

5N-8 発現プロファイルの類似性を用いた DNA チップデータの解析

渡辺恒彦 野崎康行 田村卓郎 中重亮
日立ソフトウェアエンジニアリング(株)

1. 大量に生成される DNA チップデータ

生体の組織や細胞が活動するとき、その活動に対応するいくつかの遺伝子が働くことで細胞の機能を実現する。このことを生物学では遺伝子が発現するという。

DNA チップは、生体組織での遺伝子の発現状態を定量的又は定性的に検出することができる。

しかし、DNA チップから得られる遺伝子発現データは 1 実験が数メガから数十メガバイトオーダーの大量データであるため、ここから生物学的に意味のある情報を効率よく抽出するためにはデータマイニングが必須である。

2. 遺伝子の発現プロファイル

DNA チップデータのデータマイニングでは、遺伝子の発現プロファイルに着目している。発現プロファイルとは、遺伝子の発現強度を縦軸にとり、その遺伝子が発現している場所、もしくは時間を横軸にとり、図1に示すようなグラフにしたものをいう。図1のグラフには、酵母のリボゾームに関する遺伝子の発現プロファイルを示した。

DNA チップを用いた遺伝子解析の一つの目的として、機能未知の遺伝子に対する機能推測の支援がある。遺伝子研究では、機能が類似した遺伝子は発現強度も類似するという生物学的事実に基づき、発現プロファイルの類似性をもとに各遺伝子をグループ分けする。その後、各グループに対して生物学的情報をもとに機能を推測していく。このとき用いるグループ分けの手法として、一般的には多変量解析のクラスタ分析を用いている。

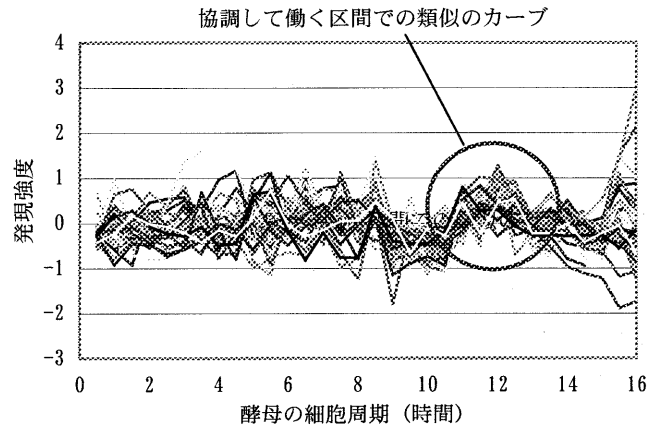


図1 協調して働く遺伝子

3. バイオ研究者の要求事項

バイオ研究者は機能が未知の遺伝子に関して実験を行いその働きを解明したいと考える。このとき実験に先立ち、DNA チップを用いて網羅的に複数遺伝子の発現を調査し、そのデータを解析することで実験対象となるような候補遺伝子の絞込みを行う。本研究では、その絞込みを実現するため、研究者の要求する次の点について検討を行った。

- (1)ある区間で協調して働く遺伝子を選別する。
- (2)遺伝子間相互作用を持つものを選別する。
(例えば、抑制関係にある遺伝子、発現量が一定量違う遺伝子、増幅する働きを持つ遺伝子等)

4. 協調して働く遺伝子の選別

図1は、ある区間で協調して働く遺伝子の発現プロファイルの実データによる例と考えることができる。これは、Web上で公開されているStanford大のDNAチップデータのうちで、酵母菌の細胞内のリボゾームに関係する遺伝子を集めてグラフにしたものである。クラスタ分析では「遺伝子発現プロファイルが類似したものは、遺伝子の機能が類似する」との仮定に基づきグループ分けを行っていたが、図1のように類似したグラフの形(以後カーブと呼ぶ)を示すのは、発

現プロフィール全体の中でも一部分であることが分かる。

従来のクラスタ分析では、協調して働く遺伝子の抽出や、相互関係のある遺伝子を大量の DNA チップデータから抽出できない。これらの問題点に対し、以下に示すような手法を検討した。

5. 部分選択による類似プロフィールの抽出

図 2 に示すように、研究者は興味の対象とする遺伝子を決め、この遺伝子の発現プロフィールの特異的なカーブを取り出し検索部分とする。これをチップデータの発現プロフィールと重ね合わせることで、①に示すように類似した発現プロフィールを持つ遺伝子を抽出する。これより協調して働く遺伝子を抽出することができる。

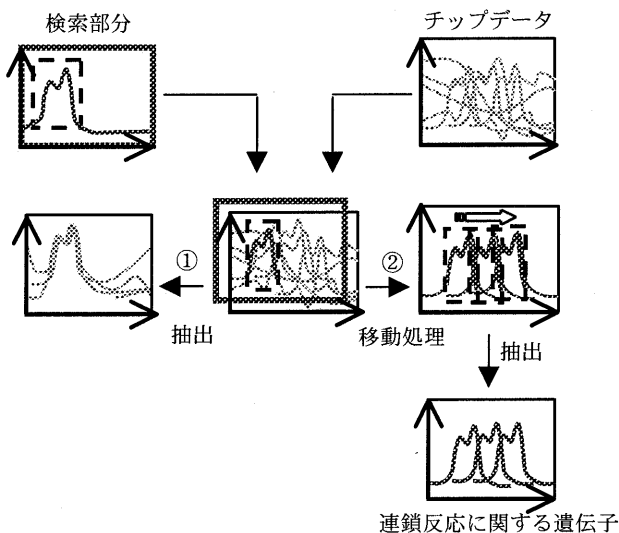


図2 類似プロフィール抽出（移動処理）

6. 移動・加工処理による類似プロフィールの抽出

遺伝子連鎖反応を例にとると、図 2 にあるように研究者の興味の対象となる遺伝子のうちで、特異的なカーブを取り出し検索部分とし、②に示すように検索部分全体を横軸方向へ移動することで、類似した発現プロフィールを持つ連鎖反応に関係した遺伝子を抽出する。これから相互作用のある遺伝子を抽出することができる。

遺伝子間相互作用には他にもさまざまあり、これらに対しては、研究者の目的に応じて加工処理を行う

と、図3に示すように抑制関係にある遺伝子、発現量が一定量違う遺伝子、増幅する働きを持つ遺伝子などを抽出できる。

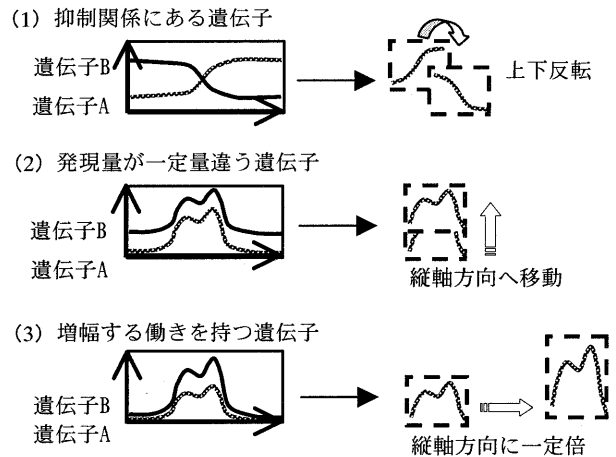


図3 類似プロフィール抽出（加工処理）

7. まとめ

本方式では発現プロフィールの特異的なカーブに注目し、協調して働く遺伝子を抽出することが可能となる。また、発現プロフィールの特異的なカーブに注目し、そのカーブに移動・加工処理を施すことで、相互作用を持つ遺伝子のみを抽出する見通しを得た。

今後は、実際の DNA チップデータに適用して、協調して働く遺伝子や相互作用のある遺伝子の抽出が可能であることを実証する。また、同一機能を持つ既知の遺伝子群の発現プロフィールを利用し、それらの最も類似性の高い区間を計算することで、その遺伝子群に特異的な検索対象とすべきカーブの候補を自動的に選び出す手法を検討する。

参考文献

- 1) 松原, 中村: ゲノムを読む: 紀伊国屋書店 (1996)
- 2) Eisen, M. et al.: Cluster analysis and display of genome-wide expression patterns: Proc. Natl. Acad. Sci. USA, 95, 25, pp.14863 ~ 14868 (Dec.1998)
- 3) Saccharomyces Genome Database : <http://genome-www.stanford.edu/Saccharomyces/>