

1 細胞レベルで薬の作用メカニズムを同定する新たな情報技術を開発 -テンソル補完アルゴリズムの創薬応用-

九州工業大学大学院情報工学研究院の山西芳裕教授らの研究グループは、1細胞レベルでの遺伝子発現データにおける欠損値や未観測値を補完するアルゴリズムを提案し、細胞タイプ特異的な薬の作用メカニズムを同定する新たな情報技術を開発しました。

(ポイント)

- 多様な細胞が混在した細胞集団に対する薬効ではなく、1つ1つの細胞に対する薬効を明らかにするための高深度かつ高精度なデータ解析技術を開発した。
- 1細胞レベルでの遺伝子発現データは欠損値や未観測値が多いため、それらを補完するテンソル^{注1}分解アルゴリズムを提案した。
- 細胞タイプ特異的な薬の作用メカニズムを解明することにより、層別化された患者に対する薬剤選択の精度向上や、より安全性や信頼性の高い精密医療の実現に貢献する。

本研究グループは、1つ1つの細胞に対する薬効を明らかにするための高深度かつ高精度なデータ解析技術を開発しました。これまでの創薬においては薬に対する細胞応答の不均一性は考慮されておらず、多様な細胞が混在した細胞集団に対して薬効を見出すことが一般的であり、1細胞レベルでの薬効のメカニズムを理解するのは困難でした。また、1細胞レベルでの遺伝子発現データには、測定技術の限界に起因する欠損値や未観測値が非常に多いという問題がありました。

本研究では、1細胞レベルでの遺伝子発現データ解析において障害となっていた欠損値や未観測値を、高深度かつ高精度に補完するテンソル分解アルゴリズムを提案し、さらに1細胞レベルでの薬の作用メカニズムを生物学的パスウェイの視点から同定する点が特色です。開発手法を、膵臓内の膵島などで観測された1細胞レベルでの化合物応答遺伝子発現データに適用し、細胞タイプ特異的な薬の作用メカニズムの同定に成功しました。

本研究により、1細胞レベルで薬の作用メカニズムを同定する高深度かつ高精度な新たなアプローチを提示することでの、薬に対する細胞応答の不均一性を考慮した医薬品開発や精密医療の展開が期待されます。

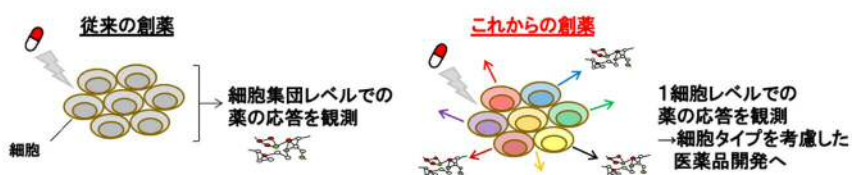
本研究成果は、2022年11月25日午前1:00（日本時間）に英国科学誌「Nature Computational Science」のオンライン版で公開されました。研究の詳細は別紙をご参照ください。

【お問い合わせ先】

九州工業大学大学院情報工学研究院
生命情報工学研究系 教授 山西芳裕
yamani@bio.kyutech.ac.jp
0948-29-7821

1細胞レベルで薬の作用メカニズムを同定する新たな情報技術を開発 -テンソル補完アルゴリズムの創薬応用-

背景



これまでの課題

- ✓ 1細胞レベルでの遺伝子発現データは欠損値(未観測値)が多い

解決方法

- ✓ テンソル補完技術を用いて欠損値(未観測値)を補完
※補完: 欠損値の値を補って、欠損値のない完全なデータを再構成すること

成果

- ✓ 1細胞レベルで薬の作用メカニズムを同定可能に

1細胞レベルでの化合物応答遺伝子発現データの予測

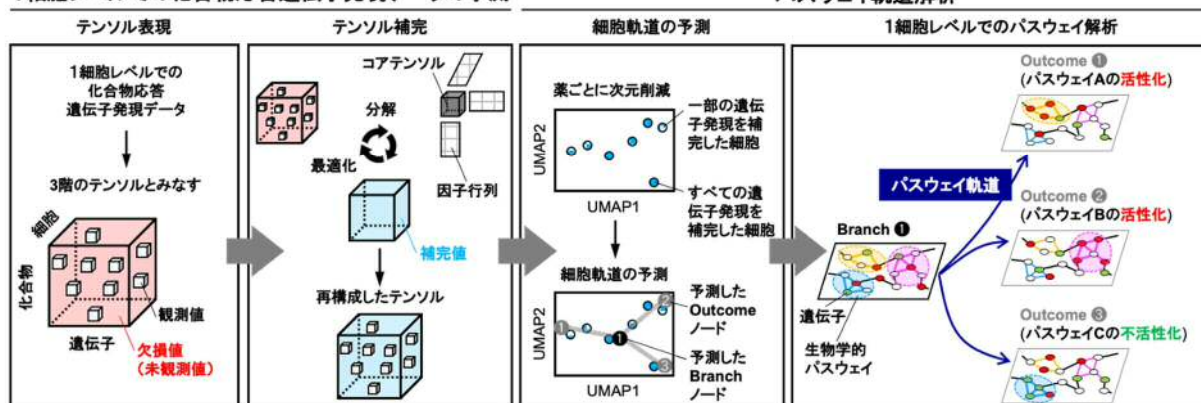


図1：本研究の概要

<研究の背景と経緯>

生命医学分野において大量に生み出されてきた医薬ビッグデータは、データ駆動型の創薬において非常に有用です。特に近年では、1細胞レベルでの測定技術が発展し、1細胞レベルでの生体分子ビッグデータを用いた生命現象の理解が重要視されるようになってきました。創薬現場でも、疾患の発症メカニズムや薬の作用メカニズムを1細胞レベルで理解し、細胞タイプ特異的な薬効や副作用を同定することが期待されています。しかしながら、そのような1細胞レベルでの生体分子ビッグデータを十分に有効活用するための情報基盤技術は整備されていません。

薬を含む生物活性化合物を細胞に添加して得られる化合物応答遺伝子発現データを解析することは、薬の作用メカニズムを明らかにする上で非常に有用です。このデータは細胞内の遺伝子の発現量が、化合物の添加によってどのように変化したのかを示しています。実際に、化合物応答遺伝子発現データはさまざまな薬の作用メカニズムを予測するために用いられてきました。しかし、応答遺伝子発現データは、通常、さまざまな特徴をもつ多様な細胞が混在した集団(バルク)に対して測定されるのが一般的であるため、細胞集団平均的な薬の作用を見出すことしかできませんでした。そこで、1細胞レベルでの薬の作用メカニズムを理解し、副作用が少なく薬効の高い薬治療法を開発することが求められています。

精密医療を志向して効果的な治療法を確立するためには、薬に対するそれぞれの細胞応答の不均一性を考慮することが重要です。細胞タイプ特異的な薬の作用メカニズムを理解するためには、1細胞レベルでの化合物応答遺伝子発現データの大規模解析が必要となります。しかしながら、1細胞レベ

ルでの化合物応答遺伝子発現の測定においては、一般的に、実験的な限界（例：細胞のバーコード化の難しさ）により、多くの欠損値（未観測値）が生じ、解析上の障害となっていました。

<研究の内容>

九州工業大学大学院情報工学研究院の山西芳裕教授、岩田通夫准教授らの研究グループは、1細胞レベルでの遺伝子発現データにおける欠損値や未観測値を、高深度かつ高精度に補完するアルゴリズムを提案し、1細胞レベルで薬の作用メカニズムを同定する新たな情報技術を開発しました（**図1**）。

本研究では、1細胞レベルでの化合物応答遺伝子発現データを化合物、遺伝子、細胞などからなるテンソル構造と捉え、テンソル中の欠損値の潜在的な値を予測（補完）し、さらに、化合物の1細胞レベルでの作用メカニズムを生物学的パスウェイ^{注2}情報に基づき同定する手法、TIGERS（tensor-based imputation of gene expression data at the single-cell level）を開発しました。

まず、1細胞レベルでの遺伝子発現データに含まれる欠損値を補完するアルゴリズムの性能を評価するため、観測値を人工的に欠損値とみなし、観測値を正しく再現できるか調べました。性能評価のため、化合物、遺伝子、細胞からなる3階のテンソルを構築しました。性能評価の結果、提案手法は、これまでの手法に比べて、高精度で欠損値を補完することができることを示しました（**図2**）。

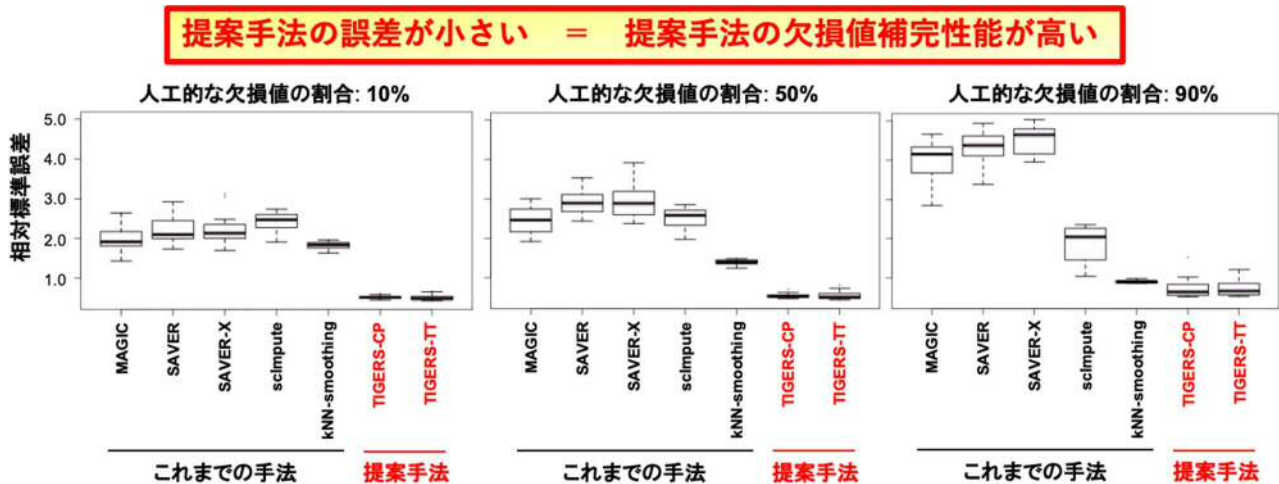
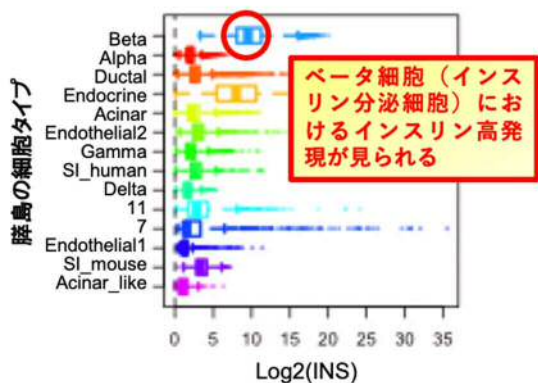


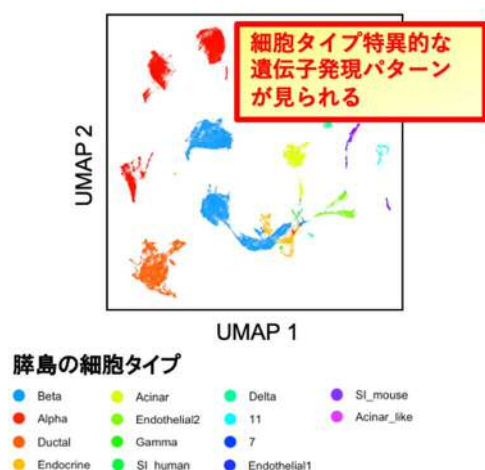
図2：提案手法は、これまでの手法に比べて、欠損値の補完性能が高い

次に、データ補完して得られた細胞タイプごとの遺伝子発現の生物学的妥当性について、臍島のさまざまな細胞タイプに対するマーカー遺伝子の発現値を用いて評価しました。例えばインスリン分泌細胞であるベータ細胞において、インスリン遺伝子が高発現であることを確認しました。これにより、提案手法は細胞のマーカー遺伝子に対して妥当な発現値を正しく予測できることを確認しました。また、次元削減手法であるUMAP（Uniform Manifold Approximation and Projection）で細胞の分布を視覚化し、提案手法は細胞タイプ特異的な遺伝子発現パターンを同定できることを示しました。例として抗マラリア薬であるアルテメテルについて、細胞タイプ特異的に制御している生物学的パスウェイを探索し、その結果、近年報告されている抗がん作用に起因するパスウェイの制御機構を解明することに成功しました（**図3**）。

マーカー遺伝子を用いた評価



次元削減による評価



薬の作用パスウェイを用いた評価

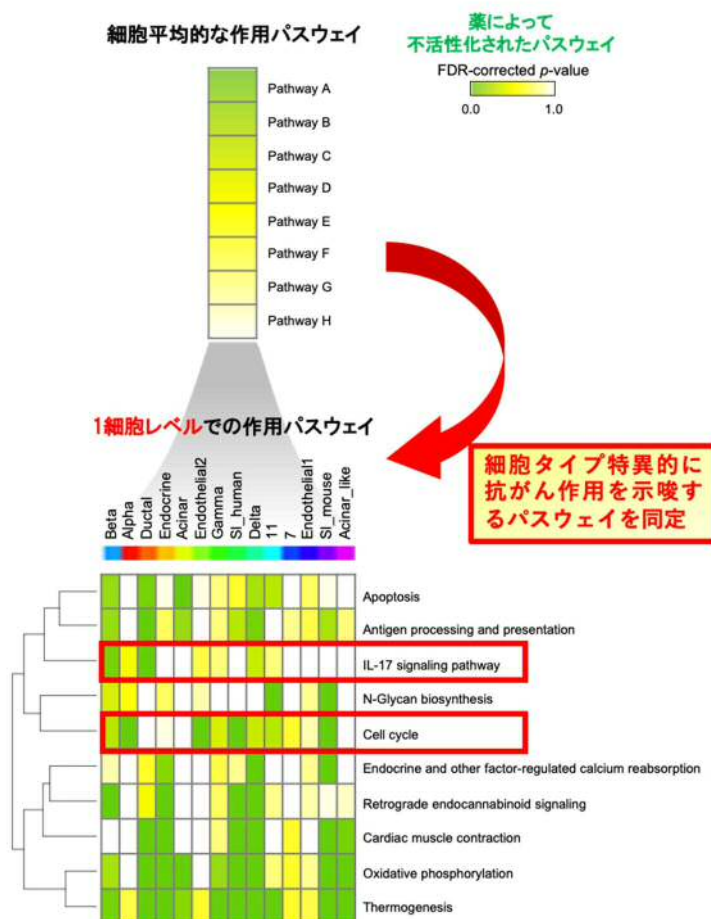


図3：提案手法は、細胞のマーカー遺伝子に対して妥当な発現値、細胞タイプ特異的な遺伝子発現パターン、薬（例：アルテメテル（抗マラリア薬））の潜在的な抗がん作用を予測できる

最後に、データ補完により得られた1細胞レベルでの化合物応答遺伝子発現データに対して、1細胞データ解析の基盤技術である細胞軌道解析^{注3)}を適用し、得られた細胞軌道に沿って、薬の作用メカニズムを生物学的パスウェイの視点から同定しました。これにより、細胞軌道を薬の作用メカニズムの観点から解釈し、薬の制御パスウェイがどのように遷り変わるのか予測できるようになりました（図4）。

<今後の展開>

本研究により、1細胞レベルで薬の作用メカニズムを同定する高深度かつ高精度な新たなアプローチを提示することでの、薬に対する細胞応答の不均一性を考慮した医薬品開発や精密医療の展開が期待されます。また、さまざまな疾患に由来する疾患特異的1細胞遺伝子発現データ、ゲノミクスやエピゲノミクスなどのオミクスデータなど、任意の医薬ビッグデータをテンソル構造として表現することで、疾患の治療標的分子やバイオマーカーの探索にも開発手法は適用可能です。複数の種類の1細胞オミクスデータを統合することで、層別化された患者に対する薬剤選択の精度向上など、より安全性や信頼性の高い精密医療が提供できる可能性があります。さらに、アルゴリズム内のさまざまなパラメータの最適化手法、および、機械学習アルゴリズムを改善して、予測の信頼性を向上させていく予定です。

本研究は、日本学術振興会・科学研究費助成事業・基盤研究(A)における研究課題「医療ビッグデータから難治性疾患の創薬標的を予測する革新的AI手法の開発」（研究代表者：山西芳裕、JSPS KAKENHI Grant Number JP21H04915）の支援を受け行われました。

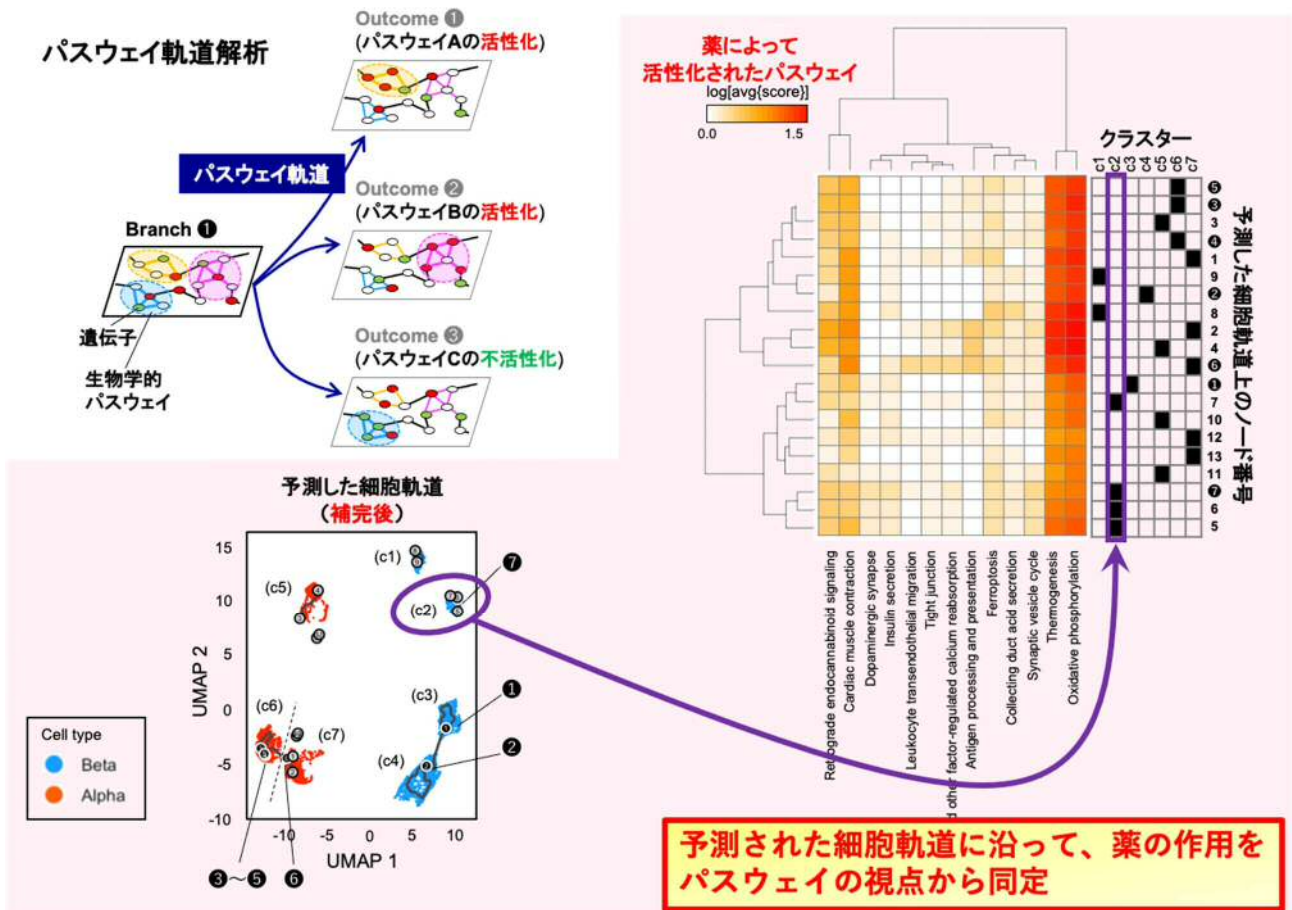


図4：提案手法により補完したデータを用いて予測した細胞軌道に沿って、薬（例：アルテメテル（抗マラリア薬））の作用をパスウェイの視点から同定できる

< 発表雑誌 >

雑誌名：「Nature Computational Science」（オンライン版：2022年11月25日午前1:00日本時間）
 論文タイトル：Pathway trajectory analysis with tensor imputation reveals drug-induced single-cell transcriptomic landscape
 著者：Iwata, M., Mutsumine, H., Nakayama, Y., Suita, N., and Yamanishi, Y.
 DOI番号：10.1038/s43588-022-00352-8

< 用語解説 >

注1) テンソル

数値を多次元の配列に配置したもの。例えば、スカラーは0階のテンソル、ベクトルは1階のテンソルとなる。

注2) パスウェイ

遺伝子やタンパク質などの生体分子からなる生体分子相互作用ネットワークの機能モジュール。

注3) 細胞軌道解析

それぞれの細胞の遺伝子発現パターンから、細胞に対して擬時間を割り当てて、細胞分化の過程などを可視化する解析。