

2023年度 研究開発実施報告

概要

研究開発課題名	jPOST prime:コミュニティ連携を基盤とするプロテオームデータベース環境の実現
開発対象データベースの名称(URL)	jPOST (Japan Proteome Standard Repository/Database) (https://jpostdb.org/)
研究代表者氏名	石濱 泰 (30439244)
所属・役職	京都大学大学院 薬学研究科 教授 (2024年3月時点)



目次

概要	1
目次	2
§1. 研究実施体制	3
§2. 研究開発対象とするデータベース・ツール等	3
(1) データベース一覧	3
(2) ツール等一覧	3
§3. 実施内容	4
(1) 本年度の研究開発計画と達成目標	4
1. リポジトリ部	7
2. 再解析部	7
3. データベース部	8
(2) 進捗状況	9
1. リポジトリ部	9
2. 再解析部	9
3. データベース部	10
○ jPOSTリポジトリ+データベース ～サーバーの現状とストレージ拡充の必要性	11
§4. 成果発表等	14
(1) 原著論文発表	14
① 論文数概要	14
② 論文詳細情報	14
(2) その他の著作物(総説、書籍など)	14
(3) 国際学会および国内学会発表	14
① 概要	14
② 招待講演	14
③ 口頭講演	15
④ ポスター発表	15
(4) 知的財産権の出願 (国内の出願件数のみ公開)	15
出願件数	15
(5) 受賞・報道等	15
① 受賞	15
② メディア報道	16
③ その他の成果発表	16
§5. 主要なデータベースの利活用状況	17
(1) アクセス数	17
①実績	17
② 分析	17
(2) データベースの利用状況を示すアクセス数以外の指標	17
(3) データベースの利活用により得られた研究成果(生命科学研究への波及効果)	18
(4) データベースの利活用によりもたらされた産業への波及効果や科学技術のイノベーション(産業や科学技術への波及効果)	18
§6. 研究開発期間中に主催した活動(ワークショップ等)	19
(1) 進捗ミーティング	19
(2) 主催したワークショップ、シンポジウム、アウトリーチ活動等	19

§1. 研究実施体制

グループ名	研究代表者または研究分担者氏名	所属機関・役職名	研究題目
石濱グループ	石濱 泰	京都大学・教授	再解析部の開発
奥田グループ	奥田修二郎	新潟大学・教授	リポジトリ部の開発
河野グループ	河野 信	北里大学・教授	データベース部の開発

§2. 研究開発対象とするデータベース・ツール等

(1) データベース一覧

【主なデータベース】

No.	名称	別称・略称	URL
1	Japan Proteome Standard Repository/ Database	jPOST	https://jpostdb.org
	Japan Proteome Standard Repository	jPOSTrepo	https://repository.jpostdb.org/
	Japan Proteome Standard Database	jPOSTdb	https://globe.jpostdb.org/

【その他のデータベース】

No.	名称	別称・略称	URL
1	jPOST repository demonstration		https://rep-demo.jpostdb.org/
2	jPOST repository development		https://rep-dev.jpostdb.org/
3	jPOST database development		https://rep-db.jpostdb.org/
4	iMPAQT		http://impaqt.jpost.org/iMPAQT/

(2) ツール等一覧

No.	名称	別称・略称	URL
1	PRESTO		https://github.com/PrestoTools/
2	iMPAQTquant		http://impaqt.jpost.org/iMPAQT/impaqtquant.html

§3. 実施内容

(1) 本年度の研究開発計画と達成目標

本研究で開発するプロテオーム統合データベース jPOST は、リポジトリ部、再解析部およびデータベース部からなる。各部の開発概要は以下のとおりである。

1. リポジトリ部の開発概要

欧米のリポジトリデータベースでは、データの登録の際に、FTP を利用することや別途ソフトウェアをインストールすることが要求されるが、jPOST リポジトリでは、ユーザーがデータの登録の際に、すべての操作をブラウザ内で完結出来るように設計されており、登録開始から終了までの一連の操作がスムーズに進行するように構築されている。実際に利用したユーザーからも、この点について非常に使いやすい、という評価を得ている。また、登録データのアップロードを高速化するための独自技術 PRESTO により、FTP を使う場合に比べ数倍以上高速にアップロード出来るため、この点もユーザーの登録作業におけるストレス軽減に役立っている。本システムはすでに糖鎖 DB チームに水平展開され GlycoPOST として運用されているが、これをメタボロームチームの MB-POST に展開する。質量分析データを対象とするこれら 3 つのデータベースに共通するオミクス MS リポジトリ UniPOST コンセプトを3つのデータベース開発プロジェクトで共有化することにより、マルチオミクス研究を行うユーザーの利便性を向上させるだけでなく、開発資金の効率的運用にもつながる。生データストレージ、ピークピッキングモジュールやメタデータ収集システムを共有・共用できる可能性があり、またマルチオミクス研究を行うユーザーにとっても、オミクス対象分子群が異なっても入り口が一つで、試料情報等のメタデータ串刺し検索ができる、というメリットは大きい。

2. 再解析部の開発概要

第 2 期 jPOST では対応しきれなかった発展著しいデータ非依存性取得 (DIA) モードや、イオンモビリティ分離による多次元化に対応した再解析法を開発する。特に MSMS スペクトルではなく、プロダクトイオンの溶出プロファイルからペプチドを同定する DIA 解析法に対応した再解析法に UniScore を適用する。また、翻訳後修飾解析について、修飾サイト同定のあいまいさに対応した HUPO-PSI における PTMeXchange プロジェクトに積極的にコミットし、再解析法の確立を目指す。また、データジャーナル Journal of Proteome Data and Methods との連携に加えて、コミュニティ連携によるメタデータ取得システムをさらに大規模展開し、若手研究者育成にも寄与するとともに、半自動化フローの実現に向けて開発を継続する。

3. データベース部の開発概要

再解析部とも連動させながら、様々なユーザーコミュニティとの連携を強化し、jPOST チームと協働でデータサイエンスを行う連携ユーザーの掘り起こしを積極的に行う。これにより、jPOST チームと連携ユーザーが一体となった双方向利活用環境 "jPOST prime" を実現させ、プロテオームデータサイエンスリソースとしての有用性を向上させる。データベース部では、すでにオントロジーによって整理されたメタデータに基づいた、生物種や組織、疾患などのカテゴリーによる絞り込み検索を入り口とし、ユーザーによる任意のデータセットの切り出しを可能としている。本データベースでは、ユーザーによりデータが自由に切り出される。このような可変情報を表示するために、新たに開発したプロテイン・ブラウザでは、同定されたペプチドのマッピングデータ・リン酸化等翻訳後修飾データ・修飾の共起データ・UniProt における既知データなどの表示・非表示を選択することで、必要なデータのみを可視化することが可能になっている。また、プロテオフォーム間に於ける複雑な共有ペプチドの関係性を表示するためのビューワーを実装することで、アイソフォームを含めたタンパク質ファミリーにおける各配列の発現の推定に利用することも可能である。さらに、2つの切り出したデータセット間を比較するためのインターフェースを準備したことで、容易にタンパク質の発現変動を調べることができる。これらはいずれも、独自性の極めて高いインターフェースである。これら各データの可視化は、DBCLS で開発され

た Java script 版 TogoStanza を用いて開発を進めている。今期も引き続き DBCLS の協力を得ながら、ユーザーの利便性の高いデータベース開発を継続する。

第 3 期となる本研究では、開発項目を大項目 3、中項目 15、小項目 20 に分け、3 グループで分担・協力しながら 5 年間で目標を達成する(表 1)。各テーマ、タスクについては、リポジトリ部、再解析部、データベース部に対し、奥田グループ、石濱グループ、河野グループがそれぞれ主担当グループとして対応するが、独立して開発するのではなく、複数のグループ間で情報を共有しながら共同して開発を行う。1 年目に当たる本年度は、大項目 3、中項目 13、小項目 9 について計画を実施する。以下、詳細を記述する。

表1 プロジェクト全体計画(大中小項目別)

実施する研究開発項目	2023 年度		2024 年度		2025 年度		2026 年度		2027 年度	
リポジトリ部の開発 (大項目)	●	●	●	●	●	●	●	●	●	●
UniPOST 基盤の開発 (中項目)	●	●	●	●	●	●	●	●	●	●
共通ピークピッキングツールの開発 (小項目)			●	●	●	●				
オミクス MS リポジトリへの展開 (小項目)							●	●	●	●
DIA データ対応 (中項目)					●	●	●	●	●	●
投稿データの品質保証 (中項目)	●	●	●	●	●	●				
リポジトリ登録数の確保 (中項目)	●	●	●	●	●	●	●	●	●	●
サーバー体制の確立 (中項目)	●	●	●	●	●	●	●	●	●	●
新潟サーバーの設置 (小項目)	●	●								
3拠点体制の確立 (小項目)			●	●	●	●				
再解析部の開発 (大項目)	●	●	●	●	●	●	●	●	●	●
PTMeXchange 対応 (中項目)	●	●	●	●	●	●	●	●	●	●
Site localization に対する FDR 設定 (小項目)	●	●	●	●	●	●				
UniScore による信頼性評価 (小項目)							●	●	●	●
Peptide-centric DIA 対応 (中項目)	●	●	●	●	●	●				
既知法の評価 (小項目)	●	●	●	●	●	●				
UniScore の適用 (小項目)							●	●	●	●
Isobaric tag 定量対応 (中項目)	●	●	●	●	●	●				
バッチ間標準化対応 (小項目)	●	●	●							
プロジェクト間対応 (小項目)				●	●	●				
検索用リファレンスデータベース (中項目)	●	●	●	●	●	●	●	●	●	●
単一生物種対応 (小項目)	●	●	●	●	●	●				
メタプロテオーム対応 (小項目)							●	●	●	●
コミュニティ連携メタデータ収集 (中項目)	●	●	●	●	●	●	●	●	●	●
応募システムの確立 (小項目)	●	●	●	●	●	●				
年間 100 報の収集 (小項目)							●	●	●	●
データベース部 (大項目)	●	●	●	●	●	●	●	●	●	●
メタプロテオームデータ対応 (中項目)	●	●	●	●	●	●	●	●	●	●
複数生物種対応 (小項目)	●	●	●	●	●	●				
可視化ツール開発 (小項目)							●	●	●	●
非典型翻訳対応 (中項目)	●	●	●	●	●	●				
連携プライムユーザー対応 (中項目)	●	●	●	●	●	●	●	●	●	●
大型国家プロジェクト対応 (小項目)	●	●	●	●	●	●				
公募型 (小項目)							●	●	●	●

1. リポジトリ部

奥田グループが主担当となって開発を進める。

① UniPOST 基盤の開発:

1年目となる本年度は、他のデータベースプロジェクトと UniPOST コンセプトの共有を図り、共同開発体制を確立する。メタデータ入力の共通システム、ピークピッキングツール、オミクス MS ストレージシステムの確立に向け、関係各所とのコンセプト共有を図る。

② リポジトリ登録数の確保:

国内外の各種学会や研究者コミュニティを通じて jPOST の宣伝活動を継続するとともに高品質な再解析データをタイムリーに公開することにより、データ投稿数をコンスタントに増加させる試みを継続する。数値目標として、本年度は累計 2,000 projects を目指す。

③ 3 拠点でのサーバー体制の確立:

1年目となる本年度は、新潟サーバーの設置を行い、データバックアップ体制を強化する。

④ 再解析フローの半自動化:

再解析を担当する石濱グループと協力して、再解析に必要なメタデータをオリジナル登録データ、JPDM 登録データおよびコミュニティ連携で取得する論文抽出データの統一フォーマットによる半自動化エクスポートツールを開発するとともに、再解析データをデータベース部に登録するにあたって河野グループと協力して半自動化ツールの開発を行う。

2. 再解析部

石濱グループが主担当となって開発を進める。

① PTMeXchange 対応:

1年目となる本年度は、HUPO-PSI、PXC 等における情報収集を中心に検討を開始する。

② Peptide-centric DIA 対応:

1年目から DIANN 等の既知法の情報収集を開始する。装置開発スピードが目覚ましく、イオンモビリティ等を取り入れた新しい次世代 DIA が開発されつつあり、またそれに伴うスキャン手法の開発も続いているので、その動向を注視しつつ、さらにはデータ解析ソフトについても、Peptide-centric アプローチを中心に、どれが主流となっていくのかを見極めながら再解析フローのツール開発を進める。

③ Isobaric tag 定量対応:

定量法として今後も汎用的に使われるのは、DIA 法および Isobaric tag 法であると考えられる。そこで Isobaric tag 定量において、大規模データ、多試料データを取得した際に問題となるバッチ間効果、プロジェクト間効果に対して行う補正法に対する標準化を検討する。1年目は、まずはバッチ間補正法についてその標準手法の選択を行う。

④ 検索用リファレンスデータベース:

様々な大規模データが登録されているリポジトリからは、ペプチドの同定・定量のためのデータベース検索で用いることができるリファレンスデータベースを生成することが可能である。さらにこのリファレンスデータは DIA においてもライブラリとして使用が可能である。とくに複数生物種が混在するメタプロテオーム解析ではリファレンスデータライブラリが威力を発揮すると考えられる。まずは1年目に情報収集から開始し、主たる単一生物種(ヒト、マウスなど)に対応したリファレンスデータベースの生成法確立を目指す。

⑤ コミュニティ連携メタデータ収集:

現在、完全にマニュアルで行っているコミュニティ連携メタデータ収集応募システムについて、半自動化で行えるシステムを3年目までに確立する。本システムを使って、年間100報からのデータ抽出・収集ができるシステムの確立を5年目までに行う。

⑥ 再解析フローの半自動化:

リポジトリ部を担当する奥田グループと協力して、再解析に必要なメタデータをオリジナル登録データ、JPDM 登録データおよびコミュニティ連携で取得する論文抽出データの統一フォーマットによる半自動化エクスポートツールを開発するとともに、再解析データをデータベース部にエクスポート登録するにあたって河野グループと協力して半自動化ツールの開発を行う。本項目は5年間にわたって試行を繰り返しながら継続的に行う。

3. データベース部

河野グループが主担当となって開発を進める。

① 複数生物種(メタプロテオームデータ)対応:

第2期からの積み残し課題のひとつとして、複数生物種データをどのようにデータベースに中に入れて可視化するか、がある。どのような機能が必要とされるか、情報収取から始め、DBCLS との協力し、技術的にはかなり高度になると思われるシステムの整備を行う。

② 非典型翻訳対応データベースの開発:

大型国プロ(学術変革領域マルチファセットプロテインズ、AMED-CREST プロテオスタシス、学術変革領域タンパク質寿命学など)を中心にして、世界規模で様々な非典型翻訳情報が集積しつつある。これらのデータセットに対し、高品質再解析データを生成し、データサイエンスに資する専門データベースの構築を目指す。次の項目とも関連するが、本項目では、様々な翻訳開始点、終始点をトランスクリプトーム、トランスレイトーム、プロテオームのレベルで同時に可視化できるツールの開発を検討する。

③ 連携 prime ユーザー対応:

上記の大型国プロをはじめとし、各種 ERATO 領域などと連携し、大規模プロテオームデータを jPOST に誘導するとともに、関連領域の研究者を巻き込みながら、魅力的なデータサイエンスの場を提供する。データベース研究者とデータサイエンス研究者および普段自分では大規模データを利用してこなかった生物系研究者に対し連携 prime ユーザーとして jPOST チームとタッグを組み、発見研

究の可能性、仮設証明の可能性などを様々な可能性をパイロット研究として探る。

④ 再解析フローの半自動化:

リポジトリ部を担当する奥田グループと協力して、再解析に必要なメタデータをオリジナル登録データ、JPDM 登録データおよびコミュニティ連携で取得する論文抽出データの統一フォーマットによる半自動化エクスポートツールを開発するとともに、再解析データをデータベース部にエクスポート登録するにあたって再解析を担当する石濱グループと協力して半自動化ツールの開発を行う。本項目は5年間にわたって試行を繰り返しながら継続的に行う。

(2) 進捗状況

1. リポジトリ部

奥田グループが主担当となって開発を進めた。

① UniPOST 基盤の開発:

1 年目となる本年度は、MassBank, GlycoPost と UniPOST コンセプトの共有化を図った。特にメタデータ入力の共通システムの開発を開始した。オミクスMSストレージシステムの確立に向け、DBCLSとの議論を開始した。2024.3 に開催した HUPO-PSI ワークショップに付随した形で半日の国際シンポジウムを開催し、関連する NBDC-DICP プロジェクトから発表していただくとともに、UniPOST コンセプトについて、PXC 関係者に対して国際アピールも行った。

② リポジトリ登録数の確保:

アジアオセアニア質慮分析学会(韓国済州島)、HUPO(韓国釜山)、日本プロテオーム学会 2023 年大会等で宣伝活動を行うとともに、2024.3 に HUPO-PSI ワークショップを京都で主催し、研究者コミュニティを通じて jPOST の宣伝活動を行った。データ投稿数をコンスタントに増加させる試みを継続した結果、数値目標としていた累計 2000 projects を大きく超え、2024.3 の時点で 2629 projects (1978 公開)を達成した。

③ 3 拠点でのサーバー体制の確立:

1 年目となる本年度は、新潟サーバーの設置を行い、データバックアップ体制を強化した。なお、データストレージ体制については新たな問題が顕在化しつつあり、別途詳細を記述する。

④ 再解析フローの半自動化:

JPDM 登録データと jPOST 登録データおよび PXC のすすめる SDRF フォーマットとの整合性を整える作業を行った。JPDM 投稿時に必要な Supplementary Table 用の jPOST エクスポートツールの修正を行った。

2. 再解析部

石濱グループが主担当となって以下の開発を行った。

① PTMeXchange 対応:

2024.3 HUPO-PSI workshop において議論を進めた。様々な懸念が積み残しポイントとして蓄積していて、当初考えていたほど簡単に PXC に導入するのは困難であるという認識に変わっている。来年以降も情報収集を続け、ある程度内容が確定した時点で導入を開始する予定である。

② Peptide-centric DIA 対応:

DIANN、MQ 等の複数の既知法について、再解析用ツールとしての作動環境を整えた。また FragPipe 中での新規のツール開発などコミュニティの動向を注視しつつ、Peptide-centric アプローチを中心に、どれが主流となっていくのかを見極めながら再解析フローのツール開発を進める。

③ Isobaric tag 定量対応:

Isobaric tag 定量における、大規模データ、多試料データを取得した際に問題となるバッチ間効果、プロジェクト間効果に対して行う補正法に対する標準化検討を開始した。本年度は国際プロジェクトである ICPC で NCI-7 試料を用いた大規模他施設間データ取得を行った。次年度に最初の論文化が予定されている。

④ 検索用リファレンスデータベース:

1年目である本年度はまず情報収集から開始した。主たる単一生物種(ヒト、マウスなど)に対応したリファレンスデータベースとして、PeptideAtlas もしくは NextProt データの確認とともに PXC-MASSIVE knowledgeDB についても調査を行った。

⑤ コミュニティ連携メタデータ収集:

石濱研究室において博士課程学生、田畑研究員を中心に、メタデータ収集応募システムについての検討を開始した。

⑥ 再解析フローの半自動化:

奥田グループによる、JPDM 投稿時に必要な Supplementary Table 用の jPOST エクスポートツールの修正をうけ、JPDM Supplementary Table 確認作業を行った。

3. データベース部

河野グループが主担当となって以下の項目について開発を進めた。

① 複数生物種(メタプロテオームデータ)対応:

遺伝研・森先生と共同で、MAG データ(MicrobeDataHub と MetaLab-MAG)を用いたメタプロテオーム解析の検討を行った。石濱グループの田畑研究員の協力のもと、再解析ワークフローに組み入れた場合の検討も開始した。さらにデータベースにおける複数生物種登録方針についての基本的なルール決めを行った。

② 非典型翻訳対応データベースの開発:

大型国プロ(学術変革領域マルチファセットプロテインズ、AMED-CREST プロテオスタシス、学術変革領域タンパク質寿命学など)を中心にして、世界規模で様々な非典型翻訳情報が集積しつつある。日本プロテーム学会と協力し、これらの DIA データ解析についての WG 発足に向け、準備を開始するとともに、上記大型国プロ領域会議において、宣伝を行った、また、末端プロテオーム DB である TopFind の調査を行った。

③ 連携 prime ユーザー対応:

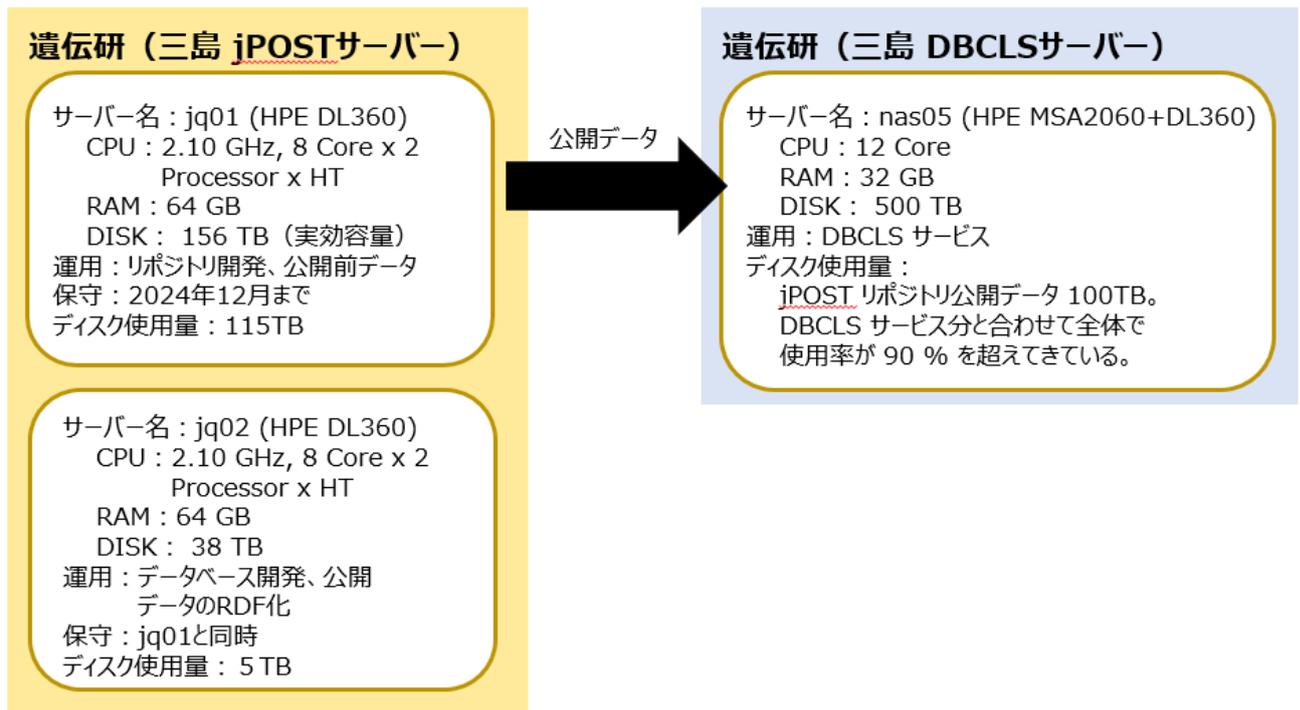
上記の大型国プロをはじめとし、各種 ERATO 領域などと連携し、大規模プロテオームデータを jPOST に誘導するとともに、関連領域の研究者を巻き込みながら、魅力的なデータサイエンスの場を提供する。本年度は日本プロテオーム学会若手 G に対する宣伝活動を中心に行った。

④ 再解析フローの半自動化:

本年度は奥田 G、石濱 G の項で述べたように、jPOST エクスポートツール開発を中心に検討を行った。

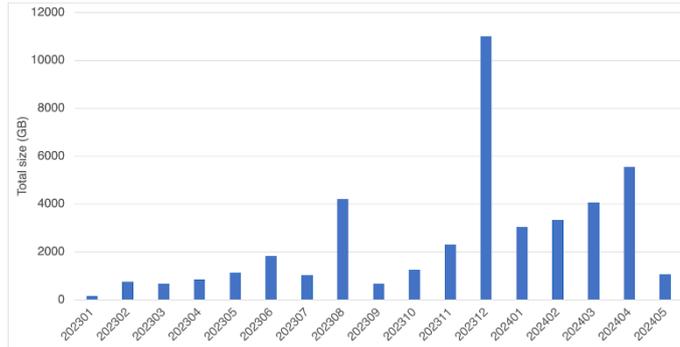
○ jPOST リポジトリ+データベース ~サーバーの現状とストレージ拡充の必要性

現在の jPOST のサーバー構成について以下に示す。

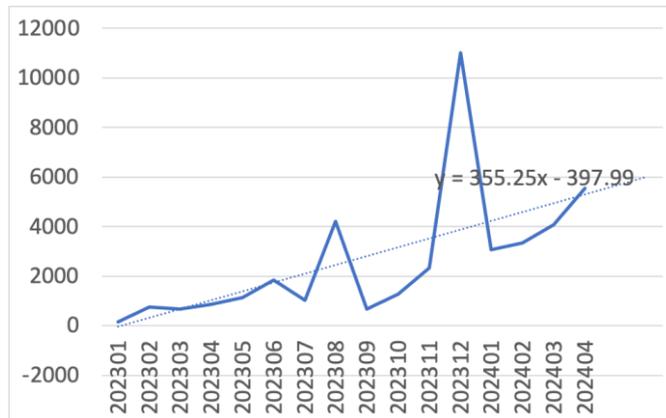


2022 年度後半に、突然の急激な投稿データの増加に対応して、追加予算をいただき、ある程度システムの整備を行ってきた。ところが、その後も予想した以上のペースでデータ量が増加してきている (下の 3 つのグラフ)。

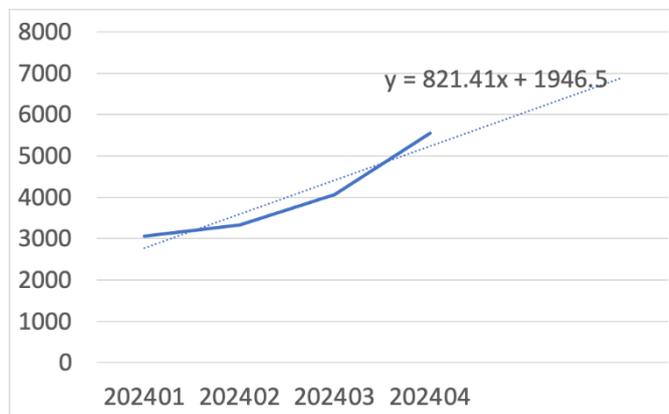
2023 年からの投稿データサイズ



2023 年からのデータで近似→月 355GB 増



2024 のデータで近似→月 821GB 増



今後 60 ヶ月で増加する容量としては、70TB から 150TB くらいが予想されるが、2023 年 12 月ように突発的に 10TB 超えもあり得る。現在は jq01 でリポジトリに登録されるデータを受け入れており、受け入れたデータのうち、公開可能となったものを nas05 に移して公開している。NBDC データベースアーカイブから提供していたものを 2022 年に DBCLS のサーバーに移行しているが、公開データサイズの急速な伸びにより、nas05 の使用量が逼迫しており、DBCLS のサービス提供に影響を及ぼしている。2024 年度に予定しているサーバーの更新に加えて、ディスク増設も行うことにより、リポジトリの公開データを DBCLS のサーバーから jPOST プロジェクトのサーバーに移行したい。ちなみに、jq02 は jPOST プロジェクトで再解析したデータを公開している。

チームからの要望:期中の追加予算配分があるのであれば、早めに準備しておきたいと考えている。

§4. 成果発表等

(1) 原著論文発表

① 論文数概要

種別	国内外	件数
発行済論文	国内(和文)	0件
	国際(欧文)	0件
未発行論文 (accepted, in press 等)	国内(和文)	0件
	国際(欧文)	0件

② 論文詳細情報

該当なし

(2) その他の著作物(総説、書籍など)

該当なし

(3) 国際学会および国内学会発表

① 概要

種別	国内外	件数
招待講演	国内	4件
	国際	2件
口頭発表	国内	0件
	国際	0件
ポスター発表	国内	4件
	国際	2件

② 招待講演

〈国内〉

1. 石濱泰、jPOST prime: コミュニティ連携を基盤とするプロテオームデータベース環境の実現、第 71 回質量分析総合討論会、グランキューブ大阪(大阪府立国際会議場)、5月15日
2. 石濱泰、プロテオフォーム大規模解析とプロテオームデータベース、第 2 回環境化学物質 3 学会合同大会、あわぎんホール、5月30日
3. 奥田修二郎、データベースからアプローチするトランスオミクス、日本プロテオーム学会 2023 年大会、朱鷺メッセ、7月25日
4. 石濱泰、jPOSTprime: データサイエンスコミュニティハブとしてのプロテオームデータベース、第 95 回日本生化学会大会、福岡国際会議場、11月2日

〈国際〉

1. Yasushi Ishihama, Exploring Public Data Repositories by Integration and Sharing of Proteome Data, HUPO World Congress 2023 (Busan), Sep 19, 2023.

2. Shujiro Okuda, Current status of the jPOST and JPDM projects and future approaches to multi-omics data integration, HUPO-PSI Kyoto Symposium 2024 on Omics Repository, Database and Data Journal, Fujita Hall-Kyoto University, Mar 21, 2024

③ 口頭講演

〈国内〉

該当なし

〈国際〉

該当なし

④ ポスター発表

〈国内〉

1. 奥田修二郎、jPOSTリポジトリから UniPOST への展開、トーゴの日シンポジウム 2023、日本科学未来館、10月5日
2. 田畑剛、コミュニティ連携基盤 jPOST prime を通じた再解析ワークフローの確立にむけて、トーゴの日シンポジウム 2023、日本科学未来館、10月5日
3. 河野信、jPOSTdb の開発、トーゴの日シンポジウム 2023、日本科学未来館、10月5日
4. 吉沢明康、小林大樹、守屋勇樹、幡野敦、高見知代、松本雅記、荒木令江、田畑剛、岩崎未央、杉山直幸、小寺義男、田中聡、五斗進、河野信、奥田修二郎、石濱泰、

〈国際〉

1. Henry Lam, Tytus D. Mak, Joshua A. Klein, Wout Bittremieux, Ralf Gabriels, Yasset Perez-Riverol, Tim Van Den Bossche, Andrew R Jones, Pierre-Alain Binz, Shin Kawano, Luis Mendoza, Nuno Bandeira, Jeremy Carver, Benjamin Pullman, Zhi Sun, Nils Hoffmann, Jim Shofstahl, Yunping Zhu, Helge Hecht, Eric Deutsch, Juan Antonio Vizcaino, Proteomics Standards Initiative (PSI) proposed peak annotation format (mzPAF) and spectral library format (mzSpecLib) standards, ASMS2023, Houston, 6月4日-8日
2. Satoshi Tanaka, Masaki Murase, Masaki Kato, Hiroyuki Yamamoto, Tsuyoshi Tabata, Maiko Kusano, Shin Kawano, Susumu Goto, Yasushi Ishihama, Akiyasu C. Yoshizawa, P Mass++ ver.4 – An open-source MS data viewer with enhanced basic functions and easy implementation of external software, ASMS2023, Houston, 6月4日-8日

(4) 知的財産権の出願（国内の出願件数のみ公開）

出願件数

種別		件数
特許出願	国内	0件

(5) 受賞・報道等

① 受賞

該当なし

② メディア報道

該当なし

③ その他の成果発表

該当なし

§5. 主要なデータベースの利活用状況

(1) アクセス数

①実績

表 1 研究開発対象の主要なデータベースの利用状況(月平均)

DB名	種別	2023年度
jPOSTrepo	訪問者数	1,595
	訪問数	7,543
	閲覧ページ数	2,417,238
jPOSTdb	訪問者数	1,106
	訪問数	7,437
	閲覧ページ数	16,828
jPOST	訪問者数	890
	訪問数	1,097
	閲覧ページ数	11,009



② 分析

利用状況の月間平均)を以下に示す。アクセス数(訪問数)約 1,400 件 (2023 年度の月間平均は現時点で約 18,000 件)で国内のみならずアジアや欧米からも多くのデータが登録されている。この 1~2 年の収載データ数と多様性の増加に伴い利用が急増している。国外の研究者からも広くデータが利用されている(国別のアクセス比率:米国 34%, 日本 14%, 中国 12%, ドイツ 5%, インド 3%)。ただ、この急増の理由のすべては説明しきれず、我々にもこれについては原因不明である。

(2) データベースの利用状況を示すアクセス数以外の指標

2023 年は 849 プロジェクトが投稿された。jPOSTrepo 論文(Okuda et al. NAR 2017)は 127 回引用、j

POST env 論文(Moriya et al. NAR 2019)は 28 回引用(Google Scholar 2023 年分)された。

(3) データベースの利活用により得られた研究成果(生命科学研究への波及効果)

2023 年 12 月現在 Scopus で調査した結果、データベースに関連する論文数は 7 報で、その引用は 1,216 報となっているが、具体的な研究成果については把握できていない。

(4) データベースの利活用によりもたらされた産業への波及効果や科学技術のイノベーション(産業や科学技術への波及効果)

把握できていない。

§6. 研究開発期間中に主催した活動(ワークショップ等)

(1) 進捗ミーティング

年月日	名称	場所	参加人数	目的・概要
2023年 4月3日	チーム内月例ミーティング (非公開)	オンライン	13	第3期 jPOST キックオフミーティング
2023年 4月25日	チーム内月例ミーティング (非公開)&DBCLSとの 打ち合わせ	ハイブリッド (京大薬)	15	対面でのキックオフ。DBCLS 開発 チームによるインタビュー、チーム内 の開発方針についてなど。
2023年 5月23日	「統合化推進プログラム」 キックオフミーティング	オンライン	30	NBDC 主催のキックオフミーティ ング。総括方針と期待、2023 年度 採択課題 研究代表者の発表、 DBCLS の発表、総合討論など
2023年 5月31日	チーム内月例ミーティング (非公開)	オンライン	13	研究進捗報告のためのミーティング
2023年 7月7日	チーム内月例ミーティング (非公開)	オンライン	14	研究進捗報告のためのミーティング
2023年 7月23日	チーム内月例ミーティング (非公開)	朱鷺メッセ	14	研究進捗報告のためのミーティング
2023年 8月31日	チーム内月例ミーティング (非公開)	オンライン	13	研究進捗報告のためのミーティング
2023年 9月28日	チーム内月例ミーティング (非公開)	オンライン	11	研究進捗報告のためのミーティング
2023年 10月6日	チーム内月例ミーティング (非公開)	オンライン	8	MicrobeDatahub との連携のため のミーティング
2023年 10月27日	チーム内月例ミーティング (非公開)	オンライン	11	研究進捗報告のためのミーティング
2023年 11月24日	チーム内月例ミーティング (非公開)	オンライン	13	研究進捗報告のためのミーティング
2023年 12月21日	チーム内月例ミーティング (非公開)	オンライン	11	研究進捗報告のためのミーティング
2024年 1月23日	チーム内月例ミーティング (非公開)	オンライン	9	研究進捗報告のためのミーティング
2024年 2月22日	チーム内月例ミーティング (非公開)	オンライン	11	研究進捗報告のためのミーティング
2024年 3月14日	チーム内月例ミーティング (非公開)	オンライン	13	研究進捗報告のためのミーティング

(2) 主催したワークショップ、シンポジウム、アウトリーチ活動等

年月日	名称	場所	参加人数	目的・概要
2023年 7月23日	JPDM-jPOST セミナー	新潟	20	JPDM Data Descriptor論文の書 き方についてのセミナー
2023年 10月5日-6日	MaxQuant Workshop	日本科学未 来館	25	MaxQuant ソフトウェアの使い方 についてのセミナー
2024年 3月18日-20 日	HUPO-PSI-workshop	京大薬	88	PSI メンバーが世界から一同に集ま り集中的に議論するための国際ワー クショップ
2024年 3月21日	HUPO-PSI Kyoto Sym posium 2024 on Omics Repository, Database and Data Journal	京大薬	107	データベース統合のための国際シン ポジウム

以上

別紙1 既公開のデータベース・ウェブツール等

No.	正式名称	別称・略称	概要	URL	公開日	状態	分類	関連論文
1	Japan Proteome Standard Repository/Database	jPOST	本DBは、国内外に散在している種々のプロテオーム情報を標準化・統合・一元管理し、データベース化したもので、多彩な生物種の翻訳後修飾や絶対発現量等の情報まで網羅したプロテオーム統合データベースです。リポジットリパート、再解析リパート、データベースリパートからなり、現在はリポジットリパートとデータベースリパートを公開しています(2018年3月)。データベースにはプロジェクトの名称、説明、サンプル情報や、関連タンパク質のパスウェイマップによる絞り込みリストをまとめた「Dataset」と、タンパク質のID、アミノ酸配列、ペプチド情報へのリンクをまとめた「Protein」の一覧があります。生物種、サンプルタイプ、疾患名などによる検索が可能です。	https://jpostdb.org		維持・発展	データベース等	
2	jPOSTrepo	Japan ProteOme STandard Repository	ユーザによる質量分析の生データ、ピークリスト、解析データを登録するためのリポジット。ProtomoeXchange (PX) Consortiumに加盟し、登録ユーザが論文投稿する際に必要となるPX IDを発行する。	https://repository.jpostdb.org/	2016/5/2	維持・発展	データベース等	Okuda S, Watanabe Y, Moriya Y, Kawano S, Yamamoto T, Matsumoto M, Takami T, Kobayashi D, Araki N, Yoshizawa AC, Tabata T, Sugiyama N, Goto S, Ishihama Y.: “jPOSTrepo: an international standard data repository for proteomes”, Nucleic Acid Research, Database Issue, 45 (D1): D1107-D1111, 2017 (DOI: 10.1093/nar/gkw1080)
3	jPOST repository demonstration		jPOSTリポジットシステムでデータ公開を気にせず試しに利用してもらうためのデモサイト。	https://rep-demo.jpostdb.org/	2016/5/2	維持・発展	データベース等	
4	PRESTO		登録ファイルを高速にサーバーに転送するために、ファイルをチャンクに分割し、並列転送を実行するツール。	https://github.com/PrestoTools/	2018/8/19	維持・発展	データベース等	
5	jPOSTdb	Japan ProteOme STandard Database	本DBは、再解析後のデータ(Cubeと呼ぶ)を一元化しデータベース化したもので(Globeと呼ぶ)、様々なフィルターによりプロジェクトを串刺しにして目的データセット(Sliceと呼ぶ)を抽出し、解析するものである。	http://globe.jpostdb.org/	2018/3/26	維持・発展	データベース等	Moriya, Y., Kawano, S., Okuda, S., Watanabe, Y., Matsumoto, M., Takami, T., Kobayashi, D., Yamanouchi, Y., Araki, N., Yoshizawa, A. C., Tabata, T., Iwasaki, M., Sugiyama, N., Tanaka, S., Goto, S., and Ishihama, Y. The jPOST environment: an integrated proteomics data repository and database. Nucleic Acids Res. 47(D1):D1218-D1224 (2019). doi: 10.1093/nar/gky899.
6	iMPAQT	in vitro proteome-assisted MRM for Protein Absolute Quantification	in vitroにおいて網羅的に合成した組換えタンパク質を用いて、MS/MSスペクトルを取得しデータベースを構築。Description, GeneID, Pathway, GOなどで検索したタンパク質のMRMtransitionファイルをダウンロード可能。	http://impaqt.jpost.org/iMPAQT/	2016/12/25	維持・発展	データベース等	Matsumoto, M., Matsuzaki, F., Oshikawa, K., Goshima, N., Mori, M., et al. (2016). A large-scale targeted proteomics assay resource based on an in vitro human proteome. Nature Methods. 14, 251-258.
7	iMPAQTquant	iMPAQT-quant is a tool for analysis of MRM chromatogram.	多重反応モニタリング法による計測に必要な情報を格納したデータベースと定量解析のためのツール	http://impaqt.jpost.org/iMPAQT/impactquant.html	2016/12/25	維持・発展	データベース等	Matsumoto, M., Matsuzaki, F., Oshikawa, K., Goshima, N., Mori, M., et al. (2016). A large-scale targeted proteomics assay resource based on an in vitro human proteome. Nature Methods. 14, 251-258.