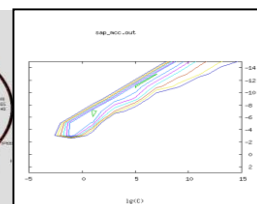
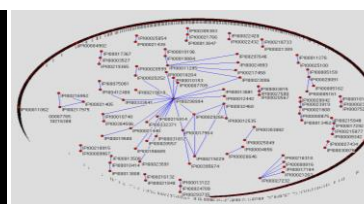
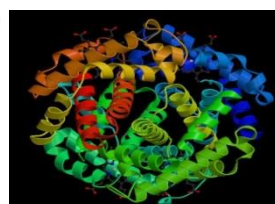
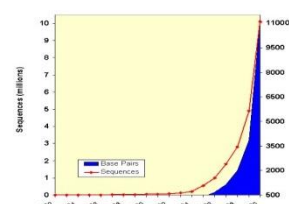
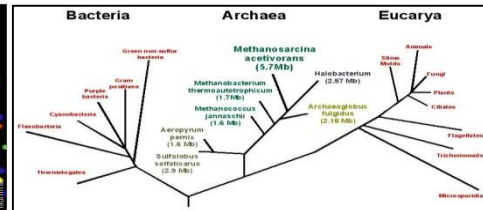


北京大学生物信息学中心 魏丽萍

Liping Wei, Ph.D.

Center for Bioinformatics, Peking University



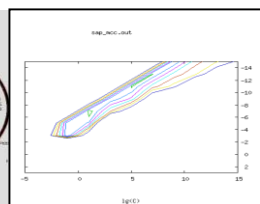


Unit 3: European Bioinformatics Institute

北京大学生物信息学中心 魏丽萍

Liping Wei, Ph.D.

Center for Bioinformatics, Peking University



EBI (<http://www.ebi.ac.uk/services>)

EMBL-EBI

Services

Research

Training


About us


Services


OverviewA to ZService teamsSupport


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.


 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


 Literature


 Other software


Popular


 Ensembl


 UniProt


 PDB


 ArrayExpress


 ChEMBL

 BLAST

 Europe PMC

 Reactome

 Train online

 Support

Bioinformatics training

Guide to resources

Service news

DNA, RNA, and proteins

Genes and Genomes

Databases	
Comprehensive annotation	Ensembl
Sequence archive	ENA, EGA, EBI Metagenomics
Large-scale sequencing project	1000 Genomes
Tools	
Sequence alignment	BLAST, Clustal Omega

Expression

Databases	
Expression profile archive	ArrayExpress, Expression Atlas

Proteins

Databases	
Protein annotation	UniProt, InterPro, Pfam
Proteomics data archive	PRIDE
Protein structure	PDBe, PDBsum, Quaternary structure
Tools	
Protein function prediction	InterProScan
Protein structure prediction	PDBePISA
Protein function prediction	ProFunc
Protein structure alignment	PDBeFold

Reactions, pathways, ontologies

Interactions, pathways & networks

Databases	
Molecular interaction	IntAct
Pathway database	Reactome
Network dynamics	BioModels

Small molecules

Databases	
Annotations of chemicals	ChEBI, ChEMBL
Metabolomics archive	Metabolights


Enzymes and reactions

Databases	
Enzyme annotation	EBI Enzyme Portal, Catalytic Site Atlas

Ontologies

Databases	
Ontologies	Gene Ontology, Systems Biology Ontology, Experimental Factor Ontology
Tools	
Search engine for ontologies	Ontology Lookup Service, QuickGO

Ensembl (http://ensembl.org)

 **BLAST/BLAT** | **BioMart** | **Tools** | **Downloads** | **Help & Documentation** | **Blog** | **Mirrors**

Login/Register


Search all species...


Search: for
e.g. **BRCA2** or **rat X:100000..200000** or **coronary heart disease**


Browse a Genome

The Ensembl project produces genome databases for vertebrates and other eukaryotic species, and makes this information freely available online.

Popular genomes

 **Human**
GRCh37

 **Mouse**
GRCm38

 **Zebrafish**
Zv9

★ [Log in to customize this list](#)


All genomes

-- Select a species --


[View full list of all Ensembl species](#)

Other species are available in [Ensembl Pre!](#) and [EnsemblGenomes](#)

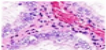
ENCODE data in Ensembl



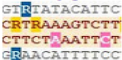
Variant Effect Predictor



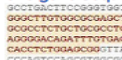
Gene expression in different tissues




Find SNPs and other variants for my gene



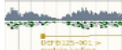
Retrieve gene sequence




Compare genes across species



Use my own data in Ensembl



Learn about a disease or phenotype



What's New in Release 73 (September 2013)

- [New search engine using Solr](#)
- [New species: Duck \(*Anas platyrhynchos*\) and Flycatcher \(*Ficedula albicollis*\)](#)
- [Updated patches for the human assembly \(GRCh37.p12\)](#)


[Full details of this release](#)
[More release news on our blog →](#)

Latest blog posts


- 19 Nov 2013: [Ensembl Reorganises](#)
- 01 Nov 2013: [Upgrade of public MySQL servers](#)
- 01 Nov 2013: [Retirement of archive 60](#)


[Go to Ensembl blog →](#)

Did you know...?



If you're using the [VEP script](#), please use [the cache](#). Our Variant Effect Predictor tool also has a [web interface](#).

 Ensembl is a joint project between [EMBL - EBI](#) and the [Wellcome Trust Sanger Institute](#) to develop a software system which produces and maintains automatic annotation on selected eukaryotic genomes. Ensembl receives major funding from the Wellcome Trust. Our [acknowledgements page](#) includes a list of additional current and previous funding bodies.

EMBL-EBI 

Ensembl release 73 - September 2013 © [WTSI](#) / [EBI](#)

[About Ensembl](#) | [Privacy Policy](#) | [Contact Us](#)

[Permanent link](#) - [View in archive site](#)

Screenshot made at http://ensembl.org on November 26th, 2013 (UTC+0800)

Lots of Species in Ensembl



Aardvark ([preview - assembly only](#))
Orycteropus afer afer
OryAfe1



Alpaca
Vicugna pacos
vicPac1



Anole lizard
Anolis carolinensis
AnoCar2.0



Armadillo ([preview new assembly Dasnov3.0](#))
Dasypus novemcinctus
dasNov2



Baboon ([preview - assembly only](#))
Papio hamadryas
Pham



Budgerigar ([preview - assembly only](#))
Melopsittacus undulatus
MelUnd6.3



Bushbaby
Otolemur garnettii
OtoGar3



Ciona intestinalis
Ciona intestinalis
KH



Ciona savignyi
Ciona savignyi
CSAV2.0



Caenorhabditis elegans
Caenorhabditis elegans
WBcel235



Cat
Felis catus
Felis_catus_6.2



Cave fish ([preview - assembly only](#))
Astyanax mexicanus
AstMex102



Gibbon
Nomascus leucogenys
Nleu1.0



Gorilla
Gorilla gorilla gorilla
gorGor3.1



Guinea Pig
Cavia porcellus
cavPor3



Hedgehog
Erinaceus europaeus
HEDGEHOG



Horse
Equus caballus
EquCab2



Human
Homo sapiens
GRCh37



Hyrax
Procavia capensis
proCap1



Kangaroo rat
Dipodomys ordii
dipOrd1



Lamprey
Petromyzon marinus
Pmarinus_7.0



Lesser hedgehog tenrec
Echinops telfairi
TENREC



Macaque
Macaca mulatta
MMUL_1



Marmoset
Callithrix jacchus
C_jacchus3.2.1



Platypus
Ornithorhynchus anatinus
OANA5



Prairie vole ([preview - assembly only](#))
Microtus ochrogaster
MicOch1.0



Rabbit
Oryctolagus cuniculus
OryCun2.0



Rat
Rattus norvegicus
Rnor_5.0



Rhinoceros ([preview - assembly only](#))
Ceratotherium simum simum
CerSimSim1



Saccharomyces cerevisiae
Saccharomyces cerevisiae
EF4



Sheep ([preview - assembly only](#))
Ovis aries
Oar_v3.1



Shrew ([preview new assembly SorAra2.0](#))
Sorex araneus
COMMON_SHREW1



Sloth
Choloepus hoffmanni
choHof1



Spotted Gar ([preview - assembly only](#))
Lepisosteus oculatus
LepOcu1



Squirrel
Ictidomys tridecemlineatus
spetri2



Squirrel monkey ([preview - assembly only](#))
Saimiri boliviensis
SalBol1.0

Ensembl Statistics for *Homo sapiens*

Summary	
Assembly	GRCh37.p12
Database version	73.37
Base Pairs	3,324,592,091
Golden Path Length	3,101,804,739

Gene Counts (Primary assembly)	
Coding genes	20,769
Short Non coding genes	9,079
Long Non coding genes	13,564
Pseudogenes	14,165
Gene transcripts	195,565
Short Variants (SNPs, indels, somatic mutations)	55,286,608
Structural variants	10,343,072

http://asia.ensembl.org/Homo_sapiens/Info/Annotation

Four Types of Data in Ensembl

- **Curated/Reference:** Data that are manually curated based on experimental results and validations. They serve as the “reference” for future studies.
- **Large-scale projects:** Data that are not curated but are produced by high-throughput experiments. They tend to have a large sample size or data volume.
- **Other studies:** Low-throughput experimental data that are not curated, possibly with some computational postprocessing.
- **Only computationally analyzed:** Secondary data that are generated by running automated algorithms over curated/reference data and large-scale high-throughput data (e.g. the genome sequence).

UniProtKB (<http://www.uniprot.org/>)

UniProt > UniProtKB

Downloads · Contact · Documentation/Help

Search Blast Align Retrieve ID Mapping *

Search in

1 - 25 of 573 results for cxcl in UniProtKB sorted by score descending

Browse by [taxonomy](#), [keyword](#), [gene ontology](#), [enzyme class](#) or [pathway](#) | Reduce sequence redundancy to [100%](#), [90%](#) or [50%](#)

Page 1 of 23 | [Next](#)

Results [Customize](#)

> Show only [reviewed \(185\)](#) (UniProtKB/Swiss-Prot) or [unreviewed \(388\)](#) (UniProtKB/TrEMBL) entries

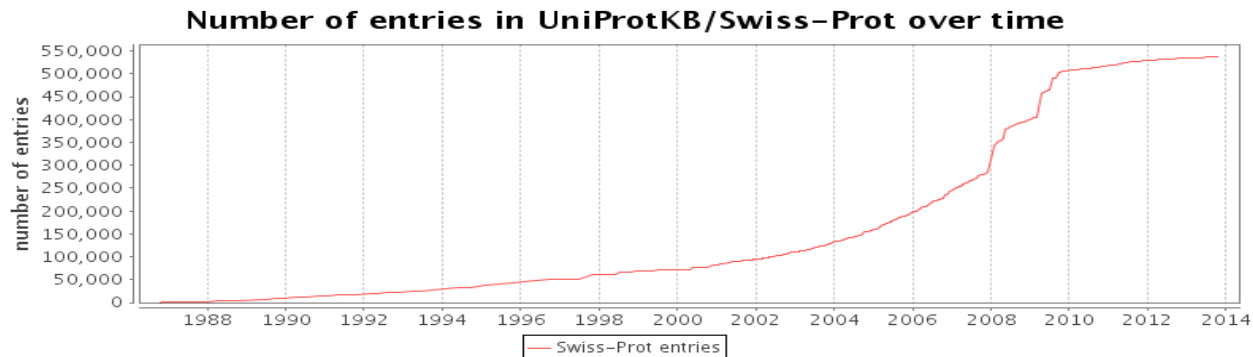
> Restrict term "cxcl" to [gene name \(358\)](#), [protein name \(70\)](#), [strain \(1\)](#), [taxonomy \(1\)](#), [web resource \(14\)](#)

Entry	Entry name	Status	Protein names	Gene names	Organism	Length
<input type="checkbox"/> Q6AXU5	CXL16_RAT	★	C-X-C motif chemokine 16	Cxcl16	Rattus norvegicus (Rat)	247
<input type="checkbox"/> Q8BSU2	CXL16_MOUSE	★	C-X-C motif chemokine 16	Cxcl16 Srpsox	Mus musculus (Mouse)	246
<input type="checkbox"/> Q29RT9	CXL16_BOVIN	★	C-X-C motif chemokine 16	CXCL16	Bos taurus (Bovine)	252
<input type="checkbox"/> Q9GKE2	CXL16_PIG	★	C-X-C motif chemokine 16	CXCL16	Sus scrofa (Pig)	250
<input type="checkbox"/> P02778	CXL10_HUMAN	★	C-X-C motif chemokine 10	CXCL10 INP10 SCYB10	Homo sapiens (Human)	98
<input type="checkbox"/> Q9H2A7	CXL16_HUMAN	★	C-X-C motif chemokine 16	CXCL16 SCYB16 SRPSOX UNQ2759/PRO6714	Homo sapiens (Human)	254
<input type="checkbox"/> P10720	PF4V_HUMAN	★	Platelet factor 4 variant	PF4V1 CXCL4V1 SCYB4V1	Homo sapiens (Human)	104
<input type="checkbox"/> Q5U5W9	Q5U5W9_MOUSE	★	Cxcl1 protein	Cxcl1	Mus musculus (Mouse)	107
<input type="checkbox"/> Q8R392	Q8R392_MOUSE	★	Cxcl11 protein	Cxcl11	Mus musculus (Mouse)	100
<input type="checkbox"/> Q6AXC2	Q6AXC2_MOUSE	★	Cxcl14 protein	Cxcl14	Mus musculus (Mouse)	61
<input type="checkbox"/> Q0P5I9	Q0P5I9_BOVIN	★	CXCL12 protein	CXCL12	Bos taurus (Bovine)	89
<input type="checkbox"/> Q8K453	Q8K453_RAT	★	BRAK	Cxcl14	Rattus norvegicus (Rat)	99
<input type="checkbox"/> B0ZYQ7	B0ZYQ7_RABIT	★	CXCL13	CXCL13	Oryctolagus cuniculus (Rabbit)	86
<input type="checkbox"/> Q6PUJ1	Q6PUJ1_PIG	★	CXCL2	CXCL2	Sus scrofa (Pig)	107
<input type="checkbox"/> Q5I0J6	Q5I0J6_RAT	★	Chemokine (C-X-C motif) ligand 13	Cxcl13 LOC498335	Rattus norvegicus (Rat)	108
<input type="checkbox"/> A9ZPF6	A9ZPF6_DANRE	★	Chemokine CXCL-C5c	cxcl-c5c	Danio rerio (Zebrafish) (Brachydanio rerio)	147
<input type="checkbox"/> Q8K4B1	Q8K4B1_RAT	★	Chemokine (C-X-C motif) ligand 9	Cxcl9 Mig rCG.60435	Rattus norvegicus (Rat)	125
<input type="checkbox"/> Q80YV8	Q80YV8_RAT	★	Protein Cxcl12	Cxcl12	Rattus norvegicus (Rat)	119
<input type="checkbox"/> Q8AV10	Q8AV10_DANRE	★	Chemokine ligand 12	cxcl12a	Danio rerio (Zebrafish) (Brachydanio rerio)	99
<input type="checkbox"/> P50228	CXCL5_MOUSE	★	C-X-C motif chemokine 5	Cxcl5 Scyb5	Mus musculus (Mouse)	132
<input type="checkbox"/> Q5G9I8	Q5G9I8_TUPRE	★	CXCL2	CXCL2	Tupaia belangeri (Common tree shrew) (Tupaia glis)	99

Screenshot made at <http://www.uniprot.org/uniprot/?query=cxcl&sort=score>
on November 26th, 2013 (UTC+0800)

Statistics (UniProtKB/Swiss-Prot)

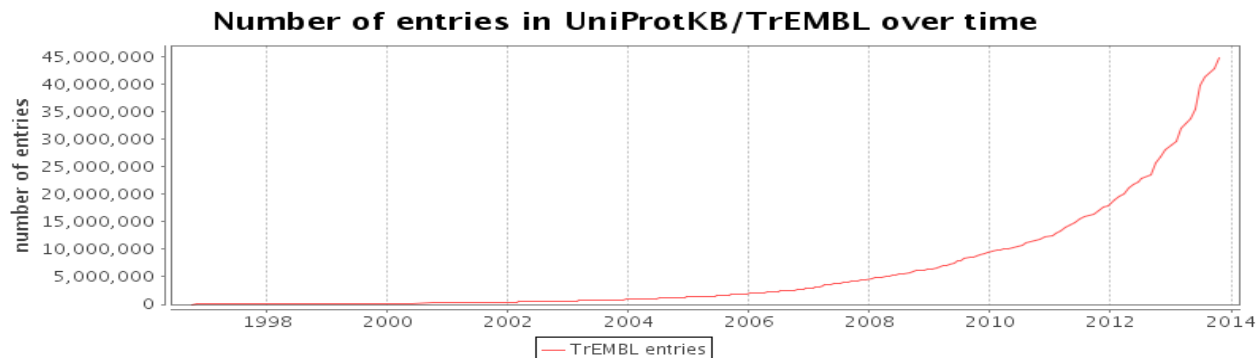
New entries	212
Updated entries	420,451
Unchanged entries	121,099
Total	541,762
Entries with updated sequences	24
With a fragmented AA sequence	9,083
With known alternative products	22,463



Data and image collected from <http://www.uniprot.org/statistics/Swiss-Prot> on November 26th, 2013 (UTC+0800)

Statistics (UniProtKB/TrEMBL)

New entries	3,500,305
Updated entries	21,817,495
Unchanged entries	22,862,624
Total	48,180,424
Entries with updated sequences	24,395
With a fragmented AA sequence	4,732,694
With known alternative products	0



Data and image collected from <http://www.uniprot.org/statistics/Swiss-Prot> on November 26th, 2013 (UTC+0800)

IntAct (<http://www.ebi.ac.uk/intact/>)

EMBL-EBI

ServicesResearchTrainingAbout us


IntAct

- Home
- Advanced Search
- Tools
- Data Submission
- Downloads
- Documentation
- Acknowledgements
- Contact Us

FOLLOW US ON
twitter

[Tweets by @intact_project](#)

v.4.1.1



Back to IntAct homepage

IMEx member

IntAct > View

Search:


SearchClear

[Show Advanced Fields >](#) [MIQL syntax reference](#)

- Free text search will look by default for interactor identifier, species, interaction id, detection method, interaction type, publication identifier or author, interactor xrefs, interaction xrefs
- For a more specific search, use MIQL syntax or advanced search
- Search based on exact word matches eg. BRCA2 will not match BRCA2B
- Search for isoforms of 'P12345' by using 'P12345'

Examples

- Gene name: e.g. [BRCA2](#)
- UniProtKB Ac: e.g. [Q06609](#)
- UniProtKB Id: e.g. [dmc1](#)
- Pubmed Id: e.g. [10831611](#)


Support
and feedback

HomeSearchInteractions (436711)BrowseListsInteraction DetailsMolecule ViewGraph

IntAct provides a freely available, open source database system and analysis tools for molecular interaction data. All interactions are derived from literature curation or direct user submissions and are freely available. To perform a search in the IntAct database use the search box above.

Publications

Experiments

Interactions

Interactors

12233

31593

436711


79940


Citing IntAct


- The MintAct project—IntAct as a common curation platform for 11 molecular interaction databases. Orchard S et al [PMID: 24234451] Nucl. Acids Res. (2013) doi: 10.1093/nar/gkt1115 [\[Abstract\]](#) [\[Full Text\]](#)


Dataset of the month: November


- Identification and Comparative Analysis of Hepatitis C Virus-Host Cell Protein Interactions.



MINT



UniProt



SiB


I2D


InnateDB


Molecular Connections


MatrixDB


MBInfo

[jk/intact/?conversationContext=d](#)

Screenshot made at <http://www.ebi.ac.uk/intact/> on November 27th, 2013 (UTC+0800)

IntAct statistics

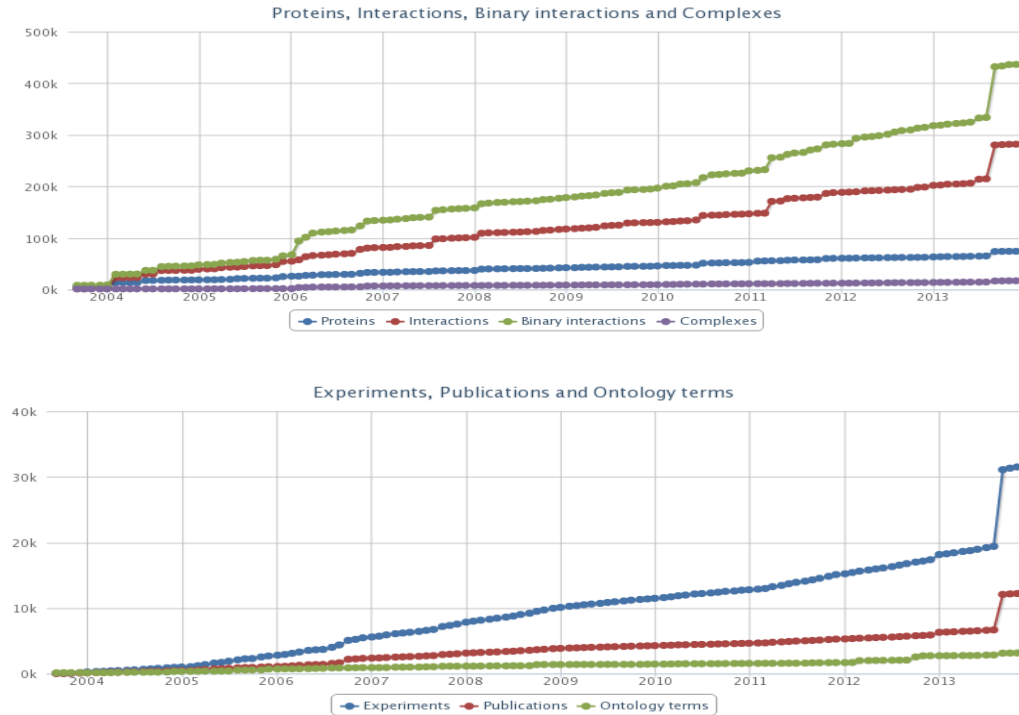
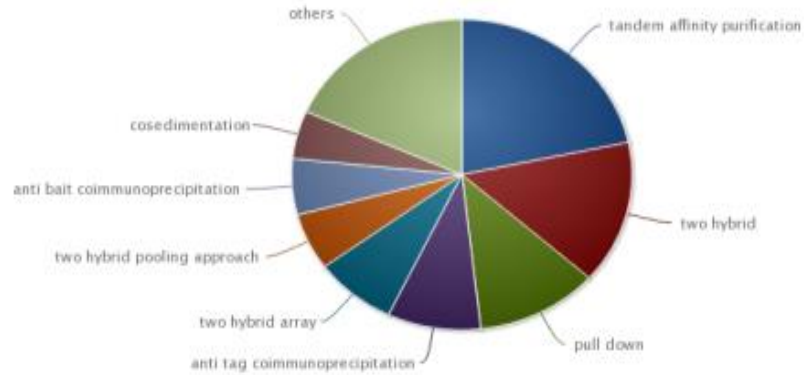
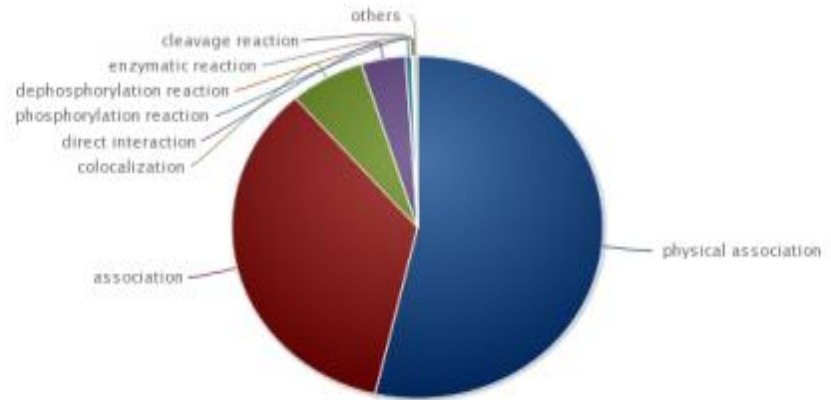


Image collected from <http://www.ebi.ac.uk/intact/pages/documentation/statistics.xhtml> on November 27th, 2013

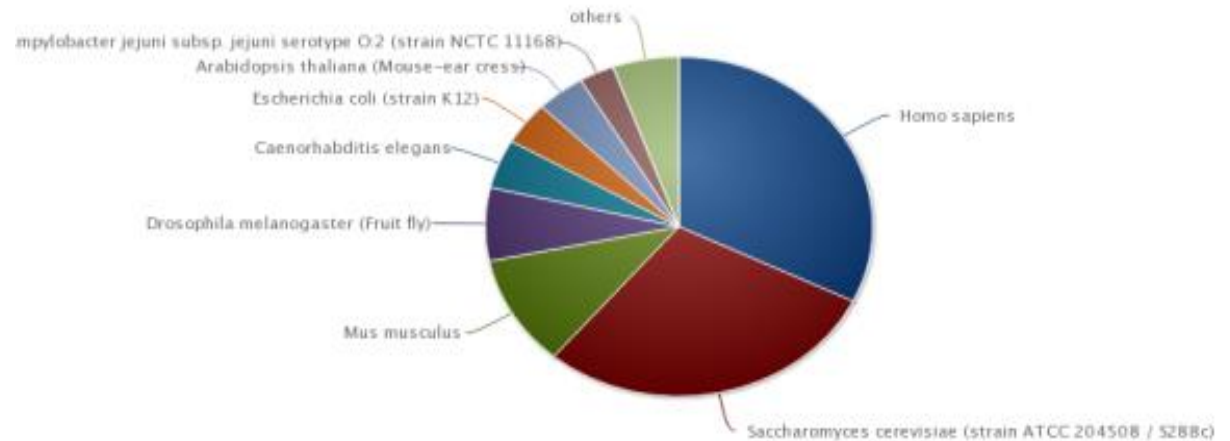
Interaction detection methods



Interaction types



Species



Clustal Omega (<http://www.clustal.org/omega/>)

EMBL-EBI 

Services Research Training About us

Clustal Omega

Input form Web services Help & Documentation

Share Feedback

[Tools](#) > [Multiple Sequence Alignment](#) > Clustal Omega

Multiple Sequence Alignment

Clustal Omega is a new multiple sequence alignment program that uses seeded guide trees and HMM profile-profile techniques to generate alignments.

STEP 1 - Enter your input sequences

Enter or paste a set of sequences in any supported format:

Or, upload a file: 選択されていません

STEP 2 - Set your parameters

OUTPUT FORMAT

The default settings will fulfill the needs of most users and, for that reason, are not visible.

(Click here, if you want to view or change the default settings.)

STEP 3 - Submit your job

☐ Be notified by email (Tick this box if you want to be notified by email when the results are available)

If you plan to use these services during a course please [contact us](#).

Please read the FAQ before seeking help from our support staff.

ClustalW2 – Phylogeny

(http://www.ebi.ac.uk/Tools/phylogeny/clustalw2_phylogeny/)

ClustalW2 - Phylogeny

[Input form](#) [Web services](#) [Help & Documentation](#)

[Share](#) [Feedback](#)

[Tools](#) > [Phylogeny](#) > ClustalW2

Phylogeny

Commonly used phylogenetic tree generation methods provided by the ClustalW2 program.

STEP 1 - Enter your multiple sequence alignment

[Enter or paste](#) a multiple sequence alignment in any [supported](#) format:

Or, [upload](#) a file: [浏览...](#)

STEP 2 - Set your Phylogeny options

[TREE FORMAT](#)


[DISTANCE CORRECTION](#)

[EXCLUDE GAPS](#)

[CLUSTERING METHOD](#)

[P I M](#)

InterPro & InterProScan (<http://www.ebi.ac.uk/interpro/>)



InterPro
Protein sequence analysis & classification

Examples: IPR020405, kinase, P51587, PF02932, GO:0007165

[Home](#) | [Release notes](#) | [Training & tutorials](#) | [FAQs](#) | [Download](#) | [About InterPro](#) | [Contact](#)

InterPro: protein sequence analysis & classification

InterPro provides functional analysis of proteins by classifying them into families and predicting domains and important sites. We combine protein signatures from a number of member databases into a single searchable resource, capitalising on their individual strengths to produce a powerful integrated database and diagnostic tool. [Read more about InterPro](#)

Analyse your protein sequence

[Search](#) | [Clear](#)


Documentation

[About InterPro](#): core concepts, update frequency, how to cite, team and consortium members.

[FAQs](#): what are entry types and why are they important, interpreting results, downloading InterPro?


[Web services documentation](#)

Protein focus



[Bee afraid, bee very afraid - neonicotinoids and the nAChRs family](#)
Bees are disappearing! The crisis is spreading around the world! To protect these important pollinators, the EU has proposed restricting the use of neonicotinoid pesticides. But what are neonicotinoids? And why were they chosen as pesticides in the first place? Let's find out more in this article.
[View PDF](#) | [Previous protein focus](#)

Publications



[InterPro in 2011: new developments in the family and domain prediction database](#)
A recently published paper describing new developments with the InterPro database (*Nucleic Acids Research*, 2012, Vol. 40, Database issue).
[HTML](#) | [PDF \(2,9Mb\)](#)

InterPro 45.0

19th November 2013

Features include:

- An update to PROSITE patterns (20.97), HAMAP (201310.09), PROSITE profiles (20.97).
- Integration of 324 new methods from the HAMAP, PANTHER and PROSITE profiles databases.

[Download](#) | [Read more](#)

InterProScan 5

Learn more >>

Latest News

- **InterProScan 5.2-45.0**
19 November 2013 - InterProScan 5.2-45.0 is now available. This release uses InterPro release 45.0 data and includes minor improvements over the 5.1-44.0 release.
[InterProScan 5.2-45.0 release notes](#)
- **InterProScan 5.1-44.0**
31 October 2013 - We are pleased to announce the release of InterProScan 5.1-44.0. This release contains performance improvements and minor bug fixes. It still uses InterPro release 44.0 data.
[InterProScan 5.1-44.0 release notes](#)
- **InterProScan 5 is out!**
26 September 2013 - We are delighted to

[Tools](#) | [InterProScan](#) | [Tools](#) | [BioMart](#) | [Tools](#) | [Utopia](#)

Screenshot made at <http://www.ebi.ac.uk/interpro/> on Nov 24th, 2013 (UTC+0800)

PRIDE (<http://www.ebi.ac.uk/pride/>)

EMBL-EBI



Services

Research

Training

About us



Proteomics Identifications Database

- Search
 - Advanced Search
 - Browse
 - PRIDE BioMart
- Submission Guidelines
- Documentation
- Tools
- Log in
- Acknowledgements
- Partners
- Contact PRIDE Support

PRIDE core version: 2.8.18

PRIDE web version: 2.8.18

Check out our new beta [website](#), tell us what you think ([feedback](#))

Introduction

The PRIDE PRoteomics IDentifications database is a centralized, standards compliant, public data repository for proteomics data, including protein and peptide identifications, post-translational modifications and supporting spectral evidence.

Submitting Data to PRIDE

PRIDE encourages and welcomes direct user submissions of mass spectrometry data to be published in peer-reviewed publications.

[Submit data to PRIDE](#)

Search PRIDE

Search

Examples:

PRIDE Experiment accession number: e.g. [8500](#)
UniProtKB Ac: e.g. [P29375](#)
UniProtKB Id: e.g. [KDM5A_HUMAN](#)
IPI: e.g. [IPI00619935](#)
Ensembl: e.g. [ENSP00000382688](#)

You can also [browse PRIDE](#) by species, tissue, cell type, GO terms and disease.

More complex searches can be performed using the [Advanced Search](#).

You can use the [PRIDE BioMart](#) for custom queries, such as linking PRIDE identifications to REACTOME pathways.

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>