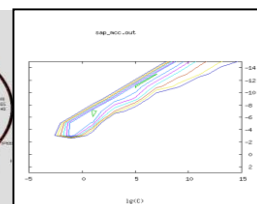
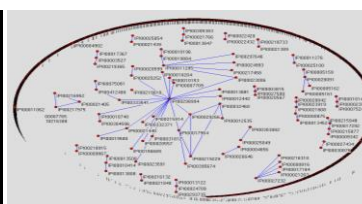
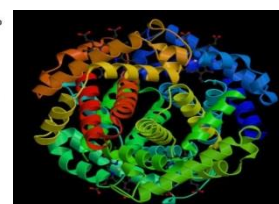
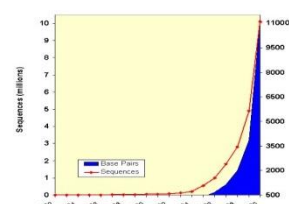


北京大学生物信息学中心 魏丽萍

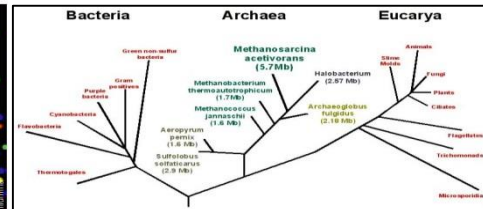
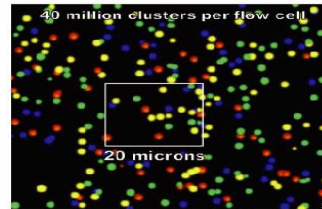
# Liping Wei, Ph.D.

# Center for Bioinformatics, Peking University





TAACCCTAACCCTAACCCTAACCCTAACCCTA  
 CCTAACCCCTAACCCCTAACCCCTAACCCCTAACCC  
 CCCTAACCCCTAACCCCTAACCCCTAACCCCTAAC  
 AACCCCTAACCCCTAACCCCTAACCCCTAACCCCTA  
 ACCCTAACCCCAACCCCAACCCCAACCCCAAC  
 CTACCCTAACCCCTAACCCCTAACCCCTAACCCCTA  
 ACCCTAACCCCTAACCCCTAACCCCTAACCCCTAA

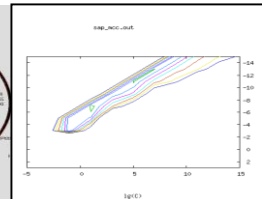
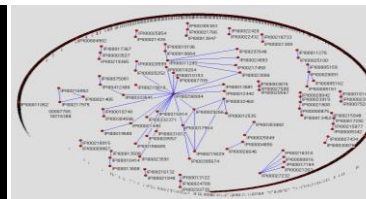
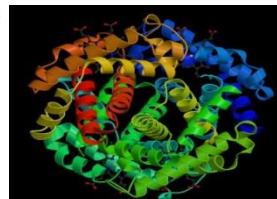
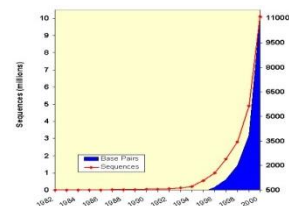


# Unit 2: National Center for Biotechnology Information

北京大学生物信息学中心 魏丽萍

Liping Wei, Ph.D.

Center for Bioinformatics, Peking University



# NCBI: <http://www.ncbi.nlm.nih.gov/>

NCBI Resources ▾ How To ▾ Sign in to NCBI

NCBI  
National Center for Biotechnology Information

Gene ▾

- All Databases
- Assembly
- BioProject
- BioSample
- BioSystems
- Books
- ClinVar
- Clone
- Conserved Domains
- dbGaP
- dbVar
- Epigenomics
- EST
- Gene**
- Genome
- GEO DataSets
- GEO Profiles
- GSS
- HomoloGene
- MedGen

Search

NCBI

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

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

Popular Resources

- [PubMed](#)
- [Bookshelf](#)
- [PubMed Central](#)
- [PubMed Health](#)
- [BLAST](#)
- [Nucleotide](#)
- [Genome](#)
- [SNP](#)
- [Gene](#)
- [Protein](#)
- [PubChem](#)

NCBI Announcements

Planned change in bacterial strain-level information management

Nov 21, 2013

Please be aware that there is an upcoming change (January 2014) in how

Exploring next-gen sequencing

Genetic Testing Registry

A portal to clinical genetics resources with detailed information about genetic tests and laboratories.

GO


1 2 3 4 5 6 7 8

# Examples of resources at NCBI

|                             |                                   |
|-----------------------------|-----------------------------------|
| <b>Data Repository</b>      | GenBank, dbEST, GEO, SRA          |
| <b>DNA</b>                  | Genome, Gene                      |
| <b>Comparative Genomics</b> | Taxonomy, HomoloGene              |
| <b>Genetic Variation</b>    | dbSNP, dbVar                      |
| <b>Disease Mutations</b>    | OMIM, dbGaP, ClinVar              |
| <b>RNA</b>                  | RefSeq, UniGene                   |
| <b>Proteins</b>             | Protein, RefSeq, Conserved Domain |
| <b>Literature</b>           | PubMed, MeSH                      |
| <b>Tools</b>                | BLAST                             |

# NCBI-Genome

<http://www.ncbi.nlm.nih.gov/genome/>

 [Resources](#) ☒ [How To](#) ☒

Genome

[Limits](#) [Advanced](#) [Help](#)



## Genome

This resource organizes information on genomes including sequences, maps, chromosomes, assemblies, and annotations.

### Using Genome

[Help](#)  
[Browse by Organism](#)  
[Download / FTP](#)  
[Submit a genome](#)

### Custom resources

[Human Genome](#)  
[Microbes](#)  
[Organelles](#)  
[Plants](#)  
[Viruses](#)

### Other Resources

[Assembly](#)  
[BioProject](#)  
[BioSample](#)  
[Map Viewer](#)  
[Protein Clusters](#)

### Genome Tools

[BLAST the Human Genome](#)  
[Genomic groups BLAST](#)  
[NCBI remap](#)  
[Genome Decoration Page](#)

### Genome Annotation and Analysis

[Eukaryotic Genome Annotation](#)  
[Prokaryotic Genome Annotation](#)  
[PASC \(Pairwise Sequence Comparison\)](#)  
[TaxPlot \(3-way Genome Comparison\)](#)

### External Resources

[GOLD - Genomes Online Database](#)  
[Ensembl Genome Browser](#)  
[Bacteria Genomes at Sanger](#)  
[Large-Scale Genome Sequencing \(NHGRI\)](#)

NCBI

Human genome overview page (Annotation Release 105)

Human genome overview page (Build 36.3)

[Map Viewer Home](#)

Map Viewer Help

Human Maps Help

FTP

Data As Table View

**Maps & Options**

Region Shown:

Go

out zoom in

You are here:

Ideogram



PubMed Entrez BLAST OMIM Taxonomy Structure

Search  Find Find in This View Advanced Search

[Homo sapiens \(human\)](#) [Annotation Release 105 \(Current\)](#) [BLAST human sequences](#)

Chromosome: 1 2 3 [ 4 ] 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 X Y MT

Master Map: Genes On Sequence Summary of Maps Maps & Options

Region Displayed: 0-191M bp

| Ideogram | Hs UniG   | Genes_seq | Symbol   | Links                                 | E           | Cyto     | Description                          |
|----------|-----------|-----------|----------|---------------------------------------|-------------|----------|--------------------------------------|
| 4p16.3   | Hs.741355 |           | FGFR3    | OMIM HGNC sv pr dl ev hm sts CCDS SNP | best RefSeq | 4p16.3   | fibroblast growth factor receptor 3  |
| 4p16.2   | Hs.741128 |           | HTT      | OMIM HGNC sv pr dl ev hm sts CCDS SNP | best RefSeq | 4p16.3   | huntingtin                           |
| 4p16.1   | Hs.741127 |           | PROM1    | OMIM HGNC sv pr dl ev hm sts CCDS SNP | best RefSeq | 4p15.32  | prominin 1                           |
| 4p15.33  | Hs.741135 |           | PPARGC1A | OMIM HGNC sv pr dl ev hm sts CCDS SNP | best RefSeq | 4p15.1   | peroxisome proliferator-activated r  |
| 4p15.32  | Hs.741137 |           | PDGFRA   | OMIM HGNC sv pr dl ev hm sts CCDS SNP | best RefSeq | 4q12     | platelet-derived growth factor recej |
| 4p15.31  | Hs.741133 |           | KIT      | OMIM HGNC sv pr dl ev hm sts CCDS SNP | best RefSeq | 4q12     | v-kit Hardy-Zuckerman 4 feline sai   |
| 4p15.2   | Hs.741136 |           | KDR      | OMIM HGNC sv pr dl ev hm sts CCDS SNP | best RefSeq | 4q11-q12 | kinase insert domain receptor (a ty  |
| 4p15.1   | Hs.655705 |           | ALB      | OMIM HGNC sv pr dl ev hm sts CCDS SNP | best RefSeq | 4q13.3   | albumin                              |
| 4p14     | Hs.570770 |           | IL8      | OMIM HGNC sv pr dl ev hm sts CCDS SNP | best RefSeq | 4q13-q21 | interleukin 8                        |
| 4p13     | Hs.733448 |           | CXCL10   | OMIM HGNC sv pr dl ev hm sts CCDS SNP | best RefSeq | 4q21     | chemokine (C-X-C motif) ligand 10    |
| 4p13.1   | Hs.731997 |           | SPP1     | OMIM HGNC sv pr dl ev hm sts CCDS SNP | best RefSeq | 4q22.1   | secreted phosphoprotein 1            |
| 4p13.2   | Hs.527104 |           | ABCG2    | OMIM HGNC sv pr dl ev hm sts CCDS SNP | best RefSeq | 4q22     | ATP-binding cassette, sub-family C   |
| 4p13.3   | Hs.570759 |           |          |                                       |             |          |                                      |
| 4p13.1   | Hs.444229 |           |          |                                       |             |          |                                      |
| 4p13.2   | Hs.125593 |           |          |                                       |             |          |                                      |
| 4p13.3   | Hs.654735 |           |          |                                       |             |          |                                      |
| 4p13.1   | Hs.458281 |           |          |                                       |             |          |                                      |
| 4p13.2   | Hs.680226 |           |          |                                       |             |          |                                      |
| 4p13.3   | Hs.734503 |           |          |                                       |             |          |                                      |

# NCBI-Nucleotide/Protein (RefSeq)

[www.ncbi.nlm.nih.gov/nuccore/](http://www.ncbi.nlm.nih.gov/nuccore/) [www.ncbi.nlm.nih.gov/protein/](http://www.ncbi.nlm.nih.gov/protein/)


NCBI Resources How To Sign in to NCBI

Nucleotide

Nucleotide

Search

Limits Advanced Help



## Nucleotide

The Nucleotide database is a collection of sequences from several sources, including GenBank, RefSeq, TPA and PDB. Genome, gene and transcript sequence data provide the foundation for biomedical research and discovery.

| Using Nucleotide                  | Nucleotide Tools                  | Other Resources              |
|-----------------------------------|-----------------------------------|------------------------------|
| <a href="#">Quick Start Guide</a> | <a href="#">Submit to GenBank</a> | <a href="#">GenBank Home</a> |
| <a href="#">FAQ</a>               | <a href="#">LinkOut</a>           | <a href="#">RefSeq Home</a>  |
| <a href="#">Help</a>              | <a href="#">E-Utilities</a>       | <a href="#">Gene Home</a>    |
| <a href="#">GenBank FTP</a>       | <a href="#">BLAST</a>             | <a href="#">SRA Home</a>     |
| <a href="#">RefSeq FTP</a>        | <a href="#">Batch Entrez</a>      | <a href="#">INSDC</a>        |

You are here: NCBI > DNA & RNA > Nucleotide Database [Write to the Help Desk](#)

| GETTING STARTED                  | RESOURCES                                 | POPULAR                   | FEATURED                                 | NCBI INFORMATION                 |
|----------------------------------|---|---------------------------|--|----------------------------------|
| <a href="#">NCBI Education</a>   | <a href="#">Chemicals &amp; Bioassays</a> | <a href="#">PubMed</a>    | <a href="#">Genetic Testing Registry</a> | <a href="#">About NCBI</a>       |
| <a href="#">NCBI Help Manual</a> | <a href="#">Data &amp; Software</a>       | <a href="#">Bookshelf</a> | <a href="#">PubMed Health</a>            | <a href="#">Research at NCBI</a> |


NCBI Resources How To Sign in to NCBI

Protein

Protein

Search

Advanced Help



## Protein

The Protein database is a collection of sequences from several sources, including translations from annotated coding regions in GenBank, RefSeq and TPA, as well as records from SwissProt, PIR, PRF, and PDB. Protein sequences are the fundamental determinants of biological structure and function.

| Using Protein                     | Protein Tools                | Other Resources              |
|-----------------------------------|------------------------------|------------------------------|
| <a href="#">Quick Start Guide</a> | <a href="#">BLAST</a>        | <a href="#">GenBank Home</a> |
| <a href="#">FAQ</a>               | <a href="#">LinkOut</a>      | <a href="#">RefSeq Home</a>  |
| <a href="#">Help</a>              | <a href="#">E-Utilities</a>  | <a href="#">CDD</a>          |
| <a href="#">GenBank FTP</a>       | <a href="#">Blink</a>        | <a href="#">Structure</a>    |
| <a href="#">RefSeq FTP</a>        | <a href="#">Batch Entrez</a> |                              |

You are here: NCBI > Proteins > Protein Database [Write to the Help Desk](#)

| GETTING STARTED                  | RESOURCES                                 | POPULAR                   | FEATURED                                 | NCBI INFORMATION                 |
|----------------------------------|---|---------------------------|--|----------------------------------|
| <a href="#">NCBI Education</a>   | <a href="#">Chemicals &amp; Bioassays</a> | <a href="#">PubMed</a>    | <a href="#">Genetic Testing Registry</a> | <a href="#">About NCBI</a>       |
| <a href="#">NCBI Help Manual</a> | <a href="#">Data &amp; Software</a>       | <a href="#">Bookshelf</a> | <a href="#">PubMed Health</a>            | <a href="#">Research at NCBI</a> |



# NCBI-Nucleotide/Protein (RefSeq): [NM\_\*, NP\_\*]

[www.ncbi.nlm.nih.gov/nuccore/](http://www.ncbi.nlm.nih.gov/nuccore/)

[www.ncbi.nlm.nih.gov/protein/](http://www.ncbi.nlm.nih.gov/protein/)

## Homo sapiens potassium channel, subfamily K, member 2 (KCNK2), transcript variant 1, mRNA

NCBI Reference Sequence: NM\_001017424.2

[FASTA](#) [Graphics](#)

[Go to:](#)

LOCUS NM\_001017424 3597 bp mRNA linear PRI 17-NOV-2013  
DEFINITION Homo sapiens potassium channel, subfamily K, member 2 (KCNK2), transcript variant 1, mRNA.  
ACCESSION NM\_001017424  
VERSION NM\_001017424.2 GI:126365744  
KEYWORDS RefSeq.  
SOURCE Homo sapiens (human)  
ORGANISM [Homo sapiens](#)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 3597)

FEATURES  
source Location/Qualifiers  
1..3597  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="1"  
/map="lq41"  
1..3597  
/gene="KCNK2"  
/gene\_synonym="hTREK-1c; hTREK-1e; K2p2.1; TPKC1; TREK; TREK-1; TREK1"  
/note="potassium channel, subfamily K, member 2"  
/db\_xref="GeneID:3176"  
/db\_xref="HGNC:6277"  
/db\_xref="CCDS:21961.1"

ORIGIN  
1 acaaacctgt gagaataagt gacgctgcca actttgatat gtcattcat aaggaagaaa  
61 ctctctgtct ttaaatctc tggaaacact gggatatagt toctcatttt aatcaggaaa  
121 ttctctctgt aaccagactg taataaacaga ggaattgagt agcacatttc acttgcatac  
181 agcagcgttt gtctcttttg agtgtttttg tacaggaaac tgatctacag ggtaacattc  
241 acacacaaaa aggaacacagt atgggacgat ggcctgttag tacaagtgcac tctgtttttg  
301 aaattacgga caagaacact tggaggaaga acggcattaa agatcaaaag catgatgact  
361 cctgatgaaa acatcaccaa atgatgaacc cagagcaaaa aagggtattc taactggcgg  
421 cacctgactt gctgatcctc aaatctgcgc ctcagaactc caaaccgagg ctctcgtttt  
481 ccacgaacac cacagtgcct gcttcocggg tggagagtga caagaccatt aatgttatga  
541

## potassium channel subfamily K member 2 isoform a [Homo sapiens]

NCBI Reference Sequence: NP\_001017424.1

[FASTA](#) [Graphics](#)

[Go to:](#)

LOCUS NP\_001017424 422 aa linear PRI 17-NOV-2013  
DEFINITION potassium channel subfamily K member 2 isoform a [Homo sapiens].  
ACCESSION NP\_001017424  
VERSION NP\_001017424.1 GI:62912487  
DBSOURCE REFSEQ: accession [NM\\_001017424.2](#)  
KEYWORDS RefSeq.  
SOURCE Homo sapiens (human)  
ORGANISM [Homo sapiens](#)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (residues 1 to 422)


FEATURES  
source Location/Qualifiers  
1..422  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="1"  
/map="lq41"  
1..422  
/product="potassium channel subfamily K member 2 isoform a"  
/note="potassium inwardly-rectifying channel, subfamily K, member 2; TWIK-related potassium channel 1; two-pore potassium channel 1; tandem-pore-domain potassium channel TREK-1; K2P2.1 potassium channel; TREK-1 K(+) channel subunit; two pore potassium channel TPKC1; two pore domain potassium channel TREK-1; outward rectifying potassium channel protein TREK-1"  
/calculated\_mol\_wt=46757  
117..189  
/region\_name="Ion\_trans\_2"  
/note="Ion channel; pfam07885"  
/db\_xref="CDD:21961.9"

ORIGIN  
1 manprakrdr ylaapdlldp ksaaqnsqpr lsfstktptvl asrvesdtti nmvkwktvst  
61 iflvvlylli igatvfkale qpheisqrtr iviqkqtfls qhscvmtel deliqqivaa  
121 inagiipnln tsnqishwdl gssffftagtv ittigfgnis prteggkifc iiyallgipl  
181 fgflilagvgd qlgtifgkgi akvedtfikw nvsqtkirrii stiifilfgc vlfvalpaii  
241 fkhiegwsal daiyfvvitl ttigfgdyva ggsdieylrf ykpvvfwil vglayfaavl  
301 gldldvrvv hthveevge frahaaewta nvtasefketr rrlsveiydk fqratsikrk  
361 gldldvrvv hthveevge frahaaewta nvtasefketr rrlsveiydk fqratsikrk  
421



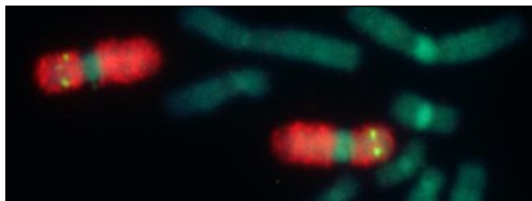
# NCBI-Gene

<http://www.ncbi.nlm.nih.gov/gene/>

 NCBI Resources ☒ How To ☒

Gene

[Limits](#) [Advanced](#) [Help](#)



## Gene

Gene integrates information from a wide range of species. A record may include nomenclature, Reference Sequences (RefSeqs), maps, pathways, variations, phenotypes, and links to genome-, phenotype-, and locus-specific resources worldwide.

### Using Gene

[Gene Quick Start](#)

[FAQ](#)

[Download/FTP](#)

[RefSeq Mailing List](#)

[Gene News](#) 

[Factsheet](#)

### Gene Tools

[Submit GeneRIFs](#)

[Submit Correction](#)

[Statistics](#)

[BLAST](#)

[Genome Workbench](#)

[Splign](#)

### Other Resources

[HomoloGene](#)

[OMIM](#)

[RefSeq](#)

[RefSeqGene](#)

[UniGene](#)

[Protein Clusters](#)

### Representative queries

Find genes by...

free text

chromosome and symbol

Search text

[human muscular dystrophy](#)

[\(11\[chr\] OR 2\[chr\]\) AND adh\\*\[sym\]](#)

# NCBI-Gene

<http://www.ncbi.nlm.nih.gov/gene/>

[Display Settings:](#) ☒ Full Report

[Send to:](#) ☒

## ABCA1 ATP-binding cassette, sub-family A (ABC1), member 1 [ *Homo sapiens* (human) ]

Gene ID: 19, updated on 21-Nov-2013

### Summary

**Official Symbol** ABCA1 provided by [HGNC](#)  
**Official Full Name** ATP-binding cassette, sub-family A (ABC1), member 1 provided by [HGNC](#)  
**Primary source** [HGNC:29](#)  
**See related** [Ensembl: ENSG00000165029](#); [HPRD:02501](#); [MIM:600046](#); [Vega:OTTHUMG00000020417](#)  
**Gene type** protein coding  
**RefSeq status** REVIEWED  
**Organism** [Homo sapiens](#)  
**Lineage** Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorhini; Catarrhini; Hominidae; Homo  
**Also known as** TGD; ABC1; CERP; ABC-1; HDLDT1  
**Summary** The membrane-associated protein encoded by this gene is a member of the superfamily of ATP-binding cassette (ABC) transporters. ABC proteins transport various molecules across extra- and intracellular membranes. ABC genes are divided into seven distinct subfamilies (ABC1, MDR/TAP, MRP, ALD, OABP, GCN20, White). This protein is a member of the ABC1 subfamily. Members of the ABC1 subfamily comprise the only major ABC subfamily found exclusively in multicellular eukaryotes. With cholesterol as its substrate, this protein functions as a cholesterol efflux pump in the cellular lipid removal pathway. Mutations in this gene have been associated with Tangier's disease and familial high-density lipoprotein deficiency. [provided by RefSeq, Jul 2008]

### Genomic context

**Location:** 9q31.1  
**Sequence:** Chromosome: 9; NC\_000009.11 (107543283..107690527, complement)

See ABCA1 in [Epigenomics](#), [MapViewer](#)

### Table of contents

[Summary](#)  
[Genomic context](#)  
[Genomic regions, transcripts, and products](#)  
[Bibliography](#)  
[Phenotypes](#)  
[HIV-1 interactions](#)  
[Variation](#)  
[Interactions](#)  
[Pathways](#)  
[General gene information](#)  
    [Gene Ontology](#)  
[General protein information](#)  
[Reference sequences](#)  
[Related sequences](#)  
[Additional links](#)  
    [Locus-specific Databases](#)

### Related information

[Order cDNA clone](#)  
[BioAssay](#)  
[BioAssay by Target \(List\)](#)  
[BioAssay by Target \(Summary\)](#)  
[BioProjects](#)  
[BioSystems](#)  
[Books](#)


# NCBI-Gene Statistics

<http://www.ncbi.nlm.nih.gov/gene/statistics/>

| Taxa ▲                          | Count of Child Taxa | Total Genes of all children |
|---------------------------------|---------------------|-----------------------------|
| <a href="#">Archaea</a>         | 193                 | 398487                      |
| <a href="#">Bacteria</a>        | 3064                | 8376042                     |
| <a href="#">Eukaryota</a>       | 4675                | 5014420                     |
| <a href="#">Viroids</a>         | 2                   | 4                           |
| <a href="#">Viruses</a>         | 3692                | 172048                      |
| <a href="#">other sequences</a> | 5                   | 80                          |


**Total Taxa: 11631; Total Genes: 13961081**

# Human genes: GeneCards <http://www.genecards.org/>



**GeneCards®**  
The Human Gene Compendium

מכון ויצמן למדע  
WEIZMANN INSTITUTE OF SCIENCE

with 

**LifeMap**  
SCIENCES


Free for academic non-profit institutions. ALL other users need a commercial license from LifeMap Sciences, Inc.

[Home](#) | [GeneCards Guide](#) | [Suite](#) | [Terms and Conditions](#) | [About Us](#) | [User Feedback](#) | [Mirror sites](#)

Search the GeneCards human gene database


Search by  for   [Advanced Search](#) [About V3 Search](#)

**Affiliated Sites**




**MalaCards**  
The Human Malady Compendium


**GeneCards Suite**



**GeneDecks**  
Annotated Gene Sets



**GeneAlaCart**  
GeneCards Batch Queries



**GeneLoc**  
Gene Location

**Explore GeneCards**

**About GeneCards®:** GeneCards is a searchable, integrated database of human genes that provides comprehensive, updated, and user-friendly information on all known and predict [more...](#)

[Extract information for many genes at once:](#) [GeneAlaCart](#) [GeneDecks](#) [Hot genes](#) [Disease genes](#)

**View Sample Gene**

**TNFRSF10B**  
tumor necrosis factor receptor superfamily, member 10b

**GeneCards Sections**

|           |                 |              |          |              |
|-----------|-----------------|--------------|----------|--------------|
| Aliases   | Drugs           | Genome view  | Pathways | Publications |
| Databases | Expression      | Interactions | Paralogs | Summaries    |
| Disorders | External search | IP/Patents   | Products | Transcripts  |
| Domains   | Function        | Orthologs    | Proteins | Variants     |

**View Random Gene**

**MAPK7**  
(GIFTS: 73)  
mitogen-activated protein kinase 7

Category

GIFTS Group

**News and Views**

**Our Trendy Genes**

published in [GEN](#)

**New in V3.11:**  
[IUPHAR ligands](#)  
[MAXQB protein expr](#)  
[Gene/protein families](#)  
[more...](#)

**In our pipeline:**  
[PathCards](#)


**Collaborations:**  
[LifeMap Discovery™](#)  
[SPIRE/MOPED](#)  
[RNAcentral](#)

**Commentary:**  
[Press releases](#)  
[User views](#)

**Versions:**  
Current: 3.11  
17 Nov 2013  
Previous: 3.10

# NCBI-SRA

<http://www.ncbi.nlm.nih.gov/sra>

 NCBI Resources ▾ How To ▾ [Sign in to NCBI](#)

SRA   [Advanced](#) [Help](#)



## SRA

The Sequence Read Archive (SRA) stores raw sequencing data from the next generation of sequencing platforms including Roche 454 GS System®, Illumina Genome Analyzer®, Applied Biosystems SOLiD® System, Helicos Heliscope®, Complete Genomics®, and Pacific Biosciences SMRT®.

### Using SRA

[Handbook](#)

[Download](#)

[E-Utilities](#)

[Factsheet](#)

### Tools

[BLAST](#)

[SRA Run browser](#)

[Submit to SRA](#)

[SRA software](#)

### Other Resources

[SRA Home](#)

[Trace Archive](#)

[Trace Assembly](#)

[GenBank Home](#)

# NCBI-SRA

<http://www.ncbi.nlm.nih.gov/sra>

[Change access](#)

**ERX033410:** Foxa1 and Foxa2 are essential for gender dimorphism in liver cancer  
1 ILLUMINA (Illumina Genome Analyzer IIx) run: 40.9M spots, 1.5G bases, 918.8Mb data

**Accession:** ERX033410

**Experiment design:** Foxa1 and Foxa2 are essential for gender dimorphism in liver cancer

**Submission:** ERA070841 by UNIVERSITY OF PENNSYLVANIA

**Study summary:** Foxa1 and Foxa2 are essential for gender dimorphism in liver cancer

[Study](#) • [All experiments \(more...\)](#)

**Sample:** Treated-Male-Control ([ERS074997](#)) ([less...](#))

*Organism:* [Mus musculus](#)

*Attributes:*

OrganismPart: liver

Genotype: Foxa1loxP/loxP;Foxa2loxP/loxP

StrainOrLine: 129J/BL6

Organism: Mus musculus

Sex: male

**Library:** MDAR ([more...](#))

**Platform:** Illumina ([less...](#))

*Instrument model:* Illumina Genome Analyzer IIx

**Spot descriptor:**

forward

**Experiment attributes:**

*Experimental Factor:* ORGANISM: Mus musculus

*Experimental Factor:* SEX: male

*Experimental Factor:* TREATMENT: Treated

*Experimental Factor:* GENOTYPE: Foxa1loxP/loxP;Foxa2loxP/loxP

*Experimental Factor:* IMMUNOPRECIPITATE: AR

**Total:** 1 run, 40.9M spots, 1.5G bases, [918.8Mb](#)  

| #  | Run                       | # of Spots | # of Bases | Size                    |
|----|---------------------------|------------|------------|-------------------------|
| 1. | <a href="#">ERR056365</a> | 40,948,224 | 1.5G       | <a href="#">918.8Mb</a> |

ID: 197484

## Run ERR056365 (Foxa1 and Foxa2 are essential for gender dimorphism in liver cancer)

[Metadata](#) [Reads](#) [Download](#)

| Run       | Spots  | Bases   | Size    | GC content | Published  | Access Type |
|-----------|--------|---------|---------|------------|------------|-------------|
| ERR056365 | 40.9 M | 1.5 Gbp | 963.4 M | 43.8%      | 2012-06-26 | public      |

Quality graph ([bigger](#))

This run has 1 read per spot:

L=36, 100%

[Legend](#)

| Experiment                | Library |          |          |         |           |        |
|---------------------------|---------|----------|----------|---------|-----------|--------|
| <a href="#">ERX033410</a> | Name    | Platform | Strategy | Source  | Selection | Layout |
|                           | MDAR    | Illumina | ChIP-Seq | GENOMIC | ChIP      | SINGLE |

[Show design](#)

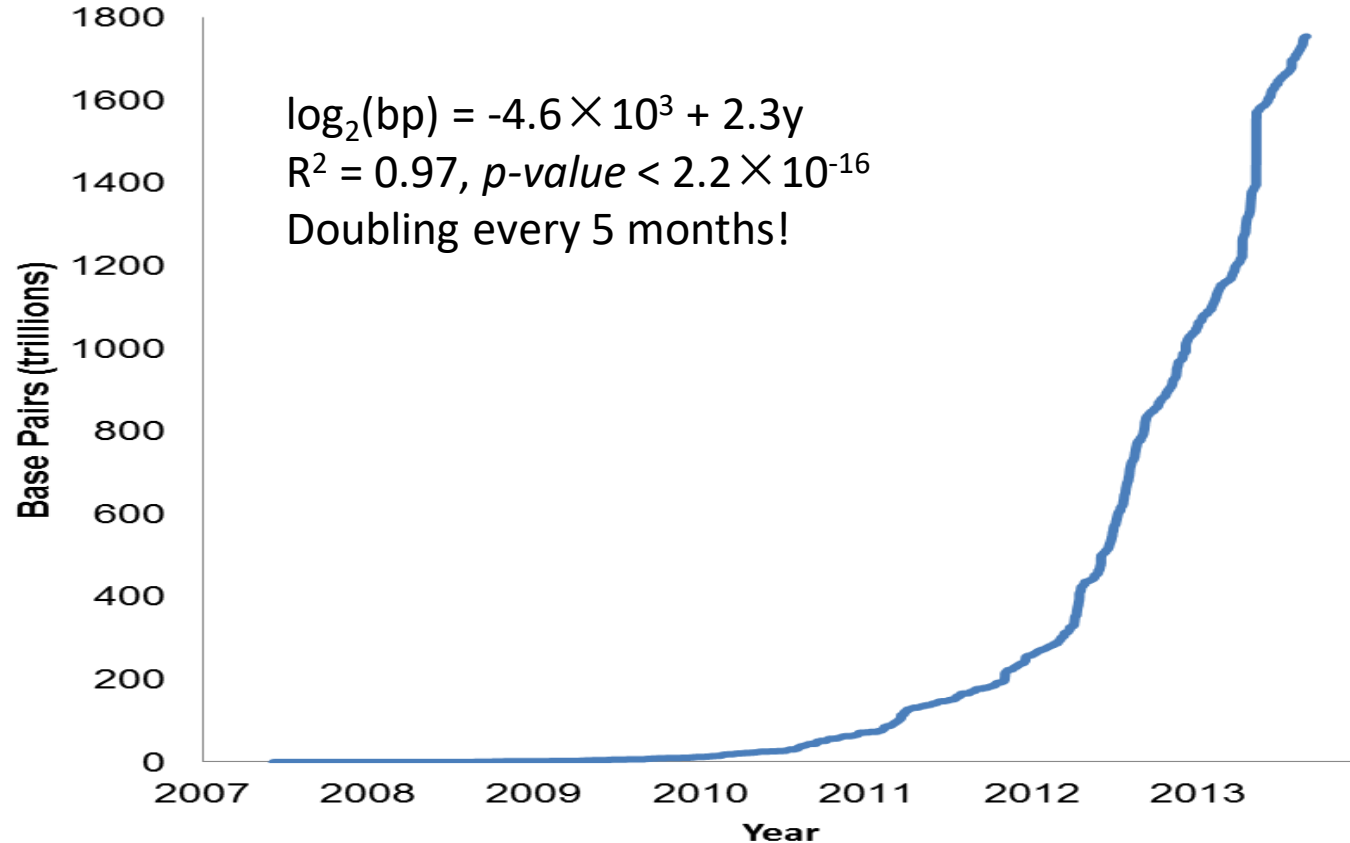
| Biosample                 | Sample name                     | Title                |
|---------------------------|---------------------------------|----------------------|
| <a href="#">ERS074997</a> | E-MTAB-805:Treated-Male-Control | Treated-Male-Control |

| Bioproject | SRA Study                 | Title   |
|------------|---------------------------|---|
|            | <a href="#">ERP001029</a> | Foxa1 and Foxa2 are essential for gender dimorphism in liver cancer |

[Metadata](#) [Reads](#) [Download](#)

| Object     |           | .sra  |
|------------|-----------|---|
| Run        | ERR056365 | 963.4 Mb <a href="#">HTTP</a> <a href="#">FTP</a> |
| Experiment | ERX033410 | 963.4 Mb <a href="#">HTTP</a> <a href="#">FTP</a> |
| Study      | ERP001029 | 12.0 Gb <a href="#">HTTP</a> <a href="#">FTP</a>  |

# SRA Growth



Data source: <http://0-www.ncbi.nlm.nih.gov.elis.tmu.edu.tw/Traces/sra/>



# NCBI-Taxonomy

<http://www.ncbi.nlm.nih.gov/taxonomy/>

NCBI Resources ☒ How To ☒

[Sign in to NCBI](#)

Taxonomy

Taxonomy

Search

[Limits](#) [Advanced](#)

[Help](#)



## Taxonomy

The Taxonomy Database is a curated classification and nomenclature for all of the organisms in the public sequence databases. This currently represents about 10% of the described species of life on the planet.

### Using Taxonomy

[Quick Start Guide](#)

[FAQ](#)

[Handbook](#)

[Taxonomy FTP](#)

### Taxonomy Tools

[Browser](#)

[Common Tree](#)

[Statistics](#)

[Name/ID Status](#)

[Genetic Codes](#)

[Linking to Taxonomy](#)

[Extinct Organisms](#)

### Other Resources

[GenBank](#)

[LinkOut](#)

[E-Utilities](#)

[Batch Entrez](#)

[INSDC](#)

# NCBI-Taxonomy

<http://www.ncbi.nlm.nih.gov/taxonomy/>

NCBI Taxonomy Browser

Entrez PubMed Nucleotide Protein Genome Structure PMC Taxonomy

Search for  as  complete name ☒ lock

Display  3 levels using filter:  none

☐ Nucleotide ☐ Nucleotide EST ☐ Nucleotide GSS ☐ Protein ☐ Structure ☐ Genome ☐ Popset ☐ SNP  
☐ Domains ☐ GEO Datasets ☐ UniGene ☐ UniSTS ☐ PubMed Central ☐ Gene ☐ HomoloGene ☐ SRA  
☐ MapView ☐ LinkOut ☐ BLAST ☐ TRACE ☐ Assembly ☐ Bio Project ☐ Bio Sa  
☐ Bio ☐ dbVar ☐ Epigenomics ☐ GEO Profiles ☐ PubChem BioAssay ☐ Protein Clusters ☐ Host

**Lineage** (full): [root](#); [cellular organisms](#); [Eukaryota](#); [Opisthokonta](#); [Metazoa](#); [Eumetazoa](#); [Bilateria](#); [Deuterostomia](#); [Chordata](#); [Craniata](#); [Vertebrata](#); [Gnathostomata](#); [Teleostomi](#); [Euteleostomi](#); [Sarcopterygii](#); [Dipnotetrapodomorpha](#); [Tetrapoda](#); [Amniota](#); [Mammalia](#); [Theria](#); [Eutheria](#); [Euarchontoglires](#); [Primates](#); [Haplorrhini](#); [Simiiformes](#); [Catarrhini](#); [Hominoidea](#); [Hominidae](#); [Homininae](#); [Homo](#)

◦ [Homo sapiens](#) (human) *Click on organism name to get more information.*

- [Homo sapiens neanderthalensis](#) (Neandertal)
- [Homo sapiens ssp. Denisova](#) (Denisova hominin)

**Disclaimer:** The NCBI taxonomy database is not an authoritative source for nomenclature or classification. Please consult the relevant scientific literature for the most reliable information.

Comments and questions to [info@ncbi.nlm.nih.gov](mailto:info@ncbi.nlm.nih.gov)

## Homo sapiens

Taxonomy ID: 9606

Genbank common name: **human**

Inherited blast name: **primates**

Rank: species

Genetic code: [Translation table 1 \(Standard\)](#)

Mitochondrial genetic code: [Translation table 2 \(Vertebrate Mitochondrial\)](#)

Other names:

common name: **man**

authority: **Homo sapiens** Linnaeus, 1758

**Lineage** (full )

[cellular organisms](#); [Eukaryota](#); [Opisthokonta](#); [Metazoa](#); [Eumetazoa](#); [Bilateria](#); [Deuterostomia](#); [Chordata](#); [Craniata](#); [Vertebrata](#); [Gnathostomata](#); [Teleostomi](#); [Euteleostomi](#); [Sarcopterygii](#); [Dipnotetrapodomorpha](#); [Tetrapoda](#); [Amniota](#); [Mammalia](#); [Theria](#); [Eutheria](#); [Euarchontoglires](#); [Primates](#); [Haplorrhini](#); [Simiiformes](#); [Catarrhini](#); [Hominoidea](#); [Hominidae](#); [Homininae](#); [Homo](#)

- [Archaea](#)
- [Bacteria](#)
- [Eukaryota](#)
- [Viroids](#)
- [Viruses](#)
- [Other](#)
- [Unclassified](#)

These are direct links to some of the organisms commonly used in molecular research projects:

[Arabidopsis thaliana](#)

[Bos taurus](#)

[Caenorhabditis elegans](#)

[Chlamydomonas reinhardtii](#)

[Danio rerio](#) (zebrafish)

[Dictyostelium discoideum](#)

[Drosophila melanogaster](#)

[Escherichia coli](#)

[Hepatitis C virus](#)

[Homo sapiens](#)

[Mus musculus](#)

[Mycoplasma pneumoniae](#)

[Oryza sativa](#)

[Plasmodium falciparum](#)

[Pneumocystis carinii](#)

[Rattus norvegicus](#)

[Saccharomyces cerevisiae](#)

[Schizosaccharomyces pombe](#)

[Takifugu rubripes](#)

[Xenopus laevis](#)

[Zea mays](#)

| Entrez records  |               |              |
|-----------------|---------------|--------------|
| Database name   | Subtree links | Direct links |
| Nucleotide      | 10,401,809    | 10,401,779   |
| Nucleotide EST  | 8,704,844     | 8,704,844    |
| Nucleotide GSS  | 1,729,196     | 1,727,870    |
| Protein         | 800,192       | 800,044      |
| Structure       | 23,839        | 23,839       |
| Genome          | 1             | 1            |
| Popset          | 15,254        | 15,254       |
| SNP             | 73,357,341    | 73,357,341   |
| Domains         | 11            | 11           |
| GEO Datasets    | 598,345       | 598,345      |
| UniGene         | 130,056       | 130,056      |
| UniSTS          | 328,873       | 328,873      |
| PubMed Central  | 15,467        | 15,448       |
| Gene            | 199,013       | 198,940      |
| HomoloGene      | 18,473        | 18,473       |
| SRA Experiments | 114,540       | 114,510      |
| Probe           | 27,841,436    | 27,841,435   |

This is the top level of the taxonomy database maintained by NCBI/GenBank. You can explore any of the taxa listed below by clicking it.

[Help]

[Search]

[NLM NIH]

[Disclaimer]

# NCBI-Taxonomy Statistics

<http://www.ncbi.nlm.nih.gov/taxonomy/>

## Taxonomy Nodes (all dates)

| Ranks:                        | <a href="#">higher taxa</a> | <a href="#">genus</a> | <a href="#">species</a> | <a href="#">lower taxa</a> | <a href="#">total</a>  |
|-------------------------------|-----------------------------|-----------------------|-------------------------|----------------------------|------------------------|
| <a href="#">Archaea</a>       | <a href="#">140</a>         | <a href="#">134</a>   | <a href="#">490</a>     | <a href="#">0</a>          | <a href="#">764</a>    |
| <a href="#">Bacteria</a>      | <a href="#">1327</a>        | <a href="#">2412</a>  | <a href="#">12002</a>   | <a href="#">783</a>        | <a href="#">16524</a>  |
| <a href="#">Eukaryota</a>     | <a href="#">19704</a>       | <a href="#">64657</a> | <a href="#">277209</a>  | <a href="#">20892</a>      | <a href="#">382462</a> |
| <a href="#">Fungi</a>         | <a href="#">1454</a>        | <a href="#">4356</a>  | <a href="#">27553</a>   | <a href="#">1043</a>       | <a href="#">34406</a>  |
| <a href="#">Metazoa</a>       | <a href="#">14278</a>       | <a href="#">43157</a> | <a href="#">133817</a>  | <a href="#">10481</a>      | <a href="#">201733</a> |
| <a href="#">Viridiplantae</a> | <a href="#">2430</a>        | <a href="#">14456</a> | <a href="#">107121</a>  | <a href="#">9126</a>       | <a href="#">133133</a> |
| <a href="#">Viruses</a>       | <a href="#">588</a>         | <a href="#">411</a>   | <a href="#">2015</a>    | <a href="#">0</a>          | <a href="#">3014</a>   |
| <a href="#">All taxa</a>      | <a href="#">21786</a>       | <a href="#">67621</a> | <a href="#">291749</a>  | <a href="#">21675</a>      | <a href="#">402831</a> |

Dates: [2004](#) [2005](#) [2006](#) [2007](#) [2008](#) [2009](#) [2010](#) [2011](#) [2012](#) [2013](#) [all dates](#)

# NCBI-PubMed

<http://www.ncbi.nlm.nih.gov/pubmed/>

NCBI Resources How To

[Sign in to NCBI](#)

PubMed.gov

PubMed



Advanced

Search

US National Library of  
Medicine National  
Institutes of Health

[Help](#)

## PubMed

PubMed comprises more than 23 million citations for biomedical literature from MEDLINE, life science journals, and online books. Citations may include links to full-text content from PubMed Central and publisher web sites.

## PubMed Commons

[click for overview page](#)

### Using PubMed

[PubMed Quick Start Guide](#)

[Full Text Articles](#)

[PubMed FAQs](#)

[PubMed Tutorials](#)

[New and Noteworthy](#)



### PubMed Tools

[PubMed Mobile](#)

[Single Citation Matcher](#)

[Batch Citation Matcher](#)

[Clinical Queries](#)

[Topic-Specific Queries](#)

### More Resources

[MeSH Database](#)

[Journals in NCBI Databases](#)

[Clinical Trials](#)

[E-Utilities](#)

[LinkOut](#)

Display Settings: Abstract

Nucleic Acids Res. 2011 Jul;39(Web Server issue):W316-22. doi: 10.1093/nar/gkr483.

## KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases.

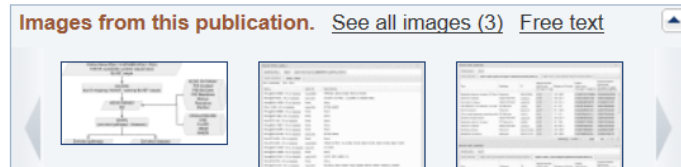
Xie C, Mao X, Huang J, Ding Y, Wu J, Dong S, Kong L, Gao G, Li CY, Wei L.

Center for Bioinformatics, State Key Laboratory of Protein and Plant Gene Research, College of Life Sciences, Peking University, Beijing, China.

### Abstract

High-throughput experimental technologies often identify dozens to hundreds of genes related to, or changed in, a biological or pathological process. From these genes one wants to identify biological pathways that may be involved and diseases that may be implicated. Here, we report a web server, KOBAS 2.0, which annotates an input set of genes with putative pathways and disease relationships based on mapping to genes with known annotations. It allows for both ID mapping and cross-species sequence similarity mapping. It then performs statistical tests to identify statistically significantly enriched pathways and diseases. KOBAS 2.0 incorporates knowledge across 1327 species from 5 pathway databases (KEGG PATHWAY, PID, BioCyc, Reactome and Panther) and 5 human disease databases (OMIM, KEGG DISEASE, FunDO, GAD and NHGRI GWAS Catalog). KOBAS 2.0 can be accessed at <http://kobas.cbi.pku.edu.cn>.

PMID: 21715386 [PubMed - indexed for MEDLINE] PMCID: PMC3125809 Free PMC Article



Send to:

Get full text



Nucleic Acids Res  
Free full text



Nucleic Acids Res  
Free full text

Save items

Add to Favorites

Related citations in PubMed

KOBAS server: a web-based platform for automated annotation ; [Nucleic Acids Res. 2006]

IPAD: the Integrated Pathway Analysis Database for Systematic Enrich [BMC Bioinformatics. 2012]

GBA server: EST-based digital gene expression profiling. [Nucleic Acids Res. 2005]

MyBioNet: interactively visualize, edit and merge biological networks on the ; [Bioinformatics. 2011]



NCBI

Resources

How To

[Sign in to NCBI](#)



US National Library of Medicine  
National Institutes of Health

PMC



Search

[Journal List](#)

[Limits](#)

[Advanced](#)

[Help](#)



## PMC

PMC is a free full-text archive of biomedical and life sciences journal literature at the U.S. National Institutes of Health's National Library of Medicine (NIH/NLM).

### PubReader

A whole new way  
to read scientific  
literature at  
PubMed Central



### Get Started

[PMC Overview](#)

[Users' Guide](#)

[Journal List](#)

[PMC FAQs](#)

[PMC Copyright Notice](#)

### Participate

[Add a Journal to PMC](#)

[Participation Agreements](#)

[File Submission Specifications](#)

[File Validation Tools](#)

### Keep Up to Date

[New in PMC](#)



[PMC Announce Mail List](#)

[Utilities Announce Mail List](#)

[Tagging Guidelines Mail List](#)

### Other Resources

[PMC International](#)

[Open Access Subject](#)

**2.9 MILLION Articles**

*are archived in PMC.*

### NIH Public Access

[NIH Public Access and PMC](#)

[NIH Manuscript Submission System](#)

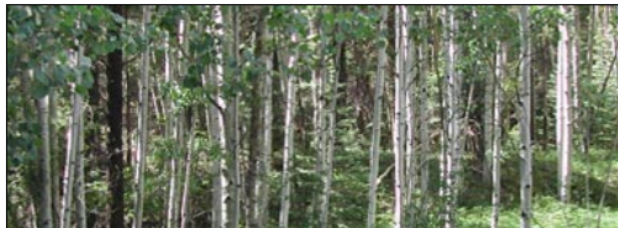
MeSH

MeSH

Search

[Limits](#) [Advanced](#)

[Help](#)



## MeSH

MeSH (Medical Subject Headings) is the NLM controlled vocabulary thesaurus used for indexing articles for PubMed.

### Using MeSH

[Help](#)

[Tutorials](#)

### More Resources

[E-Utilities](#)

[NLM MeSH Homepage](#)

You are here: [NCBI](#) > [Literature](#) > [MeSH Database](#)

[Write to the Help Desk](#)

#### GETTING STARTED

[NCBI Education](#)

[NCBI Help Manual](#)

#### RESOURCES

[Chemicals & Bioassays](#)

[Data & Software](#)

#### POPULAR

[PubMed](#)

[Bookshelf](#)

#### FEATURED

[Genetic Testing Registry](#)

[PubMed Health](#)

#### NCBI INFORMATION

















[About NCBI](#)

[Research at NCBI](#)











## MeSH Tree Structures - 2014

[Return to Entry Page](#)

1.  Anatomy [A]
2.  Organisms [B]
3.  Diseases [C]
4.  Chemicals and Drugs [D]
5.  Analytical, Diagnostic and Therapeutic Techniques and Equipment [E]
6.  Psychiatry and Psychology [F]
7.  Phenomena and Processes [G]
8.  Disciplines and Occupations [H]
9.  Anthropology, Education, Sociology and Social Phenomena [I]
10.  Technology, Industry, Agriculture [J]
11.  Humanities [K]
12.  Information Science [L]
13.  Named Groups [M]
14.  Health Care [N]
15.  Publication Characteristics [V]
16.  Geographicals [Z]

# MeSH Tree Structures - 2014

[Return to Entry Page](#)

1.  Anatomy [A]
2.  Organisms [B]
3.  Diseases [C]
4.  Chemicals and Drugs [D]
5.  Analytical, Diagnostic and Therapeutic Techniques and Equipment [E]
6.  Psychiatry and Psychology [F]
7.  Phenomena and Processes [G]
  - [Physical Phenomena \[G01\]](#) +
  - [Chemical Phenomena \[G02\]](#) +
  - [Metabolic Phenomena \[G03\]](#) +
  - [Cell Physiological Phenomena \[G04\]](#) +
  - [Genetic Phenomena \[G05\]](#) +
  - [Microbiological Phenomena \[G06\]](#) +
  - [Physiological Phenomena \[G07\]](#) +
  - [Reproductive and Urinary Physiological Phenomena \[G08\]](#) +
  - [Circulatory and Respiratory Physiological Phenomena \[G09\]](#) +
  - [Digestive System and Oral Physiological Phenomena \[G10\]](#) +
  - [Musculoskeletal and Neural Physiological Phenomena \[G11\]](#) +
  - [Immune System Phenomena \[G12\]](#) +
  - [Integumentary System Physiological Phenomena \[G13\]](#) +
  - [Ocular Physiological Phenomena \[G14\]](#) +
  - [Plant Physiological Phenomena \[G15\]](#) +
  - [Biological Phenomena \[G16\]](#) +
  - [Mathematical Concepts \[G17\]](#) +
8.  Disciplines and Occupations [H]

## Genetic Phenomena [G05]

### Genetic Variation [G05.365]

Antibody Diversity [G05.365.036]

Antigenic Variation [G05.365.073]

Genetic Heterogeneity [G05.365.331]

#### ► Mutation [G05.365.590]

Allelic Imbalance [G05.365.590.029] +

Base Pair Mismatch [G05.365.590.060]

Chromosome Aberrations [G05.365.590.175] +

Codon, Nonsense [G05.365.590.195]

DNA Repeat Expansion [G05.365.590.220] +

Frameshift Mutation [G05.365.590.265]

Gene Amplification [G05.365.590.310]

Gene Duplication [G05.365.590.320]

Genomic Instability [G05.365.590.335] +

Germ-Line Mutation [G05.365.590.350]

INDEL Mutation [G05.365.590.500]

Mutagenesis, Insertional [G05.365.590.575]

Mutation Rate [G05.365.590.612]

Mutation, Missense [G05.365.590.650]

Point Mutation [G05.365.590.675]

Sequence Deletion [G05.365.590.762] +

Sequence Inversion [G05.365.590.770] +

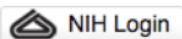
Suppression, Genetic [G05.365.590.835]

Polymorphism, Genetic [G05.365.795] +

# NCBI-My NCBI

## Sign in to NCBI

### Sign in with



*eRA Commons users should sign in via the NIH Login*

[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 retains user information and database preferences to provide customized services for many NCBI databases.

 [My NCBI Overview](#)

My NCBI features include:

- Save searches & automatic e-mail alerts
- Display format preferences
- Filter options
- My Bibliography & NIH public access policy compliance
- Highlighting search terms
- Recent activity searches & records for 6 months
- LinkOut, document delivery service & outside tool selections

### NIH funded investigator?


Extramural NIH-funded investigators looking for NIH Public Access Compliance tools should sign in using the "NIH Login" button. Use your eRA Commons credentials on the subsequent sign in page. Once signed in, navigate to the My Bibliography section.

Documentation for using these features is located in the [Managing Compliance to the NIH Public Access Policy](#) section of the NCBI Help Manual.

Information about the NIH Public Access Policy is located at  
<http://publicaccess.nih.gov>

# NCBI-BLAST

<http://blast.ncbi.nlm.nih.gov/Blast.cgi>

 **BLAST®**  
Basic Local Alignment Search Tool

[Home](#) [Recent Results](#) [Saved Strategies](#) [Help](#)

[My NCBI](#)  
[\[Sign In\]](#) [\[Register\]](#)

► [NCBI/BLAST Home](#)

BLAST finds regions of similarity between biological sequences. [more...](#)

**New** DELTA-BLAST, a more sensitive protein-protein search

[Go](#)

## BLAST Assembled RefSeq Genomes

Choose a species genome to search, or [list all genomic BLAST databases](#).

- |   |  |  |
|---|--|--|
| <input type="checkbox"/> <a href="#">Human</a>                | <input type="checkbox"/> <a href="#">Oryza sativa</a>            | <input type="checkbox"/> <a href="#">Gallus gallus</a>   |
| <input type="checkbox"/> <a href="#">Mouse</a>                | <input type="checkbox"/> <a href="#">Bos taurus</a>              | <input type="checkbox"/> <a href="#">Pan troglodytes</a> |
| <input type="checkbox"/> <a href="#">Rat</a>                  | <input type="checkbox"/> <a href="#">Danio rerio</a>             | <input type="checkbox"/> <a href="#">Microbes</a>        |
| <input type="checkbox"/> <a href="#">Arabidopsis thaliana</a> | <input type="checkbox"/> <a href="#">Drosophila melanogaster</a> | <input type="checkbox"/> <a href="#">Apis mellifera</a>  |

## Basic BLAST

Choose a BLAST program to run.

|                                  |  |
|----------------------------------|--|
| <a href="#">nucleotide blast</a> | Search a <b>nucleotide</b> database using a <b>nucleotide</b> query<br><i>Algorithms:</i> blastn, megablast, discontinuous megablast |
| <a href="#">protein blast</a>    | Search <b>protein</b> database using a <b>protein</b> query<br><i>Algorithms:</i> blastp, psi-blast, phi-blast, delta-blast          |
| <a href="#">blastx</a>           | Search <b>protein</b> database using a <b>translated nucleotide</b> query  |
| <a href="#">tblastn</a>          | Search <b>translated nucleotide</b> database using a <b>protein</b> query  |
| <a href="#">tblastx</a>          | Search <b>translated nucleotide</b> database using a <b>translated nucleotide</b> query  |

## Specialized BLAST

Your Recent Results **New**

[All Recent results...](#)

## News

### [Update to organism BLAST databases](#)

The organism BLAST pages are being updated to use top-level (chromosome + unplaced and unlocalized scaffolds) RefSeq genomic records instead of scaffold records.

Thu, 17 Oct 2013 14:00:00 EST

[More BLAST news...](#)

## Tip of the Day

### [Use Genomic BLAST to see the genomic context](#)

If you are interested in the evolution of a particular gene or gene family it is often interesting to examine the intro-exon structure even across species.

[More tips](#)

# NCBI-BLAST

<http://blast.ncbi.nlm.nih.gov/Blast.cgi>

NCBI/BLAST/blastn suite

Standard Nucleotide BLAST

blastn blastx tblastn tblastx

Enter Query Sequence

BLASTN programs search nucleotide databases using a nucleotide query. [more...](#) [Reset page](#) [Bookmark](#)

Enter accession number(s), gi(s), or FASTA sequence(s) [Clear](#) Query subrange [From](#)

Or, upload file  未选择

Job Title

Enter a descriptive [Align two or more sequences](#)

Choose Search Set

Database ☐ Human genomic ☒ Nucleotide collection

Organism  Optional  Enter an organism or taxon

Exclude ☐ Models (XMM/XP)

Entrez Query  Optional  Enter an Entrez query

Program Selection

Optimize for ☒ Highly similar ☐ More dissimilar ☐ Somewhat similar

RID [9FUE30VP01R](#) (Expires on 11-30 08:11 am)

Query ID [Id|10604](#)

Description [None](#)

Molecule type [amino acid](#)

Query Length [180](#)

Database Name [nr](#)

Description [All non-redundant GenBank CDS translations+PDB+SwissProt+PIR+PRF excluding environmental samples from WGS projects](#)

Program [BLASTP 2.2.28+](#) [Citation](#)

[Alignments](#)

Other reports: [Search Summary](#) [Taxonomy reports](#) [Distance tree of results](#) [Multiple alignment](#)

[Graphic Summary](#)

[Show Conserved Domains](#)

Putative conserved domains have been detected, click on the image below for detailed results

Query seq.  [transporter signature motif](#)  [ATP binding site](#)  [D-loop](#)  [Walker B](#)  [M-loop/switch region](#)

Superfamilies  [ABC\\_ATPase superfamily](#)

Multi-domains

Distribution of 173 Blast Hits on the Query Sequence

Mouse over to see the define, click to show alignments

Color key for alignment scores

Query [1](#) [30](#) [60](#) [90](#) [120](#) [150](#) [180](#)

[Download](#) [GenPept](#) [Graphics](#) Sort by: [E value](#)

PREDICTED: ATP-binding cassette sub-family A member 3 isoform X2 [Homo sapiens]

Sequence ID: [ref|XP\\_005255226.1|](#) Length: 1275 Number of Matches: 2

Range 1: 232 to 411 [GenPept](#) [Graphics](#) [Next Match](#) [Previous Match](#)

| Score         | Expect | Method                       | Identities    | Positives     | Gaps      |
|---------------|--------|------------------------------|---------------|---------------|-----------|
| 370 bits(950) | 5e-119 | Compositional matrix adjust. | 180/180(100%) | 180/180(100%) | 0/180(0%) |

Query 1 [RSFSLGGMRRLSLGIALIAGSKVILDEFTSGMDAISRAIWDLLARQKSDRTVLIT](#) 60

Subject 232 [RSFSLGGMRRLSLGIALIAGSKVILDEFTSGMDAISRAIWDLLARQKSDRTVLIT](#) 291

Query 61 [HFMDEADLLGRRIATMAKGEQLCCGSSFLKQRTGAGTHMTLVKEPHCNPEISQLVGGH](#) 120

Subject 292 [HFMDEADLLGRRIATMAKGEQLCCGSSFLKQRTGAGTHMTLVKEPHCNPEISQLVGGH](#) 351

Query 121 [VFNATLESSAGAEISFILPRESTRHRFGLFAKLEKQKQELGIASFGASITIMEEVFLRW](#) 180

Subject 352 [VFNATLESSAGAEISFILPRESTRHRFGLFAKLEKQKQELGIASFGASITIMEEVFLRW](#) 411

Range 2: 1084 to 1261 [GenPept](#) [Graphics](#) [Next Match](#) [Previous Match](#) [First Match](#)

| Score         | Expect | Method                       | Identities  | Positives    | Gaps      |
|---------------|--------|------------------------------|-------------|--------------|-----------|
| 116 bits(291) | 1e-28  | Compositional matrix adjust. | 64/179(36%) | 103/179(57%) | 5/179(2%) |

Query 3 [RFLSGGMRRLSLGIALIAGSKVILDEFTSGMDAISRAIWDLLARQKSDRTVLIT](#) 61

Subject 1084 [R\\_SGG+RRLS\\_GIALI\\_V+LDEP+GMD++R+WD+R++S+I+I+H](#) 1143

Query 62 [HFMDEADLLGRRIATMAKGEQLCCGSSFLKQRTGAGTHMTLVKEPHCNPEISQLVGGH](#) 118

Subject 1144 [SMEECEALCTRLATMVGQKFLGSPQHLKSKFSGTSLRAKVSQSQQEALEKFAFVD](#) 1203

Query 119 [HFMVFNATLESSAGAEISFILPRESTRHRFGLFAKLEKQKQELGIASFGASITIMEEVFLRW](#) 177

Subject 1204 [LTFPGSVLEDRGQWVHYHLPGRDL-SNAKVFGLKAKKRYGVDDTSYSISLQVFL](#) 1261

[Download](#) [GenPept](#) [Graphics](#) Sort by: [E value](#)

ABCA3 variant protein [Homo sapiens]

# NCBI-BLAST

<http://blast.ncbi.nlm.nih.gov/Blast.cgi>

- Online: NCBI-BLAST([blast.ncbi.nlm.nih.gov/Blast.cgi](http://blast.ncbi.nlm.nih.gov/Blast.cgi))
- Standalone: BLAST+
- Embedded in webpage: wwwblast



# 生物信息学：导论与方法

## Bioinformatics: Introduction and Methods

Ge Gao 高歌 & Liping Wei 魏丽萍

Center for Bioinformatics, Peking University



<https://www.coursera.org/course/pkubioinfo>