

mtDNA-encoded:

>gi|11467048|ref|NP_042524.1| ORF142 [Acanthamoeba castellanii], 142 aa, ABB .
MKIDKNWIVGVFDGEGCFYIGINKSVDSKLG YQVLP EFRV VQHKRDIKVL
YAIKDFFGHGSVVCNKSNGSEIYEYRVRKFETLHDVILPFFESNGLLTSK
KFNFLAFRDVILIMKRREHLTESGLSKIIDIKSRMNRSSIHY

>gi|11467049|ref|NP_042525.1| ORF168 [Acanthamoeba castellanii], 168 aa, 1496 .
MKNL DLKQEKALIDPFWISGFVDGEGCF C ISFNLKERLTLGIEVRPSFSI
SQTRDKEGLNLRCLQDFL NFFDCGFIRFSKRDN TWKYECRDLSDIRSKVL
PHFEKFTLR TKKLRDFELFKD VVNSVASKQHLNEVGLKRIIDISYQINMG
KRKLTKDELLSKINLKNL

>gi|11467050|ref|NP_042526.1| ORF164 [Acanthamoeba castellanii], 164 aa, 599 .
MQKNQFKKL NNEQLAYLAGFVEADGCFLVQIIPGLQYRYKHTIRISIVFY
QKKDKHWYFLQLKNLIGLGSIRFRNDGMLEYSITGLSLVNKFLEMLFPYL
ILKKNLAVLIFRIIKGLNDVKNEAGFLEVCKLVDEVDHTYSKRRKNTSL
TVKNSLLLPVETEE

>gi|11467068|ref|NP_042544.1| ribosomal protein S12 [Acanthamoeba castellanii], 127 aa, 67C .
MPTYHQILKNPRRKKMHTNNVKALEGCPQKKGV CVKLRI VPKPKPNSAQR
KVAKLRLSTRRMIIAYIPGQGHNLQ EYSSVLVSGGRAPDLPGVRYTLIKG
KYDFSWKESFERKKKLSKYGYSTKKKF

>gi|11467078|ref|NP_042554.1| ribosomal protein L6 [Acanthamoeba castellanii], 181 aa, 865 .
MELLKVIKSQDKITLKT EKVKNVFFYFKNKKKILLKITFSGELFLKERT
LITLNNLNYSNYSNYIKFMAKNLETFYDQYNYFSKLHMI GLGFKNFILR
KHLIILVGCNYYIFRIPDSLKIFCKKNQV F ILGESNVEIFNFMSNIKRV
KKS NFYKGGV LQFKNFKFTKLKVGKKQRFM

>gi|11467062|ref|NP_042538.1| NADH dehydrogenase, subunit 4L [Acanthamoeba castellanii], 103
aa, 1873 .
MTSFNIFLLFFSFVIFFLGILGIFITRKNIIIIILVSI ELMLLAVNFNFA
IFSVLLEDMFGQVFILYTLTLAGAEAAI GLAILIIFYRIRGIISVNFVTS
LKG

>gi|11467058|ref|NP_042534.1| NADH dehydrogenase, subunit 2 [Acanthamoeba castellanii], 527
aa, F14 .
MNVINFIYSIFDKKLFKCGT ISSRGLFYNDL NFFHFEIYFLLALITFLIF
FVILSNKKNYRSYYYLN VGGIFSNLLPFV VILLIILNTNTNDSYYL FAG
FYYNDAAVVFFKNIILIGFLIFTFAIKQYLSYFKYDFEFILVLFISLFS
SLLILNSNDLISLFFIIE LQSLTFYILVASKQTSSFSSTESGLKYFILGCF
SSGIILFGISLIYGFTGLLSYTDLTLFLSEVYVTNFILDSSFFSFSGFLI
GFLLLTVGFLFKLGSAPFHMWMPDVYEGSP LLITAYLSTLPKISLIFVIF
KLYYYVFFEVFLFSQGLFTLTALFS ILLGSI AAIYQVKLKRMLTYSMITN
TGYYLLGLSFGDISGIYITIFYLISYIFIMIGLFFCFLSLRDRSSG LLVK
RINLFSNLLEVNPSLSFSIFILLFSIAGIPPLLGFYSKFFLFLFSLKYKM
YWMTILFVIFS VSVFYIRLVKLMYFNRTTGWIFLYDIPFLNSLIISFI
TIINCLFFFNP NLLFKLIYNFSFYLYI

>gi|11467063|ref|NP_042539.1| ORF83 [Acanthamoeba castellanii], 83 aa, 174F .
MFKFVRDFFENKQSFNIKTFAIWLNSSIKNRAIFVYSMATFSILVRYVYI
PFTTIIGYLILLRFTRHFFPTVVLQFAYSFLVF

>gi|11467064|ref|NP_042540.1| ORF115 [Acanthamoeba castellanii], 115 aa, BE1 .
MNGAGNVI SESTTG SGLGKKILDKASRVASNSFN YVTD SKNTFTIIRGLV
LAGGVIASADYASAEGSHRTTYVTRA FEVHENGTCSPDPDTRLKAAALRR
RGLIYLIVVNQKPAV

>gi|11467087|ref|NP_042563.1| ORF349 [Acanthamoeba castellanii], 349 aa, 1883 .
MISPIILISVPVITII SLATAYFFAPDIDPAL TLLYPDIPVRRIWPMVFD
ENRDELSPYELRCFLSNVHRTELGHFTCNVKYPILAINSFDDVALFTKNY
VAACGKSINVNPEGQFSFSWPTTYHNDKNSVHVFSFTGIREILTEGDLTE
EDQIVVDWFNSWAKDNPWWTYL FVSPDDPGNSD GSPSCIKITYFTHQVYT
PEPGGVGCDGLVYDTRRLRAQYEPVMYSTADVLNQP PFTVSK EGLFPRL
CLIELLKS AIPDYHIGVIHSYSTFYTTLVHPVNNLPPQDFISVHLEGAGS
YKDFNPGPLSSAGNPADSISSVFPKILIFVACATTLALCTYLGYNYIG

>ISD11 Contig7577 NOT DETECTED, 83 aa, 53C, 83 aa, 53C .
MASRPQALRVYRQLMRNGROYNDYNMREYIQRRVRDEFKANKTMSDAEQI

DKCLDKARHELEVVKRQAIISRLYNHGKLIIEV

>Isul Contig20119 NOT DETECTED, 161 aa, F34, 161 aa, F34 .

MLRTRTRAVFGSATLASHNGGSRMAGGAQRLYHERVIDHYENPRNVGSLD
KNATDVGTVGLVGPACGDVMKQLQIEVKDQKIVNAVFKTFGCGSAIASSSL
ATEWVKGKSLDEAKVIKNSQIAKHLALPPVKLHCSMLAEDAIAKSAISDYE
KKQKAKQEAQQ

>Isal homolog (Iscal) Contig2426 NOT DETECTED, 130 aa, 4C1, 130 aa, 4C1 .

MSVAGTVRVVRKPLTSRFAAKAPITLTPAAVSRLLHSLMGGSKDKYEGIRL
DVKSRGCGGNSFTLDYAEKKERTDEVVEADAGVKVVFVGSKALLHVIGTEM
DYVEDDLQSGFVFNPNQAKSTCGCGESFTL

>Arh1 (Ferredoxin reductase) (sequence may contain mistakes ... N-terminus uncertain) NOT
DETECTED, 507 aa, 31B, 507 aa, 31B c, 507 aa, 31B .

MKVWRRGIRTSMALREVGGMGVPTSRHLKVCVVGSGPGGFYAVEALQRM
AEEAKAKGQAPEMALDVDIYERQAVPYGLVRFVGVAPDHPEVKAVQHRYDA
FMADPRLRFVGNVSIQSSSSGAAALSVEELLGAYDGVVLAAGA
QRDRPLGIEGEDLPGVYSAREFVAVYNGDPAMRSRQFELSSSREAVVIGQ
GNVALDVARVLLHEPGQLASTDMPEYAVDALRRSQIDTVHIVGRRGPAQA
SFTNKEVREILNLPVQCHFLPSGVLALNEASKAETAKERGKCRMIDLFA
ATEKKAGETKGEPPKKNLYFHLLSPLRFLANEKGHVSGIELARNELQGEA
GNQKAVPTHEKETLECDVVFKSIGYKSEPLEGVPFDHHRGVVPNEHGRVV
RPDGSTVPGLYVSGWMKRGPSGVIQGNRMDAETVASMWADLAATRPGEK
PGWRAIQDKLRERQVAVVDWKQWKAIEDEEHRRGQVLGKEREKFTSVPEM
LDVAFKQ

>mitoferrin (likely complete) NOT DETECTED Contig13335 (and 5394,17393), 310 aa, E4A .

MEFLVDEALEEELQOSSDPVHLIAGACAGLMEHCGMFPIDTVKTHQQLA
GSRTSIAATIRTIVAKNGVTGLFRGLPVVVAGSAPVHGVAFSIYEFCKRL
LGADQPGHLLASSMSGVVATLAHDACLAPVDTLKQRLQFSARPYRGVWD
CFGHILKSEGVSGFYRGYTTAAVMNLPHASIYYGAYESIKKLLKRATGKE
YESNDPVTHMLAGAAGGCLAGGLTNPLDVGKTRLQVGTDAGKSYRGMVST
LRTIYREDGWAGFTKIRPRMVFHSMASAAISWTTYEYIKHTLEPSFSSHS
PPSSNKSGPS

>g1045.t1 - pentatricopeptide (PPR) repeat-containing protein, ---- NOT DETECTED --- 441 aa,
11C, 441 aa, 11C .

MITHHSQGGGLMRGIEAKQRLVWRALYVRPGPLAQLISFAQRSGDPAF
LSQSLAALPKVRMGNTIGYASLVSTLSATNNLDEMFFVYARMKEERVPLN
IFSAGALINASLRHNKLERAFEVLHDMKLYNLRIDNGLVNSFVTFLEKR
DVDRLLSVLDALGDRSDLSIDRVSIIDELWQRLSSPHEDSSLAYRFIERL
SAAVPPASDGTQSTASGQSQQHQOQGETKEGENEGGQEGEGDGRPOE
SEADKWARVGVYWHGRMDQLLNQLVTRGQLARAHDLTELMVSRWRLQPHK
RIITVFIRACITREAAINARGGVISGDDQRRPVQLLDEVGRMVVEGRLEA
DLITFTALIDANARLRRKATALRFFELMQORGKADTTTYRVLLIMCSVL
RDKKEGLRLLDQAEREGVLEMALFTDVVRTIRALPDGATP

>g10450.t1 - pentatricopeptide (PPR) repeat-containing protein, ---- NOT DETECTED --- 584 aa,
1BF5, 584 aa, 1BF5 .

MSAGAVASSSSSAPASRQPAATAAHYREVQLPDELASAVFPSSPAFMTD
DERRAADQAERELYLDRLDRRLKGIARRREKDRRLASRLQFAGGGDAIGA
MMAGKARSADQWAVHGLNRHWEAEAEERMRREREAGILAAFREEHAEA
MKLASSSSSSASPLTAAADADDVVPILRYWRYEEGTIGGGGGDEANDDD
ETGGHRVHHQTDETRDEGDDDDDEEDRAQQRVEGSVGNPLLWDDDDGDDGG
GREPGLTERRRKGTKQEEEDGPCASVDAQFLHHTKNIAAKTPGESWRAF
LVYLASLRRLDLAFALLDDLEHHDQATPSVYLSFLSACAQVVHQEADMRA
QEIIISRMQARGLKPSASHYTALLSAATSPTRITAAALQAMTDAGVEPTPVT
YIAYFKACSRFAKVHGHGPTLHGYPHGFQKQRFGTSTLEKRTAMHLFGAVVT
SLVQFISPTTEQEVMMVLRDMGKMLKLNHFMNSLIYMYLHNDQOEKAY
ELLNRTKAGTDIEPNEFIYTTFLRYHHRREYDHCKKLLARMRHDRIPPT
HAAMSTFKLLAEQLQDPTIWAANKELNNATKPRR

>g10770.t1 - pentatricopeptide (PPR) repeat-containing protein, ---- NOT DETECTED --- 636 aa,
B09, 636 aa, B09 .

MTTEVSIKTNTAAPYVHDERPSTNTRRGERNPNHRRRRTTASAQHVASAS
VSARSTDSASVEQLKETFNRLRERLSADAGGSGKQQPWLLWDDVVAAIKQL

GGAMQRLKVRSGKLGGEIELDVDFVPLVWRLEIDQLHRLNGLMSALSTLPG
AGRVMQDIIENLYAQNSPVKPDVRYVYKGLIARYARSRQTEQAFALDLMR
AHGLHPNLRVITALLPVCTMATVEPVFDQIRRHDLRPDRVFLVGTLELYA
RSHATDKLLELLRVMPEFKVAPDVHIFNVALQGCVNAEDLAAGWRVREM
MREFKVAPDRASVVQLIALCSQKQLRWDOQRQAWGRAERQRPHYPYARPD
QQELMDRWLGDVDFADMRHRIVPDQITFTALMGFNSPDIERAWGLFQIML
DAGVEANEVAYNVLLKRCRDLAECDRGLALWETMRASGVPPSATTFALLL
QLCGRMPDAVEGAARAQRVWDSMVHEHGRQPDHGHLGALLYACGRTGDVD
AALALLDDVKLRHGVAPIEDTFTSLLDGLTAHNNNNNSNDNDDDDNEVDGD
GESEERDRQYEACLRVMRRAEAEGVTPRMQVKYTFKMSMRRDRFDEAEW
LLGCMRRKGLPDAAHQSFAASLASAKAAASVGLVTM

>g11211.t1 - pentatricopeptide repeat-containing protein, putative, ---- NOT DETECTED --- 794
aa, 1AD8, 794 aa, 1AD8 .

MPAVASILRRMQSLHDQGDVGVLDALKEELSRAGQCHPFICSALIALL
GKLQRTDAALDLVTRMTQQGYVPDVVLTALVSALGRAGEVGAQYFFNE
IQRLGMTPTLHSYNSMIMAFARAGRIDKVRDMLEVMKHKGIEPDVRTYTE
IINNSCKAGRVAEAYQTFIEMQDVGVPDLKVYSALIAGFSHVGDHLKAQ
QLKEEMEALGLEASMITYSGLMSGLARNRQLTEMERVIEMKKRNMV
RALGDLINRMSDTPMDQAFNMAQALVQKKVEANADVWHALLRVCAKADD
VEKGIRILDVMRSHSPELVSASSYCVLIDTFKAGRVSRGATLLAHMRKH
KIKPDVSIYNVFIGAMRETGYSKALQVFELMKAEGVRPSHTTFSILIDA
AGHAGRPEEAEAHFAAMGTSPRPDTNNYNSLIEALARNGQLDKAERVLDS
LLRHTHTTDRNGSTDDHVGLARPTAKSFSSLLGFCLERKDVVRARRLAEK
MRRASVAPDAIAQRQLDQIPAEQHSHHHHQHPQRPNTADMYRELSTILN
WPEPHFPTGKFIILIRDSLEADGAVLLPHFVFGYKANKRALLVGLQSF
HYSGVGRKLACNLKAAQTKQFMYINALSTPYDWSAPSTSSAPAVAATPG
LDSWSLGENEDAAFKSLYHKVKAFVESSEAMDEVCIVIDNLSVLLLETQOR
PALVLDLVHYLQAFIRRTENRASLVVLVHGDVEDDQKLVKALEHRADIIM
SASGLPSGYSKEVNGQLLVHNGVQPQSKLMHFKSHDNTVRFMSK

>g11333.t1 - pentatricopeptide (PPR) repeat-containing protein ---- NOT DETECTED ---, 558 aa,
3C1 .

MKQSNRIREAASIFDQHPTPFSASALISACGRRRDLKQALGVYAKFIASG
HQPNAVLIIGSLVSACRRCGEPGRALRLDDLDLDRFDIAIDESSFHMLAAAS
AEAGDAAAARLLVRLQEGRLPFKANSVDCGQLIKGLVTGVPDPSAGASS
RIADALALLELMEEQAIQNSCHFTPVLAACAATGEAAIGRQLHTRIVGS
QSQAVVNDYTTAALISMYAKCGFVDEAAGVFADMRRLODGRCELRVGSW
SAMIQALGQHGRGKEALALFDEMVTDRHATPDGAVANTVLSACSHAGLVE
EALRVYSGLVAGRYGAIAPDIQHQTCCVVDALGRAGRLDEAEQFIDTISR
SEVTWRTLLGACRIHGDVARAERAVARLRKLPDAPTRVLMANVYAAAG
HWEERDLERKNMDKAGVRKRPGKSWLVVDGTRHVFTVEDKRHPRIEIRA
ELAVLWQRMKNAGFVPNTSVVLRSMADDEEAKCHLCHHSEKLAITFGLMS
TPPGTTIRVFNLRVCPDCHEATKFAHLTGRDIVVRDANRFHHFHGSGQ
HCSCGDYW

>g11349.t1 - pentatricopeptide repeat-containing protein, putative,---- NOT DETECTED --- 558
aa, 79D, 558 aa, 79D .

MLHPHAGQARPAAVLRRQRYPGLDQWAARFRKLTREKTLPIILLEMLANA
PPGPTGAKYEMLIEQCFEEKRPDLATQLYDHLRAETTTTQAAAPVPLAVF
DVLLQGYGRAGDRQRLAQVQDRMREVGVRPGMSTYTTILIDAWGKAGDLRR
MWQAWEDMREQGLRPNVVVYTSVIATLGLKGDVAAMERTFAEMQRSSGDG
VEPNRTTYNAMVHSYGOQMMDKMEALVERMRADPGLGLDNFTHSAVVAA
WSRAGRVDKALEAFDSIAATGGQPSLHAWTAILHMLGTAGRSDEMLRTLD
KMKRLGVKVPSTAVYNTIINAFGKARNIHSMMDTFKAMRRDGVAAADVKTYN
TLIDTWAKTGKAVEAEKFFVLMKREGLQPTMYTIASLMDAYTRADQFEKV
LRLISRKKEGRAPDNVFNLLIDTYGRMGKPEEAKEVLCGAMKEYGIAL
ETNNFTSVIEAWARNGLDKAAEELHRMNTDYATKPDIKTLTLLSFCIP
AADVKRGRRVIHMIHQKLOKDKQARETIRQFHAATKRPDQEKKNQPTSP
STTSPDRS

>g13583.t1 - pentatricopeptide repeat-containing protein, putative ---- NOT DETECTED --- is
this continuous with 13581???????, 55, 553 aa, E21 .

ESKVKRRAILASGLPFARYAEELTALEQLPSVPANSLSTIISEYAAAGKH

EMILARLRPLVPRLRGGAGAFGLI IKALQDVRGSVEVAFVAVMHQAATRPA
RSKTDKATAPDARAYHCLIGVCARQGDYTDALRVGQLMDDNRVAFTSGTY
DCLINACARSLTDKLDLSDALVGVDSVAGANNVDTGARRHPGLDAAAFALLDD
MRQQRGLEPDRRIYCSLMRVCLSYGDVARMRRTFFDDMRQRAGIAPDIYAF
NFLMQGCAVERDLGAAMRVLADMKLARVAPVTITYNMILNLCAETAAAAA
SESADSVDSAGTQSVDSLAAAAEQVLARMEQSRLAPPDQVTFNTLLKACAN
AAASRSAAGAAAVGGFVEALDKATAYVAAMKARDIDPDVYTYNTLLQVCS
HCPRPWKRRTTVDDADTGDDAAAAEHGQAAAAARVDKVVVEVMREMNERGV
RPDGATVTLILSHFGAREDPARAFVLSDMKRANIKPYLPSFTALITACR
HYSDPAMAERSTNDLNSAARTTNCGTSAAVVGVEASSIPRSCSLCVDLK
CSD

>g15080.t1 - pentatricopeptide (PPR) repeat-containing protein,---- NOT DETECTED --- 323 aa,
E0B, 323 aa, E0B .

NNDVDRARGVLRVMEQTISNSVDSRNSLMAARNAVNDPEGALRVLGEMGG
SEPGNGNVWTYVEAAESYLRLGDRDRALALIDTIADQVRQEGSKPGGRPR
RPSARESENESESEDEDEVEEERRATTVERQGELDLWNQVLARFVQNGR
TAECFEVWRRMREGEGGLAPDNRTARQLIRACVQAKNMARAHQVATNMLKL
HKLKLDLSDAGWNEIIEGYAAEQLETMWKYVEKMYREEGTYPGATTWVVE
RTCKSHGVWDQWRARVARKQAIDSDKQKKAQERTMKRNQTKFTTRRQTT
NKKTAVKKKAKENQKVVQDNQOK

>g15695.t1 - pentatricopeptide (PPR) repeat-containing protein,---- NOT DETECTED --- 163 aa,
23C5, 163 aa, 23C5 .

MATDPVAAYHQVLRVAVDHMKQAGVHPHATFSVVIDAAGFAGLPHEAERQ
FDAMTSLHGLHPQANNFTSLIEALARNRGLDKAEAVLDSLAAARQPTSALK
HDRLTAKSFNTLLGFCLGAGDSARGQORVAAMKHQAGVEPTEVTQARLRDL
AKLCERRGGHTGG

>g15807.t1 - Pentatricopeptide repeat [Medicago truncatula],---- NOT DETECTED --- 242 aa, 32B,
242 aa, 32B .

LVLPKALLGAGSRTLQRGGLGSPGARRSYASTSGAAAAGRTGEEEEATTT
PTRKRSGGGGKQOREAEKVITDITPHLAELDAEKRQKGGVDPKHVAAVLR
MVWEGREHVSEKGREGERLLRHLKWL RAGGDDGEGGVEGYTDLISALAKF
GLVDDCFVAVFDRMKGSGVQPNWYTYTALISACVPQRDRTRAMRVFEEMRE
QREDLKPDAVTYTLLLLKANEGRTVVEVKEAMQLLTKLKKSGV

>g1661.t1 - pentatricopeptide (PPR) repeat-containing protein,---- NOT DETECTED --- 534 aa,
DE1, 534 aa, DE1 .

MKQQGLEPDMISIYALIRTCARHGDVSAALKVMAMVEADGLPYTLPLLT
LANVCSTQPFPAKKGFAVLNLMQOHGIAPDLAAFNALLKCCAIVDDIAGM
QQIFQLMKEQDLSSEFTMLSLIVGQLANDNLQAARDIVSSFRILGVTPS
VKTYHKLFSVCKRTHDLDAALNLWDQMRREEQVEPDLPIFNALMDIYHRG
DLEQVHQLLDAMKAQGLEPTPTEVACLAVMYHRDGRVTEAGYQLVEDMR
REGKVAPHLATFSPFIDILLKADKIDEAWDVRNMMARWDIQPKVLRNT
FLFALAAKRDVDKIDELFRHMEAHPREFFPHCDLIVLLEHHLARDPRSP
RVAEVLETQLRLNVAHKLSRQVFVEPALQVIDGLIDQNSFDRARKLLRYL
DPNKLETYTRLFATCRRPEHWPMLEEFWDMLMEPRRRIQPDGTVALAYLE
ALCRLGRTHDAVHFLHNMVHKQSFDPLEHYNLVLRHASSNDKWNFVRKVL
KLMRYDRVTGDNVTAAIRERLHEVDGSLEEESAE

>g5356.t1 - pentatricopeptide repeat-containing protein, putative,---- NOT DETECTED --- 161
aa, DF6, 161 aa, DF6 .

MPPHGHTPNLFVFSGLLQVLGKARQPKALHYFAEIERSGLTPNTHCFTR
SSAALGTAGRVDDAECHFGQMAQLGVAPNLYCFSALISGFAEAGRVDAAE
HWFAKITTVGRLKPDLSYNALSQRGGGAAREGSQASAGDDPQQAQTAT
DEHRVSAFPQH

>g5406.t1 - pentatricopeptide repeat-containing protein, putative,---- NOT DETECTED --- 659
aa, 1E5D, 659 aa, 1E5D .

MLQRENRRPSTERTTATTAASPOKPKSTQDVQQTQORIKHAGAIGEVDE
LVATIKRTHQLHPSAVSAVPLLLARTRQWETAVELVELMPRHGHGDNRRGR
RPPNTTTTGTAAASPOKQKQSPSDVQQTQORIKRSGAIGEVDELVATIK
RTHQLHPSAVSAVPLLLARTRQWETAVELVELMPRHGHGHTPNLFVISGLLK
VLGQARQPKALHYFAEIERRGLTPDSHCFSALITALGGAGRADEAERYF
ARMTQLGLAPNLHCF SALINALGNAGRVDAAERCFDAMDQLGVMEKLGVV

PDLHCFSTLMSALTKAGKVDAAQHVFHQIASRGLKPDLVCFNALISGFAD
AGRADEAEHWFAKITTTGRLKPDRISSALIQANEVVARPDRALALLQEM
THKGVRPDDVVSYYTIVISCLAKARGRRHMDQAQEVFNEMVKQGVNPNATTF
GALIHGHALMGDLETARSRLKQMSSTEFNLAPAAEIHNILMDAEKRVATDP
VEAYHQVLRVAVDHMKQEGVRPDHATFSIVIDAAGFAGLPHEAEHQFHAMT
SLHGLHPNANNFTSLIEALARNGQLDKAEAVLDSLAAARQATSALHDRLT
AKSFNTLLGFCIGAGDAARGQRVAAMKHQAGVEPDDIGQAKLRDLAKLCE
RRAAPPTPK

>g9903.t1 - Pentatricopeptide (PPR) repeat-containing protein-like ---- NOT DETECTED --- n=2

Tax=*Oryza sativa* RepID=Q5VS02_ORYSJ,, 168 aa, D77 .
MQAEGVESDPATTFTLIIDAHAASGDAERCFGAYADMRRDQVMPTVATYA
TLRLRFCEKHGNVEKACALWEDIKADQIRLPWNVICFLLRATLTNNKTDTA
FRLAEELSQLYKPOFVYSQLLTICDQLADVWQAERVWRVMRAKKIPVYAR
TKQLLADFGLNPDEPHHT

>Rhomboid protease Contig14413 -----NOT DETECTED -----, 353 aa, 766, 353 aa, 766 .

MRRIGSLRTSPFCAPCLASRALPLRSTLLSRPAFGPKISSGLPPRCTPK
AAPQPRVSAIVAARSFSWTAHPLHGRPQTNSGLQSLSNNTSSSASSKASG
RTSWLTCQPRPTTPLSSHNSASTSMWARRYQVWSRDQYEYRGVDVINV
LLGINALVFGGWWLAPPRFMRENFVSTLRNLREGRVWTLITPAFVSHQSPM
HLLFNMSMAMWVLGKEVMFVVGPSRTLLLYLAAGVVSCVCHLAYVNEIKPR
IMASTPRQYQLYKMGVGAEQGALGMSGVSMGISILYAAFFPSAILLLYGVV
PVRAPILVGGFIAADLLFAFSAPESSISHAGHLGGALYGLLYALYLRRRF
GRR

>Omal incomplete, with gaps ----- NOT DETECTED -----, 302 aa, 594, 302 aa, 594 .

TLPSALIQRHKSQYFHKFHHLKMRTFWSQVQKEGKRPKRLAFAVGLLL
AIPLGGYGYWQYYQDIVPHSGRRRMDVSREEEQRMGLASFHLLKQOYKD
QILPENHELVLAVKRIGRRIIEQTDLQGLHWEFLVIQSDENAFVLPGGK
VCVFTGILSVTQDEYGLATVVGHEVGHVVARHGAENGWERALGWLTAILI
AREEPGWLINQFLTFLALPFSRKFHEADFIGLMLMAEAGYDPTNAVDL
WIRFSAASPSMLKYLSTHPPNADRIVLLKRWMPQALDQYHKHTRPQATAV
IV

>pyruvate dehydrogenase kinase NOT DETECTED Contig2871 (and 5370) 441 aa, 441 aa, 1824 .

MMNRMRTTVASTSSHVHIPSSISSAASSSPVSWQRTMDDMIHRYANLK
QTPVSVKNMIVFGKDANMNTLLRSHAFRLRYELPIRLAHIAKEITALPEEL
LEVEPVQKVLNWRYSFSEVIESPVPNVADMELKTEDEQCASLHQFQDML
DHIKDRHSGVVTMAEGVLELKNRLGREMIDTSVQFFLDRLYMNRI SIRM
LITQHLELFKQAQTNNNLCAITSGASGRKIAEDAVIDDARYLCSNNYSVCP
SVQIIVPKLGTETSAGAPTLPYVPSHLYHMLFETIKNSLRVVEVHGAN
AQSEDDLPPVRVVLVKGTEDLTIKISDMGGGIPHADVPKLFYFYTTAAP
PTKDTLDSLNEGAAPMAGLGYGLPISRLYARYFGGDLKVIPMEGYGTDAY
IHLKAAGETREVLPEYSPSTYLGSQANLREWLNVDNCAFQL

>TOM7 Contig18214 NOT DETECTED, 64 aa, E0, 64 aa, E0 .

MTLGGLDLEVMsAKAARLYDKAVVVARVGFHWGFIPFIIFLGVRSMRREQ
SWLSLLVPTPYLQA

>Mdm12 Contig7451 (and 3688) NOT DETECTED, 255 aa, 134F, 255 aa, 134F .

MSFKIYWRMGEEQAAMKDYLNFAFFSSAPKPEFIGDIAITALTLGDKAP
EVRLIDVSDPPAHFSSLPMNDAPSSLDSPATTARTTSCRQFDGSDRHIVI
IASSSSSSSSSSSTTTEPEVKATDMDMMIKVFLEYSGTAKLTMRCLISL
VNWPRPRTVCLPVTFTLSDFHFEGVVCAIHHANKVHFCDKREGDDVGPL
KSLKIEAQIGDSTQOVLKNLEKVENFIVGMLKLIKDRILFPNFITVLDQ
TNKAH

>Tim9a Contig4832 NOT DETECTED, 88 aa, A34, 88 aa, A34 .

MQRGRMSKDDEASISQMMERMOMEDIITHYNRLADTCFTQCVGDFKSNAL
LQNEEVVRRRCVEKFTQFSTRFQVFDVHDQTAANKQOQ

>Tim9b Contig26620 NOT DETECTED, 127 aa, 2699, 127 aa, 2699 .

MGNRQFGMGYGMGAGQSRAGSASDEERMMQVMMASMQMODYLTMYNSTTE
KCFKCKVFNLRTPQLVEKEEVCLNRCIEKMSHYNLRFOQKVGAEASALRE
QQQKEAEQKQPPTEAAQADKEAPSSK

>Tim10a Contig2235 NOT DETECTED, 106 aa, 2360, 106 aa, 2360 .

MNGQPMPSGMSEQEREDVLAASEFITIQMDLYNRVLRGRCWDMKCI AHYWDK

DLTVGEGSCVDRCVVKYMTVQRTVAKRVLGIRAWTKREEERQKVIQKLRD
NHFFGV

>Tim10b Contig8351 NOT DETECTED, 84 aa, 14D6, 84 aa, 14D6 .

MSNEQYMTREQMLASAOQELVAFQDMYNRMQHQCKKCVLRLGESELSVS
EGLCADRCVKKYMEVHNRVGVKVLQGLQQQQPQPQ

>Tim22 Contig19250 NOT DETECTED, 189 aa, 182D, 189 aa, 182D .

MEGRDAASQGEPEPTARRQESQSPFPQAPPGTWQQKLDADFVEPFVIKYLWD
PRFAENCAFGRVQTFVVGAMGFVMGIFFSSLGGPSSMGMPDVEMKGGWK
KQTAEHFKHMGRSGVSMKAFAYVGALYATTECVVEKYRGSKSDLMNPLIA
GCISGGLLASRAGFTATAMGCGGFAAFSVGIDWLMMENH

>Pam18 Contig13309 NOT DETECTED, 143 aa, 24A0, 143 aa, 24A0 .

MQRTTAQAAKTSQPKMGPGTIFLALGVGGAGLALAPRLVSSAAKHFSKGG
SSGKGFATMGSMWKRSGGGASTGFKNFYKGGFESEMTRAEEAALILGIRQS
APKEKIRLAHRRIMLLNHPDNGGSDYMASKINEAKDVLVKDLN

>Mdm38/LetM homolog Contig3961 (and 13276, 19860) NOT DETECTED, 339 aa, 1BA, 339 aa, 1BA .

MRRLTRTTSAAAAVDGVVMRHESWAERAKAGAKAMWAQYWGGRQLWVDA
GAARALRRRVRESVASGGVAADGVALSRRETRFVRRVADDVRALVPFLA
FFMLPFSAYALPVIVRFFPGFLPSTYHSSSYRVAQMQRKQKVKVEAARQL
RDALSERLLATEGPSARERDLLLDFIHRAAAGEAIPLTQLRPLDRYFATH
LKLLDLEPRRSSGASLGSWLGSTLWLCPAELRALCTFYGLSALPLPTFDH
CANQLILHASRIQHDDKLLVQEGDRWPTTELDEELLEACDERGLPTDVTE
RDRLVAQINQWLAFRPPPPVSAELTLFAPALATNTRAK

>Mdm38 (Contig17895) NOT DETECTED, 386 aa, 9AC, 386 aa, 9AC .

MRKCPTTTTTTALGLRFARRSPSSGLLTARRLPSSGAYRSQLRPAAVAG
WRGMSTTGHGPAAPDTKDEQQPLGATTTPPPAIAPEAEAAAPAAVAAAPKW
QTLMAKAKELARGIKDGTVELYRNVKHTRRLRAVVRERIDALTRREIRL
MNRTREDLKALVPFLILEIVLPEATPFVVMLFPGILPSYQRLIPQQRK
ERERKRTVEALEQMRSELHALGGQRPINF'TDSSLDDLVASLQDESGLRE
KLEVASLPDGMVEAMTRYTSVWSRFRTQNGLRSSLQKLDALRKDDELIA
KEGGIDSLSEEEKEEAYYERRIGLAATSPASPAATVEPSSAVGPTDPDRE
RHLREWLHLSAAIRDCEQHRATVPDTSLLLLLALRH

>Imp1 (Contig19247 and 2953) NOT DETECTED, 241 aa, 12FE, 241 aa, 12FE .

WRSGGGSRLGAVLRTTIMVGAYYGVAITFCNTVLKPSSTAGPSMHPTFNA
AGDSVWVYRRIDPATDLRVGDIVHARTPTYCRLEGGKQPGVLKRIKGLPGD
TIKVTFFYSPSKVTIKVPAGHVWVEGDNPGQSTDSRMWGPLPLALIEGKVV
SRLNPFSLQTLDAPGHIKRWERQARDIADSRATGKRSGEALRAIDRLI
KSGSEASDNTLLVSDRHDADVNDTALAPPQDTVQGHERVVA

>Imp2 (Contig6343, 15292) NOT DETECTED, 169 aa, 1090, 169 aa, 1090 .

DGKQAKQITQTTGAVRTRRGCGVLLWLHARRHLPIWGHTAVEGESMYPTL
EKNDWVLTSSRRNGLHRGDVVVLRVSVKEPKERMIKRVTLGLPDDLIHTDDA
KLVRVPRGYCWIEGDNASVSLDSNVFGAVPVGLVESRALYRLDRGWRLHK
LESVPRERLQRTKKEREVR

>Coq1 (Contig16693) NOT DETECTED, 453 aa, 68F ., 453 aa, 68F .

MRKMASRSTMRRLGGRATATCATQCTVSGAARPSSLLYASPLPSMARRDL
ATATPAQTPVEGLEKFKVSKYQRTFDLADPHHDGSASSVVEGLLRALPT
ALAGGAEAREDPKVALRLVEKEVKSLNGSIQQQIEAVTERPLLAQVASY
FFKLQGKRIRPVLVLLTAKAIAVHLNPAAGLTPENADVPVSDRQFSLA
EIMEMIHTASLMHDDIIDQAE'TRRGMPAAHKVWDSKKAVALAGDFLLARAS
VKIARLREHEVTELLSTMIADLVEGEFFQMR'TNVEKEGKEMTQFDYYMRK
TYLKTASLLEKCCSCAAILGEADRATINIAYGAHLGYAFQLVDDMLDF
TGTSSDLGKPAAVDLSLGLATAPVLFAMEEF'GELRALVERGFKEEGDVEK
AFSLVQRSQGIARTHALAQEHCAKAVETIGQLAPSPYRDALARL'TEKVL'S
RSS

>Coq2 (Conti14535, 20541, 15288 ... incomplete) NOT DETECTED, 394 aa, 53E ., 394 aa, 53E .

MKRAAGRSCASLMRPTTTTTMRAGNSLQWRMRSCAEARNARVATVTTT
TAVTAMPRPTCGHLLTTQRAVLLRRRQLSSSTTSRSFFLTSEEEPPSSSTT
TGWAALKQRLTPYGALARVEKPAGTWLLLWPGWVAIGMAAPPWEHLALL
GTFGVGAFLMRGAGCTINDILDRDFDARVERTKSRPIASGALTRKQAVAF
LGLQLTAALGCLVTLNPLSIGVGAASLVLVGTYPMLKRITYWPQAFGLT
FNYGALLGWTAATGGFDPTVVLPYASGICWTLVYDTIYAHQDKKDDIMV

GVKSTALLFGDKTKRWLSGFAAGSASGLALAGYGAGMGLPYFASVAAAAGA
HLAWQVGTVNIHDPADCGRKAFAASKWLGAVVFFAGIELSNFLAVM
>Coq4 Contig5533 (and 11891) (mostly complete) NOT DETECTED, 292 aa, 213C ., 292 aa, 213C .
MCCRLCPRATALATSTSLRFASSSSSSFSSNSASTAATPQRTAAPPPIIDA
SCNLYANHRPTSLFQKGLLAAGSALAAIINPERADMVATLGETTGPSALR
ALHARMAAHPVGLLEILQTKPRITEDTLAFDRLLAMPEDTLGGAYARWMRG
HEFTTPDERPAVRVFDAAELAYVMQRYREVHDFWHVLTGVPTTVLGEIALK
WFEMYQTGLPMTALSSVVGPLRLPVSEMRVLYSEYIPWAKRSATNATFLM
NFYYERRLGDNLAAIRDELRLPWHGQCETFFHHYMLPFLG
>Coq7 Contig (possibly complete) NOT DETECTED, 253 aa, 847 ., 253 aa, 847 .
MQSRIVLRPQRLVSPGGDRTFVAVLSTLPRGTAAISARWQSAQTSPTSSEEA
EAPGAPTADSDAAASHDAAAGSDGQERLRQRLYDQIIRVNQAGELGAQQI
YAGQLFVMRGTDEPILQHMADQEKELQKFDELMERYRVRPTVLTPLWR
VAGFALGAGTALMGREAACTVAVETTVEHYNDQLRTLHGHNMNSPKD
QELRQTIRKFRDEEHLHDIGLKNDAEKAPLYGLLTKGIELGSKAAIWL
S
TRI
>Coq8 (possibly complete) NOT DETECTED, 585 aa, 12B4 ., 585 aa, 12B4 .
MLGRAIFRSPPPAMRTTGTDAATATAPQTGASAAEIAATTQDVVASSVLA
PPVVPPIPPVAPAPTTATALPPHAVHSPVPPSAPAPSALAESITTTAPQA
AKLSGAGAKVEAQAPPLEEPLTKTAPPPRPTFDRIKYLQEQQRIKENS
K
QRKVPSSQVGRLLWHFGGLGVRMGMGTLGEEAMRTVGLGSSASAHPAVTPA
NAQRLTNTLGRMRGAALKLQMLSIQDSALIPPELAEILDRVRYAADRMP
DQQLVEDVMNHELGHDRSRVKEFDPVPIAAAASIGQVHRGVLADGREVAIK
VQYPGVADSISSDLNNGSLKMLGILPKGMYLENTLRVAQDELLWETDY
HREA EATMKMKELLSDEEIVVPTVPIEMSSKKVLTTELLDGOPIEHVVK
DDQKTRDMVAESMLRLCLREIFEFYMQTDPNWSNFLYNPETRKIGLLDF
GACRDFNKSFTDEYMRVIYFAAKQDREGIIDASRKLKFLTGEPPIMNEA
HCSAVMILGEPFAKPGPFNFQAQNVQRIHELIPMLKYRLTPPTTETYS
LHRKLSGAFLLCAKLQAQFACQPMFMDLYRKCGPQ
>Coq10 (from PGP) NOT DETECTED, 184 aa, 2494 ., 184 aa, 2494 .
MHARGKVRTVGRATTRVPALRPSGVTADAGRAIEYKDERVLPYSADEVYD
VVSDVAKYQDFLPWCTHSRVLFKTPTRMDAELGIGFKMLPGQSYVSQITM
ERPRSLKVVVTPDSSLFNHLINWAFQDLSKDGKKSCTFFFDISFEFRSQ
LHKSMADAYFQEV AQNM IQVFNDRC AHVYRWGKQ
>AGGG, 74 aa, ----- NOT DETECTED ----- 2085 ., 74 aa, 2085 .
MGGDHHPYKGLDVKLPYGGFHPPPVRVHKNAALFIGTMMWFVFWRLK
HDWRHLFTSHRPEIITHEGEDEHH
>NADH dehydrogenase MLRQ ----- NOT DETECTED ----- Contig11127, 89 aa, 1839 .
MQAIRRTYNENTKAVWLTLPVVGGLWAAAYIGFRTFATHNDVVLKARG
NEPYLMRDSPKLLSRDAAITLADKWSNEMGGDDQFQPK
>Ndufaf1/Cia30 Contig5982 NOT DETECTED, 228 aa, 4EB .
MLGRSRAGVGSVKNLVREALAGPPLLREGAFGGPQARERMLYSFGAGAD
PAAWEAVTDADFGGRSRKLEPTEQGTWVFSGALDLSTEGTEMKQAGYAG
LQPRQRKTIKSLEGFDALEVRKTDGRVYIANIKTDSMVKHHLFQAFFTT
RKDEWTNVVLPFDRFTLTFQGVQVEGESLPIDPRQFQAVSFLMAERKDGEF
RMELEWVKAINTRLRKGARYTGLVHDQD
>Foxred1 homolog Contig27121 (and 22107, 24645) NOT DETECTED, 458 aa, 1851 .
MMKCRQPRHLLSARCASPAPAAVVRPVPRHRHSRLLSSTAGGAGGGSIVI
AGGGVVGSSVAYFVARALRSRGTLASPPPVGSRSEWTVTVVERDPTYRF
ASSTLSASSIRHQFSTPENILIGQYGFELRNAGKLLSISDEDEARLSGR
LQEKPOQPQHPALLESNHQLQORDLGSDIRFLGSPSEIKAQFPYLHTDDL
S
AGCFGASGEGWFDPATLLHGFKRKAALGVNFVHGEVVGTELEADGPRQR
VSGVQVTTADGVAHKLSCSHLVNAMGPAARLSQMMGLTLPVVPKRKCVF
VLRCRDEEVNQAKMPLVIDPSGVYWRPEGCGFIAGVSPLVDPDAHGDFFV
DHDLFEEVIWPVLAHRLPAFEAVKVVGHWAGHYDNTFDQNAILGTHPSV
GNYYHANGFSGHGLQQSPAVGRAVAELLLLEGAYKTLDL SMFHFERVLENR
PIIELNVV
>NDUFAF3 homolog Contig17448 (complete) NOT DETECTED, 144 aa, 1772 .
MAAAGPRLPGEYDVFVGGDGGEGSHGNKVMVDGYSGSRFSINGVLLPGPL
FLLPTLPLAWNVTSLDDLSPESLALAYAIRPRVSILILGCGQRIKMINPQ

LRDFFRQKGISIEALDTWNAVATWNILNQEDRAVAAALMPFDYE
>Lyr5 homolog Contig26357 (likely complete) NOT DETECTED, 127 aa, 1ADD .
MAQNSSRTAARNLYKQMIWVGRDYPAGLDKLRPGMIWVGRDYPAGLDKLR
PGAKRAFFATKDETDPEKIDQAFRRGEYILQELEALVKLHKYRTLKQYN
PDREDEMVASEFEKFKAEAPLPSGF
>B17.2L Contig22529 (possibly complete) NOT DETECTED, 177 aa, 268 .
MIEAKDTIYDPSSVPALWRSWLNFRREEPPTEKELEAYEAKQRNLAQKVA
ELEEEEDAKLALQEQAAARQAGVRNAGPQGATPKGMLDITLVQLAAGEIDQR
GQKSAEGKEMDVEDDDLPPFEFPSEAKEWDSRRRPRSTDDTYSRPSRAA
SFSAPPDSFADDEDEADQEWKPPSNNK
>SdhB isoform 2 (probably incomplete/incorrect ... almost no EST data) NOT DETECTED, 340 aa,
1172 .
MLPMRQRIASSAVCTARSCVTRSPLKIAPTLTTSRTLHIPPGLTKLDKTV
QDGSTCTAPSIARVAQLPTKLTTCMQSWTSKCTAGRARRARSPTLRRYV
VNLIRFVPLGGALLNSLCPFLNRNSCREGICGSCSMNIDGQTTLACTKPI
KESITNGKIKIYPLPHMEVIRDLVTDLTHFFEQHREIRPYLMKNDDAALK
QSKVEQHQSREDRRLDGLYECILCACCSSTSCPSYWWEGAQDDKFLGPAA
LLNAHRWISDTRDDFQADRLLAEADLKVYGCHQIMNCAVCPKNLNPGLSI
ARMKDLVSVQVADLDKKLEDKAAIELAIRKRMHHMPPVAGV
>Cox8a (4720) 80 aa, 2286 ---- NOT DETECTED ---, 80 aa, 2286 .
MNFLRRTPLSGVRSFALRGHGPSHGPDGFQLGQAVGTTLPFSTAQTAKPF
RVGLSTWGYFVAGIGLPTWAVWYSANKEAK
>Cox10 (incomplete) NOT DETECTED, 382 aa, 1982 ., 382 aa, 1982 .
MLPIPHLGVHALGRSAFNPVRLAVEGCSSTERSWGWNGAWSRAEQPGEAA
SGVCIFPSFPHTVAHTKRVSFLQSIRETAMYSSELGKIKLGSVVITMA
GYMAPFSSFDLKRFAITSLGTGLAVISANTFNHIIIEAKLDRMVRTSTR
PLPSGRMSSTHATMFGVITGIVGPALLMARVDNTAAALALANIVLYAGVY
TPLKQMPHPINTWVGSVVGAIIPMIGWVAVTKGKLDLGAVALGGLLFLWQI
PHFLSLSFPLKEDYARAGYKMLSVTNVDVMSLSYVFSAMLLPLGFVSSY
AGMTSWVYAIDSTLCNGYLLWMAHKFYTRRDTTSARKLFFASLIYLPLOM
SLMMLQQEQPPRQDEQRKTKGKLDLGAWATWR
>Cox15 Contig7183 NOT DETECTED, 428 aa, FAE ., 428 aa, FAE .
MRTSSSGVAPLFTSARPGIAARVFPRLATATNGAALRPATLLRRANSTVT
HQATQPGVMPKVAAATWHDVPTSMKVVIAGCTATVFGMVVVGGLTRLTES
GLSMVDWGVLMQKPPITQEEWEAEFDKYKQFPEYKKNLSKMTVEEFKSIY
YYEWAHRNLGRFLGAAAYTLPLLYFTARGHVGRPLRTRLYAIAAGIGFQGL
LGWYMKVSGLVEKDHPRVSQYRLAAHLGSAFVLYTAMLWTTTFDFVLTPLT
TTAKLTTSAASAAAASAPTVLKLOAKMRKYAHGVAGLVFLTAISGAFVA
GLDAGLIYTDFFPYMGQVVPNEILDMDPKWRNFTENDVTVQFQHRVLGIT
TFSSIVAMWLWSRKVPLPPRSRLAMNALLGMSLQVTLGIATLYTLVMTH
VAATHQAGSLTLLSIATWLMHELKRLPK
>Pet191 Contig2276 NOT DETECTED, 90 aa, 1739 ., 90 aa, 1739 .
MASSCKSLKEELVACLKISPCMQDKEKTFHECLKSKDEDEIGQCSKVRRG
YFECKRGQLDMRKRKFGNVDFIRKSPTQAPAQPPAAGGRS
>Surf1/Shyl (likely incomplete) NOT DETECTED, 284 aa, 17E8 ., 284 aa, 17E8 .
MPPRSSLGRAPLRSFANESREDIAARQODLKNFVGGGNKAAATKPASGEA
GOVMEGGLAAWFIFPAITVCLGTWQVYRWOEKKEMIRYRKERLLEPTVD
LPPHLTDETLRDFEFRHVEVSGKFDVREFVVRPRTRDQNGGLLLTPFR
RSDDGTIVLVNRGWIPAALMNDIVQRRAMLDRVSGDTRLFGLIRPGEKGG
MFSPENKLADQWYWFVVDKMRQALGLDYTPALLDFLPRHTQCRRIPT
KLRDNHLLNYIITWYSLTVALVATMLLKGRGRVRR
>mitochondrial transcription factor B-like protein NOT DETECTED, 307 aa, 1A8 .
MLKLPPMPTPRELVRLYRLSAVKELSONFILDNLVTDKLARAAGPLRGST
VIEVGPGPSLTRSLLTNGARKVIVVEKDKRFMPALETLQOASGGRLELV
FGDMLKIDERDLLKNEPKAENWADESPVRIVGNLPAVATELLLLKWLROI
PEREGPFAHGRASMTLMFQLEVGKRIEARSSTSEYGRLSVMTQOSCTAQT
CFNVPASVFPVPPKVTGTMVRIEPRVTPLAPAPVKELEVCRQVFGQRRK
MLSNAITTLGEGSLPLIARAGLDPTKRPDALTVEEWCSLARAYKEWMDM
KTSSFNI
>IF2 g476.t1 (incomplete/incorrect N-terminus ... from ESTs, PGPs and inference) NOT

DETECTED, 741 aa, 78, 741 aa, 78 checksu, 741 aa, 78 .

MIVRVCGIVFCSYASSPRHFIFHILLFRAPLPSPHRFLNAIYGVNIDGSG
ISIHSARQCAVRGRCSPIIIVSFDIFLDSDFNQEQLSIDFLMTQAGGTRR
LYATASEMEKSKMVRKRLAATRKRKLEELKERKAKPKEISIPSRATPKNL
SAVTGLSTIDILKVAIKCGHEPRTVDDVLPPELVDILCEELYFVPRRSDL
NEGTGMKLTTRTIDEADMVPRPPIVAVMGHVNHGKTSLLDALRKTSSVVDVE
EGRITQRLSASMVQLSSGQKITFMDTPGHNAFAAMRQGAAMTDMVILVV
AADEGVMQOQTEEAIEVTREANVPLLVAINKIEKEGANPSKVQRQLLQHG
ILEEFGGDVQCVELSASTGKNLDELTDAILLQAEMLDLRADPNGPAESV
IESKIDKGRGVVGTVLVRQQLKLNFFVAGGTWGRVRGLTDWLGRSTTK
ADPSTPVEVIGFKTKILPDPGDDLVVVLGEEQAKQILDYRKERLHAKEAL
TDVVQASDPAARAAAQEMEKALNVVIKADIGGTLQAVESTISQFPQDE
VKLRRIKSAVGDISTTDIGLAKDTGAVVFGFNVPMMNSVLELARVEKVDI
NNYSIIYQLADDIKRLESMLAPREVEEEVGRGDVLSVFSVTMPNREQHN
VAGLRVVRGALDSRLRVRLLRNKVDNADEEYAAALYDGPISLKKFKDDV
TLAKKGDECGLALPKWNLRPDIIISCYRMKQVPRKFGDPK

>Guf1 g3637.t1 (possibly complete ... N-terminus from Genscan) NOT DETECTED, 606 aa, CE1, 606 aa, CE1 .

MLGRRGLLASRGPAAAVAAEEAAAARGAPDLSKFPSPKIRNFCCIISHIDHG
KTTMSTRIMEKAGTISLDKEGLYLDKLOVEKERGITVRAHTTSLVYKHS
GESYLLNLIDTPGHVDFTYEVSRLKACQGAULLVDCTQGIQAQTVANYF
LAWEANLKIIPVMNKIDLPSTDPARVIKELANMGFESEVLKASGKTGEG
VEGILKAVIERLPAPSGDPAKPLKALLFENWYDQFRGVVCLVNILDGALK
KGDRIVAASSNKVYELENGVILYPERAPTALYTGQVGYIIGGMRSTKEA
RVGDTLYLEGAPVEPLPGFKSAKPMVFAGLYPADGESLDRLTDAFEKLT
NDASVTTYTKERSDALGMGFRGFLGLLHMDVFLQRLQEQYGTQDVIATAPT
VPYIVTTKEGQEIIVNNPSLFPQHDILTCTEPVAEARI IAPKEYLSPLL
RLCLERRGVQOELS IDEARLHLRYRIPLSELITDFNDSLKALTSGLRHP
VEALSAIAHQSKGVNYGKDLVQKLRKVVQRQLFQVAIQASVDGRIVARET
LQALRKDVTAKCYGGDITRKRKLLKQKEGKRMKTIGNVQLSQDAFYSI
LKKGDE

>YchF g6187.t1 (likely complete) NOT DETECTED , 435 aa, 203C, 435 aa, 203C .

MKVATRRSTTWGSSRSATSTFVGGGRSVLWPPSSSCLTSGSGVLGVR
QWFAAKAGKQOQOQAAAGGQASALS YFVARTASNGCGIVGMPNVGKSTLF
NALTSSQRARAENFPFCTIEPNSGMVWVDPRLDTLAALANSKRILPCQM
EFVDIAGLVAGASKGEGLGNKFLAHIRNVSMILQVRCFEGADGTEQVTH
VEGSVDPARDMEIINTELLADLESITTARAKLEKRTKLDAAEQKRLDLM
HRTIAVLDEGQPARSVLISAEWPTFHAFQLLTSKPVVYCCNVKEEEAAT
GNAMVEQVKQAVLKEEPGVDPNERVLVVS AKLESELANMESVEDRKSFL
LYGLSEADAGLPKVRATRQLLQHQHFFTVGETEAPRLEHTQTIAFADFV
AAGGEKRAKELGKVRFEGETSYVVQDGDIIILFHHRK

>EngB g15942.t1 (incomplete - poor genomic, poor EST data) NOT DETECTED, 134 aa, 19F9, 134 aa, 19F9 .

AFIGRSNVGKSSLINALTAQGVKRVSDTPGETQQINWYRVGKHLTLIDL
GYGFATAAEKVKAWTELMCEFLTSRKS LKRICVLIDARHGFKPADLDFI
DMLEKAKRKYQIVMTKADLVTPPDLARRYHLVQQ

>EngA g13051.t1/g13050.t1 (incomplete due to retained intron ... N-terminus uncertain) NOT DETECTED, 675 aa, D48, 675 aa, D4, 675 aa, D48 .

MPSMSCPIASARGLATSRGAGGSRGGGSGRDRAQQTTRPHSQKSSSSSS
SSSSATRGAGGKGGGKSNAARGPGAAAATKKRAPSGEASKARPAAEAAA
LGGVESDTERYTMRRDPVLALVGLPNCGKSTLFNRPMVTMGR LVRMKNKSL
VTSEPGTTRDRVYAYCDIGGRSVMVVDTGGMVGTGDDFFSPLIHEQALVAI
NEADIILFLDFDFREGVTKKDVVARVLRKRVHDKDIVLVANKCDNPDVL
ELDERNKAGPTGKPDYWLGLGPAIPVSAIHGAGAGELLDKVYDIVQTK
FPAEDQSEGDAEAEAEELARKRISISIVGRPNVGKSSFLNQVLGRERVIVT
NVAGTTHDAVDHSIMWRGKKPRKRRAAKATMGLEEEAAAATTEGKQOD
AEVVHAAEDDEVIDEQEDDLEDEQEDDLDDDEHETKVRAKKANNTKLKE
SAANPDADGVPIRIVDTAGINRRATHVKGLERSVWLWAIKSINRSDIVLQ
LIDATQGITDQDLRIAHEHILNSKATTIVVVNKWDMSESQKQWEKYVQET
LKFMKYVPIMFCSAKTGVNVKETIDLAIRITRERAHFIKTSTLMKLLESA

HTRHRPPSKNGRQLKIRYATQAKVKVPSFTFFVNDVELVHFTYQRFLENA
IREFHPTGSPMRLIFKDKEKTIVS

>Rf1 g10217.t1 (possibly complete ... from PGP), 370 aa, 1643 NOT DETECTED, 370 aa, 1643 .

MAARQRWVSMQAPLLLFHAGSPFPRLDHMKHQFDDLSATLQTNALSPEE
RGMLTKQLHALDPVMGMVRRIEQOREELRELHQLMEESKQTQDKELGALA
EEDYKRILGELMDLEKELENEIMLELLGKDEADEKSAILEVRAAAGGQES
QLFTMELFNMYSAYAEMKDWEWETLAVSDSEVGGYRESSACITGRGVFGR
LKFETGVHRVQRPETETQGRTHSTVTVAVLPVAQEADIQILPKDIRVD
TFRASGPGGQHVNKTDASAVRVTHNKDRAMKILRSRLYEIERERLASSRAD
QRRKQVGTGARHERIRTYNFPQSRITDHRCLTVHDIESMMQGHLLDTFI
DALVAHHSAEALNSIDLDSF

>EFG2 g6016.t1 (incomplete due to internal gaps ... N-terminus inferred) , 669 aa, 1A01 NOT
DETECTED, 669 aa, 1A01 .

MQLRRATLRVGPVAVTIAPRSGGRQWWCQARHSGSSNSPAAWPPQQGGG
SLVLSSSSSRLIRPSFFLTWVHRRQOHSEAEVVVERVRNIGVVAHVVDAGK
TTTCERMLHYSVTRRIGDVSNTVLDVDFMKLEQERGITIKAAAITFNWE
GHKINLIDTPGHVDFMFEVERSVRVLDGAVAFDGVAGVQAQSETVWRQA
RKYGVPVIAYINKLDREGASLERAADSMRVRLGARPLLLQLPLPTTLGVL
EGVIDVDMRVIRWTDPSGKAMQVVDLTPEEHPGLYAKALQARETLLEQL
ADADEAIMRAVLEEGASAETVAGELLHGALRRVTLAGEGVPVVCSSLRN
RGVQPLLAACRYLPSPLDRPPVVDVQVGLMVYLRVYSGELTPATEVYN
STRQVKERSRLLNIQADDMTEVGRVAGNIVAALGLKDTCTGDTLLVPA
KGNKKAHKGQHAQGGRLLEGLDPPPPVFFCSIEPESLSAQKALDDALGL
LRKEDPSFSVSLDPETGQTLLSAMGELHLEILKDRILNHVRVGARVGNVR
VSYRATVGRALSHTLAREYEIGGVKQHVLETLTDIEPTARAHGNEYQASGA
GRGMVRAHVHLPHHRLVEAEAPLKELLGYSTLFRSLTRGSGSFTMEFLRY
GDMGAAESKKLQDELRRGF

>prolyl-tRNA synthetase g100.t1 (probably complete) NOT DETECTED, 622 aa, 197A, 622 aa, 197A

MRRGLCVTARRLVASRTNALWHPQAPLLRRSLTTPGATLPVAHRVST
LLGAGTGSQAQSSGGGEQDQQLQVASHRLMLQAGIMRQAGSGIYSLLPM
GVRAVEKLTALVDREMGIGGQKLQMPPLLLSAAIWKQTGRWESSGPELFR
LKDRRGADFCCLAPTHEELITQLFANEAVSYRQLPLLLYQIGLKVRDEVRP
RFGLMRGREFIMKDMYSFDITRDDALRTYQHVVGAYKRILAQLELDYVIA
AADSGNIGGDHSHEFHVMANIGEDTILRCNSCTYAANTEKARGRLHHHE
HRRRLDHRTLDPALQAGLADPELPSVHVVRVATEAPAASAEGGAARE
WAVVAPKGRQVNELKVKALLRHALAPSGEIVDITVRPAAAADAGARPVV
LVDESLVAAEAVDAKAVGDFTLTEERDGCAGEGEGELVGSRGIEVGHV
FYLGTKYSAKLDARVATPSGTNYAEMGCFGLGVTRLLASIVETSHDERGI
VWPVAVAPYKAVVAMTPKDADVQAHARNLYADLQRSFYFANDVALDDSA
DSPGAKLHRAQLLGFPMVVLVVKAMKNEGRVEFENRKTKEKQFVPYDQMV
AHLEQLARDHARTIGAPLPVGL

>tryptophanyl-tRNA synthetase g3676.t1 (probably complete) NOT DETECTED, 354 aa, 457, 354
aa, 457 .

MLRQSCGRRGSLALAPYRGLPAEAGRIFSGIQPKGHLHLGNYLGAVVNW
LQLQEQRKGDGDRDGVVFCVVMHSYTTDVTPEDLRSTREMAIALLACGID
PEKCILYAQSEVAEHAELAWILGCRTTMGALNRMTQYKEKKASQGAMFGL
FAYPVLMAADILLYRASQIPVGGDQLQHLELAREIAEAFNRRYNRTLFPPL
POPILQOATRVMSLRDGTAKMSKSDPNDRSRINLTDKDEILFKLQKAKT
DPLPGIVYDPENRPELANLLSIFSALTNRVSVESLCHEYADAKMPDFKNAL
APHLVDTVGPIGERIRELRDKPQWVEDVLQOGASRARTIARRNLQEVKQV
IGML

>GatC g5738.t1 (possibly complete) NOT DETECTED, 146 aa, 2026, 146 aa, 2026 .

MRRACLTRPFRRGYAAVTPREPSWSVEKFFQSSDTQIDAATLQQLAKR
AQLTIAPEKEESLCREVGQILSCVHIIQEVNTEGVEPLVSPLDARQAQLR
LRADEVTDGDIAEDLMRNAPARDGSFFVVPKVKQSGGEGEGADA

>translation inhibitor protein g7864.t1 (possibly complete) NOT DETECTED, 203 aa, 1656, 203
aa, 1656 .

MKRTLHVARTSVLPISRGASCGAKMLATRATAPAEKKPRAKRTSASLPK
QAAGHHQVVSTALAPAAIGPYSQAVVATGPFLFVSGCIGLHPQTMDFPGA

DVSTQTEQVMQNMGAILEAGGSSFAEVVKCTILLTDMQHAFATVNDIYAKY
FPRDPPARATFATNGLPKGALVEIECIASLAQKTTTRTRKAGTTSGRGKAK
QEP

>ribosome-binding factor A Contig6327.t1 (possibly complete) NOT DETECTED, 849 aa, DE1, 849
aa, DE1 .

MQKTLRGRVARSGHLPPAFPTASRATHFLRFSSHNSLGSAAAAATGIRSR
LSSSSSSPLVRSCFPVAPAATYHTSASRALPRHDRGGGATKAPPSSSAAA
AAAEEKKAPSTSSSASPSSRSSSSSSSAKKGKPMLDQLLAEYQDLDLTGDR
PPRAKAGRAQKQDKPAATAVDGDDDDGDEHSASPQAQQRGPRPGDFTLDDL
LAEYDGPAGVGGPVKAQAQAKATTDPLAGREGDFTLDDLLADYDDVSGPA
AEGEDKGDARRLVQGVLAGKQADPRLAQLQGDFTLDDLLSAYDHPRRR
PRQRRPGGGDEWREDMARFGGAPDEREASLEDLVKTYQPPPAEDQSPAPA
QGASSKARVREELRHVQSMVADDLGLTSGGEADEREAMEPLRATRARAR
RHPPQATATTQHDDGVRAADVVDVDDDDADYDDAEAEEDVLRHEAVDDVDI
DEQKDDDDDELDDDEEKALDKMMAEAEAEEDQDMELLTRELEAGRTNLPPR
RTREEWERLGGDDVEPAFVGTCIERSESELEGEVETKRFDAVNDNETDEQ
YMERMIQMEEEAEMRRAEGELTGKKQGTVPVRLTKKEKAEAKLATVDPKA
KNARTKLRKERRRQEHQHEKVKRHHKEGGVKLDGSGMFAGANVPPE
LMHWFSDPEGLKTKRDAMLAKKCNKLQMAELRKTREKRIDMPSDPTTAE
VEELLGMTRRQKRMASKIQSHLEDVVFVEDLGADSALGQLDININEVMSR
DMRIARVYWKSGNEDAAERLLHARTSMRLRGLAQRAGVRFSPELVFYRE
SAESMYHEVSRMYDQLQKQEGWDEATIKAELKRIRKEREAAAAGGRTKT

>Mitochondrial mRNA processing protein PET127, g6336.t1 (incomplete N-terminus) putative n=2
Tax=Aspergillus fumigatus RepID=B0YC, 608 aa, 423 .

MQPSPEASGATTTSTSTSSSTSSSSASDIPYHPWFYRESSEAGYVYRQ
HRPPKPATSPRTVQPLPELSALWKPLELTLPOQPVPPEELPIARLAHGL
KRVVFNEGVHPLRDFQTHKPLFSFPLRKIHQPAEINYDSMPFVTTSKDK
ILHKLAGEQGVKYKGSTSSMSPTLSQIYRLLSNFKPPDFDCFSPOFKNMV
RPRQGTSRKSSIKPVCLFLRPVDGIYSIDSDPGNIPVATNQILLDLGKSL
ERMLTTEPEEFNQEFLLSGKPTGKYSGPEAYHYLKMDFLLRSQDCYEP
DLPGEHKTFFDLKTRATMGVRVNMGOYEKHLKYRLTCVVRGNMFSYEREFYD
ICRAAMLKYWYQVRIGKMHGIFVCYHNTREVFGEYMKQEEMERILFGNS
EFADVAFQAPMRLWQNILNEITSRAPPKTLRVVLNAEKLTKTLDFVVEP
VEFDGWLPEPTSLTSGDMGMISARTSGYGLSSTIQRSSMLDPTNPPNP
ELELERYEWLKTGKLQPSLNHYKHLHLATFLNGRPTLDHPISHRVGDRLDV
FYNLEDVTGTTEKSVMRMLKSYKALLSKSMLYEWALDEEEDAEENEDER
KNETINNV

>Aca_m-S33_[114_aa]_{Contig20706} NOT DETECTED, 114 aa, 13E0 .

MAVYLRSLRPAFEASAKIFGQRIGNGEHSGFKYLQALRKGEAMMKWYQED
LDQMKFPGWVSERRERKIIRTASRAERGKAPRPKKGFGKIALRREKEEKR
LAAKAAKGGGKKA

>Aca_Ribosomal small subunit Rsm22 NOT DETECTED, 455 aa, 1CE4 .

MEELAEDELGDSTDREDEDDDFVEPAAGRSEEEGERREGSGADDEHTGAL
GDAIEGSAEAHDWLDSSRRMRTQHDHLKPMKLGVMVLPDALKKRIGATLKK
VGRTGLRRTAGLLSERLRERTRLVPGQPKSVMPDPVVYRREEAMAYAAHR
LPGVYACTFRVMNEIKLRRPEWTPKSLMDFGSGPGTAVWSATEVWPELDN
ILAVEPSEAMIDVANELCKGKPKVWKRLLPDTPNERYDVVIASVVLSELP
SEEERLAKINALWKHTRRGIMILVEPGTPVGFHNIKLARSTVLGDLSHG
LPNVLAPRMQRIPVQTSKQNTSINWEDEKFSYVVLTKGNLLVDKNEAFD
RLLISPQKRGRHVIMKTCGQDGEITTRIIAKSDGAIYKEARKAAWGDGFG
FPDETSKRDLGKFRPKAFKYRSLSKVGSVAPATRSRRPPKPAADADAKE
DSQVM

>Aca_m-L17_[158_aa]_{Contig23660} NOT DETECTED, 158 aa, 1FF8 .

MRHLVAYRKLRSRPAHRQALLRNLVNAVVKHGRIRTTVPKAKELRRLAEK
AVTLAKRGDANARRSAHGMLYEKPVLDKLFREMPTRMADRNGGYTRIVRL
GRRLDGNAEMCYVGFVDVLPDRHDIEQRALEEKQRNTPPTDGKEQAIDDL
FLLSSPTQ

>Aca_m-L21_[234/236_aa]_{Contig3572} NOT DETECTED, 236 aa, 114D .

MKSLRSVAGASGVPRPTLPLPFALQTRNATAGSSLLRTMSAAHFTSSTQ
FASRPLSIFSKRQLHSSGPVSQESASQSAEVPAPTSPSYIVPESFAVVYVG

GROYKVKSGDTIITERLLVDVGETIHLKVKVLMVGSQNFTAIGQPLLTEAV
VEAVVEEQTRAAKVVVFKFKRRKNHRRRTRGHRQDITVLRISNVKLGYEQA
APDATTGLSTVFGPTVPASAVLPPRASDDLKESILS
>Aca_m-L30_[83_aa]_{Contig8613} NOT DETECTED, 83 aa, 17C3 .
MATAAGKAAVELGVRQVRVVGVTLLRSGWHQRPEVRSTLRALGLTKINR
TIYHKNIQPIRGQLLKVKHLIAVHPLKEPESIV
>Aca_m-L32_[109_aa]_{Contig2995} NOT DETECTED, 109 aa, 1F59 .
MQRFTSLGRLTAPRLPVGNLPAAVGSSFGRAVGDVFLPSPGAIGGVGDAG
GLITCAVPKVKVSISSRSRKRHRHAPGKERRNISHIEDCRACGQPKLRQHVC
LNCAGWKVS
>Aca_m-L33_[62_aa]_{4945} NOT DETECTED, 62 aa, 1FAC .
MAKGKKNQNMVVKLVSMAGTGGFFYTTTKNPAATRKLFLQKYDPMGLSHTL
FREERISRSKRK
>Aca_m-L34_[163/165_aa] {Contig10244} NOT DETECTED, 165 aa, 1B11 .
MQRVLLGSRGAVAFRPAACGPSSSRFLFTSSANVESAFNSEAFALPTLS
TSSPLPVSRRPVAAFNDVSPRMAMVPLDRTLEIAPTDRSDVLDMLVPTLP
LPLDLADAPLYPTTEGNAPMEAVKRTYQPSTVVKRHHTHGFLVVRMRSKNGRR
TIARKKEKGRHKLGI
>Aca-m-L36_[45_aa]_{Contig17064, Contig15385} NOT DETECTED, 45 aa, 224C .
MKYRAALRRLCIHCRFVLRERRLFVVCASANQRHKQOQIKAKKSSD
>Aca-m-L37_[136_aa] {Contig24432} NOT DETECTED, 136 aa, 632 .
MNRSMQQORRAVSSAASRTILARPVVGHAGALPLLSQORRGYAKPGGRGG
AVERPKKEYKGDQYIDEMLVAYGLQVPTWDDIPEWVWKAEVDLKPKKELT
PLDGPLFWKKINRERIKQFNEDRTWGVVKGKWKLG
>Aca_m-L38_[323_aa]_{Contig24117} NOT DETECTED, 323 aa, 508 .
MLRRTSTTLQACRLGSRSAAPSASNLGVSRTALQSARTSRILQNSPVL
ASEGARRNISQIQIAKKNRAALRSKPKWSPFPGWKPDPETEISKHKNI
RKKAKLHKVFPDVLPEPFTPYFMLQVEYSPEVVVRKGFHLPKAKLQTTPR
VSWPSDPTKRWTILMATPDDPDLVPDDVEIVKPGWVVPHEEGRGKEVLHW
LVTNIPGNDISQOQVLCDYLPMPKKGHGGHRYVFLLEQLDGETQFDFAF
PEGALKEFVQRKNWNTFKAQEKYGLRPKALAFFKATWDAEVSAQYSRLNL
PEPEGPAPVNLIRPRKWLDPKSW
>Aca_m-L41_[112/113_aa]_{Contig9854} NOT DETECTED, 113 aa, E62 .
MGVFGSLTRRTGLGARTTVGTTGFTRGKVSANQKGRLTPKMTPRGWYKG
VGCRLGFFHTKGGYRVVPAKVPVYVIPDLTDCPLKPYVSWSTKRVVVQP
LRLPSDKELAQLE
>Aca_m-L53_[149_aa]_{Contig23150} NOT DETECTED, 149 aa, 1AE1 .
MSTARQWRLAQKRAHYELHFLKSLNFSFCPANNHPTVAEFLRRVSNDR
RASNPDCEVTVEVLEDIKARPSIKAVYENGMEEVLDHEQORVEHLVDDFWK
RRERLQFFSELQFSMKNEEEENPFLNAAAQDEARKKQOQAKGKAGK
>Aca_m-L54_[136_aa]_{Contig24432} → possible homolog id by hmmer NOT DETECTED, 136 aa,
632 .
MNRSMQQORRAVSSAASRTILARPVVGHAGALPLLSQORRGYAKPGGRGG
AVERPKKEYKGDQYIDEMLVAYGLQVPTWDDIPEWVWKAEVDLKPKKELT
PLDGPLFWKKINRERIKQFNEDRTWGVVKGKWKLG
>Fe hydrogenase NOT DETECTED, 688 aa, 618 .
MKKHMTMRRLPWNAVARSLGCTTTSRCAAGLRVVRGTATTSSGTGSAASAA
DRLKQLTQLKADFPLGKEVEKLVNVTINGTPRQVAEGLSIVEACRLFK
VEVPTLCYHPLMTIVGQCRCLVQLANRQGLVAGCATHVEEGMDIITES
ERTHHTVLSNLQLLRCRHPNACMTCEADGQCEFQNLVYKFKVEELLPOYH
RDRSHTRDASSHSVVRDMNKCVLCTRCIRACSQVQGMHILGMIGRAEEE
VSTINFLPLNQTACISCGQCTAVCPV GALVEKTHSHEVESGLRHKAKEGK
IYVAHTAPAVRVAIQEEFNMAPGTVSTGKLVSALKELGFDYVFDTNFAAD
VTIMEEGYEFIRRLQSNLPMFTSCCPGWVNLVEKSYPELMPHLSTCKS
PQGMGTLVKTYFAQKLGVPDRIVTVSIMPVAKKDEIARPQLSMEIED
SKGEKRRVQDQDYVLTRELGHLIQRHRIPFVSLEDKEFDTPGLTTGAG
ALFGATGGVMEALRTTYEVMTGKPLPKVELTAVRGYS AVRKASVMIGDR
EVRVAVVHGTTKVVREVEAVLSGDKSFDLIEVMACPGGCVGGGGEPKTS
VDPDVIRKRIDAIYSIDERKQLRRSHENPAVQELYKTFLRGEPNSAEAAH
LLHHTYDRTSEVKSRSDEAFGRVPPALELKPADHPHH

>HydE NOT DETECTED, 464 aa, 3E8 .
MLLRRGLHDAAPVSGSPTDLHKLLYSPLLPDAPAAEDPVAARNALQAA
LSRGGKKRPVNRHLPLEEKIAPLASLVGRALTPGEPLTLEKEQIVHILQ
TADEGLENALYDYADTVTQRYFGKKIYFRGIVEFSNVCTKDCYCYCGIRKH
IQVKRYSMSKEEIVNCAIFAYESGYGSLMLQSGELPTDKRLAFLVDTIKE
VKRVTIEKDKERNVKEGVENKGLGVAISLGELTKDMYRSLREAGAHRYL
LRIESSNPDLIYAKLHPADHSWHRRVECLRELNELGFQVGTGVMIGLPGQT
LGDLAQDLLFFKSMIDMIGMPYICQKDTPLGKLWQEENSNDTKRYNE
KLFRLSTKMVALARIICGDVNI AATTALQAINPQGREVALNQGANMLMPI
LTPRKYREDYQLYEGKPCIDEGAEECRSCLKNRVQWAGKELTLHDWGDPO
HFFSRQNIPTPTRQ

>HydG NOT DETECTED, 554 aa, 64B .
MFPKRLGVSAFRYSVVSTTCRTTGLLAHAPARGLRWSVELETHKLASAE
EINEGEINRQLEETRELARDPLVIRDILAKAKQORALLQDIKGPVGEYV
QGLNLRRETATLLNLDHAHSKEPENQKLMQELFDTALAIKQNIYGNRIVLFA
PLYLANYCVNSCTYCAFRGANTHIARSALTKEELIQEVQVLEKQGHRRLL
ILTGEHPKYTFDQFLDALHTVKNVKTDPGDIRRINVEIPALSVSDFRRL
KATDCVGTYYTLFQETYHRESFRRFHPSPGPKSDYEHRLQTM DRAQIAGIDD
VGIGALFGLHDYRFEVMGLMHAQHLESTFGAGPHTISIPMRPADHAPD
AIRPPAPVDEDFKFLVAILRCAVPYTGMIILSTRESAEMRRELLNLGVSQ
MSAGSKTEVGSYHKGVDVGAEAKIKQOFEAEDQKEKKEKESGQFELADER
STEDVVKDLMQMGFLPSWCTACYRLGRTGEAFMKI AKKGD IQNFCHPNSL
LTLQEYKDYASEDTKTLGEAVIARERETISARAKKALDKKLRIDAGER
DLYF

>g10463.t1 - Mitochondrial substrate carrier family protein B n=1 Tax=Dictyostelium discoideum
RepID=MCFB_DICDI, 153 aa, 76E NOT , 153 aa, 76E .
MQYRYKIGDCLVQIVKQEGPLAFWKGLSVSLVGIAPFVAINFTTFFETLR
QEVTERHGGQMPLLWGPVCGAASGTFAMTCTYFPDLLRRRMMMLQGRGGEE
RFYSSIWDACRKIHQFEGVGGFFKGMIPYTLKVVPSVAISFGTYELCKRV
GGE

>g11200.t1 - similar to solute carrier family 25, member 14, 227 aa, 2615 NOT DETECTED., 227
aa, 2615 .
IALGQKILAGLMSGAIAAAVFNPTDVLKVRVQADPARTPELRRYKSVVSA
VVEIVRTEGLIAGLYKAHRSCCAPVVVGFADFTLASVAITTPASVVQGV
GTTVIRASLLTSAQMASYDESKHFLIDSLAFSDNFLTHFCMFSGFMTSLV
TNPVDVVRTRIMTEYASPGQPRTYSNPFTSLVRIFRAEGLVGLYKGFVPS
YLRLGSASVVVFMLEYQLRRLAGIPTL

>g11233.t1 - Mitochondrial carrier protein n=1 Tax=Tetrahymena thermophila
RepID=UPI000150A9DC, 302 aa, 1E96 NOT DETECTED., 302 b, 302 aa, 1E96 .
MLASATASMLSRI PCHPLD TVKAKLQVGAQGGRLRGVLRHTLRTEGLRGVL
RHTLRTEGLRGLYRGFGAAVVGSGPAGCLYFTTYELAKKGLLSVGLVGQS
PFLAHFGAGLLAELVSCALWVPIDVVKERMVQVSTLAAGKPSYAYTGDLH
AAATILRTEGLRGLYRGYATVLSFGPFSALYFVYYEQLKGLAEAFSASN
DSSTSASTRRPPPELHSTNVLDMAKLRMQUVERAGGERTFGYTNV FHGVAR
IVSDEGWRGIFRGAGARIAFQAPTTAIALAAFERLKVGYGR LAERVPAGR
EK

>g12189.t1 - PREDICTED: similar to Graves disease carrier protein from bovine heart
mitochondria n=1 Tax=Monodelphis domestica Re, 322 aa, 15CE .
MERRGVRLASIREERKIESWSQLWR TLLCGGIAGCVAKTTTAPLERVKIL
FQVATIHYPFKGVVPTLRRIVEREGFRGLYKGNVSSLVRIFPYAATQFAA
FDIFKAALTPKDAGISGLANFLAGAGAGATAVAFTYPLDVTRARLAVQVE
KRHYTGLVHAIQNMWRHEGGLKALYRGLQPTMFGILPYAGINFFTYDTLK
WYYSKKLR IAANGPPPIPTTLRLAFGAVAGALGQTLTYPLDVVRRMQ
IDGLMLEQTYSYKYLSTWHGLQ TIAQKEGWRTLFRGLHINYIKVVPLVSV
SFTINDLMRRWMLKTEGGVER

>g12685.t1 - mitochondrial carrier protein , 160 aa, 1E7B NOT DETECTED., 160 aa, 1E7B .
MSAAAAGGAQQQLRYRGTMHCMHTLTKTEGIPGLWRGWSTFFRELVGTA
LWFGTYETVCR TLAADAGDASDAPPLIQLLAGGCAGCIYWGVPYPMDTVK
SRIQIKQATEQTSIFKMFVEVLKEEGVRGLYRGIGPTMMRAIPGSAVALF
TYERSRMLL

>g13824.t1 - Mitochondrial carrier protein, putative n=1 Tax=Ricinus communis
RepID=B9SLZ3_RICCO, 366 aa, 1BB5 NOT DETECTED., 366, 366 aa, 1BB5 .

MVQSSLSSHIDRNEHQRPIDESGWETTLRVVASLGSMSGAVHGLGVLG
YQWSVTPRPVIGQWLPAITFSAYWNLKKEAATHPALRYGAFSAAGLWAGA
VKAISHMFLQEVYVAPLLHIATSDEELDESFEAYDPIAPIKQGDYQYRGG
WTGPVCAMASAAGAILLSLPLFLKAKFLLRTQSRPDGFSKYTGFLQSIGV
IWREQRWGLITNNLSAFISFGILYDVVGFLLTDYKLLSTLIMKIKTREGVL
HNLKHALALAIIDIIIEAAIIVAFVMPVNVVLRRLLEAQAGGVAPFSADGVV
YAGVLDVCVQQIFVREGLNGFYKGAPELLERLLPRVVFDAAVDLGLNSFF
GVPATYDNTTEHSHAD

>g14367.t1 - Mitochondrial carrier protein, putative n=2 Tax=Paramecium tetraurelia
RepID=Q6BFJ6_PARTE, 109 aa, 1D35 NOT DETECTED, 109 aa, 1D35 .

MAGLGAWVTSYPMDFIKTQLQAEPEASSKYQKNRFLLDGGFVDCFRQHVR
ENGWRSVWKGFGPCVSRAFPANAAGFLAYEIAAKLIRERQEEKDAAASAS
TAAAATTAQ

>g14606.t1 - Mitochondrial carrier protein, putative n=4 Tax=Trichocomaceae
RepID=B6Q8S5_PENMQ, 114 aa, ED7 NOT DETECTED., 114 aa, 114 aa, ED7 .

MEKARGQATTTAGRPKYSVLRGLGVILREEGWRGFRGHLLTNLLHVAPAA
AARFYSFEAYRSWLVRDQKPLPPLKRMLCGALAGITSTTLTYPLDLVRTR
LAAQTPDTPMQYRA

>g14859.t1 - Mitochondrial carrier protein n=1 Tax=Tetrahymena thermophila
RepID=UPI000150A9DC, 230 aa, 1715 NOT DETECTED., 230 aa, 230 aa, 1715 .

MSSRQPGDRPKREEGPPHALLPPSLTVTMLASATASMLSRIPCHPLDTV
KAKLQVQALGGLRGLRHTLSTEGLRGLYRGGAAAFVSGPAGCLYFTTY
ELAKKGLLSVGPVQSPFLAHFGAGLLAELVSCALWVPIDVVKERMVQVS
TLAAGKPSYAYTGDHLAAATILRTEGLRGLYRGGATVLSFGPFSALYFV
YYEQLKGLAEAFSASNDSSASTRRPPPE

>g16015.t1 - Mitochondrial substrate carrier family protein B n=1 Tax=Dictyostelium discoideum
RepID=MCFB_DICDI, 273 aa, 2066 NOT, 273 aa, 2066 .

MDKSGPRYQGVLPALRKIWAEEGFRAYWKNGTENVIRIMPSDAARFYSYD
TFKKLISTPGEPIPTMIRIMAGGLAGMVSTIATYPLDLVRTRLSVQTLPG
RGAIYAARYRGMWHCLGSIFREEGFFALYKMGVSIILGVAPYVAINFASY
ETLKQLVKTDGSETHALEGLVMGGLSGTAAVTLTYPERPPATNDDARYR
RCFEYVQRLVGRVRQDRAGGGRVLPRAHSLLPQGGARRRDRVGLHRDA
AEGRTRPLLLHQLKPQAWRFTIL

>g2171.t1 - ADP,ATP carrier protein, 287 aa, CA NOT DETECTED., 287 aa, CA .

MTLASPIQRVKLLQQAEMVAAGRLLRYPYAGVRDCFRVVAEEGLLSLW
RGNVAHLLRIVSANALGFVLKDRFKNVLHVDKERDGYWKMAANVASGSA
AGLVTVLCTFSLDYARTRLASDVLVAHPGAGGERQFSGLMDVYSQVQRR
EGTAGLYRGFTSSCIGVILYRGLYFGLYDSLKPVVLRGDLKDSVVASFVL
GWGITLGAALAAFPDITCRRMMTTTSEHRGGLMHTLTEIVEREGLAALW
KGAAANVAYALIGALLVAYDRAVARAAAPKRTKDNN

>g4338.t1 - similar to solute carrier family 25, member 44, 350 aa, 2B7 NOT DETECTED., 350 aa, 2B7 .

MESEINSAEGSVFVGDDIDWQSLDKSKFYVVSFAFFMGVRTLVIYPPILIK
TRLQVQOGRSQYRGTDFDAFRKIFKLEGFGRGFYKGFLLTTSVGIIPGQFMYI
TVYEFVRHRIKSFFPTDSLRYDALRNFIGGGTASLASSLVSVPLDVISQL
LMIQDGTVNKRKYSGGINAFCEIILKTEGVRGLYRGTASMMVYVPSSGIW
WGTYASVKGKAAAFINEQGGLLRQLDVLVFGLCGILAGSTAVVVTNPMDV
VKTRLQVLAYTNPSSPATSAEAAPPPPQQOTTSSQRPIASSTRSYQKI
SIWATAKQLLREEGWRAKKGITARMMSPVSVFLMTITYEQAKRLSVKQ

>g6456.t1 - Mitochondrial substrate carrier family protein X n=1 Tax=Dictyostelium discoideum
RepID=MCFX_DICDI, 320 aa, A11 NOT D, 320 aa, A11 .

MSTHKQSGQKAPFLAQTKEEGQASSRQATLPGERQKQVPLGAKLVVGG
IAGMVGMTSVFPIDLVKTRQLQNRWCESVSECVRRIRAEEGWRGFRGLGA
CLCGVAPEKAVKLAVNEKLRHLAARDHDQITLAHEILAGAGAGMCQATV
SNPAEIVKIRLQVQSEPNANGTKSAVGIVRELGFRGLYKGLPATLLRDVP
FSFLFFPIYSNIRQAWLHQGRDAKGEVGLLPTLAAGAAAGAVAAAATPA
DVVKTRYQVEHSPYTSLHQCARAVWREGGPRFTFFKGAVERMAIQAPLYGV
ALLAFELQKKWYLAHQOQD

>g742.t1 - PREDICTED: similar to solute carrier family 25, member 14 n=1 Tax=Hydra magnipapillata RepID=UPI00019247B0, 321 aa, 17, 321 aa, 176F .

MTSSAAEGEVVHKFLLLGMVSCAAPLPRGGHAVVNPVIEVIKTRLQLQGELO
EEKAKSGLSRIYKGERKYKGFMHGGVQILRDEGIAGLYKGVPAALRECS
YAAIRLALYDPIKTLLENRADGVKDGGLPFWKKLVAGATAGSIGAAIAT
PTDVLKVRMQAEGARDKPRYKNTLEGFVTIARTEGIRGLYKGVVPTTQRA
CILSAAMSSYDHSKHFILQKGIKHDNLYAHICAGMMAGFSMAVVSTPI
DVVKTRIMNRSAGGPAPYRGMFDCLVKTAQAEGVLGLYKGFVPTFLRLGP
HTILAFTHIYEELRWKAGIRPV

>g7680.t1 - similar to Solute carrier family 25 (mitochondrial carrier, 496 aa, 116A NOT DETECTED., 496 aa, 116A .

MTIQPNGQAREEANVERYRRVFEELDLSSGSLDEHEIQAGLQRMGLPSS
DKQVHEMMVMVDADKDGSIYGEFASFAQEREKDLRELFDEIDLNGNGLI
EANEIRIALEKIGMSTEVSDGDIIQLIDRMGVDHTKGIQDFQVFKRILMLF
PSRNISISNVFDYWRHIIDTGEVVIIPDIIDHPMKRLIAGGIAAPFDRL
KMLLQAQNSSAMLAGVATKQLAGGKPAARPGVIRPAPDAAARAAAPEY
RGIWNSLKKIYFESGWKGFYRNGTNIKIAPESAVKFWAYESIKRMLCR
DSSAPAIKEKLIAGSAAGAISQTAIYPLEITKTRLAVSAPGEYRGIMHCI
SSIVRTDGVSAFLRGLLPSVVGVIPIYAGVDFAVYSTLRDVYTRRYPNTHP
GVLTVFVCGAISSTCGQVAYPLQLVRLRQLTQGMAGRPMLYNGMSDAFF
KIWKCDGLLGFYSGILPNFMKAIPAVSISYIVYEQVSRGMGISSGL

>g845.t1 - Mitochondrial carrier protein ymc n=1 Tax=Aedes aegypti RepID=Q16YM1_AEDAE, 299 aa, 2271 NOT DETECTED., 299 aa, 227, 299 aa, 2271 .

MSDGGNGQDSGQASRERVRHFLAGTLAGAAGVFGHPFDTVKVRQLQTLPG
AFKGPLQCFVQTLRREGVRGLYKGLSSPITGDSLNTCLVFGVYGLTRRAQ
LETDQTEEHFPLLSLTQIGLAGGAAGLIGGLIQSPVELIKIKLQVQTGTG
AARLYSGPIDCIKKIVRADGLKGLTRGLGATWWRDIPGFGAYFVVYEGLR
RQWADDPLRPDDVSPVKQLVAGGLGGIAAWGISYPIDVIKSRQLQTNPEYK
GMWDCAVKSYRADGHRVFFTTGLGTTVVRSPVNAVIFLVYEVLIKTIPA

>g8625.t1 - PREDICTED: similar to solute carrier family 25 (mitochondrial carrier, Aralar), member 12, partial n=1 Tax=Ciona inte, 613 aa, 35E .

MSTTATASIPVPVEQSPATLDKIKGNIANEIRREIQQRPSLMNAAVSFS
IGAAASAIASATAVYPIDLVKTRLQNRQTLVQQRGMSMWGCLTSVVRHEGP
FGLYKGLIPQLVGVQVPEKAIRLFIVDRIRSLSATDGVNYSIELLAGLAAG
ANQVLITNPAEIVKVRMQVQGEYAKKKADAAATTAKPKGAMSILRELGI
KGMYKGASACFLRDVPFSGIYFGSYAWLKEQLRTGNEPLHSIELFFCASA
CRAGVAAASLTPADVLKTRMQVEAKKGEYANLRDCYRRVTTTEGYKAL
WKGVVPRVLRSSPOYGVMLFSYELLQRLFNNNDVQSSISMDVLEDKWKVVM
LLEDKMGILLENSWEPPTSFTFRKGGFKHISVGGSTSVWAIIRGDEIYR
WKENQWEKTTGAKLKQLSAAEDGAVWGVNNGEVYRWDGDKQTWQSIPSP
FSVLHVAVGSDREWVAVIDDEKEPTSALKQSNIAATWDGKAWIRIPGSEQM
KLVSCGADGTLWAINSERFTFRLDRESLQWEQMPGSFKQVSVGSKKHVWG
VDSKDCIYKWKKTWQRVTMADVPQHISVASDGSIVAMDKDTNGVVNVKK
LVLPVEDKEGEVH

>g8908.t1 - mitochondrial carrier protein (Rim2), putative, 335 aa, 25E4 NOT DETECTED., 335 aa, 25E4 .

MVGGVLTSPLEVVKTRLQGTYNKEGLANTQHRFGTRTVRALTHLHAEGGV
RALYRGMVPHLVGVVPARALYFGTFSWSKRELKQLELDPHGMLLNWLGG
VAAGVVTITLTSPVWVVKTRLQLQVNGKGSVAPGTGVRQYRSAWHCVTS
MMRHEGPRAFFAGMSASYVGISESSIQFMLEYKFKSLIKQRRVRQCEEA
QLELQRRRAHTAAAPFDHADPELVQLRELARSAAASSIGAWETLGISS
TAKLVASVLTYPHEVARTRLREERAKELYGGFVNCVVRVWRNEGLAGLYG
GLAPHLRLRVVPAAIMFLTYEAVLSLFDPSSTR

>g9258.t1 - Mitochondrial substrate carrier family protein B n=1 Tax=Dictyostelium discoideum RepID=MCFB_DICDI, 365 aa, EA6 NOT D, 364 aa, 873 .

MAAAREERQVVFGEGESEEEAVQRRQOIRALFDEFDRDKRGLTSLQFEQ
GLKREGLWHRIRDDAHLRRVWNATRLDPSDEAGMDFGEFYNNIMVEHYQIL
PHAHLVEVFEVQYLSKAGPQYGGVWSALVAMGRNEGWRGYFKNGVNIIL
RIMPSSAARYYAYEALKRALHPENQOPTAGVRMLSGALAGIFATGSTYPL
DLVRLRLAAQTASAKYKGLMDATRTIVKEEGVAGLYKGLWTSCLGVAPFV

AINFTSYEMLRQW AIDARQGEKPSLFMNL SIGALAGTIAMSITYPSELLR
RRMMLQGIGGAEREYKGITDAVVKIARNEGVAGFYRGIVPCYLKVVPSQA
VSWGMLELCKKLAG