

線虫の初期発生における遺伝子発現の クラスタリング解析

鈴木義実[†][†] 筑波大学

図書館情報専門学群

中山伸一[‡][‡] 筑波大学

図書館情報メディア研究科

伊藤将弘[§][§] 立命館大学

情報理工学部

真米城哲也[‡][‡] 筑波大学

図書館情報メディア研究科

1 はじめに

遺伝子は生命活動の基本要素であり、遺伝子に記述されている内容がどのように機能するかを知ることで、生命の仕組みを知ることができる。しかし、現時点では多くの遺伝子の機能は未解明である。さらに、遺伝子がどのように他の遺伝子と関係しているのかを知ることは重要であるが、遺伝子間の関係は遺伝子の機能以上に未解明の部分が多い。

本研究は、多細胞生物のモデル生物である線虫 *C.elegans* [1] の初期発生段階を対象として、(1) 線虫の遺伝子はそのような相互関係によって発生を実現しているのか、(2) 発生では、どのような分子レベルの現象が生じているのか、といった現象の解明を目的とする。

我々の研究グループでは超高速シミュレータを用いたシミュレーション実験によって遺伝子相互作用を解明する研究を進めている [2, 3]。遺伝子発現の時系列データに基づく遺伝子相互作用の推測方法はいくつか提案されているが [4, 5, 6, 7]、それらの推測精度は低く、現実的な条件ではたかだか 60% 程度である。これまでの実験で、我々の手法の予測精度は従来の手法よりも高いことが判っている。この手法は進化的計算により相互関係を推測するが、初期値として遺伝子のクラスタ構造を入力することができる。ここでは計測間隔時間の異なる発現の時系列データを用いたクラスタリングについて報告する。

2 方法

線虫 *C.elegans* の初期発生段階を対象として、2 種類の計測間隔でゲノムワイドな DNA マイクロアレイ実験を行った。1 つは 10 分間隔、もう 1 つは 30 分間隔で、どちらも計測回数は 5 回である。

線虫の培養および扱いは標準的な手法 [8, 9] を改良して行った。N2 Bristol 株の *C.elegans* を、餌として LB で 37°C で培養した大腸菌 OP50 の濃縮物を加えた S 培地で飼育した (20°C)。液体培地から産卵期の線虫をフィルターおよび遠心分離で回収し、2 段階のアルカリブリーチによって同期の精度を高めた。得られた初期胚を 15ml の遠心分離チューブに分液し、20°C で培養を行った。10 分間隔または 30 分間隔でチューブを取り出し、氷に浸して成長を止めた。全ての遠心分離チューブに 100ml の RNA 安定溶液を追加した。

これらの線虫の初期胚を *C.elegans* ゲノムアレイ (Affymetrix Part No. 900383 と 900384 GeneChip) を使って、[11] に記載のように発現量を計測した。このマイクロアレイには Dec.2000 ゲノム配列と、Sanger Centre および GenBank (release 121) から予測された遺伝子および EST から選択された 22,150 個の重複無しの遺伝子がある。Sanger Centre の 18,800 アノテーションがこのアレイに表されている。また、Sanger Centre のアノテーションとは重複しない、約 2,300 個の 3' EST および、ゲノムにアラインした 300 個の GenBank mRNA も表されている。

これらのマイクロアレイデータは計測間隔が異なる時系列データである。時系列で全ての計測点の発現量が低い (0.2 以下) の遺伝子を除外する。このようにフィルタリングされたデータを使って階層的にクラスタリング処理する。発現の時系列パターンの相関係数を遺伝子間の距離とし、クラスタリングの距離として群平均距離を用いる。これは、2 つのクラスターそれぞれに含まれる全てのペア間の距離の平均をクラスター間の距離とする手法である。

クラスタリングした 10 分間隔と 30 分間隔のデータを比較する際に、両方で一定の閾値以内の距離に属する遺伝子のクラスターを抽出し、クラスター単位で遺伝子の機能から見られるクラスターの特徴と、機能が不明の遺伝子について調べる。

¹Clustering Analysis of Initial Development Stage of *C.elegans*

²Yoshimi Suzuki, Univ. of Tsukuba

²Shin-ichi Nakayama, Univ. of Tsukuba

²Masahiro Ito, Ritsumeikan Univ.

²Tetsuya Maeshiro, Univ. of Tsukuba

3 結果および考察

最低発現量によるフィルタリングの結果、10分間隔のデータは遺伝子数 5,219 個、30分間隔のデータは遺伝子数 2,907 個であった。これらのデータをクラスタリングし、類似度 0.999 以上のクラスタを抽出すると、10分間隔ではクラスター数 41 個、遺伝子数 834 個、また 30分間隔ではクラスター数 36 個、遺伝子数 1,222 個となった。

これらのクラスター間で、10分間隔および 30分間隔の両方のクラスターに 2 つ以上の遺伝子が属する場合、1 つのグループを構築する。この処理の結果、33 個のグループが得られた。これらのグループを以下の 3 つに分類した。なお、ここで機能が判明している遺伝子とは、WormBase で機能が記載されている遺伝子を指す。(1) グループに属する遺伝子のうち、2 つ以上の遺伝子の機能が判明しており、かつ類似した機能である場合。この場合、グループに属するが機能が判明していない他の遺伝子も同様の機能を持つと推測される。(2) グループに属する遺伝子のうち、2 つ以上の遺伝子の機能が判明しているが、異なる機能であり、グループの機能が明確に定義できない場合。(3) グループに属する遺伝子のうち、1 つの遺伝子のみの機能が判明している場合。この場合、グループの機能はこの遺伝子の機能だと考えられるが、確率は高くないと考えられる。

グループの数はそれぞれグループ 1 : 5 個、グループ 2 : 12 個、グループ 3 : 16 個であった。グループ 1 の機能は、リボソーム合成、生殖器、酵素の合成、運動器管に関わっていた。ここから、5 個の遺伝子の機能が推測できる。また、グループ 3 に分類された 22 個の機能が判明していない遺伝子についても推測は可能である。

この結果は予備的な解析結果であるが、本手法の有効性を示していると考えられる。

謝辞

線虫の培養および胚サンプルの作成について伊藤氏に感謝します。

参考文献

- [1] The C.elegans Research Community, ed., *Worm-Book*, doi/10.1895/wormbook.1.7.1
- [2] T. Maeshiro, H. Hemmi, and K. Shimohara, "Ultra-fast genome wide simulation of biological signal transduction networks: Starpack", *Frontiers of Computational Science*, 243-246, Springer, 2007.
- [3] T. Maeshiro, S. Nakayama, H. Hemmi, K. Shimohara, "An evolutionary system for the prediction of gene regulatory networks in biological cells, *International Conference on Instrumentation, Control and Information Technology*, 2007.
- [4] V. Anne Smith, E. D. Jarvis and A. J. Hartemink, "Evaluating functional network inference using simulations of complex biological systems", *Bioinformatics*, S216-S224, 2002.
- [5] K. Basso, A. A. Margolin, G. Stolovitzky, U. Klein, R. Dalla-Favera and A. Califano, "Reverse engineering of regulatory networks in human B cells", *Nature Genetics*, 37, 382-390, 2005.
- [6] A. J. Hartemink, "Reverse engineering gene regulatory networks", *Nature Biotechnology*, 23, 554-555, 2005.
- [7] R. Kelly and T. Ideker, "Systematic interpretation of genetic interactions using protein networks", *Nature Biotechnology*, 23, 561-566, 2005.
- [8] J. Sulston and J. Hodgkin, *The Nematode Caenorhabditis elegans*. Cold Spring Harbor Laboratory Press, 587-606, 1998.
- [9] H. Tabara, T. Motohashi, and Y. Kohara, "A multi-well version of in situ hybridization on whole mount embryos of *Caenorhabditis elegans*", *Nucleic Acids Res.*, 24:2119-2124, 1996.
- [10] S. Brenner, "The genetics of *Caenorhabditis elegans*", *Genetics*, 77:71-94, 1974.
- [11] L.R. Baugh, A.A. Hill, D.K. Slonim, E.I. Brown, C.P. Hunter, "Composition and dynamics of the *Caenorhabditis elegans* early embryonic transcriptome", *Development*, 130:889-900, 2003.