

Supplementary material for:

Ancient eukaryotic origin and evolutionary plasticity of nuclear lamina

Ludek Koreny and Mark C. Field

School of Life Sciences, University of Dundee, Dundee, DD1 5EH, UK.

Supplementary text

Mutagenesis experiments on *P. infestans* lamin: While the lamin of *S. goreau* localized preferentially to the nuclear periphery in the human HEK293T cells, the filaments of the *P. infestans* lamin were spread throughout the nucleus with no clear perinuclear enrichment, suggesting little or no anchoring to the NE (Fig. 3). Therefore, additional sequence features are likely required for NE association and the *P. infestans* lamin is not fully compatible with the mammalian system. One possible explanation is that the prenylation signal (CaaX motif) of *P. infestans* lamin is not recognized by human farnesyltransferase and the posttranslational modification for increasing hydrophobicity of the C-terminus and attachment to the INM is consequently not taking place. The CaaX of *Phytophthora* possesses methionine at a2, which is uncommon (Supplementary Table S1, Supplementary Material online). However, mutation of the CaaX sequence to the more conventional CAIM of human lamin B1 had no impact on the localization. We also considered the possibility that *P. infestans* lamin could be processed by Zmpste24 protease since, within its sequence, we identified a hypothetical cleavage site with properties similar to that contained in pre-laminA, a natural target of Zmpste24 (Barrowman et al. 2012. PLoS One 7:e32120) (Supplementary Fig. S3, Supplementary Material online). We expressed *P. infestans* lamin with a non-cleavable mutation at this site as well as a truncated version with a premature stop codon immediately downstream of the hypothetical Zmpste24 site. None of these mutations significantly altered the localisation of the lamin fusion proteins. SDS-PAGE and western blot of the isolated wild type fusion protein failed to reveal the predicted processed form (Supplementary Fig. S3, Supplementary Material online), suggesting that, despite the presence of a likely functional prenylation signal in the *P. infestans* lamin, additional sequence features are required for NE association in the mammalian nucleus.

Supplementary figure legends

Fig. S1. Distribution of NUP-1 and NMCP proteins. (A) Phylogenetic trees of NUP-1 and NMCP proteins. Searches for NUP-1 and NMCP identified orthologs restricted to Trypanosomatida and Streptophyta, respectively. Phylogenetic relationships are based on analyses of protein sequences (Supplementary Table S1, Supplementary Material online). The edited alignments used for the phylogenetic analyses are below in the Supplementary

Material online. The numbers above branches are PhyML bootstrap/Bayesian posterior probability values. (B) Secondary structure of NUP-1 and NMCP proteins. Predicted nuclear localization signals (NLS) are marked by blue rectangles. The central portion of NUP-1 orthologs consists of repeats of variable size and number. The presence of direct repeats precludes unambiguous assembly of sequence data for many orthologs, and repeat numbers for *P. serpens* and *S. culicis* were estimated based on repeat regions assembled into several contigs, while repeat regions are absent for NUP-1 of *P. confusum* and *H. muscarum*. The four NMCP paralogs of *A. thaliana* are the result of three duplications in the plant lineage (red dots) in the ancestors of Spermatophyta (seeded plants), Eudicots and Brassicales.

Fig. S2. Sequence alignment of lamin protein sequences from representatives of major eukaryotic groups. Coiled-coil regions were predicted by Marcoil and Pcoils tools and are highlighted in green; the parts recognized as filament domains by NCBI CD-search are highlighted in darker shade of green and LTD domains in dark yellow. The light yellow delimits the regions that match the Ig-like LTD fold predicted by Phyre2. The red box marks the CDK1 phosphorylation site consensus sequence, the blue box marks the classical monopartite NLS, and the CaaX motif is in the orange box. The first and fourth positions of the predicted heptad pattern in the filament domain are marked by letters a and d above the sequence. Where they match hydrophobic and uncharged amino acids favourable for the coiled-coil formation, they are in a black box, while a letter in a white box marks the less favourable amino acid. The heptad-repeat interruptions and gaps within the coiled-coil region are marked by red dots and their hypothetical positions were estimated based on both Marcoil and Pcoils predictions and the level of conservation of the respective regions in the alignment. The secondary structural elements in the LTD domain were predicted by Phyre2 and are in grey. The arrows indicate beta sheets and spirals alpha helices. The known secondary structure of the human lamin B1 is shown using red symbols for comparison.

Fig. S3. Comparison of C-terminal regions of human lamins with lamin homologs of Oomycetes. (A) Amino acids with similar properties surrounding the cleavage site of human prelamin A and hypothetical cleavage site of *P. infestans* lamin (arrow heads) are highlighted in red. The cleavage of prelamin A by Zmpste24 protease results in loss of the hydrophobic C-terminus and mature lamin A therefore localizes to the nuclear interior. A-type lamins are unique to vertebrates, suggesting that there would be no selection pressure for other eukaryotes to avoid a Zmpste24 cleavage site (Barrowman et al. 2012. PLoS One 7:e32120). Additional to several amino-acid groups surrounding the cleavage site, the spacing from the farnesyl-cysteine to the cleavage site is important for optimal cleavage (Barrowman et al. 2012. PLoS One 7:e32120). We noticed that the C-terminal region of *P. infestans* lamin contains F-L (500-501) sequence 10 amino-acids upstream of the farnesyl-cysteine, similar to the Y-L (646-647) cleavage site of human prelamin A. Furthermore, arginine is present in the -4 position of the putative F-L cleavage site of *P. infestans* lamin, while the presence of arginine in the -3 position of the cleavage site in human

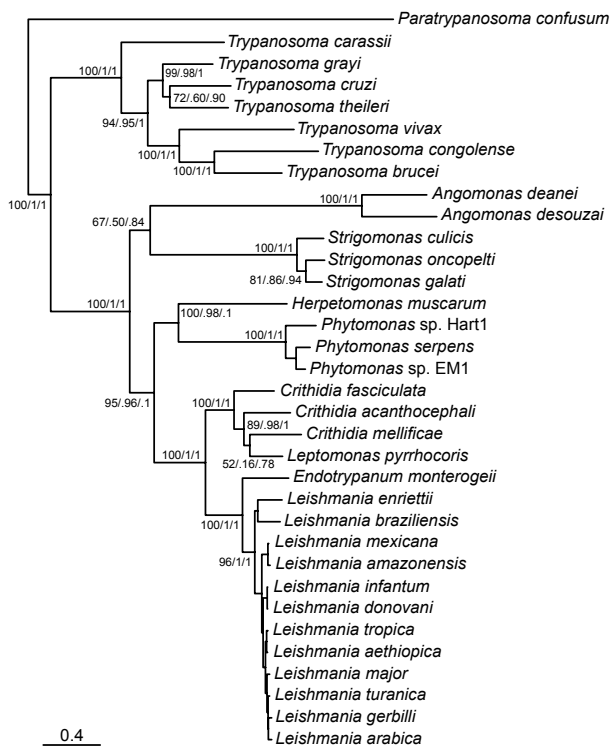
prelamin A is essential for cleavage. The L647R mutation in human prelamin A effectively abolishes cleavage by Zmpste24 (Barrowman et al. 2012. PLoS One 7:e32120). We therefore created similar mutations in the *P. infestans* lamin, i.e. L501R and L501M (methionine is the most common amino-acid in that position in other *Phytophthora* species), and a version containing a STOP codon after F500. The intranuclear localization of all mutated lamins was indistinguishable to the wild-type lamin. (B) Western blot detection of EGFP-Lamin constructs in lysates of HEK293T cells. The wild-type (WT) lamin of *Phytophthora infestans* is compared with mutagenised lamins. There is an apparent shift in size of the lamin L501Stop compared to the wild-type lamin and the L501R and L501M mutations that were all detected as proteins of the same size by a western blot, which clearly demonstrates that the lamin of *P. infestans* is not cleaved by the human Zmpste24. The constructs were detected using Roche mouse monoclonal anti-GFP antibody, version 06 (1:1000) and secondary goat anti-mouse HRP conjugated antibody (1:10000). Samples were separated alongside the Prestained Protein Marker, Broad Range (New England Biolabs; 7-175 kDa). The calculated theoretical molecular weights (kDa) of the eGFP-lamin constructs for lamins of *Symbiodinium goreau* (S) and *P. infestans* (P) are shown next to the protein bands.

Fig. S4. Specificity in searches for lamins. E-value distribution of the top blastp hits in representative genomes/transcriptomes of eukaryotic taxa that possess lamin homologs identified in this study (A) and in representative genomes of related taxa, where no lamin homologs were found (B). The lamin protein sequences of metazoa, namely of *Homo sapiens*, *Branchiostoma floridae*, *Priapulus caudatus*, *Hymenolepis microstoma*, *Aplysia californica*, *Drosophila melanogaster*, *Daphnia pulex*, *Nematostella vectensis* and *Trichoplax adhaerens* were used as queries to search in protein or predicted peptide databases of selected species (supplementary table S1, Supplementary Material online). The accession numbers and annotations of the hits are listed in the tables beside the e-value distribution charts. There is a substantial difference between the e-values of the 1st hit for each taxon in A - the actual lamin orthologs (surrounded by a dashed line) and all the other hits that mostly represent various coiled-coil proteins that are clearly distinct from lamins.

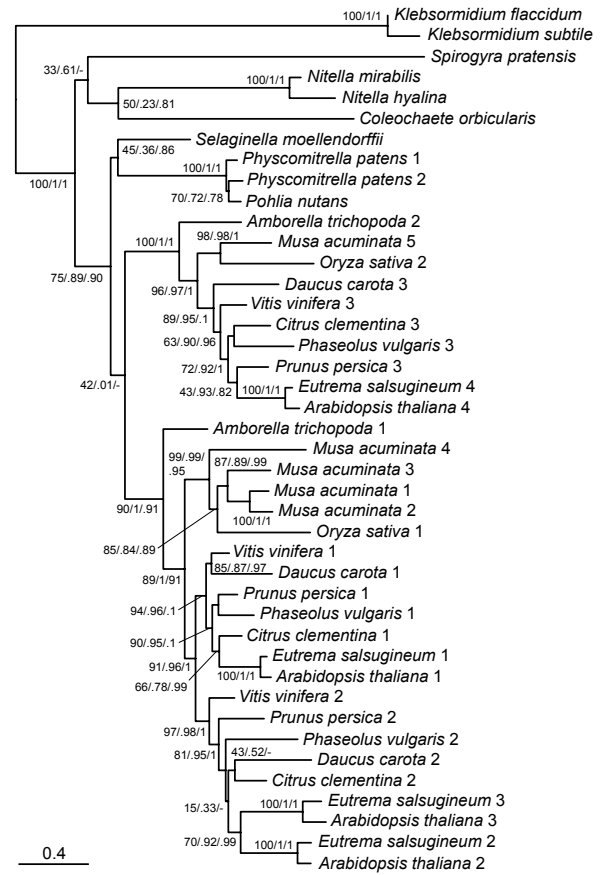
Fig. S1

A

NUP-1

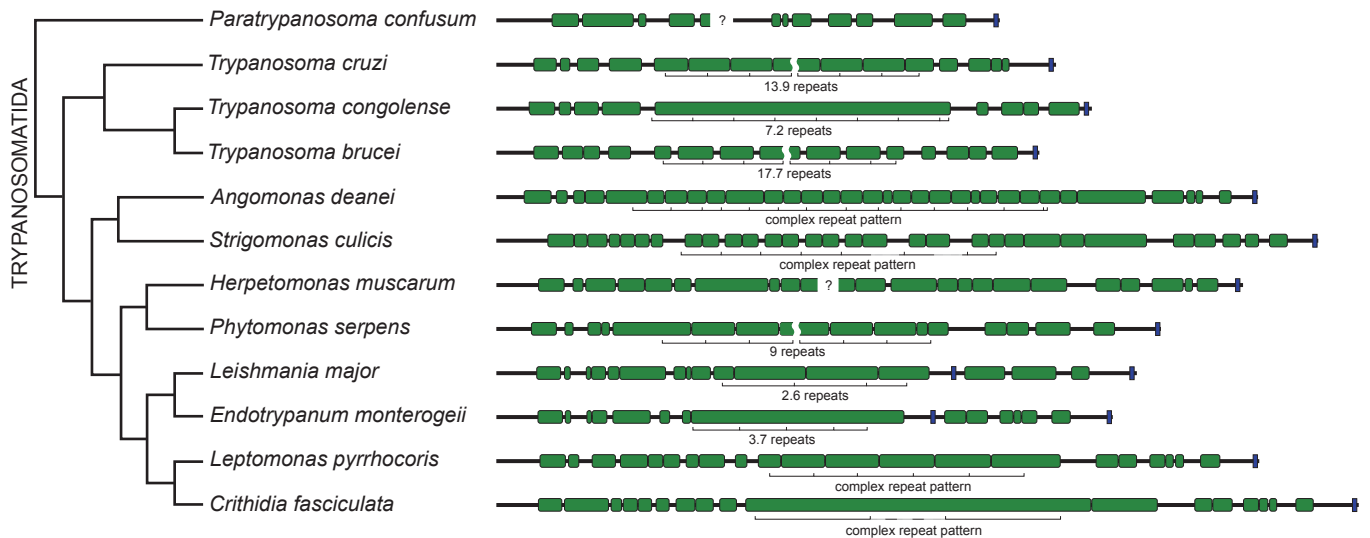


NMCP

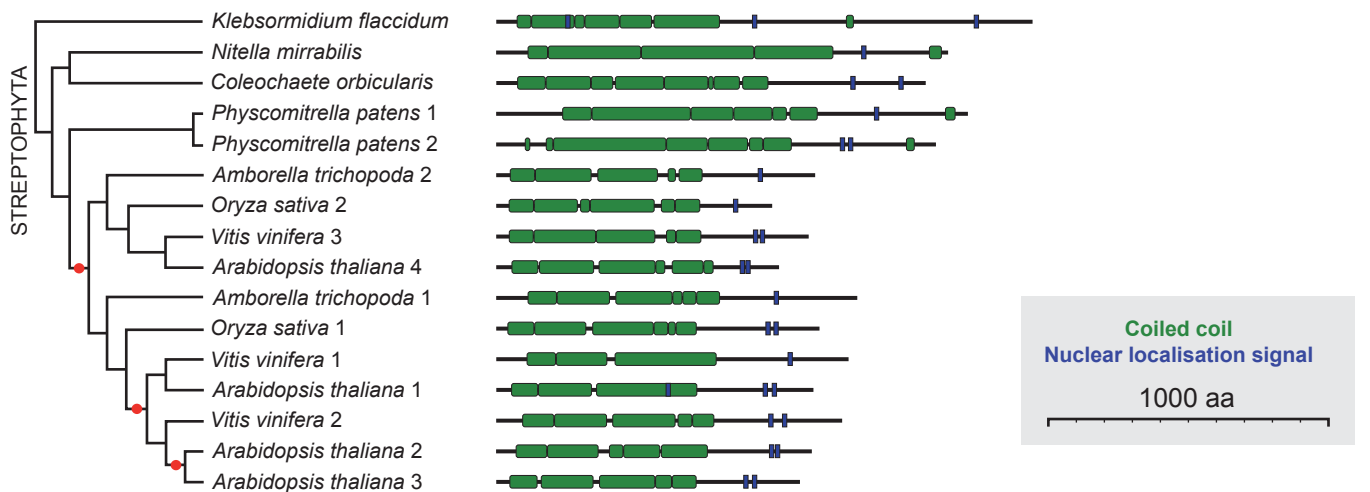


B

NUP-1



NMCP



Coiled coil
Nuclear localisation signal
1000 aa

Fig. S2

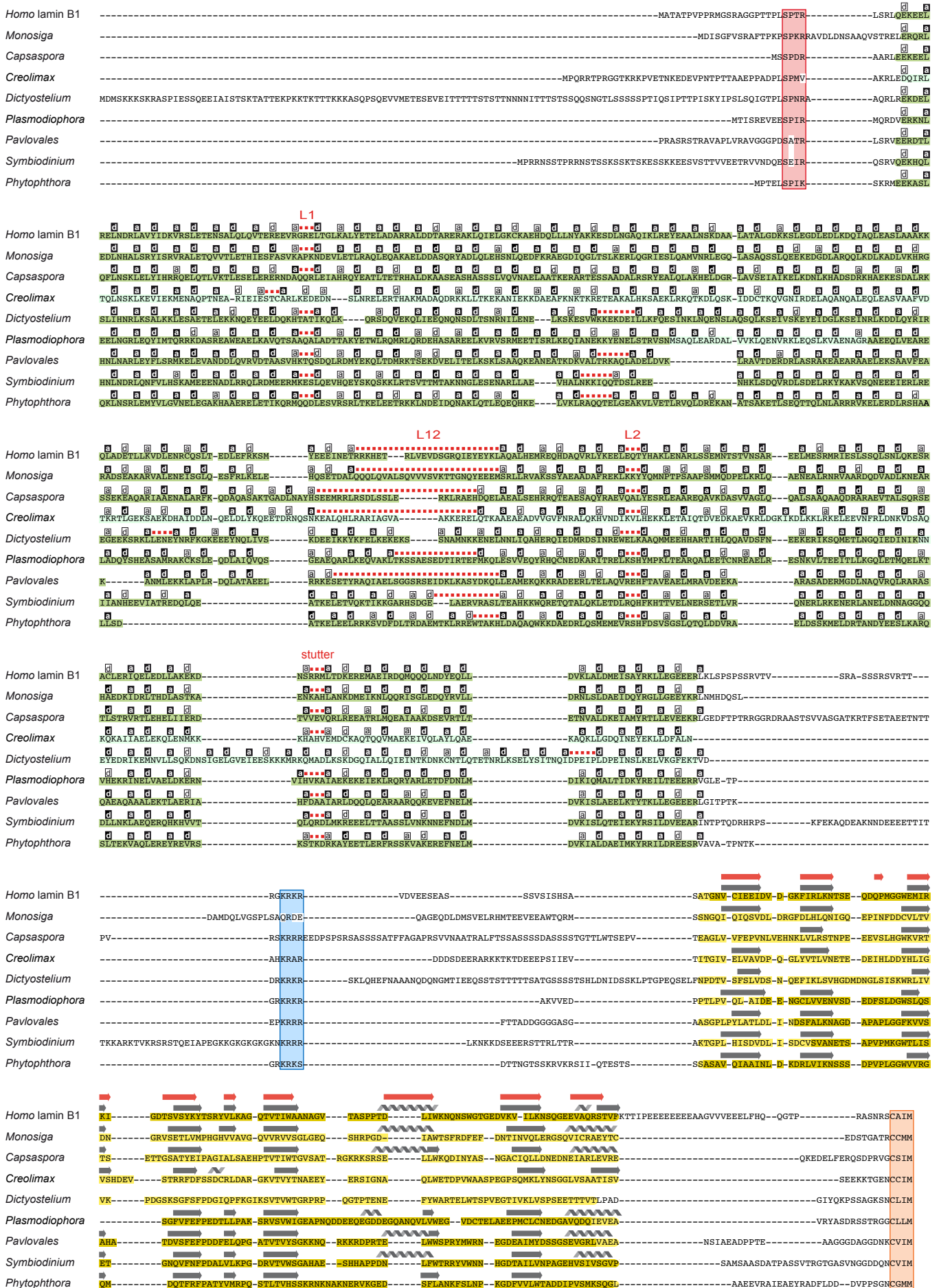


Fig. S3

A

Homo lamin B1	KVILKNSQGEVAGRS	-----TVF	---KTTIPEE	-EE-EE-EEA	-AGVVVEELFHQQGTP	--RASN	-----RSCA	IM
Homo lamin B2	RTVLVNADGEVAMRT	-----VKK	---SSVMREN	-EN-GEEEEEE	-AE-FGEEDLFHQGGP	--RTTS	-----RGCVIM	
Homo prelamin A	RTALINS TGEVAMRK	-----LVR	---SVTVVED	-DE-DEGGDL	-----LHHHGHSCSSGDPAEYNLR	LRRTVLCGTCGQPADKASASGSAQVGGP	ISSGSSASSVTVTRSYR	VGGSGGSFGDNLVTRSYLLGNSSPRTQSPNCSIM

Oomycetes	<i>Phytophthora infestans</i>	FVVLWTTADD	IPVSMK	SQGLAAEEVRA	IEAEYRADFLD	--VPPS	-----GNCGMM	
	<i>Phytophthora capsici</i>	IVFLVTDD	IPVSMSEGLSEEEVRA	IEADLGVDLMD	-DE--APPS	-----GNCGMM		
	<i>Phytophthora parasitica</i>	FVVLVNSDD	IPVSMSEGLPAEEVRA	IEAELRADFMD	-DE--VPPS	-----GNCGMM		
	<i>Phytophthora cinnamomi</i>	CVILITSD	IPVSMKSEGLPDDEVRA	IEAEFRADFMD	-DE--APPS	-----GNCGMM		
	<i>Phytophthora sojae</i>	FVVLSSSE	IPVSMKSAGLPDEVRA	IEADMRAEFMD	-DE--APPS	-----GNCGMM		
	<i>Phytophthora lateralis</i>	FVILVTSDD	IPVSMSEGLPEEEVRA	IESELRADFMD	-DE--APPS	-----GNCGMM		
	<i>Hyaloperonospora arabidopsidis</i>	FIVLVTSDD	IPVSMKSEGLSEEEARA	IEADLRADYTD	-DE--APPS	-----GNCGMM		
	<i>Pseudoperonospora cubensis</i>	FVVLTSDD	IPVSMKSEGLPEEEVRA	IEADLRADFMD	-DE--VPPS	-----GNCGMM		
	<i>Pythium ultimum</i>	FAVLLTAD	GVVPSVTRAEGLSAEEVNTLEAE	IKADLDD	-E--APST	-----EGCGIM		
	<i>Pythium arrhenomanes</i>	FAVLLTNE	GVVPSVSMKGEGLPAEIKKLEQEVFAE	LDD	-D--APAA	-----EGCGLM		
	<i>Pythium irregulare</i>	FAVLLTAE	GVVPSVTRAEGLSAEEVKALEAE	IKADLDD	-D--APPG	-----EGCGIM		
	<i>Pythium neoyamsii</i>	FAVLLTAD	GVVPSVTRAEGLSAEEVKALEAE	IKADMDD	-D--APPG	-----EGCGIM		
	<i>Pythium vexans</i>	LAVLTTD	GVVPSVLCSEGMSAADVKKFQDE	IRNDFED	-D--OPPS	-----EGCGIM		
	<i>Albugo labachii</i>	WVLYDDE	GVVPSVSLAQLPKVEEVALEAA	IRTESPE	-DDFKEGSS	-----DACWIM		
	<i>Albugo candida</i>	WVLYDSD	GVVPSVSLAQLPKVEEVALEAA	IRSESEPE	-DDFEEGSS	-----DACWIM		
<i>Bremia lactucae</i>	FVVLMTSDD	IPVSIKSEGLPEDEVRA	IEADLRADFMD	-DE--APPS	-----GSCALM			
<i>Eurychasma dicksonii</i>	IAEVRDA	KGE LVVYKYYGDISDPY	TEETDGA	IPGE	-----TACCVIM			
<i>Saprolegnia diclina</i>	VALLVKNAE	GEVVCSHAEGLNEDEYDDEQEK	-----EGNN	-----EGCGIM				
<i>Saprolegnia parasitica</i>	VALLVKNAE	GEVVCSHAEGINEDEYDDEQEK	-----EGNG	-----EGCGIM				
<i>Aphanomyces invadans</i>	VAQLTNP	SQDVVSSYAEAGMYVDDDDVDAAD	-----TPAK	-----DGCGIM				
<i>Aphanomyces euteiches</i>	VAQLTNP	AGDVVSSYAEAGMYVDDDDVDAAD	-----TPVK	-----EGCGIM				

B

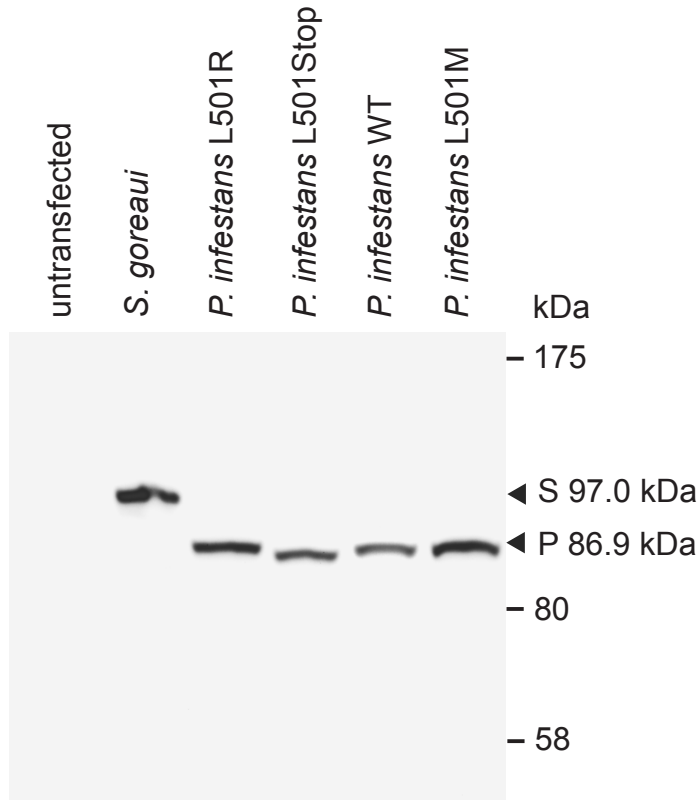
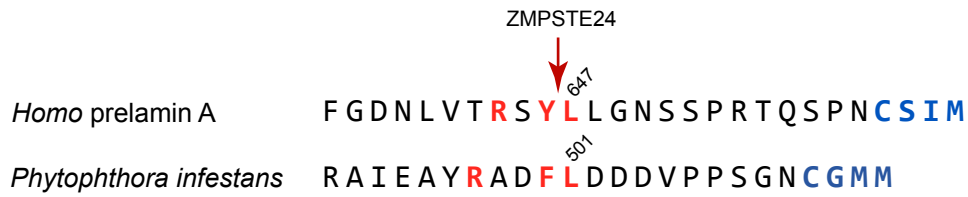
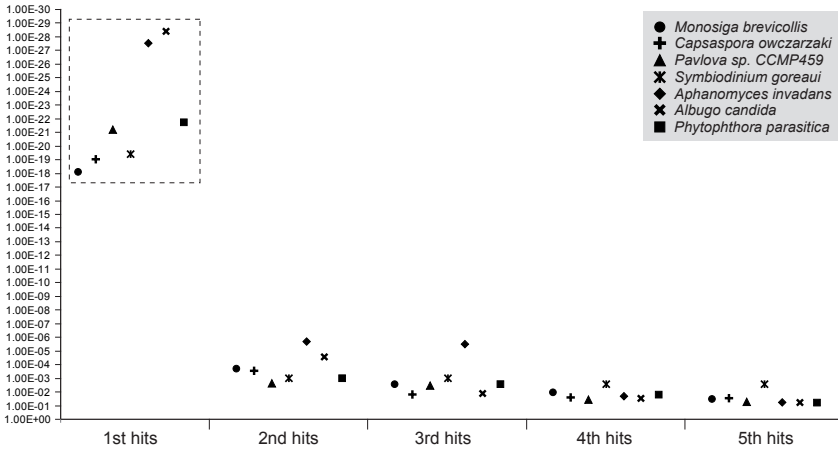


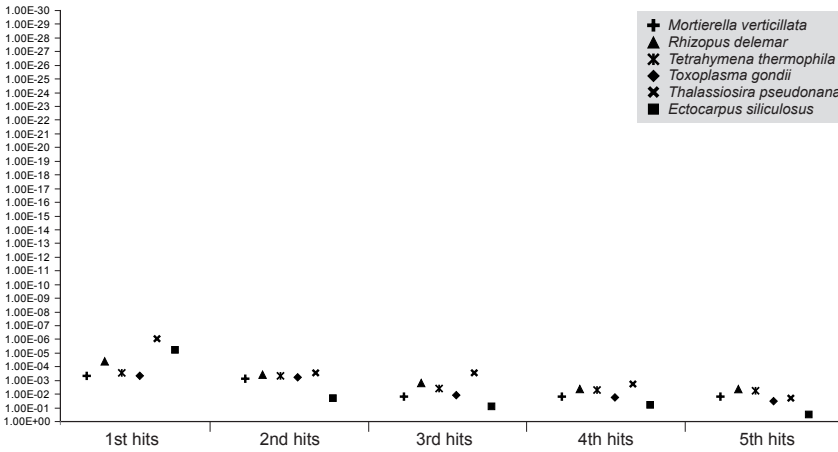
Fig. S4

A



Genome	e-value	Accession nb.	BLAST HIT Annotation
<i>Monosiga brevicollis</i>	8E-19	XP_001745977	Lamin
	2E-04	XP_001750258	SH3 domain-containing coiled-coil protein
	0.003	XP_001742372	RNR_PFL domain-containing coiled-coil protein
	0.011	XP_001745137	Calponin homology domain-containing coiled-coil protein
	0.033	XP_001746167	hypothetical coiled-coil protein
<i>Capsaspora owczarzaki</i>	1E-19	XP_004365259	Lamin
	3E-04	XP_004344350	cenpe protein
	0.017	XP_004346859	cut-like homeobox 1 (CASP)
	0.028	XP_004345230	unconventional myosin-XVIIa-like
<i>Pavlova sp. CCMP459</i>	0.031	KJE95785	pseudouridine synthase domain-containing coiled-coil protein
	8E-22	190508862	Lamin
	0.003	190510564	Choline transporter domain-containing protein
	0.004	190525724	dynein heavy chain 7
	0.047	190503290	outer dynein arm-docking complex protein-like
<i>Symbiodinium goreau</i>	0.062	190525808	LIM and Calponin homology domain-containing protein
	4E-20	199596768	Lamin
	0.001	199636648	DnaJ-domain-containing protein
	0.001	199640602	ubiquitin-like domain containing coiled-coil protein
	0.003	199570084	Leucine Rich Repeat family protein
<i>Aphanomyces invadans</i>	0.003	199596170	fas-binding factor 1-like
	3E-28	XP_008867452	Lamin
	2E-06	XP_008870094	TNFR domain-containing coiled coil protein
	3E-06	XP_008873497	WWP domain-containing coiled-coil protein
	0.031	XP_008877304	hypothetical coiled-coil protein
<i>Albugo candida</i>	0.075	XP_008869784	hypothetical coiled-coil protein
	4E-29	CC110441	Lamin
	3E-05	CC150727	putative spindle assembly protein
	0.013	CC148786	hook protein-like domain containing coiled-coil protein
	0.047	CC141228	Leucine rich repeats-containing coiled-coil protein
<i>Phytophthora parasitica</i>	0.072	CC141579	TPR/Mlp nucleoporin
	2E-22	XP_008905021	Lamin
	0.001	XP_008895485	plectin-like protein
	0.003	XP_008902902	hypothetical coiled-coil protein
	0.019	XP_008902191	WD repeat-containing protein 65
0.081	XP_008905554	coiled-coil domain-containing protein 57-like	

B



<i>Mortierella verticillata</i>	5E-04	MVEG_05989	myosin 9-like protein
	8E-04	MVEG_07393	hypothetical coiled-coil protein
	0.016	MVEG_00433	Zuo1/DnaJ subfamily C member 2
	0.017	MVEG_00637	tuberous sclerosis 1-like
	0.017	MVEG_04666	glutaredoxin domain-containing coiled-coil protein
<i>Rhizopus delemar</i>	5E-05	EIE92276	hypothetical coiled-coil protein
	5E-04	EIE76308	PT repeat-containing coiled-coil protein
	0.002	EIE85364	P-loop containing nucleoside triphosphate hydrolase protein
	0.005	EIE90623	GRIP domain-containing coiled-coil protein
<i>Tetrahymena thermophila</i>	0.005	EIE81526	Putative Dynein heavy chain
	3E-04	XP_001015907	golgin-like coiled-coil protein
	5E-04	XP_001008452	FYVE zinc finger-containing coiled-coil protein
	0.004	XP_001014796	hypothetical coiled-coil protein
	0.005	XP_001009123	EEIG1/EHBP1 protein amine-terminal domain protein
<i>Toxoplasma gondii</i>	0.006	XP_001032362	hypothetical coiled-coil protein
	5E-04	EPT28173	hypothetical coiled-coil protein
	6E-04	XP_002371471	RCC1 domain-containing protein
	0.013	XP_002366795	hypothetical coiled-coil protein
	0.019	EPT31185	surface antigen repeat-containing protein
<i>Thalassiosira pseudonana</i>	0.035	EPT29827	RCC1 domain-containing protein
	1E-06	XP_002294807	TPR/Mlp nucleoporin
	3E-04	XP_002286197	plectin-like protein
	3E-04	XP_002291803	hypothetical coiled-coil protein
<i>Ectocarpus siliculosus</i>	0.002	XP_002287281	hypothetical coiled-coil protein
	0.022	XP_002294067	hypothetical coiled-coil protein
	6E-06	CBN77699	TNFR domain-containing coiled coil protein
	0.021	CBJ48665	hypothetical coiled-coil protein
	0.081	CBJ48925	structural maintenance of chromosomes protein 6
0.064	CBJ25580	WD repeat-containing protein 65	
0.31	CBN74724	kinetochor protein NUF2	

Table S1. Protein sequences used in this study.

Species	Group	Source	Sequence
LAMINS:			
Homo sapiens	Metazoa	NCBI AAC37575	MATATPVPPRMGSRAGGPTTPLSPTRLRLQKEEELRELNDRLAYID-KVRSLETENSALQLQVTEREEVGRGLTGLKALYETELADARRALDDTARERAK-LQIQLGKCKAEHDQLLNYAKKESDLNGAQIKLREYEAALNSKDAALATALGD-KKSLGDLKDDQIAQLEASLAAAKQLADETLLKVDLENRCQSLT-EDLEFRKSMYEEIEINERTRKHETRLVEVDSGRQIEYIEYKLAQALHE-MREQHDAQVRLYKEELEQTYHAKLENARLSSEMNTSTVNSAREELM-ESRMRIESLSQLSNLQKESRACLERIQELEDLLAKEKDNRRMLTD-KEREMAEIRDQMQQLNDYEQLLDVKLALDMEISAYRKLLEGEEERL-KLSPSPSSRVTVSRASSRSRVTRTRGKRKRVDVEEASSSVSISHS-ASATGNVCIIEIDVDGKFIKNTSEQDQPMGGWEMIRKIGDTSVSY-KYTSRYVLKAGQVTIWAANAGVTASPPDTLWKNQNSWGTGEDVK-VILKNSQGEVAQRSTVFKTTIPEEEEEEEEAAGVVVEELFHQQGTPRASNRSCA-IM
Nemato-stella vectensis	Metazoa	NCBI XP_001629288	MATATKSPASSSSTPKTPVSSSRIMGSSPPGSAKFRQAQKAEQLHNLNDR-LATYIDRVKLNLEQENSKLRSEVTVSRKTVREVDMSKSLYETELADARRLLDE-TAKEKAKQIESSKNSNDAQEFKNKFDKEAAARKKAEKELNDRKLLHDKENQL-TRRNQALNLESVLRQGECEELKDALAAKYALEQETLTVRDLNRCQSLQ-EEQNFKQMYDKELSDIRSQKTVETKRVVVEVDYKDYKYEGLMAEKLQELRED-YDSDARSFKEETELLYSSKFEELRIQREERDSEALAKLREENRNLKSKVDLSSQ-VHOLEAKNNALVSRVSDQLQGLRAQDKKHDNEILLRENEIAELRTSIDDALRDY-EDLMGVKVALDMEITAYRKLLESEETREITPPASPVLGGSSVSQTRRGNKRA-RTEETESTMTTTTTAEGAQIFTEADPDGKIKIYNSGEKDEALGGWTIQRQVGT-EDPSVYKFTPKYVLSKQSHVTVWSAQGGGTHKPPSDLVFKQLPSWGSNGEAR-TALVNAAGGEEMATLLEEKVFQHYSTDIVDSGRRGRGRDRGEVAKGCAVM
Trichoplax adhaerens	Metazoa	Krüger et al. 2012. Mol Biol Cell. 23:360-70.	MSPLKIKRVQKEEELKNLDRLASVIEKSRFLSERNARLNEEIRSSRRSGDENIKSL-KALYEQELQEARKTIDEMANEKSKITITVMKYTKQVEELTSDLEKVKASKSSLQ-VALNNAEKRATELEGEVNIERNKVAELREVIGHKEKIQRSSEELASANKLAE-QQALRRVQIENEYQTLKESAEFSTEMHQEIRSLKKNLKTVDQDQ-STLQDDFQSEYDTKMTALQQLRRDNEENSKRLKOEVEELYSQ-VKELSQERDRTKIITELKNECRKFRDKSESISSEVHVLRAENSSLQT-RKRELESJIEEROKNTNKFEEYEMEMRALQLKEEQSTYEYKELLDV-KIELDNEIAAFRKLLEGEETRLSSPSPFAQRKRRIEDSDYIVSS-HATGDVQISDDVKHGLYVRLNNTSKTDCHIGNFVIKHQVDNKNIEISY-KFNTKAILKGNVTTVMAADSGQSHKPPANYLWKKQNNFGLGEE-MVTRLVNAGGEEIAIFNLTKAQEVEIKDELFSYQVKDGTGNRCCIQ
Branchio-stoma floridae	Metazoa	NCBI XP_002586730	MSRRSQKLTQKTVVVTTTTSSVSTPKVTKQVQEAFTVQLSPTRLTRM-QEKQELQWLNDRLAQYIDRVRYLEAENSRLMVQVTSSEITQREVTNIKSMFE-QELTDARLLDDTAKEKARVQIEAGKYRAEDELRAKLAKSEGALATAEKRRHQAE-SALNEKEGRLQNAIADKQRAEGELAAALRLEFANLEKQLATARKQLVEEETLLRVDLE-NKVTQLREEVEFNKQVYEQELTESRTRKVEVITVEVDAGEAAAFESRLQALREMRQ-QHELEARSMPRELETMYTKITSMRTDSEKSNALSVAREEMRESRARIDSLMSQ-VSGLQGNASLEARCKLEGMARMSEESRSSAEQYEREIQRLEEIQSMMVYD-QELMDIKIALDLEISAYRLLGEEQRLKLTPTSPPTASNHSISHDCRGRTSKRKR-VEEESGSSSSGAVAVGGVDLEGKFKVQLNNTSAEKDMSMGWLLKRTVLVGGGEE-ISYKFPSSRYLKAQGSVTVWATEGGGTHSPPSDLLFRGQASWGSQDFTKTLVND-SGEEMATRSVTRVASSGFSSTSGGTFVEVGGQELFHQQGDPASPRQRCSIM
Priapulul caudatus	Metazoa	NCBI NP_999665	MASQTSQRYEKRSTRTEFKTTSTPKASSSSQAKSLLSPAKISRHEEKEELIGLNDR-LATYIDRVRHLEENGRLLVQITTFEETQTRDIEGIVKYEKELADARKLLDE-TAKEKAKLQIECGKYKTELDSLPRPVGKLEKELNAANKRVASLEAQVAEKD-VRIRLSLNDKRSLEDELNELKDLGNKEKQLKVAKKQVEEETLLRVDLENRLQSL-KEELSFKEQLFKEELRETRTKRQVDMSTIEGQSIIEELNNSLYESLQELREQTSEQ-TLLRQETESLYFSLADLKAALAEHRNAAVNSQDEVRKLRSTVDDLTSIETIRAQ-NNALIARIKDLEKQLRQDDHLEATLDRKELQQLRDVAEQLKDYDDLNLKLSL-DNEISTYRKLLEGEETRLRITTPPRKTSTGRDPRPTKRRRVEDESSVTRPTNN-GVVAIVESDPGENFVKLHNNSDQDALGGWHLKRSVDGGSQSYKFTAKYVLLK-AGQEVTVWASGSGKSHSPKDLVYKNVESWGTGDNVETSLLDASGEIMATR-TVIKEVTTHEYREGREDRCSLM
Daphnia pulex	Metazoa	NCBI EFX71199	MSTRSKKTPASSSQASSASQSRSSPLSPTRLRLQKEVQLNDR-LANYIDRVRLQLESENNRLTLQITTTQDTISREVTSVRGSYEKELSDTRQLLDE-TAKEKARLQLEAGRARAEELGELTPKYNAKNDLLQAEKVVQLETRNNELNASL-HQLKSDMARAIIEKKDAVSRDKLAKQVPELQQQLETELLARTDLENRNMTLK-EELSFQKQMYEREMTEVRSNKQVEISEIDGRLEQEQARLQSTLQELREQYEQ-SQIQQSRDEVEVLYQNKVDELDQVKKQGAAGAHYEELIQARNRLKDALNK-LSELELANSHMNDRLAEMERNMETERQTHATAMQDASAVYSSMCQGEITRL-REQMAIQLEQYQDLMDIRLALDMEISAYRKLLEGEEETRLKLTPTAGASTSSTR-SGTPTSRRTPTRALKRKRITMLESSSYSLADFITSSAKGEVQVEEVDSEGKFI-RLINKSEKESVLSGWQLVHKAQETETIYKFRSLKVPAGGTVSVWSAGSGETAHE-PPSTLVMEKQRWVFNEMVTLQLDNSGEEAQAQRESKRAQTSLSLRQREEY-RGLGGPEELHHQQGDPPEGKDRCVVM
Drosophila melanoga-ster	Metazoa	NCBI P08928	MSSKSRAGTATPQPGNTSTPRPPSAGPQPPPPSTHSQTSASSPLSPTRHSR-VAEKVELQNLNDRLATYIDRVRNLETENSRLTIEVQTTTRDVTRETTNKNIFEAE-LLETRRLDDTDARDRARAIDIKRLWEENEELKNKLDKKTKECTTAEGN-VRMYESRANELNNKYNQANADRKLLNEDLNEALKELERLRKQFEETR-KNLEQETLSRVDLENTIQLSREELSFKDQIHSQEINESRRIKQTEYSEIDGR-LSSEYDAKLLKQSLQELRAQYEEQMQINRDEIQSLYEDKIQRLEQAAARS-NSTHKSIEELRSTRVRI DALNANINELEQANADLNARIRDRLERQLDNDRE-RHGQEIDLLEKELIRLREEMTQQLKEYQDLMDIKVSLDLEIAAYDKLLVGE-EARLNIPTATNTATVQSFQSRLRNSTRATPSRRTPSAAVKRRAVDDES-EDHSVADYVVSASAKGNVEIKEIDPEGKFRVLFNKGSEEVAGVWQLQR-LINEKGPSTTYKFRHSVRIEPNGVITVWSADTKASHEPPSSLVKMSQKW-VSADNTRTILLNSEGEAVANLDRIKRVISQHTSSSRLSRRRSVTAVDGNELQYH-QQGDPPQSSNEKCAIM

Amphimedon queenslandica	Metazoa	NCBI XP_003386754	MSQQATSSPSLPTPERRTRMKEKEDLQLLNDRFVYINRVRRLRDEKEKLVNLTLEHMQTTTQEESGAIKRMFEKELQDARLLIDE TAKEKARYQIQSSKTEDRIKDLTELKEDFREDYDRMQQELAEAKKSAETFEYSLVNDSQAKQIQDLLEEEIDDLRARLREVQDSLEQETLAKVDIQNQNSLREELAFKFKYVDEELLTIKRTGLTTHVRDGVVEEDDDFAERLQAAIDEAREEIEKETE QFKLDLEKSYTRLESLESQSQRDASARVRLDDELKNCHTTLAKYNRDIKLEKNEQLESSLREKSEDLARAKQLHSAELANLMDERMALRVNYDQMKMEYELFDLRIQLEQEIALSALLQEEEQRLNLQTTTPREKRPRVARDGEGPRSKRKKQVAATSSAVGFIQIDVDPDGRFVQIKNMSDKVTEIPKNNKLQVGTAMLIASKYEEMAYEVADVFYITDSTYSNTEIKAMERNLKTLDYSFGNPLCLHFLRRNSRAGDATPQMHTMAKFLMELCLPDYMSLEYLVA-AAALYSNKLYSDGWEPTALRHSYQYTEPDVLPVCGVMASLVLSMHTAKQAVKNKYCSSKFMRIAKEPCLQGKIMKELASVLSQ
Hymenolepis microstoma	Metazoa	NCBI CD-J06754	MSARTKKSQSQPVAEKLATPKKVERSSTTISSPVQSSDAEYRSPSPLNISRVDEKEELAHLNDRLASVYDVRKLERDKESLTRISTITEERLAKVDNARSTYENIEQSLRLVDDLAKKEAARDVELKKHMDDASDAKSKLNKRDELNRNLRQKRYDTELERDIGTYKSDHERRYQKLLPEYEHLEKSLAAKEGLKAEITLRTDLENKVASLREELDFDKRLFEERESKLVMRSLTVEEEVAETKAKEYESRLADELQAYRESANELEEQYRQLQMESTFGNLEQLQKANEANRDGGRLNEELIAMRRMNDL SHELARKSREVELDHSRIADLEKLLNKAHDDYQAQLAEREVEKRLRAELERQFAEFTDLMNTKIVLDQEILMYRKMLEGESRLNLTSPKRHSVFGGFMGVKRRRLDSGDGDFDEESNSTGQSQTAFVSTSAVSGIEFSAQDQDGSNGWIKLSNVGKEDINIGNWVWKHKSDEGEVITYKFRNVILKPGTPTVWSVDAGATHDPPEDIVMKNQRFHAGANTVTLT-NENEEQASCLVRESRPKIIPRRFRAGAARGEDKCSLM
Aplysia californica	Metazoa	NCBI XP_005094526	MKQTRKTVTTTTSSSSGGGGGAGGEEAGTSFSSSSQSRSSGGLGRPPSPARLSRIQEQUELQGLNDRLAVIDKVRSLAEANSRLALQVRSSTEEVVKREVTSVKTLYESELAEELRRLDETSKEKARLQIEANKYKTDYEDLLAKYINRDRDANALERVQVLERQLAEFSARDFEKENQNLKNEELRLEQLAAAKKQLEEEETLVRDLQNRVKTLEKEDLHFKSQVYEQELEESRVTTTQLEEVQDRIEQIEEARLQDALLREIRAGHDYDLQSVKVELETLYESKVEDLRNQLERNQSSSAWELMLARRSVDLSAEIGRLKSENAGLDLRVKDIDQLERERDEFRLRMKORDEEISIRLRAQVELEIEYALALEIKIKLDREIEAYRKLLEETELNISASMTQORSMTTASSSGGGRKRKMDIQQAVIEYTORASSGFSSATSSCGVDISDVTDCGFVKLTNTTDDKDVHLFWQIRHVSQDNETRHKFKQKNVLRKAGQTTVWSASDSD-QTHNPPSGLVMKGRWFVSGDMTTTTLLDQDETEVASCSMSRNIGSVASVSTRRRGPRDELDEELQDKDKCFVM
Monosiga brevicollis	Choanoflagellata	NCBI compound of XP_001745977 and FE885114	MDISGFVSRAPTKPKSPKRAVDLNSAAQVSTRELERQRLDLNHALSRYISRVRALTEQVVTLETHIESFASVKAPKNDEVLTLRAQLQEAQKALD-DASQRYADLQLEHNSLQEDFKRAEGDIQGLTSLKERLQGRISLQAMVNRLEGQLASAQSSLQEEKEGDGLARQQLKDLKADLVKHRGRADSEAKARVALENEISGLQESFRLKELEHQSETDALQQQLQVALSQVVVSVKTTGNQYEEEMSRLLRVAKSSYAAEADAFREKLKYYQMNPTPSAAPSMMQDP-ELKRLQAENEALNRVAARDQVADLKNEARHAEDKIDRLTHDLASTKANEKHLANKDMIEKLNQQRISGLEDQYRVLLDRNLSLDAEIDQYRGLLEEEYKRLNMHDQSLDAMDQLVGSPLSAQRDEQAGEQDLDMSELRHMTEEVEEAWTQRMSSNGIQIQSVLDLDRGFDLHLQNIQEPINFDDCVLTVDNGRVSETLVMPHGHVVAVGVVVRVSVGLGEQSHRPPGDIAWTSFRD-FEFDNTINVLQERGSQVICRAEYTCEDSTGATRCMM
Salpingoeca rosetta 1	Choanoflagellata	NCBI XP_004988923	MADTASTPGSPQFRRRVKEREQLDELNNEAAYLVRVRELQKENGELTSQLEAIRFSRKRFE STATIKLEGGVNDLLEEIRRKTTIEIGTLNHQVQDLRKLQNKTEGDALREKDDDLAKQLNEALSKVQELGQVSLQQRDVSQRDATTNLRQVRVASLEADLGRQVDDTKRERLARLEAESRAQGLEDKLSQDDMYQQQVTALEEQIKIGTSGGLSEEDLEAALTEAKEHYKQAVNNFRTQMRSYFAQQPIEPDTRGLEEARLQTEVIEWSKYEDVKAQAGATAQTEIEGLKSKLAGLEDVFAQIKKSHTEIRHRTVVEFKLTKTMQDKDDLRYRDLDERIALDAEIDRYRQLESALRMFEARSPGTGAPQMVASVRKTTRELVALSRRRAVSTKRRATAAQADGGDGGADGELGDGSPKAKMSRLAEEAHSFTVETNTQGALSVEDVDYREGQYILLKNTTASPLDADWAVTNRDRTLQGTFFPAPVTLAPGRSTRVIRGAGDGAEKLEGDVWVTEFFDLAAQPELAVYAFDPDDNESSTWAIMRD
Salpingoeca rosetta 2	Choanoflagellata	XP_004988925	MASFDGGAANNSSHAIAAVRNTMKEREEMEKLNDNLEVYIRRTSALYSSHSTLVSLQASRRTSVVASTPTASPSKSNEEERLKEVKRQQAIEDIRLQSLASQLEKQFENAVEVKDTALNQLNEECAATQQDLRTELEKQKEAYDARLASTATELEESRARAHAAERRVIAAEAVVTLKHLNHHQSSQTTVT-TLEELQACTKAMASKQAPDITIDAVAEVKKEFAAANAYRHKLROYATHAIRQMSQSSWFTEERMALQKQVFTLTKKADLSTTNASLHATVDKAEQVAAIR-QSMKDAVHHHTTEIKAKEERIAALTAHIRANEKKYADLLDERLWDAEIDEYRLLNSIQRPPRPKPKKATPTRLSPAGTPAPRTQATTAPVESDKSDASSGGPGRLEL-RDVTKEVCLLVANPSKTSVSDIDGWALTASADNDLISTLESAHVLPPGGVLR-VFQQRNEHFVENRDCLWDELTRPREWGDGQKAGSDGEGEGL
Capsaspora owczaraki	Filasterea	NCBI XP_004365259	MSSPDRAARLEEKEELQFLNSKLELYHRRQELQTLVKTLESELERERNDAAQRLEIAHRYEATLETETRALDKAAEASHASSLVQVNAELAAATKERARTES- SAADALRSRYEALQAKHELDGRLAVSIEAIKELKDNLKHADSRKHAEKESDALRKSSEKEAQARIAENALARFQDAQASAKTGADLNAYHSEEMRRLRSDLSLKERLRAEHQELAEALSEHRRQTEAESAQYRAEVQALYRSLEAAEQAVKDAVASVAGLQALSAQAQADRARAETVALSORSETLSTRVRLTEHLEHIE-RDTVVEVQRLREEAATRLMQEIAAKDSEVRTLETETNVALDKEIAMRYTLLEVEEKRLGEDFTPTRRGGRDRAASTSVVAGATKRTFSETAEETNTPTVRSKRR- REEDPSPRSASSSATFFAGAPRSVVAATRALFTSSASSSDASSSTGT-TLWTSEPVTEAGLVFEPVNLVEHNKLVLRSTNPEEEVSLHGWKVRTTSETT-GSATYIEPAGIALSAEHPTVTIWTGVSATRGKRKRSRELLWKQDINYSAGAC-IQLLDNEDNEIARLEVRQKEDELFERQSDPRVGCMSIM
Creolimax fragrantissima 1	Ichthyospora	Origins of Multicellularity db Cfra_2736T1	MPQRRTPRGGTRKRPVETNKEDEVNPTPTTAAEPPADPLSPM-VAKRLEDQIRLTLQNSKLEKIEKMEANAQPTNEARIEIESTCARLKEDEDNSLNREL-ERTHAKMADAQRKLLTKEKANIEKKDAEAFKNTKRETEAKALHKSAEKL-RKQTKDLQSKIDDCTKQVGNIRDELAAQNALEQLAESVAAFVDTKRTLGEK-SAEKDHAIDDLNQEILDLYKQEEETDRNQSNEKALQHLRARIAGVAKKERELQ-TKAAEAEADVGVFNRLQKHVNDIKVLEHKKLETAIQTVDVEDKAEVRLDKG-IKDLKLRKELEEVNFRLDNKVDSAQKQKAIIEALEKQLENMKKHAVEMD-CKAQTQVMAEKVQLAYLQAEKQKLLGDQINEYEKLLDFALNAHKKRAR-DEDSDEERARKKTKTDEEPEIIEVITIGIVELVAVDPQGLYVTVLVNETDEEHL-DDYHLIGVSHDEVSTRFRDFSSDCRLDARGKVTYVYNAEYERSIGNAQLW-ETDPVWAASPEGSPQMKLYNSGGLVSAATISVSEKKTGENCCIM
Creolimax fragrantissima 2	Ichthyospora	Origins of Multicellularity db Cfra_6445T1	MTQQTDAYKQEQRDLRELKIVLLNKVLGEKIKKYAAMNPPAQKADELKASLMEVEVKLRDNDKVIDVLEKLNITVSCDALETENRLTTDNEAAL-VAQTEAIEVEKVKLMDDIESLHEVHADLQDLLKRRKQCKEKESESTGKAIKKKR-LEETERTKAQQCDSLTNNAEKDKVLADLRNEADVKNQEMKREISDHENVITNLIE-LQKTEADAKAIEYAKFKKIASEKREDEMKALLEEKKQRTAELEALEQSLPDLKKS-IKTKQAVNKAQAEVDELSTIIKEKEASLPDVEHLKQAAEHKADIKMVKWERIDTLK-KTTEEIAELTNEKLGHEQMTQDLRYGITLEEQLNYYRDLHVDLVSAEAGDGVAD-PSHIPVYSLDSINPNTSSVGVNRNNSITDVLDAWTLTIEYDEDRSTIESAESIKYTF-GDNTMIKPEAIYVTTDPDESVDHNAVWIMEDDKVWVWGYNNVIVKLVKDSGELQ-CSAVVSANDFSSQLG

Sphaeroforma arctica 1	Ichthyosporea	NCBI SAR-C_12785/13776/13427	MARNKGSARGASRRKANSTSQATPTPEVENTKDTQENTVTNES-LLKSPAKNGKESRSPSSPLVKQRKTAKRKLALNAQLANVIHRMEHAEPTEEDRAEV-CHIVNLISSVCCXXTIEEYARLLDYASNGFKRKHGDSDAELDASFLQPTNKRPR-VENEDSFHVITITGVSLKEVDPKGTVKVLENGTEEVVDLSGWHLVGVSHDK-ATARYDFDKDATLNARVCSVLVFSYGYGYPDIPXXXWKTDPPVWSASKEGPST-MKLYHEGGLVSAATISVDESASEQSEGCNIM
Sphaeroforma arctica 2	Ichthyosporea	NCBI SAR-C_07948/15657/16962	QQLEKETGKLFKKKKKFEDTERKTAQKCDLTVIEIKDLDGLLVQRQEQADQKNEN-MKEAFKAHEATVNALKEELXXXVDLDEQLNDYQNLHLVKPSADGLFAPTTTT-PLYIDHIRPAQSSGLVNRNTERDVLHSHWMIRVRYTDETVDASGYDLPGTQH-ITYTFSGPAIVSPDQLYTLTDEHERSTHDEVEIQDDSLDVWAYRYVYLELLDA-DKLLVSTAVV
Pirum gemmata 1	Ichthyosporea	Origins of Multicellularity db c30326_fr4	MLSSFDMMRTPDASKKSPKSPNSTHNSAKGMLSPASQERKAERFERNL-NGRLKYYIRMVREKYPEELERVHNHQREMASQKDRSDNEKQEMKEHFNKIQN-LEEEVARHSASNRESLKKIDQLHREKANFEKAIENLQDVLKKNHINQLKKEEK-TSSVLVQEEELSELKELADAQDEKAGAIQIEIENLRS-QLNSQIELVDKLTTEELETKSNLENNEQELNSAYEERLE-IALAEAKVHQQNAKMLDDVERSIEKIFAEKEADIKRV-TQGLQKENDLKEQKETILQEGEEIKNNENLQQSVV-LKAKIDELEKQYNNVSKSKANLREIDSINKLLEEARL-ERDIAEKTKNKNESSVVDNTGGSRKRGDRDLEED-EDVEVPKRSDVDKLDQVDDSSGDVHKEVDAGFFIRL-QNVTSAPIQMDQWLLRSYEDDAIDFNFSKDLVLEPEQYL-TIITRLDEEEIQDSEQLNVQWDTSEPLWNNEGGFVLFNQEGASVSAQI-ELEKRTCIQQ
Pirum gemmata 2	Ichthyosporea	Origins of Multicellularity db c13682_fr2	ELQNGREVLESKIEDLVTLSSKTKQFTEEKQNLQDSDNAERLENIKTFIKDQY-EEKLEKTRQDLAEEYNARLEDALLNARDQHERENEELFELRLNNEHQEELSELNVK-MMKNESLLKQSIKALQVNDTNDLRELSNQVMKEEQEAEKKELEENLQVLD-VKKSFEQLQADYEEQVNNDSHRRLLDEVAEENSDRLREIDYTFKMIQEEARMG-MKFQVDPKTVKTNKTSRVFSDNYQPQSFGEDEDEEQDVEDVSGSKRSKNIQFQ-LQSSINQHSTTFLDEVDPEGNFFYLKNNKPKVDLSTWVLRASSASRDYVDFHF-PDNITLKGSEKAVLTSQNTSESDMKLDRREFNYLVFWTNTVSIWANDKGSVNLFK-GEKLCASAATITEEVQESPAKRFSSYIFGL
Abeoforma whisleri 1	Ichthyosporea	Origins of Multicellularity db c11593_fr4	MTSAPGTPANQITIDTNMLSPQSLAKEQKRLADLNDRLSDFSLGRKGRDDIISGY-ERKIQDLNKLLEAGEVHLNLEADIAERDAAIQKLEIVKQN-LEASREDLQKYYESKERQCEDLVEDREQKVS-RFHREIQLVQVEKSEEKAKNKEMTKELKEVKNQNLNSAQKLEKAIDEKSEAINK-AKQLEENLDFKTKELNAQIEKLEHDLKEKDEENASRLEEEAECDFFVRMQLAL-QEARQLFETDLESQLEAVREELNNDLVEKESKREAEEMNIESDLNENVSNLEA-LREELKTQAKGAKAAKKADESHAKEKKEKLEQEQKANKFENEREQNVREA-EDRYSKMKQKLEDEKEQCQKMKMEDQLQMKKLEKERKQSEKNSLSDWRKT-KKKMDEEYQREKETQKAANDLAYDELKAEYENLVQITDVTNLLLEERELGIP-IEQENRSLTTVTKLMMKHQTPSRTGGKRTAQQAEPANTSQFGKMKLRFDSD-DESENDVSGKQVQITKYGHQINFDEIKTNGQYIKIKNESEENVSLTGWSLKCTF-ESEIEITFDYDELASEDSVAIVTSSDYMSNIKGNKILFETEPFVAVDNGVVEL-YDDRGDVICSAVVEEYDNLNKS RNGPLCYQAQQ
Abeoforma whisleri 2	Ichthyosporea	Origins of Multicellularity db c11301_fr3	CTATISCLYVIALLSLLNMSVMKPLQSRNLLNSNYEATPVRKIVKDSGQ-NARLQNLNGLRKYINLVROKHPEEMDKVKNVEYVMMKDLQGGQI-NALQDEISELNSLNVLQEKELKEKNASLRQKDDL MNKQEQFTGLEQSIADLQ-VANKVLTNKRDLKAKVSALKEDYDELCEANALEKLRVLEENMMKKAADN-AAEQKLLHAIEMKLSHQDFAEKERLASEKHSNEMREMSEEFFEKERNLKNEL-KQAERQVEKIESEKATTEILGKAVSDFNGKFEQKELKEVTHKLNKLAD-SKTLLEDENAKLLQKIKDLELNSDYKLTQVETKLENKTRKYHDSSTILEQ-LDIYNLLLEEEKRIEGLPKFSEASSRFFHESPOFSSAVAAPLEDDDKYKRP-RESDENSEETEYHEKRONVETQLQSEIVSSTGSIVLEEDATGAFIRIFNN-SKHPVCLSGWNILVKNPSSGEIGQFCFPDDFLAASGKVSVTSFNDNDDD-TEVHLQWNEPSIWSDDGCTIQLFDMNGHIASTGKVEEVGVLDLSTSIWGLFSSFLG
Dictyostelium discoideum	Dictyosteliida	NCBI XP_636248	MDMSKSKSRASPIESSQEEIAISTSKTATTEKPKTKTKTKKASQPSQEVVETE-SEVETTTTTSTSTNNNNITTTSTSSQQSNGTSSSSSTPQISPTTPIKSIY-LSLQIGTPLSPNRAAQRLEKDELSLIHNRLKALKKLESAEATGKLEQKQY-EEELDQKHATATIKQLKQRSQDQVEKQLEIEQNNQNSDLTNRNLE-LENELKSKESVWKKEKDEILLKFOESINKLNQENSLAQSGLKSE-IEVSKYEYIDGLKSEINRLKDDLQYRIRREGEEKSRKLELENEYN-RFKGKEEYENQLIVSKDEEIKKFKELKEKSSNAMNKK-NELENNLIQAHERQIEDMRDSINREWELKAAQMMEEHARTI-HLQQAVDSFNEEKERIKSQMETLNGQIEDINIKNNEYEDRIK-EMNVLLSQKDNSIGELGVEIEESKMKMRKQMAADLKSQKGGQIA-LLQIEINTKDNKNTLQETNRLKSELYSITNQIDPEIPLDPEI-NSLKELVKGFECTVDDRKRKRKSLQHEFNAAANQDQNGMT-IEEQSSTSTTTTTATGSSSSTSHLDNDIDSSKLPTEGPESEL-FNPDTVSFLVDSNQEFIKLSVHGDMNDGLSISKWRLIVVK-PDGSKSGFSFPDGIQPFKGIKSVTVWTRPRPQGTPTENE-FYWARTELWTSPEVGTIVKLVSPSEETTTVTLADGIYQKPSAGKSNCLIM
Dictyostelium purpureum	Dictyosteliida	NCBI XP_003294572	MESTKKKTASKRSQSPVDVAEKEKVVEEVEVQTVAPTKTKK GKATTT-TKKSKEVVVSTPVSNTVLSADDMATESEIEPSTTTTTTTTTSTST- SPTTPTISKYIPSLNMGTPSPNRLRLKEDLENNIKSKLQL-CLKLESTENELEKKNKELDDMEHEHSLAIGNLKNRNEHEVKLEIEIKTSELS- SRDALESELKTEASWKKDQDEYVTKLETINKLNQEHSTIVSGLKTDIAAREY-ES-DLTKTEVNRKEDLQYRMRESEDKSRKLENEYKRMKEKEEEFNQVLLIKDEEIK- KLKVEVKEKDLHSASSKKNELKHLIEAHERQIEDIRDSINREWVRAVQMV- EHNARTFELQQSVSEFEEKQRYKQISLYGAQIDELNIKNEFEDVKELNDLIAQ- KDSSIGVLNNELESKKRYRQAADLKNKDSQISLLQVEMNGRESKCATLQAEI- RLKSEIFITSQTPEIPLTREIEIKTLVGVFEKNVNERKRNKVDNNAANVNT- TADPNISTDIPSSAAQVDPPELEKFNPNVTFAVVDNTNQEIRLSVHGDYDNGL- SLSKWRLIVAKPDGSKAGFSFPEGIQQPFRGIKSVTVWTRPRPQGTPTENE- FYWARTELWTSPEVGTIVKLVSPSEETTTVTLADGIYQKPSAGKSNCLIM
Dictyostelium fasciculatum	Dictyosteliida	NCBI XP_004349984	MIVQKEEEMEPTKSKSSASSVKKTKTEKRDVAAEESAHEESDVQVKD- KRKKKSTAVPKDTAASSSSSSQVQDADANSPPMAIDEPNESPSSSTVT- STSSSTSTTTTSHWTKTPQHNGSISHNISTPQITPSFQKLSMTTPLASPC- GRSSTAARLKEREDLIQINSRLRDLASSEQKDNEIRRLRDEYITTSRQSDQISN- LEKLELDIARLQNESKLTVELSSQLQLAHOARSAKEQGHQAEIQNFQNKMDTEK- KITEENSOLVVALRNDNSKVMIEISLKSVDVRLQAEILLRSKAEAKAHRLKE- YNRTKLEEEELRNHINQKVAIEKLTKEKELQVEYSRKESQLQTIETGKFL- QAEDMRDSVNRWEIKCVLVEQNAKITQLEISANFSEDEKEIKYSQQLQVYQ- IDDLNIKHEFEDRISLQKDLDTKDDTISQLAEIDEVKKAAARKVQADSKVKDIA- IMLQDEINQKDVKNLSQSEMGLRKEIMFTNASPEQDIPITREIEQLKSLVNF- RRVKIKREDEDEESNODQEMNGGQQQQQOQHQHQHQHQKQSSDFQV- NTSLSIDAISECIRLTVSGDFEDGVSISGKIVVTKPSSRIGFSFFENIVFKGLN- INLFTGKTRPGTTPNEFVWSRPNWESPEGTIVKLLTPTDEVLANVTLPSNGLYHK- DDKQSNCLIM

Dictyostelium citrinum	Dictyostelida	NCBI AJWG0100 4320	MDITKKKSKRSSPIESSQEEIAIXXXKYIPSLSQMGTPSPNRAAQRLEKDEL SLIHNRKLTALKKLES AIELEKKNQYEEELDQKHATIKLQKQRSDQFEKQLMEEKQKQNSDLASNRNILENELKSKESVWVKEKDEILLKFQESINKLNQENSLQSKSDIVSKYEIDSLKSEINRLKDDLQYRIRIEGEEKSRKLENEYNRFKGEYEQLLSKDXXXSNAMNKKENELNLIQAHERQIEDMRDSINREWELKAAQMMEEHHSRTIHLQAVVDSFNEEKERKMSQMETLNGQIEDLNINNEFEDIKEMNVLSSQKDNSIGELGVEIEELKKMKRQADLAKSKDGQIALQIEINTKDNKCNTLQETNRKLKSELYSITNQIDPEIPLDPEINSLKELVGFECTVDDRKRKRSLQHEFNAAANQDQNGMIEEQSTTTTSSANSSTSHLDNIDSSKLTGPQESELFNSDTVTFSLIDNSQEFIRLSVHGWNWNGLSISKWRLVVKPDGSKAGFSFPDGIQPFKGISSVTVWTRPRPQGTPTENEFYWARTELTWTSVVEGTIVKLVSPSEETTTVTLPADGIYQKPSATGKSNCLIM
Dictyostelium firmibasis	Dictyostelida	NCBI AJWH0100 3706	KYIPSLSQMGTPSPNRAAQRLEKDEL SLIHQRKSSKKLESTEIELEKKNQEELEIDQKHTLTINKLQKRSDFERQLLEEONQNSDLTSNFINLDNELKTKESSWKKEKDEMLSKFQESINKLNQENSLVQSQLKADIVTKEYEIEGLKTEINRLKDLQYRVREGEDKSRKLENEYSRFKIEDEYHQLISSKDXSSNAMNKKENELNLIQAHERQIEDIRDSINREWELKAAQMMEEHHSRTIHLQAVVDSFNEEKERIKSQMDTLNGQIEDLNINNEYEDKVKEMNVLLSQKDNSIGELSVEIEELKKMLRQLADLAKSKDGQISLLQIEINTKDNKCNTLQETNRKLKSELYSITNQIDPEIPLDPEINSLKELVGFECTVDDRKRKRSLQHEFNAAANQDQNGMIEEQSTTTTSSNGSSAHLNNDSSNLTGPQESELFPNGTVFSFLIDSTQEFIRLSVHGDMDEGLSISKWRLLVVKPDGSKAGFSFPDGIQPFKGISSVTVWTRQRQSGTPTENEFYWSRQELWTSVVEGTIVKLVSPSEETTTVTLVVDGIYQKPSASGKSNCLIM
Dictyostelium intermedium	Dictyostelida	NCBI AJWI01001 930	KYIPSLSQMGTPSPNRAAQRLEKDEL SLIHQRKASIKKLESVEIELEKKTQEYEEELDQKHATISKQKQRSDQVEKQLEIEEQTQNSDLSSNFNILENELKSKESWKKEKDELLKQFQESINKLNQENSLVQSQLKADILNKEYEIESLKTEINRLKDLQYRIRIEGEDKSRKLEFEQSRFKTREDEYNQLILSKD
Polysphondylium pallidum	Dictyostelida	NCBI ADB-J01000003	MSETTSASKKKSSKSKRSSEKEKEKELSSSVINTSNDGDDEEMSEPSTPIEKKEKERKKSSSSSSSSKTRAEARGDKSGADKSTTVTSSSTTTSEKDTGAGHISSMQVEESTQSVTTSSNLNNNVSKTPTSHNANGSGITASSLPMTLKMFGSIGSTLQSPSRLSINRMKEKEELTEVTKLSKILTELEERNNEIURLKAEILNSKSSSVEVSAALRTRLDE TENRLTQEIQRQAEEAASLETLADLKNKEDLYMHDNRFTDARLEDTSSITKEHSAQLTVLRTDLNKRLEHEIQKDIKLLNADLLTKSKEQDETRHRLDAEYKRMKQKDELELNLISQKDDDIKVKLTLKEKELIIIANQRKEAELNQTVEAFERQVEDIRDSINREWEIKVAEAEIEQTSKYASLEMTAKSDFERETFRNQJAVIQQQTEDLVNKHLSRDRILDLSNELRNKEEQIITLTAEQEELKKNLRRSQAEKSKDKGKLSLLQDEISSKEIKFQYQSEMHTLKKELHSANNQSMDEPIGDEIRRIKELVNTLHQSRPHKRQRSGEYSGENNNNNGNELAANHEEEMSDAEQQSPSSTNSVEPAASTLKNLNTMCLSIDTTKETIRLNASGDHYNGMTDILDKWIVNKPDPGFKYGFQFPANITPLKGINVITLHTGRSRALNTAAENEFYWSRTGIWDSPEGTTIQLVSPKEVLQESVSLTNGAVVREESNPKEQGNCLIM
Polysphondylium violaceum	Dictyostelida	NCBI AJWJ01000 603	TPLNKMRESLLGTPMSPSRVMRLKEKDELNAIHKKLQOCITSLKKEDEIEKEREIEISLQKEKTKSNAYKTKLEEAEERLLSQEIEKTEFLSRADFLESELKTEATWNKDKQDMISKMDAEAINLTHEITLSSSVKSDLVKSEYDNETKKNEIARLQGEELHTRVKEYDEKSRKLLDNEYNRMRGKEDFNSLLKQKDEDI-KKLKMLKEKEKLSLVAQRKENELNQTIQSYERQLEDIRDSINREWELKTAAMVEEXXERERFKTQITVIQEQVNELNIKNQEQDLIEQLNNTLMNKENLNDLSQVENEDNKKLRLKQADLNKNSQISLLQDEINSEKTKLITYQNETTKLRTIEIQYINNQNDEPEIPLTRIEIASLKELVNGFEKVAEDSSRRKRKRXXXQEYIKLSVHGDGFKGFCISGWKLIVLPGGGKCGFSFPEINQIPVKGIIITVWTRSRPTHIQTPENEFFWRSREIWTKPYEIEITLKLAVGVSNETTAKVSLPETGVYQKDPSPGKSNCLIM
Pavlova sp. CCM-P2436	Haptophyta	MMET-SP0985 CAMN-T_0021991 105	GASRNRPPSCSPAFASMSSESGRRSSIGAGPESATRLRVEERETLQNLNSRLVANDLQVRVDTAASVHKTQSDQLRDMYKQLTDMRKTEKQVDELITELKSKLSAAQKEANEATKDKVALTRKAQLADELVDKLRVTDERDRLASRAKEAEARAAEAEKSAVFEAKANMLEKKLAPLRDQLATAEELRREKSEYTRAQIAELSGGSRSEIDKLSAYDKQLLEAMEKQKKRADERELTELAQVREHFTAVEAELMRAVDDEKAAASADERMGLNAQVRQLRARASQAEQAALAEKTLAERIAEHFDAAIARLDQQLQEARAARQQKEVEFNELMDVKISLAEEELKTYTKLLEGEERLGITPTKEPKRRRTTADDGGGGASGAASGPLPYLATLINDSFALKNAGDAPAPLGGFKVVSAAHATDVSFEFPDDELQPGATVTVYSGKKNQRKRRDPRTLWWSPRYMWRNEGDEAIMYDSSGSEVGRVVAEANSIAEADPTEAAGGDAGGDNKCVIM
Pavlova sp. CCMP459	Haptophyta	MMET-SP1140 CAMN-T_0007433 825	FSSPRNLRERRRTEHHSRPAHAMSESRRLLSTGSGSDSATRLSRVEERDTLHNLNSRLQFLSRMKELEARNTELQSKLDMGSSASQKVTDSMHALMELRQIQLRAASEKDTLIELKGLAALQASSGENIRSNQAMKAKAERAAEGLRVTIDERDLAALAKDAERRADVAEQAAVTSSEKATVLEKKLAPLRDQLQTSSEELRREKETAAYRSQISELSTANKTEVARIQADYDKLQALAAAKKRADEERAQVLAELHFNVAEQLTREAKDERASKEAALARAELGKTNDQLQSRADASEKRAAAEAKNLTERARQHEATVAELDSLHQAARAARKEKEGFENLMDVKISLAEEELKTYTKLLEGEERLGITPTKEPKRRRTTADDGGGGASGAASGPLPYLATLINDSFALKNAGDAPAPLGGFKVVSAAHATDVSFEFPDDELQPGATVTVYSGKKNQRKRRDPRTLWWSPRYMWRNEGDEAIMYDSSGSEVGRVVAEANSIAEADPTEAAGGDAGGDNKCVIM
Pavlova sp. RC-C1486	Haptophyta	MMET-SP1335 CAMN-T_0007433 825	PRASRSTRVAVPLVAVGGGPD SATRLSRVEERDTLHNLNARLEYFLSRMKELEVANDLQVRVDTAASVHKTQSDQLRDMYKQLTDMRKTEKQVDELITELKSKLSAAQKEANEATKDKVALTRKAQLADELVDKLRVTDERDRLASRAKEAEARAAEAEKSAVFEAKANMLEKKLAPLRDQLATAEELRREKSEYTRAQIAELSGGSRSEIDKLSAYDKQLLEAMEKQKKRADERELTELAQVREHFTAVEAELMRAVDDEKAAASADERMGLNAQVRQLRARASQAEQAALAEKTLAERIAEHFDAAIARLDQQLQEARAARQQKEVEFNELMDVKISLAEEELKTYTKLLEGEERLGITPTKEPKRRRTTADDGGGGASGAASGPLPYLATLINDSFALKNAGDAPAPLGGFKVVSAAHATDVSFEFPDDELQPGATVTVYSGKKNQRKRRDPRTLWWSPRYMWRNEGDEAIMYDSSGSEVGRVVAEANSIAEADPTEAAGGDAGGDNKCVIM
Symbiodinium gore- aui	Dino-flagel- lata	MMET-SP1369 CAMN-T_0049246 593	MPRRNSTPPRRNSTSSKSSKTSKESKKEESVSTTVVEETRVVNDQSEIORSRVQEKHQLHNLNDRQLNFWLHVKAMEEENADLRQLRDMEEERMKESLQEVHQEYKSKKLRSTSVMATAKNNGLESENARLAEVHALNKKIQTDSLRENNHKLSDQVRDLSDELRYKAGVSNQEEIEERLEIIANHEEVIATREDQLQEATKELETVQKTIKKGARHSDGELAEVRASLTAHKKWQRETQTALQKLETDLRQHFHTTVELNERSETLVQRNERLRKENERLANELDNNAGGQQDLLNKLAEQERQKHVVTVQLORDLMKREELTTAASSLVNKNNEFNDLMDVKISLQTEIQRHYSILDVEEARINPTQDRHRPSKFEKAQDEAKNNDEEETTTITTKARKTVKRSRSTQEIAPGKGGKGGKGNKRRRLKNNKDESEERS TTRLTTRAKTGPLHISDVLDISDCVSVANETSAPVPMKGWTLISETGNQVFNFPDALVLPKGDVTVVSGAAHESHHPDNLFWTRRYVWVWNNHGDAILVNPAGEHVSIVSGVPSAMSAASDATPASSVTRGTGASVNGDDQNCVIM

Alexandrium tamarense	Dino-flagellata	MMET-SP0384 CAMN-T_0004031397	LTQSKKSAVSOARKAGQQLASERQKWKEESDSRAAGIEAELREAFSTSSSEQLK-ERAQILNTQNVRLQEEMDSLMEELQKANEAEAKERLQELQASRASH-SEESQAWAMERSRLEKEVTSLHEAHEAEKIDFNLDLMDIKVSLQAEIEQYRAI-LEGEEARLGVDPSSVDFAGVKRDAVIADSEADDDGDHSAAGAGAGAGGGGSGL-DEDDVADEEEETPRRSKQRRGRRARKSSRAKSSRTTTRSAKRARVEVASSEVGA-AAAASATSRQLAAAASGVKLDNVDLSDCVSINKNFTSEDQPMKGVWLLKSAVGSQV-FFRPDDLVLPVPPGKRVTVWVGKTAHRHHSPPYNFLWTRRFIWNHNGDTAVLINADN-MEVSRITSAPGDDGGMRADDDASEEGEGEGEGEGEG
Kryptoperidinium foliaceum	Dino-flagellata	MMET-SP0119_2 CAMN-T_0007964289	GEARRSSRSRRTTRSTKSSSSSRHASGTSKRSRDKKGGAGAGSAGTSASTH-PRRHAPSIGIKIDNVDLSDCVSIVNNTGASQSLKGWHLTSEVGDQIFHF-PDDLVLPPGERVTVWVGKNAQRHHNPPYHLFWTKRFIWNHNGDTAVLVDANNNVTSKITSPPGGYSRDASSDGAGEGEGEGEGGDSNCSIM
Lingulodinium polyedrum	Dino-flagellata	NCBI GABP01011089	LAVKPEAPEAAEVKVVISLDEADSVVITNNGSEAVSLKGWKLVSKTGDQYAFKAKDSIDAGASLTVLSGKKAKAAARKGGADHVFWTARYIWN-NEGDEAGLLNADGEEVSSLAGTGT
Hyphochytrium catenoides	Hyphochytridiomycota	NCBI com-pound of CAFC02154760, CAFC02067206, CAFC02215379, CAFC02216946, CAFC02036197, CAFC02052195, CAFC02276476, CAFC02179373	LNNRLEAYVLRVKEMDDARSHLDHELKTISTRYQKEMDALRERLSRD-VDTLRSELEYEIDQKTVHQVKAEDLTLQVEKMTAQVRITSKNLIEEEEEEREM-DRDEKESCAFKREMTGEPPELXXXXXXXXTTPKQELFFKIKQTEDV-GREEAKRYKDEIRLEKEVSALRDSQAESEKQNDLMDVKIGLNEIRHFKA-LDMEEDRVGVFSPEKRRKKSITPKSSAKKARVAEPAFVIVVQIMDTENE-CVVIKNNSSAEKLDLHWTLNKAQTKQFEPPEGFSIPADSSVRIWNGKKNKSK-HSPPENIYWSSKFWVWADGEAALLSPSGDLAASSAGGEVAGGAETGPSQE-YEEVEVEVEEERCPVQGDQEQKCTIM
Phytophthora capsici	Oomycota	igilestExt2_fgenesh1_pg.C_PHYCAscaf-fold_160053	MAKELSPIKSKRMEEKASLQKLNRSRLEMVYLVGNELSAKHAARELETIKQR-MQDQDSVRSRLTKELEETRKKLDDEVDQNNRLQTLQEQNRELVKLRAQQLKE-FGELKVLVETLKVREDREKTNESAQAQALSEQTTQLNSARRSVKELERDLR-SHKAALTDATQELQDLRKKCANYDMAKDSLTKLRREWAAKHLEAQAIW-KDAEERLATMEMEVRSHFNSMSTLESQDDVRLTELESTKKELDRTANDY-EDSLKARQSLTEKVAQLERDYREARSKSTKDRKAYEETLERFRSSKV-AKEREFNELMDVKIALDAEIMKYRRLDREESRVAVTTPNTKGRKRKSEHT-NGTSSKKAKPSSAVQIASLDLEKDRIVQNTSNPEVSLGGWAVRGQMDQT-FRFPKTYVMRPRSTLTVLSSKRNKSAKHEKSGEAAFLATQFSLNPNGDI-VFLVTDDDIPVSMMESEGLSEEEVRAIEADLGVDLMDDEAPPSSGNCGMM
Phytophthora parasitica	Oomycota	NCBI ETO76261	MPTLSPISKSRMEEKASLQKLNRSRLEMVYLVGNELGAKHAARELETIRQR-MQLELESVRSRLTKELEETRKKLDDEMDQNAKLQTLQEQHQKELVKLRAQQL-GEAKVVLVETLRVQLDREKANAVSAKETLSEQTTQLNARRVVKELERLGRHSAAL-GDATKELEELRKKSVDFDLTRDAELTKLRREWAKHLEAQAWKKADE-DRLQSMEMEVRSHFESVSGSLQTLQDDVRLTELESTKTELDRDRTANDY-ESLKARQSLTEKVAQLERDYREARSKSTKDRKAYEETLERFRSSKVAKE-REFNELMDVKIALDAEIMKYRRLDREESRVAVTTPNTKGRKRKSESSNGTSSK-RAKRIIETETSSASAVQIASMNLKDRIVKNSSSDPVPLGGWVVRAGM-DQIFRFPATYVMRPPQSTLTVHSSKRNKNAKNERKKGEDSFLANKFTLNPK-GDFVVLNSDDIPVSMMESEGLPAEEVRAIEAELRADFMDEVPSSGNCGMM
Phytophthora cinnamomi	Oomycota	igilestExt_Genewise1-Plus.C_460078	MPTLSPISKSRMEEKASLQKLNRSRLEMVYLVGNELGAKHAARELETIKQR-MQDQDSVRSRLTKELEETRKKLDDEMDQNAKLQTLQEQHQKELVKLRAQQL-GEVVKVVLVETLRVQLDREKANADSAKETLSEQTTQLNARRVVKELERVRGH-VAALGDATQELQELRKKSAADFDLTRDAELTKLRREWAKHLEAQAWKKA-DEADRLQSMEMEVRSHFDSVSGSLQTLQDDVRAELDSSKMELDRTANDY-EESLKARQSLTEKVAQLERDYREARSKSTKDRKAYEETLERFRSSKVAKE-REFNELMDVKIALDAEIMKYRRLDREESRVAVTTPNTKGRKRKSENTNG-TSSKRVKRSIIQTESTSSASAVQIAALDLKDRIVLKNSSSDPVLGGWVVR-GQMDQTFRFPATYVMRPPQSTLTVHSSKRNKNAKNERKKGEDSFLANKF-SLNPKGDFVVLWTADDIPVSMKSGLAEEVRAIEAEYRADFLDDVPPSSGNCGMM
Phytophthora infestans	Oomycota	NCBI AATU01004774	MPTLSPISKSRMEEKASLQKLNRSRLEMVYLVGNELGAKHAARELETIKQR-MQDQDSVRSRLTKELEETRKKLDDEMDQNAKLQTLQEQHQKELVKLRAQQL-GEAKVVLVETLRVQLDREKANADSAKETLSEQTTQLNARRVVKELERDLR-SHAALLSDATQELQELRKKSAADFDLTRDAEMTKLRREWAKHLEAQAWK-KDAEDRLQSMEMEVRSHFDSVSGSLQTLQDDVRAELDSSKMELDRTANDY-EESLKARQSLTEKVAQLERDYREARSKSTKDRKAYEETLERFRSSKVAKE-REFNELMDVKIALDAEIMKYRRLDREESRVAVTTPNTKGRKRKSENTNG-TSSKRVKRSIIQTESTSSASAVQIAALDLKDRIVLKNSSSDPVLGGWVVR-GQMDQTFRFPATYVMRPPQSTLTVHSSKRNKNAKNERKKGEDSFLANKF-SLNPKGDFVVLWTADDIPVSMKSGLAEEVRAIEAEYRADFLDDVPPSSGNCGMM
Phytophthora sojae	Oomycota	NCBI AAQY02000086	MPTLSPISKSRMEEKASLQKLNRSRLEMVYLVGNELGAKHAARELETIKQR-MQDQDSVRSRLTKELEETRKKLDDEMDQNAKLQTLQEQHQKELVKLRAQQL-GEVVKVVLVETLRVQLDREKANADSAKETLSEQTTQLNSARRVVKELERMRGH-VAALS DATQELQELRKKSAADFDLTRDAELSKLRREWAKHLEAQAWKKA-EAEDRQRTMEMEIRSHYDTSVSGSLQSLQDDVNAELESTKKELDRTANDY-EESLNARQSLTEKVAQLERDYREARSKSTKDRKAYEETLERFRSSKVAKE-REFNELMDVKIALDAEIMKYRRLDREESRVAVTTPNTKGRKRKSENTNG-TSAKTRKRALVAEPVSSSSAVQIAALDLKDRIVLKNSSSDAIPVGGWV-VRGQMEQTFRFPATYVMRPPQSTLTVHSSKRNKNAKNERKKGEDSFLAN-KFSLNPKGDFVLLSSEIPVSMKSAGLPDDEVRAIEADMRAEFMDEAPPSSG-NCGMM
Phytophthora ramorum	Oomycota	NCBI AOBL01008173	MPTLSPISKSRMEEKASLQKLNRSRLEMVYLVGNELGAKHAARELETIKQR-MQDQDSVRSRLTKELEETRKKLDDEMDQNAKLQTLQEQHQKELVKLRAQQL-GEAKVVLVETLRVQLDREKANADSAKETLSEQTTQLNSARRVVKELERMRGH-VAALS DATQELQELRKKSAADFDLTRDAEMTKLRREWAKHLEAQAWKKA-EAEDRQRTMEMEIRSHYDTSVSGSLQSLQDDVNAELESTKKELDRTANDY-EESLNARQSLTEKVAQLERDYREARSKSTKDRKAYEETLERFRSSKVAKE-REFNELMDVKIALDAEIMKYRRLDREESRVAVTTPNTKGRKRKSENTNG-TSSKTRKRALVAEVSVDLSSSVQIASLDLEKDRIVLKNSSSEPVLGGWV-VRGQMEQTFRFPATYVMRPPQSTLTVHSSKRNKNAKNERKKGEDSFLAN-KFSLNPKGDFVLLSSEIPVSMMESEGLPEEEVRAIE

Phyto-phthorales	Oomycota	NCBI AOFH01002457	MPTELSPIKSRMEEKASLQKLNRSLEMYVLGVNELEGAKHAAERELETIKQR-MQQDVSRSRLTKELEETRKKLDDEMDQNARLQTLQEQHQKELVLRRAQQTEL-GEVYKLVETLRVQLDREKANADSAKETLSEQTTLQNSARRRVKLEERDMRGH-VAALSDATQLEELRKKASADFDLTRDAEMTKTRREWAAKHLEAQAALWKK-DAEDRLHMEMEVRSHFDNVVTSLQTLQDDVRELDSTKKELDRTANDY-EDSLKARQSLTEKVAQLERDYRVERAKSTKDRKAYEETLERFRSSKYSK-EREFNELMDVKIALDAEIMKYRRLDREESRVALATPKTKGRKRKSENTNG-TSKRTRKRALAVEESVTDLSSSSVQIASLDLEKDRIVLKNSTSEPVPLGG-WVVRGQMDQTFRFPATYVMRPQSTLTVHSSKRKNKNAKNERKKGEDSFL-ANKFSLNPNQDFVILVTSDDVPVSMSEGLPEEEVRAIESELRADFMDDDAPPS-GNCGMM
Bremia lactucae	Oomycota	NCBI compound of JP956599 and JP961338	MTTELSPKSRMEEKASLQKLNRSLEMYVLGVNELEGAKHAAERELETIRQR-MQQDLSVTRTLTKELEETRKYXARGFMLYDVIIVAVTYRGLYRKLDDVEDQ-NAXXXXXXXXXXXXXXXXXXDEESLKAQALTEKVAQLEREYREIRSKATDRKAY-EETLERFRSSKVAKEHEFNEMLMDVKIALDAEIMKYRRLDREESRVAVATPNT-LGHKRRKKEATRNVAASKRTRKTSTETNVLVSPVQIAAIDLKDRITLKNSTNDA-IPLGNWVVRGQMDQTFRFPETYVMRPHSTLTVHSSKRKNKNAKNERKKGED-SFLANKFSLNVKGFVVLMTSDDVPVSIKSEGLPEDEVRAIEADLRADFMD-DDAPPSGSCALM
Hyaloperonospora arabidopsidis	Oomycota	NCBI ABWE02006180	MSAELSPIKSRMEEKASLQKLNRSLEMYVLGVNELEGAKHAAERELEAIKQR-MQQDLERSRSLTKELEETRKLNNEMDQNARLQTLQEQHQKELVLRRAQQTEW-SEVKVMLETLKAQLIGEKANAKAAKETLSEQTTLNAAQRRVKELERETRQSQ-SVALSDATQLEELRKKASADFDLTRDVELTKLRREWAKQLE-SQAQMKKDAEDRLTMEVEVRSFESVSSLQNLQDLDVKS-ELESTKQELDRANDYEEESLKAQALTEKVAQLERDYREVR-SKATKDRKAYEETLERFRSSKYSKEREFEFNEMLMDVKIALDAEIMKYRRLDREESRVAGATPNSKGRKRKSENTNGFQSKRAK-RAVVFEREGLSSASAVEIASLDLEKDRIVLRNTSSKPVPL-GGWWVRGLMDQTFRFPATYVMRPHSTLTVHSSKRKNKNAK-DEKRRKGEDSFLANKFSLNQGDFVILVTSDDVPVSMKSEGLSEEEARAIEADLRAD-YTDEAPPSGNCCMM
Pseudoperonospora cubensis	Oomycota	NCBI AHJF01000192	MPTELSPIKSRMEEKASLQKLNRSLEMYVLGVNELEGAKHAAERELETIKQR-MQQDLNMRSLTKELEETRKKLDDEMDQNARLQTLQEQHQKELVLRRAQQTEL-GEAKVLIESLRVQLDREKANAESAETLSEQTVQLNSARRRVKLEERDM-RGNAAVLNDATQLEELRKKASADFDLARDALSKLRREWAKHLETQA-QWKKDAEDRLQTMEMEVRSHFESVSTLSQSQFDDVKAELSTKKELDRT-ASDYEESLKAQSLTEKVAQLERDYREVRTKATKDRKAYEETLERFRSS-KYSKEREFEFNEMLMDVKIALDAEIMKYRRLDREESRVAVATPNTSRKRKSE-STNGTSSKRAKRAVFETEFKSSPVSAVRIDSLNLEKDRIVLKNSTSEPVPL-GGWWVIRGQMDQTFRFPATYVMRPQSTLTVHSSKRKNKNAKNERKKGEDSFL-ANKFSLNPNKQDFVILVTSDDVPVSMKSEGLPEEEVRAIEADLRADFMDDEVP-PSGNCSMM
Pythium ultimum	Oomycota	NCBI ADOS01000666	MPTALSPIKSKRLDEKSTLQFLNGRLELYVLRVKEMEDAKNVAERELETIRQRMQA-DIDALRLRMSKELDETRKKLDVLELDQKARLQVLEQEQHVLEVLKLRRAQVKEL-SEIKALAESLQVELTKERSNSKASKETLSVQTTQLQSARRRIKDLERENRFAESLSD-TTKELDELRRKSSDFDLTRDLEITSIRKEMNQKHQEAALQWKKDTEERVLGVEK-VEVRYLFEVGVQGFQKQVEDLSVELDSTKKELDRTASDYEESLQVRQQLTEKVAHL-EREYREERTKFKEDRKKYESTLEHFRSSKFAKEQEFNLDMDVKIALDAEITR-RLDREETRFGLPTPTNPNAKQTAARKRSETSGASAAAAASKRIKRAPSSVQIHL-DLEKDRIVLENSDKPVSAGWEVVRGAEQVFRFPATYSMRPKSKMTPVYSSKR-NNKNAKDERKPNEDSFLATKFLSNASGDFAVLLTADGVPVSTRAEGLSAEEVNTLEA-IKADLDDEAPSTEGCGIM
Pythium arrhenomanes	Oomycota	NCBI AKXY02009383	MPTALSPIKSKRLDEKSTLQFLNGRLELYVLRVKEMEDAKNVAERELETIRQRMQA-DIDNVRSLRSHLEETRKKLDVLELDQKARLQVLEQEQHVLEVLKLRRAQVKEL-SEIKALAESLQVELTKERSNSKASKETLSVQTTQLQSARRRIKDLERENRFAESLSD-TTKELDELRRKSSDFDLTRDLEITSIRKEMNQKHQEAALQWKKDTEERVLGVEK-VEVRYLFEVGVQGFQKQVEDLSVELDSTKKELDRTASDYEESLQVRQQLTEKVAHL-EREYREERTKFKEDRKKYESTLEHFRSSKFAKEQEFNLDMDVKIALDAEITR-RLDREETRFGLPTPTNPNAKQTAARKRSETSGASAAAAASKRIKRAPSSVQIHL-DLEKDRIVLENSDKPVSAGWEVVRGAEQVFRFPATYSMRPKSKMTPVYSSKR-NNKNAKDERKPNEDSFLATKFLSNASGDFAVLLTADGVPVSTRAEGLSAEEVNTLEA-IKADLDDEAPSTEGCGIM
Pythium irregulare	Oomycota	NCBI AKXZ02005003	MPTALSPIKSKRLDEKSTLQFLNGRLELYVLRVKEMEDAKNVAERELETIRQRMQA-DIDAIRLRSLEETRKKLDVLELDQKARLQVLEQEQHVLEVLKLRRAQVKEL-SEIKALAESLQVELTKERSNSKASKETLSVQTTQLQSARRRIKDLERENR-SEASLNDVTKLELDLRRKASADFDLTRDSEITIRKEMNQKHQEAALQWKK-DKTEERLQAVEKEVRHYFEGVVGFKNQVEDLSLELDSTKKELDRTAND-YYEESLQVRQQLTEKVAHLEREYREERTKFKEDRKKYESTLEHFRSSKFA-KEQEFNLDMDVKIALDAEITRRLDREETRFGLPTPTSTPNNAKQTPRKRK-SESVNAAKRVKRAPSSVQIHLDEKDRIVLENTSDKAVSLAGWEVVRGAEQVFRFP-STYSMRPKSKMTPVYSSKRNNKNAKDERKPNEDSFLATKFLSNASGDFAVLLT-ADGVPVSTRAEGLSAEEVNTLEA-IKADLDDEAPSTEGCGIM
Pythium iwayamai	Oomycota	NCBI AKYA02003977	MPTALSPIKSKRLDEKSTLQFLNGRLELYVLRVKEMEDAKNVAERELETIRQRMQA-DIDAIRLRSLEETRKKLDVLELDQKARLQVLEQEQHVLEVLKLRRAQVKEL-SEIKALAESLQVELTKERSNSKASKETLSVQTTQLQSARRRIKDLERENR-SEASLNDVTKLELDLRRKASADFDLTRDSEITIRKEMNQKHQEAALQWKK-DKTEERLQAVEKEVRHYFEGVVGFKNQVEDLSLELDSTKKELDRTAND-YYEESLQVRQQLTEKVAHLEREYREERTKFKEDRKKYESTLEHFRSSKFA-KEQEFNLDMDVKIALDAEITRRLDREETRFGLPTPTSTPNNAKQTPRKRK-SESVNAAKRVKRAPSSVQIHLDEKDRIVLENTSDKAVSLAGWEVVRGAEQVFRFP-STYSMRPKSKMTPVYSSKRNNKNAKDERKPNEDSFLATKFLSNASGDFAVLLT-ADGVPVSTRAEGLSAEEVNTLEA-IKADLDDEAPSTEGCGIM
Pythium vexans	Oomycota	NCBI AKY-C02003929	MPTALSPIKSKRMDEKSTLQFLNGRLELYVLRVKEMEDAKNVAERELETIRQRMQA-DIDNIRLTKLEETRKNLSDELEQNARLQVLEQEQHAEVLKLRRAQVKEL-SEYKALAEQLRLRELTREQSNASKETLSEQTTLQQLARRRKLKEL-ERENRQSASHLNDTSMEEELRKKASADFDLTRDGEIANIRKEMNAKHQEA-ALAQWKKTEERLLGVESETRSYFDNVVTGYKSQVEDLSSEESTKKELD-RTATDYEESLQVRQQLTEKVAHLEREYREDRTKFKEDRKKYESTLEHFRSSKFA-KEQEFNLDMDVKIALDAEITRRLDREESRVGLPTPTNPNAKQTPRKRK-SESVNAAKRVKRAPSSVQIHLDEKDRIVLENTSDKAVSLAGWEVVRGAEQVFRFP-STYSMRPKSKMTPVYSSKRNNKNAKDERKPNEDSFLATKFLSNASGDFAVLLT-ADGVPVSTRAEGLSAEEVNTLEA-IKADLDDEAPSTEGCGIM

Albugo laibachii	Oomycota	http://protists.ensembl.org/FR824125	MPSQQFSPIKSKRLDEKASLQSLNSRLEIYVMRVKEMEDAKTVAERELETIRDR-MQDDMDVVRMLTKELDDTRKLDHELDQKTRQLQLEQEQHVLEKLRQAVKEL-GDLKTLTEQLKSELSKERSNATAAKEELSRQTTELQAVRRRIKDLERERRGM-DAALSDATTELEHLRKKSATFDLVRDTELSNIRKEMNAKHQEAALANWKKDTEDRHAEVEKVRNYFEGVVTTSKSAEELSLDELSTKKE-LDRTANDYEEESLQVRQSLADKVAQLERDYREERKKFKDDRKMVE-TMVEKLYRKLKSEEEFNLDMDVKIALDAEITAYRRILDREETRVGLPTPKTPMEKGGKRSALFSGNRKRKVKQENDDPIRIQLNLEKDFVVFENQGGDKPVSLGNWEVVRGKLETQVFRFPSTYVIKPHARVTVFSAKRNKNARADV KPEDAFMTKFSWNHSGDWAVLYDDEGVPVSSLAQGLPKKEVEALEAAIRTESPEDDFEKGSSDACWIM
Albugo candida	Oomycota	NCBI CAI-W01001050	MSSSQFSPIKSKRLDEKASLQSLNSRLEIYVLRVEMEDAKTVAERELETIRDR-MQDDMDVVRMLTKELDDTRKLDHELDQKTRQLQLEQEQHVLEKLRQAVKEL-GDLKTLTEQLKSELSKERSNATAAKEELSRQTTELQAVRRRIKDLERERRGLD-SALSDATTELEHLRKKSATFDLVRDTELSNIRKEMNAKHQEAALANWKKDTEERIHAVEKVRNYFEGVVTTSKSAEELSLDELSTKKE-LDRTANDYEEESLQVRQSLADKVAQLERDYREERKKFKDDRKMVE-TMVEKLYRKLKSEEEFNLDMDVKIALDAEITAYRRILDREETRVGLPTPKTPMEKGGKRSALFSGNRKRKVKQENDDPIRIQLNLEKDFVVFENQGGDKPVSLGNWEVVRGKLETQVFRFPSTYVIKPHARVTVFSAKRNKNARADV KPEDAFMTKFSWNHSGDWAVLYDSDGSPVSSLAQGLPKKEVEALEAAIRSESPEDDFEKGSSDACWIM
Sapro-legnia diclina	Oomycota	NCBI EQC27961	MSSSLASKRESFSPMKSRLDEKTTLQALNNRLEMYVLRVKEVEDSKEIAEKELESIRDRMQLDLMTKTRLSKELEEVRRKKEEEREAKIRVQNLKEQYVNLQRLRQVKELENTKTYLETQVSDLAKEKESCKAAKEALASNATALQSERRKLDMEKERYKTLASLHDATSELEQLKQKTSEFSMTRDSEMTALRKMNAKHSEALAAWRRESEERQQVVENEVRSHFEGVVEGLRSQLEATKETTALRTDYERTANDYDESLKLRQSLTEKLAUVESQYRADRKKFQEDRRVYENIEAARKARKEKDELFNLDMDIKIALDAEISAYRILDREETRVGVEPKATTKKRKSTPQSSRAQKRKTHASGDFRIRQLNLDGRIILENTGATPLPLNGWQITSKSSNHVFSFPDDYVIQPGGKVSVLGKNTPTPEGEQDQMDFYAVKKAIWNPKGVALVKNAEVEVCSHAEGINEDEYDDEQEKEGNEGCGIM
Sapro-legnia parasitica	Oomycota	NCBI AD-CG01001748	MSSSLASKRESFSPMKSRLDEKTTLQALNNRLEMYVLRVKEVEDSKEIAEKELESIRDRMQLDLMTKTRLSKELEEVRRKKEEEREAKIRVQNLKEQYVNLQRLRQVKELENTKTYLETQVSDLAKEKESCKAAKEALASNATALQSERRKLDMEKERYKTLASLHDATSELEQLKQKTSEFSMTRDSEMTALRKMNAKHSEALAAWRRESEERQQVVENEVRSHFEGVVEGLRSQLEATKETTALRTDYERTANDYDESLKLRQSLTEKLAUVESQYRADRKKFQEDRRVYENIEAARKARKEKDELFNLDMDIKIALDAEISAYRILDREETRVGVEPKATTKKRKSTPQSSRAQKRKTHASGDFRIRQLNLDGRIILENTGATPLPLNGWQITSKSSNHVFSFPDDYVIQPGGKVSVLGKNTPTPEGEQDQMDFYAVKKAIWNPKGVALVKNAEVEVCSHAEGINEDEYDDEQEKEGNEGCGIM
Eury-chasma dicksonii	Oomycota	NCBI compound of gbl FR844597, FR840649, FR844317, FR844598, FR840650	IREPEDETKPTCNHRVALEKYTALEKNFADHKKRWKEEKALEHTLDVNRNARVEKEKQYNTLLDQKTLDAEINXXXLRLMQRSINGSILEREESRVGLPTPKANKRKRKARTPM TDKRAKRSNEPVIHDLDFEKDFIQLRNTSDSAISFS-GWTVLVEAPDGPHTKFSFENYSLKANTVTVLGGKKNKQKGDPKDKFWTNK-YMVKDSNIAEVRDAKAGELVYKVVYGDIDSPYTEETDGAIPGETACCV
Aphanomyces euteiches	Oomycota	http://www.polebio.fr/sv.ups-tlse.fr/aphano/Ae_2AL7951	NELEQLKQKTSEFSLTRDTEITAVRKMNAKHLEALAAWRRETEERMHNIEAVERSHFEGVVEGLRTQVEETNAELDQLKVEYERTANDYDESLKLRQSLTKLSTLETQYRNERKQFQEDRKYETENIDARQARLAKENFNLDMDIKIALDAEISAYRILDREESRVGLDSAKHKKRKLSTTSPATROLKRRKSHSTGDRLITVNLQGRITLNTGATPLPLNGWQITSKSSNHVFSFPDDYVIQPGGQVSVVSGRNA-TPSEEEKEAMDFYVIKAMWNPQADIALLTNPQDEV
Aphanomyces invadans	Oomycota	NCBI ETW04496	MASKRESFSPMKSRLDEKTTLQELNNRLEMYVLRVKEVQDSRDVAEKELETIR-ERMQMDLSMTKTRLSKELEDTRKLEFEIDQKTRQLQVLEQEQHTLVKLRQVKE-FGDIRVLELQVQAEALAKEKESKAAKEALALQTTSLQSAARRKLDLKENR-RKLTSSLSDDTTNELDQLKQKTSEFSLTRDTEITLVRKEMNAKHLEALAAWRRESEER-LHSVEAEVRAHFEGQIEQLRSQVDEVNLELDSLKVEYERTANDYDESLKLRQSLTD-KLSTIETQYRNERKQFQEDRKYETENIDARQARLAKETEFNLDMDIKIALDAEISAYRILDREESRVGLDQANHSKRRKASLTPVKSSTTRQHKRRKSHSTGVLRTYVNLQGRITLNTGATPLPLNGWQITSKATNVVFAFPEDYVIQPNGRVSVISGRNAAPTEEE E E K S M D F Y V I K A M W N T Q A D V A Q L T N P S G D V V S S Y A E G M Y V D D - D V D A A D T P V K E G C G I M
Aphanomyces astaci	Oomycota	NCBI XP_009842573	MASKRESFSPMKSRLDEKTTLQELNNRLEMYVLRVKEVQDSRDVAEKELETIR-DRMQMDLSMTKTRLSKELEDTRKLEFEIDQKTRQLQVLEQEQHTLVKLRQVKE-FGDIRVELDAVQGLAKEKESKAAKEALALQTTSLQSAARRKLDLKENR-KLASSLSDTTNELDQLKQKTSEFSLTRDTEITLVRKEMNAKHLEALAAWRRESEER-LHNEAVERSHFEGQIEGLRSQVEEANLELDSLKVEYERTANDYDESLKLRQSLTDKLS-TIETQYRNERKQFQEDRKYETENIDARQARLAKETEFNLDMDIKIALDAEISAYRILDREETRVGIDHANHSKRRKASLTPVKSSTSRQHKRRKSHSTGAVRITYVNLQGRITLNTGATPLPLAGWQITSKATNVVFAFPEDYVIQPSGHVSVISGRNAAPTEEEKE-SMDFYVIKAMWNSHADVAQLTNPAGDVSYSYAEGMTVDD-DVDVDAADTPVKEGCGIM
Cafeteria robergensis	Bicosoecida	MMET-SP0942 CAMN-T_0005518387	MASPGFAGSRRRVEEMRVLGSLNNRLEQHVQRQMTLAKERDALLERVA-GLVDHSTTMRDRAAAHQEEVSRNLNRQIRSLTGADKAETAVRRLQVVK-TLQQGVDDIDVVRQQRDSLAAANLNRNTRSDLANMEELLRDAEG-DRDQYMEGLEAAKAGSKDLELQAAETAHADRTRREDLDAQTLRRLRQADA-ASEASDDEETIRALRDELVELRARKVELEDAEAELEAELEGLRGGAGQIAAL-KAELERVAELAEALSQEQQLTERANDAEEKAAAAEAARAAEAERAVVAAA-REAAAAARAAEAESVDMRRRLSALAAQADASAAPGSAGGSSMIGERVRELE-EERSVLAELKQYERILDTDMAFLGINTPQRKQxxxGASHLQLWLDLHQEVVIV-NTGSDTVAMDGWHVLSQGGQRFDLPEGQSIKAGSCIAVWSGPDGAVNAATHKA-SLLWTERPVWNNKGDVARLVDSGGDIAAERAVSDKAAPLNLAASLR
Ochromonas sp. Strain CCMP1899	Ochrophyta	MMET-SP1177 CAMN-T_0007030967	MEEQNTIDDLRRDLESYVNSVAVIESENKQFKEIIAQLEKERELEAER-WLVLNKGKESWREQAIGHRQKQKQTMIECEKVMQSFQSEL-RTQYFOYQTRDSSVLLSCLRANSTRERAEHRCVFSMALF-GAIKPTVTVQDFIEDNHVSNRDTMEVSLTGCSLKFKFGSAGAKYKFPDDLIIHPTST-ISLWYGEANSIRQAQNRGSLYWHNSAITDFAESVGLVCKKGTTFVCSLANKE-RQARHATGQDPTDNSVHPSHFLKFNQFFSLKNDDEPLVGOQKDLNDDIYKPNP-STESKHSGNKSSTSSKSSSSSARKRHYSVIASPOPQTTDAITSEDKGSISSQSS-GLSVSSALSSHFTSPVLYRPLRLRPLRSLRQGLHFNKVLKQKQEGIGKAVAV-TVNSSEEGPVALEGWRLIAGVNPSEHFFLPSDAVIGRISITLSSCTDDTLWAQIK-LSDINPTPDMDISNGSKDVKMPRSEKKNKAKISSILSRDPEGVSDRQIAVAGL-SALSAGDILCLHLIDQLGRNICMVQGTVEVDDHTKDTTVFGDNANDNNKISVLSL-TNIAKNSKDCSISQ

Nanno-chloropsis gaditana	Ochrophyta	NCBI EWM28057	MVRTRKSDAAPGTPAIGKDKQNKSRGDDMYKQVEEYVMRVKLEITNKQLREEL-RRSGGDDTVGKGRICSTRMSSPNCVKKMRDLRQHIQLTESKIKEARRENQQLRE-QNFESFEDAREKFRSQGRIHGNLMAAKKADEFSGQILYLDKSVHRELEKTYKE-LEAERRRLQOEVEKVEDELNRDLKRAEERHAMLERNVENLQDENENLELEHQ-EGRHSLGVVVEEEKPGVPELEAEIRTLKEQVEGLTEDLRLTRMSAAEETESLKDEVE-EELRASLKAATSADRRERREARDSIVTQLEVCREQARKEKDRCEQLQKRAELAV-AQKAAEERARRLEDSPSRITVKARATALESTVKSLEGELESLRRRNRKNLGDVDA-LQEDKEALEDEVRRALQIEQQRRAEAVLARETESLERKIDRLQELDTANAKS-VAAVEASASAVARALQEVGPASAAARQGRGTSLLGGTEEGSEQRLVVEENARLRR-EVDVYRELAGSQHDQISLVYIQEVNAQADYLVVENRNLNLDVSLAGWALGTSLSSES-EIQFEFPEELMLSSRASLKVWWGARNISYKSRPTRGNLFWESDRIDIFQAQND-ELLMDAEGCEMSRMRIIVSDKRRRARAGPFVGTPTPDSKFRFRSSPWSGGSGMR-DNFVVRTSILDEDDQGESDMFMMAVEGFSFNEEEE
Corallo-myxa ten-era	Rhizaria	NCBI GAB-V01000374	MDDAPKTPARVSASLFEHALASPARQTRASEKDEMQLNRRLFEYINKQRERE-ARRSTEEERLMTAKQVVSREIDRARAEYQSQLAEVRKQRDQHVASKTHLQTTIAR-LQQNVYDLEGRLLKQEQAEHKELESRRDQLSNQLKDVNKELG-QNAKELIRSTEGDLQEQQKQNDALGEIDRLRKEYEDAVKNASEAEARRISIEEDLAI-ERTKAEKASQRFDEEMKQLSNIDKVENQLRQEFNTQLNSILAERQNGYEAEKNA-LANELRTLQYDEKVSAYRAQVEKSSQEAQLEAECKRFLETTQAAGAEVNAKAIADR-LSIAIQALQOELSQARDGPEREIEHKKRIIRRMKEAYQRKEAEFDALMDVKIA-LSMEIKRYRQMLDEEEDRLGI
Reticulo-myxa filosa	Rhizaria	NCBI ETO28743	MFNKTKYKLLKKQQRHRELEEKRELDLNRQLTMLVRHQREGRSEVNHLLKLL-LAEVEVDFRQKLDSEKHYQLLQDKFMAENEQLNLHNKTLREQLDTEKLNFTF-PQYFTFVNFVFLNPKPEVKMLRADMTELELKDKKLOAMQNLQELGLMNE-LVAVKASYKASSKNDQLAGQLQEQCDQWKSQFDMTCEELAKIKSRHCKQVGDG-YQRELNGKSEENKLEQLREYVKKQQLNQIQRQEFDTKMTFVLDLNLQYQ-QLFLLVFKREIQYQSEKEEWKIFKEEYQKMTQFKEINAQLVESNKKLESEI-NDLKLRLTRVREKAEIDAETHHLEEELEKARNDTDELQRKDDLEKTLVLS-SLQESLKQKETAFFELSGAKIQLDNEIAVENEQWRDLLCAFVCLFFIQCPPP-SFKKTEKEAGYVNPLESGRKRNDDGNNCSYSTPGLNRARQAARQELQDVG-TSDKNSPNNVRDNEIMEDEEEENGESEDESEQENIFDNNLPRGNITNPLG-SNNHNRKRYPDQKVDKTMSSNSESKHNIDISVWQIEKR
Reticulo-myxa filosa	Rhizaria	NCBI ETO24379	MSESQNDQCVSSQYQDQIGDLETSDAQHVTEKELSNLNRHLEMFVRHQRD-SRNEISRLRKLALVDTEVLDLRLQRKDAEKRRHQIQDKIMAENELSQNKTLQDQLE-T-TEKAYQSSESRRASLDEKVTCKMTEIELDQQLMLQNLQLEALT-NELITVKSSVKLMVKNDHEHTAEHLQAQQCDQWKNKFDVTEELAKVTYAF-FSFLQLTPNIDHCFFRNEDFGKFTSKKKKLLFNKCLNCRNRSFSDKL-VEKLEKLEQLNESTTKFQDMEDQLRKEFDVKMLEVWSGFIVFLFWSHP-FVHVRTYKFTQKRESQYQNEKEEWKMFKEEFQKLAHFKEENAQYL-ESNKLECEIVDTKARLARVRREKLELEAEKRSLVHFLKDVVETKGGKNIY-IRIFFQKKEEVEKSHSDVNLRQKDEEIEHEKTESLRLNDTIRQDIQYQQLSS-AKVQLDNEIA
Reticulo-myxa filosa	Rhizaria	NCBI ETO09514	IANLQNSIKQKDCYQKLSAAKVQLDNEIALYRNILSNAEKETDHFNPLGSALKKRK-TSDGKKHIDTPDAAYSRTAGNNVQHIEDKSSKTKYNKVESENINKSGNKGK-D-NGNKRKHLFEYEAKEIGETQQRANDTNEHNETDSYVSFAYLSLALFCSPPPP-KNKLFFVQKTPGQLCPLQFSAMDLSSTLEVQNTSEEPNLNGLFYLTNTDTSKQF-HLPSNRIAHIGFVFADERIKIIVGDTPGLKVKTGDLRWHDNVWSGLKDLVRLYNPQ-HEEIIARIEIAPEMPLPKPERTRENCLIIKLLSNLKMFFSTIFSSVSAHQIS-HFELAFHLVFLTKLNQ
Reticulo-myxa filosa	Rhizaria	NCBI ETO02958	EEEELEKSHNDLDELQRKDEEIREKTLIMSLQEQRQKDIQFDALSAGIKQLDNEIA-LKQLLEETEFNFRQRIKDESEKRFQLLQDKMQAENEQLTVQNNSLREQFENAEKYYII-IFVVVPPNRRARQNAESRCANLDQTVARKSLELEVKDQKLQALQKHLDSLNTDL-AVKASAKASLKNDEQLTGLQEQCDQWKAQKAFESNCEDLAKVLDCEFKARHKTQ-VEAYQHSLSDKSEIEKLEKLEKCEWKKSQDMENQMRKEVDIKMSE
Reticulo-myxa filosa	Rhizaria	NCBI ETO17410	LYRDILNEAEKDAECFFQTSNGCKKRVISINDKPRVAVTTPGLTRARQAANKDL-SRVVNSNSKCATPKNTEKKNETNVICKNENVYEESSNLGENDENQNRNRFNONS-GLHKSSSNKSSSVNSSHKRSFEFDKTNKSKINETDYSPTRVKLSSTFFVQVCPDP-PLQFSGMDLCSSTLELQNVSEEPVLDLKGFLTNSNKSIFILPHHHVLM-PKPKFNADPPFFVTTLDKIKIIVGNDGLNVGKDDLWVQENVWNGLEDE-IARLFNSKNEEVATVNITTEMIPEKIRARNKCLFM
Reticulo-myxa filosa	Rhizaria	NCBI ETO27405	MKRNNVECPFNRRDEIDPSIQQMDDKHKIRDLNTRLKMFVRHQEGRSEVAR-LKQLLEETEFNFRQRIKDESEKRFQLLQDKMQAENEQLTVQNNSLREQFENAEKYYII-IFVVVPPNRRARQNAESRCANLDQTVARKSLELEVKDQKLQALQKHLDSLNTDL-AVKASAKASLKNDEQLTGLQEQCDQWKAQKAFESNCEDLAKVLDCEFKARHKTQ-VEAYQHSLSDKSEIEKLEKLEKCEWKKSQDMENQMRKEVDIKMSE
Reticulo-myxa filosa	Rhizaria	NCBI ET-N99247	TPGEPNPLQFSTMDLCSSTLEVQNVSDHAILKGFYLTNLDKSKKFLLPGRVLPKS-FVIFYFYCFMFFFFIIIVSCMIMAVDQVQLNYAGERVKIIVGNSNLRVAEGDL-WWNESVWSGKTDEIARLYDSSNQEISRVESPEMLHEKVRSCKNCLIMCLIFFV
Ammonia sp. 2	Rhizaria	MMETSP 1384 CAMPEP_0 197081440	MSDEPGSSQLPIAKQALAQKRNVRDLNTRLEMFVRAQAEKTRAINQLKESLAR-QEMDYKSKLTKQRVQFDNVKLRKENQNLAYENKCTQEQFTNALSKEAFEEERL-LKSEKNAALSSSENSIIRTDLSRTQEQFDDMRKAFATMKYKYDSYEAERCFSFESR-MAKQKSKTEQLKELTKQKATLKEKATFERLIVEKNEIEHFTEAVKKLELEN-STTQSRLRQEFDNKLAEFVQKREEQYKQKEDWWRIFKEEFNRKLSRFKEA-NQELSHSNIKQSEIEQDLRARIKQKQKTELEVTVNRNNEEIEKLRNDLDDL-RRTKDAEIKLNSLIQERDRFAKELQFDELQAGMLQLDAEIELYRNILNEAE-RDCGYNSPMVMKTNITGTRNSRKRKRMINNYNMTPIFRKIITPGVSRAAKJQA-KDLKAFESDGDHEDQDMKEQESFDTYETPSGIEGALQFSGLDLNLKGM-I-EIQNMGEPEVSLDGFALTNQSGSMIYDLPTDMLQAKNTRLRIYVGEALFKDM-CGSESEKQQRRSRKFIQDYDGAUYVFWGNDVWTGTDTCARLYNPSQ-EEIARIEISPDMDIGVDKKNCLMM
Ammonia sp. 1	Rhizaria	MMETSP 1384 CAMPEP_0 197029186	MSKPEPGSVDRDVPAPKALQKRTVRLNTRLEMYVRAQAEKTRNIN-QLKEALARNEMDYRNKLLKQKIHQKVEKQKQIENLEYDKKCTQEQLEENAM-SAKTEYEERLISSEGRNAALASECQLKTDIGRLSKELEGLR-KTYASLKYKYSYIEIRNSFGEKMQKLSKSNQELLRELTANEQHLKSEKETFORVIE-KNDELNLKQKLELTHNSTQTRLRQEFQKLAEFVTRNREQQYETEKAEMWRI-FKEEFNRKLSRFKEANQELSHANVQKSEIIDLTRMSKLLKQKTELEVLRNRIE-DAEKLRDNLNRQKDEEIKQKNVILQQRDAYKAKELQFDELASIKMLQDEIEL-YRNILNDAEQEGYTSPLATHGGDKGTRNSRKRKRNYAMIHGTPMGPTFAENS-NREDDGMNTSNSSNSNSAALKVETPGITRAAKRAKKEVQKVLQSDSAGVQSS-DEMEEEEQVNEEEMKDEEDDSEFDEYSTPGNAEGARLQFSGLDLLKGMIEIQ-NQEQPISLRGYTSLNDSGTQFPLPNDVLDHNQKIRIYCGKQLYQEMQFDEED-AAPQIDKQKQIIGDFNGIYVSWGQDVTGQSNDCARLYNPNHIEARIEISPD-MVDMNASSKSGCFVM

Elphidium margaritaceum	Rhizaria	MMETSP 1385 CAMPEP_0 202725738	VRDLNTRLEMFVRAQCEKSREVNTLKEALARSEMEYRNKMKQQKLQYQDQIEK-SRKAIEINLEYDNKCTQQLDNGIASKKELEDRLMSSESARNASMSSENDDTLSTELQR-LKTELDTVRAYAAKYKYESYELERNSFDDKVNALKSKNDKLLRELTQTSTHKKSEKETFSRVLVEKNDEIETLRTENKQLKLANITTTQTRLRKEFDSKLAFFVQKREE-QYKQEKDEWVRIFKEEFNRKLSRFKEANQELKLANCKQLEENTELRERISKLKQ-LKMELEVQNRNNEEETEKLRTDLNLRQKDEELKQKNAIQQRDAYKAKELQF-DELAGIKMQLDSEIELYRSILNEAEQACGYVSPFNADKHGCTSGGPPSRKRKRVRH-MPTTGGHSHATHMTLGGHPFGGQHSHNSNTNGSSSNTTTAKQTTDGDG-GKKVETPGIGRAAKQAHTDIKQNSNTNGSSSNTTTAKQTTDGDGKKVETP-GIGRAAKQAHTDIKQIKQALESDTDGDANEAEKDAEEDDEEDCQSVSGSLGT-PGDLEGSALQFSGLDLVKGMIEIQNVGEHAMPLKAHALSNKEGTQQFPLPSHITL-DSHQKLRIVYGERYKAVSSAEDDAQEEERKRIVGSDGHYVFWTKDVTWGTGDS-KDCARLYNAQQQQIQHIEISPDMDVHVHSGKNGCFIM
Cercozoa sp. strain D1 2	Rhizaria	MMETSP CAMN- T_0037965 289	DSLNLGQEIIRADAGSANLKEIANKNARLQDQVVEYQAADT-DTKALLATERQKRKSLQAEINLNLARLKEQVLSPTQKSDAAGPLMYTQQVDFDKR-LQSVIADQCSAAERNRQIAMTEQKLYFQEKVKGYQAQLQVAGCEIEQQKQRCRE-LNENRLEAAAEDYNRREAEAMAFANALEVRIFGLEEAIKKEKSPKSIKAEKN-EIIRKQVEVLKEREYDELMIETKINLQVELQNYQSLQEQESRCGWENA-TNNEEHEEQGNELQKVVGHKRKRGGSHVSRRTKVSRRGKGVQSRRCVE-DTETQQLADNKRIRCVLSAGSDHRGLPLVLSSEFLDISEWFSFQVENSQ-QRLVYTPRDFSIIEGAGSFRVFISSPEVSVEEKTEAWWNWPEFAVVDTTSPDD-LD
Cercozoa sp. strain D1 1	Rhizaria	MMETSP CAMN- T_0037988 625	MPRTPASRRKGGLETPQPKPPNTPGSTTRKARLEEKSELODLNKRLEFY-ILKQREKDNAGSIQRTISENKEYEYENIKKLSLHETQVNNFRKQRDEL-SEIQLSTEENKQASMLKSYTSQLSADKTRGTDLDRVSQLTEEL-LEAQAATAVAEADAKSAHGLKKAESEIKHLQKELKDSQARDKDHAV-ELAAQSLAKNLQDDINIMSQKASESAKVEELESALSSTRANAEDK-LRALFGAQLKEILAEQRQFDEKAEGLQELRDSYEGKLEANEKCH-LLEKLEEAHKAEEAAGASSVVDTDASDAVAKVWENRIATLQNLQAE-ORTQADQIAQQTLREHSKEVAALKAIEVLEEQAANSASSTNSKLT-SEENSFRLSERTKLNKINTQKKEKSAEQARKIAALQKQAALFLE-QCESHKQENQTLNNEYISLMDVKVGLDMEISEFRRLSDEESRLQ-QUEEPAPIEEEEEEEEQKEVTRTNKKRGRQKASTPAPKAKRARAS-TAKRGGKAAVEEEEEPAKPTLIISGIDSKKWIETNSVKEAVCLDGVSL-KVVSKECFPPDVAPLKPKGTLKINLGSKAKKGKNAVWAKDDIFKEEDKV-YLMSPEGVGHSSVAWVND
Plasmodiophora brassicae	Rhizaria	NCBI CE095498	MTISREVEESPIRMQRDVERKNLEELNGRLEQYIMTQRRKASREAWAEALKA-VQTSAAQALADTTAKYETWLRQMLQRDEHASAREELKVRVSRMEETISRLKE-QIANEKYENELSTRVNSMAQLEARDLVVQLQENVRKLEQSLKVAENA-GRAAEEQLVEARELADQYSHEASAMRAKCKSLEQDLAIQVQSGEA-EQARLKEQVAKLTKSSAESEDITRTEFMKQLESVVEQYRHOQCNEDK-ARITRELKSHYMPKLTARQALEETCNREAELESKNVLTETEILLK-GQLETMQELKTVHEKRINELVAELDKERNVIHVKAIAEKEIEIEKLR-QRYARLETDFDNLMIDIKQMALTDIKYREILTEEEERRVGLTEPGRKR-KRAKVVEDPPTLPVQLAIDEENGCLVVENVSDDFSLDGVSLQSSG-FVFEFPEDTLLPAKSRVSVWIGEAPNQDDEEQEGDDEGQANQVL-VWEGVDCTELAEPEMCLCNEDGAVQDQIEVEAVRYASDRSSTRGCLLM
Spongospora subterranea	Rhizaria	UniProt A0A0H5R4 D7	MATNVNVEESPIRHPVERKNLEELNGRLQYIMKQRAKASREAFKEDV-VAIQQARVAIHANTKYYEEQLRQMRQQRDEHASAREELQVRVARQEAATVTR-LKDQIAEKKFQSDNLNGRLSSSLSSQLEAANAQVQRLQEQLRKTEHNFKSAE-SAGRAAEEQLIARDAADQYSHEASSLRAKLSFEQDMTINAKASEAEISLAK-IEQLSAGVTITTEDSCRQEFASQLSSIVEDYRKCNDNDKTRITRELKSHFVPLKAE-LRAAYEESCQHESELREINVTVEQVNLKKNQTLAEMKSLHEKRIAELESE-LDTERNVTHNDIAAKDKIEKLRHTCQRLETDFFDLMDVQIQLALTIKRYELLS-EEEHRLGLETPGRKRKRKSLPAPQTPALKLWVNDVENCLFVKNTSDKIVN-LNGMILQSGSFDPPDAYVGGQCEVCWIGETSNSLEASSEDRTNISWEGV-NPDLTAAPMQLIDQNGTVIDQFEVETISAGDDKNACLLM
Cercozoan plant parasite (EST of Beta vulgaris)	Rhizaria	NCBI KM- S94577	VRPRLSISDMTAAADVGESPVRLQRDVEQKNLEELNARLEKYILKQRAKAS-REAFKEDLAKIQQARQAIHANTKYYEEQLRMLRQQRDEHASAREELQIRV-SRQEAIVTRLKDLTDEKQYQNDLNQVSGLSQLEEARGONSLQDQLRKT-VHALKVAESAGRSAAEQLAAREGADKSSHEASALRAKCKSL
Cercozoan plant parasite (EST of Festuca arundinacea)	Rhizaria	NCBI GT045861	MTVADVGESPVRLHRDVEQKNLEELNARLEKYIFKQRAKASRESFEKD-LANIQQSARLAIHNSNTKYYEEQLRMLRQQRDEHASAREELQIRVSRQEAIVTR-LKDQLEEKYQNDLNQVSSLKTVEEAHQGNSKLDQDLRKTALHVAESAGR-SAAEQLAAREADTADKFSHEASSLRAKCKSYEQDTSISANENETIKLETKIDK-LQAQIATEDNCRSEFSQSLAKIVEDYRKCNDNDKTRITRELKSHYMPKLTREL-LSLDEAMRRENDLREANIALTDQVSLKREMQSLNDIRTVQEKRIQELTEELDRE-RNFTHSDVIGAKDKIEIHLRHECQRLECFDLDMDIKIQLALTIKRYELLS-EEEHRLGLETPGRKRKRKSLPPEGPSLQGLVHEQLDNCLLNKASGETMNVNGFVLK-SGSLDFHFPDNIQLPAYGEICVWIGEEVPEGMWVAS
LTD-like domain-containing proteins:			
Nannochloropsis gaditana	Ochrophyta	Uniprot tr W7TNN4	MRRLEIPLWDECDQDGTARSQCNPRRRSILYRSPPIYSDVDVKREIITIS-NPESVPDLSGHSCLKDQGHRYEFEEGYTLAPGAELHVY-CAPGKYGVVEEERPGELQILLWTTREGRPRRREVLNOEGDRVTLIDQNG-MEIAALEVTVNGDDEEEGEREKRGTKAERREIEIKRLYLRLVLEVLQLYKASVVT-WLQGSRLMLVLAIFVLLNDLHSHYVLEVGGFALDMAARYLDFVRLPQRSRQVFA-LVADRDSASLLAVLMTGLVETGGPLYGPNPVHVRVRLGLGMLLIEVAFWFEVLE-VSGETRLAADAAPVYWRQGPAPAGEEVSQGTGPYRKLWLRQHLGRCPALITFATGG-NHLFLVLAVVRLAHAQAGGSEGLVVRVSVASRIMAYLPLPRAIVSNWQFLLIACL-SAMYRCILIAWELLDLLESKL
Ectocarpus siliculosus	Ochrophyta	Uniprot tr D8LHR1	MATAAFHVSPVPSRLLPLGLDTHQVQEEEDVPTDTPPGSHWLFRRPPPIA-HADLEAETVTRNPSVWRGANLAGYTLTDRLQRHTYHFFPERFVLRARTSLTYL-CYCPGKSEFAGLHNLPOPFLLWHNQDGLRRKEVMENSGETLVLS-DAKGNVAVLEVPDGTGDEETSRTPGVKRSQVLAARLTVLVLSCYL-RVCTGLAAARVTVHPDVFLLVTAHALVFDILSRWASSRPGVPMDFGVVATMGDRFSYLVLLGSLVLLDKDPRHSQIFGGMLALDMTANWLQLSVASISNGSHPLAGTLVSGRAAAPDFATHIMLTHPVELTIVL-SLGSFAFLVWSYLVASSDPLGLRAMVAKAGDFLSSLFTELLSSFS-SSSSSSSSSLSSDIADDEEVAISLISGEETTTAAEDGGVFGFEPNRI-DSAALGLGEDLVREAEEDDGGSPFSTAWKAVWMTLMVCCAAAR-QLLSCIQALISMSLSSGVIADAAGKDMQRRHRPRSSSSGGSSRRSRRAR

Auxenochlorella proto-thecooides	Chlorophyta	NCBI XP_011395 790	MCRSSTSRVPHPGYQLTVSKASGKASAWTFPAGVSVPGNGYLLIADGLDVTSG-SEVHAGIKLGGSSQGLSLDASGASIGSWDLPELADDQAYGVAEGDVALDGPSTSRVLVYLQPTTANSANGAATGPVWARVEDPAPRPNGGSAIPISAVVTPQQAT-VAGATLTYVIGYGSPVSDMAATGGTYTADIPAAEAAAGQLVRYVYATDAD- GGRTNVVPAKDSNGAVQYFTIVADPTDSASLPLELYSENDRAFSTGPSS- NAIAVPGGGKGDVTCGSVWFNFTFYDNVSIIRRGATSLGWPKPKIKIKSENG- KVFVLLGLDYEVSFAFLNSNWFEFGANTFTREPLVWEAFNQMGVSELLS- FHVHVRFNKGKYFRFSLQEEWSKDALERNFDFTSSTGGYGFVWKSNGIG- TNLRFVTTTDLPVVYELESSKQDAAVQALEAFAGLAGGGPVARSKYL- FDVNLQVINYLAAQTLILNQDRCTKNFYVYLDASSGQWSMLPWVDES- FGIDRGLGGQPAPDYCLACEQWNSPLYCDRNHTQDLVISTPYGRISQVPT- GATSGTPASRRRRSSSLQALGTPRLGRKLLQSLFQDQTTVDTGAGQDL- SDYDANLTSGRPASGAIGSNYLNADAVLSLPRTRAMYMRRLRTLMDEFYDG- RLAGIVTTLHTQIREAAIRDNAKWNWSGNPDTGYQLIKEQLPLRQKQD- VYGPQGDHPPIGAAQAGAGLTIGQAHTDAGSWLELTPGAEAVDWSG- WALTDGVDFTFRPGTVVDAGDVIIVAAIDILQFKASYGGGGYVVGPLPQPLTSS- SPSVSLVKA
Helico-sporidium sp.	Chlorophyta	NCBI KD-D75179	MVYISEILLDSSESSSGWELGNSGTNPVTLDTLSIGVLDGSPSAARLPNGNITH- PQSFHIVSLGRHGIAGSVPTLRVPAKGGGSIWDDQQLAGRLTYP- PILLKQDQAFGLTLPPTYNESLIQAASVNSNPSGSLLOSPTGAPNSRLRQGP- PAIVGTTSGLAALPNQGNATISAVALPQGAIEQSQVLIYQQGYSEPQTLMLP- CGHTATLPSHLFTPGALVWVQVAATDSQRTALDPPDSEGAARKHGIIRVADGET- FTPLLLQSCPDNAPWSTGPEINAAGRTGGKGGVMGCSMVFSDVDFYVVR- RRGASSTAWPKPKIKITSNKQKGLFRYPQGTYRIKALYNSNWMEPGENSYRE- PLMWLDALGAMGVPAALGAFHVVFNHAYFRFSLTEEWSLDALAEARLWSR- APKPVWLKANSIWSNLRWDVSRRIEYQYGLEESSPSGTAGAVTELEQLMGLAG- GDSLPRSEYLFQALHLPVINYMAAQTLLVNQDRCTKNYFMLEAATGQWSVLP- WDVESGFDIRGYGGVPAPDYCLACEQWNSPLYCDRQHVQDVPLAGTEGTWD- FYGGPVPTGGGQFFSFLFTQTPQPQQSQFQVQVQSSPQVQSSPQTLLQ- PATVSTAEAVAPQDGRARAPQPQDPGAFMRWIFGRRRLSSTGRSLKQAPET- VAPEPESSPQTIQSDAPDFYNVDQTKNRNAMYNYSSSFVWLDLAVLGH- RTRSMYMRRLRSLMDEFYSGKLEALVHKWNNVSSMEIARDCEKWNCPDRDPTL- YQQLIVEQLPIRAQLYDYSVGGTHLPLPAAAAEAQQTVAEVMFMECLSSGL- LRNPADAVLDSHKLVAQLDYLPGTVIPANDTLIVLTDVVGFLRRHPGK- GYFVVGPAQRGTERAPKIVPK
Chlamydo-monas reinhardtii	Chlorophyta	NCBI XP_001699 673	RCTSGWELRLLDHDEEDGNDSSRNSTWTFPPDTSIAPNAYLVIIYASGNKVSAG- NATAGSAAAPASHASFPLPAAAGVGQVLLGPGGVVVERVQL- PRLVVLQQAAPLAGPPASHNLTSSRRQYANVSFVGPAAADGVRDAATAAPLTYL- SAATPQADNAPLAIGPIITSVSRPTGDRPPADGADLVNITLRLPNLVNPDGAGVSL- YVYNWQGEQELAAQEGPAGADGELLYTAIAPAAAFKAGDMVWRARAADTAGY- LRSALPRPPPSAGTAAAEWPPAMAAAAALGQAQGGEGQEQEPEHYGGYTAIAVLE- DQVITPDVPILWFSNAAEATSVDGLGQVLVYGGTLDHNVARRRGTVALSWPKPK- KFLPSRDFRYSDTAFEVSEFGLSQFVWYELGERSYMKEPLALQFMREAGVIAPTSF- HVHVRNLNGOYGLFAFVEVLDDTYLQRHLGPPSGPLFKSVSGELSRLWDLPLKE- YPSFWKEENRDKVAADWDALRNFSRGLAGGGPEPRSEFVQSNVLPSPMINNMAA- QTLVNNMDRCKTFFMYLHPHTREWYMLPMDMDGSGFGQDNGLGKPKDLNDYCV- LACEQWNSPLYCDSEHPQDLRTRTPWGTAVKLDNSYQGAARVKNPSNAPGRS- RALLQDEADTAANINKSPMIAAGPTTITKPPPRGWDNDPRVTLTQTPNGPAGTY- NHLDAILDNATRAMYLRLRLTLADTLFASGRIAGMANATYTRIKPLADLAKKWS- SGISVDRGQQTTFPIRTDQLLGSLYGPTGRRPLPTSQKPEQGTAEVDLSGW- QLRPAAAAAAGANLTLPAGTVVAPSSVAVTDLATARELQRRAGAAAAFAGGA- AAAAEYVELSTSAQSGSKAGGSAPGSGGGAGKVELVDPGSAVAGAGGGV- AMQTKK
Mono-raphidium neglectum	Chlorophyta	NCBI XP_013901 328	MPNYWGKANRKNNPEDWQLLANLAKLAGGGLVPRSRFIFDALDLPVINE- MALETLGNMDRCKTNNYVWYVNPQKQWTRIPWDLAASMGQDNLGGAPGNLY- CLLAGEQWSSPLYCDSEHPQDLVTRTPWGTAVVGLNYSYFPGPGRGAAGGSTAQT- GRRLLQGGGAGGKLAEGNQQLSFPAPRGWDDPDRVFTGTPSPNGAGPSYN- HLTDAVLDVNAATRAMYLRLRLTLADKLYGPEPKGRRLRQIIDEWTARIKDAAAR- DNAKWARGDTRGAKQLLQTYAPGGQRLPPLGQPDSSKAKISVTKVETSPPA- GEHYIQIANGSPDEAVDASGWTLSGGDASWALPPGTILPGGASAFVTDWPPAF- KARKASPCKGEGLLVQVQAPQMMLALAGPGQWSVKDAAGAEVSKASL
Ostreo-coccus tauri	Chlorophyta	NCBI CE-F99931	MDARARGARPRRRRANEGRLGAAVVVTCALLGTGTGAGAITVDVAVS- GAGLGSQDAVEVRVVEGGGAAATARATCALGWGAERTTGMRYDNGDERGRM- RFTGWCDDTNGAPEGERVVKVSVAAGTARATSARWGAALERRVKEETPLPIL- HVVTPDYDKITDAGERVAVYFEGRYDDVFMRRRRGNSRNEAVIGVELSAANW- PKRKFKDLFGRRRRFKLNANMSRVSEINLNSHYQEPGEETYMRENLFMAILRRAG- VPAPLTRHVRVNVNEYYGLWSLIEQVDRFLKRNLFTRDGYSLYKAVNWKYS- NLRRAGDPSLPCPYATPDYRREWMNDGCEPIYKASKPRDKWDDLWQLTNDVIE- RVRRNPEGEAYLLFDHLNLPSLVNEMAAQTLMLGADRCTKNYMHDRDWTGEW- RIPWVDEDFPDKRYGTSKPCSECAATANLYCILSCEKFNSPLYCDRNHPQDIF- YVWDGSRPEQDPRSTYNVLIDVMLSVVWPKEMYLTRLRLTMDQILATSFVDDYVR- RKVNLRKDALDRSCKWVGVGAGRAIDQGVNLLDVIIVPTRRQKLEFYQSYMIPPS- IPSNARVYVSHAQRSGDESIVKLSNPNFGAVDVSDDVIVSTPGWTTWTLKPGSV- MGPGRVLFVVKNAKEFRNRATWARREYPQGLFVQGNFVRDLDVDDPRMFTVTKPA- RL
Ostreo-coccus lu-cimarinus	Chlorophyta	NCBI XP_001420 675	MTTDEGERVGVYFEGRYDDVLMRRRGNSRKEAVIGVDLSAANVPKHKFLDFH- GRVFKFDENRERKVEEINLQSHYQEPGEETYLRNLGFAILRRAGV- PASLTKHVQVRNNEFYGLFSLIEQVDSTFLRNNLDPEGALYKAVNWKYSNL- RAGDPNLPCPYATPDYKREWMDGCEPIYKASKANRDNWDDLWELTQ- VIERVRRNPGEAYLLYDHTNLPALVNEMAAQTLMLGADRCTKNYMH- KDWTEGWSRIPWVDEDFVPGDKRYGIDLCKSCECKKSTAYCILSCEKF- NSPLYCDRNHPQDIFYAQSPQDPRSTYNVLIDVMLAVWVWPKEMYFTLR- LMDLQILATSFIDYARRTLNIIRKDALRDKWVGVGVAIDQGVLLQLLQ- SQ-IVPKRRDQLFNQYSMIPSTSPSNARVYVSHAQKNGVEAYVKISPNNGYA- VDVSFVIVETRGWYKYLKPGTVLGAGRTLFWVRDAKEFRRRATWARR- EYEGVVFQGNFQSDLDVDDVNMSVTRTSERVAALNNGRGAFRRFRP- RVGFNTSCALPDVGPWSEYGDCEAGKGLWCQVQNAARVRDRD- VNLGLDNPDLPQCLVHSYEIDRRCASLECGTTPRPNSTACRIPSNESFVR- LSRNDQSVIFSPIGLNTRTWEDWYSYKVEVAPHDAGVDGASKSLKILVH- SGGLAFRATARNDTRTSAHFWIARHAPPQPLPRLVDFLESTSDNTTTT- RAAPGEFLRGIARAQSDWNAGCYVEIJADFAGFDDSPPTAAREPSSVKIV- LKSAFGSPTRLWLGRAVAVFSSHENAP

Emiliania huxleyi	Haptophyta	NCBI XP_005763 134	MPPPPPRGWGGGLAGSLHRVDPAAARPAVLARAVRDASSGLAVHLCGVSHVD-PASATVRYAEALAAAPGGLGAVAVEDEPRALLTRAARAALHGLPPERI-RAEGEGLVRHALFGLAEVHAWAAEAGATLDGAERAEAAEAAEELGARVACGGT-PAPPPHVLTRRGAPHLTVHPRVRRRAAVRLAYWLRVRLADPGHDERSPAEL-HVVASNRLAEAFRRASPPVAPSAHEAVLLRKKDAHFAAALRTVCADAAARAGGEPFP-AVLAVVGATHVEGIAERLQAGDGTPRNSVPLLLVCSVANVIINEVSDKGLTGCDG-ADWVVLNVGADVDLDGWMCLDSCGGQDNVNGTIAMAGTIAPGEYRLVCPQAQ-PKWWGDSDTLSLVDEHALEADITLGGDGEFGKSWARWPDSTGPFYRYVPTGAPYADSVIINEVSDKGLTGACDQDQWVVLNVGAEVLDLGDWVSLDSDGCDNDDITL-SAMVAVPSGYLLVCPQARGWMMHRKWISDSDTITLINAQGMQADATMLGGDGEF-EKTWARWPDSTGSGFRYTFHTPTGAPYAKSVIINEVSDKGLTGCDGADWVVLNV-GADVDLDGWMCLDSCGGQDNVNGTIAMAGTIAPGEYRLVCPQAQPKWWGDSDT-ISLVDEHALEADITLGGDGEFGKSWARWPDSTGPFYRYVPTGAPYAEAVETLQ-VNEVADKGDPTDPCDGDWVWELRNPTHGALPLSGLLLSDDKGPHNEKALGMDGPG-GCFFSLGTGELLRLCRNGNASMGGANAGCGFAFGIGKNDVSVLHQDAAADRVLV-DEAVGCCAGDANLSYGRPPRGGAFSILPWRTGTANAVAAARQSPRVTPLPSVL-RPLSLSSPPFSSVVFEEGGGAGFGKDRFPEVVLGPPQAGNAGSLDVLVSLGE-AGEVLRFDGIELVDGDPDLLVFENPFTNWYETGVVAVSADGMTWTFEWCNDEP-DATSLYPGCAGVGLVWSNPNDIDPTDPEAGGDAFDLADLIPPAHLADLGVRG-FRYVRIRDSGRNSYGGISGFDLDAIAAVNGEPLAVSTALEPVAEQIMRSDCQDPA-DLDDGDAFRADAVYVAEAVWSGSKQWRSRAMPCCGDPAGDFDGDGAFRLADAVY-LAEVWAGKVLGPTSDHRLIAWKIPMEITEHEVETAGTTVAFEWQNYHNVMLSTQ-AEYDACSRTGEIVAGTEAPQNKVKIVVK
Emiliania huxleyi	Haptophyta	NCBI XP_005784 342	MAAVTCTSTPPPAEYDDACTEPRVIIEVAASGADGEPCHDDWVLFNPHPTR-SAPLLGLVLSDDHGDFRSDRLALGASGCPSTLANRSYLLCRADSPCGFAFGI-GRHDVVLFNASGSPDETACCCANSQRSFGRSPGGAWTELDRTPGAPNTG-VAHAPPAPPTAPLAGSALGLDRLHVATLVPVPPRPGAAAVDLSGVAAAWED-LSSGGWRATLQTVVNSEPAVIEVEVATSAASPPVLRRTTFLLEGFEDTEGIAL-PSGQVAIAQERRMAVSVLLPPASSSEGDASHLETRLGLGAANTFDRLPAERN-KGLEGVAYDPEQGLLYAEKSPIRVVALNLTGAVADLFDAPSVLGRHSDLAGAV-FAPRQRVLLLSQESRLVVEVRAVQRISGTOPEGVALSPDLSTFLASEPNELLYV-SAEKKGTFIFKAVGASGQTAGVEARSSYA
Emiliania huxleyi	Haptophyta	NCBI XP_005771 646	MAAVTCTSTPPPAEYDDACTEPRVIIEVAASGADGEPCHDDWVLFNPHPTR-SAPLLGLVLSDDHGDFRSDRLALGASGCPSTLANRSYLLCRADSPCGFAFGI-GRHDVVLFNASGSPDETACCCANSQRSFGRSPGGAWTELDRTPGAPNTG-VAHAPPAPPTAPLAGSALGLDRLHVATLVPVPPRPGAAAVDLSGVAAAWED-LSSGGWRATLQTVVNSEPAVIEVEVATSAASPPVLRRTTFLLEGFEDTEGIAL-PSGQVAIAQERRMAVSVLLPPASSSEGDASHLETRLGLGAANTFDRLPAERN-KGLEGVAYDPEQGLLYAEKSPIRVVALNLTGAVADLFDAPSVLGRHSDLAGAV-FAPRQRVLLLSQESRLVVEVRAVQRISGTOPEGVALSPDLSTFLASEPNELLYV-SAEKKGTFIFKAVGASGQTAGVEARSSYA
Emiliania huxleyi	Haptophyta	NCBI XP_005789 335	MTTTPVSVLLMFLNAAVGMRRLLSSAVLINEVTNQGSETDLCAGED-WVELNVVVGAPSSVAGMALDDSSSPDLGIVSDLPAEVAAGKALVLGGQSC-PAEIPAHGYLVCKGEKAFTFDAAGAVVASYLEGCGFGFNGNTDDVLFSSAEG-NVTADMIVDRTGVWVNDALGNNPTAAQSLGRNGAGDSAVLINEVTNQGSE-TDLCAGEDWVVLNVVVGAPSSVAGMALDDSSSPDLGIVSDLPAEVAAG-KALVLGGQSCPAEIPANGYLVVCKGEKAFTFDAAGAVVASYLEGCGFGF-NGNSDDVLFSSAEGNVTADMIVDRTGVWVNDALGNNPTAAQSLGRNGA-GDSAVLINEVTNQGSETDLCAGEDWVVLNVVVGAPSSVAGMALDDSSSP-SDLGIVSDLPAEVAAGKALVLGGQSCPAEIPAHGYLVVCKGEKAFTFDAAG-AVVASYLEGCGFGFNGNTDDVLFSSAEGNVTADMIVDRTGVWVNDAL-GNNPTAAQSLGRNGAGDSAVLINEVTNQGSETDLCAGEDWVVLNVVVG-APSSVAGMALDDSSSPDLGIVSDLPAEVAAGKALVLGGQSCPAEIPAHG-YLVVCKGEKAFTFDAAGAVVASYLEGCGFGFNGNTDDVLFSSAEGNVT-ADMIVDRTGVWVNDALGNNPTAAQSLGRINWAESAVVVFDETRTGLVAN-PFFVEQWDLSESSKAEDLSGVAIIPGGAGDIDSIIFCENNPLNSDVAKGE-CLVRLFEYTLPTPSSPPLMRMVRMDGDDVEGLALVEKSTGVRLVTE-ENRNLNVLITLPLMGAADVVEATVDYAAVPPEDMYAVSLVDDAKSTNR-GLGVAYDPAGFFYTLTEELPMRVLVDMRTRETVLLSGPPRLASGELRI-QSHSSIRNDEQALLSGVATDIAGMTYDYKSRSLIFVSHGIDDGSGIPAT-VFTSDLSGKLLSGLPLQGAQPEGIVMKDASTMYIIGEPAEVGMVYSKKA-GVLAASPPSSPLRGDCPPGLEVACRVWTVNPAFA
Dictyochophyceae sp.	Stramenopiles	MMETSP CAMN-T_0037720 169	MSSSGMHRAFQSRMIFLFPERTVVEDELPGWFLRILAPMGYAPDIIWVHEISL-DRGFVIMGSSWEVDLTFGHVNEERRQRFDFPDGFLPANGRVKIVCYGTA-LSPSRSAAEVELANGACLAPLTYWEHERHEPVLNRHGDAILYVDRYARIVTA-LAKCGDGASKRYTPSRLELHNSFQNLGLLGLLRAFVAGGCVFGLWPLLGLH-ASAQARATPFPSPNTPESAVSGGFDALAFGSLNLEAVLVCWVAVAGGLDLAERLCA-ERCGLPLLAHVALGDRLNAFQLYGLSALAYDGNASLAFGLGLLDFTTWVW-LRPGSSVSPSTHDESAATTTAAATSTAVVKKPKFLLALASAPHLPALCLAS-EGFLVLCVSHPLVWPTLSTAPTPLLWSGVASRLAAFCARHLLQVLQLQVALA-TLLGGGTAAGAASYPAGGGGRGASASSYLSGSRGSSPSSPSSRRLGDHYSD-APAPVGSSEMWAIGALVKKGRKTLTSFDGRVLSATAPVKHDSRGRGRAPTPAA-AAALPRRQGLDRTSQSHVTHGASIASQSPGSSSTNVESEGEDSSGDDGND-DEDEGMADVQTKPAGQSSVRRRGRPNRNSPVPRLNKTMRATRLTSLN
Dictyochophyceae sp.	Stramenopiles	MMETSP CAMN-T_0037655 291	VCCWQDPMMAQGTMPRRSTTDQSEGLDPAFLDAFQTRGLTRTIEVYERQR-RILPFQATWLGSLFQKKHGVRIYDIDLQEDWVLIENFNSPVMGAYLTCDEEREN-KLEFPVDLQPGRYLKVICAPGRKLNLDVFEEDGRTSLFLWKTTPRGAPRSE-HVLREDGDCMYLCDLSNEVVSFAFKTDDGAFKEYGALNQAFQKLRWGLGLM-FVFAIRALLGTAGYCAIHTNLSSEPSAHPFKVVGCMVGAVALDLLQREALLYAH-LHDFFGALAIADRLHAMVIYGLTALFHGDYCALFLLFLTLDLASCWFLQFASD-CGELEMKAVAGAIARAPSLTLIGLASEAFLLLYMDHGHVQFRLYGQIAPAAA-AIGWSSFPVLTLPVVKDVLMAFLTKQIVLSQLWFCMGLMNTTFFCVGDDA-KTYLQHALESQRPNLELLRNSGSLQMQPSE
Trichosphaerium sp.	Amoebozoa	MMETSP CAMN-T_0008558 515	TKKAKIDNKTCPIALHKLKDDVVTLINKTTKEVLSGWSLKSQSGGQEYSF-PDDATLKPQCTLSVWSGAKNSNKAQGGKSSLFTKSYVWKSQDGCACVLY-NAEGEPVDEVEVTKEREPPSIAIQKLDADSSQCSIIINRGGEDVDIGGYFLTRMRL-HTFTFPKKTVLKAGDVTTVWSGTHKTKNVPSSFCWKSQSHDGTDTDAVLYD-NHCVLVHKVTEMIHLGELTERNRVPKYTSEPGEYPCII
Stereomyxa ramosa	Amoebozoa	MMETSP CAMN-T_0015333 301	MQAEKRIARKSKDNTQEGGATKAKIDSASVLSLSDLIADSVTLNLSGLEEVDSG-WIIDSQVGDQSYEIPAGTVIGPGATLTVWSGAKNKGKDNPPNSFFWTAKYL-WNNKGDVSVLKNADGDEVDKKEEPEIEEPVAIHSLLDKRECAVINQNSNT-DQDISGWLNSVVGDDTFLFPENTILKAGQSIWISGPESDQKVPFSSWTRRY-IWNDNGDTAALYNKAGNLVSKRVEFPNVTSPHVEPTTKTG

Tannerella sp.	Bacteria	Uniprot tr W2CEI4	MASPKGAKGLPETDYVELYNTTDDHTISLKGWSFIYDGTVIRLPDVLPAANRYAV-LYRKGSPLNKKVGLGFSNPLNMSDEGKQLTKDSSGTVIHSYTYPKAKAGR-SIERGEGDKWHLSSDPRGGTPGEENSVPASVPTPKPDDPATSPDTPSAPSSPS-DTVKSGDIVINEVMASPKGAKGLPETDYVELYNTTDDHTISLKGWSFIYDGTVIR-LPDVPLAANHYAVLYRKGSPLPKDKVGLGFSNPLNMSDEGKRLTKDSSG-TVIHSYTYPKAKAGRSIERGEGDKWHLSTDPGGTPEEENSEGAPDKPEKE-KSSPGDVLINEMADPHGLTKLPATEYVELHNTTDEHEINLEGWAFVYDKTSIPL-PDAELSAGGYAVLYKAGREISVADGAAEVAVKRFPANMINAGKPLALKDPSTV-IHSYTYPKAKAGRSIERGEGDKWHLSTNPRGGTPEEENSEGAPDKPEKE-KSSPGDVLINEMADSRGLTKLPATEYVELHNTTDEHEINLEGWAFVYDKTSIPL-PDAELPAGGYAVLYKAGREISVADGAAEVAVKRFPANMNAGKPLTKDPSGTV-IYSYAYPKAKAGRSIERGEGDKWHLSSDPRGGTPEEENSEGAPDKPEKE-KSSPGDVLINEMADPRGLTKLPATEYVELHNTTDEHEINLEGWAFVYDKTSIPL-PDAELPAGGYAVLYKAGREISVADGAAEVAVKRFPANMINAGKPLALKDPSTV-IHSYTYPKAKAGRSIERGEGDKWHLSSDPRGGTPEEENSEGAPDKPEKE-KSSPGDVLINEMADPHGLTKLPATEYVELHNTTDEHEINLEGWAFVYDKTSIPL-PD-ELPAGGYAVLYKAGREISVADGAAEVAVKRFPANMINAGKPLALKDPSTV-IHSYTYPKAKAGRSIERGEGYKWHLSDDPRGGTPEEENSGTPDRPDESEKPKA-SPGDVLINEMADPRGLTKLPATEYVELHNTTDEHEINLEGWTFVYDKTSIPLPD-TELSAGGYAVLYKAGREISVADGAAEVAVKRFPANMINAGKPLALKDPSTV-IHSYTYPKAKAGRSIERGEGDKWHLSTNPRGGTPEEENSEGAPDKPEKE-KSSPGDVLINEMADPRGLTKLPATEYVELHNTTDEHEINLEGWAFVYDKTSIPL-PDAELPAGGYAVLYKAGREISVADGAAEVAVKRFPANMNAGKPLALKDPSTV-IHSYTYPKAKAGRSIERGEGDKWHLSTDPGGTPEEENSEGALDKPEDEPKV-SPGDVLINEMADPRGLTKLPATEYVELHNTTDEHEINLEGWTFVYDKTSIPLPD-DAELPAGGYAVLYKAGREISVADGAAEVAVKRFPANMNAGKPLALKDPSTV-IHSYTYPKAKAGRSIERGEGDKWHLSTDPGGTPEEENSEGALDKPEDEPKV-SPGDVLINEMADPRGLTKLPATEYVELHNTTDEHEINLEGWTFVYDKTSIPLPD-DAELPAGGYAVLYKAGREISVADGAAEVAVKRFPANMNAGKPLALKDPSTV-IHSYTYPKAKAGRSIERGEGDTWHLSSDPRGGTPEEENSEGAPDKPEDEPKV-NEPNATEQVEPREIVLNEILFDPQPRGSEYIELYNRSDRTLSTHGLAIALRKS-DGHLGTRHSLTSLATLAPGDYLVLTSDPNGVTSIIRTALDVRIRFKLPALNNQGA-TIVLLRTADSTVDEVTYSAKWHSSAVKIRRGVALERISPDGSSQEAANWTA-SSEETGYGTPGYKNSQSGTSSQIEEGATISEPEYNASTRDYLIRYRMDKPDYR-CQMAVYSSNGQKVAVIANNQLLTPEGEIRWDGAGLTPGVYIFVYVELYHPDGSSQH-RKPLLVH
Vibrio nigripulchritudo	Bacteria	Uniprot tr U4KER2	MALTIQIYQEQKLTLEIRLATALTKQQLDELWEDLNRIQETSQVDIGELDFKQGT-DHLDEYIDIVNRGGLAIDLTSWKILAGSPDQEFEPGGSVLAPYKIRVATSGDSEFS-FQSDKPIWNNHGDTALLNPHGQSVSTLAYGGDAYPDVLT-NIYFDGEEKHTEGDEYVEISNVSDNTVDIGAWRVESVNRQNTTTFPEGIRLQAQSS-VKVFTNKEHLESEGEFSFCSPRAIWNNEQGHCKLFDYLDHEVGSYQY
Microcystis aeruginosa	Bacteria	Uniprot tr I4IMT5	MTKSTITGGNELKTRGVSDSRSNFYLVLDLNVPSGKGVLSWEIWDKVT-PVQLVIYRKANAWSVYKSDLKTPALGLNQFALTTAIAVQPNDFAGVYYPOT-GSVSNRKRTEQEAWDLGNLGGSVLFTSTGAAETAFAFGSSNRVYSLKWHGID-DAPFVPIEKPANGRIVHNDEWALSADAGFQVSDTATFALNVATYFVGTIGK-FHVL SNNFGLYGATLERTMTGAGHTWTKGMNIAINAVALAQYDAIFVGGDPVD-NQVLDIYKNGGKYL CAGTGHGGSQAQAEAGQVWNTFLATGLKFDGVMYKISG-SLAVTSPSHPLFSNVKTLYQNNNSISLIPNATANIVMTGSNQGLIATADYIKPN-APARSEPTSPATSEPPKPKTIEEFSTIAEDDDVFMNSLVIYKGLVKKTSQDEY-VELSNRGNHPVNISGKWKITSVGSARQQTFFPANTVLA SGKTFRVYTVNEVHSE-TGGFSFASKTAIWNDA GDELNLYDAMGKISTLAYGIARTKKS
Pseudoalteromonas marina	Bacteria	NCBI WP_010556046	MTIKPLLFYSLITSTTFLYCGGTEQTTDSSTEQTDSSTITSTGLVINEIVASPSDT-TYDWIELYATEDIADLVSVLVDNDADREPAALPAVALSQEIEFVIAIDETDTP-PDNGYVYVTKLGSDDAVLYEDGTAISVLDWEEGAAEAGFSYGLYTDGTG-TAQLTLPTEGAANQTANTSTLVTLIAEDAPLRINEVAVKDDSSGGEDWIELVYTS-SDVYLADYTLSDNNQFALPDITLSPGEFYRIYASTDDLGDLPVAFKLGASDT-FHVL SNNFGLYGATLERTMTGAGHTWTKGMNIAINAVALAQYDAIFVGGDPVD-NQVLDIYKNGGKYL CAGTGHGGSQAQAEAGQVWNTFLATGLKFDGVMYKISG-SLAVTSPSHPLFSNVKTLYQNNNSISLIPNATANIVMTGSNQGLIATADYIKPN-APARSEPTSPATSEPPKPKTIEEFSTIAEDDDVFMNSLVIYKGLVKKTSQDEY-VELSNRGNHPVNISGKWKITSVGSARQQTFFPANTVLA SGKTFRVYTVNEVHSE-TGGFSFASKTAIWNDA GDELNLYDAMGKISTLAYGIARTKKS
Dorea formicigenerans	Bacteria	NCBI WP_005339862	MKRAVRMLCTFLSATMLVPLQVLGADAVSKPQVEQPSIVINEVESDAPNKGND-WVEITNIGTEAVDISGWYLTDDKGEERKTEGKTTPLAKGTILEPAGFLVLEETKNFD-FGFGKVDAILYDAASTEIDSAYTVAAGTWSRQTDGDFSDVEATPGTANAAP-TAPEPSKPEQTEPSAKPSLVLEINSSPDDWMLMNTGAETLDSGFELRDNSD-DHRWRFPDGSIDAGKLMVVDKASNLVYDDQTKSFADGTFEAIKIGSGDSIR-LYDKGKLLDEYSWTEHASDYGDPKASYGRYPDGTGEFRTOETQGTGANAY-APT VVINEVESNGDITDWEIMNIGTQAVDISGWYLLDNDPVGHKGDVTPADG-TTLEPGALYVFDQNKDFTFLGKDDKAVIYDAGGSIVAEYAWEAHANGVYARIP-DGTGEFQDFATATKNNKKNVMPVINEVQSNDPSPGGPDDWELANPTADELSI-SGLVLKDNKDKDPYTIPEGTTIPANGYLVYQNDGATGFGFLGKGDVSRVLF-DEGEIAATTWPDGSHNTPTWGLYPDVNGNSYRNTLEATPGAANKFAGIFDVA-WPGSDQVRIFDTPFLEDSSGLDFANGKLYAVDNGTATFWMDVAKDGTFTFA-DGFEQKGRVCFQKSDNEKAKGPAEIGITVDGSDMVYLASERDSSKGVNDY-D-TILMVNPSSESGTRLVARKQWDLTATLPQVSANMGVEAVEWVANSNDVNGKLDQ-NTGSFAFDAKYPNAVAGGVFFVALEDNGHYVYILNEDETVDVIADIDSKLGG-AMALDYDTEYKLLWVAADDGYGNRMAQITLNGTENPEIVHVLPAAGVDTTANNEG-FAIADASYTVNGQRPAFYRCQDGVTSALGTIGSINCRYCGTEQPPDSEPIDPGK-DNPKPSEPTRPSTPEKPTTPSKRPAVPSAGTSVSTEKTTNPKTGDSSMTMGLV-ITLAVLSVAVCFATLLLRNQKKH
Alistipes sp.	Bacteria	NCBI CCZ98558	MLVINEVYTYADGSEKDDDFIELYNAGGSDIDLGLKLWESGGSAEAWSFPEGRS-VAAGGFFVVVCDKDNAWYADPVNYPGWGLSKGPDVYVTLADAE MNEIDC-VACPSMKRGESYGRVTDGAPEWQIFAAFTKGTPEEGPARQPVNT-MGLWVNEVFTNNQDTAQLPWNESVDFFIEFNSSDITAIIDGGYVIKDDKGD-DESYTVPAGTVIPAGGFLTYDVCKKNAEGSPFGLGKSGDWWVYDPAGVLA-ELEIPAFADDEVMSYGRMPDGGDVLKMEPTKNSNGGVEHPEVAVVINEI-TNGSQDEDWVEFYNGGSADADLEGVYLYDDGGVEKAFTPAGTVVPAGGY-LVMVAKEEGSFDFLGGKGDALTLLDASGAVTDEVEIPALDDDETYGRRT-DGAAEWTVFGTSTRGASNAGGTVRE

Micro-bacterium sp.	Bacteria	NCBI WP_029266 970	MPRLHRAAAVAVCALSAALLATPIAAIGADPGIRPSVATSTAPSLVLNEIVYDDAAT- GLADQVEIYNAGTEVVDLAGWKIADKRDTFGAAPDGTSLAPGEFLVLDVDV- FAFGLGKGEDEVVDFPDGTEVDSYAYANTAPLWVARCPDGTGAWAPATQVTP- GAANDCTVAPVAGSVINEVDSQPADWVFHNPSTAALDISGYEIRDNSDDH- RWQFLPGTQIGAGQFLVVEEGTVGVSGVGEAAFRPREGIIGSADRIRLYDTS- AMIDDLTPWQGHAAIDGFAAATLARCPDGVGVSFLAHPHTPGATNSCVMPPDV- VINEIESNGDTTDVWEVNTGSTAVDLGSGWTVMDSDPVGHAGETTLPGT- ILQPGGYVFDQPAFVFLGNGDTSIRDANGNTVEHVYAAHATGVLAR- CADGTGDFVDIAVSTKGLRNACGNPVRINEVESDGGSPDDWVELVNPTAL- DVSIVKDDDDTHAYAIPAGTLLGAGEYLVIERAQFGFGLGDGDAVRVFDH- DLLUDETTWAGHAATTWGRCPDPTTGAFVTAAPTKGANVCPGEVAVSP- WPGSAEVRVVDGIPTFLLEDSSGLDQETADGAFWVAVDNGEGRIWKLDAH- ADGVSKEAEGWDAGKRVRFQKDAANPGAAGPDTEGITVDDGDFVVASER- DNSAKGVTQNVVLKVDPEASDGLVAQQEWDLTALLPAVGANLGMEAVQW- VPDAALAGKLFDDRTGAAYDPRDYAGHGDGLFFVAVEDNGHVYAFALGADG- SATLVSEIAPGLTGMALDYDTRNTLWAVCDDGCQGRSAEITLNGTGQPSL- VHYARPAGMPDINNEGFATAPASLSVDGQRPVWWFADGFASEALRTGTLPG- VDDGTPGGETPLPGTGLVDDNRHGLTVDPSVATRGQKVITVGGAGAGT- DVSVMWYSDPRTIASGTLTDAGTISVTPADAPLGAHRIAVFDASGALLGWA- DLRVAAGNGAGAGAGAGAGAGAGAGAGGLATTGAELPVAALALALLLMTGTA- VAVRRKRTA
Verrucomicrobia bacterium	Bacteria	NCBI WP_038132 244	MKLLAQTFALLLITAFRAQAEIVAHWPLDANAKDLLGNHDAESGVIFGVEGAAN- HTGTAAEFNGSSITVFPDSAMNPESFTLMMWVADSNFASVPTSRD- DTPTSVMHGYLYYNDSSGNWFWTGTGGPSSGAWNQMSSGGKVEIGSWTHIAVSYD- DGSQTKKLYVNGSLAGNSAPNLYDPNGPQSENLI- IGSADSGGAFFFDLIDVALWDHALTSVEIKDLMT- NGVPGGPPSITVFEASPPFIDLQDDVTLTSWEIQNAT- SVSPNVGSAADRGSIWVTPSKTTTTYITAIAGESSPAAT- SQITVGDVDESLDPIITEFLADNKTGLIAPDGNRSDW- VELHNPFPFAIDIGWHLSDSTLQEKFTFSPGQIPA- GDYKIFFADSSLEALNFKLAKTGDYALSDPNGNLS- SEFSPAYPAQFDDVSYGLSNSTLYLEPTPGVANG- SSRSEIGPKVEGLTRNPLPPTPGEFLVIANITPRIG- IASSKFLYRVFGFDRQLTMTKGADDOFTTIPASAY- EAGEMVRWYLIATTSGESTREPPFDLTESAEYFG- TVISDPSISVDQPMHWFVKNIGAADTRGGTRGSLY- FEGQFYDNIFCRIRGQSTANWPKHYKDFYRGDH- FRWKPPEAKRVEENVNSHYRDSYVRENTIFAFLNE- AGAMAPETRYLWKRNKSDMGLFTFVEQVDEEFLE- RGINPTGSMYKAINVPATLSPTVNSSLYRKLVRKNEPY- TELRELTSGINISNPNRFEFVADVAVNPYINVMMA- MCVFPNHDQLTKNYVYVHDLDRGEWFRIAWDGD- QGLPTGRNNGNENWSSPLYGDALHTQELVGGNPN- PIWQNHLLHAAILDNPTREMYMRRVRTLMDEYLAIPE- TGPSTTLDRGARLRYLAPIDASLESSWHLPEFDDSD- WLSGIAGLGYENNPGLYGLIETRVKPSSELNPGTS- VYQRHFNVDSPSSDVLVRMYDDGFVYLVNGKEI- ARDNINGAVRYNSTASSHPDSQAINVFEFPLPGVSLIP- GENVLAIQIINQSSGSSDLLCEPELVDRPGANGGYF- EGGLEGRNITQRDVVVDQALWSGAGITNFFNNGYN- GVLNLSLPNRRNALFKTYGPSGSLIPQEQSSGLVIN- FGNIESNPDGNGQDEEFIELKPNNEAVDLSGWNISG- GVELSLPPGTVIPSNSLYLSPDVHDFRLRPASPTGR- EGHYVIGNYDGLHNSFTEILSLSDSGGLISEVITEDQ- PSDAQRFVISEIMYHPGEGAGEEFIEIMNISKSVTLNL- GGISFTDGINYIPAGTLLAPGNRMVISASDFKNGTAL- SNGGELLKLEDASRSTISEFRYDDQAPWIPSPDGGG- TSLILINPTLKPDPASLNWRPSTLTGNGPMTDAISF- RGOANGDLNNGISDLIDYALGNMGMTQMTGDSFTY- LKRLGADDATFQIEASTNLINWTDASALLIEQSRDTQGDTALIKSKLSPAVSGTRW- FFRIRVTLNLLQN
Clostridium straminisolvens	Bacteria	NCBI GAE86954	MKKVVSSTVVIAMLLAIVPAYFFGAFAEETQTLFINEVMSSNESTIRGDVDTP- KHGSKGGAYSDWIEIYNSGSTVDLTYGILEDSSAEWTFPGQIVPAGGYLLVWASD- KNMVAQDGGQLHNFKLSASGETVTLKKPDGTVDSVAVIGLGDGQSY- GRIRDAGEPEFVFSKATPGAANVYDAPGTTTTLINEVMASNASTIRDGDVDDP- KDGSKGGAYSDWIEIYNSGSTVDLTYGILEDSSAEWTFPGQIVPAGGYLLV- WASDKNKVAADGQLHTNFKISSGSEALTKDPDGNVIDMLVTINLLDDESIGR- KSDGSSSELALLRPTPGAANYDRLIPVSEPAFESHKGFIPMRI
Limnochorda pilosa	Bacteria	NCBI BAS28019	MLRSSALLFLAGLALVGNLGERQVWDPALAAAGAPSGNG- PDDLVRAGVVEIDGDTFVVAFAGTPSARARLVGVDAPEVHP- TRGVPEYGLAAAFAADLRLPGLIELRLEFVDAVQNGRLLVYAYLPDGRMVNEVL- LEAGYAQIYTFPPNVRYVERFRAAQAGREQQGRGLWAIPAEGDGEPGRVIV- IASVDLEAEQVEIWNGTARALEMTGWRLISVQGGQYEFPEFGVLGPEAVV- VTSGSFAHQHPATLWTRRSVWVANRWDPAELVDPISGRVAAFF
Desulfo- sporosinus orientis	Bacteria	NCBI WP_014184 986	MIMRSRLFYVLVICITLLVTGCGNTSTTTATSTSSQTSVDQGEELNQNQPTM- NAHGSLQVHFIDVGQADSLIKASDGTAVLIDGGNPDGPNVNYLKSQVKKE- LAAVIATHPHEDHIGGLDVTINSFVKAVYMPDASTTTKTFEDFISAVAAS- GAKRIQAKVGVKLDVPLGTGQFIAPNGSGYEDLNYSAVLRLTYGKVSFLTGDA- EEVSETEMLGSGQIVKADVLRVGHGGSSSTSSAFKAVSPKYAVISGADNDYD- HPTASTLSKLSGAVKVRFDLGTIVATCDGESVTFNKKGSAPASVITSIDLAGE- IVTLVNDGSSAVDLGWKLVSEKGNQTYSFPSGITIPAGGTLLKVLSSGSAESSTGA- LLWTKSDVWNNWGDGPGSLYNAQGQVSVK
Petrogoba mobilis	Bacteria	NCBI KUK80215	MKKVLLIFCILLPFSLSLNFPSFIWNYVSSYQTDEQIYWVKNYQKLIK- LIEDYLPFLEKSGISLPITSNVEVHFIDVGGQDSVYIKLNNVDLIDGGN- NWKYDQVVYTKSQDVEDIYLIATHPADHIGGLDVLKAFVENYVAPDVSHT- TKTNYDFVLEVRNEGTYIKTKAGMTLINSALAAVIGCPSFSVYFLGPVKNYGTL- NSWSAVMKLDFGNSSYLFYGDADRISEKDMIDNMGFLKADVLKVGHHGYSST- KEFLKEYSPKYAVISGKNNYGHQAQEILEGEMKYVDFRDLGGHIIAISDGNK- INFNVEPINTVNLMEDAQESAPQKTSILITLNDVSEKVTICNNTDEDVDLGTWVW- SEVGNQKFNFPDGYVLRVGEVNLISGRNAIDEPVNLKWTGVYIWNNDSDI AVL- YDAEGTLISRLEF
Methanoculleus marisnigri	Archaea	NCBI KUK63359	MSSIPGYLITIALCTCIIACFITAGCSAPIAQPEIASGLSGDLVHFID- VGQDSILIEFRDKTMLIDAGERMGGERVIAYLDERNVEELDVVVAHSHDHG- GLGSVVSAYVGRFVDAAPQHSTATEDLVLVVEEQGIPYTAERGGTIALDPDELE- VLNPTPQPLGDINESSVLMVMTYGEISYLFMGDAGNAEESSMMEA- GLDLADLVKVGHHASRYASSAEFLAAVSPAISVIVPVGEGNDYGH- EEAVERIEATGSRIRYTDLDGTVIVATDGRALTVAAGGAPSATVATARA- TTSTPAATSSGVYITDLDIQDEWITVNAETAAVNLGTWITIDEGTRN- TYTFPLFSLSPGADVTVHSGPNDTVSDLYWGRETVPVWNNWGDGLATLA- DANGTVVSTLER

Methanoculleus marisnigri	Archaea	Uniprot tr A3CW90	MSSIPARYLICILCACTVACLVTAGCSNPSGQQPEIGSDLSGLDVLVHFID-VGGQDSILIEFRDKTMLIDAGERGMGERVIAYLDERNVETLDVVATHAHDHIG-GLSDVISAYPVGRFVDAAPHPATYENLLVLEDLGPYTAERGGQVALD-PDLEILINPGAELPLGDINEDSVMLMVTYGEISYLTGADGTPAEESMA-EAGLDLADLVKGVGHASRYASSAEFLSSVSPAISVIEVGEENNDYGHF-HEEAVERLEATGSRIYRTDLDGTIVATDGTALTVAAGGAPAAVTAGAT-TAAATATPTATAPAPSSGVYISDLGLQEEWIAVANAETAVALNTGWTIT-DEGTRNTYTFPVYLDPGADVTVHSGAGNDTATDLYWGRGSAVWNNGDGLA-TLADANGTVVSTMER
Haloferax volcanii	Archaea	NCBI WP_004043458	MVRRATLVVVVAVLAGCLGGGAGAATTVESTATETASPTAQSGSPT-TAAAAGDVTVEVVDGDTIKVMPDGARETVRLGVDTEVHAENTPDE-FEGVPETEAGRTCLRAAGEDASAYAKSRLADRTVELRYDEKAGERGYYGRL-LAYVVVDGAEFNYDLITEGHARLYESSFEERERYERAERDARERGVGLWSCATEG-AAAGGADATDDGLSLIVADAPGNDNDNLNEEYVTLRNDGDAIDLSGWTVSDA-AGATYTFAGGTLDPGASLTLHTGSGTDDEDDVYWGRGSAVWNNGGDTVTVRD-ATGDEVLAYTYE
Methanomethylovorans hollandica	Archaea	NCBI WP_015324112	MQIKLNRILICLSILLFIVMTSGIGSADEVGVFDNGQWAVGNDPATATIFGFWIN-TRPMVGDWNGDGKEEVGVYNSPGKNFVQDANGYKIVGLGWTGVIPVWNGWNG-DRADEVGVYNTTEGTWALWNTDINSADIVGFG-WVGTPEVVDGDDVTEVGIYNNRRGNLFLKQSGFDIHLGWTGVTPVVD-WNGDGADEVGVYNNAGTVALWNTDINSAEIVGFGWAGTEPIVGDWNGDGIFEVG-IYNTAGNNFLITENGFDVIGLWGSRVTPVIGSWIDASTIYKPKQDKSPPEQPPTQQT-PTQIPSPVQDSSLTVHFIDVGGQDSILLEYGGKMLIDAGENNMGYTVSAFLRERK-ISSLDYYVATHPHADHIGGMLTLNNYPVGGFIDSGYPHTSKTYENMLMAIDAQNL-FHVAERGETINLAPGTIRVLNPKKEQSEELNENSVVLKITHGDVYFLLMGDAGLEA-EENIMAAGYDVVDILKVGHHGSGTSASGYEFISAVSPDISVIEVAGNDYGHPHANI-LQRLQGASTVYRTDYDGTITITDGSRYTVTEKTGSSGTSTSTSEDVAYSDDL-RNEFVIITNPGSSTVDLTGWKIKDEGNKHTYTFYFQLESAGATVITYTSKGLDSTKL-YWGSNGPFWNNDGDTASLYDSNGNLVDSLVG
NMCPs:			
Klebsormidium flaccidum	Klebsormidiophyceae	http://genome.microbedb.jp/klebsormidium kfl00193_0080	MFDSPTGAILAEELHAREAARSIRVWTVAVNESVGVSDRREENTERALTAVNPS-GEVEVWRRFRGAGALDEEVLVREKEELQRQLDDVERELADYRYNL-GVLIIEHKKCKQIDELKALQRTREELQKEGRSLQLALDDVTRREDGLRSLKA-ERTVIADLKQSLSEEMHAKLQAEKESRERLARAQELAEAGVDRERSEE-LFEKAARDLKGAAQLEELGVRESELDGKLRGLAEREDDAASREKALNE-GEAALNAGKRELTEKGKILAEKRKRVGEGELEERKLGREREEEELRVEA-LRRSVEGEREQMQKEMQAEKAAEHDMSTRKAEISVKEDSLAAQEKAL-QMSQDELERKQEDANKRDQDLQREKLLGDEKREMLRTEADLRQEQ-SMNEERARLAEQSEWQRIEAKIDARVKEIDAREAAVQAFESTAESSAKE-LEEGAKALRERQAGLEEKAAVQKKEEDLQKLEKQKGRAQMLSEERA-ALERRKGEDVLEREQAEKAAQLEMEAAALLDKEAQLLEEQTRQLASKDK-DLAAREHELLRESGTDENLRQKKEELEAQRQAEADVAKLQQLQGD-RAAKGEALLQASRAQLDKERAAMQLELQETRASLEIEFQTKMLTADAATQR-AAKSEVEKRVKDLEREAEIRAEERATVRLQERLAGVDLEVERRVSAQLKE-RVSELEVRNREEVERRVSEGSKKGVREVKQAVKDAVKKAKEEAKAKYD-QERGAAIQAVRDAERKAEKRAAIDKAVKEERSRTAGQARAVSAEAG-SEARNEEGERQILVDGPPLEQIDGLLSQVGMETPPVAVNPVGSSEAA-GVEGASADEVAVHAEPTPVVSAVNGAPSPKRPKLQSKRKQPETPPPEEA-PEERPLKQRRIIDAGPDKPVLEEKHGGGVEALGAAVTAVGRAADTITPG-FGFVLLDDTDLGFVPGHTDRKAPREVSVKPRWLSLWPGFLSPRKAKAE-GGQDQAGTSPSGEEGAEKNLPGGQVEGQSSPLENARNDGGEVGPV-LMDVQQGPLPTENETPPEMDLRGDETRAAEERATRVAKTPNTRVGRAVL-PTPKPARTPKPPLSAVRTSSRLRERTLSSDLKGTLEPTSATQLAGSSGKR-GRSSPPGADVGPDAQSQEPKQTERVSPLDGTRNQAPGKSAEIAIGR-QRDPSPAKTPTGTGPARRLQALTAEEQAKIERIMARLEAEERAAIADEAAV-GPLNARTSPDLAPERFLNAGSADGAKTSVPEREAANADIVDGATVKGL-EKVRGAPETGNSGLKKAIVAVKRRSGERAGASGKTDETLIEAAAANKQ-ELPNDDPQEAAPNPMPAPGGASAAVDAELPTVSPSPGPTSTLPPKPT-KSAFGRAFDGFRAAVGFGAGSSRGPASEGDEEGSLEGGVYRSQVGL-RSIVGDLVAAADRDEMEMGGPERLEFGTPEGVPFDGPEGEERGR-EAAQGGEGLLNTTETVVEKSEALREVVEKGGQKAGSVIVKAKATAAFDL-NETAMELADAAGVVSASDVADDVTRSAADLADDVTALGGAVEATVGGVAD-ELTGNIGELADELGAVHTRDAQGKSLVETGSDLVGTALGSLQAGAGAG-SSALAADPSRGAEGGARSTRGGAKVQSKKEGGNGEERRGSKRKAPES-GKGGKVTDSSEEDVGPKRKVKRRVKQADSSEDAQLVDRPVVEFAT-KFRGEVEGEDTPSEGGGKRHLRAVDPRDGEDMSAGLRRYNLRKSTLV-KMHIWEPRGADFGTVSTGSEGHTRGGASSTPRNLQRSVFGGDNPFLED-IAPGNPHSLEIIRTPRSRGRHGARDDEKSGGAAGGTADDEEADQANAGRWLGF-LGW
Klebsormidium subtile	Klebsormidiophyceae	NCBI JG442173	GEAALKAGKREVIENGIELAGRKRKVEGELEERKLEWKREELERIAL-RRSVEGEREQMRKEMQEEREAAERDLSARKADISTKEESLAAQEKALQASQEEL-ERKHKNAIKRQDLDQREKLLGDEKREMLRKADLREQEKSMEEKARLTEQE-SEWRRIEAKIDARVKELDGRAAVQAKESAESSTKELEEGGAALRERQAVL
Spirogyra pratensis	Zygnemophyceae	NCBI GB-SM01021289	MRNAFSPRLRLRTQRDADENMSNYNRIGSESIPSSSIMPRHRAQEVWITDAQG-PVIDEVS VIRSERDALSIRLQLELEYGDKLQTEAFAFDERDRCLSQFS-TIQSQLKEMVEKFKLEKNAHQSTIASNTEKERKLRDALHTQRTADELELALTEKN-LFEDNFRQSLTSENETIRKCKIAEEKILECEISLKEFSLKGVEYKVKVEY-ESKLQETQITESNLRLLEYLRKDSQNEFQIQVSLKEQALKEMEERSKLI-SEYEQMNEEMEAAMKTIKEEEMESKEKELIHFENETKIERNLELENLEK-LKLEMQERLNEQEKLVIAKENSFSETKMHLTSFEEALNFRQESIELKLNEL-ESKEEMNEEKIKLLSLEENLKHKQVAVDEHKTALIKEINELQFQREEMEK-EKFLFDREKELNHLNHLRQREGLNLERDALDSKESNLIDLERKLGQKE-SMVYIKKESVEEDNLTNVKMSASKMMRELMEEKRQLQEDLSRLGEDT-TVLRQERELVKLMRESMDAAYSVMNEEKERIEQSAEENERVRRREAEKAL-DEARREKEDVFRKDMVEIEREELREEKRFRFHAYEELDAMKAILDEEKDEAKR-MHEVRRRELEMQVMEINQMK

<p>Nitella mirabilis</p>	<p>Charo-phyceae</p>	<p>NCBI JV767595</p>	<p>MKSAKKASSMNGSASPKNPNPDNNMISAGSFEIRKMTTTRTLRESASCEKKKK-TYKEVEEKNAGHNFGGAFSREALALDLSGLPSDLKPIELGAWRRFQDSGALSEQ-GLEKKERNFMVDHIKQLEDELYKSQYHMLLLMERKSLLDVAKLDHTAKEYEIEIKREREESMTRIEYEEIETTKENLLAEQQCVKEVEKAIRDVQKEKDETTTLAELR-FSEARALMSEVEKKEHGVESIRKNVEATRAEFTRKKVIEESRARELTKESLVRRE-REQVRIEESSELEMKRQELIERENVAREWEAKLASREQSVRKMEDRMKGRDVTL-RKERDVEEGMARLDGKRKGLSEMEDDLRLREKDLAMKEESLVRREVNVOQEQ-AQQEIEERLSMKDEEMRSAKERLLESEHDVKLERERLKELEATLRRQDIEQKQ-QMSKENALMERELSSASAVHNKLEREDLERKDRDLRLREEIEEQKALLNDSRR-DAQQWEAKLEMRERSIAKEESLRERQQDVELRGKGLDKERELKMDRRRIEAS-KKIEEEREVRLDMELAKKDREALANEREEVKEIEEDCRRLKVEADNVRADALAKR-AEVKKEREDLARDREDFMRKREDMRKERDAVSKDREELKREKELLEQKRQDQV-ESENVDAEREEVEREREELDKEREVVKQREENSKELEGLMMLKDDLNEKDRSLK-IESRRVEIELRTVKEDRDVARRSEELRKDKKEVQQLKIEVKNEKTLEDIWKEEIETE-RERMEREKHEVRRDRELLIRQEEVHVVKLIEIEESQRTFKMDKKKFELEDHRAKKEE-RDDIKEREELARERDLVDIERESLKKERDEVERDRLELELERERLEAKSELFEKEE-EELETREEMVKKDMESLRSQESRYSSSEVLQKETELESLEKRRWKEGLEENRVTRT-LIREKEEFREKEDAAKQRAENQREKEDLERDREELDRDRIELTVKRDDELKRAMEL-HLRQEEELAKTMEMLAKERMMLKSERKLESASAKLKEEREEVSGEANEVKKERE-EVLRERESVRKEKDELLQDREDLERGKEDVRRERDEMRRERDVRRRERIQLLRE-EETLSMDRDELMRERANTAKLELELTQVEELKRQREELVLERLCKDQKAAALSQ-NMELTKQGEERYRRKMDDLGRVDTMRRRRREAMIVEADAAVNHSSGLVATLLANVA-HVEDGATAVALPQALTAGEDALHGARSREERSLQERVSEGGREETDKISWRSE-TRVSSKLRGRGSYFLRQHAASENRRSSVRAQAIEDERRRKRKRGHSELSTRDL-QKGESLVAVRSEEADEQANTDRRVDSVSSAGRSCTPPSGADASPSSRRRTSMQDV-INKGKEMWMTSTIGDEVDDITIGSKGVGKLNVSQKRRRKRVRAMVDADEDTTDED-DVGEDSEVARKTEGKVKYNLKRTTVYHQATRVIESRDEQPSAGGAEAMALQAVT-SAGLIPSPAMGWKLTRSLGLKQQVPRRRPSEGEHTSSRDTSQETVALGARKSG-GEEDASGRSREEEEEEEEEENNEDEGGDEEENEAEKDEGADAEEDVVERGR-ENSSWLEWLLPVPWNV</p>
<p>Nitella hyalina</p>	<p>Charo-phyceae</p>	<p>NCBI com- pound of: HO490484, HO531334, HO566387</p>	<p>LFLMEKRSLLAETSANDTVEGYEIIKRRREWSIMDFEELTQLKDTYLKEQRD-VEEVKIDXXXXXXXXXXXXXXXXXETKENLLRREREQVRLGADMETKQRELTERE-NIAREWEAKLATREQQVKKTEERMKAREDTVLRKERDTEEGLAR-LEGKRRKGLSEMEDDLRLREKDLSTKEEGLXXXLNMKDEELRSAKERLLESEHDVK-MERERLKELEAMLARRQEIIDQKQQLSSRENAVTERELSSASAAHNMKLDREDI-E-RKERDLRIEIEEQKALLNDSRRKHSSGKRSVELRERNXXXXXXXXXXXXXXXXXX-XXXXXXXXXXXXXXXXXXXXXXXXFLRQHAASENRRSTVTRTVTEEAARKRRKRGHS-ELSTRDLQRGESVAVAGSALEEQGNADRRVDGVSSAGRSCTPPSGTDAASPSSR-RTSMQDVIINKGNEIWGASLVQGDMEAAQAVKAGKXXSYIPKRRRKGQAMVDA-XDADEDTTDDDTGKDDTEVARKIDRSVKYNLRKSTVYHQAAFSEQNRDAQAYE-GCGVLRQVAGTSDRLVPPSPVMGRKLDLSVLGLKQQASWRQPSEREL</p>
<p>Coleo-chaete or-bicularis</p>	<p>Coleo-chaeto-phyceae</p>	<p>NCBI GB-SL0103124 2</p>	<p>MFOHQKRTGPRPSTGIFRNGPISQSPPLPGIPLQGEAPSAREVATPVGR-RRELEGEGHEDEAGVWQRFYEVGALDEESIEQKLEDLMN-QLQSKENLKYQFEMGVLLNTNLRLDELVAELAAKDRLELRAEQNEHREM-VAQMRTRTEOQLHKSVEAERRCTADLERAVREMQRNESMEAKKAEVQMEESM-ALWKEAERKRQEAQAQVRSQAALHREAQKMGIEIRREEDVAAREVNRHEI-AKRLAEFEAEKAEVDRRQREVASWDASLKECDGKLEAQQRLRQLEKLSLEERS-ASILSNEKALEHRDSSLVAREENAKYKEEVEKAWIEIRRAWIALEELKAVGQQL-NTEKWAEEVGARERAVVEAGDRLAVAEADVTRQSGELVQAQAEAREHERLT-ELSTRLLSDEAGLASQILKHDSDWEQRHSHLRTLESEIEKLEAEARAKESAVA-REESLTVALSDEQRRRAEITAARSACEAAQAQLQEKEENLRKREEAVERSEN-MDKKLEKVKEREDLGERERARLQAKKAVELDRQSLLEAGRASMKKADIEDL-LEELGKRAEFALVDGLQTKHALLEDKAALERENAELAGLQGELEKKEEG-EELAKIRQNLAEAREELKAERTRFEDKVVVLDQREKELAEEMEERLAGEVKSER-LQLQEERVGMQAVRDAEARATEAERAAAAARVEQVAAAKMEMEKERVAERE-LMDRQRAAELLKEESGEKYRKDLEVEFATKLEEARLQMLQETVAEADKRMKD-KVAEAEEEKRVAANAAGLAVVERKELEKEREALAEORRRVMAEAHAAAVIVAT-KVVEVQKNLESARMALLKEVEDLKROAKEDLDKEREVLACARASAEQMKVREN-AVETEHAQVKAALKELRFERSKLVALKEELQAHOQAQQLAGAVATGEIDGMHLPN-VDAEAVYLHGGEPDAALLTPLAASQPMDAVGEKERAMDVAGKRAEGSVERFN-QMRSPPLFPSPGGTINWLQOCTRFFTPKKNGDEQSPAPEQAQADAVDDNHNE-EDQRMVMSVCPVQGRSGSSGSLRRKTTIEVTVAEAALFLGSGGTASTHDLGP-QEQGHAEATASGCKERQDVRPIPRRTISETIEEGRFLGLDAAAADDEDDDD-MRVSMEMVMGSPVGALETVEEVDEEGEGAEQLGIELSAADLVAVRESGMPM-EGERRGEDSDAITSAAKSALRVTKRTRSMRDVVDEGRKFFVGISTDATDADDTTQ-IPLSGRTRRSKPTAAASMDNALLEASGPVSSAAAEPIGNRLRRSNRPHDI-KGGIDTSETPSMQFVGVGVSVNRRSRRNRPTHTKAGPDTSEAPSIQPG-EESAQGSRRRLQRRHSMADYSKGNTDSESEGMRRPKSRRFSTIGPSETPAMS-ADTATAHGSRYNLKRSKRFVWMDGCVIAGRHRSSGHTDVESESEATSAHAETS-PPKALD</p>
<p>Selaginella moellen-dorffii</p>	<p>Embryo-phyta</p>	<p>NCBI XP_002993 584</p>	<p>MFTPHRRGATPNRGAGFSVSTERREVRFASSPPDGRQQHQSADGTLAGN-GAGDGAGKSSSEIWWQTFREAGALDQESLELKDRNALLAHISKLETELY-DYQYQMGLLLLLENKLRGSESLKSVIDEITRDGLKREQSAHMIALQEAERRED-SLKRAVITTEKCKVADLEKALKEMHEEVAEAKAAAATQFQQGKATAMSAEEKLLE-AESKLSAEALLAKANRKHADAERKQVEVESREDALRRQRHSFLAECGAHKLLE-LEHEKQNLKGWERTLEESQARFVENEKLLNKREEYMQQRDDALTKLERLDDE-ARKVLEKDRSALRQEQAEYSALLSALSLEEAVERENAATKKEQELLLQEKLA-SRDRAFEQHEQMVRELEQANAKEKERLVDLEASLSTRENLLAVSKQSLVN-IVFYVPHALICFCMVGGLYLEDPG</p>

<p>Physcomitrella patens 1</p>	<p>Embryophyta</p>	<p>gij Phy-pa1_1 169197 estExt_fgenesh1_pg.C_2000036</p>	<p>MILSSGFSSGMLQELASAGLTGYRLVEYLSRGLFLVFREADGTSNLLLKFFLVLQFVCRIGVLSFESIDVPEFLPCSTMALHSPSEYVPLSAYRSATLSAGTSVHFNCPLDIRNWTAAHKVIGFLYLKERSPRYLRLWRKVAEELVKGRRGSPQQRVPTTRSLAREKGGKPPITNSTIGALTITTTTTLAAAMEGEEAMALMDPLMGSPDMIGVADAVPTEVWKRQFNEGALDMPSLERKDRALHARIAALEAEALYDYQYNNMGLLLLQRKTWTSQADDLKAADAVADQETLQREKAAHLELSEVMRREEAAKKALETQKQVADLEKALKEFQDTESEVROAADKQLAQARELVASIEERSVQADLKAQVQVLRADANRKLQESLRLQEVEAREVALRERHSLMADVARKEQVASEEASLKEWKRLEEGRARLQEGERLLNERENSLSKQREDEALKQTSRELAETRSYIENERALIKQTDADLNARVISLSERERTLSERELKILTKEQDLLAEERIAEQTRFENRELVKQKTEKYVEQERARLDQYVESALKFOETTLEEQKMELESEMEALLKIHTSDVDSSKKAELLAEEELRSVRKTLAAEKEEVETLKAEEAREARSRHLETAITAREEELKLRVQEIVDREVDLNRRLLEEVSNLEQGIRVEEKKYENEHRIAELKEEIRKSKKEEMENKLELQKQLEEEEREHLRRECELEQRQIEEEREKVRKRDWEEEREWEQQRLVQKQDVEYKAEQLEFEKERLREELKAEREKQSAELERMRVNLHNEALEALREKLKAEVEFERDALRKEIETDQERVAELREESRRAIAEERQVDEERSRIRRELEVEEERQQAEESEERAAHAILERRKIEDEQEKLKALEQERGELVRIQVQLKQEID-EIRARKQFVDEEAELKQKDRFEREWELLDEKREATRKEEREFEEESKRMAEWMQDEEERLKETRRVQEQSRRMTEELQKERESEWRSRLERERNQLYTQLDVERQALNRNLELQREDLRRELERDFAFEQFEEREAQRAEVEWQEKEDLRKNRGSVIGELEQLRAERAKLEKERQELLKQRVDAEKVEWIKKDIEQLQGEKLRQRESLHLERQNTMREAERLQKLRQEMKQGSEGSMSMRVPEQPMRMEEVEVSPHPQGLVTRDTRQAVGGRPAPEGTHKPPSSHISSRRMIARTPSRLAWLQRCASRASLLFSSTKLLTGQEPVEEEEEEEQVKQDPNAPSSSNQSQLGQVEGNIDDGPRFRRTSIRQVVEEANAALIGIVGEEETSESNNRNSNADAFITTPAESAEATRKQKRRAGNVDDDDANPLEADAHGGTQRKRIKDIMVEIETNEDSLHTPHSRVSTPATKRYNFRPSTVNM-TAASENSVRHHDRTSKKAASATSQPTAAAVDVLPEVISPQVQETEAAMHEAPAVGETLEAQEEGDVAEASPHSGTTTTVRVDVHTETQVVTETTIVTETVKELAVFDLNVVELDMAIEAEVPTAEAVHLSRSELVESGEEETAGDSSPMQAGEVDVGVTLDDSCDEVEDVDAMEGNDEVEDDEGAAENGDNQAVPENGELLSEEESEAESEAEVEEESEAESEAVEEKSEAVEVEEDV-GEVEDDENEDTREDDPEDEPDDEGEKPSIRAKIWDFLTT</p>
<p>Physcomitrella patens 2</p>	<p>Embryophyta</p>	<p>gij Phy-pa1_1 233301 estExt_fgenesh2_pg.C_760029</p>	<p>MSGSLSPMYTPQGMRGSPHQRETPIRSAREKGTSPATATTGVTMTTTLT-AGGEEGLVLMNPLTGAPDTNAGEADGVPTDVKWRFOSEGALDISLERKDRALHARIAALEAEVKIATRTPVLELMDKCYESSFTLGSSTLGGQRLCKDDAAILYDQYNNMGLLLLQRKTWTSQADDLKAADAVADQETLQREKAAHLELSEVIRREEAAKSALETQKQVADLEKALKEIQADESEVROAADKQLAQARELVASIEERSIQADLKAQVQVLRADANRKLQESLRLQEVEAREVALRERHSLADVAREKQV-ESEEAASLREWEKRLLEDGRMLRQEGERLLNERENSLKERDEALQINREVAEAR-SYIEKERVLIQKSDVDLNARAVAFSEKERALSERELEILKDDQDLIAERIAADKTR-EFETREQQVRETEVYFGQERTRLSDFETALKFREESLEEKHELAMEKFLKSH-TSDVDSKKAELLTAAEELRSVRKVLAAEKEEVEETLKLVAESREARLRLHLEAIT-AREEELELTVQEVVDREKVLERRLEEVSNLEQGIRFEKKEKYENEHQRIALKEEIKRKAQEMEENKRLKLEQKQIEEEREHLRRECELEQRQIEEEREKVRKRDWEEE-REWEQQRLVQKQDVEYKKEQLELEQERLRGELKAERERQGAELNLRVNLH-NNLEVLREKLKAEVEYERDALRKEIEADQERVAAGLREESRKAIEAEREQVDEGR-SRIRRELELERQQLIESEERAAHAIYLERQKFEDEQEKLKALEQEREELVRIQVQ-LKQIEIDARARKQFVDEEAQELKQKDRFEREWELLDEKREATRKEEREFEEE-FKRVTWMRDEEERLKETRRFQEQSRRMTEELQKERESEWRSRLERTERNQY-LAQLDAERQELNRNLERQREDLRRLLESEAFAKQFEEREAQRAEVEWQEKED-LRKNRGSVIGELEQLRAERSKLEKEROELLQORADAKEWDEIKKDHQVQVQ-EKLREQRQLSHERQNTLLETERLQKLRDQMGSEGSMSMRVSEQPMRMDEE-VLWVSPHSHGLVTRDTRLSALPFVLGTHHASSQTPSRMIARTPSRLAWLQRCAS-RASQLFLSPNKLTTGQEPILLEKTDDEGEPRLGANDPCSSFNQSQLGQVLDL-TEDSHRFKRTWSTQRVVEEANTIPGLQEEKNFESNRNSNVVTFPTDASGDR-KKKRARGNVDDVPLQSQPADQDGGTKRKKRIKDIMVESETNGDSDLTPR-SRVGTPATKRYNFRPTTIVNMMGASENESSRHHHSNKAASAANQFPAASVD-RLPDASSQFVQDTEIDMHEAPTVEEFGFAEREQNNVAEDDQERDITAAVAHF-PITQVVTETTTVTETIREQAVFDLNGVENVIAIRTVPEQEGIPTAGEVILSRV-WLPGEAEATAGDSLVQADEADARDAQNSQSDRSDEEVAEVEDDSAEVEDEDG-NGTESGENDAASVVGESLSEEESEAEVEEVEEDVGEVEDEDEYDTR-EDPEDDGPTPTIREKIWDFLTT</p>
<p>Pohlia nutans</p>	<p>Embryophyta</p>	<p>NCBI compound of: GACA01014689, GACA01016482, GACA01042116, GACA01016308, GACA01016301, GACA01041377, GACA01007203, GACA01001674</p>	<p>EGRQMSGLSPMYTPQGRGSPQQRGTPLRALAREKGPNTTTPVGAVTTTVA-VATAEAMALMDPLLGSPDMINGADGHTHESEVWKRFOSEGALDMPSLERKDRALHARIAALEAEALYDYQYNNMGLLLLQRKTWTSQVDELKA- AVVDAQETLQREKAAHLELSEVMGEEAAKNALETEKECVADLEKALKEVQAD-ESEVROAADKQLAQARELVASIEERSVQADLKAQVQVLRADANRKLQESHL- QVEVEAREVALRERHSLMADVGARKEQVEAEASLREWEKRLLEGVRLPDGE-RLLNEREDALKMRDEALQTSRELAETRSFLERKRALIQQSDVDLNARVLSLSE-RDRALSERDLEVLKKEQDILLFEERIAADKTRFETREQQVKETEEYVQERARLN-DFESALKREETLEEQKIELAEMERLKSHTSDVDSKKAELMAEELRSVRKTL- AAKEEVESSLTAAEAREARLRLHLEAITAREAEFELRVQEIVDREQVLERLRLLE- VSNLEQGIRVEEKKYENEHRIAELKEEIRKAKEELENMGERGELVRIQVQLKQEID-EIRARKQFVDEEAQELKQKDRFEREWELLDEKREATRKEEREFEEESKRMAE-WMHDEEERLKETRRFAEQSRRATDELQKERESEWMSRLERQNLVQLDSE-RQELVNLRESQRAADLRRLQERDFAVLFEEKAEQRLRAEVDQEKEDLRMINRGSV- IAELEQLRAERAKFEKERQELLKQRDFAEKQWSEIKKDIEQLQVQGEKLRQKRS-LHLERQNTLRESERLQKLRDQLKQSEGSMSMRVSELPRIIEEEVSPHHGQLLA-RTEQTQTPGLPPPLTFRASSHTQPSSRRMIARTPSRLAWLQRCASRASQLLGS-PVRLLTGQEPVLEENAEELKQKQDANAPSSSNQSQLGQVVTGTEDGPKFKR- TRSIVRVEEANAFTRHGHGEEIEISESNNRNDFFTPPAAANADTRRKRGRGQNV- DDDPKPEAIEIHGSTRKRIKIDILVENETNGDSDPADTPHGRALTPAAKRYNFRPTT- IVSMMAASENESPRHHSTSKRAASATSQPVAIDELPEVSSQPNKEMETEEHG- PAVGEASDAQVGAVENTEVEAT</p>
<p>Amborella trichopoda 2</p>	<p>Embryophyta</p>	<p>NCBI XP_006855781</p>	<p>MLSPRKKENSGTLKSPGSRVSLQSPSIVNGDEALWTRLREVLDEETLKQRD-KAALISYITKLESEMFQYHMGLLULEKKEWTSKYEQIKASADSAD- KYKRQQAALLSALAEAQREENLQALGVKEKCVASIEKALHEMRAECAETK-VAAETKLAEVRCVLEDQAQKLLAVETKQHTVEALQVETSWQ-HAVAERKLEVEAREDELRRQVSLKSEMAKEKDLLNEKE-SLRELEKVIQGGQKLFEGQTLNQRQEQIKERSDRLSRLEK- EVQAAATVKLQEDLEILKEEKANCLTSVALLTREEAIVQREVSF-DKKEQELLLLQKLSREQDEIRRLTIEHOTAIELRESQFEE- ELHEKHKSFEADLQQRHALDRDAELKHQEDLMDHDKHE- LDQLSELEKREKLETLGLKSLVEKQSLDAREKKIEMERN- LEKENQELDVIKKELDVRVNSLENERKQILEEQRKLVMN- DRKDLLELETLEKVEVDNLRAEKVKILAEADNLATEKEKFE- KEWEQIDEKRELQKAEVWAERMEKSKFLKTEHILLN- EKDSLREQAKRDADSLCEREAFLESEMHGHEWFTRIQ- RERADVFHDIEMQKREFQKGVDKRNNEIQRYLREDDTFQL- ERLREFQYIDAKELVRKELGISELMKKLENERKNIALDRE- QRDKEVSELKDDIELQVQREKLEKQRELLHQDRDILKRI- EDLKKLEDLVPSSETLMLPEMQSTGLNLNEVKTTPANYLVG- CATKAAVEVHADECNENANIGAKSELLEQKESDSDVPTPKS- WLKRCAEKLFNTSLEKIVASANNYETHFSRHKEGQGPL- SFSLRQKSHRTRDARRVKTFLSRSTRPVPDEKNAVLEGLP- VREEKDHQDFDAETNVSANKNGSCNIKSSVDFSERQTS- ANNGGECEFLYGRKRSRQYTSIEDADAQFSRKQSKRQQQ- APTAEHPRMGTSAELLVHSPVHPGANGKPCSHIPDVR- EVVDGGPSNGPAKVREGGEASGIVLEGLNNSKKDALES- QTSIDIEGKAYEVATQPHDNGVLAAKFDEQIGANSSSPEVTG- WTFVMPALFIFAIEAIKVLNLSLWHRFACADGNLIDVPSGAVALPF</p>

<p>Amborella trichopoda 1</p>	<p>Embryo- phyta</p>	<p>NCBI XP_006849 769</p>	<p>MFTPQQQPTRKAGGIWPFSPNSNKGPPVTRLASDPDPNITGILLHGGD- KGRAGKGVSSSFPLQEEEQATPPPPRASLLSDENGRPRPYSSSASEVWRHF- REAGSLDLDLQKKEKDALSHLAKLEDELFDYQYNMGLLIEKKEWTSKYD- DLKQALVEAQESLQKREQASHLIALSEVEKREENLRKALGVEKQCVADLENALHEM- RTEFAEIKFTADKLLAEARSLVASIEEKSLEAEAKLRSADAQLAEAASRKSNNLRO- LQEIETRESVLRERERQSLKAEREAHETTFNRERENLRNWERKLEKQERLVE- GLLNQREELANEKEMFLTKKEDKLEVAVAKFEKGNLNLKDEKEMNMRLRSLTA- QEEEAARVRRNLDEAQQLHLLQEKLNAREKEGQKLLDEHNAVLELRKRFDIE- MDQKRKSLSEEEFEKQVVEQKLEVDLKEEKINKKEQLLEKRETEKTEKEDLE- LTKSLKEKEKFLKIEQKDLDTDKKKMVEKADLHSLKLELERIKAAVEEKEKIVK- EQENLKVTEDDKRELLRLQSELKQEIDEFRLQKLAVEREDLKLKDEKFEREWE- VLDVVRDEVNKEVELHNVEKDEFKRRKEEELKLRKREKQKSEKFORREYEALEL- QKNSFTENMNHRSVQLNARRERDDMIREFELQKNALESSIQNRREDMQKFL- EKERDFQEVRRERMWKEIAQRELAQKEMEEMKLERLGRERQEVALSKKHVE- GERLEIQKDEQLHILTKLKEQREELRRERDRILSRIEHLKRGQGGSDIVDTGLAL- SELQSFKEFENNGNLLPRLLDGYMKESMQGRSNVGPNSLMEETPPLGAVLNS- TSPARFVSWLQKCKSIFKLSGKRLDEQVTNQEKSPPSDVEADADQILENDSGGLV- SGGANYPEISVGIQSQAVDFHRAASPEISGRGDEEETVTPSAADQEQSDM- LEMQEGPSASAEIHSPPAAAAGGRARKPRRGAAPLRRTRRSVQDVVSKAILG- ESSEELKTEEEESAQANVDSKGGQIVKGGRRKRQHPTTSRTMSEQKSDVDR- SESVTRGRSKRRQIEPQSHIQQPGRRYNLRHSTLEKHVENPVGSQALASQVITD- ADENHSQHTKSPGEVVEGQTSNHHHPDEPSIESLENAHGGGEAKTDVRLMLQHT- KFESIVEIHFREFSTQKVIIMETGGALEETDVNDPQPNSSQEQPQANQGGANDLLEY- DEDGGGSGRGGEDDDGNDDDDDGYVNDENQEASIGKLLWTFFTT</p>
<p>Musa acuminata 1</p>	<p>Embryo- phyta</p>	<p>Uniprot tr M0RRY8</p>	<p>MVLKIGDRFESNRTGASFPTRRPDRRRIACIKGDGVDKRYWVIGSPIW- DREALPFVLFVFLKFLDEELGFLMFTPQRKGVSPSPRYGDGVDNRMTTNAVNTRT- GSGVAFKKGKGSVAEALPPPPPLQALLGENSGVGVQGDGDAEAVWRSF- REAGLLDESSLQRKDRDALVQRISELEKELHEYQYNMGLLIEKDWASKYEEI- RQALAEVDETLKKEKSAKLASIEFAKREENLQKALGVEQQCVSDLEKALREMR- SELAEVKFTSDKKLDDAHALEIGLEEKYLEVEQKLHAADAKLAEASRKSDDVDR- KLEDVAREHKLQKEYLLFDQSRGLHEKDIHQREHLRDEWEQKLDQSQKRLVE- TORYLNREDRTNEADRVLKKKEADAEARKMIEATKKSLLKTEEEITKRLGSL- AAKEKEVDVVESENKEDLISREELNAREVEIQKLLDDHNLISSKKEEFEL- DLEKRRKSLKSEIECKIREVEKRRREIDSMEEQITKREQALQMNQLKMLKDEED- VDLKSNDLKKWEEVQNDKLEKERQQLASDSEEFKSKSDLESKAAIESRK- EQIMKEEENLRLTKGEREEHLLQSNLQOESDCRILKESLLRDTEDLQOQKREK- FEEVEVLEDEKRLALEAERKFNDEREKFEKWRHDEEERLNNEALVARANFER- LEELNQKTEAFGEIMEHERLEALEVLKREERADMALELELCKHELEMDMQKRL- EDTEKLLDKENDFQRKRLDFNQMSLSSNDLKIQLKMEEDRLEREKEDLS- SYRKRLEIDRLEIQKIDALRMLSRNLKEQREEFMKEKERFLAQAEQKCTCKNG- LLVGDLDTFICQDAGDVQLPNLGFEELNDTNAETTNAKVSPPAASGGGRMSWLQ- KCSRLFNLSGKGVLDSSQHPDNSNLYSSLDREAFDGEASHKPAASVYVDD- SDSQRAQSVTIGDQNVESKRLCGVVEPEPSPFVANNSHIMIRTQTQMDNGVR- DVVDQLAMPVSVLNDREKYAPAGSDNLRVSKQRQSQPGRGRPKAVKRTHTI- KAVVDAKALEQSSDEKNGHPHNGEAKDPRRVAQTSVGTNSDPPDAEDSEAH- SISLGGRRKRQILASAVPEKRYNFRRTIAATTTAAQTMSDQTKGFIAGYD- RQLTGNELKEIGGESSRPAVEPVSVDVNSIASNMLQKTAAVGIAEVEISSQK- IVQAESNDTVKSVVEVSYQSGEDGHILDDAATGSRPATPSDEDEDEEECEQQNA- SVGRKLWTFFTT</p>
<p>Musa acuminata 2</p>	<p>Embryo- phyta</p>	<p>Uniprot tr M0TQE3</p>	<p>MFTPQKKGWPGWSPSPRVGDGVDNGMTTPVNTRSGSVLAFKKGKGGKGN- NTAEALPLPLQASLGENGDTVVVGGGDAEAVWRNFRFAGLLDE- SALQNKDREALVQRILALEKELHEYQYNMGLLIEKDWALKYEEIQRALMD- VEETLKRKLAHLASISIEFEKREENLQKALGVEQQCVSDLEKALREMH- ELAEVKFTSDKKLDDAHALEAGLEEKYLEVEQKLHSAADAKLAEASRKS- VANRKLDEVAREHKLQKEYLLSSEWKLHEKGITEQREHLCYWEKLL- QDSQKRLVESQRFLNQREYQANEADRFHKKKEALEEESRKMIEATKKS- LKSKEEDITIKLSIAAKEIDVVKIESLGGKEDLFSREETLNARERVEIQK- LDDHNLISKREEFELNLEKRRKSFADLEKGVHEVEEKKREIDCME- DQVKKREQALEINLQKMDKEKELDSKASKKWEESVKNNDERKLEK- RQHLASECEELLKCNSELESKAAIESSKQIINEEENLRLTKVEREDH- LLQSNLQOELDCRLMKELLRDTEDLQLRKFFEEVEVLEDEKRLALEA- EIKFNDEREKVEKQWQHEKERLNEALIAKANFERELEELSQKEEAL- KAMEHERLEAFELLKREHADMDRELELRKHELQMDMQMKMGMEKLL- LDKENEFRTRDLELSQMSLSSLDNSKSKRLKMEEDRLEREKEDILSH- KRLEVEQLEIEKIDALCMLSRNLKEQREEFMKEKEHFLDQAEQKTC- KNCGHPLGDMGYCILDAGNVLLPNLVFEERSNMMNAKSSPNAMVSV- PAASGGGRMSWLQKCSRLFPSPGKTSCKPVVSHGVADFSYRQENKEP- RLGAEAGEPEPSEVADNSIDIMRTWMDNGAREVDDYVMPVFAQNER- ENFAPAESDTLPESLQRRSQPRRRGRPKAVKRTGTTKAVVTDVKAIL- GKSSNEKNHGSQDLVLANSTTSAGQKRCVAQISGMTTSDLNLGDSEAH- SISLGGRRHKKRQILAPAAQIPGEKRYNFRHSAIAAVTTAAQTIFERTK- GPKAGGHEDSTGNEIPMQSGGEEGSAARPVVEPVSVDVSKKASNMLQK- TAVESTTEVHEIFPNKIVQAESNDVKSVEHSDQSEDFVDDAATGTD- PATPSNGGCSEDEDEEYDQLNASIGKLLWTFFTT</p>
<p>Musa acuminata 3</p>	<p>Embryo- phyta</p>	<p>Uniprot tr M0T513</p>	<p>MFTPQNKGWSLSPRIRGGADGSGSTANPRGGLGLLSTKGGKSVVEAAPP- PQALLGDDGEDAFGGSTEVAWRFRFAGLLDQSVLQRKDRREALVQRITELKEL- HEYQYNMGLLIEKESIARYEEVQALAEAEIILKREQTALH- AISEYKREETWLKDLGVEKQKVSALKEDLREVRFEISEVKFSSE- RKLSEAHALETGLEEKYLEIARMHAADAKLAEAGRNSRNRK- LEDIEAHERKLQRDCLSTSERKAHEKDLLEQREHLFDWEKRLQ- ESQRRVLEEQRLLNEREDSANEADHILKKEKTELEETREAIASKR- SLKLEEDDITIRLSSLASKEKEAEIKMGSLEKERELFAREEKLS- RERVEIQKLLDDHNAMLDSKKEHEFELEMENQRKSFEEMKAKID- EVEEIKKELDHKEEQILEREHALENMKLKEKMEKNLESKSQALK- RWEESVQIYEKLEEDKQQLDRDADIVKSISELSLKVITIAAKE- QIKKEEELRLTKEREENLQKSKLQOEIEDYMIMKDSLCRDS- EDLRQOREKFEVEWQLLDEKQLALELETKQINDERVRFKQWQY- DEEERIRNEEKAKRSIATELEDLRMKKQAFKTEMEHERLNVHEM- LTRERSAVAREFLRQDELEMDMRKQOAMEKLDQDRESEFQR- KMTIELDEIRSVSSDFELKSRNLEMEQDRLEKEDLSAFRESLK- TDQLEIQKDITLRLVRELKQREKVEEERDRFLGLANQKFK- KNCGSSVCNLDLLGLQNTDVVQLPSLTFEDRLEAKDSETSPRH- MVSFVSSGGRLSWLRKCSGFFSFPKGSSEDTAQNVKNPISL- DVRLAREALDGEASDEPAPSQGFIFAKSFDTRTQSDSGIRDNE- VSKRLGRAREELESFGLSVPPRNESQEPSPNEKPRQPKRSGR- PRKISRTRTKAVVEEAQAILGETSMGKNGQPNGLAKRSLNIOE- STEGNLVHAGQKRLTHISVAAASELDGEDSETRSESISLGGRR- KRRQINIPETQTPGEKRYNFRHTIAAARSISDQTKGHKRG- HQQPSGDESLRGDGDGEGTSLRLDVEPASSFAESLKSVD- QKMAAENVLDVQEIFQKPVSHIEIECHADDAGKSVFSKQGTI- EGVMADGATAVEREPATPDDGCSEDDSDAEENSDDQNESS- IAPMEKTEQEAALYMIQGFVQGRWNGSDLYPDCGWTOIQGVS- CDLFDGLWYVITLISGPILENSLECTEKSEFTNLGQLRNLRLSLV- VDNSLVGELPMEELGNLIQLKRLMLSGNRFSGQIPASLGINLQLLI- LDLGGNSLTGSLSSCGLSSLKLDINSNRLHGSLLPGLNLSH- LALLDLRNNLSGVPSKSLAGMESLQVLLSYNPWGGSLLEFEW- KNLRLNLTLDLHMLGEGTIPETIASLKRRLYALDNNHLSGIVSSK- FAALPSLTALYLNGLNLTGELEFFPERFYRRMGRFASVWNNPNL- CCNAAMATGSAPHGVAQCKQDQEPSANGSNANERVDRNDP- QNSGLSTSFLFPASSISGFVWVIVQEIQTWVWVLEL</p>

<p>Musa acuminata 4</p>	<p>Embryo-phyta</p>	<p>Uniprot tr MOSIR4</p>	<p>MLAPQKKGLFLSPRAVAARRNGPAQSSFGNWWGGVVPGRK- GREYVAGNAIPQPEELPLCGDGEDREKEQSEAQVWRRFREAGFLDEAVLQR- DREALVRRRISELEKELYQYHMGLLLIEKKWAVKYDKLRQEM- SEAAQLQKCMQAAHIVAVAEFEKSEGNLRRAMGFQRQSIHLEKALNDMHAIEA- EVLKDSQKLLSEAHTLEATIEEKLEIKEKQHSLDARLAKVSRKSSEVDRLEDV- EAREHELFKQTSFSAIEKKAFAEKDLRQRENLRWEQQLQDNQKLGKWHST- ENQREMETNERDNTFRKKEKELEEARKTLEISNELIKLEEDMCMRIGALDAKE- KEALLKQEFLEKKENELLAIEEQLNKKERVEIQKITDFHNSILESQKDEFELETK- KKRAVDEQLQGRIEEEVAHKEIILENRELERFKKEQLLEIREIGNLKNREKENDIMLS- AVKVSIEENEKEEMRQGRGKLEKEWELLGERRLSLEEGLKQLFDEKERFDQWR- CTEEERLRKENPEVSIHAQMLEDSDISEEAFKDKTTHQKMDVLEVFNSENAH- VVEYIMQRIPEKVKETLLEKEDNSNRRSNIVLNNCKLSSLDESNIKLEKQEDQL- KSEKQLVLGKKEAGQSTSGTLNRNNKQVVEPAEEGGLYPASAEQLKACRY- CGFEDGGDTALSGGSVEVSDQGTCPGSKLEARIPCMQRCSRLNFPSPGKKA- TEHSEKSVCLDGEPLHEHDNLEPGLPGLVDFVAFQWAQAGGQVQYNAEPERS- NNDDDATKRDSQIADRSADILIFELNDRVRDLEETLHSDVEQKYREGCSIRPEL- NSLLWPLKQKQSGRSVRRKSLVKKSRVSNALVEDANLEEASQIKHSEQSTCRA- QCLIKDKCLEEKYSLNDDVTVCSKKRCLDKYGMMSLEGECAEAHTEDVSSL- GCCFQMENIPGTEIPGLKRYNLRHSTIVRAVAASQALACRTKQKRKGELELSL- NKLVKARRDGEAESHASASETLRSDIRNS</p>
<p>Musa acuminata 5</p>	<p>Embryo-phyta</p>	<p>Uniprot tr MOTBW8</p>	<p>MTSPRPGGTPASRAVNASKTPATGGTPLGDDAIWKRLKESGLDEESVKKRRDKAAL- ISYITKLESEIYEQHHMGLLIERKELVSKYEQVKAASSD- SAEIAKREAEAKRSSALAEARKRELNLKLLGIQKECVANIEKALHONLVESAERKL- GYESKIAEAHAMMTAAQEKLEDAEKLLAAESLQAEANRTRNTAIRTLDDVEARE- ELRRRLATFKSQDCAKENEISIQRALYEQKTLHQQERFLEGOTLLNQREYF- ERTKELNRIEKELEESKANIEEESRTLLKERSNLDLIEAALRNREEVIVKRESMLDKR- ERELLILQEKIACREHDEIQRIEMEEHOSILEKKSELEADIEQRHLLKLEAEKIAE- IREADLCSREISLQEKHEAIELQSSVLAKKQEDVANLRLLEDKEHNLSTTKREAEIE- VQNMOKEREIFLKMVDLEKTKAVLEDEKKEIILAEKFEITLGERNELLLENLKEE- IDSLRAQKALVAEADILKAEKFEIWEWIDEKREDLQKEAERIDEERKTLAQYLK- NEHDSIKLEENLHNFKRDVERLSCEEEFICEMDRQHSDFWTKMQRENF- KDIQIORNELNSINERREEIETYLREKEESEFKDKVKELQLINSQKDMIQKLEHVA- SEMQLNTERLEIAQDREQREREWADIKRFTALDLQCEKQKQRELLHAEIREEN- QKIQQLKLEELQIENRALSVMQTDKCDASVGGKSCQINGADRHIATPNGVSTM- KLLPQGTNPSTPTSVTKSWIKKCTEAMFKHSPEKSDDTGHEENVESKMLAKRS- FRFSEMDLQGHGNFAEGKEVSVQEMDNFTPKRTKSNRQEKVNGQEKVCRNCF- EQNMISDARPAKSAQSPSEVANSIKFNOALEDSGQKSRFLFSSINSVWRRRKR- NDMLSHDHADMSEPMPKQKPRQNGNSDVEGDSNGLAEQQVQPIDDECEPV- LNRQTSQCEQLHAVAFAKQDQHENMVVPAEPIESSQHKLAVSNFDIVENGFKF- EHSPLAGVGAATSSDANEISMOKQVFDKEHIAKPSQETSVSASDLIVEDNDKLE- EQDRYNEVLEDELEDDGSGLSVKEKIWNFLIT</p>
<p>Oryza sati-va 1</p>	<p>Embryo-phyta</p>	<p>NCBI BAC78596</p>	<p>MFTPQKGGWGTWSTPAPANQRSGGGAPAAASAPLGGKGGTTLRVAELEQEL- HEYQYNNMGLLIEKKEWTAKLDEINQALTQKEEILKREQAHHNAISEYERRRESM- RKALGVEKQCVTDLEKALREIRGEIAEVKFMSEKKTDAQSLEASLEEKRLIEGLK- HAADAKLAEANRKSQADRDLEEEVEARQRRLEKEKLYFENERKAGEDRIK- RQEDSLRDWKKLKESSQRILDLQRSNDREREANENDKLFKIKQEELLEE- AKKALEHTKATLKIKEDDINKRLAELHLQEKAEAKNRKLEEREKIAEREE- KVSAREKVLGKQLLEDHNVKLESKRDFDLQLENEKSFDMVLVQKEADL- VQREKDVRSSEELSKKEQVLNESKKELEEWQNDLDTKSNALKKWEESL- QNDEKQLSEQLQIENERKQAEYKLELESKATVAEKEKILQEQNNLKL- TEEERQEHIMLTAQLKKEIDEYRMRNSLSSEETEDLRKQKQKFEFEEWQL- DEKTRHLEEEAKLNNKKNLRLWHDNEEKRLKDREDELDIKYKEQGENL- ALKKSLIDNIDHQRLNEEELLKRERADLQRNLQHRHELEMEMEKKQASK- ERELEEKENELNRKMDPVENELKRAAELNESKIQKLEKQKQKEKEVLY- EDRQKLETDKADIRRDIDSLNLSKSLKERRERAYNRDRNLDIFEKYKVCNK- CGVIFEGLDALAKDSTDIIEYPLAVEADDRSPNPDTLAQETGALVNSGGR- LSLLQKCSRIFKFSRPKKAEQSSQAVKNTDFGARLEEAQSDDDYDEPT- PVYQVAYNSFDAEDLPSSEGAFFENESERQDIADDVQMESSLGVADNCVD- IHGTQSFQDGMTDMVVDTTIVDQNGKDSAVLPVVDLEPETSQKGRQR- NRKGRAGKGVKTRSVLAVVEDAKEILGENLEVKKDDGGQDSVTVGGR- KRRFAGATISEQDEDSAEHSESVSLGGQRKRRTAAAVTQAPGEKRYN- RRTTVANAATAAQNTKKAAKKQKQVTEATADDTGETSKAEAPATGSKG- ASQSDASQLPEYSQAEAGDTHGPVEVTSAEQVDIVDGDIAAPDAMPMT- PSGSELGAEQDDEEDDSERRNQSIGKKLWSFFT</p>
<p>Oryza sati-va 2</p>	<p>Embryo-phyta</p>	<p>NCBI BAC78597</p>	<p>MASPRAGGVGGGGGGGGGGGAAAGDDAIWSKLREAGFDEESLKRDRKAAL- AYISRESEIYQYQHNGLVLMERKELTSKHEQLRAASEAIEIMHKR- ERAATQSQALAEARKKEENLKKSLGIQKECVANLEKALHDMRGE- TAEATKVSYESKLAELQLMEAAHKKFDEAEKLLAKSLAEASIRTHNAALRSLH- DIDDREDQLRRDRISCELENAKEKEISLQRKSLNDMKKILHEKEVLLKEQALL- NORDENILRLAYVTHSEKRVEEKKNLEAERKYLLEEKYKLELMEIAVSEAL- IQKESLLDKRESELLIQUETIASKERAIEJERLNOEQAIALERRKHDFESEMANKQ- MSFDAEMVTRNALHQRECALSEQESVYVQRSONLDLQALASKEALAGR- SDELKEEEELKLLHRAIHNELQKEREIEQRIKSLDEKEKAFFEEKREIAQOQD- LAITQADRDLELTLQMKLEEIDSLRAQKRELMAADRLQAEKERFEIWELEDE- KKEELQKEAIRAEERRAITEYLKNESDIIEKQKDNLRVQKSNSETLSREHKEF- MSKMQQEHASWLSKIQQERODLKRDIQIRVELLSAKARQMEIDSYLREER- EFEQKKAKELEHINSQKEMINTKLEHVAVELQKLDKERKEATLERERREQLSEIK- GTIALNNQREKLOEQRKLHSDREAITVQIQQLNVLLELKDSENKQLSLQHD- KSKLGSIDINVKDHHNDSHSPKQRFGRKLDLSPVSTPISVWRKCAQVIFKRSP- EKSASHDQFVQNGVPKVGDSDVDVNDLFAKVGQKRLNHLVSCDQTEVLE- PKRKHRRSTIQVNGGEITSNCLSALEEKSKNEHDEAPLGLSNTCKEHEYGD- KGPENLTKPGEPASSVDVPYVNGIVDSDSVQEEPSVEATVSATETSNDVQPE- DNNDSDDEEEDDEEETSSAKKLWRFLIT</p>
<p>Daucus carota 1</p>	<p>Embryo-phyta</p>	<p>NCBI BAA20407</p>	<p>MGRVEDMGLNAKLMKLELFDYQYNNMGLLIEKKEWTSKFEELQQVYETK- DALKQEQEAHLIAISDAEKREENTKALGVEKQCVLDLEKALRDMRSDYAEIKFTSD- SKLAEASALITKVEEKSLEVESKLSHADAKLAEALSRKGSDIERKSHELEARESALR- RERLALNAEREAALTDNISRQREDLREWERKLEDEERLAEVRRLNQREERA- NENDRLYQKQSELDEGQKKEIIMVSLKNKEDDSSRIAKLNIKEKEADAVKHS- LEVKEKDLTEFEQKLNAREQSEIQKLDEHKAILEVKKQSFEMEMDKRKNDFE- NDLQNRAVEVEKKEVEVKHLEAKLAKREHALDQKHEKLEKEEQYLASKLQDL- NEREKSMKLEENKIEDERNQLSDKQEMCLKAEIEKDRASTEERQLKLESEIE- RLKITEERLELARLQSELKQEIENCRHQRELLKEEDELKQEKMRFEKEWED- LDERRTALMKDLKDITVQKENFEKLKHSEEDRLNKKLDTESYVQKELDARL- TKDSFAATMEHEKAVLAERTSSEKQMLNDFELWKRLETKLFNEREDMENA- LRLREKQFDEEREKELNNINYIKEVSKEREDIKLERSRIAKEQKIELMHQKHL- EQHVVMQKDIGQLVSLSEKLDQREQFFKERECEFRFVSEKQSKCKNGEMTS- EFVSDLSLAELENLKAISVPLAENYLRQDLQGTDPKLNLSVTPGAVGLGS- PASGGTKSVLQKCTSKIFIFSASKNNNSPDQNTSRRLHVEASPNKLLNTEV- ELPSVAGETLEMQNMVQVNSNREMSNLNLSGTEQSNIDSKALDVEDSQ- SDVRAGNRKPKGRKAKGRVRRKRSKAEVAAEAKTVLADPIELNENEHSLAS- AYTNEGRDSSLVGKRTRNSRKRNPSPQSSQAAGDVGADSEGHSDSVTAG- GRQKRRRVVPAVQAPTGRYLNRRHKTAAPLVANALSDPNKGEKEIDGG- GIGEEIPDEVGNTHLVQVTLKRRINVVNEFSSAGFHGINATSESQDRDAANQ- LVSDTMLSSEVNGTPEQSRGYQKQGDTSAGEGEDGDEVEHPGEVSMRKKVW- KFLTT</p>

<p>Daucus carota 2</p>	<p>Embryo-phyta</p>	<p>Uniprot tr N0DLR1</p>	<p>MFTPQKSNTRNSLIPTTMSHTNPRSTNKAKSVFVNDPAPPRALLGGDY- VAVERGEEEDWRRFRREAGLLDEAAMERRDRDAVVEKVAKLERELFDYQYN- MGLLLMEKTEWTLKYEEMRRAQVELKEVLEQEQTTHLLLSESEKREENLRKALD- MEKCCITDLEKALRDSGADNAQTKQSSSEAKMVKANALLSGFKEKSMVDVETK- LHVADAKLEEYVYKTSLELERLQEVETRDLSLQRRERMSFIAEREAHEATFSIQ- KDDLQEWKLLQEAERLCEIRRTTSGREVKNMEMALNLKQELNKAQK- ENDLSTSVLKEADDINHRLANLTAQEHKAETLRNELEMRDKELLALAEKLA- RESVEIQTLLDEQAVLDAMQEFVDMGKRKSLDEEMRSLDVAQYKKADE- ITHIEEKNRLELLENKSERIKEKEDLESKLRRLTKDKESLLKSDKRLDLEK- KHMADKDTLQTLKDEIEKTRADISQQQSKIQEEIVKLKISEDERAEYIRLRESEL- KEEIEKCRFEKELLLKAHKNLKEKDRKSFEEKWEALDERSNALSREIKLIGEEK- EKFEKFRLSMEEKIKNDRLATEDYIRRELELETEKETFATITRQEQSLISEKA- ELEYSQLHEFELRRKDLVDIQKRRDELESHMSEREREFEEEREKEHNHNS- RLKEVAQKDMEEELRSEKRRIEKDRQEIALLKKEKELKEHQLMHHKDIIDEVLN- KVKIQREQFIKERDRFLFDVDTLKSCNYCGGCTREYELSDQLLLEKIDNSPI- VELGPGVSYESQDRINLRSSNSGGHISWLQKCTSKIFKYSFGKAAQDSEFQ- SDMLATVEEDERPDSGHLETRGLNIANDGPEPFSFIANESCEIHLASNDNK- RDADQRHEICTDELSNIDSKAPVAPEDSQSSELSSGRRRPGKTRSGSVAV- GTTKRKRQAQPSGVMKSAVTAHSEEHSESVSEVGRRRKQQSVTSSVQTP- GKRYNLRNKIVGTSGSALASVDLKVSEVDVNTKTVQDYALASSQVIA- SEKDNPTGPLENMTCRSLIEYDLSTEGDVELKTSKSRDKSIDPAIMGNIEFNE- EVNSTIPECSIENGRGSTLHEDRDNEVEVEVLNEDEDLDIDSEGDSVSIHKKLWTLFTT</p>
<p>Daucus carota 3</p>	<p>Embryo-phyta</p>	<p>NCBI BAI67718</p>	<p>MASPRLTVIQSEKTTVTSSRVSRSSMSDDDIWKRLQEAQFDEDSIKRRDKASLIAY- ITKLEAIYDHQYQMGLLIMERKEWGSKFVERVEAALNSAELMRKHDKNLKLD- LAEAKKREENLKAIEIERECLANIEKTLHELRAEYAEKVMAD- SKLVEARSMIEDALKKLEADAKKHAASL- EAEASRYHSAERKLEHEVEAREDDLRRRAT- SFKTECDTKEEELHERRLLNERQKALQOS- QORLVQDQLLNKRESHIFERTQELNRKEK- ELEASLQEEELQALVEQANLETKASSLS- LREEVITKSELEVKKREEELCVLQEKLEKKE- SERIQQLANYEASLMMKSEFEALVVKR- KSVHDDIENKRRDWELREVDLHREELILE- KEHELEMQSRAVVDKERDLAQRFLSLEEKEN- RLHAVEKEIESKEALLQKEKEIISKLDIQ- RSLDALEDEKQKOLHHAEEKMEAMKSETNEL- CVLESKLKEEITIRAQKQELETEADEMKEL- KLKFEIEWQSIDKREKELKQEAECINEQRE- SLETLKDERNSLKLKEDAMRDEYMRNN- ESLSRDREDFMKKMEHERSEWFSKIQKERS- DYLLAIEVQSKDLEDRLAKRREEIESYLAER- ERAFEEKKEKLMRMDTLRETLARETEQV- NAELNRLDTERREINLDRERRDRREWALNT- LIEELKVQRQKLEKQRELMRADKEEILVQIEH- LKQLEDLKVVPDRIALTDIQQSDLQPSKR- VSARRSLKRQSGLDSGCAEDNGNASSGN- GSVILSPPLSSPFSWLKRCASSLLEQKVSNNK- MRHSEIITPSTIPARLNAPDDEHAVISANQQT- VHAKETTVYIDKIITIREVTSFNDAIVD- GNNQNLEALSQRAEEKLEDNNIESEKLEKNGEVDPKIMQASLTEQ</p>
<p>Vitis vinifera 1</p>	<p>Embryo-phyta</p>	<p>Uniprot tr F6HSF7</p>	<p>MELLSEIKKNSVHVRGLGFLHLSHSASPTLRVVEIVMFTPQRKVVWSGWSLT- PRSDAQKNAAGSGSNLSPRNGGVDGDSVSKGSAAFVEPVTGGENGN- MVERPGEVADLEALVAKVSKLESEIFEYQYNMGLLIEKKEWTSKYDELQOALVD- VKDALKREQDAHLVAMSEVEKREENLRKALGIEKQCVLDLQKALHEMRSEY- AEIKFTSDSKLAEANALVTSIEERSFEVEAKLHAADAKLAEVSRKSSEIERK- SQEVDARENALRRERLRFNAEREAHETTLKQREDLREWEKLEQEEERL- GEGRRILNQREERANENDKIFTQKEKDLLEAAQKKNEMTHLTKKEDDISG- RLSNLTKKEKETDAVRQSLIEIKELELEELKLCARERVEIQKLVDEHNIIDA- KKREFELEIEQKRKSLLEELKSKVVEVEKETEFTFNHMEAKVAKREQALEK- LEKFEKEKEFEFSKSKALKEKESIRAEKNELEAKKHILADKEDLLSLKVAE- KIRVIEEQKLKVHEEREQLEITEEERSEFLRQSELKQIEIKYRLEKEVLL- EVEDLKLQRETFEREWEVLEDEKRAIEKDLIDVSEQRKLEKLEKHEEERL- KTEKLAQDYIQREFESLKLAKESFAASMEHEQSVLSEKAQSEKSMIHFDE- LLKRELETDIQNRQEELEKQLQERKEVFEERERELNNVNYLREVARQEM- EEVKLERLRIEKEKQEAANKHLDHQQFEMRKDIDELVSLSRKLDKQRELE- FSKERERFIAFVEQQKCKNCGEITCEVFLSDQLPLPIEINVEVPPPLPRLAD- RYFKGVSQGNMAASERQNNEMTPGIVGSGSPTSGGTISFLRCKTSKIFNL- PGKIEVAIQNLTEAPEPSRQAIPEPSKRLGSTEDEPEPSFRINDSFDVQ- RQSDNSIKEVEAGQDLSIDESNIDSKALELQQHSQHSDLKGARRKPGKRS- KQRIHRTSRVAVVRDAKAILGESLELSENEHPNGNPEDSAHMNDDESQGE- SSFADKGTNRNGRKRQRAYTSQTMVSEQDGDSEGRSDSVMMARRQGR- RQKVPVAVQTLQGERYNLRRPKTTVTVAAAKSSTNLHKKRKTETIDGSGAG- GTGEEIPDCNAAPATSVGLISENGGSTHVLQVETFKTIVDVHFPVSDRVVRL- EAAEDTQDDNADVTKELVENMALSEEVNETPDEGPMYERNGGDEDED- NEDDEEYEHFGEVIGKLLWTLFTT</p>
<p>Vitis vinifera 2</p>	<p>Embryo-phyta</p>	<p>Uniprot tr A5BQE9</p>	<p>MAAFKERNSESGFPFYFRHKVSRKAWTGLSLTPRSEAQKSGGGAVSNPVG- GKGSVAFVDGPPPLGSLSGKAMLTGIDGGDMEDWRRLREAGLLDEAAMERK- DREALVEKVKLQNELFDYQYSMGLLIEKKEWTSKYEELSQUALAEAEIKLKREK- SAHFIAIEVEKREENLRKALGVERQCVAELEKALGEIHAHESQIKLSSETKLSA- NALVAKIEKRSLEVEEKLLAADAKLAEASRKSSELERKLEVEARESVLRRERL- LNAEREAHEATFHQKQEDLREWERKLEGEERLCEGRRINQREKANEIDRTL- KLKERNLEEAQKIDLDLSLVKVKEDDINNRLAELTVKEKQAESMRGILEVEKEL- IVLQEKLSARERVEIQKLLDEHRAILDTKKQFELEMEQKRNVSDEELRSKVHE- VEQKEVEVLHREEKLGKREQALEKRLERVKEKEKLEAKLTKLEKEKLSKAE- KRVGEKQMLADKESLHLLKDELEKIRADITEQELQIHEETERLKVTEERSEH- HRLQLELQKQIDKCRHQEEMLQKEREDLKQERIMFEKDWALDEKRAVITKEM- EIGDEKLEKHLSEEEERLKKELAMEEHIQRELEAVRIEKESFAAIMKHEQVTL- SEKAQNDHSQMLRDFELRRKRDLEIEMQNRQDEIQKRLQERERAFEEERERELN- NINHLKEVARREIEEMKTEERRIEKEQEVLLNKRQLEGHQLEMRKDIDELGIL- SRLKQDQREQFIKERDRFLTVDKHKTKCNCGEITREFVLDLQPEMEVEAFPL- PNLADEFLNSPQGNMAASDGTNVKIXTGEIDLVSQSGGGRMSFLRKCATKIFNL- SPSKSEHVGVQLREESPLLDLQVNLKAEGPSIVGQSAIEDELEPSFGIANDS- FDIQLHSDSVMREVDGHAQSVQVGVSNMGSKEQEGPEDSQSSELKSGRRK- PGRKRRTGVHRTSRVKNVLDNGDERPNDSTYTNEEGERETSHAEEAASITTRK- RQRAPSSRIESEQDAADSEGRSDSVTAGGRKRRQTVPVQVTPGKRYNL- RRHKTAGTVATAQASANLPKRDEKGGDGGDDNTLQTKANPKAASSPLADSD- NPKTTPLVHVTTLKSVIERYSPDRVVRFTKTVDIVGGNDSARLANMELRQEI- PGNPGDTPGYEDENGSMSEEDDSDESEHPGDASIGKLLWNLFTT</p>

Vitis vinifera 3	Embryo- phyta	Uniprot tr D7TG95	MASPPARFSAATPGSRVLSQPLSDDAIWKRLRDAGFDEESIKRRDKAALIAYIAK- LEAEIFDQHHMGLLILERKEWATKYEQIKTEAESAEIVYKRDQSAHSSALAEAR- KREDSLKKALIEIEKCIANLEKALHEMRQCAETKVAAEIKLAEAHSMVEDAQKR- FVEAEAKLHAAEAFQAEICFRRTAERKQVEAREDDLRRLISFKSDCCDEKE- KEIILERSLQSERQKVVQQGQERLIDGQALLNQREEYIFSRSQELNRLKELEK- ASKSNIKELRALNEEKSNLELKLASLTTRREEDVVKREALLNKEHEILIQEI- ASKESDEVQKLMALHEIALKTRKAEFAELETKRKLVDEIEAKRRASELREV- DLSNREDFALEREHELEVQSRALAEKEDVTEKLNLSLDEKEKYLNAAEKDVE- LEKIHLEKEKEEINKMKNIEKLSLSELEDKKQVDAHEKVEAMKSETSELLV- LEMKLKEEDVIRAQKLELMAEAELEAQAQKAFEAESIDEKREELRMAEER- IAEERLAKSKFLKDERDSLKLEKAMDQYKQVEVSLSREREDFMKSMVHE- RSEWFSKIQQERADFLDIEMQKKELENCIDNRREELSYFKEREKTFEKEK- MKELQHISSMKERVAKELHEVASEMKRLDAERMEINLDHERRDRWAEALSNS- IEELMKQRQKLLKQRELLHADRKEIHTQIEHLKLEDLKIADSNIALAEMQQS- NQEPSQRKVVYKRYKAQNTIPNADFESHQKINVVKNGSGFNLPALPDSSSP- STATPFSWFKRCAELIFKLSPEKPSIKHGEKSSISSENANLTLAGNLDSDG- FDREVHDRNEKTHSISDRQPTRYALGEPKIVILEVPSGVEDVKGLHTLESEIKK- DTSSENSHSFSEKELLAGRKRVRVNSSNDVDTTLEQRQKNNKRRQKQED- AADPCGVSIQSDAREGQDVSISLNTQGGAAETNLLITDEIHKISEVTENVV- FQDQAKPNALQNSVVELGDDIQHGGTNGLADSNACVLSDFKAQEKIKGK- EVLVFDVGGVIEHSQPQDESISEKQQELQEQGVPKSDDDKKLEKVGRRMR- SRQKS
Citrus clementina 1	Embryo- phyta	Uniprot tr V4T3E0	MFTPQKKAWSGWSLTPRGEKNGTGSVSNPTTVDGLTGKGSIVAFTEPRT- PQNGVLADDVESLAEKYSKLENELEFYQYNNMGLLIEKKEWSSKYEELKQT- FGEAKDALKREQAHLIAITDVEKRENLKALGVEKQCVLDLEKALREMRSE- NAEIKFTADSKLAEANLTVSVEEKSLEVEAKLRSDAKVAEI- NRKSSIEERKSHELESRESALRMERASFIAREAEHGTFS- QQREDLREWERKLDGEEERLAKGQRIVNQREEKANEKEKI- FKQEKDLEEAQEKIDATNLSMRKEDDINKRLANLITKEKA- SEYDAARKSLEMKEELRQLEELKNAREKVEVEKLLDEHKA- SLDAKOREFDLEIEQKRAFDDDLKSKVVEVEKKEAEINHK- EKIAKREMALEKRLKCKDKEKDVESKLDLNGREKTMKSE- EKNLETEKKQLLADKEDILTEKAELEKIRDANEQQLLKIYEEK- NQLRISEEERAEYLRQSELKEEIGKCRLOEEMLLKEAED- LQKQENFEKEWEQLDDKRAETEKLEKEKLESEERIKRDK- QLAEDHIKREWEALVAKESFKATMDHEQSMITKAESERR- QLLHDFELQKRKLESDMQNRQEELEKDLKEKERLFEEER- ELSNINYLRIARKEMEEMKLERLKEKEKQEVDSHRKLEG- EQVGIRKIDIMLVGLTKMLKEQEQIVKERDRFLNFVEKQK- CEHCAEITSEFVLDLQVEIVKSEVPPLPVANDYVNEKKN- SEMSPDVLASGSPASAGTISWLRKCTSKIFKLSPKKGEN- VVRELTETPSSGGQTKLQESSRRLGQTNPEPLSFAIVND- SFDQRYHSETSTREVEADQHKQVDGQNNLNGKAPEVQ- ENSQPSDLNHGRQPRKRGRPRVSRTRSVKAVVQDAKAIL- GEGFELTESENLGNADDSVQEAESRGEPLDDDKGTSR- NARKRNHAQSSQITSEHDVDDSEAAQSGSVVVGOPRRR- QKVDPAEQTPVPTRYNLRPKTGAPAAAVSEPNKEKEEVS- EGVGALEDEIVNSKAAPPNSVGVFSDNGRSSQLVRCGA- VDNNDASKQFVENMAMTMSEEVNGTPEGAGDYGDADE- FRSESPGEDASGFDGSDDECEHPGEASIGKKIWTFFT
Citrus clementina 2	Embryo- phyta	Uniprot tr V4TH87	MFTPQRRIPATKLTPRGTEAQRSGAISNARNIKGKAVAFETPSVPPPPVNSLL- DYNSSGATVFPFAESEDWRRFRAGLLDEATMERKDRREALMEKYSKLEKLE- DYQYNNMGLLIEKKEWTSKIEELRQSFEEQELKREQSAHLIAFSAEKREYD- RALSMKQCVADLEKALRDMGEEHAQTKLFEKTLTDANTLLGGIEKSL- EVEEKFAAAEAKLAEVNRKSSLEMLKLELESRESVIRERLSLVTREKA- HEAAFYKQREDLREVEKLLQIGDERLSELRRTLNQREVKANENERILKQK- ERDLEELKIDLSSSKLKEREDEINSRLAELVVKERADCLRSTVEMKEKR- LITIEELNARERVEIQKLLDDQRAILDQKQFELEELKRSKIEEEMRSK- ISALDQQEFESHREEKLERREQALDKKSDRVKEKENDLAKRKSIVKER- KFAVAAEKLELEKQKLIADKESLQILVKEIDKIESENAQQELQIQEECQK- LKEEKSSELLRLQSQKQOQIETRYHQEQLLKEHEDLQDREKFEKEWEV- LDEKREINKEQEKIADKLEKLEKQHSAAEERLKEECAMRDYVQREIEAI- RDKAEAFATMRHEQLVSEKAKNDRKMLEEFEMQRMNQEAELLNRR- DKMKELEQERTRTFEEKRERLVNDIAHLKEVAEIGEIEKISERDQLEKE- HEVKVNRKLEQEQQLGMRKIDELDILCRRLYGDREQFKREKERFLEFV- EKHTSCKNGGEMMRAFVISNLQLPDDEARNDIPLQVAERCLGNLQGDV- AAPYDSNISNSHGMNLGRADSGGRMSWLRKCTSKIFISPIKSEHIST- SMLEEEEPQSAVPTIMQEKAEAGPVLVSKAIGYSSPEDEPQSSFRVND- STNREVDDEYAPSVDGHSYMDSKVEDVAEDSQQSELRSKRRPGRKR- KSGVNRTRSLKAAVEDAKLFLGESPEGAGLNASFQAHESSQGISSHTQE- ASNMAKRRRRPQTSKTTQSEKDGAGSEGYSDSVTAGGRRRKRRTVA- TVSQTPGERRYNLNRHKTSSAVLAEASADLSKANKTVAEVTNPVEVNS- PKSASTFPFAVLNENRKSTHLAQVTSVKSMELSQDRAVRFKSTNIVDEN- ADAPKSIENTVLSSEVNGTSEYVDEDENGGRVLEDEEDDDSDHPGEASIGK- LWFFTS
Citrus clementina 3	Embryo- phyta	Uniprot tr V4SQK6	MASPPSSGRLSITPSSRVLSQPLSDESIVKRLKEAGLDEESIKRRDKAALIAYIAK- LETEIFEHQHHMGLLILEKELASKYEQIKASAEAAELLQKHDQASHLSAIAEAR- KREESLKTGLVEKECIASLEKAVHEIRAESAETKVAADSKFAEARCMVE- NAQKFAEAAEAKLHASELQAEANRYHRSERKLDQVAREDDLRRRIASFAD- CEEKEREIIRERQSLSDRKILQQEHERLLDAQTLNREHDHILSKLQELSRKE- LEASRANVEEKFKALNEEKSNLDTLVLLKREEAVIEREASLQKKEQKLLVQET- LASKESNEIQIIANHESALRVKQSEFAELAIKYKLAEDIEIKRRRAWELRDL- QREESLLEREHLEVQSRALVDKEDLVERSHELLKENKLIAFEKADLKKLLQ- KEKEEVNIIKSDLQKLSLSELEKQVNCADKLEAMKSEAGELSVLEIKLKEELD- VVRAQKLELMVEITDKLQLEKAKFAEWEIMIDEKREELRKAESVAVERVVSKL- KDERDSLQERDAMRDQHKRDVDSLNREREEMNKMVHEHSEWFTKIQQERA- DFLLGIEMQKRDLENCIEKREELSSFREREKAFEEEMKRELQQISSLKEKAEK- LEQVTEIKRDLERMEINMDRQRDRDREWALNNSIEELKVQRQKLEEQQLLH- ADREIEIQAESERLKKLEDLKIADVMAVSEMQRSRLEHSQKISAKRHLNQQTSL- AHADLGSQKFDVTNNGDRFNTPSVQKTASASPPSLARFVWIKRFADLVFKHSG- ENSIENDEEKSPSDHEDASLTINSRKRQPVRYSFGEKPVILEVPSENEVVKRTVD- LESENNAQAQKCKQSVSEGGIHAARKRRVDVDCVPSSELLMQNNKRRKQED- FPRNSSEEAINHGAQAEQSNLPEDQHTLTSKNKSNVPEGLHTLTSNNHTGGNE- EASLIVDKIIEVTECMPDADNFINQEKIDGSONSVAESVQDIVKVGGTNDHST- PAHTDDVLPYVSEIDGMVQEKQMGVNDLTECGQAQNEIGEHLCELV- QSDNSKKNKELIAYRTRSKQK

Phaseolus vulgaris 1	Embryo-phyta	Uniprot tr V7B0B4	MFTPQKVVWSGLTPNKSQVGRGGTGSGLGPNSSGDGVSASKEQGVAVVWENG- GNNLDRGVLVERVSNLEKELYEQFNMGLLIEKKEWTSKYTEQSQDLVEVK- DALEREKAAHLIALSEAEKREENLRKALGVEKECVLDEKALREIRSENAKIKF- TAESKLAEANALVASVEEKSLEVEAKLRSADAKFAEISRSKSEFDRKQSDLESQ- SSLRDRDLRFIAEQEAHESTLSKQREDLWEWEKLGQEEERLAKGQRINIERE- QRANENDKLCRQKEKDLEEAQKIDATNITLRSKEDDNNRLADIALKEKEYDSL- GINLDLKEKELSAWEKLNKAEKVMQKLLDEHNAVLDVKKQEFELNEKRSK- FEDGLKDLVELEKKEAEINHMEEKVQKREQALEKKAELKEKEKEYEQVKAL- KEKEKSISSEERSLETTKKKIESEREELVTDKAEVEKIRSNNEQELLRINEIEERL- VTEERSEYLRQLSQKHEVDQYRHKELLVKESDLRQKQESFERWEDELDL- KRADVEKELKSVIQQKEEILKQQFEEELKNEKQAAQDHIKRETLALAKESF- AAEMELEKSSLAEKASQORNQMLDFELQKKELEADMNQNLQEQKEDLIERKN- LFEEKRESENLNINFLREVANREMDKMLQRSLKEKEKQETDENKHHLESQRM- EMQEDIDLVDLNRKLNQREQFIVERQRFIEFVEKLRSCQNCGEIIEFVLDL- QSSDDIENLEVPSPKLAGDIIIGDSIENLASSRKNIGASPATDQKSPVSAQTISW- LRKCTSKIFKISPIKFESEDSGLTRDVMNLSVEKTNMDSRHENEAEELSFVAVND- SLDGRRRASGNDITEVEAVDQDPPSVENQSNIDSKTPEESKAEQQKSRGGGRT- RIKRTHTVKAVLKEARGILGEEAELLPGESVDNHETEFPNGNAEDSANVENSESQ- GLSNRRIPMNVKRNRVQTSQMTVSEHDGEASEGHSQSVIPGQRKRKRQKAA- AAPAQTAGETRYNLRPRRTGATTSSARATSAGGKESQGEVHRVKDTEEEIVDSK- ISHSLVGITNEDGGSVHLEQSMKGVETRDGYGGDTTGFANNITLSEEVNGTA- DDAEENDAEYRSESHGEDAGVEIDDEDEDYQHPGEASIGKLLWNFFTT
Phaseolus vulgaris 2	Embryo-phyta	Uniprot tr V7C8S5	MFTPQQKAWPNAVPFPHRGGAAATVSASAKGKAVADGPPPPP- PLGSLTETTVAVGFDTGNAEDWKRFTLGLLDESVMQRKDHEALMEKVL- LERELFDYQYNNMGLLIEKKEWNSKFDQLRQELAEETEEILKRE- QSAHLIALFEVEKREENLRKALSTERQCGADLERALRAMQEEHAQIQSKSHTKL- AASALVDGIEEKSSVVDKLLDAEAKLAEVNRKNAELGMKQVEVARESSLQK- ERLSLVDRELFDATFYKEREDLKEWERKQQRENMLCNGRQNGEKENIVK- TEKNLKQKERDLEVEKINSNSILKEEAEIIRRTADLNMEKVKVDSLKSMLEK- KEKELFALESKLSRREREGIQKLLGEQKATLDLQKQVEFEMEHRKRSVVEEFS- SKEEALQREVEVNHREKQVEKEEQALSKKAERLKEQSKIEAKLKSLEKETE- MKIKELEKDNQQLADRESLENLNAELQKIKAEISLQELQICEETKNKLTITEDD- RLENSRLQLELQKQENTRLQKDSLVEAESELEREQRFKEWEVLDERRREIT- RKQHDIDEEKESLRKLNSEEEERLRSKQNMQEHKKELEKLEKESFRDMSM- NQEKHLLSEKVKNEQDKMLQDFESKTRNLENIQKQEEIEKDLQERERNFQE- EMRKLDNINILKDVTEKEWEAKAEKIRLENERKELELNKQKQKSGQEMHE- DSEMLMNLQKVKKERQLVAERKHFLELVENLKSRCVCGEVGDFVSDIELP- DFKESMAIPSPISPVLYKSPKNSQDIVASSDINNSSGSRVPSVWIRKCTSKIFL- PNKRAEAVSALDTAGTSLPSDVNVSVKADPAEPLNIEGARVILDEROPASGR- AYHSLDTPLLQSENIDKELDDEYSQSVGDHRSRVDLVDGDRDSDHQSPKLR- GRPKKSKSGIARTSRVAVVEAREFLGKTPKKNENASLQSLTDDHIKEDSRE- DSSHVEKAVGNTGRKRQRAQTSRVTESEQNAGDSEGGSESITAGGRKRKQA- LAPPAQVTSEKRYNLRQHKIAGKDSSTDLNPNATKSVVKEAAGGNKLGEMSP- EVTETSLAAADNADKSMVQDSTTKTVEVSDERVVRFVPRDIVDNGAAT- DSLNPAAEENGTEPHQNGENGTIHDFFEDDDDDDEEEDGDEEHPGEV- SIGKIFRFFTT
Phaseolus vulgaris 3	Embryo-phyta	Uniprot tr V7CFW3	MELSTPNSSKPLSITPGSRVLKSPILDEQIWKRLRDAGFDEESIKHKDKAALIAYIAK- LEAEIYDQHMHMGLLIMEKDLASKYEQLAALAESELMHKKHDSAMNKSALAES- RRREESLKKTVSVKDACIASLEKALHELRTESAETKVAAESK- FAEAHLQIDEAQKITEAEAKVRAAESLQTE- ANRYHNAAERKLRDVEAREDNLRKIMSFK- ADCDEKDKEMIFERQSLSERQKGLQEEQER- LLQSQSLLNQREEHFLSRSQELNRLQKELED- TKAKVEKEHETLHDEKTTLMKMEATLMQREE- ELAKWKTLSKKEQELLEFAKLSIRESDET- KKVIAGQEAALTKKYNLEVELQMQRKWVEN- DIETKRRAWELKEVDLKHCKDEILEKQHELEA- LSRSLSEKEDLKDLSALEEKDQKLSAAEKE- FELNKVLLQKEKDTIEQAKQDLQKSLASLENK- RRQVDIDKERFEAVKNETGDLSEVLKKEEID- LVRSQKFELLAADKLAEKAKFEAEWELLDE- KKEELQKEAEFIKEREAVSTFIKNERDQLK- EKENLRYQYTQDLGLASERESFMNKMA- QEHAELFGKMQQERADFLREIEMQKQELN- LIEKRREEVESYLKEREKAFEEEEKNTLHYI- NARKEKAKELDQVSELMKRLQTERAEINL- DRERRNREWAELTNCIEELEVQRDKLQKQRE- LLHADRVIFAQTEELKLEDLKAUSDNNAIT- EMLKSDMESNRKISSRKNLKRQTLTQGGD- KISNGFDPFVERSSAGSPSPVRFVSWIKRC- SELIFRNSPVASDADTGSNSQKHLNDKPLG- IGKGGQMGFSFEESKIVVEVPSRDDARRRE- IESEAKNVNGKSALLFPDGHLAGRRRGR- GNVTSKVGDPVLDLQGNKKSRAEGQTTENP- IDQGTTRRVVSTQSDVLKVVQVLTSSNQ- GNTTETRVVMVVKVHVSEVTSEKVDALPID- SQEPGDNPNPALAEDHYGETIDQINSKTKREDILPRVSRVLGSTEISK- GNNGQDSENC
Prunus persica 1	Embryo-phyta	Uniprot tr M5WQZ9	MFTPQRWSGWSLTPKTAETGTGSGSNMKSQTPNFNSGDGV- VAKGKGLSLFEPRTPASGSVLENGGNMQVESGEGATDRELAQRVSE- LENELFEYQYNNMGLLIEKKEWTSRHEELRQSLTEAKDAVRRQEAHLIAI- SEIEKREENLRKALGVEKQCVHDLEKALHEIRSENAIEKFTADSKLAEANALVASIEE- KSLLEAKSRAADAKLAEVSRKSSEFERKSKDLEDRESALRRDRLSFSNSEQEAHE- NSLSKRREDLLEWERKLEGEERLAKGQRILNQREERANENDRIFKQEKDLEDA- QKKIDATNETLKRKEDDISSRLANLTKKASSEYDTRMNLMEKKEKLLALEEKNL- RERVELQKIDIEHNAILDAKKCEFELEIDQKRKSLDDELNRNLDVVEKKESEINMHEE- KVAKREQALEKKEKVEKEDKDFESKMKSLKEKESIKSEEKDLSEKQLIADKE- DLVRLLAEVEKIRANNEEQKISSEKDRILVSEEEKSEYHRLQSELKQOIDKYM- QKELLLKEAEDLQKQKELFEREWEELDDKRAEIEKELKNVNEQKEVEEKWKHVE- ERLKSEKVMQAQDHIQREQDDLKLAESFEAHMEHEKSVLDEKAQSRMSQMLHELE- TRKRELEIDMQRNLEEMKPLREREKSFEEERELDNVNYLREVARREMEIEK- ERLKIEKEREADANKEHLERQHEIRKIDIDLDLQKLRDQREQFIERESFISFIE- FKKCTNCGEMISEFVLSNLRPLAEIENAEVIPPRLGDDYKLGGFENLQARQNN- ISLGDSSRSPVSGGTISWLRKCTSKIFNLSPGKIEFGSPQNLANEAPFSGEQNVA- SKRGCGEIENAEELSFVASDSFDVQRVQSDNRIREVAVQYPSDDEHNSNMSE- DLPEDSQPSDLKGGCQKPSRRGGRRGRPAVKRTRSVKAVVDAKALGEAFETND- SEYANGTAEDSDMHMESHGGSSALDKRSARNGRKRGAQTSQIAVSGGDDEE- RSDSVMGAQRKRREKVIPEAQAPGESRYNLRPKTGVTVAASASRDLVDKNEE- EVDNARATEHYSKAAPATSIGVSENGGSTHFRVRCGLDGTQDGEADAINKLENT- AVSEEVNGSTEGGQEVVDGDEYRSESQNGTPIEEDDDDEEHEHPGEASIGK- KLWTFFTT

<p>Prunus persica 2</p>	<p>Embryo-phyta</p>	<p>Uniprot tr M5Y1X5</p>	<p>MMFTPQRKALNAQSLTPRSGAVVSNPRTAGKGVAFVDPGPPPLGSLSESQPK-TIPDFDTGDMDDWRRFKEVGLLNEAAMERKDRQALADKVKLQKELYDYQYN-MGLLIEKKEWALKHEELGEALAEATQEIILKREQSAHLISISEVEKREENLRKVL-VAEKQCVALEKALREMHEEHAQIKLKSEAKLADANSLVVGIEE-KSLETDAKFLAAEANIAEVNRKSTELMRLQEVARESVLRRRE-HLSLSAEREAHKTKFYKQREDLQEWERKLEGEERLCKLRRL-NEKEKANENDLIMKQKEKELDEVQKIELSNTILKEKKADVNR-LADLVSKKEADSVGKIWELKEKELHELEKLSRENAEIQ-VLDKQALCNTKMQEFELEMEERRKSLDKELSGKVEVVEQKE-LKINHREELKQEQALHEKSERLKEKNKELETKSKNLKENEK-TIKVNEEMLEVERQVVLADLESFQNLKEEQIKKDENVQLELQIR-EEREKLVITQEERSEHLRLQSELQEQEIKTYRLQNELLSKEAED-LKQQRKEFEWENLDERKAEISRGLEKIVVEEKELEKLGQT-EEERLKEEKHAMQDYIKRELDNLNLEKESFAAKMRNEQFAIA-EKAQFQHSQMVQDFESQKRELEVDMQNRQOQEMEKHLQEME-RAFEKEDREYTNINFLKEVAEKKSEELRSEKYRMEKEREEL-ALNKKQVEVNLQEMRKDIDLQAMLKSKIKHQREQLIEERGRFL-AFVEKIKSCKDCGEMTREFVLSDLQVPGMYHIEAVSLPRLSD-EFLKNSQADLSAPDLEYPESGWGTSLLRCKCKMSVSKVPIKMM-EHITDAVSTELPPLSTMVNEGARGHIGHEDEPEPSFRMPNDA-ISQPLPNDTTEKVDGAPSIDHSDSFKVVDVDDSEQSE-LKSYQCKPGRGRKSRLSRTRTVKATVEEAKIFLRTLEEPSNA-SMLPNDSSNIHEESRGSDFVEKANTSIGRKRRAQSSRITE-SEQDDCSEGRSGSVTTAGRRRKRQSIASSVQAPGEQRYN-LRHRKTAGSVTAAAPAAADLKRKRKEEAGGGAEFPNPEVSS-LGMAGETGQTAQLMQVTTSKSVEFSQERVVRFPEDIVDGN-ADAATAKVENTELSGEDNGTPESGSGNNTVGESDDDYDDEERPGEASIR-KKIWNFLT</p>
<p>Prunus persica 3</p>	<p>Embryo-phyta</p>	<p>Uniprot tr M5WL04</p>	<p>MASQSELFARTPGSGRALSIPTGARILQSPFSDEAIWKRLKEAGFDEESIKRRD-KAALAIYAKLEAEIFDQHMHGMLLIMERKELASKYEEVKASNET-TELLHKRDQAAYVLSALAEARKREELKVVGVKKECISSEKSMHEMRAESAETK-VAAESKLAEARNMVGEAQKFTAEAKLHVAESLQAEASRFHRVAER-KMQEVEAREDALRRNILSFKTDCDTKEKESLERSQSLCERQKTLQQEQ-DRLLDAQALLNOREDFIGRSQELNRLKELEDEVKANIEKERRALDDGK-LNLELTEASLVNREEALTRREALLNKKEQELVLEKLVSKESDEIRKAL-ASHEVELRKKKFEFDSELDVKRKLFEDEIAKRRAWELREVDL NORD-DLLQEREHDLVQLRTLVDREKDVAEVMSNLVDEKELTRDAEKEFEL-NVLLQREKEEIIKMKVQLQCSLSDLEDKRQLDCAREKFEVLTETSEL-SDLEMKLKEEIDLVAQKQELMAEADKAVEKAKFESEWELIDEKREE-LQKAEHVAEERLAFSKFIKDEHDLNROKQEMRDOHQKRDVELLVSE-REDFMNMVHERSEWFGMKQKERADFLLEIEMRKRELENCIDKKHEE-LECSLKEKIEAFQEKKNFQINSLKEEAKEERQVALERKRLETERI-ENLDRERRDRWAELNNSIEELRVQREKLKEQRELLHADREILGOIQ-HLKELESKLAALDSAVSEMQSDLVPRSRKTSRRYLKQLTSVREAD-HNSHNEENVANISNSSIMLKSFGSPSSARFSLKRCRELLFKQSP-KHOTEYEENHVISREETSMTVTEQVDTSSKYDGHRYTGNNGSPRFFS-KRONAFGEKPVIVEVFPVGETVKGHTHESEIKFEFDGESCSPLISEHVC-QGGRKRVDKSLNDGFDPLLEPRONLKKRRQQQDATVNSSEHANT-HCIVSTQEKVLEDQNSMPLPSDQICEGAEESGALIVDKIIVSEVIFEE-TGTGSLGNEGLKLEAQNISVEAHHGQNGVFGQAVGQVTEHCQ</p>
<p>Eutrema salsugineum 1</p>	<p>Embryo-phyta</p>	<p>Uniprot tr V4KN99</p>	<p>MSTPLKVVQRWSTPTKATNPDSNGKGPANMVTVPVGRVSEIYYDDPRILPEKVS-LEKELFFEYQHNLGILLIEQKQVSSKYEEELQEFEEVNECLKRERNAHLIAVAD-VEKREEGLRKALGIEKQCAVDLEKALRELRSENAIEIKFTADSKLLEANALVRSVEEK-SLEVEAKLRAVDARLAEVSRKSSSEVERKSKEVEARESSIQRRFYSIAERDADEATL-SKQREDLREWERKLEGEERVAKSMIVKQREDRANEQDKIIOKQKELLEEAQK-IDAAANLAKKEDDISSRIKALAFREQETEVLKKSIEKTERELLALQEKLDAREKVAV-QQLIDEHQAKLEAAQREFELEMEOKRSIDDSLRSKVVVEVEKREAEVWKHEEKVVA-KREQALDRKLEKHEKEKELRLKGVSSREKALKSEEKALETEKRLKLEDKDIILN-LKAEVEMKMTENEVQLSEIHKKEGLRVTEERSEYLRQLTELEKQLEKCRSQEEL-LLKEVEDLKAQRECFEKEWELDERKAEIESELKNITDQEKLERHSHLEDERLKK-EKQAANDNMKRELELEVAKASFAETMEYERSVISKKAESKQLLHDIEMLKRNL-EADMQTKLEEREKELQAKEKLFEEEREKELSNINYLRDVARRMADMQRERQRIE-KELEVDASKKHLEEQQTEIRKDVDDLVALTCKLKEQREQFISERNRFLSSMESNR-NCNPGCELLQEVLPDIDNLEMTNLKSLTNILENAPRQEMRDISPTAAGLGLPVP-GETVSWLRKCTSKILKLSPIKMAETSATRNLAGQEPQSTEQANVNSGPTMLQAQS-VSDTREVEVNNADSDGQNSINSAQVEANSLSTLNADQSSRIRGKARARVRRIT-HSVKAVVEDAKAIYGKSIENEPEDSTENVEDSSKANDGNTGEPDHSKGASKNG-RKRGRVGLRQCTTEQDGTESDGKSDSVTGGGERQVRKRQKVTSEQQEVGGQR-YNLRPRRGGAGKTALGKKNETVTVQQEEGIYSAQTIATASVGVAVSDNGASANV-VQSETMADSEDTAGSPKRTCESAAMSEEDVNKTPQRAHSGNEVYDGEDE-SESEHPGKQSIGKLLWTLTT</p>
<p>Eutrema salsugineum 2</p>	<p>Embryo-phyta</p>	<p>Uniprot tr V4MU99</p>	<p>MFTPQRKPVISPAVTPRSETRKIGGVSNPNRDDRKGAIAISEDVISTLPPP-PIGTLTGEVYRQAEEMDMGDWRRRFREVGLLDEASMERKDRALLEK-ISTLEELYQYHNMGLLLENKEMWAKHEELNQAQFAEQEILKRE-QSSHYALTVEQREENLRKALGLEKQCVEEELKALREIQEEN-NKIRLTSEAKLAENALVASVTGRSSDVENKIYSAESKLAET-RKSSLEMLRLKEVETRESVLQOERLFAKERESYEGIFHKQ-REYLHEWEKLLQEKESIQKRSNLQREEKVTKEKLNKL-KAKQLEEWDRKVELSVSKETEDDMNKRLOELAAEKESCT-LQSMVAKESLRALEELIVREGTEIQKLIDDOKEALAAKM-LEFELECEERRKSLDRELQKKEIEVERORVEINHSEKLGQRN-EALNKKFDRVNEKEIELEARVKTIKEKIMOAEEKLSLDKQ-QLLSDKENLKDLOQELNIRSEMMRKEEMIQEEHKSLEIKKEE-REEYLRQLSELKSQIEKSRLEHEFLSKEVENLKQEKKEFEKE-WEILDEKQAEYKERMIISEEQAKFORFOLLEGERLKNENAL-RAQIQELDDIRLQRESLEANMDHERSALHEKAKLEHSHKVLLED-IEMMRRNIEIELQKRKEQDEKDRQDRLOFEDKRMKELSD-LNHQKQALNREMEEMVSKRSALQKESSEIAKHMKLKEQQV-EMQNDISELGTLSNNLKKRREEFARERARFLAFVOKLDCES-CGQLANEFALSDLQPYNEEATLPPNGVLDLPSSDASDS-CNIKSLDGDAPASGGSGRPTMSILQKCTSLFSPSKRAEHG-MDTGKPEHLSSVAVSKEIKVEKPLVDLRPRPSSSIPED-EEYTVSRVQETSEGSQLEFQSAKRGGRGRPRKPKPALN-PSSSVKHASPESSEKDEAGGHVSVTSEKTTGRGGRKRQHI-EDTTTGGRRKROQTVAVLPQTPGQRRYNNLRNRVTDQAP-ADDEDNAAGGEYDADIAALAFSKDNVEETSSESVESLRARR-LESSEVVRVETVETVTDADVAANNVGVSVANEELAFNI-ARSPSVEDEQRQRTVDEDKNEEYEDGDEEVHDDQDD-DDEGGDDDDDDGDDGLKAGEGSIKKLWTFFTT</p>

Eutrema salsugi- neum 3	Embryo- phyta	Uniprot tr V4KM52	MHVDTHLYDYKSIMHWTSTNEELQQAIDEASEILKRERMSKLIALNEAEKREENL- RKALISEKQFAAEERDLKYLQEQHSEVSTSEAKLAENALVMGIKENSLEV- DRKRAIAEELKSVINRKSSELEKLVKKEVE TREKVLQREHLSLVTEREA- HEAKKLTLEEDRLSEAKRSVNHIEERIESEKTIKKKKELKEEEFINSMLNDISM- KEKAFEMKTNIDMKELHELEEKLVREQMETGKLFDDQNAVLDLRSRHEFEM- ELEQMRLSNEELERKKSEVQLVEVISHKGEKLAKKESTLEKMEIVKEKEDL- EARQEVVKEKALKAEEKLMHMKRLLLEDKESLRQLKDEIEEIGAKTTQKESR- IRREEHSLRITKEERLFKRLQSELKQIDRVEQEEELLLKEREELKTAGRLEK- EWALDEKRADKTREQKVEVTEENLRSQTSSEKHLKREDIILRDNKREVD- DEVMQRESFEAGIETKFLSHDNANIQTREMEVEVQYKELALKRERIEISVRRK- TLYKQSVVYMGVDDSLRISLKEKREKQICCAKRFALFLKENLKNCSGKGFHK- FVQSNRAPDIETMKSEMKLDPKNQTTDPVENGDKLSDNSKASLIGTLAAIK- LPESWQHVDLTDLDTLDTLDTVAGNDHEASGTQSFIAIKSRRRGRGRSK- SVRGRQATKAASRDKTSDEEIVKVEAETFKNDRGRKRPVQDPQFEAGSS- GEKKEDDGNISMIIEENKGEEEEEEETERPDEASIGKKIWAFLLI
Eutrema salsugi- neum 4	Embryo- phyta	Uniprot tr V4KV78	MATSRSERFPITNTASNRLTITPGRSRLKSPLEEVMMWRKLKAGFDEQSIKNRD- KAALIAIAKLESEVYDQHNMGILLILEKDELLSKYEEVKASVNEAD- LAHRDQASYSALAEAKKREEDLKDVDGIAKCEISSLEKTLHEMRAE- CAETKVSAGSKMSEAHLMIEDLKKYADAEAKMRAAEALQ- AEANRYHRIARLKEVESREDDLARRLASFKSDSETREN- EIDIEROTLSERRKSLQQEHERLDAQASLNQREDHIFGRS- QELAELEKGLSEAKTTFEEERRALEDKISNLIALASLAKR- EEAVSERESSVLKKEQELLVAEEKIATKESLIQVLANQEV- ILRKRKSDVEAELESCKLVEDEIESKRRAWELREVDIQR- REDLVGEKHELDVQSRAEKEKIDITERSYNLDEKEKNLN- AREKDINLKTLLLENEKERLQQLDLQQLMSLLEEKRRKV- DCATRKLKALKSETSDLFLEMNLLKELDDLRAHKLLELA- EADRLKVEKAKFEAEWEHIDVREELRKEAEYITRQAEAF- SMYLLKEERDNIEREERDALRNQHKNDVFNREEFMN- KMVEEHSEWLKIQRERADFLGIEMQKRELEYCIETKRE- ELENSSRDREKVFQEKLEERIQSLKESSEKLEHVVQV- ELKRLDAERLEIKLDRERREREWAELKDSVEELKQVREKL- ETQRHMLRAEERIEIRVEELKKLENLKVTLDDMSMAKM- QLSNLERSWEKVSALKQKVTRDDELVFQNGVSTVSN- DDGYNFMRQNGSTPSSGTPFVWIKRCTNLIFKASPEK- SPPMDPHQGGPLLENLKLDSRREERAYTEGLSIAVER- LEAGRRGRGNTGRDTPSSNKKRKHDDVTQKPKPSPDE- TDPHVSISPPQNPEDKHELPSQQTTPSGMVISSETVKIT- KVTCEVINKVNTNIDCSENSEAGTMVEEQHDSGCNETVNVSETVTRKEAE- SDNRKEQSDDGGVVA
NUP-1:			
Paratrypa- nosoma confusum	Trypano- somatida	Non-public short-read genome assembly, kindly shared by Prof Julius Lukes, Ceske Budejovice, Czech Re- public	MITDSDLHRAGMFFTRARATGDGDPQRASQGAHGHGAAGSSHGPPRRGSMR- FVASVEAAPWLAQQDAAGASAAAAAHRDRRPVDRDRSGSEHASQVAAAASET- PLAHSOPLRGPFPMYQALAGLSQTTTPGQRDPRTDAGESSHDRHLLPS- GRSTAASAGLDTGSPGVGATRTGHAAGRASFFHNFSEFRPSALEMDELSDELRS- YASRLLEDNVGIASDNLQAQFSLRDSYKDEAARLRAELDHQRLLKRYRDRDIERY- KESSSQORDIAFLKOKLDEANGTVRLLQGQLDAAHDLPSPKYYKEEIRLDRASLE- HATNTISELQEAENHRLRTNLARVETLELRAAQVELRESVQRNHELEAALHSSQ- ROITAKDGDITRLKHEMDRLRAVDHENTOLRSQIQQVQDKHSSAQOYEQILLDT- LQHRSTLTVVEEARLQHVADITKELQRAHTTLKDKCHNVDLALAMNLISQOJQISL- LEAKQDVPNRLNPHEASADDNAHRSHTIHNLMEYVTSLERQIYDQAKLLE- NYENPLQTKDEKPESSNRHKQNKPPPTTTTCSQHSDEEAADGPTDAYPNRT- SNTTISNAQDTDQTLPALSPSEQTLHKKPSPDENHSELANCLKTIQOMQRELVG- MEQRLREL DATVAQLERAAEPEPTVVAALATELRHAQEQLQEACARAAAAAGA- RDDALRTVDELLAQRRDRSATPRPSKGGSGEGDAVTPADLAAATARATAEA- EAEAAARKEADALAADRLATXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX- XXEAEARKEK- VLSTDGCEKSGLFRGFDTAEVGVCNFSCF6PGCRKECFCSGCLRVVEICREVG- LFCNSEVGCSAVNDAVAASRFVAMRTALVVDVDFGDSRMGDDGTLRDIRR- ADESSCTEVSTYVVGVAASLQATDALAARVEELSSALETEREVSQRLRGIDE- MIDGLAEATSDADALREIGDELAGLRESLLMRDRDGGSLHESDEDARRVALQ- DALDERSVQVSLTRALESLAVEYELMKAECAQKDSRASDLERKVERIIRDAAA- EALGAQASSVVMLMKAALRNTGLTAEAAAIVRSVTRDSLVGSSDELRLLSGVIT- RFDSDVHGTGLTAEIARDLVLQIEQLAADRAYHEECVQSLREVEAAKGAALQALE- LQOASDEKVTVEQVAHERSEKILTTLSTLQSEFQRMKERLDPRSPRPSAV- MMDLVEAKRAVSRANERGRRLAERLLDDREGELHLLRQSLAEVTEAKVRIEAELR- AVNYSATHLRETAELRPPAPRGAAPAAASSQAADDPAGPAPDPTLASRFS- PFNSTSDAVAASAAAVWQOQQRAMNVLEQSHVQLSTLHAEHLTROEAASLR- QRLKQVAAQVREKHSLLTEANRLREEVKTNDRVQEEVLVQMCVYEAQAAR- LSDASSEALKQLRLKYDRVCSL SAARDSLARLEREAREIDARSQAEEFAK- TFRNVYEQALHALREMDAQIQSAPAPPSGMRPPRASALSPDARLNRVTD- LQELASRDEEVSQLQLLALREKIEAQLEADLKRAVRALETYSDGEQEAALRLA- ADRLMRLQAEEVKLADILAAASAGAAGGPRFTRHRLRSSLAVRAAGOPDASG- DGALPPAPQKPERSRADQAEAPPRDDDGAGKALTAGSGPQGAARRSRPRA- SASSVTSPTRRGRHPSKALGARRSGVAKRPREV
Trypano- soma carassii	Trypano- somatida	Non-public transcrip- tome as- sembly, kindly shared by Dr Steve Kelly, Ox- ford, UK	MFTTGDARRYSSFFRSRWTPPPDGVNNGRGLHSATPPPIPRLGAPIHVRGLVGT- DPMRYSGPVTLASAGLANRALVGMRSRDEPNMTDSPYVSPYVSTLDA- GRGGRIIPDDVLYNARLEEDVGHVNNLDRAYHLRDHYKEEAARLHKELM- RKQMLDLSLQCDHQHCGGMILSRQEIDSLRQLLDEAEGEVQRLKIALASDGHQ- GLSVMYCKSELVVRRETEIENLVQVENEKLDKLDLQELQEDIFVKRVSDTSDAKSNVD- ETQQSEITIKALCDELRLDKSQIENERQEYEGRINSLTVERNALSAEVTSLNHLHS- TRTLQCTHADTVDELRESFIQRSSVTQEQYVVLQRRYKEIQEELSQRHSQDNTQX- XXSNLESTLKENQNLQIQIILRQSEMEKEELLHRLQKLDKQSLSQNLHQNLDL- LHKTHALQKTHLLEDTEKLIQSQKDNEEANRVLTDFFRRRLLEDSCSTVNLAAARGSEQ- RGYKSTISHENIEPTIRPPPFDISKAKSPNTNESCPTQDIDISTNTQREYKTSQNIHID- DEVHTLRSRLAEADALRSRLAEADDTIEKMSAERVAQSAKVDELTAAMVSRQLQGSSE- TTTSDMIGALTAEIRDTQDRLEHEAAEACRTAADLHAADAIVRITELTEQLSRDERNAR- ETQSPLQLQLSSALSAVDCAAEREAASKAADLEDRVGALEDELAARSRADGIID- EADALRSRLXXXNNLILLLEQYANIPEDVIGALAEALRDTQDRLVAEGERATGLA- SNRRVMTGSVERNDRGSVLEKDLRYDVSSGCVGDDLSLRSQALATALAAIIRLGED- YDTMCEKYRSSEGQVSLKMDLLAERRSLAEVNEISVHREEIDMKSRALAAALVA- LDRVAEYADAQVQEEERSANLLSTFEALYAQFEATKRSLEEQRTRNAEASAAELI- LRRNVQAEARARRRAEDHLEAECELELEREHEAHLRQREHSTLERMSRRGDR- SEPMESESTAVDESNLVAPVQAEYVDHTQFLQISSLQADMLLTRRRRCRLEGIREE- MQSELERGTVGGNDTVSARAELRRLRGQFELRVEHRVLDQDQHNTMLRDH- AALDKLLEQNTRLTHELGENQNALASLAKHNHSSQPHPLNAPTREHSEALDRAF- RQVLDQMHVLRREEVRHSTRVSPRRSASAPRRSAPPDNRVAALQVAVRRGEE- EIQRVQEELEERREEVDALLEELVATERQNSSKRSPLHYEIEVQLQANEGPQE- SLRQVSVQRLTHVQSASVSRIRDRTRISGREDSQGNIAATSSRPTEYRKRSSRVSR

<p>Trypano- soma vivax</p>	<p>Trypano- somatida</p>	<p>NCBI com- pound of: CCC46700 and CCD21021</p>	<p>MFTAGDARRYPGFFTRTWTPPEGVSGVGRNSTSDSGIRYATPPLLTPLRALPHI- HVRGLTSSSPVPLNSPPSASSQGLLTSRALVGMSSREPCISPNVSTLNAAGDTESF- STRQCRRRAVTEMSREEQSNYMRLEEDIRQMSNALHVSQYQVRD- HYKGEAARLHKELMDRQSFDRLEQEHKNCRDVIHQQRNTEQELRERLDAS- QTEVRYLKSIVGSSGAQKATCANQAHEIEKKREEILELQRQNKALVQLGGI- GGQLSKEQTLAFQQQSEINASVAAWERGRRELKDTENLIKVMGAEMNDLREQ- IQKERVQHEESLWAVTQERNLQMLKMSLAEEVDRLRELCOAQSGTIDSLNET- ILKRPALPVNSDHDEEAVQNTIANLNNEGQSLIRDRAVETTSLEKQTLHSMC- ALRQDVEDSNALQRAHIEMAQFKKHITLLEDEISTSKQEIANYEQKLEAQI- LSEAKETSRESESVIRDFKMHEDQNATYELREKIQIFQLESKGTNYSETGAH- DDNKIKQNSSQGYTERQRTDMAHEHMKRQDGSHTIENTEAQKQTEVAQKQL- CAELRETQARLREAEQVARLTAEANAGATAASGSESASAADVVRSTAEYVSV- ASQLSAALAALDKLAEEERAMEHQAAEHARVLEEEQARVWREARSREVD- ADSVQQQLDRALATISEMAAERDAHAAKLAELSDTVSRLEGVETVPEHAVEAL- CVLRETQARLREAEQVARLTAEANAGATAASGSESASAADVVRSTAEYVSV- ASQLSAALAALDKLAEEERAMEHQAAEHARVLEEEQARVWREARSREVD- ADSVQQQLDRALATISEMAAERDAHAAKLAELSDTVSRLEGVETVPEHAVEAL- CAELRETQARLREAEQVARLTAEANAGATAASGSESASAADVVRSTAEYVSV- ASQLSAALAALDKLAEEERAMEHQAAEHARVLEEEQARVWREARSREVD- ADSVQQQLDRALATISEMAAERDAHAAKLAELSDTVSRLEGVETVPEHAVEAL- ANSVQQQLDRALATISEMAAERDAHAAKLAELSDTVSRLEGVETVPEHAVEAL- CVLRETQARLREAEQVARLTAEANAGATAASGSESASAADVVRSTAEYVSV- ASQLSAALAALDKLAEEERAMEHQAAEHARVLEEEQARVWREARSREVD- LSDTVSRLEGVETVPEHAVEALCVLRETQARLREAEQVARLTAEANAGATA- ASGSESASAADVVRSTAEYVSVASQLSAALAALDKLAEEERAMEHQAAEHAR- VLEEEQARVWREARSREVDADSVQQQLDRALATISEMAAERDAHAAKLAEL- LSDTVSRLEGVETVPEHAVEALCAELRETQARLREAEQVARLTAEANAGATA- ASGSESASAADVVRSTAEYVSVASQLSAALAALDKLAEEERAMEHQAAEHAR- LERVQSDKTSYLLSMVDLNEVSDLRQLKNNDDVAVTEILSCLGDVSAECRG- GNSSGDSSALLKAIEDLYSRLSSAKRNFELYSTCNSETGVTELIARRSADRM- NTARRHAXXXXXXMMNQECLFLRRELRLHGICFSYNSLDVEGFSARLSAA- NVNDECSDETQLLQISSLQAEILLCRKRCRRFEVQNEELLTLEHHVSAQEV- ATNMAELRLDLAQLRLQFTQLQEEHSGLQDRFEDVNHSQLDELDRLLQNDR- LDCDLKEKSEKISAMEKMRDAERSAKGRVNELLESCLIMEAQMGDILLKGA- SSVDSSESQQSDRVQGRALAAARVFLQTKRKEEIEIQLQDELLRKEQDLD- QEQRVTDSTQELDNVKKVTHMESVQRLQSAKDSLQEEENRLLRAREVALT- QHPSPSTAHVaelGALNQRPPRTGSAATVNMTRTDNTAVVPTALPSV- SVGRRSTPARTSGLTQARKRPRS</p>
<p>Trypano- soma con- golense</p>	<p>Trypano- somatida</p>	<p>NCBI CC- C89495</p>	<p>MFSAGDARRYPGFFTRTWAPPDPPGDIKNNRSSSIPQTGISHPPLLTPLRALP- PINVRGLAASDPMSRVGSLSSPAGLLPDHALVGLAREEPSFGYMCRC- RGLNDMSHEERLEYTSQLDGDMKHVSSALDRAHQ- RDEYKNEAARLHRELMEKQORSLDRVREYGECKNVLHFRFTENE- LQQLDQSQGEVVKHLSRVISANVQELKSGAGQLLGDKNQER- QLLEENRRLERQLQELTEQLEKEQAYNRRWKNSETATTTDQDLSA- VVEQGRQLSLNVAMAEIDDLKNEQQRKNDYSSRLEQLTNERNSLHQ- GHLLTSETLEKLEDTCKEQOQTIEDTVRLLQEQSTNRKERNSHNEQ- HIGYQHVHSEENTDQLRAKCAIELEKHALQOQLFHLRQEGERKDEI- KKANTDIRLKDSEATMSRKIGEVVWQNEKLEQRIEETTRQLHISQN- QNEEDQYVIKNFHKRLEEQIKLTHNTQTENNNTTENQSDSQSRKQER- QTTEIKKNQELGKKTAEHENPRETEQKHINKENECTATITKEDR- AAVGAPSSDLATQLSSALAALNRLAEEREAVSERAEMEDRLGVLE- DELLKARTALDASSGNEEALKEQLEQALEATESLAAERNAHATKLE- GTIARLEGCGDSSAKVIEALKSELAHTQKRLREAEQDCARLTRD- AVKERDAAVGAPSSDLATQLSSALAALNRLAEEREAVSERAEMEDRL- GVLEDELLKARTALDASSGNEEALKEQLEQALEATESLAAERNAHAT- KLEGTIARLEGCGDSSAKVIEALKSELAHTQKRLREAEQDCARLT- RTDAVVKERDAAVGAPSSDLATQLSSALAALNRLAEEREAVSERA- EMEDRLGVLEDELLKARTALDASSGNEEALKEQLEQALEATESLAAE- RNAHATKLEGTIARLEGCGDSSAKVIEALKSELAHTQKRLREAE- QDCARLTRDVAVKERDAAVGAPSSDLATQLSSALAALNRLAEERE- AVSERAEMEDRLGVLEDELLKARTALDASSGNEEALKEQLEQALEA- TESLAAERNAHATKLEGTIARLEGCGDSSAKVIEALKSELAHTQ- KRLREAEQDCARLTRDVAVKERDAAVGAPSSDLATQLSSALAALN- LEEEHEAVSERAEMEDRLGVLEDELLKACTALDASSGNEEALKEQ- LEQALEATESLAAERNAHATKLEGTIARLEGCGDSSAKVIEALKSE- LAHTQKRLREAEQDCARLTRDVAVKERDAAVGAPSSDLATQLSSA- AALNRLAEEREAVSERAEMEDRLGVLEDELLKARTALDASSGNEE- ALKEQLEQALEATESLAAERNAHATKLEGTIARLEGCGDSSAKVIE- ALKSELAHTQKRLREAEQDCARLTRDVAVKERDAAVGAPSSDLATQ- LSSALAALNRLAEEREAVSERAEMEDRLGVLEDELLKARTALDASS- GNEEALKEQLEQALEATESLAAERNAHATKLEGTIARLEGCGDSS- AKVIEALKSELAHTQKRLREAEQDCARLTRDVAVKERDAAVGAPSS- DLATQLSSALAALNRLAEEREIALEEDAFEGPIVDVLRDILFLKSSLG- ECRRYISGLNDLNLSCDDGSSSMVNLFRVGGGLSDDLPNKCDG- TATLGATRVMLGSDCVDEVSCVADLVAARRSLQRSNDARRRLEERN- IELEERELEKRSMDIAGLEKECRDLEKQVRRGSSYMRSHIDECDFR- SSVSGGVCMETIDLAQFLQISSLQADLMLSRRTCQLESNQEELL- SVEQGVQSONLLESVEELRMEVHNLRRARDELLCERRLLTERVND- LELEREDVVLKQNRWLTQIEEAGDKLSAAEKSRREAEALWEKQ- AEELIKAFNDLNAQMTHLRSEGAPSRNSLSRSLRSRSDVSEDELN- QVRNARISFLETTILRKDGDYIRLQELIKKEDNIDKLRQEIIRMSN- KIKNATRKSEKLESLQLHEKKEEMNKELYLLKTKAPQKIVQPRFSTVT- NQQTPEINTPSKRTSLVIIEEDSKQLHDISTGRRKAQTL- SKEIKQEKIPKRESRKRGRN</p>

Trypano- soma bru- cei	Trypano- somatida	NCBI XP_951625	<p>MFSAGDARRYPGFFTRTWPPPENIGHVRSNRSASSIQGGLTHETPPLLPRLAAP- INVRGLAATDSIPRLNPLPSPGTLLTNSALVGMSSREEPSIMPLLSKQTPMGMHP- TIFAGRRGLADMSEEEERMEYTNRLGDMTHVHNTLSRAYQLRD- DYKNEAARLHRELQDKNHRFDCLLREHSACNDVYRCKRENEELRQKLDESEG- EVRQLRDLKLVSVNSQGGKYVPSGGERHVGQRQESALEEKNKLEEELELTKLE- RERECIRHHAVAAEMGKSENTSHEEEEAQSRYLQVTRTEIDLQQLRKRERED- YEESLREAIQARNNLHQONTALQEKEQLQEMCDEQHRITIEDLTSQLLRKTE- QAVQRGAPDTQMETTENDKTDNTNNDDEVYRMLLQHTLQQQFLLRREG- EAKDILLQKASEEIFNLQNLQQQLEAALQKSREHAALTKLSHTQNLQTAQER- ITEDSYVINNFHHQLREKIQISGSISGEKNIPQGGNKEESIELVRETQMPSRSG- NDSQYITANVQHEKLNQPKADSGHNATGNNKELSSAQNDEYEQAIHKHMT- EGLTEVIEALKTELQHTQKCLREAGEENVQLTNKLNAAAGARGRSTSTRSGSLT- PNDTEGSLRTYNAGLTKQLSSALAALQLAEQHDATLARATEMEERVSTLEEEEL- RTAHSTTKMSAERELHVTKLQLEETVSRLESYGTTPEQTVAAFTTELQHTQQ- RLREAEIIIQLTNKLNAAAGVVRVTSQSDKDGNAARALVSDVAVRNADTDLGTQ- LASALVALERLAEEEREAALKEATEMEERVSTLEEEELRTAHSTTKMSAERELHVT- KLQLEETVSRLESYGTTPEQTVAAFTTELQHTQQRLREAEIIIQLTNKLNAAAG- VVRVTSQSDKDGNAARALVSDVAVRNADTDLGTQLASALVALERLAEEEREAAL- KATEMEERVSTLEEEELRTAHSTTKMSAERELHVTKLQLEETVSRLESYGTT- PEQTVAAFTTELQHTQQRLREAEQEIQLTNKLNAAAGVVRVTSQSDKDGNAARAL- VSDVAVRNADTDLGTQLASALVALERLAEEEREAALKEATEMEERVSTLEEEELRTA- HSTTKMSAERELHVTKLQLEETVSRLESYGTTPEQTVAAFTTELQHTQQRL- EAEEIIIQLTNKLNAAAGVVRVTSQSDKDGNAARALVSDVAVRNADTDLGTQLASA- LVALERLAEEEREAALKEATEMEERVSTLEEEELRTAHSTTKMSAERELHVTKLQ- LKETVSRLESYGTTPEQTVAAFTTELQHTQQRLREAEIIIQLTNKLNAAAGVVR- VTSQSDKDGNAARALVSDVAVRNADTDLGTQLASALVALERLAEEEREAALKEA- TEMEERVSTLEEEELRTAHSTTKMSAERELHVTKLQLEETVSRLESYGTTPEQTV- AAFTTELQHTQQRLREAEIIIQLTNKLNAAAGVVRVTSQSDKDGNAARALVSDV- AVRNADTDLGTQLASALVALERLAEEEREAALKEATEMEERVSTLEEEELRTAHST- TKMSAERELHVTKLQLEETVSRLESYGTTPEQTVAAFTTELQHTQQRLREAE- EIIIQLTNKLNAAAGVVRVTSQSDKDGNAARALVSDVAVRNADTDLGTQLASALVA- LREAEEREAALKEATEMEERVSTLEEEELRTAHSTTKMSAERELHVTKLQLE- TVSRLESYGTTPEQTVAAFTTELQHTQQRLREAEIIIQLTNKLNAAAGVVRVTS- QSDKDGNAARALVSDVAVRNADTDLGTQLASALVALERLAEEEREAALKEATE- MEERVSTLEEEELRTAHSTTKMSAERELHVTKLQLEETVSRLESYGTTPEQTVAA- FTTELQHTQQRLREAEIIIQLTNKLNAAAGVVRVTSQSDKDGNAARALVSDVAV- RNADTDLGTQLASALVALERLAEEEREAALKEATEMEERVSTLEEEELRTAHSTTK- MSAERELHVTKLQLEETVSRLESYGTTPEQTVAAFTTELQHTQQRLREAEIII- QLTNKLNAAAGVVRVTSQSDKDGNAARALVSDVAVRNADTDLGTQLASALVALER- LAEEEREAALKEATEMEERVSTLEEEELRTAHSTTKMSAERELHVTKLQLEETV- SRLESYGTTPEQTVAAFTTELQHTQQRLREAEIIIQLTNKLNAAAGVVRVTSQSD- KDGNAARALVSDVAVRNADTDLGTQLASALVALERLAEEEREAALKEATEMEERV- STLEEEELRTAHSTTKMSAERELHVTKLQLEETVSRLESYGTTPEQTVAAFTTE- LQHTQQRLREAEIIIQLTNKLNAAAGVVRVTSQSDKDGNAARALVSDVAVRNAD- TDLGTQLASALVALERLAEEEREAALKEATEMEERVSTLEEEELRTAHSTTKMSAE- RELHVTKLQLEETVSRLESYGTTPEQTVAAFTTELQHTQQRLREAEIIIQLTN- KLNAAAGVVRVTSQSDKDGNAARALVSDVAVRNADTDLGTQLASALVALERLA- EEEREAALKEATEMEERVSTLEEEELRTAHSTTKMSAERELHVTKLQLEETVSR- LESYGTTPEQTVAAFTTELQHTQQRLREAEIIIQLTNKLNAAAGVVRVTSQSD- KDGNAARALVSDVAVRNADTDLGTQLASALVALERLAEEEREAALKEATEMEERV- STLEEEELRTAHSTTKMSAERELHVTKLQLEETVSRLESYGTTPEQTVAAFTTE- LQHTQQRLREAEIIIQLTNKLNAAAGVVRVTSQSDKDGNAARALVSDVAVRNAD- TDLGTQLASALVALERLAEEEREAALKEATEMEERVSTLEEEELRTAHSTTKMSAE- RELHVTKLQLEETVSRLESYGTTPEQTVAAFTTELQHTQQRLREAEIIIQLTN- KLNAAAGVVRVTSQSDKDGNAARALVSDVAVRNADTDLGTQLASALVALERLA- EEEREAALKEATEMEERVSTLEEEELRTAHSTTKMSAERELHVTKLQLEETVSR- LESYGTTPEQTVAAFTTELQHTQQRLREAEIIIQLTNKLNAAAGVVRVTSQSDK- GNAARALVSDVAVRNADTDLGTQLASALVALERLAEEEREAALKEATEMEERV- STLEEEELRTAKEKLEERSVEEISFLKDEVLSNRLLVDSVSSLNGKVGDSGAVGADV- ERLSRVDELHAQVSATKRGFEFEDRRSEGCVTILIVARRSVDRSNDARRRL- EERNVRLQDLERKCLEVVKLQKCEQRLEQFVRAKDVIRGAVSVLGVGDSVDV- SSVGAEPVDLEAVDLAQFLQISSLHADLMCRKTCRQLESNQEEILLLEQNS- QSNAYLEDLDIRQQLVEMRQREELIERTLTERVDELGRERGEVSRKQ- QNNLSAQLQASRNKLSALEASKREGELAAARQAAEELAKAFSLMEAQVQTLRE- EVASTSGSPKRGSSSRQKAVVEGDEARIRMSQARVTFLEKALQRKDEEVQR- LQDELVQKDEQLDQYEQDAAKAAQDAENASRKTQLLESVQKLGDKKGLD- ELRYAKTRVVTYGGRRVSSVAQHSPPPEQQIRGSPVLGAGRTTRERVSLSVES- SHHSRTEQTRQVRQVMDIRSTRKRSRSANAVS</p>
------------------------------	----------------------	--------------------	---

<p>Angomonas deanei</p>	<p>Trypanosomatida</p>	<p>NCBI compound of: AUXM0100 0739 and AUXM0100 0607</p>	<p>MSSNWGNRGYNTSVNNYKSSFDILREQRLAAQNQSTPSAGRETPSGFYDHTP- SPLERSRYEVLHNPI SMAVQSGIEPPYYTRAPPTPAVR- PQLDLMTRDDLYNECAKLRWSFDNAQELRERSEQQRIYYRDL- VAKYQAEKDLLNSRLQSAIEAKSRAEETGNYRRSCESLRNQLNDLEEFFRRA- KLDLDRAHADLRGGQFLLEGEQGGQAVVVGMRKEDLREKDACIMDL SRQNSE- LEAELERLRNALSPADQAAOPPASLSEYPRSTYDNDPLARQRLEDELOAARE- ELDQLRNTLDDERTAAERELQHARSESSQMGHKELENALRNSDALQETIDQL- AAELAKRAPVSLSDWEDYQKYSDDLQKNKQSESTVRQLERQLENGPTPNEE- ELAINENKDLQIQTYSEEVELRAKKEALLQELEAVSAQLQOSTKQERLARVQ- KNRIVELELSEKDLKEANKLEEEIEAEKENALVDAAAATKRLQALDASEQEK- LDLEEQLRELKQSSDDQKADEEIRRLQAALDEMADRVRDEDIDELQRLQDER- TAELEAEVQRNRDLAANGRTLETEIVELNSQVAALESRLLAAKETEVGVSSQQS- EAAVAGLEQELANAVDAAERLSQALDASEQEKRDLEAQLHELKQKSNNNQR- AADDSAEEIRRLQAALDEMADRVRDEDIDELQRLNERTAELEAEVHNDRLE- AQLRELKQAGNSVRAVDDSAEEIRRLQAALDKLADRVRDVEELQLQLDE- RTVELESEVERNRLGAEQLRELQNTTNNQKNADDSAEIIRRLQAALDELAD- RVRDEDVDELQRLLDERTAELEAEVARNRDLATNGRTLEAEIHELCSQVALES- RLAAKEAEVGVSSQQSEAAQAGLEQELANAVDAAERLSQALDASEQEKRDLE- AQLYKLGKSSDNQNAEEIIRRLQAALDEMADRVRDEDVDELQRLNERTSE- LEAEVARNRDLLEEQLRELKSTNSQKAADDSDEEIRRLQAALDELADRVL- EDVDGLQRLGERTAELEAEVQHNDRDLAANGRTLEAEIVELNSQVAALESRLA- AKETEVGVSSQQSEAAVAGLEQELANAVDAAERLSQALDASEQEKRDLEEQ- RELQKEAGSSQKAADDSAEIIRRLQAALDEMADRVRDEDIDELQRLNERTAE- LDAEVARNRDLAANGRSLEAEIVELCSQVAALESRLAAKEAEVGVSSQQSEA- AVVSLQEELANAVDAAERLSQALDASEQEKRVLEAQLRELKQKSNNNQKSD- DSAEIIRRLKAALDELADRVRDEDIADLSASLAQEKDNTARLATELNALKSRLA- GLEETVSLRDESIRSLVSDLKACDEDRLLDTLKQVSDQKEVAKLQKVHADS- EQEIRRLQALDEMADRVRDEDVDELQRLNERTSELEAEVARNRDLLEEQLH- ELRKKSTNNQKAADDSDEEIRRLQAALDELADRVRDEDVDELLRQLGERTVE- LDAEVARNRDLAANGRSLEAEIVELCSQVAALESRLAAKEAEVGVSSQQSEAV- KAGLEQELANAVDAAERLSQALDASEQEKRDLEAQLRELKQKSNNNQKSD- DSAEIIRRLKAALDELADRVRDEDIADLSASLAQEKDNTARLAAELNALKSRLA- GLEETVSLRDESIRSLVSDLKACDEDRLLDTLKQVSDQKEVAKLQKVHADS- QEIRRLQAALDEMADRVRDEDVDELQQLDERKAELEAEVARNRDLAANGRT- LEAEIVELNSQVAALESRLAAKETEVGVSSQQSEAAVAGLEQELANAVDAAER- LSQALDASEQEKRDLEEQRELQKEAGSSQKAADDSAEIIRRLQAALDEMAD- RVRDEDIDELQRLNERTAELEAEVARNRDLAANGRSLEAEIVELCSQVAALE- SRLAAKEVEVGVSSQQNEAAQAGLEQELANAVDAAERLSQALDASEQEKRVL- EAQLRELEKQKSNNNQKSDSAEIRRLKAALDELADRVRDEDIADLSASLAQ- EKDNTARLAAELNALKSRLAGLEETVSLRDESIRSLVSDLKACDEDIKLLGQLQ- KENSVLTEKGGKALSDLSKLTSTLDKAEKRLAEQQQYAAAEVKRSACLQVSLD- EMAITVESTKDALQQEQEKLFKMDHLRQENDSLMAYSQKLLERSEELVLEE- KTQLCDVSVSNYNALRRSLEEECLKSVTLTEGEKDLLAEVASLTSQLEETAEQ- YAEKLLDVEYEAADRKAARWKVADRRLAQLKELLADYKMSVESIVEALRYTLQV- DAKKENSALVLSHVEAALTRADALRNTLNTIAVKDDDDENDSNPAATAGDDN- LENFSQETAAALTEYKESLQITISEMQVRMAQQAADVLAERGYDMDEAKRRI- ENDIRVETEANRRHQEHRDLLFELASANKARERAEAEARREADERATLLEESD- LQVMRLQQGDAASGSPSEEEVQLRSRYAQLAADYETAMNRLQLLERSGVER- AGDLYQQLLETEARCQVIQRELDLYKDNVARRAALAKSNTLAKYEEEMRYSDT- LHTNLDKVVQSLRATFGASDAKPSGTDPLSKLQDALQRNLQLEEQEQIEWKSRY- ERVEEKLHAIIEENNRRLSRLSGVVERLSTSQRAPTLTRSMPPPSTPSFVHMLG- SPVTPGPLAAPSPLPQQPLFSNRDVSSEDEEEEEEQTSRRSSRSFKSPQSR- KRTRSASSKR</p>
<p>Angomonas desouzai</p>	<p>Trypanosomatida</p>	<p>NCBI compound of: AUXL01001 657, AUXL01006 132, AUXL01002 005</p>	<p>MSRTPSPFAPSTAASRSPRTTPAPSPPPAPRPALETMTREDLYNECSKLRWSFD- DALERRRSEQAQYRYRDLVGYQSELERMQDQKGE- FLEAKRKAEEETERYKQSCDSLRRQLEDMDLEAYEARAERDRLGRDLRSQSLM- NGHRGAETVVESLKGELRERDAAVFDLSRENERLEKELEALREELRROAASPS- PTAAPVPFLAQDRLDSEVEELRDTVEALQAKLAETEQOHAELRQEAQVQDQ- RKRELDQNKDKDMENLLKRNALQEVTDNLATQLSKRAPISSEDDWQFQNDYD- NLAKKIKSLEEAEEKATTGLDAVNRQQQDELDSYKKEVELYQQKEEALLKEISN- LQAQHQSTEANLAELRRGMDEMOKEQSLLKENEQRKTLLETSIQATATEDAA- PQLEEMENRLQALQKAVDMLATSLSEQSEQRHELEQQHLHNGQSTVSAETIE- SAKGNAEENEALRAALDVMASKVSOEDVDALMKQIAELTNELEAERENSRAME- TEFASLQAEALAAKDAEVGVTSSQSEENQRHLEEDLQSANATADRLGEALAESE- EQRLDLQRLDELQORLADAKKQADEAKGSTDEVEALRAALDVMASKVREED- VARLEKLLADRTNELEAERDHSRSLNENGRALAEIVELASQIXXXXXXXXXXVAV- LAKKHADEAKGSTDEVEALRAALDVMASKVREEDVARLEKLLADRTNELEAER- DHSRSLNENGRALAEIVELASQIASLQAEALAAKDAEVGATSSQSEENQRHLEE- DLQSANATADRLGEALAESEEQRLDLQRLDELQORLADAKKHADEAKGSTDE- VEALRAALDVMASKXXXXXXXXXXLADRTNELEAERDHSRSLNENGRALAE- EIVELASQIASLQAEALAAKDAEVGATSSQSEENQRHLEEDLQSANATADRLGEA- LAESEEQRLDLQRLDELQORLADAKKHADEAKGSTDEVEALRAALDVMASKV- REEDLAALRANYDALKEKNVELNDAVKQLRLKIGDLESTVSLRDSRSLVSDLK- ACDDDLRKLGEKQENEGKLSLRLGEKETSLSLTKSNASDVALKNSEEATAALK- AKLVASTEAGDRLAEQVVLQSEKADLLQIDVLSLKDKEGVEVKKYGEELLEK- SEMILATLLEQRSRECNELANAVSLKLEVEEELRNATLDQELKEKESALAAK- EQLEDTVLYEELLEVEGATEKCEWRKTAEKRFQDSYRMAKEVETCVVGLV- EAVGCATDPDSLKTTSSALSKNVEGAVKNIRDTLAKMDTAAAEELISPSVPA- ADAFSSKMNDVSLCKKHISEIEEMQVRLTHLTAAVDVLAEERYDMGEAWKKE- QEIQTAEANRRNQEHDRLLFELDSANKAREKMDREARREADERATLLESEVQ- VLRQQSDARSVSPGELDRLRARNALQVADYEAIVDRILRVLERAGGRHEDLY- QQLVEAEARCQVLQRELDLYKSDAERRSALAKSNTLAKYDGLRYSHSLQSNL- HKMQDLRAAFDPSAPHKSTSKDPTETIAIORNEALEQERIELQTRYSRAEKLSV- LEEEENRRLAQLGVAERLSTSRRTLLPKDMMPPPSTPGFVDVPLSPAPTGPY- NIPSPLPQQPLFRITLYADDEDEEEVSRPSTVRSNTKSPSTQQGKAKRPRSK- SSRA</p>

<p>Strigomonas culicis</p>	<p>Trypanosomatida</p>	<p>NCBI compound of: AUX-H01001038, AUX-H01002460, ATMH01010215, AUX-H01000948</p>	<p>MFRSTYKSTSSSSFFPYQTSRVPTEGVEGGPHPTTTAPHPYVGDLEGPISVR-DAPKWWGGSSSLIPRASAGGDTSSPLSAGRRTGGDGRPPFPYNDPDDSSYYRYR-PEPATGVLPPPEPGHSTHGVRAPRDALLWATMASGREPPPAFAAGPSARQECLEDE-FLTTAAEGDELAKHTPAEVAAYVERLSNYHYFGEQLRAACERAQHLYKSEVARL-REELETRRLKQKQFVDVDAQDGVQHALQEDQVRRRLRGQLEEQEGLYSLRQEL-VARSQATSTATSQEIMWYEAQRVRAAEEARETQQALRDLAAELQSSKAQVRL-VEVQQQQQQQQPSELARRLQDLERENAQLRQSLAAEREEHEQQKQQLLEEA-LAAQAQVQLQADTDAMALALTRRAPLSLDEYDAQQQQLAALQREVAQTA-EAAQKKAIVAPLEATALELTELQRVAAAHDRTTLERNCHQRDEEIVRLQELL-QCYKTNLAAQEVSRVTEQYEHLLLEQVDAEAARHAAEEAVDDLQRQLQG-ERQGLVAEVAQLRQAAADAPAPQERQDLTEARAKIATLLQQVQNLGEALEVQTA-ERDALQVQLSASGATTSQEQSVQTIQEGRAAFTEVQRSASHDHTTETHIQHP-DATSDTASSSPYSHTRDRDPHQVRESQKAPAPKPVAELEALRSLAAEKTALQ-EQLASAGDAQAQAAVAAALQQENAYLGAEIQRAAAIVLRLGELKQLAALKAQKTEA-DAANAALAKVRELEAARAQADADMQLAAARKSDEKRARKEETAAKATPKQ-DEKEXXXXXXXXXXEDALRSLAAEKTALQEQLASAGDAQAQAAVAAALQQENAYL-G-AEIQRAAAIVLRLGELKQLAALKAQKTEADAANAALAKVRELEAARAQADADM-QMLAAARKSDEKRARKEETAAKATPKQDEKEPRESDGKAPAPKPVAELED-ALRSLAAEKTALQEQLASAGDAQAQAAVAAALQQENAYLGAEIQRAAAIVLRLGEL- KQLAALKAQKTEADAANAALAKVRELEAARAQADADMQLAAARKSDEKRA- RKEETAAKATPKQDEKEPRESDGKAPAPKXXXXXXXXXQEQLASAGDAQAQAAV- AALQQENAYLGAEIQRAAAIVLRLGELKQLAALKAQKTEADAANAALAKVRE- LEAARAQADADMQLAAARKSDEKRARKEETAAKATPKQDEKEPRESDGKAP- AAPKPVAELEDALRSLAAEKTALQEQLASAGDAQAQAAVAAALQQENAYLGAEQRA- AAVLRLGELKQLAALKAQKTEADAANAALAKVRELEAARAQSDIDLELSS- YAALLADEQTLISALKELKAQSVAGETASLQWEENVAAAVRAVEERSATEASSR- ATVDSIAAELSRTEAHVVMQIKQDQNTSGKRMREEETVAKAAKVPKQDEKE- PRESDGKAPAPKXXXXXXXXXEDAANAALAKVRELEAARAQSDIDLELSS- SYAALLADEQTLISALKELKAQSVAGETASLQWEENVAAAVRAVEERSATEASSR- ATVDSIAAELSRTEAHVVMQIKQDQNTSGKRMREEETVAKAAKVPKQDAK- EPRESGKAPAPKPVAELEALRSLAAEKTALQEQLASAGDAQAQAAVAAALQQEN- AYLGAEIQRAAAIVLRLGELKQLAALKAQKTEADAANAALAKVRELEAARAQ- AEERAAWQEEALALGKELQRQNGKEVGDREAAVSVVAGEIEKLEEVQLLAAE- ECTSLQTSQEAALRVDREALVATVRKMRREEAEDAERVRAKSDERHAELGRMES- RTRAKVAEFQAALDTLAREKVVHVTADLADREQLTALRDELKTAERRVALEST- LKQAAQETERLRMVVEERDAALREQLSYNVLHADQALLINALQMQMKTASL- EGARREAEQVAVQHAMEERAALEKRLLEEGDSVAGMQAMLESTMQGAA- GWREAFERAEKELAREGELQELTASTRAAAEESLMLQSMELQHTESLAKVE- ARKEEFTALTRVKNEMQYAKMDAEKAQREVVQQQLREVEEVNAQLRQALSRL- ETQREAEKANADEALRQLAVKYTAGREEVSVLTAQRTALLATEKEAASLRKAAE- LESTLRATNARMEEQEQLIEEALQAAADQEEALLLQVQELAEERGAELAAEK- AQHLLQVEALKKTLTLEDTATSYNDRIVVEERLDRMEEYRQLAAARQLRLTAA- VWLQATESLRDGLRVAEAAAGSLKGPATPLKLPLEKAAAWYADVALRETAEE- ARRSSQDLSDDSRRASTVQGGEGTHQEEVRLVQRNAAAGLCAEVGDQLRQ- LWELGRAARREELTRRLKLELQQLRAALDTLAVENVDLKEAGRNYQSLFNET- QAALQQITNDFQQQLLEEKAASVEVVTARRDMQRAGESRRRAEQRAQLLQ- ELEQLQHVKGELEGQLLAMSEERDELTRATETVEAVLAAEPAGAGSGDGLD- VRSASRYSTAMAQEAQVQLVTSLSQSDLLLVRRLREMEAQDAALRRAASAE- REVAALRTERVRLQRELEHVTTRRYRQSLLEQTDAGAAAELRELRLKVRLE- GVEVFRFEERERELARVRSQQEVQAGRRQEAVALTRLSMSSRMTDVRRLN- DGGERASRGGLTATDPLQVAEVQQLQHDLENYKLAVEHHLRLESQAQLAAE- EREARAMADKAVRQLSRHNERLRTTMQSERQSTLLALSPFTPPPPLTVGASP- LPRTPPTPTPEEEEEEEVARTPTQSRKVEDDGPAGSGGRGRGSAKAKAPA- PEGGTPTPRKSSKAGRATGSRKRMRSHTSSS</p>
<p>Strigomonas galati</p>	<p>Trypanosomatida</p>	<p>NCBI compound of: AUXN01001911, AUXN01004474, AUXN01001425</p>	<p>MFRSTYMTSSSSFFPYQTNRNPTSTGEGYISTPQPNNNNRQEMPAFSSLEG- PISVRQAPKWGGGAAPPPLRTSAGGMADVSSPASGR- RTGGGGGDNLSNPTYPQQQQQQHQPPFNPDSSYYRYQQQLPPP- PQQQQASASLTPRGATRPAYTSVGGGDSVVQAPHDALLWATMASGREPPVFSH- PYQQNQQQQQPQRQECLEDFLLTAAEGEGLEHKTAEVYKAYVLRLSNYHYFG- DQLREACERAQHYKLEVARLREALETQRLKQRYNLDAAEQDSVQHALQEEQV- RLRGQLEEQGLLALRQQLARSSQTTSTATEIMRYEAERVKAASELRSAAQ- LRDLAADLQRSKAQVRALEVVQQQQQPSELRLKLEKLENAQLQESLAAERAE- LEVVQRLEESLQADTDIMALALTRAPISLEEDYDAQRRLAGOORQLLELQ- VEVEQCKAAVAPLEATALELNEHQHVAASHEKAAALQREGRKDEELTRLQELL- QCYKAETDAAEEANRVRTKEYEAALVEQVEAAEAARREAAEAVALRQAAQ- AQQTQTMEEGARSSQANNHQHHTDASSSSSSPSPSTADGTTTTTAMSGML- LGAGNEAHLRVAMGKIAVLLKQVQELGAALQTOQTEHETVQGTLEGGLELRE- ENAYLQGEIQRGADVLRLEARLKEQEKASKQSEKSAERRASTAAAHKEEQEG- LRVAAEKVAILSRQVEELGAALQQLTEKQELQEQLLAEESALQEQGASQQE- ALAALKEENKYLCEEIKRGAAVVLRLEARLKELEKELSKQAQQPELAASNK- LEKSNKLEERVRELEAAQAQAEALQALTVAEREQEQVRELEQGVVQAAQ- STSQEQSAAVKQVHQEQEDLRVAAEKVSVLSQQVEELGALEQLAAEKAL- QEQAVADEKTDKQTVSALKEDNAYLSKEIKRGAAVVLRLEALLKEQEKELRRAK- QPEIAAENKLEKSNKALEERVRELETAQAQAAEVQALTVAREEQEQHVRELEK- GVKAEEERVSTSQEQAAAVRQAQAEQEDLRVAAEKIALSKQVEELGALEQL- AAEEKAMQEQASADQKSQDETSSSLKDNAYLSEEIKRGAAGVLRLESHLKELE- KKLAKNEYLPDIAADNKKLEKSNKALEESVRELEAAQAQAAEAERVLTVAEQ- QHVRELEQNLVRLVLEQLAAEKASLQQRVATLEDELVQAQTDAAEELALEEKND- LQALQEAHDELQKDEQALISALKQVKAESAEEHSTTEQLQQEQADLRVAAEKI- AAVLRSQVEELGEALQTLAHKAELQAQLLEAXXXXXXXXXXNLRALQEAALASK- QQRVATLEDELVQAQTDAAEELALEEKNDLQALQEAHDELQKDEQALISALK- VKAESAEEHSTTEQLQQEQADLRVAAEKIAVLSRQVEELGEALQTLAHKAELQ- AQLLEAESALQEQGASESQOETLSTLKEENKYFCEEIKRGAAVVLRLEKEL- EKELAQQAQVVAESKQLQERVRELEAAQVQAAEALQTVAAAREEQEQHVRELE- QGVKAAEERASTSQEQAAVAAQQTQQEQQAQALQEQQLMSAVESVAVYEEA- LNALQENALQGAIEKRGAAVLRLEARLKEQEAQKSEKAEVAATKALRESRQEL- EALAKAEGASSADVANLRAALEQLAAEKASLQQRVATLEDELVQAQTDAAEEL- ALLEEKNDLQALQEAHDELQKDEQALISALKQVKAESAEEHSTTEQLQQEQAD- LRVAAEKIAVLSRQVEELGEALQTLAHKAELQAQLLEAESALQEQXXXXXXXXV- YEEALNALQENALQGAIEKRGAAVLRLEARLKEQEAQKSEKAEVAATKALRESR- VQELLESALAKAEGASSADVANLRAALEQLAAEKASLQQRVATLEDELVQAQTD- AELELELEEKNDLQALQEAHDELQKDEQALISALKQVKAESAEEHSTTEQTEKLV- ENEKHEARLVYDQLKEQYTLTAVLTKKALKEEAEEERQAMQAAVKSSEKITT- LENTIAKAQTDAAELETALNDKDGALQALQKSYDELKDEQALIVALKMKMEE- SAALHSATEKSEAQAAAEVRKVEERSRMELESNLHAALDKLASEKALADTLANKEAL- NAVRTELKQAQASIKEGEKVATLEDDLKVVQTDAAEELAALEEKDDSLKALQES- YDELLKDEQALIVALKMKMETTASLEVTSRQTEEQREEKLHVAEARAFAEARL- RDEGTSVAVQALLETQAQEAASWQELAKAEQTLAEKDALVQLTEAARTAAAE- LAKLEEMEQLHSAASLAREVARKEEFVTLTKLKNERYAKMDAEKAMQQAQTEL- RTLQETHAQLQETLARVEQQREKAKANADEALRRLAEKYAEGKEIDVLTAQTE- LLVKEKAERLRAKAAQLEATLRADNERLEEQGLIEEALARTSGEEQQRVLA- QVQELADRGTELAAEKERLEAEAFKTLDDTARSYNDRIVVEDRLRDRVEQYR- QLAARLELRTAAAVWLQATESVRDGLREVESGAAGRTAPHLPLQIQEAVAW- YRVVALREPPVEATRSGDLSDESSRRSSTAASAPDEVKSLFEHNATALCQ- EYVGDTRQLCEAGRAAARRETLARKLEEQQLRAALDTLAVENLDLKEAGRNYL- TLFNETQAALQQITNDFQQQLLEEKAASLEVITARRDMEERAEARRRAEERLQ- LTRELAQLQEANRELEGLIEIGEERDELTRVTVANEAMRAEASLEGSPSRNG- GGDIIHVRASRYSAVAQEQQAQMLQVTSLSQSDLLLARRQLRELAQESALRRA- TLAEREAVLRTENTRLQKELVQEREHTTALKRYSQLLDQTDIGAAEEELRLRLK- VRELEGVVFRFEERERELARVRSQQEVQAGRRQEAVALRRLSMSSRMTDVRKQ- LEAGEGGGERASTGSARSRRGSLTATDPLQVAEVKHLQHDLETYKLAVEHHL- QLESQAQLSQEREARTTADKTVQQLSRHNEQLRQLGHERQSAQLALSFPTA- PFLQGLGAPSPPLRTPPTPTPDDNNEEEEEEEETTRPLRSIHTQEEEFDEQ- VVVRKSATLVAVSNKRTSSNRRTTATKQTKIEEEETKKAASITPVKKTESR- KRNRSKSASS</p>

<p>Strigomonas oncopelti</p>	<p>Trypanosomatida</p>	<p>NCBI compound of: AUXK01001 933, AUXK01001 270</p>	<p>MFRSTYVSTNTSSSFPHQTDNRNTPQGVAFSSSLEGPISVREAPRWGGHNS-GLPARPLDVTSSPASAGHRTGGGDLSDLGQQPLFNPDDDSYYRYQQQ-PLPAPLPPPSAQSASFAAAAAPARYRGGDSIVHAPRDTLLWATMAGSREPPAF-SPPSAERPAQRQECLEDFLLTAAEGEALDHKTPAEVAAYVQRLRANVHYFGE-QLREACERAQHYKAEVARLREALQRLQRQKDFIDCEAQDQVQHAALEEQV-RRLRGQLEEQEGLLLRLRQQLAARGAEGQSTSTSEIMRYEVERVKAEEAR-WAQQAMRDAAELQRSKAQVKALEVQQQQEQQRPAELARQALAALEKENAQ-LHQSLAAEREEHETQKQLEEAALQADTDTMALALTKRAPISLEEYEAQQQ-RLAALQRQIEQLATETAQAQAAVPLEATAVELTELQHVAAAGHEKAALEA-RAQKDEEITRLELLQCYKTEADAAGEANRVRYERYEALVEQVEVAEAARQA-AETALLRQQLLEEAQASAHQHEHQEGEENHDAFFAASGMLGGAGNEAHD-LRVAKGIATLLTLVQLEASLAAQQAERQALQEQLATTAASRSHDAASLQEEL-QQLRHEKAYLGEVQRAATVVLRLAARLSEQATTTQQQATAATEEEKEEELR-AVMEKLGLLQQQLQALGKLETQQAQEKATLVEQLLAAEGVVEQQRGTLTQNTA-ELEALRQENALFGDEIKRGAAVVRLAALRERERAEARQDEVMQNTLEK-QNMTLKKSNAALEKRLLEEEAHAQAEADVQALATAHEKEQQRVRELEQGVQ-AAEQRVLASQEEQVAEMHKEQERHASEDAHAQHLHATIAQYEGEELKNER-AYLAEVVKHGAAVVLRMEGLLKEHEKREARHPDVITANTELEKANKALKQLH-HLEHVHSQSESHLQALQESSDELAKDKQALVAVLQKQTTREESAAALRGAGEAAE-ERARAELANLRAALDVLAREKAEALTEALSAQEARAAQEQEQGARAAEERVSA-SQOQLAAEAQREGAHAAQLQAQAQYEAQVEALRQENAVFGDEIKRGAAVV-LRLEAALRERERAEQALPEALAAKEALEQRVQLEEAARAQAEADVQSLQESH-DALLADEQALVAALQKMKEEASASLNAATLQAEERATAEAGTLRAALEQLAVEK-TALQADAVTALTRDKEALEQSVRAGEAEEALEDRAVQAQTNAEVLEVALKEKNS-SLRALQESSDELAKDKQALVAVLQKXXXXXXXXXXXXXXXXXXXXXXXXXXXXX-XXADVQSLQ-ESHALLADEQALVAALQKMKEEASASLNAATLQAEERATAEAGTLRAALEQLAVE-KTQLADAVTALTRDKEALEQSVRAGEAEEALEDRAVQAQTNAEVLEVALKEE-KNSSLRALQESSDELAKDKQALVAVLQKTTREESAAALRGAGEAAEERARAELA-NLRAALDVLAREKAEALTEALSAQEEALHVARSEAEAAHKLKARETTAATLEDA-IAKERSDVAALTEELETKAASLEALQSSCDELAKDKQALVAVLQKTTREACANLE-ENSHEAAEQEELKVVVEARAAALEARLRDEGVSVAAMHAELEAKAQEAQW-ERLAKAERLEAEKDALVQRLSESAAAAAELAKLEELERLHTVSLANAVE-RKEEAFVTLTKLKNERYAKMDAEKALREARLRASEEAHAQLQALAAKVE-KQRDADRNVNADAAALRLAEKYTEGKAEIDVLTQAQRAELAKEKDAETLRQKAA-QAATLRADKERLEEQEGLIAELEALRTSGEQQEQVQVQRLADRATLTA-ERERLQMEADALRKTLEDTARSYNDRIVVEERVDRAKRQTLQAARQLERL-TAAAVWLQATESIRDSSLAVEGAPDGRKDRPAPALPLEPREAVAVYVAVLRE-PAIESPRDSYELSDDDSSRRSSTAGEGAATPQQPDPVREFEHNAAGLCAEV-GDRLRLWEAGRAAARREESLQRQLEEQQLRAALDVLAAENLDLKEAGRN-YLSLFNETQAQVQITNDFQQQLLEEKASSLEVITARRDMERAEERARRAEQR-LELLEELQQLSANRELEGLVAVGEERDVLTRVTRVVEREALLAEVSAETSP-SRDVNVVPSRYSAVAQEQQAQMLQVTSLQSDLLARRQLREMAQETALRR-LRQAQERETVLRAEATRLQEELVQEREQTVALKRYSQLDQSDVGAQVEELR-QLRLKVRLELEVFRFEKERELARVQSQEQVQAGRQEAVALRLRSMSSR-MTDVRRQLEVGGERDAGVSSGGLSLATDPQLVAEVQHLQDLENYEKLAVEH-HLRVESLQAQLTHEREARTADRAVQQLTRSNEQLRLTLGRERQSLQLGSPF-LTAPFAQMGAPSLPRTPLPPTTPDDEEEEEEEEEETTRTPVRSULTQEGD-EVIVVSRSRSSRTKKPTKKEKTEERRTKPNKTGVSVTPKLAASHKRT-RSGSAASQK</p>
<p>Herpetomonas muscarum</p>	<p>Trypanosomatida</p>	<p>NCBI compound of: AUXJ01004 927 and AUXJ01009 581</p>	<p>MMPLNDPFRGPRASSSSAATYRQPTLAPELARPISVRELAYG-GRSRQQHQHQQRQSTFTPPPPPLPDPAPLQTAPTAAPPVYAATSSSLEGEY-DMHTPGGRVSGARPSTSSLLTAAADRGGISISPLGDRQLRDFSSAQKSIDMM-SPRELYDHTSRLQDEARQTRGVYAAAYARDQYREEVGRLLR-QELHSRYLQSQALERDRGRYNDVIHKLQGDIAALRRDLDDAE-ARENALQGRLARANTAMAGPSSSSSGSGGTRPPAAPEADPA-QLAYATQLAAKDGQLELEAENTLRAALRDEQQQQQQQ-RSAPALTGEDQVAAQLQALQEDLQENAVLADALATERESRA-ALLAETAEEKTRMDADATRAAASNAALQKELDDLLEAFSSRAP-ITNEHYEHLRRELAEQRTIAALESEANRAAPVAEAVERPAD-GAALEASALRESVATLRAELADALKQAKTQELVAEFKEITWL-AETANTVROGLIDDLRANLQAATAAREAAEAALAAERTAHEAL-AGRTDADDVHSAVAELAAEVAALRATGQSLGEGKAALVSLA-STQAALDAMAERNSALEESEQRHGRAIEELQALQREQRAAA-AVTEGEPATQSAAYDALSGETALEAGVAKAADERAQRAEELA-ATQKALEDMAGLLDEERQORAGALADTAALDRSVAEL SARLE-GPTPPPPAAQLLDSGDQQNQQQQMAALQSSHAELRALSE-AESATAQAHAAMASLTELQAAQAAAQKAVLLEAAVDTMAGQ-LEAAEARARELVAGLAAEQSVSLAKEQLASKEAHLAAASAAVE-GPAGDQRTIIDLRTQDALSTANTALSERLEAAEATLQAAAKD-RDDQAEQLRRTSSELAEEAREADRHRRAAETTAQQLAAEAQ-VAALQLAIDSLAEQVDASQQRVAVLEETLREVQSVGDGARA-EARLTTEAEALQAAAQGGQGHSTDAVELNSQLAAARAHNEEL-TRRLQEMDAATSAVAERGGQPHQQQEEESASALASAQQLC-EKAREAAEERARREKADADVTALQAALQADELEKAKADA-SERQGVYDITTSMLQGLQLEEVRTNNAALSARLQSSAGVAST-ATERDQQSTAVEDLRKRLSVYAEALASRXXXXXXXXXXDLR-RRLEAAEADAVTKEAAVSAQTAAVAKAEADASALQAALDDVAV-RLAETESVVVDQQQALDSTASTLQRQLEELRGQNAALLERL-G-SSAADCDGHPTAVEDLRSRLAAEAADAAARQALDVAQELD-AAKAEVEDLRASLADARAHATTLQHEAARRESAVSHLTGAVVSR-NEILHSTAGQSAISEALRRQDAAVEEKLRTVTEAALADAH-D-ARVELGQLQAKARSSKNIAAAEAQAQAAYRTAQEALN-RMADEVARQEQARAAEATAAAQQGNSASLVATTAESASASAA-LQSDLDVSTALDAERAHAKELGVDLEAARLEEGAARLSLA-ATDAETSALRAEVEALAAKATLEKALADAEKSLADAEKVTR-QLTAERDQLAAARAQAREAAETRAE GAARASAVENGQLLQA-AVNTLAGDLDAERARTAEALAESALRARAAL EADLAAAKT-SAALESTLAGEARSTEDRLRGFTEELTAQRDTAAQLQRDKAAL-QERLDALQRSASTATAVEELANRVLSAEEAALRSVYAEV-AVERQRSDQTRRCEALEKELATATAASQAAAAQVEELKRRIG-DLETGAGAAAAKATAATALATAQQRVVDLEALTAQAAVELSDA-TAQLKELAEERRAAEKKEHASALAAIDDLAKQLQEAQEAAGEL-GTVRRRWEEQRRCASLTERATQLDITLADVMDRLDTSALL-ARSRREMLAATDESKALEAALAAKKEQLTTSERQRVDAEERL-DAALLQNTQTAAMYNDNRNDEISRTQARATDLKSRMLMDRDSAL-ELAAHLCDCAKAVQVALLGVLROPTTPAGRSSGSTARGAASQ-IAKDLRPFALDVIATVREDVADDDRVVALPDDCTEDEATAA-ALRRRSSARASSSSLSELAADKEGNAQRMSSLTSALEELLNE-LRQAGVATGLEETLAATEERLRAMDEALVRLSEDKLSEVERA-DSLYAALNVQRAVSRTOAQYEQQMRRETESATADMVAQRA-AERAEAEKRAEERLSLVEQLEEFQRAELRGLQANRDLGL-ENNRLSRMLPNSHRGPHGHDDGADGRSEPGSATARRGFGD-VDQAEAAQMLHVATMNVLMQTRRRVRELEANDTLRRAYTV-LETQLEELRDEVGNERLRGELDRLEEHADLALRYDALVETT-DATEGGVAQELKSLRQQRSLLEAELEWKARYRDLNLRNLP-KVEASQREAALESEQLRGLTSELDAQAQYARPSTANGSGAG-RGGDAAQARLAHLEASNREKARLQQAARLQDQSEIAELE-DEVARKTALEEAERLAAQYARQVDSLTTNAKALRRSTQVGRA-SSALRLSPPGTAQPGPSPQPRAPSEEGTRRGHSNSPSTST-RSPRASTVGSRTPRPSLLGAGRGGAKRRTTSPENPPATKRTKA</p>

<p>Phytomonas sp. Hart1</p>	<p>Trypanosomatida</p>	<p>NCBI compound of: CAVR02000 1000 and CAVR02000 1443</p>	<p>MYSRHSRFRALHPASTKEVRMTSVFSPDLPRPISVRDLAYEDASSGVR-LVRSRTQPISHSRAPHITNPPCLGGEGGGRMRTRSEIRVGAPTAGVYATGR-LAEAPSSWRRNWARITPPQPDSTSLAACSLDEMNEGELYTFASR-MQDEARHARGHIAQAYELRDQYIQEVARLKQELHSHKYLQTAELERNRKGYEDTIQS-LREDIQLRRELDASEGREQVMRLRLAAGTSVFPPTLASGSDIPDGGSGQYFIVKY-LRQTELTTKNQIQIHLKEHIEHIELEKQLQNEQRIQTPEVFPKSPSRSAANSEM-YTIGTDJIFSRKVLGSAVHEPGICWDPNELLRQIQSLEATNSKLRQDLLEESVQRQ-STIQLEAERSRSELQASSQTRINNLEEEAEQLRTLNLISLQEVDMLTITLANQAPLTR-GAYETFDQENETLKRIFSDLEAKYQSLERENLVAAALETDMLESLTKVRLHEHRD-TTEQLKLNKLELVLVSTFEETKLSMTANNLRKTQVESFHKQIEELTSELNALQSL-NEKKEIEGAAEAQVMKNQLEAKLVELEEKCDNAEKLNNYTIQIEDERKRKRKQ-EREMLNQTKQHKRLTQTVVXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX-XX-LKYRS-NAENIIASLKSQISLLSAELSDHGFLLPVQSIESGSDTILHDSIQFDDSSLVHLTTIKA-SSVQRNYSLQYALALRDVISAFCLTSDLNRFVKNVEHLNDSSTLSTDTFVLPSSG-AFEACLEPVAMASRWAAQMDILKKNSTDELGSCSEFDRSLRIPKTNLVDIAEQE-AWSLISAIPEFVSLVEARKASDIVCLRDLLSRVDSLVYSLQEEKEHSRSLSAAMDQI-AEALDDEKERSDSLFGTIADINQQLKADKDFEQIRKQEVASSVELIAARRAEGRAL-EAKRRAEEQLVLETELEERMLRKLNDYNGELKIEAERLRLTGSISSANQASQ-QDWRATEAENTDTHNQIGLRDDQLEVLQQAQLQISTLQAEELQMRKMRVREASR-EAVLNQKNRVEQETLAEQHEEVHKTALLMDFDQLQSEYTDLLNRYKNLVESST-TEISTIEELKQLRRQVRLLETELAEEKAHTRAIELAKNTPDLEAMLRQEALETGL-TSTHLAAVQAQYARPSVNTSASPNDHFSFGVEPKSNDDVSGSGVAVGSNEDDTIFR-TPNPSIDLHRSRITQLESSIKENKKEVYAADILRQQTVIETLEDNLTKLTLDRTE-QLNEQYLSIDALTRSSKALRASAPVRLVVENTKTAAYASDKHSNRSSRSLRITTS-GEDASRPLGSLPRTQQQFVVESSAQSDATESDDVSVPSKSIMKVVSPALTKK-STGARSRPRLKDLKDNLIGGVEKGHRRSNRRSESQSNSESPRVSTTTSRTSN-LALHTSSARILTKRKRGRGIS</p>
<p>Phytomonas sp. EM1</p>	<p>Trypanosomatida</p>	<p>NCBI compound of: CAVQ0100 00701 and CAVQ0100 00700</p>	<p>MFSRHQDFVLHSGSAKQVEMTPPFSPDLPRPISVRDLAYEDLSSGVRPVR-RGDPPVLRSRTPATGRPSRAVPPVTSNPAERLWGGNRLPRISAPTAGEVYSATSR-LAETPWPVRRSQQEPLADSLARGLDDMDGELYNLASR-LQDEARHARGHIAQAYELRDQYIEVARLKQELHSHKYLQTELEERNRDRYLDTI-QSLREDIQLRRELDVSEEREQAMRLRLAAGASVFPPTIAGSDVSESSAQH-AFVYKYQTEITTKNKIQQMKHEIEMLEKQLQKEQHQITPESVFPKSPSRSA-ANAHSETYITGIGNISSRELALGDRPVQELSTGPDLDQLQQAQLEGANAKL-VQLEESVQRQSTIQLESERTQSELKMASQTRIVELETEVMQLRELNLQRE-VDTLAPLANQAPLTRGAYETQEEHESLKMRFSELEIYHSLCEARVLAATA-LEDMRLRESLTKLREHRDTEQLKLNKNEKLVLIATFEETKLSRTANDLRKQVQ-VESELHKEIQLDLSLRLRNHSTHEQVEMRVAEEGKQVKTDELETKLAMLDEQ-RINAEMKQKAMQIQDERKRCEDLEKEVLDRTTQLENLXXXXXXXXXXXXX-XX-XXXXXXXXXXXXXXXXXXXXXXXXXFGYRSNAESLIASLKSEVSLRLEIIDLRLSILA-SQKSLQVTSYNDQIELDDSRQALKAMRESFEDRHNSLYAVAYLRDVIAS-FSRRLTDVNCVFKFSVARLEDPSLSPDSTSELSSVFDACLQVPAEASRWAAQ-THDTLHKVVADELGNDFPDESNTAKSNVSGASIDKEASVLAALVPEFVKA-EAQKALDVIACQDVVSRADRLAHSLEQEEKERSRSLTAAMDQIADLLDEKER-SDSLFSTISDINQQLKTKTYFDEQMRQAAAASVELIAARRAEGRAHEARRR-AEEQLVLVEAELEEEERMALRKLRTYGEKVEAERLTRMTSSISDVNQTSRQ-GWRAAEADDDASRHIGLRDARLEVLQQAQMLQISTLQAEELLTRKMRVREAS-REAILSQNRILEQLTEQHEDVQTTTLRVELAQQLQSEYTDLLERYKNLVESST-AEISAIIEELKQLRRQVRLLETEEKAHTRAIELAKNTPGLEAMLRQQALETGL-TGLSTHLAAVQAQYARPSVNSASPYRFSLPVGLKANDGISGSGVTSGN-SEEGANLRTNPSIDLQSRHIAHLETVLKEKDKAMEKYASDNLRQQAIEIALED-NLASKLVALNRTEQLNEQYLGTDALARSSKALRASAAVRLSVKTNKAVCTAS-DQHLTRSPRASLTAASGGVSTGSGPSPPPRTPQPLVETGSSYSEGSSESDVT-STPFEIEKVVSRISPKKSPRARLRARREVEDDDPSGDIVRVRRKSNIRSE-SQGSSVSPRVSTTTSRASNVARNSSLAQTLQRRKRGRKIS</p>
<p>Phytomonas serpens</p>	<p>Trypanosomatida</p>	<p>NCBI compound of: AI-HY0100182 7, AI-HY0101464 4, AI-HY0101944 4, AI-HY0102015 3, AI-HY0100994 0, AI-HY0102020 7, AI-HY0102013 1, AI-HY0100260 5</p>	<p>LVRRGDDPPPSRTPATGRPPPPRLSAPTAGEVYSATSRALAEPTWP-PRRGRGGEGGPPQVDFSPLEPRGLDLDLLENELYNVASR-LQDEARHARGHIAQAYELRDQYIEVARLKLELHNNKYLQTAELERGRGDYED-TIQSLREDIQLRRELDNAEGREQAMRLRLAAGVSVFPPTMDAGSDVAESSAQ-HAFVYKYQTEITTKNQIQIHLKEHIEHIELEKQLQKEQHQITPESVFPKSPSRSA-AAHTHSETYITGIGNISSHELALGERPVHESSSGQSDSLLQQAQLEGANAKL-VQLEESVQRQSTIQLEAERNQMLKTTSEDRIQLEKEIEKRELNTSLQREVD-TLAPLANQAPLTRGAYETQEEYELKIRFSELEAEYHSLEREGRVLAALAE-LEDMRLRESLTKLREHRDTEQLKLNKNEKLVLIATFEETKLSRTANDLRKQVQ-VESELHKEIQLDLSLRLRNHSTHEQVEMRVAEEGKQVKTDELETKLAMLDEQ-RINAEMKQKAMQIQDERKRCEDLEKEVLDRTTQLENLXXXXXXXXXXXXX-XX-XXXXXXXXXXXXXXXXXXXXXXXXXFGYRSNAESLIASLKSEVSLRLEIIDLRLSILA-SQKSLQVTSYNDQIELDDSRQALKAMRESFEDRHNSLYAVAYLRDVIAS-FSRRLTDVNCVFKFSVARLEDPSLSPDSTSELSSVFDACLQVPAEASRWAAQ-THDTLHKVVADELGNDFPDESNTAKSNVSGASIDKEASVLAALVPEFVKA-EAQKALDVIACQDVVSRADRLAHSLEQEEKERSRSLTAAMDQIADLLDEKER-SDSLFSTISDINQQLKTKTYFDEQMRQAAAASVELIAARRAEGRAHEARRR-AEEQLVLVEAELEEEERMALRKLRTYGEKVEAERLTRMTSSISDVNQTSRQ-GWRAAEADDDASRHIGLRDARLEVLQQAQMLQISTLQAEELLTRKMRVREAS-REAILSQNRILEQLTEQHEDVQTTTLRVELAQQLQSEYTDLLERYKNLVESST-AEISAIIEELKQLRRQVRLLETEEKAHTRAIELAKNTPGLEAMLRQQALETGL-TGLSTHLAAVQAQYARPSVNSASPYRFSLPVGLKANDGISGSGVTSGN-SEEGANLRTNPSIDLQSRHIAHLETVLKEKDKAMEKYASDNLRQQAIEIALED-NLASKLVALNRTEQLNEQYLGTDALARSSKALRASAAVRLSVKTNKAVCTAS-DQHLTRSPRASLTAASGGVSTGSGPSPPPRTPQPLVETGSSYSEGSSESDVT-STPFEIEKVVSRISPKKSPRARLRARREVEDDDPSGDIVRVRRKSNIRSE-SQGSSVSPRVSTTTSRASNVARNSSLAQTLQRRKRGRKIS</p>

Crithidia fasciculata	Trypanosomatida	NCBI AOD-S01000118	<p>MWNSSFYAEPHRAGGSAPSSTSLPQSSRKPPPVGLSPHLARPISVRDLAQS- RVPTVAPPRSSHSFSGGAHPTSPRTSGGFTGLAVPKAPVVAATS- GVSLGEHSGRPTDVSLSLQNTVEASDLTSLAARNIDHFNFDLYTYASHLQ- DAAKNYSNNLSAAYTKRDMYRSQVAELKQELQERYLQVDLLRRDHERADDALL- RLRENDTQLRQLLAESEGLKRNVSARLSLRNGADPAAAMRAYQAQLAQKDAQL- RELVERLAQQTAEEAQRGSEASAAGQRRSHLSAAADEETRANEQQQHAE- EEARLQAKVRAMEERVREAREATASIAEATTVLQAEIDALHQQHATLQSQLA- AETASHTQERTELLETTIAALRREKGDMESEDAIAALTRAPVSKADFESLQAS- YAELOASLANSEASVSLREQLSKAVVQEGYTEVLQQLSTTTQAERQELVTKTE- QLQALVDNLTRELQLADEANQLRQQQMEETAVANSRDRWQQSQHTIQELQERE- ASQVAELQRVRAEMDSLRLDRLQASRSAMTTTTTATTNNGDITHERTELMAE- LETLRNTSGDLVRQDALMSEVRHLEERVQAAKEELSHHSSTPPALSPASQEG- EAIADTAALQVAQCCIMALQHDLVANSVAARAEADMQAQLAAQEKQVADLAE- LRVAQERQREAEEDLNHARAAAAAEHHLKDVVKALKEELETQQRTELLQR- DKTHQEHHAQLHVTEQLRERIVDLEKRASEVRPHEPTPPSTQPTHISARTSTA- TEAETEKVRLTTMSAIDRMAEQLAASEDRVRDLEAERQRQADELEHARRAHA- LEDRVVAHERGSRVAVSDGGASAPAPEHAMRVLAERDAACAEMAEELRSL- LQSHAEHYAADLAAMEQRISVGEELRVSKAALEQMADLTATTATHLVGLSTAEK- QLEAEKVDLSTALDAAQSSAVNAQKENDALVATDTERQNKLDAAQAAVSAAMQ- ARNAEQRATTAEEEVRAAKAALEQMAGDAADTAQQCAALSAANTELEAKVLS- LASEFAAAQDMLMLTNSERDSNAEQESTASQKVQAAEARASAAAAELSAVTHR- VSASEDEVRAAKAALEKMGDLSAEGAVQILTAAKSTLEAESASLKAALFAEQ- TAATLAAERSSHAGTESEARAMEAAEAESVVAELSTAQQRLAASEEEVRAA- KAALQMAGDAADTAQQCAVLSAANTELEAKVLSLASELAAAQETVVALTSE- RRAESESTASQKLAARAAELASKASELEQVGTTKELRVAKAALEKVVAD- FTESTRLAELLSAGKADVEVENAALKGDIAAAQALFSLRAERDKONGSDSTV- AKVAAEAEMASAAAAELSAMTHRVAASEDEVRAAKAALEKMGADLFDATQLVEY- LTAESQLEAEKARLYSALSLEESVATLRKEHNSCSMKDSEVRARLEAAEAMA- LNAAEVRALQQRATTTEELRVRAAVDSMAGMTASAHNSDSALARQTSLQ- DDVVAEAERIASLTVCESSARDEAAQLFVQAQAAEARSAVATDRATMEERMN- AAVEELRVSKAALEQMAADFTTATRLVGLSTAEKTLLEAEKIKLSTALDAAQSSAV- DARKENDALVATDTERQDKLDAQAQAAVSAAMQARNAAEQRATTAEEEVRAAKA- ALEQMAGDAADTAQHCAVLSAANTELEAKVLSLSELAAAQETVMTFTSERDS- RAERESTAVQKVDAAEARASAAAAELSAVTHRVSASEDEVRAAKAALEKMGAD- LSAEGAVQILTAAKSTLEAESASLKAALSAEQTAATLAAERSSHAGTESETRA- MLEAAEAESVVAELSTAQQRLAASEEEVRAAKAALEQMAGDAADTAQQCAV- LSAANTELEAKVLSLSELAAAQETVVALTSE- RDNHAESESTASQKLAQAAETRAA- ELASKASELEQVGTSTAEELRVAKAALEKVVADFTSTQRANDLQANTALEAL- RSQDVADLASAQERVAATEAKLVSESAVEKVAEELNSAKRLEFLAVKADVE- EQYKQRSALQLVEQRVALESRLGDAERRSADAATAEQLLSTYTKCDELEK- ANARLREQLSEARSVNTLAEECTSATAQSSTLAQKGETEAEELRKTAAEAHVTL- QEELSSAKAAIAQLDDTLTAQRDYNDTVQTARAAQEEAMAVLVSKAAEDSAELA- EVRRRGEELQTMMLDITLRLAQLKELLLEEVNDNAELHKERAAALADDVTLRSA- LETAVVERQDLQERLAAANEGLERTAEYNERILEDDRMQEELNAVSAFAAQT- QEMRAGLAAMTEMSAGFCNVLETVMAPAKETLLAITDAAASSTDGGAGGGAA- SARSSQLISAITIALAPWSAALDWAGKMERRLDALERGRAAVDGEDAKADTDA- ASTPPSSPLSHASDPATARSALLDFVQYNEAALGHLHHIRRTMADRDELRL- RSALDSMADATVEERERVDHLLSAIQDAKDHVARAEREFDEQLRLRSEAITEL- VVARRAEERATDARVRAETELAATAEELRESHLALRKLQEEHRLKTRERELSR- YSARLTDTSAAARGRGLSGSVPTISAMLAETATNSYAVTASAVEALEQSLLQV- SSLQDQLLRRRHARDLEGREALLRQNHSTLEQEVVSLRVEAAEAADLREEVAT- LRADNAELLEERCEQLERMLAAAGGGTMQELKSLRTLQKKELELADLRMRD- FLLNRSTSEFEVVRREGLDRETLGIGITLQAAQYSGGRTARERGTALLSG- GAGGTAGDEASPGVLSRIDHLENVVRERDAVIEKMQASQLESRTVDGLQE- NLSEKAALDRANTLLAQYATETIETHSAGVAGTTTRQLPATPRRLTSESPHSAR- SRNNTASVPPPTPLGMSMHDDTAGSISRAGVTAPMASPAPRMPNLDVAGEE- EGENSSAGDEAEAPSESPASVRRRTGNSGQRSSRRAGGRRSSEERTSADAP- VPSRSATRSASATTTAAAGRKRGRPPTRK</p>
-----------------------	-----------------	--------------------	--

<p>Crithidia acantho- cephali</p>	<p>Trypano- somatida</p>	<p>NCBIJ com- pound of: AUX- I01001182, AUX- I01003068, AUX- I01001621</p>	<p>MWDSTFNPAPSRPPSDGSSLSASMSATAPRPPSSVLSPLDARPIVSRDLAQSRLV- PAVPPPRASANPNYLSGTPNAHRYSSGDGERANEAPSPRPSS- GFGFGVPRAPVVAATSGVLPTWPPSSRVSSGLDTPPELSTHGEASDLT- SLAYRHIDHLSDDLYNYASHLQDAAKHYSSHLSEAYAKRDYRSQ- VAELNHRHQEEYKTDLLRRNRKADKDALQKLRDNETELRHLLAEA- EGQTRNLSAKLSLQSGAEPVSVIRALQAQLAQKDAQLREMSERLLH- ATAEHESAVEALQQRSSSSGGGAHPRASNASSAAGLEGEDGTR- DPPQHEQQTDFDMEVEERVQRKVTETTTEQAVAAATAHLRAAMEDL- QQRHDLAQTLTQATAHAAALGRAAAQRAELVDAMTLLKQEKAE- LQANEDDLVAALTQRAPVSKADFESFQAAYAELOASLRTAESAAAT- LQEQQNASAAQEGYVSALQHSLLTTQAERQALADANRQLQALVDN- LQKELELAGTSLRLQEQVASHANEEQLAAQLQRTLAAEVMSTA- QQELERRCRQLEEEVRRQLQVDNDGLHQQQQQQQQQPPQQQ- VVEEQVEEQEPLTPASNNNNNAADATHQNSSSSAPLEEAANMA- REKAELEAELETLRISRDSLVVSRDSLVEAVRDLTEQQQQQEEEMA- RVQAGTATAAAAAAAAAAQMERDAEVDVATLRAAQECIALQRDL- AAANADLARREAAAQSQLQTLTNALAEQETLLAAAMERQNGAEE- ADRQRAAAAAAQQHAEDEVIEALKEELETQARDHELRESQSAYK- AQYEAKVREAAQLRERVDELERPATVALENTYAAAVHAQAVQQTGA- ATANAALDRLAELQADAEERTASLEREKHHLADELARAQRSVAVQ- KERSSAAVQGAEEQASDHSAPAVQASANNVTDLLRQRDTALED- ASRLREMHAAAEARAATVAQERDAAQQRATAAADELRTATAALDR- MVAELSEVTQRVEALSSSNNDLQREAAQLAAARDAEAENLSIAKAA- LERVVEESAALTECVNNLAATNAHLTDECNTERRALAAEADARTA- QAALKEKMAEESANVAQPVALTSAITATLEATNAAMQTELLTARDTVA- ALTKHEDTLCDANVELQQLTNAVAHAHAASAATAAEQRALQLRLD- AAEDQLRVAEAAVAVAEQADAVQRAQALANTNTSLEAASAAQQA- MELQRVLAHTNALRPVGVVAVSVAADADVDVSIQELTSSIAFLM- AEGTRMKAELSVARGAVASLADTNTNSDADAELKQTLTAEARA- ATVAQERDAAQQRATAAADELRTATAALDRMVAELSTASQEAATLTA- TVRALEAERDAAQCCASTAAEELRTAKAALDAMAELSEVTQRVE- ALSSSXXXXXXXXKQLTTAEARAATVAQERDAAQQRATAAADELRT- ATAALDRMVAELSTASQEAATLTVRALEAERDAAQCCASTAAEEL- RTAKAALDAMAELSEVTQRVEALSSSNNDLQRDAAQLAAARDAEA- AENLSIAKAAEKIAGELSEVTQQRATLTAVAHLTEERDEARQGES- AAAAREVFHVQAALDSAKAEELSATAQRNSEVVAAMAQLEAERDA- AQLNVAATAECRTAQAALKEKMDRLSEISQGDARMAVVEELER- QHTQLIADRNRMQDQLTAAEAERVTAKTALERMMADEHADAVDRLR- TSDASNSLEGGMAQATAEARLQQLSDAEARTATLERRNAELAT- ASAAQVTAHTAQVEELKENTRLKTDLSVARGAVASLADTNTNSD- ADAELKQTLTAEARAATVAQERDAAQQRATAAADELRTATAALI- AWWRSCRRRHRRRQRPSPRLCVRWRRRGTRQSSVRAPLQRSC- ARAKAARGWRVGRGXXXXXAAATVAQERDAAQQRATAAADELRTA- TAALDRMVAELSTASQEAATLTVRALEAERDAAQCCASTAAEEL- RTAKAALDAMAELSEVTQRVEALSSSNNDLQRDAAQLAAARDAEA- QKRAETAEAVHVAQAARQEMTKALAAATQHNDTLSNSNQTFEKE- IAQQTAELEVRVAAAARAAAEARRGAEALAAVEKSVAEGETSK- IAEVEAERTRLKEQLSVARTAVSTLADNCTQESSQIADLKQLETEL- HEHSAETTVAVKRDLANAQAQAVALTDLAVFSANEEAMKASVA- AQEQGIALLQRKHDDHAELAQVQRQGEELQGLLNTTLDRLTQTE- ALVKELNDNAEALHDERAALAAQLATVQAAFETVHERQDLQERL- AAAQGLEQTTDLYNERILEENRVQDELRAVAKAFAAQTEQVKVGI- AHVLMSEGFGRGAVTAVMEHATSTTQATAELALASGAANADAAT- RLTASSHNTAAAVTAVNAAFAPWAVTMEWVAQMQRRLQKLEASQ- VDENGERCSGGDADLAAAAAATLTTTRSDVSDTRNALLDFVEQQ- TMSLRHLQTVQHALNHRDEELQLLRSEAVAADLVARRAEERATDA- RLRAETQLETVAEELRTTQAALQKLEEEKRQLTRETRELSRYSTRV- VDVSARLRGAANGSVTPPLTALYAEESTHYTMNAAVALEQAHL- LQVSNLQQELMLCRRQVRELEGGQETTLRQAYMVLEQQVVDLRTD- AAEVEDLREERDTRAEANAEWEEKNAQLQRVMDATQRSTHQELS- QLRQQVVRVLESELEEKTRLQVLMIGKSTPEFEALRRQEFKENS- GLTTELAATQAQYSGGSNAASPQNVYITRIEHLQTAIKERDVIK- LQAAQLQSRVAKDSLEEKLEKAAALSRVSELNAQYLETIEALRSAG- VSAATRAATPRRNTAESLTRSQHDSVMPPTPLSLHTSTSEKTKG- AGLTARVRWQCRPLHHACRISTTWRPRKATKQVRSKRWRRRRL- LPPRAGVPLWEKAAPVAVVSVVRVAGPAPRLRQASHLVPRR- MLARQRQPANVVVSKFTPAQTNVTTGDPFEGDF</p>
---	------------------------------	--	---

<p>Crithidia mellificae</p>	<p>Trypano- somatida</p>	<p>NCBIJ com- pound of: AHI- J01002057 and AHI- J01000179</p>	<p>MWNSDFYAARPHRVGSPSSSTAPPPPKEVLSADLARPISVRELAQSRVPAVPP- PRASTSSFFSSGNSSGATYLN GAHLPPSPRTSGDFAGFVGPAPRVVYAAATSGVLPN- WAPARTTTVTPPAELSTYGEASDLASLAEHRDRFSDHLLYNYASHLQD- TAKRYSNLSEAYTARNMYREQVAQLKQELQKQTLQVLDLQRRDRERCDAAV- QKLRREDNQLRQLAEAEQGTRNIAARSLQNGADPSVIRAYQAQLAQDRDAQ- LRELTRLACESSVAEARGQSGDGDGEADTRPSRALGAEESNDDRLOQRQQQ- LLEEEIKAATAQTVAAATAELQAEFLNDLQQRKAALERQLTAQETAHAALARC- AERTEQLQVTFIERARQEQAEMQENEDALIALAQRAPVSKANFEALQTSFAETQ- AALTKAEEAVATLQEQQRVAAAGQYVEVLQQAALATGAERQELTDQNRQLQ- SLVDNLQREVELADTANQLRQQQLLQETALQGGEQLAEQVQHAEAMHAAHMSEE- IGALQERCAASEAEMQRLQSENEELHRQLQEPVVSQTATHASHLVASDMANL- EREKAELAAEVELTRGARDSLVYRDALVNEVRELEERQRCCSAELEEAAQLR- NTAAQLEAEPVSAAPPSPLDDEEDTEEHVMLQAARQCIALQDQLETERA- DHAHAEDAHAQLLALERTVAEQRAALRTAEEOQRTHAEAAQNTVVERLTNL- NTMHETLMTAAAVTAEERDALARAAAEEVEVRTVKALEELAGDLSSTTQR- ADTAAANALQAENAAAMKTEL SVARSAASIASERDSRREDDGAHERLEA- AEAHAARAAATECEGMQRLRTAEENDRTAVAALAEAMAELSDAEARVEALT- ENGALQAKTLTAERDDAQRRVAAAGENERTLKAAVEELAAKLAADQHV- LTAQGTSLAERAQLCSERDTRQRVSDAAEELQSVKAALEDLAECSDTK- RADALAAQVAGLEGERLKLAELHAEHVLAEVEVRLADAEHRDSERTAAME- HLSAAQLAQIASLEEQTSKQKELSTARAAISTMAEHRADTETQISDLQSQLNH- ASEQRKTEEKASALRSELAVARAVAQLADDLTHAADDTLRAAAQAQDEA- VALLRTHHEEDNAELMELRRRAEELRGVLSSTLDRLAQKEELEKQLNDNAGE- LAANNAALTEELATLRQXXXXXXXXXXXXXXXXXXXXXXXXXXRAAAATECEGM- QQRLRTAEENDRTAVAALAEAMAELSDAEARVEALTSENGALQAKTLTAER- DDAQRRGAAAGENERTLKAAVEELAAKLAADQHVDELTAQGTSLAERAQLCS- SERDTRQRVSDAAEELQSVKAALEDLAECSDTKRADALAAQVAGLEGERL- KLAELHAEHVLAEVEVRLADAEHRDSERTAAMEHLSAAQLAQIASLEEQTS- KQKELSTARAAISTMAEHRADTETQISDLQSQLNHASEQRKTEEKASALRSE- LAVARAVAQLADDLTHAADDTLRAAAQAQDEAVALLRTHHEEDNAELMEL- RRRAEELRGVLSSTLDRLAQKEELEKQLNDNAGELAANNAALTEELATLR- QMLAEERLELRLDANRGVEQTAASYMYNERILDMRDLQGLQEDAVTQAFAQ- TEYFKSGLTRVVGMSDGFRRVAAEMAPAAATTVRATAKLAADASSYVFEDCE- GREAGAACSVSAAVAAVKAALTPWSETLDWVAATQRIEELEGSANAGAGH- EDFTASADGAAATTDARSQAGEDSPASSYLSEVQSARSALLDFVEQQTSS- LAHLRRLRHALHERDQAADVQAALAEVAGGVAQERADGFRATEDARHR- ILOAEREWEELQQLKSQATAELIVARRAEERANDARTRAEAQAAAELKE- TQAAALRTTQEEKQLALQTERLSLSQQSDSSALQRRRSGVSDSVTPPTAA- VTEGSRTRTRYVSSSTEAVEQTHLLQVAGLQQLVRCRRQIRELESEATLR- QYMALEQEVSDLRSDAAEVVNLRAEMATLRADYADSEERYERLMLMVDVQ- DGTAQENRRRLRQELKLDAALEEKYTRQENSLVSRGAPDFEQRLLEEAQ- MSGIATQTLAKQAEFGGGHGGNLSASPQDVLSNRDHLQTLVREERDAVSK- MKAADLQSRVCKEGLEELKAEKTAALNGAEELLAQYMERLDAMRAAGLAAV- TTRATTTQPRNTGKSSRERAKRDSALPPTPLSLNVTGDSSTSHKGATPAGLA- MPSPAPMPNLDDVAAGSGEEDEEVEVRKINGSEDDASATPPATSRRTGT- ATGENSSCGNRRTRREVHSSSGTTPPRASPKVKGKTVSTRKRAR</p>
<p>Leptopmo- nas pyrrho- coris</p>	<p>Trypano- somatida</p>	<p>Non-public short-read genome assembly</p>	<p>MWNSDYYTAPPHRVGGTTASSSVSAHSHPPPPPAGVLPDLARPISVRDLA- QTRALAGPPRASAATSAPFSFGSGRDRSDYVPLPRTSGGGFAGFVGP- KAPVYVAATSGLPSWSSQHADSSGGASELSTHGEASDLT- SLAERNIDHLSNDDLNYASHLQDTAKRYSHHLEAEHAKRDMYRNLVAVQKRE- LQEKYVQIDRQHHEQETAALRRLREDNQLREQVATAEGKVRNIAAKLSLQ- NGADPSVAIRYQAQLASKEAQVRDLRMRVHVHSAERQSAAEQRRSS- GANPRQSDTLGAAGTRDHDQQQQLHQWPEEEHPCTIDERVRRELAENA- AAAASEMAKLAELNHLHQRNAALQEQQLAQQDKTHAEALASMTAEHTLKETI- AQLKQGHAEEMQENEDALVAALTRAPYSKTDFAEFQATYAEMQAALTTEATV- ATLQAAQNASAVQDGYIEVLQQLSATSQAERQELTEKNRQLQALSENLRQ- QLADTANRLRQQLELAVAGGEQLAQEVQRTQEARTHVEHAEIVALRCAE- ADAELQRLRAENEELSRQMQDARTQSETNHHEQEVVSADNTRERSALE- ETLNRNARDNLVVRSDLLGELRELEERRQTAERDAPQSTNAQHASETHITV- AAAAASSVQHDRDTAEDTAMLQAAQCCEALERDLETTNSAHARMEVDMQ- VQLRAELTKMAEQADALRSAEQQQKRDAEAYHNRTIAAAAEQKHLLEVIQAL- KLEEFTQTNRELROSEASRGLYSQLRQAELRERVVELEERRASDQV- LTPSTQPASVGTGCGSTTADAPDAKMNPSTMSVAIDRLAQQQLAEAQDRVH- LTERHRQAELQVQVAATLEDRSVSEEHSSKAAGSSATSTRPSAQKGT- RYANGVKEGIVADQVATLIAQQDAAAHQTAQLRDLTLEAAEARTSAVATERD- AAQAAAAAEEVVRARAALKEKMAEELSDSAHRVDEL SADI GELATVKAQAAV- QRRASVAEDEL SAAKAVLDRFTKEVSGVTRCANNVDVMAAVEAQWDAV- QEDASEAELRTAKAALERVVEEASASAQHASSLAAATSOLEADKTRLTSE- AMQRRAAAAEEVVRARAALKEKMAEELSDSAHRVDEL SADI GELATVKAQLA- AVQRRASVAEDEL SAAKAVLDRFTKEVSGVTRCANNVDVMAAVEAQWDAV- QQREDASEAELRTAKAALERVVEEASASAQHASSLAAATSOLEADKTRLTSE- RDAMQRRAAAAAEEVVRARAALKEKMAEELSDSAHRVDEL SAAKACWESEKR- QRDAELARLESTAAEWETRTADVERRSAAEAAAASEKRIATLAAEAALEKETV- RLKARL SAARGAVSLSAEDCTDVNTQDGLTRHVERVEQQRKSAEAMLSALK- RELVTQAQAVVQLTEDELMRADNDAAMRATLKAEFEAFALLKSERDAAQQA- AAAAEEVVRARAALKEKMAEELSDSAHRVDEL SAAKACWESEKRQORDAEL- ARLESTAAEWETRTADVERRSAAEAAAASEKRIATLAAEAALEKETVRLKARL SA- ARGAVSLSAEDCTDVNTQDGLTRHVERVEQQRKSAEAMLSALKRELVTQAQAV- VVLTEDELMRADNDAAMRATLKAEFEAFALLKSERDAAQRRAAAAAEEVVR- ARAALKEKMAEELSDSAHRVDEL SAAKACWESEKRQORDAELARLESTAAEW- ETRTADVERRSAAEAAAASEKRIATLAAEAALEKETVRLKARL SAARGAVS- LSAEDCTDVNTQDGLTRHVERVEQQRKSAEAMLSALKRELVTQAQAVVQLTE- DELMRADNDAAMRATLKAEFEAFALLKSDHQDDRAELSEVRRRGEELQRLD- TDLTDLDRLEEEELVKGFLNDAADLHAERAALETVEVVTVRLALETVLERQDIERD- AAKLGLEQTAEYNERILEDDRLQDELDAVTQAFSAQTMQVQKQMRHMLQ- MSAGFDALAEVMAALAAATVAAAADLAASCTAASKTMTGSTERKPAQATQD- KFRAVTAIQAALTPWSAAELWVATMQRVVALEGGSSADTSEDPRIISAALSDG- ANDPSHMPLSVSLAAASTVADAARKARGALDVFVLEALLTHLRLGHALDER- DYLRLRLSAVDEMANSVAEKERADHLLAEIQDAKHNHVSLAEKFEFDDQL- LRSEAAVELVARRAEERATDARLRAESQAAVGAELTDALALHKTTEEERKM- SREVERLSRFSARLADPGARVHSGGTVGSVTPPSSGLFAEGVSNHYTMTASA- VEALEQAHLQVSSLQELLSRQIRELEROEVLVIRANVTLEQQVLDLRAD- AEEAVDLREEVARLRAENADLEEKIEQLERILAAASGGGVAHEVQLRQQLK- LREAELEEEKARMQGLLVSKTAEFAAARRQAALRESLGGITTLAATAQAGVSSG- RSRPGSLSAALSSNSAATSAPLDVLSKRIIDLQAMKERDALKEMKAQQLQ- SHVTVDGLERLSEKAAALERANLVTHYAEALDGLRAARVSAARTLPSRGN- HSSSSAAQSPPPAAPPTPLALQTSVDSNVKGGVGDSPQHMTAPVSPAP- RMPHLAAVEAAEVEGSLTPPATSRRTGTGESSRDRGRRSRGTRSTSS- ASSAHRSPASSGTREGASHRKRTR</p>

<p>Endotrypanum mon-terogeeii</p>	<p>Trypano-somatida</p>	<p>NCBI AOF- S01000490</p>	<p>MWVNSSASPSGQQGSSLLCDERSPALCGAGNRLSSYSAPPAHSLPAGVLSAD-LARPISVRELAQSRVPVPPVPAVAAAAPSMPQRPCDSSSTAGFS-DASKQFTAHLQEAYAKRDQYRSEVAHLKQELQSRVVEVDLLRRERERASKA-LLGVHEENTTLREQLAEVEGRLRNMQAKMSVLSDDSSLALHFYQSQLTK-EAQMRELQQDYERVAASLQRCSTSTAGAA SARHSINIPVEDGAATECSLS-RDPLNSGGGQDSEEVAVELRATLEELRSLNAGLSHRLEEEERAAHDAALAAH-TEAIAAAAADRTELLDITIFQLRQRCTEMQASEDELVAALTQRAPISKSDYAVF-QSSYNDVTAALAKAEAQITELREKERHEYASARQAQEEHLHSSALGMEEKN-TLQARNEQMQLSVHHLKQELAEANQLRQKQLDAAVNGNQGGLANNLRM-AQQQLSMSSEVEDLRNTRTNLEDDQVNELRHMDEFENDDRLQRRSDEQ-RGDSQTDVDTRTASPLRESDEAVERSALKNELATLRAQRDALLRRDRDAIL-ELRQLNETRHPGKLLSPSQMERPTDSSDSRNTTDPRENERRGVFRQPEEV-AALEEARQCIERMNDTISTVQLEHLRVENNYKLSALEAELESQSRALAEKD-AAGVAQLAHLEEVIMMALRDELVATQTEKELRASEDAQRVRIAEQQHAVEQL-QAEVASLTTENADLKEKLSSTRAAVTTLASRQTDVEGAAVSLLEERCASEVQR-REAAEAEEVRLREALDAANVTAAGIEHLASLEAAREEIAEMQANRDRVEAE-LESQSRALAEDKAAGVAQLAHLEEVIMMALRDELLATQTEKELRASEDAQRV-RIAEQQHAVEQLQAEVASLITENADLKEKLSSTRAAVTTLASRQTDVEGAAVS-LEERCASEVQRREAAEAEEVRLREALDAANVTAAGIEHLASLEAAREEIAE-MQANRDRVEAELESQSRALAEDKAAGVAQLAHLEEVIMMALRDELVATQTE-KELRASEDAQRVRIAEQQHAVEQLQAEVASLTTENADLKEKLSSTRAAVTTL-ASRQTDVEGAAVSLLEERCASEVQRREAAEAEEVRLREALDAANVTAAGIEH-LASLEAAREEIAEMQANRDRVEAELESQSRALAEDKAAGVAQLAHLEEVIMM-ALRDELVATQTEKELRASEDAQRVRIAEQQHAVEQLQAEVASLITENADLKEK-LSSTRAAVTTLASRQTDVEGAAVSLLEERCASEVQRREAVEIAAVKEALKH-REGELSEVLVRGEAVQALDEALDRLSEEEVLEELNGDAAQHSKEMKALA-ESVAMLKQSLSTAIVDKEDLSAQLVATQKEVDRAIGQDVLLEAERLKDLE-ATTSFAAQTREVRGMAHIVDMSNVFCGVLEVVAGAASARGADGMEQE-ETATSRGSAAPIGGRCTTAKIGEALSPWDEALEWVATMEQRLEVLVLDGAI-ITSGLNDGGGRPARRRRRRRSRAVTEGHEAIEAEDGCQECGLSASFMCRA-VPRSRAEATLARLFCEQASAKRTIADLRVSLAEKDVELGNVRFAMDALASEA-ITERDRADNLSAIDDATKIIVSQAEFEELKQRMATTAEVVEARAEARA-TASRLCAEEQLSAVEDDMKEQALVRLQDENRRLTHDVRLLRSLRLEG-SSRVAEEVRESVTPASLSPAVSARIASQPNTSPSIAETIEQAQILQVSSQLA-DLMRTRRQTRDLKGREATLQSCATLEKELSELPRAREAEQLRGDLDALR-ADYAELELRHDQLERMATAAGDGTMQELKQLRQQLKVRAELEELKFRD-LLSQATPELEVARRQALRESLTDITHLVATQAQYRAGSEDSGAIRSRRS- LAMSCRGSASLTPPPSGPSATDVLTRDMNYLQALVRDREVELEKVVAAQLE-SRAATDTRDQLVAKAAALERAELVAQYTDINTLTTTAMRGGGECSGRLN-TTPSLSNEDALNTKVSSPSVPLNQLTPRRLSATLEAYTSVAMASPVPLAP-CLDEAVKIHENGDMRKNVGP SADNEVLPATSRRTGGTNTQRSTRGSSVRRS-SLSTNRVSSASSSPRGSTGKVSAGTSSRKRST</p>
<p>Leishmania enriettii</p>	<p>Trypano-somatida</p>	<p>NCBI ATAF01000 382</p>	<p>MWDNSVYKAGAHQQSSLRPDGLPIFSSIGTALSPHSHPARLARPPPAVLSAD-LARPISVRELAQSRVPVPPVPAVAAAAPSMPQRPCDSSSTAGFS-DASKQFTAHLQEAYAKRDQYRSEVAHLKQELQSRVVEVDLLRRERERASKA-LLGVHEENTTLREQLAEVEGRLRNMQAKMSVLSDDSSLALHFYQSQLTK-EAQMRELQQDYERVAASLQRCSTSTAGAA SARHSINIPVEDGAATECSLS-RDPLNSGGGQDSEEVAVELRATLEELRSLNAGLSHRLEEEERAAHDAALAAH-TEAIAAAAADRTELLDITIFQLRQRCTEMQASEDELVAALTQRAPISKSDYAVF-QSSYNDVTAALAKAEAQITELREKERHEYASARQAQEEHLHSSALGMEEKN-TLQARNEQMQLSVHHLKQELAEANQLRQKQLDAAVNGNQGGLANNLRM-AQQQLSMSSEVEDLRNTRTNLEDDQVNELRHMDEFENDDRLQRRSDEQ-RGDSQTDVDTRTASPLRESDEAVERSALKNELATLRAQRDALLRRDRDAIL-ELRQLNETRHPGKLLSPSQMERPTDSSDSRNTTDPRENERRGVFRQPEEV-AALEEARQCIERMNDTISTVQLEHLRVENNYKLSALEAELESQSRALAEKD-AAGVAQLAHLEEVIMMALRDELVATQTEKELRASEDAQRVRIAEQQHAVEQL-QAEVASLTTENADLKEKLSSTRAAVTTLASRQTDVEGAAVSLLEERCASEVQR-REAAEAEEVRLREALDAANVTAAGIEHLASLEAAREEIAEMQANRDRVEAE-LESQSRALAEDKAAGVAQLAHLEEVIMMALRDELLATQTEKELRASEDAQRV-RIAEQQHAVEQLQAEVASLITENADLKEKLSSTRAAVTTLASRQTDVEGAAVS-LEERCASEVQRREAAEAEEVRLREALDAANVTAAGIEHLASLEAAREEIAE-MQANRDRVEAELESQSRALAEDKAAGVAQLAHLEEVIMMALRDELVATQTE-KELRASEDAQRVRIAEQQHAVEQLQAEVASLTTENADLKEKLSSTRAAVTTL-ASRQTDVEGAAVSLLEERCASEVQRREAAEAEEVRLREALDAANVTAAGIEH-LASLEAAREEIAEMQANRDRVEAELESQSRALAEDKAAGVAQLAHLEEVIMM-ALRDELVATQTEKELRASEDAQRVRIAEQQHAVEQLQAEVASLITENADLKEK-LSSTRAAVTTLASRQTDVEGAAVSLLEERCASEVQRREAVEIAAVKEALKH-REGELSEVLVRGEAVQALDEALDRLSEEEVLEELNGDAAQHSKEMKALA-ESVAMLKQSLSTAIVDKEDLSAQLVATQKEVDRAIGQDVLLEAERLKDLE-ATTSFAAQTREVRGMAHIVDMSNVFCGVLEVVAGAASARGADGMEQE-ETATSRGSAAPIGGRCTTAKIGEALSPWDEALEWVATMEQRLEVLVLDGAI-ITSGLNDGGGRPARRRRRRRSRAVTEGHEAIEAEDGCQECGLSASFMCRA-VPRSRAEATLARLFCEQASAKRTIADLRVSLAEKDVELGNVRFAMDALASEA-ITERDRADNLSAIDDATKIIVSQAEFEELKQRMATTAEVVEARAEARA-TASRLCAEEQLSAVEDDMKEQALVRLQDENRRLTHDVRLLRSLRLEG-SSRVAEEVRESVTPASLSPAVSARIASQPNTSPSIAETIEQAQILQVSSQLA-DLMRTRRQTRDLKGREATLQSCATLEKELSELPRAREAEQLRGDLDALR-ADYAELELRHDQLERMATAAGDGTMQELKQLRQQLKVRAELEELKFRD-LLSQATPELEVARRQALRESLTDITHLVATQAQYRAGSEDSGAIRSRRS- LAMSCRGSASLTPPPSGPSATDVLTRDMNYLQALVRDREVELEKVVAAQLE-SRAATDTRDQLVAKAAALERAELVAQYTDINTLTTTAMRGGGECSGRLN-TTPSLSNEDALNTKVSSPSVPLNQLTPRRLSATLEAYTSVAMASPVPLAP-CLDEAVKIHENGDMRKNVGP SADNEVLPATSRRTGGTNTQRSTRGSSVRRS-SLSTNRVSSASSSPRGSTGKVSAGTSSRKRST</p>

<p>Leishmania braziliensis</p>	<p>Trypano- somatida</p>	<p>NCBI XP_001565 950</p>	<p>MWGNDFYKAGAHQQPSLRHDGLPTSFASPSYTAQPARPLPAGVLSAD- LARPISVRELAHSRVPVPPPPAVPPARRNPAMTGFAGFGIPTSPIIYAATS- GVSTLPPSTRADASDLMSLAHRDIDNLSGDDLYNYASHLQGGASKQYSKHL- LEAYTRRDQYRGEVARLKEELHSKYLQIDVLRREQERASEALVKVREDNA- QLHQKLAESSEGHVVRNMQAKVSVLSGADPSAAVHFYQSQVLKDAQLRE- LQQQCEREAAAASSQRHSSSATAAEHGRRSINVSGEAAVPOISPLRDLPL- STRKVHNGSEAASTALHTALEELQHLNLSLRRLEEEHAAHKAALAVHAE- ENATADRAELLNTISQLRQCTEMQSTDELVAALTRAPISKEGYAAL- QTSYNDLTAALAKAEDQIAALHVKEQQYHQSQAQQAEELRHSLTTGLEE- RQELQARNEQMQLATDLQRELELAERANQLRQEQLDAAVSGNQQLAE- HLRTTQEQQLGATAELQDLNTRGSLRQELRQRREEFQEQDQRRRQ- QHRDVEGGATFSYTPAASAESETVKTREALEDELATLRSQRDALVGDRE- IISEVRRQLQEDRLALLRSLPQAEVVLQKHLSTVKATLGHADDEELPQEE- DIAALEAARHCIEQLKETIATLQSERVRVEAELQAKIKVLEAELASQSRALS- DAEDAAAQRAHLEGIVVALRDELVATQEQMQFITSEEAQRERCAEQQ- RAVEHLQAVQVQLQMGQDSANSLATTWHAQSPVPAKSDLSTMNAALE- SLAHQLEDAQYRAEELSAEHARTLEELARATRAAAELEARAAEAESRRE- AAHAAAEGTEVELTARIAALATENAQLKEKLSASRAAVTTFASKQTDGEST- AASLEERCATEARRREAAEAELVRETLDRTGAEELTEVLRQSEGVQSTL- TRTVDRLAEQEALVGVLAGENMRLQDELTAVERISILLTSDERDALVEE- TAQ- LHGKLAASERMSDAAAAQQRLVRELEEAAVRSKAVELGTVFAALDQ- TASEVCGAEQQLRVYAEAKEELSAAEHARTLEELARATRAAAELEARAAEA- ESRREAAHAAAEGTEVELTARIAALATENAQLKEKLSASRAAVTTFASKQTT- DGESTAASLEERCATEARRREAAEAELVRETLDRTGAEELTEVLRQSEGV- QSTLTRTVDRLAEQEALVGVLAGENMRLQDELTAVERISILLTSDERDALV- EETAQ- LHGKLAASERMSDAAAAQQRLVRELEEAAVRSKAVELGTVFAALDQ- TASEVCGAEQQLRVYAEAKEELSAAEHARTLEELARATRAAAELEARAAEA- ESRREAAHAAAEGTEVELTARIAALATENAQLKEKLSASRAAVTTFASKQTT- FASKQTDGESTAASLEERCATEARRREAAEAELVRETLDRTGAEELTEV- LRQSEGVQSTLTRTVDRLAEQEALLEQLNRNASELNDARAVLTGEVTR- LQ- AALAAVTDKSDVMAQLVATQRELDRAVEQQDAQYCEEVRLQDALTTT- M- SAFVGQTRQVQLGMAHVVDMSKAFYRVVEVVAKGAASATVQGEIVDKE- AAAQSFASVGVGSSVKGASPTATSVREILLPWSRALEWVWASMEERLES- L- GTGSTSSDTTSSDSQCERRRRRRGRSTVADEEEEAIDEEETSGADAVST- SFLCRAVPASHAAEFARWIWEQAAAASTIADLRSALADKEVELNLRSA- M- NDTLANDAMAEQDRADNLSAIGDATEKIALVQAEFEELRQRQVATTA- E- VVEARRAERARAAQLRAEEHLLTVENELKEQQAVLRKQDENHRLTRE- AD- LRLRRSSRHLESDSRALNKISDSVTPQPLSYAGVSAVTEVESFTAAS- A- VEVVEQAQFLQVFLQEDLMLSRQARELEGREATLRQTCATLEQQVSO- Q- LREATDAEKLREELATLRDYAELEERYDQLERMTTAAGGGTMQELKQ- L- RQQLKARETELDELKARVRLVLSKVAPELEAARRQELRESLGGITTT- Q- LAATQAQYGGGTGSSNAGSRRASLASALSKPDHLSLSSSPSSQDLT- S- RIEHLQILVHDHQEAFKVAQALRESRATMDTLEDQLAAKTAALERAEG- L- VEYADTINVLTMTAVVGRSGGAHLSTVQPLRNEVANLVAEVPSSLVPLT- P- PLTPRPVADAGQPRAAVAVASPTPRMPNLDEVAANNDGSVTDVEGDS- A- GSSADTALPSATHRRRTDAEGSRRSARGTPARRSSSSSSNNNTARKSFVGT- M- IAKARKRARSSTTA</p>
<p>Leishmania mexicana</p>	<p>Trypano- somatida</p>	<p>NCBI XP_003876 756</p>	<p>MSGSRSTLSSPSYAAQPARPPPAGVLSADLARPISVRELAQSRVPVPPPPATQT- PRPSATTAGFAGFGVPTSSAVYAATSGVPTLQLLSTHTDAGDLTSLAHRD- VDHLS- GNELNYASHLQESSKQYSGVLYSEAYAKRDQYRGEVARLKEELHNRMQIDVLR- REQARATDALVKVREDNAQLRQKLAEEAGHTRNMQAKMSVLSGADPS- IAVHFYQSQALKDAQLRELQECERAAAAASSQRSSSTTAAEHGRRS- ANISGEGAAAPHTSPSCDLPSTEELQERAEAAASALPTALEELQRINISL- SHRLEEEHAAHKAALAAHAEASATRADRAELLDITIAQLRQCCADMQSA- EDELVAALTRAPISKDYAALQASYDDLTAALAKAEQVAALQEQEQH- DHIAQAEFEELRHLLTTQEEERQSMQARNEHMQSVLYHLELELAE- R- ANQLRQEQQLDAAISGNQQLAESLHVAQEQLLSATAELQDLRDTCE- S- LQRELQRIEELQADQHQHGHGDAEAERTLIHTSAAASEALQ- Q- RQELAEALASLCSQRDALVGDREDDIGEMRRLODERLALLTSPPAQAE- L- QORRSVAVKAPQDVRADGGEWQAQEDIAALEAAHQCEQMKTIAAL- Q- SESARVEADLKAENVTLKAEASQSRALTGAQAAAAAAQAHLEEVMLA- L- RDELVATQAEKRLRASEEAHRERCVEQQRVMEQLQHQVVEELERAA- R- DAASSLSTPPDQPTTAAESELAMNAALDSLANELEDAQLRAEQ- L- AEHARTVEELARGTRAVAELEARAADTEGKREAMHTAAEATEAELMAR- V- VAALAAENAQLKEKLSRAAVTTFASRQTDGESTAAPLEERCAAAARR- S- SEAAEAASALREALSRSAELVEVLQRSEVDQATLERTVARLAEQESL- V- GVLSAESVRYQDELASAQRVLMMLTSDERDAGEAAGLHGKLSAAD- R- ASDAAAQKRLQMLREEAEAARVSKVDELSTFAALDRTASELCAE- D- AQLQAYSEAKEQLAAEHARTVEELARGTRAVAELEARAADTEGKREAM- H- TAAEATEAELMARVAALAAENAQLKEKLSRAAVTTFASRQTDGEST- A- AAPLEERCAAAARRSEAAEAASALREALSRSAELVEVLQRSEVDQAT- L- LERTVARLAEQESLGVLSAESVRYQDELASAQRVLMMLTSDERDAGE- E- EAAGLHGKLSAADRASDAAAQKRLQMLREEAEAARVSKDVELSTFA- A- AALDRTASELCAEQRLQAYSEAKEQLAAEHARTVEELARGTRAVAE- L- EARAADTEGKREAMHTAAEATEAELMARVAALAAENAQLKEKLSRAA- V- TTFASRQTDGESTAAPLEERCAAAARRSEAAEAASALREALSRSAE- L- VEVLQRSEVDQATLERTVARLAEFEELVEQLNYDASELNVARAALS- D- VTRLQRALAVALTDKSDVVAQLLATQKELDRAAEQDAQYRDEERLQDA- L- LATTTSAFDVQTKQVRLGMAHVVDMSRAFCRVAEVAVQWSAAQTDGL- A- ASTGASVQEMLPWSRALWVAGMEQRLALDIGAAMLNSFSRTDGR- R- RERRRRRRRRGRTTESEDDAAEEEPSSGSNALSLSSLCRVAPANRAEV- A- AFARLMDAQAAAMNTIADLRRLADKEADLNLRSAMDALASDLMGER- D- RADNLSAIDDAASEKIGSAQAEFEELRQRRTAATAEVVEARRAEAR- A- TAAQLRAEKQLAAVENDRKEQAVLRKQDENHRLAREADRLRTSRL- E- PSSRVRSNGSGSVTPKSLSATSAEDAVQRFTTAPSTVEAAEQAIL- Q- VSSLOADLMLSRQARELEGREATLRQACAAEQVNLREVAEAEL- H- LREDLATLRADYAELEQRYSELERMATAAGGGTMQELKQLRQQLKVR- E- EALEELKARMQDLVLSKAAPELEAARRQELRESLGGITTLAATQA- Q- YGGGTGGSNKARSRRASLTDALNSDLLSLSSSRPSPQDVLTSDIEH- L- QTLVHDREVALEKVAQALRESRAAMDMLDQLAAKTAALERSEGLVAE- Y- ADTINTFTMTAVVRRGNHGARPPTAPPLSSGMGSPDAEALSTSVPLT- P- PLTPRTAAVNGQSHAVAMASAPRMPNLDEVADSDGSGVEGKHAD- S- SAGPSTDSAPPVTHRRRTGASGSWRSKSCPAHHPSSSTERSAPSSP- S- STVPGRKSSAGASMPAATRKTRRSSTNA</p>

<p>Leishmania amazonensis</p>	<p>Trypanosomatida</p>	<p>NCBI AP-N- T01003049</p>	<p>MWGNFSYKAGTHQQSSSLCHDGPVTMSGSRSTLSPSPSYAAQPARPPPAGVLSAD-LARPIVRELAQSRVPVPPPPATQAPRPSATTAGFAGFVPTSSAVYAATS-GVSTLQQLSHADAGDLSLAHRDVLHLSGNELYNASHLQESSKQYSGYLSEAYAKRDQYRGEVAHLKEELHNRYMQIDVLRREQARATDALVKVREDNAQLRQKLAEEAGHTRNMQAKMSVLSGADPSIAVHFYQSQLAKDAQLRELQCECERAAAASQRRSSSTAAEHGRRSANTSGEAAAAPHTSPSCDLPSTEELQEREAASA-VLPTALEELQRLNLSLHREELHAAHKAALAAHAEASATADRALLDIAQLRQ-QCADMQSAEEELVAALQRAPISKDYAALQVSYDDLTAALAKAEQVAALQEQHQDHIARQAEELRHLLTTGQEERQSMQTRNEHMQSLVYHLERELAEARANQLRQEQDAAISGNQQLAESLHVQEQQLSATAELQDLRDTRESLESQRLRELQRIEEL-EQADQHQRHQHGDAEERTLIHTSAASAESEALQQRQEELEGLASLRSQRDVLV-GDRDDIISGMRRQLQDERLALLTSPPAQAEELQQRSSAVKAPQDVCADDGEWQA-QEDIVALAAHQICQEMKQSIQALQSEARVEADLKAENVNLEKELASQSRALAGA-QAAAAQQAHLVEVILALRDELVATQAQEKRLRASDEAHRERCEVQQRAMEQLQK-HVEELERAARDTSSFTSTPPDQPSSTAAESELAMNAALDSLAEHLEDAQLRAEQ-LAAEHARTVXXXXXXXXXXXXXXXXXXQATLERTVARLAEQESLVGVLSAESVVRV-QDELAQAQCVLMLTSEADALGEEAAGLHGRLSAAEDRASDAAGAQRKLMQRLLE-EAAAVRSKDVSTAFALDRTASELCEAEQRLQAYSEAKEQLAAEHARTVEELAR-GRTRAVAELEARAADTEGKREAMHTAAEATEAELMARVAALAAENAQLKEKLS-TS-RAAVTTFASQTDGESTAASLEERCAAAARRSEAAEAEASALREALSRSAELVEV-LQRSEDVQATLERTVARLTEEEALVEQLNYDASELNVARAALSDEVTRQLKALVALV-DKSDVVQALLATQKELDRAAELQDAQYRDEERLQDALATTTSAFDVQTKQVRLG-MAHVIDMSRAFRCVAEVTVQVWSAAQTDGLASTGASVQEMLSPWVRALEWVAGM-EQRLEALDIGAMLSFRTDGRRERRRRRRRGRRTTESEEEEDAAEEEPSSGNSLS-LSSL CRAVPANRAEVAFARLMDAQAAMNTIADLRRLTAEKEADLTNLRATDALA-SDLMGERDRADNLSAIDDAEKIASAQAEFEEQLRQRTAATAEVVEARRAEARA-TAAQLRAEKQLAAIENDRKEQEAVALRKLQDENHRLAREADRLRTRLEASSRVRS-NGSGSVTPKLLSLSATSADAVQRFTTAPSTVEAVEQAQILQVSSLQADLMLSRQ-ARELEGREATLRQACAALQVQVNEELRVEAAEAHLREDLATLRADYAELEQRYSEL-ERMATAAGGGTMQELKQLRQQLKVRAEAELEELKVRMRDLVLSKAAPELQAEVKA-EALRELSGITLQATQAQYGGGTGGSNKARSRRASLTDALNSDDLSSSSRPS-PQDVLTSRIEHLQTLVHDREVALEKVVAAQLESAAAMDMLGDQLAAKTAARESEG-LVAEYADTINTFTMTAVVRRGNHGRAPTTAPPLSSGMGSPDAEALSASVPTLPTP-TRAAVNGQSHAVAMTSPAPRMPNLEVADSDGSGVEGKHADSAGPSTDSAPP-PVTHRRTAGSGSWRSKSCPAHRPSSSTERSAPSSPSTVPRKSLAGASMPAAT-RKRTRSSNTA</p>
<p>Leishmania infantum</p>	<p>Trypanosomatida</p>	<p>NCBI XP_001466 421</p>	<p>MWGNFCKAGAHQSSALCCDGPATMSGSRSTLSPSPSYAAQPARPPPAGVLSAD-LARPIVRELAQSRVPVPPPPATQAPRPSATTAGFAGFVPTSSAVYAATS-GVST-RSHADAGDLSLAHRDIDLHLSGSELYNASHLQESSRQYSGYL- SAAYAKRDQYRGEVAHLKEELHNRYVEIDALRREQERASDALLKVRREDNAQLRQK-LAAEAGHARNMQAKMSVLSGADPSIAVHFYQSQLAKDAQLRELQCECERAAAAS-SQRRSSSTAAEHGRRSANISGEAAAAPHTSPSRDLPSEAEELQDGAEAASAHT-LALEELQCLNLSLHREELHAAHKAALAAHAEASATADRALLDIAQLRQQA-DMQSAEDELVAALQRAPISKDYASLQASVDDLTAALAKAEQVAALQEQHQH-HSAQQAQEEELRHSLTTGLEERQSMQARNEQMQSLVNLHLELAEARANQLRQEQ-DLVAISGNQQLAESLHVQEQQLSATSELDQLRDTRESLESQRLRELQRIEELQ-GEQHQHRHGGGAEERTLIHTPAASAESEASQQRQEESEALASLRSQRDVAVSD-RDDIIGEVRRLPDERRALLISPPAHAEELQQRSSAVRAPLDVRAADDREWQKQEDI-AALEAAHQICQEMKQTIQALQSERVVEADLKAENVNLEKELASQSRALAEAAEA-AAAQQAHLVEVMAALRDELVATQAQEKRLRASEEAHRERCAEHQRAVEQLQAHVE-ELERAARDAASSLTPPPAHPSTAAESELAMNAALDSLAYLEDAQHRYKELSAE-HARTVEELARATRAVAELEARAADAEGKHEVMHTAAEVEAELMARVAALTAENA-QREKLSRAAVTTFASQTDGESTAASLEERCAAEARRSEAAEAEASALRETL-NGTAAELAEVLRSEDVQATLERTVARLAEQESLVGVLSAESLRLQDELATQRSV-LLLTSEADALGEEAAGLHGKISAAEDRASDAQAQQLQACLEEAEEAARVRSKDV-ELASTALALDRTASEVCDAEQRLQVYSEAKEQLAAEHARTVEELARATRAVAELEA-ADAEGKHEVMHTAAEVEAELMARVAALTAENAQLREKLSRAAVTTFASQTD-GESEAAASLEERCAAEARRSEAAEAEASALRETLNGTAAELAEVLRSEDVQATL-ERTVARLAEQESLVGVLSAESLRLQDELATAQRSVLLTSEADALGEEAAGLHGKISAA-EDRASDAQAQQLQACLEEAEEAARVRSKDVSTAFALDRTASEVCDAEQRLQV-YSAEKEQLAAEHARTVEELARATRAVAELEARAADAEGKHEVMHTAAEVEAELM-ARVAALTAENAQLREKLSRAAVTTFASQTDGESTAASLEERCAAEARRSEAAE-AEASALRETLNGTAAELAEVLRSEDVQATLERTVARLAEQALVEQLNYDASEL-N-DARAVLTDVTRQLKALAAVSDKADVVAQLLATQKELDRAAEAQDAQYRDEERL-QDALATTTSAFDVQTKQVRLGMAHVVDMSRAFRCVAEVAVQSRASQTDGASG-TSVQEILSPWTRALEWVAAMEQRLEALDRGAMPNSFTRTNGRERRRRRRGRRTT-ESENDDAVEMESNSGSYVLTSPLCRAVPANRAEAFARLIDEQAATNTIADLRRL-READRLRTRLEPSSRARSNVSGSVTPKLSLATSASANAARQFTTAPSTLEAIE-QAQILQVSSLQADLMLSRQARELEGREATLRQACAALQVQVNEELRRAKAAEAHL-REDLATLRADYVELEQRYSELELMATAAGGGTMQELKQLRQQLKVRAEAEVKA-RMRDLVLSKAAPELQAARRQEALERESLGGITLQATQAQYGGGTGGSSKARSRR-DSLTAALNSDDLSSSRPSQDVLNSRIEHLQTLVHDREVALEKVVAAQLESRAA-MDLEDQLAAKTAALERAEGVAEYADTINAFMAAVARRGNRGAVSTVPLSSG-MGSPDAEALSASVPTLPTPRTAAVNGQPHAAVAMVSPAPRMPNLEVADSDG-SGEGKQADSAGPSTDSAPPVTHRRTAGSGSWRGAKSCPAHRPSSSTERSAP-SSPNVPRKSFAGASVPATRRKRSSNTS</p>
<p>Leishmania donovani</p>	<p>Trypanosomatida</p>	<p>NCBI AVPQ01001 269</p>	<p>MWGNFNYKAGTHQQSSALCCDGPATMSGSRSTLSPSPSYAAQPARPPPAGVLSAD-LARPIVRELAQSRVPVPPPPATQAPRPSATTAGFAGFVPTSSAVYAATS-GVST-RSHADAGDLSLAHRDIDLHLSGSELYNASHLQESSRQYSGYL- SAAYAKRDQYRGEVAHLKEELHNRYVEIDALRREQERASDALLKVRREDNAQLRQK-LAAEAGHARNMQAKMSVLSGADPSIAVHFYQSQLAKDAQLRELQCECERAAAAS-SQRRSSSTAAEHGRRSANISGEAAAAPHTSPSRDLPSEAEELQDGAEAASAHT-LALEELQCLNLSLHREELHAAHKAALAAHAEASATADRALLDIAQLRQQA-DMQSAEDELVAALQRAPISKDYASLQASVDDLTAALAKAEQVAALQEQHQH-HSAQQAQEEELRHSLTTGLEERQSMQARNEQMQSLVNLHLELAEARANQLRQEQ-DLVAISGNQQLAESLHVQEQQLSATSELDQLRDTRESLESQRLRELQRIEELQ-GEQHQHRHGGGAEERTLIHTPAASAESEASQQRQEESEALASLRSQRDVAVSD-RDDIIGEVRRLPDERRALLISPPAHAEELQQRSSAVRAPLDVRAADDREWQKQEDI-AALEAAHQICQEMKQTIQALQSERVVEADLKAENVNLEKELASQSRALAEAAEA-AAAQQAHLVEVMAALRDELVATQAQEKRLRASEEAHRERCAEHQRAVEQLQAHVE-ELERAARDAASSLTPPPAHPSTAAESELAMNAALDSLAYLEDAQHRYKELSAE-HARTVEELARATRAVAELEARAADAEGKHEVMHTAAEVEAELMARVAALTAENA-QREKLSRAAVTTFASQTDGESTAASLEERCAAEARRSEAAEAEASALRETL-NGTAAELAEVLRSEDVQATLERTVARLAEQESLVGVLSAESLRLQDELATAQRSV-LLLTSEADALGEEAAGLHGKISAAEDRASDAQAQQLQACLEEAEEAARVRSKDV-ELASTALALDRTASEVCDAEQRLQVYSEAKEQLAAEHARTVEELARATRAVAELEA-ADAEGKHEVMHTAAEVEAELMARVAALTAENAQLREKLSRAAVTTFASQTD-GESEAAASLEERCAAEARRSEAAEAEASALRETLNGTAAELAEVLRSEDVQATL-ERTVARLAEQESLVGVLSAESLRLQDELATAQRSVLLTSEADALGEEAAGLHGKISAA-EDRASDAQAQQLQACLEEAEEAARVRSKDVSTAFALDRTASEVCDAEQRLQV-YSAEKEQLAAEHARTVEELARATRAVAELEARAADAEGKHEVMHTAAEVEAELM-ARVAALTAENAQLREKLSRAAVTTFASQTDGESTAASLEERCAAEARRSEAAE-AEASALRETLNGTAAELAEVLRSEDVQATLERTVARLAEQALVEQLNYDASEL-N-DARAVLTDVTRQLKALAAVSDKADVVAQLLATQKELDRAAEAQDAQYRDEERL-QDALATTTSAFDVQTKQVRLGMAHVVDMSRAFRCVAEVAVQSRASQTDGASG-TSVQEILSPWTRALEWVAAMEQRLEALDRGAMPNSFTRTNGRERRRRRRGRRTT-ESENDDAVEMESNSGSYVLTSPLCRAVPANRAEAFARLIDEQAATNTIADLRRL-READRLRTRLEPSSRARSNVSGSVTPKLSLATSASANAARQFTTAPSTLEAIE-QAQILQVSSLQADLMLSRQARELEGREATLRQACAALQVQVNEELRRAKAAEAHL-REDLATLRADYVELEQRYSELELMATAAGGGTMQELKQLRQQLKVRAEAEVKA-RMRDLVLSKAAPELQAARRQEALERESLGGITLQATQAQYGGGTGGSSKARSRR-DSLTAALNSDDLSSSRPSQDVLNSRIEHLQTLVHDREVALEKVVAAQLESRAA-MDLEDQLAAKTAALERAEGVAEYADTINAFMAAVARRGNRGAVSTVPLSSG-MGSPDAEALSASVPTLPTPRTAAVNGQPHAAVAMVSPAPRMPNLEVADSDG-SGEGKQADSAGPSTDSAPPVTHRRTAGSGSWRGAKSCPAHRPSSSTERSAP-SSPNVPRKSFAGASVPATRRKRSSNTS</p>

<p>Leishmania tropica</p>	<p>Trypanosomatida</p>	<p>NCBIJ compound of: ATAT01001 197 and ATAT01001 196</p>	<p>MWGNDFHQAGTHQQSSSLCCDQGPVTMPGSRSTLSPSPAYAAQPARPPAGVLSAD-LARPIVRELAQSRVPVPPPPATQTPRRPRATTAGFAGFGVPTSSAVYAATS VGVST-SQPLSAHADAGDGLTSLAHRDIDRLSGRELYNYASHLRESSQQYGGYLSEAYAKRDQYRGEVARLKEELHNRYMEIDVLRQQERASDALLRVRENAQLRQKLAEEGHARNMQAKMSVLSGADPSLAVHFYQSQLAKDAQLRELQQECERAAASSQRSSSTTAAEHGRRSANISGGCAAAPHTSPSCDLPAAEELQDGTAEASA-LHTALEELQCLNLSHRLDEHTAHKAALAAHAEASATATADRAELLDIAQLRQCCADMQNAEDELVAALTLRAPISKDYAALQASYDDTLAALAKAEQAALQKEQEQYHHSAAQQAEELRHSLTTGLEERQSMQVRNEQMQSLVNHLELELAERANQLRQEQQLDAAISGNQQLAESLHVAQEQLLSATSELQDLRDTRESLESQRLRQRIELEQQDQHQRHQHGDAEAESEASQQRQQLAEELARLRSQRDAPVGNRDDIIGDVR-LRPEERSALLTSTPAHAEELQSPSSAVRAPLDVDRADNGEWQKQEDIAALEAHRICIEQMKTIAALESARMVEADLKAENVNLEAEASQSRALAEAEAAAAQAHLLEE-AMALRDELVATQAQEKLLRASEEAHRERCAEQQRAVDQLQAHVEELERAGRN-ASLSTPPPAQATAAESELIAMNAALNSLAYQLEDAQHRVEELSAEHARTVEELARATRAVAELARAAGVEGKREVHTAAEAELMARVAALTTENAQLKEKLSRAAV-TTFAXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXVLRSEDVQATLERTVARL-EEEEALVEQLNFDASELNDARAALTEEVTRLQKALAAAVTDKADVVAQLATQKELD-RAAEVQDAQYRDEERLQDALATTTSAFVDTKQVRLGMAHVVDMSRAFCRVAEV-AVQRSAAQMDGAASGASVQEMLSPWTRALEWVAAMEQRLEALDMGAMSNCF-RTDGRERRRRRRGRATELEDNNAVEMEANSQSHALSSSLCRAVSANCAGAAF-ARLMDQAAANTIAIDLRRTLADKEVDLTLNRSAMDALASDLMEQDRADNLAAI-DDASEKIASAQAEFEELRQRTAATAEVVEARRAEARATAAQLRAEEQLAAVEKD-LKEQEGVLRKLDENHRLTREADRLLRTRLEPSSRVRNSVSGSVTPKSLSLSAAS-AEDAAGRFTTAPSTLEAVEQAQILQVSSLQADLLSRRQARELEGREATLRQACAL-LEQQISELRVAAEAHLREDLRLADYAELEQRYSELERMATAAGGTMQELKQ-LRQQLKVRVELEELKARMDLVSKAAPPELAARRQELRESLGGITTLAATQAQYGGG-TGGSSKVRSRASLTAALNSDLSLSSSRPSPQDVLTSRIEHLQTLVNDREVALEK-VQAAQLESRAAMDLEDQLAAKTAALERAELVAEYADRINNAFTMAAVVRR-SSHGAGPTTVPPLSSGMGSPDAEALSASVPLTPTLPRTRAAVNGQPHAAVAMV-SPAPRMPNLDVADDSDGSGEGNQAGSAGPSTDSAPPPVTHRRTRASGSWRDA-KSCPAHRPSSSTERRAPLSPSNVPGRKSFAGASVPATTKRARSSNTA</p>
<p>Leishmania aethiopia</p>	<p>Trypanosomatida</p>	<p>NCBIJ compound of: AUM-B01001107 and AUM-B01001108</p>	<p>MWGNDFHQAGTHQSSSLCCDQGPVTMPGSRSTLSPSPYAAQPARPPAGVLSAD-LARPIVRELAQSRVPVPPPPATQTPRRPRATTAGFAGFGVPTSSAVYAATS VGVST-SQPLSAHADAGDGLTSLAHRDIDRLSGRELYNYASHLRESSQQYGGYLSEAYAKRDQYRGEVARLKEELHNRYMEIDVLRQQERASDALLRVRENAQLRQKLAEEGHARNMQAKMSVLSGADPSLAVHFYQSQLAKDAQLRELQQECERAAASSQRSSSTTAAEHGRRSANISGGCAAAPHTSPSCDLPAAEELQDGTAEASA-LHTALEELQCLNLSHRLDEHTAHKAALAAHAEASATATADRAELLDIAQLRQCCADMQNAEDELVAALTLRAPISKDYAALQASYDDTLAALAKAEQAALQKEQEQYHHSAAQQAEELRHSLTTGLEERQSMQVRNEQMQSLVNHLELELAERANQLRQEQQLDAAISGNQQLAESLHVAQEQLLSATSELQDLRDTRESLESQRLRQRIELEQQDQHQRHQHGDAEAESEASQQRQQLAEELARLRSQRDALVSNRDIIGDVR-LRPEERSALLTSTPAHAEELQSPSSAVRAPLDVDRADNGEWQKQEDIAALEAHRICIEQMKTIAALESARMVEADLKAENVNLEAEALQSRALAEAEAAAAQAHLLEE-AMALRDELVATQAQEKLLRASEEAHRERCAEQQRAVNLQAHVEELERAGRN-ASLSTPPPAQATAAESELIAMNAALNSLAYQLEDAQHRVEELSAEHARTVEELARATRAVAELARAAGVEGKREXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXGLNRTA-TELEAVLRRSEDVQATLERTVARLEEEALVEQLNFDASELNDARAALTEEVTRLQ-KALAAVTDKADVVAQLATQKELDRAAEQAQYRDEERLQDALATTTSAFVDTK-QVRLGMAHVVDMSRAFCRVAEVAVQRSAAQMEGAASGASVQEMLSPWTRALEW-VAAMEQRLEALDMGAMSNCFRTDGRERRRRRRGRATELEDNNAVEMEANSQSHALSSSLCRAVSANCAGAAF-ARLMDQAAANTIAIDLRRTLADKEVDLTLNRSAMDALASDLMEQDRADNLAAI-DDASEKIASAQAEFEELRQRTAATAEVVEARRAEARATAAQLRAEEQLAAVEKD-LKEQEGVLRKLDENHRLTREADRLLRTRLEPSSRVRNSVSGSVTPKSLSLSAAS-AEDAAGRFTTAPSTLEAVEQAQILQVSSLQADLLSRRQARELEGREATLRQACAL-LEQQISELRVAAEAHLREDLRLADYAELEQRYSELERMATAAGGTMQELKQ-LRQQLKVRVELEELKARMDLVSKAAPPELAARRQELRESLGGITTLAATQAQYGGG-TGGSSKVRSRASLTAALNSDLSLSSSRPSPQDVLTSRIEHLQTLVNDREVALEK-VQAAQLESRAAMDLEDQLAAKTAALERAELVAEYADRINNAFTMAAVVRR-SSHGAGPTTVPPLSSGMGSPDAEALSASVPLTPTLPRTRAAVNGQPHAAVAMV-SPAPRMPNLDVADDSDGSGEGNQAGSAGPSTDSAPPPVTHRRTRASGSWRDA-KSCPAHRPSSSTERSAPLSPSNVPGRKFFAGASVPATTKRARSSNTA</p>
<p>Leishmania major</p>	<p>Trypanosomatida</p>	<p>NCBIJ XP_003721 989</p>	<p>MWGNDFHQAGTHQSSSLCCDQGPVMPSSHTSLPFPSPYGEQPARPPAGVLSAD-LARPIVRELAQSRVPVPPPPATQTPRRPRATTAGFAGFGVPTSSAVYAATS VGVST-SQPLSAHADAGDGLTSLAHRDIDRLSGRELYNYASHLRESSQQYGGYLSEAYAKRDQYRGEVARLKEELHNRYMEIDVLRQQERASDALLRVRENAQLRQKLAEEGHARNMQAKMSVLSGADPSLAVHFYQSQLAKDAQLRELQQECERAAASSQRSSSTTAAEHGRRSANISGGCAAAPHTSPSCDLPAAEELQDGTAEASA-LHTALEELQCLNLSHRLDEHTAHKAALAAHAEASATATADRAELLDIAQLRQCCADMQNAEDELVAALTLRAPISKDYAALQASYDDTLAALAKAEQAALQKEQEQYHHSAAQQAEELRHSLTTGLEERQSMQVRNEQMQSLVNHLELELAERANQLRQEQQLDAAISGNQQLAESLHVAQEQLLSATSELQDLRDTRESLESQRLRQRIELEQQDQHQRHQHGDAEAESEASQQRQQLAEELARLRSQRDAPVGNRDDIIGDVR-LRPEERSALLTSTPAHAEELQSPSSAVRAPLDVDRADNGEWQKQEDIAALEAHRICIEQMKTIAALESARMVEADLKAENVNLEAEALQSRALAEAEAAAAQAHLLEE-AMALRDELVATQAQEKLLRASEEAHRERCAEQQRAVNLQAHVEELERAGRN-ASLSTPPPAQATAAESELIAMNAALNSLAYQLEDAQHRVEELSAEHARTVEELARATRAVAELARAAGVEGKREVHTAAEAELMARVAALTTENAQLKEKLSRAAV-TTFAXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXVLRSEDVQATLERTVARL-EEEEALVEQLNFDASELNDARAALTEEVTRLQKALAAVTDKADVVAQLATQKELD-RAAEVQDAQYRDEERLQDALATTTSAFVDTKQVRLGMAHVVDMSRAFCRVAEV-AVQRSAAQMDGAASGASVQEMLSPWTRALEWVAAMEQRLEALDMGAMSNCF-RTDGRERRRRRRGRATELEDNNAVEMEANSQSHALSSSLCRAVSANCAGAAF-ARLMDQAAANTIAIDLRRTLADKEVDLTLNRSAMDALASDLMEQDRADNLAAI-DDASEKIASAQAEFEELRQRTAATAEVVEARRAEARATAAQLRAEEQLAAVEKD-LKEQEGVLRKLDENHRLTREADRLLRTRLEPSSRVRNSVSGSVTPKSLSLSAAS-AEDAAGRFTTAPSTLEAVEQAQILQVSSLQADLLSRRQARELEGREATLRQACAL-LEQQISELRVAAEAHLREDLRLADYAELEQRYSELERMATAAGGTMQELKQ-LRQQLKVRVELEELKARMDLVSKAAPPELAARRQELRESLGGITTLAATQAQYGGG-TGGSSKVRSRASLTAALNSDLSLSSSRPSPQDVLTSRIEHLQTLVNDREVALEK-VQAAQLESRAAMDLEDQLAAKTAALERAELVAEYADRINNAFTMAAVVRR-SSHGAGPTTVPPLSSGMGSPDAEALSASVPLTPTLPRTRAAVNGQPHAAVAMV-SPAPRMPNLDVADDSDGSGEGNQAGSAGPSTDSAPPPVTHRRTRASGSWRDA-KSCPAHRPSSSTERSAPLSPSNVPGRKFFAGASVPATTKRARSSNTA</p>

Leishmania turanica	Trypanosomatida	NCBI compound of: AT-BU0100126 9 and AT-BU0100127 1	MRRSEPGDAHTPHGQNRGEEEDKGGELQRSISIGIESD-SIHRAGHGLRMWGNFHNAGTHQQSSSLCCDGGPQAIMPGSRTSLPSPSHAAQ-PARPPPPAGVLSADLARPIVRELAQSRVAVPPPPPLATQT-PQPRATTAGFAGFVPTSSAVYAATS-SGVSTSRPLSAHADAGDLTSLAYRIDHDHLS-GRELYNYASHLRESSQQYGGYLSEAYAKRDQYRGEVARLKEELHNRYMEIDVLR-REQERARDALLKQEDNAQLRQKLAEEAEGHARNMQAKISVLSGADPSLAVHFYQ-SQLALDKDAQRELQQRERAAAAASSQRRSSSTTAAEHGRRSANISGEGAAAPHT-SPKCNLPSAEELQDQTEAASAAALHTALEELQCLNTSLSHRLEDEHTAHKAAALAAH-AEASATATADRAELLDTIAQLRQCCADMQNVDELVAALTRAPISKDYAALQAS-YDDLTAALAKAEQAVALQEKEQHLHSAQQAQEEELRRSLTTGLEERQSMQVR-NEQMOSFVNHLERELKLAERANQLRQEQLYAAISGNQQLAESLHVAAEQQLSATS-ELQDLRNTRESLESQRELHQRIEELQGDGHQRHQHGDADAEASASSQRRQEL-EVELARLRSQRDPLVGDGDDIVSDVRRLLPDERRALLTSPPAHAEELQQRSSAV-RAPLDVRAADDGEWQKQEDIAALEAAHRCIEQMKQTTIAALQSERVVEADLKAKV-NTLEAELASQSRALAEAAAAVQQAHLLEEVMMALRDELVATQAQEKLLRASEEA-HRERCAEQQRAVEQMQARVEELETAARNAASSLSTPPPAQPATAAESELIAMNA-ALNSLAYQLEDAQHRVEELSAEHARTVEELARATRAVAELERAAGAEQKREVMH-TAAEVTEAELMARVAALTTENAQLKEKLSRAAVTTFXXXXXXXXXXXXXX-XXAEVTEAELMARVAALTTENAQLKEKLSRAAVTTFASRQTDGEGAAASLEER-CAAEARRSEAAEAEASALRGTNRNRTAAELAEVLQRSEVQATLERTVARLAEEEA-LVEQLNYDASELNDARTALTEEVTRLQKALAAVTDKADVLAQLLATQKELDRAA-EVQDAQYRDEERLQDALATTTSAFDVQTKQVRLGMAHVVDMSRAFCRVAEAVV-QQSAAGVQDGAASGASVQEMLLPWTRALEWVAAMEQRLLETMDGAMSNFSFR-TDGRRRRRRCRRGRRTTELEDDNAVEMEANSGSQALSSSLCRAVPANRAEAFA-ARLMDQAAATNTIADLRRLTADKEVDLTLNRSAMDALANDLISERDRADNLSAAI-DDASEKIASQAEFEHLRQRTAATTAEVVEARRAEARATAAQLRAEEQLAAVEK-DLKEQGVLRKLDQENHRLTREADRLLRASRLSESSRARSNVSGVTPKSLSLSA-ASAEDAQRFTTAPSTLEAVEQAQILQVSSLQADLLLRRQARELEGREATLRQ-ACAALQVSELVVEAAEHLREDLALRADYAELEQRCSELERMAVAAGGGT-MQELKQRRQQLKAREVELEELKARMRDVLSKAAPLEAARRQALRESLSGIT-TQLAATQAQYGGGTGGSSKVRRAASLTAAVNSDLSLSSSRPQDVLTSRIE-HLQTLVNDREVALEKVAQAQLESRAAMDLEDQLAAKTAALERAELVAEYADSV-NTAFTMAAVVRRGNRGAGPTTVPPLSSGMGSPDAEALSASVPLTPTLPTPRAA-VNDGQPDAAVAMASAPRMPNLDVAVDDSDGSGEGKQAGSAGPSTDSAPPVVT-HRRTRASGWRDAKSCPAHRPSSSTERSAPSSPNVGRKSFASASVP-ATTRKRARSRTNT
Leishmania gerbilli	Trypanosomatida	NCBI AT-BK0100060 8	MRRSEPGDAHTPHGQNRGEEEDKGGELQRSISIGIESDSIHRAGHGLCMWGN-FHNAGTHQQSSSLCCDGGPQAIMPDSTSLPSPSHAAQPARPPPPAGVLSAD-LARPIVRELAQSRVAVPPPPPATQTPRPHATTAGFAGFVPTSSAVYAATS-SGVST-SQPLSAHADAGDLTSLAHRDIDHLSGRELYNYASHLRESSQQYGGYLSE-AYAKRDQYRGEVARLKEELHNRYMEIDVLRREQERARDALLKVRDDNAQ-LRQKLAEEAEGHARNMQAKISVLSGADPSLAVHFYQSQLALDKDAQRELQ-QRERAAAAASSQRRSSSTTAAEHGRRSANISGEGAAAPHTSPSCDLPSEA-EELQDQTEAASAAALHTALEELQCLNTSLSHRLEDEHTAHKAAALAAEASA-TATADRAELLDTIAQLRQCCADMQNVDELVAALTRAPISKDYAALQAS-YDDLTAALAKAEQTHVAALQENQKHLHSAQQAQEEELRHSFTTALREERQS-MQARHEQMQLVNHLERELKLAERANQLRQEQLYAAISGNQQLAESLHV-AEQQLSATSSELQDLRNTRESLESQRELHQRIEELQGDGHQRHQHGD-ADAESESSQRRQEL-EVELARLRSQRDALVGDVDRDDVGDVRRLLPDERRT-LTSPPAHAEELQQRSSAVRAPLDVRAADDGEWQKQEDIAALEAAHRCIEQ-MKQTTIAALQSERVVEADLKAKVNTLEAELASQSRALAEAAAAVRAQ-LEEVMMALRDELVATQAQEKLLRASEEAHRERCAEQQRAVEQMQAHVVEE-LTAARNAASSLSTPPPAQAAAADSELIAMNAALNSLAYQLEDAQHRVE-ELSAEHARTVEELARATRAVAELERAAGAEQKREVMHTAVGVAEELM-AVRAALTENAQLKEKLSRAAVTTFASRQTDGEGAAASLEERCAAEARR-SEAAEAEASALRGTNRNRTAAELAEVLQRSEVQATLERTVARLAEEEA-LVQDLNRSAMDALANDLISERDRADNLSAAIDDAEAKIASVQAEFEELRQ-RTAATTAEVVEARRAEARATAAQLRAEEQLAAVEKDLKEQGVLRKLDQ-ENHRLTREADRLLRTRLEPSSRARSNVSGSVTPKSLSLSAASAEDAQR-FTTTPSNLEAVEQAQILQVCSLQADLLLRRQARELEGREATLRQACAAL-EQQVSELVVEAAEHLREDLALRADYAELEQRYSELERMAIAAGGTM-QELKQLRQQLKVREVELEELNARMRDVLSKAAPLEAARRQALRESLS-GITTLAATQAQYGGGTGGSSKVRRAASLTAAVNSDLSLSSSRPQ-DVLTLSRIEHLQTLVNDREVALEKVAQAQLESRAAMDLEDQLAAKTAAL-ERAELVAEYADSVNTAFTMAAVVRRGNRGAGPTTVPPLSSGMGSPDAE-ALSASVPLTPTLPTPRAAVVNDGQPDAAVAMASAPRMPNLDVAVDDSD-EGKQAGSAGPSTDSAPPVTHRQTRASGWRDAKSCPAHRPSSSTERS-APSSPNVGRKSFASASVPATTRKRARSHTNT
Leishmania arabica	Trypanosomatida	NCBI compound of: ATB-H01000268 and ATB-H01000267	MWGNFHNAGTHQQSSSLCCDGGPQAIMSGSRISLPSPSHAAQPARPPPPAGVLSAD-LVRPIVRELAQLQVAVPPPPPATQTPQPRATTAGFAGFVPTSSAVYAATS-GVSTSQPLSAHADAGDLTSLAYRIDHLSGRELYNYASHLRESSQQY-GYLYSEAYAKRDQYRGEVARLKEELHNRYMEIDVLRREQERARDALLKVR-REDAQRELQQRERAAAAASSQRRSFSTTAAEHGRRSANISGEGAAAPHTSPSCDLPSEA-EELQDQTEAASAAALHTALEELQCLNTSLSHRLEDEHTAHKAAALAAHAEASATATADRAELLDTIAQLRQ-CCADMQNVDELVAALTRAPISKDYAALQAS-YDDLTAALAKAEQAVALQEKEQHLHSAQQAQEEELRRSFTTGLEERQSMQVR-NEQMOSFVNHLERELKLAERANQLRQEQLYAAISGNQQLAESLHVAAEQQLSATSSELQDLRNTRESLESQRELHQRIEELQGDGHQRHQHGD-ADAESESSQRRQEL-EVELARLRSQRDALVGDVDRDDVGDVRRLLPDERRT-LTSPPAHAEELQQRSSAVRAPLDVRAADDGEWQKQEDIAALEAAHRCIEQ-MKQTTIAALQSERVVEADLKAKVNTLEAELASQSRALAEAAAAVRAQ-LEEVMMALRDELVATQAQEKLLRASEEAHRERCAEQQRAVEQMQAHVVEE-LTAARNAASSLSTPPPAQAAAADSELIAMNAALNSLAYQLEDAQHRVE-ELSAEHARTVEELARATRAVAELERAAGAEQKREVMHTAVGVAEELM-AVRAALTENAQLKEKLSRAAVTTFASRQTDGEGAAASLEERCAAEARR-SEAAEAEASALRGTNRNRTAAELAEVLQRSEVQATLERTVARLAEEEA-LVQDLNRSAMDALANDLISERDRADNLSAAIDDAEAKIASVQAEFEELRQ-RTAATTAEVVEARRAEARATAAQLRAEEQLAAVEKDLKEQGVLRKLDQ-ENHRLTREADRLLRTRLEPSSRARSNVSGSVTPKSLSLSAASAEDAQR-FTTTPSNLEAVEQAQILQVCSLQADLLLRRQARELEGREATLRQACAAL-EQQVSELVVEAAEHLREDLALRADYAELEQRYSELERMAIAAGGTM-QELKQLRQQLKVREVELEELNARMRDVLSKAAPLEAARRQALRESLS-GITTLAATQAQYGGGTGGSSKVRRAASLTAAVNSDLSLSSSRPQ-DVLTLSRIEHLQTLVNDREVALEKVAQAQLESRAAMDLEDQLAAKTAAL-ERAELVAEYADSVNTAFTMAAVVRRGNRGAGPTTVPPLSSGMGSPDAE-ALSASVPLTPTLPTPRAAVVNDGQPDAAVAMASAPRMPNLDVAVDDSD-EGKQAGSAGPSTDSAPPVTHRQTRASGWRDAKSCPAHRPSSSTERS-APSSPNVGRKSFANASVPATTRKRARSHTNT

MSC domain-containing proteins:

Perkinsus marinus	Alveolata	Uniprot tr C5KFV5	MSSSPGNYAAMTIAELRSILTRNGVKVPQSSIKKAVLHVFEVGGSRFRKFRIFYGE-SATTNAVELWRKESRRGSSKGDGRSRVSKSSRVGRESTSRPT-TNSRGSVAITRRSSRARTSISSAVDSEGTDLVDVARMFSSDEEEMSDSMGLPG-PRPIYCDITKEDCEPCPPNAECRGRSMVCKPLFKRDGDVCRDERL
Tetrahy- mena ther- mophila	Alveolata	NCBI EAR90926	MNSDQNNRYRGGGNSISVSSMRKELIKNNIDIDKADLADLASLYQQFLQEKGVYQPN-QFSTPSEQNRENGQRQSHNLPPPKERQQSIQGLNIMIDHSEKKN-NQNYVFDVQHNSDVTGGQSSNQMPNRRNDIASNQNMQIPEEAIEINQSSARQR-SFOAFNFQQNQSGRNSLQNNRASSASMNNSRLSGNQNGLGLQNGQQRF-SNEGGRNSRQEQISYSPVNYGNFNFRMSNQIYSPGINNGIRSSNGMMNN-QQGYFINENSKRVSNOQSYNNFNQIESANRRSQQLNSPVPQTGKKNLWKN-GEEINDFAGSAVFNHRHQYTPNYPKSTISITDKWKGLLDTARELSQKKQRQS-NDMMNISQYNIFSNKKSQVESRKSSENIAEVDTSFYLSQQQKLNKDKVLIFF-SKIEISHIVMFMVGLFCIFFLQNIYDFLTHTEPLEYCDSTLRQEQIEGSSCPCP-QNAICLNNNIKCKHSFVLIDKCCVTNPQIEPNTIRYLSLVQNLQYLQQAQHC-YDSNSSGYGNLSRYPYKEHKWESIVNLITHHSKDLLFYDTLTKVKDYLSKKS-IREQFDIYDLSLRTYTYVTFNPSYCRIRLWSIQNVIIIAVLCIFFLIFLLYQVA-AVRRFIYFKLYRIQQKIVNENKIQQQNNSLQDLDNKSIDNQQLYQQGGQV-SQRFASPNQNSNEQSFSQRYSQQQNPSQQQKIKLLSQQQFLQVPQNNQNYL-QIPSNQQQQQQQLKLDNLSLNLQRRGKLS
Tetrahy- mena ther- mophila	Alveolata	NCBI XP_001007 953	MSKPVNFDLIIKPNKIQNKQPIKGSNSNTTTPKEAVIPPNLAGAQKNEWLE-QQQKNSQANAKKDDDDVQILDYDNPFKNNQEAQKDSLKKQIQKPNQQAAPP-KNISEGNKINNTYQIISKPNQLQANSPQQQINNNVNIIGDQVLFRTQNNKQD-KNAIQPTIEISKNSIQFNSDSSNSKNNIINQNNKANQISSNTNTHNNKIV-SNNQRVQSNKSGSIFSSNKSNSTSNKSAGMSKQEMLKAKYSLVNSQSKDEYKIF-SKIVKEYSDTENLTYNPSKINYSLFYFVCFVFFVMVSLDYVASNTPKFCNEVK-YDTPKIVCPENAAQCKNYVMKCNDFGIEQGYKCKLNPNLISSESFIEKITQIQY-DNGHRMCFAGQESDIDQKQKEEEEVIKISRAQLLEELQQLADKKAQVQKQKILEK-FLVSTENQELGFKYDQESQTLTSSRLRFTIFCKCKQLIKNNQKSIASISLTLVSLIIV-GILKHFKKIRIANEIEYELAHQIQETSQDKQINVSYYKEGINRFRKSDFNEDVWIL-VDRKRVRNEQKINVYRDCDGLYWSQI
Toxo-plas- ma gondii	Alveolata	NCBI XP_002365 595	MLLVIFRRSCSRVQRDITLALPEQPASLASLSPRVRRASSPPSSSSSSSSSS-APSSSSAPSSSSSSSSAAPPSSSRCSSSSPMQRVLTSDGGGGNTAGALASAR-RRAGRTEASKTGGDEQLDDEASRVHTAGEQPQEQRRRARCRVGSKGGRR-RRGNLSLRGARRHSEERHMCPPSPSFASSLSPAEEERRQAATPSSLPLASSPLSS-CPFLRPLRVSLISLLIFALGLVLLFLATPLTLRNGARDSQLFLPYCNKTRNLPLA-LVDNSPAPPPALAAARSDCVCPAHAVCLNGEMRCVPPYQASTRASALESLLAR-FGVRATFLRSFFVLSGLSRGEDASSCDEAAKETAKRVTDVAVLFRSKAGKACQ-ASFAPLGSQQSQSPRPAAPPCEEGVESGRGARHTVCGPDKSVFQGRDRDATF-ARRTEKFAETRTGTEIEEHSGRSSLSGRTSKSSGALSERSARESPSESATRW-IEIQAQVLDLLEDNNVYAYVNFHFLDEETAERNFSVRVEQTSYFSQSLDVLVPTKS-AAARHEQTEALLPLFLREQRRSRGITYTYTYTGLHALRPWSTCLYIAVMRLAPLAV-ALLLLLLSLPLAFWLRRVYLRHLREVLLSRERAEASELFTLQRNALATAIFRPTSS-KELSHLLRLSRTLRTGRFLRWLQRLARNTAEVDTLCHDLLVDKYIRLHTARGIEQ-NFWWAETESPPEAQPKHPSPNPVPPGASGANVAGSWDAHSQFSSAGPPEHGV-VTGGCQDELDFQQSARPTLLEMCNGSNGPTNWCYSYLRGDGMCGETDMFASGG-NRTVADGWLGERVRRSSYSLACRLADSLHAETSLSGQSGKSEMRRPTLGGSDA-SWREQGYSYQTAGDTLRQATLGGGIDQRLGVLLRTEDTSEIERTTSLGKALA-TLMGGDQVRSLDPGSRNVRLQGLSVSSEGGESGTVSSRQSSVKERAARDYSH-AAFFSRASRRREGSRGATEQTDGYTAERGRDDLDFMSGEVADNHTINGDYKGTG-PYGQTHLSRRARRRSEERRSSGMRRVDDVDEDHETGHSVLC
Ectocarpus siliculosus	Strameno- piles	Uniprot tr D7G5B7	MARPSKRKPAGAAGSTSPSSKPKRTPAKPAKNTKAEDRKSGRKNAAK- WAEVTLKKKRTSGEGTSSSEDDNVSDAMETSIDLNDTPGSPVRKPKAPRPA- SRNGRTRKPSLPPMSETISETRAPPSTSWPESGSGSGGRRAQASQMSNGSP- RAPPPTPTATPKRATGRRLGKTYTPSRKSSFLRHANGTEGDTASATAASAD- GPSLNSDNLGGPKTRGQARRGPSFLDRERAGRDVATAGAERTPDMKGR- GQDKSKSVDSVPOPAQATDAGADADGDDARPRRGKSLWTFKCAVWVLA- PVVLLLVSGLLVAVPHVQVAAAGAWGAFRGLTAKTPSEKCFLOPGQEEDIV- CKSGTEAMDCPDHATCGAGRILHCNLPYTLRKDRACVLAEAKEAKVAEAL- RTLTAELCKGKGLMPASARGSFITLDAATAQSRGPLLLAGMPPGNVGLI- ELEGEQARVFGGQDGAEGCAGSGMERSVARLSESAHDVAMRTLGWRCARL- AAGSWFMTYVWQYCLAMYCLVGVAVCLDLRRNSYRHWVQVCTIRDLYQTL- DHASKDSSPMQVSHIQSSVTDWTRQVAKNRATIRHAGLWPLVERVEHDE- RIRRCYKSLAGKRSVACWVKWVGTGTRFETIGALNPGRAPRLLLLWCQERA- RPMFWCQLCWLLQA
Phaeo- dactylum tricornutum	Strameno- piles	Uniprot tr B7G9F2	MPSTAASRKRKARAPESPETAPARRSRAERESQKQAAREWAAQVSHTRPET- SAADKNEVLQSPPKRRLTPVVSATTSNRKAPPSEVSKKAPAASQASELTS- SRAQAPTTPPVKLSAVEQARAARYMQQSSSQSPKPPPNVSSSETSRAPAAT- PPRPTTTKAASTHVTKPVFGSPVPTSRQRHGDAARNRRSE- ETHVDSLPLDEVVKTLLHTRSQANEDEISLPSDDNIDEESFL- DSEDEDEDYEDDDDEAEQEELAQARARERRLQVANDASRR- GWCGSLLGLVIVSMALASVFESGLSLETLAEVWWSPIEPT- QPCFLDTFPPGPEDDVDPDLPHCANRQPCPDGGFCRDG- HLQRCQSKYEQVSDHGDACILSADTNQTIANMVDLLRSWTI- DYYCGLDERAQDLHLREDPTTGRPLFHYSRVTGELEQGYD- WNLVKTANAKDPQIFITAVDADLFIGLHPDQASRPLICVAK- QWVFGGLGMLGTFVYSLGEFCLMFYGRRAAPAATVIVSVT- GWLLVIFISRWQRREAEERQILLDVTIRQAAFDRLLSAPGT- ALSALQLRDQLMWEHFQTSRQGRERLVKRIWPKIVHDVRO- DNRVRKSLGVRDQKQDLVWTIATTTKSPATEDRRVAPTAVRFG
Phytoph- tho-ra par- asitica	Strameno- piles	NCBI ET- P40127	MEHSTTHIKRPPSSSTASSNKRHRPPASPTPAPSAPLSRPKPKTRRLVT- SNKVAESKPKERKRKVEQLTFSRYADPLVQSPKQINHCSTGATILETP- PVRHTEPGSVNGGTVDLATVPDMTKLEIVGEDEILETMLRASTRPVKRF- PSEEAALDSEQKDRERILRRRRVKLRSQYSAMKDDQVDRDVGFTVPTQ- SLNIERNLFQSEDKTEQIAKEEVERAQLLEEEAEETDEETTESKINRLS- KAHLNWRQFCWLLSGAFLLCVVVTAFFVKLLLEPSLPCDSEW- EANEGSFVLADPADYDFDRSKALQPFIEISTDSTQTSRPLCQCPVY- NCLNGSVISCAAPPYLLHSGVCKENPEIQENLDQALTIQKFVVDKAA- KNVCDNVSLWSYLNPGSKAESTRDLASPIEVLDFVQAFVTRTISFG- KAVAMLPREYVFNALDMALRDLKDIVVTEQQRQLVGGSVVWVSCR- AKLQLYSHFKLIALAVALGTALVFSYRQFLYRTERQLVDRFVKEVRF- LLDRTRKRDRFYPSDHLRDDLFKQSLQNRWLCKSVWPKVAVVW- DESIRTRVIVRQDQDLVWVWVSSSPHRRPVGGR- GGFHAASRQQVTGDRSRSTHVQSRKVT
Naegleria gruberi	Heterolobo- sea	Uniprot tr D2V886	MSTSNRASLNNDEDESPTRLTIKELKSILTEHDVELPAQDKKRFYVDLYNKHK- SKKRKNDEVNDLSDDENPEQKVKVKSSTEEPAESSTTTTTTEISQPTVEEPK- VATPMAEVKAKQKTKSFAPSSLEAETPQVKVSDVRRRTVGNRDLQYROR- PARNEINQGLLQTPVSSNRAPPPTLSPDSTLVTQDIRPSEIDQEONLQKVLTS- ILMASIHALFFCYVLFQGRDPHLPYGGRYCQSEDLGLLQKPRDNLCPCPE- NGYCSDGRLVLCQKPYVIKNAVCVDSRITLNALEFITEQLNELSSKSGKFCGR- AESNSTSLEDIKAKLTVKLAGSGVAVEDVLEQVTKLVKSSPQNFQMINIAPVGS- DKEMIFSLNPTLFPDCRIQKATQENKGVITILTAVAILIANFYKKNVEREQNVED- LFRISITRIREEEIEVQLRDELKEKYGHEIDVEKYYSRVEDLLASDSRLSQ- IPSNSTAYQWAVATDEQE

Giardia intestinalis	Metamodada	Uniprot tr A8BJR2	MDPNDYLNTPSRLRVIDLKRAITEFGGVPEERATKATLIKQFAQLRSNYLRTVSRG-PRSRPQSEFSDIGFVDAIDEQLVAVNQNLYSKVGRGDLTQSAFDTGAFNDYAL-HEQLVRENEVDAAKPVSLPKRMKNPYGVERTPGERKDSASAPSTGSSKLDGK-SQTEKSKTTLTVSPVDQOEIARSVLRKSSYIDKGRDKHSSSQMQVHAEMK-HSSKNTTKKHKSERHRSKSKEKRVESQKELCKSEAGQFPVKTPTRIEA-PTNPYPHIGSSIPRDASHAYGRVQLERSVHSHRGSTVHVPDKNLTKHT-KRRDRLTLETQPLKQKQYEDHNAYNFKTIHSHPTLVRSLFVFLVLLVSYQ-LFSLADSPVKPQIPVISHVISVLFACIFSGPVYETLTHSAQYGITTEPCTFG-TCSGGLTSCMPGYQAPSNNSFYQVSVFVARNFFSAIGNNAMARWMP-GAYKHMSDAASRSSDINGDDSYLSQIQYSYGFMTAPVLRMRRCRQA-SGVVTKLSSLLQALDHELQAYKGSFLCRSRGRKIVTDSLPLLQTRTISRFL-RNSLIFSVVNTSTNGILSADSLKDLVTTYTDLQNDVLEALTAQILHEGTDN-EVYLPNNIKFRLPEEARINRSVVCQSDSDSDFQDCMHKQYIYLSNTNVFTHS-LPRFSLACKFKFWFHRDPAALGKILQCTVSIAGSLYIATKIGVYIAGYIARR-IKTYLHDYKVDMLQVAKPGDAKLPGFAAVTLVQLRDLFLASYPRVIRLILWKK-VLKNHSDSRITITSSIDLTRGIAEFVAVNGLLLSNSFGNDAPIAGRHFKNHVTPVS-LPPPSDNE
Trichomonas vaginalis	Parabasilids	NCBI XP_001299627	MNNDYLSAVPRGQLFDALKKGVTTPRYDAPQEEVLNLIKQKIQGTDIQRDKELL-GETFDHLLDSIVNKLFIALFVALSAIVLIANFYKPSKKLFCDSNPATGLCKICPK-SNAICSEGKAKCHEGFKLMHRRCIYDDDDSIYSSLYEYKLEKQAGRFDRCLDK-YKXITRNDLHLIRHSKFNVDIEYLSNKAVDHLVYENTIGNVTEFNQELIYSTNMQ-KPISICVROEIESHIFLTISSILLSTAIYIYVQSVKNRRRQSFLLFVQSMINSYRNHHIP-VNEDTIRSNLVSIDPNPDSLMPVYLEELRRNPVKVTSYICGSLTRFRDQK
Trichomonas vaginalis	Parabasilids	NCBI XP_001313092	MRTAGPLDDAFVPTQRRILGNYTREQLVDLIDKYDEAFDINASKEKILDLIAIMKN-TKPVSQPETKRSDIYLYLWSWIKYVLLFIAYFAFLAVSLILFYLVFVKPKQFC-NTNKQSYICITCPKFAKCKSNVTCETGFTFDGGRYCVKTDHVVHISKMLDFATTELR-KRAGKHRCGKCLSNLYTDLQDALIYQNFVQYDFVFKMIQILQQYVNIQSTQ-YYDDRTTIIASTIADRPFSCLIRRLISSTFLFILCICYLIKYSIKRTKQKFIENKAQYM-AIKVINEMKNHDSIDENTLKHKFSKNPEIDLISLWLIIDSYLKKSPYVIARIS-KGVTYRYFIL
Acanthamoeba castellanii	Amoebozoa	NCBI XP_004337553	MSYQVDPDPSLTIPLKLSVLTQHGFQDQLPAHQEKKGEYVDLRYKHIVPLLEARN-NAASQSLTPGPTSTTTAASAKERATKTDLVDRRRTIHSLSAPETEAAPTTLTQTKL-VAKSAVPTGTPTTAAKADAAPAAMVPRRLFPEQPTVTPSSSSSSSSSSSAPIL-SPKASASRRTTIATGAVSYPKLGAAPAAITAKAITSAGPDLRQIALQHARQSA-CGSSAAMWLFGLVLLAVAGAGLFMYGQLPSLLQAEKPYFCDDTGTGSDNRQNP-REPIVPCVDCPNFGVCKDGRLIKCMRGYVSGGVHCARDKEIALAFEFCKEAH-KTLALRAGEHCCGYPVAPALPMEQELRDILQKMKDAALFPWDEAKFNEFLDL-KKDLQRTSISHRNVIRASNGNYTFDITDVPYACVLRVLEHIMYVIGTLVLAAL-VLRFVRLQQTAKAREADIETLVESTKTLIKEKKGKAGPPSSVLHIREQREKQTEL-KKRYPTVKDWLAIWVAENVRADTRFKEYSQMVAGQQQQVWEFVGTDAAMA-EKILSPSRRSYPH
Galdieria sulphuraria	Plantae	NCBI XP_005703659	MASTPDKEAQQGSTSVEKRSKTSASKVVASGDRGATGKARLEFRKGFYKNDTKD-SLDRSSKSTGSRARGENIHSAAKHEDEQVVSPPISPKLDEVESKSPFFVSRNL-GTEFLSVEGTPOPKSAPEHFSGIKLLSITILAVSIFFAFTYSPPIQLPFC-D-SNNKVLVEDCRPCPSHGTCAGSAGELTCDHGFIESAGKCVEDKTSFVVTDKNRIEA-ILRKRAGMSLCTGTEVQKEMDSACLKAVLLSFGTKLPLEEIRFNRAFTYATVIVN-DSVSSDEFCEEEGERIWCALHPSSLSCSRLRFVQLVEWVLLITMSILIGISRLY-FWFRLLRRWIDKQVKTVEEYVYRQLTASSQLTSPLSERKVVVTRLDLHDQSSL-KWKDVIWRKVVVERVDSRSRVQKREIVNDLPQLVWVWHDKIV
Cyanidioschyzon merolae	Plantae	merolae.bi-ol.s.u-tokyo.ac.jp Locus_CM-P246C Location_c16f0016147420 CDS_#642	MVHSPPLGSKYKTFSTADEKRRVNPQRPIRKHGRWIRALSRNAFRNLPASRL-STPMTDQWQKQVRELKQYLLHEDPSAMPAQQRPTKAELEAVEVQVHQRLQR-RQKTQGNPPRALPKEEAREPRRPGPRRQSGALGERKLPTRAPRTP-KKAERQQQDSNRGVGKSEAPERPPSSSESATSRKTRQKRSARKTEAVQSGSSGV-DGSSSGRKRVRRSPLRGEAEARRLVGGDDIPEDRSPDGGAWAESQHPLLWVSRP-SEAQLSPASLIQDLRRRPFANATSTRTSSPDRHRSASPRASSTWQDVTREFFRPT-KQGRSPSPVTSASNGAGPSTSSGLGLREALRDVLSMVKSHDQLEPMEYVEPEQ-VLPAEASHARWHALTHQVSAVTPAEGDASRSSLHSSPVAGRDVWVLRQPRNA-AAQTKAESRITLLRDTCDPLAQSQHMPPREDVVALSTGSPGELTSGGTLAQRSTP-TRAYTVQDQEAASGFMTGSMGTRLASLHEQVSAVLERIRAEQANLVSPEAATS-TALSRNADSLATDALHQRLGDSGAAAPAVASLENRSPGSAALTRGTLTVAPEPP-FAQSPQLQAFRTLQNRERMPVSSAATLGSRLRGRPVHLEQRSQGSAAERSTAGT-TDPTRSPGAFHASWLSPLSTDAHSGDQPSISESQRYLARPETHAAVQETSAGTPGN-ASRVSRRDALLGNRALEHRHWLRSHWRVVALCLLLWFAVRYLVLPSPFCDSS-SSGSPVLCRCPCEHGICVGGVLCRDGYVVLGSTCAPDRDLNRYAHIIQGVHR-LLGESAGRYRCGERAIRWVWTERELREALEQHPYMERLVRSSKWKAAFRRALDT-LDDPVVQKPRSAHDTVHRSRPSVDDANDTVPAEAFQRGGAERPLLADDSGTNS-NSNSNISPNMASNESMQPDSDATQTGQLADDELLYVSTDAQRPMWCRLWQLTE-AHLGSLIWLGSITAVTLMFATTLRRKRLRQRQIRSLQAVHGRLEQEKRSYLVALQR-GRADVVTPLIDVHLRDELLGDIANIYDRQLWEVGSARARTDSRLQLRSETFTDGR-RRALVWEVTGLGSAENAAPHLVLSGGTTSVE
Physcomitrella patens	Plantae	Uniprot tr A9RFH2	MPLQSSPVCVAKPSHYASPVKFETKYEPASTPTSSGPKDKDRRHKKKQHSS-GKKRSGTFLSLKEQVQLIALLTAAVVYVNTLTAIGWFEEARRPFCDDQDQPLRD-NCRPCPENGICRGGELRCIPGFRRQGCALCVPDKQIDRNAGNLIPEHQIWD-DASEEEKMGLDAENVLVHEKALQLVRDKLNIRIDDFGLREFRCPALALTYQ-PIGQRIRLIKANIYGIFASFVIFVSYLRYLQSRRLVTRAEELYMQVCEVL-EERSTGSSGETKWVVASRLRDHLLPSEKRVATLWKEVERLVQEDSRIDQ-YPKLVNGDSRVVWEVQAATVSRIRICVAYGLRGS
Amborella trichopoda	Plantae	Uniprot tr W1P105	MKRSRASSSSSSIDLLPSTKEQFLQLLGVAAIVSVAMSCNYFLNYSLLTITLPCFD-NGADSSDFCEPCPSNGLCSEGKLSCIDGYRKHGRLCIEDGELNETAKRL-SKMVELQACGSYAHFLCHGTGAIVFQEEELRSVWESNLKVNYGIDSHILEH-TIQKRVREIYESSFDIKTSPDGMRELKCPDWLAEQYKPLPCILRIWFLNHIHYILPIS-LLLVAFVRLWVRVHMHRKLSTRAEQIYEQVCEALEENAVMVKNSRQDCGAWV-VASRLRDHLLPNERKNASLWKKVGLVQEDSRIDRYPLLIKGETKIVWVWQ-VGQDQLVRFGLRKLADPDIELDILWDFVCGKVPVQDPQLDYCDDWFLFRG-GLSTFKDENNTKSEQDETF
Arabidopsis thaliana	Plantae	Uniprot tr F4KH17	MDSIPRKRPKSETRTGRTPKSSSSSSPIRSMLEPPQSLFSPKGEFFTLKVLVLA-CAVAFTCNLSKLSLSSNPKSFCDSNFNPIDSLDICEPCPINGECYGGKLCQNL-GYKNQRNLCEVDEGINESTKLVGYFERKVCESYAHNECYGTGTIWPEND-VWTELRSNSFLSNLDESAYNFLKGAKEVGVTELEKRTNSNGIDELKCPESYAK-SYKPLTCRLHQWILRHILISSCAMLVGSAMLRRIQRKQCFRRVEELYDQVC-DFLEENAVASNSAETSNECEPWVIAVWLDYLLLPRERDRDPLWTKVEELIKEDSR-IDRYEKLLKGEKKVWVWVQVQVGLSLSKLLKQKRETQKVKRKSIDSSSTLQKEY-YNRRIAETSS

Homo sapiens (Man1)	Holozoa	Uniprot sp Q9Y2U8	MAAAAASAPQQLSDEELFSQLRRYGLSPGPVTESTRPVYLLKLLKLEEEE-QQQRHSRGGGRGNKTRNSNNNTAATAAATAAGPAAAAAGMGRPVSGDLSYL-RTPGGLCRISASGPELGLGPGGASAAAPAGSKVLLGFSSEEDSDEA-SPRDQAGGGGRKDRASLQYRGLKAPPAPLAASEVTNSNSAERRKPHSWGARR-PAGPELQTPPKKGDGAVEDEEGEGEDGEEEDPEETEEPLWASRTVNGSRLVPSYSCR-ENYSDSEEDDDVASSRQVLKDDSLSRHRPRRTHSKPLPPLTAKSAGGRLETSV-QGQGGAMNDRAAAAGSLDRSRNLEAAAAEQGGGCDQVDSSPVPYRVNAK-LLPPLPPLTMDSTLDSSTGLSKTNHHIGGGAFAVSDSPRIYSNLSPPSAWAAS-SLRINHANTHGNHTYLKNTYKPKLSEPEEEELLQQFKREEVSPGTGSFAHYLSMF-LTAAACLFLLGLTYLGMGRGTGVSDEGELSNIENPFGETFGKIQESEKLTMMNTLYKL-HDRLAQLAGDHECGSSSQRTLSVQEAAYLKDLPGEYEGIFNTSLQWILENGKDVG-IRCVGFGPEEELNITDVLQFLQSTRPLMSFVCRFRRAVFTVTHRLLLCLGVMMV- VVLRMYKYRWTKKEEETRQMYDMVVKIDVLRSHNEACQENKDLQPMPIPHVRD- SLIQPHDRKMKKVVDRADVFLAANESRVRTETRRIGGADFLVWRWQPSASCDKI- LVIPSKVWQGQAFHLDRRNSPPLSLPCLKIRNMFDPVMEIGDQWHLAIEAILEK- CSDNDGIVHIAVDKNSREGCVYVKLSPEYAGKAFKALHGSWFDGKLVTKYLR- LDRYHHRFPQALTSNTPKSPNKHMSMSHLRLRTGLTNSQGS
Homo sapiens (LEMD2)	Holozoa	Uniprot sp Q8NC56	MAGLSDELRELEALGFQPGPITDTRDVYRNLRLRGEARLRDEERL- REEARPRGEERLREEARLREDAPLRARPAASAPRAEPWLSQLPASGSAYTP- GAYGDIRPSAASVWGSRLGAYPARPAQLRRRASVRSSEDEEDARTP- DRATOGPGAARRWVAASAPARLPSSLLGPDPRPGLRATRAGPAGAARARP- EYGRRLERVLRLWASLGLLVLGILWVKMGKPSAPQEAEDNMKLLPVD- ERKTRDEFQAQKQAALLELHELNYFLAIQAGNFECGNPENLKSCKPMEAQEY- IANVTSSSSAKFEAALTWILSSNKDVGIWLKGEDQSELVTVDKVVCLESAPHR- MGVGCRSLRALLTAVTNVLIFFWCLAFWGLLILLKYRWKLEEEEEQAMYEMVK- IIDVVDHYVDWEQDMERYPVYVILHVRDSLIPPSRRMRKRVWDRAVEFLAS- NESRIQTESHRVAGEDMLVWRWTKPSSFSDSER
Hydra vul- garis	Holozoa	Uniprot tr T2M302	LFNVGPPSKDHYRINMAEINVELSNAELRKLTEFGKSVGPITKSTRVTLEN- KLLKLINESEQSISNNSNHHKKVVRKSNVSSRRSSLSRSESSNLKFAFFSDDEE- LFLKLLKSKNSEKLLDLEKPKLVNKNELIEAKVQIPSTYTSASPISQIDIKSSP- GKKYNSSNDYQSGCDLNRKNIAEYSDDEHFLEKLNEEKRQNIIRLSKKEIKSPSME- NNVFRTPSPVIVSRSTSMNAIDKNIRGEVDACKAEVRNFTSKKPSPTISRRVQLS- SYNNYNAKMKVEGFKDREANKNEHSEKHFFLIKLSLFEKQIKCIACILFGLSAV- LVGLYWNEDPLKLTLYLKDQSGKEKDJINIEICLYKDLITRKGMEVCGSYSGPDYL- RREDFPQLAEGCLKNOKDIQFNVIANDTWEAAKSYTEFFRFNETMTVTSIGSKPF- HCRVGOALYAVLFRIFILLTIIFGMAIYFYMKRWSAEDKTRQVLYVHVHIREVVRQ- HDTQCLNNKELSPYLIPIHVRDMLIPLEKRKEMAKTWQKAVEFLSSDSRIVERTOR- ISGEDFEVWRWIGVRTPGIKEKINSFNHLNSDQDKWQGFADQIEKVIIRLPIITPA- PCLKIRCMHEGPDEREEDWVKVENAVLEKCEVDGARILHIHVDINSRECVVYK- CDTLESARKAFRSMYGNWFDGRVLVIVKLVTLARYHQRFDPALCKVQALSPSGSNA- VVNSNKDI
Capsaspo- ra ow- czaraki	Holozoa	NCBI XP_004364 881	MPVLSRSSGSRSSASASAPAAATSSVPAPASPTRRRWTKRLPSPRSRGGEPVGV- PSVATPTTAASSHGREDRYDDDDDDDDDDNEQVQVQRAPRSRT- VAVPRRSADGVAKRARGSGAGVPGAHAALQQGGVSDDDVAANSDDAANELQI- GRIGGOIGRRRAGRTATLPGIGWDFVKSIVQGRQASVSHQNEADAASSRRAVRG- SVHASKSAASSSLSRVFSSTAVLKAFALTCLLLSVMVVSVSHHQMPYCPTRST- AVASTKQCKPCPEHAVCLGGPTFTCNPDYQGGWRGCVLNTALEKLDNIAHE- IVEAQRVYRQRECESSSVTSLAVPREAMRELVKARMETDAQFTTAWNVFC- ADPEKYFGDAVEVRMDGFLAKAGIKPFECQLREELLYLDDCAVAQVDFFWYVW- QLLWLPPLLWFFTLAVSYFKWQKSVKQNEQDALNTMVLAIMTLREQAISQPQ- PAVSVSHLREIDVSDRKRLAARWDEACKIVRRDRSVAQSVQSIHGIPQDMLS- WVSPVVASPMAAQRQASPLVDNPDNADNGGGDENDRQP- AIRSPQAQHRPYASSIYPIADI
Saccharo- myces cerevisiae (Src1)	Fungi	Uniprot sp Q03707	MNSDLEYLEDGFDPNMVKVATLRRILVENNVDFPNSARK- NALVGLFDEKVKQPQLRKMVLNVRPDSGEGIVKMDRPSSSPSAPRRSRAR- REKASAPMAKQFKNRILDVSNDDDDDDDDDDNDKDDPLVPSGDTDEVD- DEEDDVTSSSNKSDTNDQNSDRTRKRRKDDPDSDWSESNKKNIDKHLNL- LSSDSEIEQDYQAKKRKTSDLNQHENGSAIGLKSVMKPIKNTNRKPVSMDFN- NDSLTSSTENDPFPVNIHNPKEKLTANGTGHSTPLSKLVKSAFADKLQKVE- PSTILYPEVEQEPSQSERTPSLFSSEGSSESEAPLPEITTPGHQPMGNTSN- NVVEMIDTSSNLVSDDEVELVPTRIETPQLTEKDVKECEARVQELQEEVNEQ- HENSGSEFDVQKSGKVGNRHFKRALKFLSKLALLFLFCIFVPLFLGLVYRE- QRLLIGYCGHEVPSHRVSGNSFEIQKLDNLLQDYRPKCIPCPNGICYPYLKLC- KPDYKLAPSRDLFLEIIPAGKCVKDDKQQLVSEVVEKSLFLRAKNAKISCD- GKDDIESGMTEDALYQIFNEARAPWIRDDEFEDLWIKDLTEEPILWRLQSPT- DNNIGGNSNNIKTNDVPROKRHLPEKFIKTRNFRSTSKYIGMKCFEREIYQ- TYKFORPIWLMFLLVISKVIEIKNYRKKARIEELVTQMEKLFQKIKSMSDPK- ENAYLSIVQLRDLIFLSDVLDKYNQLWSEVVKYLEHNSNIKSNLTERIGEIMKCV- EWIGPMELNKPKDASAENKI
Saccharo- myces cerevisiae (Heh2)	Fungi	Uniprot sp Q03281	MDHRNLDPKTLKVSQLRRLVENDVAFPANARKPVLVLFEEKVRQRLQSS- PEASKVRTSIQKVVKSGAKNADRKTLKSKKLESSSESSTVKDENVETNKRKRE- QISTDNEAKMQIEEESPKKRRKRSSKANKPPSPQSKSDGKATSADLTSE- LETVEELHKKDSSDDKPRVKELPKPELNLKVSNEFLAQLNKLASAATENY- DHSIKSTDLSSIRIEETEPVGPSTGAETRNESEVMENILEVQPEVKEAKEEL- TKISFTDNQDEEDTSLRSSKNIRSPKGRTRHFIAKTKRGIIMKPIAHL- FIVLWNGAIFLSIICPIFLGLWYREQRIQVGYCGHEKPLKSLAISAFPQTERV- DSVLQAYRPNCELECPHEGICSSFMNVECEPGEYEPKSSILETYGIIPPKYCA- KDSEKEVDELVWKVNEYLKKKNAQHECGEGENLFESGETETKLYDIF- HSRPSWESQREFNHWNVLEILKDDIWLPLDFETNGKREKSKSNNTN- YYRSTSKWVTLQCHLEGDIEYITKYGSLFITLGLVFLIKIQSTLDNYQ- GEQIIKLVKEADKLDVKKNGEPEFLTQVLRATLLSDIPNIEQNQLWAQ- TKEIMKEQSENIYLLLEENGEIMTCWEWKE
Schizo- saccharo- myces pombe (Heh1/ Lem2)	Fungi	Uniprot sp Q10109	MDNWEDPNFELRNL RVIDLKKILHESGVFPVNARKIEYIRMVDRIRKKNLSSGPKH- LLSHLQKEENSNTSKASSSEDEIAPKYLYSPSPKSTKPKHNTEPILLSPOFIDK- SNIEPTVKIEMPHVSNNTFQSYSELSPNVETSLTMTKPPAHPASTPKFRSHK- SHRVAVPMFMDSSALHTSPAFSERLKLSSNNFSPQLRSP- KISHRLQTSATSSLPQHKRPFNTNPERVSRDIEFAPLDSARP- SESSSPYSEVDSAEEDDELQNYVLLQQRKESKLWSFIKVF- HDIKYANYRLLHNLRAFGISAISYLVHIFMILLGVVAIFLA- LLREKMTAGFCDSGASGSSASILGISFPLCRTPPNAICPS- PNYVECKPGYVLEPYWSSLGFWPSKYCVSDTSREESVNI- REECLSVLRSWNAILHCSNNSDLERNMSYNAHPYVADNLN- ISSDHSFPKPFALGLHDTLLERKSPTLGLEMFEDLFKASLA- VLSETNEVMDSKLYCDYSWAGIPLRCLKQQLKIFVWRNKVFL- FGILALSGVIFKLNFRTRISIVAKYLPASARFCVESLKRQKANY- QMSRSQEPVILIEHDILFHGNGPLEQIHMTKATARTLWEAI- VERVEQVGSVTRRESEVDGEWTRVWEVGTNTLDFQTDRSFINTSPLRE

Schizosaccharomyces pombe (Man1)	Fungi	Uniprot sp O13712	MEVPSYFDPDYDPSLRVVDLRNLTETQIYYPSTAKKAQLITLFSKLR- RAKNGLISMTLEQKQNVPPSSRSRPRRRVAGVTNNVTARISSKRKINMVE- ANDTEISKTQFEDNVNMGMLQDENVQVLTNTITISEESEFHASKIADSRNEEI- THIPFETQTELNAAVNLDNSMSESSIVQNLTKDSSVDATYDFSAEAGNIV- TPASKFLDYDQSYLVNASVSGDPTPVKVLNNTSPKSENLNQSSFLSFLGLENL- KPKFTSRSSSVYASPIKSSLSNLECNPSNLLSVRKNFQQSSDYLKSNKSFQD- LNNLVGLSTGNSENFTPENNSFWTHPKKNSSSPLPQSQSSSIFVEHLNQLYE- ANASIHRRPVPNPAFSTNFGLEASNTSTPEKKKFDQKPDSDSVNEISDLGTT- GIDRVEENISLTKDRQPKRPFYSLGFSLIFSFVKVNSLWLVLLVPLLGFVG- FWHQEVQRVFGCGVPAEPYPSLLYLQPGVLRSSIESAYSFAHSLGIEASCQ- PCPENAEFCGNRQLFCKEGLKASFPFLADFLKPYPRCIPNTVKVNVKVEEMV- QAFMSIIGKWWYKAPKEFATFESAKNLNGKSFVDNFKDRYMYKQDIDNVVL- KDFVYVLTLLNRLYNSKLRKVLVYLFPLFTLELWKLVRGALSKFPTNCLR- SVYSHTVSLMKYLTSAVISCWRILYLLIGLAAITGTVVWRIRVYAKKHVVKAGV- VCVSHCIAKLQKTKLSTLDFSVNPRVEVQLRSDCFVSGVADDKGLFELVHLP- LSIQLEIWEKVVSLGEMVSKVWVDSERLAKNRAWEWIGVSDIAL
Encephalitozoon romaleae	Fungi	NCBI XP_009265 646	MPKDMVPEDYLNKDFDYTKLTKQLRKIMYENGVEAIPPLTAKKSELI- DAYKENYKRIDVLMSSRRKNISMENSFQNVSPKKKYFEEVDPDVTNSKGED- PYKPYSEASMSDIGTPRKDVKEIKESTSRNFSKSSSFEKAPSNQEKRNS- VFDEAPGGSSLVAGMSAANNSTIGFSGSSFSFSSFIHSSSSGRSSISRNEKOGSS- SDVRKRAFTRSDPTLDKPELQSEHSSPRSSSNFTEINATRRFRSSKKNKWLVLII- GFAFMSIIGKWWYKAPKEFATFESAKNLNGKSFVDNFKDRYMYKQDIDNVVL- HEKEMSKKIKELKILERRSGDYMGPIPTKSVVPLKSSDPEIIRLKEKVEGIVIS- EMIVSNARVSMKTFFRYYFRMIMISTPLAIIIMIKIIGMFRKKKERLQAAQKIVKDI- ADVLVRQIYSTRANFPSSHVYIEQLKDCFGVDRKVVREAEKMIRRNNSNIRESCIE- GSAWEVWGPILYKPEFNGSLF
Reticulomyxa filosa	Rhizaria	NCBI ETO27229	ESSPMPANLDTTPLSASKKETTATKIKEDDEVKEAAYKAPLITKIEDDEVK- GATYKDPFSTSKVPSQHWNEGDLSAMAFPSNPQNSVLPNKITRRASVSD- KEGLGKVKVEQRFFYSQAQESQKHQLQDNMVFPKIDVFPYETNKTFDFTLEN- NEQIFTPKASQNAHKSVAASMTPLMDEGDDHASARLMOFSF- DELPHVKNQKRTDAIAATATVTDATASISVQVKKDETNLKSSSPRKT- IGKICTVLTIVRLFLFALVVVGLVIAQFQWDRYFTLSDSNNSP- RFVGDQLPECVESPLLYEGQPPEANGVRLCTDEELVLLLETKIHA- AAQLMLSQYGGYECGDETVRYKMEMTEIEDNVRDVLALSEAE- KSRWKEAFEKFDKAVRENKISGVIQYDDNDKTKCKTKVVKPLGCI- LSQWSKGHSVEIFATSTLLVIVSGFIFWYQKRKVNKKTQGVNYICP- FLCFLIYTYTRMYIIVIVIVIVLCKWTDGSGVVH
Reticulomyxa filosa	Rhizaria	ETN97409	MNSSALLSDGFSKGNKSQSQKNTNNGWCRCMVPISVMLVTVLVAIVN- VPNYAIENWNKWLLEGFLADEYTDSTLFCDTNRNTVFGQRDAEGNPLRC- VACPVALQCVNGIALCEKGYILEEYECVVQDVIVR- MAQEYIAQVKDGFVLPFNIFVFFVFFPQRGNYNNAIQAQKFLG ERKKGKHCYGD- VESDMSEALKRLVQMYVLI
LEM domain-containing proteins:			
Capsaspora owczaraki	Holozoa	Uniprot tr E9C9D3	MSAVSYNHLTKDKEIHELLRYGEDMVVAGSTRKLVSKLQKLQAAGASSPARSKSP- SRSKSPSRTAAPKSPARAKSPSRSRSPSRAAAAAEESPKRSTAATRV- GALAAEADEIAAVPTPRKSRTSASPSRSPARVSKTAAAGRAKQSEEFEE- VIVEKIKPISPAKSRGKTAAVDNASNNAAQIGYTLVAFVAFVLLLVFVDFGLQ- CLRQWVESFKASGAMAIENASTKQ
Amphimedon queenslandica	Holozoa	Uniprot tr I1G4M5	MEETYKSLYTTMLQEHEAVFAQQDLGQLRDFLAAIQNCNGSVFVHCGGRE- GIAARAFAMRLTHLGRKYWLWDDTTPGMHAGDLFILVNGSAGINMYMYE- QAKQTGAAVAMITGDPLRCPKDVQHTLFLPAMVYKGTCPDVVPSAQPMGNLFE- QHLFLFDVMAMMLETQMLSHAEMESRHRNIE
Nemostella vectensis	Holozoa	Uniprot tr A7RQP3	MYRVVLAIRKVKVEKPPTRPQPKLVIRKPVKGLPYDVHSLSDGLARQLKSYGQV- PITETTRPLYQKLVKLLQESKPKAGSQTPDKKPRPWEYSDNEEEAKEEVM- DAKPGNVTMENGDDHDEGDADDEIDAVPEPQLTKPYPRRTKKAVERKPEEVEV- SAPQCVRKRETIQETMYTTRKRTAAEKASSKAETVTAKEVTPNVPKQAKSSSFW- WKVIIFLILIVSIALLVVYFMEGNPVKPSIAAGRSH
Hydra vulgaris	Holozoa	Uniprot tr T2M302	LFNVGPPSKDHYRINMAEINVELSNAELRKKLTFEGKSVGPITKSTRVLEN- KLLKLINESEQISNSNSNHKKVVRKSNVSSRRSSLRSSESNLKFAPFSSDEDE- VLFLLKKSNSKSEKLLDLEKPKLVNKNELIEAKVQIPSTYTSASPIQDIKSSPLS- GKIKYNSNDYQSGCDLNRKNAIEYSDDEHFEKLEKLNEEKRQNIIRL SKKEIKSP- MENNVRTSPSPVVISRTSMNAIDKNIRGEVDACKAEVRNSFTSKKPSPTISRRVQLS- SYNNYNAKNKVEGFKDREANKNEHSEKHFFLIKCLKSLFSEKQIKCIACILLFGLSAV- LVGLYWYNEEDPLKLYLKDSSQGEKIDINIFIECLYKDLITRKGMEVCEGYSSGPDY- RREDFFPQALAEGLKQKDIQFNVIANDTWEAAKKSYTEFFRFNETMTVTSIGSYKPF- HCRVGGQALYAVLFRIFILITIFGMALYFYMKRKWSAEDKETRQVLYFVHRIVEVRK- HDTQCLNNKELSPYLPPIHVRDMILIPLEKREMAKTWQKAVEFLSSSDSRIRVETQR- ISGEDFEVWRWIGVRTPGIKKINSNFHLNSDQDKCWQGFADQIEKIVIRLPIITPA- PCLKIRCMHEGPDEREEDWVKKVENAVLEKCECENDGARILHHIHDINSREGCVYV- CDTLESARKAFRSMYGNWFDGRLVIVKFTLARYHQRFDPDALKCVQALSPSGSNA- VNSNNDI
Trichoplax adhaerens	Holozoa	Uniprot tr B3RIA0	MKSDLAKDPSVLTKDRLKALLSTNGVPLPEGDQRKDVYVYKLMQHIQNKNDTRN- TEPNVSEFSSDDTQSVKSAKRRKPAKRRKSTDTNNVFNIDSLSEELLN- LKKYDSAAGPIIDSTRNVYKRLKAFIQRSEATVKDGEFSDPDGVAEDST- SKKVVKDTPTPRRGRKTSRRSVSRKEFVKSEVKTSDSSDRKDDIAILNTSH- SVKPSQSNVESQSQDLQETNLWYFIIILVLAIAIYTGVIHLEIGRKH
Homo sapiens (Lap2β)	Holozoa	Uniprot sp P42167	MPEFLEDPSVLTKDKLSELVANNVTLPAGEQRKDVYVQYLQHLTARNRP- PLPAGTNSKGGPPDFSSDEEREPTVLGSGAAAAAGRSRAAVGRKATKTD- KPRQEDKDDLDTVELTNEDDLQLVKYGVNPGPIVGTTRKLYEKLLKRE- QGTSRSSTPLTISSAENTRQNGSNDSDRYSDNEEDSKIELKLEK- EPLKGRKTPVTLKQRRVEHNQSYSAQITETEWTSGSSKGGPLQAL- TRESTRGSRRTPRKRVETSEHFRIDGPIVSESTPIAETIMASSNESLV- NRVTGNFKHASPILPITEFSDIPRRAPKPLTRAEVGEKTEERRVERDI- LKEMFPYEASTPTGISASCRRPKGAAGRPLELSDFRMEESFSKYYV- KYVPLADVSEKTKGRSIPVWIKILLFVVAVFLFLVYQAMETNQVNFPSNLFHVD- PRKSN
Homo sapiens (LEMD1)	Holozoa	Uniprot sp Q68G75	MVDVKLSDCKLQNLQLEKLGFSPPGILPSTRKLYEKLVQLLSPPCAPPV- PRELDGAQSDSDEELNIIQLQNIILSTEKSKLKKWPEASTTKRKAVDYTCYDKP- SKGRRWAARAPSTRITYGTITKERDYCAEDQTIIESWREEGFPVGLKLAVALGIF- IVFVYLVTKENKSLFG

Homo sapiens (LEMD2)	Holozoa	Uniprot sp Q8NC56	MAGLSDELRLRELQALGFQPGPITDTRDVRNKLRLRGEARLDEERL-REEARPRGEERLREEARLREDAPLRARPAASPRAEFWLQSPASGSAYATP-GAYGDIRPSAASWVGSRLAYPARPAQLRRRASVRGSEEDEDARTP-DRATQGGPLAARRWAAASAPARLPSSLLGPDPRPGLRATRAGPAGAARARPEVGRRLERWLSRLLWASLGLLLVFLGILWVKMGKPSAPQEAEDNMKLLPVDC-ERKTDFEQCAKKAALLELLELYNLAIQAGNFECGNPENLKSCKIPVMEAQEY-IANVTSSSAKFAALTWLSSNKDVGWIKGEDQSELVTTVDKVVCLSAHPR-MGVGCRSLRALLTAVTNLFFWCLAFWGLLILLKYRWRKLEEEQAMYEMVKK-IIDVVQDHYVDWEQDMERYPYVGLHVRDSLIPQSRRRMKRVRVRAVEFLAS-NESRIQTESHVAGEDMLVVRWTKPSSFSDSER
Homo sapiens (Man1)	Holozoa	Uniprot sp Q9Y2U8	MAAAAASAPQQLSDEELFSQLRRYGLSPGPVTESTRPVYLKLLKLRREEE-QQHRSGGRGNKTRNSNNNTAAATVAAAAGPAAAAAGMGVVRPVSGDLSYL-RTPGGLCRISASGPESLLGGPGGASAPAAAGSKVLLGFSSDESDEVEA-SPRDQAGGGGRKDRASLQYRGLKAPPAPLAASEVTNSNSAERRKPHSWWGARR-PAGPELQTPPGKDGAVEDEEGEGEDGEERDPETEPLWASRTVNGSLRVPYSCR-ENYSDSEEDDDVASSRQVLKDDSLSRHRPRRTHSKPLPPLTAKSAGGRLETSV-QGGGLAMNDRAAAAGSLDRSRNLEAAAAEQGGGCDQVDSVPYRVRNAK-LTPLLPPLTDMDSLDSSTGSLKTNHIGGGAFVSDSPRIYSNLSPPSAVAASS-SLRINHANHTGSHNTYLNKTYNPKLSEPEEELLQQFKREEVSPSTGFSAAHLSMF-LLTAAACFFLILGLTYLGMRTGTVSEDEGELSIENPFGETFGKIQESEKLMMNLYKL-HDRLAQLAGDHECGSSQRTLSVQEAAYLKDGLPEYEGIFNTSLQWILENGKDVG-IRCVGFGPEEELNITDQFLQSTRPLMSFWCRFRRAFVTVTHRLLCLGVMMVC-VVLYRMYKRWTKEEEEERQMYDMVVKIIDVLRSHNEACQENKDLQPMPIPHVRD-SLIQPHDRKMKVWDRADVFLAANESRVRTETRRIGGADFLVWRVIQPSASCDKI-LVIPSXWQGGQAFHLDRRNSPPNSLTPCLKIRNMFDPVMEIGDQWHLAIQEAILEK-CSDNDGIVHIAVDKNSREGCVYVKLSPEYAGKAFKALHGSWFDGKLVTVKYLR-LDRYHHRFPQALTSNTPLKPSNKHMMNSMHLRLTGLTNSQGGSS
Homo sapiens (Emerin)	Holozoa	Uniprot sp P50402	MDNYADLSDTELTLRLRYNIPHGVPVVGSTRRLYEKKIFEYETQRRLSPSS-SAASSYFSDLNSTRGDADMYDLPKKEDALLYQSKGYNDDYEESEYFTR-TYGPESAGPSRAVRSQVTSFDPADAFHHQVHDDLLSSSEEECKDRERPMY-GRDSAYQSIHYRVPVSAASRSLDLSYPTSSSTFSMSSSSSSSWLTRAIRPENR-APGAGLQDRQVPLWGQLLLFLVIVLFFIYHFMQAEENPF
BAF:			
Homo sapiens	Holozoa	NCBI NP_848572	MDNMSPLRAFLSEPIGEKDVWVDGISHELAINLVTKGINKAYILL-GQFLLMHKNEAEFQRWLICCFGATECEAAQQTSHCLKEWCACFL
Salpin-goe-ca rosetta	Holozoa	NCBI XP_004995959	MAERTTSKKHDFEVFRERMEDKAATELAGIGEKGGEQLAALGFATAKQVL-GQFLVMGEDPDPVFGDWLQQNVSKLNSKHRGDLIFCLQEWIKRNL
Capsaspora ow-czarzaki	Holozoa	NCBI XP_004343428	MAETTSQKHQNFKEQMRDKPVNQVAGIGEKISGELNDQGFKYAQLL-GQFLILNKNEGDFLDWFKEVAPSANSKHRQDAYECLSKWCLEML
Phyto-ph-thora para-sitica	Strameno-piles	NCBI ET-M47535	MADGYDPQKSRVAEDTLADFLRAPLTGDLTEVPGIGKAAVAKLSEAEDGEDAVTNT-FQLIGKFLMLKANSDDNDGVIDCAAHCDAFWFLWLSKSGITAYRSGIVMA-IAEKVNTMLPGIYDAAEFQ
Thalassio-sira pseu-do-nana	Strameno-piles	NCBI XP_002294632	MMDFIRGSVTGDITEVPGIGPAAAKLANVEGEEGITNTYQLIGKFLMLKGPDT-DTNKVESEFHECKFWYWLQAMGISAHRSIAVKAIAQKVNGLLPGVYDPLEYED-DEEE
Ectocarpus siliculosus	Strameno-piles	NCBI CB-J48376	MAEVGYDPKRSVSDDKLAEFIRSAVTGDLTEVPGIAARGAEILADGEGDDKVT-NTYQLIGKFLALRGPDKTDHEVDSVEHCDKFWYWLQAKGINSNRSGIVNAIAEKVN-TWIPGIYDADSYTDTA
Aplano-chytrium kergue-lense	Strameno-piles	jgi Aplke1 103957 es-tExt_fge-nesh1_pg.C_610016	MATGFDVNRSKVSDDALAEFLSKPVEEDITVPGIGPKAAEILAKNCENDPGIET-TFQLIGKFLSLKKGKMGSKHECDAMWFLQAKGVNSFRAGIVHSIAERT-NIMMPGIYESNED
Emiliana huxleyi	Haptophyta	NCBI XP_005762802	MLISMVPLGDBGIRKLNAGYATTYQLIGKFLPLNRNEEFLLKQGNTRPHVH-NTAPLAL
Guillardia theta	Cryptophyta	Uniprot tr L1JLR7	MPSCSRSQLALCQVGRVSCDGSDDTTAADKLKEANIDTAEKLMGNFMVM-GRSEEMAKWLEDVCEVRSVEAKKISAALVEKGDKICSM
Lamin-B receptor:			

Strongylo-centrotus purpuratus	Holozoa	NCBI XP_786536	MPSTTSFGDGDVMSRWPGSTLWFKSKILRVSEEDGYKVFQEDGTETEEIPLTDVK- SESYFTRSRSRSRSPRRSRSPARKSSPARTQQSRKPRSPGRK- PAAVKKEVETPVSRTQEKRTSEKLSRTQVTKTEHHSYTRAQTRS- GKQQLPELVKMKAGKVAKTTHYEFGGPIGTFLMIFGLPLVYVLYFTCLPSCGCKLV- YNNPFLDWRDYDQEAFLYVGVWFLFQAILALLPFGKVVQGPRLSRGQRLSYR- NGLFALIVTCATFGGMIIYKCPVTLVDKILPLMTASALFLLSTLLYIKARCGPNS- ALATGGNSGNFFYDFMGEHLNPLGSLDLKFFCELRLPGLFWALINACDLTKVW- TEFPDNPWNILVCFQFLYVFDALLYESAILTTMDIIQDGFGLMVFGLDWTWVFP- TYTLQARFLADHPAPFDYCLIPVALLFLSYGIFIRMSNSEKNAYRQNPYKGNVAG- LQTIPTDGTGRLLVSGWVGFVRKPNYLDLLMALSWSLCTGFVSIVPYFYPIYFFV- LLVHRERDDASCRQYGGAWTKYCATVKYRIPIYI
Petromyzon marinus	Holozoa	Uniprot tr S4R4L3	MVHSKFTIGEEMARWPGSCMWYKQVLSYSEESSIYN- VKYEDGTELELETSIKKIGSFRTPRKRSRSRSRSRSPGRRSRNSPGRRRPRAR- SPIRSPKVNDEKAKSRRAKPEEPPVQTTNHTAVMQVPQAITKVKKEISKEKN- GAETGKQAAVVVEKAMKEEKEEMKKKMGAVASESSPLVYRLSRQGIN- QKPETPPVPELPLKSSKPEEICAEPGLGVKLDPAAVCGSAGTALLPLALPLAL- LYVLHVCCSDGCSLLSPPAPPGLASLWDMCTLVLCGLWCLFQALLCHLPL- GYVTERVPPNTEQRLKYRLNGIHALVLSISAFSLAKYKEAYVYVHDKFLQLA- VSAVLLAYLLSAYLVRARACASHIAVATTRRPGNIFYDFFGRELNPRIGNMDE- LKSFFELRPGFIGWAIISLVMVLEDLQKNEAFSPALLLYVALQFIYVADVLMW- ERSLSTGDIPHKYFGLLAFNLAWIPFASLLPAYLVVHHHQLTAYWAAAI- LLYLVGYVSRRAHTQKNAQSDSAGDRGKPADSRSSPVLGSWGRLRHP- LGELMLLAWSLPCGFSHPLPYIPVYVSLVMLLHREAR
Callorhynchus milii	Holozoa	Uniprot tr V9KH74	MPAKHYEDGDMVGRWPGSSLYEYVQVTFGDKISQLYGVKYDGTLELKE- DMKPLNYFKSKSRSRSRQRGRSRSRSPARAPSSKRSTSR- GRETKEVEKQIVPEVKPAPQKLQENSNKSTTLVAELNEEQENHTKT- STYKITEHKIRSEIVTDELSRYNLRTRKEEKEAKPIKLETDKDLYKDDVPQ- TQLEFGGRIGAFFMVFLLPATVIFLLMCGQEDSSILSFPQVPPLETLWAC- LALGIFVWFLFQALLYLLPIGKIVEGIPLNPKTKYRINAFYALLSAAAVG- ALYNGVNLVSIHANLQVAISAMVFLSLLSVLYLRSWRVPTDELAPGGNSG- SFIYDFGIRELNPRIRNFDLYFCELRPGLMGWVIINLGMVFAEMNVQNL- DMPSLAMILVNSFQLLYLDALWNEEAILTMDIVHDGFGYMLAFGLDVLVW- FIYSLQAFYLVKHPSELTPVAAAIALNTVGYIIFRANSQKNAFRNRP- KLAHLKTIPTATGKNLLVSGWVGFVRHPNYLDLIMALAWSLPCGFNHVLPY- FYVIYFTCLLIHREARDEHQCRKKGVAWDKYCQQRVYRIFPYI
Homo sapiens	Holozoa	Uniprot sp Q14739	MPSRKFADGEVVRGRWPGSSLYEVEILSHDSTSQLYTVKYGDTLELKE- DIPTRFRQKGGSTSSSPRRRGRSRSRSRSP- GRPPKSAEVSASASHQADIKARREVEVKITPLIKPFGNSISRYNGEPEHIERN- DAPHKNTQEKESLQSESSYIATQYSLRPREVEKLEKIDSKKEEYVAKELAVRT- FEVTPIRAKDLFEGVPGVFLIMFGLPVFLLLMCKQKDPSSLNFPPLPALY- ELWETRVFQVYLVWFLIQLVYLLPIGKVVVEGTPIDGRLLKYRLNGFYAIFLTS- VIGTSLFQGVFHYVYSHFLQFALAAVFCVWLSVYLYMRSKAPRNDLSPAS- SGNAVYDFGIRELNPRIGTDFDKYFCELRPGLIGVWVIVLMLLAEMIKQDRAV- PSLAMILVNSFQLLYVDALWNEEAILTMDIHDGFGFMLAFGLDVLVWVPIY- SFQAFYLVSHPNVSWPMASLIVLKLGVYVIFRANSQKNAFRNPSDPKLAHLK- TIHTSTGKNLLVSGWVGFVRHPNYLDLIMALAWSLPCGFNHILPYFYIYFTM- LLVHREARDEYHCKKKGVAWEKVCQRVYRIFPYI
SUN-domain proteins:			
Ostreo-coccus tauri	Plantae	NCBI XP_003079 402	MKTTAVLNKVPFAVAASLVIATAGCATAALGWHARETALALASARVSATSVDLGPA- TRRIRVLEEQVEALQRNREDRGRKSAERVVREALLDAQVKALGTQKSGKAVGLFKR- DADTPALKKDLKAIETTLASLTKSQGAFATSQAKELVDAVDALRKSSSGYASS- DVTVELTKVSVTELTNKTSGSMKEVAALRKALEELASTQKSVDA- RATSEVKSRLDDITSLKNIQETIYASKSALQTLDVSSSELKSEAA- IAGLSKDRDSYATATQLKLLDAQVKALGTQKSGKAVGLFKRDADT- ATLKKDLKAIETTLASLTKSQGAFATSQAKELVDAVDALRKSS- GYASSDVTVELTKVSVTELTNKTSGSMKEVAALRKALEELAST- QKSVDA RATSEVKSRLDDITSLKNIQETIYASKSALQTLDVSSSEL- KSEAAIAGLSKDRDSYATATQLKLLDAQVKALGTQKSGKAVGL- FKRDADTATLKKDLKAIETTLVSLTKSQGAFATSQAKELVDAV- DALRKSSSGYASSDVTVELTKVSVTELTNKTSGSMKEVAALRK- ALEELASTQKSVDA RATSEVKSRLDDITSLKNIQETIYASKSALQ- TLDVSSSELKSEAAIAGLSKDRDSYATATQLKLLDAQVKALGTQ- KSGKAVGLFKRDADTATLKKDLKAIETTLVSLTKSQGAFATSQAK- KELVDAVDALRKSSSGYASSDVTVELTKVSVTELTDTKTKSGSG- SASKNLRDLQVMERIEGYINAIPKDTSMKLLKHKIEKAATLWF- ADRTGRQDFALSTGGRRVYVHSQLTPFVARGDGPLMSAVSF- LRSVGHPKSDEWMLTPSMEQGDCLALHGSYGYVDVRLRQP- VKVDVLEHTNSLNAYDMHSAPRDIQIFGWYAHGNCKHSPK- PKSLIAMGNYTYHTAGDSVQSFEVSAPTHVDHVRLVKNHGHARWTCYRFRV- HGIPSS
Arabidopsis thaliana	Plantae	Uniprot tr Q9FF75	MSASTVSITANTAAATRRTPILAGEKSNFDFYQSESLANGGVGEAGGTSRDL- SRGEATLDRSQDQLGPVTRRSVSAATGNTTATQRRTRKVTATPKSEKARW- TVVRFKALQGLALLIIVGLIQLTRKMLKASSPSSISSYETEMAFSGLESRI- AEVDGLVKATTNSMQVQVELLDKMKEREAKVLRQEIERSKASAFQSELK- KIESRTELESKSVDEVNAKPPWTKDELERIYEELKGNVDDSAFSEISID- ELRAYARDIMEKEIEKHAADGLGRVDYALASGGAFVMEHSDPYLVGK- GSSWFATTMRRRHTNAVKMLSPSFGEPGQCFLPKGSEGYVQIRLRGPI- IPEAFTLEHVAKSVAYDRSSAPKDCRVSGSLQGPSSAETENMQLTTE- FTYDLDRSNAQTFNILESSSSGLIDTVRLDFTSNHGSDSHTCIYRFRVHGRA- PDPVPVVGNTLDDQSSPESE
Galdieria sulphuraria	Plantae	Uniprot tr M2YAH2	MNLSPLTRAKARELALAQSPSRVLTPELYTSPRRSLDNTDDSDAGLSRRK- TISRSGRKSNTKYFSSDD- KENISIHRSPPKREESSLSQKGGKAKESYQIQNLGNFQKDGTTSYSTLSSP- PYFSKTSNSERVEDSSPLTNSLDPKNSSLHSNSNERKNRSMQSSIDSKK- RNHSHKANVFKLLILFLAFILIGTFIVSLRNTTLESSRRQFTSYRNMGLWFR- FYRPNKNERDSFETALSFMTQITDLAISLEQKKNHPEAILHSIQDIMEKTE- NILEKIPYRELREWRETYRQLFNFINLHEVNDISSNDIARLLESKRLRYKLTQDYA- LSSEGAKVICSTPSKLRQSQRYWLRMASFLDPSINTTVLRYPKGPETILSDN- VSVGNCFWAFAGKTATVTIQLSKTIQPIAFSLEHIIINLDSKNVLPKFLR- VYIDKIRLEIVDNY- GGAYTCLYRFRVHGHEQESSLQ
Phaeo-dactylum tricorutum	Strameno-piles	Uniprot tr B7FQ97	MRSKLSQLLLGVVALGTGCVLLVSSNADDTGLAGERHAAYQALASAFSIVET- WHRATQLIDRTVDLHVKYDAVEESVFEAEQAIRLFNAQQLKVVIEGDAVHEAL- ERRAIQQQDLEDVPMPLTKDEFNRNAILDSIVDPSDVRMIEHWVVDYIDRVL- NERATPPTQVTTPTSPHSCVTPQKAVQEVHAALVRHATDGGIMEDHARGARI- VHEMTTTTTYPPQSHQRLGNVWRRRFPQDWESFLPSGWEEVDARVPLF- FSHTFGSKAPVSAKPELILPITPGACWPMDGSGNHVTLALAYPVAVSAITID- HVSKYLLNEPSEQLTAPKDFRIVAYPPCIEHCHGLSFDVNDPFDLAQGTFER- DGTTVQTFATQNLDPPLGREHTMEEGSCSAAAAATTCGVPDANQGLVAAVQV- QILSNWGNEDYTCYRIRRVHGESADL

Phaeo-dactylum tricornutum	Stramenopiles	Uniprot tr B5Y466	MSKRKGEDLPVSPDDFSPPGVKTRSIKSGKRQRRIQVAEES- DAEPSSPSVAFAFNPPSSSSSIPDRQSNNNNNNNKDDDDRRDSS- DPEDPMEDSNRRRTNPSKTHDARRKAPPAASGSPKEAGENVPLSRRLDYRET- DTRKESPTLLPTESVLRERDVLENGHDKTTDDVDAEEHRRNNDHDDVRPHVS- VHIRVYQVVELVAGPTSAQVETEMDVPVNTDTASAGSRTWMLQSWVWVLLI- LAWHICFPPLAINPGLVTVSSTGNLYTTVYRAQAKLAQAVQGLQTVQHQLLARL- TARIALAQEAENAFRSQQLAAEENLARLEQSWETEASSVMERLEKEEATARTLNH- WIEQVLEVPDDEEEEEEYEAQVVPPEIRNVLGPTQDALLDSSFITLWDVPEPVI- CETPDVSLGAGGLYKEDVEQAISDLITDIVQVDEEMEMVRKVVENYLDTKAG- DAMTTTTTADANIPPLDGVVDADALKKRAFIDGRMEVERADQGLDGLDASLLNGA- RIIRVGDRTSMLVDQLPVFNRLAALLSLRFYGHGPEAALLPTYPNPAALGQCW- SFEPASRRSGPFGVLTVQLSRPIHVQSVSIEHPPELTDKQTAIRSFRIEGFE- DTQTHGKAHSLGSFEYDGGKQLRQDFVDRNVPRLQSLVLDVTDNWGEPYACLY- RFRVHGQE
Phyto-ph-thora infes-tans	Stramenopiles	Uniprot tr D0NXG7	MADGNTYTRRLRSRRSDSTSSEEEEDPQRVTRSGSRRYGIYTPPEVQRTLELRS- DEFEDDDEEDSDFEELDDYGETVYRSTYRPPQVYEQDVVELEDVHEDDE- VDEIVQETPEHEAQRSELKRAAGAAAYFKSNKVDRMWQKVTDSKAMKTKSKYL- RRFWRFMLRNSFMVAVNLWLLAPLCCFVVAITVPHHLLTAFQYVDDLSSWGGH- GNADAGFEKAGMRSVQVEIVDMKLVGLNEEIGMLRQTVQTEHEIEALKLWTLFL- LDLEQRQKFSLSEPDASAINVHIEKVTKTEELWEKIIDRTSOLQDQLONATKQSS- VISSVLKEQEEKMDSVQTIKETAAPAPDAASENARAMKKEFTQWRQSGFIELOS- EMQRKVAIESRMSRVLQDEKDALSSADALRGLDADTPGILRVIEVAVQAVEIKKT- GRVDHAALANGASVIHSERDLYQDSSSPVLAQLVGLSDSDGDSRFTSPSYRR- APAPFGLQLSSGENPWVLRHNGRPEALSETMEIGSCWGISGSSGRLSVKFA- QQIVADAITIDHIAQIASDFSSAPNQFRVLGISGHLRETVELISFGNFYSASNGPAS- QTFKLTSSLQRSIDGITLEVLNHNPEYTCYRFRVHGQPA
Guillardia theta	Cryptophyta	Uniprot tr L1J5Q2	MSSSHFPAWTSSTVTPAEDSKNKNKYIALGALAALLVIAAFAMSSGSKPAISM- DADTHKLVSDMNKRFEEMKTYNENVGLLKTAKQNGQVVDKFNDEALSQIKST- MAKTEESMKKISEQRNFESVKTLSDDIQLKRYMLDQIEEQKKSSEAKTADLEK- SLSKLREKAKMLSDAESLNSRLQLSELTAKTEIDSQMKEVNSLLATERKAIEKKS- EASQSMVAAKSEILAMAIKMNELNEKIKQAVQLVQRQSDLEKQIKHVLRAL- FQERIGRIDWLELVNGASITNHSSETFPACKNTSVFYLKCSSTQQLNPKTDIRSHVA- GLSMIDSPMMLGNCWAMSGSQGFVEIRLSKLLRVTEAVIQHIPKRISPDFRSAPRG- FRILGSLSGAETEEGSKVLEGEYEVREDASEEELLPVQTFVAVKNSPLVDITRIEY- LSNQKQDEFCTCLYHMLRHGEEATPMVSS
Trichoplax adhaerens	Holozoa	Uniprot tr B3S288	MALNNQDFPQVSKIPDNEGLFAQGSLYCSMKRDSLGSVA- GRKSVRSNNWYRKYGGMTYDHGPLRALIGLVRVFIITISALIIYDVAIILIKIGIK- TAFMLPLYPNLNKTSHKTPQTTLNALSIAIIVGLTCTIAVFIYVPTFQK- IQSVSSTLTKSSGDPVIKDNTNNTNC LNGIGDIRIRELEAVIMQLKSEIQEIKSYPRQKVLSDITFNIPQTLNRYVLMINYLE DNLKELKSMISKEIVEQVNSALKLYDEDKIGLADYALYPAGGRVISIGN- TKPYNLSEGKPFHSPNIMIQLPQGNCAWAFEGRMGEVITQGLDDIYAEHKKLLGS- TFEDSNVMNLQRFTVQHFPNRPFLIKIIFTSNHGSSYTCVYRFRVHGFEKSTL
Monosiga brevicollis	Holozoa	Uniprot tr A9V320	MSVVRASPTPQRRTGPRTPPGFATNTLFSDDADRSVASEYENQAE- FASVLHSRINIATPNTSSSIRHVQTTTELRRVDNDLGLSEDELELGNETLLYDDDG- PIVYERRVHRVLRNTEDESSASAOETESVGRSPHAFAPRIIDSALLAARYQERT- SLDDVDKALQSLIQARTQAEASHSQTERPKGHFTICKQIKACVSS- NLIIMLRDLARLEAQQTNFTHQQLHVRDALNSQAQIMRQTMFT- MAEAQVFNKQKASNASRASAACPKCPSCTPTCPTCPNCPNTNA- PTKACPETAAPPSPVNLASVARGARIETHCARARSSWRSR- WESLFSMQPDSARAMLSDSMEPKQCWAFRGAATAALQLAAP- THVESVALSHVAQALPASHNQSSAPRRFRVWAAGNSAAPAMPA- AHHMLLEADFPARSPQSFAPAVHQHEARFIKLEIQSNHNGNEYTCVVSFQV- NGRATLVA
Capsaspo-ra ow-czarzaki	Holozoa	Uniprot tr E9BZX3	MAPRVGPESTPSARRASILVAADAATDEAPRSVRRSTRQLQQLQQQSHSQA- GAGQHESGDETVSVSNTPARARAPHSAVAPRSRTAAAAAREDTQLQLQVPAV- GRASRSLSLRAAHDNSSADLTTPTTKRSRSARSASTATAAAAAASAAAAD- NGAYAVNEDEAPSTAASSTSRWTASSKGRSESLRSRALDLSFTTNAOSPDLDL- SVLHMHSAMEDADEDEGDDAENASMTAARARLSPTRSPRSQATKSPAAAAA- AATDNRRKGLSPTSFSSPSSLRSGSPKAAAGSTPINHSQSPATPSTLRTRIVNVL- TSPFRRAEAGDSSKSAASSPGLRGAALAEVHRPQRSADDDVFLGLDSS- DTEDDGSKSDKSRARAPNVAATAAAGSTTTTTFVRLQPRSLDGLRISGIL- FYLFTTSMILNVFLAKGGQALRRIRGGKPSARSAATAATPSPQSPRSLSYS- PHYRVGRSNDLDDASSDEELLHDQHRGQASIRRLQGEVVPALARRLWKGAL- VALLILVLLALRRLPAGVQHASSPLLDHQVDDRLKLVQASAHDGNLASGKSV- EELVQLVQLLSAELTLEKFLAAQKDDRRLEQLSADAAAAAQKLTQAAQK- AQPVVDDSAKGANLQAQAKELDTLKGEMERLKAALDKESNNYQADSKLQ- EIRSCVTLDVLESIAPQVNSIVERRLSEATLPKSTDSVTAVQQAIDPTELRQ- LHIEEVAARPTTLEEEREKELSNLTVRVAIMEPRNREALDLEVSRLNTRVSTLH- SEVVQFQQQKQHQQQQQQQQASVDSQVRLALREAIYLDADKIGVVDL- ALASAGGSVSECTASFSVPLRALFGMPVMMQSTSPLLALQRDINSVGNCAW- MAGQSGQLTVRLARSISITGFVSEHVAKSISLNTMKSAPRAFVRVALNACDTEG- TLGQSYDIDGTPQLQFAVQSKPSRSFEYVRLIASNYGEEDYTCLYRFRVHGVA
Encephali-tozoon in- testinalis	Fungi	Uniprot tr E0SA25	MNRDRDLQVKRTPDSTLNLGMDETILQAAAPKARANKSVEPMEG- NGYQGERKGIKDYPIYMAIPIYVLFYMAIKRPMDSMILTNLMEEINILRENSRIS- SQEIMMKIHEVNYAKIEEGARIRIESMSQLFSYGFGRKHKEPSTIFDENVIGIE- CLAFKGAQCKFSIDLEKAAISIKIGLHPVTKDTSIAIRFEFVFSNPSPEGNL- LLGRFEYDSTCGQTFWEETPISSVEIVRNSNGNKKYTCIYKLYVLFNGK
Batracho- chytrium dendroba- tidis	Fungi	Uniprot tr F4P736	MGTAAADSGEVSGDLSLTDSTSGVRRSKRAHRNINYSPIRTASPKRIGHPAAN- HESSIVNTIRRTSISSTASDCSRSSRTSRATHVTSISSNNTLSPITTPARSYSHRN- VLFSTPQDPIPSDLSVLIKPDTEQPTSAWNHPFRSTNNGIDSGHEINSTKSF- SSTESTKRTPTPKIARHKMAKSNPAASDSELDNLIPTKPNSSQLNPSKTPESFV- EITPSKFLTSFFSPFRSPMNLKHILENVGMTPRTHKRKVAKSWLGRMATSSDQIE- VEYEDLCKSDDELGENRILFPDIFAEAEVIEEPGLLRFIDWLSDFYIVQLVIPFS- WIASHALLIIVGCVCTLMWNVIVYLCGKPLGFIWTLFRVAGFIFFTFRVLGYHRNM- CVAGLLCGLGVYLLTKPGHSDILLASLDPVNSPQLQSMFTPNISMIFELHRSFT- ASTETEHEIENSQDQSNTPRSPENWWSQKQLVMLEKRVATVEKLLLVTOGLV- EVAGDMKSQHSRLEKSEIHTIAGQDSLKDTLSAQIADVTQIFDLRNLTDALHDAT- EKTITQESAYIKHLHDSLEIHKSSIVELEKQIHEDVQNLVSNVNEVSKGLAEKLTSE- TVTRLGQLEAYGSVDDVVKRVMdTIRTSTDFPGFIATKSSSETSEIELPELWKA- IESKLNMFQDKDQIDVSNTAEARLKTFLAQIDDIRVYATNADVDVRIRMAKDAI- KDYTQSGQNDASLAARVLDKMDVAKALLQTKVEELNLEQETREKLGQAFSGTK- KEYTQLDLHLKTLTEQIASHTTSLKEVMSQDRLGVFYDNQKREFELIKTTVDTS- TWHHNFLEQNRQALSTIIEHQVDTHPNSHILTQEQTLSAIEVKLESIVSKSSQDMAL- IMNRLSDSVSSYAHKGDAPLSDRDATELHRIHVASALEEYRADVLAIPIYALE- SAGARVVYNTSSTFTTHFKPVLGLFARVMGIRKASGRSPSTALTQDISPNCW- AMSDSSGTLAISLAEPPIPTDMTIEHSHIQTSISDRHTSAPRQIELWAVDAVEFAL- DLNNDQVRLQTLGTLSSNKKTKQAGILLGDFEFNPMTAALKTYPLHRLNCKVNM- VVVRIKNNWGNPKWTIYRVRHGRE

Reticulo- myxa filosa	Rhizaria	Uniprot tr X6M7T6	MRNKDLLLLKQVYNNKQINLRMAPKNLITCNIWLTTVKKTFCRCM- FAAKKKKKKGGKSDSNKELIKEWTGSVFRSHQHFLLGISIVAVTYMFTYGMVTL- AANMMDNKKYARAFSLNDFDMYTDGQSTSTILDIEIKDLGSIHDLDFIQHSGK- PITDNSAKRSDSKSTMEPNTVNDSPINWSQIHIGSKVTLYLDAGDVTVAIRN- VDVEGQRIFVKYINCWWKCSHWIDLKSGNLVLYAKDNLGISNGKLLKHQTKVL- DDKISVEDIITNGSFFFFRGCPCENEILKSRGIQKLESKIEHVLEQSKPSSQ- NNGAKGREPMDFLVLKQVVEEQQALDLLRGSVDALRLKSSAEHSSLPLKT- SSTPTTSTSAKVGGEIENRPTLQEFKRELTNLFKILLIFLFTIITKLNKYHQ- EVKSEQQLAIEGVITSVKELRSSVHSLQQKAMEKRKEKDGAEKDNETLL- SLSSPNDWSYQVIKHSVMKSETWIDRLLDYVTLNGWRSYRQQAIMEVTS- KTPGECIPLDFQHVLSNNDDQLQKLDLDRNPNLQYALHVELYRPVVIKQ- FSLFHLHSDLIPDDITRRSAPKLFHVLASNDKERWFNLGLFVYDYDDVKRTS- PDRFDGVHRYVAQHFDGNGYNGALKFKFVFRILTNGGSDYTCYRVMVHGTE
Bigelowiella natans	Rhizaria	jgi Bigna1 89433 est- Ext_fge- nesh1_pg.C _490068	MPTTRRSSRRSASEQRELPGDAQSKRKRKTPGRRSSSTSSSTRKRSTRKSS- RKSSKKAPLMESVPEDREQEDNSQVVGIGVNLVESFDDVSNEGNEEGLEGIDE- VEEEEEQEEEDLSSAAHAGLRRRVRQGRDTTTTAAPTAPRPHHTSPYTLITVTP- KEDRDRRRKHIFISIIILLLLGIYYILQLPPSGHNREAIKDFSKMAELNRIKQK- MKTGLGSDYDRTVVKLKSEVDEMMKNLHEKYENFGNDNTDVVKATDFAKEQI- RKNPKFHKAIKEEVQKRMNEYTEKIHKKIDQLQAKQNGFRAEMEKQESALSTL- KANVDTTHAKDYTEKKMKEIQSELEKMKREFMLEVKNFKKTLNQPAGEQMSQ- LKIWTEQLAGDTLDMINKYLEKFAADRTGIVDYAMAATRGRVVDHSPYTKQK- KRGVTNIMNFKSSQPISRNPPAQVLKPGIEPGNCWSMEGADGFLRLIRLKR- VPDKVSLHASPILLSTGSAPKDFEVYGTNAGKYTDDMEKTLGRGTDLNG- KSSLQQFNLKAGGKSFHHTLFKVKSNYGEAYTCYRIRIHGSEL
Bigelowiella natans	Rhizaria	jgi Bigna1 84919 est- Ext_fge- nesh1_pg.C _10366	MSQRTRTGLWSLIRARFLPSQGNMVEVENEGNLRSSQPERKRRRRGSP- SRAHTRDGTLRPLNPSHRYQTRHKENMLHMAASLDVAVRFRRRRSRSP- SRESKRQRQIAAPTGRIPGTDIIVAGMRGNIGNLPEDQTEAVEQRLAQLHQP- INGDEDEINVEATPRRLSSRLKKLKSPATAELSSLLGEEAQEDIMVQGHQ- TKHAHNLKAPATVARRGSPDTIFELLRFFIILLLLTTALVARIYYKSTIKGTGEY- FNDVDASLDENLARSQRQISKWNSEERTLRHEVTSGGYRSQREIKENFLQEV- KKSKEWVDKDSQILKILLEGNVIARDHISLNPLEPLGSQLEKMEGIFDM- KNQFMVEEDSLLERVEKAIDEVINSKRSIAASKFATMKENVKSEFADSERILAAI- EKYAADRVGEIDYASKLWNSRIIGKSESASNTAKSFLPRVWSQKDFVNPSP- RKGPVILENNMNPQDCWALNGPAGYVTVLGRKVKYSKYSVQHSKGIAPK- STALKFKFVWGLSGKDKESIQIGHASYQFGSTFSVPTIQYAEIYCEWYDIT- IEFTENYGGSYTCYRVRVHGGESARNLDGKTMNERSERLNNVTPSNHKS

Edited alignments used for phylogenetic analyses:

Lamin:

>Dicdis

SPNRAAQRLREKDELSLIHNRKLSALKKLESAAETELEKKNQEYEELDQKHTATIKQLKQR
SDQVE----KQLIEEQNQNSDLTSNRNILEN---ELKSKESVWKKEKDEILLKFQESINK
LNQENSLQSQLKSEIVSKEYEIDGLKSEINRLKDDLQYRIRREGEEKDEEIKKYKFEKKE
EKSSNAM-----NKKENELNLIQAHERQIEDMRDSINREWELKAAQMMEHHARVDSFN
EEKERIKSQMETLNGQIEDINIKNNEYEDRIKEMNVLLSQKDNRKQADLKSQKDGQIALLO
IEINTKDNKCNTQIDPEIPLDPEINSLKELVKGFEKTV-----NPDTVSFSLVDSNQE
FIKLSVHGDNGLSISKWRLIVVKPKSGFSFPDGIQPFKGSVTVVWTGRPRQGTPTENEFY
WARTELWTSEGTIVKLVSPSEET--TTVTLPA

>dicpur

SPNRNALRLKEKDELNNIKSKLQCLCKKLESTENELEKKNKELDDMEHEHSLAIGNLKNR
NEHVE----KQLIEEIKTKSELMSSRDALES---ELKTKEASWKKDQDEYVTKLTETINK
LNQEHSTVSQKTDIAAREYESDTLKTEVNRLKEDLQYRMRESEDKDEEIKKLVKEVKEK
DKLHSAS-----SKKENELKHLIEAHERQIEDIRDSINREWEVRAVQMVVEHNARVESFE
EEKQRYKQISLYGAQIDELNIKNNEFEDKVKELNDLIAQKDSRKQADLKNKDSQISLLO
VEMNGRESKCATQTEPDIPLTREIEEIKTLVGVFEKNV-----NPNTVTFAVVDTNQE
FIRLSVHGDNGLSLSKWRLIVAKPKAGFSFPEGIQPFKGSVTVVWTGRPRDSTTTENEFY
WSRQGLWSQEGTIVKLVSPSEET--TTVTIPE

>dicfas

GRSSTAARLKEREDLIQLNSRLRDLASSEQKDNEIRRLRDEYITTSRQSTDQISNLEKK
LEDIE----ARLGNESKLTVELSSQLQLAHQ---DARSKEQGHQAEIQNFQNKMDETIKK
ITEENSQVVALRNDNSKKVMEIESLKSQVRLQAEALLTRSKEAEKDAEIKKLTTELKEK
EKLOVEY-----SRKESQLQQTIEGYERQAEDMRDSVNREWEIKCVQLVEEQNAKANSFE
DEKEIYKSQLVYQGGQIDDLNKNHEFEDRIISLQKDLDTKDDRKVADSKKDIAMLO
DEINGKDVKCNSSPEQDIPLTREIEQLKSLVNNFERRV-----QVNTISLLSIDAISE
CIRLTVSGEDGVSISGYKIVVTKPRIGFSFPENIVPFKGNINLFTGKTRPGTTPPNEFY

WSRPNIWESEGTIVKLLTPTDEVLAVNTLPS

>diccit

SPNRAAQRLEKDELSTLIHNRKLTALKKLES AEIELEKKNQ EYEELDQKHTATIKQLKQR
SDQFE----KQLMEEQKQNSDLASNRNILEN---ELKSKESVWKKEKDEILLKFQESINK
LNQENSLQSQLKSDIVSKEYEIDSLKSEINRLKDDLQYRIREGEEKDXXXXXXXXXXXXXXXXXX
XXSSNAM----NKKENELNLIQAHERQIEDMRDSINREWELKAAQMMEEHHSRVSDFN
EEKERMKSQMETLNGQIEDLNKNEFEDKIKEMNVLLSQKDNRKQADLKSKDGOIALLO
IEINTKDNKCNTOIDPEIPLDPEINSLKELVKGFECTV-----NSDTVTFSLIDSNQOE
FIRLSVHGDNGLSISKWRLIVVKPKAGFSFPDGIQPFKGS SVTVWVGTRPRQGTPTENEFEY
WARTELWTSEGTIVKLVSPSEET-TTVTLPA

>dicfir

SPNRAAQRLEKDELSTLIHQRLKSSLKKESTEIELEKKNQELEEIDQKHTLTINKLKQR
SDQFE----RQLLEEQNQNSDLTSNFFNILDN---ELKTKESSWKKEKDEMLSKFQESINK
LNQENSLQSQLKADIVTKEYEIEGLKTEINRLKDDLQYRVREGEDKXXXXXXXXXXXXXXXXXX
XXSSNAM----NKKENELNLIQAHERQIEDIRDSINREWELKAAQMMEEHHSRVSDFN
EEKERIKSQMDTLNGQIEDLNKNEYEDKVKEMNVLLSQKDNRKQADLKSKDGOISLLO
IEINTKDNKCNTOIDPEIPLDPEINSLKELVKGFECTV-----NPGTVSFSLIDSTQOE
FIRLSVHGDEGLSISKWRLIVVKPKAGFSFPDGIQPFKGS SVTVWVGTRQRQGTPTENEFEY
WSRQELWTSEGTIVKLVSPSEET-TTVTLPV

>dicint

SPNRAAQRLEKDELSTLIHQRLKASIKKLESVEIELEKKTQEYEELDQKHTATISLKLQR
SDQVE----KQLIEEQTQNSDLSSNFNILEN---ELKSKESSWKKEKDELLKFQESINK
LNQENSLQSQLKADILNKEYEIESLKTEINRLKDDLQYRIREGEDKXXXXXXXXXXXXXXXXXX
XX
XX
XX
XX
XX
XX
XX
XX
XX

>polys

QSPLSINRMKEKEELTEVTKKLSKILTILEERNNEIVRLKAE LND SNKSSSVEVSALRTR
LDETE----NRLTOEIQRARAAASTLELTLA---DLKNKEDLYMHDRNTFDARLEDTISS
ITKEHSALTIVLRTDLNKRLEHEIEQLKGDIKLLNADLLTKSKEQDEKDDDIKVKLTLKEK
EKIIIAN----QRKEAELNQTVEAFERQVEDIRDSINREWEIKVAEAEIEEQTSAKASFD
FERETFRNQUIAVIQQOTEDLNVKNHELSDRILDLSNELRNKEQRRSAELKSKDGKLSLLO
DEISSKEIKFQQSMDPEIPIGDEIRRIKELVNTLTHQS-----KLNTMCILSIDTTKE
TIRLNASGVNGMTITDWKIVNPKPKYGFQFPANITPLKGNVITLHTGRSRNTTAAENEFEY
WSRTGIWDSEGTIQLVSPKEVLQESVSL-T

>polio

SPSRVMRLKEKDELNAIHKKLQOCITSLKEKDEEIEKKEREIESLQKEKTKSTNAYKTK
LEEAE----RLLSQEIKEKTEFLSRADFLES---ELKTKEATWNKDQDMISKMDEAINK
LTHEHTLLSSVKSDLVKSEYDNETKNEIARLQGELHTRVKEYDEKDEDIKKLMELKEK
EKLSLVA----QRKENELNQTIQSYERQLEDIRDSINREWELKTAAMVEEXXXXXXXXXX
XERERFKTQITVIEQVNELNIKNQEQDLIEQLNNTLMNKENRKLADLKNKDSQISLLO
DEINSKETKLITQNDPEIPLTREIASLKELVNGFEKSVXXXXXXXXXXXXXXXXXXXXXXXXQOE
YIKLSVHGKDGFCISGWKLVKPKCGFSFPENIQPVKGHTITVWVGTRSRHIQTPENEFF
WSREEIWTKEEITLKLVASNETT-AKVSLPE

>trichop

MSPLKIKRVQEKKEELKNLTDRLASYIEKSRFLESRNARLNEEIRSSRRSGDENIKSLKAL
YEQELQEARKTIDEMANEKSKITITVMKYTKQVEELTS DLEKVKASKSSLQVALNNAEK
ATELEGEVVNERKNVAELREVVIHKEKIQRSS EELESANKLAEQHQQEIRSLKKNLKT
VDDQORSTLQDDFQSEYDTKMTEALQQLRRDNEENSKRLKQVEVEELYRSQVKELESQITELK
NECRKFRDKSESISSVHVLAENSSLQTRKRELEESLIEERQKNTNKFEEYEMEMRALO
LKLEEQSTHEYKELLDVKIELDNEIAAFRKLLEGEETRLSSSPSPATGDVQISDVDKHGL
YVRLNLTSKTDCHIGNFVIKHQVNEISYKFNKAILKGNVCVTTVMAADSGQSHKPPANYL
WKQNNFGLEEMVTRLVNAGGEEIAIFNLT

>brachi

LSPTRLTRMQEKQELQWLNDRLAQYIDRVRYLEAENSRLMVQVTSSEEITQREVTNIKSM
FEQELTDARKLLDDTAKEKARVQIEAGKYRAEADELRAKLAKEGALATAEKKRHOAESA
LNEKEGRLQNAIADKQRAEGELAAALRLEFANLEKQLATARKQLEEYEQELTESRTR---K
EVSITEVDAG-EAAFESRLOEALREMREQHELEARSMEELEMTYTQKITSMRTDLSVAR
EEMRESRARIDSLMSQVSGLQGNASLEARCKELEGMARMSEESRSSAEQYEREIRQLR
EEISQMMVDYQELMDIKIALDLEISAYRTLLEGEEQRLKLTPTPSSSGAVAVGGVDLEGG
FVKLQNTSEKDMMSGWLLKRTVGEISYKFPSTRYVLKAGQSVTVWATEGGGTHSPPSDLL
FRGQASWGSDDTKTLLVNDSGEEMATRSVTR

>priapul

LSPARITRLOEKEDMINLNDRLANYIDKVRNLETENSQLSIQVGESEETITREITNVKAL
YERELADARKLLDETSREKAKLLIDCNKYQSENDELRTKLARLTRDLTGAQRSLTTAETO
VPDLTIKYDTANKDRKKADELRLQKKDMADFQTLQSAARSELENYENEMKEVRAK---K
ELAVTEVDSRLQEEYEHKLADALRELRDQYEEQSAIHRDEIEQMYESKLADLKHVARLAR
AQIKEMRSEYDSLKAQAESFESERAALARMGLERALDREREHHRDQMSQRDEIKNLY
AQIAQNLOEYQDLMDIKIALDMEIAAYRKMLEGEEARLNIDTPPSSKGEVVRMECDPDGK
FVKLSNTSDKDIHIGTWQLKRITGEITYKFHRNIIKAGQHINWSSDVKVTHNPPSDLV
MKNQ-KWYSESMKTILIDSEGEVGERENFR

>drosop

LSPTRHSRVAEKVELQNLNDRLATYIDRVRNLETENSRLTIEVQTTTRDVTRETTNIKNI
FEAELLETRRLLDDTARDRARAIEDIKRLWEENEELKNKLDKKTKECTTAEGNVRMYESR
ANELNNKYNQANADRKKLNEDLNEALKELERLRKQFEETRKNLEQHSQEINESRRI---K
QTEYSEIDGRLSSEYDAKQSLQELRAQYEEQMQRINRDEIQSLYEDKIQRLQEATHKSI
EELRSTRVRIDALNANINELEQANADLNARIRDLERQLDNDREHRHGQEIIDLLEKELIRLR
EEMTQQLKEYQDLMDIKVSLDLEIAAYDKLLVGEEARLNITPATNAKGNVEIKEIDPEGK
FVRLFNKGSEEVAIGGWQLQRLINSTTYKFHRSVRIE PNGVITVWSADTKASHEPPSSLV
MKSQ-KWVSDNTRTILLNSEGEAVANLDRIK

>nematos

PGSAKFTRAQEKAELOHLNDRLATYIDRVKNLEQENSKLRSEVTVSRKTVEREVDSMKSL
YETELADARRLLDETAKEKAKQOIESSKNSNDAQEFKNKFDKEAAARKKAKEKELNDVRKL
LHDKENQLTRRNQEALESVLRLEQGECEELKDALKAAKYALEQYDKELSDIRSOLKTV
ETKRUVVETDYKDKYEGMAEKLOELREDYDSDARSFKEETELLYSSKFEELRIQLAKLR
EENRNLSKSVDELSSQVHOLEAKNNALVSRVSDLQGLRAQDKEKHDNEILLRENEIAELR
TSIDDALRDYEDLMGVKVALDMEITAYRKLESEETRLEITPPASAEGAIQFTEADPDGK
YIKIYNSGEKDEALGGWTIQRQVTPSVYKFTPKYVLKSQSHVTVWSAQGGGTHKPPSDLV
FKQLPSWGSNEARTALVNAGGEEMATLLEEK

>amphimed

PTPERRTRMKEKEDLQLLNDRFVTYINRVRRRLRDEKEKLNVTLEHMQTTTQEESGAIKRM
FEKELQDARLLIDETAKEKARYQIQSSKTEDRIKDLETELKEFREDYDRMQQELAEAKKS
AETFESLYKSSLVDNQSAKQKIQDLEEEIDDLRARLREVQDSLEQYDEELLTIKRT--GL
TTVHRDGVVEETDDDFAEERLQAAIDEAREEIEKETEQFKLDLEKSYKTRLESLESQVRVLD
GELKNCHTTLAKYNRDIKLOEKNEQLESSLREKSEDLARAKQLHSAELANLMDEMRALR
VNYDQMKMEYEELFDLRIQLEQEIATLSALLQEEEQRLNLQTTPRAVGFIQIIDVDPDGR
FVQIKNMSDKVTEIPKNKLQLVGLIASKYEEMYAPEVADFVYITDSTYSTEIKAMERNI
LKTL-DYSFNPLCLHFLSRAGDATPQMHTMA

>daphnia

LSPTRLTRLOEKVELQNLNDRLANYIDRVRQLESENRLTLQITTTQDTISREVTSVRGS
YEKELSDTRQLLDETAKEKARLQLEAGRARAELGELTPKYNKAKNDLLQAEKKVQOLETR
NNELNASLHQLKSDMARAI EKKDAVSRDKLAKQVPELQQOLETYEREMTEVRSN---K
QVEISEIDGRLOEQYQARLQSTLQELREQYESQIQQSRDEVEVLYQNKVEDLEDQAGAHY
EELIQARNRLKDANLKLSELELANSHMNDRLAEMERNMETERQTHATAMSMCQGEITRLR
EQMAIQLOEYQDLMDIRTALDMEISAYRKLEGEETRLKLTPTAGAKGEVQVEEVDSEGK
FIRLINKSEKEVSLSGWQLVHKAQETIYKFHRSLKVAPGGTVSVWSAGSGTAHEPPSTLV
MKEQ-RWFVNEMVTQLLDNSGEEMAQRESKR

>hymenol

PSPLNISRVDEKEELAHLNDRLASYIDYVRKLERDKESLTRRISTITEERLAKVDNARST
YENEIQSLRRLVDDLAKKAARDVELKXHMDASDAKSKLNKRDVELRNLQRKYDTLERD

IGTY-----KSDHERYQKLLPEYEHLEKSLEAAKEGLKAFEEERSKLVMR---S
LTVEEEVAETKAKAYESRLADELQAYRESANEELEQYRLQOMESTFQNKLEQLQKAGGRLN
EELIAMRRMNDLDSHELARKSREVELLHSRIADLEKLLNKAHDDYQAQLAAEREEVKRLR
AELEQRF AEFTDLMN TKIVLDQEILMYRKMLEGEE SRLNL TSPKRAVGSIEFSDQDGSN
WIKLSNVGKEDINIGNWVLKHKSDDEVTYKFG RNVILKPGTPCTVWSVDAGATHDPPEDIV
MKNQ-RFHAANTTVTLTNEEEEEQASCLIVR

>aplysia

PSPARLSRIQEQEELQGLNDRLAAYIDKVRSLAENSRLALQVRSTEEVVKREVT SVKTL
YESELAELRRLDDETSKEKARLQIEANKYKTDYEDLLAKYNKRDRDANALERRVQVLERQ
LAEFSAR-----DFEKENQNLKNELNRLEQQLA AAKKQLEEYEQELEESRVR---T
TTQLEEV DGRIEQYEARLQDALREIRAQHDYDLQSVKVELETLYESKVEDLRNQSSAAW
EELMLARRSVDDLSAEIGRLKSENAGLDLRVKDIDKQLERERDEFRLRMKQRDEEISRLR
AQVEELEREYAALLEIKIKLDREIEAYRKLESEETRLNISASMTSSCGVDISDVDTDGK
FVKL TTTDKDVHLFWQIRHVSGETRHKFKQN-VLKAGQTTVWSADSDQTHNPPSDLV
MKGK-RWFVGDMTTLLDQDETEVASCSMSR

>laminB1

LSPTRLSRLQEKEELRELNDR LAVYIDKVRSL ETENSALQLQVTEREEVVRGREL TGLKAL
YETELADARRALDDTARERAKLQIELGKCKAEHDQ LLLNYAKKESDLNGAQIKLREYEA
LNSKDAALATALGDKSLEGLDLEDLKDQIAQLEASLAAAKQLADYEEEEINETRRK---H
ETRLVEVDSGRQIEY EYKLAQALHEMREQHDAQVRLYKEELEQTYHAKLENARLSVNSAR
EELMESRMRIESLSSQLSNLQKESRACLERIQELEDLLAKEKDNSRRMLTDKEREMAEIR
DQMQQQLNDYEQLLDVKLALDMEISAYRKLEGE EERLKLSPSPSATGNVCIEEIDVDGK
FIRLKN TSEQDQPMGGWEMIRKIGSVSYKYTSRYVLKAGQTTVIWAANAGVTASPTDLI
WKNQNSWGTEDEVK VILKNSQGEEVAQRSTVF

>salpin

GSPQFRRRVKEREQLDELNNELAAYLVRVRELQKENGELTSQLEAIRFSRKRFE STATIK
LEGQVNDLLEEIRRKTT EIGTLNHQVQDLRQKLNKTEGDALEATREKDDLAKQLNEALS
VQLEGGVSLGQRDVSQRDTT VTNLQRRVASLEADLGRQVDDTKRYQQQVTALEEQIK--
-----IGGTSGGLSEEDLEAALTEAKEHYKQAVNNFR TQMRSYFAQQ--PIIEPRARLQ
TEVIEWRSKYEDVKAQGATAQTEIEGLKSKLAGLEDVFAQIKKSHTDEIRHKEDFIFKLT
KTMQDKDDLRYRDLDERIALDAEIDRYRRQLESALRMFEARS PGTTQGALSVEDVDREGQ
YILLKNTTASPLDLADWAVTVTNRLGTFEFP-AVTLAPGRSTRVIRGAGDGAEKLEGDVV
WTEFPDLAAPELAVYAFDPDDNESSTWAIMR

>salpin2

AIAAVRNTMKEREEMEKLNDNLEVYIRRTSALYSSHSTLVSLQASRRTSVVASTPTASPS
KSNE-----ERLRKEVKRQQA EIDRLQSLASQLEKQFENAVEVKDTALNQLEECAAT
QQDLRTEL---EKQKEAYDARLASTEATISSMTAELEESRARAHAHETT VTTLEEE---L
QACTKAMA--SKQAPDTTIDA AVAEVKKEFQAAANAYRHKL RQYYATHARGPQAMRMALK
QQVFTLTKEKADLSTTNASLHATVDACKQEVAAIRQSMKDAVHHHTTEIKAKEERIAALT
AHIRANEKKYADLLDERLWLD AEIDEYRRLNLSIQRRLSLPAAPRGPGRLELRD VDTKVE
CLLVANPSTSSVDIDGWALTASADLISTLEFSAHVLP PGGVLRV FQQRNEENRDCLWDEL
TRPR-EWGDQKXXXXXXXXXXXXXXXXXXXXXXXXX

>monos

PSPKRRTRELERQRLEDLNHALSRYISRVRALETQVVTLETHIESFASVKAPKNDEVLET
LRAQLEQAKAELDDASQRYADLQLEHSNLQEDFKRAEGDIQGLTSLKERLQGRIESLQAM
VNRLEGQLASAQSSLQEEKEDGLARQQLKDLKADLVKHRGRADSHQSETDALQQQLQVA
LSQVVVSVKTTGNQYEEEMSRLLRVAKSSYAEAA DAFREKLKYYQMNPTPSAAPLKRLO
AENEALRNVAARDQDVADLKNEARHAEDKIDRLTHDLASTKAENKAHLANKDMEIKNLQ
QRISGLEDQYRVLLDRNLSLDAEIDQYRGLLGEEYKRLNMHDQSLSNQIQIQSVDDRGF
DLHLQNIQEPINFDCVLTVDNGSETLVMPHGHVAVGQVVRVVSGLGEQSHR-PGDIA
WTSFRDFEFNTINVQLE--RGSQVICRAEY T

>capsa

SSPDRAARLEEKEELQFLNSKLELYIHRRQELQTLVKTLESELERERND AQORLEIAHRQ
YEATLTETRHALDKAASEASHASSLVQVNAELAATKERARTESSAADALRSRYEALQLA
KHELDGRLAVSEIAIKELKDNLKHADSDRKHAEKESDALRKSSEKGADLNAYHSEEMRRL
RSDLSSLERKLR AEHDQELAEALSEHRRQTEAESAQYRAEVQALYESRLEAAREQVAGLO

QALSAAQAAQDRARAQVTALSQRSETLSTRVRTLEHELI IERDTVVEVQRLREEATRLMQ
EAIAAKSEVRTLTETNVALDKEIAMYRTLLEVEEKRLGEDFTPTEAGLVVFEFVNLVHN
KLVLRSTNEEEVSLHGWKVRTTSESATYIEIPAGIALSAEPTVTIWTGVSAGKRKRSSELL
WKQDINYASNGACIQLLDNEDNEIARLEVRE

>creol

LSPMVAKRLEDQIRLTQLNSKLKEVIEKMENAQPNTNIEIESTCARLKEDEDNSLNRLEL
THAKMADAQDRKKLLTKEKANIEKKDAEAFKNTKRETEAKALHKSAAEKLRKQTKDLQSK
IDDCTKQVGNIRDELAQANQALEQLEASVAAFVDTKRRTLGEKSAENKEALQHLRAR----
---IAGVAKKERELQTKAAEAADVVGVFNALQKHVNDIKVLHEKKLETAIQTDVEDK
AEVKRLDGKIKDLKLRKELEDSAQKQKAI IAELEKQLENMKKKHAVEM---KAQTQQVM
AEKEIVQLAYLQAEKAQKLLGDQINEYEKLLDFEER-----ITGIVELVAVDPQGL
YVTLVNETEDEIHLDDYHLIGVSHTRRFDFSSDCRLDARGKVTVYTNAEYERSIGNAQL
WETDPVWAAEGPSQMKLYNSGGLVSAATISV

>sphero

SSPLVKQRKTAKRKLEALNAQLANVIHRMEHAEPTEXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XX
XX
XX
XX
XX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXTIEEYARLLDYSDAELDASFQPTITGVVSLKEVDPKGT
KVLLENQTEEVVDLSGWHLVGSHTARYDFDKDATLNARCSVLVFSYG--YGYDIPXXXX
WKTDPVWSAEGPSTMKLYHEGGLVSAATISV

>creol2

MEEVEKLRTDNKVQIDVLHEKLNITVSQCDALLETENAALVAQTEAIEVEKVKLMDDIESL
HEVH-ADLQDLLKRRQECHELESETGKAIKKKRLEETERKTA-----
-----QOCDSLTSNNAEKDKVLADLRNEADVMKREISDHENVITNL
TIELQKTEDAACAQKFKKIASEKRDEMKALEEKKKQRTAELEAELQSLKKS VNKAQ
AEVDELSTIIKEKEASLPDVE---HLQAAEHIKADIKWKERIDTLKKTTEEIEAELT
NEKLGHDEQMTQLDRYGITLEEQLNNYRDLLHVAAGDGVDP--HIPVYSLDSINPNTS
SVGVRNNSITDVLAGWTLTIEYTSIKYTFGDNMTIKPEAIYTVTTDPDEVTHN---AV
WIKVDVWGYNNVKIVLKDTSGELQCSAVVSA

>sphero2

XX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXLEKETGKLFKFKKKKFEEDTERKTA-----
-----QKDDLTVIEIKDLGLLVQRQQEADQMKEAFKAHEATVNAL
KEELXX
XX
XXXXXXXXXXXXXXXXXXXXXXXXVLDLEQLNDYQNLLHVKPSADGLAP---TTPSLYIDHIRPAQS
SLGVRNNTERDVLHSMIRVRYTHITYTFSGPAIVSPDQLYTLTTDEHESHTD---EV
WESLDVWAYRTVYLELLDADKLLVSTAVVXX

>pirum1

LSPASQERKAEPFERLCNLRGLKYYIRMVREKYPEENQHOREMASQKDRSDNEKQEMKEH
FNGKIQNLEE-----EVARHSASNRESLKKIDQLHREKANFEKAIENLQTD
LVKKENHINQLKEEKKTSSVLVQEELEELSELKKELEADAQDEKAGLTEELETTSN----
---LENNEQELNSAYEERLEIALAEAKKVHQONADKMLDDVERSYEKI FAEKEADTQGLQ
KENDELKEQKETILQEGEEIK-----NENNELQQS VVVLK
AKIDELEKQYNNVSKSKANLTREIDSINKLLTEEARLERDIAKESGDV I I KEVDAEGF
FIRLQNVTSAPIQMDQWLLRSYIEAIDFNFSKDLVLEPEQYLTIIITRLDEEEIQPDSEVQ
WDSEPLWNEGGFVELFNQEGASVSAAQIEL

>pirum2

XXELQNGREV
LESKIEDLV-----TLSSKTQKFTEEKQNLQDSDLNAERE
LNEIKTI-----YEEKLEKTRQD----
-----LAEEYNARLEDALLNARDQHERENEEFELRLNNEHQEELSELNVKLKQSI
AKLKQVNDTNDELRRLELSQKEAEKKELENQLVDVKKSFEQLQADYEEQVNENDSHRLLD
-----EVAEENSRLRTEIDYTFKMIEQEARMGMKFDPKHSTFTLDEVDPGN

FFYLKNNSSKPPVDLSTWVLRASSAYVDFHFPDNI TLKGSEKVAVLTSQNTTEESDMKLDLF
WTVTSIWANDKGSVNLFK--GELCSAATITE

>abeofor1

LSP---SLAKEQKRLADLNDRLSDLFSGLRGKR-----DDIISG
YERKIQDLNKQLLEAGEVLHNLEADIAERDAAIQKLEIVKQNLASREDLQKVYESKERQ
CEDLHREIQLOVEKLSEEKAKNKEMTKELKEVKNQNSASQKLEKLNAQIEKLEHD---L
KEKDEENAEAEACDFDVRMQLALQEARQLFETDLESQLEAVESKREAEMKNLKTQDESHA
KEKKLEDQLEQLKANFEKLEDEKEKEMEDQLLQMKKELEKERKQSEKNLKMDEEYQREKE
TQKAANDLAYDELKAEYENLVQTTITDVTNLLLEEEERLGIPITPSYGHQINFDEIKTNGQ
YIKIKNESEENVSLTGWSLKCTFEEITFTF--DDYELASEDSVAIVTSSDYMSNIKGNKIL
FETEPFWAVDNGVVELYDDRGDVICSAVVEE

>abeofor2

ATPVRKKDSGQNARLQNLNRLKYYINLVRQKHPEE-----MDKVKNE
YEVKMKDLQGOINALQDEISELNSLNLVQEKELKEKNASLRQKD----DLMNKQEQFTEG
LEQS---IADLQVANKVLTNKERDLKAKVSALKEDYDELCEANEAHAIEMKLSHDQ----
-----AEKERLASEKHSNEMREMSEEFFEKERNLKNELRQVEKIKIESEKATFEQKE
NELKEVTHKLNKLADSKKTLEDENAKLLQKIKDLELNSEDLYK-----KLTKVE
TKLENKTRKYH---DSSQTILEQLDIYNNLLLEEEKRIEGLPKFSSTGSI VLEEIDATGA
FIRIFNNSKHPVCLSGWNILVKNPIGQFCFPDDFTLAASGKVSVTSPS FNDDDDTEVHLQ
WNEPSIWSSDGCTIQLFDMNGIIASTGKVEE

>hypoch

XXXXXXXXXXXXXXXXLNNRLEAYVLRVKEMDDARSHLDHELKTI STRYQKEMDALRER
LSRDVDTLRSELEYE IDQKTVHQVKAEDLTLQVEKMTAQVRITSKLNLEEE---EERERE
MDREKESCAFKREXX
XX
XXREEAKRYKDEIRRLEKEVSALR
DSQAESEKQFNLDMDVKIGLDNEIRHFKA ILDMEEDRVGVFSPERAVPDVVVQ MIDTENE
CVVIKNNSSSEAKDL SHWTL SNKAGTQKFEFPEGFSIPADSSVRIWNGKKN SKHSPPENIY
WSSKFVWNADGEAALLLSPSGDLAASSAGGE

>phytcap

LSPIKSKRMEEKASLQKLN SRLEMYVLGVNELES AKHAAERELETIKQRMQQDLESVRIR
LTKELEETRKKLDDEVDQNNRLOTLEQEQNRELVKLRAQLKEFGEL----KVLVETLKV
RDREKTNAESAKQALSEQTTQLNSARRSVKELERDLRSHKAALTDATQELQDLRKKCANY
DMAKDSELTKLRREWA AKHLEAQAIWKKDAEERLATMEMEVRSHFNSMSTSLESQ LDDVR
TELESTKKELDR TANDYEDSLKARQSLTEKVAQLERDYREARSKSTKDRKAYEETLERFR
SSKVAKEREFNELMDVKIALDAEIMKYRRILDREESRVAITTPNTPSSAVQIASLDLEKD
RIVIQNTSNEPVSLGGWAVRGQMD-QTFRFPKTYVMRPRSTLTVLSSKR NHESKKGEAAF
LATQFSLNPNGDIVFLVTD DDIIPVSMMS EGL

>phytpar

LSPIKSKRMEEKASLQKLN SRLEMYVLGVNELEGA KHAAERELETIRQRMQLELESVRSR
LTKELEETRKKLNDE IDQNAKLQTLQEQHKELVKLRAQQTELGEA----KVLVETLRVQ
LDREKANAVSAKETLSEQTTQLNLARRHVKELELERLGHSAALGDATKELEELRKKSVDF
DLTRDAELTKLRREWTAKHLEAQAWKKAEDRLQSMEMEVRSHFESVSGSLQTLDDVR
TELESTKTELDR TANDYEE SLKARQSLTEKVAQLEREYREVRSKATKDRKAYEETLERFR
SSKVAKEREFNELMDVKIALDAEIMKYRRILDREESRVAVATPNTSASAVQIASMNLEKD
RIVIKNSSSDPVPLGGWVVRQMD-QIFRFPATYVMRPQSTLTVHSSKRNNERKKGEDSF
LANKFSLNPKGDFVVLVNSDDIPVSMMS EGL

>phytcin

LSPIKSKRMEEKASLQKLN SRLEMYVLGVNELEGA KHAAERELETIKQRMQQDLDSVRAR
LTKELEETRKKLDDEMDQNARLQTLQEQHKELVKLRAQQVELGEV----KVLVETLRVQ
LDREKANADSAKETLSEQTTQLNVARRRVKDLEREVRGHVAALGDATQELQELRKKSAFD
DLTRDAELSKLRREWTAKHLDAQAQWKKAEDRLLSMEMEVRNHFFETVSGSLQGLDDVR
AELDSTKKELDR TANDYEE SLKARQSLTEKVAQLEREYREVRSKSTKDRKAYEETLERFR
SSKVAKEREFNELMDVKIALDAEIMKYRRILDREESRVAVSTPNTSSSAVQIAALDLAKD
RIVLKNSSSDKVPLGGWEVVRGQMD-QTFRFPPTYLMRPQSTLTVHSNKR NHERRKKGEDSF
LANKFSLNPKGDCVILITSDGIPVSMKSEGL

>phytinf

LSPIKSKRMEEKASLQKLNRSRLEMYVLGVNELEGAKHAAERELETIKQRMQQDLESVRSR
LTKELEETRKKLNDEIDQNAKLQTLQEQHQKELVKLRAQQTELGEA----KVLVETLRVQ
LDREKANATSASETLESEQTTQNLARRRVKELERDLRSHAALLSDATKELEELRRKSVDF
DLTRDAEMTKLRREWTAKHLDAQAQWKKDAEDRLQSMEMEVRSHFDSVSGSLQTQLDDVR
AELDSSKMELDRTANDYEESLQARQSLTEKVAQLEREYREVRKSTKDRKAYEETLERFR
SSKVAKEREFNELMDVKIALDAEIMKYRRIILDREESRVAVATPNTSASAVQIAAINLDKD
RLVIKNSSSDPVPLGGWVVRGQMD-QTFRFPATYVMRPQSTLTVHSSKRNNERKKGEDSF
LANKFSLNPKGDFVVLWTADDIPVSMKSQGL

>phytso

LSPIKSKRMEEKASLQKLNRSRLEMYVLGVNELEGAKHAAEKELETIKQRMQQDLDSVRTR
LTKELEETRKKLDDEMDQNARLQTLQEQHQKELVKLRAQQVELGEV----KVLVETLRVQ
LDREKANADSASETLESEQTTQNLARRRVKELEREMRGHVAALSDATQELEELRKKSAF
DLTRDAELSKLRREWTAKHLEAQAQWKKAEADRQRTMEMEIRSHYDTSVSGSLQSLDDVN
AELESTKKELDRTANDYEESLNARQSLTEKVAQLEREYREVRKSTKDRKAYEETLERFR
SSKVAKEREFNELMDVKIALDAEIMKYRRIILDREESRVTVSTPNTSSSAVQIAALDLVKD
RIVLKNSSSDAIPVPLGGWVVRGQME-QTFRFPPTYVMRPQSTLTVHSSKRNNERKKGEDSF
LANKFSLNPKGDFVVLSSDIPVSMKSAGL

>phytra

LSPIKSKRMEEKTSLQKLNRSRLEMYVLGVNELEGAKHAAERELETIKQRMQQDMESVRSR
LTKELEETRKKLDDEMDQNARLQTLQEQHQKELVKLRAQQTELGEA----KVLVETLRVQ
LDREKANAESAKETLESEQTTQNLARRRVKELERDMRGHVAALSDATQELEELRKKSAF
DLTRDAEMTKTRREWATKHLEAQAQWKKAEADRQRTMEMEIRSHYDTSVSGSLQSLDDVN
TELDSTKKELDRTANDYEDSLKARQSLTEKVAQLERDYRVERAKSTKDRKAYEETLERFR
SSKVSKEKREFNELMDVKIALDAEIMKYRRIILDREESRVALATPKTSSSSVQIASLDLEKD
RIVLKNSSSEPVPLGGWVVRGQMD-QTFRFPATYVMRPQSTLTVHSSKRNNERKKGEDSF
LANKFSLNPDGDFVILVTSDDVPVSMSEGL

>phytlat

LSPIKSKRMEEKASLQKLNRSRLEMYVLGVNELEGAKHAAERELETIKQRMQQDVESVRSR
LTKELEETRKKLDDEMDQNARLQTLQEQHQKELVKLRAQQTELGEV----KVLVETLRVQ
LDREKANADSASETLESEQTTQNLARRRVKELERDMRGHVAALSDATQELEELRKKSAF
DLTRDAEMTKTRREWAAKHLEAQAQWKKAEADRQRTMEMEIRSHYDTSVSGSLQSLDDVR
TELDSTKKELDRTANDYEDSLKARQSLTEKVAQLERDYRVERAKSTKDRKAYEETLERFR
SSKVSKEKREFNELMDVKIALDAEIMKYRRIILDREESRVALATPKTSSSSVQIASLDLEKD
RIVLKNSSSEPVPLGGWVVRGQMD-QTFRFPATYVMRPQSTLTVHSSKRNNERKKGEDSF
LANKFSLNPNDFVILVTSDDVPVSMSEGL

>hyalop

LSPIKSKRMEEKASLQKLNRSRLEMYVLGVNELEGAKHAAERELAIKQRMQQDLERVRSR
LTKELEDTRRKLNNEMDQNARLQTLQEQHQKELVKLRAQQTEWSEV----KVMLETLKAQ
LIGEKANAKAAKETLESEQTTLLNAAQRRVKELERETRSQSVALSDATQELEELKKSADF
DLTRDVELTKLRREWTAKQLESQAQMKKDAEDRLTMEVTEVRSHFESVSSLQNLDDVK
SELESTKQELDRTANDYEESLQARQALTEKVAQLERDYREVRKATKDRKAYEETLERFR
SSKVSKEKREFNELMDVKIALDAEIMKYRRIILDREESRVAGATPNTSSASAVEIASLDLEKD
RIVLRNTSSKPVPLGGWVVRGLMD-QTFRFPATYVMRPHSTLTVHSSKRNDERKKGEDSF
LANKFSLNQKDFIVLVTSDDDVPVSMKSEGL

>pseucub

LSPIKSKRMEEKASLQKLNRSRLEMYVLGVNELEGAKHAAERELETIKQRMQQDLENMRSR
LTKELEETRKKLDDEMDQNARLQTLQEQHQKELVKLRAQQVELGEA----KVLIESLRVQ
LDREKANAESAKETLESEQTVQLNSARRRVKVLERDMRGNAAVLNDATQELEELRKKCADF
DLARDAELSKLRREWTAKHLETQAQWKKDAEDRLQTMEMEVRSHFESVSTSLQSQFDDVK
AELESTKKELDRTASDYEESLQARQSLTEKVAQLERDYREVRTKATKDRKAYEETLERFR
SSKVSKEKREFNELMDVKIALDAEIMKYRRIILDREESRVAVATPNTPVSAVRIDSLNLEKD
RIVLKNSSSEPVPLGGWVVRGQMD-QIFRFPATYVMRPQSTLTVHSSKRNNERKKGEDSF
LANKFSLNPKGDFVVLTSDDIPVSMKSEGL

>pythult

LSPIKSKRLDEKSTLQFLNGRLELYVLRVKEMEDAKNVAERELETIRQRMQADIDALRLR

MSKELDETRKKLDVELDQKARLQVLEQEQHVLEVKLRAQVKELSEI----KALAESLQVE
LTKERSNSKASKETLSVQTTQLQSARRRIKDLERENRSYEASLSDTTKELDELRRKSSDF
DLTRDLEITSIRKEMNQKHQEALAQWKKDTEERVLGVEKEVRLYFEGVVQGFKSQVEDLS
VELDSTKKELDRITASDYEESLQVRQQLTEKVAHLEREYREERTKFKEDRKKYESTLEHFR
SSKFAKEQEFNDLMDVKIALDAEITAYRRILDREETRFGLPTPTNAPSSVQIVHLDLEKD
RIVLENSSDKPVSLAGWEVRGRAENQVFRFPATYSMRPKSKMTVYSSKRNDERKPNEDSF
LATKFSLNASGDFAVLLTADGVPVSTRAEGL

>pytharr

LSPIKSKRLDEKTSLQTLNSRLELYVMRVKEMEDAKNNGERELETIRERMQKDIDNVRSR
LSHELEETRKKLDVELDQKARLQVLEQEQHVLEVKLRAQVKELSEI----KALAESLKA
LSKEKANAQLSKEELSTLTTQLQAARRRIKDLERENRAHEASLNDATQELEHLRKKSA
DLTRDAEISNIRKEMSLKHEALAQWKKDTEERLKGVEKEVRNYFESVVSGLKNQVEDLS
SELDSTKKELDRTANDYEEESLKVRQSLTDKVAQLEREYREDRAKFKEDRKYETNLEKLR
SWKVTKEGEFNDLMDVKIALDAEIIETTYRRILDREETRVGLPTPTHTAPTATIAHLDLEKD
RIVVENTSDQAVSLGGWEVRGRAENQVFRFPSTYSMRPKSKMTVYSSKRNDERKPNEDSF
LATKFSLNQSGDFAVLLTNEGVPVSMKGEGL

>pythirr

LSPIKSKRLDEKSTLQLLNGRLELYVMRVKEMEDAKNVAERELETIRQRMQADIDAIRLR
LSRELEETRKKLDIELDQKARLQVLEQEQHVLEVKLRSQVKELSEI----KALAESLQVE
LTKEQANSKSSKETLSVQTTQLQSARRRIKDLERDNRSHEASLNDVTKELEDLRKKSADF
DLTRDSEITTIKREMNOQKHQEALAQWKKDTEERLQAVEKEVRHYFEGVVQGFKNQVEDLS
LELDSTKSELDRTANDYEEESLKVRQSLTEKVAHLEREYREERTKFKEDRKKYESTLEHFR
SSKFAKEQEFNDLMDVKIALDAEITTYRRILDREETRVGLPTPTPSAPSSVQIVHLDLEKD
RIVLENTSDKAVSLAGWEVRGRSENQVFRFPSTYSMRPKSKMTVYSSKRNEERKPNEDSF
LATKFSLNAGGDFAVLLTAEGVPVSTRAEGL

>pythiwa

LSPIKSKRLDEKSTLQLLNGRLELYVMRVKEMEDAKNVAERELETIRQRMQADIDAIRLR
LSKELDETRKKLDIELDQKARLQVLEQEQHVLEVKLRSQVKELSEI----KALAESLQIE
LTKERSNSKSSKETLSVQTTQLQSARRRIKDLERDNRSHEASLNDVTKELEDLRKKSADF
DLTRDSEITTIKREMNOQKHQEALAQWKKDTEERLQAVEKEVRHYFEGVVAGYKNQVEDLG
QELDSTKKELDRTANDYEEESLQVRQKLTEKVALLEVEYREERNKFKEDRKYESTMESFR
SSKFAKEQEFNDLMDVKIALDAEITTYRRILDREEVVGLPTPTPSAPNTVHIVHLDLEKD
RIELENTSDKPVSLAGWEVRGRSENQVFRFPSTYSMGPKSKMTVYSSKRNEERKPNEDSF
LATKFSLNASGDFAVLLTADGVPVSTRAEGL

>pythvex

LSPIKSKRMDEKTSLQTLNSRLELYVLRVKEMEDAKNVAERELETIRQRMQLDIDNIRTR
LTKELEETRKNLSDELEQNARLQVLEQEQHAELVKLRAQMKELSEV----KALAEITLRL
LTREQSNAKASKETLSEQTTQLQLARRRLKELERENRSQSAHLNDTSMELEELRKKSA
DLTRDGEIANIRKEMNAKHQEALAQWKKDTEERLLGVESETRSYFDNVVTGYKSQVEDLS
SELESTKAELDRATDYEEESLKVRQQLTEKVAQLEREYREDRTKFKEDRQFEITIERAR
SSKFAKEQEFNLMVVKIALDAEITAYRRILDREESRVGLPTPNTPSVQIVRLDLEKD
SVYIENTGDKPVSLAGWEVRGQLEGQLFRIPSTYTIKPKTTCCVVISSMRNKERKKGEDSF
LATKFSLNPKGDLAVLTPDGVPVSLCSEGM

>albugo

FSPIKSKRLDEKASLQSLNSRLEIYVMRVKEMEDAKTVAERELETIRDRMQQDMDVVRLR
LTKELDDTRKKLDHELDQKTRLQILEQEQHVLEVKLRAQVKELGDL----KTLTEQLKSE
LSKERSNATAAKEELSRQTTTELQAVRRRIKDLERERRGMDAALSDATTELEHLRKKSATF
DLVRDELSNIRKEMNAKHQEALANWKKDTEEDRIHAVEKEVRNYFEGVVTTSSKSQAEELS
LELDSTKKELDRTANDYEEESLQVRQSLADKVAQLERDYREERKKFKDDRKYETMVEKLR
YAKLSKEEEFNDLMDVKIALDAEITAYRRILDREETRVGLPTPKTNDPPIRIAQLNLEKD
FVVFNQGDGKPVSLGNWEVRGKLETQVFRFPSTYVVKPHARVTVFSAKRNADVCPDEDAF
MTKKFSWNHSGDWAVLYDDEGVPVSSLAQGL

>albcan

FSPIKSKRLDEKASLQSLNSRLELYVLRVKEMEDAKTVAERELETIRDRMQQDMDVVRMR
LTKEVEDTRKKLDHELDQKTRLQILEQEQHVLEVKLRAQVKELGDL----KTLTEQLKSE
LSKERSNATAAKEELSRQTTDLQAVRRRIKDLERERRGLDSALSDATTELEHLRKKSATF

DLVRDTEISSIRKEMNAKHQESLANWKKDTEERIHAVEKEVRNYFEGVVQSFKSQVEELS
LELDSTKKELDRTANDYEESLQVRQSLADKVAQLEREYREERKKYKDDRKYETTVEKLR
YAKQSKEEENFDLMDVKIALDAEITAYRRILDREETRVGLPTPKTNDGP IRIAQLNLERD
FVVFENQSEKPVALGNWEVRGKLETQVFRFPSTYVIKPHARVTVFSAKRNADV KPEDAF
MTKKFSWNHSGDWAFLYSDSGSPVSSLAQGL

>sapdic

FSPMKSRLDEKTTLQALNNRLEMYVLRVKEVEDSKEIAEKELESIRDRMQDLTMTKTR
LSKELEEVRRKKYEEEREAKIRVQNLEKEQYNELVQLREQVKELNNT----KTYLETVQSD
LAKEKESCKAAKEALASNATALQSERRKLS DMEKEYRKT LASLHDATSELEQLKQKTSEF
SMTRDSEMTALRKEMNAKHSEALAAWRRESEERQQVVENEVRSHFEGVVEGLRSQLEEAT
KETTALRTDYERTANDYDESLKLRQSLTEKLAVVESQYRADRKKKFQEDRRVYESNIEAAR
KARKEKDLEFNDLMDIKIALDAEISAYRSILDREETRFQVVEPKASASGDFRIRQLNLDEG
RI ILENTGATPLPLNGWQITSKSSNHVFSFPDDYVIQPGGKVS VLSGK-NEGEQDLMDFY
AVKKAIWNPKG DVALVKNAEGEVVCSHAEGL

>sappar

FSPMKSRLDEKTTLQALNNRLEMYVLRVKEVEDSKEIAEKELESIRDRMQDLTMTKTR
LSKELEEVRRKKYEEEREAKIRVQNLEKEQYNELVQLREQVKELNNT----KTYLETVQNE
LTKEKESCKAAKEALASNATALQSERRKLN DMEKEYRKT LASLHDATSELEQLKQKTSEF
SMTRDSEMTALRKEMNAKHSEALASWRRESEERQQVVENEVRSHFEGVVEGLRSQLEDEAT
KEITDLRTDNERTANDYDESLKLRQSLTEKLAVVESQYRADRKKKFQEDRRAYEANI EAAR
KARKEKDLEFNDLMDIKIALDAEISAYRSILDREETRFQVVEPKATASGDFRIRQLNLDEG
RI ILENTGATPLPLNGWQITSKSSNHVFSFPDDYVIQPGGKVS VLSGK-NEGEQDQMDFY
AVKKAIWNPKG DVALVKNAEGEVVCSHAEGI

>eurch

XX
XX
XX
XX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXEKYTALEKNFADHKRWEKKALEHTLDVNR
NARVEKEKQYNTLLDQKLTLDAEINXXS ILEREESRVGLPTPKASNEPVI IHDLDFEKD
FIQLRNTSDSAISFSGWTVLSDGPTHKFSFPENYS LKANTVVTVLGGKKND---PDKDF
WTNKYMWK-DSNIAEVRDAK GELVVYKVYGD

>bremia

LSPIKSKRMEEKASLQQLN SRLEMYVLGVNELEGAKNAERELETIRQRMQQDLDSVRTR
LTKELEETRXX
XX
XX
XXXXXXXXXXXXXXXXXXXXYEEESLKARQALTEKVAQLEREYREIRSKATRDRKAYEETLERFR
SSKVAKEHEFNELMDVKIALDAEIMKYRRILDREHRVAVATPNTLVSPVQIAAIDLK KD
RITLKNTSNDAIPLGNWVVRGQMD-QTFRFPETYVMRPHSTLT VHSKRNNERQKGEDSF
LANKFSLNVKGFVVLMTSDDVPVSIKSEGL

>aphinv

FSPMKSRLDEKTTLQELNNRLEMYVLRVKEVQDSRDVAEKELETIRERMQMDLSMTKTR
LSKELEDTRKLLFEIDQKTRLQVLEQE QHTELVKLR TQVKEFGDI----RVELEQVQAE
LAKEKESCKAAKEALALQTTLSQARRKLDLDKENRKLTSSLSDTTNELDQLKQKTSEF
SLTRDTEITLVRKEMNAKHLEALAAWRRESEERLHSVEAEVRAHFEGQIEQLRSQVDEVN
LELDLKVYERTANDYDESLKIRQSLTDKLSTIETQYRNERKKFQEDRKM YEANIDNAR
QARLAKETEFNDLMDIKIALDAEISAYRSILDREESRVGIDQANHSTGVLRTITLNL EQG
RITLENTGSDALSLSGWQVSSKATNVVFAFPEDYVIQPNGRVSVISGR-NEEEKESMDFY
VIKKAMWNTQADVAQLTNP SGDVVSSY AEGM

>aphast

FSPMKSRLDEKTTLQELNNRLEMYVLRVKEVQDSRDVAEKELDTIRDRMQMDLSMTKTR
LSKELEDTRKLLFEIDQKTRLQVLEQE QHTELVKLR TQVKEFGDI----RVELDAVQGE
LAKEKESCKAAKEALALQTTLSQARRKLDLDKENRK LASSLSDTTNELEQLKQKTSEF
SLTRDTEITLVRKEMNAKHLEALAAWRRESEDR LHNIEAEVRS HFEGQIEGLRSQVEEAN
LELDLKVIEYERTANDYDESLKIRQSLTDKLSTIETQYRNERKKFQEDRKTYEINIDNAR

XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

>retic2

XXXXXQQHVTEKKELSNLNRHLEMFVRHQDRSNEI SRLRKALVDTEVDLRQRLKDAEKR
HQI IQDKIMAENEQLS IQNKTLDQDLETTEKAYQSSESRRASLDEKVTCKMTEIELLQNO
LEA---LTNELITVKSSVKLMVKNDEHTAEHLQAQCQDQWKNKFDVTCEELAKFSFLLNES
TTKFQDMEDQLRKEFDVKMLEFTQKRESQYQNEKEEWIKMFKEEFQKKLAHFKESNKKLE
CEIVDTKARLARVRREKLELEAEKRSLEEEVEKSHSDVNELRQKKDEE IHEKTESLRTLN
DTIRQKDIQYQQLSSAKVQLDNEIAXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XX
XX

>retic3

XX
XX
XX
XX
XX
XX
XSIKQKDCYQKLSAAKVQLDNEIALYRNILSNAEKETDHFNPLGQLCPLQFSAMDLSSS
TLEVQNTSEEPNLNGFYLTNTDTSKQFHLPSNRIAHIGFVFAIIVGDTPLGKVKTGDLR
WHDNVWSGLKDELVRLYNPQHEE IARIEIAP

>retic4

XX
XX
XX
XX
XX
XX
EQIRQKDIQFDALS GAKIQLDNEIALYRSIWSEAEKEAGYSCPLGXXXXXXXXXXXXXXXXXX
XX
XX

>retic5

XX
XX
XX
XX
XX
XX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXLYRDILNEAEKDAECFFQTNDPFPPLQFSGMDLCS
TLELQNVSEEPVDLKG FYLTNSNKS KIFILPHHHVLMPPKPFNIIVGNDGD LNVGKDDLW
WQENVWNGLEDEIARLFNSKNEEVATVNITT

>retic6

PFNSIQQMDDKHKIRDLNTRLKMFVRHQQEGRSEVARLKQLLEETEFNFRQRIK DSEKR
FQLLDQKMQAENEQLTVQNNLSREQFENA EKARQNAESRCANLDQTVARKSTELEVLQKH
LDS---LTNDLIAVKASAKASLNDEQLTGQLQEQCQDQWKAKFESNCEDLAKIKARLKEC
WKKSQDMENQMRKEVDIKMSEXX
XX
XX
XX
XX

>retic7

XX
XX
XX
XX
XX
XX
XX
XX
XX
WNESVWSGKDEIARLYDSSNQEISRVEISP

>ammonial

PGSIKQALAQKRNVRDLNTRLEMFVRAQAEKTRAINQLKESLARQEMDYKSKLTQQRVQ
FQDNVEKLRKENQNLAYENKCTQQEFTNALTSKAEFEERLLKSEEKNAALSSSENSNTQQE
FDD-----MRKAFATMKYKYDSYEAERCSFESRMAKQKSKTEQLLKELTQKEATVKKL
ELENSTTQSRRLRQEFDNKLAEFVQKREEQYKQEKDEWMRIFKEEFNRKLRSEFKHSNIKQS
EEIQDLRARIKQKTELEVTRNRNNEEEIEKLRNDLDDLRRTKDAEIKLKNSLIIQER
DRFKAKELQFDELAGMKLQLDAEIELYRNILNEAERDCGYNSPMVEGGALQFSGLDLNGK
MIEIQNMGEPEVSLDGFALTNQSGSMIYDLPDMELQAKNTLRIYVGEKDESEKDDQVVF
WGNVWTGTDTDCARLYNPSQEEIARIEISP

>ammonia2

PGSPAKQALTOQRTRVLDLNRLEMYVRAQAEKTRNINQLKEALARNEMDYRNKQKQKI
FQEKVEKQRKQIENLEYDKKCTQQELENAMSAKTEYEERLISSEGRNAALASECEQLSKE
LEG-----LRKTYASLKYKYDSYIEIERNSEFGEKMQKLSKNEQLLRELTANEQHLKEL
ELTHNSTQTRLRQEFQKLAEFVTNREQQYETEKAEWMRIFKEEFNRKLRSEFKHANVKQS
EETDLRTRMSKQKTELEVLRNRIEEDAELRDLNLRQKDEEIKQKNAVILQOR
DAYKAKELQFDELASIKMQLDSEIELYRNILNDAEQECGYTSPLAEGARLQFSGLDLLKG
MIEIQNMCEQPISLRGYTLSDTGTSGFPLPNDVLDHNQKIRIYCGKQEDDEDAAPQVS
WGQDVWTGQSNDCARLYNPNHHEIARIEISP

>elphi

XXXXXXXXXXXXXXXXXVRDLNTRLEMFVRAQCEKSREVNTLKEALARSEMEYRNKMKQKQLO
YQDQIEKSRKAIENLEYDNKCTQQELDNGIASKKELEDRLMSSESRNASMSSENDTLKTE
LDT-----VRTAYAALKYKYESYELERNSEFDDKVNALKSKNDKLLRELTQTTSTNKQL
KLANTTTQTRLRKEFDKLAEFVQKREEQYKQEKDEWMRIFKEEFNRKLRSEFKLANCKQL
EENTELRERISKLKQKMELEVQNRNNEEETEKLRTDLNLRQKDEELKQKNAIIIQOR
DAYKAKELQFDELAGIKMQLDSEIELYRSILNEAEQACGYVSPFNEGSALQFSGLDLVKG
MIEIQNVGEHAMPLKAHALSNKEGTQGFPLPSHITLDSHQKLRIYVGEYKAEDDAQQEVF
WTKDVWTGDSKDCARLYNAQQQIQHIEISP

>D1_289

XX
XXQ
IIR-----ARADAGSANLKEIANKNIARLQDQLVEYQAADTDTKALLATERQKVLSP
STQKSDAAGPLQVFDKRLQSVIADQCSAAERNRQIAMTEQKLYFQEKVKGYQAQLQVAG
CEIEQQKNRLEAAAEDYNEAMAFANALEVRIFGLEEAIKKEKSKPSKIIAEKNEIIRKVQ
EVLKEREYDELMETKINLQVELQNYQSLQEEESRCGWENATNADNKICVLSAGSDHR
GLFLVLSSEFLDISEWSFQVSGQRLVYTPRDFSIEGAGSFRVVISSEPVSVVEEKTEAW
WXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

>D1_625

PKPTRKARLEEKSELQDLNKRLEFYILKQREK DANAGSIQRTISENKEYYENEIKKLKSL
HETQVNNFRKQRDEL DSEIQSLTEENKRQASMLKSYTSQLSADKTRGTDLED RVSQ LTEE
LLEAQAATAVAEADAKKSAHGLKKAESI KHLQKELKDSQARDKDHAVELAAAQSLESAL
SSTRANAEDKLRALFGAQLKEILAERQRFDEKKAEGLOELRDSYEGKLEANEKLEEAS
KEVAALKEQANSASSTNSSEESERTKLNKINTQKKELSKAEGQARKQAAFLEEQCESHK
QENQTLNEEYISLMDVKVGLDMEISEFRRLSDEESRLQPQEEPPAKPTIISGID-SKK
WIEITNSVKEAVCLDGWSLKVVSKECFFFPDVA PLKPGKTLKINLGS--KAKKGKNAV
WAKDDIFKEE-DKVYLMSP EGVGHSSVAWND

>spongos

ESPIRIHPEVERKNLEELNGRLQOYIMKQRAK DASREAF EKDVVAIQQARVAIHANTKK
YEEQLRQMRQORDEHASAREELQVRVARQ EATVTRLKDQIAEEKKFQSDLNGRLSSLSQ
LEANAQVQRLQEQLRKTEHNFKSAESAGRAAEEQLIEARDAADQYSHEASSLRAKIEQL
SAQVTITEDSCRQEFASQLSIVEDYRQKCNDDKTRITRELKSHFVPKLAELRAAYEESC
QHESELREQVNQLKKNQTLAEMKSLHEKRIAE LSEELDTERNVTHDAIAAKDKEIEKLR
HTCQRLETDFDDLMDVKIQALTI EK YRELLSEEHRLGLET PGRQTPALKLWVND EEN
CLFVKNTSDKIVNLNGMILQS--GSFSFDFPDDAYVGGQCEVC IWIGETSNSLEASSEIS
WEGVNPEDLTAAPMQLIDQNGTVIDQFEVET

>plasmod

ESPIRMQRDVERKNLEELNGRLEQYIMTQRRK DASREAWAE LKAVQTSAAQALADTTAK
YETWLRQMRQORDEHASAREELKVRVSRMEETISRLKEQIANEKKYENELSTRVSNMSAQ

RQNERLRKENERLANELDNNAGGQODLLNKLAEQERQHKHVVTQLQRDLMKREEELTTAA
SSLVNKNNEFNDLMDVKISLQTEIEKYRSILDVEEARINTPTQDRKGTPLHISDVDLISD
CVSVANETSAPVPMKGWTLISETGNQVFNFPDALVLKPGDRVTVWSGAHASHHAPPDNLF
WTRRYVWNNHGDTAILVNPAGEHVSIVSGVP

>pavlov

ESATRLTRVEERETLQNLNSRLQYFLLRMKELEARNGELQTKVDSGSAGAKAHNDSMRAV
FEKQIAELRSAQETDVELIAELEAKASENGKVASAYKKAERCDEAEVKLRAVIEERDKA
LVL----FKNTEERA EKAEQAAATFEAKASLLEKKL----APLREQ LSTAEELRRKIDEL
SGANRGE LSRMQAAYDAQMQDALDVAKLRAEEGSR AELEQLAAHFTSVESEL TREVAEEK
AAKESALKRANDLAAN-KESI QARV SIAERVTE LERTLADNSRES DGKCAKLEQALEVAR
SERKTKEGEFNE LMDVKISLDEELKTYNMLL GREEARLGITPTQDTQGPMYVSQNLNDD
CVAVKNAGDEPALLGGYTIKSHSEDLVVFVFPSEFELLPGASVTIWSGKKNRRRDPQTELW
WSARYMWSDDGDMATLFDTTGAEVSKLEAAA

>pavlosp

DSATRLSRVEERDTLHNLNSRLQLFLSRMKELEARNTELOSKLDMGSSASQKVTDSMHAL
MEKQIQDLRAASEKDTELIAELQASSGENIRSNQAMKAKAERAAELEGKLRVITDERDRL
AAA----LKDAERRADVAEQAAVTSESKATVLEKKL----APLRDQLQTSEELRRKISEL
STANKTEVARIQADYDKQLQAALAAK KRADEERAAEVAQLAEHFNVVEAQLTREAKDER
ASKEAALARAAELGKT-NDQLQSRADASERAAAAEKNLTERARQHEATVAELDSLHQAR
AARKEKEGEFNDLMDVKISLDEELKTYRLLLEQEEQRLGYTPEPVKSGPVYLAQVDLEND
CFTVQONAGDEACVLGGYSVKAADSEGAYEFPADFELLPGSSVTVWSGKKNKRRDPQAE LW
WTARYMWNNEGDTGILLPGGEEVNRVTAPA

>pavl

DSATRLSRVEERDTLHNLNARLEYFLSRMKELEVANDDLQVRVDTAASVHKTQSDQLRDM
YEKQLTDMRKTSEKDV ELITELQKEANEATKDKVALTRKAQLADELDVKLRAVTDERDRL
ASR----AKEAEARAAELEKSAAVFEAKANMLEKKL----APLRDQLATAEELRRKIAEL
SGGSRSEIDK LKASYDKQ LLEAMEKQKKRADEERLTEL AQVREHFTAVEAELMRAVDEEK
AARASADERMGDLNAQ-VRQLRARASQAEQAAALEKT LAERIAHFDAAIARLDQQLQEAR
AARQQKEVEFNE LMDVKISLAEELKTYTKLLEGEERLGITPTKEASGPLYLATLDLIND
SFALKNAGDAPAPLGGFKVVS HATDVSFEFPDDFELQPGATVTVYSGKKNKRRDPRT ELW
WSPRYMWRNEGDEAIMYDSSGSEVGR LVAEA

NMCP:

>daucus1

EDMGLNAKLMKLETELFDYQYNMGLLLIEKKEWTSKFEELQQVYTETKDALKQEQEAHLI
AISDAEKREENLTKALGVEKQCVLDLEKALRDMRSDYAEIKFTSDSKLAEASALITKVEE
KSLEVESKLSHADAKLAE LSRKGS DIERKSHELEARESALRRERLALNAEREALTDNISR
QREDLREWERK LQEDEERLAEVRRLLNQREERANENDRLYQKQSELDGEQKKIEIIMVS
LKNKEDDISSRIAKLNIKEKEALEVKEKDLTEFEQKLNAREQSEIQKLLDEHKAILLEVKK
QSFEMEMDKRKNDFENDLQNR AVEVEKKEVEVKHLEAKLAKREHALD

>daucus2

DRDAVVEKVAKLERELFDYQYNMGLLLMEKTEWTLKYEEMRRAQVELKEVLEQEQTTHLI
LLESEKREENLRKALDMEKCCITDLEKALRDSGADNAQTKQSSEAKMVKANALLSGFKE
KSMDVETKLHVADAKLEEVYKTSLELERK LQEVETRDSLLQRRMSFIAEREAHEATFSI
QK KDLQEW EKKLQEAERLCEIRRTTSGREVKNEMEMALNLK KQELNKAQKENDLSTSV
LKKEADDINHRLANLTAQEHKALEMRDKELLALAEKLTARE SVEIQ TLLDEQQAVLDAKM
QEFEVMDGKRKSLDEEMRSKLD AVQYK KDEITHIEEKLNRLELSLE

>daucus3

DKASLIAYITKLEAEIYDHQYQMG LLI MERKEWGSKFERVEAALNSAELMRKHDKNLYLK
DLAEAKKREENL KKAIEIERECLANIEKTLHELRAEYAETKVMADSKLVEARSMIEDALK
KLSEADAKKHAASLEAEASRYHSAERKLHEVEAREDDLRRRATSFKTECDTKEEEILH
ERRLLNERQKALQOSQQR LVDGQDLLNKRESHIFERTQELNRKEKELEASKLKQEEELQA
LVEQQANLETKASSLSREEVIVKKREEELCVLQEKLEKKE SERIQQLLANYEASLSMKK
SEFEAELEVKRKSVHDDIENKRRDWELREVDLHHREELILEKEHELE

>citrus1

DVESLAEKVSLENELFEYQYNMGLLLIEKKEWSSKYEELKQTFGEAKDALKREQAHLI
AITDVEKREENLRKALGVEKQCVLDLEKALREMSENAEIKFTADSKLAEANALVTSVEE
KSLEVEAKLRSVDAKVAEINRKSSEIERKSHELESRESALRMERASFIAEREAHEGTFSQ
QREDLREWERKLDQGEERLAKGQRIVNQREEKANEKEKIFKQKEKDLEEAQEKIDATNLS
LMRKEDDINKRLANLITKEKEYLEMKEEELRQLEEKLNAREKVEVEKLLDEHKASLDAKQ
REFDLEIEQKRKAFDDDLKSKVVEVEKKEAEINHKEEKIAKREMALE

>citrus2

DREALMEKVSLEKELYDYQYNMGLLLIEKKEWTSKIEELRQSFEETQEIILKREQSAHLI
AFSEAEKREDNLRRLSMEKQCVADLEKALRDMGEEHAQTKLFSEKTLTDANTLLGGIEG
KSLEVEEKFHAAEAKLAEVNRKSSELEMKLQFELESRESVIKRERLSLVTEREAHEAFYK
QREDLREWEKKLQIGDERLSELRRTLNRQEVKANENERILKQKERDLEEELEKKIDLSSSK
LKEREDEINSRLAELVVKEREAVEMKEKRLTLTIEEKLNRARVEIQKLLDDQRAILDAKQ
QEFELELEEKRKSIIEEMRSKISALDQQEFEISHREEKLERREQALD

>citrus3

DKAALIAYIAKLETEIFEHQHMGLLLIEKKEKELASKYEQIKASAEAAELLQKHDQASHLS
AIAEARKREESLKKTLGVEKECIASLEKAVHEIRAESAETKVAADSKFAEARCMVENAQK
KFAEAEAKLHASESLQAEANRYHRSAERKLDQVVAREDDLSRRIASFADCEEKEREIIR
ERQSLSDRKKILQGEHERLLDAQTLNREDHILSKLQELSRKEKELEASRANVEEKFKA
LNEEKSNLDTLVSLKREEAVLQKKEQKLLVSQETLASKESEIQKI IANHESALRVKQ
SEFEAELAIKYKLAEDEIEKKRRAWELRDLDLGQREESLLEREHDLE

>prunus1

DREELAQRVSELENELFEYQYNMGLLLIEKKEWTSRHEELRQSLTEAKDAVRREQAHLI
AISEIEKREENLRKALGVEKQCVHDLEKALHEIRSENAEIKFTADSKLAEANALVASIEE
KSLELEAKSRAADAKLAEVSRKSSEFERKSKDLEDRESALRRDRLSFNSEQEAHENSLSK
RREDLLEWERKLDQGEERLAKGQRILNQREERANENDRIFKQKEKDLEDAQKKIDATNET
LKRKEDDISRLANLTLKEKEYLEMKEKELLALIEEKLNRARVELQKIIDEHNAILDACK
CEFELEIDQKRKSLDDELNRNLVDVEKKESEINHMEEKVAKREQALE

>prunus2

DRQALADKVS KLQKELYDYQYNMGLLLIEKKEWALKHEELGEALAE TQEILKREQSAHLI
SISEVEKREENLRKVLVAEKQCVAELEKALREMHEEHAQIKLKSEAKLADANSLVVGIEE
KSLETDAKFLAAEANIAEVNRKSTEMRLQVEVEARES VLRREHLSLSAEREAHKKTFYK
QREDLQEWERKLDQGEERLCKLRRILNEKEEKANENDLIMKQKEKELDEVQKKIELSNTI
LKEKKADVNRKRLADLVSKEKEAWELKEKELHELEEKLS SRENAEIEQVLDKQRALCNTKM
QEFELEMEERRKSLDKELSGKVEVVEQKELKINHREEKLLKQEQALH

>prunus3

DKAALIAYIAKLEAEIFDHQHMGLLLIMERKELASKYEEVKASNETTELLHKRDQAAYVS
ALAEARKREECLKKVVGVKEECISSIEKSMHEMRAESAETKVAAESKLAEARNMVEGAQK
KFTEAEAKLHVAESLQAEASRFHRVAERKMQEVEAREDALRRNILSFKTDCDTKEKEISL
ERQSLCERQKTLQEQDRLLDAQALLNQREDFIFGRSQELNRLEKELEDVKANIEKERRA
LDDGKLNLELTEASLVNREEALLNKKEQEIILVLQEKLVSKESDEIRKALASHEVELRKKK
FEFDSELVDVKKLFEDEIEAKRRAWELREVDLNQRDDLLQEREHDLE

>vitis1

DLEALVAKVSKLESEIFEYQYNMGLLLIEKKEWTSKYDEL RQALVDVKDALKREQDAHLV
AMSEVEKREENLRKALGIEKQCVLDLEKALHEMRSEYAEIKFTSDSKLAEANALVTSIEE
RSFEVEAKLHAADAKLAEVSRKSSEIERKSQEV DARENALRRERLSFNAEREAHETTLSK
QREDLREWEKKLQEEEEERLGEGRILNQREERANENDKIFTQKEKDLEEAQKKNEMTHLT
LKKKEDDISGRLSNLTLKEKETLEIKEKELLELEEKLCARERVEIQKLVDEHNIILDACK
REFELEIEQKRKSL EELKSKVVEVEKKE TEFNHMEAKVAKREQALE

>vitis2

DREALVEKVS KLQNELFDYQYSMGLLLIEKKEWTSKYEELS QALAEAEQEILKREKSAHFI
AISEVEKREENLRKALGVERQCVAELEKALGEIHAHSQIKLSSETKLS DANALVAKIEK
RSLEVEEKLLAADAKLAEASRKSSELERKLDQVEVEARES VLRREHLSLSAEREAHEATFHK
QKEDLREWERKLDQGEERLCEGRRIINQREEKANEIDRTLK LKERNLEEAQKKIDLDSL N
VKVKEDDINNRLAELTVKEKQALEVKEKELIVLQEKLSARERVEIQKLLDEHRAILD TTK
QEFELEMEQKRNSVDEELRSKVHEVEQKEVEVLHREEKLGKREQALE

>vitis3

DKAALIAYIAKLEAEIFDHQHHMGLLILERKEWATKYEQIKTEAESAEIVYKRDQSAHSS
ALAEARKREDSLKKALEIEKECIANLEKALHEMRQECAETKVA AEIKLAEAHSMVEDAQK
RFVEAEAKLHAAEAFQAE AICFRRTAERKLQVEAREDDLRRRLISFKSDCDEKEKEIIL
ERQSLSERQKNVQQQERLIDGQALLNQREEYIFSRSQELNRLEKELEASKSNIKELRA

LNEEKS NLELKLASLT TREEDVLNKKEHEILILQE KIASKESDEVQKLMALHEIALKTRK
AEFEAELETKRKLVEDEIEAKRRASELREVDLSNREDFALEREHELE

>phase01

DRGVLVERVSNLEKELYEQFNMGLLLIEKKEWTSKYTEQSQDLVEVKDALEREKAAHLI
ALSEAEKREENLRKALGVEKECVLDLEKALREIRSENAKIKFTAESKLAEANALVASVEE
KSLEVEAKLR SADAKFAEISRSKSSEFDRKSQDLESQESSLRDRLSFIAEQEAHESTLSK
QREDLWEWEKKLQEGEERLAKGQRIINEREQRANENDKLCRQKEKDLEEAQKKIDATNIT
LRSKEDDVNNRLADIALKEKEYLDLKEKELSAWEEKLNAKEKVMQKLLDEHNAVLDVKK
QEFLEVELNEKRKSFEDGLKDKLVELEKKEAEINHMEEKVKGREQALE

>phase02

DHEALMEKVLRLERELFDYQYNMGLLLIEKKEWNSKFDQLRQELAEETEEILKREQSAHLI
ALFEVEKREENLRKALSTERQCGADLERALRAMQEEHAQIQSKSHTKLAEASALVDGIEE
KSSVVDKLLDAAEAKLAEVNRKNAELGMKLQVEVEARESLLQKERLSLVTDRFLDATFYK
EREDLKEWERKLLQORENMLCNGRQNI GEKEENIVKTEKNLQKERDLEVLEKKINSSNSI
LKEKEAEIIRRTADLNMEKKVLEKKEKELFALESKLSREREGIQKLLGEQKATLDLQQL
QVEFEMEHRKSLVEEFSSKEEALQREVEVNHREKKVEKEEQALS

>phase03

DKAALIAYIAKLEAEIYDHQHHMGLLIMEKKDLASKYEQLEALAESELMHKHDSAMNKS
ALAESRKREESLKKTVSVKDACIASLEKALHELRTESAETKVA AESKFAEAHQLI DEAQK
KITEAEAKVRAAESLQTEANRYHNAERKLRDVEAREDNLRRKIMSFKADCDEKDKEMIF
ERQSLSERQKGLQEEQERLLQSQSLLNQREEHFLSRSQELNRLQKELEDTKAKVEKEHET
LHDEKTTLMKEATLMQREEELLSKKEQELLE FQAKLSIRESDETKKVIAGQEAALKTKK
YNLEVELQMQRKQWVENDIETKRRAWELKEVDLKHCKDEILEKQHELE

>arab1

DPRI LPEKISELEKELFEYQHS MGLLLIEKKEWSSQYEALQOAFEEVNECLKQERNAHLI
AIADVEKREEGLRKALGIEKQCALDLEKALKELRAENAEIKFTADSKLTEANALVRSVEE
KSLEVEAKLRAVDAKLAEVSRKSSDVERKAKEVEARESSLQRRERFSYIAEREADEATLSK
QREDLREWERKLOEGEERVAKSQMIVKQREDRANESDKIIKQKGKELEEAQKKIDAANLA
VKLEDDVSSRIKDLALREQETIETKARELQALQEKLEAREKMAVQQLVDEHQAKLDSTQ
REFELEMEOQRKSIDDSLKSKVAEVEKREA EWKHMEEKVAKREQALD

>arab2

DQEALLEKISTLEKELYGYQHNMGLLLMENKELVSKHEQLNQAFQEAQEILKREQSSHLI
ALTTVEQREENLRKALGLEKQCQVELEKALREIQEENSKIRLSSEAKLVEANALVASVNG
RSSDVENKIYSAESKLAEATRKSSELKLRLEKEVETRESVLQOERLSFTKERESYEGTFQK
QREYLNWEWEKKLQKKEESITEQKRNLNQREEKVNEIEKKLKLKEKELEEWNRKVDLSMSK
SKETEEDITKRLEELTTKEKEALLAKENELRAFEEKLIAREGTEIQKLIDDQKEVLGSKM
LEFELECEEIRKSLDKELQRKIEELERQKVEIDHSEEKLEKRNQAMN

>arab3

DRDALIEKILKLEKELFDYQHN MGLLLIEKKQWTSTNNELQOAYDEAMEMLKREKTSNAI
TLNEADKREENLRKALIDEKQFVAELENLKYWQREHSVVKSTSEAKLEEANALVIGMKE
KALEVDRERAI AEEKFSVMNRKSSSELERKLKEVETREKVHQREHLSLVT ERAEAHEAVFYK
QREDLQEWKKTLEEDRLSEVKRSINHREERV MENERTIEKKEKILENLQOKISVAKSE
LTEKEESIKIKLNDISLKEKDFVDIKEKELHEFEENLIEREQMEIGKLLDDQKAVLDSRR
REFEMELEQMRRSLDEELEGKAEIEQLQVEISHKEEKLAKREAALE

>arab4

DKAALIAYIAKLESEVYDYQHNMGLLLLEKNELSSQYEEIKASVDES DLTHMREKSAYVS
ALAEAKKREESLKKDVGIAKECISSEKTLHEMRAECAETKVSAGSTMSEAHVMIEDALK
KLADAEAKMRAAEALQAEANRYHRIAERKLKEVESREDDLTRRLASFKSECTKENEMVI
ERQTLNERRKSLQOEHERLLDAQVSLNQREDHIFARSQELAELEKGLDTAKTTFEERKA
FEDKKS NLEIALALCAKREEAVLLKKEQELLVAEEKIASKESELIQNVLANQEVILRKRK

SDVEAELECKSKSVEVEIESKRRRAWELREVDIKQREDLVGEKEHDLE

>oryza1

KGKGTTLRVAELEQELHEYQYNMGLLLIEKKEWTAKLDEINQALTOKEEILKREQAHLN
AISEYERREESMRKALGVEKQCVTDLEKALREIRGEIAEVKFMSEKKITDAQSLEASLEE
KRLEIEGKLHAADAKLAEANRKSQADRDLEEVEARQRRLEKEKLYFENERKAGEDRIKR
QEDSLRDWDKLLKESQNRILDLQRSLNDRERANENDKLFKIKQEELEEAKKALEHTKAT
LKIKEDDINKRLAELHLQEKEALEEREKKAEREEKVSAREKVGLQKLEEDHNVKLESKR
RDFDLQLENEKKSFDAMLVQKEADLVQREKDVRSSEEKLSKKEQVLN

>oryza2

DKAALIAYISRLESEIYQYQHNGLVLMERKELTSKHEQLRAASESAEIMHKRERAAQOS
ALAEARKKEENLKKSLGIQKECVANLEKALHDMRGETAETKVSYESKLAEALQMEAAHK
KFDEAEKLLAKSLEAESIRTHNAALRSLHDIDDREDQLRRDRISCELENEAKEKEISL
QRKSLNDMCKILHEKEEVLLKEQALLNQRDENILERLAYVTHSEKRVEEEKNILEAEKRV
LLEEKYKLELMEAIVSREEALLDKRESELLILQETIASKERAEIERLNQEQAIALERRK
HDFESEMANQMSFDAAMEVTRNALHQRECALSEQESVVVQORSQNLN

>musa1

DRDALVQRRISELEKELHEYQYNMGLLLIEKKDWASKYEEIRQALAEVDETLKKEKSACLA
SISEFAKREENLQKALGVEQQCVSDLEKALREMSELAEVKFTSDKKLDDAHALEIGLEE
KYLEVEQKLHAADAKLAEASRKSSDVRKLEDVEAREHKLQKEYLLFDSGRKLHEKDITE
QREHLRDWEQKLQDSQKRLVETQRYLNEREDRTNEADRVLKKKEADAEARKMIEATKKS
LKTKEEETIKRLGSLAAKEKEVLENKEKDLISREEKLNARERVEIQKLLDDHNLISSKK
EEFELDLKRRKSLSKEIECKIREVEKKRREIDSMEEQITKREQALQ

>musa2

DREALVQRILALEKELHEYQYNMGLLLIEKKDWALKYEEIRQALMDVEETLKREKLAHLA
SISEFEKREENLQKALGVEQQCVSDLEKALREMHSELAEVKFTSDKKLDDAHALEAGLEE
KYLEVEQKLHSADAKLAEASRKSSVANRKLLEDVEAREHKLQKEYLSLSSEWKLHEKGITE
QREHLCYWEKKLQDSQKRLVESQRFLNQREYQANEADRFBKKKEAELEESRKMIETKKS
LKSKEEDITIKLRSIAAKEKEILGKKEKDLFSREETLNARERVEIQKLLDDHNLALLISKR
EEFELNLEKRRKSFADLEGKVHEVEEKKREIDCMEDQVKKREQALE

>musa3

DREALVQRITTELEKELHEYQYNMGLLLIEKKESIARYEEVRQALAEAEIILKREQTAHLI
AISEYEKREETWLKDLGVEKQKVSALAKDLREVRFEISEVKFSSERKLSEHAHALETGLEE
KYLEIEARMHAADAKLAEAGRRNSETNRKLEDIEAHERKLQRDCLSLTSERKAHEKDLE
QREHLFDWEKRLQESQRRLVEEQRLNLEREDSANEADHILKKKETELETEATREAIASKRS
LKLEEDDITIRLSSLASKEKEALERKERELFAREEKLNARERVEIQKLLDDHNLMLDSSK
HEFELEMENQRKSFEEEMKAKIDEVEEIKKELDHKEEQILEREHALE

>musa4

DREALVRRRISELEKELYQYQYHMGLLLIEKKEWAVKYDKLRQEMSEAAQLQKCMQAAHIV
AVAEFEKSEGNLRRAMGFQRQSIHLEKALNDMHAIEIAEVKLDQKLLSEAHTEATIEE
KCLEIKEKQHSLDARLAKVSRKSSEVDRRLEDVEAREHELFFQOTSSFIAEKKAFKDLR
QRENLRRAWEQQLQDNQKKGWHSTENQREMETNERDNTFRKKEKELEEARKTLEISNEL
IKLKEEDMCMRIGALDAKEKEALEKKENELLAIEEQLNKERVEIQKITDFHNSILESQK
DEFELETEKKKRAVDEQLQGRIEEVAHKEIILENRERELFKKEQLLE

>musa5

DKAALISYITKLESEIYEQYQHMGLLILERKELVSKYEQVKASSDSAEIAYKREEAKRSS
ALAEARKRELNLEKLLGIQKECVANIEKALHDNLVESAERKLGYESKIAEAHAMMTAAQE
KLDEAEKLLAAESLQAEANRTRNTAIRTLDDVEAREDELRRRLATFKSQCDAKENEISI
QRQALYESQKTLHQQERFLEGQTLNQRREEYIFERTKELNRIEKELEESKANIEEESRT
LKLERSNLDLEIAALRNREEVILDKRERELLILQEKIACREHDEIQRIMEEHQSILEKKK
SELEADIEQRHLLLKNELEAKKIACEIREADLCSREISLQEKEHAIE

>eutrem1

DPRILPEKVSELEKELFEYQHNGLGLLLIEQKEWSSKYEELQOEFEEVNECLKRERNAHLI
AVADVEKREEGLRKALGIEKQCAVDLEKALRELRSNAEIKFTADSKLTEANALVRSVEE
KSLEVEAKLRAVDARLAEVSRKSSEVERKSKEVEARESSIQRERFSYIAERDADEATLSK
QREDLREWERKLOEGEERVAKSQMIVKQREDRANEGDKIIKQKGKELEEAQKKIDAANLA
LKKKEDDISSRIKALAFREQETIETKERELLALQEKLDAEKVAVQQLIDEHQAKLEAAQ

REFELEMEQKRKSIDDSLRSKVVEVEKREA EWKHMEEKVAKREQALD

>eutrem2

DREALLEKISTLEEELYGYQHNMG LLLL MENKEWVAKHEELNQAFQEAQEILKREQSSHLY
ALTTVEQREENLRKALGLEKQCVEELEKALREIQEENNKIRLTSEAKLAEANALVASVTG
RSSDVENKIYSAESKLAEATRKSSELEMRLKEVETRESVLQOERLSFAKERESYEGIFHK
QREYLHEWEKKLQEKES IPEQKRSLNQREEKVTEKEKNLKLKAKQLEEWDRKVELSVSK
SKETEEDMNKRLQELAAKEKESLVAKESSELRALEEKLI VREGTEIQKLIDDQKEALAAKM
LEFELECEERRKSLDRELQKKIEEVERQ RVEINHSEEK LQKRNEALN

>eutrem3

XXWTSTNEELQQAIDEASEILKRERMSKLI
ALNEAEKREENLRKALISEKQFAAELERDLKYLQQEHSEVKSTSEAKLAEANALVMGIKE
NSLEVDRKRAIAAEKLSVINRKSSELEKKLKEVETREKVLQREHLSLVTEREAHEA----
-----KKLTL EEDRLSEAKRSVNHI EERIIESEKTIKKKEKILK-----
--EKEEFINSMLNDISMKEKAFIDMKEKELHELEEKLI VREQMETGKLFDDQNAVLD SRK
HEFEMELEQMRI SLNEELERKKSEVEQLEVEISHKGEK LAKKESTLE

>eutrem4

DKAALIAYIAKLESEVYDYQHNMG LLLILEKDEL LSKYE EVKASVNEADLAHRRDQSAYVS
ALAEAKKREEDLKKDVGIAKECIS SLEKTLHEMRAECAETKVSAGSKMSEAHLMIEDALK
KYADAEAKMRAAEALQAEANRYHRIAERKLKEVESREDDLARRLASFKSDSETRENEIDI
ERQTL SERRKSLQOEHERLLDAQASLNQREDHIFGRSQELAELEKGL ESAKTTFEERRA
LEDKISNLEIALASLAKREEAVVLKKEQELLVAEEKIATKESELIQKVLANQEVILRKRK
SDVEAELESKCKLVEDEIESKRRAWELREVDIRQREDLVGEKEHDLE

>ambor1

EKDALLSHLAKLEDELFDYQYNMGL LLLIEKKEWTSKYDDLQALVEAQESLKREQASHLI
ALSEVEKREENLRKALGVEKQC VADLENALHEMRTFAEIKFTADKKLAEARSLVASIEE
KSLEAEAKLRSADAQLAEASRKS SNLERQLQEIETRESVLR RERQSLKAEREAHETTFNR
ERENLRNWERK LKEGQERL VESQGLLNQREELANEKEMFLTKKEKDLEVAWAKFEKGNLN
LKDKEVEMNMR LRS LTAQEE EALDEAQOELHLLQEKLNAREKEGIQKLLDEHNAVLELRK
REFDIEMDQKRKSL EEEFEKKQVVVEQKLV EVDLKEEKINKKEQLLE

>ambor2

DKAALISYITKLESEMFDYQYHMG LLLILEKKEWTSKYEQIKASADSAEDKYKRDQAALLS
ALAEAEQREENLQRALGVEKECVASIEKALHEMRAECAETKVA AETKLA EVRCLVEDAQK
KLLAVETKQHTVEALQTETSWQHAVAERKLKEVEAREDELRRQOVSLKSEMEAKEKDLLN
EKESLRELEKVIQQGQEKLFEGQ TLLNQREQC I KERSDRLSRLEKEVQAATVKLQEDLEI
LKEEKANLCLTSVALTTREEA I IDKKEQELLLLQEK LTSREQDEIRRLTIEHQTAIELRE
SQFEEELHEKHKSFEADLGLQRHALDLRDAELKHQEDLMHKDKHELD

>PhySc1

DRAALHARIAALEAELYDYQYNMGL LLLLQRK TWTSQADDLKA AVADAQETLQREKAAHLL
ELSEVMRREEA AKKALETEKQC VADLEKALKEFQTDESEVRQAADKQLAQARELVASIEE
RSVQADLKLAVQVVRADANRKLQ ESELRLQVEAREVALRRERHSLMADVDARKEQVAS
EEASLKEWEKRLEEGRARLQEG ERLLNERENSLKQRDEALKQTSRELAETR SYIENERAL
IKQTDADLNARVISLSERERTLILTKEQD LLLAEERIAEQTREEMEALLKIHTSDVDSKK
AELEEEELRSVRKTLKEEVETLKLAAEAREARSRHLETAITAREEEELK

>PhySc2

DRAALHARIAALEAELYDYQYNMGL LLLLQRK TWSSQVDELKAAVADAQGT LQREKAAHLL
ELTEVIRREEA AKSALETEKQC VADLEKALKEIQADESEVRQAADKQLAQARELVASIEE
RSIQADLKLAVQVVRADANRKLQ ESEHRLQVEAREVALRRLERHSLIADVEARKEQVES
EEASLREWEKRLEDGRMLQEG ERLLNERENSLKERDEALKQINREVAEARSYIEKERV L
IQKSDVDL NARAVAFSEKERALILKKDQDLILAERIADKTREEMEKF LKSHTSDVDSKK
AELEEEELRSVRKVLKEEVETLKLVAESREARLRHLEAAITAREEEELE

>selag

DRNALLAHISKLETELYDYQYQMG LLLLLESNKL RGESERLKSVI DETRDGLKREQSAHMI
ALQEAERREESLKRAVTTEKKCVADLEKALKEMHEEVAEAKAAAATQFQQGKATAMSAEE
KLLAEESKLHSAEALLAKANRKHADAERKLQEVESREDALRRQRHSFLAECEAHKLELEH
EKQNLKGWERTLEESQARFVENEKLLNKRE EYMQQRDDALTKLERDLDEARKVLEKDRSA

LRQEQA EYSALLSALS LREEAATKKEQE I LLLQEKLASRDRXXXXXXXXXXXXXXXXXXXX
XX
>pohlia
DRAALHARIAALEAE LYDYQYNMGLLLLQRKTWTSQVDELKAAVVDAQETLQREKAAHLL
ELSEVMGREEAAKNALETEKECVADLEKALKEVQADESEVRQAADKQLAQARELVASIEE
RSVQADLKLAVQVVRADANRKLQSEHKLQVEAREVALRRERHSLMADV GARKEQVEA
EEASLREWEKRLEEGRVRLPDGERLLNEREDALKMRDEALQOTSRELAETR SFLEKERAL
IQQSDVDLNRVVSLSERDRALVLKKEQDLILFEERIADKTREEMERLLKSHTSDVDSKK
AELEELRSVRKTLKEEVESLTLAAEAREARLRHLEAAITAREAEFE
>nitell
ERNFMVDHIKQLEDELYKSQYHMGLLLMERKSLLADVAKLDHTAKEYEEI IKREREESW
TRIEYEEIETTLKENLLAEQQCVKEVEKAIRDVQKEKDETTTLAELRFSEARALMSEVEK
KEHGVEISRKNVEATRAEFTRKKVEIESRARELETKE SLVRREREQVRIE----SSELEM
KRQELIERENVAREWEAKLASREQSVRKMEDRMKGREDTVLRKERDVEEGMARLDGKRKG
LSEMEDDLRLREKDLAMKEESLLSMKDEEMRSAKERLLESEHDDRQDI IEQQKQOMSSKE
NALELDLERKDRDLREEIEEQKALLNDSRRDAQQWEAKLEMRRSIE
>nithil
XXXXXXXXXXXXXXXXXXXXXXXXXXXXLFLMEKRSLLAETS KSDNTVEGYEEI IKRRREEW--
SIMDFEELQTLKDTYLKEQR--DVVEVIKDXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XX
KRQELTERENIAREWEAKLATREQGVKKTEERMKAREDTVLRKERDTEGLARLEGKRKG
LSEMEDDLRLREKDLSTKEEGLLNMKDEELRSAKERLLESEHDDRQEI IDQQKQQLSSRE
NAVELDIERKERDLREEIEEQKALLNDSRRKHSSGKR SVELRERNXX
>spirog
ERDALSIRLQELELEYGDYKLQTEAAFDERDRCLS QFSTIQSQLKEMVEKFKLEKNAHQ
TIASNTEKERKLRDALHTQRTADELELALTEKNLEFDNFRQSLTSLENETIRKKCIAEE
KILECEISLKESFSLKGEYKVKVEYESKLQETQITENLRLEYLRKDSQNEFQIQVSL
KEQALKEMEERSRXX
ILEKEEEMESKEKELIHFENETLENLEKRLKEMQERLNEQEKLETMKHLTSFEEALNFRQ
ESIELELESKEKEMKIKLLSLEENLKHKQVAVDEHKTALIKEINELQ
>coleo
LKEDLMNQLQSKENELKRYQFEMGVLLLENTNLRDELVAEELAAKDRELRLAEQNEHRE
MVAQMRTRREQQLHKSVEAERRCTADLERAVREM QNESMEAKKKAEVQMEESMALWKEAER
KRQEAQAQVRSQAALHREAEQKMGEIRRREEDVAAREREVNREHAKRLAEFEAEKAEVDR
RQREVASWDASLKECDGKLKEAQQLRQLEKSLEERSASILSNEKALEHRDSVLRAREEN
AKYKEEIEALEEKLA VQQLNTV GARERAVVEAGDRLAVAEADELSTRLLSDEAGLASQI
LKHSDWEQRHSHLESEIEKKRSACEAAQAQLQEKEENLRKREEAVE
>klebsi
EKEELQRQLDDVERELADYRYNLGV LILEHKKCKPQIDELEKALQRTREELQKEGRSLQL
ALDDVTREDGLRASLKAERTVIADLKQSL EEMHAKLQAEKSSRERLARAQELEAGAVK
DRERSEEDLKGAQQLEEGLVRESEL DGKLRDAASREKALNEGEAALNAG----KRELTE
KGIKLAERKRKVEGELEE--ERKKLGREREEEELRVEALRRSVEGEREQMQEMQAEKEA
MSTRKAEISVKEDSLAAQEKALLERKQEDANKRDQDL DQREKL----LGDEKREMLRT--
---EADLREQQKSMRRLAEQESEWQRIEAKIDARVKEIDAREAAVQ
>klebsub
XX
XX
XX
NGIELAGRKRKVEGELEE--ERKKLEWKRQEEELRIEALRRSVEGEREQMRKEMQEEREA
LSARKADISTKEESLAAQEKALLERKHKNAIKRGQDL DQREKL----LGDEKREMLRK--
---DADLREQEKSMRRLTEQESEWRRIEAKIDARVKELDGREAAVQ

NUP-1:

>Imaj

PSYGEQPARPPPAGVADLARPISVRELALTSLAHRDIDHLSGRELYNYASHLRESSQOYG
GYLSEAYAKRDQYRGEVARLKEELHNRYMEIDVLRREQERARDALLKVREDNAQLRQKLA
EAEGHARNMQAKISVLSGADPSLAVHFYQSOLALKDAQLRELOQECERAAAASSQRRSSS
TTAAEHGRRSANISGEGGTEAASAALHTALEELQCLNTSLSHRLEDEHTAHKAALAAHA
EASATATADRAELLDITIAQLRQOCADMQNVDELVAALTLRAPISKKDYAALQASYDDL
AALAKAEAQVAQENEQQHHSAQQAQ-EELRRSLTTGLEER---QSMQVRNE---QMQS
LVNHLEKELKLAERANQLRQEQLDAIISGNQQLAESLHVAQEQLLSATSELQDLRDTRES
LERSQVQATLERTVARLAEELVEQLNYDASELSDARAALTEEVTRLQKALAAVTDK
ADVLAQLLATQKELDRAAEVQDAQYRDEELLQNALATTTSAFDVQTKQVRLGMHVVDMS
RAFRCRVAEVAVQRSATGANVQEILSPWARALEWVAAMEQRLETLDAAFARLMDEQAAATN
TIADLRRTLVDKEVDLTNLRSDALASDLISERDRADNLSAAIDDASEKIASAQAEFEE
QLRQRTAATTAEVVEARRAEARATAAQLRAEEQLAVVEKDLKEQEGVLRKLDENHRLTR
EADRLLR-TSR-LEPSEDAVQRFPTTAPSTLEAVEQAQILQVSSLQADLLSRRQARELE
GREATLRQACAALQVSELRVEAAEAHLREDMATLRADYAELEQRYSELERMAIAAGG
GTMQELKQLRQQLKVREVELEELKARMRDLVLSKAAPELEAARRQEALRESLSGITTTOLA
ATQAQYLLSSSRPSPQDVLTSRIEHLQTLVNDRDVALEKVQAAQLESRAAMDTLEDQL
AAKTAALERAEGLVAEYADSVNAFTMAAVVGRGNRGAGPTASASVPATTRKRAR

>Linf

PSYAAQPARPPPAGVADLARPISVRELALTSLAHRGIDHLSGSELYNYASHLQESSRQYS
GYLSAAYAKRDQYRGEVAHLKEELHNRYVEIDALRREQERASDALLKVREDNAQLRQKLA
EAEGHARNMQAKMSVLSGADPSIAVHFYQSOLALKDAQLRELOQECERAAAASSQRRSSS
TTAAEHGRRSANISGEQDGAEEAASAALHTALEELQCLNTSLSHRLEEEHTAHKAALAAHA
EASATATADRAELLDITIAQLRQOCADMQSAEDELVAALTQRAPISKKDYASLQASYDDL
AALAKAEAQVAQEKEQQHHSAQQAQ-EELRHSLTTGLEER---QSMQARNE---QMQS
LVNNLERELELAERANQLRQEQLDVAISGNQQLAESLHVAQEQLLSATSELQDLRDTRES
LERSQVQATLERTVARLAEQEALVEQLNYDASELNDARAVLTDEVTRLQKALAAVSDK
ADVVAQLLATQKELDRAEAQDAQYRDEERLQDALATTTSAFDVQTTQVRLGMHVVDMS
RAFRCRVAEVAVQRSATGTSVQEILSPWTRALEWVAAMEQRLEALDAAFARLIDEQAAATN
TIADLRRTLADKEVDLTNLRSDALASDLMGERDRADNLSAAIDDASEKIASAQAEFEE
QLRQRTAVTTAEVVEARRAEARATAAQLRAEEQLAAVEKDLKEQEAFLRKLQDENHRLTR
EADRLLR-TSR-LEPSENAQRFTTAPSTLEAIEQAQILQVSSLQADMLLSRRQARELE
GREATLRQACAALQVSELRKAAEAHLREDLATLRADYVELEQRYSELELMATAAGG
GTMQELKQLRQQLKVREAELEEVKARMRDLVLSKAAPELQAARRQEALRESLSGITTTOLA
ATQAQY-LNLSSSRPSPQDVLNSRIEHLQTLVHDREVALEKVQAAQLESRAAMDTLEDQL
AAKTAALERAEGLVAEYADTINAFTMAAVARRGNRGAVSTAGASVPATTRKRAR

>Ldon

PSYAAQPARPPPAGVADLARPISVRELALTSLAHRDIDHLSGSELYNYASHLQESSRQYS
GYLSAAYAKRDQYRGEVAHLKEELHNRYVEIDALRREQERASDALLKVREDNAQLRQKLA
EAEGHARNMQAKMSVLSGADPSIAVHFYQSOLALKDAQLRELOQECERAAAASSQRRSSS
TTAAEHGRRSANISGEQDGAEEAASAALHTALEELQCLNTSLSHRLEEEHTAHKAALAAHA
EASATATADRAELLDITIAQLRQOCADMQSAEDELVAALTQRAPISKKDYASLQASYDDL
AALAKAEAQVAQEKEQQHHSAQQAQ-EELRHSLTTGLEER---QSMQARNE---QMQS
LVNHLERELELAERANQLRQEQLDVAISGNQQLAESLHVAQEQLLSATSELQDLRDTRES
LERSQVQATLERTVARLAEQEALVEQLNYDASELNDARAVLTDEVTRLQKALAAVSDK
ADVVAQLLATQKELDRAEAQDAQYRDEERLQDALATTTSAFDVQTTQVRLGMHVVDMS
RAFRCRVAEVAVQQSATGTSVQEILSPWTRALEWVAAMEQRLEALDAAFVRLIDEQAAATN
TIADLRRTLADKEVDLTNLRSDALASDLMSEDRADNLSAAIDDASEKIASAQAEFEE

QLRQRTAVTTAEVVEARRAEARATAAQLRAEEQLAAVEKDLKEQEAFLRKLQDENHRLTR
EADRLLR-TSR-LEPSENAQRFTTAPSTLEAIEQAQILQVSSLQADMLLSRRQARELE
GREATLRQACAALQVSELRKAAEAHLREDLATLRADYVELEQRYSELELMATAAGG
GTMQELKQLRQQLKVREAELEEVKARMRDLVLSKAAPELQAARRQEALRESLSGITTTOLA
ATQAQY-LNLSSSRPSPQDVLNSRIEHLQTLVHDREVALEKVQAAQLESRAAMDTLEDQL
AAKTAALERAEGLVAEYADTINAFTMAAVARRGNRGAVSTAGASVPATTRKRAR

>Lmex

PSYAAQPARPPPAGVADLARPISVRELALTSLAHRDIDHLSGSELYNYASHLQESSKQYS

GYLSEAYAKRDQYRGEVARLKEELHNRYMQIDVLRREQARATDALVKVREDNAQLRQKLA
EAEGHTRNMQAKMSVLSGADPSIAVHFYQSQLALKDAQLRELOQECERAAAASSQRRSSS
TTAAEHGRRSANISGEQERAEASAALPTALEELQRLNISLSHRLEEEHAAHKAALAAHA
EASATARADRAELLDITIAQLRQOCADMQSAEDELVAALTQRAPISKKDYAALQASYDDL
AALAKAEAQVAQEKEQQHDHIARQAE-EELRHLLTTGQEER----QSMQARNE---HMQS
LVYHLERELELAERANQLRQEQLDAAISGNQQLAESLHVQEQLLSATAELQDLRDTCS
LERSQVQATLERTVARLAEELVEQLNYDASELNVARAALSDEVTRLQRALAVALTDK
SDVVAQLLATQKELDRAAELODAQYRDEERLQDALATTTSAFDVQTKQVRLGMAHVVDMS
RAFCRVAEVAVQWSATGASVQEMLSPWSRALEWVAGMEQRLEALDVAFARLMDAQAAAMN
TIADLRRTLADKEADLTNLRSAMDALASDLMGERDRADNLSAAIDDASEKIGSAQAEFEE
QLRQRTAATTAEVVEARRAEARATAAQLRAEKQLAAVENDRKEQEAVLRKLQDENHRLAR
EADRLLR-TSR-LEPSEDAVQRFTTAPSTVEAAEQAIQVSSLOADLMLSRRQARELE
GREATLRQACAALQVNELEVEAAEAHLREDLATLRADYAELEQRYSELERMATAAGG
GTMQELKQLRQQLKVREAELEELKARMQDLVLSKAAPELEAARRQEALRESLSGITTTQLA
ATQAQYLLSLSRRSPQDVLTSRIEHLQTLVHDREVALEKVQAAQLESRAAMDMLDQ
AAKTAALERSEGLVAEYADTINTFTMTAVVRRGNHGARPTAGASMPAATRKRTR

>Lbra

PSYTAQPARPLPAGVADLARPISVRELALMSLAHRDIDNLSGDDLYNYASHLQGASKQYS
KHLLEAYTRRDQYRGEVARLKEELHNSKYLQIDVLRREQERASEALVKVREDNAQLHQKLA
ESEGHVRNMQAKVSVLSGADPSAAVHFYQSQLVLDKAQLRELOQECERAAAASSQRHSSS
ATAAEHGRRSINVSGEHNGSEAASTALHTALEELQHLNTSLSRLEEEHAAHKAALAVHA
EENATATADRAELLDITISQLRQOCTEMQSTEDDELVAALTQRAPISKEGYAALQTSYNDLT
AALAKAEDQIAHVKEQQYHQSAQQAQ-EELRHSLTTGLEER----QELQARNE---QMOS
LATDLQRELELAERANQLRQEQLDAAVSGNQQLAEHLRTTQEQLLGATAELQDLCNTRGS
LERSQVQSTLTRTVDRLAEQEALLEQLNRNASELNDARAVLTGEVTRLQAALAAAVTDK
SDVMAQLVATQRELDRAVEQQAQYCEEVRLQDALTTTMSAFGVQTRQVQLGMAHVVDMS
KAFYRVVEVVAKGAATATSREILLPWSRALEWVASMEERLES LGAAFARWIWEQAAAAS
TIADLRSLADKEVELNLRSAMDTLANDAMAEQDRADNLSAAIGDATEKIALVQAEFEE
QLRQRTVATTAEVVEARRAEERARAAQLRAEEHLTVVENELKEQAVLRKLQDENHRLTR
EADRLRR-SSRHLESDSVTEVESFTTAASAVEVVEQAQFLQVFSLQEDLMLSRRQARELE
GREATLRQTCATLEQQVSQLRVEATDAEKLREELATLRDYAELEERYDQLERMTTAAGG
GTMQELKQLRQQLKARETEDELKARVRLVLSKVAPELEAARRQEALRESLSGITTTQLA
ATQAQYHLSLSASSPSSQDLLTSRIEHLQILVHDHQEAFQVQAAQLESRATMDTLEDQ
AAKTAALERAEGVVAEYADTINVLMTAVVGRGSGGAHLSVGVMTIAKARKRAR

>Lger

PSHAAQPARPPPAGVADLARPISVRELALTSLAHRDIDHLSGRELYNYASHLRESSQOYG
GYLSEAYAKRDQYRGEVARLKEELHNRYMEIDVLRREQERARDALLKVRRDQNAQLRQKLA
EAEGHARNMQAKISVLSGADPSLAVHFYQSQLALKDAQLRELOQERERAAAASSQRRSSS
TTAAEHGRRSANISGEQDGTAAASAALHTALEELQCLNTSLSHRLEDEHTAHKAALAA
EASATATADRAELLDITIAQLRQOCVDMQSVDELVAALTTLRAPISKKDYVALQASYDDL
AALAKAETHVAQENQKHLHSAQQAQ-EELRHSTTTALEER----QSMQARHE---QMOS
LVNHLERELKLAERANQLRQEQLYAAISGNQQLAESLHVQEQLLSATSELQDLRNTRES
LERSQVQATLERTVARLAEELVEQLNCDASELNDARAALTEEVTRLQKALEAAVTDK
ADVLAQLLATQKELDRAAEVQDAQYRDEERLQNALATTTSAFDVQTKQVRLGMAHVVDMS
RAFCRVAEVAVQQPATGASVQEMLLPWTRALEWVAVMEQRLETLDAAFARLMDQAAATN
TIADLRRTLADKEVDLTNLRSAMDALANDLISERDRADNLSAAIDDASEKIASVQAEFEE
QLRQRTAATTAEVVEARRAEARATAAQLRAEEQLAAVEKDLKEQEGVLRKLQDENHRLTR
EADRLLR-TSR-LEPSEDAQRFTTTTPSNLEAVEQAQILQVCSLQADLLSRRQARELE
GREATLRQACAALQVSELEVEAAEAHLREDLATLRADYAELEQRYSELERMAIAAGG
GTMQELKQLRQQLKVREVELEELNARMRDLVLSKAAPELEAARRQEALRESLSGITTTQLA
ATQAQYLLSLSRRSPQDVLTSRIEHLQTLVNDREVALEKVQAAQLESRAAMDTLEDQ
AAKTAALERAEGVVAEYADSVNAFTMAAVVRRGNRGAGPTASASVPATTRKRAR

>Lama

PSYAAQPARPPPAGVADLARPISVRELALTSLAHRDIDHLSGNELNYNYASHLQESSKQYS
GYLSEAYAKRDQYRGEVARLKEELHNRYMQIDVLRREQARATDALVKVREDNAQLRQKLA
EAEGHTRNMQAKMSVLSGADPSIAVHFYQSQLALKDAQLRELOQECERAAAASSQRRSSS

TTAAEHGRRSANTSGEQERAEASAVLPTALEELQRLNTSLSHRLEEEHAAHKAALAAHA
EASATARADRAELLDTTIAQLRQOCADMQSABEELVAALTQRAPISKKDYAALQVSYDDL
AALAKAEAQVAQEKEQQHDHIARQAE-EELRHLLTTGQEER----QSMQTRNE---HMOS
LVYHLERELELAERANQLRQEQLDAAISGNQQLAESLHVVEQLLSATAELQDLRDTRES
LERSQVQATLERTVARLTEEEALVEQLNYDASELNVARAALSDEVTRLQKALAVALTDK
SDVVAQLLATQKELDRAAEQDAQYRDEERLQDALATTTSAFDVQTKQVRLGMAHVIDMS
RAFCRVAEVTVQWSATGASVQEMLSPWSRALEWVAGMEQRLEALDVAFARLMDAQAAAMN
TIADLRRTLAEKEADLTNLRSATDALASDLMGDRADNLSAAIDDASEKIASAQAEFEE
QLRQRTAATTAEVVEARRAEARATAAQLRAEKQLAAIENDRKEQEAVLRKQLDENHRLAR
EADRLLR-TSR-LEASSEDVQRFTTAPSTVEAVEQAQILQVSSLOADLMLSRQARELE
GREATLRQACAALQVNELELVEAAEAHLREDLATLRADYAELEQRYSELERMATAAGG
GTMQELKQLRQQLKVREAELEELKVRMRDLVLSKAAPELEAARRQEALRESLSGITTTQLA
ATQAQYLLSLSSSRPSPQDVLTSRIEHLQTLVHDREVALEKVQAAQLESHAAMDMLGDQL
AAKTAALERSEGLVAEYADTINTFTMTAVVRRGNHGARPTAGASMPAATRKRTR

>Ltur

PSHAAQPARPPPAGVADLARPISVRELALTSLAYRDIHLSGRELYNYASHLRESSQOYG
GYLSEAYAKRDQYRGEVARLKEELHNRIMEIDVLRREQERARDALLKLQEDNAQLRQKLA
EAEGHARNMQAKISVLSGADPSLAVHFYQSQLALKDAQLRELOQERERAAAASSQRRSSS
TTAAEHGRRSANISGEQDGTAAASAALHTALEELQCLNTSLSHRLEDEHTAHKAALAAHA
EASATATADRAELLDTTIAQLRQOCADMQNVDELVAALTTLRAPISKKDYAALQASYDDL
AALAKAEAQVAQEKEQQHLHSAQQAQ-EELRRSLTTGLEER----QSMQVRNE---QMS
FVNHLERELKLAERANQLRQEQLYAAISGNQQLAESLHVVAQEQLLSATSELQDLRNTRES
LERSQVQATLERTVARLAEAAALVEQLNYDASELNDARTALTEEVTRLQKALAAAVTDK
ADVLAQLLATQKELDRAAEVQDAQYRDEERLQDALATTTSAFDVQTKQVRLGMAHVVDMS
RAFCRVAEVAVQQSATGASVQEMLSPWTRALEWVAAMEQRLETLDAAFARLMDEQAAATN
TIADLRRTLADKEVDLTNLRSAMDALANDLISERDRADNLSAAIDDASEKIASAQAEFEE
HLRQRTAATTAEVVEARRAEARATAAQLRAEEQLAAVEKDLKEQEGVLRKQLDENHRLTR
EADRLLR-ASR-LESSEDAQRFTTAPSTLEAVEQAQILQVSSLOADLLSRRQARELE
GREATLRQACAALQVSELVEAAEAHLREDLATLRADYAELEQRCSELERMAVAAGG
GTMQELKQRRQQLKAREVELEELKARMRDLVLSKAAPELEAARRQEALRESLSGITTTQLA
ATQAQYLLSLSSSRPSPQDVLTSRIEHLQTLVNDREVALEKVQAAQLESRAAMDTLEDQL
AAKTAALERAEGVVAEYADSVNAFTMAAVVRRGNRGAGPTASASVPATTRKRAR

>Larab

PSHAAQPARPPPAGVADLVRPISVRELALTSLAYRDIRLSGRELYNYASHLRESSQOYG
GYLSEAYAKRDQYRGEVARLKEELHNRIMEIDVLRREQERARDSLLKVREDNAQLRQKLA
EAEGHARNMQAKISVLSGADPSLAVHFYQSQLALKDAQLRELOQERERAAAASSQRRSFS
TTAAEHGRRSANISGEQDGTAAASAALHTALEELQCLNTSLSHRLEDEHTAHKAALAAHA
EASATATADRAELLDTTIAQLRQOCADMQNVDELVAALTTLRAPISKKDYAALQASYDDL
AAHAKAEAQVAQEKEQQHLHSAQQAQ-EELRRSFTTTGLEER----QSMQVRNE---QMS
FVNHLERELKLAERANQLRQEQLYAAISGNQQLAESLHVVAQEQLLSATSELQDLRNTRES
LERSQVQATLERTVARLAEAAALVEQLNYDASELNDARAALTEEVTRLQEALASAVTDK
ADVLAQLLATQKELDRAAEVQDAQYRDEERLQDALATTTSAFDVQTKQVRLGMAHVVDMS
KAFCSVAEVAVQQSATGASVQEILLPWTRALEWVAAMEQRLEMLDAAFARLMDEQAAATN
TIAELRRTLADKEVDLTNLRSAMDALANDLISERDRADNLSAAIDDASEKIASAQAEFEE
QLRQRTAATTAEVVEARRAEARATAAQLRAEEQLAAVEKDLKEQEGVLRKQLDENHRLTR
EADRLLC-TSR-LEPCSEDAQHFTTAPSTLEAVEQAQILQVSSLOADLLSRRQARELE
GREATLRQACAALQVSELVEAAEAGEHLREDLATLRADYAELEQRYSELERMAIAAGG
GTMQELKQLRQQLKVREVELEELKARMRDLVLSKAAPELEAARRQEALRESLSGITTTQLA
ATQAQYLLSLSSSRPSPQDVLTSRIEHLQTLVNDREVALEKAQASQLEYRAAMDTLEDQL
AAKTAALERAEGVVAEYADSVNAFTMAAVVRRGSCGAGPTANASVPATTRKRAR

>Ltrop

PAYAAQPARPPPAGVADLARPISVRELALTSLAHRDIDRLSGRELYNYASHLRESSQOYG
GYLSEAYAKRDQYRGEVARLKEELHNRIMEIDVLRREQERASDALLRVREENAQLRQKLA
EAEGHARNMQAKMSVLSGADPSLAVHFYQSQLALKDAQLRELOQECERAAAASSQRRSSS
TTAAEHGRRSANISGQDGTAAASAALHTALEELQCLNKSLSHRLEDEHTAHKAALAAHA
EASATATADRAELLDTTIAQLRQOCADMQNAEDELVAALTTLRAPISKKDYAALQASYDDL

AALAKAEAQVAQEKEQQYHHSAQQAQ-EELRHSLTTGLEER----QSMQVRNE---QMOS
LVNHLERELELAERANQLRQEQLDAAISGNQQLAESLHVVAQEQLLSATSSELQDLRNTRES
LERSQVQATLERTVARLAEVEALVEQLNFDASELNDARAALTEEVTRLQKALAAAVTDK
ADVVAQLLATQKELDRAAEVQDAQYRDEERLQDALATTTSAFDVQTKQVRLGMHVVDMS
RAFRCRVAEVAVQRSATGASVQEMLSPWTRALEWVAAMEQRLEALDAAFARLMDEQAAATN
TIADLRRTLADKEVDLTNLRSDALASDLMSQDRADNLSAAIDDASEKIASAQAEFEE
QLRQRTEATTAEVVEARRAEARATAAQLRAEEQLAAVEKDLKEQEGVLRKQLDENHRLTR
EADRLLR-TSR-LEPSSDAAQRFTTAPSTLEAVEQAQILQVSSLQADLLSRRQARELE
GREATLRQACAALQQISELRVEAAEAHLREDLATLRADYAELEQRYSELERMATAAGG
GTMQELKQLRQQLKVREVELEELKARMRDLVLSKAAPELEAARRQEALRESLSGITTTQLA
ATQAQYLLSLSSSRPSPQDVLTSRIEHLQTLVNDREVALEKVQAAQLESRAAIDTLEDQL
AAKTAALERAEGLVAEYADRINAFTMAAVVRRSSHGAGPTAGASVPATTRKRAR

>Laeth

PPYAAQPARPPPAGVADLARPISVRELALASLAHRDIDRLSGRELYNYASHLRESSQOYG
GYLSEAYAKRDQYRGEVARLKEELHNRYMEIDVLRQQERASDALLRVREENAQLRQKLA
EAEGHARNMQAKMSVLSGADPSLAVHFYQSQALAKDAQLRELOQECERAAAASSQRRSSS
TTAAEHGRRSANISGEQDGTAAASAALHTALEELQCLNTSLSHRLEDEHTAHKAALAAHA
EASATATADRAELLDITIAQLRQOCADMQNAEDELVAALTLRAPISKDYAALQASYDDL
AALAKAEAQVAQEKEQQHHSAQQAQ-EELRHSLTTGLEER----QSMQVRNE---QMOS
LVNHLERELELAERANQLRQEQLDAAISGNQQLAESLHVVAQEQLLSATSSELQDLRDTRES
LERSQVQATLERTVARLAEVEALVEQLNFDASELNDARAALTEEVTRLQKALAAAVTDK
ADVVAQLLATQKELDRAAEQDAQYRDEERLQDALATTTSAFDVQTKQVRLGMHVVDMS
RAFRCRVAEVAVQRSATGASVQEMLSPWTRALEWVAAMEQRLEALDAAFARLMDEQAAAAN
TIADLRRTLADKEVDLTNLRSDALASDLMSQDRADNLSAAIDDASEKIASAQAEFEE
QLRQRTEATTAEVVEARRAEARATAAQLRAEEQLAAVEKDLKEQEGVLRKQLDENHRLTR
EADRLLR-TSR-LEPSSDAAQRFTTAPSTLEAVEQAQILQVSSLQADLLSRRQARELE
GREATLRQACAALQQISELRVEAAEAHLREDLATLRADYAELEQRYSELERMATAAGG
GTMQELKQLRQQLKVREVELEELKARMRDLVLSKAAPELEAARRQEALRESLSGITTTQLA
ATQAQYLLSLSSSRPSPQDVLTSRIEHLQTLVNDREVALEKQAAQLESRAAMDTLEDQL
AAKTAALERAEGLVAEYADSIINAFTMAAVVRRGNHGAGPTAGASVPATTRKRAR

>Lenri

HSHPARLARPPPAGVADLARPISVRELALTSLAHRDIDHLSRNDLYNYASHLQDASKOYS
GYLLEAYTKRDQYRSEVAHLKEELHNRYMQIDVLRREQERASDALLSVREENAQLREKLA
EAEGHARNMQAKMSVLSGADPSVAVHFYQSQALAKDAQLRELOQGCERAAAASSQRRGSS
TTAADRGRLSTNISSEQDGTDAASAPLHITTEELQCLNTSLGRRLEEEQISHKAALVAHA
EASATAAADRAQLLDITIAQLRQOCTEMQAAEDELVTALTQRAPVSKKDYALQASYDDL
AALAKAEAHVQGEKAQQHHQSAQNAQ-EELRHSLSIGLEER----QALQARYE---HMES
LVNHLQRELELAERANQLRQEQLDAALNGNQQLAEHLRVAQEQLLSATSSELQDLRKTHTGD
LERSQVQATLLECTLTRLAEQEALVERLNREALEFHQRPALEREVAELRVALAVAMTDK
VDVVAQLVATQKELDRAAEVHDEHYREEERLHEALSTTTSAFAVQTAQVRLGMHVVDMS
RAFCGVAEAVAQRAASVKCLREMLLPWTQALEWVAAMEKRLDALDAAFARWMDVQSAAN
IIAELRSALAEAEVNSLRSAMDALATELMDEQDRADNLSAAIDDATEKIASAQAEFEE
QLRQRTEATTAEVVEARRAEARATAAQLRAEERLTVVEDDLKEQQALLRNLQDENRRLTR
KADCLLH-SSRLTDPDSEGAAKRFATAASAVEAVEEALILQVSSLQADLMLSRRHARELE
GREATVRQACLALQQVSELRVEAADTGQREDFAALRADYTDLEQRYDELERMTTAAAGG
GTMQELKHLRQQLKVQEAELCESKARLRGLVLRKATPELAATTRQEALRESLSGITTAQLA
ATQAQYPLNRSAPSPSSQDVLTSRIEHLQLVVRNREMAIDKVQAAQLESRAAADTLEDQL
AMKTAALERAEGLVAEYADTIDTLTMTVLAGRGRGARPTAGATMRATTRKRDR

>endotr

SSYSAPPAHSLPAGVADLARPISVRELALMSLAQRDIDRLSSDELYIYASHLQDASKQFT
AHLQEAAYAKRDQYRSEVAHLKQELQSRVVEVDLLRRERERASKALLGVHEENTTLREQLA
EVEGRLRNMQAKMSVLSDDSSLALHFYQSQTLKQAEQMRLEQDYERVAASLQRCSTS
TAGAASARHSINIPVEQDSEEVVAVELRATLEELRSLNASLSHRLEEEERAHDAALAAHT
EAIAAAAADRTELLDITFQLRQRCTEMQASEDELVAALTQRAPISKSDYAVFQSSYNDVT
AALAKAEAQITREKERHEYASARQAQ-EELHSSLALGMEEK----NTLQARNE---QMOS
LVHHLQKELELAEEANQLRQKQLDAAVNGNQGLANNLRMAQQQLLSMSSEVEDLRNTRTN

LERGEAVQORALDEALDRLSEEEVLVEELNGDAAQHSKEMALAESVAMLQKSLSTAIIVDK
EDLSAQLVATQKEVDRAIGQODVLLDEAERLKDLEATTSAFAAQOTREVRMGMAHIVDMS
NVFCGVLEVAVAGAAHTTAKIGEALSPWDEALEWVATMEQRLEVLDTLARLFCEQASAKR
TIADLRVSLAEKDVELGNVRFAMDALASEAITERDRADNLSAAIDDATKKIVSVQAEFEE
QLKQRMATTTAEVVEARRAEARATASRLCAEEQLSAVEDDMKEQQALVRTLQDENRRLTH
DVRLLR-SLRSLEGSRIASQPNTSRPSIAETIEQAQILQVSSLQADLMRTRRQTRDLK
GREATLQLSCATLEKELSELRPRAREAEQLRGDLDALRADYAELELRHDQLERMTAAAGD
GTMQELKQLRQQLKVREAELEEQLKFRDLLLLSQATPELEVARRQEALRESLTDITTHLV
ATQAQYSLTPPPSGPSATDVLTDNRMYLQALVRDREVELEKVQAAQLESRAATDTLRDQL
VAKAAALERAEGVLAQYTDITNTLTAMRGGGECGRLNGKVSAGTSSRKRSR

>leppy

SAHSSHPPPPAGVPLARPI SVRDLALTS LAERNIDHLSNDDLYNYASHLQDTAKRYS
HHLAEAHAKRDMYRNLVAVQKRELQEKYVQIDRQRHEQETATAALRRLREDNTQLREQVA
TAEQVNRNIAAKLSLQNGADPSVAIRAYQAQLASKEAQVRDLMERVVHASAEASQRRSSE
--GGANPROSDTLGAAAAAASEMAKLOAELNALHQRNAALQEQLAAQDKTHAEAL----
---ASMTAEHTELKETIAQLKQGHAEQENEDALVAALTQRAPVSKTDFEAFQATYAEMQ
AALTKTEATVATLQAQNASAVQDGYIEVLQQLSATSQAER----QELTEKNR---QLQA
LSENQREVQLADTANRLRQQQLETA VAGGEQLAQEVQTQEARHTEVHAEIVALRARCA
ADRGEELQRLDITLDRLAQEEELVKGLNDNAADLHAERAALTEEVVTVRLALETVVLER
QDIQERLDAAKLGLQTAELYNERILEDDRLQDELDAVTQAFSAQTMQVKQGMHMLQMS
AGFHDALAEAVMALAAAVTAINAALTPWSAALEWVATMQRRVHALEGALDFVELQEALLT
HLRRLGHALDERDYELRLLRSVDEMANSNVAEKERADHLLLEAIQDAKNHVS LAEKEFDD
QLQLRSEAVAAELVARRAEERATDARLRAESQLAAVGAELTDALAAHLKTEERRKMSR
EVERLSRFSARLADPGAEGVSNHYTMTASAVEALEQAHLQVSSLQOQELLSRRQIRELE
RQEVILRQANVTLEQQVSDLRADAEAVDLREEVARLRAENADLEEKIEQLERILAASGG
GVAHEVKQLRQQLKREAELEEEKARMQGLLVSKTAPEFAAARRQAALRESLSGITTTQLA
ATQAQY--NSAATSASPLDVL SKRIDHLQSAMKERDALIEKMKAAQLQSHVTVDGLERL
SEKAAALERANALVTHYAEALDGLRAARVSA-----ARSGTREGASHRKRTR

>Crifas

TSLPQSSRKPPPVGVPHLARPISVRDLALTS LAARNIDHLNFNDLYTYASHLQDAAKNYS
NNLSAAYTKRDMYRSQVAELKQELQERYLQVDLLRRDHERADDALLRLREDNTQLRQLLA
ESEGKLRNV SARLSLNGADPSAAMRAYQAQLAQKDAQLRELVERLAQQTAEAAQRLGSE
A--SAAGQRRSHTLSAAASAI EAATTVLQAEIDALHQHATLQSQLAAETASH-----
-----TQERTELLETIAALRREKGMQESDALIAALTLRAPVSKADFESLQASYAELQ
ASLANSEASVASLREQLSKAVVQEGYTEVLQQLSTTTQAER----QELVTKTE---QLQA
LVDNLTRELQLADEANQLRQQQMEETAVANS DRWQQSQ-----TIQELQEREAS
QVRGEELQTMLDITLDRLAQKELLLEEVNDNAEELHKERAALADDVVTLR SALETAVVER
QDLQERLAAANEGLERTAE LYNERILEDDRMQEELNAVSAFAAQTOEMRAGLAAMTEMS
AGFCNVLETVMAPAKLISAITIALAPWSAALDWAGKMERRLDALESALLDFVQYNEAALG
HLHHIRRTMADRDELRLQRSALDSMADATVEERERVDHLLSAIQDAKDHVARAEREFDE
QLRLRSEAITAELVARRAEERATDARVRAETE LAATAEELRESHLALRKLQEEHRKLTR
ETERLSRYSARLTD TSAETATNSYAVTASAVEALEQSNLLQVSSLQODLLSRRHARDLE
GREALLRQNHSTLEQEVVSLRVEAAEAADLREEVATLRADNAELEERCEQLERMLAAAGG
GTMQELKSLRTQLKLETTELADLRARMDFLLNRSTSEFEVVRREGDLRETLSGITTTQLA
ATQAQYAGGTAGDEAS PQVLS SRIDHLENVVRERDAVIEKMQASQLESRVTVDGLQENL
SEKEAALDRANTLLAQYAETIETLHSAGVAGTTTRQLPATSATTTAAAGRKRGR

>CritAcant

SASMSATAPRPPSSVPLARPI SVRDLALTS LAYRHIDHLSDDLYNYASHLQDAAKHYS
SHLSEAYAKRDAYSQVAELNHRLEEEYKTDLLRRNREKADDALQKLREDNTELRHLLA
EAEQTRNLSAKLSLQSGAEPVSVIRALQAQLAQKDAQLREMSERLLHATAELQRRSSS
SGGGAHPRASNASSAAEQAVAAATAHLRAAMEDLQQRHDALQTQLTTQATAHAAAL----
---GRAAAQRAELVDAMTLLKQEKAELOANEDDLVAALTQRAPVSKADFESFQAAYAELQ
ASLRTAESAAATLQEQNASAAQEGYVSALQHSLSLTTQAER----QALADANR---QLQA
LVDNLQKELELAGTSLRQQEQVASAHANEEQLAAQLQTLAAEV--MSTAQQELERRCRO
LERGEELQGLLNTLDRLTQTEALVKELNDNAEELHDERAALAAQLATVQAAFETVVHER
QDLQERLAAAQLGLEQTTDLYNERILEENRVQDELRAVAKAFAAQTEQVKVGI AHVLHMS

EGFRGAVTAVMEHATAVTAVNAAFAPWAVTMEWVAQMQRLOKLENALLDFVEQQTMSLR
HLQTVQHALNHRDEELQLLRSALDTMASTNLDEKQRADHLLQAIQDARRRVTMAEQDFDA
QLRLRSEAVAADLVVARRAEERATDARLRAETQLETVEAELRTTQAALQKLEEEKRQLTR
ETERLSRYSTRVVDVSAESTSTHYTMNAAVEALEQAHLLOVSNLQQELMLCRRQVRELE
GQETTLRQAYMVLEQQVVDLRTDAAEVEDLREERDTLRAENAWEWEEKNAQLQRVMDATOR
STHQELSQRQQRVLESELEEKTRQLQVLMIGKSTPEFEALRRQEFKENLSGLTTELA
ATQAQY-----AAASPQNVYITRIEHLQTAIKERDVVIEKLOAAQLQSRVAKDSLEEKL
SEKAAALSRVSELNAQYLETIEALRSAGVSAA-----TRHLVPRRMLARQRT

>Crimelif

SPSSSTAPTPPPKEVADLARPI SVRELALASLAERHIDRFSHDDLNYASHLQDTAKRYS
NNLSEAYTARNMYREQVAQLKQELQOKTLQVDLQRRDRERCDAAVQKLREDNTQLRLQLA
EAEGQTRNIAARLSLQNGADPSVIRAYQAQLAQRDAQLRELTERLACESSVAEARQGS
DGGEADTRPSRALGAAQTVAAATAELQAE LNDLQQRKAALERQLTAQETAHAAAL----
---ARCAAERTELQVTIERARQEQAEQENEDALIAALAQRAPVSKANFEALQTSFAETQ
AALTKEAAVATLQEQRTVAAQEGYVEVLQALATGQAE-----QELTDQNR---QLQS
LVDNLQREVELADTANQLRQQLELALOGGELAEQVQAEAMHAAHMSEEIGALQERCAA
SERAEELRGVLSSTLDRLAQKEELEKQLNDNAGELAANNAALTEELATLRQTLEMALAE
RELLERLDAANRGVEQTASMYNERILDMDRLQGE LDAVTQAFAAQTEQLKSGLTRVVGMS
DGFRRAVEAAMAPAAA VAAVKAALTPWSETLDWVAATRQRIEELLESALLDFVEQQTSLA
HLRRLRHALHERDQAAVDKQAAL EAVAGGAVDAQERADGLFRATEDARHRILQAEREWEE
QLQLKSQATAAELIVARRAEERANDARTRAEALAAAEKELKETQAALRTTQEEKQLAL
QTERLSLLSSQSDSSAGSRTRYTVSSSTEAVEQTHLLQVAGLQQLVRCRRQIRELE
SEEATLRQAYMALEQEVSDLRSDAAEVNLRAE MATLRADYADSEERYERLRLMVDVAVGD
GTAQENRRLRQELKLDAAELEEYKTRQENSLVSRGAPDFETQRLEEA FRQSMGIATQLT
AKQAEF-----SLSASPQDVLNSRIDHLQTLVRRERDAVISKMKAADLQSRVKCEGLEEKL
AEKTAALNGAEELLAQYMERLDAMRAAGLAAAVTTRA---PKVGKTVSTRKRAR

>Herpet

PRASSSSAATYRQPTPELARPI SVRELAFASSAQKSIDMMSPRELYDHTSRLQDEARQTR
GYVAAAYAARDQYREEVGRRLRQELHSRYLQSQALERDRGRYNDVIHKLQGDIAALRRDL
DAEARENALQGRLARANGTRPPAALAYYATQLAAKDGQLELEAENTALRAAQQQQ-----
-----RSAPALTGE---DQRVAAQLQRALEDLQENAVLADALATERESRAALL----
---AETAEEKTRMDADATRAAASNAALQKELDDLLEAFSSRAPITNEHYEHLRRELAEQQ
RTIAALESEAAAAEAVERADGAAELEASALRESVATLRA-----ELADALKQLAKTQE
LVAEFEKETWLAETANTVRQGLIDDLRANLQAAATAAREAAEAALAAERTAEALAGRTDA
DDRATQLDTTLADVMDRLEDTSALLARSREEMLAATDESKALEAALAAMKEQLTTSERQR
VDAEERLDAALLQNTQTAAMYNDNRNDEISRTQARATDLKSRLMDRDSALELAAHLCDCA
KAVQVALLGVLRQPTAASQIAKDLRPF AEALDWIATVEDVADDDRSAL EELLNERLQGAV
RATGLEETLAATEERLRAMDEALVRLSEDKLSEVERADSLYAALTNVQRAVSRTOAQYEQ
QMRERTESATADMIVAQRAAERAEAAKRRAEERLSLVEQELEEQR AELRGLQEANRDLGL
ENNRLSRMLPNSHRGPHEPGSATARRGFGDQVQAEAAQMLHVATMNVELMQRTRRRVRELE
ANDATLRRAYTVLETQLEELRDEVEGENERLRGELDRLREEHADLALRYDALVETTTDATEG
GVAQELKSLRQQRSLLEAE LAEWKARYRDLNLARNLPKVEASQREAAALSEQLRGLTSELD
AAQAQY-----GGDAAQARLAHLEASNREKEARLQQQAAERLRDQSEIEALEDEV
ARKTAALEEAERLAAQYARQVDSLTTNAKALRRSTQVGRATTSPENPPATKRTK

>StriCul

PYQTSRVPTTEGVEGGDLEGPI SVRDAPLTAAEGDELAHKTPAEVAAYVERLRSNYHYFG
EQLRAACERAQH YKSEVARLREELETRRLQKQQFDVDVEAQDGVQHALEDQVRRLRGQLE
EQEGLIYSLRQELVAR-SAQTSTAIMWYEAQRVRAAEEARETQQALRDLAAE-----
--SSKAQVRALVQVQQQQQQQQPSELARRLQDLERENAQLRQSLAAEREH-----
-----EQKQQL EELAAQRAAQVQLQADTDAMALALTRRAPLSLDEYDAQQQQLAALQ
REVEAQTAEEAQRKAAVAPLEATALELTELQRVAAAHDRTTLERNCHQRDEEIVRLQE
LLQCYKTENLAAQEVSRVRTEQYEH TLLEQVDAAEAARAAEEAVDDLQRQLQGERQGLVA
EVKAAELESTLRATNARMEEQEQLIEELEAALQAAADQEEALLLQVQELAERGAELAAEK
AQHLLQVEALKKTLEDTATSYNDRIVVEERLRDRMEEYRQLAAARQLELR TAAAVWLQAT
ESLRDGLRVAEAAAAGPATPPLKLPLEKAAA WYADVALRETAEEAEVGDQLRQLWELGR
AAARREETLRRKLELEQQLRAALDTLAVENVDLKEAGRNYQSLFNETQALQQITNDFQQ

QLLEEKEAASVEVVTARRDMQRAGESRRRAEQRAQLLQEELEQLQHVKGEGLEGQLLAMSE
ERDELTRTRATETVEAVLGS DLSVRASRYSTAMAQEQAVLQVTSLSQSDLLLVRRLREME
AQDAALRRSAAAASEREVAALRT---ERVRLQREL--ER-EHVTTORRYSQLEQTD--G
AAAEELRELRLKVVRELEGGVFRFEERERELARVRSQQEVQAGRRQEA VLTRLRSMSSRMT
DVRRL--RGGLTLATDPQ--LVAEVQQLQ-----HDLENYEKLAVEHHLRLESLOAQL
AAEREARAMADKAVRQLSRHNERLRTTMSERQSTLLALSSKAGRATGSRKMR

>StriGal

PYQTNRNPTSTGEGVSSLEGPI SVRQAPLTAEEGEGLEHKT PAEVKAYVLRRLRSNYHYFG
DQLREACERAQHYKLEVARLREALETQRLQKQRYNLDAAEQDSVOHALEEQVRRRLRGQLE
EQEGLILALRQQLLAR--SSEQTTSIMRYEAERVKASEELRSAQOGLRDLAAD-----
--RSKAQVRALEVQQQ----QQQPSELALRLKELEKENAQLQESLAAERAH-----
-----EVVKQRLEESLAT-----LQADTDTMALALTKRAPI SLEEYDAQORRLAGQO
RQLEELQVEVEQCKAAVAPLEATALELNELRQHVAASHEKAALQREGQRKDEELTRLQE
LLQCYKAETDAAEEANVRTERKEYEALVEQVEAAEAAREAEAAVAELRQQLAQQAQTQT
QMKAAQLEATLRADNERLEEQEGLEIELEAALRTSGEEQORLVQOVQELADRGT ELAAEK
ERLELEAEAFKKTLD D TARSYNDRIVVEDRLRDRVEQYRQLAAARQLELR TAAAVWLQAT
ESVRDGLREVESGAAGRTAPHLPLQPLQEA VAWYVRVALREPPVEQEVGDTLRQLCEAGR
AAARREETLARKLELEQQLRAALDTLAVENLDLKEAGRNYLTLFNETQAALQQITNDFQQ
QLLEEKEAASLEVITARRDMERAEERARRAEERLQLTREELAQLQEANRELEGQLIEIGE
ERDELTRTRVTVANEAMRGGDIHVRASRYSAVAQEQAQMLQVTSLSQSDLLLARRQLRELE
AQESALRRSATLAEREAVLRT---ENTRLQKELVQER-EHTTALKRYSQLLDQTDI--G
AAAEELRQLRLKVVRELEGGVFRFEERERELARVRSQQEVQAGRRQEA VLARLRSMSSRMT
DVRKQLRRGSLTLATDPQ--LVAEVKHLQ-----HDLETYEKLAVEHQ LQLESLOAQL
SQEREARTTADKTVQQLSRHNEQLRQLTGLHERQSAQLALSITPVKKTESRKRNR

>StriOnc

PHQTDNRN--TTPQGVSSLEGPI SVREAPLTAEEGEALDHKT PAEVAAYVQRLRANYHYFG
EQLREACERAQHYKAEVARLREALETQRLQKQFDIDCEAQDGVQHALEEQVRRRLRGQLE
EQEGLLLLRLRQQLAAR--GAEQSTSIMRYEVERVKAEEARWAQQAMRDLAAE-----
--RSKAQVKALEV--QQQEQORPAELARQLAALEKENAQLHQSLAAEREEH-----
-----ETQKQQLLEEALAA-----LQADTDTMALALTKRAPI SLEEYEAQQORLAALQ
RQIEQLATETAAQKAAVAPLEATALEVT ELRQHVAAGHEKAALERAGQRKDEEITRLL E
LLQCYKTEADAAQEAANVRTEREYEAALVEQVEVAEAAARAETALALLRQQL EEAQASAQH
QEKAAQLEATLRADKERLEEQEGLEIAELEEAALRTSGEEQQLVEQVQRLADRAT ALTAER
ERLQMEADALRKTLED TARSYNDRIVVEERVDRAKRQTQLAAARQLELR TAAAVWLQAT
ESIRDSLLAVEGAPDDRPAPALPLEPLREAVAWYARVALREPAIEAEVGDRLRLLLWEAGR
AAARREESLQRQLALEQQLRAALDVLAENLDLKEAGRNYLSLFNETQAAVQQITNDFQQ
QLLEEKEAASLEVITARRDMERAEERARRAEQRLELTLEELEQLQSANRELEGQLVAVGE
ERDVLTRTRVTVEREALL--DVNVRPSRYSAVAQEQAQMLQVTSLSQSDLLLARRQLREME
AQETALRRSLAQARETVVLR--EATRLQEELVQER-EQTVALKRYSQLLDQSDV--G
AAVEELRQLRLKVVRELEGGVFRFEERERELARVRSQQEVQAGRRQEA VLARLRSMSSRMT
DVRRL--GGLSLATDPQ--LVAEVQHLQ-----QDLENYEKLAVEHHLRVESLOAQL
THEREARTTADRAVQQLTRSNEQLRLTLGRERQSLQLGLSVTP-KLAASHKRTR

>Pserp

-----FSPLEPRGLDDLDENELYNVASRLQDEARHAR
AHIARAYEMRDQYMEVARLKL ELHNKYLQTAELEGRGDYEDTIQSLREDIQKLRRELD
NAEGREQAMRLRLAAGVGS DVAESVKYKTEITTKNQIILQMKHEIEMLEKQKEQHAQTP
ESSKSPSRSAATHSEESSGQSDQLLQQAQLEMANAKLVQEELEEQTTIQLEAERNQM
ELKTTSEDRIVQLEKEIEKLRELNTSLQREVDTLAPVLANQAPLTRGAYETFQEEYESLK
IRFSELEAEYHSLEREGRVLAAALETDMRESLTRLRLEHRDTTEQLKLNKNE---KMVE
LIATFEEETKLSRTANDLRKVQVETLHKQIEELNN-----ELKALRNLNTQ
EQRCAELSEKLEDACRRA---KAAELDKAQAKEFNEYRKNTE SRITSLLESVRSSEI
EHLRLSLLTTQNSLEKVTVM LNDHMELDDSRQSHMKVMRESLMNKDNSLR YAVTHLCDVI
NAFSCFTLDVNRVFKSPVFDACLEPVAEALRWATKTHDTLNQVAEKALDVVACVRDVVS
RADRLANSLQEERERSSSLMAAMDQIATDLLDEKDRSDSLFSTIADINQQLEKTKAYFDE
QMRQQAASSVELIAARRAEGRAQEA KRRAEEQLALVEAELEEEERMALRKL RDSYSELKV
EAERLTRLTSSISDANQTDASRHIGLRDARLEVLEQAQMLQISTLQAE LLQTRKMVREAE

SREAVLSQRNRILEQTLTEQHEDVQTATTLRIDLAQLQSEYTDLLERYKNLVESSTAEI
STIEELKQLRRQVRLLETELTEEKAHNRALAKNTPGLEAMLRQQALGETLTGLSTHLA
AVQAQYEGAILRTPNPSTDLLQSRIAHLETVLKEKEKAMEKYASDNLRQQAEIEALEDNL
ASKLVALNRTEQLNEQYLGITIDALARSSKALRTSAAVRLSSSLAQTLP RRKRGR
>Phart1

HPASTK-----PDLPRPISVRDLASTSLAACSLDEMNEGELYTFASRMQDEARHAR
GHIAQAYELRDQYIQEVARLKQELHISKYLQTAELERNRGKYEDTIQSLREDIQKLRRELD
ASEGREQVMRLRLAAGTGS DIPGVKYYQTELTTKNQIQHLKEHIETLEKQNEQRIQTP
ESSKSPSRSAANANSEEPGICWDPNELLRQISQLEATNSKLRQDLEEQSTIQLEAERSRS
ELQASSQTRINNLEEEAEQLRTLNI SLQQEVDMLTITLANQAPLTRGAYETFQDENETLK
IRFSDLEAKYQSLERENLVAAAALETDMRESLTKVRLEHRDTTEQLKLNKNE---KLLE
LVSTFEEETKLSMTANNLRKTQVESFHQIEELTS-----ELNALQSLNEK
KE-----TYRSNAENIIASLKSQISLLSAEL
SDHGFLLPVQSIESIGPSDTILHDSIQFDDSSLVHLTTIKASSVQRNYSLQYALALLRDVI
SAFCCLTSDLNRFVKPSGAFEACLEPVAMASRWAAQMDILKKNRSKASDIVVCLRDLLS
RVDSLVSLSQEEKEHSRSLSAAMDQIAADILDEKERSDSLFGTIADINQOLEKTKADFDE
QIRKQEVASSVELIAARRAEGRALEAKRRAEEQLVLVETELEEERMALRKL RDNYGELKI
EAERLTRLTGSI SDANQTD TNHQIGLRDDQLEVLEQAQILQISTLQAE LLQMRKMVREAE
SREAVLNQKNRVQEQTLAEQHEEVHKTAAALLMDFDQLQSEYTDLLNRYKNLVESSTTEI
STIEELKQLRRQVRLLETELAEEKAHTRAI ELAKNTPDLEAMLRQEALGETLTGLSTHLA
AVQAQYDDTIFRTPNPSIDLLHSRITQLESSIKEKNKKVEKYAADILRQQTVIETLEDNL
TNKLTTLDRTEQLNEQYLSITIDALTRSSKALRASAPVRLVTSSARILTKRRKRGR
>Pem1

HSGSACK-----PDLPRPISVRDLASASLTARGLDDMDEGELYNLASRLQDEARHAR
GHIAARAYELRDQYIEEVARLKQELHISKYLQTTTELERNRRDYLDTIQSLREDIQKLRRELD
VSEEREQAMRLRLAAGAGSDVSESVKYYQTEITTKNKQIQQMKEHIEMLEKQKEQHIQTP
ESSKSPSRSAANAHSEELSTGPDLDQLLQOIAQLEGANAKLVQELEEQSTIQLESERTQS
ELKMASQTRIVELETEVMQLRELNNSLQREVDTLAPALANQAPLTRGAYETFQEEHESLK
MRFSELEIEYHSLECEARVLAATALETDMRESLTKLRLEHRDTTEQLKLNKNE---KLVE
LIATFEEETNLSRTANDLRKAQVESLHKQIEDLTS-----ELRTLNRNHSTH
EQXXXXXXXXXXXXXXXXXX---XXXXXXXXXXXXXXXXXFGEYRSNAESLIASLKSEVSLRLSEI
EDLRLSILASQKSLEQVTSMYNDCIELDDSRQEALKAMRESFEDRHNSLRYAVAYLRDVI
SAFSRLTLDVNCVFKSSVVFDAQLQPVAEASRWAAQTHDTLHKVVQKALDVIACVQDVVS
RADRLAHSLOEEKERSRSLTAAMDQIAADLLDEKERSDSLFTISDINQOLEKTKTYFDE
QMRQOAAAASVELIAARRAEGRAHEAKRRAEEQLVLVEAELEEERMALRKL RDTYGELKV
EAERLTRMTSSISDVNQTDASRHIGLRDARLEVLEQAQMLQISTLQAE LLLTRKMVREAE
SREALSQRNRILEQTLTEQHEDVQTTTTLRLVELAQLQSEYTDLLERYKNLVESSTAEI
SAIEELKQLRRQVRLLETELTEEKAHTRALELAKNTPGLEAMLRQQALGETLTGLSTHLA
AVQAQYEGANL RTPNPSIDLLQSRIAHLETVLKEKDKAMEKYASDNLRQQAEIEALEDNL
ASKLVALNRTEQLNEQYLGITIDALARSSKALRASAAVRLSSSLAQTLP RRKRGR
>Tbru

PGFFTRTWTPPPENIPRLAAPINVRGLATIFAGRRGLADMSEEERMEYTNRLEGDMTHVH
NTLSRAYQLRDDYKNEAARLHRELQDKNHRFDCLLREHSACNDVIYRCKRENEELRQKLD

ESEGEVRQLRDKLVSV-NSQGYVPSGGERHV--GRQEISALEEKNKKLEERERECIRH
HA-----MGKSENTSHEEELAQSRYLLQVTRTEITDLQQLRKEREDYEESLREAI
QARNNLHQONTALQEQKEQLQEMCDEQHRTIEDLTSOLLQRKTEQAVQORGAPDTQMETTD
ENKTD T-----NTNNDDEVYRMLELQOHTLQQQFFLLRREGEAKDILLQKASEEIFNLQ
LQQQLEAALQKSREHAAELTKSLSHTQNLQTAQERITEDSYVINNFHHQLREKIQISGS
ISTTKKMSAERELHVTKLTQLEETVSRLESYGTTP EQTVAAFTTELQHTQQRRLREAE EEEI
IQLTNKLNAAAGVRVRTS-----QSDKDG NARAALVSDVAVRNADTDLG--TQLA
SAL-----VALERLAEEREAALEKATEMEERVSTLEEEELRTAKEKLEERSVE
EISFLKDEVLVSNRLLVDSVSSLN---GKDGAVGADVERLSRVVDELHAQVSATKRGFEE
FYDRRSEGCVTE LIVARRSVDRSNDARRRLEERNVRLEQDLERKCLEVVKLQKECQRLEQ
FVRAKDVRGAHSV LGVDSVDVSSVGAEPVDLEAVDLAQFLQISSLHADLMLCRKTCRQLE
SNQEELLS---LEQNSSQSNAYLEDLDEIRQQLVEMRQOREELIAERRTLTERVDELGR

ERGEEVSRLKQONNSLSAQLQASRNKLSALEASKREGELAAARQQAEEELAKAFSLMEAQVO
TLREEVQKAVVEGDEARIRMSQARVTFLEKALQRKDEEVQRLQDELVQKDEQLDQYEQDA
AKAAQDAENASRKTQLQESAVQKLOGDKKGLEDELRYAKTRQVMDIRSTRKRSR

>Tcongo

PGFFTRTWAPPPDSPRRLAPPINVRGLASGYMCRRGLNDMSHEERLEYTSQLDGMKHVS
SALDRAHQLRDEYKNEAARLHRELMEKQRSLDRVVREYGECKVLRFRFTENEELQOKLD
QSQGEVKHLRSVIQSA-NVQOELKIGAGQLLGDKGNQERIQLEESNRRLERQKEQAYNRR
WKT-----ATTTDQDSLAVRVEQGRLSLNVAMAEIDDLKNELOQRKNDEYSRRLEQLT
NERNSLHQHGLTLSETLEKLEDTCKEQOQTIEDLTVRLLQEQSTNRKERNSHNEQHIGYQ
HVHSEE-----NTDQLRAKLCAIELEKHALQQQLFHLRQEGERKDEI IKKANTDIRKLD
SETAMSRKIGEVWKQNEKLOQERIEETTRQLHISQONQNEEDQYVIKNFHKRLEEQIKLTHN
TQATESLAAERNAHATKLKELEGTIARLEGC GDSSAKVIEALKSELAHTQKRLREAEQDC
ARLT-----RTDAVVKERDAAVGAPSSDLA--TQLS
SAL-----AALNRLAEEREIALEEDAFEGPIVDVLKRDILFLKSSLGECRR
YISGLNDLLSCDDGSSSMVNLFNVRVGGGLPNKCDGTATL-----GATRVMLQ
GSDCVDEVSVADLVAARRSLQRSNDARRRLEERNIELEERELEKRSMDIAGLEKECRDLEK
QVRARG--SSYMRSIHDECDFRSSVSGGVCMETIDLAQFLQISSLQADLMLSRRTCRLQLE
SNQEELLS---VEQGVAQSQNLLESVEELRMEVHNLRRADELLECCERRLLTERVNDLEL
EREEDVEVLRKQNRWLTSQIEEAGDKLSAAEKSRREAELAWEKQAEELIKAFNDLNAQMT
HLRSE-RSAVDSEDETLNQVRNARISFLETTILRKGDGIYRLQEELIKKEDNIDKLQREI
IRMSNKIKNATR KSEKLESNLQQLHEKKEEMNKELYLLKTQEKIPKRESRKRGR

>Tvivax

PGFFTRTWTPPEGVPRLAPPIHVRGLTTIRQCRAVTEMSREEQSNI MRLEEDIROMS
NALHSVYQVRDHYKGEAARLHKELMDRQRSFDRLEQEHKNCRDVIHQQRNTEQELRERLD
ASQTEVRYLKSLVGSS-GAQQKATCANQAHEIEKKREEIILELQRQNKALEVQKEQTLAFQ
QQ-----ASVAAWERGRRELKDTENLIKVMGAEMNDLREQIQKERVQHEESLWAVT
QERNNLQMLKMSLAEVDRLELRCQAQSGTIDSLNETILKRPALSDHDEEAVQNTIANLN
NEGQ-----SLIRDRAVETTSBLEKQTLHSMCALRQDWEDSNSALQRAHIEAQAQFKK
HITLLEDEISTSKQEIANYEQKLI EAQIQLSEAKETSRESESVIRDFKMHEDQNATIYE
LRTISEMAAERDAHAAKLAELS DTVSRLEGVETVPEHAVEALCAELRETQARLREAEQV
ARLTAEANAGATAASGS-----ESASA AVDVRSTAEYVVA--SOLS
AAL-----AALDKLAEEREALDQRVVELSSDLERVQSDKTSYYLSMVDLNN
EVSDLRLQLKNNDVAVTEILSCLGDVSAEGGNSGDSALLKAI EDLYSRLSSAKRNFEL
YSTCNSETGVTELIAARRSADRMTARRHA-----XXXXXXXXMNQECLFLRR
ELRLHGICFSYNSLDVEGSFASRLSAANVNDECSDETQLLQISSLQAEELLCRKRCRRFE
VNQEELTT---LEHHVSQAEVSATNMAELRLDLAQLRLQFTQLQEEHSGLQDRFEDVNH
SQLDELDRRLQNDRLLCDLKEKSEKISAMEKEMRDAERSAKGRVNELLESLCIMEAQMG
DLLKGASSESQQSDRVQGRALAARVSFLEQTSKRKEEIQRLQDELRLKEDQLDQLEQRV
TDSTQELDNAVKVTHMESVVRQLQSAKDSLQEEENRLLRAARTSGLTQARKRPR

>Tcruzi

PGFFTRTWTPPPD-VPRLATPIGVRGLAATHASRHGWEDMPRGERMNYAARLEDDVAQMQ
NALDRAYHLRDHYKEEAARVNKNELMQKQRALDCLERDHRQCGEVICRHEREAVDLRHRLD
CAEGEVQRLKYLQORSE-DPNVRVPLLDQORELRLKQTEVDSLLEEKQLEEHGEHYAARV
HERP-----PNVAHVEEGGEELTRAQHVIQVLC TELQDMRQQLQNERCRNEGTLDEAI
QQRNGLQGVNTSLREELEGLKEMCCSQNRTIDELRDTLLQRAPVTHHEHYALVRRCEKLE
EEIERC-----QTRGGDSVNKAHELEKQAIQQQLLVLRADGEETNTLLKTRADVADMR
YISTLETDELTRKAKTLEEEELAAAKEKLQYATEQNKEAEQVIVNFNRRLQDQFSTIQV
QKAIQDLSAERSAQAEVAELTAAVARLESSADAPERVVEALS AELRETQDRLRQAEESA
SQLAERIDSSEGLRESA-----VATRSAETS LA--VRLS
SAL-----KALEQLAEERETLARGSAELEERVRELEQALRDAENSLCVREG
ELLKLGEECSAKSALLSNALAAVDQVVQEETSFEKQTS SLVKTFEVLHGQFLTKKGFEE
QWERRAEACTAELLAARRSAQRANEARKRAEENAAICMRKVEEVNAELMKAQRDYRALER
LLRSSGRLSISRVENNDEMSFTPTFLR--NTEGAELAQFLQISSLQADLMLSRSTCHRLE
ARQLELQTA---LEQCEWRLASLPPGVEESHAE LLVRGQLENHQLEYHRLEGRYKALKR
QRAEEVEQLQRQNERLVQQLREKRDKIGLSRQMRDSEVAAKQQAEEELAQAFNLL ETQMQ
ALREEVRQGV EETTEAEGKMLRGRITFLEHTLRQKETE MHRLOEELGRNEAQLDQFEDQV

TVATQRLNTARRSTHLEQKVHELKRTNQELREELGAVRA-----PRAATKRPR
>Tcara
SSFFSRSWTPPPDGVPRLGAPIHVRGLV-----DDVLNYTARLEEDVGHVN
NTLDRAYHLRDHYKEEAARLHKELMRKQOMLDSLQCDHQHCGGMILSLRQEIDSLRQLLD
EAEGEVQRLKIALASD-GHQGGLSVMYCKSELVRRETEIENLQVENEKLLKDVKKRV---
-----SDTSDAKSNVDETOQSEITIKALCDELDRDLKSQIENERQEYEGRINSLT
VERNALSAEVTSLNERLHSTRTLCTHADTVDELRESFIQRSSVTQQEYVVLQRRYKEIQ
EELESQRHSDTQNXSNLESTLKLLENQNLQIQI I ILRQSEMEKEELLHRLQKDLKQOS
LSQNLQNELDLLHKTALQKTHLLDTEKKLIQSQKDNEEANRVLTDFFRRRLESDCSTVNL
AARADGIIDEADALRSRLAXNNLILILLEQYANIPEDVIGALAAELRDTQDRLRVAEGER
ATLASNRRVTMGSVERN-----DRGSLVEKDLRYDVSSGCVGDDLSLRSQLA
TAL-----AAIRRLGEDYDTMCEKYRSSEGQVSLKMDLLAERRSLAEVNE
ISVHLREEIDMKSRALAAALVALDRVSAEEVAQEERSANLLSTFEALYAQFEATKRSLEE
QRTRNAEASAAELIILRRNVQRAEEARRRAEDHLAECELELEREHEAHLRLQREHSTLER

MSRSRGDRASEPMSESTAVDESNLVAPVQAEYVDHTQFLQISSLQADLMLTRRRRCRQLE
GIREEMQSE---LERRGTVGGN--DTVSAARAELRRLRGQFEQLRVEHRVLQDQHNTMLR
DHDAALKDLLEQNTRLTHELGENQNALASLAKHNHSQQAPTRHESEALDRAFRVLDAQMH
VLREEVRSASAPRRSAPPD--NARVAALQQAVRRGEEIQRVQEEELERREEEVDAL EEEL
VSATERQNSSKRSPLHYEEIVEQLQANEGPQESLRQVSVTSSSRPTEYRKRSR

>Thei
PGFFTRTWTPPEGIPRLVTPISVRGLADTQRCRSGWDDMTPEERMDYTARLES DMSNVK
NTLDRAYHLRDHYKEEAARLRGEVMQKQALFDRLERDHRKCSEVIRSHQLEVQELRRQLD
LSEGEVYRLKQVLQSK-GENEFFSPMDYERRFSQNETEMARLQEENRSLNQAKSSLTVQ
KTQVDILDESLOKMGQYEEKEEQLAQSOHLIQVLAEMQDLRQQLQKERLDNESSNELI
TQRNSTQSINTSLREEKEALDDLRRQNKIEQELREDLLQRAPVSHQDYHALTKRYEELK
EEIEHL-----KEKHDNSLDIALELEKQTLQQLFNLRHEGEEKERLLQNSLKDASRLRE
LNENFQKELELSRMTKELLEKELSKTKKELQSTEQSNREAENI IEDLNKRVSTQLSTINQ
LT-----EVAETQLRTRDTERRO
SATADREAA-----EAAREHDAALS--EQLS
SAL-----TALERLAEEREEAAEKALQMEACIASLEESLHELRLMLLSESRN
EALSLEDLDVKGKLLSDALSALDRVGADEIEWGERSERFLRTFEELQSQFAVTKKGFEE
QWQRSEACTAELIAARRSAHRATEARKRAEERAALCMQEWEMHGELLRLQREYRVVER
IVRAHGLYSTPHVDDIDESGETPPGVRPWNSEGVDLAQFLQISSLQADLMLSRKKCRLE
SRQEELQAA---LERSELQLASLPASVEESHIELRRVRDQHETLLHEHQRL EEKYEAVKK
QRIEEVEQLRQQSERLMQQLREKRDKLTTLTRHMRESELSAKQQAEEELSDAFNVLDGQMR
LLRSEVQHNEKVADAETKMLQGRIFLQRTLQKQEEIIRRLQEELERKEELLD SFEDQV
TSATHDLENTSRKSAQLEQTVHHLKRNNELEQEEIRILKSHQTESERRGRKRSR

>Tgray
PGFFTRTWSAPPDDIPHLAAPIHVRGLAAVLHNRPGWEEMSAERMDYTARLEGDVSQAK
DALDRAYHLRDHYKEEAARLHRELMRKHGSLERIERDHQLCGDVIYRHQMEAQELRHQLD
EANGEVQRLKALLHTTEHRGGLSVIEYGVGLEQRQAEVNELREENRRLLDQNERSTARE
REH-----STTTRTAGEEELKETRHVLQVLAEMQDLRQMQQERLDHESL EEVI
KQRNSLOGLNTSLREEMGLQTL CRAQNMTNEELQETLLQRASVTHHEYHALNRRCEELQ
KHVELL-----QQEQSPSLDTTLELEKQTLQEQLLTLRREGEEKNEIIRKSNAEIDTMRQ
FTSNLEVELDSRRAKTLIEQDLLVTKEQLLSLQARCQESDNVIAAFHVRLEGQMSTIQALN
LN-----AEXXXRETQDRLREADEL
RAAAELRKA-----
-----AALGRFADEREILNDRAAGLVDRVAALEEELDETSIRASVDSS
TVVRLREELNVKSRLADVLAALDRLTVEEVSYDERTENLLKTFEALQQQFEVTKKTFDE
QWERKSEACTAELIAARRSAQRAIEARKRAEERAAQCARDLEREHSSELLRVQRECRALER
GVRIQSRVLTSHDEVFEDTDVIQSRRRVSNAEGLDLAQFLQISSLQAEMLLARRHIRQLE
ARQEELQLV---LEKTEWELASLPASAEESIVELRWVRAQHKLQVEHRRLOEKFDALKN
QRAEELEEMRQONERLMQQLRERRDKLASVSRQMRESELAAKEQAEELSQAFNLLDTQMR
ALRAEVRHSTKEPEDAQGKLLQGRITFLEHTLKQKEAEIHHLQOELERKEEQLDNLEEQV
GSATRNVESTTRRSALLEKTLEQLRRTNDDLQDELRVVKGRRTPRATSDRKRSR

>Paratryp

GMFFTRA-----RATPWLARPVRDRSGSFSEFRPSALDMLTSDDELRSYASRLEDNVIAS
DNLQQAFSLRDSYKDEAARLRAELDHQRLLRKRYDRDIERYKEISSSQORDIAFLKQKLD
EANGTVRLLQGLDAAH-----SVKYKEEIRLRDASLEHATNTISELQEARVETL---
-----ELELRAAQVELRESVQRNHELEAALHSSQRDAKGDITRLKHEM
DRLRAVDHENTQLRSQLQQVDKHSSAQKQYIEQLLDTLQHRSTLTVEEEARLLQHVDIT
KELQRAHTTLKDKCHNVDLALAMNLI SQTQISLLEAKQQEASADDNAAHRSHDTIHNLM
YVTSLERQI-----VDQAKLLENYENPLQTKDEKEPSSNRHKQNKPKPTTTTCSQH
SDKVERIRDAAAELGAQASSVWLMKAALRNTGTTLTAEAAAIVRSV--TRDSLGVSSDER
LR-----RSGIVTRFSDVHGTGLT
AELRDVL-----QQIEQLAADRAYHEE
CVQSLREEVAAKGAALQRALEALGQASDEEQVAHERSEKILTTLSTLQSEFQRMKERLDP
RPSRSSAVMMDLVEAKRAVSRANEGRRLAERLLDDREGELHLLRQSLAEVTAEKVRIEA
ELRAVNSQSSQAADDPAAASAAVWQQQORAMNVLEQSHVLQLSTLHAEHLTRQEAASLR
QRAAQVREKHSLLTEANTRLEEEVKTNDRVQEELVQ-----MCEVYEAQLQAARLSDA
SSEEALKQLRLKYDRVCSDSL SAARDLARLEREKREAEIDARSQAEEFAKTFRNVEAQLH
ALRREMRPPRASALSPDATRLNDTVRDLQEALASRDEEVSQLOQLLALREKEIAQLEADL
KRAVRALETVSDGEQALRLAARLMRLQAAEVKLADILAALGARRSGVAKKRPR

>Angom

QLLAAQNQSTPSAGR-----PTPAVRPQLDLMTRDDLYNECAKLRWSFDNAQ
ELRERSEQGRIYYRDLVAKYQAEKDLLNSRLQSAIEAKSRAEEETGNYRRSCESLRNQLN
DLEEFRRAKLDLDRAHGAQVV--VGRMKEDLREKDACIMDL SRQNSELEAPADQAAQP
PASLSEYPRSTYDNDP-----LARQRLDELQAAREELDQLRNTLDDERTAAERELQHA-
-----RSESSQMGHELENALRNSDALQETIDQLAAELAKRAPVSLSDWEDYQQKYSDDL
QKNKQSESTVR-----QLERQLENGPTPNEEEELAIN-----NKDQ
LIQTYSEEVELRAKKEEALLQELEAVSAQLQQSTKEQEASEQEKLDEEQLELQKQSSD
DQENDSLMAYSQKLLERSEELVALLEEKTLCDVSVSNYNALRRSLEEECLKSVTLEGEK
KDLLAEVASLTSQLEETAEQYAEKLLDVEYEADRKAARWKVADRRLAQLKELLADYKMSV
ESIVEALRYTLQV--KENSALVLSHVEAALTRADALRTNLNTIADNLENFSQETAAL
EYKESLQTISEMQVRMAQQKAAIDVLAERGYDMDEAWKRIENDI-----RVETEANR
RHQEHRLDLLLFELESANKARERAEFEARREADERATLLES DLQVMR-----
-----LQQGDAASGPSSEEVEQLRSRYAQLAADYETAMNRLQLLERSGVERAG
DLYQQLLETEARCQVIQRELDLYKARRAALAKSNTLAKYEEMRYSDTLHTNLDKVQSLRA
TFGASD-----AKPSGTDPLSKLQDALQRNLQ-----LEQEQIEWKSRYERV
EEKLHAI EENRSLRSLSGVVERLSTSQRAPTLTRSM-----SRSFKSPQSRKRTR

>AngDes

-----PPPAPRPALETMTREDLYNECSKLRWSFDDAL
ERRERSEQAKQYYRDLVGKYQSELERMQDQKGEFLEAKRKAEEETERYKQSCDSLRRQLE
DMDEALYEARAERDRLGGAETV--VESLKGELRERDAAVFDLSRENERLEKERQ-----
--AASPSPTAAPVDPF-----LAQDRLDSEVEELRDTVEALQAKLAETQEQHAAELRQEA
EKVDQRKRELDQNKDDMENLLKRNAELQEVTDNLATQLSKRAPISSEDKQFQNDYDNLA
KKIKSLEEAKE-----KATTGLD-----AVNR--QQQD
ELDSYKKEVELYQQKEEALLKEISNLQAQHQST-----
--EGVEVKKYGEELLEKSEMLATLLEQRSRECNEHLANAVSLKEVLEEECLRNATLDQEL
KEKESALALKEQLEDTVLYEEKLLEVEGATEKCE TWRKTAEKRFQDSYRMAKEVETCV
VGLVEAVGCATDP--TKTSSALS LKNVEGAVKNIRD TLAKMDTAAPAADAFSSKMNDAVS
LCKKHISEIEMQVRLTHLTAAVDVL AERGYDMGEAWKKIEQEI-----QTEAEANR
RNQEHRLDLLLFE L DSANKAREKMDEARREADERATLLESEVQVLR-----
-----LQQSDAR--SVSPGELDRLRARNALVADYEIAVDRLRVLERAGGRHGE
DLYQQLVEAEARCQVLQRELDLYKERRSALAKSNTLAKYDGLRYSHSLQSNLHKMQDLRA
AFDPSA--PHKSTSKDPTEIAIQRNEALE-----QERIELOTRY SRA
EEKLSVLEENRRLAQLAGVAERLSTSRRTLLPKDMMP--SPSTQQGKAKKRPR