2005－BIO－1（3）
2005／7／25

遺伝子発現量ダイナミックスにおける自己相似性について

落合　友四郎[1], J.C. Nacher[1], 阿久津 達也

京都大学 化学研究所 バイオインフォマティクスセンター

〒 *611-0011* 京都府宇治市五ヶ庄

（概要）今回、ヒトと酵母の遺伝子発現量の時系列データを解析することにより、遺伝子発現のダイナミックスを記述する基礎方程式を導いた。（ここで、ダイナミックスとは遺伝子発現量が時間とともにどのように変化するかということである。）この方程式には興味深いことに、ある種の不変性（自己相似性、スケール不変性）が存在する。この自己相似性により、観測されている遺伝子発現量のスケール則（分散と平均の関係）を説明することに成功した。

Self-similarity symmetry in gene expression dynamics

T. Ochiai[1], J.C. Nacher[1], T. Akutsu

Bioinformatics Center, Institute for Chemical Research, Kyoto University,

Uji, 611-0011, Japan

Abstract

In this article, we analyze the gene expression dynamics (i.e., how the genes expression fluctuates in time) by using a new constructive approach. This approach is based on only two fundamental ingredients: symmetry and the Markov property of dynamics. First, by using experimental data of human and yeast gene expression time series, we found a symmetry in short-time transition probability from time t to time $t+1$. We call it self-similarity symmetry (i.e., surprisingly, the

[1]These authors contributed equally to this work.

Corresponding authors: ochiai@kuicr.kyoto-u.ac.jp, nacher@kuicr.kyoto-u.ac.jp

gene expression short-time fluctuations contain a repeating pattern of smaller and smaller parts that are like the whole, but different in size). Secondly, the Markov property of dynamics reflects that the short-time fluctuation governs the full-time behaviour of the system. Here, we succeed in reconstructing naturally the global behavior of the observed distribution of gene expression (i.e., scaling-law) and the local behaviour of the power-law tail of this distribution, by using only these two ingredients: symmetry and the Markov property of dynamics. This approach may represent a step forward toward an integrated image of the basic elements of the whole cell.

The final goal in Systems Biology is the understanding and modelling of the cell's system. In the cell, the expression level of genes plays a key role, since gene expression is a complex transcriptional process where *mRNA* molecules are translated into proteins, which control most of the cell functions. Recently, gene expression profiles for different types of cells of several organisms have been measured, and some experiments have also provided data about the fluctuation of expression level of thousands genes in time. Here, we propose a stochastic approach by extending our previous research [1, 2] to gain insight into gene expression time series data, in order to uncover fundamental principles that govern the gene dynamics.

Although many complex systems may be governed by non-stochastic processes, in the gene expression problem the random variation is reasonable, plays a relevant role in cellular process, and furthermore stochastic noise (e.g., *intrinsic and extrinsic*) have recently been measured and studied theoretically [3, 4, 5]. For example, the expression level of thousands genes is very low, which creates intrinsic uncertainties in the number of expressed genes in the cells. Furthermore, the number of molecules which are involved in signal transduction pathways fluctuates from 10^2 to 10^4. Therefore, the randomness connected with elementary molecular interactions and their amplification in the signaling cascade generates significant spatio-temporal noise. Therefore, the stochastic approach seems more plausible than the deterministic approach. Finally, we also note that the current experimental techniques also generate an additional source of fluctuation, which come from the ubiquitous instrumental noise (which may be around 30% or more) from chip to chip with the current GeneChips technologies.

Currently, DNA microarray/GeneChips (like the pendulum clock of Hyugens used by Newton to uncover the dynamical laws written in his famous *Principia Mathematica*) offers the ability to monitor in time changes in ex-

pression level of large subsets of genes from a variety of organisms on a scale unattainable by other methods. In particular, experiments done on time series of absolute value of gene expression level studying the yeast mitotic cell cycle and transcriptional regulation during human cell cycle [6] have provided a huge wealth of data to uncover general principles of gene expression dynamics.

In the article [7], by using these experimental data [6], we have found the following interesting phenomena: 1) "Mean- Reverting" and "Extreme Value- Jumps-More" mechanisms. These mechanisms are present in short-time fluctuation (i.e., how the gene expression level changes from time-step t to time-step $t+1$). Surprisingly, these two mechanisms are governed by a symmetry: Self-similarity. By inserting this information (i.e., two mechanisms and the self-similarity symmetry) into the general stochastic partial differential equation (SPDE) [8] which spontaneously emerges from Markov property, we found the fundamental equation for gene expression dynamics. This equation has a high predictive power. Here, we enumerate a couple of results. For example, the distribution solution of this equation $\rho^m(x)$ gives the distribution of expression level x of genes which fluctuate around m (i.e., their average expression is in the vicinity of the value m). This solution agrees with the experimental data. Furthermore, we rebuild the observed global behaviour of gene expression distribution, which is characterized by the following relation $\rho^{cm}(x) = \frac{1}{c}\rho^m(x/c)$ (or as the more symmetric form $\rho^{cm}(x)dx = \rho^m(x/c)d(x/c)$). We call it scaling-law formula. This formula indicates that distribution of gene expression level also has the self-similarity structure (We can observe a repeating pattern of smaller and smaller parts that are like the whole, but different in size, which is a direct consequence of the self-similarity symmetry in short-time transition probability. In a sense, the self-similarity governs all the dynamical structure of gene expression phenomena, by means of Markov property. On the other hand, our model can predict sensitive aspects of gene expression systems. More precisely, our constructive model predicts that the tail of the distribution has a power-law tail $\rho^m(x) \propto x^{-4}$ (when $x \to \infty$), which is observed in experimental data. This self-similar symmetry may help to understand other properties found in complex systems. For example, preliminary results shown in a companion work [9] indicate that the universal property of fluctuation (i.e., coupling between the average flux and dispersion follows a scaling-law with exponent one) reported in [10] can be re-built by using the self-similar symmetry.

References

[1] T. Ochiai, J.C. Nacher, T. Akutsu, Physics Letters A **330**, 313 (2004).

[2] T. Ochiai, J.C. Nacher, T. Akutsu, Physics Letters A, **339**, 1 (2005).

[3] M.B. Elowitz, A.J. Levine, E.D. Siggia, P.E. Swain, Science **297**, 1183 (2002).

[4] J. Paulsson, Nature, **427**, 415 (2004).

[5] W.J. Blake, M. Kaern, C.R. Cantor, J.J. Collins, Nature **422** 633 (2003).

[6] R.J. Cho et al., Nature Genetics **27**, 48 (2001).

[7] T. Ochiai, J.C. Nacher, T. Akutsu: e-print archive *http:arxiv.org/abs/q-bio.BM/0503003*, (2005).

[8] N.G.van Kampen, *Stochastic processes in physics and chemistry,* Elsevier Science B.V. (1992).

[9] J.C. Nacher, T. Ochiai, T. Akutsu, e-print archive *http:arxiv.org/abs/q-bio.BM/0503004*, (2005).

[10] M. Argollo de Menezes, A.-L. Barabasi, Phys. Rev. Lett., **92**, 028701 (2004).