

バイオインフォマティクス人材育成カリキュラム (次世代シークエンサ) 速習コース

2. 配列インフォマティクス

2-2 バイオ系データベース概論

小野 浩雅



大学共同利用機関法人 情報・システム研究機構
ライフサイエンス統合データベースセンター

2014年9月5日(金)

小野 浩雅

<https://sites.google.com/a/dbcls.rois.ac.jp/hono/>

- ふりがな: おの ひろまさ
- 所属: ライフサイエンス統合データベースセンター
- twitter: h_ono 
- もともと実験系
 - 脂肪細胞とか使った分化・脱分化@マイクロアレイ解析
 - DBCLS RA第一号 兼 RAから職員第一号
 - 統合TV番組作成人 → 編集人
 - 発現データリファレンス RefEx 開発
 - 統合データベース講習会AJACS 講師 (遺伝子発現DB)

講習・実習に際してのお願い

- 資料を見ながら自力で進められそうな方はどんどん先に、そうでない方は講師と一緒にすすめていきましょう。
- サイトの反応が悪い時は一呼吸置いてから実行してみましょう
- 反応が無いからと言って何度もクリックするとますます繋がらなくなってしまいます。おおらかな気持ちで臨みましょう。
- わからないことがあったら挙手にてお知らせください。講師の説明が足りないせいかもしれません。



基本的な各種バイオ系データベースの理解 統合DBの利用法

- 困ったときは、**統合TV**を見る
- 最低限知つておくべきバイオ系DB
 - NCBI
 - UCSC Genome Browser
 - Ensembl, Biomart
 - IGV
- そもそも、**統合DB**とはなにか
- 統合DBのサービスを活用する

分からぬときの合言葉は ま、とにかく、まずは「ググれ」

<http://www.google.co.jp/>

分からぬ単語
エラーメッセージ
謎のID
etc..

その疑問は、過去に多くの人が
感じた疑問かもしれない（いや、たぶんそう）

とは言っても、モノ知りな人に聞きたい

Hiromasa ONOさんはTwitterを使っています 『テキスト中にマッチする単語の「個数」をgrepとwcで数えたいですがどうしたらよいですか。』というご質問。会場では回答出ませんでしたが、どなたか分かりますか?? #AJACS

興味があるものを検索する

Hiromasa ONO
@h_ono

『テキスト中にマッチする単語の「個数」をgrepとwcで数えたいですがどうしたらよいですか。』というご質問。会場では回答出ませんでしたが、どなたか分かりますか??

#AJACS

返信 リツイート お気に入り 登録 その他

リツイート お気に入り
1 2

22:20 - 2014年9月1日

崖っぷちのボニョ(蚊帳の外) @bonohu · 9月2日
@h_ono 単語って?

返信 リツイート お気に入り 登録 その他

Hiromasa ONO @h_ono · 9月2日
. @bonohu 今回のケースは塩基配列のデータから、任意の「ATGGCCT」みたいな文字列を取りたい場合です。

返信 リツイート お気に入り 登録 その他

崖っぷちのボニョ(蚊帳の外) @bonohu · 24 時間
@h_ono ググるなかなかw

返信 リツイート お気に入り 登録 その他

崖っぷちのボニョ(蚊帳の外) @bonohu · 24 時間
@h_ono EMBOSSTのdreg使いますね、UNIXコマンドであれば

返信 リツイート お気に入り 登録 その他

げんきいただききゅうけつき @n0rr · 24 時間
@h_ono @bonohu sedとtrは使用可ですか?

返信 リツイート お気に入り 登録 その他

中村保一 博士 (猫) @yaskaz · 24 時間
@h_ono @bonohu なるほど。普通の英文+ワードなら「はてな」の最後の解答が正解っぽいけど、塩基配列みたいに区切られてなくてもカウントしたいってことですね #AJACS q.hatena.ne.jp/1135146114

返信 リツイート お気に入り 登録 その他

中村保一 博士 (猫) @yaskaz · 24 時間
@bonohu @h_ono ですね。わたしならせっかくなのでfuzznucをご紹介するところです

統合TV

<http://togotv.dbcls.jp>

The screenshot shows the homepage of the Togo TV website. At the top, there's a navigation bar with links for 'はじめての方へ', '番組ランキング', 'ほかの便利な方法', 'よくある質問', 'スタッフ', and 'お問い合わせ'. Below the navigation is a search bar with a placeholder '検索窓にキーワードを入れると、入力のたびごとに即座に候補の番組が絞り込まれます。先頭のタイトル行をクリックすると、昇順・降順で並び替えができます。' and a dropdown menu set to '10 エントリを表示'. The main content area features two video thumbnails. The first thumbnail is for 'Wolfram Alpha を高機能関数電卓として使う その3：微分・積分・方程式 編' (140711版), with a description about Wolfram Alpha being a question-answer system developed by Wolfram Research. The second thumbnail is for 'Integrative Genomics Viewer IGVを使い倒す ~マッピングデータを可視化する~' (140711版). On the left side, there's a sidebar with a 'DBCLS Database Center for Life Science' logo and a '目的別に検索！' section containing links for various bioinformatics topics like 'ゲノム・核酸配列解析', 'タンパク質配列・構造解析', '発現制御解析', '文献検索・辞書 情報収集 PC環境構築', 'DBCLSサービス 講演・講習動画', and 'データベース別 分類'. At the bottom left, there are Creative Commons Attribution (CC BY) license icons.

生命科学分野の有用なDB
やツールの使い方を 動画
で手取り足取り解説!!

最低限知っておくべき
バイオ系DBを網羅!!
さらに 中級者・上級者
向けの番組も!!

困ったら、 統合TV !!

データベース別分類 in 統合TV

<http://togotv.dbcls.jp/ja/contents/category/database/>

The screenshot shows a web browser window for 'togotv.dbcls.jp/ja/contents/category/database'. The header features the 'TOGO TV CURATED' logo and the text '生命科学系DB・ツール使い倒し系チャンネル'. Below the header is a navigation bar with links: はじめての方へ, 番組ランキング, ほかの便利な方法, よくある質問, スタッフ, お問い合わせ. The main content area has a title 'データベース別分類' and a sidebar with links for NCBI, Ensembl/BioMart, and UCSC. A large red box highlights the statistics: 'NCBI 26 本', 'Ensembl, BioMart 16 本', and 'UCSC 10 本'. Below these are tables for each category.

データベース別分類

NCBI 26 本
Ensembl, BioMart 16 本
UCSC 10 本

番組タイトル	概要	作成日
NCBI bookshelfの使い方 2013	医学、生物学に関する教科書を検索・閲覧する	2013.5.29
MMDBを使ってタンパク質の構造を知る	タンパク質やRNAの3次元構造のデータベースMMDBの使い方	2012.11.05
NCBI GEOの使い方5～GEO2Rを使う～	マイクロアレイ実験データを、データ解析環境Rをベースにして解析するGEO2Rの解説	2012.5.24
PubChemを利用して化学物質やアッセイの結果を調べる	BioAssayを利用したアッセイの結果検索、およびCompoundを利用した化学物質の情報取得	2012.3.30
NCBI GEOの使い方4～データセットブラウザの使い方2～2012	データ解析ツールを用いた目的遺伝子の検索。クラスタリング結果の閲覧、統計解析、発現分布の見方	2012.2.27
NCBI GEOの使い方2～遺伝子プロファイルの検索・処理済みデータの取得～	興味のある遺伝子のプロファイルを検索および実験データの正規化されたファイルの取得方法	2011.10.20
NCBI GEOの使い方2～遺伝子プロファイルの検索・処理済みデータの取得～	興味のある遺伝子のプロファイルを検索および実験データの正規化されたファイルの取得方法	2011.10.20
NCBI BioProjectの使い方	ゲノムに関する様々なプロジェクトを調べる	2011.8.5
NCBI GEOの使い方1～マイクロアレイデータの検索・取得～2011	目的のマイクロアレイ実験データセットへのアクセス方法および生データ取得方法	2011.7.11
PSI-BLASTで類縁の配列を調べ倒す 2011	PSI-BLASTで類縁のアミノ酸配列を検索する	2011.6.7
NCBI Gene Expression Omnibus(GEO)にマイクロアレイデータを登録(Submit)する 2011	自分のマイクロアレイデータをGEOに登録する方法について	2011.4.22
ESTデータベース Entrez Unigeneを使い倒す 2011	EST配列の取得	2011.3.2
遺伝子のRefSeq IDを調べる 2011	目的遺伝子のRefSeq ID取得	2011.2.18
Entrez Books検索の使い方 2010	医学、生物学に関する教科書の中身を含めた検索と閲覧方法	

どんな番組がよく見られているか

<http://togotv.dbcls.jp/ja/rankings/total>

累計再生数一覧

togotv.dbcls.jp/ja/rankings/total

TOGO TV CURATED 生命科学系DB・ツール使い倒し系チャンネル

はじめての方へ 番組ランキング ほかの便利な方法 よくある質問 スタッフ お問い合わせ

使い倒し系チャンネル
統合TV

旧 統合TVはこちらから!

DBCLS Database Center for Life Science

目的別に検索！

- ゲノム・核酸配列解析
- タンパク質配列・構造解析
- 発現制御解析
- 文献検索・辞書 情報収集 PC環境構築
- DBCLSサービス 講演・講習動画
- データベース別 分類

累計再生数一覧

順位	番組名	累計再生数
1	Copyright in the Digital Age and Its Impact on Scientific Data Sharing (日本語字幕付き)	37995
2	Open Refineの使い方（旧Google Refine）の使い方～導入・基本編～	29661
3	ImageJを利用して画像を処理・解析する	29578
4	PCRプライマー設計ツール Primer3の使い方 2011	23627
5	NCBI BLASTの使い方～基本編～ 2010	19175
6	Primer3でPCR用のプライマーを設計する。	19126
7	MEGAを使って配列アライメントおよび系統解析をする	17338
8	ライフサイエンス新着論文レビュー FirstAuthor's を使い倒す	16276
9	PCRプライマー設計ツール Primer3の使い方	16138
10	Primer BLASTの使い方	15544
11	Local BLAST の使い方～検索実行・オプション編～ 2011	14038
12	GSEA softwareの使い方～基本編～	13672
13	CINIIを使って日本語論文を検索する	11757
14	パワーポイントの図形描画機能でイラストをつくる方法 その2	11689
15	遺伝子発現バンク(GEO)目次を使い倒す～その壹	11530
16	統合TVの作り方 (Windows版)	10768
17	Local BLAST の使い方～導入・準備編 (MacOSX版)～ 2011	10589
18	ImageJを利用して画像を処理・解析する-2	9891
19	生命科学分野のデータベースを統合する仕事：落ちこぼれ大学生がDB（Doctor of the database）にいたるまで	9544
20	PyMOLを使い倒す	9425
21	Local BLAST の使い方～導入・準備編～ 2011	9136
22	新着論文レビューをテキストマイニング2.0：破	

©2014 DBCLS Licenced Under CC 表示 2.1

統合TV YouTube 支店

<https://www.youtube.com/user/togotv>

The screenshot shows the YouTube channel page for 'togotv'. The channel banner features a cartoon character in a lab coat standing behind a desk with a computer monitor displaying 'TOGO TV IN YouTube'. Below the banner, the channel name 'togotv' is displayed, along with a 'チャンネル登録' button and a subscriber count of 641. The main content area shows a video thumbnail for '生命科学分野の有用なDBやツールの使いかたを動画で手取り足取り解説します' (A video guide for using useful databases and tools in the life sciences field). The video has been viewed 0:30 of 0:59. To the right of the video, there is a description box for '統合TV プロモーションムービー' (Promotional movie) which has 1,304 views and was uploaded 9 months ago. Below the video, there is a section titled '人気のアップロード' (Popular uploads) showing two other video thumbnails: 'PCRプライマー設計ツール Primer3の使い方' (How to use the PCR primer design tool Primer3) and 'パワーポイントの図形描画機能でイラストをつくる方法' (How to create illustrations using the drawing function of Microsoft PowerPoint).

オンライン人体地図サービス「アнатモグラフィー」を使い倒す

6年前。再生回数 11,751 回

高画質はこちら <http://togotv.dbcls.jp/20080606.html>

アнатモグラフィーは統合ホームページで提供されているツールの一つで、器官別の発現...

©2014 DBCLS Licenced Under CC 表示 2.1



統合TV YouTube 支店 統計情報

アナリティクス - YouTube
https://www.youtube.com/analytics?o=U#dt=lt,fs=13733,fe=16309,fr=lw-001;r=views,fcr=0,rpm=t,rpa=a,rps=

チャンネル togotv
作成日: 2007/08/08 動画: 835 全期間の再生回数: 398,429

全期間 (2007/08/08~2014/08/27)

△ このレポートのデータは不完全か不足しています。
"推定再生時間 (分)" のデータは 2012年9月1日 以降に限り使用できます。

再生回数レポート 再生回数

概要 収益レポート 推定収益額 広告の掲載結果

ユーザー層 再生場所 トライフィック ソース 端末 視聴者維持率

視聴者の反応レポート 登録者 評価 お気に入り コメント 共有 アノテーション

作成 フィードバックの送信

再生回数 398,429 推定再生時間 (分) 599,240*

統計情報の比較 月別 増加率を表示

再生回数 * 推定再生時間 (分) *

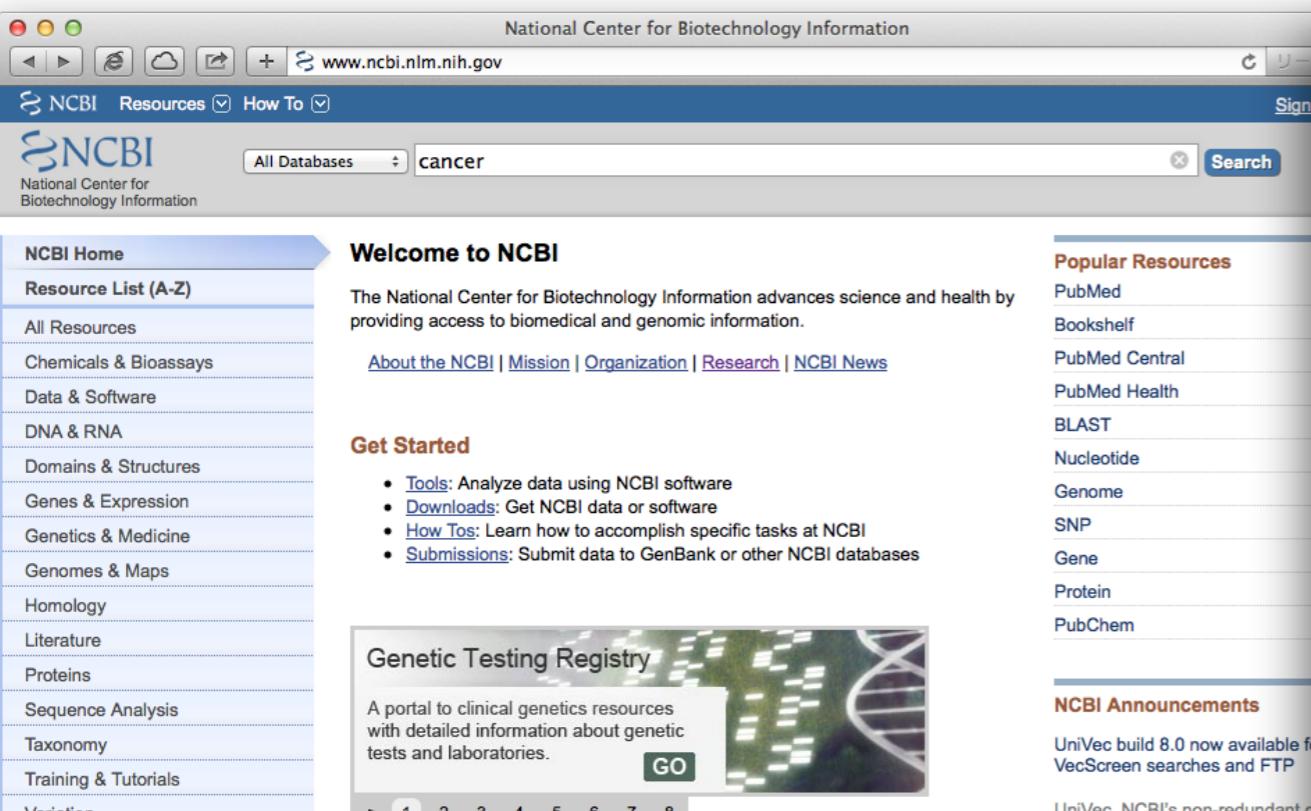
* 選択した期間に不完全な月が含まれています。

動画 地域 日付 字幕

動画	再生回数	推定再生時間 (分)
PCRプライマー設計ツール Primer3の使い方	20,857 (5.2%)	7,468 (1.2%)
パワーポイントの図形描画機能でイラストを...	18,930 (4.8%)	62,313 (10%)
オンライン人体地図サービス「アナトモグラ...	11,571 (2.9%)	692 (0.1%)
PCRプライマー設計ツール Primer3の使い方 ...	8,728 (2.2%)	22,342 (3.7%)
パワーポイントの図形描画機能でイラストを...	8,138 (2.0%)	28,548 (4.8%)
統計解析ソフト「R」の使い方 ~導入編~	7,624 (1.9%)	3,171 (0.5%)
PyMOLを使い倒す	6,152 (1.5%)	4,552 (0.8%)
統計解析ソフト「R」の使い方 ~ヒートマッ...	5,931 (1.5%)	

©2014 DBCLS Licenced Under CC 表示 2.1

- National Center for Biotechnology Information
(米国立生物工学情報センター) の略称
- PubMed、核酸(GenBank)、アミノ酸配列、発現データ、
PubChem、タンパク質構造、ゲノム、Taxonomyといった
NCBIの主要な30以上のデータベースに対して、テキスト
ベースで統合的に検索できる



The screenshot shows the NCBI homepage with a search bar containing "cancer". Below the search bar, there's a "Popular Resources" sidebar with links to PubMed, Bookshelf, PubMed Central, PubMed Health, BLAST, Nucleotide, Genome, SNP, Gene, Protein, and PubChem.

Welcome to NCBI
The National Center for Biotechnology Information advances science and health by providing access to biomedical and genomic information.
[About the NCBI](#) | [Mission](#) | [Organization](#) | [Research](#) | [NCBI News](#)

Get Started

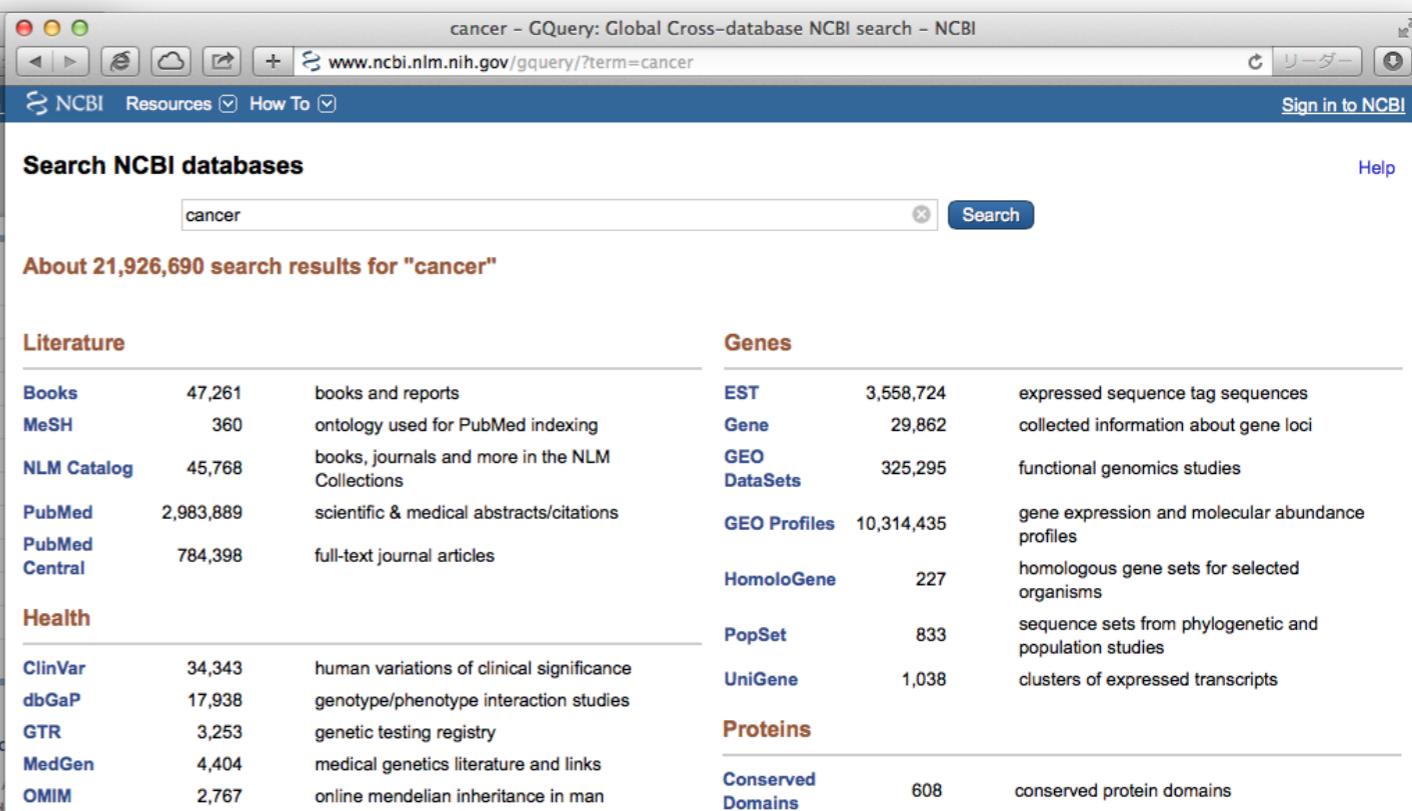
- [Tools](#): Analyze data using NCBI software
- [Downloads](#): Get NCBI data or software
- [How Tos](#): Learn how to accomplish specific tasks at NCBI
- [Submissions](#): Submit data to GenBank or other NCBI databases

Genetic Testing Registry
A portal to clinical genetics resources with detailed information about genetic tests and laboratories. [GO](#)

NCBI Announcements
UniVec build 8.0 now available for VecScreen searches and FTP
UniVec, NCBI's non-redundant database of vector sequences, has been updated.

General Research Use collection
streamlines access to patient-level data in dbGaP

In response to many requests for
NCBI/CDC/FDA/USDA collaborate using whole genome sequencing to improve food safety in humans.



The screenshot shows the GQuery search results page for "cancer". It displays approximately 21,926,690 search results across various databases. The results are categorized into Literature, Genes, Health, Proteins, and Genomes.

Literature

Category	Count	Description
Books	47,261	books and reports
MeSH	360	ontology used for PubMed indexing
NLM Catalog	45,768	books, journals and more in the NLM Collections
PubMed	2,983,889	scientific & medical abstracts/citations
PubMed Central	784,398	full-text journal articles

Genes

Category	Count	Description
EST	3,558,724	expressed sequence tag sequences
Gene	29,862	collected information about gene loci
GEO DataSets	325,295	functional genomics studies
GEO Profiles	10,314,435	gene expression and molecular abundance profiles
HomoloGene	227	homologous gene sets for selected organisms
PopSet	833	sequence sets from phylogenetic and population studies
UniGene	1,038	clusters of expressed transcripts

Health

Category	Count	Description
ClinVar	34,343	human variations of clinical significance
dbGaP	17,938	genotype/phenotype interaction studies
GTR	3,253	genetic testing registry
MedGen	4,404	medical genetics literature and links
OMIM	2,767	online mendelian inheritance in man
PubMed Health	8,512	clinical effectiveness, disease and drug reports

Proteins

Category	Count	Description
Conserved Domains	608	conserved protein domains
Protein	886,753	protein sequences
Protein Clusters	14	sequence similarity-based protein clusters
Structure	7,065	experimentally-determined biomolecular structures

Genomes

Category	Count	Description
Assembly	27	genomic assembly information
BioProject	9,024	biological projects providing data descriptions of biological sources
BioSample	207,459	descriptions of biological sources

©2014 DBCLS Licensed Under CC 表示 2.1 

NCBIの総エントリー数は?

<http://www.ncbi.nlm.nih.gov/gquery/?term>All%5Bfilter%5D>

All[filter] – GQuery: Global Cross-database NCBI search – NCBI

NCBI Resources How To Sign in to NCBI

Search NCBI databases

All[filter]

About 1,226,289,377 search results for "All[filter]"

All[filter] で検索!!

Literature			Genes		
Books	304,403	books and reports	EST	75,563,489	expressed sequence tag sequences
MeSH	252,518	ontology used for PubMed indexing	Gene	17,307,298	collected information about gene loci
NLM Catalog	1,506,157	books, journals and more in the NLM Collections	GEO DataSets	1,254,935	functional genomics studies
PubMed	24,068,783	scientific & medical abstracts/citations	GEO Profiles	91,392,791	gene expression and molecular abundance profiles
PubMed Central	3,169,617	full-text journal articles	HomoloGene	141,268	homologous gene sets for selected organisms
Health					
ClinVar	111,144	human variations of clinical significance	PopSet	205,778	sequence sets from phylogenetic and population studies
dbGaP	163,233	genotype/phenotype interaction studies	UniGene	6,473,284	clusters of expressed transcripts
GTR	31,919	genetic testing registry	Proteins		
MedGen	258,064	medical genetics literature and links	Conserved Domains	49,641	conserved protein domains
OMIM	23,634	online mendelian inheritance in man	Protein	145,632,691	protein sequences
PubMed Health	48,786	clinical effectiveness, disease and drug reports	Protein Clusters	820,546	sequence similarity-based protein clusters
Genomes					
Assembly	32,010	genomic assembly information	Structure	101,427	experimentally-determined biomolecular structures
BioProject	132,407	biological projects providing data to NCBI	Chemicals		
BioSample	2,674,287	descriptions of biological source materials	BioSystems	612,889	molecular pathways with links to genes, proteins and chemicals
Clone	36,916,420	genomic and cDNA clones	PubChem BioAssay	1,048,167	bioactivity screening studies
dbVar	4,126,864	genome structural variation studies	PubChem Compound	52,773,083	chemical information with structures, information and links
Epigenomics	6,634	epigenomic studies and display tools	PubChem Substance	149,338,355	deposited substance and chemical information
Genome	10,107	genome sequencing projects by organism			
GSS	37,613,354	genome survey sequences			
Nucleotide	143,865,973	DNA and RNA sequences			
Probe	31,883,086	sequence-based probes and primers			
SNP	394,164,614	short genetic variations			
SRA	927,243	high-throughput DNA and RNA sequence read archive			
Taxonomy	1,278,880	taxonomic classification and nomenclature catalog			

You are here: NCBI > GQuery

GETTING STARTED RESOURCES POPULAR FEATURED NCBI INFORMATION

©2014 DBCLS Licenced Under CC 表示 2.1

Ensembl

<http://www.ensembl.org/>

- ・ 英国 Wellcome Trust Sanger InstituteとEBI(European Bioinformatics Institute)が共同開発しているゲノムブラウザ
- ・ 遺伝子の注釈(アノテーション)情報をゲノム上で閲覧(ブラウズ)することができる

The image shows two side-by-side screenshots of the Ensembl Genome Browser.

Left Screenshot (Home Page):

- Header: Ensembl Genome Browser
- Search bar: Search: All species for e.g. BRCA2 or rat X:100000..200000 or coronary heart disease
- Browse a Genome section: Human (GRCh37), Mouse (GRCm38), Zebrafish (Zv9)
- Popular genomes: Human, Mouse, Zebrafish
- Links: ENCODE data in Ensembl, Variant Effect Predictor (VeP), Gene expression in different tissues, Find SNPs and other variants for my gene, Retrieve gene sequence, Compare genes across species, Use my own data in Ensembl, Learn about a disease or phenotype
- What's New in Release 75 (February 2014): New VEP interface, New 'Age of base' track for human, New GENCODE basic renderer for human and mouse
- Latest blog posts: Retirement of archives 63, Variation annotation for G, The Ensembl Regulation Analysis Pipeline
- Did you know...? FAQs

Right Screenshot (Gene Page: BRCA2):

- Header: Ensembl genome browser 75: Homo sapiens - Summary - Gene: BRCA2 (ENSG00000139618)
- Search bar: Search all species...
- Location: 13:32,889,611-32,973,805 Gene: BRCA2
- Gene-based displays menu:
 - Summary
 - Splice variants (6)
 - Transcript comparison
 - Supporting evidence
 - Sequence
 - Secondary Structure
 - External references
 - Regulation
 - Expression
 - Comparative Genomics
 - Genomic alignments
 - Gene tree (image)
 - Gene tree (text)
 - Gene tree (alignment)
 - Gene gain/loss tree
 - Orthologues (61)
 - Paralogues
 - Protein families (2)
 - Phenotype
 - Genetic Variation
 - Variation table
 - Variation image
 - Structural variation
 - External data
 - Personal annotation
 - ID History
 - Gene history
- Gene: BRCA2 ENSG00000139618
- Description: breast cancer 2, early onset [Source:HGNC Symbol;Acc:1101]
- Location: Chromosome 13: 32,889,611-32,973,805 forward strand.
- INSDC coordinates: chromosome:GRCh37:CM000675.1:32889611:32973805:1
- Transcripts: This gene has 6 transcripts (splice variants) Show transcript table
- Summary: BRCA2 (HGNC Symbol), BRCC2, FACD, FAD, FAD1, FANCD, FANCD1
- CCDS: This gene is a member of the Human CCDS set: CCDS9344
- RefSeq: Overlapping RefSeq Gene ID 675 matches and has similar biotype of protein_coding
- LRG: LRG_293 provides a stable genomic reference framework for describing sequence variations for this gene
- Ensembl version: ENSG00000139618.10
- Gene type: Known protein coding
- Prediction Method: Annotation for this gene includes both automatic annotation from Ensembl and Havana manual curation, see article.
- Alternative genes: This gene corresponds to the following database identifiers: Havana gene: OTTHUMG0000017411 (version 4)
- Buttons: Configure this page, Add your data, Export data, Bookmark this page, Share this page
- Region in Detail: Go to Region in Detail for more tracks and navigation options (e.g. zooming)
- Panel: Genes (GENCODE...) showing the genomic region for BRCA2-001, spanning 104.19 kb from 32.88 Mb to 32.98 Mb, on the Forward strand. The panel includes tracks for Genes (GENCODE...), BRCA2-001, and IFIT1P1-001 > processed pseudogene.

BioMart

<http://www.biomart.org>

- Ontario Institute for Cancer Research (OICR) と European Bioinformatics Institute (EBI)、San Raffaele Scientific Institute (SRSI) が共同で開発しているデータ管理システム
- 様々なデータを取得するためのインターフェースとして秀逸

The image shows two screenshots of the BioMart interface. On the left is the main BioMart website at www.biomart.org, featuring a sidebar with links like HOME, TOOLS, COMMUNITY, PUBLICATIONS, NEWS, CREDITS, DOCUMENTATION, VERSION 0.7, and CONTACT. The main content area displays the BioMart logo, a brief description, and four tool icons: BROWSE DATA, ID CONVERSION, SEQUENCE RETRIEVAL, and ENRICHMENT ANALYSIS. A yellow box at the bottom left contains text about the BioMart community and its data federation technology. On the right is the BioMart Central Portal at central.biomart.org/martwizard/, showing a search interface for Ensembl 75 Genes (WTSI, UK). The portal includes sections for ATTRIBUTES, GENE:, and PHENOTYPE, each with a list of selectable items. A summary panel on the right provides details about the database, datasets, filters, and attributes.

BioMart

[3rd July 2014] BioMart Portals (0.7 and 0.9) updates: urgi, vectorbase, phytozome, uniprot, ensembl genomes.

BioMart

is a community-driven project to provide unified access to distributed research data to facilitate the scientific discovery process. The project adheres to the open source philosophy that promotes collaboration and code reuse.

BROWSE DATA ID CONVERSION SEQUENCE RETRIEVAL ENRICHMENT ANALYSIS

46 DATABASES, 4 CONTINENTS

A large number of servers that provide access to a wide range of research data have been set up by the BioMart community. Using BioMart's unique data federation technology, a Central

BioMart Central Portal

Home > Search by database name (A-Z)

Datasets → Filters → Output ↻ Restart ← Previous Results »

ATTRIBUTES

Features Structures Transcript Event Homologs Variation Sequences

GENE:

ENSEMBL

Ensembl Gene ID Ensembl Transcript ID
 Ensembl Protein ID Ensembl Exon ID
 Description Chromosome Name
 Gene Start (bp) Gene End (bp)
 Strand Band
 Transcript Start (bp) Transcript End (bp)
 Associated Gene Name Associated Transcript Name
 Associated Gene DB Associated Transcript DB
 Transcript count % GC content
 Gene Biotype Transcript Biotype
 Source (gene) Source (transcript)
 Status (gene) Status (transcript)

PHENOTYPE

Phenotype description Source name
 Study External Reference Strain name
 Strain gender P value

SUMMARY

Database: Ensembl 75 Genes (WTSI, UK)

Datasets: Homo sapiens genes (GRCh37.p13)

Filters: no data

Attributes: no data

©2014 DBCLS Licensed Under CC 表示 2.1

BioMart

<http://www.biomart.org>

- ・ 【実習】 BioMart v0.9を使って遺伝子情報や配列を取得する
- ・ 【統合TV】
BioMart v0.9を使い倒す ~遺伝子情報や配列を取得する~
http://youtu.be/Cc43P2fz8_k



UCSC Genome Browser

<https://genome.ucsc.edu/cgi-bin/hgGateway>

- 米国カリフォルニア大学サンタクルーズ校(UCSC)が開発しているゲノムブラウザ、ゲノムアノテーション閲覧システム
- 前述のEnsemblと双璧をなす
 - どちらが使いやすいかは、その人次第!?

Human (Homo sapiens) Genome Browser Gateway

The UCSC Genome Browser was created by the [Genome Bioinformatics Group of UC Santa Cruz](#). Software Copyright (c) The Regents of the University of California. All rights reserved.

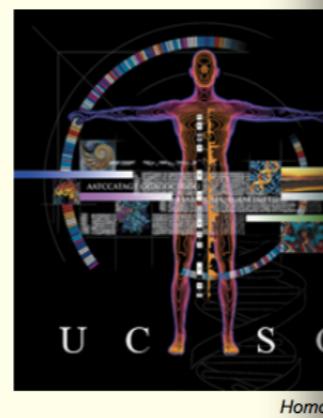
group	genome	assembly	position	search term
Mammal	Human	Feb. 2009 (GRCh37/hg19)	chr21:33,031,597-33,041,570	enter position, gene symbol or search terms

Click here to reset the browser user interface settings to their defaults.

track search add custom tracks track hubs configure tracks and display

Human Genome Browser – hg19 assembly (sequences)

The February 2009 human reference sequence (GRCh37) was produced by the [Genome Reference Consortium](#). For more information about this assembly, see [GRCh37](#) in the NCBI Assembly database.



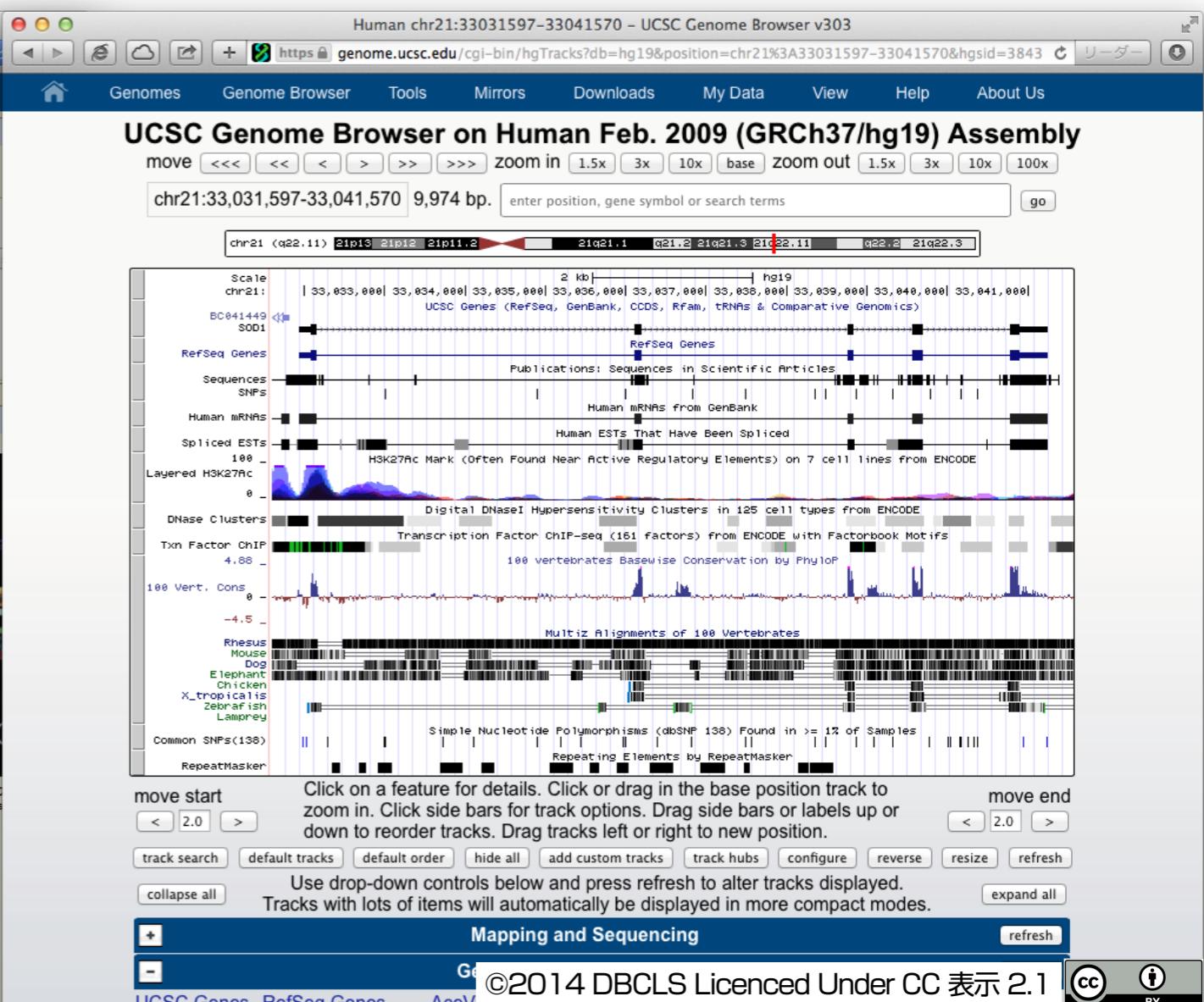
Sample position queries

A genome position can be specified by the accession number of a sequenced genomic clone, an mRNA or EST or STS marker, a chromosomal coordinate range, or keywords from the GenBank description of an mRNA. The following list shows examples of valid position queries for the human genome. See the [User's Guide](#) for more information.

Request: Genome Browser Response:

chr7	Displays all of chromosome 7
chrUn_gl000212	Displays all of the unplaced contig gl000212
20p13	Displays region for band p13 on chr 20
chr3:1-1000000	Displays first million bases of chr 3, counting from p-arm telomere
chr3:1000000+2000	Displays a region of chr3 that spans 2000 bases, starting with position 1000000

RH18061;RH80175
15q11;15q13
rs1042522;rs1800370 Displays region between genome landmarks, such as the STS markers RH18061 and RH80175, or chromosome bands 15q11 to 15q13, or SNPs rs1042522 and rs1800370. This syntax may also be used for other range queries, such as between uniquely determined ESTs, mRNAs, refSeqs,



UCSC Genome Browser

<https://genome.ucsc.edu/cgi-bin/hgGateway>

- ・【実習】 wig形式のファイルをトラックとして追加して表示する
- ・【統合TV】 UCSC Genome Browserの使い方～wig形式のファイルをトラックとして追加する～
- ・GSE21365
- ・【呪文】 track type=wiggle_0

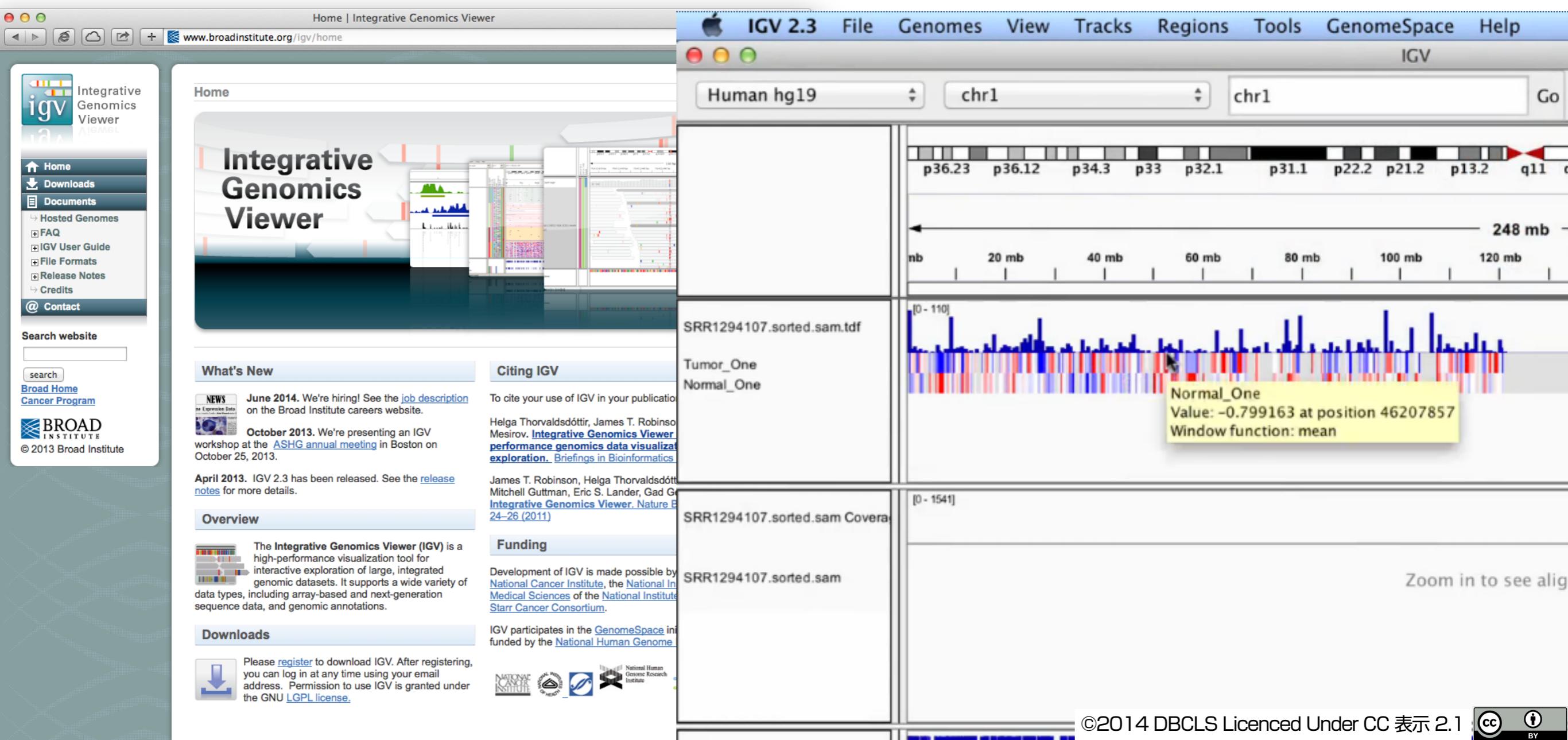
<http://www.youtube.com/watch?v=St00ZcfEeo>



Integrative Genomics Viewer (IGV)

<http://www.broadinstitute.org/igv/>

- ・米国 Broad Institute が提供するゲノムデータ可視化ツール
- ・個別のマシンにインストールして使うので、手元に必要なデータがあれば、それらを高速に表示、閲覧することが可能



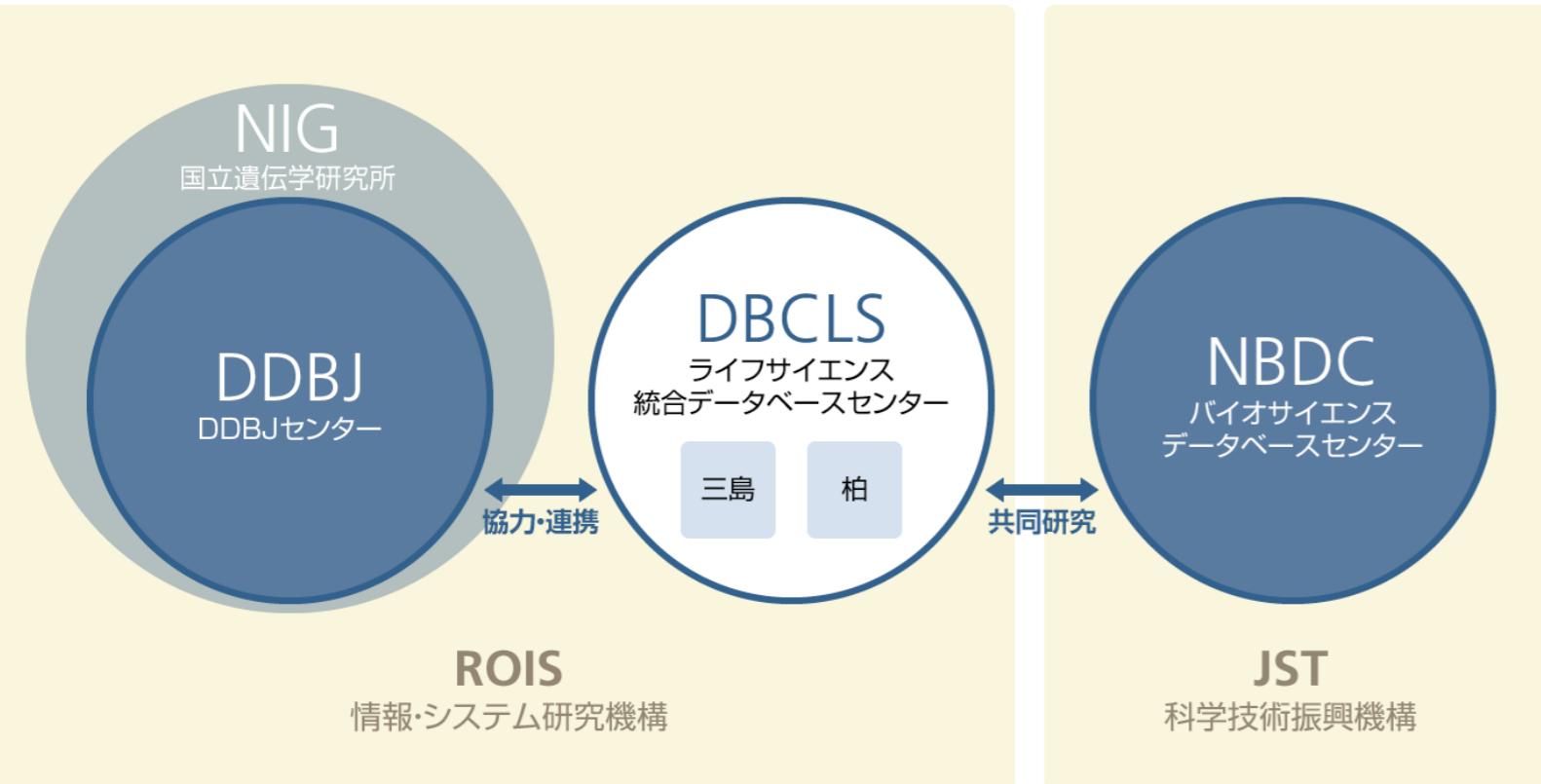
Integrative Genomics Viewer (IGV)

<http://www.broadinstitute.org/igv/>

- ・ 【実習】アノテーションデータの読み込み、表示の切替、検索、特定の塩基配列の取得
- ・ 【統合TV】
Integrative Genomics Viewer IGVを使い倒す～基本編～
データ元: http://www.nimblegen.com/downloads/annotation/hg19/human_hg19_september_2011.zip
<http://youtu.be/kHlga3aca0k>



NBDC と DBCLS と 統合DB



<http://dbcls.rois.ac.jp/about>

2006年9月	文部科学省ライフサイエンス分野の統合データベース整備事業 開始
2007年4月	ライフサイエンス統合データベースセンター 設立 初代センター長 高木利久 就任 統合データベースプロジェクトの中核機関として事業を受託（期間4年）
2010年10月	第2代センター長 米澤明憲 就任
2011年4月	JSTバイオサイエンスデータベースセンター 設立 JSTライフサイエンスデータベース統合推進事業 開始 「基盤技術開発プログラム」開始（期間3年）
2012年4月	第3代センター長 小原雄治 就任 ライフサイエンス統合データベースの推進方策について（報告）を発表
2014年4月	NBDC「統合データベースにおける基盤技術開発とデータベース運用に係る共同研究」開始

<http://dbcls.rois.ac.jp/about/history>

- 必要なデータベースが見つからない
- 使い方がわからない
- データを組み合わせた高度な解析ができない



データベースを効率よく利用するための環境整備



統合データベース

ライフサイエンス統合データベースセンター (DBCLS)

<http://dbcls.rois.ac.jp>

The screenshot shows the homepage of the DBCLS website. The header features the DBCLS logo and the URL <http://dbcls.rois.ac.jp>. The main content area displays a large image of a modern office hallway. Below the image, a section titled "news" lists five recent articles:

- 2014/08/21 **Open Science Award 2014を開催します**
国内の生命医学分野におけるオープンな成果物を、コミュニティによる投票を元に表彰するオープン・サイ...
・イベント
- 2014/08/21 **全サービス一時停止のお知らせ<9月27日(土) 10:00～28日(日) 19:00>**
DBCLSが運用するサーバ群が停電作業を行うため、2014年9月27日(土) の10:00～翌日28日（日）の19:0...
・サービス
- 2014/08/21 **「NBDC / DBCLS BioHackathon 2014」を11月9～14日に開催します**
ライフサイエンス分野におけるデータ統合の最新技術の共有や、標準化・互換性についての議論と開発を合...
・イベント
- 2014/08/20 **SPARQLthon23を開催しました**
ライフサイエンス分野におけるセマンティック・ウェブ技術の話題を中心として、国内の技術者/研究者が集...
・イベント
- 2014/08/12 **統合データベース講習会：AJACS十勝 (2014年9月12日@帯広畜産大学) の参加者募集中です**
帯広畜産大学 情報処理センター1階 第1実習室 にてバイオサイエンスデータベースセンター (NBDC) 主催 D...
・イベント
・募集

A "more" button is located at the bottom of the news list. The sidebar on the left contains links to "センターについて", "研究開発", "サービス", "イベント", "メンバー", "アクセス", and "お問い合わせ". It also includes social media icons for Twitter and RSS, and a search bar.

At the bottom, there is a "site map" link and a footer containing the DBCLS logo and copyright information: "©2014 DBCLS Licenced Under CC 表示 2.1".



大学共同利用機関法人 情報・システム研究機構
ライフサイエンス統合データベースセンター

©2014 DBCLS Licenced Under CC 表示 2.1



バイオサイエンスデータベースセンター (NBDC)

<http://biosciencedbc.jp>

The screenshot shows the NBDC homepage with a red circle highlighting the '生命科学全体のデータベース統合' (Integration of all life science databases) section. This section includes links to Integbio, a cross-database catalog, and a search function for both domestic and international databases.

The page also features several other sections:

- 日本語や動画でわかりやすく**: Includes links to new literature reviews and domain fusion reviews, as well as a combined TV section.
- 論文をもっと読みやすく、書きやすく**: Includes links to Allie, inMeXes, and TogoDoc.
- 大量の配列データを扱いやすく**: Includes links to DBCLS SRA, Sequencing (β), RefEx, and a combined gene search for GGRNA.
- さまざまな統合コンテンツ**: Includes links to biological icons, main project overviews, web resource portals, genome analysis tool link collections, HOWDY, and GenLibi.
- 開発ツール**: Includes links to TogoDB, TogoWS, DBCLS Galaxy, BodyParts3D, and Anatomography.
- 統合のための連携**: Includes links to integbio.jp (4 prefecture joint portal site) and BioHackathon.

On the right side, there is a news section with recent updates, a Twitter feed, and an RSS link. At the bottom, there are logos for Gender-Equal Society and Research Integrity, along with a Creative Commons license notice.

Integbio データベースカタログ

<http://integbio.jp/dbcatalog/>

ホーム - Integbio データベースカタログ

検索結果 - Integbio データベースカタログ

- 生命科学系データベースを一覧から探す -

Integbio データベースカタログ

全条件をリセット

一覧内を検索する

一覧を絞り込む

生物種

- + 動物 (516)
- + 植物 (231)
- + 原生生物 (40)
- + 菌類 (73)
- + 真正細菌 (128)
- 古細菌 (43)
- ウイルス (47)

カテゴリ

- <対象>
- ゲノム (176)
- 遺伝子 (294)
- cDNA (187)
- + 続きを見る

<データの種類>

- 配列 (515)
- 構造 (198)
- 遺伝子発現 (147)
- + 続きを見る

稼動状況

- 稼動中
- 休止
- 運用終了

地域

- 日本
- 日本以外の国・地域

その他の条件

- LSDBアーカイブ収載のDB

検索機能について

データベースのレコード一覧

1374 件

最初へ 前へ 1

Raphanus sativus Genome

運用機関: かずさディー・エヌ・ベイ

生物種: *Raphanus sativus*

説明: ダイコン (*Raphanus sativus*) のデータベースです。EST-SSR、Isozyme marker の配列情報を収録... 詳細へ

追加

Carnation DB

運用機関: かずさディー・エヌ・ベイ

生物種: *Dianthus caryophyllus*

説明: カーネーション (*Dianthus caryophyllus*) のデータベースです。遺伝子、UniProtなどの... 詳細へ

追加

日本人由来B細胞株・DN

運用機関: 独立行政法人 医療機能評価機構

生物種: *Homo sapiens*

説明: 約2,000名の日本人由来B細胞株を分譲しています。 詳細へ

追加

ヒト組織バンクデータベ

運用機関: 独立行政法人 医療機能評価機構

生物種: *Homo sapiens*

説明: 患者への十分な説明と同様の手続きを厳格に行った上

追加

KNAPSAck Bicycle: KNAP

運用機関: 奈良先端科学技術研究センター

生物種: *Stramenopiles* | *Cyanobacteria*

説明: 藻類のタンパク質と二方向で互いに作用するタンパク質を在する96,641個のタンパク質の情報を調べられ... 詳細へ

追加

TradMPD: 伝統医薬データ

運用機関: 富山大学 和漢医学研究所

生物種:

説明: 生薬や漢方薬、およびその成分の構造情報を用いて、生薬の学術情報や関連植物の成分情報を検索することができるデータベースです。生薬の学術情報や関連植物の成分情報を調べられます。 詳細へ

追加

NCBNカタログデータベ

運用機関: 国立高度専門医療研究センター

生物種: *Homo sapiens*

説明: 国立高度専門医療研究センターが運営することができるデータベースです。疾患、治療法、薬理作用、副作用、感染症・代... 詳細へ

追加

病気の解説 難病情報セン

運用機関: 難病医学研究財団

生物種: *Homo sapiens*

説明: 病気の解説、難病情報センターアーカイブなどを収載することができるデータベースです。 詳細へ

追加

配列 (62)

構造 (6)

生物種: ヒト

カテゴリ: 遺伝子

配列

62 件

最初へ 前へ 1 2 次へ 最後へ

NBDCヒトデータベース

運用機関: 独立行政法人科学技術振興機構 バイオサイエンスデータベースセンター

生物種: *Homo sapiens*

説明: 本サイトはヒトに関する様々なデータを共有するためのプラットフォームです。次世代シーケンサーをはじめとした解析技術の発達に伴い産生されつつある膨大な量のヒト関連データを整理・格... 詳細へ

ASPicDB: Alternative Splicing Prediction Data Base

運用機関: Consorzio Interuniversitario per le Applicazioni di Supercalcolo per Universita e Ricerca

生物種: *Homo sapiens* | *Mus musculus* | *Rattus norvegicus* | *Gallus gallus* | *Bos taurus* | ...

説明: オルタナティブスプライシング (選択的スプライシング) を受ける遺伝子とタンパク質に関するデータベースです。オルタナティブスプライシングパターンによる条件絞込み検索が可能で、それぞ... 詳細へ

miRTarBase

運用機関: Institute of Bioinformatics and Systems Biology, National Chiao Tung University

生物種: *Homo sapiens* | *Mus musculus* | *Rattus norvegicus* | *Gallus gallus* | *Bos taurus* | ...

説明: 世界最大規模および最新のmiRNA-ターゲット相互作用 (miRNA-target interaction; MTI) を蓄積しているデータベースです。MTIの収集は、miRNAの機能解析研究に関する文献からテキストマイニング... 詳細へ

CCDB: Cervical Cancer Gene Database

運用機関: Institute of Cytology and Preventive Oncology

生物種: *Homo sapiens*

説明: 子宮頸癌に関わる遺伝子のカタログデータベースです。実験結果をもとに子宮頸癌の様々なステージに関与するであろうと考えられる遺伝子及びmiRNAの情報を、手作業のキュレーションにより収集... 詳細へ

Mutation Database

運用機関: 独立行政法人科学技術振興機構 バイオサイエンスデータベースセンター

生物種: *Homo sapiens*

説明: 疾患原因遺伝子や関連遺伝子の変異に関する情報とそれに伴う疾患の臨床情報をアーカイブし共有することを目的としたレポジトリデータベースです。ラボでの独自の解析結果と文献から抽出した... 詳細へ

Human Variation DB: human genome variation database

運用機関: 独立行政法人科学技術振興機構 バイオサイエンスデータベースセンター

生物種: *Homo sapiens*

説明: 次世代シーケンサーによって検出された健常者変異、疾患関連変異、及び、文献に発表されている疾患関連変異を抽出・登録し、GUIにより閲覧できるようにしています。また、日本人健常者の参... 詳細へ

ASAP II: the Alternative Splicing Annotation Project II

運用機関: Universitatis Regensburg

生物種: *Homo sapiens*

説明: ©2014 DBCLS Licenced Under CC 表示 2.1

CC BY

生命科学データベース横断検索

<http://biosciencedbc.jp/dbsearch/>

講習会用

The screenshot shows the 'Life Science CROSS' search interface. At the top, there's a search bar with the placeholder '検索' (Search) and a magnifying glass icon. Below it, a navigation bar includes 'HOME', 'ABOUT', 'CONTACT', 'SEARCH', 'SEARCH HELP', and 'SEARCH HISTORY'. On the left, there's a sidebar with 'Information' and 'What's New' sections, and a 'DATABASE' selection button. The main content area displays search results for '高血圧' (Hypertension), with a red circle highlighting the search term in the header.

This screenshot shows the detailed search results for '高血圧' (Hypertension). The header shows the search term '高血圧' highlighted with a red circle. Below it, a blue bar indicates '全てのデータベース(95857)' (All databases (95857)). The results list several entries, each with a title, a brief abstract, and a URL. Some specific terms in the abstracts are highlighted with red circles. A yellow circle highlights the '統合DBプロジェクト' (Integrated DB Project) link. Another yellow circle highlights the '用語解説' (Glossary) link under a section about primary pulmonary hypertension. A third yellow circle highlights the '遺伝子発現・転写制御 - 転写モチーフ' (Gene Expression - Transcription Control - Transcription Motif) link under a section about hypertension-associated homologs.

http://biosciencedbc.jp/dbsearch_new/

– 関連遺伝子で検索

acyl-CoA synthetase medium-chain family member 3
[GENE:6296][GNP:6296][UNP:]
[HGNC:10522]

Hypertension, essential, susceptibility to
[GENE:50986][GNP:50986][UNP:]
[HGNC:]

Hypertension, essential, susceptibility to
[GENE:444980][GNP:444980][UNP:]
[HGNC:]

Hypertension, essential, susceptibility to
[GENE:387575][GNP:387575][UNP:]
[HGNC:]

primary pulmonary hypertension 2
[GENE:89873][GNP:89873][UNP:]
[HGNC:15733]

Hypertension, essential, susceptibility to
[GENE:117191][GNP:117191][UNP:]
[HGNC:]

hypertension with brachydactyly
[GENE:8080][GNP:8080][UNP:]
[HGNC:]

Renal failure, progressive, with
hypertension
[GENE:59331][GNP:59331][UNP:]
[HGNC:]

– 関連語句で検索

検索クエリ
高血圧 | "hypertension"

翻訳結果

[高血圧]
hypertension
hypertensive
ht



生命科学系データベースアーカイブ

<http://dbarchive.biosciencedbc.jp>

The screenshot shows the homepage of the LSDB Archive. At the top right, there are language selection buttons for Japanese and English, and a link to the depositor login page. Below the header, there's a search bar for cross-archive search and radio buttons for similarity search and image search. The main content area includes a brief introduction about the archive's purpose and a diagram illustrating the submission process between data users and database creators. A sidebar on the right lists new archive information.

いくら良質なデータベースでも、説明が十分でない、利用条件が明確でない、ダウンロードできないなどの理由で十分に利用され、引用され、相応しい評価をうける機会を逃していることがあります。

生命科学系データベースアーカイブは、国内のライフサイエンス研究者が生み出したデータセットをわが国の公共財としてまとめて長期間安定に維持保管し、データ説明（メタデータ）を統一して検索を容易にすると共に、利用許諾条件などの明示を行うことで、多くの人が容易にデータへアクセスしダウンロードを行えるようにするサービスです（[詳細説明](#)）。

データを長期にわたり保全し、データベース作成者のクレジットを明示する一方、公的機関や民間等様々なユーザが利用しやすい形にすることで、それぞれの研究の生命科学へのいっそうの貢献を支援します。データベースの寄託を随時募集しています（[寄託応募要領](#)）。

新アーカイブ情報

- 2014/07/23 「おもしろい糸状菌」（農研機構 北海道農業研究センター）を追加しました
- 2014/07/17 「PLACE」（農業生物資源研究所）を追加しました
- 2014/06/10 「SEVENS」（青山学院大学 諏訪牧子教授）を追加しました

アーカイブデータベース一覧 (ヘルプ)

一覧内検索

全 87 件 (1 件から10件)

10 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	生物種	要約 (キーワードを太字表示)	利用許諾
 おもしろい糸状菌 ダウンロード 簡易検索	独立行政法人 農業・食品産業技術総合研究機構 北海道農業研究センター	-	図鑑	菌類	海洋性の糸状菌17種について、顯微鏡写真、学名、形態の説明をまとめたデータベース	CC 表示-継承 詳細
 PLACE ダウンロード 簡易検索 オリジナルサイト	農業生物資源研究所	肥後 健一	発現制御	植物	高等植物遺伝子のプロモーター等の発現制御に関わる領域の塩基配列モチーフを論文から収集し、解説等を付加したデータベース	CC 表示-継承 詳細
 SEVENS ダウンロード 簡易検索 オリジナルサイト	産業技術総合研究所 ゲノム情報研究センター	諏訪 牧子	タンパク質	真核生物	7本膜貫通ヘリックス型タンパク質のGPCR遺伝子を網羅的に収めたデータベース	CC 表示-継承 詳細
 KNApSAcK Family Bicycle ダウンロード 簡易検索 オリジナルサイト	奈良先端科学技術大学院大学 情報科学研究科 計算システムズ生物学研究室	金谷 重彦	代謝	各種藻類	藻類のタンパク質と二次代謝反応の情報を集めたデータベース	CC 表示-継承 詳細

統合データベース講習会 AJACS, MotDB

<http://motdb.dbcls.jp/>



MotDB – MotDB

トップ 編集 凍結 差分 バックアップ 添付 リロード 新規 一覧 単語検索 最終更新 ヘルプ リンク元

MotDB

MotDB

ようこそMotDB（「もっとでいーびー」と発音、Master of the database【日本語訳：データベースの達人】の略）へ。このサイトでは、ライフサイエンス分野のデータベース統合を担う人材 AJACS(All Japan Annotator/Curator/System DB administrator)向けの教材を提供しています。このサイトは、文部科学省委託研究開発事業「統合データベースプロジェクト」（平成18年度～平成22年度）の元で運用されてきましたが、引き続き、ライフサイエンス統合データベースセンター(DBCLS: Database Center for Life Science)により維持管理されています。

- 統合データベース講習会 AJACS
 - 募集！あなたの地域で「統合データベース講習会」を開催してみませんか？
 - 今年度の募集は終了しました
- 今後の予定
 - 統合データベース講習会：AJACS阿波（2014年8月20日開催・徳島大学）講習会資料



AJACS48/nakazato – MotDB

トップ 編集 凍結 差分 バックアップ 添付 リロード 新規 一覧 単語検索 最終更新 ヘルプ リンク元

MotDB

AJACS48/nakazato

「次世代シーケンサー（NGS）と関連するデータベース・ツール」 担当：仲里 猛留

AJACS信濃 > 「次世代シーケンサー（NGS）と関連するデータベース・ツール」

おことわり

- 次世代シーケンシングデータのコマンドラインによる解析は今日は演習はやりません。
 - ↓時間がない。マシンパワーとディスク容量が必要なため。データの転送に膨大な時間がかかるため。
 - 目的別に解析の方法が違う
- 時間があれば紙芝居（あるいは3分クッキング方式）でお見せします
- ↓かわりに参考サイト

参考

- (Rで)塩基配列解析（主に次世代シーケンサーのデータ）（門田幸二さん：東京大学 大学院農学生命科学研究科 アグリバイオインフォマティクス教育研究ユニット）
- お家でできるMac Bookでやる次世代シーケンスデータ解析（PDFへのリンク）（緒方法親さん：東京農工大学応用昆虫学研究室（当時））
- 統合TV：R+Bioconductorを使ったNGS解析1（二階堂愛さん：理化学研究所 発生・再生科学総合研究センター（当時））

ライフサイエンス分野のデータベース統合を担う人材
AJACS (All Japan Annotator/Curator/System DB administrator) 向けの教材

「もっとでいーびー」と発音、**Master of the database**【データベースの達人】の略

実践的な講習資料、手順書

新着論文レビュー、領域融合レビュー

<http://first.lifesciencedb.jp/>

次世代シーケンサーによる日本人全ゲノム配列のはじめての包括的な解析 : ライフサイエンス 新着論文レビュー

ホーム 「新着論文レビュー」とは 掲載記事一覧 編集人 著作権・クレジット

ライフサイエンス 新着論文レビュー

FIRSt AUTHOR'S

次世代シーケンサーによる日本人全ゲノム配列のはじめての包括的な解析

2010年11月8日

角田 達彦
(理化学研究所ゲノム医科学研究センター 情報解析研究チーム)
email : 角田達彦

Whole-genome sequencing of Japanese individuals reveals massive parallelism in the variation of individual genomes. Akihiro Fujimoto, H. Nagasaki, Rui Yamamoto. Nature Genetics, 42, 1021-1025 (2010).

野外の環境における生物の環境応答の理解にむけて : トランск립トームデータと気象データの統合 : ライフサイエンス 領域融合レビュー

ホーム 「領域融合レビュー」とは 掲載記事一覧 著作権・クレジット 引用にあたって 編集人

LEADING AUTHOR'S

野外の環境における生物の環境応答の理解にむけて : トランスク립トームデータと気象データの統合

2014年7月22日

永野 悟・工藤 洋
(京都大学生態学研究センター)
email : 永野 悟

領域融合レビュー, 3, e009 (2014) DOI: 10.7875/leading.author.3.e009
Atsushi J. Nagano & Hiroshi Kudoh: Integration of transcriptomics and meteorology.

PDF Download

目 次

DBCLS Database Center for Life Science

このサイトについて

3文字以上入力すると補完されます。

このサイトについて

生命科学において注目される分野・学問領域における最新の研究成果について、第一線の研究者の執筆による日本語のレビューを、だれでも自由に閲覧・利用できるよう、無料で公開します。くわしくは、「領域融合レビュー」とはをご覧ください。

統合TVにて解説動画を公開しました。

<http://leading.lifesciencedb.jp/>

トップジャーナルに掲載された**日本人を著者とする**生命科学分野の論文について、**論文の著者自身の執筆による日本語のレビュー**を、だれでも自由に閲覧・利用できるサービス

生命科学において注目される分野・学問領域における最新の研究成果について、**第一線の研究者の執筆による日本語のレビューを、だれでも自由に閲覧・利用できるサービス**

Allie: 生命科学分野の略語／展開形検索

<http://allie.dbcls.jp/ja/>

The screenshot shows the Allie search interface. At the top, there's a navigation bar with icons for back, forward, and search, followed by the URL 'allie.dbcls.jp'. Below the bar, the title 'Allie (アリー): 生命科学分野の略語／展開形検索' is displayed. A large red 'Allie' logo is on the left. The main content area has a dark header 'Allie A Search Service for Abbreviation / Long Form' and a sub-header '■ 略語／展開形の検索'. It includes a search input field, a '検索' button, and a 'リセット' button. Below the search area, there are links for 'Allie Data Portal', 'SOAP/REST APIについて', and 'English'. A sidebar on the left contains sections for 'ヘルプ' and '略語一覧', with buttons for 'SPFで試す' and '検索オプション'. The main body of the page contains a section titled 'Allieとは?' with a detailed explanation of what Allie is and how it works.

The screenshot shows the search results for the term 'NGS'. The title is 'NGS - Allie (アリー): 略語からの検索結果'. The search conditions are set to '検索語: NGS' and '検索方法: 完全一致'. The results table has columns for 展開形 No., 展開形, 分野, 共起略語, and PubMed/MEDLINE情報 (発表年, 領域). The results are:

展開形 No.	展開形	分野	共起略語	PubMed/MEDLINE情報 (発表年, 領域)
1	next-generation sequencing (DNA集団全体の配列を短時間で決定する新技術) 次世代シーケンシング (1170回)	Molecular Biology 分子生物学 (170回)	SNPs (34回) miRNAs (28回) SNP (26回)	2008 Next-generation sequencing: the race is on.
2	next-generation sequencing technologies 次世代シーケンス技術 (16回)	Medicine 医学, 医薬 (4回)	AC (1回) ARNSHL (1回) CLIA (1回)	2010 Whole methylome analysis by ultra-deep sequencing using two-base encoding.
3	natural guide star (8回)	Ophthalmology 眼科学 (8回)	MCAO (4回) LGS (3回) WFS (3回)	2001 Methods for correcting tilt anisoplanatism in laser-guide-star-based multiconjugate adaptive optics.
4	normal goat serum (7回)	Allergy and Immunology アレルギー・免疫 (3回)	B cell (1回) BDNF (1回) CDV (1回)	1979 In vivo effects of antiserum to IgD on surface immunoglobulins, serum immunoglobulins and lymphocyte blastogenesis in rhesus monkeys.

MEDLINE（米国立医学図書館が開発・維持している生物医学文献書誌情報データベース）を利用して、文献中に登場する**略語**とその正式名称の組およびその付随情報を検索できる

ある略語が最初に文献に登場した時期を知る

新たな略語をつくろうとしたとき、その略語がすでに存在しているかどうかを調べる

inMeXes: 逐次PubMed/MEDLINE表現検索

<http://docman.dbcls.jp/im/>

The screenshot shows the inMeXes search interface. At the top, it says "Incremental PubMed/MEDLINE Expression Search: inMeXes (インメクゼス)". Below that is a search bar with the query "be associated". The results section contains a list of terms and their associated categories, such as "Neoplasms / Neurology / Medicine" and "Biochemistry / Neoplasms / Neurology". There are also links for "LSD共起リスト" and "Permalink(?)". A QR code and W3C 1.0 validation icons are visible at the bottom.

生命科学系の文献
(PubMed/MEDLINEに含
まれるタイトルとアブスト
ラクト) に頻出する
英語表現を、1文字の入力
ごとに高速に検索

検索結果から用例や
関連情報を容易に取得する
ことができる

This screenshot shows the same inMeXes interface as the first one, but with a different search term: "be associated". The results list is much longer and includes many entries related to medical specialties like "Neurology", "Medicine", and "General Surgery". The interface includes standard search controls and a footer with DBCLS and W3C validation information.

RefEx

<http://refex.dbcls.jp>

The screenshot shows the RefEx website interface. At the top, there's a logo and a search bar with placeholder text "ex) troponin, ALB". Below the search bar are sections for "組織特異的に発現する遺伝子を見る" (Organ-specific expression genes) and "遺伝子オントロジー" (Gene Ontology). The "遺伝子オントロジー" section lists categories like cellular process, biological regulation, metabolic process, multicellular organismal process, response to stimulus, and developmental process. To the right, there are two heatmaps titled "ヒートマップ on Bodyparts3D" and "組織40分類別データ". The first heatmap shows a 3D human body model with colored organs representing gene expression levels. The second heatmap displays four panels for EST, GeneChip, CAGE, and RNA-seq across 40 tissue categories, with a color scale from blue (low) to red (high).

4つの異なる実験手法 (EST、GeneChip、CAGE、RNA-seq) によって得られた**40種類の正常組織における遺伝子発現データを直感的に比較**することが可能なリファレンスデータセット

遺伝子発現解析などで見出された不詳な遺伝子群の関係性を知るためのツールとしても有用

GGRNA、GGGenome

<http://ggrna.dbcls.jp/ja/>

<http://gggenome.dbcls.jp/>

The screenshot shows the GGRNA search interface. At the top, there's a navigation bar with links for Help, Advanced search, English, and Old version. Below it is the GGRNA logo and a search bar. A sidebar on the left provides search examples like 'homeobox' and 'RNA interference'. At the bottom, there's a news section with links to recent updates.

The screenshot shows the GGGenome search interface. It has a similar layout to GGRNA, with a navigation bar, logo, and search bar. The sidebar contains search examples and a news section with links to recent updates. The main content area shows a search result for 'Human genome, GRCh37/hg19 (Feb, 2009)'.

【GGRNA】 NCBI RefSeqに登録された転写産物の情報高速全文検索

遺伝子名や各種ID、タンパクの機能など、どのようなキーワードからでも簡単に遺伝子を検索

【GGGenome】 ゲノムや転写産物から塩基配列を高速に検索

BLASTやBLATでは検索の難しい20塩基以下の短い配列の検索が得意

ミスマッチやギャップを含む塩基配列にも対応

テキスト比較ツール diff《デュフ》

<http://diff.jp>

The screenshot shows the diffff.jp web interface. At the top, there are two text input boxes. The left box contains the original text:

```
Betty Botter bought some butter,  
But, she said, this butter's bitter;  
If I put it in my batter,  
It will make my batter bitter,  
But a bit of better butter  
Will make my batter better.  
So she bought a bit of butter  
Better than her bitter butter,  
And she put it in her batter,  
And it made her batter better,  
So 'twas better Betty Botter  
Bought a bit of better butter.
```

The right box contains the modified text:

```
Betty Botter bought some butter,  
But, she said, the butter's bitter;  
If I put it in my batter,  
That will make my batter bitter.  
But a bit of better butter,  
That will make my batter better.  
So she bought a bit of butter  
Better than her bitter butter.  
And she put it in her batter,  
And it made her batter better.  
So it was better Betty Botter  
Bought a bit of better butter.
```

Below the text boxes is a button labeled "比較する" (Compare). Underneath each text box is a "diff" output box. The left "diff" box shows the differences between the two versions:

```
下記の文章を比較してください。  
Betty Botter bought some butter,  
But, she said, this butter's bitter;  
If I put it in my batter,  
It will make my batter bitter,  
But a bit of better butter  
Will make my batter better.  
So she bought a bit of butter  
Better than her bitter butter,  
And she put it in her batter,  
And it made her batter better,  
So ''twas better Betty Botter  
Bought a bit of better butter.
```

The right "diff" box shows the differences between the two versions:

```
下記の文章を比較してください。  
Betty Botter bought some butter,  
But, she said, the butter's bitter;  
If I put it in my batter,  
That will make my batter bitter.  
But a bit of better butter,  
That will make my batter better.  
So she bought a bit of butter  
Better than her bitter butter,  
And she put it in her batter,  
And it made her batter better,  
So it was better Betty Botter  
Bought a bit of better butter.
```

At the bottom of the page are several buttons: "結果のみ表示 (印刷用)" (Printable), "カラー1" (Color 1) (selected), "カラー2" (Color 2), and "モノクロ" (Monochrome).

diffコマンドを使った比較結果をWeb上に表示するツール

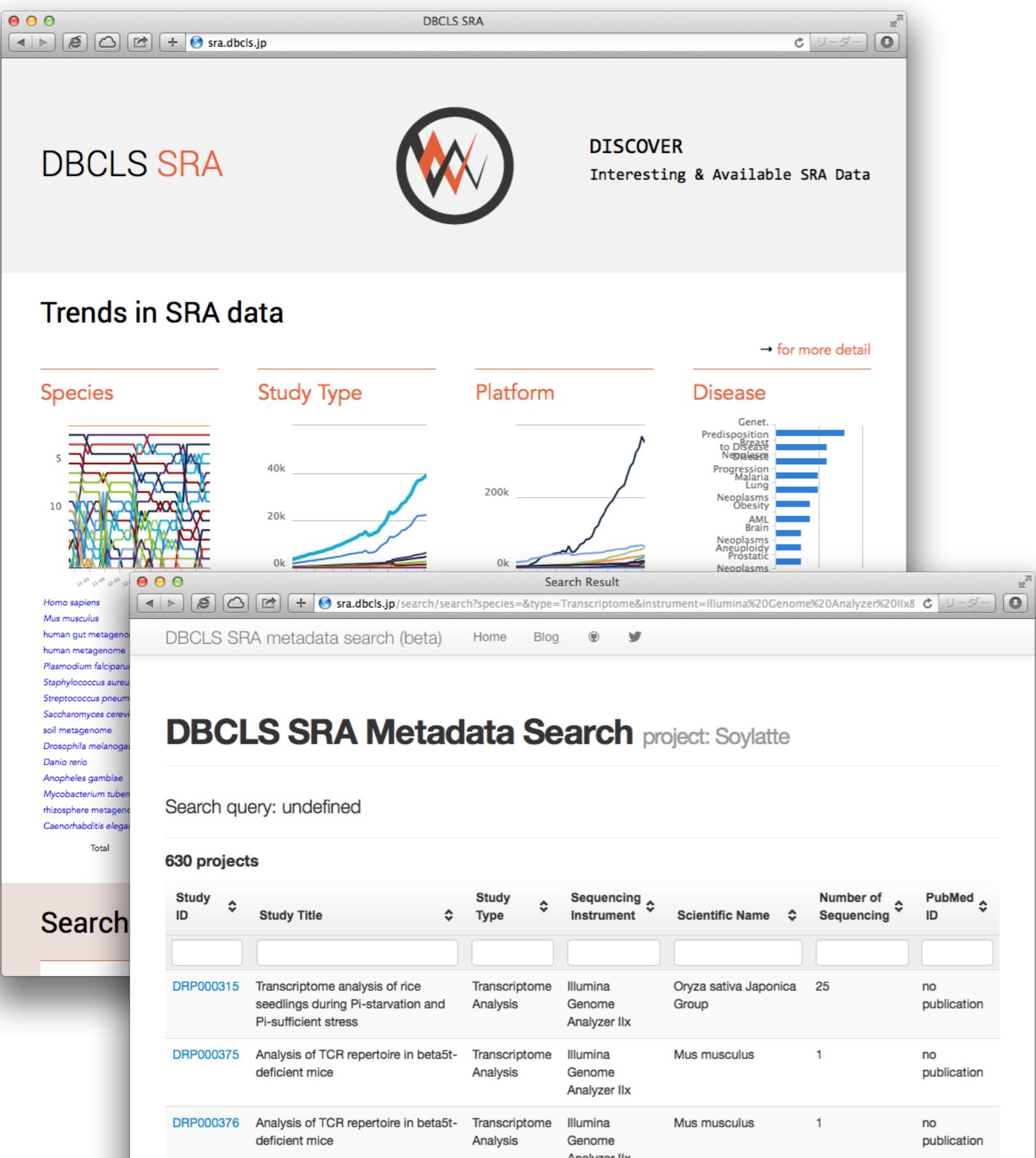
テキストボックスに比較したい文章をコピペしてボタンを押すだけで**2つの文章でどこが変更されたのか、差分の確認ができる**

日本語のテキストも対応

Word文章はもちろんソースコード、遺伝子リストなどの比較も可能

DBCLS SRA

<http://sra.dbcls.jp>



公共データベース (SRA [NCBI], ENA [EBI], DRA [DDBJ]) に登録された
「次世代シーケンサ」データについて、さまざまな
統計情報から閲覧、比較、
データのダウンロードができる目次サイト

複数に分かれた登録でも同じ
プロジェクトの実験ならば一
度に表示

目的別、機器別、生物種別に
登録データを検索可能

ライフサイエンスQA

<http://qa.lifescienceedb.jp/>

The screenshot shows the homepage and a detailed view of the website.

Homepage: The top navigation bar includes 'ログイン', '概要', and 'よくある質問'. Below it is a search bar and a '質問する' button. The main area displays a list of questions under 'すべての質問':

- アライメントについて (3回答, 714閲覧)
- データベースにあるデータの再解析 (1回答, 950閲覧)
- Togo picture gallery リクエスト (1回答, 737閲覧)
- 統合TV番組リクエスト (4回答, 2.2k閲覧)

Detailed View: This view is for the question 'RNA-seq データの定量と可視化の、ツールの組み合わせ' (3回答, 714閲覧). It includes:

- A sidebar with a message: 'ただいまベータテスト中です。そのため通知無く停止更新されることがあります。質問するならライフサイエンスQA'.
- A '概要' section with a 'よくある質問' link.
- A right sidebar for following the question via email or RSS, and a note about beta testing.
- A 'タグ' section listing: bowtie (x6), cufflinks (x2), express (x1), rna-star (x1), rpkm (x1).
- Information at the bottom: 質問日 Apr 10 at 19:06, 閲覧数: 709回, 最終更新日: Apr 16 at 05:06.

One Answer: A single answer is shown, starting with: 'いくつか私の誤解があるかもしれない前提で回答します。 (つまり鵜呑みにされても責任は持てませんという前提です。私は一介の技術者で研究者ではありません。) 結論から言えば'.

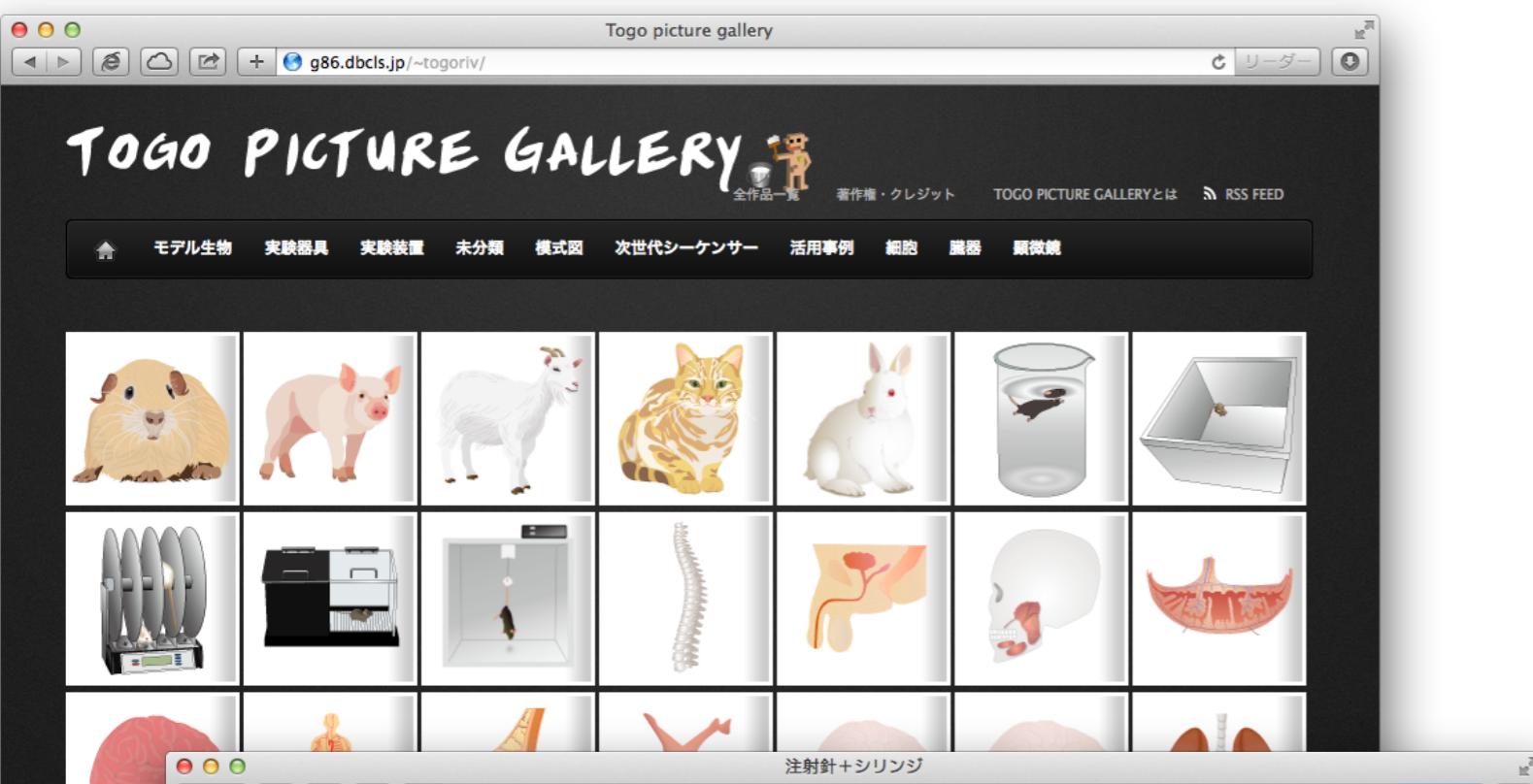
ライフサイエンス研究にまつわる**疑問や質問**、それへの**回答を共有**するウェブサイト

- ・ ライフサイエンス研究
- ・ バイオインフォマティクス研究
- ・ DDBJのサービス
- ・ DBCLSのサービス

対象ユーザーは、
ライフサイエンス研究に
携わっている人を想定

togo picture gallery

<http://g86.dbcls.jp/togopic/>



A screenshot of a specific image page from the Togo Picture Gallery. The main image is a detailed illustration of a medical syringe with a needle. To the right of the image is descriptive text: "注射針+シリンジ" (Needle + Syringe), "シリンジ+注射針" (Syringe + Needle), and instructions for use. Below this is a section about licensing: "Togo picture gallery by DBCLS is Licensed under a Creative Commons 表示 2.1 日本 (c)" and "あるいは" (Or) followed by "Togo picture gallery by DBCLS is licensed under a Creative Commons Attribution 2.1 Japan license (c)". At the bottom left is the Creative Commons BY logo, and at the bottom right is the Japanese Creative Commons license text: "この画像データは、クリエイティブ・コモンズ (CC) 表示 2.1 日本 ライセンス のもとに利用を許諾されています。" (This image data is licensed under the Creative Commons Attribution 2.1 Japan license). The footer also includes "帰属: Database Center for Life Science (DBCLS)" and "あなたは以下の条件に従う場合に限り、自由に" (Under the following conditions, you may use it freely).

だれでも自由に閲覧・利用できるように、Web上にて無料で公開している**ライフサイエンス分野のイラスト集**

モデル生物、実験器具、実験装置、模式図、次世代シーケンサー、細胞、顕微鏡があり、幅広い資料に利用可能

Creative Commons ライセンスで提供されているので、クレジットを記載するだけで、どのように利用してもOK