



ライフサイエンスデータベース統合推進事業
統合化推進プログラム
2026年度公募 募集説明会

2025年12月26日（金）
国立研究開発法人 科学技術振興機構（JST）
情報基盤事業部 NBDC事業推進室

注意事項

本公募は2026年度予算によるもので、国会における2026年度政府予算の成立を前提として行います。

予算編成の状況によっては、公募の内容を予告なく変更する場合があります。予めご了承ください。



統合化推進プログラムについて

統合化推進プログラムとは

- 統合化推進プログラムは、ライフサイエンスに関する国内外のデータを統合的に扱うための統合データベースの開発を目的としたプログラム。
- 国際的なデータ共有の推進とオープンサイエンスへの貢献を目指し、国際的に運用されている一次レポジトリや、新たな知識創出を目的とした二次データベースなど、幅広いライフサイエンス分野のさまざまなデータベースの研究開発を支援してきた。

「統合化推進プログラム」の現採択課題

区分	研究開発課題名	研究代表者	所属機関	開発対象データベース	採択年度
本格型	① ヒトゲノム・病原体ゲノムと疾患・医薬品をつなぐ統合データベース	金久 實	京都大学	KEGG MEDICUS	2022年度
	② 異分野融合を志向した糖鎖科学ポータルのデータ拡充と品質向上	木下 聖子	創価大学	GlyCosmos Portal	
	③ 蛋白質構造データバンクのデータ駆動型研究基盤への拡張	栗栖 源嗣	大阪大学	PDBj	
	④ バイオイメージングデータのグローバルなデータ共有システムの構築	大浪 修一	理化学研究所	SSBDデータベース	
	⑤ 統合的な転写制御データ基盤の構築	柏川 雄也	理化学研究所	INTRARED	
	⑥ マイクロバイオーム研究を先導するハブを目指した微生物統合データベースの特化型開発	森 宙史	情報・システム研究機構	Microbiome Datahub	
	⑦ 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	石濱 泰	京都大学	Japan Proteome Standard Repository/Database (jPOST)	
	⑧ 次世代低分子マススペクトルデータベース シン・マスバンクの構築	松田 史生	大阪大学	Shin-MassBank (開発中)	
育成型	① 非モデル植物のための遺伝子ネットワーク情報活用基盤	大林 武	東北大学	ATTED-II	2023年度
	② 日本人塩基配列情報の公開可能なゲノム・オミクス情報基盤による双方向型研究教育データベース開発と国際連携	長崎 正朗	九州大学	Joint Open Genome and Omics Platform (JoGo)	
	③ 空間オミックスデータ解析用データベースの開発	VANDENBON Alexis	京都大学	DeepSpaceDB	
	④ AI駆動型データキュレーションによる持続可能な中分子相互作用統合データベースの開発	池田 和由	理化学研究所	MIIDB-AI (開発中)	2024年度
	⑤ 細胞レベルの機能・表現型と遺伝子発現を関連付ける「Cell IO」データベースの開発	尾崎 遼	理化学研究所	Cell IO (開発中)	
	⑥ 創発的再解析のためのメタボローム統合データベース	早川 英介	九州工業大学	integMET (開発中)	
	⑦ 大規模言語モデルを活用した病的スプライシング変異データベースの自律的構築	白石 友一	国立がん研究センター	SSCV DB	2025年度
	⑧ 4Dゲノム状態の理解と可視化を支援するデータベースの構築	新海 創也	理化学研究所	PHi-C database (開発中)	
	⑨ 化合物と個体をつなぐ毒性病理画像データベースCuraToxiiの開発	水野 忠快	東京大学	CuraToxii (開発中)	

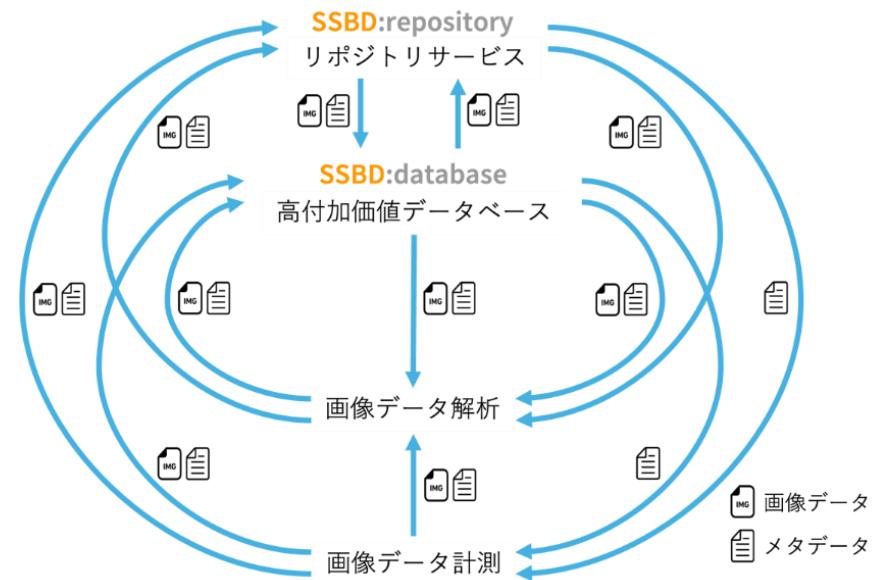
※本格型は1年間に直接経費35,000千円/課題（+間接経費30%）×5年間、育成型は1年間に10,000千円/課題（+間接経費30%）×3年間を支援

統合化推進プログラムのデータベース構築例（本格型）

○ 生物画像データベース SSBD [理研 大浪 修一] **SSBD:database** (<https://ssbd.riken.jp/database/>)

バイオイメージングデータと生命動態定量データの統合データベース。日本の全ての生命科学研究者が利活用し、かつ日米欧などの国際連携に基づく国際的データ共有基盤として再構築する。情報科学的な画像解析手法の開発・共有促進プラットフォームの構築、持続的な運用と発展を可能にするためのシステム開発を実施する。

- **SSBD:repository** : 生物画像データの公共リポジトリ。
- **SSBD:database** : メタデータリッチな高付加価値データベース。



○ 統合的転写制御解析基盤 Intrared (理研 粕川 雄也) **Intrared**

(<https://www.intrared.org/>)

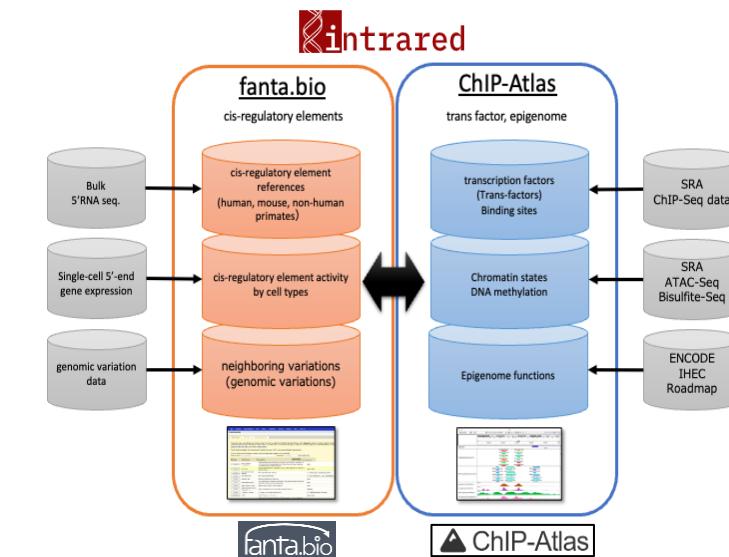
シスエレメントデータベースfanta.bioとエピゲノミクスデータベースChIP-Atlasを統合し、統合的転写制御解析基盤を構築する。



- **fanta.bio** : 遺伝子発現調節に関わるゲノム上のシス転写調節配列（シスエレメント）と細胞の種類・状態ごとの活性状態を収録するデータベース。ヒト、マウスのシスエレメント情報と疾患関連多型や変異情報等を収載。



- **ChIP-Atlas** : ChIP-seq、ATAC-seq、DNase-seq、Bisulfite-seqのデータをそれぞれ統一プロトコルで再解析して抗原別・生物種別・細胞種別に整理し、Hi-C, ChromHMM, GWAS SNPs, eQTLsデータなどの様々なエピゲノミクスを統合したデータベース。



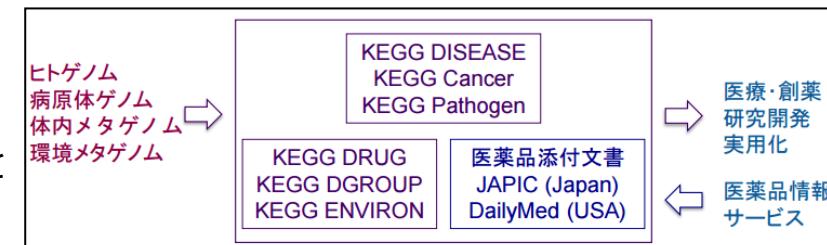
統合化推進プログラムのデータベース構築例（本格型）

○ ゲノム・疾患・医薬品の統合データベース KEGG MEDICUS [京都大学 金久 實]

(<https://www.kegg.jp/kegg/medicus/>)

個々の遺伝子や複数の遺伝子から構成された機能モジュール、遺伝子、タンパク質、環境因子、医薬品等から構成された相互作用ユニットに関する知識をデータベース化し、ゲノム情報を有効利用するためのリソース。

- ・遺伝子、タンパク質、また代謝やシグナル伝達などの分子間ネットワークを統合したデータベースKEGGで、ヒトゲノム、病原体ゲノムなどのシークエンス解読と有効利用を促進する統合リソースKEGG MEDICUSを開発し提供。
- ・KEGG MEDICUSの月間平均利用者数は約280万人、利用件数は600万件を超える。Googleで医薬品名を検索すると上位にヒットし、一般市民からも広く利用されている。



KEGG MEDICUSの構成。日米の医薬品添付文書、疾患情報、医薬品情報を整備し、ゲノム、メタゲノム情報と統合して提供している。

○ 糖鎖情報統合データベース GlyCosmos Glycoscience Portal [創価大学 木下 聖子]

(<https://glycosmos.org/>)

国際コンソーシアムGlySpace Allianceと連携し、糖鎖構造のIDや構造表記、フォーマット等の共通化を図る。

- ・リポジトリに登録された様々な糖鎖関連データをキュレーションして公開する仕組みを構築し、分野を超えた糖鎖情報の活用を促進して、糖鎖科学と構造生物学、微生物、植物、プロテオミクスなどとの融合を推進する。
- ・これまで構築してきた糖鎖構造、複合糖鎖、糖鎖分析データなどリポジトリに加え、レクチンマイクロアレイデータや微生物関連糖鎖情報および糖鎖関連パスウェイデータなどにもリポジトリを拡張する。



統合化推進プログラムのデータベース構築例（本格型）

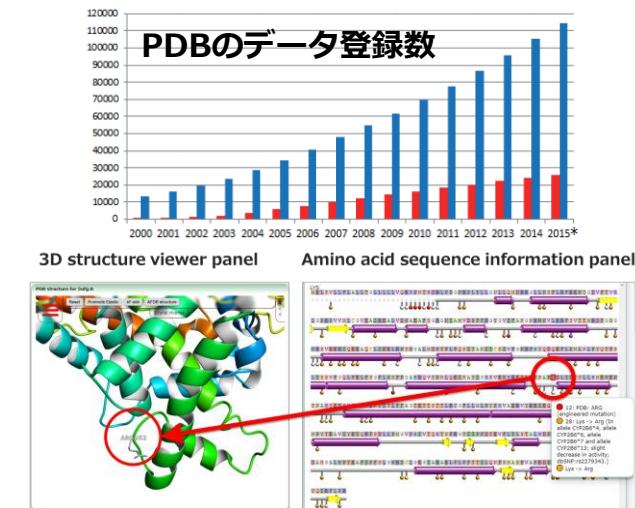
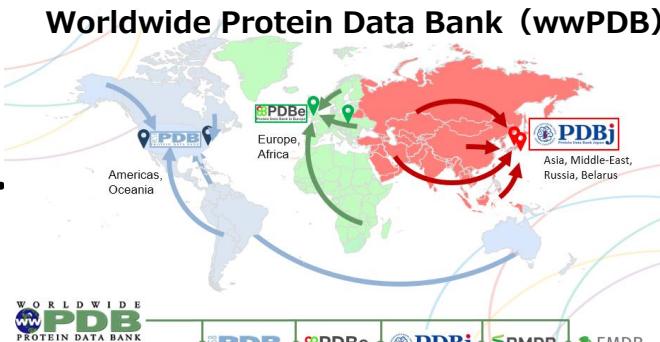
○ タンパク質構造データバンク PDBj [大阪大学 栗栖 源嗣]



(<https://pdbj.org/>)

PDB（タンパク質構造データバンク）とBMRB（NMR実験情報データバンク）を構築・公開し、日米欧3極で構成するwwPDBの主要メンバーとしてアジア地区からのタンパク質立体構造の登録を処理。独自のアノテーションやツールを提供し、構造生物学データのさらなる高度化を図る。

- ・世界的な生命科学研究基盤データベースとして認知されており、2022年、Global Biodata Coalition (GBC)のGlobal Core Biodata Resourcesに認定。
- ・世界全体で登録された約24万件のデータのうち約3割をPDBjが登録。
- ・全ての実験手法でデータ品質の検証レポートを確立しRDF化して利用者に提供。
- ・近年、急増している予測構造を含む構造にも対応。
- ・日本人ゲノムバリアント情報をPDB構造上にマップし、生物学・化学の視点からの検索結果を表示する「PDB統合利用ポータル」を2024年10月に公開。
- ・構造ベース創薬（SBDD）など応用分野でも広く利用されている。



○ マイクロバーオーム統合データベース Microbiome Datahub (国立遺伝研 森 宙史)

(<https://mdatahub.org/>)

これまで構築してきた微生物ゲノム統合データベース MicrobeDB.jpを発展させ、国際的に隆盛をきわめる微生物叢（マイクロバイオーム）研究に特化した国際的なデータハブ Microbiome Datahubとして再構築する。微生物の表現型やメタゲノム由来のゲノム（Metagenome Assembled Genome; MAG）などの新しいデータを統合化し、マイクロバイオーム関連のデータ検索・解析機能を大幅に強化して、自由度の高い検索・データ取得を実現する。

統合化推進プログラムのデータベース構築例（本格型）

○ プロテオーム統合データベース jPOST [京都大学 石濱 泰]



(<https://jpostdb.org/>)

国内外に散在するプロテオーム情報を標準化・統合・一元管理し、多彩な生物種

・翻訳後修飾・絶対発現量も含めた統合プロテオームデータベースを開発。

- ・プロテオーム国際コンソーシアム（ProteomeXchange）のアジア初メンバー。
ProteomeXchangeではデータIDを共通化、各サイトに登録されているデータ検索を可能にした。

- ・独自に開発した品質評価指標UniScoreを用いた再解析によりデータセットを跨ぐ統合解析を可能にした。

- ・2022年、jPOSTを含むProteomeXchangeがGlobal Biodata Coalition (GBC)のGlobal Core Biodata Resourcesに認定、国際的に重要なバイオデータベースの一つとして認知された。
- ・アジアや欧米から多くのデータが登録され、国内外の研究者により広くデータ利用がされている。

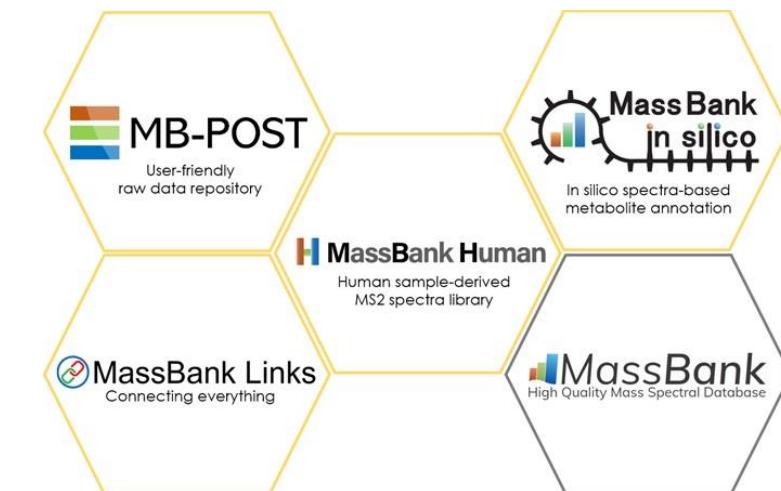


○ メタボローム解析パイプライン Shin-MassBank [大阪大学 松田 史生]

(<https://shin.massbank.jp/>) (開発中)

生体由来スペクトルを集積していくための解析パイプライン Shin-MassBankを構築し、高品質な実測マススペクトルをMassBankに保存していく。ユーザー主導の開発体制で、生命科学研究者などのユーザーニーズを満たす次世代データベースを構築する。

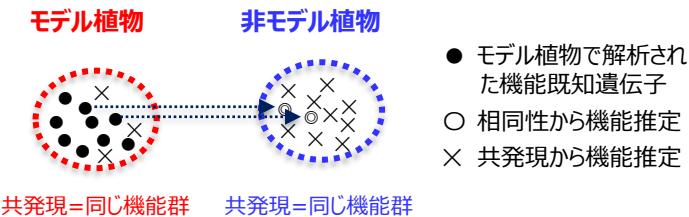
- ・MB-POST : 質量分析生データリポジトリ
- ・MassBank human : ヒトサンプル由来のスペクトラムDB
- ・MassBank *in silico* : 予測マススペクトルを収載
- ・MassBank Links



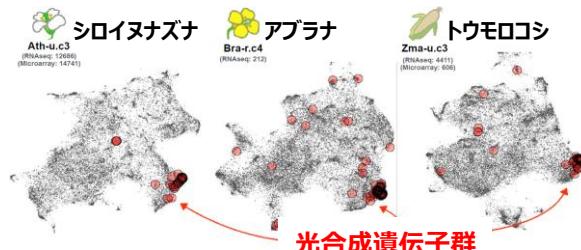
統合化推進プログラム 2023年度 採択課題（育成型）

植物遺伝子の共発現DB ATTED-II 東北大学 大林武

遺伝子の共発現情報から、機能未知遺伝子の機能を推定するDB。モデル植物で明らかにされた知見を実用植物を含む非モデル植物の解析に役立つ重要な情報基盤となる。



共発現データをオオムギ・コムギ・コメなどを含む19植物種20系統に拡張とともに、共発現情報を比較する機能を高度化して、種間比較解析ツールを新たに開発。また、2つの遺伝子が共発現する環境や条件を調べたり、共発現遺伝子が担っている生理的機能や制御関係を調べる新たな機能を実装。



共発現マップを使った異種間遺伝子の比較解析

ヒト2万遺伝子の470万ハプロタイプDB JoGo Platform 九州大学 長崎正朗

長鎖リードシーケンスで決定された5大陸258例（うち国内105例）のゲノム配列を収載し、これを基に同定した約2万遺伝子の約470万ものハプロタイプをカタログ化。さらに、同一検体で取られた遺伝



子発現情報を収載し、遺伝子のハプロタイプと発現の関係を解析できるようにした。従来の短鎖リードシーケンスでは同定が困難だった構造多型を含む遺伝子の型の違いを精緻に捉え、多因子疾患や希少疾患などの関連リスクを、ゲノムワイドでハプロタイプの違いとして評価できるようにした世界初のデータベース。2025年11月に公開。

Joint Open Genome and
Omics Platform (JoGo)
<https://jogo.csml.org>

対象ヒト遺伝子数 19,194

ACTG-ハプロタイプ 総数 4,656,478

ACTG-ハプロタイプ記法

Gene | Transcript Separator

H B B : a 1 c 1 t 1 g 1

Amino acid sequence level
(Missense, Stop, Frameshift)

Coding sequence (CDS) level
(Synonymous)

Transcript level
(5'UTR or 3'UTR in exon region)

Genebody level
(Intron)

オンラインハプロタイプ エクスプローラー[®]
ハプロタイプ エクスプローラー[®]
ハプロタイプ QTLP[®] ユーアー[®]
ClinVar / GWAS Catalog / GTExの変異
のハプロタイプ 情報へのアノテーションツール
ヒト近縁種とのハプロタイプ 比較ヒューリズム

HBB_chr11_5220464_5232071

...
a 0001c0001t0001g0001: G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0003: G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0007: G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0010: G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0011: G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0012: G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0014: G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0016: G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0021: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0022: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0023: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0026: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0027: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0028: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0029: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0030: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0031: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0032: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0033: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0034: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0035: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0036: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0037: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0038: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0039: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0040: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0041: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0042: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0043: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0044: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0045: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0046: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0047: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0048: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0049: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0050: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0051: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0052: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0053: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0054: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0055: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0056: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0057: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0058: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0059: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0060: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0061: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0062: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0063: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0064: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0065: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0066: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0067: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0068: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0069: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0070: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0071: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0072: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0073: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0074: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0075: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0076: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0077: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0078: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0079: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0080: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0081: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0082: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0083: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0084: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0085: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0086: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0087: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0088: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0089: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0090: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0091: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0092: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0093: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0094: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0095: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0096: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0097: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0098: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0099: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0100: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0101: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0102: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0103: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0104: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0105: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0106: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0107: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0108: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0109: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0110: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0111: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0112: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0113: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0114: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0115: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0116: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0117: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0118: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0119: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0120: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0121: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0122: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0123: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0124: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0125: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0126: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0127: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0128: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0129: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0130: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0131: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0132: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0133: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0134: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0135: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0136: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0137: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0138: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0139: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0140: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0141: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0142: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0143: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0144: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0145: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0146: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0147: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0148: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0149: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0150: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0151: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0152: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0153: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0154: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0155: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0156: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0157: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0158: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0159: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0160: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0161: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0162: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0163: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0164: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0165: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0166: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0167: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0168: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0169: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0170: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0171: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0172: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0173: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0174: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0175: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0176: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0177: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0178: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0179: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0180: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0181: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0182: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0183: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0184: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0185: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0186: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0187: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0188: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0189: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0190: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0191: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0192: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0193: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0194: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0195: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0196: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0197: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0198: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0199: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0200: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0201: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0202: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0203: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0204: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0205: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0206: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0207: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0208: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0209: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0210: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0211: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0212: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0213: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0214: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0215: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0216: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0217: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0218: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0219: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0220: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0221: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0222: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0223: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0224: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0225: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0226: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0227: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0228: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0229: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0230: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0231: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0232: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0233: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0234: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0235: T G A C G T C G T T A C A C G A G G
a 0001c0001t0001g0236: T G A C G T

統合化推進プログラム 2024年度 採択課題（育成型）

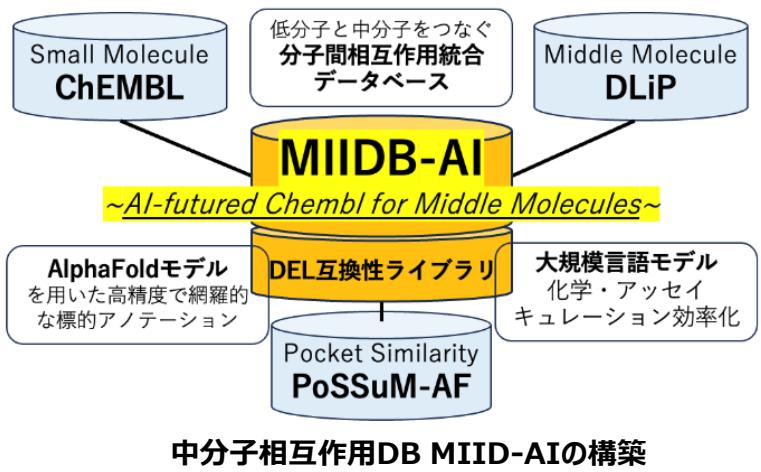
中分子相互作用の統合DB

MIIDB-AI (開発中)

理化学研究所 池田 和由



構造ベースの創薬が困難とされる標的分子情報を収載することを目的として、ペプチド、非ペプチド、核酸などの中分子と標的タンパク質との相互作用情報を収載したDB MIIDB-AIを構築する。本DBでは、リガンド結合部位の類似性データに基づき、中分子の標的相互作用部位を特定し、独自に開発したAI技術を使って高精度に標的とリガンド間の相互作用を予測する。中分子の知識基盤DBによって、新規の中分子医薬品候補を効率良く見つけ出せるようになり、次世代創薬研究への貢献が期待される。



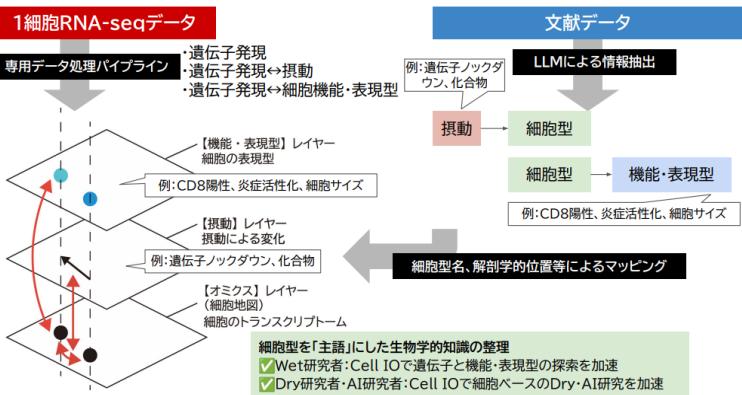
細胞の機能・表現型と 遺伝子発現を関連付ける

DB Cell IO (開発中)

理化学研究所 尾崎 遼



文献情報と遺伝子発現情報から、細胞の機能・表現型を整理したDB Cell IOを構築する。ヒトやマウスなどの1細胞トランск립トームデータと紐づく細胞の機能・表現型と摂動情報を整理する。文献からは細胞の由来、機能・表現型、摂動の種類、生物種等の情報を抽出し、検索やデータ統合が容易にできるようにする。遺伝子発現データからは、細胞の機能・表現型を解釈し、1細胞解析から明らかにされる新たな細胞タイプを整理する。細胞の機能・表現型を予測する研究や、所望の機能・表現型の細胞をデザインする研究などへの貢献が期待される。



Cell IO (Input/Output)の構築

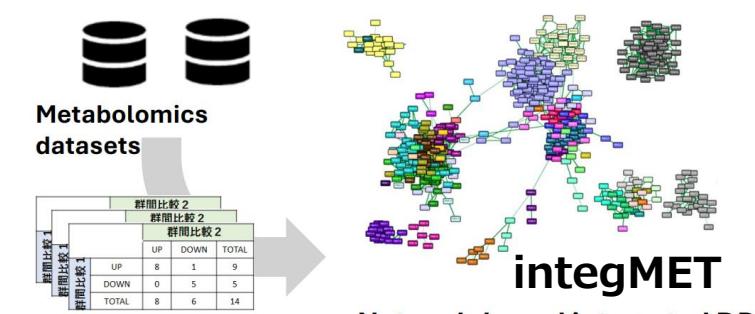
創発的再解析のための メタボローム統合DB

integMET (開発中)

九州工業大学 早川 英介



質量分析で取られたメタボロームデータの再解析や異なる実験間での比較解析は困難なことが知られている。代謝物の変動傾向とメタデータの類似度からメタボロームデータを知識ネットワークとして整備し、再解析や実験間での比較を可能とするintegMETを構築する。メタボロミクスデータをネットワークとして体系化することで、大規模なデータ再解析が可能となり、代謝と生命現象との関係を俯瞰的に理解できるようになる。さらに研究情報を統合したメタボロームの知識グラフを使い、データサイエンスやAI研究に応用するなど、分野横断的な新しい研究・知識の創出につながるDBの構築を目指す。



タボロミクスデータの再解析基盤となるintegMETの構築

統合化推進プログラム 2025年度 採択課題（育成型）

病的スプライシング変異

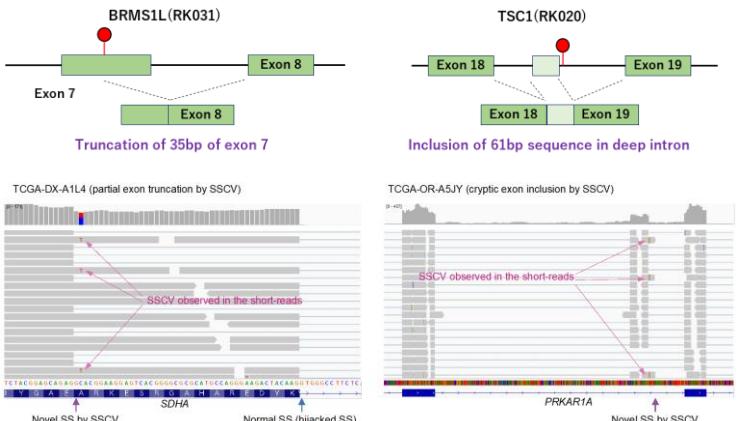
DBの自律的構築

SSCV DB

国立がん研究センター 白石 友一



公共RNA-seqデータからスプライスサイト生成変異 (SSCV) を同定し収載したSSCV DBを発展させ、大規模言語モデルを用いて病的 possibilityを予測し、疾患との関連性情報を提供する。また、トランスクリプトームデータから自動的にSSCVを検出し、DBを自動更新するパイプラインを開発する。遺伝性疾患やがんの原因ともなるSSCVは、従来のゲノム解析では難しく見落とされがちだったが、本DBを使って疾患とSSCVとの関連性を明らかにすることにより、SSCVの検出による疾患診断法の開発や、SSCVを標的とする核酸医薬の研究開発などへの寄与を目指す。



4Dゲノム状態を可視化

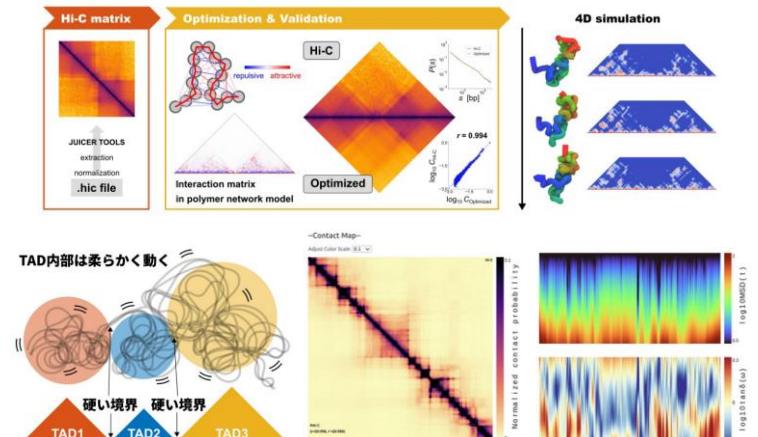
するDB

PHi-C database (開発中)

理化学研究所 新海 創也



独自に開発した解析手法 (PHi-C) を用い、公共の3次元ゲノム構造 (Hi-C) データに対し、動的な3次元ゲノム構造 (4Dゲノム状態) を網羅的に解析・可視化するPHi-C databaseを構築する。GPUサーバーを用いたPHi-C計算の高速化と大規模化を進め、複数のHi-Cデータを直感的に比較・理解できる4Dゲノム情報の統合的可視化基盤を実現する。本DBを通じて、真核生物における生命現象の解明に不可欠な細胞核内におけるゲノムDNAの動的な高次構造情報を提供し、シミュレーションや理論解析、ゲノム統合解析を通じた新たな仮説創出や知識発見への貢献を目指す。



毒性病理画像DB

CuraToxii (開発中)

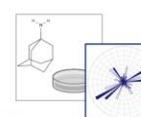
東京大学 水野 忠快



毒性病理の専門家と連携して収集・体系化した毒性病理画像DB CuraToxiiを構築する。毒性病理画像およびメタ情報の提供の他、深層学習モデルの開発や毒性評価への応用を支援するためのベンチマークデータセットの構築、リーダーボード機能、動的データ取得機能等を実装する。本DBによって毒性病理画像のファイル容量に起因する操作性の低さやデータ入手性といった既存の課題を解決し、化合物が個体に与える影響の理解に重要な情報を提供する。医薬品安全性評価の効率化・高度化とともに、化合物と個体への影響の理解深化に基づく創薬研究の加速化への貢献を目指す。

オミクスデータのパターン①

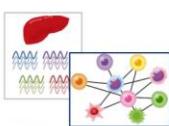
入力：培養細胞のマイクロアレイ
手法：因子分析
出力：化合物の複合的効用



Mizuno T., Sci Rep., 2019
Nemoto S., J Nat Prod., 2021

オミクスデータのパターン②

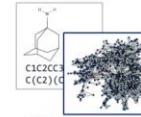
入力：組織RNAseq
手法：潜在的ディクレ配分
出力：組織中の細胞比率



Morita K., Toxicol Sci., 2023
Azuma I., Brief Bioinform., 2024

化合物構造のパターン

入力：化合物構造
手法：ニューラル機械翻訳
出力：化合物構造の数値表現



Nemoto S., J Cheminform., 2022
Yoshikai Y., Nat Commun., 2024

毒性病理画像のパターン

入力：毒性病理画像
手法：画像自己教師あり学習
出力：個体応答の潜在表現



Maedera S., Comput Biol Med., 2024
Iwasaka T., biorxiv., 2024

統合化推進プログラム
研究総括方針

九州大学 生体防御医学研究所 特任教授
伊藤 隆司

統合化推進プログラムの公募開始にあたって

伊藤 隆司 (九州大学 生体防御医学研究所)

統合化推進プログラム（DICP）とは

統合化推進プログラムは、ライフサイエンスに関する国内外のデータを統合的に扱うためのデータベース（統合データベース）開発を目的としたプログラムです。

具体的な研究開発は、研究データの収集・整理、標準化・公開・共有・品質管理、運用等で、他のデータベースとの連携・統合化や、収録データを円滑に利用するために必要なインターフェース・ツールの開発等も含みます。

本プログラムは、これらの研究開発を通じて **公共データ利活用のための情報環境整備**を行うとともに、利用者の知識発見や課題解決への寄与および国際的なオープンサイエンスへの貢献を目指します。

☞ **健全なバイオデータサイエンスの発展こそが究極の目的**

統合化推進プログラム (DICP) の歴史

- 2011年に第1期が開始
 - ・過去3期で計31課題を支援 (<https://biosciencedbc.jp/funding/project/>)
- 2022年に第4期が開始
 - ・「つなぐ」から「使う」へ
 - ・広範なユーザーの知識発見支援
 - ・国際的なプレゼンス
 - ・新しい動向への対応

【研究アドバイザー】 (50音順)

鎌田 真由美	北里大学 未来工学部 教授
坂井 寛章	農業・食品産業技術総合研究機構 高度分析研究センター ユニット長
清水 佳奈	早稲田大学 理工学術院 教授
白井 剛	長浜バイオ大学 バイオサイエンス学部 教授
瀬々 潤	(株)ヒューマノーム研究所 代表取締役社長
夏目 やよい	医薬基盤・健康・栄養研究所 AI 健康・医薬研究センター センター長
馬場 健史	九州大学 生体防御医学研究所 主幹教授
本郷 裕一	東京科学大学 生命理工学院 教授
山本 一夫	お茶の水女子大学 ヒューマンライフサイエンス研究所 客員教授
吉田 哲郎	アクセリード(株) 経営企画部 シニアディレクター

※ 新任

第4期の課題に目指して頂きたいこと

- 当該研究コミュニティに支持され愛用されるDB
地に足の付いたDB
- 広範な非専門家ユーザーのクロスボーダーを支援
玄人相手はむしろ簡単、より難しいのが素人相手の「敷居の低さ」
- 國際的なプレゼンス
協調 and/or 突出
- 課題間の相乗効果
データの統合と技術の共有 — 実質的な連携
- 新しい動向を先取りした対応

第4期でDICPとして目指したいこと

ライフステージに応じたDB支援の複線化

- 萌芽期 試験的支援による発掘・育成 **育成型**
- 発展期 本格的支援による成長・確立 **本格型**
- 円熟期 別枠の支援による維持・継続 ?
- 引退後 データの保全 **DBアーカイブ**

統合化推進プログラム (DICP) の歴史

- 2011年に第1期が開始
 - ・過去3期で計31課題を支援 (<https://biosciencedbc.jp/funding/project/>)
- 2022年に第4期が開始
 - ・「つなぐ」から「使う」へ
 - ・広範なユーザーの知識発見支援
 - ・国際的なプレゼンス
 - ・新しい動向への対応
 - ・2023年より育成型を開始
(DICPの新しい挑戦)

【研究アドバイザー】 (50音順)

鎌田 真由美	北里大学 未来工学部 教授
坂井 寛章	農業・食品産業技術総合研究機構 高度分析研究センター ユニット長
清水 佳奈	早稲田大学 理工学術院 教授
白井 剛	長浜バイオ大学 バイオサイエンス学部 教授
瀬々 潤	(株)ヒューマノーム研究所 代表取締役社長
夏目 やよい	医薬基盤・健康・栄養研究所 AI 健康・医薬研究センター センター長
馬場 健史	九州大学 生体防御医学研究所 主幹教授
本郷 裕一	東京科学大学 生命理工学院 教授
山本 一夫	お茶の水女子大学 ヒューマンライフサイエンス研究所 客員教授
吉田 哲郎	アクセリード(株) 経営企画部 シニアディレクター

※ 新任

今回の募集：育成型の4回目の公募

新たなデータベースの構築を目指す萌芽的な研究開発

技術動向や研究ニーズへの対応を目指し、**未整備分野**の萌芽的な統合データベースの構築や、**斬新な切り口**によるデータ統合など、独創的な発想に基づいた知識発見・課題解決の支援に資する先駆的なデータベースを研究開発する提案を募集します。

新たにデータベースを構築しようとしている、既にデータベースを構築しているが未公開の状態にある、データベースを公開していく将来的に国際的な地位確立を目指している場合などを対象とします。

いざれは国際基準のデータ基盤へ発展していくことを期待しています。

育成型に求めること

課題全般に求めること

- ・「つなぐ」から「使う」へ
- ・広範なユーザーの知識発見支援
- ・国際的なプレゼンス
- ・新しい動向への対応

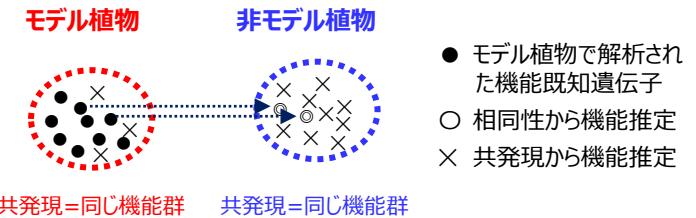
育成型に求めること

- ・構想に新規性・独自性・挑戦性・将来性が感じられること
- ・実績は問わないが、実現可能性が感じられること
DB構築の戦略に不安を感じさせない陣容
- ・既存DBの新展開の提案に、独自性・挑戦性・将来性が感じられること
従来の延長線上に留まらない飛躍・新機軸

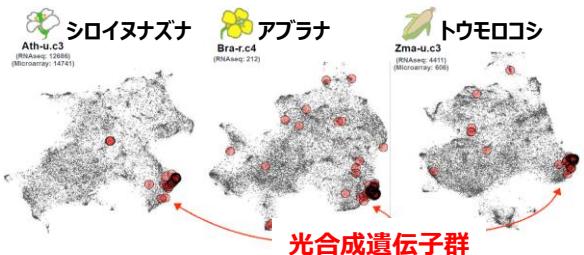
統合化推進プログラム 2023年度 採択課題（育成型）

植物遺伝子の共発現DB ATTED-II 東北大学 大林武

遺伝子の共発現情報から、機能未知遺伝子の機能を推定するDB。モデル植物で明らかにされた知見を実用植物を含む非モデル植物の解析に役立つ重要な情報基盤となる。



共発現データをオオムギ・コムギ・コメなどを含む19植物種20系統に拡張するとともに、共発現情報を比較する機能を高度化して、種間比較解析ツールを新たに開発。また、2つの遺伝子が共発現する環境や条件を調べたり、共発現遺伝子が担っている生理的機能や制御関係を調べる新たな機能を実装。



共発現マップを使った異種間遺伝子の比較解析

ヒト2万遺伝子の470万ハプロタイプDB JoGo Platform 九州大学 長崎正朗

長鎖リードシーケンスで決定された5大陸258例（うち国内105例）のゲノム配列を収載し、これを基に同定した約2万遺伝子の約470万ものハプロタイプをカタログ化。さらに、同一検体で取られた遺伝子発現情報を収載し、遺伝子のハプロタイプと発現の関係を解析できるようにした。従来の短鎖リードシーケンスでは同定が困難だった構造多型を含む遺伝子の型の違いを精緻に捉え、多因子疾患や希少疾患などとの関連リスクを、ゲノムワイドでハプロタイプの違いとして評価できるようにした世界初のデータベース。2025年11月に公開。

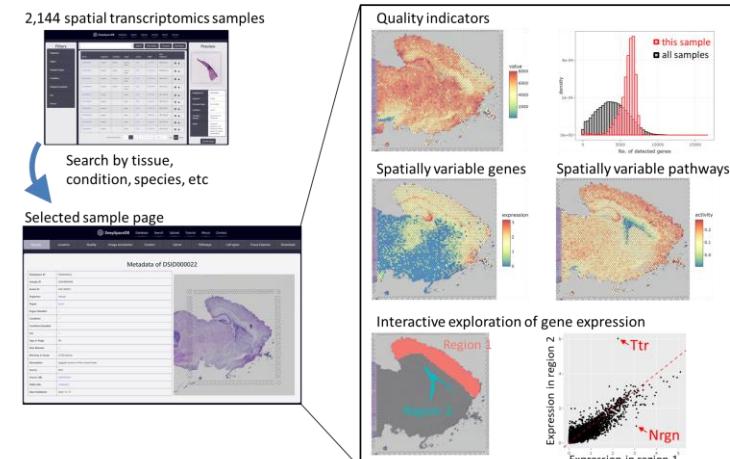
Joint Open Genome and Omics Platform (JoGo)
<https://jogo.csml.org>
対象ヒト遺伝子数 19,194
ACTG-ハプロタイプ 総数 4,656,478
ACTG-ハプロタイプ記法
HBB : a1 c1 t1 g1
Amino acid sequence level (Missense, Stop, Frameshift)
Coding sequence (CDS) level (Synonymous)
Transcript level (5'UTR or 5'UTR in exon region)
Genebody level (Intron)
HBB_ch11_5220464_5232071
オンラインハプロタイプ イクスプローラー
ロカルハプロタイプ イクスプローラー
ハプロタイプ QTBLi ユア
ClinVar / GWAS Catalog / GTExの変異のハプロタイプ 情報へのアノテーションツール
ヒト近縁種とのハプロタイプ比較ヒュア

ACTGハプロタイプ表記法とHBB遺伝子の例

空間トランскルiptームDB DeepSpaceDB 京都大学 A. Vandenberg

組織の微細領域ごとに取られた遺伝子発現データを、誰もが簡単に高度な解析ができるようにしたDB。2024年9月公開。主要プラットフォームの10x VisiumとXeniumで取られたヒトとマウスの

ほぼ全ての公開データを収載し、任意の遺伝子の空間的発現分布、空間的に発現が変動している代謝やシグナル経路、発現遺伝子から予測された細胞の種類の空間的分布などの表示、さらに、異なるサンプル間での比較を含む任意に選択した領域間での遺伝子発現比較ができる。また、自分のデータをアップロードしてDB中のデータとの比較解析が可能。



DeepSpaceDBインターフェイスでの表示例
(特定遺伝子の発現状態や選択領域で特徴的な遺伝子を可視化)

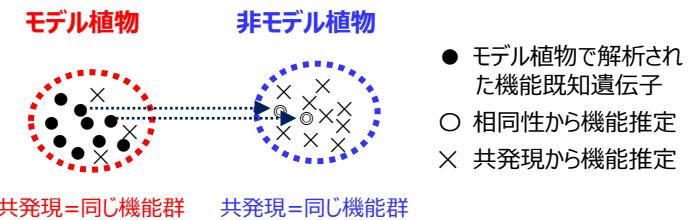


統合化推進プログラム 2023年度 採択課題（育成型）

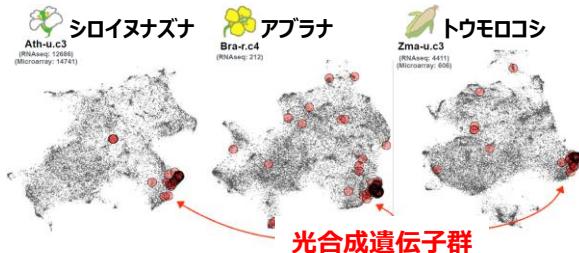
植物遺伝子の共発現DB ATTED-II

東北大学 大林武

遺伝子の共発現情報から、機能未知遺伝子の機能を推定するDB。モデル植物で明らかにされた知見を実用植物を含む非モデル植物の解析に役立つ重要な情報基盤となる。



共発現データをオオムギ・コムギ・コメなどを含む19植物種20系統に拡張するとともに、共発現情報を比較する機能を高度化して、種間比較解析ツールを新たに開発。また、2つの遺伝子が共発現する環境や条件を調べたり、共発現遺伝子が担っている生理的機能や制御関係を調べる新たな機能を実装。



共発現マップ[®]を使った異種間遺伝子の比較解析

トト工場に ヨコハマ

九間トランスクリプトードB

京都大学 A. Vandenberg



オミクス領域の新しい動向への対応が目立った

同定した約2万遺伝子の約470万ものハプロタイプをカタログ化。さらに、同一検体で取られた遺伝子発現情報を収載し、遺伝子のハプロタイプと発現の関係を解析できるようにした。従来の短鎖リードシーケンスでは同定が困難だった構造多型を含む遺伝子の型の違いを精緻に捉え、多因子疾患や希少疾患などとの関連リスクを、ゲノムワイドでハプロタイプの違いとして評価できるようにした世界初のデータベース。2025年11月に公開。




**Joint Open Genome and
Omics Platform (JoGo)**
<https://jogo.csml.org>

対象ヒト遺伝子数 19,194

ACTG-八ヶ岳・白山・縦走 総数 4,656,478

ACTG-ハプロタイプ記法

ne | Transcript Separator

amino acid sequence level

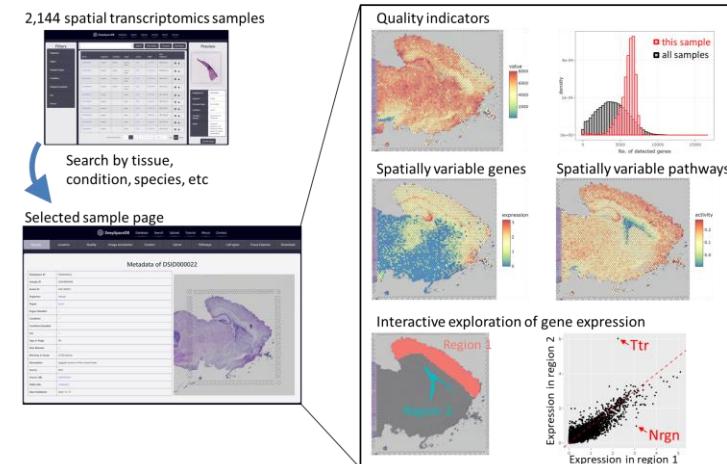
Translate

Transcript

anchovy level

ACTGハプロタイプ表記法と*HBB*遺伝子の例

たDB。2024年9月公開。主要プラットフォームの10x VisiumとXeniumで取られたヒトとマウスのほぼ全ての公開データを収載し、任意の遺伝子の空間的発現分布、空間的に発現が変動している代謝やシグナル経路、発現遺伝子から予測された細胞の種類の空間的分布などの表示、さらに、異なるサンプル間での比較を含む任意に選択した領域間での遺伝子発現比較ができる。また、自分のデータをアップロードしてDB中のデータとの比較解析が可能。



DeepSpaceDBインターフェイスでの表示例

(特定遺伝子の発現状態や選択領域で特徴的な遺伝子を可視化)

統合化推進プログラム 2023年度 採択課題（育成型）

Nucleic Acids Research, 2025, gka1232
<https://doi.org/10.1093/nar/gka1232>

NAR Breakthrough Article



OXFORD

JoGo 1.0: the ACTG hierarchical nomenclature and database covering 4.7 million haplotypes across 19,194 human genes

Masao Nagasaki^{1,2,*}, Toshiaki Katayama³, Yuki Moriya³, Yayoi Sekiya¹, Shuichi Kawashima³, Ryo Teraoka¹, Shuto Machida¹, Taichi Matsubara¹, Hiroki Hashimoto¹, Akihiro Asakura¹, Akio Nagano⁴, Riu Yamashita^{5,6}, Toyoyuki Takada⁷, Nobutaka Mitsuhashi³, Mayumi Kamada⁸, Yasuyuki Ohkawa¹, Katsushi Tokunaga⁹, Yosuke Kawai^{9,10}, Variant Information Standardization Collegium¹

¹Medical Research Center for High Depth Omics, Medical Institute of Bioregulation, Kyushu University, Fukuoka 812-8582, Japan

²Department of Center for Genomic Medicine, Graduate School of Medicine, Kyoto University, Kyoto 606-8507, Japan

³Database Center for Life Science, Joint Support-Center for Data Science Research, Research Organization of Information and Systems, 178-4-4 Wakashiba, Kashiwa, Chiba 277-0871, Japan

⁴PENGE Inc., 2-20-11 Kojima, Taito-ku, Tokyo 111-0056, Japan

⁵Division of Translational Informatics, Exploratory Oncology Research & Clinical Trial Center, National Cancer Center, 6-5-1 Kashiwanoha, Kashiwa, Chiba 277-8577, Japan

⁶Department of Computational Biology and Medical Sciences, Graduate School of Frontier Sciences, The University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8561, Japan

⁷Integrated Bioresource Information Division, RIKEN BioResource Research Center, 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan

⁸Department of Data Science, School of Frontier Engineering, Kitasato University, 1-15-1, Kitazato, Minami, Sagamihara, Kanagawa 252-0373, Japan

⁹Genome Medical Science Project, National Institute of Global Health and Medicine, Japan Institute for Health and Security, Shinjuku-ku, Tokyo 162-8655, Japan

¹⁰Bioinformatics and DDBJ Center, Research Organization of Information and Systems, National Institute of Genetics, 1111 Yata, Mishima-shi, Shizuoka 411-8540, Japan

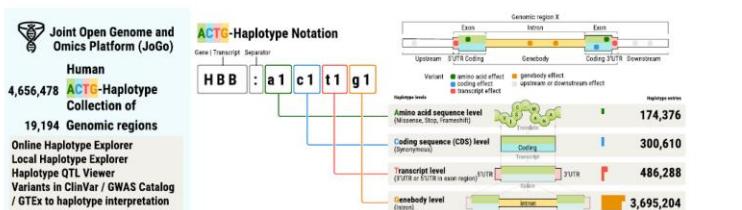
*To whom correspondence should be addressed. Emails: nagasaki@csml.org & nagasaki.masao.822@m.kyushu-u.ac.jp

[†]Full group members are listed in the appendix.

Abstract

The Joint Open Genome and Omics Platform 1.0 (JoGo) is a global, long-read-based human haplotype database covering 19,194 MANE-standardized protein-coding genes. JoGo introduces a novel ACTG hierarchical nomenclature—A (amino acid), C (coding), T (transcript), and G (gene body)—that assigns numeric identifiers in descending order of global frequency. Using high-fidelity long-read sequencing, we assembled haplotype-resolved contigs for 258 globally sampled genomes, including 108 sequenced in-house. We cataloged 174,376 A, 300,610 C, 486,288 T, and 3,695,204 G-level haplotypes (4,656,478 in total). Haplotype IDs are assigned once globally across all sequences, including those originating from GRCh38 and CHM13v2 reference assemblies, embedding reference haplotypes within the same frequency-ranked space and enabling direct cross-assembly comparison. JoGo maps functional variants from ClinVar, GWAS Catalog, and GTEx onto their corresponding ACTG-haplotypes and provides haplotype-expression QTL results from 1280 HapMap RNA-seq samples across three independent studies. The web portal provides flexible search by gene name, variant ID, or ACTG code. It offers both an interactive online viewer and a privacy-preserving local viewer for secure integration with user data. JoGo enables high-resolution exploration of haplotype diversity, facilitating the identification of functional variants relevant to gene regulation, disease associations, and precision medicine. JoGo 1.0 is freely accessible at <https://jogo.csml.org>.

Graphical abstract



Received: August 15, 2025. Revised: October 8, 2025. Accepted: October 18, 2025.

© The Author(s) 2025. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Nucleic Acids Research, 2025, gka1117
<https://doi.org/10.1093/nar/gka1117>

Database issue



OXFORD

DeepSpaceDB: a spatial transcriptomics atlas for interactive in-depth analysis of tissues and tissue microenvironments

Vladyslav Honcharuk^{1,2}, Afeefa Zainab¹, Yoshiya Horimoto^{○3,4}, Keiko Takemoto¹, Diego Diez^{○5}, Shinpei Kawaoka^{6,7}, Alexis Vandenberg^{○1,8,9,*}

¹Laboratory of Tissue Homeostasis, Institute for Life and Medical Sciences, Kyoto University, Kyoto 606-8507, Japan

²Department of Neuroscience, Graduate School of Medicine, Kyoto University, Kyoto 606-8507, Japan

³Department of Breast Surgery and Oncology, Tokyo Medical University Hospital, Tokyo 160-0023, Japan

⁴Department of Human Pathology, Juntendo University Faculty of Medicine, Tokyo 113-0033, Japan

⁵Immunology Frontier Research Center, Osaka University, Osaka 565-0871, Japan

⁶Department of Integrative Bioanalytics, Institute of Development, Aging and Cancer (IDAC), Tohoku University, Sendai 980-8575, Japan

⁷Institute for Life and Medical Sciences, Kyoto University, Kyoto 606-8507, Japan

⁸Institute for Liberal Arts and Sciences, Kyoto University, Kyoto 606-8507, Japan

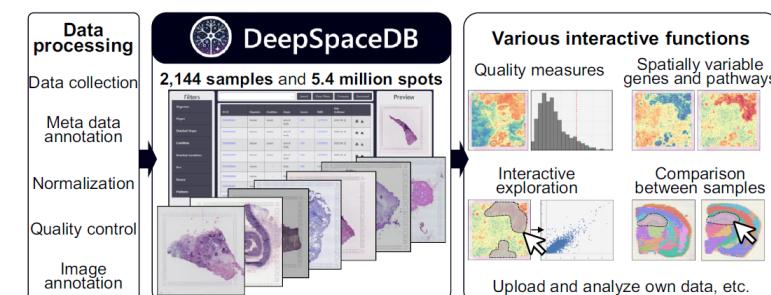
⁹Department of Mammalian Regulatory Network, Graduate School of Biostudies, Kyoto University, Kyoto 606-8507, Japan

*To whom correspondence should be addressed. Email: alexisvd@infront.kyoto-u.ac.jp

Abstract

Spatial transcriptomics enables detailed mapping of gene expression within tissues, revealing spatial organization of cellular and molecular processes. However, generating such data is costly and technically challenging, and analysis requires advanced bioinformatics skills. Although public datasets are growing, existing databases offer limited tools for interactive exploration and cross-sample comparison. Here, we introduce DeepSpaceDB (www.deepspacedb.com), a next-generation spatial transcriptomics database designed to address these challenges. The current version of DeepSpaceDB focuses on 10X Genomics Visium samples, ensuring higher-quality analyses and enhanced tools. This distinguishes it from databases that prioritize broad platform coverage over functionality. Emphasizing interactivity and advanced analytics, DeepSpaceDB enables flexible exploration of spatial transcriptomics data. Users can interactively compare gene expression across regions within or between tissue slices, such as between hippocampal areas of an Alzheimer's model mouse and a control. The database also offers quality indicators, database-wide trends, and interactive visualizations like zoomable plots and hover-based info. Moreover, these functions are not restricted to samples in our database but can also be applied to samples uploaded by users. Combining advanced tools with interactive features, DeepSpaceDB is a powerful resource for spatial transcriptomics, enabling deeper insights into tissue organization and disease biology.

Graphical abstract



Received: July 18, 2025. Revised: October 6, 2025. Accepted: October 7, 2025.

© The Author(s) 2025. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

統合化推進プログラム 2024年度 採択課題（育成型）

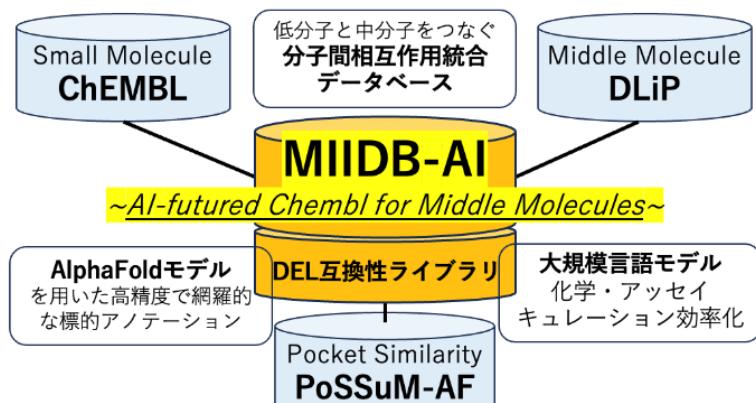
中分子相互作用の統合DB

MIIDB-AI (開発中)

理化学研究所 池田 和由



構造ベースの創薬が困難とされる標的分子情報を収載することを目的として、ペプチド、非ペプチド、核酸などの中分子と標的タンパク質との相互作用情報を収載したDB MIIDB-AIを構築する。本DBでは、リガンド結合部位の類似性データに基づき、中分子の標的相互作用部位を特定し、独自に開発したAI技術を使って高精度に標的とリガンド間の相互作用を予測する。中分子の知識基盤DBによって、新規の中分子医薬品候補を効率良く見つけ出せるようになり、次世代創薬研究への貢献が期待される。



中分子相互作用DB MIID-AIの構築

細胞の機能・表現型と

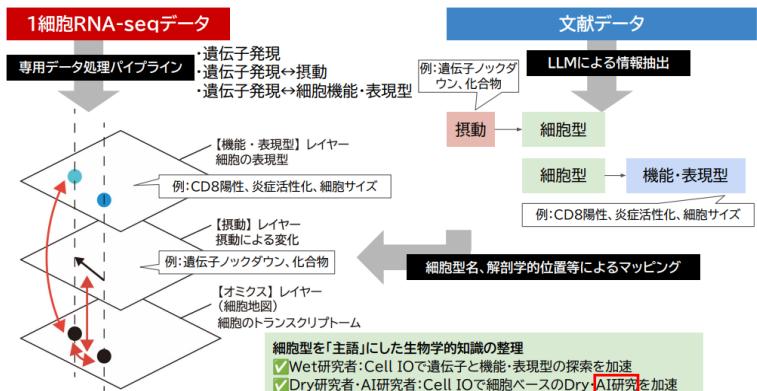
遺伝子発現を関連付ける

DB Cell IO (開発中)

理化学研究所 尾崎 遼



文献情報と遺伝子発現情報から、細胞の機能・表現型を整理したDB Cell IOを構築する。ヒトやマウスなどの1細胞トランскルプトームデータと紐づく細胞の機能・表現型と摂動情報を整理する。文献からは細胞の由来、機能・表現型、摂動の種類、生物種等の情報を抽出し、検索やデータ統合が容易にできるようになる。遺伝子発現データからは、細胞の機能・表現型を解釈し、1細胞解析から明らかにされる新たな細胞タイプを整理する。細胞の機能・表現型を予測する研究や、所望の機能・表現型の細胞をデザインする研究などへの貢献が期待される。



Cell IO (Input/Output)の構築

創発的再解析のための

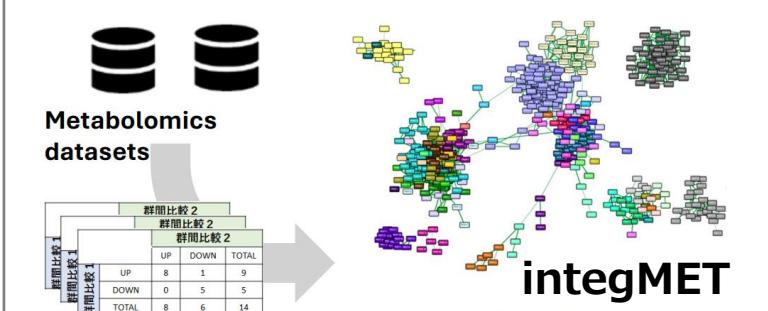
メタボローム統合DB

integMET (開発中)

九州工業大学 早川 英介



質量分析で取られたメタボロームデータの再解析や異なる実験間での比較解析は困難なことが知られている。代謝物の変動傾向とメタデータの類似度からメタボロームデータを知識ネットワークとして整備し、再解析や実験間での比較を可能とするintegMETを構築する。メタボロミクスデータをネットワークとして体系化することで、大規模なデータ再解析が可能となり、代謝と生命現象との関係を俯瞰的に理解できるようになる。さらに研究情報を統合したメタボロームの知識グラフを使い、データサイエンスやAI研究に応用するなど、分野横断的な新しい研究・知識の創出につながるDBの構築を目指す。



メタボロミクスデータの再解析基盤となるintegMETの構築

統合化推進プログラム 2024年度 採択課題（育成型）

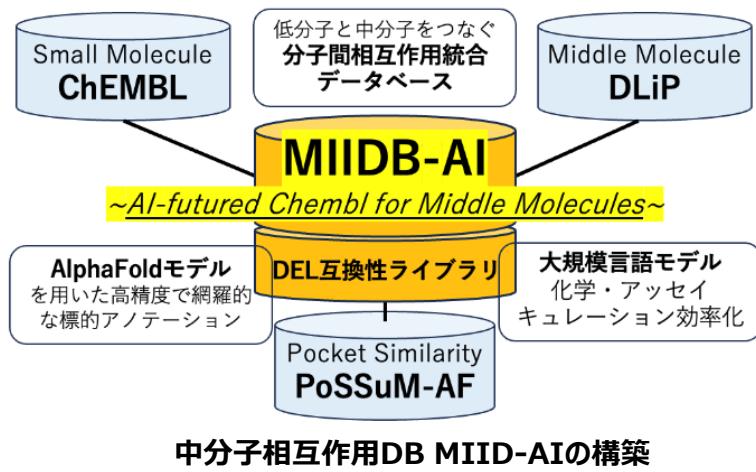
中分子相互作用の統合DB

MIIDB-AI (開発中)

理化学研究所 池田 和由



構造ベースの創薬が困難とされる標的分子情報を収載することを目的として、ペプチド、非ペプチド、核酸などの中分子と標的タンパク質との相互作用情報を収載したDB MIIDB-AIを構築する。本DBでは、リガンド結合部位の類似性データに基づき、中分子の標的相互作用部位を特定し、独自に開発したAI技術を使って高精度に標的とリガンド間の相互作用を予測する。中分子の知識基盤DBによって、新規の中分子医薬品候補を効率良く見つけ出せるようになり、次世代創薬研究への貢献が期待される。

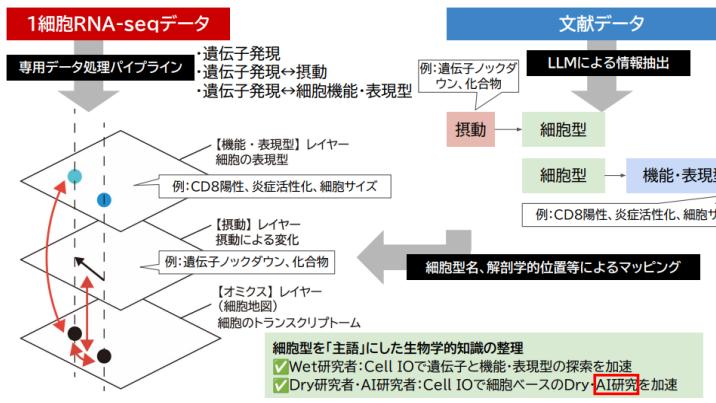


細遺理

LLMなどAIを利用する提案が急増した

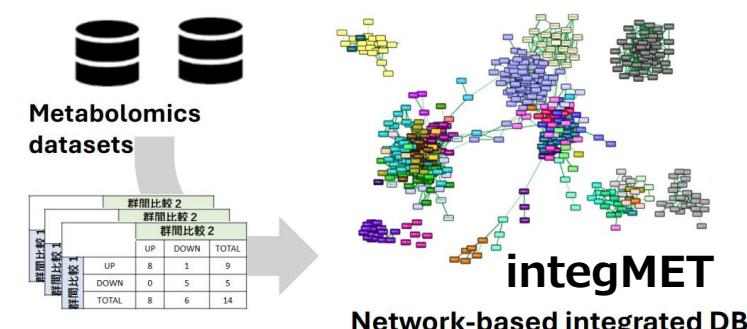


文献情報と遺伝子発現情報から、細胞の機能・表現型を整理したDB Cell IOを構築する。ヒトやマウスなどの1細胞トランск립トームデータと紐づく細胞の機能・表現型と摂動情報を整理する。文献からは細胞の由来、機能・表現型、摂動の種類、生物種等の情報を抽出し、検索やデータ統合が容易にできるようにする。遺伝子発現データからは、細胞の機能・表現型を解釈し、1細胞解析から明らかにされる新たな細胞タイプを整理する。細胞の機能・表現型を予測する研究や、所望の機能・表現型の細胞をデザインする研究などへの貢献が期待される。



Cell IO (Input/Output)の構築

質量分析で取られたメタボロームデータの再解析や異なる実験間での比較解析は困難なことが知られている。代謝物の変動傾向とメタデータの類似度からメタボロームデータを知識ネットワークとして整備し、再解析や実験間での比較を可能とするintegMETを構築する。メタボロミクスデータをネットワークとして体系化することで、大規模なデータ再解析が可能となり、代謝と生命現象との関係を俯瞰的に理解できるようになる。さらに研究情報を統合したメタボロームの知識グラフを使い、データサイエンスやAI研究に応用するなど、分野横断的な新しい研究・知識の創出につながるDBの構築を目指す。



メタボロミクスデータの再解析基盤となるintegMETの構築

統合化推進プログラム 2025年度 採択課題（育成型）

病的スプライシング変異

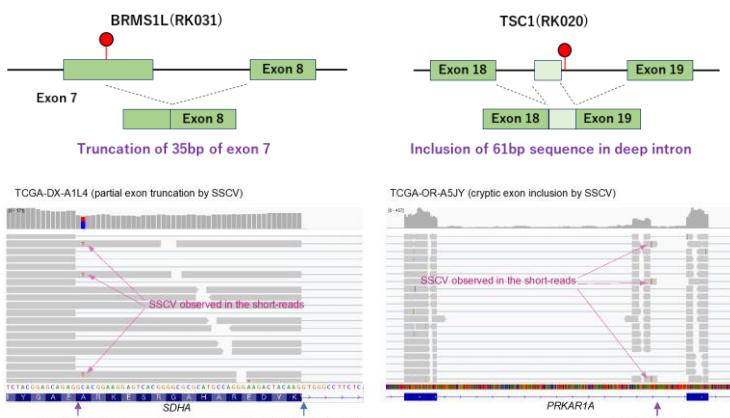
DBの自律的構築

SSCV DB

国立がん研究センター 白石 友一



公共RNA-seqデータからスプライスサイト生成変異 (SSCV) を同定し収載したSSCV DBを発展させ、大規模言語モデルを用いて病的 possibilityを予測し、疾患との関連性情報を提供する。また、トランスクリプトームデータから自動的にSSCVを検出し、DBを自動更新するパイプラインを開発する。遺伝性疾患やがんの原因ともなるSSCVは、従来のゲノム解析では難しく見落とされがちだったが、本DBを使って疾患とSSCVとの関連性を明らかにすることにより、SSCVの検出による疾患診断法の開発や、SSCVを標的とする核酸医薬の研究開発などへの寄与を目指す。



4Dゲノム状態を可視化

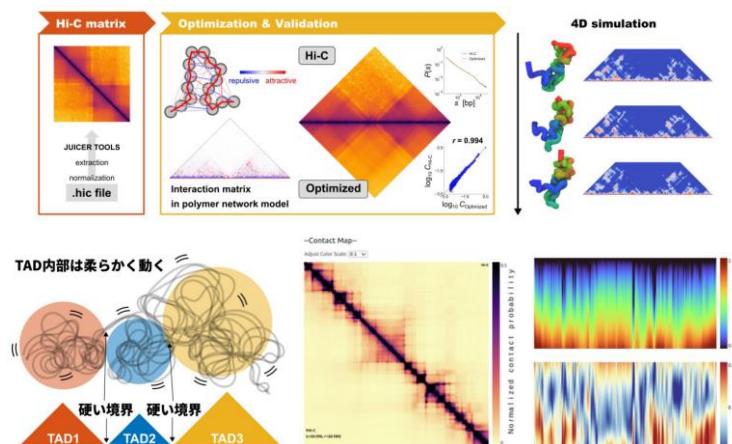
するDB

PHi-C database (開発中)

理化学研究所 新海 創也



独自に開発した解析手法 (PHi-C) を用い、公共の3次元ゲノム構造 (Hi-C) データに対し、動的な3次元ゲノム構造 (4Dゲノム状態) を網羅的に解析・可視化するPHi-C databaseを構築する。GPUサーバーを用いたPHi-C計算の高速化と大規模化を進め、複数のHi-Cデータを直感的に比較・理解できる4Dゲノム情報の統合的可視化基盤を実現する。本DBを通じて、真核生物における生命現象の解明に不可欠な細胞核内におけるゲノムDNAの動的な高次構造情報を提供し、シミュレーションや理論解析、ゲノム統合解析を通じた新たな仮説創出や知識発見への貢献を目指す。



毒性病理画像DB

CuraToxii (開発中)

東京大学 水野 忠快



毒性病理の専門家と連携して収集・体系化した毒性病理画像DB CuraToxiiを構築する。毒性病理画像およびメタ情報の提供の他、深層学習モデルの開発や毒性評価への応用を支援するためのベンチマークデータセットの構築、リーダーボード機能、動的データ取得機能等を実装する。本DBによって毒性病理画像のファイル容量に起因する操作性の低さやデータ入手性といった既存の課題を解決し、化合物が個体に与える影響の理解に重要な情報を提供する。医薬品安全性評価の効率化・高度化とともに、化合物と個体への影響の理解深化に基づく創薬研究の加速化への貢献を目指す。

オミクスデータのパターン①

入力：培養細胞のマイクロアレイ
手法：因子分析
出力：化合物の複合的作

Mizuno T., *Sci Rep.*, 2019
Nemoto S., *J Nat Prod.*, 2021



オミクスデータのパターン②

入力：組織RNaseq
手法：潜在的ディリケル配分
出力：組織中の細胞比率

Morita K., *Toxicol Sci.*, 2023
Azuma I., *Brief Bioinform.*, 2024

化合物構造のパターン

入力：化合物構造
手法：ニューラル機械翻訳
出力：化合物構造の数値表現

Nemoto S., *J Cheminform.*, 2022
Yoshikai Y., *Nat Commun.*, 2024



毒性病理画像のパターン

入力：毒性病理画像
手法：画像自己教師あり学習
出力：個体応答の潜在表現

Maedera S., *Comput Biol Med.*, 2024
Iwasaka T., *biorxiv*, 2024

統合化推進プログラム 2025年度 採択課題（育成型）

病的スプライシング変異

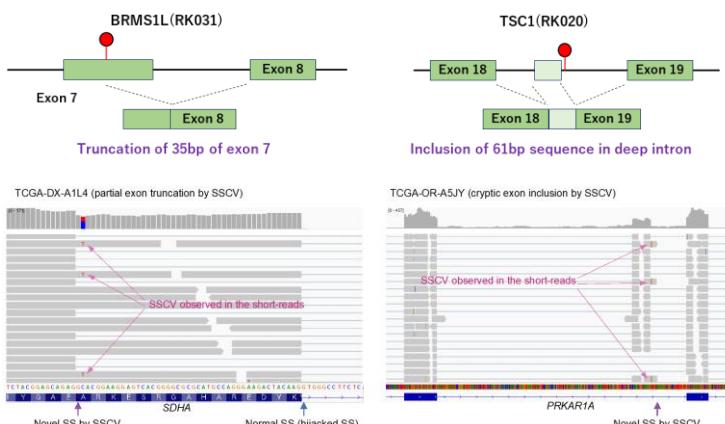
DBの自律的構築

SSCV DB

国立がん研究センター 白石 友一



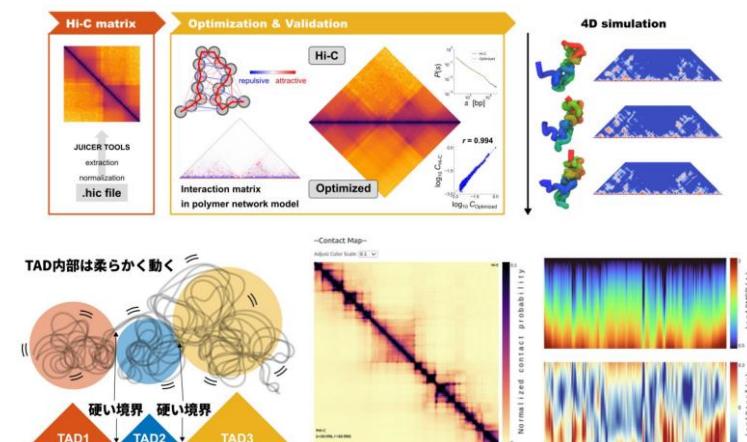
公共RNA-seqデータからスプライスサイト生成変異 (SSCV) を同定し収載したSSCV DBを発展させ、大規模言語モデルを用いて病的 possibilityを予測し、疾患との関連性情報を提供する。また、トランスクリプトームデータから自動的にSSCVを検出し、DBを自動更新するパイプラインを開発する。遺伝性疾患やがんの原因ともなるSSCVは、従来のゲノム解析では難しく見落とされがちだったが、本DBを使って疾患とSSCVとの関連性を明らかにすることにより、SSCVの検出による疾患診断法の開発や、SSCVを標的とする核酸医薬の研究開発などへの寄与を目指す。



4
す
P
理

応募課題がカバーする分野が更に拡大

独自に開発した解析手法 (PHi-C) を用い、公共の3次元ゲノム構造 (Hi-C) データに対し、動的な3次元ゲノム構造 (4Dゲノム状態) を網羅的に解析・可視化するPHi-C databaseを構築する。GPUサーバーを用いたPHi-C計算の高速化と大規模化を進め、複数のHi-Cデータを直感的に比較・理解できる4Dゲノム情報の統合的可視化基盤を実現する。本DBを通じて、真核生物における生命現象の解明に不可欠な細胞核内におけるゲノムDNAの動的な高次構造情報を提供し、シミュレーションや理論解析、ゲノム統合解析を通じた新たな仮説創出や知識発見への貢献を目指す。



毒性病理の専門家と連携して収集・体系化した毒性病理画像DB CuraToxiiを構築する。毒性病理画像およびメタ情報の提供の他、深層学習モデルの開発や毒性評価への応用を支援するためのベンチマークデータセットの構築、リーダーボード機能、動的データ取得機能等を実装する。本DBによって毒性病理画像のファイル容量に起因する操作性の低さやデータ入手性といった既存の課題を解決し、化合物が個体に与える影響の理解に重要な情報を提供する。医薬品安全性評価の効率化・高度化とともに、化合物と個体への影響の理解深化に基づく創薬研究の加速化への貢献を目指す。

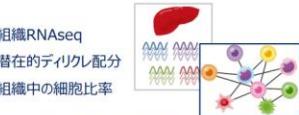
オミクスデータのパターン①

入力：培養細胞のマイクロアレイ
手法：因子分析
出力：化合物の複合的効用



オミクスデータのパターン②

入力：組織RNaseq
手法：潜在的ディリケル配分
出力：組織中の細胞比率



化合物構造のパターン

入力：化合物構造
手法：ニューラル機械翻訳
出力：化合物構造の数値表現



毒性病理画像のパターン

入力：毒性病理画像
手法：画像自己教師あり学習
出力：個体応答の潜在表現



2期生・3期生による開発課題に関する動画

トーゴーの日シンポジウムのホームページ (<https://biosciencedbc.jp/event/symposium/>)

AI駆動型データキュレーションによる中分子相互作用統合データベースの開発

11:05～

中分子は、これまで創薬が難しかった標的に対する新しい治療手段として注目されています。ペプチドや核酸など多様な中分子データを統合し、一元的にアクセスできるようにすることで、これらの研究開発が加速する可能性がある。我々は、近年発展している構造予測技術を用いた相互作用のアノテーション法の確立や自然言語処理技術を用いたデータ収集にAI技術を活用することで、キュレーションの精度向上と効率化を目指す。本開発によって、創薬研究における中分子相互作用の総合的なデータベースの構築を目指す。

› 講演スライド (3.0MB) □ › 質疑応答一覧 › 講演動画 | 統合TV □



▶ 池田和由

理化学研究所 計算科学研究センター 上級研究員

細胞レベルの機能・表現型と遺伝子発現を関連付ける「Cell IO」データベースの開発

11:15～

多細胞生物における細胞レベルの機能・表現型に関する知識は、現在文献データと1細胞トランск립トームデータという2種類の公共データの中に散逸しており、気軽にアクセスしづらい。そこで本研究では、細胞型を「主語」とした生物学の知識の利活用を促進するため、細胞の機能・表現型を整理したデータベース「Cell IO（セル・アイオー）」（Cell Input/Output）の構築を目指す。具体的には、1細胞トランск립トームデータと紐づく細胞の機能・表現型と撰謄情報を整理するとともに、文献から細胞の機能・表現型に関する情報を抽出し、統合して提供する。本発表では、Cell IOの構想について紹介し、開発の方向性について議論する。

› 講演スライド (1.6MB) □ › 質疑応答一覧 › 講演動画 | 統合TV □



▶ 尾崎遼

筑波大学 医学医療系 准教授

創発的再解析のためのメタボローム統合データベース

11:25～

メタボロミクスは生物の代謝プロセスの結果として生じる代謝物群を包括的に分析することで生理現象の背後にある詳細で膨大な代謝情報を提供する。近年多様な研究分野における分析データが急速に蓄積されているが、研究間でのデータ統合およびその活用は依然として困難である。そこで我々は大規模メタボロミクスデータの研究横断的な統合再解析プラットフォームの開発を行っている。代謝物の変動情報と研究メタデータでのネットワーク化を行うことで広範な研究を統合し柔軟な再解析を可能にしている。様々な背景をもつ研究データを統合したネットワークは、単独の研究では見落とされた新しい発見や様々な仮説生成を可能にすることが期待できる。

› 講演スライド (3.8MB) □ › 質疑応答一覧 › 講演動画 | 統合TV □



▶ 早川英介

理化学研究所 環境資源科学研究センター 客員研究員 / 九州工業大学 情報工学研究院 准教授

大規模言語モデルを活用した病的スプライシング変異データベースの自律的構築

14:25

SSC DBは、Sequence Read Archive等の大規模な公共トランスクryptオームデータを、独自のアルゴリズムで再解析し、同定したスプライスサイト生成変異を格納するデータベースである。

本研究課題では、大規模言語モデルを活用した病的予測アノテーションの追加を行う。さらに、シークエンスレポジトリへの新規データの追加に伴い、自律的に再解析を実行し、データベースを更新する基盤の開発を進める。

› 講演スライド (2.7MB) □ › 講演動画 | 統合TV □



白石友一

国立がん研究センター 研究所 ゲノム解析基盤開発分野 分野長

4Dゲノム状態の理解と可視化を支援するデータベースの構築

14:35

独自に開発したPHi-C（ファイシー）法を用いて、公共の3次元ゲノム構造データ（Hi-Cデータ）に対し、動的な3次元ゲノム構造（4Dゲノム状態）を網羅的に解析し、その可視化を実現する「PHi-C database」を構築する。GPU計算による高速化と大規模化を進め、複数のHi-Cデータを直感的に比較・理解できる4Dゲノム情報の統合的可視化基盤を提供する。これにより、細胞核内におけるゲノムDNAの動的高次構造の理解を促進し、シミュレーションや理論解析、統合ゲノム解析を通じて新たな仮説の創出と知識発見に貢献することを目指す。

› 講演スライド (5.2MB) □ › 講演動画 | 統合TV □



新海創也

理化学研究所 生命機能科学研究センター 上級研究員

データサイエンティストフレンドリーな毒性病理画像データベースCuratoxiiの開発

14:45

毒性病理画像は医薬品等の安全性評価にて欠かせない他、化合物の個体での作用を恣意性なく記述するデータである。

一方、疾患画像に比してデータベース(DB)の整備が進んでおらず、当該データを対象とするデータサイエンティストは限られている。

上記課題解決に向け我々は、データサイエンティストフレンドリーな毒性病理画像データベースCuratoxiiの開発に取り組んでいる。

本発表では毒性病理画像解析分野の背景となぜDB開発がその課題解決に資するかを中心に、プロジェクトの概要を紹介する。

› 講演スライド (2.2MB) □ › 講演動画 | 統合TV □



水野忠快

東京大学 大学院薬学系研究科 助教 / 統計数理研究所統計思考院 特任助教

育成型に求めること

課題全般に求めること

- ・「つなぐ」から「使う」へ
- ・広範なユーザーの知識発見支援
- ・国際的なプレゼンス
- ・新しい動向への対応

育成型に求めること

- ・構想に新規性・独自性・挑戦性・将来性が感じられること
- ・実績は問わないが、実現可能性が感じられること
DB構築の戦略に不安を感じさせない陣容
- ・既存DBの新展開の提案に、独自性・挑戦性・将来性が感じられること
従来の延長線上に留まらない飛躍・新機軸

バイオデータサイエンスの流れ

- ・ 実験技術・研究手法の急速な発展
- ・ 研究スタイルの変化
- ・ データ(ベース)の量的・質的変容
- ・ データ(ベース)駆動型研究への期待
- ・ AIの急速な進歩が期待を更に加速

Nobel prize

- | | |
|------|-----------------------------|
| 2002 | MS, NMR |
| 2006 | RNAi |
| 2007 | KO mouse |
| 2008 | GFP |
| 2014 | Super resolution microscopy |
| 2017 | Cryo-EM |
| 2020 | CRISPR/Cas |
| 2024 | AlphaFold2/protein design |

Progress in science depends on
new techniques, new discoveries,
and new ideas, probably in that
order. Sydney Brenner

Highly accurate protein structure prediction with AlphaFold

<https://doi.org/10.1038/s41586-021-03819-2>

Received: 11 May 2021

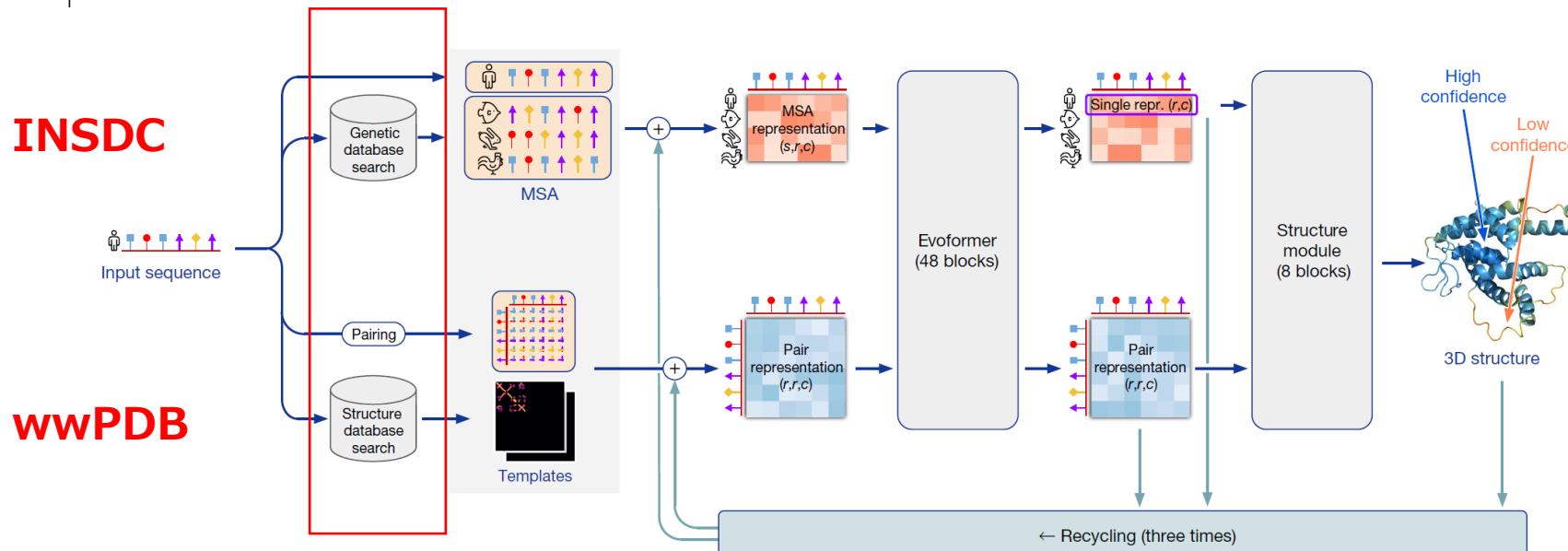
Accepted: 12 July 2021

Published online: 15 July 2021

Open access

 Check for updates

John Jumper^{1,4}✉, Richard Evans^{1,4}, Alexander Pritzel^{1,4}, Tim Green^{1,4}, Michael Figurnov^{1,4}, Olaf Ronneberger^{1,4}, Kathryn Tunyasuvunakool^{1,4}, Russ Bates^{1,4}, Augustin Žídek^{1,4}, Anna Potapenko^{1,4}, Alex Bridgland^{1,4}, Clemens Meyer^{1,4}, Simon A. A. Kohl^{1,4}, Andrew J. Ballard^{1,4}, Andrew Cowie^{1,4}, Bernardino Romera-Paredes^{1,4}, Stanislav Nikolov^{1,4}, Rishabh Jain^{1,4}, Jonas Adler¹, Trevor Back¹, Stig Petersen¹, David Reiman¹, Ellen Clancy¹, Michal Zielinski¹, Martin Steinegger^{2,3}, Michalina Pacholska¹, Tamas Berghammer¹, Sebastian Bodenstein¹, David Silver¹, Oriol Vinyals¹, Andrew W. Senior¹, Koray Kavukcuoglu¹, Pushmeet Kohli¹ & Demis Hassabis^{1,4}✉



バイオデータサイエンスの流れ

- ・本格的なデータ(ベース)駆動型研究の時代
- ・マルチモーダルデータの統合が求められる時代
- ・データ統合による知識発見プロセスの革新
- ・AIによる科学研究自体の一大変革期



AI時代に相応しいDBの作り方とは？

AI時代のDBの使われ方とは？

AI時代におけるバイオDBのあるべき姿とは？

これからのDICP

- どんな時代にあっても基本は「しっかりとしたコンテンツ」
 - “Garbage in, garbage out”的防止
 - 知識グラフの統合によるAIの強化
- 「AIを活用すること」と「AIによる利用を意識すること」は当然
- その先を行くような斬新な提案を期待

我が国のバイオサイエンスの状況

- 研究環境の格差が拡大し、研究の機会均等性の喪失
先端技術・先端機器へのアクセシビリティには大きな格差が存在
- 研究(者)の多様性の喪失
トレンドに適応した研究者のみが繁栄、それ以外は絶滅危惧種...
多様性こそが進化の源泉、狭い裾野に高いピークは立たない
- 研究(者)多様性保全策としての公共データの統合的利活用環境
研究環境に恵まれない研究者にとっての生存戦略のひとつ
👉 公共性に富む本プログラムのもうひとつの意義

DBを評価する風土・文化の醸成

- 目立つ成果に対する直接的な貢献しか評価できない未熟さ
無駄な研究活動の回避という貢献に思い至らない「想像力の貧困」
 - データベースの構築と維持というエッセンシャルワークの軽視
華やかなAIの影にDBが隠れてしまうことへの危機感
- ☞ 出口しか評価できない風土・文化からの脱却を（飲水思源）
評価軸の成熟・複線化へ

バイオデータリソースのサステナビリティは世界共通の課題



The screenshot shows the Global Biodata Coalition's website. At the top, there is a navigation bar with the logo 'GLOBAL BIODATA COALITION' (a blue globe icon) on the left and links for 'WHAT WE DO', 'MEMBERSHIP', 'NEWS & RESOURCES', 'CONTACT', and social media icons for X and LinkedIn on the right. Below the navigation is a red banner with the text 'SIGN THE GBC OPEN LETTER'. The main content area has a dark blue background with a network of glowing blue dots and lines. The text 'GLOBAL BIODATA COALITION' is centered in blue. Below it, a paragraph in English discusses the risk to biodata resources and the need for sustained and open access through financial support. Another paragraph describes the GBC's role in convening funders and providing a forum for data resource managers. At the bottom of this section, there is Japanese text: '国際協調なくしてバイオデータリソースの維持発展は不可能' (It is impossible to maintain and develop biodata resources without international cooperation). The bottom of the page features a white footer with text in Japanese: '世界中のライフサイエンスおよび生物医学研究にとって
長期的な持続性が重要と考えられるリソースを認定' (Recognize resources that are important for the long-term sustainability of life sciences and biomedical research around the world), '日本が関与するものとしては、DDBJ、PDB、PXCが選定' (Selected as involved in Japan: DDBJ, PDB, PXC), and '将来これらに続くような新しいDBの芽を育成型で支援したい' (Want to support the growth of new DBs like these in the future).

WHAT WE DO ▾ MEMBERSHIP ▾ NEWS & RESOURCES ▾ CONTACT X in

SIGN THE GBC OPEN LETTER

GLOBAL BIODATA COALITION

Biodata resources and the critical ecosystem they form are at risk. As these data continue to grow, we must ensure they are sustained and remain open into the future through effective, long-term financial support.

The GBC convenes the world's research funders to exchange knowledge and share strategies for supporting biodata resources, and provides a discussion forum for the managers of data resources, with the aim of developing principles and models for the coordinated funding of global core biodata resources.

国際協調なくしてバイオデータリソースの維持発展は不可能

世界中のライフサイエンスおよび生物医学研究にとって
長期的な持続性が重要と考えられるリソースを認定

日本が関与するものとしては、DDBJ、PDB、PXCが選定

将来これらに続くような新しいDBの芽を育成型で支援したい

運営方針

- 強化した研究アドバイザーに加えて、適宜、外部有識者の助言・協力も仰ぎます。
- サイトビジットに加えて、オンライン個別面談も活用して、十分な意思疎通を図ります。
- 本格型も含む全課題参加の対面交流会を開催して、課題間連携の促進による相乗効果の創出を図ります。
- 資金獲得以外の面においても、参加してよかったですと思って頂けるようなプログラムを目指します。

最後に

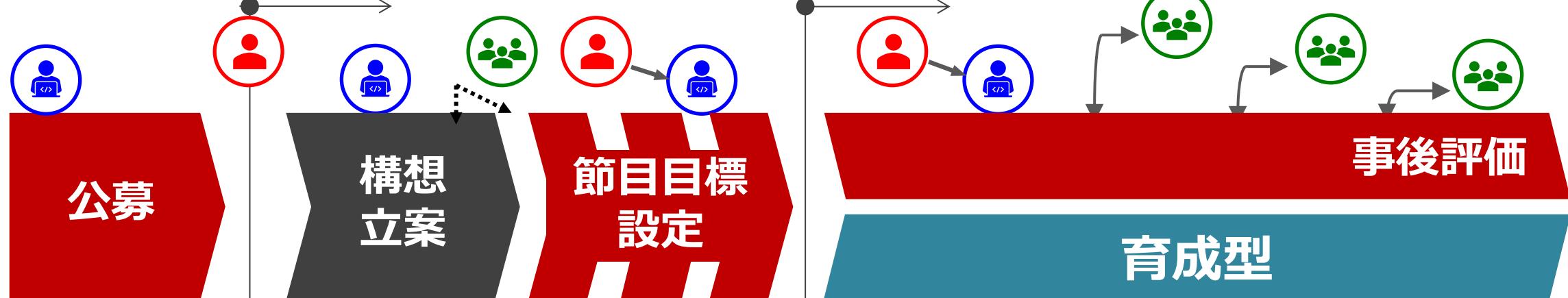
- 生命科学の研究動向を見据えて、広範なユーザーの知識発見・課題解決に資するようなデータベースを、高い志と情熱をもつて構築してください。
- 優れた構想、情熱と公共心、そして着実な実行力を兼ね備えた若々しい提案を期待しています。
- 皆さまの積極的な応募をお待ちしております。

提案・選考に関する注意事項など

プログラムのマネージメント



選考・採択 研究開始 2年次 3年次



成果像の明示

- ✓ 利活用・価値の視点で明確化

目標の細分化・密なフィードバック

- ✓ 研究の節目ごとの目標を設定
- ✓ 節目目標は国内外コミュニティのニーズ・利用状況を踏まえ設定
- ✓ 研究開始前にマネジメント側との節目目標を合意

年次レビュー、隨時の検証も実施

- ✓ 進捗遅延・目標に向けた道筋からの逸脱を早期把握・介入
- ✓ 有効性検証・利活用例の創出

研究総括・アドバイザリーボード

【研究総括】

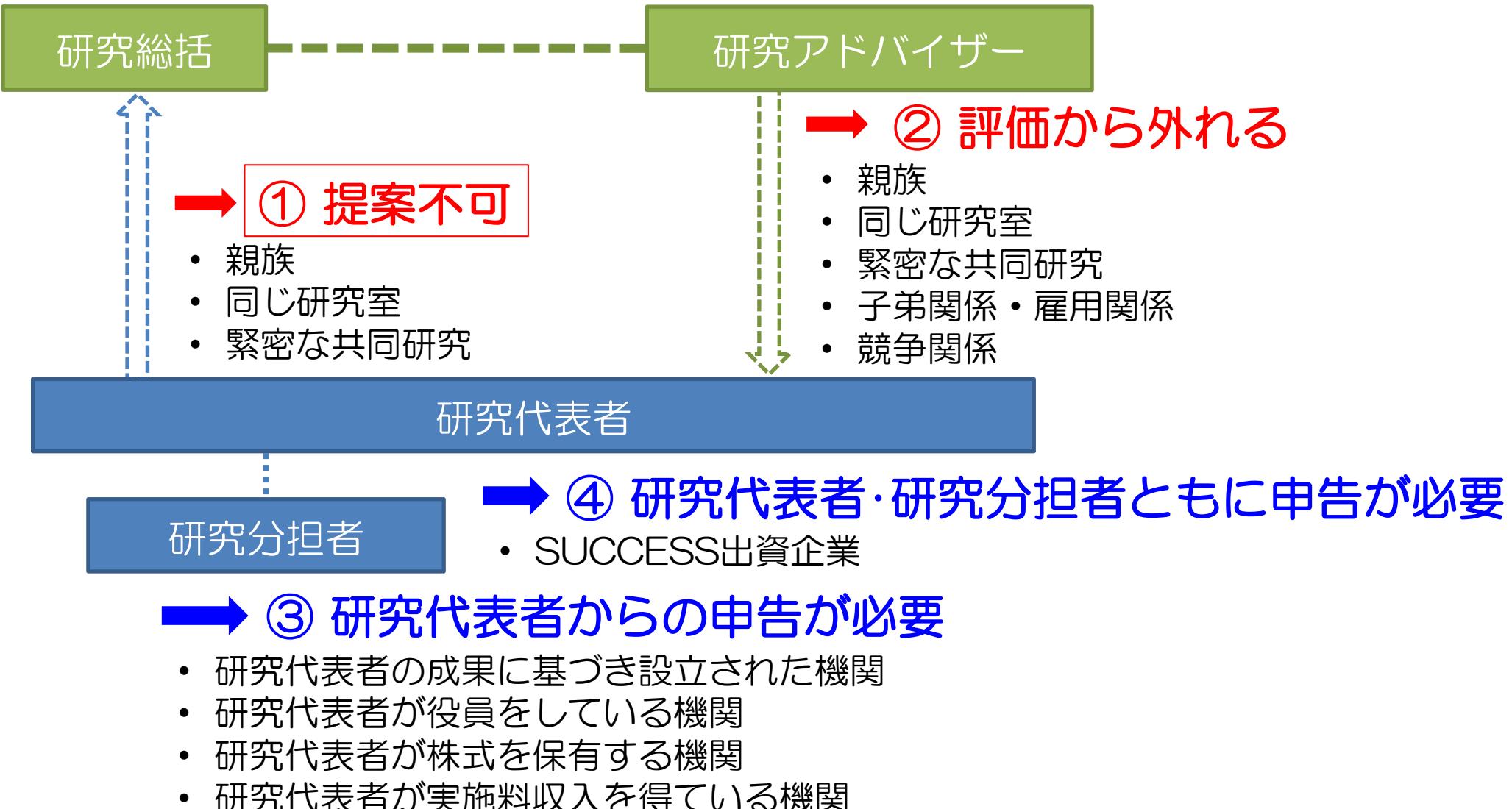
伊藤 隆司 九州大学 生体防御医学研究所 特任教授

【研究アドバイザー】 (50音順)

鎌田 真由美 北里大学 未来工学部 教授
坂井 寛章 農業・食品産業技術総合研究機構 高度分析研究センター ユニット長
清水 佳奈 早稲田大学 理工学術院 教授
白井 剛 長浜バイオ大学 バイオサイエンス学部 教授
瀬々 潤 (株)ヒューマノーム研究所 代表取締役社長
夏目 やよい 医薬基盤・健康・栄養研究所 AI 健康・医薬研究センター センター長
馬場 健史 九州大学 生体防御医学研究所 主幹教授
本郷 裕一 東京科学大学 生命理工学院 教授
山本 一夫 お茶の水女子大学 ヒューマンライフサイエンス研究所 客員教授
吉田 哲郎 アクセリード(株) 経営企画部 シニアディレクター

※ 必要に応じ、上記以外の外部有識者に加わって頂く場合があります。

利益相反について



研究開発期間・研究費・採択予定課題数

- 研究開発期間

- 育成型

2026年4月から3年以内（予定）

- 研究開発費（直接経費）

- 育成型

～1,000万円／年度程度（予定）

※ 委託研究契約に基づき、研究機関に対して直接経費とは別に、原則として直接経費の30%の間接経費を支払います。

- 採択予定課題数

- 育成型

～3課題程度（予定）

※ 今回は「本格型」の募集は行いません。

公募期間・選考スケジュール

1. 公募締切：

2026年 1月27日(火) 正午 厳守

【注意事項】

締切り間際はe-Radへのアクセスが集中し、入力できない等のトラブルが発生するおそれがありますので、e-Radでの手続きはなるべく早めに完了するようにしてください。

2. 書類選考結果の通知： 2026年 3月2日(月)までに通知

【ご注意】書類選考結果通知から面接選考会までの日にちが短いのでご注意ください。

3. 面接選考会： 2026年 3月6日(金) (都内での対面開催)

4. 選定課題の通知・発表： 2026年 3月下旬以降 (※)

5. 研究開発の開始： 2026年 4月 (※)

※ 課題選考手続きの関係で予定よりも遅れることがあります。

募集説明会の動画とQ&Aの公開

- 質疑応答のパートを除いた募集説明会の動画と説明資料、ならびに本日の質疑応答でご紹介した質問と回答は、近日中に公開いたします。

下記のNBDC公募サイトをご覧ください。

<https://biosciencedbc.jp/funding/calls/2026.html>

皆様からの多数の応募をお待ちしております

問い合わせ先

● JST NBDC公募担当：

- ・本プログラムの内容、応募手続き・利害関係等についての問い合わせ
- ・応募後に他の競争的研究費等に採択された場合の連絡

E-mail : nbdc-funding@jst.go.jp

Tel : 03-5214-8491 (受付時間：10:00～12:00/13:00～17:00)

※緊急時以外は、メールにて問い合わせ願います。

※土、日、国民の祝日及び年末年始（12月29日～1月3日）を除く

● JST 法務・コンプライアンス部 研究公正課：

- ・不正経理、研究不正、研究倫理教育に関するプログラム等

E-mail : rcr-kousyu@jst.go.jp

● 内閣府 e-Radヘルプデスク：

- ・e-Radの操作方法

<https://www.e-rad.go.jp/contact.html>

Tel : 0570-057-060 (受付時間：9:00～18:00)

※土、日、国民の祝日及び年末年始(12月29日～1月3日)を除く