

質量分析インフォマティクス研究会・第1回ワークショップ（2016年度）

## 質量分析データ解析のための計算科学

開催日時： 2016年10月7日(金)午後1時～午後5時50分  
(午後12時30分開場)

開催場所： JST 東京本部(サイエンスプラザ)地下1階大会議室（東京都千代田区  
四番町5-3）

主 催： 質量分析インフォマティクス研究会※（日本バイオインフォマティクス学  
会(JSBi)）

協 賛： JST/NBDC 統合化推進プログラム「プロテオーム統合データベースの構  
築」課題(jPOST) (<http://jpost.org/>)

※ 本ワークショップは、質量分析インフォマティクス研究会が、JSBiの公募研究会として活動する  
一環として開催しています。

## プログラム

1. 開会挨拶	13:00 ~ 13:10 (10分)
-----	
2. 質量分析/メタボロミクス	13:10 ~ 13:50 (40分)
「MassBank: 質量分析インフォマティクスのデータ基盤」	
	西岡 孝明 (京都大学名誉教授)
-----	
3. ソフトウェア	13:50 ~ 14:30 (40分)
「Made in Japan のオープンソース・ソフトウェアを世界へ・・・ 質量分析データ表示・解析ソフト Mass++ の現状と今後の展望」	
	田中 聡 (Trans-IT)
-----	
休憩	14:30 ~ 14:45 (15分)
-----	
4. プロテオミクス	14:45 ~ 15:25 (40分)
「定量プロテオミクスプラットフォームのためのインフォマティクス」	
	松本 雅記 (九州大学生体防御研)
-----	
5. データベース	15:25 ~ 16:05 (40分)
「プロテオーム統合データベース jPOST:リポジトリの開発と今後」	
	奥田 修二郎 (新潟大学医歯学総合)
-----	
休憩	16:05 ~ 16:20 (15分)
-----	
6. グライコミクス/グライコプロテオミクス	16:20 ~ 17:00 (40分)
「Endoglycosidase と LC/MS/MS を用いたグライコプロテオーム解析」	
	○太田 悠葵, 川崎ナナ (横浜市大生命医科学)
-----	
7. 総合討論	17:00 ~ 17:40 (40分)
-----	
8. 閉会挨拶	17:40 ~ 17:50 (10分)
-----	
懇親会	18:15 ~ 20:30 (予定)

○…講演者(複数著者の場合)

(司会・会場進行: 吉沢 明康(京都大学化研), 河野 信(DBCLS), 守屋 勇樹(DBCLS))

## 講演要旨

### 質量分析／メタボロミクス

MassBank : 質量分析インフォマティクスのデータ基盤

西岡 孝明 (京都大学名誉教授)

MassBank が収集したスペクトルデータは、もともとは、メタボローム解析で検出された代謝物の同定を目的としていましたが、質量分析の高精度化にともなって、イオンの解裂機構をはじめ代謝物の化学構造推定、異なるスペクトルデータベース間の相互参照などの研究に利用されてきました。ワークショップではこれらのトピックスについて紹介させていただく予定です。

## ソフトウェア

Made in Japan のオープンソース・ソフトウェアを世界へ・・・

質量分析データ表示・解析ソフト Mass++ の現状と今後の展望

田中 聡 (Trans-IT)

Mass++は 2006 年に CREST プロジェクトで始まり、2010 年に FIRST プロジェクトで開発を継続、2015 年に FIRST プロジェクト終了と共にオープンソース・ソフトウェアとして公開された質量分析データの表示・解析ソフトである。質量分析データを扱う為には通常、機器付属のソフトウェアが用いられていたが、それらは

1. 機器によってソフトウェアが異なり、使用方法やできる解析も異なる。
2. 外部からの操作が通常許されておらず、大量データの処理が困難。
3. 自分達で新しい解析方法を思いついても、その実現には手間がかかる

という問題点があった。それらを解決する為に Mass++は

1. 複数の機器データを読み込む事ができるマルチベンダー対応ソフトとする。
2. 大量処理と相性の良い、コマンドラインプログラムを用意する。
3. プログラムの知識があれば簡単に機能を追加できる様にプラグイン構造のソフトウェアとして実現したり、スクリプト機能を持たせる。

といった方法で対応し、Mascot や X! Tandem を用いたタンパク質同定、MassBank を用いた代謝物同定、定量機能や糖鎖解析機能等を開発してきた。本セッションでは現在のオープンソース・ソフトウェアとしての Mass++開発の取り組みや、今後の展望について紹介・解説する。

## プロテオミクス

定量プロテオミクスプラットフォームのためのインフォマティクス

松本 雅記 (九州大学 生体防御医学研究所  
トランスオミクス医学研究センター プロテオミクス分野)

質量分析計の感度や速度、さらには分解能の向上によって網羅的なタンパク質同定・定量が可能となっている。しかしながら、現在の主流である data-dependent acquisition (DDA) による手法は技術的な制約から高深度の分析を多検体に対して行うことは困難である。近年 MRM などのターゲット・プロテオミクスの手法が確立され、ハイスループットに任意のタンパク質の正確な定量が可能になり、新たなプロテオーム解析基盤を提供しつつある。その一方で、ターゲット・プロテオミクスを実施するためには事前に取得したスペクトルやクロマトグラムなどの膨大な情報の管理・運用からクロマトグラムのピーク検出に至るまでの過程が極めて煩雑であることから、その有効性に反して普及が遅れている。本ワークショップではこれまでにわれわれが開発してきたターゲット・プロテオミクス解析プラットフォームを、情報管理および処理にフォーカスをあてて紹介し、今後の定量プロテオミクスにおけるインフォマティクスの重要性について議論したい。

## データベース

プロテオーム統合データベース jPOST:リポジトリの開発と今後

奥田 修二郎 (新潟大学 大学院医歯学総合研究科)

質量分析に基づいたプロテオミクス分野が近年めざましく進展している。その結果、様々なプロテオームのデータセットが大量に生成されつつある。通常、これらのデータセットはリポジトリに登録され、そのアクセッション番号が論文などで引用される。このプロテオームデータに対するリポジトリは、欧米のグループが中心に立ち上げた ProteomeXchange コンソーシアム ( PXC ) の下で整備が進められてきたため、日本を含むアジア諸国からはデータ登録の際の転送速度などで問題が多かった。そういった問題を解決するため、ライフサイエンス統合データベースの枠組みの中で 2015 年より jPOST ( Japan ProteOme STandard ) プロジェクトが立ち上げられ、リポジトリの開発が進められている。その結果、jPOST リポジトリは、2016 年 5 月に公開され、現在すでに百を超えるプロジェクトが登録されている。本リポジトリの主な特徴として、(1) 実験プロトコルに即した入力インターフェイス、(2) 高速データアップロードシステム、(3) 柔軟なファイル管理システム、などが挙げられる。また、jPOST は PXC の正式メンバーになっていることから、データ登録ユーザーは自身のプロジェクトに対して現在プロテオームではデファクトになりつつある PXC の ID ( PXD ) を紐付けることが可能である。

一方、jPOST では、リポジトリに登録されたデータを統一的な手法に基づいて再解析する予定である。また、その再解析データは、統合化されデータベースとして公開することになっている。そのため、jPOST データベースが完成すると、研究者ごと・実験ごとのバラ付きがないデータをあらゆる切り口で俯瞰することができるようになる。本講演では jPOST リポジトリ・データベースの現状と今後について紹介する。

## グライコムクス/グライコプロテオミクス

Endoglycosidase と LC/MS/MS を用いたグライコプロテオーム解析

○太田 悠葵, 川崎 ナナ (横浜市立大学院生命医科学研究科  
プロテオーム科学研究室)

N 結合型糖鎖は重要なタンパク質の翻訳後修飾であり、近年、その構造及び機能の解析が急速に進められている。このような糖タンパク質の解析方法として、LC/MS/MS を用いた方法が広く用いられるようになってきている。LC/MS/MS を用いた糖タンパク質の解析アプローチには大きく分けて3つのアプローチ: 遊離糖鎖を測定するアプローチ、インタクトタンパク質を測定するアプローチ、糖タンパク質を消化し、糖ペプチドを測定するアプローチが存在する。遊離糖鎖を測定するアプローチは糖鎖の不均一性や詳細構造情報を取得できるものの、N 結合型糖鎖の結合位置の情報が失われてしまうという欠点がある。インタクトタンパク質を測定するアプローチは試料の処理が簡便であるという利点があるものの、高感度に測定できないという欠点がある。糖タンパク質を消化し、糖ペプチドとして測定する手法は、遊離糖鎖の測定のような異性体の識別はできないものの、N 結合型糖鎖の糖鎖の結合位置と糖鎖構造を同時に得ることができる点で有用な手法となっている。

しかし、プロテオーム解析と比較して、糖ペプチドを網羅的に LC/MS/MS で解析するグライコプロテオーム解析の試料処理手法、データ解析手法は発展途上であり、当研究室では、この LC/MS/MS を用いたグライコプロテオーム解析手法の開発を行っている。

一般に、LC/MS/MS を用いたグライコプロテオーム解析では、peptide-*N*-glycosidase (PNGase F) を用いてタンパク質から糖鎖を切り離し、糖鎖結合位置のアスパラギン (Asn) がアスパラギン酸 (Asp) に変化することを利用して糖鎖結合位置を決定する。次に、糖ペプチドの構造を決定するためには、該当糖ペプチドのプリカーサーイオンを MS<sup>n</sup> 分析に供さなければならないが、夾雑物であるペプチドのイオンと糖ペプチドのイオンを明確に区別して、糖ペプチドのピークだけを選択的に MS<sup>n</sup> 分析に供することは困難であった。

当研究室では、PNGase F の代わりに、endo-*N*-acetylglucosaminidase F1, F2, F3 (Endo F1, F2, F3) で処理して、N 結合型糖鎖結合位置の Asn 上にキトビオースコアの *N*-acetylglucosamine (GlcNAc) を一つだけ残して (この時 GlcNAc には core Fucose が結合していてもよい) 糖鎖を切断し、N 結合型糖鎖の結合位置を同定する手法を開発した。本手法の利点は、N 結合型糖鎖結合位置の Asn が Asp となり、ペプチドの電荷が変化する PNGase F 処理と異なり、Endo F1, F2, F3 で糖鎖を遊離させたペプチドは電荷の変化が起こらないため、保持時間が糖ペプチドの保持時間と近接するという特徴がある。この特徴により Endo F1, F2, F3 処理したペプチドの質量、保持時間から、糖ペプチドの保持時間、およびプリカーサーイオンの *m/z* 推定が可能となることで、より効率的に MS<sup>n</sup> 分析に糖ペプチドのイオンを供し、糖ペプチドを同定することが可能となった。