

令和 4 年 6 月 12 日現在

機関番号：12605  
研究種目：挑戦的研究（萌芽）  
研究期間：2018～2021  
課題番号：18K19155  
研究課題名（和文）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と代謝制御機構の解明

研究課題名（英文）Identifying epimetabolites and its elucidations by computational mass spectrometry

研究代表者  
津川 裕司（Tsugawa, Hiroshi）  
東京農工大学・工学（系）研究科（研究院）・准教授

研究者番号：30647235  
交付決定額（研究期間全体）：（直接経費） 4,400,000円

研究成果の概要（和文）：本研究では、既知代謝経路では定義できず酵素変化や非特異的な反応から生じるエピメタボライトといった新規代謝物の同定に向けた新たなノンターゲットメタボローム解析プラットフォームの構築を行なった。本研究により行われたタンデムマススペクトル（MS/MS）の機械学習により、未同定スペクトルが「どのような代謝物クラスに由来するものか」を予測することが可能となった。また、脂質理論MS/MSスペクトルの開発により、合計117脂質クラスを同定ための基盤構築を行い、生体試料から8000種を超える脂質多様性を捉えることに成功した。

#### 研究成果の学術的意義や社会的意義

本研究で構築した新たなMS/MS解釈論であるfragment set enrichment analysis (FSEA)は、過去に報告例が無いことはもちろん、既知構造とリファレンススペクトルの情報を必要としない全く新しい網羅的な化合物構造推定手法である。当該研究で構築される手法は疾患代謝研究だけでなく、腸内細菌叢を含めた微生物研究、および植物化学研究においても化合物構造決定の効率が飛躍的に向上し、全世界の代謝研究に貢献できると期待される。また、10万種を超えるとされる生命の脂質多様性を明らかにすることは、脂質異常症の背後に潜む動作原理を分子レベルで明らかにするための基礎的なデータを提供する。

研究成果の概要（英文）：In this study, we developed novel untargeted metabolome analysis platform that can identify previously unknown metabolites. Characterizing epimetabolites, which are not defined by known metabolic pathways but arise from enzymatic changes or nonspecific reactions, is the main aim of this study. The machine learning of tandem mass spectra (MS/MS) performed in this study enabled us to predict the metabolite class from unknown MS/MS spectrum information. In addition, the development of lipid theoretical MS/MS spectral library has laid the foundation for the annotation of a total of 117 lipid classes, successfully capturing more than 8,000 lipid molecules (lipid diversity) from biological samples.

研究分野：システムバイオロジー

キーワード：システムバイオロジー メタボロミクス リピドミクス 質量分析 インフォマティクス

### 1. 研究開始当初の背景

本研究では、既知の代謝制御機構では説明できない、遺伝子や酵素変異によって新しく生成される代謝産物を「エピメタボライト (Curr. Opin. Chem. Biol. 36, 70-76, 2017)」と定義し、その包括的な同定を行うための質量分析インフォマティクスの技術開発を行う。研究代表者は 2018 年度までの研究で、腸内細菌叢特異的に産生される新たなメチル化体である *N*-methyl alanine や乳がん細胞特異的に発現する新規メチル化体である *N*-methyl UMP のような新規エピメタボライトを、自身のインフォマティクス技術で明らかにしてきた。しかしながら、代表者らが理化学研究所で取得する様々な植物や動物組織のメタボロームデータの 80% 以上は、「未同定イオン」として残されているのが現状である。未同定イオンとは、液体クロマトグラフィータンデム型質量分析 (以下、LC-MS/MS) で得られる代謝物イオンの物理化学的プロパティ (保持時間、*m/z*、そして MS/MS スペクトル) のみがデータとして格納されていることを意味する。このような「質量情報」を円滑に「化合物情報」に変換することが、生物学的な解釈やシステムバイオロジー研究を加速させるために必須と考えられる。

### 2. 研究の目的

現状のノンターゲットメタボローム解析基盤では、ゲノム情報からは説明できないような未知代謝産物の同定は決して容易なことではない。そこで本研究では、既知の代謝制御機構では説明できない、遺伝子や酵素変異によって新たに生成される代謝産物を「エピメタボライト」として定義し、その包括的な同定を行うための質量分析インフォマティクスの技術構築を目的とした。

### 3. 研究の方法

エピメタボライトの包括的同定法のための研究戦略として大きく 2 つの方法論を展開する。1 つ目は、リファレンス MS/MS スペクトルや化合物構造情報を必要としない全く新しい概念であり、fragment set enrichment analysis (FSEA) という新 MS/MS 解釈論を構築するものである (GSEA と概念が類似するためこのように命名した)。流れとしては、未知 MS/MS を fragment ontology のセットへと変換し、ライブラリー化された chemical ontology 及び MS/MS fragment ontology セット中の検出頻度から、元構造の chemical ontology をフィッシャー検定による *p* 値に基づき推定する、というものである (図 1)。本手法により、未知の MS/MS スペクトルがどの化合物クラス (chemical ontology) に属するか客観的に判断することが可能となる。

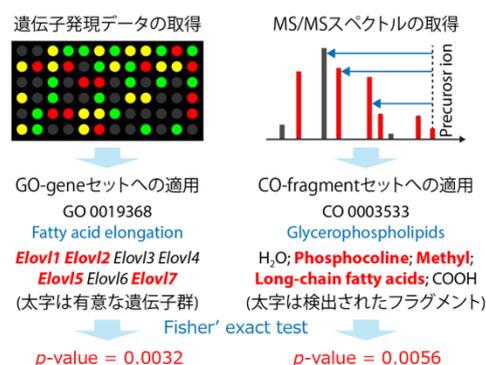


図 1. GSEA (左) と FSEA (右) の概念図

2 つ目は、代表者がこれまで構築してきた理論マススペクトル手法を更に拡張し、腸内細菌や臓器特異的な脂質分子種の網羅的なアノテーションを行うためのプラットフォーム構築に着手する。脂質理論 MS/MS スペクトル構築における基本的なアイデアは、1 つの脂質クラス内でのマスフラグメンテーション (MS/MS) 機構が体系化できることに起因する。そのため、ある脂質クラスと MS/MS の関連性が明らかになれば、脂質構造情報に対する MS/MS をインフォマティクスにより自動生成することが可能となる。本研究では、標準品が利用可能な脂質クラスに加え、文献情報の調査を詳細に行い (時には、文献の MS/MS 情報の電子化を行なった)、脂質アノテーションのプラットフォーム構築を行なった。また脂質研究に関してはライブラリー作成だけでなく、「MS/MS スペクトルの質」を評価することで、質量分析装置から得られる情報として適切な脂質表記を出力するアルゴリズムの開発にも着手した。たとえば図 2 は、ネガティブイオンモードにおけるスフィンゴミエリン (SM) 分子の酢酸イオン付加体から生成される MS/MS スペクトルの例を掲載している。リン酸イオン (*m/z* 78.959) やホスホコリン基に由来する *m/z* 168.043 のような断片化イオンは、SM クラス特異的なイオンとして観測される。このようなフラグメントイオンが観測されていれば、実際のプリカーサーイオン *m/z* 761.581 の値から、SM 34:1:20 とアノテーションできる。更に、強度は小さいものの *N*-acyl 基のペプチド結合が開裂されることで観

測される  $m/z$  449.315 が観測できた場合、SM 18:1;20/16:0 と記載する。このように、脂質研究の分野では、得られる MS/MS スペクトルの質を評価し、それに適切なアノテーション結果を出力することが国際的にも推奨されている。本研究では、作成した全 117 種類の脂質クラスに対して上述したようなアノテーション法を実装しており、MS/MS スペクトルのクオリティに基づいた脂質アノテーションを行うことが可能である。

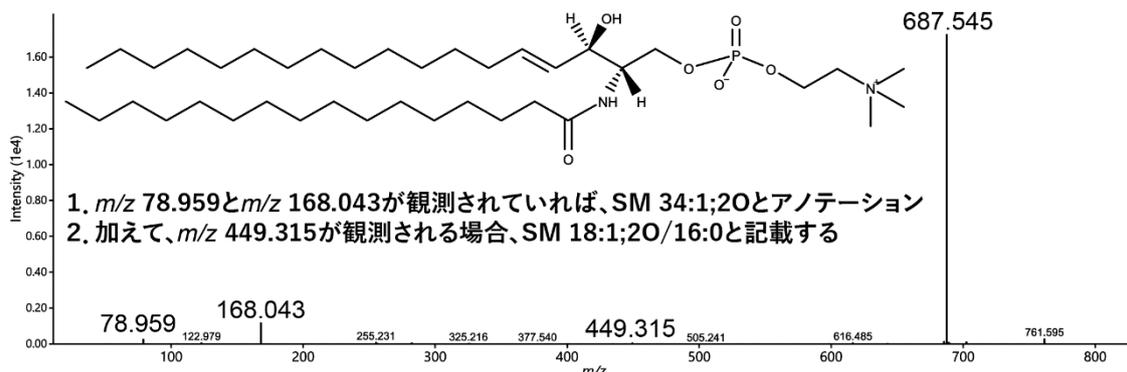


図2. スフィンゴミエリンであるSM 18:1(4E);1OH,3OH/16:0のMS/MSスペクトル

#### 4. 研究成果

FSEA と理論 MS/MS、およびスペクトルネットワークの手法を駆使して、植物天然物の網羅的なアノテーションを行った。スペクトルネットワーク法とは、「構造が類似している化合物から得られる MS/MS スペクトルは類似する」という仮説に基づき、MS/MS が類似している代謝物イオン（ノード）同士を繋げる（エッジ）ことで、分子ネットワークを生成する。これにより、データベースにより同定されたイオンの情報から、同定されなかったイオンの部分構造を推定することが可能となる（連想解析）。

本基盤技術を用いて、モデル植物であるシロイヌナズナや、農作物として重要なイネ、トウモロコシ、トマト、ジャガイモや、薬用植物として知られる甘草、タバコ、そして抗がん剤として用いられるカンプトテシンを含むアルカロイドを多く産生するチャボイナモリを含む合計 12 植物の葉、根、果実などさまざまな部位の LC-MS/MS データを解析した。その結果、69 個の新しい代謝物が本研究によって見出された。

また、上述したスペクトルネットワークのアイデアと、実際に得られた MS/MS スペクトルと予測された代謝物クラスや部分構造をもとに、各植物代謝物をひもづける「植物代謝物-MS/MS スペクトルネットワーク」形成のためのアルゴリズム開発を行った。これらの機能を、2015 年より開発が行われている質量分析データ統合解析プログラムである MS-DIAL (MS-DIAL 3.0) に組み込み、解析を行ったところ、合計で 824 の植物代謝物の構造を割り当てることに成功し、このうちの 505 種類（組成式のみ決定できたものを含めると 721 種類）が植物の「科（イネ科、ナス科など）」単位で独立して産生されていることを見いだした。植物研究ではこれまで、「各植物は、その植物種特異的な代謝物を作り出す」と考えられてきたが、本研究によりそれが具体的に証明された (Tsugawa, H. et al.: Nat. Methods 16, 295, 2019)。

構築した理論 MS/MS スペクトル法の可能性を更に実証するため、脂質メタボローム (リピドミクス) 解析への適用を試みた。生体内には多くの種類の脂質分子種が存在し、その多様性がさまざまな生命現象や機能制御に関わっている考えられており、その多様性を解明する意義は大きい。我々はまず、「脂質の分子構造と MS/MS スペクトルの関連性」を定式化し、MS/MS スペクトルから脂質構造を網羅的に推定するためのアルゴリズム開発に着手した (上述)。最終的には、現在脂質国際会議で定義されている脂質クラスの全 301 のうち、未報告の新しい構造を含む合計 117 脂質クラスを包括的に捉えるためのアルゴリズムを構築した。この数は、世界で公開されているどのプログラムよりも広い分子種をカバーしており、従来に比べて約 2 倍の脂質クラスを捉えることが可能である。

また、化合物構造を決定する基準として重要なパラメーターに、LC-MS/MS および IM (イオンモビリティ) -MS/MS における化合物の「溶出時間」がある。溶出時間とは、分析装置内で生体分子が「分離部」に入り、そこから溶出するまでの時間を表す。特に、同じ質量を持つ分子を分離するためのイオンモビリティによる化合物の溶出時間は、化合物自身の物理化学的性質である衝突断面積 (CCS) に依存するため、全ての研究機関で共有できる普遍的な物性であり、MS/MS スペクトルだけでは決定できない「異性体」を識別するためにも有用な化学情報である。そこで本研究では、全部で 3,000 種以上の脂質構造と溶出時間の情報をさまざまな機械学習器 (ディープラーニングや勾配ブースティングなど) に学習させ、最も高い予測を出力する学習器を選定した。その結果、勾配ブースティングに基づく XGBoost を用いることとし、117 脂質クラスに含まれる約 60 万種の脂質構造 (生物では未発見の推定構造を含む) の溶出時間の予測値を、誤差 2.5% 前後の予測精度で出力した。そして、脂質分子の構造推定テストを行ったところ、構築した MS/MS スペクトルのパターン認識、化合物の液体クロマトグラフィーにおける保持時間、イオンモビリティ

ティにおける CCS の情報全てを統合することで、分析データから正しい脂質分子構造を 99%前後の精度で出力できることが分かった。

構築したアルゴリズムを、質量分析データ統合解析プログラムである「MS-DIAL」に実装し、「MS-DIAL 4」として、世界に向けて公開した(<http://prime.psc.riken.jp/comps/index.html>)。MS-DIAL 4 を、我々が過去 5 年間に計測してきたさまざまな細胞、マウスおよびヒト由来検体の高精度なノンターゲット質量分析計測データに適用しところ、8,051 の脂質分子種の構造多様性を捉えることに成功した。この結果は、これまで 500~1,000 種ほどの分子種しか捉えられなかった既存の研究に比べ、およそ 10 倍の脂質構造の多様性を捉えるものである。さらに、未報告の新しい脂質クラスの構造を六つ推定し、そのうち一つの構造（生体に豊富に存在するスフィンゴミエリンに、新たにアシル基が一つ付加されたもの）に関しては、化合物の全合成を行い、二重結合の位置やシス/トランス異性体の帰属を行った (Tsugawa, H. et al.: Nat. Biotechnol. 38, 1159, 2020)。

## 5. 主な発表論文等

〔雑誌論文〕 計16件（うち査読付論文 16件 / うち国際共著 7件 / うちオープンアクセス 10件）

1. 著者名 Okahashi Nobuyuki, Ueda Masahiro, Yasuda Shu, Tsugawa Hiroshi, Arita Makoto	4. 巻 2
2. 論文標題 Global profiling of gut microbiota-associated lipid metabolites in antibiotic-treated mice by LC-MS/MS-based analyses	5. 発行年 2021年
3. 雑誌名 STAR Protocols	6. 最初と最後の頁 100492 ~ 100492
掲載論文のDOI (デジタルオブジェクト識別子) 10.1016/j.xpro.2021.100492	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -
1. 著者名 Mi-ichi Fumika, Ikeda Kazutaka, Tsugawa Hiroshi, Deloer Sharmina, Yoshida Hiroki, Arita Makoto	4. 巻 6
2. 論文標題 Stage-Specific De Novo Synthesis of Very-Long-Chain Dihydroceramides Confers Dormancy to Entamoeba Parasites	5. 発行年 2021年
3. 雑誌名 mSphere	6. 最初と最後の頁 e00174-21
掲載論文のDOI (デジタルオブジェクト識別子) 10.1128/mSphere.00174-21	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 該当する
1. 著者名 Yamamoto Hiroyuki, Nakayama Yasumune, Tsugawa Hiroshi	4. 巻 11
2. 論文標題 OS-PCA: Orthogonal Smoothed Principal Component Analysis Applied to Metabolome Data	5. 発行年 2021年
3. 雑誌名 Metabolites	6. 最初と最後の頁 149 ~ 149
掲載論文のDOI (デジタルオブジェクト識別子) 10.3390/metabo11030149	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -
1. 著者名 Matsuzawa Yuki, Higashi Yasuhiro, Takano Kouji, Takahashi Mikiko, Yamada Yutaka, Okazaki Yozo, Nakabayashi Ryo, Saito Kazuki, Tsugawa Hiroshi	4. 巻 69
2. 論文標題 Food Lipidomics for 155 Agricultural Plant Products	5. 発行年 2021年
3. 雑誌名 Journal of Agricultural and Food Chemistry	6. 最初と最後の頁 8981 ~ 8990
掲載論文のDOI (デジタルオブジェクト識別子) 10.1021/acs.jafc.0c07356	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Rai Amit, Hirakawa Hideki, Nakabayashi Ryo, Kikuchi Shinji, Hayashi Koki, Rai Megha, Tsugawa Hiroshi, Nakaya Taiki, Mori Tetsuya, Nagasaki Hideki, Fukushi Runa, Kusuya Yoko, Takahashi Hiroki, Uchiyama Hiroshi, Toyoda Atsushi, Hikosaka Shoko, Goto Eiji, Saito Kazuki, Yamazaki Mami	4. 巻 12
2. 論文標題 Chromosome-level genome assembly of <i>Ophiorrhiza pumila</i> reveals the evolution of camptothecin biosynthesis	5. 発行年 2021年
3. 雑誌名 Nature Communications	6. 最初と最後の頁 405
掲載論文のDOI (デジタルオブジェクト識別子) 10.1038/s41467-020-20508-2	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Yasuda Shu, Okahashi Nobuyuki, Tsugawa Hiroshi, Ogata Yusuke, Ikeda Kazutaka, Suda Wataru, Arai Hiroyuki, Hattori Masahira, Arita Makoto	4. 巻 23
2. 論文標題 Elucidation of Gut Microbiota-Associated Lipids Using LC-MS/MS and 16S rRNA Sequence Analyses	5. 発行年 2020年
3. 雑誌名 iScience	6. 最初と最後の頁 101841 ~ 101841
掲載論文のDOI (デジタルオブジェクト識別子) 10.1016/j.isci.2020.101841	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Tsugawa Hiroshi, Ikeda Kazutaka, Takahashi Mikiko, Satoh Aya, Mori Yoshifumi, Uchino Haruki, Okahashi Nobuyuki, Yamada Yutaka, Tada Ipputa, Bonini Paolo, Higashi Yasuhiro, Okazaki Yozo, Zhou Zhiwei, Zhu Zheng-Jiang, Koelmel Jeremy, Cajka Tomas, Fiehn Oliver, Saito Kazuki, Arita Masanori, Arita Makoto	4. 巻 38
2. 論文標題 A lipidome atlas in MS-DIAL 4	5. 発行年 2020年
3. 雑誌名 Nature Biotechnology	6. 最初と最後の頁 1159 ~ 1163
掲載論文のDOI (デジタルオブジェクト識別子) 10.1038/s41587-020-0531-2	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する

1. 著者名 Nothias Louis-Felix, Petras Daniel, Schmid Robin, Duhrkop Kai, Rainer Johannes, Sarvepalli Abinash, Protsyuk Ivan, Ernst Madeleine, Tsugawa Hiroshi et al.	4. 巻 17
2. 論文標題 Feature-based molecular networking in the GNPS analysis environment	5. 発行年 2020年
3. 雑誌名 Nature Methods	6. 最初と最後の頁 905 ~ 908
掲載論文のDOI (デジタルオブジェクト識別子) 10.1038/s41592-020-0933-6	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する

1. 著者名 Tada Ipputa, Tsugawa Hiroshi, Meister Isabel, Zhang Pei, Shu Rie, Katsumi Riho, Wheelock Craig E., Arita Masanori, Chaleckis Romanas	4. 巻 9
2. 論文標題 Creating a Reliable Mass Spectral-Retention Time Library for All Ion Fragmentation-Based Metabolomics	5. 発行年 2019年
3. 雑誌名 Metabolites	6. 最初と最後の頁 251 ~ 251
掲載論文のDOI (デジタルオブジェクト識別子) 10.3390/metabo9110251	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 該当する

1. 著者名 Nakabayashi Ryo, Mori Tetsuya, Takeda Noriko, Toyooka Kiminori, Sudo Hiroshi, Tsugawa Hiroshi, Saito Kazuki	4. 巻 92
2. 論文標題 Metabolomics with 15N Labeling for Characterizing Missing Monoterpene Indole Alkaloids in Plants	5. 発行年 2020年
3. 雑誌名 Analytical Chemistry	6. 最初と最後の頁 5670 ~ 5675
掲載論文のDOI (デジタルオブジェクト識別子) 10.1021/acs.analchem.9b03860	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Naoe Satoko, Tsugawa Hiroshi, Takahashi Mikiko, Ikeda Kazutaka, Arita Makoto	4. 巻 9
2. 論文標題 Characterization of Lipid Profiles after Dietary Intake of Polyunsaturated Fatty Acids Using Integrated Untargeted and Targeted Lipidomics	5. 発行年 2019年
3. 雑誌名 Metabolites	6. 最初と最後の頁 241 ~ 241
掲載論文のDOI (デジタルオブジェクト識別子) 10.3390/metabo9100241	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Tsugawa Hiroshi, Satoh Aya, Uchino Haruki, Cajka Tomas, Arita Makoto, Arita Masanori	4. 巻 9
2. 論文標題 Mass Spectrometry Data Repository Enhances Novel Metabolite Discoveries with Advances in Computational Metabolomics	5. 発行年 2019年
3. 雑誌名 Metabolites	6. 最初と最後の頁 119 ~ 119
掲載論文のDOI (デジタルオブジェクト識別子) 10.3390/metabo9060119	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 該当する

1. 著者名 Tsugawa Hiroshi, Nakabayashi Ryo, Mori Tetsuya, Yamada Yutaka, Takahashi Mikiko, Rai Amit, Sugiyama Ryosuke, Yamamoto Hiroyuki, Nakaya Taiki, Yamazaki Mami, Kooke Rik, Bac-Molenaar Johanna A., Oztolan-Erol Nihal, Keurentjes Joost J. B., Arita Masanori, Saito Kazuki	4. 巻 16
2. 論文標題 A cheminformatics approach to characterize metabolomes in stable-isotope-labeled organisms	5. 発行年 2019年
3. 雑誌名 Nature Methods	6. 最初と最後の頁 295 ~ 298
掲載論文のDOI (デジタルオブジェクト識別子) 10.1038/s41592-019-0358-2	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する

1. 著者名 Hiroshi Tsugawa	4. 巻 54
2. 論文標題 Advances in computational metabolomics and databases deepen the understanding of metabolisms	5. 発行年 2018年
3. 雑誌名 Current opinion in biotechnology	6. 最初と最後の頁 10-17
掲載論文のDOI (デジタルオブジェクト識別子) 10.1016/j.copbio.2018.01.008	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Hiroshi Tsugawa, Ryo Nakabayashi, Tetsuya Mori, Yutaka Yamada, Mikiko Takahashi, Amit Rai, Ryosuke Sugiyama, Hiroyuki Yamamoto, Taiki Nakaya, Mami Yamazaki, Rik Kooke, Johanna A Bac-Molenaar, Nihal Oztolan-Erol, Joost JB Keurentjes, Masanori Arita, Kazuki Saito	4. 巻 16
2. 論文標題 A cheminformatics approach to characterize metabolomes in stable-isotope-labeled organisms	5. 発行年 2019年
3. 雑誌名 Nature Methods	6. 最初と最後の頁 295-298
掲載論文のDOI (デジタルオブジェクト識別子) 10.1038/s41592-019-0358-2	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する

1. 著者名 Tsugawa Hiroshi, Rai Amit, Saito Kazuki, Nakabayashi Ryo	4. 巻 38
2. 論文標題 Metabolomics and complementary techniques to investigate the plant phytochemical cosmos	5. 発行年 2021年
3. 雑誌名 Natural Product Reports	6. 最初と最後の頁 1729 ~ 1759
掲載論文のDOI (デジタルオブジェクト識別子) 10.1039/D1NP00014D	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

〔学会発表〕 計1件（うち招待講演 0件 / うち国際学会 1件）

1. 発表者名 Hiroshi Tsugawa, Ryo Nakabayashi, Amit Rai, Tetsuya Mori, Yutaka Yamada, Hiroyuki Yamamoto, Makoto Arita, Masanori Arita, Kazuki Saito
2. 発表標題 PlaSMA: decoding plant specialized metabolome by the suite of mass spectrometry cheminformatics
3. 学会等名 Metabolomics conference (国際学会)
4. 発表年 2018年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

Computational mass spectrometry <a href="http://prime.psc.riken.jp/comps/index.html">http://prime.psc.riken.jp/comps/index.html</a>
--

6. 研究組織

氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------