

ライフサイエンスデータベース統合推進事業
(統合化推進プログラム 2018 年採択課題)

研究開発終了報告書

「物質循環を考慮したメタボロミクス情報基盤」

有田 正規

情報・システム研究機構 国立遺伝学研究所 教授

研究開発期間： 2018年4月～2023年3月



©2023 有田 正規(国立遺伝学研究所) Licensed under CC BY 4.0 国際

目次

§1. 研究開発実施の概要.....	4
§2. 研究開発実施体制.....	5
1. 研究グループ.....	5
(1) 「遺伝研」グループ（研究代表者グループ）.....	5
(2) 「奈良先」グループ（主たる共同研究者グループ(1)）.....	5
(3) 「かずさ」グループ（主たる共同研究者グループ(2)）.....	6
(4) 「理研」グループ（主たる共同研究者グループ(3)）.....	6
(5) 「農工大」グループ（主たる共同研究者グループ(4)）.....	7
(6) 「府大」グループ（主たる共同研究者グループ(3)）.....	7
2. 有識者会議等.....	8
(1) DNA データ研究利用委員会.....	8
(2) DDBJ サービスに関する専門部会.....	9
§3. 研究開発の目的、実施内容及び成果.....	11
1. 研究開発の背景.....	11
2. 研究開発対象のデータベース・ツール.....	12
(1) データベース.....	12
(2) ツール等.....	12
3. 達成目標及び実施計画.....	13
(1) 当初の実施計画・達成目標.....	13
(2) 期間中に追加・削除・変更した実施計画・達成目標.....	13
4. 実施内容.....	14
(1) 実施内容.....	14
(2) 実施内容のうちの特項目の詳細.....	20
§4. 主要なデータベースの利活用状況.....	23
1. アクセス数.....	23
(1) 実績.....	23
(2) 分析.....	23
2. データベースの利用状況を示すアクセス数以外の指標.....	23
3. データベースの利活用により得られた研究成果（生命科学研究への波及効果）.....	23
4. データベースの利活用によりもたらされた産業への波及効果や科学技術のイノベーション（産業や科学技術への波及効果）.....	24
§5. 今後の展開.....	25
§6. 自己評価.....	26
§7. 外部発表等.....	28
1. 原著論文発表.....	28
(1) 論文数概要.....	28
(2) 論文詳細情報.....	28
2. その他の著作物（総説、書籍など）.....	31
3. 国際学会発表及び主要な国内学会発表.....	32
(1) 概要.....	32
(2) 招待講演.....	32
(3) 口頭講演.....	32
(4) ポスター発表.....	33
4. 知財出願.....	35
(1) 出願件数.....	35
(2) 一覧.....	35

① 国内出願	35
5. 受賞・報道等.....	35
(1) 受賞.....	35
(2) メディア報道	35
(3) その他.....	35
§8. 研究開発期間中の活動	36
1. 進捗ミーティング	36
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.....	39

§1. 研究開発実施の概要

本研究課題ではメタボローム生データの公共リポジトリである **MetaboBank** を構築、公開した。公共リポジトリとは、受け付けるデータをパブリック・ドメインにおいて恒久的に参照可能とする活動であり、登録者からの信頼および立場の中立性が重要となる。そのため、**MetaboBank** を国際塩基配列データベース連携 (INSDC) に並ぶ **DBJ** サービスとして位置づけ、登録手続きも **BioProject**, **BioSample** という **INSDC** サービスと連携、**DBJ** と同じユーザーアカウントから実施する形式にした。**INSDC** との連携サービスと呼べる **Version2** は 2021 年 10 月より開始し、開発者以外からのデータ登録受付も開始した。

本リポジトリの初期データセットとして、理研 **CSRS** およびかずさ DNA 研究所が 10 年以上にわたって測定してきたデータセットを精査し、メタデータを付与する作業を実施した。かずさデータからは、これまで **MassBase** および **Metabolonote** データベースに収載されていた合計 104 の研究プロジェクト (植物 79, 微生物:3, 食品:16, その他:6。生物の種数は 157) を整備し、3,428 の解析生データと紐付けた。理研データからは 59 の研究プロジェクト、9 千以上の解析生データを整備し、そのうち再解析対象とした **GC-MS** データ 54 プロジェクト分について、4 つの解析ソフトウェア (**MS-DIAL**, **eRah**, **AI-Output**, **IP4M**) を用いて再アノテーションを進めた。整備したメタデータは **Resource Description Framework (RDF)** の書式に変換し、**MetaboBank Version2** より公開している。

これらデータセットを再解析するためのソフトウェアとして **PowergetBatch** および **MS-DIAL** というソフトウェアの開発を実施した。かずさデータからは、改善した **PowergetBatch** を用いて化合物の分類レベル (フラボノイドやアミノ酸といったレベル) で予測できるピーク数が自動解析で 2 倍、手動解析では 4 倍に増加させることができた。また理研で開発した **MS-DIAL** はイオンモビリティ解析など様々なメタボロミクス研究に対応できる統合プラットフォームとして発展を続け、世界中に利用者を持つ標準的な解析ツールとして認識されるまでに発展した。

こうして整備された代謝物データを、文献より得られた **KNAPSAcK** データベース (141,486 の生物種-代謝物関係、代謝物総数 57,906 種) と照合できるように、**CAS** 番号、**InChI** コード、**PubMed** 番号を付与する作業を実施した。このデータは **RDF** 化され、統合化推進プログラムにおける植物ゲノム統合データベース、微生物ゲノム統合データベースにも用いられている。また米国 **PubChem** からも参照されている。更にアルカロイドの生合成に関するメタ代謝マップも整備した。これらの情報は **KNAPSAcK** および **MetaboBank** に関連した **wiki** サーバより公開した。

§2. 研究開発実施体制

1. 研究グループ

(1) 「遺伝研」グループ(研究代表者グループ)

・ 人員構成

氏名	所属機関	役職	研究開発項目	参加時期
有田 正規	国立遺伝学研究所	教授	総括	2018.4～2023.3
川島 武士	国立遺伝学研究所	助教	リポジトリ構築	2018.4～2022.1
佐藤 充治	国立遺伝学研究所	特定有期 雇用職員	代謝マップ作 成	2018.4～2023.3
多田 一風太	総合研究大学院大学 先端生命科学研究所	大学院生	リポジトリ構築	2018.4～2023.3
櫻井 望	国立遺伝学研究所	特任准教 授	リポジトリ構築	2018.4～2023.3
時松 敏明	同上	特定有期 雇用職員	データ受付、 キュレーション	2018.4～2023.3
吉本 美和	同上	技術補佐 員	データ収集・整 理	2018.4～2023.3
パク ジェヒョ ク	同上	技術補佐 員	データ収集・整 理	2021.10～ 2023.3
児玉 悠一	同上	特定有期 雇用職員	データフォーマ ット作成	2021.4～2023.3
福田 亜沙美	同上	特定有期 雇用職員	データフォーマ ット作成	2021.4～2023.3
高木 佳苗	同上	特定有期 雇用職員	データフォーマ ット作成	2022.4～2022.8

・ 担当項目

MetaboBank リポジトリの構築

MetaboBank メタデータの策定と受付業務

MetaboBank Wiki データの作成

(2) 「奈良先」グループ(主たる共同研究者グループ(1))

・ 人員構成

氏名	所属機関	役職	研究開発項目	参加時期
金谷 重彦	奈良先端大学院大学 先端科学技術研究科	教授	メタ代謝マップ 構築	2018.4～2023.3
小野 直亮	同上	准教授	メタ代謝マップ 構築	2018.4～2023.3
森田 晶	同上	研究員	メタ代謝マップ 構築	2018.4～2023.3

大橋 美名子	同上	技術補佐員	メタ代謝マップ構築	2018.4～2023.3
黄 銘	同上	助教	メタ代謝マップ構築	2020.4～2023.3
城戸 孝士郎	同上	大学院生	代謝経路収集	2020.4～2021.3
石谷 勇人	同上	大学院生	代謝経路収集	2020.4～2021.3
宮崎 優	同上	大学院生	代謝経路収集	2020.4～2022.3
江口 遼平	同上	博士研究員	生理活性情報	2018.4～2020.3
田中 徹士	同上	大学院生	生理活性情報	2018.4～2020.3

• 担当項目

メタ代謝マップ構築

(3)「かずさ」グループ(主たる共同研究者グループ(2))

• 人員構成

氏名	所属機関	役職	研究開発項目	参加時期
平川 英樹	かずさDNA研究所	ゲノム情報解析施設長	レポジトリ登録	2018.4～2023.3
長崎 英樹	同上	研究員	アノテーション	2018.4～2023.3(2020年度除く)
櫻井 望	同上	特別客員研究員	アノテーション	2018.4～2023.3
大澤 祥子	同上	研究補佐院	レポジトリ登録	2018.6～2022.9

• 担当項目

食品データアノテーション高度化

(4)「理研」グループ(主たる共同研究者グループ(3))

• 人員構成

氏名	所属機関	役職	研究開発項目	参加時期
福島 敦史	理研環境資源科学研究センター	研究員	アノテーション	2018.4～2023.3
津川 裕司	同上	研究員	ソフトウェア開発	2018.4～2023.3
高橋 みき子	同上	テクニカルスタッフ	アノテーション	2018.4～2023.3
小林 紀郎	理化学研究所・情報システム本部	研究補佐院	データベース開発	2018.4～2023.3
長崎 英樹	理研環境資源科学研究	研究員	アノテーション	2020.4～2021.3

センター				
神谷 健	同上	パートタイマー	データ整理、再解析	2018.4～2020.3
青野 佑亮	同上	パートタイマー	データ整理、再解析	2018.4～2023.3
加藤 雅樹	理化学研究所・情報システム本部	技師	データベース開発	2021.4～2023.3
佐野 瑞希	理研環境資源科学研究センター	パートタイマー	データ整理、再解析	2021.4～2020.3
轡田 圭又	理研環境資源科学研究センター	パートタイマー	データ整理、再解析	2021.4～2022.3
朽方 ひかり	理研環境資源科学研究センター	パートタイマー	データ整理、再解析	2022.4～2023.3

- 担当項目

植物データアノテーション高度化

(5)「農工大」グループ(主たる共同研究者グループ(4))

- 人員構成

氏名	所属機関	役職	研究開発項目	参加時期
津川 裕司	東京農工大学	准教授	ソフトウェア開発	2021.4～2023.3
松沢 佑紀	同上	技術補佐員	ソフトウェア開発	2021.4～2023.3

- 担当項目

MS-DIAL 開発

(6)「府大」グループ(主たる共同研究者グループ(3))

- 人員構成

氏名	所属機関	役職	研究開発項目	参加時期
福島 敦史	京都府立大学 生命環境学部	教授	アノテーション	2022.4～2023.3

- 担当項目

植物データアノテーション高度化

2. 有識者会議等

(1) DNA データ研究利用委員会

会議概要

名称	DNA データ研究利用委員会
目的	<p>委員会は、センターが実施している日本 DNA データバンク事業を推進 するため、国内の研究機関との連携、ENA/EBI(欧州)、GenBank /NCB I(米国)との国際連携協力のもとに次の各号に掲げる事項について審議する。</p> <p>一 センターの事業計画に関すること 二 センターの関係事業経費に関すること 三 センターの総合的な評価・調整に関すること 四 その他 DNA データに関する調査・検討等に関すること</p>
委員数	11 人(2020-2022)

委員一覧

氏名	所属機関	役職	備考
伊藤隆司	九州大学 医学研究院	教授	
笠原雅弘	東京大学大学院 新領域創成科学研究科	教授	
二階堂愛	理化学研究所 生命機能科学研究センター	チームリーダー	
平井優美	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	
岩島真理	JST バイオサイエンスデータベースセンター	室長	2022 年より
松岡聡	理化学研究所 計算科学研究センター	センター長	
水島洋	国立保健医療科学院 研究情報支援研究センター	センター長	
菅野純夫	東京医科歯科大学	特任教授	
有田正規	国立遺伝学研究所	教授	
大久保公策	国立遺伝学研究所	教授	
黒川顕	国立遺伝学研究所	教授	
井ノ上逸郎	国立遺伝学研究所	教授	

開催歴

年月日	場所	参加人数	主な議題
2018 年 2 月 20 日	国立遺伝学研究所 (当時の DDBJ センター長は高木利久教)	27 人	第35回 DNA データ研究利用委員会 メタボローム・リポジトリの試験運用と題し、DDBJ サービスとして MetaboBank の運用を打診。理研 CSRS,かずさ DNA 研, 遺伝研の

	授)		3 研究所間での「生命科学データの共有に関する覚書」締結を報告。予算制約のなかでの運営への懸念。その他、指摘事項は特になし。
2019年2月20日	国立遺伝学研究所 (センター長 有田正規)	23人	第36回 DNA データ研究利用委員会 2018年11月のメール審議における、メタボローム・リポジトリの開設承認を確認。理研およびかずさのメタボローム・データを整理中であることを報告。公的リポジトリで欧米とのデータ交換を確認。
2020年2月13日	国立遺伝学研究所 (センター長 有田正規)	委員11名 (オブザーバ数不明)	第37回 DNA データ研究利用委員会 メタボローム・データのリポジトリである MetaBoBank の公開準備を報告。DDBJ ホームページで Wiki を運営する困難も説明。Wiki は学会等に変更させなければ可能ではないという指摘(伊藤剛委員)。
2021年3月2日	オンライン	29人	第38回 DNA データ研究利用委員会 DDBJ サービスに関する専門部会の設置を決定。MetaboBank Version1 の状況を報告。
2022年2月17日	オンライン	25人	第39回 DNA データ研究利用委員会 DDBJ サービスに関する専門部会の進捗を報告。
2023年2月13日	オンライン	25人	第41回 DNA データ研究利用委員会 DDBJ サービスに関する専門部会の進捗を報告。

(2) DDBJ サービスに関する専門部会

名称	DDBJ サービスに関する専門部会
目的	専門部会は、委員会の諮問に応じ、次の各号に掲げる事項について専門的に調査・検討する。 1. DDBJ サービスの妥当性、有効性、安全性、効率に関すること 2. DDBJ サービスの設置と廃止に関すること 3. その他、生命情報・DDBJ センターの運営に関して必要な事項
委員数	8人(2021)

委員一覧

氏名	所属機関	役職	備考
石濱泰	京都大学 薬学研究科	教授	専門部会長
三枝大輔	帝京大学 薬学部	准教授	
平山明由	慶應義塾大学 先端生命科学研究所	特任講師	
松田文生	大阪大学 情報科学研究科	教授	

鈴木謙一	アステラス製薬 創薬推進研究所	主任研究員	2021 年度 で退任
大川友之	塩野義製薬 新薬研究所	主任研究員	
益田勝吉	サントリーグローバルイノベーション研究センター	主席研究員	
岩橋福松	住友化学株式会社 健康農業関連事業研究所	主任研究員	

開催歴

年月日	場所	参加人数	主な議題
2021 年 6 月 24 日	オンライン	19 人	第 1 回 DDBJ サービスに関する専門部会議事: DDBJ サービスに関する専門部会の運営について等
2021 年 7 月 26 日	オンライン	17 人	第 2 回 DDBJ サービスに関する専門部会議事: 「統合化推進プログラム」中間評価について等
2021 年 9 月 8 日	オンライン	9 人+ 外部	統合化推進プログラム サイトビジット 課題: 「物質循環を考慮したメタボロミクス情報基盤」
2021 年 10 月 21 日	オンライン	21 人	第 3 回 DDBJ サービスに関する専門部会議事: MetaboBank Ver.2 の開始について等
2022 年 9 月 5 日	オンライン	不明	第 4 回 DDBJ サービスに関する専門部会議事: MetaboBank、MetaboBank Wiki の進捗、国内外との連携、MetaboBank Ver.2 における入力必須項目について

§3. 研究開発の目的、実施内容及び成果

1. 研究開発の背景

研究を開始した 2018 年度当時、国内にメタボローム研究の生データリポジトリは無かった。当時日本が世界に発信していたリソースは、代表者らが作成に関わり JST のサポートによって国際的な広がりをもせた MassBank である。日本で 2006 年に開始されて以降、ドイツ Norman-Network による欧州版 MassBank、日米 SICORP 事業で作成した MoNA (MassBank of North America) といったスペクトルサーバが世界に普及した。従来、質量分析計による代謝物同定に用いるスペクトルライブラリは有償が基本であった。それを CC ライセンスを付与したデータ投稿により集積し、無償で配布する仕組みとして MassBank はデータ共有のさきがけを作った。

MassBank が対象としたデータは市販される標品を測定したスペクトルであり、生体サンプルを測定した生データではない。ゲノミクスと異なりメタボロミクス分野には主要な機器ベンダーが6社以上あるうえ、各社の生データ形式は測定装置にも依存する。結果として、測定生データは再利用が極めて難しい状況にあった。しかし欧米では主に研究の公正さを担保するためのデータ共有義務付けが開始されており、米国では保健省のセンターグラントによる Metabolomics Workbench (<https://www.metabolomicsworkbench.org/>)、欧州ではバイオインフォマティクス研究所による MetaboLights (<https://www.ebi.ac.uk/metabolights/>) が、恒久的なデータ・アーカイブとして生データの保管を開始していた。しかし発足直後から、測定条件の詳細をデータ投稿者に義務付けることが極めて難しいこと、複雑な記載を達成してもなお再利用が難しいことが明らかにされていた (Salek et al. 2013 Database)。

この状況において、代表者らは当初生データのリポジトリを時限プロジェクトとして開始することは極めて困難と判断していた。そのため本研究の前年度に、研究分担者の開発する KNApSACk データベースと MassBank スペクトルを統合する「メタ代謝マップ」を構築する提案書を提出したが採択されなかった。当時の JST 事業総括の意見では、日本に必要なのは国内研究グループを束ねて研究振興に資するリポジトリの開発であり、マススペクトルや代謝マップのような二次的なリソースではないというものであった。この意見は極めてまっとうであり反論の余地がないものであった。そのため、翌年度に MetaboLights および Metabolomics Workbench に並ぶ日本からの生データリポジトリとして MassBank を拡張する申請書を提出し、採択された。

当時より MassBank は日本質量分析学会の公式サーバとして機能しており、採択時点ではサーバの構築を本事業で引き受け、新しく wiki サーバを併設して学会によるデータ入力を受け付ける想定であった。しかし、開始直後の 2018 年 5 月、学会側から本研究開発計画に賛同できない旨の意見が提出された。JST も含めた調整の結果、ドメインの継続性の観点から massbank.jp ドメインおよび MassBank の名称はすべて学会側の所有であるため、譲渡することになった。また、学会が利用するサーバとしてドイツ Norman MassBank と同じシステムを希望したため、2018 年度は massbank.jp ドメインに学会側が希望するサーバを本研究にて構築し、稼働させた状態で全て譲渡する手続きを取った。一連の処理が終了したのは 2019 年 4 月であり、その後 MetaboBank という名称で別システムを構築開始した。

2. 研究開発対象のデータベース・ツール

(1) データベース

- 主要なもの

正式名称	略称	概要
MetaboBank	MetaboBank	DDBJ センター事業の一環として実施するメタボローム・データ リポジトリ。Metabolonote および RIKEN PMM に記載されるメ タデータを含む。

- 上記以外のもの

正式名称	略称	概要
RIKEN Plant Metabolomics MetaDatabase	RIKEN PMM	理研 CSRS で取得した植物メタボロームデータを体系的に 整理して公開するサービス。本データベースでは、メタデータの世界標準技術であるセマンティックウェブや Resource Description Framework (RDF) と呼ぶ枠組みに基づいてデータを公開している。 http://metabobank.riken.jp/
食品・植物・万物メタボロームリポジトリ	食レポ・植レポ・万物レポ	食品と植物にあるメタボロームピークを検索, 確認できるデータベース http://metabolites.in/foods /plants と /things
KNApSAcK	ナップザック	様々なデータを含むが, 本研究ではアルカロイド等二次代謝物の生合成マップおよび植物界における二次代謝物の分布をデータベース化。

(2) ツール等

正式名称	略称	概要
MS-DIAL MS-FINDER		質量分析データ解析用ソフトウェア。MS-DIAL は代謝物由来イオンの検出、同一化合物由来イオンのグルーピング、代謝物 アノテーション等の一連のデータ処理を行う。MS-FINDER は未 知のスペクトルから構造解析(structure elucidation)するための ソフトウェアツール。
PowerGetBatch		メタボロームデータ解析ツール http://www.kazusa.or.jp/komics/ja/tool-ja/235-powergetbatch.html

※データベース、ツールの詳細は別紙参照。

3. 達成目標及び実施計画

(1) 当初の実実施計画・達成目標

申請当初は遺伝研、かずさ、理研、奈良先の4グループからなる実施計画であり、第3年次までに以下の目標を達成する予定であった。

○遺伝研（研究代表者）グループ

- ・ メタボローム・実測データの公共リポジトリを新設し、データの受付を開始。
- ・ 欧米の MassBank サーバーが収載するスペクトルを精査し、MassBank コンソーシアムにおける一次データと二次データを整理。
- ・ 二次データにバージョンを付与し、MassBank ライブラリ 2020 といった名称でデータ頒布を開始。必要なライセンスを整備。
- ・ データを metabolomeXchange に登録し、MetaboLights, MetabolomicsWorkbench に並ぶ国際リポジトリの発足。
- ・ MassBank 事務局を開設し、事務局担当者の人件費は自己収入で維持できる仕組みの実現。

○かずさグループ

- ・ 保有する公開可能なメタボローム測定データ（114 生物種 3000 サンプル以上）を精査して、実験的に同定推定された代謝物リストを作成、公開。
- ・ 代謝物リストを MassBank より公開。再アノテーションにより研究開始時のアノテーション数より 3 割多い化合物（またはその派生物）の同定を実現。
- ・ 遺伝研グループが作成する公共リポジトリに、実測データのコンテンツを登録。
- ・ 公共リポジトリのデータ解析ソフトウェアとして PowerGet 等をアップデートし、MassBank グループとして提供。

○理研グループ

- ・ かずさグループ同様、保有する公開可能なメタボローム測定データを精査して、MassBank より公開する。再アノテーションにより研究開始時のアノテーション数より 3 割多い化合物（またはその派生物）の同定を実現。
- ・ 遺伝研グループが作成する公共リポジトリに、実測データのコンテンツを登録。
- ・ 公共リポジトリのデータ解析ソフトウェアとして MS-DIAL 等をアップデートし、MassBank グループとして提供。

○奈良先グループ

- ・ 代謝物の生体活性を整理し、代謝物の母核構造と生理活性の関連（構造活性 オントロジー）を作成。
- ・ 得られた母核構造のうち、まだ収集していない物質に対して文献 から新たに情報を取得し、構造活性関係の妥当性を検証。文献情報やオントロジーは KNApSACK より公開。
- ・ 作成したオントロジーの例示となるメタ代謝マップを 20 例作成し、MassBank より公開。

(2) 期間中に追加・削除・変更した実施計画・達成目標

MassBank を日本質量分析学会に譲渡したため、マススペクトルに関しては MassBank が投稿を受け付けるサーバとして残された。MetaboBank はスペクトルに関して MassBank と連携し、生

データを専門に受け付けるリポジトリとして役割分担が生じた。開発するサーバ等の名称は全て MetaboBank に統一し、MassBank という名称は利用しないことになった。

4. 実施内容

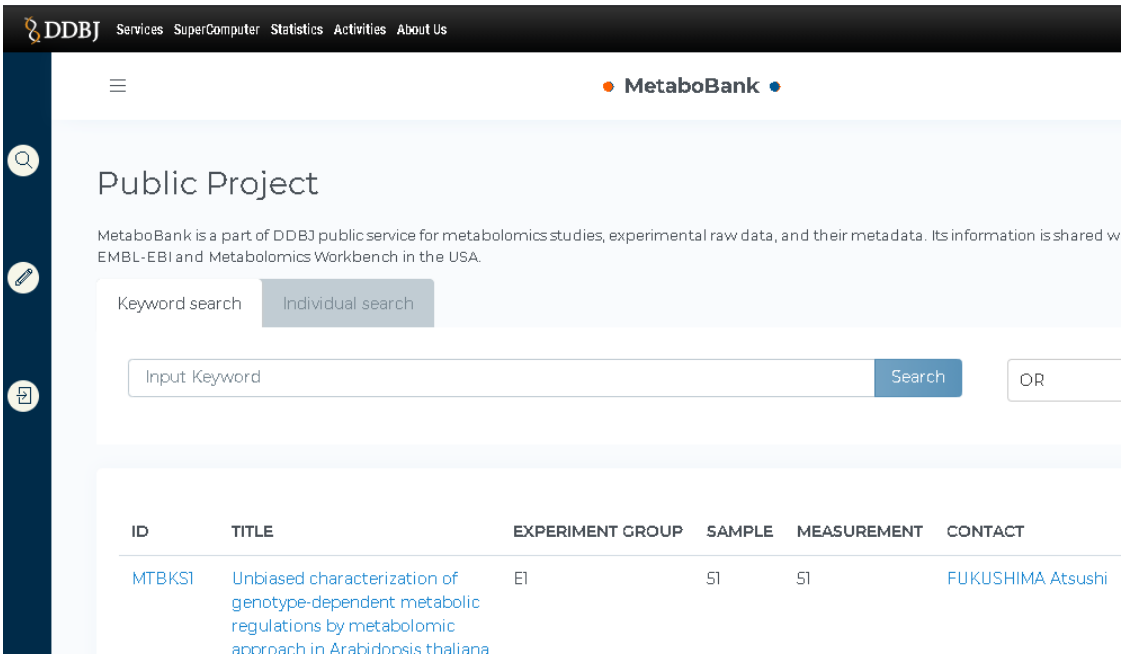
(1) 実施内容

生データのリポジトリとして最も重要なのはメタデータの記載法や MassBank を含むほかデータとの連携法である。統合化推進プログラムでは登録データを RDF 形式で公開することが義務付けられているため、利用するオントロジーの策定から開始する必要があった。

○2019-2020 年度

MetaboBank の主サーバは遺伝研グループが開発担当であったが、そのプロトタイプとして理研グループが「植物メタボロームメタデータベース(理研 PMM)」を作成することとした。理研では情報システム本部が理研全体の公開データを統合する「理研メタデータベース」の構築を進めており、本計画の分担機関を含め各部署は公開データを「理研メタデータベース」(<https://metadb.riken.jp/>) に登録する必要があったためである。このメタデータベースは登録内容を RDF として公開する機能を備えていたため、本計画の方向性とも一致するものであった。そのため、理研グループが保有するメタボローム生データについて、そのメタデータを理研 PMM として公開するべく作成を開始した。

遺伝研では研究開始時にかずさ DNA 研究所に所属していた櫻井望が国立遺伝学研究所に移籍したため、かずさと遺伝研グループを橋渡しする重要な役割を担うことになった。理研データは理研 PMM として整理を進める方針となったため、主にかずさデータのメタデータ作成と RDF 化を実施する中心的役割を櫻井が担った。



ID	TITLE	EXPERIMENT GROUP	SAMPLE	MEASUREMENT	CONTACT
MTBKSI	Unbiased characterization of genotype-dependent metabolic regulations by metabolomic approach in Arabidopsis thaliana	EI	SI	SI	FUKUSHIMA Atsushi

Figure 1 MetaboBank Version 1 のトップページ

当初の公開アドレスは <http://mb.ddbj.nig.ac.jp>。MetaboBank が Version 2 に移行し現在は閉鎖した。

その結果として 2020 年 10 月に公開したのが MetaboBank リポジトリの Version 1 (図 1)およびエクセルによる入力エディタ(図2)である。初期データとして理研およびかずさグループで整理

してきた情報を収載し、塩基配列と重ならないアクセッション番号を付与した。

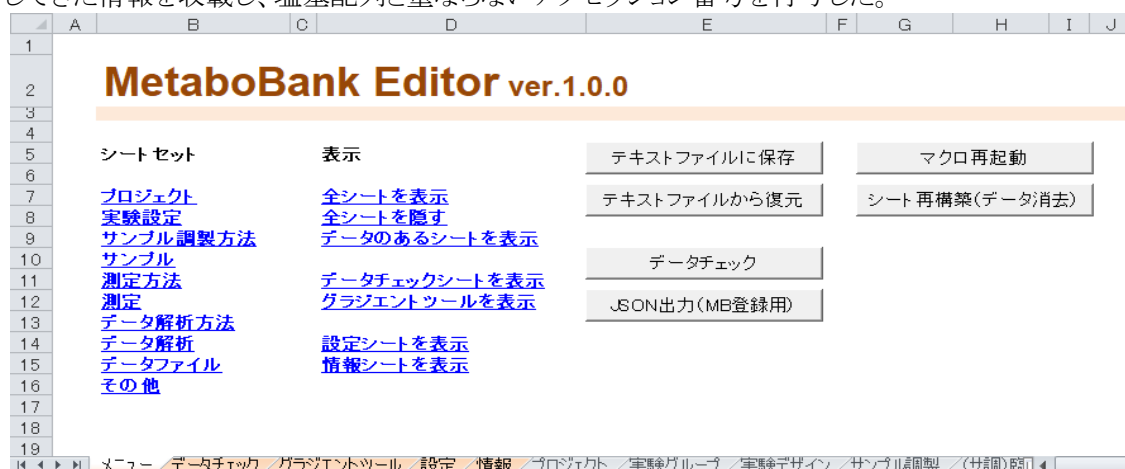


Figure 2 MetaboBank Version 1 のエクセル・入力エディタ。その機能や編集方法は、羊土社「メタロミクス実践ガイド」(馬場ほか編, 2021) [その他著作 1]にも詳述された。

Version1 システムは国際メタロミクス学会の標準イニシアティブ (MSI)が推奨するガイドラインを完全に満たしている点で、欧米のリポジトリよりも再利用性が高くなっていた。メタデータが詳しくれば詳しいほど実験条件の詳細がわかる。付属のエクセル・エディターはクロマトグラフィー条件を GUI 形式で入力・登録させる機能まで有しており、論文等では文字情報で書かれ概要を掴みにくい条件を視覚的に確認できた。このツールを用いた入力はフォーマットが揃うため、レポジトリ上でグラジエントの類似性を比較・検索することも可能となった。

同時並行で理研グループが開発したのが理研 PMM である。植物の GC/MS データを中心にメタデータを作成、登録・公開した。理研メタデータベースの機能を利用したが、生データおよび解析済みデータも格納しダウンロード可能とした(図3)[論文 23]。



Figure 3 理研 PMM のトップページ。http://metabank.riken.jp/ で公開していたが、現在サーバ不正アクセスのため公開を停止している。

これらのデータベースに格納するデータを整理したのは理研およびかずさグループのメンバーである。2020 年末には理研より 56 件、かずさより 104 件のメタロームデータセットを公開できた。

奈良先グループはその間に KNApSAcK データベースの構築を続け、代謝物数は 53,032 種、生物種と代謝物の関係として 128,771 組を収載するまでになった。また代謝物情報をも

とに、オントロジーを作成し、活性情報の集積を、KNpSack Metabolite Activity DB として公開した (<http://www.knapsackfamily.com/MetaboliteActivity/top.php>)。2020 年度

Group of SS	Starting substance (SS)	3-phospho glycerate		Pyruvate		Phosphoenol pyruvate			Oxaloacetate		alpha-Ketoglutarate			Terpenes			TCA cycle	Fatty acid	Nucleic acids																			
		Gly	L/Ser	D/Ser	L-Cys	D-Ala	L-Ala	D/Len	L/Val	D/Trp	L/Trp	L/Phe	Anthranilate	N Formyl anthranilate	Indole-3-glycerol phosphate	O-Methyl-L-tyrosine	L/Asp	L/Thr	L/Lys	L/Arg	L/His	L/Pro	L/Glu	IPP	DMAPP	Scopolanin	GGPP	Cholesterol	Acetyl CoA	Oxaloacetate	Malonyl CoA	Acetyl CoA	Adenine					
Arg, Pro, Asp, Pyrrolidine/Pyrrrolidine Alkaloids	1	●																																				
Trp, Ergot Alkaloids	2							●																														
Trp, Monoterpenoid Indole Alkaloids	3							●																														
Trp + DMAPP, Trp + Pro: Trp + Oxaloacetate: Trp (Simple Indole Alkaloids), Trp + His	4							●																														
Trp + DMAPP, Trp + Pro: Trp + Oxaloacetate: Trp (Simple Indole Alkaloids), Trp + His	5							●																														
Trp + DMAPP, Trp + Pro: Trp + Oxaloacetate: Trp (Simple Indole Alkaloids), Trp + His	6							●																														
Trp + DMAPP, Trp + Pro: Trp + Oxaloacetate: Trp (Simple Indole Alkaloids), Trp + His	7							●																														
Trp + DMAPP, Trp + Pro: Trp + Oxaloacetate: Trp (Simple Indole Alkaloids), Trp + His	8							●																														
Trp + DMAPP, Trp + Pro: Trp + Oxaloacetate: Trp (Simple Indole Alkaloids), Trp + His	9							●																														
Trp + DMAPP, Trp + Pro: Trp + Oxaloacetate: Trp (Simple Indole Alkaloids), Trp + His	10							●																														
Tyr: Betanin group: Tyr + Phe, Calchicine	11																																					
Tyr: Isoquinoline Alkaloids (Benzylisoquinoline Alkaloids)	12																																					
Tyr + Scopolanin (Tetrahydroisoquinoline monoterpene alkaloid), Tyr + Ser	13																																					
Tyr + Scopolanin (Tetrahydroisoquinoline monoterpene alkaloid), Tyr + Ser	14																																					
Lys: Quinolizidine, Piperidine Alkaloids, Indolizidine Alkaloids Lycoperdium Alkaloids	15																																					
Phe + Ser	16																																					
Ala + Acetyl-CoA + Malonyl-CoA	17																																					
His: Imidazole alkaloid(histidine derivatives)	18																																					
Anthranilate(Acridine Alkaloids)	19																																					
Anthranilate + Trp + Anthranilate + Trp + Ala: Anthranilate + Phe (PKS)	20																																					
Anthranilate + Trp + Anthranilate + Trp + Ala: Anthranilate + Phe (PKS)	21																																					
Anthranilate + Trp + Anthranilate + Trp + Ala: Anthranilate + Phe (PKS)	22																																					
Anthranilate + Trp + Anthranilate + Trp + Ala: Anthranilate + Phe (PKS)	23																																					
Anthranilate + Trp + Anthranilate + Trp + Ala: Anthranilate + Phe (PKS)	24																																					
Anthranilate + Trp + Anthranilate + Trp + Ala: Anthranilate + Phe (PKS)	25																																					
Anthranilate + Trp + Anthranilate + Trp + Ala: Anthranilate + Phe (PKS)	26																																					
Anthranilate + Trp + Anthranilate + Trp + Ala: Anthranilate + Phe (PKS)	27																																					
Anthranilate + Trp + Anthranilate + Trp + Ala: Anthranilate + Phe (PKS)	28																																					
Anthranilate + Trp + Anthranilate + Trp + Ala: Anthranilate + Phe (PKS)	29																																					
Indole-3-glycerol phosphate (Indole-terpene alkaloids)	30																																					
Steroidal alkaloids	31																																					
Purine alkaloids	32																																					

Figure 4 KNA p Sack データベースに記載されたアルカロイドの生合成とアミノ酸の関係表。
<http://www.knapsackfamily.com/CobWeb/top.php> 左側のパスウェイ名をクリックすると、クリック可能なメタ代謝マップに移動する。

には活性種として 10852 件の代謝物と活性の関係を整理し、140 の包括的活性種情報に分類した [論文 44-46]。

また奈良先グループはアルカロイドの生合成経路を、生物間相互作用を含む物質循環として示したメタ代謝マップを 30 枚以上作成し公開した(図4)。67 報の原著論文をもとに 515 種のアルカロイドの生合成を示してある。この情報をもとに、共生菌種と植物についての新たな関係性を取得できるようになった。

○2021 年度-2022 年度

2020 年末の公開により、かずさグループが Metabolonote として公開してきたメタデータ情報はほぼ MetaboBank サーバ上に移行できた。ただ適切な食品オントロジーがないために、食品関連のみ RDF 化ができず MetaboBank への移行が完了していなかった。そのため食品データの取り扱いが課題として残された。

まだ RDF 化を実施できない食品等データの公開先として作成されたのが、櫻井による食品メタボロームレポジトリ (<http://metabolites.in/foods/>)、植物メタボロームレポジトリ (<http://metabolites.in/plants/>) および、万物メタボロームレポジトリ (<http://metabolites.in/things/>) である [国内招待講演 2-7]。これらはあらゆる物質のメタボロームをクロマトグラフィー結果から比較可能なように同一条件で分析し、精密質量やスペクトルで検索可能にして再解析を容易にしたサイトである。二次元クロマトグラムおよびスペクトルを全て公開するかわりに、ピークのアノテーションはユーザが API を通じて検索や同定を実施する仕組みになっている(図5)。研究終了時点において NBRP リソー

スの計測数もあわせて 1000 件以上の測定結果を公開している。これらのデータは RDF 化の整備やメタデータの調整を実施しており、MetaboBank にまとめて収載を予定している。

食品メタボロームレポジトリ



Figure 5 食品メタボロームレポジトリのトップページ。上のバナーから食品一覧を選ぶことで生データのダウンロードも可能。

食品オントロジーの策定以上に問題として浮上したのが、MetaboBank Version 1 のシステム構成であった。Java Spring という新しい機能を用いて構成されていたが、その改修とメンテナンスが DDBJ に所属するスタッフおよびキュレータには難しかった。そのため、DDBJ が所有する既存システムに近い構成で再構築せざるを得なくなり、MetaboBank Version 2 を新たに構築、公開した。Version 2 では投稿受付の内部手順、担当者を定め、実験計画は BioProject データベース、試料情報は BioSample データベースに記載する。これらは公的リポジトリである DDBJ が塩基配列情報向けにサービスする内容と共通しており、オミックス研究の登録に適しているうえ研究内容を検索しやすくする。

Version 2 への移行に伴い、メタデータ形式も変更せざるをえなかった。これは実験計画や試料情報を別データベースと連携させることにとどまらず、詳しく記載したメタデータの区分けを他の一次レポジトリ（とりわけ遺伝子発現量を収載する Genomic Expression Archive）と揃える必要も生じた。メタデータ入力はトランスクリプトーム解析のデータ登録と同じ形式とし、実験デザインは IDF、サンプル情報は SDRF、そして同定した代謝物リストを MAF と呼ばれる形式でエクセルを介して登録するシステムとした。形式の詳細は <https://www.ddbj.nig.ac.jp/gea/metadata.html> における記載と共通化し、案内および登録法などは <https://www.ddbj.nig.ac.jp/metabobank/> より公開になった（図 6）。

また、最終年度に KNApSACk データベースと MetaboBank を統合するインターフェースととして、MediaWiki を基盤とした Wiki システムを構築した。MediaWiki の最新バージョンにこれまで MassBank Wiki として構築してきた機能を移植し、代謝物情報

として KNApSAcK データベースの中身を投入している。サンプルページは <https://mb.metabolomics.jp/wiki/Compound:C00000001> にある。

Wiki には化合物、スペクトル、およびマップ情報を収載した。以下にその一部をリストする。

- MetaboBank メタデータ
<https://mb.metabolomics.jp/wiki/Category:MetaboBankStudy>
- KNApSAcK マップデータ
<https://mb.metabolomics.jp/wiki/CobWeb:No1>
- MS-DIAL スペクトルデータ
<https://mb.metabolomics.jp/wiki/Spectrum2:Splash10-014i-0090000000-37388e20bda94bb654f1>

DDBJ サービス スパコン 統計 活動 センターについて

DDBJ Web Sites

D-way 停止 10月12日 10:00-12:00

MetaboBank

Home Submission FAQ Search Download Contact

登録の流れ

1. 登録アカウントの取得
2. 登録申し込み
3. BioProjectの登録
4. BioSampleの登録
5. メタデータの作成
6. データファイルの準備
7. ファイルのアップロード

アクセッション番号
データ公開
Reviewer access
更新
MDS チェックサム値

ホーム > [metabobank](#) > MetaboBank への登録

MetaboBank への登録

登録の流れ

MetaboBank は、塩基配列や遺伝子発現データと関連づけられるように BioProject と BioSample と連携しています。メタデータは実験種別毎に用意されたエクセルファイル（以下に記載するMAGE-TAB の IDF と SDRF）に記載します。解析済みのデータ（同定・推定された化合物に関する情報）は所定の形式である Metabolite assignment file (MAF) に記載します。

- 1. 登録アカウントの取得**
 - D-way 登録アカウントを作成します。マニュアル
 - 公開鍵と center name をアカウントに登録し MetaboBank 登録を可能にします。
- 2. 登録申し込み**

MetaboBank 登録申し込みフォーム から登録を申し込みます。申し込み内容に応じて担当者が個別に登録方法をご案内します。フォームにアクセスできない場合、申し込みファイルをダウンロードし、必要事項を記入のうえ metabobank@dbj.nig.ac.jp にメールで申し込んでください。
- 3. BioProjectの登録**

研究プロジェクトの内容を BioProject に登録します。プレフィックス PRJOB のアクセッション番号がプロジェクトに対して発行されます。
- 4. BioSampleの登録**

データを得るために使われた試料情報を BioSample に登録します。メタボロミクス用サンプルには Omics パッケージを使用します。非生物サンプルの場合、生物名には NCBI Taxonomy [metagenomes](#) などから適切なものを選んでください。

Figure 6 新しい MetaoboBank のデータ登録ページ。初心者でもわかるように詳述しており、登録フォーマットも DDBJ の他のリポジトリと揃えられた。

かずさグループは再解析を順調に進め、MassBase および Metabolonote データベースに収載された合計 104 の実験セット(スタディと呼ぶ。植物 79, 微生物:3, 食品:16, その他:6。生物の種類数は 157)を整備した。またそれらを 3,428 の解析生データと紐付し、メタデータは Resource Description Framework (RDF) の書式に変更した。

また 2021 年度、2022 年度には理研グループの2研究員がそれぞれ独立し、農工大グループと府大グループを構成することになった。担当業務は理研グループにおける内容と一致しているため、大きな変更はない。農工大グループは MS-DIAL ソフトウェアを開発している。全研究期間を通じてこのソフトウェアは大きく発展し、2022 年時点で Version5 の公開に至っている。MS-DIAL は月間ダウンロード数が千を超える国際的にも人気のソフトウェアに成長した [論文 20,21,31 など多数]。

理研がソフトウェアを公開するウェブサイトは、月間訪問者数 7000 ユニーク IP アドレスに成長した。

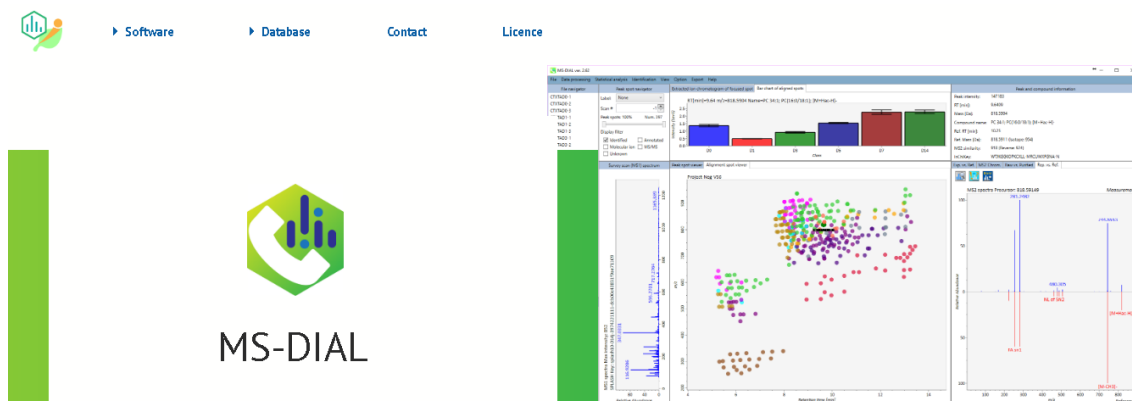


Figure 7 MS-DIAL の公開ページ <http://prime.psc.riken.jp/compms/msdial/main.html>。スペクトルライブラリもこのページからダウンロードできる。

府大グループは遺伝研および理研グループと連携して、データの再解析とリアノテーションを継続している。植物関連代謝物の再解析および他オミックスデータとの統合作業を円滑に進めるため、京都府立大学に計算サーバを複数台設置し、データ解析環境を整えた。再解析対象とした理研 GC-MS データ 54 プロジェクト分について、4 つの解析ソフトウェア (MS-DIAL, eRah, AI-Output, IP4M) を用いて再アノテーションを進め、MS-DIAL では 50 プロジェクト分が完了 (解析済み割合 93%), eRah および AI-Output については共に 46 プロジェクトが完了 (解析済み割合 85%) している。

(2) 実施内容のうちの特定項目の詳細

① 研究コミュニティを含むデータ提供者との連携・協業

○統合化推進プログラム内での連携

植物ゲノム統合データベース Plant GARDEN (<https://plantgarden.jp>)

微生物ゲノム統合データベース MicrobeDB.jp (<https://microbedb.jp>)

と密接に連携している。

とりわけ 2021 年度は統合化推進プログラム追加実施予算を受け、一般利用者が SPARQL 言語を入力せずに SPARQL 検索を実施できる検索システムを DBCLS および NBDC からの協力を得て作成している。検索システムは、同年に DBCLS から公開された TogoDX システムを利用し、植物・微生物・メタボロームの主要データを格納した TogoDX-TPP を新たに構築した。格納データは次の 6 種になる。

1. 植物(系統、ゲノム配列、遺伝子・化合物キュレーション)、
2. 微生物(系統、BioSample)、
3. メタボローム(KNApSACk の化合物、生物種、文献情報)、
4. DDBJ(系統、BioProject、BioSample)、
5. TogoGenome(系統、ゲノム)、
6. 化合物データ(EBI CheBI、NCBI PubChem)

検索の効率化を図るため、利用者が TogoDX-TPP に格納されているデータを予め検索することができるユーザインタフェース TPP APP(開発版)を作成した。現在、システム開発を委託したインフォ・ラウンジ社のテストサーバ上 (<https://tpp-app.il3c.com>) にて公開しているが、システムの修正が終わり次第、かずさ DNA 研究所のウェブサーバから公開する予定である。

○質量インフォマティクス研究会、日本バイオインフォマティクス学会との連携

本課題に従事するメンバーは国内で質量インフォマティクス研究を実施する中核でもある。それもあるって質量インフォマティクス研究会 (<https://ms-bio.info/>) と、とりわけソフトウェア講習会や国際シンポジウムにおいて連携しながら研究開発を続けてきた。また代表者は日本バイオインフォマティクス学会の幹事や理事も務めるため、学会と連携したチュートリアルセッション等を複数企画した。

○ナショナルバイオリソースプロジェクト NBRP との連携

万物メタボロームリポジトリ (<http://metabolites.in/things/>) は、NBRP の各リソースセンターと連携して、バイオリソースのメタボローム情報を公開する作業を開始した。2022 年時点ではトマトデータが多いが、今後はより多くのバイオリソース情報を充実させる予定である。

② データベース利用者への周知、利用者との連携・協業

MetaboBank Version2 の公開が 2021 年度になったため、本格的な周知活動は 2022 年度より開始している。これまでに日本質量分析学会、メタボロームシンポジウム等で紹介した。

1. 株式会社 Reifycs との包括連携協定
主要な質量分析計ベンダーの生データ形式を読み取る技術を持つ Reifycs 社と遺伝研は

2022 年度に包括連携協定を結び、MetaboBank に提供されるデータを同社の HIVE フォーマットに変換、MetaboBank より公開する予定としている。HIVE フォーマットは ABF フォーマットの後継にあたる。現在の Metabobank は生データを ABF フォーマットでも公開しているが、これを同社の協力のもとに HIVE 形式へ移行を決定し、2023 年度より HIVE フォーマットでもデータを提供予定である。

2. JFE エンジニアリングとの共同研究

2020 年度、株式会社 JFE エンジニアリングはメタボロミクス分野の受託解析を実施する計画を立て、国立遺伝学研究所、群馬大学、横浜国立大学との四者間共同研究を開始した。JFE エンジニアリングが受託したデータを MetaboBank に円滑に登録する取り決めを進めていたが、社内の事業担当者が秘密情報を社外に持ち出して独立する不祥事にいたり共同研究は 2021 年度に解消せざるを得なかった。

3. ツムラとの共同研究

2018-2020 年度に国立遺伝学研究所と株式会社ツムラは漢方処方メタボローム解析に関して共同研究を開始した。製薬会社が自社製品の有効性を研究発表で示すことが難しい世相のため、ツムラの漢方処方データを公的なリポジトリ MetaboBank より公開する計画もあった。しかしデータ公開を懸念する社内意見を説得しきれず、共同研究成果は論文発表にとどまった。

4. さくら科学との包括連携協定

遺伝研の櫻井は「さくら科学」という受託分析会社を立ち上げ (<https://sakura-kagaku.com/corp/>)、一般からの分析依頼を受け付けている。その成果を万物メタボロームレポジトリ等に登録することで、民間セクターからの情報を公開する努力を続けている。

③ 利用者にとって有用なデータ基盤の構築

上記「①研究コミュニティを含むデータ提供者との連携・協業」に記載。RDF サーバを介して連携する研究グループとはデータ共有が達成できている。また理研、かずさ、遺伝研が測定するメタボロミクス生データはいずれも大規模で再利用価値が高い内容である。

④ 持続的なデータベース運用体制の構築に向けた取り組み

当初目的にあった「MetaboBank の事務局」を設置することはできなかった。ただ国立遺伝学研究所の「DNA データ研究利用委員会」「運営協議会」の承認を通じて MetaboBank リポジトリが DDBJ 事業として認められている。今後も DDBJ サービスの一環としてメタボローム生データを受け付け、保管することは担保できた。

今後、DDBJ 事業と連携する企業コンソーシアムを設置して協賛金を募ることで、DDBJ 事業の一部運営費を賄えるように努力する。

⑤ 人材の育成

理研グループの研究者が独立した研究室主催者となるなど、アカデミアの育成に関して顕著な成果を挙げている。また多くの学生がメタボロミクス研究に携わった。

⑥ 国際連携・国際貢献

○欧州 EBI、INSDC との連携

MetaboBank v2 の登録フォーマットは発現解析で用いられている MAGE-TAB を基に作成しており、European Bioinformatics Institute (EBI) の MetaboLights 担当チームとは登録情報についてオンラインで協議を継続している。MetaboBank は DDBJ サービスの一貫となるため、MetaboLights のメタデータを MetaboBank 内で閲覧可能にする予定である。

また、MetaboBank に登録した公開データには BioProject、BioSample レコードからリンクを張れる。そのため BioProject、BioSample を介して、NCBI、EBI 管轄の ENA のウェブサイトから MetaboBank データが閲覧可能になる。

○GNPS 国際連携への参加

米国カリフォルニア大学サンディエゴ校が中心となって推進する FBMN (Nature methods 17, 905-908, 2020) や IIMN (Nature communications 12, 2021) プログラムに、MS-DIAL ソフトウェアが参加している。それらのデータ規格に直接インポート可能な出力を、MS-DIAL 側より提供している。

○LSI(Lipidomics Standards Initiative)への貢献

メタボロミクスの一分野であるリポドミクスは、他分野のように標準品が充実していない点や正確な構造推定が難しいことから、分子構造や名称を含む成果の記述形式が研究者によってばらばらしている。その国際標準化を目指したコンソーシアム LSI に参加している。

○PubChem との連携

2022 年 12 月より、KNApSack Core DB は PubChem に組み込まれ、双方向リンクとなり、世界で認知されたデータベースとなった。

<https://pubchem.ncbi.nlm.nih.gov/source/KNApSack%20Species-Metabolite%20Database>
本プロジェクトが世界でも有効活用された例といえる。

⑦ その他

DDBJ のアナテーター、メタボロミクス、ゲノミクスの研究者、ソフトウェア開発者など MetaboBank の開発には様々なバックグラウンドの人々が関わっており、研究の意義やデータの扱いについて異なる意見が多かった。そのため開発の優先順位や問題意識の共有が困難な場面が多く見られた。コロナ禍により対面の会議が困難な状況となり拍車がかかった。このためデータベースの構成を刷新した MetaboBank v2 の開発開始以降(2021 年度以降)、チーム内の進捗会議を月 1 回から 2 回にオンラインではあるが増やした。同様に 2021 年度の有識者会議も 6 月から 10 月にかけて毎月実施し、問題点の洗い出しに努めた。

§4. 主要なデータベースの利活用状況

1. アクセス数

(1) 実績

表 1 研究開発対象の主要なデータベースの利用状況(月間平均)

名称	種別	2018 年度	2019 年度	2020 年度	2021 年度	2022 年度
MetaboBank Version 2	訪問者数	公開前	公開前	公開前	57	NA
	訪問数	公開前	公開前	公開前	NA	NA
	ページ数	公開前	公開前	公開前	NA	NA
MetaboBank Version 1	訪問者数	公開前	公開前	公開前	51	閉鎖
	訪問数	公開前	公開前	公開前	NA	閉鎖
	ページ数	公開前	公開前	公開前	NA	閉鎖

(2) 分析

- MetaboBank Version1 は 2020 年 10 月より開始し 2022 年初まで開設していた。ユニーク IP アドレス数は月平均 50 アドレスしか無く、ほとんど訪問されなかった。
 - MetaboBank Version2 は 2021 年 10 月より開始し、Version 1 と同程度のアクセス数であった。2022 年度に入ってデータ登録用ページを作成し、月平均 70 アドレスほどに増えている。

2. データベースの利用状況を示すアクセス数以外の指標

第三者より3プロジェクトを受付済みでアクセッション番号 MTBKS201 ー216 を発行した。2022 年の登録件数は 3 プロジェクト、4 スタディ。
メタボバンク登録件数は以下より公開。

<https://docs.google.com/spreadsheets/d/16ZF79i1X17Zfn3x6vnj2elmWXb3ToHt9nZIDTtg-zGA/edit#gid=987043383>

3. データベースの利活用により得られた研究成果(生命科学研究への波及効果)

○ターゲットメタボロームの解析によるオカラミンの同定 [文献 13]

第三回メタボロミクスソフトウェア講習会の動画を YouTube にアップしてあり、現在までに視聴回数が 700 を超えている。ターゲット分析から新規の化合物を見いだせたことで、データを集積する意義を端的に示した。

4. データベースの利活用によりもたらされた産業への波及効果や科学技術のイノベーション(産業や科学技術への波及効果)

○MS-DIAL による 10 倍以上の脂質分子同定 []

MS-DIAL 4 では、Lipidomics Standard Initiative (LSI) に準拠した脂質解析が可能で、世界中のユーザが代表的なメタボロミクス・リポドミクスソフトウェアと認識するツールに成長した。イオンモビリティ、オールイオンフラグメンテーション、電子誘起分解 (EAD) など、様々な新技術にも対応している。

§5. 今後の展開

1. MetaboBank の宣伝につとめ、できる限り質の高い、大規模で再利用価値のあるデータセットを寄託してもらおう。その再解析を積極的に実施することで、バイオインフォマティクスが既存のゲノム情報に基づいて研究を実施するように、公開データに基づくメタボロミクスを可能とする。
2. 遺伝研 DDBJ サービスを支援する企業コンソーシアムを設立し、賛助金などの収入を確保する。その収入によってキュレーション業務等の一部を賄えるようにする。
3. 現在も理研 PMM、万物メタボロームレポジトリなど、複数のシステムが存在する。しかし公共リポジトリと呼べる仕組みは MetaboBank にしかないため、生データの一次リポジトリとして MetaboBank の地位を確立する。その他のデータベースやサイトは、一次リポジトリに準拠する形にそろえていく。

§6. 自己評価

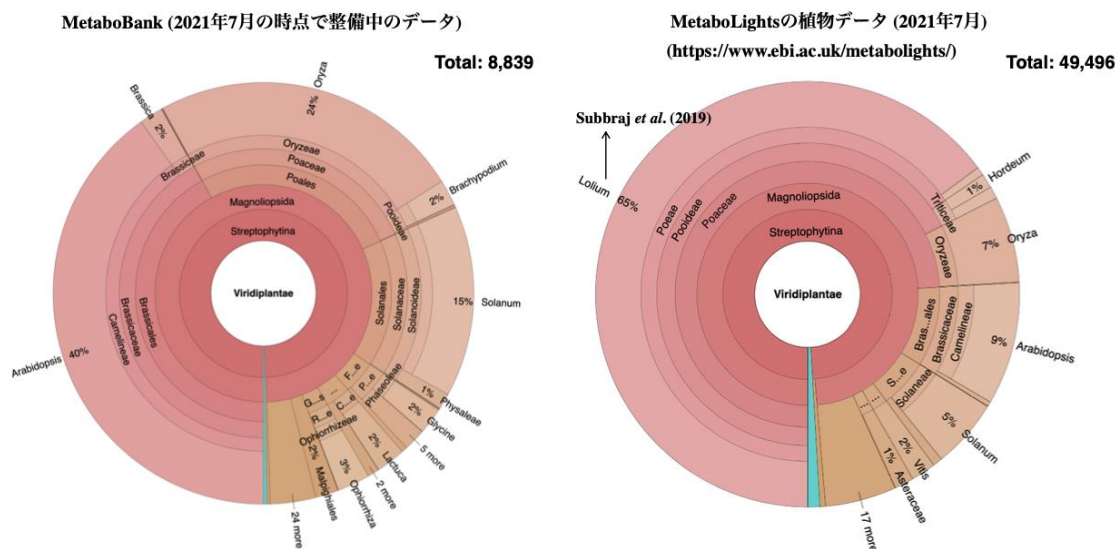
○MetaboBank リポジトリ構築の達成状況と代表者自身の評価

2018 年度に始まった研究計画だが、一般の研究者がデータを投稿できるプラットフォームの構築と公開が 2021 年度にずれ込んでしまった。そのため実質的な登録依頼や宣伝が 2022 年度になり、当初計画よりも大幅に遅れてしまった。

遅れの主要因として代表者の研究マネジメントの至らなさもあるが、世の中全体におけるリポジトリへの理解の薄さも大きな原因と考えている。リポジトリの構築や運営は、論文成果を重視する研究活動とは大きく隔たる活動になる。研究者の多くは論文を受理してもらうためにデータをリポジトリに登録するだけで、論文化されたら後の面倒を見ようとはしない。再利用を可能とするアノテーションを付与し、データを維持管理する努力は受け付けるキュレータの肩にかかる。しかし一般にキュレータの地位は低く、研究者からも活動があまり評価されていない。結果としてキュレータ業務の観点から必要となるシステム開発と、研究者が成果を求めて実施する開発には大きな隔たりがある。この違いを 2018 年度より正しく認識できていなかったため、本プロジェクトは研究面（例えば論文、ソフトウェア開発、研究者の育成）では優れた結果を残せた反面、将来持続可能なリポジトリの構築という点では不十分な結果となってしまった。具体的には自己収入を得る手段を構築できておらず、第三者からのデータ登録数も始まったばかりという状況である。

○理研とかずさによる再アノテーションの状況

かずさと理研の植物サンプル数を EBI の MetaboLights の植物サンプル数を比較すると MetaboBank の総数が 8,839 に対して MetaboLights が 49,496 と数としては MetaboLights が約 5.6 倍になる。しかし MetaboLights のデータの 65%である約 3,200 が *Lolium* (ドクムギ属) のサンプルである(kazusa_Fig1)。このため、数では劣るものの MetaboBank の植物データ構成は、国際的にも有用なリソースと考える(下図はリポジトリにおける植物種分類)。

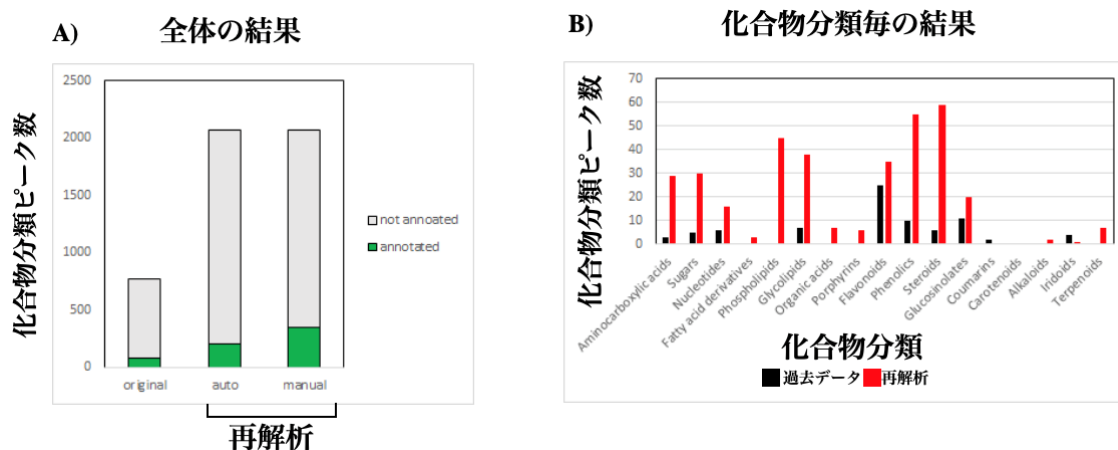


○当初目標とした再アノテーション効率について

データの再解析はマスマスペクトルのシグナルピーク検出とデータ間のアライメントという 2 つのステップをもとに実施する。そのためデータ量が多くなればアライメントによる計算機負荷は大きいものの同定精度を向上させられる。当プロジェクトでは、スパコンを用いた再解析を実施した。結果を

人手で精査してシグナル感度に関するパラメータの数値調整をおこない、化合物分類できたピーク数は自動解析で 2 倍、手動解析後は 4 倍に増加させられた(下図の左緑色部分)。

また、シロイヌナズナの葉の LC-FT-ICR による化合物分類結果ではアミノ酸、糖、脂質、フェノー



ル性化合物、ステロイドなど多様な分類群でアノテーションできたピーク数が増加した(上図右の赤色部分)。現在はゲノムワイド関連解析(GWAS)によって、化合物と特定のイネ品種の遺伝子座との関連性を見出せるデータセット(MetaboBank 登録データ: MTBKS52)の再解析をおこなっており、概算であるが過去の発表論文(Matsuda et al., 2015)で決定された代謝産物の 3 倍強の 300 個の代謝産物を同じデータから特定できている。これらデータの再解析はこれから発表する予定である。

§7. 外部発表等

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)	44 件
未発行論文	国内(和文)	0 件
	国際(欧文)	2 件

(2) 論文詳細情報

1. Burla B, Arita M, Arita M, Bendt AK, Cazenave-Gassiot A, et al. “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)
2. Fukushima A, Hikosaka S, Kobayashi M, Nishizawa T, Saito K, et al. “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”, *Front Plant Sci.* 9:1439, 2018 (DOI: 10.3389/fpls.2018.01439)
3. Kitazaki K, Fukushima A, Nakabayashi R, Okazaki Y, Kobayashi M, et al. “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)
4. K Liu, AH Morita, S Kanaya and Md. Altaf-Ul-Amin “Metabolite-content-guided prediction of medicinal/edible properties in plants for bioprospecting” *Curr Res Complement Altern Med* 130, 130, 2018 (DOI: 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 Bioinform* 19(1), 264, 2018 (DOI: 10.1186/s12859-018-2251-x)
6. Tada I, Tsugawa H, Meister I, Zhang P, Shu R et al. “Creating a Reliable Mass Spectral-Retention Time Library for All Ion Fragmentation-Based Metabolomics” *Metabolites* 9(11), 251, 2019 (DOI: 10.3390/metabo9110251)
7. Lipidomics Standards Initiative Consortium “Lipidomics needs more standardization” *Nat Metabolism* 1(1), 745-747, 2019 (DOI: 10.1038/s42255-019-0094-z)
8. Tsugawa H, Satoh A, Uchino H, Cajka T, Arita M, Arita M “Mass Spectrometry Data Repository Enhances Novel Metabolite Discoveries with Advances in Computational Metabolomics” *Metabolites* 9(6), 119, 2019 (DOI: 10.3390/metabo9060119)
9. Tsugawa H, Nakabayashi R, Mori T, Yamada Y, Takahashi M, et al. “A chemoinformatics approach to characterize metabolomes in stable-isotope-labeled organisms” *Nat Methods* 16(4), 295-298, 2019 (DOI: 10.1038/s41592-019-0358-2)
10. Vos RA, Katayama T, Mishima H, Kawano S, Kawashima S, et al. “BioHackathon 2015: semantics of data for life sciences and reproducible research” *F1000 Res* 9(1), 136, 2020 (DOI: 10.12688/f1000research.18236.1)
11. A Fukushima, T Kuroha, K Nagai, Y Hattori, M Kobayashi, et al. "Metabolite and

- Phytohormone Profiling Illustrates Metabolic Reprogramming as an Escape Strategy of Deepwater Rice during Partially Submerged Stress", *Metabolites* 10, 68 (2020). (DOI: 10.3390/metabo10020068)
12. NT Vu, K Kamiya, A Fukushima, S Hao, W Ning, et al. "Comparative co-expression network analysis extracts the SIHSP70 gene affecting to shoot elongation of tomato", *Plant Biotechnol* 36, 143-153, 2019 (DOI: 10.5511/plantbiotechnology.19.0603a)
 13. Sakurai N, Mardani-Korrani H, Nakayasu M, Matsuda K, Ochiai K, et al. "Metabolome analysis identified okaramines in the soybean rhizosphere as a legacy of hairy vetch" *Front Genet* 11, 114 2020 (DOI: 10.3389/fgene.2020.00114)
 14. Ohbuchi K, Sakurai N, Kitagawa H, Sato M, Suzuki H, et al. "Differential annotation of converted metabolites (DAC-Met): Exploration of Maoto (Ma-huang-tang)-derived metabolites in plasma using high-resolution mass spectrometry" *Metabolomics* 16(6), 63, 2020 (DOI: 10.1007/s11306-020-01681-3)
 15. Hiraga Y, Ara T, Nagashima Y, Shimada N, Sakurai N, et al. "Metabolome analysis using multiple data mining approaches suggest luteolin biosynthesis in *Physcomitrella patens*" *Plant Biotechnol (Tokyo)* 37(3), 377-381, 2020 (DOI: 10.5511/plantbiotechnology.20.0525b)
 16. Kusano M, Fukushima A, Tabuchi-Kobayashi M, Funayama K, et al. "Cytosolic GLUTAMINE SYNTHETASE 1;1 modulates metabolism and chloroplast development in roots", *Plant Physiol* 182(4):1894-1909 (2020). (DOI: 10.1104/pp.19.01118)
 17. Miyazaki Y, Ono N, Huang M, Md.Altaf-Ul-Amin, Kanaya S, "Comprehensive exploration of target-specific ligands from compounds using a graph convolution neural network" *Mol Inf*, 39:e1900095 (2020) doi: 10.1002/minf.201900095
 18. Hijikata A, Shionyu-Mitsuyama C, Nakae S, Shionyu M, Ota M, Kanaya S, Shirai T. "Knowledge-based Structural Models of SARS-CoV-2 Proteins and Their Complexes With Potential Drugs" *FEBS Lett*, 594: 1960-1973 (2020) doi: 10.1002/1873-3468.138
 19. Ono N, Eguchi R, Morita AH, Katsuragi T, Nakamura S, Huang M, Md Altaf-Ul-Amin, Kanaya S "Classification of alkaloids according to the starting substances of their biosynthetic pathways using graph convolutional neural networks" *BMC Bioinf*, 20, 380, (2019) doi: 10.1186/s12859-019-2963-6
 20. Tada I, Chaleckis R, Tsugawa H, Meister I, et al. "Correlation-based deconvolution (corrdec) to generate high-quality MS2 spectra from data independent acquisition in multisample studies" *Anal Chem* 92(16), 11310-11317, 2020 (DOI: 10.1021/acs.analchem.0c01980)
 21. Tsugawa H, Ikeda K, Takahashi M, Satoh A, Mori Y, et al. "A lipidome atlas in MS-DIAL4" *Nat Biotechnol* 38(10), 1159-1163, 2020 (DOI: 10.1038/s41587-020-0531-2)
 22. Takeshi Ara, Nozomu Sakurai, Hideyuki Suzuki, Koh Aoki, Kazuki Saito, Daisuke Shibata, "MassBase: A large-scaled depository of mass spectrometry datasets for metabolome analysis" *Plant Biotechnol*. 38(1): pp.167–171, 2021 (DOI: 10.5511/plantbiotechnology.20.0911a).
 23. Atsushi Fukushima, Mikiko Takahashi, Hideki Nagasaki, Yusuke Aono, Makoto Kobayashi, Miyako Kusano, Kazuki Saito, Norio Kobayashi, and Masanori Arita, "Development of RIKEN Plant Metabolome MetaDatabase" *Plant Cell Physiol*. 63(3), pp.433–440. 2022 (DOI: 10.1093/pcp/pcab173).
 24. Mostafa Abdelrahman, Nur Aeni Ariyanti, Yuji Sawada, Fumitada Tsuji, Sho Hirata, Tran Thi Minh Hang, Mami Okamoto, Yutaka Yamada, Hiroshi Tsugawa, Masami Y

- okota Hirai, Masayoshi Shigyo. Metabolome-Based Discrimination Analysis of Shallot Landraces and Bulb Onion Cultivars Associated with Differences in the Amino Acid and Flavonoid Profiles. *Molecules* **25**, 5300, 2020.
25. Hiroyuki Yamamoto, Eisuke Hayakawa, Hiroshi Tsugawa, Yuki Moriya, Eiichiro Fukusaki, Susumu Goto, Tomohisa Hasumura, Nobuaki Miura, Akiyasu C Yoshizawa. Japan Computational Mass Spectrometry Meeting 2020 Activity Report. *Journal of Proteome Data and Methods* **2**, 5, 2020.
 26. Naoto Katakami, Kazuo Omori, Naohiro Taya, Shoya Arakawa, Mitsuyoshi Takahara, Taka-aki Matsuoka, Hiroshi Tsugawa, Masahiro Furuno, Takeshi Bamba, Eiichiro Fukusaki, Iichiro Shimomura. Plasma metabolites associated with arterial stiffness in patients with type 2 diabetes. *Cardiovascular Diabetology* **19**, 1-8, 2020.
 27. Shu Yasuda, Nobuyuki Okahashi, Hiroshi Tsugawa, Yusuke Ogata, Kazutaka Ikeda, Wataru Suda, Hiroyuki Arai, Masahira Hattori, Makoto Arita. Elucidation of gut microbiota-associated lipids using LC-MS/MS and 16S rRNA sequence analyses. *iScience* **23**, 101841, 2020.
 28. Ophelie Fraiser-Vannier, Justine Chervin, Guillaume Cabanac, Virginie Puech-Pages, Sylvie Fournier, Virginie Durand, Aurelien Amiel, Olivier Andre, Omar AbdelAziz Benamar, Bernard Dumas, Hiroshi Tsugawa, Guillaume Marti. MS-CleanR: A feature-filtering approach to improve annotation rate in untargeted LC-MS based metabolomics. *Analytical Chemistry* **92**, 9971-9981, 2020.
 29. Kazuma Uesaka*, Hiroya Oka*, Ryuji Kato*, Kei Kanie*, Takaaki Kojima*, Hiroshi Tsugawa*, Yosuke Toda*, Takaaki Horinouchi*. Bioinformatics in bioscience and bioengineering: Recent advances, applications, and perspectives. *Journal of Bioscience and Bioengineering* in press 2022.
 30. Fumika Mi-Ichi, Hiroshi Tsugawa, Makoto Arita, Hiroki Yoshida. Pleiotropic Roles of Cholesteryl Sulfate during Entamoeba Encystation: Involvement in Cell Rounding and Development of Membrane Impermeability. *mSphere* **7**: e00299-22, 2022.
 31. Haruki Uchino, Hiroshi Tsugawa*, Hidenori Takahashi, Makoto Arita*. Computational mass spectrometry accelerates C=C position-resolved untargeted lipidomics using oxygen attachment dissociation. *Communications chemistry*, accepted, 2022.
 32. Zhixu Ni , Geoff Jukes , Karla Mendivelso Espinosa , ... , Hiroshi Tsugawa , et al. Guiding the choice of informatics software and tools for lipidomics biomedical research applications. *Nature Methods*, accepted, 2022.
 33. Jeremy P Koelmel, Wan Y Tan, Yang Li, ..., Hiroshi Tsugawa, et al. Lipidomics and Redox Lipidomics Indicate Early Stage Alcohol - Induced Liver Damage. *Hepatology Communications* doi.org/10.1002/hep4.1825, 2021.
 34. Pei Zhang, Christopher Carlsten, Romanas Chaleckis, ..., Hiroshi Tsugawa, et al. Defining the Scope of Exosome Studies and Research Needs from a Multidisciplinary Perspective. *Environmental science & technology letters* **8**, 839-852, 2021.
 35. Robin Schmid, Daniel Petras, Louis-Félix Nothias, Mingxun Wang, ..., Hiroshi Tsugawa, et al. Ion identity molecular networking for mass spectrometry-based metabolomics in the GNPS environment. *Nature Communications* **12**: 3832, 2021.
 36. Nobuyuki Okahashi, Masahiro Ueda, Shu Yasuda, Hiroshi Tsugawa, Makoto Arita. Global profiling of gut microbiota-associated lipid metabolites in antibiotic-treated mice by LC-MS/MS-based analyses. *STAR protocols* **2**:100492, 2021.
 37. Fumika Mi-Ichi, Kazutaka Ikeda, Hiroshi Tsugawa, Sharmina Deloer, Hiroki Yoshida, Makoto Arita. Stage-Specific De Novo Synthesis of Very-Long-Chain Dihydroceramid

- es Confers Dormancy to Entamoeba Parasites. *Mosphere* **6**: e00174-21, 2021.
38. Hiroyuki Yamamoto, Yasumune Nakayama, Hiroshi Tsugawa. OS-PCA: Orthogonal S smoothed Principal Component Analysis Applied to Metabolome Data. *Metabolites* **11**: 149, 2021.
 39. Hiroshi Tsugawa, Amit Rai, Kazuki Saito, Ryo Nakabayashi. Metabolomics and complementary techniques to investigate the plant phytochemical cosmos. *Natural Product Reports* **38**, 1729-1759, 2021.
 40. Yuki Matsuzawa, Yasuhiro Higashi, Kouji Takano, Mikiko Takahashi, Yutaka Yamada, Yozo Okazaki, Ryo Nakabayashi, Kazuki Saito, Hiroshi Tsugawa*. Food Lipidomics for 155 Agricultural Plant Products. *Journal of Agricultural and Food Chemistry* **69**, 8981–8990, 2021.
 41. Amit Rai, Hideki Hirakawa, Ryo Nakabayashi, Shinji Kikuchi, Koki Hayashi, Megha Rai, Hiroshi Tsugawa, et al. Chromosome-level genome assembly of *Ophiorrhiza pumila* reveals the evolution of camptothecin biosynthesis. *Nature communications* **12**, 1-19, 2021.
 42. Kusano M, Worarad K, Fukushima A, Kamiya K, Mitani Y, et al. Transcriptomic, Hormonomic and Metabolomic Analyses Highlighted the Common Modules Related to Photosynthesis, Sugar Metabolism and Cell Division in Parthenocarpic Tomato Fruits during Early Fruit Set. *Cells* **11**(9), 1420, 2022
 43. Desmet S, Saeys Y, Verstaen K, Dauwe R, Kim H, Niculaes C, Fukushima A, et al. Maize specialized metabolome networks reveal organ-preferential mixed glycosides. *Comput Struct Biotechnol J*, 19, 1127-1144, 2021
 44. Nasution AK, Wijaya SH, Gao P, Islam RM, Huang M, Ono N, Kanaya S, Altaf-Ul-Amin M. Prediction of Potential Natural Antibiotics Plants Based on Jamu Formula Using Random Forest Classifier. *Antibiotics (Basel)*, 11(9):1199, 2022
 45. Karim MB, Kanaya S, Altaf-Ul-Amin M. Antibacterial Activity Prediction of Plant Secondary Metabolites Based on a Combined Approach of Graph Clustering and Deep Neural Network. *Mol Inform.* 41(7):e2100247, 2022
 46. Okada T, Namiki T, Tohge T, Kanaya S. Cheminformatics modeling of the correlation between Bupleurum Root-formula medicines and Excess and Deficiency pattern in the diagnostic criteria of Sho in Kampo (traditional Japanese medicine) by non-targeted direct infusion mass spectrometry with machine learning. *J Nat Med.* 76(1):306-313, 2022

2. その他の著作物(総説、書籍など)

1. 櫻井望、津川裕司、「メタボロミクス実践ガイド」実践編 1, 3, 4, 9, 10 章 (馬場、津川ほか編)、羊土社 2021
2. 有田正規、金谷重彦、櫻井望 「食品分野におけるメタボリックプロファイリング」1-1, 1-2, 2-2 節 (福崎 編) NTS 2021
3. Miyako Kusano and Atsushi Fukushima. Plant Metabolomics: The Great Potential of Plant Metabolomics in Big Data Biology. In "Plant Omics : Advances in Big Data Biology (Edited by Hajime Ohayanagi et al.)". CABI, Nosworthy Way, Wallingford OX10 8DE, UK. (in press).

3. 国際学会発表及び主要な国内学会発表

(1) 概要

種別	国内外	件数
招待講演	国内	7件
	国際	2件
口頭発表	国内	7件
	国際	4件
ポスター発表	国内	13件
	国際	3件

(2) 招待講演

〈国内〉

1. 金谷重彦 「コロナの時代の食品データサイエンス、データベースの体系化、深層学習」日本バイオインフォマティクス学会年会、9/2, 2020、オンライン
2. 櫻井望 「有望な未知成分をトップダウンに探索するための古くて新しいデータ基盤」名古屋大学×島津製作所 最新質量分析 Webinar Bioinformatics が拓く、質量分析の夜明け - 質量分析と Bioinformatics の融合からみえる世界- 2021年2月17日
3. 櫻井望 「未活用な植物特化代謝成分を トップダウンに探索するための データ基盤の構築」日本植物生理学会年会 シンポジウム 2021年3月15日
4. 櫻井望 「有用な化合物を トップダウンに探索するための メタボロームデータ基盤の構築」日本分析化学会 表示・起源分析技術研究懇談会 第24回講演会 2021年3月17日
5. 櫻井望 「食品成分の網羅解析(高付加価値化ツール)」2020年度グローバル化のためのSDGs勉強会 2021年3月26日
6. 櫻井望 「未知食品成分の同定と活用に向けたデータ基盤の構築」第169回 質量分析関西談話会 2021年9月4日
7. 櫻井望 「LC-MS で取得したノンターゲットメタボロームデータから有用情報を引き出すコツ」山口大学総合科学実験センター年次セミナー「代謝物情報解析の現状と展望」2021年11月16日
- 8.

〈国際〉

1. Masanori Arita “Computational Aspects in Metabolomics and Lipidomics” Virtual Podium in Metabolomics and Lipidomics in Asia Pacific region (online) 2020年11月13日
2. Masanori Arita “Computational aspects in metabolomics” Online Lecture at INBIOSI S, Malaysia, Mar 26, 2021

(3) 口頭講演

〈国内〉

1. 長崎英樹、大澤祥子、荒武、福島敦史、高橋みき子、藤澤貴智、小林紀郎、櫻井望、平川英樹、有田正規 メタボローム統合データベース MetaboBank 構築に向けた植物メタボローム解析メタデータの RDF 化と測定生データの再解析: RDF Conversion of Plant Metabolome Metadata and Raw Data Reanalysis Toward Built of MetaboBank, the Metabolome Integrated Database 第 38 回日本バイオテクノロジー学会(つくば)大会、筑波大学、2021 年 9 月 10 日
2. 長崎英樹、荒武、福島敦史、大澤祥子、高橋みき子、藤澤貴智、時松敏明、児玉悠一、福田亜沙美、諏訪和夫、小林紀郎、櫻井望、金谷重彦、平川英樹、有田正規 植物由来のメタボローム測定生データの再解析とメタボローム統合データベース MetaboBank の構築: Reanalysis of Metabolome Raw Data from Plants and Construction of MetaboBank, an Integrated Metabolome Database 第 39 回日本バイオテクノロジー学会(堺)大会、大阪公立大、2022 年 9 月 13 日
3. 櫻井望 メタボロームデータの公共レポジトリ MetaboBank 第 69 回質量分析総合討論会 2021 2021 年 5 月 20 日
4. 福島敦史、高橋みき子、長崎英樹、小林誠、草野都、斉藤和季、小林紀郎、有田正規、"植物メタボロームデータの標準化と共有に向けた統合メタデータベースの開発", 第 38 回日本植物バイオテクノロジー学会(つくば)大会、オンライン、2021 年 9 月 10 日
5. 長崎英樹、メタボローム統合データベース MetaboBank 構築に向けた植物メタボローム解析メタデータの RDF 化と測定生データの再解析、第 38 回日本バイオテクノロジー学会つくば 2021、筑波大学、9 月 10 日
6. 福島敦史、高橋みき子、長崎英樹、小林誠、草野都、斉藤和季、小林紀郎、有田正規、"植物メタボロームデータの標準化と共有に向けた統合メタデータベースの開発", 第 63 回日本植物生理学会年会、オンライン、2022 年 3 月 23 日
7. 第 15 回メタボロームシンポジウム、2021 年 10 月 14 日
 - 1) 小野直亮 "深層学習を用いた二次代謝物質の合成経路予測と活性予測"
 - 2) 福島敦史 "メタボロームデータの標準化と共有のための統合メタデータベース開発"
 - 3) 津川裕司 "質量分析データ解析プログラム MS-DIAL の現状と課題"

〈国際〉

1. 津川裕司、Computational mass spectrometry to deepen the understanding of metabolisms、ASHBi Seminar (オンライン)、2022 年 2 月 14 日
2. 津川裕司、Decoding mass spectrometry data to understand the metabolisms of living organisms、BioC Asia 2021 (オンライン)、2021 年 11 月 1 日
3. 津川裕司、MS-DIAL、LIPIDMAPS spring school (オンライン)、2021 年 4 月 15 日

(4) ポスター発表

〈国内〉

1. 福島敦史、高橋みき子、櫻井望、時松敏明、長崎英樹、平川英樹、荒武、小林誠、草野

- 都, 齊藤和季, 有田正規, 小林紀郎, " RIKEN Plant Metabolome MetaDatabase の開発", 第 13 回メタボロームシンポジウム 筑波大学, 2019 年 10 月 16-18 日
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」、トーゴの日シンポジウム 2018、東京、2018 年 10 月 5 日(金)
 3. 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)
 4. 長崎英樹、大澤祥子、荒武、福島敦史、高橋みき子、小林紀郎、櫻井望、平川英樹、有田正規 植物・食品メタボローム解析メタデータの RDF 化と測定生データの再解析 第 13 回メタボロームシンポジウム、筑波大学、2019 年 10 月 16-18 日
 5. 長崎英樹、大澤祥子、荒武、福島敦史、高橋みき子、小林紀郎、櫻井望、平川英樹、有田正規 植物・食品メタボローム解析メタデータの RDF 化および測定生データの再解析に向けて トーゴの日シンポジウム 2020、オンライン、2020 年 10 月 5 日
 6. 平川英樹、藤澤貴智、森宙史、ゲルフィ アンドレア、市原寿子、中村保一、磯部祥子、田畑哲之、黒川頭、生物間相互作用の解明に向けた植物・微生物統合データベース間の連携、トーゴの日シンポジウム 2020、オンライン、2020 年 10 月 5 日
 7. 長崎英樹、荒武、福島敦史、大澤祥子、高橋みき子、小林紀郎、藤澤貴智、櫻井望、平川英樹、有田正規、有田正規 植物メタボローム解析メタデータの RDF 化および測定生データの再解析 第 62 回日本植物生理学会年会、オンライン、2021 年 3 月 16 日
 8. 長崎英樹、荒武、大澤祥子、福島敦史、高橋みき子、小林紀郎、藤澤貴智、時松敏明、櫻井望、金谷重彦、平川英樹、有田正規 メタボローム統合データベース MetaboBank の開発、トーゴの日シンポジウム 2021、オンライン、2021 年 10 月 5 日
 9. 平川英樹、藤澤貴智、長崎英樹、森宙史、福島敦史、ジェルフィ アンドレア、市原寿子、中村保一、金谷重彦、有田正規、黒川頭、磯部祥子、田畑哲之、植物ゲノム統合データベース Plant GARDEN における微生物、メタボローム統合データベースとの連携、トーゴの日シンポジウム 2021、オンライン、2021 年 10 月 5 日
 10. 長崎英樹、大澤祥子、荒武、福島敦史、高橋みき子、藤澤貴智、時松敏明、小林紀郎、櫻井望、平川英樹、有田正規 メタボローム統合データベース MetaboBank 構築と測定生データの再解析、第 15 回メタボロームシンポジウム、オンライン、2021 年 10 月 14-15 日
 11. 長崎英樹 統合データベース MetaboBank は超分野植物科学の夢を見るか? 第 1 回超分野植物科学研究会、オンライン、2021 年 6 月 4 日
 12. 長崎英樹、荒武、福島敦史、高橋みき子、大澤祥子、小林紀郎、藤澤貴智、時松敏明、福田亜沙美、櫻井望、諏訪和夫、金谷重彦、平川英樹、有田正規 メタボローム統合データベース MetaboBank、トーゴの日シンポジウム 2022、オンライン、2022 年 10 月 5 日
 13. 平川英樹、藤澤貴智、守屋勇樹、信定知江、長崎英樹、森宙史、福島敦史、Andrea Ghelfi、市原寿子、中村保一、金谷重彦、有田正規、黒川頭、田畑哲之、磯部祥子、植物ゲノム統合データベース Plant GARDEN、微生物、メタボローム統合データベース間の連携検索システムの開発、トーゴの日シンポジウム 2022、オンライン、2022 年 10 月 5 日

〈国際〉

1. **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

The 15th International Conference of the Metabolomics Society, Hague, June 23-27, 2019

2. Katsuya Obuchi, Nozomu Sakurai, Hiroyuki Kitagawa, Hirotaka Kushida, Akinori Nishi, Masahiro Yamamoto, Kazuhiro Hanazaki, Masanori Arita "DAC-Met : Exploring the metabolic fate of the herbal components in "maoto" decoction through a differential annotation strategy"
3. Nozomu Sakurai, Kunihiro Suda "Food Metabolome Repository: A database for cross-sample specificity-based peak prioritization in untargeted metabolomics"

4. 知財出願

(1) 出願件数

種別	件数	
特許出願	国内	0 件
	国外	0 件
その他の知的財産出願		0 件

(2) 一覧

- ①国内出願
- ②海外出願
- ③その他の知的財産権

5. 受賞・報道等

(1) 受賞

- ・ 理研 CSRS 奨励賞、高橋みき子, 山田豊、2020 年 6 月 10 日
- ・ 理研梅峰賞, 津川裕司, 2020 年 3 月 12 日
- ・ 国立遺伝学研究所森島奨励賞, 多田一風太, 2020 年 9 月 29 日
- ・ 総合研究大学院大学研究科長賞, 多田一風太, 2020 年 9 月 29 日

(2) メディア報道

櫻井望「メタボローム解析を用いた玄米中の未知成分の探索」東洋ライス株式会社 記者発表 2022 年 8 月 8 日

(3) その他

§8. 研究開発期間中の活動

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 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
2019年 5月30-31日	チーム内ミーティング (非公開)	遺伝研	7人	メタボロームリポジトリ MetaboBank の研究開発進 捗報告のためのミーティング
2019年 7月18日	チーム内ミーティング (非公開)	かずさ DNA 研究所	9人	研究開発進 捗報告のための ミーティング
2019年 7月26日	チーム内ミーティング (非公開)	かずさ DNA 研究所	4人	研究開発進 捗報告のための ミーティング
2019年 8月29日	チーム内ミーティング (非公開)	理研東京連 絡事務所	6人	メタボロームリポジトリ MetaboBank の研究開発進 捗報告のためのミーティング
2019年 9月25日	チーム内ミーティング (非公開)	理研横浜	8人	同上
2019年 11月28日	チーム内ミーティング (非公開)	遺伝研	8人	同上
2020年 2月4日	チーム内ミーティング (非公開)	理研横浜	8人	同上
2020年 4月21日	チーム内ミーティング (非公開)	オンライン開催	9人	同上
2020年 5月21日	チーム内ミーティング (非公開)	オンライン開催	9人	同上
2020年 6月25日	チーム内ミーティング (非公開)	オンライン開催	10人	同上
2020年 8月24日	チーム内ミーティング (非公開)	オンライン開催	10人	同上
2020年 11月19日	チーム内ミーティング (非公開)	オンライン開催	10人	同上

2021年 1月15日	チーム内ミーティング (非公開)	同上	10人	同上
2021年 2月2日	チーム内ミーティング (非公開)	同上	10人	同上
2021年 4月20日	チーム内ミーティング (非公開)	同上	4人	MetaboBank 開発進捗報告の ためのミーティング
2021年 5月10日	チーム内ミーティング (非公開)	同上	5人	同上
2021年 5月19日	チーム内ミーティング (非公開)	同上	4人	同上
2021年 6月2日	チーム内ミーティング (非公開)	同上	4人	同上
2021年 6月16日	チーム内ミーティング (非公開)	同上	4人	同上
2021年 6月22日	チーム内ミーティング (非公開)	同上	4人	同上
2021年 6月30日	チーム内ミーティング (非公開)	同上	5人	同上
2021年 7月5日	チーム内ミーティング (非公開)	同上	5人	同上
2021年 7月14日	チーム内ミーティング (非公開)	同上	6人	同上
2021年 7月20日	チーム内ミーティング (非公開)	同上	7人	同上
2021年 7月27日	チーム内ミーティング (非公開)	同上	6人	同上
2021年 8月3日	チーム内ミーティング (非公開)	同上	7人	同上
2021年 8月10日	チーム内ミーティング (非公開)	同上	4人	同上
2021年 8月18日	チーム内ミーティング (非公開)	同上	7人	同上
2021年 8月25日	チーム内ミーティング (非公開)	同上	6人	同上
2021年 9月2日	チーム内ミーティング (非公開)	同上	10人	同上
2021年 9月8日	チーム内ミーティング (非公開)	同上	8人	同上
2021年 9月21日	チーム内ミーティング (非公開)	同上	7人	同上
2021年 9月27日	チーム内ミーティング (非公開)	同上	8人	同上
2021年 10月13日	チーム内ミーティング (非公開)	同上	11人	同上
2021年 10月27日	チーム内ミーティング (非公開)	同上	13人	同上
2021年	チーム内ミーティング	オンライン開催	7人	MetaboBank 開発進捗報告の

11月1日	(非公開)			ためのミーティング
2021年 11月4日	チーム内ミーティング (非公開)	同上	9人	同上
2021年 11月24日	チーム内ミーティング (非公開)	同上	10人	同上
2021年 12月8日	チーム内ミーティング (非公開)	同上	11人	同上
2021年 12月22日	チーム内ミーティング (非公開)	同上	11人	同上
2022年 1月13日	チーム内ミーティング (非公開)	同上	10人	同上
2022年 1月27日	チーム内ミーティング (非公開)	同上	5人	MetaboBank wiki 関連ミーテ ィング
2022年 2月3日	チーム内ミーティング (非公開)	同上	8人	MetaboBank 開発進捗報告の ためのミーティング
2022年 3月16日	チーム内ミーティング (非公開)	同上	9人	同上
2022年 3月30日	チーム内ミーティング (非公開)	同上	7人	同上
2022年 4月13日	チーム内ミーティング (非公開)	同上	8人	同上
2022年 4月27日	チーム内ミーティング (非公開)	同上	9人	同上
2022年 5月11日	チーム内ミーティング (非公開)	同上	9人	同上
2022年 5月25日	チーム内ミーティング (非公開)	同上	9人	同上
2022年 6月8日	チーム内ミーティング (非公開)	同上	9人	同上
2022年 6月22日	チーム内ミーティング (非公開)	同上	10人	同上
2022年 7月6日	チーム内ミーティング (非公開)	同上	3人	MetaboBank Wiki/KNApSAc K ミーティング
2022年 7月13日	チーム内ミーティング (非公開)	同上	10人	MetaboBank 開発進捗報告の ためのミーティング
2022年 7月27日	チーム内ミーティング (非公開)	同上	8人	同上
2022年 8月10日	チーム内ミーティング (非公開)	同上	6人	同上
2022年 8月31日	チーム内ミーティング (非公開)	同上	10人	同上
2022年 9月21日	チーム内ミーティング (非公開)	同上	9人	同上
2022年 10月12日	チーム内ミーティング (非公開)	同上		同上
2022年 11月9日	チーム内ミーティング (非公開)	同上		同上

2022年 12月8日	チーム内ミーティング (非公開)	同上		同上
2023年 1月24日	チーム内ミーティング (非公開)	同上		同上
2023年 2月22日	チーム内ミーティング (非公開)	同上		同上
2023年 3月15日	チーム内ミーティング (非公開)	同上		同上
2023年 3月29日	チーム内ミーティング (非公開)	同上		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

n	名称	場所	参加人数	目的・概要
2019年10月17日	13回メタボロームシンポジウム 特別セミナー	筑波大学	80名	NBDC 統合プロジェクトにおけるメタボローム活動の紹介
2019年6月22日	LipoQuality International Workshop on Lipidomics	理研横浜	60名	新学術領域リポクオリティ, シンガポール国立大学と合同でリポミクスの技術に関する国際ワークショップ
2020年9月2日	日本バイオインフォマティクス学会年会 (IIBMP2020)チュートリアル	オンライン	各セッション 200名	DDBJ 主催の利用チュートリアル(3セッション)においてメタボロミクスも紹介
2020年10月22-29日	日本インドネシアバイオインフォマティクス・バイオリソースワークショップ	オンライン	30名	DDBJ 主催の利用チュートリアル(英語。2時間ずつ5日)においてメタボロミクスも紹介
2021年7月1日	第3回メタボロミクスソフトウェア講習会 ONLINE	オンライン	50名	質量インフォマティクス研究会、JSBiの企画でアノテーションについて講習。 https://www.jsbi.org/activity/event/others/detail--id-300.html
2022年11月21日	第4回メタボロミクスソフトウェア講習会	理研横浜	50名	質量インフォマティクス研究会との企画でアノテーションについて講習予定。

以上

別紙 研究開発対象のデータベース等

No.	研究開発課題名	正式名称	別称	概要	URL	公開日	状態	分類	生命科学系データベースアーカイブ	NBDCヒトデータベース	NBDC RDFポータル	関連文献 (論文リストに記載があれば、その番号でも可)
1	物質循環を考慮したメタボロミクス情報基盤	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/KNApS AcK/		維持・発展	データベース等				
2	物質循環を考慮したメタボロミクス情報基盤	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)のみを専門的に取り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)をすることで誰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することが可能です。	http://metabolonote.kazusa.or.jp/		維持・発展	データベース等				
3	物質循環を考慮したメタボロミクス情報基盤	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				21
4	物質循環を考慮したメタボロミクス情報基盤	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				
5	物質循環を考慮したメタボロミクス情報基盤	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				
6	物質循環を考慮したメタボロミクス情報基盤	RIKEN Plant Metabolome MetaDatabase	RIKEN PMM	理研とかずさが蓄積したメタボローム情報のメタデータを整理したRDFサーバ	http://metabobank.riken.jp/		維持・発展	データベース等				
7	物質循環を考慮したメタボロミクス情報基盤	MetaboBank		メタボローム・データを恒久的に保管する公共リポジトリ	http://www.ddbj.nig.ac.jp/metabobank/	2021年10月5日	新規	データベース等				