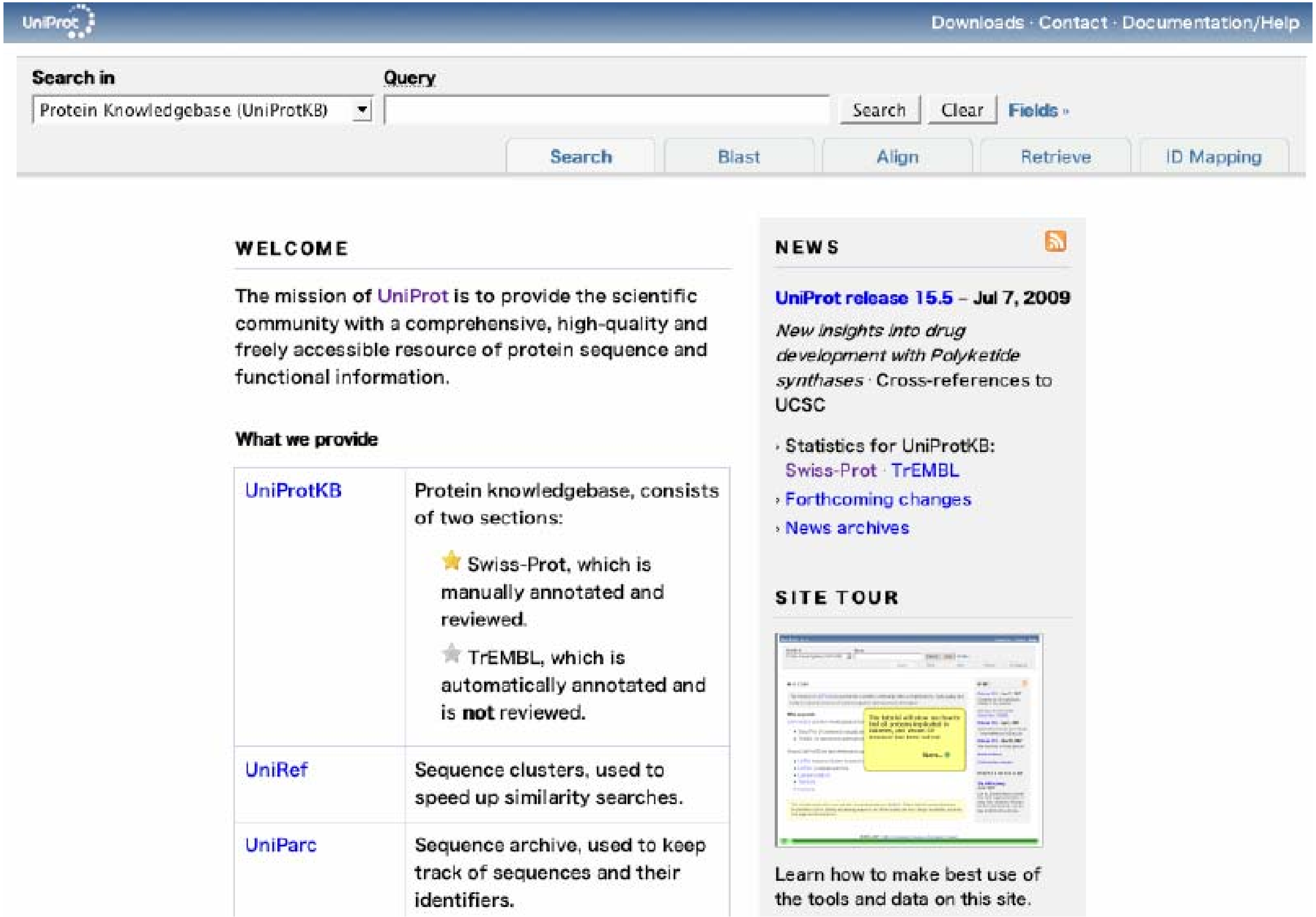
タンパク質配列のデータベース

UniProt　http://www.uniprot.org/

InterPro　http://www.ebi.ac.uk/interpro/

UniProtのトップページ



WELCOME

The mission of UniProt is to provide the scientific community with a comprehensive, high-quality and freely accessible resource of protein sequence and functional information.

What we provide

UniProtKB	Protein knowledgebase, consists of two sections: ★ Swiss-Prot, which is manually annotated and reviewed. ☆ TrEMBL, which is automatically annotated and is **not** reviewed.
UniRef	Sequence clusters, used to speed up similarity searches.
UniParc	Sequence archive, used to keep track of sequences and their identifiers.

NEWS

UniProt release 15.5 – Jul 7, 2009

New insights into drug development with Polyketide synthases · Cross-references to UCSC

- Statistics for UniProtKB: Swiss-Prot · TrEMBL
- Forthcoming changes
- News archives

SITE TOUR

Learn how to make best use of the tools and data on this site.

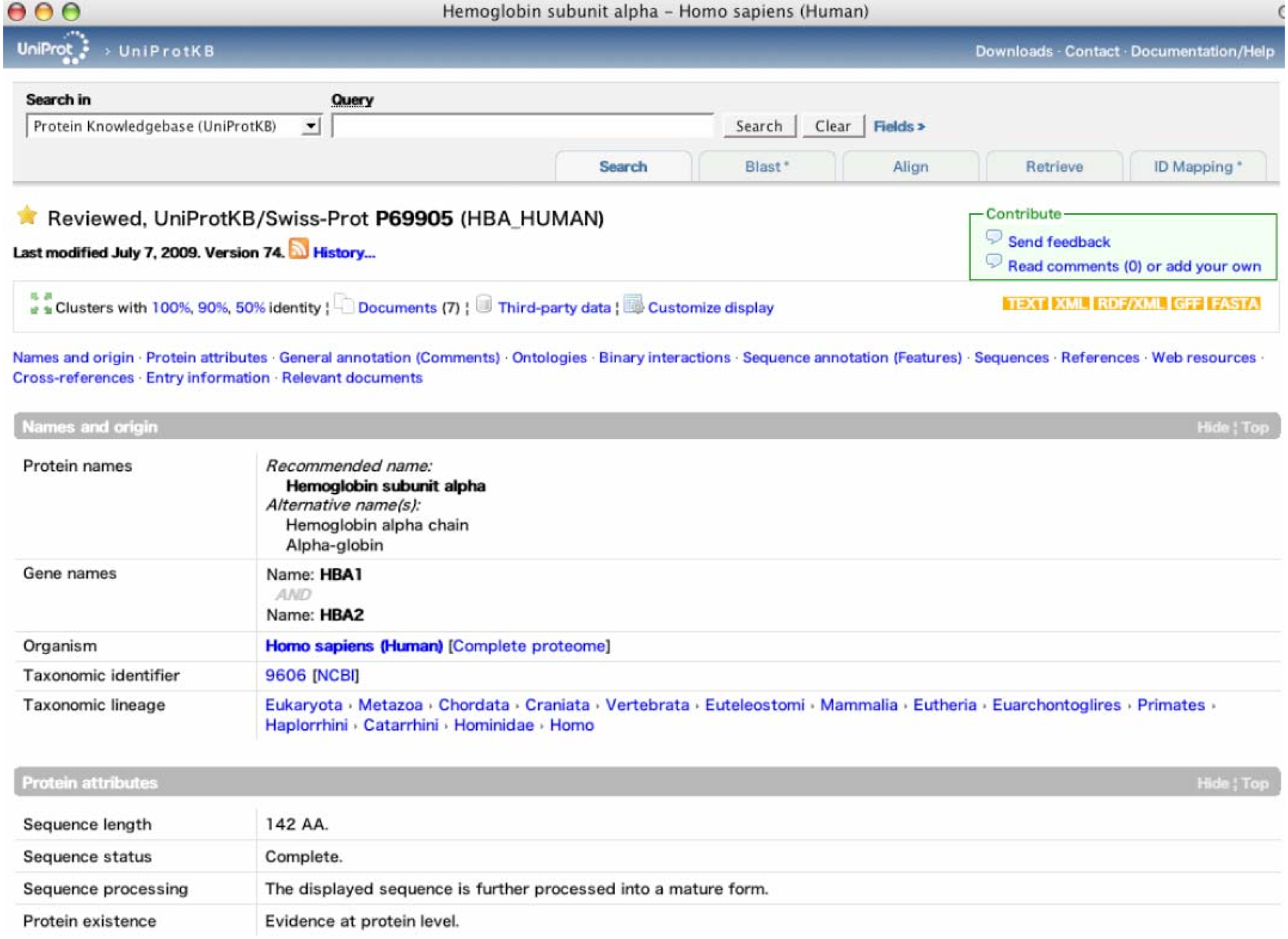
UniProt

タンパク質の配列と機能に関する網羅的で高精度の情報を、無料で提供するデータベース。3つのデータベースで構成されている。

- UniProtKB (UniProt Knowledgebase)
  - Swiss-Prot: マニュアル（手動）でアノテーションを行い、レビュー(チェック＋修正)されたデータ。
  - TrEMBL: 計算機を使って自動でアノテーションされたデータ。レビューを受けていない。
- UniRef

  配列のクラスター。相同性検索の対象データベース等に用いるために、膨大な量のデータを圧縮してある。
- UniParc

  配列データとデータIDのアーカイブ。



Reviewed, UniProtKB/Swiss-Prot **P69905** (HBA_HUMAN)

Last modified July 7, 2009. Version 74. **History...**

Clusters with 100%, 90%, 50% identity ¦ Documents (7) ¦ Third-party data ¦ Customize display

Names and origin · Protein attributes · General annotation (Comments) · Ontologies · Binary interactions · Sequence annotation (Features) · Sequences · References · Web resources · Cross-references · Entry information · Relevant documents

Names and origin Hide ¦ Top

Protein names	*Recommended name:* **Hemoglobin subunit alpha** *Alternative name(s):* Hemoglobin alpha chain Alpha-globin
Gene names	Name: **HBA1** *AND* Name: **HBA2**
Organism	**Homo sapiens (Human)** [Complete proteome]
Taxonomic identifier	9606 [NCBI]
Taxonomic lineage	Eukaryota › Metazoa › Chordata › Craniata › Vertebrata › Euteleostomi › Mammalia › Eutheria › Euarchontoglires › Primates › Haplorrhini › Catarrhini › Hominidae › Homo

Protein attributes Hide ¦ Top

Sequence length	142 AA.
Sequence status	Complete.
Sequence processing	The displayed sequence is further processed into a mature form.
Protein existence	Evidence at protein level.

UniProtのデータ例（ヒトのヘモグロビンA）

```
ID   HBA_HUMAN          Reviewed;       142 AA.
AC   P69905; P01922; Q1HDT5; Q3MIF5; Q53F97; Q96KF1; Q9NYR7; Q9UCM0;
DT   21-JUL-1986, integrated into UniProtKB/Swiss-Prot.
DT   23-JAN-2007, sequence version 2.
DT   07-JUL-2009, entry version 74.
DE   RecName: Full=Hemoglobin subunit alpha;
DE   AltName: Full=Hemoglobin alpha chain;
DE   AltName: Full=Alpha-globin;
GN   Name=HBA1;
GN   and
GN   Name=HBA2;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;
OC   Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [GENOMIC DNA] (HBA1).
RX   MEDLINE=81088339; PubMed=7448866; DOI=10.1016/0092-8674(80)90347-5;
RA   Michelson A.M., Orkin S.H.;
RT   "The 3' untranslated regions of the duplicated human alpha-globin
RT   genes are unexpectedly divergent.";
RL   Cell 22:371-377(1980).

..
..

SQ   SEQUENCE   142 AA;  15258 MW;  15E13666573BBBAE CRC64;
     MVLSPADKTN VKAAWGKVGA HAGEYGAEAL ERMFLSFPTT KTYFPHFDLS HGSAQVKGHG
     KKVADALTNA VAHVDDMPNA LSALSDLHAH KLRVDPVNFK LLSHCLLVTL AAHLPAEFTP
     AVHASLDKFL ASVSTVLTSK YR
//
```

アクセッション番号は、
統合などのため別名が複数ある

UniProtのデータ構造

ID	データ識別名
AC	アクセッション番号
DT	公開日
DE	タンパク質名
GN	遺伝子名
OC	生物種の分類
R*	引用文献（reference）
CC	コメント
DR	外部データベースの対応データ
KW	キーワード
FT	タンパク質の特徴（feature）
SQ	アミノ酸配列
//	

InterPro

- InterProは、(1) タンパク質のファミリー分類、(2) ドメイン、(3) リピート、(4) 機能サイト、に関する記述を集めた統合データベースである。
- タンパク質の特徴を集めた複数のデータベースから、ひとつのデータベースを構築した。
- 合計18,843のエントリー(Release 22.0)

InterPro; http://www.ebi.ac.uk/interpro/

InterProのメンバーデータベース（1）

The UniProt (Universal Protein Resource) is the world's most comprehensive catalog of information on proteins. It is a central repository of protein sequence and function created by joining the information contained in Swiss-Prot, TrEMBL, and PIR.

PROSITE is a database of protein families and domains. It consists of biologically significant sites, patterns and profiles that help to reliably identify to which known protein family (if any) a new sequence belongs.

HAMAP, stands for High-quality Automated and Manual Annotation of microbial Proteomes. HAMAP profiles are manually created by expert curators they identify proteins that are part of well-conserved bacterial, archaeal and plastid-encoded (i.e. chloroplasts, cyanelles, apicoplasts, non-photosynthetic plastids) proteins families or subfamilies.

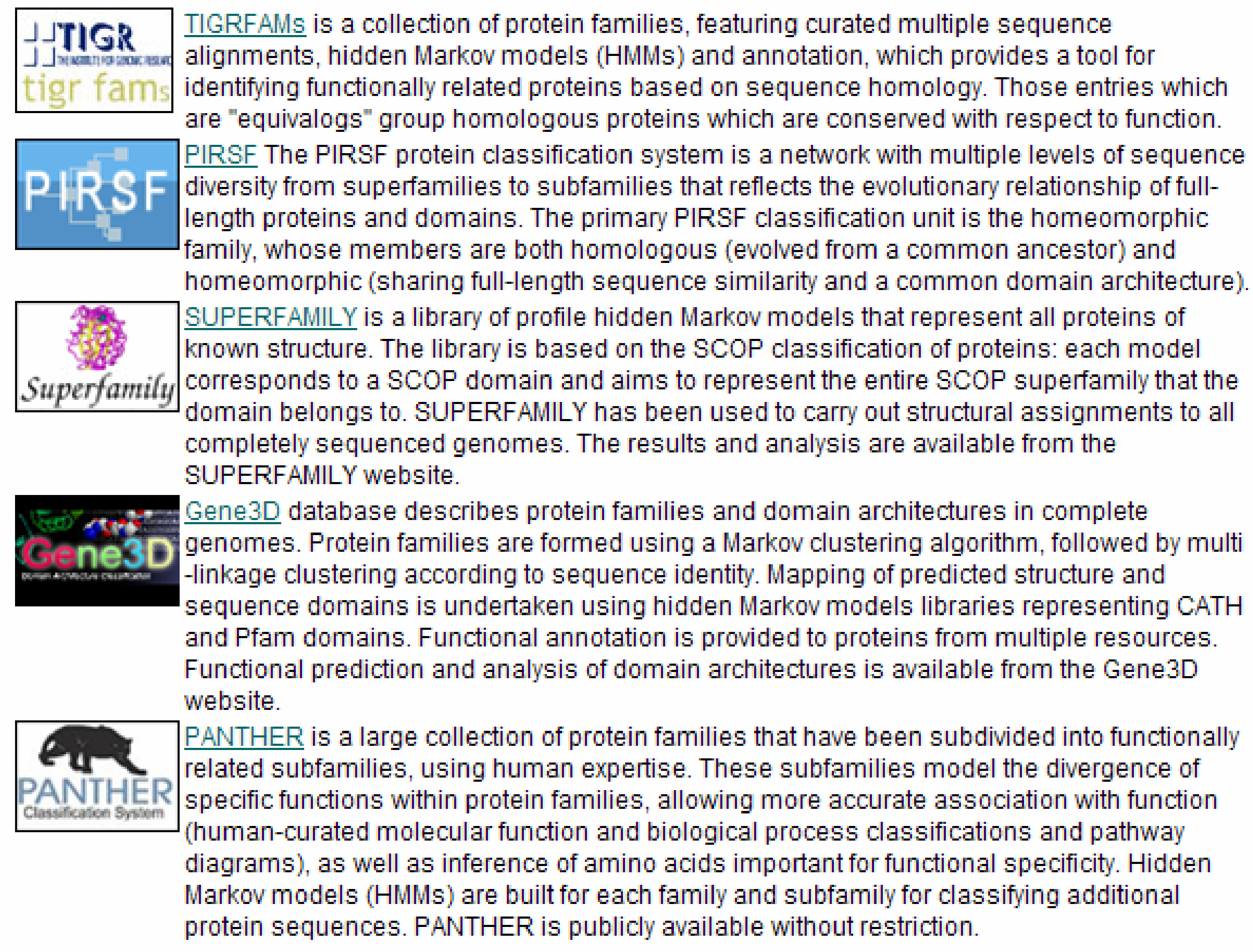
Pfam is a large collection of multiple sequence alignments and hidden Markov models covering many common protein domains.

PRINTS is a compendium of protein fingerprints. A fingerprint is a group of conserved motifs used to characterise a protein family; its diagnostic power is refined by iterative scanning of UniProt. Usually the motifs do not overlap, but are separated along a sequence, though they may be contiguous in 3D-space. Fingerprints can encode protein folds and functionalities more flexibly and powerfully than can single motifs, their full diagnostic potency deriving from the mutual context afforded by motif neighbours.

The ProDom protein domain database consists of an automatic compilation of homologous domains. Current versions of ProDom are built using a novel procedure based on recursive PSI-BLAST searches (Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W & Lipman DJ, 1997, Nucleic Acids Res., 25:3389-3402; Gouzy J., Corpet F. & Kahn D., 1999, Computers and Chemistry 23:333-340.) Large families are much better processed with this new procedure than with the former DOMAINER program (Sonnhammer, E.L.L. & Kahn, D., 1994, Protein Sci., 3:482-492).

SMART (a Simple Modular Architecture Research Tool) allows the identification and annotation of genetically mobile domains and the analysis of domain architectures. More than 500 domain families found in signalling, extracellular and chromatin-associated proteins are detectable. These domains are extensively annotated with respect to phyletic distributions, functional class, tertiary structures and functionally important residues. Each domain found in a non-redundant protein database as well as search parameters and taxonomic information are stored in a relational database system. User interfaces to this database allow searches for proteins containing specific combinations of domains in defined taxa.

InterProのメンバーデータベース（2）



TIGRFAMs is a collection of protein families, featuring curated multiple sequence alignments, hidden Markov models (HMMs) and annotation, which provides a tool for identifying functionally related proteins based on sequence homology. Those entries which are "equivalogs" group homologous proteins which are conserved with respect to function.

PIRSF The PIRSF protein classification system is a network with multiple levels of sequence diversity from superfamilies to subfamilies that reflects the evolutionary relationship of full-length proteins and domains. The primary PIRSF classification unit is the homeomorphic family, whose members are both homologous (evolved from a common ancestor) and homeomorphic (sharing full-length sequence similarity and a common domain architecture).

SUPERFAMILY is a library of profile hidden Markov models that represent all proteins of known structure. The library is based on the SCOP classification of proteins: each model corresponds to a SCOP domain and aims to represent the entire SCOP superfamily that the domain belongs to. SUPERFAMILY has been used to carry out structural assignments to all completely sequenced genomes. The results and analysis are available from the SUPERFAMILY website.

Gene3D database describes protein families and domain architectures in complete genomes. Protein families are formed using a Markov clustering algorithm, followed by multi-linkage clustering according to sequence identity. Mapping of predicted structure and sequence domains is undertaken using hidden Markov models libraries representing CATH and Pfam domains. Functional annotation is provided to proteins from multiple resources. Functional prediction and analysis of domain architectures is available from the Gene3D website.

PANTHER is a large collection of protein families that have been subdivided into functionally related subfamilies, using human expertise. These subfamilies model the divergence of specific functions within protein families, allowing more accurate association with function (human-curated molecular function and biological process classifications and pathway diagrams), as well as inference of amino acids important for functional specificity. Hidden Markov models (HMMs) are built for each family and subfamily for classifying additional protein sequences. PANTHER is publicly available without restriction.

InterProを用いた配列解析ツール: InterProScan

InterProScanは、ウェブサーバーで提供されるツールで、InterProに登録されている各種の機能ドメイン・構造ドメインを検索したり、ユーザーが入力した機能未知の配列を解析して特徴を発見することができるツールである。

http://www.ebi.ac.uk/Tools/InterProScan/

InterProScanの使い方

InterProScan Sequence Search

This form allows you to query your sequence against InterPro. For more detailed information see the documentation for the perl stand-alone InterProScan package (Readme file or FAQ's), or the InterPro user manual or help pages.

Please Note: Due to resource limitations the InterProScan service will not accept nucleotide sequence submissions until further notice. Please see the Help for more information.

Download Software

RESULTS	YOUR EMAIL
interactive	

APPLICATIONS TO RUN ○ Clear all ● Check all

☑ BlastProDom ☑ FPrintScan ☑ HMMPIR ☑ HMMPfam ☑ HMMSmart
☑ HMMTigr ☑ ProfileScan ☑ ScanRegExp ☑ SuperFamily ☑ SignalPHMM
☑ TMHMM ☑ HMMPanther ☑ Gene3D

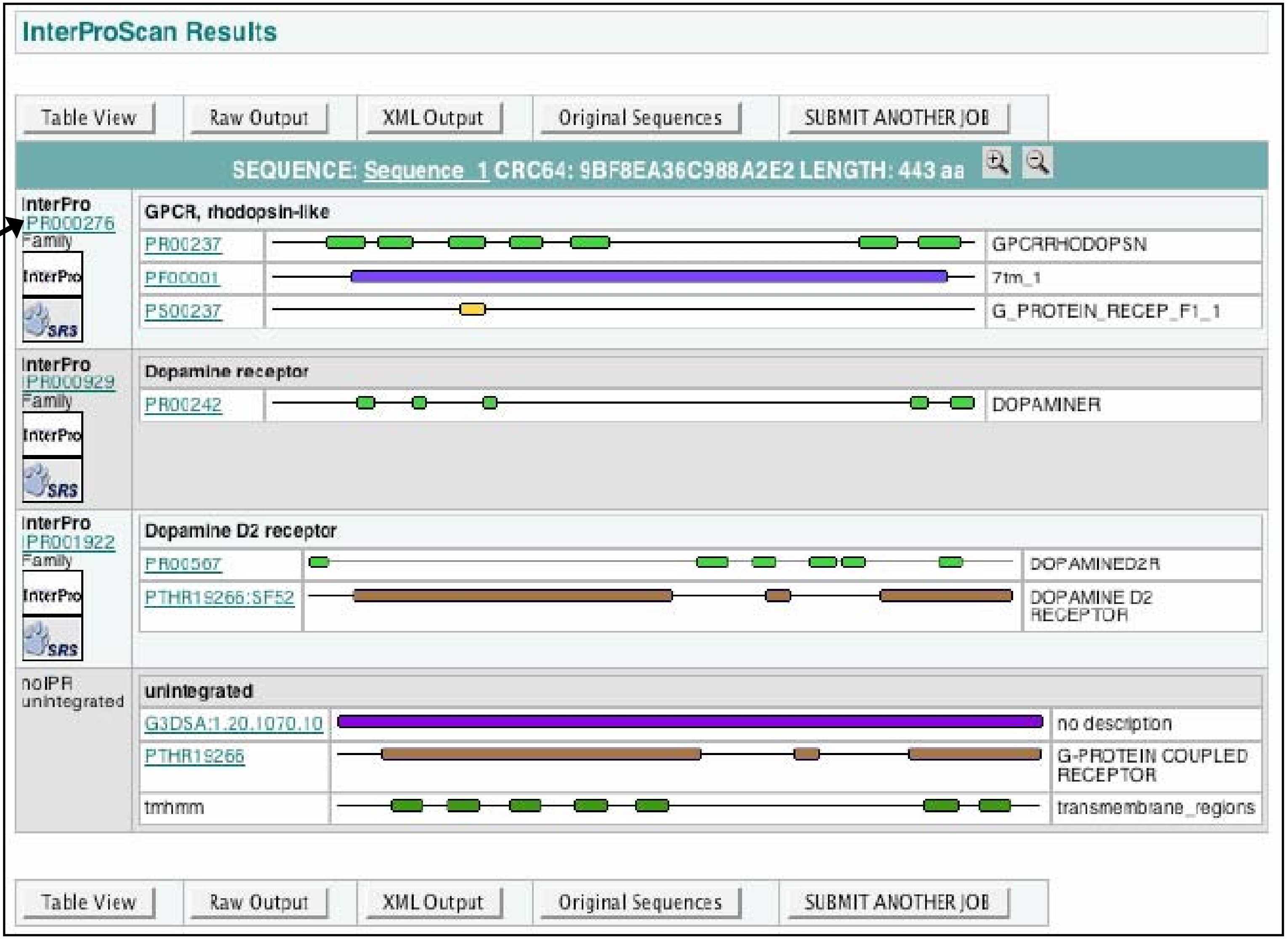
Enter or Paste a PROTEIN Sequence in any format:

PLEASE NOTE: Interactive job results are stored for 24 hours, email job results are stored for one week.
If you plan to use these services during a course please contact us using the email below.

配列を入れる

実行ボタン

結果画面



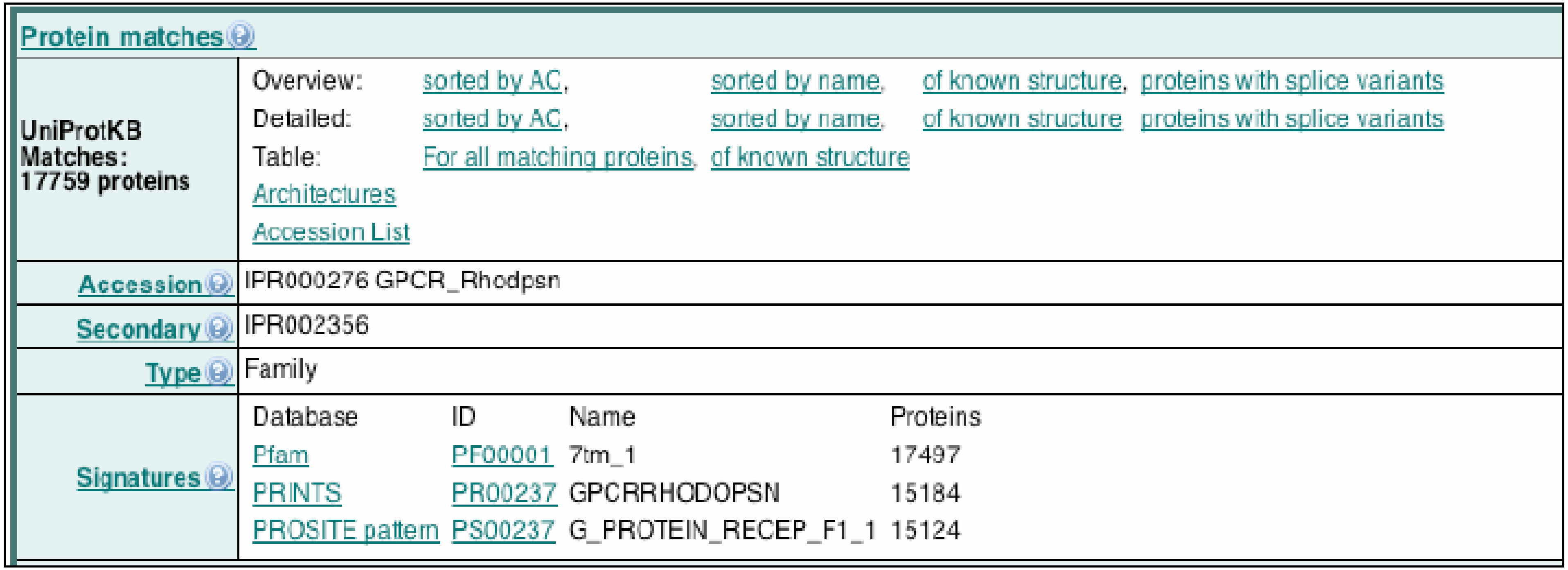
InterProScan Results

SEQUENCE: Sequence_1 CRC64: 9BF8EA36C988A2E2 LENGTH: 443 aa

InterPro IPR000276 Family	GPCR, rhodopsin-like	
	PR00237	GPCRRHODOPSN
	PF00001	7tm_1
	PS50237	G_PROTEIN_RECEP_F1_1
InterPro IPR000929 Family	Dopamine receptor	
	PR00242	DOPAMINER
InterPro IPR001922 Family	Dopamine D2 receptor	
	PR00567	DOPAMINED2R
	PTHR19266:SF52	DOPAMINE D2 RECEPTOR
noIPR unintegrated	unintegrated	
	G3DSA:1.20.1070.10	no description
	PTHR19266	G-PROTEIN COUPLED RECEPTOR
	tmhmm	transmembrane_regions

click here

InterProのアノテーション
（ファミリー）

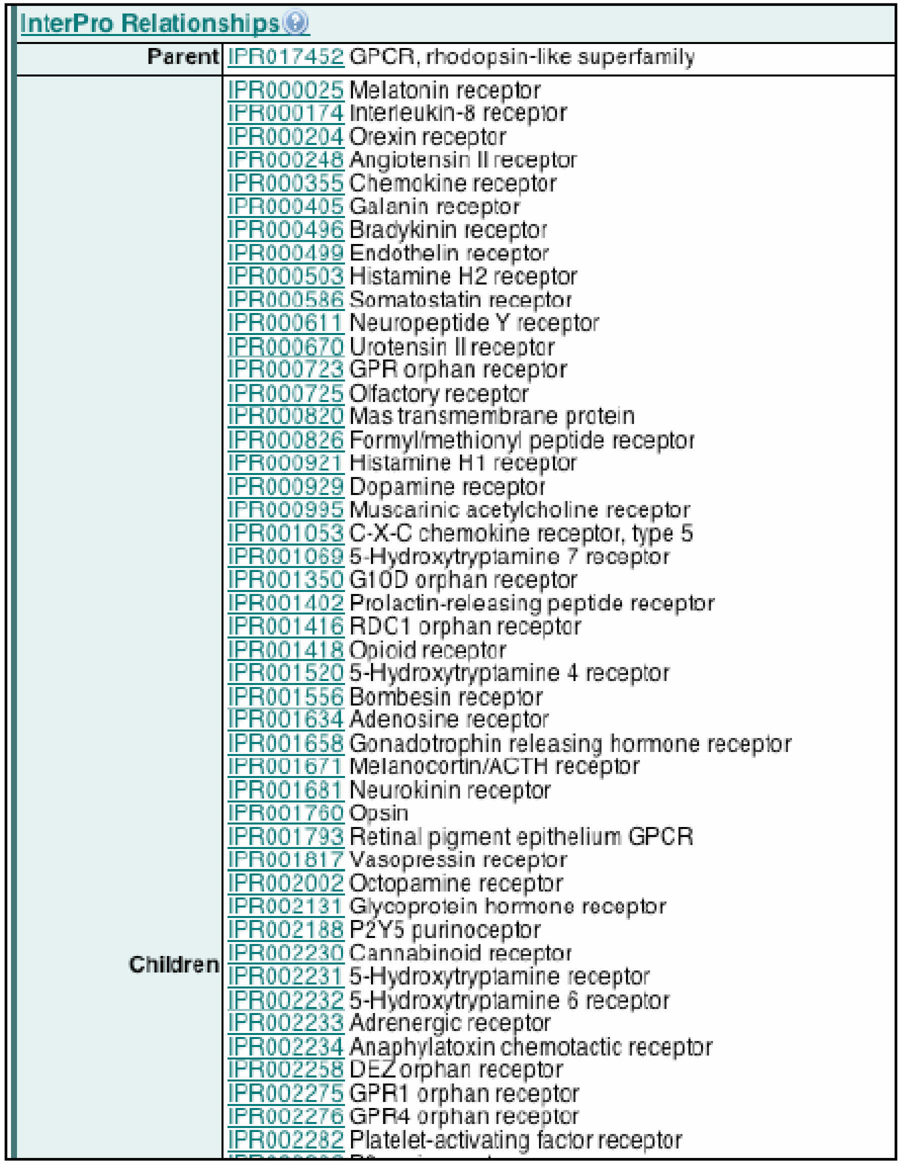


Protein matches

UniProtKB Matches: 17759 proteins	Overview: sorted by AC, sorted by name, of known structure, proteins with splice variants Detailed: sorted by AC, sorted by name, of known structure proteins with splice variants Table: For all matching proteins, of known structure Architectures Accession List
Accession	IPR000276 GPCR_Rhodpsn
Secondary	IPR002356
Type	Family
Signatures	

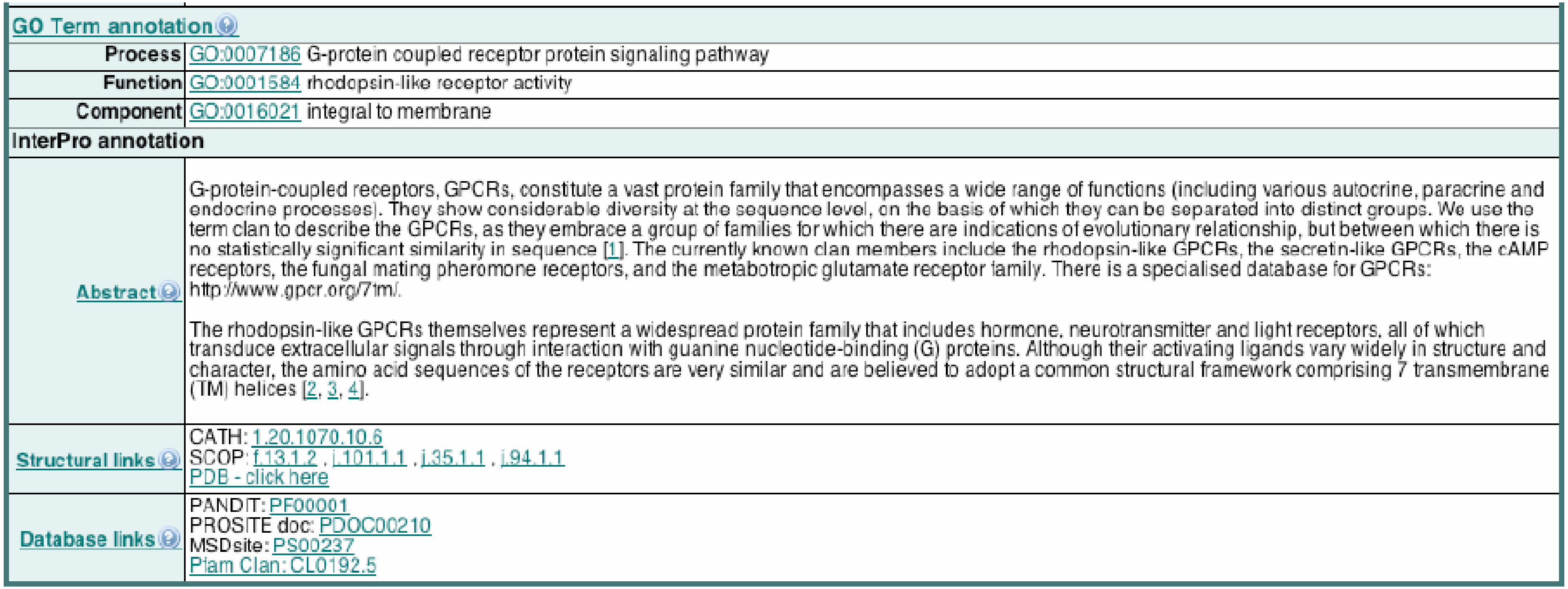
Database	ID	Name	Proteins
Pfam	PF00001	7tm_1	17497
PRINTS	PR00237	GPCRRHODOPSN	15184
PROSITE pattern	PS00237	G_PROTEIN_RECEP_F1_1	15124

InterProのアノテーション
（ファミリー）



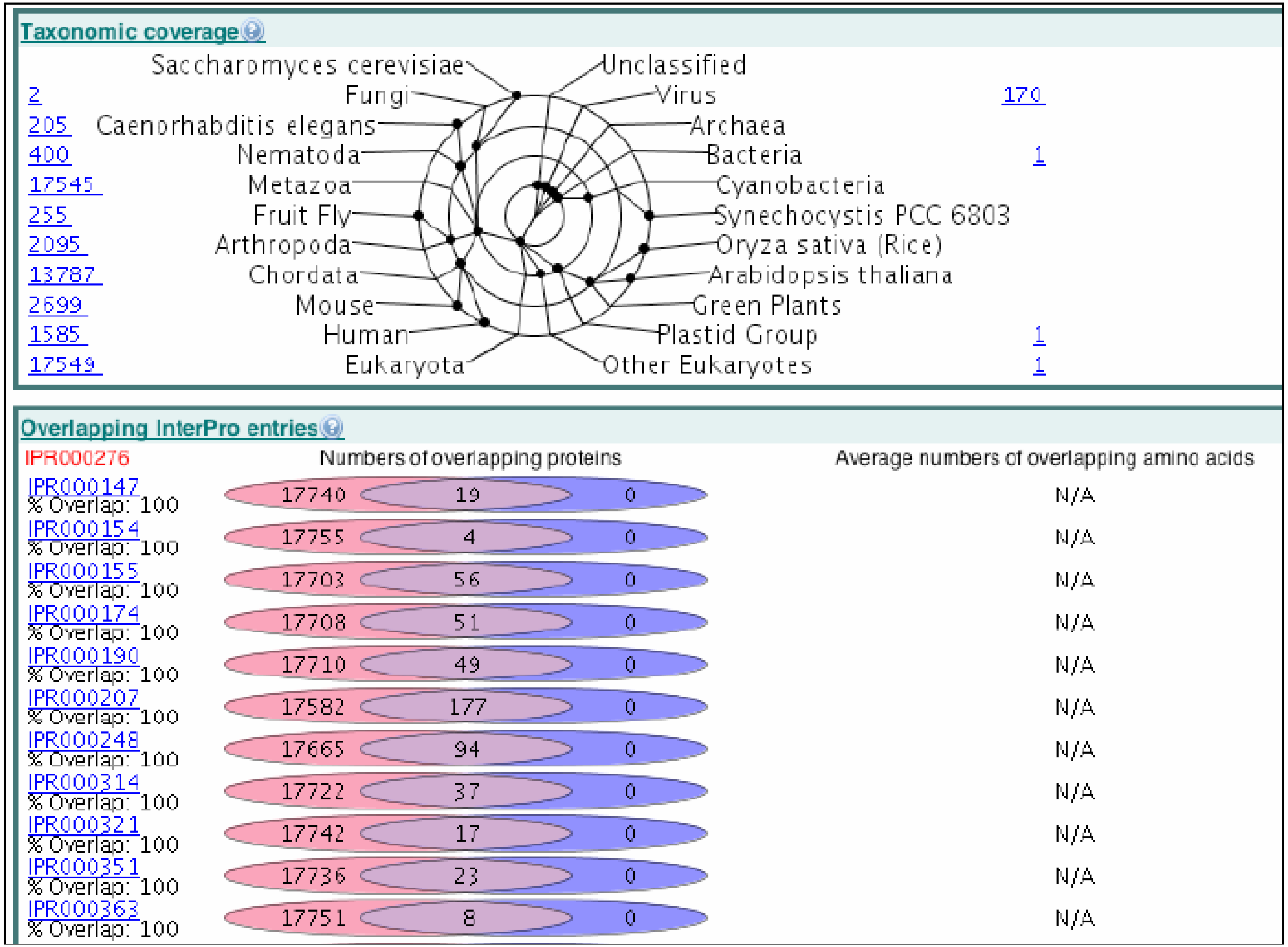
InterPro Relationships	
Parent	IPR017452 GPCR, rhodopsin-like superfamily
Children	IPR000025 Melatonin receptor IPR000174 Interleukin-8 receptor IPR000204 Orexin receptor IPR000248 Angiotensin II receptor IPR000355 Chemokine receptor IPR000405 Galanin receptor IPR000496 Bradykinin receptor IPR000499 Endothelin receptor IPR000503 Histamine H2 receptor IPR000586 Somatostatin receptor IPR000611 Neuropeptide Y receptor IPR000670 Urotensin II receptor IPR000723 GPR orphan receptor IPR000725 Olfactory receptor IPR000820 Mas transmembrane protein IPR000826 Formyl/methionyl peptide receptor IPR000921 Histamine H1 receptor IPR000929 Dopamine receptor IPR000995 Muscarinic acetylcholine receptor IPR001053 C-X-C chemokine receptor, type 5 IPR001069 5-Hydroxytryptamine 7 receptor IPR001350 G10D orphan receptor IPR001402 Prolactin-releasing peptide receptor IPR001416 RDC1 orphan receptor IPR001418 Opioid receptor IPR001520 5-Hydroxytryptamine 4 receptor IPR001556 Bombesin receptor IPR001634 Adenosine receptor IPR001658 Gonadotrophin releasing hormone receptor IPR001671 Melanocortin/ACTH receptor IPR001681 Neurokinin receptor IPR001760 Opsin IPR001793 Retinal pigment epithelium GPCR IPR001817 Vasopressin receptor IPR002002 Octopamine receptor IPR002131 Glycoprotein hormone receptor IPR002188 P2Y5 purinoceptor IPR002230 Cannabinoid receptor IPR002231 5-Hydroxytryptamine receptor IPR002232 5-Hydroxytryptamine 6 receptor IPR002233 Adrenergic receptor IPR002234 Anaphylatoxin chemotactic receptor IPR002258 DEZ orphan receptor IPR002275 GPR1 orphan receptor IPR002276 GPR4 orphan receptor IPR002282 Platelet-activating factor receptor

InterProのアノテーション
（ファミリー）



GO Term annotation	
Process	GO:0007186 G-protein coupled receptor protein signaling pathway
Function	GO:0001584 rhodopsin-like receptor activity
Component	GO:0016021 integral to membrane
InterPro annotation	
Abstract	G-protein-coupled receptors, GPCRs, constitute a vast protein family that encompasses a wide range of functions (including various autocrine, paracrine and endocrine processes). They show considerable diversity at the sequence level, on the basis of which they can be separated into distinct groups. We use the term clan to describe the GPCRs, as they embrace a group of families for which there are indications of evolutionary relationship, but between which there is no statistically significant similarity in sequence [1]. The currently known clan members include the rhodopsin-like GPCRs, the secretin-like GPCRs, the cAMP receptors, the fungal mating pheromone receptors, and the metabotropic glutamate receptor family. There is a specialised database for GPCRs: http://www.gpcr.org/7tm/. The rhodopsin-like GPCRs themselves represent a widespread protein family that includes hormone, neurotransmitter and light receptors, all of which transduce extracellular signals through interaction with guanine nucleotide-binding (G) proteins. Although their activating ligands vary widely in structure and character, the amino acid sequences of the receptors are very similar and are believed to adopt a common structural framework comprising 7 transmembrane (TM) helices [2, 3, 4].
Structural links	CATH: 1.20.1070.10.6 SCOP: f.13.1.2 , i.101.1.1 , i.35.1.1 , i.94.1.1 PDB - click here
Database links	PANDIT: PF00001 PROSITE doc: PDOC00210 MSDsite: PS00237 Pfam Clan: CL0192.5

InterProのアノテーション
（ファミリー）



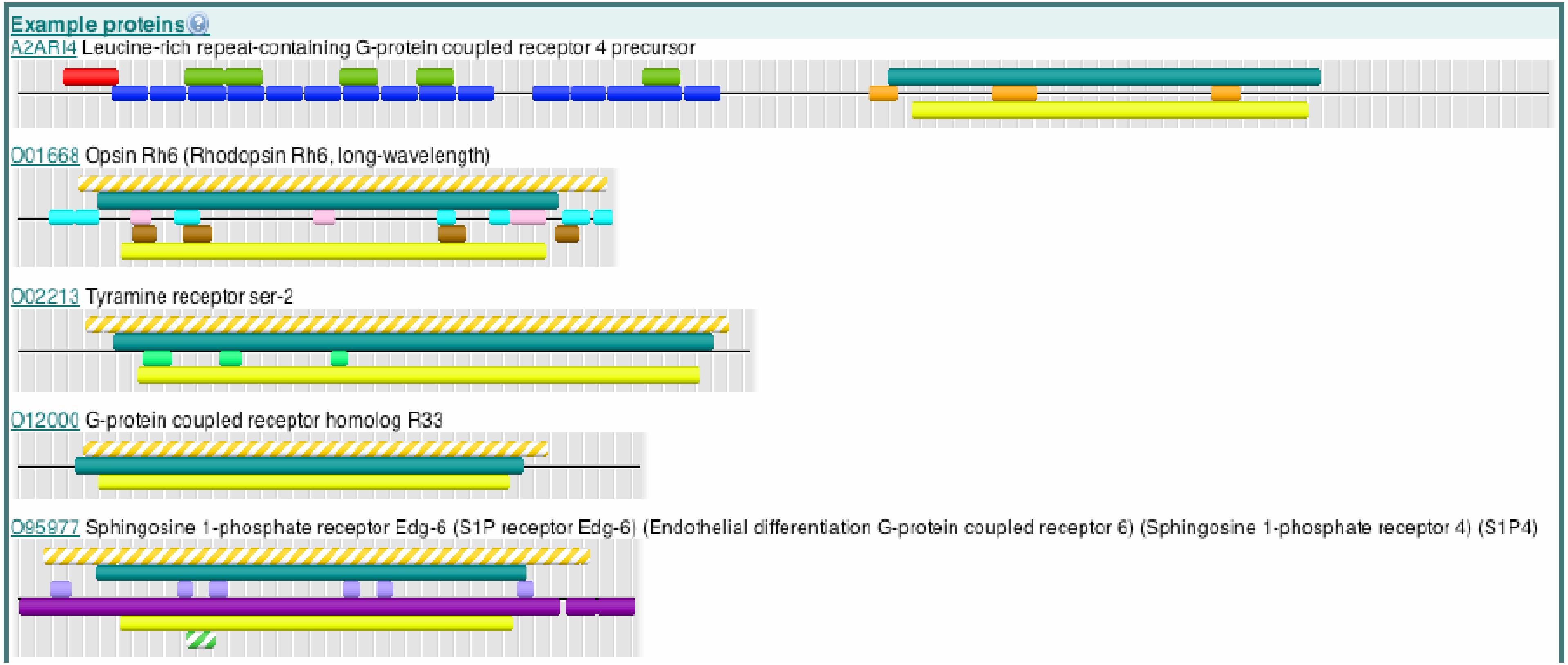
Taxonomic coverage

	Saccharomyces cerevisiae	Unclassified	
2	Fungi	Virus	170
205	Caenorhabditis elegans	Archaea	
400	Nematoda	Bacteria	1
17545	Metazoa	Cyanobacteria	
255	Fruit Fly	Synechocystis PCC 6803	
2095	Arthropoda	Oryza sativa (Rice)	
13787	Chordata	Arabidopsis thaliana	
2699	Mouse	Green Plants	
1585	Human	Plastid Group	1
17549	Eukaryota	Other Eukaryotes	1

Overlapping InterPro entries

IPR000276	Numbers of overlapping proteins			Average numbers of overlapping amino acids
IPR000147 % Overlap: 100	17740	19	0	N/A
IPR000154 % Overlap: 100	17755	4	0	N/A
IPR000155 % Overlap: 100	17703	56	0	N/A
IPR000174 % Overlap: 100	17708	51	0	N/A
IPR000190 % Overlap: 100	17710	49	0	N/A
IPR000207 % Overlap: 100	17582	177	0	N/A
IPR000248 % Overlap: 100	17665	94	0	N/A
IPR000314 % Overlap: 100	17722	37	0	N/A
IPR000321 % Overlap: 100	17742	17	0	N/A
IPR000351 % Overlap: 100	17736	23	0	N/A
IPR000363 % Overlap: 100	17751	8	0	N/A

InterProのアノテーション
（ファミリー）



Example proteins

A2ARI4 Leucine-rich repeat-containing G-protein coupled receptor 4 precursor

O01668 Opsin Rh6 (Rhodopsin Rh6, long-wavelength)

O02213 Tyramine receptor ser-2

O12000 G-protein coupled receptor homolog R33

O95977 Sphingosine 1-phosphate receptor Edg-6 (S1P receptor Edg-6) (Endothelial differentiation G-protein coupled receptor 6) (Sphingosine 1-phosphate receptor 4) (S1P4)

InterProScanの出力結果は4つのタイプに分類される

Type defines the entry as a Family, Domain, Repeat or PTM (Post-translational modification).

An InterPro family is a group of evolutionarily related proteins, that share one or more domains/repeats in common. A InterPro entry of type=family may contain a signature for a small conserved region that is representative of the family, and need therefore not necessarily cover the whole protein.

An InterPro domain is an independent structural unit which can be found alone or in conjunction with other domains or repeats. Domains are evolutionarily related. An InterPro entry of the type=domain is diagnostic for a domain but not necessarily define domain boundaries exactly.

An InterPro repeat is a region that is not expected to fold into a globular domain on its own. For example 6-8 copies of the WD40 repeat are needed to form a single globular domain. There also many other short repeat motifs that probably do not form a globular fold that have type=repeat.

A post-translational modification includes for example, an N glycosylation site. The sequence motif is defined by the molecular recognition of this region in a cell. This may group together proteins that need not be evolutionarily related.

タンパク質立体構造の
データベース

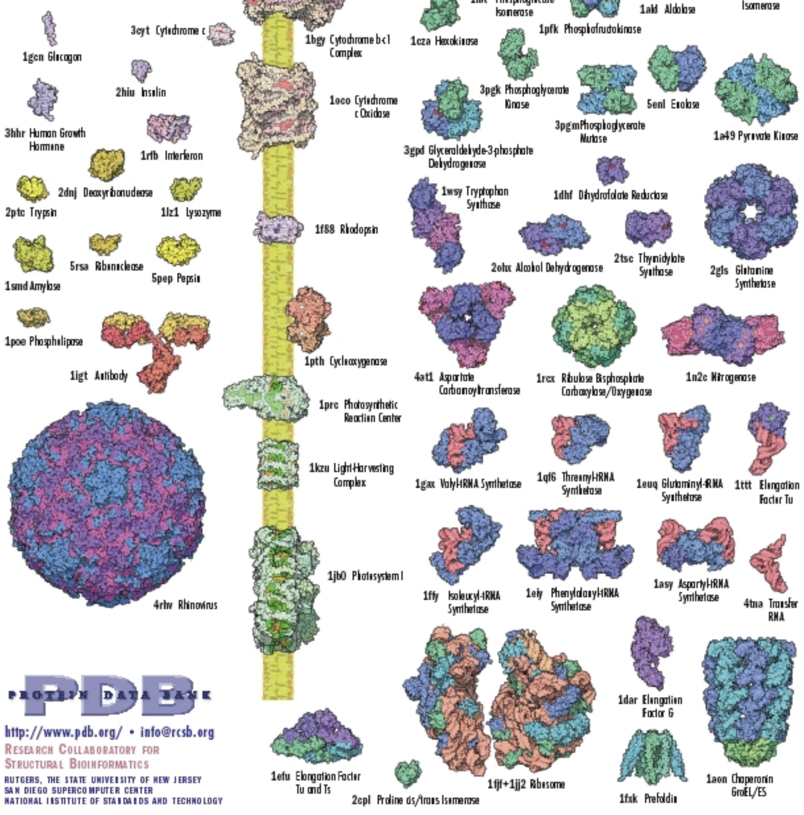
Protein Data Bank (PDB)

by Helen M. Berman

PDBとは

- Protein Data Bank (PDB)は、タンパク質と核酸の3次元構造データのデータベース
- 3次元構造データは、X線結晶解析法、NMR法 (核磁気共鳴法) などによって実験的に決定されたデータ
- Worldwide Protein Data Bank (wwPDB)
  - PDBのデータの登録、処理、配布を行う
  - 以下の組織による共同組織
    - Research Collaboratory for Structural Bioinformatics (RCSB PDB), USA
    - the Protein Databank in Europe (PDBe), Europe
    - 日本蛋白質構造データバンク (PDBj), Japan
    - Biological Magnetic Resonance Data Bank (BMRB), USA

MOLECULAR MACHINERY: A Tour of the Protein Data Bank

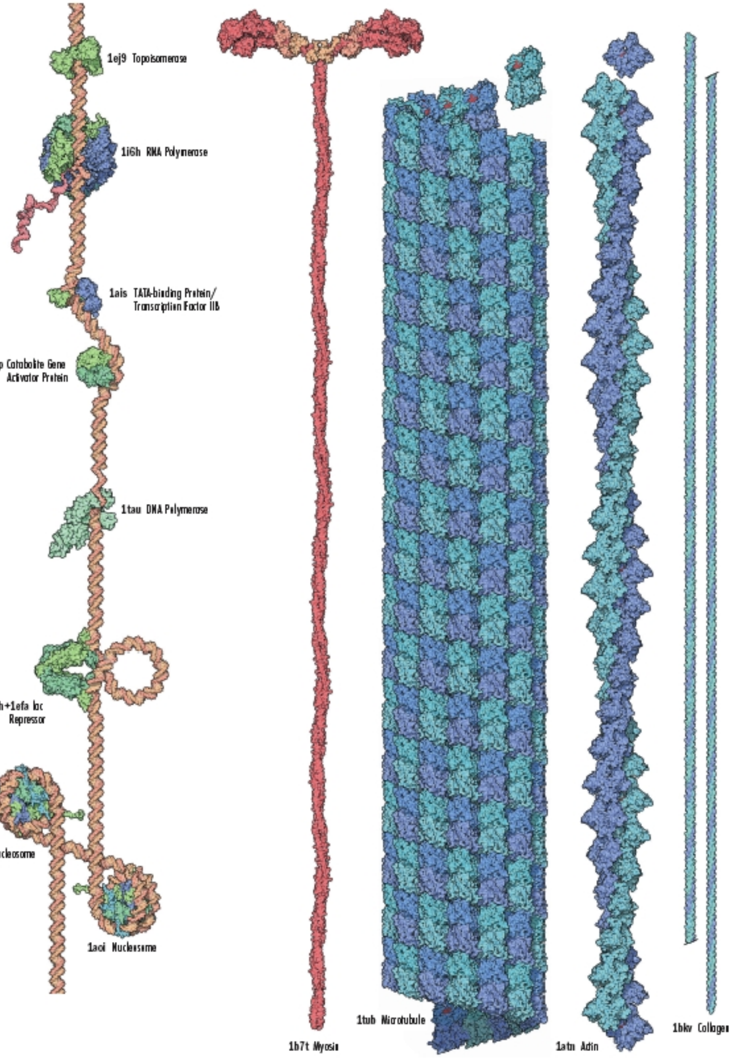


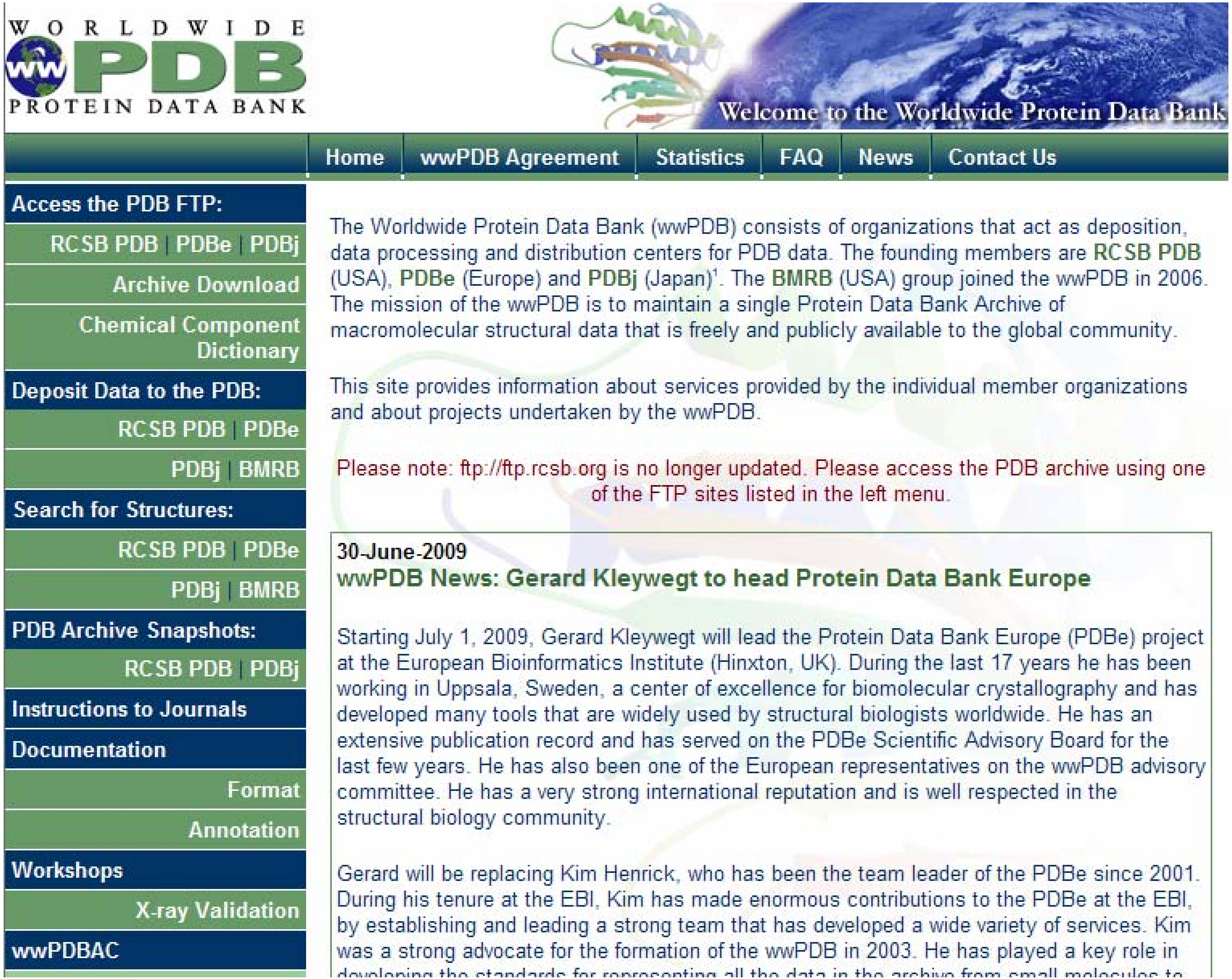
PDB
PROTEIN DATA BANK

http://www.pdb.org/ • info@rcsb.org

RESEARCH COLLABORATORY FOR STRUCTURAL BIOINFORMATICS

RUTGERS, THE STATE UNIVERSITY OF NEW JERSEY
SAN DIEGO SUPERCOMPUTER CENTER
NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY





The Worldwide Protein Data Bank (wwPDB) consists of organizations that act as deposition, data processing and distribution centers for PDB data. The founding members are RCSB PDB (USA), PDBe (Europe) and PDBj (Japan)[1]. The BMRB (USA) group joined the wwPDB in 2006. The mission of the wwPDB is to maintain a single Protein Data Bank Archive of macromolecular structural data that is freely and publicly available to the global community.

This site provides information about services provided by the individual member organizations and about projects undertaken by the wwPDB.

Please note: ftp://ftp.rcsb.org is no longer updated. Please access the PDB archive using one of the FTP sites listed in the left menu.

30-June-2009

wwPDB News: Gerard Kleywegt to head Protein Data Bank Europe

Starting July 1, 2009, Gerard Kleywegt will lead the Protein Data Bank Europe (PDBe) project at the European Bioinformatics Institute (Hinxton, UK). During the last 17 years he has been working in Uppsala, Sweden, a center of excellence for biomolecular crystallography and has developed many tools that are widely used by structural biologists worldwide. He has an extensive publication record and has served on the PDBe Scientific Advisory Board for the last few years. He has also been one of the European representatives on the wwPDB advisory committee. He has a very strong international reputation and is well respected in the structural biology community.

Gerard will be replacing Kim Henrick, who has been the team leader of the PDBe since 2001. During his tenure at the EBI, Kim has made enormous contributions to the PDBe at the EBI, by establishing and leading a strong team that has developed a wide variety of services. Kim was a strong advocate for the formation of the wwPDB in 2003. He has played a key role in

www.wwpdb.org

日本タンパク質構造データベース PDBj

http://www.pdbj.org/

English Japanese Chinese Korean 統計情報 ヘルプ お問い合わせ

トップページ

データ登録 >>

- ADIT: PDB Deposition
- ADIT-NMR

検索 >>

- Search PDB (xPSSS)
- Latest Released Search
- Sequence-Navigator
- Structure-Navigator
- SeSAW
- Ligand Binding Sites (GIRAF)
- EM Navigator
- Search NMR Data (BMRB)
- Status Search

サービス&ソフトウェア >>

- jV: Graphic Viewer
- Protein Globe
- ASH
- MAFFTash
- **Structure Prediction >>**
  - CRNPRED
  - Spanner
  - SFAS

派生データベース >>

- eF-site/eF-seek/eF-surf
- eProtS
- ProMode
- Molecule of the Month

日本蛋白質構造データバンク(PDBj: Protein Data Bank Japan)は、JST-BIRDの支援を受け、米国RCSBおよび欧州PDBeと協力して、生体高分子の立体構造データベースを国際的に統一化されたアーカイブとして運営するとともに、様々な解析ツールを提供しております。

データ登録

データ登録のご案内 >>

PDB登録 NMRデータ登録

検索

PDB検索 NMRデータ検索

● PDB ID ○ Keywords [Go]

● Accession number ○ Deposition code [Go]

詳細条件検索 >>

最新情報

2009/5/27
PDB登録の主要ページに日本語版を追加しました。

2009/3/18
PDBアーカイブ バージョン3.15が公開されました (詳細...)

2009/2/13
PDBアーカイブバージョン3.15が、2009年3月18日に公開されます (詳細...)

2009/1/23
FTPサイトのディレクトリ構造を変更します (詳細...)

2008/12/8
PDBアーカイブ バージョン3.15が公開されます (詳細...)

60173
entries available
on 16 Sep., 2009

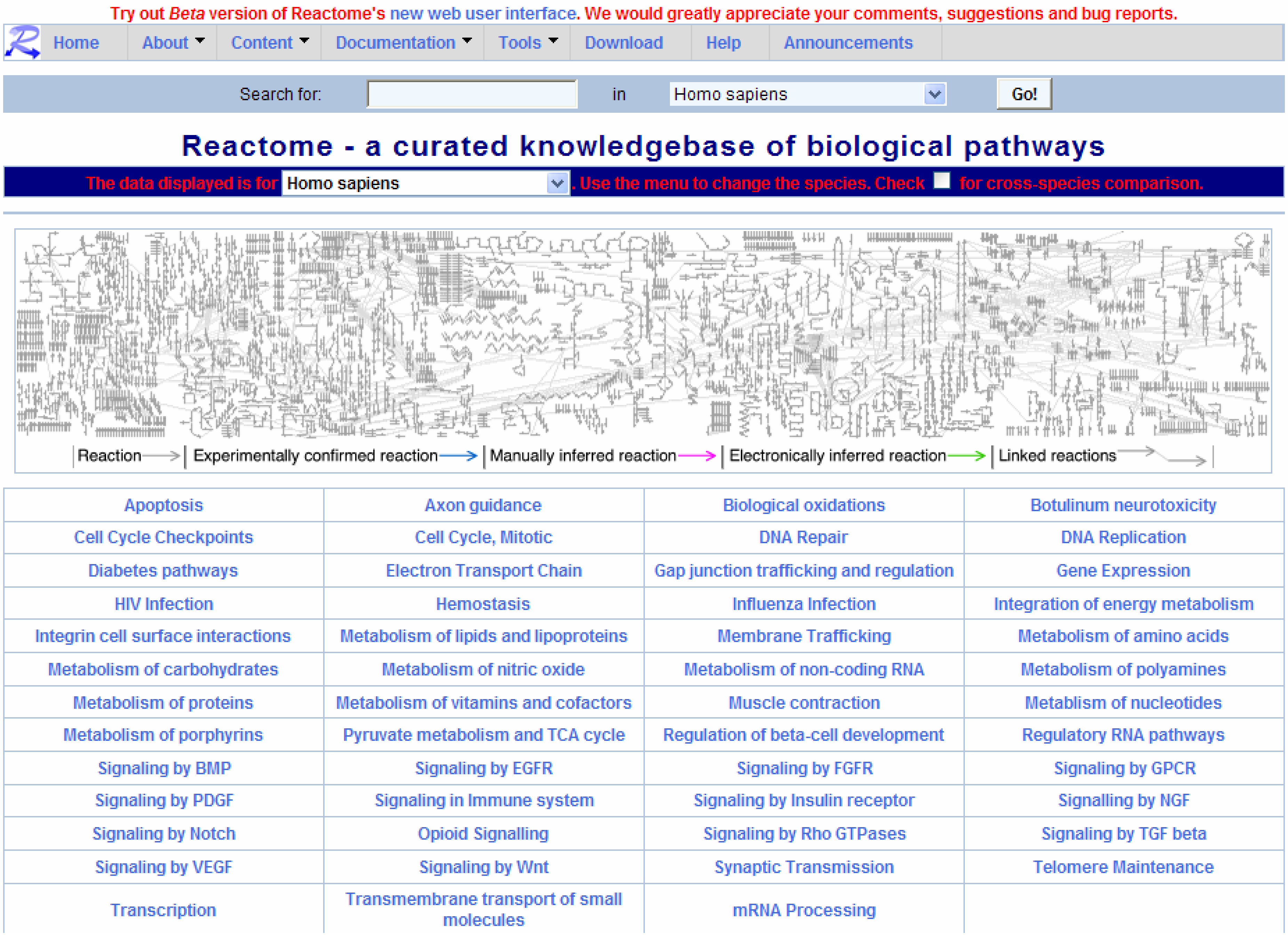
反応経路のデータベース

Reactome

KEGG

Reactome
http://reactome.org

- Reactomeは、ヒトの主要反応経路(pathway)や生化学反応(reaction)の精査されたデータベース
- ヒト以外のデータも格納している（23種）
- タンパク質間相互作用データもある
- 外部へリンクしているDBとしては、NCBI Entrez Gene, Ensembl and UniProt databases, the UCSC や HapMap のゲノムブラウザ, the KEGG Compound や ChEBI の低分子DB, PubMed, GO などがある
- Cold Spring Harbor Laboratory, EBI, GOコンソーシアムによる開発



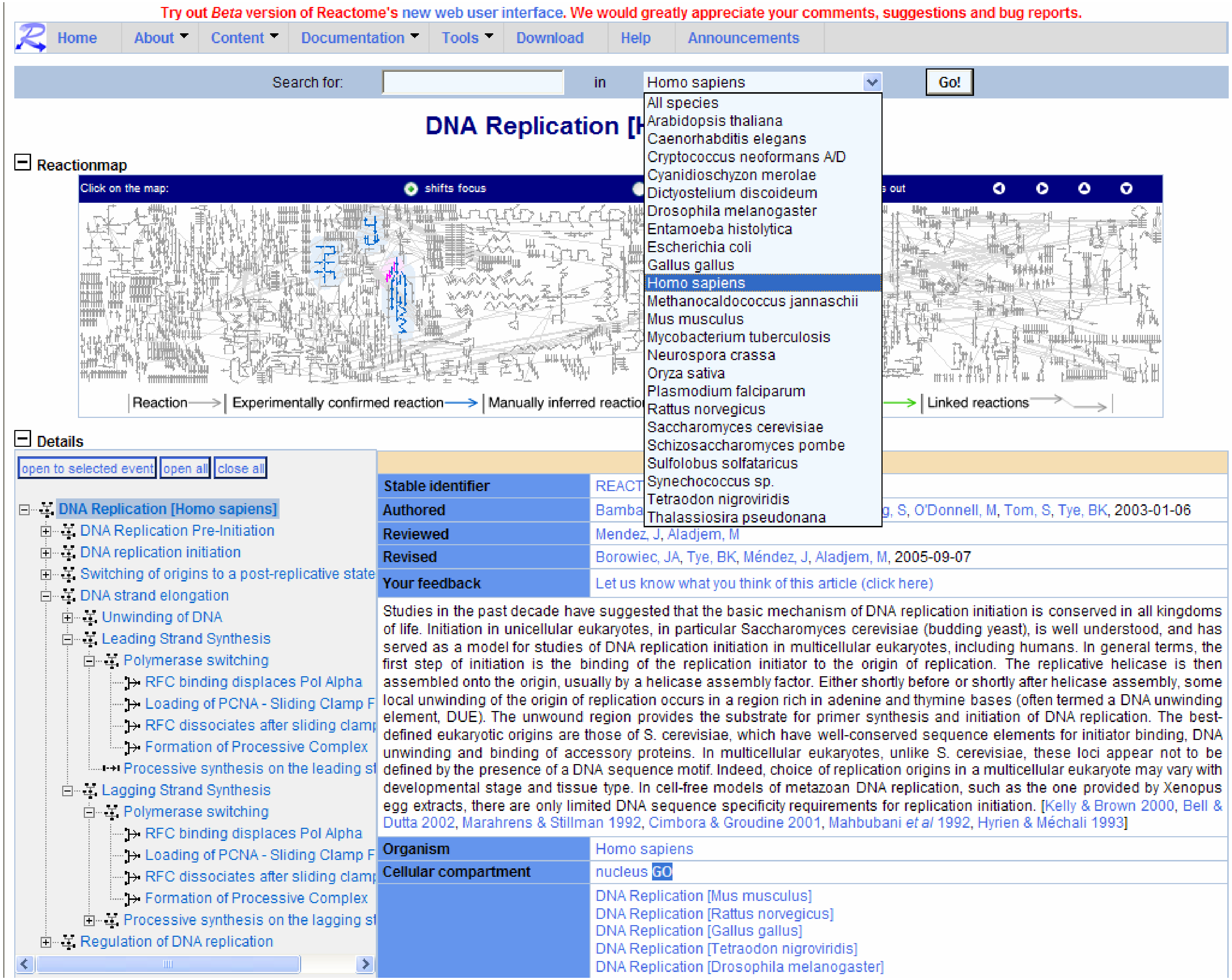
Try out *Beta* version of Reactome's new web user interface. We would greatly appreciate your comments, suggestions and bug reports.

Search for: in Homo sapiens

Reactome - a curated knowledgebase of biological pathways

The data displayed is for Homo sapiens. Use the menu to change the species. Check ☐ for cross-species comparison.

Apoptosis	Axon guidance	Biological oxidations	Botulinum neurotoxicity
Cell Cycle Checkpoints	Cell Cycle, Mitotic	DNA Repair	DNA Replication
Diabetes pathways	Electron Transport Chain	Gap junction trafficking and regulation	Gene Expression
HIV Infection	Hemostasis	Influenza Infection	Integration of energy metabolism
Integrin cell surface interactions	Metabolism of lipids and lipoproteins	Membrane Trafficking	Metabolism of amino acids
Metabolism of carbohydrates	Metabolism of nitric oxide	Metabolism of non-coding RNA	Metabolism of polyamines
Metabolism of proteins	Metabolism of vitamins and cofactors	Muscle contraction	Metablism of nucleotides
Metabolism of porphyrins	Pyruvate metabolism and TCA cycle	Regulation of beta-cell development	Regulatory RNA pathways
Signaling by BMP	Signaling by EGFR	Signaling by FGFR	Signaling by GPCR
Signaling by PDGF	Signaling in Immune system	Signaling by Insulin receptor	Signalling by NGF
Signaling by Notch	Opioid Signalling	Signaling by Rho GTPases	Signaling by TGF beta
Signaling by VEGF	Signaling by Wnt	Synaptic Transmission	Telomere Maintenance
Transcription	Transmembrane transport of small molecules	mRNA Processing	



DNA Replication [H

Stable identifier	REACT
Authored	Bamba g, S, O'Donnell, M, Tom, S, Tye, BK, 2003-01-06
Reviewed	Mendez, J, Aladjem, M
Revised	Borowiec, JA, Tye, BK, Méndez, J, Aladjem, M, 2005-09-07
Your feedback	Let us know what you think of this article (click here)

Studies in the past decade have suggested that the basic mechanism of DNA replication initiation is conserved in all kingdoms of life. Initiation in unicellular eukaryotes, in particular Saccharomyces cerevisiae (budding yeast), is well understood, and has served as a model for studies of DNA replication initiation in multicellular eukaryotes, including humans. In general terms, the first step of initiation is the binding of the replication initiator to the origin of replication. The replicative helicase is then assembled onto the origin, usually by a helicase assembly factor. Either shortly before or shortly after helicase assembly, some local unwinding of the origin of replication occurs in a region rich in adenine and thymine bases (often termed a DNA unwinding element, DUE). The unwound region provides the substrate for primer synthesis and initiation of DNA replication. The best-defined eukaryotic origins are those of S. cerevisiae, which have well-conserved sequence elements for initiator binding, DNA unwinding and binding of accessory proteins. In multicellular eukaryotes, unlike S. cerevisiae, these loci appear not to be defined by the presence of a DNA sequence motif. Indeed, choice of replication origins in a multicellular eukaryote may vary with developmental stage and tissue type. In cell-free models of metazoan DNA replication, such as the one provided by Xenopus egg extracts, there are only limited DNA sequence specificity requirements for replication initiation. [Kelly & Brown 2000, Bell & Dutta 2002, Marahrens & Stillman 1992, Cimbora & Groudine 2001, Mahbubani *et al* 1992, Hyrien & Méchali 1993]

Organism	Homo sapiens
Cellular compartment	nucleus GO
	DNA Replication [Mus musculus] DNA Replication [Rattus norvegicus] DNA Replication [Gallus gallus] DNA Replication [Tetraodon nigroviridis] DNA Replication [Drosophila melanogaster]

KEGG: Kyoto Encyclopedia of Genes and Genomes
http://www.genome.jp/ja/gn_kegg_ja.html

KEGG: 生命システム情報統合データベース

KEGGは、ゲノムの情報から生命システムの構築原理を理解し、環境との相互作用を理解して、創薬・医療・環境保全へとつなぐための、知識のレファレンスデータベースです。1995年より金久研究室（京都大学化学研究所バイオインフォマティクスセンターおよび東京大学医科学研究所ヒトゲノム解析センター）の研究活動の一環として構築されています。

KEGG ウェブサイト

KEGG - トップページ [日本語版]

KEGG2 - 目次のページ

データ内容別のエントリーポイント
- KEGG Atlas - パスウェイマップとBRITE機能階層の新しいビュア
- KEGG PATHWAY - パスウェイマップ・モジュール
- KEGG BRITE - 生命システムの機能階層・オントロジー
- KEGG ORTHOLOGY - オーソログアノテーション
- KEGG GENES - ゲノム・遺伝子・タンパク質
- KEGG LIGAND - 化合物・医薬品・糖鎖・生体内化学反応

特定分野別のエントリーポイント
- KEGG Organisms - 生物種ごとのリソース
- KEGG DISEASE - 病気に関するリソース [日本語版]
- KEGG DRUG - 医薬品に関するリソース [日本語版]
- KEGG GLYCAN - 糖鎖研究のリソース
- KEGG COMPOUND - メタボローム研究のリソース
- KEGG REACTION - 環境科学への応用を目指したリソース
- KEGG PLANT - 植物・天然物のリソース

デスクトップアプリケーション
- KegHier - Brite機能階層ブラウザ
- KegArray - マイクロアレイデータ解析ツール
- KegDraw - 化合物・糖鎖構造描画ツール

ソフトウェア開発支援ツール
- KGML - KEGGパスウェイのXML表現
- KEGG API - KEGGシステムのSOAP/WSDLインターフェース

リリース情報
- KEGG Release Notes - 更新履歴
- KEGG Update Notes - 新規パスウェイほか（RSS feedsでも提供）

KEGG FTP サイト

KEGG FTP - アカデミックユーザのみ

京都大学化学研究所バイオインフォマティクスセンター

- 分子相互作用を扱うKEGG pathway が有名だが、それ以外にも化合物を扱うKEGG LIGANDなど、多くのデータがある

多型データベース

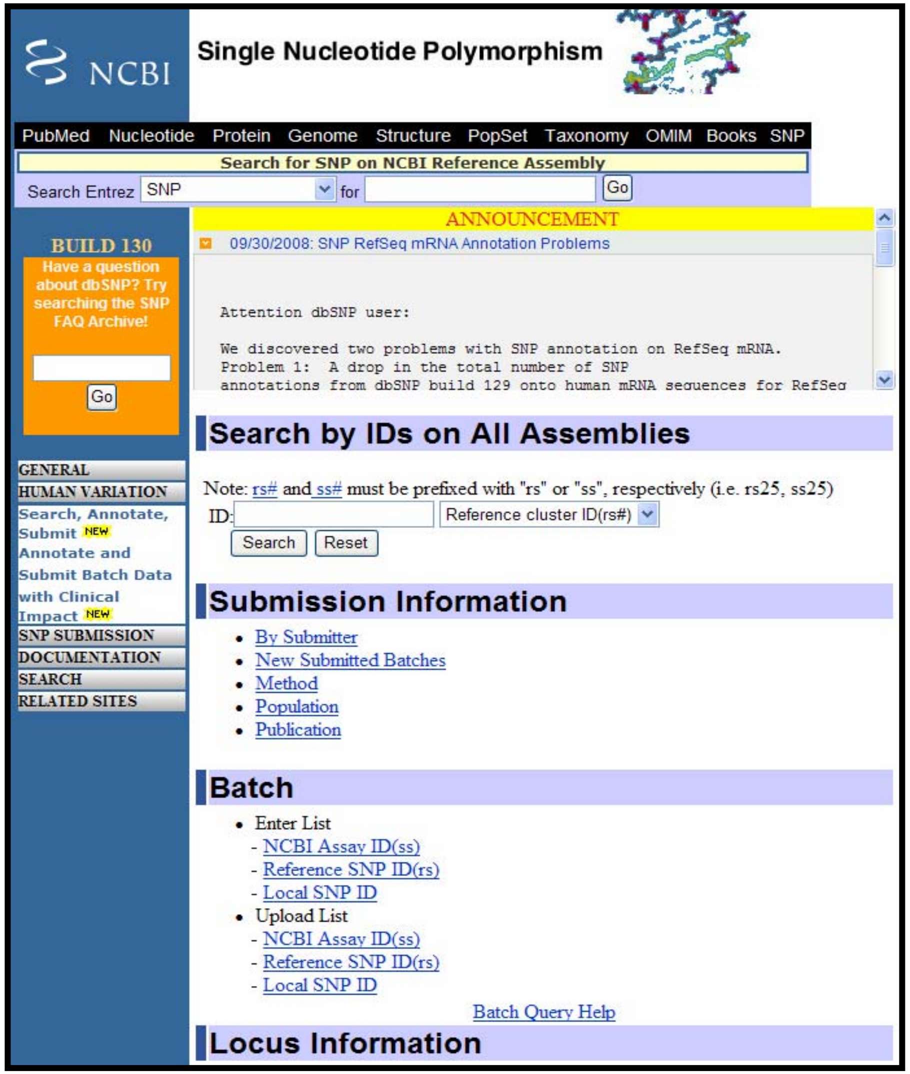
dbSNP

SNP

- 1塩基多型（single nucleotide polymorphisms, SNPs)は、最も共通な変異
- ゲノム中に、100から300塩基に1つの割合で存在する
- ゲノムと表現型をつなぐ基礎データ
- dbSNP(NCBI)は、置換だけでなく、挿入・欠失のSNPを含む

NCBIのdbSNP (トップ画面)

http://www.ncbi.nlm.nih.gov/projects/SNP/



Single Nucleotide Polymorphism

Search for SNP on NCBI Reference Assembly

BUILD 130

Have a question about dbSNP? Try searching the SNP FAQ Archive!

Search, Annotate, Submit NEW

Annotate and Submit Batch Data with Clinical Impact NEW

ANNOUNCEMENT

09/30/2008: SNP RefSeq mRNA Annotation Problems

```
Attention dbSNP user:

We discovered two problems with SNP annotation on RefSeq mRNA.
Problem 1:  A drop in the total number of SNP
annotations from dbSNP build 129 onto human mRNA sequences for RefSeq
```

Search by IDs on All Assemblies

Note: rs# and ss# must be prefixed with "rs" or "ss", respectively (i.e. rs25, ss25)

ID:

Submission Information

- By Submitter
- New Submitted Batches
- Method
- Population
- Publication

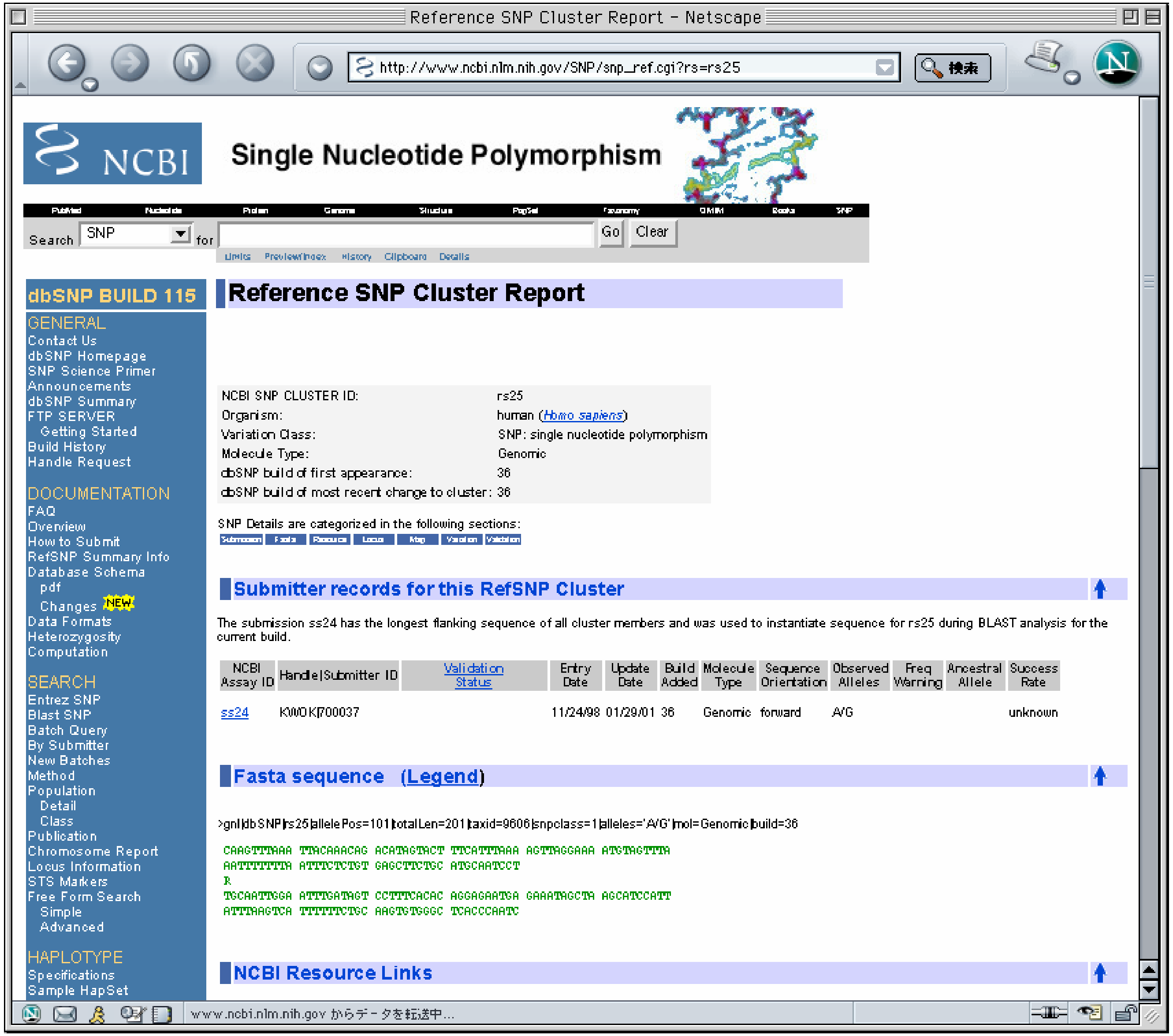
Batch

- Enter List
  - NCBI Assay ID(ss)
  - Reference SNP ID(rs)
  - Local SNP ID
- Upload List
  - NCBI Assay ID(ss)
  - Reference SNP ID(rs)
  - Local SNP ID

Batch Query Help

Locus Information

NCBIのdbSNPのデータ



Reference SNP Cluster Report

NCBI SNP CLUSTER ID:	rs25
Organism:	human (*Homo sapiens*)
Variation Class:	SNP: single nucleotide polymorphism
Molecule Type:	Genomic
dbSNP build of first appearance:	36
dbSNP build of most recent change to cluster:	36

SNP Details are categorized in the following sections:

Submitter records for this RefSNP Cluster

The submission ss24 has the longest flanking sequence of all cluster members and was used to instantiate sequence for rs25 during BLAST analysis for the current build.

NCBI Assay ID	Handle|Submitter ID	Validation Status	Entry Date	Update Date	Build Added	Molecule Type	Sequence Orientation	Observed Alleles	Freq Warning	Ancestral Allele	Success Rate
ss24	KWOK|700037		11/24/98	01/29/01	36	Genomic	forward	A/G			unknown

Fasta sequence (Legend)

```
>gnl|dbSNP|rs25|allelePos=101|totalLen=201|taxid=9606|snpclass=1|alleles='A/G'|mol=Genomic|build=36

CAAGTTTAAA TTACAAACAG ACATAGTACT TTCATTTAAA AGTTAGGAAA ATGTAGTTTA
AATTTTTTTA ATTTCTCTGT GAGCTTCTGC ATGCAATCCT
R
TGCAATTGGA ATTTGATAGT CCTTTCACAC AGGAGAATGA GAAATAGCTA AGCATCCATT
ATTTAAGTCA TTTTTTCTGC AAGTGTGGGC TCACCCAATC
```

NCBI Resource Links

遺伝子発現データベース

GEO (NCBI)

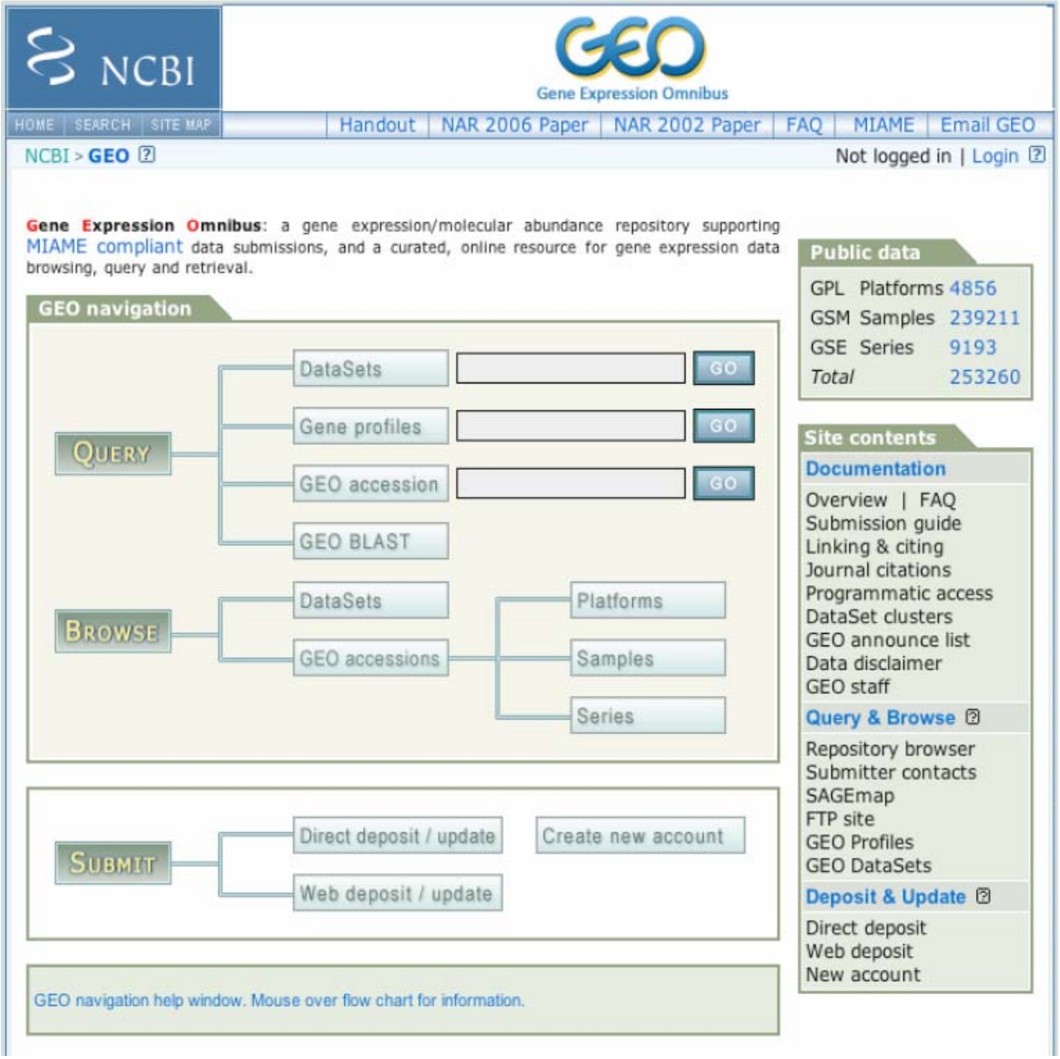
ArrayExpress (EBI)

CIBEX (DDBJ)

Gene Expression Omnibus (GEO)

http://www.ncbi.nlm.nih.gov/geo/

- GEOは、MIAMEに準拠した遺伝子発現データのリポジトリ（倉庫）



Gene Expression Omnibus: a gene expression/molecular abundance repository supporting MIAME compliant data submissions, and a curated, online resource for gene expression data browsing, query and retrieval.

Public data

GPL	Platforms	4856
GSM	Samples	239211
GSE	Series	9193
Total		253260

GEO navigation help window. Mouse over flow chart for information.

GEO and MIAME (Minimum Information About a Microarray Experiment)

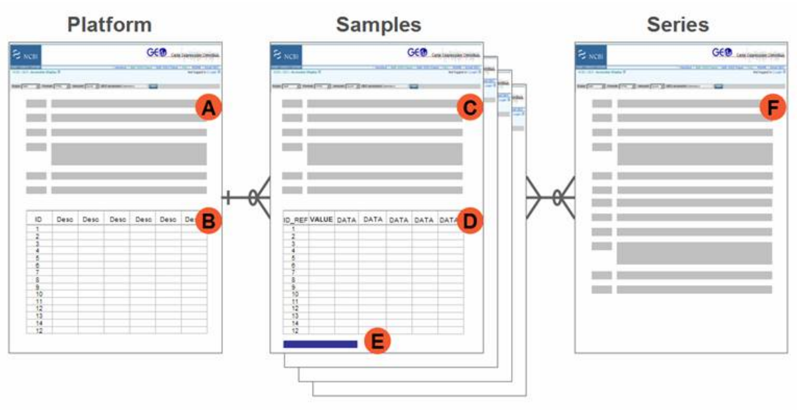
- MIAME：マイクロアレイ実験を記述する上で最低限の情報を規定したガイドライン
- データの形式や手段ではなく、中身を規定する

- 中身の例：生データ、生物種、どの臓器の試料か、実験デザイン、プローブ配列、データ処理プロトコール等

http://www.ncbi.nlm.nih.gov/geo/info/MIAME.html

GEO基本データ構造

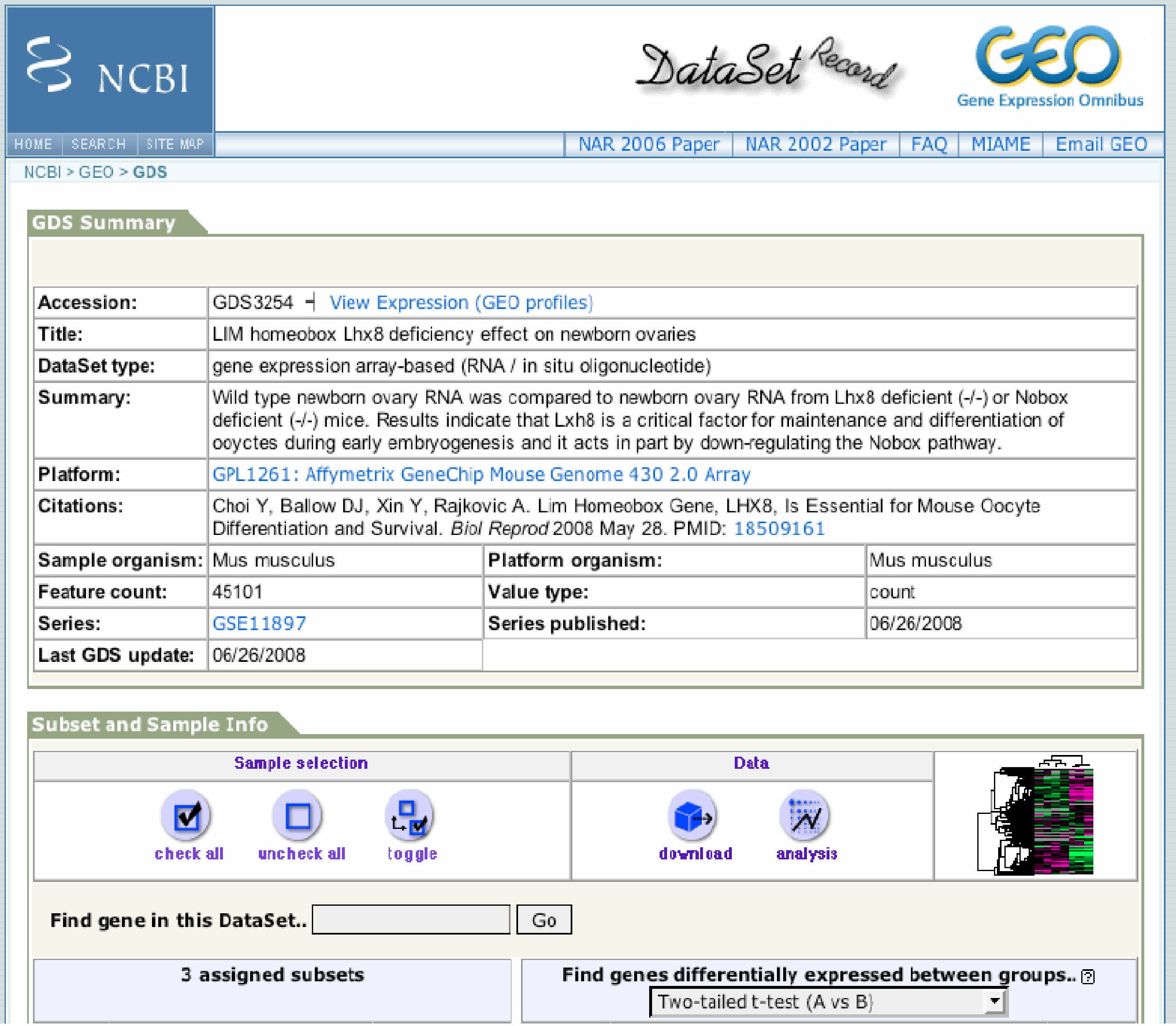
- Platform (GPLxxxx):アレイタイプ、プローブリスト等
- Sample (GSMxxxx):実験条件、測定値。Platformは1つだけ。多くのSeriesに属しうる。
- Series (GSExxxx):登録者が関連するサンプル(GSM)をグループ化したもの



- DataSets (GDSxxxx):統計的に比較可能なGSMをGEOが独自に再編成したもの
  - すぐ解析に使えて便利

http://www.ncbi.nlm.nih.gov/projects/geo/info/overview.html

GEOのデータ表示画面



GDS Summary

Accession:	GDS3254 View Expression (GEO profiles)		
Title:	LIM homeobox Lhx8 deficiency effect on newborn ovaries		
DataSet type:	gene expression array-based (RNA / in situ oligonucleotide)		
Summary:	Wild type newborn ovary RNA was compared to newborn ovary RNA from Lhx8 deficient (-/-) or Nobox deficient (-/-) mice. Results indicate that Lxh8 is a critical factor for maintenance and differentiation of ooyctes during early embryogenesis and it acts in part by down-regulating the Nobox pathway.		
Platform:	GPL1261: Affymetrix GeneChip Mouse Genome 430 2.0 Array		
Citations:	Choi Y, Ballow DJ, Xin Y, Rajkovic A. Lim Homeobox Gene, LHX8, Is Essential for Mouse Oocyte Differentiation and Survival. *Biol Reprod* 2008 May 28. PMID: 18509161		
Sample organism:	Mus musculus	**Platform organism:**	Mus musculus
Feature count:	45101	**Value type:**	count
Series:	GSE11897	**Series published:**	06/26/2008
Last GDS update:	06/26/2008		

Subset and Sample Info

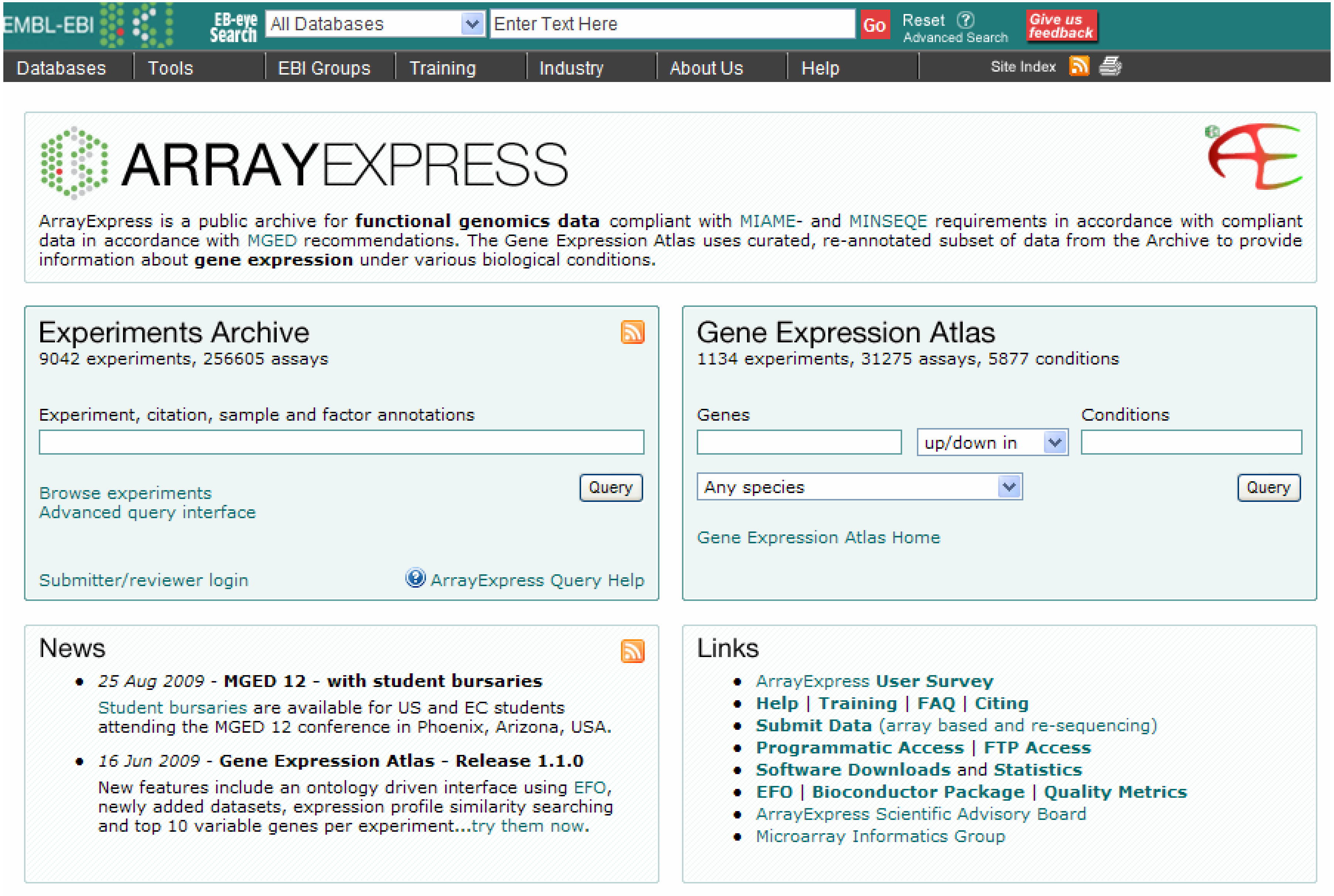
Find gene in this DataSet.. Go

3 assigned subsets

Find genes differentially expressed between groups..

Two-tailed t-test (A vs B)

ArrayExpress (EMBL-EBI)



ArrayExpress is a public archive for **functional genomics data** compliant with MIAME- and MINSEQE requirements in accordance with compliant data in accordance with MGED recommendations. The Gene Expression Atlas uses curated, re-annotated subset of data from the Archive to provide information about **gene expression** under various biological conditions.

Experiments Archive

9042 experiments, 256605 assays

Experiment, citation, sample and factor annotations

Browse experiments
Advanced query interface

Submitter/reviewer login

ArrayExpress Query Help

Gene Expression Atlas

1134 experiments, 31275 assays, 5877 conditions

Genes

Conditions

Gene Expression Atlas Home

News

- *25 Aug 2009* - **MGED 12 - with student bursaries**

  Student bursaries are available for US and EC students attending the MGED 12 conference in Phoenix, Arizona, USA.

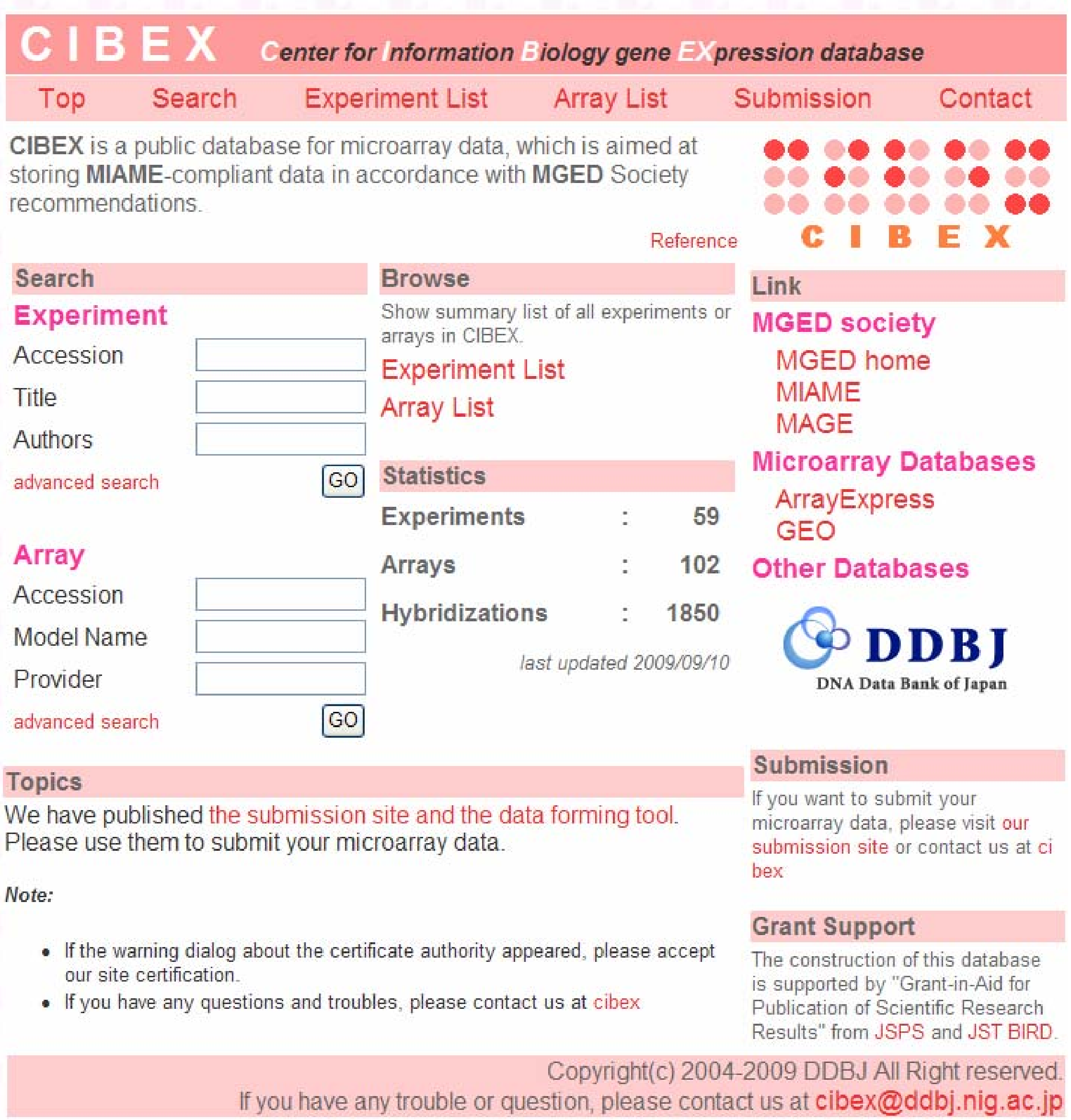
- *16 Jun 2009* - **Gene Expression Atlas - Release 1.1.0**

  New features include an ontology driven interface using EFO, newly added datasets, expression profile similarity searching and top 10 variable genes per experiment...try them now.

Links

- ArrayExpress **User Survey**
- **Help** | **Training** | **FAQ** | **Citing**
- **Submit Data** (array based and re-sequencing)
- **Programmatic Access** | **FTP Access**
- **Software Downloads** and **Statistics**
- **EFO** | **Bioconductor Package** | **Quality Metrics**
- ArrayExpress Scientific Advisory Board
- Microarray Informatics Group

CIBEX (DDBJ)



CIBEX Center for Information Biology gene EXpression database

CIBEX is a public database for microarray data, which is aimed at storing **MIAME**-compliant data in accordance with **MGED** Society recommendations.

Reference

Search

Experiment

Accession
Title
Authors
advanced search

Array

Accession
Model Name
Provider
advanced search

Browse

Show summary list of all experiments or arrays in CIBEX.

Experiment List
Array List

Statistics

Experiments	:	**59**
Arrays	:	**102**
Hybridizations	:	**1850**

last updated 2009/09/10

Link

MGED society
MGED home
MIAME
MAGE

Microarray Databases
ArrayExpress
GEO

Other Databases

Topics

We have published the submission site and the data forming tool. Please use them to submit your microarray data.

Note:

- If the warning dialog about the certificate authority appeared, please accept our site certification.
- If you have any questions and troubles, please contact us at cibex

Submission

If you want to submit your microarray data, please visit our submission site or contact us at cibex

Grant Support

The construction of this database is supported by "Grant-in-Aid for Publication of Scientific Research Results" from JSPS and JST BIRD.

Copyright(c) 2004-2009 DDBJ All Right reserved.
If you have any trouble or question, please contact us at cibex@ddbj.nig.ac.jp

オントロジー

Gene Ontology (GO)

Gene Ontology (GO)

- Gene Ontology Project とは、生物種やDBを超えて、遺伝子・遺伝子産物に関する表現（用語）を標準化するプロジェクト
- [効果]異なるDBの結合、比較が可能になる
- Gene Ontology Consortiumが主催するプロジェクト。
- 主なゲノム・バイオインフォマティクス研究機関が参加（Sanger, TIGR, EBIなど）。
- マウス、ショウジョウバエ、酵母などが対象。
- EC番号、MIPSの分類、InterPro、UniprotKB/Swiss-Protなどとの対応づけがある。

GO termの構造

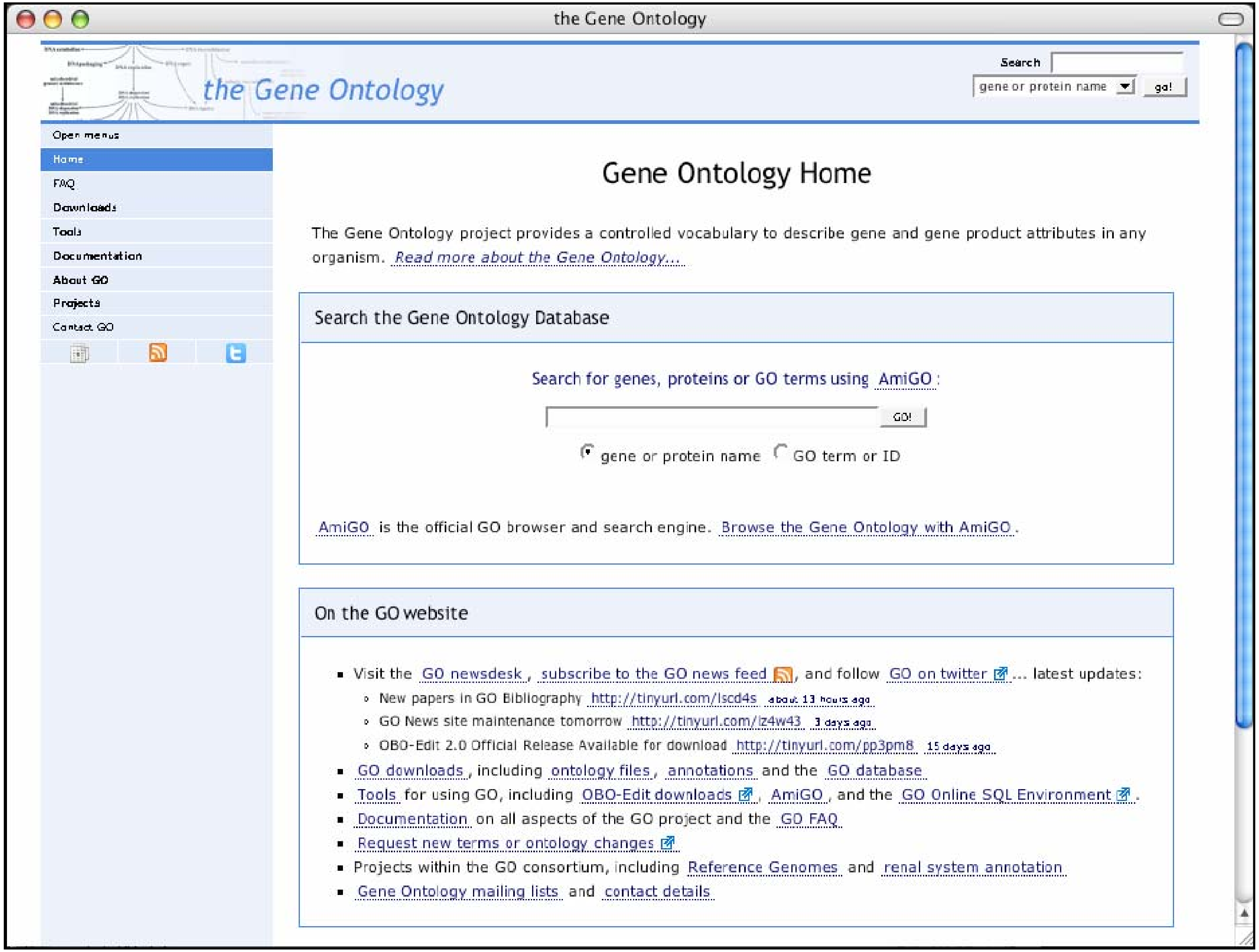
- 用語全体は階層構造を持つ
- 第1階層は3つ
  １．分子機能（molecular function）
  ２．生体内における役割（biological process）
  ３．細胞内構造や分布（cellular component）

GO termの例

ID: GO:nnnnnn

Name: cell, fibroblast growth factor receptor binding

http://www.geneontology.org/



Gene Ontology Home

The Gene Ontology project provides a controlled vocabulary to describe gene and gene product attributes in any organism. *Read more about the Gene Ontology...*

Search the Gene Ontology Database

Search for genes, proteins or GO terms using AmiGO :

GO!

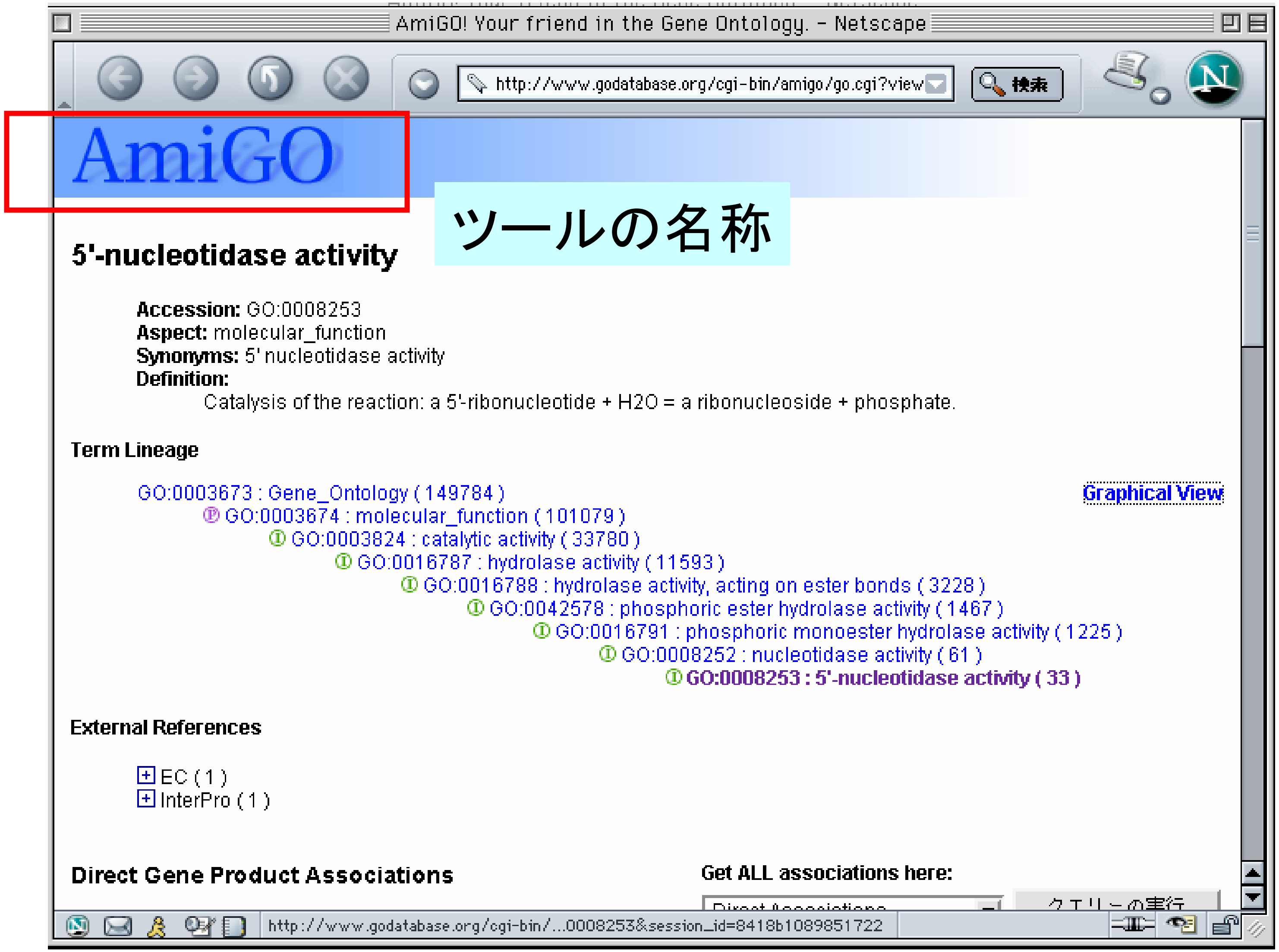
● gene or protein name ○ GO term or ID

AmiGO is the official GO browser and search engine. Browse the Gene Ontology with AmiGO .

On the GO website

- Visit the GO newsdesk , subscribe to the GO news feed , and follow GO on twitter ... latest updates:
  - New papers in GO Bibliography http://tinyurl.com/lscd4s about 13 hours ago
  - GO News site maintenance tomorrow http://tinyurl.com/lz4w43 3 days ago
  - OBO-Edit 2.0 Official Release Available for download http://tinyurl.com/pp3pm8 15 days ago
- GO downloads , including ontology files , annotations and the GO database
- Tools for using GO, including OBO-Edit downloads , AmiGO , and the GO Online SQL Environment .
- Documentation on all aspects of the GO project and the GO FAQ
- Request new terms or ontology changes
- Projects within the GO consortium, including Reference Genomes and renal system annotation
- Gene Ontology mailing lists and contact details

Molecular functionの例



ツールの名称

5'-nucleotidase activity

Accession: GO:0008253
Aspect: molecular_function
Synonyms: 5' nucleotidase activity
Definition:
Catalysis of the reaction: a 5'-ribonucleotide + H2O = a ribonucleoside + phosphate.

Term Lineage

- GO:0003673 : Gene_Ontology (149784)
  - Ⓟ GO:0003674 : molecular_function (101079)
    - Ⓘ GO:0003824 : catalytic activity (33780)
      - Ⓘ GO:0016787 : hydrolase activity (11593)
        - Ⓘ GO:0016788 : hydrolase activity, acting on ester bonds (3228)
          - Ⓘ GO:0042578 : phosphoric ester hydrolase activity (1467)
            - Ⓘ GO:0016791 : phosphoric monoester hydrolase activity (1225)
              - Ⓘ GO:0008252 : nucleotidase activity (61)
                - Ⓘ **GO:0008253 : 5'-nucleotidase activity (33)**

External References

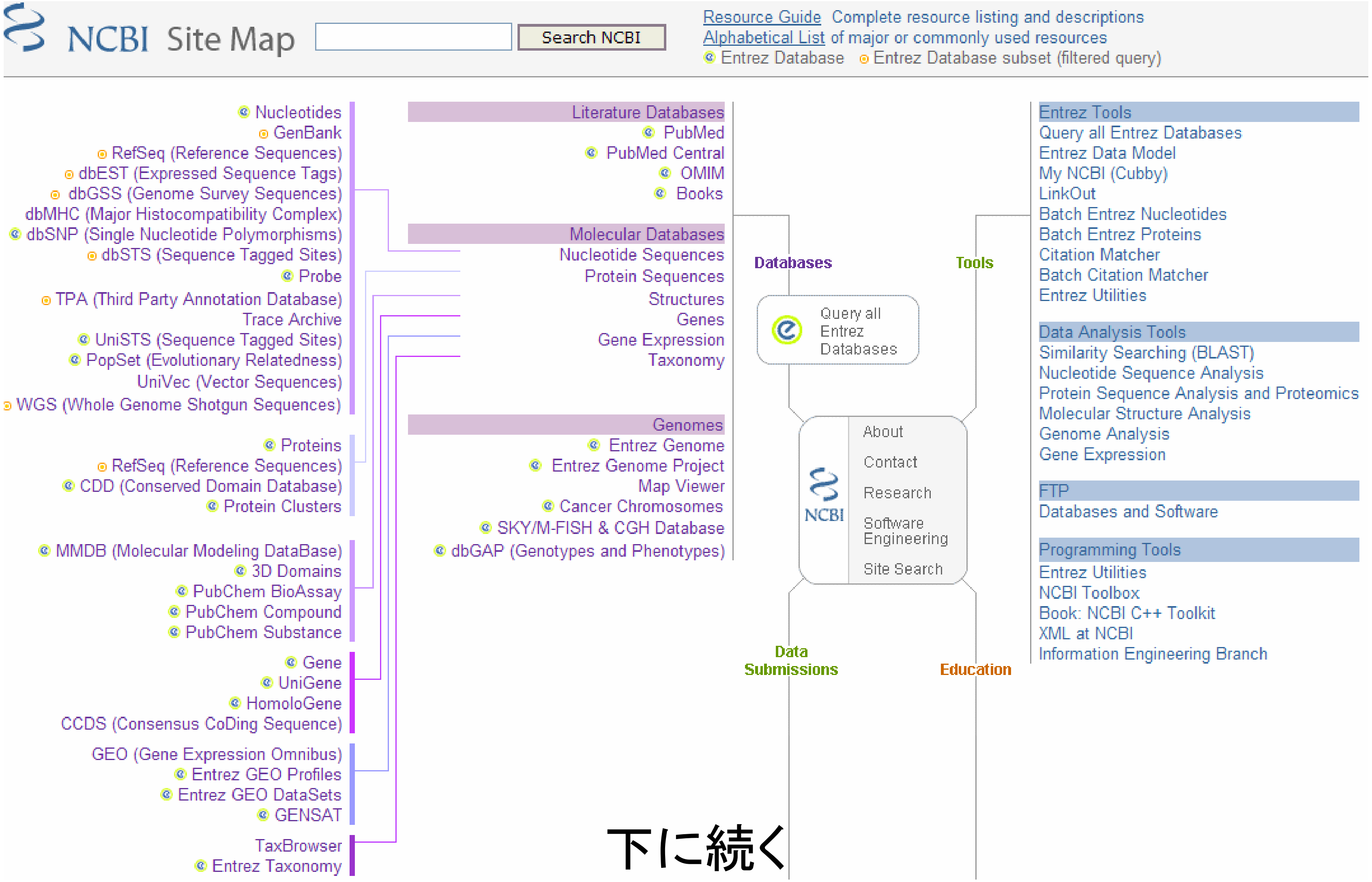
⊞ EC (1)
⊞ InterPro (1)

Direct Gene Product Associations

Get ALL associations here:

NCBI　その他

NCBI site map (1)



NCBI Site Map [ ] Search NCBI

Resource Guide Complete resource listing and descriptions
Alphabetical List of major or commonly used resources
Entrez Database Entrez Database subset (filtered query)

- Nucleotides
  - GenBank
  - RefSeq (Reference Sequences)
  - dbEST (Expressed Sequence Tags)
  - dbGSS (Genome Survey Sequences)
  - dbMHC (Major Histocompatibility Complex)
  - dbSNP (Single Nucleotide Polymorphisms)
  - dbSTS (Sequence Tagged Sites)
  - Probe
  - TPA (Third Party Annotation Database)
  - Trace Archive
  - UniSTS (Sequence Tagged Sites)
  - PopSet (Evolutionary Relatedness)
  - UniVec (Vector Sequences)
  - WGS (Whole Genome Shotgun Sequences)
- Proteins
  - RefSeq (Reference Sequences)
  - CDD (Conserved Domain Database)
  - Protein Clusters
- MMDB (Molecular Modeling DataBase)
  - 3D Domains
  - PubChem BioAssay
  - PubChem Compound
  - PubChem Substance
- Gene
  - UniGene
  - HomoloGene
  - CCDS (Consensus CoDing Sequence)
- GEO (Gene Expression Omnibus)
  - Entrez GEO Profiles
  - Entrez GEO DataSets
  - GENSAT
- TaxBrowser
  - Entrez Taxonomy

Literature Databases
- PubMed
- PubMed Central
- OMIM
- Books

Molecular Databases
- Nucleotide Sequences
- Protein Sequences
- Structures
- Genes
- Gene Expression
- Taxonomy

Genomes
- Entrez Genome
- Entrez Genome Project
- Map Viewer
- Cancer Chromosomes
- SKY/M-FISH & CGH Database
- dbGAP (Genotypes and Phenotypes)

Entrez Tools
- Query all Entrez Databases
- Entrez Data Model
- My NCBI (Cubby)
- LinkOut
- Batch Entrez Nucleotides
- Batch Entrez Proteins
- Citation Matcher
- Batch Citation Matcher
- Entrez Utilities

Data Analysis Tools
- Similarity Searching (BLAST)
- Nucleotide Sequence Analysis
- Protein Sequence Analysis and Proteomics
- Molecular Structure Analysis
- Genome Analysis
- Gene Expression

FTP
- Databases and Software

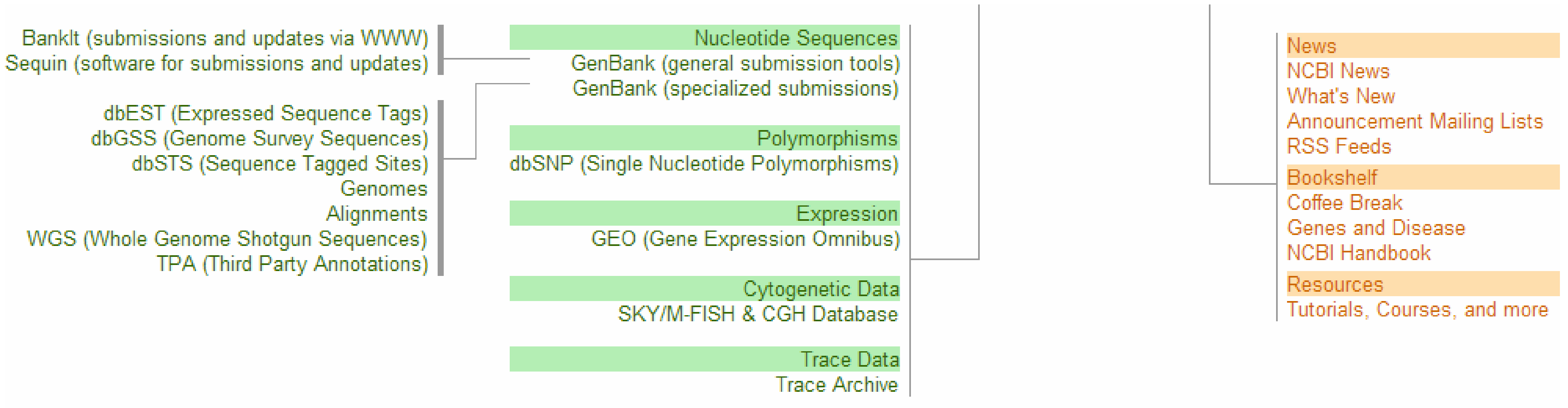
Programming Tools
- Entrez Utilities
- NCBI Toolbox
- Book: NCBI C++ Toolkit
- XML at NCBI
- Information Engineering Branch

下に続く

NCBI site map (2)

データ提出

教育



BankIt (submissions and updates via WWW)
Sequin (software for submissions and updates)

dbEST (Expressed Sequence Tags)
dbGSS (Genome Survey Sequences)
dbSTS (Sequence Tagged Sites)
Genomes
Alignments
WGS (Whole Genome Shotgun Sequences)
TPA (Third Party Annotations)

Nucleotide Sequences
GenBank (general submission tools)
GenBank (specialized submissions)

Polymorphisms
dbSNP (Single Nucleotide Polymorphisms)

Expression
GEO (Gene Expression Omnibus)

Cytogenetic Data
SKY/M-FISH & CGH Database

Trace Data
Trace Archive

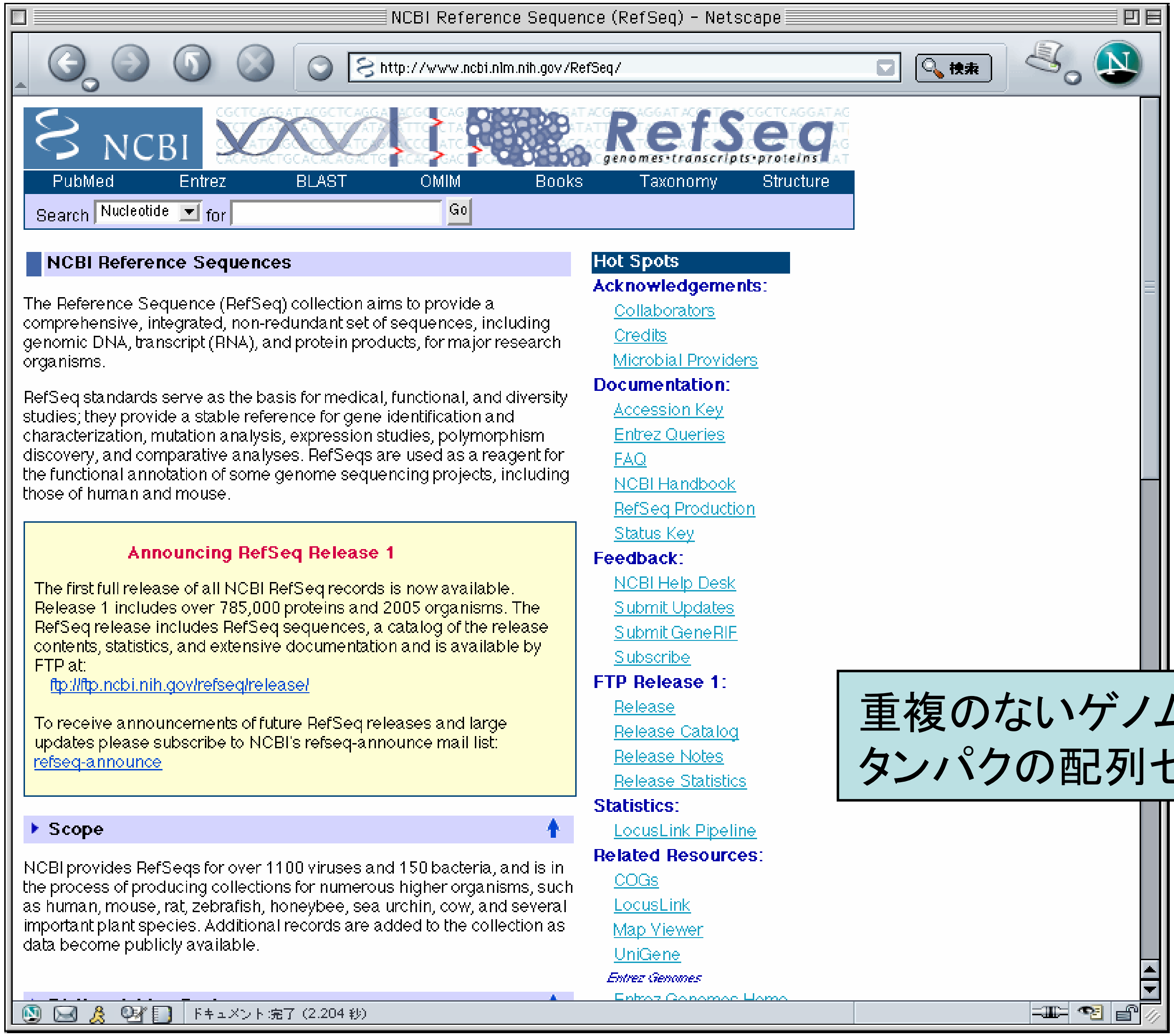
News
NCBI News
What's New
Announcement Mailing Lists
RSS Feeds

Bookshelf
Coffee Break
Genes and Disease
NCBI Handbook

Resources
Tutorials, Courses, and more

NCBIのRefSeq

http://www.ncbi.nlm.nih.gov/RefSeq/



NCBI Reference Sequences

The Reference Sequence (RefSeq) collection aims to provide a comprehensive, integrated, non-redundant set of sequences, including genomic DNA, transcript (RNA), and protein products, for major research organisms.

RefSeq standards serve as the basis for medical, functional, and diversity studies; they provide a stable reference for gene identification and characterization, mutation analysis, expression studies, polymorphism discovery, and comparative analyses. RefSeqs are used as a reagent for the functional annotation of some genome sequencing projects, including those of human and mouse.

Announcing RefSeq Release 1

The first full release of all NCBI RefSeq records is now available. Release 1 includes over 785,000 proteins and 2005 organisms. The RefSeq release includes RefSeq sequences, a catalog of the release contents, statistics, and extensive documentation and is available by FTP at:
ftp://ftp.ncbi.nih.gov/refseq/release/

To receive announcements of future RefSeq releases and large updates please subscribe to NCBI's refseq-announce mail list:
refseq-announce

Scope

NCBI provides RefSeqs for over 1100 viruses and 150 bacteria, and is in the process of producing collections for numerous higher organisms, such as human, mouse, rat, zebrafish, honeybee, sea urchin, cow, and several important plant species. Additional records are added to the collection as data become publicly available.

Hot Spots

Acknowledgements:
- Collaborators
- Credits
- Microbial Providers

Documentation:
- Accession Key
- Entrez Queries
- FAQ
- NCBI Handbook
- RefSeq Production
- Status Key

Feedback:
- NCBI Help Desk
- Submit Updates
- Submit GeneRIF
- Subscribe

FTP Release 1:
- Release
- Release Catalog
- Release Notes
- Release Statistics

Statistics:
- LocusLink Pipeline

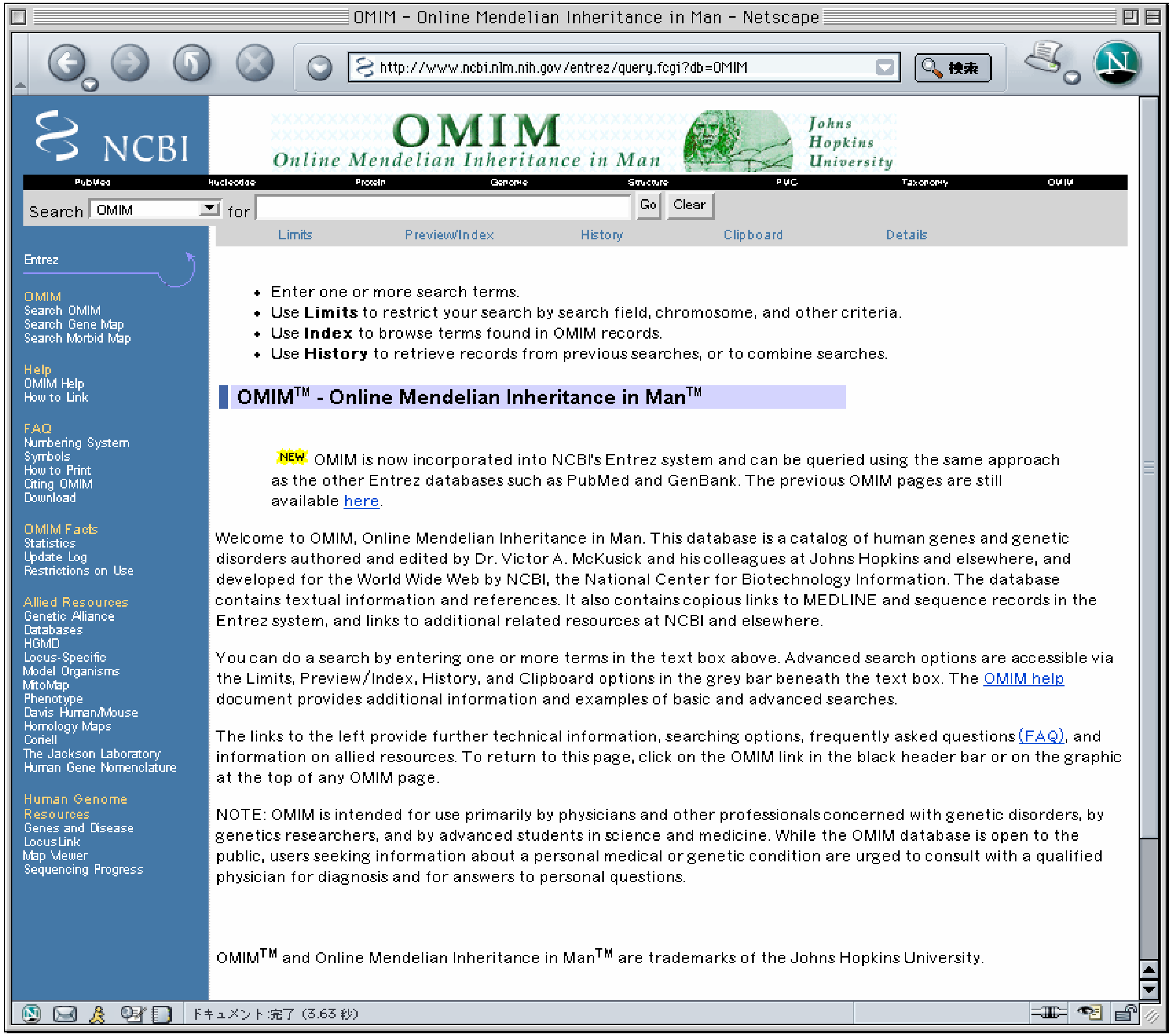
Related Resources:
- COGs
- LocusLink
- Map Viewer
- UniGene

Entrez Genomes

重複のないゲノム、DNA、RNA、タンパクの配列セットのDB

NCBIのOMIM（遺伝性疾患情報）

http://www.ncbi.nlm.nih.gov/sites/entrez?db=omim



- Enter one or more search terms.
- Use **Limits** to restrict your search by search field, chromosome, and other criteria.
- Use **Index** to browse terms found in OMIM records.
- Use **History** to retrieve records from previous searches, or to combine searches.

OMIM™ - Online Mendelian Inheritance in Man™

NEW OMIM is now incorporated into NCBI's Entrez system and can be queried using the same approach as the other Entrez databases such as PubMed and GenBank. The previous OMIM pages are still available here.

Welcome to OMIM, Online Mendelian Inheritance in Man. This database is a catalog of human genes and genetic disorders authored and edited by Dr. Victor A. McKusick and his colleagues at Johns Hopkins and elsewhere, and developed for the World Wide Web by NCBI, the National Center for Biotechnology Information. The database contains textual information and references. It also contains copious links to MEDLINE and sequence records in the Entrez system, and links to additional related resources at NCBI and elsewhere.

You can do a search by entering one or more terms in the text box above. Advanced search options are accessible via the Limits, Preview/Index, History, and Clipboard options in the grey bar beneath the text box. The OMIM help document provides additional information and examples of basic and advanced searches.

The links to the left provide further technical information, searching options, frequently asked questions (FAQ), and information on allied resources. To return to this page, click on the OMIM link in the black header bar or on the graphic at the top of any OMIM page.

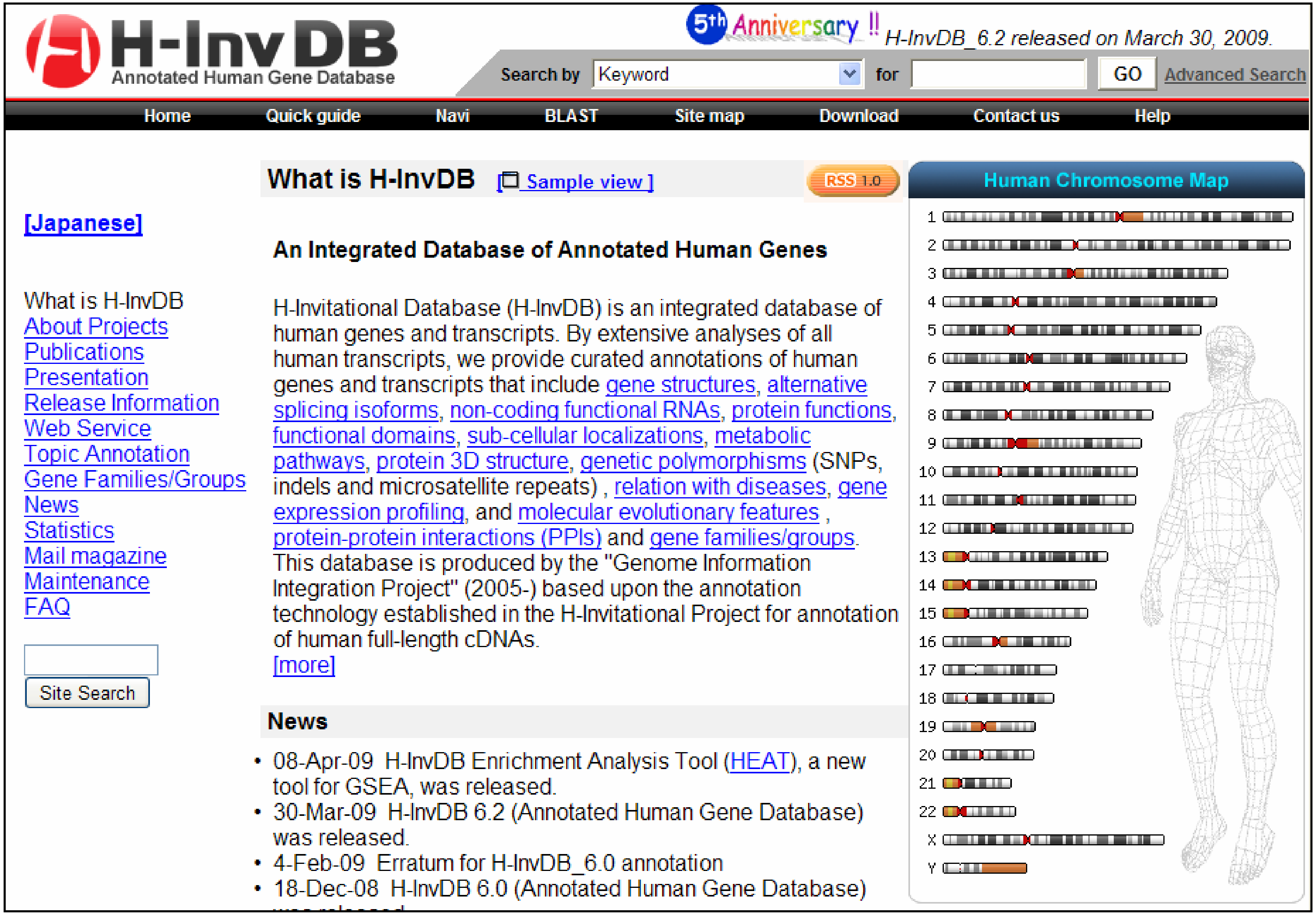
NOTE: OMIM is intended for use primarily by physicians and other professionals concerned with genetic disorders, by genetics researchers, and by advanced students in science and medicine. While the OMIM database is open to the public, users seeking information about a personal medical or genetic condition are urged to consult with a qualified physician for diagnosis and for answers to personal questions.

OMIM™ and Online Mendelian Inheritance in Man™ are trademarks of the Johns Hopkins University.

その他

ヒト遺伝子統合データベース：H-InvDB

http://h-invitational.jp/



H-InvDB_6.2 released on March 30, 2009.

What is H-InvDB

An Integrated Database of Annotated Human Genes

H-Invitational Database (H-InvDB) is an integrated database of human genes and transcripts. By extensive analyses of all human transcripts, we provide curated annotations of human genes and transcripts that include gene structures, alternative splicing isoforms, non-coding functional RNAs, protein functions, functional domains, sub-cellular localizations, metabolic pathways, protein 3D structure, genetic polymorphisms (SNPs, indels and microsatellite repeats) , relation with diseases, gene expression profiling, and molecular evolutionary features , protein-protein interactions (PPIs) and gene families/groups. This database is produced by the "Genome Information Integration Project" (2005-) based upon the annotation technology established in the H-Invitational Project for annotation of human full-length cDNAs.

[more]

News

- 08-Apr-09 H-InvDB Enrichment Analysis Tool (HEAT), a new tool for GSEA, was released.
- 30-Mar-09 H-InvDB 6.2 (Annotated Human Gene Database) was released.
- 4-Feb-09 Erratum for H-InvDB_6.0 annotation
- 18-Dec-08 H-InvDB 6.0 (Annotated Human Gene Database)

日本の統合データベースプロジェクト

http://lifesciencedb.jp/

日本語で各種DBの横断検索

経産省関連成果物

http://medals.jp/

連携関係

Guest Account | アカウント | ログイン | English

LSDB 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

ホーム | データベース | 検索 | ツール | ダウンロード | About us

統合ホームページへようこそ

横断検索

生命科学系 データベース カタログ
生命科学系 学協会カタログ
ゲノム・ポストゲノム主要プロジェクト一覧
生物アイコン
WingPro (JSTのDBポータル)
Webリソースポータルサイト (JST解析ツールポータル)

検索
生命科学データベース横断検索
蛋白質核酸酵素 全文検索
文科省「ゲノム」研究報告書 全文検索
学会要旨統合検索
OReFiL (オンラインリソースファインダー)
Allie (略語の正式名称を検索)
inMeXes (文献中の英語表現を軽快に検索)

データベース
DNAデータベース総覧と検索 (DDBJ/EMBL/GenBank)
遺伝子発現バンク(GEO)目次
かずさアノテーション & Navigation (かずさDNA研究所)
ゲノムネット医薬品データベース (京大)
統合医科学データベース (東京医科歯科大グループ)
疾患解析から医療応用を実現するDB開発 (東大グループ)
持続可能型社会への貢献遺伝子データベース (長浜バイオ)
tRNADB-CE:エキスパートがキュレートしたtRNADB (長浜バイオ)
日本糖鎖科学統合データベース (JCGGDBワーキンググループ)
生体分子の熱力学データと構造データの統合 (九工大)
MDeR (ライフサイエンス分野のメタデータ要素レポジトリ) (JST)

生命科学系データベースアーカイブサービス
DDBJトレースアーカイブ (遺伝研 DDBJ)
DDBJリードアーカイブ (遺伝研 DDBJ)

ツール & 解析サービス
アナトモグラフィー/BodyParts3D
Wired-Marker
MiGAP (微生物ゲノムアノテーションパイプライン)

基盤技術開発
TogoDB (誰でもデータベースが構築できる)
TogoWS (ウェブサービスの標準化)
OpenID 認証システム
統合DB情報基盤サイト (CBRC)

教材・人材育成
統合TV(DBやツールの動画教材)
MotDB (教育・人材育成のサイト)

統合DB事業
シンポジウム・講演会
外部評価
文科省 統合データベース整備事業サ
運営委員会等の資料
日本の生命科学データベース政策
過去のコンテンツ

連携
ゲノム解析ツール リンク集
経済産業省統合データベ
Human Gene and Protein D
BioHackathon 2009, 2008

新着情報
「生命科学系データベースアーカイブサービス」アップデートのお知らせ 2009-09-14 (Mon)
索」アップデート
on)
ーカイブサービ
2009-09-11

LSDBブログ
EMBOSS dottupを使い倒す 2009-09-18 (Fri) 16:08:55

MEDALS
METI database portal for life science

English 使い方ビデオ 検索

プロジェクト便覧を公開、成果報告書を3件追加 [09.09.08]

経済産業省ライフサイエンス統合データベースポータルサイト

このウェブサイトは経済産業省統合データベースプロジェクトのポータルサイトです。"MEDALS"は、METI Database portal for Life Scienceの略であり、METIとは経済産業省(Ministry of Economy, Trade and Industry)の意味です。これまで経済産業省関連機関により、様々なライフサイエンス分野の研究開発プロジェクトが実施されてきました。それらの成果物が効果的に利用されることを目的に、1989年以降の成果物(データベース、ツール等)の情報を整理・提供していきます。今後も内容を更新・追加していきます。もっと見る

データベース便覧 全リスト(45件)
・様々なデータベースの情報があります
DNA・ゲノム(18件) RNA(15件)
蛋白質(22件) その他(28件)

ツール便覧 全リスト(36件)
・データ解析に役立つツールの情報があります
DNA・ゲノム(11件) RNA(14件)
蛋白質(14件) その他(10件)

プロジェクト便覧 全リスト (14 件)
プロジェクトー成果物一覧
統合プロジェクト
ベースプロジェクト

ードページ
ウンロードできます
ツールダウンロード(2件)
成果報告書ダウンロード(4件)

統合データベース
・各種のデータを統合したDBです
ヒト遺伝子統合データベース (H-InvDB)
文科省統合データベースプロジェクト (連携先) DBCLS

サイト内検索
・MEDALSサイト内のキーワード検索です
サイト内の詳細な検索

MEDALSツール 全リスト(3件)
・MEDALSで開発した研究ツールです
リンク自動管理システム
ID一括変換システム
新規関連文献お知らせツール(PubMedScan)

リンク集
・バイオ産業関連情報へのリンク集です。
日本のサイト
米国のサイト
カナダのサイト

その他の情報リソース

- *Nucleic Acids Research*（NAR）のDB issue
  - 1993から現在まで、毎年1月に特集号
  - 1,170個の運用中DB (2009まで)
  - 繰り返し（1, 2年ごと）載るDBもある
    - 2009年は、１７９個のうち８４個（47%）がupdate版
- Wikipedia、特に英語版
- Google scholar（文献の引用、論文本体）
- ISI Web of knowledge, SCOPUS（文献引用）

バイオデータベース参考書

- **JSTのサイトにある参考図書**

http://www.jsbi.org/modules/jsbi/index.php/nintei/index.html

- **バイオリソース&データベース活用術**—Webでキャッチ!!実験材料・インフォマティクス (細胞工学 別冊) (大型本)
  - ナショナルバイオリソースプロジェクト情報運営委員会
  - 4830円

- **バイオデータベースとウェブツールの手とり足とり活用法** — 遺伝子の配列・機能解析、タンパク質解析、プロテオミクス、文献検索、検索エンジン…etc.真に役立つサイトを使い倒す！
  - 中村 保一 (編集), 石川 淳 (編集), 礒合 敦 (編集), 平川 美夏 (編集), 坊農 秀雅 (編集)
  - ４４１０円

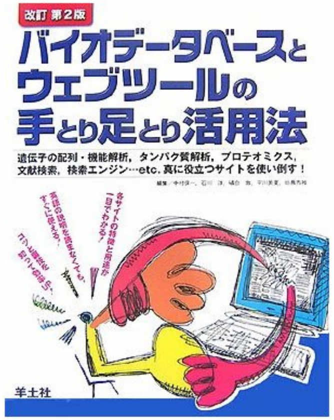
細胞工学 別冊

バイオリソース&
データベース活用術

Webでキャッチ!!
実験材料・インフォマティクス

監修●ナショナルバイオリソースプロジェクト情報運営委員会

秀潤社



改訂 第2版

バイオデータベースと
ウェブツールの
手とり足とり活用法

遺伝子の配列・機能解析，タンパク質解析，プロテオミクス，
文献検索，検索エンジン…etc. 真に役立つサイトを使い倒す！

羊土社

バイオデータベース参考書

- **バイオデータベースとソフトウェア最前線**—DNA解析からRNA･タンパク質の機能解明、エピジェネティクス研究、システム生物学と誰もが使う文献検索の新機能 (実験医学増刊 Vol. 26-7)
  - 森下 真一 (編集), 阿久津達也 (編集)
  - 5670円

EXPERIMENTAL MEDICINE
実験医学 増刊
Vol.26-No.7 2008

生命研究への応用と開発が進む
バイオデータベースとソフトウェア最前線
DNA解析からRNA・タンパク質の機能解明，エピジェネティクス研究，
システム生物学と誰もが使う文献検索の新機能

編集＝森下真一，阿久津達也

羊土社