

平成30年6月8日現在

機関番号：82401

研究種目：基盤研究(C) (一般)

研究期間：2015～2017

課題番号：15K01812

研究課題名(和文) LC/MS/MSを用いた次世代メタボロミクス研究のための基盤技術開発

研究課題名(英文) Development of next generation untargeted metabolomics methodologies using LC-MS/MS

研究代表者

津川 裕司 (Tsugawa, Hiroshi)

国立研究開発法人理化学研究所・環境資源科学研究センター・研究員

研究者番号：30647235

交付決定額(研究期間全体)：(直接経費) 3,600,000円

研究成果の概要(和文)：近年、遺伝子・タンパク質から生成される代謝産物(メタボローム)が生体の恒常性維持に深く関わっていることを強く示唆する報告が相次いでおり、質量分析を用いたノンターゲットメタボローム解析の重要性が高まってきた。しかしながら、質量分析より得られる複雑なビッグデータを解析するための手法、および得られる質量スペクトル情報を紐解き化合物構造を推定する手法が開発されてこなかったため、得られる代謝物情報が極端に少ない(数十種類)という問題があった。本研究により開発されたデータ解析ソフトウェアにより円滑な解析を可能にただけでなく、1000種類を超える化合物種の一斉解析が可能となった。

研究成果の概要(英文)：There are a lot of reports providing that the 'metabolome' is definitely relevant to the biological functions and physiological phenotypes in living organisms. The mass spectrometry based untargeted metabolomics is the popular platform to profile the metabolites, which can be contributed to the discovery of novel metabolite-phenotype relationships. However, the methodology smoothly decoding the complicated 'big data' of mass spectrometer has not been reported yet, and unfortunately, only several dozen metabolites are grasped in metabolomics due to the lack of mass spectrum annotation technology. New methodologies developed in this study enables us to easily and precisely perform the data processing and to capture >1000 metabolites from the biological samples.

研究分野：メタボロミクス

キーワード：質量分析 インフォマティクス 化合物同定 ケムインフォマティクス バイオインフォマティクス

### 1. 研究開始当初の背景

セントラルドグマは生命活動の「流れ」であり、メタボロミクスで扱う低分子代謝物はその「結果」として位置づけられる。そのため、生物・医学研究におけるメタボロミクス研究の役割の1つとして「マーカー探索」が挙げられる。しかしながら未だに、分析データから得られた代謝物由来イオンを網羅的に同定するための技術基盤は確立されておらず、検出された代謝物由来イオンの2-3割しか同定できないのが現状である。バイオロジーの理解を深めるためには、質量分析より得られる「質量情報(イオン情報)」を「代謝物の構造情報」に変換することは必須であり、質量分析データを理解することは代謝を理解することに繋がるといっても過言ではない。また近年、炎症の惹起や収束に関わる新たな脂質メディエーターや癌特異的に蓄積するオンコメタボライトが、質量分析を用いた「ノンターゲット解析」により見出されてきたという事実から、メタボロミクスの重要性が更に高まってきている。

### 2. 研究の目的

本研究では、質量分析より得られる複雑なビッグデータから、「網羅的に」代謝物情報を抽出するためのデータ処理プログラムの構築を目指す。また、質量分析装置における代謝物の構造決定は、化合物イオン構造特異的に観測される「マスフラグメンテーション」と、それから得られる「マススペクトル」の解釈が、最も重要である。これまでは、代謝物の同定は、標準品化合物を購入し、そのリファレンスマススペクトルと実測値を比較することで行われてきたが、この方法では、「標準品が購入可能なもの」しか化合物同定できないといった問題がある。そこで、マススペクトルを紐解くことで代謝物構造を推定するためのプログラム開発を行うことで、低分子化合物の網羅的解析を目指すメタボロミクス研究を革新的に推進させるための「次世代メタボロミクス技術基盤」を本研究により確立することを目指す。

### 3. 研究の方法

(1) ガスクロマトグラフィー質量分析(GC-MS)、液体クロマトグラフィー質量分析(LC-MS)、液体クロマトグラフィータンデム質量分析(LC-MS/MS)、そして近年開発されてきた data independent MS/MS acquisition (LC-DIA-MS) など、様々なタイプの分析手法から得られる生データを統合的に解析するためのアルゴリズム(コア)開発を行う。

(2) 本法の実用性・汎用性を高めるため化合物同定、定量、並びに統計解析に供するデータ行列(サンプル名×化合物名×定量値)作成のための自動データ処理、及び円滑なデータ整理と結果の直感的解釈が可能なユーザーフレンドリーなグラフィカルユーザー

インターフェイス(GUI)を搭載したソフトウェア開発を行う。

(3) 脂質などの体系化可能な分子種については理論マススペクトルを構築する。また、化合物構造とそこから得られるマスフラグメンテーションを解釈するための「理論」を導き出し、それをを用いた構造推定プログラムの構築を行う。

### 4. 研究成果

以下に、3つの主な研究成果を記載する。

【質量分析データの統合解析ソフトウェア MS-DIAL の開発と、理論マススペクトル構築の最初の研究成果 (Nature Methods, 12, 523-526, 2015)】

液体クロマトグラフィータンデム質量分析装置(LC-MS/MS)で化合物を同定するためには、MS/MS スペクトルの情報が必要不可欠である。しかし、従来の LC-MS/MS では走査スピードに問題があり、すべての代謝物由来イオンについて MS/MS スペクトルを取得することは困難であった。つまり当時、化合物同定の網羅性の欠如は、「質量分析から得られる情報量」に起因していた。当該研究では、検出される全ての MS/MS スペクトルを取得する技術である Data Independent MS/MS Acquisition 法(DIA法)に着目し、研究を行った。DIA法は、「理論的には」検出される全 MS/MS スペクトルが取得可能である。しかしながら本手法は、25 - 50 Da 範囲のイオンをまとめてフラグメント化させるため、得られる MS/MS スペクトルは複数の成分に由来する混合スペクトルであるという問題点があり、手法としての実用レベルまで到達していなかった。

そこで本研究では、複数の化合物が重なりあった MS/MS スペクトルを個々の MS/MS スペクトルに分離する解析手法である「デコンボリューション法」の開発に取り組んだ。また、生体を構成する主要な脂質分子種とフラグメンテーションの関係性を明らかにし、構造から予測される保持時間と理論 MS/MS スペクトルを全 77,962 分子種に対して網羅的に構築した。本研究により開発されたプログラムが MS-DIAL であり、申請者の最初の脂質理論スペクトル構築の研究である。

本プログラムの有用性を検証するため、9種の藻類を LC-MS/MS に供した計測データを MS-DIAL で解析したところ、全 1,023 個の脂質分子種を同定することに成功した。さらに、同定した 1,023 個の脂質分子種それぞれの有無 0・1 の値を持った脂質フィンガープリントを調べ分類した結果、古典分類学の系統樹と完全に一致する結果を得た(図1)。当時、藻類株 UTEX2341 はクロレラ種に属するかナンノクロロプシス種に属するか議論が分かれていたが、脂質フィンガープリントによる系統樹はクロレラ種に属することを示唆し、

形態観察や 16S rRNA 系統解析による系統樹

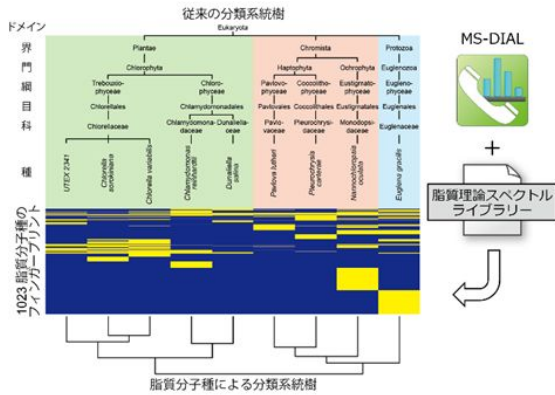


図1. 同定できた1023脂質分子種による藻類系統樹解析に存在する曖昧さを補う系統樹解析法としても応用可能であると期待される。

【既存スペクトルデータを学習セットとして用いた構造推定研究の最初の成果 (Analytical Chemistry 88, 7946-7958, 2016)】

メタボロミクス最大の難関は、未知代謝物由来ピークの構造推定である。生体内で存在が確認もしくは存在するものとして定義されている代謝物数は 50 万種類を超えている一方、標準品によるマススペクトルの確認が行われているのは 10,000 種にも満たない。そこで、当該研究では、未知スペクトルから組成式を予測し、化合物候補と MS/MS の妥当性を評価することで代謝物構造候補を「ランク付け」することで化合物構造を提案するソフトウェア MS-FINDER の開発を行った(図7)。また本論文の本質は、化合物構造とマスフラグメンテーションの関係性を解き明かし、マススペクトルを解釈するための統一的理論である「水素再配置則 (Hydrogen rearrangement rules)」を提唱したことにある。MS-FINDER を用いることで、頻用される飛行時間型質量分析レベルの質量精度において、98%以上の組成式予測能を達成し、正解となる構造がTop3に入る確率が80%となる精度で未知スペクトルの解析が可能となった。

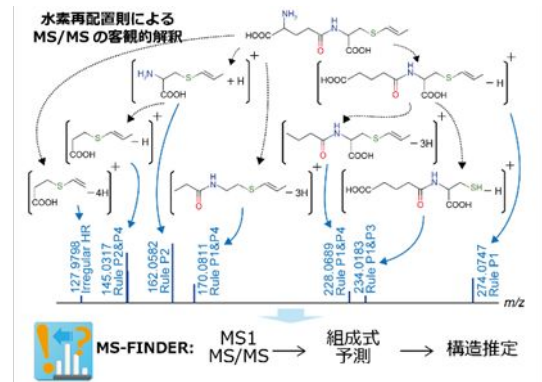


図2. 水素再配置則を用いた構造妥当性の評価と候補化合物の絞り込み

【質量分析ケムインフォマティクスによるエピメタボライトの同定 (Nature Methods 15, 53-56, 2018)】

当該研究では、既知の代謝経路には含まれない未開拓の代謝物群を「エピメタボライト」と定義し、質量分析ビッグデータからエピメタボライトを包括的に捉える「質量分析ケムインフォマティクス」の新手法を開発した。新たな質量分析ケムインフォマティクスの手法には、大きく分けて三つの技術革新が寄与している。

一つ目は、Oliver Fiehn 研究室で 13 年にわたって蓄積され、既知および未知代謝物の発現量やマススペクトル情報を納めたメタボロームデータベース (BinBase) から、エピメタボライト候補を抽出するプログラム「BinVestigate」の開発である。本研究では、構造が未解明、かつ特定の検体で有意に発現しているという条件での検索により、乳がん、糞便、微生物、光合成生物、特定の薬を投与した臨床検体のそれぞれにおいて特異的に発現する代謝物を合計五つ、エピメタボライト候補として抽出した。

二つ目は、MS-DIAL を改良し、ガスクロマトグラフィー質量分析装置 (GC-MS) と液体クロマトグラフィータンデム質量分析装置 (LC-MS/MS) といったどのタイプの装置からでも、搭載されたデコンボリューション法により高精度のマススペクトルを抽出できる技術を開発した。これにより、前述したエピメタボライト候補の高精度スペクトルを円滑かつ高解像度で抽出でき、化合物同定の精度を飛躍的に向上させることに成功した。

三つ目は、MS-FINDER の改良である。今回、メチル化やアセチル化といった修飾を含むエピメタボライト構造を検索可能にし、構造情報が豊富に含まれる GC-MS スペクトルに対しても、妥当性の高い候補化合物を効果的に絞り込む手法を構築することで、エピメタボライトの構造候補を絞り込みに大きく貢献した。予測された構造は実際に標準品の合成をすることで確認し、本手法の有効性の検証を行った。本研究では、「BinVestigate」、「MS-DIAL」、「MS-FINDER」の三つのケムインフォマティクス技術を組み合わせることで、これまで報告がなかった、乳がん特異的に検出される N-methyl UMP や腸内細菌が産生すると考えられる N-methyl alanine など五つのエピメタボライト発見に成功した(図3)。

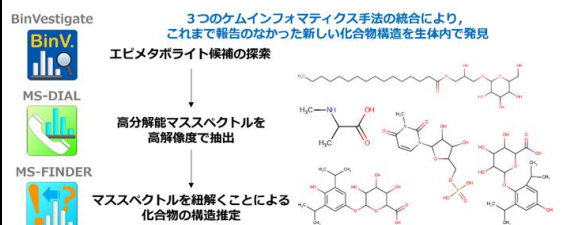


図3. ケムインフォマティクス手法の統合によるエピメタボライトの発見

5. 主な発表論文等  
(研究代表者、研究分担者及び連携研究者には下線)

〔雑誌論文〕(計 8件)

学術雑誌(以下,すべて査読付き学術雑誌であり,申請者が筆頭著者もしくは責任著者)  
1. Hiroshi Tsugawa. Advances in computational metabolomics and databases deepen the understanding of metabolisms. *Current Opinion in Biotechnology* **54**, 10-17, 2018.

2. Zijuan Lai<sup>§</sup>, Hiroshi Tsugawa<sup>§</sup>, Gert Wohlgemuth, (8名省略), Peter Beal, Masanori Arita<sup>\*</sup>, Oliver Fiehn<sup>\*</sup>. Identifying epimetabolites by integrating metabolome databases with mass spectrometry cheminformatics. *Nature Methods* **15**, 53-56, 2018 (<sup>§</sup>contributed equally, <sup>\*</sup>co-corresponding authors)  
3. Hiroshi Tsugawa<sup>\*</sup>, Kazutaka Ikeda, Makoto Arita<sup>\*</sup>. The importance of bioinformatics for connecting data-driven lipidomics and biological insights. *BBA-Molecular and Cell Biology of Lipids* **1862**, 762-765, 2017 (<sup>\*</sup>co-corresponding authors)

4. Hiroshi Tsugawa<sup>§</sup>, Kazutaka Ikeda<sup>§</sup>, Wataru Tanaka, Yuya Senoo, Makoto Arita<sup>\*</sup>, Masanori Arita<sup>\*</sup>. Comprehensive identification of sphingolipid species by in silico retention time and tandem mass spectral library. *Journal of Cheminformatics* **9**: 19, 2017 (<sup>§</sup>contributed equally, <sup>\*</sup>co-corresponding authors)  
5. Teruko Matsuo<sup>§</sup>, Hiroshi Tsugawa<sup>§\*</sup>, Hiromi Miyagawa, Eiichiro Fukusaki<sup>\*</sup>. Integrated strategy for unknown EI-MS identification using quality control calibration curve, multivariate analysis, EI-MS spectral database, and retention index prediction. *Analytical Chemistry* **89**, 6766-6773, 2017 (<sup>§</sup>contributed equally, <sup>\*</sup>co-corresponding authors)

6. Hiroshi Tsugawa, Tobias Kind, Ryo Nakabayashi, Daichi Yukihiro, Wataru Tanaka, Tomas Cajka, Kazuki Saito, Oliver Fiehn<sup>\*</sup>, Masanori Arita<sup>\*</sup>. Hydrogen rearrangement rules: computational MS/MS fragmentation and structure elucidation using MS-FINDER software. *Analytical Chemistry* **88**, 7946-7958, 2016 (<sup>\*</sup>co-corresponding authors)

7. Ryo Nakabayashi<sup>§</sup>, Hiroshi Tsugawa<sup>§</sup>, Tetsuya Mori, Kazuki Saito. Automation of chemical assignment for identifying molecular formula of S-containing metabolites by combining metabolomics and cheminformatics with 34S labeling.

*Metabolomics* **12**, 168, 2016 (<sup>§</sup>contributed equally)

8. Hiroshi Tsugawa, Tomas Cajka, Tobias Kind, Yan Ma, Brendan Higgins, Kazutaka Ikeda, Mitsuhiro Kanazawa, Jean VanderGheynst, Oliver Fiehn<sup>\*</sup>, Masanori Arita<sup>\*</sup>. MS-DIAL: data-independent MS/MS deconvolution for comprehensive metabolome analysis. *Nature Methods*, **12**, 523-526, 2015 (<sup>\*</sup>co-corresponding authors)

〔学会発表〕(計 3件)

9. 津川裕司. 質量分析を用いたメタボロミクス研究に資する情報処理および化合物同定のための技術開発.

<https://www.youtube.com/watch?v=HmKDhzhapNg>

10. Hiroshi Tsugawa (keynote). Identifying epimetabolites by mass spectrometry-based cheminformatics. 13th annual conference of the metabolomics society 2017, Brisbane, Australia, June 25-29.

11. Hiroshi Tsugawa. Hydrogen rearrangement rules: a study for computational MS/MS fragmentation and for structure elucidation using MS-FINDER. ASMS annual conference, June 5-9, 2016.

〔図書〕(計 1件)

11. 津川裕司, 早川英介, 三浦大典. メタボロミクスにおける化合物同定 生体内低分子代謝物の構造推定ガイド *J. Mass Spectrom. Soc. Jpn.*, **65**, 203-209, 2017

〔産業財産権〕

出願状況(計 0件)

名称:  
発明者:  
権利者:  
種類:  
番号:  
出願年月日:  
国内外の別:

取得状況(計 0件)

名称:  
発明者:  
権利者:  
種類:  
番号:  
取得年月日:  
国内外の別:

〔その他〕  
ホームページ等

研究代表者が開発したプログラムは、すべて以下の理化学研究所 PRIme website からダウンロード可能である。

<http://prime.psc.riken.jp/>

## 6. 研究組織

### (1) 研究代表者

津川 裕司 (TSUGAWA, Hiroshi)

国立研究開発法人理化学研究所・環境資源科学研究センター・研究員

研究者番号：30647235

### (2) 研究分担者 なし

( )

研究者番号：

### (3) 連携研究者 なし

( )

研究者番号：

### (4) 研究協力者

有田 正規 (ARITA, Masanori)

国立研究開発法人理化学研究所・環境資源科学研究センター・チームリーダー

研究者番号：10356389

斉藤 和季 (SAITO, Kazuki)

国立研究開発法人理化学研究所・環境資源科学研究センター・グループディレクター

研究者番号：00146705

中林 亮 (NAKABAYASHI, Ryo)

国立研究開発法人理化学研究所・環境資源科学研究センター・研究員

研究者番号：30586160

有田 誠 (ARITA, Makoto)

国立研究開発法人理化学研究所・生命医科学研究センター・チームリーダー

研究者番号：80292952

池田 和貴 (IKEDA, Kazutaka)

国立研究開発法人理化学研究所・生命医科学研究センター・副チームリーダー

研究者番号：496541

オリバー フィーン (FIEHN, Oliver)

カリフォルニア大学デイビス校・West Coast Metabolomics Center・教授

ジジュアン レイ (LAI, Zijuan)

カリフォルニア大学デイビス校・West Coast Metabolomics Center・学生

トビアス キント (KIND, Tobias)

カリフォルニア大学デイビス校・West Coast Metabolomics Center・Assistant Project Scientist

トーマス チャイカ (CAJKA, Tomas)

カリフォルニア大学デイビス校・West Coast Metabolomics Center・Associate Specialist

ゲルツ ウォールゲムート (WOHLGEMUTH, Gert)

カリフォルニア大学デイビス校・West Coast Metabolomics Center・リードプログラマー

サジャン メータ (MEHTA, Sajjan)

カリフォルニア大学デイビス校・West Coast Metabolomics Center・プログラマー

マシュー ミュラー (MUELLER, Matthew)

カリフォルニア大学デイビス校・West Coast Metabolomics Center・プログラマー