

***** globins *****

>Pdu_Egb_A1a MNGITVFLILAMASASLADDCTQLDMIKVKHQWAEVYGVESNRQEFGLAVFKRFFVIHPD
RSLFVNVHGDNVYSPEFQAHVARVLAVGVDILISSMDQEAIFKAAQKHAYADFHKSKEFGEVPLVEFGTAMRDVLP
KYVGLRNYDNDSWSRCYAYITSKVE >Pdu_Egb_A1b MKGLLVFLVLASVSASLASECSSLDKIKVKNQWA
RIHGSPSNRKAFGTAVFKRFFEDHPDRSLFANVNGNDIYSADFQAHVQRVFVGGDLILIVSLDQDDLFTAASH
YSEFHKKLGDVPPFAEFGVAFDLTSLDFLPLRDYNDPWSRCYNYIIS >Pdu_Egb_A1c MNTVTVVLVLLG
CIASAMTGCNTLQRTKVKYQWSIVYGATDNRQAFGLTVWRDFFGLYPDRSLFSGVRGENIYSPEFRAHVVRV
FAGFDILISLLDQEDILNSALAHYAAFHKQFPSIPFKEFGVVLEALAKTIPEQFDQDAWSQCYAVIVAGVTA
>Pdu_Egb_A1d_alpha MYQILSVAVLVLSCLALGTLGEEVCGPLERIKVQHQQWVSVYGADHDLKRVSTL
VWKDFFEHHPERARFERVNSDNIFSGDFRAHMVRVFAGFDLLIGVLNEEEIFKSAMIHYTKMHNLDLGVTTET
IKEFGKSIARVLPPEFMDGKPDITAWRPCFNLIAAGVSE
>Pdu_Egb_A1d_beta MYFSYFTAAASYLSVAVLVLSCLVQGILGEEVCGPLEKIKVQHQQWASAYRGDHD
RLKMS TLVWKDFFAHNPEERARFERVHSDDIYSGDFRAHMVRVFAGFDLLIGALNQEDI FR SAMIHYTKMHKK
LGVTYEIGIEFGKSIGRVLPEFIDGKLDITAWRPCYKLIATGVDE
>Pdu_Egb_A1d_gamma MYLSVAVLVLSCLALGTQGEEVCGPLEKIKVQHQQWASAYRGDHDRLKMS TLVW
KDFFAHHPERARFERVHSDDIYSGDFRAHMVRVFAGFDLLIGVLNQDEIFKSAMIHYTKMHNLDLGVKTEIVL
EFGKSIARVLPDFIDGKPDITAWRPCFKLIAAGVSE >Pdu_Egb_A2 MNLVILVGLLCLGLTSATKCGPL
QRLKVKQQWAKAYGVGHERLELGLIALWKSIFAQDPESRSIFTRVHGDDVRHPAFEAHARVFNQFDRIISSLT
DEDVLQAQLAHLKAQHIKLGISAHHFKLMRTGLSYVLPQAQLGRCFDKEAWGSCWDEVIIYPGIKSL
>Pdu_Egb_B1 MLVLAVFVAALGLAADQCCSIEDRNEVQALWQSIWSAENTGKRTIIGHQIFEEELFDINP
GTKDLFKRVNVEDTSSPEFEAHVLRVMNGLDTLIGVLDDPATGYSLITHLAEQHKAREGFKPSYFKDIGVALK
RVLPQVASCNFPEAWDHCNFGFVEAITNKMNAL
>Pdu_Egb_B2 MLVLVLSLAFVLSALAECCSAAADRKTVLRDQSVWSAEFTGRRVAIGTAIFEELFAIDA
GAKDVFKNVAVDKPESAEWAAHVIRVINGLDLAINLLEDPRALKEELLHLAKQHRERDGVKAVYFDEIGRALL
KVLVQVSSNFNSGAWSRFCFSRIADVKS DLP
>Pdu_gb_IIA MGCSATKTLTLKKGSVLTEITKKARGDRVGS PNDQDTNGDAIQEKEELDPRLPMTARQQY
GITKSWKAISRNMKTSTINMFVRLFETNTEVWDLFATFKDLETVSDLRENRSLENHAMMVMCTIDETITNLDD
LDYVIDMLQVRGKTHTRFQKFRADLFWKIEEPFLLAIKETLGDRTVSMEFYTKRTIRFIVENLSQGFKKQD
>Pdu_gb_IIB MGCSASEHIYPAADGMLAKPGFKDSLKNGIKLPLTEREIFALTKTWKAI SKNMVSAGVT
MFLRMFESRQEVRTMFEQFRITDDVASLYTSSALENHSLIVMNALDESMNMDDEEYLI EFLLTGRSHQRFE
NFSETVFWAIEEPPFITAVRQTLGDKFTTGLERIRYLAIKFILTTLIIGFNQRDRGESNTY SNGY
>Pdu_gb_IIC MGCRQKQKSAANGPENLDTIPEPSTPPPVDPRLPINARQVFKLKKSWKGIKRNMEATGV
EMFIRLFRNNNELISLFDKFDITSVDDITRENEALEHALLVMSVIDEAI CHIDEVDRVIELCTRVGATHSR
FTGFTSDLFWNMEEPFLASVKLILGDRYS DSDMVIYQLTIKFILSNLTGFKFAASS >Pdu_gb_IID MIMG
CGCSQSSTLKI SRKVADISVSNEVLEKRGQARARSVASTISTALGLDKTSIKMPLTEREKHTLMRTWEHMH
IVEHGVTMFIKMFETSPOVKS VFEKFNREGENSSDLYNSDVLKIHGLSVMSAIDDIANFDDKDVALELIINQ
QSHAFFGFQDDMAADI FWAIEAPFLHAVKETLHGQYGEKVS KIYEKTI RFI LNSFVQAFKAGKAQKALLEKNQ
NDKRLSVATAFSSSVSRHSSIAAVY
>Pdu_gb_IIIA MLQYLMQGYRKMGNRVSQSPEERRRSVLCKKDVTMSKSSLEEQCIENPHKPQIKDV PRA
DRPPFTEEQKTLVRKSWKVLQEDMSRVGVVMFIFLFEHPDVQDLFLPFRNLTTADMKHNAQLKTHALKVMGT
VEKALARLDEPKKLEDMLHSLGRRHSTYNIKPEYVDLRLN >Pdu_gb_IIB MGAHSSTPKSKHRPKVKEK
DRKLSLEVEENISGSSQKEAPPSTSSPLEATRGLHQGEWRGDLKNSDSINANDIEFTPSPLMIRQAKISELKD
NRPSVSNEEKQLILDTRWRIIEHTANLGFQMFSTSLFETYPTDKGAFSIFRALSQDDAQFEMELRMHGTRVMTT
VREVLERIDDLGCVVEHLHELGRKHVIFNAKADYIDLIGPQFLFAIEPLLGEHWTPEVESAWANLFKVMGYIM
REAMIL
>Pdu_gb_IVA MGCSASLDLMGAPKGIPTTRDRVFEEEGAQSDAILGWKLPPEQKKGHASRFMQAVGAVID
NVDDLNTIYPVLTNLGKRHVTFEDFEGHFFDAFEQAMLLVWSEELGTFEERVRDAWQLVFKYIIRTLKAGY
DGEKELMNSAKDKNTSAQTSQLSLQ >Pdu_gb_IVB MGGVCSSSNQSVASGQDSNHKMPTNCKEEGQLLS
AKQLKVVQETWAMLDKDMSGRGIRVFQOIFTLAPETKGLFNFRYIPDELHNNLLFKAHAGRFMQIICAVMDN
ANDLDTQVKPTLHNLGQAHVVKFGVNMEYFDLFKTAMLSVWSKDLGKAFTVEAREAWTIVFEYVLEEMKVG
SANHNHNSNIVMNNTESASTQMADEKQVSMEMADNGNTS >Pdu_Sp_gb MGCSTSHPKSPPKNQSLGDCLTE
KQMKLVRETWKGVDDDLGEGTKMVFKRFLFELEKGFMKLFLRSGSMGDIKQKDIEMSEERLGRHVTIVMQALGA
AVASLDDSRYLTSVLQNLGEMHTNYRVTDMLPKLWPAIDHALKTKLGDYSTEARES WRVVFVFFICKMREG
MEAASKVSS >Pdu_Ngb MFTVEPEVMKQFSFVPGVNTNPEELKSSARFLRHAKNLIATVSNVNDLDDMED
LSKTLNNLGRRHKYKTKTEYFPIVGRSLTHAISTATGDAFTPETAAAFSQFFAMITFYTNEGLMEEA
>Cte_98019 e_gw1_817_19_1

MGICSTHLRPVVVPESNELFTDRQKAIITKTWRHMGNDLTGRGSKVFLKIFNLHPEVKQL
FPSLKNEDQLLKNPCFRGHASRFMQSVGAVVENLDTPGDLSPLLDLGRKHVLFGGFT
PEYFAAFTEGMMCIWSEELGKGFTDEVSVAWKTVFDFIMSQIQDGYAKASADSTVSSNRE
*

>Cte_147415 e_gw1_637_4_1
MFEERSDVKSMFEQFRTVEKQDLGTSQSLENHALLVMNALDEAIANMDDPEYLIEMLLTT
GKSHRRFDDFVPSDFWAIEEPFIQATRESLGDRFTNHMDGIYRKTIRFVLEVLIFGLKNN
FEESNTENALPAFGSQR*

>Cte_188542 fgenesh1_pg_C_scaffold_271000004
MGTEISSVDCGQTSADLSITRPPVNQEVSEEESSPPSPWQPNTPPSNWKRLWSSLRRTN
SRKTKSVDSDEEQGPTSLEIRRAHLQQLAEARPQLTDRQMTLIEDTWSIVQRDISTVGLD
MFSRMFEAFPGIKESFGPLSSMVPEDHRYRKEISEHGVRVLTVDITLRLRHDPDKTIET
LHDLGGKHISFNAKVYDIDLLGQHFLFAMEPVLKQHWTPVEVEQAWADLFRMLSHVMKEGM
VL*

>Cte_144794 e_gw1_580_36_1
KLSAEHKTTIRDTWPLISHSLQDNGIVVFEKIFEVSPSIRTVFAASFGFPASPIPDAYEL
SRASNLRDHSVTRFMQAVGWSVQHMDLDTVTTVFVNLGKRHIHLKSLEPDDFRVFSGALM
YVWRSTIGPDLFTAVERGAWCKLFEFMLQHLAHGYNQAKQA*

>Cte_166020 estExt_Genewise1_C_17170004
MRLVQMSWDIVQEDLAALGHGAFDRLFMDHPEIKDAFGPFKELSKDNIHFDRELRLHGVR
VLRTVETVLDRCYDCVRLVRLHNLGRKHVNYRANADYFEIVGRQFILVIASVVGDKWTP
EVEESWSHLFTFVAYVMREAMLLNSLSNP*

>Cte_110047 e_gw1_303_7_1
MTAETSLSNCGPFDQSVFLTDKDKLLIRRNWKHLACNLTERGARVFLRIFVDNPSVKDL
FPFKKLQGEELTRDVNFRGHASRFMQAVGAVVDNIDDFEQSLAPLLNGLGRKHIDFHGFT
PTHFNAFQDAMLAVWSEDLGGKLTPEGRDAWIKVFGFIMRELKKGYSQAEGERVKDDPER
I*

>Cte_194549 fgenesh1_pg_C_scaffold_307000005
MFETNRQVKEAFEKFRSMDTPSELWQSSVLETHGMVVMNAIDEIICNFDDRETVELIIE
QGRSHVRFGLDTEDLFWSIEEAFHVKETLDKNYPQHLQVIYKKAIRFIINLLVIGFKS
AMREANKGYCVPEGPLFGNHGAP*

>Cte_27881 gw1_714_5_1
MFEENHDVQYYFCKFAKLET'SADLRSSRQLRAHALQVMETLDDAISNLDDIDYVINMLKA
VASTHVNKFDASNLQIFWVIRDPFLLAIKESLGDRFSLSEIATYRICIGFILDMMVKG

>Cte_227018 estExt_fgenesh1_pg_C_5830012
MGNTSSHKQHQSLEYDQNKAVDVIIIRDDHLLSPHSDADDHSSLTVPVALRRQHMEEV TANR
PIPSEESFQCAEITWAILSENDRDGLGTEVFVRMFESYPDLKSAFGPLRHMNKKDAGYEDV
LRAHGIRVLSIVEQVLSKRHNMEEVLSILHDLGRKHLTFSKAVEYIDIVSQMFLFAIESA
LKEKWNNSTEKSWGEIIRFVTYVMKETMVL*

>Cte_198978 fgenesh1_pg_C_scaffold_681000004
MGNAQPFVSCAGGQPPDQKPSNIPQHEFLTQNKLGAKSTWEFLCHTSTPTERGMRVFLR
IFEIAPVTKTLFPFKDMPNEDLHRNSLFKGHATRFMKSVEFTMQNLDALDVIVNPTLVS I
GNKHVHIKGFHPDYLDTFQTALIDIWDEDLGGKFSKETKEAWIKIFALITRKVFEGFQEE
TTRFRPPLPYEGKQNGVQSPEISANNNSCVSPDEMECELVNGDSNVVTLDA*

>Cte_192439 fgenesh1_pg_C_scaffold_120000029
MHHHYKKAANNWLP RRDMTDKDLKNL FVNLPNEDVNQSSVHLD
ELLLKKHAAHVMEALGAAVECLEDSVFLSNVLVALGQIHATYHVKPMYLPRLWPAIRHGL
KEVLQDVFTQVEDCWRIIFNFIISKMKEGIRIEKHHHTQAHHHQQQQQHORAQQNS
TATLHTEGASDGAATKASQNRRESLFSVRENMSLVGSKRRASRVAPLVDTGQNNKDGA VL
EGATSGV*

>Cte_181579 estExt_Genewise1Plus_C_270128
MVDLTEQQKTLIANGWAVIKKHLKQNGVDAFIMLFTMKPAYQDYFSTFKGMSMEEIALSG
KMAHAGSMFMGALAGI IENLDDLECAAEILRSKVQSHEKRNIGRQHFHDL LYAVLPAFLA
QKLGNEFTDEAGAAWMTAVDILMSVIDDELAKMAQLSAA*

>Cte_214116 fgenesh1_pg_C_scaffold_6000096
MGCLQGKEKEAEAVVTELPESKPKQVDPRLPFD TYRQLFNLRNSWKT VARNLTETSKDC
LIRYLKYPHKAMYRGCANLEDEEAMRASTSFENAAVPVFNLFDDVMENSENV DIAIAQ
IGMGATPHKKIKGFR TDYKAMEEPFIEAVSVTLGDRFTDPTEQNFRKLYQFVIQEMIKA

LGGETPPEEPQVEADKESLPTDQEDVQIKL*
>Cte_218767 estExt_fgenes1_pg_C_890010
MVDLTEQQKTLIANGWAVIKKLNKQNGADAFILLFTMKPAYQDYFSTFKGMSMEEIALSG
KMPRAHGSFMFGALAGI IENLDDLECAAEILRSKVQSHEKRNIGRQHFHDLLYVVFPAFLA
QKLGKDFTDDEAGAAWMTAVGILMSVIDDELAKMEQLSVA*
>Cte_190593 fgenes1_pg_C_scaffold_325000008
MGCGSSKTVALTNSDHVYTYTASYSESRTTRDADQLLTPEEIVLVRVTWEQLKTNLTLANL
GKKVFLRIFNLKPDIKKLFPSVWGDLLIRHPKFLHSERFMLVVDCCVQNLECIKSEH
GEMLANLGRAHVNYKGFSSRENFEVFMKAIWYVWYHQLKDSMDSEVECAWKKLLLF IIVQQ
RAGYDAEKEAPPNGLS*
>Cte_195701 fgenes1_pg_C_scaffold_9000065
MTITRGFRFCGYQLLQDADFASAMGRLKYTMGCNFSKPKGFRVRVVRKQEDASARRIPQR
KHRELLKTEQVALLKSSWQQLCVKRSFYFLGRQIFLRFVFNPEIKKSFQFGEFHGNDLI
NNPMFKIHVKNFVSVIDSSIRSVDLSKTVLAPTLLHTLGGTHQSVEGFNKNLEIFLKAML
LVLRQEFKSALDVEDLEVEVAWRKLLFIVYQIHIGYRSAISTTPNKQKGFQAQESP*
>Cte_188798 fgenes1_pg_C_scaffold_6191000001
MGRLKYTMGCNFSKPKGFRVRVVRKQEDASARRIPQRKHRELLKTEQVALLKSSWQQLCV
KRSPYFLGRQIFLRFVFNPEIKKSFQFGEFHGNDLNNPMFKIHVKNFVSVIDSSIRSV
DSLKTVLAPTLLHTLGGTHQSVEGFNKNLEIFLKAMLLVLRQEFKSALDVEDLEVEVAWR
KLLFIVYQIHIGYRSAISTTPNKQKGFQAQESP*
>Cte_200756 fgenes1_pg_C_scaffold_608000004
MGCSSSAEVKVLNAPTPLPVEPPPTKPKGHSTFLTDEEVEILKASWNDLNDSSDLSSIG
KRVFLQAFEMRPEMKKIFPFDCWGDKLLQHPKFQAHQSFMVIIENSVEQVDNESSDFS
DSLTLGQSHSDRIGFTRENVQVFLKAILAVWHDLKSSDDRTEKIWSKFLAHVQIMRN
GYEDAIDEETTSKIPD*
>Cte_229341 estExt_fgenes1_pg_C_6370009
MVLSSAQAKAVQKNWANVKAHAQKYGNLDFVQYLTQNPQDIQIFAKFRDANLGDLSNAE
FNKQTKTVIDALSKIVDNLGDLNAGCSFLRERVRTHHPRGISMAQFERLLDLMPMFLQE
AGADGVVADAWRILVADLMPMRDEFKCSQ*
>Cte_219237 estExt_fgenes1_pg_C_1600024
MVLSSAQAKAIQANWKTVRARIQDYGNLDFVRYFIQNPQDVEYFGKFCDIPLTEVRSNPE
FQKQTKTVLEAVGKIVDNLDELQTAANYLRERVRTHHPRRISMAQFERLLDLMPMFLQEV
AGADGATADAWRILVADLMPMRDEFARCSH*
>Cte_60954 gw1_19692_1
LSEQTIQIVKATAPVLAEHGETITRHFYQQMFSQNPPELLNIFNATNOKTGRQQAALANAV
YAYAANIDNLGVLTAAVQRIAHKHASFNILPEQYPIVGKHLLEAVAHVLDAAATDEILNA
WKEAYGFLAALFIEVEEGIYRESELAEGGWRGTREFRVNVKQVESELITSF
>Cte_193618 fgenes1_pg_C_scaffold_124000014
MGCSNGRPMGRRSDSTVPMVDPRLPFTQAQIKTIRSVWNLVVRKRFEDSARENLVIFFH
LNPIFQDLFPHLNDLKSEQDMRHSFVMMQALAI FGVYDDVIEGLTHDVGAIKLEEVG
RLHAKIDAFHVDFQFMEEPFIQ
TLRENFADDFQDAEIELYRLLFHWMQAVMTTSIETGNFGCEDPSAVFCLSNIYLYTFLAH
TQEARGLLTRQ*
>Cte_21023 estExt_fgenes1_kg_C_6850003
MGLTTAQRAAIQNNWATVSNANMQEFGDALFMRYLLANPGDIQFFPKFQSAGVGAQLRSNE
AFQEQTTLTVFQFLGQIVAKLGDLDAGKMLQERVRVSHKPRGITMAQFERLLDLLPRFLQE
NAGANGPCADAWRVAIANLMPYMRDQFAK*
>Cte_5226 fgenes1_pm_C_scaffold_95000005
MGLTIGQRSIIQNNWVTVASNIQEYGDDELFSRYLSANPGDIDFFPQFVGDGEIFNFADLR
SKPEFQDHTLTVMLFLSKIVACLTEIEVAGSLLQERVRTHFGRGISMAQFERMLDLMPRF
LQETAGANGQCADAWRVAIATLLPFMRQEFSSRCQAQ*
>Cte_225570 estExt_fgenes1_pg_C_3920012
MGLTTAQRAAIQNNWATVSNANMQEFGDALFMRYLLANPGDIQFFPKFQSAGVGAQLRSNE
AFQEQTTLTVFQFLGQIVAKLADLDAGKMLQERVRTHKPRGITMAQFERLLDLLPRFLQE
NAGANGPCADAWRVAIANLMPYMRDQFAK*
>Cte_1470 fgenes1_pm_C_scaffold_396000001
MVLTTAQRAAIQNNWATVSNANMQEMGDALFMRYLLANPGDIQFFPKFASAGVGAQLRSNE
AFQEQTTLTVFQFLGQIVAKLNDLDAGKMLQERVHSHKPRGITMSQFERLLDLLPRFLQE

NAGAHGPCADAWRVAIANLMPYMRSEFKKC*
>Cte_216611 fgenes1_pg_C_scaffold_358000016
MGLTTAQRRAAIQNNWATVNVANMQEMGDALFMRYLLANPGDIQFFPKFASAGVGAQLRSNE
AFQEQTLTVFQFLGQIVAKLADLDAAGKMLQERVHSHKPRGITMAQFERLLDLLPRFLQE
NAGAHGPCADAWRVAIANLMPYMRSEFKKC*
>Cte_5378 fgenes1_pm_C_scaffold_336000002
MVLTTAQRRAAIQNNWATVDANMQEMGDALFMRYLLANPGDIQFFPKFASAGVGAQLRSNE
AFQEQTLTVFQFLGKIVAKLGDLDAAAGKMLQERVHSHKPRGITMAQFERLLDLLPRFLQE
NAGAHGPCADAWRVAIANLMPYMRSEFKKC*
>Cte_192931 fgenes1_pg_C_scaffold_173000024
MGLTKAQIAAIQNNWARISNNLQDFGDTLFMRYLTIYPGDLAFFPKFEHEGVDHLRHNA
DFQAQTLVVCQFLSKVIASLSMDAAKAMLQERVTRTHAPRGIAMAQFERLLDLLPRLVQD
ASAASGPTADAWRVAVASLMPAMRQEFQFAKV*
>Cte_21508 estExt_fgenes1_kg_C_101120005
MGLTTAQRRAAIQNNWATVNVANMQEFGDSLFMRYLLANPGDIQFFPKFASAGVGAQLRSNE
AFQEQTLTVFQFLGQIVGKLNLDLDAAGKMLQERVKTHKPRGITMAQFERLLDLLPRFLQE
NAGANGPCADAWRVAIANLMPYMRSEF*
>Cte_221680 estExt_fgenes1_pg_C_7710011
MGLTTAQRRAAIQNNWATVNVANMQEFGDALFMRYLLANPGDIQFFPKFASAGVGAQLRSNE
AFQEQTLTVFQFLGQIVGKLNLDLDAAGKMLQERVKTHKPRGITMAQFERLLDLLPRFLQE
NAGANGPCADAWRVAIANLMPYMRSEF*
>Cte_227604 estExt_fgenes1_pg_C_5710023
MGLTTAQRRAAIQNNWATVNSNMQEFADSLFMRYLSANPGDIQFFPKFASAGVGAQLRSNE
AFQEQTLTVFKFLGQIVAKLNLDLEAAGKMLNERVVTHKPRGITMAQFERLLDLLPRFLQE
NAGANGPCADAWRVAIANLMPFMRDAFAK*
>Cte_21094 estExt_fgenes1_kg_C_440003
MGLTTAQRRAAIQNNWATVNVANLQEFGDALFMRYLLANPGDIQFFPKFASAGVGAQLRSNE
AFQEQTLTVFKFLGQIVGQLNLDLDAAGNMLKERVASHKPRGITMAQFERLLDLLPRFLQE
NAGANGPCADAWRVAIANLMPYMRSEF*
>Cte_132766 e_gw1_47_150_1
MGLTKAQIAAIQNNWARISNNLQDFGDTLFMRYLTIYPGDLAFFPKFEHEGVDGHEHLRHNT
DFQAQTLVVCQFLSKVIASLSNMDAAKMLKERVHTHAPRGIAMAQFERLLDLLPRLVQD
ASAASGPTADAWRVAVANLMPAMREEFAKV*
>Cte_18433 estExt_fgenes1_pm_C_1730003
MGLTPAQVASIQKNFATINSDLQGYGNKFLRYLGANPGDMVFFPKFENVAYNDLVSNSA
FNAQTLVVMFELGKVVVNLGDLNKAGAMLQERVKTHKPRISMAQFERLLDLLPRFLQEE
GKASGAVADAWRVAVASLMPFMRAQFAK*
>Hro_171404
MGCQQSTSKSLEVDYSKEDWFAFLVVRTQNMGANGFKKVKSEPLLNNLNYNLNDVLTIRE
KQLVRESWTLISIKLKSGLKQVFLRIFELRPSTKNLFPFKTVWGDKLIKHPFLTHSKRF
VKVIGCVDRDLQYQRECAQPLIELGKKHVSIEGFLPDYDYVYIRAIISIWKQELKDVYT
NELSEAWHKVLVYIVSKLKEGYETEWKVATYFNPO*
>Hro_176761
MGCICSMIRKRNFGDSDSVYLDAPLTPDQKRKLNWALLSIEIEKKGTRMFKLLFHKNPF
IKNLFPPRNLDGDELESDPSFVAHAHRFMNTLGQVMHNLDRYKEIVAPQLVELGRRHANF
IGFKPNYFNHFEEAMMDVFSDELGPVFDVKAVEAWRVVFRFILIELKRGFTLALRDRGK
QMRECCEQQYLQYQREQLKELRQKQQQIEEMQKIQQQLKRLEAQQEQKMRDIKEEQLEM
DILIHNEQMEKLEMEKRQFTSSSSQMTSTNS*
>Hro_181863
MAARATGTHVINNTAHVTSQIITSRLSYVITGTLQSNPLVKNTFEKFRQMDPMSDFTDS
SVFSTHAMVMSAFEDIIDNLDSEIVKDILEQGKSHGKFSEDFAPETFWAIEEPFMSM
KDILGRKMSSQLEKIYKTIKFIILSVLIKGLRDATMA*
>Hro_159415
MNNTGKTTEDTTLQSTSTEKQGVTLQVLEIFRNFCEGCFNKKPTLFGNQVEPSQPI SMPKAV
ITSRTDQPVVXHLTEGEIYNFEKIWKHVQKNLPEIGKNMFKELFTKNPALSEKFDKFRGM
KGEALLESVALDVHGMVVMNTMDEMVSMDGSKFIEEILEEVGMSHARFGSQIKIDDFW
ELEDPLVNAIVKVTPGINQALKLTRKVIIRFIITGITKGFKDEWIKNGEILEKIDNKDPS
TFSSFYSGSHKKPDLSEIEKDMEEKIQTNEMLHHHHHGHGGHHHHHHGGDHRHRHHENELV

MDSVSGVWSVASYSNNRKSSPSGHLSTATRKSEARNSTVSHLSGRSLAVTSPNQYHP*
>Apo_N60237
MKILIVLAACLVAACECGPLERLKVKAQWEMVYGSNAAARETLGEEVWTHVFLHEPKVR
EHFGRVRGDAVTSYEFQAHATRFFGGFDICISLLNDPDTLNAELMHLHDQHNERGIPHEY
FDVFRHSLLYIIISTHVEHFDADAWRDCYNVIKNGITEGLTE
>Apo_N65733 ORF sequence | 160 aa
MMSLQLVLVACLVAATAATECGPLERLKVKSQWDEVYGASAEKRETFGEEVWTHIFLHEP
KARELFTRVRGDSVGSHEFRAHAVRVFGGLDICISLLNDEDTLNAELSHLHDQHVERAIP
PEYFDVFRHALMHVMSQHVEHFEQDPWVACFDVIKSGITG
>Apo_N44255 ORF sequence | 242 aa
MGCAPSKEGATVKAVDLQAVNSSQSRSSGASKSILREAVETGSYSKMASYKPDPRCPLTE
RQLYSITKSWKAINREMASTAVNMFIRLLEHDGIRSFFTKFKDHKTVAELRASKVFESHA
LMVISVIDDVITNLDDMDYVMSLLQATGESHSIKFKNFNPDLWNVEGAFLWAVKETLGD
RYTISIENIYTTIRYIILQSLHDAFTKHRERQNSTNNDXXXKTNLLNQELSTADRKTAPDS
KD
>Apo_N32970
VGECQVALSVFEDLFEFERPDAKFLFTRVNVENLQSPFEKAHCIRVVNGLDTAISLLDDPF
TLINQLDHLGKQHQIRDGVKKEHFELMSRSFLKVMQVSSCFNADAWSRCFDGIARRISS
HLSN
>Apo_N58141 ORF sequence | 159 aa
MKTLLIILFAFVAVATCTECGPLQLKVKQQWSVAFGTDHHRIDFGIAIWRGLFRQVPDAR
XLFKRVNGDDLYSGAFRAHSMRVLGGLNMIISAIDNEDIAKFXLNXLHDQHVDHRHVAASY
YQAMKNSLMKVI PAAIGRCFDEDAWNACMDVIIHGISGN
>Apo_N33329 ORF sequence | 130 aa
MDIFVRLFELHPVYKSYFQKLRDVEDIEDLRQSGKLRVHSTSVMKSIITDLVETLDHPPDLR
DMAIKIAHPHFDRGVRPSQYRELFAIILEYLKDKAKVVFNDEAEAAWQKLFYVLDITAA
VMDLQIEKMG
>Apo_N60246 ORF sequence | 172 aa
MFRFVALLAVLAITVAENPHAHDFQCCSTEDRKEMQALWHEIWSAQFTGRRVQVALSVFE
DLFEREPDAKNLFKRVNVDDMNSPEFKAHCIRVVNGLDTAISLLDDPFVMLHQLEHLGKQ
HQVRDGVKKEHFDMARSYLKVMQVSSCFNADAWSRCFDGIAHKIASYLP
>Apo_N61352 ORF sequence | 161 aa
MSELTEERREAVVESWKEICKDVRGNGVQLFLRYFGKFPAYQEYFKDLRGLSLEELKTSK
HMRIHGTRVLHAITSMVDALDELVAAGIMSKTVDTHFDNFVKNIKQYEDLFSVMPDFMQ
ACAGDKFTPTAAEGWTIVLNTLLAVGKERIAEIEKQKAAMA
>Apo_N64919
MGCAPSKEGATVKAVDLQAVNSSQSRSSGASKSILREAVETGSYSKMASYKPDPRCPLTE
RQLYSITKSWKAINREMASTAVNMFIRLLEHDGIRSFFTKFKDHKTVAELRASKVFESHA
LMVISVIDDVITNLDDMDYVMSLLQATGESHSIKFKNFNPDLWNVEGAFLWAVKETLGD
RYTISIEGLTERMNESRKATERTLSKTTLCIASPRE
>Apo_N67641
MGCAPSKEGATFTSSDPQANSSQSRSSGASKSILREAVETGSYSKMASYKPDPRCPLTE
RQLYSITKSWKAINREMASTAVNMFVRLLEFDGIRSFSSKFKDTQTVAELRANKVFEGHA
LSVISIIDEVITNLDDMDYVISLLQATGESHSIKFENFNPDFLWKVEGAFLWAVKETLGD
RYTISIENIYTTTIRYIILQSLHDAFTKHRKENPRESTETEKAGELPTNETETHGETKAIS
DNSLSVDFIGTNGATKHFVAAEE
>Apo_N74712
MNTSLMVVCFVAMATSALADCGPLQLKVKQQWTSAFGTAHHRIDFGTAMWRSIFRQAP
QAVDELFTTRVEGHDLVSGKQAHSMRVLGGLNMVFSVMDDEAVLQSVLGHLRDQHKERAI
PRNYFGVFKSALMKVVPATIGRCFDSDAWDACFDVVVDTLDLADY
>Apo_N81834
MDRVVTLTLTFFTLTTFVTVVMATNRKCGPLQILKVKGQWEKAYGVGRTRDDFGLAIWRSI
FAQDPSVLPFRRVHGEDLYHPDFIAHSARVFGAMESLIALLDHQDTYDMQVAFILRSVHK
YQGLAANHLWLMKNALMRVLPVQLQFCFDEEAWDGCLNKIIEDMNGQRQPPGERQVEIK
>Ama_62990081 emb|CAI56309_1| haemoglobin B2 chain [Arenicola marina]
MLRFVALLALVGLAVCDDCTTEDRKEVQTLWSEIWSAQFTGRRVQVAQAVFEDLFRDRP
ESKNLFKRVNVDDMNSPEFHAHCIRVVNGLDTVIGLLDDPDTLKSQLEHLAQQHKERDGI
HKTHFDEMASHAFGAVMPQVSSCFNPDAWNRCFGSIATKIASLLED

>Ama_62990079 emb|CAI56308_1| haemoglobin A2 chain [Arenicola marina]
MKSLVVLVLFALVAMVAEACGPMQRLLVKTKQWNKVYGTSKVRDEAGHVLWKAIFAQDPETRA
LFKRVNGDDIYSPEFMAHSARVLGGLDIAISLLDNQADLDVALAHLHVQHVERHIPTRYF
DLFKNALMEYAPSALGRCFDKTAWSSCFDVIANGIKE
>Ama_84618895 emb|CAJ32740_1| haemoglobin A2b chain precursor [Arenicola
marina]
MKFLVVLVLFALVAMVAADC GPMQRLLVKAQWNKVYGTSKVRDDAGHVLWKAIFNQDGETRA
LFNRVHGDDIYSPEFMAHSARVLGGLDIAISLLDNQAELDAVL AHLKEQHIERGIPDRYF
DLFKNALMEFAPSALGRCFIKDAWSSCFDVIANGIKGQ
>Ama_84618897 emb|CAJ32741_1| haemoglobin A2c chain precursor [Arenicola
marina]
MKVLIIVLMACLAYVAADC GPLQRLKVKHQVWVQVYSGHGYEREA FGREVFLEMVYNQAPKAK
DLFTRVRGENVFSPEFGAHMVRVLGGLDMCIALLSDDTVLNAQLAHLSTQHKDRGIPNEY
FDVMKVALMKVVPGHVSHFDFDAWSACYDVIANGIKH
>Ama_84618899 emb|CAJ32742_1| hemoglobin B1 chain precursor [Arenicola
marina]
MMSVVFLLGLVAYASASSCCSYGDQQKVKQWNSLWNTPDSSSTKIIIFGKEVFFARFFEVD
PESKSLFGRVKVEDPDSPEFAGHVIRVLTGLDLIINLMGDDAMD AEL AHLNTQHLAREGI
TGTHFTEMFKVLDGSLRQVLEEYDLSWRVYCFRGLGAALRDGLPA
>Lte_56967016
ADDEDCCSYEDRREIRHIWDDVWSSSFTDRRVAIVRAVFDDL FKHYP TSKALFERVKIDE
PESGEFKSHLVRVANGLKLLINLLDDTLVLQSHLGLADQHIQRKGV TKEYFRGIGEAFA
RVLPQVLSCFNVDAWNRCFHRLVARIKDL P
>Lte_56967017
KKQCGVLEGLKVKSEWGRAYGSGHDREAFSQAIWRATFAQVPESRSLFKRVHGDDTSHPA
FIAHADRVLGGLDIAISTLDQPATLKEELDHLQVQHEGRKIPDNYF DAFKTAI LHVVAQ
LGRCYDREAWDACIDHIEDGIKGGH
>Lte_56967018
DEHEHCCSEEDHRIVQKQWDILWRDTESSKIKIGFGRLLLTKLAKDIPEVNDL FKRVDIE
HAEGPKFSAHALRILNGLDLAINLLDDPPALDAALDHLAHQHEVREGVQKAHFKKFGEIL
ATGLPQVLDYDALAWKSKLGILTKISSRLNA
>Lte_56967019
ECLVTESLKVKLQWASAFGHATERVAFGLELWRDIIDDHPEIKAPFSRVRGDNIYSPEFG
AHSQRVLSGLDITISMLDTPDMLAAQLAHLKVVQHVERNLKEPFFDI FLKHLHLVLDRLG
THFDGAWHDCVDQIIDGIK
>Lgi_72750 gw1_61_132_1
ECEPHEDETKPSLTDROKELVLESWQIVQGDLDKVGVD MFMRLFQTQPDVQDVFPFKGK
SADDLKDSKQLKSHATRVMGTV EKSLANIEDPKRLEQMLS DLGARHV MYNAKV DYIDLIG
PQFIWAIQPVVDEKWTPEMEQAWSDFKYISHIMK
>Lgi_73010 gw1_61_134_1
LVLESWNIVQQDISKVGVMFMNLFETHPDVQDVFMFPFKGSKEDLKQSTQLRSHALRVM
GTVEKCLARINEPQKMD EMLHQ LGSRHV MYNAKV DYIDLIGPQFIWAIQPTLGDKWTP EL
EAAWSDFKHVAYIIKRAMV
>Lgi_159879 fgenesh2_pg_C_sca_22000090
MGCTISATIRNKGIKDLPGYLTQREVDIVQKTWAILSTDMMG TGVMFQRLFEIQKDLQK
LFRKLLLPSETADYFFDEVK LQSHAMIVMQGLGA AVESLDDSVYLTNILLTMGEKHSQYN
VQPEMISLLWPAIRDALKHKLKEDFTPETELAWRHVFDYISSNMAEGIRNGQNKVKS NKA
VKS*
>Lgi_118731 e_gw1_29_229_1
MDLVRETWNIAREDIARVGVVAFSLFEKYPETQNI FL PFRGLTVDELQHSARLREHGLR
VMLTVEKCIARLDKPEKLQDLLHDLGQKHVEFN TKVDYIDMVGPPQFIQAIRPVVKS DWT P
EVEDAWEDFFKLLIYFMKEAMCF*
>Lgi_63395 gw1_124_28_1
KKENPPVLS EEEKTLILNSWKFIKDDIAKVGIFTFIGLFESH PDIKHAFVSFRGLQPKEL
NSSVLR AHALRVM TTVDKCLFRFDNLEKVEELMKS LGCRHGQSYQVVHQHLDLMAPHFN
YAIKHNLDQWSPELES AWDKL FKLMIHYMKSGM
>Lgi_128314 e_gw1_61_318_1
MFKNNSGVQQLFRELRDLKTTDEL RMSEVLEKHASKVMSILDD S INNIDSVDIT LLELLHR

TGASHKGYQGFTAELFWEIEQPFLDSVKLTLGDRYTDNMDSIYKITIKFILENLIKGMKN
S*
>Lgi_125519 e_gw1_50_153_1
RLFEANPGVKHTFVTFRSVPSKELTNSLLRAHALRVMTMVDKCVSRLDELDKVEETMKS
LGNRHTRYQVMPHEVDLMGPHFVQAIQKCMESEWNKETEDAWYQMFKLMTFYMKEGLKYH
EKA*
>Lgi_233216 estExt_fgenes2_pg_C_sca_360046
MGCDNSKQQLQPNIPDSQEEGTIGFTETQIDTIRSTWPLLSRNMVRVGTDFVFRIFTEVP
TVKELFSSFNIVDNDLHKMPTFRAHAEMFMQVLHLVVDNLETPYSELNHELMVLGARHA
TFSGFKPEYFKFYVKCLIQVWELELGEFILEVRDCWKIVDFDLVDNMTEGYELALREST
TKHCQNGVKTLTVSDNGNSKSNQNGHMTSQNKRNTPNPAVKKNKSACIMDKNKKNDTLCE
VKIT*
>Lgi_128580 e_gw1_61_320_1
AENEKLLKGHVAAVMLTIDEAITTLNDADQVIDTLTHIGATHVRFSGFKPEWFLLIKEPFL
VAVKDTLGERYTASMESNYRKAIRFILETLIKGYENGSAVANGQG*
>Lgi_231953 estExt_fgenes2_pg_C_sca_240135
MACAITGLTAAEKSLVAEAWKNISPKKKDYGVELFMSLFTAHPDALDHFKNFPSKNLDEI
KNTADMRAHAAAVMYALDSLVDNLEDPECMVGLIRKIARNHKSARGIGKSRFEQLRGIFGN
FLDTSLGGKSNATTKGAWDKFLHFTNNQIETMQA*
>Lgi_232639 estExt_fgenes2_pg_C_sca_300121
MAVYYTSRSRYQNKRFVYWTITRTCFYRFFETEPDMKALFPKIVQMNESNQLEWQMDR
DMLQKHAVTVMIEGLGAAVESLEDSDFLNSVMVSIQQTHVRRNVKPKQYVKKLWPSLHYGLG
VCLGDHYNKEVSEAWRKVYIYISAQMTRGMKNPNLKTNDELS*
>Lgi_233247 estExt_fgenes2_pg_C_sca_370005
MSADNMGCGSSSGVSTGMPADLTEKDDELVKSSWAKFNEGDVIADGAHIYYKLFKAPKA
KEKFGFAKDGEVSLNKFKAHVRLVDVFEVSVREIDQLEGLLPVLNNDLGARHKSYPV
LKYEILGSCIMYAWDRKLLKMDADTKKAWGKLYGVVQTEMKKGNA*
>Lgi_203115 estExt_fgenes2_kg_C_sca_40019
MGLVMSSIQYIFGGGEQGLDIPDKATGLTIREKRLIVNSWEQIKVNIKQNGVALFIGYFI
AYPYTQDYFHKFKGKDLKQVEKSAQMRSHGTTVMHALNALVENLDDPECLVDLLQKNAV
HFKLGIRYQQYKELFDMFPGYLKDRIGAQCTEQTVVAVNKATNVMLSI IETELTKQSNVT
K*
>Lgi_231954 estExt_fgenes2_pg_C_sca_240136
MTCDVTGLTDAEKSTIRDSWQLLSSKKKDNMGALFMTLFSSYPQSLHYFHEFDGKSIEEL
QGMSDMRAHATAVAMYAMDSLVDNLDPHCMLGLIKKISRNHKARNIGKKDFEELRALFGG
YLDKQLGDQATSGIKAAWDKFLTVLNNQIEKDQA*
>Lgi_167450 fgenes2_pg_C_sca_66000094
MDDNQESLKENYRFRCHVGLFCETIRIAVEEMREIEEVLLFLKDLGRKHRMYGATPTYIK
TAGEGIVYAIRKLGNEFTSRMKTSSWKKFFFTILQDSILEGLEFG*
>Obi_961126590 |XP_014784435_1| PREDICTED: neuroglobin_like [Octopus
bimaculoides]
MMGCGLSFKSEKPKKKQKEVTSNTCTILSDRQNLIRETWSIIQDEILTIGVYIYYSFFQ
SEYDLREYFSSIFQINKEKTKMHMNQDKLERHAMLLMETLGFVTKLDEIEDLRDFLNDL
GGKHYWRKVKPCMIRRLWPIIDDSFRQVLCVYTTTEVREAWQTFIEFIFESMKEGIDAMS
LAQSLSTECEINIINNQSNDKPILNRLEENSDAELTISDENVDTKLLNK
>Obi_961080890 |XP_014768482_1| PREDICTED: non_symbiotic hemoglobin_like
[Octopus bimaculoides]
MGRNSSIEKKLSLETDNRNLAQTNTRPIKERPPFNENQKQLIRKSWKIIQADIKVGVI
TFLRLFEEKYPDVQDIFIPFKGQSLDLSNNDRMREHGLQVMRTVEKAIARLDQPAKLESM
LFELGQKHVMFDTKVDYMDLIGPQFILAVKPTMKDKWTIETEEAWSDFFKYITSIMKEAM
IF
>Obi_A0A0L8I8V4 A0A0L8I8V4_OCTBM Uncharacterized protein OS=Octopus
bimaculoides GN=OCBIM_22028113mg PE=4 SV=1
MGPOFMMAIKPHLESQWSDEMEEAWTHLFRIISYHMKKGLRPSKAEQI
>Obi_961133878 |XP_014786974_1| PREDICTED: non_symbiotic hemoglobin_like
[Octopus bimaculoides] >gi|918288843|gb|KOF68082_1| hypothetical protein
OCBIM_22008093mg [Octopus bimaculoides]
MGSDYSKNRKLKQRKVSSKVQPEPSSHVISEVSRGLSESNNVANYITDKQKELVLETW

KIVQCNMARVGVVFMNLFETHPDVRDVFMPFRGLATEDMKQNSRLISHAFRVMGTVEKC
LARINEPRRLEDMLKNL GARHVMYNAKVDYIDLIGPQFILAIKPEVGDHWSLEVEEAWSD
LFKLITHIMKAAMAF

>Obi_961129975 |XP_014785608_1| PREDICTED: neuroglobin_like [Octopus
bimaculoides]

MLCLQRAEAMADGLPYPPPPPTDPRRLPLSPLQVFKLKKSWKGIKRSIELTGVEMFVRMF
RTEPFLKDMFKDFRNLVTDDEMRENMALEKHATMVMNLLDEAINNIDNVDLLLDLHRVG
KNHLRFEGFDVSYFWLAEQPLLEAIKITLGDRTYENMDIYKLVIRFLLTEITKACRNDV
S

>Obi_961129978 |XP_014785609_1| PREDICTED: neuroglobin_like [Octopus
bimaculoides]

MGRESKEQCGPNLTERRKSNVCQEPNSMESAIQVDPRLPLNARDIFKLLKQSWKGIKRN
METTGVEMFIRLFQTEEEIGYLFSDFKNVHSDALRQNEVLEAHALLVMATLDKAITSLDD
YDAVRESLLRTGGKHYRLKNYKPQYFRLMEQPFLSAIQITLGDRTYSESMIYKKTIDFI
ITNLQGTGILEAIAEAARTNNAENSQPAPIENTNHVTSTPTVQPLPVVS

>Obi_961125540 |XP_014784070_1| PREDICTED: cytoglobin_1_like [Octopus
bimaculoides] >gi|918295466|gb|KOF72022_1| hypothetical protein
OCBIM_22000178mg [Octopus bimaculoides]

MGNQQTKSEVRPHRISINSVIYHRRQNSRQLNKYLTTTRQIQLIQDSWKLKDLTFVGT
TFKHFFETDQELKTFPPKIIRISEKNELEWDVDNDMLQRHGITVMRGLEAAVESLDDSQF
LNNILFRIGQAHVYKNIKPHMLKRLWPSLNRSFKEVLKDRYNKDIAEAWRLTYQYICSQM
KNGMENSSEQ

>Pfu_297949 72256_t1

ELRPELKQLFPFRDVTGDELKMNPTFKGHASRFMQAVGAAVDNIDDWEAAAFAPLLLGLGR
QHMFNFGGFKPEYFNAFADSMFWEQVLGDRFTTEARCAWQGVFEFIMLKLKEGYAEADS
ERNDNQNAKV

>Pfu_10603 24769_t1

MILSERTDQFAAWTGHARFGNSNLVSDPQSKRCNIPKECENHYVRSNLLXLFKPNPDLK
DLFVFPFRGMDMNQLRQHEGLREHGLRVMGSIEKCVARIDSPPRFDAMLTELGHKHVVFNI
KPDYFELLGPQLITAIKPAIGEKWTPVEVEAWMNFYHLIYTMRVAMVS

>Pfu_11519 46391_t1

MMGCDVSKSVDMPEPQQSECTSAEFTDTQIDTIRSTWPLLSADMTDIGGKVFRLTFYEEP
RTKDVFPQFSDLEDDEFIAXHPFFKGHVAKFMQVVDVAVDNLGPKASIQQLLLMLGARHA
TFQGFHLEYFDVFTKVLIDVWETEIGEEFIPEVKECWVVFAYIVKYLRQGYFLYVNECT
DDVIEEKS

>Pfu_19774 04035_t1

MVCNVRRRMFEIHEDIPTYFNITGTKAFLDIRQNDKFMHDHILVMHTLDEAITGLGEVDY
VQEMLRGVGRSHKQLQNFHFLIFERIEEPFIFAVKETLGDRFSANMDNIYRKTIKFIVKE
LAMGFNEPSLCRSSGDR

>Pfu_2532 30156_t1

LFETHPDVQDVFMYPYKGLSQEDLRHSTTLRDHALRVMGTVEKCLARINEPKKLQELLHDL
GARHVMYSKVDYMDVSVRNFPPLSRVDRTQHVALSNCLFLHQW

>Pfu_12427 25029_t1

MFKRKEEIKGIFTKFEHLKNDDEMMDDETFEKHATLVMGTLDEIISNIDNVDYIFSIIKT
MGERHVRIPNFKQDFIMEIEQPFLAVKLTGLDRTYDNMDSIYKITIKFILETFVKSIGE
AEKARAAER

>Pfu_464 00313_t1

MEDGNVEFQWNDITNSEQMNTIISDRQKALVTESWKSFTKHGDLPLGIPMFLKLFEDHP
EVKSLFSFMRAEGISSKHPQFRRHANRVFEMIGSAVGSIDDLDTLGDVVKELGANHYKYG
TQSAHLPAVGNALLFALEKELREEFTNEVRNAWLSFYSVVSKCFSEGLEKKIKAMEDDNS

>Pfu_7779 67746_t1

MQEHLLKLFASMSSQKNGHHMFIDDDKLRSHGKIVMEALGSAVESLDDSVFLTNLLIAMG
TKHSVYDVKTDMVVYFWPAIRDTMKDKLGDDFTPEVERAWSHVFEYIASKFQEGIRKGGK
NDRNKAEKVTGI

>Pfu_7847 45741_t1

MTLSITRHYIYSFYAILSRLKERTIVVLNPLIVRGPHNTAKRRGSPNTDEYVNGYKK
SAYVIQSVLYRLERAHSFGIHTDSDLEVPIYWI PKMHKNPYKHLFIAGVPEGFCRKVR
STQSFLQIRKFLLSKFLKISETFKKYPKRKNVIVDIRQIILFVRFEEIEPRMKLLFPKI

VHLSDSNKLAIEPDVHRVLRHAASVMYSLGAAVECLDQAELEFGVIAEDIGKLHADRHIK
INTIHRLWPSLNYGLKQILGDSYTKDVAKAWRKVFYICRRMEQGMAGKISSESSDRSSG
ESGNY

>Pfu_2070_44228_t1

KKPGSTKNALNLRLLDFGLMVPLFWASVSPQFTTIHLSFQIRRLTLNEFSFVRVKKNVGK
ELGIIGSHLDADYSLEDFVLCSKRFFKREDNETLRLFPKVIQNEGNGGSMKIHENLQYH
VMLVMKGIGKAVENLDYPSSLISYLNLYLGRRLQHNVPYLLLECLWPAIDEGFASVLTDV
YTEELRTAWKMFIDMMVTRMQSVMTESRDVTSQQNVVKSFRKSSSLSDNSEQTQNGQVK
RNSGSSKH

>Pfu_97_57971_t1

MKKAGIFSRSCITDTSVRQCVPRLFSYVATCCLERC SVTEPLKCYIVI ICIVAFKEILIL
RPRSEFVKTMGCTSGRMVAPSMTEKTLIKDSWSLIAGEKDLGDVGLIMFNKLFQDNSS
LRSLSYFLRDDNVNDDVNGKLRSHGKGVMEVVDIAVSSIDDLRGLHPVVLDGCRHYKYG
VQKS

>Pfu_15076_68815_t1

FFETEPLKSLFPKIVQMENSENQLLEWDIDKDMLQKHAVTVMEGLGAAVESLEESDFLNSV
LVSIGQTHIKRQVKPQMLKVRNRSRGSNDNFALPSGL

>Pfu_56872_12873_t1

MNSENQLEWDIDKDMLQKHAVTVMEGLGAAVESLEESDFLNSVLVSIGQTHIKRQVKPQM
LKAYQCICRIVLFVYFF

>Pfu_3172_37526_t1

MAANGSEQNGREEVDPVPTGLLPSEVDALLESWSLLFRKEHRKKNAIELFMILFQDPNNG
RDAQNLFALKDVPTDKLRTNKT MAGHAVTVFHAIDSFMEYIEDSDTLVELVKKQAINHI
NRKIGAKEMAWLIPVIKLLDQVLEDDCTEKMKSWEKFLGVLVAVTESCEKEMAK

>Lan_25933_t1 unnamed protein product

MGCSSSSDGLDNHYDEPPRKVDPRLPYTTYRQVFNLKNWSKAINRKLEDTAKDNLMRFF
RKHREYQAMFPQLKMDDEEIMQTSDAFEDQSLRIFNIFDEVMEYIDCNVDNAIDILHSTG
KQHARIEGFKPEMFSEMEESFMEA AKEVLGDRFTESADDNLRKLYQFCIKHINAGYAAAT
Q

>Lan_6164_t1 unnamed protein product

MGSHISKGKFSKKPVNGPKCTEPTASTSEGQANGHAHKYVQTEDELPTKSDDADIQET
STGSTETQKHLNPKNRPSEADVPFVNEQKKIIVKTNWIKEDIKAVGVITFTKLFETHP
DVKEAFLHLKDMSQEELEYS DILKRHTLRVMQSVDKAVSRIHKPVMIHNVFLELGERHID
YHAEPRIVDLMVVQFIEALHSRLPAEHWNETEQAQQWFLEYTMFFLKVPILCMRKT

>Lan_19222_t1 unnamed protein product

MSTNYDEVCDNMLLQLYERVRYVKDKSIHLSLDANQRMMIKLSWNAFVAGDPSNRGMQMF
LKMFSMQPKTQSVFEFARGSSAAQMNSTRIIFHVTRVVKYITKVVENLESLEDMAVLR
RLGGRHGTSGYNVPTDYFPHLGVAMRELMKAHIGNWTKQHDQLWDTLYTWIVARMREGQA
TYGGKR

>Lan_18236_t1 unnamed protein product

MANLLLRKSRRESVSYDEQCEVMLIQLYERVRYVRYGEKLNFTPSERTSLRLSWNAWIGGN
PSQRGFQMFRLRMFKSHPETQLAFEFAGSSLAYLQNSSRLLFHVNRVVKYIGKV IENLDS
PEEVVPVLLQLGAMHGPRRFNVPSGYFPHLGVAMRELMRESLGNWTEDLNKLWDELYQWI
IQRLIEGQQRVES

>Lan_6155_t1 unnamed protein product

MGAIWSMFFGYGSDTVDEKTLTEREKNILRYTWKGISTKPRELFTKIPETKKYFRSLK
DKSNEELRKSFVLRAGATVVNSLTSVVE SLDDADCLVSLKIGTNHQKIGIPVALFGP
EFCELIISWLENSLGSSRFGPATERATWDKALSVIMSVIKSAYDEK

>Sme_15011987

MYVMGHVHSSSKLSNDGNISNNPNDSKSSSKLKS CIDTFDEVNLNDLSSTQKCLISESW
TIIRNDIERMGSIMFLELFAHNVDFQDAFSIFKGRHMEDIKSDSRLASHGIRVMSIVDKV
VSRINESEKAARVLFELGQRHCAYDVDEKFI PVLGRQFLITIKPTLDTNNLWTEELEKAW
TIIIVIFICQKMT EGIRTNKQVNN*

>Sme_15030189

MYVMGHVHSSSKLSNDGDISNNPNDSKSSSKLKS CIDTFDEVNLNDLSSTQKCLISESW
TIIRNDIERMGSIMFLELFAHNVDFQDAFSIFKGRHMEDIKSDSRLASHGIRVMSIVDKV
VSRINESEKAARVLFELGQRHCAYDVDEKFI PLLEYS DKMKELMSIKLNNSN*

>Sme_15031672

MGVAKSFLVFHKKSIITNKQKQFIDSTPEKYLIQLTDEYKLLIENSWKKLKTDVESMGSL
IFLQLFAENKDLKSNFRYFKDLSIEQIKADSNLADHGLRVMSSVEKLI SRIHSPDKMRSM
IQDLGNRHKDYGAIVHFVPMVGERFLKIIIEPRLIEHDMWCEKTEKAWKILIKYLCTELMT
SMTS*

>Sme_15019624

MSISIEDKEILWKEWSEIVSTKEDKIKLGQKIFQRLFTIHPEYIKLFKRFFENMSSLDEII
SSSAMKAHSLRFINSLNGVFETLDTPEELGDMLKHVGSSSHKLNRVTCNHFVATVPILNEV
LAENLKSCVKEMVKNILNIAAPIILEGCSE*

>Sme_15013088

MHTHHLLFKSCINIQNSINFKMTLTAGSREILWKEWSELVNPKEKKILLGQILFRRLFAA
HPDFIKLFRRFDMNSLDEIIQSTELKAHALRFMNSLNGVFETLDNAEELGDMLKHVGV
HKLRSVTCDHFAATVPVLQEIILVENLKSKEAKEAVKFLTLAAPIILEGCRE*

>Sma_T1JGQ8 T1JGQ8_STRMM

MGVESSKTSKLCRHGQNAAPACVNSGMSTSMGTGKEFCENENAYDDVESWVLDEREIEHVI
FTWKLVERNIKAVGVITFLGLFETHPAVQSVFLPLSHMSREQLGSSAKLEAHALKVMNFI
QKIIARIDNPSKNWQNTRESLWDHVGRYRGDEVI SEGLGEVPAAESATWGQAEKGNFSEC
RG

>Sma_L8460 ORF sequence

MGCSEFVKHNSGAGEEGRSAKSSAGKSGVVNAVDTPAAPAPVDPRLPLNARQLFQIGKSWK
GISRAMEYTGVMNFIKLFEEHNE LLNLF TKFSDLKTKEQQAESLELQE HATLVMTTLDES
IQALENVDAFTAYLHQVGRSHTRVPGYKKEYFWRIQKPFLEAVSETLGDRYTENMETIYT
VTIQFILETLVKGFEIGEKEKGV

>Sma_T1JI21 T1JI21_STRMM

MDDIFHHIKSIVTMGGWISYFWPQKSAEFDVSPGLDEVESASGLTLRQKKVVTEIWDLVK
IDIKQNGIDFFIEFFKAFPLNLFKAFQNMDDQLRKS KLEAHATNVMYAI STVVDNL
QDVECLTELLSTIGRNHIKRKITPVQFDQVGITFIKFL ENKLGSRITPFCRNAWEVTFKV
MNSIIVAGLQSNDD

>Pte_1009569340

MGVNLSKVLAVLQKKGPEEAGAAEI PAPLED PPTPPAPDPRLPLTAKQLFNISKSWKGIA
RAMEPTGVIMFVRLFEENEDI LHLFEKFKKRITELHRDSMELAQHASIVMNTLDES IKK
LHNVDYFMDYLHAVGKLHTKIPGFQRDYFWRIERPFLAAVQETLGDRYTDNMENIYKITI
RYILDTVVKGFDLHSHKKEPEPSDKVRPPKASPEPSPSPHDDDTSTRPENKTEIQKKT DGLS

>Pte_1009569338

MGVNLSKVLAVLQKKGPEEAGAAEI PAPLED PPTPPAPDPRLPLTAKQLFNISKSWKGIA
RAMEPTGVIMFVRLFEENEDI LHLFEKFKKRITELHRDSMELAQHASIVMNTLDES IKK
LHNVDYFMDYLHAVGKLHTKIPGFQRDYFWRIERPFLAAVQETLGDRYTDNMENIYKITI
RYILDTVVKGFDLHSHKKEPEPSDKVRPPKASPEPSPSPHDDDTSTRPENKTEIQKKT DGLS

>Pte_1009596185

MGCTFAKAQKNGSVQDLNHATEAPAAPPQDPRIPLTARQKFSISKSWKAIARAMDTTGV
AMFVKLFEENAELLELF EKFKHLKSRQE QE ESEELKEHAV SVMNSLDEGIN TLENVDQCI
DYLR SVGKRHRKINGFKSEYFWKMEAPFLAAVKETLDDRYTENMESIYKITIHFILQTVI
DGFEGNPSQNNV

>Pte_1009596183

MGCTFAKAQKNGSVQDLNHATEAPAAPPQDPRIPLTARQKFSISKSWKAIARAMDTTGV
AMFVKLFEENAELLELF EKFKHLKSRQE QE ESEELKEHAV SVMNSLDEGIN TLENVDQCI
DYLR SVGKRHRKINGFKSEYFWKMEAPFLAAVKETLDDRYTENMESIYKITIHFILQTVI
DGFEGNPSQNNV

>Pte_1009596181

MGCTFAKAQKNGSVQDLNHATEAPAAPPQDPRIPLTARQKFSISKSWKAIARAMDTTGV
AMFVKLFEENAELLELF EKFKHLKSRQE QE ESEELKEHAV SVMNSLDEGIN TLENVDQCI
DYLR SVGKRHRKINGFKSEYFWKMEAPFLAAVKETLDDRYTENMESIYKITIHFILQTVI
DGFEGNPSQNNV

>Pte_1009553765

MGFIFSKLWTPANYDDDPATGLTPRQRDTVQNTWKIVRSNIKENGLTFFVKFFQKYPEY
QKLF PFADVPLEKLRDDKKVLAHAMAVMYALNSIVDSLGDVDCLVQILVRI GSGHKPRS I
QPIHFENLASFLVSLIEALGKSVMDSSAVEAWRTAFKAANGI IINALQNA

>Pte_1009561538

MGFIFSKLWTPANYDDDPATGLTPRQRDTVQNTWKIVRSNIKENGLTFFVKFFHKYPEY

QKLFPPFADVPLEKLRDDKKVLAHAMAVMYALNSIVDSLGDVDCLVQILVIRIGSGHKPRSI
QPIHFVIFLLEIFYLFLRYVTIRYLQLEKKKTAKICWLCINYVC
>Dpu_14524_Nterminal
MTTGPTTSTVPAKENSQGPTPKLDCFDHSIIRNTWDQAKKNGEFAPQVLLRFIKAHPEYQ
QMFQKFAVPHYNLLRNGDFLAQAYTISAGLNVVIQSLSSQELLAAQFNLLGSAYQPRGV
TPAMFEEFSVILEQVLEETLGSTFNVEARKAWNKGMAIIAGI
>Dpu_14524_Cterminal
SKTLKNPEDLADAQSNLTRPQIRNVQRSWESMKSGRNSLVS AIFIKLKFKETPRVQKHFAK
FANVPVDSLGRNGDYIQQVALVADRLDTLISAMDDQLQLLGNINYLYTHAKRSIPRKTW
EDFARLLVELLPTRGVSASDVESWKGVTTVLVNGIAPKNX
>Dpu_1462_Nterminal
MQVLSLALFIGIAAAVSAYAPGTVTTVTVTTLDEESTGILSSHRSIIRKTWDQA
KKDGDVAPQVLYRFIKAHPEYQKFKSFADVPQSELLSNGNFLAQAYTILAGLNVVIQSL
SSQELLANQLNALGGAHQPRGVTPAMFEEFGVIVEQVLEEEELGSTFNAEARDAWKNGIRA
LVGGV
>Dpu_1462_Cterminal
SKTLKNPEDLVDPQTKLTLHQIRDVQRSWETIRNDRNAMVSSIFIKLKFETPRIHKHFAK
FSGVAVDALSANGDYNQQVALVADRLDTIVSAMDDKLQLLGNINYMKYSHIKRGIARQTF
EDFGRLLMDVLGAKGISSDDLDSWKGVLTVFVNGVSPKNX
>Dpu_16194_Nterminal
MAFKLALLFGVIAFASACSYAPGTTVTTVTTAVTTVSADEGEEGILSSHDRSVIRKTWDQ
AKRDGDVAPQILFRFVKAHPEYQKMFQKFAVNPQNELLSNGNFLAQAYTILAGLNVVIQS
LSSQELMASQLNALGAAHQPRSATPIMFEQFGAVLEEVLAEEELGSTFNSEAQQAWKNGIA
ALVAGI
>Dpu_16194_Cterminal
SKTLKNPDDLVDLPQTKLSAHQIRDVQRSWENVRGGRNAMVSAIMIKLKFETPRIQKYFAK
FGKVAVDSLTGDAEFNKQVALVADRLDTIVSAMDDKLQLLGNVNYMRYTHARTARSIPRSAW
EDFGRLLMDSL GASGVSSDDLASWKGALAVLVNGISPKNX
>Dpu_23322_Nterminal
MAFKLVLLFGVIAFASACSYAVSQPGTSVTTAVTTVSADEGEEGILSSHDRSVIRITWDQAK
RDGDVAPQILFRYVKAHPEYQKMFQKFAVMPQNELLSNGNFLAQAYTILDGLNVAIQSLS
SQELMAGQLNALGAAHHPRGATPIMFEQFGTILEEGLAEELGSIFNAEAKQAWKNGLAAL
VAGI
>Dpu_23322_Cterminal
SKTLINQEDFADPQTKLSAHQIRDVQRSWENIRSVRNTLVSSIMIKLKFETPRIQKYFAK
FGKVAVDSLTGDFEFNKQVALVADRLDTIVSASMDDKLQLLGNVNYMRYTHAARCI PRSA
WEDFGRLLIDSLGASGVSSDDLAAWKGT LAVLINGISPKNX
>Dpu_1460_Nterminal
MATPGKTSYAIAMSIMTTEDDEMGSGLLSTQERAIIRTTWNKARKDGDVAPKLLFKFLKA
YPEYQKFKSFADVPQSNLLSNGNFLAQAYTILAGLNVVIQSLSSQELMANQLNALGGAH
QPRGVTTTILELKEFGVILIQVLEEEIGSAMTIDARQAWK
>Dpu_1460_Cterminal
NGIHELIGGLSQT LKNPEDLPDPQTRLT PQQIKEVQRTWASMRSDRNSIVSAIFIELFRE
NPRSQKYFAK FASLPLESLTSNTDFNQVALVANRLDTIISAMGDKLQLLGNINYMRYSH
EQRIYSPRNAVDRDFEDFGRLLLDTLIAKGIAGDDLDSWKSVLKIFIDGIAPEQX
>Dpu_1455_Nterminal
MQFLKIALFFALVALASSPSCSQAPGTTITSVTTT VTTVTADESDNGLLSSHDRSVIR
KTWDQAKKGDVPPKILFRFIVANPEYQKFKSFAAVPQNELLGNGNFLAQAYTILAGLN
VVIQSLSSQELLANQLNALGGAHQPRGITPIMFEQFGTVTEEVLAEEELGNAFNAEARQAW
KNGIRALVAGI
>Dpu_1455_Cterminal
SKNLKKNPEDLADPQTKLTPHQIHDVQRSWENIRANRNSLIS AIFVKLKFETPRVQKHFKV
FANVAVDLSLGNADYEQIALVADRLDTIISAMDDKLQLLGNINYMRYTHQPPRAIPRQT
FEDFARLLIDGLSASGVSGDDMDSWKGVLTVFVNGVSPKQX
>Dpu_1454_Nterminal
MQFFNVALVFGVLAIASAYSQAPGTTTTT VTTT VTTVTADEGTD SGLLSAHERSVIRKTW
DQAKKGDVPPKILFRFIVANPEYQKMFKS FATVPQNELMGNGNFLAQAYTILAGLNVVI
QSLSSQELLANQINALGGAHQPRGATPIMFEQFGAITEEVLAEEELGIAFNAEARQAWKNG

IRALVAGI

>Dpu_1454_Cterminal

SKNLKKAEDLADPQTKLTPHQIRDVQTSWENLRSDRNSLVS AIFIKIFKETPRAQKHFTK
FANVAVDSLSGNADYEKQIALVADRLDTIIISAMNDKLQLLGNINYMRYSHQPPRAIPRER
FEDFARVLLDVLVSKGVSADDMDMSWKGVLTVFVNGVSPRQX

>Dpu_1459_Nterminal

MAFKFALLFGLVAFASACSQAPGTTTTTTVTTT VTTVSADEGDEGILSSHDRSVIRKTWDQ
AKKDSVAPQVLYRFVTAHPEYQKMF SKFASVPQNELMGNFNFLAQAYTILAGLNVVVQS
LSSQELLASQLNSLGGAHQARGATPIMFEQFGEILTGVLAEEELGSAFNAEAQSAWKSGLA
ALVAGI

>Dpu_1459_Cterminal

SKTLKKPEDLVDPQTKLSGHMIGDVQRSWENIRGGRNAMVSDIFIKL FKETPRIQKFFAK
FATVAADALPGNADYEKQVALVADRVDTIIISALDDKLQLLGNINYMRYTH TARSIPRGAW
DDFGRLLVQSLAAKGVSSDDLDSWKS VLSVLVNGISPKNX

>Dpu_1456_Nterminal

MASFKIVFLLSVLAFACAYKPGTTTTTTVTTT VTTVSADEGNEGILSSHDRSVIRKTWDQA
KKDGDVPPQVLFVTAHPEYQKMF SKFATVPQNELMGNFNFLAQAYTILAGLNVVVQSL
SSQELLANQLNALGGAHQARGATPIMFEQFGEILTGVLAEEELGSAFNAEAQSAWKSGLA
LVAGI

>Dpu_1456_Cterminal

SKTLKKS EDLVDPQTKLSGHMIGDVQRSWENIRGDRNAMISSIFVKL FKETPRIQKFFAK
FANVAVDALAGNADYEKQVALVADRLDTMIAAMDDKLQLLGNVNYMRYTHAARSIPRGVW
EDFGRLLLDV LNAKSVSSDDLDSWKGVLAVFVNGISPKNX

>Dpu_1461_Nterminal

MAFKFALLFGLVAFASACSQAPGTTTTTTVTTT VTTVSADEGDEGILSSHDRSVIRKTWDQ
AKKGDVAPQVLYRFVTAHPEYQKMF SKFASVPQNELMGNFNFLAQAYTILAGLNVVVQS
LSSQELLANQLNALGGAHQARGATPIMFEQFGEILTGVLAEEELGSAFNAEAQSAWKSGLA
ALVAGI

>Dpu_1461_Cterminal

SKTLKKS EDLADPQTKFTGRQIRDAQRTWENIRGGRNAMVSSIFIKL FKETPRVQKYFAK
FANVAVDALAGNADYEKQVALVADRLDSMIAALDDKLQILGNMNYMRYTHAARSIPRGVW
EDFGR LIFDV LSSKGLSADELNSWRGV LGVFLNGIAPKKX

>Dpu_1457_Nterminal

MAFKFALLFGLVAFASACSQAPGTTTTTTVTTT VTTVSADEGDEGILSSHDRSVIRKTWDQ
AKKGDVPPQVLFVTAHPEYQKMF SKFATVPQNELMGNFNFLAQAYTILAGLNVVVQS
LSSQELLANQLNALGGAHQARGATPIMFEQFGEILTGVLAEEELGSAFNAEAQSAWKSGLA
ALVAGI

>Dpu_1457_Cterminal

SKTLKKS EDLADPQTKLSPHMIGDVQRSWENIRGGRNAMVSDIFIKL FKETPRIQKFFTK
FATVAADALPGNADYEKQVALVADRVDTIIISALDDKLQLLGNVNYMRYTHIARSIPRGPW
EDFGRLLVQSLAAKGVSSDDLDSWKS VLSVLVNGIAPX

>Dpu_29

MDFVVEAGANIVRERLEQI LLLFETHPDMQSVFLPFTGVVLDLKKSKLLSEHALRVMGA
VQRAVHRLQEPEKLHAF LSELGRKHEKNGAKLEYIDYIGPQFLCAIRPILGDDWTL DTEK
AWTLLLDYMTATMKESLVEARNASAAESSKPLTLPPSSSSSSSSAATDD

>Dpu_5909

MIRFVAVAAVCDRLFEEHKELNLNLF TKFHKLTRDEQAGSEELAEHAVSVMTTLDESIRS
LDNVDTFILYLHQVGQSHYKIPGFQKEYFWKIRNPFLEAVKMTLGDRYTDNIENIYK VSI
NLVIETLVEGYEKAHQHLGAPSS

>Dpu_8981

MSHTSPNGS QLLEDLQREQPETIASSDDNPIDPVTGLSQRERDYIQQSWHHVRQDLKAAG
LGFFQAFFKAHPDYQLKFKKFADVPADQLADNKSFLVHAMSVMNAV TMVVDSLDDIPKLV
NELKNLGNHGRHNIK TENFRNLTVVLVAFLESALGSQLFPEDVKQSWIKALDVVGVVA
TGLPQPS PDDDAGSAMX

>Dpu_18472

MDVLKSVNVA AVQSTWAI VKADLNTHAPKFYVALLTAHPEYQPMFPTIANVPAGELLNNA
ALKT LSVNVL SKLSELIDGMSNP DGLNAQLVELAKQHKNRGTTRTHFDNLAKVLVDFLAA
NLGAAFTPDAKQAWTATMQGINTVVEANAX

>Dpu_23124

MDVLNSVNVAAVQSTWAVIKSDINTFAPQFYVALLTAHPEYQAMFPTIANVPSGQLLNA
ALITLSVNVVTKLSEIIDSLGNPGALNGKLVDLANQHKQRGTTTRAHFDNMATVLLGFLAA
TLGSAFTPEAKQAWTSTMQGINTVVEASAX

>Dpu_25981

NTSPGSGCPWGFNKQSSVAGTKCPCFGYDNAFASKSGCPYGFKNSDKGCPLGVKRLVEGS
KCPFGYKKEISIGDGKCPAGFSVSTGKGCPIGFKRQLVEGSLCPCFGYTSEDAAAPGCPHN
FKKGASGCPLGLKKGLVAGSKCPYGGTSSGSGNGCPLGLKKQLVAGSKCPYGFDTADASA
SGCPHGFKKSANGHGCPCFGIQKSSLKLSNCPYSHGHGSKCPYSQQGSKCPSQHGTKCP
HKCGHSAKTQTREYAMKEADRTLQGTWRIAKKNGNIAPKAFIRYFKLKPEAQKQFAAFA
DVELADLPTNSHFLNQVYTCLAGLNAYMENLGNPKQCPHLNSPVFKAVKPDLLKLFGEV
MFTVMEEELGQSFSTEARKAWKDGLIACDVAFRKSX

>Dpu_13500

MLVHYILTALVVFLETRPGRAECPRGYKASSSQRDPSTKSESQRILENFLNERDEATI
RSTWNTAKKNGNIGPKTFLRYFELKPEAQKMFPAFAEVDHMKLPTNEDFLAQANCVSGL
NSYVEHLGKNPKNCPPIAKAKGKYHHEDLKLGLVTLMGVLEELGKGFTDETKEAWKKGL
RAMNEAVTKRPNPSRRX

>Dpu_18245

MSMFILVLCVAVLSLSAGQSILFKEGTQGPFIATTTVTTFDFNPAGVPRARSSACDRTHY
DSDKKYGSLSQMDVDIVNSWNLKKGNGFAPKVFIRYFKAKPESQKLFPAIANVSI TDL
PTNPDFLNSAFTCVNSLNYLIPFLKYDHPERCPSFPKQIKDNYNEVDVKKLGSIWMMAMQ
EEMGSDFTNDVRDAWKAVMAVIEYVSKX

>Dpu_18401

MSWLILVFCSILSLSAGQSPFNEGIPGGVESGGYGDSTMRGSGPFVTTTTVTTILDFNPA
GMSRSMRSYPSASERSYHYGDKKFGSFSQKDVVDIVNTWNTLKRGRDFAPKVFIRYFKAK
PESQKMFPAFANVPITELPTNHDFLNSAYTCITSLNYLIPYLKFDHPERCPAFPKHLKDK
YNAVDLKKLGSIWMTAMQEEMGNAFTNDVRDVKKAVMAVIEYASKX

>Dme_524369_1 globin 1

MNSDEVQLIKKTWEIPVATPTDGAAILTQFFNRFPNLEKFPFRDVPLEELSGNARFRA
HAGRIIRVFDESIQVLGQDGDLEKLDEIWTKIAVSHIPRTVSKESYNQLKGVILDLVLTAA
CSLDESQAATWAKLVDHVYGIIFKAIDDDGNK

>Dme_649669_3 globin 2

MSQISKLTHISRISQNNQSDGSDKDFRRANFPVYPKPLPDRDLSYKADENEFTMVEKAS
LRNAWRLIEPFQRRFGKENFYSFLTRNEDLINFRRKDGKINLSKHLGHAMAMMKLMSKLV
QTLDCNLAFRLALDENLPTHKNGIDPDYMRMLATALKSYYLASSVIENHNCSLSNGLA
RLVEIVGEYAVVDEARKRAMSTALRRTVDDAGNRIVKVALGT

>Cel_500166_3

MEALNRFSEQEEDI LRRSWKVL DKNLNHTAYNIFEMIFNQSPDTRQLFPFMKFNTGGRSK
EIEFHALRFMQVLESVVKTLDNPETLNPLCDNLGRVHGRLSESRGFRTHHWGVFIECTLF
HFRKVLGQDITYFHRMDALDKVIINWRIIRLLIKQMKRGFNTDIKNRQASRELEESNKAS
TSSSPISQESCSLGTGLRKNRQLMFLAVPGAAAATMSQSLTLPAINNNNYRKGSDSRLS
TVSALSATSVSPAIERAPSTGLLRPLKEFARRRFINHF

>Cel_001129775_1

MPSAARQLVSNMLSFLSSPQTSVSEKGFSTPQANRKNKVNNGNRRMSDSNNVAYSNG
SKKVFEGDIEKWEPNVYEKELLRRTWSDEFDNLIELGSAICYIFDHNPCKQLFPFISKYQD
YQGDEWESKEFRSQALKFVQTLAQVVKNIYHMERTESFLYMGQKHVKFADRGFKHEYWDIFQ
DAMEFALEHRLSIMTDLDDNQKRDVAVTVWRTLALYTTVHMRNGFIDGLKGVNRFPP
LV

>Cel_492074_1

MLLTTRPRFLFRSHSASSVSEKGFSTPQANRKNKVNNGNRRMSDSNNVAYSNGSKKVF
FEGDIEKWEPNVYEKELLRRTWSDEFDNLIELGSAICYIFDHNPCKQLFPFISKYQD
EWESKEFRSQALKFVQTLAQVVKNIYHMERTESFLYMGQKHVKFADRGFKHEYWDIFQ
DAMEFALEHRLSIMTDLDDNQKRDVAVTVWRTLALYTTVHMRNGFIDGLKGVNRFPP
LV

>Cel_497442_2

MESLSHLTPIDREILNKSXGIVSKDMQQVAVNIFQMI FEQAPDAKLMFSFMMKYKEDK
SNEFIFHAVRFLQVIESTMTHLEDPAQLDAVFLNLGKI HAKHEEQLGFSAHYWSVFEKCV
LFHFRKAMKSHNKFHKRNEMSF AEIDSAIILWREVLRFIIDRMKVGYSSES GAIRKSNKQA
INGQHSISGDDSGLSGLSVETKQDLTQVKISAFSGRQHRTASTRHS LAYITATPAEPMT

SSTCNFLKTLASMSPKCVRRRFSASPNARMHV

>Cel_492188_1

MGSSTSTPAPPPKKNKPEGRKADNQILNSYQKSIVRNAWRHMSQKGPSNCGSTITRMMAR
RKSTIGDILDRSTLDYHNLQIVEFLQKVMQSLDEPDKISKLCQEIGQKHAKYRRSKGMKI
DYWDKLGAEAITETIREYQGWKIHRESLRAATVLVSYVVDQLRFGYSRGLHVQGSRETKED
DEE

>Cel_001300599_1

MISLSVPRTSRSLSPAMARSAPSSPMTRYGPGGLEPMSGRDRSGSISLSPRNVIPVCSQL
TPSQVSVVRRSWRHINTKGLIIVLTRCFSRLESNCPIVSQCFQSATYSLSTNPNNGVRTVA
DHAKYLLQLLDKIIIEGDVDSEFLREIGANHVCLKHESGFSTQEWDRFQEIMVEVILKQDG
VKQSKETSRAWRL LICSFIELIRDGFDAQVRQFRRKHSFNHVVQYFENIEKRVGVCPSRK
ISLNVDPRTPPNGVRKYSQY

>Cel_495806_1

MSGRDRSGSISLSPRNVIPVCSQLTPSQVSVVRRSWRHINTKGLIIVLTRCFSRLESNCP
IVSQCFQSATYSLSTNPNNGVRTVADHAKYLLQLLDKIIIEGDVDSEFLREIGANHVCLKHE
SGFSTQEWDRFQEIMVEVILKQDGVKQSKETSRAWRL LICSFIELIRDGFDAQVRQFRRK
HSFNHVVQYFENIEKRVGVCPSRKISLNVDPRTPPNGVRKYSQY

>Cel_001300484_1

MLKQKAQSVKYPRNSRPRKEMISLSVPRTSRSLSPAMARSAPSSPMTRYGPGGLEPMSGR
DRSGSISLSPRNVIPVCSQLTPSQVSVVRRSWRHINTKGLIIVLTRCFSRLESNCPIVSQ
CFQSATYSLSTNPNNGVRTVADHAKYLLQLLDKIIIEGDVDSEFLREIGANHVCLKHESGFS
TQEWDRFQEIMVEVILKQDGVKQSKETSRAWRL LICSFIELIRDGFDAQVRQFRRKHSFN
AHVVQYFENIEKRVGVCPSRKISLNVDPRTPPNGVRKYSQY

>Cel_506867_2

MYRQLIEQHIEEVNVQRDARPQNSLKEPTIEHKEQKKSSEKPVTIIVEKPAKVAEKPKT
TSSEKENAVPVDPLARKIIDETSRLSDRQRDVLQKTFAPILQDCVRNGLKIFVRLFSEYP
RYKLIWPQFRAIPDSSLMNAVELRRHASVYLNGLGKIDSMRDEEALGKSMSRIAVAHIK
WNVQRNHVIHMIEPVLEVVKCNGYQLDDETRQAWTVLYQVIADLIEVFRCLAND

>Cel_492107_2

MLNSYRNGKRKSYSAGSITSSSNSTKKNPDFENFFEEPQIKVMLDARRESFLRHSTSAEV
FPVAPPDLEVLERSTKTTNPPTPRTVRKQLRFDIPEVHISFETNEPRKVTFEGGSTSNS
IVPPPIIDWDSTPKMNSVPEIEPILCFDDEVSVTTRRLSANERQNVIAQRRQSNVQLV
QSMGRTTTTTLIPLTCAQIHLVLRALWRQVYTTKGPTVIGASIIYHRLCFKNVMVKEQMKQ
VELPPKFQNRDNFIKAHCKAVAELIDQVVENLDHLDNVTGELMRIGRVHAKVLRGELTGK
LWNTVAETIIDCTLEWGDRRCRSETVRKAWALIVAFVIEKIKAGHHEQRKMLGTRQSLP
SIRTPSMERF

>Cel_001129755_1

MQKIWEWIVDKKNTIRKWRSSNIRRNTEPDDL SAYRNGKRKSYSAGSITSSSNSTKKN
PDFENFFEEPQIKVMLDARRESFLRHSTSAEVFPVAPPDLEVLERSTKTTNPPTPRTVR
KQLRFDIPEVHISFETNEPRKVTFEGGSTSNSIVPPPIIDWDSTPKMNSVPEIEPILCFD
DEVSVTTRRLSANERQNVIAQRRQSNVQLVQSMGRTTTTTLIPLTCAQIHLVLRALWR
QVYTTKGPTVIGASIIYHRLCFKNVMVKEQMKQVELPPKFQNRDNFIKAHCKAVAELIDQV
VENLDHLDNVTGELMRIGRVHAKVLRGELTGKLVWNTVAETIIDCTLEWGDRRCRSETVRK
AWALIVAFVIEKIKAGHHEQRKMLGTRQSLPSIRTPSMERF

>Cel_001254331_1

MSVAHNKKGLFRRSQSTRSPPEHSPPCSPLIRRQRPATLYVKHSPRASHVAQFPKIKISI
PTVNVHKMSLTNAECLTVHQIIPPPQQQQLYPNSMVPPSDRFTRSASSSPRRGGANGGGA
VVVGPDPQLRRRESVLQMLNRIWKSDECTRRPSVQKSNALDVPTEPLSRQLSLSDTNVTP
QISRCALNLSVKQKLLRQSFNAMNSGGTFLKLMKIFRLETKCPDMRSIFLTAFVNSLSRERQ
TPPLVKTEYDHCKCMVGI FERLIENLENINEQLTMI RHYGEKHAQMAESGFTG
AMIEQFGEISV FVIGSQDVVKFNHETVKAWRLLLACVTDEMKGVGFDRMSRINGRRNSCNP
PSTVPSGN

>Cel_001254332_1

MSLTNAECLTVHQIIPPPQQQQLYPNSMVPPSDRFTRSASSSPRRGGANGGGA VVVGPDPG
QLRRRESVLQMLNRIWKSDECTRRPSVQKSNALDVPTEPLSRQLSLSDTNVTPQISRCAL
NLSVKQKLLRQSFNAMNSGGTFLKLMKIFRLETKCPDMRSIFLTAFVNSLSRERQ
TPPLVKTEYDHCKCMVGI FERLIENLENINEQLTMI RHYGEKHAQMAESGFTGAMIEQFG
EISV FVIGSQDVVKFNHETVKAWRLLLACVTDEMKGVGFDRMSRINGRRNSCNP PSTVPSG

N

>Cel_001254334_1

MLTLLGKIAHFREPLSRQLSLSDTNVTPQISRCAKLNLSVKQKLLRQSFNAMNSGGTFL
KLMEKIFRRLETKCPDMRSIFLTTAFVNLSRERQTPPLVKTEYDHCKCMVGI FERLIEN
LENINEQLTMIRHYGEKHAQMAESGFTGAMIEQFGEISV FVIGSQDVVKFNHETVKAWRL
LLACVTDEMKGVGFDRMSRINGRRNSCNPPSTVPSGN

>Cel_001254333_1

MLNRIWKSDECTRRPSVQKSNALDVPTPEPLSRQLSLSDTNVTPQISRCAKLNLSVKQKLL
LRQSFNAMNSGGTFLKLMEKIFRRLETKCPDMRSIFLTTAFVNLSRERQTPPLVKTEYD
HCKCMVGI FERLIENLENINEQLTMIRHYGEKHAQMAESGFTGAMIEQFGEISV FVIGSQ
DVVKFNHETVKAWRLLLACVTDEMKGVGFDRMSRINGRRNSCNPPSTVPSGN

>Cel_510079_2

MGOENSKCPHQSLAEKRYKVERPKTKKVSSGSATERCLSTQSDEKNAARRLTVSSCDVSA
EDDLPEIKKLSVCEPNEDEATSMTNAAAAGGAKSKCKHFLTRRERILLEQSWRKTRKTG
ADHIGSKI FFMVLT AQPDIAIFGLEKIPTGRLKYDPRFRQHALVYTKTLDVFIRNLDYP
GKLEVYFENLGRHVAMQGRGFEPGYWETFAECMTQAAVEWEANRQRPTLGAWRNLSICI
ISFMRRGFDEENGKKKQYSYVQGFSSNRARRSISPYAPGVH

>Pca_014676299_1 PREDICTED: neuroglobin_like [Priapulus caudatus]

MGGTCSAAEAPVTEETDKRVCCVVTEQERRLLRKTWRYMTNQYDLMELGCEVFLMIFELN
PEAKKMFPCR DVEGEELLTNHDFKGHASRFMQAVGAAIDHLDDLQLELEPLLCTLGRTHTA
NYRNRGFVPHFFDEFTAAMIQTWQKKGKGRKFTPDVRAAWLTMLS FVAAAMKSGYNEYSQA
YSRY

>Pca_014672816_1 PREDICTED: hemoglobin subunit mu_like [Priapulus
caudatus]

MGASCSSSAVKVTEKQNCQCLVSEQERKLIRRSWRYLTNHYDLTELGCEVFLAIFELNPEV
KKMFPCR DVEGEELLKNPDFKGHASRFMQAVGAAVDHLDNLQLELEPLLCTLGKTHNTYK
DLGFAPDFFDEFTAAMIQTWQKKGKGRFAADVCAAWLNMLS FVAAACMKCGYRKTATTIGY
ENCHENVQLNSKQVFNAGMVNLGRT

>Pca_014676644_1 PREDICTED: neuroglobin_like [Priapulus caudatus]

MGGGCSVVATSGRPPDIELTEREKALIRRSWKLLEDSTMTMIGGEIFLKIFELNNEVKAL
FPFRDVTGEALKHDPDFKGHASRFMQALGAAVDNLDNLRHSLEPLLSALGKTHTRYTDRG
FTPAFFDEFTASILHVWRERLGRKFTPEVRAAWLAMVSFIATAMKDG YRKYRVRWARVEEGN
ERRGSATGPVDRHKFDEEILKLMAEKETK

>Pca_014677944_1 PREDICTED: uncharacterized protein LOC106817732
[Priapulus caudatus]

MGCNQCKPCSRGSAKKAIVFGNGHTFSEKEKRRFFKHNRTTANGIATIDATHYNNMASG
SALADGTPDDHERKESTSVKSLGFSAGSESVTAAQTLHLFSDVKILPEMSGEEKAVLL
NTWKIIRENVARVGVIAFMGMFEDNPNVKA AFISLRNMDEVELRNSKELRGHALRVMGV
DKVTRRLNEPEKIEQLLRTLGTKHVGYGAKIQYIDLLGSQFILAVKPILEDHGSDASVE
NIWKKLFSIIGFHLKNGMQDNLQRSKDSPGGRLLODKRASHDSRS

>Pca_014670957_1 PREDICTED: globin_3_like [Priapulus caudatus]

MGCATSASPGNRSGAQSTNGSVDAPLPEPVDPRPLPTARQRF SIQKSWKAIARNMETTGV
TMFVKLFLQHEELLAVFKKKFKDVPVDQLAKSPELENHASQVMSTLDEAIHSLNNLNFFF
ELLYTIGVTHRDIAGFRPEYFWASALDNSW

>Pca_014670958_1 PREDICTED: neuroglobin_like [Priapulus caudatus]

MPVYLFPSKLD FCKTDLFAFCRLFLQHEELLAVFKKKFKDVPVDQLAKSPELENHASQV
MSTLDEAIHSLNNLNFFFELLYTIGVTHRDIAGFRPEYFWRIEQPFLDAVKETLGERYTH
NMEQIYQVTIKFIIDTLYKGYTKGAVK

>Pca_014679465_1 PREDICTED: hemoglobin subunit epsilon_like, partial
[Priapulus caudatus]

DLGDLLRRDPRLKGGGRFMQVVGAVLENIDNADTAIRPLL FQLGERHAYYDGFKPEYF
EQYKGM MYMCEQNFQDKLTPKLRDLWTKVFNFIIDSVRDGCVEALRHREDDRQYSVARD
AQVTRGTI

>Pca_014674382_1 PREDICTED: uncharacterized protein LOC106814565
[Priapulus caudatus]

MTLLTNPDPYRKYATFAEFGKDESQMRSSSELEVFGT SVLMAIDKMVASFDDVDAGVAHL
QSVGKLHAKLDGFTPNMFANMKDAFMYAASVAFEDRYNDQVEQCFQTLDFDCIKYMQEGL
QS

>Pca_014664117_1 PREDICTED: leghemoglobin_2_like [Priapulus caudatus]
MATLTSKEKDLIRESWAVLHANKAQNVAFFVLFRRKNPTYKDKFSHFRDQSLDELAKGG
NPKLQVHASSVFDSFEAMVENIDDVAMTSSLWDNTSRRHYVSHGIDFHHFEDMETAFLEI
LKNKLGAKMTPEIDASWRKLFGIMNVNFKSNAAKLGKSTT
>Sko_30018631m
MGCSASTSTENGGSSFAHLDDSKDLFNGRQRRIVRKTRWRPLANDMTGIGSKVFLRIFELNP
TVKQLFPCRDLNNGEELQKDLNFKGHASRFMQSVGAAVDNLDNLETSLAPLLTNLGRSHIH
FRGFEPHYFDAFTEAMMYVWEQELQDRFTSEVRDAWKMMFEYMMGKMKVGYIMSKEEQMN
ETNSQKIKTIEX
>Sko_30018621m
AQNQTICLGKGLLLLTGLCRLFAYSIYIYSTYVLRMGCSVSTSTENGNFVQLSDSV
DIFTERQRRIVRKTRWRPLANDMTGNGTKVFLHIFEMNPKVKQLFPCRDKTGEELLKDLNF
KGHASRFMQSVGAAVDNLDNLETSLAPLLMNLGKSHNHFSGFELNYFDSFTGAMLHVWEL
ELQDRFTPEVMEAWKLVFDYMMGKMKDGYITRRDEKLNETNEQNKIIVX
>Sko_30022689m
MGCTPSINERDFQTPVDDKHLDDRQKRIVRKTRWRPLANDMTENGQKIFINIFESHPEIK
YMFPTRDIEGRDNLNANPHFRMHSSRFMQSVGAAIDNLDNLDNALRPLLKLAETHVRFK
GFKPDYFDAFEEAMLSVWQEEELGQRFTEVEESWKLFFYIKDCLKEGYDIAMNEKTSGE
LNNSDFINQX
>Sko_30033150m
MGCSNSSHNCVSPKKEDESMQLPPTSLTDQHRVILLDSWKVIQEDIKVGVMFMGLFET
HPECKEVFMPFKELQGDDLWSSALKAHGLRVMVIERVLARIDSDEKIEEHLKALAKKH
VEYGANSDLVRLFGPQFIFGSMKRQLHKSWSDEMQUAWTVLFDIIYHMTTNMVPQEPENN
NIAKRQKSSRKSRTKYMIDNGHSQX
>Sko_30033142m
MSRFTSRLSSSTLDNFEAISNLGWEKKLYEHSSTRTRFRTRKRKLTIHNYTIYPAGLFLLT
VYNAMGCGSSKINGNVVEEKPELTKEQKDTLIQTWQNLHADLERIGMLMFMGLFEHNPEI
KEFFVGADSRDMKTEELRYNEKLQEHGIRVMGLVEKIISSMGFEDEKIDQMVVDLGRHL
GYDVHIPFIDLFGRFVFAIKPTLHTHWTANVEEAWTQLFKYIGYLMRYGYHTKLQQVQK
KNSX
>Sko_30015006m
SEQISELRRTPKPLACDLTGNGAQVFLQIFAINENIKILPFRYVPVDILSQNEVFRGHS
RRFMQAVGACVENLENLDGDVTTLFVGLGKKHIFHEGFKVDYFSTYVTSMQTVWDIALTG
HHYDKQTKQSWTQIFEFVITRMAEGYHIAMDEQEAKKKNQLAHENGKIIIX
>Sko_30027052m
MTRNGGKIFLQIFAVAPHVKDLFPFRYVPNDMLQONEIFKMHGRRFMQSVGAVIENIDNL
DGDISILLHNLGKRHTDFDEVDGAYFDIYTDMMHTWRSSLGNELFTPDVGGVWHKLFDF
IINCIKDGFLAMKNKKNNSDEVKKIX
>Sko_30033155m
MGNEVAKSSRSSTSQSLSKEQEKILVQTWLSIRGDLERIGLLMFTGLFEHHPEAKVMFGL
SDTAMSPKDKENTALIKEHGLRFMNVVRDVLTLISEKNGSQAECVLIDLGRRHCSYNADI
NLIDVFGQQFIASIQPTLTGSDWKKVEDAWIQLFKYIAFTMKQGLAAELIDKSLKLNKGP
X
>Sko_30035097m
MGCTSSAASDRPSKNDPLDPPPPQELDPRIPLTARQKFSIQKSWKAIQRNMEGVGMDIF
IRLFKAHPEYQDLFPEFKGMSEEKLRNSINFETHVGI FMNVIDECIDSLEDADHVINLLT
KKGRKHANYGVKPEFISNRFX
>Sko_30033204m
MFGTHPQTRFEFFNFRGTSDDPKNTQRLREHGLRFMSLVKKILVFIDEKPRLDAMLLDLGR
RHQEYKADFNLIDVFGQFILLSVRPTLKHSWNPDESAWAQLFKYISYMMKKGMMQTDKN
KX
>Sko_30007810m
MESTESNTPVKETTEINPDDIPDEVITILTPKEVKAISESWKVVYAKKKENGVALFIRLFQ
SVPGSKSLFKNLDGIDDEEKLNRHPRKHAHGFVMSSVNSLIESLEEGLLVQLKDLGS
SHSKNKVTSSHF DALGPV I I WLLQKENGDSFTPAVKNAWLKGWGMKSVIVGSLEEAYAK
MKT X
>Sko_30007813m
MDRNTNCSDNWTASSHCHISTSYYTKTQFPSTKMSLSAGEIKLVKDSWAPVYANKKESGI

ALFVRLFSENPGFQSQFRYLDGVSGLAAIEKTPALADHGVKVM DTVNSWVGS LGDAPALV
KQLTALGTSHIALKVT PANFDAMGPVLLWTLQEKAGGAFT PAAKDAWAKGWDLMKSHIVK
ALQGX

>Sko_30007818m

MALSAGEIKLVTDSWTAVYANKKANGVALFVRLFSENPGFQSQFRYLDGVSGLAAIEKTP
ALGDHAVKVM DTVNSWVGS LGDSSAMVAKLTALGTSHIALKVT PANFDAMGPVLLWMLQE
KAGGAFT PAAKDAWAKGWDLMKSHIVKALQGX

>Sko_30007821m

AKLAKLEVPNAVTTLTPSEAIQSTWLFVYEDKEENGVELFVKLFTEHPDYQALFGYLE
GIVGIENIKNVPFLRVHASHVLIYLNMTLES LNDGTILVELLKT LGYTHVGLNLTPEHFD
ALGPILISLLQEKGGD

>Sko_30018543m

MADPVTTLTSDEVAIKSSWSAVYDKKKESGVTLFVKLF TENPSFKSQFGYMSGVADGDM
KTLPALENHGKVM DTVNSWVGS LGDSSAMVAKLTALGTSHIALKVTEDNFNAMDSVLMY
TLQEQGGS AFT PAAKAAWQKAWGVMKSVIVGALKGX

>Pfl_1g287

MGCNVSSASQAREITNRD TDI FTDRQRRI VRKTWRPLANDMTGN GTKVFLRIFEME PKVK
QLFPCRDK EGEELKDMNFKGHASRFMQSVGAAVDNLD SLETSLAPLLLK LGRSHTNFSG
FKPDYFDIFTRAML DVWEQELKDRFTSEVKESWLT VFEFMMGKMKEGYMQAYTEEMNAKN
LQKSNGTVVDE

>Pfl_1g25972

MGCTPSTKDQDCHDPASDNSLLDERQRRI VRRTWRPLANDMTENGKRI FLRIFESNPEIK
YLFPTRDVEGKENLYANPHFRMHASRFMQSVGAAIDNLDNLDNALKPLLIK LAKTHVRFK
GFKPDYFDVFEEAMLYVWREELA QWFTEEVVEEAWKLLFFYIKECLKTGYEEAMEEKGTLI
QGNENTE

>Pfl_1g22622

MYSCERTDLTRNAFITFRLPQSKGESLTRPIHRQTGT LKEKQIAINFFLVTSIETKSDL
KGMGCAESIAMATNGKVVKTNAGKFM SKAPRFTQEEIHELRRTWPKLACDTTGN GGQVF
LQIF SINPAVKNLFPFRYIPNDILSQNETFKMHSRRFMSAVGACVENLES LDGEVTELF I
DLGKKHVKFDGQVDYFTSYIASMQFVWDEL TGHHTNHTRELWTQIFQFVVTRMKEGY
NIAMEEKKAQETVQETSKI

>Pfl_1g11486

MGCKNSTEQEDGELKLSNEDRSLIAESWKELHVDLERIGMLLMFMGMFDTHPETRSFFGFS
NNSTIAEDPKHVKRLREHGLRFMGLVKLLT CIDDNDRFDSILLELGKRHIDYNADISLV
DVFGEQFVISIRPTLKHCWNPKIEEAWTQLFKYIASMMKRGMLERAAAT

>Pfl_1g11488

MGCRTSRADGEKPRLSPEQKLLVADSWKELHIDLERIGMLMFMGMFDTHPETKSYFQFAE
EESESTHAANVQRLREHGLRFMSLVKLLCHLDNKEQFDNMLIELGKRHHNYHANIDL VV
VFGECFILSIRPTLKHSWNPVVEEAWTQLFKYISYMMKRGILRQDKSEKKDKSSK

>Pfl_1g11487

MGCRNSRTEGERPKLTPGQKALVADSWKELHIDLERIGMLMFMGMFDTHPETKRYFDFED
KVKATESTHAENIQKLRHGLRFMSLVKLLCHIDKKEKFDNMLLELGRRHGHYQANVDF
FAVSRYLPFDST

>Pfl_1g11484

MMSSEDPKVVQKIQEHLRFMSIARKLVANLDDPEKFDAILLDLGRHHTYQADINLVDTF
GQHF IASIRPTLKNWSEAVENAWEQLFKCIAYRMKQGLTEAATEAIAAEDILQ

>Pfl_1g11494

MFKNLGEFDDLEDMRESQQL ENHASLVMYTIDEAIASLEDVDFVVELLKKIGRTHKRMDF
HPELFWDIYHADIRLPEKLLMTPVDNSVYVAYRPVHF

>Pfl_1g698

MSSKWNNYISRSLKLEDSFEEVFVRQSSDLKVD TYRRSSTWKALRDTWSVIYANKREN
GAAFFLKLFEVHPDYKKLFKNLEGINLDEELAKHPRLKAHGLRVMASFNSLVENLDDAEV
LVQLLVDIGISHAKHKVKEENFNELGGIVLWLIETKSGNSYSEYAKEAWTKAWGVMKPIM
VKALNDASEEQQAQLKTG

>Pfl_1g699

MLIGLPEDTIVLCRCRNLSPMTLEYSAYYRSFQRSVQDLKKKNYLNLYLFRIRIXLFSE
YPETKQRFKNLFLDDEKLKGHPMTKAHAFRVMASLNSLVETLDDAEVLVELLVNLGRT
HKVKNVTESEDFDDTIPNHPRYLSTRMIYPCR

>Spu_007045
MDTIINNHDDDDVIDETTGLTKQQKALIKKSWTYVLEDKLRIGVIFIKLKFKAFFPASQQLF
EKLKDYTDFEELARNKKMKAHATRVMMAALTSLVENIDQPDILDELLRNTSVTHYRMRMP
HYFEDLGGVIIIEALVENLGDKFTPKTKEAWLIYYGYMCRIMLEEMEELEPSNDDD
>Bfl_96953 fgenesh2_pg_scaffold_295000036
MGSLSAKEDGTPDDVTGLTANQIRHIRETWQVVLSNKRANGFAIFRILFTDYPFTKKLFR
SMDQVDIDVPEQFEKNIALRAHITRFLHSFDTYVSNLDEPADLQQLLYDTGKSHLRHSVK
PEYFDALGNVLMKGLTAVLGKDFTEEVQGAWGTAWGFFVIHLKQGLEAMRHGAETNGNA
AGTDE*
>Bfl_92354 fgenesh2_pg_scaffold_222000023
MGGALGKPLSLVKTLLWKVLFSSWVVKPIETPNDVTGLTPTQVRLQQTWKVILLHKKQNG
FLIFKILFTDYPMTKKLFGIDKVDPEQYKTTSMRAHVTRFINSFDSFMECLEDEPEALK
SLLYDTGKAHLRHNTKPEHFDDLEVMMKSLKAVLGLKFTESEVEAWKTAFAFFVHLLKM
GVEDGLRGREKNTSVVETVEEREIRNDVSLICGKRTAR*
>Bfl_72350 fgenesh2_pg_scaffold_38000007
MPKGPSISRGDHMGCSASMTGMGRAGPALPEPEAPPVDPRLPLDARQKFHLEKSWKSVA
RNIDRAGMFMFLRFRDCPEMLEKYPELRGMDDQEELRNSQFLQEHSQRVLDADFHTIDS
LDDVDYVIQLLKIGQMHADLELKPDDMWKLEQPFLLAAVAECLEDRYTPKFQEIYSKLIT
FIIIEHVNGFDPH*
>Bfl_99970 fgenesh2_pg_scaffold_351000006
MGCSASMTGMGRAGPALPEPEAPPVDPRLPLDARQKFHLEKSWKSVERNIDRAGMFMFL
RFRDCPEMIEKYPELRGMDDQEELRNSQFLQEHSQRVLDADFHTIDSLDDVDYVIQLLK
KIGQMHADLELKPDDMWKLEQPFLLAAVAECLEDRYTPKFQEIYSKLITFIIIEHVNGFDP
H*
>Bfl_77082 fgenesh2_pg_scaffold_68000096
MGANMGCSNSKKMSHESESANSNGDSTPPKSSTPSALDERLPLTQKQKFLLLKSWKGVARQ
ISQCGKTMLIRLFKDDPQLMAVFNQKFRHLRERDADVLYQDAILDAAHAATVMEALHEAIT
HLDDSVFVMKVLHDVGMKQRYNVDPVSFLKVEKPFLLTAVSEVLGDRYTKNMEEIYTITI
KFILATLSEGATMELTEDEQKNLGRLLWRPPGRVHKFVRPEKVAIIVDAQSEENGVH*
>Bfl_111239 fgenesh2_pg_scaffold_861000006
MGCEMSTDGQALSSVIRKDRSDLYKSPGVGDREDWRLPLDAWQRFYLQKSWKTVARSDQ
AARTVFLRMLQDNPGLRQKWPRISLLTEEEIPTSPYIKFLGERIFDCLDYIIDNLGDLDH
VISELTKLGRQHSMDMNMVTPEDVWAIEAAFLAGVQECLEDRFTIKYEEIYSRFIVFVIET
MVIGFDPH*
>Bfl_66937 fgenesh2_pg_scaffold_12000092
MGCEMSTDGQALSSVIRKDRSELYKSPGIGDREDWRLPLDAWQRFYLQKSWKTVARSDQ
AARTVFLRMLQDNPGLRQKWPRISLLTEEEIPTSPYIKFLGERIFDCLDYIIDNLGDLDH
VISELTKLGRQHSMDMNMVTPEDVWAIEAAFLAGVQECLEDRFTIKYEEIYSRFIVFVIET
MVIGFDPH*
>Bfl_74222 fgenesh2_pg_scaffold_49000078
MGSGASRPTPRKRKAKKGPLPSPQPPKPLDPRLKLDAAKEKFFLEKSWKTVARNEVAAMA
MFINLFRSSPEIKDKWPQLRKLSEDEMRSPLYLQKLSVRILGAMDHVIDSLDDPDYLIIPA
LEKLGQMHADMTNPIILPEDLWVNKAFLRQQ*
>Bfl_74626 fgenesh2_pg_scaffold_52000051
MGTIADGEGTELNGYGGDKEPGGGHGGPLTQEQVHGITETWAILAQDPVERGVDLFMKIF
EEDPDLKLFYFADDGRELSREDQRMRSHERVMEAVGAAVDSLGLDTAVVPVLTTEL GAL
HHKYGVQPSYFDTVGAALIYILETNLGDKLTPSIRQGVVLVYAIVGATMKKGMQQAMDHQ
NMAKTRP*
>Bfl_84888 fgenesh2_pg_scaffold_136000004
MSTDRSAVVSLETEGEKATIRRTWAVASRDMMGNANILLKMF E INPDTKKVFAKFRNIPD
DQLRSTPRFRAHVTRVMASISTVNSLDDQEVLLDLFKDIGKKHYPARVPTEYFDVIAGA
ILCMLQRCLGTGYTAEVDSAWTKLYGSLGRHAKDGLREAAAMGTP*
>Bfl_18056 Branchiostoma_floridae_proteins_unnamed_protein_product
MGSWWGKPVNDTPDDITGLTANQIQIRDTWQIVYKKNKRENCFAIFRILFTDHPSTKSLF
RLMDAVDLDVPGFEKNAARAHMVRFMHSFATFMDTLDEPAELRQLLYDLGKNHAKHQV
GPELFDALGPILMKALPIVLDGKFTPEVKTAWLTAYTFMSTHLKEGVEEGQRQLADSKX
>Bfl_16013 Branchiostoma_floridae_proteins_unnamed_protein_product
MSTDRSAVVSLETEGEKATIRRTWAVASRDMMGNANILLKMF E INPDTKKVFAKFRNIPD

NQLQSTPRFRAHVTRVMASIGTVVNSLDDQEVLLDLFKDIGKKHYPARVPTEYFDVIAGA
ILCMLQQCLGTGYTAEVDSAUTKLYGSLGRHAKDGLREAAAMGTPX
>Bfl_15997 Branchiostoma_floridae_proteins_unnamed protein product
MATTMCMQTNKVGLPQHTSQKGTIAGSLHKQSYRRLDSTSGAADIYSKDSMGAFLLTKPFSL
VGRLLWKVLFVSWVVKQIETPSDVTGLTPTQSRVLKESWKMFLSKKRENGFVIFRVLFTDY
PVTRKLFKGVQDLGAPGQLESIITLRAHVTRFMHSFDTYMESELDDEPDLKQLLYDTGK
SHLIHNIKPEYFDVLETVLMKSLRMVFGSKLTPQLEEAWQTAYSHLKVTIKQGLEDAIQK
RDQADTSVVDPNGX
>Bfl_18782 Branchiostoma_floridae_proteins_unnamed protein product
MATTMCMQTNKVGLPQHTSQKGTIAGSLHKQSYRRLDSTSGVADIYSKDDMGAFLLTKPFSL
VGRLLWKVLFVSWVVKQIETPSDVTGLTPTQSRVLKESWKMFLSKKRENGFVIFRVLFTDY
PVTRKLFKGVQDLGAPGQLESIITLRAHVTRFMHSFDTYMESELDDEPDLKQLLYDTGK
SHLIHDIKPEYFDVLETVLMKSLRIVFGSKLTPQLEEAWQTAYSHLKVTIKQGLEDAIQK
RDQADTSVVVTVEVX
>Bfl_98913 fgenesh2_pg_scaffold_331000046
MGTIADGEGTELNQYGGKEKEPGGGHGGPLTQEQVHGIKETWAILAQDPVERGVDFLKMIF
EEDPDLKFLFYFADDGRELSREDQMRMRSHGERVMEAVGGAVDSLGLTAVVPVLTTELGA
HHKYGVQPSYFDGVQLDDRPX
>Bfl_18055 Branchiostoma_floridae_proteins_unnamed protein product
MGSLSAKEDGTPDDVTGLTANQIRHIRETQVVLVSNKRANGFAIFRILFTDYPFTKKLFR
SMDQVDIDVPEQFEKNIALRAHITRFLHSFDTYVSNLDEPADLQQLLYDTGKSHLRHSVK
PEYFDALGNVLMKGLTAVLGKDFTEEVQGAWGTAWGFFVIHLKQGLEDAVRHGAETNGNA
AGTDEX
>Bfl_18783 Branchiostoma_floridae_proteins_unnamed protein product
MGGALGKPLSLVKTLLWKVLFVSWVVKPIETPNDVTGLTPTQVRLQOTWKVILLHKKQNG
FLIFKILFTDYPMTKKLFGKIDKVDPEQYKTTSMRAHVTRFINSFDSFMECLEDEPEALK
SLLYDTGKAHLRHNTKPEHFDDLEVMMKSLKAVLGLKFTESVEEAWRTAFAFFVHLKM
GVEDGLRGREKNTSVVETVEERAICNAVSLITLYGKRRX
>Bfl_23417 Branchiostoma_floridae_proteins_unnamed protein product
MGSLSAKEDGTPDDVTGLTANQIRHIRETQVVLVSNKRANGFAIFRILFTDYPFTKKLFR
SMDQVDIDVPEQFEKNIALRAHITRFLHSFDTYVSSLDEPADLQQLLYDTGKSHLRHSVK
PEYFDALGNVLMKGLTAVLGKDFTEEVQGAWGTAWGFFVIHLKQGLEDAVRHSAETNGTA
AGGSGKVAKNIQANSVISRFSPIQMLRLX
>Bfl_14796 Branchiostoma_floridae_proteins_unnamed protein product
MGLTSEDKSAVLDSWAKMSGPTFQDAGEKVFLLLLKTDSTKALFPKFRDIPYDQLAGHPD
VRDHGGKVMQVLDLDFIKGLDNGGDGAVQKVGLLHKGVGVSVDNINLMKPVLMTLLGELGC
SSAAGAWENLWARFMDVHRTCYX
>Bfl_26332 Branchiostoma_floridae_proteins_unnamed protein product
MSLSAADKKAVALDSWAKMSKPSFQDAGERVFLKLLKDKSTKAMFKKFKDIPRERLPGNAA
LREHGGKVVQALDDFIKGLDGSGETVNRVGRVHKAAGMTNDNINLMKPVLELLELDEAGC
GDAKAAWDLWNLFMTVHGDGCX
>Bfl_31701 Branchiostoma_floridae_proteins_unnamed protein product
MALSAEELATVKQAWAKLTASSFEDAGEKVFLLALLKDPNIKANFKKFKDIPASLPGNTD
MRAHGKVKCTVLDKFIKGDGAASKSTGTMHKGLGMSNDQIGAMRGALVAVLNDAGEGGAV
PAWNKLFDFHFMVHKTGYX
>Cin_210686 fgenesh3_pg_C_chr_01q001220
MPLTKIEIEGVQESWEKVSSGGPKTTGLILMEKLFNTYPASIAVFSHLGIPSKPDGAIIV
SDLASIGGVSNAVSLASRIGKLVGLLNNETELKESSTEVGRIHVKYGVVSEHVDDLGSV
LLSVISENQGLSNTSELIGWWSKTWNIIGNYVKTGLKE*
>Cin_248158 estExt_genewise1_C_780125
MPFTDEELKLLRNSWDEVKKGKLMKEVGLHIFTGLLNAAPSLRDLFYITIDLPEEELTIDV
MRENKVVVAHATRIANAISKFIKFLDQPEELEKLLTSLGESHARRQVDPESEFYVAPVIL
SVIGHHLKLPNSPTLQAWVKAYGVLNRNGIVSAMEA*
>Cin_288135 estExt_fgenesh3_pg_C_chr_03q0645
MSLTSEQVLLRSSWQTIGKLGMSNVGLAVLHRLFNVDVPELTPFFHSVLSPTQQTEIEVL
KSNKVVVRHASRVGLSIDKIINLLDNGEELVKYLLFLGKVVHVKRSIPRKYFSAMGPVLLS
VISAVLEKDLDPVMQAWATAAYGVIEKGIIDGM*
>Cin_295168 estExt_fgenesh3_pg_C_780038

MGLTTEEIGLLRSSWNEMKTIGMKELGLLIFHRLFSDVPRIRKMFYNLELPDDETLTLEA
MRSNQKMSRHATRIATSI STYLKLVLDQPEELKTFLNGLGELHAGHNVEPEDFEYLAPVML
AVIGGQLNLNSNSAILQAWVKAYGVLNRNGIVRGMAYK*
>Cin_281017 gw1_02q_1920_1
MKIICGLILFSTFAIIFVSGLNCWTCNVLGGNNVCRRSGGLRTCFNNQVCYNEVRRRGNT
INIRKGCKNSGVCENHIMQSMSTPNPNNQCVNGPNYFCSCCCGNSVCNSNWLTCVTAGSV
APTSTTVPPADEGLKRSDIINIQDSWNTLKGFGYETVGMLVLHRLFNDAPQTRYLFSQL
SLSSNESFTLEQMRNNSRVVYHANRVARAVGRLVDLIELPTNFTDHLVWLGQRHAYHGVA
PVNFDYMGPVLLETIKVNLELPSDSPTLSAWAKAYGVKNGIKDAIATYAEG
>Cin_281665 gw1_1740_1
QVCYNEVRRRGNTINIRKGCKNPGVCENHIMQSMSTPNPNNQCVNGPNYFCSCCCGNSMC
NSNWLTCVTAGSVAPTSTTVPPADEELKRSDIINIQDSWNTLKGFGYETVGMLVLHRLF
NDAPQTRYLFSQLSLSSNESFTLEQMRNNSRVVYHANRVARAVGRLVDLIELPTNFTDHL
VWLGQRHAYHGVA PVNFDYMGPVLLETIKVNLELPSDSPTLSAWAKAYGVKNGIKDAI
ATYAEG
>Hsa_Ngb 10864065 |NP_067080_1| neuroglobin [Homo sapiens]
MERPEPELIRQSWRAVSRSPLEHGTVLFARLFALEPDLLPLFQYNCRQFSSPEDCLSSPE
FLDHIRKVMLVIDAAVTNVEDLSSLEEYLASLGRKHRAVGKLSSTVGSLLYMLEKC
LGPFTPATRAAWSQLYGAVVQAMSRGWDGE
>Hsa_Mgb 44955885 |NP_976311_1| myoglobin [Homo sapiens]
MGLSDGEWQLVLNVWGKVEADIPGHGQEVLI RLFKGPETLEKFDKFKHLKSEDEMKA
DLKKHGATVLTALGGILKKGHEAEIKPLAQSHATKHKIPVKYLEFISECIIQVLQSKH
PGDFGADAQGAMNKALELFRKDMASNYKELGFQG
>Hsa_Cgb 19549331 |NP_599030_1| cytoglobin [Homo sapiens]
MEKVPGEMEIERRERSEELSEAERKAVQAMWARLYANCEDVGVAILVRFFVNFPSAKQYF
SQFKHMEPLEMERSPQLRKHACRVMGALNTVVENLHDPDKVSSVLALVGKAHALKHKVE
PVYFKILSGVILEVVAEEFASDFPETQRAWAKLRGLIYSHVTAAYKEVGVWVQVNPATT
PPATLPSSGP
>Hsa_Hgb_Alpha 4504345 |NP_000508_1| hemoglobin subunit alpha [Homo
sapiens]
MVLSPADKTNVKAAGKVGGAHAGEYGAEALERMFLSFPTTKTYFPHFDLSHGSAQVKGHG
KKVADALTNVAHVDDMPNALSALSDLHAHKL RVDPVNFKLLSHCLLVTLAAHLP AEF
AVHASLDKFLASVSTVLT SKYR
>Hsa_Hgb_Theta 4885395 |NP_005322_1| hemoglobin subunit theta [Homo
sapiens]
MALSAEDRALVRALWKKLGSNVGVYTTALERTFLA F PATKTYF SHLDLSPGSSQVRAHG
QKVADALS LAVERLDDLPHALSALSHLHACQLRVD PASFQLLGHCLLVTLARHYPGDFSP
ALQASLDKFLSHVISALVSEYR
>Hsa_Hgb_Zeta 4885397 |NP_005323_1| hemoglobin subunit zeta [Homo
sapiens] >gi|530407994 |Tad_5255345_1| PREDICTED: hemoglobin subunit zeta
isoform X2 [Homo sapiens]
MSLTKTERTIIVSMWAKISTQADTIGTETLERLFLSHPQTKTYFPHFDLHPGSAQLRAHG
SKVVAAVGDAVKSIDDIGGALS KLS ELHAYILRVD PVNFKLLSHCLLVTLAARFPADFTA
EAHAAWDKFLSVVSSVLTEKYR
>Hsa_Hgb_Mu 51510893 |NP_001003938_1| hemoglobin subunit mu [Homo
sapiens]
MLSAQERAQIAQVWDLIAGHEAQFGAELLRLFTVYPSTKVYFPHLSACQDATQLLSHGQ
RMLAAVGA AVQHDNLRAALSPLADLHALVLRVDPANFPLLIQCFHVVLASHLQDEFTVQ
MQAAWDKFLTGVAVVLTEKYR
>Hsa_Hgb_Beta 4504349 |NP_000509_1| hemoglobin subunit beta [Homo
sapiens]
MVHLTPEEKSAVTALWGKVNVDVGG EALGRLLVVYPW TQRFFESFGDLSTPDAVMGNPK
VKAHGKKVLGAFSDGLAHL DNLKGT FATLSELHCDK LHVDPENFRLLGNVLVLCVLAH HFG
KEFTPPVQAAYQKVVAGVANALAHKYH
>Hsa_Hgb_Delta 4504351 |NP_000510_1| hemoglobin subunit delta [Homo
sapiens]
MVHLTPEEKTAVNALWGKVNVDVGG EALGRLLVVYPW TQRFFESFGDLSSPDAVMGNPK
VKAHGKKVLGAFSDGLAHL DNLKGTFSQLSELHCDK LHVDPENFRLLGNVLVLCVLA RNF

KEFTPQMQAAYQKVAVANALAHKYH
>Hsa_Hgb_Gamma2 6715607 |NP_000175_1| hemoglobin subunit gamma_2 [Homo sapiens]
MGHFTTEEDKATITSLWGKVNVEDAGGETLGRLLVVYPWTQRFFDSFGNLSASAIMGNPK
VKAHGKKVLTSLGDAIKHLDDLKGTFAQLSELHCDKLHVDPENFKLLGNVLVTVLAIHFG
KEFTPEVQASWQKMVTGVASALSSRYH
>Hsa_Hgb_Gamma1 28302131 |NP_000550_2| hemoglobin subunit gamma_1 [Homo sapiens]
MGHFTTEEDKATITSLWGKVNVEDAGGETLGRLLVVYPWTQRFFDSFGNLSASAIMGNPK
VKAHGKKVLTSLGDATKHLDDLKGTFAQLSELHCDKLHVDPENFKLLGNVLVTVLAIHFG
KEFTPEVQASWQKMVTAVASALSSRYH
>Hsa_Hgb_Epsilon 4885393 |NP_005321_1| hemoglobin subunit epsilon [Homo sapiens]
MVHFTAEEKAAVTSLSKMNVEEAGGEALGRLLVVYPWTQRFFDSFGNLSSPSAILGNPK
VKAHGKKVLTSLGDAIKNMDNLKPAFAKLSLHCDKLHVDPENFKLLGNVMVILATHFG
KEFTPEVQAAWQKLVSAVAIALAHKYH
>Nve_A7SFV0 A7SFV0_NEMVE Predicted protein OS=Nematostella vectensis
GN=v1g211636 PE=3 SV=1
MGCSSSLSQANLPRMTPLSEAQKYLVRTWETIEPQKQTVGKKAFLRFFDMNPDYQNLFP
EFKSLSYEELQKANALHGHAKRVMKAVENAVMSIDDVMSFSAYLEELGRRHKTRALKPSY
LEAMHGALMDTLRNLQSQWTEETAEAWNKLFSFISTTMVRGLQSRD
>Nve_A7RHV8 A7RHV8_NEMVE Predicted protein OS=Nematostella vectensis
GN=v1g197347 PE=3 SV=1
MGCSSSTFKPPREPVKIPLSVAQKYLVRTWETIEQHSKAVGKKTFLRFFEMNPDYQKLFPEFATLDQVELEQ
ANAL
HGHAKRVMKAVENAVSAMDDAESFAAYLENL GARHKARALKPAYLDAMQVAYTDTIQDLL
KTQWTDGTAEAWNKLFRFIADTMKHGLSS
>Nve_A7RZB2 A7RZB2_NEMVE Predicted protein OS=Nematostella vectensis
GN=v1g204383 PE=3 SV=1
MGCVVSKNPSTVAKIVPGGEEKLFETSRIPLDAKETQLVRKTWAILGDRQVEVGKSLFL
RFFEEHPTSKDLFPEFRNISNEKIAESPALYGHARRVMKSVDNAVASIENVQVYSAYLYE
LGTRHQTRQLSEEQLKFMGGAFLFAMRLHLRKEWSRATSKAWEKIFSFADAMMRGCKG
>Nve_A7RGQ4 A7RGQ4_NEMVE Predicted protein OS=Nematostella vectensis
GN=v1g196903 PE=3 SV=1
MGCASSTVRPFFIRQPASDTEENTLTSVPLSTRKKLVRESWELIEPVKITIGKRLFTRL
FDVNPNMQDTFPNFKGKELKDILNSRSLYLHAKRVMVAVENAVTVLDDAETFESYLINLG
GRHLPWGVTKDHFGVVGAEAFI WALQDVLGEGCTSDVAEAWIDLYGYIVQAMLEGLQQAKK
GR
>Nve_A7S5B8 A7S5B8_NEMVE Predicted protein OS=Nematostella vectensis
GN=v1g207046 PE=3 SV=1
MGCASSLATQTKLLKGLHPTCETQYLTQVDQLPLTETQKYYIKQSWMGLESNKGELGIEI
FLRLFSENPTLQLMFPEFREYSTLEELKESRSLQGHTRKVMKVVENAVNSLEDGHALMEY
LQELGRRHKTRQIKPTVSNLQEISQAINETFEENLGIKWTVIEIAESWKL LLDYVMAMIIR
GLRSP
>Nve_A7RJ19 A7RJ19_NEMVE Predicted protein OS=Nematostella vectensis
GN=v1g197838 PE=3 SV=1
MGCASKTLTTPHGTEEHLTKKSQSENGNQSFVGNRPRLTERQIKLVQDTRWLLIPSQKK
TAMIFYLKLF TLDPIFKVFSFHTENEGQLEQDERFLFQSRKFMEMINS AVDR LNDISLL
VMILKSLGEVHWT KFKIKPEYYEPVGKALIYSISKGLGSLFNDEIGEAWQAMYDLMSGAM
ISGTKAVQARSQNSL
>Nve_A7RY39 A7RY39_NEMVE Predicted protein OS=Nematostella vectensis
GN=v1g203852 PE=3 SV=1
MGCASASLKAAKAQEAKKKAKVVISLTPEEKAIILETWKLI EPYQRIVGKTLFLRFFKE
HPTYQDLFPFRGLSRDDLKSTRVLYGHASRVMKAIETAVASMDNIQAFSAYLEDLGARH
NKRALKAAHLMASIVDQ
>Nve_A7RWR6 A7RWR6_NEMVE Predicted protein OS=Nematostella vectensis
GN=v1g203304 PE=3 SV=1
MGCSSSVVTQGNMPHL CGLKLECDMTYEQKYLIRETWKFLVSKKEIGVSVYKRFLNMHP

GLQTYFSEFKHIKIDNINGSHGHPRLLMAIDNAV TALGDSDFSAYLVELGRRHHGMNF
RPGPTHFN DL RKCFLSVIEEILATASLWDFQVEEAWNRLFDSITAMILRGIQLAKV
>Nve_A7RWR5_A7RWR5_NEMVE Predicted protein OS=Nematostella vectensis
GN=v1g203303 PE=3 SV=1
MGCGSSVVTQGNMPHLCGLKLECDMTYEQKYLIRETVDNRECVNEKDFLAWRYVCELA AI
FLNMHPGLQTYFSEFKHIKIDNINGSHGHPRLLMAIDNAV TALGDSDFSAYLVELGRR
HHGMNFRPGPTHFN DL RKCFLSVIKEILATASLWDFQVEEAWNRLFDSITAMMLRGIQLA
KLNAFRSCFISVVEILAMAYLWDPEVEQAWNRLFDYITISMLRGIQLAKE
>Sci_5513 scgid8208|
MGTQPSK LQKGAKKFEKGMMLCISGGGDEDCNEQMTSYQLQRYTHSPSENRPNRS PARTL
TPRPCDDLMPSSLSTSPCSLGHSSQHRRLSSTKSSFSSSWLED TSSSV DTRSELSRSCVS
PPSRREQQLPVADVTNFRHIQQHQGR LAAGSRSPSTASVASTGPTGPLKTTWNRLET DLE
YYGMILLDKIIEQEPQACYLMTFVRHTDIAALLADLAQQRILAVRIMETVGLMVECADDW
KTAEFVLSLARRHNVSARDLRYHIQV LNSSLFAMLHIAMGPEWNAHMATAWKALLAQVC
STLFNQPDGHRSSSSSPSHALRTATACSSAYSTTNNLAWTSVKWKSTDSKDASRKTSSPL
PFKSLDDLTRAWHASSRSARLQSLKLRTRTSSSSPTKEQCNSSRKASASAGLSHTALPKSP
LKRSWSSR
>Sci_61721 scgid1626|
MGQNQSSRATKLANKAPKLN RNPLVANNAEKLKAEKTAQLQLILHNESKDEPCTSTPQTH
LI IHSAPCKTRLGR TKDAEPTTSAQRSSCSCECGGGKTS DQDRNDASAAPSTCPSRNS
SSSLSDLRPAPSVSSLQRQK GKFKDLSPPSVPNKKQRKAISEAESSQRPEPISESAISS
KQPSQQSKRKRKQRKTLKSLKLRGNSSSLVSRRRSSRRKREMTQLSKSWLAVSNNLEQL
GMDMFRRVADLCPDITRTISGLPEGAQISPHILLSHPRARQHARRVMETVGMVVDSYGEF
DRVRPLLVM LGRHHAKYGV EERQLMVRYELTWTYGV C
>Sci_44518 scgid1626| Neuroglobin
MGQNQSSRATKLANKAPKLN RNPLVANNAEKLKAEKTAQLQLILHNESKDEPCTSTPQTH
LI IHSAPCKTRLGR TKDAEPTTSAQRSSCSCECGGGKTS DQDRNDASAAPSTCPSRNS
SSSLSDLRPAPSVSSLQRQK GKFKDLSPPSVPNKKQRKAISEAESSQRPEPISESAISS
KQPSQQSKRKRKQRKTLKSLKLRGNSSSLVSRRRSSRRKREMTQLSKSWLAVSNNLEQL
GMDMFRRVADLCPDITRTISGLPEGAQISPHILLSHPRARQHARRVMETVGMVVDSYGEF
DRVRPLLVM LGRHHAKYGV EERQLMMIGEALIFTLQDYLEEEWTAMLADTWASLYAMIYG
ALLEGIKQRQAPPPGMERGHHLRTSFRQAPGHAI VSPATSSRSSPTGN NRKVGLSRKLSR
RNTSSSTQKNGIAVQSERKVSA
>Tad_2113619_1 hypothetical protein TRIADDRAFT_57230 [Trichoplax
adhaerens]
MDQAQTDSVQTPPQPSL TEEQKAI IRENWQDVEENMSEVGLYLFSKLFTIAPEYREVFPF
ETTTDNVRLRVHATGVMKTVGKAVQNL DQFSELQSALSTLGQFHHRKAIKFENFQAVGQA
LIQTLSDKLQENFTPEVHEAWSKTFDMITAAMKSGMN
>Tad_2115923_1 hypothetical protein TRIADDRAFT_59832 [Trichoplax
adhaerens]
MAPTAQDLQTIRETWALVAPDLKKGHTVLF LRLFEQHPDVQRLF EKIKDVP HDQLATNEN
FVFHTTRVMETIDHAVK GIDNLPAL TVLLKQLGSSHAQYNVKKEYFKIGLR ISEF
>Tad_2118316_1 hypothetical protein TRIADDRAFT_62364 [Trichoplax
adhaerens]
MTKIDSENHVKSNVVTGNDIKSYLNYQERQAI IDSWNAISTEKQKYGTILFLKLFEL EP
RVKSLFTIFDFNEPLEDIIQSPHFRSHAMRFMQSLETGVLMGFDKESCDLFLFKSLGSRHH
FYDLKSEFLDVIPECILHTIKKGCNNWSNETADAWKIATKVLCELFREGLETKPKK
>Tad_2110660_1 predicted protein [Trichoplax adhaerens]
MKDLIKDPLVRSHGLRFMKAIETMLEIEFDSNGCIFLFS AIGNRHCSYGIEADYLDYVPQ
AFRFMLTKALGN NYTDKIASVWDEILSHIIKAMQDKVREGTKLKEDKEEVARRISSAYLT
DKKREDCKSTTINGSEDS PNMV
>Tad_2110659_1 hypothetical protein TRIADDRAFT_54901 [Trichoplax
adhaerens]
MVLVNNYSLIKLSPATKIYFHGVDFEKRDSYLA KNTFLRNHAARFMEAINVIIGQDMDIF
SVESYFRVVGSKHHSYNL KLEHVQDISDAFLEMARNALKKKFTKSTEAAWRSFFQMVTDA
IKNGIMKAQNRN

***** HBL linkers *****

>Pdu_L1

MIQLLCVLGLVVTGALAGALRPNDGCNCPNRPPPPWGP PPPYQRGPSADLAEGRLLNQEA
RLEQLESYFQGLVDKYEKYASGKDERIARWNILQDRVWGLEAHHCDDHFS CRDNTYSCI
GHNLVCDGDQDCLNGRDEDEDTCRVVSDVGSADF GILLEEDPCTSRKPSGFRFIVTSVDI
NPQFTQEPKIKAAVI IRSRNDQGEVTEALTTEGIYDYTHRITLYSPDNDSLVECTFS
RYDDDHCDGEIRRQSGATCAKFGRLRLDD

>Pdu_L2

MSALWLLLGAALCAIASADHHADCSCPGGRQWGSVAAKTDAQEARINRLAGKVEAVA EK
LRKGKRVKNFLEEFQELE YRVNEIEGNGCETRHFQCGGDFPYCISDLLCCDGSND CPNG
ADENNATTCHIPI PAGTVLIGHLNTDHD FCTKRKPT EIDLII TSIVRPEYLKSRLLV KAN
LRLKFFAEGAEQEDVLPVKGYNF CNHQLVILPPE SDRLGLVCNFRAGNDHRCWPP LCTR
PPSPTAEMTSSLSRKTLGWVFEIYYLIS PFCGAV

>Pdu_L2b

MSALWLLLGAALSAITSAEECVCPGSRQASSVVS RANAQELRVNRLAGRIEAVADRLRT
GGERVKTF LGEFRELE YRVDQLEGNGCEPRHYQCGGDT PYCVSDLLTCDGSND CPNGSDE
NEDTCHIPI PAGTVLIGXLPNDHDFCTKRKPT EMDLIITSITR PKYLP SRLRVANLRLK
YHAEGADQEDVLPVSGYNF GNHQLIILPPE TDRGLGLVCNFRAGNDHRC LAAI VHEASLT
HCGDDFIFVKEEH

>CAJ00868.1 extracellular hemoglobin linker L2 precursor, partial
[Alvinella pompejana]

MRILIVLCLAALVRAGARFDEIGEEIAELTAEV VAFENMIEDRIEQAEARHAQWHV FSSRIDAVEGSGCD
DAIQMQCGGDT PDCVSRLLICDGENDCLNGADES QCRVFTPEGSKWKA EFTWDTCTKRKPKEL SVTFTHY
EVPEELPSFP EVKAEAYMSENDDFSY SANLILSGYMDL KPGQNKLYLRPPESDGLGLQCIFDGVNDDHC
PG

>CAJ00867.1 extracellular hemoglobin linker L1 precursor [Alvinella
pompejana]

MAGLVALAIILLA AVQVESHERRDVNDMRLNDLQ GKIEELQQT LDERAKTRDQRLREFAELTARVHKLKE
SHCGPREFECTESANHC IHDILVCDGANDCPDGSDEKNCGNPAHAGATFKGVILNSQCQTEN VAKNMQID
IVGEKRYSDFTP TISVLELLVTLDDHQDLYNGIYSYGRKALV SFGKGGGGLGMVCYFD TDDGKFKCAEFLS
IVSKEVCGTAILTSD

>CAJ00869.1 extracellular hemoglobin linker L3 precursor, partial
[Alvinella pompejana]

DVDKEMEVL EEIIRQAEP SGCNDRFEFQCGGHHPKCI SRLLVCD SYADCENAA DENEQCRVYTPAGSTWE
GRVESDTCTKR RPKEVRLTIDSYKIFDYLR SFPRVDGHTGDRSALAGLRPDEQRQVGGI HGSDAWRQHPV
LRCSGAGRCP PPVRLRRH

>CAJ00866.1 extracellular hemoglobin linker L2 precursor [Arenicola
marina]

MKSYVLVCCLVGAVAYPHEVMH HAVGANRMCKCDAPAGNAETSADREQSHTLDELTHQLHMLQQAYDTGMGR
VDDVMED
MDDL SHRIADHEKEHCKKYREFQCGGDHPKCI SNLLVCDGDND CDNGADEARCDVLT EAGSSWTGTVVYDHCT
KRRPETM
KLSIKSVDTPFFTT HPKVRGTVLMEKHTKDYSEVINEPVSGYWSSADRSAAMP PDSAGHLGFVCI FHGHDDH
TCTGLLT
KGKVTDACAEFTFHRD

>4D0I_1 Chain 1, Extracellular Hemoglobin Linker L2 Subunit [Lumbricus terrestris]
MLRLLLLSALSGLILADHHQPSGGGGGSYGGGGGGGPFGRFLSDQLDPRLGANAFLLIIRLDRIIEKLRCLKLDEAEKIDP
EHFVSEIDARVTKIEGTHCEKRTFQCGGNEQECISDLLVCDGHKDCRNAHDEDPDVCDSVVKAGNVFSGTSTWHGCLAR
EDHVTRITITASKRRKFFFTARIWLRALVESELERHGENVTSSFNAKGYYNFASRRLLILLPTDDHDDHLAVVCSFNRGDNE
RAECHRVTATLHQCADLFFVTLEEHDHDDHDDHDDHDDHGGKHH

>4D0I_0 Chain 0, Hemoglobin Linker Chain L1 [Lumbricus terrestris]
MWYVLGLMLVGLAAGASDPYQERRFQYLVKNQNLHIDYLAKKLHDIEEEYNKLTHDVKKTIRQLKARISNLE
EHCDEH
ESECRGDVPECIHDLLFCDGEEKDCRDGSEDEPETCSLNITHVGSSYTGLATWTSCEDLNPDHAIVTITAAHRK
SFFPNRV
WLRATLSYELDEHDHTVSTTQLRGFYNGKRELLLLAPLKGQSEGYGVICDFNLGDDHADCKIVVPSLFFVCA
HFNAQRY

>2GTL_N Chain N, Extracellular Hemoglobin Linker L2 Subunit [Lumbricus terrestris]
LDPRLGANAFLLIIRLDRIIEKLRCLKLDEAEKIDPEHFVSEIDARVTKIEGTHCEKRTFQCGGNEQECISDLLV
CDGHKDC
RNAHDEDPDVCDSVVKAGNVFSGTSTWHGCLAREDHVTRITITASKRRKFFFTARIWLRALVESELERHGENV
TSSFNAK
GYYNFASRRLLILLPTDDHDDHLAVVCSFNRGDNERAECHRVTATLHQCADLFFVTLEEHD

>5M3L_N Chain N, Extracellular hemoglobin linker L2 subunit [Lumbricus terrestris]
LDPRLGANAFLLIIRLDRIIEKLRCLKLDEAEKIDPEHFVSEIDARVEKIEGTHCEKRTFQCGGNEQECISDLLV
CDGHKDC
RNAHDEDPDVCDSVVKAGNVFSGTSTWHGCLAREDHVTRITITASKRRKFFFTARIWLRALVESELERHGENV
TSSFNAK
GYYNFASRRLLILLPTDDHDDHLAVVCSFNRGDNERAECHRVTATLHQCADLFFVTLEEHD

>ABB71123.1 extracellular hemoglobin linker L3 subunit precursor [Lumbricus terrestris]
MKSLGLLLAALAVVVTLASADSPPAQSHDEIIDKLIERTNKITTSISHVESLLDDRLLDPKRIRKAGSLRHRVE
ELEDPSC
DEHEHQCGGDDPQCISKLFVCDGHNDCRNGEDEKDCDLPKAGDKFIGDVVFDHCTKRRPEHMTLAFESSSIA
AFFTPIA
DLHVHIEIESETDEDESEVSM PADGEYSFADHRLTIHPPEEDGLGLVGEFDGYNFDRFVGHIVHELSEEVCAE
FIFHRKK

>2GTL_0 Chain 0, Extracellular Hemoglobin Linker L3 Subunit [Lumbricus terrestris]
QSHDEIIDKLIERTNKITTSISHVESLLDDRLLDPKRIRKAGSLRHRVEELEDPSCDEHEHQCGGDDPQCISKL
FVCDGHN
DCRNGEDEKDCDLPKAGDKFIGDVCFDHCTKRRPEHMTLAFESSSIAAFFTPIADLHVHIEIESETDEDESE
VSM PADG
EYSFADHRLTIHPPEEDGLGLVGEFDGYNFDRFVGHIVHELSEEVCAEFIFHRKK

>4D0I_2 Chain 2, Extracellular Hemoglobin Linker L3 Subunit [Lumbricus terrestris]
MKSLGLLLAALAVVVTLASADSPPAQSHDEIIDKLIERTNKITTSISHVESLLDDRLLDPKRIRKAGSLRHRVE
ELEDPSC
DEHEHQCGGDDPQCISKLFVCDGHNDCRNGEDEKDCDLPKAGDKFIGDVCFDHCTKRRPEHMTLAFESSSIA
AFFTPIA

DLHVHIEIESETDEDESEVSM PADGEYSFADHRLTIHPPEEDGLGLVGEFDGYNFDRFVGHIVHELSEEVCAE
FIFHRKK

>5M3L_O Chain O, Extracellular hemoglobin linker L3 subunit [Lumbricus
terrestris]
QSHDEIIDKIIERTNKITTSISHVESLLDDRLLDPKRIRKAGSLRHRVEELEDPSCDEHEHQCGGDDPQCISKL
FVCDGHN
DCRNGEDEKDCITLPTKAGDKFIGDVCFDHCTKRRPEHMTLAFESSSIAAFFTPIADLHVHIEIESETDEDESE
VSM PADG
EYSFADHRLTIHPPEEDGLGLVGEFDGYNFDRFVGHIVHELSEEVCAEFIFHRKK

>2GTL_M Chain M, Hemoglobin Linker Chain L1 [Lumbricus terrestris]
RFQYL^VVKNQNLHIDYLAKKLHDIEEEYNKLTHDVDDKKTIRQLKARISNLEEHHHCDEHESECRGDVPECIHDLL
FCDGEKD
CRDGSDEDPETCSLNITHVGSSTGLATWTSCEDLNPDHAIVTITAAHRKSFFPNRVWLRATLSYELDEHDHT
VSTTQLR
GFYNFGKRELLLAPLKGQSEGYGVICDFNLGDDDHADCKIVVPSSLFVCAHFNAQRY

>ABB71124.1 extracellular hemoglobin linker L4 subunit precursor
[Lumbricus terrestris]
MRGPFIVVVVLA AVACLQVDAAAEEDNRARDISERIDKLTAEAFKLGRNLDARLDPIRIKKAGTLKARVD
AIAEPTC
DEHEYQCGGDDPQCVGDLVCDGITDCRNGDDEKHCVL PFAKGDTFVG DQEFDHCGRFNPDHITLHIDSVTTI
PFFTSHP
KVTGRVDIHVDRDDWAVSTPSFGFYSFATHRIIFRTPDKDSL^VYLAQFDGYNFDRFVGETLRVGTGLPCARF
IYKRQH