

質量分析インフォマティクス研究会・第11回ワークショップ (2026年)

古都で考える 質量分析データ活用の未来

開催概要

開催日時： 2026 年 9 月 11 日 (金)

【午前の部】 9:30 ～ 11:30 (9:15 開場)

【午後の部】 13:30 ～ 18:00 (13:00 開場)

開催場所： 【午前の部：ポスター発表会場】

京都府立大学下鴨キャンパス構内・京都府立 歴彩館 小ホール

〒606-0823 京都府京都市左京区下鴨半木町1-29

【午後の部：講演会場】

京都府立大学下鴨キャンパス構内・

京都三大学・教養教育共同化施設 稲盛記念会館 104号室

〒606-8522 京都市左京区下鴨半木町1-5

主 催： 質量分析インフォマティクス研究会 (JCompMS)※ (<https://ms-bio.info/>)
(日本バイオインフォマティクス学会 (JSBi)) (<https://www.jsbi.org/>)



主 催： 日本質量分析学会スペクトルデータ部会

主 催： 京都府立大学大学院生命環境科学研究科

※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

開場	9:15
ポスター発表	9:30～11:30
休憩	11:30～13:30
開会挨拶	13:30～
ポスター ショートプレゼンテーション	13:35～
招待講演	14:30～
MB-POSTとMassBank Human：メタボロームデータの共有とアノテーションを促進 する新基盤の紹介	
松田 史生	(大阪大学)
招待講演	15:15～
血液プロテオームをどこまで読めるかー臨床データ統合から由来臓器の推定まで	
夏目 やよい	(医薬健栄研)
ポスター賞表彰	16:00～
休憩	16:10～16:30
講習会	16:30～
オミクスデータの公開とメタボロミクス	
吉沢 明康	(新潟大学)
メタボロミクス・データリポジトリMB-POST利用講習会	
高橋 悠志	(新潟大学)
閉会挨拶	18:00～
閉場	18:05

※講習会はハンズオンセミナーではありません



講演・ポスター発表要旨

招待講演

MB-POSTとMassBank Human：メタボロームデータの共有とアノテーションを促進する新基盤の紹介

○松田史生

(大阪大学)

Keywords: メタボローム, 質量分析, MassBank, データレポジトリ

質量分析を用いたメタボロミクスにおけるデータ共有および代謝物のアノテーションを推進するには、インフォマティクス基盤のさらなる発展が不可欠である。そこでJST統合化推進プログラム、シン・マスバンクプロジェクトでは、MB-POSTやMassBank Humanなどの基盤構築を進めている。2025年より運用を開始したMB-POSTは、迅速かつ簡略化されたデータ登録ワークフローを提供する。一方、MassBank Humanは、十分なアノテーションが付与されたヒト由来MS/MSスペクトルのライブラリを提供し、代謝物アノテーションの深さと幅を向上させる。

メタボロームデータレポジトリであるMB-POST (<https://repository.massbank.jp/>) は、新潟大学の奥田らが中心となり標準化とユーザビリティに重点を置いて構築を進めている (Anal. Chem. (2026) 98 (31): 22609–22617)。フロントエンドにReact、PHPバックエンドにSlim Frameworkを用いた。メタデータの選定は一般的な解析ワークフローに基づいて行い、基盤となるアーキテクチャは将来的なりポジトリ間の統合を追及している。プロテオームリポジトリであるjPOSTの技術的枠組みと直感的な操作感 (look and feel) を継承することで、研究者にシームレスなインターフェースと最適化されたデータ送信機能を提供する。

MassBank Human (<https://human.massbank.jp/>) は慶應大学の平山らが中心となつての開発を進めている。MB-POST、Metabolomics Workbench、およびMetaboLightsなどのリポジトリから、ヒトのDDAメタボロームデータセットを取得した。すべてのデータ処理手順は、自作のPythonスクリプトを用いて実施した。公開されているヒトメタボロームデータセットのDDAデータファイルに含まれるMS2スペクトルを平均化してノイズや測定誤差を低減することにより、高精度なMS2スペクトルライブラリを構築した。MS-FINDERやSIRIUSなどのツールを用いてアノテーションを行った結果、スペクトルの7%が信頼性レベル3 (confidence level 3) に達した。データはMassBank HumanおよびGitHubにて公開されている。MassBank Humanのデータにより、代謝物アノテーションの深さと幅が向上した。

招待講演

血液プロテオームをどこまで読めるか – 臨床データ統合から由来臓器の推定まで○夏目やよい^{1,2,3}(医薬基盤・健康・栄養研究所¹, 大阪大学大学院薬学研究科², 大阪大学蛋白質研究所³)**Keywords:** 細胞外小胞, 創薬標的探索, プロテオーム

特発性肺線維症（IPF）は進行性で治療選択肢が限られる疾患である。しかし本邦では外科的肺切除が行われる機会が少なく、病変部組織を得ることは容易ではない。そこで我々は、血清中の細胞外小胞（extracellular vesicle, EV）のプロテオームを病変の分子状態を反映する低侵襲な代替指標と位置づけ、電子カルテ・血液検査値・CT読影所見といった臨床情報と対応づけて解析してきた。大規模かつ異種の臨床表現型と定量プロテオームを統合する際の難しさは、両者の対応がデータ全体ではなく「一部の項目どうし」に局限する点にある。この構造を捉えるために subset binding 法を開発し、IPF患者が濃縮される部分集団の層別化条件と、その集団に対応する血清EVタンパク質群を同定した。さらに上流因子の推定を通じて創薬標的候補を絞り込み、培養細胞を用いた検証まで進めている。

本発表では、この一連の研究から派生して現在進めている二つのアルゴリズム開発について紹介する。第一に、subset binding が依拠する Apriori 型の網羅探索は、属性数の増加と、大きな頻出項目集合の出現という二つの経路で計算が破綻する。そこで探索を微分可能な目的関数の勾配最適化として再定式化した Differentiable Subset Binding (DSB) を開発した。任意サイズの項目集合を固定長の重みベクトルとして表現することで組合せ爆発を回避し、より高次元のオミクスデータに適用できる。

第二に、血中EVは複数の臓器に由来する信号の混合であり、測定されたプロテオームだけでは個々の分子の由来臓器も、各臓器の寄与割合も分からない。この課題に対し、半教師あり非負値行列因子分解によるタンパク質単位の由来臓器推定と、ベイズデコンボリューションによる検体単位の臓器割合推定を並行して行う EV-STAR-Protein を構築した。公開データを用いた検証では、既知の臓器マーカーを高い特異性で回復し、疾患群において標的臓器由来EVの割合が上昇する傾向を捉えられることを確認している。臨床の課題を起点に解析手法を開発することで見えてきた設計上の判断と、現時点での限界についても議論したい。

講習会

オミクスデータの公開とメタボロミクス

○吉沢明康

(新潟大学 医学部メディカルAIセンター)

Keywords: オミクス解析, データ公開, データリポジトリ, FAIRデータ原則, Shin-MassBank

2024年以降、日本学術振興会の科学研究費（科研費）を用いた研究で、採択研究者がデータマネジメントプラン（DMP）を作成し、研究データの管理・利活用のための公開が強く推奨されるようになったことは、多くの研究者の知るところである。しかしこのような“研究データの公開”の潮流は、最近突然に始まったことではなく、古くは20世紀末のヒトゲノム計画まで遡ることができる[1]。

本講演ではまず、メタボロミクスに限定せず、プロテオミクスやグライコミクスなど質量分析を用いたオミクスの研究者にも関係のある一般論として、(i) データ公開の流れが特にオミクス解析分野で拡大していく経緯を概観しつつ、その理念や動機について考察、更に (ii) 質量分析に基づくオミクス解析でのデータ公開方針の雛形となった、プロテオミクスに於けるデータ公開、そして (iii) メタデータの必要性和標準化について論じる[2]。

その上で、(iv) メタボロミクスのデータ公開と現在進行中のShin-MassBank計画について紹介し、メタボロミクス・データリポジトリMB-POST講習会の導入としたい[3]。

【文献】

- [1] Contreras, J. L.; Knoppers, B. M. The Genomic Commons. *Annu Rev Genomics Hum Genet* **2018**, *19*, 429–453.
- [2] 吉沢 明康; 小林 大樹; 河野 信. データ公開とデータ論文～データ論文とは何で、何故必要で、何が求められるのか～. *Proteome Letters* (日本プロテオーム学会誌) **2025**, *10* (1), 11–26.
- [3] Takahashi, Y.▽; Yoshizawa, A. C.▽; Kiuchi, T.; Hayasaka, R.; Torigoe, T.; *et al.* Reviving Data Potential—The Reanalysis-Driven Metabolomics Data Repository MB-POST. *Analytical Chemistry* **2026**, *98* (31), 22609–22617.



講習会

メタボロミクス・データリポジトリMB-POST利用講習会

○高橋悠志^{1,2}

(新潟大学 医学部メディカルAIセンター¹, 新潟大学 ひと脳研究資源イニシアチブ推進センター²)

Keywords: メタボローム, 質量分析, データリポジトリ

Shin-MassBank プロジェクトにおいて開発されている MB-POST はメタボロームデータのための新たな質量分析データリポジトリである。プロテオームデータのための jPOST リポジトリやグライコームデータのための GlycoPOST から定評のある使いやすいインターフェースや高速なファイルアップロードシステム、メタデータ入力の手間を省くプリセットの仕組みなどを継承しており、さらにウェブブラウザ上のみで完結し外部プログラムのインストールを必要としないという特徴を持つ。

本講演では、ユーザー登録からデータアップロードおよび公開に至るまでの一通りの MB-POST の使い方についてステップバイステップで紹介し、メタボローム研究におけるデータ共有のための新たな選択肢として MB-POST がどのように活用できるかを示す。

大規模言語モデル (LLM) を利用したLC-MSメソッドデータベースの構築

○西野耕平

(徳島大学・技術支援部)

Keywords: LLM, LC-MS, データベース

近年、液体クロマトグラフィー質量分析法 (LC-MS) は、低分子から高分子まで幅広い分析対象に利用されている。一方、分析対象に応じて前処理、分離カラム、移動相、質量分析条件、データ解析法などを選択する必要がある、分析系の構築には多くの文献調査を要する。そこで、本研究では、大規模言語モデル (LLM) を用いてオープンアクセス (OA) 論文からLC-MS分析法に関する情報を構造的に抽出し、分析条件の検索・比較を支援するLC-MSメソッドデータベースを構築した。

情報の抽出にはオープンウェイトモデルQwen3.6-35B-A3Bを用いた。要旨執筆時点で2,234報の論文を処理し、2,795件のLC-MSメソッドを抽出した。主な項目としてサンプルの前処理法、抽出溶媒、分析カラム、移動相、グラジエント条件、MS装置、イオン化法、極性、データ取得モード、解析ソフトウェア、サンプルの生物種、組織、疾患、処理条件、測定対象化合物などを構造化した。分野別に分けるとリポミクス1,098件、メタボロミクス935件、プロテオミクス318件、グライコミクス62件、その他382件であった。また、測定対象として計52,213件の化合物情報 (重複あり) を抽出した。現在、化合物名の抽出は検索に利用可能な水準である一方、低分子化合物の同定レベル、特に標準品を用いて確認したか否かについては追加検証を進めている。

データベースの妥当性を評価するために、200報の論文をランダムに抽出し、LCおよびMS条件に関して2,794項目を元の論文と照合した。その結果、2,664項目 (95.35%) を正しいと判定した。なお、本評価は値が抽出された項目の正確性を評価したものであり、未抽出情報を含む再現率は評価していない。サンプル情報については50報から抽出された617項目中614項目 (99.5%) が正しいことを確認した。さらに、Vision-Language Model (VLM) を併用して、論文中に図の存在を判定させたところ、192報を対象に検証するとクロマトグラムでは適合率96%、再現率90%、マススペクトルでは適合率99%、再現率88%であった。

本データベースにより、分析対象やサンプルに応じて過去に使用されたLC-MS分析条件を横断的に検索できるほか、各分析分野で用いられているカラム、前処理、装置、取得モードなどの傾向を解析できる。今後はデータベースの拡充と抽出精度の検証を進め、LC-MS分析系構築時の文献調査を支援するデータベースとしての活用を目指す。

☆ポスター 2

Real-time machine-learning classification of glycopeptides enables enrichment-free glycoproteomics

○Bingyuan Zhang¹, The Huong Chau¹, Kristina Mae Bienes¹, Hiromu Arakawa¹, Christopher Ashwood², Hiroyuki Kaji¹, Yusuke Matsui¹, Rebeca Kawahara¹, Morten Thaysen-Andersen¹

(iGCORE, Nagoya University, Nagoya, Japan¹, Protea Glycosciences Pty Ltd, New South Wales, Australia²)

Keywords: 質量分析, グライコプロテオミクス

本研究では、質量欠損 (mass defect) 情報を利用した機械学習型糖ペプチド分類器について、ヒト血清を対象とした実時間 (real-time) 測定への実装を検討した。今回、2 種類の概念実証 (PoC) を構築した。第一に、事前測定で予測した糖ペプチド前駆体を選択リストとして次回測定に用いるオフライン方式、第二に、装置制御用インターフェース (iAPI) を介して各 MS1 スキャンで前駆体を即時に判定し、DDA における MS/MS 選択へ直接反映するオンライン実時間方式である。実時間予測では、1 回の MS1 スキャンにつき前駆体を 0.1 ミリ秒未満で処理でき、測定周期への影響はほとんど認められなかった。提案した方法とも従来の強度依存型 DDA と比較して、糖ペプチドスペクトルの割合および glycoPSM 数が増加した。また、分類閾値を高く設定することで、糖プロテオームの解析範囲を維持しつつ予測精度が向上した。さらに、既知の血清 N 型糖鎖組成との比較から、従来 DDA に比べて糖鎖組成の偏りが小さいことも確認された。以上より、iAPI を利用した糖ペプチド前駆体の実時間選択は、事前濃縮を必要としない血清糖プロテオーム解析を実現する有効な手法であることが示された。

☆☆ポスター 3

公共プロテオームデータ駆動型解析による新規 O-フコシル化タンパク質の発見

○清水勇希^{1,2}, 木村美遥^{1,2}, 鈴木健裕¹, 堂前直^{1,2}

(理化学研究所環境資源科学研究センター生命分子解析ユニット¹, 埼玉大学大学院理工学研究科²)

Keywords: 質量分析, プロテオーム, 翻訳後修飾, O-フコシル化

プロテオミクス技術の発展により、公共データベース上では大量のプロテオームデータが公開されている。そこで我々は、公共データベースからの翻訳後修飾探索法の確立を試みており、中でも哺乳類において特異的な質量であるタンパク質の O-フコシル化に着目してきた。O-フコシル化は EGF-like ドメイン、TSR ドメインの適切なコンセンサス配列中に POFUT1/2 がそれぞれフコースを転移させることが知られている。しかし、コンセンサス配列に基づき 100 個以上の O-フコシル化タンパク質が予測されているが、大部分の実在は未検証である。また近年、POFUT3/4 による O-フコシル化が報告されたが、実例は少ない。本研究では公共ヒトプロテオームデータについて再解析を行い、膜侵襲複合体構成因子である Complement component 9 (C9) を新規 O-フコシル化タンパク質として同定した。さらに修飾部位が従来のコンセンサス配列を持たない EGF-like ドメインに存在することを見出したので解析結果について紹介する。

★ポスター 4

湛水栽培したダイズの根圏土壌におけるメタボローム解析

○岡市和也, 佐藤友昭, 杉山暁史

(京都大学生存研)

Keywords: メタボローム, 根圏, 微生物代謝, 植物微生物相互作用

根の周辺領域である根圏では、植物が根から分泌した、糖やアミノ酸、二次代謝産物などが土壌微生物に影響を与えることで、特徴的な微生物叢が形成される。また、植物は環境に応じて根から分泌する代謝物の組成を変化させ、それに伴って根圏微生物叢を変化させることが知られている。一方、根から分泌された代謝物が土壌微生物によって変換されることも報告されており、根圏に存在する代謝物は、植物微生物間の相互作用や環境応答を反映していると考えられる。特に湛水環境では、植物におけるストレスに加え、根圏の酸素環境も変化するため、植物側の代謝や物質分泌と土壌微生物側の代謝の双方が変化すると考えられる。そこで、栽培作物であるダイズを対象に、湛水栽培および非湛水栽培における根圏代謝物を質量分析によるメタボローム解析で比較した。湛水・非湛水条件間で根圏代謝物プロファイルに違いが認められ、湛水環境における根圏の代謝変化を網羅的に捉えることができた。

☆☆ポスター 5

同位体パターンスクリーニングを活用したセレノメタボロミクス研究

○山岸由和

(千葉大学)

Keywords: メタボローム, 質量分析, 同位体比, セレン, 腸内細菌

セレン (Se) は天然に6種類の同位体 (^{74}Se : ^{76}Se : ^{77}Se : ^{78}Se : ^{80}Se : ^{82}Se = 0.9: 9.4: 7.6: 23.8: 49.6: 8.7) が存在し、その同位体存在比はSe含有化合物を特徴付ける強力なフィンガープリントになっている。特に、未知のSe含有化合物の同定には、Se含有化合物のスぺンエーションで用いられるLC-ICP-MSだけでは不可能であり、このSeの特徴的な同位体存在比に基づいて、ESI-MSで得られるマススペクトラムにより可能となる。しかし、多様なマトリクスで構成される生体試料中のSe存在量は微量であるために、今日においても未知のSe含有化合物の同定は難易度が高い。そこで高感度であり、Se含有化合物に対する特異性を高めることが可能なESI-Orbitrap-MSを用いた新たなSe含有化合物の同定法について検討し、Se同位体パターンスクリーニング (Se-IPS) と名付けた新規分析法を開発した。そこで本発表では、我々の研究グループが開発したSe-IPSや、そこから得られたSe含有化合物の同定の取り組みについて紹介する。

☆☆ポスター 6

プロテオミクスデータにおけるlog2FC空間の構造的特徴

○西崎愛花¹, 荒木令江², 河野信¹(北里大学¹, 熊本大学²)**Keywords:** プロテオミクス, Tandem Mass Tag (TMT), 品質管理 (QC), log2 fold change (log2FC)

近年、質量分析を用いたプロテオミクスでは、Tandem Mass Tag (TMT) による多重化定量データが広く利用されている。TMTを用いたプロテオミクスでは、定量値が参照サンプルに対するlog2 fold change (log2FC) として提供されることが多い。log2FCは解析間比較に有用である一方、参照サンプルの変動や欠損構造の影響を受け、分布シフト、外れ値、符号反転など特有の統計構造を生じうる。しかし、log2FC空間がどのような構造を持ち、それらが何によって形成されるのかは十分に検討されていない。一方、従来の品質管理 (QC) は主としてタンパク質方向の評価に依存しており、実験条件単位の品質を体系的に評価する手法は十分ではない。本研究では、log2FCデータを対象とした実験条件単位の多指標QCスコアリング手法を開発した。本研究では、実験条件ごとの特徴を多面的に記述する指標群 (descriptor) を設計し、log2FC空間の統計的・構造的特徴を記述できるか検討した。

CPTACで公開されているTMT10/11プロテオミクス49プロジェクトを解析対象とした。まず、欠損率、CV、SD、MAD、IQRに基づきタンパク質方向の品質フィルタリングを行った。続いて、実験条件ごとに欠損率、欠損改善量、逆符号頻度、外れ値頻度、順位構造一貫性、中央値シフトを算出し、中央値およびMADを用いてロバストに標準化した0-20スケールのdescriptorへ変換した。シミュレーションデータを用いた性能評価に加え、CPTACデータに対して相関解析、主成分分析、階層クラスタリングを実施した。また、Hallmarkタンパク質群除去、欠損タンパク質除去、ランダムタンパク質除去など複数の条件でdescriptorの変化を比較し、各descriptorが反映する構造について検討した。

人工データによるシミュレーション解析では、提案したdescriptorは技術的劣化サンプルを高精度に識別した。また、劣化の程度に応じて寄与するdescriptorが異なることが確認された。CPTACデータでは、欠損関連、外れ値・符号関連、順位関連など複数の統計的特徴が異なる変動軸を形成し、実験条件ごとの特徴を多面的に記述できることが示された。さらに、Hallmarkタンパク質群の除去ではdescriptor構造は大きく変化せず、一方で統計的に不安定なタンパク質を除去する品質フィルタリングでは一部descriptorに特徴的な変化が観察された。これらの結果は、descriptorが単一の生物学的経路ではなく、log2FC空間に残存する統計的・構造的特徴を反映している可能性を示唆する。以上より、本手法はlog2FCデータに特有の構造的異常を検出するQC手法として、公開プロテオミクスデータの品質評価に有用であると考えられる。

☆ポスター 7

実運用に基づく質量分析データ管理基盤の構築と外部リポジトリ連携への取り組み

○小池仁美¹, 東裕介¹, 前田真秀¹, 小松輝久¹, 鈴木健裕², 大月香³, 中川れい子⁴, 堂前直², 宮坂信彦³,
清水義宏⁴, 俣賀宣子¹, 高橋悠志⁵, 吉沢明康⁵, 奥田修二郎⁵, 大浪修一¹

(理化学研究所最先端研究プラットフォーム連携 (TRIP) 事業本部¹, 理化学研究所環境資源科学研究センター², 理化学研究所脳神経科学研究センター³, 理化学研究所生命機能科学研究センター⁴, 新潟大学医学部メディカルAIセンター⁵)

Keywords: 質量分析, メタデータ, データ管理, 自動化, 電子ラボノート

近年、プロテオミクス分野では、ProteomeXchangeを中心とした国際的なデータ共有基盤の整備が進んでいる。一方で、試料情報や測定・解析条件などのメタデータが十分に整備されていないことや、記述方法の違いは、データの再利用や統合解析、異なるシステム間での相互運用性を妨げる要因となっている。

我々は、複数の質量分析計や解析ソフトウェアから得られるデータを集約し、測定ファイルからのメタデータ抽出と統制語彙 (Controlled Vocabulary) を用いた入力支援により、メタデータを半自動的に付与するプロテオームデータ管理システムを開発している。これにより、研究者の入力負担を軽減するとともに、多様な試料や解析手法に対応したメタデータを効率的かつ一貫して管理できるようになった。また、認証・アクセス制御やバックアップ環境を整備し、複数拠点で安定して運用できる基盤を構築した。

さらに、本システムでは電子ラボノート (ELN) との連携により、実験情報と測定データを一元的に管理し、研究データ管理の効率化と重複入力の削減を図っている。現在は、取得したメタデータをELNへ自動登録する機能の実装を検討するとともに、LLMを活用したメタデータ入力支援や開発支援にも取り組んでいる。

また、jPOSTチームが整備したセキュアなデータ転送環境および専用のプロジェクト管理コマンドを用いて、外部リポジトリとの連携に向けた通信試験を実施した。現在は、メタデータ変換方法や運用ルールの整理を進めており、今後は実データを用いた検証を行い、システムへの実装につなげる予定である。

本発表では、実運用を通じて得られたメタデータ管理システムの知見と、電子ラボノートおよび外部リポジトリとの連携に向けた取り組みについて紹介する。

☆☆ポスター 8

ダイズの硫黄欠乏代謝応答

○家田愛菜, 渡邊むつみ, 峠隆之

(奈良先端科学技術大学院大学)

Keywords: 硫黄, ダイズ, メタボロミクス

硫黄 (S) は植物に必須である。S欠乏はS代謝のみでなく、炭素・窒素代謝にも影響を及ぼし、マメ科植物では収量低下や種子の品質低下が報告されている。そこで本研究では、ダイズとシロイヌナズナを対象にS欠乏応答を比較し、共通および種特異的な応答を明らかにすることを目的とした。植物をFull Nutrient (FN) 条件とLow S (LS) 条件で21日間生育し、シュートと根の代謝物をGC-MS、LC-MSおよびUPLC-PDAにより解析した。S欠乏マーカー代謝物のO-アセチルセリンはシロイヌナズナで増加し、ダイズの根でも増加傾向を示した。代謝物の蓄積パターンは、シロイヌナズナではS欠乏による処理間差が顕著な一方、ダイズでは器官間差が顕著であった。両種で共通してフェニルプロパノイド類は増加し、種特異的な代謝物では、シロイヌナズナのグルコシノレート類やダイズのソヤサポニン類も蓄積パターンが変化した。以上より、S欠乏応答には種間で共通する変化に加えて、種や器官に特異的な応答が存在することが示唆された。これらの知見は、ダイズのS欠乏適応機構の理解に繋がることが期待される。

★ポスター 9

クズ(葛)が産生する特化代謝物の器官特異的な産生機構の解析

○HUANG ZHIXUAN, 渡邊むつみ, 峠隆之

(奈良先端科学技術大学院大学)

Keywords: 植物二次代謝, クズ, フラボノイド

クズ (*Pueraria lobata*) はマメ科のつる性植物で、主に日本、中国および東南アジアに分布し、その根は漢方薬として利用されてきた。近年、クズ根に含まれる成分が心血管疾患や糖尿病などに対して生物活性を示すことが報告されている。クズ根にはフラボノイド類、ポリフェノール類、サポニン類などの代謝物が豊富に含まれ、特にイソフラボノイド配糖体であるプエラリン (8-C-グリコシルイソフラボン) が主要成分として知られている。また、プエラリンにはメチル化体や配糖体など多様な誘導体の存在が報告されている。一方、葉にはフラボン類、花にはアントシアニン類、つるにはイソフラボノイド類が多く蓄積することが知られているが、器官や生育段階による代謝物の蓄積パターンや、それらの生合成経路および修飾酵素遺伝子については十分に解明されていない。そこで本研究では、液体クロマトグラフィーと紫外検出器および質量分析計 (UPLC-PDA-MS) を用いて、クズの各器官および生育段階におけるフラボノイド類を包括的に解析し、その蓄積パターンを明らかにするとともに、クズにおけるフラボノイド生合成経路を構築した。

☆☆ポスター 10

ヒト腸内メタプロテオミクスにおける再利用可能な検出可能ペプチドデータベースの構築

○孫富韜, 吉沢明康, 高橋悠志, 瀧原速仁, 白井福寿, 奥田修二郎

(新潟大学医学部メディカルAIセンター)

Keywords: ヒト腸内メタプロテオミクス, MS再解析, データベース検索

ヒト腸内メタプロテオミクスは、腸内微生物が発現するタンパク質および機能を直接評価できる手法であるが、検索データベースの構築はいまだ大きな課題となっている。メタプロテオミクスでは、検索データベースの設計において、検索感度と微生物多様性の保持との間に本質的なトレードオフが存在する。経験的に選択した小規模データベースは検索空間を縮小できる一方で微生物多様性を損なう可能性があり、逆に大規模なタンパク質データベースは多様性を保持できるものの、検索空間の増大やペプチド共有により同定感度の低下を招く。

本研究では、この課題を解決するため、ヒト腸内メタプロテオームにおいて実際に検出されたペプチド・タンパク質空間に基づく再利用可能な縮小データベースの構築を目的とした。複数の独立したデータセットから得られた約100名分の欧米人糞便メタプロテオミクスデータに対し、1,000種類以上の腸内細菌ゲノムを対象として、MaxQuantによるゲノム単位の個別検索を実施した。各検索でそのペプチドが同定されたタンパク質を抽出・統合し、実際に観測可能な腸内微生物タンパク質のみを集約したデータベースを構築した。このデータベースを用いて追加データセットの二次検索を行うことで、ゲノム単位検索を繰り返すことなくペプチドおよびタンパク質同定を可能にすることを目指した。さらに、ペプチドおよびタンパク質群の同定数、相対タンパク質量、個体間出現頻度、分類学的多様性、機能プロファイルを評価した。

本手法は、大規模なゲノム単位検索で得られた微生物多様性を維持しながら検索空間を縮小できる実用的な解析フレームワークを提供する。より大規模なデータの蓄積は今後も必要であるものの、本研究は実際のメタプロテオミクスデータで観測されるペプチドに基づいて縮小データベースを構築できる可能性を示し、標準化されたヒト腸内メタプロテオミクス解析ならびに集団・疾患別リファレンスデータベース構築の基盤となることが期待される。

★ポスター 11

代謝物構造予測プログラムのアンサンブル学習による高精度モデルの開発

○廣瀬泰樹¹, 鳥越大平², 松田史生³, 和泉自泰², 津川裕司^{1,4}(東京農工大学¹, 九州大学², 大阪大学³, 京都大学⁴)**Keywords:** メタボロミクス, インフォマティクス, 機械学習, アンサンブル学習

近年、質量分析 (MS) を用いたメタボローム解析では、組成式や化学構造を推定するアノテーションプログラムが多数開発されているが、構造推定精度は依然として不十分である。一方、MS-FINDER、SIRIUS、MetFrag、msbuddyは異なる特徴量およびアルゴリズムを利用しているため、正解候補の提示にはプログラム間で相補性が認められる。本研究ではこの相補性を活用するため、複数ツールの予測結果を統合するアンサンブル学習基盤MS-Emblatorを構築した (<https://github.com/systemsomicslab/msemblator>)。組成式推定にはMS-FINDER、SIRIUS、msbuddy、構造推定にはMS-FINDER、SIRIUS、MetFragを用い、NIST26由来17,401スペクトルを用いて学習した。XGBoost、LightGBM、CatBoostの3つをCASMI 2022のスペクトルデータにより評価した。最良の単一ツールと比較して、組成式推定で2.2ポイント、構造推定で1.3ポイント高い正答率を示した。さらに、不正解スペクトルについても、正解化合物のfingerprintおよびTanimoto係数解析から、正解構造に高い類似性を持つ候補を提示できることが確認され、プログラム間の相補性を利用したアンサンブル学習の有効性が示された。

☆☆ポスター 12

構造に基づくターゲットデコイ法によるFDR制御可能な代謝物アノテーションの理論武装(FDR estimation for metabolite identification using structure-based target-decoy approach)○倉田明咲^{1,2}, 津川裕司²(東京農工大学¹, 京都大学²)**Keywords:** メタボロミクス, FDR予測, Target decoy approach

質量分析による化合物アノテーションはスペクトルライブラリーとの一致度に基づいて行われる。ただし、一度の分析で数百のピークが観測され、繰り返しアノテーションが付与されることから、結果には偽陽性が含まれる。プロテオミクスにおいてはtarget-decoy approachによるFalse Discovery Rate (FDR) 推定が、同定結果の信頼性を定量的に評価する手法として広く利用されている。一方、メタボロミクスでは対象となる低分子化合物の構造が非常に複雑かつ多様であることからFDR推定法が十分に確立されていない。

本研究では、低分子化合物のスペクトルライブラリー検索における高精度なFDR推定を目的として、新たなデコイライブラリー構築法を開発した。まず、Natural product likeness score(NP score)を指標として、繰り返し化合物構造の組み換えを行うことで自然界には存在しないような構造(デコイ構造)を生成した。次に、スペクトル予測モデルであるFIORA[1]を用いて、疑似的なマススペクトルを生成することによってデコイライブラリーを構築した。

構築したデコイライブラリーのFDR推定性能を、既存手法(Fragmentation tree based method[2], XY-Meta[3])と比較した。推定FDRと実測FDRの一致度を評価した結果、既存手法と比較して実測FDRに近い推定値を示した。さらに、本手法によるFDR制御を公共メタボロミクスデータに適用したところ、コサイン類似度0.7以上を閾値とする従来の検索条件と比較して、平均X%多くのピークにアノテーションを付与することができた。

[1] Nowatzky, Y. *et al.* *NATURE COMMUNICATIONS*, **16**, 2298 (2025).

[2] Scheubert, K. *et al.* *NATURE COMMUNICATIONS*, **8**, 1494 (2017).

[3] Li, D. *et al.* *ANALYTICAL CHEMISTRY*, **92**, 5701-5707 (2020).

ポスター 13

高脂肪食負荷時に応答する腎臓内Gb3の発現制御機構の解析

○浅川秀太郎¹, 津川裕司^{1,2}, 時吉花菜子², 長友洸樹²

(京都大学¹, 東京農工大学²)

Keywords: 腎臓, リピドミクス, RNA-seq, 質量分析

高脂肪食負荷は腎臓の脂質代謝を変化させ、腎障害の発症・進展に関与することが知られている。また、ヒトの腎障害では発症・進展に男女差が存在し、代謝応答の違いもその一因となる可能性がある。我々は高脂肪食負荷マウス腎臓のリピドーム解析から、スフィンゴ糖脂質の一種である globotriaosylceramide (Gb3) を含む複数の脂質分子が食餌条件に応答して変動し、その応答には雌雄差や年齢差が認められることを見出した。さらに、同条件のマウス腎臓を対象としたRNA-seqデータから、脂質代謝および炎症に関連する遺伝子発現の変化を明らかにした。現在は、ヒト腎近位尿細管上皮細胞HK-2を用い、高グルコース、脂肪酸および炎症刺激による脂質プロファイルの変化をリピドミクスで解析し、あわせて炎症関連遺伝子の発現をqPCRで評価している。本発表では、高脂肪食負荷に応答して変動する腎臓Gb3とその雌雄差・年齢差を報告し、培養細胞を用いた検証を含めて、腎障害における糖脂質代謝の制御機構について議論したい。

★ポスター 14

電力駆動型CO₂資源化プラットフォーム構築に向けたギ酸資化大腸菌の代謝工学

○直林大生^{1,2}, 竹川悠里¹, 田代洋平³, 平野伸一⁴, 津川裕司^{1,2}

(京都大学大学院 薬学研究科¹, 東京農工大学 工学府², COLA株式会社³, 電力中央研究所⁴)

Keywords: 代謝工学, メタボロミクス, 合成生物学

カーボンニュートラル社会の実現に向け、CO₂を資源として有用物質へと変換する技術の開発が進められている。電気化学的にCO₂を還元する手法は、ギ酸などのC1化合物を生成できる一方、いわゆる低分子化合物の選択的合成には不向きである。そこで我々は、CO₂電気還元と微生物代謝による物質生産を組み合わせたハイブリッドバイオものづくり研究を行なっている。

本研究では、炭素固定経路としてカルビンベンソンサイクルと比較して高いエネルギー効率と代謝可塑性を有する還元的グリシン経路 (RGP) に着目している。我々は近年、RGP関連遺伝子をゲノム上に持ち、ギ酸独立栄養増殖を可能にする大腸菌株の構築に成功した。さらに、クエン酸回路に関連する代謝関連遺伝子の改変を行うことで、最大ODは約0.1から0.2へ向上し、倍加時間を約半分に短縮することに成功した。今後は、さらなる増殖能の向上に加え、リナロールをはじめとするモノテルペンをモデル化合物とした有用物質生産の実現を目指す。本発表ではこれまでの研究進捗に加え、メタボロミクス解析を活用したRGPのボトルネック同定および代謝改変戦略など、代謝工学をさらに加速させるための方法論について議論したい。

★ポスター 15

脂質分子間の相対的溶出順序に基づくアノテーション自動キュレーション手法の開発

○木村英, 松沢佑紀, 津川裕司

(京都大学)

Keywords: リピドミクス, LC-MS/MS, 保持時間, 保持順序, 脂質アノテーション

LC-MS/MSを用いたメタボロミクスでは、プロダクトイオンスペクトルがアノテーションに広く利用される。しかし保持時間(RT)は化合物由来の構造情報を明確に含むにもかかわらず十分に活用されていない。また絶対的な保持時間は分析条件に強く依存するが、分子ペアの相対的な溶出順序は異なるLC分析系間でも比較的保存されやすいことが報告されている(Bach et al., Bioinformatics 2018)。本研究では、リピドミクスにおいて分子の炭素数・二重結合数がRTと直接的に対応する性質に着目し、脂質分子間の相対的な溶出順序の一貫性からアノテーションの妥当性を自動判定するキュレーション手法を開発している。本手法はRTそのものに依存しないため、条件や装置が変わっても転用可能な脂質アノテーション精度向上基盤となることが期待される。

★ポスター 16

MS-DIAL駆動型リピドミクスデータ解析に資するMCP開発とAIエージェントとの連携

○亀谷優歩, 松沢佑紀, 津川裕司

(京都大学)

Keywords: リピドミクス, 質量分析, MCP, LLM, LC-MS/MS

リピドミクス研究において、液体クロマトグラフィー-タンデム質量分析 (LC-MS/MS) から出力されるデータの解析にはMS-DIALが頻用される。一方、MS-DIALから出力される結果の解釈は解析者の経験に依存し、再現性の確保と解析時間の長さに課題が残る。そこで我々は、MS-DIALを用いたリピドミクスデータ解析を、AIエージェントから対話的に操作可能にするModel Context Protocol (MCP) 基盤の構築を進めている。

本研究では、MS-DIALが出力するバイナリーファイル5形式 (.arf, .arf2, .pai2, .dcl, .EIC.aef) および標準形式mzTab-Mの読み取り・解析に加え、MS-DIAL Consoleの実行・成果物管理などを行う合計51種のMCPツールを実装した。対応する処理は、品質管理 (QC) 付きデータ前処理、主成分分析、二群間比較解析と結果出力、抽出イオンクロマトグラムの可視化、精密質量・アダクトイオン種によるアノテーション結果の検証とMS/MS根拠の確認、および解析目的・知識・レポートの記録である。解析状態と注意事項を機械可読に保持し、LLMの推測による誤解析を抑制する設計とした。さらに、mzTab-Mの構造検証と定量種別の推定、および解析条件や成果物のハッシュ値を記録するジョブ情報を介した結果受け渡しにも対応した。今後は、定量値の意味、極性、来歴の検証を強化した標準連携へ拡張するとともに、公開研究情報やスペクトルデータベースとの接続を目指す。

本発表では、実装と検証結果に加え、専門家の判断根拠を保ちながら解析の負担と属人性を低減する方法論、および質量分析データ解析の標準化・自動化の展望を議論したい。

★ポスター 17

マウス脳の脂質代謝理解に資するマルチモーダルリピドミクス情報計測基盤の構築

○山元らな¹, 坂本七海¹, 清水邦善², 杉浦悠毅^{2,3}, 本田真也⁴, 宮本潤基¹, 林青順⁵, 今見考志⁵, 津川裕司^{1,2}

(東京農工大学¹, 京都大学², 慶應義塾大学ヒト生物学-微生物叢-量子計算研究センター³, 東京科学大学⁴, 理化学研究所生命医科学研究センター⁵)

Keywords: 質量分析, リピドミクス

昨今、質量分析 (MS) を用いた脂質メタボローム解析 (リピドミクス) は生命科学研究で広く利用されているが、トランスクリプトミクスと比較して空間分布や細胞タイプ情報を同時に捉えることは困難である。本研究ではまず、マウス脳を対象に、数理モデルによる統合解析に適した3つの相補的なリピドミクス基盤の構築を試みた。具体的には、(1) マトリックス支援レーザー脱離イオン化 (MALDI) 法を用いた高空間分解能リピドミクス (MALDI-MS)、(2) レーザーマイクロダイセクション (LMD) と液体クロマトグラフィータンデム型質量分析 (LC-MS/MS) を用いた低空間分解能リピドミクス、(3) 細胞分画による細胞タイプ別リピドミクスの3つである。各分析モダリティについて、材料・前処理・分析条件の検討を行った。その結果、MALDI-MSでは手動マトリックス塗布により自動塗布と同等のイメージングデータを取得し、低空間分解能リピドミクスでは96ウェルプレート上で工程を一元化し、 2.5×10^3 cells相当から283分子の脂質代謝物を検出した。また、細胞分画ではニューロン、オリゴデンドロサイト、アストロサイト、ミクログリアに対する抗体と分画手順の最適化により877分子を検出し、細胞種ごとのリピドーム情報の取得を可能にした。本発表では、各モダリティにおける材料選定、前処理および分析条件の最適化について主に報告し、各モダリティデータの統合解析を志向したデータの質・特徴について紹介する。

ポスター 18

Structural elucidation of *N*-acyl ornithine lipids in *Akkermansia muciniphila* by EAD MS/MS

○汪念^{1,2}, 轟善勝^{1,2}, 津川裕司^{1,2}(京都大学大学院 薬学研究科¹, 東京農工大学 工学府²)**Keywords:** *N*-acyl ornithine, electron-activated dissociation, lipid structural analysis

The gut microbiota produces structurally diverse lipid metabolites that regulate host immune and metabolic functions. *Akkermansia muciniphila* is a promising probiotic, and its ornithine lipids have been reported to modulate host immunity. However, conventional collision-induced dissociation tandem mass spectrometry (CID MS/MS) cannot resolve structural isomers differing in hydroxyl position and chain branching, leaving the precise structures of microbial lipids largely uncharacterized. Here, we applied electron-activated dissociation (EAD) MS/MS to systematically characterize *N*-acyl ornithine lipids (NAOrn) produced by *A. muciniphila*. After solid-phase extraction (SPE) enrichment and method validation with synthetic standards, we found that EAD MS/MS yielded diagnostic fragment ions for hydroxyl position and chain branching that were absent in CID spectra. Applying this workflow to *A. muciniphila* lipid extracts revealed that hydroxylation occurs predominantly at the C-3 position of odd- and straight-chain fatty acids. This multidimensional approach, integrating SPE, EAD MS/MS, and chromatographic retention behavior, enabled the first systematic characterization of the structural diversity of *A. muciniphila*-derived NAOrn, providing an analytical platform for understanding microbial lipid biosynthesis and host immunomodulation.

★ポスター 19

開裂ルールを外部から定義できるプロダクトイオンスペクトル予測手法の構築

○小川誠寛¹, 津川裕司²

(東京農工大学¹, 京都大学²)

Keywords: 機械学習, in-silico MS/MSスペクトル, フラグメンテーション

質量分析に基づくメタボロミクスでは、得られるMS/MSスペクトルの多くが標準品ライブラリーに収載されておらず、未知化合物のアノテーションが課題となっている。そのため、膨大な化合物構造情報と未知スペクトルを関連付ける計算的アプローチが求められている。そこで我々は、化合物構造からin-silico MS/MSスペクトルを生成する機械学習手法を開発している。NIST23に収載された約100万件のCID MS/MSを学習に用い、化学結合の開裂反応をSMIRKS形式で定義してその生成確率を推定することとした。単結合開裂のみを対象にCASMI2022データセットで評価した結果、実測スペクトルとのcosine類似度は0.26であり、既存ツールの0.29に近い性能を示した。また、本手法は任意の開裂ルールを追加できるため、フラボノイドでみられる1,3-retrocyclizationをはじめ、各化合物クラスに特徴的なフラグメンテーションにも拡張可能である。今後は、柔軟かつ解釈可能なスペクトル生成基盤の改良に加えて、未知化合物の構造候補探索やアノテーション支援への応用を目指す。



日本バイオインフォマティクス学会 入会のお誘い

日本バイオインフォマティクス学会（JSBi）は、我が国においてバイオインフォマティクスという学問分野を発展させ、その技術および関連事業の振興、並びにその教育基盤を確立するために、平成 11 年に設立されました。この学会は、バイオインフォマティクス分野最大の国際学会である ISCB（International Society for Computational Biology）の地域グループとして、またアジアの関連学会の連合体である AASBi（Association of Asian Societies for Bioinformatics）のメンバーとしても、活動しています。学会の主な活動としては、**年 1 回、年会を開催して会員の研究発表の場を提供しているほか、各種の研究会および地域部会活動、日本バイオインフォマティクス学会賞、Oxford Journals-JSBI Prize の授与、JSBi Bioinformatics Review 発行、ニュースレター発行**などを通して、多くの会員にとって意義のある学会であるように努めています。

例えば、以下のような方のご入会をお待ちしております。

- ☐ バイオインフォマティクス分野の**最先端研究に関する研究討議**を行いたい方
- ☐ バイオインフォマティクス分野での**研究交流やネットワーキング**を行いたい方
- ☐ バイオインフォマティクス分野の**知識や技術を新たに身につけたい方**
- ☐ バイオインフォマティクス分野の**共同研究者や人材を探されている方**
- ☐ バイオインフォマティクス分野に**関連する製品やサービスを開発・提供されている方**
- ☐ その他、バイオインフォマティクス分野に関する**ご関心をお持ちの方**

ぜひ多くの方にご入会いただき、今後のさらなる活動の発展にご協力いただければ幸いです。

入会方法のご案内

入会申し込みは学会ウェブページ（www.jsbi.org）から随時受け付けております。入会金は1,000円、年会費は5,000円（正会員）・2,000円（学生会員）、また賛助会員は1口50,000円となります。



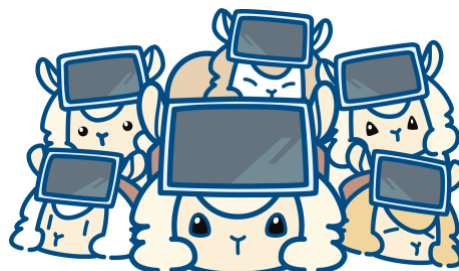
アンケートへのご協力をお願い

このたびは、日本バイオインフォマティクス学会（JSBi）の公募研究会にご参加いただきありがとうございました。つきましては、簡単なウェブアンケート（<https://www.jsbi.org/activity/koubo/questionnaires/>）にご記入のうえ、ご意見・ご感想をお聞かせくださいますようお願いいたします。

携帯からもご回答いただけます



ウェブアンケートページ



バイオインフォひつじ（案内人）