

# 生命科学データベース統合に関する調査研究

## DB統合が望まれる背景 ～研究内部の変化と研究環境の変化～

- ・ 巨大観察データの登場によるライフサイエンス全体のデータ(ベース) 依存型への変貌
- ・ 高額な巨大データに関して大学・企業研究開発の公共データ(ベース)への依存増
- ・ 科学の民営化 (データ占有による知財獲得、データベースサービスも競争的)

多面的かつ  
複雑な背景

## 課題：「生命科学データベース統合に関する調査研究」

実施内容：①国内外のDB調査、産業界等からのヒアリング

②DB統合に向けた技術的FS：省をまたいだDB連携

③DB統合に向けた制度調査：DB統合が進む米国等の制度調査

技術的課題は  
十分対応可能

制度上の問題  
を明確化

NCBIを代表とする米国での統合DBの成功例とそれを支える社会の仕組み(制度)の変化  
OECD等国际的な公的機関によるデータ共有の必要性の認識

## DB統合化に向けた提言に資する試案

1. 「政府資金によるデータ産生型プロジェクトのデータを我が国の研究社会で早期に共有するためのルール」が必要
2. 公共財としてのデータを保全・管理し、長年にわたるデータの育成と共有を行う公的機関を設ける(ルール運用の際の受け皿)

# 我が国における統合DBプロジェクト

- 内閣府主導の統合データベースプロジェクト(H18～)
  - 文科省、経産省、農水省、厚労省で実施
  - H23年12月に四省連携のポータルサイト
- 文科省の統合データベースプロジェクト(H18～)
  - データを公共財化・コモンズ化し、その価値を最大化
  - 中核センターの設立
    - H19～ライフサイエンス統合データベースセンターDBCLS
    - H23～バイオサイエンスデータベースセンターNBDC
  - CCライセンスの採用
  - フォーマット、辞書、統合技術、動画教材などの開発
  - カタログ、横断検索、アーカイブの構築など
  - 分野ごとのデータベース統合化進行中(ファンディングによる)
  - ヒト由来データの共有・セキュリティガイドラインの作成

RDFによる  
すべてのDBの統合

連邦型統合DB

DB統合化基盤技術の開発

目的別統合DB

微生物DB

植物DB

ヒトDB

データバンク的DB

ゲノム

タンパク質立体構造

メタボローム

糖鎖

DBをまとめることで  
使いやすく

統合化推進プログラム

個々のDBを  
使いやすく

カタログ

横断検索

アーカイブ

標準化DB3  
標準化DB4b

データベース

DB1a

DB1b

DB2

DB3

DB4a

DB4b

プロジェクト

P1

P2

P3

P4

測定装置

顕微鏡

MRI

chip

NGS

質量分析

NMR

NBDCは、日本の生命科学研究を推進するために、データベースをつなげて使い易くします。そのためにNBDCや協力機関は、以下のようなサービスやウェブサイトを作成・提供しています。

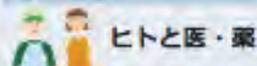
生命科学全体のデータベース統合

[Integbioデータベースカタログ](#)

データベース横断検索

[生命科学系データベースアーカイブ](#)

分野ごとのデータベース統合



ヒトと医・薬

[NBDCヒトデータベース](#)

[ヒトゲノムバリエーションデータベース](#)

[ヒト疾患脳画像データベース](#)

[KEGG MEDICUS: 疾患・医薬品統合リソース](#)



生命を支える分子

[DDBJ: 日本DNAデータバンク](#)

[PDBj: 日本蛋白質構造データバンク](#)

[TogoProt: 蛋白質関連データベース統合検索](#)

[JCGGDB: 日本糖鎖科学統合データベース](#)

[MassBank / Bio-MassBank / KNApSACk Family](#)



ゲノムから個体へ

[PGDBj: 植物ゲノム統合データベース](#)

[MicrobeDB.jp: 微生物関連データベース](#)

日本語や動画でわかりやすく

[新着論文レビュー / 領域統合レビュー](#)

[統合TV](#)

論文をもっと読みやすく、書きやすく

[Allie / inMeXes / TogoDoc](#)

大量の配列データを扱いやすく

[DBCLS SRA](#)

[RefEx / 統合遺伝子検索 GGRNA](#)

さまざまな統合コンテンツ

[生物アイコン](#)

[生命科学系主要プロジェクト一覧](#)

[Webリソースポータルサイト](#)

[ゲノム解析ツールリンク集](#)

[HOWDY / GenLibi](#)

開発ツール

[TogoDB / TogoWS](#)

[DBCLS Galaxy](#)

[BodyParts3D / Anatomography](#)

ライフサイエンスDBに関するアンケート実施中

2015/09/05更新



NBDCパンフレット

(PDF: 3.73MB /

2014/09/05更新)

新着情報

twitter RSS

2015/01/21

【メンテナンス】2015年2月20日（金）18:00～23日（月）10:00の間、生命科学横断検索とTogoProt、アーカイブ内横断検索が停止します。

2015/01/19

2015年度 統合データベース講習会の受け入れ機関の募集を開始しました。

2015/01/14

「日化辞InChI対応表」（独立行政法人 科学技術振興機構）を「生命科学系データベースアーカイブ」に追加しました。

2015/01/14

【横断検索】12件のデータベースが検索できるようになりました

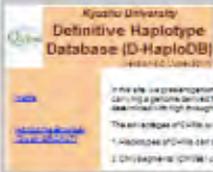
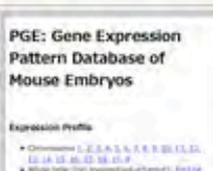
2015/01/08

## アーカイブデータベース一覧 (ヘルプ)

一覧内検索

全 93 件 (56 件から60件)

5 件を表示 Previous Next

データベース	データベース運用場所	代表者	データベースカテゴリー	生物種	要約 (キーワードを太字表示)	利用許諾
 <p><b>D-HaploDB</b> ダウンロード   簡易検索 オリジナルサイト</p>	九州大学 生体防御医学研究所 附属遺伝情報実験センター ゲノム構造学分野	林 健志	多型	ヒト	全胎状奇胎を用いて直接決定した <b>日本人ゲノムのハプロタイプ</b> と連鎖不平衡構造のデータベース	CC 表示-継承詳細
 <p><b>RIKEN SSBC</b> ダウンロード   簡易検索 オリジナルサイト</p>	独立行政法人 理化学研究所 生命分子システム基盤研究領域	横山 茂之	立体構造	ヒト、マウス	高等動植物等 <b>タンパク質の結晶観察画像・X線回折画像</b> ・付随実験情報のデータベース	CC 表示-非営利-改変禁止詳細
 <p><b>Micro-Tom BAC End Sequence Database</b> ダウンロード   簡易検索 オリジナルサイト</p>	筑波大学大学院生命環境科学研究科遺伝子実験センター	浅水 恵理香	DNA配列	トマト	トマト品種 <b>マイクロトムのBACライブラリー</b> を構成する <b>クローンの末端配列情報</b> のデータベース	CC 表示-継承詳細
 <p><b>PGE</b> ダウンロード   簡易検索</p>	独立行政法人 科学技術振興機構	ERATO土居バイオアシンメトリプロジェクト	発現	マウス	<b>マウス胚の初期発生</b> における <b>遺伝子発現パターン</b> データベース	CC 表示-継承詳細
 <p><b>HHDB</b> ダウンロード   簡易検索 オリジナルサイト</p>	日本原子力研究開発機構	玉田 太郎	立体構造	生物全般	水素及び水和水を含む <b>タンパク質や核酸の立体構造データ</b> (中性子回折法や高分解能X線回折法による) に含まれる <b>水素結合</b> のデータ	CC 表示-継承詳細

# 公募要領にデータ提供協力依頼記載

- 文科省ライフ課委託プロジェクト(H20～)
- JST戦略事業(CREST、さきがけ)(H23～)
- 厚労科研費(H24～)
- 文科省科研費(H25～)

## 6.8 バイオサイエンスデータベースセンターへの協力

ライフサイエンス分野の本事業実施者は、論文発表等で公表された成果に関わる生データの複製物、又は構築した公開用データベースの複製物を、バイオサイエンスデータベースセンター(※)に提供くださるようご協力をお願いします。提供された複製物は、非独占的に複製・改変その他必要な形で利用できるものとします。複製物の提供を受けた機関の求めに応じ、複製物を利用するに当たって必要となる情報の提供にもご協力をお願いします。

# データ共有に関わる今後の課題

- 省間連携の司令塔機能の再構築
  - AMEDとの連携協力、センター間の連携、国際的な共有の枠組み構築
- ファンディング機関によるデータ共有の義務化
  - 公募要領のデータ提供協力依頼記載では不十分
- ヒト由来データなどの機微情報の扱い
  - 制限アクセスの導入とそのためのガイドラインの策定
  - 種々の医学研究、ヒトゲノム研究の指針との整合性
  - 外国からの利用への対応
- データ量の爆発、データ種類の多様化への対応
  - 捨てるもの vs. 残すもの、再利用性の高い vs. 低いもの
  - 生活習慣情報や環境情報などへの対応
  - 持続可能な予算確保の仕組みやデータ受入れの省力化
- フォーマットや用語の標準化、データの統合、品質管理、技術開発
- 知識DBの整備とデータとの統合
- バイオインフォマティクス人材の育成、確保
- 出てきたデータの共有からプロジェクトの立案、実行時への貢献
- 研究不正への対応