

# MassBankにおける(MS)<sup>n</sup>スペクトル検索法の開発

蓬萊尚幸<sup>1,5)</sup>\* 有田正規<sup>1,3,4,5)</sup> 二瓶義人<sup>1,5)</sup> 池田奨<sup>1,5)</sup> 諏訪和夫<sup>3,5)</sup>  
尾島雄也<sup>1),5)</sup> 嘉数勇二<sup>1,5)</sup> 曾我朋義<sup>1,5)</sup> 西岡孝明<sup>1,2,5)</sup>

<sup>1)</sup>慶應義塾大学先端生命科学研究所 <sup>2)</sup>京都大学大学院農学研究所

<sup>3)</sup>東京大学大学院新領域創成科学研究科 <sup>4)</sup>理化学研究所植物科学研究センター <sup>5)</sup>JST-BIRD

## 概要

MassBankは、代謝物を中心に多くの化合物のマススペクトルの収集を目的とした公開データベースである。様々な実験条件(イオン化手法、衝突分解手法、分析機器、など)の下で測定されたマススペクトルの包括的な収集を特徴とする。特に、高分解能ESI-(MS)<sup>n</sup>のスペクトルを多く収集する。現時点でのMassBankの登録状況を表1に示す。10月31日現在での総スペクトル数は12,727である。

本ポスターでは、MassBankにおけるスペクトル検索アルゴリズムと(MS)<sup>n</sup>スペクトル検索結果の効果的な表示方法、MS/MSスペクトルを用いた評価、および、(MS)<sup>n</sup>スペクトル検索の適合度改善のアイデアについて発表する。

データ供給機関	分析法	スペクトル数	化合物数
慶應大IAB	QqTOF-MS/MS	4,510	695
	QqQ-MS/MS	4,275	
	IT-MS <sup>n</sup>	515	
理研PSC	GC-TOF-MS	241	334
	LC-TOF-MS	85	
	LC-QqQ-MS/MS	87	
	CE-TOF-MS	20	
日本ウォーターズ	LC-Q-MS	2721	577
	QqQ-MS/MS	273	
かずさDNA研		近日公開予定	
東大医学系研究科		近日公開予定	
京大薬学研究科		近日公開予定	
鳥取大農学研究科		近日公開予定	

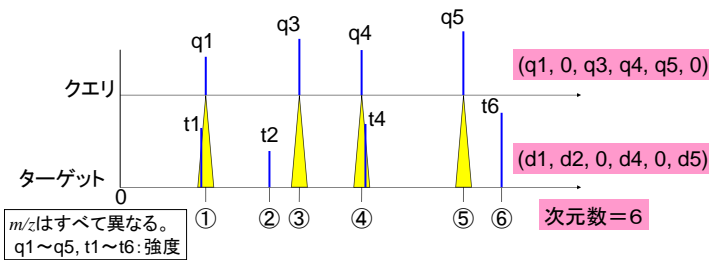
表1: MassBankの登録状況 (2007年10月31日現在)

## MassBankのスペクトル検索アルゴリズム

現在、MassBankでは、NIST MS Searchなどで利用されている整数 $m/z$ のためのコサイン相関を用いたスペクトル検索法を実数 $m/z$ 用に拡張した検索法を用いている。ユーザ指定可能な許容範囲内の $m/z$ を同一視することでクエリとターゲットが張るベクトル空間の次元を確定し(図1)、コサイン相関を求めている。許容範囲内に複数のピークが存在する場合は、クエリのピークを基準に次元を決定する。

また、強度について、 $m/z$ の2乗をかける重み付け、100で割り平方根をとる正規化を行っている。ユーザが強度のカットオフ値を設定し、それ以下のピークは無視することも可能である。

図1: 許容範囲を用いた次元の決定



## MS/MSスペクトルの3D表示

MassBankでは複数のスペクトルを三次元表示するビューを提供している。図2では、Tyrosine methyl esterのMS/MSスペクトルを衝突エネルギー順に並べて表示している(手前が低衝突エネルギー、奥が高衝突エネルギー)。衝突エネルギーの増大によりフラグメンテーションが起こっていく様が明確に表現されている。また、MassBankに登録されている複数段階(図2では5段階)のスペクトルの連続性/不連続性の度合いを知ることができる。

図2: MS/MSスペクトルの3D表示



## MassBankデータベース://www.massbank.jp

このデータベースは(独)科学技術振興機構バイオインフォマティクス推進事業「メタローム・マススペクトル統合データベースの構築」によるものである。

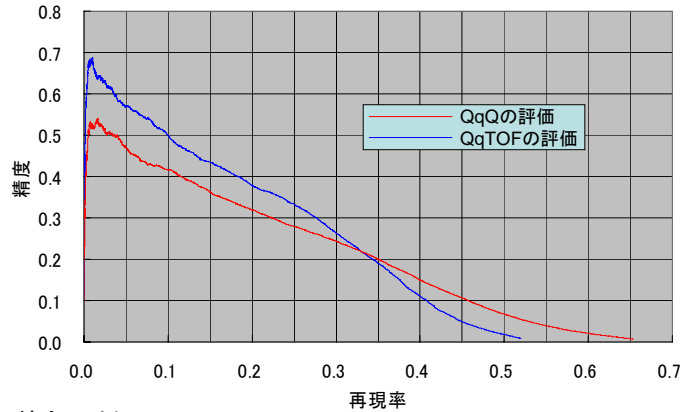
このデータベースは、日本質量分析学会協力事業である。

## MS/MSスペクトル検索の評価

MassBankに登録されているスペクトルのうち、慶應大IABで同一化合物について測定したQqTOF-MS/MSスペクトルとQqQ-MS/MSスペクトルを評価対象とクエリ集合として用い、正しく化合物を同定したかを基準にして評価を行った(図3)。

高スコア領域(グラフの右端)では、QqTOFとQqQの間での検索は有効であることが判った。また、高スコア領域ではQqTOFの方がQqQよりも検索精度が高い。これによりTOFのような精密スペクトルの収集は、低精度スペクトルでの測定結果の同定に役立つことが判った。

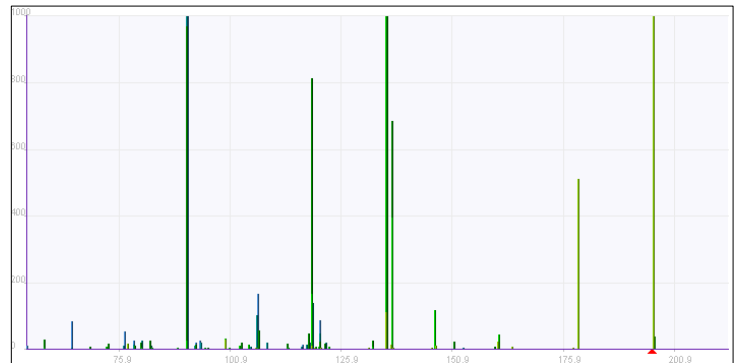
図3: MS/MSスペクトルの検索の評価



## 統合スペクトル

データベース内の複数の衝突エネルギーで測定されたスペクトルを重ね合わせることで、収録されていない過渡的な衝突エネルギーのスペクトルに対する適合度をあげる研究を実施している。複数のスペクトルを重ね合わせる方法は複数考えられる。もっとも単純な重ね合わせ方として、ピークの $m/z$ に注目して、どれか1個のスペクトルだけに現れるピークはすべて集める、複数のスペクトルに表れるピークについては最も強度が高いものを選ぶ、という方法がある。この方法で図2のスペクトルを重ね合わせたものが図4である。

図4: 統合スペクトル



## ピーク同定と検索適合度向上

MassBankプロジェクトでは、QqTOF-MS/MSやIT-(MS)<sup>n</sup>を中心にフラグメントイオンピークの同定を行っている。この同定結果を検索にも利用し、同定できたピークの重みを高めたり、同定された分子式から計算される $m/z$ の理論値を検索時に用いることで、検索能力を向上させる研究を実施している。

たとえば、同定できたピークについてはその $m/z$ の値を理論的な精密質量で補正することでスペクトルのデータクリーニングを行うことによる適合度向上、同定できた特徴的なピークに重み付けを行うことによる精度向上、などが考えられる。

## 天然物のマススペクトルの提供にご協力ください

MassBankでは、これまでに日本の研究機関で単離、精製されてきた貴重な天然物のマススペクトルを収集、公開することによって、天然物科学研究に寄与することも目的の一つとしています。

データの提供方法や著作権などについてのお問い合わせは、[massbank@iab.keio.ac.jp](mailto:massbank@iab.keio.ac.jp)にメールをお送りください。