Bioinformatics & Biostatics

Lecture 2: Bioinformatic Databases



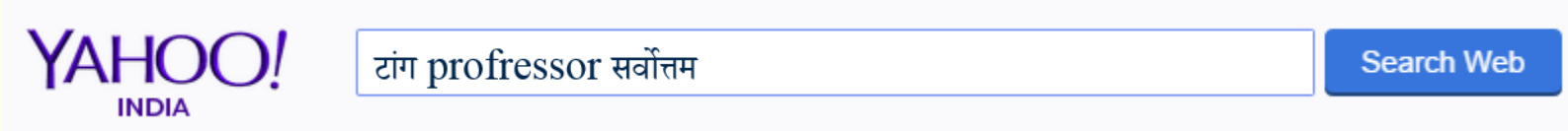




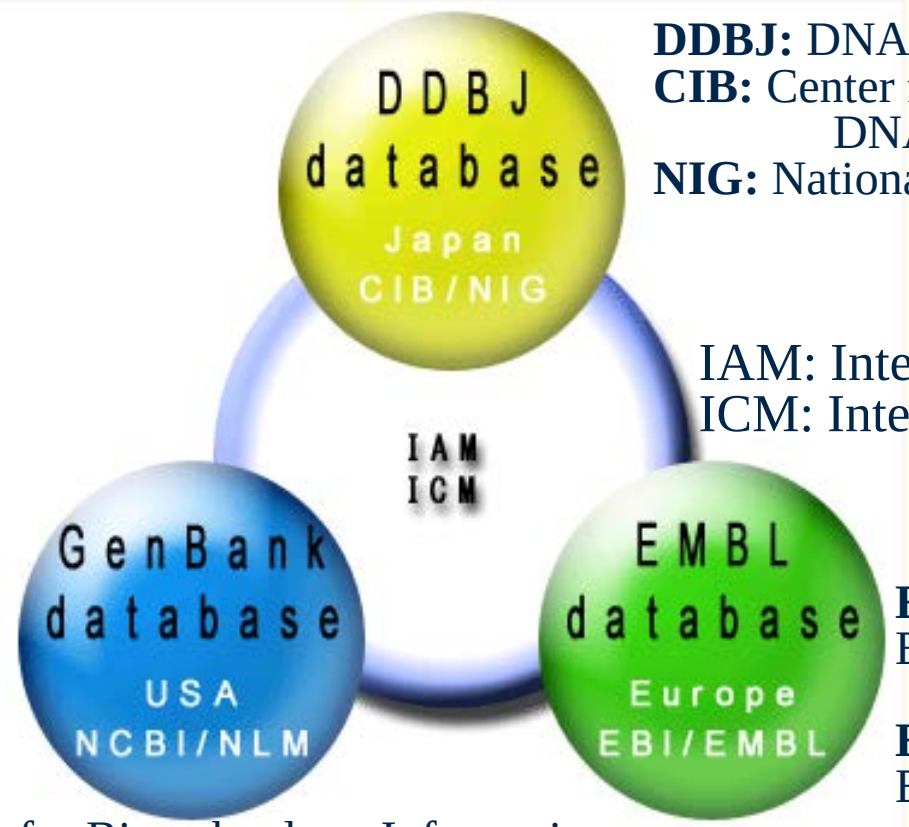








GenBank/EMBL/DDBJ International Nucleotide Sequence Database



DDBJ: DNA Data Bank of Japan
CIB: Center for Information Biology and DNA Data Bank of Japan
NIG: National Institute of Genetics

IAM: International Advisory Meeting
ICM: International Collaborative Meeting

EMBL:
European Molecular Biology Laboratory
EBI:
European Bioinformatics Institute

NCBI:
National Center for Biotechnology Information
NLM:
National Library of Medicine

NCBI Resources How To

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

Sign in to NCBI

GenBank Nucleotide Search

GenBank Submit Genomes WGS HTGs EST/GSS Metagenomes TPA TSA INSDC

GenBank Overview

What is GenBank?

GenBank ® is the NIH genetic sequence database, an annotated collection of all publicly available DNA sequences (*Nucleic Acids Research*. 2013 Jan;41(D1):D36-42). GenBank is part of the International Nucleotide Sequence Database Collaboration , which comprises the DNA DataBank of Japan (DDBJ), the European Molecular Biology Laboratory (EMBL), and GenBank at NCBI. These three organizations exchange data on a daily basis.

The complete release notes for the current version of GenBank are available on the NCBI ftp site. A new release is made every two months. GenBank growth statistics for both the traditional GenBank divisions and the WGS division are available from each release.

An example of a GenBank record may be viewed for a *Saccharomyces cerevisiae* gene.

Access to GenBank

There are several ways to search and retrieve data from GenBank.

- Search GenBank for sequence identifiers and annotations with Entrez Nucleotide , which is divided into three divisions: CoreNucleotide (the main collection), dbEST (Expressed Sequence Tags), and dbGSS (Genome Survey Sequences).
- Search and align GenBank sequences to a query sequence using BLAST (Basic Local Alignment Search Tool). BLAST searches CoreNucleotide, dbEST, and dbGSS independently; see BLAST info for more information about the numerous BLAST databases.
- Search, link, and download sequences programatically using NCBI e-utilities .

GenBank Resources

GenBank Home

Submission Types

Submission Tools

Search GenBank

Update GenBank Records

GenBank Data Usage

The GenBank database is designed to provide and encourage access within the scientific community to the most up to date and comprehensive DNA sequence information. Therefore, NCBI places no restrictions on the use or distribution of the GenBank data. However, some submitters may claim patent, copyright, or other intellectual property rights in all or a portion of the data they have submitted. NCBI is not in a position to assess the validity of such claims, and therefore cannot provide comment or unrestricted permission concerning the use, copying, or distribution of the information contained in GenBank.

NCBI Resources How To

NCBI
National Center for Biotechnology Information

Search All Databases

Search Clear

Resources

- **NCBI Home**
- All Resources (A-Z)
- 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.

More about the NCBI | Mission | Organization | Research | RSS



Popular Resources

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

NCBI News

MyNCBI supports OpenID and InCommons IDs
22 Sep 2010
A list of supported organizations can be found on the MyNCBI

Personalized settings in My NCBI
30 Aug 2010
Personalized preferences such as record format and results display

More...

Secondarily Databases

EMBL-EBI EB-eye Search All Databases Enter Text Here Go Reset Advanced Search Give us feedback

Databases | Tools | EBI Groups | Training | Industry | About Us | Help | Site Index

- InterPro:Home
  - Advanced Search
    - InterProScan
    - BioMart
  - Help / Documentation
    - About InterPro
    - Release Notes
    - BioMart Manual
    - Tutorial
    - Publications
    - Contributors
  - Web Services
  - Downloads
  - Protein Focus
    - Collagen

EBI > Databases > InterPro

InterPro protein sequence analysis & classification

InterPro is an integrated database of predictive protein "signatures" used for the classification and automatic annotation of proteins and genomes. InterPro classifies sequences at superfamily, family and subfamily levels, predicting the occurrence of functional domains, repeats and important sites. InterPro adds in-depth annotation, including GO terms, to the protein signatures.

Current release: **28.0 11th August 2010** (see Release Notes for further details)

Search InterPro: >

Do a sequence search of InterPro, via InterProScan

Extract large datasets by querying our BioMart

You can access our data programmatically, via Web Services

Secondarily Databases



FlyBase
A Database of the Drosophila Genome

SGD Saccharomyces Genome Database

The Genome Database
An international collaboration in support of the Human Genome Project.
GDB

Mouse Genome Informatics
The Jackson Laboratory

BDGP
Berkeley Drosophila Genome Project

Sanger Centre | Acedb | Acebrowser | Ensembl | Trace Server | Library
The Sanger Centre
Info | Database Searches | HGP | Projects | Software | Teams | Search
Data Release Policy | Conditions of Use

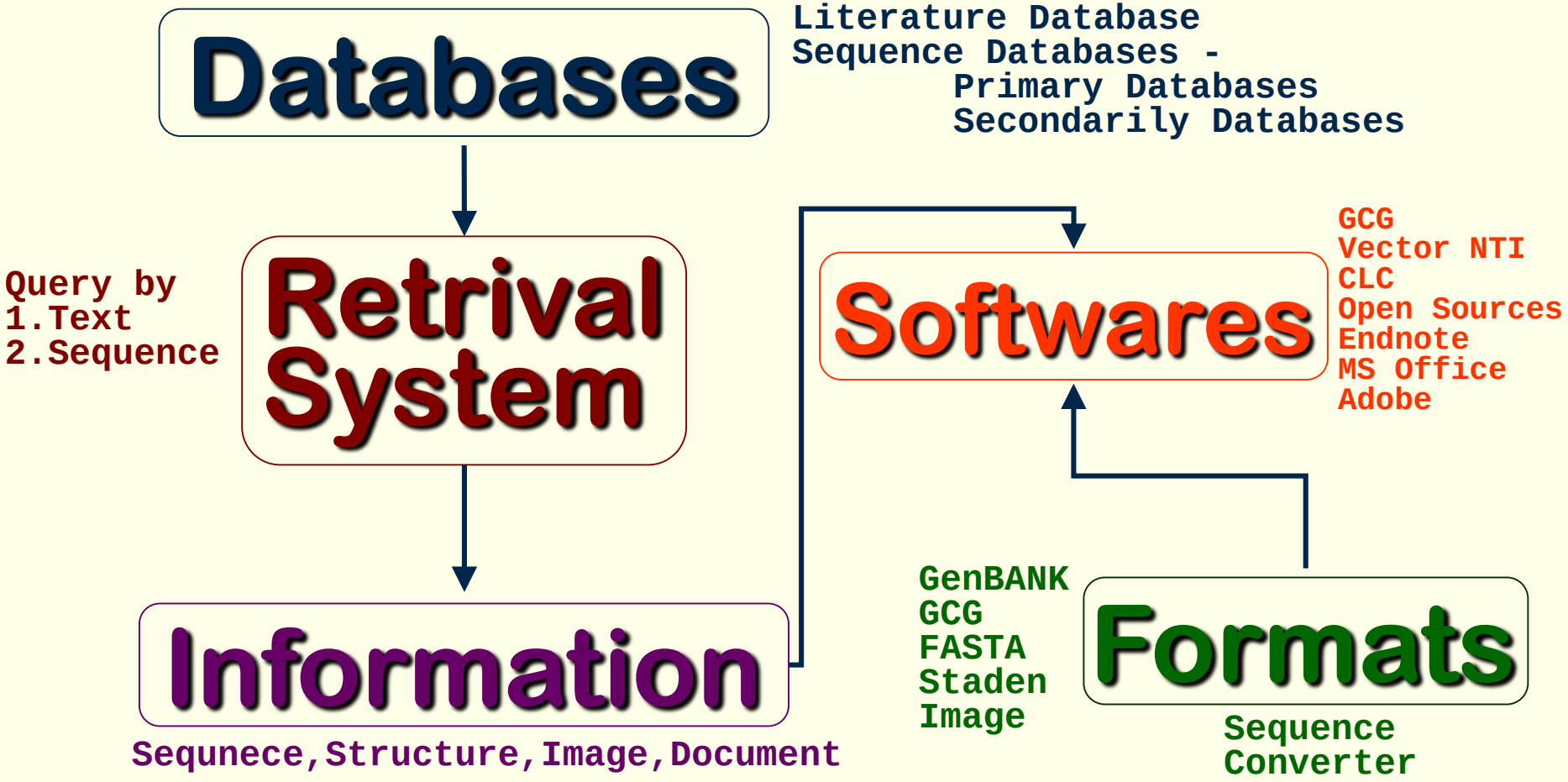


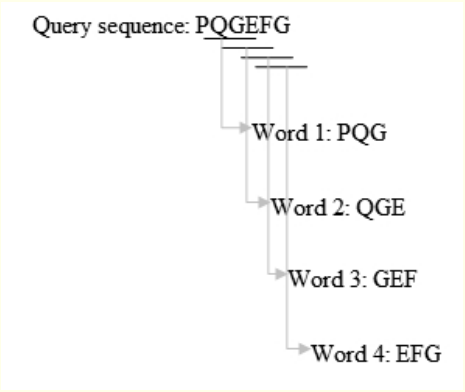
Projects Home

The *C. elegans* Genome Project

The *C. elegans* sequencing project is a collaboration between the Sanger Centre, Hinxton, Cambridge and the Genome Sequencing Center at the Washington University School of Medicine, St. Louis. The sequence is now essentially complete, as published in *Science* (additional notes here).

Database Retrieving and Manipulation Network





fuzzy search
（approximate string matching）

Google 台灣

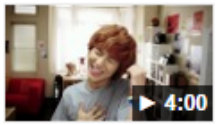
Hello

網頁 圖片 影片 地圖 新聞 更多 搜尋工具

約有 108,000,000 項結果 (搜尋時間：0.42 秒)

HELLO! Online: celebrity & royal news, magazine, babies ...
www.hellomagazine.com/ ▾ 翻譯這個網頁
HELLO! Online brings you the latest celebrity & royal news from the UK & around the world, magazine exclusives, celeb babies, weddings, pregnancies and ...
Royalty - William and Kate - Celebrities - This week's issue

Hello - 维基百科，自由的百科全书 - Wikipedia
zh.wikipedia.org/zh-tw/Hello ▾ 轉為繁體網頁
Hello. Hello是英語中常用的問候語，在英語中意思相同的問候語還有Hi和Hey，早在1883年即有詞典收錄此詞。除了英語，許多歐洲語言中也有類似的詞彙，像是 ...

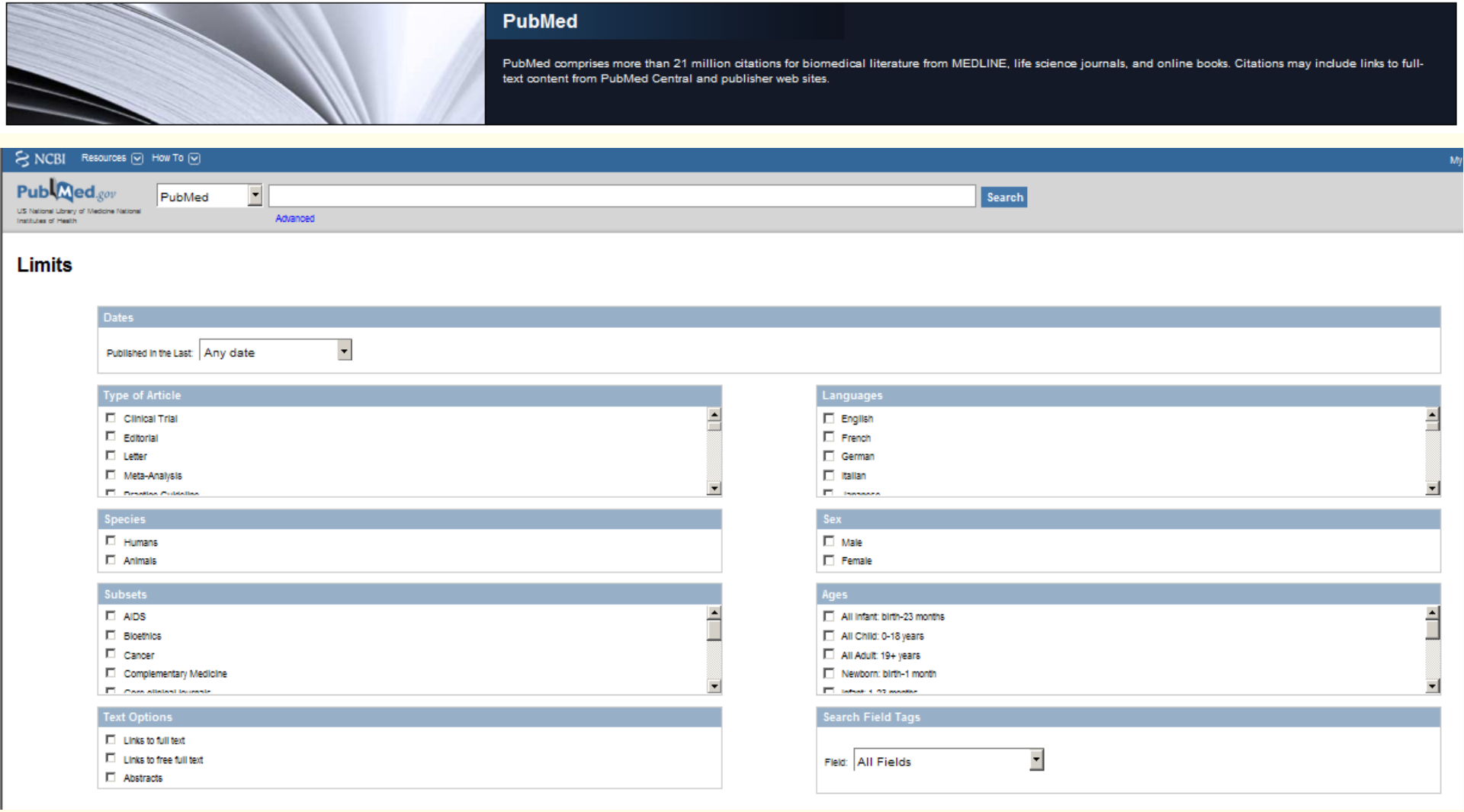
SHINee(샤이니) _ Hello _ MusicVideo - YouTube


www.youtube.com/watch?v=skZxb5sBoiU ▾
2010年10月4日 - 上傳者：sment
SHINee(샤이니) _ Hello _ MusicVideo ♪ Download on iTunes : http://itunes.apple.com/us/album/hello-the ...

Gina hello ⁂ | Facebook
https://zh-tw.facebook.com/Gina.Hello ▾
Gina hello ⁂， 臺中市. 186163 個讚 · 26347 人正在談論這個. ❤MIMIG 品牌主理http://mimig.tw ❤BLOG: http://ginahello.com ❤邀約請mail:design.gina@gmail.com...

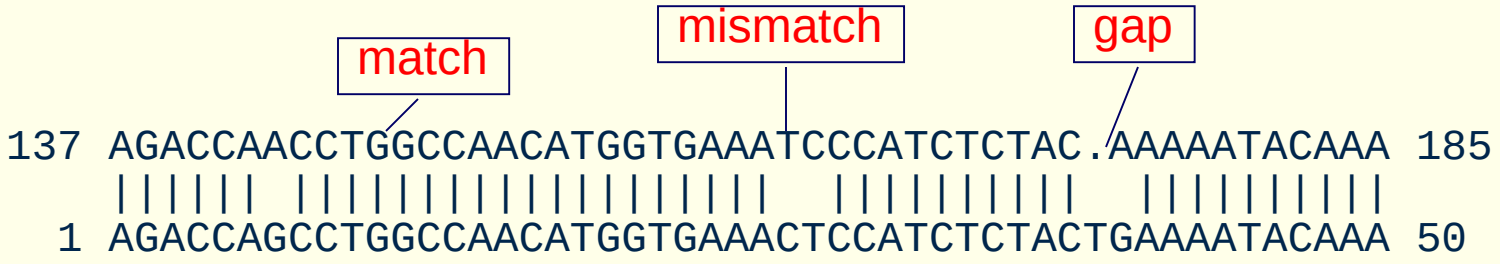
Hello Kitty Jet Travels with You! | 首頁
evakitty.evaair.com/ ▾
10月29日起長榮航空Hello Kitty彩繪機直飛巴黎 · 長榮航空首架777 Hello Kitty Jet — 牽手機，以三麗鷗家族串連全世界... Hello Kitty Jet加持祝Jolin上海演唱會大成功 ...

Literature Databases

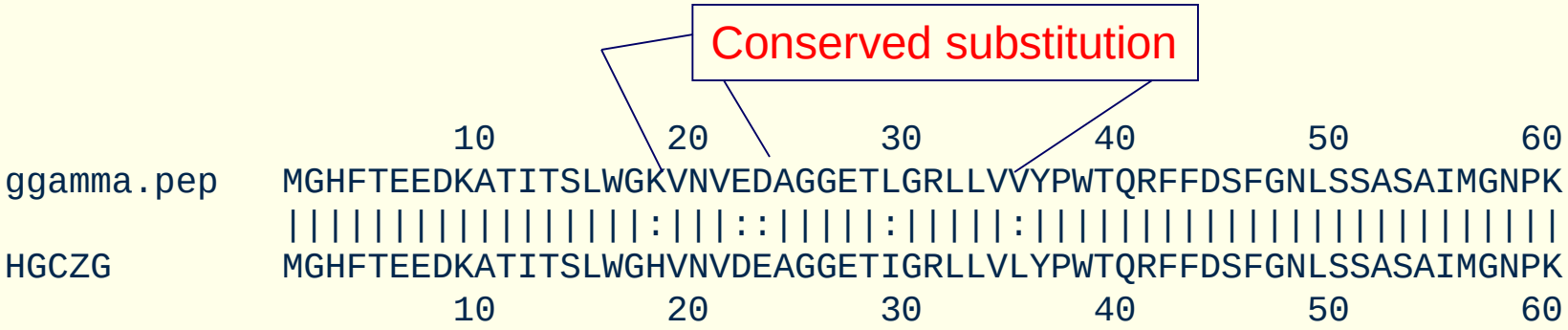


Sequence Comparison

Nucleotide sequence alignments



Protein sequence alignments



Residues with shared chemical properties can substitute for each other
Size, charge, hydrophobicity, polarity
scored less than a match, but better than a mismatch
Conservative changes scored as better than non-conservative

Pairwise Comparsion

Local Alignment

compares regions within two sequences and can return several matches



BLAST

VS

Global Alignment

compare entire sequences



FASTA

Query by sequence

Program	QUERY	Database
blastp	amino acid sequence	protein sequence database.
blastn	nucleotide sequence	nucleotide sequence database.
blastx	nucleotide sequence translated in all reading frames	protein sequence database (use this option to find potential translation products of an unknown nucleotide sequence)
tblastn	amino acid sequence	nucleotide sequence database translated in all reading frames
tblastx	six-frame translations of a nucleotide sequence	six-frame translations of a nucleotide sequence database. (tblastx program cannot be used with the nr database on the BLAST Web page because it is computationally intensive)

BLAST principle

1: Query Preprocessing

Break query into words

`DTLVRAIP ->  DTL, TLV, LVR, VRA ...`

Make a table of similar words

`TLV -> TLI, TIV, SLV`

2: Search query words in indexed table of database words

Find exact match between table word & db.

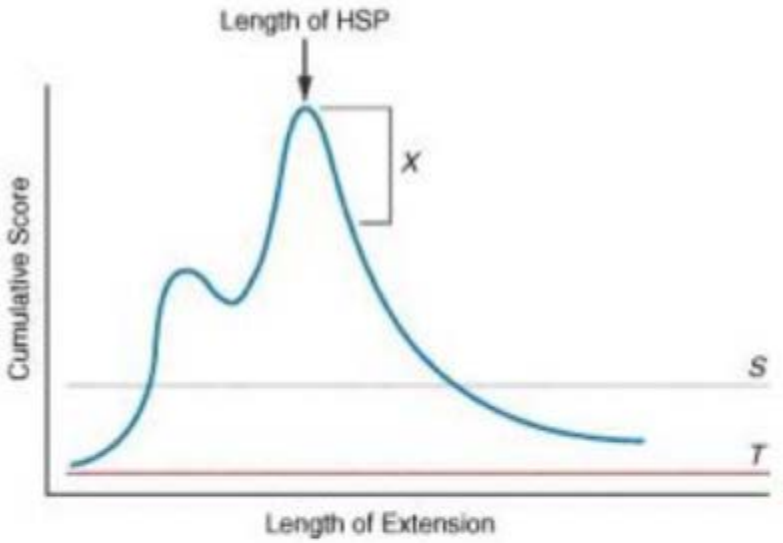
```
          TIV
SDTDGDKNADGWIETIVRALPTSD
```

3: Extend query match to HSP

– keep HSPs of significant quality.

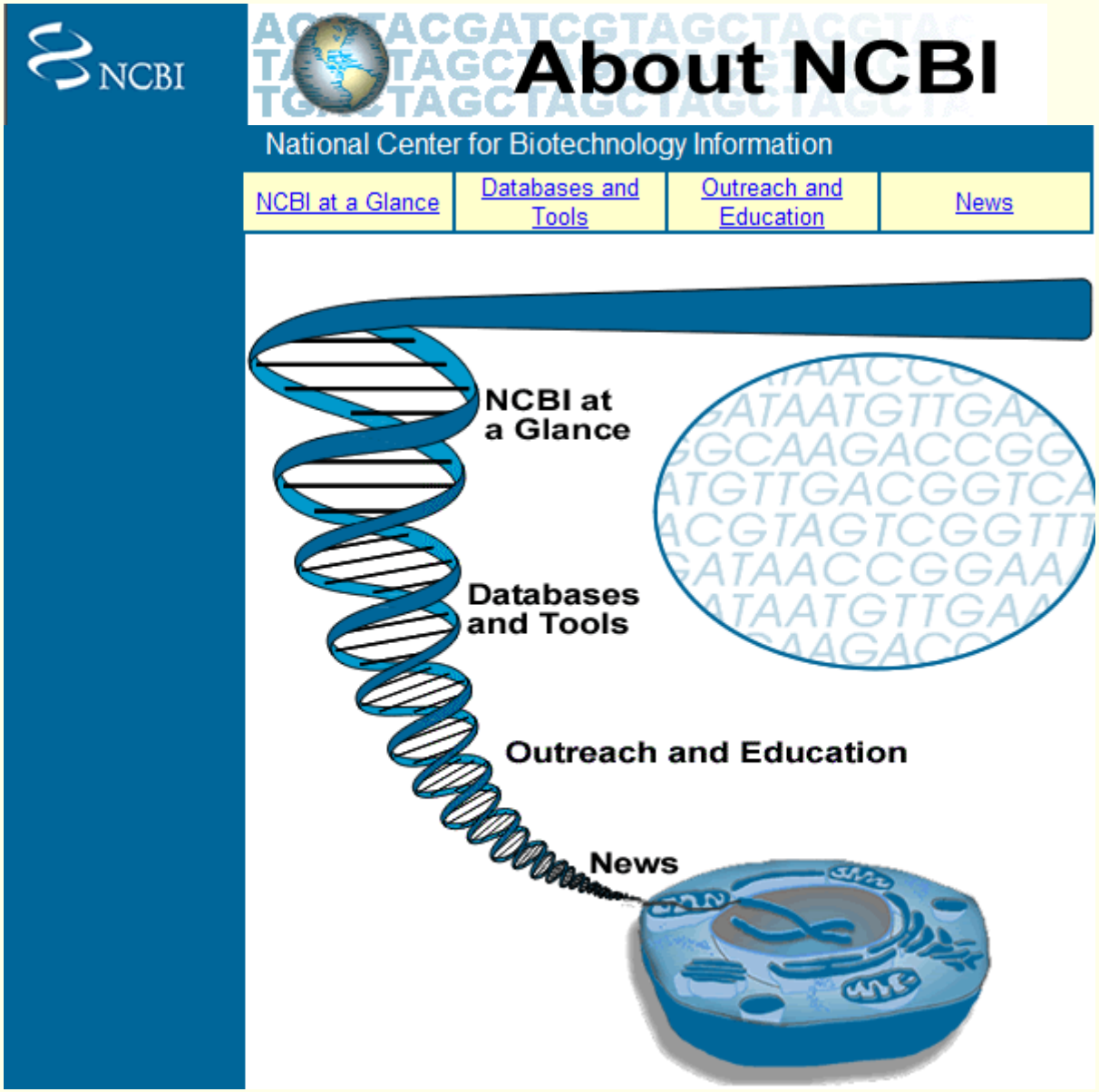
```
        DTLVRAIP
SDTDGDKNADGWIETIVRALPTSD
```

4: Assemble HSPs into gapped alignment



HSP (High Scoring Segment Pair): **An ungapped, high-scoring, local alignment**

http://www.ncbi.nlm.nih.gov/About/glance/index.html



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

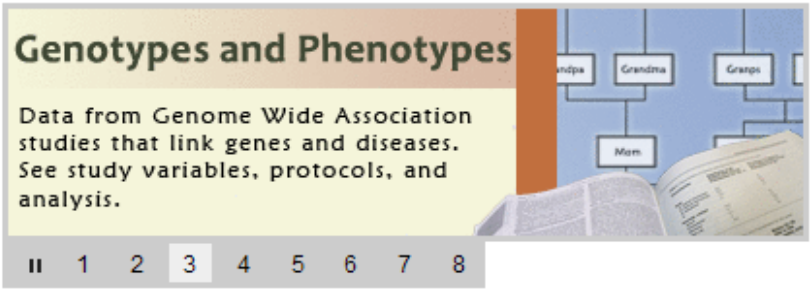
Welcome to NCBI

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

About the NCBI | Mission | Organization | Research | NCBI News

Get Started

- Tools: Analyze data using NCBI software
- Downloads: Get NCBI data or software
- How-To's: Learn how to accomplish specific tasks at NCBI
- Submissions: Submit data to GenBank or other NCBI databases



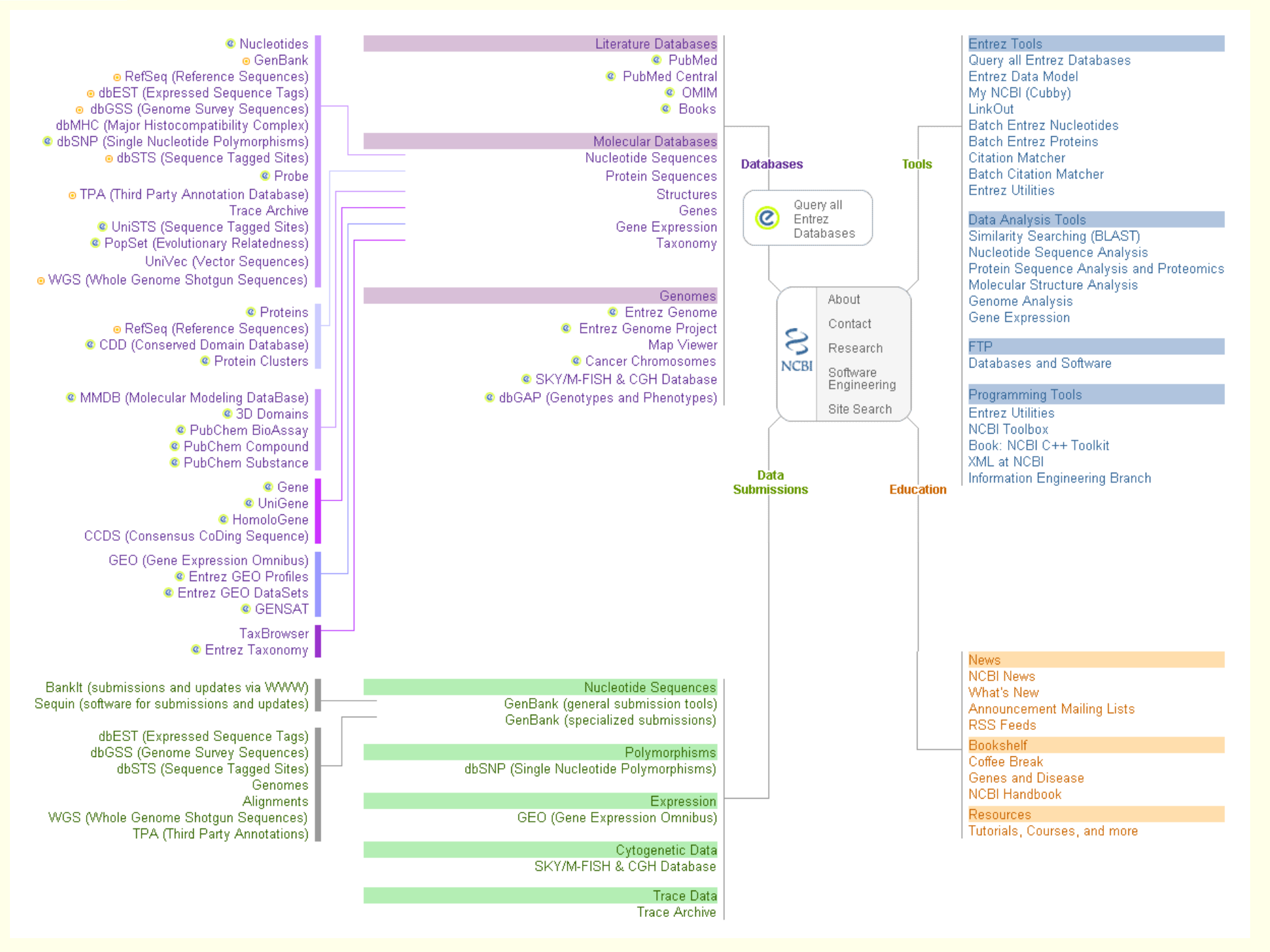
Popular Resources

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

NCBI Announcements

Try the new My NCBI Feature: SciENcv
Sep 26, 2013
The National Institutes of Health has issued an invitation to researchers

Comments Requested: NIH genomic data sharing policy
Sep 20, 2013
The National Institutes of Health (NIH) is


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
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)
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)
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
Databases
Query all Entrez Databases
NCBI
About
Contact
Research
Software Engineering
Site Search
Data Submissions
Tools
Education
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
News
NCBI News
What's New
Announcement Mailing Lists
RSS Feeds
Bookshelf
Coffee Break
Genes and Disease
NCBI Handbook
Resources
Tutorials, Courses, and more

http://www.ncbi.nlm.nih.gov/sites/gquery

Search NCBI databases Help

Search

Literature

Books	books and reports
MeSH	ontology used for PubMed indexing
NLM Catalog	books, journals and more in the NLM Collections
PubMed	scientific & medical abstracts/citations
PubMed Central	full-text journal articles

Health

ClinVar	human variations of clinical significance
dbGaP	genotype/phenotype interaction studies
GTR	genetic testing registry
MedGen	medical genetics literature and links
OMIM	online mendelian inheritance in man
PubMed Health	clinical effectiveness, disease and drug reports

Genomes

Assembly	genomic assembly information
BioProject	biological projects providing data to NCBI
BioSample	descriptions of biological source materials
Clone	genomic and cDNA clones
dbVar	genome structural variation studies
Epigenomics	epigenomic studies and display tools
Genome	genome sequencing projects by organism
GSS	genome survey sequences
Nucleotide	DNA and RNA sequences
Probe	sequence-based probes and primers
SNP	short genetic variations
SRA	high-throughput DNA and RNA sequence read archive
Taxonomy	taxonomic classification and nomenclature catalog

Genes

EST	expressed sequence tag sequences
Gene	collected information about gene loci
GEO DataSets	functional genomics studies
GEO Profiles	gene expression and molecular abundance profiles
HomoloGene	homologous gene sets for selected organisms
PopSet	sequence sets from phylogenetic and population studies
UniGene	clusters of expressed transcripts

Proteins

Conserved Domains	conserved protein domains
Protein	protein sequences
Protein Clusters	sequence similarity-based protein clusters
Structure	experimentally-determined biomolecular structures

Chemicals

BioSystems	molecular pathways with links to genes, proteins and chemicals
PubChem BioAssay	bioactivity screening studies
PubChem Compound	chemical information with structures, information and links
PubChem Substance	deposited substance and chemical information

NCBI Resources How To Sign in to NCBI

Search NCBI databases

Help

p53 Search

About 2,054,751 search results for "p53"

Literature

Books	1,434	books and reports
MeSH	162	ontology used for PubMed indexing
NLM Catalog	108	books, journals and more in the NLM Collections
PubMed	73,213	scientific & medical abstracts/citations
PubMed Central	98,959	full-text journal articles

Health

ClinVar	296	human variations of clinical significance
dbGaP	22	genotype/phenotype interaction studies
GTR	133	genetic testing registry
MedGen	72	medical genetics literature and links
OMIM	584	online mendelian inheritance in man
PubMed Health	78	clinical effectiveness, disease and drug reports

Genomes

Assembly	1	genomic assembly information
BioProject	665	biological projects providing data to NCBI
BioSample	319	descriptions of biological source materials
Clone	0	genomic and cDNA clones
dbVar	1,659	genome structural variation studies
Epigenomics	0	epigenomic studies and display tools
Genome	6	genome sequencing projects by organism
GSS	36	genome survey sequences
Nucleotide	47,635	DNA and RNA sequences
Probe	3,518	sequence-based probes and primers
SNP	6,593	short genetic variations
SRA	462	high-throughput DNA and RNA sequence read archive
Taxonomy	0	taxonomic classification and nomenclature catalog

Genes

EST	797	expressed sequence tag sequences
Gene	8,300	collected information about gene loci
GEO DataSets	9,570	functional genomics studies
GEO Profiles	1,749,571	gene expression and molecular abundance profiles
HomoloGene	38	homologous gene sets for selected organisms
PopSet	97	sequence sets from phylogenetic and population studies
UniGene	414	clusters of expressed transcripts

Proteins

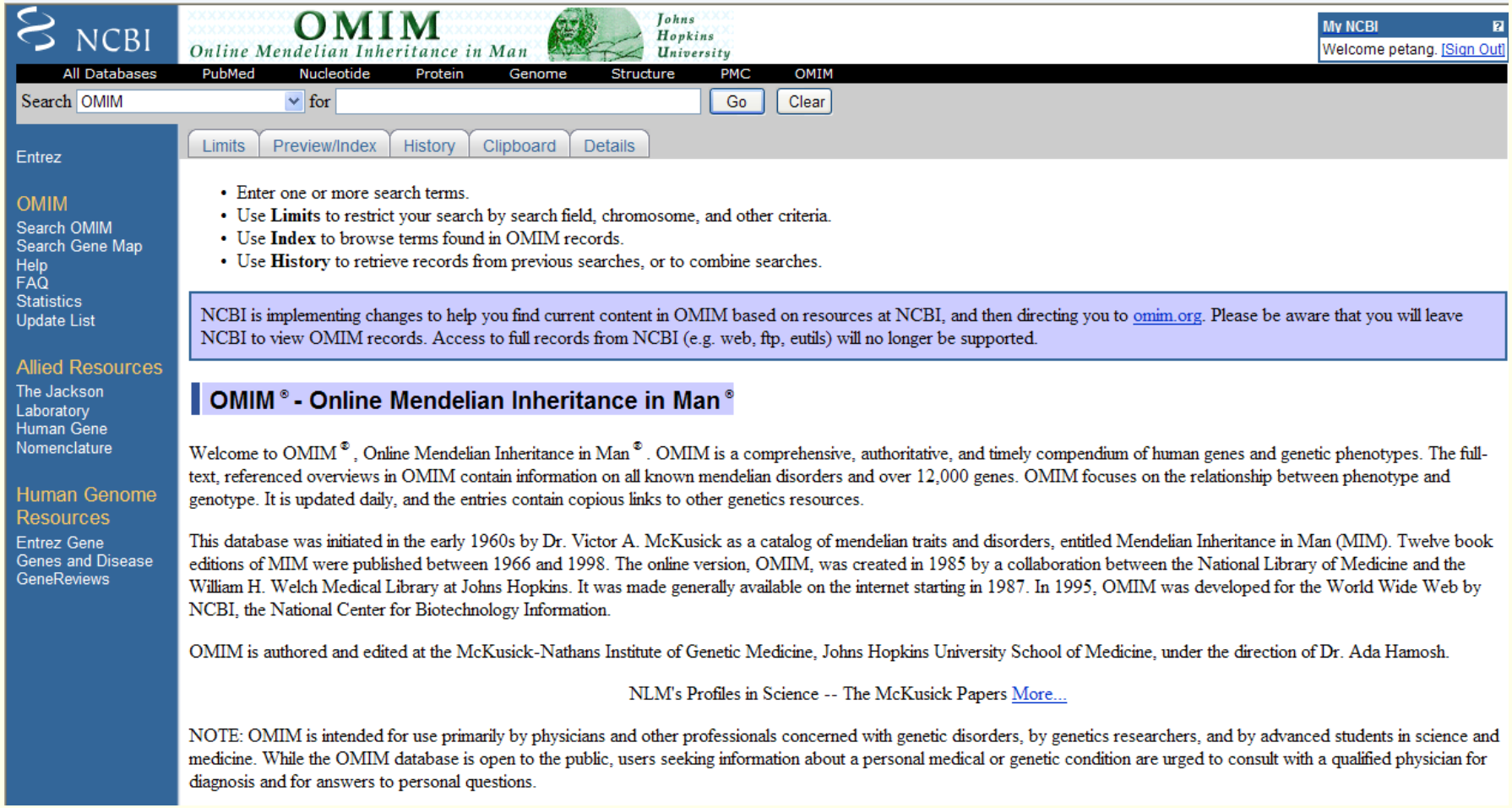
Conserved Domains	120	conserved protein domains
Protein	32,168	protein sequences
Protein Clusters	15	sequence similarity-based protein clusters
Structure	1,146	experimentally-determined biomolecular structures

Chemicals

BioSystems	4,040	molecular pathways with links to genes, proteins and chemicals
PubChem BioAssay	11,806	bioactivity screening studies
PubChem Compound	9	chemical information with structures, information and links
PubChem Substance	664	deposited substance and chemical information

Literature Databases

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



OMIM | OMIM | p53 | Search

Save search | Limits | Advanced | Help

Display Settings: Summary, 20 per page | Send to:

Results: 1 to 20 of 584

<< First | < Prev | Page 1 of 30 | Next > | Last >>

1. *191170 - TUMOR PROTEIN **p53**; TP53
   Cytogenetic locations: 17p13.1
   OMIM: 191170
   Gene summaries | Genetic tests | Medical literature

2. *605426 - TUMOR PROTEIN **p53**-REGULATED APOPTOSIS-INDUCING PROTEIN 1; TP53AIP1
   Cytogenetic locations: 11q24.3
   OMIM: 605426
   Gene summaries | Genetic tests | Medical literature

3. *614502 - CHROMOSOME 9 OPEN READING FRAME 116; C9ORF116
   Cytogenetic locations: 9q34.3
   OMIM: 614502
   Gene summaries | Genetic tests | Medical literature

4. *605157 - **p53**-RESPONSIVE GENE 1
   Cytogenetic locations: 19q13.31
   OMIM: 605157
   Gene summaries | Genetic tests | Medical literature

5. *605159 - APOPTOSIS-INDUCING FACTOR, MITOCHONDRION-ASSOCIATED, 2; AIFM2
   OMIM: 605159
   Gene summaries | Genetic tests | Medical literature

6. *606185 - TUMOR PROTEIN **p53**-INDUCIBLE NUCLEAR PROTEIN 1; TP53INP1
   Cytogenetic locations: 8q22.1
   OMIM: 606185
   Gene summaries | Genetic tests | Medical literature

7. *605160 - **p53**-RESPONSIVE GENE 4
   OMIM: 605160
   Gene summaries | Genetic tests | Medical literature

Filter your results:

All (584)
OMIM UniSTS (67)
OMIM dbSNP (135)

Manage Filters

Find related data

Database: Select

Find items

Search details

p53[All Fields]

Search | See more...

Recent activity

Turn Off | Clear

p53 (584) OMIM

See more...

p53 Search

Advanced Search ▾ | Display Options ▾

*191170

TUMOR PROTEIN p53; TP53

Alternative titles; symbols

P53

TRANSFORMATION-RELATED PROTEIN 53; TRP53

HGNC Approved Gene Symbol: TP53

Cytogenetic location: 17p13.1 ***Genomic coordinates (GRCh37): 17:7,571,719-7,590,867*** (from NCBI)

Gene-Phenotype Relationships

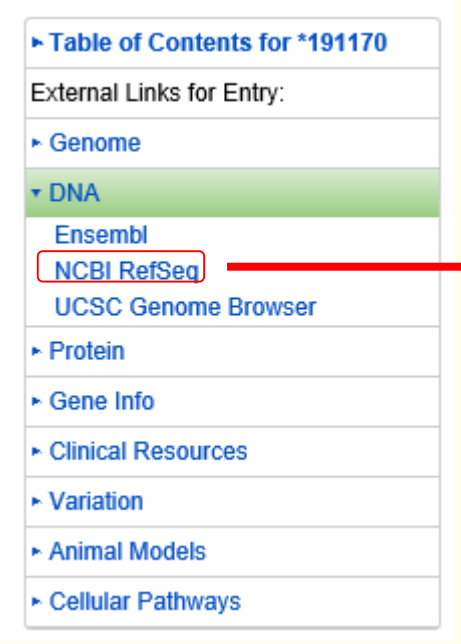
Location	Phenotype	Phenotype MIM number	Phenotype mapping key
17p13.1	Adrenal cortical carcinoma	202300	3
	Breast cancer	114480	3
	Choroid plexus papilloma	260500	3
	Colorectal cancer	114500	3
	Hepatocellular carcinoma	114550	3
	Li-Fraumeni syndrome	151623	3
	Nasopharyngeal carcinoma	607107	3
	Osteosarcoma	259500	3
	Pancreatic cancer	260350	3
	{Basal cell carcinoma 7}	614740	3
	{Glioma susceptibility 1}	137800	3

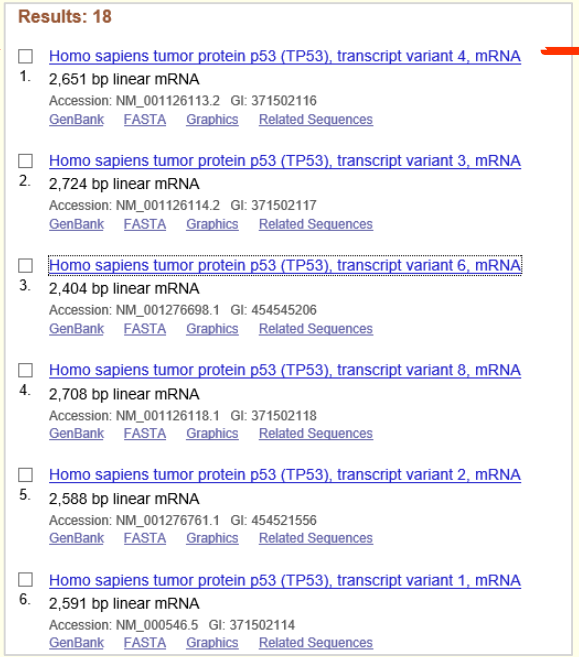
▾ Table of Contents for *191170

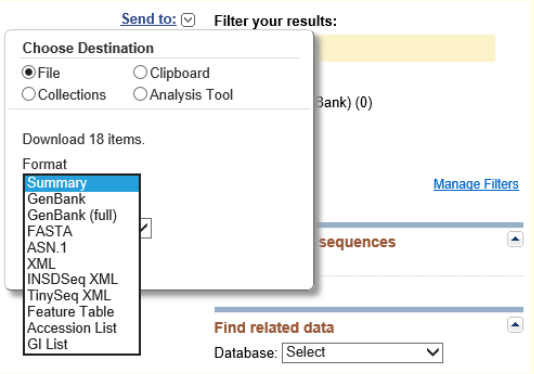
Title
Gene-Phenotype Relationships
Text
Description
Cloning and Expression
Gene Structure
Mapping
Gene Function
Biochemical Features
Molecular Genetics
Animal Model
Allelic Variants
Table View
See Also
References
Contributors
Creation Date
Edit History

External Links for Entry:

▸ Genome
▸ DNA
▸ Protein
▸ Gene Info
▸ Clinical Resources
▸ Variation
▸ Animal Models
▸ Cellular Pathways







Homo sapiens tumor protein p53 (TP53), transcript variant 4, mRNA

NCBI Reference Sequence: NM_001126113.2

FASTA Graphics

Go to:

```
LOCUS       NM_001126113            2651 bp    mRNA    linear   PRI 04-MAY-2014
DEFINITION  Homo sapiens tumor protein p53 (TP53), transcript variant 4, mRNA.
ACCESSION   NM_001126113 XR_243565
VERSION     NM_001126113.2  GI:371502116
KEYWORDS    RefSeq.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;
            Catarrhini; Hominidae; Homo.
```

Customize view

Basic Features
- (•) Default features
- ( ) Gene, RNA, and CDS features only

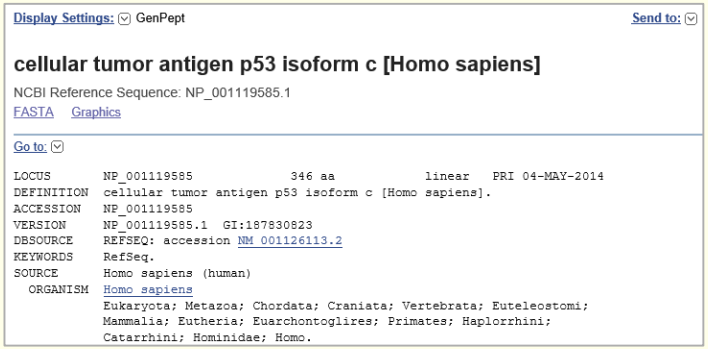
Features added by NCBI
- [ ] 174 SNPs
- [x] 2 conserved domains

Display options
- [x] Show sequence
- [ ] Show reverse complement

Update View

```
FEATURES             Location/Qualifiers
     source          1..2651
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="17"
                     /map="17p13.1"
     gene            1..2651
                     /gene="TP53"
                     /gene_synonym="BCC7; LFS1; P53; TRP53"
                     /note="tumor protein p53"
                     /db_xref="GeneID:7157"
                     /db_xref="HGNC:HGNC:11998"
                     /db_xref="MIM:191170"
     exon            1..174
                     /gene="TP53"
                     /gene_synonym="BCC7; LFS1; P53; TRP53"
                     /inference="alignment:Splign:1.39.8"
     misc_feature    95..97
                     /gene="TP53"
                     /gene_synonym="BCC7; LFS1; P53; TRP53"
                     /note="upstream in-frame stop codon"
     exon            175..276
                     /gene="TP53"
                     /gene_synonym="BCC7; LFS1; P53; TRP53"
                     /inference="alignment:Splign:1.39.8"
     CDS             203..1243
                     /gene="TP53"
                     /gene_synonym="BCC7; LFS1; P53; TRP53"
                     /note="isoform c is encoded by transcript variant 4; tumor
                     protein 53; mutant tumor protein 53; cellular tumor
                     antigen p53; phosphoprotein p53; transformation-related
                     protein 53; p53 tumor suppressor; antigen NY-CO-13"
                     /codon_start=1
                     /product="cellular tumor antigen p53 isoform c"
                     /protein_id="NP_001119585.1"
                     /db_xref="GI:187830823"
                     /db_xref="CCDS:CCDS45605.1"
                     /db_xref="GeneID:7157"
                     /db_xref="HGNC:HGNC:11998"
                     /db_xref="MIM:191170"
                     /translation="MEEPQSDPSVEPPLSQETFSDLWKLLPENNVLSPLPSQAMDDLM
                     LSPDDIEQWFTEDPGPDEAPRMPEAAPPVAPAPAAPTPAAPAPAPSWPLSSSVPSQKT
                     YQGSYGFRLGFLHSGTAKSVTCTYSPALNKMFCQLAKTCPVQLWVDSTPPPGTRVRAM
                     AIYKQSQHMTEVVRRCPHHERCSDSDGLAPPQHLIRVEGNLRVEYLDDRNTFRHSVVV
                     PYEPPEVGSDCTTIHYNYMCNSSCMGGMNRRPILTIITLEDSSGNLLGRNSFEVRVCA
                     CPGRDRRTEEENLRKKGEPHHELPPGSTKRALPNNTSSSPQPKKKPLDGEYFTLQMLL
                     DLRWCYFLINSS"
```





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

The European Bioinformatics Institute

Part of the European Molecular Biology Laboratory

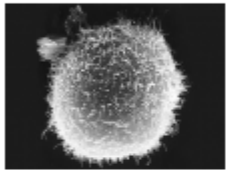
EMBL-EBI provides freely available data from life science experiments, performs basic research in computational biology and offers an extensive user training programme, supporting researchers in academia and industry.

Find a gene, protein or chemical:

Search

Examples: blast, keratin, bfl1...

News from EMBL-EBI



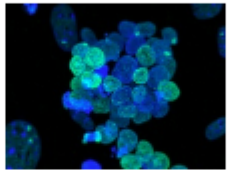
BLUEPRINT: new perspectives on white blood cells

Major international initiative to study the epigenomics of blood publishes its first results



Taking the guesswork out of genetic analysis

New method by EMBL-EBI and Microsoft Research improves statistical power of genome-wide association studies



Advance in stem-cell technology

Scientists revert human stem cells to pristine state

Read more press release

Our mission

- To provide freely available data and bioinformatics services to all facets of the scientific community
- To contribute to the advancement of biology through basic investigator-driven research
- To provide advanced bioinformatics training to scientists at all levels
- To help disseminate cutting-edge technologies to industry
- To coordinate biological data provision throughout Europe.

Bioinformatics services — Services

Research at EMBL-EBI — Research

Bioinformatics training — Training

Industry Programme — Industry

ELIXIR — European Coordination

EMBL Alumni Relations — EMBL ALUMNI

European Molecular Biology Laboratory

Visit EMBL.org

EMBL 40

Popular

- Services
- Research
- Training
- News
- Jobs
- Visit us
- EMBL
- Contacts

Events

- Introduction to EBI Resources - Cambridge University UK — Oct 6 2014
- Bioinformatics: EMBL-EBI - Protein Databases and Tools — Oct 7 2014
- Bioinformatics Workshop: Small molecule resources with EMBL-EBI - Cambridge, UK — Oct 10 2014
- UniProt: Protein Databases & Tools (OtST, 2014) — Oct 14 2014
- 1000 Genomes and browsing Ensembl — Oct 14 2014
- EMBL-EBI Careers Day 2014 — Oct 16 2014
- UniProt: Protein Databases & Tools (Tokyo, 2014) — Oct 16 2014
- Bioinformatics: Transcriptomics Data and Tools - Cambridge, UK — Oct 17 2014
- Ensembl Highlights Workshop - ASHG, 2014 — Oct 18 2014
- Interactions & Pathways - IntAct - Cambridge, UK — Oct 20 2014

See all courses and conferences

See other events at EMBL-EBI

- ENA Home
- Search & Browse
- Submit & Update
- About ENA
  - Browser
  - CRAM
  - Submissions
  - Download
  - Formats
  - Standards
  - Policies
  - Training
  - News
  - Forthcoming changes
  - Statistics
  - Sequence Search
  - Events
  - Software
- Contact
- FAQ

About the European Nucleotide Archive

The European Nucleotide Archive (ENA) captures and presents information relating to experimental workflows that are based around nucleotide sequencing. A typical workflow includes the isolation and preparation of material for sequencing, a run of a sequencing machine in which sequencing data are produced and a subsequent bioinformatic analysis pipeline. ENA records this information in a data model that covers input information (sample, experimental setup, machine configuration), output machine data (sequence traces, reads and quality scores) and interpreted information (assembly, mapping, functional annotation).

Data arrive at ENA from a variety of sources. These include submissions of raw data, assembled sequences and annotation from small-scale sequencing efforts, data provision from the major European sequencing centres and routine and comprehensive exchange with our partners in the International Nucleotide Sequence Database Collaboration (INSDC).

Provision of nucleotide sequence data to ENA or its INSDC partners has become a central and mandatory step in the dissemination of research findings to the scientific community. ENA works with publishers of scientific literature and funding bodies to ensure compliance with these principles and to provide optimal submission systems and data access tools that work seamlessly with the published literature.

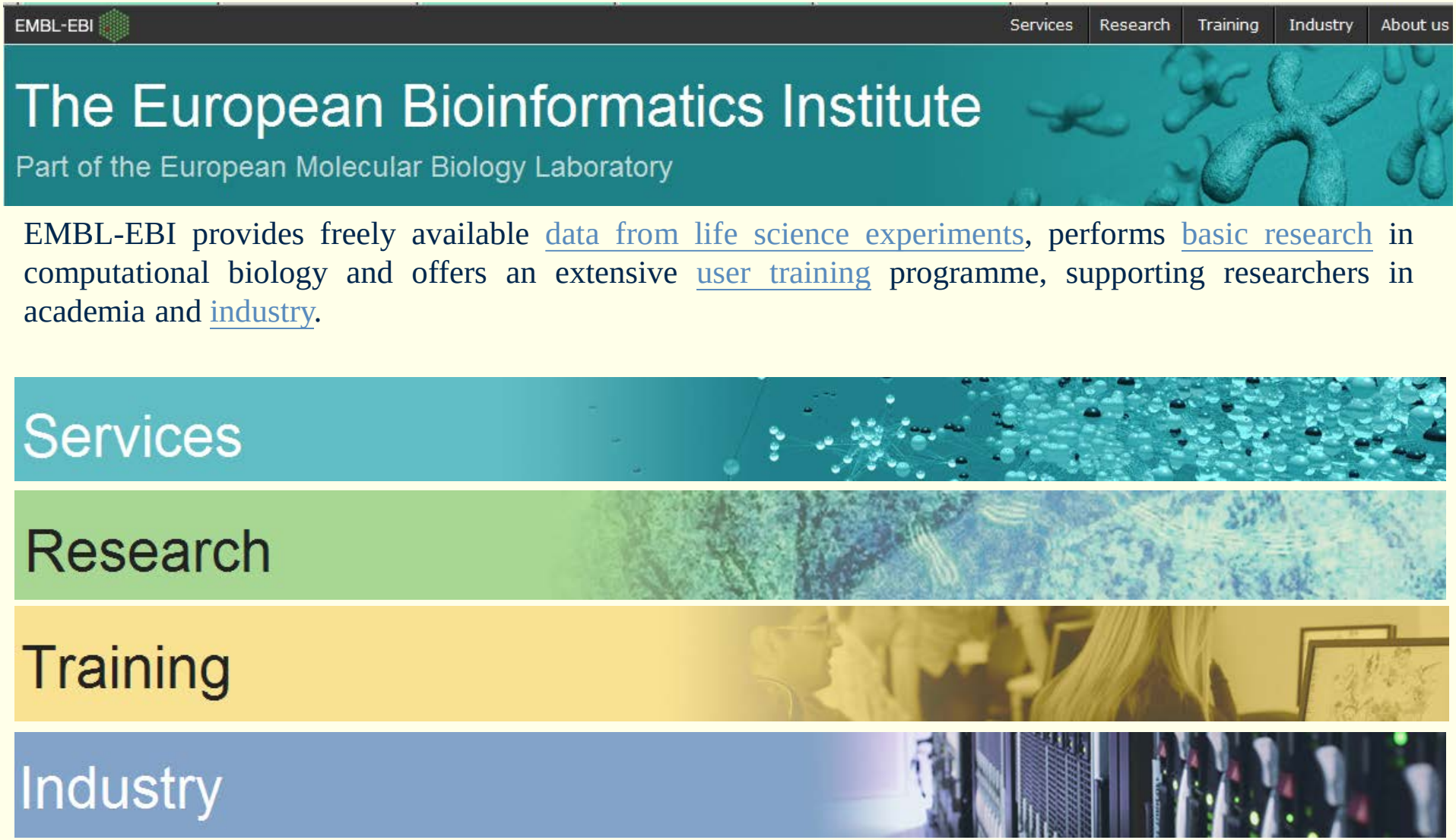
ENA is made up of a number of distinct databases that includes EMBL-Bank, the newly established Sequence Read Archive (SRA) and the Trace Archive each with their own data formats and standards. ENA data classes and formats are described here.

Although the ENA has almost 30 years of history, the data and services are constantly changing to reflect growing volumes of data, ever improving sequencing technology and the broadening of applications to which sequencing is now put. Latest developments and changes to services are announced here.

As part of the global effort to improve access to and usability of nucleotide sequencing data, we collaborate extensively in the development of our services and technologies and in standards activities.

The ENA is developed and maintained at the EMBL-EBI under the guidance of the INSDC International Advisory Committee and a newly formed Scientific Advisory Board.

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



EMBL-EBI provides freely available data from life science experiments, performs basic research in computational biology and offers an extensive user training programme, supporting researchers in academia and industry.

Services

Overview | A to Z | Service teams | Support

Bioinformatics services

We maintain the world's most comprehensive range of **freely available** and up-to-date molecular databases. Developed in collaboration with our colleagues worldwide, our services let you share data, perform complex queries and analyse the results in different ways. You can work locally by downloading our data and software, or use our web services to access our resources programmatically.

Popular

- Ensembl
- BLAST
- UniProt
- Europe PMC
- PDBe
- Reactome
- ArrayExpress
- Train online
- ChEMBL
- Support

DNA & RNA

genes, genomes & variation

Gene expression

RNA, protein & metabolite expression

Proteins

sequences, families & motifs

Structures

Molecular & cellular structures

Systems

reactions, interactions & pathways

Chemical biology

chemogenomics & metabolomics

Ontologies

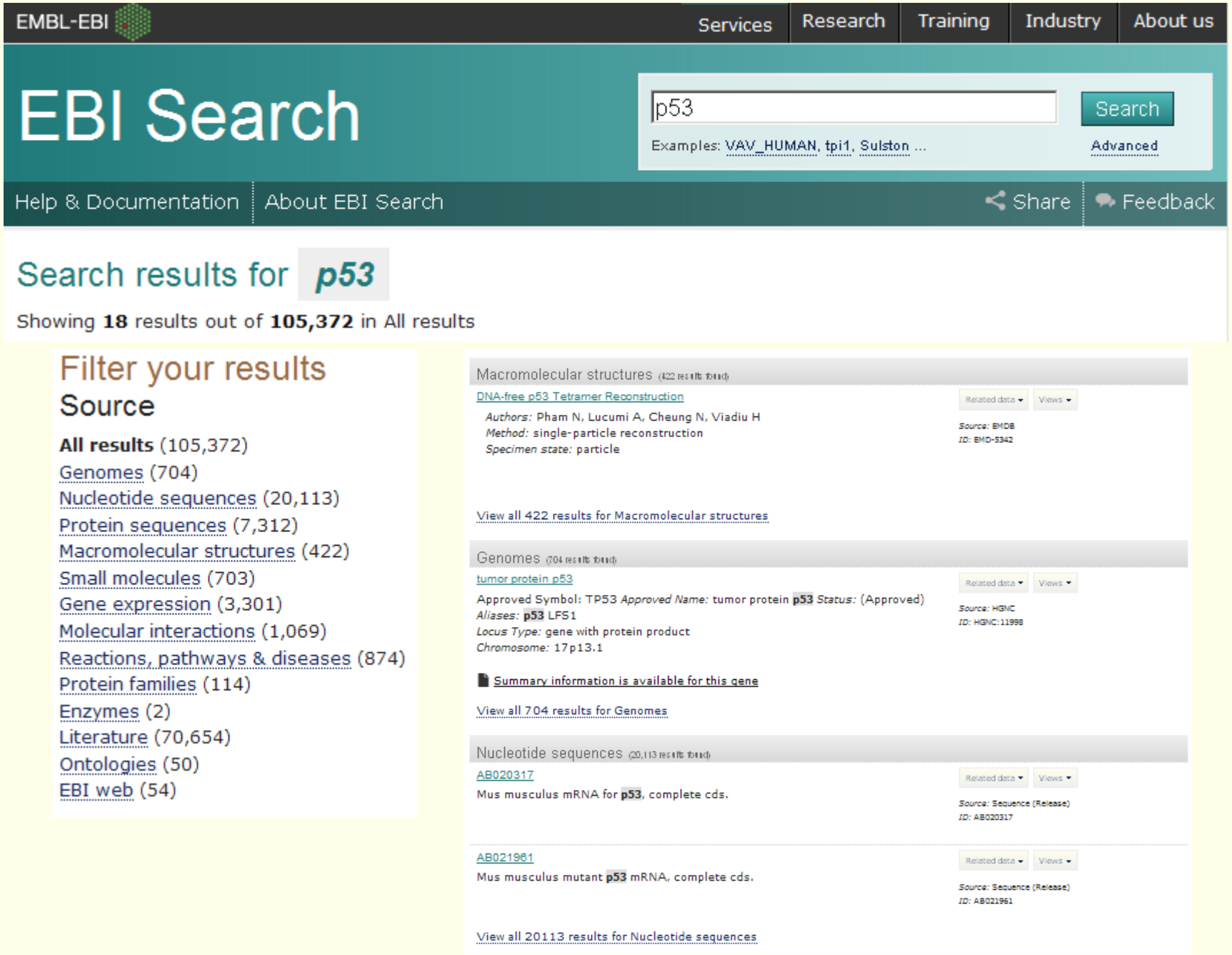
taxonomies & controlled vocabularies

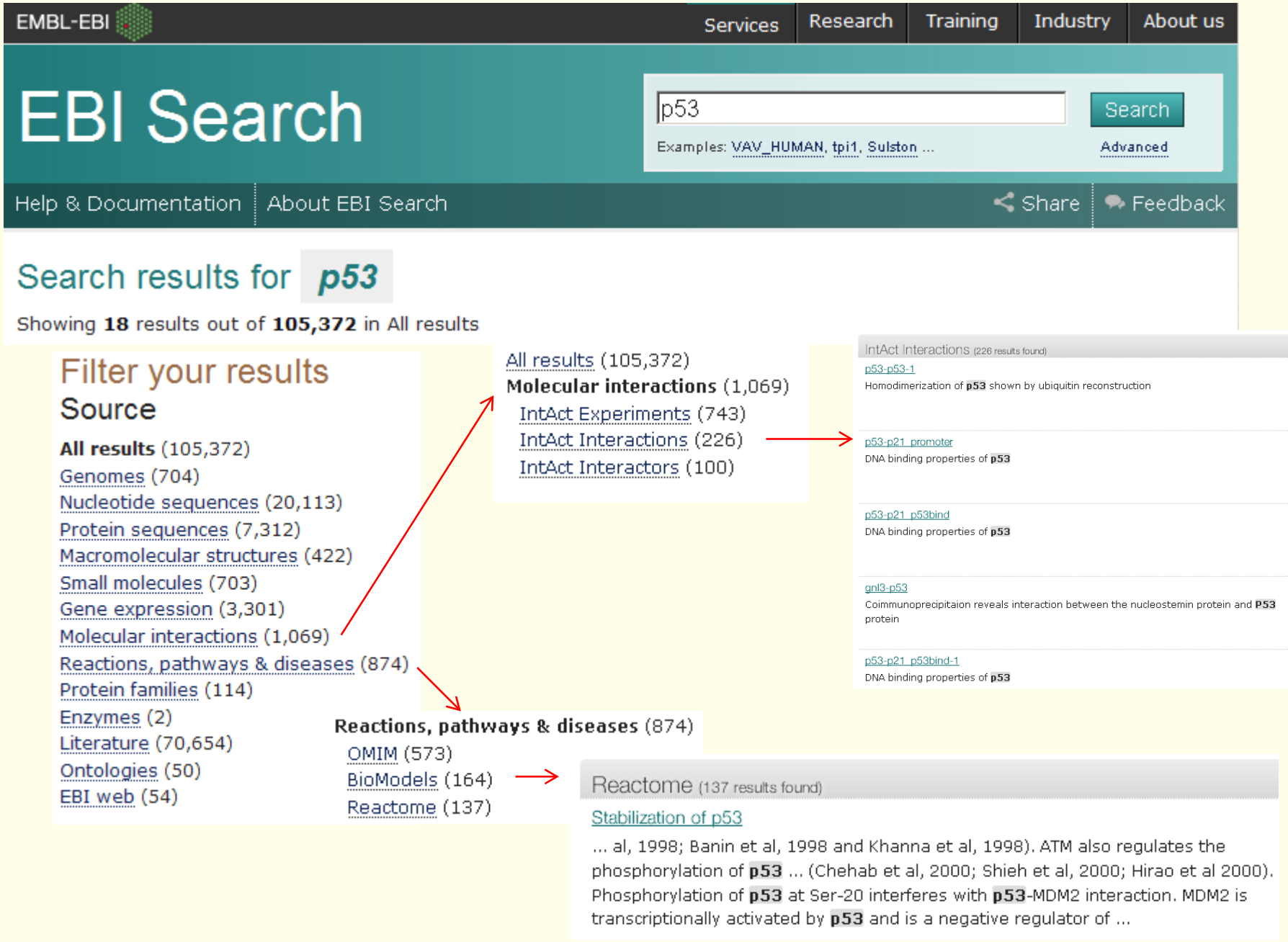
Literature

Scientific publications & patents

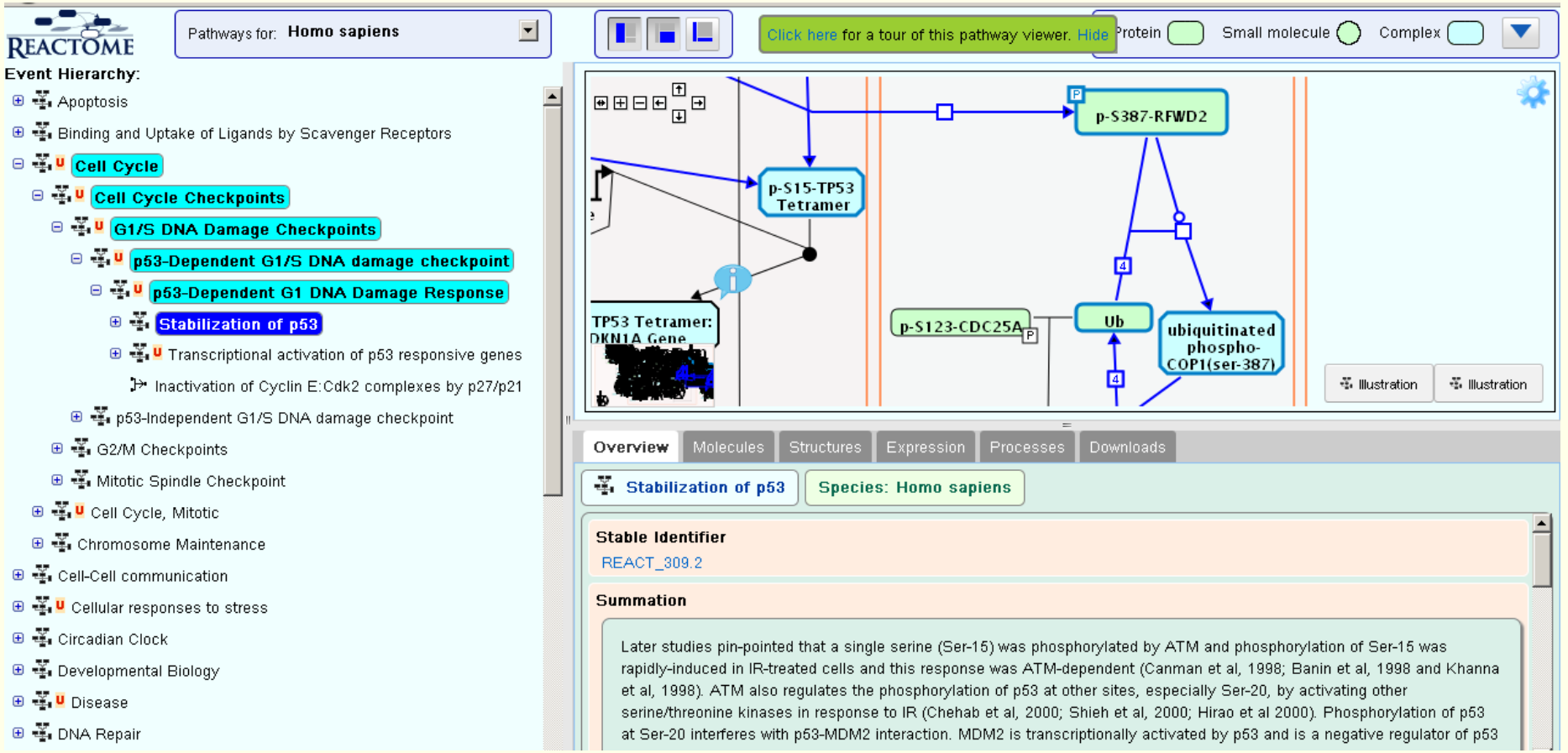
Other software

cross-domain tools & resources


EMBL-EBI
Services
Research
Training
Industry
About us
EBI Search
p53
Search
Examples: VAV_HUMAN, tpi1, Sulston ...
Advanced
Help & Documentation
About EBI Search
Share
Feedback
Search results for p53
Showing 18 results out of 105,372 in All results
Filter your results
Source
All results (105,372)
Genomes (704)
Nucleotide sequences (20,113)
Protein sequences (7,312)
Macromolecular structures (422)
Small molecules (703)
Gene expression (3,301)
Molecular interactions (1,069)
Reactions, pathways & diseases (874)
Protein families (114)
Enzymes (2)
Literature (70,654)
Ontologies (50)
EBI web (54)
Macromolecular structures (422 results found)
DNA-free p53 Tetramer Reconstruction
Authors: Pham N, Lucumi A, Cheung N, Viadiu H
Method: single-particle reconstruction
Specimen state: particle
Related data
Views
Source: EMDB
ID: EMD-5342
View all 422 results for Macromolecular structures
Genomes (704 results found)
tumor protein p53
Approved Symbol: TP53 Approved Name: tumor protein p53 Status: (Approved)
Aliases: p53 LFS1
Locus Type: gene with protein product
Chromosome: 17p13.1
Related data
Views
Source: HGNC
ID: HGNC:11998
Summary information is available for this gene
View all 704 results for Genomes
Nucleotide sequences (20,113 results found)
AB020317
Mus musculus mRNA for p53, complete cds.
Related data
Views
Source: Sequence (Release)
ID: AB020317
AB021961
Mus musculus mutant p53 mRNA, complete cds.
Related data
Views
Source: Sequence (Release)
ID: AB021961
View all 20113 results for Nucleotide sequences


EMBL-EBI
Services
Research
Training
Industry
About us
EBI Search
p53
Search
Examples: VAV_HUMAN, tpi1, Sulston ...
Advanced
Help & Documentation
About EBI Search
Share
Feedback
Search results for p53
Showing 18 results out of 105,372 in All results
Filter your results
Source
All results (105,372)
Genomes (704)
Nucleotide sequences (20,113)
Protein sequences (7,312)
Macromolecular structures (422)
Small molecules (703)
Gene expression (3,301)
Molecular interactions (1,069)
Reactions, pathways & diseases (874)
Protein families (114)
Enzymes (2)
Literature (70,654)
Ontologies (50)
EBI web (54)
All results (105,372)
Molecular interactions (1,069)
IntAct Experiments (743)
IntAct Interactions (226)
IntAct Interactors (100)
IntAct Interactions (226 results found)
p53-p53-1
Homodimerization of p53 shown by ubiquitin reconstruction
p53-p21_promoter
DNA binding properties of p53
p53-p21_p53bind
DNA binding properties of p53
gnl3-p53
Coimmunoprecipitaion reveals interaction between the nucleostemin protein and P53 protein
p53-p21_p53bind-1
DNA binding properties of p53
Reactions, pathways & diseases (874)
OMIM (573)
BioModels (164)
Reactome (137)
Reactome (137 results found)
Stabilization of p53
... al, 1998; Banin et al, 1998 and Khanna et al, 1998). ATM also regulates the phosphorylation of p53 ... (Chehab et al, 2000; Shieh et al, 2000; Hirao et al 2000). Phosphorylation of p53 at Ser-20 interferes with p53-MDM2 interaction. MDM2 is transcriptionally activated by p53 and is a negative regulator of ...

http://www.ebi.ac.uk/intact/pages/interactions/interactions.xhtml?query=EBI-1799550&filter=ac



Metabolic & Signalling Pathways

Kyoto Encyclopedia of Genes &Genomes

http://www.genome.ad.jp/kegg/



Search KEGG for p53 Go Clear

Database: KEGG - Search term: p53

KEGG PATHWAY

map04115
p53 signaling pathway

KEGG BRITE

ko00001
KO; KEGG Orthology (KO)
ko00002
Module; KEGG modules
ko01000
Enzyme; Enzymes
ko01001
Kinase; Protein kinases
ko01009
Protein phosphatases and associated proteins; Protein phosphatases and associated proteins
• • • » display all

KEGG ORTHOLOGY

K04451
TP53, P53; tumor protein p53
K05627
BAIAP2, IRSP53; BAI1-associated protein 2
K08851
TP53RK, PRPK, BUD32; TP53 regulating kinase [EC:2.7.11.1]
K10128
RPRM; reprimo, TP53-dependent G2 arrest mediator candidate
K10133
TP53I3; tumor protein p53-inducible protein 3 [EC:1.-.-.-]
• • • » display all

KEGG GENES

hsa:27296
TP53TG5, C20orf10, CLG01; TP53 target 5
hsa:8925
HERC1, p532, p619; HECT and RLD domain containing E3 ubiquitin protein ligase family member 1; K10594 E3 ubiquitin-protein ligase HERC1 [EC:6.3.2.19]
hsa:112858
TP53RK, BUD32, C20orf64, Nori-2, Nori-2p, PRPK, dJ101A2; TP53 regulating kinase (EC:2.7.11.1); K08851 TP53 regulating kinase [EC:2.7.11.1]
hsa:90313
TP53I13, DSCP1; tumor protein p53 inducible protein 13
hsa:10458
BAIAP2, BAP2, FLAF3, IRSP53; BAI1-associated protein 2; K05627 BAI1-associated protein 2
• • • » display all

KEGG DGENES

T10002:9129
K08851 TP53 regulating kinase [EC:2.7.11.1]
T10003:9362
K08851 TP53 regulating kinase [EC:2.7.11.1]
T10004:9514
K08851 TP53 regulating kinase [EC:2.7.11.1]
T10005:1048
K10133 tumor protein p53-inducible protein 3 [EC:1.-.-.-]
T10005:2716
K10136 TP53 apoptosis effector
• • • » display all

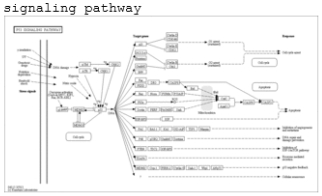
KEGG MGENES

T30027:GL0163842
[Lack_3'-end][KO:K01409][COG:COG3642]; K08851 TP53 regulating kinase [EC:2.7.11.1]
T30027:GL0239786
[Lack_5'-end][KO:K08851][COG:COG3642][KOG:KOG3087][eggNOG:fuNOG06542 meNOG17527 veNOG18025]; K08851 TP53 regulating kinase [EC:2.7.11.1]
T30079:GL0055487
[Complete][KO:K08851][COG:COG3642][KOG:KOG3087]; K08851 TP53 regulating kinase [EC:2.7.11.1]
T30126:GL0061334
[Lack_3'-end][KO:K00344]; K10133 tumor protein p53-inducible protein 3 [EC:1.-.-.-]
T30146:73081
similar to uniprot|P53189 saccharomyces cerevisiae YGL028c SCW11
• • • » display all

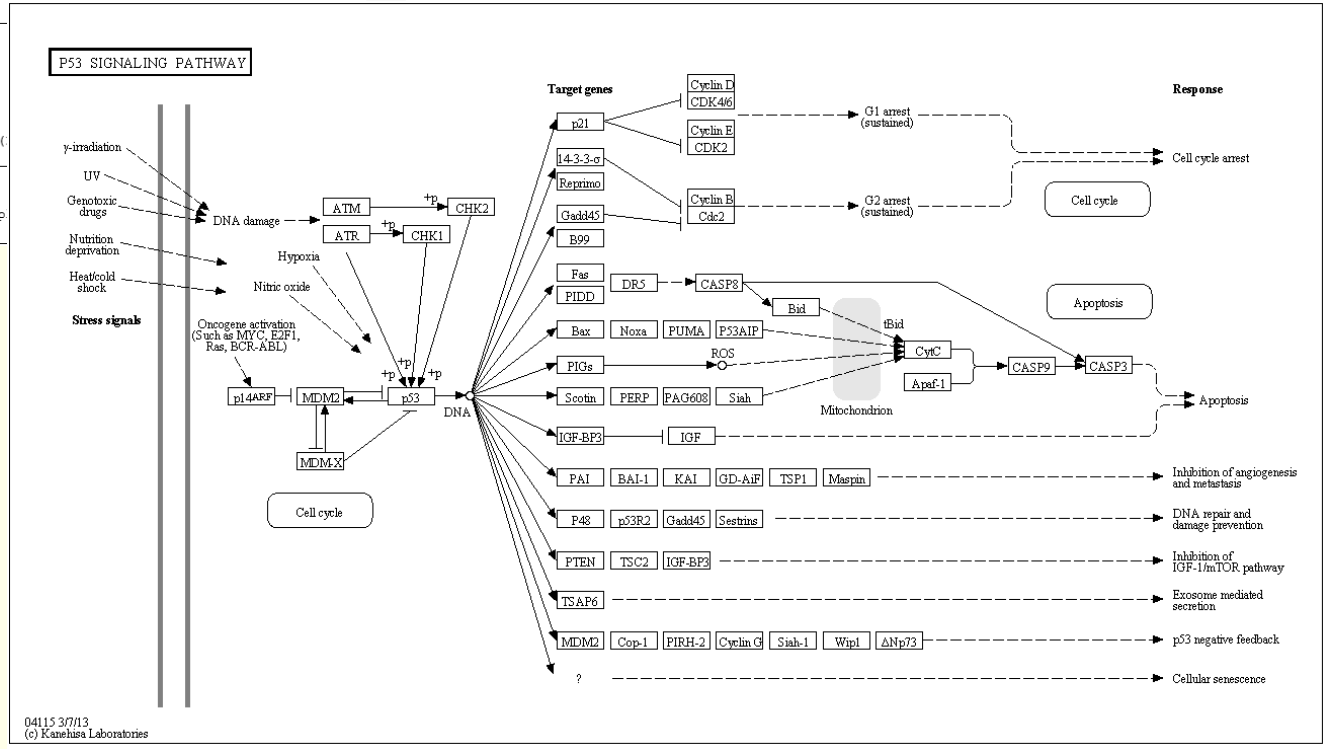
KEGG

PATHWAY: map04115

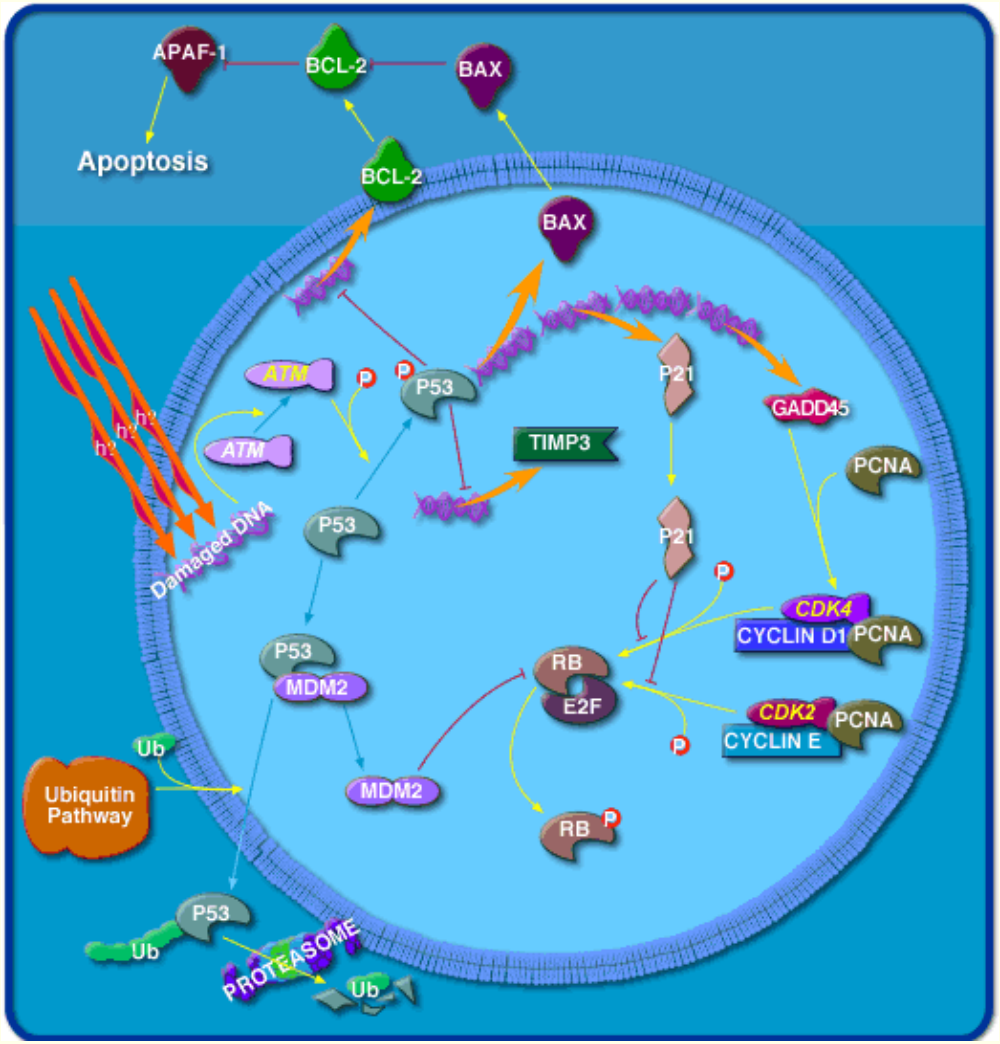
Help

Entry	map04115 Pathway
Name	p53 signaling pathway
Description	p53 activation is induced by a number of stress signals, including DNA damage, oxidative stress and activated oncogenes. The p53 protein is employed as a transcriptional activator of p53-regulated genes. This results in three major outputs; cell cycle arrest, cellular senescence or apoptosis. Other p53-regulated gene functions communicate with adjacent cells, repair the damaged DNA or set up positive and negative feedback loops that enhance or attenuate the functions of the p53 protein and integrate these stress responses with other signal transduction pathways.
Class	Cellular Processes; Cell growth and death BRITE hierarchy
Pathway map	map04115 p53 signaling pathway  All organisms Ortholog table
Disease	H00409 Type II diabetes mellitus H00848 Ataxia with ocular apraxia (AOA) H00881 Li-Fraumeni syndrome H00896 Lymphangioleiomyomatosis (LAM) H00915 Tuberous sclerosis complex (TSC) H00992 Seckel syndrome H01007 Choroid plexus papilloma H01106 Plasminogen activator inhibitor type 1 (H01195 VACTERL/VATER association
Reference	PMID:16557269
Authors	Levine AJ, Hu W, Feng Z.
Title	The P53 pathway: what questions remain to be exp
Journal	Cell Death Differ 13:1027-36 (2006)

http://www.genome.jp/kegg-bin/show_pathway?map04115



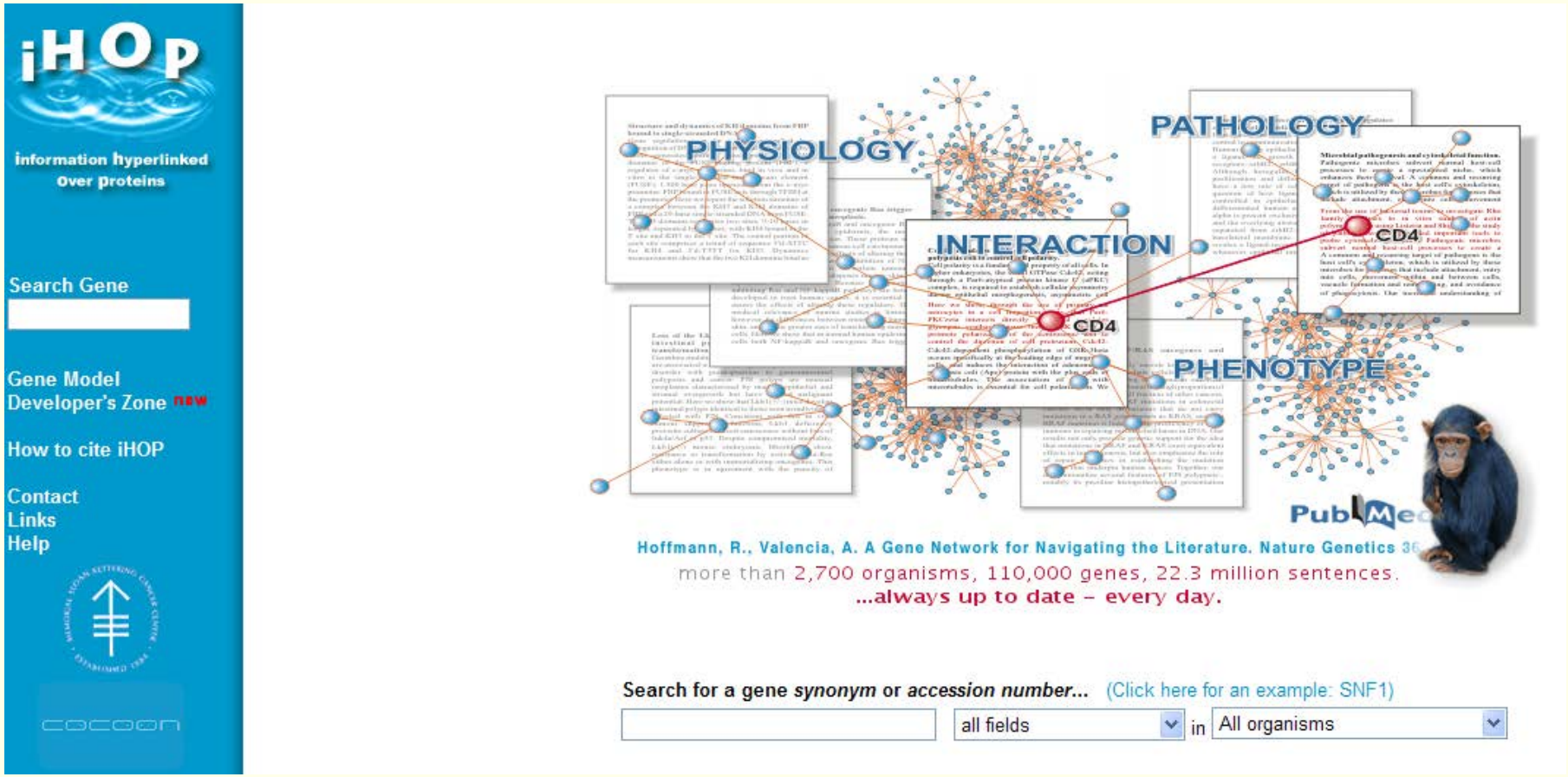
Metabolic & Signalling Pathways



BIOCARTA

Biocarta
(*c*)

http://www.ihop-net.org/UniPub/iHOP/



iHOP

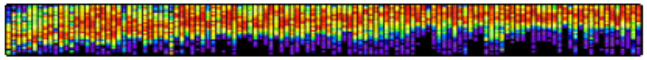
information hyperlinked over proteins

Search Gene

Gene Model Developer's Zone new

How to cite iHOP

Contact
Links
Help

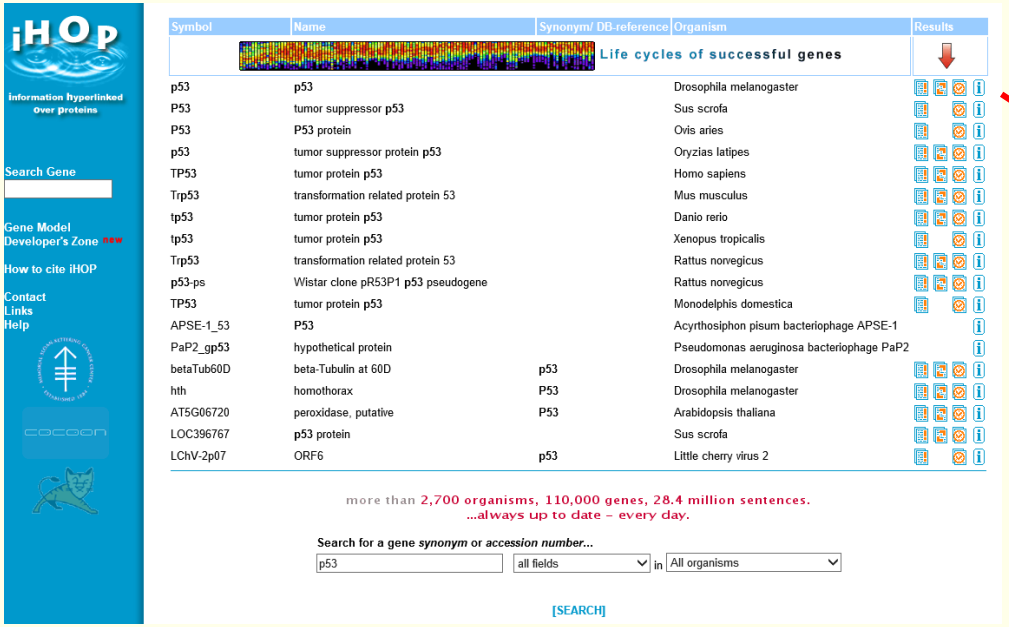
Symbol	Name	Synonym/ DB-reference	Organism	Results
	 Life cycles of successful genes			
p53	p53		Drosophila melanogaster	
P53	tumor suppressor p53		Sus scrofa	
P53	P53 protein		Ovis aries	
p53	tumor suppressor protein p53		Oryzias latipes	
TP53	tumor protein p53		Homo sapiens	
Trp53	transformation related protein 53		Mus musculus	
tp53	tumor protein p53		Danio rerio	
tp53	tumor protein p53		Xenopus tropicalis	
Trp53	transformation related protein 53		Rattus norvegicus	
p53-ps	Wistar clone pR53P1 p53 pseudogene		Rattus norvegicus	
TP53	tumor protein p53		Monodelphis domestica	
APSE-1_53	P53		Acyrthosiphon pisum bacteriophage APSE-1	
PaP2_gp53	hypothetical protein		Pseudomonas aeruginosa bacteriophage PaP2	
betaTub60D	beta-Tubulin at 60D	p53	Drosophila melanogaster	
hth	homothorax	P53	Drosophila melanogaster	
AT5G06720	peroxidase, putative	P53	Arabidopsis thaliana	
LOC396767	p53 protein		Sus scrofa	
LChV-2p07	ORF6	p53	Little cherry virus 2	

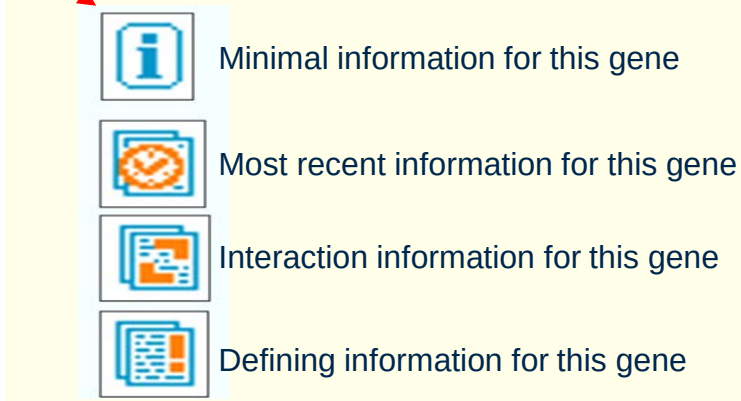
more than 2,700 organisms, 110,000 genes, 28.4 million sentences.
...always up to date – every day.

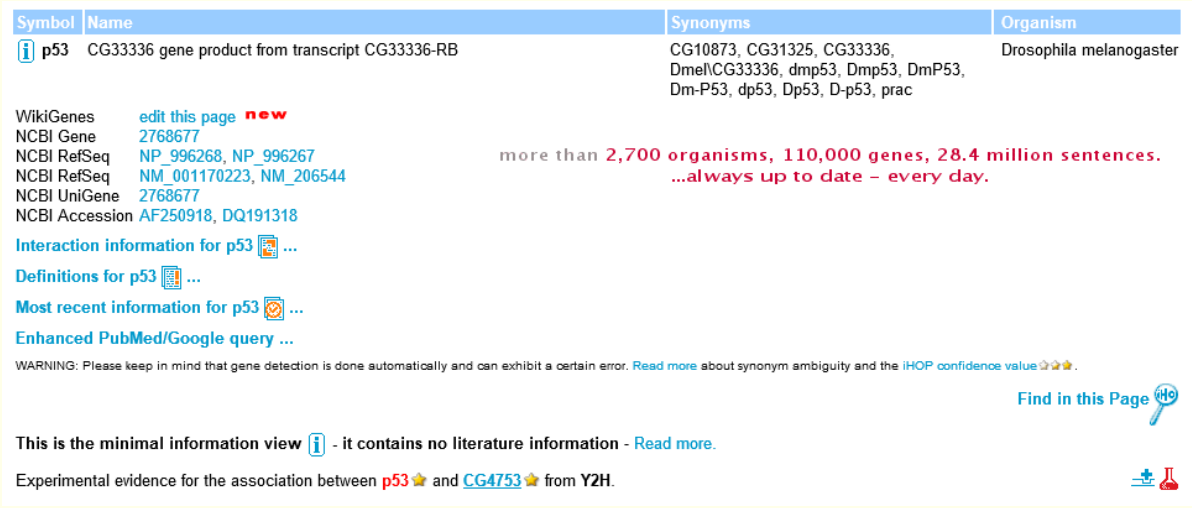
Search for a gene *synonym* or *accession number*...

p53 | all fields | in | All organisms

[SEARCH]


iHOP
Information hyperlinked over proteins
Search Gene
Gene Model Developer's Zone new
How to cite iHOP
Contact
Links
Help
Symbol
Name
Synonym/ DB-reference
Organism
Results
Life cycles of successful genes
p53 p53 Drosophila melanogaster
P53 tumor suppressor p53 Sus scrofa
P53 P53 protein Ovis aries
p53 tumor suppressor protein p53 Oryzias latipes
TP53 tumor protein p53 Homo sapiens
Trp53 transformation related protein 53 Mus musculus
tp53 tumor protein p53 Danio rerio
tp53 tumor protein p53 Xenopus tropicalis
Trp53 transformation related protein 53 Rattus norvegicus
p53-ps Wistar clone pR53P1 p53 pseudogene Rattus norvegicus
TP53 tumor protein p53 Monodelphis domestica
APSE-1_53 P53 Acyrthosiphon pisum bacteriophage APSE-1
PaP2_gp53 hypothetical protein Pseudomonas aeruginosa bacteriophage PaP2
betaTub60D beta-Tubulin at 60D p53 Drosophila melanogaster
hth homothorax P53 Drosophila melanogaster
AT5G06720 peroxidase, putative P53 Arabidopsis thaliana
LOC396767 p53 protein Sus scrofa
LChV-2p07 ORF6 p53 Little cherry virus 2
more than 2,700 organisms, 110,000 genes, 28.4 million sentences.
...always up to date – every day.
Search for a gene synonym or accession number...
p53
all fields
in
All organisms
[SEARCH]


Minimal information for this gene
Most recent information for this gene
Interaction information for this gene
Defining information for this gene


Symbol
Name
Synonyms
Organism
p53 CG33336 gene product from transcript CG33336-RB
CG10873, CG31325, CG33336, Dmel\CG33336, dmp53, Dmp53, DmP53, Dm-P53, dp53, Dp53, D-p53, prac
Drosophila melanogaster
WikiGenes edit this page new
NCBI Gene 2768677
NCBI RefSeq NP_996268, NP_996267
NCBI RefSeq NM_001170223, NM_206544
NCBI UniGene 2768677
NCBI Accession AF250918, DQ191318
more than 2,700 organisms, 110,000 genes, 28.4 million sentences.
...always up to date – every day.
Interaction information for p53 ...
Definitions for p53 ...
Most recent information for p53 ...
Enhanced PubMed/Google query ...
WARNING: Please keep in mind that gene detection is done automatically and can exhibit a certain error. Read more about synonym ambiguity and the iHOP confidence value.
Find in this Page
This is the minimal information view - it contains no literature information - Read more.
Experimental evidence for the association between p53 and CG4753 from Y2H.

Most recent information for this gene.

Symbol	Name	Synonyms	Organism
p53	CG33336 gene product from transcript CG33336-RB	CG10873, CG31325, CG33336, Dmel\CG33336, dmp53, Dmp53, DmP53, Dm-P53, dp53, Dp53, D-p53, prac	Drosophila melanogaster

WikiGenes edit this page new
NCBI Gene 2768677
NCBI RefSeq NP_996268, NP_996267
NCBI RefSeq NM_001170223, NM_206544
NCBI UniGene 2768677
NCBI Accession AF250918, DQ191318

more than 2,700 organisms, 110,000 genes, 28.4 million sentences.
...always up to date – every day.

Interaction information for p53 ...

Definitions for p53 ...

Enhanced PubMed/Google query ...

WARNING: Please keep in mind that gene detection is done automatically and can exhibit a certain error. Read more about synonym ambiguity and the iHOP confidence value.

Find in this Page

Sentences in this view contain the most recent information on p53 - Most recent information is available whenever you see this symbol - Read more.
For a summary overview of the information in this page click here. new

Show all
Order by relevance

We report here that dE2F1 **promotes** IR-**induced** p53-independent apoptosis in larval imaginal discs. [2010]

DmChk2 **induced** cell death was not suppressed by Dmp53 mutant, revealing for the first time that in Drosophila, over-expression of DmChk2 can **induce** cell death which is independent of Dmp53. [2010]

E2F1 and E2F2 have opposite **effects** on radiation-**induced** p53-independent apoptosis in Drosophila. [2010]

Moreover, our results demonstrate that although DmChk2 and Dmp53 affect different stages of ovarian development, loss of ovarian stem cells by p53 expression and mid-oogenesis cell death **induced** by DmChk2 do not require caspase activity. [2010]

Furthermore, a rough eye phenotype because of overexpression of DREF in Drosophila eye imaginal disks could be **suppressed** by half dose reduction of the dmp53 gene. [2010]

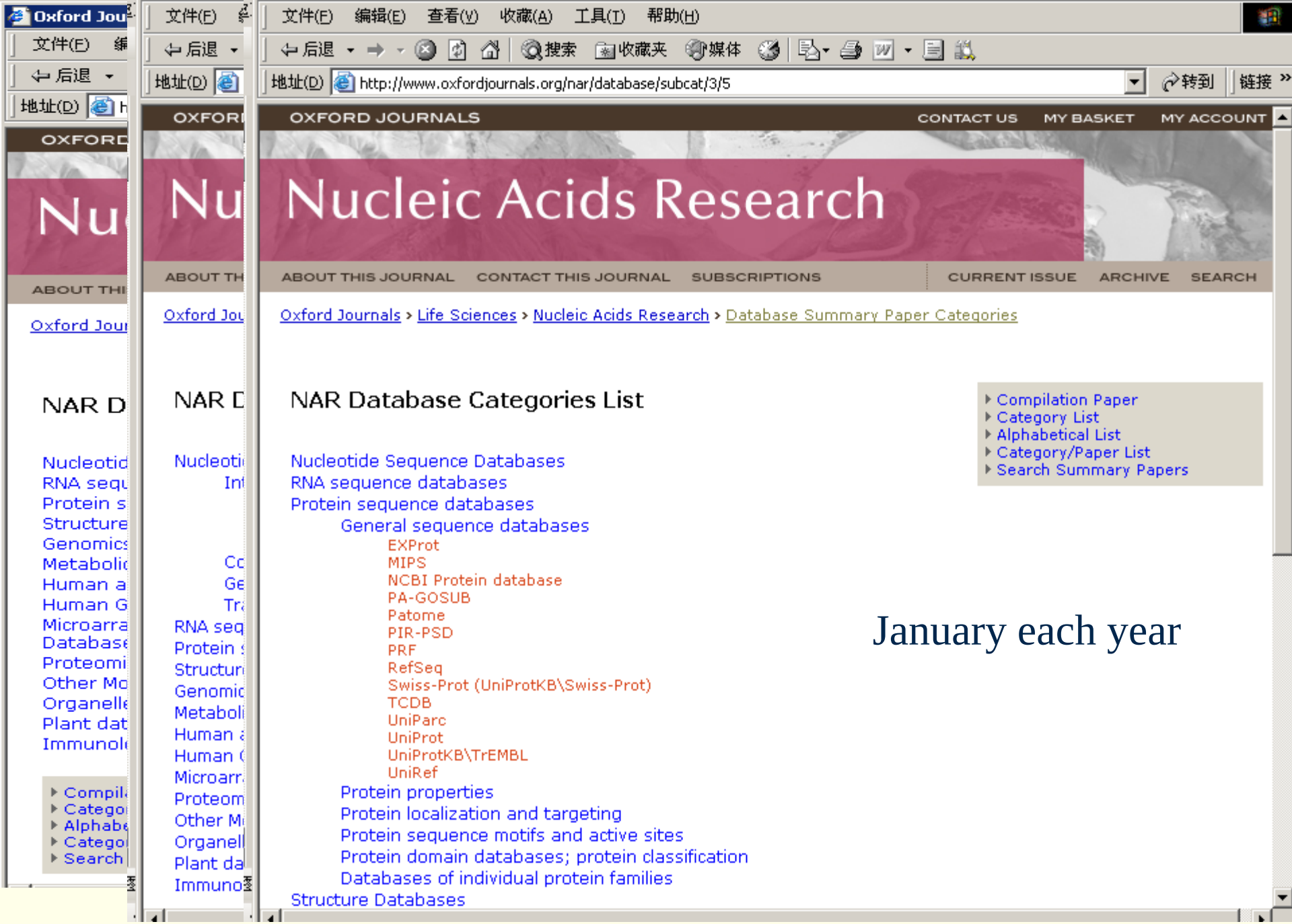
DREF **binds** specifically to DRE sequences in the Drosophila p53 (dmp53) gene promoter as shown using anti-DREF antibodies in chromatin immunoprecipitation assays. [2010]

The data demonstrate that wild-type p53 over-expression can **regulate** life span independent of foxo, and suggest that foxo acts in males to **produce** sexually antagonistic life span **effects** of p53. [2010]

We propose that in response to UV irradiation, Dmp53 **enhances** the expression of the dKDM4B histone demethylase, which demethylates H3K9me3 preferentially in heterochromatin regions. [2010]

In addition, the level of mRNA of dmp53 was decreased in DREF-knockdown cells and transient expression of the luciferase gene under **control** of the wild-type dmp53 gene promoter showed strong promoter activity in S2 [?] cells, but this was almost completely abrogated with a DRE-mutated promoter. [2010]

The ability of ionizing radiation (IR) to induce apoptosis independent of p53 is crucial for successful therapy of cancers bearing p53 mutations. p53-independent apoptosis, however, remains poorly understood relative to p53-dependent apoptosis. [2010]


文件(F) 编辑(E) 查看(V) 收藏(A) 工具(T) 帮助(H)
后退 搜索 收藏夹 媒体
地址(D) http://www.oxfordjournals.org/nar/database/subcat/3/5
转到 链接
OXFORD JOURNALS
CONTACT US MY BASKET MY ACCOUNT
Nucleic Acids Research
ABOUT THIS JOURNAL CONTACT THIS JOURNAL SUBSCRIPTIONS
CURRENT ISSUE ARCHIVE SEARCH
Oxford Journals › Life Sciences › Nucleic Acids Research › Database Summary Paper Categories
NAR Database Categories List
Compilation Paper
Category List
Alphabetical List
Category/Paper List
Search Summary Papers
Nucleotide Sequence Databases
RNA sequence databases
Protein sequence databases
General sequence databases
EXProt
MIPS
NCBI Protein database
PA-GOSUB
Patome
PIR-PSD
PRF
RefSeq
Swiss-Prot (UniProtKB\Swiss-Prot)
TCDB
UniParc
UniProt
UniProtKB\TrEMBL
UniRef
Protein properties
Protein localization and targeting
Protein sequence motifs and active sites
Protein domain databases; protein classification
Databases of individual protein families
Structure Databases
January each year

Exercise

Use the database searching techniques you learned today to retrieve the reference sequence

Homo sapiens LEGUMAIN

and the amino acid sequence of

ALL LEGUMAIN

(Store all the fasta sequences in your USB or Notebook)

AND

Everything about LEGUMAIN

There are many different ways to DO it.