

親水性代謝物の包括的構造推定法の開発

鳥越, 大平

<https://hdl.handle.net/2324/7182317>

出版情報 : Kyushu University, 2023, 博士 (システム生命科学), 課程博士

バージョン :

権利関係 : Public access to the fulltext file is restricted for unavoidable reason (3)



氏 名 : 鳥越 大平

論文名 : 親水性代謝物の包括的構造推定法の開発

区 分 : 甲

論文内容の要旨

近年の高分解能タンデム質量分析 (high-resolution tandem mass spectrometry, HRMS/MS) 技術の発展によって、血液や生体組織などの試料から数千種類の分子量関連イオンピークを検出できるようになった。しかしながら、この中から同定される代謝物は約 20%未満に過ぎず、残された約 80%の情報は多くの場合活用されてこなかった。これらの標準品が存在しない未知代謝物ピークの構造アノテーション法として、標準品の HRMS/MS スペクトルとサンプルの HRMS/MS スペクトルの類似性に基づく化合物クラスのアノテーション法 (部分構造推定) や構造候補の *in silico* MS/MS 予測スペクトルへの照合による構造アノテーション法が開発されてきた。現行の構造アノテーション法は、多数の偽陽性構造を含む場合がある。これらの偽陽性構造は、構造異性体の関係にある場合が多く、ブリカーサーイオンの精密質量と HRMS/MS スペクトルによる解析だけでは不十分である。LC/HRMS/MS を用いたノンターゲットメタボローム解析において、未知代謝物ピークの実測保持時間と構造候補の予測保持時間の照合によって、代謝物の化学的性質に基づき偽陽性構造を削減することができる。しかし、幅広い親水性代謝物を網羅的かつ一斉に分離検出できる親水性相互作用/陰イオン交換クロマトグラフィー (unified-hydrophilic-interaction/anion-exchange liquid chromatography, unified-HILIC/AEX) は、親水性相互作用 (HILIC) と陰イオン交換 (AEX) の分離モードを段階的かつ連続的に作用させる複雑な分離のため、本手法による代謝物ピークの保持時間予測には至っていなかった。さらに、現行の構造アノテーション方法における構造推定の網羅性は、構造候補を収録したデータベースの化学空間に依存するため、未知代謝物の構造候補を *in silico* で算出することによって化学空間を拡充した新しいデータベースを開発する必要がある。しかしながら、親水性未知代謝物に着目した網羅的な構造候補の算出はほとんど実施されていないのが現状である。そこで、本研究では、親水性代謝物の包括的構造推定法の開発を目的とした。

第二章では、unified-HILIC/AEX の保持時間予測モデルを構築した。トレーニングデータ (203 種類の極性代謝物) とその分子記述子 (12,420 種類) に対し、ランダムフォレストを用いて特徴量重要度をランク付けした。特徴量重要度に基づいて選出した 26 種類の分子記述子を用いて保持時間予測モデルを構築した。テストデータ (51 種類の極性代謝物) を用いて保持時間予測モデルの精度を評価した結果、 ΔRT (実測保持時間と予測保持時間の差) の 86.3%が ± 1.50 分以内、平均絶対誤差は 0.80 分であったことから、高い保持時間予測精度を示した。開発した保持時間予測モデルと *in silico* MS/MS 予測を用いて、グローバル品質管理 (global quality control, QC) 血漿として使用されている NIST SRM 1950 ヒト血漿サンプルから取得したノンターゲットメタボロームデータを解析した結果、標準物質に基づいて同定された 62 種類の既知代謝物に加え、216 種類の極性代謝物の構造アノテーションに成功した。本章において提案された保持時間予測モデルは、メタボローム研究における重要な課題である未知の親水性代謝物の構造アノテーションを加速することに役立つと考え

られる。

第三章では、(1) 安定同位体標識、(2) unified-HILIC/AEX/HRMS/MS 分析、(3) 各種データマイニング技術 (第二章の unified-HILIC/AEX 保持時間予測モデルを含む)、(4) *in silico* エピメタボライトデータベース (*in silico* epimetabolite database, IEMDB) の 4 つの技術を組み合わせた方法論を開発し、内因性の親水性未知代謝物の包括的構造推定法の開発に取り組んだ。最終的に、大腸菌抽出試料のノンターゲットメタボローム解析データを用いて本研究の方法論の実用性を評価した。その結果、炭素数および窒素数を決定した 2,021 種類の代謝特性を検出し、プリカーサーイオンと HRMS/MS スペクトルの情報から 90 種類の代謝特性に構造候補を割り当てた。さらに、各代謝特性の実測保持時間と構造候補の予測保持時間の照合によって、44 種類のエピメタボライトの帰属に成功した。これらの新規エピメタボライトは、安定同位体標識とペアードピークスフィルタリングによって、炭素数と窒素数が決定されており、その組成式が高精度に予測されている。新規エピメタボライト候補は、既知の代謝物と代謝反応によって算出されていることから、その信頼性は高く、当該 *in silico* 代謝反応情報は生化学的解析に利用することが可能である。さらに、ペアードピークスフィルタリングと組み合わせた *in silico* MS/MS 予測によって、高品質なプリカーサーイオンに由来するプロダクトイオンを最大限活用した構造アノテーションを行っており、高精度な保持時間予測によって偽陽性エピメタボライト候補が低減されている。したがって、本章の方法は、内因性の親水性未知代謝物の包括的かつ高精度な構造アノテーションに有用であることが示された。

本博士論文では、親水性代謝物の高精度かつ包括的な構造アノテーション方法を開発した。今後、本研究の方法を応用することで未知代謝物の構造アノテーションが推進され、メタボロミクスの発展の一助になることが期待される。