Global relationships in fluctuation and response in adaptive evolution

Chikara Furusawa¹ and Kunihiko Kaneko²

¹Quantitative Biology Center, RIKEN, 6-2-3 Furuedai, Suita, Osaka 565-0874, Japan
²Research Center for Complex Systems Biology, University of Tokyo, Komaba, Meguro-ku, Tokyo 153-8902, Japan

Cells change their internal state to adapt to environmental changes, and evolve in response to the new conditions. The phenotype changes first via adaptation in response to environmental changes, and then through mutational changes in the genomic sequence, followed by selection in evolution. Here, we analysed simulated adaptive evolution using a simple cell model consisting of thousands of intracellular components, and found that the changes in their concentrations by adaptation are proportional to those by evolution across all the components, where the proportion coefficient between the two agreed well with the change in the growth rate of a cell. Furthermore, we demonstrate that the phenotypic variance in concentrations of cellular components due to (non-genetic) noise and to genomic alternations is proportional across all components. This implies that the specific phenotypes that are highly evolvable were already given by non-genetic fluctuations. These global relationships in cellular states were also supported by phenomenological theory based on steady reproduction and transcriptome analysis of laboratory evolution in Escherichia coli. These findings demonstrate that a possible evolutionary change in phenotypic state is highly restricted. Our results provide a basis for the development of a quantitative theory of plasticity and robustness in phenotypic evolution.

1. Introduction

Biological systems change their state to evolve and adapt to changes in environmental conditions. Despite the recognized importance of characterizing the biological capacity to adapt and evolve, studies on biological evolvability and plasticity have thus far remained at a qualitative, rather than quantitative, level. After an environmental change, cells may first respond by changing the abundance of cellular components, including proteins and metabolites, without changing the genome sequence. The typical timescale of such environmental adaptation is generally shorter than several generations. However, over the long-term, i.e. over many generations, the abundance of cellular components is gradually changed by evolutionary dynamics, in which the genome sequence is altered by mutations and individuals with higher fitness are generally selected. This leads to the important question of whether there is a general relationship between short-term phenotypic changes in adaptation and long-term phenotypic changes in evolution. Note, however, that these two types of phenotypic changes occur over distinct timescales and are generally caused by different mechanisms. Thus, any relationship between them, if its existence is confirmed, would be non-trivial [1–3].

This question can now be addressed experimentally from a quantitative perspective. Recent advances in high-throughput experimental analysis enable us to characterize cellular state in terms of abundances of a huge number of components [4–6], which constitute a high-dimensional state space. The high-dimensional phenotypic changes due to environmental change or to changes in genomic sequence throughout the course of evolution are now measurable separately, for example, by using transcriptome and genome resequencing analysis. Such phenotypic and genotypic data are available, for example, from the results obtained from the experimental evolution of Escherichia coli [7–11].

© 2015 The Authors. Published by the Royal Society under the terms of the Creative Commons Attribution License http://creativecommons.org/licenses/by/4.0/, which permits unrestricted use, provided the original author and source are credited.
Considering the availability of these quantitative data, the above question we posed can be rephrased to ask whether a quantitative relationship between such high-dimensional phenotypic changes exists, one due to adaptation in a single generation and the other due to genetic evolution over generations. Owing to the high-dimensionality in phenotypic changes as well as the difference in timescales between adaptation and evolution, one might assume that it would be difficult to adequately address this question. However, our recently published theoretical study, as well as other recent studies, suggests the possible existence of such a relationship [12–16].

It should be noted that the essence of cellular dynamics is reproduction in which the abundance of each cellular component is roughly doubled, and this constraint imposed by cellular reproduction imposes a restriction on the time development in phenotype. Such a restriction can be represented by a high-dimensional cellular state space, in which each axis represents the abundance of each cellular component. Here, it is possible that a cellular state would be restricted to a sub-space in the high-dimensional state space, described by a relatively smaller number of variables. Indeed, in non-genetic adaptation, we have experimentally confirmed that such a restriction to low-dimensional dynamics provides a global relationship in the change in the expression levels across thousands of genes [12].

Whether such restrictions also provide a non-trivial link between the phenotypic changes occurring in non-genetic adaptation and the long-term changes occurring over the course of evolution is a related, valid question. Some experimental studies have suggested that there is a common trend observed over the thousands of gene expression changes in adaptation and evolution, in which genes whose expressions exhibit a larger response to environmental change tend also to show a larger response in their expression at the evolutionary scale [14–16].

It is naturally expected that few macroscopic quantities extracted from global cellular behaviour, such as the growth rate and fitness of the system, play dominant roles in adaptation and evolution, which govern the entire (high-dimensional) dynamics of phenotypic change. Therefore, there may be a possible link between high-dimensional phenotypic dynamics as represented by gene expression profiles and a macroscopic variable such as the growth rate, which may lead to a relationship between the expression changes due to non-genetic adaptation and genetic evolution. The purpose of the present study is thus to investigate such a relationship by in silico and in vitro evolution, as well as by theoretical formulation.

The relationship between phenotypic changes of non-genetic and genetic origins has also been discussed in the fluctuations in phenotype. As has been recently uncovered [17–22], the abundances of cellular components generally exhibit fluctuations even without genomic alternations, which originate from the stochastic nature of intracellular chemical reactions. The proportionality between such non-genetic fluctuations and the evolutionary rate of fitness has been demonstrated in bacterial experimental evolution and in simulations of toy-cell models, which are supported by phenomenological theory [23–26]. The relevance of the non-genetic fluctuation to environmental adaptation [27] and robustness of regulatory networks [25,28] has also been analysed. Studies to date have primarily focused on the fluctuation and evolutionary response of fitness. However, considering a possible link between the fitness and abundances of thousands of components through evolution, there possibly exists a relationship between non-genetic fluctuations and evolutionary variation of component abundances, which warrants further investigation. Here we examine proportionality between phenotypic fluctuations of component abundances due to non-genetic noise and due to genetic changes.

The suggested proportionality between the response of cellular states to environmental change and genetic change, as well as between the response and fluctuations of component abundances, if confirmed, implies the proportionality among two-by-two quantities, namely, fluctuations and responses induced by environmental (noise) and genetic changes (mutation). The goal of the present study is to explore this grand relationship. This should be of critical importance to evolutionary biology, as it would indicate that if the abundance of a component (or phenotypic trait) is more variable by environmental changes, it is also variable by genetic evolution. Thus, a general restriction for the direction of phenotypic evolution in a high-dimensional space is stipulated. Specifically, we demonstrate a theoretical basis for the quantitative study of the phenotypic plasticity underlying adaptive evolution.

To date, experimental confirmation of this grand relationship remains premature, and further analysis is required, such as by using transcriptome analysis. Accordingly, it is important to examine such a relationship by employing an integrative approach combining in silico evolution of a cell model consisting of thousands of cellular components, laboratory evolution of bacteria under environmental stress, and phenomenological theory for time development in a high-dimensional state space. In this study, we aim to uncover statistical laws underlying the fluctuation and response of high-dimensional state variables occurring through adaptive evolution and to connect them with changes in growth rate or fitness.

2. Results

2.1. Evolutionary simulations of a simple cell model

We employed our previously established mutually catalytic reaction network model, as this model is capable of capturing the basic characteristic of cells such as the power-law abundance and lognormal fluctuations of cellular components, fluctuation–response relationship of fitness, adaptation with fold-change detection, and so forth, in spite of its simplicity [20,24,29,30]. In the model, the cellular state is represented by numbers of $K$ chemical species, and the internal chemical reaction dynamics is governed by a catalytic network among these chemical species. Some resources (nutrients) are supplied from the environment by transportation through the cell membrane with the aid of some other chemicals that are termed ‘transporters’. Through catalytic reactions, these nutrients are transformed into cell-component chemicals, and a cell divides when the amount of component chemicals reaches a certain threshold. Here, to achieve a higher growth rate, the synthesis of the cell components, transporters and chemicals that catalyse the synthesis of those components need to be harmonized with the nutrient uptake. We allowed the above toy-cell model consisting of catalytic reaction
networks to evolve by rewiring the network paths with a given mutation rate and selecting the pathways with a certain fraction of cells that showed a higher growth rate (see Material and methods for details). For a given environmental condition, evolution progresses so that the cell growth rate, i.e. the inverse of the average division time, is increased (figure 1). To study the response to environmental change, we then switched the nutrient condition after evolution under a fixed condition for 3000 generations (denoted by the arrow in figure 1). The growth rate initially decreased following this environmental change, and then recovered through genetic evolution over generations. Next, we explored the phenotypic state changes in response to the environmental and evolutionary changes to evaluate the relationship between non-genetic and genetic responses to environmental change.

As phenotypic state variables for the cell, we computed the abundances of each chemical $N_i$ at the division event. Here, it is convenient to choose $X_i = \log N_i$ as a phenotypic variable, as the abundance generally increases exponentially over time through cellular growth, and perturbation in a network is also generally amplified exponentially. Indeed, this choice of logarithmic abundances is also relevant to the theoretical argument presented below, as well as to transcriptome analysis of gene expression. Note also that the abundances $N_i$ are distributed by cells, even for those sharing the same reaction network, due to stochasticity in reaction dynamics; thus, the average abundance over all cells is required to study the mean response of cells, denoted by $\langle \cdots \rangle$.

After the change in nutrient condition, the abundances of all the components change. Let us denote the average change of these abundances by:

$$\delta X_i^{\text{Env}} = \langle X_i(1) \rangle - \langle X_i(0) \rangle = \log \left( \frac{\langle N_i(1) \rangle}{\langle N_i(0) \rangle} \right),$$

where generation 1 refers to the time point immediately following the environmental change, and generation 0 denotes the generation right before this nutrient change. Similarly, we define the response by genetic evolution after $m$ generations by $\delta X_i^{\text{Gen}}(m) = \langle X_i(m) \rangle - \langle X_i(0) \rangle$. Figure 2 shows the plot of $\delta X_i^{\text{Env}}$ versus $\delta X_i^{\text{Gen}}(m)$ for $m = 5, 10$ and 50. Interestingly, proportionality was found between the environmental and genetic responses over all components.

Let us now define this proportion coefficient $r(m)$ for $\delta X_i^{\text{Gen}}(m) / \delta X_i^{\text{Env}}$ across components $i$. This proportion coefficient $r(m)$ is initially close to 1, but with the increase in generations $m$, it decreases towards zero, in conjunction
with the recovery of the growth rate. In other words, evolution shows a common tendency to reduce the changes in components introduced by the environmental change. This common proportionality across all chemicals suggests that the proportion coefficient \( r(m) \) is a ‘global variable’ over a huge number of chemical species. A reasonable candidate for such a global variable is the cell growth rate \( \mu \). Hence, it is natural to compare the coefficient \( r(m) \) with the growth rate. Toward this end, we again computed the change in the growth rate \( \delta \mu^\text{Env} = \mu(1) - \mu(0) \) and \( \delta \mu^\text{Gen}(m) = \mu(m) - \mu(0) \) at the \( m \) generation. The ratio \( \delta \mu^\text{Gen}(m)/\delta \mu^\text{Env} \) gives an index for the recovery in this growth rate from the decrease caused by the environmental change, with 0 and 1 representing full and null recovery, respectively. In figure 3, the proportion coefficient \( r(m) \) is plotted against this growth rate recovery \( \delta \mu^\text{Gen}(m)/\delta \mu^\text{Env} \). The proportionality between the two is clearly discernible.

Recalling the possible relationship between fluctuation and response, as is typical in statistical physics, we then evaluated whether there also exists a common relationship among the variances of all the components. Here, as previously reported [20], the distribution of each \( N_i \) follows an approximately lognormal distribution, as confirmed experimentally in the protein abundances in the present cells. Hence, it is again relevant to adopt \( X_i = \log N_i \) as a phenotype variable [31], so that the distribution of \( X_i \) follows a roughly Gaussian distribution [23]. The phenotypic variance \( V_{ip}(i) \) for each component \( i \) is defined as the variance of \( X_i \) in an isogenic population. On the other hand, the variance due to genetic change \( V_{ig}(i) \) is defined as the variance of mean \( X_i \) over a heterogenic distribution, where the mean is computed across clones of a given genotype (i.e. a network), while the heterogenic distribution is related to different genotypes (networks) that exist at a given generation.

The results of simulations are given in figure 4, which shows proportionality between \( V_{ip}(i) \) and \( V_{ig}(i) \) across the components \( i \) for the evolved population. As the mutation rate is increased over evolution, \( V_{ip}(i) \) increases as the genotype distribution is broadened, whereas \( V_{ig}(i) \) remains at the same level, so that the ratio \( V_{ip}(i)/V_{ig}(i) \) is increased while maintaining the proportionality. This observed proportionality means that the components that are more variable owing to noise in the reaction dynamics are also more variable owing to mutation.

So far, we have confirmed the existence of common proportionality between non-genetic and genetic variances, as well as between the environmental and evolutionary responses. As the proportionality between the fluctuation and response is a natural outcome in statistical physics, we compared the response and fluctuations in more detail. However, direct comparison of the phenotypic variance \( V_{ip}(i) \) with the environmental response did not show a clearly discernable proportionality. This is probably because of the discrepancy in the definitions of the two quantities: the variance originates from very high-dimensional dynamics without any specific directional change, while in the environmental response, only one specific environmental change considering only a few nutrients is applied. To make a more direct comparison, we then sampled the environmental responses against a variety of external changes introduced by different nutrient conditions to define the average environmental response \( R^\text{Env}(i) = \langle (\Delta X^\text{Env})^2 \rangle \) with \( \langle \cdots \rangle \) over \( 10^4 \) environmental conditions (see Material and methods). As shown in electronic supplementary material, figure S1, this average environmental response showed clear proportionality with \( V_{ip}(i) \) and \( V_{ig}(i) \), respectively. Hence, the proportionality relationships among two-by-two quantities, i.e. genetic and non-genetic responses and fluctuations, hold all components (figure 5).

**Figure 3.** The relationship between growth recovery rate \( \delta \mu^\text{Gen}(m)/\delta \mu^\text{Env} \) and the proportion coefficient \( r(m) \). The proportion coefficient \( r(m) \) was obtained by using the least-squares method for the relationship of \( \Delta X^\text{Gen} \) and \( \Delta X^\text{Env} \) for \( m = 1 \) to 200 (see the blue lines in figure 2 for examples). The black solid line is \( y = x \) for reference. (Online version in colour.)

**Figure 4.** The relationship between \( V_{ip}(i) \) and \( V_{ig}(i) \). The variances were computed by using the network and environment at generation 0 (before the environmental change) shown in electronic supplementary material, figure S1 with various mutation rates. \( V_{ip}(i) \) and \( V_{ig}(i) \) were calculated based on the simulation results of randomly generated \( 10^6 \) networks. The solid line is \( y = x \) for reference. (Online version in colour.)

**Figure 5.** Proportionality relationships among genetic/non-genetic fluctuations and responses hold over all components. The arrows indicate the proportional relationships.
Thus, the degree of plasticity required to achieve an adaptive response to a new environment is characterized by phenotypic fluctuations $V_{ij}(t)$, i.e. those that do not consider environmental or genetic changes. On the other hand, when cells are exposed to a novel environment, the potential of adaptation is expected to increase. Therefore, when placed in a novel condition, it is expected that the phenotypic fluctuations would increase to allow the cells to adapt to the new environment. In electronic supplementary material, figure S2, the variances ($V_{ij}(t)$) are plotted before and after the environmental change (arrow in figure 1). In this case, all of the variances increase while roughly maintaining their proportionality. After this increase, the variances decrease over generations under a fixed environmental condition, while the proportionality between $V_{ij}(t)$ and $V_{k}(t)$ is maintained.

### 2.2. Theoretical argument

By using a simple cell model, we have confirmed the common proportionality over thousands of components for genetic and non-genetic variances, and environmental and genetic responses as summarized in figure 5. The results suggest the existence of a global variable that governs adaptive evolution. Here, the growth rate $\mu$ of a cell is a candidate for such a variable, as, for a cell to maintain its composition, every component has to be synthesized in conjunction with the growth rate. Indeed, in [12] we considered the dynamics of concentration of component $i$

$$\frac{dx_i}{dt} = f_i(x_i) - \mu x_i,$$  

where $x_i$ is the concentration of the component $i$, i.e. $x_i = N_i/V$ with the cell volume $V$, $f_i(x_i)$ is a function that governs the synthesis and degradation of the component $i$, and $\mu x_i$ gives the dilution of the component $i$ by the increase in cell volume. By using $x_i = \log x_i$, and $F_i((X_i)) = j_i((x_i))$, the original stationary state is given by

$$F_i((X_i^0)) = \mu.$$  

Now, with the change in environmental condition $E$ and genetic change $G$, the logarithmic concentration $X_i$ is shifted to $X_i^0 + \delta X_i$ and $\mu$ is shifted to $\mu + \delta \mu$. Assuming that $\delta X_i$ is not so large, and taking only the linear part of the changes and using the Jacobi matrix $J_{ij} = \partial F_i/\partial X_j$, we get

$$\sum_{j} J_{ij} \delta X_i(E, G) + \gamma^E_i \delta E + \gamma^G_i \delta G = \delta \mu(E, G),$$

where $\gamma^E_i = \partial F_i/\partial E$ and $\gamma^G_i = \partial F_i/\partial G$, respectively.

Here, $G$ is a coordinate introduced to represent the genetic change. It is not evident that the genetic change is represented by only a single variable. However, considering that under this scenario evolution progresses under a stressed environmental condition, one could project high-dimensional genetic change in the direction required to increase fitness (growth rate) under the condition, indicating that a single variable $G$ can be introduced; indeed, several studies conducted to date support this assumption [23,24,26]. Accordingly, the variable $G$ has the same dimensions as $E$, and can be scaled so that $G$ and $E$ induce the same degree of change in expression. The genetic evolution following the initial stress $\delta E$ is expected to diminish the environmental stresses, so that evolution occurs in the direction $\delta G < 0$, if the environmental change $\delta E$ is positive.

A relationship between the environmental $\delta X_i(E, 0)$ and evolutionary response $\delta X_i(E, G)$, suggested in simulations (see experimental verification section below), has not yet been established theoretically. To further evaluate the relationship between environmental and evolutionary dynamics, we assume that evolutionary change assimilates the change induced environmentally, as proposed by Waddington [32]. In our theory, this genetic assimilation is formulated by the introduction of the variable $G$, which has a similar effect on the environment (or compensates for the environmental stress), such that $\gamma^E_i = \partial F_i(x_i)/\partial E \sim \gamma^E_i = \partial F_i(x_i)/\partial G$.

Recall that $\gamma^E_i$ represents the direct responsive change in $x_i$ in response to the environmental change. Considering that evolution occurs in the projected space in the direction of $\delta E$, and also following the concept of genetic assimilation, in which the genetic phenotypic change imitates the genetic change, it can be assumed that $\gamma^E_i = \gamma^G_i$ (however, proportionality alone could be acceptable, as $\delta G$ can be rescaled to match the two).

This postulate is not directly derived from a particular principle, and could therefore be a crude approximation. However, it may be better understood by considering the reaction process in a cell. With the change to the novel environment, reaction processes to a few components $m$ are added or removed from the environment. For these components, $\gamma^E_m \neq 0$, while for most others, $\gamma^E_i = 0$. In fact, most components do not directly respond to environmental changes, but rather change only in response to changes in other components $\delta X_j$ (given by the Jacobi matrix $J_{ij}$).

In this case, the evolutionary change of such components could be facilitated by other components, but no direct change will be requested for the reaction path from the environmental component to the component $i$, by the genetic change and thus $\gamma^E_i = 0$ is expected, at least statistically. Alternatively, for components $m$, with $\gamma^E_m \neq 0$, there will be advances to change the path from the environment to $m$ by genetic changes so that $\gamma^E_m \neq 0$. Also, as the environmental influence on the component is larger, genetic changes to make a larger influence on the path will be selected, so that the correlation between $\gamma^E_i$ and $\gamma^G_i$ is predictable. Therefore, even though $\gamma^E_i = \gamma^G_i$ might be a strong postulate, a high correlation between the two is statistically expected.

Under the linear conditions of interest, the change in $\mu$ is proportional to $\delta E$ or $\delta G$, with $\delta \mu(\delta E, \delta G) = a(\delta E + \delta G)$. It should be noted again that the direction of $\delta G$ is opposite to $\delta E$. Thus, we obtained

$$\delta X_i(\delta E, \delta G) = \sum_{i} L_{ij}(\delta \mu(\delta E, \delta G) - \gamma(\delta E + \delta G))$$

$$= \delta \mu(\delta E, \delta G) \sum_{i} L_{ij}(1 - \frac{\gamma_j}{a}).$$

Then, over the course of evolution $\delta G = 0$ to $\delta G(m)$ under a given environmental condition $E$,

$$\frac{\delta X_i(\text{Gen}(m))}{\delta X_i(\delta E, \delta G(0))} = \frac{\delta \mu(\delta E, \delta G(m))}{\delta \mu(\delta E, 0)}.$$  

Specifically, all expression changes are proportional, as confirmed in the present simulations. To check the validity of the theory, we compared the proportion coefficient in the expression change (LHS of equation (2.5)) with the change in growth rate (RHS) numerically through the course of the evolution simulation. As shown in figure 3, the relationship...
The theoretical argument and the simulation results demonstrated that the evolutionary dynamics with growth rate recovery 
\( \delta \mu_{\text{Gen}}(m)/\delta \mu_{\text{Env}} \). To verify this relationship, we analysed the time-series transcriptome data obtained in an experimental evolution study of *E. coli* under conditions of ethanol stress [33,34]. In this experiment, after cultivation of approximately 1000 generations (2500 h) under 5% ethanol stress, six independent ethanol-tolerant strains were obtained, which exhibited an approximately twofold increase in specific growth rates in comparison to the ancestor. For all independent culture series, mRNA samples were extracted from approximately 10^7 cells at six different time points, and the absolute expression levels were quantified by using microarray analysis. All mRNA samples were obtained from the cells in exponential growth phase, which means that the changes in cellular state over the timescale of several generations were negligible, and each expression level represented cells in a steady-growth state (see [34] for details of materials and methods).

Using the time-series expression data of bacterial adaptive evolution, we analysed the common proportionality in expression changes. The environmental response of the *i*-th gene \( \delta X_{\text{Env}}^i(n) \) is defined by the log-transformed ratio of the expression level of the *i*-th gene obtained 24 h after exposure to the stress condition to that obtained under the no-stress condition. Similarly, the evolutionary response at *n* hours after the exposure to the stress \( \delta X_{\text{Gen}}^i(n) \) is defined by the log-transformed ratio of the expression level at *n* hours to that of the no-stress condition. We found a common trend between the environmental and genetic responses over all genes, as shown in figure 6. Furthermore, we also found that the proportion coefficient \( r(n) \) for \( \delta X_{\text{Gen}}^i(n)/\delta X_{\text{Env}}^i \) is roughly proportional to the growth recovery ratio \( \delta \mu_{\text{Gen}}(n)/\delta \mu_{\text{Env}} \), as shown in figure 6b, where \( \delta \mu_{\text{Gen}}(n) \) and \( \delta \mu_{\text{Env}} \) are the growth rate differences of *n* hours and 24 h after the exposure to stress, respectively. The results demonstrated that the evolutionary dynamics with growth recovery were accompanied by gene expression changes to the proportionality in equation (2.5) (as well as in equation (2.7)) is not perfect, but the data are scattered around this common proportionality line. This would reflect on the deviation from \( \gamma_f = \gamma_i^f \) for each component *i*.

Note that in the present theory, we did not assume any optimization explicitly. The increase in the growth rate \( \mu \) through evolution itself was not imposed to derive equation (2.5). The equation is simply a consequence of the dynamic systems in (i)–(iii). If we further assume that the growth rate is recovered by the evolution, then we can conclude that the ratio in equation (2.5) is between 0 and 1. Therefore, as the growth rate is recovered, the component concentrations tend to revert to the original through evolution. This is not trivial, as the system after evolution is still under new environmental conditions, which implies that all the concentrations (statistically) revert to the levels before the environmental change was applied. This strong homeostasis is primarily because of the global constraint across components set by the growth rate (i), and then because of genetic assimilation to cancel out the change in specific components as a result of the environmental change.

### 2.3. Experimental verification

The theoretical argument and the simulation results demonstrated the existence of a common proportion coefficient \( r(m) \) for \( \delta X_{\text{Gen}}^i(m)/\delta X_{\text{Env}}^i \) across components *i*, and its proportionality to the growth rate recovery \( \delta \mu_{\text{Gen}}(m)/\delta \mu_{\text{Env}} \). To verify this relationship, we analysed the time-series transcriptome data obtained in an experimental evolution study of *E. coli* under conditions of ethanol stress [33,34]. In this experiment, after cultivation of approximately 1000 generations (2500 h) under 5% ethanol stress, six independent ethanol-tolerant strains were obtained, which exhibited an approximately twofold increase in specific growth rates in comparison to the ancestor. For all independent culture series, mRNA samples were extracted from approximately 10^7 cells at six different time points, and the absolute expression levels were quantified by using microarray analysis. All mRNA samples were obtained from the cells in exponential growth phase, which means that the changes in cellular state over the timescale of several generations were negligible, and each expression level represented cells in a steady-growth state (see [34] for details of materials and methods).

Using the time-series expression data of bacterial adaptive evolution, we analysed the common proportionality in expression changes. The environmental response of the *i*-th gene \( \delta X_{\text{Env}}^i \) is defined by the log-transformed ratio of the expression level of the *i*-th gene obtained 24 h after exposure to the stress condition to that obtained under the no-stress condition. Similarly, the evolutionary response at *n* hours after the exposure to the stress \( \delta X_{\text{Gen}}^i(n) \) is defined by the log-transformed ratio of the expression level at *n* hours to that of the no-stress condition. We found a common trend between the environmental and genetic responses over all genes, as shown in figure 6. Furthermore, we also found that the proportion coefficient \( r(n) \) for \( \delta X_{\text{Gen}}^i(n)/\delta X_{\text{Env}}^i \) is roughly proportional to the growth recovery ratio \( \delta \mu_{\text{Gen}}(n)/\delta \mu_{\text{Env}} \), as shown in figure 6b, where \( \delta \mu_{\text{Gen}}(n) \) and \( \delta \mu_{\text{Env}} \) are the growth rate differences of *n* hours and 24 h after the exposure to stress, respectively. The results demonstrated that the evolutionary dynamics with growth recovery were accompanied by gene expression changes to

\[ \langle \delta X_i(\delta Y)^2 \rangle = \langle \delta \mu(\delta Y)^2 \rangle \left( \sum_j L_{ij} \left( 1 - \frac{Y_j}{a} \right) \right)^2, \]

where \( \delta Y \) is either \( \delta E \) or \( \delta G \), i.e. phenotypic changes induced by variation in the environment (i.e. noise) or by genetic change (e.g. mutation), and \( \langle \cdots \rangle \) is the average over the distribution induced by phenotypic noise or genetic variation. The variance \( V_{ip}(j) \) and \( V_{ip}(i) \) are \( \langle \delta X_i(\delta E)^2 \rangle \) and \( \langle \delta X_i(\delta G)^2 \rangle \), respectively, such that

\[ \frac{V_{ip}(j)}{V_{ip}(i)} = \frac{\langle \delta \mu(\delta E)^2 \rangle}{\langle \delta \mu(\delta G)^2 \rangle}. \]
increase the traits, and observed proportionality between the two. The correlation between the isogenic variances in trait expression and variance due to mutation (but without selection) was measured across a few thousand genes in Saccharomyces cerevisiae. Correlation between the two variances was observed [36–38], whereas proportionality was not so clear. This is possibly because evolution without selection was applied in the experiment, and therefore only the variance resulting from random mutation was measured.

3. Discussion
In this study, we have shown that the proportionality between the abundance of most intracellular components during adaptation and evolution changes. This was confirmed by simulated evolution of catalytic reaction network models of cells, laboratory experiments of bacterial evolution, and phenomenological theory. To check the generality of the simulation results, we performed evolutionary simulations using several different models, including those with different rules of chemical reactions and cell division. For example, we changed the rules of reaction in the model, such that chemical species were classified into two classes: (i) those with catalytic activity and (ii) those without catalytic activity, which was analogous to enzymes and metabolites. In another example, we performed simulations using differential equations for the time development of chemical concentrations instead of stochastic simulations. All these models showed the same proportionality as presented here, which strongly suggests that the results are general and independent of the details in the model.

As summarized in figure 5, the proportionality among two-by-two quantities, particularly, fluctuations and responses induced by environmental changes (noise) and by genetic changes (mutation), was demonstrated. Although the present theoretical argument is phenomenological, it is rather general, therefore, we expect that the observed relationships obtained from the simulation and laboratory experiments are universal, and independent of the specific models or organisms considered.

3.1. Phenotypic change in adaptive evolution is highly constrained to a restricted direction within a high-dimensional state space
The proportionality across thousands of components implies that there is a strong constraint on phenotypic evolution. In particular, the expression of different components cannot evolve independently, but rather change together, for the most part, along a one-dimensional path provided by equation (2.5). This is the strong constraint imposed on all possible, high-dimensional changes in phenotypes, and is a consequence of a steady-growth state and gradual evolutionary process parametrized by one parameter $G$, characterizing the evolutionary process under a given environment $E$. Here, the general restriction or feasibility of the direction of phenotype changes in future evolution is quantitatively formulated.

3.2. Quantitative representation of genetic assimilation
To derive equation (2.5), we assumed $\bar{y}_G = y_0$. Environmental inputs and genetic alterations are two ways to

eliminate the phenotypic changes introduced by the new environment, and agreed well with the simulation results of the simple cell model shown in figure 2 as well as the theoretical argument presented above.

Furthermore, there is some indirect experimental support for the proposed relationships between the variances. Stearns et al. [35] measured the isogenic variance $V_{ip}$ of five life-history traits (such as body weight, lifespan, etc.) in Drosophila melanogaster, as well as the genetic variance $V_g$ between different genetic lines observed during laboratory evolution to

![Figure 6. Response by environmental change versus response by evolution in E. coli adaptation to ethanol stress.](http://rsif.royalsocietypublishing.org/)
change each intracellular reaction process. Both change the rate (or existence) of each reaction, and it would be natural to assume a correlation between the two, which ultimately leads to proportionality between environmental and genetic responses. This could be regarded as the quantitative representation of genetic assimilation by Waddington [32], in which the response induced by the environmental change is ultimately shaped as a result of genetic evolution.

3.3. Homeostasis in high-dimensional phenotypic space through evolution

Considering that the growth rate is recovered by the evolution under environmental stress, the result of equation (2.5) implies that the changes in the concentrations of most components that are induced by the environmental change become relaxed through the evolutionary process. This suggests a phenomenon of strong homeostasis, in which the original, adapted, intracellular state is recovered through genetic changes. This homeostasis can be considered similar to the Le Châtelier’s principle in thermodynamics, which states that changes introduced by external perturbations are relaxed by subsequent temporal evolution.

3.4. A variety of genetic changes could lead to a given or similar directional change in phenotypes

There may be a huge variety of genetic changes that yield the phenotypic changes required for adaptation. In our simulations, there were a variety of network structures that could achieve same phenotypic adaptation. When the simulation was run a second time with a different seed of random numbers for mutations, the resulting network (i.e. genotypes) was different for each run. However, the change in concentrations (phenotypes) followed the proportionality given by equation (2.5), independent of the specific genetic changes occurring during evolution. Furthermore, in bacterial evolution experiments, the results from different strains tended to follow the same proportionality law described by equation (2.5). It is interesting to note that such correlated changes in expression levels by genetic changes have also been suggested in several other studies [14–16]. Of course, it will be important to further confirm the relationship shown in equation (2.5) in more laboratory evolution experiments, and to also unveil the underlying genotype–phenotype map that achieves the common, restricted change in expression levels observed in the experimental data.

3.5. Proportionality between fluctuations of non-genetic and genetic origins—a possible means of characterizing the feasibility of evolution

We have also demonstrated proportionality in the fluctuation in expression levels across components. As expected from Fisher’s fundamental theorem of natural selection [39–42], the higher the genetic variance, the higher the evolutionary rate. Hence, the proportionality between \( V_{\text{pp}}(i) \) and \( V_{\text{g}}(i) \) suggests that the higher isogenic variance of a given expression level because of noise would be accompanied by the higher rate of change in the expression level because of evolution. Components that are more variable by noise are also more variable by genetic change, and are more feasible to evolution, which is analogous to the fluctuation–response relationship described in statistical physics [43,44]. Our results suggest that the direction of evolutionary change in phenotypic space is likely to be predetermined by the isogenic variance of expression level because of noise.

According to our theoretical framework, the responses and fluctuations in expression levels are represented by the macroscopic growth rate and its fluctuation. Therefore, the relationship between the responses and fluctuations, analogous to thermodynamics, is represented by the landscape of the growth rate as a function of phenotype (expression level) and the environment. This is in contrast to the established fitness landscape represented in genetic space proposed by Wright [45]. The present study provides a basis for the development of a future macroscopic theory for phenotypic evolution.

4. Methods: model simulations

The cellular state was represented by a set of numbers \( (N_1, N_2, \ldots, N_3) \), where \( N_i \) is the number of molecules of the chemical species \( i \) with \( i \) ranging from \( i = 1 \) to \( K \). For internal chemical reaction dynamics, we chose a catalytic network among these \( K \) chemical species, where each reaction from some chemical \( i \) to some other chemical \( j \) was assumed to be catalysed by a third chemical \( \ell \), i.e. \( (i + \ell \rightarrow j + \ell) \). To represent the reaction matrix, we used the notation \( \text{Con}(i, j, \ell) \), which takes the value 1 when the reaction from chemical \( i \) to chemical \( j \) was catalysed by \( \ell \), and 0 under other conditions. This catalytic network was randomly generated, in which the probability that chemical \( i \) is generated from chemical \( j \) is given by the connection rate \( \rho \). For simplicity, all reaction coefficients were chosen to be equal. Some resources (nutrients) are supplied from the environment by transportation through the cell membrane with the aid of other proteins called ‘transporters’. Here, we assumed that the flux uptake of chemicals \( i \) from the environment is proportional to \( D()N_{w}, j \), where chemicals \( w \) act as the transporter for chemical \( i \), \( c_i \) is the concentration of chemical \( i \) in the environment and \( D \) is a transport constant. The concentrations of nutrient chemicals in the environment were kept constant, and they have no catalytic activity to prevent the occurrence of catalytic reactions in the environment. Through catalytic reactions, these nutrients are transformed into other chemicals, including the transporters. Here, we assume that all of the \( K \) chemical species are necessary for cell division. Therefore, cell division was assumed to occur when the minimum number of species exceeded a threshold of \( M \), i.e. \( \min_{\sum_{i=1}^{K} N_i} \geq M \) (in all analyses \( M \) was set to unity). Chosen randomly, the parent cell’s molecules are evenly split between the two daughter cells. In our numerical simulations, we randomly picked up a pair of molecules in a cell and transformed them according to the reaction network. In the same way, transportation through the membrane was also computed by randomly choosing from molecules within the cell and from nutrients in the environment. The parameters were set as \( K = 1000, p = 0.01 \) and \( D = 0.001 \). There were two nutrient chemicals, each was associated with one transporter chemical, and the concentrations of these two nutrients in the environment \( (c_1, c_2) \) were set to \( (0.5, 0.5) \).

We studied the evolution of replication dynamics by generating slightly modified networks and selecting those that grew faster. First, \( n \) parent cells were generated, and catalytic...
reaction networks were randomly generated using the connection rate $p$. From each of the $n$ parent cells, $L$ mutant cells were generated by randomly replacing $mpK^2$ reaction paths, where $mpK^2$ is the total number of reactions and $m$ is the mutation rate per reaction per generation. Reaction dynamics were simulated for each of the $nL$ cells to determine the rate of growth of each cell; i.e., the inverse of the time required for division. Within the cell population, $nL$ cells with faster growth rates were selected to be the parent cells of the next generation, from which $nL$ mutant cells were again generated in the same manner. Throughout the simulation, the parameters were set as $n = 1000$ and $L = 5$. In the simulations shown in figure 1, the mutation rate $m$ was set to $1 \times 10^{-3}$.

The environmental change is given by changing the nutrient concentration ratio in the environment. In the evolutionary simulation shown in figure 1, before adding the new environmental condition (generation $\leq 0$), the concentrations of these two nutrients in the environment ($c_1$, $c_2$) were set to $(0.5, 0.5)$, while after the environmental change (generation $> 0$), they were set to $(0.9, 0.1)$. In the result shown in electronic supplementary material, figure S1, to add a variety of environmental changes, we randomly selected a nutrient chemical and a transporter chemical for this nutrient among $K$ total chemical species. Then, the concentrations of the new nutrient $c_{\text{New}}$ and the original nutrients were set to $(c_1, c_2, c_{\text{New}}) = (0.45, 0.45, 0.1)$. We iterated the random addition of a nutrient 104 times to obtain the average environmental response $R^{\text{Env}}(i)$.

**Authors’ contributions.** C.F. and K.K. designed the study. C.F. performed the simulations and data analysis. K.K. performed the theoretical analysis. C.F. and K.K. wrote the manuscript.

**Competing interests.** The authors declare no competing financial interests.

**Funding.** This work was also supported in part by Grant-in-Aid for Scientific Research (S) (15H05746 to K.K.) and (B) (26290071 to C.F.) from JSPS, and Grant-in-Aid for Scientific Research on Innovative Areas (25128715 and 26119719 to C.F.) from MEXT, Japan.

**Acknowledgements.** The authors thank T. Yomo for stimulating discussions and constructive comments.

---

**References**


