

Dataset S1: FASTA sequences

FULL LENGTH FASTA SEQUENCES

aminoacid sequences downloaded from VectorBase: <https://vectorbase.org/vectorbase/app>

In yellow: predicted Signal Peptides according to SignalP Server:

<http://www.cbs.dtu.dk/services/SignalP/>

In light blue: predicted pro-peptides according to ProP 1.0 and Spider P (mature peptides are listed in red):

<http://www.cbs.dtu.dk/services/ProP/>,

<http://www.arachnoserver.org/spiderP.html>

```
>hyp6.2 AGAP006495
MGRVMCLRLRLMSTLLVVLIVGKKTNAAPQVTEAPGNVGSTYSPMADIGRLATGATKLFQFWNTGTRFGTELSRRTFDF
LRVKK
>hyp17 AGAP000151
MKLYSGYRLLVLLVMTVCCLLLFIAPTGAADPLPGQTQRTLGYRGNDKRATPPMHSLSGSGARMAMTGGGILGGIFSAL
>hyp15 AGAP000152
MKFYSVGKLVKVLVMVAVCCLLLCTAPTGAADPLPGRDRNTIANKSKDKKASAPKHSLSGTGARMALTGGGVLGGLVTNM
>hyp12 AGAP008306
MRFCCVALIGLLCSVQSVTAANDPVDALGACSGNLFGLLMTRLQQMVEDFTACRQEATANDPQHDCSDSIQRAKVDLQQQ
LVNYSYCTKNIQ
>hyp10 AGAP008307
MRFLTTLTVGLLVWVGVFATVNAEDPRTELIGCGSVLFHLAANRLSLQLEEFVAVCKRSNPGYDCSDSIHRAISDLQQGLF
DLNHCTKDIR
>sg2a AGAP006504
MKLFLTLTLTSLVVMVFAALPAHHHSRGGDGSSANSTGQSYNNSGSPNNSAGVPDFGFNSQSNVPGFGNGQQPGQQQQQQQ
GQGFPPFGQSGFSGFSGNRQPFQGNQQGQDGNQQQGRGIPFFGQGGGQGGIPSGSGQQNGGIPSLGNGQAQSGFSPF
GNGQQGGNFPFFG
>sg2 AGAP006506
MKSILVTFATLSVALVVVAIPTSFNYGGGGGYFINGTGQSFNFSGESNGTSIPGVPDFGSFLPNLGNLTQQFGSSGAF
PQFSIPSWTNFTDAFSSIFPFFGNVQGGGFPPFG
>hyp55.3 AGAP005822
MKYRFALLTVLVAVVPRPSAALLPDFAVPARMTGVASGLDSSGNVVQNTARVLHLLDQFQSITESLYGLQTPPLYRLATG
LRVVLDSLVESGSPIFQGLSNAARLSTGNITTVFDGIRRSINGTIALHELHQSALNGTVLLLGTGGVHNISTVLHQLVRN
VANLTGALDEIEPAIVAIQEQLSRPTQAQVDGRYPRAAIRRLNGILLDYVNIGQATVPQINAVVNRIRMMDGFIGRLAA
VTGGLWSDLNRTLAAVNGSVYAGVQQLHTTARTIRNSFNATVLGTARKLKLFFLDDEELQLVARNASALLVKRLTNVT
GALDGLVNRTEVRLEGYSELVLNRTIAAVTDLAWRLALSVTSSVPGADICYARYNYEFDKIPRLIYGLVACGQGEART
LQSVGGGLSGALSLTQAQLDAETSNEYEQCLGGLVPGSSDVLKRQRIVCLRGAERFSGVLGDAVVDAQLSAFQALLQKELS
YSAERHRQCVRITSGHLMVAEVKRLWSSIANCLS
>hyp14.5 AGAP004883
MLCSRSTLYAVFAFIFVSVLHDAMALQCRNCLSSKSFAECEKMGEVVCNSTIVNQNHQAFLPDNPTLPQGNWTEFKCY
QLQVARLHQNNNTDSGLRGYARGCTFLHVNFCNGWTSTLNLTSCTSTSDNCEQHLPTVAPELTTTTTTTKPSSNGANVLT
WRNLSSLGCLSAVVLLLAAR
>hyp14.5-1 AGAP001174
MKLFLLSVFFFLASFAAQAALRCMQGTEIVLPSSDAEKDKDKVPEFTECSILTTGASAASLLVLKPNTATLPSTDYKCY
HLRIEESKATIVKGCYKAQNVCDGKFKVDTVREVFCSCDQDECNSAHRFASNWKILSLTLFTILCLFMK
>hyp13 AGAP003474
MKNVFFALLLVVLVCCLVSVQGENEIQNVVKRSIPLRQLILQHNAALDDSDSDSGSQ
>LysC1 AGAP007347
MKVFSTVLLAIVACCAVAEAKTFTGKCELAKALANNGIAKASLPDWVCLVQNESAFSTSATNKNKNGSTDYGFQINNKYW
CDSGYGSNDCKIACKNLLNDDITDDIKCAKLIHCRHGFNAWYGWKNHCNGKKLPNVSSCF
>Cec AGAP000693
MNFSKIIFVVLAVLLCSQTEAGRLKKLGKKIEGAGKRVFKAAEKALPVVAGVKALG
```

PRECURSORS WITHOUT SP

>hyp6.2_AGAP006495
APQVTEAPGNVGSTYSMPADIGRLATGATKLFGQFWNTGTRFGTELSRRTFDFLRVKK
>hyp17_AGAP000151
DPLPGQTQRTLGYRGNDKRATPPMHSLGSGARMAMTGGGILGGIFSAL
>hyp15_AGAP000152
DPLPGRDRNTIANKSKDKKASAPKHS�GTGARMALTGGGVLGGVLTNM
>hyp12_AGAP008306
NDPVDALGACSGNLFGLLMTRLQQMVEDFTACRQEATANDPQHDCSDSIQRAKVDLQQQLVNYSYCTKNIQ
>hyp10_AGAP008307
TVNAEDPRTELIGCGSVLFHLLAANRLSLQLEEFVCKRSNPGYDCSDSIHRAISDLQQGLFDLNHCTKDIR
>sg2a_AGAP006504
LPAHHHSRGGDGSSANSTGQSYNNSGSPNNSAGVPDFGFNSQSNVPGFGNGQQPGQQQQGQGQGFPPFFGQGQSGFPSFG
NRQPFFGQNQQGQDNGNQQGRGIPFFGQGGGQGGIPSGSGQQNGGIPSLGNGQAQSGFPSFGNGQQGGNFPFFG
>sg2_AGAP006506
IPTSFNYGGGGGYFINGTGQSFNFSGESNGTSIPGVPDFGSFLPNLGNLTQQFGSSGAFPPQFSIPSWTNFTDAFSSIFP
FFGNVQGGGFPPFFG
>hyp55.3_AGAP005822
LPDFAVPARMTGVASGSLDSSGNVVQNTARVLHLLDQFQSITESLYGLQTPTLYRLATGLRVVLDLSDVESGSPIFQGLSN
AARLSTGNITTVFDGIRRSINGTIALHELHQALNGTVLLLTGGVHNISTVLHQLVRNVANLTGALDEIEPAIVAIQEQ
ELSRPTQAQVDGRYPRAAIRRLNGILLDYVNIGQATVPQINAVVNRIRMMDGFIGRLAAVTGGLWSDLNRTLAAVNGSVY
AGVQQRLHTTARTIRNSFNATVLGTARKLKLFFLDLLEELQLVARNASALLVKRLTNVTGALDGLVNRTEVRLEGYSEL
VLNRTIAAVTDLAWRLALSVTSSVPGADICYARYNYEFDKIPRLIYGS�VACGQGEARTLQSVGGGLSGALSLTQAQLDA
ETSNYEQCLGGLVPGSSDVLKRQIRIVCLRGAERFSGVLGDAVVDALSAFQALLQKELSYSAERHRQCVRTSGHLMVAEV
KRLWSSIANCLS
>hyp14.5_AGAP004883
LQCRNCLSSKSFAECEKMGEVVVCNSTIVNQNHQAFLPDNP TL PQGNWTEFKCYQLQVARLHQNN TDSGLRGYARGCTFL
HVNFCNGWTSTLNL TSCSTCTSDNCEQHLP TVAPEL TTTTTTKKPSSNGANVLTWRNLSSLGCLSAVVL LLAAR
>hyp14.5-1_AGAP001174
LRCMQGTEIVLP SDDSAEKDKDKVPEFTECSILTTGASAASLLVLKPNTATLPSTDYKCYHLRIEESSESKKATIVKGC IY
KAQNVCDGKFKVDTVREVFCSQCDQDECNSAHRFASNWKILSLTLFTILCLFMK
>hyp13_AGAP003474
NEIIQN VVKRSIPLRQLILQHNALDDSDSDSGSQ
>LysC1_AGAP007347
KTFGKCELAKALANNGIAKASLPDWVCLVQNESAFSTSATNKNKNGSTDYGIFQINN KYWCDSGYGSNDCKIACKNLLND
DITDDIKCAKLIHKRHGFNAWYGWKNHCNGKKLPNVSSCF
>Cec_AGAP000693
GRLKKLGKKIEGAGKRVFKAAEKALPVVAGVKALG

MATURE PEPTIDES

>hyp6.2_AGAP006495
LATGATKLFQGFWNTGTRFGTELSRRTFDFLRVKK
>hyp17_AGAP000151
DPLPGQTQRTLGYRGNDKRATPPMHSLGSGARMAMTGGGILGGIFSA
>hyp15_AGAP000152
DPLPGRDRNTIANKSKDKKASAPKHS�GTGARMALTGGGVLLGGVLTNM
>hyp12_AGAP008306
NDPVDALGACSGNLFGLLMTRLQQMVEDFTACRQEATANDPQHDCSDSIQRAKVDLQQQLVNYSYCTKNIQ
>hyp10_AGAP008307
TELI GCGSVLFHLLAANRLSLQLEEFVCKRSNPGYDCSDSIHRAISDLQQGLFDLNHCTKDIR
>sg2a_AGAP006504
GIPFFGQGGGQGGIPSGSGQQNGGIPSLGNGQAQSGFPSPFGNGQQGNFPFFG
>sg2_AGAP006506
IPTSFNYGGGGYFINGTGQSFNFSGESNGTSIPGVPDFGSFLPNLGNLTQQFGSSGAFPPQFSIPSWTNFTDAFSSIFP
FFGNVQGGGFPPFFG
>hyp55.3_AGAP005822
LAAVTGGLWSDLNRTLAAVNGSVYAGVQQRLHTTARTIRNSFNATVLGTARKLKLFFLDDLEELQLVARNASALLVKRLT
NVTGALDGLVNRTEVRLEGYESELVLRNTIAAVTDLAWRLALSVTSSVPGADICYARYNYEFDKIPRLIYGSILVACGQGE
ARTLQSVGGGLSGALSILTQAQLDAETSNYEQCLGGLVPGSSDVLKRQRIVCLRGAERFSGVLGDAVVDAQLSAFQALLQK
ELSYSAERHRQCVRTSGHLMVAEVKRLWSSIANCLS
>hyp14.5_AGAP004883
LQCRNCLSSKSFAECEKMGEVVVCNSTIVNQNHQAFLPDNPTLPQGNWTEFKCYQLQVARLHQNNNTDSGLRGYARGCTFL
HVNFCNGWTSTLNLTSCTCTSDNCEQHLPTVAPELTTTTTTKKPSSNGANVLTWRNLSSLGCLS AVVLLLAAR
>hyp14.5-1_AGAP001174
LRCMQGTEIVLPSDDSAEKDKDVPEFTECSILTTGASAASSLVLKPNTATLPSTDYKCYHLRIEESSESKKATIVKGCY
KAQNVCDGKFKVDTVREVFCSCDQDECNSAHRFASNWKILSLTLFTILCLFMK
>hyp13_AGAP003474
QLILQHNLDDSDSDSGSQ
>LysC1_AGAP007347
KTFGKCELA KALANNGIAKASLPDWVCLVQNESAFSTSATNKNKNGSTDYGFQINNKYWCDSGYGSNDCKIACKNLLND
DITDDIKCAKLIHKRHGFNAWYGWKNHCNGKKLPNVSSCF
>Cec_AGAP000693
GRLKKLGKKIEGAGKRVFKAAEKALPVVAGVKALG