

Characterization of p53 family homologs in evolutionary remote branches of Holozoa

Václav Brázda, Martin Bartas, Jiří Červeň and Petr Pečinka

Supplementary material 12: FASTA sequences of eleven non-Metazoan p53 family homologous sequences and twenty-seven other representative closest relative Metazoan species for protein tree construction

>XP_020604944.1 cellular tumor antigen p53-like isoform X1 [Orbicella faveolata]

MFLPQNPSTSPPADPPSTAVSRDSNVMVTCGPKDRPGYTFTGEAFRKILTNPDLNKMADQVEPTIGPLINEKQLKQ
LMEDVRGSDDGNCFLSMGGGEDSQDLSFGVQSDMRGYGPQYHYISSQAVHDLPTSSSPQPPAYPHNIPYSPGP
VAGHRSSSTCISISTTPYETCSNYSEASTSGPSTPPPGLTSSYMPIHARLPTIPSTTEYPGNWNFQISFGPLTESASKSA
TWTYSDVCKKLYVNLASFCAPIKFTNSPPPPGTYLRAVAVFKGSTNLHDIVKRCPNHMESSNDGSAPNNHFIRSNNP
AARYHTCPESRRHSVLLPYNGPQVGTEYVTEMFAMCFSSCASGPSRRPVEVIFTLEQNGQTLGRRVVEIRVCACPG
RDRKSDEKVACGEPSSSQGVRRTKSGSGNKGPSKRRRTSSRSQDEEFIITTKNRHVYDILLEMKDSLEGKFESLNIQS
PEHSLVLSTSEADQPSTSGSPGAVFKRVGLSSM

>ABF72158.1 transcription factor p63 [Nematostella vectensis]

MSQSSSQESQILSQGTYEVLRSFVAEAEAGAEYQYNVAPYSVDLTSCNEQENLEINIEKVALIPSDFLDYQGVNGFHTE
PVESREQYHHHVPSAITPPPDYPTTNLVPANTDFPGEYGFDFVGFDKENGPTPKSAPWTYSHQLQKLLCRMKCLVP
VRLVFRSKVPEGFYIRAVVVYKQPEHFRVVERCANHITRQDDGHTAPKHLLRCENTKTLYRTCNTLGRHELMFPT
RKPDAGMDYFKDMFQFMCFNSCPGGLNRRPIIVFTLELSGVVYGRKVLDRVCACPGRDKADQESLETRKKNKM
KNLNTPPPTPDENKPTSGETPAGSHNGVKEELESSHKSSDKRDIRQCTKKGFELHISKRIKREKEDEEEFTITVRGR
DNYEMLLKIKESLDLMNMVPQSVVDNYRKTCPPCGPTWFLDQKPKPIQQVESSTDSELECSPPNSLSQLSRTSSN
ASSQGRRAAAEYSAVRFSKRTVTYPPSRH

>XP_015761542.1 PREDICTED: cellular tumor antigen p53-like isoform X1 [Acropora digitifera]

MGVSVLCDFVFRVGIHVTCRGNHKTCCQRTAAEGIRSFSTREVNMAEEQVEPTISLNEKQFKHLLADVARGENSS
SCFLSMGEDSHDSLNFQSDMRPFTQQFHYSVSSSQSPTSTSPQPSYSPQPPAYPNIPASYSPTSNPANRTSISSSP
YDNCNTYNDGSATASSCSQGIYMHARLPTIPSTTEYPGDWGFELSFGPLTESASKSATWTYSDVCKKLYVNLASFCA
PIKFKTRMQPPAGTVLRAVAVFKGSTNLHDVVKRCPNHMESSNDGSAPKDQFIRSNNPAAQYHTCPESMRHSVLL
PYNGPQVGTEFVTEMFAMCFSSCASGPSRRPVEVIFTLERDQTLGRRVVEIRVCACPGRDRKSDEKAVSGEPTSS
QGIVKRARRGSGNKGPGAKRRRTSSRTSSEEEYVITTRNRHVYELLMDFKKLEDKFEYVHQNLDNSLVLSTSEDET
EQPSTSSVAPGSVFKRGGLSTM

>RMX37105.1 hypothetical protein pdam_00016598 [Pocillopora damicornis]

MVDNVEPTINTLLHDKHYKHLDDVRGDENPSCFLSPLETGEDSQDNFLPSDMRSYGQFPYIPSSQAPVDLPTSTS
PNPPAYPHNIPQSPGLVSVHRNSSISPINTAHYETSSTYSVDVSTPGTPATSSPNMVHTFMHPSRLPQIPSTTEYPGEL
NFQLSFGPLTESASKSATWTYSDVCKKLYVNLNSYCAPIKFTSCPLPGTYLRAVAVFKGSTNLHDVVKRCPNHMEN
SNDGSAPNAHFMRSSNNPAAHYHTCPDSSRHSVPIPYSGPQVGTEYVTEMFAMCFSSCASGPSRRPVEVIFTLERD
GQTLGRRVVEIRVCACPGRDRKSDERIACGEPSSSQGAKRPRKSGKGPMSKRRRTSSRSQEEFTVTTKNYHVYEILL
DMKEKLEEFESWNIQSPDHSMVLSTSEDEGEQPSTSSSPGAVYKRTSLASM

>XP_028403575.1 cellular tumor antigen p53-like [Dendronephthya gigantea]

MESRPDHPNSLIPQEQMFFNNIDPNEAGRTFFANNPGMAMPQDDFMRQPDGNMEECCDPHQYLPPTSSPLSF
PPVLTSGISNTRHQISTAVTSSDSSCSRPHYQQSPYRDNQVSSSFPMNTNGFYQRAQLPTVPTSKQFSGDFNFEISFGP
PTESAPKSATWTYSPKTSKLYTNIGSYCPFQFKTNGKSMSPAFIRAVAMFKGSSNLHDIVKMCANHAENNTDDSQ
MRNNFFIRSNNANAVYLECPQTGRHSVLVPYSGPQAGTNFVTEMFAFKCFSSCASGHGRKPIEIIFTLEDAETGNIL
GRCVVEIRVCACPGRDRRSDEVNLDRIASNNNSNKRASKQKSYNRKKLKKASHEEYTTTTTSKEVHDLVLTIKIALENK
LSYKSEFDAVAEEENSSAKRVQPLGASTSHEASIENKQTYTSCI

>XP_022798783.1 cellular tumor antigen p53-like isoform X1 [Stylophora pistillata]

MRHDARKSDSYLPFDSSSRVEINFGSVKRGVLGDLASRTVQNMVDNVEPTINTFLQAKQYKSLDDVREENTSCFL
SPLGTGEDSQDNFLPSDMRPYQGSFPFIPSSQAPVDLPTSTSPNPPAYPHNIPCSPGLVSVHRSSSSISPINTAPYETQ
SSTYSDASTPGTPAASSPNMVPNYMYQPQPQSRLPQIPSTTEYPGDLNFQISFGPITESASKSATWTYSDVCKKLYV
NLNSYCPKFKTSQAPPHKTYLRVAVVFKGSTNLHDVVKRCPNHMENSNDGSAPNNHFMRSNNPAAGYHTCPRESS
RHSIIIPYSGPQVGTEYVTEMFAFMCFSSCASGPSRRPVEVIFTLEREGQTLGRRVVEIRVCACPGRDRKSDERACGE
PSSSQGAKRTRKSGKGPLSKRRRTSFRSHEEEYSVSTRNYHVVEILQDMKDNLDEKFESWNIQSPEHSTVLSTSEDEG
EQPSTSGSPGALYKRVGQSSM

>XP_020898669.1 cellular tumor antigen p53 isoform X1 [Exaipasia pallida]

MTNDIVFHHVYSRCHISVILSRLSQKRGNETPCKHSIIYYSIILFNASHLTMKRKEELEMVDKLMDSFFNVGGEETD
FNPEIRSMFPGYPIPYGSHGDFPPQTTGPPAYPQNIQGPMYTNVNTAGRRTPVSPVSNQSDMHYDNKSDLSPGHS
PYNHLPTIPSTTEYPGEMGFAISFGPATESASKSATWTYSECKKLYVNLASFCAPIKFKTDLGPPGLGCYLRAVAVFKGS
TNLHDIVKRCPNHSQTSQDSSDNNLHFIIRSNNPAAHYHTCPESGRHSVVIYPYHGPQVGTEYVTEMFAFMCFSSCPS
GPSRRPVEIIFTLEKDGITLGRVVEIRVCACPGRDRKSNESLSGVSSTSASSKRKKGQKSSKKQRATSSGDEGDKE
WTITGKDIRVYNLLMTMKEQLESQFMEETPIHSEDEDPVFGPSTSTSTATSYPNLRYKRVGHSSM

>XP_004209954.2 PREDICTED: cellular tumor antigen p53-like [Hydra vulgaris]

MVVWNTMESSIRGMIHEPPTYLNQLMSNHAEQTSVLKSEDLNNNTNSVVPFLNGNCSFYGQEQSHYNQVY
PFDYSNGCSININTNLVNTLQTLSSCGANTYARKTEQSLSSREAISYEPNLYLEMQNHSRIQTNIHLPDPTTNFPGEF
EFQLSLGPHTESAASADFTFSSILKKLYVQRASQCPFMFKTEKKPPSNSYVKILPVFKGTHVLEPVRRCNSHAEVDD
EGGPRYHFIRASNQSCYNHSDDGRLFVTVPYNGPQVGAEQIELLSFMCFNCSGHSSHSKRRIIDVIFILECEGYVLG
RQVIEVRVCACPGRDRTNEEFSSPNYILTKSCKSRNNESEKKRKSNDNNQDYTITTSNPVFYQILRQILVTLETQLPQ
AELILESSTSPISSVSSPDNVS

>THD27310.1 Cellular tumor antigen p53 [Fasciola hepatica]

MFASLTMDREVLSAVREPDRLLFLARQHSHTGAELSLIDNTLSTYTDNTVPYLITDPTVESGTTSTQTAQDLKQAYL
SLHHLQPQSQAHPIPHRHHHRLHQQTQQKPEQEREQQQSSTDSATGSSNNERARTGLVVTMDLLCAPPHTELSGI
KEPSQGTPIQSATQSTPVSPSVSSTPSARQSGFNVSPSAMTELDVSGPIRQDSGVDTSMPPKKLYSDDIPVLDSFPG
FYGFDLFRPLCGDTEAAETKPSTSTFFFKDEQGWNTLYSKKTPAWWTLSYWCQRQPPEGSFLRLTPVYGTSDKQQ
EVIQRCFEDFMAMPTNSLARYSIIVVGNSADYFYDPNTERLCVTLPYERPKGCEYSQFNGKFCFNLSCLHNGGQ
GNKKPLFLIVTLERLVAGTEPSKGQCEVLGRQCIFRSCACPKRDKENSERRTGDSICGSTQSGGKRKKDLERGDRHN
KKWCPQTQDFGATGTHKSFHLNGRWQSPGHelfahdgsnplGEDDNDSDPRHSRLFVDQDTGSSVIQYNGE
SYHMLLIPTGLPGGLETLAGVRFGLLRWLYAQSVMMKQSSPSTNRSTNVALKPEVLENTVTENENRDDPSQTYL
PGRNMEPDTVLHQLLTIESQLAHVIRERIYPTNGCHGPYPYGGESTDSQPNSNGAVVGLSLGSVDETAHANHEFST
AANVGPLSSDSGVSSDAIQHHTPLLMNGNFPVSLAHRPGLYPTSGSQTSTGYDPDYGNRFQNHLSVGQLYPPHL
MNNTQLRNSIGLSGYFPEITDNPVLTHTIPHSQATSPISDRTGASSNLNSVVVDHADAAKSTTTTTTSAQVRGS RTE

GVTLPTSLSAFTSTVFYPTSYFGSRPTDNSDQTNPNSTSLDPFSRGHLTTGAAVGMSTSTNGVSTGLCFYPAVGPTT
LSESLIHEEVISNYLDREKSYDETDGGHQSVLLTRNPIHRRCSSFGSSSPQSPGSSCCDQTTLLEV

>XP_018646355.1 cellular tumor antigen P53, putative [Schistosoma mansoni]

MYGTQHRFKISPNTYHDLENGNSLTTTSIGPRPDTDLDSAFPKKLYTDDIPVLFHFHKLKLYFKDEQGWNTNLYAKKTP
TWWTLSYWCQRKPPEGSFIRLTPVYGASDKQQEVIQRCFEDFMAIPTSGMGRYSIVMVGNSLAEFFDPITERLCV
TLPYETPKEGCEYSQFNGKFMCFNSCLYNGGQGNKKPLFLIITLERLITGSDPSKGHCEVLGRQCIKFRSCACPKRDKE
NNERRTGDSLVIPTSLSKRRKESKRERINLLNMYQQTYHFGSDSTDSQPNSTLSTGAMGFNGTGNLDEPVNINL
QDYFNSSNMNGSVGSDLMNSNLPRNLSNHQTNIMQNMAAYILSSTSTNTTTSTTTTATSINNIGNRGFIYNTNPA
VDTNWSYETDSDLRSGQTSHTDQLYESNHQGLIIGNNNNNSTLHVVSSELIDPYALVQHLPQSLSSSSPVHQHHHHI
SHHLHSAHVQSQHNLHQIHLHQGMNNSTSIDTNDNHSSAFTTTALISDNTMNNKSLTLNTCLLPQITPTTTATVTD
TGTIGTTNFGHDFNTSNFYTSAYFNLTNSNTNADNNSNGSNNVMDCNTNVNQSNDITDQNSTLIDSLRSSIPSNL
NHRISQNSTLCNYQLTNHTNLSESLLLSEELLHNCLDNKKKCYDEPDGVHSILT KHLRNTSFCLSSSTRSPESCGAEKFV
KLEI

>CDI97221.1 tumor protein p63 [Echinococcus multilocularis]

MSPGGGSGNGSGEMEGKMGFVVDHQPPPKLFSDDIPVLDPPFGYYGFDLYRPLCEDAYDSIETKPSTAYFFFKDD
QGWNTNLYAKKAPTWWTLSYWCQRQPPDGCIFRLTPVYGTSDKQQEVIQRCLEDFMVSPTPGLSRYSMILVGNSN
AEYFYDPTTERLCVTLPYERPKEGCEYSQFNGKFTCFNSCLYNGGQGNKKPLYLIITLERPIKNGSATEGPCEVLGRQC
LKFRSCACPKRDKDNNERRMAQDMGDGNLASNSVLSKRRREGVSQRSSRGSRGKNFPSTSDRVGTPQFFQQDE
ISSSSGYDDLVDLQIDGESFHLILMPAYLPGGAHTLRQMRSALLKIFQTARDCEDDRKISNSAASGYSGERQEELLQALN
SMEGRTASLLRDLQTSSTFVNPNNSCSTDSQPNSSSTQAFTCPTGGSSGALDPIGSSTLDLLAYDHLHSQHHKSENSFL
QPFSGQTNEYHRSRMNYPSTLTTTLESAHQQRVSSCFESNTGGTDLPTFNGISMDSAAAAYAMMPSAFQSFRRDS
DSIHYYQPNVITSSQNTVNSSVPSVVISQDNNGALVAAFGAAVSSSQSSGKLFEEIANSRQQQCHQTSQTDVSM
NSSCGTASPIPASSNKN

>TGZ61581.1 hypothetical protein CRM22_007927 [Opisthorchis felinus]

MGDTAHPKLTSDIPVLDSFPGYYGFDLYRPLCGDTYEASDTKPSTAFFFFKDEHGWTNLYSKKTPTWWTLSYW
CQRQPPEGSFIRLTPVYGTADKQQEVIQRCFEDFMSIPTSAMSRYIVLVGNLADYFFDPVTERLCVTLPYERPKEG
CEYSQFNGKFMCFNSCLYNGGQGNKKPLFLIITLERLVAGTEPSKGQCEVLGRQCIKFRSCACPKRDKENSERRTGEL
SNYHMSGFGKRKKDGDRTDRSGKKGRSHSIDFTTSTAKATRHLSHWGSPTTESHIHGSNMDDGFEVRHNADQY
GDAEAGQTVINYNRSYHLLVPTGLPGGIETLSGVRLGLLRWLYTSANSIMDKKNSGLISTGPANQTAIKPEVDTD
KETTLPANDSDDVYSSPVNCRASVDSALQLLVAEMQLAHLVLRERFVQNGQPYHHSRDSTDSQPNSDGAVGLGL
GSLDESVPNGQEFQACAVGPLSSDSGISTDLGQQHHHLSHPQQLGHRNSMGFTQQSHLISDMFSVAPTTHRGS
VLPGTAVNTSSAGSYEVDYASRFSAPNSHLYTSNQSSEHPADLRSSLGYAYPAELIETYALVQQASSSPGSTISSGHP
SLPAKPQSVSVRNAGNRQNGSSMSVPESVATASVNLNHAFTSSAFYPHTYFTARPNGSSDQTSLLDNFSRNQLA
ASVSGAQSNFCFHPVGVNGALAEITLLPEEVLSSYLERDKGYDEANGVTHTSIGRHPEST

>GAA29806.2 cellular tumor antigen p53 [Clonorchis sinensis]

MERVLSRPGSDPILAAHLHSLRPLQQSFDPDRISQLINGAQNNTVSGGGSEPLSGGTTTSFNVGERARTGLALAM
DLIRMPNVVDKTEHPTSVLSTINRTTTPVCSPVSTPGTYTRHSSGFNVSPNGYTDLESAGPVRADAMGDTAHP
KKLTSDIPVLDSFPGYYGFDLYRPLCGDTYEASDTKPSTAFFFFKDEHGWTNLYSKKTPTWWTLSYWCQRQPPEG
SFIRLTPVYGTADKQQEVIQRCFEDFMSIPTSAMSRYIVLVGNLADYFFDPVTERLCVTLPYERPKEGCEYSQFNGK
FMCFNSCLYNGGQGNKKPLFLIITLERLVAGTEPSKGQCEVLGRQCIKFRSCACPKRDKENSERRTGELGSYHMSGF
GKRKKDGDRTDRSGKKGRSHSIDFTTSTAKSTRHLTSRWGSPTTESHIHGSNMDDGFEVRHNADQYGDAEAGQTV

VNYNGRSYHLLLVP TGLPGGIETLSGVRLG LLRTWLYTSANSIMDKKNGGLINTGPTNQTAIKPEVDADKETTLPAN
DSDDVYSSPVNNGRASVDSALQQLVAEMQLAHVLRERFVQNQGPYHHSRDSTDSQPNSDGAVGLGLGSLDES VN
PGQEFSQACAVGPLSSDSGISTDLGQQHHHLSHPQQLGHRNSMGFTQQSHLISDMFSAAPTTHRG SILPGTAVNT
SSAGSYEVDYASRFSAPSNSHLYTSNQSS EHPADLRSSLGYAYPAELIETYALVQQASASP GSTISSGHPSIPAKPQSVS
VRNAGNRQNI GSSMSVPESVATALSVNLNHAFTSSAFYP THYFTARPNSSSDQTSSLLDNFSRNQLVASVSGAQSN
FCFHPSVGN GALAETLLLPEEVLSSYLERDKGYDEANGVTHTSIGH RHPYRRSPSYSSSPVSPNSACCEKVILDM

>VDM22031.1 unnamed protein product, partial [*Hydatigera taeniaeformis*]

MEALSKLSSGFHDFAAAVSGHTFTHDDVGALPTPPSILPVPD TTINSGDGSSVSGPLMNVKMSPGGGSGSGSGE
IEGKMSFVVDHQPPPKLFSDDIPVLD PFPGYGFDLYRPLCEDAYDSVETKPSTAYFFFKDDQGW TNLYAKKAPT
WWTLSYWCQRQP PDGCFIRLTPVYGTSDKQQEVIQR FMANQCLED F MVSPTPGLSRYSMILVGNSNAEYFYDPT
TERLCVTLPYERPK EGCEYSQFNGKFTCFNSCLYNGGQG NKPLYLIITLERPIKNGSATEGPCEVLGRQCLKFRSCAC
PKRDKDNNERRMAQDLGEGSLSSSSVLSSKRRREGVSQRSSRSSRSKNFPSANDSVGGTPQFFQQDEISSSSGYDD
LVQIDGESFHLILMPAYLP GGA

>RJW71291.1 Cellular tumor antigen p53 [*Clonorchis sinensis*]

MPPEGSTRAGILPGCSSLDREVERQRSGSNHGPCANLHSGLRPLQQSF PDRISQLINGAQNN TVSGGGSEPLSGGT
TTSFNVGERARTGLALAMDLIRMPNVVDKTEHPTSVLSTINRTTTPVCSPPVVSTPGTYTRHSSGFNVSPNGYTDLE
SAGPVRADAMGDTAHPPKKLTSD DIPVLDSFPGYGF DLYRPLCGDYEASDTKPSTAFFFFKDEHGW TNLYSKKTP
TWWTLSYWCQRQPPEG SFIRLTPVYGTADKQQEVIQR CFEDFMSIPTSAMSRY SIVLVGN SLADYFFDPVTERLCVT
LPYERPK EGCEYSQFNGKFMCFNSCLYNGGQG NKPLFLIITLERLVAGTEPSKGQCEVLGRQCIKFRSCACPKRDK E
NSERRTGELGSYHMSGFGKRKKDGDRTDRSGKKGRSHSIDFTTSIAKSTRHLTSRWGSPTTESHIHGSNMDDGFEV
RHNADQYGDAEAGQTVVNYNGRSYHLLLVP TGLPGGIETLSGVRLG LLRTWLYTSANSIMDKKNGGLINTGPTNQ
TAIKPEVDADKETTLPANDSDGVYSSPVNNGRASVDSALQQLVAEMQLAHVLRERFVQNQGPYHHSRDSTDSQPN
SDGAVGLGLGSLDES VNPGQEFSQACAVGPLSSDSGISTDLGQQHHHLSHPQQLGHRNSMGFTQQSHLISDMFS
AAPTTHRG SILPGTAVNTSSAGSYEVDYASRFSAPSNSHLYTSNQSS EHPADLRSSLGYAYPAELIETYALVQQASASP
GSTISSGHPSIPAKPQSVSVRNAGNRQNI GSSMSVPESVATALSVNLNHAFTSSAFYP THYFTARPNSSSDQTSSLLD
NFSRNQLVASVSGAQSNFCFHPSVGN GALAETLLLPEEVLSSYLERDKGYDEANGVTHTSIGH RHPYRRSPSYSSSPV
SPNSACCEKVILDM

>VDD76345.1 unnamed protein product [*Mesocostoides corti*]

MARQRM TLTAAVKPAHDSTIFGSQNA STSPALAPRANSSPDWSTASTRPRLP HSTMEALSKLVSPVLSCLVFHPN
GASSSLSLCTDARLSCGEAILERRRLVPTCWRCSPQLGFGSSVCRPPSSNGHTPLGFHDFAAAAA AVSAPAYHHHH
HHHHDAVGS LPTPPPLPLATDSVDVSGGGGGALLGGSGLLNVKMSPGGGSGGGGGELEGKMGLPEHQPPPKKL
FSDDIPVLD PFPGYGFDLYHPLCEDAYDSMETKPSTAYFFFKDDQGW TNLYAKKAPTWWTLSYWCQRQPPEGCL
IRLTPVYGTSDKQQEVIQR CLEDFMTSPTPGISRYSMILVGNSNAEYFYDPLTERLCVTLPYERPK EGCEYSQFNGKFT
CFNSCLYNGGQG NKPLYLIITLERPMKNGFVVFVSGTATEGPW EVLGRQCLKFRSCACPKRDKDNNERRMAQDL
NDGSLSGGLPSSKRRRDGASQRAARVNRSGFTPAGDRIGTPQFFQHEDVSSIAGYDDL VQIDGESFHLILMPAYL
PGGANTLRQIRSALLKV FQTARGGEEERKTHESVASGYSAERHEELLQALNSIEGRTATLLRDLQTSTGFINPNNCSS
DSQPNSSTSQTFPGPN GGSAGAI DLIGSSTLDLLAYDHLHSQHLLKEGSFLQPFSTQTNDYHRRMSYPSTGATSLD
THQQRVGGCFDGGSGGGDL SAFNGSMESSAAAYATMPSAFQSFRDN GGMHYQPNVFITTSQTTGHSAA SVAVS
HDTDGTIVSVFGGGGG STASQGGKLFNEVTDGRQQQQQHQTATSTATDVSMLESCRLLSAGVQC LILFLLSLN
REFRSLDYNVEVLNSAQRGYLV PGLPIMSRRFTPYRTGTHV

>XP_011405395.1 PREDICTED: cellular tumor antigen p53-like isoform X1 [Amphimedon queenslandica]

MEEPHSDPDSVIHSDEEDLPGSQESLGLSQGTNLNELMEAIKSNNGSITEEFYVCEESESSVTENSSSMDSITIQRSLTPA
SFMNSQQQQQQSQHMLPISSSYSYPQPPPPPLYHVYGDGSAVSSANTTPTRVTPPGDSPAAMSDCSGSYSMH
GMPFIRKKVPSNIPMPGEYGFTLILNDDDNSKPPKTPFTYSNLMKRAYIKRDSTVGMTFSFSKVPPNAVIRAMAI
HKSPDLIGDILQCCPKHIEDQKKRGHQFPHFICGAAKTETIYCEDPASGRSLITMPISSLQAKSLTSGSVQAFFVFPCF
TSELHKGPGQVAQLIFTLEIGGVLYGRAVVDIRVCASTGRDRDNDEQKQSTPSTIDPLVDLKKSRKPALTQNVECAI
PSKVSRTSLSAQSFTVETYPELFPFVEVCVRMMERVHQVNISEANLSQLNIGATATDVPDDPLPSASSQLHGSQ
DIRNVESWLQQLNLSQYKVLLDNGGICDLAQVISLDNNQLKDLGVSNSHREQLLNGISLINTAATQQPSREDETFT
YSQSTQPYKPQLLSQSSSTASGMSNIISPNTLEVLRIQVKKTYSSSSGSESN

>VDP80469.1 unnamed protein product [Echinostoma caproni]

MDREVL SAVSEPDRLLAHQHLHSTGAELSLNNGNSSSSASSSSSSSSSSSSASSASTSAAASYADRSLSYLITGTGTVSH
HKPQQQHLP LHS LQTQQT LAH PPPALPLYPQH PPPPPPLPLPSQSTQH HLSQSHH THHIQQQSH PQPDTDCATGV
STHERARTGLVLTMDLLCASHPDLSGIKEPQGTPVRSATQSTPVSSPSVASTPSARQSGFNVSPSALTELDVSGPMR
QDSGVDTSMPKKLYSDDIPVLDSPGYGFDLFRPLCGDTYEAETKPSTSF FFKDEQGWTLNLSKKT PAWWTL
SYWCQTQPPEG SFLRLTPVYGTSDKQQEVIQRCFEDFMAMPTNSLARYSIIVVGNSSADYFDPNTERLCVTLPYER
PKGEFTSHLTDY

>VDQ00547.1 unnamed protein product [Trichobilharzia regenti]

MDVKFSLVDDDKENSSSELDVEEVIQNSSIITKNSLQKVYVTFMRAWPINFVFRPPITPHDFLIRITPMFTYAERCNEV
VQRCKKH EAQTLSE DQALMGKSFVEILGVNCQYMIGNTIPNCPTGQISVCTSDVDLQLLHQSKPEPPSSSVDS DISQ
SLLPEVKGLLHCR LSCYNSCFGSTSRGEVDLIITLEAKSND SHVGEETYKVFGLSKISVRCCSPARDAKSDCKQSDSPD
KNSRRLSLNLSSQSNIRLKN DSSSIRRSCPSKI QKVENINQIDSCSSVIIDGKKYYITVVS LI

>XM_001745968.1_Monos

MASQDTLPASQETVPDAETGVFPMLTRHGSSYKRMSMRDFHEY
LSRINKDGDMA SGSMVRRDCMMHKDGWQQVDAEFDVDIDILGGQLPMDTELVDLETPL
PGTGLDDTPQASAKPCGALEDSGEPNPAGFRANLADSSVAAGPGARAIGWTYSPILNT
LFTPMDYSCPIRFATNESVPDLSRIVAHLEYTQTNRNFVVRCDMHRQGDSGPFAEH
VLRVNNPQANYHQRQERLAVSVPVASTRSGKVEQNELFEWHCLTSCAGGINRRKIRVV
FRLIDPDQNV LGVQHINVRVCACPVRDRRTHEKAHAKREQANDKRNVIASDPTPSKKP
KLKDHPISMALRKMPSVSTLCQIGSGFDHDQDVQYLAICGRANYEVALKVRDQLNQSA
GLGRVAASSVHLPRSEVTLIKQ RVAEATNLDEFLRACNLQSLKRTLENLG YTDLPLIA
TLDAREIEELTLTKQDQRR LKLALVTLNETDKAQDPYLARRRET KLMRVTVRSTPGQA
QAGGAGPTSPFAAAS

>XM_001747604.1_Monos

MEMSLFDDVDGEVSSSSSSSSSSSSSSSSSRFPPLATNMNA
LTMIDWATPILSPLLSFLDSIMTSVVP GSVADGNERGPGQFQVAVETITLQHSDSDLM
VDIPGYGTTAVKRAPMTERDFEMLPSSFRGPYDLQLDLIDENSLKPRNTSPWTVSTPG
LDLTLQNGSAPIISSNSLHVPCRWP HFDVLLMIQYSPQLGRFFVHV NADVVLKIVL
ARAPPKGTDLVFR LRYALPEHRKTRVETCVTHQQAGSHFFGAPHNHLMSINREHV TYD
TDSTGHHYARVALDQFPFTDNVYSVPLRFHCFSSCPGSIARRMMQLMVYLEHSEHILG

ITSVDCRCCACPGRDRLSAEKRQQRNSAKRHPAASGTKSKRTRRSNGDNMALASSQESN
EGESSQPSEDL SIHILPIHGKTHYRIIEMIEGMSSARILTPNQTA VTAKANLKLQQR
LREHSQSHSNPPNLELLQRELSTGVSESPSLHKASARAMA IKEEPSAFKADGPTSGNN
SAVEAGDQGSTVPTMDDQQSGDTEHTPHLDLVPEADALYRNDWKFLQVGTARAHEPGQARTW

>XM_004991339.1_Salp

MQAGARREDSLGRDIKVDIEDLLSTREASTIGISSFDVPFDQLE
LDRTCDFLIETVQLNDGFEDPTPVVDLP GKGRVKLPRVPTTSAELEHTPSKYPGPFNF
QLQFIDDSGGKPVPRATPFTYSPQAQRLFVQMNHKVPFRFSISDRVPNGTWLRI RTF
TQPEFRNRQD

>XM_004991340.1_Salp

MEVTTLVTHRADGMHAYHLVRCDMPEAVYEKDATGHYSIRIPFD
RCAAQNGTHVFYIELLCFSSCPGGLNRRPF ETVSLEHSSRILGCDSIDTRVCACPGR
DRGAAEKRLRGDIGKKRGRKKKKPPT EPTVSMQDDNKIHTLLVQGRTHYAILHHLVEG
LDATVASSSSPPSSSSSSAAA AVLSSSAARNQAPSADKPLSEASFPAGND AEETPGGD
TEPESEADDND EEEDEQQLTPRQQV KYAFLNRLQEAVRAWLSALRLEAYADM FMS
KGYDDL DVVCDLTQDDLDCLGISRPGHRKKLLAAQKLKSL LHQARAVEAVPDRRLHK
TLSISRTTYRQALPSAD

>XM_004994533.1_Salp

MSGHARLSRQASFANLLRGSSTESISTWISSLGRAPSLTRCSSS
FIRHVTPDASAPLSWAESTT NIEIAASDVPGFSEEFNLNTRSPTRTGLVAAQQQERQQ
QQQQEGEPQHYQQHHHQQQQQQQQQQQQQQQHHEQQSSALTNLNMEADTRGSAKASS
WMYSGTLHQLFTNIDQVFPVQITSS ESVPESAKLHVQLQFSKDP SHPIPRCPHLAED
TAPHAEHVLQTTMPTADYHMDTSGRAYLVVPIERSKDEGLLCQTL LLMRCFSSCSGG
MNRRRPVDIVFTLINGAQLLAQVSVAVRVCA CPARDCRNMERQQAKRAAKSHQGKDVRH
PAKKGKVRGITAICKRQDADLRDSDTFSLNIKGRANYHIIRRIAEALKTNQVNPWNR
SAGAIVPDNIHDVASVPEFLARLGLTQYQPLFMGSGFDDLEVLHLDLNDLSAIGVE
DGIHQDTLLYAAAKLARSQQHHGEGCRPRGRCTSLHSPDVVDTHDHNNSTISSNCH
VFPDTSQQMQYTTQPHFPPMQLAPSN GVPHSLPSAAPSVAHAAKPSSGPGTSTSSAVA
TTAPVQQQQQQQQQQQQQQQQQQQQQQQQQQQQHSGPLIHLVPSSSSA

>Caps

MDMVRLTRGQSFQQQQQQQKSASQQNASGDGAAAAGAGAGAGAGAAAGADADADSKGSHDPADLNSTLLL
GDSQFGETTGTSLLNMSGRLGFNTPSSSSSHSHSHSHSNSSNSSFLGRPRTL PVQKPLSILQGSAP EPF
VEPLTIGEP IYASPNLGEAGFLLSVDVSNARHSAISSAYSSEALGTLFTNFDVGVFPVFRVAKPPPTV
HPLHIRATLRYKQM QFMKEPVRRCP LHLSIDSDLHLLRACDQDTVYSVDYHGRASIAVPFTPTMQPLVPV
LINTLKDAHVPLVTRHPSSTHGEASRNPYVCWSMTWFLKFCYTSCTGGMNRRATEIVFTLED SQGLIY
GAQALDFRTCASPSRDRKQLEKSPPIPVGSIIPPEVTAHQGSALADPGTTASTPTRRSAQPVVPSPVV
AASPTPAPGLLSSGRKAGAIPTGVTPRPLVVIPAASSAQTA PTAPAAAIAAPSPALTEAHAAPTLLRD
RSITAIARRRHADDDDEIFTISIRGRDNFVLVSKIVEGLNALAQRTSA AVASKLERTTTPPVVEDDPSAP
PALSALDTLAQAAAGHV RAPDEAAPATHERKRARLADDDTKAGRATEASGESISTWLGSFGLKQYAVNFA
KNGYTNT PQILNLCEDDLKSIVRVKAHRVRIAAAIDELADAQAEPTSTPPMSSDGDVTVVRMTIRRTSS
TAK

>Sphae

MYNLPLQDHAPAAEPSTSDPQLNIQPELTPPPRQTTPVLDSSHTPMPGIYNFRVRLGHFDSTNPTGKLT
GSTTYSLLDRLFINMGVYTPITLFSHTPPPEGAVVRCFVVFSSQEKCTTTVTRCPNHRSATDHSNTGND
PQKAPLTHILQVSHPPAQYLTPSGVESAVVPYNARESSVKNSAAKKERTAGFGAFQDTSIRYMCYSSC
AGSINRENIHTVFRLELGSILGVASLDTKICSCPSRDRTKAEKLTRQQEASGTTGSSRKRSEKGRDSPS
HSTDIASTGATKDSHTSERIQAPSIVHPHSSQATLVGKRSQRLIQNVQHKLPTPRINTADHTAHMPMRWST
AQSTRIATADSTVGIDEDAVYVAVRGRQNYEVLRSVADGLNALAQRPFRARSFRQHNPLPTRTPSLLVT
SNPPPPPTTTANTAAPTNTNTNTNTNTSTDTKTVFSRPGQNSTIPSLSSRRNPSLVSTTSVPPAQPRTA
ASSRSLRTVPSRERRVSGTVPSDVPVSESPARNT

>Creol

MTESESADMTGKLGQPSMSFSQIFNSFAENGSTVVGEHNEVGQEVMFVSMRRSTSNAQATMATRDYNTNAQ
QTFQPDGQATNAGYYASEPLEDNGLYSYGRIEEDGGYQYIMNPMNNPLGILPATLPSYLDNVASASYDAHGSLANLIG
LPFVNDGRAADSNAQGIQSHPELVPPAAQSPPLPLDSSHTPVPPIYGFVNVGHLDSNPTGKLNSTTYSKLLDRL
FININVYTPITIATNKQPPPGAMVRCFVVFSSQEKCRDIVTRCPNHRSPDPTNQSVANEVPLSHMLHVSHPKALY
CPSNTGIESVLIPYRPTISFAKASKIKRPSGTTDSTDDAPTLEHTYVRFMCFSSCPGSINRENIHAVFRLELGDEILGV
ASLDTKICSCPSRDKAKAEKLARAHVGMDSRNSNAKSMKHEVRGSDGAGTVKRRSPITHSVSSASTPSNQKGFVGP
ASVLTRSASQRLQENNMQTRTKGQADSSMYKSNIHSHPNGININGSTKLMARGREDSNRLGSGVDNGLEHTNGG
GSTRYYNVSAMLNSNTREPGGHNKFDDELYYVTVRGRKNEYILSRVAEGLTCTSQKPQRALSYYRRHPGAPAKVQSI
VVSNHSSKEPVKPISSERARATAIPRVVTSSITSPHIPTSQGDITTPKRSKRLVSSEAKIQSPARRAIGKGVDTVNGSE
VRQNNPKSRVNGQNKPATRQGNKTHRAVQNNADGSSTEEPELADRSIIPPLTSKSRISVPTRSHQTRRSSARLN

>Ichth

MHNLHNLPRQNTSALGKTKDNESNRTGRENQLQHDTLELSLMDVDIDLSNNMFGLPSFPNSGLTNY
DLTSTTNSVDMLEYLDPILTSPVLDPSWALPQAQHQAPQSQQLPLPMLTRSQSNAEFEPNFEOPYH
FNISFVTKTESGGKSDAKGATWLYGSQRDLVYTMNTTVPINLTIEGNPPADANVRVKVYANPEDYMER
VVRCPHADRKEQATKPFDPQIRVDHLAATNELDRLNHYSVLIPYSKLKSRPQLRLRYLCYSSCVASIN
KRPVSTVILEDASGRLLGQRSVSTKICACPRDLKKDNPIAEKDKWKNNRNVKGQQTNDIDPHPSRTT
FVNPKEILQSTKSIVQDNNNMVTVNRTIRGRANFEMINKIIDSLNMDNLQNDINPNHSSVDPPALHR
QFNLRDMCPTLALIPPIIQNHKTKEQRMVSREKTPTIESDEEEIVRDTKPTKKFKQGSFPSPGPSLQK
TNSFSRMQ

>Chromos

MSFPNQNRVPGSEQTWLENLVMRFPSEGGGNEFMLNNRHHHPNKENEMLDEVEFKFEAVNCFSAVPDTF
QAGGSSNSNMSGVGDKDSSCHSDSVDRYIQGHLQPVGNHQSNNSNNNRCRAPLVPTNSLGRESTSSNTTA
PEILGRNYVRGALSIEFPHNVKGESGAHDKLSLDAGRTAAMDASARSSPLLQSPKAELSSLVDENALQH
PPETSNISTCPSEGRFARLEQVLDDVAVDGTSAQHSDVYDGNAQVREQIPFDVCFEGSSLQSKATTWGY
STLNKVLNLDLTCPIPLSIEQENIPRGFILRATARFTKPQHYNETVVRCPNHEQKDPNFKSSHDKKIVP
TSHVVRCEPNSAVYETDAASGHHSVYIPIHTRTKSGSSDLSSFADDSKANDNGNDDKPKKTNIIVPFKFM
CFSSCPGGINRRPFELQLTIEENGVVFATRSGIRVCACPLRDKRSEELTFMQKNKRLVDCEKEEESTG
AHGSDEKSSKWLVSRRKTNIPSSLSGNSLASLSKTVPTIDKHGNEIFNIKVKGRQNYDIVARVVEGLN
ALYGDITHGRQNAEKKAIGSKPGDYSGTGTDAAESDEEESEMGRKGQEDKLMTSYLEVLSKYSPEAL
KCMPPVREFLKAVALDYYQSFKDNQYDDIQLVTCIDKSDLDILNISLGGHRKKLVAAARALFSRLKGPVE
TTRRRGSIVYTQTTKRGSTMKKTSATSSQPSGSIFPADFPNISSLGGKDDGSL

>Corall

MSFQQRGPMKAPPPPLDLQGSGANRPELMRQQSQDGAERAITALAQVAFQESVALVEGRIQFVDRNQVLN
QQTNNGPGNANGEEAKPQSTEPKSDSSALLLDVPMADIGLLLNKEQEEELKKEVKRKVTYFDVDIPTHKP
FGAFSKKLEKLFALPHVPYLFsfYTPITEGSIQIKLKFNPPYYQHEPVERCEQCSKKEEPQPDAPKHFVT
CTADERGESNAEHTITDDGPIVTMRFNLRQEVAPRVNTGSGNVHCVQRNLKFGCLNSHLGGKNKRDLEIN
LSLLSTEGVQLASKTIDLVCKRPQRDADLEEQGEAMESVDAMGMYGNVPQGGQRPPEEKDKSKEDWHLI
KVRDFEDWKLLNQMAEALSFFREHVHNAYSQKLLSGQKGEDQDIGSKRGPPDADGPAAKLRKLSPEENQR
TLIEWLERNGLKQLQPVLASNNVHSFDDLWSICASEQVPFLPGDVKQKMKTILEKDNPRADAVRDAFENA
TRLLKMEITKVFDLNDSSQGNAFKGLQRNAVPTPAAPLSAILPSPARFTTQNADWLLQAGSLDELNVRF
NPFSPPLITPGLQAELLSGQNHFFGQNSLDMIHQQQQTNERLGIGLNPTHVQQHQEQPTPDIF