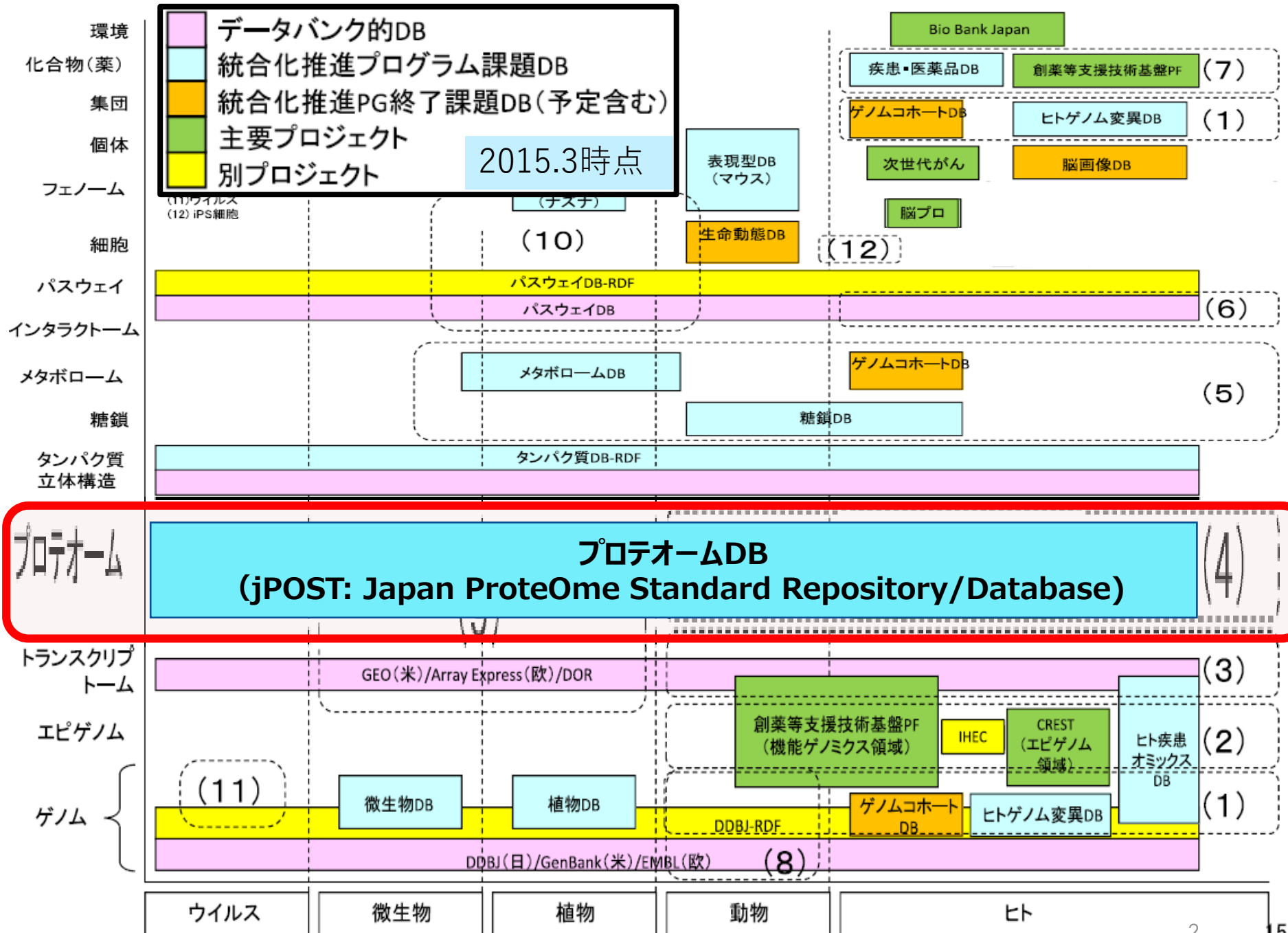




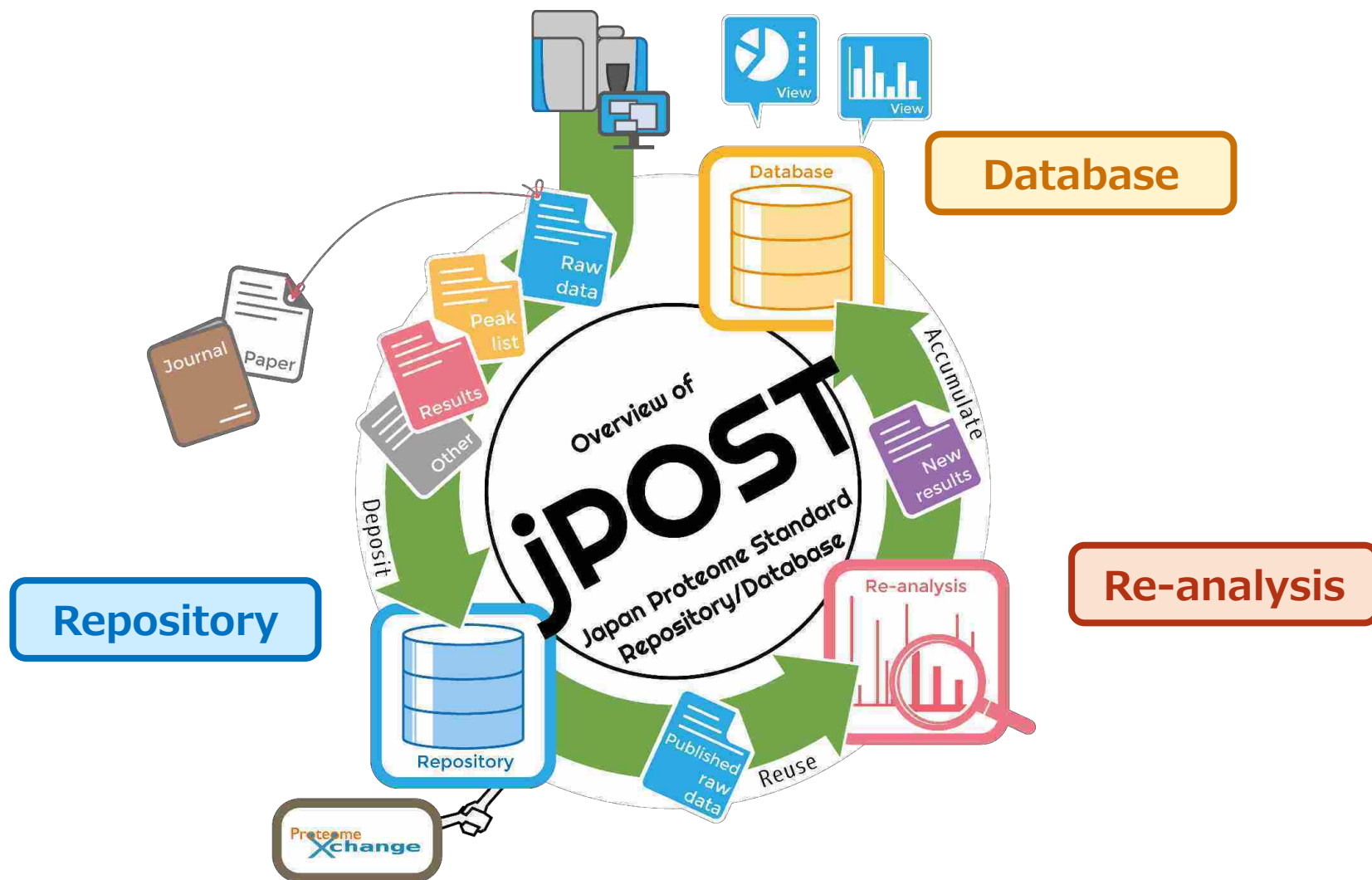
# プロテオームデータベース jPOSTの挑戦

京大院薬 石濱 泰  
& jPOST TEAM





## Data Integration & Sharing in Life Science



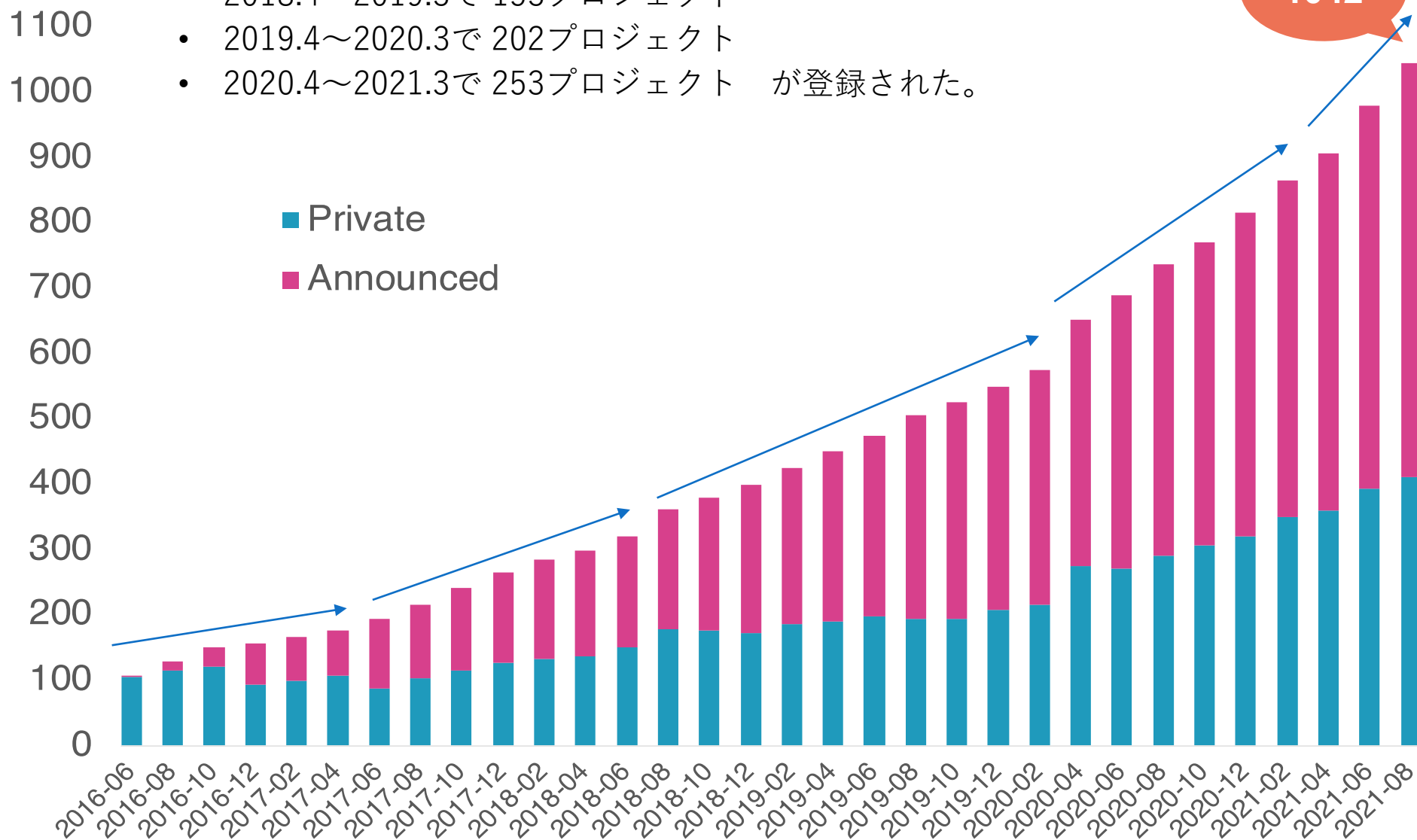
# jPOST repository



# jPOST repository – Current status

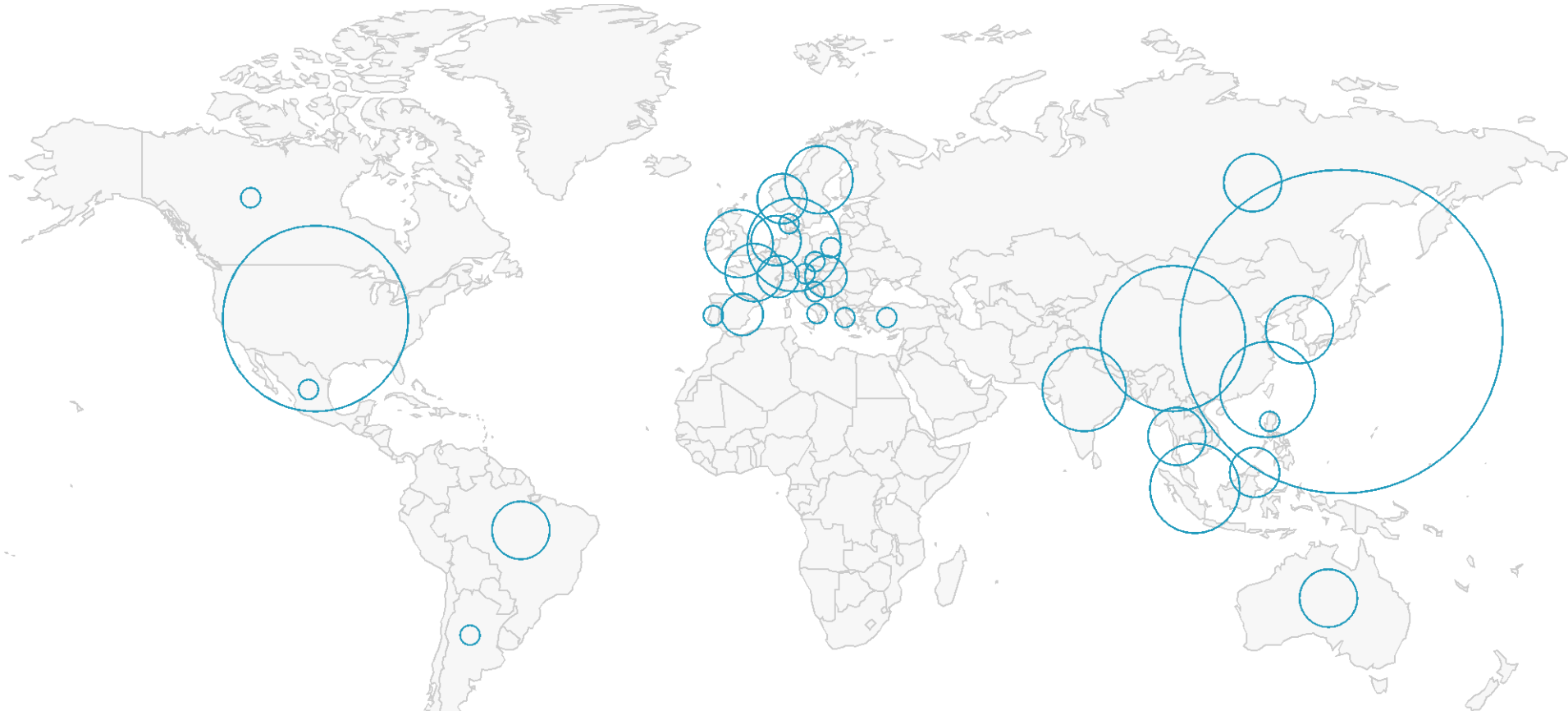
データ登録数は目標としていた年間100プロジェクトを大きく上回り、

- 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。



# jPOST User Distribution

User distribution



- July 2019
- July 2020
- Sep 2021

**265 users from 28 countries**  
**391 users from 37 countries**  
**543 users from 37 countries**

# Statistics

Nov 2020

798 projects are registered. 479 are opened.

65623 files amount to 26.3 TB.

126 species.

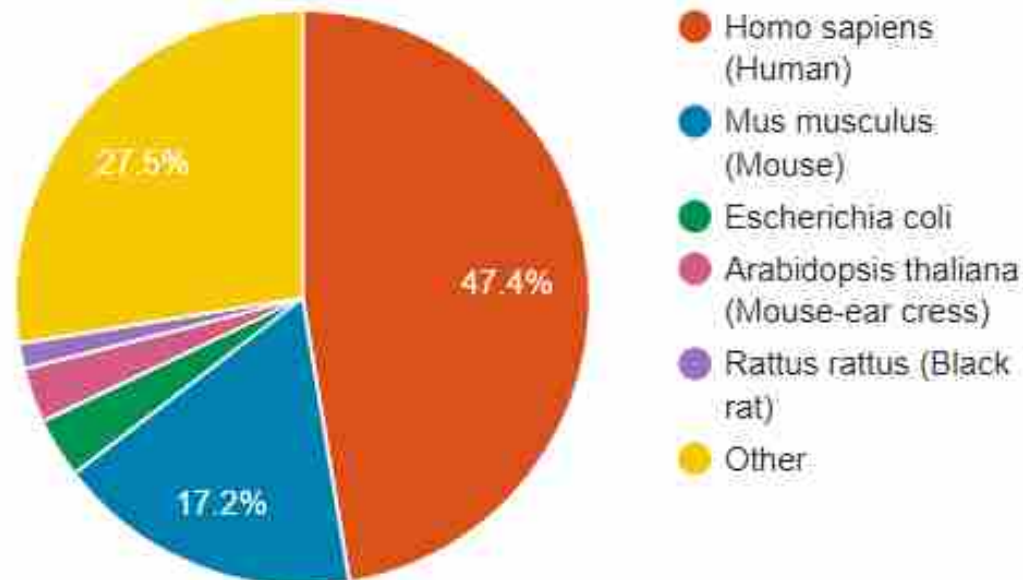


Sep 2021

1050 projects are registered. 643 are opened.

80195 files amount to 35.6 TB.

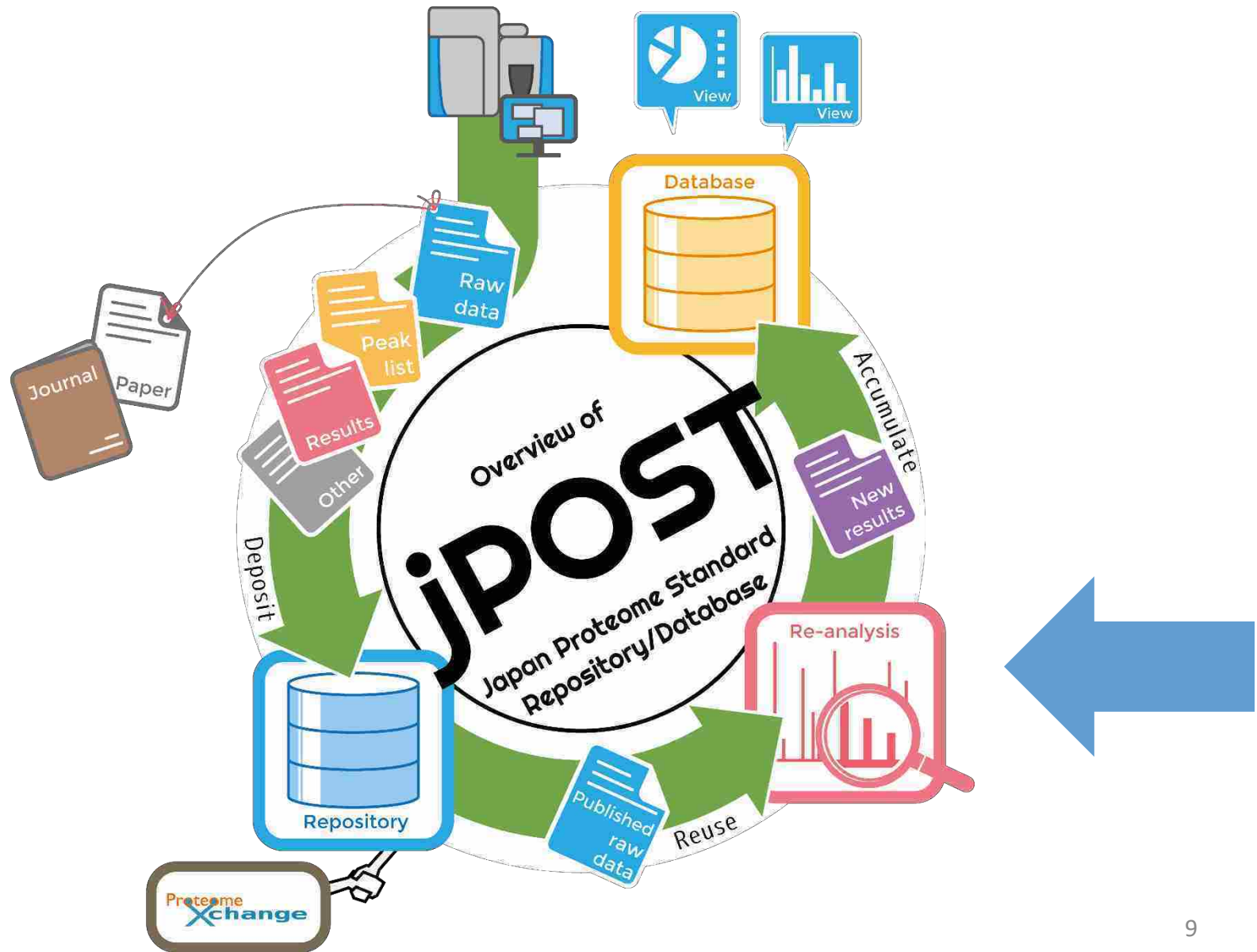
172 species.



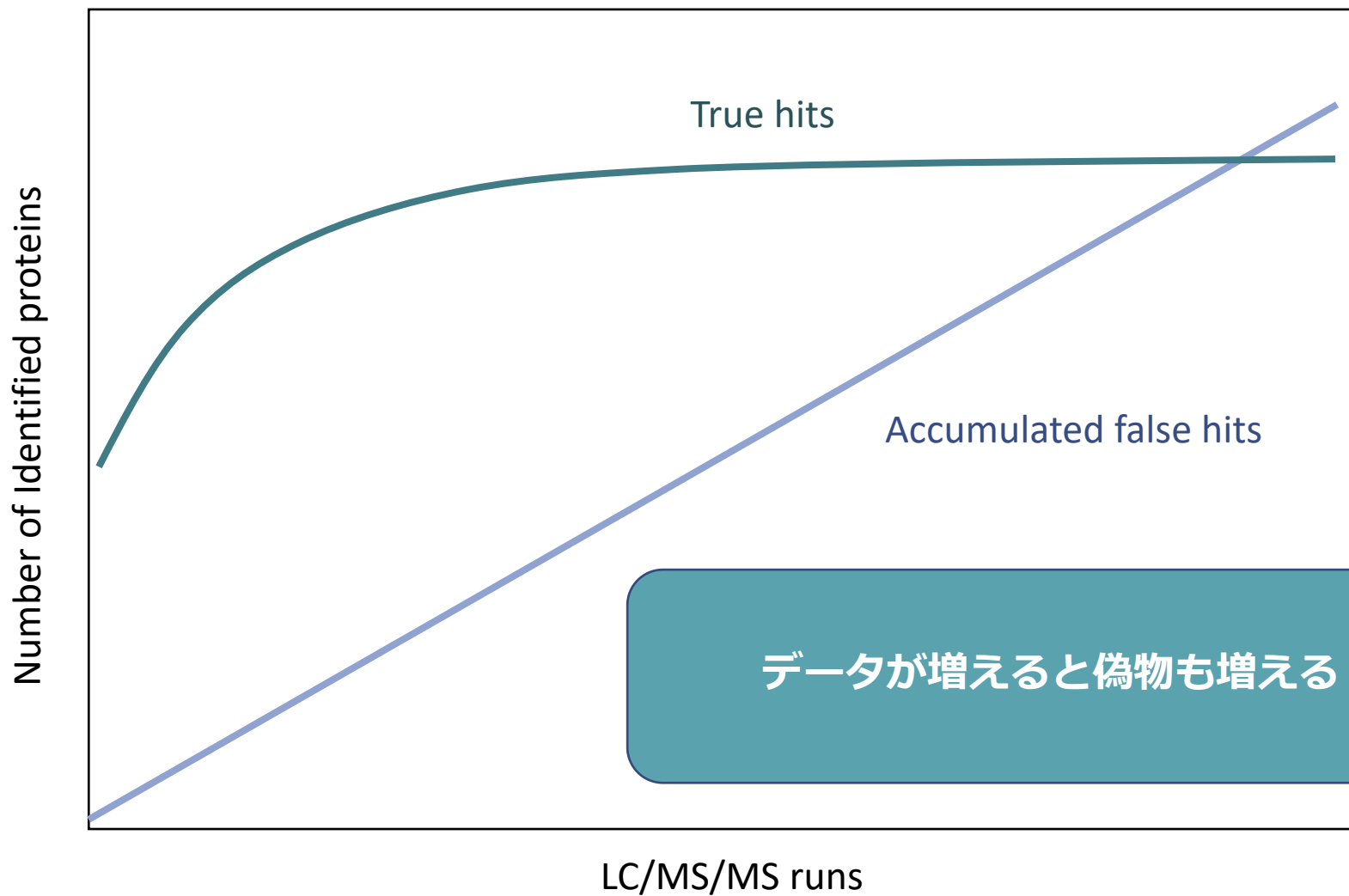
## 1. フレッシュなデータがどんどん勝手にたまる



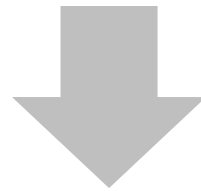
# jPOST Re-Analysis



# 偽ヒットのコントロール

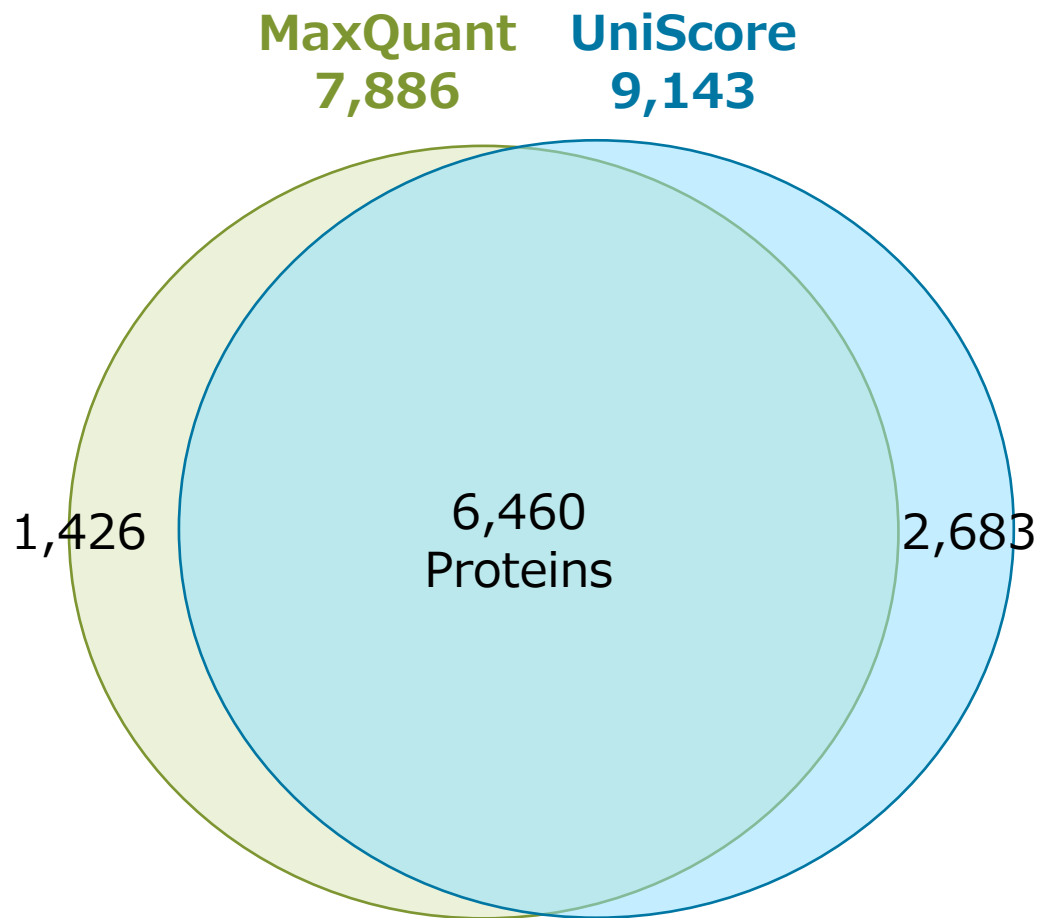


How can we merge the results  
from different sources?



**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

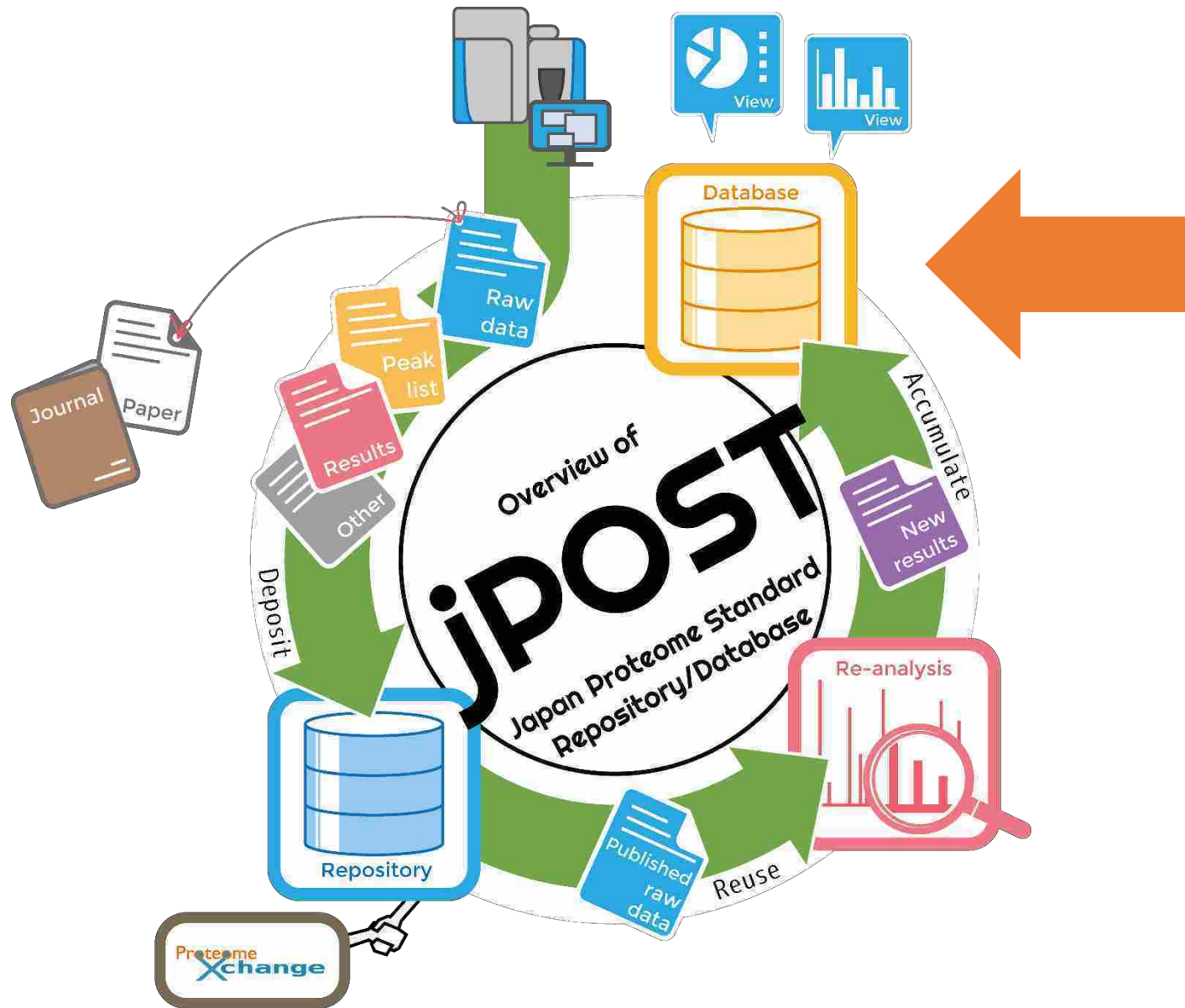


**Dataset:JPST000203 – 206**  
**Colorectal cancer samples (used in Wilhelm et al Nature 2014)**

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

# jPOST customizable database 'Slice'



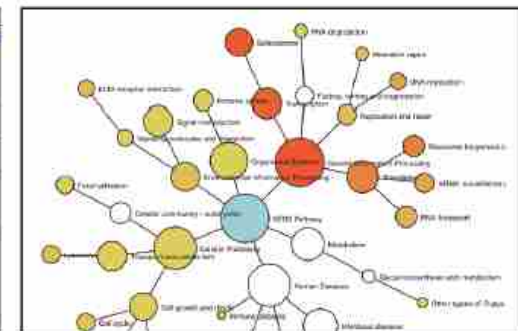
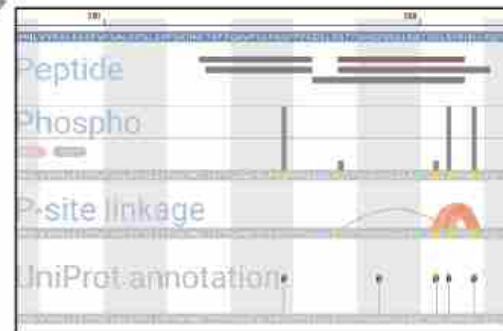
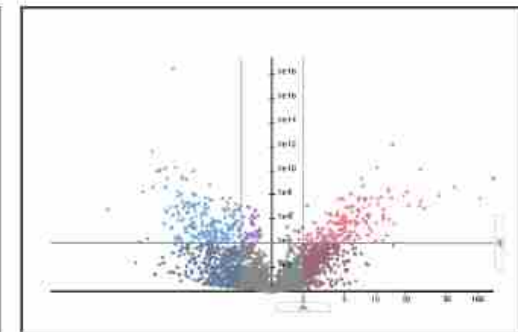
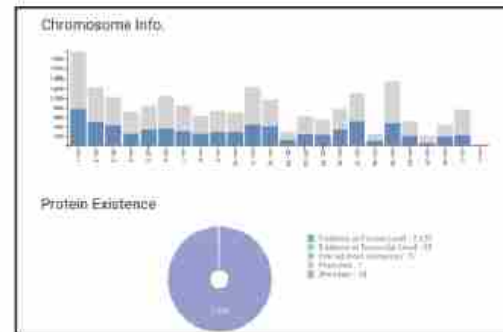
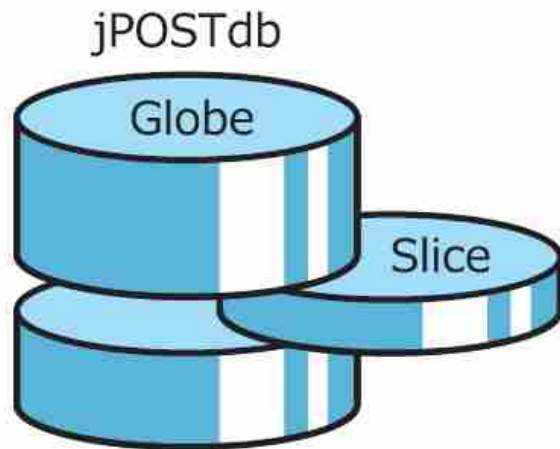
# The jPOST Environment

The jPOST environment: an integrated proteomics data repository and database

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

jPOSTrepo

jPOSTdb



# UniProtとの連携：



## Cross-references between UniProt and jPOST

UniProtKB - P40925 (MDHC\_HUMAN)

Protein: **Malate dehydrogenase, cytoplasmic**

Gene: **MDH1**

Organism: *Homo sapiens (Human)*

Status: Reviewed - Annotation score: ●●●●● - Experimental evidence at protein level<sup>1</sup>

Modified residue<sup>1</sup>: 332 Phosphoserine (By similarity)

Modified residue<sup>1</sup>: 333 Phosphoserine (Combined sources)

Post-translational modification<sup>1</sup>

ISGylated. 1 Publication

Acetylation at Lys-118 dramatically enhances enzymatic activity and promotes adipogenic differentiation. 2 Publications

Keywords - PTM<sup>1</sup>: Acetylation, Methylation, Phosphoprotein, Ubl conjugation

Proteomic databases

EPD <sup>1</sup>	P40925
jPOST <sup>1</sup>	P40925
PaxDb <sup>1</sup>	P40925
PeptideAtlas <sup>1</sup>	P40925
PRIDE <sup>1</sup>	P40925
ProteomicsDB <sup>1</sup>	55384
	55385 [P40925-2]
TopDownProteomics <sup>1</sup>	P40925-1 [P40925-1]

Protein: P40925

Protein Name: Malate dehydrogenase, cytoplasmic

Protein ID: MDHC\_HUMAN

Gene Name: MDH1

Accession: P40925

Length: 334 aa

Sequence: MSEPIRELVTGAAGGIAYSLLSYIGNSGVFGKDDPIILVLLDITPPRRVLDGLVMEIQQCALPFLKQVZATDKEDVAFKDLQVALVLSMPREGEKEDLLYAWKFLKSGDRAALDNYAPKSNVILVYDAPARTWELTASKSAFLIPERSELTHLDNRARQALKLSGTANDVKNVTLWGNHSSLYPPDVAHWKLVKLGKEVGVYKAKKQVSNLRIGEPVTTVQDRGAAYLRKRLSSAASAAKACDHRVQIDVFGTPEGEFVSMGVLSZGNSYGVDDLLYSFPVVKIKKTKPVEGLPINDPSREKMDLAKELTEEKESAFELSSA

Location: Chromosome 2

Statistics

# peptides:	42
# spectra:	1,834
# unique peptides (UniProt entry level):	31
# unique peptides (gene name level):	42

for Abdolet

GeneID: 4190

Quantification: search in IMPACT

Protein Browser

Peptide Sharing

shared in isoforms: 1



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

→ ↻ 🔒 https://jpostdb.org



[About](#) [Repository](#) [Database](#) [Workflow](#) [Gadgets](#) [COVID-19](#) [Help](#) [FAQ](#) [Contact](#)



Japan Proteome Standard  
Repository/Database



## Repository [Help](#)

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



## Database [Help](#)

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

## 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



## Workflow [Help](#)

This provides the re-analysis workflow used in jPOST.



## Gadgets [Help](#)

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



## 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.



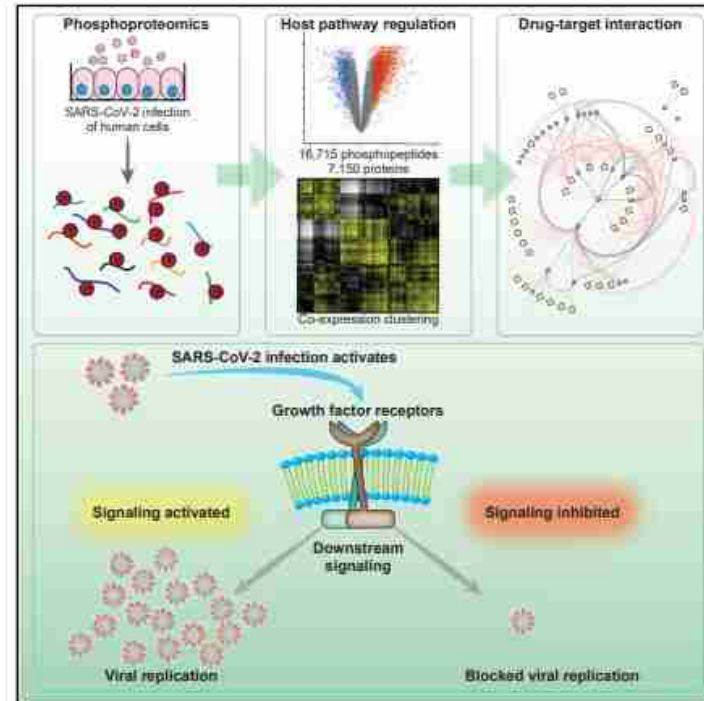
## Help

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 downstream signaling prevent viral replication

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: 20

### COVID-19 datasets

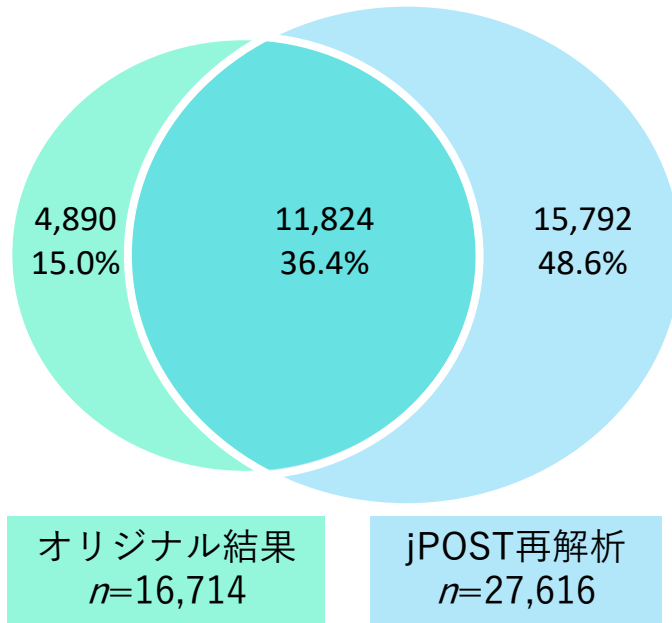
#### Analysis target

- PXD019113 The Global Phosphorylation Landscape of SARS-CoV-2 Infection
- 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
- PXD017710 Proteome and Translatome of SARS-CoV-2 infected cells

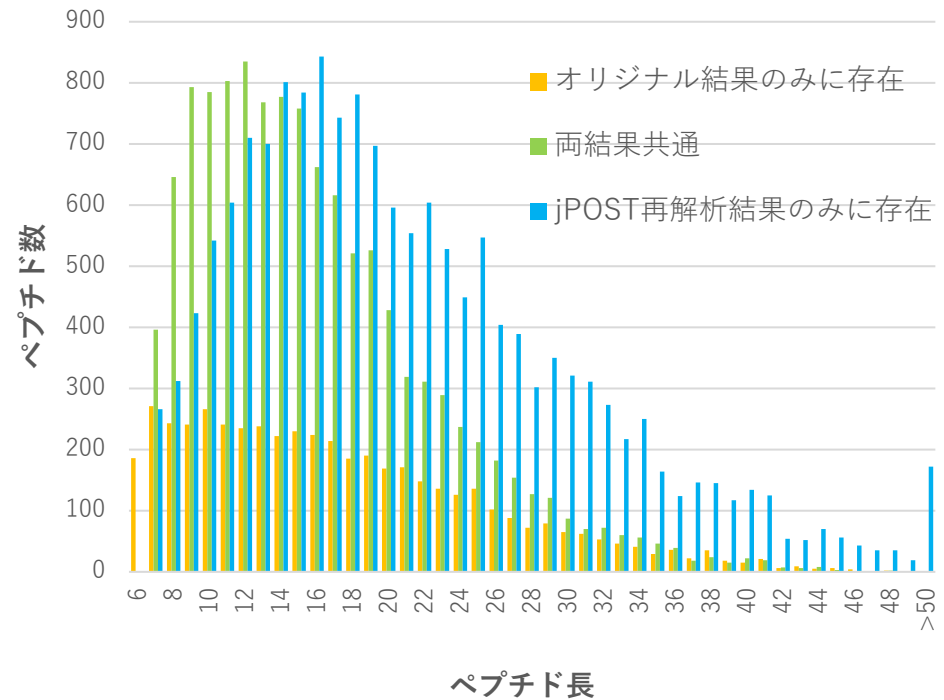
# Re-analysis of COVID-19 paper

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

## 同定リン酸化ペプチド数の比較



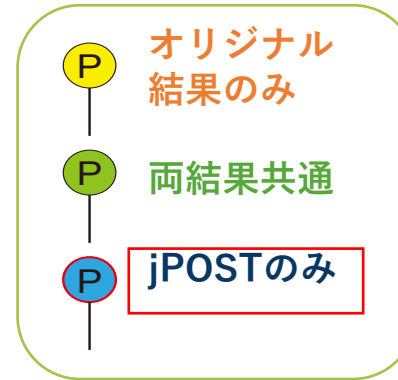
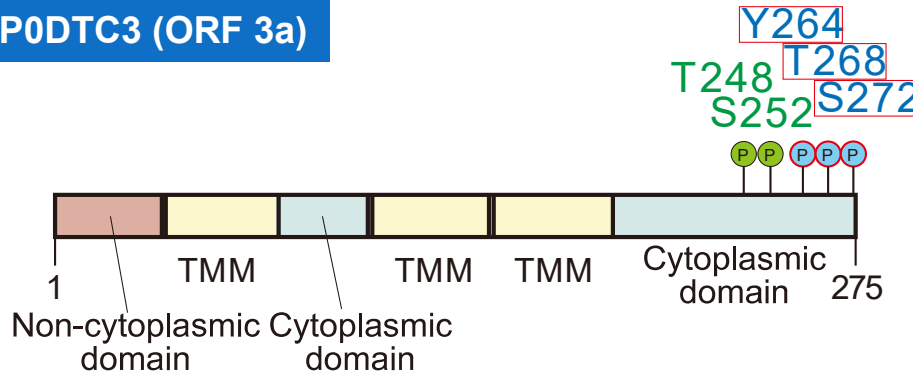
## 両解析結果のペプチド長ごとの比較



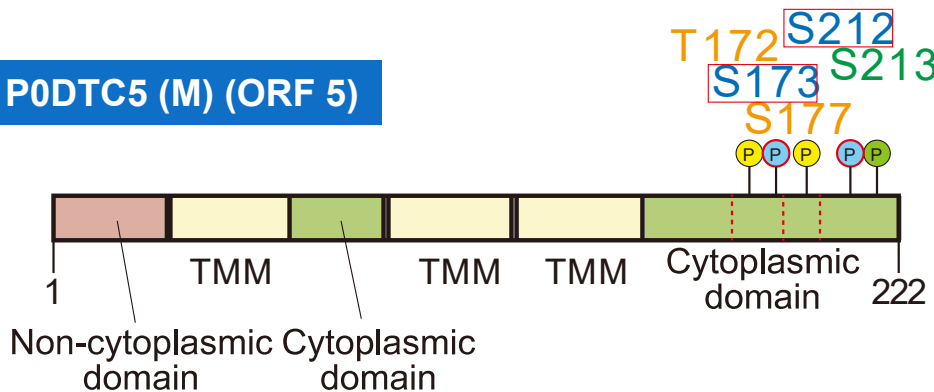
# Phosphorylated sites on SARS-CoV-2 proteins

## 同定されたリン酸化サイトの比較

### P0DTC3 (ORF 3a)



### P0DTC5 (M) (ORF 5)



T172 vs S173  
 ↓ ↓  
 EITVATSR

b+	#	Seq	#	y+
359.2128	1	E	8	
472.2968	2	I	7	827.4022
573.3445	3	T	6	714.3182
672.4129	4	V	5	613.2705
743.4500	5	A	4	514.2021
844.4977	6	T	3	443.1650
1011.4960	7	S	2	342.1173
	8	R	1	175.1190

[Click] to move table

Add to N-term: 229.1629

Variable Modifications:

S: 79.9663 [7]

# 使ってください jPOST

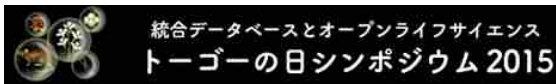
提案：データサイエンス研究者とデータベース研究者の融合



“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！





  バイオサイエンス  
データベースセンター



Japan Proteome Standard  
Repository/Database

面白いアイデ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。