

令和 元年 10月 21日
国立大学法人浜松医科大学
国立研究開発法人 日本医療研究開発機構

機械学習を用いた質量顕微鏡解析の自動化手法の 開発と小脳に限局して分布する分子集団の発見

<概要>

浜松医科大学・細胞分子解剖学講座・国際マスイメージングセンターの南平眞理君（学部6年生）、堀川特任助教、瀬藤教授らは、機械学習（注1）を利用した質量顕微鏡解析の自動化法を開発し、この手法を用いて小脳に限局した分布を持つ分子集団を発見することに成功しました。この成果は、国際学術誌『Scientific Reports』に掲載されました（日本時間9月13日から、Online版で公開されました）。

<研究の背景>

質量顕微鏡法は質量分析（注2）を応用した解析法の一つであり、多数の生体分子に関して組織や細胞でそれぞれどのような分布を示すかを、一度の測定で解析できる方法です。研究グループはこの方法を用いて、これまでにがん組織やアルツハイマー病患者の脳において増加または減少する、病気の原因と推定される生体分子を数多く発見してきました。これまでの測定では飛行時間型質量顕微鏡装置（注3）を用いており、観察する分子をあらかじめ決めておく解析（ターゲット解析）が主流で、一度に測定できる分子の分布も数十個程度と、手動および目視による解析も可能でした。しかし、近年利用が進みつつあるフーリエ変換イオンサイクロトロン共鳴質量顕微鏡装置：FT-ICR-IMS（注4）を用いた解析では、一度に1000を超える分子の分布が観察できるため手動や目視による解析は困難で、このデータ解析が測定全体のボトルネックとなっており、膨大な分布情報を迅速かつ簡易に解析する手法の開発ニーズが高まりつつありました。そこで、研究グループは機械学習を利用して質量顕微鏡法によるデータ解析の自動化を試みました（図1）。

<研究の成果>

研究グループは、まずFT-ICR-IMS（Solarix, Bruker Daltonics社）を用いて、ラット脳の矢状断（縦切り断面）における約500個の生体分子の分布データを取得しました。このデータに対して、教師無しの機械学習である主成分分析（注5）を行い、ラット脳の生体分子の多くが灰白質または白質のいずれかに特異的な分布を示すことが明らかになりました（図2左）。この結果は、生体分子の分布データを事前に選別しなくても、灰白質または白質といった脳の解剖学的な構造を自動的に再構築できたことを示しています。さらに、同じデータに対して階層クラスタリング（注6）を用いて脳の生体分子の分布パターンを約10種類に大別し、その中には灰白質や白質に特異的な分子の分布だけではなく、新たに小脳灰白質のみに分布するリン脂質の分子集団も発見することができました（図2右）。この結果は、階層クラスタリングの解析により自動的かつ機械的に分布パターンを分類でき、さらにこれまでに知られていない分布を示す分子集団を発見できる可能性を示しています。

＜今後の展開＞

今回の研究成果は、FT-ICR-IMS を用いた測定により取得される膨大な生体分子の分布データを、機械学習により自動的・機械的・迅速に解析できることを示しており、機械学習は質量顕微鏡による解析のボトルネックを解消できることを強く示唆しています。また、従来の目視による主観的な判断によるところが大きかった質量顕微鏡の解析を、機械学習を用いてより客観的・ノンバイアスな解析手法へと昇華できることが期待されます。これにより、従来の解析手法と比べてより早くより多くの病変組織を質量顕微鏡で解析することが可能になり、さらに、これまでは検出が困難であった未知の病因分子の発見にもつながることが期待されます。

＜用語解説＞

1. 機械学習：サイズの大きいデータ（ビックデータ）をコンピュータに学習させることで、データが持つ特徴を自動的に発見させるデータ解析手法の総称。
2. 質量分析：分子の質量を測定してその分子が何であるか分析する手法。
3. 飛行時間型質量顕微鏡装置：質量顕微鏡の解析に最適化された質量分析計であり、生体小分子からタンパク質までの幅広い範囲で、分子の質量分布を分析できる装置。
4. フーリエ変換イオンサイクロトロン共鳴質量顕微鏡装置：飛行時間型質量顕微鏡装置に比べて、より精密かつ高感度に分子の質量を分析できる質量顕微鏡装置。
5. 主成分分析：データが持つ特徴・情報を集約する分析法であり、人による主観的な分析が難しいビックデータの分析などに利用される。
6. 階層クラスタリング：データをその類似性でグループ分けする手法。この研究では分子分布が似ているもの同士をグループ化するために利用。

＜発表雑誌＞

Scientific Reports:

Scientific Reports は、自然科学と臨床科学のあらゆる領域を対象としたオープンアクセスの査読付き電子ジャーナルです。

＜論文タイトル＞

Unsupervised machine learning using an imaging mass spectrometry dataset automatically reassembles grey and white matter

Doi: 10.1038/s41598-019-49819-1

＜著者＞ (* equal contribution, + 研究責任者)

Makoto Nampei *, Makoto Horikawa *, Keisuke Ishizu, Fumiyoshi Yamazaki, Hidemoto Yamada, Tomoaki Kahyo, Mitsutoshi Setou +

(南平 眞理 *, 堀川 誠 *, 石津 啓介, 山崎 文義, 山田 秀元, 華表 友暁, 瀬藤 光利 +)

＜研究支援＞

本研究は浜松医科大学細胞分子解剖学講座および国際マスイメージングセンターの研究で、国立研究開発法人日本医療研究開発機構（AMED）の革新的先端研究開発支援事業（AMED-CREST）「画期的医薬品等の創出をめざす脂質の生理活性と機能の解明」研究開発領域（研究開発総括：横山信治）における研究開発課題「光による脂質の同定制御観察技術すなわちオプトリポミクスの創生」（研究開発代表者：瀬藤光利）、新学術領域研究

「リポクオリティ」(科研費)、共用プラットフォーム事業「原子・分子の顕微イメージングプラットフォーム」(JST)からの研究費を受けて得られた研究成果です。

<本件に関するお問い合わせ先>

国立大学法人浜松医科大学・細胞分子解剖学講座
国際マスイメージングセンター
〒431-3192 浜松市東区半田山 1-20-1
教授 瀬藤光利
Tel: 053-435-2086/FAX: 053-435-2468
E-mail: setou@hama-med.ac.jp

<AMED 事業に関するお問い合わせ先 >

国立研究開発法人日本医療研究開発機構 (AMED) 基盤研究事業部 研究企画課
〒100-0004 東京都千代田区大手町 1-7-1 読売新聞ビル
TEL : 03-6870-2224 E-mail : kenkyuk-ask@amed.go.jp

<参考図>

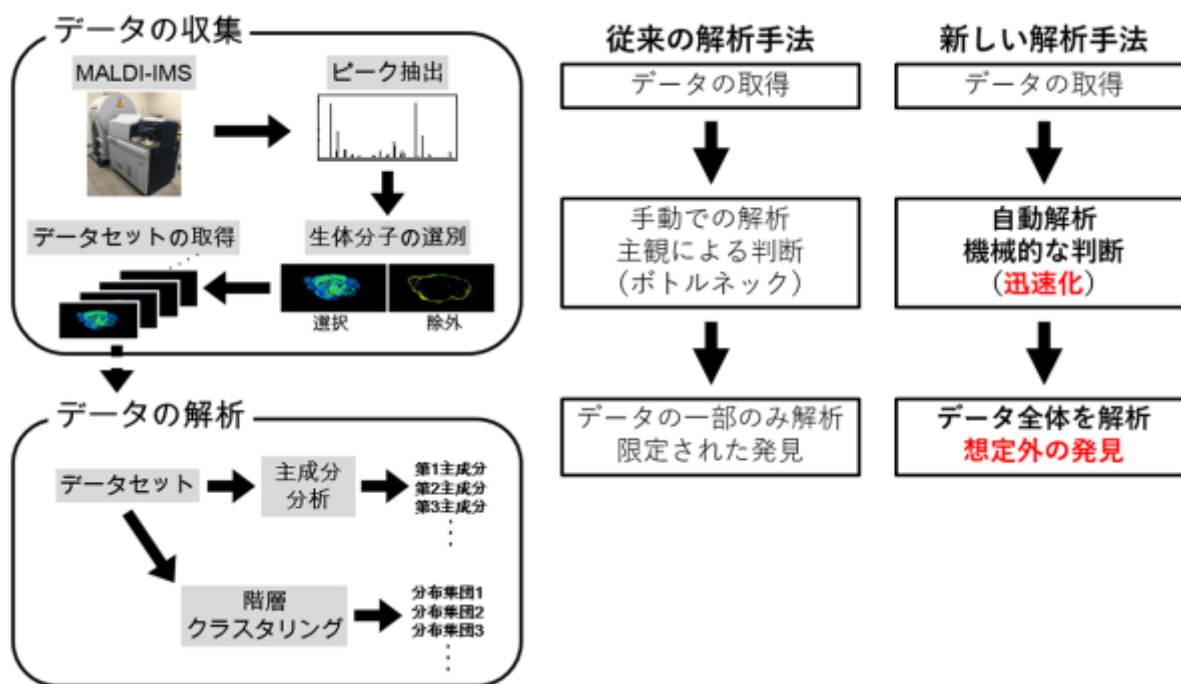


図1. 今回の解析手法のワークフローと従来法との比較

質量顕微鏡装置によって測定された様々な分子の分布から、生体由来の分子の分布を選別してデータセットを得る (左上)。このデータセットに対して主成分分析を行う事で、組織における生体分子の分布の特徴が得られる (左下)。さらに、階層クラスタリングにより類似した分布を持つ分子同士が自動的にグループ化できる (左下)。この手法により、従来法と比較して解析の迅速化が可能であり、想定外の分布を持つ分子を発見することも可能である (右)。

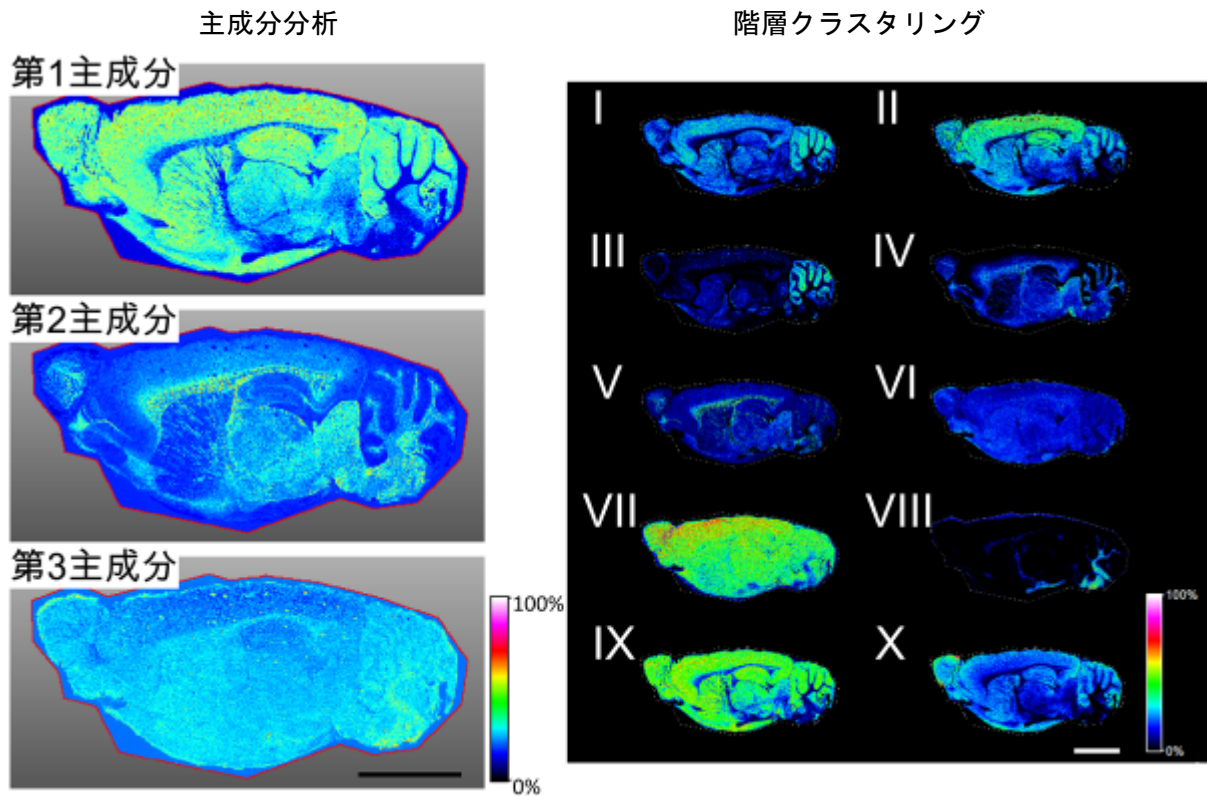


図2. 主成分分析および階層クラスタリング

主成分分析により第1主成分として脳の灰白質（細胞体が集まる領域）に対応する分布、第2主成分として白質（神経線維が集まる領域）に対応する分布が自動的に再構築された（左）。階層クラスタリングにより脳の分子の分布パターンが概ね10種類程度に分類でき、さらにその中に小脳に限局して分布するリン脂質のグループ（III）を発見につながった（右）。