

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 (https://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 (2021年3月時点)

□目次

§ 1. 研究実施体制.....	2	③ 口頭講演.....	9
§ 2. 研究開発対象とするデータベース・ツール等.....	3	④ ポスター発表.....	9
(1) データベース一覧.....	3	(4) 知的財産権の出願 (国内の出願件数のみ公開)	10
(2) ツール等一覧.....	3	① 出願件数.....	10
§ 3. 実施内容.....	4	② 一覧.....	10
(1) 本年度の研究開発計画と達成目標.....	4	(5) 受賞・報道等.....	10
(2) 進捗状況.....	5	① 受賞.....	10
§ 4. 成果発表等.....	8	② メディア報道.....	10
(1) 原著論文発表.....	8	③ その他の成果発表.....	10
① 論文数概要.....	8	§ 5. 研究開発期間中に主催した活動(ワークショップ等) ...	10
② 論文詳細情報.....	8	1. 進捗ミーティング.....	10
(2) その他の著作物(総説、書籍など).....	9	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.....	11
(3) 国際学会および国内学会発表.....	9		
① 概要.....	9		
② 招待講演.....	9		



§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	情報システム研究機構 国立遺伝学研究所・ 教授	メタボローム・リポジトリの開設と運用
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
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
奈良先	金谷 重彦	奈良先端大学院大学 情報科学領域・教授	生理活性オントロジーの構築

§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.ddbj.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	食レポ	食品メタボロームレポジトリ	http://metabolites.in/foods/
3	RIKEN Plant Metabolomics MetaDatabase (RIKEN PMM)	理研植物メタボロミクスメタデータベース	http://metabobank.riken.jp/
4			
5			

(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	PowerGetBatch		http://www.kazusa.or.jp/komics/software/PowerGetBatch
3			

§3. 実施内容

(1) 本年度の研究開発計画と達成目標

[遺伝研グループ]

リポジトリデータの MetabolomeXchange 登録と国際連携

収集したスタディのメタデータを国際インデクスである MetabolomeXchange に登録した。しかしその動作が遅く便利とは言い難いため、「食品メタボロームリポジトリ」(食レポ) や「理研植物メタボロームメタデータベース」(PMM) といった自前の検索・公開サービスを充実させている。欧州の MetaboLights リポジトリとは連携を続けているが令和 2 年度に予定した合同ワークショップ等は新型コロナウイルスのために中止となった。令和 3 年度以降に、合同ワークショップを企画する予定である。

リポジトリの支援体制構築

MetaboBank は遺伝研のデータベース事業として正式承認され、恒久的なデータのアーカイビングを実現した。事務局を担うスタッフは現在本予算で雇用しているが、DDBJ キュレータのメンバーとしても活動している。ただし現在公開する MetaboBank サーバ・システムは大幅な改訂が必要となるため、関連企業や関連研究組織に対する有償化サービスはまだ計画できていない。

[かずさおよび理研グループ]

植物二次代謝物や食品のスペクトルライブラリを構築

MetaboBank に収録されたデータの再解析を通して、生理活性のある化合物同定に役立てることができた。かずさと理研を中心に 2019 年度に構築した「理研 PMM」は検索インターフェースを充実させた。理研 PMM は植物に特化しているため、別途「食品メタボロームリポジトリ」を準備し、様々な検索機能を充実させた。両システムから、MetaboBank の生データを利用した解析の重要性を国内外に発信している。

[奈良グループ]

メタ代謝マップを用いた植物共生系の解析

KNpSAcK データベースを充実させ、生物活性の情報を整理してオントロジーとした。また代謝物全体を俯瞰するため、植物に限らない生物種から代謝物情報を収集しており、これらの代謝物が複数種の共生によって合成・利用される事例を蓄積している。

(2)進捗状況

[遺伝研グループ]

MetaboBank リポジトリおよび入力エディタの開発

DDBJ サービスの一環として 2020 年 10 月に MetaboBank の公開を開始した (Fig. 1)。

初期データには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、塩基配列とは重ならないアクセッション番号を付与する規則を整備した。

開発状況は EBI/NCBI にも適時連絡した。データの登録は DDBJ 統合アカウントシステムを通して実施する形態をとる。同アカウントシステムはゲノム等、他のデータ投稿も可能なデザインで、現在も改訂が継続中である。

現行の MetaboBank システムでは論文投稿時のレビュー用公開 URL 設定や、一般公開への切り替え操作をユーザ自身が設定する形になっている。一般公開を選択した場合に DDBJ の担当キュレータが内容を確認して必要項目が正しく入力されていることを確認するが、いつまでも作成中で残る情報の扱いや、データ撤回要求に対する規則等が未整備である。今後は遺伝研のキュレータが流れ作業で情報処理できる必要があり、改善点が多々残されている。

現在のデータ登録は「MetaboBank エディタ」と呼ぶ Microsoft Excel のマクロを用いる形式である。ユーザにとって大変親切な設計になってはいるが、複雑なシステムをどうメンテナンスするかという問題も抱えるため、残りの研究実施期間で入力形式の再検討も実施する。

国際的な汎用メタボローム・リポジトリ MetaboLights と MetabolomicsWorkbench との違いは、国際メタボ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満たすほどメタデータが詳しい点である。

Metabolonote からのデータ移行

上記 MetaboBank のメタデータとして、これまで Metabolonote に記載されていた事項の移行は完了した。ただし適切な食品オントロジーがないために、食品関連のみ MetaboBank への移行が完全には進んでいない。食品部分は 2021 年度に移行を完了する予定である。

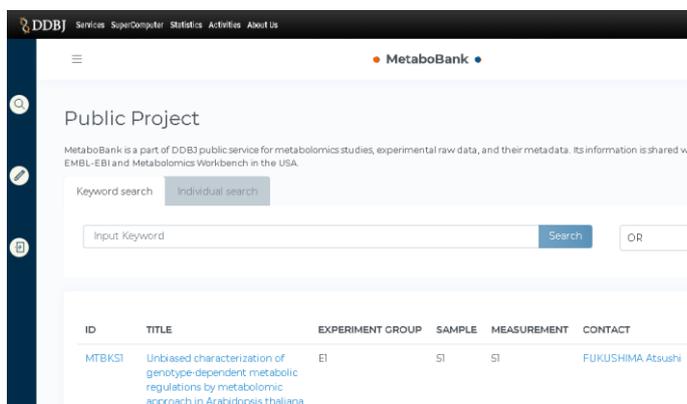


Figure 1 MetaboBank のトップページ

<http://mb.ddbj.nig.ac.jp>

[かずさおよび理研グループ]

「理研メタボロームメタデータベース」のアップデート

MetaboBank の初期データ（生データおよび解析済みデータ）は、「理研 PMM」の公開内容をフォーマット変換して作成している（Fig. 2）。今後は MetaboBank が恒久的なりポジトリとなるため、理研 PMM の位置づけは逆転し、MetaboBank データのうち理研やかずさで取得したものを使いやすく検索できるシステムとなる。



Figure 2 理研メタボロームメタデータベース（理研 PMM）

2021年5月時点では63件の理研データセットおよび88件のかずさDNA研データセット（ただし植物種のみ）を公開している。かずさグループは、データの再解析用にJavaで開発されたソフトウェアPowerGetBatchを使うコマンドラインを工夫し、サーバにおける大量処理パイプラインを構成した。これを用いて大量データを再解析している。

食品メタボローム・リポジトリの構築

MetaboBankより2021年度公開予定のデータから222食品について利用可能にしたシステムが「食レポ」である。HMDBやLipidMapsなど各種データベースへの検索機能を充実させ、精密質量やスペクトルで食品データを検索すると、そのヒット結果がKNApSAcKを含むデータベースに存在するか一目でわかるようになった。またピークエリアから代謝物量を評価する機能も充実させ、様々な利用法について多く発表した。

MS-DIALの開発

Version 4.0を発表し、代表的なスペクトルをもとに脂肪酸の長さを変更した理論スペクトル・ライブラリを117脂質クラスについて作成した。脂質クラスの分類や名称は脂質の国際コンソーシアムと連携して標準化し、それらの存在分布をマウス臓器および培養細胞について検証して、国際共著論文として発表した。

[奈良グループ]

代謝物情報の蓄積と活性オントロジーの整理

KNApSAcK Core DB の開発を継続し 2021 年 4 月 8 日現在、56929 代謝物、24249 種の生物に対する 135, 156 組の生物種-代謝物の関係 KNApSAcK Core DB に格納した。2020 年度には植物科における代表植物を 1 件ずつ選び、少なくとも 100 代謝物を蓄積する作業を実施した。これらの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、KNApSAcK Metabolite Activity DB として進めた。2020 年時点、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。メタボローム研究では、通常 50 程度の代謝物に関わる遺伝子発現、代謝経路解析がなされることが多い。植物に限ってみると、KNApSAcK Core DB に登録された植物種について、植物におけるすべての分類目に分布するに至っており、植物メタボローム研究における中核DBとしての役割を担っている。本 DB を公開したときの学術論文(Plant Cell Physiol., 53: e1, 2012)は、5%最多引用論文として評価されており、たゆまないデータ蓄積が高く評価されている。植物メタボローム以外についても、ヒト(347 種)、海洋生物としては Alcyonacea 属(124 種)など 4 種、カビ類では Aspergillus(583 種), Penicillium 属(569 種)など 6 種、細菌類では Storeptomyces 属(1903 種)を含む 7 属について 80 以上の代謝物が蓄積されている。植物についてみると、198 の被子植物属については、80 以上の代謝物数が蓄積されるに至っている。これらの植物属は 36 種の分類目に帰属される(図 1)。さらに、植物界全体では 230 種の植物属について代謝物の充実を図るに至っている。

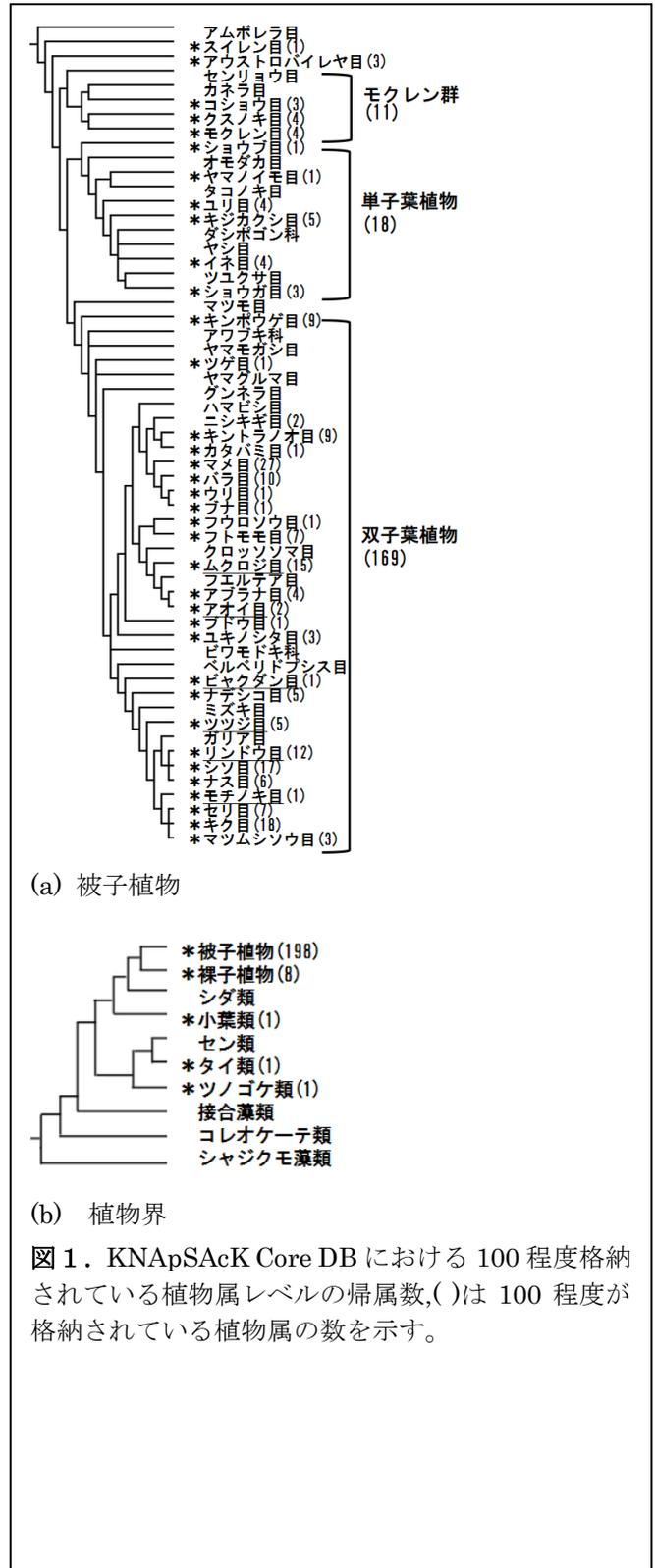


図 1. KNApSAcK Core DB における 100 程度格納されている植物属レベルの帰属数、()は 100 程度が格納されている植物属の数を示す。

§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)	10 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

(直接的な成果論文のほかに開発対象データベースを利用した間接的な成果論文を含む場合があります)

1. Hiroshi Tsugawa, Kazutaka Ikeda, Mikiko Takahashi, Aya Satoh, Yoshifumi Mori, Haruki Uchino, Nobuyuki Okahashi, Yutaka Yamada, Ipputa Tada, Paolo Bonini, Yasuhiro Higashi, Yozo Okazaki, Zhiwei Zhou, Zheng-Jiang Zhu, Jeremy Koelmel, Tomas Cajka, Oliver Fiehn, Kazuki Saito, Masanori Arita, Makoto Arita, "A lipidome atlas in MS-DIAL 4.", *Nature biotechnology*, vol. 38, No. 10, pp.1159-1163, 2020 (DOI: 10.1038/s41587-020-0531-2; PMID:32541957).
2. Ipputa Tada, Romanas Chaleckis, Hiroshi Tsugawa, Isabel Meister, Pei Zhang, Nikolaos Lazarinis, Barbro Dahlén, Craig E Wheelock, Masanori Arita. "Correlation-Based Deconvolution (CorrDec) To Generate High-Quality MS2 Spectra from Data-Independent Acquisition in Multiple Studies.", *Analytical chemistry*, vol. 92, No. 16, pp 11310-11317, 2020 (DOI: 10.1021/acs.analchem.0c01980; PMID:32648737)
3. Katsuya Ohbuchi, Nozomu Sakurai, Hiroyuki Kitagawa, Masaru Sato, Hideyuki Suzuki, Hirotsugu Kushida, Akinori Nishi, Masahiro Yamamoto, Kazuhiro Hanazaki, Masanori Arita. "Differential annotation of converted metabolites (DAC-Met): Exploration of Maoto (Ma-huang-tang)-derived metabolites in plasma using high-resolution mass spectrometry", *Metabolomics*, vol. 16, No. 5, pp 63-na, 2020 (DOI: 10.1007/s11306-020-01681-3).
4. Ili Nadhirah Jamil, Juwairiah Remali, Kamalrul Azlan Azizan, Nor Azlan Nor Muhammad, Masanori Arita, Hoe-Han Goh, Wan Mohd Aizat, "Systematic Multi-Omics Integration (MOI) Approach in Plant Systems Biology.", *Frontiers in plant science*, vol. 11, No. na, pp 944-944, 2020 (DOI: 10.3389/fpls.2020.00944; PMID:32754171)
5. Kusano M, Fukushima A, Tabuchi-Kobayashi M, Funayama K, Kojima S, Maruyama K, Yamamoto YY, Nishizawa T, Kobayashi M, Wakazaki M, Sato M, Toyooka K, Osanai-Kondo K, Utsumi Y, Seki M, Fukai C, Saito K, Yamaya T, "Cytosolic GLUTAMINE SYNTHETASE1;1 Modulates Metabolism and Chloroplast Development in Roots", *Plant Physiology*, vol. 182, No. 4, pp 1894-1909, 2020 (DOI: 10.1104/pp.19.01118)
6. Hiraga Y, Ara T, Nagashima Y, Shimada N, Sakurai N, Suzuki H, Kera K, "Metabolome analysis using multiple data mining approaches suggests luteolin biosynthesis in *Physcomitrella patens*", *Plant Biotechnology*, vol. 37, No. 3, pp 377-381, 2020 (DOI: 10.5511/plantbiotechnology.20.0525b)
7. Utsumi Y, Tanaka M, Utsumi C, Takahashi S, Matsui A, Fukushima A, Kobayashi M, Sasaki R, Oikawa A, Kusano M, Saito K, Kojima M, Sakakibara H, Sojikul P, Narangajavana J, Seki M, "Integrative omics approaches revealed a crosstalk among phytohormones during tuberous root development in cassava", *Plant Molecular Biology*, vol. na, No. na, pp 0-0, 2020 (DOI: 10.1007/s11103-020-01033-8)
8. Hossain SF, Huang M, Ono N, Morita A, Kanaya S, Altaf-Ul-Amin M, "Development of a bio marker database toward performing disease classification and finding disease interrelations", *Database (Oxford)*, vol. na, No. na, pp 0-0, 2020 (DOI: 10.1093/database/baab011)
9. Higa Y, Kim YS, Altaf-Ul-Amin M, Huang M, Ono N, Kanaya S, "Divergence of metabolites in three phylogenetically close *Monascus* species (*M. pilosus*, *M. ruber*, and *M. purpureus*) based on secondary metabolite biosynthetic gene clusters", *BMC Genomics*, vol. 21, No. 1, pp 679, 2020 (DOI: 10.1186/s12864-020-06864-9)
10. Desmet S, Saeys Y, Verstaen K, Dauwe R, Kim H, Niculaes C, Fukushima A, Goeminne G, Vanholme R, Ralph J, Boerjan W, Morreel K, "Maize specialized metabolome networks reveal

(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	5 件
	国際	2 件
口頭発表	国内	0 件
	国際	0 件
ポスター発表	国内	1 件
	国際	0 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

〈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 INBIOSIS, Malaysia, Mar 26, 2021

③ 口頭講演

〈国内〉

1. 有田正規 「メタボバンク」 分子生物学会年会フォーラム(NBDC 企画) 2020 年 12 月 2 日

④ ポスター発表

〈国内〉

1. 長崎英樹、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、トーゴーの日シンポジウム 2020、オンライン開催、2020 年 10 月 5 日

〈国際〉

該当なし。

(4) 知的財産権の出願

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

該当なし。

2) 海外出願

該当なし。

3) その他の知的財産権

該当なし。

(5) 受賞・報道等

① 受賞

1. 理研 CSRS 奨励賞、高橋みき子、山田豊、2020 年 6 月 10 日
2. 理研梅峰賞、津川裕司、2020 年 3 月 12 日 https://www.riken.jp/pr/news/2020/20200319_1/index.html
3. 国立遺伝学研究所森島奨励賞、多田一風太、2020 年 9 月 29 日
4. 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

該当なし。

③ その他の成果発表

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
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	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
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	9 人	同上
2021 年 1 月 15 日	チーム内ミーティング (非公開)	オンライン 開催	11 人	同上
2021 年	チーム内ミーティング (非公開)	オンライン 開催	11 人	同上

年月日	名称	場所	参加人数	目的・概要
2月2日				

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
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

以上

