

## 科学研究費助成事業（科学研究費補助金）研究成果報告書

平成25年 6月10日現在

機関番号：14401

研究種目：基盤研究（B）

研究期間：2010～2012

課題番号：22310125

研究課題名（和文）細胞分化過程の解明のための遺伝子ネットワーク解析技術の開発

研究課題名（英文）Development of Gene Regulatory Network Analysis Methods for Revealing Cell Differentiation Processes

研究代表者

松田 秀雄（MATSUDA HIDEO）

大阪大学・情報科学研究科・教授

研究者番号：50183950

研究成果の概要（和文）：複数の細胞分化系列にまたがる遺伝子制御の関係を表現する数理モデルを、遺伝子制御のネットワークに基づいて作成した。この数理モデルをもとに、実際にマウスの脂肪細胞・骨芽細胞の分化にまたがる制御遺伝子を探索できた。この探索は、複数の細胞分化過程を考慮するため膨大な時間がかかるため、マルチ CPU・マルチコアの計算サーバを多数利用する並列実行方式を開発し、処理時間の大幅な短縮を達成した。

研究成果の概要（英文）：We constructed a mathematical model for expressing gene regulatory relationships, based on gene regulatory networks, on multiple cell differentiations. We successfully explored several regulatory genes appeared in both of the two cell differentiations of mouse adipocytes and osteoblasts. We developed a parallel processing method on massively parallel servers to reduce a large amount of computational costs of the processing.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
2010年度	6,400,000	1,920,000	8,320,000
2011年度	5,500,000	1,650,000	7,150,000
2012年度	2,500,000	750,000	3,250,000
年度			
年度			
総計	14,400,000	4,320,000	18,720,000

研究分野：システムゲノム科学

科研費の分科・細目：ゲノム科学・システムゲノム科学

キーワード：遺伝子ネットワーク、細胞分化、ベイジアンネットワーク

### 1. 研究開始当初の背景

細胞分化とは、同一の幹細胞が異なる分化誘導を受けることで、構造的・機能的に変化して別の細胞になることであるが、その過程で転写因子による遺伝子ネットワークが関与していることが知られている。例えば、同一の間葉系幹細胞に対して分化誘導を行うことで、脂肪細胞や骨芽細胞などに分化させることができるが、脂肪細胞分化では PPAR $\gamma$ 、骨芽細胞分化では Runx2 と Osterix (SP7) が

重要な働きをする。さらに、いったん一方の細胞分化が誘導されてある時点を過ぎると、他方の細胞分化への誘導ができなくなることから、細胞分化の過程で、一方の分化のネットワークを促進し、他方を抑制するようなクロストーク遺伝子が存在し、細胞分化過程での分化能の消失に、遺伝子ネットワークの構造変化が関与している可能性が示唆されていた。このため、細胞分化過程での遺伝子ネットワークの構造変化を推定し、それによ

り体細胞のリプログラミング機構を解明し、クロストーク遺伝子を探索する方法論が求められていた。

## 2. 研究の目的

本研究は、細胞分化の過程での遺伝子ネットワークの構造変化とそれを制御する遺伝子を検出することで、細胞分化過程での分化能の消失が生じる時期を推定し、細胞分化の機構、さらには体細胞のリプログラミングを説明できる数理モデルを作成することを研究目的としている。従来の遺伝子ネットワーク解析手法は、ネットワークの構造変化を想定していないため、新たな情報解析技術の開発が必要となる。

細胞分化とは、同一の幹細胞が異なる分化誘導を受けることで構造的・機能的に変化して別の細胞になることであるが、その過程で転写因子による遺伝子ネットワークが関与していることが知られている。さらに、いったん一方の細胞分化が誘導されてある時点を超えると、他方の細胞分化への誘導ができなくなることから、細胞分化の過程で、一方の分化のネットワークを促進し、他方を抑制するようなクロストーク遺伝子が存在し、細胞分化過程での分化能の消失に、遺伝子ネットワークの構造変化が関与している可能性が示唆されていた。

そこで、本研究では次の2点を研究目的として推進した。

- (1) 複数の細胞分化過程における遺伝子ネットワークの間のクロストーク遺伝子を表現し、推定するための数理モデルを作成する。
- (2) 推定精度を保ったまま、遺伝子ネットワークの構造変化が推定可能な、新規の遺伝子ネットワーク解析技術を開発する。
- (3) 細胞分化過程に関与する多数の制御因子による遺伝子ネットワークの推定を効率的に行うための並列処理方式を開発する。

## 3. 研究の方法

- (1) クロストーク遺伝子の数理モデルの構築

細胞分化過程における遺伝子制御関係を表現する数理モデルを構築した。同じ幹細胞から別々の細胞へと分化していくとき、ある時点を過ぎると別の細胞への分化ができなくなることから、一方の細胞分化を制御する遺伝子ネットワーク中に、他方の細胞分化を制御する遺伝子ネットワークを抑制する遺伝子(クロストーク遺伝子)が存在するという図1のような状況を考え、これに基づいてクロストーク遺伝子を表現する数理モデルを立てた。具体的には、遺伝子間には時間遅延を考慮した相関係数で表される制御関係が存在すると仮定し、1から相関係数の絶対値を引いた値を制御辺の重みとして、重みの和

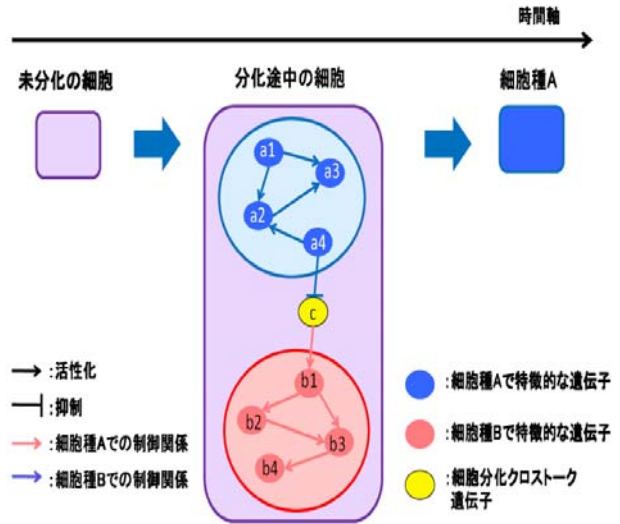


図1 クロストーク遺伝子の概念図

が最小となる、すなわち、相関係数で表される制御関係が最も強い(相関係数が正の時は活性化の制御、負の時は抑制の制御)と考えて、複数のネットワークにまたがる最短経路問題として定式化した。

(2) 遺伝子ネットワーク解析技術の開発  
ベイジアンネットワークによる遺伝子ネットワーク推定では、最適解を探索しようとすると遺伝子数のべき乗に比例する計算時間がかかるため、グリーディ法などの近似的な探索手法が使われてきた。しかし、グリーディ法は局所探索であるため、ネットワークの推定精度が十分でない可能性があった。そこで、ベイジアンネットワークによる新たな遺伝子ネットワーク推定手法を開発した。図2に手法の概念図を示す。

まず、与えられた遺伝子セットから、任意の3遺伝子を選ぶ組合せをすべて行い、得られた3つ組の遺伝子で、遺伝子ネットワーク推定を行う。

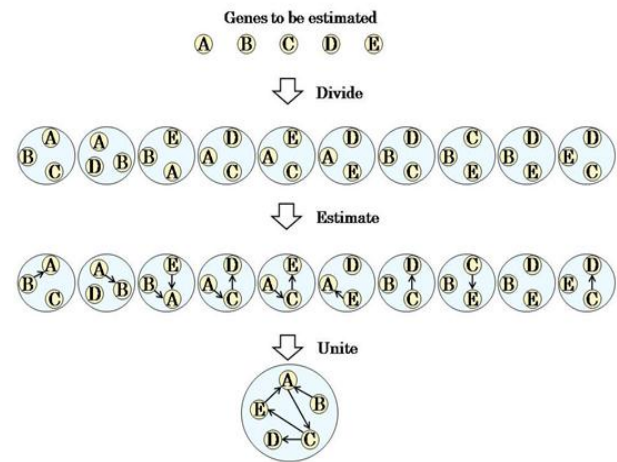


図2 3つ組による遺伝子ネットワークの構成

3つ組遺伝子のネットワーク推定では局所探索は行わずに全探索を実施する。次に、得られた3つ組遺伝子ネットワークを統合して、全体の遺伝子ネットワークを構成する。グリーディ法とは異なり、最初に3つ組を選ぶとき、それぞれの遺伝子ネットワーク推定の部分では局所探索を行っていないので、推定精度の低下を防ぐことができると考えられる。

### (3) 遺伝子ネットワーク推定のための並列処理方式の開発

本研究での遺伝子ネットワーク推定は、複数の細胞分化過程を考慮して行うため、逐次的に実行すると膨大な時間がかかる。そこで、マルチCPU・マルチコアの計算サーバを多数利用する並列実行方式を開発し、処理時間の大幅な短縮を図った。図3に手法の概要を示す。

3階層の並列処理を実現しており、まずCPU内で複数のネットワークの推定処理をマルチコアによるスレッド並列で並列化する。次にそれらをマルチCPUでMPIによりプロセス並列により並列化する。さらにプロセス並列での並列処理を、同時に複数実行し、それらの間の動的負荷分散を行う。この方式により、多数のCPUおよびコアがある並列処理サーバ上で、効率的に遺伝子ネットワーク推定を行うことができる。

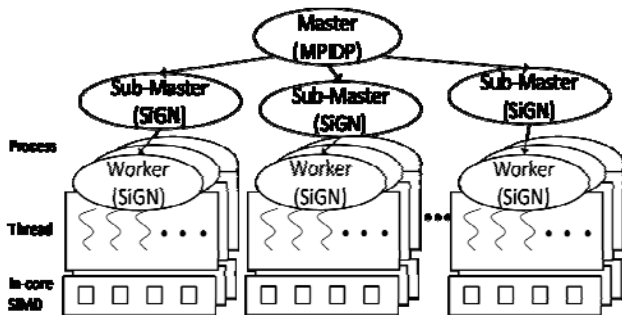


図3 遺伝子ネットワーク推定の並列処理方式

## 4. 研究成果

### (1) 数理モデルによるクロストーク遺伝子の探索

3. (1) のクロストーク遺伝子の数理モデルを使い、最短経路問題を解くアルゴリズムとしては、最短経路1個だけでなく、上位K個までの最短経路を求められる Yen のアルゴリズムを実装した。

実際に、マウスの脂肪細胞、骨芽細胞の細胞分化過程での、各遺伝子発現プロファイルを Affymetrix GeneChip Mouse Genome 430 2.0 Array で計測した。

骨芽細胞分化で代表的な制御因子である

Runx2 や Osterix (Sp7) を siRNA によるノックダウンさせたときに、発現が大きく変動した脂肪細胞分化関連の遺伝子をあらかじめ調べておき、前述のクロストーク遺伝子の数理モデルを元に推定した結果とどれくらい一致するかを調べた。その結果、モデルから推定された発現変動遺伝子 83 個のうち 16 個が、siRNA の実験で発現の変動が見られ、このモデルの有効性が示された。

### (2) 遺伝子ネットワーク解析技術の開発

マウスの脂肪細胞分化と骨芽細胞分化での遺伝子ネットワークを、3つ組による推定手法で推定を行った。推定結果を図4に示す。図4で(a)は既知の遺伝子ネットワークであり、グリーディ法と3つ組法の推定結果をこれと比較して評価している。グリーディ法では、探索範囲を指定するパラメータがあるため、デフォルトの推定結果を図4(b)に、パラメータを変えて探索範囲を拡大し推定精度をあげたときの結果を図4(c)に示す。探索範囲を増やすと、デフォルトのときよりも計算時間は大きく増大する。図4(d)が3つ組法による推定結果である。図4で、(b)、(c)、(d)の青い辺は、それで結ばれた遺伝子間に既知の制御関係はあるが、制御の向きが逆になっていることを示し、赤い辺は制御関係も向きも既知のものとは一致していること、黒い辺は既知のものとは制御関係が一致していないものを示す。これからわかるように、3つ組法は、探索範囲を広げたグリーディ法と比べてもより精度が高い。3つ組法の計算時間は、デフォルトのグリーディ法とほぼ同じくらいであることから、3つ組法は精度が高く、計算時間も大きくはない優れた探索手法であることが示された。

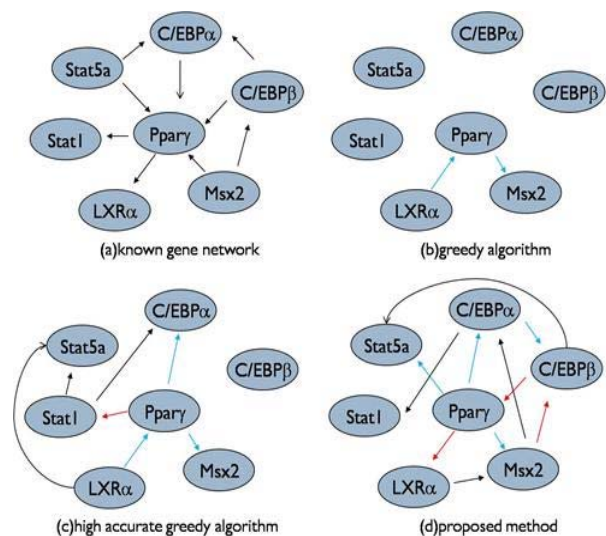


図4 グリーディ法と比較した3つ組法でのネットワーク推定結果

(3) 遺伝子ネットワーク推定のための並列処理方式の開発

3. (3) の並列処理方式を、東京大学情報基盤センターのデータ解析拠点クラスタ上で実装した。このクラスタの720コアの計算機資源を使って、マウス脂肪細胞分化過程における60時点の時系列遺伝子発現プロファイルで、26857個の遺伝子による大規模ネットワークを推定した結果、360コアで43069秒かかっていたものが、720コアで23044秒と短縮され、速度向上比1.87、並列化効率は93.4%と良好な加速を達成できた。

#### 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計7件)

①Ryo Araki, Shigeto Seno, Yoichi Takenaka, Hideo Matsuda, An estimation method for a cellular-state-specific gene regulatory network along tree-structured gene expression profiles, *Gene*, 査読有, Vol.518 (2013) 17-25. doi: 10.1016/j.gene.2012.11.090

②Tomoshige Ohno, Shigeto Seno, Yoichi Takenaka, Hideo Matsuda, A method for isoform prediction from RNA-Seq data by iterative mapping, *IPSJ Transactions on Bioinformatics*, 査読有, Vol.5 (2012) 27-33. doi: 10.2197/ipsjtbio.5.27

③Yukito Watanabe, Shigeto Seno, Yoichi Takenaka, Hideo Matsuda, An estimation method for inference of gene regulatory network using Bayesian network with uniting of partial problems, *BMC Genomics*, 査読有, Vol.13 (2012) S12. doi:10.1186/1471-2164-13-S1-S12

④Tomoyoshi Nakayama, Shigeto Seno, Yoichi Takenaka, Hideo Matsuda, Inference of S-system models of gene regulatory networks using immune algorithm, *Journal of Bioinformatics and Computational Biology*, 査読有, Vol.9 (2011) 75-86. doi: 10.1142/S0219720011005768

⑤吉澤陽志, 瀬尾茂人, 竹中要一, 松田秀雄, 細胞分化クロストークのモデル化と細胞分化クロストーク遺伝子の推定手法, *情報処理学会論文誌 数理モデル化と応用*, 査読有, 4巻 (2011) 59-68.

[学会発表] (計14件)

①Tomoyoshi Nakayama, Yoshiyuki Kido,

Hiromi Daiyasu, Shigeto Seno, Yoichi Takenaka, Hideo Matsuda, Estimate Dynamic Gene Regulatory Networks in Adipocyte Differentiation for Detecting Changes of Gene Regulations by Splitting Time Course Data, 23rd International Conference on Genome Informatics (GIW2012), 2012.12.13, National Cheng Kung University, Tainan, Taiwan.

②Hideo Matsuda, Large-Scale Gene Regulatory Network Analysis for Adipocyte Differentiation on High-Performance Computer, The 23rd International Conference on Genome Informatics (GIW2012), 2012.12.12, National Cheng Kung University, Tainan, Taiwan (基調講演).

③Ryo Araki, Yoichi Takenaka, Tomoshige Ohno, Yukito Watanabe, Shigeto Seno, Hideo Matsuda, Direct unique gene regulatory networks along dendrogram of cell differentiation, InCoB/ISCB-Asia Joint Conference 2011, 2011.11.30, Renaissance Hotel, Kuala Lumpur, Malaysia.

④Tatsuya Kitaguchi, Shigeto Seno, Yoichi Takenaka, Hideo Matsuda, Network analysis for time-series expression profile using nested effects models, The 2011 Joint Conference of CBI and JSBi (CBI/JSBi2011), 2011.11.8, Kobe International Conference Center, Hyogo.

⑤Tomoshige Ohno, Kiyoshi Yoshizawa, Shigeto Seno, Yoichi Takenaka, Hideo Matsuda, A Directed Graphical Gaussian Model for Inferring Gene Regulatory Networks, The 19th Annual International Conference on Intelligent Systems for Molecular Biology (ISMB2011), 2011.7.19, Austria Center Vienna, Vienna, Austria.

⑥Yukito Watanabe, Shigeto Seno, Yoichi Takenaka, Hideo Matsuda, A Method for Inference of Gene Regulatory Network based on Bayesian Network with Uniting of Partial Problems, The 19th Annual International Conference on Intelligent Systems for Molecular Biology (ISMB2011), 2011.7.18, Austria Center Vienna, Vienna, Austria.

⑦Kiyoshi Yoshizawa, Shigeto Seno, Yoichi Takenaka, Hideo Matsuda, Revealing

regulatory relationships of crosstalk with multiple time-series gene expression profiles, The 2010 Annual Conference of the Japanese Society for Bioinformatics (JSBi2010), 2010.12.13, Centennial Hall, Kyushu University School of Medicine, Fukuoka.

⑧ Hideo Matsuda, Reconstruction of Gene Regulatory Networks for Drug Target Discovery, The 3rd National Seminar and Workshop on Computer Aided-Drug Design: Virtual Screening (CADD2010), 2010.12.3, Gurney Hotel, Penang, Malaysia (基調講演).

## 6. 研究組織

### (1) 研究代表者

松田 秀雄 (MATSUDA HIDEO)  
大阪大学・情報科学研究科・教授  
研究者番号 : 50183950

### (2) 研究分担者

竹中 要一 (TAKENAKA YOICHI)  
大阪大学・情報科学研究科・准教授  
研究者番号 : 00324830