

多様な細胞における薬物応答遺伝子発現を高精度に予測する新技術を開発
 ～医薬ビッグデータの有効活用へ～

九州工業大学、理化学研究所の共同研究により、多様なヒト細胞における薬物応答の遺伝子発現パターンを高精度に予測する新たな情報技術を開発しました（図1）。

多様な細胞における薬物応答遺伝子発現を網羅的に予測する新技術を開発

薬物応答遺伝子発現を解析することの意義

- 薬物が細胞に及ぼす影響を遺伝子発現レベルで評価可能
- 遺伝子発現パターンの類似性から、薬物の類似性を予測可能

これまでの課題

- 薬物応答遺伝子発現は、限られた細胞においてのみ測定
- 欠損値（未観測値）があり、**網羅的な解析が困難**

解決方法

- 薬物応答遺伝子発現データを、薬物、遺伝子、細胞などからなるテンソル構造とみなす
- **テンソル補完技術を使用**して、テンソル中の欠損値を補完

※ここで、「補完」とは、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を意味する。

成果

- 欠損値を補完した薬物応答遺伝子発現データを用いることで薬物の適用可能疾患を高精度に予測可能に！

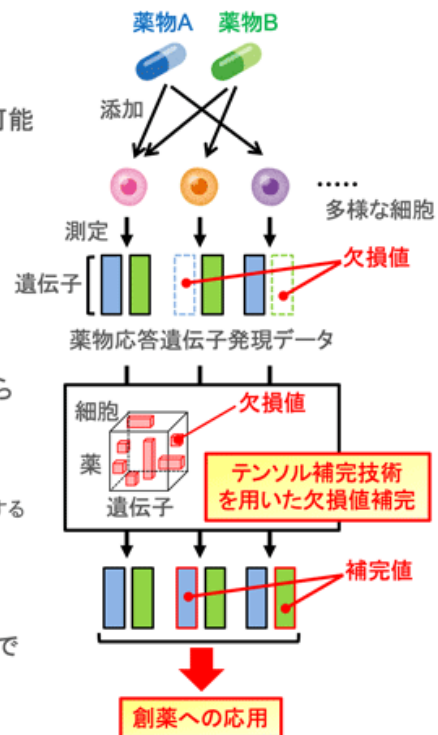


図1 本発表の概要

本研究では、薬物応答遺伝子発現データを、薬物、遺伝子、細胞、時系列からなるテンソル構造と捉え、新規のテンソル分解手法である `Tensor-train weighted optimization` アルゴリズムを用いて、さまざまな細胞における未観測の薬物応答を予測する機械学習手法を開発しました。薬物の効能予測に応用し、開発手法が効能予測の性能の向上に大きく寄与することを示しました。

まず、薬物、細胞、遺伝子からなる3階のテンソル構造である薬物応答遺伝子発現データにおける未観測値予測性能を評価するため、観測値を人工的に欠損値と見なし、観測値を正しく再現できるか調べました。その結果、提案手法は、従来手法に比べて最大約1.5

倍の精度で欠損値を補完することができることを見いだしました（図2）。また、薬物応答遺伝子発現パターンは時間的にも変化するので、薬物、細胞、遺伝子、時系列からなる4階のテンソルを構築し、同様の性能評価を行いました。その結果、提案手法は、特に高階テンソル構造のデータに対する有効性を示しました（従来手法に比べて最大約2倍の精度向上）。

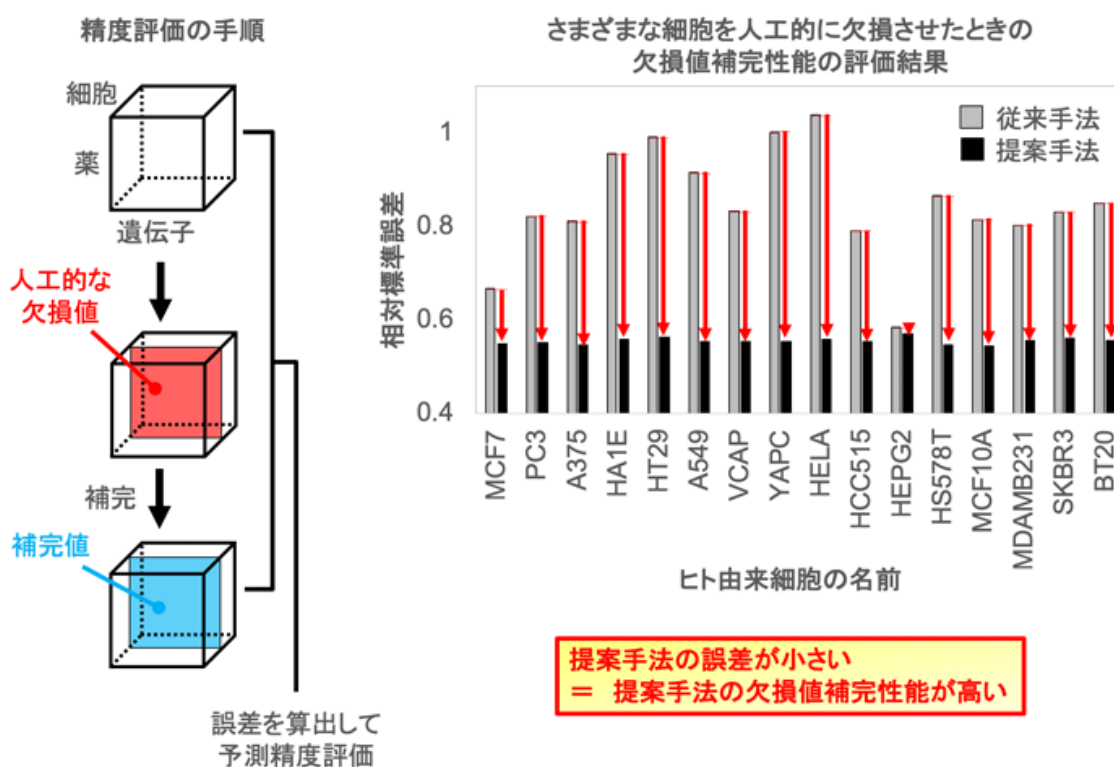


図2 提案手法は、従来手法に比べて、テンソル中の欠損値補完性能が高い

次に、欠損値を補完した薬物応答遺伝子発現データを、薬物の効能の予測に応用しました。これにより、欠損値が多い細胞ほど、効能予測の性能が大きく向上し、最大約20%性能が向上することを示しました（図3）。さらに、1483個の薬物を16種類のヒト細胞に添加して得られた薬物応答遺伝子発現データに提案手法を適用し、駆虫薬であるニコロサミドの成人T細胞白血病に対する効能など、さまざまな疾患に効能があると期待される薬物を選び出しました。そして、その新しく予測された薬物の効能についての妥当性を、近年の文献や臨床報告で確認することができました。

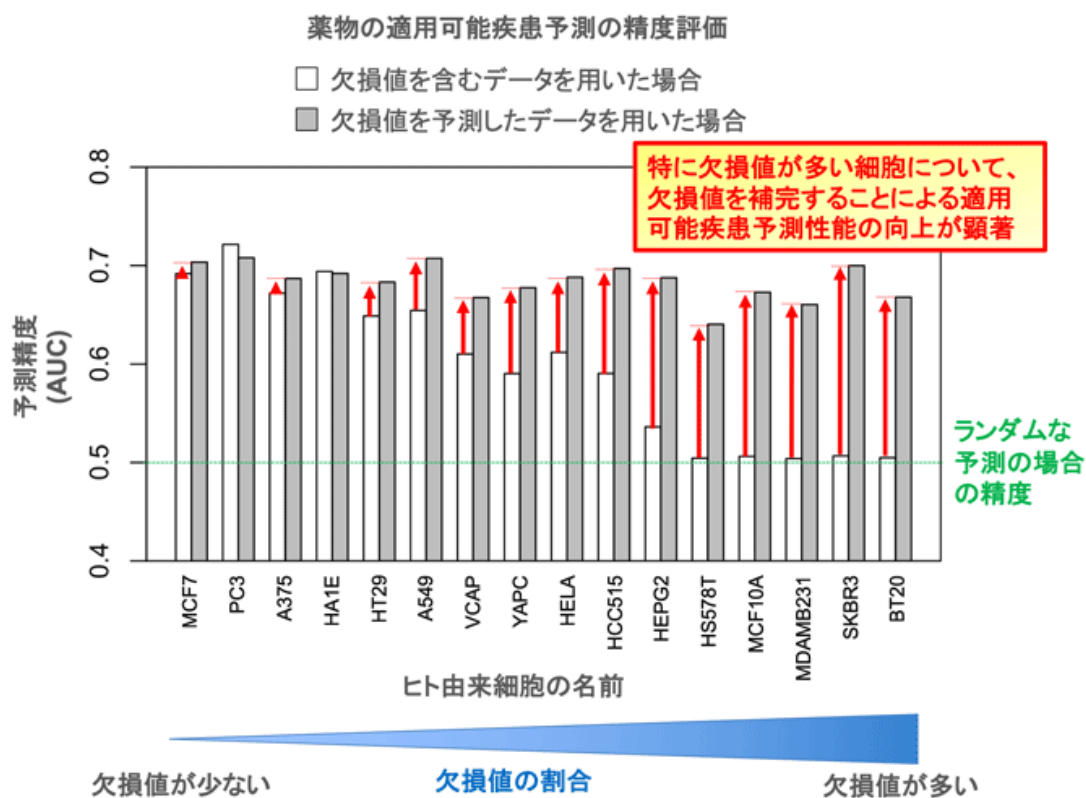


図3 提案手法により欠損値を補完した遺伝子発現データは、創薬においても有用

多様なヒト細胞における薬物応答を予測する新たなアプローチを提示することにより、薬物の作用メカニズム解明や薬効予測、ツール化合物の探索など、創薬現場での医薬品開発の促進が期待されます。今後は、アルゴリズムを改善して予測の信頼性や計算効率を向上させていくとともに、既承認薬だけでなく合成化合物や天然化合物などあらゆる化合物にも解析対象を広げ、さまざまな疾患に対する医薬品候補化合物を探索していく予定です。

(日文新聞发布全文 <https://www.jst.go.jp/pr/announce/20190708/index.html>)

文：JST 客观日本编辑部翻译整理