

令和 6 年 6 月 16 日現在

機関番号：13901

研究種目：挑戦的研究（開拓）

研究期間：2021～2023

課題番号：21K18327

研究課題名（和文）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

研究課題名（英文）Creation of data-driven direct reprogramming by AI and avoidance of tumorigenic risk

研究代表者

山西 芳裕（Yamanishi, Yoshihiro）

名古屋大学・情報学研究科・教授

研究者番号：60437267

交付決定額（研究期間全体）：（直接経費） 19,900,000円

研究成果の概要（和文）：iPS細胞を介さずに目的の臓器の細胞に直接変換するダイレクトリプログラミングが革新的な再生医療技術として注目されている。しかしながら、ダイレクトリプログラミングを誘導する因子セット（転写因子や低分子化合物など）を同定するのは極めて困難である。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や低分子化合物を予測する情報技術を構築した。さまざまな多階層オミクスデータを基に、細胞の直接変換を誘導する転写因子セットや、それを代替する低分子化合物セットを予測する最適化アルゴリズムを開発した。皮膚線維芽細胞から神経細胞や心筋細胞などへの直接変換において有用性を示した。

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の標準的な方法は、レトロウイルスを用いた転写因子の遺伝子導入であり、ウイルスに起因する発がんリスクという深刻な問題があった。そこで、本研究では、近年蓄積されている医薬関連の様々なデータから、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することが目的であり、そこに学術的意義がある。通常の細胞の直接変換は、ウイルスを用いて必要な転写因子の遺伝子を元細胞に導入するため、ウイルスに起因する発がんリスクなどの問題があるが、それを回避することによって、再生医療の臨床応用を促進することができるという社会的意義がある。

研究成果の概要（英文）：Direct reprogramming, direct cell conversion without the use of iPS cells, is attracting attention as an innovative regenerative medicine technology. However, it is extremely difficult to identify the set of factors (e.g., transcription factors, small molecule compounds) that induce direct reprogramming. In this study, we developed information technologies to predict the transcription factors and small molecule compounds that induce direct reprogramming. Based on various multi-level omics data, we developed an optimization algorithm to predict the set of transcription factors that induce direct cell conversion and the set of small molecule compounds that replace them. We demonstrated the usefulness of this algorithm in the direct conversion of skin fibroblasts into neurons, cardiomyocytes, etc.

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 腫瘍化リスク データ駆動 細胞直接変換 AI

様式 C-19、F-19-1 (共通)

1. 研究開始当初の背景

ダイレクトリプログラミング（細胞の直接変換）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ある。iPS 細胞を経由した細胞の作製法より短期間で低コストに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。しかしながら、従来の細胞の直接変換は、レトロウイルスなどのウイルスを用いて必要な遺伝子を元細胞に導入するため、レトロウイルスなどのウイルスに起因する腫瘍化リスクなどの問題があった。そこで近年では、腫瘍化リスクを回避するため、低分子化合物（薬剤や化学物質など）の添加による細胞の直接変換が切望されている。

細胞の直接変換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ことは多くの実験コストがかかり、極めて困難である。また低分子化合物による細胞の直接変換のメカニズムは不明であり、添加すべき低分子化合物の数が多ければ、臨床応用の際に重篤な副作用に繋がる可能性が高まる。そのため、低分子化合物を用いた細胞の直接変換のメカニズムの解明や少ない低分子化合物での誘導が求められる。しかしながら、細胞の直接変換の分子生物学的メカニズムに基づいて低分子化合物を予測する手法は未だに確立されていない。

2. 研究の目的

本研究では、転写因子だけでなく低分子化合物で制御するダイレクトリプログラミングを提案する。ダイレクトリプログラミングを担う転写因子セットだけでなく、それを代替する低分子化合物セットを予測し、腫瘍化リスクを回避できるようなダイレクトリプログラミングを支援するためのインシリコ手法を開発するのが目的である。

3. 研究の方法

まず、細胞分化誘導に関与する転写因子や低分子化合物のオミックスデータ、細胞分化誘導に関与するパスウェイのデータを整備した。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融合解析し、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を開発した。次いで、オミクス情報と化学構造情報を融合解析し、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手法を開発した。膨大な低分子化合物とタンパク質の相互作用データを基に、細胞変換に関連している生物学的パスウェイを推定し、細胞変換に関わる生物学的パスウェイを効率的に制御できるように組み合わせを最適化することによって、細胞の直接変換を誘導する低分子化合物の新たな組み合わせを予測するアルゴリズムを開発した。最後に、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を基に、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する手法を開発した。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し、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実装した。

4. 研究成果

(1) 多階層オミクス情報に基づく転写因子の解析結果

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融合解析し、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を開発した。本研究では、細胞変換に関与するエピジェネティックな修飾の変化を介して転写を活性化するパイオニア因子に着目し、クロマチン構造から標的の細胞種それぞれのパイオニア因子を推定した。次いで、様々なマルチオミクス情報や細胞間の系統関係情報を統合し、パイオニア因子を考慮した上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する世界初のコンセプトを提唱した。提案手法をヒトゲノムにコードされている約千個の転写因子に対して適用し、各細胞直接変換に対して転写因子セットを予測し、それを制御する低分子化合物を大規模推定するためのモデルを実装した。提案手法は、オミクスデータが入手できるさまざまな細胞に対して適用できる。

本提案手法を用いて、線維芽細胞から肝臓、骨、脳、心臓、脾臓、腸の6種類の細胞への直接変換を誘導する転写因子を予測した。その結果、全ての標的細胞に対して転写因子の予測を行うことができ、また、予測された転写因子にはパイオニア因子を含む転写因子の組み合わせが同定された。予測された転写因子には、先行研究で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既知の転写因子を多く含んでおり、予測の妥当性を確認できた。例えば、線維芽細胞から肝細胞への直接変換では、予測上位10個のうち既知の転写因子6個を正確に再現できた。このことから、本提案手法を適用することで、既知の転写因子を含む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が可能であることが明らかになった。

次いで、本提案手法と3つの先行研究手法の予測精度の比較を行った。予測スコアの高い上位10個の転写因子のうち、各標的細胞に対して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の数を比較した (図 1)。その結果、本提案手法は、3つの先行研究手法より多くの既知の転写因子を再現することができ、最大で 67%の精度向上を達成するなど、高い予測精度を確認した。

クロスバリデーション実験を行い、シミュレーションでの予測精度で、既存手法を上回ることを確認した。評価指標として、AUC (area under the receiver operator characteristics curve) と AUPR (area under the precision-recall curve) を用いた。AUC とは、ROC (receiver operator characteristics) 曲線の下部分の面積のことであり、ランダムな予測の時 0.5 を示し、完全に予測されたとき 1.0 を示す。AUPR は、予測の適合率と再現率から算出される指標である。どちらの指標も、予測性能が高ければ、高い数値を示す。AUC と AUPR を用いて、標的細胞ごとにそれぞれの予測手法の性能を評価したところ、非常に高い値を示した。AUC と AUPR のどちらの指標についても、複数のオミクス情報を融合したトランスオミクス手法が一番高い予測性能である傾向を示した。この結果から、複数のオミクス情報を融合することで、性能の向上ができることが明らかとなった。例えば、皮膚線維芽細胞から肝細胞への変換に関与すると予測された転写因子の詳細について生物学的な考察を行なった。予測上位の転写因子について、文献報告と比較し、その妥当性を検証できた。

これらの研究成果は、バイオインフォマティクス分野のトップジャーナルである *Bioinformatics* 誌で出版された (Eguchi et al, *Bioinformatics*, 2022)。

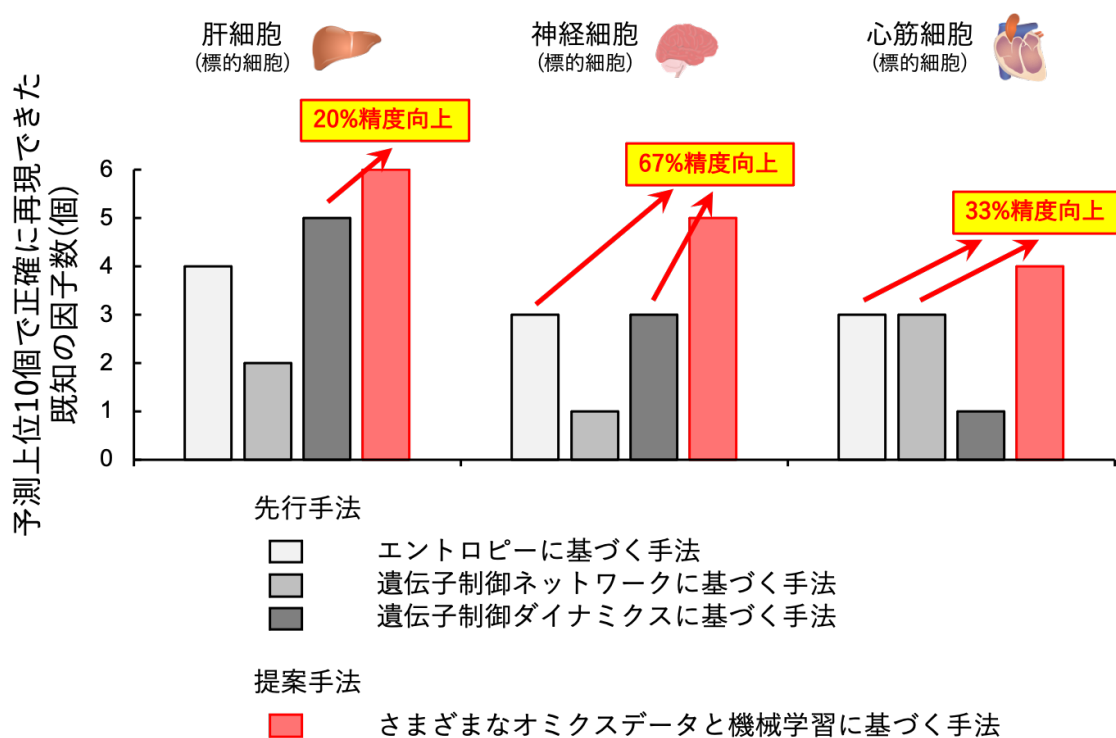


図 1: ダイレクトプログラミングを誘導する因子候補の予測性能の比較 エントロピーに基づく手法 (Entropy-based method; D' Alessio et al., *Stem Cell Reports*, 2015)、遺伝子制御ネットワークに基づく手法 (Mogrify; Rackham et al., *Nat. Genet.*, 2016)、遺伝子制御ダイナミクスに基づく手法 (DGC-based method; Ronquist et al., *Proc. Natl. Acad. Sci.*, 2017) を用いた。本提案手法を用いることにより、線維芽細胞からさまざまな標的細胞への直接変換に必要な転写因子を、より精度良く予測することができる。

(2) パスウェイ情報に基づく低分子化合物の解析結果

今回開発した手法を用いて、皮膚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の組み合わせを予測した。その結果、細胞の直接変換を誘導することが証明されている既知の低分子化合物 (バルプロ酸やフォルスコリンなど) を含む組み合わせを高い頻度で再現することが確認できた (図 2)。また、既知の細胞変換誘導化合物だけでなく、既知の細胞変換誘導化合物と類似した機能を持つ低分子化合物も多く含まれていた。このことから、本開発手法は細胞の直接変換を誘導するために重要な低分子化合物を含む新規な組み合わせを予測できることを示した。

次いで、本開発手法により予測された組み合わせを構成する低分子化合物の個数を確認した。皮膚線維芽細胞から神経細胞を誘導する 6 例の既知の化合物組み合わせ及び心筋細胞を誘導する 7 例の既知の化合物組み合わせに対して、代替となる新たな化合物組み合わせを予測した。その結果、全ての報告例に対して、予測された化合物の個数は既知の化合物の個数より少なくすることができた。例えば、神経細胞を誘導することが知られている 4 つの化合物の組み合わせに代わる化合物として、2 つの化合物からなる新たな組み合わせが予測された。また、心筋細胞を誘

導することが知られている7つの化合物の組み合わせに代わる化合物として、4つの化合物からなる新たな組み合わせが予測された。このことから、本開発手法は既知の組み合わせよりも少ない最適な化合物の組み合わせを予測できることが示唆された。このことは、化合物の数が増えると交互作用による副作用の可能性が高まるため、少ない化合物での細胞変換は安全性の向上に繋がることを期待される。

これらの研究成果は、バイオインフォマティクス分野のトップジャーナルである *Bioinformatics* 誌で出版された (Nakamura et al, *Bioinformatics*, 2022)。

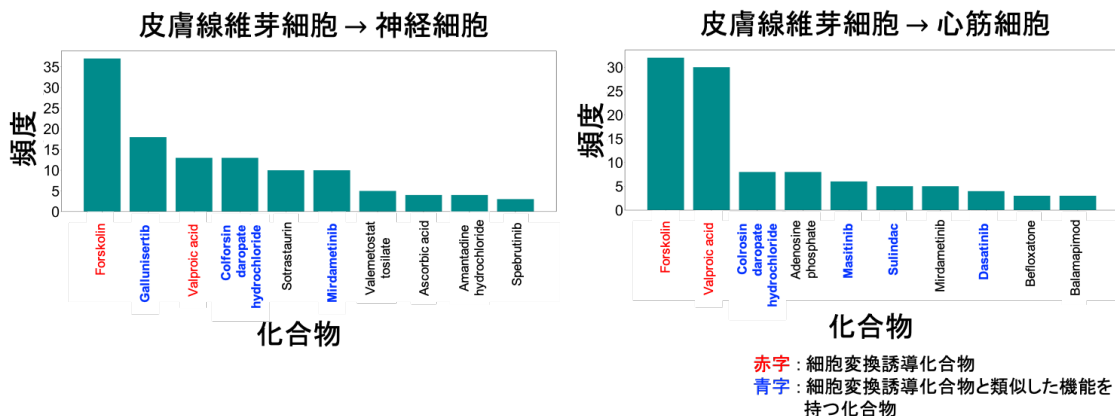


図2: 開発した手法を用いて予測された組み合わせに含まれる化合物 皮膚線維芽細胞から神経細胞と心筋細胞への直接変換を誘導すると予測された化合物の組み合わせに含まれる低分子化合物の頻度を示す。赤字は、実験的に細胞の直接変換を誘導することが証明されている既知の細胞変換誘導化合物であり、青字は、既知の細胞変換誘導化合物と類似した機能を持つ化合物を示す。

(3) トランスクリプトーム情報に基づく低分子化合物の解析結果

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た。その結果、10個の化合物の組み合わせが予測された。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った。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た。

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た (図3)。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ある。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る。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った。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る。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た。

これらの研究成果は、バイオインフォマティクス分野のトップジャーナルである *Bioinformatics* 誌で出版された (Hamano et al, *Bioinformatics*, 2024)。

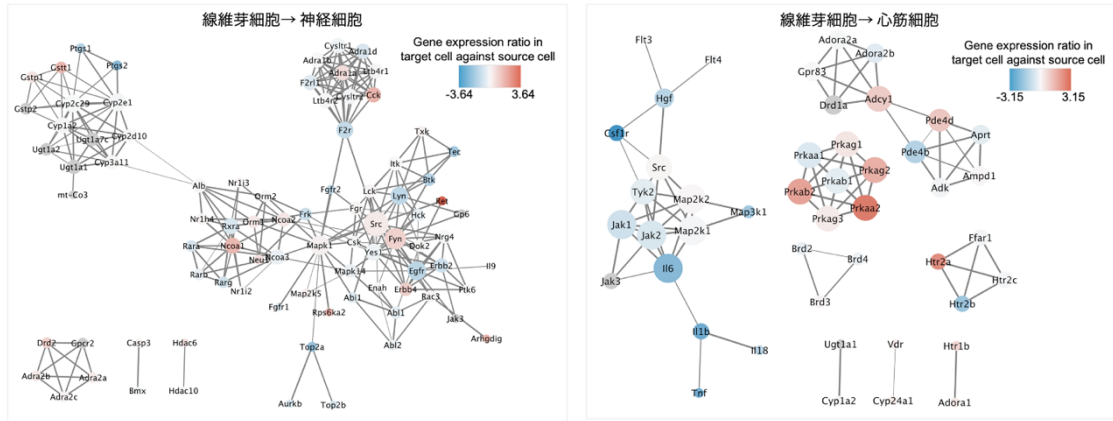


図3: 予測された化合物が制御する標的タンパク質のネットワーク 神経細胞を誘導すると予測された化合物群及び心筋細胞を誘導すると予測された化合物群が作用するタンパク質のネットワーク (左が神経細胞、右が心筋細胞)。ダイレクトリプログラミングが誘導される際の各タンパク質の遺伝子発現レベルを色で示している。

5. 主な発表論文等

〔雑誌論文〕 計3件（うち査読付論文 3件/うち国際共著 0件/うちオープンアクセス 3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 btac048 |
| 掲載論文のDOI（デジタルオブジェクト識別子） 10.1093/bioinformatics/btac048 | 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 Nakamura Toru, Iwata Michio, Hamano Momoko, Eguchi Ryohei, Takeshita Jun-ichi, Yamanishi Yoshihiro | 4. 巻 38 |
| 2. 論文標題 Small compound-based direct cell conversion with combinatorial optimization of pathway regulations | 5. 発行年 2022年 |
| 3. 雑誌名 Bioinformatics | 6. 最初と最後の頁 ii99 ~ ii105 |
| 掲載論文のDOI（デジタルオブジェクト識別子） 10.1093/bioinformatics/btac475 | 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 1件/うち国際学会 1件）

| |
|--|
| 1. 発表者名 山西芳裕 |
| 2. 発表標題 機械学習によるデータ駆動型細胞リプログラミング |
| 3. 学会等名 第14回 がんゲノム・エピゲノム、数理統計解析についての研究会（招待講演） |
| 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 ダイレクトリプログラミングを誘導するパスウェイ制御機構の同定と最適な低分子化合物の組み合わせ予測 |
| 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 伊藤緑風, 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 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 ダイレクトリプログラミングを誘導する低分子化合物組み合わせの予in silico予測 |
| 3. 学会等名 第22回日本再生医療学会 |
| 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アプローチ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 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 TRANS DIRE: バ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トリプログラミング |
| 3. 学会等名 第11回生命医薬情報学連合大会 |
| 4. 発表年 2022年 |

| |
|---|
| 1. 発表者名 Nakamura, T., Iwata, M., Hamano, M., Eguchi, R., Takeshita, J., and Yamanishi, Y. |
| 2. 発表標題 Small compound-based direct cell conversion with combinatorial optimization of pathway regulations |
| 3. 学会等名 The 21st European Conference on Computational Biology (ECCB2022) (国際学会) |
| 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, 山西芳裕 |
| 2. 発表標題 トランスオミクス解析によるダイレトリプログラミング誘導転写因子の予測 |
| 3. 学会等名 第21回日本再生医療学会 |
| 4. 発表年 2022年 |

| |
|---|
| 1. 発表者名 濱野桃子, 中村透, 岩田通夫, 江口凌平, 山西芳裕 |
| 2. 発表標題 トランスクリプトームデータを用いたダイレ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予測 |
| 3. 学会等名 第10回生命医薬情報学連合大会 |
| 4. 発表年 2021年 |

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

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 |
|-----------|---|--------------------------------------|----|
| 研究 分担者 | 鈴木 淳史 (Suzuki Atsushi) (30415195) | 九州大学・生体防御医学研究所・教授 (17102) | |
| 研究 分担者 | 味八木 茂 (Miyaki Shigeru) (10392490) | 広島大学・病院(医)・特定教授 (15401) | |

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 |
|---------|---------|
|---------|---------|