

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://www.ddbj.nig.ac.jp/metabobank/)
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 (2022年3月時点)

□目次

§1. 研究実施体制	2	③ 口頭講演	8
§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	8
(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) 8	
(2) ツール等一覧	3	(5) 受賞・報道等	8
§3. 実施内容	4	① 受賞	8
(1) 本年度の研究開発計画と達成目標	4	② メディア報道	9
(2) 進捗状況	5	③ その他の成果発表	9
§4. 成果発表等	7	§5. 研究開発期間中に主催した活動(ワークショップ等)	9
(1) 原著論文発表	7	(1) 進捗ミーティング	9
① 論文数概要	7	(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	10
② 論文詳細情報	7		
(2) その他の著作物(総説、書籍など)	7		
(3) 国際学会および国内学会発表	7		
① 概要	7		
② 招待講演	7		



§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・上級研究員	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	金谷 重彦	奈良先端科学技術大学院大学・先端科学技術研究科・教授	メタ代謝マップ構築
農工大	津川 裕司	東京農工大学グローバルイノベーション研究院・テニュアトラック准教授	MS-DIALの開発

§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	https://www.ddbj.nig.ac.jp/metabobank/

【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	RIKEN Plant Metabolomics MetaData base (RIKEN PM M)	理研植物メタボロミクスメタデータベース	http://metabobank.riken.jp/

(2) ツール等一覧

No.	名称	別称・略称	URL
1	MS-DIAL MS-FINDER	RIKEN CompMS	http://prime.psc.riken.jp/compms/index.html
2	PowerGetBatch		http://www.kazusa.or.jp/komics/software/PowerGetBatch

§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]

- **MetaboBank Wiki サーバの構築と公開**

MetaboBank Wiki サーバのシステムおよび維持管理体制を決定し、代謝物情報、スペクトル情報、メタ代謝マップ情報を掲載する。

- **MetaboBank メタデータの策定とシステム改訂**

ヒトや環境メタボロームデータを収載できるようにメタデータ項目を策定し、MetaboBank でデータを受け付けられるように実装、また入力エディタ等のシステムを改訂する。

[理研グループ]

- **理研メタデータベースの作成とシステム改訂**

遺伝研・かずさグループと連携して理研植物メタデータベースを植物以外のデータへ拡張する。MetaboBank 本体との整合性を整え、双方におけるメタデータ項目のすり合わせを実施する。

- **再解析によるスペクトルデータの整理とモデルメタボロームの構築**

同定可能なメタボロームのリストを作成、利用するスペクトルデータをライブラリとして整える。再解析結果は MetaboBank Wiki サーバより公開する。

[かずさグループ]

- **理研メタデータベースの作成強力とシステム改訂**

理研グループ同様、かずさにおいて測定したデータについて、MetaboBank 本体との整合性を整え、双方におけるメタデータ項目のすり合わせを実施する。

- **再解析によるスペクトルデータの整理とモデルメタボロームの構築**

解析プログラム PowerGetBatch のパラメータ設定を検討し、過去解析結果と比較をしながらデータ再解析する (京都大学 生存圏研 荒氏と共同開発)

[奈良グループ]

- **メタ代謝マップと文献データのすり合わせ**

代謝物には InChI コードを付与し、できる限り文献の PMID を取得する。この InChI コードを介してメタ代謝マップ、代謝物、文献情報をリンクさせる。

- **メタボロームデータとメタ代謝マップの統合**

理研およびかずさグループが再解析で同定する化合物について、InChI を介してメタ代謝マップとリンクさせる。

[農工大グループ]

- **MS-DIAL の開発とスペクトルライブラリの構築**

MS-DIAL を、MetaboBank データを直接利用できるソフトウェアとして開発、普及させる。

(2) 進捗状況 [遺伝研グループ]

○ MetaboBank Version 2 の構築

前年度に作成した MetaboBank Version 1 はレビュー用インターネットアドレスの設定や一般公開切り替えをユーザー側で実施できる仕様になっていた。またシステムも Java Spring を用いて構成されており、その改修とメンテナンスが DDBJ 所属のスタッフおよびキュレータには難しかった。そのため、既存システムにできるだけ近いシステムで再構築せざるを得ない状況となり、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得なかった。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揃える必要も生じた。とりわけ食品データはゲノムデータベースで対応する術がなかったため、メタゲノム解析の一種として扱っている。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/> より公開する。

○ MetaboBank Wiki の構築と公開

KNpSack データベースとの統合およびスペクトル情報を閲覧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ている。また代謝物情報は KNpSack データベースと共通化した。サンプルページは <https://mb.metabolomics.jp/wiki/Compound:C00000001> に示してある。

Wiki のコンテンツ掲載は予定よりも遅れており、メタ代謝マップは 2022 年度に移植する予定である。スペクトル表示機能は既に実装してあるが、日々更新されるデータを反映させる手順がまだ定められていないため、これも 2022 度に完成させる予定である。

[理研グループ]

○ MetaboBank Version 2 への(メタ)データ寄託

理研 CSRS 由来の植物メタボロームデータを中心に、MAGE-TAB 形式でのメタデータを整備し直し、MetaboBank の初期データとして登録した。具体的には、MetaboBank v2 に登録していた 59 個の研究プロジェクト(MTBKS1~59)の IDF、SDRF、MAF ファイル、対応する生データファイルおよび処理済データファイルを新たに作成し、DDBJ へ寄託した。

○ データ再解析および再アノテーション

農工大グループが開発するソフトウェア (MS-DIAL, Aloutput) と理研グループが開発したワークフロー (eRah) を組み合わせて、理研 CSRS 植物メタボロームデータ (GC-MS) について再解析をおこなった。結果として再解析の対象データセット 55 個 (RIKEN PMM プロジェクト単位) のうち、MS-DIAL では 45 個、eRah で 44 個、Aloutput によって 41 個の再解析を終えた。さらに利用可能なソフトウェアツールを探索し、来年度に再アノテーションの比較等を行うための準備を進めた。

○ 「理研植物メタボロームメタデータベース」のアップデート

今年度、論文発表がなされた理研 CSRS の植物メタボロームデータの公開に加えて、理研 IMS の脂質データについても MetaboBank v2 への登録準備 (IDF/SDRF/MAF ファイルの作成および生データの整理)を進めた。また、理研植物メタボロームメタデータベースに関して、学術論文 (Plant Cell Physiol., 63: 433-440, 2022)として公表した。

[かずさグループ]

○ 解析プログラム PowerGetBatch のパラメータ設定とデータ再解析

かずさグループの解析では LC-MS の FT、Orbitrap、QTOF 等複数種の質量分析計の結果を扱うため各機器用にパラメータ設定をおこなった。

Orbitrap の解析結果において、論文のキュレーション結果とで比較したところ、1) 化合物アノテーションされたピークの 94.3%を検出。2) 1 万以上の新規ピーク候補を抽出 3) 976 個の新規化合物アノテーション候補ピークを抽出した。現在他の質量分析計の結果を査定、精査評価方法の効率化を計るとともに再解析を進めている。

○ かずさ-所有データの MetaboBank Version 2 への移行

MetaboBank V1 では「ヒト臨床」「動物」「植物」「微生物」「環境」等 9 種合わせて 398 あった入力項目を数十に減らす検討をおこなった。V2 へのデータ移行のため IDF, SDRF フォーマットへの変換作業を年度末から実施しており、2022 年 5 月でほぼ完了する。

[奈良グループ]

○ メタ代謝マップと文献データのすり合わせ

メタ代謝マップ中に含まれる化合物と他の RDF データベースとのすり合わせを実施するため、昨年度に引き続いて植物種・代謝物情報の収集を進めた。コアデータベースには 141,486 レコードの生物種-二次代謝物の関係があり、二次代謝物の総数は 57,906 種、収録されている生物種は 24,690 種である。これらのデータに CAS 番号、InChI コード、PubMed 番号を付与する作業を実施し、メタ代謝マップ中のデータと、かずさや理研等が扱うデータを RDF 統合する準備を整えた。また二次代謝物の構造類似性を収載したデータベースとして Twins セクションを設け(2020 年 1 月スタート)、3 次元グラフマッチングアルゴリズムに基づいて化合物を検索する機能を追加した。

RDF 化したコアデータを公開するとともに、遺伝研グループと共同で代表的な 164 植物種における代謝物情報の精査を実施した。これは研究上よく用いられる植物種(および比較対象として大腸菌)のメタボローム標準リストとして利用できる。対応する文献数は 3394 件、30542 化合物について InChI コードや SMILES 情報も取得した。このうち 7800 化合物は KNApSack と異なると思われるため、2022 年度に再精査して統合する。

[農工大グループ]

○ MS-DIAL 5 シリーズの制限付き公開

MS-DIAL は現在、世界でもトップシェアを持つ質量分析データ解析プログラムである。今年度は、これまでサポートされていなかった直接導入法(DI-MS)や直接導入+イオンモビリティ分離法(IM-MS)データに対応し、事実上、すべての計測手法に対応するサポート体制を整えた。また昨今、新しいマスフラグメンテーションによる構造解析手法が提案されており SCIEX 社の electron activated dissociation (EAD)法から得られる MS/MS スペクトルのアノテーションプログラムを実装した。以上を MS-DIAL 5 として論文投稿準備中である。一方、論文掲載に先立ち、ユーザーからフィードバックを受けてソフトウェアをより良いものにするため、MS-DIAL 5 のベーシック認証付きの制限付き公開をおこなった。MS-DIAL 5 は、理化学研究所 PRIME ウェブサイトの、CompMS セクションからダウンロード可能である。

§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. Atsushi Fukushima, Mikiko Takahashi, Hideki Nagasaki, Yusuke Aono, Makoto Kobayashi, Miyako Kusano, Kazuki Saito, Norio Kobayashi, Masanori Arita, "Development of RIKEN Plant Metabolome MetaDatabase.", *Plant & cell physiology*, 2021 (DOI: 10.1093/pcp/pcab173).
2. Mostafa Abdelrahman, Rie Nishiyama, Cuong Duy Tran, Miyako Kusano, Ryo Nakabayashi, Yo zo Okazaki, Fumio Matsuda, Ricardo A. Chávez Montes, Mohammad Golam Mostofa, Weiqiang Li, Yasuko Watanabe, Atsushi Fukushima, Maho Tanaka, Motoaki Seki, Kazuki Saito, Luis Herrera-Estrella, Lam-Son Phan Tran, "Defective cytokinin signaling reprograms lipid and flavonoid gene-to-metabolite networks to mitigate high salinity in Arabidopsis.", *Proceedings of the National Academy of Sciences*, vol. 118, pp. e2105021118, 2021 (DOI: 10.1073/pnas.2105021118).
3. Altaf-Ul-Amin M, Kanaya S, Ono N, Huang M, "Recent Trends in Computational Biomedical Research.", *Life (Basel)*, vol. 12, 2021 (DOI: 10.3390/life12010027).
4. Altaf-Ul-Amin M, Hirose K, Nani JV, Porta LC, Tasic L, Hossain SF, Huang M, Ono N, Hayashi
5. MAF, Kanaya S, "A system biology approach based on metabolic biomarkers and protein-protein interactions for identifying pathways underlying schizophrenia and bipolar disorder.", *Scientific reports*, vol. 11, 2021 (DOI: 10.1038/s41598-021-93653-3).
6. Hossain S.F, Huang M, Ono N, Morita A, Kanaya S, Md Altaf-Ul-Amin, "Development of a biomarker database toward performing disease classification and finding disease interrelations.", *Database (Oxford)*, 2021 (DOI: 10.1093/database/baab011).

(2) その他の著作物(総説、書籍など)

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	0 件
	国際	0 件
口頭発表	国内	6 件
	国際	4 件
ポスター発表	国内	4 件
	国際	0 件

② 招待講演

なし

③ 口頭講演

〈国内〉

1. 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
2. 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
3. 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
4. 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. 発表者、タイトル、学会名、場所、月日
2. 津川裕司, Computational mass spectrometry to deepen the understanding of metabolisms, AS HBi Seminar(オンライン), 2022年2月14日
3. 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
4. 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〉

発表者、タイトル、学会名、場所、月日

1. トーゴーの日シンポジウム 2021, 10 月 5 日(オンライン)
 - 1) 長崎英樹, 荒武, 大澤祥子, 福島敦史, 高橋みき子, ほか. メタボローム統合データベース MetaboBank の開発
 - 2) 平川英樹, 藤澤貴智, 長崎英樹, 森宙史, 福島敦史, ほか. 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
2. 長崎英樹, 大澤祥子, 荒武, 福島敦史, 高橋みき子, ほか, 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, 10 月 14-15 日
3. 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, 6 月 4 日

〈国際〉

なし

(4) 知的財産権の出願 (国内の出願件数のみ公開)

なし

(5) 受賞・報道等

① 受賞

なし

② メディア報道

なし

③ その他の成果発表

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

2021/6/24	第1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	オンライン	19 人	議事: DDBJ サービスに関する専門 部会の運営について等
2021/7/26	第2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	オンライン	17 人	議事: 「統合化推進プログラム」中 間評価について等
2021/10/21	第3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	オンライン	21 人	議事: MetaboBank Ver.2 の開 始について等
2021/9/8	統合化推進プログラム サイトビジット	オンライン	9+外部	課題: 「物質循環を考慮したメタボ ロミクス情報基盤」
2021/4/20	チーム内ミーティング (非公開)	オンライン	4 人	MetaboBank 開発進捗報告の ためのミーティング
2021/5/10	同上	同上	5 人	同上
2021/5/19	同上	同上	4 人	同上
2021/6/2	同上	同上	4 人	同上
2021/6/16	同上	同上	4 人	同上
2021/6/22	同上	同上	4 人	同上
2021/6/30	同上	同上	5 人	同上
2021/7/5	同上	同上	5 人	同上
2021/7/14	同上	同上	6 人	同上
2021/7/20	同上	同上	7 人	同上
2021/7/27	同上	同上	6 人	同上
2021/8/3	同上	同上	7 人	同上
2021/8/10	同上	同上	4 人	同上
2021/8/18	同上	同上	7 人	同上
2021/8/25	同上	同上	6 人	同上
2021/9/2	同上	同上	10 人	同上
2021/9/8	同上	同上	8 人	同上
2021/9/21	同上	同上	7 人	同上
2021/9/27	同上	同上	8 人	同上
2021/10/13	同上	同上	11 人	同上
2021/10/27	同上	同上	13 人	同上
2021/11/1	同上	同上	7 人	同上
2021/11/4	同上	同上	9 人	同上
2021/11/24	同上	同上	10 人	同上
2021/12/8	同上	同上	11 人	同上
2021/12/22	同上	同上	11 人	同上
2022/1/13	同上	同上	10 人	同上

2022/2/3	同上	同上	8人	同上
2022/3/16	同上	同上	9人	同上
2022/3/30	同上	同上	7人	同上
2022/1/27	同上	同上	5人	MetaboBank wiki 関連ミー ティング

(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
なし

以上

