

## S9 Appendix. Flatworms talin alignment.

Species	Accession number (gene)	Protein name	Database
<i>Clonorchis sinensis</i>	CLF_105939	talin	UniProt
<i>Echinococcus granulosus</i>	EgrG_000736000	talin	UniProt
<i>Echinococcus multilocularis</i>	EmuJ_000736000	talin	UniProt
<i>Hymenolepis microstoma</i>	HmN_000220000	apolipoprotein A I binding protein (similar to talin)	UniProt
<i>Schistosoma japonicum</i>	Sjp_0006570	talin	GeneDB
<i>Schistosoma mansoni</i>	Smp_037860	putative talin	UniProt
<i>Taenia solium</i>	TsM_000902500	talin	GeneDB

### *C. sinensis* vs *S. japonicum*

Score	Expect	Method	Identities	Positives	Gaps		
2767 bits (7172)	0.0	Compositional matrix adjust.	1513/2497 (61%)	1848/2497 (74%)	120/2497 (4%)		
Query	20	NLGGVGATIVQHQRLP	PPRRVGGPKDGFVNGALEETI-QPKMDDSIQMATGY-MQTNG-WS-HQ--NLHTD-RTDFVDGQ----	WAERPMSPDNFMGGSSHTFHYQEITPAQRALLVTIIE	128		
Sbjct	436	NVQAERATIVQHRVPVRGR--GKDQLVNGVVEKPFDPQYHKTDS----	SGHQIHTNGVWSIHQQNNLIGDENSTVIDGQSMMLTTTTRQMSPDNFMFGGTHIFHYQEMSPAQKFLLV	549			
Query	129	DVETLQAAKEQLYLSPEDEDR-IIGIGSDEASKRWLSESIGASQAKVTDEVGAMNAAVAQALRSANRAEAYDSTGMSQVQV--LGVSDP	GGDDLMMMQSFRVVTIHFFPAFVDDVVKRVAVLR	245			
Sbjct	550	D+ETLQAAK+ L + P +R ++ +G+DEASKRWLSES+GASQAKVTDEVGAMNAAVAQALRSANRAE+ D + Q G L S+DPGDDLMMMQSFRVVTIHFFPAFVDD+KRVAVLR	DLETLQAAKDNLDIGPPTERHLLNVTGDEASKRWLSESMGASQAKVTDEVGAMNAAVAQALRSANRAESAD-INVVQHGGSLDQSRDPGGDDLMMMQSFRVVTIHFFPAFVDDLKRVAVLR	668			
Query	246	REVGAIRADQGTGEPVIAAQSEESSQNLLSAARHVADAFTDLLESARPLATGQPREEV--VTDESTIRSQAVTT-----GPVSSSRKAIIDAASRVGEASNDLLRHVMHEGDGEDFEA	RE G ++ADQGTG+P+IA ++EES+QNLLSAARHVADAFTDLL+SARPLA GQ EE+ +D+ ++ VTT GPVSSSRKAI+AAASRVG ASNDLLRHVMH GD ++ F	357			
Sbjct	669	RESGVLKADQGTGQPIIAEKTEESTQNLLSAARHVADAFTDLLQ SARPLAMGQTHEEIIYPTSDDGNYQANRVTTTQQAQQGPVSSSRKAIIEAASRVGAASNDLLRHVMHGGDDDLFNT	788				
Query	358	SSLMLTEEERLYKDQLLSLAKAVANTASLVKAKNLATQTSLDPEAQQHVIAAATQTGLCTSQLVACTKVLAPTIYQPSQQQLSEAAREVSWAVDGVVQASRAAGARTADQPVVQQP	+ L+LT+EEER+Y+DQLL LAKAVAN TA LVVKAK LA+Q + DPEAQQ VIA+ TQTGLCTSQLVACTKVLAPTI+QP+CCQQQLSEA REV+WAVDGVVQASRAA +D P VQQ	477			
Sbjct	789	NYLILTDEERIIYQDQLLGLAKAVANATAGLVVKAKVLASQITSDPEAQQRVIASVTQTGLCTSQLVACTKVLAPTIHQPTCQQQLSEAGREVTWAVDGVVQASRAAIHVPSPDPQMVMQ-	907				
Query	478	VRTAVAETEMAATEVRDALDQLNAHLSKASAKTSSGDALDNFQLAYDQLQQYHHTDQGRMVASARRMAQATAQMIADIKAQAEALAADDPDRQSRLFAAAKQLADATTILIASAKVCSSTNP	++ ETE AATEVRDALDQLNAHL K S K DAL+NFQ AY+ L+QY+T DQGRMVASARR+AQATA MIADIKAQAE + DD DRQ+RLF AAKQLADATT LI +AK CS+NP	597			
Sbjct	908	---SIETETAATEVRDALDQLNAHLLKGSYKGYQSDALNNTQAYEDLKYTYTDDQGRMVASARRLAQATALMIADIKAQAEESGDDSDRQNRFLAAKQLADATTNLINAACKSCSSNP	1024				
Query	598	ENPVTQEELRQAAESNLIVHSAAADLLHRRLIRNLQAAARAAVTGATQLVNVSVQAAKRSRGNTYQVISDGLKLVNGLIPRTVISIRESRACPDPLTQVELISACERFVLPESLVRSS	+N Q ELR AA+ LN + +SA + RLI NLQ+AARA VTGATQLVN+SQ A K SRGN+YQ+ +D K V LIP+TVI+IRESRA PDDP+T +ELI+ACERFV PC ++VRS	717			
Sbjct	1025	DNIRLQTELRDAADELNTLAYSAGGEQFQTRLIHNLQSAARATVGTGATQLVNISQAAIKYSRGNSYQLQNDTKNVTNLIKPTVITIRESRANPDDPMTLLELINACERFVQFCTAMVRSS	1144				
Query	718	RSVAPTSDPTTQAALDNSTSQLSGAIEVLRTCLARLTPLTRMQVVDGALARLARLAVEAAAIEGLKQGTPLAPLPEKIDDCFNLLGTSVQDANVSTHELHSIVETMAQQQRDMMSRSD	+S+APT+SDP+ QAAAL+NS SQL+ A+E L CL + PL R Q+DGAL RLARL+ E AA+ET L QGTLA +P++KI+DC +LL SV+D S L SI+ TMA+ +++ +	837			
Sbjct	1145	KSLAPTSLSDPSIQAALENSNSQLASALETLIACLIKAAPLARHFQLDGALTRLARLSDEVAAMETSIIQGTSLASIQEKIEDCCHLLNVSVRDGTESAKNLDIIISTMAKMDSTSLNQL	1264				

Query 838 VCGVPASMLATAIAGLVHATRGI VAHDAYDAADAGRSGAVGPSSDSTR LAVTQDTRQAVQLAYELVKAARDARFAYDDGQMEALATV VSHSEQLTVQLLDTLGRCLRGLPGHREINEATN 957  
 + G A LA+A+ GLVHATRGI+A + G SG +++STR AV +DTRQAVQLAYELV+AA+D R A D Q ATV SHS+QLT QL++TL RCLR LPHGREI+EAT+  
 Sbjct 1265 LAGQSADHLASAMTGLVHATRGI IADQNSLSYIGSSGINTTAAESTRRRAVIRDTRQAVQLAYELVRAAKDTRIACDAKQPANAATVQSHSQQLTAQLIETLTRCLRSLPHGREISEATH 1384

Query 958 LIEKRREDLVQLSHEAPTR-RVRWVEPSVYERAQSELTRA AVEFNQATGDLSSSYSPGAFRRTRRFTGAYDTLTVKGV D LCM-SGVPPNEVPLPEPPPINRDLVDGLVNVSNHSYALMT 1075  
 LIE+RR L+Q S T R+V+P Y+R Q++LT+AAVEFNQATGDLSSSYSPG FRRTTRRFTGAYD LT KG++L M S N P +N DL++GL+NV S+HSY LM  
 Sbjct 1385 LIEQRRLQLIQSYQPNTSLSARFVQPEYQRKQTDLTQA AVEFNQATGDLSSSYSPGTFRRTRRFTGAYDHLTDKGIELSMCSEFSSNNQ--QSGPAVNSDLLNGLINVSDHSYVLMN 1502

Query 1076 EAGRVCARPDEQEIRDRFQTAARDVTE S ISQLLTICTSGVTAEQRDCEVALRRLEALRPLENPNRPNVQKTYCECVDEVAR SITPLADSLR TMSNAAKEKHTQEFGA A VRQCSNSMCQ L 1195  
 EA RVC +P+ + +R +FQ+AARDVTE S ISQLLT+CTSGV+ +QR+CE A+RRL++LRPLL NPNRP+NQ +Y +CVD VA+ + PLA+SLR+MS +A+E+ EF +AVRQCS+S+CQL  
 Sbjct 1503 EASRVCQPEGENLRGKFQSAARDVTE S ISQLLTCTSGVSPDQRECEAAIRRLDSLRPLL NPNRPLNQHSYD CVDTVAKCLEPLAESLRSMSTSAREQKPNFSSAVRQCSDSVCQL 1622

Query 1196 IEETAQAAYLIGLADP-----RSEPGRPSLVDPQLFMRSQREIQVCDAICDPSITNRQIITLSTEMARSAKTLCEACSAVSAQTNNP DARHQLNALTRETMQSITALI 1299  
 +EETAQ AYLIGLAD S PG S++D LF++SQ+EIQ CDAIC+P++ NRQ+I+LSTEMARSAK LCEACS +S+ T N DAR +LN LT++TMQSIT+LI  
 Sbjct 1623 VEETAQTAYLIGLADR FENRRVSGATSMSSNISPPGGSSMIDLDFIQSKEIQSACDAICEPNVNNRQVISLSTEMARSAKILCEACSQLSHTTNIDARQLNLLTKQTMQSITSLI 1742

Query 1300 QRRGNVSTAAAGSQLAEVNGEWEANRQVTLANARGVSSNVARLVHMVTQGPQF----IGQAARVTD EAREAQPVCSAGLSSLNAAQAVLRAAQTLITNARLQGPQAFLSFSTASKEL 1415  
 Q RG ST A +VNGE+ E+NRQ TL NA+ VS+NVA LV ++T PQF I Q RVT +E Q+PVC AGL SLNA QAVLR AQ L++ +R+G E +F++FS+ASKEL  
 Sbjct 1743 QHRG--STVA-----TDVNGEFSESNRQQTLLINAQTVSANVANLVKLITTEPQFTDLTISQ-ERVSTIKETQRPVCVAGLES LNATQAVLR TAQNL LSTSRMGPSSENSFMAFSSASKEL 1854

Query 1416 SESIKALAAAMREHAPGQLEQCQ RVLQNINGLLQELQRAKMASMEGRQLPRELNEEGFQKQLATCCRALLDAPNVGRAARSEAEQLGHSVRVDSYLPGITSSAIAAASRSPLSSTQLV 1535  
 S S Q ECQ V NI LL ELQ A++ASME RL PRRELNEEGFQ+QLAT RALLD+AP+VGRAA SEAEQLG ++ DSYLPG+ S +IAAASRSP+ STQLV  
 Sbjct 1855 SNS-----QRECQHVFNKIKLLHELDARIASMENRLLPRELNEEGFQRLATSVRALLDSAPS VGRAASSEAEQLGRTILT WDSYLPGLVSGSIAAASRSPMPSTQLV 1960

Query 1536 YLEHASTILEAADQLVRVARDAGGNPRATHLPHLDEAVRGLMESCEDLLTALDDVASRQGHVSTLIDTINRSLAQTEEIIQVPVDARFADYQARLLRIARHMEQLTQAIQLRARQPPSD 1655  
 YL+HA T+LEAA QL+ VA+D+GGN PRA HL L+E+VRGL++SCE+ LT LDDVASRQGHVSTLIDT+NR+ AQ EEIIQVP+DARF+DYQ R+LRIARHM+QL QAIQLRARQP SD  
 Sbjct 1961 YLDHACTVLEAAGQLITVAKDSGGNPRAIHLQDSLNE SVRGLVDSCEFLTILDDVASRQGHVSTLIDTLNRAQAQAE EIIQVPMDARFSDYQTRMLRIARHMDQLNQA IQLRARQFASD 2080

Query 1656 GELAPLAHTLAQEYQEMCQTCKGAAATLPDARQADELRGAVRAVGMATAELVQATTAARIRAYDS-HGDQQMLLSRNQGIAYLDTRAEALDSKLRDLIALLDTQGPNTQACLQSASTVSG 1774  
 GE+APLAH L QEYQEMCQTCKGAAATLPD QAD+LR AVR VG A L+QATT +R Y S G Q +LLSRN+ IA LD RAE LDSKLR+LI LL+T GP TQACLQSASTVSG  
 Sbjct 2081 GEIAPLAHNL TQEYQEMCQTCKGAAATLPD VQADQLRS AVRGVGTACISLIQATTSRSVDYVSTGGQNL LLSRNK NIAELDRRAEFLDSKLRNLINLLETHGPGTQACLQSASTVSG 2200

Query 1775 I IADLDTTILFASSGTLHGPVHDDVDALLGPDYRPPYALPGAGEAIADRG-----SRFGNREYPE---EGFGTVRDSIVRTARALVDDTQSLVSGTGEDQTRLATTAHVAVERVT 1882  
 II+DLDTTILFASSGTLH PV DD DAL +++ + +A G SR G+R PE E F +R+SI RTARALVDDTQSLVSGTGEDQ RLA+TAH AVER+T  
 Sbjct 2201 IISDLDTTILFASSGTLHPPVRDDFDALWADNEFGHSRLTDRSDHYLALSGNTDGTGIPSRAGSR--PEMHN AENFTPLRESIARTARALVDDTQSLVSGTGEDQARLASTAHTAVERIT 2318

Query 1883 QLADVVKRGA AVIGPGQPD TQVEVLSACRDVATGLRDVLLAASQTPGRDASDPVHEQVRNNVQLTSLNIGALLQVKTIAD DENRGIQALVSAAKYCRDQAAQLAPRPEPESLLTPVSSR 2002  
 QLADVVK+GA AVIGPGQPD TQVE+L++CRD++ LR+V L++S+ GR + DPV+++VR+N QL +SNIG LLQ +++I +DE RGI++L AA YCR+QA L +PE +LTP SSR  
 Sbjct 2319 QLADVVKQGA AVIGPGQPD TQVELLTSRDISNALRN VFLSSSKAHGRPSLDPVYDEVRSNAQLVISNIGQLLQTLRSIEEDERRGIRSL ELAATYCREQA KHLPPSSRDPEGILT PPSR 2438

Query 2003 RSAPSVANVSAVSGSSSLIARYLAPDDLARAASGPVQTAVSKAILAGNTQTQRDVLSTAATTRDVVTDLVSAATALLRCPEAVTETRTACVASTKGLAE EFAGLLDALKAVNSKPADHER 2122  
 +S + S SGSSSLIARYLAPDDLARAASGPVQ+AVSKAILA NTQTQRD++ +TTRD VTDLV+AA LLR + +ETR++C T+ E GL K S P E  
 Sbjct 2439 KSG--AYSTSVA SSSSLIARYLAPDDLARAASGPVQSAVSKAILANN TQTQRD VQATSTTRDAVTDLVAAARCLLRYSDIPSETRSSCAIVTR---EXXGLCTPRK---SGP---EN 2546

Query 2123 IGNAARRIADLSHTLLSQLDNLRES PRLV MYFRASSPEWRDVAEKFIGRHVAYHHHYVPTYTI---VGPGSTNRGLYVRPICYPFPRQKAIPEQDEDRSGDEAEFRIATAIDQLK-SILD 2238  
 + N ARRIAD+SH+L++ +DNLRE PRL+MYFRASSPEWRDVAEKFIGRHVAYHHHYV Y + G GS+++ +RRP CYP Q + + E DEAE RI AI+QL S+  
 Sbjct 2547 VSNARRIADISHSLINLV DNLREGPRLIMYFRASSPEWRDVAEKFIGRHVAYHHHYVSNYALGGSGSGSSHKSFIRPTCYPL--QVSHDQGGEQIINDEAE CRIVDAINQLNLSLEK 2664

Query 2239 C--EPPSDPTILQSAHSVAVATQSLIHTARSIGLAPPSDAVDRARRAGT----SHHGASLWRTELTV ALATQDMCQLAQLCSAVLQEENAH PDSKDLFGTSTRLVPSKERLLAAVRRVAA 2352  
 C + ++L SA SVAV+ Q+LI TAR++ PP+DAV+RAR A + S A LWRTEL V+LATQ +CQL++LCS + + + S L+ S+ERLLAAVR+VA

Sbjct 2665 CLVDNKLAQSLVARSVAVSQNLIRTARTVVSVPND AVERARTAESASQKSTSSAVLWRTELNVSLATQHLCQLSKLCSETVPSSTSTKLENNTGELSGGLILSRERLLAAVRQVAT 2784

Query 2353 ASAQLLSAKSRKSSADIQKLVAGQVKEATDRLSGVLHQAN--FTCDSTGLVITPMPATLQNVIEETHSKIHSQSELDVLAQRLAQLQDNQ 2447  
 A AQLL++AK R+ S+DI+KL+ AGQ VKE+TD L+ + +++ T T + T +L I T ++I S Q EL+ L+ LA+LQ Q

Sbjct 2785 AGAQLLLAAKLRKEALMSSDIRKLAAGQAVKESTDLAEAVQKSDAISTTQRTEITWPTYTNQSLIQGIHTKARIQSQQIELEALESHLARLQQTQ 2881

**C. sinensis vs S. mansoni**

Score Expect Method Identities Positives Gaps  
 491 bits(1265) 4e-158 Compositional matrix adjust. 286/557(51%) 380/557(68%) 20/557(3%)

Query 1903 QVEVLSACRDVATGLRDVLLAASQTPGRDASDPVHEQVRNNVQLTSLNIGALLQKVKTIAADDENRGIQALVSAKYCRDQAAQLAPRPEPESELLTPVSSRRSAPSVANVSAVSGSSSLIA 2022  
 +VE+L++CRDVAT LR V L++ + GR + DP++++VR+N Q+ +SN+G LLQ +K++ +DE RGI++L AA YCR+QA L +PE LLTP SSR+S P S SGSSSLIA

Sbjct 5 KVELLTSACRDVATALRSVFLSSGKVQGRPSLDPIYDEVRSNAQVVISNVGQLLQTLKSVVEEDERRGIRSLELAANYCREQAKLLPSSKDPPEGLTPNSRKSGRPRST--SVASGSSSLIA 122

Query 2023 RYLAPDDLARAASGPVQTAVSKAILAGNTQTQRDVLSTAATRDRVVDLVAATAALLRCPAEVTETRTACVASTKGLAEFAGLLDALKAVNSKP--ADHERIGNAARRIADLSHTLLSQ 2140  
 RYLAPDDLARAA GPVQ+AVSKAILA NTQ RD++ T++ TRD V DLV+AA LLR + ETR++C T+ LAEEFA LLDALK P + E + N ARRIAD+SH+L++

Sbjct 123 RYLAPDDLARAAGGPVQSAVSKAILANNTQIHRDIVQTSSATRDAVADLVAAAKCLLRYSDIPPESTRSSCAIVTRELAEEFAALLDALKNGLYIPRKSHPENVSNIARRIADISHSLINL 242

Query 2141 LDNLRESRLVYFRASSPEWRDVAEKFIGRHVAYHHHYVPTYTIVVPGSTNRGLYVRPICYPFRPQKAIPEQDEDRSGDEAEFRIATAIDQLKSILD-CEPPSD--PTILQSAHSVAVA 2257  
 +D+LRE PRL+MYFRASSPEWRDVAEKFIGRHVAYHHHYV Y + G +++ ++RP CYP Q + + E DEAE RI AI+QL S L+ C + ++L +A SVAV+

Sbjct 243 VDLSREGPRLIMYFRASSPEWRDVAEKFIGRHVAYHHHYVSNYALGGSSKSHKSI FIRPTCYPL--QISSEQTTEQTMNDEAECRIVDAINQLNSLEKCSADNKLAQSLLV TARSAVAVS 360

Query 2258 TQSLIHTARSIGLAPPSDAVDRARRA-GTSHH---GASLWRTELTVLATQDMCQLAQLCSAVLQEEENAHPSKDLFGTSTRLVPSKERLLAAVRRVAASAQLLSAKSRKSSADI 2373  
 Q+LI TAR++ PPS+AVDRAR A G S A+LWRTEL V+LATQ +CQL++LCS + + + S L+ S+ERLLAAVR+VA A AQLL++AK R+ SADI

Sbjct 361 VQNLIRTARTVISVPPSNAVDRARAEAGASQKSTSSAALWRTELNVSLATQHLCQLSKLCSETVPSITSQKLENNAGELSGGLILSRERLLAAVRQVATAGAQLLLAAKLRKEALMSADI 480

Query 2374 QKLVAGQVKEATDRLSGVLHQANFTCDSTG---LYVITPM---PATLQNVIEETHSKIHSQSELDVLAQRLAQLQ 2444  
 +KL+ AGQ VKE+TD L+ + + + +S+ + + P A LQ I T ++I S Q EL+ L+ RLA L+

Sbjct 481 RKLKTAGQAVKESTDLAETVQRGDIVSNSSTSPQIEITWPTYTNQALLQG-IHTKARIQSQQMELEALESRLASLK 556

**C. sinensis vs E. granulosus**

Score Expect Method Identities Positives Gaps  
 1734 bits(4491) 0.0 Compositional matrix adjust. 1063/2421(44%) 1503/2421(62%) 185/2421(7%)

Query 100 DNFMMGGSHTFHYQEITPAQRALLVTIIEDVETLQAAKEQLYLSPEDEDRI-IGIGSDEASKRWLSESIGASQAKVTDEVGAMNAVAQALRSANRAEAYDSTGMSQVGLGVSQ--DPGDD 216  
 D+F+ GG+ + +TPAQRAL VTI E++ L+ K++L + P +R+ + +G DEAS+RWL ES+G ++ +V DE+GAMNAVAQA+RSANRA DS G V GV++ DP DD

Sbjct 546 DDFITGGTQYYVCKTMTPAQRALHVTIEENINALEEGKQRLDIGPPAERVLVNLGHDEASRRWLYESMGVTEVVRVADEMGAMNAVAQAVRSANRA---DSAG-GGVHFGVTEVTDPSDD 661

Query 217 LMMMQQSFRVVTIHFPFVDDVVKRVAVLRRREVGAIRA-DQTGEPV---IAAQSEESSQNLLSAAHVAFAFTDLLESARPLATGQPREVVTDSTIRSQAVTT--GPVSSSRKAIIDA 330  
 LM +QQSFRVVT+HFPFVDDVVKRVAVLRRRE G +RA D+ E + I QSE ++QNLL AAR VAD+FT LL++A PL T Q TD + VT V+ SSRKAI++A

Sbjct 662 LMRVQQSFRVVTVHFPFVDDVVKRVAVLRRREGLLRAEDEPDERLRDITDQSEANAQNLLGAARTVADSFTALLKAATPLTTRQ-----TDYNLDTVNVYVTEQGSQVAGSSRKAILEA 775

Query 331 ASRVGEASNDLLRHVMH-----EGD-----GEEDFEASSLMLTEERLYKQLLSLAKAVANTTASLVVKAKNLATQTSLDPEAQQHVIAAATQTGLCTSQLVACTKVLAPTIYQPSQQ 440  
 A+RVGEASND LR+V+ EG+ E ++A L L Y+D+LL LAK VAN+TA LVVKAKNLATQTSLDPE+ QQ V+ AAT TGLCTSQLVACTKVLAPTI+Q SCQQ

Sbjct 776 ANRVGEASNDFLRYVLQDTEVLEGEIVPEVDERAYQADLLAL-----YQDKLLELAKEVANSTAVLVVKAKNLATQTSLDPEHQQLVVMMAATNTGLCTSQLVACTKVLAPTIHQASCQQ 889

Query 441 QLSEAAAREVSWAVDGVVQASRAAGARTADQPVQVQQP---VRTAVAETEMAATEVRDALDQLNAHLSKASAKTSSGDALDNFQLAYDQLQYHHDGQRMVASARRMAQATAQMIADIK 556  
 QLSE+ EV+ AVD VV+ +R AG + A + + +P + +AV + E AATEVR +LD+LNAHL + S + GD LD FQ AYD+L+Q DG R+VA+ARR++Q TAQMI+D+K

Sbjct 890 QLSESTCEVATAVDDVVRVARGAG-QAAHDTINLSEPEVHEIDSVMQVESAAATEVRTSLDRLNAHLRGSIRPYQGDITDLDFQQAYDELQ--ETDGVRLVAAARRLSQVTAQMSIDLK 1006

Query 557 AQAELAADDPDRSRLFAAAKQLADATTILIASAKVCSTNPENPVTEELRQAAESLNLIHVHSAAADLLHRRIRLNQAAARAAVTGATQLVNVSVQAAKRSRGNITYQVISDGKLVNGLI 676

Sbjct 1007 QAE DP+RQ+RLFAAAKQLADATT LI AK+CS+NP++ Q +L+ AA++L ++ +++AA+LL+ R+IR+LQAAARA V+ + L+ S+ AA +SR N Y ++ + K V+ L+ VQAETVEGDPERQTRLFAAAKQLADATTELIQAKICSSNPDSLTPQADLKHAAADNLVVVAYASAAELLNARVIRSLQAAARAVVSASNNLITTSRYAANKSRANNYHILINEKAVSDLL 1126

Query 677 PRTVISIRESRACPDPLTQVELISACERFVLPCESLVRSSRSVAPTVSDPTQAALDNSTSQLSGAIEVLRCLARLTPLTRQMVDGALARLARLAVEAAAIEGLKQGTLPAPLDEK 796  
P+ V +IR R P+DPL Q+ELI ACE + PC L RS RS+ PT+ DP Q+ALDNSTSQ+ A+E L+ CLAR++P+ RQ+Q+DGALA L R EA +E K GTL LPD+K

Sbjct 1127 PKMVS AIRFIRRDPEPLAQLELICACEHAIQPCGQLARSCRSMVPTIGDPAVQSALDNSTSQMVVAVETLKACLARVSP IARQLQMDGALASLLRTVREAELEDL SKAGTLTALPDDK 1246

Query 797 IDDCFNLLGTSVQDANVSTHELHSIVETMAQ-----QORDMMSRSDVCGVPASMLATAIAGLVHATRGIVAHDAYDAADAGRSGAVGPPSSDSTR LAVTQDTRQAVQLAYELVKAARDARF 911  
++D + +L S++DA +T+E+ S V+++ Q Q + D G A+ LA A++ L+ +TR + + +G S A R A QLAY+LV AR

Sbjct 1247 VEDRYRILKVSIRDAQNATNEVTSAVDSLEQDYAAPQSPNCDQGEDWLGSANRLAGAMS DLLRSTRSCLVSEDRPIDSSGASAA-----RTAAQLAYQLVCEAR---- 1346

Query 912 AYDDGQMEALATVVSHSEQ-----LTVQLL-DTLGRC-----LRGLPGHREINEATNLEKRREDLVQLSHEAPTRRVRVWVPSVYERAQSELTRA AVEFNQATGDL LSSYSP 1013  
++EAL VS +Q + +L+ D L R L+ LPG +E++EA++L+ +RR+DLV+ + + PT +E + ER Q+E AAVEFNQAT DL S Y+P

Sbjct 1347 -----KLEALQPTVSSPDQADVLESNRNVAINCKLIGDNLNRAFELCKLALPGQKELSEASDLVGRRRKDLVRFA-DNPTVFEYPIEETRVERTQTEFAGAAVEFNQATADLTSCYTP 1460

Query 1014 GAFRRTRRRFTGAYDTLTVKGV D L C M S G V P P N E V P L P E P P P I N R D L V D G L V N V S N H S Y A L M T E A G R V C A R P D E Q E I R D R F Q T A A R D V T E S I S Q L L T I C T S G V T A E Q R D C E V A L R R R L E A L R 1133  
FRR++ RF GAYD+L KG+ L + P + P L+ GLV VS+ S ++ +A +VC++P+ + +R + AAR VT+SIS LL++ SGV DC+ LRRLEA+R

Sbjct 1461 SNFRSSVRFAGAYDSLVDKGLQLSRAKSPKD-----PTGPQLITGLVEVSDRSSEMLEDQVCSQPEAEPLRQKHLAARSVTDSISHLLSVSVSGVVPVADCD AELRRLLEAMR 1572

Query 1134 PLEENPNRPVNQKTYCECVDEVARSI TPLADSLRTMSNAAKEKHTQEFGA A V R Q C S N S M C Q L I E E T A Q A A Y L I G L A D R P S E P G R P S L V D P Q L F M R S Q R E I Q Q V C D A I C D P S I T N R Q I I T L 1253  
P L L E P + R P N Q T Y C V A + S + P L + D + R + K + T + F + + R + + + C Q + + E E T + Q A A Y L I G + A P S E P G R V D F R Q R + I + + C A + P + T R Q + + L

Sbjct 1573 P L L E Y P D R P T N Q H T Y Q L C V S S F A Q S L A P L S D G I R G T L D G I K S H N T E V F T S N I R K V T D C L C Q V V E E T S Q A A Y L I G V A H P T S E P G R G G F V D I A F F E R L Q R D I K N I C R A M K S P E V T E R Q V V N L 1692

Query 1254 STEMARS AKTICEACSAVSAQTNNP DARHQLNAL TRETMQSITALIQRRGNVSTAAAGSQLAEVNGEWDEANRQVTLANARGVSSNVARLVHMVTQGPQFIGQAARVTDEAREAQQPVCS 1373  
+ +M+ + ++L +AC+ ++AQ NN +AR +L L ++++S++ LIQR + +W + R++T N V+ N+ RLV +T +F G+ AR++ EAR AQQP+C

Sbjct 1693 ANDMSANVRSLRDACNTIAAQANNQEARRRLQNLADDSVKMSSELIQR-----SADWSDEGRRI TANNIHAVNGNIDRLVSFITSISEFGGEPARISSEARSAQQPICL 1796

Query 1374 AGLSSLNAAQAVLRAAQT LITNARL-GQPEQ-AFLSFSTASKELSES IKALAAAMREHAPGQLECRVQLQNINGLLQELQRAKMASMEGR LQPRRELNEEGFQKQLATCCRALLD AAPNV 1491  
AG + L + + V+ A++ ++ A G+P AF +F+ AS++L+++ K L A + E P G Q E C Q R L N I L L E L + R K M A M + G P R + N E E G F K Q L R A + D A V

Sbjct 1797 AGQACLESGRGVILASKHMLQTAETSGEPSNTAFGAFTAASRDLT DNTKHL LAVLNEQGPGQAECQRALLNIRRLHLELERDKMAMMDGVFTPRHDSNEEGFLKQLTISARAVRDMATPV 1916

Query 1492 GRAARSEAEQLGHSVRVVD SYLPGITSSAIAAASRSP LSSTQLVYLEHASTILEAADQLVRVARDAGGNPRATHLHPHLDEAVRGLMESCEDLLTALDDVASRQGHVSTLIDTINRSLAQ 1611  
GR A SEAE+LGH+VR +D LPG+ SSA+ AASR+P SS QL+Y+EH T+LE+ +QL+ A+ AGGNPR+ + + + V+ ++SCE+LL+A++ VAS+Q ++ L+ + S A

Sbjct 1917 GRGATSEAERLGHAVRELDLLL PGLVSSALGAASRAPNSSAQLIYVEHTRTVLESVEQLISA AKQAGGNPRSDPIREAVGDGVKAQIDSCEELLSAIEGVASQQSFMAKLMAVLEESRAL 2036

Query 1612 TEEIIQVPVDARFADYQARLLRIARHMEQLTQAIQLRARQPPS--DGELAPLAHTLAQEYQEMCQTCKGAAATLPDARQADELRGAVRVMGATAELVQATTAARIRAYD SHGQQMLLS 1729  
++ + P DA F DYQ+RLLR+ARHMEQ T+A AR+P E+ L LAQ+Y E+C + A TL D +A++LR AV+ +G T L+ AT + S D +

Sbjct 2037 VDKPGRAPYDAHFTDYQSRLRLARHMEQQTEATVRAARRPQESISQEMTGLVQNLAQDYAEICMISR DACETLRDPHEAEKLRWAVKNLGTTRGLIMAT----VNVASSRPDTSL--- 2149

Query 1730 RNQGI-----AY---LDTRAEALDSKLRDLIALLD TQGPNTQACLQSASTVSGI IADLDTTILFASSGTLHGPVHDDVDALLGPDYRPPYALPGAGEAIADRGSRFGNREYPEEGFGT 1840  
N+G AY LD A A++ +LR+LIALL+ QGP TQACLQ ASTVSGI IADLDTTILFASSGTLH P+ DD D +R Y+ A++ FG RE + F +

Sbjct 2150 -NRGSVGAKPPAYLRSLDFSAGAMNDR LRELIALLEMQPGPTQACLQGASTVSGI IADLDTTILFASSGTLHPPIRDDADVHQLAQAHR--YS-----ALSSGYDDFG-RE-AQGNFIS 2258

Query 1841 VRDSIVRTARALVDDTQSLVSGTGEDQTRLATTAHVAVERVTQLADVVKRGAAVIGPGQPDTQVEVLSACRDVATGLRDVLLAASQTPGRDASDPVHEQVRNNVQLT LSNIGALLQKVKT 1960  
+R+SI RTA ALV DT+SLV GEDQTRLA + +AV ++QLADVVK+GAA+IGPGQ D+QV +L A +D A LR LLAASQ GR +DP +++ S++ LL+ V +

Sbjct 2259 IRESIHRTADALVQDTRSLVMCAGEDQTRLAGSTQMAVRNISQLADVVKQGAALIGPGQSDS QVLLILGAAKDAAVALRSTLLAASQVQGRPKTDPAFKEMETYE GSVESVAFLLKSVDS 2378

Query 1961 IADDENRGIQALVSAKYCRDQAAQLAPRPEPESLLT PVSRRSAPS VANVSAVSGSSSLIARYLAPDDLARAASGPVQTAVSKAILAGNTQTQRDVLSTAATTRDVVTDLVSAATALLR 2080  
I + RG +AL++ A CR QLA R P S T + + +S +SSL++RYLAPD+L+RA GP++T + +A+ +T+ Q D L A T R V DLV+A L R

Sbjct 2379 IDMESTRGTKALIATAHLCR----QLANR-MPSSAATAGQAVSLNDLGSATLLSSTSSLSVRYLAPDELSTRATEGPLRTTMDRAMATASTRRQDDALIFANTARHAVVDLVNACKNLQR 2493

Query	2081	CPEAVTETRTACVASTKGLAEFAGLLDALKAV--NSKPADHERIGNAARIADLSHTLLSQLDNLRSPRLVYFRASSPEWRDVAEKFIGRHVAYHHHVPTTYTIVGPGSTNRGLYVR	2198
Sbjct	2494	PE ER + + + EA LLDA+K V + P D R+ +A +I LS+ L+ +D +R++PRLVYFR S+PEWRDVAEK++GRH+A+HHHY G +NR +	2610
Query	2199	PICYPFRPQKAIPEQDEDRSGDEAEFRIATAIDQLKSILDCEPPSDPTILQSA-----HSVAVATQSLIHTARSIGLAPPSDAVDRARRAGTSHHGASLWRTELTVLATQDMCQLAQLC	2313
Sbjct	2611	C Q + +G EA + ID +L+ + QS V TQ ++ AR+I LA + V + A TS A+ R +L+VALAT+ +C+L +LC	2719
Query	2314	SAVLQEENAHPSKDLFGTSTRLVPSKERLLAAVRRVAAASAQLLMSAKSRKSSADIQKLVAGQVKEATDRLSGVLHQAN--FTC----DSTGLYVITPMP--ATLQNVIEHISK	2425
Sbjct	2720	S E P + L R++PS+ERLL AVR+VA AS +LL + +S ++ +SA++ +Q A ++VKE TD++S ++ + F+ G + TP+P + Q + ++	2834
Query	2426	IHSHQSELDVLQORLAQLQDN 2446	
Sbjct	2835	IHQ+ Q+ LD L+ + +L N	
Query	2426	IHTEQAALDDLRNEVRLNRN 2855	

### *C. sinensis* vs *E. multilocularis*

Score	Expect	Method	Identities	Positives	Gaps	
1727 bits(4474)	0.0	Compositional matrix adjust.	1070/2502(43%)	1524/2502(60%)	210/2502(8%)	
Query	29	VQHQRLLPPRRVGPDKDFVNGALEETIQPKMDDSIQMATG-----YMQTNGWSHQNLHTDRDTDFVDGQWAERPMSPDNFMMGGSHTFHYQEITPAQRALLVTIIEDVETLQAAKEQLYLS	143			
Sbjct	480	++ QRL P++ P +NG ++ + D+++ G Y+Q + R D + D+F+ GG+ + + +TPAQRAL VTI E++ L+ K++L +	589			
Query	144	IEKQRLQQPQALPP--LLNGTAQQMVVNDNMMVGYIGTSERRYIQAGSIPTLSSEIGRIKVD-----TTDDFITGGTYVCKTMTPAQRALHVITIEENINALEEGKQRLDIG	589			
Query	144	PDEDRII-GIGSDEASKRWLSESIGASQAKVTDEVGAMNAAVAQALRSANRAEAYDSTGMSQVGLGVSQ--DPGDDLMMMQSFRVVTIHFFAFVDDVVKRVAVLRREVGAIRA-DQTGEP	259			
Sbjct	590	P PAERVLANLGHDEASRRWLDESMGVTEVRVADEMGAMNAAVAQA+RSANRA DS G V GV++ DP DDLM +QQSFRVVT+HFFAFVD+VKRVAVLRRE G +RA D+ E	705			
Query	260	V---IAAQSEESSQNLLSAAARHVADAFDTLLESARPLATGQPREEVVTDESTIRSQAVTTG--PVSSSSRKAIIDAASRVGEASNDLLRHVMH-----EGD-----GEEDFEASSLMLTE	364			
Sbjct	706	+ I QSE ++QNLL AAR VAD+FT+LL++A PL T Q TD + VT V+ SSRKAI++AA+RVGEASND LR+V+ EG+ E ++A L L	817			
Query	365	LRLDITDQSEANAQNLGAARTVADSFTTELLKAAATPLTTRQ-----TDYNLDTVNVYVTEQGLQVAGSSRKAIIEAANRVGEASNDFLRYVLQDTEVLEGEIVPEVDERAYQADLLAL--	817			
Query	365	EERLYKDQLLSLAKAVANTTASLVVKAKNLATQTSLDPEAQHVIAAATQTGLCTSQLVACTKVLAPTIYQFSCQQQLSEAAAREVSWAVDGVVQASRAAGARTADQPVVQQP----VRT	480			
Sbjct	818	Y+D+LL LAK VAN+TA LVVKAKNLATQTSLD + QQ V+ AAT TGLCTSQLVACTKVLAPTI+Q SCQQQLSE+ EV+ +VD VV+ +R AG + A + +P + +	932			
Query	481	----YQDKLLELAKEVANSTAVLVVKAKNLATQTSLDLDHQQLVVMMAATNTGLCTSQLVACTKVLAPTIHQASCQQQLSESTYEVATSVDDVVRVARGAG-QAAHDTINPSEPEVHEIDS	932			
Query	481	AVAETEMAATEVRDALDQLNAHLSKASAKTSSGDALDNFQLAYDQLQQYHHDGQRMVASARRMAQATAQMIADIKAQAEALAADDPDRQRSRLFAAAKQLADATTILIASAKVCSTNPENP	600			
Sbjct	933	AV + E AATEVR +LD+LNAHL + S + GD LD FQ AYD+L+Q DG R+VA+ARR++Q TAQMI+D+K QAE DP+RQ+RLFAAAKQLADATT LI AK+CS+NP++	1050			
Query	601	AVMQVESAAATEVRTSLDRLNAHLLRGSIRPYQGDTLDLFLQQAYDELQ--ETDGVRLVAAAARRLSQVTAQMI+D+K QAE DP+RQ+RLFAAAKQLADATTILIASAKVCSTNPENP	1050			
Query	601	VTQEELRQAESNLNLIHVHSAADLLHRLRIRNLQAAAAAVTGATQLVNVSVQAAKRSRGNTYQVIVSDGKLVNGLIPRTVISIRESRACPDPLTQVELISACERFVLPCESLVRSRVS	720			
Sbjct	1051	Q +L+ AA+L ++ +AA+LL+ R+IR+LQAAAA V+ + L+ S+ AA +SR N Y ++ + K V+ L+P+ V +IR R P+DPL Q+ELI ACE + PC L RS RS+	1170			
Query	721	TPQADLKHAADNLVVVAYASAAELLNARVIRSLQAAAAAVVSANNLITTSRYAANKSRANNYHILINEKTVSDLLPKIVSAIRFIRRDPEPLAQLLELICACENAIQPCNQLARSCRSM	1170			
Query	721	APTVDSPDTQAALDNSTSQLSGAIEVLRCTLARLTPLRQMQVDGALARLARLAVEAAAIEGLKQGTLPDEKIDDCFNLLGTSVQDANVSTHELHSIVETMAQ-----QQRDMMRS	835			
Sbjct	1171	PT+ DP Q+ALDNSTSQ+ A+E L+ CLAR++P+ RQ+Q+DGALA L R EA +E K GTL LPD+K++D + +L S++DA +T+E+ S V+++ Q Q +	1290			
Query	836	VPTIGDPAVQSALDNSTSQMVVAVETLKACLARVSPARQLQMDGALASLLRTRVREAELEDLAKAGTLTALPDDKVEDRYRILKVSIRDARNATNEVTSAVDSLEQDYAAPQSPNCDQG	1290			
Query	836	SDVCGVPASMLATAIAGLVHATRIGVIAHDAADAGRSGAVGPPSSDSTRLAVTQDTRQAVQLAYELVKAARDARFAYDDGQMEALATVVVSHSEQ-----LTVQ	934			
Sbjct	1291	D G A+ LA A++ L + R + + +G A R A QLAY+LV AR ++EAL VS +Q L	1385			
Query	1291	EDWLGSSANRLAGAMSDLSRSIRSLVSEDPRIDSSGAFAA-----RTAAQLAYQLVCEAR-----KLEALQPTVSSPDQQADVLESNRNVAINCKLIGDNLNCA	1385			

Query 935 LLDTLGRCLRGLPGHREINEATNLIKRRREDLVQLSHEAPTRRVRVWVPSVYERAQSELTRAAVEFNQATGDLSSYSPGAFRRTRRRFTGAYDTLTVKGVDLCSGVPPEVPLPEPPP 1054  
 LL+ L L+ LPG +E++EA++L+ +RR+DLV+ + + PT +E + ER Q+E AAVEFNQAT DL S Y+P FRR++ RF GAYD+L KG+ L + P + P  
 Sbjct 1386 LLECL---LKALPGQKELSEASDLVGRRRKDLVRFA-DNPTVFYEYPIEETRVERTQTEFAGAAVEFNQATADLTSCYTPSNFRRSVRFAGAYDSLNVNKGQLSRAKSPKD-----P 1493

Query 1055 INRDVLVGLVNVSNHYSALMTEAGRVCARPDEQEIRDRFQTAARDVTEISISQLLTICTSGVTAEQRDCEVALRRLEALRPLENPNRPVNVKTYCECVDEVARISITPLADSLRMTSNAAK 1174  
 L+ GLV VS+ S ++ +A +VC+++P+ +R+ AAR VT+SIS +L+I SGV DC+ LRRLEA+RPLLE P RP NQ TY CV A+S+ PL+D +R + +  
 Sbjct 1494 TGPQLITGLVEVSDRSSEMLEDKQVCSQPEAEPLRQKLHLAARSVTDSSISHILSISVSGVVPVADCDLAEALRRLEAMRPLLEYPGRPTNQHTYQLCVSSFAQSLAPLSDGIRGTLDGVR 1613

Query 1175 EKHTQEFGAAVRQCSNSMCQLIEETAQAAYLIGLADPRSEPGRPSLVDPQLFMRSQREIQQVCDACDPSITNRQIITLSTEMARSAKTLCEACSAVSAQTNNPDARHQLNALTRETMQS 1294  
 +T+ F + +R+ ++ +C ++EET+QAAYLIG+A P SEPGR VD F R QR+I+ +C A+ P +T RQ++ L+ +M+ + ++L +AC+ ++AQ +N +AR +L L +++S  
 Sbjct 1614 SHNTEVFTSNIRKVTDCLVVVEETSQAAYLIGVAHPTSEPGRGGPVDIAFFERLQRDIKNICRAMKSPEVTERQVVNLANDMSANVRSLHDACNTIAAQASNQEARRRLQNLADDSVKS 1733

Query 1295 ITALIQRGNVSTAAAGSQLAEVNGEWEANRQVTLANARGVSSNVARLVHMVTQGPQFIGQAARVTEDEAREAQPVCSAGLSSLNAAQAVLRAAQTLITNARL-GQPEQ-AFLSFSTAS 1412  
 ++ LIQ + +W + R++T N V+ N+ RLV +T +F G+ AR++ EAR AQQP+C AG + L + + V+ A++ ++ A G+P AF +F+ AS  
 Sbjct 1734 MSELIQH-----SADWSDEGRRITANNIHAVNGNIDRLVSFITSISEFGGEPARISSEARSQQPICLAGQACLESGRGVILASKHMLQTAETSGEPSNTAFGAFTAAS 1837

Query 1413 KELSESIKALAAAMREHAPGQLEQCRVLQNINGLLQELQRAKMASMEGRLOPRRELNEEGFQKQLATCCRALLDAAPNVGRAARSEAEQLGHSVRVDSYLPGITSSAIAAASRSPLSST 1532  
 ++L+++ K L A + E PGQ ECQR L NI LL EL+R KMA M+G PR + NEEGF KQL RA+ D A +VGR A SEAE+LGH+VR +D LPG+ SSA+ AASR+P SS  
 Sbjct 1838 RDLTDNTHKLLLAVLNEQGGQAEQCRQLLNIRLLHELERDKMAMMDGVFTPRHDSNEEGFLKQLTISTRVTRDMATSVGRGATSEAERLGHAVRELDLLLPGLVSSALGAASRAPNSSA 1957

Query 1533 QLVYLEHASTILEAADQLVRVARDAGGNPRATHLHPLHDEAVRGLMESCEDLLTALDDVASRQGHVSTLIDTINRSLAQTEEIIQVFPVDARFADYQARLLRIARHMEQLTQAIQLRARQP 1652  
 QL+Y+EH T+LE+ +QL+ A+ AGGNPR+ + +D+ V+ ++SCE+LL+A++ VAS+Q ++ L+ + S A ++ + P DA F DYQ+RLLR+ARHMEQ T+A AR+P  
 Sbjct 1958 QLIYVEHTRTVLESVEQLISAQAGGNPRSDPIREAVDDGVKAQIDSCEELLSAIEGVASQQSFLAKLMAVLEESRALVDKPGRAPYDAHFTDYQSRLRLARHMEQQTEATVRAARRP 2077

Query 1653 --PSDGELAPLAHTLAQEYQEMCQTCCKGAAATLPDARQADELRGAVRAVGMATAELVQATTAAIRRAYDSHGDDQMLLSRN----QGIAY---LDTRAEALDSKLRDLIALLDTQGFNTQ 1763  
 + E+ L LAQ+Y E+C + A TL D +A++LR AV+ +G T L+ AT + S D SR + AY LD A A++ +LR+LIALL+ QGP TQ  
 Sbjct 2078 QESTSQEMTGLVQNLADQYAEICMISRDACETLRDPHEAEKLRWAVKNLQGTTRGLIMAT---VNVSSSRPDPSS--FSRSGVGPKPAYLRSLDFSAGAMNDRLELIALLEMQPGTQ 2191

Query 1764 ACLQASASTVSGIADLDTTILFASSTLHGPVHDDVDALLGPDYRPPYALPGAGEAIADRGRSFRGNREYPEEGFVTRDSIVRTARALVDDTQSLVSGTGEDQTRLATTAHVAVERVTQ 1883  
 ACLQ ASTVSGIADLDTTILFASSTLH P+ DD D +R PG + FG RE + F ++R+SI RTA ALV DT+SLV G GEDQTRLA + AV ++Q  
 Sbjct 2192 ACLQGASTVSGIADLDTTILFASSTLHPPIRDADVHQLEAHHRYSALS PGYDD-----FG-RE-AQGNFISIRESIHRTADALVQDTRSLVMGAGEDQTRLAGSTQKAVRNISQ 2301

Query 1884 LADVVKRGAAVIGPGQPDQVEVLSACRDVATGLRDVLLAASQTPGRDASDPVHEQVRNNVQLTSLNIGALLQKVKTIADDENRGIQALVSAAKYCRDQAAQLAPRPEPESELLTPVSSRR 2003  
 LADVVK+GAA+IGPGQ D+QV +L A +D A LR LLAASQ GR +DP +++ S++ LL+ V +I + RG +AL++ A CR QLA R + + ++ +  
 Sbjct 2302 LADVVKQGAALIGPGQSDS QVLILGAAKDAAVALRSTLLAASQVQGRPKTDPFAKEMEAYEGSVESVTFLLKSVSDIDMESTRGTKALIAATAHLR---QLANR-----MPSSAATAK 2412

Query 2004 SAPSVANVS AV----SGSSSLIARYLAPDDLARAASGPVQTAVSKAILAGNTQTRDVLSTAATTRDVTDLVSAATALLRCPEAVTETRTACVASTKGLAEFAGLLDALKAV--NSKP 2117  
 A S+ ++ + S +SSL++RYLAPD+L+RA GP++T + +A+ +T+ Q D L A T R V DLV+A L R PE E R + + + E A LLDA+K V + P  
 Sbjct 2413 QAVSLNDLGSATLLSTSSLVSRYLAPDELSRATEGPLRRTMDRAMATASTRQDDALIFANTARHAVVDLVNACKNLQRQPEIKEEFKGSSTVTRKVVLETAELLDVAKCVIAEANP 2532

Query 2118 ADHERIGNAARRIADLSHTLLSQLDNLRES PRLVMYFRASSPEWRDVAEKFIGRHVAYHHHYVPTYTIVPGSTNRGLYVRPICYFPRPQKAIPEQDEDRSGDEAEFRIATAIDQLKSIL 2237  
 D R+ +A +I +S L+ +D +R++PRLVMYFR S+PEWRDVAEK++GRH+A+HHHY G + R + C Q + +G+EA + ID +L  
 Sbjct 2533 NDLNRVSASAGKITMSSQLMRLIDRMRDNPRLVMYFRVSNPEWRDVAEKYLGRHIAFHHHYESCSDYDG---SYRATSSQISCL---QSQLDSTGRLLTGNEAALEASRRIDVALGLL 2645

Query 2238 DCEPPSDPTILQSA----HSVAVATQSLIHTARSIGLAPPSDAVDRARRAGTSHHGASLWRTELTVLATQDMCQLAQLCSAVLQEBNAHPDSKDLFGTSTRVLVPSKERLLAAVRRVAA 2352  
 + + QS V TQ ++ AR++ LA + V + A TS A+ R +L+VALAT+ +C+L +LCS E P + L R++PS+ERLL AVR+VA  
 Sbjct 2646 EGD-----FYQSMQRTTNQDVVHHTQPILSMARAVTLA-TREFVQTVQLASTSKAAANFGRDLDSVALATEQLCELTRLCSETRSE---GPYTAPL-DDPVRVMPSRERLLNAVRQVAT 2754

Query 2353 ASAQLMSAKSRRSKCSSADIQKLQVAGQLVKEATDRLSGVLHQAN--FTC--DSTGL--YVITPMPATLQNVIEHS--KIHSHQSELVQLQRLAQLQDN 2446  
 AS +LL + +S S+ +SA++ +Q A + VKE TD++S ++ + F+ +GL + TP+P +T+S +IH+ Q LD LQ + +L N

Sbjct 2755 ASGELLYAVQS-NSQITSANMTAVQTAARAVKEQTDKMSTMVRNKDPIFSLHPSQSGLEASITPIPIVSPQQEDTNSSAQIHTEQEALDDLQNEVERLNRN 2855

### C. sinensis vs H. microstoma

Score	Expect	Method	Identities	Positives	Gaps	
1696 bits(4392)	0.0	Compositional matrix adjust.	1030/2357(44%)	1468/2357(62%)	155/2357(6%)	
Query 98	SPDNFMGGSSHTFHQEIETPAQRALLVTIIEIDVETLQAAKEQLYLPDEDRI-IGIGSDEASKRWLSESIGASQAKVTDEVGAMNAAVAQALRSANRAEAY-DSTGMSQVGLGVSQDPGD	215				
Sbjct 533	S D+F GG+ + +++TPAQRAL +TI E++ LQ K++L + P +R+ I +G DEAS+RWL ES+GA++ +V DE+GAMNAAVAQA+RSANRA+A S G + G+ DP D	STDDFCGGTQYYICKKMTPAQRALHITIEENISALQEGKQRLDIGPPTERVLINLGHDEASRRWLDESMGATEVVRVADEMGAMNAAVAQAVRSANRADAVVRSVGPMTYGVAEITDPAD	652			
Query 216	DLMMMQQSFRVVTIHFPAFVDDVKKRVAVLRREVGAIRA-DQTGEPV---IAAQSEESSQNLLSAAHVADAFDTLLESARPLATGQPREEVVTDDESTIRSQAVTTGPFVSSSRKAIIDAA	331				
Sbjct 653	DLM MQQSFRVVT+HFPAFVD+VKKRVAVLRRE +RA D+ E V I +SE ++QNLL AAR VAD+FT+LL++A PL T Q ++ D ++ P SSRKAI++AA	DLMRMQQSFRVVTVHFPAFVDNVKKRVAVLRRESALLRAEDEPDERVRHGITDESEANAQNLLGAARTVADSFTTELLKAATPLTTRQTEYDL--DVVNYVTEQGVQAP--GSSRKAILEAA	768			
Query 332	SRVGEASNDLLRHVMHEG---DGE---ED---FEASSMLTTEERLYKQQLLSLAKAVANTTASLVVAKKNLATQTSLDPEAQHVIAAATQTGLCTSQLVACTKVLAPTIYQPSCQQQ	441				
Sbjct 769	+RVGEASND LR+VM + DGE ED ++A L + ++D+LL LAK VAN+TA LVVAKKNLATQT+LD E QQ V+ ATQTGLCTSQLVACTKVLAPTI+Q SCQQQ	NRVGEASNDFLRYVMQDSEVTDGEIVPEEDEGVYQADQLAI-----FQDKLLELAKEVANSTAVLVVAKKNLATQTNLDSEHQMVVMNATQTGLCTSQLVACTKVLAPTIHQASCQQQ	882			
Query 442	LSEAAREVSWAVDGVVQASRAAGARTADQ---PVQVQPVRTAVAETEMAATEVRDALDQLNAHLSKASAKTSSGDALDNFQLAYDQLQYHHTHDGQRMVASARMAQATAQMIADIKA	557				
Sbjct 883	LSE+AR+VS AVD VV+ +R++G D+ VQ+ V AV E E AA EVR +LD+LNAHL SA+ GD+LD FQ AYD LQQ DG R+VA+ARR++Q TAQMI ++K	LSESARDVSNVAVDSVVRTARSSGQTAHDRMDLSDADVQE-VNAAVIEVEDAAVEVTRSLDRLNAHLLSGSARPYQGDSDLDFQQAYDALQQ--ETDGIRLVAAARLSQVTAQMINNLKV	999			
Query 558	QAELAADDPRQSRFLFAAAKQLADATILIASAKVCSSTNPENPVTQEELRQAAESLNLIVHSAADLLHRLRIRNLQAAARAAVTGATQLVNVSVQAAKRSRGNTYQVISDGKLVNGLIP	677				
Sbjct 1000	QAE DP+RQ+RLFAAAKQLAD+TT LI AKVCS++P++PV + L+ +A+ L ++ +++AA+LL+ R+IR+LQAAARA V+ + L+ S+VA+ +SR N Y + SD + VN L+P	QAETLEGDPERQTRLFAAAKQLADSTTDLINQAKVCSDDPSPVRKAALKHSADLVVVAYASAAELLNARVIRSLQAAARA VVSASNNLITTSRVASSKRSRNNYHMASDERAVNDLLP	1119			
Query 678	RTVISIRESRACPDPLTQVELISACERFVLPCELSVRSSRSVAPTVDPTQAALDNSTSQLSGAIEVLRTCLARLTPLTRQMQVDGALARLARLAVEAAAIEGLKQGTLPAPLDEKI	797				
Sbjct 1120	+ V +IR R P+DPL Q+ELI C++ V PC L RS R + PT+SD Q ALD S+SQ+ A+E L+ CLAR +P+ RQ+Q+DGALA L R EA IE K GTL LPD+++	KMVSARIRFIRRPEDPLAQLELICTCKQAVQPCSQLARSCRQIVPTISDAALQNALDGSSSQMVVAVETLKAACLARASPIARQLQMDGALASLLRTVREAEIEKLAAGTTLTALPDDR	1239			
Query 798	DDCFNLLGTSVQDANVSTHELHSIVETMAQ-----QQRDMMSRSDVCGVPASMLATAIAGLVHATRGIVAHDAYDAADAGRSQAVGPFSSDSTR LAVTQDTRQAVQLAYELVKAARDARFA	912				
Sbjct 1240	+D + L V+D S E+ S V+T+AQ Q + D G A+ LA+ ++ LV +TR +A + DA + V R + QLAY+LV AR	EDRYQTLKVGVRDTQNSISEVLSAVDTLAQDYASPPQSPNRDQSEDWLGSSANRLASGVSDLVSTRSTRCLAEEER-PVDASLASCV-----RSSAQLAYQLVCEARKLEPV	1343			
Query 913	YD-----DGQMEALAT---VVSHSEQLTVQLDLDLTLGRC-LRGLPGHREINEATNLIKRREDLVQLSHEAPTRRRVWVPEPSVYERAQSELTRAAVEFNQATGDLSSYSPGAFRRTRRF	1023				
Sbjct 1344	D Q + L + V ++ + + L TL C L+ LPG +E++EA+NL+++RR+DLV+ S E P+ +E + E+ QSE T AAV+ NQAT DL S Y+P F +++ RF	QQANLTPDQDFVLESNRNVNANCKLVGDNLNRTLSECLLKALPGQKELSEASNLVQRRRDLVRFSEE-PSVFEYPIEETRIEKIQSEFTGAAVDLNQATADLTSCYTPVNFCKSSVRF	1462			
Query 1024	TGAYDTLTVKGVLDLMSGVPPNEVPLPEPPPINRDLVDGLVNVSNHYSALMTEAGRVCARPEQEIRDRFQTAARDVTESISQLLTICTSGVTAEQRDCEVALRRLEALRPLENPNRPV	1143				
Sbjct 1463	GAYD L KG+ L + + P + L+ GLV+VS+ S ++ A ++C++P+ + +R R AAR VT+SIS LL++ SGV DC+ ALRRLE++RPLLE P+RPV	AGAYDNLVDKGIQLSRTRSSKD-----PSSNQLISGLVDVSDRSDEMLESAKQLCSQPEAEPLRQRLGAARSVTDSISHLLSVSASGVIPGVADCDALRRLESMPRPLEYDPRV	1574			
Query 1144	NQKTYCECVDEVARSIITPLADSLRTMSNAAKEKHTQEFGAAVRQCSNSMCQLEIETAQAAYLIGLADPRSEPGRPVLDVDPQLFMRSQREIQVCDIACDPSITNRQIITLSTEMARS AKT	1263				
Sbjct 1575	N Y CV A+S+TPL++ +R + K +T+ F + VR+ ++S+CQL EET+QAAYLIG++ P SEPGR L+D F R QR+I+ +C +I P +T R++++L+ +M+ + ++	NSHIYQTCVHACAQSLTPLSEGIRGTLDGVKSHNTEVFTSNVRRVTDLSLQLAEEETSQAAYLIGVSHPTSEPGRVGLIDIAFFERLQRDIRNICRSISSPQVTEREVS LANDMSANMRS	1694			
Query 1264	LCEACSAVSAQTNNPDARHQNLALTRETMQSITALIQRGGNVSTAAAGSQLAEVNGEWEANRQVTLANARGVSSNVARLVHMVTQGPQFIGQAARVTDEAREAQQPVCSAGLSSLNAAQ	1383				
Sbjct 1695	L +AC+++ A++ NP+A+ +L AL+ + QS++ LIQR + +W + R+ T +A V++N+ RL++ ++ +F + +R+ E + QQP+C+AG + L + +	LRDACNSIRAKSTNPEAQRKLQALSNDATQMSMDLIQR-----SADWSDEGRATNNHAHVNTNIDRLMNYLSSTNEFTSEPSRMGGEIKAVQQPICAAGQACLES GK	1798			
Query 1384	AVLRAAQTLITNAR-LGQP-EQAFLSFSTASKELSES IKALAAAMREHAPGQLEQQRVLQNINGLLQELQRAKMASMEGRQPRLNEEGFQKQLATCCRALLDAAPNVGRAARSEAQ	1501				

Sbjct 1799 V+ A++ ++ A QP + AF +F+ AS++L+++ K L + + APGQ ECQR L NI LL EL+R KM+ M+G PR + EEGF KQL+ RAL D AP VG A SEAE+ GVILASKHMLQTAESKAQPSDTAFAAFTAASRDLDNTKHLLSVLNAQAPGQAECQRALLNIRLLHELERDKMSIMDGAFAPRHQATEEGFLKQLSVNSRALRDIAPQVGHGATSEAEK 1918

Query 1502 LGHSVRVVDVSYLPGITSSAIAAASRSPSSQLVYLEHASTILEAADQLVRVARDAGGNPRATHLHPLHDEAVRGLMESCEDLLTALDDVASRQGHVSTLIDTINRSLAQTEEIIQVPPVD 1621  
LGH+VR +D LPG+ SSA+ AASRS S Q+ YLEH T++E+A+QLV A+ AGGNPR+ +L +++ V + SCE+LL+A++ ++S+Q ++ L+ ++ S ++ ++P D

Sbjct 1919 LGHTVRELDLLLPGLVSSALGAASRSNPSAQMTYLEHTRTVIESAEQLVSAAKQAGGNPRSDNLRVTVNDVCVNAQVASCEELLSAIEGISSQQNFINKLMAILDESQNLVDPKGRIPSD 2038

Query 1622 ARFADYQARLLRIARHMEQLTQAIQLRARQP--PSDGELAPLAHTLAQEYQEMCQTCCKGAAATLPDARQADELRGAVRAVGMATAELVQATTAARIRAYDSDHGDQMLLSRN-----QG 1733  
A F DYQ+RLLR+ RHMEQ T+ AR+P + E+ L +LAQ+Y E+C + A T+ D +A++LR AV+ +G T L+ AT + A S D L +

Sbjct 2039 AHFTDYQSRLRLVRHMEQQTEVTVRTARRPQETASQEITGLVQSLAQDYAEICMISRDACETIRDPHEAEKLRWAVKNLQQTTRGLILAT----VNAASSKSDTSSLTRKALVRPPAS 2154

Query 1734 IAYLDTRAEALDSKLRDLIALLDLTQGPNTQACLQSASTVSGIIADLDTTILFASSGTLHGPHVDDVDALLGPDDYRPPYALPGAGEAIADRGSRFGNREYPEEGFGTVRDSIVRTARALV 1853  
I LD A A++ +LR+LIALL+ QGP TQACLQ ASTVSGIIADLDTTILFASSGTLH P+ DDVD + + + Y+ +G DR + + F ++R+SI RTA ALV

Sbjct 2155 IRSLDFSAGAMNERLRELIALLAQAQPGTQACLQGASTVSGIIADLDTTILFASSGTLHPPIRDVD--MQNLESQQRYSAISSGYDAFDREA-----QSHFISIRESIHRTADALV 2264

Query 1854 DDTQSLVSGTGEDQTRLATTAHVAVERVTQLADVVKRGAAVIGPGQPDTQVEVLSACRDVATGLRDVLLAASQTPGRDASDPVHEQVRN--NVQLTSLNIGALLQKVKTIADENRGIQA 1971  
DT+SLV G GEDQTRLA + +AV ++QLADVVK+GAA+IGPGQ D+QV++LS +D A LR+ LL A+Q DPV +V+ N T SN LL+ V + + NRG +A

Sbjct 2265 QDTRSLVLGAGEDQTRLAGSTQMAVRNISQLADVVKQGAALIGPGQSDSQVILSTAKDTAVALRNTLLTAAQVQSGSKKDFVFNVEVKTYENEVGTCSN--RLLKAVDAVDMEANRGSKA 2382

Query 1972 LVSAACYCRDQAALAPRPEPESELLTPVSSRRSAPS VANVSA----VSGSSSLIARYLAPDDLARAASGPVQAVSKAILAGNTQTQRDVLSTAATRDRVVDLVAATAALLRCP EAVTE 2087  
L+S A CR QLA R P SAPS ++ + +S +SSL++RYL PD+LARAA GP++ V KA+ A +TQ Q D L A T R + DLV+ L + E E

Sbjct 2383 LISTAHLCR----QLASR-----MP----NSAPSFNDLGSATTLMSTTSSLVSRYLTPDELARAAEGPLRAIVDKAVTAASTQRQDDALVCANTARHAILDLVNTCKNLQKQEIKEE 2487

Query 2088 TRTACVASTKGLAEFAGLLDALKAV----NSKPADHERIGNAARIADLSHTLLS QLDNLRESPLRVMYFRASSPEWRDVAEKFIGRHVAYHHHVPTTYTIVGPGSTNRGLYVRPICYP 2203  
R + K + E A L+D++K V N+K D R+ N+A +I LS L+ +D +R+SPRLVMYFR +SPEWRDVAEK++GRH+A+H+HY + ++ + +R C

Sbjct 2488 FRKQSGDVTKKIVLETAELMDSVKGVIAGENTK--DSSRVNSAAKIGALSQMLRLIDKMRDSPRLVMYFRVNSPEWRDVAEKYLGRHIAFHNYH---ESCTEQDGFSGKSI SIRTSCLO 2602

Query 2204 FRPQKAIPQEDEDRSQDEAEFRIATAIDQLKSILDCEPPSDPTILQSAHSVAVATQSLIHTARSIGLAPPSDAVDRARRAGTSHHGASLWRTLTVALATQDMCQLAQLCSAVLQEEENAH 2323  
+ + + EA RI A+ L+S L+ + +P + V TQ ++ ARS LA + V + A S +L R L+VALA + +C L++LCS E

Sbjct 2603 SQLDTKVFLMSGAAEAIEAIHRIDIALGLLES DLN-KSKQEPI---NQDIVVHTQPILSLARSTTLA-IREFVQTVQALASRESTNLGRNLNSVALAIEQLCDLSRLCSETRSEGPYT 2716

Query 2324 PDSKDLFGTSTRLVPSKERLLAAVRRVAAAASQLLMSAKSRKSCSSADIQKLQVAGQLVKEATDRLSGVLHQANFT 2400  
D TR++PS++RLL AVR+VA+AS +LL + +S + SS ++ +Q A + +KE TDR+S ++ + T

Sbjct 2717 APVDD---PTRVMPSRQRLNNAVRQVASASGELLCAVQS-NPQISSTNLASVQAAARAIKEQTDRMSQLIRNKDPT 2788

### C. sinensis vs T. solium

Score	Expect	Method	Identities	Positives	Gaps	
1721	bits(4457)		0.0	Compositional matrix adjust.	1064/2416(44%)	1500/2416(62%) 166/2416(6%)
Query	100	DNFMGGSHTFHYQEITPAQRALLVTIIEDVETLQAAKEQLYLSPDEDRI-IGIGSDEASKRWLSESIGASQAKVTDEVGAMNAAVAQALRSANRAEAYDSTGMSQVGLGVSQ--DPGDD				216
Sbjct	610	D+F+ GG+ + +TPAQRAL VTI E++ L+ K++L + P +R+ + +G DEAS+RWL ES+G ++ +V DE+GAMNAAVAQA+RSANR DS G V GV++ DP DD				726
Query	217	DDFITGGTQYYVCKTMTPAQRALHVTI EENISALEEGKQRLDIGPPTERVLVNLGHDEASRRLWDESMGVTEVRVADEMAMNAAVAQAVRSANRV---DSMGGGGVRFVGVAEVTDPSDD				332
Sbjct	727	LMMMQQSFRVVTIHFPAFVDDVKRVAVLRREVGAIRA-DQTGEPV---IAAQSEESSQNLLSAAHVADAFDLDLESARPLATGQPREVVTDDESTIRSQAVTTGVPSSSRKAIIDAAS				842
Query	333	LMRQQSFRVVT+HFPAFVD+VKRVAVLRRE G +RA D+ E + I QSE ++QNLL AAR+VAD+FT+LL++A PL T Q D T+ V+ SSRKAI+AA+				444
Sbjct	843	RVGEASNDLLRHHMHEGDGEEDFEASSLMLTEEER-----LYKDQLLSLAKAVANTTASLVVKAKNLATQTSLDPEAQHVI AATQTGLCTSQLVACTKVLAPTIYQSCQQQLSE				958
		RVGEASND LR+V+ + + E ++ +ER LY+D+LL LAK VAN+TA LVVKAKNLATQ+LDP+ QQ V+ +AT TGLCTSQLVACTKVLAPTI+Q SCQQQLSE				
		RVGEASNDFLRYVLQDTEVVE----GEIVPEVDERAYHVDLLALYQDKLLELAKEVANSTAVLVVKAKNLATQTNLDPDHQQLVVISATNTGLCTSQLVACTKVLAPTIHQASCQQQLSE				



Query	445	AAREVSWAVDGVVQASRAAGARTADQPVVQQP----VRTAVAETEMAATEVRDALDQLNAHLSKASAKTSSGDALDNFQLAYDQLQQYHTHDQGQRMVASARRMAQATAQMIADIKAQAE	560
Sbjct	959	+A EV+ AVD VV+ +R AG D + + +P V +AV E E AATEVR +LD+LNAHL + S + GD LD FQ AYD+L+Q DG R+VA+ARR++Q TAQMI+D+K QAE	1075
Query	561	LAADDPDRQRSRLFAAAKQLADATTILIASAKVCSSTNPENPVTQEEELRQAESLNLIHVHSAADLLHRRILIRNLQAAARAAVTGATQLVNVSVQAAKRSRGNTYQVISDGKLVNGLIPRTV	680
Sbjct	1076	DP+RQ+RLFAAAKQLADATT LI AK CS++P++ Q +L+ AA+ L ++ +++AA+LLH R+IR+LQAAARA V+ + L+ S AA +SR N Y +++D K V+ L+P+ V	1195
Query	681	ISIRESRACPDPLTQVELISACERFVLPCELSVRSSRSVAPTVDPTTQAALDNSTSQLSGAIEVLRCTCLARLTPLRQMQVDGALARLARLAVEAAAIEETGLKQGTLPPLPDEKIDDC	800
Sbjct	1196	+IR R PDDPL Q+ELI ACE + PC L RS RS+ PT+ D Q+ALDNSTSQ+ A+E L+ CLAR +P+ RQ+Q+DGALA L R EA +E K GTL LPD+K++D	1315
Query	801	FNLGTSVQDANVSTHELHSIVETMAQ-----QORDMMSRSDVCGVPASMLATAIAGLVHATRGIVAHDAYDAADAGRSGAVGPPSSDSTR LAVTQDTRQAVQLAYELVKAARDAR-----	910
Sbjct	1316	+ +L S++DA +T+E+ S V+++ Q Q + D G A+ LA A++ L+ +TR + + G S A A QLAY+LV AR	1419
Query	911	FAYDDGQMEALAT---VVSHSEQLTVQLLDTLGRCLRGLPGHREINEATNLIKREEDLVQLSHEAPTRRVRWVEPSVYERAQSELTRAAVEFNQATGDLSSYSPGAFRRTRRFTGA	1026
Sbjct	1420	+ D Q + L + V + + + L L C L+ LPG +E++EA++L+++RR+DLV+ + + PT +E + ER QSE T AAVEFNQAT DL S Y+P FRR++ RF GA	1538
Query	1027	YDTLTVKGVDLMSGVPPNEVPLPEPPINRDLVDGLVNVSNHSYALMTEAGRVCARPDEQEIRDRFQTAARDVTEISISQLLTICTSGVTAEQRDCEVALRREALRPLENPNRPVNQK	1146
Sbjct	1539	Y+TL KG+ L + P + P L++GLV VS+ S ++ +A +VC++P+ + +R + AAR VT+SIS LL++ SGV DC+ LRRLEA+RPLLE P+RP NQ	1650
Query	1147	TYCECVDEVARSIPTLADSLRMSNAKEKHTQEFGAAVRQCSNSMCQLIEETAQAAYLIGLADPRSEPGRPSLVDPQLFMRSQREIQVQVCAICDPSITNRQ----IITLSTEMARSAK	1262
Sbjct	1651	TYQLCVTSFAQSLVPLSDGIRGTLDGIRSHNTDVFVTSNIRKVTDCQLQVVEETSQAAYLVGVAHPTSEPGRGTVPDIALFERLQRDIQISICRAMKSPEVTEREDRARVVNLANDMSANVR	1770
Query	1263	TLCEACSAVSAQTNNPDARHQLNALTRETMSITALIQRGNVSTAAAGSQAELVNGEWDEANRQVTLANARGVSNVARLVHVMVQGPQFIGQAARVTDEAREAQQPVCAGLSLNA	1382
Sbjct	1771	+L +AC+A++AQ + +AR +L L ++M+S++ LIQR + +W + R+VT+ N V+ N+ RLV +T +F G+ AR++ EAR AQQP+C+AG + L +	1874
Query	1383	QAVLRAAQTLITNARL-GQP-EQAFLSFSTASKELSESIKALAAAMREHAPGQLECRVQLNINGLLQELQRAKMASMEGRLQPRELNEEGFQKQLATCCRALDAAPNVGRAARSEAE	1500
Sbjct	1875	+ V+ A++ ++ A GQP + AF +F+ AS++L+ + K L A + E PGQ ECQR L NI LL EL+R KMA M+G PR + NEEGF KQL RA+ D A VGR A SEAE	1994
Query	1501	QLGHSVRVVDYSLPGITSSAIAAASRPLSSTQLVYLEHASTILEAADQLVRVARDAGGNPRATHLPHLDEAVRGLMESCEDLLTALDDVASRQGHVSTLIDTINRSLAQTEEIIQVPV	1620
Sbjct	1995	+LGH+VR +D LPG+ S + + AASR+P SS QL Y+EH T+LE+ +QLV A+ AGGNPR+ + +D+ V+ ++SCE+L++A++ VAS+Q ++ L+ + S A ++ + P	2114
Query	1621	DARFADYQARLLRIARHMEQLTQAIQLRARQP--PSDGELAPLAHTLAQEYQEMCQCTCKGAAATLPDARQADELRGAVRAVGMATAELVQATTAARIRAYDSHGDQQMLLSRNQGIA---	1735
Sbjct	2115	DA F DYQ+RLLR+ARHMEQ T+A AR+P P++ E+ L LAQ+Y E+C + A TL D +A++LR AV+ +G T L+ AT + S D LSR G+	2227
Query	1736	-----YLDTRAEALDskLRDLIALLDTPGNTQACIQSASTVSGIIADLDTTILFASGTLHGPVHDDVDALLGPDDYRPPYALPGAGEAIDRGSFRGNREYPPEGFGTVR-----	1842
Sbjct	2228	LD A A++ +LR+LIALL+ QGP TQACLQ ASTVSGIIADLDTTILFASGTLH P+ DD D +R Y+ AI+ FG RE + F ++R	2337
Query	1843	--DSIVRTARALVDDTQSLVSGTGEDQTRLATTAHVAVERVTQLADVVKRGAIVIGPQPDTQVEVLSACRDVATGLRDVLLAASQTPGRDASDPVHEQVRNNVQLTSLNIGALLQKVK	1960
Sbjct	2338	+SI RTA ALV DT+SLV G GEDQ RLA + +AV ++QLADVVK+GA +IGPGQ D+QV +L A +D A LR LLAASQ GR +DPV ++++ S++ LL+ V +	2457

Query 1961 IADDENRGIQALVSAAKYCRDQAAQLAPRPEPESELLTPVSSRRSAPSVANV-SAVSGSSSLIARYLAPDDLARAASGPVQTAVSKAILAGNTQTQRDVLSTAATTRDVVTDLVSAAATALL 2079  
 I + RG +AL++ A CR QLA R S +TP S + + + +S +SSL++RYLAPD+LAR GP++ V KA+ A +T+ Q D L +A T R V DLV+A L  
 Sbjct 2458 IDMESTRGTKALIAHAHLR----QLANR--MPSCVTPTGQAVSLNDLGSATLLSTSSLSVRYLAPDELARTTEGPLRMTVDKALGAASTRRQDDALISANTARHAVVDLVNACKNLQ 2571

Query 2080 RCPEAVTETRACVASTKGLAEFAGLLDALKAV--NSKPADHERIGNAARRIADLSHTLLSQLDNLRESPRLVMYFRASSPEWRDVAEKFIGRHVAYHHHYVPTYTIVGPGSTNRGLYV 2197  
 R E E R + + + E A LLDA+K V + P D R +A +I +S+ L+ +D +R++PRLVMYFR SSPEWRDVAEK++GRH+A+HH Y + + + R  
 Sbjct 2572 RQQEIKEELRKGADTVRKVVLETAELLDAVKCVVAEANPNDINRASASAAKITAMSNHLMCLIDRMRDNPRLVMYFRVSSPEWRDVAEKYLGRHIAFHQY---ESCIEDGSGYRATTS 2688

Query 2198 RPICYPFRPQKAIPEQDEDRSGDEAEFRIATAIDQLKSIL--DCEPPSDPTILQSAHSVAVATQSLIHTARSIGLAPPSDAVDRARRAGTSHHGASLWRTELTVALATQDMCQLAQLCSA 2315  
 + C Q + +G EA + ID +L D +I Q VA TQ ++ AR++ LA + V + A TS A+L R +L+VALAT+ +C+L +LCS  
 Sbjct 2689 QLSCL---QSQLDSTGRLLTGTEAALEASRRIDAALGLLEGDLYQSMQKSINQ---DVAHHTQPILSMARAVALA-TREFVQTVQLAFTSKTSANLGRDLDSVALATEQLCELTRLCSE 2800

Query 2316 VLQEENAHPSKDLFGTSTRLVPSKERLLAAVRRVAAASAQLLMSAKSRRSKCSSADIQKLQVAGQLVKEATDRLSGVLHQAN--FTC--DSTGLYV--ITPMP--ATLQNVIEHSKI 2427  
 E D R++PS+ERLL+AVR+VA AS +LL + +S +SA++ +Q A + VKE TD++S ++ + F+ +GL V +TPM ++ + T ++IH  
 Sbjct 2801 TRSEGPTYTAPMDD---PIRVMPSTRERLLSAVRQVATASGELLYAVQS-NVHITSANMIAVQTAARAVKEQTDKMSTMVRNKDPIFSLPSQSGLEVSLLTTPMSLVSSQEENTNTSTQIH 2915

Query 2428 SHQSELDVLQORLAQL 2443  
 + Q+ LD LQ + +L  
 Sbjct 2916 TEQTVLDDLQNEVERL 2931