Supplementary File 2

AmyCo: the Amyloidoses Collection

Katerina C. Nastou[#], Georgia I. Nasi[#], Paraskevi L. Tsiolaki[#], Zoi I. Litou and Vassiliki A. Iconomidou*

Section of Cell Biology and Biophysics, Department of Biology, School of Sciences, National and Kapodistrian University of Athens, Panepistimiopolis, Athens 157 01, Greece

[#]These authors contributed equally.

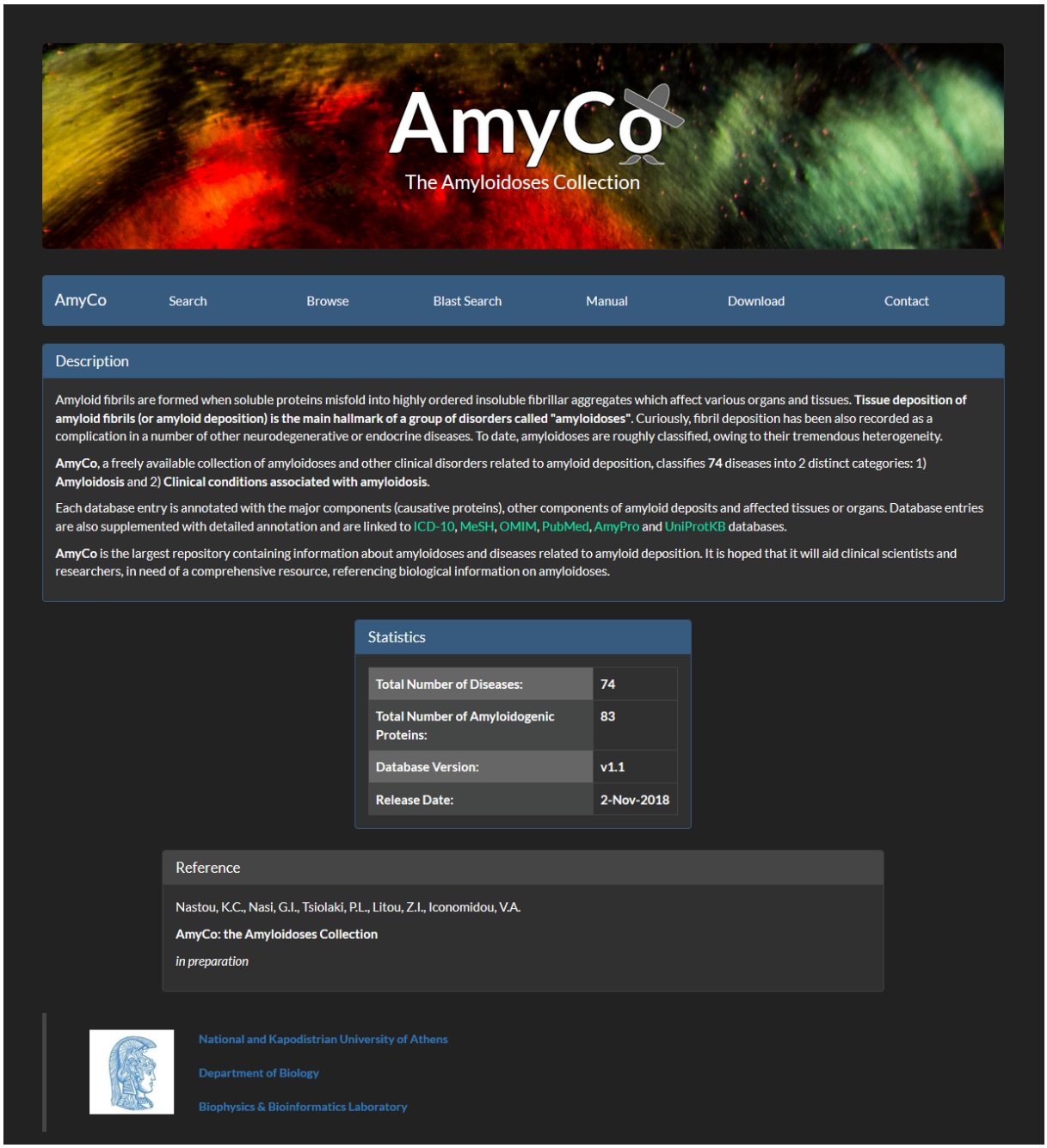
*Corresponding Author: veconom@biol.uoa.gr

Welcome to *AmyCo*, a freely available, literature-curated collection of amyloidoses and other pathological conditions associated with the deposition of amyloid fibrils.

Manual Contents

Home

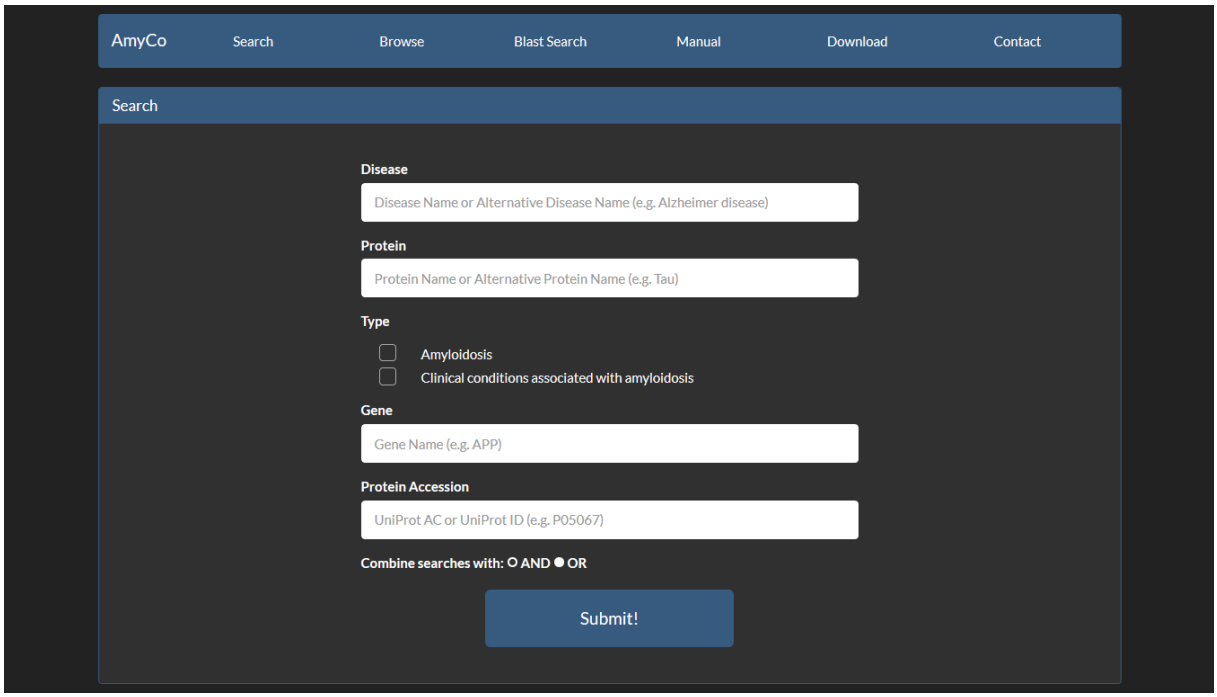
In order to visit *AmyCo*[*], the user should enter the following address: http://bioinformatics.biol.uoa.gr/AmyCo/. The page contains a short description of the repository and the *AmyCo* statistics.



[#] Logo background: A photomicrograph of Congo Red stained IAPP, taken in polarized light by Dr. P.L.Tsiolaki.

Search

The search tab allows the navigation of the database. A form with multiple options appears.



The search options are:

- by Disease Name

  The user may use a Name or an Alternative Disease Name, including the ISA name when available (e.g. Alzheimer Disease or Alzheimer Syndrome)

- by a Protein Name associated with a disease

  Proteins can be either a major (precursor protein) or a minor (co-deposited protein) component.

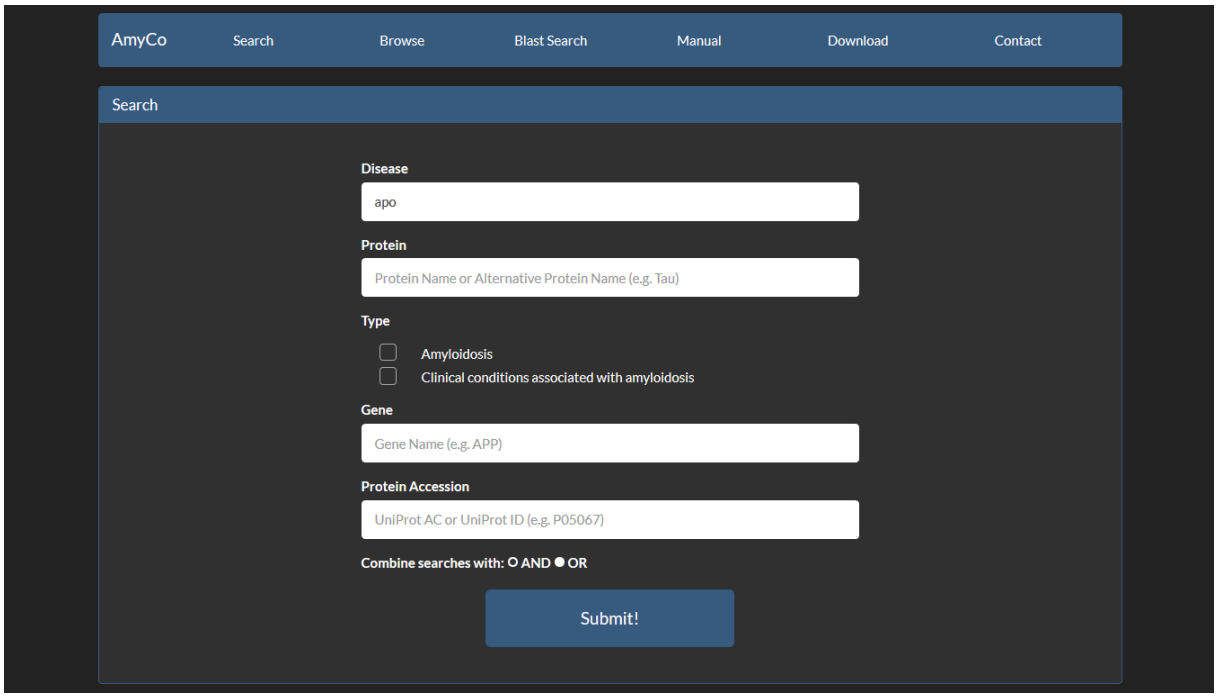
- by Disease Type

  AmyCo Disease entries are classified into two categories: **1) Amyloidosis**, when amyloid deposition is the main disease cause (e.g. AL amyloidosis) or amyloid deposits are present in tissue and organs (e.g. Alzheimer disease), and **2) Clinical conditions associated with amyloidosis**, when amyloidosis is a clinical feature of a disease or a syndrome (e.g. amyloid deposition in Waldenström's macroglobulinemia)

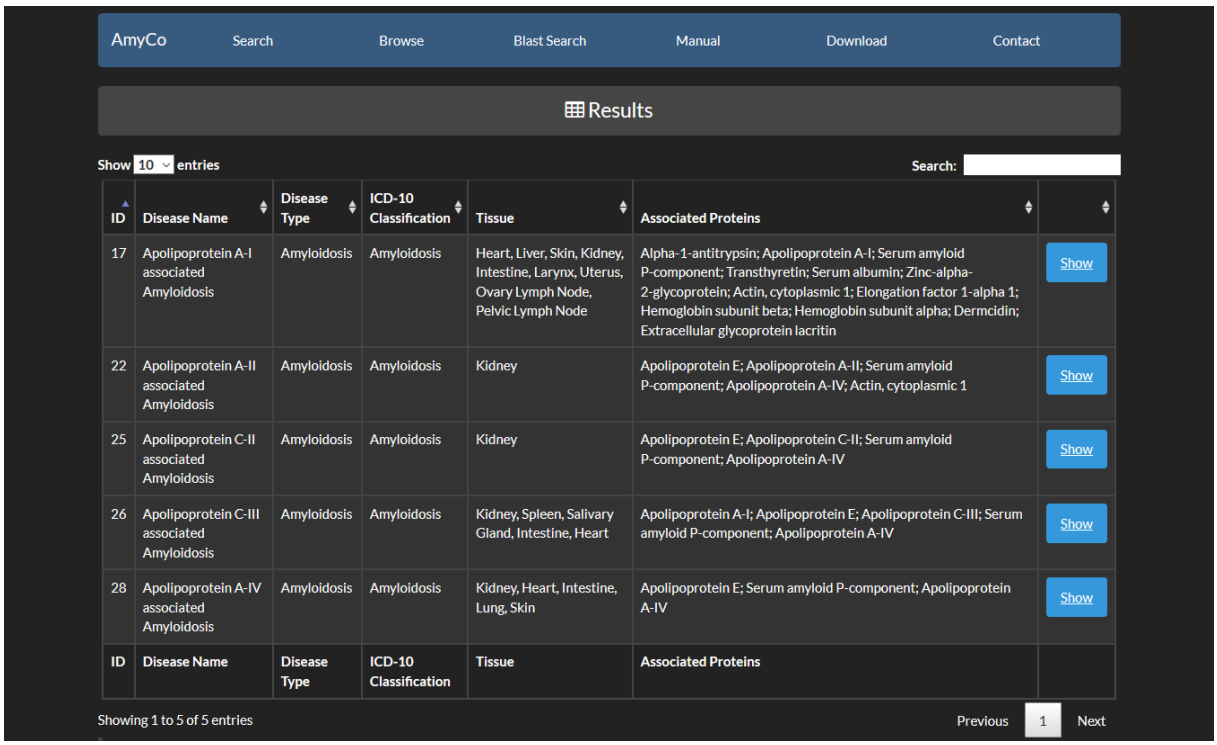
- by Gene Name (e.g. APP)

- by Protein Components based on a UniProt Accession or a UniProt Identifier (e.g. P05067, ***Two human UniProt ACs (P0DOX7, P0DOX8) corresponding to amyloidogenic light chains and four human UniProt ACs (P0DOX2, P0DOX4, P0DOX5, P0DOX6)corresponding to amyloidogenic heavy chains, available in AmyCo, are universal protein representatives annotated by the UniProtKB database in March 15th 2017**)*

The search, based on disease name, protein name and gene name does not require specific words. For example a user enters the word “apo” in the **disease search field**.



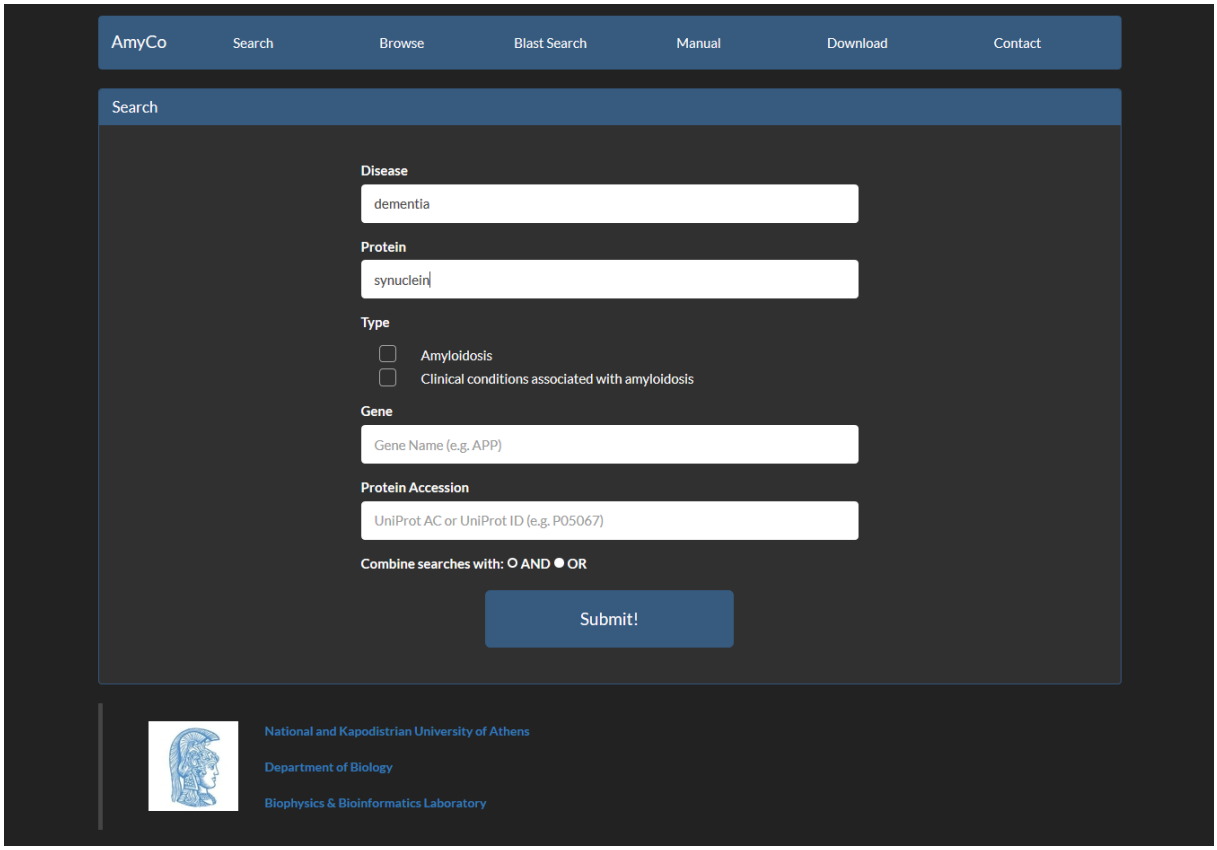
The result is all diseases containing the word “apo” in the **disease search field**.



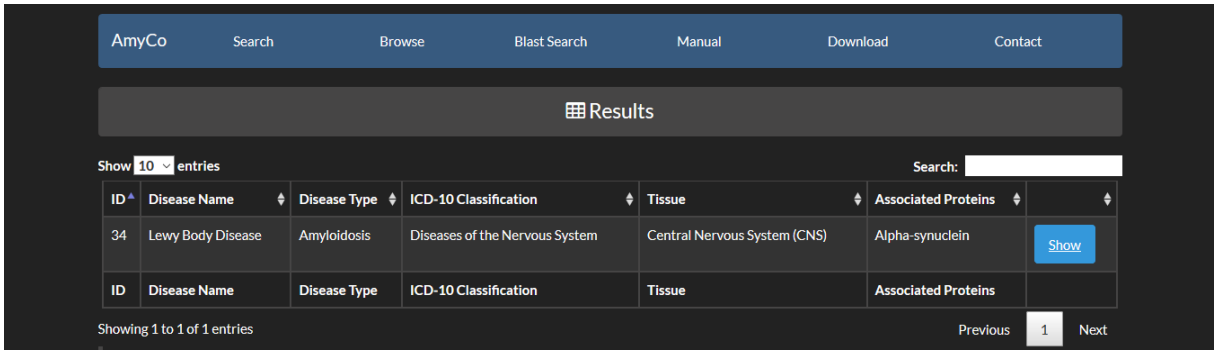
ID	Disease Name	Disease Type	ICD-10 Classification	Tissue	Associated Proteins	
17	Apolipoprotein A-I associated Amyloidosis	Amyloidosis	Amyloidosis	Heart, Liver, Skin, Kidney, Intestine, Larynx, Uterus, Ovary Lymph Node, Pelvic Lymph Node	Alpha-1-antitrypsin; Apolipoprotein A-I; Serum amyloid P-component; Transthyretin; Serum albumin; Zinc-alpha-2-glycoprotein; Actin, cytoplasmic 1; Elongation factor 1-alpha 1; Hemoglobin subunit beta; Hemoglobin subunit alpha; Dermcidin; Extracellular glycoprotein lacritin	Show
22	Apolipoprotein A-II associated Amyloidosis	Amyloidosis	Amyloidosis	Kidney	Apolipoprotein E; Apolipoprotein A-II; Serum amyloid P-component; Apolipoprotein A-IV; Actin, cytoplasmic 1	Show
25	Apolipoprotein C-II associated Amyloidosis	Amyloidosis	Amyloidosis	Kidney	Apolipoprotein E; Apolipoprotein C-II; Serum amyloid P-component; Apolipoprotein A-IV	Show
26	Apolipoprotein C-III associated Amyloidosis	Amyloidosis	Amyloidosis	Kidney, Spleen, Salivary Gland, Intestine, Heart	Apolipoprotein A-I; Apolipoprotein E; Apolipoprotein C-III; Serum amyloid P-component; Apolipoprotein A-IV	Show
28	Apolipoprotein A-IV associated Amyloidosis	Amyloidosis	Amyloidosis	Kidney, Heart, Intestine, Lung, Skin	Apolipoprotein E; Serum amyloid P-component; Apolipoprotein A-IV	Show
ID	Disease Name	Disease Type	ICD-10 Classification	Tissue	Associated Proteins	

All searches can be combined with logical operators (AND/OR), in order to make the search result as specific as possible.

For example if we proceed to the following combined search:

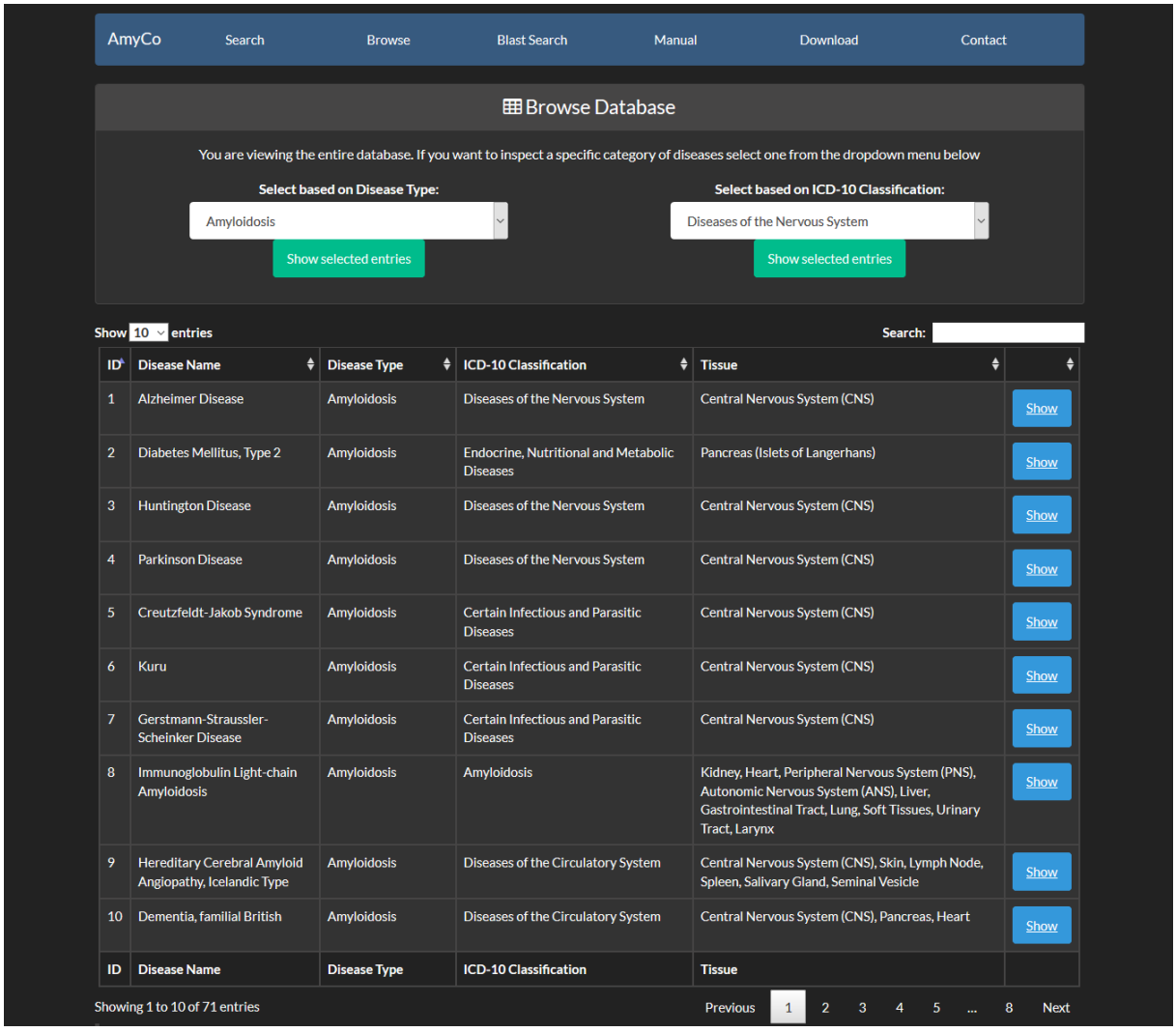


We will get only the disease, which is classified as dementia and is associated with synuclein.

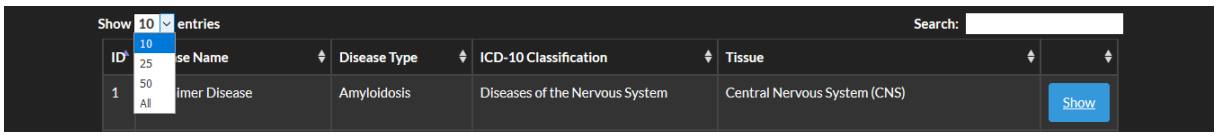


Browse

The **browse tab** allows the browsing of the database. User can apply filters and browse the database by selecting the disease Type or/and the ICD-10 Classification.



At first all entries appear. Each page shows 10 entries by default but this can be altered from the dropdown menu at the top left corner of the table to 25, 50 or all entries. Moreover, users can perform non-specific searches using the search option at the top right corner of the data table.

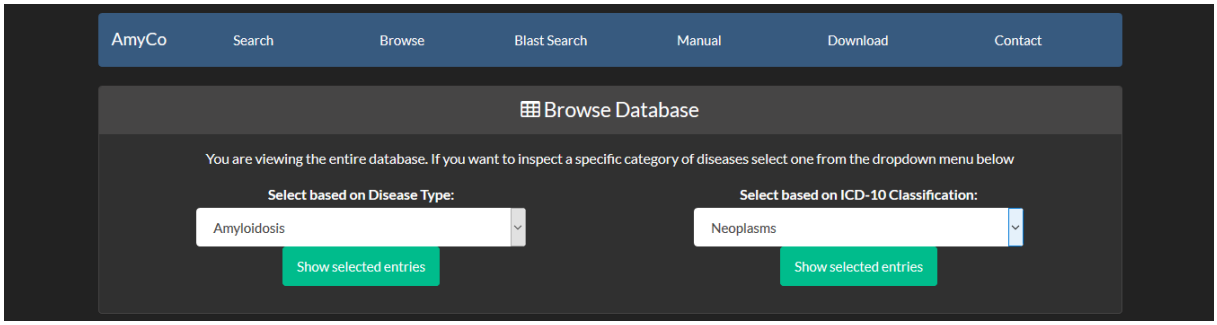


Filters and browsing can be applied according to the classification below.

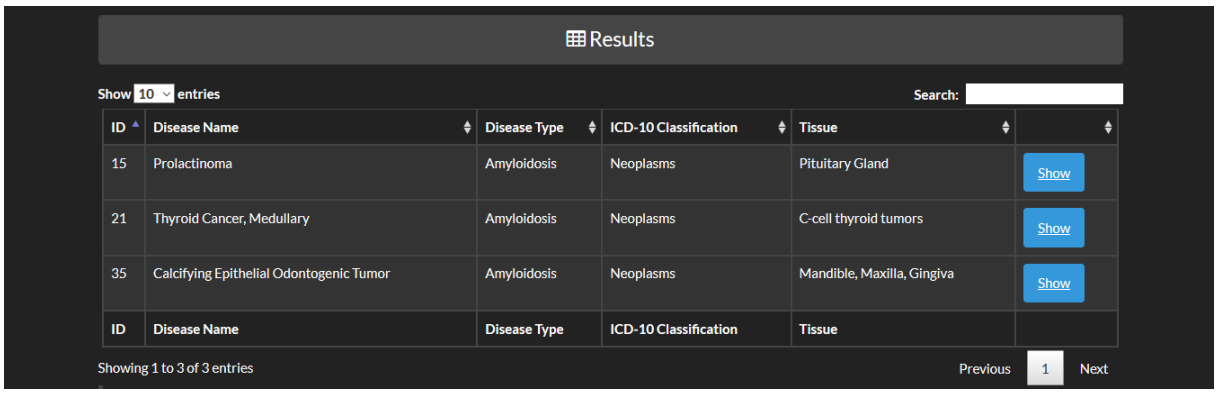
Disease type and ICD-10 classification

	Disease Type	ICD-10 Classification
1	**Amyloidosis**	Amyloidosis
		Endocrine, Nutritional and Metabolic Diseases
		Diseases of the Circulatory System
		Diseases of the Nervous System
		Neoplasms
		Certain Infectious and Parasitic Diseases
		Diseases of the Eye and Adnexa
2	**Clinical conditions associated with amyloidosis**	Endocrine, Nutritional and Metabolic Diseases
		Diseases of the Respiratory System
		Chromosomal Abnormalities
		Diseases of the Musculoskeletal System and Connective Tissue
		Neoplasms
		Diseases of the Skin and Subcutaneous Tissue
		Diseases of the Digestive System
		Diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism
		Other

For example, if the user selects **Amyloidosis** as the disease type and **Neoplasms** as the ICD-10 classification and presses *Show Selected Entries*.



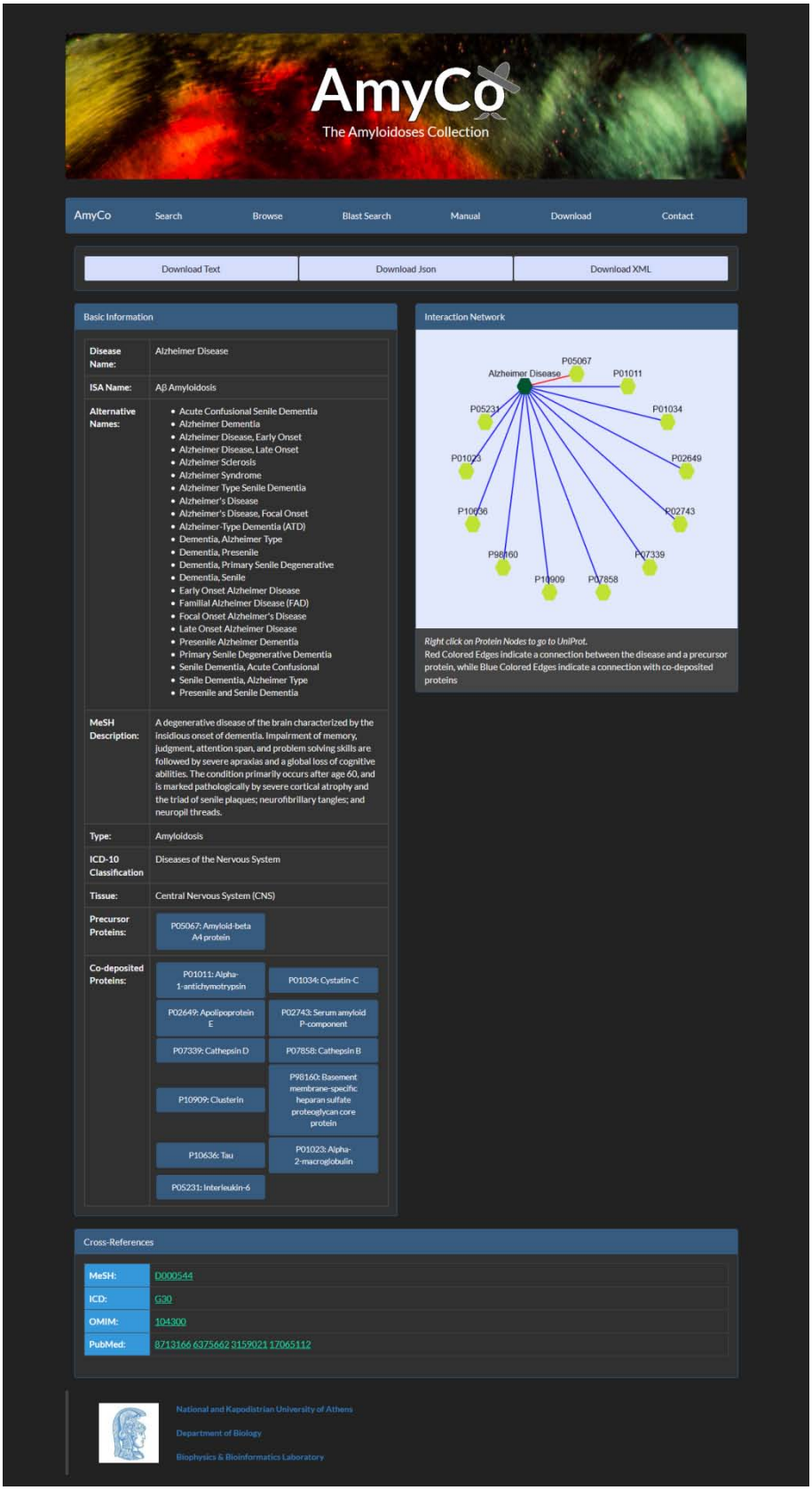
The selected diseases are:



Through browsing the database or after a search is submitted, a list of diseases appears, as the shown above. The list contains the Disease Name, the ICD-10 Classification, the disease type and the tissue(s), in which deposits are located. When the user presses the Show button they are redirected to the Entry page of a Disease.

Disease Entry

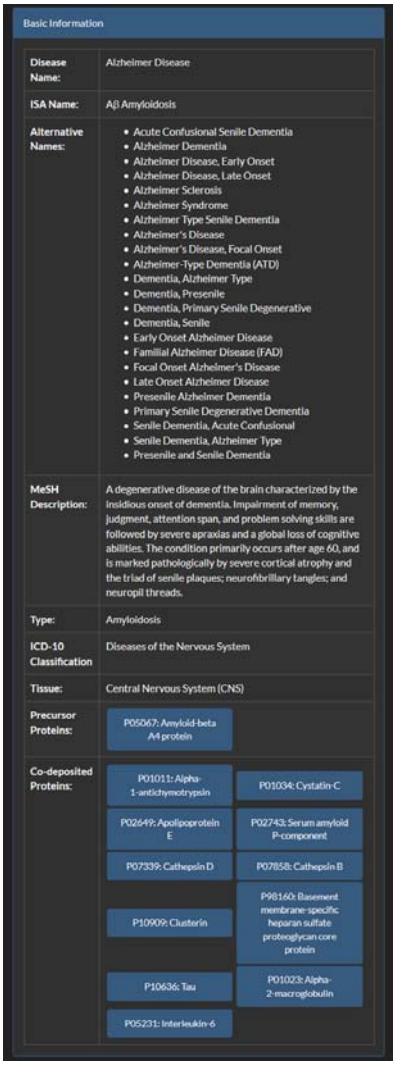
The Entry page contains information about the disease. CytoscapeJS is integrated to visualize the relationship between the disease and proteins found on amyloid deposits.



At the top of the page the user can find three buttons; *Download Text*, *Download Json* and *Download XML*. By pressing these buttons the user can download all page information in text, Json or XML format respectively.

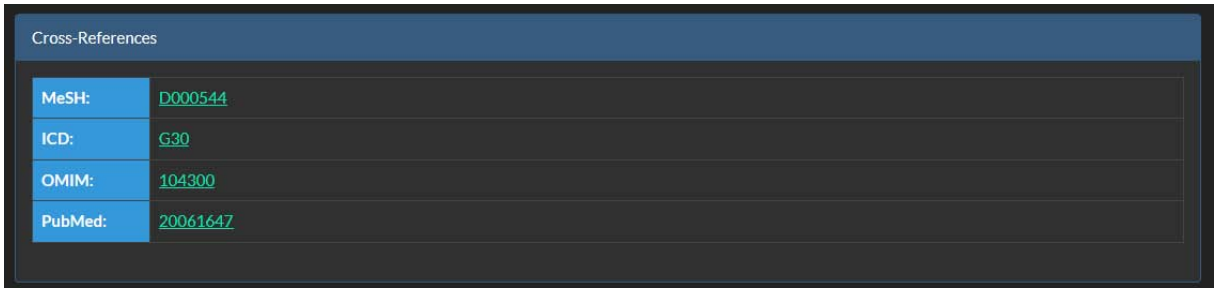
The basic disease information available is:

- ✓ Disease Name (MeSH name, when the disease is a MeSH entry)
- ✓ ISA Name (International Society of Amyloidosis name, when available)
- ✓ Alternative Names
- ✓ Disease Description (a short description from MeSH)
- ✓ Disease Type (see here)
- ✓ Disease Association (from ICD-10, when available)
- ✓ Tissue(s) where amyloid deposits are located (Tissues corresponding to "Clinical conditions associated with amyloidosis" were manually collected from the scientific literature)
- ✓ Precursor Proteins of Amyloid Deposits
- ✓ Co-deposited Protein of Amyloid Deposits



The panel at the bottom of the page contains references to other databases. The databases are:

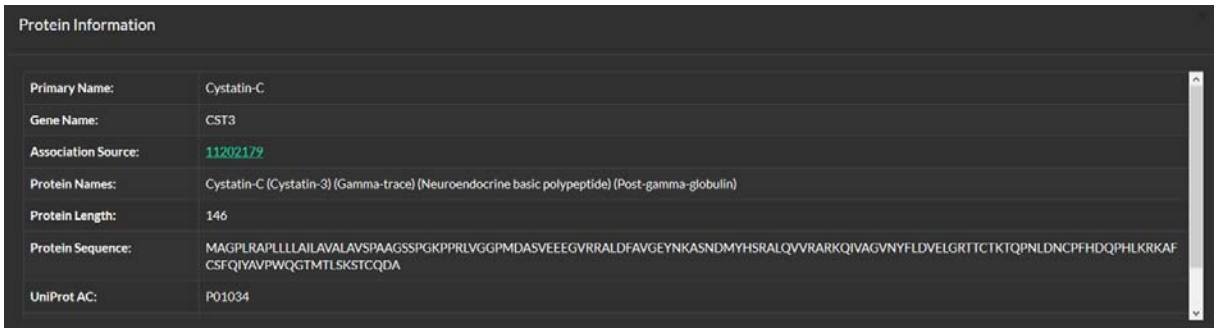
- MeSH
- ICD
- OMIM
- PubMed



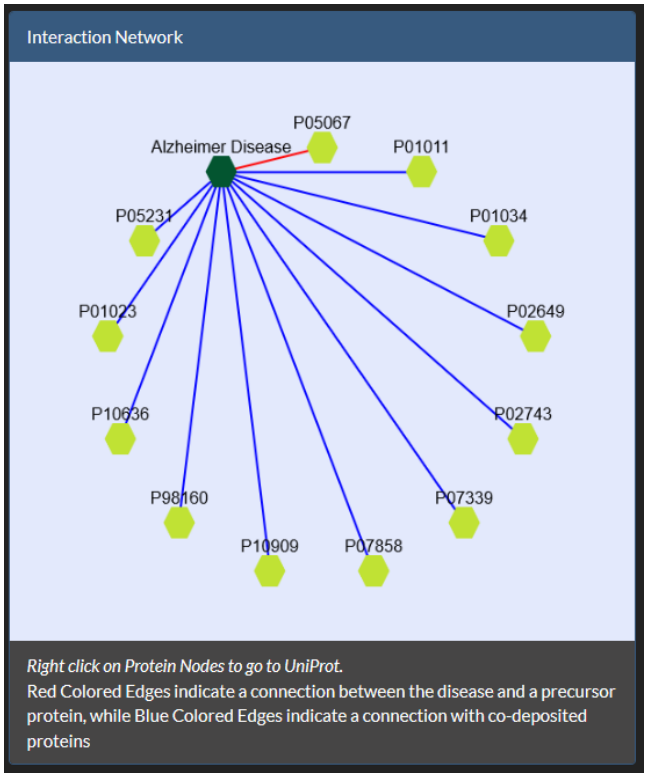
When the user presses on the buttons of precursor or co-deposited components a new page opens with information about the proteins.

- Primary Name
- Gene Name
- Other Protein Names
- Protein Sequence
- Protein Length
- Uniprot AC
- Uniprot ID

For precursor proteins an external link to AmyPro is given when available. For co-deposited proteins the peer reviewed publication (*Association Source*) of disease is also provided.



A CytoscapeJS viewer is integrated in the page for the visualization of bipartite graphs, showing the association between **each disease** with **precursor and co-deposited proteins**.



The diseases are colored green and the protein components yellow. Red Colored Edges indicate a connection between the disease and a Precursor Protein, while Blue Colored Edges indicate a connection with Co-deposited Proteins. Each protein node is also a hyperlink to UniProt.

If your browser prevents you from opening pop-up windows, please select *Allow pop-ups for this site*.

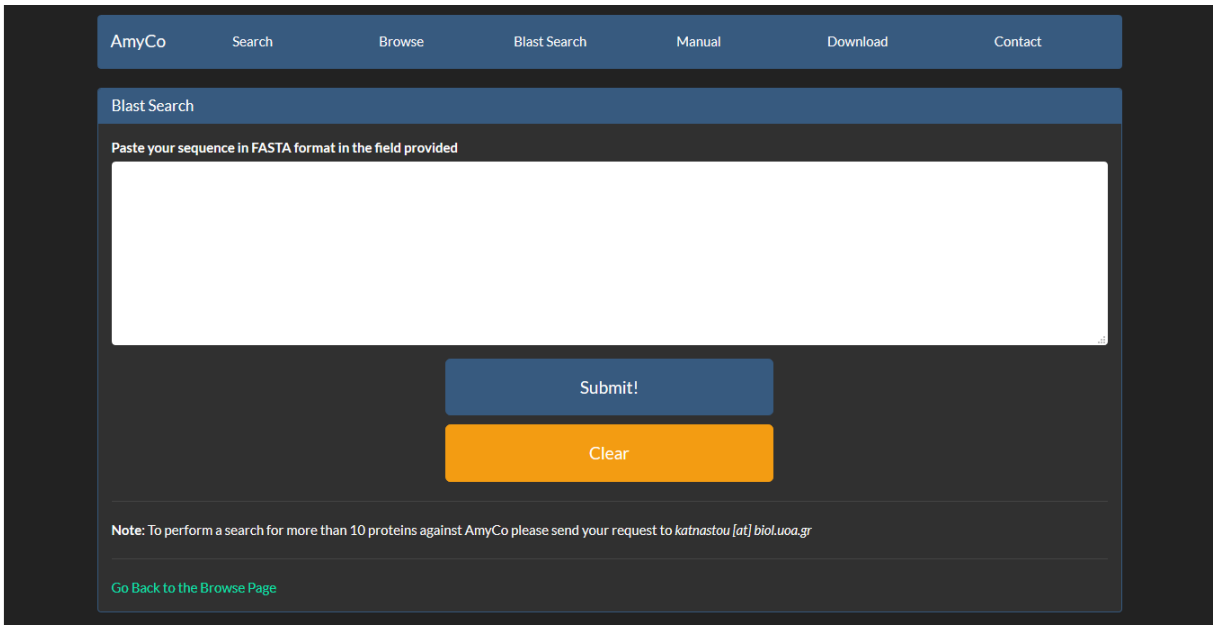
If during scrolling through the page you accidentally lose the network view, please reload the page to see it again.

BLAST Search

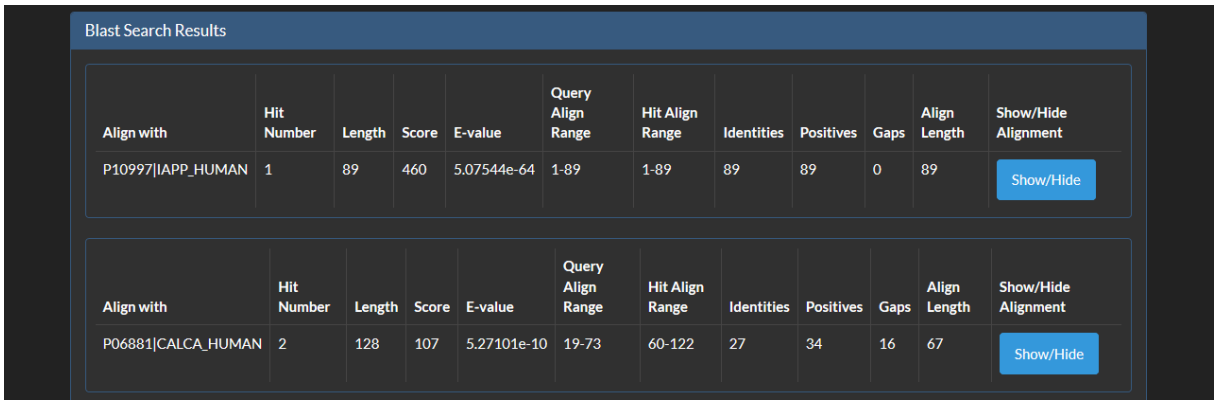
With the BLAST search tool, the user may submit a sequence and search the database for homologous protein molecules.

The input for the BLAST application is a sequence in the standard FASTA format.

```
>sp|P01034|27-146
SSPGKPPRLVGGPMDASVEEEGVRR
```

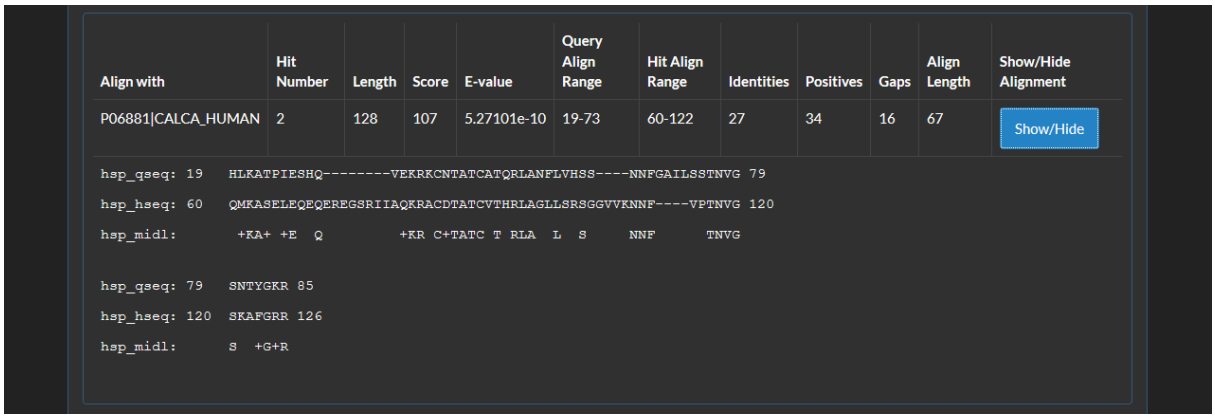


The result page of the BLAST search shows a list of the Blast hits with significant alignment on the query sequence. The list is in a table format including the target protein, the Length of the target sequence and the Query and finally, Target align range.



The BLAST results can be compared through the Score, the E-value, the Identities and Positives.

Furthermore, the user can have a more detailed view of each alignment through the Show button at the end of each line:



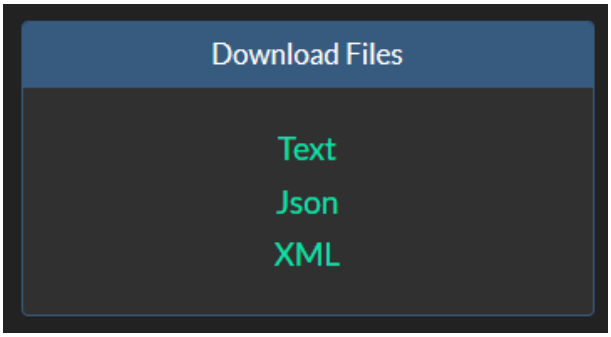
Align with	Hit Number	Length	Score	E-value	Query Align Range	Hit Align Range	Identities	Positives	Gaps	Align Length	Show/Hide Alignment
P06881|CALCA_HUMAN	2	128	107	5.27101e-10	19-73	60-122	27	34	16	67	Show/Hide

```
hsp_qseq: 19   HLKATPIESHQ--------VEKRKCNTATCATQRLANFLVHSS----NNFGAILSSTNVG 79
hsp_hseq: 60   QMKASELEQEQEREGSRIIAQKRACDTATCVTHRLAGLLSRSGGVVKNNF----VPTNVG 120
hsp_midl:       +KA+ +E  Q         +KR C+TATC T RLA  L  S     NNF      TNVG

hsp_qseq: 79   SNTYGKR 85
hsp_hseq: 120  SKAFGRR 126
hsp_midl:      S  +G+R
```

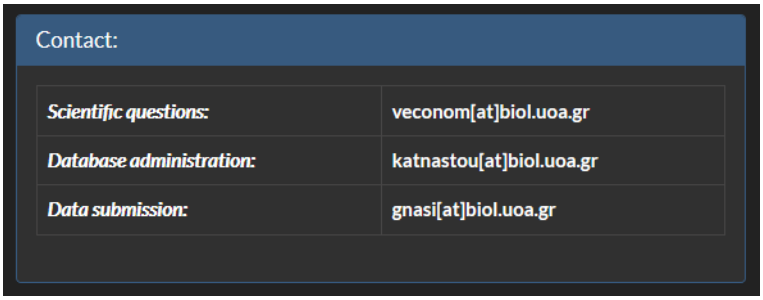
Download

User can download all database files in Text, JSON or XML format.

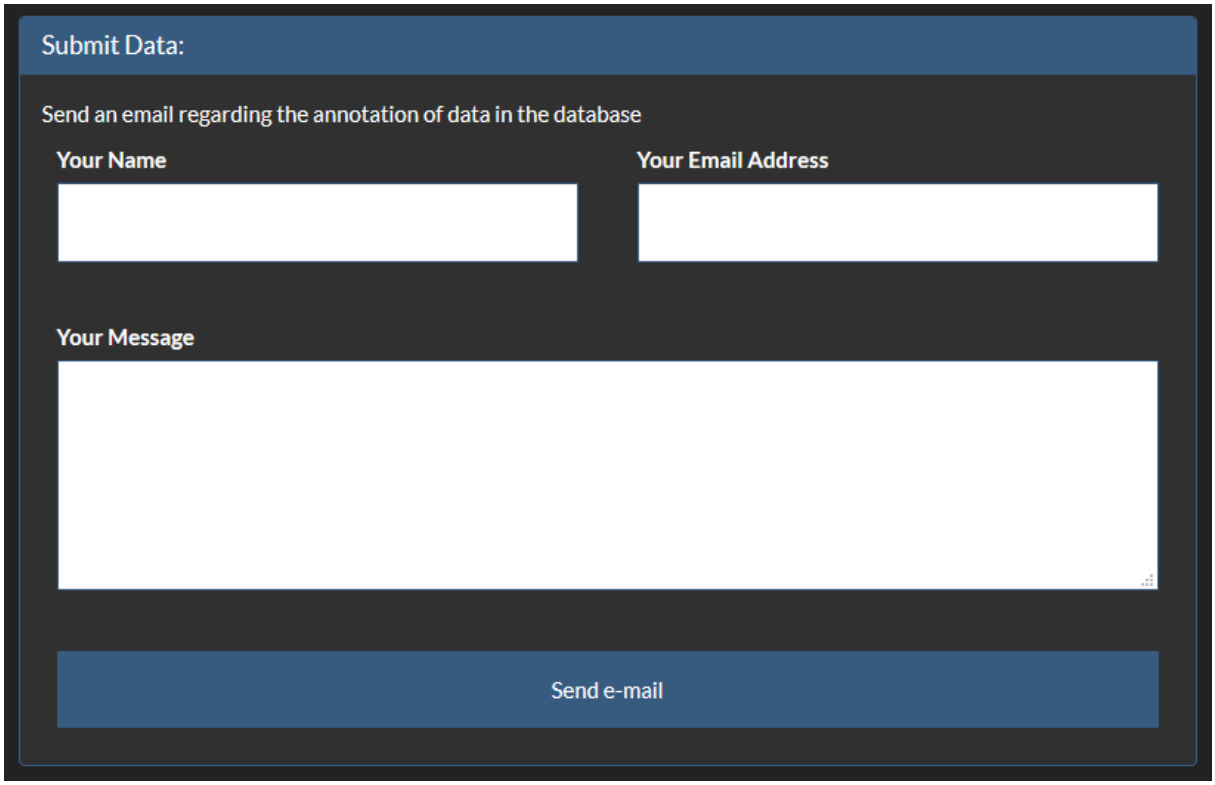


Contact

Users can contact us for more information at the emails specified at the contact page.



Users are encouraged to submit data by using the form below. Data will be reviewed and later will be added to the database by the authors.



Related publications to the current work are also presented.

Database Technologies

AmyCo is based on modern technologies. User should have Javascript enabled on the web browser.

AmyCo follows the General Data Protection Regulation (EU) 2016/679 ("GDPR") regulation. See the Privacy page for more information.

References

Hazenberg, B. P. (2013). "Amyloidosis: a clinical overview." Rheum Dis Clin North Am**39**(2): 323-345.

Winklhofer, K. F., J. Tatzelt, et al. (2008). "The two faces of protein misfolding: gain- and loss-of-function in neurodegenerative diseases." EMBO J**27**(2): 336-349.

Valastyan, J. S. and S. Lindquist (2014). "Mechanisms of protein-folding diseases at a glance." Dis Model Mech**7**(1): 9-14.

Agorogiannis, E. I., G. I. Agorogiannis, et al. (2004). "Protein misfolding in neurodegenerative diseases." NeuropatholApplNeurobiol**30**(3): 215-224.

Sipe, J. D., M. D. Benson, et al. (2016). "Amyloid fibril proteins and amyloidosis: chemical identification and clinical classification International Society of Amyloidosis 2016 Nomenclature Guidelines." Amyloid**23**(4): 209-213.

Soto, C. (2001). "Protein misfolding and disease; protein refolding and therapy." FEBS Lett**498**(2-3): 204-207.

Eisenberg, D. and M. Jucker (2012). "The amyloid state of proteins in human diseases." Cell**148**(6): 1188-1203.

Pepys, M. B. (2006). "Amyloidosis." Annu Rev Med**57**: 223-241.

Cohen, A. S., A. Frensdorff, et al. (1962). "A study of the fine structure of the amyloid associated with familial Mediterranean fever." Am J Pathol**41**: 567-578.

Fam, A. G., A. J. Lewis, et al. (1981). "Multiple myeloma and amyloid bone lesions complicating rheumatoid arthritis." J Rheumatol**8**(5): 845-850.

Ogami, Y., M. Takasugi, et al. (1989). "Waldenstrom'smacroglobulinemia associated with amyloidosis and crescentic glomerulonephritis." Nephron**51**(1): 95-98.

Tschang, T. P. (1976). "Nodular malignant lymphoma and amyloidosis. A case report." Cancer**38**(5): 2192-2196.