

Text_s1-1.txt

>NvecA

MSSDEEELSDAKGEAESKLLDGLTPRPYQSELVRIALEKNIIVCLGTGTGKTFISVMLIKELSFQVRKSRKEGGKRTFFL
VNTVPLAKQQKEVIAKNTDLKVKEFIGDMGVDFWSDEQWKKELDDNQVLVMTAQIFLDLLGHARLRLSDVNLIVFDECHH
ARKGHPYKQIMGYFEESDYPRILGLTASIINTKETHKMEQLICELEATMRAACQTSDDITVSQFATRPEEIIYPYSNDKT
EVEQDLNQLLVETVLPWQEFVDDL RASGQKEVLKGVKSLISDCRQILD EMGASAANRVAEYFLKDLASTPDISLCGIVF
VERRYTAKILCEQLNIIATRDPSELEFIKCFVVGHGASGINSSSDTQMGSTKQEKVLMFRKHQINLLVATSVEEGLDV
PKCNFVVRFDPPQNFRRSYVQSKGRARAKNSAFIVLVEDKDESEWAGLEVEGIVMPRKMLSRQAAAYEACKRLHQMGIEIDN
LRPVVSLSESEDEGTESEDDDKSKGKTGTTKRRRCYNRKRVRNMSERINFLVNRKKQSRKKSNNHVVHFIPELCSHDAIPAS
LLSNAQILPSVLHRMSTLDDKEPDPLVGDMTMQGILNGSINVLIVCRRAGSLRRRRAYKCKIPCVFTNSLPDPGQSCYVYAV
EFQEISQSLQAMFSNERTVLLHPDSELLLQALTTKSAGDAFDLERLEMLGDAFLKQAVSIFLYCCYKDGDEGKLTKRKTS
QLSNRALYRVACRLGLPEYLQNTQLDRVSWTTPSLQAAESASRQPENRPEMSTLEKPEPETARDKALHQSIPDKSVADSVEA
LIGAYLTTTCGYQGALHFLDWLGVKVLKRFVAVQLQGF EKRIGYEFSNKLYLIQALTHASYSQNRLTDCYQRLEFLGDALLD
FLVTELIYSRNTNLDPGELTDLRQALVNNNIFAEIAVKHGFNKSLFQMSPEWFNKIGVFDVHVKEREARGIQISKDPFLT
LSNDDEEGIEAPKVLGDI FESVAGAI FLDSGMDLVL TWRVYRMLKPYIDAYSADV PKNPVRLIHEKDADALFSEAKVLP
SGKVECSLTVYWGNGFQGRAANKRNAKMAASKLALSALQPSR

>NvecB

MEVTFGVQAI VSSDLLAVNKYGERPEETLVVASSMTNDIINKNVLECLHEARDTLSFANNGESVDVFNRLSSMLAECEDI
LIDLGTFFSSLKALSFTSREFNKEGKKAIAKAKQEDLATTLQICAAKLDEVVNTVKKEFDGSADILLTDVTRRIFQFLTLLD
KENSTSPTEQVELRNKNFERVHFQKQNDISKIEIANHTTRPMQRPVAVILAKSELQAHAINSLISCISDFTHPYLKTRI
LSNDKINQMINTHMSLDDKEPDPLVGDMTMQGILNGSINVLIVCRRAGSLRRRRAYKCKIPCVFTNSLPDPGQSCYVYAV
HLVLT KPDSHERAVRSQGRTRLDPEKRPLSLGII LNKSLPKLPAFIIYDRAGEMTVSLKHCVNTMLTSQQLALIQRFHRY
VFTDVLNLEKGESSLNFEPKENKNGCFLVLLRNTDQTTNPDRGQTKNPNYEEGQTTNPDRGGQTTNPDEAQENIRKGF DG
FPKL SQDVHASGVERCDKIPLDGANGSATVTSNELRRTPCNANGNKYDIRPEMVNEENVPILAKQNTHTNPSWKIAFGF
LYSLED SVTSFDDPCDPPPIITSDFEVFRDSVTTAYAPSRRLRYVADIRYDESPTSPFPSPSVAPTFGDYKDRYGV DL
SRDQPLLDVHVSSRLNFLRPYKTSKGGKDLALPDEKSKRSQRSRVSLVPEMCSIHP IPGSLWRDITNMP SILYRVDALL
IAEELRQKLAVEAGIGMPVWPEGHDL PQLSIEEMVAGPIHPAAASSDVTVCSSFSDPNTAAKTC SVNTSVPLETRLGS
MILSDSTDACTWETRRSHSNKGS ETPVGLNGKRVNSAGTCLTGDVDANVSKDRETQFDPGSPSTLILRALTTT SANDII
DHERLEIYGDAFIKYAISELLYSDFPNVNEGRLSFLRGLRVSNOQLFYLGRSRLAPLILTKPFEPVGSWVPPGFSSSSM
RNVHETLDVSPQHGVNDVNLIDDEDSLDHSEIQDISGNECRGNVEPMLLPHTFTVISDKTVADTIEALIGACLMSSTGN
AALRLMRWLDMDVSSCPDSKLSNPSTESSSEDKARSYQEPICINTHIPSATEL SVTDIKPPLTPSFSGKMDVQLFKRAVQ
EPLSTPFISKQTDLLRVRTANDQPPATVCSTAGQGLFENFESSTSVSKQMKELSVHAVDEKLDNAYSTSEHDL SI
NKQHQTAVLSEHLGELALVEASLGYTFRDRDILLQALHTSYPRAYSVNSSYEQLEFLGDALLDYLVRHVLQRFP
LSPGAI TDLRS AVNNYS LACLTIQRDFHKHMRSMSMALFKVRRFVRRLSEIEEKETRPDQVYSKLLLTTVEDGRGGEA
IEVPKVLGDLFEALLGALYVDCGRDEELAWSVIYPLLP TPIIEHYSVHL PISPTRQLVEMMPGAATFTCLSQNVIGHHACE
LVVK

>HmagA

MINKIGPKTDPCGTPLVTELHAI VFTQRRCTAVVL SKLINYQVEKDETYKGLNSDFVLGHAFDSNICLVDGSMRSEKQNL
ILSNFKQDKFNVLVSTSVVEEGLDVRKCNLVVRFDGINNYREYAQSKGRARASGSKFIIMTEEGQFDTTKQHIVMQQIE
EILFEQCQNRVLP TDEEIEIEMMDEEDEDLPVYF SHNGYDAPRITKDNVSLLYRYCSRLPSDQFCKTNPKFLYHVLKNQ
FCAECMLPLNCSLPQRIYKLP SFDLFSRVGKLNIELIKIPSRIRFECEFNIDDVTFHNMVITESCSNLVDGDIYFHIND
ENATKALFVPLHSDVFKAVRKNCEYEFIPVIEKRLNCTKEKPKDNLGNEE IYENCVVSKNYGESKNRMYVTAISKLNT
SSFIEGKTFVQYYLEKYNIEISP NKPLIDVISCSQQISFTNCESENTNKKKNVNMEKLVPELCKKVSFLPADLHFQLPC
LPRIIYFYEYFLALELKD MIAEKIKINKSENQGRYQSSVLDLENLLQLKNYVTENVKSPVLI FEALTRQASMPFDL
ERLEVLGDSF LKQAVSIFVFFQNP SFHEGKLTIRRKNCISNENLCKVGKSAFTHLSYLSNSVTGICYQYEF LGDAILD
LITLHIQSNTQNLSPGQLTYFRSAVNNNTFGLLAIKHSFYKYIYSSSNELFSSYNSFIEIVKDDHGLKDEYFESPFVIC
PTGSEEGYQAPKVLGDLFESVAGAI FLDSGLDLVTWVKIYPLLPKVIDKYMIKLPRHPAVDFFEKYPKAKIIFTNGEVV
TCTVKFGKDNTEIKACGQNKKI AKAACKRALEV

>HmagB

MEVFTISDIKSI PVSEISYSQNKLSGNDKQYQESKSDSNSSYNGILGSVEQESYVDDDKFTNNTEDLQIASEDLNTK
DKHHTYVYVVSFSDSKTESDLIGLCVNAPLPFIFPFVLFHDYQPNVVKIKFGFEYNMKSSEMELIKKFMHFIHFAVLKM
TVMSYDGYLCVLLHLENGVHKIDFNKMSSMIAAQMNLCETKINLLKNRQGFKTYDRNIYEVTNKTNEIKIVENICNNTLK
VSLNGPSCLTDCVKSCHHML ENKDIRIKKFYEKHYNTLFIQIHVSKECFHLDRAKSEKTI LASDPFPDNSKAVSYKEYF
MKRYNLFINDDDLMAPLNYFKPKFNALNKKVSNKKRTI IYLP SFLCPVDFIGLGLLCCFCLMPSVVFYIKLNLNVFHIH
SSINNF SKEKCFSLDYF THEECDISLTLGTAPNNVSSNKS VYSESQSEDIKTHHFFLQSLVAANAALDINNDLLEFLGD
SVLKYIITAEHLAFPSVKSAQLVFYRNKILSNKYLSELGFLRHMDKIIIFEI FEPAKAFSSNNSSKIRTILEKLSL NDA
NERLLANVVEAFIGACFLTLPQDSL YNVFRYFGFNIMKLSLLHKNGYRANYPLECASFKRVEAFKSLELKLGYKFQNP
LYFNACLHHSYADVLYNGKKNFIDIAAIKLTNECLELIGDAVLDILVTKEIFSRNNISGANKSGNLSFIRSAIVNTY
TFSVLATDLELHKHLIFSSYKLSQKIIDWLVNFEEMKKTGNLYNKVFFF

>HmagC

MKLINTDFRNKKNQMVLEIVSKGRKFSTSVNEYTVVHCQLLPSINHTWESLRDPTNISSRD TLLFAGKLAKRIFGFQGDK
KTSDSGDSEGHQTS PASIFLQEELMLMMDLLL RKEKATSASTLPSFDLFSVVGKLIKIELIKIPSRIRSECDINIDDIKNF
HINVVTESCLKLV DGLMSFNINDAKVTKALFVPLHSDVFKAIRDDCYEFMFPVINKRLLNFTKEKPKDNL DDEENYVNCV
VSKNYGESKNRMYVTAISSLNTSSFMEGNKT FVQYYLEKYNIEISP NKPLIDAISYSQQISFTNFESNTNKKKKNINTE
KLVPKLCKKVSFFRADLHLQLPCLPRIIYFYEYFLALELKD MIAEKIKINKSENQGRYQSSVLDLENLLQLKNYVTENV
KSPSVLKI FEALTRQANMPFDLERLEILGDSF LKQAVSIFVFFQNP SFHEGKLTIRRKNCISNENLCKVGKSVGIQNI
CNSQFGIKSSNSEN RNPWTWVLP PGYLRAELDDIKSGGLGLLGAQGYFCWRPFYGPPIKIKIKESLIQCGSIFPKKKQI
NKPSI

>Tadha

Text_s1-1.txt

MDATSQDNVDDHKEELLKPRAYQSELLEAALRSNAIICLNTGSGKTFIAVMLIKKLQHEILPKPLDQGGKKTFFLVNTVP
LARQQSDVLRKETPFKVGVDYVGGDDVDLWNTKWKVYELDSQQVLVMTADILAQILQHGYIKVSQINLLIFDECHHAVGNHT
YVQIMRDYVYKEQKSKPRIFGMTASVFGKCRIGQIENKIRSLEAILLSRVTTTSYRQDLANHSTKAEEMLIPIYAINLLD
NHRHKLNSRNFCIAIFTEKRCVAYALSSILDKVAEHDNAFSCIKSDFFVVGHGTRVSSSLVATNMSHNQQTRAINKFRQQD
VNILVATNVLEEGIDVPKCNLVVRFDEIKTFGSYLQAKGRARAQHAHYIILVEKDKIADARDQFSTWSYLERYLQDQCQC
RAVGTEEDYLADDSLNLVPPYEVESTGAKVTMSNTTIQLPSSSPLPDSITGPLAITEKRAKASTALLVCQELHEIGQLDD
NLVPVKYLQSESFSKNIPAVLNDCHPKEKQDLYLALVMKALPPYKAVNIFDKVNFGIILSHKLPEVSMIYRNTQSETSS
KITELEPELCKRYFLSASVWKVITCIPSLWRMETLLLAHELKDVLFSGIPINGISISKNTFTTLKLEALTAEAEEDDI
SLERLEFLGDSFLKFIIAANLFI SHPKYHEGQLTNELTKIVRNDNLLKLGKSTPICEYIIGKPFNLLKNEGWIIPPGYKLL
SSKVKNLDEKVTSDLDGDKAVTHQKFSSKSIADVMKAILAAILLEFDERHAKYFISWLDKDIKFTINASAASLLPDDNI
SSNEINKIYETNNLERLEKVIQYFQNKLFMIRALTHSSYHHHDIAIESNQKLEFLGDTILQYIIARHFFNSHPDAKPAK
LHDFQEAIVNNKCLAVIGIAIGVSKYVFQESPDLFRDIDQFAKIITCDYIMIDDEMNYNDRTEALFENFTEEVKIPKPVSD
VVEALIGAIYIDTGESIAEVTKALMPLLQPRISLVL RDIPI SPIRQLYEKVPPEGVSFV

>Tadh

MAIDDRLLNGVQYYHRADKQTSCHYLALSSAKVCTSVTQLIFQELNTRMKLSPRKYQLDLAVIAQERNITVYLETGSG
KTLIAVLVIRQLYRELSKSLQAGGKRAIFLVNTVPLVHQQAKFIADQTSFCVGLKFGKNSDKRRSNWKDFDFKNNVIVG
TPQTLRLLELNVNMQMINLIVFDECHHATKHPYSRIMDAVKALEEQGRPKILGLTACIFKGECKPSKVPQRLAELEE
KLSSTIAIPGYLTYIDKNKTPPTEYVIDYETTETSIFEQVEALSDCSFLKGFDMYSNSFESKAQREQQSTSLQLVLT
FDEAILTVNTFGLRIFPLVAESIVKCIKELRMFRGKGDIDYKLSASNSWFRPVIRILEAKLSTKIDESHPCYTCSHFA
LLDVLTTFKEGDTGIIFVDRRVTAFLVHELLVRISHTDPNLAHLNVGCVIGRSSIKSQLGQAQMSRYEQEVMNRFRRH
KYNLLVATNVLEEGVDVSECNVIRFDRVNHFSYLSQRGRARADDSSVLMINRENLD AFMEEYNSWHSVEKFLRSSPE
LLNTIVSGISKSLFSDNNTDASTISSKFVETTGAENLFSASVLLRYSQSMINGKSPFEFETNDDRTRAKIKLIF
SNPPEKTTIEGPMMSKIKAKKASALEACKFLHEVGEITDNL EPNVGKYHQSIDKSQYKQFPKIPDNLSMPIEDGDNFY
LHMKIELKNSTIESIDTVSCSSDFGLICKTKLHEIEPTYLQTS CGDILFSLQSLAVPMHYSCEDELSISKFVEYLFTNI
TPCSTITDLIYDPTNSYAKYMIVPLNYAKSSRSMYSINFELMATLLKPRKYWPFASNLGDGCKNVNDLLIVKNYGHVVTN
QVDVYALRNICYDLSPLSRFPTDKRLTYKAYYHRKYHTRIEENSQPLFAVKTLHRFPVALVHENCISPCNNTTESKKSNS
GFVHLVPELCLVHPFSASMMWLSQFLPLTLTKIEKNCLAYEMIDYIQCASILKVKTEPRRKKLLRETTNLVDRITQSLT
LPIAEEQCNFNKIKYIGKVFIKLVAVNL YQTCDDNSYAL TMKRRQSSSEFYLFNLAAKTNI PSYIISKPFNLK CWIP
PGFTFSIRCKDDSSDSALAERTIIESITDNFLSVKSIAECISAIVGALLFPDNETEGLNFKLWLGFAIEPFSSISNGFD
VWNEKITHMIEESNSQSWLEKHLKFEENLGYKFINIDYLLAFTHSSEFGDSFNYEKLEFVGD SILNYLVIKYLF L ATD
QDITRSKFFYGAISNEVYGLLSYAYNFHLLFYQSPFVKEMILGFEKKTADLLEGGSI PRNGNKHDWVYTI FSKSIDDDF
SIPKELADV FESLAAAIYFDSGQNLSTVWNIYSNLLKPFLLDKLLKDNQKNFQELNGED

>Tadh

MAKSVFAFHNSDLQFEASNSYSGYLVPPIRSETSPIEDRWEGDIDYDFLSRIASYNKTFIPIDLGNLDILNDAIITRNYG
NNPSYVEVAQICRHRTVNDLWGESNLNFITYYQEKY GIMIKNTTQPLLRPKYLSKKNRMLRSNSGQKSQFSVKSCNKSQ
TQSVADTITHSHKSGKSVIELIPELCKRSIFPASVWKS LACIRSILWRMESMLSVKEFENKMQESWIVGSSRNERQLQNN
SLTMTILKALTLARANDDFNLETLEFLGDSYLKYAVSISFFLSYK FHEGQLSSLKSKIVSNENLYQIGIKMGLPGYIIG
HNVDQNDIWWPPGFSIAKKSISGLEQLSTLNL SMTSRVHPTQKLKNKCIADVMEALLGGILLEGENSANNFITNNLG
IKVGRLSHNI AKLDDARKEISKIYNAHNLKDLEAVL GYSFKDRKLLISAL THATYNNHRIAEGVECYQRLEFLGDAILQ
YIVVRHCYHANNNYKPADLHDFVEASLQNESLAVMTMVYGIHKYFFYNPSPLFSQINQLLQIVDTNEECEQRL STEEKAL
VYSNLKDGVMIPKPLGDLFESLIGAI FIDCGRSIEVVDLVLRLQLPQHAIFAQGVKPSPIRILHEKFP SGLLFKKEKT
EDNSLRKLSLQVPDGSYFTATGINYKHCKKVL CMMALRCYDENSN

>Tadh

MLKISKLNKQGKRRQQLHKREVFNAKVADVITSNEYKPGEPYVVEITLLEHPFQLVNGAIVQKLSTYGILTKQQLPK
LPRIIIFTPMGKAHVNFIAREKPLFLTEEELDNCEK FHRFLFENVIKTRKDRTRFIKNESIKYYLVVPLRENGELHYELIT
SCINFLRNENYPNRQVCKEVLQFPHQVVVKPSYSLNSQSLNWMLIRKVS DATPMSEFHDT SIAATFYDYKNQYSIDIS
DKEQLVCLRKIPSNMMLKFKHRILTLSKSKQSPHFVPELLAIHPLSASRFMTTVCLPSILFRVESLLSVDELRLKIA
RETGLGLDDSRMDNSFIPLLDASTYTKKIGILPHVFHQIKLKDRYTLITLNGKDNYLHISNEEKRRLVDRKLKRGRE
KKQLVINHCKLRFRTISKLTTRKWLKIVPISVSKRLFPKHRRLLHSKLSKSNRNKAVLKRPGSPVSMALLQALTSLSA
DDFFNHERLEILGDSYLKFITTADYFFRNPNMKEGAMSLMRAINVSNSHL SRTSKLKLTLTYQFNTPFKQWLPPGFTC
IRNNANLVANNMLTDGSKQRDESFKQVMSVKNISDTVEGLGVVYSLASGPVDAARLLPWIGLDIDPNVLEKDKSVNFE
IPTTDEDIQLSSCILRNTDYQFN FYWLRIAFTHSSVASENYQRLEYLGDAILEYLLIEFVYCRKPYFGPGKISHFKSVM
TRNQTFSLIAIAFQFYTHIRCHSKLKDIMDSIQNFKEILLDNHQDDHQPDI GHFANLAYQLYTSNRYQKLPKVCADVFE
AFIAAIYLD CGKLRVWVVKVYKQRI LADILNDVGLTSTIHDMFPNRIYQVIKRSNSSKKHRRMFLDHPKLGMLTS
EGIGTSNTSAKDDAIRNMLSHLECSDEIESSNKS

>Tadh

MTTRGKIVVKVADYFEKKLDQDAISQIHNFSKFVFEEMITCANSSIFQFDALNGLK KYLLVPILRSDSRDYLD FEFVNL
FCKEEHKL YVTPASHDYIIQYENAMVARNYRGLDTLEWFFVSNVCS DMSPSSKFTNPVASTYSEYYKIKYDIDIQDANQ
PLLEVARFKMHITLLNAKQSGSITHDIRTL SVNSNSNSIVL IPELCLIHVIPAELWFLSQYALVALVYVESILLAFELR
ELLRSQANICDIYSVDYLITAVGSAKPCLLEAITLSAQDHFNLERLEFLGDRYLNKYITLTTLEHPSWSINEVINRMR
FLTKNKLLFRLGAGRRIPYYIVGCPFNQKSWVPPGFVINTVSEL SIAMSR EYTHQPLTKKAIADCEVLEGLIGYLCVCG
RNSVLRFMRWLGF EISECLDGIQELIPSIDFRTNLYSETANSVETILNYSFRNRSLLMLAL THKSYSGLESISYEKLAFL
GDTLLIYQICLFLYTQYTDYQPKDYHLLLEYLTCNRMYGIVAIASDLCKFLLHKSRELTLQIDEIYQALEADDVQSFSS
VQSGSKYMKTTLLLEKLSIEQFVPELGD TFAAVVAAVFLDNYPDEDTAGRIINSLLNPIFDVKYTCEDNLPVSLTQQM
YESLPNRVRTFFVDNKSGSYSICNVSVLKNST EKT FAGKGDHEHAKTLAMAALTRYIDENPNQFQ

>Cte1

MKNLRQDQIHSNMFTPRTYQVALLENALKHNTIVCLASRKT FIAVMLIKELAHQVRQPLQLGGKRTLFLVYSDDDVEAQV
KVLGLHTDLSVGECCADSTVAQWQEEFVKHQVLVISAQVFVEILLQGAQISQTNLLIFDDCQHAIEDPAYEKIMQFYTD

Text_s1-1.txt

QTNGQRPRILGLTTSILSEQISCPIALEKTLATLEKVLHSTVETSSDIINIDLYGNKPQESVIECVAYDDPTGLVAEFST
ILDSCMAFLDDCNIAFEEDKDPRAIPKTVVTECQIILRELGPWCAGVAHIFLQKLSKLEHETEKMRMFMQYIFTQLTI
IHSTVENVYDNQILCLEEFLQLMSPRIIQLIKILHNYKPDNFMVIMGNDNFMDAIYDSEGSTDISDEENSEPKPQQVHST
TSNLKRAVTHHRRQDNPEDES LCGVVFERRHTAFVLNKLLEVELCNWDTDLFFVQSHHICGGAAGPTNNSQQTPLYKK
QEEILKRFRQKDLNLLVSTSVLEEGIDFPKCNLVKFDAPTSYRSYVQSKARARALKSKYYILTEPTNQGPFHHTLSTYK
GIENILISKCLQSRENPLRDDDDASL PDNFMPVYVSPGGAKVTMSSAIALVNRVYCGKLP SDAFTHLTPKQVQAINADQ
ETWRATLQLPINSPIKRPIYGEAMP SRKLAMMAVALKTCEILHQQGLDDHLLPIGKEMIKYEDEENEWEDHETQGHARP
GTTKRKQYLLKKTASTLLGQPQNPCHLYIFDTRLTCPITEEQNTRGRKIHPEDSARGFGIITVRKIPLVPHPVFVTR
SGEVTVSIDPVDSEFTTAEQLPQICHFHRFI FTRVLRLEKDPVFDPKKAMSSYFIVPLLKDADGIRVDWDFVGEIQKS
PLHLTRRKYL PDAKKEAFVFNEDDFHDAVVMPSYRNIDQPQYFYVAEIRKDLNPKCPFPSPELYDTFELYKSKYGLGIT
NLDQPLLDVDHTSARLNL TPRYMNQKGVALPTSSAETRRARRENLQKQILPELCDIHFPASLWRKAVCIPAILYRM
NYLLVAEEMRLMISTAAHIGVPQVPVDFR FSELEFGFSTKPGDGMTKEEALESKLSRMKIRDEANGNLTDTHRIKFDEE
FDLVNFIGPSPCLILQTLTMSNANDFFNLERLETIGDSFLKFAITVYLYCEYPGIHEGKLSYLRSKQVSNYNLYKLGKKK
GLPDCMVAVKFEPTENLTPGYVIRKEGAYKGLNVRIASIPAPLLAPAFNESGSKKEEQLPDGQMEKFTKELECTKAED
EKLKGS DIPHMLIPYNLQTQHSLPDKSIADGVEALIGCYLVTGCGRAALQFMSWLGLRVL PNVKPAKDNKNNLAMSQSCLP
PPPSPLLTHVSNSEDVLKHL LDGYEEFEKICYKFRDRSYLLQAFTHASYHYNTVTDYQRL EFLGDAILDYVITRHLFE
DSEKHS PGVLTDLRSALVNNNIFAALAVKWNFHKYFRAISPQLFYVIDKFVTWQKDRNDEINLDEEFGLEEDGEGTEED
IEIPKALGDI FESVAGAIYLD SGMSLDVTWQVYYRMMK PQIDAFTERIPKSPVRELLEMEPETAKFEKPERTMDGKIRVT
VNVVGKGNFRGIGRNYRIAKSAAAKRALKHISIQFHHSASIMPPPHLM

>Hrob

MTSSEFQKHEEVLKRFRQQLNLLVTTNALQTGIDL PKCNLVVRFSPASSYKSYVQSKARARSNIAKFVHLVDSHLKDQF
GSDIKMYKGVKILYHRCRHYRHNINRHIPNRDVISAKFPPIYAPNCLTLPHPKVTVLTAIKLVNKYCARLPSDNFTHL
TPKCEITEHYDKNVKMYKARLKL PINSSVKQTVEGPPMPDEQQAQMAVALEMCKVLHQSEELDDNLQPIGKEILPLKRS
RDSDEDDDDGGYEPGSTKRKQHYRKRVARALLGSVRPNAPCHLYAFNMTLTGPISEEQNTGRRIIHPEDTSKGFGLVLP
RSLPKIPHPVFVTRSGEVTVSVDPMTPSNLHSDLHFDSDQLKRLYAFHKFLFSDVLKLEKCPMIFDPERAKSCCIVVPI
RGHSDHHLAKQSDHLVKQDAGRLNAQQQLQEQQNFKFRVEDFIDAVVTPSYRNVDQPQYFYVAEIRRDL SLTSPFP
SPELYETFRHYSTKYNLHL TDHTQPLLDVDHTSARLNL TPRYVQKGVPLPTSAENRRAKRESLQKQILVPELCYI
HPFPASLWRKAVCVPTILYRVNYLLLAQELLERIAKEAGVYACVRVCVVCVCMCVTITNQPIPIKPNNHIDHLQIN
QSHQPTANQPIIHNPTNHKQGLSFKKNGNSQALQFRFAKDGEASSTFCGGCGFGDDDDNDDDDAAAATTTTATTNTAFTS
MYGPGACDVIQTLTMSNANDFFNLERLETIGDSFLKFAITAYLYCTYSIGIHEGKLSYLRSMQVSNYNLYRLGKRKGLPGL
MLAAKFEPELENWLPCCYLVKRAPKSYIWNWKEQDGISQSDKNRD DDDDDGVDVCGVGRGDDNDEDEEEDDEDEA
DFFFINNVDKDDDDDEDEADYDDDDDDGGDFGGKTT SAAAAVACKNAANNNNTTTTSSNNNNLLIPYNLQTQHSL
PDKSIADCEALIGCYLACSGQNSALKLMLWMGLKVL PVSSSNDRSGDHSNDHSDHSGSSKDNFSNNANFLTSPSPPLL
RHFESIGYKFKDRSYLLQAFTHASYHYNEVTDCYQRL EFLGDAILDYVITKHLFKDSIEHSPGILTDLRSALVNNNIFA
LLAVKWNFHKFLKYIGPQLFYVISRFVSWFNEQGEDAILDELEGESHNPLDADNDDEAQPPTNRRNNNNQPACLEDIEI
PKVLGDIFESVAGAIYLDANLSLDVWVWVYKRLMDKYITAFLEKIPKSPVRELLEAEPETAKFEKPERTIDGRVRVCVNV
VGKGSFRGIGRNYRIAKTAAAKRALKYIKMISGSSC

>Lgig

MKVDLVDSALQKNTIVCLSSSSEKLFIALMVIKELSTVLRKPLSEGGKRTIYVNSDDTAEFQCQILSRHTNLKIGYYTI
DKDTENWAKDKWQEEFTSKHVLVFTA EVIDLVTQNSLPLPQINMLIFDDCHLAVGDHSYVQIMHMIENEKVEPHVIGLT
ASILGSRIPPEPHLKHITITSL ENIFHCTAETSMLVISERFDSRPTESVLQCFTDQSDEVVLKLNDIIDSALFFLYECNIS
PDTEEEPFFKDPRLGIVNLSEIRNILHQLGTWCTSRIAEILIPQIEKNEKSESACINKRFIRFTLTQLRLIVKVFANF
NPDYDEVELLMYSTIRL VHMLRKYKPDWFMIISSDDLSDDEL SNLSDDNEDADSLNLSDDNEDDHSKPLIHVA
VKRSTTDADGQIISNGDNTDAKQLCGLVFDVQYVALANKFL EEVCSWDENMCFKLSQHIVGQYRSAPNKV SIGVYR
KQEEVLRKFRMQELNLLIATKVL EHGVDIPKCNLVVRFDPDRDYKSYTL SKGRARARDSDYVILLNEENEKLFSEDLKLF
KEIEEILIGKEKEKDIV EEDSSDYLDIVPKYPTNHNIDVHVS LNNSIALINRYCAKLPSDAFTHLTPQCTISTIQHGDN
ISYIGHIRLPINSPIKQIEGIAMP TKKLAQAAAQKACEMLHKTGELDEQLQPVGKEMFLQEQQEEDDEDVTSQGRPG
TTRKQYFFKVAEAFDKSHPDVNKKCYLYIINMVL TGRITEEQNTRGRSIYAPEETTRCLGILSSKQIPQVPMFPVYTR
SGEVTVSVDLQSDITLDNQLKQLQDFHLFLFTQVHLHLDKPMENCSGLADVGYLIVPVDPSLDALADIEWSFVEYVTS
NNSDQKKNVFD SKRDI FMDK EAFLD AVVMPSYRNIDQPQNFYVAEIRYDL SPLSSFPSELYKTFNDYYSTKYGLVITN
TDQPLLDVDHTSARLNL TPRYMNQKGVALPTSSAETKRARRENLQKQILPELCDIHVFPASLWRKAACLPAILYRMN
YLLLAEEIRLRIARETGIGVIDLPEGFRFPQLDFGFETSPDKIKLNEKSDSSSDSGIASSDSCLVRSSESDIVSSASDA
SIKLEDKLVDGTHTNHVEIFNDTMDVISNKL IATNINSESN TDKPEASIPSDSSDKTVIDISDSSPKDCKQCQKNILQ
YQNCESNVNDLKNLQDCVQNASD TFDVGVWNSPDVKIIEGVADDKNFSLPEISLDEDRDLDTFVGPSPCLILQAL TMS
NANDFFSLERLETIGDSFLKYAITVYLYCSYPGIEGKLSYLRSKQVSNYNLYRLGQKRGVAECMVS SKFEPYENWLPCC
YIVNEDRRRGVPVKVLIAGTKNDKASITSNFVLEESKPSKNHFP IGT FNTGEGKVNDFNQELKEAEELDELEKASEANQL
LIPYSLQTQHGI PDKSVADCSEALIGCYL TCGKKAALMFMSWLGLKVL PKKERVGDSIKKEGQEVRTL DHEYEELSCPV
SPLITDENSSRKKLEMLLNGYDEFEESEYKFQDRSYLLQAFTHASYHYNTVTDYQRL EFLGDAILDYVITRHL YEDSQ
KYSPIGLTDLRSALVNNNIFAALAVKWNFHQYFKAISPALFQVIEKFVTRQKDREREIDIDDEFNDKDEEDDDDEEVE
MEIPKALGDI FESVAGAIYLD SGMSLDAVWTVYRLMKPQIEKYLKICIPKSPVRELLEMEPETAKFEKPERTMDGKIRVT
VNVVGKGVFSGIGRNYRIAKSAAAKRALRCMKTMQQGLI

>Smed1

MPFKFPNKFNLNFPSPNLHNINRNAIRILGIPDPVNKL SANRFLFRVVPEVTTSDSSNENFNKLCCLIVNCSFTAYALS
RLLDEL CIWDPDLFFLHPGHADNQG VKNNDATGSVSLGQEETINKFRKGQINCLVTTQILDNFKFDLQKCNLVIRFDP
VLNLQQYFCSKSRVGIITNAASMLRKANSVADAI SHLSDRSKAKLIHMISKEDKHSILSHLEDYRKVETGIINRINKVS
SHKSDDHMNQNPVNVIDVNSLENIQSTGIDSQEVYHEIHRNESCIPETIMCLVEFTDTCRSNLDIYFALIIINAINRQV
TLSQAIALVNL YCARLPTDSTFKLTPELSIECDKNSQALRYCRLRLPINSPHRLVTKGKWMPTISTAKMSALNTLQKL
YDLRELDNHFYPWNKEDNREVVSIAESNKLDENIGGMLMGSSKCRYCYERQFAKCI STTPI SPNTPCYMY SINLELSRKI

Text_s1-1.txt

PDEQNLKGRKIMISADEAKDFGIICGQTGLDKTLKFPIFTRSGEEEVQIKQLLPLNSCDETWTFNEKEFEQKILEFHHLI
TSVFRKLKDKGMVLDLEDAVIKLLLVPLNKSTPFLILNLDNFLIRSYNSVENVCRLERTRKPNGMFEFNELDYQNAVVMV
SYRNFDRHQYYYVADIWMDKRNIAKILSSKELAFYFYDSWNMRPTTDFPGEHFSTFAEYTTKKYGAKITNYQQPLLDADF
SVMRLNLTIPRHFLQRKRNFASLTVKNNFLYRNCAIFTRFRQHVLIPIFIDELEYLSHLVWHKAASMPITMHRINQLL
AEELRSQIAIDVGLGLNFPINEKLIMDYNLSSGGLRNEEEVNRNTNLLIGSYLNETKIIIEGYYYMKLRMASEFLGENE
QEYLESVSDGRLSYLRSKQVSNLNLRAKEKNLAGFLNSQKFNPDQNWMLPGYKYTENIDVSQIINENYSNFCPRIQH
LTDKNLADVVESTGCVLARGEISALRLLQWFNINCIPIDEKATPGDGTGKVSDDLWIKPESPITVKLETDNDKICYNNLI
KDVIGLEAVIGYRFREKAYLLQAMTHSSYYLNVFTDCYQRELLGDAVLDTITRFLYEDCMRHSPPGDLTDLRSALVNN
IFASLVVRYNIHQYFKCMSPLWLKIDKFVYYQKIVMKDLDLDETQNNVDENLHNIDEEVEIPKCLGDVFEFLAGAIFL
DSGMDLNIWVVFHPLFHERIGKNLIHSMVDSERYTSLIPKSPVRQLLEAMPEIVNFELDLICSLKRSKILDLSLFRYP
KRTADGKTRVCVRIIGKGRFYGIGRNYRIAKSVAALKARLKLQIDNFN

>Smed2

MDGRKIIYITKALENLPKIDIPVYILIDPEPIAPPLNVKNNLVQDISKSHVQNNFELHCFNSEESKDKCDFIKHINY
EITNVVSNVLESTNLIAQSDFNDDFINSVDIISDPTDINDYDIKTPEYSSLYHDTNDSFMTDFYECLRPDSSLSEN
ESQVKLGQSLSSSTAVSKTLESFINIPKNSFPTSSNNLFNLNQEIEPIFVHEKNIANSNHSIIDSLFVADTCKENFNSNI
SDDTKFYCEEDYKISEQENTLSSNIINIKPDNIESNKILYQKQDLQVTKPFEPNGSKCIENYDLPTQFQFDESKFYSPKE
SSTVSNNFVSSNSGKSIKMPNPQASGYQLIQASHKLNKNNLHSIWNINTSLPQIDISKPPPTFVNSLSHPHLPQSFSV
PPPIDEPPTVKSEEIAKIKSRRSAKRVPKRPLCFKEYQYVMDINERYLVHNLKKKKTIKTLGTLMTLLEDKFEILSR
RETALDGVPKWEALDRSFTNTSTLFDNENEDDKITDKSNGSDNDSTISDLIHSDDDDLGESYKESNFTKYQISQLAH
IEKKVAATRKRKLNLSTESQIPKNKNNFNKRKRPNKHFSDHHVPAKKAQSTGYLTIQKNSLQNNAPVISKGRITIEQQA
EINHKLKHPKLKHPDIWYNEANELNDGPACSCSPKYQIGPLHNKFPGEKSIPLCQSNSNNGKLYHYRLIISPPINFLSL
DPTSITYKDKKEYSFTGYSIFTHKPLVDIPPCQVRYNIVYFLSLVKEEFPIGFTVRTLDLITDYVFKELLELLDLKWWPF
DTPDKTSCRVIHLLPRFERQYSANSVELLSANVILKWLKHAEKPLFDVMDLPRIHAMSQSEWSEYIDSLRGLATFPFGK
KPSTIRIDQIDRISIESKQNDVMIFFPVIVHMSMAPMKLGLTREPAYKRLLRNFLKLQYLILNRPEVTQEDRAELNNLAR
LLENFEMNGIQRREFTVEICSDGFYRTGFGTDVQFALCIPSIAMAHIRFHLSELENRMKYVFKDKSLLHQALTHPSFR
NTNYGTNPDHYQNTVTSGLRHRYGDKLNLRYSRFRKKGLSKMVQVMSHLPYWHETRSNIFGNERLECLGDVIELITSV
HLFYMPELSEGHLDAFRQSLVQNHLYLAELSVRLGFHNYILYTHNMDFSVDSTITHARSDGFVAVMGMKKIEFSQLFSS
VLAAITLDGGLDEADRIFGSALFADDPRIHKVWCEIPLHPLQEQLNGDRHVIRNSPLFENLTKLEDRLGIRFKHIRLLA
KAMTIRKTGYNLFTLGDNRLEFLGDSLLKYLTTDYLKFKFPRHHEGHISLLRNTLVNRYTQATVCTELGLDEFIICKDN
NKSNTILNKSNEKPKADLLEAFLGALFVDKDLKWVERFCQVCFWPRLVDFILNQSNDPKSRLQCCLYRSLNEEPEIA
HYKLITTFGSDHREFKVAVYFRGDRLATGFGRSIIQIAQMNAAASQALETHKTEFRQLNYQEKVITDHYDYNTREKIMNQF
ENWDELIDKFVTSMDPMPKPMHRKLVPTSNDMVSTKNPLLKSEKPKDRLNIAFTSSSLGELLKNMKKT

>Sman1

MILENLILCTFLSVGLTILVDSSLFLQDFNFQKWFSSGLCYFYLVVVDECHMVLDSNHPLSIVFGPNGPMLQNTEDFGPT
ATRQNMDSFNVSVTGKVRILCFTSALLPREITDPEKARLRLQTLKRTSCRLETASELLTMLSGLGARPQERVLCSKSS
IAAIPFHQFMLGVFREIKEFIIDIESNLPCTKSSIGDSTSVINSGGFRICVLDYCKRAISQCEEILNELGLWCAAQIV
RVFAKHLIALDRQRSEILKNSDILKDNEYSITIPKINDNEDRISYLLRLTVTQMCFLSRLQMEFDSILTLEEFQRMISP
KVLNLEIQLKLYKPSMNFRIEVAELPTAKNPPITITNNKRYKGGSRQNSRCQTSKISNSSLSSMGSICDVSLSDSM
SDTMSSLSDDDDDTRSVRSLNSTKSRLSTQSMKNLNSSSKSRKRKNRSNLSINSFNLHDLHFVPASTLNNGGIRPDVPS
TLVYRAVLNTDNRHNKHNRSRSGFIGQESSDMTASTNNNDVSNRNL CGLILVPCQFSAYALSRLIDELCIWDVDLYFIKI
GHLFCRQTLLKEDNDLSTTTNQTNLKTKHGTFNSYKTLVKFSNEPVNNDSSSTNNNSNCRNPFVSNQEETITNFRRA
INLLVATQAAISTVTTSGTELPRCNLVIARLPLNSLAEYLSKARSRLVNYGAKVIYLMHDHDSIEKINNNSNNNNNN
NNNNVKNQSNLNDKRSKQTDHSTVDSMYSNHNLDQFLGNFQQLQLLLQRCRGNVFADEHAIDPTVVDKILPPIF
PRGPKPKFCLSKAINIINQYCARLPSDYITNLTPKWYKIFPQPQGFTPSSTGPGSTCGSLKDLQPDGTFNLYQCVR
PINSSVKETIVGEPMVCKKLAKYSAAFNAIHLVLSGEMDSKWLINRETNPMHHLTSENYGNLKLSSQHSMPMKYSFSS
LNGSANIDSESDYGDYASSIGCCESADDEGEMKSNIQNTVKRRQYRKFPTQISNCLPQPNEPSSNYLYYIDMHLVKPF
AEQQYLGRVCHHPENEPIGFGLLTTKPLHHPIFIPIFSRSGEEKIRFIELWSPKSNYQFNDQYLPQGPILTQEIDRL
IKFHRLFQEVLRFEKDSVLEFNHFKAYLQVLPVARRDTCSDWDFINLVTSSCDKSIHLLPRRMNEISQEWKERI
DQIKLQAGVTSNGSKTFNTSNTHHHNRKVGSDSVASILVNLRLQKSGIIESNSVNNKPSIFEFMRFEFVNAVVTSPGYR
NLDQPQYYYVAIIRNDMSPLSPFSPDKFRNFAAYYINKYNALITTNQPLLDVDTLVLRLNLLIPRYMNIYGHNTNAND
NNGDINNSKSNVYTHKQYQQNHDDHQGEDLTNKQLLPELFCRHSFPASVWRKAVCIPSILYRLEHLLLAEELRHRI
ACETNLGCAYLPECNKILKDDFVNLSSHYNKNGDVNDPDTTSCLFESLNIHLPIAPDHVVNNESTSKSNNTYNNNERKP
TRTNSKFLGSSGGKSRGGRRKQNCRNTRAKNHNVIIGTTTKVKSTENVPDLSDTCELNKKMSTIELNENKDKSSYS
VNSHTSTDQINKEQNVVISSDNDNIYDDEDHRNSIDAETKENHDSNDNLVVNEDDGDGDLLEKPTILKLHATSAS
MVHNNKHRRHQSSAQSTTTALSSLPYDHFSLSNHNLESLLQETKVIELNSNEKPSNIRKRAYQPGPTTILQALTMSCS
NDFINLERMETIGDSFLKFVVTVHLYLTYPEAHEGKLSHLRSRIVCNLSYLRLGKAKDLQNRMIGCKFEPHENWIPPGY
VRQDKRLNNEIIKKFNENRLVIWSTDTLMDDEVLRNIDFIDENKIKPIENFPSEWDPNDPKVLAHQHLNNQYLITIQ
AIPDKSIADCVEALIGCYLTTRGERSALRLMQWFGIDCLHKSNSQPTARAPWSLPSKNYLDTDENRANLNEARLVWRF
ELESSLYTFKDPSSLIQAFTHPSYHQLRVLSTSNLSQSNLMFSTDLDCYQREFLGDAVLVDYITRFLYEDSKQHSP
GVLTDLRSALVNNIFAALAVRIGLHKYFRASSPOLLHTIDAVFVRYQKDVAKDDLDFITNESQSKPQHQTIEQHLK
SGKSNANNTIPTNTGHLTTNRLSDDVEIPKALSDIFESLAGAIFLDSNFSLDTVWQIFYPIMKRIERYTACIPKSPVR
QLLELEPEGTKFERARRMVDGRISVCAHVLGKGRFYGVGRNYRLAKSLAALKRALRVLRLNLPQSQTTPTVTTNGDYDQSS

>Sman2

MWIPPVSDNPTPFNIERELLPTVISEPSSKTSLSSTGTLDRDSLKENCPQYTPQNVRYQHYVPKVRCLSFPYQTLPS
CGSAVYLYTWLPRSDIMTQCLSLAAACFQDTSTTVGLLFSKPIKSKGYVCRIPLYLTRGMARSVRNLTKGFTLSEKEL
CVESAHHVIAELITNIAHMSHVGYKPSDPLLRAHFDRQNASTNIPESLRVVLDFRFDLSFNPEKANRLSLVTLIRLPDY
TYDYEALGALLEWSLRCRDKGISTADDNFPTHVKPVHGVIPKWRIRLSQIPQDNWVGLVVKPNHLPSDPMGFSISKVSSE
TASSPVPETLASKLGDIQNVKGASPLTYSYYRIKYQMKTISFTIDPSLPLFECLRITRHQNSTQVNAGLRKNKASIKP

Text_s1-1.txt

SLFLSDACL VHPLSSWIWFQVSLVPTILYQMSRALLASQLFTELNYELKSPQFQSIRINSLTISNDSLHTTILVDPRL
SVPVFLNGEPMEQNI FDFEVD RATSEIDKELDNKAHSESELEKSTICHPNNLIEPTTLGARDAVDLERLEFYGDSFLHF
IATLCVYGTNPQDADEGLSSKRGSLSVNAHLCDIACDLKWYEYCTGQTFSPPEHFLPPCYSVASEACKYDPRLYTRLTD
KSLADMVEALIGCFLLRSGLSAAFNLHYFQICPVSWSTLKNKHYKVEAPWHFFLQPKLYSELNCGNALNASCRSLHIS
KEIILELNQRRFFQLQEKLGYYFKGIELLA EAMTHQSSINKQYWGNYQRLEFLGDSILGHIISNYLFQKCPNSSPGALTTA
RSTVVSINLANVVVEHDYIPFIDFGNCSLKACIDDIKSIHRQTDSHFERIELITKKVSSGLNVKVLADVFE SILGAI FV
DSNGNHQVASSIIHRFLGKMTTQLLCHQSYMKIFMIRLHRLMVSAYV

>sjap1

MVFRDVKYQIAKKILENVPNDHFLPPPYQLDLLDAAYERNTIICSSNELNKMFFVNLIREMRYDNPGKQVIYFTDDTK
MNSINNFITHHTGLSCRGFYDEHNNMKIDFTDWLNSIVQKYDIVLVSPWLIPVIVMLSSINFFIKFIQSSYLLVDEC
HMVLDPNHPLSVVFGPNGPVLRNTDDHEQTVGKQNIHDHMSLLGKVRILCFTSALLPREITDPGKARLRLQTLEKRTNC
QLETASELLTMLSLGAQPQEQVIFCRKSPLYATTPFHFKFMLGVFQEIKEFINDIESIPLIGDSTPAIINAGGFRI CVLDY
CKRAITQCEEILNELGLWCAAQIVRVFAKHLIALDRQRSEMFKNFKLVNNTNENEDESSCKKEELKDSKYLVITPKI
NDVENRISYLLRHTVYLMCFLSRLFQMEFDSILTLKEFQCMSIPKYLHRIEQLKSYKPNINFRIEVAELPTAKHSPITIT
TNNKRYRNGGGYRRRHRGLSQTSSMSMGSFISCDADSLSDSMSDTMSSLSDSDVDEETRSVRSLSLTKSNLSLNSF
NLRDLHFVPASTLNNGGIRPDVDPSTLVYRAVLNTDNRHKNNRSKLGKESSDITSTTNASNTNNISSNRLCGLILVPC
QFSAYALSRLIDELCIWDVDLYFIKIGHLCFRQTLKDDNDITGQLLNKTKYGTNLNSYKTLVKFSNEPVNNDHSTNN
NNRNPFVSNQEETITSFRGAINLLVATQAASAVTISGTELPRCNLVLALRPLNSLAEYLSKARSRLVDYGAKVIYLM
DSVENENSQSKLDDKQIETQLSNIISSTNNPSTYDNTINDQFLGNFQQIEQLLVQRCRGYGVFSDEHAVDPTVVDKLPP
GYRNLDQPQYYYVAAIRNDMSPLSPFSENYRNFASYINKYNALITTNKQPLLDVLTVLRLNLLVPRYMNIHGHLTN
RLPINSSVKQTI VSEPMVCKKLAKFSAAFNAIHLLYISGEMDTKWELTRNKISPNSLNDNNLKLTYDSPLKYQSFSSNDP
TNNVDGESDYGDCYASSTYCESADDDGDELKSIPTNSAKRRRYRKFPTQLSNCLPQPGSEKSPNYLYIDVHLVKPF
SEHQNVRSRVCHHPPEDEPIGFGLLTTKPIHHIPIFPIFSRSGEEKVRFVLELWSPKPNHQSNQYLP IQSPSLTQE QIDQL
VKFHRILFQEVLRFEKDSVLEFNFEKAYLQVLVPPARRDTC SIDWDFINLVLTSCFDKSI CRLLPRRMIEVSQEEWKEQI
DQIKLQAGVVTGGCKTFNINNNNRQQHHHHRTRGGGSIASILVNRLOKSGIVESNMVNNKPSIFEFRVEEFNNAVVT
GYRNLDQPQYYYVAAIRNDMSPLSPFSENYRNFASYINKYNALITTNKQPLLDVLTVLRLNLLVPRYMNIHGHLTN
NTTSTDDKSNISNSSVIGSSHKQHNRQKQHQYQYDHHEDLAHKQLLPELFCFRHSFPASVWRKAVCIPSILYRLEQLLL
AEELRHRIAYETNLGYAYLPECDKITKDDFINFSHHYKNGMSDADSDCYLFEPLNIHLPIIQDQVDVSVNNEFTSKYL
TNNSDNERQSTKPSKFLTNVKGKRGAYTRRNQKRKTRKEDSMDINKMKLPTENVQNDQSDVFGLEKT VATMGLNHNKNE
SYLFSNDNDINNPLVPKNEQNNDMISSDSDENDTISFCNEEDHNTCDVERRKDNDDILGEDDSSL SERKPTILKLHAT
SASMVNNKNHTIHTSSIPPTLYNQFSLNGLNETLLQETKVIELSDNDVNQLNNSDDGDELRVVENYTNMDEKLELDMS
DTSVDSFEGNIRFFPSEDDDENVTNNVSEAYILTKLQEKSTSI NSIRRAYQPGPTTILQALTMSCNDFINLETI
GDSFLKFVVTVYLYLTYPTAHEGKLSHLRSRIVCNSNLYRLGKTKDLQNRMIGGKFEPHENWVPPGYVVRQDKRLNAEIA
RKPTPNRNLMIWSTNTLMDDEVLRNIEFIDDKKIQPVENFSISEWDPNDPKVLHAQHTIIGCYLTTRGERSALRLMQWFG
IDCLNISGSTQSPGTGAPWSLPKSNYLDTYENRAHLAEARLVWRFDELESSNYTFKDPSSLIQAFTHPSYHLRVLTTSD
QHQRMTFPTDLDCYQRLFLGDAVL DYVITRFLYEDSKQHSPIGLTDLRSALVNNNIFAALAVRIGLHRYFRASSPQLLH
TIDVFRYQXEIEIRQPLADTFIETVNNLTNSGLASSTSDFLPTTIDKQKQSTFTKQRHHDSGNDEVTVVANECND
EDLDVDEVEEEDHAPLLNEAKECESKWQAKDTGLTGLFVEDTSQKQSEPPQHL SKPNNENNSTFNNTTSMNTTII
LNNLTSNRLSDDVEIPKALGDI FESLAGAIFLDSNFSLDTVWQVFYIPIMKERIERYTACIPKSPVRQLLELEPEGTKFE

>sjap2

MWTPPICDNPFFNLKELLPTVISESSSKSGVSSSTTDGLDYDPSQGNCTQYTMQNI RYQFYPIPKDYVCRIPLYLTRG
MVRSRVRLVRKILNENELKCIESTHSVIAELITNIGHMHSVGYKPSDPLLRAYFDKQGHLLNNVPECFRVTISNGFADLS
FNPEKAIILSLMTLIQLPDYTDYEA LNALLESWQCRANASSADGDL SVNVKPVHGIVPKWRIRLSQIPQDDWVGLVVR
PNHL PDGDPGMFVSVAVCSETALS NVPRTLASKLNEARSPEGASQVTFDYTYIKYQMKSISSINLSLPLFKCMRITR
HQNSAQVNAGLRKNKLSVNPSIFLSDACL VHPLSSWIWFQVSLIPVILYQMSRALLASQLCTELNLELKNPQPFQSIRIN
SLNISYDSLSDITVLPDRLSVSVFVSGERMEQNI FDFEVD EATSEKDKELDNKAYNESVKESKICHPNNLIEPTTLG
ARDAVDLERLEFYGDSFLHFVATLSVYGTSAHDADEGLSSKRGSLSVNAHLCDIACDLKWYEYCTGKIFNPPEHFLPPC
YTVASEVRKHDPRLYTRLSKSLADMVEAIGCFLLRLGLPAFNLHYFQICPVNWNTLNDKHFVDVAGWHLLLQPNL
YSKQNVLLNSSRSLYISLPETIHGLNKKFSHLGKGYFKRIELLA EAMTHQSSPNKQFWGNYQRLEFLGDSILGHIIVS
NYLFRKCPSLPPGKLTAA RSAIVSNNLANTVVEHEIYPFIDFGTCSAKACIDEILNIHHQTELYCERIELLTKKVSSGL
NIKVLADVFE SILGAVFVDTGNHQIIQITIVNRFLGKSMHTVIINLPVQPLQLNSPYPNTRYKDSIAIVSSNLDDSN SVS
DAVWTSSKRLCATESVRDETHLRF AEDFDRLLNDKENRGNPNF

>Egra1

MPRKHLIVRKILENVPNHFMRPYYLDLLDSAEARNTIICTDEDVTLQFVAQTLFGDMAIKYPGKCMVIVISDASSVKS
AAETFKYVTDTPVLVADASSKIPTTTEVIFMSAETAANLCSGEEPVFSIHERAFLILDDCELVLSPGHPYQRLFYYSK
DPLSEHQIEDIWRNDPFSYSSTSASYRILALTSTILPVGCSHLLAKRYICSLEDRLGCRLETTSELINLLSLGAQPEEE
EIPAPDTTNGDNDQHSIRGTILRLLVEARDFLSDVDFTGSEENVATTGQRVINVPYYCLRSVAQCEAITKDLGVYCGA
LIARVFLRHL YRLERARASILEARRQLEHGEDLLTRILRYTATQLSMIARLFQTVCDNCMTREEVRTLASPRVLQLVDRL
RTLKPTGYSYRIEAVWRTEGVSED TGSDSDAESSGLTSHGWSSSRRRSRLRGRSTSSGFVAPHGFEAMSGDDDEDDYGE
GYDDSSLFSGVDSVYSEDDNDFKDCAYNDSNGKDSADDNDGESGGESDRLDVSMGLSSGIRRKSTRKPCWRPHKIRSPVI
RGSRGSLVYRVL PDNPKENTVPPLCGLLVGGCQFSAYALPRFIEEYCNWDPEIAFVRPGYFVFP GASDQVKTLLS
ADV GSTQEETVANFRCGGEVNLVTTQSGLAAASSLPRCNLIALFHPSTSLATYLA AKSRLRLVAVKSTKAHFHFLESD
GFSLRTRFQEMERMLVQRCNSYLSISDEHDVDADYVNEQFTFLPKRKACRAIDVVNHYCARLPCNFMTRLAPRWHLTTRR
LAADSRGAVEGGDHTALLTCPDP SHAATASNMIYRICALQMPINSPVREVVWSEWSLCKKMAKYSAAVRVGQLL METREI
DARGRPLRHTPIA AVAAASRRLRATRERSPPKHKSSYLKTPDLSLVKLAKEADRGSIGQNRLLYLDIRHLDLTA S RLQR
EHRSQPTSSHIDFPEAENRGFAMLTTKPLPFIEPESIFTRSGEETVHIYEAESSCPFLNIDQLNRLSRFHRIIFKSURLF
EKDCLMKFNLVDAPCQLLIVPVKRD SQKIDWLVNCLNSYASVGTDCRLPVRRRPAPPLQLPTGPVLTQLLSQTSSPR
TVFHFVQDDFVNAVLTPTYRNVEYPPRYVVAVRNDLSPATEFPSPQYATFADYYACKYGLRLSGAPSDQPLLDAAAYAPL

Text_s1-1.txt

RLAPLLVPRGCVARQKSNNNGVASGNATSTSKHVLPRHIQLLVPEFCHRHPATVWQKAICLPSILYRLQHLLLAEELRS
EIARETGLGVVTPPHRFPCFPPLRCYFTQVPSPPSPSPSSKKQRDPEDRDSEPEVVKLTLSPLIQSKPRNPVTDFOA
LLYESDRLESDEGDLPEAEVEVEVDAKSIDSEENATNMCFPSSEDDDEGEDEKQIVYKPGPTMILQSLTLLGAGD
FINMERLETIGDSFLKFAVTMDLFLRYPEAHEGHLSSLRSQIVGNSNLRYLGCENLPGRIVGVAFEPHENWLPCCYVCT
TDPTNIIASTRIKNQCPSLTHQFLTNKSVADCVESLVGCYLTERGERSALKLLQWFGIECLPPAGEMQPAGSPWKLAEEK
PLSKVEEATVERVYRAGRFSKLEEIIGYTFKNRRLLEAFTHSTCRDLHGPPSDDDDGGFHGSATSISGGYERLEFLGDAV
LDYAVTRFLFTTDATARLSPGALTDLRSALVNNVVFALAVRHGLHAFLRAAAPPLTRTMALFLRHHHDVVRGDLVDLIA
PEVNVCRENKGATASQAEGDRANEEVEVPKALGDVFESLAGAVFLDSGLCLNTLWRVFFPLLRERIERYSTCVAKSPVRR
LLELCPERVKFERPMVRPDGKMRVVVRVAGIGRYVGVGRTYRLAKSAAADLAYRRVTEASTEQU

>Egra2

MKRSLPAFRETLLLEPRPYQVEIANMARNESIIVKLSTGLGKSFIAAMVIKDHLPETYRPVTEVHLVYQMAEFLRSQLPV
PPNEIGIFHGVVAPSVIDSWTKASFWPLLPFLRIFDENESFDGFGQIWAQLEKHKVLVATAGVLRDVLNHRKLSLRD
VCLIIFDECHHADPNKSDYTDICQHLHAFPPGRWL PDYSRGPVKVGLSASLVNLLKRGESMETKVRLLEQLMRARVVT
TDSSINAVMGQRHLEIKHGCGRTRALNDPYNFSQVLLRGIEKVKQLHSGVVGEDCIDLESLLDENKIKKEAFLSVMEL
LHEFNSAKIKRILLQCVRVMEEFVGYCAAHACEQFEKVLKGLLLSSPKVITNPQCVNIIQMCLASMAEALETYRQIRQT
LIAKLTQPLASNPTDPAWGEWIVAHFPMSSKMVSLRLLHYYHKVVGKNNETLSALLLVKERIAAASVAHLITEVSAMAP
LYTGLKASHCVSANPTNPVASTNEQLNVLDGFRNTAFNILLVATNVVEEGLDVRSCNVVVIKADEL TNFRSFIQSQGRAR
AKTSYFVLLTDDVSKCRTNVESFFRMSKVIASVFWLVCSFDPFLTLNLVIDDFLECRICIDSMPDEEEEEEVEVSAIT
DIARQSNLAYMPYGTGPRITITGASSVLMRYTSLKTDTSYPLKILYSVHKEFGGYRVQMLMPPPSPLQDIIKGPLAST
RKLAMQLARLEACKLLHNAVICGFGSSSELETGCVTSEAAPKNVGVKVIASARLQELMESLINKISNLDLGDPSKDGEL
KPPSIMSQARTLMTVPVGEVSTDAASSDHLVDQLSSSDSDNVQNEVDASGQDLGVSCPELPPGIVYCGDSPLPEPSFTS
TRTVDLVVDGGARPHLFDNCITDGELPVKPKKPPASRPLNIARFYPIHKLSSLAPYQPIEANRPVHLYTWHMPLPEHV
AYLTKVDAAYFRHTDQTIIGLLFSKPIKSDFCRMPLYMQSGMRRVLRRRATVTLTEDQVELALRTHQILSQLSVDINK
PVNFDLNFNSNNQFSFFSTPFSTVSSFSPTPVVSSCVSTISPFNTDFAIDPFSELKVDPLNAQICGILLIVRLSDMVFDE
EASHALVRWAEERDFHVSQRQKQKAMPPLRNQSLQSGICSHYGVQLSTIPTAQWGLLVRPIDLPEDPGSYAVLGPNS
SGLCASSPVPDYMAQHLPEELRQRGATSYLDYAKFKHGTRLRLVQETRGLDPSTPLVTATRISRHRNAANVMMSGVGVSGK
EVIKEERIAQLCIVHPLNTWLWLTCLIPTVLHQVYRSLSIGELANHLREILYSDNHAQSAHPPHPTPLDLPAGEFLMPD
RHSVQPTCIPSIRKFLVQDQDQSGCRFKGPVVVFLDSETS LDRAVEDRCTSKAPAEVPTYVGLYEAFSTVNAFEAVNLER
LELLGDSFLKFAASLLLFATSPAHTDEGQTYARIAHISNANLHRISMKFGFLFRYFCSCVFKPEAYLPPYYAIADPARA
CQRHDMRMFVKLYDKSIADSMESLLGLCLLNLEAPRVGRLNLLFKLSNEPNPLSVLVRAESVGREAYPSWVPLLSPEHR
VIDPTFAVEDDSLMAKRKAEVESMLLASIPPLPKSSATGETAAEAPEAVKSTPKEISEMMEERLELQPLEDIIGYFRRI
RILLQAITHPSSRLAFIWGCYQRLFLGDAILDVFTVQRIYKDHPNMDPGLTDLRICALVAVVAVRSGIYKFLLEY
LDPNLWSFINSFSDAVSKGVCNVWKL EHDNFNERNEMLSYKVLGDMLEAIIIGAVFVDSGGVTSVVTGVIYHLLGREFEAYG
KDLPMDPVRMMHELYPDLEISVAECLPMNNASGNSEKAKAEANAQRQRVRSANYKGRRLSGEGFNLRATAKLIKIAQQLG
IGFA

>Emu11

MPRKHLIVRKILENVPSNHFMPPRYLDLLDSAEARNTIICTDEDVTLQFVAQTLFADMAIKYPGKCMVIVISDAGSVKT
VAETFKYVTDTPVLVADASSKIPTTTEVIFMSAETAANLCSGEAPVLSIHERAFLILLDDCELVLSPGHPYQRLFPYSK
DPLSEHQIEEIIWNRDPSFYSSTSASYRILALSTILPVGCSHLLAKRYICSLEDRLGCRLETTSELINLLSFGAQPEEE
EIPAPDNTNNGDNDNRNSIRGTILSLLVEARDFLSDVDFTGSEENVATTGQRVINVPYYCLRSVAQCEAITKDLGVYCGA
LIARVFLRHLYRLERARASILEARKYLEHGEDLLTRILRYTATQLSMIARLFQTVCDNCMTREEIRTLASPRVLQLVDRL
RTLKPTGYSYRIEAVWRMEGVSEDGSDSDAESSGLTSHSSWSSRRRSRLRGRSISGGFVAPHGFEAMSGDDDEDDYD
EGYDSSFSFSGVSDYSEDDNFKDACYNDSNGKDSADDNDGSEGGESDADTVSMGLSSGIVRKFSTRNPNCRDDEKIRSSP
VIRGGRRGSLVYRVLDPNDPKENTVPPLCGLLLVGGCQFSAYALPRFIEEYCNWDPEIAFVRPGYFFVPGASDQVKTTL
LSADVSTQEETVANFRCGGEVNLVTTQSGLAASSLPRCNLIALFHPSTSLSTYLAAKSRLRLVAVKSTKAHFHWFLE
SDGFSLRTRFQQMERMLVQRCSNSYLISDEHDVAADYVNEQFTFLPKRKACRAIDVNVHYCARLPCNFMTRLAPRWHLTT
RRLTADSRGAVEGGDHTALLTCPDPASHAAIASNMIVYKRALQMPINSPIREVVWSEWSLCKKMAKYSAAVRVQGLLMETR
EIDARGRPLRHTPIAAVAASRRLRATRERSPPKHSSYLIKTPDLSVKLAKEDRGSIGQNRLLYLDIRHLDTPLSRL
REQRSQITSSHIDFFPEAENRGFAMLTTKPLPFIPEFSIFTRSGEETHVIYEASSCPPLNIDLRLSRFHRIIFKSVL
RFEKDCLMKFNLVDAPCQLLIVPVKRDSSQKIDWDLVNLCLNSYVASVGDCLPVRRRPAPPLQLPTGPVPLTQLLSQTAS
PRTVVFHVRDDFVNAVLTPTYRNVEYPPRYVAVRNDLSPATEFPSPQYATFADYYACKYGLRLSGAPSDQPLDAAAYA
PLRLAPLLVPRGCVARQKANNGVASGNATSTSKHAI PRHIQLLVPEFCHRHPATVWQKAICLPSILYRLQHLLLAEEL
RSEIARETGLGVVTPPHRFPCFPPLRCYFTQVPSPPSPSPSSKKQRDPEDRDSEPEVVKLTLSPLIQDKPRNPVTDFOA
QALLHESDRLESDEGDLPEAEVEVEVDAKSIDSEENATNMCFPSSEDDDEGEDEKQIVYKPGPTMILQSLTLLGAGD
GDFINMERLETIGDSFLKFAVTMDLFLRYPEAHEGHLSSLRSQIVGNSNLRYLGCENLPGRIVGVAFEPHENWLPCCYV
CTTDSTNIIASTRIKNQCPSLTHQFLTNKSVADCVESLVGCYLTERGERSALKLLQWFGIECLPPAGEMQPAGSPWKLAEE
EKPLSKVEEATVERVYRAGRFSKLEEIIGYTFKNRRLLEAFTHSTCRDLHGPPSDDDDGGFHGSVTSISGGYERLEFLGD
AVLDYAVTRFLFTTDATARLSPGALTDLRSALVNNVVFALAVRHGLHAFLRAAAPPLTRTMALFLRHHHDVVRGDLVDLIA
IPEVNAACRENKGATASQAEGDRANEEVEVPKALGDVFESLAGAVFLDSGLCLNTLWRVFFPLLRERIERYSTCVAKSPV
RRLLELCPERVKFERPMVRPDGKMRVVVRVAGIGRYVGVGRTYRLAKSAAADLAYRRVTEASTEQU

>Emu1 RNC3.1

MKRSLSDSRETLLLEPRPYQVEMVNMARNESIIVKLSTGLGKSFIAAMVIKDHLPETYRPVTEGGKRIIFVAKTVHLVYQ
MAEFLRSQLPVSSNEIGIFHGVVAPSVIDSWTKGIWQAQLEKHRVLVATAGVLRDVLHHRKLSLRDVLCLIIFDECHHAD
PNKSDYTDICQHLHAFPPGRWL PDYSRGPVKVGLSASLVNLLKRGESMETKVRLLEQLMRARVGTSTDSSIDAAMGRRH
LEIKHGCGRTRALNDPYNFSQVLLRGVEKVKQLHSGVVREDYIDLESLLDENKIKKEAFLSVMELHEFNSAKIKRIL
LQCVRVMEEFVGYCAAHACEQFERVLKGLLLSSPKVITNPQCVNIIQTCCLASMAEALETYRQIRQTLIAKLTQPLASN
TDPWGEWIVAHFPMSSKAVSLRLLHYYHRVVGKNNETLSALLLVKERIAAASVAHLITELSAMAPLYADLKASYCVSA
NPTNPVASTNEQLSVLDGFRNTAFNILLVATNVVEEGLDVRSCSVVIKADEL TNFRSFIQSQGRARAKTSYFVLLTDDV

Text_s1-1.txt

SKCQANVDSFFRMSKMIDDFLECRCIDSMPEDEEEVGSAITDITRQSNLAYMPYIGIDGRITITSASSVLMRYTSLKTD
TSYPLQILYSVHKEFGGYRVQMLMPPPSPLQDIIKGPLASTRKLAVQLARLEACKLLHNAAGLLGADLLPINFRQLVELSK
HKASSTGTEQEIEGLISRMSNLDLGNPSKDGETEPPSIMSQAHLMLMPEEDINTDSAVSSDHLVDQLSSSDSDNVQNG
DVDSGIGLGDVCEPLPPGIVYCGDPPMVEQSFTSTRTVDPVVDGGARLHLFDNYITDDEL PVK K K K K K PASRPLNIARFYP
IHKLSSLAPYQQPIEANRPVHLYTWQMPLEPHVAYLIKVDAAYFRHTDQTIIGLLFSKPISKSDFVCRVPLYMQSGMMRV
RLRRRATWNWLSAHTKSSPNSPWTSTSPSTSTSVITSFHFSRPPSLLSPPSLPPLWSPVSVVDPLNAQICGILLIV
RLSDMVFDEKASQALVRWAEERDFHVSQRQQTAMPPPRNQSLQSGICSHYGVQLSTIPTAQWSSGLVLRPIDLPPEDPG
SYAVLGNSSGLCASSPVPDYMAQHLPEELRQRGATSYLDYAKFKHGTRLRLVQETRGLDPSTPLVTATRISRHRNAANV
TSGVGVSGKEVIKEERIAQLCIVHPLNTWLWLTCLTPTVLHQVYRSLSIGELANHLREILYSDNHAQSAHPPLPTPLDL
PAGEFLMPDRHSHVHTCIPSIKRVLVQDRDGSRRFKGPVVCLSLDSETFLDRAVEDRCTS KAPA E V P Y M V G L Y E A F T S E N
AFEANLERLEFLLL FATSPAHTDEGQTYARIAHISNANLHRISMKFGLFRYFCSWSIADSMESLLGLCLLNLEAPRVG
RLLNLFKLSNEPNLSALVRAESVGREAYPSWVPLLLSPKHRVIDPTFAVEDDSLMAKRKA E V E P M L L A S I P L L P K S S A T
GETAAAEAVKSTPKEISEMMEERLGLQPLEDIIGYFRIRIRILLQAITHPSSRLAFI W G C Y Q R L E F L G D A I L D F V V T Q
RIYKDHPNMDPGLTDLR I C L V S N I N L A V V A V R S G I Y F L E Y L D P N L W S F I N S F S D A V S K G V C N V W K L E H D F N E R N E M L S
YKVLGDMFEAIIGAVFVDSGGVTSVVTGVIYHLLGREFEAYGKDL P M D P V R M M H E L Y P D L E I S V V E C L P A N N S S G N S N S T
T L S E M E G V N A G G Q P C Q R V R V S A N Y K G R R L S G E G F N L R T A K L K I A Q Q L G I G F T

>Emu1 RNC3.2

MVNVARNESIIVKLSTGLGKSFIAAMVIKDHLPETYCPVTEVHLVYQMAEFLRSQLPVSSNEIGIFHGQVAPSVWIDSWT
KASFWPLLPFLRIFDKNESFDGFQGIWQAQLEKHRVLVATAGVLRDVLHHRKLSLRDVCIIIFDECHHADPNKSKSDYTD
ICQHLHAFPPGRWLPDYSRGPVKVGLSASLVNLLKRGESMETKVRLLQLMRARVVTSTSSIDAAMGRRHLEIKHGCGD
TRLALNDPYYNFSQVLLRGVEKVKQLHSGVVREDYIDLESLLDENKKIKKEAFLSVMELHEFN SAKIKRILLQCVRMEE
FGVYCAAHACEQFERVLKGLLLLSPKVIITNPQCVNIIQTCLASMAEAL E T Y R Q I R Q T L I A K L T Q P L A S N P T D P A W G E W I
VAHFPMSSKAVSLLRLLLHYHRVVGKNNETLSALLLVKERIAAASVAHLITELSAMAPLYADLKASYCVSANPTNPVAVSL
STNEQLSVLDGFRTNAFNVLVATNVVEEGLDVRSCSVVIKADEL TNFRSFIQSQGRARAKTSYFVLLTDDVSKCQANVDS
FFRMSKVIASVSACLFRPISNAVKPLKMIDDFLECRCIDSMPEDEEEVGSAITDITRQSNLAYMPYIGIDGRITITS
SSVLMRYTSLKTDTSYPLQILYSVHKEFGGYRVQMLMPPPSPLQDIIKGPLASTRKLAVQLARLEACKLLHNAVIRGFG
SSSKLETDGVFTFEAAPKNVGVKVIASARLQGLL GADLLPINFRQLVELSKHKASSTGTEQEIEGLISRMSNLDLGNPS
KDGETEPPSIMSQAHLMLMPEEDINTDSAVSSDHLVDQLSSSDSDNVQNGDVDSGIGLGDVCEPLPPGIVYCGDPPMVE
QSFTSTRTVDPVVDGGARLHLFDNYITDDEL PVK K K K K K PASRPLNIARFYPIHKLSSLAPYQQPIEANRPVHLYTWQMP
LPEHVAYLIKVDAAYFRHTDQTIIGLLFSKPISKSDFVCRVPLYMQSGMMRVRLRRRATVLTEDQVELALRTHQILSLS
VDINKPVNFDLNFSSNQFSFFSTPFSTVSSFSSTPVSSCVSTISPFNTDFAIDPFSELVDPVLPNAQICGILLIVRLSD
MVFDEKASQALVRWAEERDFHVSQRQQTAMPPPRNQSLQSGICSHYGVQLSTIPTAQWSSGLVLRPIDLPPEDPGSYAV
LGNSSGLCASSPVPDYMAQHLPEELRQRGATSYLDYAKFKHGTRLRLVQETRGLDPSTPLVTATRISRHRNAANVTSGV
GVSGKEVIKEERIAQLCIVHPLNTWLWLTCLTPTVLHQVYRSLSIGELANHLREILYSDNHAQSAHPPLPTPLDLPAGE
FLMPDRHSHVHTCIPSIKRVLVQDRDGSRRFKGPVVVFLDSETSLDRAVEDRCTS KAPA E V P Y M V G L Y E A F T S V N A F E A
ANLERLELLGDSFLKFAASLLL FATSPAHTDEGQTYARIAHISNANLHRISMKFGLFRYFCSCVFKPEAQYLPYYAIA
DPVRACQRHDMRMFVKLYDKSIADSMESLLGLCLLNLEAPRVGRLNLFKLSNEPNLSALVRAESVGREAYPSWVPLLL
SPKHRVIDPTFAVEDDSLMAKRKA E V E P M L L A S I P L L P K S S A T G E T A A A E A V K S T P K E I S E M M E E R R L G L Q P L E D I I G Y
RFRIRIRILLQAITHPSSRLAFI W G C Y Q R L E F L G D A I L D F V V T Q R I Y K D H P N M D P G E L T D L R I C L V S N I N L A V V A V R S G I Y
KFLEYLDPNLWSFINSFSDAVSKGVCNVWGLEHDFNERNEMLSYKVLGDMFEAIIGAVFVDSGGVTSVVTGVIYHLLGRE
FEAYGKDL P M D P V R M M H E L Y P D L E I S V V E C L P M N N A S G N S E K A K A E A N G Q R R Q R V R V S A N Y K G R R L S G E G F N L R T A K L K I
A Q Q L G I G F A

>Hmi c1

MGRKHFIVRKILENVPPNQFIPRPYYLDLLDAVAERNSIICKDKDITLQFVIQSLYRDIASKNPKKMLVIINDSTLIDD
SAKIFQYTTDDNVFIADITSTSPENVDVAFMSAETAAKLCAGVAPAFSILERSVFLVLDCCDLALKSDHPYHTLFPVAVK
CELTEEEINTVWDRPAEFYSFVTPYRILGLTFKILPDGCSDFHFLAKRYICSLEHRLNCRLETTSELYNVLSFGAQPEELE
IYTQAPSDNGSDQHPICRTLLRLLQEAHDFLSDLDFTDEANVSTGGERIINVPIYCLRAVAQCESITRELGVCYCGAIIA
RVFLRHLRYLERARAATLECKNQL EHGEDLLTRILRYTATQLSMVARLQVVS ENCTTKEEVRTLASPRVCQLIDHLRTV
KPTGTYRIEAVRQTEGFSEVEDSDNVNDSDDSDTSSRSTRSVRSRGGRRGSRKRSECFVAPHGLEDDDDGSSQQQSD
SDDGFSSTFSPSNTDSDDESASNAGRGSEVEDEMGSESGAESDVTDVSMCLSTGLHRKSKSKSNKRSNKTRNNSLMNRC
QRREKIVYRIVAYQPEGVPVPPPLCGLVLVGGCQFSAYALPRFLEDFCNWDSDIAFVRPGYFFVPKTIKSKSTDLSTAN
TQEETVANFRCGGEVNLVTTQSGLAAASTL PKCNL V A S F H P S N S I E T Y L L A K S R L R L V S I K T T K A R L W H F L E P S K I S M R
NQFQEIEMLIQRCSRSLSLNEHDVDLKFVDEAFDFL PKRNASRAIDVINHYCARLPCNFMTRLAPRWNLISRPLVDVK
SELGDSGKILTCPDTHKDPKVAYRCALQMPNSPINQVWGDWSVCKKMAKYSASVRVGHLLLEAREISNSGRAILRHS
SIIGPSRIIRESRCRSPVKTKSSYL I K T P D A F V R C L P K E D L N S T T S Q N Y L Y L I D V R H L D S V L T H L Q Q R K H K G T S S A S L L
FYPESENRIFAFLTSKRLPYVPDFSI F T R S G E E T V S I Y E A N S C P C L N I N Q L N R L S N F H R I L F K S V L R V E K D T L M T F C L P E
AQQCIL I I P V K L D S Q E I D W D L V D L C L S T F D P A V G D C R M P I R R R P T P P F R I S E E P F L T Q L L S T P D S S Q T V F D F V P E D F A N
AVLTPTYRLVDFPPRYVVTIRNDLSPASRFSPQFSNFADYRYSKYGLHLSGS AKDQPLLDAAAYAPLRLASLLVPRCLA
AAAAAQRKAANS GGGGK KENTNSSRGAPRHVQLLVPEFCYRHALPATVWQKAVCLPSILYRLHLLLAELERSEIARET
MGIVDPPSRACFPALRCHALPKLASLRSNSPANVREESTVHEQDNLEVVKLTNRPSKGDTHKTGHDFDTLLADFSGD
SMDLDDSNIDEIEEEEIDAESVDSFEENVGKLTFFPPQDDDTDL ENRILYRPGPAMILQSLTRLAAGDFINMERLEITIGD
SFLKFAVTTYLYLYKYPDAQEGSLSLFRSHVVENSNL Y R F G C G K K L P G R I V G V A F E P Q E N W L P P C Y V C N S D N S S A D D S S R L
QMVQTHQLSNKSIADCVEALVGCYLTERGERSALKLLQWFGVECLPPSSEMQPTGAPWKMPDERGLTEAEKIEIDRIYH
VSRYDKLENTIGYRFKNRRLIEAL THSTCRDLHSNPSVNNT PSSVFFGNGGYERLEFLGDAVLDTVTRVL F E S T D G E T
VRELSPGGLTDLRSALVNAVFGALAVTHCGLHAYLRGTAPYLTEGSAFLRHIRNVS R G S L N S K R L L T P D S N T V A R L N D
AENQADSGAVRPSNEMAEV PKALGDVFESLAGAVFLDSGLCLNTLWRIFPPLRERIERYSTCVAKSPVRRLLLEQYPER
VKFEKPMVPRDGKIRLVVRVVGIGRYVIGRTYRLAKSAAADLAYERAKEANLNSR

>Hmi c2

Text_s1-1.txt

MTLNEICIELHGNFLQSQCXANLLLTTPMSGDIFTDDSI LSDPFVVDSTAPISPGASCMSDVEEELPPGIVLCESTSLER
RNIPHISDGEVEYQKTPPHSRVNLDDVYDNDIVDEELPMESRSKRKRGPSNKSRYFYQIHNLYSL SOPYQEPIQPGRTCYL
YTWRLPSPENVATCTKVDAACFRHTNQTLLGLFSKPIAKSDFVCRVPLYMQFGIIRVRLRKTGTMTVLNEEQFSLAQDTHS
ILAKFTVSVNEAENLDRFDLSGKSNVTFYSTPSESFSDLGRSPVGNSSASNFKPTDPFSELQIDPLNTRFCLFLIIK
QPEMTFDEAASKDLVTWAKERDYLLQRAAATSGEGSSTFTTLPPPPPPSSRNEPLNLRSGICPRYGLDLSLPIEWR
GRLVVRPINALPPKDPGGYAVSGPNNGLCGSSDIHPEMLKHLPEEMRLRGRTSFFDYQSFKYGRVNMVLETRGLDPSLPL
VTIMRISRHQANAVTAGACATKKDEIKRAEDRLAQLCVVHPLNTWLWLTICLTPTVLHQVYRCLSTMEFSNRLHELIYT
CDGGDNWSSSNPEHLFPEGNFLMPDRESVPPCNTPSFSKICVKEGCGDNPSDWFKESCVPVFTCEEEELDGEVSR
ISQKNADMEVPSFAKLQEAFTSVNAFEAVNLERLELLGDSFLKFVSSLYIFATSPRTTDEGLTFARVAYISNSHLHRV
AVRHGLFRYLATGTFKPEIMYTPPYALVDKQHARENHDQRQFVKIYDKGIADSMESLLGVSLTIAPPRIALLRLYSL
SNEPNPLSQLKVPSPRVQLGDRYPSWVPLILESSYRAIDPEFARDDWESMSKRQVEVDRIILLANNPKTKADAAAASVIGG
TSSNLLNEALDGFGEEMGGDFHTTLENRRSSLAPEEMIGYKFRTRILLQAITHPSEMAFKWGCYQRLEFLGDAI
LDFVVTQIRFRDHPNMPGELTDLKIALVSNINLAVICIRKGIYKFLHEMDNNLWTFINNFKNAALQNFNSIWQLEHECN
KRNESLYKVLGDMVEAIGAIYVDSGGVTIVTSIYNLLGNEIREYSKSPMDPIRLMHKTYPDLKISQVEGDPSPHD
SPSDSEGGQMPAWLNNQNGDQSSRTRVKVIAKCGNCTLTGEGFNLRATAKLEIARQLGIIVC
>Tso11

MELIVTELDDLDSAEARNTIVCTDEDVTLQFLAQLFGDMAVKYPGKCVVIVISDSDSVKSAEETFKYVTDTPVSVVSDA
FTDIPNTAKVVFMSAETAANLCCGEVPVFSIPERAFLLILDNCELVLSPGHPYQRLFPYSKDPLSEHQIEEISWRDLSFY
SSITPSYRILALTSTILPVGSDHLLVKRYICSLERLGRLETTSELINLLSLGAQPEEVEIPLPPAKINGDIDGQHP
IRETVLRLLEEARDLSDVDTGSEENVATTGRVINVPYCLRSVAQCESITKELGVYCGALIARVFLRHLYRSLERARA
SILEARRQLEHGEDLLTRILRYTATQLSMIARLFQTVDCNCTREEIRTLASPRVLQLVDRLRLLKPTGSYRIEAVWTE
GVSEDTGSDSDGGSSGLASHGSWSSSGRRSRLRGRSSSSSFIAPHGFEAMSGDDGEGDDCTEDCEDSSSFGVDSVCS
DGDNDSEACNGESNDKDSVDDNDGGESGSDATDVSMLSSGIRRKSTQKPHRRPRKTRSSPIIRGSRGGSRLVYRVL
PDNPKENTVPPLCGLLVGGCQFSAYALPRFIEEYCNWDPEIAFVRPGYFFVPGASEQLKTLT SADVGSTQEETVANF
RCGGEVNLVTQSGLAAASSLPRCNLIALFHPSTSLATYLSAKSRLRLVAVKSTKAHFHFLGSDGLSLRTRFQGMERM
LVQRCNNSYLIAGEHDVDADYVNEKFTFLPKKNAYRAIDVNHYCARLPCNFMTRLAPRWHLTRQLVGDSEGEALDP
AALLTCPDPSHAAAASNTTVYRCALQMPINSPVRTSEWSLCKKIAYSAAVRMGQLMEAREIDVHGRPILRHTPIAAVA
AASRRLRATRERSPPKHSSYLIKTPDSLTKCLAREERGFLGQNRLLYIDIRHLDLALSRQREQRSQGTSSHIDFFPEA
ENRGFAMLTTPKIPFIPEFSIFTRSGETVHIYEASSCPPLNIDQLNRLSRFHRIIFKSVLRFKDCMLQFNLVDAPCQL
LIVPVKQDSQKIDWDLVNLCLNSYIASVGDCLPVRWRPAPPLHFPIEPVPLTQLLSQTASPRAVFHFPDDFINAVLTP
TYRNVEYPPRYVVAIRNDLSPATAFPSSQYATFADYACKYGLHLSGAPSDQPLLDAAYAPRLAPLLVPRGCVTRKS
NNDGASGNANSSSKHALPRHIQLLVPEFCHRHAMPATVWKAICLPSILYRLQHLLLAELRSEIARDTGLGVVTPPHRF
CFPPLRCYFTPPQSSPSPSPSPSSKKQGDLEDRDSEPEVVKLTLSPPTQPEPRNPINDFQALLQESDIMESDGEGLD
PGVEEEEEVEVDAKSVDSFEENATNLCFPPEDDDDDVEVEGAKGQIVYKPGPTMILQSLTLLGAGDFINMERLETIGDSF
LKFAVTMDLFLRYPDAHEGLSSFRSQIVGNSLYRLGCEKNLAGRIVGAAFEPHENWLPCCYVCASDPTANIASTRVKS
QCPSLTHQFLTNKSIADCVESLVGCYLTERGERSALKLLQWFGIECLPPAGEMRSTGSPWKLPEEKPLSKAEQATIERVY
RAGRFPKLEKIIGYTFKNRKLLEIAFTHSTCRDLHGPPSSDDGGFHGVSVAISGGYERLEFLGDAVLDYAVTRFLFTTDA
AARLSPGALDRLSALVNVVFGALAVRHGALHAFRAAAPPTRTMALFRHHHEVVRGDLVLAPEVNVCRENKGATA
SQPEGDRANEVEVPKALGDVFESLAGAVFLSDCLNLTWRVFFPLLRERIERYSTCVAKSPVRRLLELCPERVKFERP
MVRPDGKIRVVVRVTGIGRYVGVGRTYRLAKSAAADLAYRRVTEAFAEQV
>Tso12

MTTSAKQVSGGIRPNVSDYSSTDDDEL PVKKKKASNRPLNIARFYPIHKLSSL SOPYQQPIGANRPVNLVYWQMLPENVA
YLKVDAAEYFRHTDQITIGLLFSKPIKSDVFCRVPLYMQSGMMRVRLRRRATVVLTEDQLELALRTHQILAQLSMDINKP
VNFDLNFSNNQFSFYSAPFSSFPPTPVSSCVSTSISPFNTDFAIDPFTTELKVDPQNAQICGILLIVRLSDMAFDEEASRA
LVHWAERNFYVSQRERKTQIPPTRNQSLQSGICARYGVQLSTIPTAQWPGLLVRPIDLPPEDPGSYAVVGNSSGLCA
SSIPIDYMAEQLPTELHQRGATSYLDYARFKHGTRIRLVEETRGLDPSTPLVAVTRISRHRNAANVTSGVGVSGKTVIKE
NRIPQLCIVHPLNTWLWLTCLTPTVLHQVYRCLSIGELANRLHEMLYSDDPVQFAHSSSSSSSSFNLPAGSFLMPDRHS
VQVPCIPSIKRYIVRDQGGSWQFKGRVVVFLDSETSLDRTEVDRNTSEVSAEVPPTVGLYEAFTSVNAFEAVNLERLEL
LGDSDLKSTAIALINRYCSKLPSDIFTRLVPHNQUIPEENGVTKYCAELLPLINSPIKHAIVLKNPMPNKKTAQMAVA
LEACRQLHLEGELEDDNLLPKGRESIAKLLHEHIDEEDPEYAPGIAAKVGSKRKQLYDKKIARALNESFVEADKECFIYAF
ELERFREAELTLNPKRRKFEDPFNYEYCFGFLSAKEIPKIPFPVFLRQGNMKVRLIVAPKKTVTAAQLQEIQLFHNYL
FTQVLQMKCTGNLEFDGTSNAPLNTLIVPLNKRKDDMSYINMKYVSEVANMENMPRIKPKDEVRRQYKFAEDYKDAIV
MPWYRNLEQPVFYYVAEILPEWRPSSKFPDTHFETFNEYFIKKYKLEIYDQNSLLDVFDTSTRNLNLLQPRIQNLPPRRSR
TVSNSSSTNIPQASASDSKESNTSVPHSSQRQILVPELMDIHPISATLWNVIAALPSIFYRVNQLLTDLELRETILVKAF
GKEKTLDDNVENSLAYATEYEEKQTIIVKKIQQLRDLNDKSIEDQERETRENDKIDDGEELFNIGVWDPEEAVRIGVE
ISSRDRMDGEDQDTVGLTQGLHDGNI SDEDEL PFMHDTYARLTSNRNGIGAWSGSESIVPSGWDWDGPEPDNSPMP

>Ce1e
MVRVRADLQCFNPRDYQVELLDKATKNTIVQLGTGSGKTFIAVLLKEYGVQLFAPLDQGGKRAFFVVEKVNLEEQAI
HIEVHTSFKVGQVHGQTSSGLWDSKEQCDQFMKRHHVVITAQCLLDLIRHAYLKIEMCVLIFDECHHALGSQHPYRSI
MVDYKLLKKDKPVRVGLTASLIKAKVAPEKLMEQLKKLESAMDSVIETASDLVSLSKYGAKPYEVVICKDFEIGCLG
IPNFDTVIEIFDETVAFVNTTTEFHPDLDLPPRPKIDSLKTTTRAVFRQLGPWAAWRTAQVWEKELGKIIKSQVLDPKTL
RFLNMAKTSMITIKRLLPEPEMKIKSIEALRPYVQVRVIRLFEILETFNPEFQKERMKLEKAHLSAIIIFVDQRYIAYSL
LLMMRHIKSWEPKFKFVNPDYVVGASGRNLASSDSQGLHKRQTEVLRFRHRNEINCLIATSVLEEGVDVKQCNLVKIFDR
PLDMRSYVQSKGRARRAGRSYVITVEEKDTAACSDSLKDFQOIEKILLSRHRTVNNPIEDDSDRFEEFDVDSQMEPYVVE
KTGATLKMSTAIALINRYCSKLPSDIFTRLVPHNQUIPEENGVTKYCAELLPLINSPIKHAIVLKNPMPNKKTAQMAVA
LEACRQLHLEGELEDDNLLPKGRESIAKLLHEHIDEEDPEYAPGIAAKVGSKRKQLYDKKIARALNESFVEADKECFIYAF
ELERFREAELTLNPKRRKFEDPFNYEYCFGFLSAKEIPKIPFPVFLRQGNMKVRLIVAPKKTVTAAQLQEIQLFHNYL
FTQVLQMKCTGNLEFDGTSNAPLNTLIVPLNKRKDDMSYINMKYVSEVANMENMPRIKPKDEVRRQYKFAEDYKDAIV
MPWYRNLEQPVFYYVAEILPEWRPSSKFPDTHFETFNEYFIKKYKLEIYDQNSLLDVFDTSTRNLNLLQPRIQNLPPRRSR
TVSNSSSTNIPQASASDSKESNTSVPHSSQRQILVPELMDIHPISATLWNVIAALPSIFYRVNQLLTDLELRETILVKAF
GKEKTLDDNVENSLAYATEYEEKQTIIVKKIQQLRDLNDKSIEDQERETRENDKIDDGEELFNIGVWDPEEAVRIGVE
ISSRDRMDGEDQDTVGLTQGLHDGNI SDEDEL PFMHDTYARLTSNRNGIGAWSGSESIVPSGWDWDGPEPDNSPMP

Text_s1-1.txt

FQILGGPGGLNVQALMADVGRVFDPPSTASSLSQTVQESTVSPKQLTKEEEQFKLQNDLLKQAKERLEALEMSEDMK
PRREEDTVNLEDYDGDQENQEDENTPTNFKPTIDEEIEELSIGARKKEIDDNAAKTDVLERENCEVLPVAINEKSRFS
FEKESKAINGRILIRQRSEYVSHIDSDIGLVSPCLLLTALTTSNAADGMSLERFETIGDSFLKFATTDYLYHTLLDQHE
GKLSFARSKEVSNCLYRLGKGLGIPQLIVANKFDAHDSWLPCCYIPTCDFKAPNTDDAEKDNIEIERILDGQVIEEKPE
NKTGWDIGGDVSKSTTDGIETITFPKQARVGNDDISPLPYNLLTQQHISDKSIADAVEALIGVHLLTLGPNPTLKVMMNW
GLKVIQKDQKSDVPSPLRFIDTPTNPNASLNFLLNLWQQFQFTQLEEKIGYRFKERAYLVQAFTHASYINNRVTGQYQR
LEFLGDAVLDYMITRYLFEDSRQYSPGVLTDLRSALVNNTIFASLAVKFEFQKHFIAMCPGLYHMIKFKVLLCSERNFDT
NFNAEMYMTTEEEIDEGQEEDIEVPKAMGDI FESVAGAIYLDSEGRNLDTTWQVIFHMMRGTIELCCANPPRSPIRELM
FEQSKVRFKMERILESGKVRVTVEVNNMRFTGMGRNYRIAKATAAKRALKYLHQIEQQRQSPSLTTV

>Hxy]

MRQGMTRVFSFIRTTQKKSFSVDTLRKIYDFHQHVFDLIRVVKSAVEFNPEESNIPLLVPIIKQRNTDGNVTDICILNVE
LLEKEFNKQLGKPSLEERQKMTFEKEYEDAVVSPWYRVNEHPSYYYVARLASESNPGSKFPDEKYPNFNEYKLYDIA
IIHQDQPLLEVDRASERMLIFPRHVSSRARKQAYDPTQQQVLPVELVHVHPLSANYWTMIIALPTIFYRLNQFLLEEL
REILQAFRKPPIPYDQYNWHDNPLEYDCNSIHSMTQLKQRKEIMSTSSVPATELEEYTDSTDESGDFDIGNVDPK
AKERYQVDLILSANGAGSVNERINEEVTVSRKNIDLDGSDAEAEYEIITDFKLNIGIAPQNLDLFTPRSDIVEASDVF
DSVGTVHNGCLSIISEPKTGSSINMEKLNNDLLSVLSSPDPIIVQENQSETGPEKKKAKVSDVVEETSKEDESLPREL
YEKLPVAVPSLPQNFKALNPSRPHGLCVTDSDFDDYINGQWECKAPELSQNFVFEQDFLDQKPEGVPPALMLQAVTTASA
SDGVNLERLETIGDSFLKMAVTIYLFNEHADHHEGRLSFSRSKEVSNANLWLKGYSKGIHSIMETAKFDPNVSWLPPCYR
CEAVFHANGWLKNDENTNEETAWDDNETDQTFKEVDGVTIDLKKITPGPDETFFYNPMTQQCINDKSVADAVEALIGA
HLISLGPRAALQFMWGLKVLVEPPIIPASPIRFGDSEEDPNASKREVVELYKCNQFEKVEKTINYTFANKAFLVQAF
HASYFKGRATGQYQRLLEFLGDAVLDYMITRYLFEHPDKFAPGVLTDLRSALVNNTIFASLAVKYEYFHLKFLNYSPLAKK
IESFVKSCGTNSIFNCPNFNDELFKVTEEEIEDGQEEVEVPKALGDI FESLAGAVYLDENRDLVWVWVYVYVNLMAALIE
QCCRDPPQSPIREL YEMKDVHASFSKLERIEKKNVCRCTVEVGDKLVFHGMGRGYKIAKATAAKRALVYLKMKKAQKEKK

>Ppac

MGCAKTRNDNSSRFGKYMHNFNHGGDPPIGGSINEYLLKESRVVRQAGERNFHI FYQMLAGLPDDRLEKLGTLKSPA
YRYLNQGDSSVPTLNDKNLMEVQALKAHNSFNDDITNQIWSILAGILHLGNLSFIDIDSTEGGSSVKERATLAAA
MLALSEKELAEALCGKTVAAAGDIVSKLHSSSEAEYTRDAMAKSIYERLFSKIVALINDSILIANGHRSASVIGVLDIY
FEIFGTNSFEQLCINYCNEKLQQLFITLVLKQEQEYMREGITWMHIDYFNNQIICDMIESPKKGLAILDEACASIGKV
TDTVFLKELDKAFGGHKHYSSRGVNLSDKTLKHEQFRITHYAGDVTYNVEGFMDKNRDLYLDIKRSLYHSHKGLLSSLF
QDGAKSRAEVGRRPPTAGFLFKQSMSALATQLASKQPFYIRCIPKPNDEKSARSFDKERVDHQVRYLGLLENVRRRAGFA
YRMEYGRFLQRYKLLCKNTWPNPRRGGPKENTQMLGEMGLVKDCEAGTKIFIRSPQTVFAL EEERAKRVVDVVVFLQK
MTRGTFARRRYRRMLAVRRIIGAYRRYKLSYIFSVIDA FRGVRTMPDLGKSIRWPAPPLVNTFVARMRKMHRWRAAV
IVSRIPPHLRECFHPKIAAFEVFNRRTHWGYHRRWMDYLTSPNELPNLSNVSYSGAMSQIKTQHPYSRVLFSSLFHKY
NKFNKSARILVITDKFIAKIDAKMKMLMKDPI SLQKVSRSVSPDHNGLVVHVDDNDLVACL RNEKNEERNRNETTYN
QYSAYSTVYSGKQAEISNIRIGESLGELDSGDIDLGGGGDDERLADSSKRDRVVELEKSGLLELLKEDDALSLMSAGEDDK
HGSGGDRLEGLHLRNTHSLDENLTVSSVLGSSDLLGDELGDRGNGLLGDRLRLVQGGSTLVVELLDKAIKKNITVQL
GTGSGKTFIAVLLIKEFTYVQLRLPLAENGRRAFFVNAVALVDQQAHEVHTALSVGRAHGNLNTDIWTSKAAFDAFFQ
AHHVVMTAQVLYDLLCCGYLKFPSIALMIFDECHHALGGNHPRKIMDLYLSEPKVEQPRILGLTASLNDKTDPDFTLE
KKLYKLERILDCCIETSCDLVSVMKYGARPKHEVIKCKNYRYNPNYSSETLRLLESVKEFCSTTQEFHPELDKDVRS
DALTKTISVIKQLGPWTALKVAGIRETEFKFMKRHTKIFGVKQKDFLALGETTMNTMKFQKIQTLESLPDKVVNLLQAL
DTYHPDKQMARKGKVDHLSGIVFVEQRYIAFALHLVLKHAIKLDPDRFGFINPDYVIGNSSNSLESKASEETHKRLEKVL
LQFRKGQNLNLIATNVLEEGDVVHKCNLVVDFRPTNFRSYIQSRGRARKSGASYLVVVEKDAKNCADDLADFVAKIEKL
LLNRAKTVHNPSEIGIELTIDDLIPPYFVEKSGAKLFLSNAIQILNKYCGKLPDSVDFTRLVQDCRLLPSERNMGHWYAE
LLLPMNSPLKTAIYLDENEAVSNKLAQMLVAFKACQELHIGELNDHLLVPGKVEAML SLLDDDPDENTDSYAVRVGSM
KRKQLYDKKTASVLSALPREDEELLYVLDIDLIVPLPPDEIRSKKREILNPKTAEMCFGLSRTKIPKIPPPAFLRQ
GESRVNIVLSPKTLTLDDETLTYLKL FHHYIFRSVLQLTKDCLFRLDEATPLQTIIVPLDRITGPNNGSEYSLKMEYVR
EYVHNMEEMPRVPKEDDRKHFKFVASEWEDRVVMPWYRVNQPTFFVYIDIHSSQSPSPFPDDDFNSFNEYFKKKYNLE
IYDQQSLLVEYTSNRNLNLLPRHRPSRQRTTKSGGSPNQKQIFVPELDRHPISARLWHSISAI PAFFYRMLNLVAD
EYRARI MREAIGLSEDEATPDEMMWKPLAYPTTAEERSPTMVKLNQLRVDTKEEYIPSKKRKGGSSIASGHEKQFEIG
VWDPALAAEFVPPDAAIAIKVDGWLEETAGLHGTEIPQKGGDLSDDEENYEHALYFDRSNFIRDRHIGMPSDDVQELGWGE
LPSEIVVAPFMILDGGEQSINLNLGMADIEKVL PSTSPSSPKEEVTTTAPGDTKYIPKQLELWDPVPEVDDREKGR
KEEVIDLIAFDNDVEPFDKSKASPTAREEKKEGTIDDELSQLERGEEEKARRDHEAAMIVSEKEIKEGEVLASPLADPRF
SFHHPSRSVAMGGYFIGREEGAVGEEKAPFGVPSLILAALTTSNANDGMNLERLETIGDSFLKYTVTDFL FSHSPLDHE
GKLSFARSKEVSNCLYRLGKIGLPSIIIGSKFDVQDGLWPPCYVPSINFKAPNDDAEKKDEIMERMNGEEMIKENI
PVTGWERPDSGIEDGVETINLTKPAPDALDDL SPLPYNLLTQYISDKCIADSVEALIGAHLLSLGPQRTLQMMKWLGL
KVMTEKIPIAPPLRYIDTPLYPHRSEEMLEFFAKFLAQLLEEKIGYRFNKNAYLLQAFTHASYPNRITGQYKQYSP
GVLTDLRSALVNNTIFASLAVKYDFHKHFVAMCPGLHFMIEKFKVLLCRERNFLDANFHCEMYMTTEDEIDEGQEEDVEV
PKAMGDI FESVAGAIYLDGCGDLVWVWVFFNLMRETIEECCRNPPRSVRELLEREHDKAKFSKLERIMEKGVVVTVE
IVSGKENFRFTGMGRNYRIAKTTAAKRALKHLKGIDEDRRKAERKKQF

>Srat

MTNRSFEVSDEVFTPRPYQVTFLEKLEI LEKAKKRNIVQLPTGSGKTYIGILMIKEVQYTVRKSISEGGKRIFVNNVC
LVEQQARHIKNECEL VVGELHGESSITVHDTDKIKTFLEQHQVIVLTAQILLDLICFSRISMEDISLLIFDECHHSMGKA
HPYSNLSRYDKVPNEKPVILGLTASLFNQRLKKNQIENLLKLEYKMHAEVVTADDFSQVLKYSIAIPNIKLEICSDFN
AEKFDIAQKTL SFVDKLNELVREMEKLLKDSRDFDPSQFIKTDINSTIANSIYEKIDTSSKDIRKKISEFVYITNLLGP
FTIYKICDFWTGDLKKNYKNDQTSVKNANYVVEAFDVFDSIKKEYTKKFSTMTVFEEIEPYLSNKLKQVIEIETFNKCC
IEDPSVDFQSMCCIIFVERRSAAYAI FEVLKELKLLDKARFGFIKADYIVGVNNKSVEETTISKIKQDKKLQSRFQSGSLN
VLVSTVLEEGCFIDQCNLIRYKFPKNFRSYVQGRGRVRHKQGLYILLSDSKSEDDKAELKNFEECETLLIQRFRTPN
NFVDNKKSNPCPDFYVDSL YDSYVVESTGATAVL SNAITIIINRYCQKLPCDNFTVLSPTPDIENEDGTYECSLLL PAN

Text_s1-1.txt

PYRNIIRSKKPLPSAKLSKMAAAVEACKILHQIKELNDNLLPNGREKLLIETKFDDYEDEVDIKEVIVGSAQRKQLYDK
KSSIYLYDILPLENEESYLYIFDIDLKQPIHSYMLNKRKIDNPQDYECAFGFLSKKVIPDMPSFPIFLRYGCAEVTVKK
HPKKYLISQENWDKAIKFHEYIFSDILRVTDVATFSAQAIVKIVTLPKKKIKLESDNEFDYEIDFNMYDYILETLEGMF
KPPLEEDRKKYVFNESLYENAIVYPWYRSEENRNYIIGEIMHDVDPSSDFPDNNFKTFEEYFLEKYKITIYNKDQPLLD
VDYTSNRLNLLVPKYPSKSSRPLKNSRSSQRQILVPELVSIIHPIKSTHWSIIGALPSIFYRYNCLLMANELRETIVKNAL
GKTNVYRGPFEPLSYKTCLEKDLKDEKNLNEIEEETDQVEEDITDFEIGVWDPQYIDTFTDLKPVKEIPDQLNNDVQCL
KASYPDDEKMSDIDDIDEEKMGTFKFLSVDERPEIKPVTESSTLQKLESLSGWDDPCIDNQEIKSVLHISDVASIDK
KALVRDMNDVNELEETEYLDEKFSTYPIKKKHLVDIEKKFVDLREYEQSESPPPIITKDDFNIHMKLDEFEDLGNWKTIV
RDNYELLIEQVPMYKLDDEGIDTDDLYSKGVNPAILLCALTSRGANDGIDLERLETIGDSFLKFATTDYLYHKHVNAHE
GCLSLRSKEVSNMPLYFLGKKRNIHNIINDKFEPHTNWIPPGYVVSSEFKPNYVQONADKEDKDAEAILNCFDESDDK
LSELCEMEEGKNSKKSSEFKFKKVRQFNELSNIVYNPYLQQQICDKTIADVVEALIGCHLIHIGFDGTLKFMQWGMIS
VFSHHYEAVIDPMIKILDFFPQNPNSMQCLYGFYKKNLFDIFEQKIGYIFKNKAYLIQAFSHASYINRVGTGYQRLEF
LGDVLDVFITRYLFSHEASFNPGTLTDLRSALVNNTMLASVAVDHDHFKYLMFSPKLEHVVRKFTVFIKSHENELINL
DSDLFMVNEDEEDCDGTEDVEPKALGDI FESVGAVALDSGRNLNVWALYKMMKPHLERVYKTEPPISPIRELTESYP
DKVQFSKVERDPNTQRVKVYVEVSGTKITGGGRNFKIAKNAAKRALKYIKQLEKERAKNQNL

>Bma]

MVVRTSDVDKNFFTDRDYQVELLDKACKRNIIVPLGTGSGKTFIAVLLIKEYTTKLVTWPKNGGKRAFFLVDKVSLVEQQ
AAHIEHHTTLNVGKMHGHLNQDIWSEPAKFDTFIALHEVTVLTAQIFLDLLDHGFFNMSNAAVIIFDECHHVLGSKHPYR
LIMHRYGQLTEVDRPRIIGLTASLISKIPPSNLEHLEKLERIMHSSIETASDLVSISKYGARPKYVIMCHDFECCCTC
ETSKYVISTLESRTFCLKTEFHPEFDVDPKVPLEAVSRTKSVLEQLGPWCWAKLCLFQRQLKKGQSGGFLPEKQII
FLQMAVTTMRFIKRLLDVKVIRCFSDVKPILPDRLARLFEILKFFSPSNMEKVPDFTFCGIIIFVEQRYVAVVNTLI
RAISRWDNDKFGYLVSDVFIGYNSANIGTEETMALHQRQELVLRKFRQRHLNLLIATSVLEEGVDVRCNVVIRFDRPTD
YRAYVQSKGRARKDGASYFLLEVERDREQCSCDLKDFLQIERMLLKRYQNVHNPPEPMISPNLETVDIIAPYTVESTGA
QVTLTTAISLVNRYCAKLPDIFTRLPQNTIVPETIGDRVMYRAELLLPINSPIKETIKLKKPLESKKLAQMAVALEAC
RRLHRRKELNDYLLPVGKDTIMLTALDEDPDEFIPNMSYKVGSAARRRQLYDKRMALHNAI PRAGEECYIYVMMEMDLIK
AVTGAANPKNRRRIINPLDTEFCFGFLSNKKIPKVPSPFLFLRQGRMQANIFLVKSRLLDVTQMLELLKAFHHYLFDNVLR
LVKGGFLVFPDKAPVNLIVPLRRERNSETSEVDFKLDYAVRVNVSSIDELPRIPTAERLAFKFAAKFQDAIVMPWY
RDRDHPSFYVAEIIDAKPSSKFPDDKFTVFNDFYFIQYNIIIYDQQQPLLDVDTSSRLNLLMPRHWSRSKSRVTEES
SESGGISQGGILVPELVDVHPAAASLWNVIAALPTLLYRINSLLLADELRELVREAFSNPNYNTSDDVYWLPLDYPTPM
DDLEMKSQKICDLKKKHVEQKSQENKEMEAKDSGAEMTDFEIGVWDPELAKGLDDFCLNRERHDEIKGFEDIDRDALGL
MNGSALRQHGDMSDDDEDDAVLDFINSVHERLKGESGDI FAPRENITSSGWDDLIVIEESAPNGINMPLSVNSGDSHI
DSRGLMADLSRMSWLLDLPTVL PSTAVDTANQNDFGIKKETVLLSNKRHPSVARKPAQLYLDLSLERLEDSDRGSSKSNRQ
EECIDLIDFCDEVEEILSNVYDGTAPFDLRYFNGCLLDTDELDTAEVLSPKKISMRQDRKLINEEMVRSSSDLELFLQI
SHDNDVIKLVSSTTDTNAVDRKQEISAPVLDWMTFSFEEDTFTDHPDGVSPCTLLQALTLNASDGINLERLETVGDSTL
KYAVTDYLFHTNPEQHEGKLSFARSKEVSNCLYRLGRKHNLPSLIIGSKFDPNDGWLPPCYAPTSDFKAPNTLDAEERD
KFIENVLEGGKAVEGQETVKIPTGWDEADRNGQVRRIANGIETIEFPKNMSTTSWDGEEITPLPYNLLTQSLGDKSIADAV
ESLIGAHLELGPATLKFMKWLGKVLTEPVMEPLLRFIDTDDQPKSLRKLNDLWIQFQFSKLEDCIGYRFHDRAY
LLQAFTHYKLYNNRDSGIYQRIIGLGDVAVLDYVITRFLFQSAHYSPPGVLTLRSALVNNTIFASLAVKYNFHKHFIAMC
PRLHMHIEKFVCLCAEKNLSSANFNEEMYMTTEEEIDEGEEEDIEVPKAMGDI FESVAGAIYLDGRSLNTVWRVFNL
MKETINECCSNPPRSPIRELLMEMEPERARFSKLERILETGKVRVTVDIQGGKCRFTGMGRSRYIAKCTAAKRALRYLRSLK
KEKERVAGKE

>Tspia

MGMAASSTNGDAMEDESEMEAFDVRIADDVIQRLRFGSQDIFKSEDPVIDDMLIESAYRIQYETLRQDLIATQEEVARL
KRQLQERDYAERKKGRSGTVEVGTKEKHCSRRENQKLSNELVKLLLEIFMFTVQYSCLSNSATFRKNNIDMVYGLN
LVRASKDCVTELERHVSQVLYDLKHSQDLQYYVKEDLTFVSAKEIDVGNRAVIVFVPPFPQYKQYRRILVRLIHEVEKKF
GGKHVVFAKRRILPKPTRGKRKITQKQKRPRSRTLAAVHDEYLNIDVFAEIVGKRIRVRLDGSRLKVLHLDNRNQA
EKKVNRMEKEQSGRLENVRLYQKELLECAKKENVIVTLGTGTGKTFIAVMLIREMSESVHQLKEGGKRSFIVDKVPLV
KQAEHIRINTNLKVGEFHGALGVDCWSAKIWMFFDKFHVLMVTAEIFRNILTHGFIKFDIVNLIVFDECHHATKQHPY
KKIMELYKLYNNRDSGIYQRIIGLGDVAVLDYVITRFLFQSAHYSPPGVLTLRSALVNNTIFASLAVKYNFHKHFIAMC
NKYDWHELDFFQHLLVRSSTQKSLERCGRCEGNVSLYIQHVCDVLAKLDSIMGQLGPWCALQVCKQMLQSTTNLATRIST
LLADDEDNNL FSLFTSGSDASSVWLEVLKTCLKVTSVKKRHHIGLVDVSVQKLCVFIHQVFDQVQVAVWNCIRQSTPG
SLLYSGELNFQLEKNNSRK FENTNGANGPPFQILIDTAKKLQPRSSYLEVDYVVGSRMSAEVAEPSFARQEEVLKNFRH
GKLNLLAATSILEEGIDVRHCYVIRFDTPLTFRSFVQSKGRARQKIAYTILVQDRFLESFQEMLSNFVETEKFLKVN
RDLNILLTRTKDRDRSYSGNNDQVDCRKEVNL PNGDVDSLVEPYTYFEENGLVKKAACLVLSAHNVIKRYCNKLRKDR
FADSSPKFSVQITLNDLSISYVATVQLPTTSPKKEIEGKPMQNAKLAEMAAAFETAKMLHAMGELNEFLIPPVTVRKE
IELKEDEQFFSISRSYLLKEVPDLYKAI PRADQKSYLYAISISSIESFLPRLGILVSKPIGNLPGFSVFTDESIDVE
IEFVEETSYSAHQLDILTSFHCCI FKNYL FREEEDFI FDPENAACSYLIVPLKCEYNSKCSVDLSTAEKIINWSKDVSHI
PQQRSEKFINNSSSYFNAVVPWYKKNDRDFYFVTAVDNTCTPMSFPFLKMYRSFAEYFEVTKQVQVLDKNQPLL SVK
MVSFKRLNLLCEKPISILLDFDADDANSCHCKLVPELVIIHPMPASMWRCILFPTVLYRMNHLLIAEQLRLQILREAMF
PGEVLEENCEIQLNKDWYLLASRLNDMCIVDNQLPVKLSL SVNSNSAAAATATGSIDPETQALL EEFDVRFCNACPTV
PVDDQLDILTGNSVAESHTAESNDECLFDTHFQLLSSIQPPQPLEKVDGFIVESDGESDFVAAPALPESELNLFVE
PSDVSDTATAPAQRHPGQCRDVLRALTLRKAQDMFDLESMEALGDSFLKFI VLSLHVFIKETNWNNEGRLLTSRSEIVSNT
NL FNLGKQKLLQAKLTAVPFDPTAQWLPPCFRSLAALESGYESVNLIDEGDSKKNKNEALKANTPTPVVAERYEMCSLN
RTHQVIYDKSIADCVEALVGCYLL EAGMRPAIKLLKWFIDIDGNLMNLLSSSSSSSSSSSTAFCSENCVLIGPEAKIHS
IWTAYDLNSFEAKIGYRFTNKAYLIQAL THSSYNEVETPVTD SYERLEFLGDAILDYLSRHLYSSKIRSPGLLSDLRA
ALVNNYFLYFNAELLVNERFVLA MRGLKESVNFHNL YMMEEEQDDDEKETSEFAEHVEVPKPLGDI FESVAGAIFLDS
HCSLATVWQVYYNMI AEEIGKQRSDFTSFNKCLCHPPI SPVRHLLLEPERVQFNILDREEGNSGVHVQVVVTGKGSFIG
SGKSYRTAKHAAAKAVRELYPTS NHFII

>TspiB

MNRLYDFETDFFTPRDHLTRLVDIGRRCNLIAPLGSADVRLYVSVMLIREFSSMLRKPCSKNRRWCVYLVDRATSIKVA
DRIRIYTNLNVGEIFCDIHLESEAKQYCAAQMEGNEIFVTTGTHFMKLLDLSFIPRNCYCLVIFEDCHLAIRSHPYRNIV
KEFLNIEKEYRPRLLGLTMSLINDEVKGDCLQYGIDTMQEVLC SKVVLTMLSPA KFKLKRDTKVVLLISYPSTSLFSWDMF
DNYILHLLVSPTFILENITL PNDDDMICM QHVIESLNKVKWIHDEMGPWCAWKVCQKQEMQLNRMKRIKPGTMQYLLIEM
GQTYLRCLRKVFENHVKNLKNFSSLCAYITGHVYSFIELLSHNKMRSVFADDENFPENFCCIVFFKHRYIAYVYKILLKT
LQNLWPFDMFGYLVDFLVGYDSEADASKEALHERQHEVLKFKRTKELNLLTTRVLAKGIELRGANCVIHYDEPELSRS
FIYVKNRANKPN SHYFILLSECTSAL SASVINTFVDIDKIIRNYSLLDYEDDNLLEPDDLDDTFKPYYPKKSEEVGDDDE
EADKLKNNNVCATLANSIRIVNHYCQRLPCDI FSRLVADCHVEKVMMSGECAPSPKYRATLKL PINS PGPTVDSV MLARRS
VALETVKLLHRRGELYDSLMPIGKEMVASLLIAD EDD EEW PVTGKACPGSSKRRQY YNKAVRAGACKSLIECLKRAIPVF
DEPGYFHAI IIREVSEICNGEANISQPEVGKRILGIFSSKPIQIPAFMIYEKTKSFLVEIRPCASTITV NKGQMAVIAE
FNMYIFSEILKLEKYSMKYQPEEAENAFFIVPSISTETDYILDWQFIVEVVNYWDRVPRRPDEEQRKNFVFDISNLVKIG
VFVFNFRYKDAVMPWYRSQEMSHCYFVLNVDEGLTPLTKFPDDEYESFKSYFWQKYGLEIYHDDQPLLNVDYTASRLN
RFPRRQARRESLEADYDTMSVVAQKQKLVPELVDIHPI PASTWRCLQYLPSILYRLSSLLLANELRLQVLLDEYNNDP
PDGYCWAPIDKAWLKETMRMTDHGCPWAKMVKNSKNSGINCSSNSIDSGKNKNNNNNNNSNLSNSNIKSEDSARFEQM
DFEISVWEPVQEFVFNFGFTDDGKHDCLDANSADFGSDQLPMSDGQSFEDCLMQMGDYDGSMLDSDEEYFDTEFSRKF
NMVSLHGSYGQHPFGLFGQPIEPSGWEVDDVPTMPAADEHCDVGLHFVSAGQSLNVSTLLADVRDADEKVRGLQQQP
AKEEDTVPTS NATAQKVPVDRNGTEAKITSRHKI QAVDTLAE EETVDLNMFKRAVYNRKG NVMNGNESWKNFTNEQMLT
LDSSWYEWGPEFR TAALLPAGETCTFLEIANEETPFPGNPREV LHAITASAVAETFNLEGLEILGDSFLKYVTTVYCYKA
YSKMHEGKLSLLSRMISNYNLYKLGKRKNI PQYMI AIKFDP SVDTLWPCYVPLNCEVQDTAIEEEDKLMQEVVIRFDRPTD
VEQQRNVETTGRQATTGSKSTKWVPEDLSQLVPFNLLAQGGISDKGVADCVEALIGAYLLFCGTRRTLDLFLHWLGLKVED
EVSLRGFSRCIVSPTPISFANFDLYGVTKSALINNSADSEACLKTLWNRFSLSQFEDIIGYRFKDRSLLVQALHTSTYFY
NEVTDCYQRL EFLGDAVLDHLITRHL YEDKRMHSPGMLTDLRSALYFMYLCPGLMMMI EKFKVKT LNI IKENANFDREV
HLLNLYLLEGEENAEQVEVPKALGDI FESVAGAIYLD SGCSLHTVWCIY YNMLREEIEKCCLNPPISPIRDLLELEPD
RVKFSRVERNAVEGKVKVAVTVEGKGRFV GAGRSYRIAKSTA AKRALRYLKGSPQMSLISKIH

>L1oa

MVVRASDV DKNFFTPRDYQVELLDKACKRNIIVPLGTGSGKTFIAVLLIKEYTSKLVIPWKS GGGKRAFFLV DDKVSLVGGQ
AAHIEHHTTLNVGKMHGHLNQDIWSEPAKFADAFIALHEVTVLT AQIFLDL LDHGFFNMSNAAVII FDECHHVLGSKHPYR
LIMHRYGQLTEVDRP RILGLTASLISSKIPPNSLEQLLEKLERIMYSSIETASDLLSISKYGA KPKEYVIMCHDFFCCSC
EISKKVVDLTENLRTFCLKCTEFHPEFDVDRPKPVL EAVNRTKSVLEQLGPWCAWKVCQLFQRQLKKHSGQGFLPEKQVI
FLQ MAYTTMRFTKRLLDAKVASVRCFNDMKPILPDRLSRLFEILKFYSPSNMEKVDPDFTCGIVFVEQRYVAYVNLTLI
RAISRWDSDKFGYLVSDFIIGYNSANISAEETMALHKRQELVLRKFRQRHLNLLIATSVLEEGVDVLRQNVVIRFDRPTD
YRAYVQSKGRARKDGASYFLLVEERDREQCSRDLKDFLQIERMLLKRYQNVHNPPEPTISPSIEAMDDIIAPYTVESTGA
QVTLTSAIGLVNRYCAKLPSDIFTRLVPQNTIVPETVGDHVMYRAELLLPINSPIKETIKLKKPLGSKKLAQMAVALEAC
RRLHKRKLNDYLLPVGKDTIMLTALDEDPDEFIPNMNYKVG SARRRQLYDKRMALHNAIPHVDEECYIYVMEMDLVK
AVTGAANPKNRRIIINPLDTEFCFGFLSNKKIPKVPSPFLFLRQGRMQANI FPKVSRLLVDAQMLELLKAFHHYLF DNVL
LVKGLVFPVDPKAPVNLVPLRRERNKNTNEVDFKLDYAYVRNVSSIDELPRIPTEVERRAFKFDAANFQDAVMPWY
RDRDHPSFYVAEIIDAKPSSKFPDDKFATFNDFYFIQKYGIVIYDQEQLLDVDYTSRNLNLLMPRHWSRNRNRTTEERG
SESGGMSQGQILVPELVDPVHPIAASLWNVIAALPTLLYRINSLLLLADELRELIMREAFSNPDYRTSDDVYWMPLDYPTPM
DDEMKPVQRICDLKKQVEEKSL ENNDAKDSGSGMADFEIGVWDPELAKGLEGFCPNRGRRDEIRGFEDIDGDALGLMVN
GSTLRQHGDMSDDDDEDAIVLFD FINSVHERLKGSGDIFAPRENITSSGWDDLIVIEESAPNGINMPLSVNSGDSQIDSR
GLMADLSRMSWLLNMP TALPLAAVDTTNQSVDVKKIEIMKNGKKQHHP SIARKPAQLYLD SERLEEDSRGSLKSNRQEE
CIDLIEFCDEIDEIVSSEYDSTL PFDLRYFGKCLLDTDELNTVEMLSPTKILHQQDGKLMDEEMVRSSDGLLEFLQMSH
NSDVAKLACSTANAYEVNRKQEISAPVLDWMTFSFEEDTFNDHPDGVSPTLLQALTLNASDGINLERLETVGDSFLKY
AVTDYL FHTNPEQHEGKLSFARSKEVSNCNL YRLGRKHNLSPLIIGSKFDPNDGWLPPCYAPTSDFKAPNTLDAERDKF
IENVLEGKTV ERQEPVRVPTGWDEADRNSQVRR IADGIETIEFPKNMITSWDGEEITPLPYNLLTQQSLGDKSIADAVES
LIGAHLLELGPATLRFMKWLGKVLTEPVQVEPPLRFRIDTTEQPDKSLRKLNDLWVQFQFSKLEDFIGYRFNDRAYLL
QAFTHASYKNRITGCYQRL EFLGDAVLDYVITRFLFQHSSHSYSPGVL TDLRSALVNNTIFASLAIKHNFKHFIAMCPR
LHHMIEKFNCLTEKLNSSANFEEMYMTTEIEEIDEEYIEV PKAMGDI FESVAGAIYLD SGSRSLDIVRWFY NLMK
ETIDECCSNPPRSPIRELLEMEPERARFSKLERILETGKVRVTVDI IQGKCRFTGMGRSYRIAKCTAAKRALRYLRSLKKE
KERAAGKQ

>Dpu11

MAPMTRRSSSAGRPNYPAGKQQT LHEVELFESAKVKN TIACLKADLGMFIAVMLIREKAFEVCNSLLL PKGGFCSVYL
TPTIKLALQRSAYLRHHTPLKVGIFHEKKGVLQWSSERWTEELGKSQVLVMTYSIFLEVIHKNFLKMNKLNLLILDCHN
VMNNEALKGVLTLYENAASSVHPRI LGLTASIVNNICKPMELSQLITILETKLKS VISTSN NVLSALQHSTRPKELICEI
SSKYEATELQQEIMGRILLNDYL RDHRYDPAEY YGEQDYLKEIPNCPWGADRAALVYL IHLEK LKTKAVYERHYLLC
LVFTVLL EIRSLCEEEFQKYSEYERIVKFSHPKVLRLDLILRIYRPP IPSVLPTS AELVEVSDSLVSSEGQTVQTPKLAT
ETSQTAE LIENGPAILRQSTTQPGHTSRRRRPDFSNRHRNREDAYQLCGMI FTERRTTAKLLYHLLK DASRCDPQLAYL
SPLYTVDRMEGVSVKEMENEERKQEDVMKFRFRSRDCN ILVATSVLEDGIDVPACHLVIRYDL PQSYRAYVHSKARARAKK
AHYILMVENERMDFLVDLSHFHATEQ ILLSKSGYHNSVAEFNHNE DMLNCIQAPYPSDETSNSYASLQNAISIINRYC
AKLPDSCFTRLTPVWRIRTIKRELIEPSVLDKWIGL FKNSGKAVENFLDVAYICSLQLPINSPLRKT VLSQPMP SKVTAK
RSAALEACRL LHQKKELDH FYPTGKENLRLEEEEDYNADLEENVPENL PRPGTTKRRQY YKQVADPFINGIPRPDQP
CYLYAITMVLSEPIPEEQNTRGRKIYKPETS IQSLGILNSKPLSQICFPPIFTRSGEVEVEIVKISDAIRVTEGQLKQIS
DFQRYIFRNVLSLEKDPMLLDVYNADCSYLIVPLRRSSFSEHSLSLDFV FIDLISSVRDSKPQLIAEEVRS GYVDFPRL
FSDAVVMKWRDQEHQCFYVAEIFYELNPQSKFPDNNYPTFEKYRYRHKYGIQIQNLKQPLLVDVHTSARLNFLTPRYVN
RKGVMPLTSEHTKTKTQRESLQQKQILVPELCSIH FYSASLWRQAVLPCILYRLNGLLLAEQLRCKVATEISLGRSQLE
PGFCW PALRFDWKCSDDPALPPLQRENRTETADEYDNDLQAVALLT NDSGLDFDWSKPVVDITGDGSLPAFGVPT
PQSYLEPWGESKTKQFSQYRPQSPSELISNFS DGSYNSGQSANDSDTDDVSE EDES GKVKVEFGPKFVCRGD TAAEIEEP

Text_s1-1.txt

INLKARKFKSPEEEEELEELTEDDEPYENIEDLLSMKMQNVNDIPNRDLEVKCENLLILHKALIVDQSVVPMKDIPIMT
PFTHRITENPENSTLVPEECYKPLDPSGPAAKPNRRLASESSHMLELTRLLL E EYTSECSNDSESPDPPIRETPIPKK
DDTGLVITELTVPGDFPFSFDEQQGLESHPGSPSVILQAFMSSANDGINLERLETIGDSFLKYAITTYLYCKYSSIH
EGELSHLRSRQVSNLHLYQLGKKKLFGGCMVATKFNPHENWLPPGYVIPDAL E EALISSGIPLNYWVVDLQHSIPDKSI
ADCVEALIGAYLRACGPRGALLFMSWLGLKVLVINEGSGYGYWTPPISPLVDQPLDIHVELDRMLAGFESFEKICYKFN
DRSYLLQAFSHASYLNRLLTDCYQRLEFLGDAVLDYLITRYLYENLPRHPPGALDRLRSALVNNNTTFAVLAERYEFHRYF
KHLSPSLNQIMDKFIKAQEENGHSINEEYMI EAEDVEVPKVLGDVFESVAGAIYLD SHMSLNAVWRVYNNMMKKEIDF
SKNVPKSPIRELELETRDVKFGKAERLPDGRLRVTVDVIGKGTFKGIGRNYRIAKCTAAKYALRTLEKRKMLTKRGEQ
>Dpu1B

MQAGGGFSNFQMTSTVGLLIGMATAGVLVTVGLVVTWRFRLLCKKKPAANQNGQSQSPSDVAVNDDNKDRDRNSYSGQREY
IELGTKDPDVI AKEAQGIFIKSDMLEL YE HARRENTIVYLPTGSGKTMIAVLLIRDYAAQVNKPLAQDQGRVFLVPTVI
LAKQQAAYIRRHTCLEVGEY YGELGVDLWKGDRAWAKELEKNNVLMVTAQLFVNAVNHGFI SVGQMA LLVIDECHRAVKDS
PMKHSILVLFNEFCQRHGAGSVPRVLGLTAALFLEKCKPKDVTKNIRKICEFIFGQLSPIERLRLFTTPHVTRLMDILRQF
SPGEIDTTAKFTPLCSIVFVERSSANVLYHILREASQCDPLSYLNPLFTMGQASGNPKGKMKENHHLNQSQSKIIRQ
FRDGT CNLLVATSVLEEGVDVRACNLVIRFDGIKTCDYVQSKGRARSNPAFYILMVPTKLSPLDFTLAGHFHIEQSLM
DPSSVMMSSVTDDEPSDPQLEQLWQKIIPPYCPYDGPVRTLLSSISLLNWYCATLPSDGSAPLAPYVVNRISAGSLAE
MDRDHTNHLRRLGSTIQNPFRTLLLPLNCLRDELIGRVMP SKRLAKQEVSLQACIRLHQLGELDDQHLMPLSKSVVET
DENADVDFDCQFEHSQGSDFYLRGVARLLDGVIASPFYLYLLTYEVVKDSAPENQFDPNRAQRMMGLSSQKLPVAVNPF
GLYSPAGQINVGLFVDSFLRLTLDDVKLCHRFFQQLFEKILQLPGEEQYNPSRGSYLVVPTNNRIDVAFMNEHLDSTV
PDWQPHPPSSLSSSQVVDYFKVYGDVAVTAHHDQTKGKSSCYVNAVREVTPLSRFPNPAYATYANYFRDRHNAIIANL
SQPLLEVRKESFKTFNFLT PKYVNRKEAADEGDGEIQSLPKGSIYFVPEFCEVHRFPASMWRQAVWIPSI FYRLNSLLLA
DEL RMEIQAKTGLGLSVLPPSMCNTK FYFGQYTPILRSHSLFRSSMGTSTDENQWQMPAATVTEVINQINSPVKSAGLKF
NKKRLSTRNSSAAFVNFRRDGDPMNCPKGPSPGEVLRALTTSVANAGYDLERLEVIGDSFLKLAASIRVYCDSPSHFH
EGKMLHRLMLQVCNRNLFKLGKIKSIPRFVATKFAKENWLPCCYFPNNTDDDCMELETDEKPNAKLWQRISKKSVA DS
VEALFGLYTLHGIGKALKVMRWGMVNPDSGDEDS SSSSSSSFGGTPPEPLLLDSPAQHLLDDHLAGFETLERSLNFR
FGNRFYLLQAFSHASYHNRMTSCYQRLEFLGDAVFDYLITKYLDCRLFQLTPGALSMDRSALVNNVTFAVLAVRNGFH
RYLKHLPVDVHQ AIDRFVQQQEMCDHATPDEFFLIGTHDDPMAEAVEVPKVLGDIFESLVGAVFLDSDMSLDAVWRVFYP
LIRQIEACTRVVPKSPIRTIYEKYPGKVQFKTLENLPDGRICIAVEVLNNGRLYKGVGKNSRIAKNTAAKYALSQLGI
>Dpu1C

MSKVRVVKSLNLTKNPIMESPSKSCNVDT EFTPRDYQIELFQAALDENIIVYLPTGSGKTFIAALLIKEKSHEVTKPLDS
GGKRTVFLVPTVVLAIQQAAYLRRHTYLKVKEFYGSMGVDIWEKDRWNIEFDSNHVLMVTAQIFVDILNHAFFSPHQLNL
LVFDECHA AVKDAPMKQVLSKLNCEVSRPKILGLTAALFRMRCKPEKVPDLIQLSETMGCVVRI PSDVETVARSSTK
PCETILQYEDKNSNPNDADELSRIIFVMVSSLMQQL EQQYGE GNSKDGPDVVKQPKIDVLF DALNDIDYLLSTIGPYG
AFHCAHVYIRILERTK SRAKTAPFL TDLIRVSALCQEVREVCTPYFAKYSVPDRLLL FVQPRDI AKHDTTLEFLKPLFTM
GKAAGRNSSLKETQILNMKQNEIMQYFREGLCNLLVATSVLEEGIDIPDCNL IIRFDRIKTYCDYVQTKGRARSKKAFYC
ILVSRGETESCLADLAQFHSIEQQLSTGNFEDERD NAIDNLSQIIPPYVPFGEDGPRITL PSSISLINRYCGQLSSDP
DVSLAPQLTTKKAESDVPDIQLQKIKDILKLPDASKYSSINLLFQCHLFLPLNSPLRDEVIDGMPTKRLAKRAVAL
KASIKLHELKELDDNHLFPVPRLNVIDDESSTEDEENSAAPQLNKENIAVYNNRRLPVCFSNCRPLPGQPCFVY AIDFTLT
KPCLDITKLYPFAVDTKLA ILTSKVIPAICFPFVTRAGEFQVNL CGADSVILDQSQLGKLERFHQFVFDVFLWKRQ
LDFDVLASDLQYL VVPLQTANDKIDFALVENMINAPAINWEKPPYAGDHFIFDSSRLIDS VIVPSYKPLGTLNAFYVDRV
SDLTPLTPFPNNEFKTYASYFQLKYKLTLTNQQQLLQVSREISGKNFLVNRVASTREETPMHLVAELCHVHPLAGSIWK
QVWVLP CILHRLDRMLVAEELRVLVYKEANIGIEFLPGSVNVSSVWCLRLRPQK SRAESRCVADPTLLASTLKP SAFMD
YLDL DVNMMWSEFYEDESKLSPNQSSIEQTAEITFLLSPKPSLPLPNSPGYSAFKEFPDSIDEQRNKWTDLANGN
PDDLPTCPPETFNLKFDGSGNSGMLSNFGSPGMILEALTLAKTHEGFLMERSETIGDSILKLVISIVYVGETGSDRCDE
GRLSLMRMRQINNKHLLFKLGAKD IGEFTVAQR FELMANFLPPGFKTPTDPTDINAHVQYVLMKNVADCM EALIGVYL
TTTG IKGAIKLMDWMLKTVPRLEIATFNEMNGFPILPSTFGSSTLNDTQLKDQEEALSQ LFAGLESFEKRLRYTFKNKA
LLIEALTHASYIPNRITNCYQRLEFLGDAVLDYL VTRYYYDNPCQYTPAILDRLSAMVNNETFAVMAVQNRFHLYLKH
LSLNVILDRFVRSQEENGHLLMHNYFVLEETSEPHQSLASNIDVPKVLGDI FESVAGAI FVDSGMSLDAVWKS YLPFLH
DALEKFNEKIPISALRVLHERYPNALKFRKSESLADGRLCVAIEIQGTRVFKGAGGNSKTAKSAAKYALSVLNKQDDNQ
VEEC
>Phum

MAYHLIDNVHKS LTPKEPNVELLDASCVRNIIIVCLGNIGKTFIASKLIQNTSNNISVNNGTITLYVADSSVVVSRKAA
YFQRLTNMLVAEYHKGNCNKIKKIKKIPKLDQIVVAETETANYMLENELIVMSDINLLILHDCHCVLKNGIVPQKELMKF
YYNSSTKPHILSLASSL FSSNLDLSQIKSCLNKMEDILNCKAETASNVLSLLRYGTHPHEVLI EYGLHIPSSIDEIIEGE
ILKAIQFIKDHRYDPLEMYESFTEELQDIPDPKKEPEEIFLELLNVLHTLGPWSTDRAALVLLLQIEKLVTTTPYERHFL
LLCMAS TVLVKIRAILDDFFKDIDEKERIYKYSSPKILRVIEILRK FVPEKLDVKLDDNLSVKNQSNKLS DASKTISKAI
ETS AEKCTILNESEMKKQCLNCDCTIMTEYLLKRLDSSDIKNFTYDNFAKLEKEGYSNFIMNEDLSK SATQIENNERN
LKELGDKNEISLLSNNF SNNCNESLNQSNNNKPIKNFFYENNNKSKKEHSNISILRNKNGFLPIAKENILRAVNESLQMKK
RKINETHTIERENQDDFIKQLKNFGSDVPLCGLIFVEKKFTAKVLFHV LNDLQKQADEDFSLNVQFTVGVVDTGKDP
KGAENEHRKQELMLKRFRIRACNLLIGTSFLEEGIELKCNLVIRFDV PKHYRSYINTKERARAANSLFCLLVGPNEKDD
FVHELAKY YTEAVLLLEKCGHLEPEEEENEADLYTLVQVFCPLLKRECPHVDLSSAVMLLNRYCAKLP SDTFTRLTPI
CNIKTVKEFERKMYVCTIRLPINSPIKQIEGCPMPTRVLAKRFAAFQCCQLLHLSGELDDTLQPI SKESFWIREEGRLT
VTNDETEEIVSRDSLEPRPGTTKRRQY YYYKKIADSLTDCLPMADKPSFLYDISMVLSCPLPEEQNTRGRKIHPPENSSQ
FGILTKKRIPKIGPFPFIFTRS GEVQVKKLAQT TVQLTEIEINKIVAFIHYTFTSVLRQLKYLMFDPNSSENS YFIVPT
KKVKDSDSVEVDWAFLELIYEKRNQVPTFIPEENRKNFVFEPSK FEDAVVSPWYRNQDQPQYLVAEICTNLNPKSAFPG
TEYQTFEEYYLKKYNI VIQNLDAQLLD VDHTSARLNFLTPRYVNRKGVALLPTTSEEKRAKRENKQKQKILQIPELCTIHP
FPATLWRKAVCLPCILHRINALLADQVRMIVAAEEIGVGSIDLNEEFNGNPFNLGWSWADVLKQTKKILKQREETPEAVK
SIEMKDES DHVNKNKDTQNCNDNDRENSDLKVTLNDSCNTEIVLEESNESDKDDKQWEIGTWSNDMADDKLENIYDD

Text_s1-1.txt

SHDPDEIDDLRLPKNLTMIEETDGKEINDWDSKIFSEMKS SVRYCSPTSWLGEPKMDDEVLDGLSESTYSDDYFDGYMNS
DTDSECGGLKIEFKDDNVAEAVENHTATNFIKELNEVEESEPWYDSEKYSKYVDEFYANVQNRKEKIINSMDMYSEE
KSFDFSRSHSKNYQKTYPGCKNDSQSETSENFSGSADNRLPRLNENLQDNDFKSLPYFKIIEVLRKREENKQANPKNNVGLG
FELKTLPSNYKPFEEIDEICQKFSFDEQPDLSGHPGSPSLILQALTMSNANDGINLERLETIGDSFLKYAITTYLYCIY
ENVHEGKLSHLRSKQVSNFNL YRLGRKKMLGESMVATKFEPHDNWLPPCYVPRELERALIEAGVPASHWNRADLP SLKN
LSKEEICQLVKEKGEKLGIVNAEKKPINVTIPCLNLENMPCFIPYNLMTQHSIPDKSIADCVEALIGAYLVACGPRGALL
FMSWLGKIKVLPKEEKFESSENDPDIKKGVGASRPKIVMTNNGEVTFGKLSKPKRPLICHIPNPEGELEKLLRGDFIFEEES
IKYRFRDRSYLLQALTHASYYPNRLTDCYQRLEFLGDAVLDYLITRHL YEDSRHSPGALTDLRSALVNNTIFASLAVRH
DFHKYFRHLCPSLNEVVDRFVRFQQANGHAISEEFYLI EEEEEEEAEDEVVPKALGDVFESVAGAIFLDSGMSLDAVWRV
YYRIMKNEIELFSTNVPKSPIRELFELPETAQFGKPEKLADGRRVRVSVVEVFGKGFNGIGRNYRIAKCTAAKCALKQL
KRGKLELFFTNVGYMKSF

>Tcas1

MACYLNENVYHTFTPREYQVELLDSAKKRNTIVCSSASSAKAFITIKLLQEF SHKMRVPHGKQALFVLDGPNVPMITSH
VKLLTDLTVTSIDKEENPPSLKASNVIVTTAEVCLLCKKNFVHLD SYALIVIDCLYGGQQSLVREIMARYQAIQAPRPR
ILGLTAGLLGSEMOPDRLEAELQRLEKLLSSSVDTSS EILTIRLSCRPRERIVECFKPIPSPLQDKIKATITSCQDFLK
DHRYPSEIYDDDLLEEFKQVPDPKEQPLSFFDDFLEILDDLGPWSADRAAYGMLIKIEKLVKVPYERHYLLL CVASSV
LVSIRALCELEFQDYDKEKVFSTPKVLRFLQVLKQFKPTGDKPETCDKLPDLKDPKKGKGNKYGPRRPIYSRAQSD
EMLCALVFVKNRYKAEALFALLCVMSKSDEEYWWVSVSFSVNKIADPVREPRAEAESEHKRQEEVLRKYRSHECNIMIATS
ALEQGDCLKCNL VIRFDLPQS FHSYIHSKARANEAFHLLANENEVSDFVENLAEYNEVENTLLKRCYSLEPKNEE
LVADASSLQCRPYQPSAEPGANSVLSNAIALVNRYCAKLPSDTFTRLTPIWHEEKVENGYICSIRLPINSPVKKTVTSP
PMINTLLARRAAAFMICQLLHKAGELDDNLQPIGKENFKVNEEDWNSSALEESDEENLDPRPGTTKRRQYKYKQVADALL
DCHPIIGQPTYFYKIVMKLTCPLPEEQNTRGRKIYPPEPSQGFILTSKEIPKISAFPIFTRSGEVSDVLQCSQLIVT
ENQICKIREFLNYFTTSLRLQKYLTLFNP DASANSYLIVPTIDGATTTVDWDFIDL IYANLTVLPEI IPEEVRKSYEFD
PEKYRDAVVMWYRNQDQPQYFYVAEICSNLNPASDFPGSDYATFEYYLRKYSIQIQNKSQHLLD VDHTSARLNFLT
YVNRKVALPTSSEATKRAKREKLEQKQILVPELCAHPFASLWRKAVCLPCILYRINALLADQIRRTVALELNLGKI
ELDSEFKWPPLNFGWSLADVLKKSDEEKKKQEKIEPVIEEIPCTEIAKIEDFDQDDDEEMIEIGTWSNDMAQLNSDQE
FPVVRYASPTSWMDLQNTYDSSFSDSYSGDESESEWGLRIEFTGDNVAEAVDDENKDDDFELVDYSNVWVVEDESE
ITQTLRKQFHDACARNKDHILSSGILVSKSEQFQKCSDCDNTTTKDSQVANSYDFDFGKLFIELDQH KALQIDLAPQDAR
NEYDISETMTFKFDEQPNLVEHPGSPNVLLQALTMXXXXXXXXXXXXXXXXXXXXXKLFIELHQHKALQIDLAPQDARNEYDI
SETMTFKFDEQPNLVEHPGSPNVLLQAL TMSNANDGINLERLETIGDSFLKYAITNYLYSKYENVHEGKLSHLRSKQVS
NLNL YRLGRRKGLGEYMIATKFDPHDNWLPCFYVPKLEEALIDAQFPANCWTVADMAATRDMTLD DICSMVRQGESL
LSNSIIPYNLVTQHSIPDKSIADCVEALIGAYLIECGPRGALLFMAWLGRVLPQLEDGTYGEIELPKSPLSNHLTYPRE
ELDMLLDGYDQFERHIGYKFRDRSYLLQALTHASFSPNTL TDCYQRLEFLGDAVLDYLITRHL YEDTRMHSPGALTDLRS
ALVNNTIFASLAVRNGFHRYFRNLSPSLNEVVEKFVRLQEDSGHTLVDEL YLVVETEVEDEVVPKALGDVFESVAGAI
FLDSGMSLDAVWVYNNMMKSEIEQFSNKVPKSPIRELLELEPETAKFGKPEKLADGRRVRVTVVEVFGKGVFKGIGRNYRI
AKCTAAKCALKNLKKRGLIKKIDES

>Tcas2

MDEDELKPRNYQVNLMEIAIRENTIYLP TGSGKTFIAIMVLKQLCAPILRPYSDGGKISVILVNSVALVDQHGYVRD
HATFSVGTYTGMNVDFWSEAEWEQQFNKYQVVIMTSQIMVNLINNR FIDLKGVNLMIFDECHHGVEDQPMRQIMKH FHS
CTDKPRVLGLTATLLNGNCKLSKVMDEIRSLEVT FHSKVATVEGLDVVVGYSTNPQELFKVCQPGALS LDAKQVLNLRQ
LINDLEHINIKDEQNSVNLQSETLKP LEPSDVLKSLRNLISDLMIHIEMLGAFGGHIACVAHMIQIERIKKH CQNHLQF
IVLNYVMTIMGTTKLLLEETMAGYEPLKIRKFKSSDKVLKVFEILDEYKTKSDEELCCLVFTKRRFTA KVLHHIIDKASQ
VDPKFYHIKSNFVGNKNPNYNDTRENLYITKKNVNLVFSKEINVLVSSNVLEEGVDIPKCTLVIKFDKSEDRSYI
QSKGRARHIKSLYYTIVETTDVAKYDKKYSFAKIEENLVNDLLIGKN SERDHPNLSEIRNMYNDEKLE PYYVNGPNSAQV
NMTSAVALLCRYCSNLASDKYTTYAPEWYEEEDSSAKLRVVI FLPVVCLIDPIVGPYMHNNKDAKRAAALVACIKLHQ
CGELDNNLLPWKKQLDEADVSYL FTHWPQEKESEDAGNKKKRLHDKEIAPSVKSAIQPDRVLYLHTININPQYKRSDDLK
NAVTIYDLYKTP LKFGLLSPKPLPDLCKFP LFD SNGTLEIEIRNNVREVEFAANEMKEMREFFHLVFN DLLEILKEFLIF
DNTGMNSEMLLVVPVQDRCGDVCVDFR VIRDNKNLKNKLEPAATERINLVNTEETYLHKIVSPWYRSPPKMYVVTKVCPD
KSALSRFPNHEYPNFVSYSEKHSLSILDPSQPLLLVKGLSERLNAFKPRGAGGKRKKEKMYEELEYLIPELVIKQEF
SCLWIQARFLPSILSR LAYLLKLQQLQVDIARGIGAKAEYLKDCPPLELNLHLLHYEPNDPQLTQESDKSTPLIDNCLAL
ECPKNLRTIQYNKDFAAKMLEAEYYWKTIEEPKDIERNINVTVM DIEYYETFISHQPSKTGRLLKNDSVPKQONVPAITY
DCQFEAKQLQILDVQFDNQSPNLCQIYQAL TAAEANDIVNLERLET LGDSFLK FVASLYIIFKFPTYNEGKSTTLKGLV
SNKNLYYLGVKRLGGILKNSDLSPSDWVPPCF CIPQTISKAIGNKEYSVVSLFNCCISPEEQVSGNLNRKTLSDMTTEE
IAPDEENSYGNM CNFLNKQYVGDKSIADSVEALLGAYFLSGGIQGGIKFMEWIGILPLSEIQRLIETTQVDPV LNKKST
KTDVDFHMPQWREIEQRLGYTFTNRAFLLQAL THSSYSPNRITLSYERLEFLGDAVLDFLITCYIFEHC HLEPGQVTDL
RSSLVNNNTFASLVVRCGFHKFLMMNSNLQGHIDKFADYLASKNYVIDDEVLLLEEDEM NIAEYVDVPKVLGDI FEAL
AGAIYLD SNKDLKT VWRVYKIIWREIDLFSKNVPKNVIRRLYECHTVYPPQFSKALEVGNQKTMVSLDFMCEGRKKRVH
GFGTNKILAKRAAKIALRALKL

>Nvit1

MAFFFNDHIHTKSFTPREHQVELL YSAKEKNIICLGKITEQTFIVTKVIQELAINNRRTL SHDGKRTIYLL ENEEACTE
KANHIKQLTDLKVLKCNVDKDFDSKEFSSNTQVLILTVDMCSKLL LANKINPNQINLLIIDKCHNLVTNCKLSKILNIFR
NCDSTTKIIGFAVPLYSLTKEPGALS YEIERIESLLQCQIETASDLLSILRYSPPKPEYLLAYKIEDKKVVEKDITDFTS
FTLNFLSDHRYDPTEIYSEEFLEDIKKIPNPVHKPCQMIEEFLRILNLTGLWAADKAALVLLMLIEKLIKTPYERHYLL
FSVVATLFLKIRAYCDDIFSGL TEAEKILKFSTPKIHR LIEVLKTFSPPPKEETVTPNNRVEKNDSANTKDQNHDKSVKN
LHKKTDDGGFKFRPFRHSRGLTDPDLLCGVIFVDDAVTAKVLFYLLNEMSKCEESLHFLSPLFTSEKSNDELFCGRDLE
EYRQEEVLLKFR IHECNLLISTAIIEGIDIPKCNFVMRFEFPKNYQSVYQCKSRARATDALHVLNPEIESEVYIQRL
AHYHYIEKILLSKTSEKAEDECETEADAYDVLPEYKPLDEEDAPKVTLSAISLVNRYCALKPSDTFTKLTPEWIKINN
ITVENKQMYICSLRLPINSPLKYSISSYPMPNKALAKRMAALQMCVCLHKENEIDDYLLPVGKENFKAHPEDSEVPALPD

Text_s1-1.txt

DENMDL TEARPGTTKRRQYKYKIIADALIDCKPELEEPCYLYHINMVLSCPLPEEQNTRGRKIYPPEESDIGFGILTRKK
IPKVCFPFIYTRSSEVRVDLKLKETVVLDESKIDKVVSLNYFTTNVLRQLKYLMLFDPNVSENSYIIVPVKKQPNASD
ICVDWDFVECIYANRDSMPNRRISEENRQNFVFEASKYHDAVIMPWYRNQDQPQYFYVAEICNHLNPKSSFPGADYSTFEE
YYSKKYGIQIQNLEQPLLDVDHTSARLNFLTTRYVNRKGVALPTSSEETKRAKRENLEQKQILIAELCAIHPFPASLWRQ
AVCLPCILYRINALLADQIRRHVAQSISLQGETLDEFNWPPPLDFGWSLADVLKKSKEKNTNSNKPEKEKSDKIKEEA
PAVKEKPENIKAERLDDSGNDEMEIGTWSNEMVNMAPVVVADAGLAVQSNEYWGNIRYGSPTYDSEFDDASINSSDDDFG
DDSDSDSQYDKEEGLTIAYTGDNVAEAFDKKGAFAEAEKEKQNCVELQEKSSELMHWNFEGDLSEFYQLHRLAHIDKAKM
NVEIILSQDIFIPCDKPVAFKRKTSKVLNNSYVERKHEEYVSVSNYFKHELIESAKKRRNKQGAELTKLVPKYNKGSF
SFDYQPELGDHSGPSPSLILQALTMSNANDGINLERLETIGDSFLKYAITTYLCTYNNIHEGKLSHLRSKQVSNLNLRYR
LGRQKMLGESMIASKFEPHDNWLPPCYVPKELEQALIESGVPSAMWNQADIPPFSEINDLVRETEHKLEIKSELS
QSEHGALNGLNGLTGDPLRSFIPYNLITQHSIPDKSIADCVEALIGAYLIACGPRGALLFMSWLGIIHVLPTTEVVITSDS
KPTDRLPGSTPYVELGEEGGSTKWTQLRYKKLQEPSPFLYVNPQPEIELEIMLDGYDSLERSIGYTFRNRSYLLQAFTH
ASYQPNRLTDCYQRLFLGDAVLDYLITRHL YEDPRQHSFGALDRLSALVNNITIFASLAVRCGFHKYFRHLSPLNTVI
DRFVRIQEENGHSISEEYLIIGEREFEAEAEEDVVKALGDVFESLAGAIYLDNSMSLDAVWAVYFEIMKSEIEQFSA
NVPKSPIRELLELEPETAKFGKPEKLADGRRVRVTVDFVGKGSFKGIGRNYRIAKCTAAKCALKKLKSQSCIKGRK
>Nvit2

MDQTPEKEFKARPYQIYLYEKTIEQNSILYLPTGSGKTYIAVLLVKRLSGDVQRQYTEGGKRTIFIVNTVALVVQQTAF
TRHTGLVCKGYSGDMGVDFWSKEEWRKEINTNQVLVMTAQIFLNLTHGYISLKDINLLIFDECHRAVKDHPMRQIMQRF
QDYPEKELPRVLAMSASLLNANVPLGKIETTLRELVTFQAKIITVESLAYVTDYATNPKEFVEYETPHKISLLNEISA
IDYASSILKQVVLPNRMENPESSMIFKPVSKTIKLRILSDVEEHLTDMGLYGGSESVLQHIIQLECLRFGDEKEAIA
MFDIITQLVKIQKLLSDAMNEVQLPDTIQRFSKVLKMEVLKTFYNNMENKRNFCIIIFVKRRFTAKVLYQLLSKLS
SCDEDFQLKQYMGYSNPNYKNARETVCIKWNDEVLTFRNGVANCVVATDVVDEGVDIPSTLIVRYYPMDFRAY
IQSKGRARHSTSHFIILASSEDDYISRYRSFQQTERFLRQALHGKSDYRIKPSNNDVDALLYQYVIQPYIVKDANGTISV
ITEQSAISLINEYCANLVKSKFIMLTPTWVKEDVSPSAYRYPEALCSAFPKNESLYLHIIDMKPNYPRPNDNRHLVFN
LLSNSQTYGILSAKKIPEIPSPFIFMNIQDLQINIKSNNQLKSLTEMEIEQLKLFHTLVFSKILKVIKEFMIFDASNMEN
NFLVVPINERGQINWDIVNRYKIIASISPSENLVKSEYDLALVTPNYRASMYIVTQVCEYLKAESSFPTSDYNSYVH
YFKERHYIEIKNPAQPMLEVKPISSKINCKPRSIKANLSKRKRASLTEDFEHLVPELCDRIDFSLYLWKAATLPSIL
HRISQLIAAEELRVKIAHEAQLHISSLEAGKKWEPQIVDQFSQNSSESNDTTFDESMLDEEDLEAENLLVPDTSLEVDI
LSQERSSYSWSKEQEPTDLDRNAQQIQLIDIEYQCQFMSGKVSAPVLKMLTSSNPYGPSPVDILQSLTTKVKNDVFNLER
VETLGDSEFLKFAISLFLYQAYPTCGEGLTHLKGKLVGNLNLFYCSKQKNIAGRMHVEDFAPTGNFVTPAYAAHQVQLQI
LRAEKVSANILYEIRVPAERFSGCISNNTDMMQDKVLAWPDSVEKVTHTGMEHFLGIQVVSDDKSVSDCTEALIGTYLLH
LGIKGALQILKWFELPKSLNVDQYLYSEVENPQLGDGVNVNHPWAQTMEEERLGYKFRNRAFLLQAFTHPSYMNRQTA
SYQRLFLGDAVLDLFTIHIYETCGNLSPGELTDLRSALVNNITFACLAVRYGLHTALLALAPKLDLIDKFVKFQENR
NYKIDDELWVLEEEECNMAEHVDVVKVLDGIYESVIGAIFLDSGKNLEVVWQVIYNLMKNEIDLFSKNVPKQPIRVIH
ETQGANPKFS
>ApisA

MALTFDDYVNTKTFTPQEYQVEILSSAKNKNVVISLEGITNKVFVILKLIQELAFKMRKSEKRKWTLLVLEEAEDVQYM
FSIRHLTDLKPLEIRSVDHSDFINVDVITVTENLCLIEFEKNCLHFDQLSLAIFNNCQKIIVENSTYFKICNFIKSNSQC
RILGIVHPLFELSDPLFIENQLKFISHKLFSSETASDIVSLLRSCCKPKEVLEICAVQSEDFPNIGIKSIVQSLTDFL
NDHYNPENIYGDDEELREELKKIPCKNVPLQMISEFSDVLKTMGPYCADKVALTLLYRIENLKVKTPYERHYALLCIV
STLLIQIRAKFDNIFYNYSNIDKVLKFSSPRVLRINTLKKQFKPKSFNSSNNLAEVCKDQEGNKDKIISKSAKRCKLT
RYYNKTHRQRQSRDPCENVECAIILVNSKMTASVLYHFLDLKNSNEEDFGWIEPQYMDNADPITEPEEFIERKT
QENILKRFQRNCNVLIGTKELETGINLPRCNLVINYNLPLSYKLSMSKSKARTLDAYVLMFDEHATSKILSCCQLYK
DINQMLLSHCTLKIDMEKEIYADRNFNFVSPFKPGKEDSAFVDMATAIGLVNRYCAKLPSTFTTKSPIWKIQSVIYE
NQPMFVCSIRLPINSPIKHEIHGHPMPCEILAKRIAALVCRQIYSQGLDDTLQPITKESFHALFNTDEIIPDTEDSL
TCKTKDSMPGTYKRRQHYKRTAEALNDCRPVEGKPCPLYAILNITCPIPEQENTKSRRIHPPESIHGFGILTTKHIP
SIPAFPIYPRCGEVYVRLKRIKTDMLTQHQLDNIAFTLNYTFSNVLCFKKCRVEFNPTTTHNCCYIVPTLRDASGSIDV
DWKILDNVQTFRDNKLSIITNEDRIHFQPDIEVLKDAVVTWYRNQEQPEYFYVEEICTHLPKSPFPDTEKYKTFEDF
YYHKYSLKIQNLEQPLLDVDPTSARLNFLTTRYMNCKGEYLTSSDATKRAKRYDLSHKQILVPELNCNIHLFASMWRDA
VSLPSILYRINALLDAEIRTIVAKDTSGLVNLDFHIWKPLDFGWPIKDVLIIRKEQEKEISTIKVIEPVETVKRFIDE
VEKEEKIDELNDPNWMEIGTWSNDMADFAMNNIDNTTKESEINEVNDLQETETFSFDFQPDLYHHPGSPSILLQALTANA
IDGYDLERLETIGDSFLKYAVTAYLYCTHNNVHEGILSDLRSNHVSNLNLRYLGLKLMFGERMISTKFEPRSNWLPPCYI
GPHKLEKAVINSNISITQHITNDKSIADCVEALIGAYLISCGTRGALLFMSWLGITVNLTDSDKLGYLKLPSPLLRN
NDPEGELTKLMDGFESFEQHLGYFQDRSYLLQAMTHASYPNCLTDCYQRLFLGDAVLDYLITRHL YEDKRRHPPGAL
TDLRSALVNNITIFASLAVKNGFHKYLKHLSPGLSEVVYRFVAIYQEENGHTIDEEFCFSGEDDFEDAEDVEAPKVLGDIFE
SVAGAIYLDNSMSLDAVWVYHKIMENEIKQFSKNVPISPVRELLELEPNTAKFSKPEKLVLDGRRVRVIVEILGKGVFKG
IGRNYRIAKCTAAKCALNHLKEIKK

>ApisB
MDQPPKDTSEIVPRKYQLELLEDCKKANTILYLPTGSGKTYIATMLVKSMGDCLTKPIQGKKWTFIVQSVPLVSQQAG
NLRRHLWNIGTFSGDMNVDFWSQQQWNEILEKCHILVMTAQIYLNLRHGYMHIKDANLLIFDECHHAVALHPFKQIMQ
VLHDTDLKTDERPRILGLTATLINANTKNVKEELTKLQITFNATIKTRYDDNIQIFSARPKEFIRLYDEYILDDEIKAI
KRLNNTCNLRIQAPMKIEPIKENNETYYLEYQKQASKNFANYFLDIEANLVDGGAYIAYLVTFWHFVMEIEKKRICKD
AKRLTIMSLVSELITRKMLYDYMNKHLIHKGHSTKLHNTSPKLLKLEIILSSIKLTDCLVFDVDRRTAKILYHYIK
DYIEEFQVNIYCEFIVGARGIFSPDCKEMVYKKQNNDIKKFNNDVINVLTITSEVLEEGVDIQTNCYVIRYDSPKNFP
SYIQSKGRARSTDSKFIILMVPNQLKFKQTQTEYNKMEEEIVKLLVDNDIDDENDEKMTVEVFTHTDRALLTHEIAISII
NRYCYSLPQRDFDLSPWEYMAQNHNMSKKYKLLPINSVIKTPIDNGLFCTNKKNAKKSAAFNACIALYKIGALDENLMP
INIKNTAIFNDLWFPWDEGDVEATKYKLAGNTKMKRMVHIESPTYLHGSYPQVNSQSYLHVLHCVDPYIKLEESKYE
TFDKLLRSNEEFGILTSNKLQVSNFPIFLPFGNVNVSIEVNVEILYLDHKSHELENFHNLKFDVLDVGVKDFIARSYDN

Text_s1-1.txt

GDNSYL VVPILSTDNIYNVDWNVINVS KFVETEPTMEQRKSKFYEEYLKTPNVISPWYRNIVPVQRYIVTDVYWDMTPE
SSFPTSEYDYATYFSDKYSIQVAHKNQPMIKVKS LGVSKINYLVPRISTGKIDKRSEYIEILVPEFCTWHKFPSVYWLK
ALMLPTLLYRLDKLLLAEDLLVKINSLCNI EVEDNANKDVTMIDNFSGNHEGKKNIILPISETLKYLENSHDNNMVDWEW
NSNNFPIDIERQKNTTMLDVLNYHTFMYAINKSKSNESIDSQKSKSSSLNEFISKSKHKSPDGKSKHITECKKRRLLSSYN
KEQIEEINTFMIQSKPEPMIFQKLVELNIVNFVEMNDILKNVLTENMNTLNL TNNSPPKLSLSIEVSKQNIGPKQSDILK
VITPPFANDMFNYERMETYGDSFLKFAVSLVLFDAF PSDNEGVLTELKMKLVGNRNLLYVGRNLLNGSYLVVNVFDPNMD
WIPPCFGVPIKLEIIEEGNFQSDVLHQITIPRDNQTTGILSDETWIEIETELLKFKRLSDMPI SDSDEEDGIPQIGHSQ
SDLFLHKQSVSDKMVADCVESLIGTYVYKRGVEVGFKVLQGLGIIPKQFIDTFKPKPIDDLYDGLDFSKILPGYELLERR
IGYSFKHKHLLAQAL THPTYQFGFSECYQRLEFLGDAILD FLITTYIIEHCYHKTPGEITDIRSSLVNNITFASLSARIG
LHRFILAKSVQMTEAIDRFYEHQQKNNHKIGQEVLYLIEESDCYAAESVDVPKVLGDLFESLIAAIYLD CNRDLNFVWAI
CYRFLKEIKEIFCDNVPKNPIRILHETK LQPNFSKPDASDKAADKGLGTMMRL EIMLNNKIMRIYGFQGNKKEAKVAAK
MALKNMKMMLS

>Dme11

MAFWCDNNLHTTVFTPRDFQVELLATAYERNTIICLGHRS SKEFIALKLLQELSRARRHRGVS VYL SCEVGTSTEPCS
IYTMLTHLTDLRVWQE QPDMQIPFDHCWTDYHVSILRPEGFLYLL ETRELLLSSVELIVLEDCHDSAVYQRIRPLFENHI
MPAPPADRPRILGLAGPLHSAGCELQQLSAMLATLEQSVLCQIETASDIVTVLRYCSRPHHEYIVQCAPFEMDELSLVLAD
VLNTHKSFLLDHRYPYEIYGTDQFMDELKDIPDPKVDPLNVINSLLVVLEHMGPWCTQRAAHHFYQCNEKLVKTPHER
HYLLYCLVSTALIQLYSLCEHAFHRHLGSGSDSRQTIERYSSPKVRRLLQTLRCFKPEEVHTQADGLRRMRHQVDQADFN
RLSHTLESKCRMDQMDQPPETRALVATLEQILHTTEDRQTNRSARVTPPTPAHAKPKPSSGANTAQPRTRRRVYTR
RHHDRHNDGSDTLCALIYCQNHTARVLFELLAIESRRDPDLKFLRCQYTTDRVADPTTEPKAELEHRRQEEVLRKFRM
HDCNVLIGTSVLEEGIDVPKCNLVVRWDPPTTYRSYVQCKGRARAAPAYHVILVAPSYKSPTVGSVQLTDRSHRYICATG
DTTEADSDSDSAMPNSSGSDPYTFGTARGTVKILNPEVFSKQPPTACDIKLEIQDEL PAAAQLDTSNSSDEAVSMSNT
SPESSTEQKSRRFQCELSL TEPEDTSDTTAEIDTAHSLASTTKDLVHQMAQYREIEQMLLSKCANTEPPEQE QSEAER
FSACLAAYRPKPHLLTGASVDLGSALVNKYCARLPSDTFTKLTALWRCTRNERAGVTLFQYTLRLPINSPLKHDIVGL
PMPTQTLARRLAALQACVELHRIGELDDQLQPIGKEGFRAL EPDWECFELEPEDEQIVQLSDEPRPGTTKRRQYKYKRIA
SEFCDCRPVAGAPCYLYFIQLTLQCP IPEEQNTRGRKIYPPEDAQGGFGLITTKRIPKLSAFSIFTRS GEVKVLSLELAKE
RVILTSEQIVCINGFLNYFTTNVLR LQKFLMLFDPDSTENCVFIVPTVKAPAGGKHIDWQFLELIQANGNTMPRAVPDEE
RQAQPFDPQRFQDAVMPWYRNQDQPPQYFYVAEICPHLSPLSCFPGDNRYTFKHYYLVKYGLTIQNTSQPLLDVDHTSAR
LNFLTPRYVNRKGVALPTSSEETKRAKRENLEQKQILVPELCTVHPFASLWRTAVCLPCILYRINGLL LADDIRKQVSA
DLGLGRQIEDED FEWPMLDFGWSLSEVLKKSRESKQKESL KDDTINGKDLADVEKKPTSEETQLDKDKSKDDKVEKSAIE
LIEEGEEKLQEADD FIEIGTWSNDMADDIASFNQEDDDDEDADFHL PVL PANVKFCDDQTRYGSPTFWVDSNGESGFKGPK
SSQNKQGGKGA KGPAPKPTFNYYDSDSL GSSYDDDDNAGPLNYMHNNYSDDDDVADDIDAGRIAFTSKNEAETIETAQ
EVEKRQKQLSIIQATNANERQYQQTKNLLIGFNFKHEDQKEPATIRYEESI AKLKTEIESGGMLVPHDQQLVLRSDAAE
AQVAKVSMME LLKQLLPYVNEDVLAKKLGDRRELLLSDLVELNADWVARHEQETYNVMGCGDSFDNYNDHHRLLNLDKQL
KLQYERIEIEPPTSTKAITSAILPAGFSFDRQPD LVGHPGPSPSIILQAL TMSNANDGINLERLETIGDSFLKYAITTYL
YITYENVHEGKLSHLRSKQVANLNL YRLGRRKRLGEYMIATKFEPHDNWLPCCYVVPKELEKALIEAKIPTHHWKLADLL
DIKNLSSVQICEMVREKADALGLEQNGGAQNGQLDSDNSDNDVFCFPIPNLV SQHSIPDKSIADCV EALIGAYLIECGP
RGALLFMAWLGVRVLPITRQLDGGNQEQRIPGSTKPN AENVTVYGAWPTPRSPLLHFAPNATEELDQLLSGFEEFESL
GYKFRDRSYLLQAMTHASYTPNRLTDCYQRLEFLGDAVL DYLITRHL YEDPRQHSPGAL TDLRSALVNNTIFASLAVRHG
FHKFFRHLSPGLNDVIDRFVRIQQENGHCISEEYLLSEECDDAEDVEVPKALGDV FESIAGAI FLDSNMSLDV VWHVY
SNMMSPEIEQFSNSVPKSPIRELLLEPETAKFGKPEK LADGRRVRVTVDFVCKGTFRIGRNYRIAKCTAAKCALRQLK
KQGLIAKKD

>Dme12

MEDVEIKPRGYQLRLVDHLTKSNGIVYLP TGSGKTFVAILVLKRFSQDFDKPIESGGKRAL FMCNTVELARQQAMAVRRC
TNFKVGFYVGEQGVDDWTRGMWSDEIKKNQVLVGTAQVFLDMVTQTYVALSSLSVVIIDECHHG TGHPREFFMRLFTIA
NQTKLPRVVGLTGVLIKGNEITNVATKLKELEITYRGNII TVSDTKEMENVMLYATKPTEVMVSFPHQEQVLT VTRLISA
EIEKFYVSLDLMNIGVQPIRRSKSLQCLRDP SKKS FVKQLFNDFLYQMKEYGIYAASIAIISLIVEFDIKRRQAETLSVK
LMHRTALTLCEKIRHLLVQKLQDMTYDDDDDNVNT EIVIMFSTPKVQRFMLSLKVSFADKDPKDI CCLVQVRRYTCCK
IYGLLLNYIQSTPELPLITRQLDGMVGRNNI PGFESVLE ERKWQKSAIQFRDGNANLMICSSVLEEGIDVQACNHVFELD
PVKTFNMVYQSKGRARTTEAKFVLF TADKEREKTIQQIYQYRKAHNDIAEY LKDRVLEKTEPELYEIKGHFQDDIDPFTN
ENGAVLLPNNALAILHRYCQTIPTDAFGFVIPWFHVLQEDERDRIFGVS AKGKHVISINMPVNCMLRDTIYSDPMDNVKT
AKISAAFKACKVLYSLGELNERFV PKTLKERVASIADVHFEHWNKYGD SVTATVNKADKSKDRTYKTECPLEFYDALPRV
GEICYAYEIFLEPQFESCEYTEHMYLNLQTPRNYAILLRNKLPR LAEMPLFSNQGLHVRVANAPLEVIIQNSEQLLELLH
QFHGMVFRDILKIWHPPFFVLD RRSKENSYL VVPLILGAGEQKCFDWELMTNFRRL PQSHGNSNVQREQQAPRPEDFEGK
IVTQWYANYDKPMLVTKVHRELTPLSYMEKNQDQKTYEFTMSKYGNRIGDVVHKDKFMI EVRDLTEQLTFYVHNRGKFN
AKSKAKMKVILIPELCFNFPNGDLWLKLI FLPSILNRMYFLLHAEALRKR FNTYLNHLHLLPFNGTDYMPRPLEIDYSLK
RNVDP LGNVIPTEDIEEPKS LLEPMPKSI EASVANLEITEFENPWQKYM EPVDLSRNL LSTYPVELDYHHFVSVGNVCE
MNEMDFEDKEYWAKNQFHMP TGN IYGNRTPAKTNANVPALMPSKPTVRGKVKPLLILQKTVSKEHITPAEQGEFLAAITA
SSAADVFDMERLEILGDSFLKLSATLYLASKYSDWNEGLT EVKSKLVSNRNL FCLIDADIPKTLNTIQFTPRYT WLP
GISLPHNVLALWRENPEFAKII GPHNLRDALGDEESLVKGNCSDINYNRFVEGCRANGOSFYAGADFSSEVNFVGLVT
IPNKVIADTLEALLGVIVKNYGLQHAFKMLEYFKICRADIDKPLTQLLNLELGGKMRANVNNTTEIDGFLINHYLEKNL
GYTFKDRRYLLQAL THPSYPTNRITGSYQEL EFIGDAILD FLISAYIFENNTKMNPGAL TDLRSALVNNTTLACICVRHR
LHFFILAENAKLSEIISKVNFQESQGHRVTNYVRILL EADVQPTPLDLDELDMTELPHANKCISQEA EKGVPPKGEF
NMSTNVDPKALGDVLEALIAAVYLDCRDLQRTWEVIFNL FEPELQEFTRKVPINHIRQLVEHKHAKPVFSSPIVEGETV
MVSCQFTCMEKTIK VYGFSGNKDQAKLSAAKHALQQLSKCDA

>Agam1

MSLFHWTDGNIHTTALTPRDYQTELLATAREENLIVCIAHNSAKEFLAVKLIQSMRTNRWSSHPEAPGKAIYLRMDRSL
LSSMVSNTDLQVANVDDVEDSEGSHEPDGASNTPVTDVASADVLFFGSETLLQYIEQGTVRVQDISLLIVDECHKNYG

Text_s1-1.txt

RQELWEICARLTHQAPSSDRPAQRTRILGLAGLPHGAGCTPERLXCWELHYLERCLRARIETASDITSVLRFRSTKPTTELIL
ECIPPKPSNL TQLLRMLIQRQIAFLKQHRYEPLAVYGLDGNDSASDKPDTDGEPEKDEENDDLRELKSIDPPTVGPPLSY
LKQYLELLEDFGPGWADRGALELLTTIDQEKVKVPYDRHFLLFCMVYTTTTLQARATVASVFAQHDTLELRIKRYSTPKVR
RLLEVLAWFGQNRNPKDRNPATLHHHQQVHNQQRILYCFCRNVECKELEKSYHTFQAQIADVDERIKQLDGLLASVRK
RTERLNLRKHKVVDNAAAEGGGTLGVQSPRHGHESTRANFRRRKRFAGGGSHHRSNDTTDALCGLIFCNRRAMARILYVLL
YEVRSQREFEFISPYTVDKVATNPQNCLKQTTIEHRKQEEVLKRFMRHECNLLIGTSVLEEGLPKCNLVIRWNSPA
NYRSAQCKGRAKAPGAYHVLVFTPENAAASRNEQQEDMASIEDVSLDGMIEPGPKDDNTRETIDDQDRRMIEATDAMIE
QVAIYREVEKLLAKCRNGETPPDWELKHADC FNHLEEVYRPSGGAVALQSNGT CASLWLGNAIQT LNKYCAKLP SDTFT
KLTP IWRCATTVRKGRKLYQYTI RLPINS PWKED I LGLPMP TET LARRLAAYMTCRMLHAAGELDNSFQPFGEAFRAFE
ADWENFELEESDAKILSENSDPRPGTTRRRQYKRIASVFNECRPDAETVAYLYHIRMELICPIPEEQNTRGRKIYAPE
ESAQGFGLTTKLIPKISSFPITFRSGEVKVS LDCPQRVKLSAHLQEMVNCV KYTF TKVLR LQLKSLMLYDANATENC
FIVPTVKQAVAGAAGKDEPPLQSDDVMDWFEVKEIATNVHRSGPTFIPDEARKGYTFDVGKFRDAVVM PWRNRDQPQY
FYVAEICNHLSPKSTFPGSNYATFEYYHRKYKIHQIQNQPLLDVDHTSARLNFLTPRYVNRKGVALPTSSEETKRAKR
ENLEQKQILVPELCTIHPFPASLWRAAVCLPCVLYRINALLADEIRRQVARDLR LGWENDELQEQGFQWPMLSFGWNL
ADVLRKTKEQKIAQAEIDASAPEVEDEVELDKEAPNRLDAEVEDEEDGLKMEGVIAEVEKSQVDGEDDTGDKKTDSD
GTLLEIGTWSNEMAVGVGTNDMGEEGRRGASSPSFLRYSDCSSNSSANFYSSDEYDEEDDYLYDYGSEKPKNAIEPSQ
EAVSGTDNANDSGDKPGSRNRITITQSQDTVVNLGGRLKIEFKSETDAEAIDSECDLQRQRTQQSIIERSRQNDMLYQSSK
NAVDGFCYSPSDRQCAEEREQAEQRFERQKNTKDTIRQQGLSVRWNEPLSVSHWRSKLDESEAATL GALMDDLGGQPFT
ELVPYVEPEQLFELLVRNGSSGERWLRHL DYCLNR PFFPERYTFI FRGGSQFNDLDESCQEGEGASSPKVIELTICDP
FPAIAAGQMACNYPKQHSMSRSDANKNGHDSAGVLSTTDEAGGVDFSFYDYQDLSQHPGPSPAIILQAL TMSLNDGI
NLERLETIGDSFLKYAITTYLYCRYDNVHEGKLSHLRSKQVSNLNLYRLGRRKRLGDCMIAAKFEPHDNWLPPCYVPKE
LEQTLIDAKIPACHWNLADLPDIKRLSCAEICQLVKERARAKRREDVDRFDLQHADSNNGGEDGNEDDDDDDDDDEDDTTT
GEGNEHEADNGSCYIPYNLVTQHSIPDKSVADCVEALIGAYLIECGPRGALLFMAWLGI RVLPIREPPVKLN SNNETALT
PYKATGQNDGPLSTGVTIAEYGHVVAPSPMVRANITFGGIETGAAATSRELARLLQGFEFEQALGYRFRDRSYLLQAM
THASYSNRLLTDCYQRLFLGDAILDYLI TRHLYEDRRQHSPGALDLRSALVNTTI FASLAVRHGFH KYFLH LSPGLQE
VIDRFVRIQQENGRITTEEYYLPDEDELGEYGMAGEDGPGEGRGVGEAEDVEVPKALGDVFESIAGAIFLDSDMSLDT
VWVYRKMMGPEIEKFSSV PKSPIRELLEM EPETAKFGKPEKLTDGRRVRVTVEVFGKGTFRGIGRNYRIAKCTAAKCA
LRQLK LKLYANHHKRR

>Agam2

SQNKPEMEDFAPRNYQVQMKEICLAKNTIIFLPTGSGKTYIALMVMKEISHQLRNTVHEGGKRTFFLANTVALAKQQAQF
FARHMPFNVRLYTSEVNDAWKS DRWHEEFSEGQVIICTAQILLDVL RHGYMSPANINL IVFDECHRAVGHQPMHAIMKE
IVAAPASERPRVLGLSGTLLFKELKMASQVPDELERLENTFSSTIATVANYDDYATVASFSTNPNEVLVTYSKPAVHLM
LVKEMWPRIDAFVWELLQVYLPEYSTQSTRTLQKSFCFKPLKEVKRALTEFKHQLEEHGMYAGSLAILAVIVQLEVSKRQS
PCDKARQVYRSAISFCESIRHELVEAMSGLRGTHQILSFSSDQARKLLKYLEDSYRTAEDKKNQALVFKRRTAKVLYH
LIRIYFHL SKRDNDELVEPIVKPDFIVGANAAL EESIDAILVREDRRVIENFRKRKINVLCATNVL EEGIDLQMCNM
VIMYDAPLSYASF MQSKGRARMKTSTYLMMPAADLQFAKRMKLYRDIENR LKEELVGKTINRPEL ENDRVREKELDDL
IPPFYTPFKAKLDALSAIQLLNRYCMSMPRDLTGNSVNTWERTIDRSPTEIIVTVKLPLOSTVREVIHGQTMKNLKLAKQS
AAFNACKRLFEVGLNMYLLPIATKDKVEELSEYFKLWRKMSDEPNPKAGTMKYVRGHI VYPEETVCPQADGEQC
YVYIVRMAHF DANTHLENVRI FQELYSSANNFGILTRKRLPRLARMKLFVTLGAI GVEIVPEPVCITLAPDSGELQRLK
RFHLL FRDLLK VVKPFTVLDALPEENGFLIVPMLRSQSIDWELMGKFPYLRPAAETSTRARQHLRFETEYQLLRVVHPW
YKNDPDQNYVVRVDDL RPTSPFPNAKYDSYEQYFAQEHQVVRHEDQFLIEVKGITTSLNRLHPGA EADGGASTRSR
YWEFQEILPEL VHNFEFPADYWLKATLLPSALH RVHYLLA EGIRV DLARNAQV GSEHCERVEDVIIDRVLSIEHQKST
ALYGGEDDDDEDEDEDEE ETEEGARSMQLNQHSDSISLADRVQY PWEENEQPKDLERNWDLERNVDLSVIDIDYASF
VKKYQAVTIVTDVNRSTETNIDRMLANLN MNKNEAASPVKALPAL EDGHPVAQIAMLR L TMDNTANVALQSDLL
QALTTKSSADVNLERFEVLGDAFLKFAVS VYILFRHTSWHEGYLTTCKGRMVSNRNLLYCAMGYLPGKIKAHFPDPKN
DWVPPLSTVPGAVRRAMVDANESP ALLYNLKLTEEEIQSGVVAQATVDKFLPLIEQAPAPSQSTLHTV LQQVQIRDKVVA
DVTEALLGVCKTVGYERSFRFLSHLGIIPKGADVPMLLRTTTFPIGDYFVVRHKVDQLLCNPERIEATLGYFRNRYTL
LQAFTHTSYTNSL TGSYQLEFLGDAVLDFLVSMYIYERNPSMSPGQLDLRSALVNVNLACVLRNELHRRHLSQSP
MLTDAIGKFAVHRHQNQSNWVRLTTEESDTPMAEYDVPVKVLDVLEALIGAIYLD SGNDLAATWEVCFRLLRDEIA
DFTRKTPIQVVRQLYEHPEASPHFSAPFVEEYVVKLSYTHRSQRQTYGFGKNKDDAKRAAAKIALSKIM

>Aeeg1

MDMIMPQDDFI PRDYQRTMKTICMQKNTIIYLP TGAGKTHIALMVIKEMGKDL DKPLTEGGKRTFFVNTVALAKQQA
FLSHNLYDTSIYTS DRNVDAWQDKWLEEFAYQVIVCTCQILLDVLKHGYLSVKHINLLIFDECHHGVGEHPMHGIME
QFLRVPKSDHPRVIGL SGMLLYKQIKSVALVSP ELERLENTFNATIVGSYDAFTEVCKFSTDPNELLVSYSTLR LSPV
MADIVNNINAFSQTIEEFHLPKYLNQNKALLKDRPKPLKEIRKLFTELIYQLGDTGLFGSIALLLGLIVQFELDQRQSDS
SMLRLALRSCITFCESLRHQIEKLMSGLDMKTKLTKFSSLKVQRIDQLEKLYEENRDKAKT LIFVQRRFS AKVLYHLL
KIYFAETEDANLIVPDFMVGNNGSMPESIEQIL SAKKDRRVLERFKNETNIVITTNVLEEGIDLQMCNTVVKYDHPQTF
ASYQSKGRARMKNSQYVMVLDNENRHIFLEKYRLYKSIEEELRRLIGKTINRPDPLDADVHKELYNEIIPFF TAKGA
KLDALSAIQLLNRYCMGMPRDAFTNTNVTWERIDLKDGRIIVEVLLPLQSTVREKISGNPMRNIK LAKRSAAFNACRKLY
ENKELNEHLIPIDCKYQLNLDVYFRHWKDFDADLGLAGTKCIRTHAIQYPKQTTECFPQPGKPCYIYVLRIAGFA
QDPTNDNVNI FHSLYSEN NFGLMTTKPLPALAKMKFFVTLGLINVHIEETPIVL PNGGSEIELALLRQFHVTVFRDVLK
LWKEFLCCDYDNEENSFLV VPLKNSTHLDWKLIREFQNLSEPPSEISTIARNKMEFEADKYRHKVILPWykNNKEQPYVV
TMVHEHLTPESPFPNPEYGSYANYFSQAYHLAVVKPDQFLIEVKGITSYLNR LNPGVEDDGKSTRSKHWRFNEILIPELC
HNYQFPADYWLKATLLPSALHRLHYLLLAENIRVDLATGANVGCLENHTIEDVDVEYKERKKGKQLEELQLMEFEDEDED
DEFDLEEAKRSLVAPENLSELARNQMC SITGDIPLPWQEDEEPVDIERNWDQVSKLDLDY YNVFVNKFSDLSMREKAAER
ISNAYTSAVYRRAAGSPKREMAI LDPVVDQKFAIKLQLTPANTVNVNLQQKNIKALTTKSSSDVFDLERYELLGDAF
LKFSISLYLVKHYKEWHGFLTAVKGGIVSNRNLYCAIKYGLPGMLKIHKFDPKNDWPPLATVPKNIKR TMQSVNHS
RVLRYRLTLEEEIKTG VVTAKNSDDFIAQLELHGNNMPDPSPMANYLSQQTMGDKTPADAMEALLGVCVQSVGIERSFKLL

Text_s1-1.txt

PHFGILPKTHNVLLKADKIENQRLKTHIDIREVDAFLKNYRRIESILGYKFKDRTYLLQALTHASYPTNRTITGSYQLE
FLGDAVLDLISMYIFEQNPMTSPGQLTDLRSALVNNVTLACILVRHGLHYLAESASFTDTVSKFVLFQEONKHEITD
QVNLLEVESDRKMAEFVDPKALGDVFESLVAAVFLDSGNDFAATWQVIYGMNGNEILTFTENTPIQIVRQLYEFKPSCK
PTFSRAIPDEDTVLRVYKIRNQQHEAYGFGQNKDDAKRAAAKALQVLRKHYSRAK

>Aaeg2

MAYHWTNDNIHTTALTPRDYQVELLASAKERNLILCLAHNSKFEIALKLIHELGSQLRKPAGKRKRRTIYISQNDVFNLI
RDLTDLKVINVNDLEDADEDYDWEQIVDDYQVIITDERKCLDAIICGYLDLNEVNLVIDDCHKVYGNEEISKLFIDYYN
VCREKPKILGLAGPLHNAGCIPGRLSAELEQLEYCLQAKAETASDIVTLRYCTKPKAILQCAPPANSNLASYLKEIVL
SQITFLEDHRYDPSEIYEDDEFLEELKNIPDPRADPLKFLHEFLAVLEEMGPWCADRAALAMIVQIEKQKIKTPYERHFL
LLCLVSTVVFVQIRSHCDQIFQQYSNEKEKIQAHSTPKVLRLLLEIFRFSPEGTRNKEKAISEIISGEPDTCQSESCQNL
LEEIKLIDFKKLIADVDVTKSIDSITESIANLKTSLKQISEPSIAEADVATQSKKLRSPKNGIRQLRQRRKPGMPNRP
YRSNYYNQNDPDALCALIFCNSKFIKILYSLFYESIRSDPQLAYINVQYTVDKTADPIKDPKAEIEHRKQEEVLKRFR
MHECNLLIGTSVLEEGIDLKPCNLVIRWNEPVSYSYVQCKGRARASTAYHILFVTPKSDAQNCANLCEDLSERMNHYIC
EKSIACLNTVNEDECEILDKVEETNPNSKQTCCEVMSNSNEFESSHKLNLVMENTNEMIEKAIYMEIEKLLKGCCE
NMEPPESNQYADCFSHYLEPYRPLPETDENSNVSNAIQMVNKYCAKLPDSTFTKLTPLWRGAKTVRNGQELYQCTI
RMPINSPLKEDIVGLPMPTEILARRMAAYTACRVLHKAGELDDSLQPIGKESFRAFEADWEDFELESNDQAQVIENSEPR
PGTTKRRQYKYKRIASAFSECRPTIGSKVYLYHINMTLQCPIDEEQNTGRKIIYPPEDSPQGFILTTKLIKPVSAFPIF
TRSGEVKVALNLCDERPILNQEQLERIHMFINYFTFKVLRLLQKYLMLYDPDAMENCFIVPTITRESVTVDFWDFIDTIA
KNVDKMPFTFIDENRKNYTFDSNQFKDAVMPWYRNQDQPYFYVAEICYHLSPKSSFPGLNYSTFEEYKYKYSIQIN
CKQLLDVDHTSARLNFLTPRYVNRKGVALPTSSEETKRAKRENLEQKQILPELCTIHPFPASLWRAAVCLPCILYRIN
ALLLADLIRKEVAHDLGLGTTDIADESFEWPILNFGWSLADVLKKSREAKLNAQSAVEMPPSETNKIEVLEEEITEKESD
KSAEEEEKKEGEIEKEEKANDLLEADKLLKAEGFQIGTWSNEMAETFEDEMMLDGDNDWGLPANVDLCSRNRNTNIRY
GSPTSWEVGTKEQSSDLRYTDSGYSVSDDDFDSGSGDSDSKNYGLKIEFKSDNVAEAIESADDILKRQKQDIIINST
NINEQLYQNSKNISSGFDCTKGSLEQNLKHEEQFERSTNEFKERIRSCGTLVKFDEPLTLENSLQKSNQAKKESNNQP
NIYELFPFASKETIDQYIISNTDGTKYVNLDGLYEMNKPYFPETYTVLGTGDIIDFRFNENLLTIDKNGQKTIVLNKINP
FAGESSPTPVNSSEKIDTLTVSTANLNLTDNPDFSFDYQPNLNNHPGSPSIIQLALTMSNANDGINLERLETIGD
SFLKYAITTYLYCTYENVHEGKLSHLRSKQVSNLNLRYLGRRKVLGESMIATKFEPHDNWLPPCYVPKLEQUALIDAKI
PACHWNLADLPDIKQLSSEEICRLVKERALALGLLDDSDNEEGETMKEIQLETRNDEFIKNAEYFACFIPYNLVTQHSI
PDKSVADCVEALIGAYLIECGPRGALLMAWLGIRVLPVYEVYDENNPKVPGSSQRYEKDGIAMQSVYGYWVPPKSPLL
MYAPNPQQTLEHLLDGYSAFEDSLGYQFTDRSYLLQAMTHASYCPNRLTDCYQRLLEFLGDAVLDYLITRHLIEDPRQHSP
GALTDLRSALVNNTIFASLAVRHDFHXYFRHLSPGLNDVIDRFVRIQQDNGHIISEEYLLISEDECDEAEDVEVPKALGD
VFSVAGAIFLDSNMSLDAVWVYRSMGPEIEQFSNSVPKSPIRELELEPETAKFGKPEKLADGRRVRVTVEVFGKTFRG
FRGIGRNYRIAKCTAAKCALRQLKMMGLINKRR

>Cpip1

MAHSWTDIAHTTSLTPRDFQVELLASARERNLILCLAHNSKFEIALKLIHELGFQLRSTAQRKRRTLYISNNDVYSLM
RNLTDLVVHVDGDEPDWERIVADYQVIANERKCLDAVVCYGLLELDEVNLLVIDECHKIYGNVEIAELFSEYERCRER
PKILGLAGPLHNAGCIPGRLSAELEQLERCLKAKAETASDIVTLRYCTKPKAILQCAPPQSEALALYREIVETQMTF
LQEHRYDPSEIYEDDEFLEELKHIPDPKVDPLNFLREFLTVLDEMGPWCADRAALAMIVQIEKRIKTPYERHFLLLCLV
STVVFVQIRSYCDLVFQQYATELERIEAHSTPKVLRMLLEIFRFRPDGANRGTVSSKAPIKEVTEESCQLPSCQSLLKEIKT
VDFQKFVDRVETVTSSIAIITEGLESKTSVKSISDAPPLKQNLVFNEKKPRSPFKNGIRQIRQRKKIGFTHRNRSNFHN
QNDPDALCALIFCNSKFTTKILYSLLYEAVRSDPQLSYINVQYTVDKSANPITEPKEAEVEHRKQEEVLKRFRMHECNLL
IGTSVLEEGIDLKPCNLVIRWNEPVSYSYVQCKGRARAPSAYHILIVTPKVDQLLAANGARDDLSEQVHRKVCERTSRV
RMEEEESDDDEEKRPTLSGSNLCLVKNSETIAEETKQLGLQVMENCTNEMIEKAIYMEIEKLLLRKCNEMPEPESEL
KHADCFSRFFEIYRPLPETEENSNGVGLANAIQINKYCAKLPDSTFTKLTPLWRCAKTTTRKGHDLYQYTIIMPINSPLK
EDILFHSRADGKGRVLGDLVDGVDSVRNKRDATAAGGIARNRFGVLLLQFHTELVVVAKKLSGCAAQGLSVKLHQAYDRF
RRGSFRYAGNGGGFNEAGHFEGLPMPTEILARRMAAYIACRVLHNAGEIDDSLHPGKESFRAFEADWEDFELESIDAQI
VSENSEPRPGTTKRRQYKYKRIASAFSECRPTVGCKVYLYHINMTLQCPIDEEQNTGRRIYPPEDSPQGFILTAKLIP
KVSAPPIFTRSGEVKVALNCAERTILNQEQLDKINLFINYFTFKVLRLLQKYLMLYDPMATENCFIVPTITRQSVTVVD
WDFIDIIAKSVDKMPTFIDENRKNYTFDNIQFRDAVMPWYRNQDQPYFYVAEICYHLSPKSSFPGLNYSTFEEYKYK
KYNIIQIQNCKQLLDVDHTSARLNFLTPRYVNRKGVALPTSSEETKRAKRENLEQKQILPELCTIHPFPASLWRAAVCL
PCILYRINALLLADLIRKEVAHDLGLGATDISDAQFEWPILNFGWSLADVLKKSREAKLTAKAVLEEVPELTEPEETDKK
DDPVEGKADGVQKSEDPLSDEKTANDLLEADAKLLKAEGFQIGTWSNEMAESVQSEFMDDGENWGLPGNVNICSNRNTNI
RYGSPTSWDVDNKEQPTDARYYSSEGSYGSDDDELDSGGEENDSDQGRYGLKIEFKSDNVAEAIESEDDIMKRQKQDII
NSTNMNEQLYQTSKNVSSGFDCTRNDELEINVKLHEEQFERSTNQFKEQIKNCGTLVKCDEKLTVENHLKQGRNAATVQS
DERDIYELFPYASKQEVDRFLFTNDRGVTYMNLGDFYDLNRPYFPEKYSIIIGTGDVDFRFDNENLLLEVNENNEKTITFNI
EYPLVVRNQTKDNQKRTLPSVNVATAINGSLGEFSFDYQPDLTNHPGSPSIIQLALTMSNANDGINLERLETIGDSFL
KYAITTYLYCTYENVHEGKLSHLRSKQVSNLNLRYLGRRKVLGESMIATKFEPHDNWLPPCYVPKLEQUALIDAKIPAC
HWNLADLPDIKQLSSEEICKLVKERAQLLGLLDDSDNEEGEHMKEIQLEIRNDEFIKNSDYFACFIPYNLVTQHSIPDK
SVADCVEALIGAYLIECGPRGALLMAWLGIRVLPVYEVYDENNPKVPGSSAPYEVNGKQIQSIYGYWIPRSPLLTHA
PNPEKSLEHLLDGYSAFEDSLGYRFQDRSYLLQAMTHASYCPNRLTDCYQRLLEFLGDAVLDYLITRHLIEDPRQHSPGAL
TDLRSALVNNTIFASLAVRHDFHXYFRHLSPGLNDVIDRFVRIQQDNGHIISEEYLLISEDECDEAEDVEVPKALGDVFE
SVAGAIFLDSNMSLDAVWVYRSMGPEIEQFSNSVPKSPIRELELEPETAKFGKPEKLADGRRVRVTVEVFGKTFRG
IGRNYRIAKCTAAKCALRQLKMMGLINKRR

>Cpip2

MQTIEETQKEDFVPRFYQSQMKDVCIKNTIVYLP TGAGKTHIALMAIRELGRRGHLEKPLSQGGKRTFFIVNTVALAKQ
QAEMIGRNVVFKTSVYTSDRDVTWKQDRWLEEFACYQIIVCTCQILLDVLKHGYLAMSHINLLVFDECHHGVDHPMHG
IMEQYLARKEDRPRVIGLSGMLLYKELKMKQVAQELERLENTFDSTIATVGSYDAYTEVCRFSTDPKEGLLSFHVVQS
SDVMRNLQGGINDFINKVALFDLPKLLNQNKSLMRDMPKPKKVVITKYFKELNYQFEDLGLFGGAIALLGLIVQFELDRE

Text_s1-1.txt

SDHTMLRLLYRSCITFCENLRHQLENVMOGLEIKKKLTLFSSIKARQLIAQLEGFYKEGRSRKTKTLIFVQRRFSAKVY
HLLKIYFSETENAETILPDFMVGCGMTMPESIEQILSAKKDRRLERFKKNETNVIIVTTNVLEEGIDLQMCNSVIKFDYP
ETFASYEQSKGRARMKDSTYTVMLDSDKREKFLKKYLLYKEIEEELRRALVGKTINRPEPLEADVTKELTNELIPPFYTK
KGAKLDALSSIQLLNRYCMGMPRDAFTNTNVTWERHDDKYTGKIVVAVLLPLQSTVRDKVFGQPMTNVKLAKRSAAFEAC
RKL YEAGELNDHLIPIDSKRQLANVSEVYFRHWKQFEESAKQAGTQKNVRNHQIRYPSQTSGCCPQPGKPCYIYVLRIA
AGFNSDVQENIETFHTLYSSENNFGIMTTKPLPVLARMKFFVSLGLINVHLDPTPIRVENAGSDADLAALKQFQLMLFR
DVLRLWKEFLVLDSSNEANSFLVVPLAQRSQIDWQVVKDFPFLAQPSELSTVARSRMVFDAKQYRHRVILPWYRTDRERA
YVTVAVHEHLTPGSPFPNEKYQTYEDYFGTVYGGQICNKKQFLIEVKGITTLNRLSPGDEDDGKSATRAKNWTFHEILI
PELCHNFAPADYWLKATLLPSALHRINFLLLAENIRVELATEANVGCLECRTVEDVDVEYKNRSTDLANGMDELKFDSD
DEDEEDDDDDFDYEEAKKALVGPEDLNSLMRIQMNSFAEEVEMPWKEQDEPVDIERNWDNVSALDLDYYDSFVRKFSSLN
VGEQITTKMAQSYSSVIHKQATRASPQKVAGAILDMSLDEKFRIALLKLTPENTVYVPLQQCAIICALTTKSAHDVYDLE
RYELLGDAFLKFSVSLFLVKKHKEWHEGHLTAVKGGQLVSNRNLLYCAMLINLPGMMKIHTFDPKNDWSPPLASVPHAVKQ
KMASINHSARVLYKLMSEEEIQAGTIEKDKYEDFIAELDVQSRSLDTSMPQNYLSQQAMGDKMPADSMEAMLGVCVSTV
GIKRSFRLLSFLGILPKKENLTTLLDDRIANRMLKADVTLNKNEVDSLISPGRIERILDYTFKDRTYLLQALHTASYVNR
ITGSYQQLLEFVGDAVLDLFLISAYIYEQNPTMSPGQLTDLRSALVNNITLACILIRHGLHLYILSESASLTDTVSKFVAFQ
EAHKHEITDQVNLSEESERCQGAMAEFVDVPKALGDVFESLIGAVFLDSGNDLEVTWRVIYGLMHHEIATFSVETPIQI
VRRLYEWKPPCYPKFSQAIVDDDTVLVKLRYRIRNVDEHAYGFGQNKEDAKRAAAKAAIQKLRKQ

>Spur

MDEPPSSPMGLEVIDVETPGGSSVTSQKEYTPILSFTPRHYQLELLETALEQNTTVCLQSGTEKTFILAVMLIKELSGSVR
LRELDGGKRTIFLVNSESYLEYGNVIHTHTDLNVKEYRNSDKMISWDRTRWQDEIEHCHVAVMDSEIFLKLQMSYVGL
SMLNLLILDECHHTLQHLHPYHQIMLAFKQYPRDDRPRILGLTESILKGFKPCALEEHIDMLETSLQSKAKIATNLTSI
SKYGTKPKAEAVIVCRKYEDCIELVSKLSKLLTDGLDQLNQFEYSFKARDFEGELVDCRDPVLPARQALMECLSTLRITIG
PLGVKILIPMLVREILKLEKHEFLDVHLLFLQHATTQLRMVERICDNALREGGNFYDLKFLTPKVRRLLEVLSGFKPKEV
SPEEKAQSTNRERANKFASRRKAHRSHKDIKNPRWRPIVPEDQDEDEVSDEEEEEEEDEEDVGLQQTGRSTTNLCGIIIF
VEQRYTAVVMSRLFKKLSKRDPNLHFLCSCIAMGSGRLGTTKSAIQVQNRKHEEVLKFRRREVNVLVATSSVEDG
VELPRCNGCNLVVRFDRPTSYSYMQSKAKARAPTSHYMLLIYEPDVQLFLGDVQCYQEFEREQEMLRNLNVVGLLSTRAQ
VCLLPVRFAKKAKKPKQKVKVLRNRSKSGKTQPQEEVLDLLVADKQMPAYIPKNEEGSPRASMTTIAIQIINRYCAKLPSDV
FTHLTASCKMKERKGDVSPEFQATLQLPINSPIRKPIITGPWMKRKKWATMAVAVKAATILHKAGELNDYLQPMGKESILK
ATDNFNLHAFADGQTRPGTTKKRQYYDKEVAKCLQECLPQADQPCYLYSVDMSLASPLPDILNIRERRLRPEETQQCFG
IITSKQIPKVPGFVYTRAGELAVSIMLESDEIRLSSSQIRRAQSFHGCMFSEVLRDLKPNLELNPEKSQANYLIVLKE
GPSGGPSVIDWSFLEKVSLSGHLDQPTQLPDYSQSQFEFSTDLADAVVTPYRNIDQPQRYFMADILYLPVTSPPFS
EKYETFVDYYFERYDIQISNFQQPLVDVDCMSSRLNLLTPRYLHKKGKALPISTGQNKKFKLQKQYLLVPELCYIYPIPA
SLWRKAVCLPSILYRLNALLIAEELRVQVAEEAGIGLKTLPQEYYPNLSFGWEGLDLEALEKQEKEAEAAAARCEEDKD
NAIHVEPKCNNNEQLKEMERAAELVVRTNKSEEHKVNEQLNSPNSKGEQHSGGDMSVNSTIESKEIFDDGNFHLDDDGG
DPQMLGPSPTIILQSLTMSNSDGFNLERLEMLGDSFLKQAVTAYLYCTYPHLDGKLSFLRSKQVSNFNLYCLGDKKA
LAHKMQVSLFDPSINWLPCLLVKESSRPVPEETETDPPDAGFVIDTWPSTVEGVDDFDDDDFSDQCWEENSDDFSDE
EIDETKAECPPDFAREEKDKQWFSVGDPPDDVPLTYTGAKTQDLGVPQLPYEIHQHSMSDKSIADCVEALIGCYL
VSCGFRSALLIMAWMGLDVLPTIDGQSNKRNDQSSDLPAADLPVSLYGLYKQPEPILLSVPNAEDVLQHLVGYDGF
KTINYRFRDRAYLLQAFTHSSYHRNSITDCYQRLFLGDAVLDLITRHLVDHHTNLSPGALTLRSALVNNTIFASLAV
KFNFKYFKAYTANMFQVIDNFVTFVDAKNEALGMDSQLKYSEMDGEQEESEDI EVPKALGDI FESVAGAIYLD SGMSLN
AVWKVYYPMMKQIEQYASDLPISPVRELLEMETARFGPPEKTIDGKTRVTVNVVKGKQFKGIGRNYHIAKATAARFA
LRHLKSQPKK

>Bf10

MAPAPAQHTSQPQTCHLQREPQHLSIFTPRLYQMELLDAALEKNTIVCLGTGSGKTFVAVMLIKELSSQTRAALSDGGKR
TFFLVNNVPLVSQQAIVITHTNL SIGEYVYGAMGVDLWTRDRWQEFDTHHVLVMTAQIFLDILQHGFLPLSKVNLIID
ECHHAVGHHPYREIMKTFDTCQVQDYPRVLGLTASILQGKCAPDSLQFRVRNLEVTLRSSAETATDLVGVDRYTTQPNEV
VIESGPAQNEPGLSQUALQTLVDDALAFISSCKVRVSVEEGERDPCQVPKQAIQECRAILEVLGTWATRRAAAILNRELEK
LIKHEWSEQHQLFLRWVHTVMHHVHLLCDEAFKDVEDVDVQFVTPKVRRLLEVLRECKPNEVQSDVQSNGLGVPEAPRNG
PRGYKGTNNARNGAPKSKSVSNMNSKNHYHSSRRNDLNDQDPNTSFQLCGIVFVERRYTAVVNLKLLQEFKASDPDLAY
ISSSCITGHGLSSRGMSRETEMAFRRQEEILRRFRMHENLLIGTSVVEEGVDVPCNLVVRFDLPKDYRSYVQSKGRA
RAQGSYVMLVQQHELDSEFKEDLVQFKGIEKILRQKWAHEVPKEEIEWEHLCDDMVPPYMPRQVDDGGPRVTMTSAIGLV
NRYCSKLPDVFTHLTPKSRVDSVSDADGNKYMATLYLPINSSLRDPIQGVPMPTRRLAEKAVALKACQLLHQAGELDDN
LLPVGKEVIHYEEEEEEFEVEDSDGHARPGTTKRHVYKQKIPKAFQQLPVPAASCYLYIIDMVLTEPLPEELNIRGRK
LHPPEATTRCFGLLTSKHVPQIPCFPVYTRSGEVSISLRLCASGINVNQNLRLVQNFHGYVFSHVLRLEKKPLYVVPAD
LQSWYFVPLNKGLLADNADQLFIDFNLEHIDQSLDKEKPKYSHDNPFKFEADDFIDAVVTPSYRNADQPQRFYVAEIC
YTLNPRSEFPASADYATFDEYLLKRYGEAITNLDQPLLDVDHTSSRLNLLTPRHLNQGKALPTSSAETRAKRENLSKQ
ILVPELCEIHPVPASLWRKAVCLPSVLRYVTTLLIAEELRVQIAQEAIGIGQAVLPEGYAHPRLFPHLKKVEEVFEPCEP
APARSEKSRITLEDKEEESWFEISAWNPDVDDDDMEDEFGDRWDISEEWFRAFQDNVMLLDTPLVPALDKKGGQNGA
KVNHTNRDLNLDLDSKASLDDDDFPLTVPGENLPAGDRLSKASMSNGQIWPEEGRKGEQPVHNGQVKPDRSIAPGNN
ATSGNGISGEGTSPLSNMPDLNSVPESTSFLPMLDLDLDSHPGPSPLILQALTLNAGDFNLERLETIGDSFLKHAIT
TFLYCTYSRVHEGKLSYMRSRQVSNLNLRYLGRKGLASRMVASIFDPAVNWLPPGFICIQNNREDKEDAGSSSKDSSL
VEPGNGMTTSKNPIAQVTGYDVTPLKASLVNGDATKRLLSYDLHMEHCIAADKSIADCVEALLCGYLTTCPRAAQLFL
CWLGVKVLPKTNADPTEDLDQDFAEPTDPFDDL SGLDSSYEDWDQNPSTVCDLDTFMKVS CAEEYGYIAPPKPLFHH
VAHAQDKLSHMLSGYQQFEETVQYFNFKAYMLQAFTHPSYHYNTITDCYQRLFLGDAVLDLITNHLFKDPQQHSPGA
LTDLRSALVNNTIFASLAVKYDYHKYLKFSPEL FNIISNFVQFQEQGEEQGMDSQLKTYINDEGEESEDI EVPKALG
DVFESVAGAIYLD SGMSLETVWRVYRMMKPLIDKFSKVPSPVRELEMEPETAKFSPSERTYDGVKRVTVTVIGKQ
FKGVGRNYIAKSAARKALRALKGAKI

>Cint

Text_s1-1.txt

MLSSGSSQEPGKPSWTSNSQCNSFIPRPYQIELLECAAKQNTIICLNGDFGKCFMYTMFLREVIKHNGLNNAAPKFAVL
LLKEDCMVNEQSELLRSHLDVDIISCTHNMDSFTDESLLQQLCHDHOVLVCTPEVFKLLSISLYIAQYMDLVIFFDCHL
VNIEMHPIRKIMKILDENRYTPMILGLTSSLNNTADELNKCIHSMETVMHCKAQTATDLIAVDRHNINPPVENIILS
SDENKDGLNSQLLSDIRKILNGALEWLNQCQKEKQEEEDLDVCTVAIQLINESLMVADVLPWCVYQLIVVVVKQLQKNLH
LINVHAVMLQYVQTKFKMVYSICHWYMGDQKAVNLAFAVAPKVHRLIQVLKEYKPVVEGEEKQNTSFKYNNRNESDYVSWW
QSDAEDEDADMMEGKKEEKEDENHPSLFINTLCGIVFVEERYVAMMLQRLLRELSKEDKNLKHVSSSHLTPGVEGERRG
VDPASMDQSKQEEVLRKFRSHESNLLICTGDLEDCAYPKCNLIMRFDVPKSYRSYAESKSRARAQTSAYVMLVRETEKS
HFMEKLQEYKKIEKVLGPDWSLNEVYDNNEDDDIEEMIPPYILNRNTQVNMKNAVIRVNRVCNRLPSDPFTHLSPRCEV
VEVLNDGDSMPMFCILQLPINSHVREPQVQIVMPTLQLAQAVAMETCKVLYEAGELDENLMPIGKETVKYAEELDWPWR
EVSSITGRPGSTKKRQTYNKDSPQILRSCLPQPDINLFIYEIEMKIKNPLLELNVRRRKLFFPEETSRCFGILTTPKPIP
QIPGFPIFTRSGEESAKLKFCRQGVQLSAQQIKLVRIFHRYVFSMDLRVDRVPLSFQPEKSTSQLLVVVLNKANSSQLNI
DWSFMMKTESKRRVIEKMKSDPWYKHKMNEQNKVVFVNRQLYEDGVVAKYRSCDQPNRFYVAELKDDMTPLSSFPSSSEY
ESFLHYQLKYGLDIRCLKPLLDVDHTSSRLNLLTPRHLNQGKALPMSSEKRAKYESLHNRQILVPELCEVHPIPA
SLWRKAVCLPSVLRYRNQLLAEELRKNIAFKSGITMEQPPIGNMWNQLHFGWTDNEDSDDDNVEISSENNEVLNENKN
VFDENIPKVPNSVENDGECAQHDNLEETNHNKQNHQVDFENSTAKSDVYKTTVINHMKENHDEDDYVHVHGSNDVNVHKK
INHNDNSDRIGDTAHDYEAHFDLSDHEDENMKNNVNDLSASIRGCKKKIKDVENPIWISSDVIDDYNKENPVKNLGPVGP
LILQALTLASNADGFNLERLEMLGDSFLKHAVTVYLYCTYPESHEGKLSYMRSKKVSNNLYRIGKQFGLAGKMTGLIFD
PTVNWLPFGFVVGASGGQDSDDEDISDDVEIVDPPPSLPPSTINDEEEIYEDEIRMIIDGMLGGCIRTGGSPKGDNGL
FGLGKKISKNIHALSSHQSMHFTDKAALRRSQLNQNLDFVIDTWTVDASGKPLNNSDPLMLDLHTQHSADKSIAD
CVELLGAYLTCGHRSAQLLLCHLGLNIFYAFKETAIVCKTTPFQNTKISMINHAFTKVLPKVAEPTSKTVDRFGYL
TPPTNKMLECRPNQNKLDQLLVALSGFERKINYTFKNKAYLLQAFTHASYYNTVTDYQRLEFLGDAVLDFLITRHLF
NDERHHSPGALTDLRSALVNNTIFASLAVHYDFHXYFKSISPELHHIIEDFVMYQKQEDAQGMDAHLMRKDEEEGENED
VEVPKALGDI FESVAGAIYMDSGMDLSSVWKVYSVMMLPLIEKFSACVPRSPVRELLEMEPETAKFSCAEKRYDGKVRVT
VEIIGKGRFKGVGRSYRIAKSAAARRALRYLK

>Skow

MEEYDNMMETEAEERPQDMDIQIETSEDLQLFHQKTIHNFTPRAYVELLERALDKNTVVCLQTGSGKTFIAVMLIKELAH
QIRTSYDNGDGGKMTFFLVNSVPLVSQQANVIRTHDTLTVGEYVDAMDVDLWKKDKWNQEFKSHQVLVMTQCIFLDLLH
GMISLSSVNLMI FDECHHAIGNHPYREIMKIYDVCPKTSYPRILGLTSSILNGKCDPTRLEKKGLEMTMRGTAETATD
STLVGQYGARPKVVLICQAFQDENKLAEEINTILQARMFLEDCSFTLEDNERDPRVIPKSVLLEQSTLSVLGPWCTN
LACQVMIRELEKVEFHSTLGLSKLFI RFAMTTLRFIYMRCTEVFGSTQRVNLKYVTPKVLKLEILKQFKPQPEENPPVE
KQDCVEKNGILGNDSTADTPTSDTNMELPPNNVLVDNKLASCNAVNSVEVSAVEDNSSVSAPAVNNSRITQENNTSGRK
QQHKHKLHNYGDQHIQIGTGLSSLCGIVFVQRYTAVVLNQLLKDLSKQDPDLAYLSTSYITGRSIXKSTL SVNPDMEYKQ
EEVLRFRFRRHETNLVAVTSVVEEGVDVPCNLVIRFDSMREYRSYVQSKGRARAPESFFIMMVYDEEKEAFIKDLKQYKF
IDRILERICHGKESPAEEDIDAHMVDLLPQYTPNKQEGGARVTMSSSIPLINRYCAKLP SDAFTHLTPKCKTVQLPNTD
PHMYQSILYLPINSPLRQPVQPPMKGKKAEMAVALKTCELLHKAGELDENLLPVGKEVVPYEDRDMFEEEMDADGPRP
GTTKRRQYYCKQVADSLQNLPGINKPCYLVIKMKLALPLPDILNTRGRKLYCPEDSEQGFILISKKIPGIPSPFVYT
RSGKVTVMELASSGLRIPKDKVIQIRKFHKCIFSEVLRL EKPSMKFDPAKSLSKYYIVPLLSIDIDLYGAEIQEIDWSFL
ELISQCSGDLDNPKPVKYTNAPYTFDPAKYEDAVVTPTRYNFDQPQRYVVAEISELSPISKFPSTLYPTFAEYFQRYA
LEVTTASQHLLDVDYMSSRLNLLTPRYLNHGKALPVTSAQTKKEKQENLQNKQFLVPELCYIYPIPGSLWRKAVCLPSM
LYRLNGLL IAVEVRDLVAQEAQVIEKLSDDFKFDDL SFGWDFEINKQAVGDIRDAAGDSDEIESTGTSDDDDDEFYDF
SDAIEFQVGDETLQALNQLKMLPDGEEP IKT FNNIDHDSIENDSGWDDSVVDVQHLISFGTSSRVAPVTIPAPAIDAKD
DELDIDIFAQMDTDFDKAATGDFSKSLKHKAAATEKISSERLGTIYENDYTEPLIPDPKLDSDTNDDELTTMPGNPALIL
QAL TMSNADGFNLERLEMLGDSFLKQAVTTLYCAYPGMHEGKLSYMRSKQVSNLNLRYLKGKHLPSRMVSLFDPPV
NWLPPGYVVLDPPEKRQVPRVESGFEIGYWGNCDTVENFKAEPDVPYDLHSEQSLSDKSVADCVESLIGCYLVSCGFRGTL
LFMSWLGIKVLPSLKDEKNATSNSVGIENGIQSTGNNTTEKMSREVTENDQPNVSVGYGHLKPPNSPLFINVAHAEEERL
HLLSGFESFEERIHAFNDKAYLLQAFTHASYYHNTVTDYQRLEFLGDAILDYLLITKHYLDHHQHHS PGALTDLRSALV
NNTIFASLAVKYDYHKYFKAIAPELFHVIDNFVKFHATEETQGIDSQLNSNFSDDEDEEEDGEEREDIEVPKALGDVFE
SVAGAIYLDNMSLDAVWRVYTMMPKHIDKYSKVPISPVRELLEMEPETARFGLPEKTMGKTRVTVTVAGKGTYRGI
GRNYRIAKSAAARRALRALKAQKEAEKARAEAEVEALKSLKDEEAKDEEREAEEDDEEESPOK

>Human

MKSPALQPLSMAGLQLMTPASSPMGPFGLPWQEA IHDNIYTPRKYQVELLEAALDHNTIVCLNTGSGKTFIAVLLTKE
LSYQIRGDFSRNGKRTVFLVNSANQVAQVSAVRTHSDLKVGESNLVNASWTKERWNQEFKHKQVLIMTCYVALNVLK
NGYLSLSDINLLVFDECHLAILDHPYREIMKLCENCPCPRILGLTASILNGKCDPEELEEKIQKLEKILKSNAAETATDL
VVLDRYTSQPCEIVVDCGPFTRDRLYERLLMELEALNFINDCNISVHSKERDSTLISKQILSDCRAVLVLPWCADK
VAGMMVRELQYIKHEQEELHRKFLFTDFTLRKIHALCEEHFSPASLDLKFVTPKVIKLEILRKYKPYERQQFESVW
YNNRNQDNYSWSDSEDDDEEIEEKEKPETNFSPFTNLCGII FVERRYAVVLNRLIKEAGKQDPDELAYISSNFIT
GHGIGKNQPRNKQMEAEFRKQEEVLRKFRAHETNLLIATSIVEEGVDIPKCNLVVRFDLPTHEYRSYVQSKGRARAPISNY
IMLADTDKIKSFEEDLKTYKAIKILRNKCSKSVDTGETDIDPVMDDDDVFPYVLRPDDGGPRVTINTAIGHINRYCAR
LPSDPFTHLAPKCRTRLPDGTFFYSTLYLPINSPLRASIVGPPMSCVRALERVVALICCEKLHKIGELDHLMPVGKETV
KYEELDLHDEEETSVPGRPGSTKRRQCYKAIPELRDSYPRPDQPCYLYVIGMVLTPPELDFNRRRKLYPEDTTR
CFGILTAKPIPIPHFPVYTRSGEVTISIELKSGFMLSQMLELITRLHQYIFSHILRLEKPAEFKPTDADSAYCVLP
LNVVNDSSSTLDIDFKFMEIEKSEARIGIPSTKYTKETPFVFKLEDYQDAVIIPRYRNFQPHRFYVADVTDLTPLSKF
PSPEYETFAEYKTKYNLDL TNLNQPLLDVDHTSSRLNLLTPRHLNQGKALPLSSAEKRKAKWESLQNKQILVPELCAI
HPIPASLWRKAVCLPSILYRLHCLLTAELRAQTASDAGVGRSLPADFRYPNLDFGWKKSIDSKSFISISNSSAENDN
YCKHSTIVPENAAHQGANRTSSLENHDQMSVNCRTLLSESPGKLHVEANSADLTAINGLSYNQNLANGSYDLANRDFCQGN
QLNYYKQEIIPVQPTTSYRIQNLYSYENQPQSDLECTLLSNKYLDGNANKSTSDGSPVMAVMPGTTDIQVLKGRMDCQS
PSIGYSSRTLGNPGLILQALTLASNADGFNLERLEMLGDSFLKHAIITTYL FCTYPDAHEGRLSYMRSKKVSNNLYRLG
KKKGLPSRMVVSIFDPPVNWLPVGYVNVQDKSNTDKWEKDEMTKDCMLANGKLEDEYEEDEEEESLMWRAPKEEADYED

Text_s1-1.txt

DFLEYDQEHIRFIDNMLMSGAFVKKISLSPFSTTDSAYEWKMPKKSSLGSMFSSDFEDFDYSSWDAMCYLDPSKAVEE
DDFVVGFWNPSEENCGVDTGKQISYDLHTEQCIADKSIADCVEALLGCYLTSCGERAAQLFLCSLGLKVLVPIKRTDRE
KALCPTRENFNSQQKNLSVSCAAASVASSRSSVLKDSEYGCLKIPPRCMFDHPDADKTLNHLISGFENFEKKINRFFKNK
AYLLQAFTHASYHYNTITDCYQRLEFLGDAILDYELITKHLIEDPRQHSPGVLTDLRSALVNNTIFASLAVKYDYHKYFKA
VPELFHVIDDVQFQLEKNEMQGMSELRRSEDEEKEEDIEVPKAMGDIFESLAGAIYMDSGMSLETWQVYYPMMRP
LIEKFSANVPRSPVRELLEMEPETAKFSPAERTYDGKVR