

○時松敏明、青野英雄、李慶範、飯塚朋代、大城戸利久、川添将仁、児玉悠一、小菅武英、杉田里江、鈴木紀美子、高木佳苗、秦千比呂、福田亜沙美、真島淳、横山会美、藤澤貴智

国立遺伝学研究所 生命情報・DDBJセンター

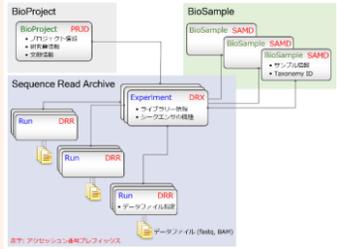
### 要 旨

DDBJは、30年以上にわたりNCBIおよびEBIと国際塩基配列データベース(INSDC)を協同運営し、遺伝研スーパーコンピュータを基盤として国内外の生命科学研究から産み出されるデータの登録・共有・解析サービスを提供している。遺伝研スパコンは、昨年三月に一新し、ゲノム解析環境の強化(Singularityコンテナおよび個人ゲノム解析対応区画)、データサイズの増加に対応する階層ストレージを導入したが、今年4月に、計算ノード数を増やすことでさらにシステムを増強した。DDBJが運営するサービスは、アノテーションを付与/アセンブルした塩基配列データ:DDBJ、次世代シーケンスデータ:DRA、機能ゲノクスデータ:GEA、アクセス制限のある個人ゲノムデータ:JGA、研究プロジェクト・サンプルのデータ:BioProject・BioSampleがあり、さらにメタボロームレポジトリ:MetaboBank およびヒトバリエーションデータベース:JVarが加わる予定である。DDBJは生命科学研究の基盤となるためのデータベースを維持・強化しつづけている。



大学共同利用機関法人情報・システム研究機構国立遺伝学研究所生命情報・DDBJセンター  
〒411-8540 静岡県三島市谷田1111 問い合わせ：<https://www.ddbj.nig.ac.jp/contact.html>

## DDBJが運営するサービス

**DRA (DDBJ Sequence Read Archive)**

次世代シーケンス生データとアライメント情報

アライメント

リード  
Quality value**DDBJ (Traditional)** アノテーションを付与した/アセンブルした塩基配列

アセンブリ



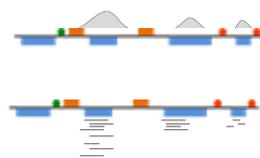
登録:

&lt;= Web版塩基配列登録システム (NSSS)

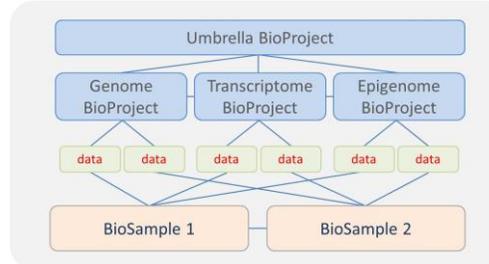
&lt;= MSS - Mass Submission System

**GEA (Genomic Expression Archive)**

遺伝子発現・エピジェネティクス解析・マイクロアレイなど

機能ゲノミクス  
データ  
タグカウント  
ピークコール**MetaboBank (10月リリース(予定))**

メタボロームデータレポジトリ

**BioProject** プロジェクト情報**BioSample** サンプル情報**Jvar (Japan Variation Database; 準備中)**

SNP, INDEL, CNV, 構造変異等のバリエーション

SNP 短いバリエーション



SV 構造変異

**JGA (Japanese Genotype-phenotype Archive)**

個人レベルの遺伝型と表現型 (匿名化情報)



ヒトデータ審査委員会が登録・利用の審査

9月に新しいグループアカウントシステムを導入

オープンアクセスデータベース

アクセス制限データベース

## NIGスーパーコンピューターシステム

<https://sc.ddbj.nig.ac.jp/ja>



- 2019年3月に新システム稼働
  - 2020年4月 計算ノードの総コア数増強
    - 11,696コア -> 15280コア
- ゲノム解析環境を強化
  - Singularity コンテナ
  - 個人ゲノム解析に対応した区画

## オープンアクセスデータベース

## アクセス制限データベース

**DRA** 次世代シーケンス  
生データと  
アライメント情報

**GEA** 遺伝子発現・  
エピジェネティクス解析・  
マイクロアレイなど

**DDBJ** アノテーション/  
アセンブルした  
塩基配列

DRA etc  
公開データセットに  
直接アクセス

利用承認された  
JGA データをコピー

個人ゲノムデータ  
をアップロード

**NIG**  
SUPERCOMPUTER  
NIG スパコン  
通常区画

**NIG**  
SUPERCOMPUTER  
NIG スパコン  
個人ゲノム解析環境

2020年4月～  
NIGスパコンシステム構成図