UniProt
the Universal protein resource

www.uniprot.org

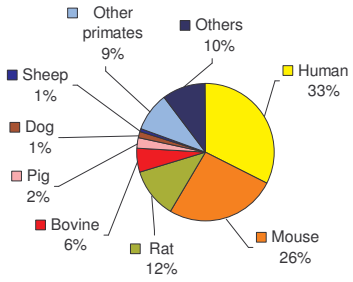
The Human Proteomics Initiative (HPI) of UniProtKB/Swiss-Prot

20 Years SIB swissprot EBI

The Human Proteomics Initiative (HPI) is devoted to the manual annotation of all known human protein sequences and their mammalian counterparts to the UniProtKB/Swiss-Prot quality standards.

Human chromosomal coverage:

14'400 human entries which represent **70% of the mapped proteins** (HGNC and Entrez Gene, May 2006).



Sequence accuracy

To avoid redundancy and increase sequence **accuracy**, the protein **sequences encoded** by the **same gene** are **merged** into a **single entry.** Sequences are manually checked, discrepancies are carefully analysed and documented in the FT keys VAR_SEQ, VARIANT, MUTAGEN and CONFLICT.

PARTICULAR EMPHASIS ON

Protein knowledge

Based on publications, internal and external scientific expertise and high-performance bioinformatic tools. This knowledge is mostly added in the appropriate topics of the comment (CC) lines and the feature (FT) lines.

Sequence diversity

20'000 to 25'000 genes for an estimate of 1 million proteins. Complexity is generated mainly by polymorphisms, alternative splicing and posttranslational modifications (PTMs).

Polymorphisms and disease mutations: 30% of human entries have at least one described polymorphism/disease mutation. Currently 54% of these polymorphisms are disease-linked and 14% human entries have an associated disease comment.

References

[37] VARIANTS DEMENTIA THR-573 AND ARG-705, AND CHARACTERIZATION OF VARIANTS DEMENTIA THR-573 AND ARG-705.
DOI=10.1002/1531-8249(200012)48:6<859::AID-ANA6>3.3.CO;2-T; PubMed=11117542 [NCBI, ExPASy, EBI, Israel, Japan]
Pickering-Brown S., Baker M., Yen S.-H., Liu W.-K., Hasegawa M., Cairns N., Lantos P.L., Rossor M., Iwatsubo T., Davies Y., Allsop D., Furlong R., Owen F., Hardy J., Mann D., Hutton M.;
"Pick's disease is associated with mutations in the tau gene.";
Ann. Neurol. 48:859-867(2000).

Features

VARIANT	573	573	1	K -> T (in a dementia resembling Pick disease; reduces the ability to promote microtubule assembly by 70%).	VAR_010344
VARIANT	705	705	1	G -> R (in a dementia resembling Pick disease; in vitro the mutation reduces the ability of tau to promote microtubule assembly by 25 to 30%).	VAR_010352

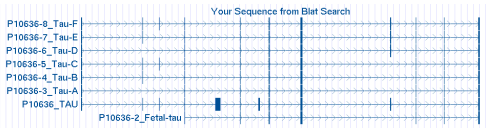
Comments

- ***DISEASE***: Defects in MAPT are a cause of frontotemporal dementia and parkinsonism linked to chromosome 17 (FTDP17) [MIM:600274, 172700]; also historically termed Pick disease. This form of frontotemporal dementia is characterized by presenile dementia with behavioral changes, deterioration of cognitive capacities and loss of memory. In some cases, parkinsonian symptoms are prominent. Neuropathological changes include frontotemporal atrophy often associated with atrophy of the basal ganglia, substantia nigra, amygdala. In most cases, protein tau deposits are found in glial cells and/or neurons.

Alternative products: 30% of human entries contain at least one isoform produced by alternative splicing, alternative initiation and/or alternative promoter usage.

Features

VAR_SEQ	1	43	AEPRQEFEVMEDHAGTYGLGDRKDQGGYTMHQDQEGDT DAGLK -> LRALQQRKR (in isoform Fetal-tau).
VAR_SEQ	44	72	Missing (in isoform Tau-A, isoform Tau-D and isoform Fetal-tau).
VAR_SEQ	73	101	Missing (in isoform Tau-A, isoform Tau-B, isoform Tau-D, isoform Tau-E and isoform Fetal-tau).
VAR_SEQ	102	103	Missing (in isoform Fetal-tau).
VAR_SEQ	124	374	Missing (in isoform Tau-A, isoform Tau-B, isoform Tau-C, isoform Tau-D, isoform Tau-E, isoform Tau-F and isoform Fetal-tau).
VAR_SEQ	394	459	Missing (in isoform Tau-A, isoform Tau-B, isoform Tau-C, isoform Tau-D, isoform Tau-E, isoform Tau-F and isoform Fetal-tau).
VAR_SEQ	591	621	Missing (in isoform Tau-A, isoform Tau-B, isoform Tau-C and isoform Fetal-tau).



Posttranslational modifications (PTMs): 50% of human entries describe at least one PTM.

References

[8] PHOSPHORYLATION AT SER-518.
DOI=10.1073/pnas.0404720101; PubMed=15302935 [NCBI, ExPASy, EBI, Israel, Japan]
Beausoleil S.A., Jedrychowski M., Schwartz D., Elias J.E., Villen J., Li J., Cohn M.A., Cantley L.C., Gygi S.P.;
"Large-scale characterization of HeLa cell nuclear phosphoproteins.";
Proc. Natl. Acad. Sci. U.S.A. 101:12130-12135(2004).

Features

MOD_RES	518	518	Phosphoserine (by PDPK).

GENOME
~25'000
human genes

TRANSCRIPTOME
~ 100'000
human transcripts

PROTEOME
~ 1 million
human proteins

Cross-references

Extensively cross-referenced to many databases within the HPI scope, such as HGNC, Ensembl, OMIM, MGI, RGD, RZPD, H-InvDB, etc.

UniProt
the Universal protein resource
Contact help@uniprot.org

Swiss Institute of Bioinformatics (SIB)
European Bioinformatics Institute (EMBL-EBI)
Protein Information Resource (PIR)

UniProt is mainly supported by the National Institutes of Health (NIH) grant 2 U01 HG02712-04. Additional support for the EBI's involvement in UniProt comes from the European Commission contract FELICS (021902) and from the NIH grant 5 P41 HG02273-06. UniProtKB/Swiss-Prot activities at the SIB are supported by the Swiss Federal Government through the Federal Office of Education and Science. PIR activities are also supported by the NIH grants for NIAID proteomic resource (HHSN266200400061C) and grid enablement (NCI-caBIG-ICR), and National Science Foundation grants for protein ontology (ITR-0205470) and BioTagger (IIS-0430743).