

オミックス解析のリスク評価への応用（メタボローム解析ツール開発）

研究分担者 有田 正規（理化学研究所環境資源科学研究センター）

研究要旨：

質量分析計およびソフトウェア性能の向上により、ノンターゲット・メタボローム解析（DIA 解析）が実用になってきた。DIA 解析法はスペクトルの精度が劣るものの、試料検体に含まれる代謝物を網羅的に測定できる。つまり、蓄積された測定データを後日に再解析し、有用情報を後から抽出することが可能になる。本研究では、DIA 解析のメリットを活かした、質量フラグメントに基づく食品の安全性解析法の構築をおこなう。研究の仮定は、安全でない代謝物成分に対応する情報が、DIA 解析で取得される網羅的な MS/MS 情報の中に存在するというものである。そうした危険成分を、MS/MS シグニチャーとして取り出すことで安全性指標の構築を試みる。具体的には以下の流れで研究を遂行する。

【2018 年度】実測スペクトルを収集し、そこに含まれる MS/MS フラグメント・ライブラリを構築した。

【2019 年度】マススペクトルを入手できない既知アレルゲン等の生理活性物質について、フラグメンテーション予測によるフラグメント・ライブラリを構築する。

【2020 年度】構築したライブラリを用いて、機械学習手法により MS/MS シグニチャーを抽出する。

A. 研究目的

メタボロミクスにおいて生データの再解析を可能にするノンターゲット・メタボローム解析結果を通して、マススペクトルから食品の安全性を判断する技術を開発する。測定されたデータを様々なリポジトリから取得、再解析することで、安全な食品に含まれるマススペクトルと、そうでないものを見分ける手法を開発する。

B. 研究方法

マススペクトルから化合物の部分構造を推定するためのフラグメント・データベースを構築した。まず、NIST14、Metlin、MassBank、ReSpect、GNPS、MetaboBase データベースから、前駆体イオンが見えており質量精度が 10 mDa 以下の、良質のスペクトルを収集した。それらのスペクトルにおける各 MS/MS フラグメントについて、精密質量に基づいて組成式を付与した。付与できないフラグメントは除いた。更に、前駆体の化学構造からフラグメンテーション規則に基づいて導出される分子構造を列挙し、先に付与した組成式と照合することで、MS/MS フラグメントの構造を推定、SMILES 表現した。こうして見出された MS/MS フラグメントに ClassyFire ソフトウェアにて化合物クラスを

付与した。

C. 研究結果および考察

元になるスペクトル全データは、ポジティブおよびネガティブイオンモードのそれぞれで 17.8 万、4.6 万スペクトルであった。そこから各スペクトルの精度を精査して抜き出すとそれぞれ 5.3 万、2.0 万の MS/MS 情報が残り、それらへのアノテーション結果として、MS/MS フラグメント 2996 件およびニュートラルロス 1644 件のライブラリを作成した。このライブラリはスペクトル中に複数出現するものを頻度順に並べて作成してある。これを用いて、標品の存在しない植物二次代謝物成分が予測できるか試みたところ、トウモロコシにおける DIMBOA 関連代謝物やフラボノイドの配糖体など、標品から糖や有機酸の修飾によって生じる二次代謝物の予測が可能であることが判明した。

出現頻度が多いフラグメントから順に精査すると、ポジティブおよびネガティブのフラグメントで頻度が高いものは多くがベンゼン環構造、ニュートラルロスでは糖や短鎖脂肪酸であった。いずれにも共通するハロゲン構造は少なく、ハロゲン化合物が共通するフラグメントを持たないことが示

唆された。

また同時に食品 220 品目と生薬 120 品目の LC-MS データを入手し、ピーク同定作業を実施した。このデータはノンターゲット解析ではないため MS/MS 情報が不十分ではあるが、インソース・フラグメント情報等を精査して抽出することで、作成したフラグメント・ライブラリとあわせて解析できると考えている。今後、見出されたフラグメントと食品および生薬データの組み合わせにより、毒性度の予測が可能かどうかを検証する予定である。

D. 結論

既存のスペクトルライブラリに収蔵されるスペクトルに対して MS/MS フラグメントのアノテーションを実施し、頻出するフラグメントおよびニュートラルロスのライブラリを作成した。またその有効性を検証した。

E. 業績

論文発表

1. Hiroshi Tsugawa, Ryo Nakabayashi, Tetsuya Mori, Yutaka Yamada, Mikiko Takahashi, Amit Rai, Ryosuke Sugiyama, Hiroyuki Yamamoto, Taiki Nakaya, Mami Yamazaki, Rik Kooke, Johanna Bac-Molenaar, Nihal Oztolan-Erol, Joost Keurentjes, Masnaori Arita, Kazuki Saito "A cheminformatics approach to characterize metabolomes in stable isotope-labeled organisms" *Nature Methods* (accepted)
2. Burla B, Arita M, Arita M, Bendt AK, Cazenave-Gassiot A, Dennis EA, Ekroos K, Han X, Ikeda K, Liebisch G, Lin MK, Loh TP, Meikle PJ, Orešič M, Quehenberger O, Shevchenko A, Torta F, Wakelam MJO, Wheelock CE, Wenk MR "MS-based lipidomics of human blood plasma: a community-initiated position paper to develop accepted guidelines" *Journal of Lipid Research*, 59(10), 2001-2017, 2018

学会発表

1. Arita M "Open genome analysis in the post-genomic era" International Workshop on

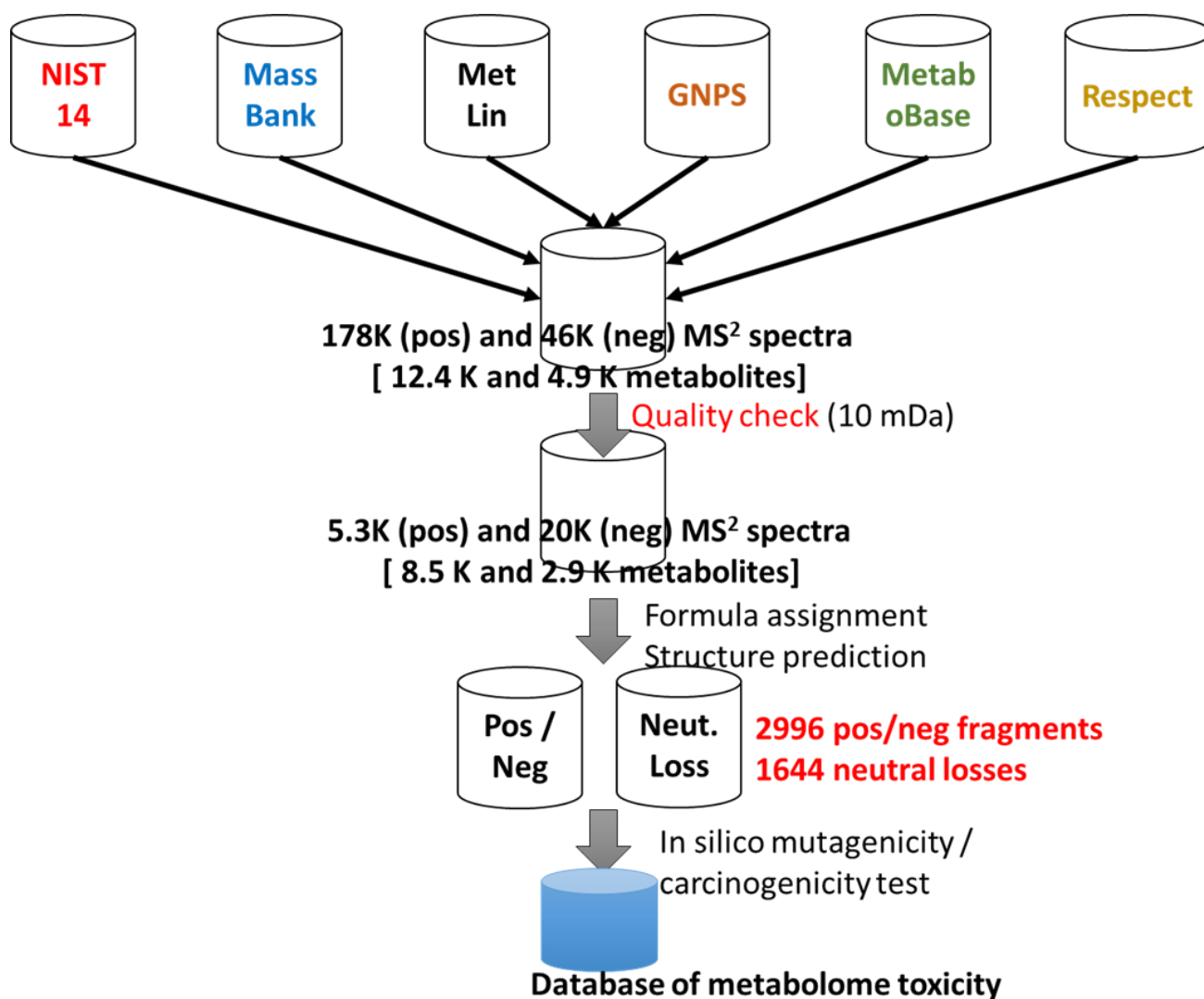
Data Science, Mishima, November 15 (12-15), 2018

2. Arita M "Computational Metabolomics" 6th Annual Korea Metabolomics Society (plenary), Seoul, Korea, April 6 (5-6), 2018

F. 知的財産権の出願・登録状況

該当なし

公共マスタデータベースをもとに Quality で選抜したデータの作成と毒性判定への応用



	A	B	C	D	E	F	G
	m/z	Formula	IonMode	Frequency	Fragment short	InChIKey	Fragment nu
1	78.959054	C3P-	Negative	8.19	ABLZXFXXLZCGV		0
2	91.054227	C7H7+	Positive	8.18	YXFVABEGXRONW;UAE		8
3	87.056386	C3H7N2O-	Negative	8.02	UYDPHMCYSNAY		0
4	151.00368	C7H3O4-	Negative	8.02	BTQAJGSMXCDDAJCRI		3
5	77.038577	C6H5+	Positive	7.62	UHOVQNZJYSORNBXD		1
6	93.034588	C6H5O-	Negative	6.02	ISWSIDIOOB;JBQZ;USLR		3
7	96.969619	H2O4P-	Negative	5.74	NBIXXVUZAFLEBC		0
8	89.024418	C3H5O3-	Negative	5.31	PEDCGBHVMGVHV;JVT		3
9	153.01824	C7H5O4+	Positive	5.26	BTQAJGSMXCDDAJCRI		2
10	59.013853	C2H3O2-	Negative	5.1	LYCAIKOWRPUZTN;QTE		2
11	79.054227	C6H7+	Positive	5.03	UHOVQNZJYSORNB;AF;		9
12	227.03498	C13H7O4-	Negative	4.12	MFVHZNTCKOGIT;RDL;		4
13	71.013853	C3H3O2-	Negative	3.92	YPFDHNVEDLHUCEXBC		3
14	152.01151	C7H4O4-	Negative	3.86	BTQAJGSMXCDDAJCRI		1
15	107.01385	C6H3O2-	Negative	3.79	YCIMNLLNPGFGHC;GHM		3
16	101.02442	C4H5O3-	Negative	3.72	ARXKVVRIIOZGF;YAXK		6
17	119.03498	C4H7O4-	Negative	3.69	UNXHWFMPAWPI;DZ;		3
18	106.02168	C6H4O2-	Negative	3.67	YCIMNLLNPGFGHC;GHM		2
19	85.029503	C4H5O2-	Negative	3.66	BMRWNKZVCUKKS;PUI		9
20	135.00877	C7H3O3-	Negative	3.4	IUNJCFABHJZSKB;DGX;		6
21	117.03459	C8H5O-	Negative	3.38	FUGYGGDSWSUORM;HX		10
22	94.041316	C6H6O+	Positive	3.34	ISWSIDIOOB;JBQZ		0
23	91.018938	C6H3O-	Negative	3.25	ISWSIDIOOB;JBQZ		0
24	87.008768	C3H3O3-	Negative	3.17	PEDCGBHVMGVHV;JVT		2
25	93.033491	C6H5O+	Positive	3.1	ISWSIDIOOB;JBQZ		0
26	89.038577	C7H5+	Positive	3.03	YXFVABEGXRONW		0
27	300.02755	C15H8O7-	Negative	2.97	REFJMTPEDEVJUY;JKHR		2
28	107.04914	C7H7O+	Positive	2.79	IWDCLRJOB;JRNH;RDC		6