

独立行政法人科学技術振興機構
バイオインフォマティクス推進センター事業

生命情報データベースの高度化・標準化

第Ⅱ期 研究開発課題

(追跡調査報告書)

研究開発課題：メタボローム MS スペクトル

統合データベースの開発

(平成18年4月～平成23年3月)

代表研究者氏名：西岡 孝明

(奈良先端科学技術大学院大学情報科学研究科 特任教授)

目次

本編

1. 研究開発による成果.....	1
1.1. 研究開発課題の目標及び新規性・独創性	1
1.2. 研究開発開始時の課題を取り巻く研究・技術水準及び分野における 課題の位置づけ	1
1.3. 研究開発終了時の成果概要	2
1.4. 研究開発の達成度	3
2. 研究開発による成果の活用状況や発展状況.....	4
2.1. 構築されたデータベース・ソフトウェア等の活用状況.....	4
2.1.1. データベース・ソフトウェア等の継続状況や発展状況	4
2.1.2. 第三者によるデータベース・ソフトウェア等の活用事例	5
2.1.3. データベース・ソフトウェア等へのアクセス数／ダウンロード数、visit 数 .	6
2.2. 課題終了後の研究開発成果の継続状況や発展状況.....	9
2.2.1. BIRD 終了から現在まで、BIRD で取り組んだ課題に関連した研究開発成果の 継続状況や発展状況（国内外の研究者との共同研究へ発展した等）	9
2.2.2. BIRD 終了後に発表された論文リスト.....	10
2.2.3. 研究開発成果の現在の国際的な評価・位置づけ	10
2.3. 現在の科学技術研究における研究開発成果の活用状況、発展状況のまとめ.....	11
3. 研究開発による成果の科学技術的、社会的及び経済的な効果.....	12
3.1. バイオインフォマティクス分野及びライフサイエンス分野の研究への貢献.....	12
3.1.1. バイオインフォマティクス分野の研究への貢献	12
3.1.2. ライフサイエンス分野（バイオインフォマティクス以外）や その他、科学技術分野の研究への貢献	13
3.2. 人材育成の面から参加研究者の活動状況	14
3.2.1. 研究開発に参加した研究者のキャリアアップ	14
3.3. 社会的及び経済的な効果	14
3.3.1. 研究開発成果が大学や公的研究機関、企業等で、応用に向けて 継承または発展した例	14
3.3.2. その他、研究開発成果が社会的、経済的な効果・効用につながる兆し、 可能性	14

資料編

1. 論文リスト	21
2. 主要論文の被引用回数.....	24

3.	学会招待講演・基調講演.....	26
4.	新聞発表等	26
5.	特許出願・成立	26
6.	学会賞等の受賞	27
7.	グラントの獲得実績.....	27
8.	書籍等の執筆実績	27
9.	総説の執筆実績	27
10.	参加研究者の活動状況.....	28

本調査報告書は平成 26 年 4 月に作成
表紙の代表研究者所属は平成 25 年 10 月時点

1. 研究開発による成果

1.1. 研究開発課題の目標及び新規性・独創性

本研究課題は、二次代謝物質の質量・化学構造式・生物種等の収集データベースと、生体成分を各種の質量分析計によって測定したマスペクトルのデータベースを統合し、ピーク同定に必要な化学構造相関知識データベースの開発に取り組むことを目標とした。

具体的には、化合物のマスペクトル (MS) データを収集した“MassBank”、二次代謝物の化学および生物学、生化学、生理活性などに関するデータや知識を収集した“KNAPSAcK”と“Metabolomics.JP”を新たに構築する。さらに、これら3つのデータベースを統合することによって、検出した二次代謝物の同定、あるいは化学構造の推定、生合成経路の推定を行うことを目指した。

これらの研究開発によって、二次代謝物をゲノム情報とリンクすることが可能になる。さらに omics 研究が大量データを出力している現状において、生データを単に公開するためのデータベースと、研究者が知的創造や知識集約をおこなって社会で共有するためのデータベースの2つを低コストで維持する学術データベースモデルを実現する。

1.2. 研究開発開始時の課題を取り巻く研究・技術水準及び分野における課題の位置づけ

メタボローム研究では1回の質量分析で数千の二次代謝物を検出し、あらかじめ測定しておいた化合物のMSデータと照合することによって、検出した二次代謝物を網羅的に同定する。しかしながら、メタボローム研究に適したMSデータベースがこれまで無かったために、同定できる二次代謝物は少なかった。また二次代謝物をゲノム情報と関連付けることを目的とするデータベースも研究開発開始時には無かった。

そこで本研究課題では、MSデータの分散型データベースを新たに開発し(“MassBank”)、同時に関連データベースの開発と充実を図った。MSデータの蓄積には、メタボローム解析や天然物化学の研究グループの協力を得ることが必須である。これらを考慮してMassBankは、研究者コミュニティで化合物MSデータを共有するpublic repositoryとし、低コストで維持するために研究グループごとに自らデータサーバを設置し、MSデータを登録、管理する分散型データベースとして設計することにした。

なお、メタボローム研究に頻繁に使用されるエレクトロスプレーイオン化タンデム質量分析法 (ESI-MS²) は、得られるスペクトルが測定条件に大きく依存するため、測定条件が異なる研究グループ間ではスペクトルデータの共有や代謝物同定ができない。そこで本研究開発課題では、公開された化合物ESI-MS²データを集約・分析することによって測定条件に依存しない化合物MSデータ(“merged ESI-MS²データ”)を人工的に作成することにした。

さらに、通常、分子の質量を電荷で割った数値 (m/z 値) で表されるスペクトル上のピーク

ークを、分子式として表現すれば、観測結果と化学構造とを関連づけることができると考え、merged ESI-MS²データについて化学的注釈を行うことにした。この化学的注釈は、新規に開発したウィキ上で作業し、**Fragmentation Library** として収集した。

国際的に見ても、MS データフォーマットの標準化は他分野に比べて遅れており、研究開発開始時からの緊急の課題であった。研究開発開始時（2006年）、世界で最も信頼され普及している MS データベースとして、米国の NIST/EPA/NIH Mass Spectral Library（有償）が、日本には産総研の提供する RIO-DB があったが、これらはいずれも EI 法（電子イオン化法）で測定したマススペクトルを収集していた。

しかしながら、メタボローム研究の主流となっている LC-MS（液体クロマトー質量分析）では、ESI（エレクトロスプレーイオン化）法または APCI（大気圧化学イオン化）法で分子のイオン化を行うため、EI 法で測定したデータは同定に使うことができない。ESI または APCI の測定データを有する DB には米国スクリプス研究所による METLIN があったが、これは Agilent 社の装置で分析した場合にのみ適用できる限定的なものであった。

このような現状を踏まえ、本研究グループは、アジア・オセアニア質量分析会議（2010年6月から毎年開催）や国際質量分析会議（第19回は2012年京都）などの機会に、新たに構築する MassBank を実質的な国際標準として提案することを目指した。

1.3. 研究開発終了時の成果概要

1) 化合物 MS データをメタボローム研究者コミュニティで共有するための、世界で最初の public repository である”MassBank”を開発した。これは、2008年から日本質量分析学会の公式 DB として認められ、19 研究グループ（日本 15、米国 2、ドイツ 1、中国 1）が 11,614 化合物について測定した 30,312 件のデータを 8 つのサーバ（うちドイツ 1、中国 1）から公開した。

MassBank は以下の特徴を有する。

- ・複数の研究機関が国際的に協力して大規模なデータベースを構築・維持するというコンセプトの上に成り立っている。
- ・特定の研究機関がサーバやデータを所有するという形を取らず、ウェブサイトにも機関等のロゴなどを掲載していない。
- ・スペクトルの諸権利はデータ提供者が所持し続けるため、データの公開だけでなく撤回もできる。
- ・学会の公式データベースとしたため、プロジェクト終了と同時に終了することができない仕組みになった。
- ・ウィキを多用してコンテンツの維持管理費を低く抑えている。

2) これらの MS データから、分析条件に依存しない化合物参照 MS データを作成し提供することによって、メタボローム研究における二次代謝物の高精度な同定を実現した。

- 3) 二次代謝物と生物種との関係を原著論文から抽出、収集することによって、“KNApSAcK”を構築した。20,018種の生物種における50,048種の代謝物について101,500対の生物種 - 代謝物の関係を整理し公開した。
- 4) 脂質DBである“LipidBank”の化合物情報や、フラボノイド類を含む二次代謝物全般を扱うウィキ形式サイト“Metabolomics.JP”を構築した。Metabolomics.JPではLipidBankに記載の脂質2,705件の化学構造式を確認し、cis/transを含めた正確な構造表記をおこなった。基礎代謝物およそ1,500件について化学構造に基づく12桁の階層分類番号を割り振り、フラボノイド、生薬や漢方情報の集約をウィキ上で行った。
- 5) 未同定代謝物の化学構造推定アルゴリズムを開発した。方法としては、まず分子イオンのm/z値を満たす候補代謝物をKNApSAcKやMetabolomics.JPから選び、次にESI-MS²データで観察されたピークについて部分化学構造式を推定し、この部分化学構造式を満たす候補代謝物を未同定代謝物の候補とする、というものであった。このようにしてESI-MS²データの化学的注釈をおこなうことによって、「ESI-MS²データと部分化学構造式との関係」を蓄積し、新しく開発したウィキ上にFragmentation Libraryとして収集した。
- 6) 大量の生データから新たな知識を抽出・創造して、それらを公開・共有するという、将来のomics研究や学会活動に関する提案を実現するためのメタボローム研究を推進し、二次代謝物をゲノム情報とリンクすることに役立つデータの収集と解析ツールを開発した。また、研究者コミュニティが維持するデータベースとして、今後のモデルとなりうるシステムを提案した。

1.4. 研究開発の達成度

二次代謝物質のDBであるKNApSAcKとMetabolomics.JP、MSデータのDBであるMassBankを構築し、これらは収集したデータ件数をはじめ当初計画をほぼ達成した。特にMassBankはこれまで存在しなかったESIマスペクトルの無償データベースを構築し、日本質量分析学会の公式DBに認定されたほか、世界中の研究者に認知されるまでに至った。

さらには、メタボロミクスや質量分析という学問分野だけでなく、栄養学、食品科学の分野においてMassBankが頻繁に使われるようになった。また、ドイツの研究協力機関(IPB, Halle)がMassBankソフトウェアの独自拡張を開始し、研究機関毎に異なるサービスを施すネットワークとしても動き始めた。

民間の質量分析機器メーカーからデータの無償提供の問い合わせが来るようになり、企業が販売するソフトウェア等にもデータが利用されることになった。

一方で未達成に終わった点としては、MassBankの「ESI-MS²データと部分化学構造式

との関係」を当初計画した以上に収集することができたため「関係の特異性の評価」を終えることができなかった。そのため未同定代謝物の自動推定アルゴリズムを提案したものの、ツールを実装するまでには至らなかった。

2. 研究開発による成果の活用状況や発展状況

2.1. 構築されたデータベース・ソフトウェア等の活用状況

2.1.1. データベース・ソフトウェア等の継続状況や発展状況

	ツール名	URL	公開状況	備考
1	MassBank	http://www.massbank.jp/	公開中 (無料)	
2	Metabolomics.JP	http://metabolomics.jp/wiki/Main_Page	公開中 (無料)	・生薬データベースや植物細胞培養プロトコルデータベースが追加されるなど、第三者によって幅広く活用されている (後述)
3	LipidBank	http://lipidbank.jp/	公開中 (無料)	・課題期間中に日本脂質生化学会の公式 DB と指定を受けた ・さらに日本動脈硬化学会とも連携して開発することになった ・また The Lipid Library へ脂質分子の和名について情報を提供
4	KNApSAcK (生物種 - 代謝物 関係データベース)	http://kanaya.naist.jp/KNApSAcK/	公開中 (無料)	・使用例は添付資料の表 1 に示した
5	Lunchbox DB (食履歴データベース)	http://kanaya.naist.jp/LunchBox/top.jsp	公開中 (無料)	
6	ESI-MS ² データと 化学構造式との 関係データベース	—	—	・269 イオンと 94 部分構造式との関係を解析した計 630 件のデータが収集され、MS/MS データから化学構造式を予測するツール (MassBank Metabolite Prediction) (http://www.massbank.jp/MetaboPrediction.html) に利用された。

MassBank は、2～3 年前までは、国内、ヨーロッパからのアクセスが大部分であった。しかし、昨年から、これらに加えて米国からのアクセスが増加している。さらに、中国、韓国、ブラジル、南アフリカからのアクセスも増えている。

また、MassBank プロジェクトで制作された、MS/MS データから化学構造式を予測する”MassBank Metabolite Prediction” (<http://www.massbank.jp/MetaboPrediction.html>) は、ESI-MS² データと化学構造式の関係を利用したツールである。

KNAPSAcK 生物種 - 代謝物関係データベースは、現在までに、100 を越えるメタボローム研究で使用されている。インドネシア生薬データベースは、プロジェクト終了後、インドネシア・ボゴール農大との共同で薬粧植物配合の最適化プロジェクトへと進展している。

2.1.2. 第三者によるデータベース・ソフトウェア等の活用事例

MassBank を利用するためのツールが第三者によって開発されている。Mass++ と RMassBank の例を紹介する。

これまで質量分析計から直接 MassBank を検索することはできなかった。それぞれの質量分析計メーカーに特異的な形式で raw data が保存されているために、互換性が無く、ユーザは MassBank を利用することや、公開をするために MassBank 形式でレコードを作成することは困難であった。このような研究者の要望に応えるため、MassBank では API を提供してきたが、それを利用する application program の開発がされていなかった。

MassBank プロジェクトは、JST-CREST「代謝調節機構解析に基づく細胞機能制御基盤技術」の研究課題「定量的メタボロミクスとプロテオミクスの融合 (2005 年採択課題)」(研究代表者、小田吉哉：エーザイ) で開発された Mass++ と連携してきた。2009 年からは「田中耕一最先端研究」に引き継がれて (<http://www.first-ms3d.jp/message/tanaka/247.html>) Mass++ の開発が進められた。

2013 年には、ほとんどの機器メーカーの raw data を入力して、ピーク検出、MassBank 検索のためのクエリデータ作成、MassBank レコード形式への変換などを容易に実施することが可能になった。とくに、メタボロミクス研究では生物試料の一回の分析によって数百から数千の MS/MS データが得られる。

これらの代謝物の同定は、MassBank の Batch API サービスを利用して Mass++ から一挙に実施することが可能になった。これはメタボロミクス研究者にとって革命的なできごとである。Mass++ は無料で誰でも download して利用することができる。

MassBank を公式データとした NORMAN Network では、質量分析計から出力される raw data を MassBank レコード形式に自動変換して MassBank にデータを登録するツールの開発をおこなっていた。RMassBank は統計パッケージである R を利用したもので、化学構造情報の獲得までを全自動でおこなうものである (M. Stravs, et al., J. Mass Spectrom., 48, 89-99, 2013)。

Mass++ や RMassBank のように MassBank の利用、普及を促進するツールが第三者パ

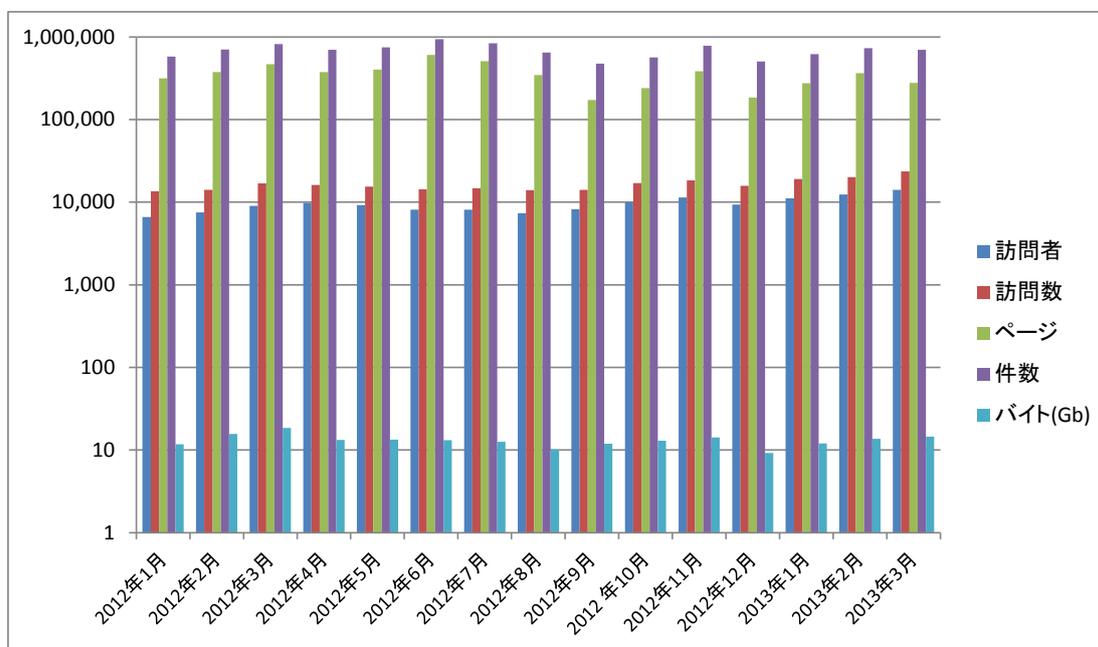
ーティによって開発されている。添付資料の表 1 で示したように、生物種 - 代謝物関係データベース KNApSAcK Core はさまざまな分野で応用活用されている。その分野は、農学、生態学、医学の広範な分野で使用されている。

2.1.3. データベース・ソフトウェア等へのアクセス数/ダウンロード数、visit 数

	ツール名	訪問者数、訪問数、参照ページ数、スペクトル検索件数	集計期間	備考
(1)	MassBank	訪問者：119,218 訪問数：202,451 ページ：4,137,430 件数：8,250,147	2012年4月～2013年3月	
(2)	Metabolomics.JP	訪問者： 11,000～15,000/月		
(3)	LipidBank	訪問者： 3,400～4,700/月		
(4)	KNApSAcK 生物種 - 代謝物関係データベース	アクセス： 10 万件/月		海外の 100 以上のドメインからアクセス

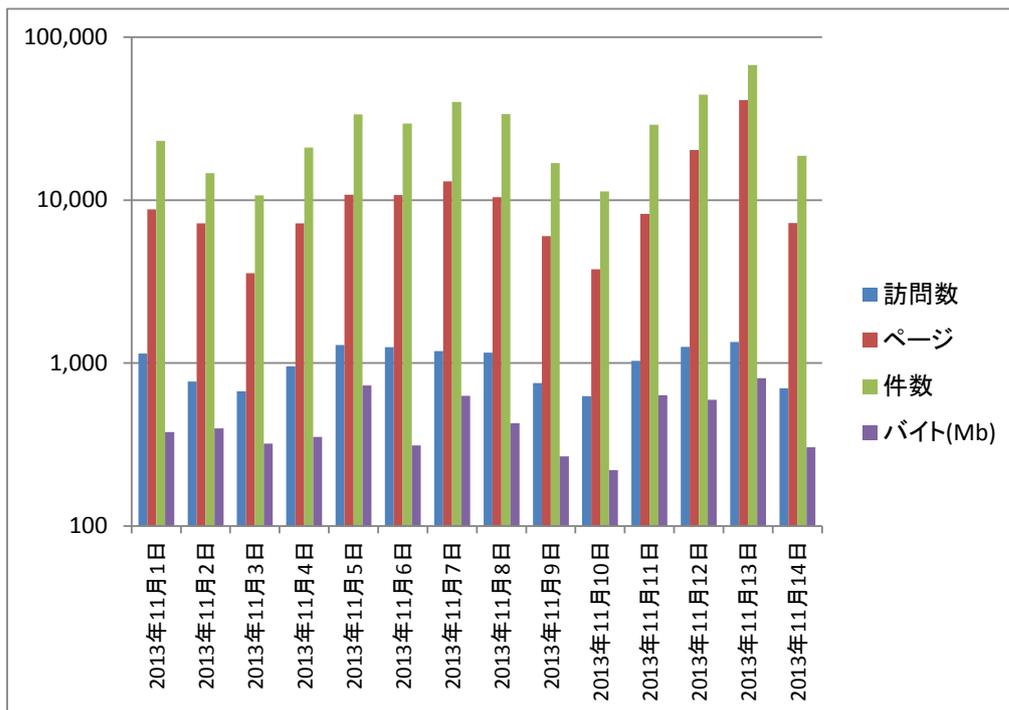
データベース MassBank へのアクセス統計は以下のようにになっている。

2012年1月～2013年3月



	訪問者	訪問数	ページ	件数	バイト(Gb)
2012年1月	6,615	13,520	315,834	578,465	11.78
2012年2月	7,525	14,102	376,838	702,028	15.73
2012年3月	9,041	16,888	469,260	817,191	18.48
2012年4月	9,815	16,101	374,276	698,479	13.29
2012年5月	9,241	15,426	403,467	748,197	13.41
2012年6月	8,085	14,292	605,229	939,026	13.14
2012年7月	8,115	14,775	507,983	838,156	12.56
2012年8月	7,361	14,005	345,139	645,887	10.15
2012年9月	8,235	14,032	172,698	475,150	11.95
2012年10月	9,979	16,976	239,891	566,856	12.97
2012年11月	11,388	18,306	385,024	781,803	14.17
2012年12月	9,361	15,739	184,910	505,903	9.21
2013年1月	11,175	19,087	275,680	619,920	12.04
2013年2月	12,420	20,078	364,658	731,854	13.72
2013年3月	14,043	23,634	278,475	698,916	14.49
合計 2012年1月 - 2012年12月	104,761	184,162	4,380,549	8,297,141	156.84
合計 2012年4月 - 2013年3月	119,218	202,451	4,137,430	8,250,147	151.1

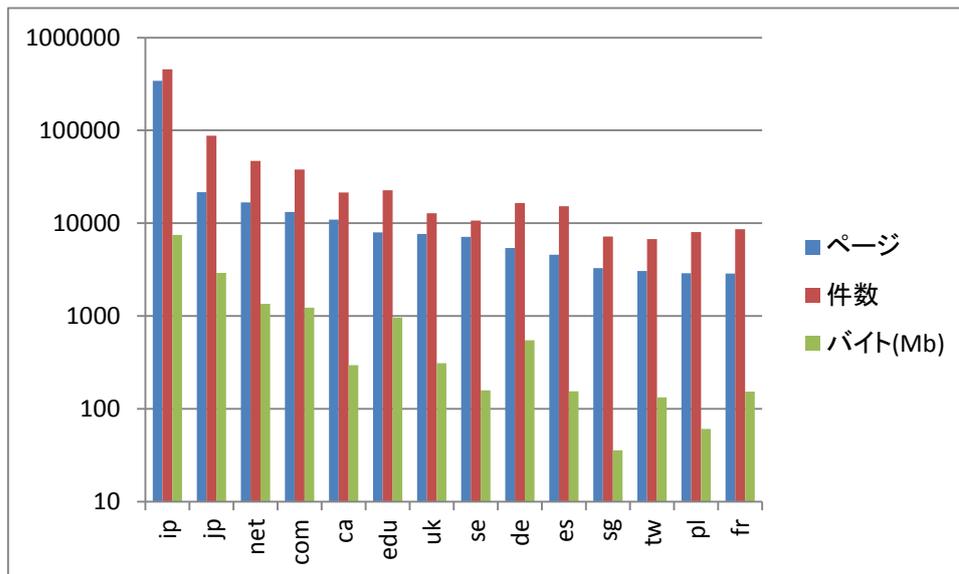
最新状況 (2013年11月1日～14日)



	訪問数	ページ	件数	バイト(Mb)
2013年11月1日	1,144	8,764	23,072	378.36
2013年11月2日	771	7,186	14,631	397.39
2013年11月3日	673	3,563	10,676	320.34
2013年11月4日	956	7,214	20,997	353.55
2013年11月5日	1,291	10,771	33,539	729.48
2013年11月6日	1,255	10,746	29,440	312.52
2013年11月7日	1,185	13,025	40,038	629.56
2013年11月8日	1,158	10,403	33,590	428.47
2013年11月9日	756	6,015	16,888	268.28
2013年11月10日	628	3,759	11,284	221.26
2013年11月11日	1,033	8,221	28,930	634.42
2013年11月12日	1,261	20,299	44,234	595.01
2013年11月13日	1,346	41,010	67,302	807.4
2013年11月14日	702	7,240	18,670	304.37

過去最多アクセス数更新

ドメイン／国名別アクセス状況 (2013年3月)



ドメイン/国名		ページ	件数	バイト
不明	ip	341,175	454,711	7.51 Gb
Japan	jp	21,650	87,491	2.92 Gb
Network	net	16,726	46,933	1.35 Gb
Commercial	com	13,275	37,816	1.23 Gb
Canada	ca	10,926	21,513	295.94 Mb
USA Educational	edu	7,955	22,575	963.28 Mb
United Kingdom	uk	7,672	12,839	311.43 Mb
Sweden	se	7,116	10,712	158.67 Mb
Germany	de	5,389	16,514	550.01 Mb
Spain	es	4,588	15,225	154.68 Mb
Singapore	sg	3,297	7,186	35.86 Mb
Taiwan	tw	3,066	6,775	133.48 Mb
Poland	pl	2,895	8,053	60.67 Mb
France	fr	2,868	8,609	154.08 Mb

2.2. 課題終了後の研究開発成果の継続状況や発展状況

2.2.1. BIRD 終了から現在まで、BIRD で取り組んだ課題に関連した研究開発成果の継続状況や発展状況（国内外の研究者との共同研究へ発展した等）

2012年11月 EU加盟国の環境科学研究機関の連合体である NORMAN Network グループが MassBank を NORMAN の公式データベースとして採用した (<http://www.normandata.eu/>)。これを記念して、2012年11月27日に NORMAN が MassBank ワークショップを VU アムステルダム大学 (アムステルダム、オランダ) で開催し、西岡の招待講演と MassBank ワークショップを開催した。参加者は、17ヶ国から39名、オーガナイザ5名であった。

2013年1月から UFZ (ライプツヒヒ、ドイツ) に MassBank サーバを設置し、NORMAN MassBank の正式運用を開始した。MassBank.jp と相互に完全に同じデータを共有している。

現在、スイス連邦水質科学研究所 (EAWAG) とヘルムホルツ環境科学研究センター (UFZ, ドイツ) を併せて 382 化合物について最新の質量分析装置を用いて分析した ESI-MS/MS データ 6,280 件が公開されている。

LipidBank データベースは日本動脈硬化学会と連携して開発することになった。また The Lipid Library (<http://lipidlibrary.aocs.org/>) と脂質分子の和名について情報を送った。

インドネシア薬用生薬データベースについては、インドネシア・ボゴール農業大学とのインドネシア薬用配合生薬 *Jamu* の共同研究を展開し、人の症状と配合の最適化研究としての成果が得られており、以下の論文で公表されている。

- [1] Afendi FM et al., Data mining methods for omics and knowledge of crude medicinal plants towards big data biology, *Comput. Struct. Biotech. J.*, 4, e20131010.1-14, 2013
- [2] Afendi FM. et al, KNApSAcK Family Databases: Integrated Metabolite: Plant Species Databases for Multifaceted Plant Research,, *Plant Cell Physiol.* 53(2): e1(1-12) (2012)

2.2.2. BIRD 終了後に発表された論文リスト

BIRD 終了後には論文は発表されていないが、*The Scientist* 誌のレポートからインタビューされ、記事の中で MassBank が引用されたことがあった (Jeffrey M. Perkel, “Name That Metabolite! A guided tour through the metabolome”, *The Scientist*, 27(7), 61-61 (2013))。

過去の論文の分野別被引用件数においては、化学・生物分野に加えて、コンピュータサイエンスや工学分野からも引用されるなど、メタボローム研究をこえた幅広い分野から注目を浴びていることがわかる。

さらに、主要論文については発表されてから年々引用回数が増えている論文があるなど、BIRD 終了後も MassBank の関連研究が発展していることが伺える。

2.2.3. 研究開発成果の現在の国際的な評価・位置づけ

Web of Knowledge として世界の論文や学会誌を格付けしている Thomson Reuters 社はデータベース (public repository) を対象とした、データの格付けをおこなって公開している (Data Citation Index: DCI, http://wokinfo.com/products_tools/multidisciplinary/dci/)。2013 年夏に Thomson Reuters 社から、MassBank の全データを対象として DCI を行いたいという提案があり、MassBank のデータを提供している研究者の意見を聴いて、DCI を受け入れることにした。

現在、Thomson Reuters 社では統計収集の準備とメタデータの収集をおこなっているところである。おそらく来年から MassBank の DCI が公開されるのであろう。DCI 化によってデータ公開を行う研究者が増えることや、データの信頼性に関する評価がユーザに歓迎されるものと期待している。

アメリカ (=世界最大) の質量分析学会である American Society of Mass Spectrometry (ASMS) の年会では MassBank を使って研究した発表が増えてきている。「MassBank は世界標準になった」と評価する研究者がふえてきていることを感ずる。このことは発表論文が国外から多く引用されていることから分かる。2012 年 9 月に京都で開催された第 9 回国際質量分析学会では、ワークショップ「Mass++ and MassBank: Tools for Data

Processing and Database on PC」が開催された。

生物種 - 代謝物データベース KNAPSAcK Core については植物並びに微生物におけるメタボローム研究における中核データベースとして位置づけられている。この数カ月は、日本国内を越えるアクセス数が米国ならびに中国において観測されており、グローバルスタンダードとなりつつある。

2.3. 現在の科学技術研究における研究開発成果の活用状況、発展状況のまとめ

メタボローム解析に用いられている市販の質量分析計の m/z 値の測定精度は研究開始時に比べて約 50 倍以上向上した。MassBank で公開されているデータは、開発開始時 (2006 年) には m/z 値は整数値であった。2010 年ごろに小数点以下 2 桁の精度 (測定誤差 50ppm)、2013 年には測定誤差 5ppm 以上の高精度なデータも MassBank で公開がはじまっている。

現在、国内のいくつかの研究室にはすでに最新型モデルの装置 (測定誤差 1ppm 以上) の設置がはじまっている。このように分析計の高精度化とそれらのメタボロミクス研究への普及は極めて早い。MassBank は public repository であるので、最新装置で分析されたデータが速やかに公開されている。

これに対して、これまでの大規模なマススペクトルデータは大型分析センターに設置した装置を使って得られたデータを収集しているものの、最新型モデルの分析機器への置き換えは (高額機器であるので) 容易ではない。従って、大規模な有料データベースで収集したデータは時代遅れとなりつつあり、ユーザの要望に対応しきれていない。現在、MassBank のアクセスが増加している背景にはこのような分析機器の進歩にある。

MassBank の高精度化に対する対応したデータ : MassBank でこれまで公開されているデータのうち測定精度が 50~30 ppm 以上のデータは 16 千件収集されている。これらもいずれは時代遅れになることは必須である。そこで、2011 年から MassBank データを数値 (m/z) だけではなく、分子式で表現する化学的注釈付けをおこなって、Chemically Accurate Data として提供する作業を進めてきた。分子式は測定誤差を含まない exact な m/z の表現である。

将来にわたって、究極的に高精度な分析計で分析したデータにも対応した MassBank データになった。2013 年内にこの作業を終了し、Chemically Accurate Data に対応した検索ツールの開発も終えたので、来年度には実用に供する予定である。

3. 研究開発による成果の科学技術的、社会的及び経済的な効果

3.1. バイオインフォマティクス分野及びライフサイエンス分野の研究への貢献

3.1.1. バイオインフォマティクス分野の研究への貢献

MS データはこれまで質量分析機器メーカー間で互換性の無い binary 形式で保存、管理されていたが、このことが MS データの交換・共有を阻害していた。これを解決するために本研究開発を推進し、MS データを MassBank で一元管理する研究室や機器メーカーが増えてきた。

このように、データベース維持のコストを研究者がデータ量に比例して分担し、データベースは研究者コミュニティがデータを共有するだけでなく、知識の集約や創造を行う場を提供する、というモデルが実現したことは、今後のバイオインフォマティクス分野において応用の可能性の高いモデルとなることが考えられる。

学術データベースのうち、分散型データベースを採用して、研究に寄与している MassBank は極めて珍しい。MassBank は最新の情報技術を利用して、大型スーパーコンピュータを採用しないで並列処理やクラウド型のデータ処理を実現して、低コストの維持管理費で運用されている。このことが評価されて、2009 年 12 月には日本バイオインフォマティクス学会 20 周年記念特別セッション 3 課題の 1 つに選ばれた。

しかし分散型データベースには大きな欠点がある。分散しているデータサーバのどれか 1 つでも（停電などによって）サービスを一時的に停止していると、検索は停止しているデータサーバに格納されているデータを検索することができない。その結果、MassBank では検索する日時によって検索結果が異なることになる。とりわけ東日本震災後の停電によって、米国の研究者たちは特にこのことを問題視していた。

そこで、文書管理システム Subversion を利用して、分散している全てのデータサーバ上にあるデータのコピーを一カ所に集めておくことにした。どこかのデータサーバがサービスを停止していると、そのサーバのコピーを検索することにした。2013 年 1 月以降、MassBank.jp と MassBank.eu は常時このようにして集めた最新の全データの同じコピーを持っている。地球上で東西にデータを置くことによって、万が一、震災や戦争にあっても、常に同じ検索結果を返すことが可能になった。このように MassBank は分散型データベースでありながら、集中型データベースとしての機能を有する「ハイブリッド型データベース」として運用している。将来の国内の大規模災害にも影響をうけない、安定なサービスを国際的に提供している。

MassBank は代謝物予測ツールの開発研究にも寄与している。MassBank から公開されているデータは化合物を質量分析したものであるが、このデータはメタボロミクス研究で質量分析によって検出される代謝物の構造予測に利用されている。

現在、MassBank で公開されている化合物の数は、生物界にある総代謝物（約 20 万化合

物と推定されている)の約6%にすぎない。そのためメタボロミクス研究では同定できない代謝物(unknownと呼ばれる)が、同定された代謝物の数をはるかに上回っている。したがって、MSデータから化学構造式や代謝物を推定するツール(代謝物予測ツール)の研究開発がすすめられている。これらツールにはマススペクトルと化学構造式との関係に関する知識の収集が欠かせない。MassBankのデータはこのような関係を解析するデータとして多くの研究に利用されている。

特に、2013年1月に最新の全MassBankデータのコピーが可能になったことによってさまざまな研究で利用されている。現在、推定ツールとして広く利用されているMetFusion(M. Gerlich and S. Neumann, J Mass Spectrom., 48, 291-298, 2013)はMassBankを利用している。

また、KNApSAcKでも候補二次代謝物質の予測ソフトウェアを開発した。<http://kanaya.naist.jp/KNApSAcK/>のサイトより無償でダウンロードできる。このソフトウェアを使ったメタボローム研究も増えてきた。

さらに、Omicsデータ処理を容易にし、マススペクトルからの構造推定をおこなう研究を促した。生物種-代謝物関係データベースの活用例を(添付資料の表1)に示す。この表における研究目的におけるDBならびにBioinfoがバイオインフォマティクス研究への応用例であり、生物種-代謝物関係データベースKNApSAcK Coreはさまざまな分野で応用活用されている。

3.1.2. ライフサイエンス分野(バイオインフォマティクス以外)やその他、科学技術分野の研究への貢献

本研究によって構築されたMassBankは2008年6月から日本質量分析学会の公式データベースとなり、国際メタボロミクス学会の中でも無償データベースの草分けとして認知されている。また、これに触発された米国NIHによるファーマコメタボロミクス研究ネットワークやEU EBIがプラットフォームや生物種を網羅する大規模なメタボロームのデータベースMetaboLightsの構築を開始した。

このように、メタボローム研究のボトルネックであった二次代謝物の同定率の向上と推定が可能になることによって、ライフサイエンス分野におけるメタボローム研究が活性化されると期待できる。さらに、二次代謝物がゲノム情報とリンクされることによって、医薬開発や診断、予防、健康維持などの科学技術分野に大きく寄与するものと予想できる。

MassBankの論文(H. Horai, et al., 2010; DOI: 10.1002/jms.1777)が、現在までの3年間で142論文で引用された(J. Mass Spectrometry誌の過去3年で最も引用回数が多かった論文に選ばれた)。引用した論文の掲載誌の学問領域は、薬学、医学、農学、環境科学、植物学など多様な分野にわたっている。このことから、質量分析を利用している多様な分野で活用されていることがわかる。国内外の学会やシンポジウムでMassBankに関する発表、講演をおこなうと、講演の後、企業からの参加者から多くの質問をうけることや、「と

でも頻繁に利用させていただいています」という感謝をされることが多い。このことから、関連分野の企業でも MassBank を利用していることが窺われる。

現在では「質量分析を利用している」ということは MassBank を利用している、ということと理解されており、Global Standard として認められていることを実感できる。2008 年 3 月に MassBank は日本質量分析学会の公式データベースとして認められた。2012 年 3 月には MassBank の開発研究の支援、データを利用することを研究することを目的として、スペクトルデータ部会が設置された。

3.2. 人材育成の面から参加研究者の活動状況

3.2.1. 研究開発に参加した研究者のキャリアアップ

本研究開発課題に参加した者で、その後新たなポジションを得た者は 7 名いた。

3.3. 社会的及び経済的な効果

3.3.1. 研究開発成果が大学や公的研究機関、企業等で、応用に向けて継承または発展した例

応用に向けて継承、発展した例としては先に記述した NORMAN Network がある。質量分析機器メーカーである島津製作所、日本電子、分析機器メーカーである GL サイエンス社、Waters 社は MassBank にデータを提供し、MassBank 利用ツールを開発することによって MassBank を継承、発展しようとしている。また、国外の質量分析機器メーカー Brucker 社は MassBank データを検索するためのインターフェースを提供している。

クックパッドとの連携により、現在、食材生物と機能性成分の関係に関するデータベース構築を進めるなどの研究へと発展がみられている。また、理化学研究所との連携で、データベースの充実を図っている。

3.3.2. その他、研究開発成果が社会的、経済的な効果・効用につながる兆し、可能性

分散型データベースとウィキを組み合わせた MassBank は、将来の研究者コミュニティ（学会）で維持するデータベースの実用化モデルである。

研究者が公開する MS データ量に応じて MassBank の維持・管理にかかる経費を分担するので、従来型データベースよりはるかに低コストのデータベースを実現している。

さらに、ウィキでは Fragmentation Library や二次代謝物質の DB を構築している点では、社会的、経済的な効果・効用につながる可能性がある。

表1 KNApSack を活用した研究例および紹介例

ID	論文	研究目的	対象生物種
'06			
K1	Shinbo Y., et al., <i>Biotechnol. Agriculture and Forestry</i> , 57 , 165-181 (2006)	DB: KNApSack DB	
K2	Kikuchi, K. and Kakeya, H., <i>Nature Chem. Biol.</i> , 2 , 392-394 (2006)	Review: Bridge between Chemistry and Biology	
K3	Oikawa, A. et al., <i>Plant Physiol.</i> , 142 , 398-413 (2006)	Bioinfo: Metabolome platform DrDMASS in FT-ICR-MS Metabolite accumulation in different herbicidal modes of action	<i>Arabidopsis thaliana</i>
K4	Shinbo Y. et al., <i>Comput. Aided Chem.</i> , 7 , 94-101 (2006)	Bioinfo: Taxonomic diversity of flavonoids	
K5	Tomiki, T., <i>Comput. Aided Chem.</i> , 7 , 157-162 (2006)	DB: Chemical biology	
'07			
K6	Gaida, A. and Neumann, S., <i>J. Int. Bioinformatics</i> , 1-8, (2007)	Bioinfo: MS Peak storage and processing	
K7	Hummel, J., et al., <i>Topics in Current Genetics</i> , 18 , 75-95 (2007)	Review: GC-MS DB	
K8	Ohta, D., et al., <i>Anal. Bioanal. Chem.</i> , 389 , 1469-1475 (2007)	Metabolite accumulation patterns caused by herbicidal enzyme inhibitors	<i>Arabidopsis thaliana</i>
K9	Moco, S., et al., <i>Trends in Anal. Chem.</i> , 26 , 855-866 (2007)	Review: Metabolomics technologies	
K11	Saito, K. et al., <i>Trends in Plant Sci.</i> , 13 , 36-43 (2007)	Review: Functional genomics research strategy of combining transcriptome and metabolome	
K11	Yonekura-Sakakibara, K. et al., <i>J. Biol. Chem.</i> , 282 , 14932-14941 (2007)	Assignment of UGT89C1 to a flavonol 7-O-rhamnosyltransferase	<i>Arabidopsis thaliana</i>
K12	Want, E. J. et al., <i>J. Proteome Res.</i> , 6 , 459-468 (2007)	Review: The role of MS in metabolomics	
K13	Nakamura, Y. et al., <i>Planta</i> , 227 , 57-66, (2007)	Light/dark regulation of metabolite activities	<i>Arabidopsis thaliana</i>
'08			
K14	Akiyama, K., <i>In Silico Biol.</i> , 8 , 339-345 (2008)	DB	
K15	Bottcher, C., et al., <i>Plant Physiol.</i> , 147 , 2107-2120 (2008)	Characterization of mutants in flavonoid and phenylpropanoid biosynthetic pathways	<i>Arabidopsis thaliana</i>
K16	Dunn, W. B., <i>Phys. Biol.</i> , 5 , 1-24 (2008)	Review: Mass spectrometry platforms	
K17	Fardet, A. et al., <i>J. Nutrition</i> , 138 , 1282-1287 (2008)	Metabolism of dietary phytochemicals	Rat
K18	Fait, A. et al., <i>Plant Physiol.</i> , 148 , 730-750 (2008)	Metabolic networks in primary and secondary pathways for achene and receptacle	<i>Fragaria x ananassa</i>

K19	Giavalisco, P., et al., <i>Anal. Chem.</i> , 80 , 9417-9425, (2008)	A combination of high-resolution mass spectrometry and ¹³ C-isotope labeling of entire metabolomes	
K20	Hagel, J. and Facchini, P., <i>Phytochem. Rev.</i> , 7 , 479-497 (2008)	Review: Metabolomics technologies and functional genomics platform	<i>Papaver somniferum</i>
K21	Hanhineva, K., et al., <i>Phytochemistry</i> , 69 , 2469-2481, (2008)	Phenolic biosynthesis pathway	<i>Fragaria x ananassa</i>
K22	Hanhineva, K., <i>Doctoral dissertation, Univ. of Kuopio</i> , (2008)	Metabolic profiling in late stages of strawberry receptacle development	<i>Fragaria x ananassa</i>
K23	Iijima, Y. et al., <i>Plant J.</i> , 54 , 949-962 (2008)	Bioinfo & DB: Metabolite annotation based on MS and MS ²	<i>Solanum lycopersicum</i> cv. Micro-Tom
K24	Malitsky, S. et al., <i>Plant Physiol.</i> , 148 , 2021-2049, (2008)	Regulation of glucosinolate biosynthesis by two clades of regulators	<i>Arabidopsis thaliana</i>
K25	Matsuda, F. et al., <i>Plant J.</i> , 57 , 555-577 (2008)	Bioinfo: identification of metabolites based on MS and MS-tagged MS ² data	<i>Arabidopsis thaliana</i>
K26	Mintz-Oron, S. et al., <i>Plant Physiol.</i> , 147 , 823-851 (2008)	Integrated analysis of metabolome and transcriptome	<i>Solanum lycopersicum</i> Alisa Craig
K27	Oikawa, A. et al., <i>Rice</i> , 1 , 83-71, (2008)	Review: technology and informatics	<i>Oryza sativa</i>
K28	Overy, D.P. et al., <i>Nature Protocols</i> , 3 , 471-485 (2008)	Protocol in metabolite fingerprints	
K29	Takahashi, H., et al., <i>Anal. Bioanal. Chem.</i> , 391 , 2769-2782 (2008)	Bioinfo: Metabolome platform DrDMASS in FT-ICR-MS. Determination of growth-specific metabolites	<i>E. coli</i>
K30	Werner, E., et al., <i>J. Chromatogr. B.</i> , 879 , 143-163, (2008)	Review: Atmospheric pressure ionization mass spectrometry	
K31	Ara, T. et al., <i>Plant Biotechnol.</i> , 26 , 445-449 (2009)	Annotation of metabolite information to MS	<i>Glycine max</i>
K32	Arita, M and Suwa, K., <i>BioData Mining</i> , 1 , 7.1-8, (2009)	DB: Embedded string-search commands on MediaWiki	
K33	Bando, K. et al., <i>J. Appl. Toxicol.</i> , DOI 10.1002/jat.1591 (2010)	Toxicological evaluation	Rat
K34	Davey M. P. et al., <i>Metabolomics</i> , 5 , 138-149, (2009)	Metabolic profiling in cold-temperature	<i>Arabidopsis lyrata</i> ssp. <i>petraea</i>
K35	Draper, J., et al., <i>BMC Bioinformatics</i> , 10 , 277.1-16 (2009)	Bioinfo: Tools for the annotation of High Resolution MS metabolomics data	
K36	Fukushima A. et al., <i>Curr. Opinion in Chem. Biol.</i> , 13 , 532-538 (2009)	Review: Integrated OMICs	
K37	Han, J. et al., <i>Bioanalysis</i> , 1 , 1665-1684 (2009)	Review: MS-based technologies	
K38	Hounsborne, N., et al., <i>Postharvest Biol. Technol.</i> , 52 , 173-179 (2009)	Changes in antioxidant compounds in white cabbage during winter storage	<i>Brassica oleracea</i> var <i>capitata</i>
K39	Kind, T. et al., <i>PLoS One</i> , e5440.1-10 (2009)	Bioinfo: Comparison of metabolite DB using rice metabolites	<i>Oryza sativa</i>

K40	Kai, K., et al., <i>Plant Biotechnol.</i> , 26 , 175-182 (2009)	Hydroxylation of fatty acids by P450 proteins	<i>Arabidopsis thaliana</i>
K41	Manach, C. et al., <i>Mol. Nutr. Food Res.</i> , 53 , 1303-1315 (2009)	Dietary phytochemicals and human	Human
K42	Matsuda, F. et al., <i>PLoS One</i> , 4 , e7490.1-10, (2009)	Bioinfo : Assessment of annotation of metabolites using FDR	
K43	Matsuda, F., et al., <i>Plant Biotechnol.</i> , 26 , 479-483 (2009)	Bioinfo : Graph representation of multiple databases	
K44	Oishi, T., et al., <i>Plant Biotechnol.</i> , 26 , 167-174 (2009)	Bioinfo : Peak detection based on MS/MS patterns	
K45	Okada, T. et al., <i>Planta Med.</i> , 75 , 1356-1352 (2009)		<i>Ephedra</i> sp.
K46	Sawada, Y. et al., <i>Plant Cell Physiol.</i> , 50 , 37-47 (2009)	Selection of metabolites	Brassicaceae, Gramineae, Fabaceae
K47	Shroff, R. et al., <i>Proc. Natl. Acad. Sci. USA</i> , 106 , 10092-10096, (2009)	Matrix-assisted laser desorption/ionization mass spectrometry	
K48	Stracke, R. et al., <i>Planta</i> , 229 , 427-445 (2009)	Determination of gene function	<i>Arabidopsis thaliana</i>
K49	Takemoto, K. and Arita, M., <i>Physica A</i> , 388 , 2771-2780 (2009)	Bioinfo : Complexity of relation between plants and metabolites	
K50	Tanaka K., et al., <i>Plant Biotechnol.</i> , 26 , 459-468 (2009)	Bioinfo : metabolic pathway prediction	
K51	Tanaka, K. et al., <i>J. Trad. Med.</i> , 26 , 179-186, (2009)	Quality assessment	Kampo medicine; hochuekkito
K52	Tianniam, S. et al. <i>J. Sep. Sci.</i> , 32 , 2233-2244, (2009)	Quality assessment	<i>Angelica acutiloba</i>
K53	Tohge, T and Fernie, A., <i>Phytochemistry</i> , 70 , 450-456 (2009)	Review : web-resources in MS-based metabolomics	
K54	Wishart, D. et al., <i>Bioanalysis</i> , 1 , 1579-1596 (2009)	Bioinfo : Metabolite annotation	
K55	Xie, Z., et al., <i>J. Exp. Botany</i> , 60 , 87-97 (2009)	Diarylheptanoid biosynthesis	<i>Curcuma longa</i>
K56	Yonekura-Sakakibara, K. and Saito, K., <i>Nat. Prod. Rep.</i> , 26 , 1466-1487 (2009)	Review : Functional genomics	
'10			
K57	Aliferis K. A., and Jabaji, S., <i>J. Agric. Food Chem.</i> , 58 , 7604-7615 (2010)	Metabolite composition	<i>Rhizoctania solani</i>
K58	Bollina, V. et al., <i>Mol. Plant Pathol.</i> , 11 , 769-782 (2010)	QTL of barley, against Fusarium head blight	<i>Hordeum vulgare</i>
K59	Bar-Akiva, A. et al., <i>J. Exp. Botany</i> , 61 , 393-403 (2010)	Changing color of flower from dark purple to white	<i>Brufelsia calycina</i>
K60	Hattori, M., et al., <i>Nucl. Acids Res.</i> , 38 , W652-656, (2010)	Bioinfo : chemical similarity search and substructure matching of compounds	
K61	Horai, H. et al., <i>J. Mass Spectrometry</i> , 45 , 702-714 (2010)	DB : MassBank, MS DB	
K62	Kind, T, and Fiehn, O., <i>Bioanal. Rev.</i> , 2 , 23-60 (2010)	Review : MS data processing	

K63	Macel, M. et al., <i>Mol. Ecol. Resources</i> , 10 , 583-593 (2010)	Review: metabolomics in plant ecology and genetics	
K64	Matsuda, F., et al., <i>Plant Physiol.</i> , 152 , 566-578, (2010)	Metabolic profiling of different tissues	<i>Arabidopsis thaliana</i>
K65	Neumann, S., and Bocker S., <i>Anal. Biol. Chem.</i> , 398 , 2779-2788 (2010)	Review: identification of metabolites	
K66	Neveu, V., et al., <i>Database</i> , (2010) doi:10.1093/database/bap024	DB: Polyphenol contents in foods	
K67	Ohta D., et al., <i>Curr. Opinion in Biotechnol.</i> , 21 , 35-44 (2010)	Review: FT-ICR-MS Reaction representation based on van Krevelen diagram	
K68	Okada T. et al., <i>Current Comput. - Aided Drug Design</i> , 6 , 179-196 (2010)	Review: Relationship among individual omics data based on multivariate analysis and DB	Medicinal plants
K69	Penn, L. et al., <i>Genes Nutr.</i> , 205-213 (2010)	Review: Dietary intake	
K70	Redestig, H., et al., <i>Data analysis</i> , 11 , 214.1-11 (2010)	Bioinfo: Multiple metabolomics platforms for different types of MS	
K71	Saito, K. and Matsuda, F., , <i>Annu. Rev. Plant Biol.</i> , 61 , 463-89, (2010)	Review: Functional Genomics	
K72	Singla, D. et al., <i>BMC Pharmacol.</i> , 10 , 4.1-8 (2010)	DB: binzylsioquinone alkaloids	
K73	Tanaka, K. et al., <i>J. Trad. Med.</i> , 27 , 210-216 (2010)	Quality assessment	<i>Glycyrrhiza uralensis</i>
K74	Tohge, T. and Fernie, A. R., <i>Nature Protocols</i> , 5 , 1210-1227 (2010)	Review: Annotation of gene function based on co-response gene and identification of metabolites	
K86	Takemoto, K., <i>BioSystems</i> , 100 , 8-13, (2010)	Bioinfo: Network analysis of species-metabolite relations	
K75	Weber, R.J.M. et al., <i>Chemometrics Intel. Lab. Sys.</i> , 104 , 75-82 (2010)	Bioinfo: MS data processing	
'11			
K76	Acharjee, A., et al., <i>Anal. Chim. Acta</i> , in press (2011)	Bioinfo: QTL informatics	<i>Solanum tuberosum</i>
K77	Krueger, S., et al., <i>PLoS One</i> , 6 , e17806.1-16 (2011)	Subcellular distribution of metabolites	<i>Arabidopsis thaliana</i>
K78	Aliferis, K. A. et al., <i>Metabolomics</i> , 7 , 35-53 (2011)	Review: Pesticide research	
K79	Kouskoumvekaki, I. and Panagiotou, G., <i>J. Biomedicine and Biotechnol.</i> , doi:10.1155/2011/525497 (2011)	Bioinfo: Metabolomics in medical purpose with systems chemical biology and chemoinformatics	Human
K80	Kumaraswamy, G. K. et al. <i>Eur. J. Plant Pathol.</i> , 130 , 29-43 (2011)	Assessment of metabolites of barley against Fusarium head blight	<i>Hordeum vulgare</i>
K81	Kusano, M. et al., <i>Plant J.</i> accepted (2011)	Metabolic responses of ultraviolet-B light	<i>Arabidopsis thaliana</i>
K82	Ohkama-Ohtsu, N. et al., <i>Plant Cell Physiol.</i> , 52 , 205-209 (2011)	Transport of 12-Oxo-phytodienoic acid-glutathione into vacuole	<i>Arabidopsis thaliana</i>
K83	Osorio S., et al., <i>J. Exp. Botany</i> , doi:10.1093/jxb-erq465 (2011)	Demethylation of oligogalacturonides by FAFE1 leads to defense against fungus <i>Botytis cinerea</i> .	<i>Fragaria vesca</i>

K84	Scalbert, A. et al., <i>J. Agric. Food Chem.</i> , 59 , 4331-4348 (2011)	DB: Food phytochemicals	
K85	Nisioka, T. , pp.9-18, バイオインフォマティクス推進事業第7回研究開発成果報告会(2011)	DB: MassBank and MS analysis	
K87	Bando, K., <i>J. Appl. Toxicol.</i> , 31 , 524-535 (2011)	Hepatotoxicity	Rats
K88	Giavalisco, P. et al., <i>Plant J.</i> , 68 , 364-376, (2011)	Bioinfo: Molecular formula annotation of polar and lipophilic metabolites	
K89	Fiehn,O., et al., <i>J. Biol. Chem.</i> , 286 , 23637-23643 (2011)	Review: metabolome DB	
K90	Kai, K., et al., <i>Plant Biotechnol.</i> , 28 , 379-385 (2011)	Cytochrome P450, CYP81F4	<i>Arabidopsis thaliana</i>
K91	Yanuar, A., et al., <i>Int. J. Computer Sci. Issue</i> , 8 , 180-183, (2011)	DB: Medicinal plants in Indonesia	
K92	Katoh, A., et al., <i>Am. J. Chinese Med.</i> , 39 , 757-777 (2011)	Review: Traditional medicinal plants	<i>Angelica sp.</i>
K93	Kaneko, Y., <i>Exp. Mol. Pathol.</i> , 91 , 614-621 (2011)	Imaging mass spectrometry	Murine
K94	Aliferis, K.A., <i>Metabolomics</i> , 7 , 35-53 (2011)	Review: pesticide research	
K95	Takahashi, H., <i>BMC Bioinformatics</i> , 12 , 259.1-8 (2011)	Bioinfo: Metabolic profiling	<i>Bacillus subtilis</i>
	'12		
K96	Sartor, M. et al., <i>Bioinformatics</i> , 28 , 1408-1410, (2012)	DB: metabolite annotation	
K97	Asano, T., <i>Phytochemistry</i> , (in press), (2012)	Camptothecin biosynthesis	<i>Ophiorrhiza pumila</i>
K98	Ahuja, I et al., <i>Trends Sci.</i> , 17 , 73-90, (2012)	Review: Phytoalexins	
K99	Okazaki, Y. And Saito, K., <i>Plant Biotechnol. Rep.</i> , 6 , 1015 (2012)	Review: Plant biotechnology	
K100	Marti, G., <i>Plant Cell Env.</i> ,(in press) (2012)	Herbivore (Spodoptera littoralis)-induced metabolites	<i>Zea mays</i>
K101	Lieberman, L.M. et al., <i>Cur. Op. Plant Biol.</i> , 15 , 162-167 (2012)	Review : Integrative system biology	<i>Arabidopsis thaliana</i>
K102	Obata, T. And Fernie, A.R., <i>Cell. Mol. Life. Sci.</i> , 69 , 3225-3243 (2012)	Review : plant responses to abiotic stresses	
K103	Afendi, F. M. Et al., <i>Curr Pharmacogenomics Personalized Med.</i> , 10 , 111-124 (2012)	Review : Systems biology in Japanese traditional Kampo medicine	
K104	Houshyani, B., et al., <i>Metabolomics</i> , 8 , S131-S145, (2012)	Natural distance	<i>Arabidopsis thaliana</i>
K105	Wahyuni,Y., <i>Metabolomics</i> , doi 10.1007/s11306-0432-6 (2012)	Molecular marker	<i>Capsicum sp.</i>
K106	Sano, R. Et al., <i>Plant Biotechnol.</i> , 29 , 175-178 (2012)	Metabolic changes during fruit maturation	<i>Jatropha curcas</i>
K107	Ohtani, M., et al., <i>Plant Biotechnol.</i> , 29 , 171-174 (2012)	Metabolites in seed kernels	<i>Jatropha curcas</i>
K108	Khan, S.A., et al., <i>J. Exp. Botany</i> , 63 , 2895-2908 (2012)	mQLT	<i>Malx x domestica</i>

K109	Alla, M.M.M. et al., <i>Plant Growth Regul.</i> , 67 , 281-304 (2012)	Salt and drought stress	<i>Atriplex halimus</i>
K110	Haug, K., <i>Nucleic Acid Res.</i> , doi:101093/nar/gks1004	Bioinfo: repository for metabolomics studies	
K111	Wagele, B., <i>PlosOne</i> , 7, e39860, (2012)	Bioinfo: Visualization of metabolome data	
K113	Peukert, <i>New Phytologist</i> , 193, 806-815 (2012)	mass spectrometric imaging	
‘13			
K112	Ballester, A.R. et al., <i>Food Chemistry</i> , 136, 178-185 (2013)	Defence against pathogens (<i>Penicillium digitatum</i>)	<i>Citrus sinensis</i>

出典：金谷ら、メタボロームデータベース：多様な研究への対応とデータの共有化、実験医学 29(15)126-36(2011)

資料編

1. 論文リスト

1. #*Horai, H., Arita, M., Kanaya, S., Nihei, Y., Ikeda, T., Suwa, K., Ojima, Y., Tanaka, K., Tanaka, S., Aoshima, K., Oda, Y., Kakazu, Y., Kusano, M., Tohge, T., Matsuda, F., Sawada, Y., Yokota-Hirai, M., Nakanishi, H., Ikeda, K., Akimoto, N., Maoka, T., Takahashi, H., Ara, T., Sakurai, N., Suzuki, H., Shibata, D., Neumann, S., Iida, T., Tanaka, K., Funatsu, K., Matsuura, F., Soga, T., Taguchi, R., Saito, K. and Nishioka, T. "MassBank: A public repository for sharing mass spectral data for life sciences", *J. Mass Spectrometry*, 45(7), 703-714 (2010). (査読あり)
2. Redestig H, Kusano M, Fukushima A, Matsuda F, Saito K, Arita M. "Consolidating metabolite identifiers to enable contextual and multi-platform metabolomics data analysis" *BMC Bioinformatics* 11:214, 2010. (査読あり)
3. Okada, T., Afendi, F., Altaf-Ul-Amin, M., Takahashi, T., Nakamura, K., Kanaya, S., "Metabolomics of Medicinal Plants: The Importance of Multivariate Analysis of Analytical Chemistry Data," *Current Computer-Aided Drug Design*, 6, 179-96 (2010). (査読あり)
4. Takahashi, H., Morioka, R., Ito, R., Oshima, T., Altaf-Ul-Amin, M., Ogasawara, N., Kanaya, S., "Dynamics of time-lagged gene-to-metabolite networks of *Escherichia coli* elucidated by integrative omics approach," *OMICS JIB*, 14, (in press, doi: 10.1089/omi.2010.0074) (2010). (査読あり)
5. Strassburg, K., Walther, D., Takahashi, H., Kanaya, S., and Kopka, J., "Dynamic transcriptional and metabolic responses in yeast adapting to temperature stress.," *OMICS JIB*, 14, 249- 59 (2010). (査読あり)
6. 蓬萊尚幸、二瓶義人、尾畷雄也、池田 奨、西岡孝明、“MassBank - メタボロームマ スペクトルデータベース- ”、「メタボロミクス：その解析技術と臨床・創薬応用 研究の最前線」(遺伝子医学 MOOK 16：田口 良編集)、メディカルドゥ、86-90 (2010). (査読あり)
7. #*Oishi, T, Tanaka, K., Hashimoto, T., Shinbo, Y., Jumtee, K., Bamba, T., Fukusaki, E., Suzuki, H., Shibata, D., Takahashi, H., Asahi, H., Kurokawa, K., Nakamura, Y., Hirai, A., Nakamura, K., Md. Altaf-Ul-Amin, Kanaya S., An approach to peak detection in GC-MS chromatograms and application of KNApSAcK database in prediction of candidate metaboltes, *Plant Biotechnol.*, 26, 167-168 (2009). (査読あり)
8. #*Tanaka, K., Nakamura, K., Saito, T., Osada, H., Hirai, A., Takahashi, H., Kanaya, S., Md. Altaf-Ul-Amin, Metabolic pathway prediction based on inclusive relation

- between cyclic substructures, *Plant Biotechnol.*, 26, 459-468 (2009) . (査読あり)
9. Arita M, Suwa K, Yoshimoto M, Hirai A, Kanaya S "Ontology checking and integration of crude-drug and kampo information" *Proceedings of the 2nd International Conference of Biomedical Engineering and Informatics (BMEI2009)*, Tianjin China, 3:1304-1307, 2009. (査読あり)
 10. Horai H, Arita M, Ojima Y, Nihei Y, Kanaya S, Nishioka T "Traceable analysis of multiple-stage mass spectra through precursor-product annotations" *Proceedings of German Conference in Bioinformatics (GCB'09) (GI-Edition Lecture Notes in Informatics)*, Halle, Germany, 173-178, 2009. (査読あり)
 11. Arita M "What can metabolomics learn from genomics and proteomics?" *Current Opinion in Biotechnology* 20, 610-615, 2009. (総説。エディターによる査読あり)
 12. #*有田 正規, 蓬萊 尚幸, 尾畷 雄也, 二瓶 義人, 金谷 重彦, 西岡 孝明 「化合物情報とマススペクトルを活用するためのウェブ基盤」 *CICSJ Bulletin*, 27(2), 48-50, 日本化学会 情報化学部会, 2009. (査読あり)
 13. #*Takahashi, H., Kai, Kosuke, Shinbo, Y., Tanaka, K., Ohta, D., Oshima, T., Md. Altaf-Ul-Amin, Kurokwa, K., Ogasawara, N., Kanaya, S., Metabolomics approach for determining growth-specific metabolites based on Fourier transform ion cyclotron resonance mass spectrometry, *Anal. Biolanal Chem.*, 391, 2769-2782 (2008) (査読あり)
 14. Horai, H. and Nishioka, T. , Automatic Generation of Structure of Phospholipids. *Journal of Computer Aided Chemistry*, 9, 55-61 (2008) (査読あり)
 15. 有田正規, 諏訪和大 「Wiki によるフラボノイドのデータベース」 *実験医学増刊*, 「バイオデータベースとソフトウェア最前線」, 26(7), 1148-1154 (2008).
 16. Suzuki, H., Sasaki, R., Ogata, Y., Nakamura, Y., Sakurai, N., Kitajima, M., Takayama, H., Kanaya, S., Aoiki, K., Shibata, D., Saito, K., Metabolic profiling of flavonoids in *Lotus japonicus* using liquid chromatography Fourier transform ion cyclotron resonance mass spectrometry, *Phytochemistry*, 69, 99-111 (2008) (査読有)
 17. 蓬萊尚幸、西岡孝明、“メタボローム解析のための質量分析スペクトル・データベース：MassBank” *実験医学増刊*、26(7), 1155-1160 (2008).
 18. Ohta,D., Shibata, D.,Kanaya, S., Metabolic profiling using Fourier-transform ion-cyclotron-resonance mass spectrometry, *Anal. Biolanal. Chem.*, 389, 1469-1475 (2007) (査読あり)
 19. Nakamura,Y., Kimura, A., Saga, H., Oikawa, A., Shinbo, Y., Kai,K., Sakurai, N., Suzuki,H., Kitayama, M., Shibata, D., Kanaya, S., Ohta,D., Differential metabolomics unraveling light/dark regulation of metabolic activities in *Arabidopsis* cell culture, *Planta*, 227, 57-66 (2007) (査読あり)

20. Shinbo, Y., Sakaguchi, S., Nakamura, Y., Md. Altaf-Ul-Amin, Kurokawa, K., Funatsu, K. and Kanaya, S., Species-metabolite Database (KNApSAcK): Elucidating Diversity of Flavonoids, J.Computer Aided Chemistry, 7, 94-101 (2006)
(査読あり)
21. 蓬萊尚幸, 西岡孝明. メタボローム解析のための MS データベース構築. 細胞工学 25(12), 1394-1398 (2006)
22. 西岡孝明, メタボローム・マススペクトル統合データベース, 日本化学会情報化学部 会誌, 24, 38-42 (2006).

: 今回の追跡調査において研究代表者が主要な論文として指定したもの (上限 30 報)

* : 研究開発期間終了後の終了報告書において研究代表者が主要な論文として指定したもの

2. 主要論文の被引用回数

論文	国内外別件数		分野別件数																				出版年別件数								
	国内件数	海外件数	農学	生物学・生化学	化学	臨床医学	コンピュータサイエンス	経済学・経営学	工学	環境・生態学	地球科学	免疫学	材料科学	数学	微生物学	分子生物学・遺伝学	複合領域	神経科学・行動科学	薬理学・毒物学	物理学	植物学・動物学	精神医学・心理学	社会科学・一般	宇宙科学	2006年	2007年	2008年	2009年	2010年	2011年	2012年
2. *Horai, H., Arita, M., Kanaya, S., Nihei, Y., Ikeda, T., Suwa, K., Ojima, Y., Tanaka, K., Tanaka, S., Aoshima, K., Oda, Y., Kakazu, Y., Kusano, M., Tohge, T., Matsuda, F., Sawada, Y., Yokota-Hirai, M., Nakanishi, H., Ikeda, K., Akimoto, N., Maoka, T., Takahashi, H., Ara, T., Sakurai, N., Suzuki, H., Shibata, D., Neumann, S., Iida, T., Tanaka, K., Funatsu, K., Matsuura, F., Soga, T., Taguchi, R., Saito, K. and Nishioka, T. "MassBank: A public repository for sharing mass spectral data for life sciences", J. Mass Spectrometry, 45(7), 703-714 (2010). (査読有り)	13	59	5	15	21	1	7	0	4	0	0	0	0	0	2	6	0	0	6	0	5	0	0	0	-	-	-	-	2	25	45
8. *Oishi, T., Tanaka, K., Hashimoto, T., Shinbo, Y., Jumtee, K., Bamba, T., Fukusaki, E., Suzuki, H., Shibata, D., Takahashi, H., Asahi, H., Kurokawa, K., Nakamura, Y., Hirai, A., Nakamura, K., Md. Altaf-Ul-Amin, Kanaya S., An approach to peak detection in GC-MS chromatograms and application of KNApSACk database in prediction of candidate metabolites, Plant Biotechnol., 26, 167-168 (2009). (査読あり)		1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	-	-	-	1	1	0	0
9. *Tanaka, K., Nakamura, K., Saito, T., Osada, H., Hirai, A., Takahashi, H., Kanaya, S., Md. Altaf-Ul-Amin, Metabolic pathway prediction based on inclusive relation between cyclic substructures, Plant Biotechnol., 26, 459-468 (2009). (査読あり)		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-	-	-	0	0	0	0
14. *Takahashi, H., Kai, Kosuke, Shinbo, Y., Tanaka, K., Ohta, D., Oshima, T., Md. Altaf-Ul-Amin, Kurokawa, K., Ogasawara, N., Kanaya, S., Metabolomics approach for determining growth-specific metabolites based on Fourier transform ion cyclotron resonance mass spectrometry, Anal. Bioanal. Chem., 391, 2769-2782 (2008) (査読あり)	12	16	1	12	6	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	6	0	0	0	-	-	0	10	10	5	3

- ・本追跡調査において研究代表者が主要な論文として指定したもの（上限 30 報）について、トムソン・ロイター社 Web of Science で調査した。調査対象は、2012 年 12 月まで。被引用情報が取得できたもののみ記載した。
- ・論文番号は「資料編 1. 論文リスト」に対応している。
- ・*：研究開発期間終了後の終了報告書において研究代表者が主要な論文として指定したもの。
- ・[国内外別件数] は、被引用文献の国内外別件数。被引用文献の著者の所属国のうち JAPAN が一つでもあれば国内としている。国内外の合計が全被引用文献数となる。
- ・[分野別件数] は、被引用文献の分野別件数。被引用文献が論文の場合のみカウントしている。
- ・[出版年別件数] は、被引用文献の出版年別件数。「-」は、データなしを表す。

論文	Web of Science 国内外別件数(2012年12月まで)			Google Scholar	
	国内件数	海外件数	合計	~2014年4月24日	~2012年
2. *Horai, H., Arita, M., Kanaya, S., Nihei, Y., Ikeda, T., Suwa, K., Ojima, Y., Tanaka, K., Tanaka, S., Aoshima, K., Oda, Y., Kakazu, Y., Kusano, M., Tohge, T., Matsuda, F., Sawada, Y., Yokota-Hirai, M., Nakanishi, H., Ikeda, K., Akimoto, N., Maoka, T., Takahashi, H., Ara, T., Sakurai, N., Suzuki, H., Shibata, D., Neumann, S., Iida, T., Tanaka, K., Funatsu, K., Matsuura, F., Soga, T., Taguchi, R., Saito, K. and Nishioka, T. "MassBank: A public repository for sharing mass spectral data for life sciences", J. Mass Spectrometry, 45(7), 703-714 (2010). (査読有り)	13	59	72	237	111
8. *Oishi, T., Tanaka, K., Hashimoto, T., Shinbo, Y., Jumtee, K., Bamba, T., Fukusaki, E., Suzuki, H., Shibata, D., Takahashi, H., Asahi, H., Kurokawa, K., Nakamura, Y., Hirai, A., Nakamura, K., Md. Altaf-Ul-Amin, Kanaya S., An approach to peak detection in GC-MS chromatograms and application of KNApSACk database in prediction of candidate metabolites, Plant Biotechnol., 26, 167-168 (2009). (査読あり)	(1) 1	1	2	5	3
9. *Tanaka, K., Nakamura, K., Saito, T., Osada, H., Hirai, A., Takahashi, H., Kanaya, S., Md. Altaf-Ul-Amin, Metabolic pathway prediction based on inclusive relation between cyclic substructures, Plant Biotechnol., 26, 459-468 (2009). (査読あり)	(2) 0	0	0	4	0

14. *Takahashi, H., Kai, Kosuke, Shinbo, Y., Tanaka, K., Ohta, D., Oshima, T., Md. Altaf-Ul-Amin, Kurokwa, K., Ogasawara, N., Kanaya, S., Metabolomics approach for determining growth-specific metabolites based on Fourier transform ion cyclotron resonance mass spectrometry, Anal. Bioanal Chem., 391, 2769-2782 (2008) (査読あり)	12	16	28	45	36
---	----	----	----	----	----

・情報科学・計算機科学分野における研究開発成果では、会議予稿集での発表に引用されることが多いため、Google Scholar での被引用件数を参考情報として添付した。

・Google での調査方法

対象論文のタイトルを検索し、当該論文を「引用元」とする件数を取得した。

最新の被引用件数と、それから 2013 年以降の被引用件数を引いたものを 2012 年までとした。ただし、年が明確でないものは引かれていないため、2012 年までの被引用件数が実際よりも多くなっている可能性がある。・論文番号は「資料編 1. 論文リスト」に対応している。

・* : 研究開発期間終了後の終了報告書において研究代表者が主要な論文として指定したもの。

3. 学会招待講演・基調講演

1. Nishioka, T., Horai, H., Masanori, A. and Kanaya, S., "MassBank: A Public Repository for Sharing Mass Spectra in Life Sciences", 1st Asian and Oceanic Mass Spectrometry Conference, Tsukuba, Tsukuba International Congress Center, Tsukuba, June 16-18, 2010.
2. 西岡孝明、"メタボロミクスのためのマススペクトルデータベース MassBank"、第 311 回 CBI 学会研究講演会、産業総合科学技術研究所、東京都、2010 年 7 月 28 日。
3. 西岡孝明、"ここまで進化した MassBank"、第 35 回日本医用マススペクトル学会、金城学院大学、名古屋市、2010 年 9 月 9-10 日。
4. Nishioka, T., "MassBank: A Reference Database for Chemical Identification of Metabolites Detected by Mass Spectrometry", Biocuration 2010 (The Conference of the International Society of Biocuration), Tokyo International Exchange Center, Tokyo, October 11-14, 2010. 2009 年度
5. Horai, H., Arita, M., Nihei, Y., Ikeda, Y., Ojima Y., and Nishioka, T., "MassBank: A Mass Spectral Database for Metabolomics", Invited lecture, JSBi 20th Anniversary Special Session, The 20th International Conference on Genome Informatics, Pacifico Yokohama, Japan, December 14-16, 2009.
6. 西岡孝明、蓬萊尚幸、"LipidBank とマススペクトルデータベース MassBank"、(ミニシンポジウム；招待講演) 第 51 回日本脂質生化学会（日本脂質生化学会主催）、ウィルあいち、名古屋市、2009 年 7 月 30-31 日。2007 年度
7. 金谷重彦 "生物種一代謝物関係データベース：KNAPSKcK" 第 25 回バイオテクノロジーシンポジウム、東京都、2007 年 11 月。2006 年度
8. 西岡孝明 "メタボローム解析—CE-MS の応用と展望—メタボローム・マススペクトル DB—" 第 31 回日本医用マススペクトル学会年会シンポジウム、招待講演、名古屋国際会議場、名古屋市、2006 年 9 月 28-29 日。

4. 新聞発表等（著作権の関係により非公開）

5. 特許出願・成立

該当なし

6. 学会賞等の受賞

1. 第 30 回情報化学討論会ポスター賞
田中 健一, 真保 陽子, Md. Altaf-Ul-Amin, 旭 弘子, 黒川 顕, 平井 晶, 有田 正規, 太田 大策, 金谷 重彦「構造情報に基づいたメタボライト分類支援システムの開発」
JP32-2007
2. Journal of Computer-aided Chemistry 誌の最優秀論文賞
Shinbo, Y., Sakaguchi, S., Nakamura, Y., Md. Altaf-Ul-Amin, Kurokawa, K., Funatsu, K. and Kanaya, S., “Species-metabolite Database (KNApSAcK): Elucidating Diversity of Flavonoids”, J.Computer Aided Chemistry, 7, 94-101, (2006) (査読あり)

7. グラントの獲得実績

1. 「統合化推進プログラム」研究開発課題名「メタボローム・データベースの開発」(金谷重彦 2011-2013 年) 分担者
2. メタボローム解析のための計測技術開発とそれを用いた代謝経路推定 (富田勝 2006-2009 年 特定領域研究)
3. 枯草菌代謝制御ネットワークの物質産生系への有効活用(田泰太郎 2006-2007 年 特定領域研究)
4. 植物の 2 次代謝誘導に対する 1 次代謝物質のマネジメントに関するグローバル解析(西岡 孝明 2006-2007 年 基盤研究(B))

8. 書籍等の執筆実績

該当なし

9. 総説の執筆実績

1. 西岡孝明, “データベース MassBank”, 「現代質量分析学」第 32 章 (高山光男, 早川滋雄, 瀧浪欣彦, 和田芳直編), 化学同人, 京都, 2013.
2. 中村由紀子, 森田 (平井) 晶, 西岡孝明, 金谷重彦, “KNApSAcK Family データベ

ス：メタボロミクスから展開する植物の多目的活用”、バイオサイエンスとインダストリー、70 (4), 267-272 (2012).

3. 金谷重彦, 平井晶, 旭弘子, 高橋弘喜, 中村建介, Md. Altaf-Ul-Amin,、二瓶義人, 池田奨, 尾嶋雄也, 西岡孝明、”メタボロームデータベース：多様な研究への対応とデータの共有化”、「使えるデータベース・ウェブツール」実験医学、29,(15), 2460 (2011).

10. 参加研究者の活動状況（個人情報が含まれるため非公開）