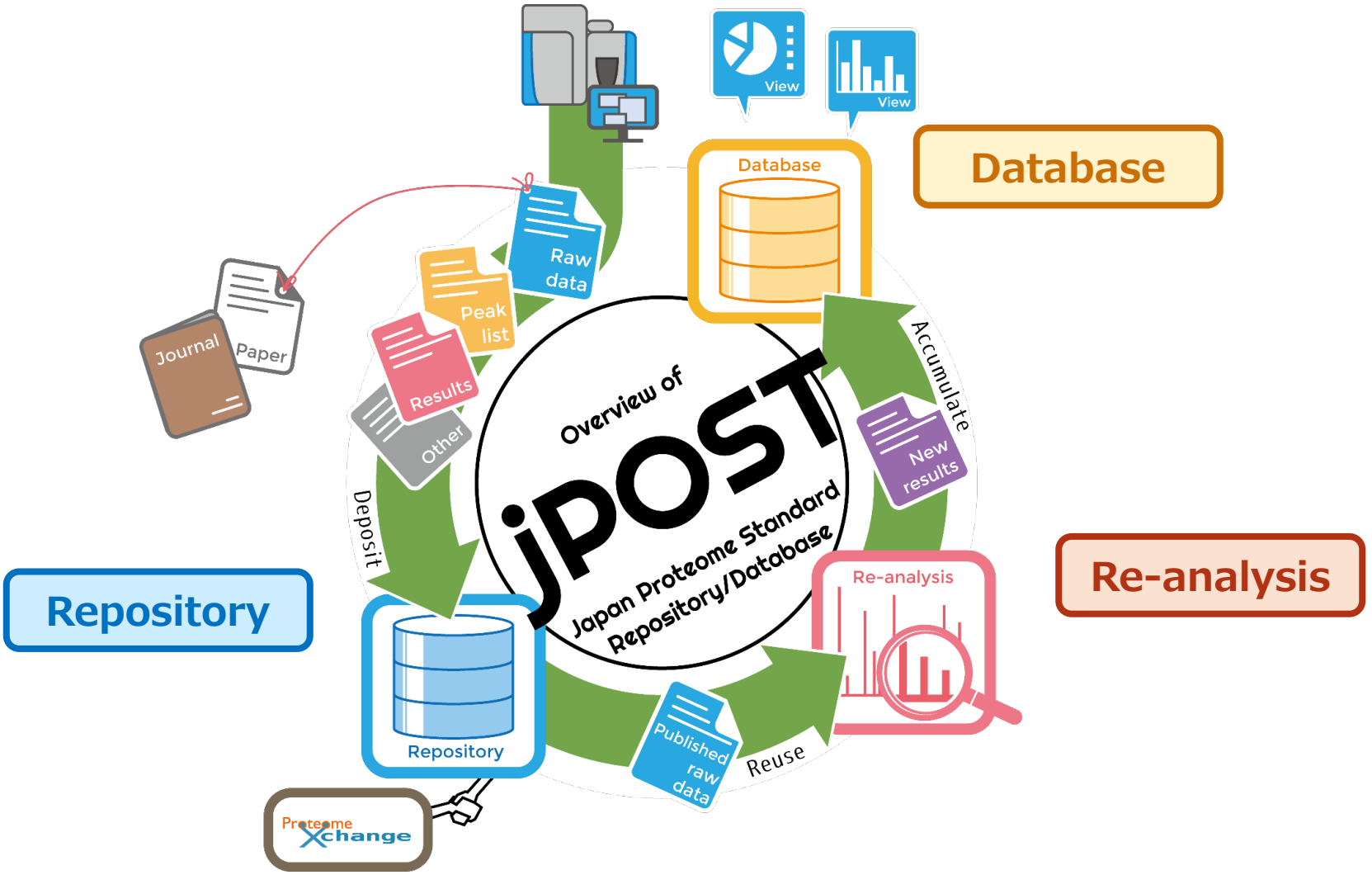
第45回日本分子生物学会年会　フォーラム「生命科学のデータベース活用法」(2022.11.30)

プロテオームデータベース jPOSTの挑戦

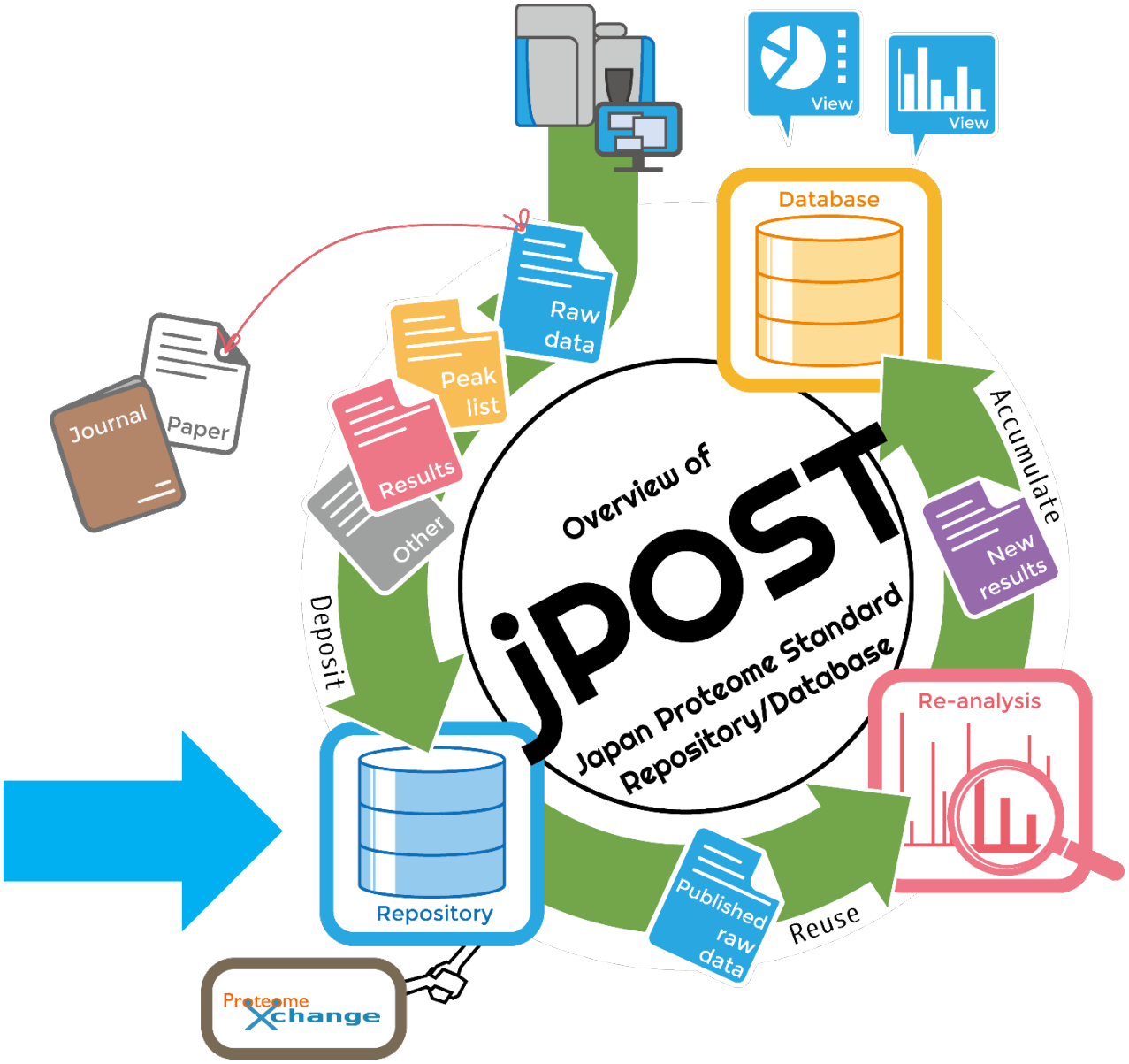
京大院薬　石濱 泰
&　jPOST TEAM

jPOSTとは

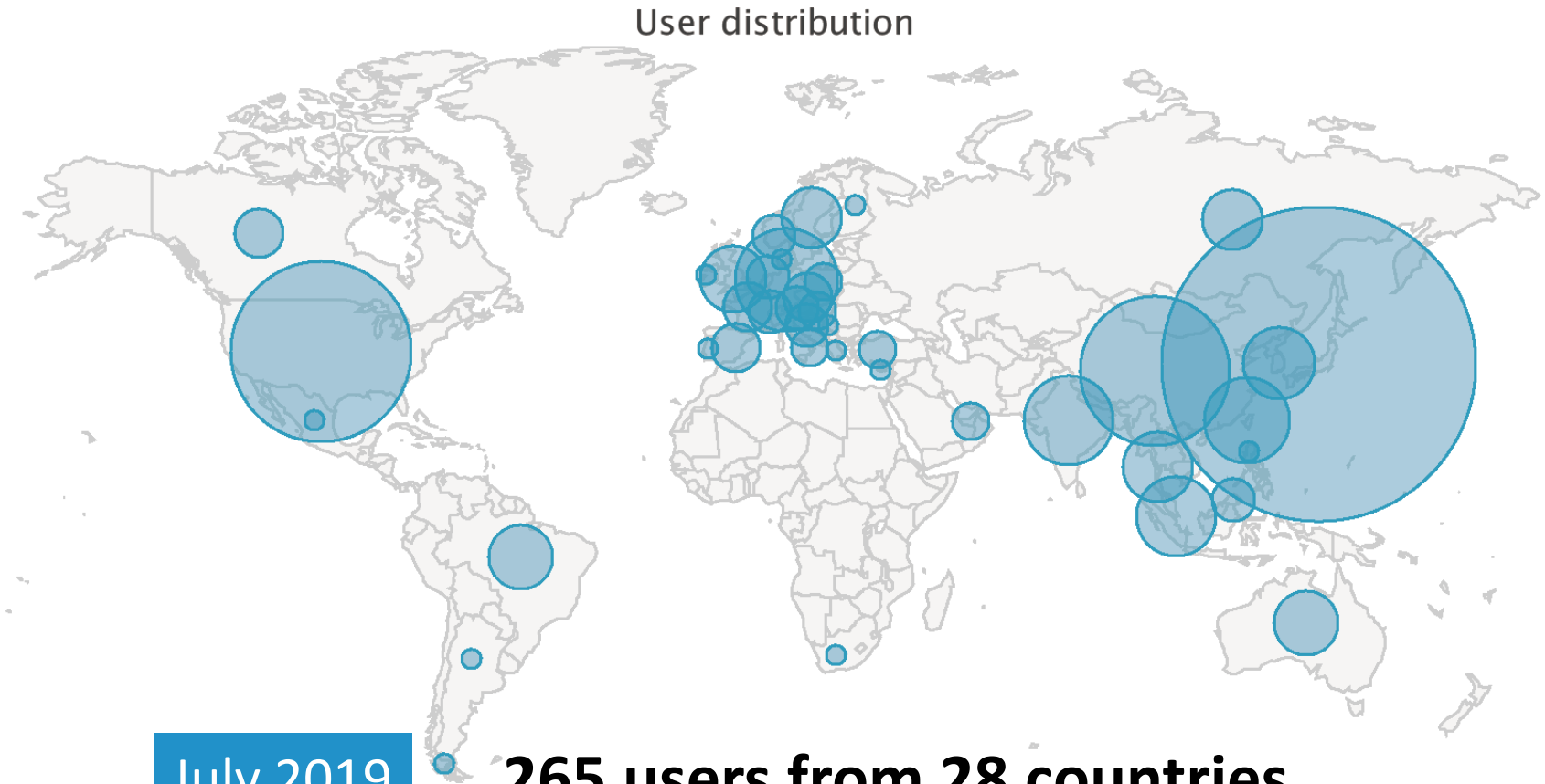
Data Integration & Sharing in Life Science



jPOST repository

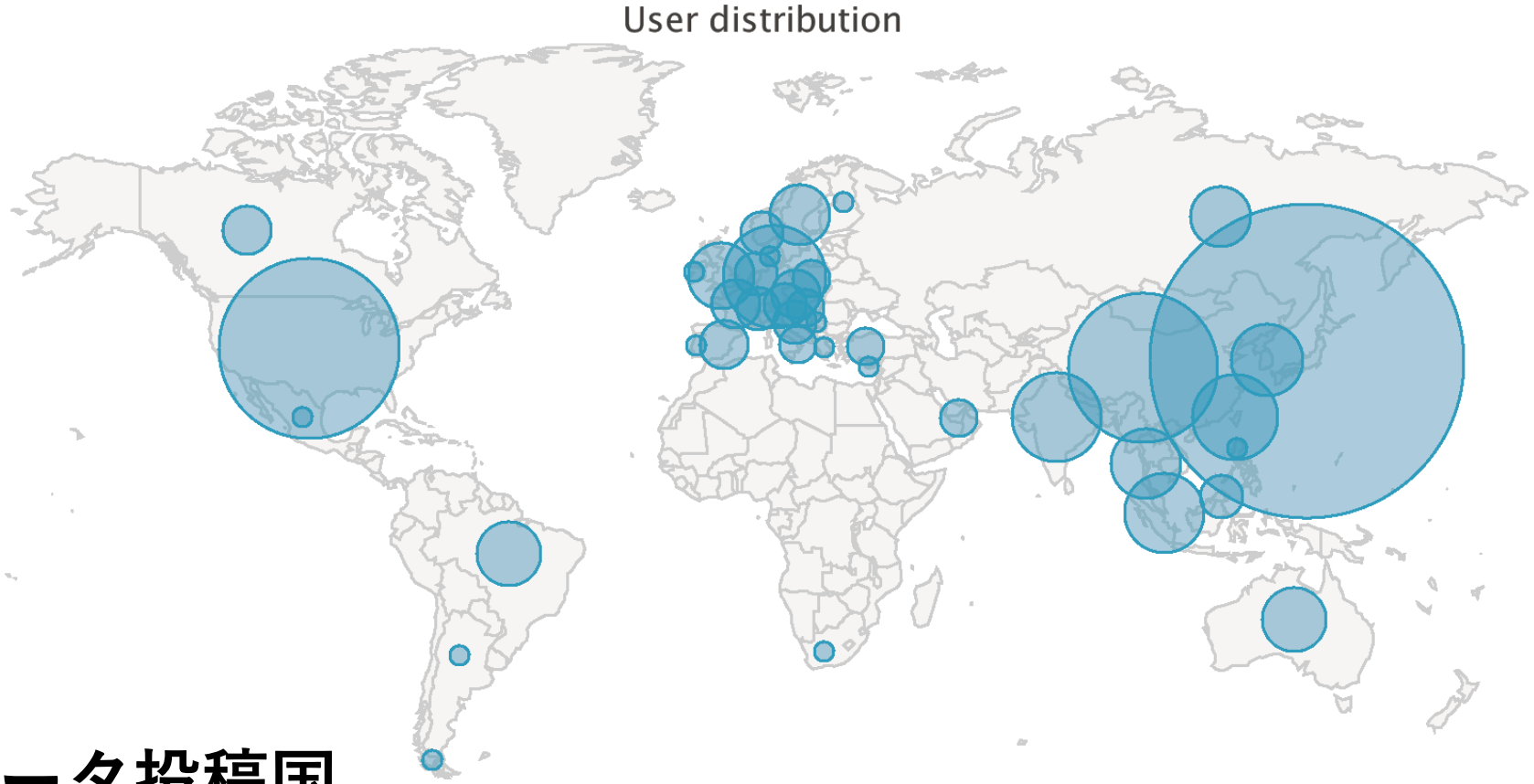


jPOST User Distribution



July 2019	**265 users from 28 countries**
July 2020	**391 users from 37 countries**
Sep 2021	**520 users from 37 countries**
Sep 2022	**682 users from 43 countries**

jPOST Submitter Distribution



データ投稿国

2019年度は日本、アメリカからの投稿がそれぞれ64%, 10%。
2022年度にはそれぞれ13%, 34%となり、急激に国際化が進んでいる。

2022年度のTop 5はアメリカ34%, 日本14%, 中国12%, ドイツ5%, インド3%。

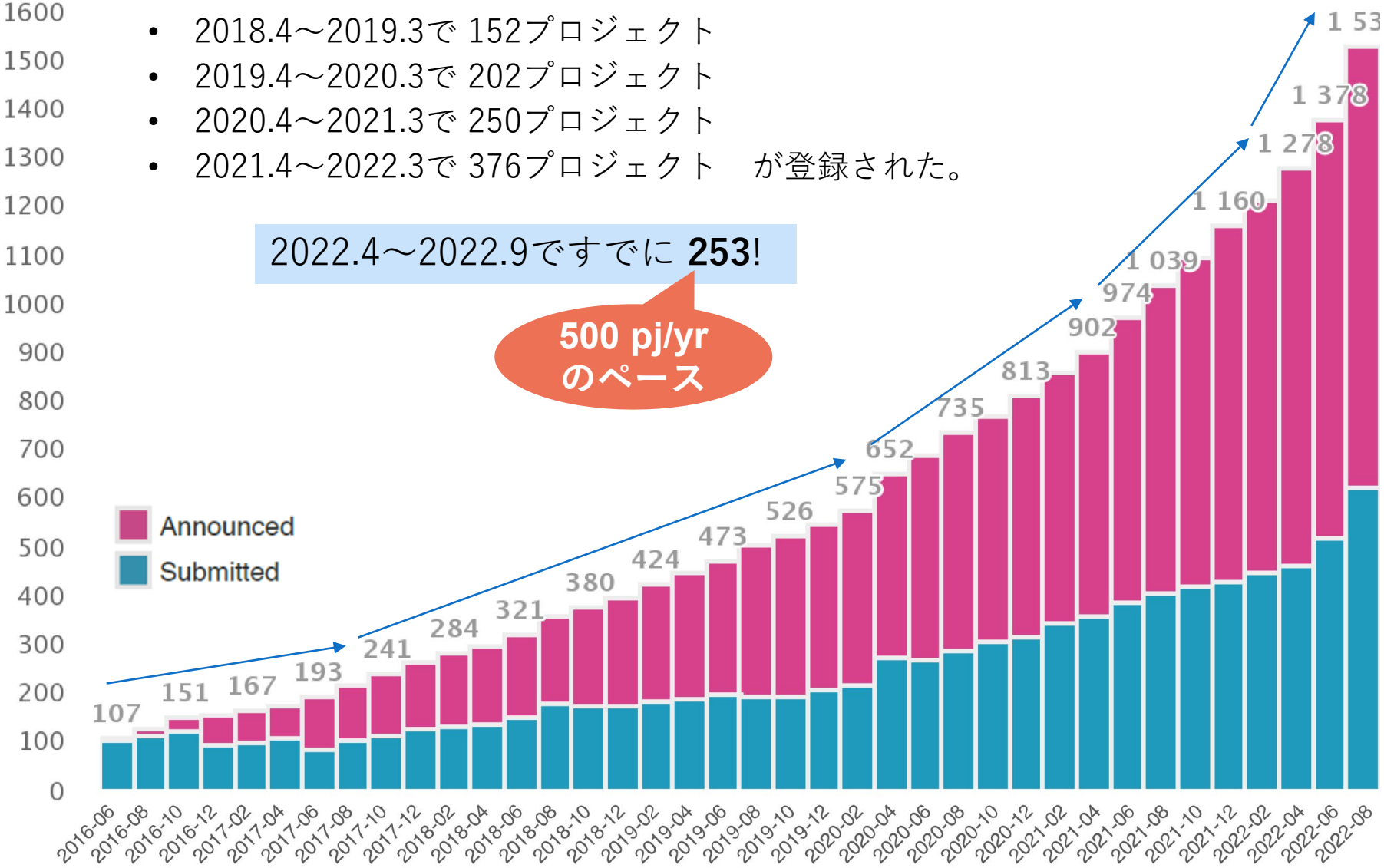
jPOST repository – Current status

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、

- 2017.4～2018.3で 108プロジェクト
- 2018.4～2019.3で 152プロジェクト
- 2019.4～2020.3で 202プロジェクト
- 2020.4～2021.3で 250プロジェクト
- 2021.4～2022.3で 376プロジェクト　が登録された。

2022.4～2022.9ですでに **253**!

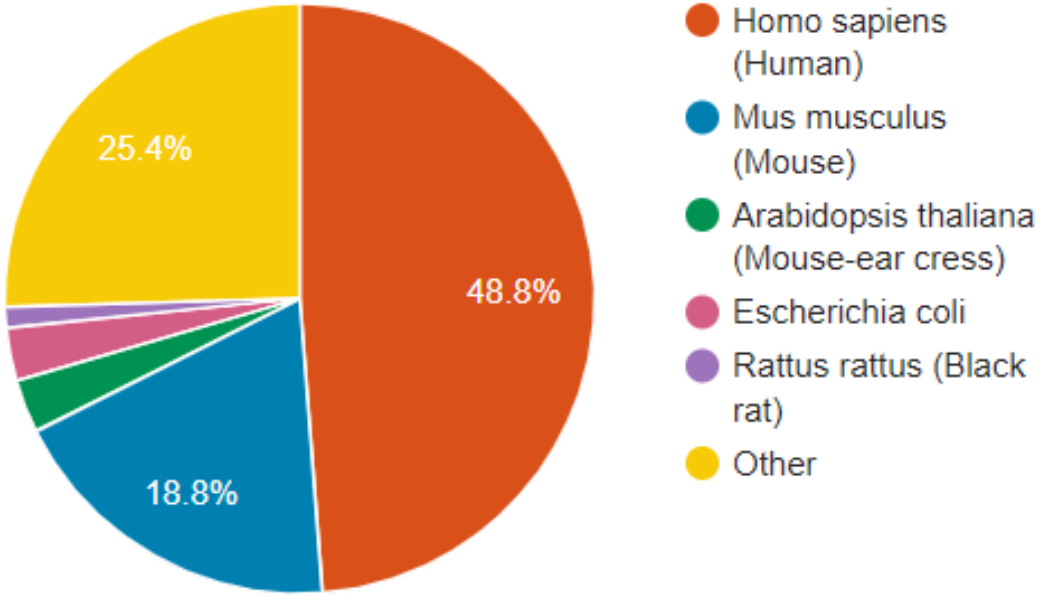
500 pj/yr のペース



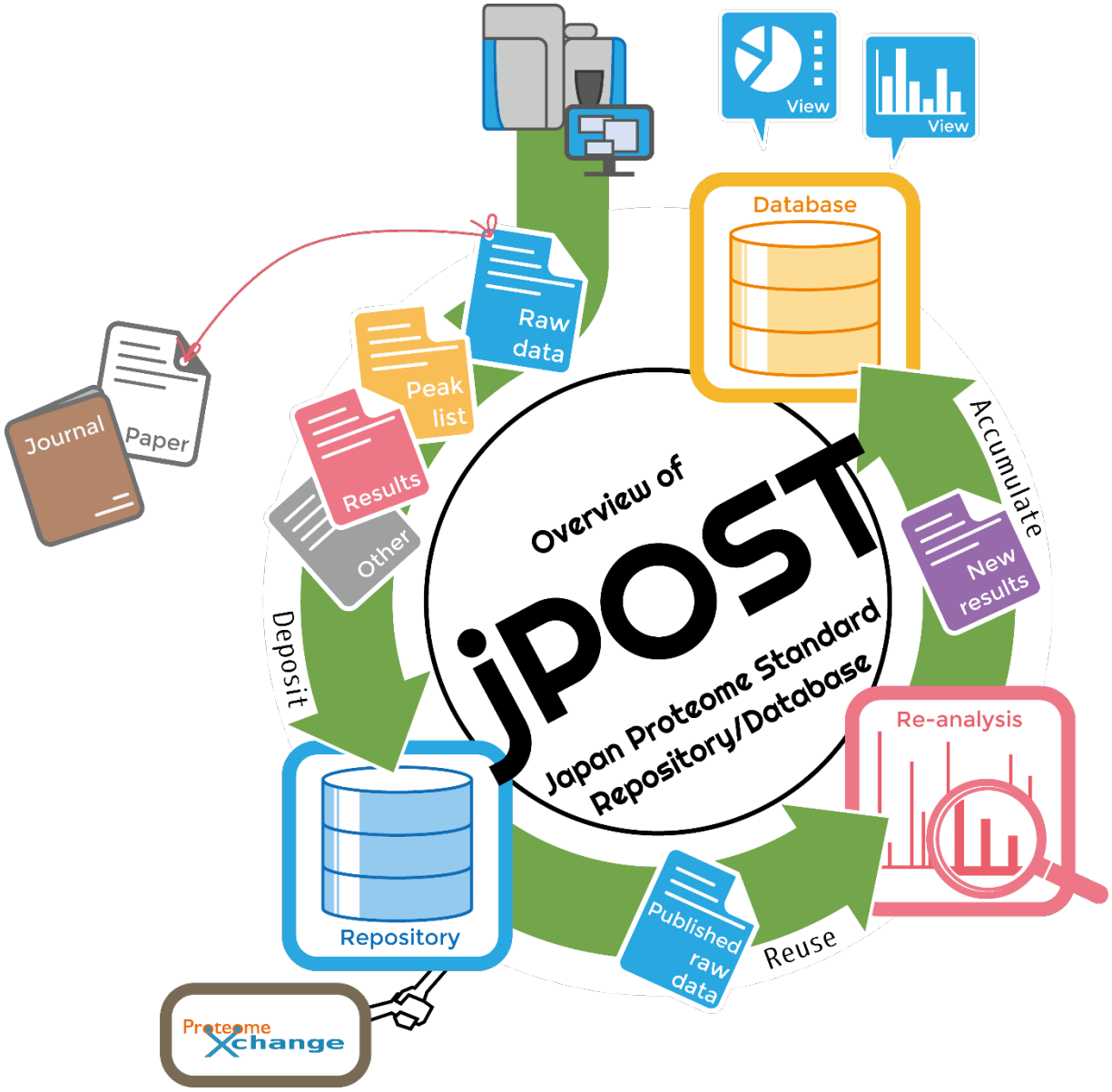
Statistics

Nov 2020	Sep 2021	Sep 2022
798 projects are registered. **479** are opened.	**1050** projects are registered. **643** are opened.	**1587** projects are registered. **934** are opened.
65623 files amount to **26.3 TB**.	**80195** files amount to **35.6 TB**.	**108783** files amount to **55.1 TB**.
126 species.	**172** species.	**225** species.

400 MB/file → 667 MB/file (2020-22)



jPOST Re-Analysis



other

jPOST will publish the re-analyz

2020-07-8 jpost

jPOST will publish the re-analyzed proteome data related to COVID-19 as

Current progress: 100% [Last modified date: 2(

COVID-19 datasets

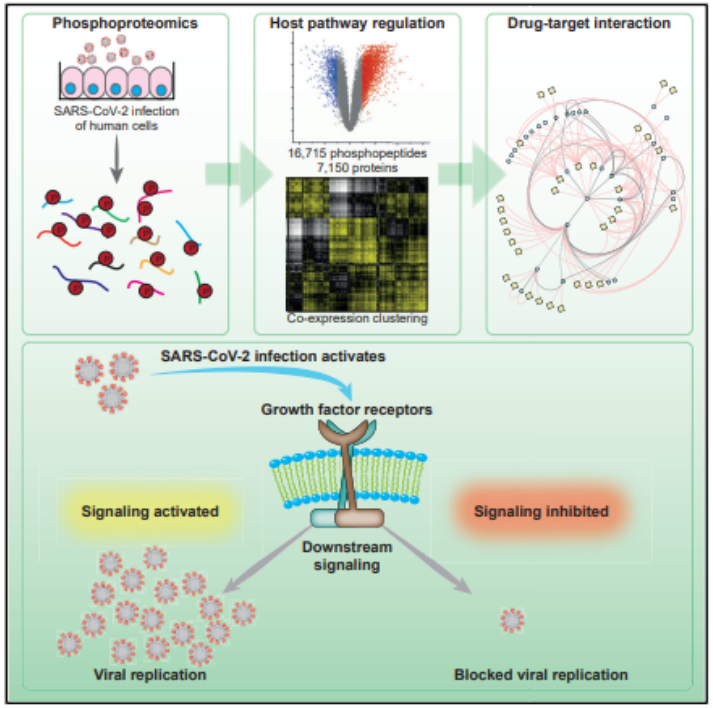
Analysis target	
PXD019113	The Global Phosphorylation Landscape of SARS-CoV-2 Infectio
PXD019645	Data, reagents, assays and merits of proteomics for SARS-CoV-
PXD019423	MS analysis of SARS-CoV2 proteins from patient samples
PXD018804	Extensive proteomic dataset of Vero E6 cells infected by Italy-I
PXD018594	Shotgun proteomics of Vero E6 cells infected by Italy-INMI1 SA
PXD018357	Inhibition of growth factor signaling prevents SARS-CoV-2 repli
PXD018117	A SARS-CoV-2-Human Protein-Protein Interaction Map Reveals
PXD018581	Proteomics of SARS-CoV and SARS-CoV-2 infected cells
PXD018241	Characterisation of the transcriptome and proteome of SARS-C frame deletion in the spike glycoprotein that removes the furin-
PXD017710	Proteome and Translatome of SARS-CoV-2 infected cells

Resource

Molecular Cell

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

Graphical Abstract



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

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.)

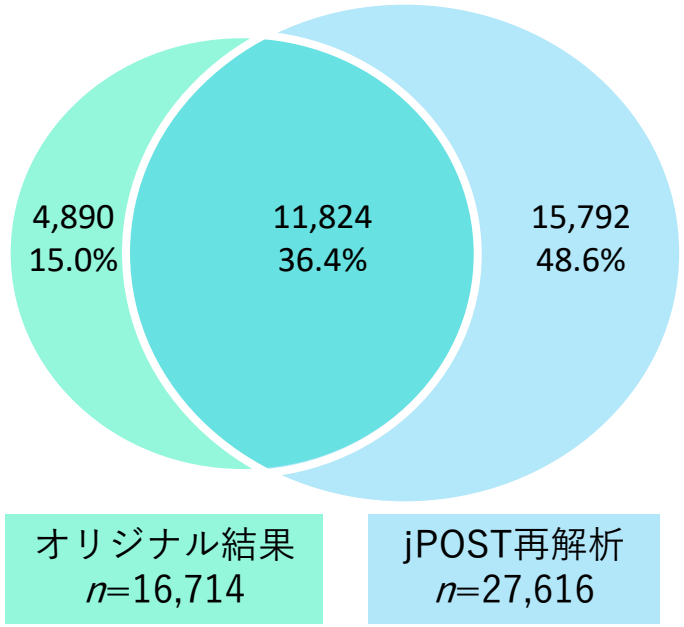
In Brief

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.

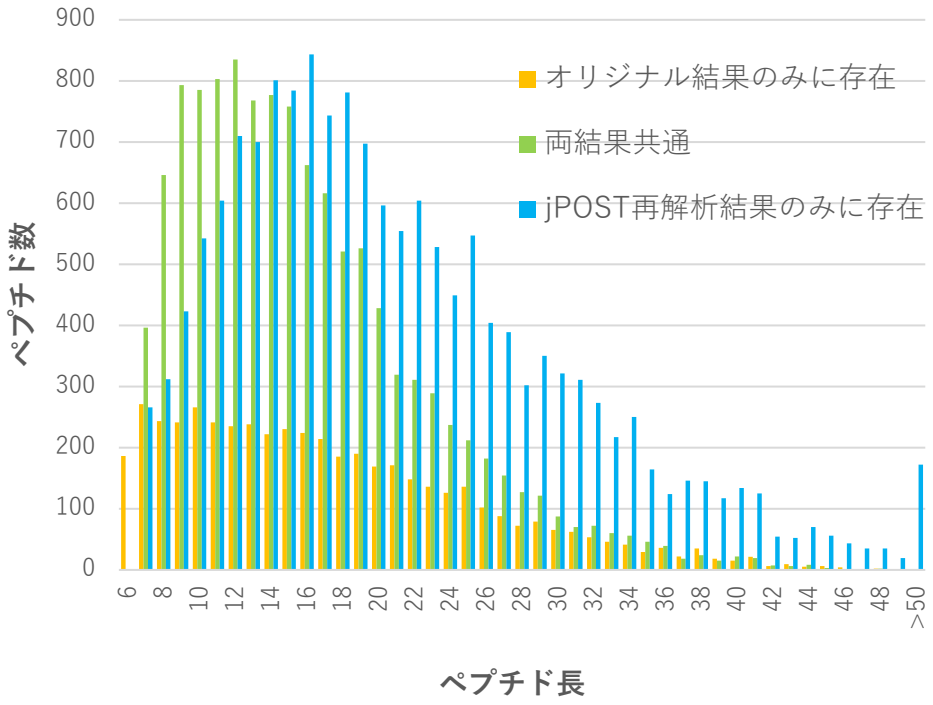
Re-analysis of COVID-19 paper

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

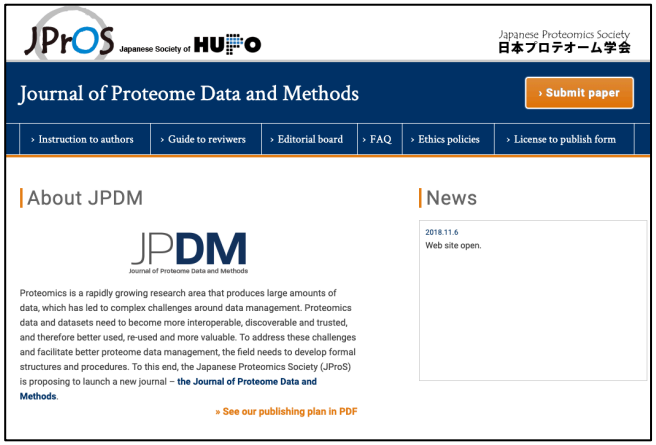
同定リン酸化ペプチド数の比較



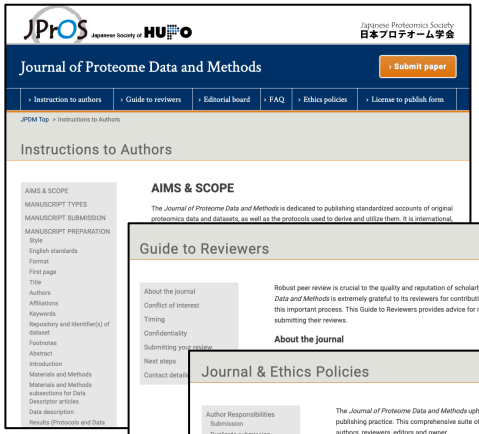
両解析結果のペプチド長ごとの比較



キュレーションの深化と強化
Journal of Proteome Data and Methods (JPDM) 創刊 (2019.9.30)



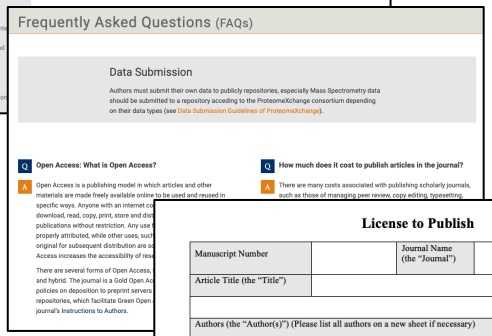
ウェブサイト
https://www.jhupo.org/jpdm/



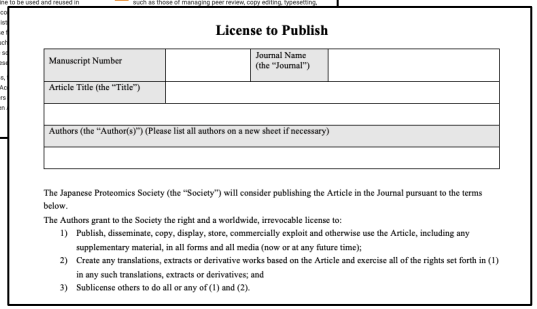
Instructions to Authors

Guide to Reviewers

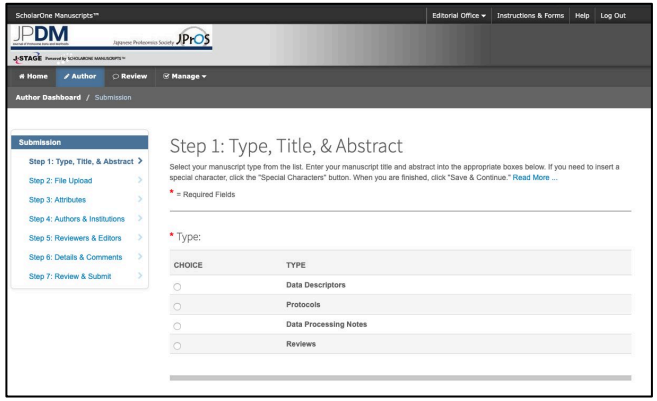
Ethics Policies



FAQ



License to Publish form



投稿システム（ScholarOne）
https://mc.manuscriptcentral.com/jpdm

- 各種ドキュメントの整備
- ウェブサイト構築
- 投稿システム構築
- JST J-stage システムから提供開始

JPDM論文によるメタデータ収集

Sample Article

Data Descriptor

https://dx.doi.org/10.14889/jpdm.2020.xxxx

JPDM Journal of Proteome Data and Methods

Data for proteomic analysis of DNA-binding proteins

Taro Suzuki[1], Jack Smith[1], Sun-Moon Lake[2], Andrew Baker[1*]

1. Laboratory of Proteome Research, National Institute of Biomedicine, Osaka, Japan
2. Graduate School of Science, University of Proteomics, New York, NY, USA

ORCID: Andrew Baker: https://orcid.org/0000-000X-XXXX-XXXX

*Corresponding author.
E-mail address: abaker@nib.go.jp

Keywords

Cell line, DNA-binding, Transcription factor

Dataset summary

Specific subject area	DNA-binding proteins and mass spectrometry
Data acquisition	Data-dependent acquisition acquired on Q-Exactive (Thermo)
Dataset repository	jPOST
Dataset identifiers	JPST100999

Abstract

Interactions between DNA and DNA-binding proteins are required for most cellular processes. Thus, it is clearly important to identify and quantify these interactions for understanding basic cellular mechanisms. We used an oligonucleotide probe to enrich specific DNA-binding proteins. As a result, we identified 512 proteins, including 120 novel DNA-binding protein candidates. The data accompanying this paper have been deposited to jPOST with identifier JPST100999.

1. Materials and Methods

In the present work, we provide the DNA-binding protein catalog obtained by proteomics experiments using the affinity purification with the oligonucleotide probe followed by LC/MS/MS analysis [1].

1.1. Samples

HeLa-S3 cells were grown in DMEM with 10% fetal bovine serum plus antibiotics in 10% CO_2 at 37 °C. For SILAC labeling, HeLa-S3 cells were cultured in DMEM supplemented with 10% dialyzed fetal bovine serum and either 28.0 mg/L normal isotopic abundance arginine and 48.7 mg/L normal isotopic abundance lysine (Light) or 28.0 mg/L arginine with six ^{13}C and four ^{15}N atoms and 48.7 mg/L lysine with six ^{13}C and two ^{15}N atoms (Heavy) [2]. Labeling efficiency was confirmed after five passages.

1.2. Sample pretreatment for MS analysis

Cell lysate was diluted with reaction buffer (20 mM HEPES pH 7.5, 150 mM NaCl, 0.1% Triton X100) to a final protein concentration of 4 mg/mL. $MnCl_2$ was added to 250 µL of lysates at a final concentration of 20 mM. Affinity purification of DNA-binding proteins was carried out at different concentrations of the oligonucleotide probe (1 nM – 10 µM) with SILAC-labeled "Light" lysate at room temperature with gentle shaking for 10 min. The control samples were prepared in the same way using the control probe with SILAC-labeled "Heavy" lysate. After the incubation, these samples were mixed, denatured by 5 M urea, reduced with DTT (5 mM final concentration), and alkylated with iodoacetamide (20 mM final concentration). After the alkylation step, the solution was substituted by digestion buffer by gel filtration followed

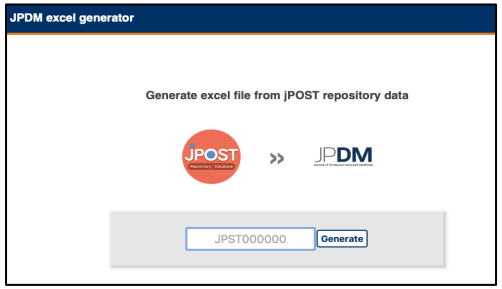
Received 19 June 2099; Received in revised form 7 October 2099; Accepted 15 October 2099; Published 26 October 2099

より**詳細・正確なメタデータ**（Excel形式）を supplementaryとして（必ず）提出する

- Sample 詳細
- 実験の条件
- **機器の詳細**
- Fileの対応関係

手作業での入力

https://repository.jpostdb.org/jpdm-excel/



jPOSTからExcelファイルを**自動生成**

JPDM投稿セミナー開催（3月、8月、11月、1月）

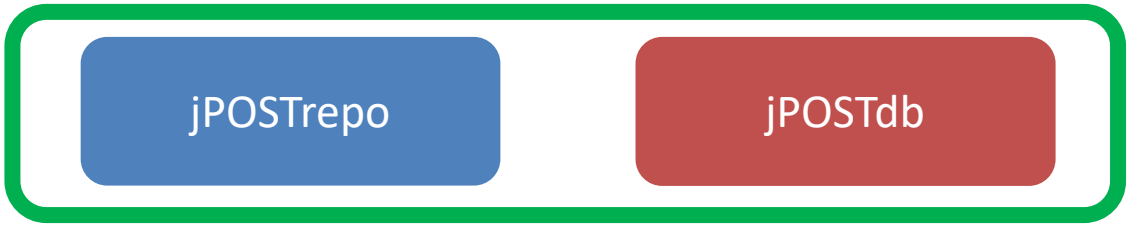
トーゴーの日シンポジウム2019・ポスター35, 同2020・ポスター53

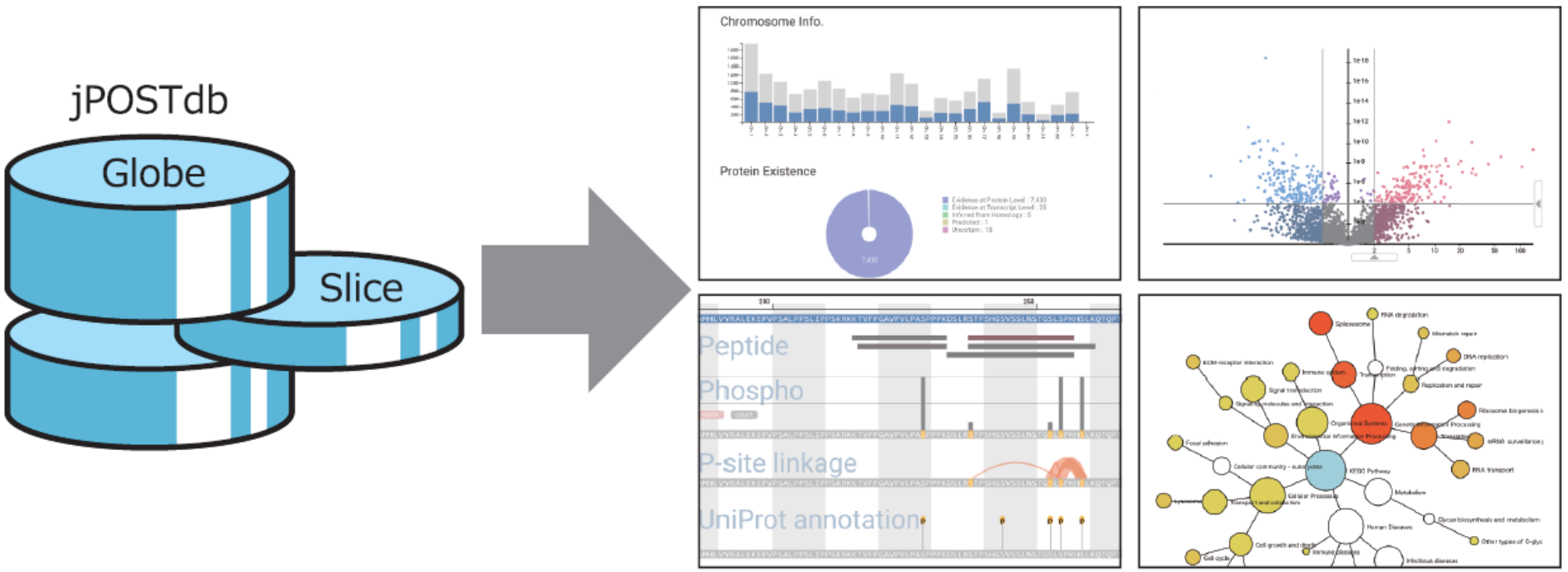
jPOST customizable database ‘Slice’

The jPOST Environment

The jPOST environment: an integrated proteomics data repository and database

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





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

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！

JST NBDC バイオサイエンス データベースセンター

JPOST Repository/Database Japan Proteome Standard Repository/Database

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。