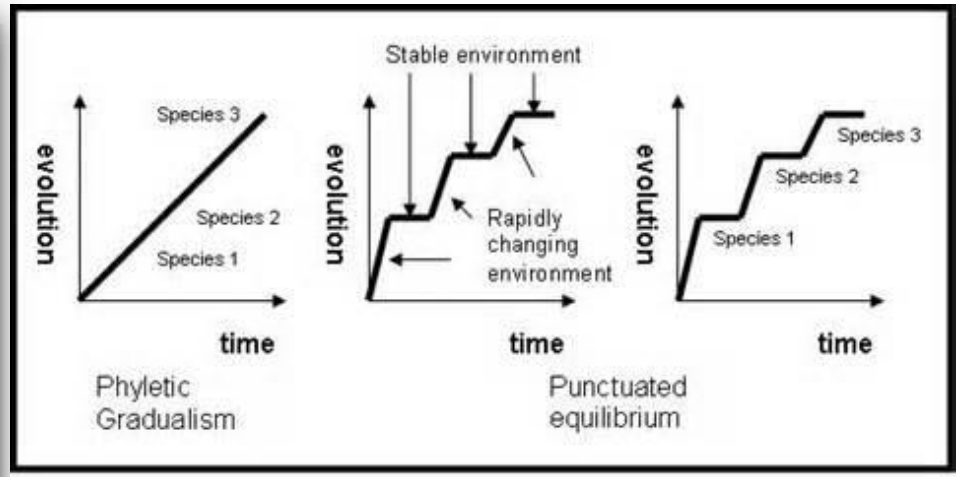
Fig. 3. Contrasting expectations of species-level evolution, the classic phyletic gradualism model (A), and the punctuated equilibrium model (B). Modified from various sources.

rubanko.myblog.it



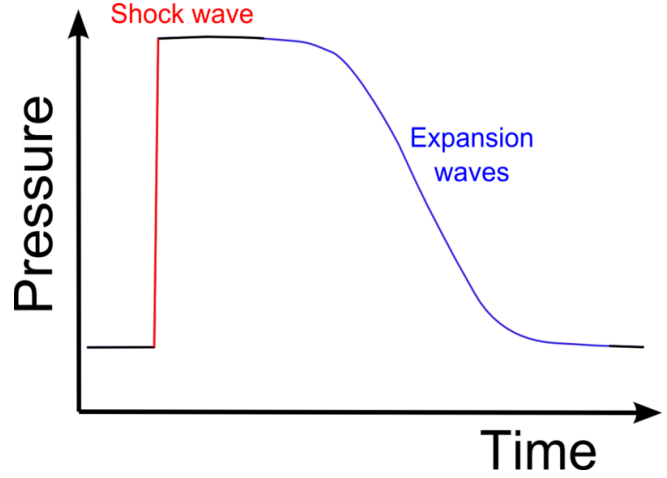
Beginning in the early 1970's, this model was challenged by **Stephen J. Gould, Niles Eldredge,** and a few other leading paleontologists . They asserted that there is sufficient fossil evidence to show that some species remained essentially the same for millions of years and then underwent short periods of very rapid, major change. Gould suggested that a more accurate model in such species lines would be **punctuated equilibrium** (illustrated by the graph on the left).



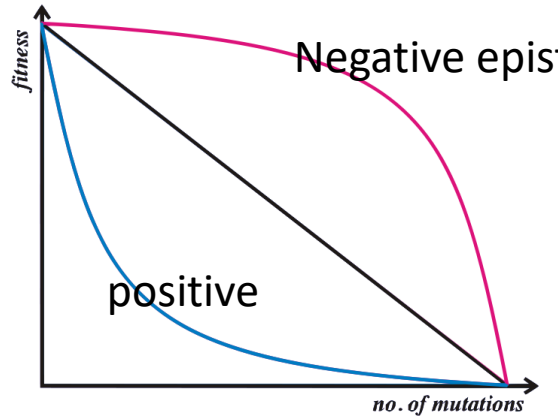


D. B. Saakian, M. H. Ghazaryan, and **Chin-Kun Hu**: Punctuated equilibrium and shock waves in the molecular model of biological evolution, PRE 90, 022712 (2014).

Abstract. We consider the dynamics in infinite population evolution models with a general symmetric fitness landscape. We find shock waves, i.e. discontinuous transitions in the mean fitness, in evolution dynamics even with smooth fitness landscapes which means that the search for the optimal evolution trajectory is more complicated. These shock waves appear in case of positive epistasis and can be used to represent punctuated equilibria in biological evolution during long geological time scale. We find exact analytical solutions for discontinuous dynamics at the large genome length limit and derive optimal mutation rates for fixed fitness landscape to send the population from the initial configuration to some final configuration in the fastest way.



Shock wave: (General Physics) a region across which there is a rapid pressure, temperature, and density rise.



Models of Asexual Biological Evolution

The genome configuration is specified by a sequence of L spin values $s_k = \pm 1$, $1 \leq k \leq L$,
purines: R=G, A $\rightarrow +1$, pyrimidines: Y=C, U, T $\rightarrow -1$.

There are $M = 2^L$ spin states, labeled by $0 \leq i \leq M-1$.

We denote the i-th genome configuration by $S_i \equiv (s_1^i, s_2^i, ..., s_L^i)$ and the probability of the i-th genome at time t is given by $p_{S_i} \equiv p_i(t)$ and the fitness r_i is the average number of offsprings per unit time. In our language, the chosen fitness r_i is a function f that operates on the genome configuration S_i, i.e., $r_i = f(S_i)$.

For the single-peaked fitness function, we take

$$f(S_0) = A >> 1, \text{ and } f(S_i) = 1 \text{ for } i \neq 1, \tag{1}$$

with $S_0 \equiv (+1, +1, ..., +1)$, which is equivalent to choosing

$$f(S_i) = 1 + (A-1)[\frac{\Sigma_{k=1}^{L} s_k^i}{L}]^p \quad \text{for} \quad p \rightarrow \infty. \tag{2}$$

Eigen Model: connected mutation-selection scheme

In the Eigen model, elements of the mutation matrix Q_{ij} represent probabilities that an offspring produced by state j changes to state i, and the evolution is given by the set of M coupled equations for M probabilities p_i of the i-th genome configuration (state):

$$\frac{dp_i}{dt} = \sum_{j=0}^{M-1} Q_{ij} r_j p_j - p_i \left(\sum_{j=0}^{M-1} r_j p_j \right). \tag{3}$$

The mutation matrix is

$$Q_{ij} = q^{L-d(i,j)} (1-q)^{d(i,j)}.$$

$d(i,j) \equiv (L - \Sigma_{k=1}^{L} s_k^i s_k^j)/2$ is the Hamming distance between S_i and S_j, $(1-q)$ is the mutation rate per site, and Q_{ij} satisfy

$$\sum_{i=0}^{M-1} Q_{ij} = 1 \quad \text{and} \quad \text{thus} \quad \sum_{i=0}^{M-1} p_i = 1$$

Error Threshold

The simplest choice of the fitness function f is the single peaked function, which corresponds to $r_0 = A$ and $r_0 = 1$ for $i > 0$. For this simple model, Eigen used qualitative arguments to find the error threshold without solving the model analytically:

$$A > \frac{1}{q^L} \approx e^{\gamma}.$$

Here $\gamma \approx L(1-q)$ =number of mutation per genome per replication.

Crow-Kimura Model: parallel mutation-selection scheme

$$\begin{aligned}\frac{dp_i}{dt} &= \sum_j m_{ij}p_j + r_i p_i - p_i \sum_j p_j r_j, \\ &\equiv \sum_j A_{ij}p_j - p_i \sum_j p_j r_j, \\ A_{ij} &= \delta_{ij} r_j + m_{ij}. \end{aligned} \tag{4}$$

One can subtract a constant from fitness function without changing equation.

Here m_{ij} is the mutation rate from state S_j to state S_i, and r_i is the fitness of state S_i. State S_i state S_j have a Hamming distance $d_{ij} = (L - \sum_k s_i^k s_j^k)/2$. We choose a mutation rate $m_{ij} = \gamma_f$ when $d_{i0} > d_{j0}$, $m_{ij} = \gamma_b$ when $d_{i0} < d_{j0}$, and $m_{ii} = -d_{i0}\gamma_b - (L - d_{i0})\gamma_f$. When $d_{ij} = 1$ $m_{ij} = \gamma_0$; and when $d_{ij} > 1$, $m_{ij} = 0$ [E. Baake, et al., PRL 78, 559 (1997)]. We can write the fitness r_i, the re-scaled mean fitness R, and the mean overlap x^*, respectively, as

$$r_i \equiv L f(s_1, \ldots, s_L),$$

$$LR \equiv \sum_i p_i r_i,$$

$$x^* \equiv 1 - 2\sum_i p_i d_{i0}/L.$$

Crow-Kimura model with asymmetric mutations

In the case that all the sequences with the same l number of " -1" alleles have the same probability, it is possible to write the following system of equations for relative class probabilities $P_l \equiv \Sigma_i \, p_i \delta(l, d_{i,0}), \; 0 \leq n \leq L$:

$$\begin{aligned} \frac{dP_l}{dt} &= P_l \left[L f\left(m_l\right) - (L-l)\gamma_f - l\gamma_b \right] \\ &+ \gamma_f \left(L - l + 1\right) P_{l-1} + \gamma_b \left(l+1\right) P_{l+1}, \end{aligned} \tag{5}$$

where $m_l = 1 - 2l/L$, and P_l are relative probabilities at the Hamming distance l (l mutations); $f(x)$ is a fitness function; and γ_f and γ_b are the forward and backward mutation rates, respectively. The forward means the the mutation increases the Hamming distance from the reference sequence, and backward mutation decreases the Hamming distance to the reference sequence. In Eq.(5), for $l = 0$ and $l = L$ we omit P_{-1} and P_{L+1}.

Hamilton-Jacobi equation (HJE)

At discrete values of overlap $x = 1 - 2l/L$ we use the ansatz:

$$P_l(t) \equiv P(x,t) \sim \exp[Lu(x,t)],$$

then Eq. (5) can be written as the HJE for $u \equiv \ln P(x,t)/L$:

$$\begin{aligned} &\frac{\partial u}{\partial t} + H(x,u') = 0, \\ &-H(x,p) = f(x) - \frac{\gamma_f(1+x)}{2} - \frac{\gamma_b(1-x)}{2} \\ &+\gamma_f \frac{1+x}{2} e^{2p} + \gamma_b \frac{1-x}{2} e^{-2p}. \end{aligned} \tag{6}$$

Here $p \equiv u' \equiv \partial u/\partial x$, the domain of x is $-1 \le x \le 1$, and the initial distribution is $u(x,0) = u_0(x)$. Let us denote the location for the maximum of distribution $P(x,t)$ as $x^*(t)$. Thus at $x^*(t)$, we have $p = u' = 0$. The differentiation of Eq.(6) with respect to x at $p = 0$ gives

$$-\frac{dx^*}{dt} = \frac{[f'(x^*(t))]}{u''(x^*(t),t)} + [(1+x^*)\gamma_f - (1-x^*)\gamma_b] \tag{7}$$

where $u'' \equiv \partial^2 u/(\partial x)^2$. We see that the dynamics of the maximum depends on the mutation rates, the fitness and the curvature of the distribution.

Minimizing $-H(x,p)$ via p, we obtain the expression of the evolution potential,

$$U(x) = f(x) + \sqrt{\gamma_b \gamma_f}\sqrt{1-x^2} - \gamma_f \frac{1+x}{2} - \gamma_b \frac{1-x}{2}. \tag{8}$$

The evolution behavior is defined by the evolution potential. The mean fitness R and the surplus s are defined as

$$R = max[U(x)]|_x, \quad f(s) = R. \tag{9}$$

In [D. B. Saakian, et al., Phys. Rev. E 78, 041908 (2008)], Eq.(6) was solved for $\gamma_f = \gamma_b = \gamma$ case by a method of characteristics. For the characteristics line $x(t)$ we have a Hamilton equation $dx/dt = dH(x,p)/dp$. Using the identity

$$\begin{aligned} & k \equiv \gamma_f \frac{1+x}{2} e^{2p} + \gamma_b \frac{1-x}{2} e^{-2p}, \\ & \gamma_f \frac{1+x}{2} e^{2p} - \gamma_b \frac{1-x}{2} e^{-2p} \\ & = \pm\sqrt{k^2 - \gamma_f \gamma_b (1-x^2)}, \end{aligned} \tag{10}$$

and the Hamilton equation with the Hamiltonian H given by Eq. (6), we obtain:

$$\begin{aligned} \dot{x} &= \pm 2\sqrt{k^2 - \gamma_f \gamma_b (1-x^2)}, \\ k &= q + \gamma_f \frac{1-x}{2} + \gamma_b \frac{1+x}{2} - f(x), \end{aligned} \tag{11}$$

where $q \equiv \partial u(x,t)/\partial t$ is constant along the characteristics, like the particle energy in classical mechanics. At every point we have two characteristics, moving to the right and left.

We consider the dynamics of the population, initially having a fixed overlap x_0 with the reference (master) sequence. Let us look at the manner of change in the mean overlap of the population $x^*(t^*) = \sum_j P_j(1-2d_{i,0}/L)$ at the moment in time t^*. As time progresses, the overlap distribution spreads out and we therefore focus on the time evolution of the overlap x^* which yields the maximum of this distribution.

Integrating $dt = (dt/dx)dx$, we obtain for the large initial x_0

$$t^* = \frac{1}{2}|\int_{x_0}^{x^*} d\xi \ \ [F(\gamma, x^*, \xi)]^{-1/2}|, \tag{12}$$

where we have the following expression for F:

$$\begin{aligned} F(\gamma, x^*, \xi) &= \left[f(x^*) + \gamma_f \frac{1+\xi}{2} + \gamma_b \frac{1-\xi}{2} - f(\xi)\right]^2 \\ &- \gamma_f \gamma_b \left(1-\xi^2\right). \end{aligned} \tag{13}$$

Equation (12) corresponds to the motion along one characteristics.

For the small x_0 we should consider the motion along two characteristics: after the point x_1 we should take the characteristics with sign $-$ in Eq. (12). We get:

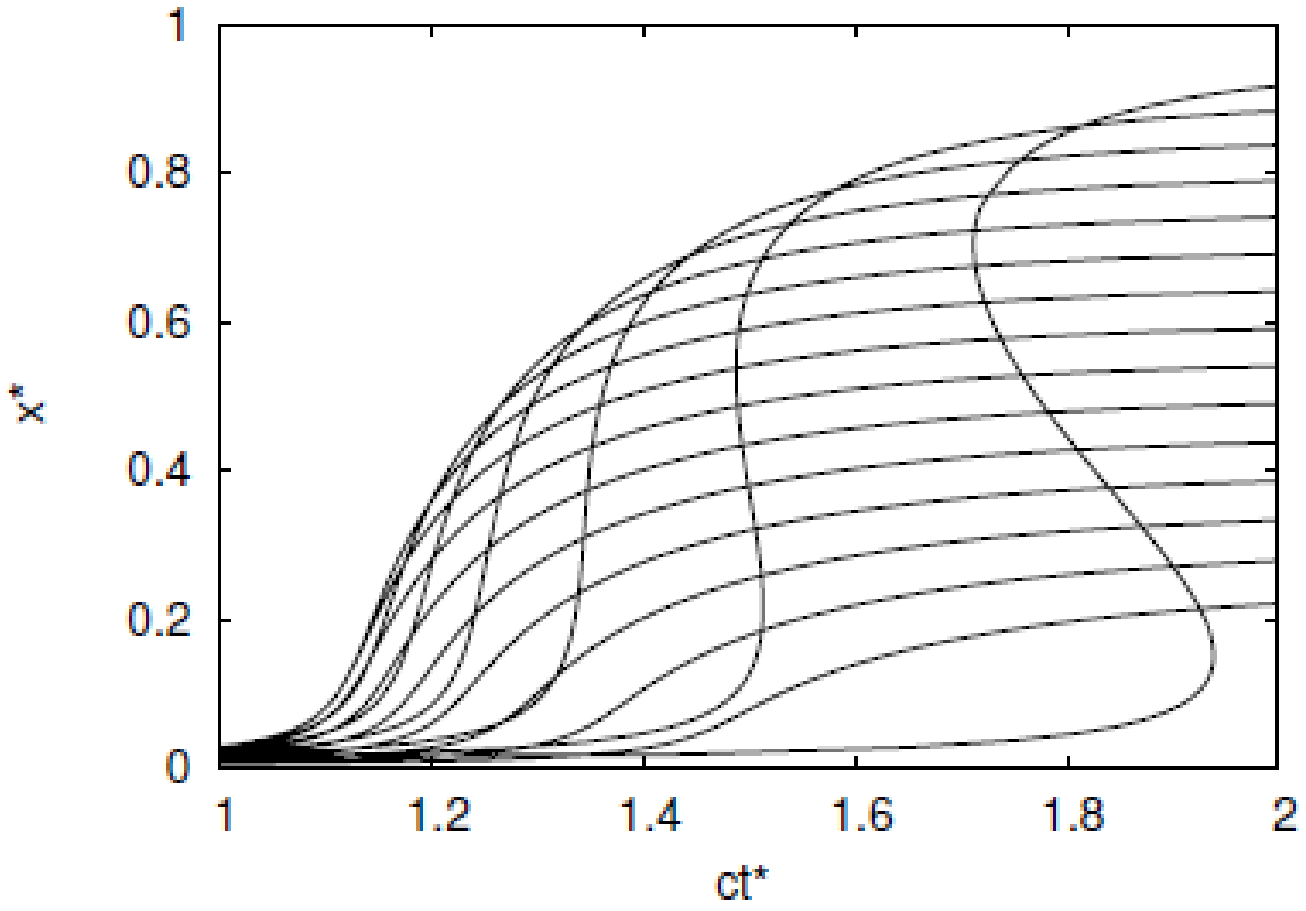
$$t^* = \frac{1}{2}\int_{x_0}^{x_1} \frac{d\xi}{\sqrt{F(\gamma, x^*, \xi)}} + \frac{1}{2}\int_{x^*}^{x_1} \frac{d\xi}{\sqrt{F(\gamma, x^*, \xi)}} \tag{14}$$

and x_1 is the solution of

$$F(\gamma, x^*, x_1) = 0. \tag{15}$$

For a symmetric mutation scheme $\gamma_f = \gamma_b = \gamma$, F is replaced by with F_s:

$$F_s(\gamma, x^*, \xi) = [f(x^*) + \gamma - f(\xi)]^2 - \gamma^2\left(1 - \xi^2\right). \tag{16}$$

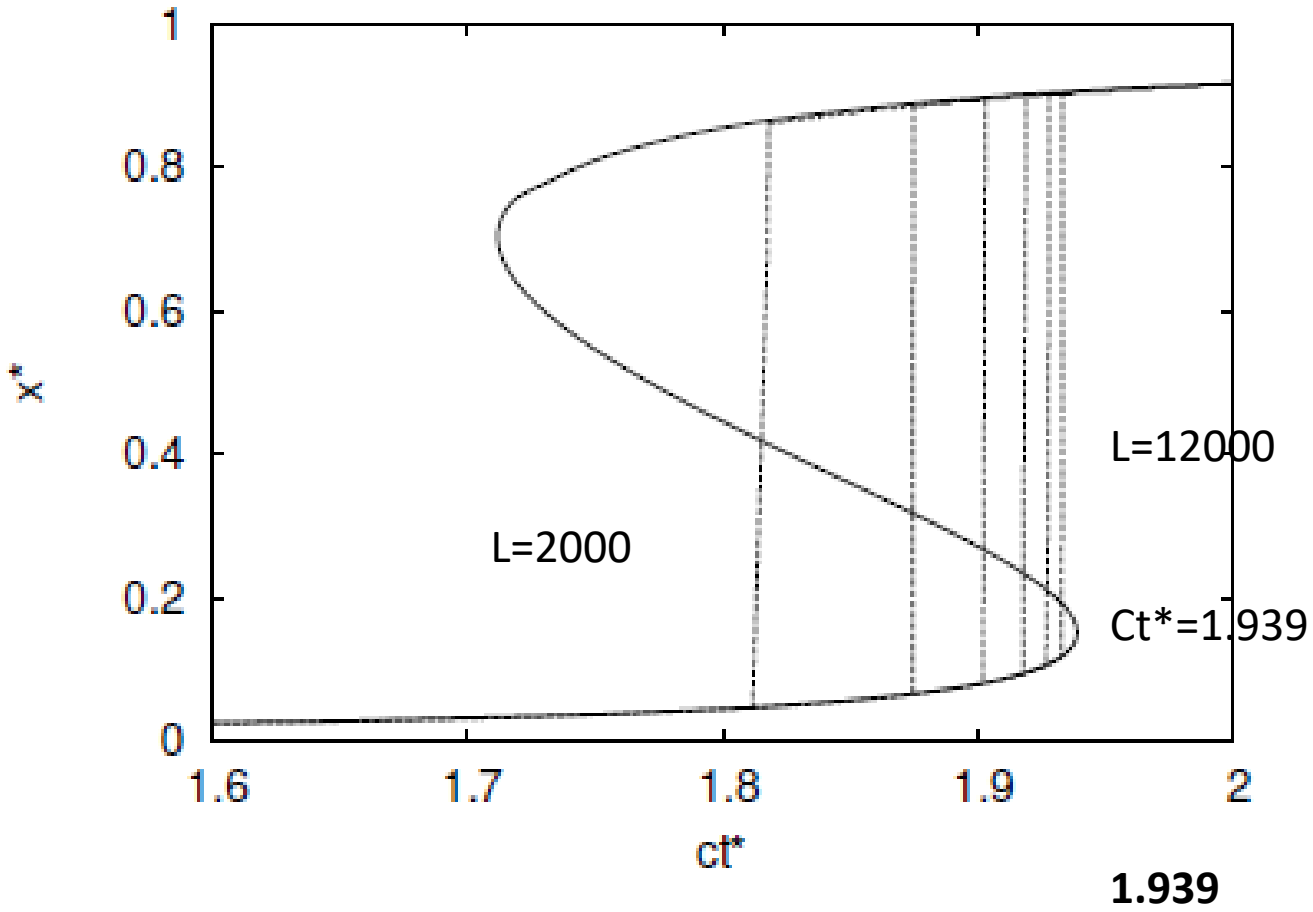


Top curve: gamma/c=0.05=1/20, c=20
2nd curve: gamma/c=0.10=1/10, c=10

$$f(x) = \frac{c}{2}x^2$$

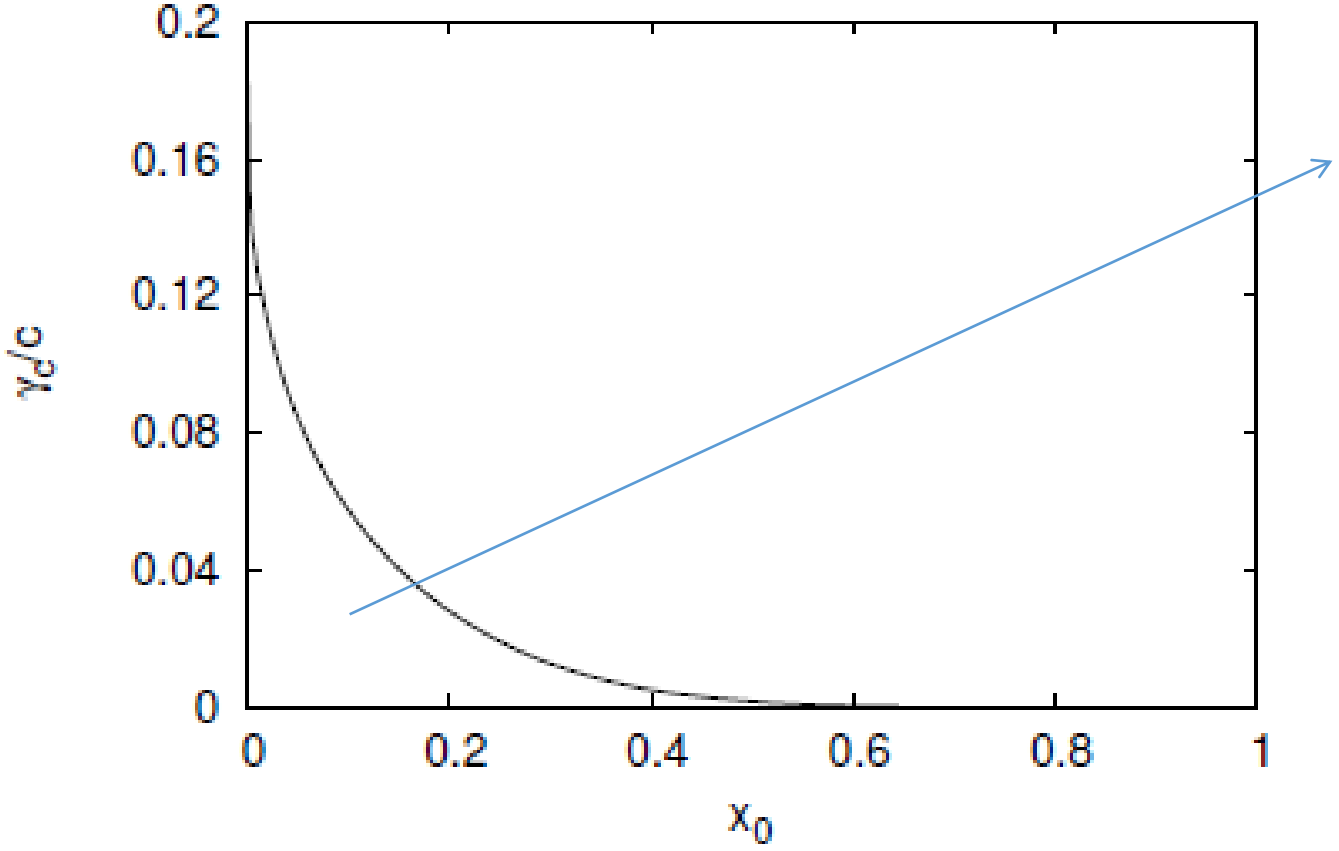
C=1/0.7=1.42…, x*-→1-0.7=0.30
C=1/0.75=1.333…., x*→ 1-0.75=0.25.

FIG. 1. The dynamics of $x^*(t^*)$ (mean overlap with the reference sequence) as function of time t^*) by Eqs. (12) and (14) for the symmetric case $\gamma_f = \gamma_b = \gamma = 1$ and (top to bottom at $ct^* = 2$) $\gamma/c = 0.05, 0.1, \ldots, 0.7, 0.75$. The initial population has overlap $x_0 = 0.01$ with the reference sequence. For $t^* \to \infty$, we find $x^* = 1 - \gamma/c$.



$$f(x) = \frac{c}{2}x^2$$

FIG. 2: The dynamics of $x^*(t^*)$ by Eq. (1) for $x_0 = 0.01$, $\gamma/c = 0.05$ and (dashed vertical lines from left to right) $L = 2000, 4000, \ldots, 12000$. For $L \to \infty$ the jump in x^* takes place at $ct^* = ct_d^* = 1.939$ and has size $\Delta x^* = 0.755$.



Region of discontinuous transition

$$f(x) = \frac{c}{2}x^2$$

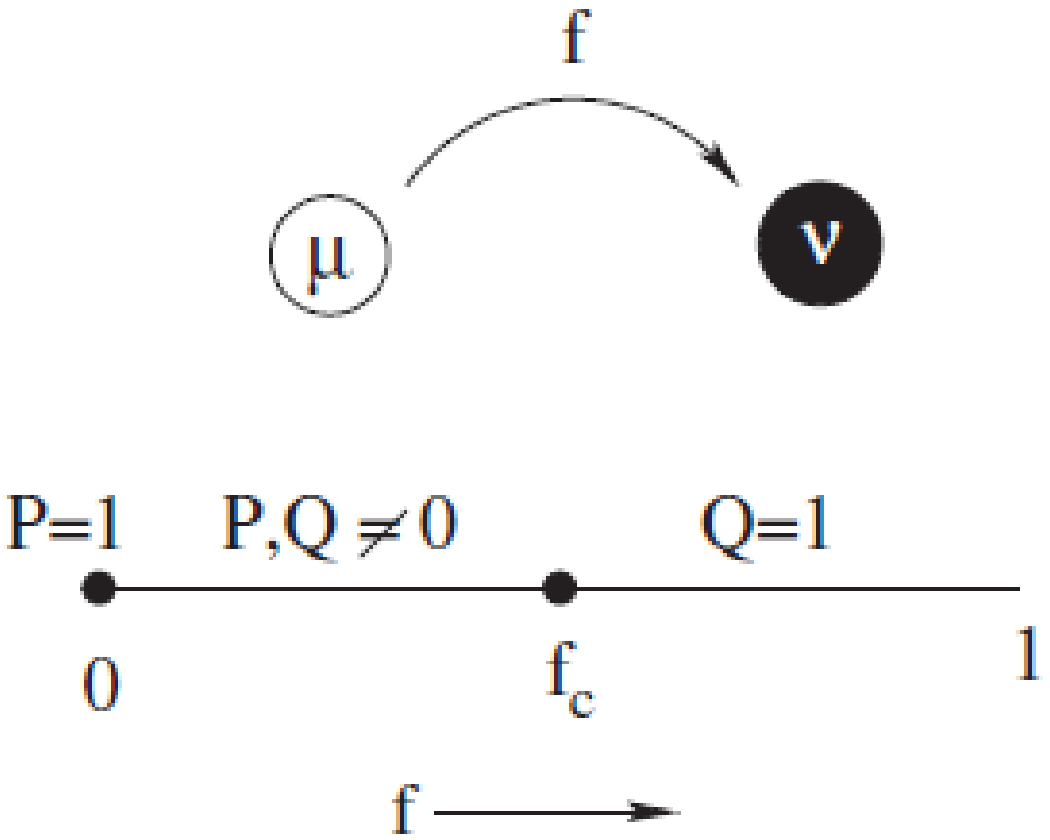
FIG. 3: The critical line γ_c/c vs. x_0 at which $\Delta x^* = 0$. The critical c is defined from the system of equations $dx^*/dt^* = 0, d^2x^*/d^2t^* = 0$. For the γ values below this curve, the most probable overlap x^* undergoes a discontinuous transition at $t^* = t_d^*$ (see Fig. 2).

Mutator gene, e.g. p53

Medical Dictionary: mutator gene
Evolutionary biology:
A gene which, under severe physiologic stress, spreads through populations of pathogens (病原體) or tumour cells.

Molecular biology:
A gene that increases the mutation frequency of other genes; DNA repair genes typically have a mutator phenotype.

Merriam Webster Online: a gene that increases the rate of mutation of one or more other genes—called also *mutator* .



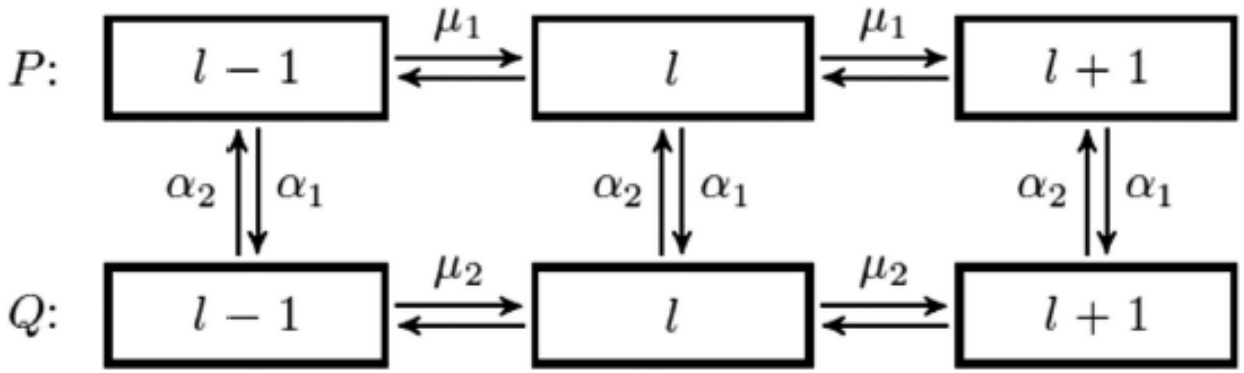
A. Nagar, K. Jain: Exact phase diagram of quasispecies model with a mutator rate modifier, Phys. Rev. Lett. 102, 038101 (2009).

FIG. 1: Schematic phase diagram of the quasispecies model with mutation rate modifier. With probability f, the nonmutators (mutation probability μ) change to mutators (mutation probability nu). The pure nonmutator phase occurs when f = 0 and pure mutator phase for f >= fc. The system is in the mixed phase for 0 < f < fc.

The rich phase structure of a mutator model

Scientific Reports 6, 34840 (2016)

David B. Saakian[1,2], Tatiana Yakushkina[3] & Chin-Kun Hu[1,4,5]



$$\frac{d\hat{P}_l(t)}{Ldt} = \alpha_2\hat{Q}_l + \hat{P}_l(f(x_l) - (\mu_1 + \alpha_1)) + \mu_1\left(\hat{P}_{l-1}\frac{L-l+1}{L} + \hat{P}_{l+1}\frac{l+1}{L}\right) - \hat{P}_l R(t),$$

$$\frac{d\hat{Q}_l(t)}{Ldt} = \alpha_1\hat{P}_l + \hat{Q}_l(g(x_l) - (\mu_2 + \alpha_2)) + \mu_2\left(\hat{Q}_{l-1}\frac{L-l+1}{L} + \hat{Q}_{l+1}\frac{l+1}{L}\right) - \hat{Q}_l R(t),$$

$$R(t) = \sum_l (\hat{P}_l(t)f(x_l) + \hat{Q}_l(t)g(x_l)),$$

$$x_l = 1 - 2l/L, \quad 0 \le l \le L.$$

$$\hat{P}_l = \frac{P_l}{\sum_l (P_l + Q_l)}, \hat{Q}_l = \frac{Q_l}{\sum_l (P_l + Q_l)}. \tag{2}$$

It is easy to show that P_l and Q_l are the solutions of the following system of linear equations:

$$\begin{aligned} \frac{dP_l(t)}{Ldt} &= \alpha_2 Q_l + P_l(f(x_l) - (\mu_1 + \alpha_1)) + \mu_1 \left(P_{l-1} \frac{L - l + 1}{L} + P_{l+1} \frac{l+1}{N} \right), \\ \frac{dQ_l(t)}{Ldt} &= \alpha_1 P_l + Q_l(g(x_l) - (\mu_2 + \alpha_2)) + \mu_2 \left(Q_{l-1} \frac{L - l + 1}{L} + Q_{l+1} \frac{l+1}{L} \right). \end{aligned} \tag{3}$$

In Eqs (1) and (2), $\hat{P}_l$ and $\hat{Q}_l$ are the probabilities under the normalization constraint, while P_l and Q_l of Eq. (3) are not normalized.

In this paper, we focus on the following characteristics of the model in a steady state: the mean fitness R, the total surplus s (the expected value of x_l), the surplus for the wild-type part of the population s_1 and for the mutator-type part s_2, and the fraction of the mutator sub-population q:

$$R = \frac{\sum_l (P_l f(x) + Q_l g(x))}{\sum_l (P_l + Q_l)}, s = \frac{\sum_l (P_l + Q_l) x_l}{\sum_l (P_l + Q_l)},$$
$$s_1 = \frac{\sum_l P_l x_l}{\sum_l P_l}, s_2 = \frac{\sum_l Q_l x_l}{\sum_l Q_l}, q = \frac{\sum_l Q_l}{\sum_l (P_l + Q_l)}.$$

Assume a smooth distribution, we obtain

$$R = qg(s_2) + (1 - q) f(s_1). \tag{4}$$

The case of forward and backward transitions of the mutator gene.

Consider Eq.(1) without nonlinear terms. We assume the following ansatz [24, 26, 29]:

$$P_l(t) = v_1(x,t)e^{Nu(x,t)}; \quad Q_l(t) = v_2(x,t)e^{Nu(x,t)}. \quad (4)$$

Our system of equation is mapped into the HJE and in the Methods we obtain the following expression for the mean fitness R using the potential function [26]:

$$\begin{aligned} R &= Max[V_+(x)]|_x, \\ V_\pm(x) &= \frac{f(x)+g(x)}{2} - \frac{1+\mu}{2} - \frac{\alpha_1+\alpha_2}{2} \\ &+ \frac{1+\mu}{2}\sqrt{1-x^2} \pm \frac{1}{2}\sqrt{A(x)^2+4\alpha_1\alpha_2} \end{aligned} \quad (5)$$

where we used $\mu_1 = 1,\ \mu_2 = \mu$, and $A(x) = f(x) - \alpha_1 + \alpha_2 - g(x) + (1-\mu)(\sqrt{1-x^2}-1)$. The numerics in support well our analytical results, see Methods.

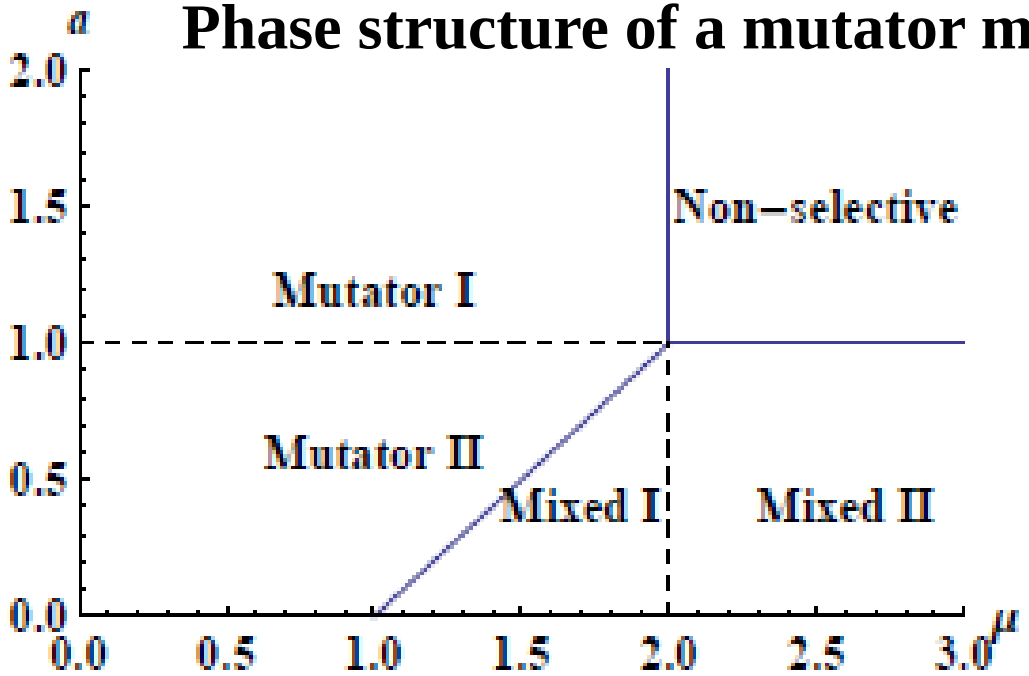
TABLE I: The comparison of the results for $f(x) = g(x) = 3x^2/2, \mu_1 = 1, \mu_2 = \mu, \alpha_1 = \alpha_2 = 1, N = 400$. R_{num} is the numerical result and R_{th} is given by Eq.(5).

μ	3.5	3.	2.5	2.	1.5	1.	0.5
R_{num}	0.1907	0.2514	0.32405	0.4127	0.5240	0.6684	0.8626
R_{th}	0.1811	0.2436	0.3180	0.4084	0.5212	0.6666	0.8615

TABLE II: The results for $f(x) = g(x) = kx, \mu_1 = 1, \mu_2 = \mu, \alpha_1 = \alpha, \alpha_2 = 0$. $K = \exp[N(\int_{s_1}^{s_3} u'(x) - N\int_{s_2}^{s_3} \bar{u}'(x))]$. R_n is a numerical result.

N	1000	1000	1000	1000
k	0.3	0.3	0.3	1
α	0.0001	0.001	0.01	0.3
R_n	0.0439	0.0430	0.0340	0.1142
R	0.0439	0.0430	0.0340	0.1142
q	0.9945	0.9460	0.530	$6/10^7$
K	1-$4/10^6$	0.9994	0.930	$1/10^5$

Phase structure of a mutator model 癌增變模型的相圖



$$\alpha_1 = a,\ \alpha_2 = 0,\ \mu_1 = 1,$$

Single-peak fitness function

f(x=1)=g(x=1)=J

f(x<1)=g(x<1)=0

$$\mu_2 = \mu$$

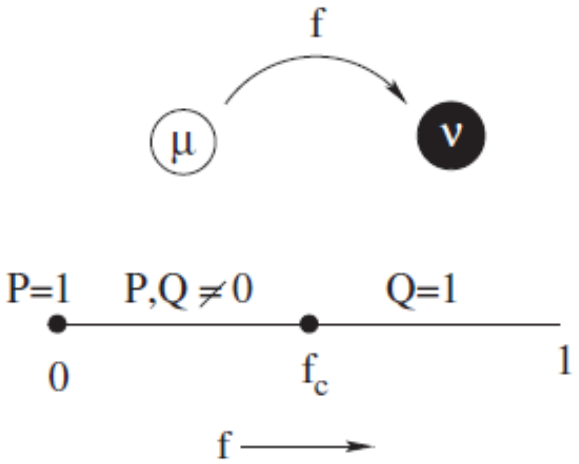


Figure 1. The phase structure of the mutator model with single peak landscape with zero value for any argument except $l=0$: $f(x=1)=g(x=1)=J$. The parameters of the model are $\alpha_1=a$, $\alpha_2=0$, $\mu_1=1$, and $\mu_2=\mu$. There are three phases: mixed phase with $0<s<1$, $0<q<1$, non-selective phase with $s=0$, $0<q\leq 1$ and mutator phase with $0<s,q=1$. The border between non-selective and mutator phases is given by $\mu=J$, the border between non-selective and mixed phases is given by $a=J-1$, between mixed and mutator phases is given by $a+1=\mu$ line. From the bio-medical perspectives we distinguish the mutator I and mutator II, mixed I and mixed II subphases. From the mutator I, the system transforms to the non-selective phase simply increasing the μ. From the mixed II the system transformers to the non-selective phase simply increasing the $a\equiv\alpha_1$. From the mutator II and mixed I subphases we need change both a and μ to transform the system to the non-selective phase. We have the same picture in case of Eigen model with a fitness A for the peak sequence and fitness 1 for other sequences. We have for the mixed phase $R=e^{-(h+\gamma)}A$, for the mutator phase $R=e^{-\mu\gamma}A$ and for the non-selective phase $R=1$. The border between non-selective and mutator phases is given by $\mu\gamma=\ln A$, the border between non-selective and mixed phases is given by $h=\ln A-\gamma$, between mixed and mutator phases is given by $h=(\mu-1)\gamma$ line.

A. Nagar, K. Jain PRL 2009

Figure 1. The phase structure of the mutator model with single peak landscape with zero value for any argument except $l=0$: $f(x=1)=g(x=1)=J$. The parameters of the model are $\alpha_1=a$, $\alpha_2=0$, $\mu_1=1$, and $\mu_2=\mu$. There are three phases: mixed phase with $0<s<1$, $0<q<1$, non-selective phase with $s=0$, $0<q\leq 1$ and mutator phase with $0<s$, $q=1$. The border between non-selective and mutator phases is given by $\mu=J$, the border between non-selective and mixed phases is given by $a=J-1$, between mixed and mutator phases is given by $a+1=\mu$ line. From the bio-medical perspectives we distinguish the mutator I and mutator II, mixed I and mixed II subphases. From the mutator I, the system transforms to the non-selective phase simply increasing the μ. From the mixed II the system transformers to the non-selective phase simply increasing the $a\equiv\alpha_1$. From the mutator II and mixed I subphases we need change both a and μ to transform the system to the non-selective phase. We have the same picture in case of Eigen model with a fitness A for the peak sequence and fitness 1 for other sequences. We have for the mixed phase $R=e^{-(h+\gamma)}A$, for the mutator phase $R=e^{-\mu\gamma}A$ and for the non-selective phase $R=1$. The border between non-selective and mutator phases is given by $\mu\gamma=\ln A$, the border between non-selective and mixed phases is given by $h=\ln A-\gamma$, between mixed and mutator phases is given by $h=(\mu-1)\gamma$ line.

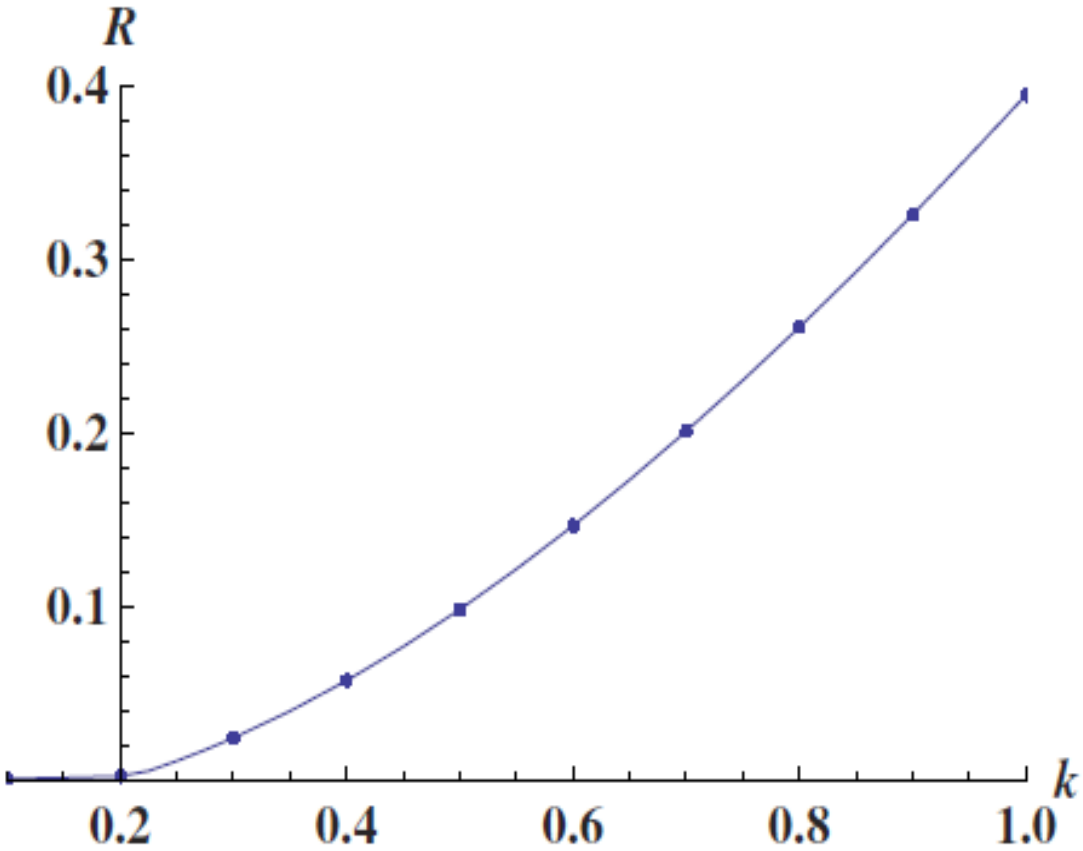
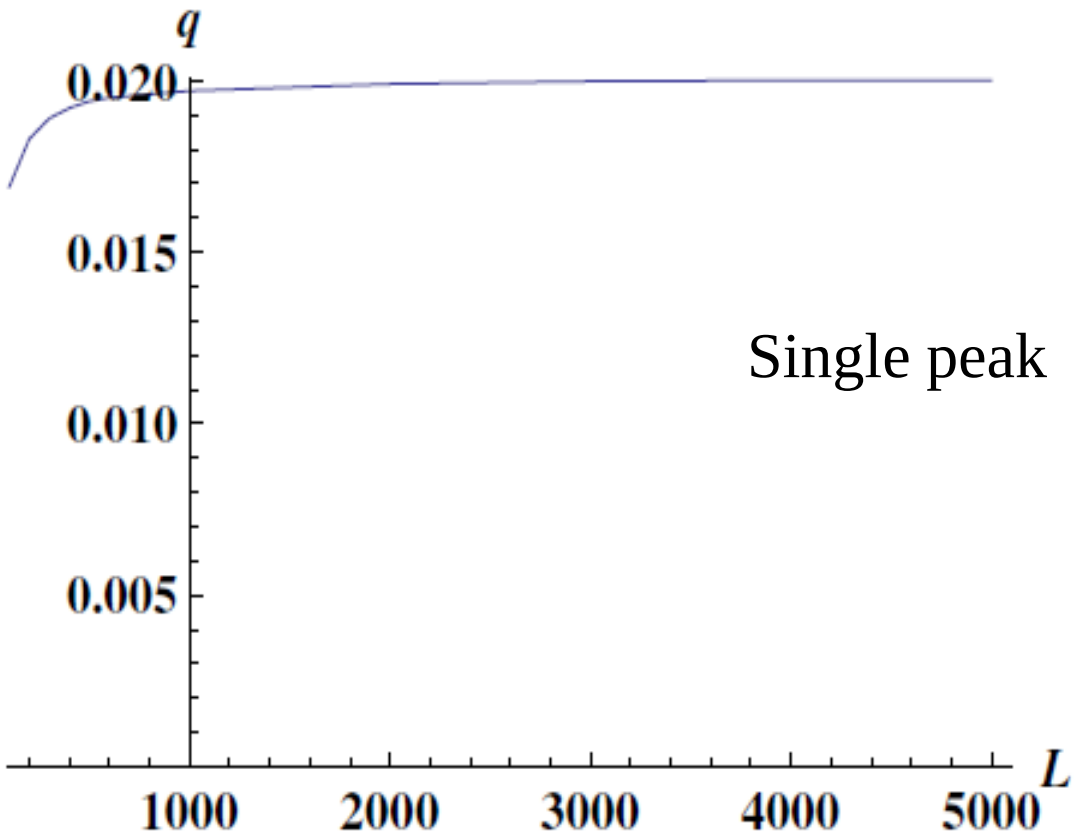


Figure 3. The mean fitness R versus k of the model with linear fitness landscape $f(x) = kx$, $\mu_1 = 1$, $\mu_2 = 10$, $\alpha_2 = 0$, $\alpha_1 = 0.02$. There are two phases in the model for the general values of parameters: mixed phase with $R = \sqrt{k^2 + 1} - \alpha_1 - 1$ and mutator phase with $R = \sqrt{k^2 + (\mu_2)^2} - \mu_2$, $q = 1$. The border between two phases is given by equation $\alpha_1 = \mu_2 - 1 + \sqrt{k^2 + 1} - \sqrt{k^2 + (\mu_2)^2}$. In our case $k_c \approx 0.212$.



Single peak

Figure 4. The dependence of the mutator probability q on the genome length L. The the single peak fitness model (smooth line) with $J = 1.05$, $\mu_1 = 1$, $\mu_2 = 10$, $a = \alpha_1 = 0.001$. For the $L = 5000$ the single peak model's numerical result coincides with the analytical result for $L = \infty$ with the relative accuracy about 0.1%.

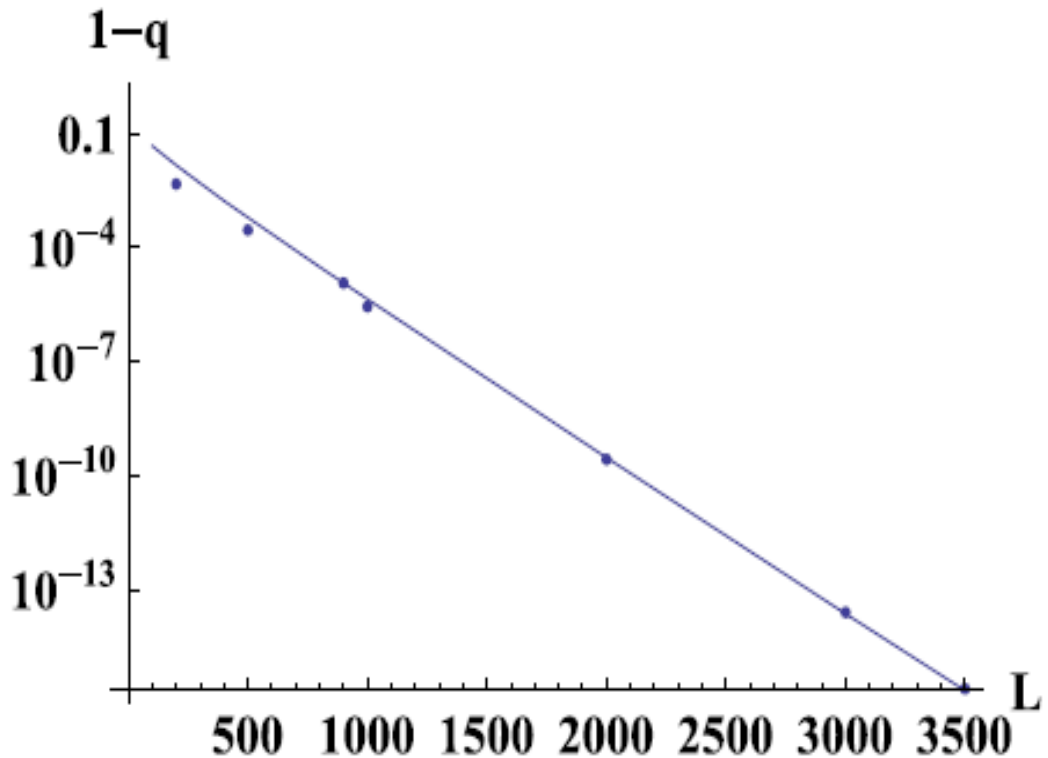


Figure 5. The dependence of the mutator probability q on the genome length L for the linear fitness model: $f(x) = kx, k = 1, \mu_1 = 1, \mu_2 = 10, \alpha_2 = 0, a = \alpha_1 = 0.3$. The smooth line corresponds to the numerics, the solid dots correspond to our analytical result.

A Stochastic Single-Molecule Event Triggers Phenotype Switching of a Bacterial Cell

Science 322, 442 (2008)

Paul J. Choi,* Long Cai,*† Kirsten Frieda,‡ X. Sunney Xie§

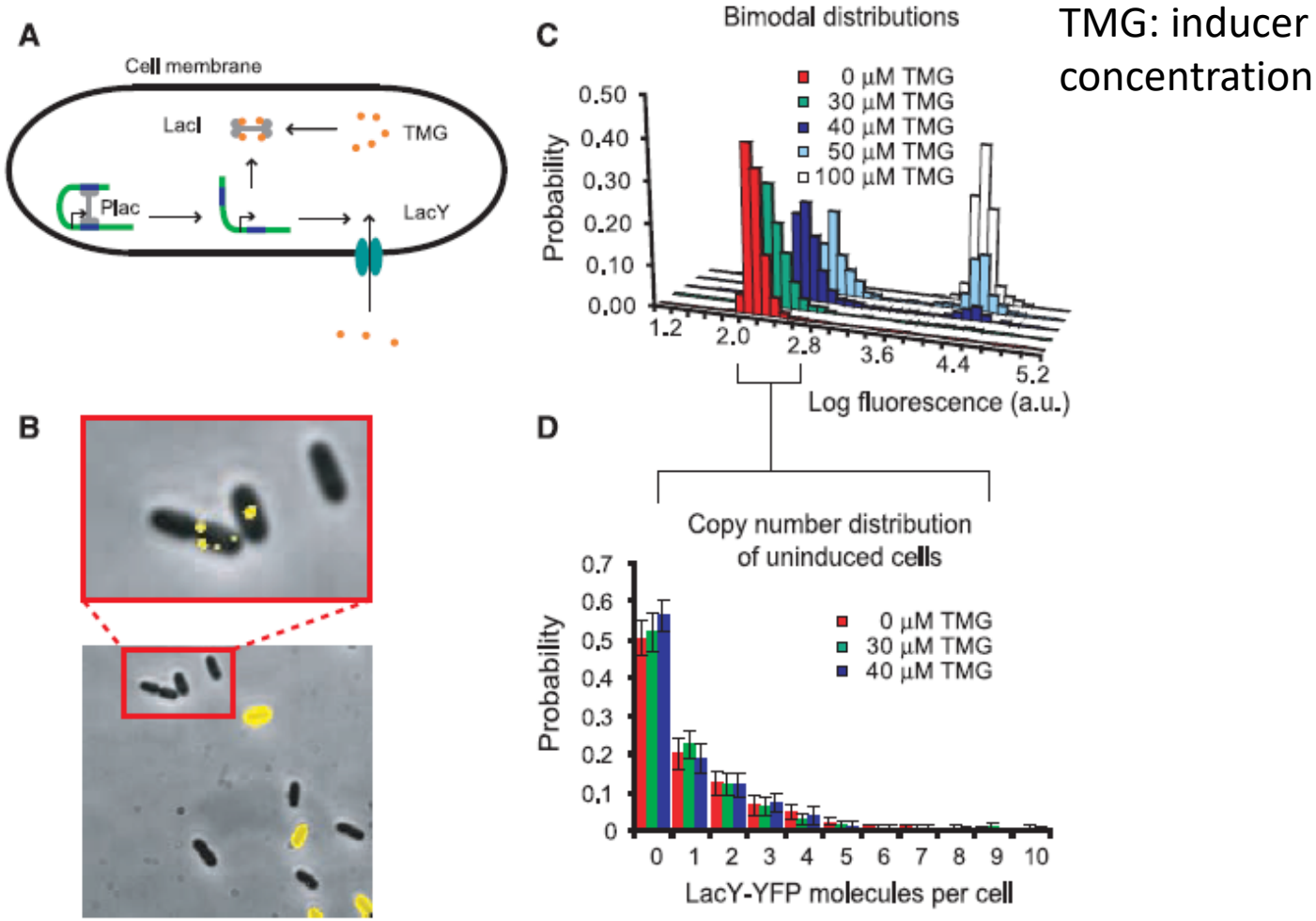


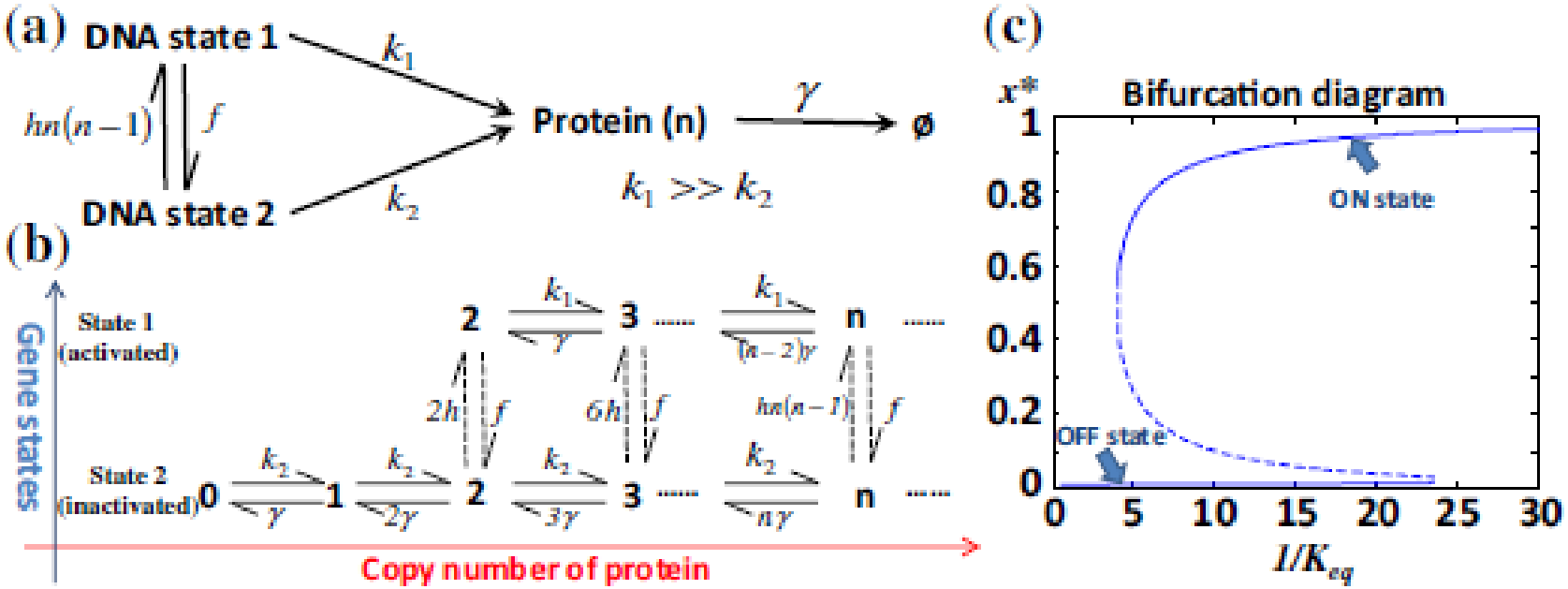
Fig. 1. The expression of lactose permease in *E. coli*. (**A**) The repressor LacI and permease LacY form a positive feedback loop. Expression of permease increases the intracellular concentration of the inducer TMG, which causes dissociation of LacI from the promoter, leading to even more expression of permeases. Cells with a sufficient number of permeases will quickly reach a state of full induction, whereas cells with too few permeases will stay uninduced. (**B**) After 24 hours in M9 medium containing 30 μM TMG, strain SX700 expressing a LacY-YFP fusion exhibits all-or-none fluorescence in a fluorescence-phase contrast overlay (bottom, image dimensions 31 μm × 31 μm). Fluorescence imaging with high sensitivity reveals single molecules of permease in the uninduced cells (top, image dimensions 8 μm × 13 μm). (**C**) After 1 day of continuous growth in medium containing 0 to 50 μM TMG, the resulting bimodal fluorescence distributions show that a fraction of the population exists either in an uninduced or induced state, with the relative fractions depending on the TMG concentration. (**D**) The distributions of LacY-YFP molecules in the uninduced fraction of the bimodal population at different TMG concentrations, measured with single-molecule sensitivity, indicate that one permease molecule is not enough to induce the *lac* operon, as previously hypothesized (*12*). More than 100 cells were analyzed at each concentration. Error bars are SE determined by bootstrapping.

TMG: inducer concentration

Stochastic Phenotype Transition of a Single Cell in an Intermediate Region of Gene State Switching

PRL 114, 078101 (2015)

Hao Ge,[1,2,*] Hong Qian,[3] and X. Sunney Xie[1,4,†]



$$\frac{dx}{dt} = g(x) - \gamma x, \qquad (2)$$

FIG. 1 (color online). (a) A minimal gene network with positive feedback and two different gene states. (b) The diagram of the full chemical master equation [see Eq. (1)]. (c) Deterministic mean-field model [Eq. (2)] with bistability induced by positive feedback, in which $k_1 = 10\ \text{min}^{-1}$, $k_2 = 0.1\ \text{min}^{-1}$, and $\gamma = 0.02\ \text{min}^{-1}$.

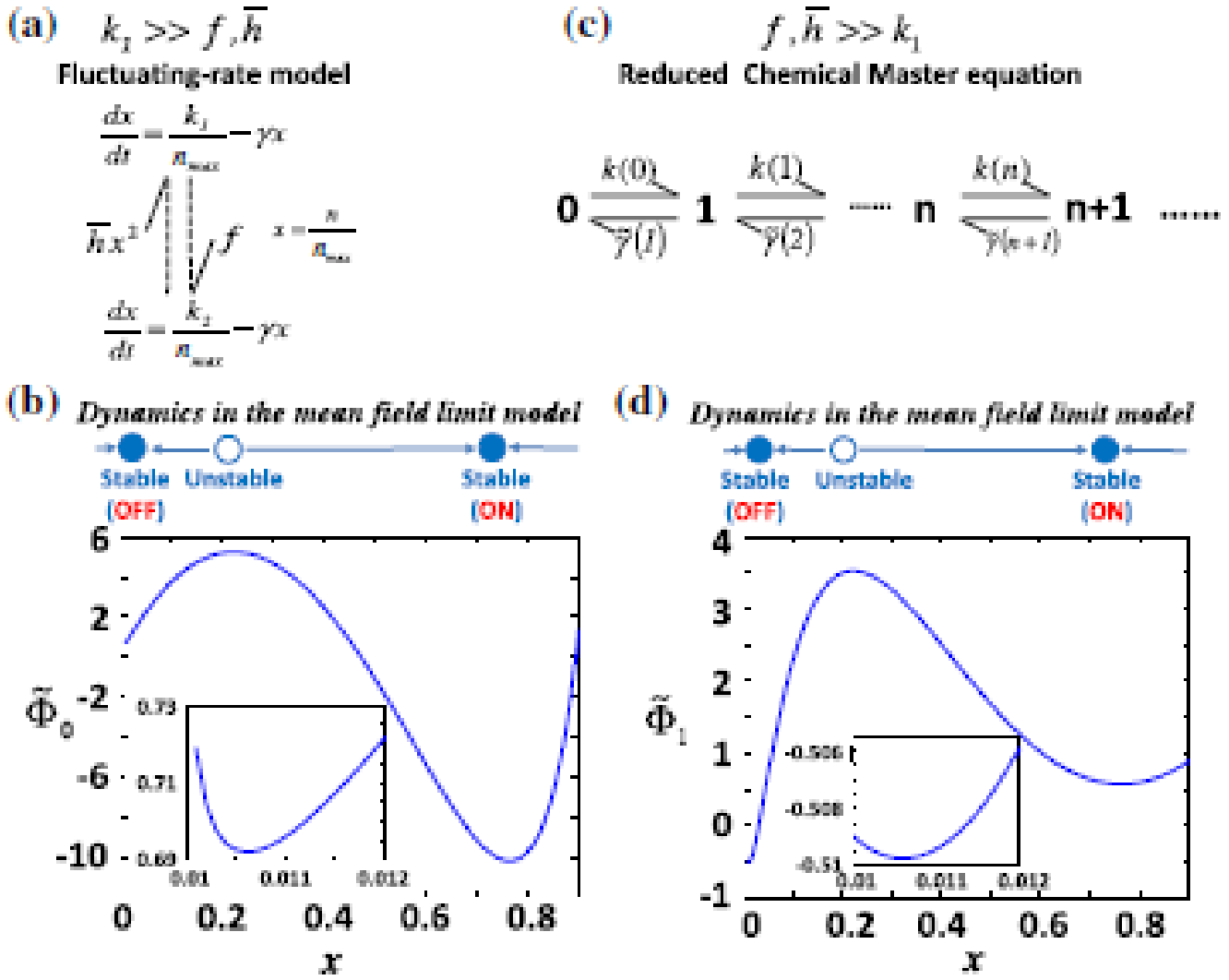


Figure 2. In the intermediate region of gene state switching, the full CME can be simplified to a fluctuating-rate model (a), the steady-state distribution of which corresponds to a normalized landscape function ˜Φ0(x) (b). In the case if the gene state switching is extremely rapid, a reduced CME (c) and a different landscape function ˜Φ1(x) (d) can also be derived. The insets in (b) and (d) are the zoom in of the functions near x=0 . The parameters in (b) and (d) are the same as those in Fig. 1 with Keq=1/5.5 .

For the condition (A): $N \equiv k_1/\gamma \gg 1$, $k_1 \gg k_2$ with parameters $\bar{h} \equiv N^2 h$, $hn(n-1) = \bar{h}x(x-1/N)$, and $x = n/N$; GQX announced that the model has been solved in the two cases [34], I: $k_1 \gg \bar{h}, f$ and II: $k_1 \ll \bar{h}, f$. In case I, Eq. (1) can be simplified as the fluctuating rate model described by (a)

PHYSICAL REVIEW E 97, 012412 (2018)

Accurate analytic solution of chemical master equations for gene regulation networks in a single cell

Guan-Rong Huang,[1] David B. Saakian,[2,3,4,*] and Chin-Kun Hu[1,4,5,6,†]

[1]*Physics Division, National Center for Theoretical Sciences, Hsinchu 30013, Taiwan*

[2]*Theoretical Physics Research Group, Advanced Institute of Materials Science, Ton Duc Thang University, Ho Chi Minh City, Vietnam*

[3]*Faculty of Applied Sciences, Ton Duc Thang University, Ho Chi Minh City, Vietnam*

[4]*Institute of Physics, Academia Sinica, Nankang, Taipei 11529, Taiwan*

[5]*Department of Systems Science, University of Shanghai for Science and Technology, Shanghai 200093, China*

[6]*Department of Physics, National Dong Hwa University, Hualien 97401, Taiwan*

$$D_1 \xrightarrow{k_1} D_1 + Pr\,,\quad D_2 \xrightarrow{k_2} D_2 + Pr\,,$$

$$Pr \xrightarrow{\gamma} \emptyset\,,\quad D_2 + 2Pr \xrightarrow{h} D_1\,,\quad D_1 \xrightarrow{f} D_2 + 2Pr\,.$$

FIG. 1. Reaction scheme of GQX's model of gene regulation network, where D_1 and D_2 are the DNA at the gene state 1 and 2, respectively, and Pr is the protein.

$$
\begin{aligned}
\frac{\partial p_1(n,t)}{\partial t} &= k_1 \ [p_1(n-1,t) - p_1(n,t)] \\
&+ \ \gamma[(n-1)p_1(n+1,t) - (n-2)p_1(n,t)] \\
&+ \ hn(n-1)p_2(n,t) - fp_1(n,t), \\
\frac{\partial p_2(n,t)}{\partial t} &= k_2[p_2(n-1,t) - p_2(n,t)] \\
&+ \ \gamma[(n+1)p_2(n+1,t) - np_2(n,t)] \\
&- \ hn(n-1)p_2(n,t) + fp_1(n,t),
\end{aligned}
$$

where k_1 and k_2 are the protein-production rates for gene states 1 and 2, γ is the decay rate of copy-number of proteins due to the cell division, and f and $hn(n-1)$ are the switching rates between the two gene states. Equation for $p_1(n,t)$ listed above was obtained from Eq. (1) in Supplementary Material of [11] by taking $\delta = 2$ for dimer formation.

For the condition (A): $N \equiv k_1/\gamma \gg 1$, $k_1 \gg k_2$ with parameters $\bar{h} \equiv N^2 h$, $hn(n-1) = \bar{h}x(x-1/N)$, and $x = n/N$, GQX [11] announced that the model in intermediate regime of parameters has been solved in the two cases: I: $k_1 \gg \bar{h}, f$, II: $k_1 \ll \bar{h}, f$

$$p_1(n,t) = v_1 e^{Nu(x,t)}, p_2(n,t) = v_2 e^{Nu(x,t)},$$

Finite size corrections.-

$$p_1(x) = v_1 e^{Nu_s + \frac{h_1}{N}} = e^{Nu_s + \ln v_1 + \frac{h_1}{N}},$$

$$p_2(x) = v_2 e^{Nu_s + \frac{h_2}{N}} = e^{Nu_s + \ln v_2 + \frac{h_2}{N}}.$$

$$\begin{aligned} P(x) &= p_1(x) + p_2(x) \\ &= (v_1 + v_2) e^{Nu_s} \\ &= (1+F) e^{Nu_s + \int^x (\frac{M_{22} h_d}{D_2} - \frac{Q_{22}}{D_2}) dx'}. \end{aligned}$$

$$h_d = \frac{Q_{11}F + Q_{12} + D_1 F' - (D_1/D_2) F Q_{22}}{M_{12} - (D_1/D_2) F M_{22}},$$

$$v_2 = \exp\left[\int^x \left(\frac{M_{22} h_d}{D_2} - \frac{Q_{22}}{D_2}\right) dx'\right],$$

$$v_1 = F v_2,$$

$$\frac{Q_{11}F - M_{12}h_d + Q_{12} + D_1 F'}{Q_{22} - M_{22} h_d} = \frac{D_1}{D_2} F,$$

$M_{11} = e^{-p} + x\, e^{p} - 1 - x - a, M_{12} = bx^2, M_{21} = a, M_{22} = c\, e^{-p} + x\, e^{p} - c - x - bx^2, D_1 = x\, e^{p} - e^{-p}, D_2 = x\, e^{p} - c\, e^{-p}, Q_{11} = \frac{p'}{2}(x\, e^{p} + e^{-p}) - e^{p} + 2, Q_{12} = -bx, Q_{21} = 0,$

$a = f/k_1$, $b = \bar{h}/k_1$, and $c = k_2/k_1$,

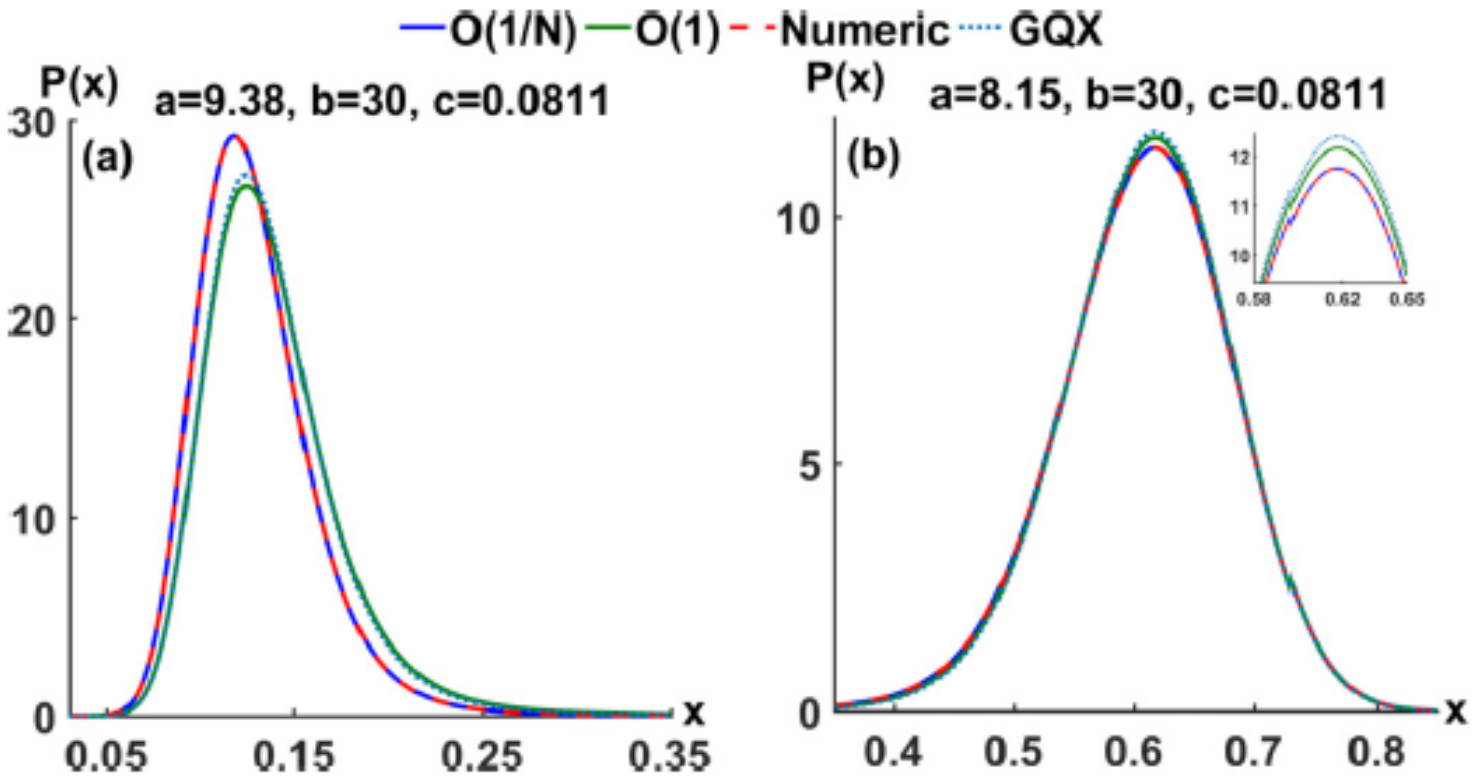


FIG. 2. $P(x)$ figures for parameters: $k_1 = 10$, $k_2 = 0.8108$, $\gamma = 0.02$, $\bar{h} = 30k_1$, $b = 30$, $c = 0.0811$, (a) $f = \bar{h}/3.2$ and $a = 9.38$, and (b) $f = \bar{h}/3.68$ and $a = 8.15$. The green solid lines are constructed by setting $v_1 + v_2 = 1$ without finite-size corrections: $P(x) = \exp(Nu_s(x))$ with $u_s(x)$ given by Eq. (14), the blue solid lines are predicted by Eq. (25), and the blue dash lines are predicted with u_s given by Eq. (18). The unit of $P(x)$ is 10^{-3} and the inset of (b) is as a guide to the eye.

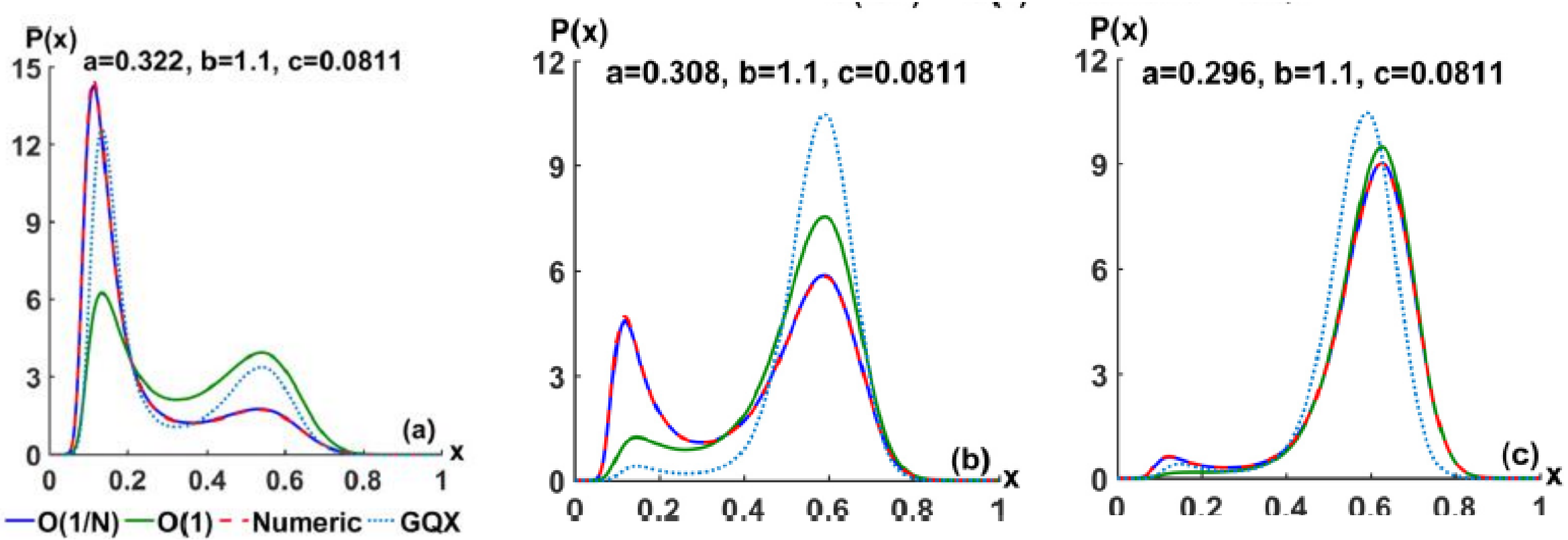


FIG. 3. $P(x)$ figures for parameters: $\bar{h} = 1.1k_1, b = 1.1$, (a) $f = \bar{h}/3.42$ and $a = 0.322$, (b) $f = \bar{h}/3.57$ and $a = 0.308$, and (c) $f = \bar{h}/3.72$ and $a = 0.296$. Other parameters and notations are the same as those of Fig. 2. The unit of $P(x)$ is 10^{-3}.

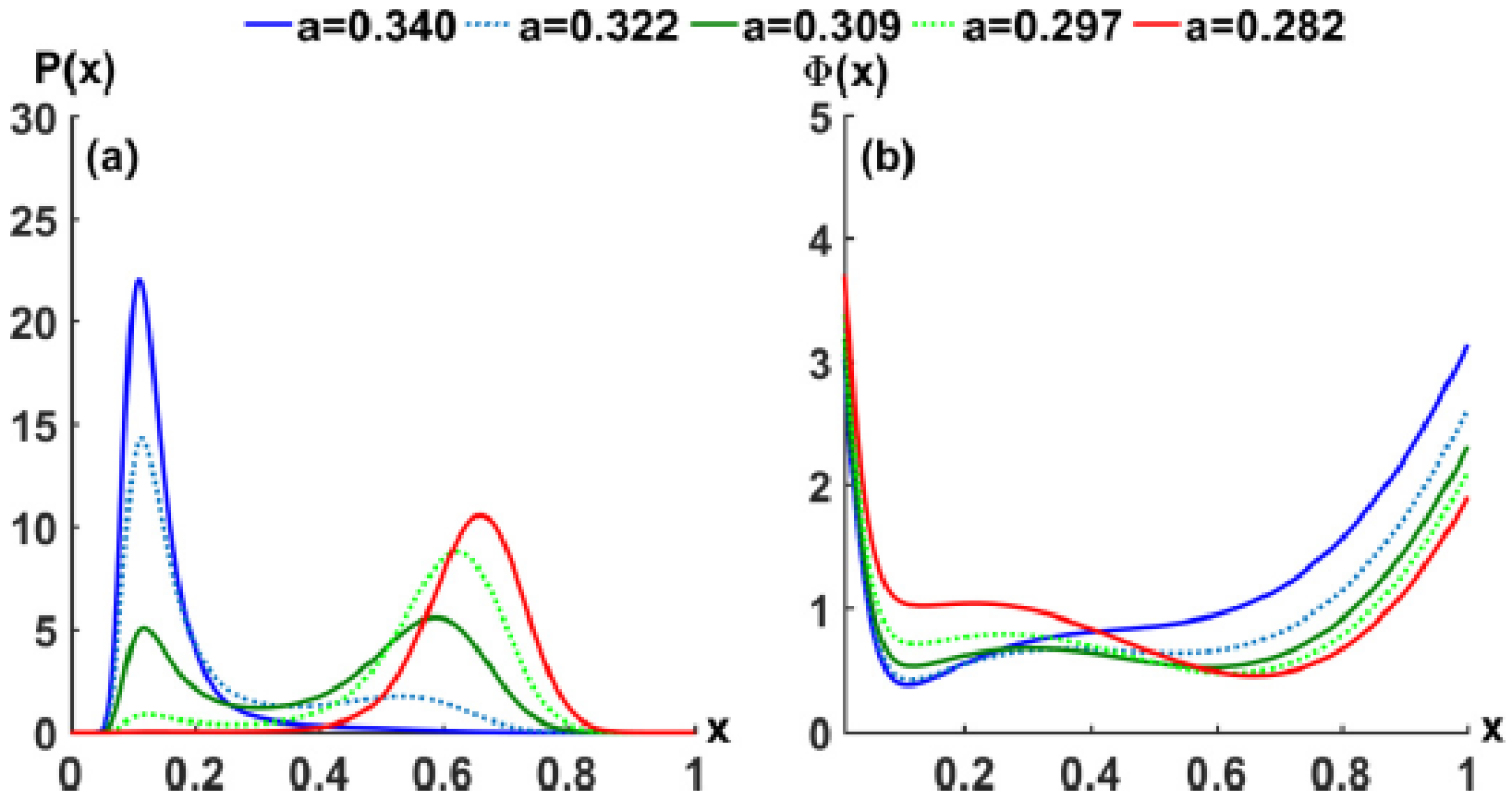


FIG. 4. (a) $P(x)$ as a function of x for different values of a; other parameters used here are the same as those in Fig. 3. The data for curves are calculated from Eq. (25) and the unit of $P(x)$ is 10^{-3}. (b) The corresponding generalized energy landscapes of (a) by $\Phi(x) = -\ln P(x)$ and the unit of $\Phi(x)$ is 10^{1}.

Summary and further works

- Statistical physics can be useful for understanding biological evolution or process in different time scales.
- It is of interest to have a coherent and consistent picture of biological evolution from very small time scales to very large time scales.

Summary

We solved Ge, Qian, and Xie (GQX) model [Phys. Rev. Lett. {\bf 114}, 078101 (2015)] for the molecular process in a single cell accurately by Hamilton-Jacobi equation (HJE) method with finite size corrections, and it is verified by numerics. For the large parameters, our method works more accurately than the method by GQX, while for intermediate parameters our formulas are close to the numerics in a high accuracy and very different from their results, yields more accurate landscape functions. Our HJE approach can explain physically the phenotypic transitions of lac operon in E. coli. [P. J. Choi, et al., Science 322, 442 (2008)] and can have wide applications.

Analytic Solution of a Phenotype Transition Model for a Single Cell

Guan-Rong Huang[1], David B. Saakian[2,3,4],* and Chin-Kun Hu[1,2,5†]

[1]*Physics Division, National Center for Theoretical Sciences, Hsinchu 30013, Taiwan*

[2]*Institute of Physics, Academia Sinica, Nankang, Taipei 11529, Taiwan*

[3] *Theoretical Physics Research Group,*

Ton Duc Thang University, Ho Chi Minh City, Vietnam

[4]*Faculty of Applied Sciences, Ton Duc Thang University, Ho Chi Minh City, Vietnam and*

[5]*Department of Systems Science, University of Shanghai*

for Science and Technology,Shanghai 200093, China

[2] P. J. Choi, L. Cai, K. Frieda, and X. S. Xie, Science **322**, 442 (2008).

[11] H. Ge, H. Qian, and X. S. Xie, Phys. Rev. Lett. **114**, 078101 (2015).

(a) $k_2 >> f, \bar{h}$

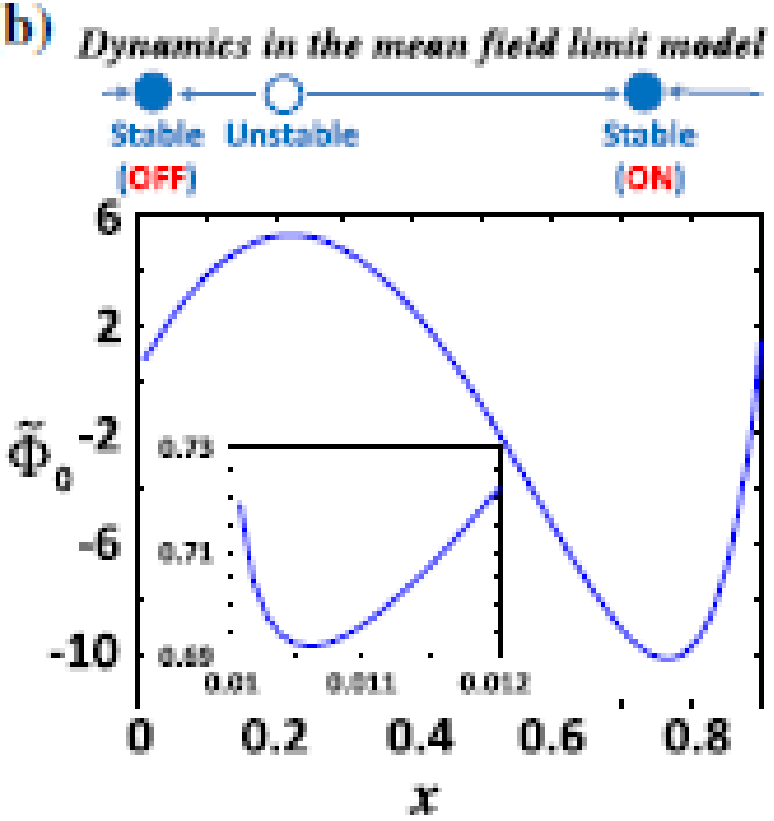
Fluctuating-rate model

$$\frac{dx}{dt} = \frac{k_2}{n_{max}} - \gamma x$$

$$\bar{h} x^2 \left\{ \begin{array}{l} \uparrow \\ \downarrow \end{array} \right. f \qquad x = \frac{n}{n_{max}}$$

$$\frac{dx}{dt} = \frac{k_2}{n_{max}} - \gamma x$$

(b) *Dynamics in the mean field limit model*



Hao Ge, Hong Qian, and X. Sunney Xie: Stochastic Phenotype Transition of a Single Cell in an Intermediate Region of Gene State Switching, Phys. Rev. Lett. 114, 078101 (2015).

Figure 1

(a) A minimal gene network with positive feedback and two different gene states. (b) The diagram of the full chemical master equation [see Eq. (1)]. (c) Deterministic mean-field model [Eq. (2)] with bistability induced by positive feedback, in which $k1=10\ \ \mathrm{min}-1$, $k2=0.1\ \ \mathrm{min}-1$, and $\gamma=0.02\ \ \mathrm{min}-1$.

Figure 2

In the intermediate region of gene state switching, the full CME can be simplified to a fluctuating-rate model (a), the steady-state distribution of which corresponds to a normalized landscape function $\tilde{}\Phi 0(x)$ (b). In the case if the gene state switching is extremely rapid, a reduced CME (c) and a different landscape function $\tilde{}\Phi 1(x)$ (d) can also be derived. The insets in (b) and (d) are the zoom in of the functions near $x=0$. The parameters in (b) and (d) are the same as those in Fig. 1 with $Keq=1/5.5$.

Analytic Solution of a Phenotype Transition Model for a Single Cell

- Chin-Kun Hu (胡進錕)

  Institute of Physics, Academia Sinica, Taipei 11529, Taiwan and NCTS, Hsinchu
- E-mail: **huck@phys.sinica.edu.tw**
- http://proj1.sinica.edu.tw/~statphys/

Collaborators:
Guan-Rong Huang, National Center for Theoretical Sciences, Hsinchu, Taiwan
David B. Saakian, Institute of Physics, AS, Taipei and Yerevan Physics Institute, Armenia