

令和 2 年 4 月 31 日

海外特別研究員最終報告書

独立行政法人日本学術振興会 理事長 殿

採用年度 平成 30 年

受付番号 201860029

氏 名 神元 健児

(氏名は必ず自署すること)

海外特別研究員としての派遣期間を終了しましたので、下記のとおり報告いたします。

なお、下記及び別紙記載の内容については相違ありません。

記

1. 用務地（派遣先国名）用務地：ミズーリ州セントルイス（国名：アメリカ合衆国）
2. 研究課題名（和文）※研究課題名は申請時のものと変わらないように記載すること。
遺伝子人工回路を用いたパイオニア転写因子操作による細胞運命制御機構の単一細胞解析
3. 派遣期間：平成 30 年 4 月 1 日 ～ 令和 2 年 3 月 31 日
4. 受入機関名及び部局名
ワシントン大学医学部セントルイス校, Department of developmental biology
5. 所期の目的の遂行状況及び成果…書式任意 **書式任意 (A4 判相当 3 ページ以上、英語で記入も可)**
(研究・調査実施状況及びその成果の発表・関係学会への参加状況等)
(注)「6. 研究発表」以降については様式 10－別紙 1～4 に記入の上、併せて提出すること。

次ページ以降に記載

研究内容、目的

ヒトを含むほとんどの多細胞生物は多種多様な細胞種で構成される。その細胞の運命や分化の状態がどのように制御されているのかを知り、利用することは発生学、細胞生物学といった基礎的領域に留まらず、薬剤探索、発生工学、再生医療、がん治療など、多様な領域にとって重要であり、様々な方面から研究が行われている。iPS 細胞研究に代表されるように、近年、重要な転写因子を組み合わせることで、人為的に発現させることで、人工的に細胞の運命を転換できることが証明された。転写因子の発現は細胞の運命制御に重要な役割を持ち、適切な転写因子の過剰発現により、分化した細胞から別の分化した細胞（神経・肝細胞・免疫細胞・表皮細胞など）が得られることが明らかとなっている。このように転写因子を人為的に操作し、特定の細胞種を誘導する現象は細胞のリプログラミングと呼ばれ、とくに分化した細胞から別の分化した細胞を直接誘導する現象はダイレクトリプログラミングと呼ばれている。

ダイレクトリプログラミングの過程では少数の転写因子の発現の変化が細胞全体の性質の変換を誘導する。この過程では多くの遺伝子の変動が複雑に絡み合っており、どのような機構で細胞運命を返還させるかについては多くが明らかになっていない。派遣先の Samantha Morris らの先行研究により、ダイレクトリプログラミングで導入する転写因子の発現量に依存して全く違った性質の細胞が誘導されることがわかっており、ダイレクトリプログラミングは用いる転写因子の種類だけでなく、転写因子の発現量が重要なことが示唆されている。

また、ダイレクトリプログラミングの過程やその結果生じる細胞は非常に多様で不均一であり、仮に均一な細胞を用いて実験を始めた場合においても生じる不均一性・不確定性は排除できないことが大きな問題であった。

本研究ではダイレクトリプログラミング時の転写因子の量的な変化がどのような機構で細胞の運命制御に影響するのかを明らかにすることを目的とする。

また、ダイレクトリプログラミングにおいては導入した転写因子の変化が別の遺伝子の変動を誘導し、さらにその遺伝子が別の遺伝子に作用していく。よって導入した転写因子がどのように細胞の性質を変化させるのかをシステムレベルで解析するには、少数の遺伝子の変化だけでは無く、関連する多数の遺伝子を網羅的に調べ、それぞれの依存関係（遺伝子ネットワーク）を解析する必要がある。細胞がリプログラミングを行う過程では細胞の転写パターンは大きく変わることがわかっているが、その変化を引き起こす要因である遺伝子ネットワークがどのように変化していくのかを定量的に解析することを目指す（図 1）。

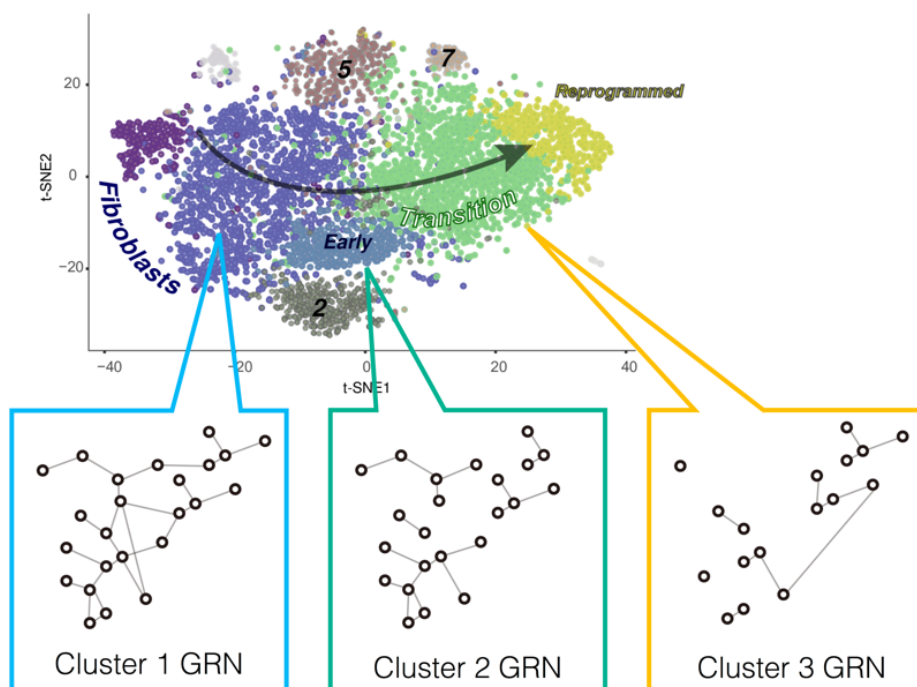


図1 ダイレクトリプログラミング時の遺伝子ネットワークの変遷

また上記したようにダイレクトリプログラミングの結果得られる細胞集団は不均一な細胞集団であり、多くの細胞状態が誘導されるため、細胞集団をまとめて解析する従来の手法で無く、細胞ひとつひとつの違いを検出できる実験系を用いる必要がある。最後に、転写因子の量的な違いが及ぼす影響を正確に解釈するには、数理モデルによる定量的な結果の解釈が必要であると考えた。

これらの条件を考慮し、細胞のダイレクトリプログラミングにおける転写因子の量的な制御機構を解明するため、本研究では single cell RNA-sequencing (scRNA-seq) 法を実験系として用いる。これは細胞をひとつずつ分画し、単一の細胞内の mRNA を網羅的に測定する技術である。単一細胞レベルでかつ網羅的な遺伝子発現の変動を測定し、その変動を説明できる数理モデル作成することで、転写因子の量的な細胞運命の制御機構を明らかにすることを目指す。

作成した数理モデルの妥当性は再度実験系によって証明する。具体的にはダイレクトリプログラミング時の遺伝子ネットワークの状態をモデル化した際に、導入した転写因子が細胞の運命制御を行う際に重要だと思われる中継遺伝子をノックアウトし、ダイレクトリプログラミングの効率や方向性が変化することを確認する。また、作成した数理モデルを利用してダイレクトリプログラミングの効率や不均一性などを向上させる条件を推測し、それを実験によって証明することを目指す。

研究の実施状況

(1) scRNA-seq を利用した転写因子の数値モデルの作成

細胞のリプログラミングのモデルとして、本研究ではマウス胎児線維芽細胞 (Mouse Embryonic Fibroblas, MEF) から転写因子 Foxa1 と Hnf4a によって内胚葉前駆細胞 (induced Endodermal Progenitors, iEPs) を誘導する実験系を利用する。

MEF にレトロウイルスによって Foxa1, Hnf4a を過剰発現させ、そのあと様々なタイミングで scRNA-seq を行い、データを取得した。また、研究室内の別の研究者とともに行った本プロジェクトの関連プロジェクト (Biddy BA, Kong W, Kamimoto K, Guo C, Waye SE, Sun T, Morris SA., *Nature*, Dec;564(7735):219-224.) において取得した scRNA-seq の時系列データも利用して解析を行った。

転写因子が細胞内で作り出す遺伝子ネットワークを定量的にモデリングする新規手法、(図 2) を作成した (以下、この新規手法を CellOracle と呼称する。) CellOracle 利用して前述のダイレクトリプログラミングモデルにおける細胞内転写因子ネットワークの変化を解析した。

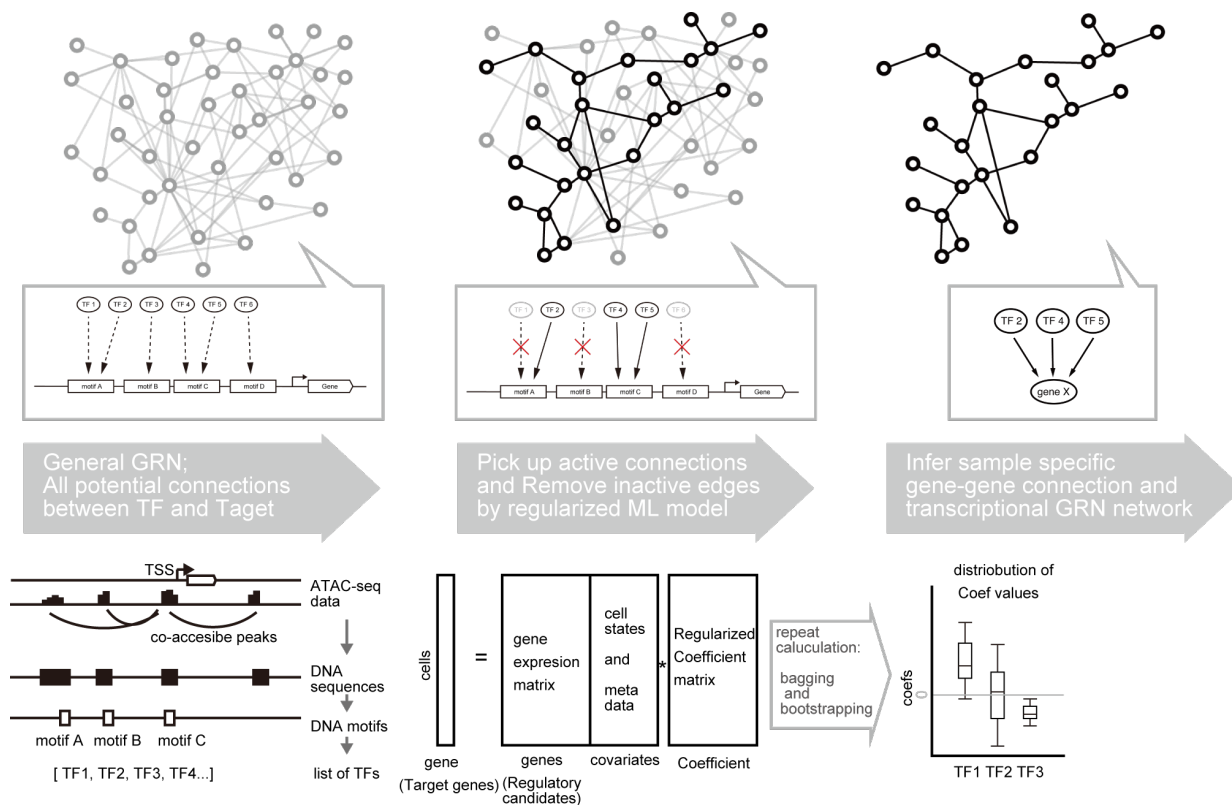


図 2 : scRNA-seq のデータから遺伝子ネットワークを推定するプログラムを作成した

(2) Perturb-seq 法による遺伝子発現阻害実験

前項で記したように、scRNA-seq データから遺伝子ネットワークの数値モデルを作成することにより、リプログラミング時に転写因子が細胞の中でどのような仕組みで分化を制御するのかを定量的に解明するのが本研究の目的である。このネットワークの解析によってリプログラミング時に重要な遺伝子パスウェイの推定や、転写因子の遺伝子発現を変動させた時のリプログラミングへの影響のシミュレーションを行うことができる。これらのモデルでの予測が妥当であるのかを実験的に検証するため、遺伝子発現の阻害実験を行った。

CellOracle の遺伝子ネットワークの数値モデルにより、リプログラミングの過程で重要だと思われる遺伝子群を予測し、その遺伝子群を一つずつノックアウトした際にどのような変化が起こるのかを調べた。

本研究では実験系の一部として scRNA-seq の応用技術である Perturb-seq 法を利用する。これは scRNA-seq の実験系に CRISPR-Cas9 法と DNA-Barcoding 技術を組み合わせたものであり、この技術によって多数の遺伝子ノックアウト実験を単一の培養皿で行うことができる。

ここでは、遺伝子ネットワークや scRNA-seq の解析でピックアップした 15 遺伝子について阻害実験を行った。15 遺伝子のそれぞれについて 3 種のノックアウト用コンストラクト (single guide RNA) を設計し、そのレンチウイルス系 45 種類を作成した。このウイルスを Cas9 タンパクが発現している細胞に感染させることで、目的の遺伝子を破壊することができる。このとき、それぞれのウイルス DNA には single guide RNA の種類に応じた固有の DNA バーコードが組み込んであるように設計を行っている。のちに細胞を scRNA-seq によって解析を行った際に DNA バーコードが内在の mRNA と共に読み込まれるが、このバーコードが guide RNA と対応しているため、細胞ひとつひとつがどのノックアウト用コンストラクトを保持するのかを判別することができ、単一の実験でも 45 の遺伝子についてのノックアウト実験が行え、その結果は single cell レベルで解析することができる (図 3)。

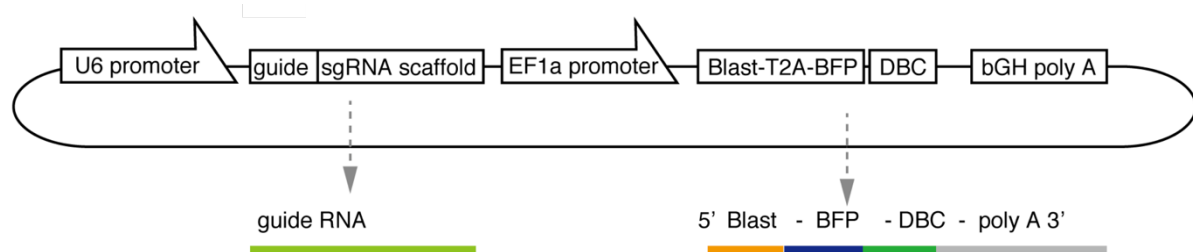


図3 DNA バーコードを組み込んだノックアウトコンストラクトの概要

まず本研究の実験モデルである iEP に Cas9 をレトロウイルスを用いて導入し、Cas9 発現 iEP を作成した。次に前述のノックアウト用レンチウイルス 4 5 種をまとめて感染させる。このとき感染効率を低くするように設定し、1 細胞に 1 種以上のウイルスが感染しないようにする。その後ウイルスに組み込んだ BFP によってノックアウトコンストラクトが感染した細胞のみを FACS で分取することで、ノックアウト細胞のプールを得る。この集団は 4 5 種類のノックアウトコンストラクトのいずれかを持つ細胞が混合してできている状態である。

この細胞を scRNA-seq 解析を行い、遺伝子発現情報と共に得られる DNA バーコード情報からどの遺伝子がノックアウトされたものなのかをラベル付けする。

Perturb-seq の解析の結果、9900 個の細胞について情報を取得し、ノックアウトの結果を解析した。

その結果、CellOracle で予測していたいくつかの遺伝子が実際にダイレクトリプログラミング時に重要な働きを持つことがわかった。他の実験手法も利用してこの遺伝子機能の解析を行い、数理モデルでの予測結果を実験的に証明した。また、この研究で新たに同定した重要な転写因子群を追加で利用することで、ダイレクトリプログラミングの効率を上昇させることに成功した (図 4)。

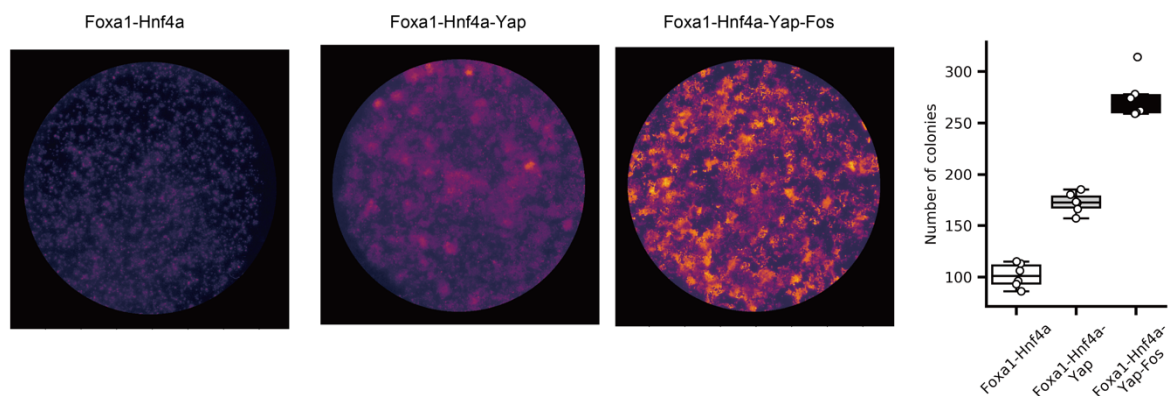


図4 新規転写因子によるダイレクトリプログラミングの効率上昇結果

これらの結果を論文としてまとめ、プレプリント論文として公表した ([Kamimoto K](#), Hoffmann CM, Morris SA. CellOracle: Dissecting cell identity via network inference and in silico gene perturbation.

BioRxiv doi: <https://doi.org/10.1101/2020.02.17.947416>).

また、現在査読付き化学ジャーナルに投稿し、査読を受けている。

研究発表、関係学会への参加状況

Kamimoto K, Morris SA.

INFERRING GENE REGULATORY NETWORKS FROM SINGLE CELL RNA SEQUENCE DATA

Nature Genomics meeting. New York University, August

Kamimoto K, Morris SA.

Inferring Gene Regulatory Network from Single Cell RNA-seq.

2019 Inaugural Joint Retreat by the Developmental, Regenerative and Stem Cell Biology (DRSCB) Program and the Department of Developmental Biology Washington University in St.Louis, April, Missouri

Kamimoto K, Morris SA.

Inferring Gene Regulatory Network from Single Cell RNA-seq.

2018 Chan Zuckerberg Science Initiative meeting, October, Massachusetts