

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24年 6月 8日現在

機関番号：82626

研究種目：基盤研究（B）

研究期間：2009～2011

課題番号：21360409

研究課題名（和文）細胞分化の人工誘導経路を短時間で最適化する方法の研究

研究課題名（英文）Study on Rapid Optimizing Cell Inductive Differentiation Routes

研究代表者

藤渕 航（FUJIBUCHI WATARU）

独立行政法人産業技術総合研究所・生命情報工学研究センター・研究チーム長

研究者番号：60273512

研究成果の概要（和文）：間葉系幹細胞とそこから分化する様々な細胞の遺伝子発現データベースを利用して、細胞分化誘導候補遺伝子を選択する情報学的手法を2種開発した。1つは実際に上位から7つの候補因子のうち2つが脂肪細胞分化を促進する能力を持つ事を示した。もう1つの手法は、ヒト全遺伝子約20,000個のうち数100程度の少ない既知データから細胞分化能を持つ因子を高精度で予測することができた。さらに複数因子を組み合わせて誘導する手法についてタンパク質を用いる実験法を確立した。

研究成果の概要（英文）：Using a gene expression database of mesenchymal stem cells and various cells differentiated from them, we have invented two informatics methods to select candidates of cell differentiation induction factors. In one method we have shown that there are two out of top five candidates actually promote the adipocyte differentiation. In another method we have shown that we use only a few hundreds known data out of about 20,000 human genes and predict cell differentiation factors with high accuracy. Furthermore, we have established an experimental method to use proteins as combinatorial induction system of multiple factors.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
2009年度	4,400,000	1,320,000	5,720,000
2010年度	4,100,000	1,230,000	5,330,000
2011年度	4,200,000	1,260,000	5,460,000
総計	12,700,000	3,810,000	16,510,000

研究分野：工学

科研費の分科・細目：プロセス工学、生物機能・バイオプロセス

キーワード：発生分化、データベース、バイオテクノロジー、生体生命情報学

1. 研究開始当初の背景

幹細胞技術が目覚ましい発展をとげており、山中らによって体細胞に4つの因子を導入してiPS細胞を作成する手法が発表されていたが、逆に分化制御技術の発展が遅れており、制御遺伝子を大量の候補から絞り込むシステムティックな研究が立ち後れていた。

2. 研究の目的

山中らが行った（脱）分化制御遺伝子発現過程をスケールアップして網羅的また確実に行えるプロトコルを確立し、安全で最適な誘導法を効率的に見つけることで、今後の幹細胞技術の再生医療への実用化に格段の加速をすることが狙いである。

3. 研究の方法

まず、細胞分化DBを開発し、間葉系幹細胞から脂肪細胞を誘導する転写因子候補を選択する。siRNAにより、分化に重要な10転写因子を絞り込む。さらに組み合わせ効果を網羅的に調べ、最適な誘導遺伝子セットを確定する。上記の実験結果から脂肪分化に関わる遺伝子ネットワークを推定し公開する。

4. 研究成果

(1)細胞分化DBの開発

本研究で必要とされる細胞データベースを開発し「ヒト細胞・細胞分化データベースCELLPEDIA」としてインターネット上で公開し、国際雑誌に論文として発表した。本データベース中にはDigital Differential Display(DDD)を行うための基礎となる情報である様々な体細胞の公共遺伝子発現データ878件が搭載されている。また、幹細胞の公共遺伝子発現データは92件の計970件が搭載されている(図1)。

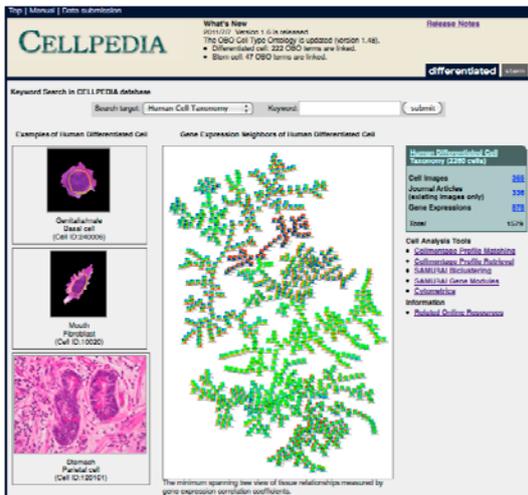


図1 ヒト細胞・細胞分化データベースCELLPEDIA

(2)分化誘導因子候補の統計解析

上記CELLPEDIAから、本プロジェクトで対象とする間葉系幹細胞とそこから分化する各種細胞(脂肪、軟骨、骨芽)の遺伝子発現データをDDD比較して細胞分化制御遺伝子候補を選択する統計手法を考案した。統計手法には、Fisher Omnibus検定と呼ばれる手法を応用し、脂肪分化に特異的に制御していると見込まれる遺伝子の候補のみ67個を選択した(図2)。図2には、DDDの概念図とそれぞれの細胞種とその遺伝子発現のヒートマップを示した。

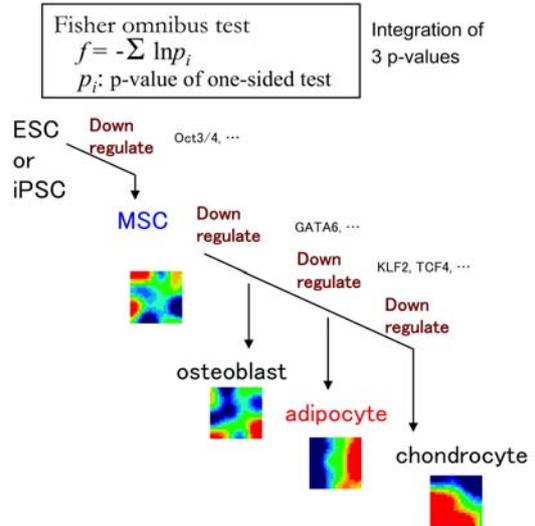


図2 Fisher Omnibus 検定によるDDD解析

(3)機械学習による細胞分化因子予測

siRNA実験はコストがかかるため、少ない細胞分化実験から機械学習により、新規細胞分化制御因子を予測することが必要とされている。本研究では転写因子をランダムに実験した結果から、サポートベクター回帰法を用いて、分化制御効果の高い遺伝子を予測する手法を開発した。

サポートベクター回帰では、データを \mathbf{x} とする $f(\mathbf{x}) = \mathbf{w} \cdot \mathbf{x} + b$ を、以下の様に解く事である。

$$\min \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^l (\xi_i^+ + \xi_i^-),$$

$$\text{s.t. } (\mathbf{w} \cdot \mathbf{x}_i + b) - y_i \leq \varepsilon + \xi_i^+,$$

$$y_i - (\mathbf{w} \cdot \mathbf{x}_i + b) \leq \varepsilon + \xi_i^-,$$

$$\xi_i^+, \xi_i^- \geq 0, i=1, 2, \dots, l$$

ξ_i^+ 、 ξ_i^- は ε -insensitive lossとよばれるものであり、 ε との関係を図示すれば図3の様になる。

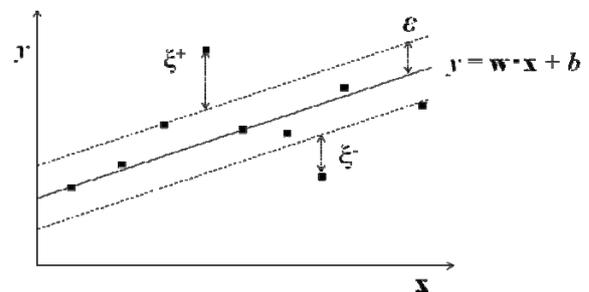


図3 SVRにおける ε -insensitive loss

C がモデルの複雑さを調整するregularization parameterになる。SVRは、SVMと同様、カーネルを使うことによって非線形に拡張することが可能である。本報告で

は、linear カーネルに加えて、RBF カーネル、Maximum entropy (ME) カーネルを用いた。サンプル間の距離行列が与えられたとき、それをカーネルに対する制約条件としながらカーネル行列 K のエントロピー $-\text{tr}(K \log K)$ の最大化を行うことで、最適なカーネル行列を得るものである。

(4)細胞分化における実験検証

①Fisher Omnibus DDD による検証

開発した Fisher Omnibus 検定を基盤とする DDD によって選択された 67 遺伝子候補のうち、トップ 7 遺伝子について脂肪分化誘導能を siRNA 実験による遺伝子サイレンシングで検証した。

市販の脂肪分化キットに含まれる 4 試薬のうち、インドメタシンとデキサメタゾンの 2 試薬のみ残した試薬ではわずかにしか脂肪分化が起こらなくなる。しかし、これに候補遺伝子 7 つを導入すると脂肪分化が起こり、我々の選択した因子が細胞分化能を持つ事が示された (図 4)。

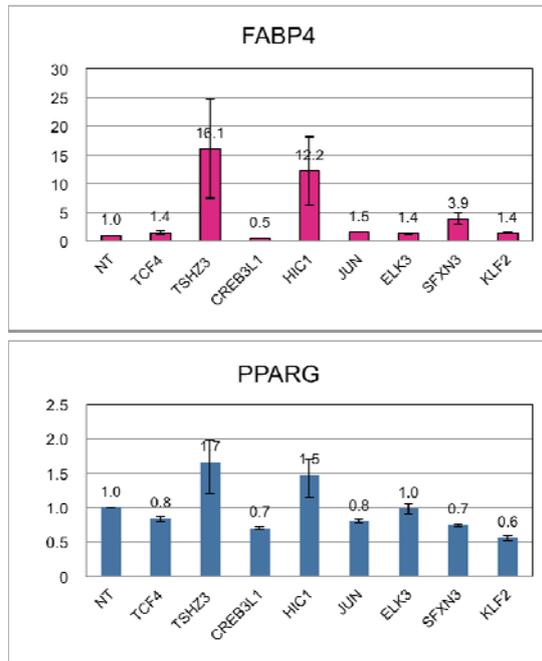


図 4 ID+DX 試薬に候補遺伝子 (左からコントロールとトップ 7 遺伝子) の siRNA を導入した時の脂肪分化誘導能

②SVR 機械学習による検証

上述した機械学習による新規遺伝子の分化能予測の実験検証は多大な研究費 (約 2 億円) がかかり、本研究費では充当できないため、既報の ES 細胞分化制御遺伝子を siRNA で網羅的に検証した論文 (Nature 2010) を利用した。CELLPEDIA から 65 種類の細胞と、ES 細胞の遺伝子発現データを用いて、分化能を学習させ、学習に用いなかった遺伝子の分化能を予測させたところ、287 遺伝子の学習で、

未知の遺伝子 20 個の分化を相関係数 $R=0.85$ と高精度で予測できた (図 5)。

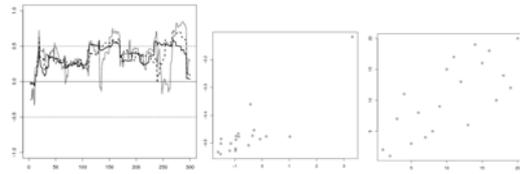


図 5 学習サンプル数を変化させた結果
左図:学習サンプル数(横軸)と相関係数(縦軸)、テストサンプル数は 20(灰実線)、40(黒破線)、60(黒実線)
中図:学習サンプル数 287 でテストサンプル数 20 の場合の実験値(横軸)と予測値(縦軸)
右図:中図と同様の条件,実験における順位(横軸)と予測における順位(縦軸)

未知の遺伝子 20 個の分化を相関係数 $R=0.85$ と高精度で予測できた (図 5)。

(5)脂肪分化遺伝子ネットワークの推定

本研究で得られた結果から脂肪分化の遺伝子ネットワークを推定した (図 6)。

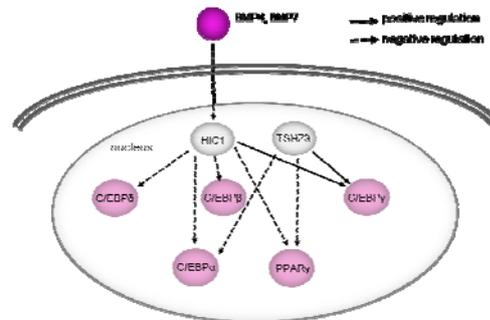


図 6 脂肪分化制御に関わる遺伝子ネットワークの推定図

(6)組み合わせによる分化因子探索法の開発

配分子算が予定より少なかったことと、前分担者 (三宅) の人事異動により平成 23 年度は同等実験であるドミナントネガティブタンパク質のタンパク質導入法による新しい遺伝子機能阻害法を設計した。我々が開発したタンパク質導入法では、カチオニックペプチド、脂質とネイティブタンパク質の複合体としてタンパク質導入するため、ネイティブタンパク質を活性型で高効率に導入可能であった (図 7)。

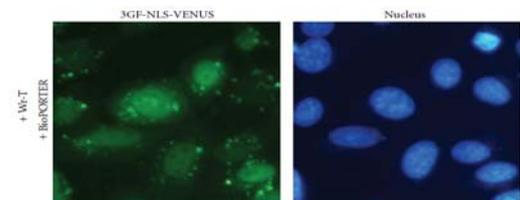


図 7 3GF-NLS-Venus タンパク質の HEK293 細胞への導入後の蛍光画像。3GF-NLS-Venus を Wt-T ペプチドと BioPORTER 脂質を用いてタンパク質導入した後、Venus 蛍光を検出した顕微鏡画像 (左) と同細胞の核を

Hoescht33342 で染色した画像 (右)

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3件)

- ① Kei Yamaguchi, Mari Inoue, Naoki Goshima, Efficient protein transduction method using cationic peptides and lipids, Vol. 2011 ID 872065, J BIOMED BIOTECHNOL. 2011、査読あり、DOI:10.1155/2011/872065
- ② Hatano, A., Chiba, H., Moesa, H. A., Taniguchi, T., Nagaie, S., Yamanegi, K., Takai, T., Tanaka, H., Fujibuchi, W. CELLPEDIA: a repository for human cell information for cell studies and differentiation analyses Database 2011, bar046 (2011)、査読あり、DOI:10.1093/database/bar046
- ③ 千葉啓和、藤渕航、細胞情報解析に役立つツール—幹細胞研究の進展とその創薬応用に向けて、実験医学、28:3171-4, 2010、査読なし、DOI:なし

[学会発表] (計7件)

- ① 千葉啓和、藤渕航、Predicting novel inducers of cell differentiation from gene expression data, Cell Functional Analysis in silico Workshop, 2011/8/19、群馬・群馬大学草津セミナーハウス
- ② 田中卓、千葉啓和、藤渕航、三宅正人、分化阻害因子の抑制によるヒト間葉系幹細胞の脂肪分化誘導、日本再生医療学会総会、新宿京王プラザホテル、2011年3月1日
- ③ 千葉啓和、田中卓、三宅正人、藤渕航、細胞分化制御のためのデータマイニング技術の開発、LS-BT合同研究発表会、2011年1月12、産業技術総合研究所つくばセンター
- ④ Hirokazu Chiba, Taku Tanaka, Masato Miyake, Wataru Fujibuchi, Prediction of novel efficient inducers for mesenchymal stem cell differentiation, CBRC2010, CBRC/AIST, 2010/7/29
- ⑤ Taku Tanaka, Hirokazu Chiba, Wataru Fujibuchi, Masato Miyake, Adipogenesis induced by short interference RNAs predicted by informatics, Systems Biology: Global Regulation of Gene Expression, Cold Spring Harbor Laboratory, New York, March 25, 2010
- ⑥ Hirokazu Chiba, Taku Tanaka, Masato Miyake, Wataru Fujibuchi, A method for

optimizing gene combination to induce adipocyte differentiation from mesenchymal stem cells, Systems Biology: Global Regulation of Gene Expression, Cold Spring Harbor Laboratory, New York, March 24, 2010

- ⑦ 幡野晶子、Harry Amri Moesa、千葉啓和、谷口丈晃、永家聖、山根木康嗣、藤渕航、細胞分化解析を目指した網羅的ヒト細胞データベース、情報処理学会・バイオ情報学研究会、北陸先端大学、2010年3月5日

[図書] (計1件)

- ① 藤渕航、Humana Press Inc., U.S.A., Methods in Molecular Biology, Vol. 577, pp. 55-65, 2009

[産業財産権]

○出願状況 (計1件)

名称: プロテインタグ、タグ化タンパク及びタンパク精製方法
発明者: 五島直樹、森正敏
権利者: 産業技術総合研究所
種類: 特許
番号: 特願 2012-087214
出願年月日: 2012/4/6
国内外の別: 国内

[その他]

ホームページ等
細胞分化データベース
<http://cellpedia.cbrc.jp/>

6. 研究組織

(1) 研究代表者

藤渕航 (FUJIBUCHI WATARU)
独立行政法人産業技術総合研究所・生命情報工学研究センター・研究チーム長
研究者番号: 60273512

(2) 研究分担者

2009~2010年度
三宅正人 (MIYAKE MASATO)
独立行政法人産業技術総合研究所・バイオメディカル研究部門・連携研究隊長
研究者番号: 60344173

2011年度
五島直樹 (GOSHIMA NAOKI)
独立行政法人産業技術総合研究所・バイオメディシナル情報研究センター・主任研究員
研究者番号: 70215482