食品成分のメタボローム解析と未知成分同定のための “食品メタボロームレポジトリ”の提案

Food Metabolome Repository: a New Database for Identification of Unknown Compounds in Food Metabolome Analyses

櫻井 望

大学共同利用機関法人 情報・システム研究機構
国立遺伝学研究所生命情報・DDBJ センター

〒411-8540
静岡県三島市谷田 1111

Nozomu SAKURAI
Bioinformation & DDBJ Center, National
Institute of Genetics, Research Organization of
Information and Systems
1111 Yata, Mishima, Shizuoka 411-8540, Japan

論文要旨：質量分析計によるメタボローム解析技術を用いて、食品をはじめとする身の周りの試料を分析すると、存在は確認できるがすぐには正体が判明しない未知の成分が多数含まれていることが分かる。これらの未知成分の正体を同定・推定することは、新規の有用成分やバイオマーカーの発見、生命現象の理解などに役立つと期待される。本稿では、質量分析計によるメタボローム解析の概要と課題を紹介し、成分推定の新しいアプローチとして、成分の試料特異性（検出実績）が分かるデータベース「食品メタボロームレポジトリ」を紹介する。

Abstract: A significant number of “unknown” metabolite peaks are usually detected in foods and other samples by mass spectrometry (MS)-based untargeted metabolome analyses. Identification and annotation of these unknowns are useful for discovery of novel functional molecules and biomarkers, and for understanding the mechanisms of living organisms. In this article, we summarize current status and issues in the MS-based metabolomics and propose a database, Food Metabolome Repository (<http://metabolites.in/foods>), as a new data resource for improving metabolite annotations based on the sample specificity of the unknowns.

Key words: food metabolome repository, liquid chromatography-mass spectrometry, compound database, metabolite annotation

1 はじめに

よく、魚を食べると良質な脂肪を摂取できると言われるが、メタボロームデータを見ると、本当に脂質ばかりを食べていることが分かる。Fig. 1a はかつおの切り身に含まれる成分を液体クロマトグラフィー質量分析 (LC-MS) で網羅的に検出したメタボローム解析の結果であり、画面右側に見えている主要なスポットはグリセロリン脂質など疎水性の高い成分と推定される。この様相は、薬味である生姜では全く異なっている (Fig.

1b)。生姜では、画面中央から右側にかけて1万個以上のスポットが検出されており、疎水性が中程度から高いものまで、実に多様な成分が含まれていることが分かる。日本食品標準成分表2015年版(七訂)では、炭水化物、たんぱく質、脂質のほか、脂肪酸、アミノ酸、有機酸、ミネラル、ビタミン類など、約140の成分に関する定量値が記載されている。しかし、Fig. 1で見たように、成分表からは読み取れない多数の成分が食品ごとに様々に含まれていることは、意外と認識されていない。それでは、これらは一体どんな物質なのだろうか？ 実は、すべてを直ちに言い当てることはほぼ不可能であるが、これに挑戦できる技術がメタボロミクスである。

連絡者：櫻井 望
E-mail : sakurai@nig.ac.jp

Fig. 1 各食品の成分プロファイル例。a かつお。b しょうが。c かつおだし。各試料のメタノール抽出物を C18 カラムと水-アセトニトリル系による逆相液体クロマトグラフィーで分離後、ESI ポジティブモードで質量分析を行った結果を二次元のマスプロファイルとして表示した。スポットは成分の検出位置と強度を示す。食品メタボロームレポジトリ (<http://metabolites.in/foods>) の MassChroBook より転載。

本稿では、LC-MS を中心とした質量分析によるメタボローム解析の技術背景や課題を解説し、上記のような「存在は確認できるが同定ができない成分」を明らかにするための新規な情報リソースとして、筆者らが近年構築した食品メタボロームレポジトリ (<http://metabolites.in/foods>, 「食レポ」と略す)¹⁾ を紹介する。Fig. 1 で示したような LC-MS の二次元マスプロファイルは、現在収録されている 222 の食品について、食レポのトップページからパワーポイントファイル (MassChroBook) としてダウンロードできる。結果は示さないが、野菜は種類ごとに個性豊かな多数の成分を含む一方、魚は、イワシからマグロまでほとんど大差なく成分が少ない。ちなみに、かつおを原料に作られるかつお節から抽出したかつおだしでは、不思議なことにたくさんの成分が検出されており、製造や抽出などの過程で大きな成分変化を受けていることが分かる (Fig. 1c)。なお、この分析条件では、ほとんどのタンパク質や、疎水性の強い単純脂質、イオン化されにくいカロテノイドやイソプレノイドなどは測定範囲外で、検出される成分に偏りはある。しかしそれでも、成分の多様性をこのように可視化すると、目で比較するだけでも様々な示唆が得られるので、ご興味があればご参照いただきたい。

2 メタボローム解析と成分同定のアプローチ

メタボロミクスは、試料に含まれる低分子化合物を網羅的に定性・定量するための技術であり、対象となる低分子化合物の集合をメタボロームと呼ぶ²⁾。化学成分を検出対象とするため、生物学や医学分野だけでなく、河川、土壌などの環境分野、農業生産分野、加工調理、発酵、流通、品質管理を含めた食品分野などに幅広く期待されている。いわば、地球上の物質循環の全体像を明らかにしその情報を有効活用し得る技術である。しかし、化合物の多様性が大きいために単一の分析手法では網羅できないことや、化合物の同定に精製標品が必要なために成分の同定・推定が難しいことが課題となっている。このため解析の方針も、対象を同定しやすいいくつかの既知成分に限定するもの (ターゲット解析, targeted analysis) と、限定しないもの (ノンターゲット解析, untargeted analysis) が、用途に合わせて使い分けられている。以下、順を追って解説していこう。

2・1 成分の幅広い多様性

メタボロミクスが対象とすべき化学成分は、種類が多く、性質や存在量も多様である。Table 1 に、代表的な化合物データベースでの登録数を示した。天然物の報告を集積した Dictionary of Natural Products には 30 万件

Table 1 主な化合物データベースと登録数。エントリー数は2018年12月現在のもの。

データベース	エントリー数	説明	URL
KEGG	18,456	京都大学で開発された生物代謝マップのデータベース	www.genome.jp/kegg
KNAPSAcK	51,179	奈良先端大学の金谷教授が開発した化合物-文献情報のデータベース	kanaya.naist.jp/KNAPSAcK
LIPID MAPS	43,307	米国 NIH の支援で開発され現在はウェルカム・トラストが維持する、脂質のデータベース	www.lipidmaps.org
HMDB	114,100	カナダのアルバータ大学で開発された、ヒトに存在する / 存在が予想される化合物のデータベース	www.hmdb.ca
Dictionary of Natural Products	278,246	英国 Taylor&Francis グループの CRC 出版が編集する天然物質のデータベース	dnp.chemnetbase.com
PubChem	96,890,507	米国 NIH が開発する化合物のデータベース	pubchem.ncbi.nlm.nih.gov
Chemical Abstract Service	>144,000,000	米国化学会の情報部門による化合物情報を集積したデータベース	www.cas-japan.jp

近い成分が掲載されている。また、植物界全体では、100万種類以上の生合成能力があると推定されている³⁾。脂質は、ステロイド、カロテノイド、フラボノイドなども含め、43,000件以上がLIPID MAPSに登録されている⁴⁾。人工的に合成された化合物まで含めると、PubChem⁵⁾には9千万以上、Chemical Abstract Service (CAS)には1億4千万を超えるエントリーがある。これらには、一部の元素が同位体で置換された化合物や、塩や水和物、チャージの状態が異なるものが独立に登録されていることがあるが、その重複を除去しても、PubChemには7,600万件もある⁶⁾。化学成分は、数が多いだけでなく、Fig. 1で見た疎水性の違い、揮発性、電荷のもちやすさなど、物理化学的性質も様々である。また、試料中の存在量も幅広い。ごく微量で機能するホルモンやシグナル分子がある一方で、オリーブオイルのトリグリセリドのように大量に蓄積するものもある。このため、目的に応じて適切な分画や抽出、濃縮、場合によっては誘導体化を行う必要がある。

このようにメタボロミクスでは、単一の分析手法で全成分を網羅することは難しく、網羅性を広げるためには複数の分析手法を組み合わせる必要がある。

2・2 使用される分析機器

メタボロミクスの主要な分析機器として、クロマトグラフィー-質量分析と核磁気共鳴が相補的に用いられている。すべての物質が共有する唯一の特徴量である「質量」を測定するのが質量分析装置 (mass spectrometry, MS) で、その前段に、性質によって成分を分離するクロマトグラフィーを接続することにより、試料中の複数の成分を個別に検出できる。クロマトグラフィーとしては、液体クロマトグラフィー (liquid chromatography, LC)、ガスクロマトグラフィー (gas chromatography, GC)、キャピラリー電気泳動 (capillary electrophoresis, CE) などが用いられる。MSは検出感度が高く、クロマトグラフィーと接続してハイスループットな多成分の分離・検出が行えるが、化学構造を直接知ることではでき

ない。一方、核磁気共鳴 (nuclear magnetic resonance, NMR) は、精製した未知物質の化学構造を決定でき、また、特定の物質について検出系ができ上がれば、抽出やクロマトグラフィーを必要とせず多検体を短時間で測定できる場合がある。このため、クロマトグラフィー-MSとNMRは相補的に用いられている。NMRについては本特集の他稿で詳しく紹介されるので、そちらをご参照いただきたい。

ノンターゲット解析ではLC-MSが多用されるが、その理由としては、分子の質量に加えて「マスペクトル」を得られることが挙げられる。マスペクトルは、MS装置内で分子を断片化させ、生成した複数の断片の質量電荷比 (m/z) を測定したものであり、元の分子の化学構造の特徴を反映しているため、成分を見分ける有力な手がかりとなる。MSは、成分を荷電した粒子 (イオン) として検出するため、MSの導入部分にはイオン化装置がある。GC-MSでよく用いられる電子イオン化 (EI) では、与えた電気エネルギーによりイオン化と同時に断片化が起こる。LC-MSやCE-MSで主に用いられるエレクトロスプレーイオン化 (ESI) では、元のイオンを断片化させずに測定できるが、さらにMS装置によっては、壊れやすい化学結合を不活性ガスへの衝突などで解離させ (collision induced dissociation, CID)、EIとは異なったマスペクトルを得ることができる。CE-MSはキャピラリーに注入できる試料の量が少ないため、十分な感度でCIDマスペクトルを得ることが難しい。これらのことから、ノンターゲット解析ではLC-MSがよく用いられている。

ESIでは、分子内に電荷を持たない中性分子などは、溶媒等に含まれる水素イオンやアンモニウムイオン、ナトリウムイオンなどと結合し、これらの付加体 (アダクトイオン) として検出される。アダクトイオン種の正しい判別は、物質同定に極めて重要である。

2・3 成分同定の難しさ

クロマトグラフィー-MSで検出されたピーク成分の

同定は、想定される精製標品と比較することで行う。試料と精製標品を、同じ装置、同じ条件で測定したときに、クロマトグラフィーの保持時間、質量電荷比 (m/z)、EI や CID によるマススペクトルなど、計測される特徴がすべて一致した場合に、同定したとみなす。これは、クロマトグラフィーの保持時間や CID スペクトルなどが、装置の組み合わせや設定条件で変化しやすいためである。なお、クロマトグラフィーで分離できない光学異性体などは、区別して同定できないことがある。

あらかじめ検出対象を決めたターゲット分析では、精製標品を入手できる化合物を対象にすることが多いので、標品に依存したこの同定プロセスは問題ない。一方、検出対象を限定しないノンターゲット分析では、これが大きな問題となっている。検出された成分ピークについて、 m/z 値や判定されたアダクトイオン種をもとに化合物データベースを検索したり、CID スペクトルをスペクトルライブラリに検索したりして、候補の化合物を見つけようとしても、該当する候補が見つからなかったり、逆に多数の候補が得られる結果となりやすい。運よく、ひとつあるいは少数の候補に絞込むことができたとしても、市販されている精製標品はせいぜい 1000 種類程度であり、Table 1 に示した成分数と比較すると圧倒的に少なく、標品をすぐに入手できないことが多い。この場合は、有機化学合成をして標品を得るか、試料から精製して NMR で構造を決めるなど、大きなコストと時間が必要となる。

このような同定の難しさから、ここ数十年の間は、精製標品が入手しやすいアミノ酸や糖、有機酸類など、一次代謝を中心としたターゲット分析が盛んに行われてきた。定量性や再現性、高感度検出を追及する技術が進歩し、がんのメカニズム解明や変異体を用いた遺伝子機能解析などで多くの成果が報告されている。しかし近年、表現型と相関する無視できない数の未同定成分が存在する現状から、新規な有用物質やバイオマーカーの探索、子細な現象理解などを目的に、改めてノンターゲット解析の進展を期待する声が高まってきているように思う。

2・4 成分推定のアプローチ

それでは、成分の同定・推定を改善するには、どのようなアプローチが考えられるだろうか？

一つめは、既知化合物のデータベースの拡充を行い、候補化合物が見つからないという状況を減らすことである。しかし Table 1 で見たように、すでに既知化合物のデータ化は飽和しているので、あまり大きな改善は望めないだろう。

二つめは、CID のマススペクトルから、元の化学構造を推定するコンピューター技術の開発である。実はこ

の部分が、メタボロミクスにおける情報処理の一番ホットな研究領域であり、推定ソフトウェアの性能を競うコンテストが開かれるほどである^{7, 8)}。マススペクトルを比較するリファレンスとして、標品を実測したライブラリは有用だが、入手可能な標品の数に制限があるためこれ以上の増強は期待できない。そこで、化学構造情報をもとに、仮想的なマススペクトルのデータベースを作る方策が考案された。実測データをもとに機械学習でフラグメント化のモデルを作る方法や、化学結合エネルギーを計算して壊れやすい部分を見つける方法、フラグメント化の際の水素原子の分配ルールから妥当なフラグメントが生成する構造を選抜する方法、などが提案されている。これら情報科学的方法とは対照的に、特定の化合物群について、経験則から仮想フラグメントを構築する方策も報告されている。最初の成功例は脂質であった (Lipid Search, LipidBlast)^{9, 10)}。筆者らも、植物の大きな二次代謝物グループであるフラボノイドについて、手作業による仮想フラグメントの構築を行い (Flavonoid-Search)、ありふれた野菜であるパセリから新規のフラボノイド候補を多数発見した¹¹⁾。このような、既知の化学構造から仮想フラグメントを構築するアプローチは、一方で、未知の化学構造を持つ物質に対処できない。そこで、既知構造の組み合わせで未知構造をひも解く手法の開発や、研究者コミュニティの知恵を使って妥当な化学構造を推定していく取り組み (GNPS)¹²⁾ なども進められている。

三つめは、意外と盲点であるのだが、ピーク検出精度を上げることである。近年、質量分析装置の性能は急速に進歩しており、小数点以下 3~4 桁まで正確に質量を測定できる装置も普及してきている。しかし、データ解析に用いるソフトウェアのパラメーターチューニングの不十分さなどから、意図せずに、ハードウェアの性能を活かしきっていない結果を得ている場合がよくある。ピーク質量が正確に見積もられているかどうかや、大事なピークを見落としていないかなどは、Fig. 1 のような二次元マスプロットグラム上に、検出されたピークの位置をプロットしてみるとすぐに確認できる。筆者らが提供しているソフト MassChroViewer¹⁾ などで確認しながら、分析条件とともにデータ解析条件を十分に最適化したうえで、ノンターゲット解析を実施することが望ましい。

2・5 問題は、どのピークに的を絞るか

以上のような研究開発により、成分の推定精度は向上してきているが、成分推定ソフトウェアには、ひとつの避けがたい問題がある。それは、返される結果がいつも候補リスト、すなわち、複数の候補化合物を何かしらの

評価スコア順に並べたものであることである。どんなに優れたプログラムでも、トップスコアのものが必ず正解だとは限らない。結局、候補化合物の中から、「この試料には存在して妥当だろう」とか、「こんな特殊な化合物は天然に存在しそうにない」とか、分析データ以外の情報を含めて総合的に判断して、人間が優先順位をつけない必要がある。この状況は、この先どんなにプログラムが進化しても、解決が難しいものではないだろうか。

このように考えると、表現型に相関する興味深い未知ピークがいくつか得られたときの課題の本質は、完璧な推定ソフトがないことではなく、むしろ、どのピークに的を絞ってじっくり時間をかけた総合判断をすればよいか分からない、ということになる。つまり、自分の貴重な研究時間を、どのピークに投資するか。その優先順位付けをするための手がかりが存在しないことが、課題の本質であると言える。

そこで開発したのが、食品メタボロームレポジトリ¹⁾である。

3 研究の方向性を与えるリソース：食品メタボロームレポジトリ

3・1 試料特異性の情報が、重要な手がかりとなる

「その未知ピークは、他にどのような試料で検出されているのか？」という試料の特異性（検出実績）の情報があると、解析の優先順位をつける手がかりとなり、研究の方針決定に役立つ。

例えばの話だが、ヨーグルトの研究をしていて、特定の乳酸菌で作った製品に、高血圧の改善効果が見られたとする。そして、ノンターゲット解析から、効果と相関する成分ピークの候補が20個見つかったとする。推定ソフトなどから候補化合物はいくつか得られるものの、確証はない。それでは、この20個のうちどれを深く掘り下げて研究すればよいだろうか？ これまでは、しらみつぶしに検討するしかなかった。ここで、食品メタボロームレポジトリに m/z 値で検索すると、同じ未知成分が、他のどの食品に含まれる可能性があるかが分かる。例えば候補成分Aは、ヨーグルト類に含まれており、この乳酸菌で作った製品に特に蓄積していて、原料の生乳や、チーズ類など他の食品には含まれていなかったとする。こういう情報が得られると、成分Aはヨーグルトを作る過程で、この乳酸菌で特徴的に生産された成分かもしれないということで、解析の優先順位は上がるだろう。一方、候補成分Bは、野菜や魚、肉などにもランダムに検出されており、強度も特に高いわけではないとする。すると、成分Bはあまり追究する価値がないものとして、解析の優先順位は下がるだろう。

筆者の活用事例では、土壌の分析で特定の試料に特異的に蓄積していた成分を食レポで検索したところ、野菜をはじめ食品には一切検出されていなかった。この情報が逆にヒントとなり解析を進めたところ、土壌微生物が生産する物質であることを同定できた。

3・2 食レポの機能と活用

現在食レポには、222の食品をLC-MSのESIポジティブモードとネガティブモードで分析したデータが掲載されている（Fig. 2）。 m/z 値は、高分解能のフーリエ変換型MSを用いて、5 ppm以内の m/z 検出誤差（運用値）で計測されている。強度の高い成分ピークについては、CIDによるMS²スペクトルをイオントラップ型MSで取得しており、ポジティブモードでは、さらにMS³スペクトルが得られている場合がある。ピーク情報としては、PowerGetBatchソフト¹³⁾で正確に評価された精密質量値およびアダクトイオンの情報をもとに、MF-Searcher⁶⁾というツールで化合物データベース（KEGG, KNApSAcK, HMDB, LIPID MAPS, フラボノイドのデータベース）に検索した結果が閲覧できる。また、ポジティブモードのデータでは、MS², MS³スペクトルからFlavonoidSearch¹¹⁾を使ってフラボノイド非糖部（アグリコン）を推定した結果も付与している。これらのピークリストや生データは、すべてダウンロードすることができ、MassChroViewer¹⁾を用いた詳細な解析に活用することも可能である。また、データの取得や検索などの機能は、すべてAPI（Application Programming Interface）で提供しており、他のプログラミング言語から利用することができるので、大量のデータの自動検索や集計を実施できる。APIを活用すると、例えば小麦製品にだけ含まれるピークを簡単に見つけることができる（<http://metabolites.in/foods/search/mz/379.295/5/ppm/pos/57/1>）。

試料特異性が、解析するピークの優先順位付けに役立つことは上述した通りであるが、食品データそのものの、有用物質の探索などに役立つだろう。例えば上記の小麦ピークのように、特定試料（食材）に存在するマーカー成分で、調理加工などの変化にも安定な成分が見つければ、アレルギー食材の高感度検出などに活用できるかもしれない。また、食事は究極の個人差なので、血液や尿の成分を分析するコホート研究などにおいては、データを正しく解釈するための基盤として役立つ可能性がある。血液や尿には、摂取した食事の成分が、腸内細菌や肝臓で代謝を受けたものが多く含まれている。特に、グルクロン酸や硫酸抱合体はよく知られているが、これらの多くは化合物データベースに記載されていない。しかし、抱合体化反応の質量差分を考慮して、摂取した食品

a トップページ

食品メタボロームレポジトリ

ホーム 食品一覧 検索 ヘルプ

主要な食品に含まれる成分を、液体クロマトグラフィー-質量分析 (LC-MS) によるメタボローム解析で分析したデータを蓄積・公開しています。

特徴1

LC-MSで検出される化合物ピークのパターンを、2次元マップで視覚的に確認できます。

b 食品一覧

ホーム 食品一覧 検索 ヘルプ

食品リスト

食品ID	食品名	食品名	ピーク数 Pos	ピーク数 Neg	イメージ写真
0001	豆類	ひよことうまめ乾	3,329	2,609	
0002	豆類	レンズ豆の煮干し	3,884	3,188	
0003	雑穀類	カシウチナツツライチ	4,433	5,881	
0004	雑穀類	かぼちゃの煮干し	3,827	4,117	
0005	雑穀類	ぜんざい	3,369	3,889	
0006	雑穀類	くすり豆 日本産	3,742	4,553	
0007	雑穀類	くるみ	4,738	4,361	
0008	雑穀類	ごま (白)	7,960	9,208	
0009	雑穀類	ごま (黒)	7,139	8,555	
0010	雑穀類	ビスネ	5,016	6,737	

c ピーク一覧

d ピーク検索機能

e ピーク詳細情報

Fig. 2 食品メタボロームレポジトリの画面と機能。a トップページ。b 分析された食品の一覧。c 各食品で検出されたピークの一覧。d 質量値によるピーク検索画面。e ピークの詳細情報。化合物データベース検索結果、MS²、MS³ スペクトル、FlavonoidSearch によるフラボノイド非糖部推定結果などが表示される。

の成分を解析すると、対応関係が取れる成分が見つかり、体液中の未知化合物を同定・推定できる場合がある。

3・3 実測データの重要性

意外なことに、食レポのような情報リソースはこれまでに存在していなかった。なぜだろうか？ 代謝物とそれが検出された試料について、文献で報告された情報はよく整理されている。天然物ではKNAPSAcK⁽³⁾、ヒトの体内に含まれ得る成分ではHMDB⁽¹⁴⁾ などである。食品に含まれる成分については、食品摂取のバイオマー

カーを探索するFoodBallプロジェクト (<http://foodmetabolome.org>) でも収集されており、現在4つのデータベースが公開されている。しかしこれらはどれも、存在が確認 (あるいは予想) された既知の成分だけを対象としている。一方、未知成分を含めたメタボロームデータを広く蓄積しているレポジトリも存在する。欧州ではMetaboLights⁽¹⁵⁾、米国ではMetabolomicsWorkbench (<http://www.metabolomicsworkbench.org>) が代表的である。これらは言わばアーカイブであり、比較解析でき

る十分な機能を提供していない。この背景として、クロマトグラフィー-MSのデータ(特にLC-MS)は、装置が違えば、溶出時間、イオン化の特性、検出感度などが異なるため、他人が取得したデータを直接利用しにくい現状が挙げられる。日本では、筆者らによる Komic-Market や MassBase¹³⁾、理研による PRIME データベース¹⁶⁾が、それぞれの機関で蓄積してきたデータを公開しているが、実施したプロジェクトの都合上、特定の生物(主に植物)のデータに集中しているなど、試料の多様性に問題があった。このように、成分の試料特異的な存在を、未知成分ピークを含め、実験データをもとに、幅広い試料で集約し、比較・検索できるデータベースはこれまでに存在しなかった。そこで、筆者らは食レポを構築した。なぜ食品なのか? 食品は、様々な生物(植物、動物、魚、微生物など)を含み、また酸化やメイラード反応などが起こりうる加工調理や発酵の過程も含まれるため、成分の多様性が幅広いと予想される。食品だけでなく多くの分野で応用できると考え、採用することにした。

食レポは、異なる装置で測定したデータでも活用できる。食レポ内ではもちろん、統一した手法で測定が行われているため、LCの溶出時間まで考慮して、成分の特異性を試料間で比較できる。それでは上記で指摘したように、異なる装置で取得されたデータは使えないかというと、実はそうでもない。近年では、精密質量の測定は日常的となってきたので、溶出時間を考慮せずに質量の一致だけで検索しても、かなりの成分候補を絞ることができる。実際に食レポを構築して分かったことであるが、このような検索を行うと、何種類の異性体がどのような試料に存在するかが、溶出時間の分布やマススペクトルの差から推測できるというメリットが明らかとなった。化合物データベースには、既知の異性体はたくさん登録されているが、実際に何個くらいを想定すべきかを実測データから確認できることは、未知物質の推定に役立つだろう。

4 今後の展開

2013年12月、和食がユネスコの無形文化遺産に登録された。和食の特徴のひとつとして、「健康的な食生活を支える栄養バランス」があり、長寿や肥満防止との関連性が知られているが、成分からの裏付けは解き明かされていない部分が多い。食レポのような情報基盤は和食の効果を解き明かすためにも役立つかもしれない。また、メタボロームデータを記録しておくことは、和食文化の特徴を成分データとして記録に残すことでもある。未知

の成分を既知に変えていき、より広範な分野で活用するためにも、データの拡充を続けていければ幸いである。特に、脂質の分離に適したLC-MSの分析条件や、官能性に重要な揮発性成分を対象としたGC-MSのデータを追加したい。試料についても、食品以外の廃棄部位や品質劣化したもの、環境試料を拡充することで、未知成分の同定の迅速化を促し、有用成分探索やその遺伝資源の確保、食の安全に役立てられるのではと考えている。

謝辞

本研究の大部分は、前職、公益財団法人かずさDNA研究所で実施されたものである。関係者および、下記の研究支援に感謝申し上げる。千葉県(かずさDNA研究所)、ライフサイエンスデータベース統合推進事業(NBDC, JST)、平成29年度「循環器疾患・糖尿病等生活習慣病対策実用化研究事業」(AMED)、平成29年度 科研費 基盤(C)(17K07810, 日本学術振興会)

文献

- 1) N. Sakurai, D. Shibata, *Carotenoid Science*, **22**, 16-22 (2017).
- 2) R. Hall, *New Phytol.*, **169**, 453-468 (2006).
- 3) F. M. Afendi, T. Okada, M. Yamazaki, A. Hirai-Morita, Y. Nakamura, *et al. Plant Cell Physiol.*, **53**, e1 (2012).
- 4) E. Fahy, S. Subramaniam, R. C. Murphy, M. Nishijima, C. R. Raetz, *et al. J. Lipid Res.*, **50** Suppl., S9-14 (2009).
- 5) S. Kim, P. A. Thiessen, E. E. Bolton, J. Chen, G. Fu, *et al. Nucleic Acids Res.*, **44**, D1202-D1213 (2016).
- 6) N. Sakurai, T. Narise, J. S. Sim, C. M. Lee, C. Ikeda, *et al. Bioinformatics*, **34**, 698-700 (2017).
- 7) I. Blaženović, T. Kind, J. Ji, O. Fiehn, *Metabolites*, **8**, 31 (2018).
- 8) F. Hufsky, K. Scheubert, S. Böcker, *Nat. Prod. Rep.*, **31**, 807-817 (2014).
- 9) R. Taguchi, M. Ishikawa, *J. Chromatogr. A*, **1217**, 4229-4239 (2010).
- 10) T. Kind, K. H. Liu, Y. Lee do, B. DeFelice, J. K. Meissen, O. Fiehn, *Nat. Methods*, **10**, 755-758 (2013).
- 11) N. Akimoto, T. Ara, D. Nakajima, K. Suda, C. Ikeda, *et al. Sci. Rep.*, **7**, 1243 (2017).
- 12) M. Wang, J. J. Carver, V. V. Phelan, L. M. Sanchez, N. Garg, *et al. Nat. Biotechnol.*, **34**, 828-837 (2016).
- 13) N. Sakurai, T. Ara, M. Enomoto, T. Motegi, Y. Morishita, *et al. BioMed Research International*, **2014**, 1-11 (2014).
- 14) D. S. Wishart, Y. D. Feunang, A. Marcu, A. C. Guo, K. Liang, *et al. Nucleic Acids Res.*, **46**, D608-D617 (2017).
- 15) K. Haug, R. M. Salek, P. Conesa, J. Hastings, P. de Matos, *et al. Nucleic Acids Res.*, **41**, D781-D796 (2013).
- 16) T. Sakurai, Y. Yamada, Y. Sawada, F. Matsuda, K. Akiyama, *et al. Plant Cell Physiol.*, **54**, e5 (2013).