

分野: 生命科学・医学系

キーワード: 単一細胞 RNA シーケンス解析、免疫細胞分類、T 細胞亜集団、精密医療、SDGS

＼階層的分類システム×機械学習／

## 新技術 scODIN により 細胞の高速・高精度解析を実現

### 【研究成果のポイント】

- ◆ 大規模な単一細胞 RNA シーケンスデータから免疫細胞を高速・高精度で自動分類する解析手法「scODIN」を開発し、膨大な数の細胞データでも迅速に、かつ高い精度で分類することが可能に
- ◆ これまで既存の自動分類手法では処理速度と精度に限界があり、希少細胞や移行期の細胞は「未分類」として見逃されていたが、階層的分類システムと機械学習を組み合わせた独自アルゴリズムにより克服し、65 万細胞を 6 分で高精度に分類
- ◆ あらゆる疾患の免疫細胞解析に汎用的に適用可能で、創薬標的の探索、治療効果のモニタリング、病態メカニズムの解明など、精密医療の実現に向けた基盤技術としての応用に期待

### ❖ 概要

大阪大学感染症総合教育研究拠点の Jonas Søndergaard 招へい研究員、James Badger Wing 教授らの研究グループは、単一細胞 RNA シーケンス解析<sup>\*1</sup>において、新たな解析手法「scODIN」を開発しました。

「scODIN」は、階層的分類システム<sup>\*2</sup>と機械学習を融合した手法で、single-cell RNA-seq データにおける名前の最適化された検出と推論(optimized detection and inference of names in scRNA-seq data)を可能にします。

これにより、これまで見逃されていた希少細胞<sup>\*3</sup>や移行期細胞<sup>\*4</sup>を高速・高精度で自動分類できることを世界で初めて明らかにしました。本手法は、数十万個の細胞を含む大規模データをわずか数分で処理でき、従来の手法では困難だった細胞状態の連続的な変化を正確に捉えることが可能です。

これまで単一細胞 RNA シーケンス解析では、処理時間と分類精度がトレードオフの関係にあると考えられており、希少細胞や中間状態の細胞の同定については解明されていませんでした。今回、研究グループは独自の階層的分類システム(Tier system)と機械学習(k 近傍法)を組み合わせることにより、65 万個の細胞をわずか6分で処理しながら、従来法を大幅に上回る分類精度を達成できることを解明しました。この成果により、疾患の病態解明や創薬標的の探索、個別化医療の実現が期待されます。

本研究成果は、2025 年 8 月 21 日に米国科学誌「The Journal of Immunology」(オンライン)に掲載されました。

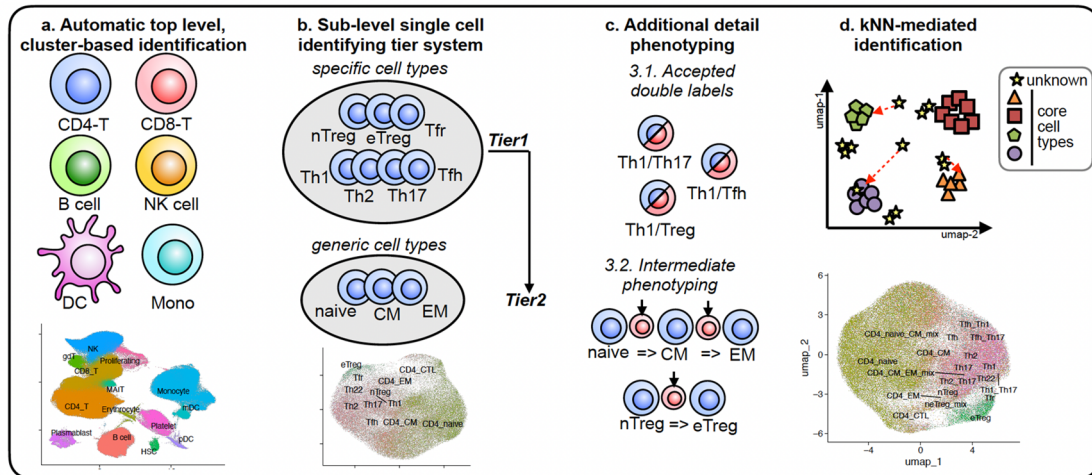


図1 scODIN (Optimized Detection and Inference of Names in scRNA-seq data) の概要

- a, 主要な細胞クラスター(CD4 T細胞、B細胞、単球)を自動で高速かつ正確に同定し、さらなる詳細な分類を行う。
- b, 階層的な分類システムにより、詳細なレベルで細胞を同定(Tier 1で制御性T細胞(Treg)やヘルパーT細胞(Th細胞)などのサブセットを識別し、Tier 2でセントラルメモリー(CM)やエフェクターメモリー(EM)などの細胞型を識別)。
- c, さらに二重標識ラベルを追加することで、移行期細胞を同定。
- d, 最後のステップではk近傍法(kNN)推論を用いることで、すべての細胞をカテゴリー化されます。

#### [Wing 教授のコメント]

免疫細胞の多様性を正確に理解することは、疾患の本質的な理解に不可欠です。scODINの開発で最も苦労したのは、計算速度と精度の両立でした。従来は「速いが粗い」か「精密だが遅い」の二択でしたが、階層的アプローチにより両方を実現できました。この技術が、研究者の方々の新たな発見を加速し、最終的には患者さんの治療改善につながることを期待しています。特に、これまで見逃されていた希少な細胞集団が、実は重要な役割を果たしている可能性があり、そうした発見の扉を開く技術になれば幸いです。

#### ❖ 研究の背景

これまで、単一細胞 RNA シーケンス解析は、個々の細胞の遺伝子発現を網羅的に測定できる強力な技術であることが知られていました。しかし、データが大規模化するにつれ、細胞の自動分類において処理速度と分類精度のトレードオフ、希少細胞の見逃し、移行期細胞の誤分類などの課題がありました。特に免疫細胞は多様な亜集団を持ち、疾患により微妙に変化するため、詳細かつ正確な分類が求められていました。

#### ❖ 研究の内容

研究グループでは、階層的な分類システム(Tier system)と機械学習(k近傍法)を組み合わせた独自の方法により、大規模な単一細胞 RNA シーケンスデータでも高速・高精度な細胞分類を可能にする「scODIN」を開発しました。これは、まず主要な細胞タイプ进行分类(Tier1)した後、さらに詳細な亜集団を同定(Tier2)する二段階アプローチで、従来法では困難だった希少細胞(Th22細胞など)や移行期細胞(ナイーブ→メモリー移行中など)の正確な同定を実現しました。

特筆すべき点として、本手法は約65万細胞を含む大規模データセットをわずか6分で処理でき、従来法で数時間を要していた処理時間を劇的に短縮しました。

さらに、COVID-19患者データへの適用により、疾患特異的な免疫細胞の変化を詳細に解析することを実現し、臨床応用への可能性を示しました。

## ❖ 本研究成果が社会に与える影響(本研究成果の意義)

本研究成果により、あらゆる疾患の免疫応答を詳細に解析することが可能となり、病態メカニズムの解明や新規治療標的の発見が期待されます。特に、がん免疫療法の効果予測、自己免疫疾患の早期診断、感染症における重症化予測など、個別化医療の実現に貢献することが期待されます。また、本手法はオープンソースとして公開予定であり、世界中の研究者が利用可能となることで、免疫学研究全体の加速に寄与すると考えられます。

## ❖ 特記事項

本研究成果は、2025 年 8 月 21 日に米国科学誌「The Journal of Immunology」(オンライン)に掲載されました。

タイトル: “Optimized Detection and Inference of immune cell type Names in scRNA-seq data”

著者名: Janyerkye Tulyeu, David Priest, James B. Wing\*, Jonas Nørskov Søndergaard\*

DOI: <https://doi.org/10.1093/jimmun/vkaf183>

## ❖ 用語説明

### ※1 単一細胞 RNA シーケンス解析

個々の細胞から抽出した RNA(遺伝子の発現情報)を網羅的に測定する技術。従来の解析では細胞集団の平均値しか得られなかったが、この技術により一つの細胞の特徴を詳細に調べることが可能。

### ※2 階層的分類システム

Tier system. scODIN で採用された段階的な細胞分類手法。第 1 階層(Tier1)で主要な細胞タイプ(T 細胞、B 細胞など)进行分类し、第 2 階層(Tier2)でより詳細な亜集団(Th1、Th2、Th17 など)を同定する。

### ※3 希少細胞

全体の細胞集団の中で非常に少ない割合(1%未満など)でしか存在しない細胞。Th22 細胞や Tfr 細胞など、重要な機能を持つにも関わらず、従来の解析手法では見逃されることが多かった。

### ※4 移行期細胞

ある細胞状態から別の細胞状態へ変化している途中の細胞。例えば、ナイーブ T 細胞からメモリー T 細胞への分化過程にある細胞など。連続的な変化の中間状態を捉えることで、細胞の運命決定メカニズムの理解につながる。

## ❖ SDGs目標



## ❖ 参考 URL

James Badger Wing 教授

<https://www.cider.osaka-u.ac.jp/researchers/james-wing/>