

RESEARCH ARTICLE

Fitness and entropy production in a cell population dynamics with epigenetic phenotype switching

Hong Qian*

Department of Applied Mathematics, University of Washington, Seattle, WA 98195, USA

* Correspondence: hqian@u.washington.edu

Received September 10, 2013; Revised March 6, 2014; Accepted March 16, 2014

Motivated by recent understandings in the stochastic natures of gene expression, biochemical signaling, and spontaneous reversible epigenetic switchings, we study a simple deterministic cell population dynamics in which subpopulations grow with different rates and individual cells can bi-directionally switch between a small number of different epigenetic phenotypes. Two theories in the past, the population dynamics and thermodynamics of master equations, separately defined two important concepts in mathematical terms: the *fitness* in the former and the (non-adiabatic) *entropy production* in the latter. Both of them play important roles in the evolution of the cell population dynamics. The switching sustains the variations among the subpopulation growth, thus sustains continuous natural selection. As a form of Price's equation, the fitness increases with (i) natural selection through variations and (ii) a positive covariance between the per capita growth and switching, which represents a Lamarchian-like behavior. A negative covariance balances the natural selection in a fitness steady state — “the red queen” scenario. At the same time the growth keeps the proportions of subpopulations away from the “intrinsic” switching equilibrium of individual cells, thus leads to a continuous entropy production. A covariance, between the per capita growth rate and the “chemical potential” of subpopulation, counteracts the entropy production. Analytical results are obtained for the limiting cases of growth dominating switching and vice versa.

Keywords: population dynamics; fundamental theorem of natural selection; diversity

INTRODUCTION

Darwinian evolution theory states that variations among subpopulations contribute to the increasing *fitness* of the total population. One of the important issues in this theory is how to maintain the variations while *natural selection* is continuously “purifying” the fittest. In population dynamics with Mendelian inheritance, this problem has been solved by the work of Hardy and Weinberg, and the theory of population genetics developed by J.B.S. Haldane, R.A. Fisher, S. Wright, and many other pioneers [1–3]. See excellent texts [4–6] and a recent review [7] together with the references cited within.

From a population dynamics perspective [8], the key to sustain the variations among the subpopulation within a population is to maintain a stationary frequency, or

proportion, for each subpopulation. In the theory of Markov processes, the concepts of relative entropy and entropy production are quantitative measures for such behavior. In recent years, a rather complete *stochastic thermodynamic theory* of Markov processes has emerged [9–13].

An integration of these two theories is a conceptual problem for a general population dynamics with intra-population transformations among subpopulations. We note that one of the key differences between the classical Lea-Coulson theory of bacterial population growth with mutations [14] and the problem of cell population growth with epigenetic switching is that in the former a back mutation can be safely neglected, while in the latter bi-directional switchings among a small number of possible states occur, thus the notion of dynamical equilibrium among subpopulations becomes meaningful in principle.

Indeed, Markov dynamics with *ergodic stationary state* among subpopulations also has emerged in laboratory studies [15,16]. Such dynamics occurs in the population biology of cellular epigenetic differentiation [17,18], or certain tumor progression [19] and cancer cell development involving cancer stem cells and non-stem cancer cells [16]. According to our current understanding, biochemical regulatory networks within an individual cell give rise to multiple attractors which yield different phenotypic states [20–22]. These subpopulations of cells grow with different rates, but with the possibility of spontaneous transition (switching) from one phenotypic state to another [23]. In the simplest mathematical model, one has [24]

$$\frac{dx_i}{dt} = x_i r_i(\vec{x}) + \sum_{j=1}^n (T_{ij} x_j - T_{ji} x_i), \quad i = 1, 2, \dots, n, \quad (1)$$

in which x_i is the number density of the subpopulation i , r_i is its per capita growth rate, and T_{ij} represents the individual cell switching rate from state j to state i . We shall use the notation $\vec{x} = (x_1, x_2, \dots, x_n)$.

Following the pioneering work of Eigen and Schuster [25], biological evolution has been extensively discussed in the context of genotype-to-phenotype maps in terms of DNA sequence space and structural based functions of proteins. It has been recognized that there are great degree of degeneracy in this map [26]. For microbial population dynamics, Kussell and Leibler have studied bet-hedging mechanism a bacterial population with diversity can utilize when encountering a fluctuating environment [27]. They discovered that while the information concerning the environmental variations can cause direct epigenetic switching in an organism, it acts through natural selection on the organism population with spontaneous stochastic switchings. More recently, Mustonen and Lässig developed a highly sophisticated, time-inhomogeneous stochastic theory of fitness flux [28], which generalized Fisher's fundamental theorem of natural selection (FTNS).

There is a long recognized parallelism between the theories of population genetics and the theory of statistical thermal physics of heterogeneous substances. Certainly both systems consist of a large collection of heterogeneous individuals with random behavior, with emergent properties on the systems level. This domain of research has been very active in the 1970s, with multiple analogous relationships being identified and proposed. In fact, one of the most fascinating parallels is the histories of the two fields: While Boltzmann's theory identified $\frac{dS}{dt} \geq 0$ as the Second Law where S is the entropy of an isolated system, Eckart, Bridgman, and

Prigogine recognized that $dS(t) = d_i S + d_e S$, in which the Second Law is represented by entropy production $d_i S \geq 0$ while total dS no longer has a definitive sign in general. In population genetics, Fisher's theory identified $\frac{d\bar{r}}{dt} \geq 0$ as Darwinian law of natural selection where \bar{r} is the Malthusian parameter of a single locus with constant fitness. However, Price, Ewens, and Edwards later recognized that $d\bar{r}(t) = \partial_{ns} r + \partial_{ec} r$ [7,29,30], in which FTNS is represented by $\partial_{ns} r \geq 0$ while total $d\bar{r}$ no longer has a definitive sign in general.

In the present paper, we do not attempt to deal with the parallelism, nor important issues such as equilibrium and non-equilibrium statistical theory of evolution. In fact, the very notions of "equilibrium" and "non-equilibrium" defined through traditional detailed balance can be challenged [31]. We have also neglected many biological factors such as genetic drift due to finite size populations. Our goal is simply to illustrate that Fisher's fitness in terms of the Malthusian parameter together with relative entropy provide a quantitative characterization of continuous natural selection and sustained diversity. It is noted that, in many recent stochastic theories of evolution, fitness of a population has been replaced by Wright's fitness landscape, which is then represented mathematically by a probability density function. We only study the deterministic population dynamics in Eq. (1) with constant per capita growth rates r_i and transition rate T_{ij} . This is certainly a gross simplification; however, our aim is to focus on the interplay between the fitness in the population growth part and the entropy production in the phenotypic switching part of the dynamics. While both are motivated from the theory of probability, the biological relationship between the entropy production used in the present work and those in stochastic theories [28,32,33] remains to be clarified.

Even though the issue of bi-directional switching remains controversial in experimental investigations and pre-clinical studies [24], there is a little doubt that (i) within a given environment, transcription regulations of gene expression are fundamental biochemical processes that determine phenotypic states of a cell; and (ii) biochemical reactions involved in gene regulations are highly nonlinear with feedbacks. Therefore, considering nonlinear biochemical reaction network dynamics, and its attractors, as a basis of cellular epigenetic states is highly logical; and bi-directional switching is a necessary prediction of such a mechanism. In reality, some transitions are fast and some transitions are slow; being able to observe bi-directional, reversible switching or not is a matter of time scale; and elucidating the specific molecular mechanism(s) for a particular phenotypic switching is a matter of details.

FITNESS, NATURAL SELECTION, AND ENTROPY PRODUCTION

In this section, we first give a brief summary of the two separate mathematical theories, one for fitness and natural selection in population with differential growth, and one for stochastic thermodynamics in populations with switching transitions.

Since the work of Haldane and Fisher more than 80 years ago [1,2,7], the mean growth rate of a total population that consists of subpopulations with differential growth

$$\bar{r} = \frac{\sum_{i=1}^n x_i r_i(\vec{x})}{\sum_{i=1}^n x_i}, \tag{2}$$

has been the mathematical quantification of the fitness of the population. Note that \bar{r} is actually the per capita growth rate, i.e., the Malthusian parameter, for the entire population [4]:

$$\bar{r} = \frac{\sum_{i=1}^n x_i r_i(\vec{x})}{\sum_{i=1}^n x_i} = \frac{d}{dt} \left(\frac{\sum_{i=1}^n x_i}{\sum_{i=1}^n x_i} \right). \tag{3}$$

We shall follow this established terminology in the present work. One can also find other related mathematical concepts in the literature. For example, Wright developed the notion of “fitness landscape” [3] and Iwasa introduced a novel mathematical quantity called “free fitness” [34], both were proposed to represent more broadly or faithfully Darwin’s original idea and evolutionary processes [35].

For the rest of this paper, we shall assume r_i to be a constant, independent of $\{x_i\}$. This assumption, which is not unreasonable for cell subpopulations growth, is essential for the mathematical analysis carried out below.

In the absence of switching transitions, e.g., $T_{ij} = 0$, the change of \bar{r} as a function time is

$$\frac{d\bar{r}}{dt} = \frac{\sum_{i=1}^n x_i (r_i - \bar{r})^2}{\sum_{i=1}^n x_i} = \sigma_r^2 \geq 0. \tag{4}$$

Therefore, variations in the per capita growth rates among the subpopulations, σ_r^2 , is responsible for the

increase of fitness \bar{r} by natural selection. This simple, elegant, and insightful result has been known as Fisher’s fundamental theorem of natural selection [4,29,30].[†]

However, in the absence of T_{ij} , it is also clear that the variation eventually disappear, and only the fittest with the largest r_i dominates.

We now denote the frequency of the i th population

$$z_i = \frac{x_i}{\sum_{i=1}^n x_i}. \tag{5}$$

Then Eq. (1) becomes

$$\frac{dz_i}{dt} = \left(r_i - \sum_{j=1}^n r_j z_j \right) z_i + \sum_{j=1}^n (T_{ij} z_j - T_{ji} z_i). \tag{6}$$

Eq. (6) with all $r_i = 0$ has a very different kind of dynamics. Let us assume that T_{ij} is irreducible, and $T_{ij} = 0$ iff $T_{ji} = 0$, e.g., the dynamics is “bi-directional” with no “absorbing state”. Then there exists a unique equilibrium $\{z_i^* | z_i^* > 0, \sum_{i=1}^n z_i^* = 1\}$ such that

$$\sum_{j=1}^n (T_{ij} z_j^* - T_{ji} z_i^*) = 0, \quad i. \tag{7}$$

And the relative entropy

$$H[z_i || z_i^*] = \sum_{i=1}^n z_i \ln \frac{z_i}{z_i^*} \tag{8}$$

has the important properties [32,37–40]:

$$H[z_i(t) || z_i^*] \geq 0, \quad \frac{d}{dt} H[z_i(t) || z_i^*] \leq 0. \tag{9}$$

In stochastic thermodynamics, H is known as generalized free energy, and

$$e_p^{(na)} = -\frac{dH}{dt} = \sum_{i,j=1}^n T_{ji} z_i \ln \frac{z_i z_j^*}{z_j z_i^*} \geq 0 \tag{10}$$

is called free energy dissipation, or non-adiabatic entropy production [10,11]. In Eq. (10), we used the convention $T_{ii} = -\sum_{j=1, j \neq i}^n T_{ji}$.

The thermodynamic theory of Markov processes also identified an adiabatic entropy production [10,11]

$$e_p^{(a)} = \sum_{i,j=1}^n T_{ij} z_j \ln \frac{T_{ij} z_j^*}{T_{ji} z_i^*} \geq 0. \tag{11}$$

[†]It is interesting to point out that this result is intimately related to a general equality which plays an important role in many other branches of mathematics [36]: If $P_n(\cdot)$ is a sequence of probability measures, for any two sets A and B even they are not disjoint,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \ln P_n(A \cup B) = \max \left\{ \lim_{n \rightarrow \infty} \frac{1}{n} \ln P_n(A), \lim_{n \rightarrow \infty} \frac{1}{n} \ln P_n(B) \right\}.$$

We see that if T_{ij} satisfy detailed balance, i.e., $T_{ij}z_j^* = T_{ji}z_i^*$, then $e_p^{(a)} = 0$, and

$$e_p^{(na)} = \sum_{i,j=1}^n T_{ji}z_i \ln \frac{T_{ji}z_i}{T_{ij}z_j} = \frac{1}{2} \sum_{i,j=1}^n (T_{ji}z_i - T_{ij}z_j) \ln \frac{T_{ji}z_i}{T_{ij}z_j}. \quad (12)$$

In L. Onsager's theory of irreversible thermodynamics of inanimated processes [41], detailed balance is a fundamental assumption. The right-hand-side of (12) is then interpreted as Onsager's "thermodynamic flux times force". We do not assume detailed balance in the present work, since for biological processes, there is clearly no fundamental reason for such a constraint. (Such a constraint in classical physics is due to the impossibility of perpetuate motion machine.)

In summary, population dynamics with differential growth alone has increasing fitness with rate σ_r^2 that reflects Darwin's natural selection due to variation, and population dynamics with switching transformation alone has non-negative $e_p^{(na)}$ and $e_p^{(a)}$ reflecting the Second Law of thermodynamics. In the next section, we shall study the dynamics in Eq. (6) that combines these two fundamental theories.

INTERPLAY BETWEEN DIFFERENTIAL GROWTH AND SWITCHING TRANSITIONS

If population changes following the full Eq. (6), we have a *dynamic equation for fitness*:

$$\frac{d\bar{r}}{dt} = \sigma_r^2 - \sum_{i,j=1}^n r_i (T_{ji}z_i - T_{ij}z_j) \quad (13)$$

$$= \sigma_r^2 + \overline{(r-\bar{r})(\phi-\bar{\phi})}, \quad (14)$$

in which the last term

$$\overline{(r-\bar{r})(\phi-\bar{\phi})} = \sum_{i=1}^n z_i (r_i - \bar{r})(\phi_i - \bar{\phi}), \quad (15)$$

with

$$\phi_i = \frac{1}{z_i} \sum_{j=1}^n (T_{ij}z_j - T_{ji}z_i), \text{ and } \bar{\phi} = \sum_{i=1}^n z_i \phi_i = 0. \quad (16)$$

ϕ_i is the per capita rate of increase of the i th subpopulation

due to switching transition. On average $\bar{\phi} = 0$. The term in Eq. (15) can be considered as the covariance between the per capita growth rates due to growth and due to switching [42]. A positive covariance can be viewed phenomenologically as a Lamarchian selection. Eq. (14) shows that, at least in this very simple case, one can mathematically distinguish the covariance effect from σ_r^2 , the natural selection. An equation like (14) is known as Price's theorem [29,42].

On the other hand, from Eq. (8) we have

$$\begin{aligned} \frac{dH}{dt} &= \sum_{i,j=1}^n (T_{ji}z_i - T_{ij}z_j) \ln \frac{z_i(t)}{z_i^*} \\ &+ \sum_{i=1}^n \left(r_i - \sum_{j=1}^n r_j z_j \right) z_i \ln \frac{z_i(t)}{z_i^*} \\ &= -e_p^{(na)} + \overline{(r-\bar{r})(\eta-H)}, \end{aligned} \quad (17)$$

where

$$\eta_i = \ln \frac{z_i}{z_i^*}, \text{ and } \bar{\eta} = \sum_{i=1}^n z_i \eta_i = H. \quad (18)$$

The first term in (17) is never positive: $e_p^{(na)} \geq 0$. The second term in (17) represents the covariance between the per capita growth rate and the "chemical potential" of a subpopulation. In classical statistical physics, $-\ln z^*$ and $-\ln z$ are called internal energy and entropy, respectively. And internal energy minus entropy is called free energy. It measures the deviation of z_i from equilibrium z_i^* [39].[‡]

When the dynamics in (6) reaches stationarity, one has $dz_i/dt = 0$, i.e., $(r_i - \bar{r}) = -(\phi_i - \bar{\phi})$. In this case, $dH/dt = 0, d\bar{r}/dt = 0$, but

$$\sigma_r^2 = -\overline{(r-\bar{r})(\phi-\bar{\phi})} > 0, \quad (19)$$

and

$$e_p^{(na)} = \overline{(r-\bar{r})(\eta-H)} = -\overline{(\phi-\bar{\phi})(\eta-H)} > 0. \quad (20)$$

The differential growth continuously keeps the population frequency out of equilibrium, $e_p^{(na)} > 0$; and an anti-Lamarchian switching continuously keeps the variation finite, $\sigma_r^2 > 0$. Such a non-equilibrium steady state [43,44] has been termed "the red queen" scenario in evolution [33,45,46]. In the simple cell dynamics, a continuous evolution is necessary to oppose the ever-present relapsing into inferior growth states due to spontaneous epigenetic switching.

[‡]Curiously, covariances also figured prominently in Gibbs' theory of ensemble fluctuations. Consider a stochastic system with Gibbs distribution $Z^{-1}(\beta)e^{-\beta E_i}$ where $Z(\beta) = \sum_i e^{-\beta E_i}$, E_i is the energy of of state i and O_i is an observable associated with the state. Then the ensemble average $\langle O \rangle = Z^{-1}(\beta) \sum_i O_i e^{-\beta E_i}$, whose rate of change with respect to β : $\frac{d\langle O \rangle}{d\beta} = -\langle (O - \langle O \rangle)(E - \langle E \rangle) \rangle$.

TWO ASYMPTOTIC LIMITS

$$\vec{x}(t) = [e^{\mathbf{A}t}] \vec{x}(0), \tag{21}$$

The solution to the linear Eq. (1) has the general form in which

$$A = \begin{pmatrix} r_1 & 0 & 0 & \dots & 0 \\ 0 & r_2 & 0 & \dots & 0 \\ \dots & \vdots & \ddots & \vdots & 0 \\ 0 & \dots & \dots & r_{n-1} & 0 \\ 0 & \dots & \dots & 0 & r_n \end{pmatrix} + \begin{pmatrix} -\sum_{j=2}^n T_{1j} & T_{12} & T_{13} & \dots & T_{1n} \\ T_{21} & -\sum_{j \neq 2} T_{2j} & T_{23} & \dots & T_{2n} \\ \vdots & \vdots & \ddots & \vdots & 0 \\ T_{n-1,1} & \dots & \ddots & -\sum_{j \neq n-1} T_{n-1,j} & 0 \\ T_{n1} & \dots & \dots & 0 & -\sum_{j=1}^{n-1} T_{nj} \end{pmatrix}. \tag{22}$$

We shall order the growth rates such that $r_1 > r_2 \geq \dots r_n$. In the limit of t tending ∞ , we have

$$x_i(t) = v_i e^{\lambda_{\max} t} \sum_{j=1}^n u_j x_j(0), \tag{23}$$

where λ_{\max} is the largest, positive eigenvalue of \mathbf{A} , and (u_1, u_2, \dots, u_n) and $(v_1, v_2, \dots, v_n)^T$ are the left and right eigenvectors corresponding to λ_{\max} . Therefore,

$$z_i^{ss} = \lim_{t \rightarrow \infty} \frac{x_i(t)}{\sum_{i=1}^n x_i(t)} = \frac{v_i}{\sum_{i=1}^n v_i}. \tag{24}$$

When all the $r_i = 0$, the $z_i^{ss} = z_i^*$.

Two special cases are particularly interesting and solvable with the eigenvalue perturbation method.

Growth dominates switching. When the switching rates are much smaller than the growth rates, $T_{ij} \ll r_k, i, j, k = 1, 2, \dots, n$,

$$\lambda_{\max} \approx r_1 - \sum_{j=2}^n T_{1j}, \tag{25}$$

and the corresponding right eigenvector $\{v_i\}$ is

$$v_i = \frac{T_{i1} v_1}{r_1 - r_i}, v_1 = \left(1 + \sum_{i=2}^n \frac{T_{i1}}{r_1 - r_i} \right)^{-1} \tag{26}$$

Therefore,

$$\bar{r} = \sum_{i=1}^n v_i r_i = \frac{r_1 + \sum_{i=2}^n \frac{r_i T_{i1}}{r_1 - r_i}}{1 + \sum_{i=2}^n \frac{T_{i1}}{r_1 - r_i}}, \tag{27}$$

and

$$\sigma_r^2 = \sum_{i=2}^n T_{i1} (r_1 - r_i) + \frac{1}{2} \sum_{i,j=2}^n \frac{T_{i1} T_{j1} (r_i - r_j)^2}{(r_1 - r_i)(r_1 - r_j)}. \tag{28}$$

Switching dominates growth. If $r_k \ll T_{ij}, i, j, k = 1, 2, \dots, n$. Then as the leading order

$$\lambda_{\max} = \sum_{i=1}^n r_i z_i^*. \tag{29}$$

The corresponding right eigenvector is approximately z_i^* with a correction on the order of $\frac{r}{T}$:

$$v_i = z_i^* + \xi_i, \text{ where } \sum_{j=1}^n T_{ij} \xi_j = (\lambda_{\max} - r_i) z_i^*. \tag{30}$$

Since matrix $\{T_{ij}\}$ is singular with left null vector $(1, 1, \dots, 1)$, the solution $\{\xi_i\}$ is not unique. An additional condition is $\sum_{i=1}^n \xi_i = 0$. Hence,

$$\bar{r} = \sum_{i=1}^n z_i^* r_i, \tag{31}$$

$$\sigma_r^2 = \sum_{i=1}^n z_i^* (r_i - \bar{r})^2, \tag{32}$$

and

$$\begin{aligned}
 e_p^{(na)} &= \sum_{i,j=1}^n (T_{ij}v_j - T_{ji}v_i) \ln \left(1 + \frac{\xi_i}{z_i^*} \right) \\
 &\approx \sum_{i,j=1}^n (T_{ij}\xi_j - T_{ji}\xi_i) \left(\frac{\xi_i}{z_i^*} \right) \\
 &= \sum_{i,j=1}^n \frac{T_{ij}\xi_j\xi_i}{z_i^*} = \sum_{i=1}^n \xi_i(\lambda_{\max} - r_i) \\
 &= - \sum_{i=1}^n \xi_i r_i.
 \end{aligned} \tag{33}$$

DISCUSSION

Homeostatic biological processes in living organisms carry out energy, material, and information transformations at the expense of chemical free energy, e.g., negentropy [47,48]. Putting this intuitive chemical perspective aside, current theoretical understanding in entropy production has clearly shown its deep root in time irreversibility of dynamical systems [12,31,43,44,49]. According to this mathematical perspective, a non-stationary process with growth naturally has a positive non-adiabatic entropy production [10,11,38], thus has a certain amount of “free energy” available for driving *other* dynamical processes out of equilibrium. In the present work, we show that cell population growth causes a non-equilibrium distribution among the subpopulations with respect to their spontaneous, epigenetic switching. This distribution, in return, drives a Darwinian natural selection through the “invisible hand” of subpopulation variations. Darwinian evolution and the Second Law of thermodynamics, thus, are intimately coupled, at least in this simple population dynamics, through the certainty of mathematical analysis.

In the past, studies on population dynamics and evolution by natural selection were intentionally separated because it was felt that they are processes with widely different time scales [5]. For example, Lotka clearly introduced the notions of “inter-species growth dynamics” and “intraspecies evolution”, with the former proceeds in a rapid rate in comparison to the latter [50,51]. Roughgarden, however, has given an illuminating discussion on the ecological versus evolutionary times. In particular, he illustrated that a consideration of the latter can provide a mechanistic understanding for the parameters in the former. In the language of the present work, the switching rates T_{ij} of individual cells contribute to the overall growth rate of the entire population. This is precisely the goal of a biochemical understanding of individual cells in a cell population with a diversity in epigenetic phenotypes [52]. Being discrete nonlinear dynamic attractors [53], such phenotypic states are naturally stable, with robustness against both internal

and external disturbances. Switchings among such states, being stochastic rare events, necessarily exhibit discontinuous jumps, e.g., punctuated equilibria.

I thank Drs. Ping Ao, Joe Felsenstein, Sui Huang, Susan Rosenberg, Jin Wang, and Da Zhou for many helpful discussions and inspirations, and my colleagues M. Kot, J. D. Murray, and K. K. Tung for encouragements. I also thank the anonymous reviewers of the paper, who have all given very helpful comments and suggestions.

COMPLIANCE WITH ETHICS GUIDELINES

The author Hong Qian declares that there is no conflict of interest.

The article does not contain any studies with human or animal subjects performed by the author.

REFERENCES

- Haldane, J. B. S. (1924) A mathematical theory of natural and artificial selection. *Trans. Cambridge Philos. Soc.*, 23, 19–41
- Fisher, R. A. (1930) *The Genetical Theory of Natural Selection*. Oxford: Clarendon Press
- Wright, S. (1932) The roles of mutation, inbreeding, crossbreeding and selection in evolution. *Proc. Sixth Intern. Congress Genet.* 1, 356–366
- Crow, J. F. and Kimura, M. (1970) *An Introduction to Population Genetics Theory*. Caldwell: Harper & Row
- Roughgarden, J. (1979) *Theory of Population Genetics and Evolutionary Ecology: An Introduction*. New York: Macmillan Publishers.
- Felsenstein, J. (2011) *Theoretical Evolutionary Genetics*. Online Books, <http://evolution.gs.washington.edu/pgbook/pgbook.pdf>
- Ewens, W. J. (2013) *Mathematics, genetics and evolution*. *Quant. Biol.*, 1, 9–31
- Qian, H. (2011) Nonlinear stochastic dynamics of mesoscopic homogeneous biochemical reaction systems—An analytical theory. *Nonlinearity*, 24, R19–R49
- Qian, H., Qian, M. and Tang, X. (2002) Thermodynamics of the general diffusion process: Time-reversibility and entropy production. *J. Stat. Phys.*, 107, 1129–1141
- Ge, H. and Qian, H. (2010) Physical origins of entropy production, free energy dissipation, and their mathematical representations. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.*, 81, 051133
- Esposito, M. and Van den Broeck, C. (2010) Three detailed fluctuation theorems. *Phys. Rev. Lett.*, 104, 090601
- Seifert, U. (2012) Stochastic thermodynamics, fluctuation theorems and molecular machines. *Rep. Prog. Phys.*, 75, 126001
- Qian, H. and Kou, S. C. (2014) Statistics and related topics in single-molecule biophysics. *Ann. Rev. Stat. Appl.*, 1, 465–492
- Zheng, Q. (1999) Progress of a half century in the study of the Luria-Delbrück distribution. *Math. Biosci.*, 162, 1–32
- Chang, H. H., Hemberg, M., Barahona, M., Ingber, D. E. and Huang, S. (2008) Transcriptome-wide noise controls lineage choice in mammalian progenitor cells. *Nature*, 453, 544–547
- Gupta, P. B., Fillmore, C. M., Jiang, G., Shapira, S. D., Tao, K.,

- Kuperwasser, C. and Lander, E. S. (2011) Stochastic state transitions give rise to phenotypic equilibrium in populations of cancer cells. *Cell*, 146, 633–644
17. Huang, S. (2012) The molecular and mathematical basis of Waddington's epigenetic landscape: a framework for post-Darwinian biology? *BioEssays*, 34, 149–157
 18. Wang, J., Zhang, K., Xu, L. and Wang, E. (2011) Quantifying the Waddington landscape and biological paths for development and differentiation. *Proc. Natl. Acad. Sci. U.S.A.*, 108, 8257–8262
 19. Huang, S. (2012) Tumor progression: chance and necessity in Darwinian and Lamarckian somatic (mutationless) evolution. *Prog. Biophys. Mol. Biol.*, 110, 69–86
 20. Wang, G., Zhu, X., Hood, L. and Ao, P. (2013) From phage lambda to human cancer: endogenous molecular-cellular network hypothesis. *Quant. Biol.*, 1, 32–49
 21. Qian, H. and Ge, H. (2012) Mesoscopic biochemical basis of isogenetic inheritance and canalization: stochasticity, nonlinearity, and emergent landscape. *Mol Cell Biomech*, 9, 1–30
 22. Wang, J., Xu, L. and Wang, E. (2008) Potential landscape and flux framework of nonequilibrium networks: robustness, dissipation, and coherence of biochemical oscillations. *Proc. Natl. Acad. Sci. U.S.A.*, 105, 12271–12276
 23. Ge, H., Wu, P., Qian, H. and Xie, X. S. (2013) A gene regulation theory for spontaneous transition between cellular phenotypes under nonequilibrium steady-state conditions. (Manuscript in preparation)
 24. Zhou, D., Wu, D.-M., Li, Z., Qian, M.-P. and Zhang, M. Q. (2013) Population dynamics of cancer cells with cell state conversions. *Quant. Biol.*, 1, 201–208
 25. Eigen, M. and Schuster, P. (1979) *The Hypercycle: A Principle of Natural Self-Organization*. Berlin: Springer-Verlag
 26. van Nimwegen, E., Crutchfield, J. P. and Huynen, M. (1999) Neutral evolution of mutational robustness. *Proc. Natl. Acad. Sci. U.S.A.*, 96, 9716–9720
 27. Kussell, E. and Leibler, S. (2005) Phenotypic diversity, population growth, and information in fluctuating environments. *Science*, 309, 2075–2078
 28. Mustonen, V. and Lässig, M. (2010) Fitness flux and ubiquity of adaptive evolution. *Proc. Natl. Acad. Sci. U.S.A.*, 107, 4248–4253
 29. Price, G. R. (1972) Fisher's 'fundamental theorem' made clear. *Ann. Hum. Genet.*, 36, 129–140
 30. Edwards, A. W. F. (1994) The fundamental theorem of natural selection. *Biol. Rev. Camb. Philos. Soc.*, 69, 443–474
 31. Qian, H. (2014) The zeroth law of thermodynamics and volume-preserving conservative system in equilibrium with stochastic damping. *Phys. Lett. A*, 378, 609–616
 32. Ao, P. (2008) Emerging of stochastic dynamical equalities and steady state thermodynamics from Darwinian dynamics. *Commun. Commun Theor Phys*, 49, 1073–1090
 33. Zhang, F., Xu, L., Zhang, K., Wang, E. and Wang, J. (2012) The potential and flux landscape theory of evolution. *J. Chem. Phys.*, 137, 065102
 34. Iwasa, Y. (1988) Free fitness that always increases in evolution. *J. Theor. Biol.*, 135, 265–281
 35. Maddox, J. (1991) Is Darwinism a thermodynamic necessity? *Nature*, 350, 653
 36. Varadhan, S. R. S. (2012) Book review: Large deviations for stochastic processes. *Am. Math. Soc*, 49, 597–601
 37. Voigt, J. (1981) Stochastic operators, information, and entropy. *Commun. Math. Phys.*, 81, 31–38
 38. Mackey, M. C. (1989) The dynamic origin of increasing entropy. *Rev. Mod. Phys.*, 61, 981–1015
 39. Qian, H. (2001) Relative entropy: free energy associated with equilibrium fluctuations and nonequilibrium deviations. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.*, 63, 042103
 40. Cover, T. M. and Thomas, J. A. (1991) *Elements of Information Theory*. New York: John Wiley & Sons
 41. Onsager, L. (1931) Reciprocal relations in irreversible processes. I. *Phys. Rev.*, 37, 405–426
 42. Price, G. R. (1972) Extension of covariance selection mathematics. *Ann. Hum. Genet.*, 35, 485–490
 43. Zhang, X.-J., Qian, H. and Qian, M. (2012) Stochastic theory of nonequilibrium steady states and its applications (Part I). *Phys. Rep.*, 510, 1–86
 44. Ge, H., Qian, M. and Qian, H. (2012) Stochastic theory of nonequilibrium steady states (Part II): Applications in chemical biophysics. *Phys. Rep.*, 510, 87–118
 45. van Valen, L. (1973) A new evolutionary law. *Evolu. Th.*, 1, 1–30
 46. Ao, P. (2005) Laws in Darwinian evolutionary theory. *Phys. Life Rev.*, 2, 117–156
 47. Schrödinger, E. (1944) *What is Life?* London: Cambridge University Press
 48. Qian, H. (2007) Phosphorylation energy hypothesis: open chemical systems and their biological functions. *Annu. Rev. Phys. Chem.*, 58, 113–142
 49. Qian, H. (2013) A decomposition of irreversible diffusion processes without detailed balance. *J. Math. Phys.*, 54, 053302
 50. Lotka, A. J. (1922) Contribution to the energetics of evolution. *Proc. Natl. Acad. Sci. U.S.A.*, 8, 147–151
 51. Lotka, A. J. (1922) Natural selection as a physical principle. *Proc. Natl. Acad. Sci. U.S.A.*, 8, 151–154
 52. Kirschner, M. W. and Gerhart, J. C. (2005) *The Plausibility of Life: Resolving Darwin's Dilemma*. New Haven: Yale University Press
 53. Kauffman, S. A. (1993) *The Origins of Order: Self-Organization and Selection in Evolution*. New York: Oxford University Press