プロテオームデータベース jPOSTの挑戦

京大院薬　石濱 泰

＆　jPOST TEAM

（凡例）
- データバンク的DB
- 統合化推進プログラム課題DB
- 統合化推進PG終了課題DB（予定含む）
- 主要プロジェクト
- 別プロジェクト

2015.3時点

（11）ウイルス
（12）iPS細胞

環境：Bio Bank Japan

化合物（薬）：疾患・医薬品DB　創薬等支援技術基盤PF　（7）

集団：ゲノムコホートDB　ヒトゲノム変異DB　（1）

個体：次世代がん　脳画像DB

フェノーム：表現型DB（マウス）（ナズナ）　脳プロ

細胞：（10）　生命動態DB　（12）

パスウェイ：パスウェイDB-RDF　パスウェイDB　（6）

インタラクトーム

メタボローム：メタボロームDB　ゲノムコホートDB　（5）

糖鎖：糖鎖DB

タンパク質立体構造：タンパク質DB-RDF

プロテオーム：
プロテオームDB
（jPOST: Japan ProteOme Standard Repository/Database） （4）

トランスクリプトーム：GEO（米）/Array Express（欧）/DOR　（3）

エピゲノム：創薬等支援技術基盤PF（機能ゲノミクス領域）　IHEC　CREST（エピゲノム領域）　ヒト疾患オミックスDB　（2）

ゲノム：（11）　微生物DB　植物DB　DDBJ-RDF　ゲノムコホートDB　ヒトゲノム変異DB　（1）
DDBJ（日）/GenBank（米）/EMBL（欧）　（8）

ウイルス　微生物　植物　動物　ヒト

2

15

jPOSTとは

Data Integration & Sharing in Life Science


Database

Re-analysis
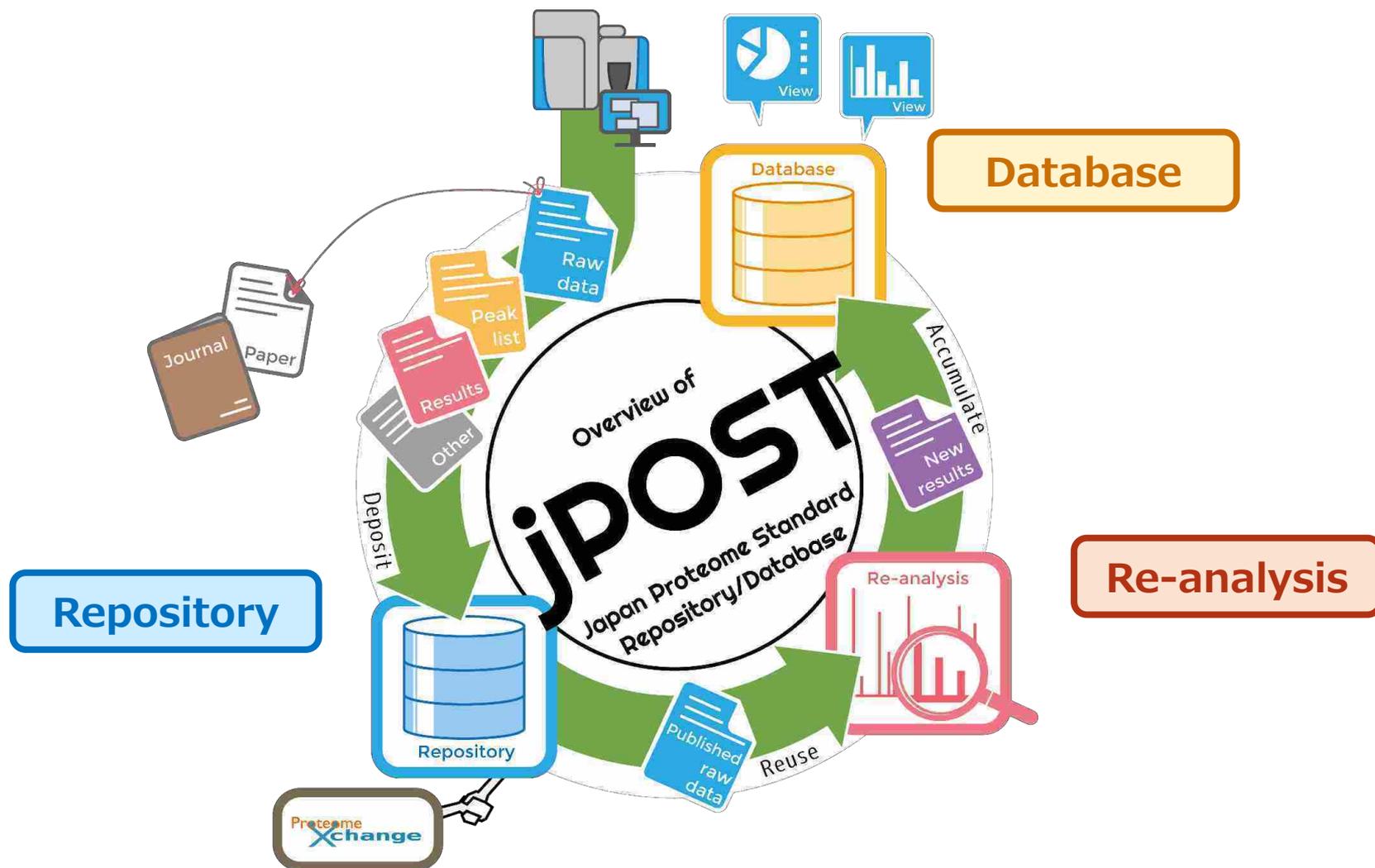
Repository

jPOST repository

4
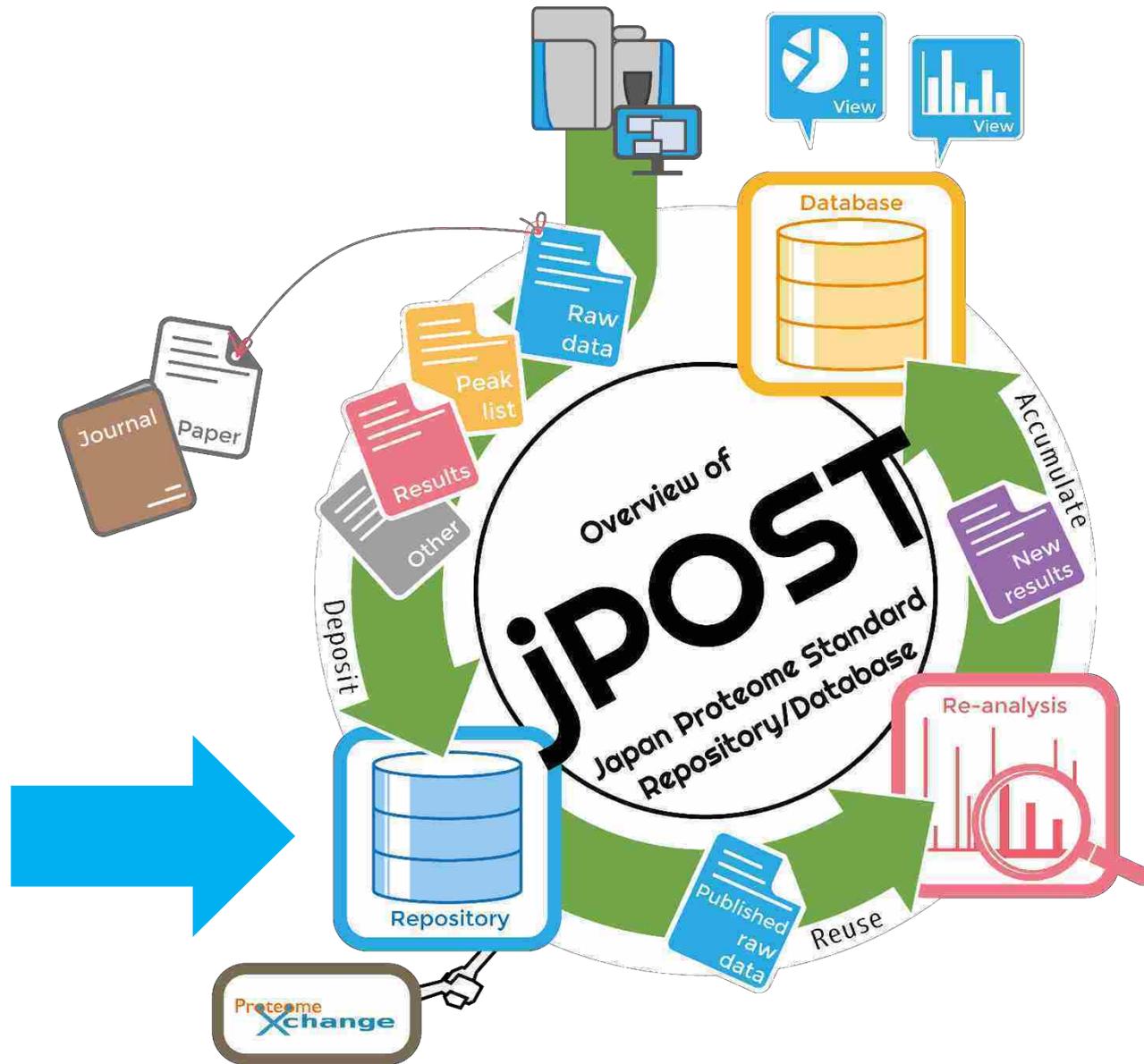
jPOST repository – Current status

データ登録数は目標としていた年間100プロジェクトを大きく上回り、
- 2018.4〜2019.3で 153プロジェクト
- 2019.4〜2020.3で 202プロジェクト
- 2020.4〜2021.3で 253プロジェクト　が登録された。

1042

Private
Announced
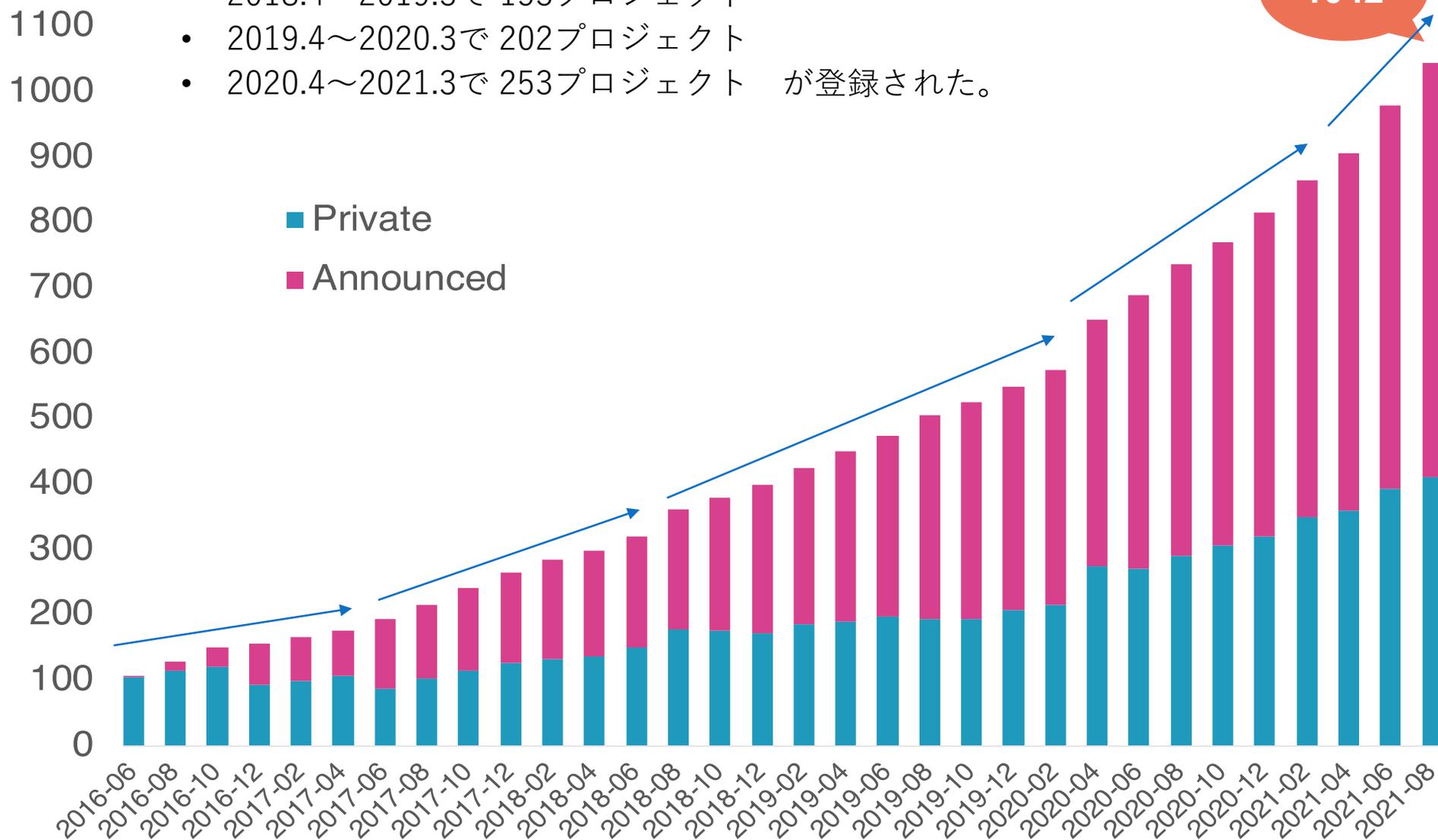
jPOST User Distribution

User distribution
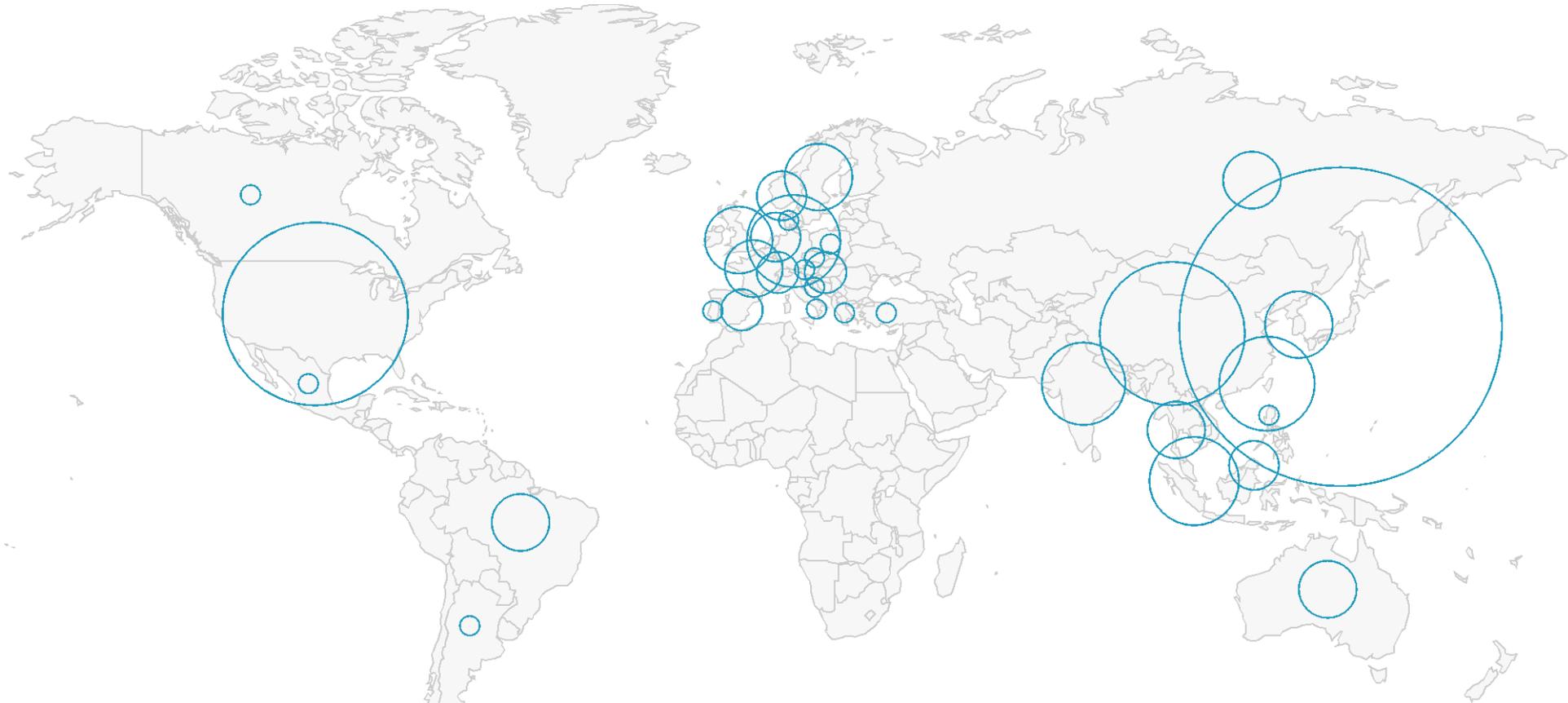
July 2019	**265 users from 28 countries**
July 2020	**391 users from 37 countries**
Sep 2021	**543 users from 37 countries**

Statistics

798 projects are registered. **479** are opened.

65623 files amount to **26.3 TB**.

126 species.
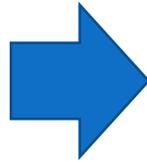
1050 projects are registered. **643** are opened.

80195 files amount to **35.6 TB**.

172 species.

１．フレッシュなデータがどんどん勝手にたまる

偽ヒットのコントロール

True hits

Number of Identified proteins

Accumulated false hits

データが増えると偽物も増える

LC/MS/MS runs
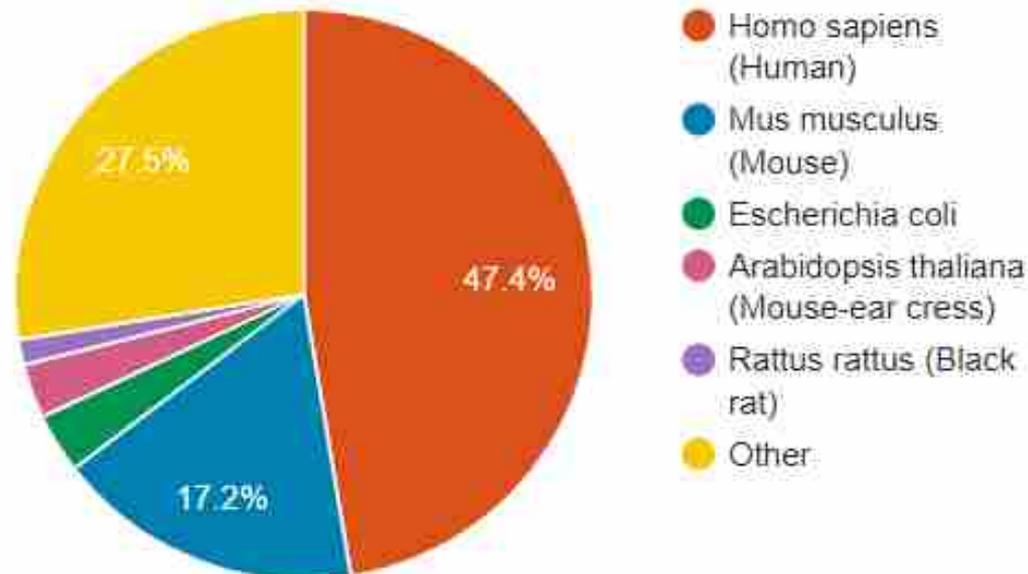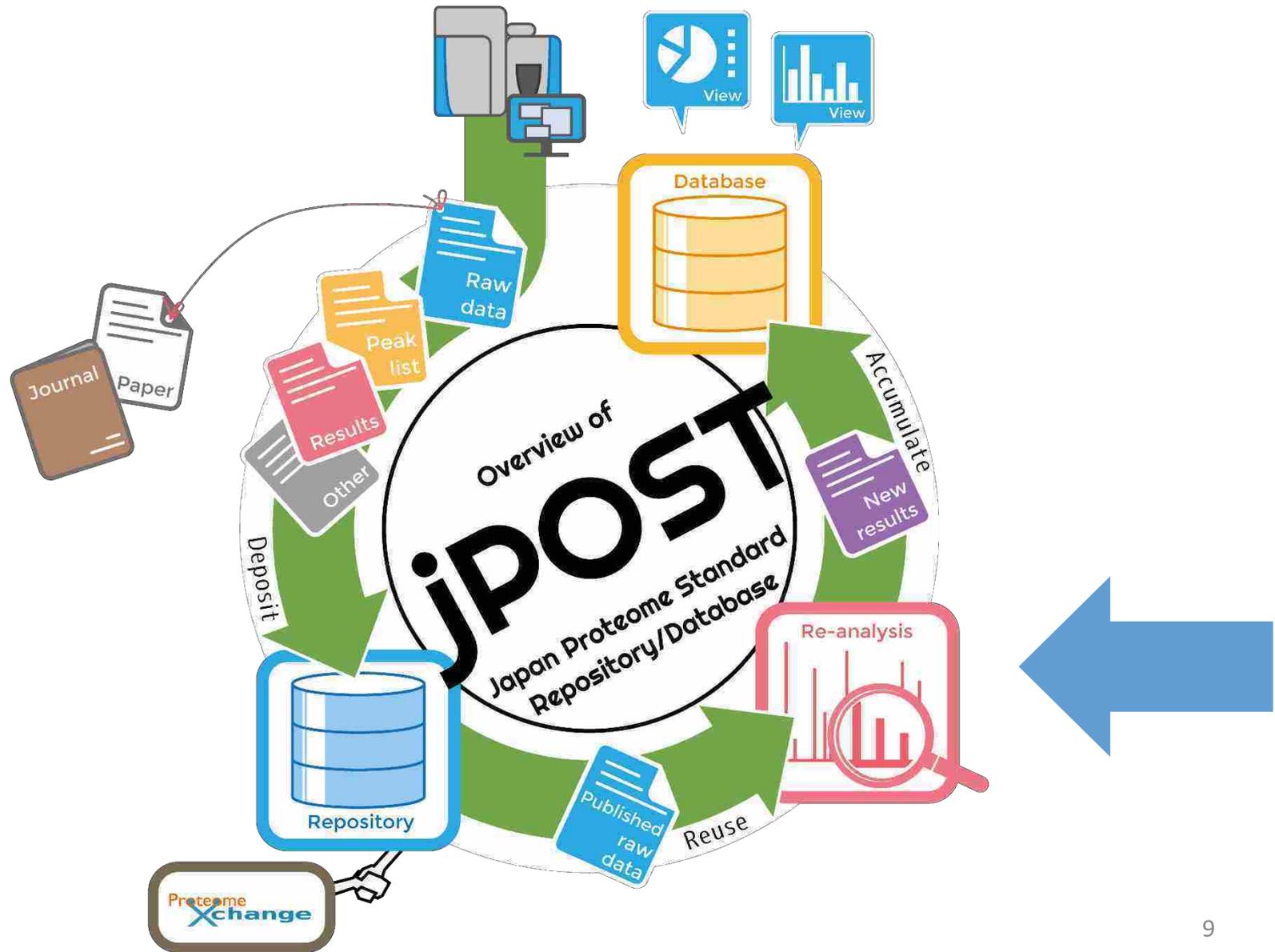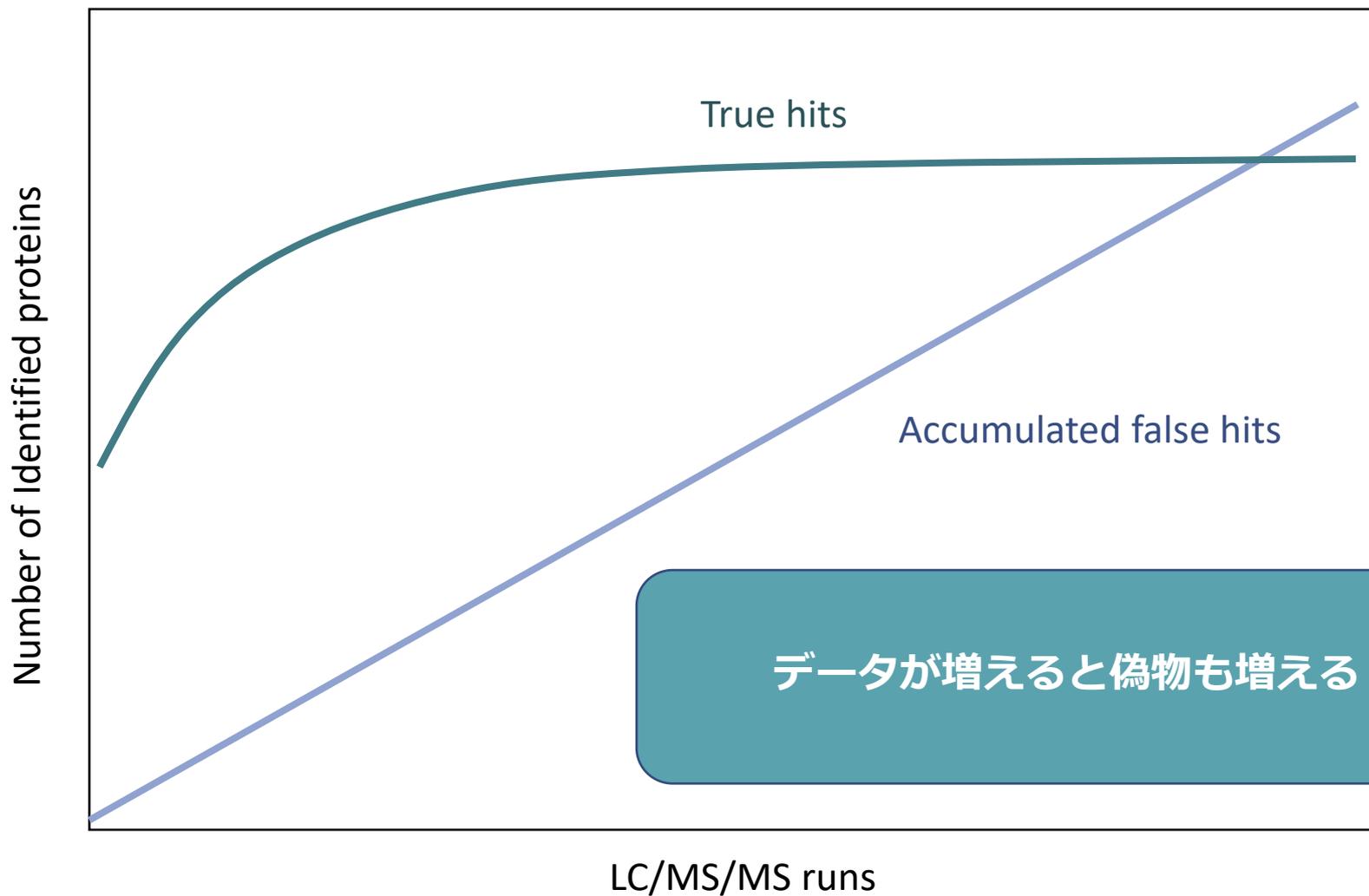
How can we merge the results
from different sources?
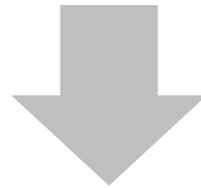
UniScore

- based on peak annotation in MSMS
- search engine independent
- MS instrument independent
- search DB independent
- can be used as universal threshold
 for peptide identification

UniScoreによる再解析 (Human Proteome Project)

MaxQuant 7,886 / UniScore 9,143

1,426 — 6,460 Proteins — 2,683
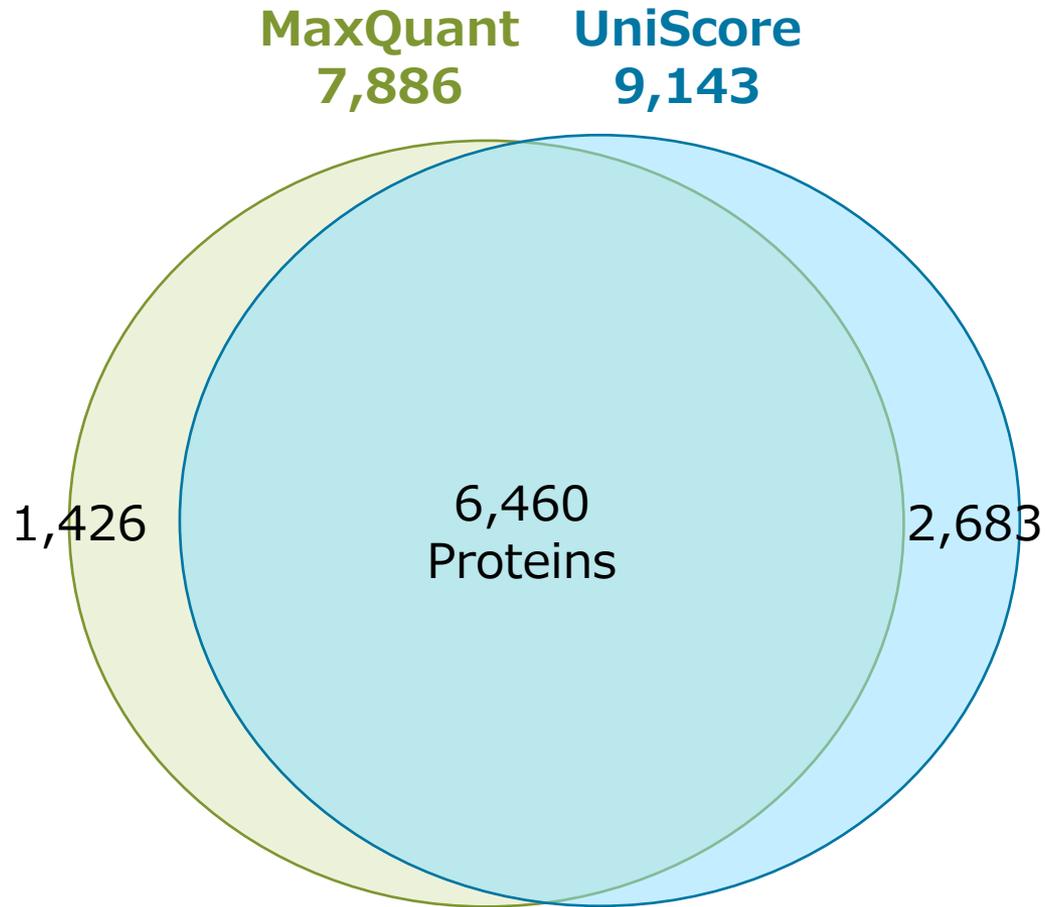
**Dataset:JPST000203 – 206
Colorectal cancer samples (used in Wilhelm et al Nature 2014)**

jPOSTの特徴

１．フレッシュなデータがどんどん勝手にたまる

２．再解析により、データの統一化が実現される
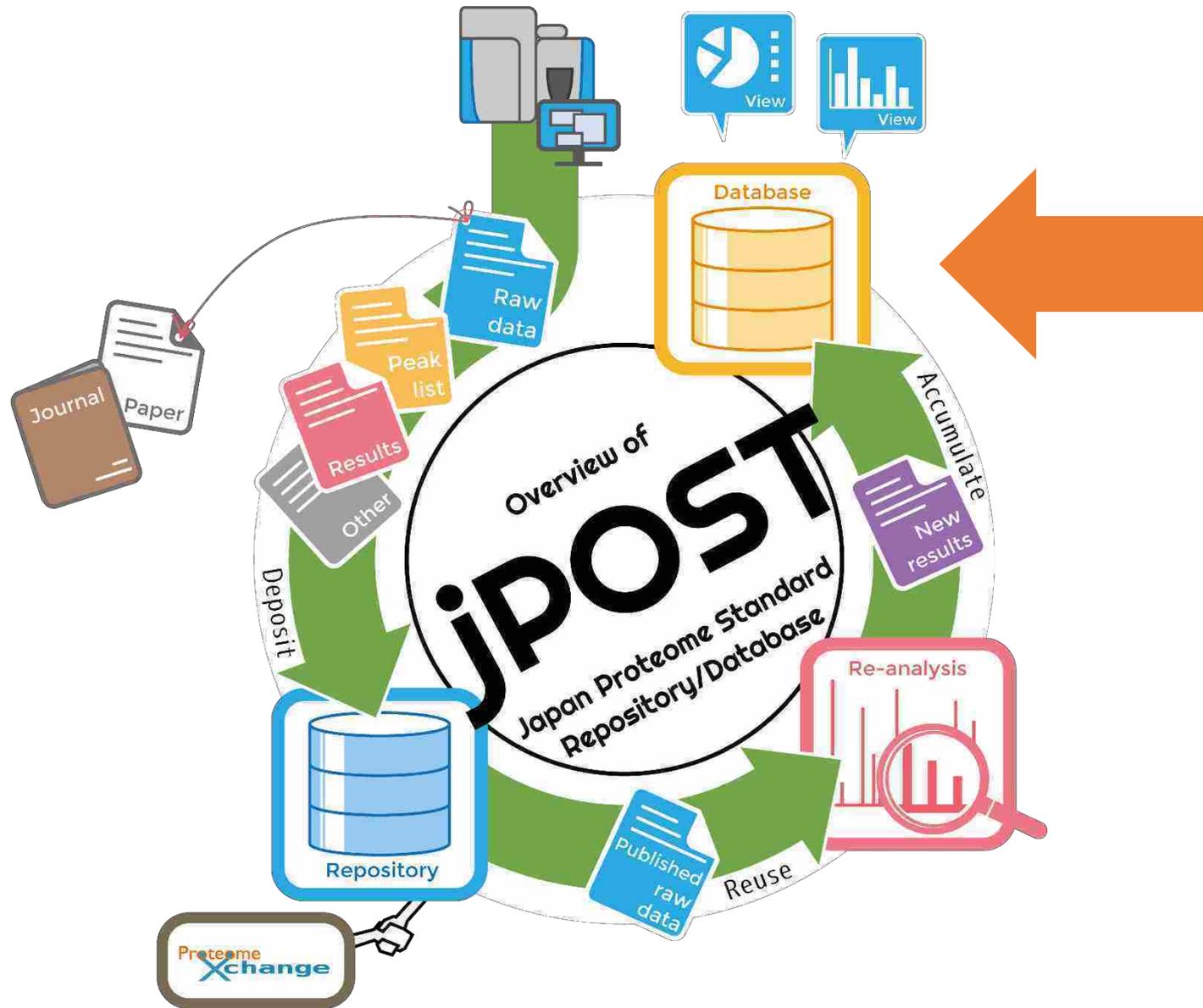
The jPOST environment: an integrated proteomics data repository and database

Moriya et al., *Nucleic Acids Res,* 2019 Jan 8;47(D1):D1218-D1224
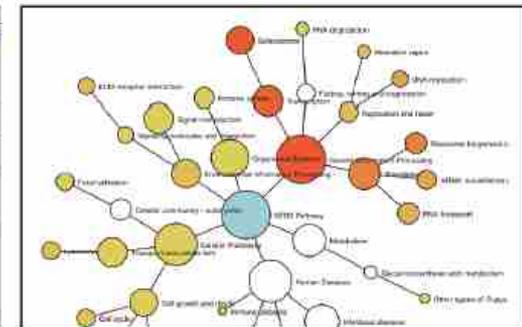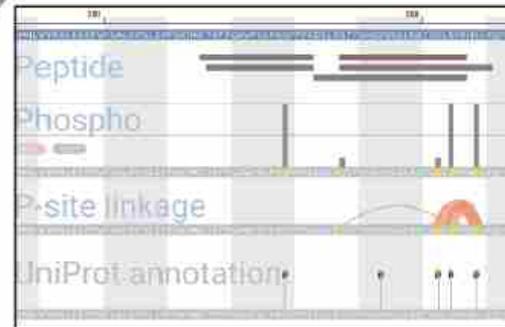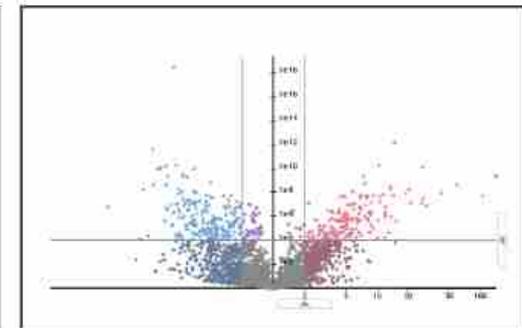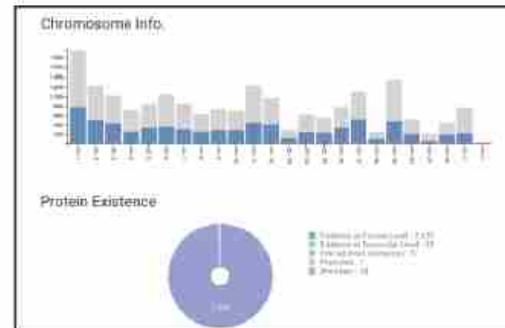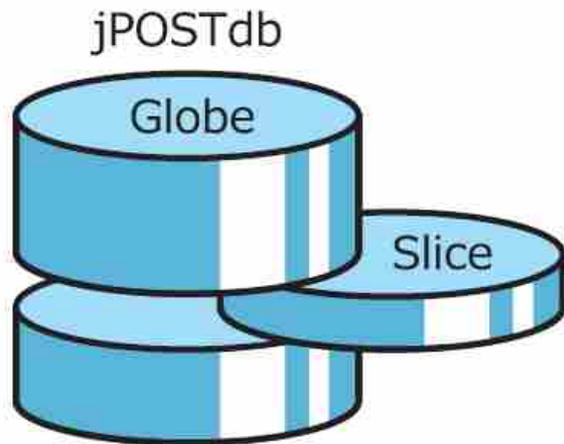
UniProtとの連携：

Cross-references between UniProt and jPOST

jPOSTの特徴

1．フレッシュなデータがどんどん勝手にたまる

2．再解析により、データの統一化が実現される

3．カスタムDBの作成ツールと可視化ツールの提供

https://jpostdb.org

About   **Repository**   **Database**   **Workflow**   **Gadgets**   **COVID-19**   **Help**   **FAQ**   **Contact**   〽

JPOST
Repository/Database

**Japan Proteome Standard
Repository/Database**

Recent posts

other

jPOST member's co-authored paper about USI has been published.

🕑 2021-06-30 👤 jpost

The Universal Spectrum Identifier (USI) is an essential mechanism for the wide use of proteomics data provided by the ProteomeXchange (PX) repositories. It has been discussed for a long time in the HUPO Proteomics Standards Initiative, to which jPOST member Prof. Kawano has contributed greatly. The USI facilitates access to the huge amount of spectral data registered in the PX repositories, and is expected to make proteomics data is more findable, accessible, interoperable, and reusable. We expect that the USI leads to the use of

Repository Help ▸

jPOSTrepo is a raw and processed data repository of mass-based proteomics.

Workflow Help ▸

This provides the re-analysis workflow used in jPOST.

COVID-19 Help ▸

This is a special page on coronavirus (SARS-CoV-2) and the COVID-19 disease data, where mass-based proteomics data in ProteomeXchange resources are analyzed by the jPOST re-analysis workflow.

Database Help ▸

jPOSTdb is a database containing re-analysis results with unified criteria for proteome data from jPOSTrepo.

Gadgets Help ▸

jPOST gadget is a collection of tools we have developed as well as other commonly used tools for mass spectrometry-based proteomics.

Help

A guide to using the jPOST resources.

Molecular Cell

Growth Factor Receptor Signaling Inhibition Prevents SARS-CoV-2 Replication

Graphical Abstract
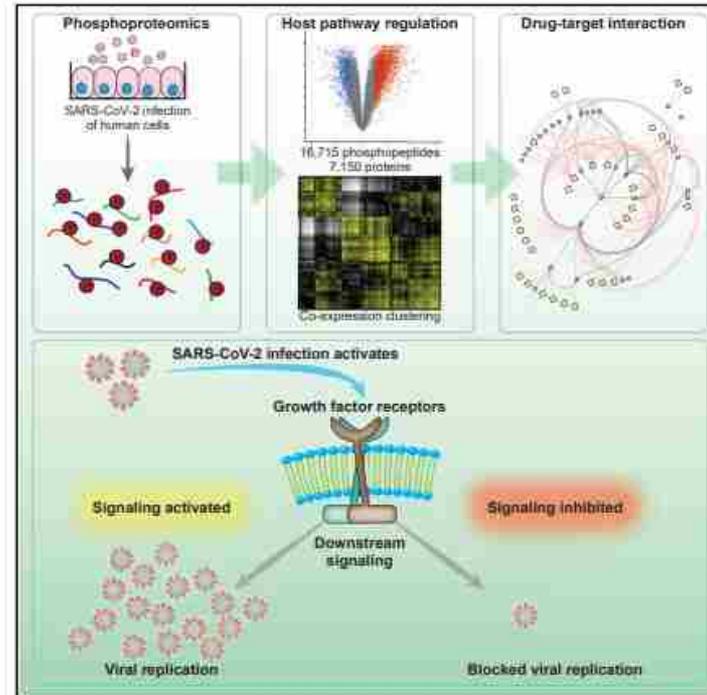
Authors

Kevin Klann, Denisa Bojkova,
Georg Tascher, Sandra Ciesek,
Christian Münch, Jindrich Cinatl

Correspondence

ch.muench@em.uni-frankfurt.de (C.M.),
cinatl@em.uni-frankfurt.de (J.C.)

In Brief

In this study, Klann et al. dissected the host cell signaling landscape upon infection with SARS-CoV-2. Mapping differential signaling networks identified a number of pathways activated during infection. Drug-target network analysis revealed potential therapeutic targets. Growth factor receptor signaling was highly activated upon infection and its inhibition prevented SARS-CoV-2 replication in cells.

Highlights

- Phosphoproteomics of SARS-CoV-2-infected cells reveal the signaling landscape

- SARS-CoV-2 proteins are extensively phosphorylated in host cells

- Infection leads to the activation of growth factor receptor signaling

- Drugs inhibiting growth factor receptor signaling prevent viral replication

Re-analysis of COVID-19 paper

Klann *et al., Molecular Cell* **80**, 164 (2020)

同定リン酸化ペプチド数の比較
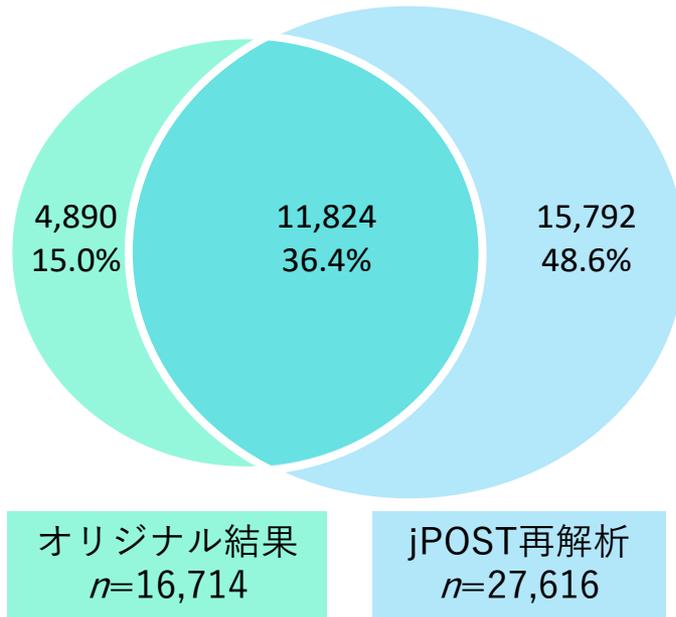
両解析結果のペプチド長ごとの比較
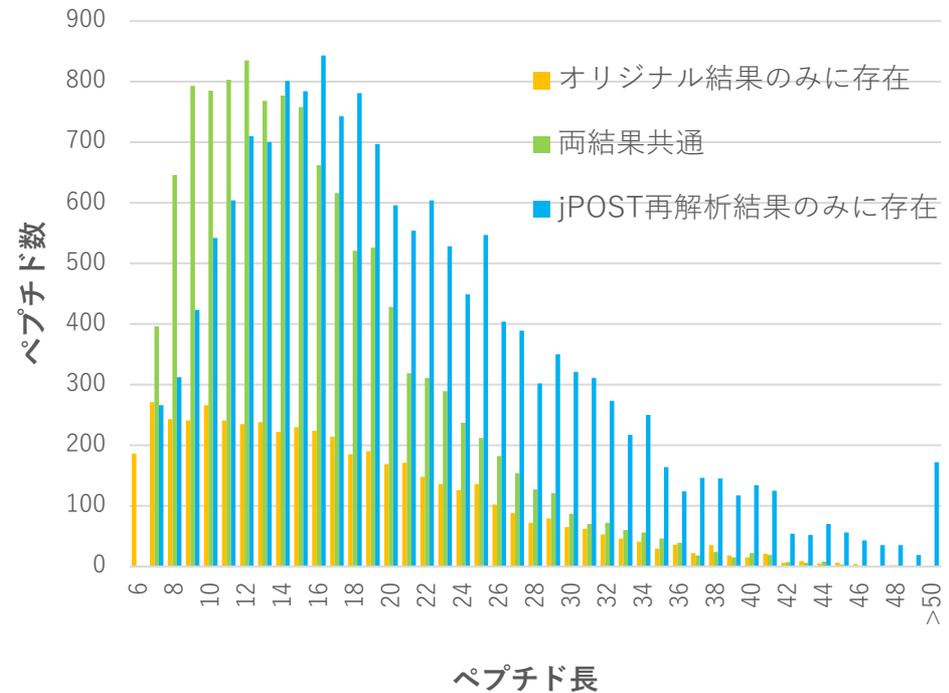
Phosphorylated sites on SARS-CoV-2 proteins

同定されたリン酸化サイトの比較
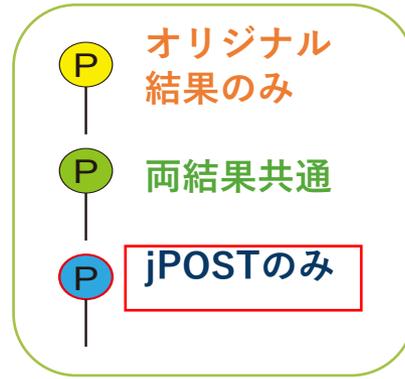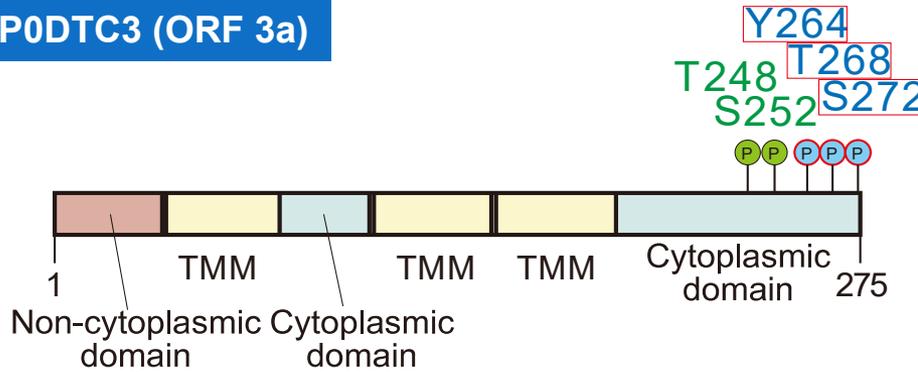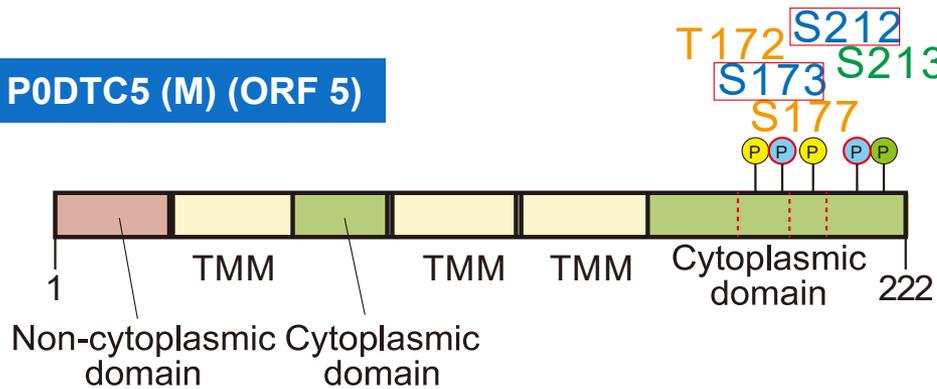
使ってください jPOST

提案：データサイエンス研究者とデータベース研究者の融合

"使う"人はいつも外部ユーザー
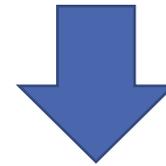
つくる人とつかう人がごっちゃになったプロジェクトをやりたい！
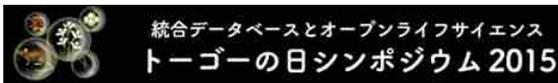
面白いアイディア持ってるデータサイエンティストの皆さん、誰か知ってる人にぜひ声をかけてください。