

## 細胞種を直接変換させる低分子化合物を予測できる 機械学習アルゴリズムを開発 ～転写因子を模倣する、より低リスクな化合物の探索～

九州工業大学大学院情報工学研究院の濱野桃子助教・岩田通夫准教授、名古屋大学大学院情報学研究科の山西芳裕教授の研究グループは、国立研究開発法人産業技術総合研究所安全科学研究部門の竹下潤一主任研究員との共同研究により、iPS細胞を介さずに、すでに分化した細胞を別の種類の細胞へと直接変換するダイレクトリプログラミングを誘導する低分子化合物(薬剤など)の組み合わせを予測する機械学習アルゴリズムを開発しました。

### (ポイント)

- すでに分化した細胞を別の種類の細胞へと直接変換するダイレクトリプログラミングを誘導する低分子化合物(薬剤など)の組み合わせを予測する機械学習アルゴリズムを開発した。
- ダイレクトリプログラミングを誘導する転写因子の働きを模倣する低分子化合物の候補を情報科学的に提案することで、細胞のがん化のリスクを回避することに貢献できる。
- 提案手法は、今後の再生医療分野における細胞治療のための細胞作製法の開発に繋がることが期待される。

本研究グループは、すでに分化した細胞を別の種類の細胞へと直接変換するダイレクトリプログラミングを誘導する低分子化合物(薬剤など)を予測する機械学習アルゴリズムを開発しました。ダイレクトリプログラミングの誘導には転写因子の遺伝子の導入が一般的ですが、細胞のがん化のリスクがあることから、低分子化合物で誘導する方法の開発が切望されていました。そこで本提案手法では、ダイレクトリプログラミングを誘導する転写因子の働きを模倣する低分子化合物を情報科学的に予測する機械学習アルゴリズムを開発しました。まず、細胞が転写因子の導入によってダイレクトリプログラミングを誘導される際の特徴を遺伝子の発現パターンから捉えました。次いで、細胞が変換する過程で変化する遺伝子発現パターンを模倣する低分子化合物の最適な組み合わせを探索しました。シミュレーティッドアニーリングの枠組みで開発した最適化アルゴリズムにより、ダイレクトリプログラミングを誘導する低分子化合物の組み合わせを予測することに成功しました。提案手法により、転写因子に置き換わる低分子化合物を簡便に予測できるようになり、今後の再生医療分野における細胞作製法の開発を促進することが期待されます。

本研究成果は、2024年1月25日にBioinformaticsで公開されました。研究の詳細は別紙をご参照ください。

## <研究の背景と経緯>

ダイレクトリプログラミングとは、iPS細胞のような未分化の細胞を介さずに、すでに分化した細胞を別の種類の細胞に直接変換する技術です。短期間で低コストに細胞を作製できることから、転写因子<sup>\*1</sup>の導入によるダイレクトリプログラミング誘導法の研究が注目されており、将来の再生医療を担う革新的な技術として期待されています。しかし、従来のダイレクトリプログラミングは、転写因子をコードする遺伝子を元細胞に導入するため、遺伝子導入に起因する細胞のがん化のリスクがありました。そこで近年では、発がんリスクを回避するため、低分子化合物<sup>\*2</sup>(薬剤や化学物質など)を用いたダイレクトリプログラミングが注目されています(図1)。

ダイレクトリプログラミングを誘導する低分子化合物の最適な組み合わせを実験的に同定するのは膨大な実験コストがかかり、極めて困難です。また低分子化合物より転写因子の導入の方が細胞の変換効率が高いため、ダイレクトリプログラミングを誘導する方法は転写因子を中心に開発されており、低分子化合物による誘導法の開発は遅れています。そのため、近年蓄積されている生命科学ビッグデータを有効活用し、ダイレクトリプログラミング誘導転写因子と同じ働きをする低分子化合物を予測する技術が切望されています。しかしながら、ダイレクトリプログラミング誘導転写因子に置き換わる低分子化合物を予測する手法は未だに確立されていませんでした。

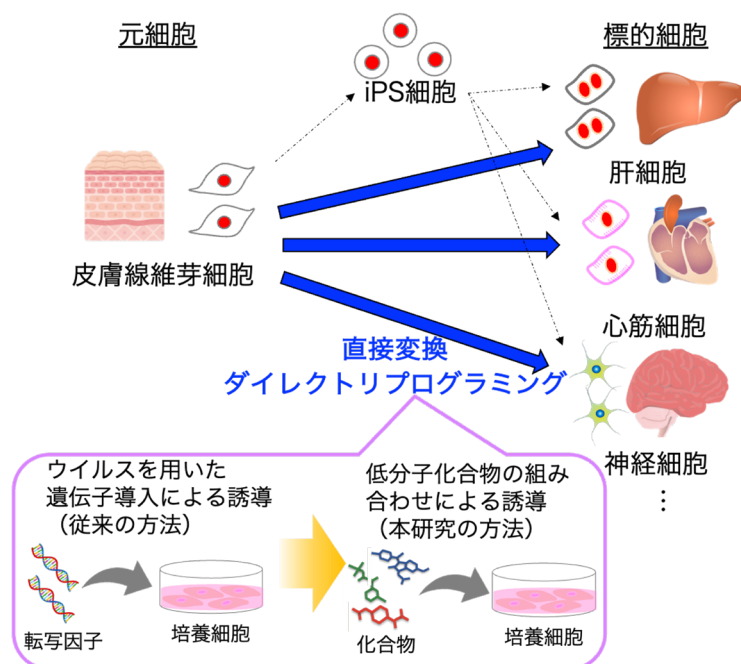


図1 低分子化合物を用いて細胞を別の種類の細胞へと直接変換するダイレクトリプログラミング

## <研究の内容>

九州工業大学大学院情報工学研究院の濱野桃子助教・岩田通夫准教授、名古屋大学大学院情報学研究科の山西芳裕教授の研究グループは、国立研究開発法人産業技術総合研究所安全科学研究部門の竹下潤一主任研究員との共同研究により、ダイレクトリプログラミングを誘導する低分子化合物を予測する機械学習アルゴリズムを開発しました。

本研究の提案手法では、まず、ダイレクトリプログラミングを誘導することが知られている転写因子が細胞にもたらす遺伝子発現パターンの変化を調べました。次いで、その遺伝子発現パターンを模倣する低分子化合物の最適な組み合わせを探索するため、シミュレーティッドアニーリング<sup>\*3</sup>の枠組みで最適化アルゴリズムを開発し、それを適用することでダイレクトリプログラミングを誘導する低分子化合物の組み合わせを予測しました(図2)。

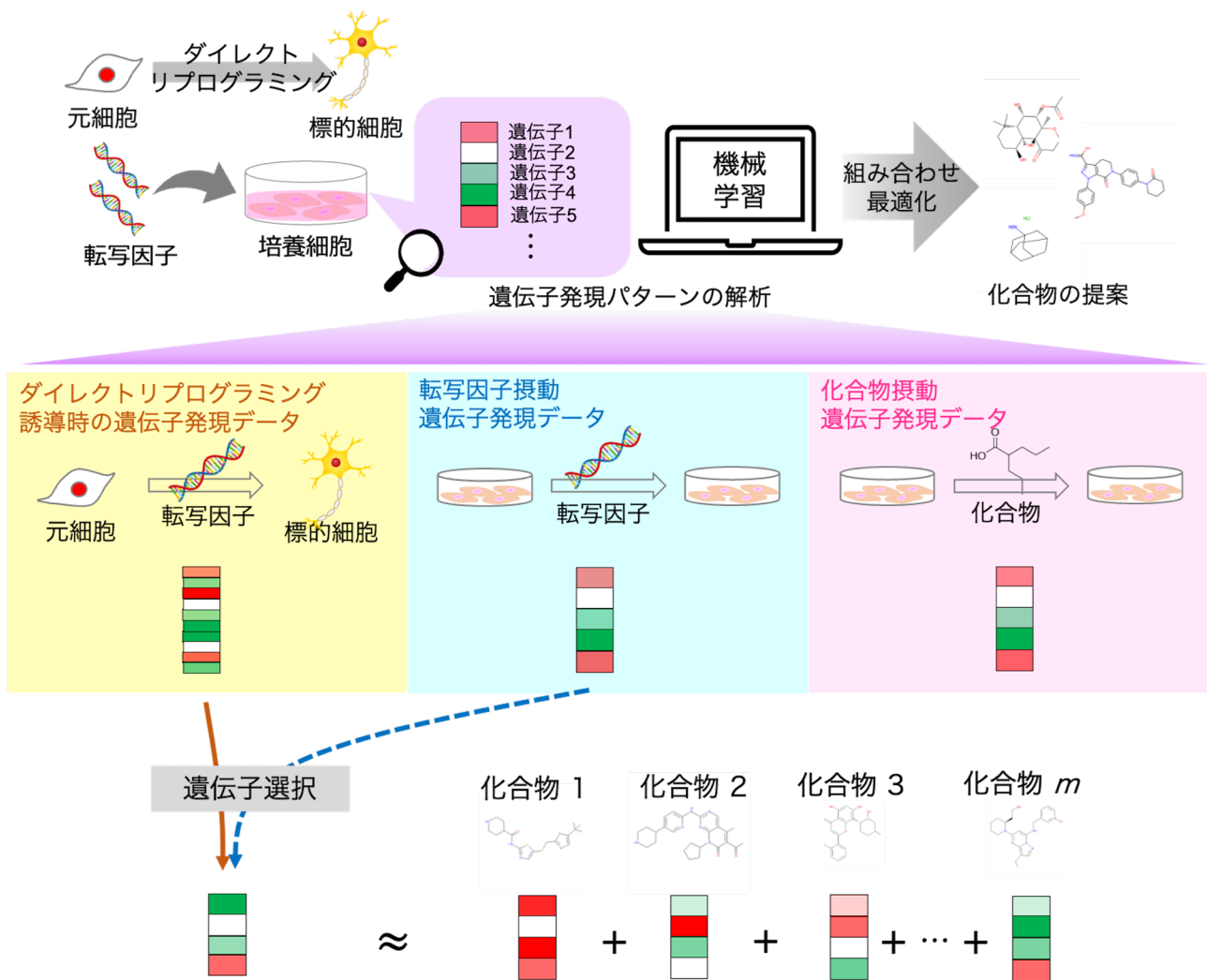


図2：ダイレクトリプログラミングを誘導する低分子化合物の組み合わせを予測する提案手法

ダイレクトリプログラミングを誘導することが実験的に証明されている転写因子の情報に基づき、転写因子と同じ働きをする低分子化合物の組み合わせを予測した。転写因子によってダイレクトリプログラミングを誘導した際の遺伝子発現データからダイレクトリプログラミング誘導時の遺伝子発現パターンを特定する。次いで、ダイレクトリプログラミング誘導時に変化する遺伝子発現パターンを模倣する化合物の最適な組み合わせを発見するアルゴリズムを開発した。

提案手法を用いて、線維芽細胞から神経細胞や心筋細胞への直接変換を誘導する低分子化合物を予測しました。その結果、10個の化合物の組み合わせが予測されました(図3)。予測された化合物群には、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既知の化合物であるTTNPB、Romidepsin、Etoposideが含まれており、統計学的有意に精度良く既知の化合物を再現できていることが分かりました。また、新規に予測された化合物には、神経活動に関与するシグナル経路であるNeuroactive ligand-receptor interactionを制御するMetaraminolやZafirlukastが含まれていました。心筋細胞を誘導すると予測された化合物として、心筋細胞へのダイレクトリプログラミングを誘導する制御パスウェイ<sup>※4</sup>として報告されているMAPK signalingやJAK-STAT signalingを制御するBnometinibやFigotinibが同定されました。これらの結果から、新規に予測された化合物群には各細胞の機能や分化を制御する化合物が含まれていることから、生物学的な妥当性が示唆されました。

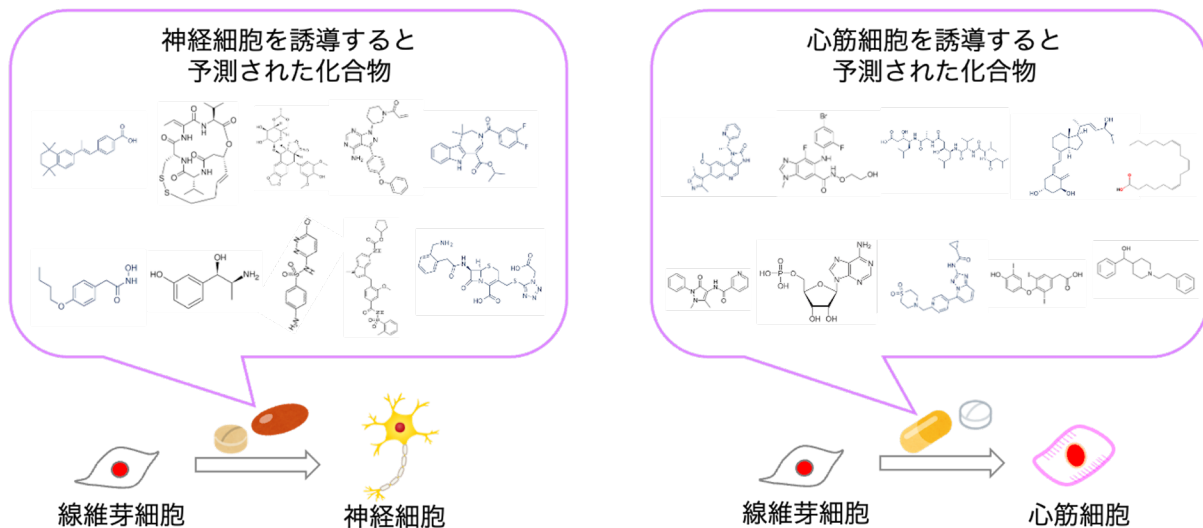


図3：提案手法を用いて予測されたダイレクトリプログラミングを誘導する可能性が示唆された化合物の組み合わせ

線維芽細胞から神経細胞と心筋細胞への直接変換を誘導すると予測された化合物の組み合わせを示す。

次いで、予測された化合物がダイレクトリプログラミングを誘導する分子メカニズムを理解するために、予測された化合物が作用するタンパク質を同定し、タンパク質間の共発現の関係性をネットワークとして確認しました（図4）。神経細胞を誘導すると予測された化合物群が作用するタンパク質ネットワークには、**Mapk1**, **Src**, **Fyn** が高い中心性を持つタンパク質として同定されました。**Mapk1** はダイレクトリプログラミングを誘導する標的パスウェイである **MAPK signaling** を構成するタンパク質です。また、**Src** と **Fyn** は神経細胞のシグナル伝達系に関与することが報告されています。これらの結果から、神経細胞のダイレクトリプログラミングを誘導する化合物群は、ダイレクトリプログラミングに重要なパスウェイのタンパク質を網羅し、神経細胞への発達を誘導するシグナル伝達経路も制御することが分かりました。心筋細胞を誘導すると予測された化合物群が標的とするタンパク質ネットワークのうち中心性の高い標的タンパク質として炎症反応を誘導するサイトカインである **IL6** が検出されました。心筋細胞へのダイレクトリプログラミングを促進する機構として、炎症反応を抑制することが報告されています。予測された化合物には抗炎症薬である **Gamolenic acid** が同定されており、それらを含む化合物を組み合わせることで、心筋細胞へのダイレクトリプログラミングを促進する可能性が示唆されました。

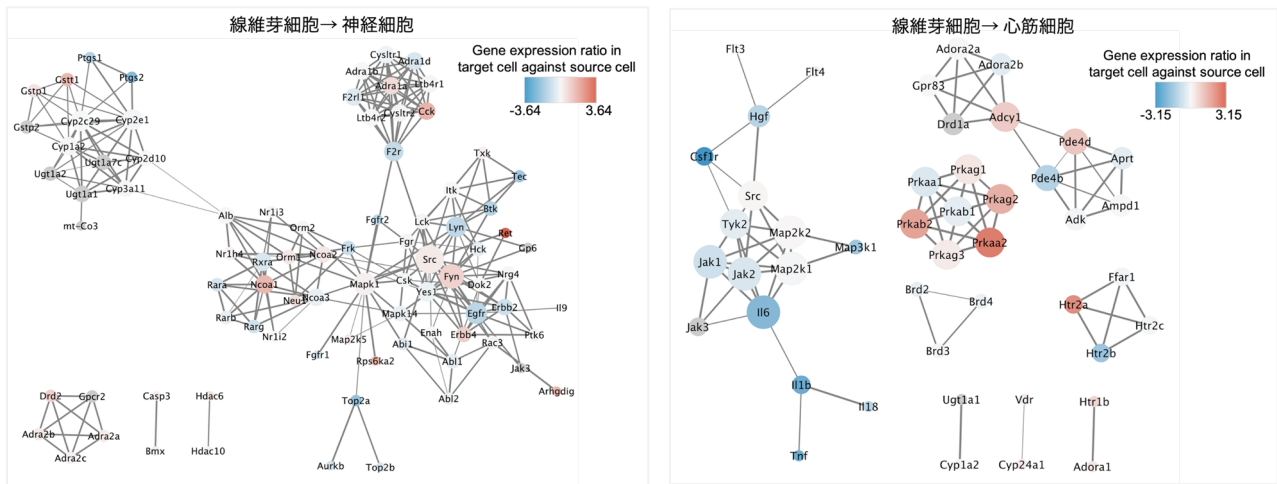


図4：予測された化合物が制御する標的タンパク質のネットワーク

神経細胞を誘導すると予測された化合物群及び心筋細胞を誘導すると予測された化合物群が作用するタンパク質のネットワーク（左が神経細胞、右が心筋細胞）。ダイレクトリプログラミングが誘導される際の各タンパク質の遺伝子発現レベルを色で示している。

#### <今後の展開>

本研究では、ダイレクトリプログラミングを誘導する化合物の組み合わせを予測する機械学習アルゴリズムを開発しました。提案手法は、ダイレクトリプログラミングを誘導する転写因子が制御する遺伝子発現パターンを模倣する化合物を見出す点が特長であり、また、その遺伝子発現パターンを基に最適な化合物の組み合わせをシミュレーティッドアニーリングによって予測する先駆的な情報技術です。

本提案手法は、ダイレクトリプログラミングを誘導する転写因子が明らかになれば、心筋細胞や神経細胞以外の細胞変換へ応用することができます。そのため、ダイレクトリプログラミングを誘導する方法が転写因子のみしか報告されていない細胞種に対しても適用することが可能です。このことから、本研究は再生医療における細胞治療法の開発への貢献に繋がることが期待されます。

本研究は、公益財団法人大川情報通信基金「安価で安全な再生医療を拓く情報技術基盤の構築」（研究代表者：濱野桃子）、JSPS 科学研究費補助金・挑戦的研究（萌芽）JP18K19930「低分子化合物による転写因子制御と細胞直接変換のためのインシリコ予測法の開発」（研究代表者：山西芳裕）、JSPS 科学研究費補助金・学術変革領域研究（B）JP20H05797「シナジー効果を有する化合物群のAIによる探索と設計」（研究代表者：山西芳裕）、挑戦的研究（開拓）JP21K18327「AIによるデータ駆動型ダイレクトリプログラミングの創生と腫瘍化リスクの回避」（研究代表者：山西芳裕）の支援を受け行われました。

#### <発表雑誌>

雑誌名：「Bioinformatics」（オンライン版：1月25日）

論文タイトル：DIRECTEUR: Transcriptome-based prediction of small molecules that replace transcription factors for direct cell conversion

著者：Hamano, M., Nakamura, T., Ito, R., Shimada, Y., Iwata, M., Takeshita J., Eguchi, R. and Yamanishi, Y.

DOI 番号：10.1093/bioinformatics/btae048

## <用語解説>

### ※1 転写因子

DNA に結合して遺伝子の発現を調節するタンパク質。様々な転写因子を細胞に導入することで標的の遺伝子を発現調節することができる。

### ※2 低分子化合物

生物活性を有する化学物質や薬剤のこと。特定の生体分子（タンパク質）を標的として、その機能を阻害もしくは活性化する分子標的薬を含む。

### ※3 シミュレーテッドアニーリング

所望の条件を満たす最適解を探索するアルゴリズムの枠組みの一つ。焼きなまし法とも呼ばれる。

### ※4 生物学的パスウェイ

遺伝子やタンパク質の機能的な相互作用を経路として表したもの。代謝や遺伝子の制御関係、シグナル伝達系などの様々な種類のパスウェイが存在する。

#### 【研究に関するお問い合わせ先】

国立大学法人 九州工業大学大学院情報工学研究院  
生命化学情報工学研究系 助教 濱野桃子  
E-mail: momoko@bio.kyutech.ac.jp  
TEL:0948-29-7831

東海国立大学機構 名古屋大学大学院情報学研究科  
教授 山西芳裕  
E-mail: yamanishi@i.nagoya-u.ac.jp  
TEL:052-789-5638

#### 【報道に関するお問い合わせ先】

国立大学法人 九州工業大学総務課広報係  
E-mail: pr-kouhou@jimmu.kyutech.ac.jp  
TEL:093-884-3007

東海国立大学機構 名古屋大学広報課  
E-mail: nu\_research@t.mail.nagoya-u.ac.jp  
TEL: 052-558-9735