



スパコン概要紹介

国立遺伝学研究所／総合研究大学院大学
大量遺伝情報研究室 教授

中村 保一

March 22-23, 2018

「先進ゲノム支援」情報解析講習会

DDBJ

Mashima J, Kodama Y, Kosuge T, Fujisawa T, Katayama T, Nagasaki H, Okuda Y, Kaminuma E, Ogasawara O, Okubo K, Nakamura Y, Takagi T. (2016) **DNA data bank of Japan (DDBJ) progress report.** *Nucleic Acids Res.*, **44(D1)**: D51-57.

The screenshot shows the DDBJ Center website interface. At the top, there is a navigation bar with the DDBJ logo, "Services" with a dropdown arrow, and links for "Login & Submit", "Contact", and "Japanese". Below this is a header section with "DDBJ Center" on the left, a dropdown menu for "DDBJ Center Web Sites", and a "Google Custom Search" box with a search icon. A yellow banner below the header reads "Please send us your feedback to our new website." The main content area features a grid of eight service icons with labels: "Search & Analysis" (monitor with graph), "Submissions" (database cylinder), "Downloads" (download arrow), "SuperComputer" (server rack), "Statistics" (bar chart), "Activities" (pencil), "Training" (graduation cap), and "About Us" (group of people).



新しいウェブサイトへのアンケート調査にご協力をお願いいたします

DDBJ センターは生命科学研究から生み出されるデータの共有・解析サービスを提供することで広く研究活動をサポートしています

検索・解析



登録



ダウンロード



スパコン



統計



活動



講習会



センターについて



DDBJ センターからのお知らせ

2018年3月1日 | メンテナンス | [DDBJ](#) [BioProject](#) [BioSample](#) [DRA](#) [JGA](#) [AGD](#) [DDBJ Center](#)
[\(3/9-15\) スパコン定期メンテナンスに伴う DDBJ センターのサービス停止のお知らせ](#)

2018年3月1日 | お知らせ | [DDBJ](#)

INSDC

Cochrane G, Karsch-Mizrachi I, Takagi T, International Nucleotide Sequence Database Collaboration. (2016) **The International Nucleotide Sequence Database Collaboration.** *Nucleic Acids Res.*, **44(D1)**: D48-50.

The screenshot shows the INSDC website. The header features the INSDC logo and the text "International Nucleotide Sequence Database Collaboration". Below the header is a navigation menu with links for "ABOUT INSDC", "POLICY", "ADVISORS", and "DOCUMENTS". The main content area is titled "International Nucleotide Sequence Database Collaboration" and contains a paragraph describing the INSDC initiative. Below the text is a table summarizing the data types and their corresponding databases. The table has four columns: "Data type", "DDBJ", "EMBL-EBI", and "NCBI". The "EMBL-EBI" column is further divided into "European Nucleotide Archive (ENA)". The table lists data types such as "Next generation reads", "Capillary reads", "Annotated sequences", "Samples", and "Studies", along with their respective database links like "Sequence Read Archive", "Trace Archive", "DDBJ", "BioSample", and "BioProject".

International Nucleotide Sequence Database Collaboration

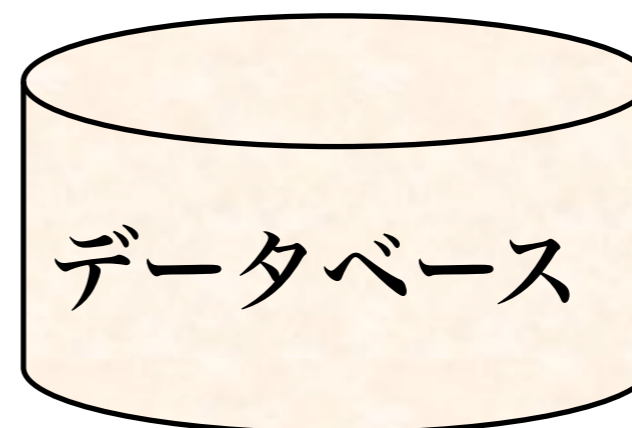
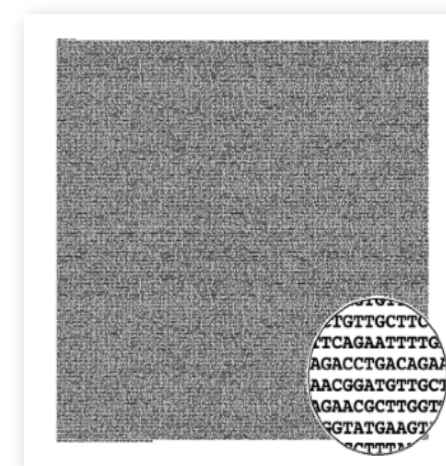
- The International Nucleotide Sequence Database Collaboration (INSDC) is a long-standing foundational initiative that operates between [DDBJ](#), [EMBL-EBI](#) and [NCBI](#). INSDC covers the spectrum of data raw reads, though alignments and assemblies to functional annotation, enriched with contextual information relating to samples and experimental configurations.

Data type	DDBJ	EMBL-EBI	NCBI
Next generation reads	Sequence Read Archive	European Nucleotide Archive (ENA)	Sequence Read Archive
Capillary reads	Trace Archive		Trace Archive
Annotated sequences	DDBJ		GenBank
Samples	BioSample		BioSample
Studies	BioProject		BioProject

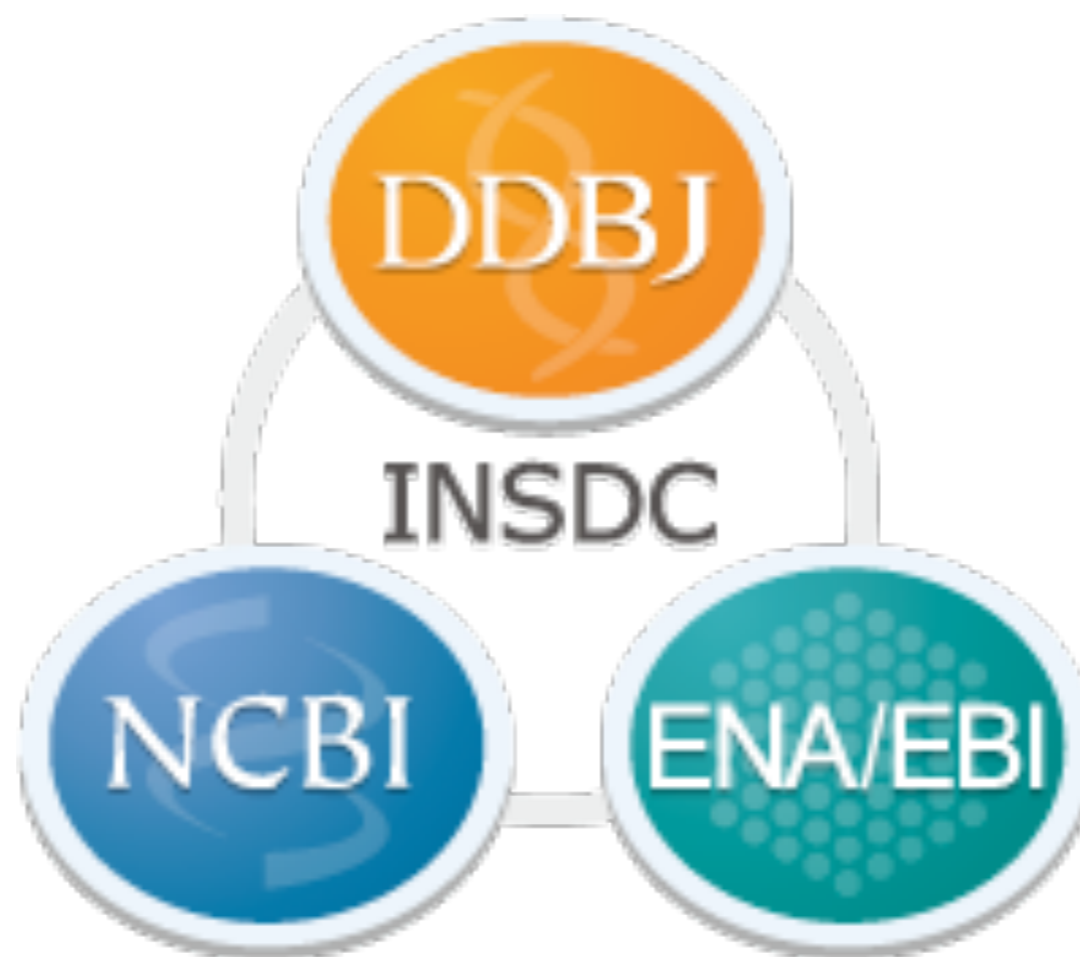
- The INSDC advisory board, the [International Advisory Committee](#), is made up of members of each of the databases' advisory bodies. At their most recent meeting, members of this committee unanimously endorsed and reaffirmed the

塩基配列データベースとはこのような事業

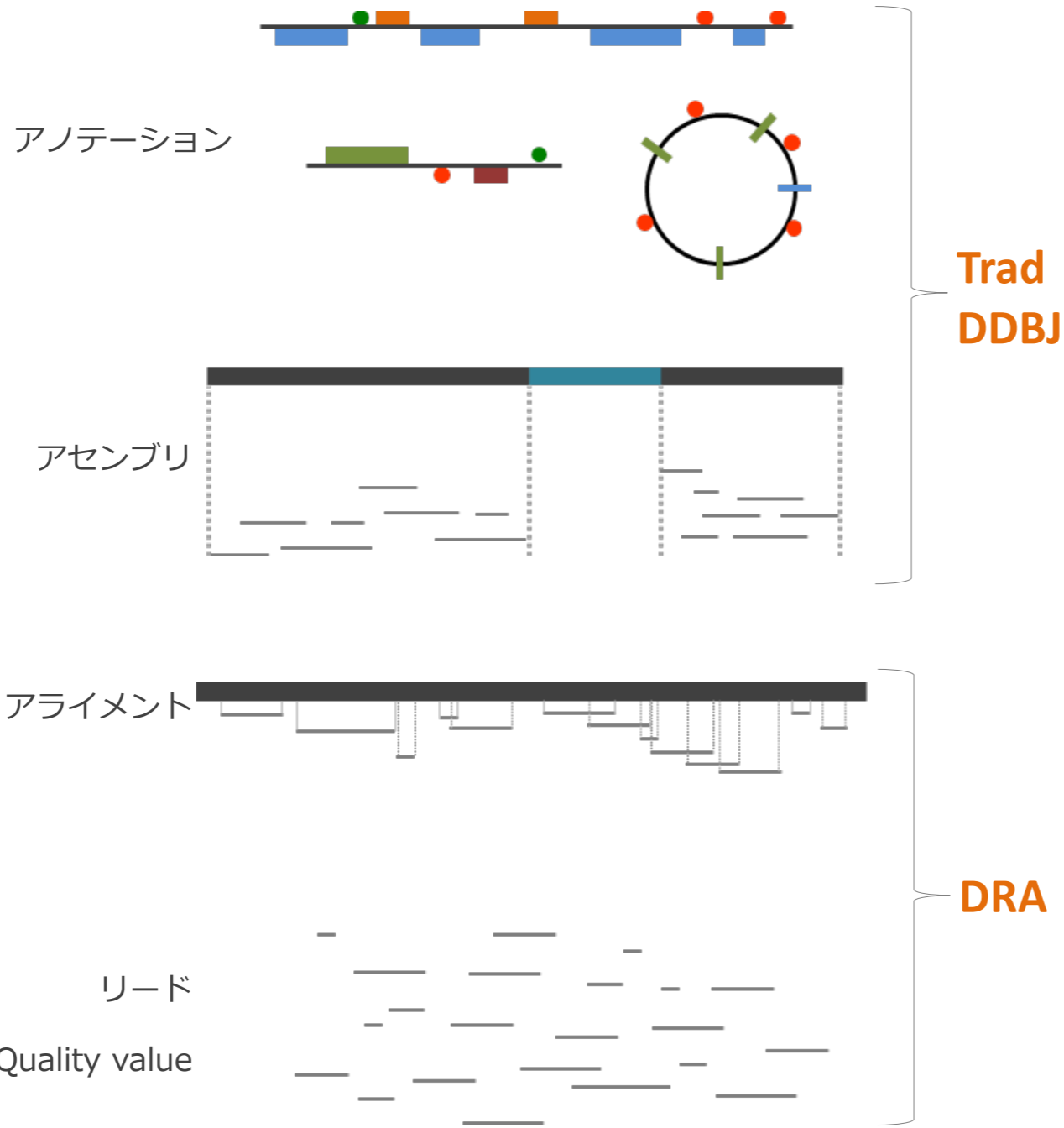
- 全世界で解読された塩基配列情報を
 - 査定して受入れ
 - データベースに蓄積し
 - 公開して共有する



- 米国: GenBank (NCBI)
- 欧州: ENA (EBI)
- 日本: DDBJ



DDBJ センター



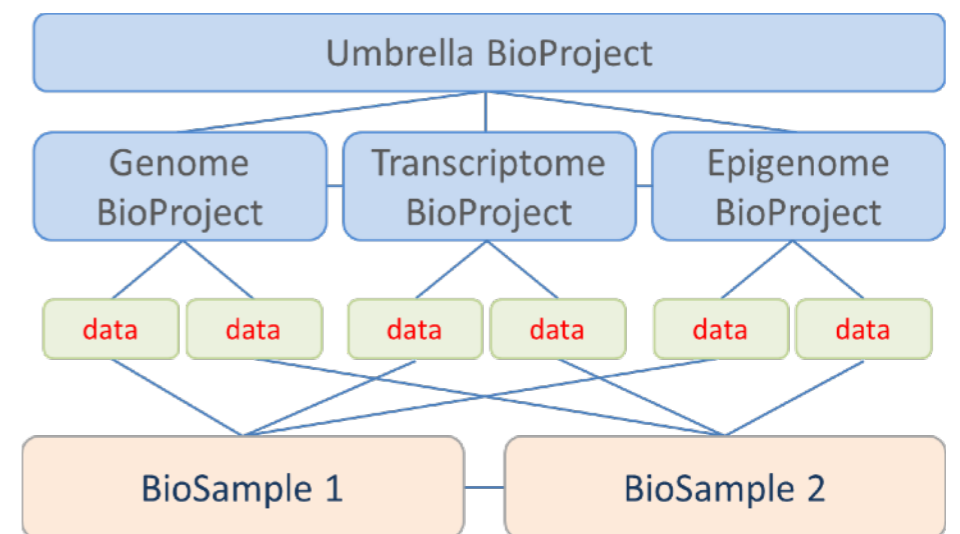
アクセス制限【制限公開】

JGA

個人の遺伝型と表現型



BioProject BioSample

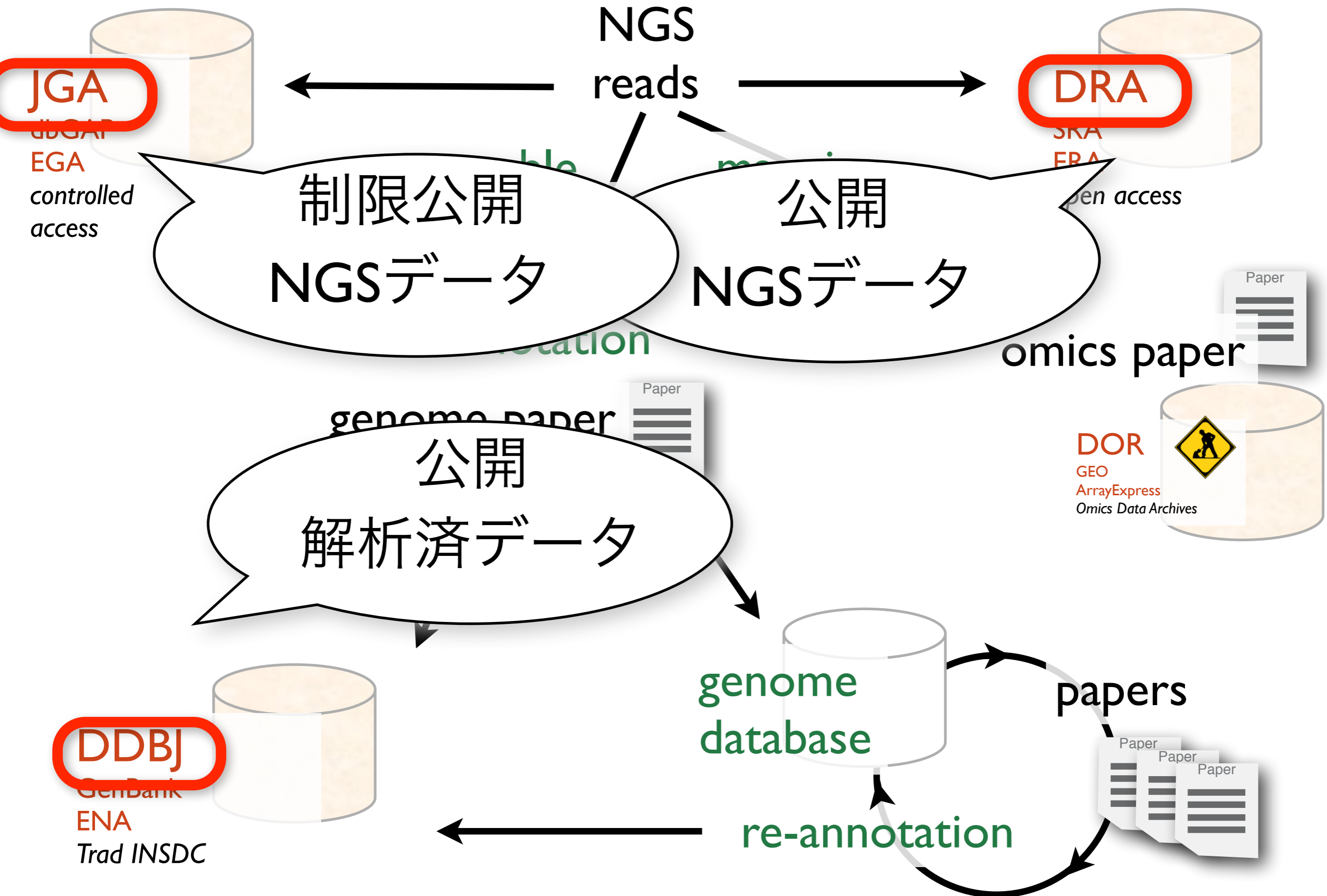


INSDC: オープンアクセス (非制限共有)

DRA: DDBJ Sequence Read Archive

JGA: Japanese Genotype-phenotype Archive

The ecosystem of sequence data



Trad INSDC

Trad DDBJ登録ファイルの例

LOCUS AB091058 2109 bp DNA linear BCT 02-SEP-2003
DEFINITION Gluconacetobacter xylinus cmcase, ccp genes for
endo-beta-1,4-glucanase, cellulose complementing protein, complete
cds.
ACCESSION [AB091058](#)
VERSION AB091058.1
KEYWORDS .
SOURCE Gluconacetobacter xylinus
ORGANISM [Gluconacetobacter xylinus](#)
Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;
Acetobacteraceae; Gluconacetobacter.
REFERENCE 1 (bases 1 to 2109)
AUTHORS Kawano,S., Tajima,K., Uemori,Y., Yamashita,H., Erata,T.,
Munekata,M. and Takai,M.
TITLE Direct Submission
JOURNAL Submitted (28-AUG-2002) to the DDBJ/EMBL/GenBank databases.
Contact:Kenji Tajima
Hokkaido University, Graduate School of Engineering; N13W8,
Kita-ku, Sapporo, Hokkaido 060-8628, Japan
REFERENCE 2
AUTHORS Kawano,S., Tajima,K., Uemori,Y., Yamashita,H., Erata,T.,
Munekata,M. and Takai,M.
TITLE Cloning of Cellulose Synthesis Related Genes from Acetobacter
xylinum ATCC23769 and ATCC53582: Comparison of Cellulose Synthetic
Ability Between ATCC23769 and ATCC53582
JOURNAL Unpublished (2002)
COMMENT
FEATURES Location/Qualifiers
source 1..2109
/db_xref="taxon:28448"
/mol_type="genomic DNA"
/note="synonym:Acetobacter xylinum"
/organism="[Gluconacetobacter xylinus](#)"
/strain="ATCC 53582"
CDS 10..1038
/codon_start=1
/gene="cmcase"
/product="endo-beta-1,4-glucanase"
/protein_id="[BAC82540.1](#)"
/transl_table=11
/translation="MSVMAAMGGAQVLSSTGAFADTAPDAVAQQWAI FRAKYLRPSGR
VVDITGGGESHSEGQYGMFLFAASAGDLASFQSMWMMWARTNLQHTNDKLF SWRFLKGH
QPPVPDKNNATDGDLLIALALGRAGKRFQRPDYIQDAMAIYGDVNLNMTMKAGPYVVL
MPGAVGF'KKDSVILNLSYYVMPSSLQAFDLTADPRWRQVMEDGIRLVSAGRFGQWRL
PPDWLAVNRATGALS IASGWPPRFSYDAIRVPLYFYWAHMLAPNVLADFTRFWNNFGA
NALPGWVDLTTGARSYPNAPPGYLAVAECTGLDSAGELPTLDHAPDYYS AALTLVYI
ARAETIK"

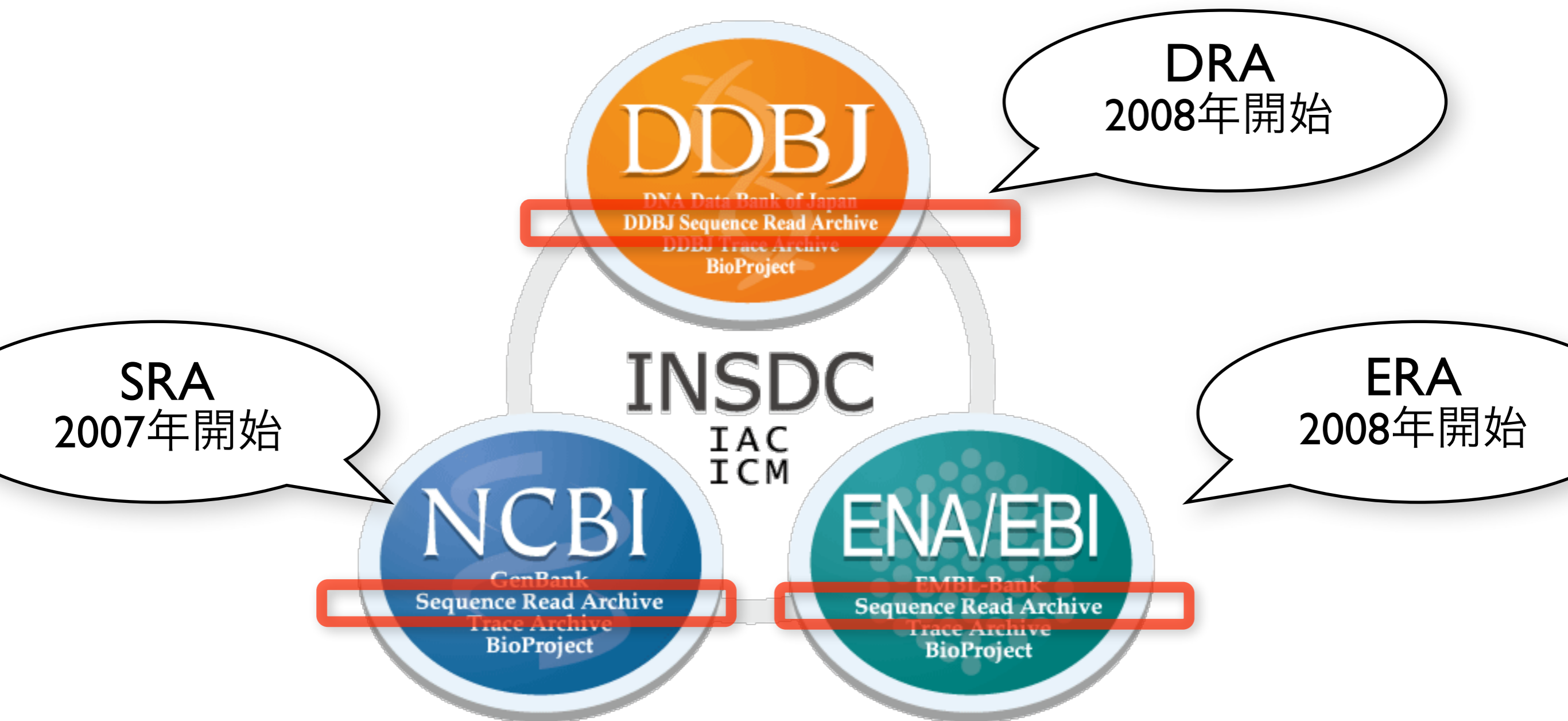
CDS 1035..2096
/codon_start=1
/gene="ccp"
/product="cellulose complementing protein"
/protein_id="[BAC82541.1](#)"
/transl_table=11
/translation="MSASGSDEVAGGGQAGSPQDFQVLRVLSFGVEGGQYSYRPFVDRS
FDVTGVPEAVERHFDQAEHDTAVEEQVTPAPQI AVAPPPPPVVPDPPAIVTETAPPPP
VVVSAPVTYEPAAAVPAEPPVQEAPVQAAPVPPAPVPPIAEQAPPAAPDPASVPYAN
VAAAPVPPDPAPVTPAPQARVTGPNTRMVEPF SRPQVRTVQEGATPSRVPSRSMNAFP
RTSASSISERPVD RGVADWSPVPKARLSPRERPRPGDLSFFFQGM RDRDEKFKFPV
ASTRSVRSNVS RMTSMTKTDTNSSQASRPGSPVASPDGSP TMAEVFMTLGGRATELLS
PRPSLREALRRRENEES"
BASE COUNT 343 a 661 c 661 g 444 t
ORIGIN
1 cgttccttta tgtcgtcat ggcggcgatg ggagggcgc aggtgcttcc atccaccggt
61 gcgttcgcag acaccgcccc cgatgcggtc gcgcagcaat gggcctatct ccgcgccaag
121 tatcttcgtc ccagcggacg tgtcgtggat acgggcaatg gtggcgaatc ccatagttag
181 gggcagggct atggcatgct ctttgcggcg tcggcggggg accttgcgtc gttccagtcg
241 atgtggatgt gggcgcgcac caacctgcag cataccaatg acaagctgtt ttcttgccgg
301 ttctcaagg ggcacagcc cccggtgccc gacaagaaca atgccacaga tggcgacctg
361 ctgatcgcgc ttgcgcttgg tcgtgcgggc aagcgtttcc agcgcgccga ttacattcag
421 gacgcatggt ccatttatgg cgatgtgctg aacctgatga cgatgaaggc gggaccgat
481 gtcgtcctca tgcccgggtg tgcggcttt accaagaagg acagcgtgat cctcaacctg
541 tcctattacg tcatgccttc gctgctgcag gcgttcgacc ttacggccga cccgcgctgg
601 cgtcaggtga tggaagacgg gattcgcctt gtttccgcog gccgtttcgg gcagtggcgc
661 ctgccccccg actggctggc ggtgaatcgc gccaccggtg cgtgtcogat cgcctgggga
721 tggccgcgcg gcttttcccta tgatgcgatt cgggtgcgcg tttatttita ttggcgcat
781 atgctggcgc cgaacgtgtt ggctgattc acccgattct ggaataatth cgggctaata
841 gccctgccag gatgggttga tctgacaaca ggggcgcgct cgccgtacaa cgccccgct
901 ggatatcttg ctggtgccga atgcacgggg cttgattctg ccggggaact cccgacactg
961 gatcatgcgc ccgattatta ttccgcagcg ttgacgctgc tcgtttacat cgcgcgggog
1021 gaggagacta taaagtgagt gcttcagggt ctgatgaggt ggctggggga gggcaggctg
1081 gaagtcgcca ggattttcag cgggtcctgc gttcttttgg tgcgaagggt gggcagttat
1141 cctaccggcc gtttgtgac cgttcctttg atgtgacagg cgtgccgag gctgtgaaa
1201 ggcaactcga tcaggcggag catgacacgg cggttgagga gcaggtcact cccgcgccac
1261 aaatcgcggt cgcaccgcca ccgccgccag tcgttcctga cccgcccgcc atcgtgacgg
1321 aaaccgcgcc cccgccgctt gtcgtggtca gcgtccggt cacgtatga cccccgctg
1381 ccgccgtgoc ggcagagcct cccgttcagg aagccccgt gcaggcggcg ccggttccc
1441 ccgcgcctgt gcccccgatt cgggagcagg ctcctccgcg ggcgccggac ccggcatccg
1501 tgccgtatgc gaacgtcgcg gcagcaccog ttccacctga tcccgcaccg gttacgctg
1561 cgccgcaggc gcgcgtgacg gggccgaaca cccgtatggt ggagcccttt tcccgccgc
1621 aggtccgcac ggtgcaggag ggggcaaccc cgtcacgtgt accttcgctg tcaatgaacg
1681 ctttcccccg cacatcagca tcgtccataa gtgagcgtcc ggtggacagg ggtgtgccc
1741 atgaatggag tcctgttccg aaggcacgcc tcagcccgcg ggagcgtccg cgtcccggcg
1801 atctgagctt tttcttcag gggatgcgcg acaccgctga tgaaaagaag ttcttcccg
1861 tggcgtccac gcgatcagtt cgttctaata tttccaggat gaccagcatg accaagacag
1921 acacgaatc ctctcaggct tctcgtcccg gcagccccgt cgctcgcct gatgggtgc
1981 ccacaatggc cgaagtgttc atgacgctgg gtggtcgtgc gacggaactc ctacgcccc
2041 gtccttcgct gcgggaggcg ctggtgcgct gtcgtgaaa cgaagaaga tcctaaggcc
2101 ctatattca

DRA

DDBJ Sequence Read Archive

DDBJ Sequence Read Archive (DRA)

新世代シーケンサから出力される配列や
アライメントデータを登録・公開



今時 (いまどき) のシーケンサー



(左) PacBio RSII System

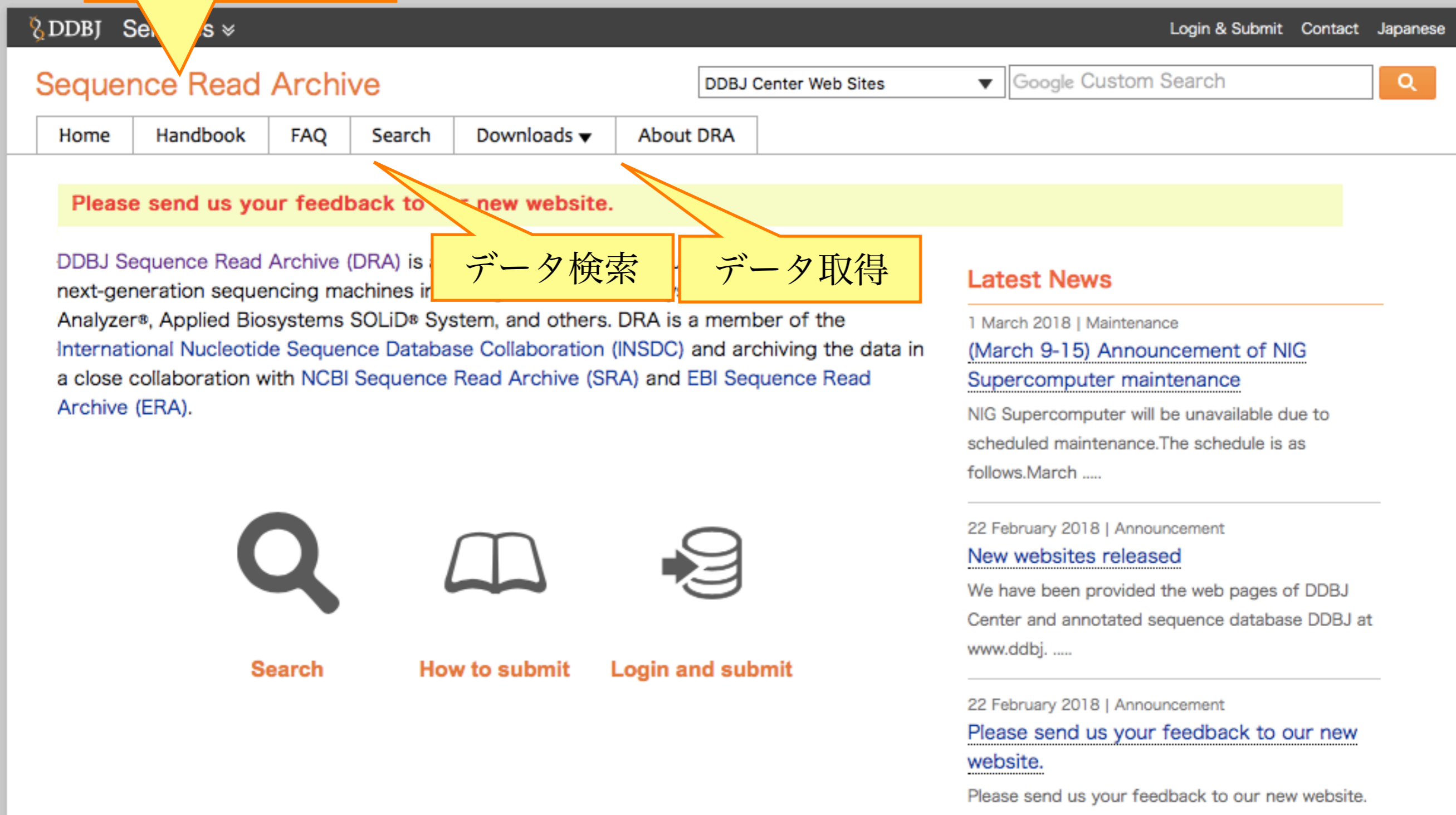
(中) illumina: HiSeq 4000 System

(右) Oxford NANOPORE MinION/SmidgION

DRAウェブサイト ⇒ [DRA] で検索

登録関係情報

<https://www.ddbj.nig.ac.jp/dra>



The screenshot shows the DDBJ Sequence Read Archive website. At the top, there is a navigation bar with 'DDBJ Services' and links for 'Login & Submit', 'Contact', and 'Japanese'. Below this is the 'Sequence Read Archive' title and a search bar with 'DDBJ Center Web Sites' and 'Google Custom Search'. A menu bar contains 'Home', 'Handbook', 'FAQ', 'Search', 'Downloads', and 'About DRA'. A yellow callout box points to the 'Search' menu item. Below the menu is a yellow banner with the text 'Please send us your feedback to our new website.' Two yellow callout boxes point to the words '検索' (search) and '取得' (acquisition) in the text below. The main text describes the DRA as a next-generation sequencing machine archive, mentioning technologies like Illumina and Applied Biosystems SOLiD. Below the text are three icons: a magnifying glass for 'Search', an open book for 'How to submit', and a database cylinder for 'Login and submit'. On the right side, there is a 'Latest News' section with two announcements: one about NIG Supercomputer maintenance and another about new websites released. A second yellow callout box points to the '検索' icon in the bottom navigation area.

DDBJ Sequence Read Archive (DRA) is a next-generation sequencing machines in Illumina® HiSeq™ X Analyzer®, Applied Biosystems SOLiD® System, and others. DRA is a member of the International Nucleotide Sequence Database Collaboration (INSDC) and archiving the data in a close collaboration with NCBI Sequence Read Archive (SRA) and EBI Sequence Read Archive (ERA).

検索 **データ取得**

Search **How to submit** **Login and submit**

Latest News

1 March 2018 | Maintenance
[\(March 9-15\) Announcement of NIG Supercomputer maintenance](#)
NIG Supercomputer will be unavailable due to scheduled maintenance. The schedule is as follows. March

22 February 2018 | Announcement
[New websites released](#)
We have been provided the web pages of DDBJ Center and annotated sequence database DDBJ at www.ddbj.

22 February 2018 | Announcement
[Please send us your feedback to our new website.](#)
Please send us your feedback to our new website.

公開データの DRA Search での検索

公開データは EBI SRA / NCBI SRA と共有されています

DRASearch Send Feedback [Search Home](#) [DRA Home](#)

Accession :
OR
Organism : StudyType :
CenterName : Platform :

Show records Sort by

Statistics

Released Entries

Type	Count
Submission	23770
Study	3423
Experiment	29624
Sample	111241
Run	71620

Organism

#	Organism Name	Study	#	Organism Name	Study
1	Homo sapiens	293	1	Whole Genome Sequencing	31
2	metagenome sequence	169	2	Transcriptome Analysis	24
3	Mus musculus	163	3	Metagenomics	14
4	Drosophila melanogaster	121	4	Epigenetics	14
5	marine metagenome	79	5	Other	14
6	Caenorhabditis elegans	39	6	Res	14
7	Arabidopsis thaliana	38	7	RNAseq	14
8	synthetic construct	37	8	Population Genomics	14
9	Saccharomyces cerevisiae	35	9	Gene Regulation Study	14
10	Panicum virgatum	21	10	Cancer Genomics	14

Show records Sort by

Search Results (358 studies) << < 1 / 18 Page > >>

#	STUDY	SUBMISSION	STUDY_TITLE	STUDY_TYPE	ORGANISM	CENTER_NAME
1	DRP000003	DRA000003	Comprehensive identification and characterization of the nucleosome structure	Transcriptome Analysis	Homo sapiens	UT-MGS
2	DRP000004	DRA000004	Comprehensive identification and characterization of the transcripts, their expression levels	Transcriptome Analysis	Homo sapiens	UT-MGS
3	DRP000005	DRA000005	Comprehensive characterization of the nucleosome structure			
4	DRP000006	DRA000006	Comprehensive characterization of the nucleosome structure			
5	DRP000007	DRA000007	Comprehensive characterization of the nucleosome structure			
6	DRP000008	DRA000008	Comprehensive characterization of the nucleosome structure			

DRP000003

Study Detail

Title	Comprehensive identification and characterization of the nucleosome structure
Abstract	Comprehensive identification and characterization of the nucleosome structure in mammalian genes were attempted. We used Nucleosome-Seq method, in which next gene sequencing technology and microcococcus nuclease digestion assay were combined.
Description	Although recent studies have revealed that the majority of human genes are subjected to regulation of alternative promoters (APs), the biological relevance of this phenomenon remains unclear. In order to understand biological significance of the presence of diverg .. [more]
Project ID	34559
Center Name	UT-MGS (University of Tokyo, Medical Genome Sciences)

Navigation

- Submission [DRA000003](#) [FTP](#)
- Experiment [DRX000003](#) [FASTQ](#) [SRALite](#)
- Sample [DRS000003](#)

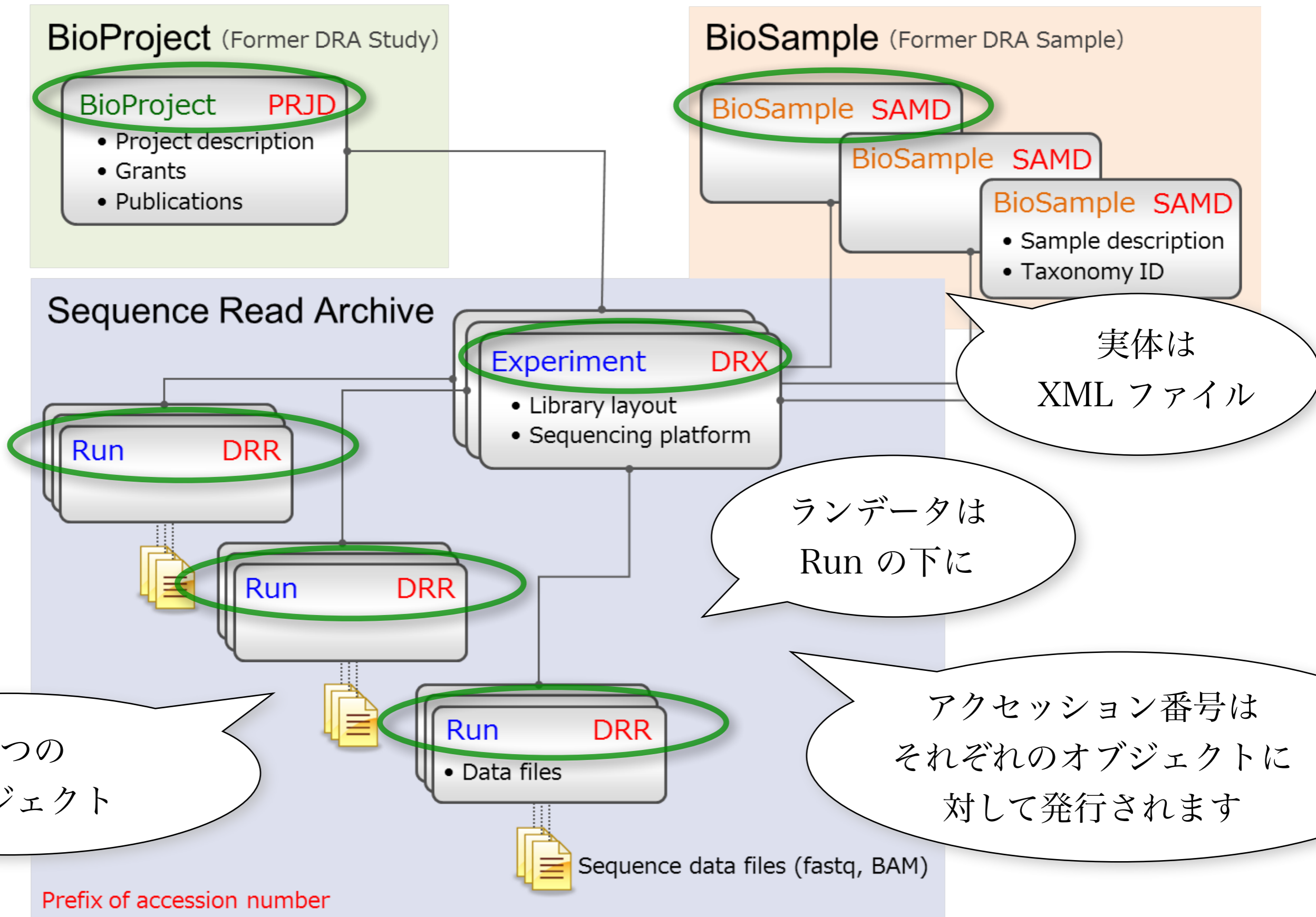
ダウンロード

詳細 (メタデータ記述)

生物名 etc での絞り込み

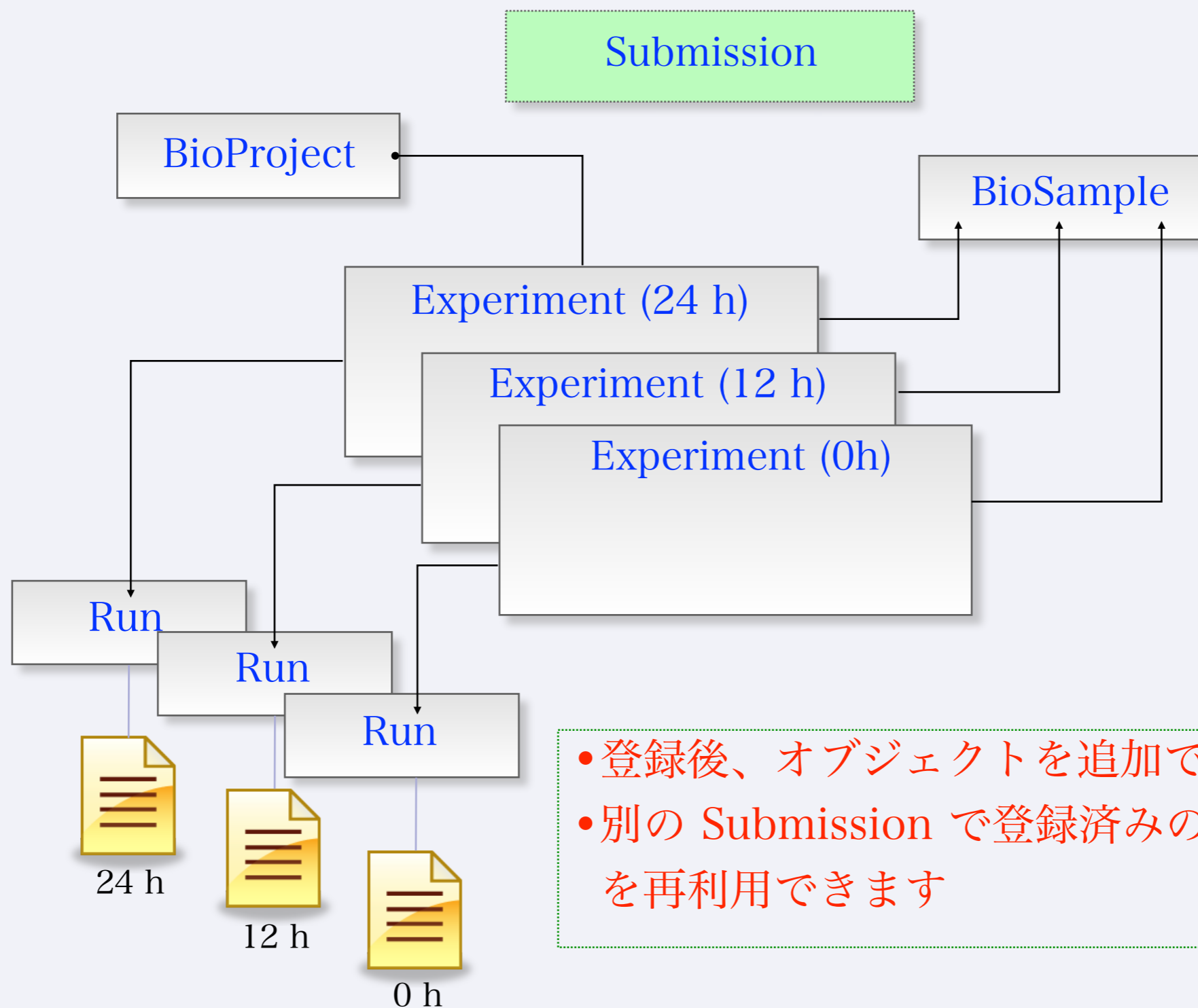
検索結果リスト

メタデータとランデータの構成



メタデータ構成の例

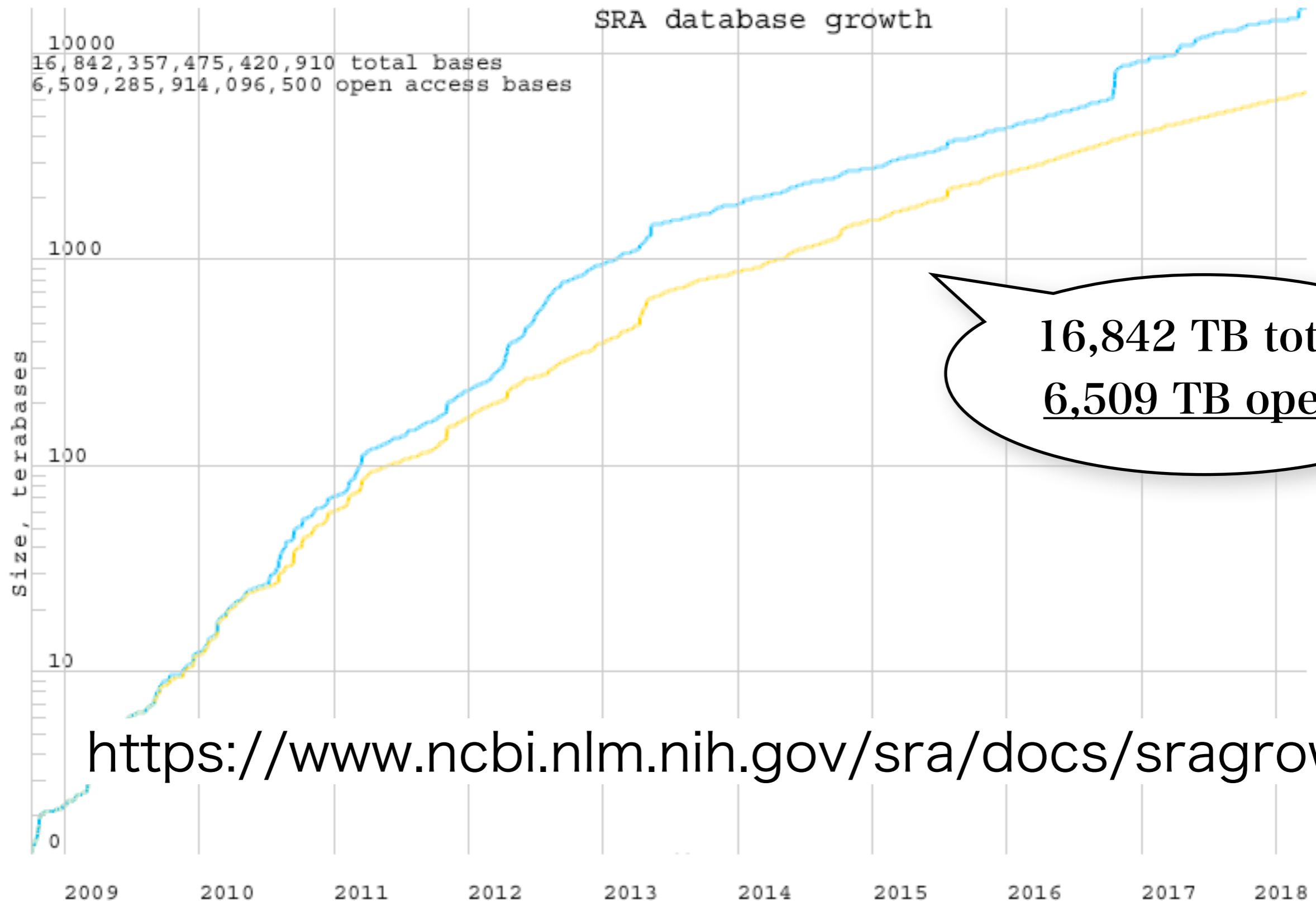
例) 培養細胞: 薬剤処理 0, 12, 24 h 後の転写プロファイル解析



- 登録後、オブジェクトを追加できます
- 別の Submission で登録済みのオブジェクトを再利用できます

SRA growth

NCBI SRA growth



<https://www.ncbi.nlm.nih.gov/sra/docs/sragrowth>

Total bases —
Open access bases —

JGA

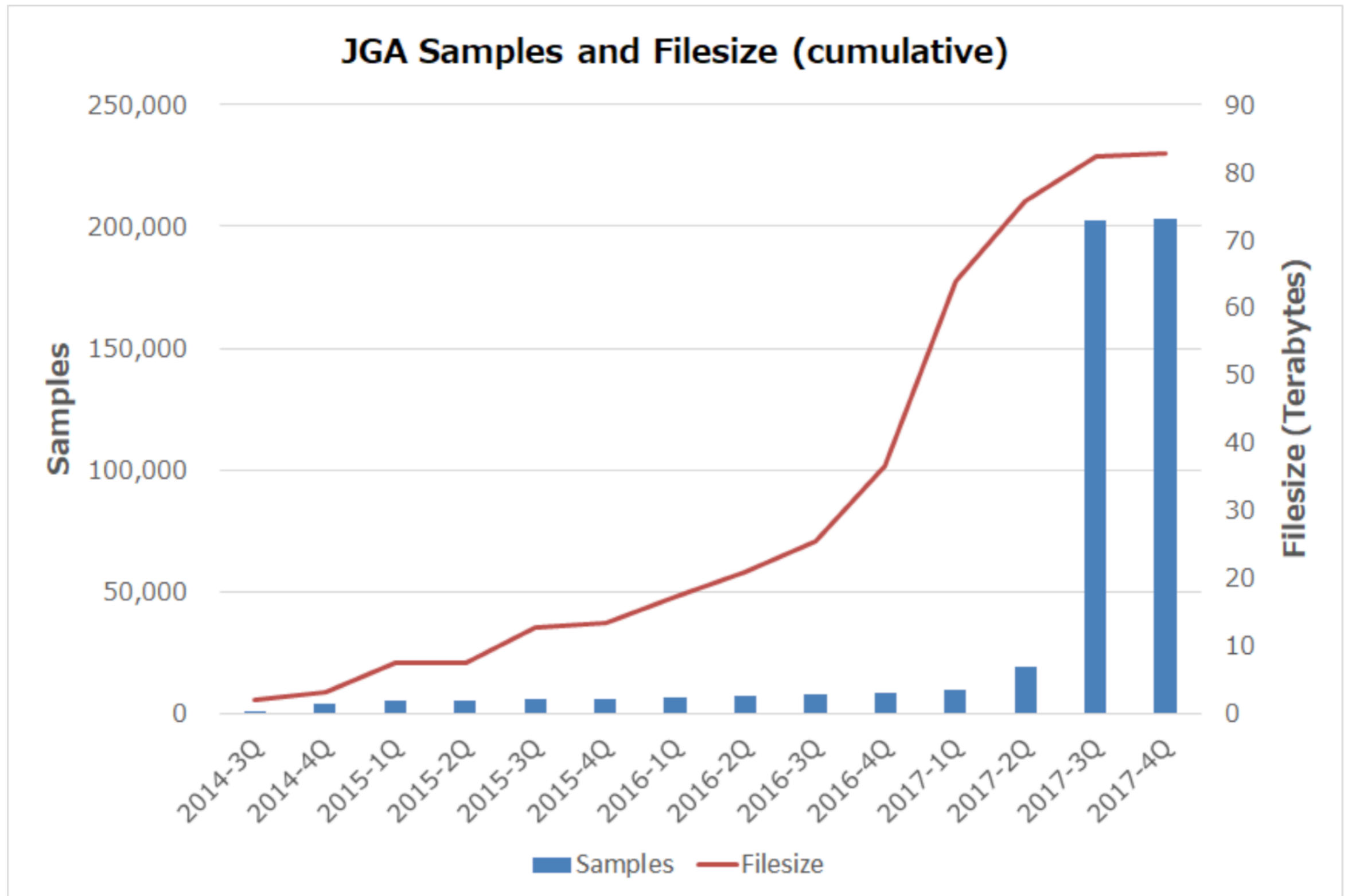
Japanese

Genotype-phenotype Archive

- 📍 Japanese Genotype-phenotype Archive (JGA)
 - 📍 個人の遺伝学的なデータと表現型情報を保存し、提供
 - 📍 匿名化されているメタデータのみを受け付け
 - 📍 Informed Consent により JGA データの利用は特定の研究目的に制限されている
 - 📍 **データの登録と利用は事前に NBDC に申請し、認可されている必要がある**
 - 📍 メタデータの体系は SRA を拡張したもの
 - 📍 NGS データに加えアレイや variation データを登録できる
 - 📍 ウェブサイト: <http://trace.ddbj.nig.ac.jp/jga/>
 - 📍 メールアドレス: jga@ddbj.nig.ac.jp

Japanese Genotype-phenotype Archive (JGA)

123 Study、203,345サンプル、83テラバイトのデータをアーカイブしており、内 72 Study を制限公開している。バイオバンクジャパンから20万サンプルの genotyping データが2017年9月に制限公開された。

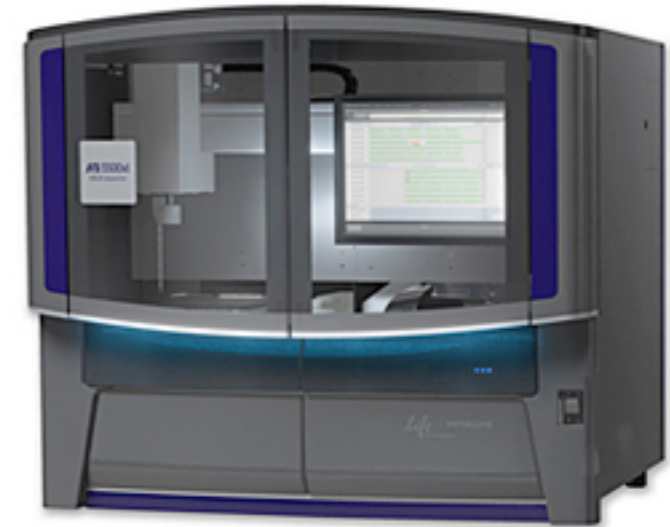


遺伝研スーパー

コンピュータ

生物学の 情報爆発

代表的 NGS 機材 (2010年頃のスライド)



(左) Roche (454): GS FLX+ System

(中) illumina: Genome Analyzer IIx System

(右) Life Technologies: 5500 xl SOLiD System

今時 (いまどき) のシーケンサー



(左) PacBio RSII System

(中) illumina: HiSeq 4000 System

(右) Oxford NANOPORE MinION/SmidgION

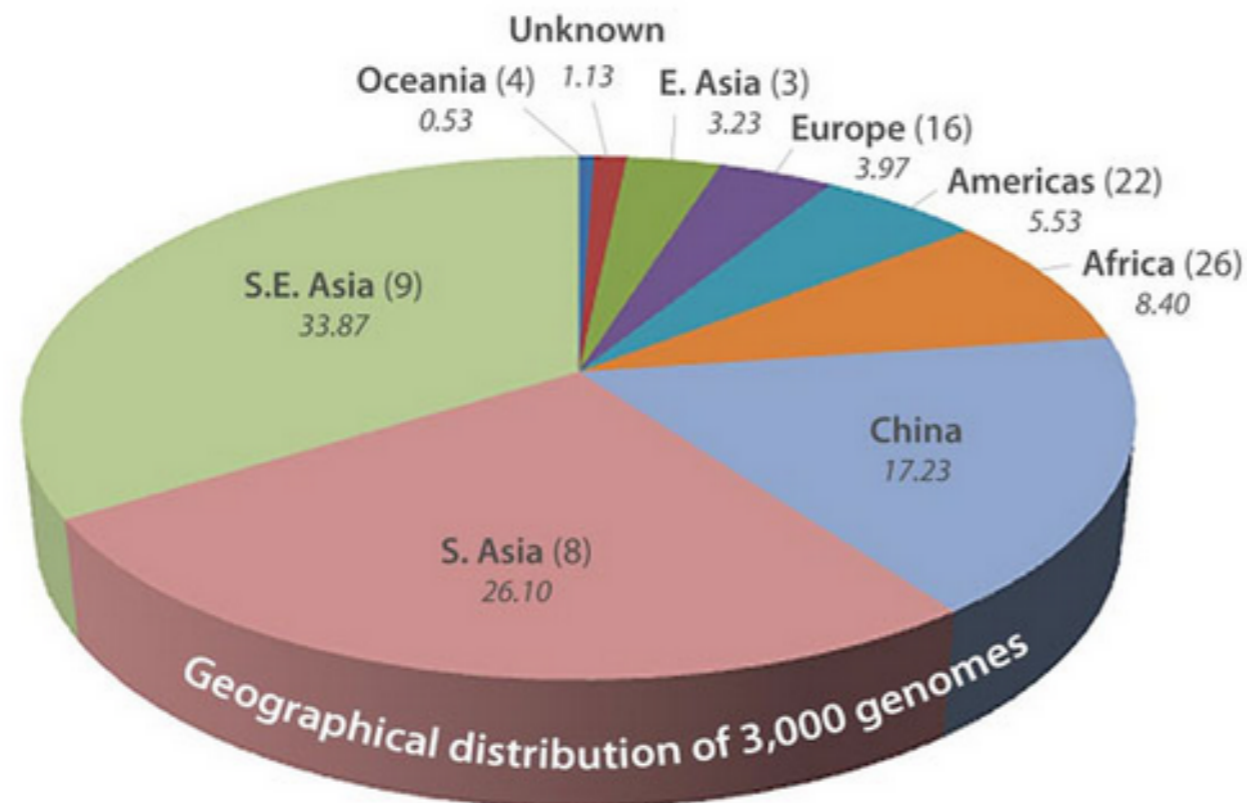
3k Rice Genomes Project

The 3,000 Rice Genomes Project

A single genome cannot reveal the large stockpile of genetic diversity in rice and hence many potentially important genes are not present in the handful of lines that have been sequenced over the last decade. So, to drastically change this dynamic, IRRI—in collaboration with BGI in Shenzhen, China, and the Chinese Academy of Agricultural Sciences (CAAS)—has completed the sequencing of 3,000 rice genomes of varieties and lines representing 89 countries (see figure) now housed in the IRGC (82%) and CAAS's genebank (18%). "This is an unparalleled development in plant science for a major food crop," says Ken McNally, senior scientist in the TTC GRC and a project team member.



[View the full issue](#)



Geographic distribution of 3,000 representative rice accessions whose genomes have been sequenced. Numbers in parentheses indicate the number of countries in each region.



UK10K

Rare Genetic Variants in Health and Disease

What is UK10K?

The UK10K project will enable researchers in the UK and beyond to better understand the link between low-frequency and rare genetic changes, and human disease caused by harmful changes to the proteins the body makes.

Although many hundreds of genes that are involved in causing disease have already been identified, it is believed that many more remain to be discovered. The UK10K project aims to help uncover them by studying the genetic code of 10,000 people in much finer detail than ever before.



Wellcome Library,
London

Project Design

Not all genetic changes are harmful or lead to disease, so the project is taking a two-pronged approach to identify rare variants and their effects:

- by studying and comparing the DNA of 4,000 people whose physical characteristics are well documented, the project aims to identify those changes that have no discernible effect and those that may be linked to a particular disease;
- by studying the changes within protein-coding areas of DNA that tell the body how to make proteins of 6,000 people with extreme health problems and comparing them with the first group, it is hoped to find only those changes in DNA that are responsible for the particular health problems observed.



The project received a £10.5 million funding award from Wellcome in March 2010 and sequencing started in late 2010. For more information, please use the links on the right hand side.

HOME
GOALS
FUNDING
INVESTIGATORS
CONSORTIUM MEMBERSHIP
STUDY SAMPLES
ETHICS
DATA ACCESS
DATA & METHODS
DALLIANCE GENOME BROWSER
POSTERS
PUBLICATIONS
CONTACT US

GGBN Global Genome Biodiversity Network

GGBN-GGI Awards Program

Mobilize your sample data!

Samples

DNA	148243
Tissues	287333
Enviros	18
Repositories	7

Vouchers

Cultures	24628
eVouchers	2
Specimens	134457
Unknown	6065
Collections	16

Taxa

Families	2819
Genera	14051
Species	38816

Total

	600746
--	--------

News

- 21 June 2017 - University of Kansas Biodiversity Institute joins GGBN
- 16 June 2017 - GGBN North American Regional Workshop Report now available
- 06 June 2017 - GGBN-GGI Awards Program, 2017 Awardees
- 22 May 2017 - GGBN ABS Fact Sheet and FAQ now available
- 02 May 2017 - GGBN Asian Regional Workshop (July 21-22), updates and RSVP deadline extended to May 31
- 26 April 2017 - Job Openings at Natural History Museum, United Kingdom
- 05 April 2017 - GGBN March 2017 newsletter now available
- 03 April 2017 - Biodiversity Research and Teaching Collections at Texas A&M University joins GGBN
- 27 March 2017 - Bishop Museum joins GGBN
- 02 March 2017 - GGBN/CETAF/SPNHC ABS Survey: Please participate by 23 March 2017
- 01 March 2017 - Denver Botanic Gardens joins GGBN
- 02 February 2017 - GGBN ABS Guidance now on Access and Benefit-Sharing Clearing-House

Explore GGBN

- Repositories
- Search
- Documents
- About

GGBN

www.ggbn.org

GGBN provides a platform for biodiversity biobanks from across the world to:

- Collaborate to ensure consistent quality standards for DNA and tissue collections,
- Improve best practices for the preservation and use of such collections and
- Harmonize exchange and use of material in accordance with national and international legislation and conventions.

DivSeek

HOME WHAT IS DIVSEEK? WORKING GROUPS NEWS RESOURCES CONTACT

What is DivSeek?

- MISSION AND GOALS
- PARTNERS
- COMMUNITY DRIVEN INITIATIVE
- HARNESSING CROP DIVERSITY
- FAQS

DivSeek

MISSION

The mission of DivSeek is to enable breeders and researchers to mobilize a vast range of plant genetic variation to accelerate the rate of crop improvement and furnish food and agricultural products to the growing human population.

GOALS

DivSeek is a community driven effort consisting of a diverse set of partner organizations that have voluntarily come together to unlock the potential of crop diversity so that it can be utilized to enhance the productivity, sustainability and resilience of crops and agricultural systems.

DivSeek will work with partners and other

DivSeek

www.divseek.org

The mission of DivSeek is to enable breeders and researchers to mobilize a vast range of plant genetic variation to accelerate the rate of crop improvement and furnish food and agricultural products to the growing human population.



GOLD

GENOMES ONLINE DATABASE

[JGI HOME](#) [LOG IN](#)

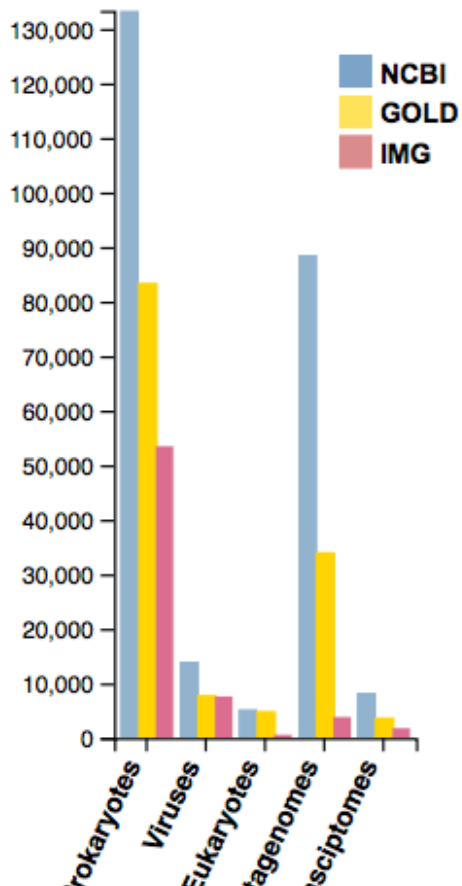
genome gold

[Home](#) [Search](#) [Distribution Graphs](#) [Biogeographical Metadata](#) [Statistics](#) [GOLD Usage Policy](#) [Team](#) [Help](#) [News](#)

Studies	32,276
Biosamples	45,952
Sequencing Projects	194,953
Analysis Projects	154,281
Organisms	297,233

↓ Excel Data file
File last generated: 21 Mar, 2018

NCBI Import Tracker



Welcome to the Genomes OnLine Database

GOLD:Genomes Online Database, is a World Wide Web resource for comprehensive access to information regarding genome and metagenome sequencing projects, and their associated metadata, around the world.

GOLD Release v.6

1. Register

Register your project information and Metadata in the Genomes Online Database

Register

2. Annotate

Annotate your microbial genome or metagenome with IMG/ER or IMG/MER

Annotate

3. Publish

Standards in Genomic Sciences

Publish your genome or metagenome in open access standards-supportive journal.

Publish

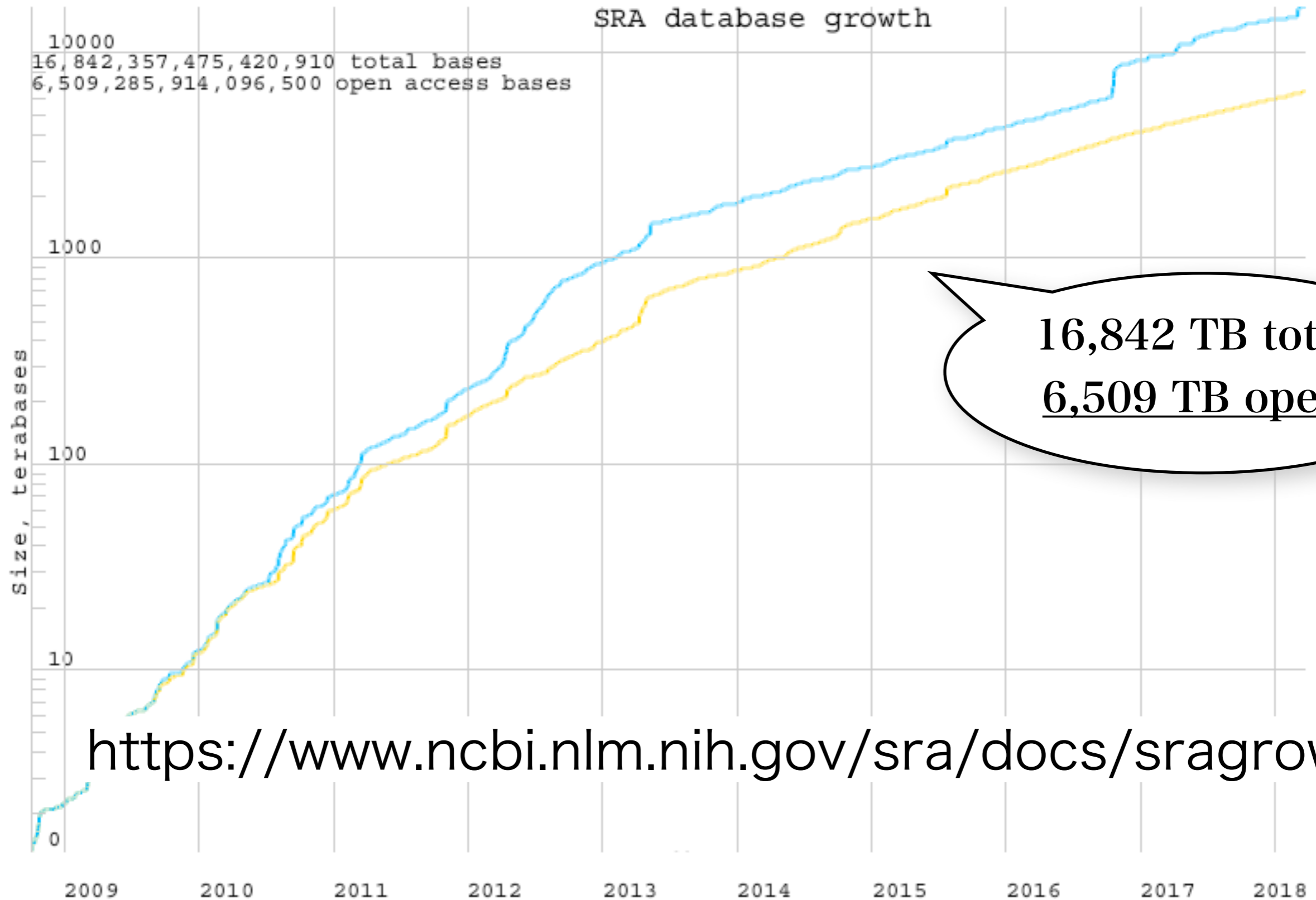
<p>Studies</p> <p>Metagenomic <u>1,431</u></p> <p>Non-Metagenomic <u>30,844</u></p>	<p>Biosamples</p> <p> <u>Classification</u></p> <p>Ecosystems</p> <p>Host-associated <u>21,668</u></p> <p>Engineered <u>4,358</u></p> <p>Environmental <u>19,924</u></p>	<p>Sequencing Projects</p> <p> Complete Projects <u>14,050</u></p> <p> Permanent Drafts <u>111,349</u></p> <p> Incomplete Projects <u>67,890</u></p> <p> Targeted Projects <u>1,295</u></p>	<p>Analysis Projects</p> <p>Genome Analysis <u>100,210</u></p> <p>Metagenome Analysis <u>29,281</u></p> <p>Metagenome - Cell Enrichment <u>935</u></p> <p>Metagenome - Single Particle Sort <u>2,757</u></p> <p>Metagenome - Assembled Genome (MAG) <u>5,301</u></p> <p>Metatranscriptome Analysis <u>3,686</u></p> <p>Combined Assembly <u>156</u></p> <p>Single Cell - Screened (SAG) <u>2,354</u></p> <p>Single Cell - Unscreened (SAG) <u>1,941</u></p> <p>Transcriptome Analysis <u>516</u></p>
--	---	--	---

<p>Special Projects</p> <p>Type Strain Projects <u>8,837</u></p> <p>Strains at Genbank <u>6,494</u></p>	<p>Seq. Projects</p> <p>Seq. Projects <u>95,165</u></p>	<p>JGI Studies</p> <p>JGI Studies <u>1,257</u></p> <p>JGI Biosamples <u>45,952</u></p>	<p>Organisms</p> <p>Organisms <u>297,217</u></p> <p>Archaea <u>2,493</u></p>
--	--	---	---

<https://gold.jgi.doe.gov>

SRA growth (再掲)

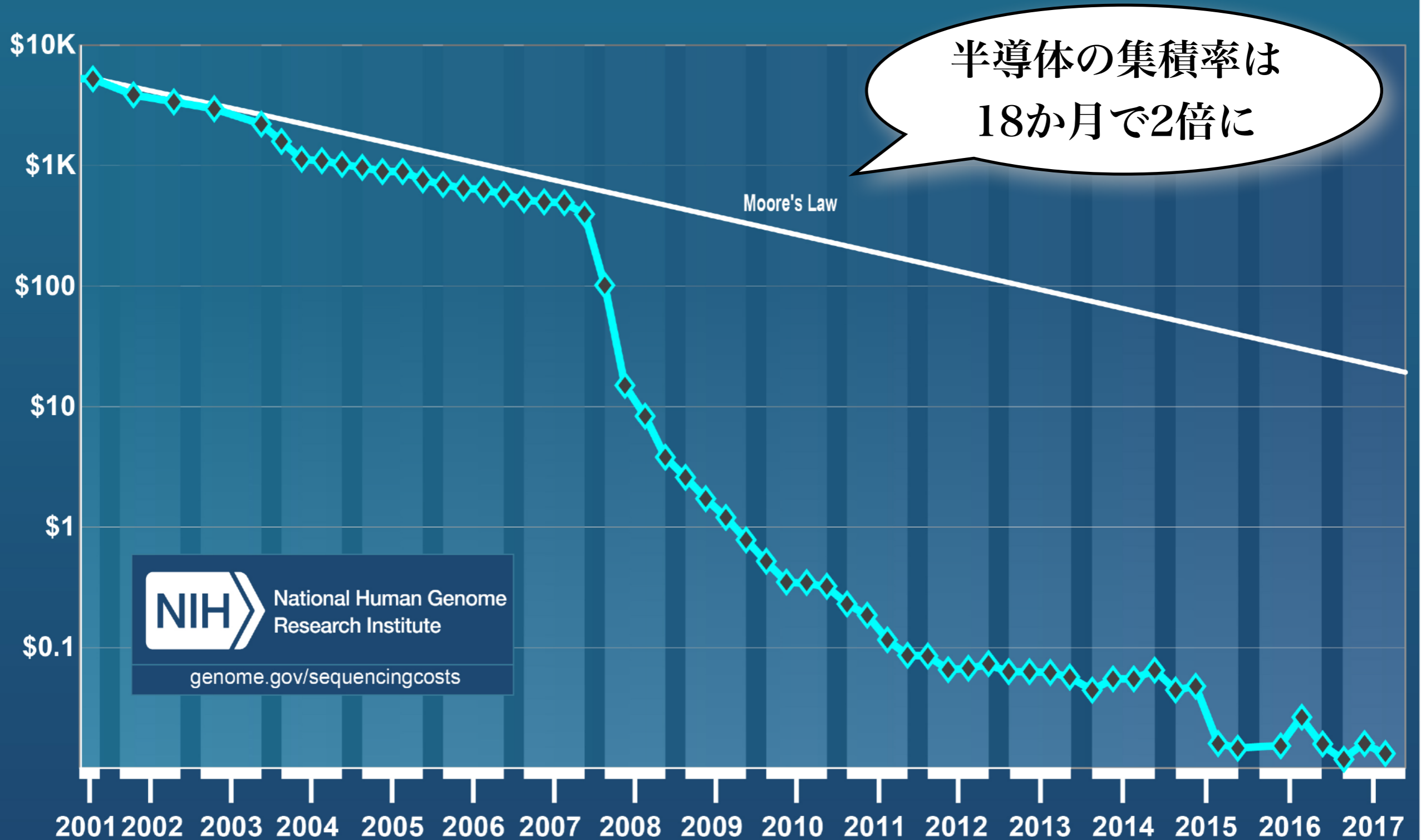
NCBI SRA growth



<https://www.ncbi.nlm.nih.gov/sra/docs/sragrowth>

Total bases —
Open access bases —

Cost per Raw Megabase of DNA Sequence



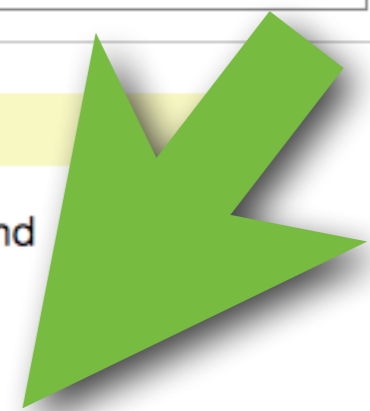
11 AUGUST 2011 WU ET AL. | 1507

FIG. 10. Same as in Fig. 9, but for the 2008 season. The 2008 season was characterized by a strong El Niño event, which led to a significant increase in precipitation over the tropical Pacific region. The map shows the spatial distribution of precipitation anomalies, with a prominent positive anomaly over the central and eastern equatorial Pacific. The color scale ranges from -10 to 10 mm day⁻¹.



Please send us your feedback to our new website.

DDBJ Center provides sharing and analysis services for data from life science researches and advances science.



Search & Analysis



Submissions



Downloads



SuperComputer



Statistics



Activities



Training



About Us



News from DDBJ Center

1 March 2018 | Maintenance | [DDBJ](#) [BioProject](#) [BioSample](#) [DRA](#) [JGA](#) [AGD](#) [DDBJ Center](#)
[\(March 9-15\) Announcement of NIG Supercomputer maintenance](#)

1 March 2018 | Announcement | [DDBJ](#)

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



サイトを検索

検索

現在のセクション内のみ

ホーム

システム利用案内

各種申請

稼働状況

成果報告

ホーム

ニュース

- 3月9日(金) - 3月15日(木)
遺伝研停電、スパコンメンテナ作業に伴うシステム停止

2018年03月05日

- 遺伝研スパコンホームページ移行について

2018年02月21日

- 2017年度 年度末更新・成果報告実施のご案内

2018年02月21日

- 2月22日(木) - 24日(土) スパコンストレージ増強工事に伴うスパコンシステム停止のお知らせ

2018年02月21日

遺伝研スーパーコンピュータシステム

国立遺伝学研究所は遺伝学を中心とした生命・医学系研究における情報処理のための大規模計算基盤として、大規模クラスタ型計算機・大規模メモリ共有型計算機・大容量高速ディスク装置を備えた最新鋭のスーパーコンピュータシステムサービスを提供しています。



新着情報

スーパーコンピュータ

- プロセッサ、メモリ、ストレージ、ネットワーク等のハードウェアと、その上で動作する OS やアプリケーションなどのソフトウェアから構成される点では一般的なコンピュータと同じ
- ただし、本来画像処理のために開発された Graphics Processing Unit (GPU) を汎用的な計算に利用する General Purpose computing on GPU ([GPGPU](#)) などのコプロセッサによって浮動小数点演算性能を稼ぐ構成も多い
- 遺伝研スパコンは CPU 重視構成（でも GPU, [Xeon Phi](#) も一部に入ってますよ）

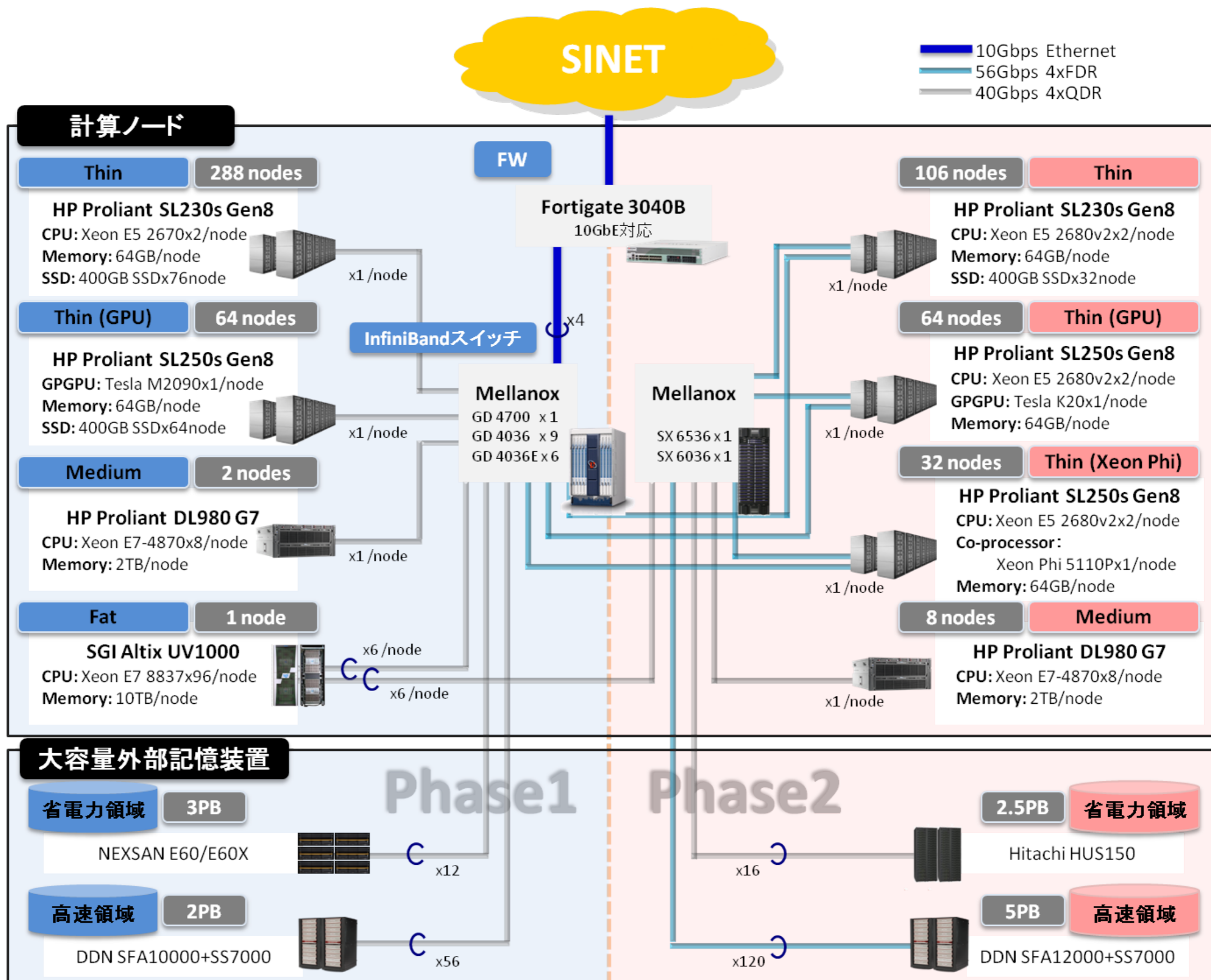
ニュース

- 3月9日(金)-3月15日(木) 遺伝研停電、スパコンメンテナンス作業に伴うシステム停止
2018年03月05日
- 遺伝研スパコンホームページ移行について
2018年02月21日
- 2017年度 年度末更新・成果報告実施のご案内
2018年02月21日
- 2月22日(木)-24日(土) スパコンストレージ増強工事に伴うスパコンシステム停止のお知らせ
2018年02月21日
- 2018年～2019年の運転予定
2018年02月21日

システム利用案内

- [ハードウェア](#)
- [ソフトウェア](#)
- [データベース](#)
- [システム利用方法のご案内](#)
- [FAQ](#)
- [サイトポリシー](#)

遺伝研スーパーコンピュータ全容



12.5 PB

> 300 TFlops



7 PB
Lustre
high-speed
HDD

5.5 PB
MAID
energy-saving
HDD



“thin”
64GB memory
x 554 nodes

“medium”
2TB memory
x 10

“fat”
10TB
memory
(SGI UV)

ニュース

- 3月9日(金)-3月15日(木) 遺伝研停電、スパコンメンテナンス作業に伴うシステム停止
2018年03月05日
- 遺伝研スパコンホームページ移行について
2018年02月21日
- 2017年度 年度末更新・成果報告実施のご案内
2018年02月21日
- 2月22日(木)-24日(土) スパコンストレージ増強工事に伴うスパコンシステム停止のお知らせ
2018年02月21日
- 2018年～2019年の運転予定
2018年02月21日

ソフトウェア

[ソフトウェア一般](#)

OS, ジョブ管理システム(Univa Grid Engine), ファイル転送ソフトウェア, プログラミング環境, MPI, 科学技術計算ライブラリ, プロファイラ, Java, シェルおよびスクリプト言語について

[プログラミング環境の詳細](#)

コンパイラ (Intel, PGI, GCC)、MPI、科学技術計算用ライブラリ、プロファイラ、Java、スクリプト言語について

[バイオインフォマティクス関連ツールの説明](#)

[ソフトウェアインストール履歴](#)

現在インストールされているソフトウェアのバージョンとインストールした日付の表。解析ソフトウェアなどは/usr/localディレクトリの下に複数のバージョンがインストールされているのでPATH環境変数などを適宜変更してご利用ください。

UNIX Operating System のはじまり



Bell 研の Ken Thompson, Dennis Ritchieらが “Space Travel” で遊ぶために（の目的だけってわけでもないけど）高価で買ってもらえなかった DEC-10 のかわりに部屋のスミに転がっていた借り物の PDP-7 で「えいやっ」と作ってみた「小さくて」「軽い」オペレーティングシステム (1968)



Life with UNIX (アスキー出版局; 1990) より引用/改変



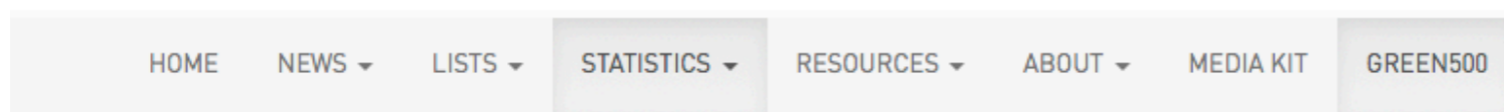
iOS の基礎部分は Darwin であり、NeXTSTEP を先祖に持つ由緒正しき Berkeley Software Distribution (BSD) UNIX の系譜に連なります。

要するに組み込み系の UNIX (POSIX 準拠) で動作。

MacOS X も勿論 UNIX。

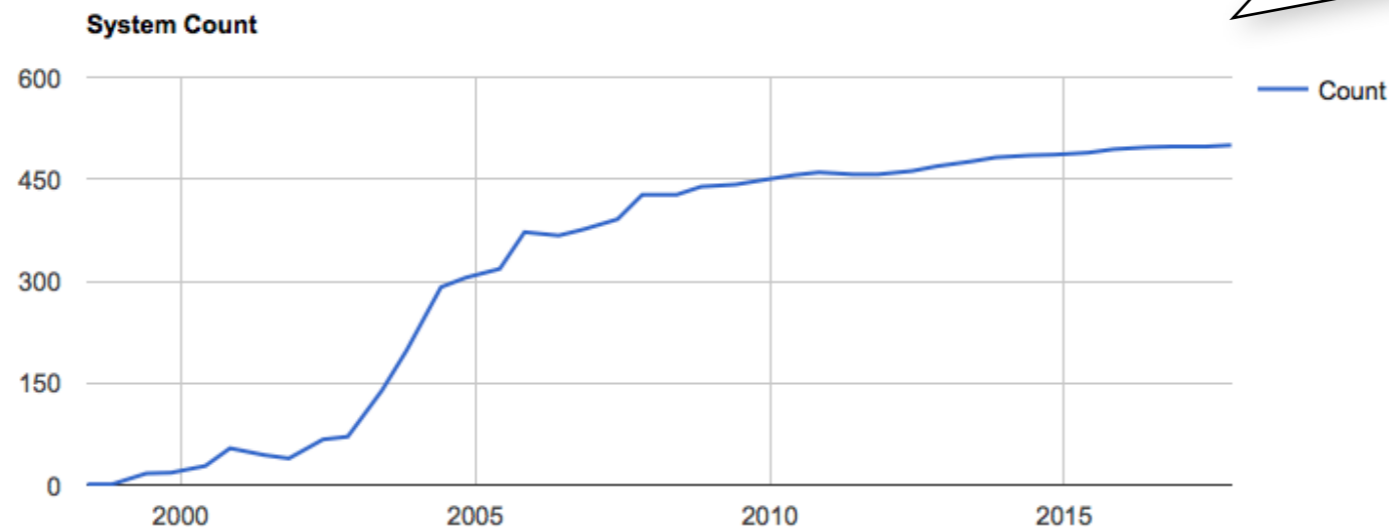
スパコンの OS だって Unix 系 Linux

- 遺伝研スパコン⇒Red Hat Enterprise Linux 6



Home / Lists / Statistics / Operating system Family / Linux

OPERATING SYSTEM FAMILY / LINUX



**Top 500 で
使われている
Linux のカウント**

List	Count	System Share (%)	Rmax (GFlops)	Rpeak (GFlops)	Cores
11月 2017	500	100	845,120,505	1,339,337,622	69,004,640
6月 2017	498	99.6	747,830,702	1,133,580,060	48,082,382
11月 2016	498	99.6	671,242,777	1,014,225,024	43,957,506

スパコン利用申請はこちら

- [遺伝研 スーパーコンピュータ] で検索

基本、無料です！



ホーム システム利用案内 **各種申請** 稼働状況 成果報告

ホーム

ニュース

- 3月9日(金) - 3月15日(木) 遺伝研停電、スパコンメンテナンス作業に伴うシステム停止
2018年03月05日
- 遺伝研スパコンホームページ移行について
2018年02月21日
- 2017年度 年度末更新・成果報告実施のご案内
2018年02月21日
- 2月22日(木) - 24日(土) スパコンストレージ増強工事に伴うスパコンシステム停止のお知らせ
2018年02月21日

遺伝研スーパーコンピュータシステム

国立遺伝学研究所は遺伝学を中心とした生命・医学系研究における情報処理のための大規模計算基盤として、大規模クラスタ型計算機・大規模メモリ共有型計算機・大容量高速ディスク装置を備えた最新鋭のスーパーコンピュータシステムサービスを提供しています。



新着情報

課金サービスもあります

ニュース

- 3月9日(金) - 3月15日(木) 遺伝研停電、スパコンメンテナンス作業に伴うシステム停止
2018年03月05日
- 遺伝研スパコンホームページ移行について
2018年02月21日
- 2017年度 年度末更新・成果報告実施のご案内
2018年02月21日
- 2月22日(木) - 24日(土) スパコンストレージ増強工事に伴うスパコンシステム停止のお知らせ
2018年02月21日
- 2018年～2019年の運転予定
2018年02月21日

各種申請

ログインアカウントの申請

国立遺伝学研究所は大学共同利用機関の一員として遺伝学を中心としたライフサイエンス系の研究・教育のためのリソース提供を目的としてスーパーコンピュータシステムの運営を行っています。ユーザーは遺伝研スパコンを基本的には無料で利用できることを原則としていますが、事業継続性、計算資源の効率的利用の観点から特に大規模データの保存、付加的なサービスの利用に対して課金によるサービスを行っています。

また、スーパーコンピュータシステムは外国為替及び外国貿易法(外為法)の輸出管理規制の対象となっており、兵器製造目的などで使用すると処罰の対象となります。このような背景から、遺伝研スパコンでは以下の方々を対象にログインユーザーアカウントの発行を行っております。

- 外為法の定める国内居住者でありかつ大学あるいは国公立研究機関の教員
- 1.の共同研究者、あるいは1.の指導の下にある研究者、学生、委託企業の研究者など(留学生、海外への転任、海外の研究者を含む)

利用目的・成果については原則として公開となっております。ただし公開時期の延期を希望される場合はその旨承ります。

遺伝研スーパーコンピュータ ログインアカウントを希望される方は「利用規定」を御一読の上、下記リンクからアカウントを申請してください。

- [利用規定](#)
 - ログインユーザーアカウント発行基準、システム利用規定
- [新規ユーザー登録申請](#)
 - ユーザー登録は随時受け付けています。新規ユーザー登録を行うと高速ストレージ上のホーム領域1TBが利用可能となります。
- [ssh公開鍵の登録](#)
 - スパコンにログインするためにssh公開鍵を登録していただきます。
- [利用可能リソース拡張申請](#)

- [ログインアカウントの申請](#)
- [課金サービス利用申請](#)
- [Webアプリケーションのアカウント申請](#)
- [その他お問い合わせ](#)

「課金サービスの利用方法」をご覧ください



Language/言語
●

ホーム

このサイトへのログイン

ユーザ名

パスワード

システム構成

- ハードウェア構成
- ソフトウェア構成
- プログラミング環境
- 利用可能バイオツール
- 利用可能OSS
- 利用可能DB

システム使用方法

- 基本的利用方法
- その他UGE利用方法
- ファイル転送方法
- システム利用TIPS
- 稼働スケジュール
- アドバンスドリザベリション コマンド利用方法
- アドバンスドリザベリションサービスの開始に伴うジョブ投入時の注意事項
- UGE設定変更に伴う注意事項(2017.07.27~)

課金サービスの利用方法

遺伝研スパコンは大学共同利用機関の事業として運営されておりユーザーは遺伝研スパコンを基本的には無料で利用できることを原則としていますが、事業継続性、計算資源の効率的利用の促進の観点から特に下記の大規模データの保存、付加的なサービスの利用に対して課金によるサービスを開始いたします。

1. ストレージの大規模利用
2. 計算ノードの優先利用

課金サービス利用の際は、[利用規定](#)をご確認いただき、下記の説明に従い、利用の申込みをしてください。利用目的を審査の上、課金サービス利用許可の案内をお送りします。遺伝研スパコンの混雑状況によってはご利用できない場合がありますのでご了承ください。

1. ストレージの大規模利用

ストレージの大規模利用に関しては以下の課金サービスがあります。30TBを超えるストレージの利用を希望する場合は、[利用申請書](#)及び下記の利用計画表をダウンロードし、必要事項を記入の上、sc-info@nig.ac.jp までメールしてください。

高速ストレージ (Quota設定のみ: 容量保証なし) 利用計画表	30TBまで無料 30TBを超えた場合10TBあたり年間3万円 (40TB: 年間12万円、以降10TBあたり年間3万円)	
高速ストレージ (容量保証) 利用計画表	10TBあたり年間20万円	
アーカイブストレージ (容量保証) 利用計画表	10TBあたり年間11.7万円	
DDBJグループクラウド アクセス制限データグループ共有 データベースサービス(*)	基本料金	1プロジェクトあたり年額1,672万円
	アーカイブストレージ (容量保証)	10TBあたり年間11.7万円
	要望対応開発	1人月140.4万円

金額は税込価格です。

(*) JGAに準じたデータ構造を持つデータをクラウド上にアップロードし、データアクセスが許可された研究者間で共有するサービス。初期導入費用、セキュリティ監視、運用の人的費用を含む。異なるデータ構造への対応など追加開発が必要な場合については基本料金の他に別途開発費が必要。状況の変化により次年度以降の価格が変更になる場合があります。

1. クォータ設定については、最大利用可能容量の設定が行われます。指定された容量を超えてファイルを保存することはできません。また、遺伝研スパコンの混雑状況によっては、最大利用可能容量に達する以前に書き込みができなくなる場合やファイルの消去をお願いする場合がありますのでご了承ください。容量保証サービスに関しては、最大利用可能容量までの書き込みが保証されます。
2. クォータ設定、容量保証サービスともに、利用開始日から年度末までの間の容量設定の指定を記した年間利用計画書を所定の書式で提出していただきます。クォータ設定に関しては月単位、容量保証サービスは半年単位で容量の変更が可能です。変更の際には年間利用計画書を再提出していただきます。年間利用計画書の提出は随時受け付けます。ただし2月1日以降は変更不可とします。
3. DDBJグループクラウドサービスでは、アーカイブストレージ容量保証サービスは半年単位での容量の変更が可能です。基本料金については利用日数に関わらず一年分の料金となります。
4. 課金対象となるストレージ利用量は遺伝研スパコンユーザーアカウントの責任者単位での合計で算出され、責任者に対して支払いの請求を

各種申請

- システムの利用規程
- ログインユーザアカウント発行基準
- 課金サービスの利用方法
- 料金のお支払い
- 各種申請窓口について

1. ストレージの大規模利用

1. ストレージの大規模利用

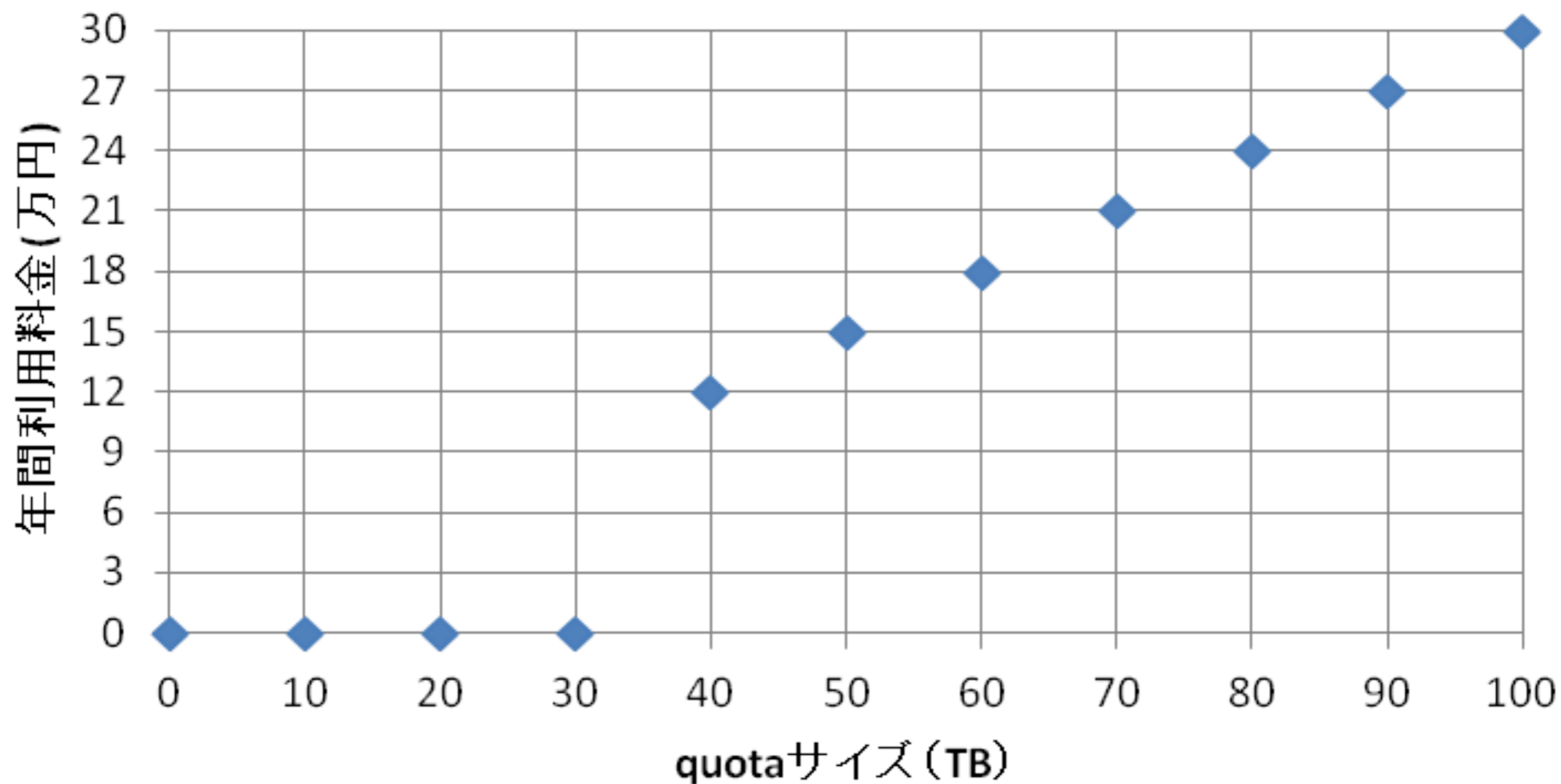
ストレージの大規模利用に関しては以下の課金サービスがあります。30TBを超えるストレージの利用を希望する場合は、[利用申請書](#)及び[料金計画表](#)をダウンロードし、必要事項を記入の上、sc-info@nig.ac.jp までメールしてください。

高速ストレージ (Quota設定のみ：容量保証なし) 利用計画表		30TBまで無料 30TBを超えた場合10TBあたり年間3万円 (40TB：年間12万円、以降10TBあたり年間3万円)
高速ストレージ (容量保証) 利用計画表		10TBあたり年間20万円
アーカイブストレージ (容量保証) 利用計画表		10TBあたり年間11.7万円
DDBJグループクラウド アクセス制限データグループ共有 データベースサービス(*)	基本料金	1プロジェクトあたり年額1,672万円
	アーカイブストレージ (容量保証)	10TBあたり年間11.7万円
	要望対応開発	1人月140.4万円

金額は税込価格です。

100 TB 確保すると年間 30 万円

高速ストレージ (quota設定)



2. 計算ノードの優先利用

2. 計算ノードの優先利用

計算ノードに対するUniva Grid Engine (UGE)を用いたジョブの投入に関してはこれまで通り無料で利用可能です。UGEのアドバンスドリ機能を利用することで計算ノードの優先利用が可能です。料金は下表のとおりです。本機能の利用を希望する場合は、[利用申請書](#)及び下表をダウンロードし、必要事項を記入の上、sc-info@nig.ac.jp までメールしてください。

Phase.1 Thinノード (64GB memory) 利用計画表	CPUコアまたはメモリのいずれか使用量の大きい方で 1CPUコアまたは4GBメモリを1単位とし、 1単位1分あたり0.05円
Phase.2 Thinノード (64GB memory) 利用計画表	CPUコアまたはメモリのいずれか使用量の大きい方で 1CPUコアまたは3GBメモリを1単位とし、 1単位1分あたり0.05円
Phase.2 Mediumノード (2TB memory) 利用計画表	CPUコアまたはメモリのいずれか使用量の大きい方で 1CPUコアまたは25GBメモリを1単位とし、 1単位1分あたり0.25円
Fatノード (10TB memory) 利用計画表	CPUコアまたはメモリのいずれか使用量の大きい方で 1CPUコアまたは13GBメモリを1単位とし、 1単位1分あたり0.13円

金額は税込価格です。

スパコン利用申請はこちら

- [遺伝研 スーパーコンピュータ] で検索

基本、無料です！



ホーム システム利用案内 **各種申請** 稼働状況 成果報告

ホーム

ニュース

- 3月9日(金) - 3月15日(木) 遺伝研停電、スパコンメンテナンス作業に伴うシステム停止
2018年03月05日
- 遺伝研スパコンホームページ移行について
2018年02月21日
- 2017年度 年度末更新・成果報告実施のご案内
2018年02月21日
- 2月22日(木) - 24日(土) スパコンストレージ増強工事に伴うスパコンシステム停止のお知らせ
2018年02月21日

遺伝研スーパーコンピュータシステム

国立遺伝学研究所は遺伝学を中心とした生命・医学系研究における情報処理のための大規模計算基盤として、大規模クラスタ型計算機・大規模メモリ共有型計算機・大容量高速ディスク装置を備えた最新鋭のスーパーコンピュータシステムサービスを提供しています。



新着情報

<https://p.ddbj.nig.ac.jp/> (ddbj pipeline で検索)



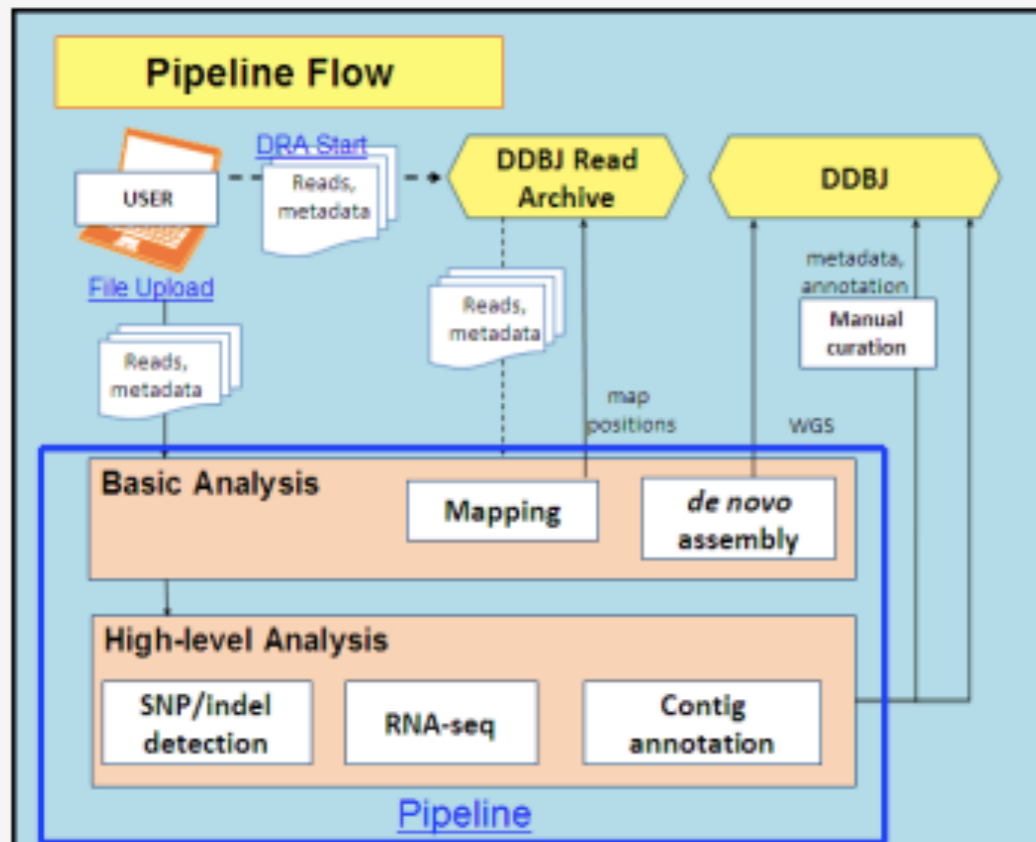
DDBJ Read Annotation Pipeline

[English](#) [Japanese](#)

DDBJ Read Annotation Pipeline is a cloud-computing based analytical platform for next-generation sequencing data.

LOGIN

[New account](#) [Login as "guest"](#)



User ID:

Password:

[Login](#)

[Check current jobs](#)
* by the guest account.

Manual & tutorial

- [Japanese Tutorial \(FAQ\)](#)
- [English manual](#)
- [DBCLS togotv Tutorial video 1 \(JP\) - Reference Genome Mapping](#)
- [DBCLS togotv Tutorial video 2 \(JP\) - De novo Assembly](#)
- [Tutorial : How to upload and register query files to DDBJ Pipeline \(JP\)](#)
- [Tutorial : How to run HGAP for PacBio sequence read on DDBJ Pipeline \(JP\)](#)

Data submission for analyzed results and sequenced

DDBJ pipeline: ソフトウェア

よく用いられる
解析用ソフトウェアを
用意。クリックだけで
実行可能

DDBJ
DNA Data Bank of Japan

ACCOUNT
login ID [guest]
Logout

ANALYSIS
Data setup
DRA Start
FTP upload
HTTP upload
DRA Import
Preprocessing Start

step-1
Preprocessing
Mapping /
de novo Assembly

step-2
Workflow
Genome (SNP/Short
Indel)
RNA-seq (Tag count)
ChIP-seq

JOB STATUS
step1.
Preprocessing
step1.
Mapping
step1.
de novo Assembly
step2-All status

HELP
HELP
TUTORIAL
Contact Us.
DDBJ Read Annotation
Pipeline.
Development Team.

Select Query Files → Select Tools → Set QuerySet → Set GenomeSet → Set Mapping

Running Status

Selecting Tools for Basic Analysis of DDBJ ANNOTATION

BACK NEXT

Reference Genome Mapping

	Tool	Help	Version	Input data			Evaluation			Analysis		Output format			Comment
				Base space	Color space	Paired end	Depth	Coverage	Error rate	SNP	Indel	.gff	.bed	SAM	
<input type="checkbox"/>	BLAT		34	✓						✓					Single-end analysis only
<input type="checkbox"/>	bwa		0.6.1	✓		✓	✓	✓	✓					✓	
<input type="checkbox"/>	Bowtie		0.12.7	✓	✓	✓	✓	✓	✓	✓				✓	
<input type="checkbox"/>	TopHat		1.0.11	✓		✓	✓	✓	✓					✓	
<input type="checkbox"/>	Bowtie2		2.2.6	✓	✓	✓	✓	✓	✓	✓				✓	For reads longer than about 50 bp, Bowtie2 is generally faster, more sensitive, and uses less memory than Bowtie1.
<input type="checkbox"/>	TopHat2		2.1.0	✓		✓	✓	✓	✓					✓	

de novo Assembly Total limit = 22 Gbp

	Tool	Help	Version	Base space	Color space	Paired-end	MSS(WGS)	Comment
<input type="checkbox"/>	SOAPdenovo		2.04-r240	✓		✓		
<input type="checkbox"/>	ABySS		1.3.2	✓		✓		Maximum K-mer value is 64.
<input type="checkbox"/>	Velvet		1.2.10	✓		✓	✓	We severe recommend when performing Velvet, total length of those reads is up to 22G bp. Maximum K-mer value is 64.
<input type="checkbox"/>	Trinity		2.1.1	✓		✓		RNA-Seq De novo Assembly
<input type="checkbox"/>	Platanus		1.2.2	✓		✓		
<input type="checkbox"/>	HGAP		Protocol3 (v 2.2.0)					HGAP Pipeline for PacBio Sequence based on SMRT Analysis v2.2.0. For bax.h5 file only. (Beta version)

DDBJ pipeline: 比較対象

イネ、マウスなど
解析比較対象となる
配列を多数用意

The screenshot displays the DDBJ pipeline interface. At the top, a workflow diagram shows the steps: Select Query Files → Select Tools → Set QuerySet → Running Status. The main content area is titled 'Specifying Database of Reference' and is divided into three sections for different organisms:

- Major genome sets (Arabidopsis thaliana):** Organisms: Arabidopsis thaliana. Genome sets: TAIR8, TAIR9, TAIR10. File selection: all.fa, chr01.fa, chr02.fa, chr03.fa, chr04.fa, chr05.fa, chrC.fa, chrM.fa.
- Major genome sets (Oryza sativa japonica):** Organisms: Oryza sativa japonica. Genome sets: IRGSP Releases Build 4.0, IRGSP Releases Build 5.0, IRGSP Releases Build 5.0 masked by RepeatMasker with tigr version5.0, tigr version6.0, tigr version6.1, tigr mitochondrion, tigr chloroplast. File selection: all.fa, chr01.fa, chr02.fa, chr03.fa.
- Major genome sets (Homo sapiens):** Organisms: Homo sapiens. Genome sets: Homo sapiens Feb. 2009 (hg19), Mar.2006 (hg18), May.2004 (hg17), NCBI build 36.1_CRA, NCBI build 36.1_Celera, NCBI build 36.1_ref, NCBI build 36.2_CRA, NCBI build 36.2_Celera, NCBI build 36.2_ref, NCBI build 36.3_CRA, NCBI build 36.3_Celera, NCBI build 36.3_ref, NCBI build 36.3_HuRef, NCBI build 37.1_CRA, NCBI build 37.1_Celera, NCBI build 37.1_GRCh, NCBI build 37.1_HuRef. File selection: all.fa, chr1.fa, chr10.fa, chr11.fa, chr12.fa, chr13.fa, chr14.fa, chr15.fa, chr16.fa, chr17.fa, chr17_ctg5_hap1.fa.

On the left, a dropdown menu for 'Arabidopsis thaliana' is open, showing a list of organisms with 'Arabidopsis thaliana' selected. The list includes: Arabidopsis thaliana, Oryza sativa japonica, Oryza sativa indica, Zea mays B73, Sorghum bicolor, Homo sapiens, Mus musculus, Pan troglodytes, Caenorhabditis elegans, Xenopus (Silurana) tropicalis, Oryzias latipes, Solanum lycopersicum Heintz 1706, and Saccharomyces cerevisiae. A red arrow points to this dropdown menu.

パイプライン利用申請はこちら

<http://p.ddbj.nig.ac.jp> [DDBJ pipeline] で検索



DDBJ Read Annotation Pipeline

English

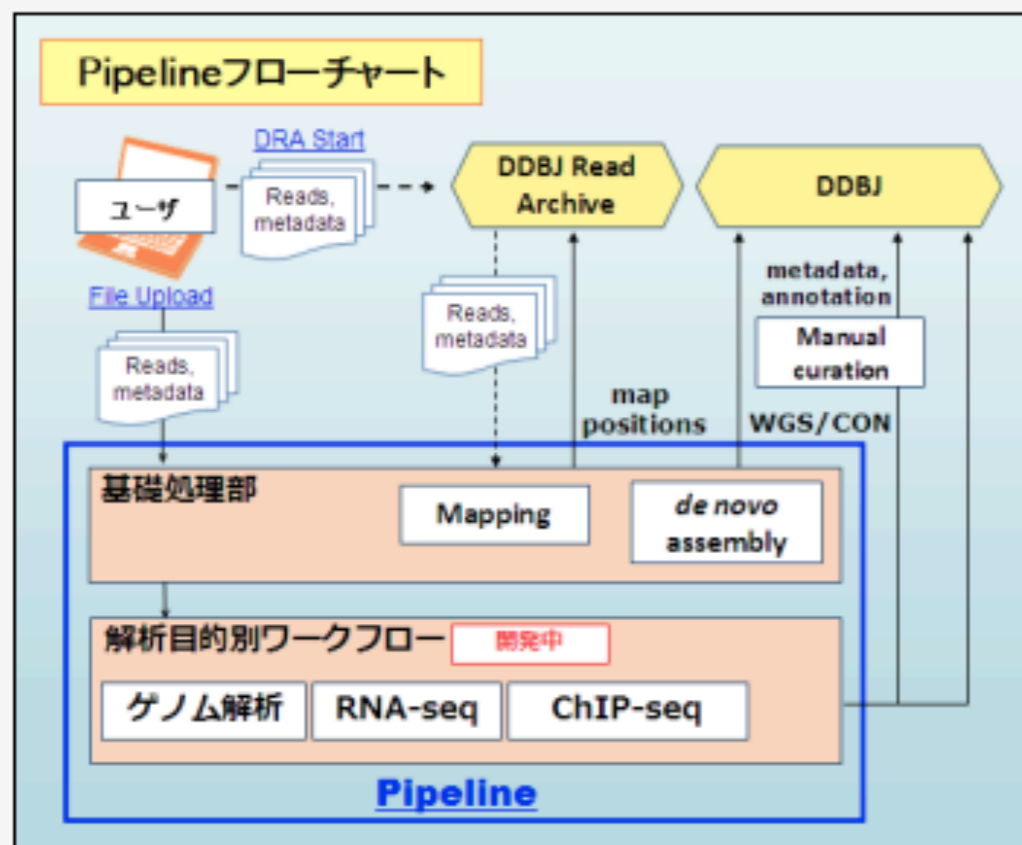
Japanese

DDBJ Read Annotation Pipelineは、次世代シーケンサ配列のクラウド型データ解析プラットフォームです。

LOGIN

新規アカウント作成

ゲストとしてログイン



User ID:

Password:

Login

動作中JOBの確認

PipelineのIDをお持ちでない場合、[ゲストとしてログイン](#)できます。

マニュアルおよびチュートリアル

- [日本語チュートリアル](#)
- [英語マニュアル](#)
- [DBCLS 統合TV チュートリアル1 - 今日からはじめるDDBJ Read Annotation Pipeline](#)
- [DBCLS 統合TV チュートリアル2 - DDBJ Read Annotation Pipelineによるde novo Assembly解析](#)

Tweets

Follow @pipeline_info



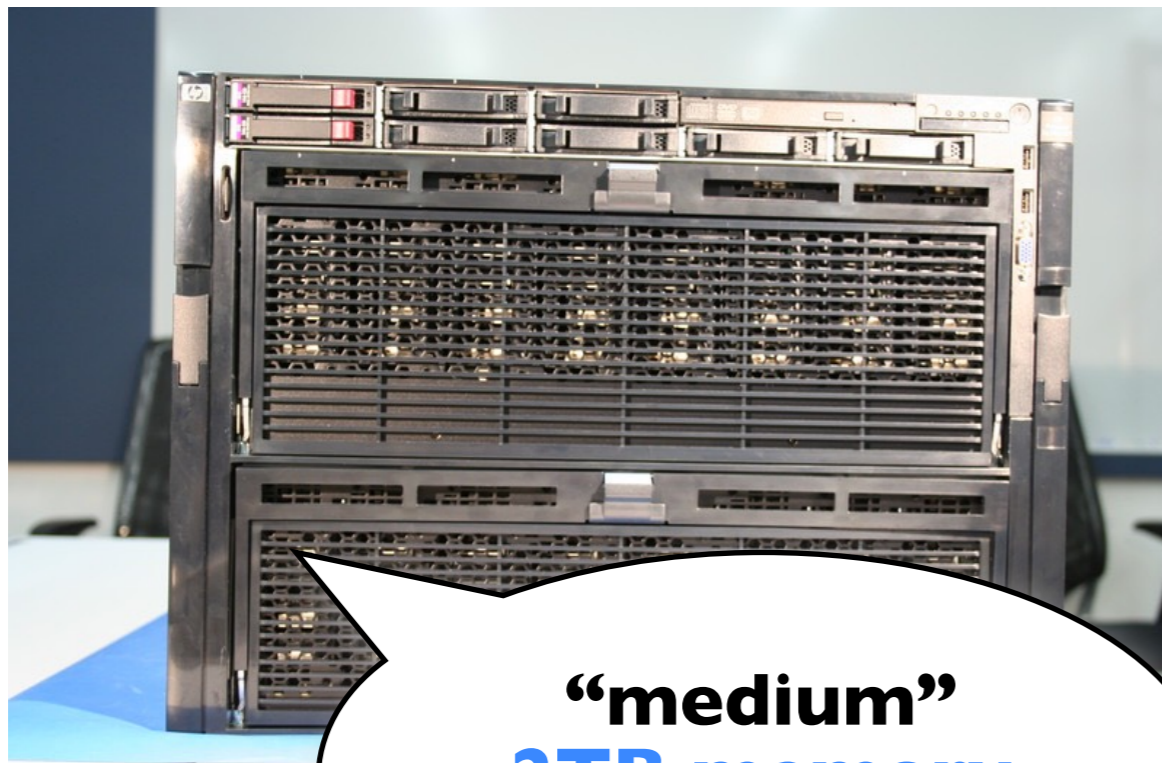
pipeline

29 Jul

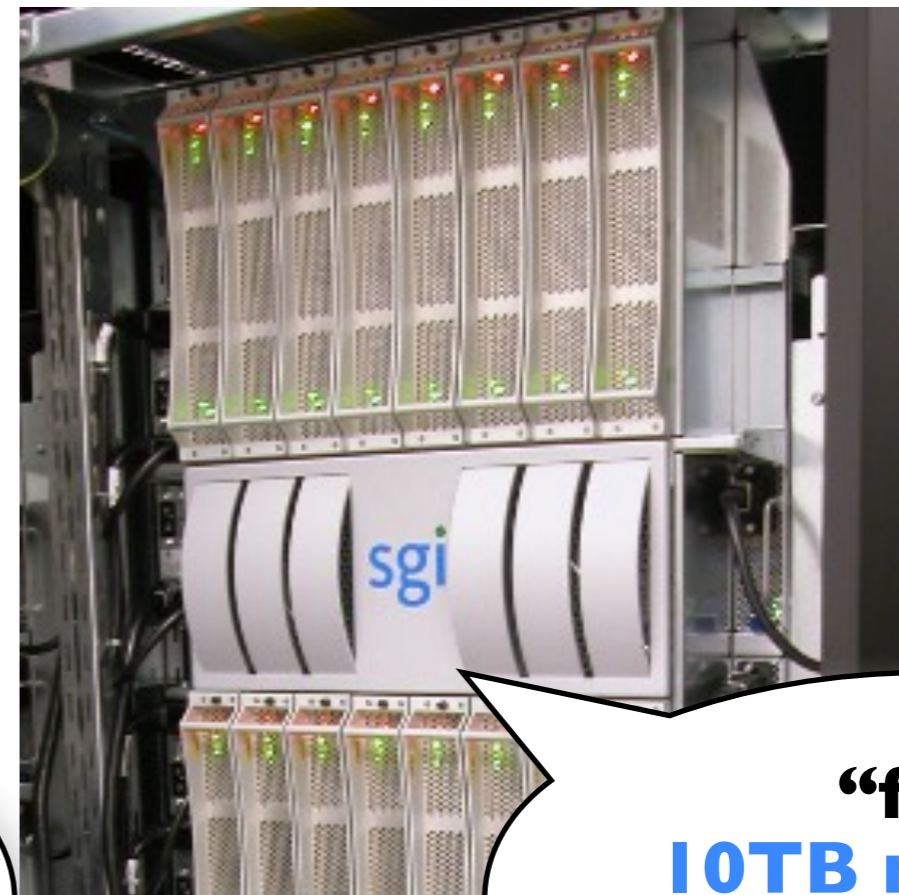
スパコンの
利用申請後に

無料です！

NGS's + SC's in Biology



“medium”
2TB memory
x 10



“fat”
10TB memory
(SGI UV)

DDBJ センター

DDBJ Center Web Sites

Google カスタム検索



新しいウェブサイトへのアンケート調査にご協力をお願いいたします

DDBJ センターは生命科学研究から生み出されるデータの共有・解析サービスを提供することで広く研究活動をサポートしています

検索・解析



登録



ダウンロード



スパコン



統計



活動



講習会



センターについて



DDBJ センターからのお知らせ

2018年3月1日 | メンテナンス | [DDBJ](#) [BioProject](#) [BioSample](#) [DRA](#) [JGA](#) [AGD](#) [DDBJ Center](#)
[\(3/9-15\) スパコン定期メンテナンスに伴う DDBJ センターのサービス停止のお知らせ](#)

2018年3月1日 | お知らせ | [DDBJ](#)



講習会

DDBJ についての講習会

DDBJing 講習会

DDBJを有効活用していただくための講習会。

on-demand DDBJing 講習会

みなさまのご要望に応じて開催する講習会。

NBDC, DBCLS, PDBj, DDBJ による合同講習会

バイオサイエンスデータベースセンター (NBDC)、ライフサイエンス統合データベースセンター (DBCLS)、日本蛋白質構造データバンク (PDBj)、日本 DNA データバンク (DDBJ) の4機関が合同で行う講習会。

All-in-one 合同講習会

生命科学研究のためのデータベースや、それらを構築・管理・運営する機関についての理解深め、活用していただくための講習会。

生命データアナリスト養成プログラム

本プログラムは、ROIS 機構長裁量プログラム (データサイエンス高度人材育成プログラム) への取り組みの一環として、ビッグデータの解析、主として次世代ゲノムシーケンス (NGS)解析需要に対応できる人材を国内で育成することを目的としています。

D-STEP 講習会

DDBJ-Supercomputer Training & Educational Program.
NGS データの様々な配列解析手法やビッグデータ解析 (機械学習、AIなど) についてハンズオンを主とした講習会。

DDBJ データ解析チャレンジキックオフ講習会

DDBJ データ解析チャレンジの参加者向けの講習会。

- 📍 DDBJ は INSDC の一員として塩基配列情報の DB を構築
 - 📍 Trad DDBJ はアノテーション付のDB (GenBank 相当)
 - 📍 DRA は NGS の生もしくはアラインメントDB (SRA相当)
 - 📍 JGA は制限アクセス付きのヒト遺伝型&表現型DB
- 📍 大量の塩基配列情報を効率的に扱うためのスパコンを提供
 - 📍 基本、無料です！
- 📍 スパコンの操作に慣れていない方のためにパイプラインを提供
 - 📍 無料です！
- 📍 ウェブサイト: www.ddbj.nig.ac.jp からアクセスしてください
- 📍 DDBJ とスパコンに特化した講習会も定期的に行っています