

- C. For human proteins, 14 sub-cellular locations (centriole, cytoplasm, cytoskeleton, endoplasmic reticulum, endosome, extracell, Golgi apparatus, lysosome, microsome, mitochondrion, nucleus, peroxisome, plasma membrane, and synapse) are provided.

Step 2. Set BLAST E-value threshold

When the BLAST E-value between the query and the most similar protein with GO annotation(s) is above this threshold, the amino acid composition of the query sequence instead of the GO terms is used for the feature vectors of the SVM classifier.

Step 3. Set multiplex threshold

A multiplex parameter θ provides multiple localizations of proteins, meaning that any predicted localization with a probability score higher than $[\theta \times \textit{highest probability of location}]$ will be assigned to the query protein as well.

Step 4. Sequence Submission (maximum 500 sequences)

Copy and paste your protein sequences in a FASTA format into the provided textbox. Alternatively, you may upload your sequences as a FASTA file.

Step 5. Click button!

Press the '**Predict**' button to obtain the result. We kindly ask users of WegoLoc to submit jobs in batches of up to 500 protein sequences at a time and not to submit another until processing of the previous run has been completed. The prediction will take approximately 2 seconds for one protein sequence.

WegoLoc Output

Prediction Results

User input options

Dataset, Kingdom	BaCellLo dataset, Fungi
BLAST E-value threshold	1.
Multiplex threshold	1.0

After table is shown below, you can download result [download](#)

No	Sequence name	Predicted locations	Probability of each location (%)	Weight, GO terms (description, evidence code[1 , 2])	Best BLAST hit (UniProtKB Accession: E-value)
1	RL17A_YEAST	cytoplasm	cytoplasm: 50.0 mitochondrion: 33.3 nucleus: 16.7 secretory: 0.0 [prob. threshold: 50.0]	6.33, GO:0005737 (cytoplasm, IDA) 5.88, GO:0005840 (ribosome, IEA) 5.85, GO:0003735 (structural constituent of ribosome, IEA) 5.50, GO:0006412 (translation, IEA) 4.98, GO:0030529 (ribonucleoprotein complex, IEA) 3.85, GO:0015934 (large ribosomal subunit, IEA) 3.84, GO:0005622 (intracellular, IEA) 0.00, GO:0022625 (cytosolic large ribosomal subunit, NAS)	P05740: 1e-100
2	SRL3_YEAST	cytoplasm	cytoplasm: 50.0 mitochondrion: 33.3 nucleus: 16.7 secretory: 0.0 [prob. threshold: 50.0]	6.33, GO:0005737 (cytoplasm, IDA) 1.73, GO:0006139 (nucleobase, nucleoside, nucleotide and nucleic acid metabolic process, IEA)	P36167: 1e-161

[Try another dataset](#)

1. The top table shows the input options selected by the user.
2. Prediction results are downloadable as a text file by clicking [download](#).
3. Descriptions of each column in the lower table:
 - A. Sequence name : the name of input amino acid sequence that is put after '>' notation in the FASTA input format.
 - B. Predicted locations : the locations predicted by WegoLoc with higher scores than a given threshold.
 - C. Probability of each location (%): the probability scores of each location predicted by WegoLoc. If the multiplex threshold θ is used, localizations with a probability score higher than $[\theta \times \text{the highest probability score in this column}]$ will also be assigned to the query proteins. This probability threshold is represented in the bottom of the cell.
 - D. Weight, GO terms (description, evidence codes): Weight means the value of the following GO term calculated by using WGO algorithm. The GO term links to the corresponding GO term information, lineage of GO terms and gene product association. Description shows the short biological explanation of the GO term. Evidence code means the attribution of the GO annotation such as a literature reference, another database or a computational analysis; detailed information on the meaning of these evidence codes can be found at <http://www.geneontology.org/GO.evidence.shtml>.
 - E. Best BLAST hit (UniProtKB Accession: E-value): When a query sequence is entered, WegoLoc searches the proteins that has GO annotation(s) for the most similar protein by BLAST and makes use of all the corresponding GO terms. UniProtKB Accession provides the accession of the most similar protein and corresponding BLAST E-value.