bioRxiv preprint doi: https://doi.org/10.1101/125278; this version posted April 7, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Formation of Dominant Mode by Evolution in Biological Systems

Chikara Furusawa*
Quantitative Biology Center (QBiC), RIKEN, 6-2-3 Furuedai, Suita, Osaka 565-0874;
Universal Biology Institute, University of Tokyo, 7-3-1 Hongo, Tokyo 113-0033, Japan

Kunihiko Kaneko†
Research Center for Complex Systems Biology, Universal Biology Institute,
University of Tokyo, 3-8-1 Komaba, Tokyo 153-8902, Japan

Constraints on changes in expression levels across all cell components imposed by the steady growth of cells have recently been discussed both experimentally and theoretically. By assuming a small environmental perturbation and considering a linear response to it, a common proportionality in such expression changes was derived and partially verified by experimental data. Here, we examined global protein expression in *Escherichia coli* under various environmental perturbations. Remarkably they are proportional across components, even though these environmental changes are not small and cover different types of stresses, while the proportion coefficient corresponds to the change in growth rate. However, since such global proportionality is not generic to all systems under a condition of steady growth, a new conceptual framework is needed. We hypothesized that such proportionality is a result of evolution. To validate this hypothesis, we analyzed a cell model with a huge number of components that reproduces itself via a catalytic reaction network, and confirmed that the common proportionality in the concentrations of all components is shaped through evolutionary processes to maximize cell growth (and therefore fitness) under a given environmental condition. Further, we found that the concentration changes across all components in response to environmental and evolutionary changes are constrained along a one-dimensional major axis within a huge-dimensional state space. Based on these observations, we propose a theory in which high-dimensional phenotypic changes after evolution are constrained along a one-dimensional major axis that correlates with the growth rate, which can explain broad experimental and numerical results.

bioRxiv preprint doi: https://doi.org/10.1101/125278; this version posted April 7, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Summary:
Cells generally consist of thousands of components whose abundances change through adaptation and evolution. Accordingly, each steady cell state can be represented as a point in a high-dimensional space of component concentrations. In the context of equilibrium statistical thermodynamics, even though the state space is high-dimensional, macroscopic description only by a few degrees of freedom is possible for equilibrium systems; however, such characterization by few degrees of freedom has not yet been achieved for cell systems. Given they are not in equilibrium, we need some other constraint to be imposed. Here, by restricting our focus to a cellular state with steady growth that is achieved after evolution, we examine how the expression levels of its several components are changed under different environmental conditions. Based on the analysis of protein expression levels in recent bacterial experiments as well as the results of simulations using a toy cell model consisting of thousands of components that are reproduced by catalytic reactions, we found that adaptation and evolutionary paths in a high-dimensional state space are constrained along a one-dimensional curve, representing a major axis for all observed changes. Interestingly, this one-dimensional structure emerges only after evolution and is not applicable to any system showing steady growth. This curve is given by the growth rate of a cell, and thus it is possible to describe an evolved system by a growth-rate function. All of the observed results are consistent with the hypothesis that changes in high-dimensional states are constrained along the major axis against environmental, evolutionary, and stochastic perturbations. This description opens up the possibility to characterize a cell state as a macroscopic growth rate, similar to a thermodynamic potential. This approach can provide estimates of which phenotypic changes are theoretically more evolvable, as predicted simply from their observed environmental responses.

INTRODUCTION

Cells contain a huge variety of components whose concentrations change through adaptation and evolution. To realize a theoretical description of such a system with many degrees, it is important to be able to characterize the system based on a few macroscopic variables, as demonstrated by the success in traditional statistical thermodynamics. In thermodynamics, by restricting the focus to the equilibrium state and its vicinity, a state can be described by a few macroscopic variables even though the microscopic degrees of freedom is huge. Thus, we sought to investigate the possibility of adopting a similar approach to a biological system that contains a large number of components.

Of course, the strategy of statistical thermodynamics cannot be directly adopted for a biological system given that cells are not in a thermodynamic equilibrium. Cells grow (and divide) over time in a manner that is driven by complex intracellular dynamics. Despite this difference, however, it may still be possible to establish a description of cells with only a few degrees of freedom by restricting our focus to only cells in a steady growth state. Indeed, for such cases, all of the intracellular components have to be roughly doubled before cell division occurs for reproduction. This restriction imposes a general constraint on the possible changes in expression levels across the many components, which opens the possibility to characterize a cellular state with few variables. Indeed, several previous studies have attempted to search for universal laws in the steady growth state, as pioneered by the work of Monod[1] followed by subsequent studies[2–5].

Currently, investigations on the existence of such constraints are generally carried out by examining changes in expression occurring over thousands of genes using transcriptome or proteome data. In fact, a general trend across changes in the overexpression of many components has been observed[6–10] during the course of adaptation and evolution[11–16]. In particular, using transcriptome analysis in bacteria, we recently found that changes in (logarithmic) expression levels are proportional across many components in given stress conditions with different strengths, and the proportion coefficient corresponds to the change in the growth rate[17, 18]. Further, a theory based on steady growth and linear approximation with the assumption of small changes in expression levels can account for these experimental findings rather well. Nevertheless, experimental data also suggest that such proportionality in the expression levels across many genes may even be extended over a wide variety of environmental conditions, which is beyond the

bioRxiv preprint doi: https://doi.org/10.1101/125278; this version posted April 7, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

scope of the simplified theory for steady-state conditions. Therefore, further experimental confirmation for the existence of such global proportionality is anticipated.

In the present study, we first analyzed the changes in protein expression levels under a variety of environmental conditions using recent experimental data. We confirmed the pattern of proportional changes across all protein expression levels, with the proportion coefficient approximately agreeing with the cell growth rate, representing a basic, macroscopic quantity for steady-growth cells. On the one hand, this finding confirms our theory, but on the other hand, the results are beyond the key assumption of the previous simple theory. That is, this proportional relationship also appears to be valid over a broad range of different environmental conditions

Such global proportionality is not generic in a dynamical system consisting of many components under a simple state of steady growth. This suggests that another factor is needed to explain this phenomenon beyond the constraint of steady growth discussed previously. At this stage, it is important to recall that biological systems are an outcome of evolution. Hence, it is natural to expect that the missing link might be provided as a result of evolution under a given environmental condition. Accordingly, we here examine the validity of the hypothesis that evolution under a given condition will constrain intracellular dynamics so that the concentration changes across many components are mainly governed by changes in the growth rate. Extensive evolutionary experiments are needed to test this hypothesis, which may be too difficult to conduct in a laboratory. However, a numerical evolution approach offers a possible method to address this question.

In fact, several models have been proposed that capture the evolution of a cellular state with many phenotypic variables. Although such models may not be able to capture the entire breadth of the complexity of cells, they are sufficiently complex to enable interpretations on the evolution of genotype-phenotype mappings[19–23]. Here, we study a cell model consisting of a large number of protein species that mutually catalyze each other to transform environmental resources for cellular growth. We demonstrate that after evolution, the changes across all catalysts are constrained along a one-dimensional curve that is governed by the growth rate.

Based on these findings in numerical simulations, with support from laboratory experiments, we formulate the following hypothesis of an evolved system: phenotypic changes of many components upon environmental or genetic changes mainly occur along a single major axis that parameterizes the growth rate. Indeed, this hypothesis can explain the results of phenotypic changes occurring by adaptation and evolution quite well. We further discuss the general consequence and biological relevance of this hypothesis.

GLOBAL PROPORTIONALITY OBSERVED IN PROTEIN EXPRESSION LEVELS

By imposing linearization approximation with steady growth, we previously demonstrated a general relationship for the change in concentration x_j across all components j in a cell under a given environmental (stress) condition, where the strength of environmental stress (e.g., temperature, osmotic pressure, degree of starvation) is parameterized by E. By representing the relationship as $X_j = \log x_j$, all the changes in δX_j against the change in the stress strength E are expressed as

$$\frac{\delta X_j(E)}{\delta X_j(E')} = \frac{\delta \mu(E)}{\delta \mu(E')}, \tag{1}$$

where μ is the growth rate of a cell in the steady-growth state, and E is the strength of a given stress such as the degree of starvation, osmotic pressure, or temperature. Since the right-hand side is independent of each component j, the changes of all components are proportional, with the slope given by the growth rate change. Here, the growth rate provides a global variable for all components, given that all of the concentrations are diluted along with the rate.

Previously, we confirmed the above relationship using transcriptome analysis for the mRNA abundance under given stress types[17]. To extend beyond the simple theory for the relationship given in Eq. (1), we here further examine this relationship not simply against the strength of a given stress but against a variety

bioRxiv preprint doi: https://doi.org/10.1101/125278; this version posted April 7, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

of environmental conditions. Moreover, to allow for direct comparison with protein expression dynamics, we examined the proteome data of growing *Escherichia coli* under various environmental conditions reported in [24]. The environmental conditions considered include culture under various carbon sources, application of stresses such as high temperature and osmotic pressure, and glucose-limited chemostat cultures with different dilution rates (which are equivalent to specific growth rates in the steady state). Measurement of intracellular protein concentrations was carried out using mass spectroscopy, which provided data of the absolute concentrations of more than thousand protein species.

From the proteome analysis in various environmental conditions, we calculated the change in protein concentrations between the standard (reference) state and that under different environmental conditions. In the following analysis, the minimal medium with glucose as a sole carbon source was selected as the standard state, and the essentially same results were obtained when the other data were used as the standard state to compute the changes $\delta X_j(\mathbf{E})$ and $\delta\mu$. The log-transformed expression changes of the j-th protein is calculated as $\delta X_j(\mathbf{E}) = \log(x_j(\mathbf{E})/x_j(\mathbf{S}))$, where $x_j(\mathbf{E})$ and $x_j(\mathbf{S})$ represent the protein concentration in a given environment $\mathbf{E}$ and the standard condition, respectively. Here, we adopted the logarithmic scale since protein expression levels typically change on this scale[25–28], and this choice also facilitates comparison with the theory described above and to be discussed below.

To examine whether the common proportionality of protein expression changes in Eq. (1) is also valid across different environmental conditions, we analyzed the relationship of protein concentration changes δX_j observed in two different environmental conditions. In Fig. 1(a) and (b), we show some examples of the relationship between protein expression changes $\delta X_j(\mathbf{E_1})$ and $\delta X_j(\mathbf{E_2})$, where $\mathbf{E_1}$ and $\mathbf{E_2}$ are two different culture conditions (see the caption of Fig. 1 for details). As shown in the figure, the expression changes in different environmental conditions are highly correlated over a large number of proteins. Furthermore, we found that the slope of the common trend in protein expression change generally agrees with that calculated by the ratio of growth rate changes as in Eq. (1) (blue lines in Fig. 1(a) and (b)). In Fig. 1(c), for all possible combinations of the two environments, we compared the slope of the common trend in protein expression obtained by linear regression, i.e., comparison of the common ratio $\delta X_j(\mathbf{E_1})/\delta X_j(\mathbf{E_2})$ with the ratio of growth rate change $\delta\mu(\mathbf{E_1})/\delta\mu(\mathbf{E_2})$. The good agreement demonstrated in Fig. 1(c) indicates that the theory based on the steady-growth state and linearization can accurately explain the protein expression data obtained under various environmental conditions.

One might assume that the present result could be a natural consequence of the theory based on the steady-growth and linear approximation. However, the experimental data suggest a much stronger relationship that goes beyond the previous simple theory developed in [17].

In the previous theory, the parameter values of E and E' represent the environmental conditions, specifically the strength of the stress of the same type; for example, a change in the culture temperature, nutrient concentration, or osmotic pressure. However, the proportionality across different types of stress conditions (e.g., temperature versus nutrient) is not derived from the original theoretical formulation. Nevertheless, the data suggest that the proportionality in Eq. (1) is still valid in this case to a certain degree.

Besides this unexpected proportionality across different dimensions, the obtained data suggest a broad range of linearity. The reaction or expression dynamics are generally highly nonlinear, with catalytic reactions or expression dynamics showing high Hill coefficients. However, the experimental data still satisfy the linear relationship even under a stress condition that reduces the growth rate below 20% or so compared to the standard, which probably occurs up to the limit beyond which steady growth can no longer be easily achieved.

These points suggest that other factors affect the present cell beyond those considered in the simple theory. The proportionality in the form of Eq. (1) is theoretically derived for a given type of sufficiently small stress, as long as the system satisfies steady reproduction while maintaining diverse components, whereas the data suggest a much stronger form of linearity.

Of course, biological systems are not only constrained by steady growth but are also a product of evolution. With evolution, cells can efficiently and robustly reproduce themselves under a given external condition. Thus, it is important to examine the validity of the hypothesis that the above global

bioRxiv preprint doi: https://doi.org/10.1101/125278; this version posted April 7, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

proportionality is a result of, and further strengthened by, evolution.

GLOBAL PROPORTIONALITY EMERGED THROUGH EVOLUTION

To validate the above hypothesis, we evaluated how reaction (or protein expression) dynamics are shaped through evolution. Toward this end, we adopt a simple cell model consisting of a large number of components, and numerically evolve it under a given fitness condition to investigate how the phenotypes of many components evolve. Here, phenotypes are given by the concentrations of chemicals, which change due to catalytic chemical reactions by enzymes that are in turn synthesized by some other catalytic reactions. Here, genes provide a rule for the possible reactions such as the parameters and structure of a catalytic reaction network.

The phenotypes of each organism as well as the growth rate (fitness) of a cell is determined by such reaction dynamics, while the evolutionary process consists of selection according to the associated fitness and genetic change in the reaction network (i.e., rewiring of the pathway). By considering a population of such cells that slightly differ in their genotypes, and accordingly in their relative fitness, the offspring are generated according to the fitness value. Through these dynamics, the evolutionary changes in the genotypes and corresponding phenotypes are traced, and the evolution of phenotypic changes is studied by using this combination of dynamical systems and a genetic algorithm, which enables testing of the hypothesis posed.

The specific cell model adopted for this purpose consists of k components whose concentrations are represented by $x_1, x_2, \cdots, x_k$[30, 31]. There are $m(< k)$ resource chemicals whose concentrations in an environment and a cell are given by $s_1, \cdots, s_m$ and $x_1, \cdots, x_m$, respectively. The resource chemicals are transported into the cell with the aid of other chemical components named "transporters". The other chemical species work as catalysts, which are synthesized from other components, including resource chemicals, through catalytic reactions with the aid of some other catalysts, while the resource and transporter chemicals are non-catalytic. These reaction dynamics are given by a network of catalytic reactions among these components, which are determined genetically. Thus, the cell volume increases with the increase of its internal chemical components. For simplicity, the volume is set to be proportional to the total abundances of the chemical components. With the transport of external resources and their conversion to other components, the cell increases its volume and divides in two when it reaches a size larger than a given threshold. The growth rate depends on the catalytic network, i.e., genotype. We adopted this model of a catalytic reaction network for simplicity, which also exhibits power-law statistics, log-normal fluctuation, adaptation with fold-change detection, which is consistent with the properties of the cells[30, 31].

At each generation, a certain fraction of cells with higher growth rates are selected to produce offspring, resulting in a given mutational change in the network (see Appendix). With the iteration of mutation and selection of reaction networks, the growth rate increases over the generations under a given original (standard) environmental condition $\mathbf{o} = \{s_1^o, \cdots, s_m^o\}$, while the network of the initial generation is provided by a random network.

We then analyzed the response of the component concentrations to the environmental change from the original condition. Here, the environmental condition is given by the external concentration $s_1, \cdots, s_m$. We then changed the condition to $s_j^{(\varepsilon,\mathbf{E})} = (1-\varepsilon)s_j^o + \varepsilon s_j^{\mathbf{E}}$, where $\mathbf{E} = \{s_1^{\mathbf{E}}, \cdots, s_m^{\mathbf{E}}\}$ denotes the vector of the new, stressed environment, and ε is the strength of the stress. For each environment, we computed the reaction dynamics of the cell, and computed the concentration $x_j^{(\varepsilon,\mathbf{E})}$ from the steady growth state to obtain the logarithmic change in the concentration $\delta X_j^{(\varepsilon,\mathbf{E})} = \log(x_j^{(\varepsilon,\mathbf{E})}/x_j^{\mathbf{o}})$; the change in growth rate μ, denoted as $\delta\mu^{(\varepsilon,\mathbf{E})}$, was computed at the same time.

We next determined whether the changes $\delta X_j^{(\varepsilon,\mathbf{E})}$ satisfy the common proportionality trend across all components against a variety of environmental changes, and found that the proportion coefficient agrees with $\delta\mu^{(\varepsilon,\mathbf{E})}$, up to its largest change. We examined the proportionality for the network both before and after evolution, i.e., the random network and the evolved network, under the given environmental

bioRxiv preprint doi: https://doi.org/10.1101/125278; this version posted April 7, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

condition.

First, we computed the response in expression against the same type of stress, i.e., the same vector $\mathbf{E}$ with different strengths ε. Fig. 2(a) shows the correlation coefficients between the changes in component concentrations $\delta X_j^{(\varepsilon_1,\mathbf{E})}$ and $\delta X_j^{(\varepsilon_2,\mathbf{E})}$ caused by different strengths of environmental change ($\varepsilon_2 = \varepsilon_1 + \varepsilon$, with $\varepsilon > 0$). This result was obtained using a variety of randomly chosen environmental vectors $\mathbf{E}$. As shown, for a small environmental change ($\varepsilon = 0.02$), the correlation is sufficiently high both for the random and evolved networks, whereas for a larger environmental change ($\varepsilon = 0.08$), the correlation coefficients become significantly smaller for the random networks. Fig. 2(b) shows the relationship between the ratio of the growth rate changes $\delta\mu^{(\varepsilon_1,\mathbf{E})}/\delta\mu^{(\varepsilon_2,\mathbf{E})}$ and the slope in $(\delta X_j^{(\varepsilon_1,\mathbf{E})}, \delta X_j^{(\varepsilon_2,\mathbf{E})})$ obtained by fitting the concentration changes of all components. Fig. 2(c) shows the ratio of the slope to $\delta\mu^{(\varepsilon_1,\mathbf{E})}/\delta\mu^{(\varepsilon_2,\mathbf{E})}$ (which becomes unity when Eq. (1) is satisfied) as a function of the strength of the environmental change ε_1. These results clearly demonstrated that for a small environmental change, the relationship in Eq. (1) is valid both for the random and evolved networks, while for the large environmental change, it is maintained only for the evolved networks.

Second, we examined the correlation of concentration changes across different types of environmental stresses. Fig. 3(a) shows the correlation coefficients of $\delta X_j^{(\varepsilon_1,\mathbf{E_1})}$ and $\delta X_j^{(\varepsilon_2,\mathbf{E_2})}$ with different vectors $\mathbf{E_1}$ and $\mathbf{E_2}$. As shown, the correlation coefficient obtained for the evolved network is much larger and closer to unity than that obtained for the random network. The relationship between $\delta\mu^{(\varepsilon_1,\mathbf{E_1})}/\delta\mu^{(\varepsilon_2,\mathbf{E_2})}$ and the slope in $(\delta X_j^{(\varepsilon_1,\mathbf{E_1})}, \delta X_j^{(\varepsilon_2,\mathbf{E_2})})$ is presented in Fig. 3(b), while the ratio of the slope in $(\delta X_j^{(\varepsilon_1,\mathbf{E_1})}, \delta X_j^{(\varepsilon_2,\mathbf{E_2})})$ to $\delta\mu^{(\varepsilon_1,\mathbf{E_1})}/\delta\mu^{(\varepsilon_2,\mathbf{E_2})}$ is shown in Fig. 3(c) as a function of ε_2. The agreement of the slope of δX with the growth rate change given by Eq. (1) is much more remarkable for the evolved networks than for the random networks.

These findings demonstrate that the global proportionality emerges as a result of evolution under a given environmental condition. This condition suggests that the data $X_j^{(\varepsilon,\mathbf{E})}$ are constrained along a one-dimensional manifold after evolution. To examine this possibility, we analyzed the changes of $X_j^{(\varepsilon,\mathbf{E})}$ across a variety of environmental changes $\mathbf{E}$ and ε using principal components analysis. Fig. 4 clearly demonstrates that in the evolved network, high-dimensional data from $X_j^{(\varepsilon,\mathbf{E})}$ are located along a one-dimensional curve, while the principal components plot agrees with the growth rate μ rather well (see Fig. 5). (Note that the contribution of the major-axis mode (i.e., the first principal component reaches 74% in the data). In contrast, the data from the random network are scattered and no clear structure is visible.

Finally, we examined the evolutionary course of the phenotype projected on the same principal component space as that depicted in Fig. 4. As shown in Fig. 6(a), the points from $\{X_j\}$ generated by random mutations to the reaction network are again located along the same one-dimensional curve. Furthermore, those obtained by environmental variation or by noise in the reaction dynamics also lie along this one-dimensional curve, as shown in Fig. 6(b). Hence, the phenotypic changes are highly restricted, both genetically and non-genetically, along an identical one-dimensional curve.

In summary, we observed emergent global proportionality beyond trivial linearity in response to even tiny environmental changes. After evolution, the linearity region is extended to a level with an order-of-magnitude change in the growth rate. Accordingly, the correlation across different environments is enhanced. Changes in high-dimensional phenotype space are constrained along a one-dimensional manifold, resulting in the correlation between environmental and evolutionary changes.

DOMINANCE OF THE MAJOR AXIS CORRESPONDING TO THE GROWTH RATE

To formulate the observed global proportionality, we established the following hypothesis and discuss its consequence: both environment- and evolution-induced changes in a high-dimensional phenotype space are constrained mainly along a one-dimensional major axis[32]. Phenotypic dynamics are slower along this axis than across it. The axis forms a ridge in the fitness landscape. Evolution progresses along

bioRxiv preprint doi: https://doi.org/10.1101/125278; this version posted April 7, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

this axis, whereas the fitness function is steep across the axis. Indeed, dominance of the first principal mode in phenotypic change emerges after evolution simulation of the cell model (see Fig. 4). Moreover, expression data from bacterial evolution studies also support this hypothesis (see Fig. 6 of [18]).

Now, by following the dynamical-systems representation of phenotypic dynamics under steady growth[17], we formulate the above hypothesis as follows. First, consider the high-dimensional phenotypic dynamics

$$dx_i/dt = f_i(\{x_j\}) - \mu x_i \tag{2}$$

where x_i $(i = 1, 2, .., N)$ is the concentration of components (e.g., proteins) in a cell and μ is the growth rate of the cell. By using $X_i = \log(x_i)$ and $x_i F_i = f_i$, the above equation is rewritten as

$$dX_i/dt = F_i(\{X_j\}) - \mu. \tag{3}$$

Thus, the stationary solution for a given environmental parameter vector $\mathbf{E}$ is given by

$$F_i(\{X_j^*\}, \mathbf{E}) = \mu(\mathbf{E}). \tag{4}$$

By linearizing the equation around a given stationary solution using the Jacobi matrix $\mathbf{J}$ and susceptibility vector $\gamma_i = \partial F_i/\partial \mathbf{E}$, we get

$$\sum_j J_{ij}\delta X_j + \gamma_i(E)\delta E = \delta\mu(E). \tag{5}$$

Thus, by using $\mathbf{L} = \mathbf{J}^{-1}$,

$$\delta\mathbf{X} = \mathbf{L}(\delta\mu - \gamma\delta E). \tag{6}$$

Diagonalizing the matrix L by $T \cdots T^{-1}$, and denoting $\delta\hat{\mathbf{X}} = T\mathbf{X}T^{-1}$, one gets $\delta\hat{\mathbf{X}} = \Lambda(\delta\mu - T\gamma T^{-1}\delta E)$. Accordingly, we get

$$\delta\hat{\mathbf{X}} = \sum_k \lambda^k(\delta\mu - \mathbf{w^k}\cdot\gamma\delta E)\mathbf{w^k}, \tag{7}$$

where $\mathbf{w^k}$ is the eigenvector corresponding to the k-th eigenvalue λ_k.

Thus, the major axis hypothesis states that the magnitude of the smallest eigenvalue of $\mathbf{J}$ is much smaller than that of the others (or the absolute eigenvalue of $L = \mathbf{J}^{-1}$ is much larger than that of the others), whose eigenmode dominates the long-term phenotypic dynamics, so that the change in δX is mainly constrained along the major axis corresponding to its eigenvector (see the schematic Fig. 7, and see Figs. 4–6 for possible numerical support). Let us then denote this eigenvector by $\mathbf{w^0}$.

Then, to study the response to a variety of environmental changes with different γ, we need to take only the largest eigenvalue of L (i.e., the smallest eigenvalue of $\mathbf{J}$), and project all the changes in the dynamics along this major axis. By introducing projection to $\mathbf{w^0}$, (i.e., a projection of γ to this major axis)[33], we get

$$\delta\hat{\mathbf{X}} = \lambda^0\mathbf{w^0}(\delta\mu - \mathbf{w^0}\cdot\gamma\delta E). \tag{8}$$

Hence, across different environmental conditions, the following relationship is obtained:

$$\frac{\delta\hat{\mathbf{X}}(\mathbf{E})}{\delta\hat{\mathbf{X}}(\mathbf{E}')} = \frac{\delta\mu(E) - \mathbf{w^0}\cdot\gamma(\mathbf{E})\delta E}{\delta\mu(E') - \mathbf{w^0}\cdot\gamma(\mathbf{E}')\delta E'}. \tag{9}$$

Now, by putting $\delta E = \delta\mu/\alpha(E)$ assuming linearity, we get

$$\frac{\delta\hat{\mathbf{X}}(\mathbf{E})}{\delta\hat{\mathbf{X}}(\mathbf{E}')} = \frac{\delta\mu(E)}{\delta\mu(E')}\frac{(1 - \mathbf{w^0}\cdot\gamma(\mathbf{E})/\alpha)}{(1 - \mathbf{w^0}\cdot\gamma(\mathbf{E}')/\alpha)}. \tag{10}$$

bioRxiv preprint doi: https://doi.org/10.1101/125278; this version posted April 7, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

This results in proportionality of δX across all components j. As for the proportion coefficient, there is correction from $\delta\mu$ by the factor $(1 - \mathbf{w^0}\cdot\gamma(\mathbf{E})/\alpha)$. However, this correction term would not be so large as long as the projection of the γ vector to the major axis is not so large. Along the vector $\mathbf{w^0}$, the original dynamics $d\mathbf{X}/dt = \mathbf{F}(\mathbf{X}; E)$ do not change substantially, while γ denotes the sensitivity of $\mathbf{F}$ to an external change, so that the two vectors are not aligned generically in the high-dimensional state space, and thus their inner product will be small. Hence, $\mathbf{w^0}\cdot\gamma$ will be approximately negligible. Thus, the proportion coefficient across different environmental conditions approximately agrees with the ratio between the growth rates.

Here, we should note that with the assumption $\mathbf{w^0}\cdot\gamma \sim 0$, we do not need the linearity between $\delta\mu$ and $\delta\epsilon$ to derive the linear relationship between $\delta\mathbf{X}$ and $\delta\mu$. From Eq.(9), we directly obtain the relationship

$$\delta\hat{\mathbf{X}} = \lambda^0 \delta\mu \mathbf{w^0} \tag{11}$$

without requiring the linear approximation $\delta\mu \propto \delta\varepsilon$. With this form, even though $\delta\mu$ as well as $\delta\hat{\mathbf{X}}$ have non-linear dependence on $\delta\epsilon$, the agreement of the slope across the component changes δX_j with the change in the growth rate $\delta\mu$ is derived. Indeed, the proportionality from the simulation results described in the last section holds up for the case of a regime in which $\delta\mu$ is no longer proportional to the stress strength $\delta\epsilon$ (Fig. 8).

To sum up, assuming the major axis hypothesis, we could explain two basic features that have thus far been observed in experiments and simulations:
(1) **Overall proportionality is observed in expression-level changes across most components and across a variety of environmental conditions**. This is because high-dimensional changes are controlled by the major axis, i.e., the eigenvector $\mathbf{w^0}$.
(2)**There is an extended region for global proportionality**. As the range in the variation along $\mathbf{w^0}$ is large, the change in phenotype is constrained along this eigenvector, so that the proportionality range of phenotypic change is extended through evolution. Further, as long as the changes are constrained in the manifold along the major-axis, global proportionality is extended to the regime nonlinear to $\delta\epsilon$.

CONSEQUENCE OF THE THEORY FOR PHENOTYPIC EVOLUTION

Congruence of phenotypic changes due to evolutionary and environmental changes

The phenotypic changes constrained along the major axis are shaped through evolution, which further constrains any subsequent evolutionary potential. As both the environmentally and genetically induced changes of phenotype, $\delta\mathbf{X}(\mathbf{E})$ and $\delta\mathbf{X}(\mathbf{G})$, respectively, are constrained along the same axis, they are inevitably correlated (recall Fig. 6). Thus, by applying the argument for $\delta\mathbf{X}(\mathbf{E})$ to evolutionary (genetic) change $\delta\mathbf{X}(\mathbf{G})$, the two changes are given by

$$\mathbf{J}\delta\mathbf{X} + \gamma(\mathbf{E})\delta E + \gamma(\mathbf{G})\delta G = \delta\mu(E). \tag{12}$$

Following the projection to the major axis again (and taking only the eigenmode), we get

$$\frac{\delta\hat{\mathbf{X}}(\mathbf{E})}{\delta\hat{\mathbf{X}}(\mathbf{G})} = \frac{\delta\mu(E)}{\delta\mu(G)} \frac{(1 - \mathbf{w^0}\cdot\gamma(\mathbf{E})/\alpha)}{(1 - \mathbf{w^0}\cdot\gamma(\mathbf{G})/\alpha)}. \tag{13}$$

This demonstrates proportionality between $\delta\mathbf{X}(\mathbf{E})$ and $\delta\mathbf{X}(\mathbf{G})$ across all components. As for the proportion coefficient, there may exist slight correction from $\delta\mu$ due to the factor $(1-\mathbf{w^0}\cdot\gamma(\mathbf{E}(\mathbf{G}))/\alpha)$. However, this would not be expected to be so large as long as the projection of the γ vector to the major axis is not so large (or the γ vectors for E and G are aligned), which might be reasonable, so that the corrections will be cancelled out between the denominator and numerator. Then,

$$\frac{\delta\hat{\mathbf{X}}(\mathbf{E})}{\delta\hat{\mathbf{X}}(\mathbf{G})} = \frac{\delta\mu(E)}{\delta\mu(G)} \tag{14}$$

bioRxiv preprint doi: https://doi.org/10.1101/125278; this version posted April 7, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

holds. This proportionality between environmentally and evolutionarily induced changes, in the form of Eq.(15), is reported both in bacterial evolution experiments and numerical evolution of the cell model[18], which is well explained by the present theory.

Fluctuation relationship

As changes due to noise and genetic variations progress along the common major axis, the phenotypic variances due to the former (V_g) and those due to the latter (V_{ip}) are proportional across all components. In fact, since relaxation of phenotypic changes is much slower along the major axis, the phenotypic fluctuations due to noise are mainly constrained along it. Thus, the variance of each component X_i is given by the means of the variance in the slow variable X,

$$V_{ip}(i) = (\mathbf{w}_i^{\mathbf{0}})^2 < \delta X^2 >_{noise} . \tag{15}$$

Likewise, the variation due to the genetic change of each component is mostly constrained along the axis, so that

$$V_g(i) = (\mathbf{w}_i^{\mathbf{0}})^2 < \delta X^2 >_{mutation}, \tag{16}$$

where $\mathbf{w}_i^{\mathbf{0}}$ is the i-th component of $\mathbf{w}^{\mathbf{0}}$ (i.e., $\delta X \mathbf{w}_i^{\mathbf{0}}$ agrees with the projection of the projection of δX_i onto the major axis $\mathbf{w}^{\mathbf{0}}$). Hence,

$$V_{ip}(i) \propto V_g(i) \tag{17}$$

holds across components.

Note that for derivation of the relationship (17), the growth rate term is not involved. What we need is just the dominance of the major-axis mode $\mathbf{w}^{\mathbf{0}}$ in the expression changes. Indeed, besides the numerical evolution of the cell model with the catalytic reaction network[18], this relationship is also observed in the gene regulation network model where the growth-rate dilution is not included[34], while it is also suggested in experiments[35, 36]. The present theory accounts for the origin of the empirical relationship as a direct consequence of hypothesis on the dominance of major-axis mode, in which changes incurred both by (environmental) noise and (genetic) mutations are constrained along the axis. Thus, agreement with simulation results can provide further support for the hypothesis.

Here, recall that the evolutionary rate of a phenotype is proportional to its variance due to genetic change, V_g[37, 38]. Hence, the $V_g - V_{ip}$ relationship suggests that it is easier for a phenotype with larger V_{ip} to evolve under selective pressure. Thus, the specific phenotype that is most feasible to evolve is predetermined according to the variability against noise prior to application of the mutation-selection process.

DISCUSSION

In the present paper, we first confirmed the global proportionality across protein expression levels against a variety of environmental conditions using the experimental data of the bacteria transcriptome/proteome. The rate of change in logarithmic concentrations agrees rather well with that of cellular growth rate. By carrying out an evolutionary simulation of a cell model with the catalytic reaction network, we showed that such proportionality is shaped through evolution, where expression changes in a high-dimensional state space of phenotype are constrained along a major axis that corresponds to the growth rate.

These results are explained by assuming that most of the expression changes are constrained along the major axis $\mathbf{w}^{\mathbf{0}}$. This axis $\mathbf{w}^{\mathbf{0}}$ is the eigenvector for the smallest eigenvalue of the Jacobi matrix in the expression reaction dynamics; accordingly, the variation in the phenotype $\delta\mathbf{X}$ is larger along this axis. Biologically speaking, this $\mathbf{w}^{\mathbf{0}}$ might be represented by a concentration change of a specific component; more naturally, this axis could be represented by a collective variable generated from the concentrations (expression levels) of many components. In other words, changes of the cellular state that progresses in a high-dimensional space are constrained along the change in this major axis $\mathbf{w}^{\mathbf{0}}$.

bioRxiv preprint doi: https://doi.org/10.1101/125278; this version posted April 7, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Note that once the dominance in major-axis $\mathbf{w^0}$ is shaped by evolution, linearization against environmental condition $\delta\mathbf{E}$ is no longer necessary. Further, the dominance can be shaped through evolution under given fitness[34], in which case the global proportionality still holds. Under the dilution by the growth-rate, then, proportionality to $\delta\mu$ is shaped as shown here.

Since the phenotypic change due to noise, environmental, or evolutionary change is constrained and larger along this major axis $\mathbf{w^0}$, the phenotypic property along this axis can be considered to be plastic. (see Figs. 4 and 6). After evolution, however, the robustness of fitness against mutation and noise also evolves[21, 22, 34], as is also discussed in nearly neutral theory[39]. Hence, the environmental or genetic change is buffered to the change along the major axis, $\delta\mathbf{X}\cdot\mathbf{w^0}$, while the fitness (growth rate) change achieves robustness after evolution. Indeed, the compatibility between plasticity and robustness remains one of the major issues in biology[40–45].

The question remains as to why such an axis, along which the change is much larger than that occurring along other directions, is shaped by evolution. As mentioned above, the robustness is shaped by evolution, so that stronger attraction is achieved in the state space. By contrast, along the axis in which phenotypes change through evolution, the plasticity, i.e., changeability, remains. Hence, phenotype dynamics against directions except or those along this axis are expected to have strong contractions as evolution progresses. This leads to the dominance of change in the major axis, as discussed here.

Note that the dominance in changes along the major axis, once it is shaped by evolution, can facilitate further evolution. If there were no such directions, and instead random mutations in many directions were the main source of phenotypic change, the changes would mostly cancel each other out, so that the phenotypic changes corresponding to specific fitness growth would be suppressed, resulting in the reduction of evolvability. Instead, if all the changes are buffered into the mode along the major axis $\mathbf{w^0}$, then a certain degree of phenotypic change would be assured to foster subsequent evolution.

In fact, the variation in X due to environmental or evolutionary change remains rather large. Both the principal components analyses in bacterial evolution[18] and the results of the present study suggest that the evolutionary course in gene expression patterns can be described by one (or a few) major degrees of freedom, while this "major axis" extracted by the principal component analysis is highly correlated with the growth rate of the cell.

Hence, the present study opens up the possibility to describe adaptation and evolution in response to environmental changes by only a few collective variables that dominantly change. With this representation, evolvability, i.e., the changeability of a phenotype through evolution, can be quantitatively characterized. This changeability is also represented by phenotypic fluctuations due to noise or environmental variances, as in the evolutionary fluctuation-response relation[28, 34]. In other words, the phenotype that is more evolvable is predetermined before a genetic change occurs, which is a manifestation of the genetic assimilation concept proposed by Waddington[40].

Of course, further studies are needed to confirm the generality of the evolution of dominance of major-axis mode, both through experiments and numerical simulations. For example, we previously studied gene regulation networks in which the expression of many genes mutually activate or suppress the expression of other genes[22, 34] with consequent dilution of the expressed protein concentrations due to cell volume growth. Preliminary results suggest that global proportionality is general to the evolved networks; that is, all expression levels change in proportion up to a large change in growth rates, where the proportion coefficient agrees with the growth rates, and the expression changes are located along a one-dimensional manifold, as represented by the principal component. Here, it is noted that the dominance in the major-axis (or slow-manifold) itself can be a general consequence of evolution to achieve robustness[22], even if the global dilution by the growth rate is not imposed. By imposing it, however, the proportionality of the component with $\delta\mu$ is achieved.

Another important future extension of our theory will be formulation beyond the steady growth state. In cells, there are states in which growth is suppressed as the stationary state[1, 46, 47], and some theoretical approaches have been proposed to capture these non-growing states[5]. Thus, it will be important to determine if (or how) the incorporation of more degrees of freedom beyond the major axis will be necessary to capture the state with $\mu \sim 0$.

bioRxiv preprint doi: https://doi.org/10.1101/125278; this version posted April 7, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Acknowledgment

The authors would like to thank Matthias Heinemann for useful discussions of his experiment. The authors would like to thank Tetsuhiro Hatakeyama, Bingkan Xue, and Pablo Sartori for stimulating discussions. This research was partially supported by the Platform for Dynamic Approaches to Living Systems from the Japan Agency for Medical Research and Development (AMED) [to K.K], Grant-in-Aid for Scientic Research (S) (15H05746 [to K.K. and C.F]), and by Grant-in-Aid for Scientific Research (B) (15H04733, 15KT0085 [to CF]) from the Japanese Society for the Prmotion of Science (JSPS). KK would also like to thank Charles L Brown Membership at Institute for Advanced Study at Princeton.

Appendix: Methods for model simulations of a replicating cell

The cellular state is represented by the numbers of k components $(N_1, N_2, \cdots, N_k)$ and their concentrations $x_i = N_i/V$ with the volume of the cell V. For intracellular reaction dynamics, we consider a catalytic network among these k chemical species, where each reaction from some chemical i to some other chemical j was assumed to be catalyzed by a third chemical ℓ, i.e., $(i + \ell \to j + \ell)$. The catalytic network was generated randomly, where the probability that chemical i is converted from chemical j is given by the connection rate ρ. For simplicity, all reaction coefficients were chosen to be equal. There are some resource (nutrient) chemicals whose concentrations in the environment and in the cell are given by $s_1, \cdots, s_m$ and $x_1, \cdots, x_m$, respectively. The resource chemicals are transported into the cell with the aid of other chemicals called "transporters". Here, we assumed that the uptake flux of nutrient i from the environment is proportional to $Ds_i x_{t_i}$, where chemical t_i acts as the transporter for nutrient i and D is a transport constant, and that for each nutrient there is one corresponding transporter, represented by $t_i = m + i$. Both the nutrients and transporters have no catalytic activity, whereas the other $k - 2m$ chemical species work as catalysts. Through catalytic reactions, these nutrients are transformed into other chemicals, including the transporters. With the intake of nutrient chemicals from the environment, the total number of chemicals $N = \sum_i N_i$ in a cell can increase; thus, we assumed that the cell volume V is proportional to the total number N. For simplicity, cell division is assumed to occur when the cell volume exceeds a given threshold. Through cell division, the chemicals in the parent cell are evenly split between 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 among molecules within the cell and among nutrients in the environment. The parameters were set as $k = 1000$, $m = 10$, $\rho = 0.075$, and $D = 0.001$.

Using this model, we analyzed the evolutionary dynamics by generating slightly modified networks and selecting those that grew faster. First, n parent cells were generated, whose catalytic reaction networks were randomly generated using the connection rate ρ. From each of the n parent cells, L mutant cells were generated by randomly replacing $m\rho k^2$ reaction paths, where ρk^2 is the total number of reactions and m is the mutation rate per reaction per generation. To determine the growth rate of each cell, i.e., the inverse of the average time required for division, reaction dynamics were simulated for each of the nL cells. Within the cell population, n 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. The parameters were set as $n = 1000$, $L = 50$, and $m = 1 \times 10^{-3}$. The simulation of evolutionary dynamics was performed under a constant (original) environmental condition $\mathbf{o} = \{s_1^o, \cdots, s_m^o\}$, where $s_1^o = s_2^o = \cdots = s_m^o = 1/m$.

The environmental change was simulated by changing the concentration of external nutrients $s_1, \cdots, s_m$. The vector representing a new (stressed) environment $\mathbf{E} = \{s_1^{\mathbf{E}}, \cdots, s_m^{\mathbf{E}}\}$ was generated, in which the values of component $s_1^{\mathbf{E}}, \cdots, s_m^{\mathbf{E}}$ were determined randomly to satisfy $\sum_j s_j^{\mathbf{E}} = 1$. Using the vector $\mathbf{E}$, the perturbed environmental condition was given as $s_j^{(\varepsilon,\mathbf{E})} = (1-\varepsilon)s_j^o + \varepsilon s_j^{\mathbf{E}}$, where ε is the strength of the stress.

bioRxiv preprint doi: https://doi.org/10.1101/125278; this version posted April 7, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

* chikara.furusawa@riken.jp
† kaneko@complex.c.u-tokyo.ac.jp

[1] Monod, J. (1949) The growth of bacterial cultures. *Ann. Rev. Microbiol.* 3:371-394.
[2] Pirt SJ. (1965) The maintenance energy of bacteria in growing cultures. *Proc. R. Soc. London B* 163(991):224-231.
[3] Scott M, Gunderson CW, Mateescu EM, Zhang Z, Hwa T. (2010) Interdependence of cell growth and gene expression: origins and consequences. *Science* 330:1099-1102.
[4] Klumpp S, Zhang Z, Hwa T. (2009) Growth rate-dependent global effects on gene expression in bacteria. *Cell* 139(7):1366–1375.
[5] Maitra A, Dill KA. (2015) Bacterial growth laws reflect the evolutionary importance of energy efficiency. *Proc. Natl. Acad. Sci.* 112(2):406–411.
[6] Carroll SM and Marx CJ. (2013) Evolution after introduction of a novel metabolic pathway consistently leads to restoration of wild-type physiology. *PLoS Genet.* 9(4):e1003427.
[7] Pancaldi V, Schubert F, Bahler J. (2010) Meta-analysis of genome regulation and expression variability across hundreds of environmental and genetic perturbations in fission yeast. *Mol. Biosyst.* 6(3):543–552.
[8] Keren L, et al. (2013) Promoters maintain their relative activity levels under different growth conditions. *Mol. Syst. Biol.* 9:701
[9] Stern S, Dror T, Stolovicki E, Brenner N, Braun E (2007). Genome-wide transcriptional plasticity underlies cellular adaptation to novel challenge. *Mol. Syst. Biol.* 3(1):106.
[10] Matsumoto Y, Murakami Y, Tsuru S, Ying BY, Yomo T. (2013) Growth rate-coordinated transcriptome reorganization in bacteria. *BMC Genomics* 14(1):808.
[11] López-Maury L, Marguerat S, Bähler J. (2008) Tuning gene expression to changing environments: from rapid responses to evolutionary adaptation. *Nat. Rev. Genet.* 9(8):583—593.
[12] Bergmann S, Ihmels J, Barkai N. (2003) Similarities and differences in genome-wide expression data of six organisms. *PLoS Biol.* 2(1):E9.
[13] Marguerat S, Schmidt A, Codlin S, Chen W, Aebersold R, Bhler J. (2012) Quantitative analysis of fission yeast transcriptomes and proteomes in proliferating and quiescent cells. *Cell* 151(3):671–683.
[14] Lehner B. (2010) Genes confer similar robustness to environmental, stochastic, and genetic perturbations in yeast. *PLoS One* 5(2):e9035.
[15] Horinouchi T, et al. (2010) Transcriptome analysis of parallel-evolved *Escherichia coli* strains under ethanol stress. *BMC Genomics* 11:579.
[16] Horinouchi T, et al. (2015) Phenotypic convergence in bacterial adaptive evolution to ethanol stress. *BMC Evol. Biol.* 16:802.
[17] Kaneko K, Furusawa C, Yomo T. (2015) Universal relationship in gene expression changes for cells in steady-growth state. *Phys. Rev. X* 5:011014.
[18] Furusawa C, Kaneko K. (2015) Global relationships in fluctuation and response in adaptive evolution.*R. Soc. Interface* 12:20150482.
[19] Fontana W, Schuster P. (1998) Shaping space: the possible and the attainable in RNA genotype-phenotype mapping. *J. Theor. Biol.* 194:491–515.
[20] Wagner GP, Zhang J. (2011) The pleiotropic structure of the genotypephenotype map: the evolvability of complex organisms. *Nat. Rev. Genet.* 12(3):204–213.
[21] Ciliberti S, Martin OC, Wagner A. (2007) Robustness can evolve gradually in complex regulatory gene networks with varying topology. *PLoS Comp. Biol.* 3:e15.
[22] Kaneko K. (2007) Evolution of robustness to noise and mutation in gene expression dynamics. *PLoS One* 2(5):e434.
[23] Soyer OS. (2012) *Evolutionary Systems Biology*. Springer.
[24] Schmidt A, et al. (2016) The quantitative and condition-dependent *Escherichia coli* proteome *Nat. Biotech.* 34(1):104.
[25] Elowitz MB, Levine AJ, Siggia ED, Swain PS. (2002) Stochastic gene expression in a single cell. *Science* 297(5584):1183.
[26] Furusawa C, et al. (2005) Ubiquity of log-normal distributions in intracellular reaction dynamics. *Biophysics* 1:25–31.
[27] Bar-Even, et al. (2006) Noise in protein expression scales with natural protein abundance. *Nat. Genet.* 38:636–643.
[28] Sato K, Ito Y, Tomo T, Kaneko K. (2003) On the relation between fluctuation and response in biological systems. *Proc. Natl. Acad. Sci. USA* 100(24):14086–14090.

bioRxiv preprint doi: https://doi.org/10.1101/125278; this version posted April 7, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

[29] Haldane JBS. (1954) The measurement of natural selection. *Proc. 9th Int. Cong. Genet* 1:480–487.
[30] Furusawa C, Kaneko K. (2003) Zipf's law in gene expression. *Phys. Rev. Lett.* 90:088102.
[31] Furusawa C, Kaneko K. (2012) Adaptation to optimal cell growth through self-organized criticality. *Phys. Rev. Lett.* 108(20):208103.
[32] Mathematically, it would be represented as slow manifold.
[33] Recall projection to collective variables in statistical physics; Mori H. (1965) Transport, collective motion, and Brownian motion. *Prog. Theor. Phys.* 33:423–455.
[34] Kaneko K. (2012) Phenotypic plasticity and robustness: evolutionary stability theory, gene expression dynamics model, and laboratory experiments. In: Evolutionary Systems Biology (pp. 249-278). Springer New York.
[35] Stearns SC, Kaiser M, Kawecki TJ. (1995) The differential genetic and environmental canalization of fitness components in *Drosophila melanogaster*. *J. Evol. Biol.* 8:539.
[36] Landry C, et al. (2007) Genetic properties influencing the evolvability of gene expression. *Science* 317:118–121.
[37] Fisher RA. (1930) *The Genetical Theory of Natural Selection*. Clarendon Press, Oxford.
[38] Hartl DL, Clark AG. (2007) *Principles of Population Genetics*, 4th ed., Sinauer Associates, Massachusetts.
[39] Ohta T. (2011) Near-neutrality, robustness, and epigenetics. *Genome Biol. Evol.* 3:1034–1038.
[40] Waddington CH. (2014) *The strategy of the Genes*. Routledge.
[41] de Visser, et al. (2003) Perspective: evolution and detection of genetic robustness. *Evolution* 57(9):1959–1972.
[42] Callahan HS, Pigliucci M, Schlichting CD. (1997) Developmental phenotypic plasticity: where ecology and evolution meet molecular biology. *Bioessays* 19:519–525.
[43] Pigliucci M, Murren CJ, Schlichting CD. (2006) Phenotypic plasticity and evolution by genetic assimilation. *J. Exp. Biol.* 209:2362–2367.
[44] Kirschner MW, Gerhart JC. (2005) *The Plausibility of Life*. Yale Univ. Press, New Haven.
[45] Hatakeyama TS, Kaneko K. (2015) Reciprocity between robustness of period and plasticity of phase in biological clocks. *Phys. Rev. Lett.* 115(21):218101.
[46] Novick A. (1955) Growth of bacteria. *Ann. Rev. Microbiol.* 9(1):97–110.
[47] Gefen O, Fridman O, Ronin I, Balaban NQ. (2014). Direct observation of single stationary-phase bacteria reveals a surprisingly long period of constant protein production activity. *Proceedings of the National Academy of Sciences*, 111(1):556– 561.
[48] Watson DI, Wright IJ, Falster DS, Westoby M. (2006) Bivariate line-fitting methods for allometry *Biol. Rev.* 82(2):259.

Figure captions

Figure 1. Common proportionality in *E. coli* protein expression changes. (a) and (b) show examples of the relationship between protein expression changes $\delta X_j(\mathbf{E_1})$ and $\delta X_j(\mathbf{E_2})$. $\delta X_j(\mathbf{E})$ represents the log-transformed expression changes of the j-th protein between an environment $\mathbf{E}$ and the standard condition. Here, we chose the M9 minimal medium with glucose as the carbon source for the standard condition. (a) shows the relationship between the glycerol carbon source condition and chemostat culture with $\mu = 0.12\text{h}^{-1}$, and (b) shows the relationship between the mannose carbon source and acetate carbon source conditions. The blue line represents the slope calculated by the ratio of observed growth rate changes. In this analysis, we used the protein expression data of genes with expression levels under all environments that exceed the threshold ($x_j(\mathbf{E}) > 50$), in order to exclude any inaccurate data (about 40% of the expression data were discarded). (c) Relationship between the slope of the change in protein expression and the change in growth rate. The abscissa represents $\delta\mu(\mathbf{E_1})/\delta\mu(\mathbf{E_2})$, while the ordinate is the slope in $\delta X_j(\mathbf{E_1})/\delta X_j(\mathbf{E_2})$. The slope was obtained by fitting the protein expression data using the major axis method, which is a least-squares fit method that treats the horizontal and vertical axes equally, and is commonly used to fit bivariate scatter data[48]. The pairs $\mathbf{E_1}$ and $\mathbf{E_2}$ satisfying $\mu(\mathbf{E_2}) < \mu(\mathbf{E_1}) < \mu(\mathbf{S})$ were plotted, where $\mu(\mathbf{S})$ represents the growth rate under the standard condition.

Figure 2. Common proportionality in concentration changes in response to the same type of stress. (a) Correlation coefficient between the changes in component concentrations $\delta X_j^{(\varepsilon_1,\mathbf{E})}$ and $\delta X_j^{(\varepsilon_2,\mathbf{E})}$ with $\varepsilon_2 = \varepsilon_1 + \varepsilon$. For the random and evolved networks, the correlation coefficients with a small ($\varepsilon = 0.02$)

bioRxiv preprint doi: https://doi.org/10.1101/125278; this version posted April 7, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

and large ($\varepsilon = 0.08$) environmental changes are plotted, which were obtained using 100 randomly chosen environmental vectors $\mathbf{E}$. Throughout the study, each component concentration δX_j was calculated by averaging 100 cell division events after the cell reaches a steady growing state, except for the results shown in Fig. 6. (b) Relationship between the ratio of growth rate change $\delta\mu^{(\varepsilon_1,\mathbf{E})}/\delta\mu^{(\varepsilon_2,\mathbf{E})}$ and the slope in $(\delta X_j^{(\varepsilon_1,\mathbf{E})}, \delta X_j^{(\varepsilon_2,\mathbf{E})})$ obtained by fitting the concentration changes of all components. (c) Ratio of the slope in the concentration changes to the growth rate change as a function of the strength of stress ε. The ratio in the ordinate becomes unity when Eq. (1) is satisfied.

Figure 3. Common proportionality in concentration changes in response to different types of stress. (a) Correlation coefficient between the changes in component concentrations $\delta X_j^{(\varepsilon_1,\mathbf{E_1})}$ and $\delta X_j^{(\varepsilon_2,\mathbf{E_2})}$. The data were obtained using 100 randomly chosen pairs of two environmental vectors $\mathbf{E_1}$ and $\mathbf{E_2}$. (b) Relationship between the ratio of growth rate change $\delta\mu^{(\varepsilon_1,\mathbf{E_1})}/\delta\mu^{(\varepsilon_2,\mathbf{E_2})}$ and the slope in $(\delta X_j^{(\varepsilon_1,\mathbf{E_1})}, \delta X_j^{(\varepsilon_2,\mathbf{E_2})})$. (c) Ratio of the slope in the relationship between concentration changes and growth rate change as a function of the strength of stress ε in the case of different types of stress.

Figure 4. The change of $X_j^{(\varepsilon,\mathbf{E})}$ with environmental changes in the principal components space. The component concentration $X_j^{(\varepsilon,\mathbf{E})}$ under randomly chosen various $\mathbf{E}$ and ε values are presented for (a) evolved and (b) random networks. Here, the contribution of the first, second, and third components is 74%, 8%, and 5%, respectively.

Figure 5. Growth rate plotted against the principal components using the computation and data in Fig. 4(a).

Figure 6. Change of the component concentration X_j by (a) mutations and (b) noise in reaction dynamics. In (a), mutations were added to the evolved reaction network by randomly replacing 0.5% of the reaction paths. The red dots show the concentrations of components after mutations, which are projected on the same principal component space as that depicted in Fig. 4(a). The gray dots represent the concentration changes caused by environmental changes for the reference, which are identical to those shown in Fig. 4(a). The red dots in (b) represent the concentration changes observed in each cell division, which are caused by the stochastic nature of reaction dynamics. These data (as well as those in Fig. 4) were obtained under the original environmental condition used for the evolutionary simulation.

Figure 7 Schematic representation of the major axis hypothesis. In the state space of $\mathbf{X}$, dominant changes are constrained along the $\mathbf{w^0}$ axis (and its connected manifold), while the attraction to this manifold is faster.

Figure 8 Change in X_i plotted against the change in μ(a) and against ε(b) in the cell model. Only a few components i among 1000 are plotted for the evolved network. The linear change is more extended in the plot against μ than in that against ε.

bioRxiv preprint doi: https://doi.org/10.1101/125278; this version posted April 7, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

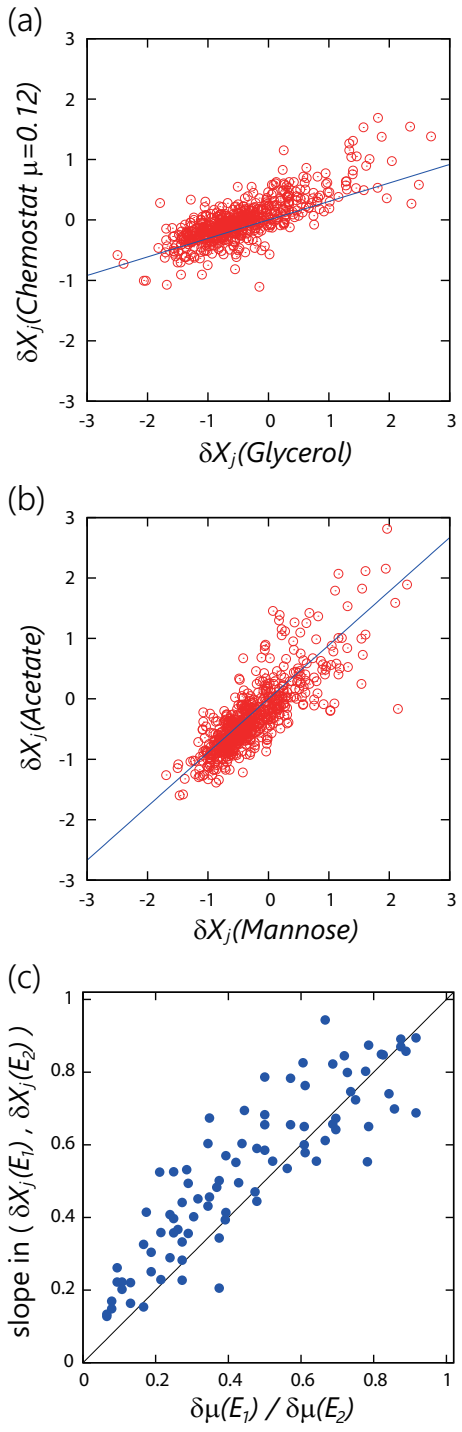

FIG. 1.

bioRxiv preprint doi: https://doi.org/10.1101/125278; this version posted April 7, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

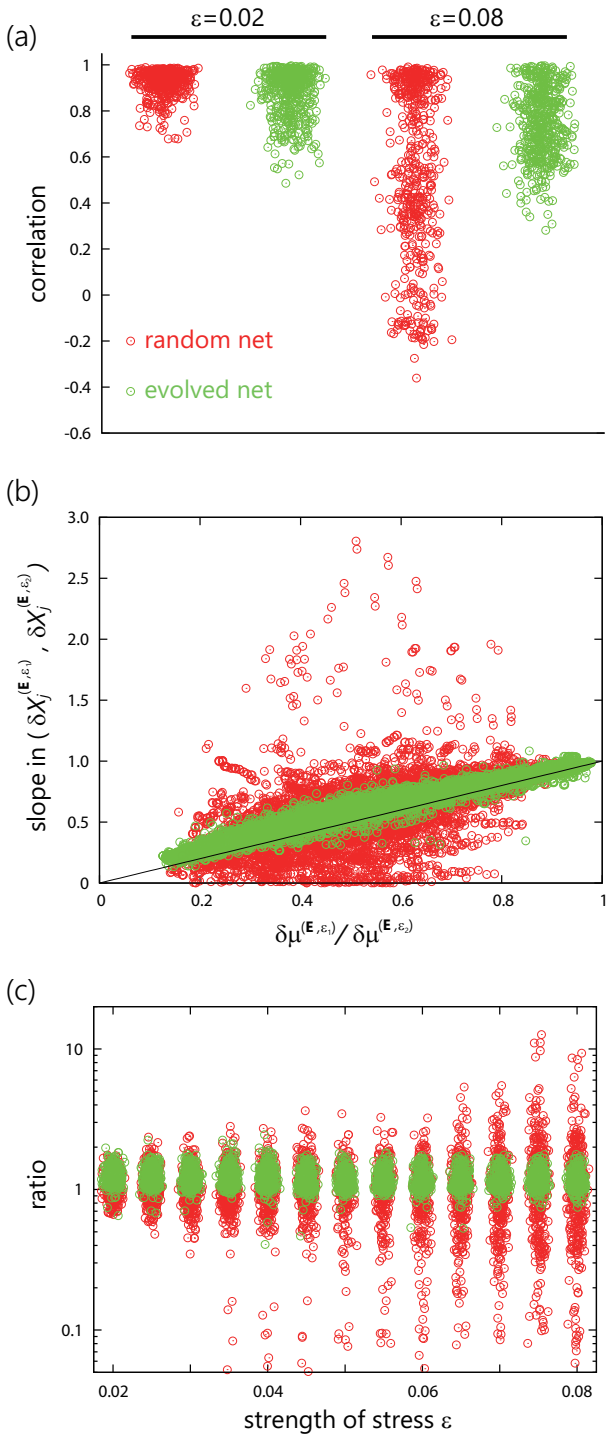

FIG. 2.

bioRxiv preprint doi: https://doi.org/10.1101/125278; this version posted April 7, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

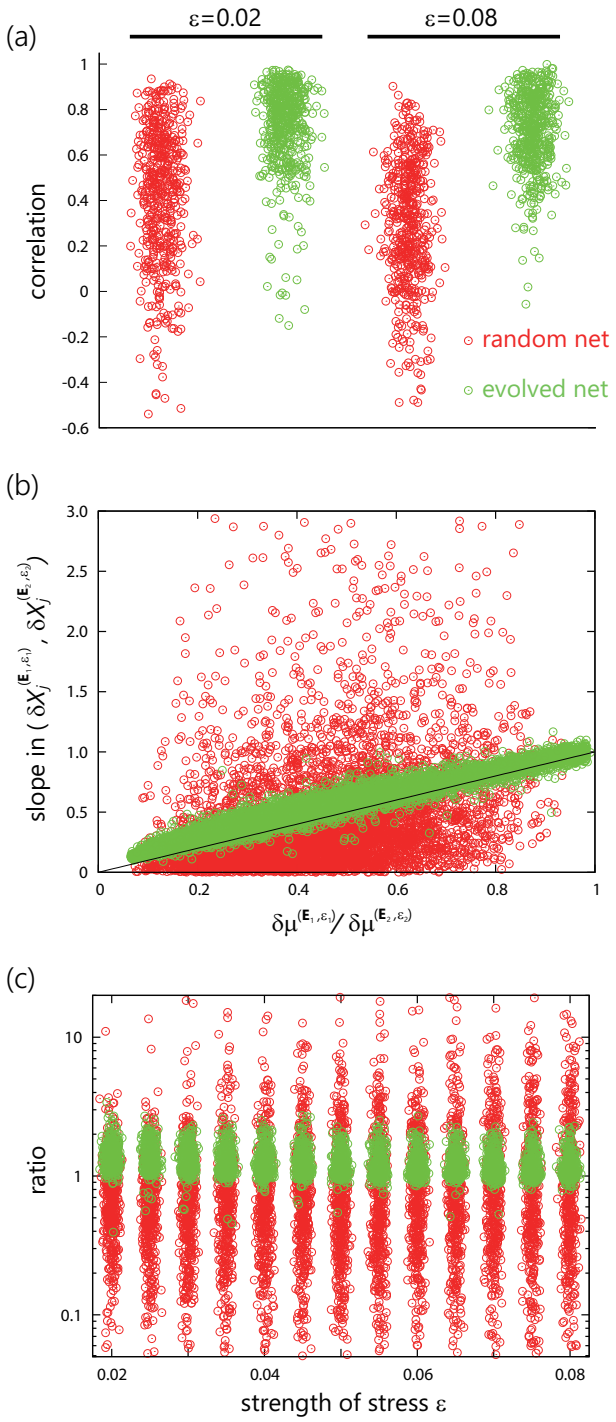

FIG. 3.

bioRxiv preprint doi: https://doi.org/10.1101/125278; this version posted April 7, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

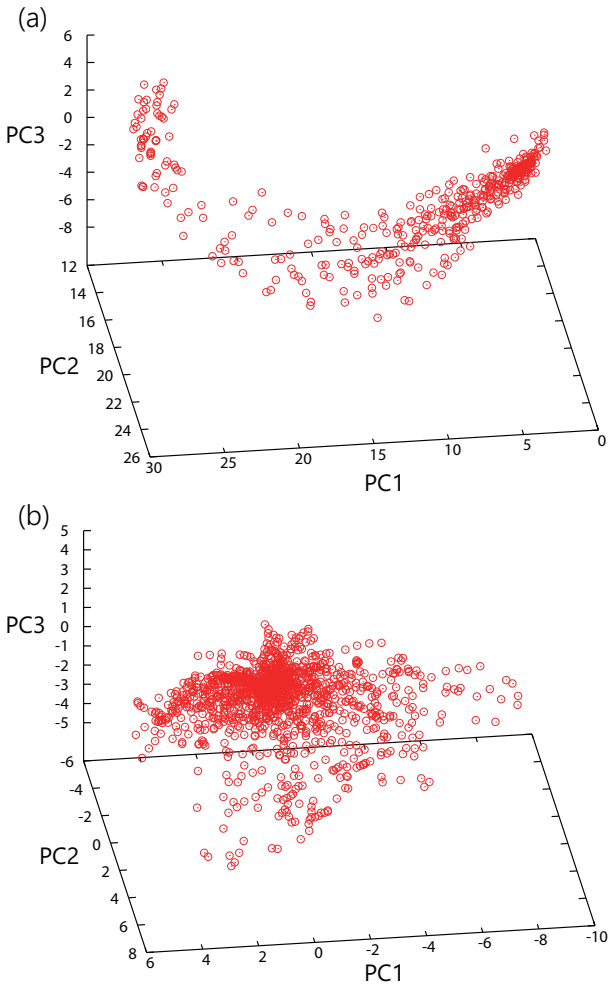

FIG. 4.

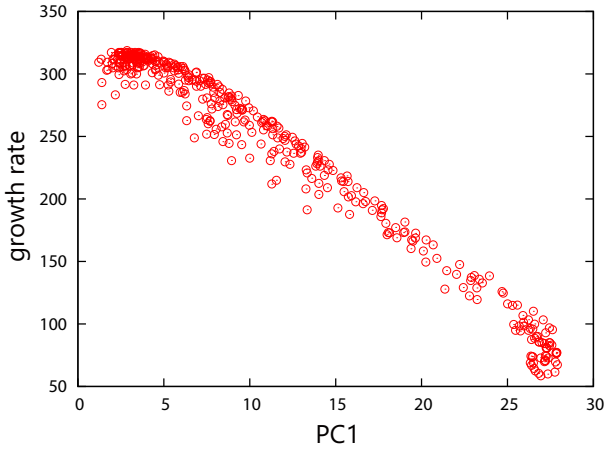

FIG. 5.

bioRxiv preprint doi: https://doi.org/10.1101/125278; this version posted April 7, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

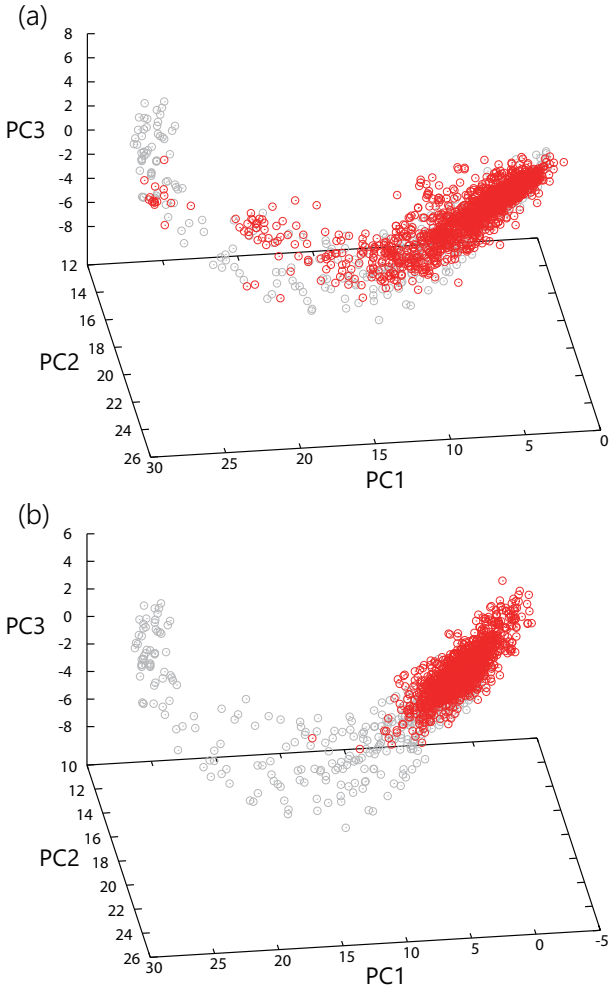

FIG. 6.

bioRxiv preprint doi: https://doi.org/10.1101/125278; this version posted April 7, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

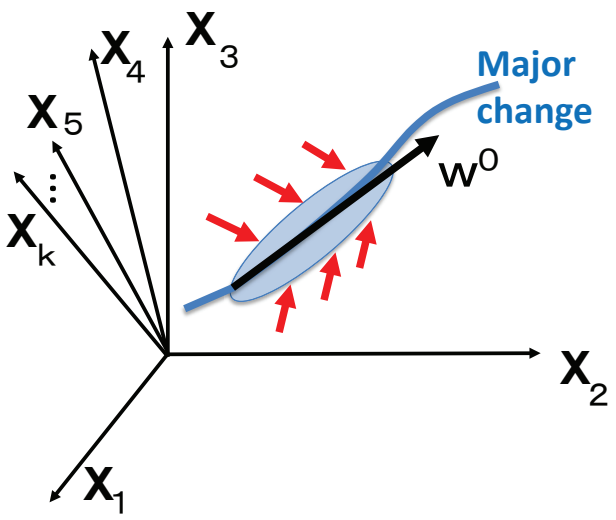

FIG. 7.

bioRxiv preprint doi: https://doi.org/10.1101/125278; this version posted April 7, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

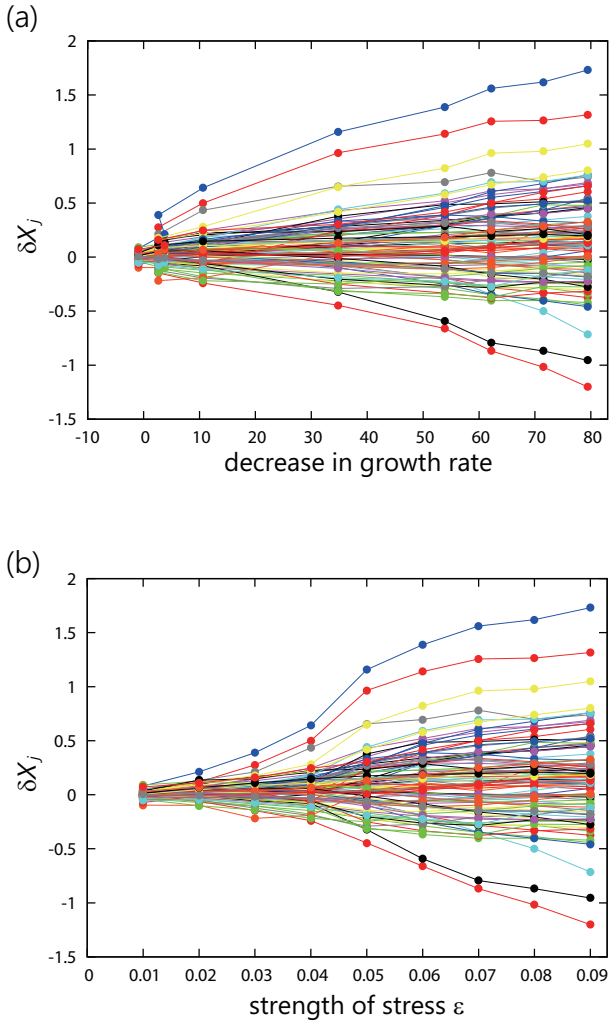

FIG. 8.