



質量分析インフォマティクス研究会・第10回ワークショップ（2025年）

生物の複雑さに挑む質量分析インフォマティクス

開催概要

開催日時： 2025 年 8 月 1 日（金）
10 時 30 分 ～ 18 時 00 分（10 時開場）

開催場所： 東京農工大学・小金井キャンパス ・工学部講義棟L0026教室
〒184-8588 東京都小金井市中町2-24-16

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/>)



※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	10:00
開会挨拶	10:30～
講習会	10:40～
シン・MassBankデータリポジトリ利用講習会	
高橋 悠志	(新潟大学)
奥田 修二郎	(新潟大学)
休憩	12:00～
招待講演	13:30～
質量分析インフォマティクスの視点からエンリッチメント解析を考える	
山本 博之	(HMT)
招待講演	14:10～
融合プロテオミクスによるがん幹細胞の分化ダイナミクスと治療標的の解析	
荒木 令江	(熊本大学)
ミキサー	15:00～
ポスター発表コアタイム	15:00～16:30
閉会挨拶・ポスター表彰	17:30～
閉場	18:00



講演・ポスター発表要旨

招待講演

質量分析インフォマティクスの視点からエンリッチメント解析を考える

○山本博之

(ヒューマン・メタボローム・テクノロジーズ株式会社)

Keywords: メタボロミクス, エンリッチメント解析

メタボロームデータの統計解析では、多変量解析によりデータを可視化し、興味のある表現型(群間差や経時変化など)と関連するスコアを抽出する。その後、ローディングもしくは群間比較により統計的に有意な代謝物を選択し、統計的に有意な代謝物がどの代謝パスウェイに多く含まれるのかを調べるために、エンリッチメント解析が用いられる[1]。

エンリッチメント解析の中でも最も単純でかつ、最も広く用いられているOver-Representation Analysis (ORA) [2] は、2×2のクロス集計表に基づいて、カイ二乗検定もしくはフィッシャーの正確確率検定を行い、代謝物と代謝パスウェイの関連を統計的に評価する方法である。ORAは、特に遺伝子発現データを用いたGene Ontologyの解析で広く用いられており、メタボロミクスでもMetaboAnalyst (<https://www.metaboanalyst.ca/>)を始めとする様々なアプリケーションに実装されている。近年メタボロミクスにおいてORAを用いたエンリッチメント解析が利用されるようになってきたが、メタボロームデータ特有の問題として、「検出されない代謝物」が多数存在し、これらの物質の扱いについては、現時点では十分に検討されているとは言えない。

例えばMetaboAnalystでは、「検出されない代謝物」がすべて“有意でない”とみなされることで、 p 値が過剰に小さく見積もられる傾向がある。この問題を回避するために最も簡単な方法は、検出された物質の情報のみを用いて解析を行う[3]ことであるが、本発表ではこれに加えて、検出されない代謝物を考慮したORA [4]について紹介する。本発表で紹介した方法は全て、Rのmseapcaパッケージ(<https://cran.r-project.org/web/packages/mseapca>)を利用して計算することが出来る。

参考文献

- [1] Yamamoto H. et al., BMC Bioinformatics (2014) 15(1):51.
- [2] Tavazoie S et al., Nature Genetics (1999) 22(3):281–5.
- [3] 山本博之、”メタボロミクス実践ガイド”、羊土社 (2021)
- [4] 山本博之、Bio"Pack"athon (2025) [[動画](#) , [スライド](#)]

招待講演

融合プロテオミクスによるがん幹細胞の分化ダイナミクスと治療標的の解析

荒木 令江

(熊本大学大学院生命科学研究部)

Keywords: Integrated Proteomics, Cancer Stem Cells, Glioma, Drug Resistance, Signal Transduction

がんは、日本人の約2分の1が罹患し、死因は男性の約4分の1、女性の6分の1を占める。死因の大部分は抗がん剤や放射線治療に対する耐性、再発・転移であり、これらはがん治療において最も難題とされている。現在のところ、ほとんどの悪性腫瘍において予後予測を高い感度と特異度で早期に診断できるマーカーや治療標的は非常に限定的であり、これらの特定と効果的な薬剤の開発は重要な課題とされている。近年、化学療法抵抗性や再発の根本原因としてがん幹細胞(CSC)が注目され、CSCとその微小環境の鍵分子を治療標的にする事によって画期的な治療法の開発が期待されているが、CSCモデルが容易でないことや技術的な問題点などから、CSC機能分子情報は非常に乏しい。

我々は、CSCの特性とがん細胞への分化誘導のスイッチングに関わる分子変動ダイナミクスを理解し、治療マーカーや標的となり得るCSC特異的な重要因子(鍵分子)を同定する目的で、特に難治性の悪性gliomaを中心に解析を進めている。様々ながん組織や樹立がん細胞から、長期の培養が可能なCSCを単離樹立する方法論を確立し、がん幹細胞(CSC-SP)、がん幹細胞からの分化誘導細胞(がん細胞前駆体: CSC-DIF)、成熟がん細胞(Parental: Par)の3相からなるがん幹細胞の分化誘導モデルの作成に成功したことから、CSC分化に関連して発現上昇及び減少する分子群を経時的に解析して大規模かつ詳細なデータを得ることが可能となった。解析には独自の統合ツールiPEACH/MANGO等を用いて、トランスクリプトーム-プロテオーム-メタボロームの融合データに加えて、リン酸化修飾分子と責任酵素の定量的同定とマッピングを行い、これらの関わるクラスターとネットワークの動的な経時変化をラージスケールで解析したところ、脂肪代謝や糖代謝に関連する転写因子を中心としたネットワークを優位に検出した。本研究会では、これらから得られた分子情報を介して、化学療法耐性・再発・転移の根源となる「がん幹細胞」に対する標的候補分子シグナルの探索を行った解析例を中心に紹介する。

参考論文

- 1) Hirayama, M. Kobayashi D et al. Mol Cell Proteomics 2013;12: 1377-1394
- 2) Kobayashi D, et al. Mol Cell Proteomics. 2019;18:245-262.
- 3) Niibori-Nambu A, et al. J. Biol. Chem. 2024; 300(3):105706
- 4) Putthisen S, et al. Exp Cell Res. 2022;410:112949.
- 5) Panawan O, et al. Cancer Science. 2023;114:3230-3246.
- 6) Araki N, Proteome Letters. 2024 ; 9 : 47-63

ポスター 1

中央値コンセンサス埋め込みによる非線形次元削減における諸問題への対処

○塘由惟¹, 米岡大輔¹

(国立健康危機管理研究機構 国立感染症研究所¹)

Keywords: 幾何中央値, 次元削減, 結果の統合

t-distributed stochastic neighbor embedding (t-SNE) やuniform manifold approximation and projection (UMAP) をはじめとする非線形次元削減法は、オミクスデータ解析において高次元データの可視化のために頻用される。しかし、これらの手法による低次元埋め込みの結果は最適化プロセスの初期値に大きく依存する。そこで我々は、複数の埋め込み結果を幾何中央値によって統合する方法として中央値コンセンサス埋め込み (Median Consensus Embedding; MCE) を開発し、MCEによって複数の初期値を用いた埋め込み結果を統合することで安定した結果が得られることを明らかにした [1]。本発表では、MCEの概要と埋め込みの一致性に関する理論解析結果について紹介したのち、MCEによる、非線形次元削減法の (1)初期値への依存性の対処, (2) ハイパーパラメータへの依存性の対処, (3) 多重補完法との組み合わせによる欠損値の対処について発表する。

[1] Tomo, Y., & Yoneoka, D. (2025). Median Consensus Embedding for Dimensionality Reduction. arXiv preprint arXiv:2503.08103.

ポスター 2

既知種の系統関係を活用した未知種のタンパク質アミノ酸配列生成法の研究

○三浦信明^{1,2}, 石野公基², 吉沢明康², 田畑剛³, 石濱泰³, 奥田修二郎²
(富山国際大学¹, 新潟大学医学部², 京都大学大学院薬学研究科³)

Keywords: メタプロテオミクス, 質量分析, アミノ酸配列, 系統関係, 生成系AI

質量分析を用いたメタプロテオミクスは、腸内細菌などに含まれる生物種とタンパク質を同時にプロファイルできる強力な手法である[1]. Metagenomics of the Human Intestinal Tract (MetaHIT)[2]や Human Microbiome Project(HMP)[3]をはじめとして腸内細菌のためのタンパク質アミノ酸配列データベース構築が盛んにおこなわれてきた。しかし、腸内細菌は種と個体間の多様性に対してデータベースの多様性が追い付いていない。

我々は、この問題を克服するために、既知種のタンパク質アミノ酸配列（以降配列と表記）の配列間にある系統的關係を意識しながら未知種のタンパク質アミノ酸配列を生成する方法を開発している[4]. 配列データベースを補強することでPeptide Spectrum Matches (PSM)の増加を目指している。我々のランダムブランチ法による配列生成と、プロテオミクスへの応用の検証として *Helicobacter pylori* F16株を題材として検証する。現在一般化に向けて生成系AIを用いる検討も進めている。

[1] Comput Struct Biotechnol J. 2023;21:1140-1150.

[2] Nature. 2010;464:59-65.

[3] Nature. 2012;486:215-221.

[4] Comput Struct Biotechnol J. 2025;27: 2313-2322

ポスター 3

異なる実験条件で得られたプロテオームデータの統合ネットワーク解析とエンリッチメント解析○西崎愛花¹, 荒木令江², 河野信¹(北里大学未来工学研究科¹, 熊本大学大学院生命科学研究部(医)²)**Keywords:** プロテオーム, ネットワーク解析, quantMS, CPTAC

プロテオームの解析を行う際には一つのデータのみならず、異なる施設や異なるサンプルから得られたデータを統合して比較できると有用である。しかしながら、これらの異なる実験条件で得られたデータを単純に比較することは難しい。メタボロームの分野では異なる実験条件のデータを、ネットワークを作成して解析する試みがなされている (Matsuta, R., et al. BMC Bioinformatics 23, 508 (2022).)。そこで、今回の研究では異なる実験条件で得られたプロテオームデータを統合してネットワークを作成し解析を行った。

PRIDE の生データをquantMS (Dai, C., et al. Nature Methods 21, 1603–1607 (2024).) ワークフローで再解析したデータを利用してネットワークを作成した。今回はラベル化定量のデータで実験群間で比率が計算されている 36 プロジェクトのデータを使用した。タンパク質変化量が2倍以上かつp値が0.01以下のデータを有意に発現変動したタンパク質であると見なし、フィルタリングを行った。実験群間のタンパク質変化量のクロス集計表を作成し、オッズ比ならびにカイ二乗検定によるp値を算出して、有意に相関した異なるプロジェクト由来の実験群の組み合わせから、cytoscapeを用いてネットワークを構築した。最終的に41ノードと77本のエッジから構成されるネットワークを再構築した。

タンパク質の発現量比をもとにクラスター解析を行った結果、タンパク質は4個のグループに分類された。これらグループのエンリッチメント解析を行った結果、がん細胞をDoxorubicinで処理した実験群では、細胞骨格・シグナル伝達に関するタンパク質の発現上昇が観察された。がん細胞をDoxorubicin処理するとアクチン関連タンパク質が増加することが知られており (梅村ら 日薬理誌156,146～151 (2021))、この結果を再現できたため、本手法は有効であると考えられる。

次に、CPTAC (Edwards NJ., et al. Journal of Proteome Research 14, 2707-2713 (2015)) からラベル化定量された49プロジェクトのデータを取得して、

quantMSのデータと統合して解析を行った。その際、コントロールと比較を行っている実験群のみを解析の対象とした。解析の結果、今回得られたネットワークは58ノード、エッジが81本で構成



されていた。本発表では、quantMSとCPTACのデータを統合して形成されたネットワークから見えてきたプロテオーム実験群間での関係性について明らかになったことについて紹介する。

ML-guided mass defect-based glycopeptide classifier for next-gen glycoproteomics

○Bingyuan Zhang¹, The Huong Chau², Rebeca Kawahara¹, Yusuke Matsui¹, Morten Thaysen-Andersen^{1,2}
(Nagoya University, iGCORE¹, Macquarie University²)

Keywords: Machine learning-based precursor selection, Glycoproteomics

Glycans, attached to proteins or lipids, play pivotal roles in many biological processes, including cell–cell communication, immune response, and protein folding. Among other glycan types, N-glycosylation is a prevalent and biologically significant post-translational modification. N-glycosylation change has been implicated in numerous pathological conditions such as cancer, autoimmune diseases, and neurodegenerative disorders, underscoring its importance as both biomarker candidates and potential therapeutic targets in biomedical research.

Challenges in Glycoproteomics

Quantitative glycoproteomics aims to comprehensively profile glycopeptides. A typical workflow involves several steps: digestion of proteins, selective enrichment of glycopeptides, and high-throughput analysis via liquid chromatography–tandem mass spectrometry (LC-MS/MS). However, one of the major technical bottlenecks is the inefficient enrichment of glycopeptides at the sample preparation stage and their inaccurate selection for MS/MS fragmentation in conventional intensity-based data-dependent acquisition (DDA) strategies. As a result, glycopeptides are often underrepresented or completely missed, limiting the depth and accuracy of glycoproteomic profiling.

Proposed Machine Learning-Based Solution

To address this limitation, we propose the integration of a informed machine learning (ML)-based precursor selection strategy into the DDA framework. By leveraging our access to large volumes of well annotated and diverse glycoproteomics datasets, we trained a predictive glycopeptide classifier capable of promptly distinguishing N-glycosylated peptides from their non-glycosylated counterparts based exclusively on accurate molecular mass (MS1) information. Conceptually, the classifier harness the unique mass defect (MD) signatures inherent to the oxygen-rich N-glycopeptides, which differ from the MD profile of non-glycosylated peptides featuring a lower proportion of oxygen. To this end, we compiled a series of high quality glycoproteomics datasets consisting of approximately 2 million peptide-spectrum matches (PSMs), including both N-glycosylated (~10%) and non-glycosylated peptides (~90%), derived from a broad range of human and mouse tissues. Due to its balance between interpretability and computational efficiency, a decision tree-based model was trained on ~70% of the curated data. When tested on the remaining 30% of glycoproteomics data, the model demonstrated a strong discriminative power to distinguish N-glycopeptides from non-glycosylated peptides even within realistic non-enriched proteomics datasets (AUC >0.95) documenting the accuracy and potential of the glycopeptide classifier.

Advantages and Future Integration

Compared to previous approaches, our model not only achieved higher sensitivity and specificity but also demonstrated high performance across a broad mass range of glycopeptides. Additionally, the ML-aided glycopeptide classifier featured swift prediction (nanosecond scale decision making) ensuring compatibility with real-time data acquisition pipelines. Our innovative informatics solution opens for intelligent DDA-



based data acquisition strategies that promise to improve the coverage and reproducibility of the glycoproteome analysis without the need for glycopeptide enrichment. This innovation paves the way for the next-generation of clinic-ready glycoproteomic workflows.

ポスター 5

時空間メタボローム解析に資する多様なモダリティ画像の統合解析基盤の構築

○坂本七海¹, 露崎弘毅^{2,3}, 内野春希⁴, 岡昂輝¹, 山元らな¹, 清水邦善¹, 松沢佑紀¹, 川上英良^{2,4,5},
有田誠^{4,6,7}, 津川裕司^{1,4}

(東京農工大学¹, 千葉大学², 理化学研究所生命機能科学研究センター³, 理化学研究所生命医科学研究センター⁴, 理化学研究所数理創造研究センター⁵, 横浜市立大学⁶, 慶應義塾大学⁷)

Keywords: インフォマティクス, 空間オミクス, 統合解析

近年、空間メタボロミクス・リポドミクス・トランスクリプトミクスの技術開発が進められている背景から、異種データの統合解析基盤の開発が望まれている。これは、単一の解析だけでは把握できない生体内の複雑な分子機構を、異なる分子レベルの情報を組み合わせることにより詳細に理解できるからである。そこで本研究では、最適輸送と呼ばれる機械学習モデルを用いることで、マウス脳切片に対する空間リポドミクスとトランスクリプトミクスデータの統合手法を提案する。

データ統合の要素技術として、空間オミクスデータの発現パターンから切片の概形を抽出し、それに対してAllen Brain Atlas における組織学情報を変形・位置合わせすることで、各ピクセルに脳組織ラベルを付与する手法を開発した。これを基に、ピクセルの座標情報と組織ラベル情報に基づき異種データの輸送コストを算出することで、異種データ間の対応関係を明らかにするための最適輸送を実現する解析ワークフローを構築した。

本解析基盤を用いて、マウス脳における脂質および遺伝子発現情報の統合解析を実施したところ、ニューロンやオリゴデンドロサイトに特異的な脂質分子としてドコサヘキサエン酸を含むリン脂質や糖脂質分子を抽出することができた。また、脂質発現量を説明する遺伝子発現情報を空間トランスクリプトミクスやシングルセルRNA-seq情報と組み合わせることで、脂質の局所発現情報を説明する遺伝子発現情報を抽出した。本発表では、液体クロマトグラフィー質量分析により取得した組織・細胞種別のリポドミクスデータとの比較を通じて、本手法の妥当性および有用性について考察する。

ポスター 6

ローディングの解釈を可能とする交絡因子を考慮した主成分分析の理論実装

○倉田明咲¹, 山本博之², 津川裕司¹

(東京農工大学¹, ヒューマン・メタボローム・テクノロジーズ株式会社²)

Keywords : 多変量解析, 次元削減, リピドミクス

オミクスデータのような多次元データの解析では、次元削減手法によって生物学的な解釈を実施することが多い。中でも主成分分析 (PCA) は、主成分得点 (スコア) に対応付けられたローディングの値からメカニズムの考察が可能であり、オミクス科学で頻用されている。一方、生物試料の背景因子 (交絡因子) の複雑性が増した場合に、その交絡因子が生物学的な解釈を困難にさせる場合がある。本問題を解決するために、データの持つ分散から交絡因子による分散を排除した値の最大化問題を解く手法であるadjustment for confounding PCA (AC-PCA) [1]が開発された。一方、本法で用いられる一般化固有値問題は固有値ベクトルの直交性が担保されず、ローディングの統計学的解釈ができないといった課題があった。そこで本研究ではAC-PCAの理論を再考し、標準固有値問題に帰着させることで、ローディングの有意差検定が可能なorthogonal AC-PCA (OAC-PCA)の理論実装に取り組んだ。本発表では、実装されたOAC-PCAについてシミュレーションデータ、および加齢変容リピドミクス公共データ (交絡因子として雌雄差や腸内細菌の有無を定義) を用いた検証結果について報告する。

[1] *PNAS*, **2016**, 113 (51), 14662-14667

ポスター 7

Transformerを用いたオートエンコーダに基づく構造とMS/MS特徴抽出による代謝物アノテーションの実現

○小川誠寛¹, 津川裕司¹
(東京農工大学¹)

Keywords: 質量分析, インフォマティクス, エンベディング, Transformer

昨今の質量分析に基づいたスペクトルアノテーションでは、標準品が利用できる代謝物数が限られており、得られるスペクトルに対して10%程度しかアノテーションできないのが現状である。そこで本研究では、標準品に依存しないスペクトルデータ解析の方向性の1つとして、機械学習に基づく代謝物アノテーション手法の開発に従事した。具体的には、Transformerを用いたオートエンコーダによって化合物構造とマススペクトルの両者を同時に学習し、共通の化合物潜在空間を構築することで、未知スペクトルから化合物の構造情報を推測する解析基盤の構築を目指した。本基盤を用いて、約1,200万の化合物と15万のマススペクトルを学習し、マススペクトルから化合物構造を推定するモデル構築を行なった。その結果、学習済みモデルによってマススペクトルから推定された化学構造と組成式情報の類似度に基づき、データベース内に存在する化合物構造の中から正解構造が効率的に抽出できることがわかった。本発表では、本基盤の現状の精度および潜在空間に埋め込まれた情報の応用可能性について主に紹介する。

ポスター 8

Activity-based protein profilingを用いた甘草由来天然物イソリクイリチゲニンの 標的探索および作用機序の解明

○酒井陽菜¹, 鄭美和¹, 磯部洋輔², 津曲和哉², 今見考志², 広川貴次³, 大淵勝也⁴, 中谷泰貴⁴, 津川裕司^{1,2}
(東京農工大学¹, 理化学研究所生命医科学研究センター², 筑波大学³, 株式会社ツムラ⁴)

Keywords: プロテオミクス、LC-MS/MS、抗炎症作用、標的探索

生薬甘草は抗炎症をはじめとする様々な薬理作用があり、漢方製剤の約7割に用いられる重要な植物資源である。甘草に含まれる300種類以上の天然物の中でも、イソリクイリチゲニン (ILG) は強い抗炎症活性を持ち、これまでの研究において、その抗炎症効果および生体内の代謝変動に特異的な変化を示すことが明らかである一方で、ILGに起因する代謝応答の詳細な分子メカニズムは未解明である。そこで本研究では、ILGの標的分子をプロテオームワイドに探索し、抗炎症作用機序を理解することを目指した。ILGがタンパク質のシステイン (Cys) 残基と求核反応を介した共有結合を形成するという仮説のもと、Activity Based Protein Profiling (ABPP) 法を用い、ILGと反応性の高いCys残基を持つタンパク質を網羅的に探索した。その結果、ILGの標的分子候補としてlipocalin-type prostaglandin D₂ synthase (L-PGDS) のCys65 が得られた。ILG添加有無におけるL-PGDSの発現量に差は無く、L-PGDS阻害剤投与でも部分的に抗炎症効果が認められた。本ワークショップで、その詳細を発表する。

ポスター 9

固相抽出及びElectron activated dissociation MS/MSを用いた超長鎖多価不飽和脂肪酸含有ホスファチジルコリン生合成機構の解析

○林田怜大¹, 清水邦善¹, 原山武士², 津川裕司¹

(東京農工大学¹, フランス国立科学研究センター 分子細胞薬理研究所²)

Keywords: LC-MS/MS, 固相抽出, 超長鎖多価不飽和脂肪酸 (VLC-PUFA), ゲノム編集

ヒトの生体内には数万種類の脂質分子が存在すると言われており、構造異性体の識別は機能理解に不可欠である。我々の研究グループでは近年、脂質の極性基・疎水性側鎖の物性に基づいて分画・濃縮する固相精製手法を開発し、electron activated dissociation (EAD) によるMS/MS法と組み合わせ生体内の微量な脂質分子の詳細な構造解析が可能な基盤を構築した。本手法を、視細胞に特異的に存在する超長鎖多価不飽和脂肪酸含有ホスファチジルコリン (VLC-PUFA PC) の構造解析に適用したところ、VLC-PUFAがグリセロール骨格のSN1位に選択的に取り込まれることが示された。我々は現在、VLC-PUFA PCの生合成経路の解明を目指し研究を行っている。本発表では、脂質代謝物の添加やHeLa細胞内での遺伝子ノックアウトや過剰発現系によって得たVLC-PUFA PCの生合成機構に関するこれまでの解析結果について報告する。

ポスター 10

CO₂資化バイオ生産プラットフォーム構築に向けたギ酸資化大腸菌のメタボリックエンジニアリング

○林 大生¹, 津川 裕司¹
(東京農工大学¹)

Keywords: 合成生物学, 代謝工学

世界的なエネルギー・環境問題の深刻化に伴い、CO₂を原料とする持続可能な物質生産技術が求められている。光合成を基盤とするプロセスではCO₂固定の効率や食糧競合などの課題があり、より効率的なCO₂固定プロセスが求められている。電気化学的なCO₂変換も検討されており、ギ酸などC1化合物への変換において高効率なプロセスが構築されている。一方、C-C結合の形成を伴う反応においてはバイオ触媒に優位性がある。

本研究では、電気化学的プロセスとバイオ触媒を融合した新規CO₂変換プロセスの構築を目指す。我々は人工ギ酸資化経路であるReductive Glycine Pathway を大腸菌において構築し、この株と電気培養系を組み合わせることで、CO₂からセリンを生成できることを見出した[1]。

現在、バイオ生産のプラットフォーム化を目標に、RGPの代謝フラックスおよび株のユーザビリティの改良を行っている。本発表では、これまでの進捗と共に、メタボロミクスを取り入れ代謝工学を加速させる方法論について議論したい。

[1] Y. Tashiro, et al. Metabolic engineering 47. 2018

RNA質量分析のための解析支援ツールの開発

○松田哲平¹, 平田章², 山上龍太¹

(愛媛大学大学院理工学研究科¹, 徳島大学大学院社会産業理工学研究部²)

Keywords: LC/MS, RNA, 核酸, エピトランスクリプトミクス

生命現象を支える遺伝子発現機構は、DNAからRNA、さらにタンパク質へと至る過程において精緻な制御を受けている。特にRNA修飾は、転写後にRNAの性質を変化させることで、遺伝情報そのものを改変することなく、遺伝子発現を時空間的に制御する役割を担っており、近年「エピトランスクリプトミクス」として生命科学の新たな潮流を形成している。現在までに150種類以上のRNA修飾が報告されており、依然として新規修飾の発見が続いている(Cappannini A *et al.*, *Nucleic Acids Res.*, 2024)。

RNA質量分析は、RNA修飾を高感度かつ高精度に同定するための強力な手法であるが、従来の解析では多くの場合、専門的な知識に依存した手動によるスペクトル解釈に基づいており、煩雑で時間のかかる作業が課題となっている。

本研究では、これらの課題を解決するため、核酸配列からの質量電荷比の計算、得られた質量電荷比および電荷数に基づく配列予測、さらに酵素処理による断片化の予測などを行うことができるGUI付き解析支援ツールを開発した。現在は、このツールを基盤として、RNA質量分析の自動解析パイプラインの構築を進めている。

ポスター 12

脳特異的な糖脂質代謝の生合成および加齢変容の分子機序の解明

○清水邦善¹, 竹田浩章², 坂本七海¹, 原山武士³, 津川裕司¹

(東京農工大学¹, 産業技術総合研究所², フランス国立科学研究センター 分子細胞薬理研究所³)

Keywords: UGT8, 脳, オリゴデンドロサイト

哺乳類の脳には多様な糖脂質代謝物が存在し、その代謝恒常性維持機構の破綻がさまざまな疾患の原因となる。我々はリポミクス技術を開発する過程において、脳特異的に発現するアシル化糖セラミド (Acylated Hexosyl Ceramide, AHexCer) をマウス脳で初めて発見した。そこで我々は、脳においてAHexCerがどのように糖脂質の代謝恒常性に関係しているかを調べるため、AHexCerの生合成に寄与する責任細胞や代謝経路の同定を目指して研究を推進している。

まず、脳における大規模なリポミクス・トランスクリプトミクスデータの再解析によって、AHexCerは脳幹部に濃縮され、加齢に伴ってその発現量が大きく変動することや、AHexCerの合成に関与すると考えられる遺伝子がオリゴデンドロサイトに高発現していることが明らかになった。そこでHEK293細胞に対して、オリゴデンドロサイトに高発現する*FA2H*や*UGT8*の過剰発現細胞を樹立したところ、これらの細胞がAHexCerを合成することを見出した。本発表では、これら過剰発現細胞に対して阻害剤・ノックアウト・ノックダウンなどによって脂質代謝に介入することで得た糖脂質代謝に対する我々の知見について発表する。

同位体標識とMS/MSネットワークによる甘草代謝物の構造解析

○澤井敬太¹, 轟善勝¹, 松沢佑紀¹, 中迎菖平², 森哲哉², Amit Rai³, 津川裕司^{1,2}

(東京農工大学¹, 理化学研究所環境資源科学研究センター², University of Illinois Urbana Champaign³)

Keywords: 植物特化代謝物, 質量分析, 分子ネットワーク

甘草は多くの薬理活性に関わる植物特化代謝物を持ち、漢方製剤の約7割に配合されている汎用的な生薬である。甘草は種間や部位間で成分が異なり、網羅的に代謝物を捉えることは甘草の品質向上や新たな有効成分発見につながると考えられる。甘草に属する2種 (*Glycyrrhiza glabra*, *G. uralensis*) の葉および根の抽出物をLC-MS/MSに供し、得られたデータをMS-DIALを用いて解析を行った。検出イオンに対し1724394スペクトルを含むライブラリ検索を行ったところ85%のイオンが未同定であった。本研究では安定同位体二酸化炭素標識を施した植物と非標識植物の質量差を比較し、3448イオンに対して炭素数および組成式推定を行った。その後、ライブラリと検出イオン間のMS/MS類似度を算出し、分子ネットワークを作成した。ネットワークを構成する化合物情報をもとに未同定イオンの構造を推定したところ、新たに186イオンの構造予測に成功し、全体で32%のイオンに対して構造情報付与が行われた。推定構造の中には窒素を有する新規構造が根特異的に存在しており、新規構造の決定を目指し甘草から目的代謝物の精製・抽出を行った。また他の植物生合成経路をもとに新規構造の生合成経路が予測され、新規構造はリジン由来のアルカロイドであると推測された。今後は予測経路から関連酵素や代謝反応について新たな知見を取得し、目的代謝物の生物学的意義の探索を行う。

ポスター 14

食餌栄養および加齢によるマウス組織脂質代謝変容の包括的解析

○時吉花菜子¹, 岡田采奈¹, 宮本潤基¹, 津川裕司¹
(東京農工大学¹)

Keywords: 脂質代謝, リピドミクス, 食餌栄養, 加齢

加齢に伴う代謝変容を捉えることは加齢関連疾患の分子機構を解明する上で重要である。本研究では、ノンターゲットリピドミクスの技術開発を通じ、月齢および食餌条件の異なるマウス組織に対して代謝変容解析を行った。大規模計測には分析の高速化と効率的なデータ解析基盤構築が必須であるため、data dependent / independent acquisitionを統合して利用する新たな手法を考案し、それらのデータを統合解析する情報基盤を構築することで、従来法の3分の1の時間で脂質多様性を捉える手法の開発に成功した (Tokiyoshi K. et al. Analytical Chemistry 2024)。

本手法を357検体の解析に供し、多変量解析を実施したところ、3,330の脂質代謝物の定性・定量値を得た。食餌および加齢の両条件で変動が大きかった臓器の1つである腎臓では、エーテル型脂質およびスフィンゴ脂質において高脂肪食 (HFD) 負荷への応答が加齢に伴い異なるという変動を捉えた。さらに、HFD群における Hex3Cerの蓄積は糖尿病性腎症と関連する報告がある一方、老齢マウスではHex3Cerの代謝応答が減弱していることが本研究で初めて明らかになった。本発表では、Hex3Cerの変動と腎臓における炎症応答との関係性に焦点を当て、研究成果を報告する。

ポスター 15

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

(東京農工大学¹, 九州大学², 大阪大学³)

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、質量分析法を利用した低分子化合物の解析を目的とし、組成式・化合物クラス・完全構造を推定する機械学習モデルの構築が盛んに行われている。一方で、2022年に開催されたスペクトル同定コンテスト (CASMI2022) の構造推定部門において、正答率が30%を上回るモデルは存在せず、その推定精度は十分ではないと言える。そこで我々は、複数の構造推定 (アノテーション) プログラムの予測結果のアンサンブル学習基盤を構築することで、より高精度の構造推定を行う基盤構築を目指し研究開発を行っている。

アンサンブル学習の適用対象として、組成式推定にはMS-FINDER、SIRIUS、msbuddyを、構造推定にはMS-FINDER、SIRIUS、MetFragを用いることとした。学習モデルとしてRandom Forest、Logistic Regression、Support Vector Machine (SVM)、XGBoostを選択、約16,000化合物の情報を学習データとして使用し、組成式推定と構造推定問題それぞれについてモデル最適化を行った。CASMI2022を用いた精度評価において、構築されたモデルでは単一プログラムよりも高い精度で組成式・構造推定結果を出力することができた。不正解であった化合物に関しても、化合物記述子とTanimoto indexを用いて正解構造との類似度を評価したところ、統合モデルが正解構造と類似した構造を出力していることも示された。本プログラムは、Github (<https://github.com/systemsomicslab/msemblator>) で公開している。

ポスター 16

質量分析インフォマティクス技術を用いたアマモとイネの代謝物プロファイルの解明

○轟善勝¹, 松沢佑紀¹, 中迎菖平², Wei He³, 竹澤浩気³, 藤田誠³, 野口恵一¹, 加藤敏代¹,
笠原博幸^{1,2}, 大津直子¹, 梅澤有¹, 津川裕司^{1,2}

(東京農工大学¹, 理化学研究所環境資源科学研究センター², 東京大学³)

Keywords: 質量分析, 保持時間予測, 分子スペクトルネットワーク, NMR, X線結晶構造解析

植物はその固着性のために様々な環境ストレスにさらされるが、独自の代謝プロファイルを発達させることで環境に適応してきた。海草のアマモ (*Zostera marina*) は約1億年前に陸生祖先から海洋環境へ再適応を遂げた種であり、その代謝物多様性の解明は海洋適応機構の理解の深化に繋がる。本研究では質量分析インフォマティクス技術を用いることで陸上植物で同じく単子葉類に属するイネ (*Oryza sativa*) とアマモの代謝物比較解析を実施した。

予測保持時間を追加したライブラリーアノテーションと分子スペクトルネットワークを利用することで最終的に合計139個の親水性代謝物に構造アノテーションが付与された。内38個のフェニルプロパノイド代謝物について、イネではtricin派生物とC-リンクフラボノイドが、アマモでは5種類の母骨格を持つ硫酸化フラボノイドが特異的な代謝物として検出できた。さらにアマモでは地下茎に豊富に存在する独自のフェノール化合物が検出されており、NMRとX線結晶構造解析を実施することでその構造が7'',8''-didehydrosalvianolic acid Bであると同定できた。

ポスター 17

マルチモーダルなリピドーム情報解析基盤の構築によるマウス脳の脂質代謝解析

○山元らな¹, 坂本七海¹, 清水邦善¹, 杉浦悠毅², 本田真也³, 宮本潤基¹, 林青順⁴, 今見考志⁴, 津川裕司^{1,4}
(東京農工大学¹, 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², 東京科学大学³, 理化学研究所生命医科学研究センター⁴)

Keywords: 質量分析、リピドミクス

昨今、質量分析 (MS) を用いた空間リピドミクス (リピドーム解析) 手法は生命科学分野で頻用されている一方、アノテーションや定量性に課題が多く残されている。また、得られるリピドーム情報がどのような細胞に由来するのか推定する基盤が存在しないことから、シングルセルRNA-seqや空間トランスクリプトミクスデータとの統合解析が困難となる。そこで本研究では、マウス脳をベンチマークとして、(1) マトリックス支援レーザー脱離イオン化 (MALDI) 法を用いた高分解能空間リピドミクス (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) を用いた低分解能空間リピドミクス、および (3) 細胞分画による細胞タイプ別リピドミクスのマルチモーダルデータを取得する基盤構築を行う。また、数理モデルや多変量解析手法を組み合わせることで各モダリティの長所を強化し、短所を補うバイオインフォマティクスパイプラインの構築も同時に行う。本発表では、マウス脳を用いた各モダリティのデータ取得に資する材料選定・前処理・分析手法についてこれまで行った検討結果を発表する。

ポスター 18

腎臓特異的なアシル化スフィンゴミエリンの生合成研究

○廣岡佑美¹, 清水邦善¹, 岡昂輝¹, 今見考志², 原山武士³, 津川裕司^{1,2}

(東京農工大学¹ 理化学研究所生命医科学研究センター², フランス国立科学研究センター 分子細胞薬理研究所³)

Keywords: リピドミクス, 質量分析, スフィンゴミエリン

本研究では、腎臓特異的な脂質であるアシル化スフィンゴミエリン(acylated sphingomyelin, ASM)の生合成経路および機能の解明を行っている。ASM は 1996 年に初めて報告されて以降¹、その詳細な生合成経路や代謝機構は明らかにされていない。しかし、近年のリピドミクス解析技術の発展により我々は以前、マウスから初めて ASM を同定した。また、腎臓由来の細胞株である KMRC-1 細胞をリピドミクスに供したところ ASM が発現していることを見出した。一方、ASM は HeLa や HEK293 などの一般的な細胞株ではほとんど検出されないこともわかった。本発表では、KMRC-1 細胞に対して阻害剤や脂肪酸添加によって脂質代謝に介入することで得た ASM に対するこれまでの知見について紹介する。

[1] Biochim Biophys Acta, 1996, 1303(1), 47-55.

ポスター 19

大規模時系列プロテオームデータを用いたペプチドターンオーバー予測モデルの提案

○石野公基¹, 劉 鈺婷¹, 吉沢明康¹, 奥田修二郎¹
(新潟大学医学部メディカルAIセンター¹)

Keywords: プロテオーム, LLM, 質量分析

タンパク質の動的ターンオーバーは細胞恒常性の維持や疾患機構の理解に重要である。大規模質量分析技術の進展により、ペプチドレベルで高精度な時系列測定が可能になったが、分解機構の多様性や細胞環境の複雑さから、動的挙動を規定する配列レベルの特徴を包括的に明らかにするのは依然として困難である。一方、Transformer ベースのタンパク質言語モデルを含む深層学習の進歩により、配列情報から生物学的特徴を抽出する手法が注目を集めている。

本研究では、事前学習済み Transformer モデルから得られるアミノ酸埋め込みを活用し、ペプチドのターンオーバー動態を予測するための2種類の時系列モデルを提案した。(1) TimeSeq は残基ごとの埋め込みを平均化し、LSTM を用いて全時系列を一括で予測する構成、(2) AASeq は残基レベルの埋め込みを保持したまま各時点で逐次的に予測を行う構成である。いずれのモデルも、大規模な公開データセットを用いた RMSE および R^2 による評価において、従来手法を大きく上回る予測精度と安定性を示した。さらに、残基ごとの寄与度を可視化する解析を行い、モデル予測に寄与する特徴的なアミノ酸パターンを調査した。これらの成果は、配列ベースの表現と時系列予測の統合が、ペプチドの動的性質を捉える上で有効であることを示しており、今後のプロテオミクス解析における新たなアプローチとしての可能性を示唆している。

ポスター 20

Exploring Lipid Biomarkers in Melanoma by Non-invasive Skin Sampling and Untargeted Lipidomics

○Skaidre Jankovskaja^{1,2,3}, Kari Nielsen^{4,5}, Gustav Christensen^{4,5}, Bujinkham Buyantogtokh¹, Tautgirdas Ruzgas^{2,3} and Hiroshi Tsugawa¹

(Department of Biotechnology and Life Science, Tokyo University of Agriculture and Technology, Tokyo, Japan¹, Department of Biomedical Science, Malmö University, Sweden², Biofilms Research Centre for Biointerfaces, Malmö University, Sweden³, Department of Dermatology, Skåne University Hospital, Lund, Sweden⁴, Department of Clinical Sciences Lund, Division of Dermatology and Venereology, Lund University, Lund, Sweden⁵)

Keywords: Melanoma, Non-invasive tape sampling, Lipidomics

Melanoma, the deadliest type of skin cancer, is currently diagnosed through visual examination followed by surgical excision of the suspected lesion for pathological confirmation. This diagnostic approach is highly subjective, has low specificity, and is resource - intensive. Non-invasive methods, such as tape sampling, which can be easily performed on large number of lesions, offer a promising approach to complement current diagnostic practices. In this study, we explore lipid profiles from melanoma lesions sampled using adhesive tape to identify potential biomarkers for melanoma diagnosis. Biological samples are collected in Sweden, Malmö University and Skåne University Hospital in Lund, while mass spectrometry-based lipidomic analysis will be carried out in Japan, Tokyo University of Agriculture and Technology.

The project consists of two parts. The first part focuses on investigating lipid profiles in samples collected from healthy human volunteers, both females and males, across different age groups (20–70 years). The goal is to explore if and how lipid profiles change with age and whether these changes can be captured using tape sampling. These healthy reference samples will also serve as a baseline for comparison when analyzing melanoma patient samples. The second, and main part of the study will analyze samples collected from melanoma patients. Tape sampling on the patients is performed before surgical excision, following Swedish national guidelines for suspected melanoma lesions. From each patient, three consecutive tape strips are collected from the melanoma-suspected lesion and an additional three from adjacent healthy skin. As a preliminary data, we introduce the optimized workflow of lipidomics for tape strip samples.

The Japan Society for the Promotion of Science (JSPS) is acknowledged for postdoctoral fellowship funding. The Mats Paulsson Foundation and Knowledge Foundation are acknowledged for supporting non-invasive melanoma patient sampling and analysis.