

令和元年6月21日現在

機関番号：14301

研究種目：若手研究(B)

研究期間：2017～2018

課題番号：17K12776

研究課題名（和文）DNAメチル化情報に基づく細胞の詳細な分類と評価手法の開発

研究課題名（英文）Development of detailed cell classification and cell evaluation method based on DNA methylation

研究代表者

森 智弥 (Mori, Tomoya)

京都大学・化学研究所・助教

研究者番号：50795333

交付決定額（研究期間全体）：（直接経費） 1,000,000円

研究成果の概要（和文）：本研究では、DNAのメチル化情報に基づく細胞の詳細な分類法の開発を行った。研究代表者は、NCBIが公開しているデータベースからヒト及びマウスの単一細胞メチロームデータを取得した後、遺伝子領域上流のCpGサイトのメチル化率を計算し、自己組織化マップによるメチル化パターンの可視化を行った。公開されている単一細胞メチロームデータは多くはないが、メチル化パターンの可視化を網羅的に行ったところ、細胞種によって異なるメチル化パターンを示す傾向にあることが分かった。本手法によって可視化された単一細胞のメチル化パターンは公開データベースにて閲覧可能な状態にしてまとめている。

研究成果の学術的意義や社会的意義

DNAメチル化は生体内における遺伝子の発現制御などと深く関連しているため、個々の細胞種やその状態を特徴づける重要な指標となりうる。近年、様々な生体組織から得られた細胞のDNAメチル化パターンに関する情報が多数公開されつつあるが、本研究結果に基づいて細胞種やその状態を詳細に分類・評価することが可能となれば、ヒトiPS細胞から作成された人工細胞の安全性に関する品質管理などに利用できると考えられるため、再生医療に大きく貢献できることが期待される。

研究成果の概要（英文）：In this study, we developed a detailed classification method of cells based on DNA methylation information. After obtaining single-cell methylome data of humans and mice from the database published by NCBI, we calculated the methylation rates of CpG sites upstream of the gene regions to visualize their methylation pattern using self-organizing map (SOM). Although there are few published single-cell methylome data, comprehensive visualization of methylation patterns showed that they tend to show different patterns depending on cell type. The methylation patterns for each single-cell data visualized by this research are summarized and made available in a public database.

研究分野：バイオインフォマティクス

キーワード：DNAメチル化 細胞種分類 単一細胞解析 自己組織化マップ 機械学習

様式 C-19、F-19-1、Z-19、CK-19（共通）

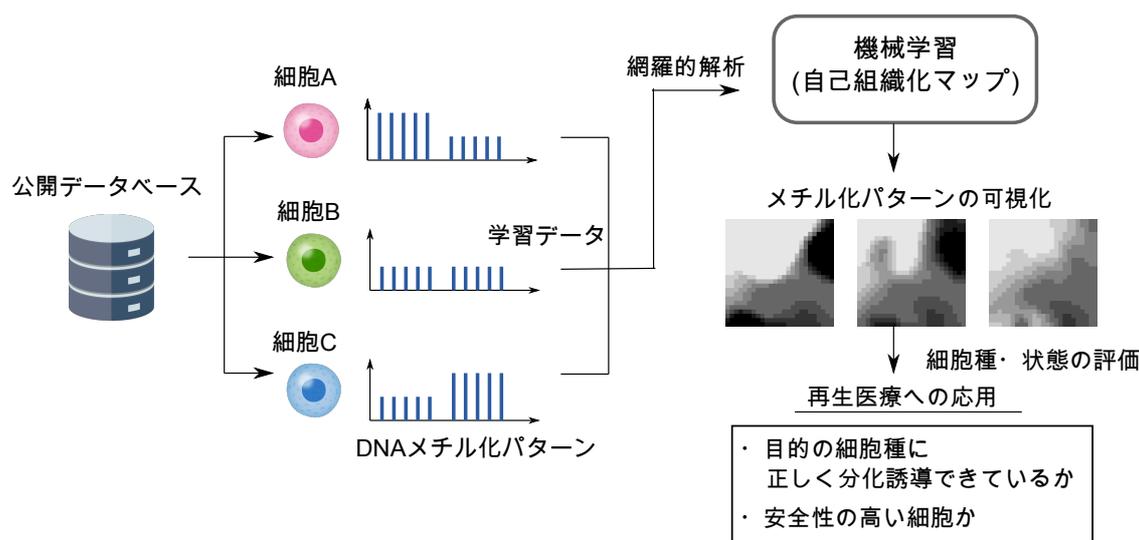
1. 研究開始当初の背景

2007年にヒト iPS 細胞が報告されて以来[1]、その臨床応用に向けた研究が進行しているが、ヒト iPS 細胞から作製された人工細胞の安全性に関する品質管理が再生医療における最重要課題の一つとなっており、細胞種やその状態への関心が高まっている。近年、個々の細胞状態を1細胞レベルで観察できる技術が普及し、トランスクリプトーム（ある時点の細胞内における転写産物の総体）を解析することで細胞種やその状態を分類する研究が多数報告され始めている[2-3]。これまで、次世代シーケンサーを用いたメチル化解析が数多く行われてきたが、あらゆる細胞種の正常細胞または疾患細胞に特異的なメチル化パターンを網羅的かつ詳細に調べた研究は未だ十分には行われていないものと思われる。特に、1細胞のDNAメチル化プロファイリングについてはいくつかの新しい技術が開発され始めたものの、公開されているデータ数はトランスクリプトームに比べると僅かである。そのようななか、2016年10月に健康と疾患理解のためにヒトの全細胞に関するリファレンスマップを作製する Human Cell Atlas プロジェクト (<https://www.humancellatlas.org>) が立ち上がったことで、ヒト細胞に関する研究が今後ますます発展すると考えられ、1細胞トランスクリプトームだけでなく1細胞メチロームのデータも充実していくことが予想される。

2. 研究の目的

本研究では、従来のようにトランスクリプトームを用いるのではなく、メチローム（DNAメチル化情報の総体）に基づいた細胞の詳細な分類方法の開発を行うことを目的とする（図1）。具体的には、細胞種が既知であるメチローム情報を学習し、細胞種特異的なメチル化パターンを抽出する手法を研究する。この手法が開発されれば、ヒト iPS 細胞から作製された人工細胞の品質管理に応用できると考えられる。例えば、iPS 細胞を目的の細胞種 A へ分化誘導させた際に、そのメチロームデータを既に公開されている細胞種 A の正常細胞のメチロームデータと比較することで、得られた細胞が正しく分化誘導できているかを評価することができる。さらに近年では、1細胞メチローム解析に関する技術が開発され始め[4]、個々の細胞種のメチル化パターンの不均一性を調べることも可能となりつつあるが、目的の細胞種に分化誘導させた際に、それが安全性が高く、正しく機能する組織を形成できるような細胞が作製できているかを示す指標として利用できるため、再生医療への応用が期待される。

研究代表者はこれまで、1細胞トランスクリプトーム解析に関する研究を行ってきたが[5]、その過程において、同じ細胞種でも遺伝子発現パターンが個々の細胞で大きく異なる際に、それが細胞の性質による違いであるのか、もしくは遺伝子発現の揺らぎ[6]が原因であるのかを正確に判断することが困難な場合が多く存在した。一方で、1細胞メチローム解析に関する研究にも携わるなかで、メチル化パターンは細胞種の決定に重要な役割を持ち、細胞分裂を経てもそのパターンが維持される機構を持つことから、メチル化パターンに基づいた解析を行えば、揺らぎに大きく影響されることなく、個々の細胞を詳細に評価可能となることが期待される。



3. 研究の方法

本研究目的を達成するため、DNAメチル化情報に基づいた細胞分類法に関して、(1) 公開データの収集、(2) 配列データ解析、そして(3) 自己組織化マップを用いた1細胞メチル化パターンの可視化を行った。

(1) 公開データの収集

公開されているメチロームデータを網羅的に解析し、ゲノム配列上のどの領域のどのようなメ

チル化パターンが細胞を特徴付けているのかを明らかとする。具体的には、NCBI(National Center for Biotechnology Information)が公開している GEO(Gene Expression Omnibus)もしくは SRA(Sequence Read Archive)からヒト及びマウスの1細胞メチロームデータの配列データを取得した。

(2) 配列データ解析

(1) でダウンロードした配列データを参照配列にマッピングすることで、転写開始点上流領域の CpG サイトのメチル化レベルを網羅的に計算した。網羅解析には同一の基準で配列データを解析することが重要であると考えられるが、公開されているデータには異なる解析方法により得られたメチル化データが含まれる。そのため、本解析では生の配列データをダウンロードし、全ての配列データに対して同一の前処理、参照ゲノム配列、そして参照配列ゲノムへのマッピング方法を用いることで、実験プロジェクト間の差を可能な限り排除し、同一の基準で評価できるようにした。

(3) 自己組織化マップを用いた1細胞メチル化パターンの可視化

(2) で計算した転写開始点上流領域の CpG メチル化レベルに基づいて、1細胞メチル化パターンを可視化した。可視化には自己組織化マップ(Self-organizing map: SOM)と呼ばれる教師無し機械学習クラスタリング手法を用いることで、事前知識を用いることなく転写開始点上流領域をそのメチル化レベルにしたがってクラスタリングし、メチル化パターンをマップ上にヒートマップとして表現した。さらに生物種及びプラットフォーム毎に分けて自己組織化マップを適用することで、プラットフォームの違いが原因で生じる測定差の影響もできる限り排除し、細胞種特異的なメチル化パターンを可視化した。

4. 研究成果

(1) DNA メチル化データとトランスクリプトームデータの揺らぎの比較

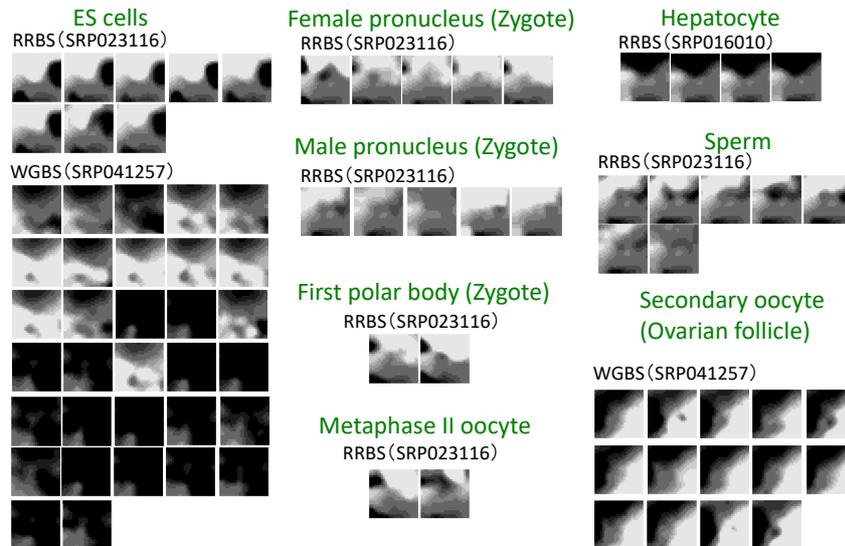
「3. 研究の方法」で示した解析によって得られた DNA メチル化パターンと遺伝子発現パターンの揺らぎ、つまり不均一性について比較した。ここで、遺伝子発現データは、メチル化パターン解析と同様、GEO や SRA によって提供されている配列データを参照ゲノム配列にマッピングさせることによって各遺伝子の発現量を計算し、さらに自己組織化マップによって発現パターンを可視化した。比較の結果として、DNA メチル化パターンと遺伝子発現パターンともに個々の細胞でパターンのばらつきは見られたものの、遺伝子発現パターンに比べ、メチル化パターンがやや安定している傾向が見られた。

(2) 細胞種特異的なメチル化パターンの検出

既に述べた通り、公開されている1細胞 DNA メチル化データはトランスクリプトームデータに比べると僅かではあるものの、網羅的に行ったメチル化パターン解析では、細胞種特異的なメチル化パターンを示すことが確認された。例えば、次世代シーケンサー HiSeq 2000 によって取得された 79 サンプルのマウス1細胞メチル化パターン解析を行ったところ、マウス ES 細胞では、RRBS (Reduced representation bisulfite sequencing) 法と WGBS (Whole Genome Bisulfite Sequencing) 法という異なった2つのメチル化解析プロトコルによって得られたメチル化パターンはそれぞれ異なったパターンを示したが、他の細胞種 (Zygote, Oocyte, Hepatocyte, Sperm, Ovarian follicle) ではそれぞれ特有のパターンを示した。さらに Zygote では同じ細胞種でも雄と雌で異なるパターンを示すことが確認された (図2)。

<参考文献>

- ① Kazutoshi Takahashi et al., Induction of pluripotent stem cells from adult human fibroblasts by defined factors, *Cell*, Vol.131, 2007, pp.861-872.
- ② Jin Li et al., Single-cell transcriptomes reveal characteristic features of human pancreatic islet cell type, *EMBO reports*, Vol.17, 2016, pp.178-187.
- ③ Gioele La Manno et al., Molecular diversity of midbrain development in mouse, human, and stem cells, *Cell*, Vol.167, 2016, pp.566-580.
- ④ Wai-Shin Yong et al., Profiling genome-wide DNA methylation, *Epigenetics & Chromatin*, Vol.9, 2016, 16 pages.
- ⑤ Tomoya Mori et al., Development of 3D tissue reconstruction method from single-cell RNA-seq data, *Genomics and Computational Biology (selected articles from the 2nd Challenges in computational biology: Gene expression data analysis, Dec.1-2, 2016, Mainz, Germany)*, e53, 2017, 4 pages.
- ⑥ Avigdor Eldar et al., Functional roles for noise in genetic circuits, *Nature*, Vol.467, 2010, pp.167-173.



SRP*: Sequence Read Archive において付与されているプロジェクト番号

図 2 : 次世代シーケンサ (HiSeq 2000) によって得られたマウスの 1 細胞 DNA メチル化パターン

5. 主な発表論文等

〔学会発表〕 (計 5 件)

- ① Tomoya Mori and Wataru Fujibuchi, Ab initio 3D reconstruction algorithm for mid-gastrula mouse embryo using machine learning, CiRA 2017 International Symposium, Kyoto, Japan, 6-8 November, 2017.
- ② 谷山 暢子, 小林 健太, 山根 順子, 森 智弥, 山下 潤, 藤渕 航, 1 細胞 DNA メチローム解析による細胞個体差のエピジェネティック解析, 日本遺伝学会第 89 回大会, 岡山, 2017 年 9 月 13-16 日.
- ③ Junko Yamane, Tomoya Mori, Nobuko Taniyama, Kenta Kobayashi, and Wataru Fujibuchi, Development of new enhanced single-cell RRBS protocol toward in-silico 3D tissue reconstruction, Single Cell Science Symposium, Yokohama, Japan, 6-7 July, 2017.
- ④ Midori Yuji, Kunie Sakurai, Tomoya Mori, Junko Yamane, Nobuko Taniyama, Kenta Kobayashi, Koji Yamanegi, and Wataru Fujibuchi, Clear classification of human cells and establishment of cell identification system, Single Cell Science Symposium, Yokoyama, Japan, 6-7 July, 2017.
- ⑤ Junko Yamane, Tomoya Mori, and Wataru Fujibuchi, Development of enhanced single-cell RRBS method toward 3D tissue reconstruction, The 12th International Workshop on Advanced Genomics, Tokyo, Japan, 27-29 June, 2017.

〔その他〕

ホームページ等

SHOGoin: ヒト細胞情報統合データベース

<http://shogoin.stemcellinformatics.org>

上記 web ページ内で 1 細胞 DNA メチル化パターン情報を提供している。

6. 研究組織

- (1) 研究分担者 : なし
- (2) 研究協力者 : なし

※科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属されます。