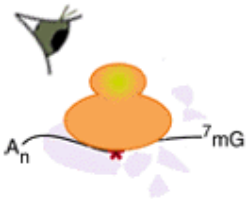
NATIONAL CENTER FOR BIOTECHNOLOGY INFORMATION
NCBI Coffee Break

13 Oct 1999

Article reference: CB7.131099
Coffee Break archives



Surveillance mechanisms detect the presence of premature stop codons in mRNA and target those mRNAs for destruction by a process known as nonsense-mediated decay.

Click here for the full figure and more details.

RNA surveillance: watching the defectives

Only those changes in DNA sequence that have functional consequences are known as disease-causing mutations. One such frequently occurring mutation causes a premature stop codon to appear in the middle of a protein-coding sequence of messenger RNA (mRNA). Stop codons (a triplet of nucleotides: UAA, UAG or UGA) normally signal the end of the stretch of mRNA that is translated into protein, so when one appears early, the result can be a truncated protein that could have nasty consequences for the host organism.

However, a mechanism known as 'nonsense mediated decay' has evolved to detect these harmful RNAs, and sequence analysis suggests that it may have been conserved in eukaryotic organisms, including humans. In yeast, three proteins have been identified that are required to seek and destroy the partly translated RNAs: Upf1p, Upf2p, and Upf3p.

Upf1p is an RNA unwinding enzyme - a helicase - that requires ATP for activity. Unfortunately, Upf1p will unwind pretty much anything, not just the problem mRNAs. So Upf2p and Upf3p are thought to be required to help Upf1p discriminate between nonsense and 'real' mRNAs.

How do the core proteins work in synergy to trigger nonsense-mediated decay? One possibility is that Upf3p, along with several other ribonuclear proteins, may first bind to an mRNA as it is being exported from the nucleus en route to the ribosome, the site of protein synthesis. If the mRNA is fully translated into protein, Upf3p and the other protein factors are displaced. However, if there is a premature stop codon, Upf3p and cohorts may sit tight and mark the mutant mRNA as one that needs disposing of. Experiments have shown that Upf3p can bind Upf2p. Once bound, Upf2p could signal to the 'termination complex' - a mixed bag of termination factors that includes Upf1p. This results in the release of the incomplete polypeptide from the ribosome, mRNA unwinding by Upf1p, and, exposure of the mRNA for total degradation by exonuclease.

While this model is attractive, more experiments are required to shown that this actually happens in a living yeast cell.

Many of the mutations that form a premature stop codon lead to human disease, for example, those in *BRCA1* that lead to breast cancer, or those in *NF1* that lead to neurofibromatosis type 1, to name just two. There are two ways by which nonsense mediated decay can play a role in the disease process. The first occurs when the machinery is functioning correctly: if mutant mRNAs are removed, then there will be a reduction in the amount of mRNA and protein available in the cell. The second is when a mutation occurs in the nonsense mediated decay process itself, such as a mutation in

Live PubMed searches
|| Upf1p || nonsense mediated decay || premature stop codon AND disease || REVIEWS ||

BLAST tutorial
|| Find relatives of yeast Upf1p || VIEW ANALYSIS ||

Related resources
|| GeneMap || OMIM || LocusLink ||

Figure Coffee Break

RNA surveillance: watching the defectives

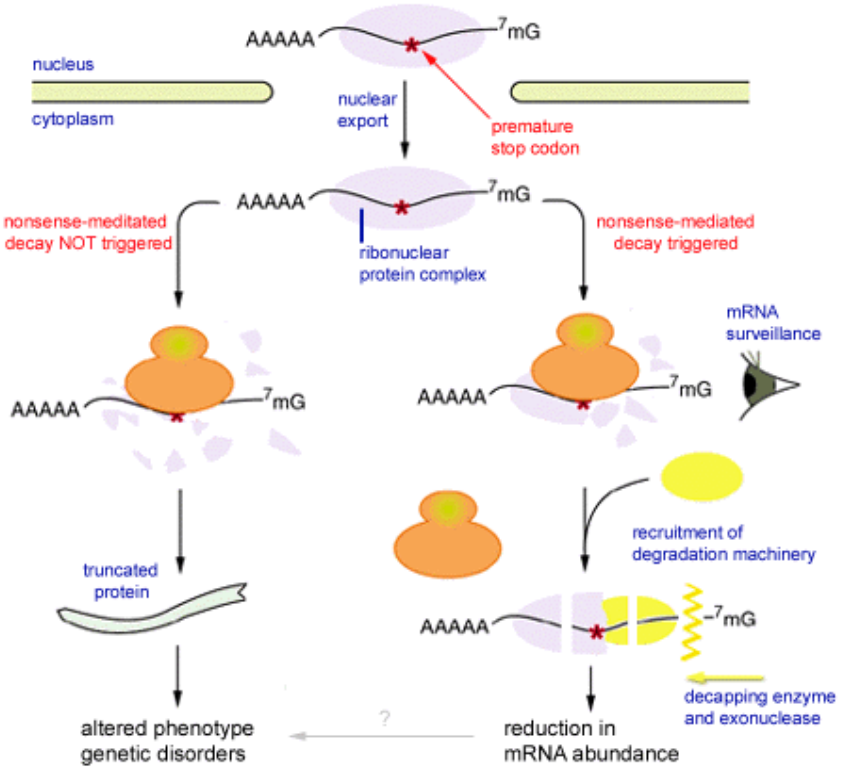


Figure 1. Nonsense-mediated decay (NMD) in yeast, as a model for NMD in humans. Ribonuclear proteins that bind to mRNAs in the nucleus remain associated with the mRNA as it becomes attached to the ribosome. When a premature stop codon is present, one of these proteins could be Upf3p. If Upf3p, or another as yet unidentified factor, is recognized by the surveillance complex (represented here by the eye), then the NMD mechanism is triggered. In yeast, this trigger may be assisted by the binding of Upf2p to Upf3p, after which the Upf1p helicase unwinds the mRNA, leaving it open for degradation by a decapping enzyme and exonuclease. Should the premature stop codon not be recognized, translation of the mRNA proceeds and results in the production of a truncated protein.

NCBI

Tutorial **Coffee Break**

This is the query page for a BLAST search. The sequence of yeast Upf1p is entered in the textbox below. The database to be searched can be selected from the following pull-down menu, as can the format that the sequence is submitted in. Click on the "Submit Query" button below the textbox to find similar sequences to this query in the database.

Database nr

The amino acid query sequence is filtered for low complexity regions by default.

Enter here your **amino acid sequence** as Sequence in FASTA format

```
GSLDKDLYIKVEV SVD FQGREK
DYIILSCVRANEQQAIGFLRDPRRLNVGLTRAKYGLVILGNPRSLA
RNTLWNHLLIHFREKGCLVEGTLD
NLQLCTVQLVRPQPRKTERPMNAQFNVESEMGDFPKFQDFDAQSMV
SFSGQIGDFGNAFVDNTELSSYIN
NEYWNFENFKSAFSQKQNRNEIDDRNLYQEEASHLNSNFARELQRE
EQKHELSKDFSNLGI
```

Submit Query Reset

Please read about FASTA format description

The options below are for advanced users. In this example, the number of descriptions has been set to the maximum of 500, with upto 500 alignments returned. If non of the other parameters below are changed or selected, then BLAST is run in default mode.

Advanced options for the BLAST server:

Expect 10 Filter default ☐ NCBI-gi ☑ Graphical Overview

Descriptions 500 Alignments 500

Expect value for inclusion in PSI-BLAST iteration 1 0.001

Other advanced options:

Submit Query Reset

Query= gi|400350|sp|P30771|NAM7_YEAST NAM7 PROTEIN (NONSENSE-MEDIATED MRNA DECAY PROTEIN 1) (UP-FRAMESHIFT SUPPRESSOR 1) (971 letters) **Database:** Non-redundant GenBank CDS translations+PDB+SwissProt+SPupdate+PIR 411,205 sequences; 125,914,940 total letters

Below are the results from a PSI-BLAST (see Altschul et al., 1997) search of the non-redundant database using Upf1p as the query sequence.

The E value is a statistical measure of likelihood that the sequences listed below are truly similar to the query, rather than found by chance alone. The lower the E value, the greater the confidence that the protein found is a biologically significant match.

E-value threshold for inclusion in PSI-Blast iteration 1: 0.001
E-value threshold for inclusion in PSI-Blast iteration 2: 0.001

PSI-BLAST constructs a similarity matrix from the significant hits found in a BLAST search and uses this matrix to search the database in a second round of BLASTing. Each round of PSI-BLAST that uses the matrix to search the database is called an iteration.

Distribution of 132 Blast Hits on the Query Sequence

Mouse-over to show defline and

Color Key for Alignment Scores

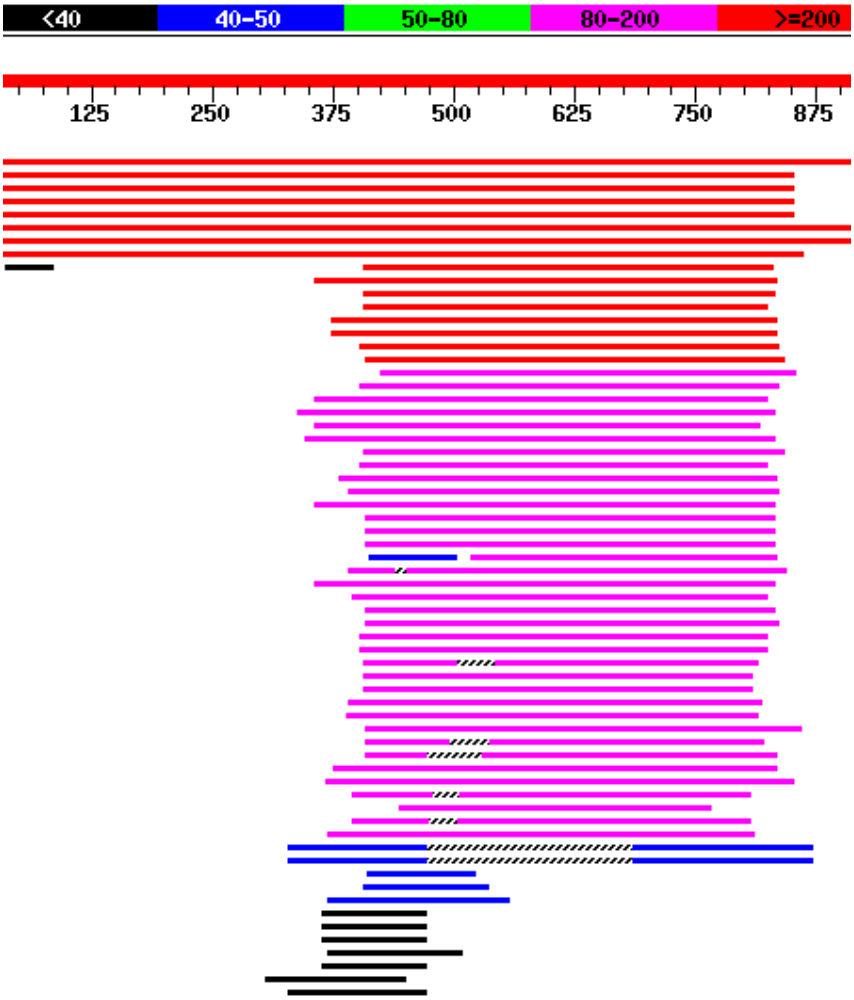


Figure legend

The above summary of the BLAST results represents proteins that match the query sequence as colored bars, with the most similar hit uppermost and appearing in red. Pink, green, blue and black bars follow, representing proteins in decreasing order of similarity. Hatched areas indicate a gap in similarity i.e., two or more distinct regions of similarity were found within the same protein hit. Moving the mouse over the bars will display the name of the matching protein found in the textbox above.

Click here to view alignments.
Click here to see the analysis of this BLAST result.

BLAST Alignments **Coffee Break**

About the alignments

There are many significant hits to Upf1p, although few that have high similarity over the complete sequence. Most hits correspond to the helicase region of the protein.

The match with the top score is to itself (reassuringly) - the lower the e-value or the higher the score, the higher the chance that the similarity is biologically relevant. Clicking on the scores will take you to the alignments for each of the top ten hits (as requested in our query). Clicking on the ID numbers on the left will display the sequence record at the NCBI site.

Back to BLAST result

Sequences with E-value BETTER than threshold

Sequences producing significant alignments:	Score (bits)	E Value
sp|P30771|NAM7_YEAST NAM7 PROTEIN (NONSENSE-MEDIATED MRNA DECAY...	1931	0.0
gi|2739355 (AC003972) pNORF1 [Homo sapiens] >gi|3328175 (AF0740...	920	0.0
gi|1885356 (U59323) type 1 RNA helicase pNORF1 [Homo sapiens]	919	0.0
ref|NP_002902.1|PRENT1| regulator of nonsense transcripts 1 >gi...	919	0.0
dbj|BAA19664| (D86988) KIAA0221 [Homo sapiens]	917	0.0
sp|Q09820|YAC6_SCHPO HYPOTHETICAL 105.6 KD PROTEIN C16C9.06C IN...	903	0.0
emb|CAA91194.2| (Z54366) putative regulator of nonsense transcr...	902	0.0
gb|AAC26789.1| (AF074017) nonsense-mediated mRNA decay trans-ac...	763	0.0
sp|P32644|YE06_YEAST HYPOTHETICAL 127.0 KD PROTEIN IN RAD24-BMH...	231	2e-59
emb|CAB11611| (Z98951) hypothetical protein [Schizosaccharomyce...	230	4e-59
gi|2984278 (AE000770) DNA helicase [Aquifex aeolicus]	214	2e-54
sp|Q92355|YDHA_SCHPO HYPOTHETICAL 192.5 KD PROTEIN C6G9.10C IN ...	205	2e-51
emb|CAB50143.1| (AJ248286) DNA helicase, putative [Pyrococcus a...	202	1e-50
gi|172574 (M74589) SEN1 [Saccharomyces cerevisiae]	202	1e-50
sp|Q00416|SEN1_YEAST TRNA-SPLICING ENDONUCLEASE POSITIVE EFFECT...	202	1e-50
dbj|BAA30003| (AP000004) 656aa long hypothetical DNA-binding pr...	200	5e-50
dbj|BAA31600| (AB014525) KIAA0625 protein [Homo sapiens]	199	1e-49
gi|2649188 (AE001009) DNA helicase, putative [Archaeoglobus ful...	198	2e-49
sp|Q57568|Y104_METJA HYPOTHETICAL ATP-BINDING PROTEIN MJ0104 >g...	197	3e-49
gi|3548803 (AC005313) putative DNA-binding protein [Arabidopsis...	193	5e-48
sp|P34243|YKB7_YEAST HYPOTHETICAL 78.3 KD PROTEIN IN RAM2-ATP7 ...	192	1e-47
emb|CAA22438| (AL034463) tRNA-splicing endonuclease positive ef...	190	5e-47
gi|2622761 (AE000922) transcriptional control factor (enhancer-...	188	2e-46
gi|3935147 (AC005106) T25N20.11 [Arabidopsis thaliana]	183	4e-45
gb|AAD35099.1|AE001689_5 (AE001689) DNA helicase, putative [The...	183	7e-45
emb|CAB43845.1| (AL078464) putative protein [Arabidopsis thaliana]	178	1e-43
sp|Q60560|SMB2_MESAU DNA-BINDING PROTEIN SMUBP-2 (INSULIN II GE...	178	2e-43
pir||A47500 Ig mu chain switch region binding protein 2 - human	176	5e-43
ref|NP_002171.1|PIGHMBP2| immunoglobulin mu binding protein 2 >...	176	5e-43
gi|2340994 (U21094) Sen1p [Saccharomyces cerevisiae]	176	9e-43
gi|908917 (L24544) DNA helicase [Homo sapiens]	176	9e-43
gi|3176714 (AC002392) putative tRNA-splicing endonuclease posit...	173	7e-42
sp|P40694|SMB2_MOUSE DNA-BINDING PROTEIN SMUBP-2 >gi|423421|pir...	169	6e-41
emb|CAA20978.1| (AL031629) cDNA EST yk384f3.5 comes from this g...	164	4e-39
sp|P23249|MV10_MOUSE PROTEIN MOV-10 >gi|110821|pir||A39611 prob...	158	2e-37
sp|P51530|Y083_HUMAN HYPOTHETICAL PROTEIN KIAA0083 >gi|1531548|...	154	3e-36
emb|CAB38508.1| (AL035637) hypothetical helicase [Schizosacchar...	141	2e-32
gb|AAD38528.1|AF144384_1 (AF144384) Dna2p [Schizosaccharomyces ...	141	2e-32
gb|AAD48967.1|AF147263_9 (AF147263) contains similarity to nons...	139	1e-31
emb|CAA90399.1| (Z50070) similar to DNA binding protein; cDNA E...	132	2e-29
emb|CAB54253.1| (Z50070) F43G6.1b [Caenorhabditis elegans]	132	2e-29
gi|2105497 (AF003740) similar to mammalian DNA-binding proteins...	130	4e-29
sp|P38859|DNA2_YEAST DNA REPLICATION HELICASE DNA2 >gi|626832|p...	128	2e-28
gi|3063473 (AC003981) F22O13.35 [Arabidopsis thaliana]	124	3e-27
emb|CAA20777| (AL031540) hypothetical ATP binding protein [Schi...	112	1e-23
emb|CAA93884| (Z70038) cDNA EST EMBL:D32579 comes from this gen...	109	7e-23
sp|Q09449|YQ12_CAEEL HYPOTHETICAL 175.7 KD PROTEIN C05C10.2 IN ...	109	1e-22
sp|P42694|Y054_HUMAN HYPOTHETICAL PROTEIN KIAA0054 >gi|473951|d...	103	5e-21
dbj|BAA25486| (AB011132) KIAA0560 protein [Homo sapiens]	101	3e-20
emb|CAB10335.1| (Z97339) SEN1 like protein [Arabidopsis thaliana]	99	1e-19
gi|3785995 (AC005499) unknown protein [Arabidopsis thaliana]	99	2e-19
gi|2648582 (AE000968) DNA helicase, putative [Archaeoglobus ful...	98	3e-19
gb|AAD36482.1|AE001793_12 (AE001793) helicase-related protein [...	97	7e-19
emb|CAB49033.1| (AJ248283) DNA HELICASE RELATED PROTEIN [Pyroco...	89	1e-16
emb|CAA99786| (Z75529) Similarity to C.elegans NAM7 protein (WP...	83	1e-14
gb|AAD07514.1| (AE000560) conserved hypothetical protein [Helic...	82	2e-14
dbj|BAA17223| (D90904) hypothetical protein [Synechocystis sp.]	82	3e-14
emb|CAB07226| (Z92790) H03G16.3 [Caenorhabditis elegans]	76	1e-12
dbj|BAA29178.1| (AP000001) 1188aa long hypothetical protein [Py...	71	4e-11
gi|183250 (M64979) glial factor-1 [Homo sapiens]	70	7e-11
gi|3329511 (AF075169) DNA helicase homolog [Schizosaccharomyces...	69	1e-10
emb|CAA15907| (AL021006) hypothetical protein Rv1251c [Mycobact...	65	2e-09

```
gi|1752701  (U60176) cag-omega [Helicobacter pylori]                  58  3e-07
gb|AAD35703.1|AE001736_1  (AE001736) conserved hypothetical prot...   57  5e-07
gi|1280167  (U55857) K08D10.5 gene product [Caenorhabditis elegans]   55  2e-06
emb|CAA86856|  (Z46828) similar to helicase domain; cDNA EST EMB...   55  3e-06
gb|AAD27693.1|AF125214_3  (AF125214) urease-enhancing factor hel...   49  1e-04
sp|P26314|RRPB_IBVB  RNA-DIRECTED RNA POLYMERASE (ORF1B) >gi|748...   48  3e-04
emb|CAA83018|  (Z30541) potential chimeric protein [avian infect...   48  3e-04
```

Sequences with E-value WORSE than threshold

```
gb|AAD07506.1|  (AE000559) H. pylori predicted coding region HP0...   45  0.004
sp|P75033|Y140_MYCPN  HYPOTHETICAL ATP-BINDING PROTEIN MG140 HOM...   42  0.019
gi|3329101  (AE001335) Exodeoxyribonuclease V, Alpha [Chlamydia ...   42  0.019
gb|AAD27694.1|AF125214_4  (AF125214) HelA [Helicobacter pylori]       42  0.024
pir||NCECXF  exodeoxyribonuclease V (EC 3.1.11.5) 67K chain - Es...   41  0.054
sp|P04993|EX5A_ECOLI  EXODEOXYRIBONUCLEASE V 67 KD POLYPEPTIDE (...   41  0.054
sp|P45158|EX5A_HAEIN  EXODEOXYRIBONUCLEASE V ALPHA CHAIN >gi|107...   41  0.054
sp|P47386|Y140_MYCGE  HYPOTHETICAL ATP-BINDING PROTEIN MG140 >gi...   39  0.21
sp|P16342|RRPB_CVMA5  RNA-DIRECTED RNA POLYMERASE (ORF1B) >gi|93...   38  0.28
gi|2621557  (AE000832) DNA helicase related protein [Methanobact...   38  0.28
gi|2641128  (AF029248) RNA-directed RNA polymerase [Murine hepat...   38  0.28
emb|CAA36202|  (X51939) open reading frame 1b (AA 1-2733) [murin...   38  0.28
pir||S52928  XSUG1 protein - African clawed frog                      38  0.47
sp|P39369|YJHR_ECOLI  HYPOTHETICAL 38.0 KD PROTEIN IN FECI-FIMB ...   37  0.62
gi|3790757  (AF099922) similar to the AAA family of ATPases; mos...   37  0.62
gb|AAD18890|  (AE001657) Exodeoxyribonuclease V, Alpha [Chlamydi...   37  0.62
sp|P29982|RRPB_CVMJH  RNA-DIRECTED RNA POLYMERASE (ORF1B) >gi|74...   37  0.81
sp|P46470|PRS8_XENLA  26S PROTEASE REGULATORY SUBUNIT 8 (SUG1 HO...   37  0.81
emb|CAB14689|  (Z99117) similar to conjugation transfer protein ...   36  1.1
dbj|BAA76763.1|  (AB023136) KIAA0919 protein [Homo sapiens]           36  1.1
emb|CAB07118|  (Z92772) recD [Mycobacterium tuberculosis]             36  1.4
sp|O50581|RECG_STAAU  ATP-DEPENDENT DNA HELICASE RECG >gi|282689...   36  1.4
gb|AAD41422.1|AC007727_11  (AC007727) Similar to gb|M87339 repli...   36  1.8
sp|P29569|YPV1_METTF  HYPOTHETICAL 40.7 KD PROTEIN (ORF1) >gi|28...   35  2.4
sp|P29570|YPZ1_METTF  HYPOTHETICAL 40.6 KD PROTEIN (ORF1') >gi|2...   35  2.4
sp|Q04913|NGFI_XENLA  NERVE GROWTH FACTOR INDUCED PROTEIN I-B HO...   35  2.4
sp|P47210|PRS8_HUMAN  26S PROTEASE REGULATORY SUBUNIT 8 (PROTEAS...   35  2.4
sp|P52915|PRS8_MOUSE  26S PROTEASE REGULATORY SUBUNIT 8 (MSUG1 P...   35  2.4
emb|CAA61864|  (X89719) put. 26S protease subunit [Sus scrofa]        35  2.4
dbj|BAA22935|  (AB000493) proteasome p45/SUG [Rattus norvegicus]      35  2.4
gi|2661071  (AF035309) similar to 26S proteasome subunit p45 [Ho...   35  2.4
pir||S47220  protein kinase C - yeast (Candida albicans)              35  3.1
gi|393034  (U01065) nonstructural polyprotein [Western equine en...   35  3.1
emb|CAA52868|  (X74892) NSP1, NSP2, NSP3 [Western equine encepha...   35  3.1
sp|P43057|KPC1_CANAL  PROTEIN KINASE C-LIKE 1 (PKC 1) >gi|832908...   35  3.1
pir||S57222  NSP1-NSP2-NSP3 polypeptide - western equine encepha...   35  3.1
sp|P37609|LCN2_LACLA  LACTICIN 481/LACTOCOCCIN BIOSYNTHESIS PROT...   35  3.1
gi|2688234  (AE001140) DNA helicase (uvrD) [Borrelia burgdorferi]     35  3.1
ref|NP_002796.1|PPSMC5|  proteasome (prosome, macropain) 26S sub...   34  4.1
pir||S51042  tat-binding protein homolog - Plasmodium falciparum      34  5.4
gi|3184291  (AC004136) putative DNA polymerase III gamma subunit...   34  5.4
emb|CAB50340.1|  (AJ248287) hypothetical protein [Pyrococcus aby...   34  5.4
gb|AAD44200.1|AF143772_2  (AF143772) IS1601-D [Mycobacterium avi...   34  5.4
sp|Q02137|ILVB_LACLA  ACETOLACTATE SYNTHASE LARGE SUBUNIT (AHAS)...   34  7.0
sp|Q56243|UVRB_THETH  EXCINUCLEASE ABC SUBUNIT B >gi|1325924|dbj...   34  7.0
emb|CAB11558.1|  (Z98866) similar to ATPases associated with var...   34  7.0
emb|CAA09623|  (AJ011482) hypothetical protein [Porcine transmis...   34  7.0
sp|P54814|PRS8_MANSE  26S PROTEASE REGULATORY SUBUNIT 8 (18-56 P...   33  9.2
gb|AAC55658.1|  (U51931) helicase [Beet yellow stunt virus]           33  9.2
sp|O18413|PRS8_DROME  26S PROTEASE REGULATORY SUBUNIT 8 >gi|2245...   33  9.2
gi|2815905  (AF043734) Pros45 proteosome subunit homolog [Drosop...   33  9.2
```

Alignments

```
sp|P30771|NAM7_YEAST NAM7 PROTEIN (NONSENSE-MEDIATED MRNA DECAY PROTEIN 1)
          (UP-FRAMESHIFT SUPPRESSOR 1) >gi|83347|pir||S23408
          prematurely terminated mRNA decay factor NAM7 - yeast
          (Saccharomyces cerevisiae) >gi|4023|emb|CAA44266|
          (X62394) helicase [Saccharomyces cerevisiae] >gi|173142
          (M76659) zinc finger protein [Saccharomyces cerevisiae]
          >gi|807962|emb|CAA89226| (Z49259) Nam7p [Saccharomyces
          cerevisiae]
          Length = 971

 Score = 1931 bits (4947), Expect = 0.0
 Identities = 951/971 (97%), Positives = 951/971 (97%)

Query: 1   MVGSGSHTPYDISNSPSDVNVQPATQLNSTLVEDDDVDNQLFEEAQVTETGFRSPSASDN 60
           MVGSGSHTPYDISNSPSDVNVQPATQLNSTLVEDDDVDNQLFEEAQVTETGFRSPSASDN
Sbjct: 1   MVGSGSHTPYDISNSPSDVNVQPATQLNSTLVEDDDVDNQLFEEAQVTETGFRSPSASDN 60

Query: 61  SCAYCGIDSAKCVIKCNSCKKWFCNTKNGTXXXXXXXXXXXXXXXXXXXXPDSDLGDTVL 120
           SCAYCGIDSAKCVIKCNSCKKWFCNTKNGT                    PDSDLGDTVL
Sbjct: 61  SCAYCGIDSAKCVIKCNSCKKWFCNTKNGTSSSHIVNHLVLSHHNVVSLHPDSDLGDTVL 120
```

```
Query: 121 ECYNCGRKNVFLLGFVSAKSEAVVVLLCRIPCAQTKNANWDTDQWQPLIEDRQLLSWVAE 180
           ECYNCGRKNVFLLGFVSAKSEAVVVLLCRIPCAQTKNANWDTDQWQPLIEDRQLLSWVAE
Sbjct: 121 ECYNCGRKNVFLLGFVSAKSEAVVVLLCRIPCAQTKNANWDTDQWQPLIEDRQLLSWVAE 180

Query: 181 QPTEEEKLKARLITPSQISKLEAKWRSNKDATINDIDAPEEQEAIPPLLLRYQDAYEYQR 240
           QPTEEEKLKARLITPSQISKLEAKWRSNKDATINDIDAPEEQEAIPPLLLRYQDAYEYQR
Sbjct: 181 QPTEEEKLKARLITPSQISKLEAKWRSNKDATINDIDAPEEQEAIPPLLLRYQDAYEYQR 240

Query: 241 SYGPLIKLEADYDKQLKESQALEHISVSWSLALNNRHLASFTLSTFESNELKVAIGDEMI 300
           SYGPLIKLEADYDKQLKESQALEHISVSWSLALNNRHLASFTLSTFESNELKVAIGDEMI
Sbjct: 241 SYGPLIKLEADYDKQLKESQALEHISVSWSLALNNRHLASFTLSTFESNELKVAIGDEMI 300

Query: 301 LWYSGMQHPDWEGRGYIVRLPNSFQDTFTLELKPSKTPPPTHLTTGFTAEFIWKGTSYDR 360
           LWYSGMQHPDWEGRGYIVRLPNSFQDTFTLELKPSKTPPPTHLTTGFTAEFIWKGTSYDR
Sbjct: 301 LWYSGMQHPDWEGRGYIVRLPNSFQDTFTLELKPSKTPPPTHLTTGFTAEFIWKGTSYDR 360

Query: 361 MQDALKKFAIDKKSISGYLYYKILGHQVVDISFDVPLPKEFSIPNFAQLNSSQSNAVSHV 420
           MQDALKKFAIDKKSISGYLYYKILGHQVVDISFDVPLPKEFSIPNFAQLNSSQSNAVSHV
Sbjct: 361 MQDALKKFAIDKKSISGYLYYKILGHQVVDISFDVPLPKEFSIPNFAQLNSSQSNAVSHV 420

Query: 421 LQRPLSLIQGPPGTGKTVTSATIVYHLSKIHKDRILVCAPSNVAVDHLAAKLRDLGLKVV 480
           LQRPLSLIQGPPGTGKTVTSATIVYHLSKIHKDRILVCAPSNVAVDHLAAKLRDLGLKVV
Sbjct: 421 LQRPLSLIQGPPGTGKTVTSATIVYHLSKIHKDRILVCAPSNVAVDHLAAKLRDLGLKVV 480

Query: 481 RLTAKSREDVESSVSNLALHNLVGRGAKGELKNLLKLKDEVGELSASDTKRFVKLVRKTE 540
           RLTAKSREDVESSVSNLALHNLVGRGAKGELKNLLKLKDEVGELSASDTKRFVKLVRKTE
Sbjct: 481 RLTAKSREDVESSVSNLALHNLVGRGAKGELKNLLKLKDEVGELSASDTKRFVKLVRKTE 540

Query: 541 AEILNKADVVCCTCVGAGDKRLDTKFRTVLIDESTQASEPECLIPIVKGAKQVILVGDHQ 600
           AEILNKADVVCCTCVGAGDKRLDTKFRTVLIDESTQASEPECLIPIVKGAKQVILVGDHQ
Sbjct: 541 AEILNKADVVCCTCVGAGDKRLDTKFRTVLIDESTQASEPECLIPIVKGAKQVILVGDHQ 600

Query: 601 QLGPVILERKAADAGLKQSLFERLISLGHVPIRLEVQYRMNPYLSEFPSNMFYEGSLQNG 660
           QLGPVILERKAADAGLKQSLFERLISLGHVPIRLEVQYRMNPYLSEFPSNMFYEGSLQNG
Sbjct: 601 QLGPVILERKAADAGLKQSLFERLISLGHVPIRLEVQYRMNPYLSEFPSNMFYEGSLQNG 660

Query: 661 VTIEQRTVPNSKFPWPIRGIPMMFWANYGREEISANGTSFLNRIEAMNCERIITKLFRDG 720
           VTIEQRTVPNSKFPWPIRGIPMMFWANYGREEISANGTSFLNRIEAMNCERIITKLFRDG
Sbjct: 661 VTIEQRTVPNSKFPWPIRGIPMMFWANYGREEISANGTSFLNRIEAMNCERIITKLFRDG 720

Query: 721 VKPEQIGVITPYEGQRAYILQYMQMNGSLDKDLYIKVEVASVDAFQGREKDYIILSCVRA 780
           VKPEQIGVITPYEGQRAYILQYMQMNGSLDKDLYIKVEVASVDAFQGREKDYIILSCVRA
Sbjct: 721 VKPEQIGVITPYEGQRAYILQYMQMNGSLDKDLYIKVEVASVDAFQGREKDYIILSCVRA 780

Query: 781 NEQQAIGFLRDPRRLNVGLTRAKYGLVILGNPRSLARNTLWNHLLIHFREKGCLVEGTLD 840
           NEQQAIGFLRDPRRLNVGLTRAKYGLVILGNPRSLARNTLWNHLLIHFREKGCLVEGTLD
Sbjct: 781 NEQQAIGFLRDPRRLNVGLTRAKYGLVILGNPRSLARNTLWNHLLIHFREKGCLVEGTLD 840

Query: 841 NLQLCTVQLVRPQPRKTERPMNAQFNVESEMGDFPKFQDFDAQSMVSFSGQIGDFGNAFV 900
           NLQLCTVQLVRPQPRKTERPMNAQFNVESEMGDFPKFQDFDAQSMVSFSGQIGDFGNAFV
Sbjct: 841 NLQLCTVQLVRPQPRKTERPMNAQFNVESEMGDFPKFQDFDAQSMVSFSGQIGDFGNAFV 900

Query: 901 DNTELSSYINNEYWNFENFKSAFSQKQNRNEIDDRNLYQEEASHLNSNFARELQREEQKH 960
           DNTELSSYINNEYWNFENFKSAFSQKQNRNEIDDRNLYQEEASHLNSNFARELQREEQKH
Sbjct: 901 DNTELSSYINNEYWNFENFKSAFSQKQNRNEIDDRNLYQEEASHLNSNFARELQREEQKH 960

Query: 961 ELSKDFSNLGI 971
           ELSKDFSNLGI
Sbjct: 961 ELSKDFSNLGI 971
```

```
gi|2739355 (AC003972) pNORF1 [Homo sapiens] >gi|3328175 (AF074016)
           nonsense-mediated mRNA decay trans-acting factor [Homo
           sapiens]
           Length = 1118

 Score =  920 bits (2352), Expect = 0.0
 Identities = 459/848 (54%), Positives = 600/848 (70%), Gaps = 20/848 (2%)

Query: 12  ISNSPSDVNVQPATQLNSTL-VEDDDVDNQLFEEAQVTETGFRSPSASDNSCAYCGIDSA 70
           + N   D +V   +QL + L  E+D+ D    ++  +            ++C+YCGI
Sbjct: 84  LQNGAVDDSVAKTSQLLAELNFEEDEEDTYYTKDLPI------------HACSYCGIHDP 131

Query: 71  KCVIKCNSCKKWFCNTKNGTXXXXXXXXXXXXXXXXXXXPDSDLGDTVLECYNCGRKNV 130
            CV+ CN+ KKWFCN +  T                     D  LG+TVLECYNCG +NV
Sbjct: 132 ACVVYCNTSKKWFCNGRGNTSGSHIVNHLVRAKCKEVTLHKDGPLGETVLECYNCGCRNV 191

Query: 131 FLLGFVSAKSEAVVVLLCRIPCAQT---KNANWDTDQWQPLIEDRQLLSWVAEQPTEEEK 187
           FLLGF+ AK+++VVVLLCR PCA     K+ NWD+ QWQPLI+DR  LSW+ + P+E+E+
Sbjct: 192 FLLGFIPAKADSVVVLLCRQPCASQSSLKDINWDSSQWQPLIQDRCFLSWLVKIPSEQEQ 251

Query: 188 LKARLITPSQISKLEAKWRSNKDATINDIDAPEEQEAIPPLLLRYQDAYEYQRSYGPLIK 247
           L+AR IT  QI+KLE  W+ N  AT+ D++ P   E    +LLRY+DAY+YQ  +GPL+K
Sbjct: 252 LRARQITAQQINKLEELWKENPSATLEDLEKPGVDEEPQHVLLRYEDAYQYQNIFGPLVK 311

Query: 248 LEADYDKQLKESQALEHISVSWSLALNNRHLASFTLSTFESNELKVAIGDEMILWYSGMQ 307
           LEADYDK+LKESQ  ++I+V W L LN + +A FTL   +S ++++  GDE+ L Y G
Sbjct: 312 LEADYDKKLKESQTQDNITVRWDLGLNKKRIAYFTLPKTDS-DMRLMQGDEICLRYKGDL 370
```

```
Query: 308 HPDWEGRGYIVRLPNSFQDTFTLELKPSKTPPPTHLTTGFTAEFIWKGTSYDRMQDALKK 367
            P W+G G+++++P+++ D   +EL+ S    P  +T  F  +F+WK TS+DRMQ ALK
Sbjct: 371 APLWKGIGHVIKVPDNYGDEIAIELR-SSVGAPVEVTHNFQVDFVWKSTSFDRMQSALKT 429

Query: 368 FAIDKKSISGYLYYKILGHQVVDISFDVPLPKEFSIPNFAQLNSSQSNAVSHVLQRPLSL 427
           FA+D+ S+SGY+Y+K+LGH+V D+     LPK F+      LN SQ  AV  VLQRPLSL
Sbjct: 430 FAVDETSVSGYIYHKLLGHEVEDVIIKCQLPKRFTAQGLPDLNHSQVYAVKTVLQRPLSL 489

Query: 428 IQGPPGTGKTVTSATIVYHLSKIHKDRILVCAPSNVAVDHLAAKLRDLGLKVVRLTAKSR 487
           IQGPPGTGKTVTSATIVYHL++     +LVCAPSN+AVD L  K+   GLKVVRL AKSR
Sbjct: 490 IQGPPGTGKTVTSATIVYHLARQGNGPVLVCAPSNIAVDQLTEKIHQTGLKVVRLCAKSR 549

Query: 488 EDVESSVSNLALHNLV-GRGAKGELKNLLKLKDEVGELSASDTKRFVKLVRKTEAEILNK 546
           E ++S VS LALHN +    +  EL+ L +LKDE GELS++D KR+  L R  E E+L
Sbjct: 550 EAIDSPVSFLALHNQIRNMDSMPELQKLQQLKDETGELSSADEKRYRALKRTAERELLMN 609

Query: 547 ADVVCCTCVGAGDKRL-DTKFRTVLIDESTQASEPECLIPIVKGAKQVILVGDHQQLGPV 605
           ADV+CCTCVGAGD RL   +FR++LIDESTQA+EPEC++P+V GAKQ+ILVGDH QLGPV
Sbjct: 610 ADVICCTCVGAGDPRLAKMQFRSILIDESTQATEPECMVPVVLGAKQLILVGDHCQLGPV 669

Query: 606 ILERKAADAGLKQSLFERLISLGHVPIRLEVQYRMNPYLSEFPSNMFYEGSLQNGVTIEQ 665
           ++ +KAA AGL QSLFERL+ LG  PIRL+VQYRM+P LS FPSN+FYEGSLQNGVT
Sbjct: 670 VMCKKAAKAGLSQSLFERLVVLGIRPIRLQVQYRMHPALSAFPSNIFYEGSLQNGVTAAD 729

Query: 666 RTVPNSKFPWPIRGIPMMFWANYGREEISANGTSFLNRIEAMNCERIITKLFRDGVKPEQ 725
           R      F WP    PM F+   G+EEI+++GTS+LNR EA N E+I TKL + G KP+Q
Sbjct: 730 RVKKGFDFQWPQPDKPMFFYVTQGQEEIASSGTSYLNRTEAANVEKITTKLLKAGAKPDQ 789

Query: 726 IGVITPYEGQRAYILQYMQMNGSLDKDLYIKVEVASVDAFQGREKDYIILSCVRANEQQA 785
           IG+ITPYEGQR+Y++QYMQ +GSL   LY +VE+ASVDAFQGREKD+IILSCVRANE Q
Sbjct: 790 IGIITPYEGQRSYLVQYMQFSGSLHTKLYQEVEIASVDAFQGREKDFIILSCVRANEHQG 849

Query: 786 IGFLRDPRRLNVGLTRAKYGLVILGNPRSLARNTLWNHLLIHFREKGCLVEGTLDNLQLC 845
           IGFL DPRRLNV LTRA+YG++I+GNP++L++  LWNHLL +++E+  LVEG L+NL+
Sbjct: 850 IGFLNDPRRLNVALTRARYGVIIVGNPKALSKQPLWNHLLNYYKEQKVLVEGPLNNLRES 909

Query: 846 TVQLVRPQ 853
            +Q  +P+
Sbjct: 910 LMQFSKPR 917
```

gi|1885356 (U59323) type 1 RNA helicase pNORF1 [Homo sapiens]
Length = 1118

Score = 919 bits (2350), Expect = 0.0
Identities = 459/848 (54%), Positives = 600/848 (70%), Gaps = 20/848 (2%)

```
Query: 12  ISNSPSDVNVQPATQLNSTL-VEDDDVDNQLFEEAQVTETGFRSPSASDNSCAYCGIDSA 70
           + N   D +V   +QL + L  E+D+ D    ++  +            ++C+YCGI
Sbjct: 84  LQNGAVDDSVAKTSQLLAELNFEEDEEDTYYTKDLPI------------HACSYCGIHDP 131

Query: 71  KCVIKCNSCKKWFCNTKNGTXXXXXXXXXXXXXXXXXXXPDSDLGDTVLECYNCGRKNV 130
            CV+ CN+ KKWFCN +  T                      D  LG+TVLECYNCG +NV
Sbjct: 132 ACVVYCNTSKKWFCNGRGNTSGSHIVNHLVRAKCKEVTLHKDGPLGETVLECYNCGCRNV 191

Query: 131 FLLGFVSAKSEAVVVLLCRIPCAQT---KNANWDTDQWQPLIEDRQLLSWVAEQPTEEEK 187
           FLLGF+ AK+++VVVLLCR PCA     K+ NWD+ QWQPLI+DR  LSW+ + P+E+E+
Sbjct: 192 FLLGFIPAKADSVVVLLCRQPCASQSSLKDINWDSSQWQPLIQDRCFLSWLVKIPSEQEQ 251

Query: 188 LKARLITPSQISKLEAKWRSNKDATINDIDAPEEQEAIPPLLLRYQDAYEYQRSYGPLIK 247
           L+AR IT  QI+KLE  W+ N  AT+ D++ P   E    +LLRY+DAY+YQ  +GPL+K
Sbjct: 252 LRARQITAQQINKLEELWKENPSATLEDLEKPGVDEEPQHVLLRYEDAYQYQNIFGPLVK 311

Query: 248 LEADYDKQLKESQALEHISVSWSLALNNRHLASFTLSTFESNELKVAIGDEMILWYSGMQ 307
           LEADYDK+LKESQ  ++I+V W L LN + +A FTL   +S ++++  GDE+ L Y G
Sbjct: 312 LEADYDKKLKESQTQDNITVRWDLGLNKKRIAYFTLPKTDS-DMRLMQGDEICLRYKGDL 370

Query: 308 HPDWEGRGYIVRLPNSFQDTFTLELKPSKTPPPTHLTTGFTAEFIWKGTSYDRMQDALKK 367
            P W+G G+++++P+++ D   +EL+ S    P  +T  F  +F+WK TS+DRMQ ALK
Sbjct: 371 APLWKGIGHVIKVPDNYGDEIAIELR-SSVGAPVEVTHNFQVDFVWKSTSFDRMQSALKT 429

Query: 368 FAIDKKSISGYLYYKILGHQVVDISFDVPLPKEFSIPNFAQLNSSQSNAVSHVLQRPLSL 427
           FA+D+ S+SGY+Y+K+LGH+V D+     LPK F+      LN SQ  AV  VLQRPLSL
Sbjct: 430 FAVDETSVSGYIYHKLLGHEVEDVITKCQLPKRFTAQGLPDLNHSQVYAVKTVLQRPLSL 489

Query: 428 IQGPPGTGKTVTSATIVYHLSKIHKDRILVCAPSNVAVDHLAAKLRDLGLKVVRLTAKSR 487
           IQGPPGTGKTVTSATIVYHL++     +LVCAPSN+AVD L  K+   GLKVVRL AKSR
Sbjct: 490 IQGPPGTGKTVTSATIVYHLARQGNGPVLVCAPSNIAVDQLTEKIHQTGLKVVRLCAKSR 549

Query: 488 EDVESSVSNLALHNLV-GRGAKGELKNLLKLKDEVGELSASDTKRFVKLVRKTEAEILNK 546
           E ++S VS LALHN +    +  EL+ L +LKDE GELS++D KR+  L R  E E+L
Sbjct: 550 EAIDSPVSFLALHNQIRNMDSMPELQKLQQLKDETGELSSADEKRYRALKRTAERELLMN 609

Query: 547 ADVVCCTCVGAGDKRL-DTKFRTVLIDESTQASEPECLIPIVKGAKQVILVGDHQQLGPV 605
           ADV+CCTCVGAGD RL   +FR++LIDESTQA+EPEC++P+V GAKQ+ILVGDH QLGPV
Sbjct: 610 ADVICCTCVGAGDPRLAKMQFRSILIDESTQATEPECMVPVVLGAKQLILVGDHCQLGPV 669

Query: 606 ILERKAADAGLKQSLFERLISLGHVPIRLEVQYRMNPYLSEFPSNMFYEGSLQNGVTIEQ 665
           ++ +KAA AGL QSLFERL+ LG  PIRL+VQYRM+P LS FPSN+FYEGSLQNGVT
```

```
Sbjct: 670 VMCKKAAKAGLSQSLFERLVVLGIRPIRLQVQYRMHPALSAFPSNIFYEGSLQNGVTAAD 729

Query: 666 RTVPNSKFPWPIRGIPMMFWANYGREEISANGTSFLNRIEAMNCERIITKLFRDGVKPEQ 725
           R      F WP    PM F+   G+EEI+++GTS+LNR EA N E+I TKL + G KP+Q
Sbjct: 730 RVKKGFDFQWPQPDKPMFFYVTQGQEEIASSGTSYLNRTEAANVEKITTKLLKAGAKPDQ 789

Query: 726 IGVITPYEGQRAYILQYMQMNGSLDKDLYIKVEVASVDAFQGREKDYIILSCVRANEQQA 785
           IG+ITPYEGQR+Y++QYMQ +GSL   LY +VE+ASVDAFQGREKD+IILSCVRANE Q
Sbjct: 790 IGIITPYEGQRSYLVQYMQFSGSLHTKLYQEVEIASVDAFQGREKDFIILSCVRANEHQG 849

Query: 786 IGFLRDPRRLNVGLTRAKYGLVILGNPRSLARNTLWNHLLIHFREKGCLVEGTLDNLQLC 845
           IGFL DPRRLNV LTRA+YG++I+GNP++L++  LWNHLL +++E+  LVEG L+NL+
Sbjct: 850 IGFLNDPRRLNVALTRARYGVIIVGNPKALSKQPLWNHLLNYYKEQKVLVEGPLNNLRES 909

Query: 846 TVQLVRPQ 853
            +Q  +P+
Sbjct: 910 LMQFSKPR 917


 ref|NP_002902.1|PRENT1| regulator of nonsense transcripts 1 >gi|1575536 (U65533) regulator
           of nonsense transcript stability [Homo sapiens]
           Length = 1118

 Score =  919 bits (2350), Expect = 0.0
 Identities = 459/848 (54%), Positives = 599/848 (70%), Gaps = 20/848 (2%)

Query: 12  ISNSPSDVNVQPATQLNSTL-VEDDDVDNQLFEEAQVTETGFRSPSASDNSCAYCGIDSA 70
           + N   D +V   +QL + L  E+D+ D    ++  +            ++C+YCGI
Sbjct: 84  LQNGAVDDSVAKTSQLLAELNFEEDEEDTYYTKDLPI------------HACSYCGIHDP 131

Query: 71  KCVIKCNSCKKWFCNTKNGTXXXXXXXXXXXXXXXXXXXPDSDLGDTVLECYNCGRKNV 130
            CV+ CN+ KKWFCN +  T                     D  LG+TVLECYNCG +NV
Sbjct: 132 ACVVYCNTSKKWFCNGRGNTSGSHIVNHLVRAKCKEVTLHKDGPLGETVLECYNCGCRNV 191

Query: 131 FLLGFVSAKSEAVVVLLCRIPCAQT---KNANWDTDQWQPLIEDRQLLSWVAEQPTEEEK 187
           FLLGF+ AK+++VVVLLCR PCA     K+ NWD+ QWQPLI+DR  LSW+ + P+E+E+
Sbjct: 192 FLLGFIPAKADSVVVLLCRQPCASQSSLKDINWDSSQWQPLIQDRCFLSWLVKIPSEQEQ 251

Query: 188 LKARLITPSQISKLEAKWRSNKDATINDIDAPEEQEAIPPLLLRYQDAYEYQRSYGPLIK 247
           L+AR IT  QI+KLE  W+ N  AT+ D++ P   E    +LLRY+DAY+YQ  +GPL+K
Sbjct: 252 LRARQITAQQINKLEELWKENPSATLEDLEKPGVDEEPQHVLLRYEDAYQYQNIFGPLVK 311

Query: 248 LEADYDKQLKESQALEHISVSWSLALNNRHLASFTLSTFESNELKVAIGDEMILWYSGMQ 307
           LEADYDK+LKESQ  ++I+V W L LN + +A FTL   +S ++++  GDE+ L Y G
Sbjct: 312 LEADYDKKLKESQTQDNITVRWDLGLNKKRIAYFTLPKTDS-DMRLMQGDEICLRYKGDL 370

Query: 308 HPDWEGRGYIVRLPNSFQDTFTLELKPSKTPPPTHLTTGFTAEFIWKGTSYDRMQDALKK 367
            P W+G G+++++P+++ D   +EL+ S    P  +T  F  +F+WK TS+DRMQ ALK
Sbjct: 371 APLWKGIGHVIKVPDNYGDEIAIELR-SSVGAPVEVTHNFQVDFVWKSTSFDRMQSALKT 429

Query: 368 FAIDKKSISGYLYYKILGHQVVDISFDVPLPKEFSIPNFAQLNSSQSNAVSHVLQRPLSL 427
           FA+D+ S+SGY+Y+K+LGH+V D+     LPK F+      LN SQ  AV  VLQRPLSL
Sbjct: 430 FAVDETSVSGYIYHKLLGHEVEDVIIKCQLPKRFTAQALPDLNHSQVYAVKTVLQRPLSL 489

Query: 428 IQGPPGTGKTVTSATIVYHLSKIHKDRILVCAPSNVAVDHLAAKLRDLGLKVVRLTAKSR 487
           IQGPPGTGKTVTSATIVYHL++     +LVCAPSN+AVD L  K+   GLKVVRL  KSR
Sbjct: 490 IQGPPGTGKTVTSATIVYHLARQDNGPVLVCAPSNIAVDQLTEKIHQTGLKVVRLCPKSR 549

Query: 488 EDVESSVSNLALHNLV-GRGAKGELKNLLKLKDEVGELSASDTKRFVKLVRKTEAEILNK 546
           E ++S VS LALHN +    +  EL+ L +LKDE GELS++D KR+  L R  E E+L
Sbjct: 550 EAIDSPVSFLALHNQIRNMDSMPELQKLQQLKDETGELSSADEKRYRALKRTAERELLMN 609

Query: 547 ADVVCCTCVGAGDKRL-DTKFRTVLIDESTQASEPECLIPIVKGAKQVILVGDHQQLGPV 605
           ADV+CCTCVGAGD RL   +FR++LIDESTQA+EPEC++P+V GAKQ+ILVGDH QLGPV
Sbjct: 610 ADVICCTCVGAGDPRLAKMQFRSILIDESTQATEPECMVPVVLGAKQLILVGDHCQLGPV 669

Query: 606 ILERKAADAGLKQSLFERLISLGHVPIRLEVQYRMNPYLSEFPSNMFYEGSLQNGVTIEQ 665
           ++ +KAA AGL QSLFERL+ LG  PIRL+VQYRM+P LS FPSN+FYEGSLQNGVT
Sbjct: 670 VMCKKAAKAGLSQSLFERLVVLGIRPIRLQVQYRMHPALSAFPSNIFYEGSLQNGVTAAD 729

Query: 666 RTVPNSKFPWPIRGIPMMFWANYGREEISANGTSFLNRIEAMNCERIITKLFRDGVKPEQ 725
           R      F WP    PM F+   G+EEI+++GTS+LNR EA N E+I TKL + G KP+Q
Sbjct: 730 RVKKGFDFQWPQPDKPMFFYVTQGQEEIASSGTSYLNRTEAANVEKITTKLLKAGAKPDQ 789

Query: 726 IGVITPYEGQRAYILQYMQMNGSLDKDLYIKVEVASVDAFQGREKDYIILSCVRANEQQA 785
           IG+ITPYEGQR+Y++QYMQ +GSL   LY +VE+ASVDAFQGREKD+IILSCVRANE Q
Sbjct: 790 IGIITPYEGQRSYLVQYMQFSGSLHTKLYQEVEIASVDAFQGREKDFIILSCVRANEHQG 849

Query: 786 IGFLRDPRRLNVGLTRAKYGLVILGNPRSLARNTLWNHLLIHFREKGCLVEGTLDNLQLC 845
           IGFL DPRRLNV LTRA+YG++I+GNP++L++  LWNHLLI ++E+  LVEG L+NL+
Sbjct: 850 IGFLNDPRRLNVALTRARYGVIIVGNPKALSKQPLWNHLLIFYKEQKVLVEGPLNNLRES 909

Query: 846 TVQLVRPQ 853
            +Q  +P+
Sbjct: 910 LMQFSKPR 917


 dbj|BAA19664| (D86988) KIAA0221 [Homo sapiens]
           Length = 1129
```

```
 Score =  917 bits (2343), Expect = 0.0
 Identities = 459/858 (53%), Positives = 600/858 (69%), Gaps = 29/858 (3%)

Query: 12  ISNSPSDVNVQPATQLNSTL-VEDDDVDNQLFEEAQVTETGFRSPSASDNSCAYCGIDSA 70
           + N   D +V   +QL + L  E+D+ D    ++  +            ++C+YCGI
Sbjct: 84  LQNGAVDDSVAKTSQLLAELNFEEDEEDTYYTKDLPI------------HACSYCGIHDP 131

Query: 71  KCVIKCNSCKKWFCNTKNGTXXXXXXXXXXXXXXXXXXXPDSDLGDTVLECYNCGRKNV 130
            CV+ CN+ KKWFCN +  T                      D  LG+TVLECYNCG +NV
Sbjct: 132 ACVVYCNTSKKWFCNGRGNTSGSHIVNHLVRAKCKEVTLHKDGPLGETVLECYNCGCRNV 191

Query: 131 FLLGFVSAKSEAVVVLLCRIPCAQT---KNANWDTDQWQPLIEDRQLLSWVAEQPTEEEK 187
           FLLGF+ AK+++VVVLLCR PCA     K+ NWD+ QWQPLI+DR  LSW+ + P+E+E+
Sbjct: 192 FLLGFIPAKADSVVVLLCRQPCASQSSLKDINWDSSQWQPLIQDRCFLSWLVKIPSEQEQ 251

Query: 188 LKARLITPSQISKLEAKWRSNKDATINDIDAPEEQEAIPPLLLRYQDAYEYQRSYGPLIK 247
           L+AR IT  QI+KLE  W+ N  AT+ D++ P   E    +LLRY+DAY+YQ  +GPL+K
Sbjct: 252 LRARQITAQQINKLEELWKENPSATLEDLEKPGVDEEPQHVLLRYEDAYQYQNIFGPLVK 311

Query: 248 LEADYDKQLKESQALEHISVSWSLALNNRHLASFTLSTFESN----------ELKVAIGD 297
           LEADYDK+LKESQ  ++I+V W L LN + +A FTL   +S           ++++  GD
Sbjct: 312 LEADYDKKLKESQTQDNITVRWDLGLNKKRIAYFTLPKTDSGNEDLVIIWLRDMRLMQGD 371

Query: 298 EMILWYSGMQHPDWEGRGYIVRLPNSFQDTFTLELKPSKTPPPTHLTTGFTAEFIWKGTS 357
           E+ L Y G   P W+G G+++++P+++ D   +EL+ S    P  +T  F  +F+WK TS
Sbjct: 372 EICLRYKGDLAPLWKGIGHVIKVPDNYGDEIAIELR-SSVGAPVEVTHNFQVDFVWKSTS 430

Query: 358 YDRMQDALKKFAIDKKSISGYLYYKILGHQVVDISFDVPLPKEFSIPNFAQLNSSQSNAV 417
           +DRMQ ALK FA+D+ S+SGY+Y+K+LGH+V D+     LPK F+      LN SQ  AV
Sbjct: 431 FDRMQSALKTFAVDETSVSGYIYHKLLGHEVEDVIIKCQLPKRFTAQGLPDLNHSQVYAV 490

Query: 418 SHVLQRPLSLIQGPPGTGKTVTSATIVYHLSKIHKDRILVCAPSNVAVDHLAAKLRDLGL 477
             VLQRPLSLIQGPPGTGKTVTSATIVYHL++     +LVCAPSN+AVD L  K+   GL
Sbjct: 491 KTVLQRPLSLIQGPPGTGKTVTSATIVYHLARQGNGPVLVCAPSNIAVDQLTEKIHQTGL 550

Query: 478 KVVRLTAKSREDVESSVSNLALHNLV-GRGAKGELKNLLKLKDEVGELSASDTKRFVKLV 536
           KVVRL AKSRE ++S VS LALHN +    +  EL+ L +LKDE GELS++D KR+  L
Sbjct: 551 KVVRLCAKSREAIDSPVSFLALHNQIRNMDSMPELQKLQQLKDETGELSSADEKRYRALK 610

Query: 537 RKTEAEILNKADVVCCTCVGAGDKRL-DTKFRTVLIDESTQASEPECLIPIVKGAKQVIL 595
           R  E E+L  ADV+CCTCVGAGD RL   +FR++LIDESTQA+EPEC++P+V GAKQ+IL
Sbjct: 611 RTAERELLMNADVICCTCVGAGDPRLAKMQFRSILIDESTQATEPECMVPVVLGAKQLIL 670

Query: 596 VGDHQQLGPVILERKAADAGLKQSLFERLISLGHVPIRLEVQYRMNPYLSEFPSNMFYEG 655
           VGDH QLGPV++ +KAA AGL QSLFERL+ LG  PIRL+VQYRM+P LS FPSN+FYEG
Sbjct: 671 VGDHCQLGPVVMCKKAAKAGLSQSLFERLVVLGIRPIRLQVQYRMHPALSAFPSNIFYEG 730

Query: 656 SLQNGVTIEQRTVPNSKFPWPIRGIPMMFWANYGREEISANGTSFLNRIEAMNCERIITK 715
           SLQNGVT   R      F WP    PM F+   G+EEI+++GTS+LNR EA N E+I TK
Sbjct: 731 SLQNGVTAADRVKKGFDFQWPQPDKPMFFYVTQGQEEIASSGTSYLNRTEAANVEKITTK 790

Query: 716 LFRDGVKPEQIGVITPYEGQRAYILQYMQMNGSLDKDLYIKVEVASVDAFQGREKDYIIL 775
           L + G KP+QIG+ITPYEGQR+Y++QYMQ +GSL   LY +VE+ASVDAFQGREKD+IIL
Sbjct: 791 LLKAGAKPDQIGIITPYEGQRSYLVQYMQFSGSLHTKLYQEVEIASVDAFQGREKDFIIL 850

Query: 776 SCVRANEQQAIGFLRDPRRLNVGLTRAKYGLVILGNPRSLARNTLWNHLLIHFREKGCLV 835
           SCVRANE Q IGFL DPRRLNV LTRA+YG++I+GNP++L++  LWNHLL +++E+  LV
Sbjct: 851 SCVRANEHQGIGFLNDPRRLNVALTRARYGVIIVGNPKALSKQPLWNHLLNYYKEQKVLV 910

Query: 836 EGTLDNLQLCTVQLVRPQ 853
           EG L+NL+   +Q  +P+
Sbjct: 911 EGPLNNLRESLMQFSKPR 928


 sp|Q09820|YAC6_SCHPO HYPOTHETICAL 105.6 KD PROTEIN C16C9.06C IN CHROMOSOME I
           >gi|2130283|pir||S62476 hypothetical protein
           SPAC16C9.06c - fission yeast (Schizosaccharomyces pombe)
           Length = 935

 Score =  903 bits (2309), Expect = 0.0
 Identities = 462/933 (49%), Positives = 614/933 (65%), Gaps = 31/933 (3%)

Query: 20  NVQPATQLNSTLVEDDDVDNQLFEEAQVTETGFR--------SPSASDNSCAYCGIDSAK 71
           N+     ++  L  ++D+ + +  +   +E G              S+  CAYC I +
Sbjct: 4   NISSCLSMSLGLQPNNDISSLVSSKNMTSENGLEHQFEELLVEKQYSEEHCAYCHIKNPN 63

Query: 72  CVIKCNSCKKWFCNTKNGTXXXXXXXXXXXXXXXXXXXPDSDLGDTVLECYNCGRKNVF 131
            ++KC  C KWFCN +  +                      S L DTVLECYNCG +NVF
Sbjct: 64  SILKCLHCNKWFCNVRGKSGASHIISHLVRARHKQVALHSHSSLSDTVLECYNCGTRNVF 123

Query: 132 LLGFVSAKSEAVVVLLCRIPCAQT---KNANWDTDQWQPLIEDRQLLSWVAEQPTEEEKL 188
           LLGF+ AK++ VVVLLCR PCA+    K+ NWD  QWQP+I DRQ L W+   P+EEE+
Sbjct: 124 LLGFIPAKAKTVVVLLCRQPCARASIAKDMNWDLTQWQPIISDRQFLPWLITPPSEEEQK 183

Query: 189 KARLITPSQISKLEAKWRSNKDATINDIDAPEEQEAIPPLLLRYQDAYEYQRSYGPLIKL 248
            A  IT  Q+ +LE  WR + +A + D+D P E +++P + LRY+DA+ YQ    PLI+
Sbjct: 184 LAIPITSQQMVRLEELWRKDPNANLEDLDKPIEDDSLPSVELRYKDAHAYQAVLSPLIQA 243
```

```
Query: 249 EADYDKQLKESQALEHISVSWSLALNNRHLASFTLSTFESNELKVAIGDEMILWYSGMQH 308
           EADYDK+LKESQ  + + V W  A+N R+ A F L   ES E+++AIGDEM L Y G
Sbjct: 244 EADYDKRLKESQTQKDVVVRWDQAINKRYTAWFLLPKLESGEIRLAIGDEMKLTYEGELR 303

Query: 309 PDWEGRGYIVRLPNSFQDTFTLELKPSKTPPPTHLTTGFTAEFIWKGTSYDRMQDALKKF 368
             W   GY++++PN+  D   LELK S   P    T  F+ +++WK TS+DRMQ AL+ F
Sbjct: 304 APWSSTGYVIKIPNNVSDEVGLELKRSDKVP-IECTHNFSVDYVWKSTSFDRMQTALRLF 362

Query: 369 AIDKKSISGYLYYKILGHQVVDISFDVPLPKEFSIPNFAQLNSSQSNAVSHVLQRPLSLI 428
           A D   +S +LY+K+LGH +        LP + S+PN  +LN+SQS AV  VL +PLSLI
Sbjct: 363 ATDGSRLSSFLYHKLLGHDIPPSFLKPKLPSDLSVPNLPKLNASQSEAVRAVLSKPLSLI 422

Query: 429 QGPPGTGKTVTSATIVYHLSKIHKDR------ILVCAPSNVAVDHLAAKLRDLGLKVVRL 482
           QGPPGTGKTVTSA++VYHL+ +   +      +LVCAPSNVAVD LA K+   GL+VVR+
Sbjct: 423 QGPPGTGKTVTSASVVYHLATMQSRKRKSHSPVLVCAPSNVAVDQLAEKIHRTGLRVVRV 482

Query: 483 TAKSREDVESSVSNLALHNLVGRGA-KGELKNLLKLKDEVGELSASDTKRFVKLVRKTEA 541
            AKSRED+ESSVS L+LH  +       EL+ LLKL+ E  ELS  D K+   LV   E
Sbjct: 483 AAKSREDIESSVSFLSLHEQIKNYKFNPELQRLLKLRSENNELSIQDEKKLRILVAAAEK 542

Query: 542 EILNKADVVCCTCVGAGDKRLDT-KFRTVLIDESTQASEPECLIPIVKGAKQVILVGDHQ 600
           E+L  A V+CCTCVGAGD+R+   KFR+VLIDE+TQASEPEC+IP+V GAKQV+LVGDHQ
Sbjct: 543 ELLRAAHVICCTCVGAGDRRISKYKFRSVLIDEATQASEPECMIPLVLGAKQVVLVGDHQ 602

Query: 601 QLGPVILERKAADAGLKQSLFERLISLGHVPIRLEVQYRMNPYLSEFPSNMFYEGSLQNG 660
           QLGPV++ +K A A L QSLFERLI LG+ P RL VQYRM+P LSEFPSN FYEG+LQNG
Sbjct: 603 QLGPVVMNKKVALASLSQSLFERLIILGNSPFRLVVQYRMHPCLSEFPSNTFYEGTLQNG 662

Query: 661 VTIEQRTVPNSKFPWPIRGIPMMFWANYGREEISANGTSFLNRIEAMNCERIITKLFRDG 720
           VT  +R   +  FPW     P+MF+AN+G+EE+SA+GTSFLNR EA  CE+I+T   R
Sbjct: 663 VTTSERIARHVDFPWIQPDSPLMFYANFGQEELSASGTSFLNRTEASTCEKIVTTFLRSN 722

Query: 721 VKPEQIGVITPYEGQRAYILQYMQMNGSLDKDLYIKVEVASVDAFQGREKDYIILSCVRA 780
           V PEQIG++TPY+GQR+YI+QYMQ NGS+ KDLY  VEVASVDAFQGREKD+IILSCVR+
Sbjct: 723 VLPEQIGIVTPYDGQRSYIVQYMQNNGSMQKDLYKAVEVASVDAFQGREKDFIILSCVRS 782

Query: 781 NEQQAIGFLRDPRRLNVGLTRAKYGLVILGNPRSLARNTLWNHLLIHFREKGCLVEGTLD 840
           +E Q IGF+ DPRRLNV LTRAKYG+++LGNP+ LA++ LW H ++H +EKG LVEGTL+
Sbjct: 783 SEHQGIGFVNDPRRLNVALTRAKYGVIVLGNPKVLAKHALWYHFVLHCKEKGYLVEGTLN 842

Query: 841 NLQLCTVQLVRPQ-PRKTERPMNAQFNVESEMGDFPKFQDFDAQSMVSFSGQIGDFGNAF 899
           +LQ  ++ L  PQ P+K +R +N Q ++          Q+  +  + SFS     + +++
Sbjct: 843 SLQKFSLTLTPPQKPQKFKRDLNVQRSLS-------PIQNAGSAMLPSFSNLPNLYSSSY 895

Query: 900 VDNTELSSYINNEYWN---FENFKSAFSQKQNR 929
           ++   + +       N   FE+F+S     +++
Sbjct: 896 LEEWNVFAQYKRRESNATDFEDFRSQVGDDESK 928
```

emb|CAA91194.2| (Z54366) putative regulator of nonsense transcript stability [Schizosaccharomyces pombe]
Length = 925

Score = 902 bits (2306), Expect = 0.0
Identities = 465/931 (49%), Positives = 618/931 (65%), Gaps = 35/931 (3%)

```
Query: 19  VNVQPATQLNS-----TLVEDDDVDNQLFEEAQVTETGFRSPSASDNSCAYCGIDSAKCV 73
           + +QP   ++S      +  ++ +++Q FEE  V +        S+  CAYC I +   +
Sbjct: 3   LGLQPNNDISSLVSSKNMTSENGLEHQ-FEELLVEK------QYSEEHCAYCHIKNPNSI 55

Query: 74  IKCNSCKKWFCNTKNGTXXXXXXXXXXXXXXXXXXXPDSDLGDTVLECYNCGRKNVFLL 133
           +KC  C KWFCN +  +                     S L DTVLECYNCG +NVFLL
Sbjct: 56  LKCLHCNKWFCNVRGKSGASHIISHLVRARHKQVALHSHSSLSDTVLECYNCGTRNVFLL 115

Query: 134 GFVSAKSEAVVVLLCRIPCAQT---KNANWDTDQWQPLIEDRQLLSWVAEQPTEEEKLKA 190
           GF+ AK++ VVVLLCR PCA+    K+ NWD  QWQP+I DRQ L W+   P+EEE+  A
Sbjct: 116 GFIPAKAKTVVVLLCRQPCARASIAKDMNWDLTQWQPIISDRQFLPWLITPPSEEEQKLA 175

Query: 191 RLITPSQISKLEAKWRSNKDATINDIDAPEEQEAIPPLLLRYQDAYEYQRSYGPLIKLEA 250
             IT  Q+ +LE  WR + +A + D+D P E +++P + LRY+DA+ YQ    PLI+ EA
Sbjct: 176 IPITSQQMVRLEELWRKDPNANLEDLDKPIEDDSLPSVELRYKDAHAYQAVLSPLIQAEA 235

Query: 251 DYDKQLKESQALEHISVSWSLALNNRHLASFTLSTFESNELKVAIGDEMILWYSGMQHPD 310
           DYDK+LKESQ  + + V W  A+N R+ A F L   ES E+++AIGDEM L Y G
Sbjct: 236 DYDKRLKESQTQKDVVVRWDQAINKRYTAWFLLPKLESGEIRLAIGDEMKLTYEGELRAP 295

Query: 311 WEGRGYIVRLPNSFQDTFTLELKPSKTPPPTHLTTGFTAEFIWKGTSYDRMQDALKKFAI 370
           W   GY++++PN+  D   LELK S   P    T  F+ +++WK TS+DRMQ AL+ FA
Sbjct: 296 WSSTGYVIKIPNNVSDEVGLELKRSDKVP-IECTHNFSVDYVWKSTSFDRMQTALRLFAT 354

Query: 371 DKKSISGYLYYKILGHQVVDISFDVPLPKEFSIPNFAQLNSSQSNAVSHVLQRPLSLIQG 430
           D   +S +LY+K+LGH +        LP + S+PN  +LN+SQS AV  VL +PLSLIQG
Sbjct: 355 DGSRLSSFLYHKLLGHDIPPSFLKPKLPSDLSVPNLPKLNASQSEAVRAVLSKPLSLIQG 414

Query: 431 PPGTGKTVTSATIVYHLSKIHKDR------ILVCAPSNVAVDHLAAKLRDLGLKVVRLTA 484
           PPGTGKTVTSA++VYHL+ +   +      +LVCAPSNVAVD LA K+   GL+VVR+ A
Sbjct: 415 PPGTGKTVTSASVVYHLATMQSRKRKSHSPVLVCAPSNVAVDQLAEKIHRTGLRVVRVAA 474

Query: 485 KSREDVESSVSNLALHNLVGRGA-KGELKNLLKLKDEVGELSASDTKRFVKLVRKTEAEI 543
```

```
            KSRED+ESSVS L+LH  +       EL+ LLKL+ E  ELS  D K+   LV   E E+
Sbjct: 475 KSREDIESSVSFLSLHEQIKNYKFNPELQRLLKLRSENNELSIQDEKKLRILVAAAEKEL 534

Query: 544 LNKADVVCCTCVGAGDKRLDT-KFRTVLIDESTQASEPECLIPIVKGAKQVILVGDHQQL 602
           L  A V+CCTCVGAGD+R+   KFR+VLIDE+TQASEPEC+IP+V GAKQV+LVGDHQQL
Sbjct: 535 LRAAHVICCTCVGAGDRRISKYKFRSVLIDEATQASEPECMIPLVLGAKQVVLVGDHQQL 594

Query: 603 GPVILERKAADAGLKQSLFERLISLGHVPIRLEVQYRMNPYLSEFPSNMFYEGSLQNGVT 662
           GPV++ +K A A L QSLFERLI LG+ P RL VQYRM+P LSEFPSN FYEG+LQNGVT
Sbjct: 595 GPVVMNKKVALASLSQSLFERLIILGNSPFRLVVQYRMHPCLSEFPSNTFYEGTLQNGVT 654

Query: 663 IEQRTVPNSKFPWPIRGIPMMFWANYGREEISANGTSFLNRIEAMNCERIITKLFRDGVK 722
             +R   +  FPW     P+MF+AN+G+EE+SA+GTSFLNR EA  CE+I+T   R  V
Sbjct: 655 TSERIARHVDFPWIQPDSPLMFYANFGQEELSASGTSFLNRTEASTCEKIVTTFLRSNVL 714

Query: 723 PEQIGVITPYEGQRAYILQYMQMNGSLDKDLYIKVEVASVDAFQGREKDYIILSCVRANE 782
           PEQIG++TPY+GQR+YI+QYMQ NGS+ KDLY  VEVASVDAFQGREKD+IILSCVR++E
Sbjct: 715 PEQIGIVTPYDGQRSYIVQYMQNNGSMQKDLYKAVEVASVDAFQGREKDFIILSCVRSSE 774

Query: 783 QQAIGFLRDPRRLNVGLTRAKYGLVILGNPRSLARNTLWNHLLIHFREKGCLVEGTLDNL 842
            Q IGF+ DPRRLNV LTRAKYG+++LGNP+ LA++ LW H ++H +EKG LVEGTL++L
Sbjct: 775 HQGIGFVNDPRRLNVALTRAKYGVIVLGNPKVLAKHALWYHFVLHCKEKGYLVEGTLNSL 834

Query: 843 QLCTVQLVRPQ-PRKTERPMNAQFNVESEMGDFPKFQDFDAQSMVSFSGQIGDFGNAFVD 901
           Q  ++ L  PQ P+K +R +N Q ++          Q+  +  + SFS     + +++++
Sbjct: 835 QKFSLTLTPPQKPQKFKRDLNVQRSLS-------PIQNAGSAMLPSFSNLPNLYSSSYLE 887

Query: 902 NTELSSYINNEYWN---FENFKSAFSQKQNR 929
              + +       N   FE+F+S     +++
Sbjct: 888 EWNVFAQYKRRESNATDFEDFRSQVGDDESK 918
```

gb|AAC26789.1| (AF074017) nonsense-mediated mRNA decay trans-acting factor [Caenorhabditis elegans]
Length = 1069

Score = 763 bits (1949), Expect = 0.0
Identities = 394/836 (47%), Positives = 534/836 (63%), Gaps = 14/836 (1%)

```
Query: 32  VEDDDVDNQLFEEAQVTETGFRSPSASDNSCAYCGIDSAKCVIKCNSCKKWFCNTKNGTX 91
           VEDD+ D+    E  +TE         +++C YCGI    CV KC  C+KWFCN+ +GT
Sbjct: 70  VEDDESDS----EKSLTEEQHEQ-KLPEHACRYCGISDPLCVAKCTVCRKWFCNSNDGTS 124

Query: 92  XXXXXXXXXXXXXXXXXXXPDSDLGDTVLECYNCGRKNVFLLGFVSAKSEAVVVLLCRIP 151
                                DS  GDT LECY CG KNVF LGF+  K + VVV++CR P
Sbjct: 125 GGHIVHHMVRSQHKEAYTHKDSPCGDTQLECYRCGSKNVFNLGFIPGKKDQVVVIICRTP 184

Query: 152 CAQT---KNANWDTDQWQPLIEDRQLLSWVAEQPTEEEKLKARLITPSQISKLEAKWRSN 208
           CA      + NW  + W+ +I ++QLLSW+   P+EE+  +AR IT +Q  ++E  WR +
Sbjct: 185 CASIAFQNDDNWSPEDWKSVIAEKQLLSWIVNVPSEEQVARARKITATQAVRMEELWRDH 244

Query: 209 KDATINDIDAPEEQEAIPPLLLRYQDAYEYQRSYGPLIKLEADYDKQLKESQALEHISVS 268
            +AT++D++ P        + LRY DA+ Y + + PL+ +EA+YD+++KES +    +V
Sbjct: 245 PEATVDDLNKPGLDREPDHVQLRYVDAHHYSKVFRPLVAIEAEYDRRVKESASQAVGTVR 304

Query: 269 WSLALNNRHLASFTLSTFESNELKVAIGDEMILWYSG-MQHPDWEGRGYIVRLPNSFQDT 327
           W   L    LA F L  F    +K+A GDE+ L +S  +   +W   G + ++P++  D
Sbjct: 305 WEQGLRQSVLAFFHLPQFADGVMKLAKGDELRLKHSQTVDGSEWTKIGSVFKIPDNHGDE 364

Query: 328 FTLELKPSKTPPPTHLTTGFTAEFIWKGTSYDRMQDALKKFAIDKKSISGYLYYKILGHQ 387
             +E++ +           FT + +W  T+++R   AL     D K+IS YLY K+LGH
Sbjct: 365 VGIEIRGAVDKSVMESRIMFTVDVVWNATTFERQYKALAALLNDSKAISPYLYQKLLGHP 424

Query: 388 VVDISFDVPLPKEFSIPNFAQLNSSQSNAVSHVLQRPLSLIQGPPGTGKTVTSATIVYHL 447
             ++     LP+  S+    +LNSSQ  AV  VL RPLSLIQGPPGTGKTV SATIVYHL
Sbjct: 425 AEEMMLKFDLPRRLSVAGLPELNSSQMQAVKQVLTRPLSLIQGPPGTGKTVVSATIVYHL 484

Query: 448 SKIHKDRILVCAPSNVAVDHLAAKLRDLGLKVVRLTAKSREDVESSVSNLALHNLVGRGA 507
            +  +  +LVC+PSN+AVDHLA K+   GLKVVRL A+SRE  E++V  L L + +
Sbjct: 485 VQKTEGNVLVCSPSNIAVDHLAEKIHKTGLKVVRLCARSREHSETTVPYLTLQHQLKVMG 544

Query: 508 KGELKNLLKLKDEVGELSASDTKRFVKLVRKTEAEILNKADVVCCTCVGAGDKRLDT-KF 566
             EL+ L++LKDE GEL   D  R+++L R  E E+L  ADV+CCTC  A D RL   +
Sbjct: 545 GAELQKLIQLKDEAGELEFKDDLRYMQLKRVKEHELLAAADVICCTCSSAADARLSKIRT 604

Query: 567 RTVLIDESTQASEPECLIPIVKGAKQVILVGDHQQLGPVILERKAADAGLKQSLFERLIS 626
           RTVLIDESTQA+EPE L+ I++G +Q++LVGDH QLGPV++ +KAA AGL QSLFERL+
Sbjct: 605 RTVLIDESTQATEPEILVSIMRGVRQLVLVGDHCQLGPVVICKKAAIAGLSQSLFERLVL 664

Query: 627 LGHVPIRLEVQYRMNPYLSEFPSNMFYEGSLQNGVTIEQRTVPNSKFPWPIRGIPMMFWA 686
           LG  P RL+VQYRM+P LSEFPSN+FY+GSLQNGVT   R +    + WP    P  FW
Sbjct: 665 LGIRPFRLQVQYRMHPVLSEFPSNVFYDGSLQNGVTENDRHMTGVDWHWPKPNKPAFFWH 724

Query: 687 NYGREEISANGTSFLNRIEAMNCERIITKLFRDGVKPEQIGVITPYEGQRAYILQYMQMN 746
             G EE+SA+GTSFLNR EA N E++++KL + GV+P QIGVIT YEGQR++I+ YM
Sbjct: 725 CSGSEELSASGTSFLNRTEAANVEKLVSKLIKAGVQPHQIGVITSYEGQRSFIVNYMHTQ 784

Query: 747 GSLDKDLYIKVEVASVDAFQGREKDYIILSCVRANEQQAIGFLRDPRRLNVGLTRAKYGL 806
           G+L+  LY  VE+ASVDAFQGREKDYII++CVR+N+   IGFL DPRRLNV +TRAKYGL
```

```
Sbjct: 785 GTLNSKLYENVEIASVDAFQGREKDYIIVTCVRSNDILGIGFLSDPRRLNVAITRAKYGL 844

Query: 807 VILGNPRSLARNTLWNHLLIHFREKGCLVEGTLDNLQLCTVQLVRPQPRKTERPMN 862
           V++GN + LAR+ LW+ L+ H++ K  L EG ++ L+   + L    P+ T R  N
Sbjct: 845 VLVGNAKVLARHDLWHELINHYKSKEMLYEGPINALKPLNLAL----PKATIRTKN 896
```

sp|P32644|YE06_YEAST HYPOTHETICAL 127.0 KD PROTEIN IN RAD24-BMH1 INTERGENIC REGION
>gi|320741|pir||S30862 DNA dependent ATPase/DNA helicase
B - yeast (Saccharomyces cerevisiae) >gi|603417 (U18922)
Yer176wp [Saccharomyces cerevisiae]
Length = 1121

```
 Score =  231 bits (583), Expect = 2e-59
 Identities = 156/470 (33%), Positives = 249/470 (52%), Gaps = 52/470 (11%)

Query: 408  QLNSSQSNAVSHVLQRPLSLIQGPPGTGKTVTSATIVYH-LSKIHKDRILVCAPSNVAVD 466
            +LN SQ  AV HVL   ++++QGPPGTGKT T   I+   + + H   IL  A SN+A+D
Sbjct: 648  KLNRSQKTAVEHVLNNSITILQGPPGTGKTSTIEEIIIQVIERFHAFPILCVAASNIAID 707

Query: 467  HLAAKLRDLG--LKVVRLTAKSREDVESS---VSNLALHNLVGRGAKGELKNLLKLKDEV 521
            ++A K+ +    +K++R+ +K +E   S    +  + LHN+V +    +++ ++  K
Sbjct: 708  NIAEKIMENRPQIKILRILSKKKEQQYSDDHPLGEICLHNIVYKNLSPDMQ-VVANKTRR 766

Query: 522  GEL-SASDTKRFVKLVRKTEAEILNKADVVCCTCVGAGDKRLDT--KFRTVLIDESTQAS 578
            GE+ S S+  +F K   +   ++++++ ++  T + AG + L    +   V++DE+TQ+S
Sbjct: 767  GEMISKSEDTKFYKEKNRVTNKVVSQSQIIFTTNIAAGGRELKVIKECPVVIMDEATQSS 826

Query: 579  EPECLIPI-VKGAKQVILVGDHQQLGPVILERKAADAGLKQSLFERLISLGHV--PIRLE 635
            E   L+P+ + G +  + VGD +QL        +    L+ SLFER++S G    P+ L+
Sbjct: 827  EASTLVPLSLPGIRNFVFVGDEKQLSSF-----SNIPQLETSLFERVLSNGTYKNPLMLD 881

Query: 636  VQYRMNPYLSEFPSNMFYEGSLQNGVTIEQRTVPNSKFPWPIRGIPMMFWANYGREEISA 695
             QYRM+P +SEFP    Y G L++GVT EQ+  P  + P       +   +     +
Sbjct: 882  TQYRMHPKISEFPIKKIYNGELKDGVTDEQKAWPGVQHPLFFYQCDLGPESRVRSTQRDI 941

Query: 696  NGTSFLNRIEAMNCERIITKLFRDGVKP-EQIGVITPYEGQRAYILQYMQMN-------- 746
             G ++ N+ E +   +II  L  D   P E+IGVITPY  QR  +   +  N
Sbjct: 942  VGFTYENKHECVEIVKIIQILMLDKKVPLEEIGVITPYSAQRDLLSDILTKNVVINPKQI 1001

Query: 747  ---------------------GSLDK---DLYIKVEVASVDAFQGREKDYIILSCVRANE 782
                                 GSL     ++   + VA+VD+FQG EK +II SCVR N
Sbjct: 1002 SMQQEYDEIELFNAAGSQGTAGSLQNNVINIINGLHVATVDSFQGHEKSFIIFSCVRNNT 1061

Query: 783  QQAIGFLRDPRRLNVGLTRAKYGLVILGNPRSLAR-NTLWNHLLIHFREK 831
            +  IGFLRD RRLNV LTRAK+GL+++GN   L + + LW   + +  E+
Sbjct: 1062 ENKIGFLRDKRRLNVALTRAKHGLIVVGNKNVLRKGDPLWKDYITYLEEQ 1111
```

emb|CAB11611| (Z98951) hypothetical protein [Schizosaccharomyces pombe]
>gi|4239673|emb|CAA20863.1| (AL031546) putative
dna-binding protein [Schizosaccharomyces pombe]
Length = 660

```
 Score =  230 bits (580), Expect = 4e-59
 Identities = 176/513 (34%), Positives = 274/513 (53%), Gaps = 54/513 (10%)

Query: 357 SYDRMQDALKKFAIDKKSISGY---LYYKILGHQVVDISFDVPLPKEFSIPNFAQLNSSQ 413
           +Y+RM+  + +F   K+SI  Y   L+Y ++G +  D+S D  L  +    N  +LN+SQ
Sbjct: 159 TYERMRHTMLEF---KRSIPEYRNSLFYTLIGRKKADVSIDQKLIGDIKYFN-KELNASQ 214

Query: 414 SNAVSH-VLQRPLSLIQGPPGTGKTVTSATIVYHLSKIHKDRILVCAPSNVAVDHLAAKL 472
             AV   +  + LSLI GPPGTGKT T   I+  L   +K RILVC  SN+AVD++  +L
Sbjct: 215 KKAVKFSIAVKELSLIHGPPGTGKTHTLVEIIQQLVLRNK-RILVCGASNLAVDNIVDRL 273

Query: 473 RDLGLKVVRLTAKSREDVESSVSNLALHNLVGRGAKGE------------LKNLLKLKD- 519
              G+ +VRL   +R  +  S+ + +L  L   G  G+            L  + K K+
Sbjct: 274 SSSGIPMVRLGHPAR--LLPSILDHSLDVLSRTGDNGDVIRGISEDIDVCLSKITKTKNG 331

Query: 520 ----EVGELSASDTKRFVKLVRKTEAEILNKADVVCCTCVGAGDKRLD-TKFRTVLIDES 574
               E+ +      K + K   KT A I++ + VV CT  GAG ++L   +F  V+IDE+
Sbjct: 332 RERREIYKNIRELRKDYRKYEAKTVANIVSASKVVFCTLHGAGSRQLKGQRFDAVIIDEA 391

Query: 575 TQASEPECLIPIVKGAKQVILVGDHQQLGPVILERKAADAGLKQSLFERLI-SLGH-VPI 632
           +QA EP+C IP++ G  +VIL GDH QL P +  ++        S+FERL+ S G  V
Sbjct: 392 SQALEPQCWIPLL-GMNKVILAGDHMQLSPNVQSKRPYI-----SMFERLVKSQGDLVKC 445

Query: 633 RLEVQYRMNPYLSEFPSNMFYEGSLQNGVTIEQRTVPNSKF--PWPIRGIPMMFW---AN 687
            L +QYRM+  +S+FPS+ FY+  L     +++R + + +      +   P+ F+    N
Sbjct: 446 FLNIQYRMHELISKFPSDTFYDSKLVPAEEVKKRLLMDLENVEETELTDSPIYFYDTLGN 505

Query: 688 YGREEISANGTSFL-----NRIEAMNCERIITKLFRDGVKPEQIGVITPYEGQRAYILQY 742
           Y  ++ S +  +F      N  EA      I+ L   G++ + I V+TPY  Q A I Q
Sbjct: 506 YQEDDRSEDMQNFYQDSKSNHWEAQIVSYHISGLLEAGLEAKDIAVVTPYNAQVALIRQL 565

Query: 743 MQMNGSLDKDLYIKVEVASVDAFQGREKDYIILSCVRANEQQAIGFLRDPRRLNVGLTRA 802
           ++  G       I+VE+ SVD  QGREK+ II S VR+N+ + +GFL + RRLNV +TR
Sbjct: 566 LKEKG-------IEVEMGSVDKVQGREKEAIIFSLVRSNDVREVGFLAEKRRLNVAITRP 618
```

```
Query: 803 KYGLVILGNPRSLARNTLWNHLLIHFREKGCLV 835
           K  L ++G+  ++   + + H  + F E+  +V
Sbjct: 619 KRHLCVIGDSNTVKWASEFFHQWVDFLEENAIV 651
```

gi|2984278 (AE000770) DNA helicase [Aquifex aeolicus]
Length = 530

Score = 214 bits (540), Expect = 2e-54
Identities = 166/494 (33%), Positives = 248/494 (49%), Gaps = 85/494 (17%)

```
Query: 408 QLNSSQSNAVSHVLQ-RPLSLIQGPPGTGKTVTSATIVYHLSKIHKDRILVCAPS----- 461
           +LN+ Q  AV   L+   + LI GPPGTGKT T    +  L++    ++L  A S
Sbjct: 48  KLNAYQKKAVERALKAEKVFLIHGPPGTGKTTTLVECIKRLAQ-EGYKVLATADSNVAVD 106

Query: 462 ---------NVAVDHLAAKLR--------------------DLGLKVVRLTAKSREDVES 492
                    NV V  +   +R                    D   K+       +E+ +
Sbjct: 107 NLVERLVRENVKVVRVGNPVRVLKTIQRHTLDYLVQFEPEFDKARKIYEEIDNLKEEQKK 166

Query: 493 SVS-------NLALHNLVGRGAKG------------ELKNLLKLKDEVGELSASDTKRFV 533
            V         L+   ++ R   G             +   +KL+++V EL     K
Sbjct: 167 YVKPEPRYRRGLSDEEILKRAKTGTPVRGLSPKILRSMAKWIKLQEKVKELYEKAKKEEE 226

Query: 534 KLVRKTEAEILNKADVVCCTCVGAGDKRLDT-KFRTVLIDESTQASEPECLIPIVKGAKQ 592
           K V K    IL++A VVC T   AG + L    F  V+IDE+TQA+EP CLIP++KG K+
Sbjct: 227 KAVNK----ILSRAQVVCTTNSTAGSEVLQNLNFDVVIIDEATQATEPSCLIPLIKG-KK 281

Query: 593 VILVGDHQQLGPVILERKAADAGLKQSLFERLISLGHVPIR--LEVQYRMNPYLSEFPSN 650
           +I+ GDH+QL P +L ++A +A L  +LFERL+ L    I   L +QYRMN  + EF +
Sbjct: 282 LIMAGDHKQLPPTVLSQEAQEA-LSYTLFERLLDLYGEEIYEILRIQYRMNKKIMEFSNK 340

Query: 651 MFYEGSLQNGVTIEQRTVPNSKFPWPIRGIP------------MMFWANYGREEISANGT 698
           MFYEG L    ++E  T+ +   P  ++ IP            ++F    G+E+     T
Sbjct: 341 MFYEGKLIADKSVENHTIKDLINPEKLKEIPEPFKSVLEPEKVVVFINVRGKEKQRRGST 400

Query: 699 SFLNRIEAMNCERIITKLFRDGVKPEQIGVITPYEGQRAYILQYMQMNGSLDKDLYIKVE 758
           SF N  EA    +I+  L + G++ E IGVI+PYE Q  ++ + +       KD   +VE
Sbjct: 401 SFYNEEEAKVAVKIVEYLMKIGLRSEHIGVISPYEDQVNFLEELL-------KDF--EVE 451

Query: 759 VASVDAFQGREKDYIILSCVRANEQQAIGFLRDPRRLNVGLTRAKYGLVILGNPRSLARN 818
           V +VD FQGREK+ II+S VR+NE+  IGFL+D RRLNV LTRA+  L+ LGN ++L+ +
Sbjct: 452 VKTVDGFQGREKEVIIISFVRSNEKGEIGFLKDYRRLNVALTRARRKLITLGNEKTLSSD 511

Query: 819 TLWNHLLIHFREKG 832
            ++   + + +  G
Sbjct: 512 EVYKQFIGYVKSIG 525
```

sp|Q92355|YDHA_SCHPO HYPOTHETICAL 192.5 KD PROTEIN C6G9.10C IN CHROMOSOME I
>gi|1644322|emb|CAB03612.1| (Z81317) DNA2-NAM7 helicase
family protein [Schizosaccharomyces pombe]
Length = 1687

Score = 205 bits (515), Expect = 2e-51
Identities = 170/520 (32%), Positives = 247/520 (46%), Gaps = 120/520 (23%)

```
Query: 408  QLNSSQSNAVSHVLQRP-LSLIQGPPGTGKTVTSATIV---------YHLSKIH------ 451
            Q+N  Q+ A+   L     +LIQGPPGTGKT T   I+         YH+++ +
Sbjct: 1129 QVNEPQAKAIMCALDNNGFTLIQGPPGTGKTKTIIGIISALLVDLSRYHITRPNQQSKST 1188

Query: 452  --KDRILVCAPSNVAVDHLAAKLR-----DLGLKVVRLTAK--SREDVESSVSNLALHNL 502
              K +IL+CAPSN AVD +  +L+     + G K +    +  + E +  SV +L+L
Sbjct: 1189 ESKQQILLCAPSNAAVDEVLLRLKRGFLLENGEKYIPRVVRIGNPETINVSVRDLSLEYQ 1248

Query: 503  -------VGRGA--KGELKNLLKLKD----------------EVGELSASDTKRFVK--- 534
                  V +GA   G L+ L + +D                +V    A DTK   K
Sbjct: 1249 TEKQLLEVNQGAIDLGSLQELTRWRDTFYDCIQKIEELEKQIDVARDVAEDTKSLGKELQ 1308

Query: 535  ----------------------------LVRKTEAEILNKADVVCCTCVGAGDKRL--- 562
                                        L +K +  IL +ADVVC T  G+G   +
Sbjct: 1309 NKINEKNLAEQKVEELQSQSFTKNKEVDLLRKKAQKAILKQADVVCATLSGSGHDLVAHS 1368

Query: 563  DTKFRTVLIDESTQASEPECLIPIVKGAKQVILVGDHQQLGPVILERKAADAGLKQSLFE 622
               F TV+IDE+ QA E + +IP+  GAK+ ILVGD  QL P +L +KAA     QSLF
Sbjct: 1369 SLNFSTVIIDEAAQAVELDTIIPLRYGAKKCILVGDPNQLPPTVLSKKAASLNYSQSLFV 1428

Query: 623  RLI-SLGHVPIRLEVQYRMNPYLSEFPSNMFYEGSLQNGVTIEQRTVP----NSKFP--- 674
            R+  +  +    L +QYRM+P +S FPS  FY+  L++G  + ++T      N KF
Sbjct: 1429 RIQKNFSNQMCLLSIQYRMHPDISHFPSKKFYDSRLEDGDNMAEKTQQVWHVNPKFTQYR 1488

Query: 675  -WPIRGIPMMFWANYGREEISANGTSFLNRIEAMNCERIITKL---FRDGVKPEQIGVIT 730
             + +RG           +E ++N  S  N  E      ++ +L   F D     +IGVIT
Sbjct: 1489 LFDVRG-----------KERTSNTMSTYNLEEVEYLVNMVDELLNKFPDVNFTGRIGVIT 1537

Query: 731  PYEGQ-----RAYILQYMQMNGSLDKDLYIKVEVASVDAFQGREKDYIILSCVRANEQQA 785
            PY  Q     RA+ ++Y        K     +++ +VD FQG+EKD I  SCV++  +
Sbjct: 1538 PYRSQLHELRRAFKVKY-------GKSFMSTIDIQTVDGFQGQEKDIIFFSCVKSYSKHG 1590

Query: 786  IGFLRDPRRLNVGLTRAKYGLVILGNPRSLARNTLWNHLL 825
```

```
            IGFLRD RRLNV LTRA+  L+I+GN  +L  + LW  L+
Sbjct: 1591 IGFLRDFRRLNVALTRARSSLLIIGNMETLKTDDLWGSLV 1630
```

emb|CAB50143.1| (AJ248286) DNA helicase, putative [Pyrococcus abyssi]
Length = 653

Score = 202 bits (508), Expect = 1e-50
Identities = 167/501 (33%), Positives = 242/501 (47%), Gaps = 83/501 (16%)

```
Query: 403 IPNFAQLNSSQSNAVSHVL-QRPLSLIQGPPGTGKTVTSATIVYHLSKIHKDRILVCAPS 461
           +P    LN SQ  A+S  L      L+ GP GTGKT T   ++    K   +++L  A S
Sbjct: 165 VPFDKSLNRSQRKAISKALGSEDFFLVHGPFGTGKTRTLVELIRQEVK-RGNKVLATAES 223

Query: 462 NVAVDHLAAKLRDLGLKVVRLTAKSREDVESSVSNLAL----HNLVG------------- 504
           NVAVD+L  +L   G+K+VR+   SR       + LA     H L G
Sbjct: 224 NVAVDNLVERLSRSGIKIVRIGHPSRVSKHLHETTLAYLITQHELYGELRELRVIGQSLA 283

Query: 505 ----------------------------RGAKGELKNLLKLKDEVGELSASDTKRFV-- 533
                                       RGA+G    L+    E  +L+    K F
Sbjct: 284 EKRDTYTKPTPKFRRGLSDEEIIKLAERKRGARGLSARLIMEMAEWIKLNRQVQKAFDDA 343

Query: 534 -KLVRKTEAEILNKADVVCCTCVGAGDKRLD-TKFRTVLIDESTQASEPECLIPIVKGAK 591
            KL  +   +I+ +ADVV  T   A  + +D   +   +IDE+TQ++ P  LIP+ K  +
Sbjct: 344 RKLEERIARDIIREADVVLTTNSSAALEVVDYDTYDVAIIDEATQSTIPSILIPLNK-VE 402

Query: 592 QVILVGDHQQLGPVILERKAADAGLKQSLFERLISLGHVPIRLE---VQYRMNPYLSEFP 648
           + +L GDH+QL P IL  +A +  L ++LFE LI     P + E   VQYRMN  + EFP
Sbjct: 403 RFVLAGDHKQLPPTILSLEAQE--LSRTLFEGLIE--RYPWKSEMLVVQYRMNERIMEFP 458

Query: 649 SNMFYEGSL-----QNGVTIE---QRTVPNSKFPWPIRGIPMMFWANYGREEIS-----A 695
           S  FY G +       G+T+    +   PN  +   +    ++ + +  + E
Sbjct: 459 SKEFYGGKIIADESVRGITLRDLVEYQSPNDSWGKILNPENVIVFIDTSKAENKWERQRR 518

Query: 696 NGTSFLNRIEAMNCERIITKLFRDGVKPEQIGVITPYEGQRAYILQYMQMNGSLDKDLYI 755
              S  N +EA    +I+ KL   GVKPE IGVITPY+ QR  I   +  +
Sbjct: 519 GSESRENPLEAEIVAKIVDKLLSIGVKPEWIGVITPYDDQRDLISMKVPED--------- 569

Query: 756 KVEVASVDAFQGREKDYIILSCVRANEQQAIGFLRDPRRLNVGLTRAKYGLVILGNPRSL 815
            VEV +VD +QGREK+ IILS VR+N+   IGFL+D RRLNV LTRAK  L+++G+  +L
Sbjct: 570 -VEVKTVDGYQGREKEVIILSLVRSNKLGEIGFLKDLRRLNVSLTRAKRKLIVVGDSSTL 628

Query: 816 ARNTLWNHLLIHFREKGCLVE 836
           + +  +  ++  F+EKGC +E
Sbjct: 629 SSHETYKKMIEFFKEKGCFIE 649
```

gi|172574 (M74589) SEN1 [Saccharomyces cerevisiae]
Length = 2112

Score = 202 bits (508), Expect = 1e-50
Identities = 178/572 (31%), Positives = 278/572 (48%), Gaps = 126/572 (22%)

```
Query: 374  SISGYLYYKILGHQVVDISFDVPLPKEFS----IPNFAQLNSSQSNA-VSHVLQRPLSLI 428
            ++ G  YY ++G Q++      P+  + +    +    +LN+SQ+ A V+ V +   SLI
Sbjct: 1178 TLEGLEYYDLVG-QILQAKPSPPVNVDAAEIETVKKSYKLNTSQAEAIVNSVSKEGFSLI 1236

Query: 429  QGPPGTGKTVTSATIV-YHLSK-----------------------IHKDRILVCAPSNVA 464
            QGPPGTGKT T   I+ Y LS                        + K +IL+CAPSN A
Sbjct: 1237 QGPPGTGKTKTILGIIGYFLSTKNASSSNVIKVPLEKNSSNTEQLLKKQKILICAPSNAA 1296

Query: 465  VDHLAAKLRDLGLK-------VVRLTAKSREDVES-SVSNLALHNLVG------------ 504
            VD +  +L+  G+          +L    R DV + ++ +L L  LV
Sbjct: 1297 VDEICLRLKS-GVYDKQGHQFKPQLVRVGRSDVVNVAIKDLTLEELVDKRIGERNYEIRT 1355

Query: 505  ---------------RGAKGEL-------------KNLLKLKDEVGELSA--------SD 528
                           R  +G+L             +++ KL+ ++ ELS          D
Sbjct: 1356 DPELERKFNNAVTKRRELRGKLDSESGNPESPMSTEDISKLQLKIRELSKIINELGRDRD 1415

Query: 529  TKRFVKLV---------RKTEAEILNKADVVCCTCVGAGDKRLDT---KFRTVLIDESTQ 576
              R    V         R  +A IL  +D++C T  G+    L T   KF TV+IDE+ Q
Sbjct: 1416 EMREKNSVNYRNRDLDRRNAQAHILAVSDIICSTLSGSAHDVLATMGIKFDTVIIDEACQ 1475

Query: 577  ASEPECLIPIVKGAKQVILVGDHQQLGPVILERKAADAGLKQSLFERLISLGHVPIRLEV 636
             +E   +IP+  G K+ I+VGD  QL P +L   A++    QSLF R+      P  L+V
Sbjct: 1476 CTELSSIIPLRYGGKRCIMVGDPNQLPPTVLSGAASNFKYNQSLFVRM-EKNSSPYLLDV 1534

Query: 637  QYRMNPYLSEFPSNMFYEGSLQNGVTIEQRTVPNSKFPW----PIRGIPMMFW-ANYGRE 691
            QYRM+P +S+FPS+ FY+G L++G  ++      +K PW    P+   P  F+    GR+
Sbjct: 1535 QYRMHPSISKFPSSEFYQGRLKDGPGMDIL----NKRPWHQLEPL--APYKFFDIISGRQ 1588

Query: 692  EISANGTSFLNRIEAMNCERIITKLFRDGVK----PEQIGVITPYEGQ----RAYILQYM 743
            E +A   S+ N  E      ++  LFR          +IG+I+PY  Q    R    +Y
Sbjct: 1589 EQNAKTMSYTNMEEIRVAIELVDYLFRKFDNKIDFTGKIGIISPYREQMQKMRKEFARY- 1647

Query: 744  QMNGSLDKDLYIKVEVASVDAFQGREKDYIILSCVRANE-QQAIGFLRDPRRLNVGLTRA 802
               G ++K     ++  ++D FQG+EK+ I++SCVRA++ + ++GFL+D RR+NV LTRA
Sbjct: 1648 -FGGMINK----SIDFNTIDGFQGQEKEIILISCVRADDTKSSVGFLKDFRRMNVALTRA 1702
```

```
Query: 803  KYGLVILGNPRSLARNTLWNHLLIHFREKGCL 834
            K  + +LG+ RSLA++ LW  L+   +++ CL
Sbjct: 1703 KTSIWVLGHQRSLAKSKLWRDLIEDAKDRSCL 1734
```

sp|Q00416|SEN1_YEAST TRNA-SPLICING ENDONUCLEASE POSITIVE EFFECTOR >gi|1078364|pir||S53416
SEN1 protein - yeast (Saccharomyces cerevisiae)
>gi|664872 (U20939) Sen1p [Saccharomyces cerevisiae]
Length = 2231

```
 Score =  202 bits (508), Expect = 1e-50
 Identities = 178/572 (31%), Positives = 278/572 (48%), Gaps = 126/572 (22%)

Query: 374  SISGYLYYKILGHQVVDISFDVPLPKEFS----IPNFAQLNSSQSNA-VSHVLQRPLSLI 428
            ++ G  YY ++G Q++      P+  + +    +    +LN+SQ+ A V+ V +   SLI
Sbjct: 1297 TLEGLEYYDLVG-QILQAKPSPPVNVDAAEIETVKKSYKLNTSQAEAIVNSVSKEGFSLI 1355

Query: 429  QGPPGTGKTVTSATIV-YHLSK-----------------------IHKDRILVCAPSNVA 464
            QGPPGTGKT T   I+ Y LS                        + K +IL+CAPSN A
Sbjct: 1356 QGPPGTGKTKTILGIIGYFLSTKNASSSNVIKVPLEKNSSNTEQLLKKQKILICAPSNAA 1415

Query: 465  VDHLAAKLRDLGLK-------VVRLTAKSREDVES-SVSNLALHNLVG------------ 504
            VD +  +L+  G+          +L    R DV + ++ +L L  LV
Sbjct: 1416 VDEICLRLKS-GVYDKQGHQFKPQLVRVGRSDVVNVAIKDLTLEELVDKRIGERNYEIRT 1474

Query: 505  ---------------RGAKGEL-------------KNLLKLKDEVGELSA--------SD 528
                           R  +G+L              +++ KL+ ++ ELS          D
Sbjct: 1475 DPELERKFNNAVTKRRELRGKLDSESGNPESPMSTEDISKLQLKIRELSKIINELGRDRD 1534

Query: 529  TKRFVKLV---------RKTEAEILNKADVVCCTCVGAGDKRLDT---KFRTVLIDESTQ 576
              R    V         R  +A IL  +D++C T  G+    L T   KF TV+IDE+ Q
Sbjct: 1535 EMREKNSVNYRNRDLDRRNAQAHILAVSDIICSTLSGSAHDVLATMGIKFDTVIIDEACQ 1594

Query: 577  ASEPECLIPIVKGAKQVILVGDHQQLGPVILERKAADAGLKQSLFERLISLGHVPIRLEV 636
             +E   +IP+  G K+ I+VGD  QL P +L   A++    QSLF R+      P  L+V
Sbjct: 1595 CTELSSIIPLRYGGKRCIMVGDPNQLPPTVLSGAASNFKYNQSLFVRM-EKNSSPYLLDV 1653

Query: 637  QYRMNPYLSEFPSNMFYEGSLQNGVTIEQRTVPNSKFPW----PIRGIPMMFW-ANYGRE 691
            QYRM+P +S+FPS+ FY+G L++G  ++      +K PW    P+   P  F+    GR+
Sbjct: 1654 QYRMHPSISKFPSSEFYQGRLKDGPGMDIL----NKRPWHQLEPL--APYKFFDIISGRQ 1707

Query: 692  EISANGTSFLNRIEAMNCERIITKLFRDGVK----PEQIGVITPYEGQ----RAYILQYM 743
            E +A   S+ N  E      ++  LFR           +IG+I+PY  Q    R    +Y
Sbjct: 1708 EQNAKTMSYTNMEEIRVAIELVDYLFRKFDNKIDFTGKIGIISPYREQMQKMRKEFARY- 1766

Query: 744  QMNGSLDKDLYIKVEVASVDAFQGREKDYIILSCVRANE-QQAIGFLRDPRRLNVGLTRA 802
               G ++K     ++  ++D FQG+EK+ I++SCVRA++ + ++GFL+D RR+NV LTRA
Sbjct: 1767 -FGGMINK----SIDFNTIDGFQGQEKEIILISCVRADDTKSSVGFLKDFRRMNVALTRA 1821

Query: 803  KYGLVILGNPRSLARNTLWNHLLIHFREKGCL 834
            K  + +LG+ RSLA++ LW  L+   +++ CL
Sbjct: 1822 KTSIWVLGHQRSLAKSKLWRDLIEDAKDRSCL 1853
```

dbj|BAA30003| (AP000004) 656aa long hypothetical DNA-binding protein [Pyrococcus
horikoshii]
Length = 656

```
 Score =  200 bits (503), Expect = 5e-50
 Identities = 172/506 (33%), Positives = 241/506 (46%), Gaps = 91/506 (17%)

Query: 409 LNSSQSNAVSHVL-QRPLSLIQGPPGTGKTVTSATIVYHLSKIHKDRILVCAPSNVAVDH 467
           LNSSQ  A++  L      LI GP GTGKT T   ++    K   +++L  A SNVAVD+
Sbjct: 171 LNSSQRKAIAKGLGSEDFFLIHGPFGTGKTRTLVELIRQEVK-RGNKVLATAESNVAVDN 229

Query: 468 LAAKLRDLGLKVVRLTAKSREDVESSVSNLAL----HNLVG------------------- 504
           L  +L   G+K+VR+   SR       + LA     H L G
Sbjct: 230 LVERLAKDGVKIVRVGHPSRVSRHLHETTLAYLITQHELYGELRELRVIGQSLAEKRDTY 289

Query: 505 -----------------------RGAKGELKNLLKLKDEVGELSASDTKRFV---KLVRK 538
                                  RG +G    L+K   E  +L+    K F    KL  +
Sbjct: 290 TKPTPKFRRGLSDAEIIKLAERRRGIRGLSARLIKEMAEWIKLNRQVQKAFEDARKLEER 349

Query: 539 TEAEILNKADVVCCTCVGAGDKRLD-TKFRTVLIDESTQASEPECLIPIVKGAKQVILVG 597
              +I+ +ADV+  T   A  + +D T +   +IDE+TQA+ P  LIP+ K   + IL G
Sbjct: 350 IARDIIREADVILTTNSSAALEVVDATDYDVAIIDEATQATIPSILIPLNK-VDRFILAG 408

Query: 598 DHQQLGPVILERKAADAGLKQSLFERLISLGHVPIRLE---VQYRMNPYLSEFPSNMFYE 654
           DH+QL P IL  +A +  L  +LFE LI     P + E   +QYRMN  + EFPS  FY+
Sbjct: 409 DHKQLPPTILSLEAQE--LSHTLFEGLIE--RYPWKSEMLTIQYRMNERIMEFPSKEFYD 464

Query: 655 GSLQNGVTIEQRTVPNSKFPWPIRGIPMMFWANY----------------GREEISANGT 698
           G +     ++  T+ +        GI    W +                  R E    G+
Sbjct: 465 GRIVADERVKNITLGDLGIKVNATGI----WRDILDPSNVLVFIDTCMLDNRFERQRRGS 520

Query: 699 -SFLNRIEAMNCERIITKLFRDGVKPEQIGVITPYEGQRAYILQYMQMNGSLDKDLYIKV 757
             S  N +EA    +I+ KL   GVK E IGVITPY+ QR  I            ++  +V
Sbjct: 521 ESRENPLEAKIVSKIVEKLLESGVKAEMIGVITPYDDQRDLI----------SLNVPEEV 570
```

```
Query: 758 EVASVDAFQGREKDYIILSCVRANEQQAIGFLRDPRRLNVGLTRAKYGLVILGNPRSLAR 817
           EV +VD +QGREK+ IILS VR+N+   IGFL+D RRLNV LTRAK  L+++G+  +L+
Sbjct: 571 EVKTVDGYQGREKEVIILSFVRSNKVGEIGFLKDLRRLNVSLTRAKRKLIMIGDSSTLSS 630

Query: 818 NTLWNHLLIHFREKGCLVEGTLDNLQ 843
           +  +  L+ H REKG  V  T D ++
Sbjct: 631 HETYKRLIEHVREKGLYVVLTKDGIR 656
```

dbj|BAA31600| (AB014525) KIAA0625 protein [Homo sapiens]
Length = 791

Score = 199 bits (500), Expect = 1e-49
Identities = 165/526 (31%), Positives = 252/526 (47%), Gaps = 112/526 (21%)

```
Query: 425 LSLIQGPPGTGKTVTSATIVYHL---------------SKIHKDRILVCAPSNVAVDHLA 469
           + LI GPPGTGK+ T   ++Y L               +KI ++R+LVCAPSN AVD L
Sbjct: 72  ICLIHGPPGTGKSKTIVGLLYRLLTENQRKGHSDENSNAKIKQNRVLVCAPSNAAVDELM 131

Query: 470 AKL------------RDLG----LKVVRLTAKSREDVESSVSNLALHNLVGRGAKGEL-- 511
            K+              LG    + +VRL  +  + + S V   +L + V    K EL
Sbjct: 132 KKIILEFKEKCKDKKNPLGNCGDINLVRLGPE--KSINSEVLKFSLDSQVNHRMKKELPS 189

Query: 512 -------------------------------------KNLLKLKDEVGELSASDTKRFV 533
                                                +N+ K+  E  EL AS  K
Sbjct: 190 HVQAMHKRKEFLDYQLDELSRQRALCRGGREIQRQELDENISKVSKERQEL-ASKIKEVQ 248

Query: 534 KLVRKTEAEILNKADVVCCTCVGAGDKRLDTKFR--------TVLIDESTQASEPECLIP 585
              +KT++ I+ ++ ++CCT   +G   L++ FR         V++DE+ Q+ E E L P
Sbjct: 249 GRPQKTQSIIILESHIICCTLSTSGGLLLESAFRGQGGVPFSCVIVDEAGQSCEIETLTP 308

Query: 586 IVKGAKQVILVGDHQQLGPVILERKAADAGLKQSLFERLISL----------GHVPI-RL 634
           ++    ++ILVGD +QL P ++  KA + G  QS+  R   L            +PI +L
Sbjct: 309 LIHRCNKLILVGDPKQLPPTVISMKAQEYGYDQSMMARFCRLLEENVEHNMISRLPILQL 368

Query: 635 EVQYRMNPYLSEFPSNMFYEGSLQNGVTIEQRTVPNSKFPWPIRGIPMMFWANYGREEIS 694
            VQYRM+P +  FPSN  Y  +L+   T  Q         WP +   ++F    G E
Sbjct: 369 TVQYRMHPDICLFPSNYVYNRNLK---TNRQTEAIRCSSDWPFQPY-LVFDVGDGSER-- 422

Query: 695 ANGTSFLNRIEAMNCERIITKLFRDGVKP---EQIGVITPYEGQRAYILQYMQMNGSLDK 751
            +  S++N ++ +     I KL +D  K      IG+IT Y+ Q+  I +       LDK
Sbjct: 423 RDNDSYIN-VQEIKLVMEIIKLIKDKRKDVSFRNIGIITHYKAQKTMIQK------DLDK 475

Query: 752 DLYIK--VEVASVDAFQGREKDYIILSCVRANE-QQAIGFLRDPRRLNVGLTRAKYGLVI 808
           +   K   EV +VDAFQGR+KD +I++CVRAN  Q +IGFL   +RLNV +TRAKY L I
Sbjct: 476 EFDRKGPAEVDTVDAFQGRQKDCVIVTCVRANSIQGSIGFLASLQRLNVTITRAKYSLFI 535

Query: 809 LGNPRSLARNTLWNHLLIHFREKGCLVEGTLDNLQLCTVQLVRPQP 854
           LG+ R+L  N  WN L+   +++G +++    N +   V++++ +P
Sbjct: 536 LGHLRTLMENQHWNQLIQDAQKRGAIIKTCDKNYRHDAVKILKLKP 581
```

gi|2649188 (AE001009) DNA helicase, putative [Archaeoglobus fulgidus]
Length = 648

Score = 198 bits (498), Expect = 2e-49
Identities = 168/496 (33%), Positives = 243/496 (48%), Gaps = 87/496 (17%)

```
Query: 404 PNFAQLNSSQSNAVSHVLQRP-LSLIQGPPGTGKTVTSATIVYHLSKIHKDRILVCAPSN 462
           P  + LN +Q  AV   +      LI GP GTGKT T   +V  L K   +R+LV A SN
Sbjct: 164 PFDSSLNRAQLKAVGCAVSTDDFFLIHGPFGTGKTRTVVEVVRQLVK-RGERVLVTAESN 222

Query: 463 VAVDHLAAKLRDLGLKVVRLTAKSREDVESSVSNLALHNLV------------------- 503
            AVD+L   L D+  K+VRL   SR  VE  +    L +LV
Sbjct: 223 TAVDNLVELLSDM--KIVRLGHPSR--VEKRLKEHTLASLVLNHPDYKRIEEIKGKIEEI 278

Query: 504 ----------------------------GRGAKGELKNLLKLKDEVGELSASDTKRFVK 534
                                        RGA+G     ++   E  E   +  + + +
Sbjct: 279 ERRMERLTKPSPQLRRGLSDEEILRLARSNRGARGVAAKKIRSMAEWIEARKALDQLYTE 338

Query: 535 LVRKTEA---EILNKADVVCCTCVGAGDKRLDTKFRTVLIDESTQASEPECLIPIVKGAK 591
           +  + E    EI+ ++DVV  T   A    L+  F T +IDE++QA+ P  LIPI + A+
Sbjct: 339 MKEEEERIVKEIIEESDVVISTNSSAF--LLEESFDTAVIDEASQATIPSVLIPINR-AR 395

Query: 592 QVILVGDHQQLGPVILERKAADAGLKQSLFERLISLGHVPIRL-EVQYRMNPYLSEFPSN 650
           + IL GDH+QL P +++ +     L ++LFE+LI L     +L  VQYRMN  L EFPS
Sbjct: 396 KFILAGDHRQLPPTVMKAEK----LSETLFEKLIELYPEKSQLLNVQYRMNEKLMEFPSR 451

Query: 651 MFYEGSLQ-----NGVTIEQRTVPNSKFPWPIRGIPMMFWANYGR-----EEISANGTSF 700
            FY G +        + + Q     ++    I G   + + +  +     E   A+ TS
Sbjct: 452 EFYGGRIVAHESCTAIALSQIAKREAEKLREILGDEPLVFIDTSKCKNRWEGKLADSTSR 511

Query: 701 LNRIEAMNCERIITKLFRDGVKPEQIGVITPYEGQRAYILQYMQMNGSLDKDLYIKVEVA 760
            NR+EA     I+T+L + G+K EQIGVITPY+ Q   + +             + VEV+
Sbjct: 512 YNRLEAEIVTEIVTELLKMGLKKEQIGVITPYDDQVDLLREK------------VDVEVS 559

Query: 761 SVDAFQGREKDYIILSCVRANEQQAIGFLRDPRRLNVGLTRAKYGLVILGNPRSLARNTL 820
           SVD FQGREK+ II+S VR+N ++ IGFL D RRLNV LTRA+  L+++G+  +L+ N
Sbjct: 560 SVDGFQGREKEVIIISFVRSNRKREIGFLDDLRRLNVSLTRARRKLIMVGDSETLSVNGT 619
```

```
Query: 821 WNHLLIHFREKGCLVE 836
           +  L+ H + KG  VE
Sbjct: 620 YARLIDHVKRKGVYVE 635


 sp|Q57568|Y104_METJA HYPOTHETICAL ATP-BINDING PROTEIN MJ0104 >gi|2129217|pir||H64312
           probable DNA helicase homolog - Methanococcus jannaschii
           >gi|1590880 (U67467) DNA-binding protein, probably DNA
           helicase [Methanococcus jannaschii]
           Length = 663

 Score =  197 bits (496), Expect = 3e-49
 Identities = 173/533 (32%), Positives = 256/533 (47%), Gaps = 90/533 (16%)

Query: 357 SYDRMQDALKKFAIDKKSISGYLYYKILGHQVVDISFDVPLPKEFSIPNFAQLNSSQSNA 416
           ++ RM++AL++FA  +  ++    Y ILG +  +      +  EF   N   LN SQ  A
Sbjct: 140 TFKRMKEALREFARKRDKLA----YIILGIEHPEKPLREDIKLEFYDKN---LNESQKLA 192

Query: 417 VSH-VLQRPLSLIQGPPGTGKTVTSATIVYHLSKIHKDRILVCAPSNVAVDHLAAKL--R 473
           V   VL R L LI GPPGTGKT T   ++    K +K ++L  A SN+A D++   L  +
Sbjct: 193 VKKAVLSRDLYLIHGPPGTGKTRTITEVIVQEVKFNKHKVLATADSNIAADNILEYLIKK 252

Query: 474 DLGLKVVRLTAKSR--------------EDVESSVSNLALHNLVG--------------- 504
              LKVVR+   +R              E+ E     LAL   +
Sbjct: 253 YPDLKVVRVGHPTRISKDLIQHSLPYLIENHEKYQEILALREKIKEIKEQRDKFLKPSPR 312

Query: 505 --RGAKGELKNLLKL-----------KDEVGELSA-----SDTKRFV----KLVRKTEAE 542
             RG   E   +LK+           K+++  ++         KR +    ++  K   E
Sbjct: 313 WRRGMSDE--QILKVAKRKKSYRGIPKEKIVSMAEWIIRNKKIKRIINNLDEITEKIMNE 370

Query: 543 ILNKADVVCCTCVGAGDKRLDT-KFRTVLIDESTQASEPECLIPIVKGAKQVILVGDHQQ 601
           IL +ADV+  T   AG + L   +F  ++IDE +QA EP CLIPIVKG K +I+ GDH+Q
Sbjct: 371 ILAEADVIVATNSMAGSEILKGWEFDVIVIDEGSQAMEPSCLIPIVKGRK-LIMAGDHKQ 429

Query: 602 LGPVILERKAADAGLKQSLFERLIS-LGHVPIRLEVQYRMNPYLSEFPSNMFYEGSLQNG 660
           L P +L        LK++LFERLI         LE+QYRMN  + EFP+ MFY   L+
Sbjct: 430 LPPTVLSENEE---LKKTLFERLIKKYPEFSSILEIQYRMNEKIMEFPNKMFYNNKLKAD 486

Query: 661 VTIEQRTVPNSKFPWPI--------RGIPMMFWANYGREEISANGTSFLNRIEAMNCERI 712
            +++  T+ +      I          IP+ F    G E       S+ N  EA     I
Sbjct: 487 ESVKNITLLDLVKEEEIDEVDRDIINEIPVQFINVEGIERKDKESPSYYNIEEAEKVLEI 546

Query: 713 ITKLFRDGVKPEQIGVITPYEGQRAYILQYMQMNGSLDKDLYIKVEVASVDAFQGREKDY 772
           + KL +  +      VITPY+ Q  Y+ +  + +        I +EV +VD FQGRE +
Sbjct: 547 VKKLVKYKIPTN---VITPYDAQVRYLRRLFEEHN-------IDIEVNTVDGFQGRENEA 596

Query: 773 IILSCVRANEQQAIGFLRDPRRLNVGLTRAKYGLVILGNPRSLARNTLWNHLL 825
           I++S VR       GFL+D RRLNV +TRAK  L+++GN   L ++ ++N ++
Sbjct: 597 IVISFVRTKN---FGFLKDLRRLNVAITRAKRKLILIGNENLLKQDKVYNEMI 646


 gi|3548803 (AC005313) putative DNA-binding protein [Arabidopsis thaliana]
           >gi|4335770|gb|AAD17447| (AC006284) putative SMUBP-2
           [mouse] DNA-binding protein [Arabidopsis thaliana]
           Length = 635

 Score =  193 bits (486), Expect = 5e-48
 Identities = 162/521 (31%), Positives = 259/521 (49%), Gaps = 40/521 (7%)

Query: 339 PPTHLTTGFTAEFIWKGTSYDRMQDALKKFAIDKKSISGY---LYYKILGHQVVDISF-D 394
           P   L T    E +    +Y RM+D L    + K  + G    L   + G +   +S  D
Sbjct: 118 PEEGLNTSLRLEKLANEVTYRRMKDTL--IQLSKGVLRGPASDLVPVLFGERQPSVSKKD 175

Query: 395 VPLPKEFSIPNFAQLNSSQSNAVSHVLQ-RPLSLIQGPPGTGKTVTSATIVYHLSKIHKD 453
           V   K F+ P    L+ SQ +A++  L  + + L+ GPPGTGKT T   IV    K
Sbjct: 176 V---KSFT-PFNKNLDQSQKDAITKALSSKDVFLLHGPPGTGKTTTVVEIVLQEVK-RGS 230

Query: 454 RILVCAPSNVAVDHLAAKLRDLGLKVVRLTAKSR---EDVESSVS-------NLALHNLV 503
           +IL CA SN+AVD++  +L    +K+VR+   +R   + ++S++        N  L N +
Sbjct: 231 KILACAASNIAVDNIVERLVPHKVKLVRVGHPARLLPQVLDSALDAQVLKGDNSGLANDI 290

Query: 504 GRGAKGELKNLLKLKDE-VGELSASDTKRFVKLVRKTE----AEILNKADVVCCTCVGAG 558
            +  K     LLK KD+    L   + +   K  RK +    ++++  ADV+  T  GA
Sbjct: 291 RKEMKALNGKLLKAKDKNTRRLIQKELRTLGKEERKRQQLAVSDVIKNADVILTTLTGAL 350

Query: 559 DKRLDTK-FRTVLIDESTQASEPECLIPIVKGAKQVILVGDHQQLGPVILERKAADAGLK 617
            ++LD + F  V+IDE  QA E  C I ++KG++  IL GDH QL P I   +A   GL
Sbjct: 351 TRKLDNRTFDLVIIDEGAQALEVACWIALLKGSR-CILAGDHLQLPPTIQSAEAERKGLG 409

Query: 618 QSLFERLISL--GHVPIRLEVQYRMNPYLSEFPSNMFYEGSLQNGVTIEQRTVPN----S 671
           ++LFERL  L    +   L VQYRM+  +  + S   Y+  +    ++    + +    +
Sbjct: 410 RTLFERLADLYGDEIKSMLTVQYRMHELIMNWSSKELYDNKITAHSSVASHMLFDLENVT 469

Query: 672 KFPWPIRGIPMMFWANYGREEISANGTSFLNRIEAMNCERIITKLFRDGVKPEQIGVITP 731
           K       + ++  A    EE      S  N  EA        +L   GV+P  IG+ITP
Sbjct: 470 KSSSTEATLLLVDTAGCDMEEKKDEEESTYNEGEAEVAMAHAKRLMESGVQPSDIGIITP 529

Query: 732 YEGQRAYILQYMQMNGSLDKDLYIKVEVASVDAFQGREKDYIILSCVRANEQQAIGFLRD 791
```

```
           Y  Q   +L+ ++      KD+    E+++VD FQGREK+ II+S VR+N ++ +GFL+D
Sbjct: 530 YAAQ-VMLLRILRGKEEKLKDM----EISTVDGFQGREKEAIIISMVRSNSKKEVGFLKD 584

Query: 792 PRRLNVGLTRAKYGLVILGNPRSLARNTLWNHLLIHFREKG 832
            RR+NV +TR++    I+ +  +++ +     ++ +F E G
Sbjct: 585 QRRMNVAVTRSRRQCCIVCDTETVSSDAFLKRMIEYFEEHG 625
```

sp|P34243|YKB7_YEAST HYPOTHETICAL 78.3 KD PROTEIN IN RAM2-ATP7 INTERGENIC REGION
>gi|422147|pir||S34700 probable purine
nucleotide-binding protein YKL017c - yeast
(Saccharomyces cerevisiae) >gi|395256|emb|CAA52266|
(X74152) orf4; homologous to human putative GTP-binding
protein [Saccharomyces cerevisiae]
>gi|486007|emb|CAA81852| (Z28017) ORF YKL017c
[Saccharomyces cerevisiae]
Length = 683

Score = 192 bits (482), Expect = 1e-47
Identities = 156/511 (30%), Positives = 259/511 (50%), Gaps = 63/511 (12%)

```
Query: 357 SYDRMQDALKKFAIDKKSISGYLYYKILGHQVVDISFDVPLPKEF-SIPNFAQ--LNSSQ 413
           +Y+RM+  ++K +     IS  +  KI+ + V +  F +P    F +I +F    LN SQ
Sbjct: 158 TYNRMESTMRKLS----EISSPIQDKIIQYLVNERPF-IPNTNSFQNIKSFLNPNLNDSQ 212

Query: 414 SNAVSHVLQRPLSLIQGPPGTGKTVTSATIVYHLS-KIHKDRILVCAPSNVAVDHLAAKL 472
             A++  +   L++I GPPGTGKT T   ++  L  K  ++RIL+C PSN++VD +  +L
Sbjct: 213 KTAINFAINNDLTIIHGPPGTGKTFTLIELIQQLLIKNPEERILICGPSNISVDTILERL 272

Query: 473 -----RDLGLKVVRLTAKSREDVESSVSNLALHNLVGRGAKGELKNLLKLKDEVGELSA- 526
                 +L L++         +   S+  L+  N + +    E+  L++   ++
Sbjct: 273 TPLVPNNLLLRIGHPARLLDSNKRHSLDILSKKNTIVKDISQEIDKLIQENKKLKNYKQR 332

Query: 527 SDTKRFVKLVRK--------TEAEILNKADVVCCTCVGAGDKRLDTKFR---------TV 569
            +    +KL+RK        T  +++ ++ +V  T  G+  + L + +R         T+
Sbjct: 333 KENWNEIKLLRKDLKKREFKTIKDLIIQSRIVVTTLHGSSSRELCSLYRDDPNFQLFDTL 392

Query: 570 LIDESTQASEPECLIPIVKGAKQ---VILVGDHQQLGPVIL--ERKAADAGLKQSLFERL 624
           +IDE +QA EP+C IP++    Q   ++L GD++QL P I   + K     L+ +LF+R+
Sbjct: 393 IIDEVSQAMEPQCWIPLIAHQNQFHKLVLAGDNKQLPPTIKTEDDKNVIHNLETTLFDRI 452

Query: 625 ISLGHVPIR-----LEVQYRMNPYLSEFPSNMFYEGSLQNGVTIEQRTVPN--SKFPWPI 677
           I +   P R     L VQYRMN  + EFPS+  Y G L    T+  R + +  +    P
Sbjct: 453 IKI--FPKRDMVKFLNVQYRMNQKIMEFPSHSMYNGKLLADATVANRLLIDLPTVDATPS 510

Query: 678 RG-----IPMMFWANYGRE------EISANGTSFLNRIEAMNCERIITKLFRDGVKPEQI 726
                 IP++++   G E      E +  G+ + N  E    +  I  L    V    I
Sbjct: 511 EDDDDTKIPLIWYDTQGDEFQETADEATILGSKY-NEGEIAIVKEHIENLRSFNVPENSI 569

Query: 727 GVITPYEGQRAYILQYMQMNGSLDKDLYIKVEVASVDAFQGREKDYIILSCVRANEQQAI 786
           GVI+PY  Q +++ + +      D+     +E+++VD FQGREKD IILS VR+NE+  +
Sbjct: 570 GVISPYNAQVSHLKKLIH-----DELKLTDIEISTVDGFQGREKDVIILSLVRSNEKFEV 624

Query: 787 GFLRDPRRLNVGLTRAKYGLVILGNPRSLAR 817
           GFL++ RRLNV +TR +  LV++GN   L R
Sbjct: 625 GFLKEERRLNVAMTRPRRQLVVVGNIEVLQR 655
```

emb|CAA22438| (AL034463) tRNA-splicing endonuclease positive effector
[Schizosaccharomyces pombe]
Length = 1944

Score = 190 bits (477), Expect = 5e-47
Identities = 167/576 (28%), Positives = 261/576 (44%), Gaps = 105/576 (18%)

```
Query: 347  FTAEFIWKGTSYDRMQDALKKFAIDKKSISGYLYYKILGHQVVDISFDVPLPKEFSIPNF 406
            FTA+ ++  T+      +L++FA  K      L  +IL   V  +  +    K+  I
Sbjct: 1215 FTAQKLFNATT------SLREFAALKSLRHLPLSQRILDANVTRLPSNFTDDKKQKIMKS 1268

Query: 407  AQLNSSQSNAV-SHVLQRPLSLIQGPPGTGKTVTSATIVYHL----------------SK 449
              +N  Q+ A+ +  +    +LIQGPPGTGKT T   ++  +                 K
Sbjct: 1269 YGVNEPQAYAIYASSVNDGFTLIQGPPGTGKTKTILGMIGAVLTSSSQGLQFNVPGQTRK 1328

Query: 450  IHKDRILVCAPSNVAV------------DHLAAKLRDLGLKV-----VRLTAKSREDVES 492
              K++IL+CAPSN A+            DH   K     ++V     + + AK     E
Sbjct: 1329 TSKNKILICAPSNAAIDEILLRIKAGVYDHEGIKFFPKVIRVGFGDSISVHAKEFTLEEQ 1388

Query: 493  SVSNLALHNL----------------------------------VGRGA---KGELKNL 514
             +  + L NL                                   G+ +   + +L+ +
Sbjct: 1389 MIKQMELTNLKKDQEANNSSDTRKKYDSIIKKRDSLREDLEKFRSTGKNSSILEAQLREI 1448

Query: 515  LKLK-------DEVGELSASDTKRFVKLVRKTEAEILNKADVVCCTCVGAGDKRL---DT 564
             K K       D++ E   S  +    L ++ + ++L +AD+VC T   +G + L
Sbjct: 1449 TKQKNMLEQSLDDMRERQRSTNRNLDVLKKQIQNQLLQEADIVCATLSASGHELLLNAGL 1508

Query: 565  KFRTVLIDESTQASEPECLIPIVKGAKQVILVGDHQQLGPVILERKAADAGLKQSLFERL 624
             FRTV+IDE+ QA E   +IP+  G +  ++VGD  QL P +L + +A  G  QSL+ R+
Sbjct: 1509 TFRTVIIDEAAQAVELSSIIPLKYGCESCVMVGDPNQLPPTVLSKTSAKFGYSQSLYVRM 1568
```

```
Query: 625  ISL-GHVPIRLEVQYRMNPYLSEFPSNMFYEGSLQNGVTIEQRTVPNSKFPW---PIRGI 680
                     L +QYRMNP +S FPS  FY   L +G  +   T      PW   P  GI
Sbjct: 1569 FKQHNESACLLSIQYRMNPEISRFPSKFFYNSKLLDGPNMSAVT----SRPWHEDPQLGI 1624

Query: 681  PMMFWANYGREEISANGTSFLNRIEAMNCERIITKLFRDGVKPE---QIGVITPYEGQRA 737
               F  N    E  +N  S  N  EA     +  +L +  +  +   +IGV+TPY  Q
Sbjct: 1625 YRFF--NVHGTEAFSNSKSLYNVEEASFILLLYERLIQCYLNIDFEGKIGVVTPYRSQVQ 1682

Query: 738  YIL-QYMQMNGSLDKDLYIKVEVASVDAFQGREKDYIILSCVRANEQQAIGFLRDPRRLN 796
             +  Q+ +  GS+   ++  +++ +VD FQG+EKD II SCVR++    IGFL+D RRLN
Sbjct: 1683 QLRSQFQRKYGSI---IFKHLDIHTVDGFQGQEKDIIIFSCVRSSMSGGIGFLQDLRRLN 1739

Query: 797  VGLTRAKYGLVILGNPRSLARNTLWNHLLIHFREKG 832
            V LTRAK  L I+GN + L +  ++  L+   + +G
Sbjct: 1740 VALTRAKSSLYIVGNSKPLMQEDIFYSLIEDAKTRG 1775
```

gi|2622761 (AE000922) transcriptional control factor (enhancer-binding protein) [Methanobacterium thermoautotrophicum]
Length = 642

Score = 188 bits (472), Expect = 2e-46
Identities = 165/500 (33%), Positives = 240/500 (48%), Gaps = 82/500 (16%)

```
Query: 408 QLNSSQSNAVSHVL-QRPLSLIQGPPGTGKTVTSATIVYHLSKIHKDRILVCAPSNVAVD 466
           +LN SQ +A+   L      LI GP GTGKT T   ++     +   R+LV A SN AVD
Sbjct: 156 ELNESQRDAIRMALGSEDFFLIHGPFGTGKTRTLHELIRQ-EVMRGSRVLVTAESNAAVD 214

Query: 467 HLAAKLRDLGLKVVRLTAKSREDVESSVSNLALHNLVGRGAKGELKNLLKLKDEVGE--- 523
           +L   +    +K VRL    R   E+    LA + +       +++   +  DE+ E
Sbjct: 215 NLLEGIAG-HVKCVRLGHPQRVSRENLRETLA-YKIENHPEYSKVREYQEKIDELIEERE 272

Query: 524 ----------LSASDTKRFVKLVRKT-------------------------------EAE 542
                        SDT+  +   ++                                 EAE
Sbjct: 273 RHQKPTPQIRRGLSDTQILINATKRRGARGISPNVMISMARWIETNQRIDDLHSKLQEAE 332

Query: 543 ------ILNKADVVCCTCVGAGDKRLD-TKFRTVLIDESTQASEPECLIPIVKGAKQVIL 595
                IL ++ VV  T   A  + +D  +F   ++DE++QA+ P  LIP+ + A + IL
Sbjct: 333 MRIADRILRESQVVLSTNSSAALEYIDGLRFDVAIVDEASQATIPSILIPLAR-APRFIL 391

Query: 596 VGDHQQLGPVILERKAADAGLKQSLFERLISLGHVPIR-LEVQYRMNPYLSEFPSNMFYE 654
            GDH+QL P IL R A++  L+++LFE LI       R L  QYRM+P + EFP+  FY+
Sbjct: 392 AGDHRQLPPTILSRDASE--LERTLFEELIKRHPGRSRMLNCQYRMHPAIMEFPNREFYD 449

Query: 655 GSLQNGVTIEQRTVPNSKFPWPIRGI--------PMMFWANYGR---EEISANGTSFLNR 703
           G ++   ++E  ++ +     P   I        P++F    G    E      TS  N
Sbjct: 450 GRIRAHPSLEDISIRDIIEDVPDSDICQKLADPDPVLFIDTSGLDGCERRLKGSTSIQNP 509

Query: 704 IEAMNCERIITKLFRDGVKPEQIGVITPYEGQRAYILQYMQMNGSLDKDLYIKVEVASVD 763
           +EA     I   L R GVKPE+IG+ITPY+ Q   I               I VEV SVD
Sbjct: 510 LEADLAVIISRSLMRMGVKPEEIGIITPYDDQVDLISSM-----------IDVEVNSVD 557

Query: 764 AFQGREKDYIILSCVRANEQQAIGFLRDPRRLNVGLTRAKYGLVILGNPRSLARNTLWNH 823
            FQGREKD II+S VR+N   +IGFL+D RRLNV LTRA+  L+I+G+ R+L+ +  +
Sbjct: 558 GFQGREKDVIIISMVRSNRNGSIGFLKDLRRLNVSLTRARRKLIIIGDSRTLSAHPSYRR 617

Query: 824 LLIHFREKGCLVEGTLDNLQ 843
           L    R++G L E  LD+++
Sbjct: 618 LTEFCRKRGFLDEAGLDDVR 637
```

gi|3935147 (AC005106) T25N20.11 [Arabidopsis thaliana]
Length = 1021

Score = 183 bits (461), Expect = 4e-45
Identities = 153/453 (33%), Positives = 216/453 (46%), Gaps = 64/453 (14%)

```
Query: 403 IPNFAQLNSSQSNAVSHVLQ---RPLSLIQGPPGTGKTVTSATIVYHLSKIHKD-RILVC 458
           +P    LN+ Q  ++  VL     P  +I GPPGTGKT+T    +  L    ++ R+LVC
Sbjct: 410 VPISPALNAEQICSIEMVLGCKGAPPYVIHGPPGTGKTMTLVEAIVQLYTTQRNARVLVC 469

Query: 459 APSNVAVDHLAAKLRDLGLKVVRLTAKSREDVESSVSNLALHNLVGRGAKGELKNLLKLK 518
           APSN A DH+  KL  L L+ VR+  K  E    + +  +   +     K E+
Sbjct: 470 APSNSAADHILEKL--LCLEGVRI--KDNEIFRLNAATRSYEEI-----KPEIIRFCFFD 520

Query: 519 DEVGELSASDTKRFVKLVRKT--EAEILNKADVVCCTCVGAGDKRLDTKFRTVLIDESTQ 576
           + + +          KLV  T   A +LN   V        G       F  +L+DE+ Q
Sbjct: 521 ELIFKCPPLKALTRYKLVVSTYMSASLLNAEGV------NRGH------FTHILLDEAGQ 568

Query: 577 ASEPECLIPIVK---GAKQVILVGDHQQLGPVILERKAADAGLKQSLFERLISLGHV--- 630
           ASEPE +I +         V+L GD +QLGPVI  R A   GL +S  ERL    +
Sbjct: 569 ASEPENMIAVSNLCLTETVVVLAGDPRQLGPVIYSRDAESLGLGKSYLERLFECDYYCEG 628

Query: 631 ----PIRLEVQYRMNPYLSEFPSNMFYEGSL------QNGVTIEQRTVPNSKFPWPIRGI 680
                 +L   YR +P + + PS +FY+G L       + V      +PN +FP
Sbjct: 629 DENYVTKLVKNYRCHPEILDLPSKLFYDGELVASKEDTDSVLASLNFLPNKEFP------ 682

Query: 681 PMMFWANYGREEISANGTSFLNRIEAMNCERIITKLF-RDGVKPEQIGVITPYEGQRAYI 739
            M+F+   G +E   N  S+ NRIE       I +L   D V+ E IGVITPY  Q
```

```
Sbjct: 683 -MVFYGIQGCDEREGNNPSWFNRIEISKVIETIKRLTANDCVQEEDIGVITPYRQQ---- 737

Query: 740 LQYMQMNGSLDKDLYIKVEVASVDAFQGREKDYIILSCVRA-------NEQQAIGFLRDP 792
              M++   LD+    +V+V SV+ FQG+EK  II+S VR+       +    +GFL +P
Sbjct: 738 --VMKIKEVLDRLDMTEVKVGSVEQFQGQEKQVIIISTVRSTIKHNEFDRAYCLGFLSNP 795

Query: 793 RRLNVGLTRAKYGLVILGNPRSLARNTLWNHLL 825
           RR NV +TRA   LVI+GNP  + ++  WN LL
Sbjct: 796 RRFNVAITRAISLLVIIGNPHIICKDMNWNKLL 828
```

gb|AAD35099.1|AE001689_5 (AE001689) DNA helicase, putative [Thermotoga maritima]
Length = 650

Score = 183 bits (459), Expect = 7e-45
Identities = 179/523 (34%), Positives = 251/523 (47%), Gaps = 94/523 (17%)

```
Query: 383 ILGHQVVDISFDVPLPKEFSIPNFAQLNSSQSNAVSHVL-QRPLSLIQGPPGTGKTVTSA 441
           +LG +  + SF+    +EF+ P    LN SQ  AVS  L      LI GP GTGKT T
Sbjct: 150 LLGKRKPEESFE----EEFT-PFDEGLNESQREAVSLALGSSDFFLIHGPFGTGKTRT-- 202

Query: 442 TIVYHLSKIHKDR-ILVCAPSNVAVDHLAAKLRDLGLKVVRLTAKSREDVESSVSNLALH 500
            + Y   ++ + + ILV A SN+AVD+L  +L    + +VR+   SR       S LA H
Sbjct: 203 LVEYIRQEVARGKKILVTAESNLAVDNLVERLWGK-VSLVRIGHPSRVSSHLKESTLA-H 260

Query: 501 NLVG-------RGAKGELKNLLKLKDEVGELSA------SDTK---------------- 530
            +         +  K EL  L+K +D   + S       SD K
Sbjct: 261 QIETSSEYEKVKKMKEELAKLIKKRDSFTKPSPQWRRGLSDKKILEYAEKNWSARGVSKE 320

Query: 531 ------RFVKL-----------VRKTE---AEILNKADVVCCTCVGAGDKRLD-TKFRTV 569
                 ++KL            RK E   + I+ +A VV  T   A  + L    F  V
Sbjct: 321 KIKEMAEWIKLNSQIQDIRDLIERKEEIIASRIVREAQVVLSTNSSAALEILSGIVFDVV 380

Query: 570 LIDESTQASEPECLIPIVKGAKQVILVGDHQQLGPVILERKAADAGLKQSLFERLISLGH 629
           ++DE++QA+ P  LIPI KG K+ +L GDH+QL P IL   A D  L ++LFE LI+
Sbjct: 381 VVDEASQATIPSILIPISKG-KKFVLAGDHKQLPPTILSEDAKD--LSRTLFEELIT--R 435

Query: 630 VPIR---LEVQYRMNPYLSEFPSNMFYEGSLQNGVTIEQRT-------VPNSKFPWPIRG 679
            P +   L+ QYRMN  L EFPS  FY+G L+    +   T       +PN    W +
Sbjct: 436 YPEKSSLLDTQYRMNELLMEFPSEEFYDGKLKAAEKVRNITLFDLGVEIPNFGKFWDVVL 495

Query: 680 IP---MMFWANYGR----EEISANGTSFLNRIEAMNCERIITKLFRDGVKPEQIGVITPY 732
            P   ++F     R    E    +  S  N +EA   + ++ KL   GVK + IG+ITPY
Sbjct: 496 SPKNVLVFIDTKNRSDRFERQRKDSPSRENPLEAQIVKEVVEKLLSMGVKEDWIGIITPY 555

Query: 733 EGQRAYILQYMQMNGSLDKDLYIKVEVASVDAFQGREKDYIILSCVRANEQQAIGFLRDP 792
           + Q   I + ++           KVEV SVD FQGREK+ II+S VR+N+   IGFL D
Sbjct: 556 DDQVNLIRELIEA----------KVEVHSVDGFQGREKEVIIISFVRSNKNGEIGFLEDL 605

Query: 793 RRLNVGLTRAKYGLVILGNPRSLARNTLWNHLLIHFREKGCLV 835
           RRLNV LTRAK  L+  G+  +L+ +  +   +   ++KG  V
Sbjct: 606 RRLNVSLTRAKRKLIATGDSSTLSVHPTYRRFVEFVKKKGTYV 648
```

emb|CAB43845.1| (AL078464) putative protein [Arabidopsis thaliana]
Length = 1311

Score = 178 bits (448), Expect = 1e-43
Identities = 167/593 (28%), Positives = 251/593 (42%), Gaps = 153/593 (25%)

```
Query: 392  SFDVPLPKEFSIPNFAQLN-------SSQSNAVSHVLQRPLSLIQGPPGTGK--TVTSAT 442
            SF   L + F+ P  A ++       +  S+ V      P +L+QGPPGTGK  TV
Sbjct: 621  SFAGHLHRSFNAPQLAAIHWAAMHTAAGTSSGVKRQDPWPFTLVQGPPGTGKTHTVWGML 680

Query: 443  IVYHLSKIH-------------------------------------------------- 451
             V HL +
Sbjct: 681  NVIHLVQYQQYYTSLLKKLAPESYKQVNESSSDNIVSGSIDEVLQNMDQNLFRTLPKLCA 740

Query: 452  KDRILVCAPSNVAVDHLAAKLRDLG-----LKVVR------------------------- 481
            K R+LVCAPSN A D L +++ D G     ++V R
Sbjct: 741  KPRMLVCAPSNAATDELLSRVLDRGFIDGEMRVYRPDVARVGVDTQTKAAQAVSVERRSD 800

Query: 482  -LTAKSREDVESSVSNLALHNLV--------------------GRGAKGELKNLLKLKDE 520
             L AKSRE++   + NL + +                      +G+ G   ++L ++D+
Sbjct: 801  LLLAKSREEILGHIHNLRVRDAQLSQDIAGLKRELTAAAFANRSQGSVGVDPDVLMVRDQ 860

Query: 521  VGE-----LSASDTKRFVKLVRKT------------------------EAEILNKADVVC 551
              +     LSA    R   LV  +                        EA   N+A++V
Sbjct: 861  TRDAMLQRLSAVVEARDKDLVEMSRLLIVEGKFRAGTSFNLEEARASLEASFANEAEIVF 920

Query: 552  CTCVGAGDK---RLDTKFRTVLIDESTQASEPECLIPIVKGAKQVILVGDHQQLGPVILE 608
             T   +G K   RL   F  V+IDE+ QASE   L P+  GA + +LVGD QQL   ++
Sbjct: 921  TTVSSSGRKLFSRLTHGFDMVVIDEAAQASEVGVLPPLALGAARCVLVGDPQQLPATVIS 980

Query: 609  RKAADAGLKQSLFERLISLGHVPIRLEVQYRMNPYLSEFPSNMFYEGSLQNGVTIEQRTV 668
            + A      +SLFER    G   + L VQYRM+P + +FPS  FY+G L +  ++   T
Sbjct: 981  KAAGTLLYSRSLFERFQLAGCPTLLLTVQYRMHPQIRDFPSRYFYQGRLTDSESVS--TA 1038

Query: 669  PNSKFPWPIRGIPMMFW-ANYGREEISANGTSFLNRIEAMNCERIITKLFRD----GVKP 723
```

```
            P+  +       P +F+  ++GRE       S+ N  EA  C  +   L R     G
Sbjct: 1039 PDEIYYKDSVLKPYLFFDISHGRESHRGGSVSYENIDEARFCVGVYLHLQRTLKSLGGGK 1098

Query: 724  EQIGVITPYEGQRAYILQYMQMNGSLDKDLYIKVEVASVDAFQGREKDYIILSCVRANEQ 783
              +GVITPY+ Q   +   ++   +L +D   ++ + +VDAFQG+E+D II+SCVRA+
Sbjct: 1099 VSVGVITPYKLQLKCL--KIEFGNALSQDELQEIYINTVDAFQGQERDVIIMSCVRAS-N 1155

Query: 784  QAIGFLRDPRRLNVGLTRAKYGLVILGNPRSLARNTLWNHLLIHFREKGCLVE 836
              +GF+ D RR+NV LTRAK  L ++GN  +L +   W  L+   + + C +E
Sbjct: 1156 HGVGFVADIRRMNVALTRAKRALWVMGNASALMKCEDWAALITDAKARNCFME 1208
```

sp|Q60560|SMB2_MESAU DNA-BINDING PROTEIN SMUBP-2 (INSULIN II GENE ENHANCER-BINDING PROTEIN) (RIPE3B-BINDING COMPLEX 3B2 P110 SUBUNIT) (RIP-1) >gi|290919 (L15625) insulin II gene enhancer-binding protein [Mesocricetus auratus]
Length = 989

```
 Score =  178 bits (447), Expect = 2e-43
 Identities = 158/510 (30%), Positives = 250/510 (48%), Gaps = 53/510 (10%)

Query: 357 SYDRMQDALKKFAIDKKSISGYLYYKILGHQVVDISFDVPLPKEFSIPNFAQLNSSQSNA 416
           +Y R++ AL          +  L   +LG      + ++P    F+  N A L+ SQ  A
Sbjct: 143 TYKRLKKALMTLKKYHSGPASSLIDVLLGGSSPSPTTEIP---PFTFYNTA-LDPSQKEA 198

Query: 417 VSHVL-QRPLSLIQGPPGTGKTVTSATIVYHLSKIHKDRILVCAPSNVAVDHLAAKLRDL 475
           VS  L Q+ +++I GPPGTGKT T   I+    K    +IL CAPSNVAVD+L  +L
Sbjct: 199 VSFALAQKEVAIIHGPPGTGKTTTVVEIILQAVK-QGLKILCCAPSNVAVDNLVERLALC 257

Query: 476 GLKVVRLTAKSREDVESSVSNLALHNLVGRGAKGEL-KNLLKLKDEV--------GELSA 526
             +++RL   +R  +  S    +L  ++ R    ++  ++ K  D+V         +
Sbjct: 258 KKRILRLGHPAR--LLESAQQHSLDAVLARSDNAQIVADIRKDIDQVFGKNKKTQDKREK 315

Query: 527 SDTKRFVKLVRKT-----EAEI---LNKADVVCCTCVGA---GDKRL--DTKFRTVLIDE 573
           S+ +  +KL+RK      EA I   L  ADVV  T  GA   G  +L  +  F  V++DE
Sbjct: 316 SNFRNEIKLLRKELKEREEAAIVQSLTAADVVLATNTGASSDGPLKLLPENHFDVVVVDE 375

Query: 574 STQASEPECLIPIVKGAKQVILVGDHQQLGPVILERKAADAGLKQSLFERLISL-GHVPI 632
             QA E  C IP++K  K  IL GDH+QL P  +  KAA AGL +SL ERL+   G   +
Sbjct: 376 CAQALEASCWIPLLKAPK-CILAGDHRQLPPTTISHKAALAGLSRSLMERLVEKHGAGAV 434

Query: 633 R-LEVQYRMNPYLSEFPSNMFYEGSLQNGVTIEQRTVPN--SKFPWPIRGIPMMFWANYG 689
           R L VQYRM+  ++ + S   Y G L    ++    + +           +P++     G
Sbjct: 435 RMLTVQYRMHQAITRWASEAMYHGQLTAHPSVAGHLLKDLPGVADTEETSVPLLLIDTAG 494

Query: 690 -------REEISANGTSFLNRIEAMNCERIITKLFRDGVKPEQIGVITPYEGQRAYILQY 742
                  E+  + G     R+  ++ + ++      GV    I VI PY  Q   + Q
Sbjct: 495 CGLLELDEEDSQSKGNPGEVRLVTLHIQALVDA----GVHAGDIAVIAPYNLQVDLLRQS 550

Query: 743 MQMNGSLDKDLYIKVEVASVDAFQGREKDYIILSCVRANEQQAIGFLRDPRRLNVGLTRA 802
           +        + + ++E+ SVD FQGREK+ +IL+ VR+N +  +GFL + RR+NV +TRA
Sbjct: 551 L-------SNKHPELEIKSVDGFQGREKEAVILTFVRSNRKGEVGFLAEDRRINVAVTRA 603

Query: 803 KYGLVILGNPRSLARNTLWNHLLIHFREKG 832
           +  + ++ + R++  +     L+ +F E G
Sbjct: 604 RRHVAVICDSRTVNNHAFLKTLVDYFTEHG 633
```

pir||A47500 Ig mu chain switch region binding protein 2 - human
Length = 993

```
 Score =  176 bits (443), Expect = 5e-43
 Identities = 145/461 (31%), Positives = 235/461 (50%), Gaps = 55/461 (11%)

Query: 409 LNSSQSNAVSHVL-QRPLSLIQGPPGTGKTVTSATIVYHLSKIHKDRILVCAPSNVAVDH 467
           L++SQ  AVS  L Q+ L++I GPPGTGKT T   I+    K    ++L CAPSN+AVD+
Sbjct: 192 LDTSQKEAVSFALSQKELAIIHGPPGTGKTTTVVEIILQAVK-QGLKVLCCAPSNIAVDN 250

Query: 468 LAAKLRDLGLKVVRLTAKSREDVESSVSNLALHNLVGRGAKGELKNLLKLKDEVGEL--- 524
           L  +L     +++RL   +R  +  SV   +L  ++ R    +  N+  ++ ++ ++
Sbjct: 251 LVERLALCKQRILRLGHPAR--LLESVQQHSLDAVLARSDSAQ--NVADIRKDIDQVFVK 306

Query: 525 --------SASDTKRFVKLVRKT-----EA---EILNKADVVCCTCVGA---GDKRL--D 563
                   S+ +  +KL+RK      EA   E L  A+VV  T  GA   G  +L  +
Sbjct: 307 NKKTQDKREKSNFRNEIKLLRKELKEREEAAMLESLTSANVVLATNTGASADGPLKLLPE 366

Query: 564 TKFRTVLIDESTQASEPECLIPIVKGAKQVILVGDHQQLGPVILERKAADAGLKQSLFER 623
           + F  V+IDE  QA E  C IP++K A++ IL GDH+QL P  +  KAA AGL  SL ER
Sbjct: 367 SYFDVVVIDECAQALEASCWIPLLK-ARKCILAGDHKQLPPTTVSHKAALAGLSLSLMER 425

Query: 624 LISL--GHVPIRLEVQYRMNPYLSEFPSNMFYEGSLQNGVTIEQ---RTVPNSKFPWPIR 678
           L       V   L VQYRM+  +  + S+  Y G +    ++ +   R +P
Sbjct: 426 LAEEYGARVVRTLTVQYRMHQAIMRWASDTMYLGQVTAHSSVARHLLRDLPGVAATEET- 484

Query: 679 GIPMMFWANYG-------REEISANGTSFLNRIEAMNCERIITKLFRDGVKPEQIGVITP 731
           G+P++     G        E+  + G     R+ +++ + ++      GV    I V++P
Sbjct: 485 GVPLLLVDTAGCGLFELEEEDEQSKGNPGEVRLVSLHIQALVDA----GVPARDIAVVSP 540

Query: 732 YEGQRAYILQYMQMNGSLDKDLYIKVEVASVDAFQGREKDYIILSCVRANEQQAIGFLRD 791
```

```
           Y  Q   + Q +          + ++E+ SVD FQGREK+ +ILS VR+N +  +GFL +
Sbjct: 541 YNLQVDLLRQSLVHR-------HPELEIKSVDGFQGREKEAVILSFVRSNRKGEVGFLAE 593

Query: 792 PRRLNVGLTRAKYGLVILGNPRSLARNTLWNHLLIHFREKG 832
            RR+NV +TRA+  + ++ + R++  +     L+ +F + G
Sbjct: 594 DRRINVAVTRARRHVAVICDSRTVNNHAFLKTLVEYFTQHG 634
```

ref|NP_002171.1|PIGHMBP2| immunoglobulin mu binding protein 2 >gi|730751|sp|P38935|SMB2_HUMAN
DNA-BINDING PROTEIN SMUBP-2 (GLIAL FACTOR-1) (GF-1)
>gi|401776 (L14754) DNA-binding protein [Homo sapiens]
Length = 993

```
 Score =  176 bits (443), Expect = 5e-43
 Identities = 145/461 (31%), Positives = 235/461 (50%), Gaps = 55/461 (11%)

Query: 409 LNSSQSNAVSHVL-QRPLSLIQGPPGTGKTVTSATIVYHLSKIHKDRILVCAPSNVAVDH 467
           L++SQ  AVS  L Q+ L++I GPPGTGKT T   I+    K    ++L CAPSN+AVD+
Sbjct: 192 LDTSQKEAVSFALSQKELAIIHGPPGTGKTTTVVEIILQAVK-QGLKVLCCAPSNIAVDN 250

Query: 468 LAAKLRDLGLKVVRLTAKSREDVESSVSNLALHNLVGRGAKGELKNLLKLKDEVGEL--- 524
           L  +L     +++RL   +R  +  SV   +L  ++ R    +  N+  ++ ++ ++
Sbjct: 251 LVERLALCKQRILRLGHPAR--LLESVQQHSLDAVLARSDSAQ--NVADIRKDIDQVFVK 306

Query: 525 --------SASDTKRFVKLVRKT-----EA---EILNKADVVCCTCVGA---GDKRL--D 563
                   S+ +  +KL+RK      EA   E L  A+VV  T  GA   G  +L  +
Sbjct: 307 NKKTQDKREKSNFRNEIKLLRKELKEREEAAMLESLTSANVVLATNTGASADGPLKLLPE 366

Query: 564 TKFRTVLIDESTQASEPECLIPIVKGAKQVILVGDHQQLGPVILERKAADAGLKQSLFER 623
           + F  V+IDE  QA E  C IP++K A++ IL GDH+QL P  +  KAA AGL  SL ER
Sbjct: 367 SYFDVVVIDECAQALEASCWIPLLK-ARKCILAGDHKQLPPTTVSHKAALAGLSLSLMER 425

Query: 624 LISL--GHVPIRLEVQYRMNPYLSEFPSNMFYEGSLQNGVTIEQ---RTVPNSKFPWPIR 678
           L       V   L VQYRM+  +  + S+  Y G +    ++ +   R +P
Sbjct: 426 LAEEYGARVVRTLTVQYRMHQAIMRWASDTMYLGQVTAHSSVARHLLRDLPGVAATEET- 484

Query: 679 GIPMMFWANYG-------REEISANGTSFLNRIEAMNCERIITKLFRDGVKPEQIGVITP 731
           G+P++     G        E+  + G     R+ +++ + ++      GV    I V++P
Sbjct: 485 GVPLLLVDTAGCGLFELEEEDEQSKGNPGEVRLVSLHIQALVDA----GVPARDIAVVSP 540

Query: 732 YEGQRAYILQYMQMNGSLDKDLYIKVEVASVDAFQGREKDYIILSCVRANEQQAIGFLRD 791
           Y  Q   + Q +          + ++E+ SVD FQGREK+ +ILS VR+N +  +GFL +
Sbjct: 541 YNLQVDLLRQSLVHR-------HPELEIKSVDGFQGREKEAVILSFVRSNRKGEVGFLAE 593

Query: 792 PRRLNVGLTRAKYGLVILGNPRSLARNTLWNHLLIHFREKG 832
            RR+NV +TRA+  + ++ + R++  +     L+ +F + G
Sbjct: 594 DRRINVAVTRARRHVAVICDSRTVNNHAFLKTLVEYFTQHG 634
```

gi|2340994 (U21094) Sen1p [Saccharomyces cerevisiae]
Length = 757

```
 Score =  176 bits (441), Expect = 9e-43
 Identities = 119/334 (35%), Positives = 189/334 (55%), Gaps = 30/334 (8%)

Query: 518 KDEVGELSASDTKRFVKLVRKTEAEILNKADVVCCTCVGAGDKRLDT---KFRTVLIDES 574
           +DE+ E ++ + +      R  +A IL  +D++C T  G+    L T   KF TV+IDE+
Sbjct: 59  RDEMREKNSVNYRNRDLDRRNAQAHILAVSDIICSTLSGSAHDVLATMGIKFDTVIIDEA 118

Query: 575 TQASEPECLIPIVKGAKQVILVGDHQQLGPVILERKAADAGLKQSLFERLISLGHVPIRL 634
            Q +E   +IP+  G K+ I+VGD  QL P +L   A++    QSLF R+      P  L
Sbjct: 119 CQCTELSSIIPLRYGGKRCIMVGDPNQLPPTVLSGAASNFKYNQSLFVRM-EKNSSPYLL 177

Query: 635 EVQYRMNPYLSEFPSNMFYEGSLQNGVTIEQRTVPNSKFPW----PIRGIPMMFWANY-G 689
           +VQYRM+P +S+FPS+ FY+G L++G  ++      +K PW    P+   P  F+    G
Sbjct: 178 DVQYRMHPSISKFPSSEFYQGRLKDGPGMDIL----NKRPWHQLEPLA--PYKFFDIISG 231

Query: 690 REEISANGTSFLNRIEAMNCERIITKLFRDGVKP----EQIGVITPYEGQ----RAYILQ 741
           R+E +A   S+ N  E      ++  LFR          +IG+I+PY  Q    R    +
Sbjct: 232 RQEQNAKTMSYTNMEEIRVAIELVDYLFRKFDNKIDFTGKIGIISPYREQMQKMRKEFAR 291

Query: 742 YMQMNGSLDKDLYIKVEVASVDAFQGREKDYIILSCVRANE-QQAIGFLRDPRRLNVGLT 800
           Y    G ++K     ++  ++D FQG+EK+ I++SCVRA++ + ++GFL+D RR+NV LT
Sbjct: 292 YF--GGMINKS----IDFNTIDGFQGQEKEIILISCVRADDTKSSVGFLKDFRRMNVALT 345

Query: 801 RAKYGLVILGNPRSLARNTLWNHLLIHFREKGCL 834
           RAK  + +LG+ RSLA++ LW  L+   +++ CL
Sbjct: 346 RAKTSIWVLGHQRSLAKSKLWRDLIEDAKDRSCL 379
```

gi|908917 (L24544) DNA helicase [Homo sapiens]
Length = 993

```
 Score =  176 bits (441), Expect = 9e-43
 Identities = 145/459 (31%), Positives = 232/459 (49%), Gaps = 51/459 (11%)

Query: 409 LNSSQSNAVSHVL-QRPLSLIQGPPGTGKTVTSATIVYHLSKIHKDRILVCAPSNVAVDH 467
           L++SQ  AV   L Q+ L++I GPPGTGKT T   I+    K    ++L CAPSN+AVD+
Sbjct: 192 LDTSQKEAVLFALSQKELAIIHGPPGTGKTTTVVEIILQAVK-QGLKVLCCAPSNIAVDN 250
```

```
Query: 468 LAAKLRDLGLKVVRLTAKSREDVESSVSNLALHNLVGRGAKGEL---------KNLLKLK 518
           L  +L     +++RL   +R  +  S+   +L  ++ R    ++         +  +K K
Sbjct: 251 LVERLALCKQRILRLGHPAR--LLESIQQHSLDAVLARSDSAQIVADIRKDIDQVFVKNK 308

Query: 519 DEVGELSASDTKRFVKLVRKT-----EA---EILNKADVVCCTCVGA---GDKRL--DTK 565
               +   S+ +  +KL+RK      EA   E L  A+VV  T  GA   G  +L  ++
Sbjct: 309 KTQDKREKSNFRNEIKLLRKELKEREEAAMLESLTSANVVLATNTGASADGPLKLLPESY 368

Query: 566 FRTVLIDESTQASEPECLIPIVKGAKQVILVGDHQQLGPVILERKAADAGLKQSLFERLI 625
           F  V+IDE  QA E  C IP++K A++ IL GDH+QL P  +  KAA AGL  SL ERL
Sbjct: 369 FDVVVIDECAQALEASCWIPLLK-ARKCILAGDHKQLPPTTVSHKAALAGLSLSLMERLA 427

Query: 626 SL--GHVPIRLEVQYRMNPYLSEFPSNMFYEGSLQNGVTIEQ---RTVPNSKFPWPIRGI 680
                 V   L VQYRM+  +  + S+  Y G L    ++ +   R +P         G+
Sbjct: 428 EEYGARVVRTLTVQYRMHQAIMRWASDTMYLGQLTAHSSVARHLLRDLPGVAATEET-GV 486

Query: 681 PMMFWANYG-------REEISANGTSFLNRIEAMNCERIITKLFRDGVKPEQIGVITPYE 733
           P++     G        E+  + G     R+ +++ + ++      GV    I V++PY
Sbjct: 487 PLLLVDTAGCGLFELEEEDEQSKGNPGEVRLVSLHIQALVDA----GVPARDIAVVSPYN 542

Query: 734 GQRAYILQYMQMNGSLDKDLYIKVEVASVDAFQGREKDYIILSCVRANEQQAIGFLRDPR 793
            Q   + Q +          + ++E+ SVD FQGREK+ +ILS VR+N +  +GFL + R
Sbjct: 543 LQVDLLRQSLVHR-------HPELEIKSVDGFQGREKEAVILSFVRSNRKGEVGFLAEDR 595

Query: 794 RLNVGLTRAKYGLVILGNPRSLARNTLWNHLLIHFREKG 832
           R+NV +TRA+  + ++ + R++  +     L+ +F + G
Sbjct: 596 RINVAVTRARRHVAVICDSRTVNNHAFLKTLVEYFTQHG 634
```

gi|3176714 (AC002392) putative tRNA-splicing endonuclease positive effector [Arabidopsis thaliana]
Length = 1090

Score = 173 bits (433), Expect = 7e-42
Identities = 143/481 (29%), Positives = 222/481 (45%), Gaps = 95/481 (19%)

```
Query: 452  KDRILVCAPSNVAVDHLAAKLRDLG-----LKVVR------------------------- 481
            K R+LVCAPSN A D L +++ D G     ++V R
Sbjct: 578  KPRMLVCAPSNAATDELLSRVLDRGFIDGEMRVYRPDVARVGVDSQSRAAQAVSVERRSD 637

Query: 482  -LTAKSREDVESSVSNLALHN--------------------------------LVGRGA 507
             L A SR+++   + NL L                                  L+ R
Sbjct: 638  QLLAISRDEILRHMRNLRLQETQISQNIAGLKRELNAAAFATRSQGSVGVDPEVLISRDQ 697

Query: 508  KGE-----LKNLLKLKDEV-----------GELSASDTKRFVKLVRKTEAEILNKADVVC 551
            K +     L  +++ +D+V           G+  A +     +     EA   N+A++V
Sbjct: 698  KRDALLQHLAAVVEARDKVLVEISRLLIVEGKFRAGNNFNLEEARASLEASFANEAEIVF 757

Query: 552  CTCVGAGDK---RLDTKFRTVLIDESTQASEPECLIPIVKGAKQVILVGDHQQLGPVILE 608
             T   +G K   RL   F  V+IDE+ QASE   L P+  GA + +LVGD QQL   ++
Sbjct: 758  TTVSSSGRKLFSRLTHGFDMVVIDEAAQASEVGVLPPLALGAARCVLVGDPQQLPATVIS 817

Query: 609  RKAADAGLKQSLFERLISLGHVPIRLEVQYRMNPYLSEFPSNMFYEGSLQNGVTIEQRTV 668
            + A      +SLFER    G   + L VQYRM+P + +FPS  FY+G L++  +I   +
Sbjct: 818  KAAGTLLYSRSLFERFQLAGCPTLLLTVQYRMHPQIRDFPSRYFYQGRLKDSESIS--SA 875

Query: 669  PNS-KFPWPIRGIPMMFWANYGREEISANGTSFLNRIEAMNCERIITKLFRD----GVKP 723
            P+   +  P+    + F  ++GRE       S+ N  EA  C  +   L +     G
Sbjct: 876  PDEIYYKDPVLRPYLFFNISHGRESHRGGSVSYENVDEARFCVGVYMHLQKTLKSLGAGK 935

Query: 724  EQIGVITPYEGQRAYILQYMQMNGSLDKDLYIKVEVASVDAFQGREKDYIILSCVRANEQ 783
              +GVITPY+ Q   +    +   +L +D   ++ + +VDAFQG+E+D II+SCVRA+
Sbjct: 936  VSVGVITPYKLQLKCLKH--EFGNALGQDELKEIYINTVDAFQGQERDVIIMSCVRAS-G 992

Query: 784  QAIGFLRDPRRLNVGLTRAKYGLVILGNPRSLARNTLWNHLLIHFREKGCLVEGTLDNLQ 843
              +GF+ D RR+NV LTRA+  L ++GN  +L ++  W  L+   R + C +E  +D+L
Sbjct: 993  HGVGFVSDIRRMNVALTRARRALWVMGNASALMKSEDWAALISDARGRNCFME--MDSLP 1050

Query: 844  L 844
            L
Sbjct: 1051 L 1051
```

Score = 33.6 bits (75), Expect = 7.0
Identities = 20/55 (36%), Positives = 28/55 (50%), Gaps = 7/55 (12%)

```
Query: 392 SFDVPLPKEFSIPNFAQLN-------SSQSNAVSHVLQRPLSLIQGPPGTGKTVT 439
           SF   L + F+ P  A ++       +  S+ V      P +L+QGPPGTGKT T
Sbjct: 457 SFVDHLHRSFNAPQLAAIHWAAMHTAAGTSSGVKKQEPWPFTLVQGPPGTGKTHT 511
```

sp|P40694|SMB2_MOUSE DNA-BINDING PROTEIN SMUBP-2 >gi|423421|pir||S35633 DNA-binding protein - mouse >gi|293806 (L10075) DNA-binding protein [Mus musculus]
Length = 993

Score = 169 bits (425), Expect = 6e-41
Identities = 152/511 (29%), Positives = 249/511 (47%), Gaps = 55/511 (10%)

```
Query: 357 SYDRMQDALKKFAIDKKSISGYLYYKILGHQVVDISFDVPLPKEFSIPNFAQLNSSQSNA 416
           +Y R++ AL          +  L   +LG      + ++P    ++      L+ SQ  A
Sbjct: 143 TYKRLKKALMTLKKYHSGPASSLIDILLGSSTPSPAMEIPPLSFYN----TTLDLSQKEA 198

Query: 417 VSHVL-QRPLSLIQGPPGTGKTVTSATIVYHLSKIHKDRILVCAPSNVAVDHLAAKLRDL 475
           VS  L Q+ L++I GPPGTGKT T   I+    K    ++L CAPSN+AVD+L  +L
Sbjct: 199 VSFALAQKELAIIHGPPGTGKTTTVVEIILQAVK-QGLKVLCCAPSNIAVDNLVERLALC 257

Query: 476 GLKVVRLTAKSREDVESSVSNLALHNLVGRGAKGEL-KNLLKLKDEV-GELSASDTKRF- 532
             +++RL   +R  +  SV + +L  ++ R    ++  ++ +  D+V G+   +  KR
Sbjct: 258 KKRILRLGHPAR--LLESVQHHSLDAVLARSDNAQIVADIRRDIDQVFGKNKKTQDKREK 315

Query: 533 ------VKLVRKT-----EAEI---LNKADVVCCTCVGA---GDKRL--DTKFRTVLIDE 573
                 +KL+RK      EA I   L  ADVV  T  GA   G  +L  +  F  V++DE
Sbjct: 316 GNFRSEIKLLRKELKEREEAAIVQSLTAADVVLATNTGASSDGPLKLLPEDYFDVVVVDE 375

Query: 574 STQASEPECLIPIVKGAKQVILVGDHQQLGPVILERKAADAGLKQSLFERLISLGHVPI- 632
             QA E  C IP++K  K  IL GDH+QL P  +  +AA AGL +SL ERL       +
Sbjct: 376 CAQALEASCWIPLLKAPK-CILAGDHRQLPPTTVSHRAALAGLSRSLMERLAEKHGAGVV 434

Query: 633 -RLEVQYRMNPYLSEFPSNMFYEGSLQNGVTIEQ---RTVPNSKFPWPIRGIPMMFWANY 688
             L VQYRM+  +  + S   Y G   +  ++     + +P        R +P++
Sbjct: 435 RMLTVQYRMHQAIMCWASEAMYHGQFTSHPSVAGHLLKDLPGVTDTEETR-VPLLLIDTA 493

Query: 689 G-------REEISANGTSFLNRIEAMNCERIITKLFRDGVKPEQIGVITPYEGQRAYILQ 741
           G        E+  + G     R+  ++ + ++      GV+   I VI PY  Q   + Q
Sbjct: 494 GCGLLELEEEDSQSKGNPGEVRLVTLHIQALVDA----GVQAGDIAVIAPYNLQVDLLRQ 549

Query: 742 YMQMNGSLDKDLYIKVEVASVDAFQGREKDYIILSCVRANEQQAIGFLRDPRRLNVGLTR 801
            +        + + ++E+ SVD FQGREK+ ++L+ VR+N +  +GFL + RR+NV +TR
Sbjct: 550 SLS-------NKHPELEIKSVDGFQGREKEAVLLTFVRSNRKGEVGFLAEDRRINVAVTR 602

Query: 802 AKYGLVILGNPRSLARNTLWNHLLIHFREKG 832
           A+  + ++ +  ++  +     L+ +F E G
Sbjct: 603 ARRHVAVICDSHTVNNHAFLETLVDYFTEHG 633
```

emb|CAA20978.1| (AL031629) cDNA EST yk384f3.5 comes from this gene; cDNA EST
EMBL:T00167 comes from this gene; cDNA EST yk287a1.5
comes from this gene [Caenorhabditis elegans]
Length = 693

Score = 164 bits (410), Expect = 4e-39
Identities = 148/468 (31%), Positives = 226/468 (47%), Gaps = 63/468 (13%)

```
Query: 395 VPLPK--EFSIPNFAQ-LNSSQSNAVSHVL--QRPLSLIQGPPGTGKTVTSATIVYHLSK 449
           +P+P   +  I N  + LN SQ  AVS  +  QR L  IQGPPGTGKT   A IV+ L K
Sbjct: 222 IPMPSIHDRKIGNLPETLNPSQVAAVSAAMNTQRNLLCIQGPPGTGKTRVIAEIVHQLMK 281

Query: 450 IHKDRILVCAPSNVAVDHLAAKLRDLGLKVVRLTAKSREDVESSVSNLALHNLVGRGAKG 509
             K ++LVCAP++VAV +  A     G  +  +  +  E     ++N           K
Sbjct: 282 -KKKKVLVCAPTHVAVRN--AMDATTGRMIEEMPEEKAEQQLCLLAN----------TKD 328

Query: 510 ELKN------LLKLKDEVGELSASDT--KRF---VKLVRKTEAEILNKADVVCCTCVGAG 558
           E +N      L ++K ++  +S SD   K+      L+R +    +    +   + +G
Sbjct: 329 EFQNHKSASKLEEIKKQLSTMSTSDPLYKKMSYEYYLIRNSIFRSIFYPKLAVFSTLGTS 388

Query: 559 DKRLDTKFR----TVLIDESTQASEPECLIPIVK--GAKQVILVGDHQQLGPVILERKAA 612
             +   ++      +++DE+ Q +EP   +P++     K++ILVGD +QL  V+L  KA
Sbjct: 389 SIQKLPEYHWNADVMIVDEAAQCTEPATWVPVLTTPSCKKLILVGDQKQLPAVVLSDKAM 448

Query: 613 DAGLKQSLFERL---ISLGHVPIRLEVQYRMNPYLSEFPSNMFYEGSLQNGVTIEQRTV- 668
               K SL E+L    S  ++ I L  QYRMN  +  + + +FYE  L    T+   T+
Sbjct: 449 KGNFKLSLMEKLAEEFSSNNINILLNEQYRMNKKIMNWSNEVFYENQLTAHSTVSDITLR 508

Query: 669 ---PNSKFPW----PIRGIPMMFWANYGREEISANGTSFLNRIEAMNCERIITKLFRD-G 720
              PN         PI  I M    +   EE  ++  SF N  E       + +L  D G
Sbjct: 509 DICPNLPEDHVLNNPIMMINMENVKDRSHEEFESH--SFTNTDELNLVTEYVNRLVVDLG 566

Query: 721 VKPEQIGVITPYEGQRAYILQYMQMNGSLDKDLYIKVEVASVDAFQGREKDYIILSCVRA 780
           + P+ I VI+PY         Y Q+   L + +  +V+V +VDAFQG E+  +I   VR
Sbjct: 567 INPKAIAVISPY---------YAQIE-KLRRSIPFRVDVNTVDAFQGHERQVVIFCLVRD 616

Query: 781 NEQQAIGFLRDPRRLNVGLTRAKYGLVILGNPRSLARN----TLWNHL 824
           N++  IGFL + RRLNV +TRA+   V++GN R +  N     L+ HL
Sbjct: 617 NDEGQIGFLNETRRLNVAVTRARRQFVLIGNGRMMKGNKHLLKLYKHL 664
```

sp|P23249|MV10_MOUSE PROTEIN MOV-10 >gi|110821|pir||A39611 probable GTP-binding protein
- mouse >gi|53169|emb|CAA36803| (X52574) GTP binding
protein [Mus musculus]
Length = 1004

Score = 158 bits (395), Expect = 2e-37
Identities = 143/472 (30%), Positives = 215/472 (45%), Gaps = 78/472 (16%)

```
Query: 410 NSSQSNAVSHVLQ---RPLS-LIQGPPGTGKTVTSAT----IVYHLSKIHKDRILVCAPS 461
           N  Q  A+ H+++   RP   +I GPPGTGKTVT       +V HL K H   IL CAPS
```

```
Sbjct: 501 NPEQLQAMKHIVRGTTRPAPYIIFGPPGTGKTVTLVEAIKQVVKHLPKAH---ILACAPS 557

Query: 462 NVAVDHLAAKLR-DLGLKVVRLTAKSREDVESSVSNLALHNLVGRGAKGELKNLLKLKDE 520
           N   D L  +LR  L   + RL A SR+                R    ++K       +
Sbjct: 558 NSGADLLCQRLRVHLPSSIYRLLAPSRDI---------------RMVPEDIKTCCNWDAK 602

Query: 521 VGELSASDTKRFVKLVRKTEAEILNKADVVCCTCVGAGDKRLDTKFRTVLIDESTQASEP 580
            GE      K+ ++  R     ++  + +V      +    +D  F  + IDE+    EP
Sbjct: 603 KGEY-VYPAKKHLQQYRVLITTLITASRLV------SAQFPID-HFTHIFIDEAGHCMEP 654

Query: 581 ECLIPIV---------KGAKQVILVGDHQQLGPVILERKAADAGLKQSLFERLISL---- 627
           E L+ I               Q++L GD +QLGPV+    A   GL  SL ERL++
Sbjct: 655 ESLVAIAGLMDVKETGNPGGQLVLAGDPRQLGPVLRSPLALKHGLGYSLLERLLAYNSLY 714

Query: 628 -----GHVP---IRLEVQYRMNPYLSEFPSNMFYEGSLQNGVTIEQRTVPNSKFPWPIRG 679
                G+ P    +L    YR +P + + P+ ++Y+G LQ    +  R          P +G
Sbjct: 715 KKGPNGYDPQFITKLLRNYRSHPTILDIPNQLYYDGELQACADVVDRERFCRWEGLPQQG 774

Query: 680 IPMMFWANYGREEISANGTSFLNRIEAMNCERIITKLFRDGVK-------PEQIGVITPY 732
            P++F    G++E   N  SF N  EA      + +L     K       P  +GVI+PY
Sbjct: 775 FPIIFHGVMGKDEREGNSPSFFNPEEAATVTSYLKQLLAPSSKKGKARLSPRNVGVISPY 834

Query: 733 EGQRAYILQYMQMNGSLDKDLY----IK-VEVASVDAFQGREKDYIILSCVRANEQQA-- 785
              Q   I   +     LD++L     IK ++V SV+ FQG+E+  I++S VR+++
Sbjct: 835 RKQVEKIRYCIT---KLDRELRGLDDIKDLKVGSVEEFQGQERSVILISTVRSSQSFVQL 891

Query: 786 -----IGFLRDPRRLNVGLTRAKYGLVILGNPRSLARNTLWNHLLIHFREKG 832
                +GFL++P+R NV +TRAK  L+++GNP  L  +  W   L   +E G
Sbjct: 892 DLDFNLGFLKNPKRFNVAVTRAKALLIVVGNPLLLGHDPDWKTFLEFCKENG 943
```

sp|P51530|Y083_HUMAN HYPOTHETICAL PROTEIN KIAA0083 >gi|1531548|dbj|BAA07647| (D42046) The ha3631 gene product is related to S.cerevisiae protein encoded in chromosome VIII. [Homo sapiens]
Length = 1077

```
 Score =  154 bits (385), Expect = 3e-36
 Identities = 133/454 (29%), Positives = 212/454 (46%), Gaps = 64/454 (14%)

Query: 409  LNSSQSNAVSHVL-QRPLSLIQGPPGTGKTVTSATIVYHLSKIHKDRILVCAPSNVAVDH 467
            LN  Q  A+  VL  +  +LI G PGTGKT T  T+V  L       +L+ + ++ AVD+
Sbjct: 643  LNKPQRQAMKKVLLSKDYTLIVGMPGTGKTTTICTLVRILYACGFS-VLLTSYTHSAVDN 701

Query: 468  LAAKLRDLGLKVVRLTAKSREDVESSVSNLALHNLVGRGAKGELKNLLKLKDEVGELSAS 527
            +  KL    +  +RL                           G+++ +     +  E
Sbjct: 702  ILLKLAKFKIGFLRL-------------------------GQIQKVHPAIQQFTEQEIC 735

Query: 528  DTKRFVKLVRKTEAEILNKADVVCCTCVGAGDKRLDTK-FRTVLIDESTQASEPECLIPI 586
             +K    L    E  + N   +V  TC+G        K F   ++DE++Q S+P CL P+
Sbjct: 736  RSKSIKSLALLEE--LYNSQLIVATTCMGINHPIFSRKIFDFCIVDEASQISQPICLGPL 793

Query: 587  VKGAKQVILVGDHQQLGPVILERKAADAGLKQSLFERLISLGHVPIRLEVQYRMNPYLSE 646
               +++ +LVGDHQQL P++L R+A   G+ +SLF+RL       ++L VQYRMN  +
Sbjct: 794  FF-SRRFVLVGDHQQLPPLVLNREARALGMSESLFKRLEQNKSAVVQLTVQYRMNSKIMS 852

Query: 647  FPSNMFYEGSLQNGVTIEQRTVPNSKFPWPIRGIPMMFWANYGREE------ISANGTSF 700
              + + YEG L+ G       V N +    ++ + + F+A+Y             N   F
Sbjct: 853  LSNKLTYEGKLECGSDKVANAVINLRHFKDVK-LELEFYADYSDNPWLMGVFEPNNPVCF 911

Query: 701  LNRIEAMNCERI----------------ITKLF-RDGVKPEQIGVITPYEGQRAYILQYM 743
            LN  +    E++                +T +F + G  P  IG+I PY  Q   I
Sbjct: 912  LNTDKVPAPEQVEKGGVSNVTEAKLIVFLTSIFVKAGCSPSDIGIIAPYRQQLKII---- 967

Query: 744  QMNGSLDKDLYIKVEVASVDAFQGREKDYIILSCVRANEQQAIG-FLRDPRRLNVGLTRA 802
              N  L + + + VEV +VD +QGR+K  +++S VR+N+   +G  L+D RRLNV +TRA
Sbjct: 968  --NDLLARSIGM-VEVNTVDKYQGRDKSIVLVSFVRSNKDGTVGELLKDWRRLNVAITRA 1024

Query: 803  KYGLVILGNPRSLARNTLWNHLLIHFREKGCLVE 836
            K+ L++LG   SL        LL H   +  +++
Sbjct: 1025 KHKLILLGCVPSLNCYPPLEKLLNHLNSEKLIID 1058
```

emb|CAB38508.1| (AL035637) hypothetical helicase [Schizosaccharomyces pombe]
Length = 1398

```
 Score =  141 bits (352), Expect = 2e-32
 Identities = 130/452 (28%), Positives = 214/452 (46%), Gaps = 77/452 (17%)

Query: 404  PNFAQ-LNSSQSNAVS--HVLQRPLSLIQGPPGTGKTVTSATIVYHLSKIHKDRILVCAP 460
            P F + LN  Q  A+   H  +   SLI G PGTGKT T ++++  L    K +IL+ +
Sbjct: 927  PEFLKCLNEDQITALKKCHAAEH-YSLILGMPGTGKTTTISSLIRSLLA-KKKKILLTSF 984

Query: 461  SNVAVDHLAAKLRDLGLKVVRLTAKSREDVESSVSNLALHNLVGRGAKGELKNLLKLKDE 520
            +++AVD++  KL+     +VRL +  +           +H LV      E        D+
Sbjct: 985  THLAVDNILIKLKGCDSTIVRLGSPHK-----------IHPLVKEFCLTEGTTF----DD 1029

Query: 521  VGELSASDTKRFVKLVRKTEAEILNKADVVCCTCVGAGDKRLDT-KFRTVLIDESTQASE 579
            +  L     K F +              +V C+ +G      +  KF   +IDE++Q
Sbjct: 1030 LASL-----KHFYE-----------DPQIVACSSLGVYHSIFNKRKFDYCIIDEASQIPL 1073
```

```
Query: 580  PECLIPIVKGAKQVILVGDHQQLGPVILERKAADAGLKQSLFERLISLGH--VPIRLEVQ 637
            P CL P+ + A++ +LVGDH QL P++   + +  GL  SLF +L+S  H      L +Q
Sbjct: 1074 PICLGPL-QLAEKFVLVGDHYQLPPLVKNSRTSKDGLSLSLF-KLLSEKHPEAVTTLRLQ 1131

Query: 638  YRMNPYLSEFPSNMFYEGSLQNGVTIEQRTVPNSKFPWPIR----GIP----MMFWAN-- 687
            YRMN  ++   S + Y G+L  G     +T+   K   P      G+P     + W N
Sbjct: 1132 YRMNEDINSLSSELIYGGNLVCG----SKTISQKKLILPKAHLSDGLPDSSSSLHWVNKL 1187

Query: 688  ---------YGREEI----SANGTSFLNRIEAMNCERIITKLFRDGVKPEQIGVITPYEG 734
                     +  ++I    S       N  EA   E+ ++     GVK   IG+I+ Y+
Sbjct: 1188 INPSHSVIFFNTDDILGVESKTNNILENHTEAFLIEQAVSSFLERGVKQSSIGIISIYKS 1247

Query: 735  QRAYILQYMQMNGSLDKDLYIKVEVASVDAFQGREKDYIILSCVRANEQQAIG-FLRDPR 793
            Q   + + ++         + ++E+ +VD +QGR+KD I++S VR+N +  +G  LRD
Sbjct: 1248 QVELLSKNLKS--------FTEIEINTVDRYQGRDKDIILISFVRSNSKNLVGELLRDWH 1299

Query: 794  RLNVGLTRAKYGLVILGNPRSLARNTLWNHLL 825
            RLNV L+RAK   ++ G+  +L+ + + +HLL
Sbjct: 1300 RLNVALSRAKVKCIMFGSLSTLSSSNIVSHLL 1331
```

gb|AAD38528.1|AF144384_1 (AF144384) Dna2p [Schizosaccharomyces pombe]
Length = 1397

Score = 141 bits (352), Expect = 2e-32
Identities = 130/452 (28%), Positives = 214/452 (46%), Gaps = 77/452 (17%)

```
Query: 404  PNFAQ-LNSSQSNAVS--HVLQRPLSLIQGPPGTGKTVTSATIVYHLSKIHKDRILVCAP 460
            P F + LN  Q  A+   H  +   SLI G PGTGKT T ++++  L    K +IL+ +
Sbjct: 926  PEFLKCLNEDQITALKKCHAAEH-YSLILGMPGTGKTTTISSLIRSLLA-KKKKILLTSF 983

Query: 461  SNVAVDHLAAKLRDLGLKVVRLTAKSREDVESSVSNLALHNLVGRGAKGELKNLLKLKDE 520
            +++AVD++  KL+     +VRL +  +           +H LV      E        D+
Sbjct: 984  THLAVDNILIKLKGCDSTIVRLGSPHK-----------IHPLVKEFCLTEGTTF----DD 1028

Query: 521  VGELSASDTKRFVKLVRKTEAEILNKADVVCCTCVGAGDKRLDT-KFRTVLIDESTQASE 579
            +  L     K F +              +V C+ +G      +  KF   +IDE++Q
Sbjct: 1029 LASL-----KHFYE-----------DPQIVACSSLGVYHSIFNKRKFDYCIIDEASQIPL 1072

Query: 580  PECLIPIVKGAKQVILVGDHQQLGPVILERKAADAGLKQSLFERLISLGH--VPIRLEVQ 637
            P CL P+ + A++ +LVGDH QL P++   + +  GL  SLF +L+S  H      L +Q
Sbjct: 1073 PICLGPL-QLAEKFVLVGDHYQLPPLVKNSRTSKDGLSLSLF-KLLSEKHPEAVTTLRLQ 1130

Query: 638  YRMNPYLSEFPSNMFYEGSLQNGVTIEQRTVPNSKFPWPIR----GIP----MMFWAN-- 687
            YRMN  ++   S + Y G+L  G     +T+   K   P      G+P     + W N
Sbjct: 1131 YRMNEDINSLSSELIYGGNLVCG----SKTISQKKLILPKAHLSDGLPDSSSSLHWVNKL 1186

Query: 688  ---------YGREEI----SANGTSFLNRIEAMNCERIITKLFRDGVKPEQIGVITPYEG 734
                     +  ++I    S       N  EA   E+ ++     GVK   IG+I+ Y+
Sbjct: 1187 INPSHSVIFFNTDDILGVESKTNNILENHTEAFLIEQAVSSFLERGVKQSSIGIISIYKS 1246

Query: 735  QRAYILQYMQMNGSLDKDLYIKVEVASVDAFQGREKDYIILSCVRANEQQAIG-FLRDPR 793
            Q   + + ++         + ++E+ +VD +QGR+KD I++S VR+N +  +G  LRD
Sbjct: 1247 QVELLSKNLKS--------FTEIEINTVDRYQGRDKDIILISFVRSNSKNLVGELLRDWH 1298

Query: 794  RLNVGLTRAKYGLVILGNPRSLARNTLWNHLL 825
            RLNV L+RAK   ++ G+  +L+ + + +HLL
Sbjct: 1299 RLNVALSRAKVKCIMFGSLSTLSSSNIVSHLL 1330
```

gb|AAD48967.1|AF147263_9 (AF147263) contains similarity to nonsense-mediated mRNA decay
trans-acting factors [Arabidopsis thaliana]
Length = 660

Score = 139 bits (346), Expect = 1e-31
Identities = 92/279 (32%), Positives = 144/279 (50%), Gaps = 7/279 (2%)

```
Query: 544 LNKADVVCCTCVGAGDKRLDT--KFRTVLIDESTQASEPECLIPI-VKGAKQVILVGDHQ 600
           L  A ++ CT  GA +   +       +++DE+ Q  E E +  + ++G    +L+GD
Sbjct: 378 LQNAHIIFCTASGAAEMTAERAGSIDMLVVDEAAQLKECESVAALQLQGLHHAVLIGDEL 437

Query: 601 QLGPVILERKAADAGLKQSLFERLISLGHVPIRLEVQYRMNPYLSEFPSNMFYEGSLQNG 660
           QL  ++       A   +SLFERL SLGH    L VQYRM+P +S FP+  FY G + +
Sbjct: 438 QLPAMVQSEVCEKAKFVRSLFERLDSLGHKKHLLNVQYRMHPSISLFPNMEFYGGKISDA 497

Query: 661 VTIEQRTVPNSKFPWPIRGIPMMFWANYGREEISANGTSFLNRIEAMNCERIITKLFRDG 720
             +++ T         + G         G+EE   +G S  N +E      I+T L +
Sbjct: 498 EIVKESTYQKRFLQGNMFGSFSFINVGLGKEEFG-DGHSPKNMVEIAVVSEILTNLLKVS 556

Query: 721 VKPE---QIGVITPYEGQRAYILQYMQMNGSLDKDLYIKVEVASVDAFQGREKDYIILSC 777
            + +    +GVI+PY+ Q + I + +    +   D    + V SVD FQG E+D II+S
Sbjct: 557 SETKTKMSVGVISPYKAQVSAIQERIGDKYTSVSDNLFTLNVRSVDGFQGGEEDIIIIST 616

Query: 778 VRANEQQAIGFLRDPRRLNVGLTRAKYGLVILGNPRSLA 816
           VR+N    IGFL + +R NV LTRA++ L ++GN R+L+
Sbjct: 617 VRSNCNGNIGFLSNRQRANVALTRARHCLWVIGNERTLS 655
```

```
 Score = 48.4 bits (113), Expect = 2e-04
 Identities = 36/104 (34%), Positives = 52/104 (49%), Gaps = 18/104 (17%)

Query: 407 AQLNSSQSNAVSHVLQR-------PLSLIQGPPGTGKTVTSATIVYHLSKIHKDRILVCA 459
           A LNSSQ +A+   L+         + LI GPPGTG T T AT+++ L  +   + +VC
Sbjct: 78  ANLNSSQESAILACLETREIRDKTSVKLIWGPPGTGNTKTVATLLFALLSL-SCKTVVCT 136

Query: 460 PSNVAVDHLAAKLRDLGLKVVRLTAKSREDVESSVSNLALHNLV 503
           P+N AV  +A+          RL A  +E   +  S   L N+V
Sbjct: 137 PTNTAVVAVAS----------RLLALFKESSSTEHSTYGLGNIV 170


 emb|CAA90399.1| (Z50070) similar to DNA binding protein; cDNA EST EMBL:Z14942 comes
            from this gene; cDNA EST EMBL:D34442 comes from this
            gene; cDNA EST EMBL:D35999 comes from this gene; cDNA EST
            EMBL:D37505 comes from this gene; cDNA EST EMB...
            Length = 1105

 Score =  132 bits (328), Expect = 2e-29
 Identities = 119/433 (27%), Positives = 199/433 (45%), Gaps = 68/433 (15%)

Query: 407  AQLNSSQSNAVSHVLQ-RPLSLIQGPPGTGKTVTSATIVYHLSKIHKDRILVCAPSNVAV 465
            A+LN+ Q  AV H L      +++G PG+GKT   + ++  L   +K ++L+ A ++ AV
Sbjct: 684  AKLNNEQRKAVVHALATEDFMMVEGLPGSGKTTLISVLIQCLVATNK-KVLLAAFTHSAV 742

Query: 466  DHLAAKLRDLGLKVVRLTAKSREDVESSVSNLALHNLVGRGAKGELKNLLKLKDEVGELS 525
            D++  KL                        +A   ++  G+   +K+ +K      +L
Sbjct: 743  DNILTKL---------------------TKEVAAEKILRLGSSSSIKDDIKKMTLKAKLE 781

Query: 526  ASDTKRFVKLVRKTEAEILNKADVVCCTCVGAGDKRLDT--KFRTVLIDESTQASEPECL 583
               ++ +   VRK    ++    +V CTC     + L +   F  V++DE++   EP  L
Sbjct: 782  NETSEEYYAAVRK----VMKTTPIVACTCHHVPRELLFSYRHFDVVIVDEASMVLEP-LL 836

Query: 584  IPIVKGAKQVILVGDHQQLGPVILERKAADAGLKQSLFERL-ISLGHVPIRLEVQYRMNP 642
            +P++  + + +LVGD +QL P+++ RKA   G   S  E+L  S   V + L  QYRMN
Sbjct: 837  LPVLATSNKFVLVGDCKQLTPLVVSRKAKQEGAGISTMEKLQQSHPGVVVSLTSQYRMNR 896

Query: 643  YLSEFPSNMFYEGSLQNGVTIEQRTVPNSKFPWPIRGIPMMFWANYGREEISAN------ 696
             +S   S +FYE  L  G     R+  +    +    + M   +++ R+ +S +
Sbjct: 897  EISVLSSKLFYENRLICGNESVSRSSLDRTGDYI---VAMDDGSDHIRKALSGDIKDSCV 953

Query: 697  ------------------GTSFLNRIEAMNCERIITKLFRDGVKPEQIGVITPYEGQRAY 738
                               G    N  EA     +  +    GVKP +IGV++ Y  Q  +
Sbjct: 954  FLDTQSTINSKMQCEDGEGGGMCNDGEAKLISELCQQFVMSGVKPHEIGVMSAYRRQVDH 1013

Query: 739  ILQYMQMNGSLDKDLYIKVEVASVDAFQGREKDYIILSCVRANEQ-QAIGFLRDPRRLNV 797
            I       G L+ D   ++EV ++D++QGREK  II S    N   +    L+D RR+NV
Sbjct: 1014 I------RGILNSD---ELEVNTIDSYQGREKRVIIWSLTWTNNSTKKSELLKDERRVNV 1064

Query: 798  GLTRAKYGLVILG 810
             LTRA+  LV++G
Sbjct: 1065 ALTRARQKLVVVG 1077


 emb|CAB54253.1| (Z50070) F43G6.1b [Caenorhabditis elegans]
            Length = 1069

 Score =  132 bits (328), Expect = 2e-29
 Identities = 119/433 (27%), Positives = 199/433 (45%), Gaps = 68/433 (15%)

Query: 407  AQLNSSQSNAVSHVLQ-RPLSLIQGPPGTGKTVTSATIVYHLSKIHKDRILVCAPSNVAV 465
            A+LN+ Q  AV H L      +++G PG+GKT   + ++  L   +K ++L+ A ++ AV
Sbjct: 648  AKLNNEQRKAVVHALATEDFMMVEGLPGSGKTTLISVLIQCLVATNK-KVLLAAFTHSAV 706

Query: 466  DHLAAKLRDLGLKVVRLTAKSREDVESSVSNLALHNLVGRGAKGELKNLLKLKDEVGELS 525
            D++  KL                        +A   ++  G+   +K+ +K      +L
Sbjct: 707  DNILTKL---------------------TKEVAAEKILRLGSSSSIKDDIKKMTLKAKLE 745

Query: 526  ASDTKRFVKLVRKTEAEILNKADVVCCTCVGAGDKRLDT--KFRTVLIDESTQASEPECL 583
               ++ +   VRK    ++    +V CTC     + L +   F  V++DE++   EP  L
Sbjct: 746  NETSEEYYAAVRK----VMKTTPIVACTCHHVPRELLFSYRHFDVVIVDEASMVLEP-LL 800

Query: 584  IPIVKGAKQVILVGDHQQLGPVILERKAADAGLKQSLFERL-ISLGHVPIRLEVQYRMNP 642
            +P++  + + +LVGD +QL P+++ RKA   G   S  E+L  S   V + L  QYRMN
Sbjct: 801  LPVLATSNKFVLVGDCKQLTPLVVSRKAKQEGAGISTMEKLQQSHPGVVVSLTSQYRMNR 860

Query: 643  YLSEFPSNMFYEGSLQNGVTIEQRTVPNSKFPWPIRGIPMMFWANYGREEISAN------ 696
             +S   S +FYE  L  G     R+  +    +    + M   +++ R+ +S +
Sbjct: 861  EISVLSSKLFYENRLICGNESVSRSSLDRTGDYI---VAMDDGSDHIRKALSGDIKDSCV 917

Query: 697  ------------------GTSFLNRIEAMNCERIITKLFRDGVKPEQIGVITPYEGQRAY 738
                               G    N  EA     +  +    GVKP +IGV++ Y  Q  +
Sbjct: 918  FLDTQSTINSKMQCEDGEGGGMCNDGEAKLISELCQQFVMSGVKPHEIGVMSAYRRQVDH 977

Query: 739  ILQYMQMNGSLDKDLYIKVEVASVDAFQGREKDYIILSCVRANEQ-QAIGFLRDPRRLNV 797
            I       G L+ D   ++EV ++D++QGREK  II S    N   +    L+D RR+NV
Sbjct: 978  I------RGILNSD---ELEVNTIDSYQGREKRVIIWSLTWTNNSTKKSELLKDERRVNV 1028

Query: 798  GLTRAKYGLVILG 810
```

```
            LTRA+  LV++G
Sbjct: 1029 ALTRARQKLVVVG 1041


 gi|2105497 (AF003740) similar to mammalian DNA-binding proteins SMUBP-2
           [Caenorhabditis elegans]
           Length = 633

 Score =  130 bits (324), Expect = 4e-29
 Identities = 126/470 (26%), Positives = 222/470 (46%), Gaps = 56/470 (11%)

Query: 392 SFDVPLPKEFSIPNFAQLNSSQSNAVSHVLQ--RPLSLIQGPPGTGKTVTSATIVYHLSK 449
           SFD  + KE  IP F+  N  Q  A+   L   R L  IQGPPGTGKT T   ++  L +
Sbjct: 61  SFD-SMEKEAGIP-FS--NEKQRTAIKMALNENRKLVCIQGPPGTGKTFTLTLLLCRLIQ 116

Query: 450 IHKDRILVCAPSNVAVDHLAAK----LRDLGLKVVRLTAKSREDVESSV--SNLALHNLV 503
             K +++V AP+  A+ ++       L+ +G+KV         +    +  S+ AL
Sbjct: 117 -QKKQVVVLAPTREALANIRMMTKKTLKRMGIKVHEHALMDTNEYRDVINKSDRALM--- 172

Query: 504 GRGAKGELKNLLKLKD--EVGELSASDTKRFV--KLVRKTEAEILNKADVVCCTCVGAGD 559
              A  E+++L K  D  E+ E    + ++ +  ++  +  AE++    V   T +GA
Sbjct: 173 ---AAEEVRDLRKAFDNGEITENVLDEMRQSIINRVRNEVGAEVIGNVRVAFAT-IGASF 228

Query: 560 KRLDTKFRTV-----LIDESTQASEPECLIPIVKGAKQVILVGDHQQLGPVILERKAADA 614
                K +       +IDE+ Q  E +   P V   K++++ GD +QL  ++   +A
Sbjct: 229 VDFVMKHKKFDPCLCIIDEAAQVMEAQTW-PAVYKMKRIVMAGDPKQLPALVFTDEAKAF 287

Query: 615 GLKQSLFERLISLGH--VPIRLEVQYRMNPYLSEFPSNMFYEGSLQNGVTIEQRTVPNSK 672
           GL+ S+ +R++   +    I LE QYR N  ++ + +  FY   L+  V   + ++
Sbjct: 288 GLQNSVMDRILEKKNNFSWIMLENQYRSNAKIATWSNTCFYHNQLKTDVKCHEYSLHTIL 347

Query: 673 FPWP--IRGI--PMMFWANYGREEI-------------------SANGTSFLNRIEAMNC 709
            P P   R +  P++        ++                   +  G S+ N  EA
Sbjct: 348 NPQPKKFRNLFDPLVLIDTSLERDVEKRLETYEHAVFDTNSINKTKQGFSYANLAEAKIA 407

Query: 710 ERIITKLFRDGVKPEQIGVITPYEGQRAYILQYMQMNGS-LDKDLYIKVEVASVDAFQGR 768
                +L + GV+P  I +ITPY+GQ + + + M+  G+      +++  + +VD+ QG+
Sbjct: 408 IGHYQRLLKYGVQPSDIAIITPYKGQTSLVTKLMEEFGAETGYTDFVQTTIGTVDSVQGK 467

Query: 769 EKDYIILSCVRANEQQAIGFLRDPRRLNVGLTRAKYGLVILGNPRSLARN 818
           E + +I + VR+N ++ +GF+ + RRLNV +TRAK   + +GN   LA +
Sbjct: 468 EYEVVIFTMVRSNPRKTMGFVSELRRLNVVITRAKRHFMFIGNGYLLAES 517


 sp|P38859|DNA2_YEAST DNA REPLICATION HELICASE DNA2 >gi|626832|pir||S48904 probable purine
            nucleotide-binding protein YHR164c - yeast (Saccharomyces
            cerevisiae) >gi|458906 (U00027) Dna2p: DNA replication
            helicase [Saccharomyces cerevisiae]
            Length = 1522

 Score =  128 bits (318), Expect = 2e-28
 Identities = 127/456 (27%), Positives = 204/456 (43%), Gaps = 75/456 (16%)

Query: 390  DISFDVPLPKEFSIPNFAQLNSSQSNAVSHVLQ-RPLSLIQGPPGTGKTVTSATIVYHLS 448
            D + D  +P  + +     LN +Q  A+  V++    +LI G PGTGKT   A I+  L
Sbjct: 1035 DANDDPVIP--YKLSKDTTLNLNQKEAIDKVMRAEDYALILGMPGTGKTTVIAEIIKILV 1092

Query: 449  KIHKDRILVCAPSNVAVDHLAAKLRDLGLKVVRLTAKSR--EDVESSVSNLALHNLVGRG 506
               K R+L+ + ++ AVD++  KLR+  + ++RL  K +   D +  V N A
Sbjct: 1093 SEGK-RVLLTSYTHSAVDNILIKLRNTNISIMRLGMKHKVHPDTQKYVPNYA-------- 1143

Query: 507  AKGELKNLLKLKDEVGELSASDTKRFVKLVRKTEAEILNKADVVCCTCVGAGDKRL---D 563
                                      K +   + K     +N   VV  TC+G  D      +
Sbjct: 1144 ---------------------SVKSYNDYLSK-----INSTSVVATTCLGINDILFTLNE 1177

Query: 564  TKFRTVLIDESTQASEPECLIPIVKGAKQVILVGDHQQLGPVILERKAADAGLKQSLFER 623
              F  V++DE++Q S P  L P+  G  + I+VGDH QL P++    A   GL++SLF+
Sbjct: 1178 KDFDYVILDEASQISMPVALGPLRYG-NRFIMVGDHYQLPPLVKNDAARLGGLEESLFKT 1236

Query: 624  LISLGHVPI-RLEVQYRMNPYLSEFPSNMFYEGSLQ---NGVTIEQRTVP---------- 669
                    +  L +QYRM   +    + + Y+  L+   N V  +   +P
Sbjct: 1237 FCEKHPESVAELTLQYRMCGDIVTLSNFLIYDNKLKCGNNEVFAQSLELPMPEALSRYRN 1296

Query: 670  ---NSKFPW------PIRGIPMMFWANYGREEISANGTSFLNRIEAMNCERIITKLFRDG 720
               NSK  W      P R +  + + N       +   +  N  EA    + +  +   G
Sbjct: 1297 ESANSK-QWLEDILEPTRKVVFLNYDNCPDIIEQSEKDNITNHGEAELTLQCVEGMLLSG 1355

Query: 721  VKPEQIGVITPYEGQRAYILQYMQMNGSLDKDLYIKVEVASVDAFQGREKDYIILSCVRA 780
            V  E IGV+T Y  Q   + +        +K++Y  +E+ + D FQGR+K  II+S VR
Sbjct: 1356 VPCEDIGVMTLYRAQLRLLKKIF------NKNVYDGLEILTADQFQGRDKKCIIISMVRR 1409

Query: 781  NEQQAIG-FLRDPRRLNVGLTRAKYGLVILGNPRSL 815
            N Q   G  L++ RR+NV +TRAK  L+I+G+  ++
Sbjct: 1410 NSQLNGGALLKELRRVNVAMTRAKSKLIIIGSKSTI 1445


 gi|3063473 (AC003981) F22O13.35 [Arabidopsis thaliana]
           Length = 814
```

```
 Score =  124 bits (308), Expect = 3e-27
 Identities = 128/462 (27%), Positives = 214/462 (45%), Gaps = 91/462 (19%)

Query: 409 LNSSQSNAVSHVLQ-RPLSLIQGPPGTGKTVTSATIVYHLSKIHKDRILVCAPSNVAVDH 467
           LN+ Q  A+  +L  +  +LI G PGTGKT T    V  L  I    IL+ + +N AVD+
Sbjct: 424 LNNDQRQAILKILTAKDYALILGMPGTGKTSTMVHAVKALL-IRGSSILLASYTNSAVDN 482

Query: 468 LAAKLRDLGLKVVRLTAKSREDVESSVSNLALHNLVGRGAKGELKNLLKLKDEVGELSAS 527
           L  KL+  G++ +R+    R++        A+H  V R +     N+  ++D
Sbjct: 483 LLIKLKAQGIEFLRI---GRDE--------AVHEEV-RESCFSAMNMCSVED-------- 522

Query: 528 DTKRFVKLVRKTEAEILNKADVVCCTCVGAGDKRL-DTKFRTVLIDESTQASEPECLIPI 586
                   ++K     L++  VV  TC+G     L + +F   +IDE+ Q + P  + P+
Sbjct: 523 --------IKKK----LDQVKVVASTCLGINSPLLVNRRFDVCIIDEAGQIALPVSIGPL 570

Query: 587 VKGAKQVILVGDHQQLGPVILERKAADAGLKQSLFERLISLGHVPIR-LEVQYRMNPYLS 645
           +  A   +LVGDH QL P++   +A + G+  SLF RL       I  L+ QYRM   +
Sbjct: 571 L-FASTFVLVGDHYQLPPLVQSTEARENGMGISLFRRLSEAHPQAISVLQNQYRMCRGIM 629

Query: 646 EFPSNMFYE-----GSLQNGVTIEQRTVPNSKFPWPIRGIPMMFWANYGREEISANGTSF 700
           E  + + Y      GS +        +  +S  PW
Sbjct: 630 ELSNALIYGDRLCCGSAEVADATLVLSTSSSTSPW------------------------- 664

Query: 701 LNRIEAMNCERIITKLFRDGVKPEQIGVITPYEGQRAYILQYMQMNGSLDKDLYIKVEVA 760
                    ++I+ +L  +GV  + IG+ITPY  Q + I   +             VE+
Sbjct: 665 --------LKKIVEELVNNGVDSKDIGIITPYNSQASLIQHAIPTT---------PVEIH 707

Query: 761 SVDAFQGREKDYIILSCVRANEQ---QAIGFLRDPRRLNVGLTRAKYGLVILGNPRSLAR 817
           ++D +QGR+KD I++S VR+ E+    A   L D  R+NV LTRAK  L+++G+ R+L+R
Sbjct: 708 TIDKYQGRDKDCILVSFVRSREKPRSSASSLLGDWHRINVALTRAK--LIMVGSQRTLSR 765

Query: 818 NTLWNHLLIHFREKGCLVEGTLDNLQLCTVQLVRPQPRKTER 859
             L   +L+  ++ G +V     +L++ +V ++     +T++
Sbjct: 766 VPLL--MLLLNKKSGDIVMMEKFHLKVSSVNVMTVDFERTKK 805


 emb|CAA20777| (AL031540) hypothetical ATP binding protein [Schizosaccharomyces
           pombe]
           Length = 1015

 Score =  112 bits (278), Expect = 1e-23
 Identities = 87/292 (29%), Positives = 145/292 (48%), Gaps = 21/292 (7%)

Query: 538 KTEAEILNKADVVCCTCVGAGD-----KRLDTKFRTVLIDESTQASEPECLIPIVKGAKQ 592
           K  A +L  A+V+  T  G        +R++ K     I+E+    E   +  +    +Q
Sbjct: 674 KRAAILLRGANVIGMTTTGLNKYRDILERINPKI--CFIEEAADVLEGPIIPAVFPSLEQ 731

Query: 593 VILVGDHQQLGP--VILERKAADAGLKQSLFERLISLGHVPIRLEVQYRMNPYLSEFPSN 650
           ++L+GDH+QL P       +     L  S+FERL+       RL +Q RM+P +    S+
Sbjct: 732 LVLIGDHKQLRPGCSTYALRQDPFNLSISMFERLVENDMEYTRLTMQRRMHPQIRRLVSS 791

Query: 651 MFYEGSLQNGVTIEQRTVPNSKFPWPIRGIPMMFWANYGREEISANGTSFLNRIEAMNCE 710
           + YE      +T    ++P         G    F+  + R E +    S +N  EA
Sbjct: 792 V-YEDLSDYEITKYWPSIPGM-------GEIRRFFLTHSRIEDNDGFASKINLFEAQMLV 843

Query: 711 RIITKLFRDGVKPEQIGVITPYEGQRAYILQYMQMNGSLDKDLYIKVEVASVDAFQGREK 770
           +    L  +GV+P++I  +T Y  Q+  I + +  + SL+++ +  ++VA+VD +QG E
Sbjct: 844 QFAVYLINNGVEPQKITCLTFYAAQKDLIERLL--SESLNREKHF-IKVATVDGYQGEEN 900

Query: 771 DYIILSCVRANEQQAIGFLRDPRRLNVGLTRAKYGLVILGNPRSLAR-NTLW 821
           D ++LS VR N++  +GFL  P R+ V L+RA+ GL I GN + +A  N LW
Sbjct: 901 DVVLLSLVRNNDRTEVGFLSSPHRVCVSLSRARRGLFIFGNAQLVAESNPLW 952


 Score = 53.5 bits (126), Expect = 7e-06
 Identities = 35/90 (38%), Positives = 52/90 (56%), Gaps = 3/90 (3%)

Query: 409 LNSSQSNAVSHVLQRPLSLIQGPPGTGKT-VTSATIVYHLSKIHKD--RILVCAPSNVAV 465
           L+SSQ  A   +L + LS+IQGPPGTGK+ VT   I   L   H     ILV   +N AV
Sbjct: 372 LDSSQLKAYQSMLTKRLSIIQGPPGTGKSFVTLKAIETLLENTHSHVLPILVACQTNHAV 431

Query: 466 DHLAAKLRDLGLKVVRLTAKSREDVESSVS 495
           D +  +L   G  V+RL +++++   ++V+
Sbjct: 432 DQILIRLLHQGASVMRLGSRTKDPEIAAVT 461


 emb|CAA93884| (Z70038) cDNA EST EMBL:D32579 comes from this gene; cDNA EST
            EMBL:D35254 comes from this gene; cDNA EST yk224b3.5
            comes from this gene; cDNA EST yk357f10.5 comes from this
            gene [Caenorhabditis elegans]
            Length = 2219

 Score =  109 bits (271), Expect = 7e-23
 Identities = 84/317 (26%), Positives = 156/317 (48%), Gaps = 23/317 (7%)

Query: 530  KRFVKLVRKTEAEILNKADVVCCTCVGAGDKRLDTKF---RTVLIDESTQASEPECLIPI 586
            + F     + +AEI+    ++  T  G    R   +    R ++++E+ +  E   +  +
Sbjct: 1191 ENFKNAQNRVDAEIMRMTMIIGATTTGCSRLRPTLEKVGPRILIVEEAAEVLEAHIISAM 1250
```

```
Query: 587  VKGAKQVILVGDHQQLGPVILERKAADA-GLKQSLFERLISLGHVPIRLEVQYRMNPYLS 645
            +   + V+++GDH+QL P     +   A GL+ S+FERL+  G    +L  Q+RMN  +S
Sbjct: 1251 ISTVEHVVMIGDHKQLRPNPAVHELGVAYGLRISMFERLVERGLPFSQLRQQHRMNLTIS 1310

Query: 646  EFPSNMFYEGSLQNGVTIEQRTVPNSKFPWPIRGIP--MMFWANYGREEISANGTSFLNR 703
            +    + +  ++ +   +         +P  ++G+   + FW++   EE S +  S+LN+
Sbjct: 1311 DKIVKLSFYDNVTDAENV-------GLYP-DVQGMATNLFFWSHTSMEE-SPDEVSWLNK 1361

Query: 704  IEAMNCERIITKLFRDGVKPEQIGVITPYEGQRAYIL-QYMQMNGSLDKDLYIKVEVASV 762
             E      ++  L +       I V+  Y  Q+  +  +Y  + GS      I VE  +V
Sbjct: 1362 HEISMTVALVKHLLKQNYTTNDIVVLATYSAQKNLMYREYANVFGSTPDSNVIPVE--TV 1419

Query: 763  DAFQGREKDYIILSCVRAN----EQQAIGFLRDPRRLNVGLTRAKYGLVILGNPRSLARN 818
            D+FQG+E+  +I+S VR++    E   IGFL    R+ V LTRA++G+ I+GN   +  N
Sbjct: 1420 DSFQGKERKIVIVSLVRSHRGGRENTGIGFLAVANRICVALTRAQHGMYIIGNGAYIMNN 1479

Query: 819  T-LWNHLLIHFREKGCL 834
            + LWN ++ + R    +
Sbjct: 1480 SELWNKIVNNLRRSNLI 1496


 Score = 37.1 bits (84), Expect = 0.62
 Identities = 24/69 (34%), Positives = 36/69 (51%), Gaps = 5/69 (7%)

Query: 409 LNSSQSNAVSHVLQRPLSLIQGPPGTGKTVTSATIVYHLSKIHK-----DRILVCAPSNV 463
           ++ SQ  A  +  +  LSLIQGPPGTGKT     IV  + +        + ILV   +N
Sbjct: 809 MDESQRLAFCNTFKYELSLIQGPPGTGKTHIGVQIVKTILQNRSYWKITEPILVVCFTNS 868

Query: 464 AVDHLAAKL 472
            +D+L  ++
Sbjct: 869 GLDNLLERI 877


 sp|Q09449|YQ12_CAEEL HYPOTHETICAL 175.7 KD PROTEIN C05C10.2 IN CHROMOSOME II
            >gi|3874068|emb|CAA88201.1| (Z48178) similarity with
            yeast NAM7 protein (Swiss Prot accession number P30771)
            and may possibly interact with DNA or RNA; cDNA EST
            EMBL:D34258 comes from this gene; cDNA EST EMBL:D37677
            comes from this gene; cDNA EST EMBL:D...
            Length = 1551

 Score =  109 bits (269), Expect = 1e-22
 Identities = 143/518 (27%), Positives = 224/518 (42%), Gaps = 77/518 (14%)

Query: 376  SGYLYYKILGHQVVDISFDVPLPKEFSIPNFAQLNSSQSNAVSHVLQ-RPLSLIQGPPGT 434
            SG   +K  G +  D       P   S P    LN  Q   V  +L   PL +   P G
Sbjct: 969  SGISEHKTSGKKSEDTPTQFYFP---STPEPLALNKYQCEYVQMLLDGNPLIIGSSPFGC 1025

Query: 435  GKTVTSATIVYHLSKIHKDR--ILVCAPSNVAVDHLAAKLR-------DL-GLKVVR--- 481
            GK++T  T    L K+ K+R  +L+   +  +V+ +    R       DL  LK VR
Sbjct: 1026 GKSMTIITAALELYKLKKNRKQLLITQSNYASVNLIDIAQRVCLSGDDDLKDLKFVRFVS 1085

Query: 482  ------LTAKSREDVE-------------------SSVSNLALHN---LVGRGAKGELKN 513
                  L +  R D +                   ++++ L  H+   +V    K +L N
Sbjct: 1086 EKNWNELPSNCRTDSDMPYLMNKLFKDWAMGRIDLTNLTCLKTHHYVQMVSHIIKNDLVN 1145

Query: 514  LLKLKDEVGELS---ASDTKRFVKLVRKTEAE-ILNKADVVCCTCVGA-GDKRLDTKFRT 568
             +   D + ++    ++D  R        EA  ++ K D+V  T   A G   +
Sbjct: 1146 PMLFGDHIAQIYDKLSADFSRAPHAQTLVEAFFMIYKPDLVMVTADSAKGLLNILRDVCA 1205

Query: 569  VLIDESTQASEPECLIPIVKGAKQVI--LVGDHQQLGPVI---LERKAADAGLKQSLFER 623
            V IDE++Q +E   L+ ++K        L+GD  QL P     LE K  D G+  ++ ER
Sbjct: 1206 VQIDEASQLAECT-LLGLLKSFNNASFGLIGDIHQLPPYCEEGLEGKLKDFGIGNTM-ER 1263

Query: 624  LISLGHVPI-RLEVQYRMNPYLSEFPSNMFYEGSLQNGVTIEQRTVPNSKFP--WPIRGI 680
             I     P+  L   YR +P  +E  S +FY+G+L +GV+   R+   +K    WP
Sbjct: 1264 AIKEKMFPVCTLRNVYRCHPKTTELLSELFYDGALVSGVSELARSDFMTKRDDFWPNPKF 1323

Query: 681  PMMFWANYGREEISANGTSFLNRIEAMNCERIITKLFRDG---VKPEQIGVITPYEGQRA 737
            PMMF  N G    +  GTS  N  E      I+  L  D    V P  IGVI+ Y  Q +
Sbjct: 1324 PMMFVNNTGAS--TKMGTSTSNSSEKSIVGEIVQNLINDPRNPVNPSDIGVISFYSAQTS 1381

Query: 738  YILQYMQMNGSLDKDLYIKVEVASVDAFQGREKDYIILSCVRANEQQAIGFLRDPRRLNV 797
             + ++++ +G         V+  +VDAFQG EK+ II+    +  ++   F++   RLNV
Sbjct: 1382 ILTEHLRGSG---------VKCGTVDAFQGSEKEIIIMC---STNERISDFMQLSNRLNV 1429

Query: 798  GLTRAKYGLVILGNPRSLARNTLWNHLLIHFREKGCLV 835
             ++RAK   +I+G+   L R   W+ ++    + G LV
Sbjct: 1430 AMSRAKQVTIIIGHLDGLRRANYWSTIVNKIEQNGNLV 1467


 sp|P42694|Y054_HUMAN HYPOTHETICAL PROTEIN KIAA0054 >gi|473951|dbj|BAA06147| (D29677)
            KIAA0054 [Homo sapiens]
            Length = 1942

 Score =  103 bits (255), Expect = 5e-21
 Identities = 131/549 (23%), Positives = 215/549 (38%), Gaps = 115/549 (20%)

Query: 368  FAIDKKSISGYLYYKILGHQVVDISFDVPLP----KEFSIPNFAQLNSSQSNAVSHV--- 420
```

```
            +A+D+   +G L+         DIS    +P    +++      +LN+ Q  AV  +
Sbjct: 605  YALDRIKDNGVLF--------PDISMTPTIPWSPNRQWDEQLDPRLNAKQKEAVLAITTP 656

Query: 421  --LQRPLSLIQGPPGTGKTVTSATIVYHLSKIHKDRILVCAPSNVAVD-------HLAAK 471
              +Q P  LI GP GTGKT T A  V H+ +  + RIL+C  SN A D       H   +
Sbjct: 657  LAIQLPPVLIIGPYGTGKTFTLAQAVKHILQQQETRILICTHSNSAADLYIKDYLHPYVE 716

Query: 472  LRDLGLKVVRLTAKSR--EDVESSVSNLALHNLVGRGAKGELKNLLKLKDEVGELSASDT 529
              +   + +R+  ++R  + V   V    L +      +      +  K+++
Sbjct: 717  AGNPQARPLRVYFRNRWVKTVHPVVHQYCLISSAHSTFQ------MPQKEDI-------- 762

Query: 530  KRFVKLVRKTEAEILNKADVVCCTCVGAGDKRLDTKFRTVLIDESTQASEPECLIPIVKG 589
                 L  +     LN +  +C   +  G       F  +L+DE+ QA E E ++P+
Sbjct: 763  -----LKHRVVVVTLNTSQYLCQLDLEPGF------FTHILLDEAAQAMECETIMPLALA 811

Query: 590  AK--QVILVGDHQQLGPVILERKAADAGLKQSLFERLISLGHVP------IRLEVQYRMN 641
             +  +++L GDH QL P +    A +  L  SL +RL    H P      I L   YR +
Sbjct: 812  TQNTRIVLAGDHMQLSPFVYSEFARERNLHVSLLDRLYE--HYPAEFPCRILLCENYRSH 869

Query: 642  PYLSEFPSNMFYEGSLQNGVTIEQRTVPNSKFPWPIRGIPMMFWANYGREEISANGTSFL 701
              +  + S +FYEG L          + + K P      P+ F+   G +    N T+F
Sbjct: 870  EAIINYTSELFYEGKL----------MASGKQPAHKDFYPLTFFTARGEDVQEKNSTAFY 919

Query: 702  NRIEAMNCERIITKLFR------DGVKPEQIGVITPYEGQRAYILQYMQMNGSLDKDLYI 755
            N  E       + +L R        +    IGV+TPY        Q  ++   L K
Sbjct: 920  NNAEVFEVVERVEELRRKWPVAWGKLDDGSIGVVTPYAD------QVFRIRAELRKKRLS 973

Query: 756  KVEVASVDAFQGREKDYIILSCVR---------------------ANEQQAIGFLRDPRR 794
             V V  V   QG++   + LS VR                     + E    GFL + +
Sbjct: 974  DVNVERVLNVQGKQFRVLFLSTVRTRHTCKHKQTPIKKKEQLLEDSTEDLDYGFLSNYKL 1033

Query: 795  LNVGLTRAKYGLVILGNPRSLAR----NTLWNHLLIHFREKGCL-------VEGTLDNLQ 843
            LN  +TRA+  + ++G+P +L         W   +    E   L       ++  L+ L+
Sbjct: 1034 LNTAITRAQSLVAVVGDPIALCSIGRCRKFWERFIALCHENSSLHGITFEQIKAQLEALE 1093

Query: 844  LCTVQLVRP 852
            L    ++ P
Sbjct: 1094 LKKTYVLNP 1102
```

dbj|BAA25486| (AB011132) KIAA0560 protein [Homo sapiens]
Length = 1421

Score = 101 bits (249), Expect = 3e-20
Identities = 88/318 (27%), Positives = 154/318 (47%), Gaps = 24/318 (7%)

```
Query: 507  AKGELKNLLKLKDEVGELSASDTKRFVKLVRKTEAEILNKADVVCCTCVGAGDKRLDT-- 564
            A+G  +++ K+  ++ E  AS+  R    + +++  ++ +A ++  TC  A  KR D
Sbjct: 930  AEGCFRHIKKIFTQLEEFRASELLR--SGLDRSKYLLVKEAKIIAMTCTHAALKRHDLVK 987

Query: 565  ---KFRTVLIDESTQASEPECLIPIV--------KGAKQVILVGDHQQLGPVILERKAAD 613
               K+  +L++E+ Q  E E  IP++           K+ I++GDH QL PVI
Sbjct: 988  LGFKYDNILMEEAAQILEIETFIPLLLQNPQDGFSRLKRWIMIGDHHQLPPVIKNMAFQK 1047

Query: 614  -AGLKQSLFERLISLGHVPIRLEVQYRMNPYLSEFPSNMFYE--GSLQNGVTIEQRTVPN 670
             + ++QSLF R + +G   + L+ Q R    L     N  Y+  G+L +   + + +  N
Sbjct: 1048 YSNMEQSLFTRFVRVGVPTVDLDAQGRARASLCNL-YNWRYKNLGNLPHVQLLPEFSTAN 1106

Query: 671  SKFPWPIRGIPMMFWANYGREEISANGTSFLNRIEAMNCERIITKLFRDGVKPEQIGVIT 730
            +   +  + I +  +   G  E   N   + N  EA     +   +   G   ++I ++T
Sbjct: 1107 AGLLYDFQLINVEDFQGVGESE--PNPYFYQNLGEAEYVVALFMYMCLLGYPADKISILT 1164

Query: 731  PYEGQRAYILQYMQMNGSLDKDLYIKVEVASVDAFQGREKDYIILSCVRANEQQAIGFLR 790
             Y GQ+  I   +      +  +    +V +VD FQG++ DYI+LS VR    +A+G LR
Sbjct: 1165 TYNGQKHLIRDIINRRCGNNPLIGRPNKVTTVDRFQGQQNDYILLSLVRT---RAVGHLR 1221

Query: 791  DPRRLNVGLTRAKYGLVI 808
            D RRL V ++RA+ GL I
Sbjct: 1222 DVRRLVVAMSRARLGLYI 1239
```

Score = 41.4 bits (95), Expect = 0.032
Identities = 25/83 (30%), Positives = 41/83 (49%), Gaps = 1/83 (1%)

```
Query: 396 PLPKEFSIPNFAQLNSSQSNAVSHVLQRPLSLIQGPPGTGKTVTSATIVYHL-SKIHKDR 454
           P P      N  Q   +Q  A+   +Q  L+++ GPPGTGKT  +  I+ ++     + R
Sbjct: 725 PYPYNQPKRNTIQFTHTQIEAIRAGMQPGLTMVVGPPGTGKTDVAVQIISNIYHNFPEQR 784

Query: 455 ILVCAPSNVAVDHLAAKLRDLGL 477
            L+   SN A++ L  K+  L +
Sbjct: 785 TLIVTHSNQALNQLFEKIMALDI 807
```

emb|CAB10335.1| (Z97339) SEN1 like protein [Arabidopsis thaliana]
Length = 555

Score = 99.5 bits (244), Expect = 1e-19
Identities = 105/379 (27%), Positives = 169/379 (43%), Gaps = 70/379 (18%)

```
Query: 444 VYHLSKIHKDRILVCAPSNVAVDHLAAKLRDLGL----------KVVRLTAKSREDVESS 493
           V + S+ ++ R+LVCAPSN A+D +  +L   GL          K+VR+  K+      S
Sbjct: 99  VVNASRKYRLRVLVCAPSNSALDEIVLRLLSSGLRDENAQTYTPKIVRIGLKAHH----S 154

Query: 494 VSNLALHNLVGRGAKGELKNLLKLKDEVGELSAS-DTKRFVKLVRKTEAEILNKADVVCC 552
           V++++L +LV +     +      K + G      D+ R   L         N    V
Sbjct: 155 VASVSLDHLVAQKRGSAID-----KPKQGTTGTDIDSIRTAILEEAAIVWGFNFRINVFA 209

Query: 553 TCVGAGDKRL---DTKFRTVLIDESTQASEPECLIPIVKGAKQVILVGDHQQLGPVILER 609
           T   +G   L   +  F  V+IDE+ QA EP  LIP+    KQV LVGD +QL   ++
Sbjct: 210 TLSFSGSALLAKSNRGFDVVIIDEAAQAVEPATLIPLATRCKQVFLVGDPKQLPATVIST 269

Query: 610 KAADAGLKQSLFERLISLGHVPIRLEVQYRMNP-------------------------- 642
            A D+G   S+FERL   G+    L+ QYRM+P
Sbjct: 270 VAQDSGYGTSMFERLQKAGYPVKMLKTQYRMHPELIPISQTKPKIVVSLILTPMIASYFT 329

Query: 643 ----YLSE-----FPSNMFYEGSLQNGVTIEQRTVPNSKFPWPIRGIPMMFWANYGREEI 693
               +L+E     FPS  FYEG+L++G  IE +T       +   G    F  + G+E
Sbjct: 330 LQTIFLTENQIRSFPSKQFYEGALEDGSDIEAQTT-RDWHKYRCFGPFCFFDIHEGKESQ 388

Query: 694 SANGTSF---LNRIE--AMNCERIITKLFRDGVKPEQIGVITPYEGQ-RAYILQYMQMNG 747
               T     L+ +E   +   R++T ++ +     Q+ +I+PY  Q + +  ++ +M G
Sbjct: 389 HPGATGSRVNLDEVEFVLLIYHRLVT-MYPELKSSSQLAIISPYNYQVKTFKDRFKEMFG 447

Query: 748 SLDKDLYIKVEVASVDAFQ 766
           +  + +   V++ +VD FQ
Sbjct: 448 TEAEKV---VDINTVDGFQ 463
```

gi|3785995 (AC005499) unknown protein [Arabidopsis thaliana]
Length = 1444

Score = 98.7 bits (242), Expect = 2e-19
Identities = 93/323 (28%), Positives = 146/323 (44%), Gaps = 30/323 (9%)

```
Query: 505  RGAKGELKNLLKLKDEVGELSASDTKRFVKLVRKTEAEILNKADVVCCTCVGAGDKRLDT 564
            R AKG   +L  +  E+ E  A +  +      +    +  +A +V  TC  A  KR D
Sbjct: 1053 RAAKGCFSHLKTVFQELEECRAFELLK--STADRANYLMTKQAKIVAMTCTHAALKRRDF 1110

Query: 565  -----KFRTVLIDESTQASEPECLIPIV--------KGAKQVILVGDHQQLGPVILERKA 611
                 K+  +L++ES Q  E E  IP++           K+ IL+GDH QL PV+
Sbjct: 1111 LQLGFKYDNLLMEESAQILEIETFIPMLLQRQEDGHARLKRCILIGDHHQLPPVVKNMAF 1170

Query: 612  AD-AGLKQSLFERLISLGHVPIRLEVQYRMNPYLSEFPSNMFYEG----SLQNGVTIEQR 666
               + + QSLF R + LG   I L  Q R  P L++   N  Y      S+     I QR
Sbjct: 1171 QKYSHMDQSLFTRFVRLGIPYIELNAQGRARPSLAKL-YNWRYRDLGDLSIVKEAPIFQR 1229

Query: 667  TVPNSKFPWPIRGIPMMFWANY-GREEISANGTSFLNRIEAMNCERIITKLFRDGVKPEQ 725
                  + + +  +P     +Y GR E + +   + N+ EA     +   +   G    +
Sbjct: 1230 ANAGFSYEYQLVNVP-----DYEGRGESTPSPWFYQNQGEAEYIVSVYIYMRLLGYPANK 1284

Query: 726  IGVITPYEGQRAYILQYMQMNGSLDKDLYIKVEVASVDAFQGREKDYIILSCVRANEQQA 785
            I ++T Y GQ+  I   +         +    +V +VD FQG++ D+I+LS VR    +
Sbjct: 1285 ISILTTYNGQKLLIRDVINRRCVPYPFIGPPSKVTTVDKFQGQQNDFILLSLVRT---RF 1341

Query: 786  IGFLRDPRRLNVGLTRAKYGLVI 808
            +G LRD RRL V ++RA+ GL +
Sbjct: 1342 VGHLRDVRRLVVAMSRARLGLYV 1364
```

Score = 42.2 bits (97), Expect = 0.019
Identities = 25/82 (30%), Positives = 42/82 (50%), Gaps = 5/82 (6%)

```
Query: 396 PLPKEFSIPNFAQLNSSQSNAVSHVLQRPLSLIQGPPGTGKTVTSA---TIVYHLSKIHK 452
           P P++    N  +   +Q  A+   +Q  L+++ GPPGTGKT T+     ++YH
Sbjct: 853 PYPQDQPKQNSVKFTPTQVGAIISGIQPGLTMVVGPPGTGKTDTAVQILNVLYH--NCPS 910

Query: 453 DRILVCAPSNVAVDHLAAKLRD 474
            R L+   SN A++ L  K+ +
Sbjct: 911 QRTLIITHSNQALNDLFEKIME 932
```

gi|2648582 (AE000968) DNA helicase, putative [Archaeoglobus fulgidus]
Length = 453

Score = 97.9 bits (240), Expect = 3e-19
Identities = 114/417 (27%), Positives = 187/417 (44%), Gaps = 45/417 (10%)

```
Query: 427 LIQGPPGTGKTVTSATIVYHLSKIHKDRILVCAPSNVAVDHLAAKLRDLGLKVVRLTAKS 486
           L+ GPPG+GKT   A     LS+  ++R+ V + +N+AVD++  KL     + +R+   S
Sbjct: 71  LVVGPPGSGKTTFIAEAARKLSE--EERVWVTSNTNIAVDNVLEKLE----RALRIGHPS 124

Query: 487 REDVESSVSNLALHNLVGRGAKGELKNLLKLKDEVGELSASDTKRFVKLVRKTEAEILNK 546
           +       + +  H++       E   L  ++       AS      + + + + +++ K
Sbjct: 125 K-----ITAGVRRHSV-------EYNLLSGIRFSDYREYASKVAEAYREIARIQNDMIRK 172

Query: 547 ADVVCCTCV--GAGDKRLDTKFRTVLIDESTQASEPECLIPIVKGAKQVILVGDHQQLGP 604
             +V  + +  G      +  F TV+IDE++       L+ + K  K V+ VGD  QL P
Sbjct: 173 GKIVVGSTILKGMMSALKNYDFDTVIIDEASNTCISTALLALEKAEKAVV-VGDPYQLPP 231
```

```
Query: 605 V--ILERKAADAGLKQSLFERLISLGHVPIRLEVQYRMNPYLSEFPSNMFYEGSLQNGVT 662
           V  +   KAA    K S +  LI++    + L   YR N  ++EF +   Y G L+
Sbjct: 232 VYEVGGHKAA----KFSAYNFLINIYGKNLWLRRHYRSNAKIAEFAARHVY-GFLEIDEK 286

Query: 663 IEQRTVPNSKFPWPIRGIP---MMFWANYGREEISANGTSFLNRIEAMNCERIITKLFRD 719
             +  +   K   P  G P   ++F    G E     G S +N  EA     I  +L  +
Sbjct: 287 CRRIKIEPCKTTIPEVGDPEKPLVFIDCNGVER--REGRSKVNEEEAEVVAMIAEEL-AE 343

Query: 720 GVKPEQIGVITPYEGQRAYILQYMQMNGSLDKDLYIKVEVASVDAFQGREKDYIILSCVR 779
            V  E +GVITPY  Q   I   +   G       + VEV++V ++QGREKD +I S
Sbjct: 344 CVGEENVGVITPYVKQAELIRGILSEFG-------VMVEVSTVHSYQGREKDVVIYSITA 396

Query: 780 ANEQQAIGFLRDPRRLNVGLTRAKYGLVILGNPRSL-ARNTLWNHLLIHFREKGCLV 835
                   F  + R  NV LTRA+   + +GN ++L  +N L    ++   ++G  V
Sbjct: 397 TGNPY---FASEKRLFNVALTRARKKFIAVGNSKALEGKNLLLTRFMVAAIKEGGFV 450
```

gb|AAD36482.1|AE001793_12 (AE001793) helicase-related protein [Thermotoga maritima]
Length = 245

Score = 96.7 bits (237), Expect = 7e-19
Identities = 56/134 (41%), Positives = 79/134 (58%), Gaps = 10/134 (7%)

```
Query: 702 NRIEAMNCERIITKLFRDGVKPEQIGVITPYEGQRAYILQYMQMNGSLDKDLYIKVEVAS 761
           N +EA   + ++ KL   GVK + IG+ITPY+ Q   I + ++           KVEV S
Sbjct: 120 NPLEAQIVKEVVEKLLSMGVKEDWIGIITPYDDQVNLIRELIEA----------KVEVHS 169

Query: 762 VDAFQGREKDYIILSCVRANEQQAIGFLRDPRRLNVGLTRAKYGLVILGNPRSLARNTLW 821
           VD FQGREK+ II+S VR+N+   IGFL D RRLNV LTRAK  L+  G+  +L+ +  +
Sbjct: 170 VDGFQGREKEVIIISFVRSNKNGEIGFLEDLRRLNVSLTRAKRKLIATGDSSTLSVHPTY 229

Query: 822 NHLLIHFREKGCLV 835
              +   ++KG  V
Sbjct: 230 RRFVEFVKKKGTYV 243
```

emb|CAB49033.1| (AJ248283) DNA HELICASE RELATED PROTEIN [Pyrococcus abyssi]
Length = 1308

Score = 89.3 bits (218), Expect = 1e-16
Identities = 126/507 (24%), Positives = 210/507 (40%), Gaps = 112/507 (22%)

```
Query: 370  IDKKSISGYLYYKILGHQVVDISFDVPLPKEFSIPNFAQL----NSSQSNAVSHVLQRPL 425
            ID +    Y + K +    ++++  V L K+  I  F  L    N  Q   V  +  R +
Sbjct: 796  IDDEEHEIYEFLKNIYEGNINVNPKVKLWKKEHIEEFLSLLDHLNEDQRKFVLDIDHRIV 855

Query: 426  SLIQGPPGTGKT---VTSATIVYHLSKIHKDR---ILVCAPSNVAVDHLAAKLRDLGLKV 479
            +L QGPPGTGKT   +  A +    S I + +    +V A S+ AV+   A +R   LK
Sbjct: 856  TL-QGPPGTGKTSGAIAPAVLARAYSTIKQGKSALFIVTALSHRAVNE--ALIRTYRLKE 912

Query: 480  VRLTAKSREDVESSVSNLALHNLVGRGAKGELKNLLKLKDEVGELSASDTKRFVKLVRKT 539
                 K  ++VE           + RG  GE + +  +K E+G +      RF      +
Sbjct: 913  ELKNIKELKNVE-----------LIRGVSGE-EAIEPMKRELGGIKVDIRARFTF----S 956

Query: 540  EAEILNKADVVCCTCVGAGDKRLDTKFRTVLIDESTQASEPECLIPIVKGAKQVILVGDH 599
            ++ +     ++  T         + +   V+IDE++    P  L+       QV+LVGDH
Sbjct: 957  KSTLFPTVKILFATPQTVFKLVKNYEADLVVIDEASMMDLPMFLLATKNAKGQVLLVGDH 1016

Query: 600  QQLGPV------ILERKAADAGL------------KQSLFERLIS-----LGHVPI---- 632
            +Q+ P+      + +RK  +  L            +  L ER +      LG  P
Sbjct: 1017 RQMQPIQVHEWELEDRKTIEEHLPFLSALNFIRFLRGELEERELKRFKRILGRNPPKWKV 1076

Query: 633  ---------RLEVQYRMNPYLSEFPSNMFYEGSLQNGVTIEQRTVPNSKFPWPIRGIPMM 683
                     RL   +R+   L+E  S +FY     +G+ +  R  P+ +    +R
Sbjct: 1077 DKEDVLPMHRLRETFRLPQALAELHSELFYSF---DGIKLRSRKEPSEEIARALR----- 1128

Query: 684  FWANYGREEI----------------SANGTSFLNRIEAMNCERIITKLFRDGVKPEQIG 727
               N G+ E                   NG+S +N +EA   + I+  L  +GV    +G
Sbjct: 1129 ---NAGKNEFLRFILDPDYPVVLIVHDENGSSKVNELEAGIVKEIVEAL--NGV---DVG 1180

Query: 728  VITPYEGQRAYILQYMQMNGSLDKDLYIKVEVASVDAFQGREKDYIILSCVRANEQ---Q 784
            V+ PY  Q+  I               + V+V +V+ FQG E+D II+S   ++     +
Sbjct: 1181 VVVPYRAQKRLIRSL------------VNVQVDTVERFQGGERDIIIVSMTSSDPSYLLK 1228

Query: 785  AIGFLRDPRRLNVGLTRAKYGLVILGN 811
             + FL DP RLNV  +RA+  L+++ +
Sbjct: 1229 VMDFLYDPNRLNVAGSRAREKLILIAS 1255
```

emb|CAA99786| (Z75529) Similarity to C.elegans NAM7 protein (WP:C05C10.2); cDNA
EST yk239g7.5 comes from this gene [Caenorhabditis
elegans]
Length = 1106

Score = 82.7 bits (201), Expect = 1e-14
Identities = 119/516 (23%), Positives = 211/516 (40%), Gaps = 76/516 (14%)

```
Query: 383 ILGHQVVDISFDVPLPKEFSIPNFAQ--LNSSQSNAVSHVLQR-PLSLIQGPPGTGKTVT 439
```

```
           + G   ++ + ++   K+F  P+     LN  Q+  VS +L   PL L   P G GK++T
Sbjct: 512 LYGGSPINSTGNINWNKQFEFPSKIPIGLNQYQNQYVSMILSGVPLILGNSPFGCGKSMT 571

Query: 440 SATIVYHLSK-------IHKDRILVCAPSNVAVD--HLAAKLRDLG-LKVVRLTAKSRED 489
             T    L K        +K ++LV   +N  V+    A KL     ++ +R  ++S  +
Sbjct: 572 IVTAAVELQKKYMRSESFNKQQLLVTQSNNAGVNLIEFALKLHKCDHIRFLRYVSESNWN 631

Query: 490 V--ESSVSNLALHNLV-----------------GRGAKGELKNLL----KLKDEVGELSA 526
              +SS + L +  L+                 G     E+K  +      K     +
Sbjct: 632 SLPDSSKTELDMPILMEDEFFEWATNTEIQKENGENLTREMKKAIVEQSTQKYHCKGILV 691

Query: 527 SDTKRFVKLVRKTEAE------ILNKA-------DVVCCTCVGAGDKRLDTKFRTVL--- 570
            + KR  + + K + E      IL +A       D++  T     +   +   R V
Sbjct: 692 GEAKRIYENILKNQREHKPSQAILKEAFFALYAPDIIVSTADSIHNLLENNVLRDVSNIQ 751

Query: 571 IDESTQASEPECLIPIVKGAKQVI-LVGDHQQLGPVI---LERKAADAGLKQSLFERLIS 626
            DE++Q  E   +  + K       LVGD  QL       L+    D G+ +++
Sbjct: 752 FDEASQVPEHTLISLLHKFPNASFGLVGDINQLPAYCDDDLQGLLKDYGIGKTMARASRD 811

Query: 627 LGHVPIRLEVQYRMNPYLSEFPSNMFYEGSLQNGVTIEQRTVPNSKFP--WPIRGIPMMF 684
                  L + YR +P  ++  S++FY G L +GV   QR       P  WP    P++
Sbjct: 812 KMFPQSILRLVYRCHPTTTKILSDLFYNGQLISGVEKNQRDEFMRMRPDIWPNSSFPILV 871

Query: 685 WANYGREEISANGTSFLNRIEAMNCERIITKLFRD----GVKPEQIGVITPYEGQRAYIL 740
             N+ ++     GTS  N  E     RI++ L +     G+    IG+I+ Y+ Q + +
Sbjct: 872 -LNHEKDGYRM-GTSVANNSEKDQVVRIVSLLTKKVKGYGLSESDIGIISFYKAQTSLLA 929

Query: 741 QYMQMNGSLDKDLYIKVEVASVDAFQGREKDYIILSCVRANEQQAIGFLRDPRRLNVGLT 800
             ++            V+  ++DAFQG E++ +I+ C   N      F+++  R+NV ++
Sbjct: 930 DALKKTD---------VKCGTIDAFQGTEREVMIVCCTNKNPSL---FMQEGTRVNVAMS 977

Query: 801 RAKYGLVILGNPRSLARNTLWNHLLIHFREKGCLVE 836
           RA+   +I+GN   L +   W+ +    +   C +E
Sbjct: 978 RARQATIIIGNIAKLKKAKHWDTIAQSAQINKCSME 1013
```

gb|AAD07514.1| (AE000560) conserved hypothetical protein [Helicobacter pylori 26695]
Length = 360

Score = 82.3 bits (200), Expect = 2e-14
Identities = 75/298 (25%), Positives = 135/298 (45%), Gaps = 37/298 (12%)

```
Query: 543 ILNKADVVCCTCVGAGDKRLDTK----FRTVLIDESTQASEPECLIPIVKGAKQVILVGD 598
           +L   + V     G  ++ L+ K    F +V++DE+ +A+  E L+ +    +++ILVGD
Sbjct: 41  LLRDYNFVFSNTTGQHNQALERKETPYFDSVIVDEAAKANPLELLMVMALAKERIILVGD 100

Query: 599 HQQLGPVI-------LERKAADA------GLKQSLFERL---------ISLGHVPIRLEV 636
            +QL   +       LE ++ DA       LK S+F++L         +      I L +
Sbjct: 101 DRQLPHYLDDEIGKKLEDESQDAQDEIEKALKDSMFKKLKERAQKLKELDGKECFITLNM 160

Query: 637 QYRMNPYLSEFPSNMFYEGSLQNGVT-IEQRTVPNSKFPWPIRGIPM--MFWANYGREEI 693
           QY M+P L E  S++FY+   ++  + ++ + +    F   +R +    + W +    +
Sbjct: 161 QYGMHPLLGELVSDVFYKPHNESFESPLKGKHLEEKPFKHNLRVLDNKPLAWIDVKESKE 220

Query: 694 SANGTSFLNRIEAMNCERIITKLFRDGVKPEQIGVITPYEGQRAYILQYMQMNGSLDKDL 753
             N      R   +   +    LF         GVIT +  Q+  + Q ++
Sbjct: 221 KRNADGSYYRESEITAIKECLDLFMKDEPDFTFGVITFFSEQKRLLEQALKG-------- 272

Query: 754 YIKVEVASVDAFQGREKDYIILSCVRANEQQAIGFLRDPRRLNVGLTRAKYGLVILGN 811
           Y  +E+ +VD+FQG+E D + LS VR    +  GFL+    L V L+R K  L++ G+
Sbjct: 273 YANLEIGTVDSFQGKEFDVVFLSSVRTRHTEGFGFLKISSCLCVALSRQKRALIVAGD 330
```

dbj|BAA17223| (D90904) hypothetical protein [Synechocystis sp.]
Length = 1118

Score = 81.5 bits (198), Expect = 3e-14
Identities = 117/468 (25%), Positives = 201/468 (42%), Gaps = 77/468 (16%)

```
Query: 409  LNSSQSNAVSHVLQ-RPLSLIQGPPGTGKT-VTSATIVYHLSKIHKDRILVCAPSNVAVD 466
            L++ Q+  +   L+ +   L+QGPPGTGKT +     V +L K++K  I + A +N AV+
Sbjct: 673  LSNDQNQIICQALEAKHYYLLQGPPGTGKTSIFLVNYVQNLLKLNKKNIFILAFTNKAVE 732

Query: 467  HLAAKLRD--LGLKVVRLTAKSREDVESSVSNLALHNLVGRGAKGELKNLLKLKDEVGEL 524
             +   L+    G  +  L   ++  V+ S   L    L G  A  +L      +++V
Sbjct: 733  QICKALKSPRYGELIDYLRLGNKTVVDES---LLAEKLTGDNA--DLWRKYLQENQVFVT 787

Query: 525  SASDTKRFVKLVRKTEAEILNKADVVCCTCVGAGDKRLDTKFRTVLIDESTQASEPECLI 584
            + S  K    L+R+   E                       F  ++IDE++Q +E + L
Sbjct: 788  TVSTFKDKYLLLREFVGE-----------------------FDQLIIDEASQLTEAD-LA 823

Query: 585  PIVKGAKQVILVGDHQQLGPVILERKAA--------------DAGLKQSLFERLIS---- 626
             IV    + +L+GD +QL  VI + +                  + L+ SLFERLI+
Sbjct: 824  GIVVLFNKFVLIGDQKQLPAVITQNEEECRVEVEPFRRYAINISDLRISLFERLINNAKN 883

Query: 627  --LGHVPIRLEVQYRMNPYLSEFPSNMFYEGSLQNGVTIEQRTVPNSKFPWPIRGIPMMF 684
                H   +L   YRM+  ++   ++ +    ++   T EQR + +S   +      ++
Sbjct: 884  KGWNHCYGQLTEHYRMHRDIASLIAHHYDRPLVEQ--TQEQRKIESS---YENNDNDLLS 938
```

```
Query: 685  WANYGR----EEISANGTSFLNRIEAMNCERIITKLFRDGVKP-EQIGVITPYEGQRAYI 739
             +  R    E  S       N  EA+    I+   + + + P  +IG+I P+  Q A I
Sbjct: 939  KLSKSRTIFIETPSETSIQRKNTNEAIIAATIVKTYWENKLLPMAEIGIIAPFRAQVAEI 998

Query: 740  LQYMQ---MNGSLDKDLYIKVEVASVDAFQGREKDYIILSCVRANEQQA--IGFLRD--- 791
             Q ++         +D    + V +V+ FQG E++ II S   +  +Q   I  + D
Sbjct: 999  NQQLEKIFAGADFSRD---SIVVDTVERFQGDERELIIFSTTISWSKQVKNIQSIADGDR 1055

Query: 792  ---PRRLNVGLTRAKYGLVILGNPRSLARNTLWNHLLIHFREKGCLVE 836
                R+L V ++RAK  LVILGN   L     +  L+ H  +   L++
Sbjct: 1056 QGTDRKLLVSISRAKSKLVILGNSSQLQFAPAYRELIKHIEQSNGLID 1103
```

emb|CAB07226| (Z92790) H03G16.3 [Caenorhabditis elegans]
Length = 646

Score = 76.1 bits (184), Expect = 1e-12
Identities = 109/433 (25%), Positives = 178/433 (40%), Gaps = 78/433 (18%)

```
Query: 430 GPPGTGKTVTSATIVYHLSKIHKDRILVCAPSNVAVDHLAAKLRDLGLKVV-RLTAKSRE 488
           G PGTGKT T   ++    K+ K +I+V AP+   ++++    +++  + + +L  K  +
Sbjct: 120 GAPGTGKTYTLTLLISCAVKLKK-QIVVLAPTFETLENI----KNMATRTMEKLGIKCSD 174

Query: 489 DVESSVSNLALHNLVGRGAKGELKNLLKLKDEVGELS---ASDTKRFVKLVRKTEAEILN 545
           +V  S S     NL+           L+ +D +G +    A+    F+  V K EA
Sbjct: 175 NVFMSTSQYI--NLIDTEP-------LQKEDMIGNVLVVLATIGADFLNFVAKQEA---- 221

Query: 546 KADVVCCTCVGAGDKRLDTKFRTVLIDESTQASEPECLIPIVKGAKQVILVGDHQQLGPV 605
                            LD  F   +IDE+      E   P V  +K  ++ GD  +L P
Sbjct: 222 ----------------LDPSF--CIIDEAGLVMASETW-PDVYKSKGFVIAGDPNELFPQ 262

Query: 606 ILERKAADAG------LKQSLFERLISL---------GHVPIRLEVQYRMNPYLSEFPSN 650
           I      +        L+Q   E  +S+            PI L+ QYR N  +  + +
Sbjct: 263 IYNHDNGNGFFNQFCCLEQKRLETQVSILDCINEKTENLSPIILDTQYRSNETIMAWSNA 322

Query: 651 MFYEGSLQNGVTIEQRTVPN----SKFPWPIRGIPMMFWANYGREEISANGT-------- 698
            FY   ++        T+ +    SKF  P+    ++   +  +E    N T
Sbjct: 323 CFYNPKVKTDFKCGSLTLRSFFKGSKFN-PLFHPFVIVDTSCEKEMEKRNETFEYLVFNN 381

Query: 699 -----SFLNRIEAMNCERIITKLFRDGVKPEQIGVITPYEGQRAYILQYMQM--NGSLDK 751
                 + N  EA        +L +  V+   I +ITPY GQ   + Q M+   N
Sbjct: 382 NNENEGYANVAEAKIAMHHYKQLLKLKVQASDIAIITPYRGQSHVVRQMMRKFANEESRS 441

Query: 752 DLYIKVEVASVDAFQGREKDYIILSCVRANEQQAIGFLRDPRRLNVGLTRAKYGLVILGN 811
           D +    V  VD   G+E D +I + VR+N +  +G   DPRRL+V +T+AK   + +GN
Sbjct: 442 DCF-DTHVGDVDKIVGKEFDVVIFTMVRSNPKLDMGKAGDPRRLHVIITKAKKHFMFIGN 500

Query: 812 PRSLARNTLWNHL 824
              LA N + NH+
Sbjct: 501 GFLLAENNM-NHI 512
```

dbj|BAA29178.1| (AP000001) 1188aa long hypothetical protein [Pyrococcus horikoshii]
Length = 1188

Score = 71.0 bits (171), Expect = 4e-11
Identities = 110/456 (24%), Positives = 191/456 (41%), Gaps = 92/456 (20%)

```
Query: 409  LNSSQSNAVSHVLQRPLSLIQGPPGTGKT---VTSATIVYHLSKIHKDR---ILVCAPSN 462
            LN  Q N       R ++L QGPPGTGKT   +  A +    S I + +    +V A S+
Sbjct: 721  LNREQVNFALDCEHRIVTL-QGPPGTGKTSGAIAPAILARAYSTIKQGKSSLFIVTALSH 779

Query: 463  VAVDHLAAKLRDLGLKVVRLTAKSREDVESSVSNLALHNLVGRGAKGELKNLLKLKDEVG 522
             AV+   A +R   LK      K  ++VE           + RG   E + +  ++ E+
Sbjct: 780  RAVNE--ALIRTYKLKEKLKDIKELKNVE-----------LIRGVSSE-EAVKPMEKELN 825

Query: 523  ELSASDTKRFVKLVRKTEAEILNKADVVCCTCVGAGDKRLDTKFRTVLIDESTQASEPEC 582
             L  + T +F      +++ +     ++  T   A     D     V+IDE++    P
Sbjct: 826  GLKVNVTNKF----SFSKSPLFLTVKILFATPQTAFKLAKDYDADLVVIDEASMLDLPMF 881

Query: 583  LIPIVKGAKQVILVGDHQQLGPV------ILERKAADAGL------------KQSLFERL 624
             +       QV+LVGDH+Q+ P+      + +RK  +  L            +  L ER
Sbjct: 882  FLATSNAKGQVLLVGDHRQMQPIQVHEWELEDRKTIEEHLPFLSVLNFIRFLRGELEERE 941

Query: 625  IS-----LGHVPIRLEVQ-------------YRMNPYLSEFPSNMFYEGSLQNGVTIEQR 666
            +      LG  P R  V              +R+   L++  S +FY     +G+ +  R
Sbjct: 942  LKRFKRILGRDPPRWNVDKDRVLPMHRLRETFRLPRALAKLHSELFYSF---DGIELISR 998

Query: 667  TVPNSKFPWPIRGIP----MMFWANYGREEI----SANGTSFLNRIEAMNCERIITKLFR 718
               + +    ++       + F  + G   I    +  G++ +N +EA     I+  + +
Sbjct: 999  KNSDREVLETLKKAGKDEFLKFILDPGYPVILIIHNEGGSTKVNELEA----EIVKDILK 1054

Query: 719  DGVKPEQIGVITPYEGQRAYILQYMQMNGSLDKDLYIKVEVASVDAFQGREKDYIILSCV 778
            + VK   +GV+ PY  Q+  I               + V+V +V+ FQG EKD II+S
Sbjct: 1055 E-VKGIDVGVVVPYRAQKRLIRSL------------VNVQVDTVERFQGGEKDVIIVSMT 1101

Query: 779  RANE---QQAIGFLRDPRRLNVGLTRAKYGLVILGN 811
             ++     + + F+ +P RLNV  +RAK  L+++ +
```

```
Sbjct: 1102 SSDPAYLSKVLEFIYNPNRLNVAGSRAKEKLILIAS 1137


 gi|183250 (M64979) glial factor-1 [Homo sapiens]
           Length = 376

 Score = 70.2 bits (169), Expect = 7e-11
 Identities = 40/120 (33%), Positives = 67/120 (55%), Gaps = 7/120 (5%)

Query: 713 ITKLFRDGVKPEQIGVITPYEGQRAYILQYMQMNGSLDKDLYIKVEVASVDAFQGREKDY 772
           I  L   GV    I V++PY  Q   + Q +          + ++E+ SVD FQGREK+
Sbjct: 32  IQALVDAGVPARDIAVVSPYNLQVDLLRQSLVHR-------HPELEIKSVDGFQGREKEA 84

Query: 773 IILSCVRANEQQAIGFLRDPRRLNVGLTRAKYGLVILGNPRSLARNTLWNHLLIHFREKG 832
           +ILS VR+N +  +GFL + RR+NV +TRA+  + ++ + R++  +     L+ +F + G
Sbjct: 85  VILSFVRSNRKGEVGFLAEDRRINVAVTRARRHVAVICDSRTVNNHAFLKTLVEYFTQHG 144


 gi|3329511 (AF075169) DNA helicase homolog [Schizosaccharomyces pombe]
           Length = 191

 Score = 69.5 bits (167), Expect = 1e-10
 Identities = 58/193 (30%), Positives = 89/193 (46%), Gaps = 31/193 (16%)

Query: 565 KFRTVLIDESTQASEPECLIPIVKGAKQVILVGDHQQLGPVILERKAADAGLKQSLFERL 624
           KF   +IDE++Q   P CL P+ + A++ +LVGDH QL P++   + +  GL  SLF +L
Sbjct: 5   KFDYCIIDEASQIPLPICLGPL-QLAEKFVLVGDHYQLPPLVKNSRTSKDGLSLSLF-KL 62

Query: 625 ISLGH--VPIRLEVQYRMNPYLSEFPSNMFYEGSLQNGVTIEQRTVPNSKFPWP----IR 678
           +S  H      L +QYRMN  ++   S + Y G+L  G     +T+   K   P
Sbjct: 63  LSEKHPEAVTTLRLQYRMNEDINSLSSELIYGGNLVCG----SKTISQKKLILPKAHLSD 118

Query: 679 GIP----MMFWAN-----------YGREEI----SANGTSFLNRIEAMNCERIITKLFRD 719
           G+P     + W N           +  ++I    S       N  EA   E+ ++
Sbjct: 119 GLPDSSSSLHWVNKLINPSHSVIFFNTDDILGVESKTNNILENHTEAFLIEQAVSSFLER 178

Query: 720 GVKPEQIGVITPY 732
           GVK   IG+I+ Y
Sbjct: 179 GVKQSSIGIISIY 191


 emb|CAA15907| (AL021006) hypothetical protein Rv1251c [Mycobacterium tuberculosis]
            Length = 1139

 Score = 65.2 bits (156), Expect = 2e-09
 Identities = 94/404 (23%), Positives = 155/404 (38%), Gaps = 79/404 (19%)

Query: 428  IQGPPGTGKTVTSATIVYHLSKIHKDRILVCAPSNVAVDHLAAKLRDLGLKVVRLTAKSR 487
            + GPPGTGKT T+A ++  L   H  RI V A S+  V++L   +   GL   ++  K
Sbjct: 764  VHGPPGTGKTYTAARVIAELVTEHAWRIGVVAQSHATVENLLEGVISAGLDPGQVAKKPH 823

Query: 488  EDVESSVSNLALHNLVGRGAKGELKNLLKLKDEVGELSASDTKRFVKLVRKTEAEILNKA 547
            +                                G   + D  ++ + +R T   ++
Sbjct: 824  DHT-----------------------------AGRWQSIDGSQYTEFIRDTAGCVIGGT 853

Query: 548  DVVCCTCVGAGDKRLDTKFRTVLIDESTQASEPECL---IPIVKGAKQVILVGDHQQLGP 604
                      G++        ++IDE+ Q     CL   I +   A  ++L+GD QQL
Sbjct: 854  ----AWDFANGNRVPKASLDLLVIDEAGQF----CLANTIAVAPAATNLLLLGDPQQLPQ 905

Query: 605  VILERKAADAGLKQSLFERLISLGHV-----PIRLEVQYRMNPYLSEFPSNMFYEGSLQN 659
            V   +      +  S    L+   H         L+  YRM+P +    S + YEG L +
Sbjct: 906  V--SQGTHPEPVDTSALSWLVDGQHTLPDERGYFLDRSYRMHPAVCAAVSALSYEGRLCS 963

Query: 660  GVTIEQRTVPNSKFPWP----IRGIPMMFWANYGREEISANGTSFLNRIEAMNCE--RII 713
                 +RT       +P     RG+               N        EA+  E  +++
Sbjct: 964  HT---ERTAVRRLDGYPPGVHTRGV-----------HHKGNSIESPEEAEAILAELRQLL 1009

Query: 714  TKLFRD--GVKP---EQIGVITPYEGQRAYILQYMQMNGSLDKDLYIKVEVASVDAFQGR 768
               + D  G +P     + V+ PY  Q A + + +   G    D    V V +VD FQG
Sbjct: 1010 GSPWTDEHGTRPLAASDVLVLAPYNAQVALVRRRLASAGLGGAD---GVRVGTVDKFQGG 1066

Query: 769  EKDYIILSCVRANEQ---QAIGFLRDPRRLNVGLTRAKYGLVIL 809
            +   + +S   ++     + I FL +  RLNV ++RA+Y  VI+
Sbjct: 1067 QAPVVFISMTASSADDVPRGISFLLNRNRLNVAVSRAQYAAVIV 1110


 gi|1752701 (U60176) cag-omega [Helicobacter pylori]
           Length = 127

 Score = 58.2 bits (138), Expect = 3e-07
 Identities = 33/85 (38%), Positives = 51/85 (59%), Gaps = 5/85 (5%)

Query: 726 IGVITPYEGQRAYILQYMQMNGSLDKDLYIKVEVASVDAFQGREKDYIILSCVRANEQQA 785
           IG+ITPY  Q+  +   ++  G  + D   ++++ +VDAFQG E D II S V+
Sbjct: 24  IGIITPYNAQKRRLRSEVEKYGFKNFD---ELKIDTVDAFQGEEADIIIYSTVKTCGN-- 78

Query: 786 IGFLRDPRRLNVGLTRAKYGLVILG 810
           + FL D +RLNV ++RAK  L+ +G
Sbjct: 79  LSFLLDSKRLNVAISRAKENLIFVG 103
```

```
 gb|AAD35703.1|AE001736_1 (AE001736) conserved hypothetical protein [Thermotoga maritima]
           Length = 1289

 Score = 57.4 bits (136), Expect = 5e-07
 Identities = 73/288 (25%), Positives = 125/288 (43%), Gaps = 65/288 (22%)

Query: 566  FRTVLIDESTQASEPECLIPIVKGAKQVILVGDHQQLGPVIL------------ERKAAD 613
            F  V+ DE++Q   PE +  +V+G KQ+I+ GD +QL P               ER+  D
Sbjct: 888  FDIVIFDEASQLRTPEAISAVVRG-KQIIVAGDPKQLPPTNFFKSYYELEDDEDEREPLD 946

Query: 614  AGLKQSL-FERLISLGHVPIRLEVQYRM-NPYLSEFPSNMFYEGSLQNGVTIEQRTVPNS 671
            + L + +   R+   G+    L   YR  +  L  F ++ FY    +N +     T P+
Sbjct: 947  SFLDECIALPRVFKQGY----LRWHYRSRDERLIAFSNHYFYG---ENPLI----TFPSP 995

Query: 672  KFPWPIRGIPMMFWANYGREEISANGTSFLNRIEAMNCERIITKLFRDGVKPEQIGVITP 731
            K+    +GI +++  N   +         +N +EA+    I+ + F+       IGV+T
Sbjct: 996  KYRNSDQGIKLVYVENGTWDRAGKR----VNTMEALKVVDIVIEHFQKH-PDRSIGVVTM 1050

Query: 732  YEGQRAYI--------LQYMQMNGSLDKDLYIKVEVASVDAFQGREKDYIILSCVRANEQ 783
               Q   I        ++Y  +   + K+      + S++  QG E+D II+S
Sbjct: 1051 NTSQSDLIENLLQRRLMEYPHLMDVIFKESNEPFFIKSLENVQGDERDTIIIS------- 1103

Query: 784  QAIGFLRDP-----------------RRLNVGLTRAKYGLVILGNPRS 814
              IG+ R P                 RRLNV +TRA+Y ++++ + RS
Sbjct: 1104 --IGYARTPSGELFYNFGPLNNEGGWRRLNVLITRARYQIILVTSLRS 1149


 Score = 35.6 bits (80), Expect = 1.8
 Identities = 26/85 (30%), Positives = 43/85 (50%), Gaps = 5/85 (5%)

Query: 410 NSSQSNAVSHVLQRPLSLIQGPPGTGKTVTSATIVYHLSKIHKDRILVCAPSNVAVDHLA 469
           +SSQ   +  V +    +I GPPGTGK+ T A I+       K  + V A    A+D +
Sbjct: 252 DSSQRKILEMVREGENVVIHGPPGTGKSQTIANIIADAVAREKTVLFVSA-KKAALDVVY 310

Query: 470 AKLRDLGL----KVVRLTAKSREDV 490
            +L++ GL      +  T KS++++
Sbjct: 311 NRLKNAGLGRFCLEIHSTRKSKQEL 335


 gi|1280167 (U55857) K08D10.5 gene product [Caenorhabditis elegans]
           Length = 602

 Score = 55.4 bits (131), Expect = 2e-06
 Identities = 63/249 (25%), Positives = 109/249 (43%), Gaps = 28/249 (11%)

Query: 592 QVILVGDHQQLGPVILERKAADAGLKQSLFERLISLG-----HVPIRLEVQYRMNPYLSE 646
           ++ILVGD  QL P   +  A  A LK++     ++L         + L   +R    +++
Sbjct: 342 RIILVGDMHQLPPYTED--ALPAELKRAAVGEPLTLAVKGRRWSSMHLTRVHRCPKMITD 399

Query: 647 FPSNMFYEGSLQNGVTIEQRTVPNSKFPWPIRGIPMMFWANYGREEISANGTSFLNRIEA 706
               +FY  +L +      R          P    P +F  NY  ++    G SF N  EA
Sbjct: 400 LLGAVFYGNTLTSTRKAVNRVPILEAMELPSHH-PFVF-VNYISQQTQV-GKSFSNENEA 456

Query: 707 MNCERIITKLFRDGVKPE---QIGVITPYEGQRAYILQYMQMNGSLDKDLYIKVEVASVD 763
               +++T L R   +      + V+  Y  Q +Y+        S+ +D    V V ++D
Sbjct: 457 RYALQLVTALKRHAREQNTTVSVAVLNFYGAQYSYVY-------SIAED---GVRVNTID 506

Query: 764 AFQGREKDYIILSCVRANEQQAIGFLRDPRRLNVGLTRAKYGLVILGNPRSLARNTLWNH 823
             QG+E D  I+   R++  +   FL +P R+NV L+R+K   V++G      R+ + NH
Sbjct: 507 GCQGQEYDVSIVLLTRSDPYEKSKFLVNPNRINVALSRSKMATVVIGQ-----RHLIGNH 561

Query: 824 LLIHFREKG 832
             +   ++G
Sbjct: 562 PNVQQAKRG 570


 emb|CAA86856| (Z46828) similar to helicase domain; cDNA EST EMBL:C08769 comes
           from this gene [Caenorhabditis elegans]
           Length = 542

 Score = 54.7 bits (129), Expect = 3e-06
 Identities = 71/262 (27%), Positives = 117/262 (44%), Gaps = 38/262 (14%)

Query: 565 KFRTVLIDESTQAS-----EPECLIPIVKGAKQVILVGDHQQLGPVILERKAADAGLKQS 619
           +F    IDE++Q +         ++P      ++ILVGD  QL P + E  A  A LK++
Sbjct: 215 RFDITFIDEASQLALYVLGSLATMLP----KSRMILVGDMHQLPPYMEE--ALPAELKRA 268

Query: 620 LFERLISLG-----HVPIRLEVQYRMNPYLSEFPSNMFYEGSLQN---GVTIEQRTVPNS 671
               ++L         + L   +R    ++E   ++FY  +L +   GVT     +P
Sbjct: 269 AIGEPLTLAVKGRRWPSMHLTRVHRCPKMITEVLGDLFYGNTLTSSKPGVT----DIPVL 324

Query: 672 KFPWPIRGIPMMFWANYGREEISANGTSFLNRIEAMNCERIITKLFR---DGVKPEQIGV 728
           K        PM+F  NY   + +A G SF N  EA    +++  L R      K     +
Sbjct: 325 KAMGLPSRHPMVF-VNYTSPQ-TAVGKSFSNEGEARYALQLVEALTRYASTANKKITAAI 382

Query: 729 ITPYEGQRAYILQYMQMNGSLDKDLYIKVEVASVDAFQGREKDYIILSCVRANEQQAIGF 788
           +  Y  Q +Y+        S+ +D   +V V ++D  QG+E D  I+   R++  +   F
```

```
Sbjct: 383 LNFYGAQYSYVY-------SMAED---EVTVNTIDGCQGQEYDVTIVLLTRSDPYERSKF 432

Query: 789 LRDPRRLNVGLTRAKYGLVILG 810
           L +  R+NV L+R K   VI+G
Sbjct: 433 LVNANRINVALSRPKIATVIIG 454
```

gb|AAD27693.1|AF125214_3 (AF125214) urease-enhancing factor helB [Helicobacter pylori]
Length = 185

Score = 49.2 bits (115), Expect = 1e-04
Identities = 43/152 (28%), Positives = 72/152 (47%), Gaps = 17/152 (11%)

```
Query: 649 SNMFYEGSLQNGVTIEQRTVPNSKFPWPIRGIPMMFWANYGREEISANGTSFLNRIEAMN 708
           S +FY+  L+NGV      + N+   +  + I  + W N   E      TS  N+ +
Sbjct: 10  SQLFYKEKLKNGV------IKNTSQFYDPKNI--IRWINVEGEH-KLEKTSSYNKNQVQK 60

Query: 709 CERI---ITKLFRDGVKPEQIGVITPYEGQRAYILQYMQMNGSLDKDLYIKVEVASVDAF 765
              +   I ++       + IG+ITPY  Q+  +   ++  G  + D   ++++ +VDAF
Sbjct: 61  IIELLEQINRILNQRKIRKTIGIITPYNAQKRCLRSEVEKYGFKNFD---ELKIDTVDAF 117

Query: 766 QGREKDYIILSCVRANEQQAIGFLRDPRRLNV 797
           QG + D II S V+      + FL D +RLNV
Sbjct: 118 QGEKADIIIYSTVKTYGN--LSFLIDSKRLNV 147
```

sp|P26314|RRPB_IBVB RNA-DIRECTED RNA POLYMERASE (ORF1B) >gi|74826|pir||VFIHB2
RNA-directed RNA polymerase (EC 2.7.7.48) - avian
infectious bronchitis virus (strain Beaudette) >gi|292953
(M95169) pol protein [Avian infectious bronchitis virus]
>gi|331173 (M94356) ORF 1b encodes a polypeptide of
potential mol. wt. 300,000. [Avian infectious bronchitis
virus]
Length = 2652

Score = 48.0 bits (112), Expect = 3e-04
Identities = 51/190 (26%), Positives = 85/190 (43%), Gaps = 37/190 (19%)

```
Query: 686  ANYGREEISANGTSFLNRIEAMNCERIITKLFRDGVKPEQIGVITPYEG--QRAYILQYM 743
            ++ G E  SA  T+ L  ++   C     K +R+ +       I+PY    QRAY +
Sbjct: 1371 SDVGHESGSAYNTTQLEFVKDFVCRN---KQWREAI------FISPYNAMNQRAYRM--- 1418

Query: 744  QMNGSLDKDLYIKVEVASVDAFQGREKDYIILSCVRANEQQAIGFLRDPRRLNVGLTRAK 803
                       + + V +VD+ QG E DY+I  CV A+ Q A+       R NV LTRAK
Sbjct: 1419 -----------LGLNVQTVDSSQGSEYDYVIF-CVTADSQHALNI----NRFNVALTRAK 1462

Query: 804  YGLVILGNPRSLARNTLWNHLLIHFREKGCLVEGTLDNLQLCTVQL--VRPQPRKTERPM 861
             G++++       R+ L++ L     +    ++GT    ++C  +   V P    T + +
Sbjct: 1463 RGILVV----MRQRDELYSALKFTELDSETSLQGT-GLFKICNKEFSGVHPAYAVTTKAL 1517

Query: 862  NAQFNVESEM 871
             A + V  E+
Sbjct: 1518 AATYKVNDEL 1527
```

Score = 35.2 bits (79), Expect = 2.4
Identities = 41/165 (24%), Positives = 64/165 (37%), Gaps = 32/165 (19%)

```
Query: 330  LELKPSKTPPPTHLTTGFTAEFIWKGTSYDRMQDALKKFAIDKKSISGYLYYK------- 382
            L  +P KT PP +    FT    +  T   ++Q  L  F  +K      +YYK
Sbjct: 1056 LSWEPGKTRPPLNRNYVFTG---YHFTRTSKVQ--LGDFTFEKGEGKDVVYYKATSTAKL 1110

Query: 383  -------ILGHQVVDISFDVPLPKEFSIPNFAQLNSSQSNAVSHVLQRPL---------S 426
                   +  H VV +      P++ +   F  L  +       V   PL         +
Sbjct: 1111 SVGDIFVLTSHNVVSLVAPTLCPQQ-TFSRFVNLRPNVMVPECFVNNIPLYHLVGKQKRT 1169

Query: 427  LIQGPPGTGKTVTSATIVYHLSKIHKDRILVCAPSNVAVDHLAAK 471
             +QGPPG+GK+  +  +  + S     R++  A S+ AVD L  K
Sbjct: 1170 TVQGPPGSGKSHFAIGLAVYFSSA---RVVFTACSHAAVDALCEK 1211
```

emb|CAA83018| (Z30541) potential chimeric protein [avian infectious bronchitis
virus]
Length = 2155

Score = 48.0 bits (112), Expect = 3e-04
Identities = 51/190 (26%), Positives = 85/190 (43%), Gaps = 37/190 (19%)

```
Query: 686  ANYGREEISANGTSFLNRIEAMNCERIITKLFRDGVKPEQIGVITPYEG--QRAYILQYM 743
            ++ G E  SA  T+ L  ++   C     K +R+ +       I+PY    QRAY +
Sbjct: 1429 SDVGHESGSAYNTTQLEFVKDFVCRN---KQWREAI------FISPYNAMNQRAYRM--- 1476

Query: 744  QMNGSLDKDLYIKVEVASVDAFQGREKDYIILSCVRANEQQAIGFLRDPRRLNVGLTRAK 803
                       + + V +VD+ QG E DY+I  CV A+ Q A+       R NV LTRAK
Sbjct: 1477 -----------LGLNVQTVDSSQGSEYDYVIF-CVTADSQHALNI----NRFNVALTRAK 1520

Query: 804  YGLVILGNPRSLARNTLWNHLLIHFREKGCLVEGTLDNLQLCTVQL--VRPQPRKTERPM 861
             G++++       R+ L++ L     +    ++GT    ++C  +   V P    T + +
Sbjct: 1521 RGILVV----MRQRDELYSALKFTELDSETSLQGT-GLFKICNKEFSGVHPAYAVTTKAL 1575
```

```
Query: 862  NAQFNVESEM 871
             A + V  E+
Sbjct: 1576 AATYKVNDEL 1585


 Score = 35.2 bits (79), Expect = 2.4
 Identities = 41/165 (24%), Positives = 64/165 (37%), Gaps = 32/165 (19%)

Query: 330  LELKPSKTPPPTHLTTGFTAEFIWKGTSYDRMQDALKKFAIDKKSISGYLYYK------- 382
            L  +P KT PP +    FT    +  T   ++Q  L  F  +K      +YYK
Sbjct: 1114 LSWEPGKTRPPLNRNYVFTG---YHFTRTSKVQ--LGDFTFEKGEGKDVVYYKATSTAKL 1168

Query: 383  -------ILGHQVVDISFDVPLPKEFSIPNFAQLNSSQSNAVSHVLQRPL---------S 426
                   +  H VV +      P++ +   F  L  +       V   PL         +
Sbjct: 1169 SVGDIFVLTSHNVVSLVAPTLCPQQ-TFSRFVNLRPNVMVPECFVNNIPLYHLVGKQKRT 1227

Query: 427  LIQGPPGTGKTVTSATIVYHLSKIHKDRILVCAPSNVAVDHLAAK 471
             +QGPPG+GK+  +  +  + S     R++  A S+ AVD L  K
Sbjct: 1228 TVQGPPGSGKSHFAIGLAVYFSSA---RVVFTACSHAAVDALCEK 1269


 gb|AAD07506.1| (AE000559) H. pylori predicted coding region HP0435 [Helicobacter
          pylori 26695]
          Length = 191

 Score = 44.5 bits (103), Expect = 0.004
 Identities = 32/119 (26%), Positives = 57/119 (47%), Gaps = 7/119 (5%)

Query: 412 SQSNAVSHVLQRP-LSLIQGPPGTGKTVTSATIV------YHLSKIHKDRILVCAPSNVA 464
           +Q  A+   L  P +++IQGPPGTGKT     I       Y   K  K +IL+CA  + A
Sbjct: 35  NQKKAIKIALNTPDIAIIQGPPGTGKTTVINAICERLFEEYPKDKNIKGQILLCAQGHDA 94

Query: 465 VDHLAAKLRDLGLKVVRLTAKSREDVESSVSNLALHNLVGRGAKGELKNLLKLKDEVGE 523
            ++   +++  GL   +  AK     E    +  L+  +   A+  ++++ K   ++G+
Sbjct: 95  TNNARERIKVGGLPTFKFGAKKNAKEEQYKQDERLNERLREFAETLIESVRKKLQKLGD 153


 sp|P75033|Y140_MYCPN HYPOTHETICAL ATP-BINDING PROTEIN MG140 HOMOLOG
            >gi|2146373|pir||S73327 MG140 homolog - Mycoplasma
            pneumoniae (SGC3) (ATCC 29342) >gi|1673646 (AE000001)
            Mycoplasma pneumoniae, MG140 homolog, from M. genitalium
            [Mycoplasma pneumoniae]
            Length = 1113

 Score = 42.2 bits (97), Expect = 0.019
 Identities = 90/408 (22%), Positives = 159/408 (38%), Gaps = 70/408 (17%)

Query: 569  VLIDESTQASEPECLIPIVKGAKQVILVGDHQQLGPVILERKAADAGLKQ---------- 618
            V+IDE++Q      +  + +GAK +I  GD +QL P    +  A+  + +
Sbjct: 699  VVIDEASQVYLERAIPILYRGAKYII-AGDTKQLKPSNFFQARAEYDVDEEFEDGNVEAA 757

Query: 619  ----SLFERLISLGHVPIRLEVQYRMNPY-LSEFPSNMFYEGSLQNGVTIEQRTVPNSKF 673
                SL   L +   +   L+  YR +   L  F +N  Y   L   + + + T     F
Sbjct: 758  VHSTSLLHFLKNRSRILTLLKFHYRSDSANLIAFTNNRIYNNEL---IFMNKATADKQVF 814

Query: 674  PWPIRGIPMMFWANYGREEISANGTSFLNRIEAMNCERIITKLFRDGVKPEQIGVITPYE 733
               +  +    W N              N  EA +  + + +L +     + +GVI  +
Sbjct: 815  I--VHDVIDGIWRNNR-----------NLQEARDVVQRLEQLTQTAEYQKSLGVIC-FN 859

Query: 734  GQRAYILQYM---QMNGSLDK--------DLYIKVEVASVDAFQGREKDYIILSCV---R 779
              +A +++YM   Q N  L++          Y+ + V +++  QG E+D II S
Sbjct: 860  KNQAELIEYMIDKQNNPLLNEWRDRVNAQGEYVGLFVKNIENVQGDERDIIIFSLGYDRS 919

Query: 780  ANEQQAIGFLRDPRRLNVGLTRAKYGLVILGNPRS-----LARNTLWNHLLIHF-REKGC 833
             N    I       RLNV +TRAK  + +    R+     L+ N+L + LL+ +
Sbjct: 920  VNSYGPISKQGGENRLNVAITRAKQRIELFKTNRASDYNGLSSNSLGSKLLVEYLLYCEA 979

Query: 834  LVEGTLDNLQLCTVQLVRPQPRKTERPMNAQFNVESEMGDFPKFQDFDAQSMVSFSGQIG 893
            +     ++L     Q   P+  K E  +  QF  E E+        F  Q  +  +   G
Sbjct: 980  MANNQGESLDFQATQKQAPK-AKYELELENQFFNELEL-------IFGEQFTIKRNVNEG 1031

Query: 894  DFGNAFVDNTELSSYINNEYWNFENFKSAFSQKQNRNEIDDRNLYQEE 941
             +  +FV       Y N   +   +F  A     +R E+ +  +Y+E+
Sbjct: 1032 AYSFSFV------FYFNENPYLAVDFNPAL--PHSRKEVSENIIYREQ 1071


 Score = 34.4 bits (77), Expect = 4.1
 Identities = 20/61 (32%), Positives = 35/61 (56%), Gaps = 2/61 (3%)

Query: 406 FAQLNSSQSNAVSHVLQRPLSLIQGPPGTGKTVTSATIVYHLSKIHKDRILVCAPSNVAV 465
           F+ L+  Q  AV       + +I+GPPGTGK+ T   I+ +L+ ++K ++L  +    A+
Sbjct: 290 FSHLDIFQQYAVKQAFDGDV-IIEGPPGTGKSETIVNILVNLA-LNKKKVLFVSEKVTAL 347

Query: 466 D 466
           D
Sbjct: 348 D 348
```

```
 gi|3329101 (AE001335) Exodeoxyribonuclease V, Alpha [Chlamydia trachomatis]
           Length = 496

 Score = 42.2 bits (97), Expect = 0.019
 Identities = 36/132 (27%), Positives = 60/132 (45%), Gaps = 13/132 (9%)

Query: 407 AQLNSSQSNAVSHVLQRPLSLIQGPPGTGKTVTSATIV-YHLSKIHKDRILVCAPSNVAV 465
           AQL+  Q+  +  VL    SL+ G PGTGKT  +  ++   L++I   +I+V +P+  A
Sbjct: 135 AQLSEEQNQVLKTVLNSCFSLVCGGPGTGKTFLAVQMIRLILAQIPSAQIVVASPTGKAS 194

Query: 466 DHLAAKLRDLGLKVVRLTAKSREDVESSVSNLALHNLVGRGAKGELKNLLKLKDEVGELS 525
            HL + L   G+            V  SV  + +H  +    +G     L L DE   ++
Sbjct: 195 AHLHSVLTSQGI------------VGDSVEVVTIHKFLKDMRRGCSPVDLLLVDEGSMVT 242

Query: 526 ASDTKRFVKLVR 537
            +     +K +R
Sbjct: 243 INLLHGLIKTIR 254


 gb|AAD27694.1|AF125214_4 (AF125214) HelA [Helicobacter pylori]
           Length = 88

 Score = 41.8 bits (96), Expect = 0.024
 Identities = 20/42 (47%), Positives = 30/42 (70%), Gaps = 1/42 (2%)

Query: 565 KFRTVLIDESTQASEPECLIPIVKGAKQVILVGDHQQLGPVI 606
           +F   +IDE+ +A+ PE LIP ++  K++IL+GDH QL P I
Sbjct: 2   EFDVTIIDETGRATAPEILIPALR-TKKLILIGDHNQLPPSI 42


 pir||NCECXF exodeoxyribonuclease V (EC 3.1.11.5) 67K chain - Escherichia coli
           >gi|42691|emb|CAA28253| (X04582) exonuclease V alpha
           subunit (AA 1-608) [Escherichia coli]
           Length = 608

 Score = 40.6 bits (93), Expect = 0.054
 Identities = 31/95 (32%), Positives = 48/95 (49%), Gaps = 6/95 (6%)

Query: 413 QSNAVSHVLQRPLSLIQGPPGTGKTVTSATIVYHLSKI---HKDRILVCAPSNVAVDHLA 469
           Q  A +  L R +S+I G PGTGKT T A ++  L ++    + RI + AP+  A   L
Sbjct: 154 QKVAAAVALTRRISVISGGPGTGKTTTVAKLLAALIQMADGERCRIRLAAPTGKAAARLT 213

Query: 470 AKLRDLGLKVVRLTAKSREDVESSVSNLALHNLVG 504
             L    L+ + LT + ++ +    S   LH L+G
Sbjct: 214 ESL-GKALRQLPLTDEQKKRIPEDAS--TLHRLLG 245


 sp|P04993|EX5A_ECOLI EXODEOXYRIBONUCLEASE V 67 KD POLYPEPTIDE (EXONUCLEASE V ALPHA
           CHAIN) >gi|882711 (U29581) exonuclease V alpha-subunit
           [Escherichia coli] >gi|1789182 (AE000365) DNA helicase,
           ATP-dependent dsDNA/ssDNA exonuclease V subunit, ssDNA
           endonuclease [Escherichia coli]
           Length = 608

 Score = 40.6 bits (93), Expect = 0.054
 Identities = 31/95 (32%), Positives = 48/95 (49%), Gaps = 6/95 (6%)

Query: 413 QSNAVSHVLQRPLSLIQGPPGTGKTVTSATIVYHLSKI---HKDRILVCAPSNVAVDHLA 469
           Q  A +  L R +S+I G PGTGKT T A ++  L ++    + RI + AP+  A   L
Sbjct: 154 QKVAAAVALTRRISVISGGPGTGKTTTVAKLLAALIQMADGERCRIRLAAPTGKAAARLT 213

Query: 470 AKLRDLGLKVVRLTAKSREDVESSVSNLALHNLVG 504
             L    L+ + LT + ++ +    S   LH L+G
Sbjct: 214 ESL-GKALRQLPLTDEQKKRIPEDAS--TLHRLLG 245


 sp|P45158|EX5A_HAEIN EXODEOXYRIBONUCLEASE V ALPHA CHAIN >gi|1074067|pir||E64116
           exodeoxyribonuclease V (recD) homolog - Haemophilus
           influenzae (strain Rd KW20) >gi|1574782 (U32811)
           exodeoxyribonuclease V, alpha chain (recD) [Haemophilus
           influenzae Rd]
           Length = 640

 Score = 40.6 bits (93), Expect = 0.054
 Identities = 51/200 (25%), Positives = 81/200 (40%), Gaps = 12/200 (6%)

Query: 371 DKKSISGYLYYKILGHQVVDISFDVPLPK-EFSIPNFAQLNSSQSNAVSHVLQRPLSLIQ 429
           D+  I  Y+   +  ++ +  S+D    K E   P   +    Q  AV+  ++ P S+I
Sbjct: 134 DEYRIVQYIKNTLKKYRTLAFSYDEIHQKLEKYFPEKQEKTDWQKVAVATAIKSPFSIIT 193

Query: 430 GPPGTGKTVTSATIVYHLSKIH--KDRILVCAPSNVAVDHLAAKLRD-LGLKVVRLTAKS 486
           G PGTGKT T   ++  L ++   K  I + AP+  A   L   +++ LG    ++
Sbjct: 194 GGPGTGKTTTVTRLLLVLQELFDCKLHIKLVAPTGKAASRLEESIKNALGFMQEKMNVSH 253

Query: 487 REDVESSVSNLALHNLVGRGAKGEL-----KNLLKLKD-EVGELSASDTKRFVKLVR--K 538
                       LH+L+G  A  +       N L+L    V E S  D     KL+   K
Sbjct: 254 SLFNAIPQKASTLHSLLGVNAFNDYTRYNSHNPLQLDVLVVDETSMIDLPMMAKLINALK 313

Query: 539 TEAEILNKADVVCCTCVGAG 558
```

```
            E  ++   D      V AG
Sbjct: 314 PETRLILLGDQAQLASVEAG 333


 sp|P47386|Y140_MYCGE HYPOTHETICAL ATP-BINDING PROTEIN MG140 >gi|1361625|pir||E64215
           hypothetical protein MG140 - Mycoplasma genitalium
           (SGC3) >gi|3844733 (U39695) conserved hypothetical
           protein [Mycoplasma genitalium]
           Length = 1113

 Score = 38.7 bits (88), Expect = 0.21
 Identities = 69/296 (23%), Positives = 117/296 (39%), Gaps = 57/296 (19%)

Query: 569 VLIDESTQASEPECLIPIVKGAKQVILVGDHQQLGPVILERKAADAGLKQ---------- 618
           V+IDE++Q    E  IPI+  A + I+ GD +QL P    +  A+  + +
Sbjct: 699 VVIDEASQVFL-ERAIPILFRADKYIIAGDTKQLKPANFFQSRAEYDVDEEFEDGNIEAA 757

Query: 619 ----SLFERLISLGHVPIRLEVQYRMNPY-LSEFPSNMFYEGSL--QNGVTIEQRTVPNS 671
               SL   L +   +   L+  YR +   L  F +N  Y+  L   N    +QR
Sbjct: 758 VHSSSLLHFLKNRSRILTLLKFHYRSDSADLIAFTNNRIYDNELIFMNKANADQRV---- 813

Query: 672 KFPWPIRGIPMMFWANYGREEISANGTSFLNRIEAMNCERIITKLFRDGVKPEQIGVITP 731
              + +  +    W N              N  EA +  + + +L       + +GVI
Sbjct: 814 ---FIVHDVIDGIWKNN------------RNLQEARDVVQRLEQLTTTNDYKKSLGVIC- 857

Query: 732 YEGQRAYILQYM--QMNGSL---------DKDLYIKVEVASVDAFQGREKDYIILSC--- 777
           +   +A +++Y+  + N  L         D   Y  + V +++  QG E+D II S
Sbjct: 858 FNKNQADLIEYLIDKQNNPLLNEWRERQNDVGEYEGLFVKNIENVQGDERDIIIFSLGYD 917

Query: 778 VRANEQQAIGFLRDPRRLNVGLTRAKYGLVILGNPR-----SLARNTLWNHLLIHF 828
              N    I       RLNV +TRAK  + +    R      L+ ++L + LL+ +
Sbjct: 918 RSVNSYGPISKQGGENRLNVAITRAKQRIELFKTNRGEDYNGLSSSSLGSKLLVEY 973


 Score = 33.2 bits (74), Expect = 9.2
 Identities = 21/61 (34%), Positives = 34/61 (55%), Gaps = 2/61 (3%)

Query: 406 FAQLNSSQSNAVSHVLQRPLSLIQGPPGTGKTVTSATIVYHLSKIHKDRILVCAPSNVAV 465
           F+ L+  Q  AV   L   L +I+GPPGTGK+ T   I+ +L  ++  ++L  +    A+
Sbjct: 290 FSHLDIYQQYAVKQALLGDL-IIEGPPGTGKSETIVNILVNL-VLNNKKVLFVSEKVTAL 347

Query: 466 D 466
           D
Sbjct: 348 D 348


 sp|P16342|RRPB_CVMA5 RNA-DIRECTED RNA POLYMERASE (ORF1B) >gi|93916|pir||S15760
            RNA-directed RNA polymerase (EC 2.7.7.48) - murine
            hepatitis virus (strain A59)
            Length = 2733

 Score = 38.3 bits (87), Expect = 0.28
 Identities = 33/132 (25%), Positives = 55/132 (41%), Gaps = 32/132 (24%)

Query: 365  LKKFAIDKKSISGYLYYK--------------ILGHQVVDISFDVPLPKE---------- 400
            L ++  DK  ++  +YY+              +  H V  +S    +P+E
Sbjct: 1134 LGEYVFDKSELTNGVYYRATTTYKLSVGDVFILTSHAVSSLSAPTLVPQENYTSIRFASV 1193

Query: 401  FSIPNFAQLNSSQSNAVSHVLQRPLSLIQGPPGTGKT-VTSATIVYHLSKIHKDRILVCA 459
            +S+P   Q N        H+  +    +QGPPGTGK+ +     VY+ +     R++  A
Sbjct: 1194 YSVPETFQNNVPN---YQHIGMKRYCTVQGPPGTGKSHLAIGLAVYYCTA----RVVYTA 1246

Query: 460  PSNVAVDHLAAK 471
             S+ AVD L  K
Sbjct: 1247 ASHAAVDALCEK 1258


 gi|2621557 (AE000832) DNA helicase related protein [Methanobacterium
           thermoautotrophicum]
           Length = 1157

 Score = 38.3 bits (87), Expect = 0.28
 Identities = 62/285 (21%), Positives = 115/285 (39%), Gaps = 57/285 (20%)

Query: 565 KFRTVLIDESTQASEPECLIPIVKGAKQVILVGDHQQLGPVILERKAADAGLKQSLFERL 624
           +F  V+ DE++Q  +PE  +     A+  +++GD  QL P              S F+++
Sbjct: 745 QFDVVIFDEASQV-KPEDALGAFLRARTAVVMGDTNQLPPT-------------SFFDQM 790

Query: 625 ISLGH----VPIRLEVQYRMNPYLSEFPSNMF---YEGSLQNGVTIEQRTVPNSK---FP 674
           IS       V    +++  ++     FP  M    Y    ++ + +  +   +++   +P
Sbjct: 791 ISAEEDSEDVATAADIESILHLCKRSFPVRMLRWHYRSRHESLIAVSNQEFYDNRLLVYP 850

Query: 675 WPIR-----GIPMMFWANYGREEISANGTSFLNRIEAMNCERIITKLFRDGVKPEQIGVI 729
            P R     G+ ++    Y  + +   G +  N +EA      I + F        +GV
Sbjct: 851 SPAREDEELGLHLV----YLPDTVYERGRTSSNPLEAAAVVDAIEEHFMRYGSSRSLGVG 906

Query: 730 TPYEGQRAYILQYMQ------------MNGSLDKDLYIKVEVASVDAFQGREKDYIILSC 777
           T    Q   IL+ ++            +N   D+  +IK    +++  QG E+D I +S
Sbjct: 907 TFSVAQMNAILEALEARLRENPELERIINQETDEPFFIK----NLETIQGDERDVIFISV 962
```

```
Query: 778 VRANEQQAIGFLR--------DPRRLNVGLTRAKYGLVILGNPRS 814
               +++    L           RRLNV +TRA+   V+  N RS
Sbjct: 963 GYGFDERGRMSLNFGPLNQEGGERRLNVLITRAREKCVVFTNFRS 1007
```

```
 Score = 36.4 bits (82), Expect = 1.1
 Identities = 30/119 (25%), Positives = 55/119 (46%), Gaps = 6/119 (5%)

Query: 410 NSSQSNAVSHVLQRPLSLIQGPPGTGKTVTSATIVYHLSKIHKDRILVCAPSNVAVDHLA 469
           +SSQ   +  V      +++GPPGTGK+ T   ++  L    K  +L  +    A++ +
Sbjct: 96  DSSQIAVIEDVKAGKNLVVEGPPGTGKSQTIVNLIAELMASGK-TVLFVSEKMAALEVVK 154

Query: 470 AKLRDLGLKVVRLTAKSREDVESSVSNLALHNLVGR-----GAKGELKNLLKLKDEVGE 523
           ++L  +GL    L   S +  +  V N     L+        A+ E + + +L+DE+ E
Sbjct: 155 SRLDSIGLGRFCLELHSHKARKKDVLNELEATLMEAEADRPEAEREFRRMERLRDELNE 213
```

```
 gi|2641128 (AF029248) RNA-directed RNA polymerase [Murine hepatitis virus]
            Length = 2733

 Score = 38.3 bits (87), Expect = 0.28
 Identities = 33/132 (25%), Positives = 55/132 (41%), Gaps = 32/132 (24%)

Query: 365  LKKFAIDKKSISGYLYYK--------------ILGHQVVDISFDVPLPKE---------- 400
            L ++  DK  ++  +YY+               +  H V  +S    +P+E
Sbjct: 1134 LGEYVFDKSELTNGVYYRATTTYKLSVGDVFILTSHAVSSLSAPTLVPQENYTSIRFASV 1193

Query: 401  FSIPNFAQLNSSQSNAVSHVLQRPLSLIQGPPGTGKT-VTSATIVYHLSKIHKDRILVCA 459
            +S+P   Q N        H+  +    +QGPPGTGK+ +     VY+ +     R++  A
Sbjct: 1194 YSVPETFQNNVPN---YQHIGMKRYCTVQGPPGTGKSHLAIGLAVYYCTA----RVVYTA 1246

Query: 460  PSNVAVDHLAAK 471
             S+ AVD L  K
Sbjct: 1247 ASHAAVDALCEK 1258
```

```
 emb|CAA36202| (X51939) open reading frame 1b (AA 1-2733) [murine hepatitis virus]
            Length = 2733

 Score = 38.3 bits (87), Expect = 0.28
 Identities = 33/132 (25%), Positives = 55/132 (41%), Gaps = 32/132 (24%)

Query: 365  LKKFAIDKKSISGYLYYK--------------ILGHQVVDISFDVPLPKE---------- 400
            L ++  DK  ++  +YY+               +  H V  +S    +P+E
Sbjct: 1134 LGEYVFDKSELTNGVYYRATTTYKLSVGDVFILTSHAVSSLSAPTLVPQENYTSIRFASV 1193

Query: 401  FSIPNFAQLNSSQSNAVSHVLQRPLSLIQGPPGTGKT-VTSATIVYHLSKIHKDRILVCA 459
            +S+P   Q N        H+  +    +QGPPGTGK+ +     VY+ +     R++  A
Sbjct: 1194 YSVPETFQNNVPN---YQHIGMKRYCTVQGPPGTGKSHLAIGLAVYYCTA----RVVYTA 1246

Query: 460  PSNVAVDHLAAK 471
             S+ AVD L  K
Sbjct: 1247 ASHAAVDALCEK 1258
```

```
 pir||S52928 XSUG1 protein - African clawed frog
           Length = 283

 Score = 37.5 bits (85), Expect = 0.47
 Identities = 36/141 (25%), Positives = 59/141 (41%), Gaps = 20/141 (14%)

Query: 370 IDKKSISGYLYYKILGHQVVDISFDVPLPKEFSIPNFAQLNSSQSNAVSHVLQRPLSLIQ 429
           ++K   S Y     L  Q+ +I   + LP +    +F  L  +Q   V         L+
Sbjct: 135 VEKVPDSTYEMIGGLDKQIKEIKEVIELPVKHP-EHFEALGIAQPKGV---------LLY 184

Query: 430 GPPGTGKTVTSATIVYHLSKIHKDRILVCAPSNVAVDHLAAKLRDLGLKVVR-LTAKSRE 488
           GPPGTGKT+ +  + +H           C    V+   L  K    G ++VR L   +RE
Sbjct: 185 GPPGTGKTLLARAVAHHTD---------CTFIRVSGSELVQKFIGEGARMVRELFVMARE 235

Query: 489 DVESSVSNLALHNLVGRGAKG 509
              S +    + ++  RG++G
Sbjct: 236 HAPSIIFMDEIDSIGSRGSRG 256
```

```
 sp|P39369|YJHR_ECOLI HYPOTHETICAL 38.0 KD PROTEIN IN FECI-FIMB INTERGENIC REGION (O338)
           >gi|1361172|pir||S56533 hypothetical protein o338 -
           Escherichia coli >gi|537149 (U14003) ORF_o338
           [Escherichia coli] >gi|1790762 (AE000501) putative
           frameshift suppressor [Escherichia coli]
           Length = 338

 Score = 37.1 bits (84), Expect = 0.62
 Identities = 24/89 (26%), Positives = 47/89 (51%), Gaps = 6/89 (6%)

Query: 726 IGVITPYEGQRAYI---LQYMQMNGSLDKDLYIKVEVASVDAFQGREKDYIILSCVRANE 782
           +GV+TP+  Q   I   L+ +++NG  ++ L   + V +V + QG E+  ++ S V +
Sbjct: 50  VGVVTPFSAQVNAIKMSLRKLEINGKDEQGL---LTVGTVHSLQGAERAIVLFSPVYSKH 106
```

```
Query: 783 QQAIGFLRDPRRLNVGLTRAKYGLVILGN 811
           +       +   LNV ++RAK   ++ G+
Sbjct: 107 EDGRFLDSNSTILNVAVSRAKDSFLVFGD 135
```

gi|3790757 (AF099922) similar to the AAA family of ATPases; most similar to 26S protease regulatory subunit 8 [Caenorhabditis elegans]
Length = 443

Score = 37.1 bits (84), Expect = 0.62
Identities = 39/159 (24%), Positives = 69/159 (42%), Gaps = 27/159 (16%)

```
Query: 427 LIQGPPGTGKTVTSATIVYHLSKIHKDRILVCAPSNVAVDHLAAKLRDLGLKVVR-LTAK 485
           L+ GPPGTGKT+ +  + +H           C    V+   L  K    G ++VR L
Sbjct: 224 LLYGPPGTGKTLLARAVAHHTE---------CTFIRVSGSELVQKFIGEGARMVRELFVM 274

Query: 486 SREDVES-----SVSNLALHNLVG-RGAKGEL-KNLLKLKDEVGELSASDTKRFVKLVRK 538
           +RE   S      + ++    + G RG   E+ + +L+L   + +L   +  + +K++
Sbjct: 275 AREHAPSIIFMDEIDSIGSSRVEGSRGGDSEVQRTMLEL---LNQLDGFEATKNIKVIMA 331

Query: 539 TEAEILNKADVVCCTCVGAGDKRLDTKFRTVLIDESTQA 577
           T     N+ D++    +  G  R+D K      DE  +A
Sbjct: 332 T-----NRIDILDSALLRPG--RIDRKIEFPAPDEKARA 363
```

gb|AAD18890| (AE001657) Exodeoxyribonuclease V, Alpha [Chlamydia pneumoniae]
Length = 493

Score = 37.1 bits (84), Expect = 0.62
Identities = 20/61 (32%), Positives = 33/61 (53%), Gaps = 1/61 (1%)

```
Query: 409 LNSSQSNAVSHVLQRPLSLIQGPPGTGKTVTSATIVYHLSKIH-KDRILVCAPSNVAVDH 467
           L+  Q+   + + Q   S++ G PGTGKT  +A ++  L K   K RI + +P+  A  H
Sbjct: 133 LSEEQNFIFNKITQGCFSIVSGGPGTGKTFLAAQLILSLVKQQPKLRIAIVSPTGKATSH 192

Query: 468 L 468
           +
Sbjct: 193 I 193
```

sp|P29982|RRPB_CVMJH RNA-DIRECTED RNA POLYMERASE (ORF1B) >gi|74827|pir||VFIHJH RNA-directed RNA polymerase (EC 2.7.7.48) 1b - murine hepatitis virus (strain JHM) >gi|331853 (M55148) open reading frame 1b [Murine hepatitis virus]
Length = 2731

Score = 36.7 bits (83), Expect = 0.81
Identities = 33/131 (25%), Positives = 52/131 (39%), Gaps = 30/131 (22%)

```
Query: 365  LKKFAIDKKSISGYLYYK--------------ILGHQVVDISFDVPLPKE---------- 400
            L ++  DK  ++  +YY+              +  H V  +S    +P+E
Sbjct: 1134 LGEYVFDKSELTNGVYYRATTTYKLSVGDVFILTSHAVSSLSAPTLVPQENYTSVRFASA 1193

Query: 401  FSIPNFAQLNSSQSNAVSHVLQRPLSLIQGPPGTGKTVTSATIVYHLSKIHKDRILVCAP 460
            +S+P   Q N        H+  +    +QGPPGTGK   S   + H       R++  A
Sbjct: 1194 YSVPETFQNNVPN---YQHIGIKRYCTVQGPPGTGK---SHLAIGHAVYYCTARVVYTAA 1247

Query: 461  SNVAVDHLAAK 471
            S+ AVD L  K
Sbjct: 1248 SHAAVDALCEK 1258
```

sp|P46470|PRS8_XENLA 26S PROTEASE REGULATORY SUBUNIT 8 (SUG1 HOMOLOG) (XSUG1) >gi|1877414|emb|CAA57512| (X81986) XSUG1 [Xenopus laevis]
Length = 461

Score = 36.7 bits (83), Expect = 0.81
Identities = 45/200 (22%), Positives = 83/200 (41%), Gaps = 30/200 (15%)

```
Query: 370 IDKKSISGYLYYKILGHQVVDISFDVPLPKEFSIPNFAQLNSSQSNAVSHVLQRPLSLIQ 429
           ++K   S Y     L  Q+ +I   + LP +    +F  L  +Q   V         L+
Sbjct: 135 VEKVPDSTYEMIGGLDKQIKEIKEVIELPVKHP-EHFEALGIAQPKGV---------LLY 184

Query: 430 GPPGTGKTVTSATIVYHLSKIHKDRILVCAPSNVAVDHLAAKLRDLGLKVVR-LTAKSRE 488
           GPPGTGKT+ +  + +H           C    V+   L  K    G ++VR L   +RE
Sbjct: 185 GPPGTGKTLLARAVAHHTD---------CTFIRVSGSELVQKFIGEGARMVRELFVMARE 235

Query: 489 DVESSVSNLALHNLVGR---GAKGELKNLLKLKDEVGELSASDTKRFVKLVRKTEAEILN 545
               S +    + ++  R   G+ G+ +    + + + +L   +  + +K++  T     N
Sbjct: 236 HAPSIIFMDEIDSIGSRLEGGSGGDSEVQRTMLELLNQLDGFEATKNIKVIMAT-----N 290

Query: 546 KADVVCCTCVGAGDKRLDTK 565
           + D++    +  G  R+D K
Sbjct: 291 RIDILDSALLRPG--RIDRK 308
```

emb|CAB14689| (Z99117) similar to conjugation transfer protein [Bacillus

```
           subtilis] >gi|2635211|emb|CAB14706| (Z99118) similar to
           conjugation transfer protein [Bacillus subtilis]
           Length = 798

 Score = 36.4 bits (82), Expect = 1.1
 Identities = 31/107 (28%), Positives = 46/107 (42%), Gaps = 16/107 (14%)

Query: 408 QLNSSQSNAVSHVLQRPLSLIQGPPGTGKTVTSATIVYHLSKIHKDRILVCAPSNVAVDH 467
           Q   SQ  A+   L  P+ L+ G PGTGKT     IV    ++H           V++D
Sbjct: 345 QYAPSQKEAIQKALSSPMLLLTGGPGTGKTTVIRGIVELYGELH----------GVSLDP 394

Query: 468 LAAKLRDLGLKVVRLT----AKSREDVESSVSNLALHNLVG-RGAKG 509
            A K +D    +V       A  R    + +  + +H L+G  GA+G
Sbjct: 395 SAYK-KDEAFPIVLAAPTGRAAKRMSESTGLPAVTIHRLLGWNGAEG 440


 dbj|BAA76763.1| (AB023136) KIAA0919 protein [Homo sapiens]
           Length = 710

 Score = 36.4 bits (82), Expect = 1.1
 Identities = 34/140 (24%), Positives = 63/140 (44%), Gaps = 18/140 (12%)

Query: 453 DRILVCAPSNVAVDHLAAKLRD-LGLKVVRLTAKSREDVESSVSNLALHNLVGRGAKGEL 511
           D++++C P    V  + +KLRD +  K      K+ E +E +      H    R  K  +
Sbjct: 170 DKLMLCLP----VLEMYSKLRDQMKTKRHYPALKTLEHLEHTYLPQVSHY---RFCKVMV 222

Query: 512 KNLLKLKDEVGELSASDTKRFVKLVRK----------TEAEILNKADVVCCTCVGAGDKR 561
            N+ KL++E+ ++S SD K F++ +RK           +A+     D +       G KR
Sbjct: 223 DNIPKLREEIKDVSMSDLKDFLESIRKHSDKIGETAMKQAQQQRNLDNIVLQQPRIGSKR 282

Query: 562 LDTKFRTVLIDESTQASEPE 581
              K   ++ D   +++ P+
Sbjct: 283 KSKKDAYIIFDTEIESTSPK 302


 emb|CAB07118| (Z92772) recD [Mycobacterium tuberculosis]
           Length = 575

 Score = 36.0 bits (81), Expect = 1.4
 Identities = 41/160 (25%), Positives = 69/160 (42%), Gaps = 19/160 (11%)

Query: 413 QSNAVSHVLQRPLSLIQGPPGTGKTVTSATIV------YHLSKIHKDRILVCAPSNVAVD 466
           Q  A    L + ++++ G PGTGKT T A ++        L+   + RI + AP+  A
Sbjct: 153 QRRAAEIALSQGVTVLTGGPGTGKTTTVARLLALVAEQAELAGEPRPRIALAAPTGKAAA 212

Query: 467 HLAAKLRDLGLKVVRLTAKSREDVESSVSNLALHNLVG--------RGAKGELKNLLKLK 518
            LA  +R    ++ +L A  R  +   +  + LH L+G        +  +  L + + +
Sbjct: 213 RLAEAVR---REMAKLDATDRARL-GDLHAVTLHRLLGAKPGARFRQDRQNRLPHNVIVV 268

Query: 519 DEVGELSASDTKRFVKLVRKTEAEIL-NKADVVCCTCVGA 557
           DE   +S +   R  + VR     IL   AD +     GA
Sbjct: 269 DETSMVSLTLMARLAEAVRPGARLILVGDADQLASVEAGA 308


 sp|O50581|RECG_STAAU ATP-DEPENDENT DNA HELICASE RECG >gi|2826896|dbj|BAA24572|
           (AB000439) RecG [Staphylococcus aureus]
           Length = 686

 Score = 36.0 bits (81), Expect = 1.4
 Identities = 33/133 (24%), Positives = 60/133 (44%), Gaps = 12/133 (9%)

Query: 387 QVVDISFDVPLPKEFSIPNFAQLNSSQSNAVSHV---LQRPLS---LIQGPPGTGKTVTS 440
           + ++I +D+   K F      +L  +Q ++V+ +   L+ P+    L+QG  G+GKTV +
Sbjct: 243 EAIEIDYDIDQVKSFIDRLPFELTEAQKSSVNEIFRDLKAPIRMHRLLQGDVGSGKTVVA 302

Query: 441 ATIVYHLSKIHKDRILVCAPSNVAVDHLAAKLRDLG--LKVVRLT----AKSREDVESSV 494
           A  +Y L        L+     +A  H  + +   G  + V  LT     K R+ +   +
Sbjct: 303 AICMYALKTAGYQSALMVPTEILAEQHAESLMALFGDSMNVALLTGSVKGKKRKILLEQL 362

Query: 495 SNLALHNLVGRGA 507
            N  +  L+G  A
Sbjct: 363 ENGTIDCLIGTHA 375


 gb|AAD41422.1|AC007727_11 (AC007727) Similar to gb|M87339 replication factor C, 37-kDa
           subunit from Homo sapiens and is a member of PF|00004
           ATPases associated with various cellular activities.
           [Arabidopsis thaliana]
           Length = 319

 Score = 35.6 bits (80), Expect = 1.8
 Identities = 20/60 (33%), Positives = 33/60 (54%), Gaps = 3/60 (5%)

Query: 424 PLSLIQGPPGTGKTVTSATIVYHL--SKIHKDRIL-VCAPSNVAVDHLAAKLRDLGLKVV 480
           P  L  GPPGTGKT T+  I + L   +++K R+L + A  +  ++ +  K++D     V
Sbjct: 43  PHMLFYGPPGTGKTTTALAIAHQLFGPELYKSRVLELNASDDRGINVVRTKIKDFAAVAV 102


 sp|P29569|YPV1_METTF HYPOTHETICAL 40.7 KD PROTEIN (ORF1) >gi|282612|pir||S26448
```

hypothetical protein 1 - Methanobacterium
thermoformicicum plasmid pFV1 >gi|44645|emb|CAA48437|
(X68366) orf1 [Methanobacterium thermoautotrophicum]
Length = 364

Score = 35.2 bits (79), Expect = 2.4
Identities = 28/96 (29%), Positives = 42/96 (43%), Gaps = 10/96 (10%)

```
Query: 403 IPNFAQLNSSQSNAVSHVL-------QRPLSLIQGPPGTGKTVTSATIVYHLSKIHKDRI 455
           +P+  Q    +  A+S  L         P  LI GPPG+GKTVT+  ++  L K   D +
Sbjct: 20  VPDTLQDRKEEVGAISQYLGYILDGATPPHLLIVGPPGSGKTVTTKYVINELEKHTSDAV 79

Query: 456 ---LVCAPSNVAVDHLAAKLRDLGLKVVRLTAKSRE 488
              +V   +   V    A+    GL  + +  K RE
Sbjct: 80  IEYIVADGTAYQVATSIARAPRRGLGFLNIVEKIRE 115
```

sp|P29570|YPZ1_METTF HYPOTHETICAL 40.6 KD PROTEIN (ORF1') >gi|282613|pir||S26458
hypothetical protein 1' - Methanobacterium
thermoformicicum plasmid pFZ1 >gi|44656|emb|CAA48438|
(X68367) ORF1 [Methanobacterium thermoautotrophicum]
Length = 364

Score = 35.2 bits (79), Expect = 2.4
Identities = 28/96 (29%), Positives = 42/96 (43%), Gaps = 10/96 (10%)

```
Query: 403 IPNFAQLNSSQSNAVSHVL-------QRPLSLIQGPPGTGKTVTSATIVYHLSKIHKDRI 455
           +P+  Q    +  A+S  L         P  LI GPPG+GKTVT+  ++  L K   D +
Sbjct: 20  VPDTLQDRKEEVGAISQYLGYILDGATPPHLLIVGPPGSGKTVTTKYVINELEKHTSDAV 79

Query: 456 ---LVCAPSNVAVDHLAAKLRDLGLKVVRLTAKSRE 488
              +V   +   V    A+    GL  + +  K RE
Sbjct: 80  IEYIVADGTAYQVATSIARAPRRGLGFLNIVEKIRE 115
```

sp|Q04913|NGFI_XENLA NERVE GROWTH FACTOR INDUCED PROTEIN I-B HOMOLOG
>gi|479404|pir||S33763 hormone receptor NGFI-B homolog -
African clawed frog >gi|64917|emb|CAA50031| (X70700)
NGFI-B [Xenopus laevis]
Length = 577

Score = 35.2 bits (79), Expect = 2.4
Identities = 18/54 (33%), Positives = 31/54 (57%), Gaps = 5/54 (9%)

```
Query: 37  VDNQLFEEAQVTETGFRSPSASDNSCAYCGIDSAKC----VIKCNSCKKWFCNT 86
           +D+ +  ++ ++ +  RSPS+++  CA CG D+A C    V  C  CK +F  T
Sbjct: 221 LDSSVLLDSPLSPSKTRSPSSNEGRCAVCG-DNASCQHYGVRTCEGCKGFFKRT 273
```

sp|P47210|PRS8_HUMAN 26S PROTEASE REGULATORY SUBUNIT 8 (PROTEASOME SUBUNIT P45) (THYROID
HORMONE RECEPTOR INTERACTING PROTEIN 1) (TRIP1)
>gi|2136007|pir||I53510 proteasome subunit p45 - human
>gi|976227|dbj|BAA07919| (D44467) 26S proteasome subunit
p45 [Homo sapiens] >gi|1096205|prf||2111282A 26S
proteasome [Homo sapiens]
Length = 406

Score = 35.2 bits (79), Expect = 2.4
Identities = 34/145 (23%), Positives = 63/145 (43%), Gaps = 23/145 (15%)

```
Query: 427 LIQGPPGTGKTVTSATIVYHLSKIHKDRILVCAPSNVAVDHLAAKLRDLGLKVVR-LTAK 485
           L+ GPPGTGKT+ +  + +H           C    V+   L  K    G ++VR L
Sbjct: 187 LLYGPPGTGKTLLARAVAHHTD---------CTFIRVSGSELVQKFIGEGARMVRELFVM 237

Query: 486 SREDVES-----SVSNLALHNLVGRGAKGELKNLLKLKDEVGELSASDTKRFVKLVRKTE 540
           +RE   S      + ++    L G G+ G+ +    + + + +L   +  + +K++  T
Sbjct: 238 AREHAPSIIFMDEIDSIGSSRLEG-GSGGDSEVQRTMLELLNQLDGFEATKNIKVIMAT- 295

Query: 541 AEILNKADVVCCTCVGAGDKRLDTK 565
               N+ D++    +  G  R+D K
Sbjct: 296 ----NRIDILDSALLRPG--RIDRK 314
```

sp|P52915|PRS8_MOUSE 26S PROTEASE REGULATORY SUBUNIT 8 (MSUG1 PROTEIN) (TAT-BINDING
PROTEIN HOMOLOG 10) (TBP10) (P45/SUG)
>gi|2137798|pir||S61923 SUG1 protein - mouse
>gi|1165125|emb|CAA90961.1| (Z54219) mSUG1 protein [Mus
musculus] >gi|1262433|emb|CAA61863| (X89718) 26S
protease subunit [Sus scrofa] >gi|1395177|dbj|BAA11938|
(D83521) proteasomal ATPase (rat SUG1) [Rattus
norvegicus] >gi|2564003|dbj|BAA22933| (AB000491)
proteasome p45/SUG [Rattus norvegicus] >gi|3193258
(AF069053) proteasome subunit SUG1 [Bos taurus]
Length = 406

Score = 35.2 bits (79), Expect = 2.4
Identities = 34/145 (23%), Positives = 63/145 (43%), Gaps = 23/145 (15%)

```
Query: 427 LIQGPPGTGKTVTSATIVYHLSKIHKDRILVCAPSNVAVDHLAAKLRDLGLKVVR-LTAK 485
```

```
           L+ GPPGTGKT+ +  + +H           C    V+   L  K    G ++VR L
Sbjct: 187 LLYGPPGTGKTLLARAVAHHTD---------CTFIRVSGSELVQKFIGEGARMVRELFVM 237

Query: 486 SREDVES-----SVSNLALHNLVGRGAKGELKNLLKLKDEVGELSASDTKRFVKLVRKTE 540
           +RE   S      + ++    L G G+ G+ +    + + + +L   +  + +K++  T
Sbjct: 238 AREHAPSIIFMDEIDSIGSSRLEG-GSGGDSEVQRTMLELLNQLDGFEATKNIKVIMAT- 295

Query: 541 AEILNKADVVCCTCVGAGDKRLDTK 565
               N+ D++    +  G  R+D K
Sbjct: 296 ----NRIDILDSALLRPG--RIDRK 314


 emb|CAA61864| (X89719) put. 26S protease subunit [Sus scrofa]
           Length = 398

 Score = 35.2 bits (79), Expect = 2.4
 Identities = 34/145 (23%), Positives = 63/145 (43%), Gaps = 23/145 (15%)

Query: 427 LIQGPPGTGKTVTSATIVYHLSKIHKDRILVCAPSNVAVDHLAAKLRDLGLKVVR-LTAK 485
           L+ GPPGTGKT+ +  + +H           C    V+   L  K    G ++VR L
Sbjct: 179 LLYGPPGTGKTLLARAVAHHTD---------CTFIRVSGSELVQKFIGEGARMVRELFVM 229

Query: 486 SREDVES-----SVSNLALHNLVGRGAKGELKNLLKLKDEVGELSASDTKRFVKLVRKTE 540
           +RE   S      + ++    L G G+ G+ +    + + + +L   +  + +K++  T
Sbjct: 230 AREHAPSIIFMDEIDSIGSSRLEG-GSGGDSEVQRTMLELLNQLDGFEATKNIKVIMAT- 287

Query: 541 AEILNKADVVCCTCVGAGDKRLDTK 565
               N+ D++    +  G  R+D K
Sbjct: 288 ----NRIDILDSALLRPG--RIDRK 306


 dbj|BAA22935| (AB000493) proteasome p45/SUG [Rattus norvegicus]
           Length = 374

 Score = 35.2 bits (79), Expect = 2.4
 Identities = 34/145 (23%), Positives = 63/145 (43%), Gaps = 23/145 (15%)

Query: 427 LIQGPPGTGKTVTSATIVYHLSKIHKDRILVCAPSNVAVDHLAAKLRDLGLKVVR-LTAK 485
           L+ GPPGTGKT+ +  + +H           C    V+   L  K    G ++VR L
Sbjct: 155 LLYGPPGTGKTLLARAVAHHTD---------CTFIRVSGSELVQKFIGEGARMVRELFVM 205

Query: 486 SREDVES-----SVSNLALHNLVGRGAKGELKNLLKLKDEVGELSASDTKRFVKLVRKTE 540
           +RE   S      + ++    L G G+ G+ +    + + + +L   +  + +K++  T
Sbjct: 206 AREHAPSIIFMDEIDSIGSSRLEG-GSGGDSEVQRTMLELLNQLDGFEATKNIKVIMAT- 263

Query: 541 AEILNKADVVCCTCVGAGDKRLDTK 565
               N+ D++    +  G  R+D K
Sbjct: 264 ----NRIDILDSALLRPG--RIDRK 282


 gi|2661071 (AF035309) similar to 26S proteasome subunit p45 [Homo sapiens]
           Length = 280

 Score = 35.2 bits (79), Expect = 2.4
 Identities = 34/145 (23%), Positives = 63/145 (43%), Gaps = 23/145 (15%)

Query: 427 LIQGPPGTGKTVTSATIVYHLSKIHKDRILVCAPSNVAVDHLAAKLRDLGLKVVR-LTAK 485
           L+ GPPGTGKT+ +  + +H           C    V+   L  K    G ++VR L
Sbjct: 61  LLYGPPGTGKTLLARAVAHHTD---------CTFIRVSGSELVQKFIGEGARMVRELFVM 111

Query: 486 SREDVES-----SVSNLALHNLVGRGAKGELKNLLKLKDEVGELSASDTKRFVKLVRKTE 540
           +RE   S      + ++    L G G+ G+ +    + + + +L   +  + +K++  T
Sbjct: 112 AREHAPSIIFMDEIDSIGSSRLEG-GSGGDSEVQRTMLELLNQLDGFEATKNIKVIMAT- 169

Query: 541 AEILNKADVVCCTCVGAGDKRLDTK 565
               N+ D++    +  G  R+D K
Sbjct: 170 ----NRIDILDSALLRPG--RIDRK 188


 pir||S47220 protein kinase C - yeast (Candida albicans)
           Length = 1097

 Score = 34.8 bits (78), Expect = 3.1
 Identities = 16/52 (30%), Positives = 31/52 (58%), Gaps = 1/52 (1%)

Query: 918 NFKSAFSQKQNRNEIDDRNLYQEEASHLNSNFARELQREEQKHELSK-DFSN 968
           +F++ + Q+Q+++  D   +  E+  H NSN  R+++ EE K E    D++N
Sbjct: 644 SFETGYGQQQHQHHRDVPQIVVEDHQHYNSNDNRDVEMEESKDEFDNFDYNN 695


 gi|393034 (U01065) nonstructural polyprotein [Western equine
           encephalomyelitis virus]
           Length = 1490

 Score = 34.8 bits (78), Expect = 3.1
 Identities = 32/121 (26%), Positives = 52/121 (42%), Gaps = 8/121 (6%)

Query: 485 KSREDVESSVSNLALHNLVGRGAKGELKNLLKLKDEVGELSASDTKRFVKLVRKTEAEIL 544
           K+R      V  + ++ + G G  G +K+ +  KD V      +    ++ VR+     +
```

```
Sbjct: 267 KTRPAAPHKVPTIGVYGVPGSGKSGIIKSAVTKKDLVVSAKKENCAEIIRDVRR-----M 321

Query: 545 NKADVVCCTCVGAGDKRLDTKFRTVLIDE--STQASEPECLIPIVKGAKQVILVGDHQQL 602
            + DV   T        +     T+ IDE  +  A     LI IVK  K+V+L GD +Q
Sbjct: 322 RRMDVAARTVDSVLLNGVKHPVNTLYIDEAFACHAGTLLALIAIVK-PKKVVLCGDPKQC 380

Query: 603 G 603
           G
Sbjct: 381 G 381
```

emb|CAA52868| (X74892) NSP1, NSP2, NSP3 [Western equine encephalomyelitis virus]
Length = 1417

Score = 34.8 bits (78), Expect = 3.1
Identities = 32/121 (26%), Positives = 52/121 (42%), Gaps = 8/121 (6%)

```
Query: 485 KSREDVESSVSNLALHNLVGRGAKGELKNLLKLKDEVGELSASDTKRFVKLVRKTEAEIL 544
           K+R      V  + ++ + G G  G +K+ +  KD V      +    ++ VR+     +
Sbjct: 267 KTRPAAPHKVPTIGVYGVPGSGKSGIIKSAVTKKDLVVSAKKENCAEIIRDVRR-----M 321

Query: 545 NKADVVCCTCVGAGDKRLDTKFRTVLIDE--STQASEPECLIPIVKGAKQVILVGDHQQL 602
            + DV   T        +     T+ IDE  +  A     LI IVK  K+V+L GD +Q
Sbjct: 322 RRMDVAARTVDSVLLNGVKHPVNTLYIDEAFACHAGTLLALIAIVK-PKKVVLCGDPKQC 380

Query: 603 G 603
           G
Sbjct: 381 G 381
```

sp|P43057|KPC1_CANAL PROTEIN KINASE C-LIKE 1 (PKC 1) >gi|832908|emb|CAA57048| (X81142)
protein kinase C [Candida albicans]
Length = 1097

Score = 34.8 bits (78), Expect = 3.1
Identities = 16/52 (30%), Positives = 31/52 (58%), Gaps = 1/52 (1%)

```
Query: 918 NFKSAFSQKQNRNEIDDRNLYQEEASHLNSNFARELQREEQKHELSK-DFSN 968
           +F++ + Q+Q+++  D   +  E+  H NSN  R+++ EE K E    D++N
Sbjct: 644 SFETGYGQQQHQHHRDVPQIVVEDHQHYNSNDNRDVEMEESKDEFDNFDYNN 695
```

pir||S57222 NSP1-NSP2-NSP3 polypeptide - western equine encephalomyelitis virus
(fragment)
Length = 1416

Score = 34.8 bits (78), Expect = 3.1
Identities = 32/121 (26%), Positives = 52/121 (42%), Gaps = 8/121 (6%)

```
Query: 485 KSREDVESSVSNLALHNLVGRGAKGELKNLLKLKDEVGELSASDTKRFVKLVRKTEAEIL 544
           K+R      V  + ++ + G G  G +K+ +  KD V      +    ++ VR+     +
Sbjct: 266 KTRPAAPHKVPTIGVYGVPGSGKSGIIKSAVTKKDLVVSAKKENCAEIIRDVRR-----M 320

Query: 545 NKADVVCCTCVGAGDKRLDTKFRTVLIDE--STQASEPECLIPIVKGAKQVILVGDHQQL 602
            + DV   T        +     T+ IDE  +  A     LI IVK  K+V+L GD +Q
Sbjct: 321 RRMDVAARTVDSVLLNGVKHPVNTLYIDEAFACHAGTLLALIAIVK-PKKVVLCGDPKQC 379

Query: 603 G 603
           G
Sbjct: 380 G 380
```

sp|P37609|LCN2_LACLA LACTICIN 481/LACTOCOCCIN BIOSYNTHESIS PROTEIN LCNDR2 >gi|2581778
(U91581) LctM [Lactococcus lactis subsp. lactis]
Length = 922

Score = 34.8 bits (78), Expect = 3.1
Identities = 43/174 (24%), Positives = 73/174 (41%), Gaps = 32/174 (18%)

```
Query: 621 FERLISLGHVP--IRLEVQYRMNPYLSEFP--SNMFYEGSLQNGVTIEQRTVPNSKFPWP 676
           FE +IS G  P  I LE  + M  ++ ++   S     G + + V         S    P
Sbjct: 245 FENVISQGENPCIIDLETMFNMPMFVKDYKNESRNIINGKIMDSVV--------STGMLP 296

Query: 677 IRGIPMMFWANYGREEISANGTSFLNRIEAMNCERIITKLFRDGVKPEQIGVITPYEGQR 736
           + GI  +F    G +     G +F         ER+I   FRD +K ++I V       R
Sbjct: 297 VLGIDSLF----GGDPSGILGGTFSKE------ERVIINPFRDDIKFQKIVV-------R 339

Query: 737 AYILQYMQMNGSLDKDLYIKVEVASVDAFQGREKDYIILSCVRANEQQAIGFLR 790
           +    ++    + ++  Y K +    D  +G EK Y I   +  N+++ +GFL+
Sbjct: 340 SVFKDHIPFFNNNNEKRYCKPKDYVNDIIKGFEKTYKI---IVKNKEKILGFLK 390
```

gi|2688234 (AE001140) DNA helicase (uvrD) [Borrelia burgdorferi]
Length = 699

Score = 34.8 bits (78), Expect = 3.1
Identities = 25/73 (34%), Positives = 39/73 (53%), Gaps = 4/73 (5%)

```
Query: 406 FAQLNSSQSNAVSHVLQRPLSLIQGPPGTGKT-VTSATIVYHLS--KIHKDRILVCAPSN 462
```

```
           F+ LN+SQ   V    + P+ ++ G PG+GKT V  A IVY +    I  + IL    +N
Sbjct: 9   FSSLNTSQEKIVFSKSKNPMLVLAG-PGSGKTRVIIAKIVYLIKYMNIDPNEILALTFTN 67

Query: 463 VAVDHLAAKLRDL 475
            A + +  ++ DL
Sbjct: 68  KAANEMNDRINDL 80
```

ref|NP_002796.1|PPSMC5| proteasome (prosome, macropain) 26S subunit, ATPase, 5
>gi|2136322|pir||S60343 Trip1 protein - human >gi|695370
(L38810) thyroid receptor interactor [Homo sapiens]
>gi|1094810|prf||2106382A thyroid hormone
receptor-interacting protein [Homo sapiens]
Length = 406

Score = 34.4 bits (77), Expect = 4.1
Identities = 34/145 (23%), Positives = 63/145 (43%), Gaps = 23/145 (15%)

```
Query: 427 LIQGPPGTGKTVTSATIVYHLSKIHKDRILVCAPSNVAVDHLAAKLRDLGLKVVR-LTAK 485
           L+ GPPGTGKT+ +  + +H           C    V+   L  K    G ++VR L
Sbjct: 187 LLYGPPGTGKTLLARAVAHHTD---------CTFIRVSGSELVQKFIGEGARMVRELFVM 237

Query: 486 SREDVES-----SVSNLALHNLVGRGAKGELKNLLKLKDEVGELSASDTKRFVKLVRKTE 540
           +RE   S      + ++    L G G+ G  +   ++ + + +L   +  + +K++  T
Sbjct: 238 AREHAPSIIFMDEIDSIGSSRLEG-GSGGSSEVQRQMLELLNQLDGFEATKNIKVIMAT- 295

Query: 541 AEILNKADVVCCTCVGAGDKRLDTK 565
               N+ D++    +  G  R+D K
Sbjct: 296 ----NRIDMLDSALLRPG--RIDRK 314
```

pir||S51042 tat-binding protein homolog - Plasmodium falciparum
Length = 435

Score = 34.0 bits (76), Expect = 5.4
Identities = 42/185 (22%), Positives = 76/185 (40%), Gaps = 28/185 (15%)

```
Query: 370 IDKKSISGYLYYKILGHQVVDISFDVPLPKEFSIPNFAQLNSSQSNAVSHVLQRPLSLIQ 429
           ++K   S Y     L  QV ++   + LP +     F  L  SQ   V         L+
Sbjct: 170 VEKVPDSTYEMVGGLDQQVKEVKEVIELPVKHP-EIFESLGISQPKGV---------LLY 219

Query: 430 GPPGTGKTVTSATIVYHLSKIHKDRILVCAPSNVAVDHLAAKLRDLGLKVVR-LTAKSRE 488
           GPPGTGKT+ +  + +H           C    V+   L  K    G ++VR L   +RE
Sbjct: 220 GPPGTGKTLLARAVAHHTD---------CTFIRVSGSELVQKYIGEGSRMVRELFVMARE 270

Query: 489 DVES-----SVSNLALHNLVGRGAKGELKNLLKLKDEVGELSASDTKRFVKLVRKT-EAE 542
              S      + ++    + G     E++    + + + +L   ++ + +K++  T   +
Sbjct: 271 HAPSIIFMDEIDSIGSQRIEGEHGDSEVQR--TMMELLNQLDGFESTQNIKVIMCTNRID 328

Query: 543 ILNKA 547
           IL++A
Sbjct: 329 ILDEA 333
```

gi|3184291 (AC004136) putative DNA polymerase III gamma subunit [Arabidopsis
thaliana]
Length = 1153

Score = 34.0 bits (76), Expect = 5.4
Identities = 20/63 (31%), Positives = 29/63 (45%)

```
Query: 412 SQSNAVSHVLQRPLSLIQGPPGTGKTVTSATIVYHLSKIHKDRILVCAPSNVAVDHLAAK 471
           S  NAV      P+ L QGP GTGKT T+      L+ +  + +  C       D ++ K
Sbjct: 472 SLMNAVKRSRIAPVYLFQGPRGTGKTSTARIFSAALNCVATEEMKPCGYCKECNDFMSGK 531

Query: 472 LRD 474
            +D
Sbjct: 532 SKD 534
```

emb|CAB50340.1| (AJ248287) hypothetical protein [Pyrococcus abyssi]
Length = 333

Score = 34.0 bits (76), Expect = 5.4
Identities = 24/83 (28%), Positives = 41/83 (48%), Gaps = 6/83 (7%)

```
Query: 418 SHVLQRPLSLIQGPPGTGKTVTSATIVYHLSKIHKDRILVCAPSNVAVDHLAAKLRDLGL 477
           SH+L  P+  +   P     VT A + Y   ++ +D+IL+   +N    ++  +L D+G
Sbjct: 48  SHLLDIPILPVN--PKELDDVTKAQL-YSYIQLGRDKILIIGNNNAVSLNVEKELEDMGF 104

Query: 478 KVVRLTAKSREDVESSVSNLALH 500
           KV R+    R +   +   LALH
Sbjct: 105 KVTRIGGADRTE---TAEKLALH 124
```

gb|AAD44200.1|AF143772_2 (AF143772) IS1601-D [Mycobacterium avium]
>gi|5524323|gb|AAD44216.1|AF143772_21 (AF143772)
transposase [Mycobacterium avium]
>gi|5524342|gb|AAD44235.1|AF143772_42 (AF143772)

```
           IS1601-D [Mycobacterium avium]
           Length = 305

 Score = 34.0 bits (76), Expect = 5.4
 Identities = 39/155 (25%), Positives = 61/155 (39%), Gaps = 11/155 (7%)

Query: 307 QHPDWEGRGYIVRLPNSFQDT-FTLELKPSKTPPPTHLTTGFTAEFI-WKGTSYDRMQDA 364
           +HPD  GR +    PN    T  T     +       LT  F+   I W+  S+ R
Sbjct: 115 RHPDLVGRDFTASAPNQLWVTDLTFVPTWAGVAYVCFLTDAFSRMIIGWRVASHMRTTVV 174

Query: 365 LKKFAIDK----KSISGYLYYKILGHQVVDISFDVPLPKEFSIPNFAQLNSSQSNAVSHV 420
           L    + +    K ++G   +   G Q   I +   L +  ++P+   +  S  NA++
Sbjct: 175 LDTIEMARWSRGKMLAGLRCHSDAGSQFTSIRYSERLAEIGAVPSIGTVGDSFDNALAET 234

Query: 421 LQ--RPLSLIQGP--PGTGKTVTSATIVYHLSKIH 451
           +       LI GP  PG  KTV    +   LS +H
Sbjct: 235 VNGYYKAELIYGPARPGPWKTVEDVELA-TLSWVH 268


 sp|Q02137|ILVB_LACLA ACETOLACTATE SYNTHASE LARGE SUBUNIT (AHAS) (ACETOHYDROXY-ACID
           SYNTHASE LARGE SUBUNIT) (ALS) >gi|486730|pir||S35138
           probable acetolactate synthase (EC 4.1.3.18) -
           Lactococcus lactis subsp. lactis >gi|2565157 (U92974)
           IlvB [Lactococcus lactis]
           Length = 575

 Score = 33.6 bits (75), Expect = 7.0
 Identities = 36/170 (21%), Positives = 75/170 (43%), Gaps = 17/170 (10%)

Query: 359 DRMQDALKKFAIDKKSISGYLYYKILGHQVVDISFDVPLPKEF----SIPNFAQLNSSQS 414
           +++Q+ L + ++ KK +       I+    ++ S  V + + F     IP  + L    +
Sbjct: 198 EQLQELLTELSVSKKPV-------IIAGGGINYSGSVDIFRAFVEKYQIPVVSTLLGLGT 250

Query: 415 NAVSHVLQRPLSLIQGPPGTGKTVTSATIVYHLSKIHKDRILVCAPS----NVAVDHLAA 470
             +SH LQ  ++ + G       +  A  + +L     DR+ V  P+    N  V H+
Sbjct: 251 LPISHELQLGMAGMHGSYAANMALVEADYIINLGSRFDDRV-VSNPAKFAKNAVVAHIDI 309

Query: 471 KLRDLGLKVVRLTAKSREDVESSVSNLALHNLVGRGAKGELKNLLKLKDE 520
              +LG K+V+       D+++++S L   N V       +K +++ K++
Sbjct: 310 DAAELG-KIVKTDIPILSDLKAALSRLLQLNKVRTDFNDWIKTVIENKEK 358


 sp|Q56243|UVRB_THETH EXCINUCLEASE ABC SUBUNIT B >gi|1325924|dbj|BAA08653| (D49912) UvrB
           [Thermus thermophilus]
           Length = 665

 Score = 33.6 bits (75), Expect = 7.0
 Identities = 19/46 (41%), Positives = 26/46 (56%), Gaps = 4/46 (8%)

Query: 430 GPPGTGKTVTSATIVYHLSKIHKDRILVCAPSNVAVDHLAAKLRDL 475
           G  GTGKTVT A ++  L +      LV AP+ +    LAA+ R+L
Sbjct: 36  GATGTGKTVTMAKVIEALGR----PALVLAPNKILAAQLAAEFREL 77


 emb|CAB11558.1| (Z98866) similar to ATPases associated with various cellular
           activities (AAA); cDNA EST yk411g8.5 comes from this
           gene; cDNA EST yk361e3.5 comes from this gene; cDNA EST
           EMBL:C13357 comes from this gene; cDNA EST EMBL:D74251
           ...
           Length = 416

 Score = 33.6 bits (75), Expect = 7.0
 Identities = 35/156 (22%), Positives = 66/156 (41%), Gaps = 21/156 (13%)

Query: 427 LIQGPPGTGKTVTSATIVYHLSKIHKDRILVCAPSNVAVDHLAAKLRDLGLKVVR-LTAK 485
           L+ GPPGTGKT+ +  + +H           C    V+   L  K    G ++VR L
Sbjct: 197 LLFGPPGTGKTLLARAVAHHTE---------CTFIRVSGSELVQKFIGEGARMVRELFVM 247

Query: 486 SREDVESSVSNLALHNL----VGRGAKGELKNLLKLKDEVGELSASDTKRFVKLVRKTEA 541
           +RE   S +    + ++    V   + G+ +    + + + +L   +  + +K++  T
Sbjct: 248 AREHAPSIIFMDEIDSIGSSRVEGSSGGDSEVQRTMLELLNQLDGFEATKNIKVIMAT-- 305

Query: 542 EILNKADVVCCTCVGAGDKRLDTKFRTVLIDESTQA 577
              N+ D++    +  G  R+D K      DE  +A
Sbjct: 306 ---NRIDILDPALLRPG--RIDRKIEFPAPDEKARA 336


 emb|CAA09623| (AJ011482) hypothetical protein [Porcine transmissible
            gastroenteritis virus]
            Length = 2121

 Score = 33.6 bits (75), Expect = 7.0
 Identities = 39/162 (24%), Positives = 70/162 (43%), Gaps = 25/162 (15%)

Query: 330  LELKPSKTPPPTHLTTGFTAEFIWKGTSYDRMQDALKKF-----AIDKKSISGY-----L 379
            L+ + SKT PP +  + FT   I K T     +   ++      ++  KS S Y     +
Sbjct: 1740 LQWEASKTKPPLNRNSVFTCFQISKDTKIQLGEFVFEQSEYGSDSVYYKSTSTYKLTPGM 1799

Query: 380  YYKILGHQVVDISFDVPLPKE--------FSIPNFAQLNSSQSNAVSHVLQRPLSLIQGP 431
```

```
            + +  H V  +   + + +E        + + N A+  ++       + ++  + IQGP
Sbjct: 1800 IFVLTSHNVSPLKAPILVNQEKYNTISKLYPVFNIAEAYNTLVPYYQMIGKQKFTTIQGP 1859

Query: 432  PGTGKT--VTSATIVYHLSKIHKDRILVCAPSNVAVDHLAAK 471
            PG+GK+  V    + Y      + RI+  A S+ AVD L  K
Sbjct: 1860 PGSGKSHCVIGLGLYY-----PQARIVYTACSHAAVDALCEK 1896
```

sp|P54814|PRS8_MANSE 26S PROTEASE REGULATORY SUBUNIT 8 (18-56 PROTEIN) >gi|1167963 (U43728) 18-56 protein [Manduca sexta]
Length = 402

Score = 33.2 bits (74), Expect = 9.2
Identities = 32/144 (22%), Positives = 63/144 (43%), Gaps = 21/144 (14%)

```
Query: 427 LIQGPPGTGKTVTSATIVYHLSKIHKDRILVCAPSNVAVDHLAAKLRDLGLKVVR-LTAK 485
           L+ GPPGTGKT+ +  + +H           C    V+   L  K    G ++VR L
Sbjct: 183 LLYGPPGTGKTLLARAVAHHTE---------CTFIRVSGSELVQKFIGEGSRMVRELFVM 233

Query: 486 SREDVESSVSNLALHNL----VGRGAKGELKNLLKLKDEVGELSASDTKRFVKLVRKTEA 541
           +RE   S +    + ++    +  G+ G+ +    + + + +L   +  + +K++  T
Sbjct: 234 AREHAPSIIFMDEIDSIGSSRIESGSGGDSEVQRTMLELLNQLDGFEATKNIKVIMAT-- 291

Query: 542 EILNKADVVCCTCVGAGDKRLDTK 565
              N+ D++    +  G  R+D K
Sbjct: 292 ---NRIDILDPALLRPG--RIDRK 310
```

gb|AAC55658.1| (U51931) helicase [Beet yellow stunt virus]
Length = 710

Score = 33.2 bits (74), Expect = 9.2
Identities = 29/102 (28%), Positives = 50/102 (48%), Gaps = 12/102 (11%)

```
Query: 405 NFAQLNSSQSNAVSHVLQRPLSLIQGPPGTGKTVTSATIVYHLSKIH-KDRILVCAPSNV 463
           +F + NSS  N +       + L + PPG GKT    T++    K+H K+RILV   +
Sbjct: 293 SFLRSNSSYRNFIFENDSCRIRLYEAPPGGGKT---HTLIASFVKMHKKNRILVLTANKS 349

Query: 464 AVDHLAAKLRD-------LGLKVVRLTAKS-REDVESSVSNL 497
           +   +  K+ D          K+++  +K+ RE+  S+ SN+
Sbjct: 350 SQVEILKKINDSLKREHETKTKLLKFASKAERENYPSADSNV 391
```

sp|O18413|PRS8_DROME 26S PROTEASE REGULATORY SUBUNIT 8 >gi|2245467 (U97538) DUG [Drosophila melanogaster]
Length = 405

Score = 33.2 bits (74), Expect = 9.2
Identities = 32/144 (22%), Positives = 63/144 (43%), Gaps = 21/144 (14%)

```
Query: 427 LIQGPPGTGKTVTSATIVYHLSKIHKDRILVCAPSNVAVDHLAAKLRDLGLKVVR-LTAK 485
           L+ GPPGTGKT+ +  + +H           C    V+   L  K    G ++VR L
Sbjct: 186 LLYGPPGTGKTLLARAVAHHTE---------CTFIRVSGSELVQKFIGEGSRMVRELFVM 236

Query: 486 SREDVESSVSNLALHNL----VGRGAKGELKNLLKLKDEVGELSASDTKRFVKLVRKTEA 541
           +RE   S +    + ++    +  G+ G+ +    + + + +L   +  + +K++  T
Sbjct: 237 AREHAPSIIFMDEIDSIGSSRIESGSGGDSEVQRTMLELLNQLDGFEATKNIKVIMAT-- 294

Query: 542 EILNKADVVCCTCVGAGDKRLDTK 565
              N+ D++    +  G  R+D K
Sbjct: 295 ---NRIDILDPALLRPG--RIDRK 313
```

gi|2815905 (AF043734) Pros45 proteosome subunit homolog [Drosophila melanogaster]
Length = 405

Score = 33.2 bits (74), Expect = 9.2
Identities = 32/144 (22%), Positives = 63/144 (43%), Gaps = 21/144 (14%)

```
Query: 427 LIQGPPGTGKTVTSATIVYHLSKIHKDRILVCAPSNVAVDHLAAKLRDLGLKVVR-LTAK 485
           L+ GPPGTGKT+ +  + +H           C    V+   L  K    G ++VR L
Sbjct: 186 LLYGPPGTGKTLLARAVAHHTE---------CTFIRVSGSELVQKFIGEGSRMVRELFVM 236

Query: 486 SREDVESSVSNLALHNL----VGRGAKGELKNLLKLKDEVGELSASDTKRFVKLVRKTEA 541
           +RE   S +    + ++    +  G+ G+ +    + + + +L   +  + +K++  T
Sbjct: 237 AREHAPSIIFMDEIDSIGSSRIESGSGGDSEVQRTMLELLNQLDGFEATKNIKVIMAT-- 294

Query: 542 EILNKADVVCCTCVGAGDKRLDTK 565
              N+ D++    +  G  R+D K
Sbjct: 295 ---NRIDILDPALLRPG--RIDRK 313
```

CPU time: 107.60 user secs. 1.06 sys. secs 108.66 total secs.

Database: Non-redundant GenBank CDS translations+PDB+SwissProt+SPupdate+PIR
Posted date: Sep 12, 1999 1:55 PM
Number of letters in database: 125,914,940

```
  Number of sequences in database:  411,205

Lambda     K      H
   0.317    0.134    0.394

Gapped
Lambda     K      H
   0.270   0.0470    0.230


Matrix: BLOSUM62
Gap Penalties: Existence: 11, Extension: 1
Number of Hits to DB: 239624468
Number of Sequences: 411205
Number of extensions: 10249470
Number of successful extensions: 30246
Number of sequences better than 10.0: 120
Number of HSP's better than 10.0 without gapping: 65
Number of HSP's successfully gapped in prelim test: 55
Number of HSP's that attempted gapping in prelim test: 29812
Number of HSP's gapped (non-prelim): 197
length of query: 971
length of database: 125914940
effective HSP length: 58
effective length of query: 913
effective length of database: 102065050
effective search space: 93185390650
effective search space used: 93185390650
T: 11
A: 40
X1: 16 ( 7.3 bits)
X2: 38 (14.8 bits)
X3: 64 (24.9 bits)
S1: 41 (21.7 bits)
S2: 74 (33.2 bits)
```

Analysis Coffee Break

RNA surveillance: watching the defectives

the potential for a conserved mechanism of nonsense-mediated mRNA decay

		I (A)		Ia		II (B)	
Human	486	LSLIQGPPGTGKTVTSATIVYHLA	5	PVLVCAPSNIAVDQLTEKIHQTG	94	ILIDE	18
Fission yeast	408	LSLIQGPPGTGKTVTSASVVYHLA	11	PVLVCAPSNVAVDQLAEKIHRTG	94	VLIDE	18
Brewer's yeast	424	LSLIQGPPGTGKTVTSATIVYHLS	5	RILVCAPSNVAVDHLAAKLRDLG	92	VLIDE	18
Nematode worm	461	LSLIQGPPGTGKTVVSATIVYHLV	5	NVLVCSPSNIAVDHLAEKIHKTG	93	VLIDE	18

	III		IV		V		VI	
Human	QLILVGDHCQL	31	LQVQYRMHPALS	114	ASVDAFQGREKDFIILS	20	VALTR	253
Fission yeast	QVVLVGDHQQL	31	LVVQYRMHPCLS	114	ASVDAFQGREKDFIILS	20	VALTR	132
Brewer's yeast	QVILVGDHQQL	31	LEVQYRMNPYLS	114	ASVDAFQGREKDYIILS	20	VGLTR	170
Nematode worm	QLVLVGDHCQL	31	LQVQYRMHPVLS	114	ASVDAFQGREKDYIIVT	20	VAITR	230

The alignment above was constructed from a selection of sequences found to be highly similar to yeast Upf1p. Illustrated here is the helicase domain only, although there is high similarity over most of the length of the proteins (see pairwise alignments). The pattern of domains I-VI is characteristic of superfamily type I helicases and sets them apart from the well-known DEAD/H superfamily, also known as superfamily type II helicases. There are two domains - I (A) and II (B) - implicated in ATP-binding and hydrolysis, and one in RNA binding, VI. Domains Ia, III, IV, and V are also characteristic of superfamily type I helicases. The number of amino acids preceding and following the helicase region is shown in blue; the number of amino acids found between each domain is shown in purple.

Back to BLAST result

Upf1p was used to search the non-redundant protein sequence database using the PSI-BLAST program [3] with standard parameters. Sequences included in the multiple alignment of the helicase domain were selected from output from the first iteration. The multiple sequence alignment was constructed using ClustalW [4].

[1] Koonin, E.V. (1992) A new group of putative RNA helicases *Trends Biochem. Sci.* 17, 495-497
[2]Culbertson, M.R. (1999) RNA surveillance: unforseen consequences for gene expression, inherited genetic disorders and cancer *Trends Genet.* 15, 74-80
[3] Altschul, S. F. et al. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389-402
[4] Higgins, D. G., Thompson, J. D. and Gibson, T. J. (1996) Using CLUSTAL for multiple sequence alignments. *Methods Enzymol* 266, 383-402

RENT1, a human homolog of Upf1p, resulting in a population of truncated proteins, which could be harmful when targeted to their site of function.

Comments?

Questions?
We would welcome feedback on NCBI's Coffee Break.
Email to: info@ncbi.nlm.nih.gov

Download this article in Adobe Acrobat PDF format