

An Integrated Approach for Microprotein Identification and Sequence Analysis

Omar Brito-Estrada^{*1}, Keira R. Hassel^{*1}, Catherine A. Makarewich^{1,2}

¹The Heart Institute, Division of Molecular Cardiovascular Biology, Cincinnati Children's Hospital Medical Center ²Department of Pediatrics, University of Cincinnati College of Medicine

* These authors contributed equally

Corresponding Author

Catherine A. Makarewich
cat.makarewich@cchmc.org

Citation

Brito-Estrada, O., Hassel, K.R., Makarewich, C.A. An Integrated Approach for Microprotein Identification and Sequence Analysis. *J. Vis. Exp.* (185), e63841, doi:10.3791/63841 (2022).

Date Published

July 12, 2022

DOI

10.3791/63841

URL

jove.com/video/63841

Materials

Name	Company	Catalog Number	Comments
Website	Website Address	Requirements	
Clustal Omega Multiple Sequence Alignment Tool	https://www.ebi.ac.uk/Tools/msa/clustalo/	Web browser	Multiple sequence alignment program for the efficient alignment of FASTA sequences (i.e. for cross-species comparison of identified microproteins)
COXPRESSdb	https://coxpresdb.jp	Web browser	Provides co-regulated gene relationships to estimate gene functions
EMBL-EBI Bioinformatics Tools FAQs	https://www.ebi.ac.uk/seqdb/confluence/display/JDSAT/Bioinformatics+Tools+FAQ	Web browser	Frequently Asked Questions (FAQs) for EMBL-EBI tools. Includes the color coding key for protein sequence alignments
European Bioinformatics Institute (EMBL-EBI), Tools and Data Resources	https://www.ebi.ac.uk/services/all	Web browser	Comprehensive list of freely available websites, tools and data resources
ExPASy - Swiss Bioinformatics Resource Portal	https://www.expasy.org	Web browser	Suite of bioinformatic tools and resources for protein sequence analysis that is maintained by the Swiss Institute of Bioinformatics (SIB)
National Center for Biotechnology Information (NCBI) Conserved Domain Search	https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi	Web browser	Search tool to identify conserved domains within protein or coding nucleotide sequences
Pfam 35	http://pfam.xfam.org	Web browser	Protein family (Pfam) database, provides alignments and classification of protein families and domains
PhyloCSF Track Hub Description	https://genome.ucsc.edu/cgi-bin/hgTrackUi?hgsid=1267045267_TEc99h2oW5QedaCd4ir8aZ65ryaD&db=mm10&c=chr2&g=hub_109801_PhyloCSF_smooth	Web browser	Detailed description of the Smoothed PhyloCSF tracks and PhyloCSF Track Hub

SignalP 6.0	https://services.healthtech.dtu.dk/service.php?SignalP-6.0	Web browser	Predicts the presence of signal peptides and the location of their cleavage sites
TMHMM - 2.0	https://services.healthtech.dtu.dk/service.php?TMHMM-2.0	Web browser	Prediction of transmembrane helices in proteins
UCSC Genome Browser BLAT Search	https://genome.ucsc.edu/cgi-bin/hgBlat	Web browser	Tool used to find genomic regions using DNA or protein sequence information
UCSC Genome Browser Gateway	https://genome.ucsc.edu/cgi-bin/hgGateway	Web browser	Direct link to the UCSC Genome Browser Gateway
UCSC Genome Browser Home	https://genome.ucsc.edu/	Web browser	Home website for the UCSC Genome Browser
UCSC Genome Browser Track Data Hubs	https://genome.ucsc.edu/cgi-bin/hgHubConnect#publicHubs	Web browser	Direct link to Track Data Hubs/Public Hubs database to search for and load the PhyloCSF Tracks
UCSC Genome Browser User Guide	https://genome.ucsc.edu/goldenPath/help/hgTracksHelp.html	Web browser	Comprehensive user guide detailing how to navigate the UCSC Genome Browser
WoLF PSORT	https://wolfsort.hgc.jp	Web browser	Protein subcellular localization prediction tool