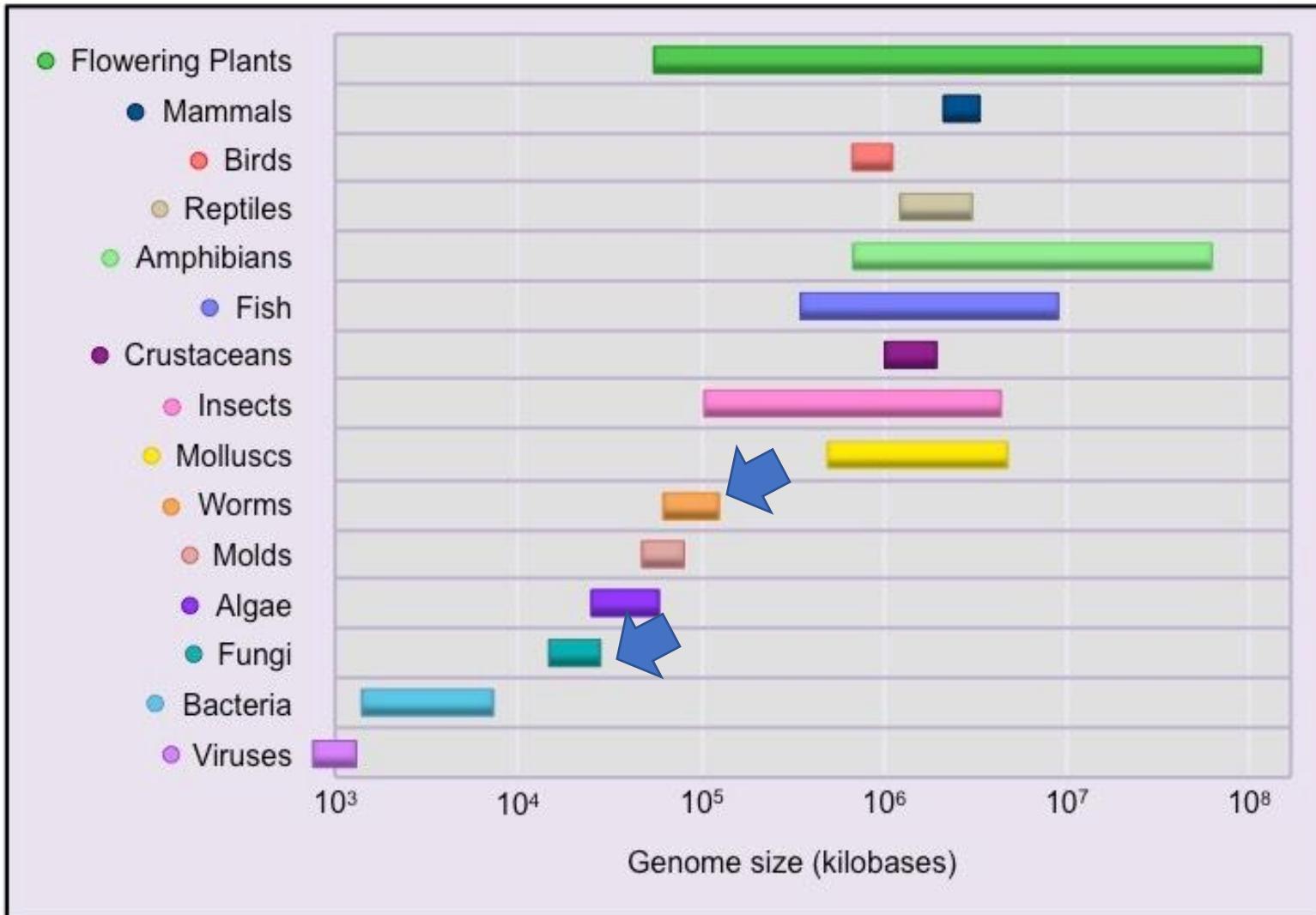


BCB 5200 Introduction to Bioinformatics

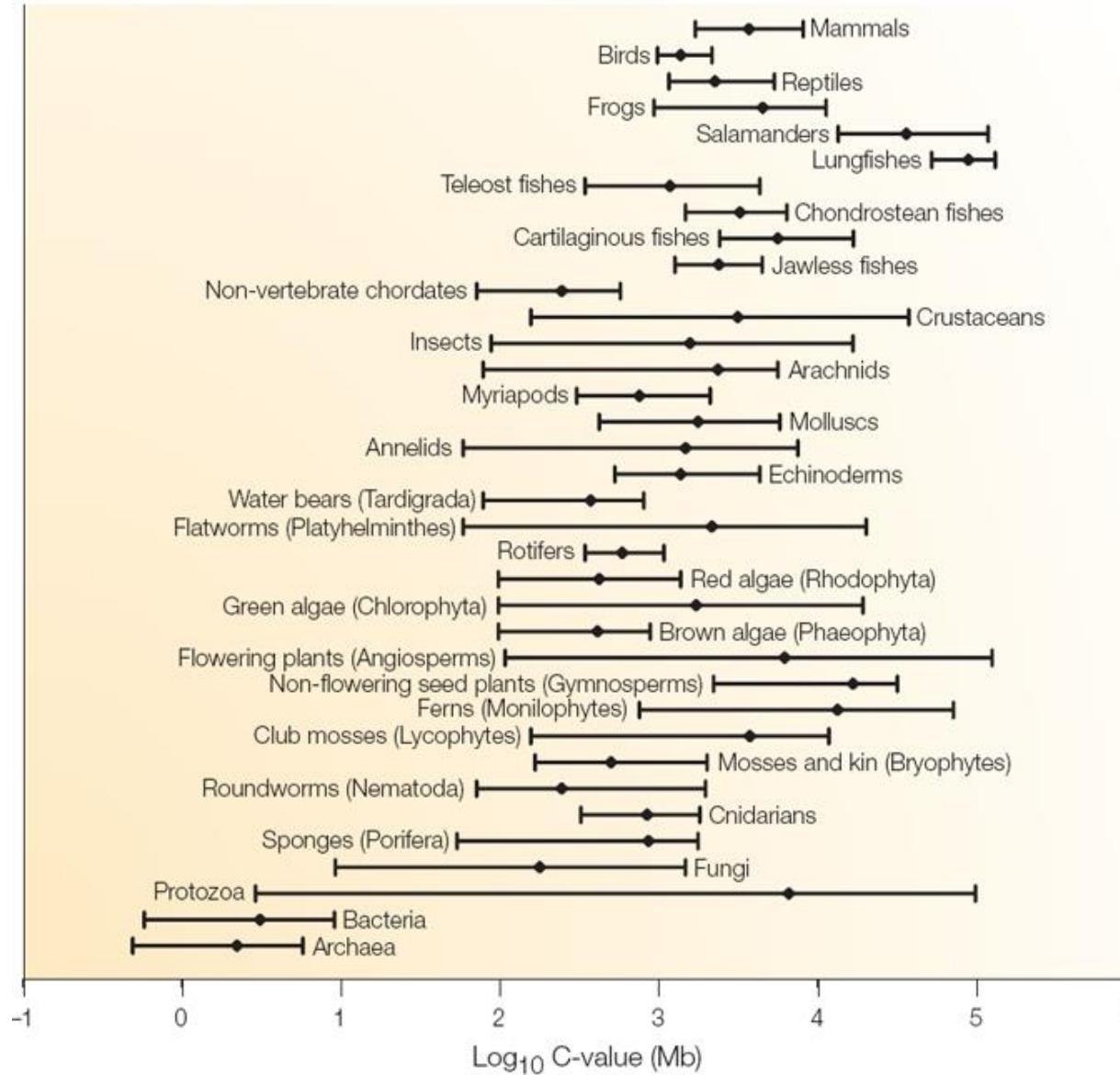
Lecture 02: Bioinformatics Resources

Bioinformatics and Computational Biology
Saint Louis University

What is the reason for big variation in genome sizes for those organisms



What is the reason for big variation in genome sizes for those organisms



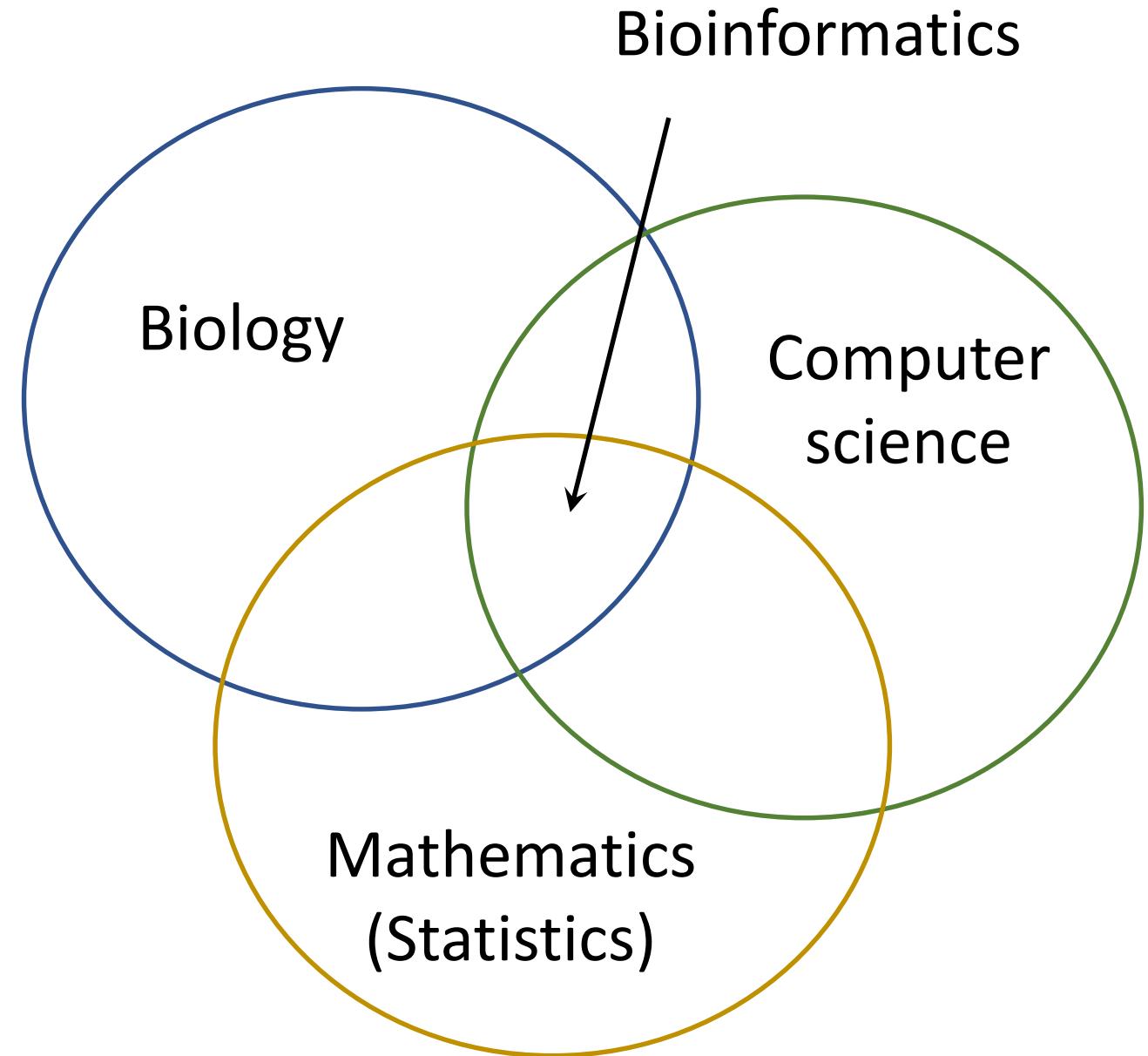
1. Content of genome
2. Timing of the expansion

Some more information
about Bioinformatics

BCB at SLU

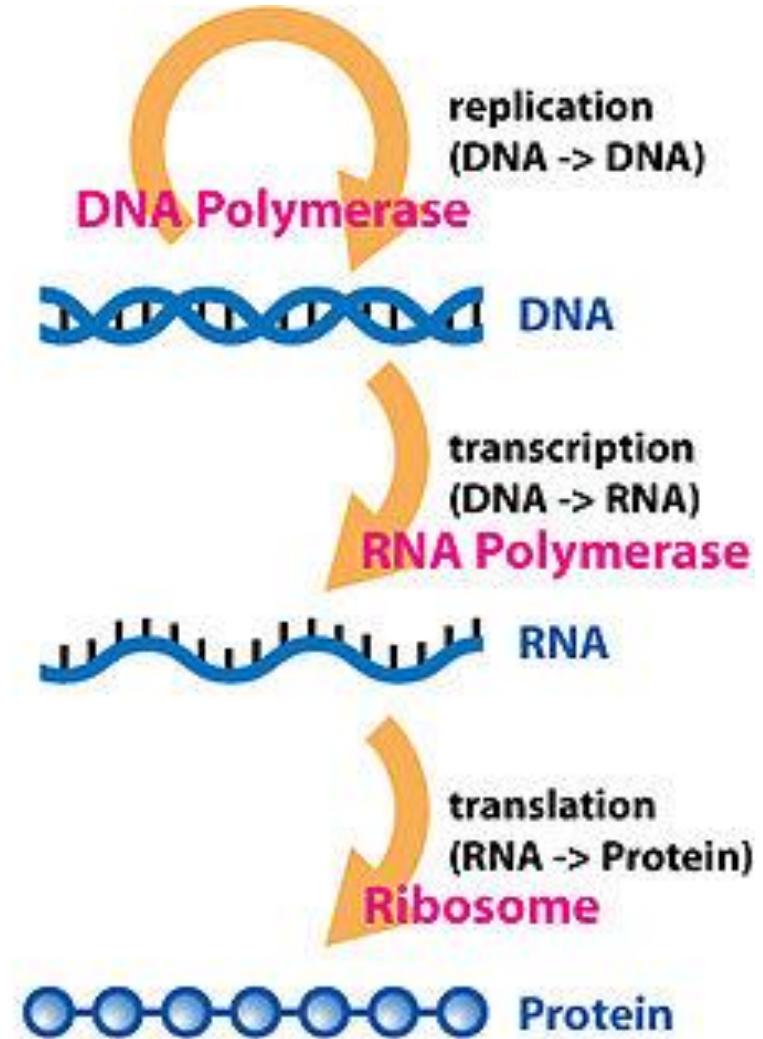
- (3 hours) **BCB 5200: Introduction to Bioinformatics I**
- (3 hours) **BCB 5250: Introduction to Bioinformatics II**
- (3 hours) **BCB 5300: Algorithms in Computational Biology**
- (3 hours) **BIOL 5030: Genomics**
- (3 hours) **Additional Biology selection** from among:
 - BIOL 5090: Biometry
 - BIOL 5100: Cellular and Molecular Genetics
 - BIOL 5700: Advanced Molecular Biology
 - BIOL 5780: Molecular Phylogenetic Analysis
- (3 hours) **Additional Computer Science selection** from among:
 - CSCI 5710: Databases
 - CSCI 5750: Machine Learning
 - CSCI 5850: High-Performance Computing
- (1 hour) **BCB 5810: Bioinformatics Colloquium**

Bioinformatics is an interdisciplinary field that develops and applies computer and computational techniques to study biological and biomedical questions.



Central Dogma of Molecular Biology

- Genome: sequence, structure
- RNA: sequence, expression level (microarray, RNA-seq), structure (3D)
- Protein: sequence (1), structure (3D), organization (domain), expression level, interactions



Examples of Bioinformatics

- Database interfaces
 - Genbank/EMBL/DDBJ, Medline, SwissProt, PDB, ...
- Sequence alignment/database similarity searches
 - BLAST, FASTA
- Multiple sequence alignment
 - Clustal, MultAlin, DiAlign, Muscle, Kalign, T-Coffee, Promals
- Gene finding/Genome annotation
 - Genscan, GenomeScan, GeneMark, GRAIL
- Protein Domain analysis and identification
 - Pfam, CDD, SMART, BLOCKS, ProDom,
- Pattern Identification/Characterization
 - Gibbs Sampler, AlignACE, MEME

Examples of Bioinformatics

- Phylogenetics
 - **Phylip, MEGA, PhyML, MrBayes, FastTree**
- Structural Bioinformatics/Protein Folding prediction
 - **PredictProtein, SwissModeler, Modeller, MD simulation**
- RNA biology
- Comparative genomics
 - **Pipmaker, Vista**
- Gene expression analysis
 - **Microarray, RNA-seq**
- Genetics and population analysis

Examples of Bioinformatics

- Mathematical biology
- Data and text mining
- Network biology
 - PPI, Cytoscape
- Systems Biology
- Metagenomics
- Microbiome
 - 16sRNA classification
- Bioimage Informatics

Bioinformatics research societies and conferences

- ISCB: International Society for Computational Biology
 - ISCB annual meeting
 - <https://www.iscb.org/iscb-conference-events>
 - Pacific Symposium on Biocomputing (PSB)
 - RECOMB conference
- SMBE: Society for Molecular Biology and Evolution
 - SMBE annual meeting
- GSA: The Genetics Society of America
- The Protein Society
 - Annual Symposium
- Biophysical Society
- EvoDir: <http://evol.mcmaster.ca/evoldir.html>

Bioinformatics Job websites

- Sciencecareers: jobs.sciencecareers.org/jobs/bioinformatics/
- Naturejobs
- ISCB
- <https://www.bioinformatics.org/jobs/>
- EvolDir: <http://evol.mcmaster.ca/evoldir.html>
- Indeed
- Washington Post

Biological Databases



Databases

- What is a database?
- A collection of related data elements
 - tables
 - columns (fields)
 - rows (records)
- Records retrieved using a query language
- Database technology is well established
 - MySQL, PostgreSQL, MongoDB, MariaDB, Microsoft SQL Server, Oracle ...
- Database (in the USA) vs Databank (in UK)
 - Solution: using the abbreviation *db*

Databases

- **Tables (entities)**
 - basic elements of information to track, e.g., *gene*, *organism*, *sequence*, *citation*
- **Columns (fields)**
 - attributes of tables, e.g. for *citation* table, *title*, *journal*, *volume*, *author*
- **Rows (records)**
 - actual data
 - whereas *fields* describe what kind of data is stored, the *rows* of a table are what the actual data is or where the data is stored

A flat-file database

Flat-file format, with fields separated by some delimiters

Jack | Kennell | St. Louis University | Department of Biology | Macelwane Hall

Tae-Hyuk (Ted) | Ahn | St. Louis University | Computer Science | Ritter Hall

Zhenguo | Lin | St. Louis University | Biology | Macelwane Hall

Dapeng | Zhang | St. Louis University | Department of Biology | Macelwane Hall

These data can be stored in a spreadsheet (columns, fields vs rows, records)

First Name	Family Name	University	Department	Address
Jack	Kennell	St. Louis University	Dept. of Biology	Macelwane Hall
Tae-Hyuk (Ted)	Ahn	St. Louis University	Computer Science	Ritter Hall
Zhenguo	Lin	St. Louis University	Biology	Macelwane Hall
Dapeng	Zhang	St. Louis University	Department of Biology	Macelwane Hall

What are the problems with this kind of database?

Relational databases

A relational database consists of **multiple tables** and defines the ***relationships*** between them. Each table contains attributes (fields or columns) and each row in a table is known as a record. Information will be “normalized” so that it is non-redundant (every row should be unique), although it is not always observed.

Professor ID	First Name	Family Name	Contact
1	Jack	Kennell	1
2	Tae-Hyuk (Ted)	Ahn	2
3	Zhenguo	Lin	1
4	Dapeng	Zhang	1

Primary key **Foreign key**

Contact ID	University	Department	Address
1	St. Louis University	Department of Biology	Macelwane Hall
2	St. Louis University	Department of Computer Science	Ritter Hall

Primary key

Why biological databases?

- Exponential growth in biological data.
- Data (genomic sequences, mRNA, protein, 3D structures, 2D gel analysis, MS analysis, gene expression data....) are no longer published in a conventional manner, but directly submitted to databases.
- A convenient way for data access

Some statistics

- More than 1000 different ‘biological’ databases
- Variable size: <100Kb to >20Gb
 - DNA: > 20 Gb
 - Protein: 1 Gb
 - 3D structure: 5 Gb
 - Other: smaller
- Update frequency: **daily to annually to seldom to forget about it.**

Categories of databases for Life Sciences

- Sequences (DNA, protein)
- Genomics
- Mutation/polymorphism
- Protein domain/family
- Proteomics (2D gel, Mass Spectrometry)
- 3D structure
- Metabolic networks
- Regulatory networks
- Bibliography
- Expression
- Specialized

Bibliography Databases

Bibliography databases:

- PubMed at NCBI, NLM, NIH
- PubMed Central: The U.S. National Library of Medicine digital archive of life science journal literature.
- Bookshelf: A collection of searchable biomedical books linked to PubMed.
- Google Scholar
- Web of Sciences (SCI)
- Elsevier

NCBI National Center for Biotechnology Information

(<http://www.ncbi.nlm.nih.gov/>)

- Entrez interface to databases
 - Medline/OMIM
 - Genbank/Genpept/Structures
- BLAST server(s)
 - Five-plus flavors of blast
- Draft Human Genome
- Much, much more...

The screenshot shows the NCBI homepage with a blue header bar containing the NCBI logo, a search bar, and links for Resources and How To. A dropdown menu is open under the Resources link, showing a list of databases including All Databases, HomoloGene, Identical Protein Groups, MedGen, MeSH, NCBI Web Site, NLM Catalog, Nucleotide, OMIM, PMC, PopSet, Probe, Protein, Protein Clusters, PubChem BioAssay, PubChem Compound, PubChem Substance, PubMed, PubMed Health, SNP, Sparcle, Sequence Analysis, Taxonomy, Training & Tutorials, and Variation. The main content area features four large cards: 'NCBI' (with a subtext about advancing science and health), 'Submit manuscripts to databases', 'Download', and 'Learn'. Below these are 'Develop', 'Analyze', and 'Research' sections. A sidebar on the right lists 'Popular Resources' such as PubMed, Bookshelf, PubMed Central, PubMed Health, BLAST, Nucleotide, Genome, SNP, Gene, Protein, and PubChem. At the bottom, there's a 'NCBI News & Blog' section with a recent entry about RefSeq release 83.

NCBI National Center for Biotechnology Information advances science and health by providing access to genomic information.

Mission | Organization | NCBI News & Blog

Submit manuscripts to databases

Download

Learn

Find help documents, attend a class or watch a tutorial

Develop

Analyze

Research

Explore NCBI research and collaborative projects

Popular Resources

- PubMed
- Bookshelf
- PubMed Central
- PubMed Health
- BLAST
- Nucleotide
- Genome
- SNP
- Gene
- Protein
- PubChem

NCBI News & Blog

RefSeq release 83 now public 21 Jul 2017

RefSeq release 83 is now accessible online, via FTP and through NCBI's programming utilities. This full release

New video on the NCBI YouTube channel: How You and Your Journal Club Can Contribute Using PubMed Commons 17 Jul 2017

The newest video on the NCBI YouTube

New releases from NCBI: Multiple Sequence Alignment Viewer 1.6, Tree Viewer 1.6.0, and Genome Workbench 2.12.0

About PubMed (Medline) ...

- PubMed is a database of over 26 million citations to published biomedical literature since 1948.
- Covers the fields of biology, medicine, nursing, dentistry, veterinary medicine, public health, and preclinical sciences.
- Most citations are for English-language articles and include abstracts.

- About 22% of the citations in PubMed link to free full text.
- Other full-text articles are available from the publisher.

About PMC...

- Separate full-text repository with more than 4 million articles
- Created to support the NIH Public Access Policy
 - Published results of research supported by NIH must be available to the public
- Publishers also voluntarily deposit articles
- Citations are created in PubMed when the article is published
- Text mining collections available:
<https://www.ncbi.nlm.nih.gov/pmc/tools/textmining/>

The Syntax ...

- Key words search
- Boolean operators: AND, OR, NOT must be entered in **UPPERCASE** (e.g., promoters OR response elements). The default is AND.
- Entrez processes all Boolean operators in a **left-to-right** sequence. You can change the order by enclosing individual concepts in **parentheses**. The terms inside the parentheses are processed first.
 - For example, the search statement: g1p3 OR (response AND element AND promoter).
- **Quotation marks:** The term inside the quotation marks is read as one phrase (e.g. “public health” is different than public health, which will also include articles on public latrines and their effect on health workers).
- **Asterisk:** Extends the search to all terms that start with the letters before the asterisk.
 - cilio* will include such terms as ciliopathy, ciliopathies, and ciliogenesis.

A search by subject:

Ciliopathy

Try to refine your searches with advanced functions

NCBI Resources How To

PubMed.gov US National Library of Medicine National Institutes of Health

PubMed Ciliopathy Create RSS Create alert Advanced Search Help

Article types Clinical Trial Review Customize ...

Text availability Abstract Free full text Full text

PubMed Commons Reader comments Trending articles

Publication dates 5 years 10 years Custom range...

Species Humans Other Animals

[Clear all](#)

[Show additional filters](#)

Format Summary Sort by: Most Recent Per page: 20 Send to Filters: Manage Filters

Search results Items: 1 to 20 of 3931 << First < Prev Page 1 of 197 Next > Last >>

1. Whole exome sequencing as a diagnostic tool for patients with ciliopathy-like phenotypes. Castro-Sánchez S, Álvarez-Satta M, Tohamy MA, Beltran S, Derdak S, Valverde D. PLoS One. 2017 Aug 11;12(8):e0183081. doi: 10.1371/journal.pone.0183081. eCollection 2017. PMID: 28800606 [Free Article](#) [Similar articles](#)

2. Primary Cilia Signaling Shapes the Development of Interneuronal Connectivity. Guo J, Otis JM, Higginbotham H, Monckton C, Cheng J, Asokan A, Mykytyn K, Caspary T, Stuber GD, Anton ES. Dev Cell. 2017 Aug 7;42(3):286-300.e4. doi: 10.1016/j.devcel.2017.07.010. PMID: 28787594 [Similar articles](#)

3. Novel Insights Into the Phenotypical Spectrum of KIF11-Associated Retinopathy, Including a New Form of Retinal Ciliopathy. Birtel J, Gliem M, Mangold E, Tebbe L, Spier I, Müller PL, Holz FG, Neuhaus C, Wolfrum U, Bolz HJ, Charbel Issa P. Invest Ophthalmol Vis Sci. 2017 Aug 1;58(10):3950-3959. doi: 10.1167/iovs.17-21679. PMID: 28785766 [Similar articles](#)

4. Interpreting the clinical significance of combined variants in multiple recessive disease genes: systematic investigation of Joubert syndrome yields little support for oligogenicity. Phelps IG, Dempsey JC, Grout ME, Isabella CR, Tully HM, Doherty D, Bachmann-Gagescu R. Genet Med. 2017 Aug 3. doi: 10.1038/gim.2017.94. [Epub ahead of print] PMID: 28771248 [Similar articles](#)

5. Nuclear/cytoplasmic transport defects in BBS6 underlie congenital heart disease through perturbation of a chromatin remodeling protein. Scott CA, Marsden AN, Rebagliati MR, Zhang Q, Chamling X, Searby CC, Baye LM, Sheffield VC, Slusarski DC. PLoS Genet. 2017 Jul 28;13(7):e1006936. doi: 10.1371/journal.pgen.1006936. eCollection 2017 Jul. PMID: 28753627 [Free PMC Article](#) [Similar articles](#)

Results by year Download CSV

Related searches ciliopathy review nephronophthisis disease mechanisms of a ciliopathy

PMC Images search for Ciliopathy See more (422)...

Titles with your search terms An elusive ciliopathy: Joubert syndrome. [BMJ Case Rep. 2017] Mutations in ARMC9, which Encodes a Basal Body Protein, Cause Jci [Am J Hum Genet. 2017] Diagnostic use of computational retrotransposon detection: Successful [Am J Med Genet A. 2017]

Refine the Query

- Often a search finds too many (or too few) sequences.
- More (or fewer) keywords in your query.
- Search only Title or Abstract.
- Search with an author name.
- Limit a query to specific organisms, a specific period of time, a journal (Cell or Science) etc.

A search with (crispr[Title/Abstract]) AND koonin e[Author]

NCBI Resources How To

PubMed.gov US National Library of Medicine National Institutes of Health

(crispr[Title/Abstract]) AND koonin e[Author]

Create RSS Create alert Advanced

Search

Help

Article types Clinical Trial Review Customize ...

Text availability Abstract Free full text Full text

PubMed Commons Reader comments Trending articles

Publication dates 5 years 10 years Custom range...

Species Humans Other Animals

[Clear all](#)

[Show additional filters](#)

Format: Summary Sort by: Most Recent Per page: 20

Send to Filters: Manage Filters

Find related data Database: Select

Find items

Search details

crispr[Title/Abstract] AND koonin e[Author]

Search See more...

Recent Activity Turn Off Clear

(crispr[Title/Abstract]) AND koonin e[Author] (53) PubMed

Ciliopathy (3931) PubMed

c9orf72 (957) PubMed

ciliopathies (3665) PubMed

See more...

Search results

Items: 1 to 20 of 53 << First < Prev Page 1 of 3 Next > Last >>

1. Peters JE, Makarova KS, Shmakov S, Koonin EV. Proc Natl Acad Sci U S A. 2017 Aug 15; pii: 201709035. doi: 10.1073/pnas.1709035114. [Epub ahead of print] PMID: 28811374 [Similar articles](#)

2. On the Origin of Reverse Transcriptase-Using CRISPR-Cas Systems and Their Hyperdiverse, Enigmatic Spacer Repertoires. Silas S, Makarova KS, Shmakov S, Páez-Espino D, Mohr G, Liu Y, Davison M, Roux S, Krishnamurthy SR, Fu BXH, Hansen LL, Wang D, Sullivan MB, Millard A, Clokie MR, Bhaya D, Lambowitz AM, Kyriides NC, Koonin EV, Fire AZ. MBio. 2017 Jul 11;8(4). pii: e00897-17. doi: 10.1128/mBio.00897-17. PMID: 28698278 [Free PMC Article](#) [Similar articles](#)

3. Polintons, virophages and transpovirons: a tangled web linking viruses, transposons and immunity. Koonin EV, Krupovic M. Curr Opin Virol. 2017 Jun 30;25:7-15. doi: 10.1016/j.coviro.2017.06.008. [Epub ahead of print] Review. PMID: 28672161 [Similar articles](#)

4. Evolutionary Genomics of Defense Systems in Archaea and Bacteria. Koonin EV, Makarova KS, Wolf YI. Annu Rev Microbiol. 2017 Jun 28. doi: 10.1146/annurev-micro-090816-093830. [Epub ahead of print] PMID: 28657885 [Similar articles](#)

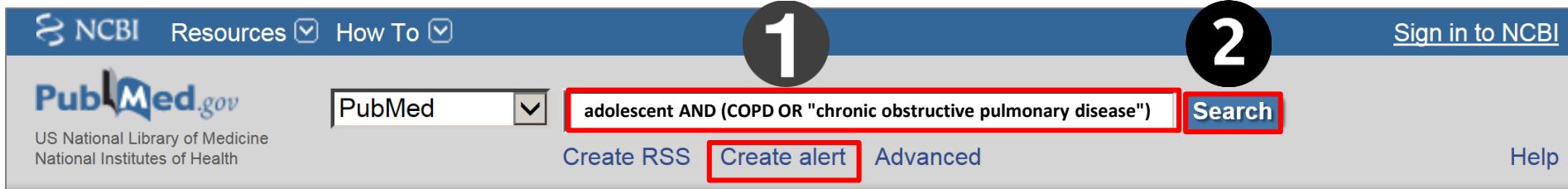
5. Diversity classification and evolution of CRISPR-Cas systems. Koonin EV, Makarova KS, Zhang F. Curr Opin Microbiol. 2017 Jun 9;37:67-78. doi: 10.1016/j.mib.2017.05.008. [Epub ahead of print] Review. PMID: 28605718 [Similar articles](#)

6. Casposons: mobile genetic elements that gave rise to the CRISPR-Cas adaptation machinery. Krupovic M, Béguin P, Koonin EV.

Search Strategies

- Start with broad keywords and narrow the search using more specific terms.
- Try variants of spelling, numbers, etc.
 - Ciliopathy, Ciliopathies, ciliary proteins or Cilio*
 - Maelstrom, MAEL, piRNA, piwi-RNA
- **Be persistent !!**

Creating Alerts



Example: Adolescent COPD
(chronic obstructive pulmonary disease)

Accessing My NCBI <https://www.ncbi.nlm.nih.gov/account>

Sign in to NCBI

Sign in with

 Google  Login  Commons

[See more 3rd party sign in options](#)

OR

Sign in directly to NCBI

Keep me signed in

Sign In

[Forgot NCBI username or password?](#)

[Register for an NCBI account](#)

[My NCBI](#) » Saved Searches

Your PubMed search

Name of saved search: **Adolescent COPD**

Search terms: adolescent AND (COPD OR "chronic obstructive pulmonary disease")

[Test search terms](#)

Would you like e-mail updates of new search results?

- No, thanks.
 Yes, please.

E-mail: sherri.bailey@nih.gov ([change](#))

Schedule:

Frequency: **Monthly**Which day? **the first Sunday**

Formats:

Report format: **Summary**

Number of items:

Send at most: **5 items** Send even when there aren't any new results

Any text you want to be added at the top of your e-mail (optional):

Name of saved search:

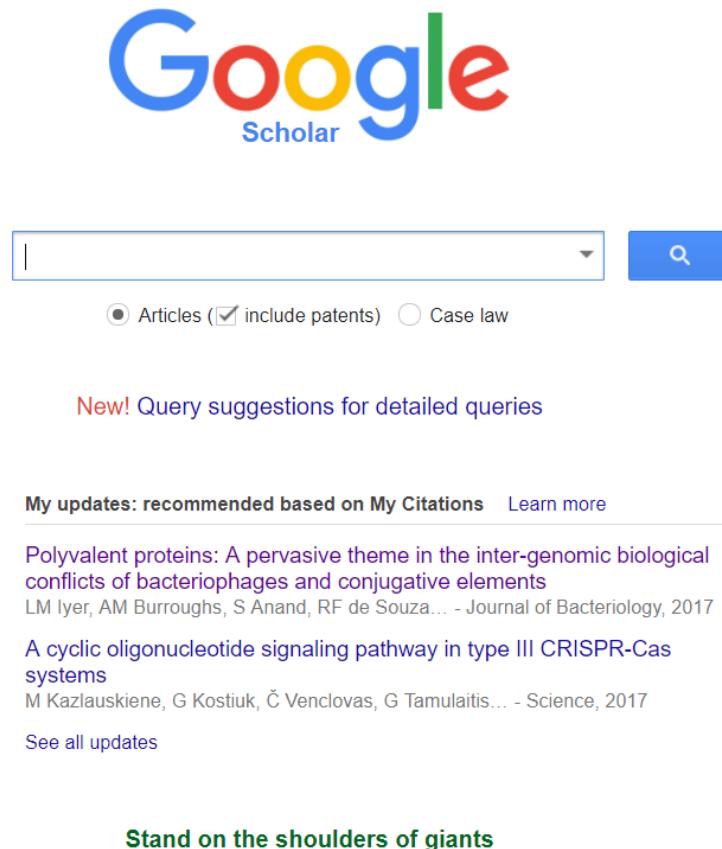
- Up to 100 characters allowed.
- The name will appear in the Subject line of the automatic email.

Save**Cancel**Skip saving and [return to your search](#), or proceed to [manage your Saved Searches](#).

Google Scholar

- Enables you to search specifically for scholarly literatures, including peer-reviewed papers, theses, books, preprints, patents, abstracts and technical reports from all broad areas of research.

 My library  My Citations  My updates  Alerts  Metrics  Settings



The screenshot shows the Google Scholar homepage. At the top is the Google Scholar logo. Below it is a search bar with a dropdown arrow and a blue search button with a magnifying glass icon. Under the search bar are two radio buttons: one selected for "Articles" and another for "Case law". A checked checkbox next to "Articles" includes "patents". Below the search area is a message about "Query suggestions for detailed queries". A horizontal line separates this from the "My updates" section. The "My updates" section header says "My updates: recommended based on My Citations" with a "Learn more" link. It lists two recent publications: "Polyvalent proteins: A pervasive theme in the inter-genomic biological conflicts of bacteriophages and conjugative elements" by LM Iyer, AM Burroughs, S Anand, RF de Souza... - Journal of Bacteriology, 2017, and "A cyclic oligonucleotide signaling pathway in type III CRISPR-Cas systems" by M Kazlauskienė, G Kostiuk, Č Venclovas, G Tamulaitis... - Science, 2017. A "See all updates" link is at the bottom of this section. At the very bottom is the tagline "Stand on the shoulders of giants".

- Google Scholar orders your search results by **how relevant** they are to your query, so the most useful references should appear at **the top of the page**.
- This relevance ranking takes into account the: full text of each article, the article's author, the publication in which the article appeared and **how often it has been cited** in scholarly literature.

Google c9orf72

Scholar About 8,400 results (0.08 sec)

Articles [HTML] Expanded GGGGCC hexanucleotide repeat in noncoding region of **C9ORF72** causes chromosome 9p-linked FTD and ALS M DeJesus-Hernandez, JR Mackenzie, BF Boeve... - *Neuron*, 2011 - Elsevier Several families have been reported with autosomal-dominant frontotemporal dementia (FTD) and amyotrophic lateral sclerosis (ALS), genetically linked to chromosome 9p21. Here, we report an expansion of a noncoding GGGGCC hexanucleotide repeat in the gene Cited by 2089 Related articles All 11 versions Web of Science: 1470 Cite Save

Case law [HTML] A hexanucleotide repeat expansion in **C9ORF72** is the cause of chromosome 9p21-linked ALS-FTD AE Renton, E Majounie, A Waite, J Simón-Sánchez... - *Neuron*, 2011 - Elsevier The chromosome 9p21 amyotrophic lateral sclerosis-frontotemporal dementia (ALS-FTD) locus contains one of the last major unidentified autosomal-dominant genes underlying these common neurodegenerative diseases. We have previously shown that a founder Cited by 1948 Related articles All 20 versions Web of Science: 1374 Cite Save

My library Any time [HTML] The **C9orf72** GGGGCC repeat is translated into aggregating dipeptide-repeat proteins in FTLD/ALS K Mori, SM Weng, T Arzberger, S May... - ..., 2013 - science.sciencemag.org Expansion of a GGGGCC hexanucleotide repeat upstream of the **C9orf72** coding region is the most common cause of familial frontotemporal lobar degeneration and amyotrophic lateral sclerosis (FTLD/ALS), but the pathomechanisms involved are unknown. As in other Cited by 491 Related articles All 14 versions Web of Science: 365 Cite Save

Sort by relevance [HTML] Frequency of the **C9orf72** hexanucleotide repeat expansion in patients with amyotrophic lateral sclerosis and frontotemporal dementia: a cross-sectional study E Majounie, AE Renton, K Mok, EGP Dopper... - *The Lancet* ..., 2012 - Elsevier BACKGROUND: We aimed to accurately estimate the frequency of a hexanucleotide repeat expansion in **C9orf72** that has been associated with a large proportion of cases of amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD). METHODS: We Cited by 507 Related articles All 31 versions Web of Science: 393 Cite Save

Sort by date [HTML] Unconventional translation of **C9ORF72** GGGGCC expansion generates insoluble polypeptides specific to c9FTD/ALS PEA Ash, KF Bieniek, TF Gendron, T Caulfield, WL Lin... - *Neuron*, 2013 - Elsevier Frontotemporal dementia (FTD) and amyotrophic lateral sclerosis (ALS) are devastating

include patents
 include citations
 Create alert

Google Scholar profiles for researchers

- label: biology

Profiles



Bert Vogelstein

Johns Hopkins
Verified email at jhmi.edu
Cited by 342080
Cancer Biology



Eric Lander

Broad Institute
Cited by 327318
Biology Genomics Genetics Bioinformatics Mathematics



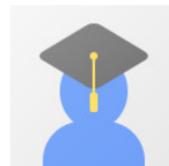
Fred H. Gage

The Salk Institute
Verified email at salk.edu
Cited by 165396
Biology



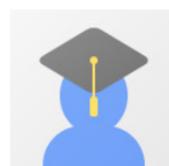
Peer Bork

EMBL
Verified email at embl.de
Cited by 163052
biology computational biology systems biology bioinformatics evolution



David Bartel

Professor of Biology, HHMI, MIT, Whitehead Institute
Verified email at wi.mit.edu
Cited by 138479
Biology Gene Regulation MicroRNAs Noncoding RNA Molecular Biology



K Barry Sharpless

The Scripps Research Institute
Verified email at scripps.edu
Cited by 133310
chemistry biology medicine materials science polymeric materials



Chris Sander

Dana-Farber Cancer Institute, Harvard Medical School, Memorial Sloan Kettering Cancer ...
Verified email at sanderlab.org
Cited by 125955
Biology Data Sciences Evolution Cancer Therapy



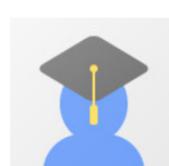
George P. Chrouzos

Professor, School of Medicine, National and Kapodistrian University of Athens
Verified email at med.uoa.gr
Cited by 123997
Clinical Medicine Biology Biochemistry



Eugene Koonin

Senior Investigator, NCBI, NIH
Verified email at ncbi.nlm.nih.gov
Cited by 143677
Biology



Mounji Bawendi

Lester Wolfe Professor of Chemistry, Massachusetts Institute of Technology
Cited by 95754
Chemistry Quantum Dots Biology Photovoltaics Spectroscopy

Google Scholar

- Try “label: Bioinformatics or Computational Biology” to see the pioneers in the field.
- Identify the classical research papers in the field that you are interested.
- Establish your own google scholar profiles.

Web of Science

- <https://www.webofknowledge.com/>
- Science Citation Index (SCI)
- Impact Factor (IF): On average, how many times an article in this journal is being cited – based on articles published in the two previous years
- ResearchID

The screenshot shows the Web of Science search interface. At the top, there's a navigation bar with links to 'Web of Science', 'InCites', 'Journal Citation Reports', 'Essential Science Indicators', 'EndNote', 'Publons', 'Sign In', 'Help', and 'English'. Below the navigation is the 'Web of Science' logo and the Clarivate Analytics logo. The main search area has a 'Search' bar with 'c9orf72' typed in, a dropdown for 'Select a database' set to 'Web of Science Core Collection', and a 'Basic Search' tab selected. To the right of the search bar are buttons for 'Learn More', 'My Tools', 'Search History', and 'Marked List'. A callout 'Check out the new citation report.' is visible. Below the search bar, there are fields for 'Title' and buttons for '+ Add Another Field' and 'Reset Form'. A 'Search' button is also present. Further down, there are sections for 'TIMESPAN' (with 'All years' selected) and 'MORE SETTINGS'. At the bottom, there are links for 'Customer Feedback & Support', 'Additional Resources', 'What's New in Web of Science?', and 'Customize your Experience'.

Bioinformatics Journals

- Bioinformatics
 - PLoS Computational Biology
 - BMC Bioinformatics
 - Briefings in Bioinformatics
 - Journal of Computational Biology
 - Molecular Systems Biology
 - PNAS
 - Nucleic Acids Research
 - Genome Research
 - Genome Biology
 - Nature Methods
 - Nature Genetics
 - Science/Nature/Cell
 - Journal of Molecular Biology
 - BMC Genomics
 - Molecular Biology and Evolution
 - Genome Biology and Evolution
 - BMC Evolutionary Biology
 - Biology Direct
 - Proteins: Structure, Function, and Bioinformatics
 - Gene
 - BMC Systems Biology
 - Frontiers in Genetics
 - Molecular & Cellular Proteomics
 - Journal of Theoretical Biology
 - Nature Communications
 - Nature Ecology & Evolution
- More journals: <http://www.bioinformatics.org/wiki/journals>

Categories of databases for Life Sciences

- Sequences (DNA, protein)
- Genomics
- Mutation/polymorphism
- Protein domain/family
- Proteomics (2D gel, Mass Spectrometry)
- 3D structure
- Metabolic networks
- Regulatory networks
- Bibliography
- Expression
- Specialized

NCBI---National Center for Biotechnology Information (www.ncbi.nlm.nih.gov)

The screenshot shows the NCBI homepage with a blue header bar containing the NCBI logo, a search bar, and user account information (dapengzhangca@gmail.com). The main content area features a sidebar with links to various biological databases and tools. The central part of the page is titled "Welcome to NCBI" and includes sections for "Submit", "Download", "Learn", "Develop", "Analyze", and "Research". Each section has a brief description and a corresponding icon. To the right, there is a "Popular Resources" sidebar listing links to PubMed, Bookshelf, PubMed Central, PubMed Health, BLAST, Nucleotide, Genome, SNP, Gene, Protein, and PubChem. Below that is a "NCBI News & Blog" section with recent articles.

NCBI Resources How To

dapengzhangca@gmail.com My NCBI Sign Out

NCBI National Center for Biotechnology Information

All Databases Search

NCBI Home

Resource List (A-Z)

All Resources

Chemicals & Bioassays

Data & Software

DNA & RNA

Domains & Structures

Genes & Expression

Genetics & Medicine

Genomes & Maps

Homology

Literature

Proteins

Sequence Analysis

Taxonomy

Training & Tutorials

Variation

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 | NCBI News & Blog

Submit Deposit data or manuscripts into NCBI databases

Download Transfer NCBI data to your computer

Learn Find help documents, attend a class or watch a tutorial

Develop Use NCBI APIs and code libraries to build applications

Analyze Identify an NCBI tool for your data analysis task

Research Explore NCBI research and collaborative projects

Popular Resources

PubMed

Bookshelf

PubMed Central

PubMed Health

BLAST

Nucleotide

Genome

SNP

Gene

Protein

PubChem

NCBI News & Blog

Yellow fever mosquito, 6 other organisms in July RefSeq genome annotations 17 Aug 2017

In July, the NCBI Eukaryotic Genome Annotation Pipeline released new

August 30 NCBI Minute: NCBI Hackathons, a framework for rapid prototyping of pipelines for bioinformatics, biomedical informatics and genomics

- Develop and maintain molecular and bibliographic databases.
- Develop software for searching, and analysis of these data.
- Provide Web access point for data and software.

NCBI Home

Resource List (A-Z)

All Resources

Chemicals & Bioassays

Data & Software

DNA & RNA

Domains & Structures

Genes & Expression

Genetics & Medicine

Genomes & Maps

Homology

Literature

Proteins

Sequence Analysis

Taxonomy

Training & Tutorials

Variation

✓ All Databases

Assembly

BioProject

BioSample

BioSystems

Books

ClinVar

Clone

Conserved Domains

dbGaP

dbVar

Epigenomics

EST

Gene

Genome

GEO DataSets

GEO Profiles

GSS

GTR

HomoloGene

MedGen

MeSH

NCBI Web Site

NLM Catalog

Nucleotide

OMIM

PMC

PopSet

Probe

Protein

Protein Clusters

PubChem BioAssay

PubChem Compound

PubChem Substance

PubMed

PubMed Health

SNP

SRA

Structure

Taxonomy

ToolKit

ToolKitAll

ToolKitBook

UniGene

Welcome to NCBI

National Center for Biotechnology Information.
Information, Tools and Resources for Life Sciences.

[NCBI](#) | [Mission](#) | [Organization](#) | [NCBI Home](#)

Submit

Upload your data or manuscripts into our molecular databases



Develop

Access our APIs and code libraries to build your own applications



Molecular Databases in NCBI

- Sequences: Nucleotide, Protein
- Gene annotation: Gene, RefSeq
- Genomes: Assembly, Genome
- Expression: GEO, EST
- Protein Domains: CDD
- Homologous Genes
- Genetic Variation: SNP, ClinVar, dbVar
- Taxonomy
- 3D Structures
- Pathways
- Literature: PubMed, PMC
- Small molecules: PubChem

NCBI Home

Resource List (A-Z)

All Resources

Chemicals & Bioassays

Data & Software

DNA & RNA

Domains & Structures

Genes & Expression

Genetics & Medicine

Genomes & Maps

Homology

Literature

Proteins

Sequence Analysis

Taxonomy

Training & Tutorials

Variation

✓ All Databases

Assembly

BioProject

BioSample

BioSystems

Books

ClinVar

Clone

Conserved Domains

dbGaP

dbVar

Epigenomics

EST

Gene

Genome

GEO DataSets

GEO Profiles

GSS

GTR

HomoloGene

MedGen

MeSH

NCBI Web Site

NLM Catalog

Nucleotide

OMIM

PMC

PopSet

Probe

Protein

Protein Clusters

PubChem BioAssay

PubChem Compound

PubChem Substance

PubMed

PubMed Health

SNP

SRA

Structure

Taxonomy

ToolKit

ToolKitAll

ToolKitBook

UniGene

Welcome to NCBI

National Center for Biotechnology Information.
Information, Tools and Resources for Life Sciences.

[NCBI](#) | [Mission](#) | [Organization](#) | [NCBI Home](#)

Submit

Upload your data or manuscripts into our databases.



Develop

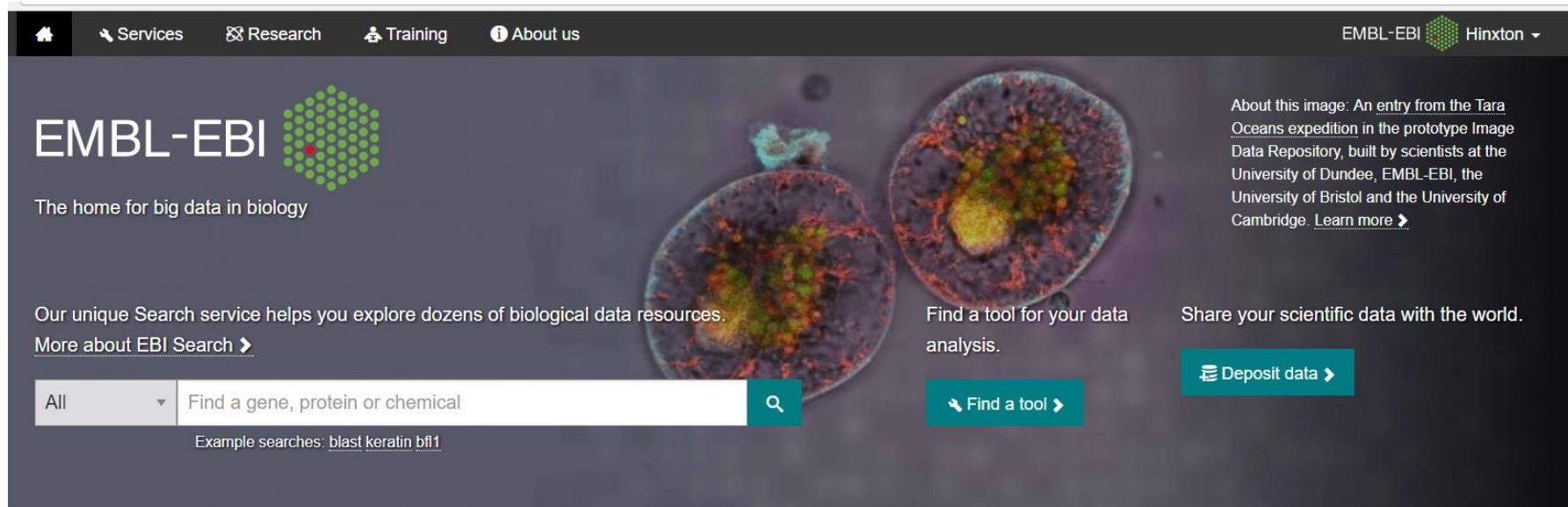
APIs and code libraries to build your applications.



NCBI Search Services and Tools

- Entrez integrated literature and molecular databases
- BLAST sequence similarity search service
- Graphical Sequence Viewer annotation viewer and analysis tool
- Genome Workbench standalone sequence analysis annotation platform
- SRA Utilities
 - SRA Run Browser: web access for viewing, searching and downloading next generation reads
 - SRA toolkit: standalone SRA manipulator and client

EBI---European Bioinformatics Institute (www.ebi.ac.uk/)



Explore EMBL-EBI and our mission

The European Bioinformatics Institute (EMBL-EBI) shares [data from life science experiments](#), performs [basic research](#) in computational biology and offers an extensive [user training programme](#), supporting researchers in academia and industry. We are part of [EMBL](#), Europe's flagship laboratory for the life sciences.

[More about EMBL-EBI and our impact](#)

Services

We provide freely available data and bioinformatics services to all facets of the scientific community [►](#)

Research

We contribute to the advancement of biology through basic investigator-driven research [►](#)

Training

We provide advanced bioinformatics training to scientists at all levels [►](#)

Industry

We help disseminate cutting-edge technologies to industry [►](#)

ELIXIR

We support, as an ELIXIR node, the coordination of biological data provision throughout Europe [►](#)

Many server-based tools:

- ClustalW
- dbfetch
- EMBOSS transeq
- EMBOSS backtranambig
- EMBOSS cpgplot
- EMBOSS needle
- EMBOSS seqret
- FASTA
- HMMER
- Kalign
- Muscle
- Mview
- PDBeFold
- Phobius
- Blast/Psi-Blast
- RADAR
- Reactome Pathways Analysis

Tools & Databases

Search all tools & databases

Tools

Assembly converter
Map your data to the current assembly. Based on the CrossMap tool

Read mapping

BiNChE
Chemical class enrichment analysis based on the ChEBI ontology

Chemical class enrichment

BioModels Database Converters
Set of converters between various modelling formats, there are also some available online to be used directly

Formatting

BLAST [nucleotide]
Fast local similarity search tool for nucleotide sequence databases.

Sequence similarity search

BLAST [protein]
Fast local similarity search tool for protein sequence databases.

Sequence similarity search

Bubastis
Tool to analyse two ontologies (typically two versions of the same ontology) to highlight differences

Ontology comparison

ChEMBL Biologicals BLAST Search
BLAST search against biological molecules (e.g. mAbs) stored in ChEMBL database

Sequence similarity search

Databases

ArrayExpress
A database of functional genomics experiments, including microarray and RNAseq expression data typically related to publications.

BioModels
A repository of peer-reviewed, published, computational models.

BioSamples
A database describing biological samples and providing links to associated experimental data

BioStudies Database
The BioStudies database holds descriptions of biological studies, links to data from these studies in other databases at EMBL-EBI or outside, as well as data that do not fit in the structured archives at EMBL-EBI. For example, BioStudies can hold supplementary material linked to published papers.

ChEBI
Reference chemical structures, nomenclature and ontological classification.

ChEMBL
An open data resource of binding, functional and ADMET bioactivity data.

Complex Portal
A manually curated, encyclopaedic resource of macromolecular complexes from model organisms.

DGVA
A repository that provides archiving, accessioning and distribution of publicly available genomic structural variants, in all species.

EBI Search
A biologically aware search that makes it easier to navigate results across different domains, via a network of cross-references.

Many (manually curated) databases:

- ArrayExpress
- DGVA
- Ensembl
- G2Pdatabase
- GWAS Catalog
- IntAct
- InterPro
- Pfam
- Rfam
- PhytoPath
- Reactome
- UniProt
- WormBase
- PRIDE
- Metagenomics

UniProtKB/SwissProt

(<http://expasy.cbr.nrc.ca/sprot/>)

- Curation!!!
 - Error rate in the information is greatly reduced in comparison to most other databases.
- Extensive cross-linking to other data sources
- SwissProt is the ‘gold-standard’ by which other databases can be measured, and is the best place to start if you have a specific protein to investigate

UCSC Genome Browser

<https://genome.ucsc.edu/index.html>



The screenshot shows the UCSC Genome Browser homepage. At the top, the UCSC logo and name are displayed, along with a stylized bridge graphic. Below the header is a navigation bar with links: Genomes, Genome Browser, Tools, Mirrors, Downloads, My Data, Help, and About Us. A large blue banner in the center features a 3D rendering of several glowing blue DNA double helix molecules against a dark background. To the right of the banner is a yellow sidebar titled "Our tools" which lists various genomic analysis tools. Below the banner, there are two sections: "Our story" and "What's new".

Our tools

- **Genome Browser**
interactively visualize genomic data
- **BLAT**
rapidly align sequences to the genome
- **Table Browser**
download data from the Genome Browser database
- **Variant Annotation Integrator**
get functional effect predictions for variant calls
- **Data Integrator**
combine data sources from the Genome Browser database
- **Gene Sorter**
find genes that are similar by expression and other metrics
- **Genome Browser in a Box (GBiB)**
run the Genome Browser on your laptop or server
- **In-Silico PCR**
rapidly align PCR primer pairs to the genome
- **LiftOver**
convert genome coordinates between assemblies
- **VisiGene**
interactively view *in situ* images of mouse and frog

[More tools...](#)

Our story

On June 22, 2000, UCSC and the other members of the International Human Genome Project consortium completed the first working draft of the human genome assembly, forever ensuring free public access to the genome and the information it contains. A few weeks later, on July 7, 2000, the newly assembled genome was released on the web at <http://genome.ucsc.edu>.

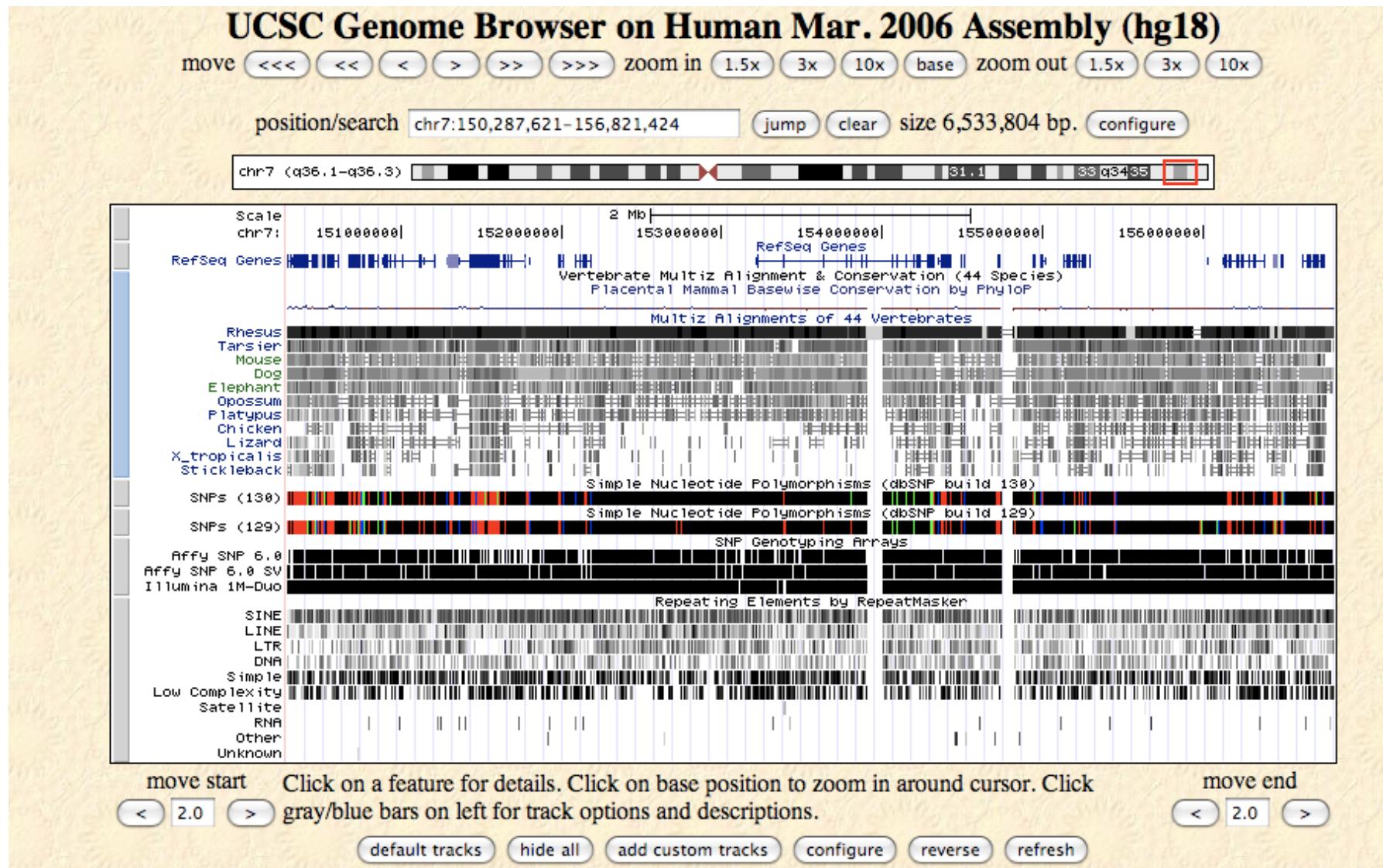
What's new

Aug. 17, 2017 - [New CRISPR track for Chinese hamster](#)

Jul. 14, 2017 - [New alternative mouse strain hub and strain-specific annotations](#)

Jul. 14, 2017 - [New video: View discontinuous regions together](#)

UCSC Genome Browser--Genome annotation



PDB---Protein Data Bank (www.rcsb.org)

RCSB PDB Deposit Search Visualize Analyze Download Learn More MyPDB Login

RCSB PDB PROTEIN DATA BANK An Information Portal to 132905 Biological Macromolecular Structures

Search by PDB ID, author, macromolecule, sequence, or ligands Go Advanced Search | Browse by Annotations

PDB-101 Worldwide Protein Data Bank Foundation

EMDataBank Unified Data Resource for 3D Structure

Nucleic Acid Database

Welcome

Deposit

Search

Visualize

Analyze

Download

Learn

A Structural View of Biology

This resource is powered by the Protein Data Bank archive-information about the 3D shapes of proteins, nucleic acids, and complex assemblies that helps students and researchers understand all aspects of biomedicine and agriculture, from protein synthesis to health and disease.

As a member of the wwPDB, the RCSB PDB curates and annotates PDB data. The RCSB PDB builds upon the data by creating tools and resources for research and education in molecular biology, structural biology, computational biology, and beyond.

Video: How Enzymes Work

VIDEO: HOW ENZYMES WORK

August Molecule of the Month

Glutathione Transferases

BioGRID--- an interaction repository (thebiogrid.org)

- Data compiled through comprehensive curation efforts:
 - **63,354** publications for **1,493,749** protein and genetic interactions
 - **27,785** chemical associations
 - **38,559** post translational modifications from major model organism species.

The screenshot shows the BioGRID 3.4 homepage with a dark red header and a white content area. The header includes a navigation bar with links: home, help, wiki, tools, contribute, stats, downloads, partners, about us, and a Twitter icon. On the right side of the header, there are two vertical buttons: 'By Gene' (green) and 'By Publication' (blue). The main content area features a large title 'Welcome to the Biological General Repository for Interaction Datasets'. Below the title is a paragraph describing the database's purpose and statistics. At the bottom of this section are two buttons: 'INTERACTION STATISTICS' and 'LATEST DOWNLOADS'. To the right of the main content area is a search sidebar titled 'Search the BioGRID' with a search input field, a dropdown menu for 'All Organisms', and a 'SUBMIT GENE SEARCH Q' button. Below the search sidebar are three links: 'Advanced Search', 'Search Tips', and 'Featured Datasets'. The bottom right corner of the page contains a section titled 'BIOGRID FUNDING AND PARTNERS' with logos for NIH, CIHR IRSC, Genome Québec, Mount Sinai Hospital, Princeton University, Université de Montréal, SGD, University of Edinburgh, and IMEx, along with a link to 'more partners'.

BioGRID 3.4

home help wiki tools contribute stats downloads partners about us |

Welcome to the Biological General Repository for Interaction Datasets

BioGRID is an interaction repository with data compiled through comprehensive curation efforts. Our current index is version 3.4.151 and searches 63,354 publications for 1,493,749 protein and genetic interactions, 27,785 chemical associations and 38,559 post translational modifications from major model organism species. All data are freely provided via our search index and available for download in standardized formats.

[INTERACTION STATISTICS](#) [LATEST DOWNLOADS](#)

AREAS OF INTEREST TO HELP YOU GET STARTED

Build and Download Interaction Datasets
Create custom interaction datasets by protein or by publication. You can also download our entire dataset in a wide variety of standard formats.

Link To Us or Submit Interactions
Send us your datasets or link to the BioGRID directly from your own website or database. Full details on how to contribute are available [here](#).

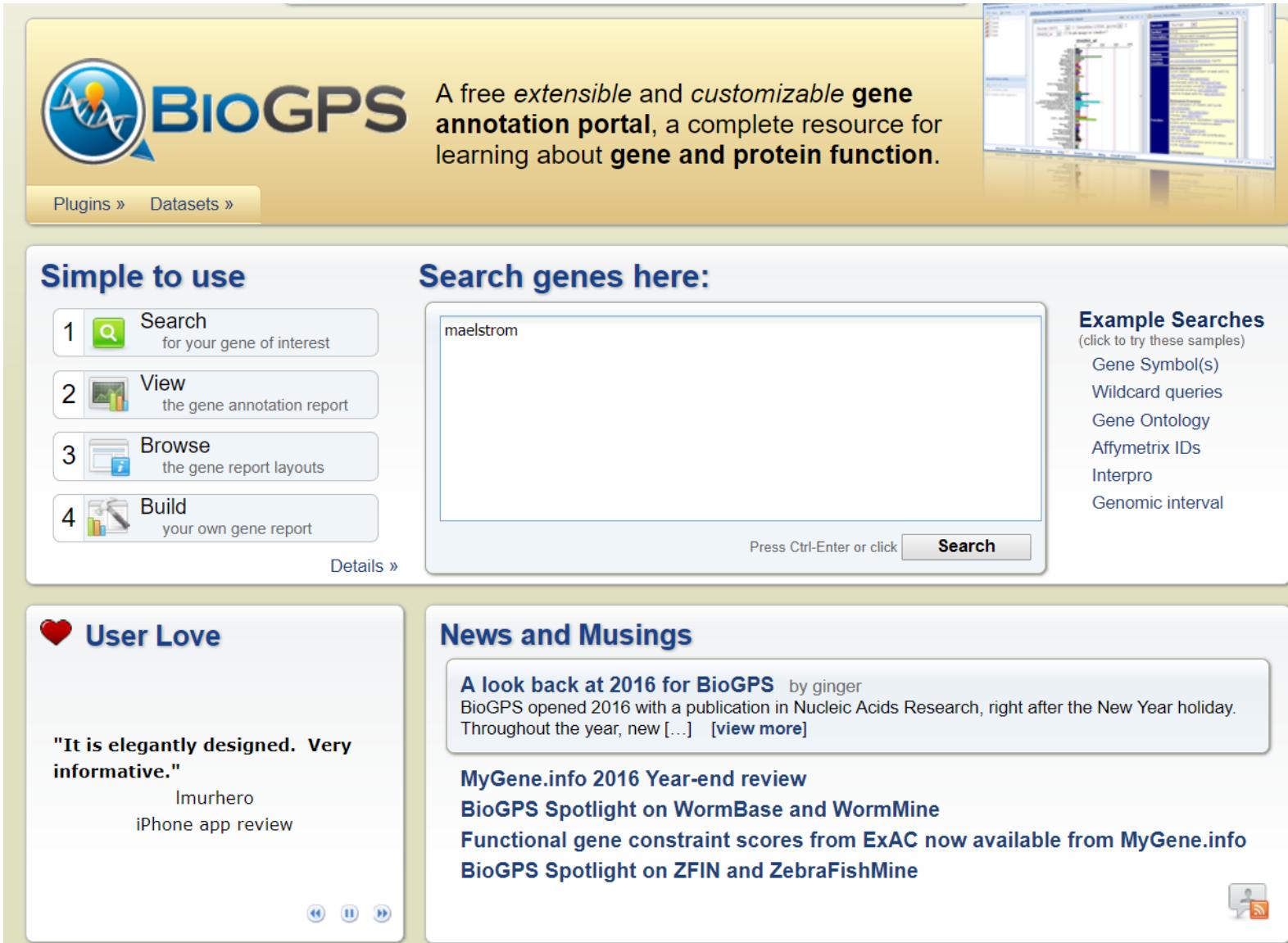
Online Tools and Resources
We've developed tools that make use of BioGRID data. Check out the list of tools to see if we can help you work with our data.

View Our Interaction Statistics
Find out how many organisms, proteins, publications, and interactions are available in the current release of the BioGRID.

BIOGRID FUNDING AND PARTNERS

[more partners](#)

BioGPS (<http://biogps.org>)



The screenshot shows the BioGPS homepage with the following content:

- Header:** BioGPS logo, tagline "A free extensible and customizable gene annotation portal, a complete resource for learning about gene and protein function.", and links for "Plugins »" and "Datasets »".
- Simple to use:** A list of four steps:
 - Search for your gene of interest
 - View the gene annotation report
 - Browse the gene report layouts
 - Build your own gene report
- Search genes here:** A search bar containing "maelstrom", a "Search" button, and a note "Press Ctrl-Enter or click".
- Example Searches:** Links to "Gene Symbol(s)", "Wildcard queries", "Gene Ontology", "Affymetrix IDs", "Interpro", and "Genomic interval".
- User Love:** A testimonial from "Imurhero" (iPhone app review) stating "It is elegantly designed. Very informative."
- News and Musings:** A summary of 2016, a link to the "MyGene.info 2016 Year-end review", and links to "BioGPS Spotlight on WormBase and WormMine", "Functional gene constraint scores from ExAC now available from MyGene.info", and "BioGPS Spotlight on ZFIN and ZebrafishMine".
- Bottom right corner:** Social media icons for Twitter and RSS.

Pfam--- Protein Domain database (pfam.xfam.org/)

[HOME](#) | [SEARCH](#) | [BROWSE](#) | [FTP](#) | [HELP](#) | [ABOUT](#)



Pfam 31.0 (March 2017, 16712 entries)

The Pfam database is a large collection of protein families, each represented by **multiple sequence alignments** and **hidden Markov models (HMMs)**. [More...](#)

QUICK LINKS

[SEQUENCE SEARCH](#)

[VIEW A PFAM ENTRY](#)

[VIEW A CLAN](#)

[VIEW A SEQUENCE](#)

[VIEW A STRUCTURE](#)

[KEYWORD SEARCH](#)

[JUMP TO](#)

YOU CAN FIND DATA IN PFAM IN VARIOUS WAYS...

Analyze your protein sequence for Pfam matches

View Pfam annotation and alignments

See groups of related entries

Look at the domain organisation of a protein sequence

Find the domains on a PDB structure

Query Pfam by keywords

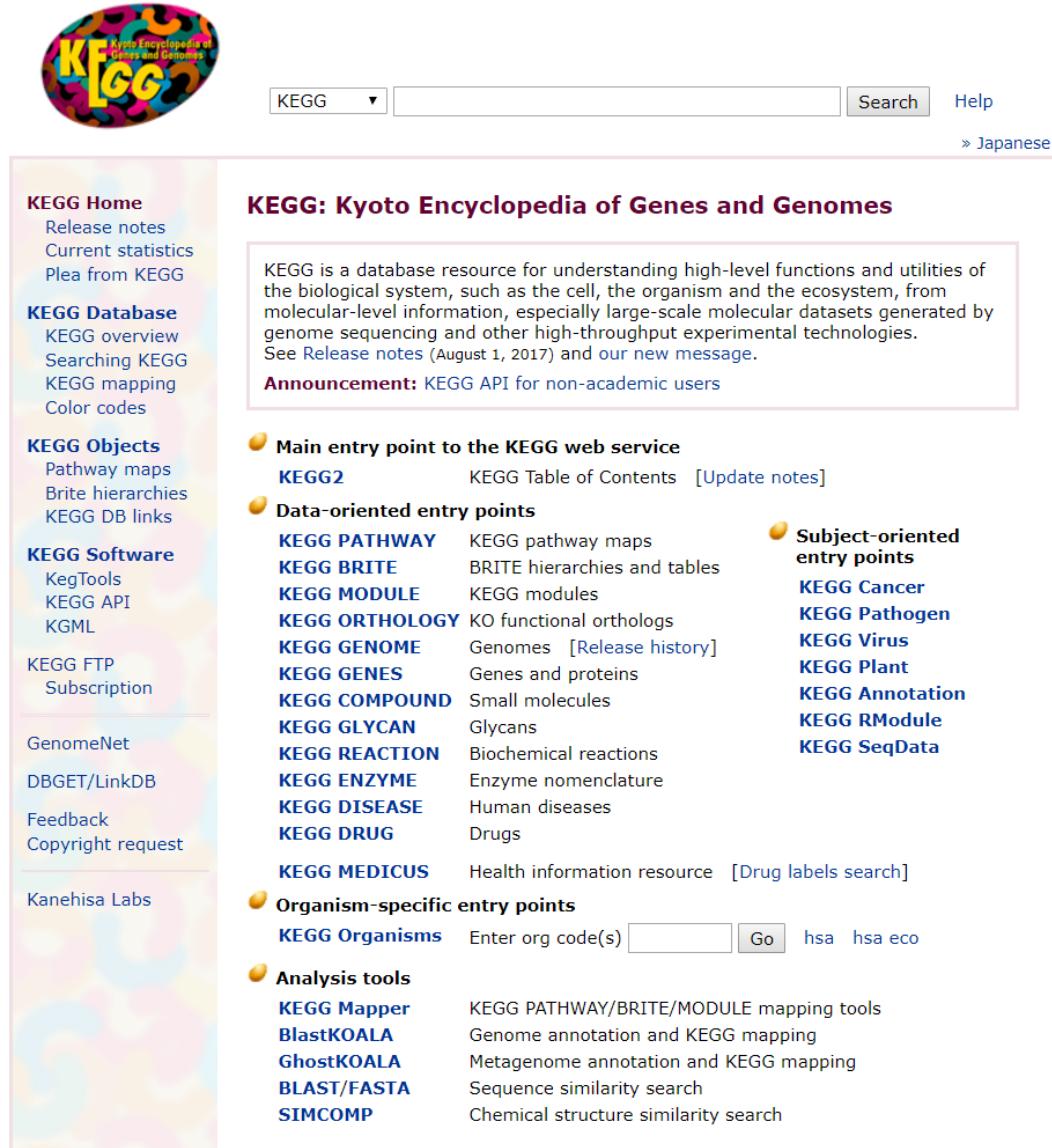
[Go](#)

[Example](#)

Enter any type of accession or ID to jump to the page for a Pfam entry or clan, UniProt sequence, PDB structure, etc.

Or view the [help](#) pages for more information

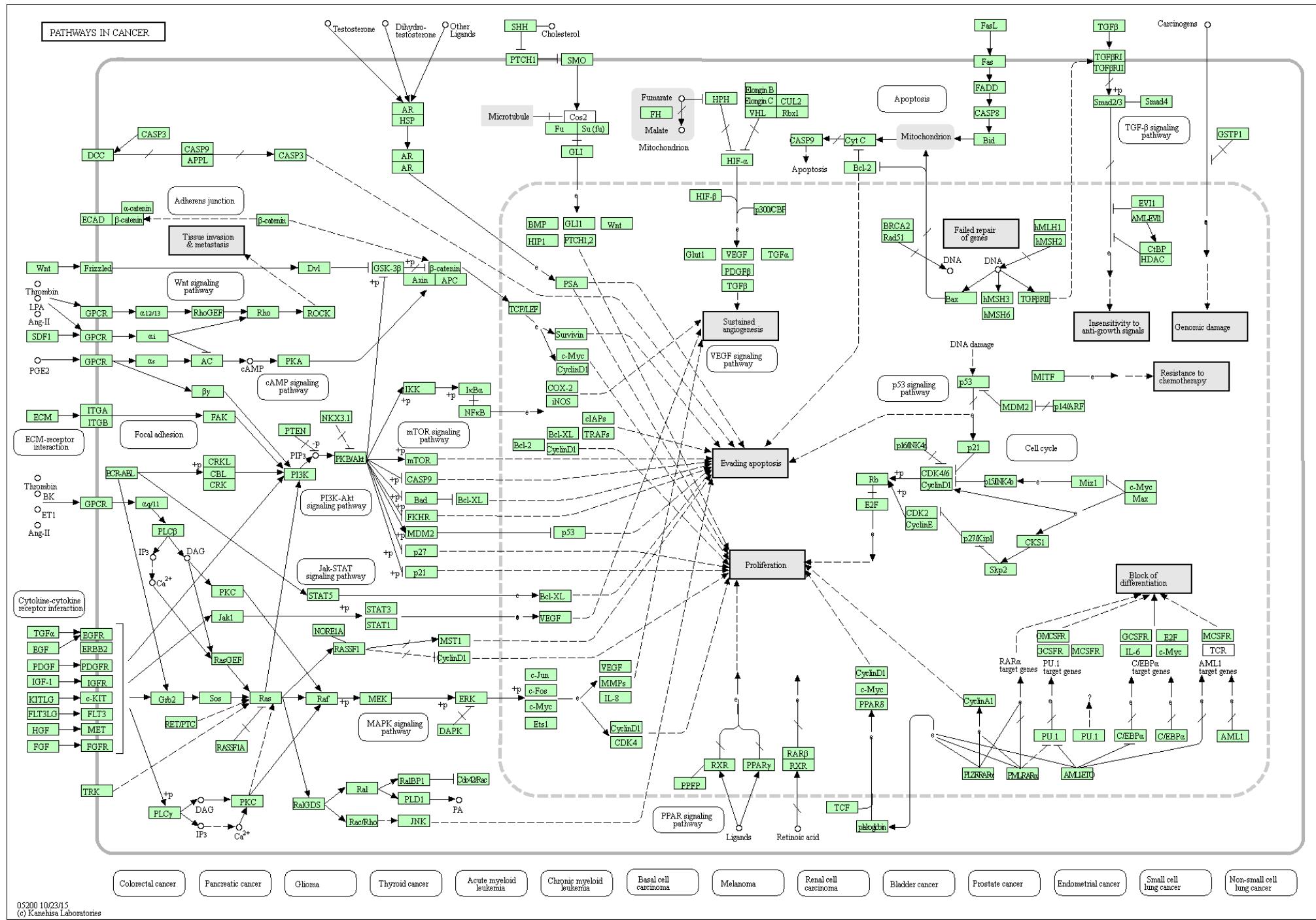
KEGG (www.genome.jp/kegg/)



The screenshot shows the KEGG homepage with a navigation bar at the top. The navigation bar includes a logo for "Kyoto Encyclopedia of Genes and Genomes" (KEGG), a search input field, and links for "KEGG", "Search", and "Help". Below the navigation bar is a link to "» Japanese". On the left side, there is a vertical sidebar with several sections: "KEGG Home" (Release notes, Current statistics, Plea from KEGG), "KEGG Database" (KEGG overview, Searching KEGG, KEGG mapping, Color codes), "KEGG Objects" (Pathway maps, Brite hierarchies, KEGG DB links), "KEGG Software" (KegTools, KEGG API, KGML), "KEGG FTP" (Subscription), "GenomeNet", "DBGET/LinkDB", "Feedback", "Copyright request", and "Kanehisa Labs". The main content area features a section titled "KEGG: Kyoto Encyclopedia of Genes and Genomes" with a brief description of the database's purpose and a message about the KEGG API. Below this are several lists of entry points categorized by subject: "Main entry point to the KEGG web service" (KEGG2, KEGG Table of Contents [Update notes]), "Data-oriented entry points" (KEGG PATHWAY, KEGG BRITE, KEGG MODULE, KEGG ORTHOLOGY, KEGG GENOME, KEGG GENES, KEGG COMPOUND, KEGG GLYCAN, KEGG REACTION, KEGG ENZYME, KEGG DISEASE, KEGG DRUG, KEGG MEDICUS), "Subject-oriented entry points" (KEGG Cancer, KEGG Pathogen, KEGG Virus, KEGG Plant, KEGG Annotation, KEGG RModule, KEGG SeqData), and "Organism-specific entry points" (KEGG Organisms, Enter org code(s) Go, hsa, hsa eco). At the bottom, there is a section for "Analysis tools" (KEGG Mapper, BlastKOALA, GhostKOALA, BLAST/FASTA, SIMCOMP).

A database resource for understanding high-level functions and utilities of the biological system, such as the cell, the organism and the ecosystem.

KEGG



ExPASy (www.expasy.org)

- SIB Bioinformatics Resource Portal

The screenshot shows the homepage of the ExPASy portal. At the top left is the SIB logo (a red square with a white 'SIB' and a stylized mountain peak icon). Next to it is the ExPASy logo with the text 'Bioinformatics Resource Portal'. A search bar at the top right contains the placeholder 'Query all databases'. On the left, a vertical sidebar menu lists categories: Visual Guidance, Categories, proteomics, genomics, structure analysis, systems biology, evolutionary biology, population genetics, transcriptomics, biophysics, imaging, IT infrastructure, medicinal chemistry, and glycomics. Below this is a 'Resources A..Z' section with a question mark icon, and a 'Links/Documentation' section. The main content area features a 'Featuring today' box for 'ProtParam', which computes physical and chemical parameters for protein sequences. A 'How to use this portal?' section at the bottom left provides links for features, updates, new users, and experienced users.

Visual Guidance

Categories

proteomics

genomics

structure analysis

systems biology

evolutionary biology

population genetics

transcriptomics

biophysics

imaging

IT infrastructure

medicinal chemistry

glycomics

Resources A..Z

Links/Documentation

Visual Guidance

Categories

proteomics

genomics

structure analysis

systems biology

evolutionary biology

population genetics

transcriptomics

biophysics

imaging

IT infrastructure

medicinal chemistry

glycomics

Resources A..Z

Links/Documentation

Examples of individual Resources

	Analyses	Databases and tools
DNA/Genome	Gene prediction	GENSCAN, Glimmer
	Genetic and Somatic Variations	COSMIC, SIFT, PolyPhen, SAPRED
Expression regulation	Transcription factors & binding	TRANSFAC
	RNA annotation	Rfam
Protein sequence and expression	microRNA annotation	miRBase
	Protein annotation	UniProt
	Protein domain	PROSITE, STRING
Macromolecule structure	Mass spectrometry	The Global Proteome Machine
	Nucleotide structure	Mfold
	Nucleotide interaction	RNAhybrid
Epigenetics	Protein structure	PDB, Swiss Model
	DNA methylation	MethylomeDB
Pathway & network	Pathway	KOBAS, DAVID, KEGG, PANTHER, PID, BioCyc
Evolution	Conservation	GERP++, PHYML

Examples of Individual Resources for NGS analysis

	Analyses	Databases and tools
Read mapping	Reads mapping (DNA)	BWA, Bowtie/Bowtie2
	Reads mapping (RNA)	TopHat/TopHat2
	Reads mapping related utilities	GATK, FastQC, RNA-SeQC, samtools, picard
Assembly	de novo genomic assembly	Velvet, SOAP de novo
	de novo transcriptome assembly	Trinity, Velvet + Oases
	reference-based transcriptome assembly	TopHat + Cufflinks
Visualisation	Genome Browser	GBrowse, JBrowse, IGV
Variant calling	SNP	GATK, SOAPSnp
	Indel	Pindel
	CNV	CNVnator
	Structural variation	SVMerge
Expression	Differential expression	Cuffdiff, DESeq/DESeq2
ChIP-Seq	Peak calling	MACS

Examples of Individual Resources

Purpose		Databases and tools
Model Organisms	Genome and gene annotations	Flybase, Wormbase, ZFIN, TAIR
Large-scale studies	Cancer	TCGA, CGP
	Epigenetics	Roadmap Epigenomics Project
	Human population studies	GTEx
	Brain	Allen Brain Atlas, Human Connectome project
Tools to assist wet-lab experiments	Primer design	Primer3/Primer3Plus, Electronic PCR
Software programming utilities	Standalone	EMBOSS
	R package	Bioconductor
	Perl package	BioPerl
	Python package	BioPython
Workflow	Workflow platform	Galaxy
	Workflow construction	Taverna

Nucleic Acids Research Database Issue

- NAR Database Issue, First issue every year
 - <https://academic.oup.com/nar/issue/>



Explore the 2017
NAR Database Issue today
54 new database descriptions

Nucleic acid sequence,
structure, and regulation

Protein sequence and
structure, motifs, and
domains

Metabolic and signalling
pathways, enzymes

Viruses, bacteria, protozoa
and fungi

Human genome, model
organisms, comparative
genomics

Genomic variation, diseases,
and drugs

Plant databases

Other databases

Erratum

Front-Matter/Back-Matter