様　式　Ｃ－１９、Ｆ－１９－１、Ｚ－１９　（共通）

科学研究費助成事業　　研究成果報告書

令和　７ 年　 ６ 月 １１ 日現在

機関番号：１７１０４
研究種目：若手研究
研究期間：2022～2024
課題番号：２２Ｋ１８１７３

研究課題名（和文）シングルセルオミクス解析によるダイレクトリプログラミング誘導転写因子の予測

研究課題名（英文）Prediction of transcription factors that induce direct reprogramming using single-cell multi-omics data

研究代表者
濱野　桃子（Momoko, Hamano）

九州工業大学・大学院情報工学研究院・准教授

研究者番号：４０７１７３３６

交付決定額（研究期間全体）：（直接経費）　　3,500,000円

研究成果の概要（和文）：細胞を分化誘導する方法として、iPS細胞を介さずに直接目的の臓器の機能をもつ細胞に変換するダイレクトリプログラミングがある。膨大な転写因子からダイレクトリプログラミングを誘導する候補を実験のみで同定することは困難なため、効率良く低コストでダイレクトリプログラミングを誘導する転写因子を予測できるin silico手法の開発が望まれている。そこで本研究では、ダイレクトリプログラミングを誘導する転写因子を予測するin silico手法を確立した。一細胞レベルのマルチオミクスデータを基に転写因子を予測する手法を構築した結果、バルクのオミクスデータに基づいた予測より予測精度が向上することがわかった。

研究成果の学術的意義や社会的意義
これまでのダイレクトリプログラミング誘導転写因子の同定は、実験者が複数の転写因子の組み合わせにより構成される膨大な実験検証により明らかにされてきた。しかし、膨大な組み合わせの探索には多大な労力と時間を要し、またすべての細胞種において十分に網羅的な知見が得られているわけではない。
本研究でマルチオミクス情報を統合した転写因子予測のin silico手法を構築したことで、誘導因子が未解明な細胞種や新規な転写因子候補の新たな開拓を可能にするという点で学術的意義がある。特に、一細胞レベルの高解像度なオミクス情報を活用することで、従来のバルクの視点の予測手法を超える高精度な転写因子の同定が期待される。

研究成果の概要（英文）：Direct reprogramming enables the conversion of fully differentiated somatic cells into various other cells without bypassing induced pluripotent stem cells. Direct reprogramming is induced by the transfection of multiple transcription factors.
However, it is difficult to identify the transcription factors
that induce direct reprogramming by only experimental validation. In this study, I developed a in silico method to predict the transcription factors that induce direct reprogramming in various human cell types. Based on single-cell multi-omics data, I predicted direct reprogramming-inducing transcription factors for cell types in six tissues.As a result, prediction accuracy was improved by using single-cell multi-omics data compared to bulk omics data.

研究分野：バイオインフォマティクス

キーワード：ダイレクトリプログラミング　転写因子　シングルセル　マルチオミクス　パイオニアファクター

※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、研究者個人に帰属します。

１．研究開始当初の背景

ダイレクトリプログラミングは、iPS 細胞のような未分化の細胞を介さずに、分化した体細胞を別の種類の細胞に分化誘導する技術である。iPS 細胞を経由した細胞分化誘導法より短期間かつ少ない誘導因子で細胞を作製できることから、将来の再生医療を担う革新的な技術として注目されている。ダイレクトリプログラミングは転写因子の遺伝子導入により誘導されるが、ダイレクトリプログラミングを誘導する転写因子の組み合わせを膨大な候補から実験のみで同定することは困難なため、効率良く低コストでダイレクトリプログラミングを誘導する転写因子を同定できる *in silico* 手法の開発が望まれている。しかし、これまでに開発されている様々な細胞種へダイレクトリプログラミングを誘導する転写因子を予測する高精度な *in silico* 手法は開発されていない。その原因として、複数の細胞種が混在したバルクの遺伝子発現データを用いることで標的細胞以外の遺伝子発現値がノイズとなり、標的細胞のみを誘導する最適な転写因子を同定できていないことが問題点としてあげられる。

２．研究の目的

ダイレクトリプログラミングを誘導する転写因子を予測するための情報基盤技術を開発する。一細胞レベルのマルチオミックスデータを活用し、転写活性を促進するパイオニア転写因子を考慮した一細胞レベルの転写制御機構を組み込むことで高精度にダイレクトリプログラミング誘導転写因子を予測可能な *in silico* 手法を開発することを目的とする。

３．研究の方法

まず、申請者が先行研究において構築したバルクのマルチオミックスデータを用いたダイレクトリプログラミングを誘導する転写因子を予測する手法（Eguchi et al., *Bioinformatics* (2022)）に基づき、ゲノム情報、エピゲノム情報、トランスクリプトーム情報、分子ネットワーク情報を収集した。次いで、脳、肝臓、小腸、膵臓、骨格筋、心臓組織由来のシングルセルマルチオミックスデータ（single-cell RNA-seq 及び single-cell ATAC-seq）を取得し、シングルセル解析により組織に含まれる各細胞種をアノテーションした。ダイレクトリプログラミングで作製したい目的の細胞種のみのデータを抽出し、そのシングルセルマルチオミックスデータを入力として、ダイレクトリプログラミングを誘導する転写因子を予測する新たな *in silico* 手法を構築した。*in silico* 手法は各細胞種に対して各転写因子が候補となる最もらしさをスコアとして算出する計算手法とし、よりスコアが高い転写因子が候補の可能性が高いとみなした。予測された転写因子候補の上位にダイレクトリプログラミングが誘導されることが実験的に証明されている正解となる転写因子がどれだけ含まれているかによって、本研究で構築した新規 *in silico* 手法の予測精度を評価した。

申請当初の計画から改良した点として、先行研究の手法との精度比較のために当初予測対象としていた軟骨細胞はシングルセルマルチオミックスデータを取得できなかったため、筋芽細胞を予測対象に変更した。

４．研究成果

（１）上位に予測されたダイレクトリプログラミング誘導転写因子候補の評価

６つの組織由来の様々な細胞種（心筋細胞、肝細胞、膵β細胞、パネート細胞、神経細胞、筋芽細胞）に対してダイレクトリプログラミングを誘導する転写因子の候補をそれぞれ予測することができた。予測された転写因子の上位には実験的にダイレクトリプログラミングを誘導することが証明されている転写因子（右図の太字）やパイオニア転写因子が含まれていた。それらの転写因子は、６つの細胞種全てにおいて統計学的有意にランキングの上位に予測できたことから、本研究で構築した *in silico* 手法の有用性が示された。

	心筋細胞
1	**TBX5**
2	HAND1
3	**NKX2-5**
4	**ESRRG**
5	ISL1
6	**MEIS1**
7	NFIC
8	**HAND2**
9	CIC
10	PBX4
P値	4.69e-5

図　心筋細胞に対して予測された転写因子の上位10（太字はダイレクトリプログラミングを誘導すると実験的に証明されている転写因子）

（２）先行研究のバルクマルチオミックスデータを用いた予測手法との精度比較

６つの組織由来の様々な細胞種のうち、先行研究における予測対象だった細胞種と共通する５つの細胞種（心筋細胞、肝細胞、膵β細胞、パネート細胞、神経細胞）に対して予測精度を比較した。その結果、５つの細胞種全てにおいて本研究で構築した予測手法の方が先行研究の予測手法より予測精度が高いことが分かった。特に、パネート細胞は飛躍的に予測精度が向上することがわかった。この予測精度の向上は全ての細胞種で確認されたことから、バルクのマルチオミックスデータから一細胞レベルのマルチオミックスデータに置き換えたことで、複数の細胞種が混在することで起こるノイズが軽減され、予測精度の向上に寄与した可能性が示唆された。

（３）シングルセルマルチオミックスデータが一部得られない細胞種に対する予測

本提案手法では２つのシングルセルマルチオミックスデータ（single-cell RNA-seq, single-cell ATAC-seq）を入力データとして候補転写因子を予測する手法を構築したが、single-cell RNA-seq もしくは single-cell ATAC-seq のいずれかしか得られない場合でも転写因子の予測が可能か検証した。その結果、得られないシングルセルオミックスデータによって計算過程で欠損値が生じても、片方のシングルセルオミックスデータから情報を補完することで、ある一定の高い予測精度を維持しつつ転写因子の候補を予測できることが分かった。このことから、シングルセルマルチオミックスデータの取得が難しい細胞種に対しても本手法を適用できることが分かった。

５．主な発表論文等

〔雑誌論文〕　計2件（うち査読付論文　2件／うち国際共著　0件／うちオープンアクセス　2件）

１．著者名 Hamano Momoko、Nakamura Toru、Ito Ryoku、Shimada Yuki、Iwata Michio、Takeshita Jun-ichi、Eguchi Ryohei、Yamanishi Yoshihiro	４．巻 40
２．論文標題 DIRECTEUR: transcriptome-based prediction of small molecules that replace transcription factors for direct cell conversion	５．発行年 2024年
３．雑誌名 Bioinformatics	６．最初と最後の頁 btae048
掲載論文のDOI（デジタルオブジェクト識別子） 10.1093/bioinformatics/btae048	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 －

１．著者名 Eguchi Ryohei、Hamano Momoko、Iwata Michio、Nakamura Toru、Oki Shinya、Yamanishi Yoshihiro	４．巻 38
２．論文標題 TRANSDIRE: data-driven direct reprogramming by a pioneer factor-guided trans-omics approach	５．発行年 2022年
３．雑誌名 Bioinformatics	６．最初と最後の頁 2839～2846
掲載論文のDOI（デジタルオブジェクト識別子） 10.1093/bioinformatics/btac209	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 －

〔学会発表〕　計9件（うち招待講演　0件／うち国際学会　0件）

１．発表者名 濱野桃子、川崎瞭太、廣瀬昌樹、山西芳裕
２．発表標題 シングルセルマルチオミクスデータを用いた生体内リプログラミング誘導転写因子の予測
３．学会等名 第23回日本再生医療学会総会
４．発表年 2024年

１．発表者名 濱野桃子、川崎瞭太、廣瀬昌樹、江口凌平、岩田通夫、沖真弥、山西芳裕
２．発表標題 シングルセルマルチオミクスデータを用いたダイレクトリプログラミング誘導転写因子の予測
３．学会等名 第46回日本分子生物学会年会
４．発表年 2023年

１．発表者名
濱野桃子、中村透、岩田通夫、江口凌平、竹下潤一、山西芳裕

２．発表標題
Small compound-based direct cell conversion with combinatorial optimization of pathway regulations

３．学会等名
第12回生命医薬情報学連合大会

４．発表年
2023年

１．発表者名
廣瀬昌樹，濱野桃子，江口凌平，岩田通夫，山西芳裕

２．発表標題
一細胞レベルのオミックスデータの統合解析によるダイレクトリプログラミング転写因子予測手法の開発

３．学会等名
第46回日本分子生物学会年会

４．発表年
2023年

１．発表者名
伊藤 綠風，濱野 桃子，山西 芳裕

２．発表標題
シングルセルレベルの細胞変換過程を考慮したダイレクトリプログラミング誘導化合物の予測

３．学会等名
第12回生命医薬情報学連合大会

４．発表年
2023年

１．発表者名
Momoko Hamano, Ryohei Eguchi, Michio Iwata, Toru Nakamura, Shinya Oki and Yoshihiro Yamanishi

２．発表標題
TRANSDIRE:パイオニア転写因子を考慮したトランスオミクス手法によるデータ駆動型ダイレクトリプログラミング

３．学会等名
第１１回生命化学情報工学医薬情報学連合大会（IIBMP2022）

４．発表年
2022年

１．発表者名
濱野桃子，江口凌平，岩田通夫，中村透，沖真弥，山西芳裕

２．発表標題
パイオニア転写因子を考慮したトランスオミクス手法によるデータ駆動型ダイレクトリプログラミング

３．学会等名
第45回日本分子生物学会年会

４．発表年
2022年

１．発表者名
川﨑瞭太，伊藤綠風，濱野桃子

２．発表標題
皮膚線維芽細胞から膵β細胞へのダイレクトリプログラミングを誘導する転写因子の予測

３．学会等名
第61回化学関連支部合同九州大会

４．発表年
2024年

１．発表者名
川﨑瞭太，伊藤綠風，濱野桃子

２．発表標題
機械学習によるインスリン分泌細胞への細胞変換を誘導する転写因子の予測

３．学会等名
日本農芸化学会2024年度西日本支部大会

４．発表年
2024年

〔図書〕　計1件

１．著者名	４．発行年
大澤　毅	2023年
２．出版社	５．総ページ数
羊土社	222
３．書名	
マルチオミクス　データ駆動時代の疾患研究	

〔産業財産権〕

〔その他〕

大学プレスリリース「細胞の直接変換を誘導する転写因子を高精度で予測するAI手法を開発」
https://www.kyutech.ac.jp/whats-new/press/entry-9065.html

6．研究組織

	氏名 （ローマ字氏名） （研究者番号）	所属研究機関・部局・職 （機関番号）	備考

7．科研費を使用して開催した国際研究集会

〔国際研究集会〕　計0件

8．本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関