

平成 26 年 6 月 2 日現在

機関番号：12601

研究種目：基盤研究(B)

研究期間：2011～2013

課題番号：23310143

研究課題名(和文)代謝物マスペクトルからの構造推定

研究課題名(英文)Structure Estimation from Metabolite Mass Spectra

研究代表者

有田 正規 (ARITA, Masanori)

東京大学・理学(系)研究科(研究院)・准教授

研究者番号：10356389

交付決定額(研究期間全体)：(直接経費) 15,100,000円、(間接経費) 4,530,000円

研究成果の概要(和文)：一次代謝に関与する物質の標準GC/MSスペクトルを国内外の5機関から集め、物質ごとにスペクトルを比較したライブラリを構築した。またESIスペクトルはMassBank、PRIME、KNAPSAck等、代表的なメタボロームデータベースにおける代謝物のIDをInChIコードやCAS番号といったIDから物質毎のスペクトルをとりまとめた。またESIデータについてはクロマトグラムのピーク比較・評価手法を設計し、統合スペクトルデータのwikiサイトを構築した。また二次代謝物に関してはフラボノイドを解析した。これらの成果を、スペクトルを閲覧・解析するための統合メタボロミクスソフトウェアの一部として実現した。

研究成果の概要(英文)：We collected standard mass spectra from GC/MS for primary metabolites from five different research institutions in the world and compared their spectra after consolidating metabolite names and identifiers. Compared results were summarized as a standard library of theoretical spectra. For ESI spectra, we consolidated metabolite names and spectra from multiple repositories such as MassBank, PRIME, and KNAPSAck. The spectra were compared and merged by considering peak distribution and abundances. In addition, methods for visualizing and comparing chromatograms were also developed. For secondary metabolites, we investigated fragmentation patterns of flavanone, flavone and flavonol and recognized similarity between CID and ESI spectra. Such results are summarized in our wiki-based data repository. The chromatographic and spectral analyses were realized as a GUI-based software program.

研究分野：複合新領域

科研費の分科・細目：応用ゲノム科学

キーワード：バイオデータベース マスペクトル

1. 研究開始当初の背景

メタボローム分野はゲノムとプロテオームに続く第三のオミクスと呼ばれ、質量分析計(MS)の技術開発とともに大きく発展を遂げてきた。主な手法はマススペクトルからの代謝物一斉分析であるが、MSで恒常的に計測される(ノイズではない)ピークのうち、物質を特定できるものはガスクロマトグラフィー(GC)-MSでも2~3割、液体クロマトグラフィー(LC)-MSに至っては1割以下にとどまる。新型のMSⁿにおけるフラグメンテーション規則がわかれば、ランニングコストの低いMSⁿを用いてシステムティックに代謝物を同定でき、真にメタボロミクスと呼べる分野を確立できる。そのためマススペクトルからフラグメンテーションにおける知識、特に規則、を抽出することを目的とした。

2. 研究の目的

目標として以下の3項目を設定した。

- (1) ESI マススペクトルの重ね合わせアルゴリズム：代謝物構造とそれを計測した複数のスペクトルを与えられた際に、構造上重要なピークを中心にアライメントをおこなうアルゴリズムを開発すること。
- (2) 重要ピークを抽出するルール発見：基礎代謝物における重要ピークの数や質量を整理し、代謝物構造との相関を計算機で網羅的に解析する。特にEI-MSにおいて特徴的なルールの抽出をめざす。
- (3) 天然物構造・スペクトルの俯瞰マップ作成：メタボロミクス分野が扱う天然物のケミカルスペースをマススペクトルと関連付けて解釈できる二次元マップを作成する。

3. 研究の方法

マススペクトルデータベースの作成に合わせて収集されたスペクトル情報に対し、上記の3目標に対応させた計算機的手法による解析を実施した。

- (1) 重ね合わせアルゴリズム：基礎代謝物におけるEI-MSならびにESI-MSⁿスペクトルを収集し、各機関におけるピーク数や強度のバイアスを解析する。更に、重ね合わせアルゴリズムを用い、共通して検出されるピーク情報を明らかにする。
- (2) ピーク抽出ルール発見：キュレーションを経たESIスペクトルから順にEI-MS向けに開発した手法を適用し、ESI用の重ね合わせアルゴリズムとあわせて典型的な開裂ルールを見つける。
- (3) 俯瞰マップ作成：ピークのバリエーションを考慮することにより、ESIスペクトルやEIスペクトルのフラグメント分布や構造に基づく偏りを視覚化する手法を考案することを試みた。

4. 研究成果

EI スペクトルデータの収集と統合：

一次代謝に関与する物質の標準GC/MSスペクトルを国内外の5機関から集め、物質ごとにスペクトルを比較したライブラリを構築した。スペクトルが安定するといわれるGC/MSでも計測機器による違いがある。それを明らかにすべく、ノイズに基づくばらつきと機器によるばらつきの違いを推定した。また、このデータと、文献からデジタル化したピーク-フラグメンテーション間の対応関係を照らし合わせることで各標準物質において特徴的と考えられるピークの同定をおこなった。実測スペクトルに対するアノテーションと比較し、機関間において同定の程度にばらつきがあることを確認した。とりわけ糖はアノテーションの方針によりばらつきがある。また一部の機関しか同定していない物質もあった。これらを考慮して、マススペクトル重ね合わせアルゴリズムに基づき、GC-MSの半実験-半理論的スペクトルライブラリ(計測結果に基づくスペクトルを理論的に補正したライブラリ)を構築した。現在は作成した重ね合わせアルゴリズムが、有効なスペクトル情報を抽出できているか、スペクトルの検索実験を通して検証中である。

ESI スペクトルデータベースと統合：

ESIスペクトルについてもMassBank、PRIME、KNAPSAcK等、代表的なメタボロームデータベースにおける代謝物のIDをInChIコードやCAS番号といったIDを手がかりに物質毎にスペクトルをとりまとめた。これらのデータを用いて、統合マススペクトルデータのwikiサイトを構築した。

スペクトルはクロマトグラムのピーク比較・評価手法を設計し、ウェブブラウザ上でクロマトグラムを重ね合わせ表示できるプログラムを開発した。また、理化学研究所の津川裕司研究員とともにクロマトグラムのピーク評価手法を設計した。特にMultiple reaction monitoring (MRM)法におけるクロマト情報の統計的評価として溶出時間とピークの高さに加え、ピーク形状およびトランジションにおける高さ比などを考慮する手法を開発した。

これらの成果はスペクトルを閲覧・解析するための統合メタボロミクスソフトウェアとして理化学研究所から公開する予定である。

フラボノイドのスペクトル解析：

代表的な植物二次代謝物であるフラボノイドについて開裂パターンの詳細な解析をおこなった。およそ30のフラバノンとフラボン、そして50のフラボノールについてスペクトル検討したところ、陽イオンについては一般的にC環が^{1,3}A⁺で開裂するパターンがフラボノイド全体で頻出し、フラボノールでは^{1,2}A⁺-および^{1,2}B⁺が特徴ピークとして検出さ

れることがわかった。またフラバノンとフラボンの違いは[M+H]-C₂H₂O という脱離の有無で判定できる。これらはCID スペクトルとESI スペクトルに類似点があることを意味し、ESI スペクトルからフラボノイドの候補絞り込み同定作業の効率化が望める。今後かずさ DNA 研究所がおこなう FT/ICR-MS スペクトルの結果とも照らし合わせるにより、フラボノイドの詳細な分析結果を報告できると考える。

糖・脂質のスペクトル解析：

理化学研究所の古橋剛研究員とともに単糖および2糖の低エネルギーEI スペクトル、FI,CI スペクトルおよそ200件を比較し、糖構造のフラグメンテーションを解析した。通常のEI スペクトル(70eV)では異性体を区別できないが、低電圧(15-20eV)およびフィールドイオン化では糖鎖の結合情報を含む構造の違いをシステマティックに検出できることがわかっている。現在、それらの情報を統計処理し、フラグメンテーションルールとして定式化することを試みている。

同様の解析は複合脂質についてもおこなっている。ただしグリセロ脂質は分析機器性能の向上により得られる情報量が大幅に向上しているため、まずはこれらの情報を視覚化し処理するソフトウェアの開発を優先している。またこれらのスペクトル解析が完了していないため、天然物構造の全体を考慮したスペクトルの俯瞰マップ作成という目標は達成できなかった。そのかわり、リン脂質を俯瞰するGUIの作成はおこなっている。

5. 主な発表論文等

〔雑誌論文〕(計 4 件)

Furuhashi T, Ogawa T, Nakai R, Nakazawa M, Okazawa A, Padermschoke A, Nishio K, Hirai MY, Arita M, Ohta D "Wax ester and lipophilic compound profiling of *Euglena gracilis* by gas chromatography-mass spectrometry: toward understanding of wax ester fermentation under hypoxia" *Metabolomics* 2014, accepted

Tsugawa H, Kanazawa M, Ogiwara A, Arita M "MRMPROBS suite for metabolomics using large-scale MRM assays" *Bioinformatics* 2014, accepted

Ogawa T, Furuhashi T, Okazawa A, Nakai R, Nakazawa M, Kind T, Fiehn O, Kanaya S, Arita M, Ohta D "Exploration of Polar Lipid Accumulation Profiles in *Euglena gracilis* Using LipidBlast, a MS/MS Spectral Library Constructed in Silico" *Bioscience Biotechnology and Biochemistry* 78(1), 2014

Tsugawa H, Arita M, Kanazawa M, Ogiwara A, Bamba T, Fukusaki E "MRMPROBS: a data assessment and metabolite identification tool for

large-scale multiple reaction monitoring based widely targeted metabolomics" *Anal Chem* 85(10): 5191-5199, 2013

〔学会発表〕(計 7 件)

Arita M "Assessing Theoretical Spectra for Flavonoids using MassBank" *ICP2014*, selected oral presentation

Arita M "MassBank and its Future" Cosmos meeting on International Data Exchange in Metabolomics, April 2-3, Hinxton, Cambridge UK, 2014

Arita M "Genetic, Protein, and Metabolic Networks: Which Choice Is Amenable to Modelling" International Conference on Mathematical Modeling and Applications (ICMMA 2013), November 26-28, Tokyo, Japan, 2013

Arita M "Cyberinfrastructure for metabolomics and synthetic biology" International Symposium on Biotechnology for Green Growth, October 24-26, Kobe, Japan, 2012

Arita M "What can metabolomics learn from genomics and proteomics?" International Conference of Research and Application on Traditional Complementary and Alternative Medicine (TCAM), June 21-13, Muhammadiyah University of Surakarta (Indonesia), 2012

Arita M "Update on Lipid Databank and other lipidomics initiatives" EMBL-EBI Industry Programme & MetaboLights Project Workshop May 22-35, Hinxton, 2012

Arita M "Wiki-based Databases: Integration of knowledge through the web" C-NAIR Seminar Nasional, Gadjah Mada University (Indonesia) Dec 5, 2011

〔図書〕(計 1 件)

有田 正規 (編) 「使えるデータベース・ウェブツール」実験医学別冊 9月 29(15), 羊土社, 2011

〔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

研究成果は<http://metabolomics.jp/> および現在も作成を継続中のマススペクトルデータベース上に掲載予定。標準品のスペクトルデータおよびそのアノテーションは MassBank データベースに掲載。

6. 研究組織

(1)研究代表者

有田正規 (ARITA, Masanori)

東京大学・大学院理学系研究科・准教授

研究者番号：10356389

(2)研究分担者

なし

(3)連携研究者

なし