

ライフサイエンス統合データベースセンター (DBCLS) の活動紹介

情報・システム研究機構
データサイエンス共同利用基盤施設
ライフサイエンス統合データベースセンター (DBCLS)
五斗 進

極域データサイエンスに関する研究集会 2023年2月20日

ライフサイエンスデータベース統合推進事業

FAIR 原則*をベースとしたバイオデータベース（特に日本のデータ）の利活用促進

■ データベースの所在を明らかにする（Findable）

- Integbioデータベースカタログ（NBDC）：約2500件のDB

■ データベースにアクセス可能にする（Accessible）

- データベース横断検索（NBDC）：780DB
- データベースアーカイブ（NBDC）：153DB
- 世界各地のデータベースセンター：米国NCBI、欧州EBI、スイスSIBなど

■ データベースを相互参照（Interoperable）、再利用（Reusable）可能にする

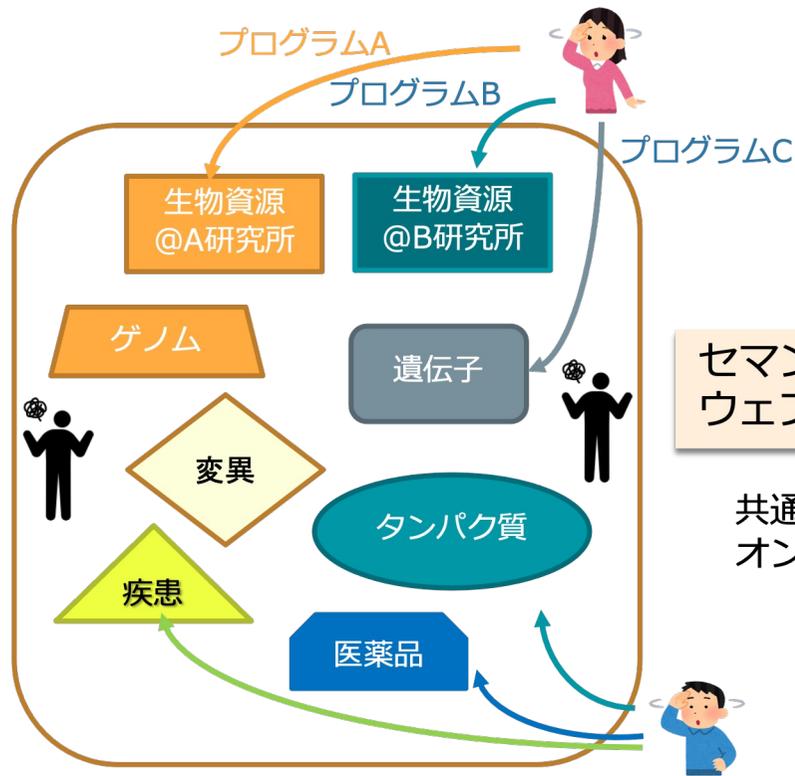
- データベースをダウンロードすると使える形になっている。
- データベース間で使用する用語やフォーマットを統一して使えるようにする。
- **DBCLS では、この部分を中心に基盤技術開発として推進している。**

* データ共有の基準としてのFAIR原則。Wilkinson, et al. (2016) Scientific Data.₂

データベース統合がなぜ必要か？

各データに固有のプログラムでアクセス
その結果をまとめる必要あり

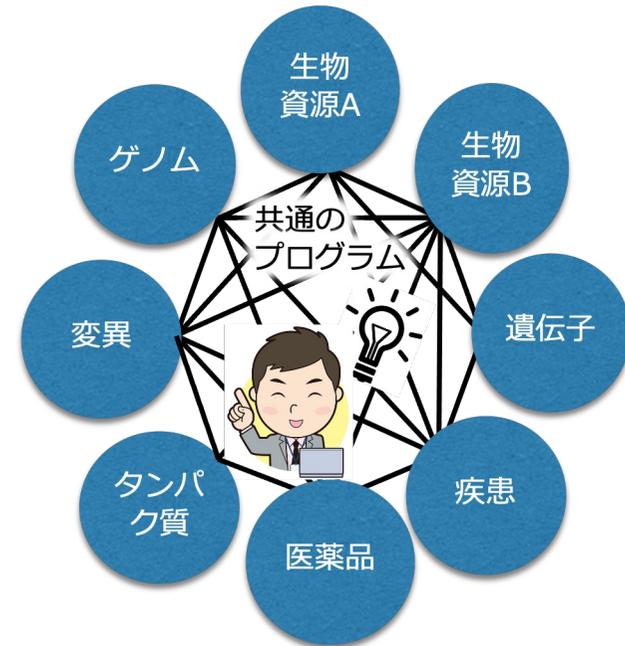
共通プログラムでアクセス
関連情報の取得も容易



バラバラのデータベース

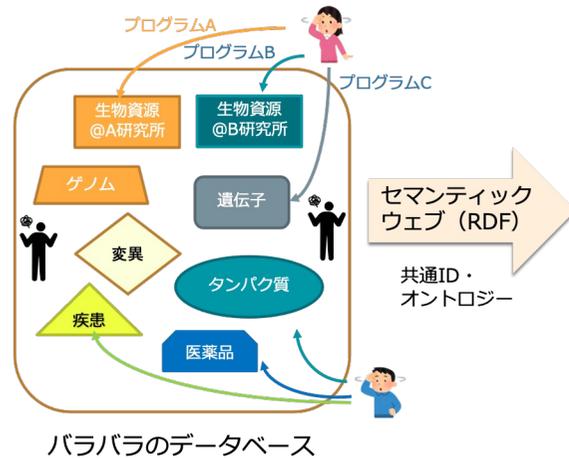
セマンティック
ウェブ (RDF)

共通ID・
オントロジー

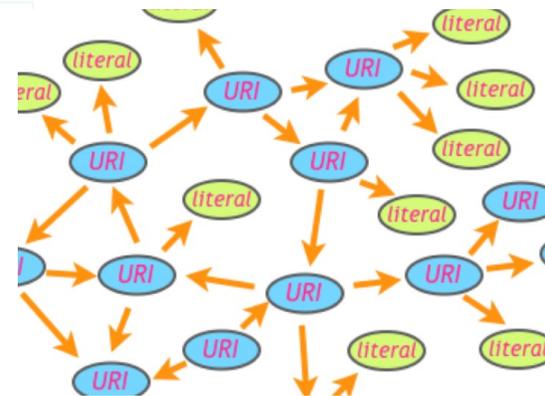
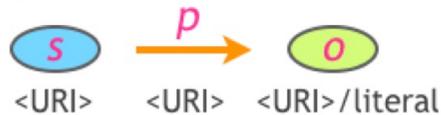


整理・関連付けがなされた
(統合化) データベース

DBCLS/NBDCにおけるデータベース統合

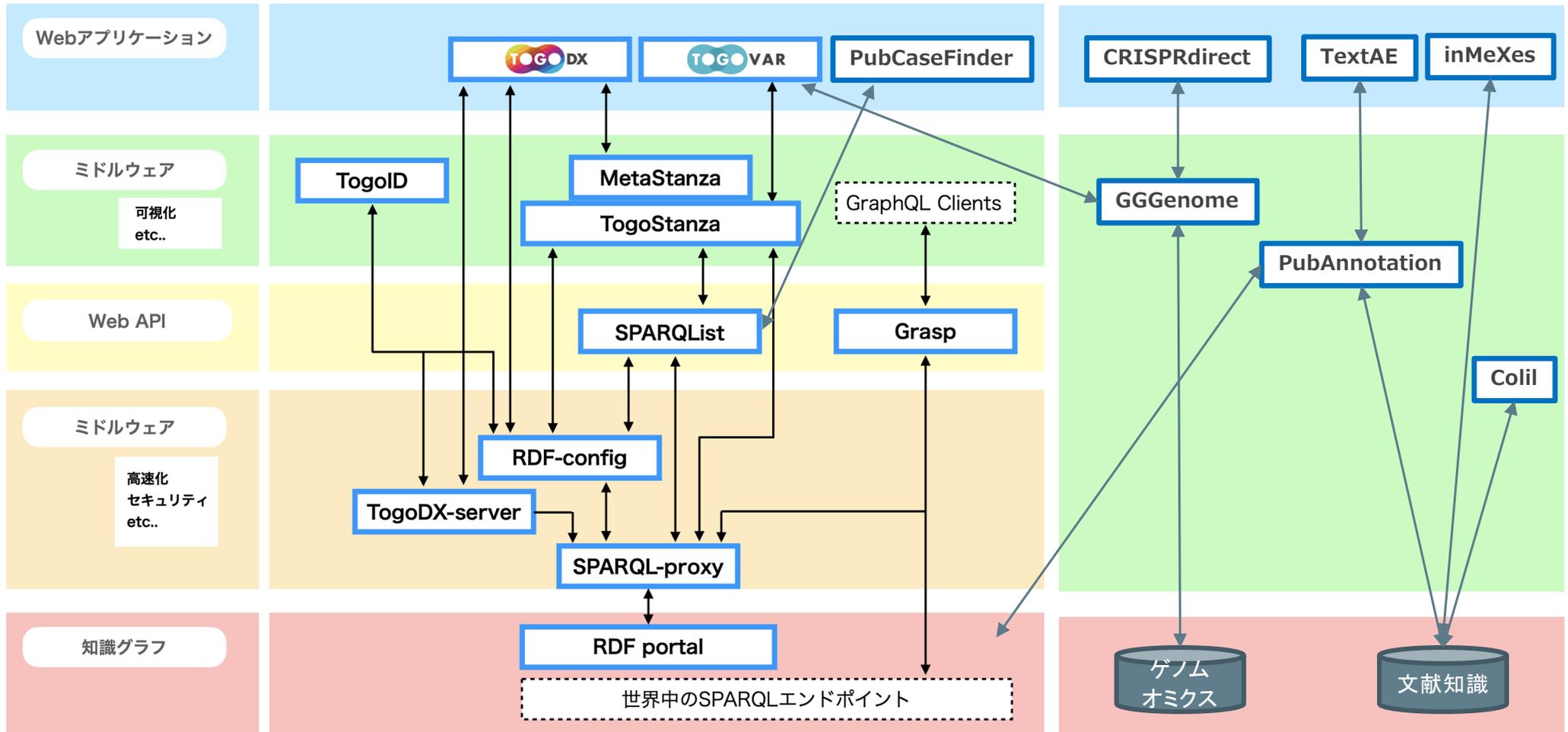


- **RDF**: Resource Description Framework
- **主語、述語、目的語**からなるデータモデル
 - **主語**:モノ(データ)の ID (URI)
 - **述語**:**オントロジー**で定義された属性 (URI)
 - **目的語**:別のモノの ID (URI) または値 (Literal)



知識グラフ化

統合化実現のために開発した基盤技術

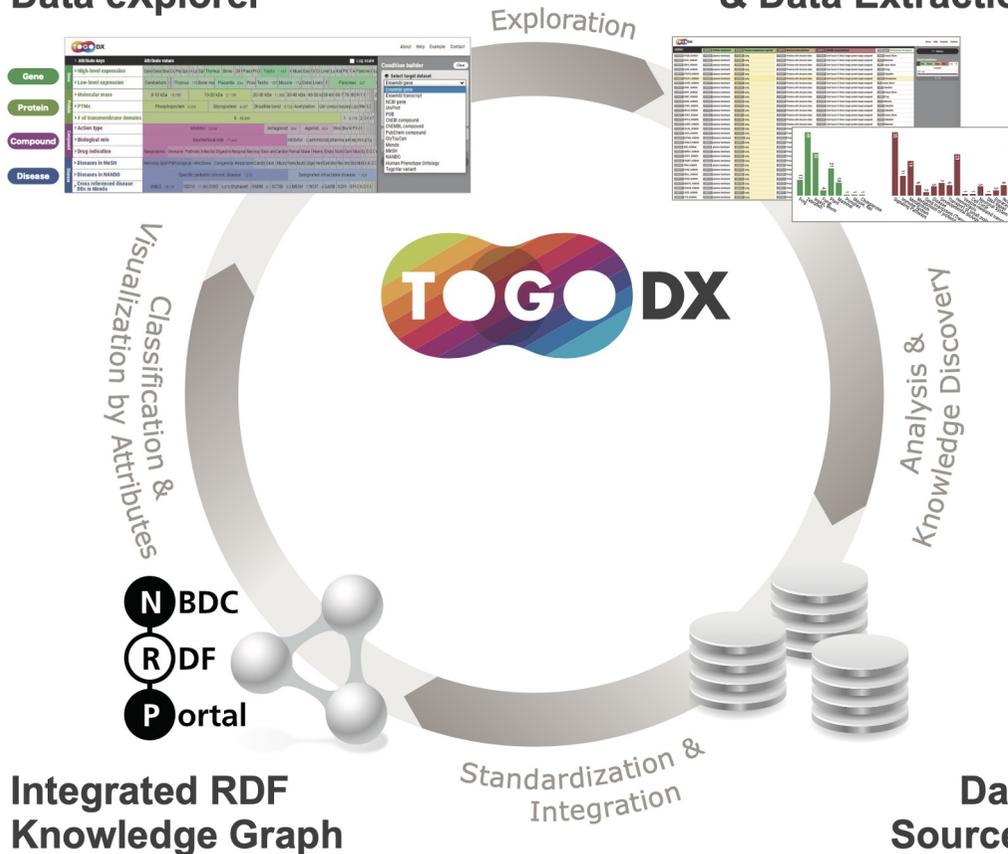


青枠はNBDC/DBCLSが開発したものの

TogoDX (Togo Data eXplore): 知識グラフの俯瞰・探索

Togo Data eXplorer

Grasping the summary & Data Extraction



さまざまな機関で作成されたフォーマットの異なるデータベースを

- 知識グラフとして標準化・統合化
- データ間のリンク情報も整理

し、データベースを横断的に利用可能にするフレームワーク

<https://togodx.dbcls.jp/human/>

TogoDX (Togo Data eXplorer) のコンセプト

- これまで統合してきたデータをどのように横断的に利用するか
 - NBDC, DDBJ, 統合化推進プログラムと作ってきた知識グラフの統合的活用事例
 - 膨大かつ多様なデータを俯瞰したい
 - 複雑なグラフの中から必要なデータを組み合わせて抽出したい
 - 抽出したデータを機械学習・AIなどデータサイエンスで活用したい
 - 統合されたデータはあらゆる生物種の様々な生命科学現象に広範囲に渡る
- 技術的なポイント
 - SPARQL: 各カテゴリにおける様々な視点からの全データ俯瞰
 - TogoID: IDの相互変換によるカテゴリ間の統合
 - フルスタックの技術開発
 - データ → 知識グラフ → RDF-configによるスキーマ → SPARQLクエリ生成
 - SPARQLListによるAPI化 → Stanzaで可視化 → TogoDXでデータフレーム化

TogoDXで知識グラフを俯瞰する

- 遺伝子
- タンパク質
- タンパク質
立体構造
- 相互作用
- 化合物
- 糖鎖
- 疾患
- バリエント

The screenshot displays the TogoDX interface with a table of attribute values and a 'Conditions Builder' sidebar. The table is organized into categories: Gene, Protein, Structure, Interaction, Compound, Glycan, Disease, and Variant. Each category has several rows of data with columns representing different tissues or conditions. The 'Conditions Builder' sidebar on the right shows a list of conditions, with an arrow pointing to a selected item.

選択したデータベースのIDを軸にデータセットを作成可能

TogoDXデータ属性ごとの俯瞰

例えば

タンパク質分子の機能

結合	触媒反応	調節	輸送	転写制御	...
----	------	----	----	------	-----

各タンパク質と相互作用するタンパク質の数

0	1	2	3	..
---	---	---	---	----

病気の大分類

臓器系の機能障害	メンデル遺伝病	身体領域による疾患分類	発生過程の疾患	癌	...
----------	---------	-------------	---------	---	-----

TogoDXでは、各データ属性の俯瞰から、分類や分布の条件を自由に組み合わせて、データの探索を行うことができる。

The screenshot shows the TogoDX interface with a sidebar of category filters and a main table of attribute values. The filters include: 遺伝子 (Gene), タンパク質 (Protein), タンパク質立体構造 (Structure), 相互作用 (Interaction), 化合物 (Compound), 糖鎖 (Glycan), 疾患 (Disease), and バリエント (Variant). The main table lists attribute keys and their corresponding values across various tissues and conditions.

Attribute keys	Attribute values
Gene	> High-level expression: Cere, Cere, Brai, Oc, Pei, Spi, V, Ly, Sp > Low-level expression: Cerebellum, Thymus, 17, Bo
Protein	> Molecular mass: 0-10 kDa, 15, 150 > PTMs: Phosphoprotein, 5, 266 > # of transmembrane domains
Structure	> # of peptides in a PDB entry > Other related molecules: Water, 42, 851
Interaction	> Compounds in pathway: Metabolism > # of interacting proteins
Compound	> Action type: Inhibit > Biological role: bioch > Drug indication: Neoplasms, Immune, Patholo, Ir
Glycan	> Molecular mass of glycans: 20, 400-600, 600-800, 800-1000, 1 > Tissues: colon, 11, 4
Disease	> Diseases in MeSH: Nervous Syst Pathological, Infect > Diseases in NANDO: Specific ped > Cross referenced disease DBs in Mondo: UMLS, 18, 101; ICD10, 11
Variant	> GWAS > Clinical significance: Uncertain significance, 29

TogoDX探索条件(Add filters)

細胞内局在として、細胞膜を選択

cellular anatomical entity	protein-containing com	Unclassified
cellular anatomical entity 39,165	protein-containing com 6,589	Unclassified 34,572
cellular anatomical entity 39,165	plasma membrane 17,308	
protein-containing complex 9,147	organelle membrane 4,508	
intracellular component of membrane 13,506	organelle membrane 4,508	
membrane 12,534	plasma membrane region 1,302	
unclassified 34,572	membrane microdomain 364	
	cytosol 7,640	
	cytoplasm 7,038	
	nucleoplasm 6,333	
	phagosome assembly site membrane 20	
	extracellular region 2,673	
	cell projection 2,232	
	cell junction 2,236	
	extracellular space 1,925	
	supramolecular complex 1,491	
	membrane-enclosed lumen 1,396	
	nuclear outer membrane-endoplasmic reticulum membrane network 2	

GTEX の組織別発現データで、肺と小腸を選択

Expressed in tissues	Low specificity	Testis	Brain - (Brain - Brain Brain Brain Pitu) Brain Brain Brain Sple Bra Bra Cel Bra Br Sk Sr Lu THN Es Li P KIM Y W O C C A F N A C L S E I L L
Expressed in tissues	Low specificity 21,661	Testis 12,404	Brain - (Brain - Brain Brain Brain Pitu) Brain Brain Brain Sple Bra Bra Cel Bra Br Sk Sr Lu THN Es Li P KIM Y W O C C A F N A C L S E I L L

PDBで、立体構造既知のタンパク質を選択

Structure data existence	Proteins without structure data	Proteins with struct
Structure data existence	Proteins without structure data 70,821	Proteins with struct

ChEMBLで、信頼性スコア9のものを選択

ChEMBL assay existence	Conf-sc Conf	No ChEMBL assay
ChEMBL assay existence	Conf-sc Conf	No ChEMBL assay 74,473

これらの選択した条件を全て満たす IDのリストを検索する。



Condition builder Clear

Select target dataset
UniProt

Map your IDs optional Try

e.g. A0A0A0MQR4, P36888-1, P25054-1, 075874, P42336, Q9Y6K1-1, P21860-1

Submit Reset

Add filters

- Protein / Cellular component
 - plasma membrane ×
- Structure / Structure data existence
 - Proteins with structure data ×
- Interaction / ChEMBL assay existence
 - Conf-score 9: Direct single protein target ass ×
- Gene / Expressed in tissues
 - Lung ×
 - Small Intestine - Terminal Ileum ×

Map attributes

View results

The screenshot shows the TogoDX search results page. The top section displays a table of search results with columns for 'Values', 'Total', and 'Values 2'. Below this, there are several filter sections: 'Cellular component', 'Expressed in tissues', 'Structure data existence', and 'ChEMBL assay existence'. The 'Cellular component' filter is set to 'plasma membrane'. The 'Expressed in tissues' filter is set to 'Lung' and 'Small Intestine - Terminal Ileum'. The 'Structure data existence' filter is set to 'Proteins with structure data'. The 'ChEMBL assay existence' filter is set to 'Conf-score 9: Direct single protein target ass'. The search results table shows a list of protein IDs and their corresponding values across different filters.

TogoDX探索結果(Add filters)

UniProt	UniProt	Cellular component	ensembl_gene	Tissue-specific high expression (GTEx)	UniProt	Structure data existence	UniProt	ChEMBL assay existence	Mondo	Diseases in Mondo
P51684	CCR6_HUMAN	plasma membrane	ENSG00000112486	Small Intestine - Terminal Ileum	P51684	Proteins with structure data	P51684	Conf-score 9: Direct single protein target assigned		
O76083	PDE9A_HUMAN	plasma membrane	ENSG00000160191	Small Intestine - Terminal Ileum	O76083	Proteins with structure data	O76083	Conf-score 9: Direct single protein target assigned		
P22732	GTR5_HUMAN	plasma membrane	ENSG00000142583	Small Intestine - Terminal Ileum	P22732	Proteins with structure data	P22732	Conf-score 9: Direct single protein target assigned		
P29972	AQP1_HUMAN	plasma membrane	ENSG00000240583	Lung	P29972	Proteins with structure data	P29972	Conf-score 9: Direct single protein target assigned		
P55087	AQP4_HUMAN	plasma membrane	ENSG00000171835	Lung	P55087	Proteins with structure data	P55087	Conf-score 9: Direct single protein target assigned		
O14649	KCNK3_HUMAN	plasma membrane	ENSG00000171309	Lung	O14649	Proteins with structure data	O14649	Conf-score 9: Direct single protein target assigned		
O14746	TEXT_HUMAN	plasma membrane	ENSG00000164362	Small Intestine - Terminal Ileum	O14746	Proteins with structure data	O14746	Conf-score 9: Direct single protein target assigned		
O14788	TNF11_HUMAN	plasma membrane	ENSG00000120659	Small Intestine - Terminal Ileum	O14788	Proteins with structure data	O14788	Conf-score 9: Direct single protein target assigned		
O15296	LX15B_HUMAN	plasma membrane	ENSG00000179593	Lung	O15296	Proteins with structure data	O15296	Conf-score 9: Direct single protein target assigned		
O43451	MGA_HUMAN	plasma membrane	ENSG00000257335	Small Intestine - Terminal Ileum	O43451	Proteins with structure data	O43451	Conf-score 9: Direct single protein target assigned		
O60603	TLR2_HUMAN	plasma membrane	ENSG00000137462	Lung	O60603	Proteins with structure data	O60603	Conf-score 9: Direct single protein target assigned		
O75015	FCG3B_HUMAN	plasma membrane	ENSG00000162747	Lung	O75015	Proteins with structure data	O75015	Conf-score 9: Direct single protein target assigned		
O75769	TRPA1_HUMAN	plasma membrane	ENSG00000104321	Small Intestine - Terminal Ileum	O75769	Proteins with structure data	O75769	Conf-score 9: Direct single protein target assigned		
O95977	S1PR4_HUMAN	plasma membrane	ENSG00000172910	Lung	O95977	Proteins with structure data	O95977	Conf-score 9: Direct single protein target assigned		
P01589	IL2RA_HUMAN	plasma membrane	ENSG00000134460	Small Intestine - Terminal Ileum	P01589	Proteins with structure data	P01589	Conf-score 9: Direct single protein target assigned		
P01730	CD4_HUMAN	plasma membrane	ENSG00000106610	Lung	P01730	Proteins with structure data	P01730	Conf-score 9: Direct single protein target assigned		
P01833	PIGR_HUMAN	plasma membrane	ENSG00000162906	Lung	P01833	Proteins with structure data	P01833	Conf-score 9: Direct single protein target assigned		
P01909	DQA1_HUMAN	plasma membrane	ENSG00000196735	Lung	P01909	Proteins with structure data	P01909	Conf-score 9: Direct single protein target assigned		
P02647	APOA1_HUMAN	plasma membrane	ENSG00000118137	Small Intestine - Terminal Ileum	P02647	Proteins with structure data	P02647	Conf-score 9: Direct single protein target assigned		
P03951	FA11_HUMAN	plasma membrane	ENSG00000088926	Lung	P03951	Proteins with structure data	P03951	Conf-score 9: Direct single protein target assigned		
P04839	CY24B_HUMAN	plasma membrane	ENSG00000165168	Lung	P04839	Proteins with structure data	P04839	Conf-score 9: Direct single protein target assigned	0010600	disease by subcellular system affected
P05121	PAT1_HUMAN	plasma membrane	ENSG00000106356	Lung	P05121	Proteins with structure data	P05121	Conf-score 9: Direct single protein target assigned	0013328	disease by anatomical system
P05167	PPB1_HUMAN	plasma membrane	ENSG00000163293	Lung	P05167	Proteins with structure data	P05167	Conf-score 9: Direct single protein target assigned		
P05362	ICAM1_HUMAN	plasma membrane	ENSG00000090393	Lung	P05362	Proteins with structure data	P05362	Conf-score 9: Direct single protein target assigned		
P06239	LCK_HUMAN	plasma membrane	ENSG00000182866	Small Intestine - Terminal Ileum	P06239	Proteins with structure data	P06239	Conf-score 9: Direct single protein target assigned	0014333	disease by anatomical system

122のタンパク質(UniProt)の一覧が得られた。これらは、肺または小腸で発現が確認され、細胞膜表面に局在し、タンパク質立体構造が明らかになっており、対応する医薬品が開発されているヒトのタンパク質である。一覧には、ACE2(COVID-19の発病原因となるSARS-CoV-2ウイルスのウイルス受容体)も含まれている。

TogoDX探索結果(Map your IDs)

手持ちのIDリストに対する各属性での偏りやp値の計算、Add filtersによる絞り込みや、Map attributesによる他の属性での分布を確認することができる。各属性の内訳ごとに該当するIDの数とp値が計算されます。

該当するID有り
(カラー)

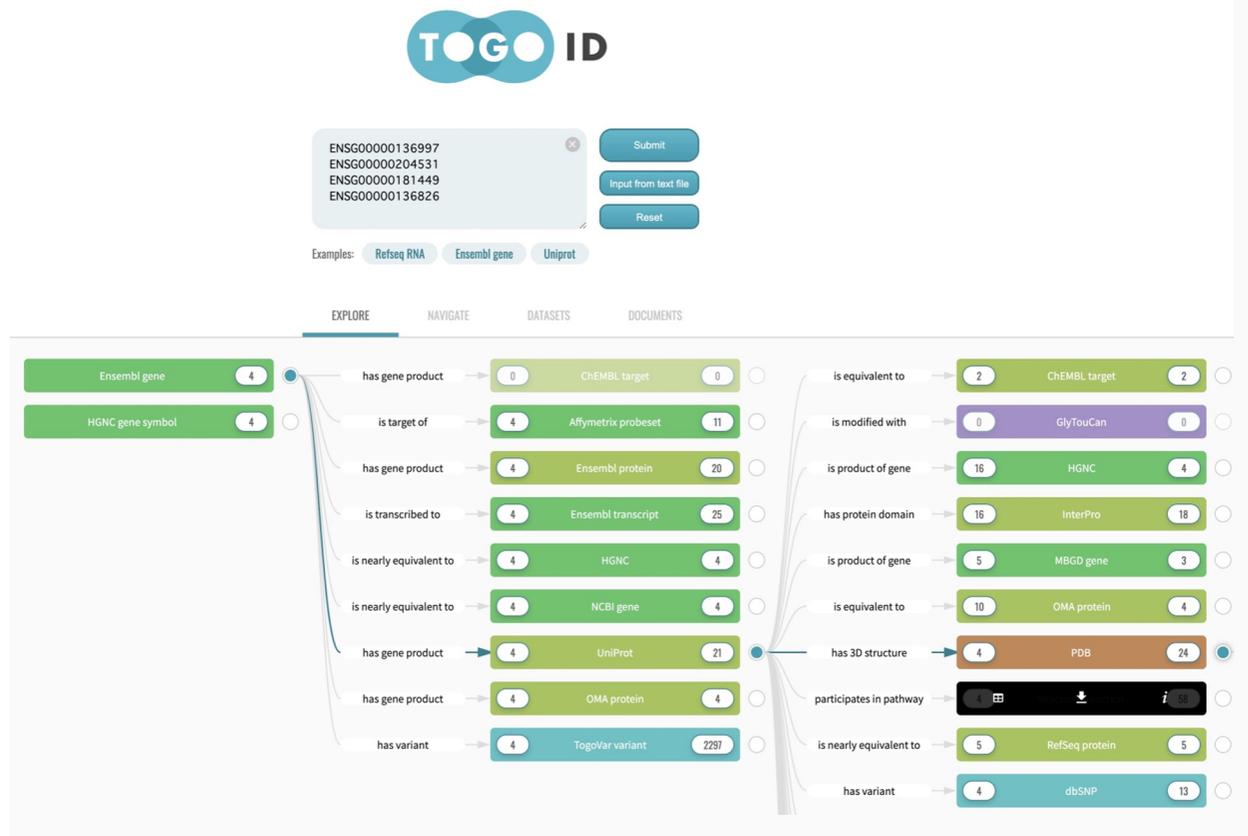
該当するID無し
(グレー)

The screenshot shows the TogoDX interface with several filters applied. A table titled 'Tissue-specific high expression (RefEx GeneChip)' is visible, showing the following data:

Values	Total	Mapped	p-value
<input type="checkbox"/> Testis	1,941	2	5.00e-1
<input type="checkbox"/> Thymus	1,217	5	7.91e-2
<input type="checkbox"/> Pancreas	1,027	1	1.00e+0
<input type="checkbox"/> Bone marrow	820		
<input type="checkbox"/> Muscle	717		
<input type="checkbox"/> Liver/Hepato	698		
<input type="checkbox"/> Placenta	694	2	1.00e+0
<input type="checkbox"/> Cerebellum	687	4	6.52e-2
<input type="checkbox"/> Cerebrum	606	1	1.00e+0
<input type="checkbox"/> Esophagus	544		
<input type="checkbox"/> Brain stem	542	1	1.00e+0
<input type="checkbox"/> Spleen	493		

A mouseover tooltip for the 'Cerebellum' entry shows: Count: 4 / 687, P-value: 6.515e-2. An orange box with an arrow points to this tooltip, containing the text: '該当数とp値を表あるいはマウスオーバーで確認'.

TogoID: データベース統合の基盤となる ID 変換サービス



The screenshot shows the TogoID interface. At the top, there's a search bar with the text "TOGO ID". Below it, a text input field contains four ENSEMBL gene IDs: ENSG00000136997, ENSG00000204531, ENSG00000181449, and ENSG00000136826. There are buttons for "Submit", "Input from text file", and "Reset". Below the input field, there are "Examples" for "Refseq RNA", "Ensembl gene", and "Uniprot".

The main part of the interface is a network diagram with nodes and edges. The nodes represent different biological categories and their associated IDs. The nodes are:

- Ensembl gene (4)
- HGNC gene symbol (4)
- ChEMBL target (0)
- Alfymetrix probeset (11)
- Ensembl protein (20)
- Ensembl transcript (25)
- HGNC (4)
- NCBI gene (4)
- UniProt (21)
- OMA protein (4)
- TogoVar variant (2297)
- ChEMBL target (2)
- GlyYouCan (0)
- HGNC (4)
- InterPro (18)
- MBGD gene (3)
- OMA protein (4)
- PDB (24)
- RefSeq protein (5)
- dbSNP (13)

The edges represent relationships between these categories, such as "has gene product", "is target of", "is transcribed to", "is nearly equivalent to", "is equivalent to", "is modified with", "is product of gene", "has protein domain", "has 3D structure", "participates in pathway", and "has variant".

入力した ID に対して

他のデータベースの同じ内容を示すデータの ID に変換

例 : NCBI Gene ID \leftrightarrow Ensembl Gene ID

カテゴリー間を跨いだ変換も可能

例 : 遺伝子が関連する疾患

幅広い生物学的カテゴリーの ID に対応

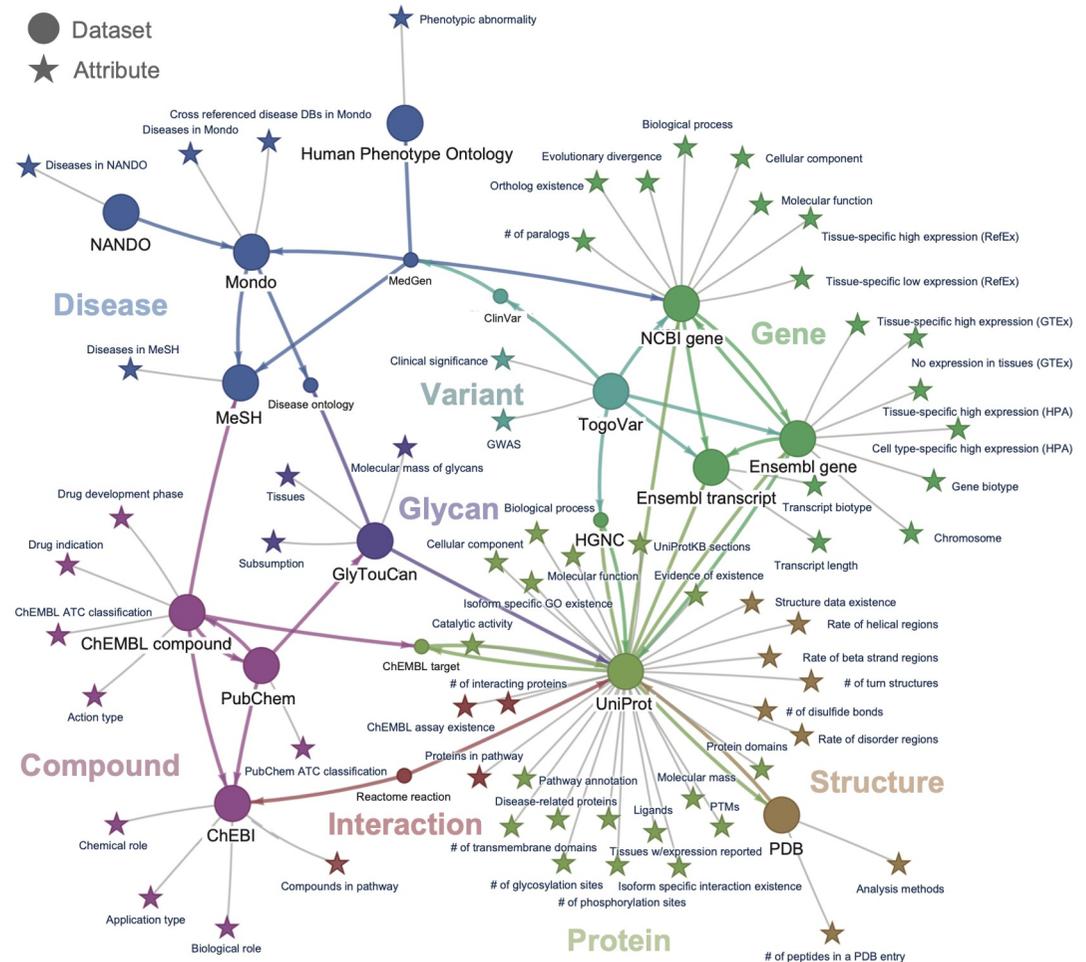
遺伝子、タンパク質、疾患、パスウェイ、化合物

TogoID: データベース統合の基盤となる ID 変換サービス

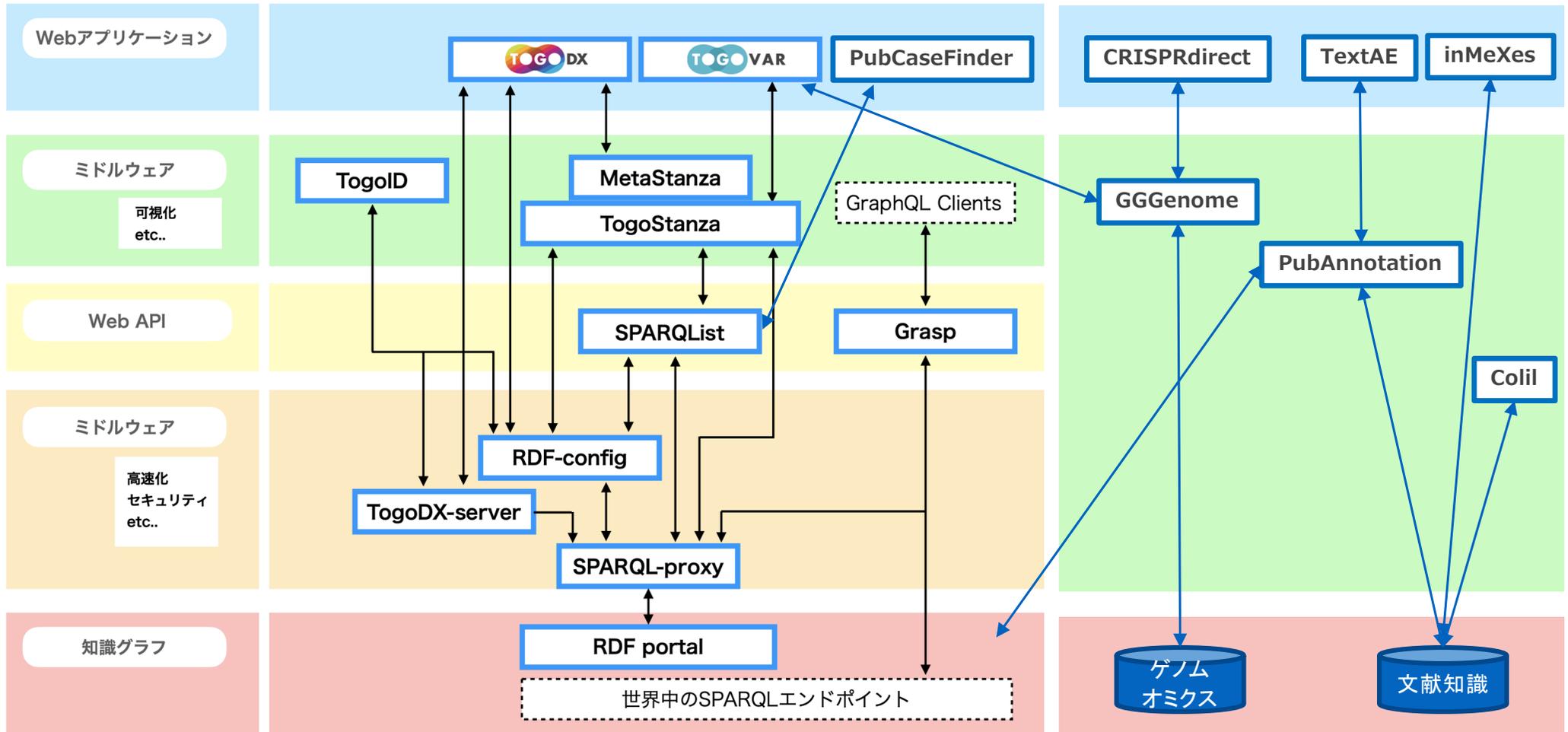
2022年度の成果

- 新規 ID ペアの追加
 - データベース数 : 48 → 64
 - データセット数 : 65 → 83
 - 190 データセットペア
- RDF Portal からデータセットを提供

TogoDX/Human でのデータセット間の関連付けでも活用



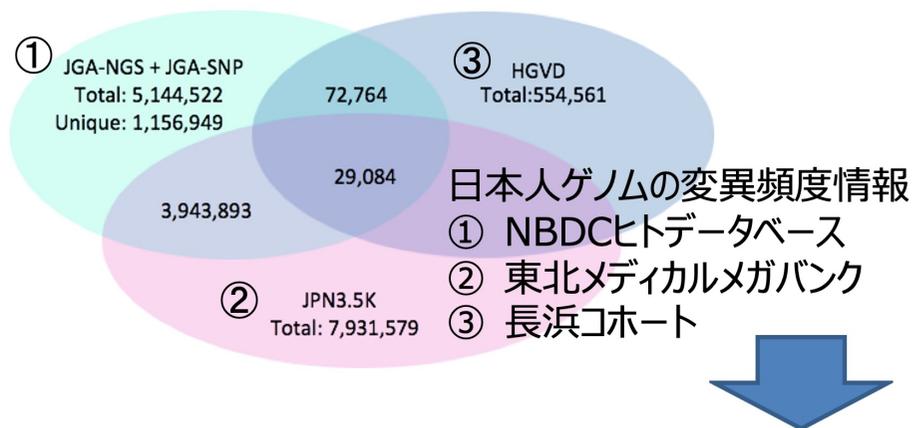
統合化実現のために開発した基盤技術



青枠はNBDC/DBCLSが開発したものの

TogoVar : 日本人ゲノム多様性統合DB

日本や海外で公開されている変異の頻度情報、変異と疾患、文献との関連情報を統合、ワンストップで検索可能に。AMEDゲノム医療プロジェクト GEM Japanの柱。



関連RDF データと統合

- ExAC: 他集団の変異頻度データ
- ClinVar: 変異-疾患関連情報
- Med2RDF: 多様な医科学関連データ
- PubMed, MeSH, PubTator, Colil: 関連文献データ
- TogoGenome: ゲノム情報

- 関連DBとの統合コストが1/4~1/6に削減：開発期間が 1年→数ヶ月
- 医科学研究者のみならず、医療関係従事者にもなくてはならないリファレンス
 - データベース検索の負担軽減、収集したデータの解釈コストの低減
- 他のデータとつなぐことによる応用への期待
 - (希少) 疾患データベースの症状と結びつけた診断支援への応用
 - 医薬品データと結びつけた創薬への応用
 - Genome Wide Association Study への応用

TogoVar検索例

疾患との関連性

TogoVar A comprehensive Japanese genetic variation database

rs114202595

Disease: Breast-ovarian cancer, familial 2 Gene: ALDH2 refSNP: rs114202595 TogoVar: tgv421843 Position(GRCh37/hg19): 16:48258198 Region(GRCh37/hg19): 10:732776976

Results The number of available variations is 2 out of 2.

TogoVar ID	RefSNP ID	Position	Ref / Alt	Type	Gene	Alt frequency	Consequence	SIFT	PolyPhen	Clinical significance
tgv30913364	rs114202595	7: 127254587	G > A	SNV	PAX4		Missense variant	0 +4	1 +4	Diabetes mellitus type 2
tgv30913365	rs114202595	7: 127254587	G > T	SNV	PAX4		Synonymous variant			

日本のデータ

- All 2
- WGS GEM-J WGA 1
- WES JGA NGS 1
- SNP JGA SNP 1
- WGS ToMMo 4.7KJPN 1
- WES HGVD 1
- WES ExAC 2
- Disease ClinVar 1

Alternative allele frequency

0 1 Invert range

Variant calling quality

Exclude filtered out variants in all datasets

Variant type

- All 2
- SNV 2
- Insertion 0
- Deletion 0
- Indel 0
- Substitution 0

Clinical significance

- All 2
- Not in ClinVar 1

個々のバリエーション情報へのリンク

DB毎頻度 (NBDCヒトDB由来データ含む)

Alt frequency

BioHackathon : 国際標準化、国際連携

- 国際版を2008年から国内版を2010年から、毎年1回開催（NBDCと共同）。昨年までの3年間はCOVIDのため中止となったが2019年の国際版は121名が参加（うち海外11カ国から37名）。



- 最先端のデータ統合技術やその応用についての世界の技術動向を調査・把握しテーマを決定。先端技術を用いてシステムやプログラム開発を行っている国内外の現場の研究者が参集する**国際標準化の場**。
- 合宿形式で分野横断的に問題解決を行い、**密度の高い情報交換と生産性の高い技術開発**が会議の場でリアルタイムに進行。



- DBCLS はこの分野で世界をリードするハブとして高く評価され、継続的にリーダーシップを発揮。
- 研究者コミュニティの共有資産となり論文発表してきた。昨年から成果を発表するbiohackrxivプレプリントサーバ (<https://biohackrxiv.org/>) を開始。
- 今年度は **RDF サミット**を開催し、RDF（知識グラフ）に基づくデータ統合の今後について欧米の研究者と議論

DBCLSメンバー

センター長
小原 雄治 / KOHARA, Yuji, PhD
 ゲノム科学
[✉ Mail](#) [👤 ORCID](#)

教授 (副センター長)
五斗 進 / GOTO, Susumu, PhD
 データベース統合
[✉ Mail](#) [👤 ORCID](#) [🎓 Google Scholar](#)

教授 (兼務)
大久保 公策 / OKUBO, Kousaku, MD, PhD
 ヒト統合、ポータル、発現統合、基盤技術
 担当サービス: BodyParts3D/Anatomography
[✉ Mail](#)

特任准教授
片山 俊明 / KATAYAMA, Toshiaki, PhD
 セマンティック・ウェブ技術によるゲノム情報
 担当サービス: TogoWS, TogoDB, TogoGenor
[✉ Mail](#) [🔗 GitHub](#) [👤 ORCID](#) [🔍 researchmap](#)

特任准教授
川島 秀一 / KAWASHIMA, Shuichi, PhD
 セマンティック・ウェブとバイオインフォマテ
 担当サービス: TogoGenome, TogoStanza, SF
[✉ Mail](#) [🔗 GitHub](#) [👤 ORCID](#) [🔍 researchmap](#)

特任准教授
金 進東 / KIM, Jin-Dong, PhD
 自然言語処理技術を利用したデータベースの生
 担当サービス: PubAnnotation, PubDictionary
[✉ Mail](#) [🔗 GitHub](#) [👤 ORCID](#) [🔍 researchmap](#)

特任准教授
島中 秀樹 / HATANAKA, Hideki, PhD
 生命科学系データベースの統合的利活用
[✉ Mail](#) [👤 ORCID](#)

特任准教授・DSコーディネーター
箕輪 真理 / MINOWA, Mari, PhD
 広報&もろもろの渉外窓口
[✉ Mail](#)

特任准教授
山本 泰智 / YAMAMOTO, Yasunori, Ph
 テキスト処理とセマンティックウェブ技術
 担当サービス: inMeXes, Allie, Colii, LSD f
[✉ Mail](#) [🔗 GitHub](#) [👤 ORCID](#)

特任助教
大田 達郎 / OHTA, Tazro, PhD
 大規模データ解析に資するデータベース技
 担当サービス: DBCLS SRA
[✉ Mail](#) [🔗 GitHub](#) [👤 ORCID](#)

特任助教
小野 浩雅 / ONO, Hiromasa, PhD
 「統合TV」編集・管理、遺伝子発現等の大
 担当サービス: TogoID, TogoTV, RefEx
[✉ Mail](#) [🔗 GitHub](#) [👤 ORCID](#) [🔍 researchmap](#)

特任助教
千葉 啓和 / CHIBA, Hirokazu, PhD
 データベース統合化のための技術開発、お
 担当サービス: SPANG
[✉ Mail](#) [👤 ORCID](#)

特任助教
内藤 雄樹 / NAITO, Yuki, PhD
 塩基配列などの大規模データを活用する技
 担当サービス: GGenome, CRISPRdirec
[✉ Mail](#) [🔗 GitHub](#) [👤 ORCID](#) [🔍 researchmap](#)

特任助教
藤原 豊史 / FUJIWARA, Toyofumi, PhD
 希少・遺伝性疾患に関わるソフトウェア開発およびデータ整備
 担当サービス: PubCaseFinder, Allie, Colii, LSD RDF Data Portal
[✉ Mail](#) [🔗 GitHub](#) [🔍 researchmap](#) [🎓 Google Scholar](#)

特任助教
守屋 勇樹 / MORIYA, Yuki, PhD
 基盤技術
 担当サービス: Endpoint browser, SPARQL support
[✉ Mail](#) [👤 ORCID](#) [🎓 Google Scholar](#)

特任研究員
申 在紋 / SHIN, Jae-Moon, PhD
[✉ Mail](#)

特任研究員
三橋 信孝 / MITSUHASHI, Nobutaka
 バリエーションを中心としたヒトデータの統合、サーバ管理
 担当サービス: TogoVar
[✉ Mail](#) [🔗 GitHub](#) [👤 ORCID](#)

学術支援技術専門員
池田 秀也 / IKEDA, Shuya
 担当サービス: TogoID
[✉ Mail](#)

学術支援技術専門員
高月 照江 / TAKATSUKI, Terue
[✉ Mail](#)

学術支援専門員
佐久間 桂子 / SAKUMA, Keiko
 広報普及、運営管理支援など
[✉ Mail](#)

准教授(兼務)
川本 祥子 / KAWAMOTO, Shoko, PhD
 データベースと文献の検索、辞書の構築と公開、オープン:
[✉ Mail](#)

特任准教授
岡本 忍 / OKAMOTO, Shinobu, PhD
 セマンティック・ウェブ技術による生物メタデータの標準化とアプリ
 担当サービス: TogoGenome, TogoStanza
[✉ Mail](#)

客員部門・研究協力者

客員教授
河野 信 / KAWANO, Shin, PhD
 広報・普及
 担当サービス: TogoTable
[✉ Mail](#)

客員教授
栗栖 源嗣 / KURISU, Genji, PhD
 タンパク質立体構造データベース
[✉ Mail](#)

客員教授
坊農 秀雅 / BONO, Hidemasa, PhD
 遺伝子発現データのアノテーションとそれらを使ったデータベース生物学
 担当サービス: TogoTV, GGenome, CRISPRdirect, GGRNA, RefEx, AOE
[✉ Mail](#) [🔗 GitHub](#) [👤 ORCID](#) [🔍 researchmap](#) [🎓 Google Scholar](#) [🌐 Website](#)

客員教授
山口 敦子 / YAMAGUCHI, Atsuko, PhD
 セマンティックウェブ技術を利用したデータベースの統合へ向けた基盤技術開発
 担当サービス: UmakaYummy, UmakaViewer, SPARQLBuilder
[✉ Mail](#) [🔗 GitHub](#) [👤 ORCID](#) [🔍 researchmap](#) [🎓 Google Scholar](#)



柏ラボ：東京大学柏の葉キャンパス駅前サテライト6F + 三島ラボ（国立遺伝学研究所内）