An Evolutionary Relationship between Genetic Variation and Phenotypic Fluctuation

Kunihiko Kaneko[1 2] Chikara Furusawa[3 2],

[1] Department of Pure and Applied Sciences, Univ. of Tokyo, 3-8-1 Komaba, Meguro-ku, Tokyo 153-8902, Japan

[2] ERATO Complex Systems Biology Project, JST, 3-8-1 Komaba, Meguro-ku, Tokyo 153-8902, Japan

[3] Department of Bioinformatics Engineering, Graduate School of Information Science and Technology, Osaka University, 2-1 Yamadaoka, Suita, Osaka 565-0871, Japan

Corresponding Author: Kunihiko Kaneko

Department of Pure and Applied Sciences, Univ. of Tokyo,

Komaba, Meguro-ku, Tokyo 153-8902, Japan

Tel/FAX: +81-3-5454-6746

E-mail: kaneko@complex.c.u-tokyo.ac.jp

Abstract

The relevance of phenotype fluctuations among clones (i.e., organisms with identical genes) to evolution has recently been recognized both theoretically and experimentally. By considering the stability of the distributions of genetic variations and phenotype fluctuations, we derive a general inequality between the phenotype variance due to genetic differences and the intrinsic phenotype variance of clones. For a given mutation rate, an approximately linear relationship between the two is obtained which elucidates the consistency between the fundamental theorem of natural selection by Fisher and the evolutionary fluctuation-response relationship (fluctuation dissipation theorem) proposed recently. A general condition for the error catastrophe is also derived as the violation of the inequality, which sets up the limit to the speed of stable evolution. All of these theoretical results are confirmed by a numerical evolution experiment of a cell that consists of a catalytic reaction network. Based on the relationships proposed here, relevance of the phenotypic plasticity to evolution as well as the genetic assimilation is discussed.

Keywords:

Evolution, Phenotypic Fluctuations, Error Catastrophe, Fluctuation-Dissipation Theorem, Stability, Phenotype-Genotype Correspondence

1 Introduction

The importance of genetic variation to evolution has been discussed for over a century and is outlined in classic books by Fisher (Fisher 1930, 1958; Edwards, 2000), Kimura (Kimura 1983), and so forth. Among the prominent results is the so-called 'fundamental theorem of natural selection' by Fisher which states that the evolution speed is proportional to the phenotypic variance brought about by the genetic variance of the population. Here it should be noted that in the established field of population genetics, the existence of a single phenotype from a single genotype under a given environmental condition is implicitly assumed in the formulation of the theory.

On the other hand, the phenotypes of clones can fluctuate from individual to individual (Spudich & Koshland 1976), even though they have identical genes, and are put in the same environment, as has been recently confirmed quantitatively for bacteria in an investigation involving the fluctuations of abundances of fluorescent proteins (Elowitz et al. 2002; Ueda et al. 2001; Hasty et al. 2000; Sato et al. 2003; Furusawa et al. 2005). Furthermore, the relevance of phenotypic fluctuations of clones to evolution has also been discussed along the lines of the fluctuation-dissipation theorem in physics (Kubo 1985) which suggests a linear relationship, or at least a positive correlation, between phenotypic fluctuations and evolution speed. This theoretical prediction has been confirmed in an artificial selection experiment involving fluorescence in proteins in E. Coli (Sato et al. 2003). Since Fisher's theorem states the proportionality between evolution speed and phenotypic variance due to genetic variation, while the above recent study concerns evolution speed and intrinsic phenotypic fluctuations of clones, the question arises whether there should be some relationship between phenotypic variance by the distribution of genes (genetic variation) and phenotypic fluctuations of clones.

In general, however, a straightforward relationship between the two is not so easily expected since the genetic fluctuations generally depend on the mutation rate and the population distribution of organisms with different genes, while the phenotypic fluctuations of clones do not. The existence of a relationship, if there is any at all, must originate in some evolutionary constraints concerning stability. In the present paper, we propose a possible relationship between the two through a stability analysis of the two-variable probability distribution function both in phenotypes and genotypes, where the existence of such distribution itself is our most important

assumption. First, we give a possible inequality between the two fluctuations, and then, we show that for the mutation rate that achieves the highest evolution speed, this inequality is replaced by an equality. From this we derive the proportionality between the phenotypic variance due to the distribution of genes and the phenotypic variance of the clones of the most dominant genotype. Through these considerations, the recent fluctuation-response relationship in evolution and the Fisher's theorem are shown to be consistently related with each other. To confirm the validity of our theory we have carried out several simulations where a simple cell model consisting of a catalytic reaction network evolves. The relationship between phenotypic fluctuations of clones and those by genetic variations proposed theoretically is confirmed, as well as the evolutionary fluctuation-response relationship.

Remark

Note that the relationship between *phenotypic plasticity* and evolution has often been discussed. The term "phenotypic plasticity" usually refers to the changeability of the phenotype under a change of the environment. Here, however, we are concerned with the fluctuations of the phenotypes of clones, in a *given fixed* environment. Only recently there are some studies on it in relationship with the Baldwin effect (Ancel 1999, 2000; Ancel & Fontana 2002) or epigenetic inheritance (Pal and Miklos 1999).

2 Theoretical Formulation

2.1 Evolutionary Stability in the Distribution Function

Let us consider a phenotype x that is controlled by a gene, while the genetic change is assumed to be represented by a (scalar) parameter (variable) a, as is often adopted in the population genetics. For example, take the concentration of some protein x whose expression is coded by some gene, represented by a variable a, that is a Hamming distance in the genetic sequence from a given sequence. In the earlier paper (Sato et al. 2003) we have studied the fluctuations of the phenotypic variable x, in relationship with a genetic change given by a parameter a, where we have discussed the probability distribution $P(x;a)$. The most dominant value of the phenotype

x for a given genotype a is written by a function $x = x_0(a)$, which is the peak position of the distribution. The phenotype is distributed around $x = x_0$.

In considering the evolution, we need to consider also the distribution of genotype a, instead of regarding it as a given parameter. Through the evolutionary process, the dominant genotype a changes, and the dominant phenotype $x_0(a)$ also changes accordingly. Now, to consider the evolution both with regards to the distribution of phenotype and genotype, we introduce a two-variable distribution $P(x, a; h)$ with h a given environmental condition. Then by changing the environment h (or selection pressure in artificial evolution experiment), the most dominant genotype a and the phenotype change in accordance with the peak in this two-variable distribution $P(x, a; h)$. Existence of such distribution function is the first assumption here. The second assumption we make here is that at each stage of evolution, the distribution has a single (sharp) peak in this (x, a) space (see Fig.1). Through the evolutionary course, the dominant genotype a changes, which is given by the peak position of $P(x, a : h)$, that changes depending on the environmental condition h. This second assumption means a kind of evolutionary stability, i.e., at each stage of evolution, neither the phenotype nor the genotype distribution is spread out, and the distribution has a clear peak around a dominant gene and phenotype. This stability condition might be a bit strong to postulate for every evolutionary process, but as a gradual evolution (without speciation), it would be a reasonable assumption. In the artificial selection experiment, gradual change of phenotype could be reinterpreted as a gradual increase of selection pressure, to increase the parameter h (e.g., as increasing the threshold value beyond which a cell (or an organism) is selected so that it has a higher concentration of some protein or function). Or, by borrowing the term in thermodynamics it may be formulated as 'quasi-static process', where a transition process is represented as a successive change over equilibrium states.

Following the above discussion on the assumption of a single peak in the distribution, let us denote $P(x, a; h) = exp(-V(x, a; h))$. For a given h (environmental condition), let us denote the peak of the distribution by $a = a_0(h)$ and $x = x_0(a_0)$. As discussed, we assume evolutionary stability in the sense that a single dominant type of species exists in phenotype and genotype through the course of evolution. In other words, the distribution $P(x, a)$ has a single peak at

$(a_0, x_0(a_0))$. Then the stability condition is given by

$$(\partial^2 V/\partial a^2)^{-1} \geq 0 \ ; \quad (\partial^2 V/\partial x^2)^{-1} \geq 0. \tag{1}$$

For a two-variable probability function to have a single peak, however, another stability stability condition is necessary. It is given by the Hessian, i.e.,

$$(\partial^2 V/\partial x^2)(\partial^2 V/\partial a^2) - (\partial^2 V/\partial a \partial x)^2 \geq 0. \tag{2}$$

Now let us recall that $x_0(a_0)$ is the peak of the distribution of phenotype x for a clone of the genotype a_0. Then, the following condition has to be satisfied:

$$\partial V/\partial x|_{x=x_0} = 0 \tag{3}$$

Writing $\partial V(x, a_0 + \delta a)/\partial x|_{x=x_0+\delta x} = 0$, and expanding up to δa and δx, as $\partial^2 V/\partial x^2|_{x=x_0}\delta x + \partial/\partial a(\partial V(x,a)/\partial x)|_{x=x_0}\delta a = 0$, and $\delta x = (\partial x_0/\partial a)\delta a$, we then get

$$\partial^2 V/\partial x^2|_{x=x_0}(\partial x_0/\partial a)|_{a=a_0} + (\partial^2 V(x,a)/\partial a \partial x)|_{x=x_0} = 0. \tag{4}$$

Using this expression, the Hessian condition (eq.3) is rewritten as
$\partial^2 V/\partial x^2 \partial^2 V/\partial a^2 - (\partial x_0/\partial a)^2(\partial^2 V/\partial x^2)^2 \geq 0$, i.e.,

$$\partial^2 V/\partial a^2 - (\partial x_0/\partial a)^2(\partial^2 V/\partial x^2) \geq 0. \tag{5}$$

If the distribution is Gaussian-type, the variance of x and a around this dominant type are represented by $< (\delta a)^2 >= (\partial^2 V/\partial a^2)^{-1}$ and $< (\delta x)^2 >= (\partial^2 V/\partial x^2)^{-1}$. Using these expressions, we then obtain

$$V_g \equiv (\partial x_0/\partial a)^2 \ < (\delta a)^2 > \leq < (\delta x)^2 > \equiv V_{ip}. \tag{6}$$

Here all the quantities are computed around the peak of the distribution located at a_0 and $x_0(a_0)$ that depends on the environmental condition h. This inequality is expected to be satisfied for any stationary probability distribution given an environment h. By assuming that evolution passes through stationary states successively by generations, the conditions should be satisfied for each generation belonging to the evolution path.

Now we discuss the meaning of eq. 6. The right-hand side V_{ip} is nothing but the phenotypic variance of the clones (of the dominant genotype a_0), that is **i**ntrinsic **p**henotypic fluctuation, while the left-hand side is (*if we borrow the term in physics*) the genetic variance multiplied by the square of the phenotypic 'susceptibility' by genetic change. This left-hand side can be interpreted as the variance of the average phenotype distributed over genes (a around a_0), i.e., the phenotypic variance by the genetic variance, which is nothing but the quantity adopted in Fisher's theorem. This correspondence can be understood as follows: Assume that the variance of a is not so large and that the deviation of a from a_0 is small. Then the average phenotype of a clone with gene a is given by

$$\overline{x}_a \equiv \int x exp(-V(x,a))dx \approx \overline{x}_{a_0} + (a - a_0)\frac{\partial \overline{x}}{\partial a}. \tag{7}$$

where $\overline{x}_a$ is the average over x for the genotype a. Now, the variance of this average phenotype due to the gene distribution, given by $< (\overline{x}_a - \overline{x}_{a_0})^2 >$, is written as

$$< (\overline{x}_a - \overline{x}_{a_0})^2 > = < (\delta a)^2 > (\frac{\partial x_0}{\partial a})^2 = V_g, \tag{8}$$

which confirms the above interpretation. Summing up, the variance of the average phenotype over genes (denoted by V_g) should be smaller than or equal to the phenotypic variance of the clones of the dominant genotype (denoted as intrinsic phenotypic variance V_{ip}).

2.2 Derivation of the evolutionary fluctuation-response relationship

Generally speaking, even though the genetic variance in the population $< (\delta a)^2 >$ is not simply proportional to the mutation rate, it is nevertheless natural to expect that it increases with the mutation rate. In other words, the above condition eq. 6 gives a limit to the mutation rate, so that the population can maintain the peak around the fittest species.

Indeed, such a limit to the mutation rate for a single-peaked distribution has already been discussed for a fitness landscape where one type is the fittest among other neutral types by Eigen and Schuster and was called error catastrophe (Eigen & Schuster 1979). They have shown that there is a critical mutation rate beyond which the fittest organisms cannot maintain their dominance in the population, (i.e., the distributions lose the single peak and become flat). Since

our formulation assumes a continuous change of genes by mutation, and explicitly takes into account phenotypic fluctuations, it differs from the Eigen's argument, but the two are similar with regards to the limit on the mutation rate required to maintain the stability of the single-peaked distribution.

Now, consider a selection experiment to increase some phenotypic characteristic (e.g., to increase x by changing the condition h). If the mutation rate is increased, the evolution speed is expected to increase accordingly until it becomes too high, when the progressive evolution leading to successively better types no longer works since the above error catastrophe occurs giving a break-down of the single peak distribution. Hence the highest evolution speed is achieved just before this catastrophe occurs. This condition is given by simply replacing the inequality of eq.6 by an equality, i.e.,

$$(\partial x_0/\partial a)^2 < (\delta a)^2 > = < (\delta x)^2 >, \tag{9}$$

or $V_g = V_{ip}$. With this mutation rate, the evolutionary path passes through marginally stable states so that the phenotypic fluctuations due to the distribution of genes equals the phenotypic variance of the clones. (As for the marginal stability hypothesis for growth to a new state, the argument could also be related to that proposed for crystal growth by Langer (Langer 1980)).

Recall that Fisher's fundamental theorem of natural selection states that the evolution speed is proportional to V_g, the phenotypic variance by genetic variance. Then, at the state with optimal evolution speed, the above relationship means that the phenotypic variance of the clones is proportional to the evolution speed. When the mutation rate is smaller than this optimal value, the original inequality is satisfied. In this case, $< (\delta a)^2 >$ and accordingly V_g generally increase with the mutation rate μ. Now V_g is written as a function of mutation rate $f(\mu)$, which is an increasing function and $f(\mu_{max}) = V_{ip}$. Hence for given mutation rate μ, $V_g = \frac{f(\mu)}{f(\mu_{max})} V_{ip}$. Since V_g is proportional to the evolution speed, it is also proportional to V_{ip}, for a given mutation rate. Thus the fluctuation-response relationship in (Sato et al. 2003) is derived. Furthermore, for small mutation rate, the function $f(\mu)$ is expanded as $f'(0)\mu + o(\mu)$ as $V_g = 0$ for $\mu = 0$. Hence $V_{ip}\mu \propto V_g \propto$ (*evolution speed*) follows.

2.3 Remark on the Distribution

It should be stressed that the derivation of our expression uses only the stability condition, and it does not depend on specific mechanisms of evolution or selection pressure, which influences only on the proportion coefficient to the evolution speed. The assumption we made is the existence of $P(x,a) = exp(-V(x,a))$, (or in other words, the existence of the 'potential function' $V(x,a)$) and the evolutionary stability, where we assume that at each generation in the evolutionary course, the distribution has a single peak, i.e., stability both in genetic and phenotypic space.

Since "fitness" (or "selection") is not explicitly referred to in the above formulation, one might wonder if it is included here. Indeed, the selection process is included in the the potential $V(x,a)$. The evolution of the distribution function in the above formulation can be rephrased as follows: By changing an environmental condition h, the selection pressure is changed, and thus the $P(x,a) = exp(-V(x,a))$ changes accordingly. Hence, if we explicitly include the process of evolution under the change of environmental condition, it is written as $P(x,a;h) = exp(-V(x,a;h))$. With the change of h, the fittest (x,a), i.e., the peak of the distribution $P(x,a;h)$ changes accordingly. With the change of h, each generation passes through such stable states 'quasi-statically'. In the present derivation eq.(1)-(8), we have not explicitly written the environment h, but it is included implicitly therein. Or, it is also possible to explicitly introduce the environmental change as another "external field" h, e.g., by introducing $V(x,a;h) = V_0(x,a) + xh$ to obtain a change of the peak $(a_0,\ x_0)$ with a change of h. With this form, the inequality (6) is also straightforwardly obtained.

2.4 A simple example

To elucidate the above argument explicitly, it may be useful to give a simple example by taking a superposition of Gaussian distributions. We set

$$V(x,a) = \frac{1}{2}\{\alpha(a-a_0)^2 + \xi(a)(x-x_0(a))^2 - log\xi(a)\}, \tag{10}$$

where the last term $-log\xi(a)$ is required to assure the normalization of $P(x,a)$ when integrated over x. Noting that the peak in a is given by $a^* = a_0 + c'(a^*)/\alpha, x^* = x_0(a^*)$, with $c(a) = (1/2)log\xi(a)$, and $'$ as the derivative by a, it is straightforward to obtain $\frac{\partial^2 V}{\partial x^2} = \xi(a^*)$, $\partial^2 V/\partial a^2 =$

$\alpha - c''(a) + \xi(dx_0/da)^2$, and $(\partial^2 V/\partial a \partial x) = -\xi(dx_0/da)$. Then it is straightforward to confirm the rewriting of the Hessian stability condition by eq.(5) which leads to

$$\alpha - c''(a^*) > 0. \tag{11}$$

In these expressions, one should note that $< (\delta a)^2 >^{-1} = \frac{\partial^2 V}{\partial a^2}$ is not equal to α. The variance of the genotype is shifted due to the a-dependence of the variance of x. Here, when considering the adopted Gaussian form $exp\{-V(x,a)\}$, it may be natural to assume that α^{-1} is proportional to the mutation rate, and accordingly that the variance of a deviates from it. Note that if $c''(a) > 0$, then for small α (corresponding to a large mutation rate μ), this Hessian condition is violated, and the critical value gives the maximal mutation rate μ_{max} discussed above. Recalling that $\xi(a)$ is the inverse of the variance of x, this loss of stability generally occurs when log(phenotypic variance) is convex at $a = a^*$.

3 Model Simulation

As an illustration of the above theory we consider a simple cell model with intra-cellular catalytic reactions allowing for cell growth and division. We employ a simple reaction network model studied earlier (Furusawa & Kaneko 2003; Furusawa et al. 2005), consisting of a variety of chemicals whose concentrations are given by $(c_1, c_2, \cdots, c_K)$, for K chemical species in a cell. Depending on whether there is an enzymatic reaction from i to j catalyzed by some other chemical ℓ or not, a reaction path is connected in the network as $(i + \ell \to j + \ell)$. Here, some chemicals including nutrients (that have no catalytic activity) are transported through the cell membrane with the aid of some other chemicals, that are 'transporters'. Transported nutrients are successively transformed to other chemicals through catalytic reactions, including transporters. When these reactions progress due to the flow of nutrients, the number of molecules in a cell increases, until it goes beyond a given threshold N and the cell divides into two. The model, in spite of its simplicity, is found to capture universal statistical behaviors of a cell as has also been confirmed in experiments(see for details (Furusawa & Kaneko 2003; Furusawa et al. 2005)).

Of course, how these reactions progress depends on the network. Here, we carry out evolution

experiments where those reproducing cells are selected that have a higher concentration c_{i_s} of a given specific chemical i_s. To be specific, we take n parent cells which evolve such that they grow recursively, starting from a catalytic network chosen randomly. From each parent cell, L mutant cells are generated by randomly adding or removing m reaction paths to the reaction network of the parent cell. Thus the mutation rate μ is given by m/M, with M the total number of paths. For each of the networks, the reaction dynamics are simulated to identify cells that continue reproduction. Among such networks the top n cells with regards to the abundances of the chemical species i_s are selected for the next generation.

As the number of molecules is finite, there are fluctuations in the abundances of each chemical. Indeed, the simulation of the reaction process is carried out by picking up molecules stochastically. Hence, for a given network, there are fluctuations in the abundances of the chemicals. Corresponding to the variable x in the theory is the concentration c_{i_s}. To be precise, we choose $x = log(c_{i_s})$, as the distribution of c_i is close to the log-normal distribution (as is also true experimentally (Sato et al. 2003; Furusawa et al. 2005), since our theory is better applied for a variable x whose distribution is close to a Gaussian distribution. (Note that logarithm of the concentration is adopted also for an experimental confirmation of the evolutionary fluctuation-response relationship in (Sato et al. 2003). As V_{ip}, we compute the variance of x, for a network that gives the peak abundances at each generation. On the other hand, the network itself is regarded as the 'genotype'. The variance V_g is computed, from the variance of the distribution of x, over the L mutant networks at each generation.

The evolutionary changes in the phenotype distribution of the clones $P(x)$ and that of the genetic variation $P(a)$ are plotted in Fig.2. As shown, the distributions evolve jointly, satisfying $V_{ip} > V_g$ as expected from the theory. Quantitatively, we first check the validity of the fluctuation-response relationship (Sato et al. 2003) that is between the evolution speed and V_{ip}. We have plotted the increment of the phenotype x (i.e., logarithm of the concentration of the selected species i_s) at each generation for the selected species successively. As shown in Fig.3, the data plotted against V_{ip} are fitted well by a linear relationship.

Next, we have plotted V_g versus V_{ip} in Fig.4. We found that the expected inequality is satisfied, and also that for each evolutionary process with a fixed mutation rate, $V_{ip} \propto V_g$ holds. As the

mutation rate μ increases, the slope of V_g/V_{ip} increases, and it approaches the diagonal line $V_g = V_{ip}$. On the other hand, with the increase of μ, mutant populations exhibiting very low values of x (the abundances of selected species i_s) increase, which corresponds to the collapse of catalytic reaction process for cell growth. As shown in Fig.5, beyond some mutation rate μ_{thr}, the distribution becomes flat, and the peak in the distribution starts to shift downwards. By comparing the $V_g - V_{ip}$ relationship, we confirmed that around $\mu \approx \mu_{thr}$, the relationship approaches the diagonal line $V_g \approx V_{ip}$. Indeed, for a higher mutation rate allowing for $V_g > V_{ip}$ the evolution does not progress, as the distribution is almost flat, and the value of x after selection cannot increase by generations. Thus the evolution speed is optimal at around $\mu \approx \mu_{thr}$. Summing up, there is a threshold mutation rate μ beyond which the evolution does not progress (i.e., an error catastrophe occurs), where V_g approaches V_{ip}. All of these numerical results support the theoretical prediction described earlier. It should also be stressed that the result here does not depend on the specific algorithm for evolution, such as the ratio of selected networks for the next generation. The selection pressure, given by the fraction of selected networks, influences only on the proportion coefficient between the evolution speed and V_g as predicted by Fisher's theorem, but it does not influence the relationship between V_g and V_{ip}.

4 Discussion

In summary, we have proposed an inequality between phenotype variation over genetically different individuals, V_g and the (intrinsic) phenotypic variance of clones V_{ip}, as a result of the stability of their distributions. It is found that there is an optimal mutation rate beyond which this inequality is violated, leading to the collapse of the single-peak distribution around the fittest type. The evolution speed is bounded by this optimal mutation rate, at which the genetic variance measured by phenotypes approaches the phenotypic variance of clone. Following this argument, a linear relationship (or correlation) between the two variances is implied, leading to the evolutionary fluctuation-response relationship in (Sato et al. 2003). Now, the consistency between Fisher's theorem and the evolutionary fluctuation-response relationship is demonstrated. By taking a simple cell model with a catalytic reaction network, this proposition is numerically confirmed.

Since in this model, V_g is the fluctuation over different mutated networks, and V_{ip} is the fluctuation through the dynamics for a given single network, there is no a priori reason that the two should be correlated. Still, we have found the linear relationship between the two as a result of the course of evolution.

Our theory is based on the existence of the probability distribution $P(x, a)$. This is not obvious at all. Since the phenotype is a function of gene, existence of two-variable distribution is a tricky assumption, because genetic and phenotypic fluctuations there are treated in the same way. Hence, the confirmation of the theory by numerical experiment is not trivial at all. Although we have not yet completely clarified why the theory is valid here, one possible reason for it is that some genetic change can give rise to the same effect to the reaction dynamics with that by the phenotypic fluctuation. For example, consider reaction process i to j catalyzed by ℓ. If the concentration c_ℓ changes by fluctuation, the rate of the reaction changes, while such change in the rate can be resulted by changing the path in the network or the catalytic activity of the reaction, which are coded by gene. In this sense there exists some genetic change that corresponds with the phenotypic change by fluctuation.

Indeed, Waddington, in his pioneering study, proposed the concept 'genetic assimilation' (Waddington 1957), in which the phenotypic change by the environmental change is later 'assimilated' by genetic changes. Here we should note that the degree of such phenotypic change as a response to the environmental change is also correlated with the fluctuations of the phenotype as a result of fluctuation-response relationship (Sato et al. 2003). Then by using the linear relationship between the evolution speed and phenotypic fluctuations, it is naturally expected that the evolution speed is correlated with the so-called phenotypic plasticity (Callahan, Pigliucci, & Schlichting 1997; West-Eberhard 2003), the response rate of the phenotype against environmental change. Based on our present study, it will be possible to reformulate Waddington's genetic assimilation or the Baldwin effect (Baldwin 1896; Bonner 1980; Ancel & Fontana 2002) quantitatively, and study the evolutionary process in terms of the plasticity (de Visser et al. 2003) measured through phenotypic fluctuations.

Another assumption in our theory is the use of (scalar) variable for the genetic change. In general, genetic change occurs in a very high-dimensional space. The choice of a (or a few) suitable

parameter(s) a is not trivial at all, whose validity should be examined in the future. In this sense also, the confirmation of the theory by the numerical model, presented here, is not obvious, since the number of degrees of freedom in the model is huge, as the change in the network paths has a huge variety of directions. In spite of this, this assumption on the existence of a and $P(x, a)$ seems to be valid in the numerical model we discussed. In the artificial selection experiment where fitter types are chosen successively, evolution can occur by accumulating mutations successively, so that a one-dimensional path along the evolutionary course acts as a parameter describing the process of the evolution. Indeed, the fitness landscape over one or a few parameters is also adopted generally in such theoretical or experimental studies of evolution (Kauffman & Levin 1987; Eigen, McCaskill, & Schuster 1989; Aita & Husimi 1996).

The inequality $V_{ip} > V_g$ concerns with the phenotypic variances due to genetic and non-genetic origin. If these variance are independent and added naively, the total phenotypic variance we observe from the wild-type population could be represented as $V_{tot} = V_{ip} + V_g$, (or $V_{ip} + V_g + V_e$, when the phenotypic variance due to environmental fluctuation V_e is also added independently). In this simple representation, the inequality we propose implies that phenotypic fluctuation by genetic variation is smaller than that of the non-genetic origin, i.e., V_g/V_{tot} is less than half. In the standard population genetics, the phenotypic variance of non-genetic origin is attributed to environmental factors, V_e, while the ratio to V_g to $V_{tot} = V_g + V_e$ is defined as the heritability h^2 (Futsuyma 1986; Maynard-Smith 1989). However, "intrinsic" phenotypic fluctuation through 'developmental noise' (Spudich & Koshland 1976; West-Eberhard 2003; de Visser et al 2003), as discussed here, contributes to the total phenotypic fluctuation as well. Although we need careful re-consideration on correlation among the variances and also on the quantitative estimate of V_g, it will be important to discuss the heritability in the light of our inequality with regards to the intrinsic phenotype fluctuation. Furthermore, studies on the environmental fluctuations and extension to multiple traits are necessary in future. It should also be noted that most of theoretical argument presented here is still valid under sexual recombination, while only the argument on the relationship between V_g and the mutation rate needs reconsideration.

Of course it is very important to verify our proposition experimentally. One good candidate is an artificial selection experiment by using bacteria, adopted in (Sato et al. 2003; Ito et al. 2004).

Indeed, preliminary data from the experiments seem to suggest the validity of our proposition.

Acknoledgement

We would like to thank T. Yomo, K. Sato, Y. Oono, M. Kikuchi, T.Sato, T. Ikegami, Y. Takagi, I.Tsuda, E. Szathmary, Y. Iwasa, and F. Willeboordse for criticisms and discussions.

References

[1] Aita T. and Husimi Y. 1996 Fitness Spectrum Among Random Mutants on Mt. Fuji-type Fitness Landscape. *J. theor. Biol.* 182, 469-485.

[2] Ancel L.W. 1999 A quantitative model of the Simpson-Baldwin Effect. *J. Theor. Biol.* 196, 197-209.

[3] Ancel L.W. 2000 Undermining the Baldwin expediting effect: does phenotypic plasticity accelerate evolution? *Theor Popul Biol.* 58, 307-19.

[4] Ancel L.W. and Fontana W. 2002 Plasticity, Evolvability, and Modularity in RNA. *J. Exprimental Zoology* 288, 242-283.

[5] Baldwin J.M. 1896 A new factor in evolution. *Am. Nat.* 30, 441-451.

[6] Bonner J.T. 1980 The Evolution of Culture in Animals, Princeton Univ. Press.

[7] Callahan H.S., Pigliucci M., and Schlichting C.D. 1997 Developmental phenotypic plasticity: where ecology and evolution meet molecular biology. *Bioessays* 19, 519-525.

[8] Edwards, A.W.F. 2000 Foundations of mathematical genetics, Cambridge University Press.

[9] Eigen M., McCaskill J., and Schuster P. 1989 The Molecular Quasi-species. *Adv. Chem. Phys.* 75, 149-263.

[10] Eigen M. and Schuster P. 1979 The Hypercycle, Springer.

[11] Elowitz, M. B., Levine, A. J., Siggia, E. D. & Swain, P. S. 2002 Stochastic gene expression in a single cell. *Science* 297, 1183-1186

[12] Fisher, R. A. 1930;1958 The Genetical Theory of Natural Selection, Oxford Univ. Press.

[13] Furusawa C. and Kaneko K. 2003 Zipf's law in Gene Expression. *Phys. Rev. Lett.* 90, 088102.

[14] Furusawa C., Suzuki T., Kashiwagi A., Yomo T., and Kaneko K. 2005 Ubiquity of Log-normal Distributions in Intra-cellular Reaction Dynamics. *BIOPHYSICS* 1, 25-31.

[15] Futsuyma D.J. 1986 Evolutionary Biology (Second edition), Sinauer Associates Inc., Sunderland

[16] Hasty, J., Pradines, J., Dolnik, M. & Collins, J. J. 2000 Noise-based switches and amplifiers for gene expression. *Proc. Natl. Acad. Sci. USA* 97, 2075-2080.

[17] Ito, Y., Kawama, T., Urabe, I. & Yomo, T. 2004 Evolution of an arbitrary sequence in solubility. *J. Mol. Evol.* 58, 196-202.

[18] Kauffman S and Levin S. 1987 Towards a general theory of adaptive walks on rugged landscapes. *J. Theor. Biol.* 128, 11-45.

[19] Kimura M. 1983 The neutral theory of molecular evolution, Cambridge Univ. Press.

[20] Kubo R., Toda M., Hashitsume N. 1985 Statistical Physics II: (English translation; Springer).

[21] Langer J.S. 1980 Instabilities and pattern formation in crystal growth. *Rev. Mod. Phys.* 52, 1-28.

[22] Maynard-Smith J. 1989 Evolutionary Genetics, Oxford Univ. Press.

[23] Pal C., Miklos I. 1999 Epigenetic inheritance, genetic assimilation and speciation. *J. Theor. Biol.* 200, 19-37.

[24] Sato K., Ito Y., Yomo T., and Kaneko K. 2003 On the Relation between Fluctuation and Response in Biological Systems. *Proc. Nat. Acad. Sci. USA* 100, 14086-14090.

[25] Spudich J.L. and Koshland D.E.Jr. 1976 Non-genetic individuality: chance in the single cell. *Nature* 262, 467-471

[26] Ueda, M., Sako, Y., Tanaka, T., Devreotes, P., & Yanagida, T. 2001 Single-molecule analysis of chemotactic signaling in Dictyostelium cells. *Science* 294, 864-867

[27] de Visser J .A.G.M. et al. 2003 Evolution and detection of genetic robustness. *Evolution* 57, 1959-1972.

[28] Waddington C. H. 1957 The Strategy of the Genes, Allen & Unwin, London.

[29] West-Eberhard M.J. 2003 Developmental Plasticity and Evolution, Oxford Univ. Press.

Figure Legends

Fig. 1

Schematic representation of the two variable distribution $P(x, a)$ on phenotype x and genotype a. The thick curve shows $x_0(a)$.

Fig. 2

Histogram of the phenotype x, that is the logarithm of the concentration c_{i_s} in the reaction-net cell model. Distributions at 5 generations in the course of evolution with mutation rate $\mu = 0.01$ are plotted with changing colors. (a) distribution of the phenotype $x = log(c_{i_s})$ of the selected clones at each generation. The distribution if obtained from 10000 cells of the clone. (b) distribution of the phenotype $x = log(c_{i_s})$ over 10000 mutants from the selected clones at each generation. The phenotype is computed as the average over 200 cells of each clone.

Fig. 3

Evolution speed versus the phenotype fluctuation of a clone. The evolution speed is measured by the difference between the phenotypes x of the two successive generations, while the fluctuation is measured by the variance of x of the clone of the selected cells, at each generation, computed by the distribution over 200 cells of the clone.

Fig. 4

The relationship between V_{ip} the variance of the phenotype x of the clone as measured in Fig.2, and V_g, the variance of the phenotype x over 10000 mutants from the selected cell. Plotted over generations in each course of evolution, given by a fixed mutation rate displayed in the figure.

Fig. 5

Distribution of the phenotype x over 10000 mutants, generated with the mutation rates 0.003, 0.01, 0.02, 0.03 and 0.05. Around the mutation rate 0.03, the distribution is flattened, and the peak position starts to shift downward.

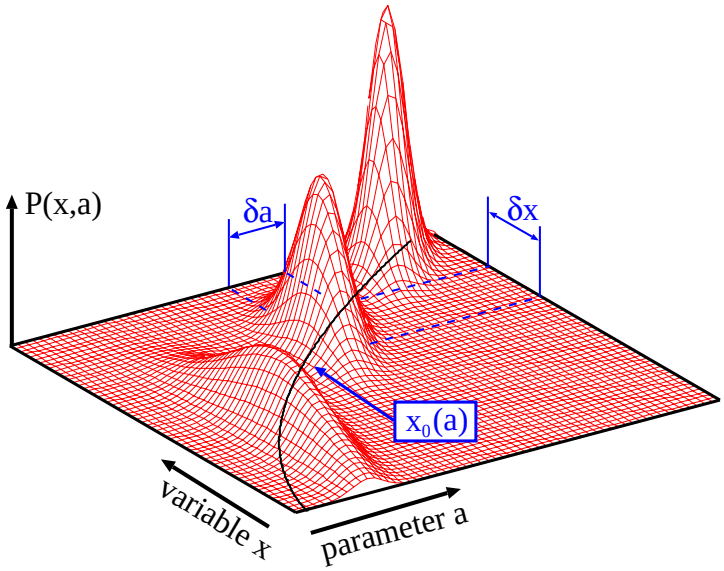

Figure 1:

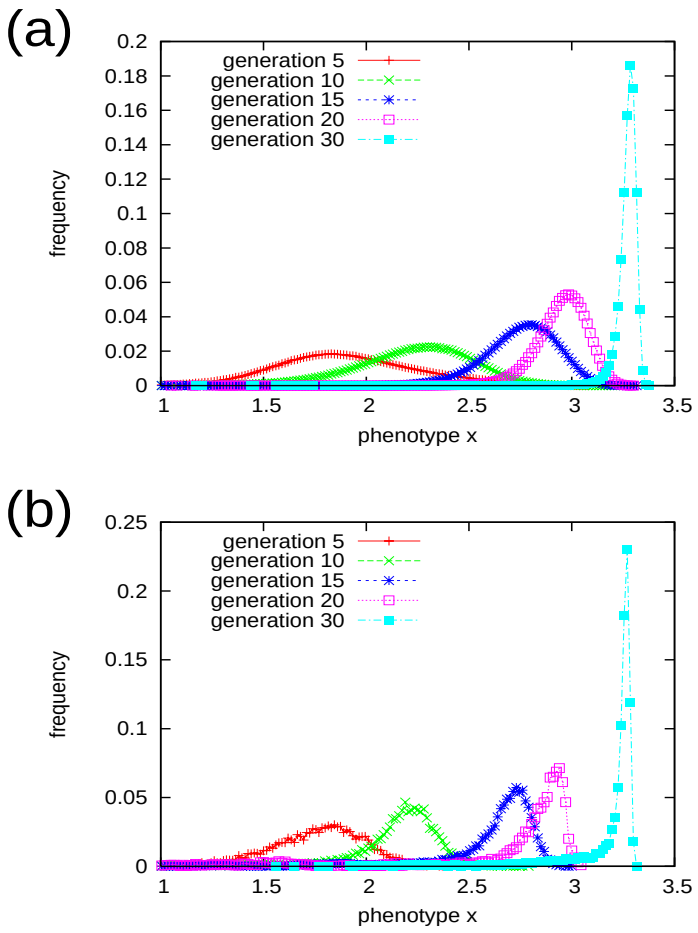
(a)
generation 5
generation 10
generation 15
generation 20
generation 30
frequency
phenotype x
0.2
0.18
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
1
1.5
2
2.5
3
3.5
(b)
generation 5
generation 10
generation 15
generation 20
generation 30
frequency
phenotype x
0.25
0.2
0.15
0.1
0.05
0
1
1.5
2
2.5
3
3.5

Figure 2:

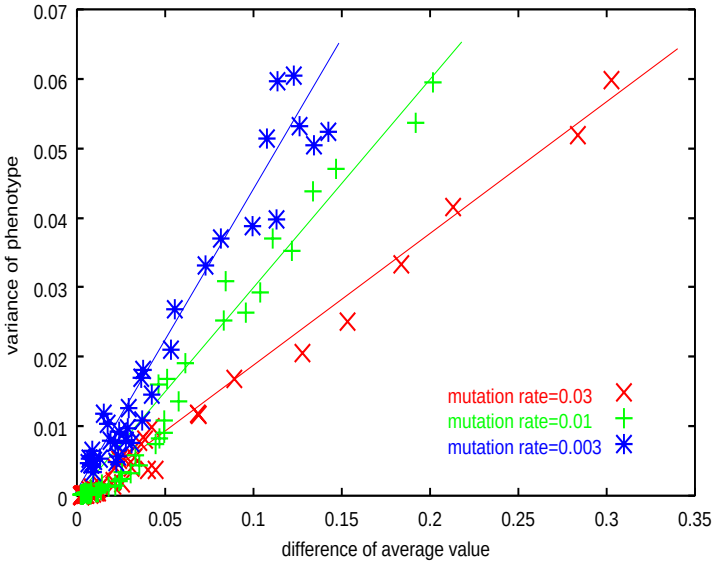

Figure 3:

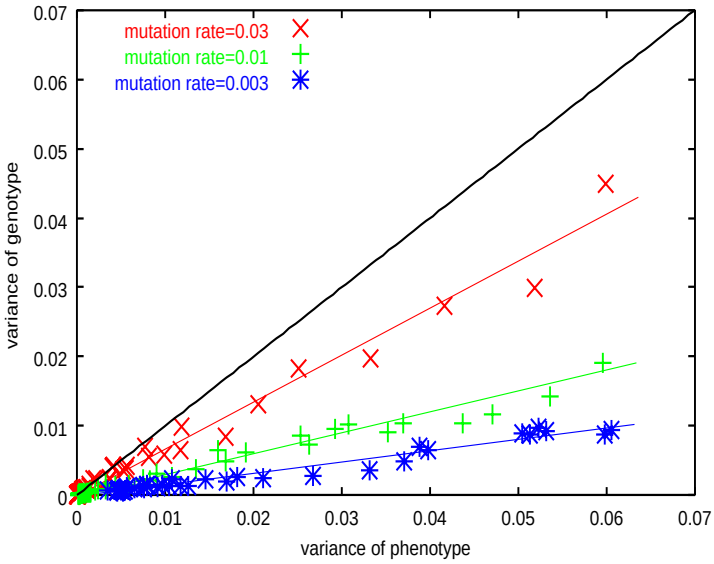

Figure 4:

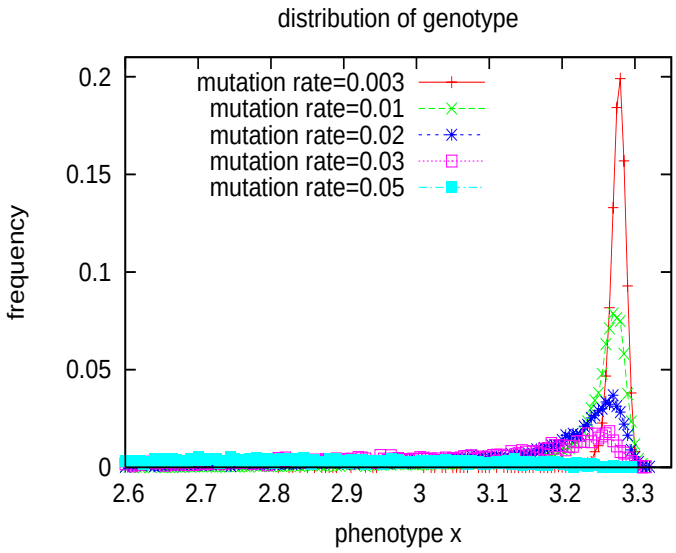
distribution of genotype
mutation rate=0.003
mutation rate=0.01
mutation rate=0.02
mutation rate=0.03
mutation rate=0.05
frequency
0.2
0.15
0.1
0.05
0
2.6
2.7
2.8
2.9
3
3.1
3.2
3.3
phenotype x

Figure 5: