

研究開発実施報告

□概要

研究開発課題名	物質循環を考慮したメタボロミクス情報基盤
開発対象データベースの名称(URL)	MetaboBank (http://mb.ddbj.nig.ac.jp)
研究代表者氏名	有田 正規
所属・役職	情報・システム研究機構 国立遺伝学研究所 教授 (2019年3月時点)

□目次

§1. 研究実施体制	2	① 概要	6
§2. 研究開発対象とするデータベース・ツール等	3	② 招待講演	6
(1) データベース一覧	3	③ 口頭講演	7
(2) ツール等一覧	3	④ ポスター発表	7
§3. 実施内容	4	(4) 知的財産権の出願	7
(1) 本年度の研究開発計画と達成目標	4	(5) 受賞・報道等	7
(2) 進捗状況	4	§5. 研究開発期間中に主催した活動(ワークショップ等)	
§4. 成果発表等	5	8
(1) 原著論文発表	5	1. 進捗ミーティング	8
① 論文数概要	5	2. 主催したワークショップ、シンポジウム、アウトリー	
② 論文詳細情報	5	チ活動等	8
(2) その他の著作物(総説、書籍など)	6		
(3) 国際学会および国内学会発表	6		

§1. 研究実施体制

グループ名	研究代表者または主たる共同研究者氏名	所属機関・役職名	研究題目
遺伝研	有田 正規	国立遺伝学研究所・教授	MetaboBankリポジトリ構築
奈良	金谷 重彦	奈良先端科学技術大学院大学・情報科学領域・教授	メタ代謝マップ構築
かずさ	平川 英樹	かずさDNA研究所・ゲノム情報解析施設・施設長	食品データアノテーション高度化
理研	福島 敦史	理化学研究所環境資源科学研究センター・研究員	植物データアノテーション高度化

§2. 研究開発対象とするデータベース・ツール等

(1) データベース一覧

【主なデータベース】

No.	名称	別称・略称	URL
1	MetaboBank	メタボバンク	http://mb.nig.ac.jp

【その他のデータベース】

No.	名称	別称・略称	URL
1	KNApSack	ナップザック	kanaya.naist.jp/KNApSack/
2	Metabonote	メタボノート	http://metabonote.kazusa.or.jp/
3	RIKEN MetaDatabase	理研メタデータベース	http://metadb.riken.jp/

(2) ツール等一覧

No.	名称	別称・略称	URL
1	MS-DIAL MS-FINDER		http://prime.psc.riken.jp/Metabolomics_Software/
2	Poweret Batch		http://www.kazusa.or.jp/komics/software/Poweret Batch

§3. 実施内容

(1) 本年度の研究開発計画と達成目標

MassBank wiki の開始(遺伝研)

- ・ Metabolonote サーバを MassBank ファミリーの一部と位置づけ、搭載されるメタデータを Metabolonote と MassBank の間で共通化。
- ・ Wiki スペクトルサーバにおいて自動取得できる情報は自動化し、InChIKey や SPLASH の登録をユーザが自分で実施しない形式にする。日本質量分析学会の MassBank と連携する。

MetaboBank 公共リポジトリの設計 (遺伝研)

- ・ 遺伝研スパコン内にストレージを用意し、MetaboBank リポジトリのプロトタイプを構築する。年度末にリポジトリを稼働させ、RDF 化した Metabolonote とリンクする。

食品メタボローム情報のリポジトリ登録 (かずさ)

- ・ 食品メタボローム情報(メタデータ)をリポジトリに登録する。
- ・ 食品メタボローム情報にアノテーションを施し、標品が入手できない化合物についてスペクトルのライブラリ化。

代謝物構造と生理活性の相関解析 (奈良先)

- ・ 二次代謝物の生物活性オントロジーの構築
- ・ メタ代謝マップの作成

C/MS 測定データのメタデータ化とリポジトリ登録 (理研)

- ・ Metabolonote 対応のメタデータ作成・拡張
- ・ 理研メタデータベース登録・公開
- ・ 遺伝研公共リポジトリへのデータ公開

(2) 進捗状況

・ MassBank wiki (遺伝研)

MassBank wiki サーバに、化合物の構造情報のフォーマット変換をおこなう機能を付加し、SMILES 形式、InChI、InChIKey の相互変換は可能になった。また SPLASH キーも自動作成できる。また脂質代謝物の名称を国際標準に基づいて整理した。これらの機能はスペクトル一次リポジトリとしての MassBank wiki のために作成した。2019 年はじめに MassBank サーバを日本質量分析学会が再稼働することが決定した。そのため一次リポジトリは MassBank.jp が担うことになり(実際は itHub MassBank がスペクトルの一次リソースになる)、MassBank wiki 登録用のスペクトルは MassBank フォーマットにて itHub に登録した。Wiki サーバは今後、二次リポジトリとして整備し、2019 年中に公開を開始予定である。

登録したスペクトル一覧

- RIKEN IMS 1140 スペクトル(リン脂質)
- RIKEN NPDEPO 1956 スペクトル(ポリケチド等二次代謝物)
- MSSJ 80 スペクトル(様々な物質群)
- RIKEN PLASMA10000 スペクトル (生体由来 1583、標品由来 8600)

・ MetaboBank 公共リポジトリの設計 (遺伝研)

MetaboBank を日米欧の INSDC 三極連携の枠組みに組み入れるため、メタボローム事業を DDBJ 事業の枠組みとして承認してもらい手続きを進めた。所内の委員会および外部評価委員会の承認を取得した。また EBI の MetaboLights 開発担当者(Claire O'Donovan)と連絡を取り、構築方針について密に連携している。当初はメタデータのモデルとして Metabolonote サーバを想定していたが、I

NSDC が管理する BioProject および BioSample とのリンクを重要視し、メタボローム登録用メタデータの中から、BioProject および BioSample 側で登録するメタデータの洗い出しを進めた。

- ・ **C/MS 測定データのメタデータ化とリポジトリ登録（理研）**
 ほぼ毎月、かずさ DNA 研および遺伝研の実務者とメタデータ RDF 作成会議を実施し、C/MS 測定データのメタデータ化をおこなった。ウェブ国際標準規格に沿った RDF 形式のメタデータ整備と共に、測定データは理研メタデータベースの枠組みを用いて、理研植物メタボロミクスメタデータベース (PMM <http://metabobank.riken.jp/>) として、公開しながら開発中である。本年度期間中に 24 件の C/MS データセット (理研 CSRS の DropMet データ含) を公開した (2019 年 5 月末時点調べ)。これら理研データ公開への取り組みは、遺伝研で構築されるメタボロームデータ公共リポジトリ (MetaboBank) との統合を経て、初期データとしての役割を担う。並行して、アップデートされた MS-DIAL による登録済 C/MS データの再解析によって、生体サンプルより抽出したマスペクトルデータの整理や代謝物プロフィールについて過去データとの比較を進めている (上記 24 件中 16 件が再解析済)。
- ・ **食品メタボローム情報のリポジトリ登録（かずさ）**
 食品メタボロームデータの情報 (メタデータ) を Metabolonote に登録し、それを含めて Metabolonote のデータ形式を新規作成する MetaboBank メタデータ形式 (RDF) に変換した。情報は理研側で作成した RIKEN PMM に登録予定である。また MassBase に登録された生データおよびそれらのメタデータを精査して、できる限り PMM 形式に変換した。
- ・ **代謝物構造と生理活性の相関解析（奈良先）**
 文献情報から、二次代謝物質と生物活性情報を、10758 対収集した。また、生物種間を生物が生成する有機気化物質 (BVOC) から関連付けるためのアルゴリズム BiClusO を開発した [1]。この BiClusO は二つの属性、例えば、生物種と BVOC の関係性を考慮して、生物種と BVOC に密に関連するクラスターを推定する方法であり、バイクラスタリング法という。この方法により、二次代謝物質と生物活性情報の関連を考慮した適切なグループを構築づける基盤ができあがった。
 また、生合成経路が既知のアルカロイド化合物を調査したところ 566 種類にとどまった。構造既知のアルカロイド化合物の生合成開始物質を推定する目的で、深層学習の一つである分子グラフ・コンボリューション・ニューラル・ネットワーク (MCNN) を活用した。その結果、認識率において 97% という高い分類性能を得ることができた [2]。二次代謝物質における生合成開始物質の予測は、代謝マップ上に代謝を射影する上で非常に重要な情報であり、代謝情報未知の化合物についての代謝マップ上の情報を提供できる。

§4. 成果発表等

(1) 原著論文発表

① 論文数概要

種別	国内外	件数
発行済論文	国内 (和文)	0 件
	国際 (欧文)	6 件
未発行論文 (accepted, in press 等)	国内 (和文)	0 件
	国際 (欧文)	0 件

② 論文詳細情報

1. Burla B, Arita M, Arita M, Bendt AK, Cazenave-assiot A, Dennis EA, Ekroos K, Han X, Ikeda K, Liebisch, Lin MK, Loh TP, Meikle PJ, Orešič M, Quehenberger O, Shevchenko A, Torta F, Wakelam MJO, Wheelock CE, Wenk MR, “MS-based lipidomics of human blood plasma: a community-initiated position paper to develop accepted guidelines”, J Lipid Res., 59(10), 2001-2017, 2018 (DOI: [10.1194/jlr.S087163](https://doi.org/10.1194/jlr.S087163))

2. Fukushima A, Hikosaka S, Kobayashi M, Nishizawa T, Saito K, Ito E, Kusano M, "A Systems Analysis With "Simplified Source-Sink Model" Reveals Metabolic Reprogramming in a Pair of Source-to-Sink Organs During Early Fruit Development in Tomato by LED Light Treatments", 9, 1439, 2018 (DOI: [10.3389/fpls.2018.01439](https://doi.org/10.3389/fpls.2018.01439))
3. Kitazaki K, Fukushima A, Nakabayashi R, Okazaki Y, Kobayashi M, Mori T, Nishizawa T, Reyes-Chin-Wo S, Michelsmore RW, Saito K, Shoji K, Kusano M, "Metabolic Reprogramming in Leaf Lettuce grown Under Different Light Quality and Intensity Conditions Using Narrow -Band LEDs", Sci. Rep., 8(1), 7914, 2018, (DOI: [10.1038/s41598-018-25686-0](https://doi.org/10.1038/s41598-018-25686-0))
4. K. Liu, A. H. Morita, S. Kanaya and Md. Altaf-Ul-Amin, "Metabolite-content-guided prediction of medicinal/edible properties in plants for bioprospecting", CRCAM-130, CRCAM-130, 130, 2018 (DOI: [10.29011/CRCAM-130/100030](https://doi.org/10.29011/CRCAM-130/100030))
5. Eguchi R, Karim MB, Hu P, Sato T, Ono N, Kanaya S, Altaf-Ul-Amin M, "An integrative network-based approach to identify novel disease genes and pathways: a case study in the context of inflammatory bowel disease.", BMC Bioinformatics, 19(1), 264, 2018 (DOI: [10.1186/s12859-018-2251-x](https://doi.org/10.1186/s12859-018-2251-x))
6. Ohnishi M, Anegawa A, Sugiyama Y, Harada K, Oikawa A, Nakayama Y, Matsuda F, Nakamura Y, Sasaki R, Shichijo C, Hatcher P, Fukaki H, Kanaya S, Aoki K, Yamazaki M, Fukusaki E, Saito K, Mimura T, "Molecular Components of Arabidopsis Intact Vacuoles Clarified with Metabolomic and Proteomic Analyses.", Plant Cell Physiol., 59(7), 1353-1362, 2018 (DOI: [10.1093/pcp/pcy069](https://doi.org/10.1093/pcp/pcy069))

(2) その他の著作物(総説、書籍など)

該当なし

(3) 国際学会および国内学会発表

① 概要

種別	国内外	件数
招待講演	国内	13 件
	国際	3 件
口頭発表	国内	2 件
	国際	0 件
ポスター発表	国内	2 件
	国際	1 件

② 招待講演

〈国内〉

1. 福島敦史、「質量分析を用いた植物メタボロミクスデータの情報解析と共有」、質量分析インフォマテイクス研究会・第4回ワークショップ、東京、2019年3月19日(火)
2. 金谷 重彦、「二次代謝物の生合成経路の体系的解析を目指したデータサイエンス」第64回日本放線菌学会学術講演会、2019.3.14(東京)
3. 金谷 重彦、「データサイエンス全般から化学の話題まで、Future Trend in Polymer Science 2018(2019.3.2 東京)」
4. 櫻井 望、「キー化合物を探索するためのメタボロームデータベース、理研シンポジウム「植物の代謝制御と化学生物学の新展開」、2018年12月20日
5. 櫻井 望、「食品に含まれる有用成分の研究、第27回異物異臭勉強会、2018年11月30日
6. 櫻井 望、「有用成分の探索による産業開拓へ向けたメタボローム解析の応用、EF 産学共創パートナーシップ第3回研究会、2018年11月21日
7. 有田 正規、「生薬情報のデータベース」薬用資源の持続的利用促進に関する研究会、11/21, 草津(滋賀), 2018
8. 櫻井 望、「多変量解析演習(主成分分析・PLS-DA・PLS 回帰)、メタボロミクスソフトウェア講習会、2018年11月12日
9. 櫻井 望、「ターゲットメタボロミクスの向上に向けて、第10回 LC/MS ワークショップ、2018年11

月 2 日

10. 櫻井 望、抱合体、フラボノイドの網羅的解析手法の紹介とその応用、CBI 学会 2018 年大会 (株) ツムラ スポンサーセッション、2018 年 10 月 11 日
11. 櫻井 望、食品に含まれる有用成分の研究、新食品会 第 2 回例会、2018 年 9 月 19 日
12. 櫻井 望、ケミカルセンサー開発によるメタボロームの高度利用、植物の栄養研究会 第 4 回研究交流会、2018 年 9 月 7 日
13. 櫻井 望、掴もう! メタボロームデータをもっと活用するためのコツ、第 51 回植物バイオシンポジウム、2018 年 7 月 18 日

〈国際〉

1. Arita M "Open genome analysis in the post-genomic era" International Workshop on Data Science, Mishima, November 15, 2018
2. Sakurai N "Databases for metabolomics-assisted life science" 2018 KSBS-BG21-GSP Joint Symposiums, July 13, 2018
3. Arita M "Computational Metabolomics" 6th Annual Korea Metabolomics Society (plenary), Seoul, Korea, April 6, 2018

③ 口頭講演

〈国内〉

1. **Atsushi Fukushima** and Kozo Nishida, "MSEAp: Development of a metabolite set enrichment analysis toolkit for plant metabolomics community", 第 36 回日本植物細胞分子生物学会(金沢)大会、金沢、2018 年 9 月 6 日(金)~8 日(日)
2. 有田 正規 「生命科学データ共有のあり方」Japan Open Science Summit, 6/19, 東京, 2018

〈国際〉

該当なし

④ ポスター発表

〈国内〉

1. 福島敦史ほか、「植物メタボロームデータの再解析・アノテーション高度化に向けた情報基盤整備」、トーゴーの日シンポジウム 2018、東京、2018 年 10 月 5 日(金)
2. 福島敦史ほか、「植物メタボロームデータの再解析・アノテーション高度化に向けた情報基盤整備」、第 12 回メタボロームシンポジウム、鶴岡、2018 年 10 月 17 日(水)~19 日(金)

〈国際〉

1. 発表者、タイトル、学会名、場所、月日
2. **Atsushi Fukushima**, et al. "RIKEN Plant Metabolome MetaDatabase: An integrated plant metabolome data repository based on the semantic web", SWAT4HCLS 2018, Antwerp, December 4-5, 2018

(4) 知的財産権の出願

該当なし

(5) 受賞・報道等

該当なし

§5. 研究開発期間中に主催した活動(ワークショップ等)

1. 進捗ミーティング

年月日	名称	場所	参加人数	目的・概要
2018年 5月1日	チーム内ミーティング(非公開)	NAIST 金谷研究室	15人	メタボロームリポジトリの会議
2018年 7月20日	チーム内ミーティング(非公開)	東京連絡事務所会議室3	8人	同上
2018年 7月27日	チーム内ミーティング(非公開)	遺伝研有田研究室	6人	同上
2018年 9月26日	チーム内ミーティング(非公開)	東京連絡事務所会議室3	8人	同上
2018年 9月28日	チーム内ミーティング(非公開)	かずさDNA研究所	3人	研究進捗報告のためのミーティング
2018年 10月12日	チーム内ミーティング(非公開)	東京連絡事務所会議室3	8人	同上
2018年11 月7日	チーム内ミーティング(非公開)	東京連絡事務所会議室3	7人	同上
2018年12 月18日	チーム内ミーティング(非公開)	東京連絡事務所会議室3	7人	同上
2019年2 月19日	チーム内ミーティング(非公開)	東京連絡事務所会議室1	7人	同上

2. 主催したワークショップ、シンポジウム、アウトリーチ活動等

年月日	名称	場所	参加人数	目的・概要
2018年 11月12日	メタボロミクスソフトウェア講習会	理研横浜交流棟ホール	100人 (企画運営除く)	質量分析インフォマティクス研究会、DDBJとの共同開催。開発するソフトウェアMS-DIAL等のハンズオン利用講習会

以上

別紙1 既公開のデータベース・ウェブツール等

No.	正式名称	別称・略称	概要	URL	公開日	状態	分類	関連論文
1	MetaboBank	メタボバンク		http://mb.ddbj.nig.ac.jp		新規	データベース等	
2	KNApSack	ナップザック	二次代謝産物データベースKNApSackをコアシステムとした統合型データベースです。遺伝子アノテーションではArabidopsis、Bacillus、Humanが公開されています。タンパク質の金属イオン結合部位のデータベースMetalMine、インドネシア生薬データベースJAMUや漢方薬のデータベースKAMPOも含まれています。更に、世界119カ国から7356の植物種データが格納されています。これらのデータは出版されている科学論文から集めているものです。	kanaya.naist.jp/KNApSack/		維持・発展	データベース等	
3	Metabolonote	メタボロノート	メタボロミクス実験の詳細な実験手法に関する情報（メタデータ）のみを専門的に扱うデータベースです。セマンティックMediaWikiを利用したシステムにより、ユーザー登録（無料）をすることでだれでも気軽に各自のメタデータを記録・編集することができます。 メタデータを実際のデータ（生データファイルや、ピークアノテーション情報、ピークのスペクトル情報など）と切り離して管理することにより、1) 実験後すぐに、さらには実験前であっても、メタデータを記載することができるため、実験設定の詳細を忘れてしまう前に記録に留めておくことができます。また、2) 一度記録したメタデータは、論文や実際のデータを管理するその他のデータベースから共通して参照できるというメリットがあります。 Metabolonoteのコアシステムは公開されているため、ユーザー独自のMetabolonoteをLIMSや公開サイトとして構築できるほか、フォーマットを独自に定義することで、メタボロミクス以外のデータ管理についても使用	http://metabolonote.kazusa.or.jp/		維持・発展	データベース等	
4	RIKEN MetaDatabase	理研メタデータベース		http://metadb.riken.jp/		維持・発展	データベース等	
5	MS-DIAL		実測データからのスペクトル抽出、アノテーション用ソフトウェア	http://prime.psc.riken.jp/Metabolomics_Software/		維持・発展	ツール等	
6	MS-FINDER		実測データからのスペクトル抽出、アノテーション用ソフトウェア	http://prime.psc.riken.jp/Metabolomics_Software/		維持・発展	ツール等	
7	PowerGet		実測データからのスペクトル抽出、アノテーション用ソフトウェア	http://www.kazusa.or.jp/komics/software/PowerGet		維持・発展	ツール等	