Evolutionary origin of power-laws in a biochemical reaction network: Embedding the distribution of abundance into topology

Chikara Furusawa[1,2] and Kunihiko Kaneko[3,2]
[1]*Department of Bioinformatics Engineering, Graduate School of Information Science and Technology, Osaka University, 2-1 Yamadaoka, Suita, Osaka 565-0871, Japan*
[2]*ERATO Complex Systems Biology Project, JST, 3-8-1 Komaba, Meguro-ku, Tokyo 153-8902, Japan*
[3]*Department of Pure and Applied Sciences, University of Tokyo, Komaba, Meguro-ku, Tokyo 153-8902, Japan*
(Received 21 February 2005; revised manuscript received 1 September 2005; published 23 January 2006)

The evolutionary origin of general statistics in a biochemical reaction network is studied here to explain the power-law distribution of reaction links and the power-law distribution of chemical abundance. Using cell models with catalytic reaction networks, we have confirmed that the power-law distribution for the abundance of chemicals emerges by the selection of cells with higher growth rates, as suggested in our previous study [Phys. Rev. Lett. **90**, 088102 (2003)]. Through further evolution, this inhomogeneity in chemical abundance is shown to be embedded in the distribution of links, leading to the power-law distribution. We analyze the mechanism of this embedding and discuss the generality of the results.

DOI: 10.1103/PhysRevE.73.011912 PACS number(s): 87.17.Aa, 87.23.Kg, 89.75.Hc

I. INTRODUCTION

Recent advances in biology have provided detailed knowledge about individual molecular processes and their functions, leading to great success in explaining life in terms of molecules. From the accumulated data, it is now important to unveil universal features. However, it is often difficult to understand the universal characteristics of the intracellular dynamics maintaining the living state solely by building up detailed knowledge of molecules, because there is such a complex and essential network of reactions among these molecules, such as proteins, DNA and RNA. Thus, constructing a model to cover all the details therein is impossible, considering the enormous diversity of cellular processes. Therefore, one possible strategy for extracting the nature of intracellular dynamics is to search for universal laws with regard to the networks of intracellular reactions common to a class of cell models—albeit simple—and then to unravel the dynamics of evolution leading to such universal features [1]. Indeed, recent large-scale studies have revealed two general features in cellular dynamics. First, the power-law distribution of links in reaction networks was discovered in metabolic and other biochemical pathways and is termed a scale-free network, where the connectivity distribution $P(k)$ obeys the law $k^{-\gamma}$ with $\gamma \approx (2 \sim 3)$ [2–6]. Second, the abundance of chemicals in intracellular reactions was also found to exhibit a power-law distribution, as confirmed at the levels of gene expression [1,7,8] and metabolic flux [9]. Here, the chemical abundances plotted in the order of their magnitude are inversely proportional to their rank.

Despite the potential importance of these general statistical laws, how they are formed through evolution and how the two laws are mutually related are still unknown. Here, we attempt to answer these questions through analysis and simulation of the evolution of a simple cell model, to demonstrate that a power-law distribution in the abundance (i.e., total amount in a cell) of particular chemicals emerges as a result of competition for greater growth of a cell. This inhomogeneity in abundance is embedded into the distribution of links, leading to a so-called scale-free network with hierarchical organization of reaction dynamics. The findings provide novel insights into the nature of network evolution in living cells.

To determine the emergence and interrelationships of the power-laws in chemical abundances and network connectivity through the process of evolution, we adopted a simple model of intracellular reaction dynamics that captures the catalytic reaction processes essential for cell growth and division, following Refs. [1,10,11]. Although this model was chosen simply to satisfy the minimal requirements of the intracellular reaction dynamics of a growing cell, it was found to capture general statistical behavior patterns as confirmed experimentally [1]. By studying a class of simple models with these features and the evolution of the network of the reaction, we can study how the power laws in abundances and network connectivity emerges inevitably.

II. MODEL

Consider a cell consisting of a variety of chemicals. The internal state of the cell can be represented by a set of concentrations $(x_1, x_2, \ldots, x_K)$, where x_i is the intracellular concentration of the chemical species i with values ranging from 1 to K. Depending on whether there is an enzymatic reaction from i to j catalyzed by some other chemical ℓ, the reaction path is connected as $(i+\ell \rightarrow j+\ell)$. The rate of increase of x_j through this reaction is given by $x_i x_\ell$, where, for simplicity, all of the reaction coefficients have been set at 1.

Next, some nutrients were supplied from the environment by transportation through the cell membrane with the aid of some other chemicals, i.e., "transporters." Here, we assumed that the rate of transport of a chemical is proportional to its concentration, and the rate of increase of x_i by such transportation is given by $Dx_{m_i}(X_i - x_i)$, where the m_ith chemical acts as the transporter for the nutrient i and x_{m_i} is the concentra-

tion of the m_ith chemical. The parameter D is a transport constant, and the constant X_i is the concentration of the ith chemical in the environment. In addition, we took into account the changes in cell volume, which varies as a result of transportation of chemicals into the cell from the environment. For simplicity, we assumed that the volume is proportional to the sum of the chemicals in the cell, which can increase by the uptake of nutrients. The concentrations of chemicals are diluted because of increases in the volume of the cell. Based on the above assumptions, this dilution effect is equivalent to imposing the the restriction $\Sigma_i x_i = 1$. When the volume of a cell is doubled because of nutrient intake, the cell is assumed to divide into two identical daughter cells.

To summarize these processes, the dynamics of chemical concentrations in each cell are represented as

$$dx_i/dt = R_i - x_i \sum_j R_j \tag{1}$$

with

$$R_i = \sum_{j,\ell} \mathrm{Con}(j,i,\ell) x_j x_\ell - \sum_{j',\ell'} \mathrm{Con}(i,j',\ell') x_i x_{\ell'} (+ D x_{m_i} (X_i - x_i)), \tag{2}$$

where $\mathrm{Con}(i,j,\ell)$ is 1 if there is a reaction $i+\ell \rightarrow j+\ell$, and 0 otherwise, while the last term in R_i is added only for the nutrients and represents its transportation into a cell from the environment. The last term in dx_i/dt with the sum of R_j gives the constraint of $\Sigma_i x_i = 1$, because of the growth of the volume.

Of course, how these reactions progress depends on the intracellular reaction network. Here, we study the evolution of the network, by generating slightly modified networks and selecting those that grow faster. First, n mother cells are generated, where the connecting paths of catalytic networks were chosen randomly, so that the numbers of incoming, outgoing, and catalyzing paths of each chemical are set to the initial path number k_{init}. From each of n mother cells, m mutant cells were generated by the random addition of one reaction path to the reaction network of the mother. Then, reaction dynamics were simulated for each of the $n \times m$ cells to determine the rate of growth of each cell; that is, the inverse of the time required for division. Within the cell population, n cells with faster growth rates were selected to be the mother cells of the next generation, from which m mutant cells were again generated in the same manner.

III. RESULT: POWER-LAWS IN ABUNDANCES AND NETWORK STRUCTURE ACHIEVED THROUGH EVOLUTION

A number of network evolution simulations were performed using several different initial networks, parameters, and various settings. We found that all of the simulations indicated common statistical properties with regard to both reaction dynamics and the topology of networks. Here, we present an example of simulation results to show the common properties of our simulations.

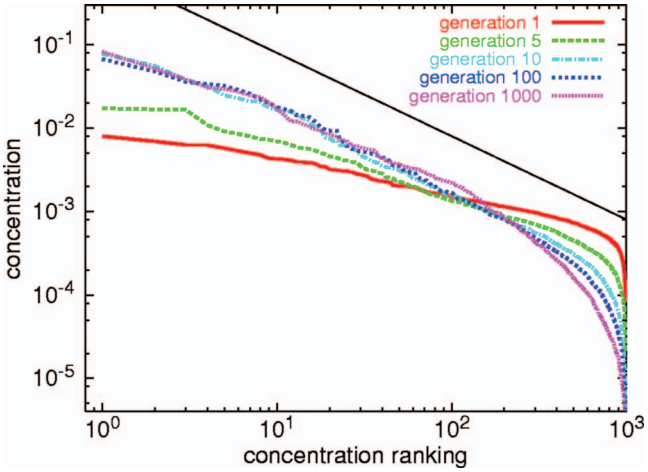

FIG. 1. (Color) Rank-ordered concentration distributions of chemical species. Distributions with several different generations are superimposed using different colors. The solid line indicates the power-law $x \propto n^{-1}$ for the reference. This power-law of chemical abundance is established around the tenth generation and is sustained for further evolutions in the network. In this simulation, growth rates of 10×2000 networks were measured, and the top ten networks with regards to the growth rate were chosen for the next generation. The parameters were set as $K=1000$, $D=4.0$, and $k_{init}=4$. Chemicals x_m for $m<5$ are considered as nutrients, and their concentrations in the environment are set as $X_m=0.2$. For each nutrient chemical, one transporter chemical is chosen randomly from all other chemicals.

The rank-ordered concentration distributions of chemical species in several generations are plotted in Fig. 1, in which the ordinate indicates the concentration of chemical species x_i and the abscissa shows the rank determined by x_i. The slope of the rank-ordered concentration distribution increased with each generation and, within a few generations, converged to a power-law distribution with an exponent -1, which was maintained over further generations. Or, equivalently, the distribution $p(x)$ of the species with abundance x is proportional to x^{-2} [12].

Indeed, the emergence of such a power-law by selecting cells with higher growth rates is a natural consequence of our previous study [1]. There, we found that there is a critical amount of nutrient uptake beyond which the cell cannot grow continuously. When the nutrient uptake is larger than the critical amount, the flow of nutrients from the environment is so fast that the internal reactions transforming them into chemicals sustaining "metabolism" and transporters cannot keep up. At this critical amount of nutrient uptake, the growth rate of a cell becomes maximal, and the power-law distribution of chemical abundance appears in the intracellular dynamics. This power-law distribution at the critical state is maintained by a hierarchical organization of catalytic reactions and, based on this catalytic hierarchy, the observed exponent -1 can be explained by using a mean field approximation. Experimentally, the power-law distributions of chemical abundances were confirmed from large-scale gene expression data of various organisms and tissues, including yeast, nematodes, normal and cancerous human tissues, and embryonic stem cells. This suggests that the intracellular re-

action dynamics in real cell systems generally lie close to the critical state (see Ref. [1] for details).

In the evolutionary dynamics of the present simulations, to increase the growth rate of cells, network changes that enhance the uptake of nutrients from the environment are favored. This nutrient uptake is facilitated by increasing the concentrations of transporters, although, if the uptake of nutrients is too large, the cell can no longer grow continuously because they exceed a critical amount, as mentioned above. Now, with the evolutionary process shown in Fig. 1, the nutrient uptakes increase to accelerate the growth rate of cells until further mutations of the network may lead the system to exceed the above critical value of the nutrient uptake. Here, successive increases in the growth rate by the "mutation" to the reaction network is possible only when the enhancement of nutrient uptakes caused by it is in step with increases in the other catalytic activities. As a natural consequence, networks are selected so that the nutrient uptake is kept near this critical point, where successive catalytic reaction processes maximize the use of nutrients and form a power-law distribution of abundance.

Next, we investigated the topological properties of the reaction networks. The connectivity distributions $P(k)$ of chemical species obtained from the network of the 1000th generation are plotted in Fig. 2(a), where k_{in}, k_{out}, and k_{cat} indicate the numbers of incoming, outgoing, and catalyzing paths of chemicals, respectively. These distributions were fitted by power-laws with an exponent close to −3. Thus, a scale-free network was approached through evolution, and this power-law behavior was maintained for further evolutionary processes.

As shown in Fig. 3, in this simple model, the evolved reaction network formed a cascade structure in which each chemical species was mainly synthesized from more abundant species. That is, almost no chemical species disrupted the flow of chemical reactions from the nutrients, as the network approached the point of optimal cell growth. It should be noted that the reaction dynamics for each chemical were also inhomogeneous, in that synthesis of each chemical species had a dominant reaction path. Such an uneven use of local reaction paths has been reported in real metabolic networks [9].

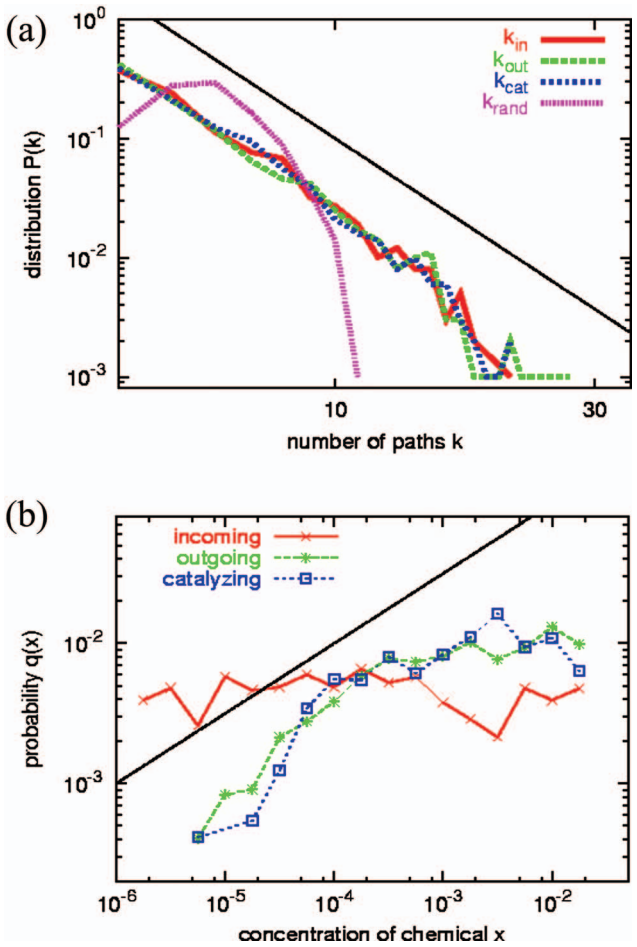

FIG. 2. (Color) Evolution of the network topology. (a) Connectivity distribution $P(k)$ of chemical species obtained from the network of the 1000th generation. The solid line indicates the power-law $P(k) \propto k^{-3}$. For comparison, we show the distribution of k_{rand}, obtained by a randomly generated reaction network with the same number of paths with the network of the 1000th generation. (b) The probability $q(x)$ is that a path to a chemical with abundance x is selected in evolution. The probabilities for incoming ($q_{in}(x)$), outgoing ($q_{out}(x)$), and catalyzing paths ($q_{cat}(x)$) are plotted. The data were obtained by 1.5×10^5 trials of randomly adding a reaction path to the network of the 200th generation, and the paths giving the top 0.05% growth rates were selected.

IV. MECHANISM: EMBEDDING THE POWER-LAW IN ABUNDANCE INTO NETWORK STRUCTURES

Why the scale-free-type connectivity distribution emerges in this evolution is explained by the selection of preferential attachment of paths to those chemicals in greater abundance. Note that the power-law distribution of chemical abundance has already been established through evolution. Here, we found that when a new reaction path is attached to an abundant chemical species, it gives a larger influence on the whole cellular state, as is to be expected from reaction kinetics. Consequently, a change in the rate of growth after the mutation of the network is also greater when a path is attached to an abundant chemical species, as shown in Fig. 4. Thus, when a certain number of cells with higher growth rates is selected from the mutant pool, the probability that those selected cells have new links to such abundant chemicals is statistically higher than those expected from random change without selection. Therefore, there is a positive correlation between the abundance of chemical species and the probability that new links will be added to such species in evolutionary dynamics: that is, preferential attachment appears to such abundant chemicals. To represent this probability, we use the variable $q(x)$, which indicates the probability that a new reaction path is attached to a chemical with abundance x after selection. For example, assume that a change in the rate of growth by the addition of a path outgoing from a chemical increases linearly with its abundance x. This assumption is natural as the degree of influence on the cellular state is generally proportional to the flux of the reaction path added to the network: that is, the product of substrate and

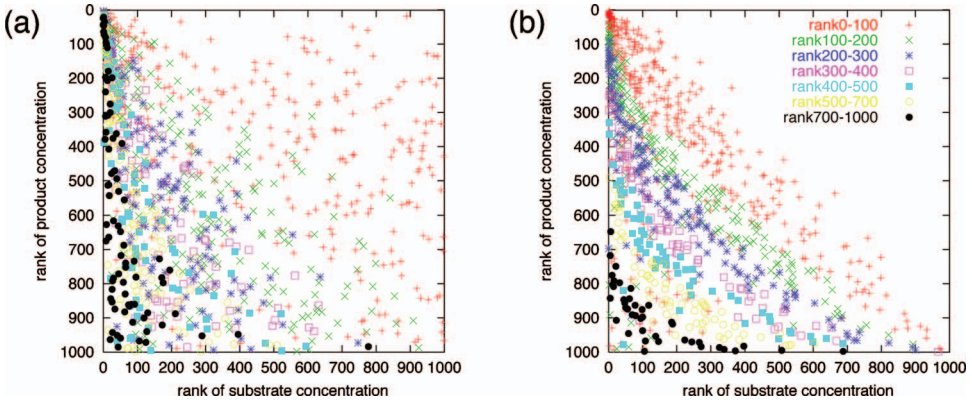

FIG. 3. (Color) Changes in the network structure. The abscissa shows the rank determined by the abundance of substrate i, and the ordinate shows the rank for the product j: the top left is the most abundant and the bottom right is the least abundant. A point is plotted when there is a reaction path $i \to j$, while the abundance of catalyst for the reactions is given by different colors determined by rank. As each product is dominantly synthesized from one of the possible paths, we plotted only the path with the highest flow. (a) The network structure at the tenth generation is rather random, even though the power-law in abundance has already been established. (b) The network at the 1000th generation. Only a small number of paths are located in the upper-right triangular portion of the figure, indicating that almost all of the chemical species were synthesized from more abundant species.

catalyst abundances. In this simple case, $q_{out}(x)$, which represents the probability of attachment for an outgoing path will increase linearly with x, even though the network change is random. Here, the connectivity distribution $P(k_{out})$ is obtained by the transformation of the variable as follows. Suppose that the probability of selection of a path attached to a chemical with abundance x is given by $q(x)$, then the path number $k \propto q(x)$. By the transformation $k=q(x)$, the distribution

$$P(k) = \frac{dx}{dk} p(x) = \frac{p(q^{-1}(k))}{q'(q^{-1}(k))} \tag{3}$$

is obtained. By applying the abundance power-law $p(x) \propto x^{-2}$, we obtain $P(k)=k^{-(\alpha+1)/\alpha}$ when $q(x)=x^{\alpha}$. Consequently, a scale-free network with an exponent of -2 should evolve $q_{out}(x) \propto x$.

Numerically, we found that the probabilities $q_{out}(x)$ and $q_{cat}(x)$ were fitted by $q(x) \propto x^{\alpha}$ with $\alpha \approx 1/2$, as shown in Fig. 2(b). Then, using the above transformation, the connectivity distribution was obtained as $P(k)=k^{-3}$. Here, it is interesting to note that the connectivity distribution observed from real metabolic and other biochemical networks follows the power-law $P(k) \propto k^{-\gamma}$ with γ being between 2 and 3, as is often seen in experimental data [2,3].

The probability $q(x)$ is determined through the evolutionary process. To clarify the reason for $q(x) \sim x^{\alpha}$ with $\alpha < 1$ in outgoing and catalyzing paths, we investigated the relationship between substrate abundance x and catalyst abundance y of a path to be selected. For this, we simulated changes in growth rates by the random addition of a reaction path to the network of the 200th generation. For 1.5×10^5 trials, paths giving 0.05% of the highest growth rates were selected and are plotted in Fig. 5 as green points on the x-y plane, while others are plotted as red points. As shown, a path with small flux is not selected because adding such a path cannot change the cellular state sufficiently, and a path with large flux is also not selected because such a large change destroys the hierarchical structure of catalytic reactions, which results in a decrease in nutrient intake or the critical point is exceeded so that the "cell" can no longer grow. Then, the fluxes of the selected paths satisfy $\Delta < xy < \Delta + \delta$, with Δ and δ being constants. We also found that the density of paths to be selected is almost constant in the above region. Consequently, for each chemical x, the probability that such a path exists is given by the probability that there is such a partner chemical with abundance y, which satisfies $\Delta/x < y < (\Delta + \delta)/x$. That is,

$$q(x) = \int_{\Delta/x}^{(\Delta+\delta)/x} p(z)dz \approx p(\Delta/x)(\delta/x). \tag{4}$$

By using Eq. (1), we obtain

$$P(k) = \frac{-p(\Delta/y)}{(p(y) + ydp(y)/dy))y^2}, \tag{5}$$

with $yp(y)=k$. Indeed, if $p(x)=x^{-2}$, the above expressions lead to $q(x) \propto x$, as well as $P(k)=k^{-2}$. This expression holds when the evolved network is just at the critical point. The evolved network is near this critical point but there is a slight deviation, as can be seen in the deviation from the power-law in Fig. 1, for a low abundance of chemicals. Note that the asymptotic behavior for a large k is also given for a small y. Then, the asymptotic behavior for a large k is given by $P(k) \approx 1/((p(y)+ydp(y)/dy))$, which depends on $p(y)$ for a small y. If the asymptotic behavior of $p(y)$ for a small y is given by $y^{-\beta}$ with $\beta < 2$, then $P(k) \approx k^{\beta/(1-\beta)}$. As $\beta < 2$, the exponent of the power is smaller than -2. For example, for $\beta = 3/2$ (which corresponds to the relationship between x and rank n as $x \sim n^{-2}$ for large n, as seen in Fig. 1), $P(k) \approx k^{-3}$ is obtained. In general, even if the behavior of $p(y)$ for a small y is not fitted by a power-law, its increase with $y \to 0$ is slower than y^{-2}. The decrease of $P(k)$ with k is then faster than k^{-2}, as is often seen in experimental data [2,3].

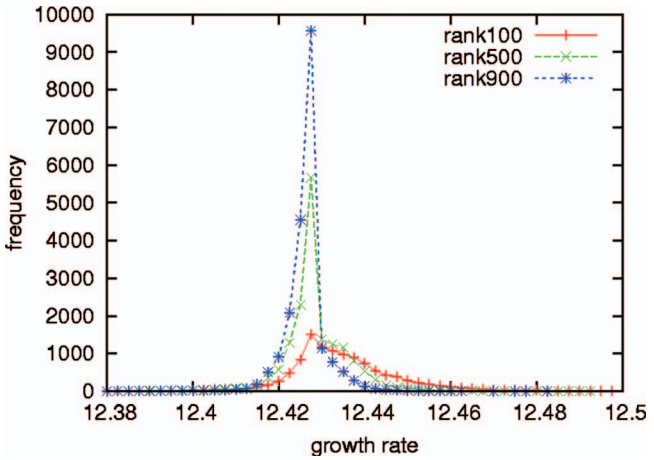

FIG. 4. (Color) Changes in growth rate with the addition of a reaction path. Reaction paths were added to the network of the 200th generation from the 100th, 500th, and 900th most abundant chemical species to investigate the changes in growth rate, whereas the product and the catalyst of the path were chosen randomly. Here, the concentrations of the 100th, 500th, and 900th most abundant chemicals were 1.80×10^{-3}, 2.03×10^{-4}, and 2.98×10^{-5}, respectively. The histograms show growth rates obtained by 20 000 trials. The growth rate is measured as the inverse of the time for a cell to divide, thus the unit of the x axis is (division/time). In some trials, the growth rates decreased markedly with the addition of a path, as the amount of nutrient uptake exceeded the limit of cellular dynamics. For the paths from the 100th, 500th, and 900th most abundant chemical species, 39%, 23%, and 4% of such trials showed growth rates of less than the given threshold (we choose 12.38), respectively. Such data are not plotted in the figure. As shown, adding a reaction path from a more abundant chemical was more effective in changing the growth rate of the cell.

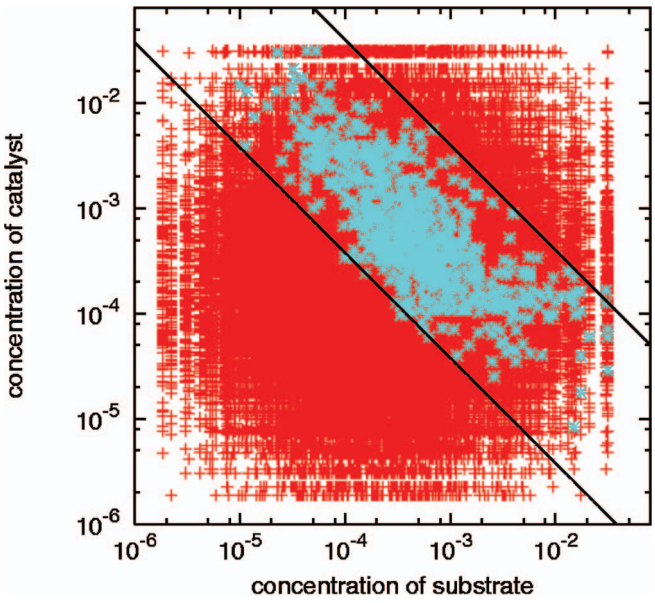

FIG. 5. (Color) Relationship between substrate abundance x and catalyst abundance y for the selected paths. A randomly chosen reaction path was added to the network of the 200th generation, and the growth rate of a cell after adding the path was simulated. For 1.5×10^{5} trials, paths giving 0.05% of the highest growth rates were regarded as being selected and are plotted as green points on the x-y plane, while others are plotted as red points. As shown, the selected paths satisfy $\Delta<xy<\Delta+\delta$, with $\Delta=3.8\times10^{-8}$ and $\delta=4.0\times10^{-6}$, respectively.

On the other hand, the probabilities $q_{in}(x)$ of having incoming paths after selection show no dependence on the chemical abundance x, and therefore, the above explanation is not directly applicable for such paths. As for incoming chemicals, we have found "hot" chemical species that facilitate the synthesis of the transporters for the nutrient uptakes, while others promote the formation of a cascade structure of reaction dynamics as shown in Fig. 3. These hot species are more likely to acquire an incoming path after selection. Such inhomogeneity of the probability among chemicals results in the inhomogeneity of the number of incoming paths as shown in Fig. 2(a). Still, further studies are necessary if such inhomogeneity results in the same power-law as $q_{out}(x)$ and $q_{cat}(x)$.

V. GENERALITY OF RESULTS

Through several simulations, we found that the emergence of two statistical features to be quite general and we expect that these do not rely on the details of our model. Specifically, we first checked the results by changing the initial conditions of the simulation, i.e., the initial concentrations of chemicals and the reaction network in the first cell, and confirmed that the results are independent of the initial conditions. Next, we studied a model by changing the parameters. Still, by restricting parameter values at which a cell reproduces efficiently, Zipf's law for abundances is generally observed. For example, we confirmed that the results are qualitatively unchanged if we change the diffusion constant D in the range $0.1\sim5.0$ and the initial path number k_{init} in the range $4\sim10$. Furthermore, we have found Zipf's law for the following class of models, based on a cell that reproduces efficiently.

(1) Generality against network structure: We studied the models with homogeneous connectivity as well as highly inhomogeneous connectivity. For example, the emergence of Zipf's law was also observed when the connectivity distribution of reaction network $P(k)$ obeys a power-law with the exponent ranging $2\sim3$.

(2) Generality against parameter distribution: In the present model, for simplicity all the reaction coefficients were chosen to be equal, and all the diffusion coefficients were also identical. Instead of identical parameter values for chemical reactions and diffusion coefficients, we studied the case with distributed parameters depending on each chemical species or each reaction. For example, even when the reaction coefficients are distributed in the range $1\sim100$, the reported results were obtained.

(3) Generality against reaction kinetics: Here, we studied the case with a higher order catalytic reaction [e.g., the reaction kinetics $x_j x_\ell^2$ instead of $x_j x_\ell$ in Eq. (2) for all chemicals] and the case with Michaelis-Menten reaction kinetics. Also, qualitatively the same results were obtained.

(4) Generality against the form of transport of nutrient chemicals: In the present model, the transport of chemicals is

mediated by some chemicals, represented by the term $Dx_{m_i}(X_i - x_i)$ in Eq. (2). Instead of this active transport, we studied the cases with a passive diffusion for the transport of nutrients, represented by the term $D(X_i - x_i)$. As a result, the emergence of Zipf's law was also observed.

(5) Generality against the condition for the cell division: In the present model, cell division occurs when the sum of all chemicals exceeds a given threshold. Instead, we studied the case that some chemicals are regarded as chemicals maintaining the cell membrane, and cell division occurs when the amount of those membrane chemicals exceeds a threshold. We confirmed that the reported results are independent of the change of cell division rules.

In any case, we found that a hierarchical structure of catalytic reactions is organized at the state with optimal growth. This hierarchical organization of catalytic reactions maintains the power-law distribution of chemical abundance, as discussed in Ref. [1]. Hence, we believe that the result is general when a reaction network system that synthesizes chemicals in a cell shows recursive growth.

Now, it is expected that Zipf's law generally emerges through evolution, for a cell system consisting of the following processes: (i) intracellular reaction dynamics within cells; (ii) intake of nutrients (that may depend on the internal chemical concentration); (iii) synthesis of chemicals through the above process leads to cell growth so that the cell divides when a certain condition is satisfied; and (iv) evolutionary processes together with this cell division, i.e., random mutations to reaction networks and selection of cells with higher growth rate. Higher growth in the cell is selected through (iv) and Zipf's law of abundance is generally reached for a cell with optimal growth. Furthermore, as the embedding mechanism is also general, the evolution to power-law in network paths is also expected to be general.

Indeed, we have performed simulations with several different evolutionary criteria, and the results are essentially the same, provided the degree of mutation is not large. For example, when we assume that the probability of being selected as parent cells for the next generation is proportional to cellular growth rate, the evolutionary dynamics are qualitatively the same as those presented here. As another example, we have performed simulations in which a fixed (large) number of cells is put in a given environment and, when a cell divides into two, a randomly chosen cell is removed to keep the total cell number constant, instead of introducing discrete generations as in the genetic algorithm rule adopted in the present paper. In such rules of simulation, cells having higher growth rates are also selected, and the power-law distribution of chemical abundance emerges as a result of evolutionary dynamics [13].

VI. SUMMARY AND DISCUSSION

In the present paper, we have shown that the power-law in abundances of chemicals and network paths naturally emerges through evolution, by taking a class of cell models consisting of catalytic reaction networks. We show that the power-law of abundance is later embedded into that of network path distribution, while the relation between the two powers is analyzed.

With regard to the evolution of reaction networks, preferential attachment to a more connected node has often been discussed [2,14]. In the previous models, preference of path attachment is simply defined as a function of the numbers of existing paths, and the origin of such a preference in evolutionary dynamics remains obscure. On the other hand, our study is different from those reports in two important respects. First, the dynamics of chemical abundance in the networks were introduced explicitly (described as node "strength" in Ref. [15]), whereas previous models generally considered only the topological structure of the network. Second, selection only by cellular growth rate results in such a preference, even though attachment itself is random. Here, we found that more abundant chemical species acquired more reaction links, as attachments of new links to such chemicals have both a greater influence on the cellular state and a higher probability of being selected. With these mechanisms, the power-law in abundance is naturally embedded in the intracellular reaction network structure through evolution, which is simply a process of selecting cells with faster growth rates.

One possible approach to show the existence of the evolutionary dynamics we present here is to investigate the positive correlation between intracellular abundance and the number of reaction links of each chemical species in a real cell. For example, analysis of intracellular metabolite concentrations might reveal both the power-law distribution of metabolite abundance and the positive correlation between the abundance and the number of reaction links of each metabolite. The positive correlation of abundance and number of reaction links, if observed, is consistent with our theory on the embedding of the power-law distribution of metabolite abundance into the power-law distribution of reaction links in metabolic networks, which is already known well.

Another way to show the existence of such evolutionary dynamics is experimental verification of the preferential attachment of paths to more abundant chemical species. Such experimental verification may be possible by comprehensive analysis of the phenotypic changes produced by the addition of metabolic pathways. For example, we can transplant metabolic pathways by genetic engineering, and analysis of phenotypic changes by such addition of metabolic pathways should allow us to extract the correlation between the degree of phenotypic changes and the abundance of metabolites to which pathways are attached. Because large phenotypic changes enhance the probability that the addition of new pathways to such chemical species is selected in evolutionary dynamics, the observation of positive correlation in such analysis suggests that the preferential attachment of links to more abundant chemical species can be maintained.

As discussed, the emergence of the power-law distribution of chemical abundance is expected to be a general feature of growing cells, as this feature seems to appear necessarily in any system having both intracellular reaction dynamics and intake of nutrients from an environment, when the cellular growth rate is maximized. Similarly, our simulations support the idea that evolutionary dynamics favoring the power-law distribution of reaction path numbers emerge when cells having higher growth rates are selected and mutations are randomly added to reaction networks. An important point here

is that the emergence of general features is independent of details of the system, provided the conditions required for such features are satisfied. The power-laws of both abundance and connectivity, which are often observed in intracellular reactions, can be simple consequences of our mechanism by Darwinian selection.

As the power-law in abundance and links in the network are general for a reproducing cell consisting of a reaction network, we trust that the theoretical origin of universality of our results is understood without assuming a specific choice of models, as, for example, has been established in renormalization group theory for critical phenomena.

We would like to thank T. Yomo and K. Sato for stimulating discussions. The work was partially supported by a Grant-in-Aid for Scientific Research from the Ministry of Education, Science, Sports, and Culture of Japan.

[1] C. Furusawa and K. Kaneko, Phys. Rev. Lett. **90**, 088102 (2003).
[2] H. Jeong *et al.*, Nature (London) **407**, 651 (2000).
[3] H. Jeong *et al.*, Nature (London) **411**, 41 (2001).
[4] S. Li *et al.*, Science **303**, 540 (2004).
[5] D. E. Featherstone *et al.*, BioEssays **24**, 267 (2002).
[6] N. Guelzim *et al.*, Nat. Genet. **31**, 60 (2002).
[7] H. R. Ueda *et al.*, Proc. Natl. Acad. Sci. U.S.A. **101**, 3765 (2003).
[8] V. A. Kuznetsov *et al.*, Genetics **161**, 1321 (2002).
[9] E. Almaas *et al.*, Nature (London) **427**, 839 (2004).
[10] K. Kaneko and T. Yomo, J. Theor. Biol. **199**, 243 (1999).
[11] C. Furusawa and K. Kaneko, Phys. Rev. Lett. **84**, 6130 (2000).
[12] The rank distribution, i.e., the abundances x plotted by rank n can be transformed to the density distribution $p(x)$, which is the probability that the abundance is between x and $x+dx$. Because $dx=dx/dn\times dn$, there are $|dx/dn|^{-1}$ chemical species between x and $x+dx$. Therefore, the abundance-rank relation is given by a power-law with exponent -1, $p(x)=|dx/dn|^{-1}\propto n^2\propto x^{-2}$. In this paper, we adopt the rank distribution rather than the density, simply because there is a relatively small amount of data (e.g., 1000 chemical species). By using such small amount of data, plotting the density distribution $p(x)$ makes the characteristics of power-laws obscure, since the bin size may become too large. Rather, by plotting the rank distribution, the power-law characteristics are detected more clearly.
[13] As for the number distribution of reaction links, the simulation has not yet reached the stage where it can show scale-free statistics in a network clearly (because such simulation requires a much longer time than the present method); however, still we found that the distribution of the numbers of such networks shows heterogeneity in the numbers of reaction links, with significant deviation from those of random networks.
[14] A-L. Barabási and R. Albert, Science **286**, 509 (1999).
[15] A. Barrat, M. Barthélemy, and A. Vespignani, Phys. Rev. Lett. **92**, 228701 (2004).