

平成 26 年度 研究開発実施報告書

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」研究代表者

[有田 正規]

[理化学研究所 横浜研究所 環境資源科学研究センター・チームリーダー]

[生物種メタボロームモデル・データベースの構築]

§1. 研究実施体制

(1)「理研」グループ(研究機関別)

① 研究代表者:有田正規 (理研 横浜研究所 環境資源科学研究センター・チームリーダー)

② 研究項目

・メタボモデルの構築

(2)「奈良先端」グループ(研究機関別)

① 主たる共同研究者:金谷重彦 (奈良先端科学技術大学院大学情報科学研究科、教授)

② 研究項目

・代謝物と生物活性の関係データベース構築

(3)「かずさ」グループ(研究機関別)

① 主たる共同研究者:櫻井望 (かずさ DNA 研究所 技術開発研究部・チーム長)

② 研究項目

・メタボローム・データベースの構築

§ 2. 研究実施内容

理研グループ

- ・シロイヌナズナ・メタボモデルの構築(有田、福島)

モデル植物のシロイヌナズナについて、連携する機関で計測した代謝物 GC-MS プロファイルを統合し、MassBank に登録された GC-MS スペクトルとの対応付けをおこなった。

- ・データ共有クローリングソフトウェアの構築(有田、Mejia)

MassBank データベース、MassBank of North America データベースよりスペクトルデータを自動ダウンロードし、JSON 形式で保有するクローリングソフトウェアを開発した。そのために必要な REST サービスの設計もおこなった。

- ・LC-MS/MS アノテーションソフトウェアの開発(有田、津川)

MRM 計測や Sequential Window Acquisition of all Theoretical Spectra すなわち SWATH 計測に対応した脂質同定用のソフトウェアを開発した。

奈良先端グループ

- ・生物種と代謝物の関係データベース拡充(金谷、平井)

KNAPSAcK ファミリーデータベースにおける二次代謝物文献情報を拡充した。また、生物種として特に、食用と薬用生物に関して、世界の 222 地域(国と島)ごとの利用情報を生物-利用国の組み合わせで 56275 件(9924 件、地域-食用植物、46351 件、地域-薬用植物)整理し、生物種-代謝物関係データベースへのリンク付けをおこなった。

- ・代謝物と生物活性の関係データベース構築(金谷、中村)

KNAPSAcK Metabolite Activity DB を構築した。現在、10758 件の情報が登録されている。

かずさグループ

- ・メタボロームデータの大規模整備(櫻井)

大規模なデータ公開のため、LC-精密質量 MS の生データからピークを抽出しアノテーションを行うソフトウェア PowerGet について、質量精度の向上、アダクト判定の高精度化を行った。また、精密質量値から化合物データベースへの検索結果を集計するモジュールについても改良を行い、より化合物候補を絞り込みやすいアノテーションを付与できるようにした。さらに、MS クロマトグラムデータの汎用フォーマットである mzXML への対応と、一連の処理をコマンドラインベースで行えるバッチ処理化を行って、大量のデータ自動処理を行うパイプラインが構築完了した。MassBase より公開した LC-精密質量 MS の生データのうち、約 500 件について、自動処理を実施し、Metabolonote より公開の準備を進めた。

- ・メタデータによる外部データベースとの連携

Metabolonote データベースのバックエンドにある MySQL データベースより、メタデータを直接 RDF 出力するインターフェースを構築した。また Metabolonote のインターフェースを改良し、ユーザがよりわかりやすく情報を入力できる仕組みにした。

§3. 成果発表等

(3-1) 原著論文発表

- ① 発行済論文数(国内(和文) 0 件、国際(欧文) 14 件):
- ② 未発行論文数(“accepted”、“in press”等)(国内(和文) 0 件、国際 (欧文) 0 件)
- ③ 論文詳細情報
 1. Nakamura Y, Afendi FM, Parvin AK, Ono N, Tanaka K, Hirai Morita A, Sato T, Sugiura T, Altaf-Ul-Amin M, Kanaya S “KNApSAcK Metabolite Activity Database for retrieving the relationships between metabolites and biological activities” *Plant Cell Physiol* 55(1), e7, 2014 (doi: 10.1093/pcp/pct176)
 2. Tsugawa H, Kanazawa M, Ogiwara A, Arita M “MRMPROBS suite for metabolomics using large-scale MRM assays” *Bioinformatics* 30(16), 2379-2380, 2014 (doi: 10.1093/bioinformatics/btu203)
 3. Tsugawa H, Ohta E, Izumi Y, Ogiwara A, Yukihiro D, Bamba T, Fukusaki E, Arita M “MRM-DIFF: data processing strategy for differential analysis in large scale MRM-based lipidomics studies” *Front Genet*, 5, 471, 2015 (doi: 10.3389/fgene.2014.00471)
 4. Fukushima A, Kanaya S, Nishida K “Integrated network analysis and effective tools in plant systems biology” *Frontiers in Plant Sci.*, 5, 1-9, 2014 (doi:10.3389/fpls.2014.00598)
 5. Amin MA, Katsuragi T, Sato T, Ono N, Kanaya S “An Unsupervised Approach to Predict Functional Relations between Genes Based on Expression Data, *BioMed Res. International* 154594, 1-8, 2014 (doi: 10.1155/2014/154594)
 6. Amin MA, Afendi FM, Kiboi SK, Kanaya S “Systems Biology in the Context of Big Data and Networks” *BioMed Res. International*, 428570, 1-11, 2014 (doi: 10.1155/2014/428570)
 7. Kibinge N, Ikeda S, Ono N, Amin MA, Shigehiko Kanaya S “Integration of Residue Attributes for Sequence Diversity Characterization of Terpenoid Enzymes” *BioMed Res. International*, 753428, 1-10, 2014 (doi: 10.1155/2014/753428)
 8. Sakurai N, Ara T, Enomoto M, Motegi T, Morishita Y, Kurabayashi A, Iijima Y, Ogata Y, Nakajima D, Suzuki H, Shibata D “Tools and databases of the KOMICS web portal for preprocessing, mining, and dissemination of metabolomics data” *BioMed Res. International*, 194812, 1-13, 2014 (doi: 10.1155/2014/194812)
 9. Ohtana Y, Abdullah AA, Amin MA, Huang M, Ono N, Sato T, Sugiura T, Horai H, Nakamura Y, Morita AH, Lange KW, Kibinge NK, Katsuragi T, Shirai T, Kanaya S “Clustering of 3D-Structure Similarity Based Network of Secondary Metabolites Reveals Their Relationships with Biological Activities” *Mol. Inf.* 33, 790-801, 2014 (doi: 10.1002/minf.201400123)
 10. Maruyama K, Urano K, Yoshiwara K, Morishita Y, Sakurai N, Suzuki H, Kojima M, Sakakibara H, Shibata D, Saito K, Shinozaki K, Yamaguchi-Shinozaki K “Integrated analysis of the effects of cold and dehydration on rice metabolites, phytohormones, and gene transcripts” *Plant Physiol.*, 164(4), 1759-1771, 2014 (doi: 10.1104/pp.113.231720)

11. Mano J, Khorobrykh S, Matsui K, Iijima Y, Sakurai N, Suzuki H, Shibata D “Acrolein is formed from trienoic fatty acids in chloroplast: A targeted metabolomics approach” *Plant Biotechnology* 31(5):535-543, 2014 (doi: 10.5511/plantbiotechnology.14.1112a).
12. Mannen K, Matsumoto T, Takahashi S, Yamaguchi Y, Tsukagoshi M, Sano R, Suzuki H, Sakurai N, Shibata D, Koyama T, Nakayama T “Coordinated transcriptional regulation of isopentenyl diphosphate biosynthetic pathway enzymes in plastids by phytochrome-interacting factor 5” *Biochemical and biophysical research communications* 443(2):768-774, 2014 (doi: 10.1016/j.bbrc.2013.12.040)
13. Kera K, Ogata Y, Ara T, Nagashima Y, Shimada N, Sakurai N, Shibata D, Suzuki H “ShiftedIonsFinder: A standalone Java tool for finding peaks with specified mass differences by comparing mass spectra of isotope-labeled and unlabeled data sets” *Plant Biotechnology* 31(3):269-274, 2014 (doi: 10.5511/plantbiotechnology.14.0609c)
14. Aoki Y, Takahashi S, Takayama D, Ogata Y, Sakurai N, Suzuki H, Asawatreratanakul K, Wititsuwannakul D, Wititsuwannakul R, Shibata D, Koyama T, Nakayama T “Identification of laticifer-specific genes and their promoter regions from a natural rubber producing plant *Hevea brasiliensis*” *Plant science* 225:1-8, 2014 (doi: 10.1016/j.plantsci.2014.05.003)

(3-2) データベースおよびウェブツール等の構築と公開

- ① 公開中のデータベース・ウェブツール等
別紙1参照。

(3-3) その他の著作物(総説、書籍など)

なし

(3-4) 国際学会および国内学会発表

- ① 招待講演 (国内 9 件、国際 0 件)

〈国内〉

1. 金谷重彦, 「バイオビッグデータに挑む: 医・薬・食用植物の機能性 DB の構築とマイニング」、7 月 28 日、北海道大学理学部セミナー、札幌
2. 金谷重彦, 「ビッグデータバイオロジーとしての生態学医食同源の体系化」、日本農芸化学会、奈良、2014 年 9 月 19 日
3. 金谷重彦, 「機能性機能食品データベース構築と活用」IT x バイオ連携促進セミナー、(株)道銀地域総合研究所、2014 年 9 月 10 日、札幌
4. 金谷重彦, 「健康に向けた食品の栄養成分データベース」、健康と長寿に貢献する遺伝学と情報学、日本遺伝学会、2014 年 9 月 20 日、長浜バイオ大学
5. 有田 正規 津川 裕司 「藻類メタボロミクス: 植物との違い」日本植物学会第 78 回大会シンポジウム「バイオリソースとゲノム情報から考える藻類研究の未来形」、東京

2014 9/12

6. 有田 正規 津川 裕司「ノンターゲット MS/MS による藻類脂質の網羅的解析とデータベース化」健康・長寿研究談話会(旧ホスファチジルセリン研究会), 品川 2014 11/7
7. Arita M "Impact of Metabolome Database in Plant Science" The 38th Naito Conference "Molecule-based biological systems", October 7-10, Sapporo, Japan, 2014
8. 櫻井望 「メタボロミクスにおける化合物推定」 CAC フォーラムー泊研修会, 湯河原 2014/10/7
9. 櫻井望 「食品メタボロミクスを拓く最先端技術」 食品開発セミナー, 千葉, 2014/11/17, 19

② 口頭講演 (国内 1 件、国際 6 件)

〈国内〉

1. 金谷重彦、オミックスプラットフォームに基づく二次代謝物情報データベース KNApSAcK Family の研究開発、第 37 回情報科学討論会、11 月 28 日、豊橋

〈国際〉

1. Arita M “Metabolome Database”, 1st Symposium on International Collaboration between NAIST and UC Davis: Data-driven genomics science towards opening a NAIST satellite laboratory at UC Davis. UC Davis, 9th Feb. 2015
2. Okazaki A “Metabolic pathway of carbon dioxide incorporation during wax ester fermentation in Euglena”, 1st Symposium on International Collaboration between NAIST and UC Davis: Data-driven genomics science towards opening a NAIST satellite laboratory at UC Davis. UC Davis, 9th Feb. 2015
3. Fukushima A “Development of a metabolite profiling database of Arabidopsis,” 1st Symposium on International Collaboration between NAIST and UC Davis: Data-driven genomics science towards opening a NAIST satellite laboratory at UC Davis. UC Davis, 9th Feb. 2015
4. Tsugawa H, “Untargeted metabolomics strategy via data independent MS/MS acquisition,” 1st Symposium on International Collaboration between NAIST and UC Davis: Data-driven genomics science towards opening a NAIST satellite laboratory at UC Davis. UC Davis, 9th Feb. 2015
5. Kanya S, “KNApSAcK Family Database towards Big Data”, PRAGMA28, Software defined e-Science, Apr,9, (2015) Nara, Japan
6. Arita M "MassBank Past, Present and Future" Norman Massbank Workshop, Duebendorf, Switzerland, 2014 9/17

③ ポスター発表 (国内 1 件、国際 5 件)

〈国内〉

1. 櫻井望、「実験条件の専用管理システム Metabolonote によるメタボロミクス実験データリソースの統合化」、トーゴーの日シンポジウム、東京、2014 年 10 月 5 日

〈国際〉

1. Sakurai N, “Metabolonote: A Semantic MediaWiki-based database for metadata of metabolomics experiments”, Metabolomics 2014, Tsuruoka, Japan, 23-26th June 2014
2. Wijaya SN, Muraoka S, Andarini LA, Afendi FM, Hirai-Morita A, Latifah K, Md. Altaf-Ul-Amin, Sato T, Ono N, Sugiura T, Kanaya S “HerbsMed: Herbal medicine apps using integrated Jamu and Kampo formulas” Metabolomics 2014, Tsuruoka, Japan, 23-26th June 2014
3. Katsuragi T, Ono N, Sato T, Kanaya S “Dynamic simulation of metabolic network in Arabidopsis thaliana using parameters estimated by a genetic algorithm with modularity” Metabolomics 2014, Tsuruoka, Japan, 23-26th June 2014
4. Li D, Ono N, Sato T, Sugiura T, Md. Altaf-Ul-Amin, Arita M, Tanaka K, Ma Z, Kanaya S “Strategy of integration between RNA-seq and metabolome in curcuminoid biosynthesis pathways in cultivars of curcuma longa” Metabolomics 2014, Tsuruoka, Japan, 23-26th June 2014
5. Nakamura Y, Parvin AK, Ono N, Hirai-Morita A, Sato T, Sugiura T, Md. Altaf-Ul-Amin, Kanaya S “A new metabolomic information search using the KNApSACk Metabolite Activity database” Metabolomics 2014, Tsuruoka, Japan, 23-26th June 2014

(3-5) 知的財産権の出願

- ① 当該年度の特許出願件数(国内 0 件、海外 0 件)
- ② 研究開始から当該年度末までの累積特許出願件数(国内 0 件、海外 0 件)
- ③ その他の知的財産権
なし

(3-6) 受賞・報道等

① 受賞

大棚 勇輝, 白井 剛, 金谷 重彦, Md. Altaf-Ul-Amin, 杉浦 忠男, 小野 直亮, 佐藤哲大, 中村由紀子, 桂樹 哲雄、「植物における二次代謝物の効能と立体構造の関係に関する研究」第 37 回情報科学討論会、11 月 28 日、豊橋(ポスター賞)

Ohtana Y, Abdullah AA, Amin MA, Huang M, Ono N, Sato T, Sugiura T, Horai H, Nakamura Y, Morita AH, Lange KW, Kibinge NK, Katsuragi T, Shirai T, Kanaya S, “Clustering of 3D-Structure Similarity Based Network of Secondary Metabolites Reveals Their Relationships with Biological Activities,” *Mol. Inf.* **33**, 790-801 (2014) (DOI: 10.1002/minf.201400123)
2014 Molecular Informatics Best Paper Award Jan, 2015

Best Young Investigator In Plant Metabolomics Award

Katsuragi T, Ono N, Sato T, Kanaya S “Dynamic simulation of metabolic network in Arabidopsis thaliana using parameters estimated by a genetic algorithm with modularity” 10th

International Conference of the Metabolomics Society, June 23-26, Tsuruoka , 2014.

§4. 研究開発期間中に主催した活動(主催したワークショップ等)

なし。

以上

別紙1 既公開のデータベース・ウェブツール等

No.	研究開発課題名 データベース、ウェブツール等の名称	概要	URL	公開開始日	関連論文 (論文リストに記載があれば、その番号でも可)	備考
01	生物種メタボロームモデル・データベースの構築	KNApSAcK Core DB	本DBは、文献情報をもとに生物種代謝物の関係を収集し、公開しています。現在までに、10万種の生物種-代謝物関係が蓄積されており、メタボロミクス研究の標準データベースとなっています。	http://kanaya.naist.jp/knap sack.jsp/top.html		
02	生物種メタボロームモデル・データベースの構築	Bio-MassBank	本DBは、植物の組織や微生物試料をLC-, GC-, CE-MSで分析して得られたマスマスペクトルを代謝物を同定できた、できないいにかかわらず収集します。未同定代謝物のマスマスペクトルをその化学構造を表現する化学的descriptorとして利用することによって、異なる試料間で同じあるいは類似したマスマスペクトルがあれば同じ未同定代謝物が含まれている、と考えることができます。このように代謝物を同定することができなくても、その存在を知ることができます。現在、シロイヌナズナの葉、ミヤコグサの花をLC-MS, MS2分析したデータそれぞれ664件、636件を公開しています。	http://bio.massbank.jp/		
03	生物種メタボロームモデル・データベースの構築	KomicMarket	本DBは、主にLC-MSIによるメタボローム解析で得られた化合物ピークについて、化合物の同定・推定(アノテーション)を行った結果を蓄積・公開しているDBです。	http://webs2.kazusa.or.jp/komicmarket/		
04	生物種メタボロームモデル・データベースの構築	MassBank	本DBは、代謝物質あるいはそれらの関連物質をEI-, FAB-, ESI-, MALDI-MS, MS2などを用いて分析したマスマスペクトルを収集、公開しています。データの公開状況(2012年3月末)は次のとおりです。21研究グループ(日本15、米国3、ドイツ2、中国1)が13,534化合物について分析した合計29,644マスマスペクトルデータを9つのデータサーバから公開しています。それらのうち、ESI-MS、MS2データは2,304化合物について測定した16,440件です。	http://www.massbank.jp/		
05	生物種メタボロームモデル・データベースの構築	MassBase	本DBは、主に質量分析の未加工データ(生データ)、生データをテキスト型式に変換したデータ、同定・推定(アノテーション)を行わないピークデータを、蓄積・公開するためのDBです。	http://webs2.kazusa.or.jp/massbase/		
06	生物種メタボロームモデル・データベースの構築	Metabolome Activity DB	本DBは、文献情報をもとに代謝物の活性情報を収集し、公開しています。現在までに、5千種の代謝物-活性の関係が蓄積されており、メタボロミクス研究の標準データベースとなっています。	http://kanaya.naist.jp/MetaboliteActivity/top.jsp		
07	生物種メタボロームモデル・データベースの構築	metabolomics.jp	本DBは、メタボロミクスを中心とした有田研究室の活動全般を対象としたポータルサイトです。フラボノイド、基礎代謝物、生薬、植物系統分類データベースのほか、ファイトレメディエーションを含めた放射線情報、講義資料も掲載しています。	http://metabolomics.jp		
08	生物種メタボロームモデル・データベースの構築	MFSearcher	本DBは、精密質量分析で得られた精密質量値から、組成式を迅速に推定するためのウェブサービスです。また、KNApSAcK、KEGG、PubChem、LipidMAPS、FlavonoidViewer等の化合物データベースに対して、精密質量値からの検索も高速に行うことができます。ハイスループットなピークアノテーションに利用されています。	http://webs2.kazusa.or.jp/mfsearcher/		
09	生物種メタボロームモデル・データベースの構築	MS-MS FragmentViewer	本DBは、フラボノイド標品のMS/MSフラグメントを化学構造に帰属させ、分子開裂モデルを提示しているデータベースです。化合物ピークのアノテーションに利用されています。	http://webs2.kazusa.or.jp/msmsfragmentviewer/		
10	生物種メタボロームモデル・データベースの構築	Metabolonote	本DBは、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)だけを専門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)をすることでだれでも気軽にご自身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ので、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ので、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/		
11	生物種メタボロームモデル・データベースの構築	KOMICS	本DBは、質量分析によるメタボローム解析とメタボロームを活用したオミクス解析のためにかずさDNA研究所でこれまで開発されてきた解析ツールやデータベースを公開しているポータルサイトです。LC-精密質量MSの強力な解析ツールPowerGetや、世界最大級のデータ量を持つ分析生データのレポジトリMassBase等、ユニークなリソースが整備されています。また、メタボロミクス関係のリソースへのリンク集など、これからメタボロミクス分野へ参入する研究者にとっても役立つ情報を提供しています。	http://www.kazusa.or.jp/komics/		
12	生物種メタボロームモデル・データベースの構築	KomicMarket	本DBは、主にLC-MSIによるメタボローム解析で得られた化合物ピークについて、化合物の同定・推定(アノテーション)を行った結果を蓄積・公開しているDBです。	http://webs2.kazusa.or.jp/komicmarket/		
13	生物種メタボロームモデル・データベースの構築	MassBank Fragmentation Library	マスマスペクトルの開裂パターンをWiki上に整理したDB	http://metabolomics.jp/wiki/Index:MassBank		
14	生物種メタボロームモデル・データベースの構築	Motorcycle DB	二次代謝酵素DB	http://kanaya.naist.jp/motorcycle/top2.html		
15	生物種メタボロームモデル・データベースの構築	Natural Activity	薬用植物、食用植物と生物への影響(ヒトへの効能)に関するDB	http://kanaya.naist.jp/BiologicalActivity/top.jsp		
16	生物種メタボロームモデル・データベースの構築	Metabolite Activity	代謝物と生物活性の関係DB	http://kanaya.naist.jp/MetaboliteActivity/top.jsp		
17	生物種メタボロームモデル・データベースの構築	World Map DB	世界で使われている薬/食用植物と使用国の関係DB	http://kanaya.naist.jp/KNApSAcK_World/top.jsp		
18	生物種メタボロームモデル・データベースの構築	KAMPO DB	漢方配合生薬における配合に関するDB	http://kanaya.naist.jp/kampo/top.jsp		
19	生物種メタボロームモデル・データベースの構築	JAMU DB	インドネシア配合生薬Jamuにおける配合に関するDB	http://kanaya.naist.jp/jamu/top.jsp		
20	生物種メタボロームモデル・データベースの構築	Tea Pot DB	ハーブとして使用される植物DB	http://kanaya.naist.jp/TeaPot/top.jsp		
21	生物種メタボロームモデル・データベースの構築	Lunch Box	食用生物DB	http://kanaya.naist.jp/LunchBox/top.jsp		
22	生物種メタボロームモデル・データベースの構築	Diet Navi	食用生物DB	http://kanaya.naist.jp/DietNavi/top.jsp		
23	生物種メタボロームモデル・データベースの構築	FoodProcessor	加工品に含まれる食用生物DB	http://kanaya.naist.jp/FoodProcessor/top.jsp		