

International Journal of Biochemistry Research & Review

26(3): 1-60, 2019; Article no.IJBCRR.49940

ISSN: 2231-086X, NLM ID: 101654445

Eukaryotic Multi-subunit DNA dependent RNA Polymerases: An Insight into Their Active Sites and Catalytic Mechanism

Peramachi Palanivelu^{1*}

¹Department of Molecular Microbiology, School of Biotechnology, Madurai Kamaraj University, Madurai - 625021, India.

Author's contribution

The sole author designed, analysed, interpreted and prepared the manuscript.

Article Information

DOI: 10.9734/IJBCRR/2019/v26i330097

Editor(s)

(1) Dr. Mudasir Hafiz Khan, Assistant Professor, Genetics and Plant Breeding, Faculty of Agriculture, Sher-e-Kashmir University of Agricultural Sciences and Technology of Kashmir, Shalimar Srinagar (J&K), India.

Reviewers:

(1) J. Y. Peter, University of Abuja, Nigeria. (2) Karen Cordovil, Fiocruz, Brazil.

(3) Maciej Gagat, Nicolaus Copernicus University, Toruń.

Complete Peer review History: http://www.sdiarticle3.com/review-history/49940

Original Research Article

Received 30 April 2019 Accepted 09 July 2019 Published 17 July 2019

ABSTRACT

Aim: To analyze the most complex multi-subunit (MSU) DNA dependent RNA polymerases (RNAPs) of eukaryotic organisms and find out conserved motifs, metal binding sites and catalytic regions and propose a plausible mechanism of action for these complex eukaryotic MSU RNAPs, using yeast (*Saccharomyces cerevisiae*) RNAP II, as a model enzyme.

Study Design: Bioinformatics, Biochemical, Site-directed mutagenesis and X-ray crystallographic data were analyzed.

Place and Duration of Study: School of Biotechnology, Madurai Kamaraj University, Madurai, India, between 2007- 2013.

Methodology: Bioinformatics, Biochemical, Site-directed mutagenesis (SDM) and X-ray crystallographic data of the enzyme were analyzed. The advanced version of Clustal Omega was used for protein sequence analysis of the MSU DNA dependent RNAPs from various eukaryotic sources. Along with the conserved motifs identified by the bioinformatics analysis, the data already available by biochemical and SDM experiments and X-ray crystallographic analysis of these enzymes were used to confirm the possible amino acids involved in the active sites and catalysis.

Results: Multiple sequence alignment (MSA) of RNAPs from different eukaryotic organisms showed a large number of highly conserved motifs among them. Possible catalytic regions in the catalytic

subunits of the yeast Rpb2 (= β in eubacteria) and Rpb1 (= β' in eubacteria) consist of an absolutely conserved amino acid R, in contrast to a K that was reported for DNA polymerases and single subunit (SSU) RNAPs. However, the invariant 'gatekeeper/DNA template binding' YG pair that was reported in all SSU RNAPs, prokaryotic MSU RNAPs and DNA polymerases is also highly conserved in eukaryotic Rpb2 initiation subunits, but unusually a KG pair is found in higher eukaryotes including the human RNAPs. Like the eubacterial initiation subunits of MSU RNAPs, the eukaryotic initiation subunits, viz. Rpb2, exhibit very similar active site and catalytic regions but slightly different distance conservations between the template binding YG/KG pair and the catalytic R. In the eukaryotic initiation subunits, the proposed catalytic R is placed at the -9th position from the YG/KG pair and an invariant R is placed at -5 which are implicated to play a role in nucleoside triphosphate (NTP) selection as reported for SSU RNAPs (viral family) and DNA polymerases. Similarly, the eukaryotic elongation subunits (Rpb1) are also found to be very much homologous to the elongation subunits (B') of prokaryotes. Interestingly, the catalytic regions are highly conserved, and the metal binding sites are absolutely conserved as in prokaryotic MSU RNAPs. In eukaryotes, the template binding YG pair is replaced with an FG pair. Another interesting observation is, similar to the prokaryotic β' subunits, in the eukaryotic Rpb1 elongation subunits also, the proposed catalytic R is placed double the distance, i.e., -18 amino acids downstream from the FG pair unlike in the SSU RNAPs and DNA polymerases where the distance is only -8 amino acids downstream from the YG pair. Thus, the completely conserved FG pair, catalytic R with an invariant R, at -6th position are proposed to play a crucial role in template binding, NTP selection and polymerization reactions in the elongation subunits of eukaryotic MSU RNAPs. Moreover, the Zn binding motif with the three completely conserved Cs is also highly conserved in the eukaryotic elongation subunits. Another important difference is that the catalytic region is placed very close to the N-terminal region in eukaryotes.

Conclusions: Unlike reported for the DNA polymerases and SSU RNA polymerases, the of eukaryotic MSU RNAPs use an R as the catalytic amino acid and exhibit a different distance conservation in the initiation and elongation subunits. An invariant Zn²⁺ binding motif found in the Rpb1 elongation subunits is proposed to participate in proof-reading function. Differences in the active sites of bacterial and human RNA polymerases may pave the way for the design of new and effective drugs for many bacterial infections, including the multidrug resistant strains which are a global crisis at present.

Keywords: Multi-subunit DNA dependent RNA polymerases; eukaryotic RNA polymerases; RNA polymerase II; Saccharomyces cerevisiae; conserved motifs; polymerase active site; polymerization mechanism; transcription slippage diseases; drug design.

1. INTRODUCTION

RNAPs (EC 2.7.7.6) are key enzymes which play a vital role in the flow of genetic information from DNA to proteins and the proteins are final players in all cellular processes. Therefore, they are found in all living organisms and play a crucial role in copying DNA sequences into RNA molecules which are subsequently translated into proteins. Thus, transcription forms the first step and a key control point in gene regulation and expression. Errors in the transcription process can potentially lead to aberrant gene products and ultimately lead to various diseases including cancer. One major class of transcription error, as transcriptional slippages, implicated in the development of a wide variety of diseases, such as colon cancer, non-familial Alzheimer's, Down's syndrome, etc. [1]. Recently the in vitro transcribed mRNAs have come into

focus as a potential new class of drugs known as 'mRNA therapeutics' to deliver genetic information through mRNAs to correct the malfunction(s) [2] and also potential vaccines against cancer treatment [3]. In fact, many anticancer drugs act by inhibiting the transcription step itself. Therefore, understanding the mechanism and regulation of RNAPs have been a key goal of molecular biologists since its discovery. Interestingly, the sequences, overall 3D framework and functions of these DNA dependent RNAPs are universally conserved in viruses, bacteria, plants and animals with small but significant differences in their active sites and catalytic amino acids. For the discussion about the discovery and dynamics of the polymerization reactions [refer to Palanivelu 4]. Instead of a single type of RNAPs in prokaryotes, Robert Roeder and William Rutter discovered in 1969 the existence of three different RNAPs in eukaryotic cells that were responsible for transcription of all types of RNAs in the eukaryotic cells [5].

Though RNAPs are found in all organisms, their number and composition vary across taxa, possibly due to evolutionary consequences, changes in the genome structure and the complexity of the transcription process. However, the MSU RNAPs display a conserved core structure across all major domains of life, viz. viruses, bacteria, archaea and eucarya. For example, viruses contain mainly two different RNAPs, viz. DNA dependent RNAPs and RNA dependent RNAPs. Both eubacteria and archaebacteria contain a single type of MSU RNAPs, while eukaryotes contain at least five distinct types of MSU RNAPs (I-V). While the eubacterial enzymes are composed of 5 different subunits, the eukaryotic enzymes are made up of up to 12-16 different subunits. Despite such differences, there are striking similarities among the transcriptional mechanisms by various types of RNAPs across the three major domains of life [6 and references therein].

Transcription and transcriptional regulation are very important in eukaryotes as it underlies all aspects of cellular metabolism including oncogenesis (cancer) and morphogenesis (development). The major type of RNA polymerase, i.e., the RNA polymerase II, which involves in the transcription of genes, is a large (550 kDa) complex made up of 12 subunits. Unlike in prokaryotes, gene expression, and therefore, RNAP II activity is controlled by several proteins known as general transcription factors. In other words, as there are no Sigma like factors in eukaryotic cells for promoter recognition and to initiate transcription, the eukaryotic RNAPs interact with a variety of protein factors known as general transcription factors. Large volumes of genomic and protein sequence data are available for most of the prokaryotic and eukaryotic MSU RNAPs. Understanding the biological principles buried in these sequencing data is a significant challenge for scientists. Therefore, the purpose of the study is to analyze the available protein sequences of the MSU RNAPs from eukaryotes to find out the highly conserved motifs among them and to understand the structure- function relationships and mechanism of action of these key enzymes. along with the X-ray crystallographic and experimental data available on these enzymes from various sources. A comparative analysis of the bacterial and human RNAPs for their

transcription mechanism will pave way to design new and effective drugs for many bacterial infections, including the antibiotic resistance, especially the multi-drug resistance, which has become a global crisis [7,8]. MSA has been proven to be very useful for assigning a function to a given sequence, by comparing the existing experimental data with the protein sequence data and thus, bridging the gap between the protein sequence and experimental data.

RNAPs catalyze the chemical reaction that synthesizes an RNA strand from a DNA template with all the 4 NTPs and a metal ion, usually a Mg²⁺ ion. RNA synthesis in eukaryotes involves four steps, viz. formation of pre-initiation complex, initiation, elongation and termination. The nucleotides are added one at a time to the growing 3' end. The newly formed RNA copies serve as blueprints for the synthesis of proteins during the next step of translation. The basic transcription unit is the distance between the sites of transcription start site (TSS) and transcription termination site (TTS), and may have one or more genes between them (e.g., mono or poly-cistronic mRNAs; poly-cistronic mRNAs are uncommon in eukaryotes).

2. TYPES OF MSU DNA DEPENDENT RNAPS OF PROKARYOTES AND EUKARYOTES

There are at least 4 different types of MSU RNAPs in living cells, viz.

- 1) MSU DNA dependent RNAPs of eubacteria
- MSU DNA dependent RNAPs of archaebacteria
- MSU DNA dependent RNAPs of chloroplasts (Plastid encoded)
- MSU DNA dependent RNAPs of eukaryotes

2.1 Types of MSU DNA Dependent RNAPs in Eukaryotes

As mentioned elsewhere, bacteria and archaea have only a single RNAP to transcribe all of its genes whereas eukaryotic cells employ different RNAPs to transcribe different types of genes, viz. rRNAs, tRNAs and mRNAs (Table 1).

There are at least 7 major types of RNAPs in eukaryotes. The nuclear genome utilizes three

Table 1. Types and functions of prokaryotic and eukaryotic RNAPs

RNA polymerase type	Type of RNA(s) synthesized
RNA Polymerase I (Nuclear)	rRNA genes
	(5.8S, 18S, 28S rRNAS from 45S pre-rRNA)
2. RNA Polymerase II (Nuclear)	mRNAs, snRNAs, microRNAs
3. RNA Polymerase III (Nuclear)	tRNAs & 5S rRNA, scRNAs, U6 SnRNA
4. RNA Polymerase IV (Plant specific)	siRNAs in plants
5. RNA Polymerase V (Plant specific)	Plant specific RNAs involved in siRNA directed
	heterochromatin formation in plants.
MSU RNAP type (Genomic)	Eubacteria and Chloroplasts (rRNAs, tRNAs mRNAs)

major types of RNAPs which are localized in the nucleus. For example, the RNAP I is localized in the nucleolus and primarily involves in the synthesis of the rRNAs and the other two RNAPs, viz. RNAPs II and III are localized in the nucleoplasm and mainly involved in the synthesis of mRNAs and tRNAs, respectively (Table 1). Plants are unique among eukaryotes in having five nuclear MSU RNAPs. Two plant-specific RNAPs, polymerases IV and V are 12-subunit enzymes that are evolved as specialized forms of Pol II. (Pols IV and V are nonessential for viability but play important roles in RNAmediated gene silencing pathways that tame transposons and defend against invading viruses [9]).

2.2 Basic Structure and Composition of the Eukaryotic RNAPs

All the 5 eukaryotic enzymes (RNAPs I-V) are MSU enzymes. The 12-subunit RNAP II is the enzyme largely responsible for transcription of protein-encoding genes and thus, forms the central component of the eukaryotic transcription machinery. Similarities between the eukaryotic and prokaryotic MSU enzymes, suggest a common lineage in the evolutionary tree. For example, counterparts for all the five core eubacterial RNAP subunits ($\alpha_2\beta\beta'\omega$) are found in the eukaryotic RNAPs I, II and III [10] (Table 2). The five orthologs include the two largest catalytic subunits Rpb1 and Rpb2, which correspond to the bacterial β' and β subunits, respectively [11]. Furthermore, Rpb3 and Rpb11 correspond to the two copies of the bacterial $\boldsymbol{\alpha}$ subunit, and the Rpb6 subunit corresponds to the bacterial ω subunit [12]. In addition to, the two large subunits of RNAP II share some sequence homology and antigenic determinants with the corresponding subunits of RNAPs I and III [13] (Refer to also Mix and Match analysis Figs. 5 and 6). Rpb3 is involved in RNAP II assembly. The subunits Rpb4 and Rpb7 form a heterodimer

and associate reversibly with the '10-subunit core polymerase' and involves in the initiation process which is otherwise defective in initiation. Rpb5, an evolutionarily highly conserved, universal eukaryotic RNAP subunit, shared by all three enzymes, facilitates communication between the RNAP core with a variety of basal and genespecific transcription factors [14]. Rpb7 is essential for the functioning of the RNAP-II as deletion is also found to be lethal. A cleft is formed by jaws; the upper jaw is formed by regions of Rpb1, Rpb2 and Rpb9 whereas the lower jaw is formed by Rpb1 and Rpb5. The jaws are thought to grab the incoming DNA template [15] (Fig. 1).

Consistent with the increased complexity of the eukaryotic genome, all three major RNAPs of the eukaryotic transcriptional machinery have several additional subunits that do not have bacterial counterparts. Table 2 shows the subunit compositions of the three major types of nuclear MSU RNAPs. Apart from the five orthologs, as discussed above, the nuclear RNAPs share four common subunits, while the remaining subunits are RNAP dependent. Thus, the three major classes of eukaryotic RNAPs: I, II and III, in addition to comprising of two large catalytic subunits are also made up of 10-14 smaller subunits (Table 2).

2.3 Characteristics of the C-terminal Domain (CTD) in the Largest Subunit Rpb1

Though exact functions of all the eukaryotic RNAPs are more or less completely understood, yet much data are not available on the individual subunits except for the larger subunits. An interesting feature of the RNAP II, a 12-subunit complex, is its unique CTD in the largest subunit Rpb1. The Rpb1 elongation subunit consists of a tandem repeat of a conserved heptapeptide repeat sequences (-YS²PTS⁵PS⁻-). These

Table 2. Composition of the three major eukaryotic nuclear RNAPs

Features	Pol I	Pol II	Poll III
Subunits unique to each RNAP (α₂ββ'ω- like)	5	5*	5
Common subunits to all 3 RNAPs	4	4	4
Additional unique subunits to each RNAPs	5	3	7
Total No. of subunits	14	12	16
Products	pre-rRNAS	pre-mRNAs 5 snRNAs^	pre-tRNAs
	(45S RNA → 28S, 5.8S, 18S)	Sno RNAs, microRNAs	5S & 7S RNAs ^{\$} , U6- snRNA
Sensitivity to α-Amanitin	Nil	High (1 µg/ml)	Moderate (10 µg /ml)
Sensitivity to Actinomycin-D#	0.05 μg/ml	0.5 μg/ml	5.0 μg/ml

NB:*The largest subunit, Rpb1, has unique Carboxy Terminal Domain (CTD); # [16].
Subunit nomenclatures RNAPs I, II & III: RP A1-A14; RP B1-B12; RP C1-C16
^U1-U5 of ~200 bases; involves in the formation of spliceosomes

\$7S RNA from the signal recognition particle (SRP), which is involved in the transport of proteins into the endoplasmic reticulum

heptapeptide repeats are completely conserved in sequence and function from yeast to human with varying lengths from 25 to 52. (The CTD domain does not exist in RNAPs I and III suggesting its special importance to mRNA processing). The Ser residues phosphorylated and dephosphorylated during the active transcription process. Therefore, the RNAP-II exists in two forms, i.e., unphosphorylated and phosphorylated forms (II and II^P), respectively [17]. In fact, the transition between the two forms facilitates different functions during transcription. (The phosphorylation of CTD is catalyzed by TFII-H, one of the six general transcription factors associated with RNAP II. TFII-H plays a dual role: one is to unwind the DNA at the transcription start site and the other is to phosphorylate the heptapeptide repeats. (TFII-H is a large protein complex that contains among others the Cdk7/Cyclin-H kinase complex for phosphorylation of CTD and an ATP dependent DNA helicase to unwind the DNA and open up the transcription bubble at TSS). It also involves transcription-coupled DNA mismatch repair. (Mutations in the human XPD kinase Xeroderma pigmentosum Trichothiodystrophy [18].

The elongation of initiation is accomplished by the phosphorylation of Ser⁵ of the heptapeptide, (-YS²PTS⁵PS-) by the TFII-H. The Ser⁵ phosphorylation recruits enzymes to cap the 5' end of the newly synthesized mRNA and the 3' processing factors to poly(A) sites. Once the second Ser is phosphorylated, i.e., Ser²,

elongation is activated. West and Corden have shown the substitution of Ala or Glu for Ser in positions 2 or 5 is lethal. In addition, changing Tyr in position 1 to Phe is also lethal [19]. In order to terminate elongation, dephosphorylation is accomplished phosphatases. Once the domain is completely dephosphorylated, the RNAP II is "recycled" and catalyzes the same process with another initiation site. Thus, the CTD acts as a platform for various transcription factors as it binds or dissociates them, depending upon their requirements during the transcription process. Thus, the phosphorylation and dephosphorylation of the CTD is important regulatory mechanism exhibited only by RNAP II [20]. Ser7 phosphorylation is required for the transcription of SnRNAs and a mutation of Ser7 to Ala causes a specific defect in snRNA expression [21].

The number of heptapeptide repeats in the CTD increases with genomic complexity; for example, 17 in Plasmodium, 26 in yeast, 32 in Caenorhabditis elegans, 45 in Drosophila, and 52 in mammals). The CTD deletion experiments have shown the CTD's central role in coupling transcription to all three of the main mRNA processing events [22]. For example, deletion of most of the CTD can result in inefficient capping, splicing, and polyadenylation *in vivo*. SDM experiments of the yeast RNAP II has found that at least 10 repeats are needed for the viability of the process [23,24].

2.4 Salient Features of the Yeast (S. cerevisiae) MSU RNAPs

To-date one of the most well-studied eukaryotic RNAPs is from the baker's yeast, S. cerevisiae. In fact, our present understanding of the eukaryotic transcription system is mainly based on the yeast system. (For solving the structure and molecular basis of transcription by the yeast RNAP II, Roger Kornberg was awarded Nobel Prize in chemistry in 2006). Like other eukaryotic cells, the yeast cells also contain 3 different polymerases as discussed above. All the three veast polymerases have five core subunits that exhibit good homology to the β , β , α and ω subunits of E. coli RNAP. RNAPs I and III contain the same two non-identical α -like subunits. whereas polymerase II has two copies of a different α-like subunit. All three yeast polymerases share four other common subunits as mentioned elsewhere. In addition, each RNAP contains three to seven unique smaller subunits as shown in Table 3.

The yeast RNAP II, which involves in the production of mRNAs, is extensively characterized and a great deal of information is available. The yeast RNAP II is composed of 12 subunits and the largest subunit (Rpb1) contains the essential CTD, containing 26 heptapeptide repeats (YSPTSPS). Thus, the Rpb1, Rpb2, Rpb3 and Rpb11 subunits are related to the subunits of the eubacterial RNA polymerase core enzyme, whereas the 5 subunits, viz. Rpb5, Rpb6, Rpb8, Rpb10 and Rpb12 are shared among yeast RNA polymerases I, II and III [25] (Table 3).

The yeast RNAP II core enzyme is composed of 12 subunits (Fig. 1). The 7 subunits, viz. Rpb1-4, Rpb7, Rpb9 and Rpb11 are unique to RNAP II

while the 5 subunits, viz. Rpb5, Rpb6, Rpb8, Rpb10 and Rpb12 are shared between the three RNAPs (Table 3). The $\triangle rpb4$ phenotypes can be suppressed by overexpression of Rpb7, and the high level of Rpb7 allows its interaction with RNAP II in the absence of Rpb4, suggesting that Rpb7 is the critical component of the Rpb4–Rpb7 complex and the role of Rpb4 is to stabilize the complex [26]. MSA analysis of the Rpb7 from various eukaryotic sources have shown 3 template binding pairs, possibly decide the right orientation of the template DNA by three-point attachment in addition to two long stretches of conserved motifs (data not shown). Rpb2, a protein of 138,750 Daltons, exists as a single copy in the haploid yeast genome and disruption of the gene is lethal to the yeast cell

During the production of the primary transcript by RNAP-II, the phosphorylation state (P) of the CTD changes to allow the transcribing polymerase to associate with the capping, splicing, polyadenylation and mRNA export machinery [22]. X-ray crystallographic results also provide evidence for RNA exit in the vicinity of the carboxyl-terminal repeat domain, coupling synthesis to RNA processing by enzvmes bound this domain to [28]. These associations are essential for normal processing of pre-mRNAs to generate mature mRNAs and to export them to the cytoplasm and also for normal termination of transcription by RNAP II. The subunit Rpb1 (RPO21) is the largest and catalytic component of RNAP II and similar to the eubacterial β'. Bacterial RNA polymerase subunit ω and eukaryotic RNA polymerase subunit Rpb6 (shared by RNAPs I. II. and III) are thesequence, structural, and functional homologs and promote polymerase assembly [12].

Table 3. Composition of the MSU RNAP II from S. cerevisiae

Subunit	Size (~kDa)	Subunit	Size (~kDa)
Rpb1 (β')	192/E	Rpb7**	19/E
Rpb2 (β)	139/E	Rpb8*	17E/S
Rpb3 (~α)	35/E	Rpb9	14/NE
Rpb4	25/NE	Rpb10*	8.8/E/S
Rpb5*	25/E/S	Rpb11	14/E
Rpb6* (ω)	18/E/S	Rpb12*	7.7/E/S

Adapted from [25]; E, Essential; NE, Nonessential; S, Shared.

*The 5 Rpb subunits, viz. 5, 6, 8, 10 and 12 are common for all 3 RNAPs, viz. RNAP I, II & III
The Rpb6 promotes RNAP assembly and contains 9 C2H2 zinc fingers [27]; ** Rpb7 is unique to RNAP
II. (Rpbs7 and 4 form a dimer and initiate transcription, not required for elongation); Rpb3 is not
absolutely equivalent to α subunit of E. coli.; Subunits Rpb1, Rpb2, Rpb3 and Rpb7 are absolutely
required for activity; Rpb1 is known to bind strongly to Rpb5; Rpb12 interacts with Rpb3

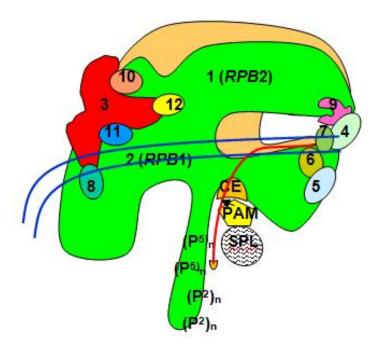


Fig. 1. A schematic diagram of the 12-subunit yeast MSU RNAP-II with the CTD and processing components

CE Capping Enzyme; PAM, Poly-Adenylation Machinery; SPL, Spliceosome; $(P^5)_n$, phosphorylation at Ser5 positions (during initiation); $(P^2)_n$ phosphorylations at Ser2 positions (during elongation)

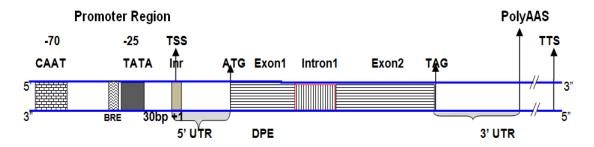


Fig. 2. Schematic diagram of a typical transcription unit in eukaryotes

NB: Blue lines, DNA strands; TATA, TATA box (Hogness box) (some promoters are TATAless), CAAT, CAAT box; The starting nucleotide (nt) is always an A in eukaryotes; UTR, UnTranslated Region); PolyAAS, Poly-A Addition Site (AATAAA); TSS, Transcription Start Site; Inr, Initiation region –pyTCACA-; DPE, Downstream Promoter Element (is seen ~ +30 bp in many genes, AGACA); TTS, Transcription Termination Site

In this communication, the yeast RNAP II is used as the model system and the structural studies of yeast RNAP II are directly relevant to RNAP II enzymes in higher organisms, as the yeast RNAP II subunits are very well conserved in sequence and function as discussed elsewhere. These analyses not only bridge the data obtained by protein sequence analysis and the experimental results obtained on the yeast RNAP II but also will pave for more SDM experiments and genetic analysis to dissect the transcription mechanism in eukaryotes in much detail. The following brief discussion on the basic structure of the transcription unit, transcription cycle and the participation of the RNAP II subunits will be useful to further understand the analysis and conclusions.

3. BASIC STRUCTURE OF THE TRANSCRIPTION UNITS IN EUKARYOTES

The transcription units in eukaryotes are slightly different from the one described in prokaryotes

[4] as most of the eukaryotic genes are interrupted with intervening sequences called introns. A typical transcription unit in eukaryotes is shown in Fig. 2.

Some transcription units have an InR start site. TATA-like site, Down Stream Promoter Elements (DPE), Upstream activator elements/ Enhancer elements, Repressor and even Insulator and Silencer sequences. The poly-A cleavage site is about 30-35 or even 100 nucleotides downstream of the STOP codon (TAG). This sequence at the mRNAs' 3' end is called poly-A signal. The sequence present in the terminal region of the gene is TTATTT (poly-A signal) is usually followed by poly-U of 20 to 35 nts downstream from the poly-A signal.

5' UTR is the portion of an mRNA from the 5' end to the first codon; the 3' UTR is the portion of an mRNA from the last codon to the poly-A site. The UTRs play crucial roles in mRNA stability, transport, translation efficiency, etcl.

4. TRANSCRIPTION PROCESSES IN EUKARYOTES

The basic transcription unit in eukaryotes is the distance between the sites of transcription start site (TSS) and transcription termination site (TTS), and the coding region is usually interrupted by one to many introns (Fig. 2)

The RNA synthesis in eukaryotes is carried out in the nucleus and involves four distinctive steps, viz. formation of pre-initiation complex (PIC), initiation, elongation and termination. It is interesting to note that the eukarvotic RNAPs do not possess any initiator subunits similar to the σ factors, as reported in prokarvotes.. However, unlike in prokaryotes where the initiation starts with the specific binding of the single σ subunit, in eukaryotes, at least about half a dozen protein factors involve in the initiation of transcription and are collectively known as General Transcription factors (GTFs). These GTFs assemble at the promoter region to form the PIC. Once the PIC is formed and the RNAP is positioned at TSS, the initiation subunit Rpb2 initiates RNA synthesis by making short RNA primers of ~10 nts, from which the elongation subunit Rpb1 extends and terminates at TTS. In eukaryotes, there is no well-defined transcriptional terminator region as

in prokaryotes and therefore, the transcription progresses well beyond 1000 or more nts downstream from the stop codon. Unlike RNAP I and III, RNAP II lacks any specific termination signals.

In my earlier communication, SSU RNAPs and MSU RNAPs (prokaryotic and prokaryotic type, i.e., plastid-encoded) RNAPs were analyzed in detail [29, 4, 30]. In this communication, the most complex MSU RNAPs of eukaryotic origin are analyzed for their conserved motifs, active sites, metal binding regions and based on these findings, a plausible mechanism of action is proposed for these MSU eukaryotic enzymes using the yeast MSU RNAP II as the model enzyme.

5. MATERIALS AND METHODS

A large number of MSU RNAPs from eubacteria and eukaryotes have been isolated, purified, characterized, cloned and sequenced [17,31,6]. Complete nucleic acid protein sequence data are available for these enzymes from different eukaryotic sources. Thus, these data have become valuable tools in analyzing and understanding the structurefunction relationships of these most complex enzymes which play a vital role gene expression. This communication presents consensus model for initiation and elongation processes and also a plausible mechanism of action for these enzymes.

The S. cerevisiae DNA-dependent MSU RNAP II is used as the model system for delineating the polymerization mechanism. Biochemically and genetically as this is the most well-studied enzyme, a large amount of data on biochemical, SDM and X-ray analyses of its subunits make this enzyme a convenient model for investigating the physicochemical aspects of transcription in eukaryotes. For MSA of various eukaryotic RNAP II, the sequences were retrieved from SWISS-PROT and PUBMED sites and analyzed using Clustal Omega, an accurate, fast and widely accepted algorithm, available on their website.

6. RESULTS AND DISCUSSION

6.1 MSA of Eukaryotic MSU RNA Polymerases from Different Sources

The eukaryotic RNAP II is composed of various subunits as described elsewhere (Table 2). Only

the two catalytic subunits, viz. Rpb2 and Rpb1 which involve in the initiation elongation processes are analyzed to find out the conserved and active site regions among them. Figs. 3 and 4 show the MSA of the two catalytic subunits of MSU RNAP II from various eukaryotic sources. To reduce the length of the article, only the relevant and highly conserved regions are shown in the Figures. The possible catalytic, template and substrate binding motifs are highlighted in yellow and the metal binding regions are highlighted in green/orange. The subunit sequences of standard organisms highlighted in yellow/magenta and used for numbering.

6.1.1 MSA of Rpb2 subunits of eukaryotic MSU RNAPs (RNAP II)

Fig. 3 shows the MSA analysis and conserved motifs of the Rpb2 initiation subunits in eukaryotic MSU RNAPs II. There are large numbers of conserved motifs are observed and some are found to be long stretches (highlighted). The yeast (S. cerevisiae) Rpb2 subunit (1224 amino acids) is used as the standard for numbering and highlighted in magenta. The catalytic region is found ~400 amino acids from the N-terminal. The invariant template binding YG is replaced by a KG pair in higher eukarvotes including human. The invariant catalytic R is placed at -9 from the template binding pair. This distance conservation is in close agreement with Pal and Luse findings that the transcription slippage abruptly stopped once the Rpb2 makes about 9 nts [32]. Interestingly, the transition from abortive to productive elongation cycle occurred once the RNAP register +10 nts [33]. The absolutely conserved R, which is implicated in NTP selection in SSU and MSU RNAPs and DNA polymerases, is placed at -5/6 positions. In fact, in all the eubacterial β subunits the catalytic R is placed at -7th position from the YG pair and completely conserved R is placed at -8th position downstream from the catalytic R. However, catalytic R is placed at -8th position from the YG pair the completely conserved R was at -4th position in SSU RNAPs and DNA polymerases [29,34]. This strongly suggests that the DNA polymerases, SSU and MSU RNAPs use the same set of amino acids for template, substrate binding and catalysis establishing a structurefunction relationship among the DNA and RNA polymerases. The immediate upstream amino acid from catalytic K in DNA polymerases is

usually a G or A [34], but in SSU viral RNA polymerases it is a Q [29] and in MSU eubacterial β subunits, it is a D in all [4] and in eukaryotic Rpb2 it is S/T, suggesting a possible role in the substrate binding and catalysis. Another catalytic like region is located in ~400 amino acids from the N-terminal but with a YG/FG pair and a catalytic R at -9 and long highly conserved stretches on both the sides. The Zn binding motif with 3 completely conserved Cs is found in the C-terminal region (highlighted in orange) which is followed by a long stretch of conservation till the C-terminal end. There are many highly conserved YG/FG/IG/LG pairs and a WG pair in the Rpb2 subunits. At least three of them may provide by three-point attachment for stereospecific binding on to the template DNA. However, the DNA polymerases show only one YG Interestingly, the RNA dependent RNAPs do not have the YG pair [35]. A long stretch of the Cterminal region is highly conserved in all Rpb2 subunits. Interestingly the higher eukaryotic subunits Rpb2 show completely а conserved motif till the end of the C-terminal whereas the lower eukaryotic Rpb2 subunits also show a completely conserved stretch but not covering all the amino acids till the C- terminal with a consensus sequence "PYAxKLLFQELMx M". A long stretch of highly conserved end suggests a possible role in the initiation process.

Fig. 4 shows the MSA analysis and conserved motifs in the Rpb1 elongation subunits in eukaryotic MSU RNAPs II. The yeast subunit with 1733 amino acids is used as standard and highlighted in magenta. Like the Rpb2 subunits, there are large numbers of conserved motifs and some are found to be long stretches (highlighted). However, the Rpb1 subunits are more conserved than the Rpb2 subunits. The 'template binding' pair is invariably an FG rather than a YG pair as reported in other MSU RNAPs. However, the catalytic R is completely conserved, including distance conservation in all the eukaryotic Rpb1 subunits and the upstream neighbour is mostly S/T but in lower eukaryotes like yeasts, it is N. Unlike in prokaryotic elongation subunit, the catalytic region is placed very close to the N-terminal (~ 100 amino acids). It is interesting to note that the distance between the catalytic R and the FG pair is 18 amino acids, i.e., placed exactly double the distance as compared to the Rpb2 initiation subunits. This distance closely agrees with the transcription bubble which extends ~ 20 bp from the TATA box [32]. The template binding and catalytic

CLUSTAL O (1.2.4) MSA of the Rpb2, initiation subunits of eukaryotic MSU RNAP II 3

```
MYDAD-EDMQYDE---DDDEITPDLWQEACWIVISSYFDEKGLVRQQLDSFDEFIQMSVQ
   spiP30876|RPB2 HUMAN
   tr|G3V8Y5|G3V8Y5_RAT
                                                MYDAD-EDMQYDE---DDDEITPDLWQEACWIVISSYFDEKGLVRQQLDSFDEFIG
   tr|A0A250Y753|A0A250Y753_CASCN
tr|A0A1U7R4C7|A0A1U7R4C7_MESAU
                                                MYDAD-EDMQYDE---DDDEITPDLWQEACW VISSYFDEKGLVROOL SFDEFICMSVQ
MYDAD-EDMQYDE---DDDEITPDLWQEACW VISSYFDEKGLVRQQL SFDEFICMSVQ
                                                                                                                                     56
                                                                                                                                     56
   tr|A0A286XIQ9|A0A286XIQ9_CAVPO
                                                MYDAD-EDMQYDE---DDDEITPDLWQEACWLVISSYFDEKGLVROQLSFDEFICMSVQ
MYDAD-EDMQYDE---DDDEITPDLWQEACWLVISSYFDEKGLVRQQLSFDEFICMSVQ
                                                                                                                                     56
   tr|I3M351|I3M351 ICTTR
                                                                                                                                     56
   tr|G7P5R6|G7P5R6_MACFA
tr|H2QP18|H2QP18_PANTR
                                                MYDAD-EDMOYDE---DDDEITPDLWOEACWVISSYPDEKGLVROOLSFDEFIOMSVO
MYDAD-EDMOYDE---DDDEITPDLWOEACWVISSYPDEKGLVROOLSFDEFIOMSVO
                                                                                                                                     56
   tr|A0A1U7V0T5|A0A1U7V0T5 TARSY
                                                MYDAD-EDMQYDE---DDDEITPDLWQEACWIVISSYFDEKGLVRQQLDSFDEFIQMSVQ
                                                MYDAD-EDMQYDE---DDDEITPDLWQEACW VISSY FDEKGLVROOL SFDEFIOMSVO
------MQYDE---DDDEITPDLWQEACW VISSY FDEKGLVROOLD SFDEFIOMSVO
------MQYDE---DDDEITPDLWQEACW VISSY FDEKGLVROOLD SFDEFIOMSVO
   tr|A0A1S2ZSL2|A0A1S2ZSL2_ERIEU
                                                                                                                                     56
   tr|A0A0D9QYL1|A0A0D9QYL1_CHLSB
tr|A0A2K5ZNR7|A0A2K5ZNR7_MANLE
                                                                                                                                     49
                                                                                                                                     49
   tr|A0A2I2ZIU3|A0A2I2ZIU3 GORGO
                                                MYDAD-EDMQYDE---DDDEITPDLWQEACWIVISSYFDEKGLVRQQLDSFDEFIGMSVQ
                                                                                                                                     56
   tr|A0A1D5QGA5|A0A1D5QGA5_MACMU
                                                MYDAD-EDMOYDE---DDDEITPDLWOEACWIVISSYFDEKGLVROOLISFDEFIOMSVO
MYDAD-EDMOYDE---DDDEITPDLWOEACWIVISSYFDEKGLVROOLISFDEFIOMSVO
   tr|A0A2J8S2N1|A0A2J8S2N1_PONAB
                                                                                                                                     56
   tr|A0A2K5K5J5|A0A2K5K5J5_COLAP
                                                MYDAD-EDMQYDE---DDDEITPDLWQEACWIVISSYFDEKGLVRQQLDSFDEFIC
   tr|A0A2J8PEW7|A0A2J8PEW7_PANTR
tr|A0A2K5CY83|A0A2K5CY83_AOTNA
                                                       ---MOYDE---DDDEITPDLWQEACWIVISSYFDEKGLVRQQLDSFDEFIOMSVQ
                                                                                                                                     49
                                                MYDAD-EDMQYDE---DDDEITPDLWQEACWIVISSYFDEKGLVRQQLDSFDEFIQMSVQ
                                                                                                                                     56
                                                MYDAD-EDMQYDE---DDDEITPDLWQEACWVISSYFDEKGLVRQQLSFDEFIOMSVQ
-----MQYDE---DDDEITPDLWQEACWVISSYFDEKGLVRQQLSFDEFIOMSVQ
   tr|A0A096NEY4|A0A096NEY4_PAPAN
                                                                                                                                     56
   tr|C9J2Y9|C9J2Y9_HUMAN
tr|G8BY61|G8BY61 TETPH
                                                                                                                                     49
                                                --MSQGEEYYADE DPYGFEDENA PISAEDSW VISSF REKGLVSOOL SFNOFVLYTLO
                                                                                                                                     58
   tr|A0A1X7QYA1|A0A1X7QYA1 9SACH
                                                                                                                                     56
                                                 ----MSNEEYYDEDPYGFEDESAPITAEDSWIVISSFFREKGLVSQQLISFNQFVIYTLQ
   tr|J7RV95|J7RV95_KAZNA
                                                                                                                                     56
   tr|H2AVJ8|H2AVJ8_KAZAF
sp|Q6FLD5|RPB2_CANGA
                                                 ----msnee yyeddpygfedesa pitaedswavissf frekglvsool s fnofvuytlo
--msadned yyde dpygfee ena pitaedtwavisaf frekglvsool s fnofvuytlo
                                                                                                                                     56
58
   tr|A0A0L8VHA5|A0A0L8VHA5_9SACH
tr|A0A0L8RB33|A0A0L8RB33_SACEU
                                                MSDLAN SEKYYDE DPYGFEDESA PITAEDSWAVISAF FREKGLVSOOLDS FNOFVDYTLO
                                                                                                                                     60
                                                MSDVENSEKYYEEDPYGFEDGSAPITAEDSWAVISAFIREKGLVSOOLISFNOFVDYTLO
MSAPGEEEYYDEDDVYGFEDENAPITAEDAWAVISSFIREKGLVSOOLISFNOFVDYTLO
                                                                                                                                     60
   tr|G0VJ71|G0VJ71_NAUCC
tr|G8ZM49|G8ZM49_TORDC
                                                                                                                                     60
                                                  -MSAAEDDYYNEDPYGYEDENSPITAEDSWAVISAFFREKGLVSQLLSFNQFVUYTLQ
                                                                                                                                     58
   tr|A0A1Q3A090|A0A1Q3A090_ZYGRO
tr|A0A0N7IS35|A0A0N7IS35_9SACH
tr|A0A212MG88|A0A212MG88_ZYGBA
                                                MSAAAVDED YYAE DPYGYEDENF PITAEDSWAVISAF FREKGLVSOOL SFNOFVLYTLO
-MSAAVNEE YYAE DPYGYDDETS PITAEDSWAVISAF FREKGLVSOOL SFNOFVLYTLO
                                                                                                                                     60
                                                                                                                                     59
                                                 -MSAAVNEDYYAEDPYGYEDENSPISAEDSWAVISAFFREKGLVSOOLDSFNQFVLYTLQ
-MSAAVNEDYYAEDPYGYEDENSPISAEDSWAVISAFFREKGLVSOOLDSFNQFVLYTLQ
   tr|A0A1S7HHE1|A0A1S7HHE1 9SACH
                                                                                                                                     59
   tr|S6ESB4|S6ESB4_ZYGB2
                                                 -MSAAVNEDYYAEDPYGYEDENSPISAEDSWAVISAFFREKGLVSQQLDSFNQFVDYTLQ
                                                tr|B6K5Q5|B6K5Q5_SCHJY
                                                                                                                                     47
   sp|Q02061|RPB2_SCHPO
tr|S9R8U4|S9R8U4_SCHOY
                                                                                                                                     47
   tr|S9W8C6|S9W8C6 SCHCR
                                                                                                                                     47
                                               RIVEDAPPIDLQAEAQHASGEVEEPPRYLLKFEQIYLSKPTHWERDGAPSPMMPNEARLR
sp|P30876|RPB2 HUMAN
                                                                                                                                      116
 rig3V8Y5ig3V8¥5 RAT
                                               RIVEDAPPIDLQAEAQHASGEVEEPPRYLLKFEQIYLSKPTHWERDGAPSPMMPNEARLF
                                                                                                                                      116
                                               RIVEDAPPIDLOAEAOHASGEVEEPPRYLLKFEOIYLSKPTHWERDGAPSPMMPNEARLF
RIVEDAPPIDLOAEAOHASGEVEEPPRYLLKFEOIYLSKPTHWERDGAPSPMMPNEARLF
tr|A0A250Y753|A0A250Y753 CASCN
                                                                                                                                      116
tr|A0A1U7R4C7|A0A1U7R4C7_MESAU
                                                                                                                                      116
tr|A0A286XIQ9|A0A286XIQ9 CAVPO
                                               RIVEDAPPIDLQAEAQHASGEVEEPPRYLLKFEQIYLSKPTHWERDGAPSPMMPNEARLI
                                                                                                                                      116
tr|I3M351|I3M351_ICTTR
tr|G7P5R6|G7P5R6_MACFA
                                               RIVEDAPPIDLQAEAQHASGEVEEPPRYLLKFEQIYLSKPTHWERDGAPSPMMPNEARLF
RIVEDAPPIDLQAEAQHASGEVEEPPRYLLKFEQIYLSKPTHWERDGAPSPMMPNEARLF
                                                                                                                                      116
                                                                                                                                      116
                                                                                                                                      116
tr|H2QPI8|H2QPI8 PANTR
                                               RIVEDAPPIDLQAEAQHASGEVEEPPRYLLKFEQIYLSKPTHWERDGAPSPMMPNEARLF
tria0a1U7V0T5ia0a1U7V0T5 TARSY
                                               RIVEDAPPIDLQAEAQHASGEVEEPPRYLLKFEQIYLSKPTHWERDGAPSPMMPNEARLF
RIVEDAPPIDLQAEAQHASGEVEEPPRYLLKFEQIYLSKPTHWERDGAPSPMMPNEARLF
                                                                                                                                      116
tr|A0A1S2ZSL2|A0A1S2ZSL2_ERIEU
                                                                                                                                       116
tr|A0A0D9QYL1|A0A0D9QYL1_CHLSB
                                               RIVEDAPPIDLOAEAOHASGEVEEPPRYLLKFEDIYLSKPTHWERDGAPSPMMPNEARLF
                                                                                                                                      109
                                               RIVEDAPPIDLOAEAOHASGEVEEPPRYLLKFEOIYLSKPTHWERDGAPSPMMPNEARLF
tr|A0A2K5ZNR7|A0A2K5ZNR7 MANLE
                                                                                                                                      109
tr|A0A2I2ZIU3|A0A2I2ZIU3_GORGO
                                               RIVEDAPPIDLQAEAQHASGEVEEPPRYLLKFEQIYLSKPTHWERDGAPSPMMPNEARL
                                                                                                                                      116
tria0a1D50GA5ia0a1D50GA5 MACMU
                                               RIVEDAPPIDLOAEAOHASGEVEEPPRYLLKFEDIYLSKPTHWERDGAPSPMMPNEARLF
                                                                                                                                      116
tr|A0A2J8S2N1|A0A2J8S2N1 PONAB
                                               RIVEDAPPIDLOAEAOHASGEVEEPPRYLLKFEOIYLSKPTHWERDGAPSPMMPNEARLF
                                                                                                                                      116
tr|A0A2K5K5J5|A0A2K5K5J5 COLAP
                                               RIVEDAPPIDLQAEAQHASGEVEEPPRYLLKFEQIYLSKPTHWERDGAPSPMMPNEARLF
                                                                                                                                      116
tr|A0A2J8PEW7|A0A2J8PEW7 PANTR
                                               RIVEDAPPIDLQAEAQHASGEVEEPPRYLLKFEQIYLSKPTHWERDGAPSPMMPNEARLF
                                                                                                                                      109
tr|A0A2K5CY83|A0A2K5CY83 AOTNA
                                               RIVEDAPPIDLQAEAQHASGEVEEPPRYLLKFEQIYLSKPTHWERDGAPSPMMPNEARLF
                                                                                                                                      116
tr|A0A096NEY4|A0A096NEY4 PAPAN
                                               RIVEDAPPIDLQAEAQHASGEVEEPPRYLLKFEQIYLSKPTHWERDGAPSPMMPNEARLF
                                                                                                                                      116
tr|C9J2Y9|C9J2Y9_HUMAN
                                               RIVEDAPPIDLOAEAOHASGEVEEPPRYLLKFEDIYLSKPTHWERDGAPSPMMPNEARLF
                                                                                                                                      109
tr|G8BY61|G8BY61 TETPH
                                               DIISEDSTLILEQLAQHTTESDNISRKYEISFGKIYVTKPNVNESDGVTHALYPQEARLF
                                                                                                                                      118
tr|A0A1X7QYA1|A0A1X7QYA1 9SACH
                                               DIIFEDSTLILEQIAQHTTEQDNISRKYEISFGKIYVTKPMVNESDGVTHALYPQEARLF
tr|J7RV95|J7RV95 KAZNA
                                               DIISEDSRLILEOLAOHTTEADNISRKYEISFGKIYVTKPMVNESDGVTHALMPOESRLF
                                                                                                                                      116
tr|H2AVJ8|H2AVJ8_KAZAF
                                               DIISEDSTLILEQLAQHTTEADNISRKYEISFGKIYVTKPNVNESDGVTHALYPQEARLF
                                                                                                                                      116
sp | Q6FLD5 | RPB2_CANGA
                                               DIISEDSTLILEQLAQHTTEQDNISRKYEISFGMIYVTKPMVNESDGVTHALYPQEARL
                                                                                                                                      118
tr|A0A0L8VHA5|A0A0L8VHA5 9SACH
                                               DIICEDSTLILEQLAQHTTESDNISRKYEISFGMIYVTKPMVNESDGVTHALYPQEARL
                                                                                                                                       120
tr|A0A0L8RB33|A0A0L8RB33_SACEU
                                               DIICEDSTLILEQLAXHTTESDNISRKYEISFGKIYVTKPMVNESDGVTHALYPQEARLF
                                                                                                                                      120
tr|GOVJ71|GOVJ71 NAUCC
tr|G8ZM49|G8ZM49 TORDC
tr|A0A1Q3A090|A0A1Q3A090 ZYGRO
tr|A0A0N7IS35|A0A0N7IS35|9SACH
                                               DIISEDSTLILEOLAOHTTETDNVSRKYEIGFGKIYVTKPMVNESDGVTHALYPOEARL
                                                                                                                                      120
                                               DIISEDSTLILEQLAQHTTEADNISRKYEISFGKIYVTKPNVNESDGVTHALYPQEARLF
                                                                                                                                      118
                                               DIISEDSTLILEQLACHTTESDNISRKYEISFGKIYVTKPWVNESDGVTHALYPCEARLF
DIISEDSTLILEQLACHTTESDNISRKYEIGFGKIYVTKPWVNESDGVTHALYPCEARLF
                                                                                                                                      120
                                                                                                                                      119
                                               DIISEDSTLILEQLAQHTTESDNISRKYEIGFGKIYVTKPMVNESDGVTHALYPQEARL
tr|A0A212MG88|A0A212MG88_ZYGBA
                                                                                                                                      119
tr|A0A1S7HHE1|A0A1S7HHE1_9SACH
tr|S6ESB4|S6ESB4_ZYGB2
                                               DIISEDSTLILEQLAOHTTESDNISRKYEIGFGKIYVTKPMVNESDGVTHALYPQEARLF
DIISEDSTLILEQLAOHTTESDNISRKYEIGFGKIYVTKPMVNESDGVTHALYPQEARLF
                                                                                                                                      119
                                                                                                                                      119
tr|B6K5Q5|B6K5Q5_SCHJY
                                               eivdddstltldqyaqhtgaqgdytrryeinfgqiylsrptmteadgstttmfpqearli
                                                                                                                                       107
                                               sp|Q02061|RPB2_SCHPO
tr|S9R8U4|S9R8U4 SCHOY
                                                                                                                                      107
                                                                                                                                      107
tr|S9W8C6|S9W8C6 SCHCR
                                                                                                                                       107
```

```
NLTYSA PLYVD TKTVIKEG-----EEQLQTQHQKFIGKIPIMLRS YCL
NLTYSA PLYVD TKTVIKEG-----EEQLQTQHQKTFIGKIPIMLRS YCL
NLTYSA PLYVD TKTVIKEG-----EEQLQTQHQKTFIGKIPIMLRS YCL
NLTYSA PLYVD TKTVIKEG-----EEQLQTQHQKTFIGKIPIMLRS YCL
  sp|P30876|RPB2 HUMAN
  tr|G3V8Y5|G3V8Y5 RAT
                                                                                                                                             162
  tr|A0A250Y753|A0A250Y753 CASCN
                                                                                                                                             162
  tr|A0A1U7R4C7|A0A1U7R4C7 MESAU
                                                                                                                                             162
                                                  NLTYSAPLYVDITKTVIKEG------EEQLQTQHQKIFIGKIPIMLRSTYCL
  tr|A0A286XIQ9|A0A286XIQ9 CAVPO
                                                                                                                                             162
                                                  NLTYSAPLYVDITKTVIKEG-----EEOLOTOHOKIFIGKIPIMLRSTYCL
  tr||T3M351||T3M351 | TCTTR
                                                                                                                                             162
                                                  tr|G7P5R6|G7P5R6 MACFA
tr|H2QPI8|H2QPI8 PANTR
                                                                                                                                             162
                                                                                                                                             162
  tr|A0A1U7V0T5|A0A1U7V0T5 TARSY
                                                                                                                                             162
  tr|A0A1S2ZSL2|A0A1S2ZSL2_ERIEU
                                                                                                                                             162
  tr|A0A0D9QYL1|A0A0D9QYL1 CHLSB
                                                                                                                                             155
  tr|A0A2K5ZNR7|A0A2K5ZNR7_MANLE
                                                                                                                                             155
                                                  NLTYSAPLYVDITKTVIKEG------EEQLQTQHQKIFIGKIPIMLRSTYCL
  tr|A0A2I2ZIU3|A0A2I2ZIU3 GORGO
                                                                                                                                             162
                                                  NLTYSABLYVDITKTVIKEG-----EEOLOTOHOKTFIGKIPIMLRSTYCL
  trlA0A1D50GA5LA0A1D50GA5 MACMU
                                                                                                                                             162
                                                  NLTYSAPLYVDITKTVIKEG-----EEQLQTQHQKTFIGKIPIMLRSTYCL
NLTYSAPLYVDITKTVIKEG-----EEQLQTQHQKTFIGKIPIMLRSTYCL
NLTYSAPLYVDITKTVIKEG-----EEQLQTQHQKTFIGKIPIMLRSTYCL
  tr|A0A2J8S2N1|A0A2J8S2N1 PONAB
                                                                                                                                             162
  tr | A0A2K5K5J5 | A0A2K5K5J5 COLAP
                                                                                                                                             162
  tr|A0A2J8PEW7|A0A2J8PEW7 PANTR
                                                  NLTYSAPLYVDITKTVIKEG-----EGOLOTOHOKIFIGKIPIMLRSTYCL
NLTYSAPLYVD TKTVIKEG------EGOLOTOHOKIFIGKIPIMLRSTYCL
  tr|A0A2K5CY83|A0A2K5CY83 AOTNA
                                                                                                                                             162
  tr|A0A096NEY4|A0A096NEY4 PAPAN
                                                                                                                                             162
                                                  NLTYSAPLYVDITKTVIKEG------EEQLQTQHQKTFIGKIPIMLRSTYCL
  tr|C9J2Y9|C9J2Y9 HUMAN
                                                                                                                                             155
                                                  NLTYSS LFVDVKRTYEATDVPGRELKYNLIAEESEDSENDKVFIGRLPVMLRSKNCY
NLTYSSLFVDVKRTYEAVDVPGRELKYELIAEESEDDSESGKVFIGRLPIMLRSKNCY
 tr|G8BY61|G8BY61_TETPH
tr|A0A1X7QYA1|A0A1X7QYA1 9SACH
                                                                                                                                             178
                                                                                                                                             176
  tr|J7RV95|J7RV95 KAZNA
                                                  NLTYSSGLFVDVKKRTYEAIDVPGRDLKYELIAEESEDDSERGKVFIGRLPIMLRSKNCY
                                                                                                                                             176
  tr|H2AVJ8|H2AVJ8 KAZAF
                                                  NLTYSSGLFVDVRKRTYEAVDVPGRDLKYELIAEESEDDSESGKVFIGRLPIMLRSKNCY
                                                                                                                                             176
  sp|Q6FLD5|RPB2 CANGA
                                                   NLTYSSGLFVDVTKRTYEAVDVPGRDLNYQLIAEESEEDSESGKVFIGRLPIMLRSKNCY
  tr|A0A0L8VHA5|A0A0L8VHA5 9SACH
                                                  NLTYSSGLFVDVKKRTYEAIDVPGRELKYELIAEESEDDSESGKVFIGRLPIMLRSKNCY
                                                                                                                                             180
  tr|A0A0L8RB33|A0A0L8RB33 SACEU
                                                  NLTYSS LFVDVKKRTYEAVDVPGRELKYELIAEESEDDSESGKVFIGRLPIMLRSKNCY
                                                                                                                                             180
  tr|G0VJ71|G0VJ71 NAUCC
                                                  NLTYSSGLFVDVKKRTYEAVDVPGRELKYELIAEESEESESGKVFIGRLPIMLRSKNCY
                                                                                                                                             180
  tr|G8ZM49|G8ZM49 TORDC
                                                  NLTYSSGLFVDVKKRTYEAVDVPGRELKYELIAEESEDDSESGKVFIGRLPIMLRSKNCY
                                                                                                                                             178
  tr|A0A1Q3A090|A0A1Q3A090 ZYGRO
                                                  NLTYSSGLFVDVKKRTHEAVDIPGRELKYEMIAEESENDSDSGKVFIGRLPIMLRSKNCY
                                                                                                                                             180
  tr|A0A0N7IS35|A0A0N7IS35_9SACH
tr|A0A212MG88|A0A212MG88_ZYGBA
                                                  NLTYSSELFVDVKKRTYEAVDLPGRELKYELIAEESE-EGDSGKVFIGRLPIMLRSKNCY
                                                                                                                                             178
                                                  NLTYSSCLFVDVKKRTYEAVDLPGRELKYELIAEESEDDSDSGKVFIGRLPIMLRSKNCY
  tr|A0A1S7HHE1|A0A1S7HHE1 9SACH
                                                  NLTYSSGLFVDVKKRTYEAVDLPGRELKYELIAEESEDDSDSGKVFIGRLPIMLRS
                                                                                                                                             179
  tr|S6ESB4|S6ESB4_ZYGB2
tr|B6K5Q5|B6K5Q5_SCHJY
                                                  NLTYSSGLFVDVKKRTYEAVDLPGRELKYELIAEESEDDSDSGKVFIGRLPIMLRSKNCY
                                                                                                                                             179
                                                  NLTYSS LYVDMRKKVMRAQDSNVPIGEEIW--LTEEEDDEPSKVFIGKIPIMLRSTFCI
                                                                                                                                             165
                                                  NLTYSS LYVDNRKKVMVAADSNVPIGEEEW--LVEEEDEEPSK FIGKIPIMLRS FCI
NLTYSS LYVDNRKKVMVAADSNVPIGEEEW--LVEEEDEDPSK FIGKIPIMLRS FCI
  sp|Q02061|RPB2_SCHPO
                                                                                                                                             165
  tr|S9R8U4|S9R8U4 SCHOY
                                                                                                                                             165
  tr|S9W8C6|S9W8C6 SCHCR
                                                  NLTYSSELYVDNRKKVMVAADSNVPIGEEEW--LVEEEDEDPSKVFIGKIPIMLRSTFCI
                                                                                                                                             165
                                                    ****:
                                                                                                 .::: * ***::*:***
                                                 LNGLTDRDLCTLNECTLDEGGYFIINGSEKVLIAQEKMATNTVYVFAKKD-SKYAYTGECLNGLTDRDLCTLNECTLDEGGYFIINGSEKVLIAQEKMATNTVYVFAKKD-SKYAYTGEC
sp|P30876|RPB2_HUMAN
                                                                                                                                             221
triG3V8Y5IG3V8Y5 RAT
                                                                                                                                            221
tr|A0A250Y753|A0A250Y753_CASCN
                                                 LNGLTDRDLCILNECPLDEGGYFIINGSEKVLIAQEKMATNTVYVFAKKD-SKYAYTGEC
LNGLTDRDLCILNECPLDEGGYFIINGSEKVLIAQEKMATNTVYVFAKKD-SKYAYTGEC
                                                                                                                                             221
tr|A0A1U7R4C7|A0A1U7R4C7 MESAU
                                                                                                                                             221
tr|A0A286XIQ9|A0A286XIQ9_CAVPO
                                                 lngltdrdlctlnectldeggyfiingsekvliaqekmatntvyvfakkd-skyaytgec
                                                                                                                                             221
                                                 LINGLIDRDLCILINECFLDFGGYFIINGSEKVLIAQEKMAINTVYVFAKKD-SKYAYTGEC
LINGLIDRDLCILINECFLDFGGYFIINGSEKVLIAQEKMAINTVYVFAKKD-SKYAYTGEC
LINGLIDRDLCILINECFLDFGGYFIINGSEKVLIAQEKMAINTVYVFAKKD-SKYAYTGEC
LINGLIDRDLCILINECFLDFGGYFIINGSEKVLIAQEKMAINTVYVFAKKD-SKYAYTGEC
tr|I3M351|I3M351_ICTTR
                                                                                                                                             221
tr|G7P5R6|G7P5R6_MACFA
                                                                                                                                            221
tr|H2QPI8|H2QPI8_PANTR
tr|A0A1U7V0T5|A0A1U7V0T5_TARSY
                                                                                                                                             221
                                                                                                                                             221
tr|A0A1S2ZSL2|A0A1S2ZSL2_ERIEU
                                                 LNGLTDRDLCELNECPLDEGGYFIINGSEKVLIAQEKMATNTVYVFAKKD-SKYAYTGEC
                                                 LNGLTDRDLCZLNECPLDEGGYFIINGSEKVLIAQEKMATNTVYVFAKKD-SKYAYTGEC
LNGLTDRDLCZLNECPLDEGGYFIINGSEKVLIAQEKMATNTVYVFAKKD-SKYAYTGEC
LNGLTDRDLCZLNECPLDEGGYFIINGSEKVLIAQEKMATNTVYVFAKKD-SKYAYTGEC
tr|A0A0D9QYL1|A0A0D9QYL1_CHLSB
                                                                                                                                             214
                                                                                                                                            214
tr|A0A2K5ZNR7|A0A2K5ZNR7_MANLE
tr|A0A2I2ZIU3|A0A2I2ZIU3 GORGO
                                                                                                                                             221
tr|A0A1D5QGA5|A0A1D5QGA5 MACMU
                                                 LNGLTDRDLCELNECPLDFGGYFIINGSEKVLIAQEKMATNTVYVFAKKD-SKYAYTGEC
                                                                                                                                             221
tr|A0A2J8S2N1|A0A2J8S2N1_PONAB
                                                 lngltdrdlctlnecpldfggyfiingsekvliaqekmatntvyvfakkd-skyaytgec
tr|A0A2K5K5J5|A0A2K5K5J5_COLAP
                                                 LNGLTDRDLCELNECPLDFGGYFIINGSEKVLIAQEKMATNTVYVFAKKD-SKYAYTGEC
                                                                                                                                             221
tr|A0A2J8PEW7|A0A2J8PEW7_PANTR
tr|A0A2K5CY83|A0A2K5CY83_AOTNA
                                                 INGLIDRDLCLINEC LDEGGYFIINGSEKVLIAGEKMATNTVYVFAKKD-SKYAYTGEC
LNGLIDRDLCLINEC LDEGGYFIINGSEKVLIAGEKMATNTVYVFAKKD-SKYAYTGEC
                                                                                                                                            214
                                                                                                                                             221
tr|A0A096NEY4|A0A096NEY4 PAPAN
                                                 LNGLTDRDLCELNECPLDEGGYFIINGSEKVLIAQEKMATNTVYVFAKKD-SKYAYTGEC
                                                                                                                                             221
tr|C9J2Y9|C9J2Y9_HUMAN
                                                 LNGLTDRDLCELNECPLDEGGYFIINGSEKVLIAQEKMATNTVYVFAKKD-SKYAYTGEC
                                                                                                                                             214
                                                 LSDATE SDLYKLKEC F FDMGGYFIINGSEKVLIAQER SAGNIVQVFKKAA PSPISHVAEI
LSDATE SDLYKLKEC FFDMGGYFIINGSEKVLIAQER SAGNIVQVFKKAA PSPISHVAEI
LSDATE SDLYKLKEC FFDMGGYFIINGSEKVLIAQER SAGNIVQVFKKAA PSPISHVAEI
LSDATE SDLYKLKEC FFDMGGYFIINGSEKVLIAQER SAGNIVQVFKKAA PSPISHVAEI
tr|G8BY61|G8BY61_TETPH
tr|A0A1X7QYA1|A0A1X7QYA1_9SACH
                                                                                                                                             238
                                                                                                                                            236
tr|J7RV95|J7RV95_KAZNA
tr|H2AVJ8|H2AVJ8_KAZAF
                                                                                                                                             236
                                                                                                                                            236
sp|Q6FLD5|RPB2 CANGA
                                                 LSDATE SDLYKLKECP FDMGGYFIINGSEKVLI AQERSAGNIVQVFKKAA PSPI SHVAEI
                                                 LSEATESDLYKLKECFFDMGGYFIINGSEKVLIAQERSAGNIVQVFKKAAPSPISHVAEI
LSEATESDLYKLKECFFDMGGYFIINGSEKVLIAQERSAGNIVQVFKKAAPSPISHVAEI
tr|A0A0L8VHA5|A0A0L8VHA5_9SACH
                                                                                                                                            240
tr|A0A0L8RB33|A0A0L8RB33 SACEU
                                                                                                                                            240
tr|G0VJ71|G0VJ71_NAUCC
tr|G8ZM49|G8ZM49_TORDC
                                                 LSDATE SDLYKLKEC PFDMGGYFIINGSEKVLIAQER SAGNIVQVFKKAA PSPISHVAEI
                                                                                                                                            240
                                                 lsdateldlyklkecpfdmggyfiingsekvliaqersagnivqvfkkaapspishvaei
                                                                                                                                             238
                                                 LSDATESDLYKLKECFFDMGGYFIINGSEKVLIAQERSAGNIVQVFKKAAPSPISHVAEI
LSDATESDLYKLKECFFDMGGYFIINGSEKVLIAQERSAGNIVQVFKKAAPSPISHVAEI
tr|A0A1Q3A090|A0A1Q3A090_ZYGRO
                                                                                                                                            240
tr|A0A0N7IS35|A0A0N7IS35_9SACH
                                                                                                                                            238
                                                 LSDATE SDLYKLKECY FDMGGYFIINGSEKVLIAQER SAGNIVQVYKKAAPSPISHVAEI
LSDATE SDLYKLKECY FDMGGYFIINGSEKVLIAQER SAGNIVQVFKKAAPSPISHVAEI
LSDATE SDLYKLKECY FDMGGYFIINGSEKVLIAQER SAGNIVQVFKKAAPSPISHVAEI
tr|A0A212MG88|A0A212MG88_ZYGBA
                                                                                                                                             239
tr|A0A1S7HHE1|A0A1S7HHE1_9SACH
                                                                                                                                             239
tr|S6ESB4|S6ESB4_ZYGB2
                                                                                                                                             239
                                                 LNGVSDSELYDLNECPYDOGGYFIINGSEKVIIAQERSAANIVQVFRKAAPSPIAYVAEI
tr|B6K5Q5|B6K5Q5_SCHJY
                                                                                                                                            225
sp|Q02061|RPB2 SCHPO
                                                 LNGVSDSELYDLNECPYDOGGYFIINGSEKVIIAQERSAANIVQVFKKAAPSPIAYVAEI
LNGVSDAELYDLNECPYDOGGYFIINGSEKVIIAQERSAANIVQVFRKAAPSPIALVAEI
LNGVSDAELYDLNECPYDOGGYFIINGSEKVIIAQERSAANIVQVFRKAAPSPVAYVAEI
                                                                                                                                            225
tr|S9R8U4|S9R8U4_SCHOY
tr|S9W8C6|S9W8C6_SCHCR
                                                                                                                                             225
                                                               *: **
                                                      :: :*
                                                                           *********
```

```
sp|P30876|RPB2 HUMAN
                                                 rsclenssrptstiwvsmlarggqgakksaigqrivatlpyikqevpiiivfr<mark>al</mark>ffysd
                                                                                                                                            281
                                                 RSCLENSSRPTSTIWVSMLARGGQGAKKSAIGQRIVATLPYIKQEVPIIIVFR<mark>ALG</mark>FVSD
tr|G3V8Y5|G3V8Y5 RAT
                                                                                                                                            281
tr|A0A250Y753|A0A250Y753_CASCN
tr|A0A1U7R4C7|A0A1U7R4C7_MESAU
                                                 RSCLENSSRPTSTIWVSMLARGGQGAKKSAIGQRIVATLPYIKQEVPIIIVFR
                                                                                                                                            281
                                                 RSCLENSSRPTSTIWVSMLARGGQGAKKSAIGQRIVATLPYIKQEVPIIIVFR
                                                                                                                                            281
tr|A0A286XIQ9|A0A286XIQ9_CAVPO
                                                 RSCLENSSR PTST IWV SMLARGGQGAKKSA IGQRIVATLP YIKQEVP I I I VFR
                                                                                                                                            281
tr|I3M351|I3M351_ICTTR
                                                 RSCLENSSRPTSTIWVSMLARGGQGAKKSAIGQRIVATLPYIKQEVPIIIVFR
                                                                                                                            GEVSD
                                                                                                                                            281
                                                RSCLENSSRPTSTIWVSMLARGGOGAKKSAIGORIVATLPYIKOEVPIIIVFR
RSCLENSSRPTSTIWVSMLARGGOGAKKSAIGORIVATLPYIKOEVPIIIVFR
tr|G7P5R6|G7P5R6_MACFA
                                                                                                                            GFVSD
                                                                                                                                            281
tr|H2OPI8|H2OPI8 PANTR
                                                                                                                            GFVSD
                                                                                                                                            281
tr|A0A1U7V0T5|A0A1U7V0T5 TARSY
                                                 RSCLENSSRPTSTIWVSMLARGGQGAKKSAIGQRIVATLPYIKQEVPIIIVFR
                                                                                                                            GEVSD.
                                                                                                                                            281
tr|A0A1S2ZSL2|A0A1S2ZSL2_ERIEU
                                                 RSCLENSSRPTSTIWVSMLARGGQGAKKSAIGQRIVATLPYIKQEVPIIIVFR
                                                                                                                            FVSD
                                                                                                                                            281
                                                RSCLENSSRPTSTIWVSMLARGGQGAKKSAIGQRTVATLPYIKQEVPIIIVFR
RSCLENSSRPTSTIWVSMLARGGQGAKKSAIGQRTVATLPYIKQEVPIIIVFR
tr|A0A0D9QYL1|A0A0D9QYL1_CHLSB
tr|A0A2K5ZNR7|A0A2K5ZNR7_MANLE
                                                                                                                            GEVSD
                                                                                                                                            274
tr|A0A2I2ZIU3|A0A2I2ZIU3 GORGO
                                                RSCLENSSRPTSTIWVSMLARGGQGAKKSAIGQRIVATLPYIKQEVPIIIVFR
                                                                                                                            GFVSD
                                                                                                                                            281
                                                RSCLENSSR PTSTIWVSMLARGGQGAKKSAIGQRIVATLPYIKQEVPIIIVFR
RSCLENSSR PTSTIWVSMLARGGQGAKKSAIGQRIVATLPYIKQEVPIIIVFR
tr|A0A1D5QGA5|A0A1D5QGA5_MACMU
                                                                                                                            GEVSD.
                                                                                                                                            281
tr|A0A2J8S2N1|A0A2J8S2N1_PONAB
                                                                                                                            GEVSD
                                                                                                                                            281
                                                RSCLENSSRPTSTIWVSMLARGGOGAKKSAIGORIVATLPYIKOEVPIIIVFR
RSCLENSSRPTSTIWVSMLARGGOGAKKSAIGORIVATLPYIKOEVPIIIVFR
tr|A0A2K5K5J5|A0A2K5K5J5_COLAP
tr|A0A2J8PEW7|A0A2J8PEW7_PANTR
                                                                                                                            GFVSD
                                                                                                                                            281
                                                                                                                                            274
tr|A0A2K5CY83|A0A2K5CY83_AOTNA
                                                 RSCLENSSRPTSTIWVSMLARGGQGAKKSAIGQRIVATLPYIKQEVPIIIVFR
                                                                                                                                            281
                                                                                                                            GFVSD
                                                RSCLENSSRPTSTIWVSMLARGGQGAKKSAIGQRIVATLPYIKQEVPIIIVFRA
RSCLENSSRPTSTIWVSMLARGGQGAKKSAIGQRIVATLPYIKQEVPIIIVFRA
RSALEKGSRFISTLQVKLYGRES-----SSTRTIKATLPYIKQDIPIVIIFRA
tr|A0A096NEY4|A0A096NEY4_PAPAN
                                                                                                                            GFVSD.
                                                                                                                                            281
                                                                                                                            GFVSD
GIIPD
tr|C9J2Y9|C9J2Y9_HUMAN
                                                                                                                                            274
tr|G8BY61|G8BY61 TETPH
                                                                                                                                            292
tr|A0A1X7QYA1|A0A1X7QYA1 9SACH
                                                 RSALEKGSRFISTLQVKLYGREG----STSRTIKATLPYIKQDIPIVIIFR
                                                                                                                            GI TPD
                                                                                                                                            290
tr|J7RV95|J7RV95 KAZNA
                                                 RSALEKGSRFISTLQVKLYGREG----SSSRTIKATLPYIKQDIPIVIIFRA
                                                                                                                                            290
tr|H2AVJ8|H2AVJ8_KAZAF
                                                RSALEKGSRFISTLOVKLYGRES-----SSSRTIKATLPYIKODIPIVIIFR
RSALEKGSRFISTLOVKLYGRES-----SSARTIKATLPYIKODIPIVIIFR
                                                                                                                                            290
sp | Q6FLD5 | RPB2_CANGA
                                                                                                                           TÆITPD
                                                                                                                                            292
tr|A0A0L8VHA5|A0A0L8VHA5_9SACH
                                                 RSALEKGSRFISTLOVKLYGREG----SSARTIKATLPYIKODIPIVIIFRALGIIPD
                                                                                                                                            294
tr|A0A0L8RB33|A0A0L8RB33 SACEU
                                                 RSALEKGSRFISTLQVKLYGREG----SSARTIKATLPYIKQDIPIVIIFRALGIIPD
                                                                                                                                            294
                                                                                                                            GILDD
tr|G0VJ71|G0VJ71_NAUCC
                                                 RSALEKGSRFISTLQVKLYGREG----SDARTIMATLPYIKQDIPIVIIFR
                                                                                                                                            294
tr|G8ZM49|G8ZM49 TORDC
                                                 RSALEKGSRFISTLQVKLYGREG----SSARTIKATLPYIKQDIPIVIIFR
                                                                                                                            GI TPD
                                                                                                                                            292
                                                 RSALEKGSRFISTLQVKLFAREG----NSSRTIKATLPYIKQDIPIVIIFR
tr|A0A1Q3A090|A0A1Q3A090_ZYGRO
                                                                                                                            GITPD
                                                                                                                                            294
                                                RSALEKGSRFISTLOVKLFAREG----NSSRTIKATLPYIKODIPIVIIFR
tr|A0A0N7IS35|A0A0N7IS35_9SACH
                                                                                                                            SILPD
                                                                                                                                            292
tr|A0A212MG88|A0A212MG88_ZYGBA
                                                RSALEKGSRFISTLOVKLFAREG----NSSRTIKATLPYIKODIPIVIIFRALGIIPD
RSALEKGSRFISTLOVKLFAREG----NSSRTIKATLPYIKODIPIVIIFRALGIIPD
                                                                                                                                            293
tr|A0A1S7HHE1|A0A1S7HHE1 9SACH
                                                                                                                                            293
                                                 RSALEKGSRFISTLQVKLFAREG----NSSRTIKATLPYIKQDIPIVIIFRA
tr|S6ESB4|S6ESB4_ZYGB2
                                                                                                                                            293
tr|B6K5Q5|B6K5Q5_SCHJY
                                                 RSALERGSRLISSMQIKLMARSTE----NSGQTIKATLPYIRSDIPIVVVFR<mark>ALE</mark>VVPD
                                                                                                                                            280
                                                sp|Q02061|RPB2_SCHPO
tr|S9R8U4|S9R8U4_SCHOY
                                                                                                                                            280
                                                                                                                                            280
tr|S9W8C6|S9W8C6 SCHCR
                                                                                                                                            280
sp|P30876|RPB2 HUMAN
                                                    <mark>ILEHI</mark>IYDFEDPEMMEMVKPSLDEAFVIQEQNVALNF<mark>IG</mark>SRGAKPGVTKEKRIKYAKE
                                                                                                                                             341
                                                 RDILEHIIYDFED PEMMEMVKPSLDEAFVIQEQNVALNFI
RDILEHIIYDFED PEMMEMVKPSLDEAFVIQEQNVALNFI
RDILEHIIYDFED PEMMEMVKPSLDEAFVIQEQNVALNFI
RDILEHIIYDFED PEMMEMVKPSLDEAFVIQEQNVALNFI
tr|G3V8Y5|G3V8Y5_RAT
                                                                                                         <mark>g</mark>ergakpgvtkekrikyake
                                                                                                                                             341
tr|A0A250Y753|A0A250Y753_CASCN
                                                                                                          5RGAKPGVTKEKRIKYAKE
                                                                                                                                             341
tr|A0A1U7R4C7|A0A1U7R4C7_MESAU
                                                                                                         SRGAKPGVTKEKR TKYAKE
                                                                                                                                             341
tr|A0A286XIQ9|A0A286XIQ9_CAVPO
                                                                                                          SRGAKPGVTKEKRIKYAKE
                                                                                                                                             341
                                                 ROLLEH LYDFED PEMMEMY KPS LDE AFVI QEQNYALNE

RD YUPFED PEMMEMY KPS LDE AFVI QEQNYALNE

RDILEH LYDFED PEMMEMV KPS LDE AFVI QEQNYALNE
                                                                                                          SRGAKPGVTKEKRIKYAKE
tr|I3M351|I3M351_ICTTR
tr|G7P5R6|G7P5R6_MACFA
                                                                                                          SRGAKPGVTKEKRIKYAKE
                                                                                                                                             341
tr|H2QPI8|H2QPI8_PANTR
                                                                                                          BRGAKPGVTKEKRIKYAKE
                                                                                                                                             341
                                                                                                          SRGAKPGVTKEKRIKYAKE
tr|A0A1U7V0T5|A0A1U7V0T5 TARSY
                                                                                                                                             341
tr|A0A1S2ZSL2|A0A1S2ZSL2 ERIEU
                                                                                                          SRGAKPGVTKEKRIKYAKE
                                                                                                                                             341
                                                           IYDFED PEMMEMVKPSLDEAFVIQEQUVALNF
tr|A0A0D9QYL1|A0A0D9QYL1 CHLSB
                                                                                                          SRGAKPGVTKEKRIKYAKE
                                                                                                                                             334
                                                           IYDFED PEMMEMVKPS LDE AFVI QEQNVALNE
IYDFED PEMMEMVKPS LDE AFVI QEQNVALNE
IYDFED PEMMEMVKPS LDE AFVI QEQNVALNE
tr|A0A2K5ZNR7|A0A2K5ZNR7_MANLE
                                                                                                          SRGAKPGVTKEKRIKYAKE
                                                  RDILEH
                                                                                                                                             334
tr|A0A2I2ZIU3|A0A2I2ZIU3_GORGO
                                                  RDILEH
                                                                                                          SRGAKPGVTKEKRIKYAKE
                                                                                                                                             341
tr|A0A1D5QGA5|A0A1D5QGA5_MACMU
                                                                                                          SRGAKPGVTKEKRIKYAKE
                                                  RDILEH:
                                                                                                                                             341
tr|A0A2J8S2N1|A0A2J8S2N1 PONAB
                                                           IYDFED PEMMEMVKPSLDE AFVI QEQ NVALNF
                                                  ROTLEH
                                                                                                          SRGAKPGVTKEKRIKYAKE
                                                                                                                                             341
                                                           IYDFEDPEMMEMVKPSLDEAFVIQEONVALNE
IYDFEDPEMMEMVKPSLDEAFVIQEONVALNE
IYDFEDPEMMEMVKPSLDEAFVIQEONVALNE
IYDFEDPEMMEMVKPSLDEAFVIQEONVALNE
IYDFEDPEMMEMVKPSLDEAFVIQEONVALNE
IYDFEDPEMMEMVKPSLDEAFVIQEONVALNE
tr|A0A2K5K5J5|A0A2K5K5J5_COLAP
                                                  RDILEH
                                                                                                          BRGAKPGVTKEKRIKYAKE
                                                  RDILEH:
tr|A0A2J8PEW7|A0A2J8PEW7_PANTR
                                                                                                          SRGAKPGVTKEKRIKYAKE
                                                                                                                                             334
tr|A0A2K5CY83|A0A2K5CY83 AOTNA
                                                  ROTLEH
                                                                                                          SR GAKPGVT KEKR TKY AKE
                                                                                                                                             341
tr|A0A096NEY4|A0A096NEY4_PAPAN
                                                                                                          SRGAKPGVTKEKRIKYAKE
                                                  RDILEH
                                                                                                                                             341
                                                                                                          SRGAKPGVTKEKRIKYAKE
tr|C9J2Y9|C9J2Y9_HUMAN
                                                  RDILEH
                                                                                                                                             334
tr|G8BY61|G8BY61_TETPH
                                                           CYDVNDWQMLEMLKPCVEDGFVIQDRETALDF
                                                                                                          RRGTALGINKEKR IQYARI
                                                  GEILEH
                                                                                                                                             352
                                                           TOVNOW MILEMIKE CVEDGEVI QDREJALDE
CYDVNOW MILEMIKE CVEDGEVI QDREJALDE
CYDVNOW MILEMIKE CVEDGEVI QDREJALDE
CYDVNOW MILEMIKE CVEDGEVI QDREJALDE
CYDVNOW MILEMIKE CVEDGEVI QDREJALDE
tr|A0A1X7QYA1|A0A1X7QYA1_9SACH
                                                                                                          rrgtalginkekriqyakt
                                                                                                                                             350
tr|J7RV95|J7RV95_KAZNA
                                                  GEILEH
                                                                                                          RRGTALGINKEKRIOYAKT
                                                                                                                                             350
tr|H2AVJ8|H2AVJ8 KAZAF
                                                  GEILEH
                                                                                                          RRGTALGINKEKRIOYAKI
                                                                                                                                             350
                                                                                                          RRGTALGINKEKRIQYAKI
sp|Q6FLD5|RPB2_CANGA
                                                                                                                                             352
                                                           CYDVNDWCMLEMLKPCVEDGFVI ODRETAL DE
                                                                                                          RRGTALGIKKEKRIQYAKT
RRGTALGIKKEKRIQYAKT
tr|A0A0L8VHA5|A0A0L8VHA5 9SACH
                                                  GETTEH
                                                                                                                                             354
tr|A0A0L8RB33|A0A0L8RB33 SACEU
                                                  GETTEH
                                                                                                                                             354
tr|G0VJ71|G0VJ71_NAUCC
tr|G8ZM49|G8ZM49_TORDC
                                                                                                          RRGTALGINKEKRIOYAKI
                                                  GEILEH
                                                                                                                                             354
                                                                                                          RRGTALGINKEKRIQYARI
                                                                                                                                             352
tr|A0A1Q3A090|A0A1Q3A090_ZYGRO
                                                                                                          rrgtalginkekriqyari
                                                                                                                                             354
                                                  GEILEH
                                                           CYDVNDWQMLEMLKPCVEDGFVIQDRETALDF
CYDVNDWQMLEMLKPCVEDGFVIQDRETALDF
CYDVNDWQMLEMLKPCVEDGFVIQDRETALDF
CYDVNDWQMLEMLKPCVEDGFVIQDRETALDF
CYDVNDWQMLEMLKPCVEDGFVIQDRETALDF
tr|A0A0N7IS35|A0A0N7IS35_9SACH
tr|A0A212MG88|A0A212MG88_ZYGBA
                                                  GEILEH:
                                                                                                          rrgtalginkekriqyari
                                                                                                                                             352
                                                                                                          RRGTALGINKEKRIOYARI
                                                  GETTEH
                                                                                                                                             353
tr|A0A1S7HHE1|A0A1S7HHE1 9SACH
                                                  GEILEH:
                                                                                                          RRGTALGIKKEKRIQYARI
                                                                                                                                             353
tr|S6ESB4|S6ESB4 ZYGB2
                                                                                                          rrgtalginkekriqyari
                                                                                                                                             353
                                                           tr|B6K5Q5|B6K5Q5_SCHJY
                                                  RDILEH
                                                                                                          KRGSTTGVTREKRLRYAHI
                                                                                                                                             340
sp|Q02061|RPB2_SCHPO
                                                  RDITER
                                                                                                          KRGSTTGVTREKRLRYAHI
                                                                                                                                             340
tr|S9R8U4|S9R8U4 SCHOY
                                                  RDILEH
                                                                                                         KRGSTTGVTREKRLRYAHI
                                                                                                                                             340
                                                                                                         KRGSTTGVTREKRLRYAHI
tr|S9W8C6|S9W8C6 SCHCR
                                                  RDILEH:
                                                                                                                                             340
```

```
sp|P30876|RPB2 HUMAN
                                          VLQKEMLPHVGVSDFGETKKAYFLGYMVHRLLLAALGRE
                                                                                          LDDRDH
                                                                                                    KRLDLAGPLLA
                                                                                                                          401
                                          VLQKEMLPHVGVSDFGETKKAYFLGYMVHRLLLAALGRE
VLOKEMLPHVGVSDFGETKKAYFLGYMVHRLLLAALGRE
tr|G3V8Y5|G3V8Y5 RAT
                                                                                          LDDRDH
                                                                                                    MKRLDLAGPLLA
                                                                                                                          401
                                                                                                    KRI.DI.AGPI.I.
tr|A0A250Y753|A0A250Y753 CASCN
                                                                                           DDRDH
                                                                                                                          401
tr|A0A1U7R4C7|A0A1U7R4C7_MESAU
tr|A0A286XIQ9|A0A286XIQ9_CAVPO
                                          VLQKEMLPHVGVSDFCETKKAYFLGYMVH
                                                                              LLLAALGRE
                                                                                                    KRLDLAGPLLA
                                                                                           DDRDH
                                                                                                                          401
                                          VLQKEMLPHVGVSDFCETKKAYFLGYMVH
                                                                              LLLAALGRE
                                                                                           DDRDH:
                                                                                                    KRLDLAGPLLA
                                                                                                                          401
tr|I3M351|I3M351 ICTTR
                                          VLQKEMLPHVGVSDFCETKKAYFLGYMVH
                                                                              LLLAALGRE
                                                                                                    KRLDLAGPLLA
                                                                                           DDRDH
                                                                                                                          401
tr|G7P5R6|G7P5R6_MACFA
                                          VLQKEMLPHVGVSDFCETKKAYFLGYMVH
                                                                              LLLAALGRE
                                                                                           DDRDHY
                                                                                                    KRLDLAGPLLA
                                                                                                                          401
tr|H2QPI8|H2QPI8_PANTR
tr|A0A1U7V0T5|A0A1U7V0T5_TARSY
                                          VLQKEMLPHVGVSDFCETKKAYFLGYMVHI
                                                                              T.T. AAT.GRE
                                                                                           יותתאחח
                                                                                                    KRT.DT.AGPT.T.Z
                                                                                                                          401
                                                                              LLLAALGREE
                                          VLQKEMLPHVGVSDFCETKKAYFLGYMVH
VLQKEMLPHVGVSDFCETKKAYFLGYMVH
                                                                                           DDRDH
                                                                                                    KRLDLAGPLLA
                                                                                                                          401
                                                                              LLLAALGRE
                                                                                           DDRDH
                                                                                                    KRLDLAGPLLA
tr|A0A1S2ZSL2|A0A1S2ZSL2 ERIEU
                                                                                                                          401
tr|A0A0D9QYL1|A0A0D9QYL1 CHLSB
                                          VLQKEMLPHVGVSDFCETKKAYFLGYMVH
                                                                              LLLAALGRE
                                                                                           DDRDH
                                                                                                    KRLDLAGPLLA
                                                                                                                          394
tr|A0A2K5ZNR7|A0A2K5ZNR7_MANLE
                                          VLQKEMLPHVGVSDFCETKKAYFLGYMVH
                                                                              LLLAALGRE
                                                                                           DDRDH'
                                                                                                    KRLDLAGPLLA
                                                                                                                          394
tr|A0A2I2ZIU3|A0A2I2ZIU3 GORGO
                                          VLQKEMLPHVGVSDFGETKKAYFLGYMVH
                                                                              LLLAALGRE
                                                                                           DDRDH:
                                                                                                    KRLDLAGPLLA
                                                                                                                          401
tr|A0A1D5QGA5|A0A1D5QGA5_MACMU
                                          VLQKEMLPHVQVSDFCETKKAYFLGYMVH
                                                                              T.T.T.AAT.GRE
                                                                                           DDRDHY
                                                                                                    KRI.DI.AGPI.I.A
                                                                                                                          401
tr|A0A2J8S2N1|A0A2J8S2N1 PONAB
                                          VLOKEMLPHVGVSDFCETKKAYFLGYMVH
                                                                              T.T.T.AAT.GRE
                                                                                           יאמאממ
                                                                                                    KRIDI AGPIJA
                                                                                                                          401
tr|A0A2K5K5J5|A0A2K5K5J5_COLAP
tr|A0A2J8PEW7|A0A2J8PEW7_PANTR
                                                                              LLLAALGRE
                                          VLQKEMLPHVGVSDFCETKKAYFLGYMVH
VLQKEMLPHVGVSDFCETKKAYFLGYMVH
                                                                                           DDRDHY
                                                                                                    KRLDLAGPLLA
                                                                                                                          401
                                                                                           DDRDH
                                                                              LLLAALGRE
                                                                                                    KRLDLAGPLL
                                                                                                                          394
tr|A0A2K5CY83|A0A2K5CY83_AOTNA
                                          VLQKEMLPHVGVSDFCETKKAYFLGYMVH
                                                                              LLLAALGRE
                                                                                           DDRDH'
                                                                                                    KRLDLAGPLLA
                                                                                                                          401
tr|A0A096NEY4|A0A096NEY4 PAPAN
                                          VLQKEMLPHVGVSDFCETKKAYFLGYMVH
                                                                              LLLAALGRE
                                                                                           DDRDH
                                                                                                    KRLDLAGPLLA
                                                                                                                          401
tr|C9J2Y9|C9J2Y9_HUMAN
                                          VLQKEMLPHVGVSDFGETKKAYFLGYMVH
                                                                               LLLAALGRE
                                                                                           DDRDH
                                                                                                    KRLDLAGPLLA
                                                                                                                          394
tr|G8BY61|G8BY61_TETPH
tr|A0A1X7QYA1|A0A1X7QYA1_9SACH
                                                                              LLLCALDRE
LLLCALDRE
                                          ILQKEFLPHITQLEGEESRKAFFLGYMIN
                                                                                           DDRDH
                                                                                                    KRLDLAGPLLG
                                                                                                                          412
                                                                                                    KRIDI AGPI LO
                                          ILOKEFLPHITOLEGESRKAFFLGYMIN
                                                                                           DDRDH
                                                                                                                          410
tr|J7RV95|J7RV95_KAZNA
                                          ILQKEFLPHITQLEGFESRKAFFLGYMIN
                                                                              LLLCALDRE
                                                                                           DDRDHI
                                                                                                    KRLDLAGPLLG
                                                                                                                          410
tr|H2AVJ8|H2AVJ8_KAZAF
                                          ILOKEFLPHITOLEGEESRKAFFLGYMIN
                                                                              LLLCALDRE
                                                                                           DDRDHI
                                                                                                    KRLDLAGPLLO
                                                                                                                          410
sp|Q6FLD5|RPB2_CANGA
                                           ILQKEFLPHITQLEGFESRKAFFLGYMIN
                                                                              LLLCALDRE
                                                                                           DDRDH
                                                                                                    KRLDLAGPLL
                                                                                                                          412
tr|A0A0L8VHA5|A0A0L8VHA5_9SACH
                                          ILQKEFLPHITQLEGEESRKAFFLGYMIN<mark>RL</mark>LLCALDRE
                                                                                           DDRDHI
                                                                                                    KRLDLAGPLL
                                                                                                                          414
                                                                                           ואמאממ
                                                                                                                          414
tr|A0A0L8RB33|A0A0L8RB33 SACEU
                                          ILQKEFLPHITQLEGFESRKAFFLGYMIN
                                                                              LLLCALDRE
                                                                                                    KRLDLAGPLLA
                                                                              LLLCALDRE
                                                                                           DDRDHI
tr|G0VJ71|G0VJ71_NAUCC
                                          ILOKEFLPHITOLEGEESRKAFFLGYMIN
                                                                                                    KRLDLAGPLLA
                                                                                                                          414
tr|G8ZM49|G8ZM49_TORDC
tr|A0A1Q3A090|A0A1Q3A090_ZYGRO
                                          ILQKEFLPHITQLEGEESRKAFFLGYMING
ILQKEFLPHITQLEGEESRKAFFLGYMING
                                                                              LLLCALDRE
                                                                                           DDRDH
                                                                                                    KRLDLAGPLLA
                                                                                                                          412
                                                                              LLLCALDRE
                                                                                                    KRLDLAGPLLA
                                                                                           DDRDHI
                                                                                                                          414
tr|A0A0N7IS35|A0A0N7IS35 9SACH
                                          ILQKEFLPHITQLEGFESRKAFFLGYMIN
                                                                              LLLCALDRE
                                                                                           DDRDH
                                                                                                    KRLDLAGPLLA
                                                                                                                          412
tr|A0A212MG88|A0A212MG88 ZYGBA
                                          ILQKEFLPHITQLEGFESRKAFFLGYMIN
                                                                              LLLCALDRE
                                                                                           DDRDHI
                                                                                                    KRLDLAGPLLA
                                                                                                                          413
                                                                              LLLCALDRE
LLLCALDRE
tr|A0A1S7HHE1|A0A1S7HHE1_9SACH
                                          ILQKEFLPHITQLEGEESRKAFFLGYMIN
                                                                                           DDRDH
                                                                                                    KRLDLAGPLLA
                                                                                                                          413
tr|S6ESB4|S6ESB4_ZYGB2
tr|B6K5Q5|B6K5Q5_SCHJY
                                          ILQKEFLPHITOLEGRESRKAFFLGYMIN
ILQKEMLPHITTLEGRETRKAFFLGYMVH
                                                                                           DDRDHI
                                                                                                    KRLDLAGPLLA
                                                                                                                          413
                                                                              MLLCALERE
                                                                                                    KRLDLAGPLL
                                                                                           DDRDH
                                                                                                                          400
sp|Q02061|RPB2 SCHPO
                                          ILQKELLPHITTMEGFETRKAFFLGYMIHR
                                                                              MLLCALERR
                                                                                                    KRLDLAGPLLA
                                                                                           DDRDH
                                                                                                                          400
tr|S9R8U4|S9R8U4 SCHOY
                                                                                           DDRDHI
                                          ILQKELLPHITTMEGFETRKAFFLGYMINRMLLCALERF
                                                                                                    KRLDLAGPLLA
                                                                                                                          400
                                                                                           DDRDH
tr|S9W8C6|S9W8C6_SCHCR
                                          ILQKELLPHITTMEGFETRKAFFLGYMIN
                                                                              MLLCALER
                                                                                                    KRLDLAGPLLA
                                                                                                                          4009
                                                             *::**:*****::<mark>*</mark>:**.** *:
  sp|P30876|RPB2 HUMAN
                                           FLFRGMFKNLLKEVRIYAQKFIDRGKDFNLELAIKTRIISDGLKYSLATGN
                                                                                                         OKKAHO
                                                                                                                         461
                                           FLFRGMFKNLLKEVRIYAQKFIDRGKDFNLELAIKTRIISDGLKYSLATGN
  tr|G3V8Y5|G3V8Y5 RAT
                                                                                                         OKKAHO
                                                                                                                         461
  tr|A0A250Y753|A0A250Y753_CASCN
                                            FLFRGMFKNLLKEVRIYAQKFIDRGKDFNLELAIKTRIISDGLKYSLATGN
                                                                                                          ОККАНО
                                                                                                                         461
  tr|A0A1U7R4C7|A0A1U7R4C7_MESAU
                                            FLFRGMFKNLLKEVRIYAQKFIDRGKDFNLELAIKTRI<mark>ISDGLKYSLATGN</mark>
                                                                                                          QKKAHQ
                                                                                                                         461
  tr|A0A286XIQ9|A0A286XIQ9_CAVPO
tr|I3M351|I3M351 ICTTR
                                           FI.FRGMFKNI.LKEVRTYAOKFIDRGKDFNI.ELATKTRTITSDGLKYSLATGN
                                                                                                          OKKAHO
                                                                                                                         461
                                            FLFRGMFKNLLKEVRIYAQKFIDRGKDFNLELAIKTRIISDGLKYSLATGN
                                                                                                          OKKAHO
                                                                                                                         461
  tr|G7P5R6|G7P5R6_MACFA
                                            FLFRGMFKNLLKEVRIYAQKFIDRGKDFNLELAIKTRIISDGLKYSLATGN
                                                                                                          OKKAHO
                                                                                                                         461
  tr|H2QPI8|H2QPI8 PANTR
                                           FLFRGMFKNLLKEVRIYAQKFIDRGKDFNLELAIKTRIISDGLKYSLATGN
                                                                                                          QKKAHQ
                                                                                                                         461
  tr|A0A1U7V0T5|A0A1U7V0T5_TARSY
tr|A0A1S2ZSL2|A0A1S2ZSL2_ERIEU
                                           FLFRGMFKNLLKEVRIYAQKFIDRGKDFNLELAIKTRIISDGLKYSLATGN
FLFRGMFKNLLKEVRIYAQKFIDRGKDFNLELAIKTRIISDGLKYSLATGN
                                                                                                          OKKAHO
                                                                                                                         461
                                                                                                          OKKAHO
                                                                                                                         461
  tr|A0A0D9QYL1|A0A0D9QYL1 CHLSB
                                            FLFRGMFKNLLKEVRIYAQKFIDRGKDFNLELAIKTRI<mark>ISDGLKYSLATGN</mark>
                                                                                                          OKKAHO
                                                                                                                         454
  tr|A0A2K5ZNR7|A0A2K5ZNR7_MANLE
                                            FLFRGMFKNLLKEVRIYAQKFIDRGKDFNLELAIKTRI<mark>ISDGLKYSLATGN</mark>
                                                                                                          QKKAHQ
                                                                                                                         454
  tr|A0A2I2ZIU3|A0A2I2ZIU3 GORGO
                                            FLFRGMFKNLLKEVRIYAOKFIDRGKDFNLELAIKTRIISDGLKYSLATGN
                                                                                                          OKKZHO
                                                                                                                         461
  tr|A0A1D5QGA5|A0A1D5QGA5_MACMU
                                            FLFRGMFKNLLKEVRIYAQKFIDRGKDFNLELAIKTRIISDGLKYSLATGN
                                                                                                          OKKAHO
                                                                                                                         461
  tr|A0A2J8S2N1|A0A2J8S2N1_PONAB
                                            FLFRGMFKNLLKEVRIYAQKFIDRGKDFNLELAIKTRIISDGLKYSLATGN
                                                                                                          OKKAHO
                                                                                                                         461
  tr|A0A2K5K5J5|A0A2K5K5J5_COLAP
                                            FLFRGMFKNLLKEVRIYAQKFIDRGKDFNLELAIKTRIISDGLKYSLATGN
                                                                                                          OKKAHO
                                                                                                                         461
  tr|A0A2J8PEW7|A0A2J8PEW7_PANTR
tr|A0A2K5CY83|A0A2K5CY83 AOTNA
                                           FLFRGMFKNLLKEVRIYAOKFIDRGKDFNLELAIKTRIISDGLKYSLATGN
                                                                                                          OKKAHO
                                                                                                                         454
                                           FLFRGMFKNLLKEVRIYAQKFIDRGKDFNLELAIKTRIISDGLKYSLATGN
                                                                                                          OKKAHO
                                                                                                                         461
  tr | AO AO 9 6NEY4 | AO AO 9 6NEY4 PAPAN
                                            FLFRGMFKNLLKEVRIYAQKFIDRGKDFNLELAIKTRIISDGLKYSLATGN
                                                                                                          QKKAHQ
  tr|C9J2Y9|C9J2Y9_HUMAN
                                            FLFRGMFKNLLKEVRIYAQKFIDRGKDFNLELAIKTRI<mark>I</mark>SDGLKYSLATGN
                                                                                                                         454
  triG8BY61iG8BY61 TETPH
                                           QLFKTLFRKLTKD I FR YMQRTVE EAN DFNMKLA INAKTIT SGLKYALATGNI
                                                                                                          OKKAMT
                                                                                                                         472
  tr|A0A1X7QYA1|A0A1X7QYA1_9SACH
                                           QLFKTLFRKLTKDIFRYMORTVEEAN DFNMKLAINAKTITSGLKYALATGN
                                                                                                         OKKAMS
                                                                                                                         470
  tr|J7RV95|J7RV95_KAZNA
                                            QLFKTLFRKLTKDIFRYMQRTVEEAHDFNMKLAINAKTITSGLKYALATGN
                                                                                                         OKKAMS
                                                                                                                         470
  tr|H2AVJ8|H2AVJ8 KAZAF
                                           QLFKTLFRKLTKDIFRYMQRTVEEANDFNMKLAINAKTITSGLKYALATGN
                                                                                                         OKKAMS
                                                                                                                         470
  sp|Q6FLD5|RPB2 CANGA
                                           OT FIRTH FREET, TWO I FRIYMOR TVE FAN DENMIKT A TNAKTITT SGT. KYAT ATGM
                                                                                                         OKKZMS
                                                                                                                         472
  tr|A0A0L8VHA5|A0A0L8VHA5 9SACH
                                           QLFKTLFKKLTKDIFRYMQRTVEEAHDFNMKLAINAKTITSGLKYALATGN
                                                                                                          OKKAMS
                                                                                                                         474
                                           OLFKSLFRKLTKDIFRYMORTVE EAHDFNMKLAINAKTITSGLKYALATGN
SLFKTLFRKLTKDIFRYMORTVE EAHDFNMKLAINAKTITSGLKYALATGN
  tr|A0A0L8RB33|A0A0L8RB33_SACEU
                                                                                                         OKKAMS
                                                                                                                         474
  tr|G0VJ71|G0VJ71_NAUCC
                                                                                                         OKKAMS
                                                                                                                         474
  tr|G8ZM49|G8ZM49_TORDC
tr|A0A1Q3A090|A0A1Q3A090_ZYGRO
                                           QLFKTLFRKLTKDIFRYMQRTVEEAKDFNMKLAINAKTITSGLKYALATGN
                                                                                                         OKKAMS
                                                                                                                         472
                                            QLFKTLFRKLTRDIFRYMQRTVEEAKDFNMKLAINAKTITSGLKYALATGN
                                                                                                          OKKAMS
                                                                                                                         474
  tr|A0A0N7IS35|A0A0N7IS35_9SACH
                                           QLFKTLFRKLTRDIFRYMQRTVEEAKDFNMKLAINAKTITSGLKYALATGN
                                                                                                         OKKAMS
                                                                                                                         472
                                           QLFKTLFRKLTRD I FR YMORTVE EAKDFNMKLA INAKT IT SGLKYALATGN
  tr|A0A212MG88|A0A212MG88_ZYGBA
                                                                                                         OKKAMS
                                                                                                                         473
                                            OLFKTLFRKLTRDIFRYMORTVEEAKDFNMKLAINAKTITSGLKYALATGN
  tr|A0A1S7HHE1|A0A1S7HHE1 9SACH
                                                                                                          OKKAMS
                                                                                                                         473
  tr|S6ESB4|S6ESB4_ZYGB2
                                            OLFKTLFRKLTRDIFRYMORTVEEAKDFNMKLAINAKTITSGLKYALATGN
                                                                                                          QKKAMS
                                                                                                                         473
  tr|B6K5Q5|B6K5Q5_SCHJY
sp|002061|RPB2_SCHPO
                                           SLFRMLFKKMTRDVYKYMOKCVETNREFNLTLAVKSNTITNGLRYSLATGN
SLFRMLFRKMTRDVYKYMOKCVETNREFNLTLAVKSNTITNGLRYSLATGN
                                                                                                          OKRGLA
                                                                                                                         460
                                                                                                          OKRAMV
                                                                                                                         460
  tr|S9R8U4|S9R8U4 SCHOY
                                           SLFRMLFRKMTRDVYKYMOKCVETNREFNLTLAVKSNTITNGLRYSLATGN
                                                                                                          QKRSMM
                                                                                                                         460
  tr|S9W8C6|S9W8C6 SCHCR
                                           SLFRMLFRKMTRDVYKYMQKCVETNREFNLTLAVKSNTITNGLRYSLATGN
                                                                                                         QKRSMM
```

: :*::: ::: * *: :: .:: **:::. |*:.**:*:***

**:

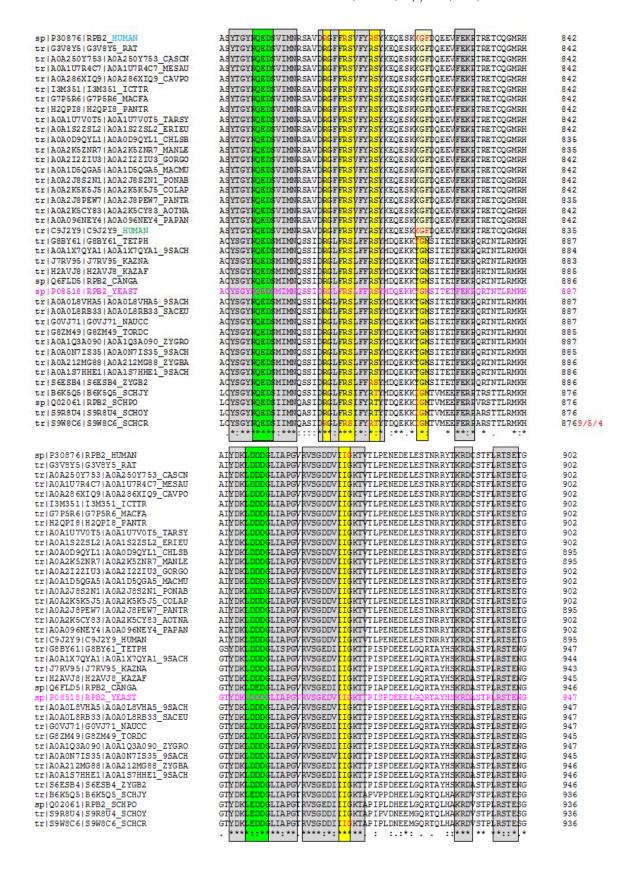
```
ARA GVS QVLNRL TFASTLSHLRR LNS PIGRDGKLAKPRQLHNT WGMVCPAETPEGHAVG

sp | P30876 | RPB2 HUMAN
 tr|G3V8Y5|G3V8Y5_RAT
                                                                                                                                                                                  521
tr|A0A250Y753|A0A250Y753_CASCN
                                                                                                                                                                                  521
tr|A0A1U7R4C7|A0A1U7R4C7_MESAU
                                                                                                                                                                                  521
tr|A0A286XIQ9|A0A286XIQ9_CAVPO
                                                                                                                                                                                  521
tr|I3M351|I3M351 ICTTR
                                                                                                                                                                                  521
tr|G7P5R6|G7P5R6_MACFA
                                                                                                                                                                                  521
tr|H2QPI8|H2QPI8 PANTR
                                                                                                                                                                                  521
tr|A0A1U7V0T5|A0A1U7V0T5_TARSY
                                                                                                                                                                                  521
tr|A0A1S2ZSL2|A0A1S2ZSL2_ERIEU
                                                                                                                                                                                  521
tr|A0A0D9QYL1|A0A0D9QYL1_CHLSB
                                                                                                                                                                                  514
tr|A0A2K5ZNR7|A0A2K5ZNR7_MANLE
tr|A0A2I2ZIU3|A0A2I2ZIU3_GORGO
                                                                                                                                                                                  514
tr|A0A1D5QGA5|A0A1D5QGA5 MACMU
                                                                                                                                                                                  521
tr|A0A2J8S2N1|A0A2J8S2N1_PONAB
                                                                                                                                                                                  521
tr|A0A2K5K5J5|A0A2K5K5J5_COLAP
                                                                                                                                                                                  521
tr|A0A2J8PEW7|A0A2J8PEW7_PANTR
tr|A0A2K5CY83|A0A2K5CY83 AOTNA
                                                                                                                                                                                  514
                                                                                                                                                                                  521
tr|A0A096NEY4|A0A096NEY4_PAPAN
                                                                                                                                                                                  521
tr|C9J2Y9|C9J2Y9_HUMAN
tr|G8BY61|G8BY61_TETPH
                                                                                                                                                                                  514
                                                               SRAGVSQVLNRYTYSSTLSHLRRTNTPIGRDGKLAKPRQLHNTHWGLVCPAETPEGQACG
tr|A0A1X7QYA1|A0A1X7QYA1_9SACH
                                                               SRAGVSQVLNRTTYSSTLSHLRRTNTPIGRDGKLAKPRQLHNTHWGLVCPAETPEGQACG
                                                                                                                                                                                  530
                                                              SRAGVSOVLNRTT YSSTLSHLRRTNT PIGRDGKLAKPROLHNT WGLVCPAETPEGOACG
SRAGVSOVLNRTT YSSTLSHLRRTNT PIGRDGKLAKPROLHNT WGLVCPAETPEGOACG
SRAGVSOVLNRTT YSSTLSHLRRTNT PIGRDGKLAKPROLHNT WGLVCPAETPEGOACG
tr|J7RV95|J7RV95_KAZNA
                                                                                                                                                                                  530
tr|H2AVJ8|H2AVJ8 KAZAF
                                                                                                                                                                                  530
sp|Q6FLD5|RPB2 CANGA
                                                                                                                                                                                  532
tr|A0A0L8VHA5|A0A0L8VHA5_9SACH
                                                              SRAGVSQVLNRT YSSTLSHLRR NTPIGRDGKLAKPRQLHNT WGLVCPAETPEGQA
SRAGVSQVLNRT YSSTLSHLRR NTPIGRDGKLAKPRQLHNT WGLVCPAETPEGQA
                                                                                                                                                                                  534
tr|A0A0L8RB33|A0A0L8RB33_SACEU
                                                                                                                                                                                  534
                                                              SRAGVSQVLNRTTYSSTLSHLRRTNTPIGRDGKLAKPRQLHNTHGLVCPAETPEGQACG
SRAGVSQVLNRTTYSSTLSHLRRTNTPIGRDGKLAKPRQLHNTHGLVCPAETPEGQACG
SRAGVSQVLNRTTYSSTLSHLRRTNTPIGRDGKLAKPRQLHNTHGLVCPAETPEGQACG
tr|G0VJ71|G0VJ71_NAUCC
                                                                                                                                                                                  534
tr|G8ZM49|G8ZM49 TORDC
                                                                                                                                                                                  532
tr|A0A1Q3A090|A0A1Q3A090 ZYGRO
                                                                                                                                                                                  534
tr|A0A0N7IS35|A0A0N7IS35_9SACH
                                                               SRAGVSQVLNRYTYSSTLSHLRRINTPIGRDGKLAKPRQLHNTHWGLVCPAETPEGQACG
                                                              SRAGVSQVLNRYT YSSTLSHLRRINT PIGRDGKLAKPROLHNT HUGLVCPAETPEGOACG
SRAGVSQVLNRYT YSSTLSHLRRINT PIGRDGKLAKPROLHNT HUGLVCPAETPEGOACG
SRAGVSQVLNRYT YSSTLSHLRRINT PIGRDGKLAKPROLHNT HUGLVCPAETPEGOACG
 tr|A0A212MG88|A0A212MG88_ZYGBA
                                                                                                                                                                                  533
 tr|A0A1S7HHE1|A0A1S7HHE1_9SACH
                                                                                                                                                                                  533
tr|S6ESB4|S6ESB4 ZYGB2
                                                                                                                                                                                  533
                                                              NRVGVSQVLNRYT FASTLSHLRR INT PIGRDGKLAKPROLHNT HGMVCPAETPEGOACG
NRVGVSQVLNRYT FASTLSHLRR INT PIGRDGKLAKPROLHNT HWGMVCPAETPEGOACG
tr|B6K5Q5|B6K5Q5 SCHJY
                                                                                                                                                                                  520
sp|Q02061|RPB2 SCHPO
                                                                                                                                                                                  520
                                                                   GVSQVLNRTT FASTLSHLRRTNT PIGRDGKLAKPROLHNT WGMVCPAETPEGOACG
GVSQVLNRTT FASTLSHLRRTNT PIGRDGKLAKPROLHNT WGMVCPAETPEGOACG
 tr|S9R8U4|S9R8U4_SCHOY
                                                                                                                                                                                  520
 tr|S9W8C6|S9W8C6 SCHCR
                                                                                                                                                                                  520
                                                                                                         · ************
                                                                                     ::******
                                                              LVKNLA LMA YISVGSQ PSPI LEF LEEWSMENLEELSPAALADA TKIFVNGGW
LVKNLA LMA YISVGSQ PSPI LEF LEEWSMENLEELSPAALADA TKIFVNGGW
LVKNLA LMA YISVGSQ PSPI LEF LEEWSMENLEELSPAALADA TKIFVNGGW
splP308761RPB2 HUMAN
                                                                                                                                                            IHKDPE
                                                                                                                                                                                  581
triG3V8Y5|G3V8Y5 RAT
                                                                                                                                                            IHKDPE
                                                                                                                                                                                  581
tr|A0A250Y753|A0A250Y753_CASCN
tr|A0A1U7R4C7|A0A1U7R4C7_MESAU
                                                                                                                                                            IHKDPE
                                                                                                                                                                                  581
                                                              LVKNLALMAYISVGSOPSPILEFLEEWSMENLEEISPAAIADATKIFVNGO
                                                                                                                                                            IHKDPE
                                                                                                                                                                                  581
tr|A0A286XIQ9|A0A286XIQ9 CAVPO
                                                              LVKNLALMAYISVGSOPSPILEFLEEWSMENLEEISPAAIADATKIFVNGO
                                                                                                                                                            IHKDPE
                                                                                                                                                                                  581
tr|I3M351|I3M351_ICTTR
                                                              LVKNLALMAYISVGS@PSPILEFLEEWSMENLEETSPAAIADATKIFVNGC
                                                                                                                                                            IHKDPE
                                                                                                                                                                                  581
                                                              LVKNLALMAYISVGSOPSPILEFLEEWSMENLEETSPAAIADATKIFVNGO
LVKNLALMAYISVGSOPSPILEFLEEWSMENLEETSPAAIADATKIFVNGO
tr|G7P5R6|G7P5R6 MACFA
                                                                                                                                                            THEOPE
                                                                                                                                                                                  581
tr|H2QPI8|H2QPI8_PANTR
                                                                                                                                                            IHKDPE
                                                                                                                                                                                  581
                                                              LVKNLALMAYISVGS PSPILEFLEEWSMENLEEISPAAIADATKIFVNGC
tr|A0A1U7V0T5|A0A1U7V0T5_TARSY
                                                                                                                                                            IHKDPE
                                                                                                                                                                                  581
tr|A0A1S2ZSL2|A0A1S2ZSL2 ERIEU
                                                              LVKNLALMAYISVGSOPSPILEFLEEWSMENLEEISPAAIADATKIFVNGO
                                                                                                                                                            IHKDPE
                                                                                                                                                                                  581
tr|A0A0D9QYL1|A0A0D9QYL1_CHLSB
                                                              LVKNLALMAYISVGS. PSPILEFLEEWSMENLEETSPAAIADATKIFVNGO
                                                                                                                                                            IHKDPE
                                                                                                                                                                                  574
tr|A0A2K5ZNR7|A0A2K5ZNR7_MANLE
                                                              LVKNLALMAYISVGSOPSPILEFLEEWSMENLEEISPAAIADATKIFVNGO
                                                                                                                                                            IHKDPE
                                                                                                                                                                                  574
                                                              LVKNLALMAYISVGSOPSPILEFLEEWSMENLEETSPAAIADATKIFVNGO
LVKNLALMAYISVGSOPSPILEFLEEWSMENLEETSPAAIADATKIFVNGO
tr|A0A2I2ZIU3|A0A2I2ZIU3_GORGO
                                                                                                                                                            IHKDPE
                                                                                                                                                                                  581
tr|A0A1D5QGA5|A0A1D5QGA5 MACMU
                                                                                                                                                            THKDEE
                                                                                                                                                                                  581
tr|A0A2J8S2N1|A0A2J8S2N1_PONAB
tr|A0A2K5K5J5|A0A2K5K5J5_COLAP
                                                              LVKNLALMAYISVGSOPSPILEFLEEWSMENLEEISPAAIADATKIFVNGO
                                                                                                                                                            IHKDPE
                                                                                                                                                                                  581
                                                              LVKNLALMAYISVGSOPSPILEFLEEWSMENLEEISPAAIADATKIFVNGO
                                                                                                                                                            IHKDPE
                                                                                                                                                                                  581
                                                              LVKNLALMAYISVGS PSPILEFLEEWSMENLEEISPAAIADATKIFVNGO
tr|A0A2J8PEW7|A0A2J8PEW7_PANTR
                                                                                                                                                            IHKDPE
                                                                                                                                                                                  574
tr|A0A2K5CY83|A0A2K5CY83_AOTNA
                                                              LVKNLALMAYISVGSOPSPILEFLEEWSMENLEETSPAAIADATKIFVNGC
                                                                                                                                                            IHKDPE
                                                                                                                                                                                  581
tr|A0A096NEY4|A0A096NEY4 PAPAN
                                                              LVKNLALMAYISVGSOPSPILEFLEEWSMENLEETSPAAIADATKIFVNGO
                                                                                                                                                            THKDEE
                                                                                                                                                                                  581
                                                              LVKNLALMAYISVGSOPSPILEFLEEWSMENLEEISPAAIADATKIFVNGO
                                                                                                                                                                                  574
tr|C9J2Y9|C9J2Y9_HUMAN
                                                                                                                                                            IHKDPE
                                                              LVKNLSLMSCISVGTDPMPIITFLSEWGMEPLEDVVPHQSPDATRVFVNGV
                                                                                                                                                            VHRNPA
tr|G8BY61|G8BY61 TETPH
                                                                                                                                                                                  592
tr|A0A1X7QYA1|A0A1X7QYA1 9SACH
                                                              LVKNLSLMSCISVGTDPTPIITFLSEWGMEPLEDYVPHQSPDATRVFVNGV
                                                                                                                                                            VHRNPA
                                                                                                                                                                                  590
tr|J7RV95|J7RV95_KAZNA
                                                              LVKNLSLMSSISVGTDPMPIITFLSEWGLEPLEDYVPHQSPDATRVFVNGV
                                                                                                                                                            VHRNPA
                                                                                                                                                                                  590
tr|H2AVJ8|H2AVJ8 KAZAF
                                                              LVKNLSLMSCISVGTDPMPIITFLSEWGMEPLEDYIPHQSeDATRVFVNGV
                                                                                                                                                            VHRNPA
                                                                                                                                                                                  590
sp|Q6FLD5|RPB2_CANGA
                                                              LVKNLSLMSCISVGAD PMPIITFLSEWGMEPLEDYVPHOS HDATRVFVNG
                                                                                                                                                            VHRNPA
                                                                                                                                                                                  592
tr|A0A0L8VHA5|A0A0L8VHA5 9SACH
                                                              LVKNLSLMSCISVGTDPMPIITFLSEWGMEPLEDYVPHQSPDATRVFVNG
                                                                                                                                                            VHRNPA
                                                                                                                                                                                  594
                                                              LVKNLSLMSCISVGTDPMPIITFLSEWGMEPLEDYVPHQSPDATRVFVNG
tr|A0A0L8RB33|A0A0L8RB33 SACEU
                                                                                                                                                            VHRNPA
                                                                                                                                                                                  594
                                                              LVKNLSLMSCISVGTDPMPIITFLSEWGMEPLEDYVPHQSPDATRVFVNGV
tr|G0VJ71|G0VJ71_NAUCC
                                                                                                                                                            VHRNPA
                                                                                                                                                                                  594
tr|G8ZM49|G8ZM49_TORDC
tr|A0A1Q3A090|A0A1Q3A090_ZYGRO
                                                                                                                                                            VHRNPA
                                                              LVKNLSLMSCISVGTDPMPIITFLSEWGMEPLEDYVPHQSPDATRVFVNG
                                                                                                                                                                                  592
                                                              LVKNLSLMSCISVGTDPMPIITFLSEWGMEPLEDYIPHQSPEATRVFVNG
                                                                                                                                                            VHRNPA
                                                                                                                                                                                  594
                                                              LVKNLSLMSCISVGTD PMPI ITFLSEWGME PLE DYVPHOS BEATRVFVNGV
LVKNLSLMSCISVGTD PMPI ITFLSEWGME PLE DYVPHOS BEATRVFVNGV
tr|A0A0N7IS35|A0A0N7IS35_9SACH
tr|A0A212MG88|A0A212MG88 ZYGBA
                                                                                                                                                            VHRNPA
                                                                                                                                                                                  592
                                                                                                                                                            VHRNPA
                                                                                                                                                                                  593
tr|A0A1S7HHE1|A0A1S7HHE1 9SACH
                                                              LVKNLSLMSCISVGTD PMPI ITFLSEWGME PLE DYVPHQS PEATRVFVNGV
                                                                                                                                                            VHRNPA
                                                                                                                                                                                  593
                                                              LVKNLSLMSCISVGTDAMPIITFLSEWGMEPLEDYVPHQSEEATRVFVNG
                                                                                                                                                            VHRNPA
tr|S6ESB4|S6ESB4 ZYGB2
                                                                                                                                                                                  593
tr|B6K5Q5|B6K5Q5_SCHJY
                                                              LVKNLSLMSYVSVGS#SAPIIEFLEEWGMESLEDYNPSAS#NATKVFVNGI
                                                                                                                                                            VHR DPV
                                                                                                                                                                                  580
                                                              LVKNLS LMS YVSVGSP SAPI IEFLEEWGLETLE DYN PSAS FNA TKVFVNGV
LVKNLALMS YVSVGSP AAPI IEFLEEWGLE SLE DYN PSAS FNA TKVFVNGV
sp|Q02061|RPB2_SCHPO
                                                                                                                                                            VHRDPA
                                                                                                                                                                                  580
tr|S9R8U4|S9R8U4 SCHOY
                                                                                                                                                            VHRDPA
                                                                                                                                                                                  580
                                                              LVKNLALMSYVSVGSPAAPIIEFLEEWGLESLEDYNPSAS HNATKVFVNGV
trlS9W8C6|S9W8C6 SCHCR
                                                                                                                                                            VHRDPA
                                                                                                                                                                                  580
                                                                                                                                                            : * : :
```

```
sp|P30876|RPB2 HUMAN
                                                    QLMNTLRKLRRQMDIIVSEVSMIRDIREREIRIYTDAGRICRPLLIVEKQK-----LL
                                                                                                                                                    634
                                                   QLMNTLRKLRQMDIIVSEVSMIRDIRERE IRIYTDAGRICRPLLIVEKQK-----LL
QLMNTLRKLRQMDIIVSEVSMIRDIRERE IRIYTDAGRICRPLLIVEKQK-----LL
QLMNTLRKLRQMDIIVSEVSMIRDIRERE IRIYTDAGRICRPLLIVEKQK-----LL
QLMNTLRKLRQMDIIVSEVSMIRDIRERE IRIYTDAGRICRPLLIVEKQK-----LL
tr|G3V8Y5|G3V8Y5 RAT
                                                                                                                                                    634
tr|A0A250Y753|A0A250Y753_CASCN
tr|A0A1U7R4C7|A0A1U7R4C7_MESAU
                                                                                                                                                    634
                                                                                                                                                    634
tr|A0A286XIQ9|A0A286XIQ9_CAVPO
                                                                                                                                                    634
                                                    OLANTLRKIROZDIIVSEVSMIRDIRERE IRIYTDAGRICRPLLIVEKOK-----LL
OLANTLRKIROZDIIVSEVSMIRDIRERE IRIYTDAGRICRPLLIVEKOK-----LL
OLANTLRKIROZDIIVSEVSMIRDIRERE IRIYTDAGRICRPLLIVEKOK-----LL
tr||I3M351||I3M351_|ICTTR
                                                                                                                                                    634
tr|G7P5R6|G7P5R6_MACFA
tr|H2QPI8|H2QPI8 PANTR
                                                                                                                                                    634
                                                                                                                                                    634
                                                    OLMNILRA ROCHDIIVSEVSMIRDIREREIRIYIDAGRICRPLLIVEKOK-----LL
OLMNILRA ROCHDIIVSEVSMIRDIREREIRIYIDAGRICRPLLIVEKOK-----LL
tr|A0A1U7V0T5|A0A1U7V0T5 TARSY
                                                                                                                                                    634
tr|A0A1S2ZSL2|A0A1S2ZSL2_ERIEU
                                                                                                                                                    634
                                                   QLMNTLRKLRQMDIIVSEVSMIRDIRERE IRI YIDAGRICRPLLIVERQK-----LL
QLMNTLRKLRQMDIIVSEVSMIRDIRERE IRI YITDAGRICRPLLIVERQK-----LL
QLMNTLRKLRQMDIIVSEVSMIRDIRERE IRI YITDAGRICRPLLIVERQK-----LL
QLMNTLRKLRQMDIIVSEVSMIRDIRERE IRI YITDAGRICRPLLIVERQK------LL
QLMNTLRKLRQMDIIVSEVSMIRDIRERE IRI YITDAGRICRPLLIVERQK------LL
QLMNTLRKLRQMDIIVSEVSMIRDIRERE IRI YITDAGRICRPLLIVERQK------LL
tr|A0A0D9QYL1|A0A0D9QYL1 CHLSB
                                                                                                                                                    627
tr|A0A2K5ZNR7|A0A2K5ZNR7_MANLE
tr|A0A2I2ZIU3|A0A2I2ZIU3 GORGO
                                                                                                                                                    634
tr|A0A1D5QGA5|A0A1D5QGA5_MACMU
                                                                                                                                                    634
tr|A0A2J8S2N1|A0A2J8S2N1 PONAB
                                                                                                                                                    634
                                                    OLMNILRA ROZDIIVSEVSMIRDIREE IRIYIDAGRICRPLIVEKOK----LL
OLMNILRA ROZDIIVSEVSMIRDIREE IRIYIDAGRICRPLIVEKOK-----LL
tr|A0A2K5K5J5|A0A2K5K5J5_COLAP
tr|A0A2J8PEW7|A0A2J8PEW7_PANTR
                                                                                                                                                    634
                                                                                                                                                    627
tr|A0A2K5CY83|A0A2K5CY83_AOTNA
                                                    QLMNTLRKLRRQMDIIVSEVSMIRDIREREIRIYTDAGRICRPLLIVEKQK-----LL
                                                                                                                                                    634
                                                    QLMNTLRKLRROMDIIVSEVSMIRDIREREIRIYTDAGRICRPLLIVEKOK-----LL
QLMNTLRKLRROMDIIVSEVSMIRDIREREIRIYTDAGRICRPLLIVEKOK-----LL
tr|A0A096NEY4|A0A096NEY4 PAPAN
                                                                                                                                                    634
tr|C9J2Y9|C9J2Y9_HUMAN
                                                                                                                                                    627
                                                    RIMETLRE LREKEDI - NEEVSMIRDIREQELKIFTDAGRVYRPLFIVDDDDDLG-HKELR
RIMETLRE LREKEDI - NEEVSMVRDIREKELKIFTDAGRVYRPLFIVEDDSELG-HKELK
tr|G8BY61|G8BY61_TETPH
tr|A0A1X7QYA1|A0A1X7QYA1 9SACH
                                                                                                                                                    650
                                                                                                                                                    648
tr|J7RV95|J7RV95_KAZNA
                                                    RIMETLRTLRRKGDI-NEEVSMIRDIREKELKIFTDAGRVYRPLFIVEDDEALG-HKELK
                                                                                                                                                    648
                                                    RIMETLRI RRKEDI NEEVSMIRDIREKELKIFTDAGRVYRPLETVEDDEALG-HKELK
RIMETLRI LRKEDI NEEVSMIRDIREQELKIFTDAGRVYRPLFIVEDDEELG-RKELK
tr|H2AVJ8|H2AVJ8 KAZAF
                                                                                                                                                    648
sp|Q6FLD5|RPB2 CANGA
                                                                                                                                                    650
tr|A0A0L8VHA5|A0A0L8VHA5_9SACH
                                                    RIMETLRTLRRK DI-NHEVSMIRDIREKELKIFTDAGRVYRPLFIVEDDESLG-HKELK
                                                                                                                                                    652
                                                    RIMETLRE LREKGDI-NHEVSMIRDIREKELKIFTDAGRVYRPLFIVEDDETLG-HKELK
RIMDTLRE LREKGDI-NHEVSMIRDIREKELKIFTDAGRVYRPLFIVEDDETLG-HKELK
tr|A0A0L8RB33|A0A0L8RB33 SACEU
                                                                                                                                                    652
tr|G0VJ71|G0VJ71_NAUCC
tr|G8ZM49|G8ZM49 TORDC
                                                                                                                                                    652
                                                    RIMETLRTLRRKGDI-NHEVSMVRDIREKELKIFTDAGRVYRPLFIVEDDETLG-HKELK
                                                                                                                                                    650
tr|A0A1Q3A090|A0A1Q3A090_ZYGRO
                                                    RLMETLRTLRRKGDI-NEEVSMVRDIREKELKIFTDAGRVYRPLFIVDDDETLN-RKELR
                                                                                                                                                    652
tr|A0A0N7IS35|A0A0N7IS35_9SACH
                                                    RIMETLRILRRKGDI-NEEVSMVRDIREKELKIFTDAGRVYRPLFIVDDDETLN-RKELK
                                                                                                                                                    650
                                                    RIMETLRE LRRKEDI NEEVSMVRDIREKELKIFTDAGRVYRPLFIVDDDESLN-RKELR
RIMETLRE LRRKEDI-NEEVSMVRDIREKELKIFTDAGRVYRPLFIVDDDESLN-RKELR
tr|A0A212MG88|A0A212MG88_ZYGBA
                                                                                                                                                    651
tr|A0A1S7HHE1|A0A1S7HHE1 9SACH
                                                                                                                                                    651
                                                    RIMETLER LERKEDI NEVSWYRDI REKELRI FTDAGRVYRPLETVDDDESIN REKEL
HLIETLESLERRIDI SÆVSIVRDI REKELRI FTDAGRVYRPLETVDDDSIN FRELR
HLIETLESLERRIDI SÆVSIVRDI REKELRI FTDAGRI CRPLFIVDNDSNSDTKGELC
tr|S6ESB4|S6ESB4_ZYGB2
tr|B6K5Q5|B6K5Q5_SCHJY
sp|Q02061|RPB2_SCHPO
                                                                                                                                                    651
                                                                                                                                                    639
                                                    HLTETLRSLRRRLDI-SAEVSIVRDIREKELRLFTDAGRICRPLFIVDNNPNSERRGELC
                                                                                                                                                    639
                                                   tr|S9R8U4|S9R8U4 SCHOY
                                                                                                                                                    639
tr|S9W8C6|S9W8C6_SCHCR
                                                                                                                                                    639
                                                    LKKRHIDQLKEREY------NNYSWQDLVASGVVEYIDT LEEETVMLAMT PDDLDE
sp | P3 087 6 | R PB2_HUMAN
                                                                                                                                                     684
                                                    LKKRHIDQLKERE!-----NN YSWQDLVAS VVEY IDT LEEE TWLLAMT PDDLQE
LKKRHIDQLKERE Y-----NN YSWQDLVAS VVEY IDT LEEE TWLLAMT PDDLQE
tr|G3V8Y5|G3V8Y5 RAT
                                                                                                                                                     684
tr|A0A250Y753|A0A250Y753_CASCN
                                                                                                                                                     684
tr|A0A1U7R4C7|A0A1U7R4C7_MESAU
tr|A0A286XIQ9|A0A286XIQ9_CAVPO
                                                    LKKRHIDQLKEREY-----NNYSWQDLVASGVVEYIDT
                                                                                                                          VMLAMT PDDLDE
                                                                                                                                                     684
                                                    LKKRHIDQLKEREY------NNYSWQDLVASGVVEYIDT
                                                                                                                    EEETVMLAMT PDDLDE
                                                                                                                                                     684
                                                    LKKRRI DQLKERE Y------NI ISWQDLVAS SVVEY IDT LEEE TYMLAMT PDDLQE
LKKRRI DQLKERE Y------NN YSWQDLVAS SVVEY IDT LEEE TYMLAMT PDDLQE
LKKRRI DQLKERE Y-----NN YSWQDLVAS SVVEY IDT LEEE TYMLAMT PDDLQE
LKKRHI DQLKERE Y-----NN YSWQDLVAS SVVEY IDT LEEE TYMLAMT PDDLQE
LKKRHI DQLKERE Y-----NN YSWQDLVAS SVVEY IDT LEEE TYMLAMT PDDLQE
tr|I3M351|I3M351 ICTTR
                                                                                                                                                     684
tr|G7P5R6|G7P5R6_MACFA
                                                                                                                                                     684
tr|H2OPI8|H2OPI8 PANTR
                                                                                                                                                     684
tr|A0A1U7V0T5|A0A1U7V0T5 TARSY
                                                                                                                                                     684
tr|A0A1S2ZSL2|A0A1S2ZSL2_ERIEU
                                                    LKKRHIDQLKEREY------NNYSWQDLVASGVVEYIDT
                                                                                                                     EEETVMLAMT PDD LOE
                                                                                                                                                     684
tr|A0A0D9QYL1|A0A0D9QYL1_CHLSB
                                                    LKKRHIDQLKEREY------NNYSWQDLVASGVVEYIDT
                                                                                                                    EEETVMLAMT PDDLDE
                                                                                                                                                     677
                                                    LKKRHI DOLKEREY-----NNYSWQDLVASGVVEYIDT
LKKRHI DOLKEREY-----NNYSWQDLVASGVVEYIDT
LKKRHI DQLKEREY-----NNYSWQDLVASGVVEYIDT
tr|A0A2K5ZNR7|A0A2K5ZNR7_MANLE
tr|A0A2I2ZIU3|A0A2I2ZIU3 GORGO
                                                                                                                    EEETVMLAMT PDDLQE
                                                                                                                                                     677
                                                                                                                          ANTIQUE TMALIMY
                                                                                                                                                     684
tr|A0A1D5QGA5|A0A1D5QGA5_MACMU
                                                                                                                          VMLAMT PDDLDE
                                                                                                                                                     684
tr|A0A2J8S2N1|A0A2J8S2N1_PONAB
                                                    LKKRHIDQLKEREY------NNYSWQDLVASGVVEYIDT
                                                                                                                          VMLAMT PDDLQE
                                                                                                                                                     684
                                                    LKKRHIDQLKEREY-----NNYSWQDLVASGVVEYIDT
tr|A0A2K5K5J5|A0A2K5K5J5_COLAP
                                                                                                                          VMLAMT PDDLDE
                                                                                                                                                     684
                                                    LKKRHIDQLKEREY------NNYSWQDLVASGVVEYIDT
LKKRHIDQLKEREY-----NNYSWQDLVASGVVEYIDT
LKKRHIDQLKEREY-----NNYSWQDLVASGVVEYIDT
LKKRHIDQLKEREY-----NNYSWQDLVASGVVEYIDT
LKKRHIDQLKEREY-----NNYSWQDLVASGVVEYIDT
tr|A0A2J8PEW7|A0A2J8PEW7_PANTR
                                                                                                                          VMLAMT PDDLDE
                                                                                                                                                     677
tr|A0A2K5CY83|A0A2K5CY83 AOTNA
                                                                                                                    EEET
                                                                                                                          VMT.AMT PDDT.OF.
                                                                                                                                                     684
                                                                                                                          MLAMTPDDLDE
tr|A0A096NEY4|A0A096NEY4 PAPAN
                                                                                                                                                     684
tr|C9J2Y9|C9J2Y9_HUMAN
tr|G8BY61|G8BY61 TETPH
                                                                                                                          VMLAMTPDDLDE
                                                                                                                                                     677
                                                    VRKGHVARLMATE YOD IEGG FEDVEE YTWT SLLNEGLVEY IDA
                                                                                                                           LIAMOPEDLOP
                                                                                                                                                     710
tr|A0A1X7QYA1|A0A1X7QYA1_9SACH
                                                    VRKGHINKLMATE YOD IEGGLDE SDE YTWS SLL SEGLVEY IDA
                                                                                                                           LIAMOPEDLEP
                                                                                                                                                     708
tr|J7RV95|J7RV95_KAZNA
                                                    VRKGHIAKLMATE YOD IEGGLDE ADE YTWT SLLNEGLVEY IDA
                                                                                                                          LIAMOPEDLEP
                                                                                                                                                     708
tr|H2AVJ8|H2AVJ8 KAZAF
                                                    VRKGHINKLMATE YOD IEGG FED SEE YTWT SLLNEGLVEY IDA
                                                                                                                          LIAMOPEDLEP
                                                                                                                                                     708
                                                    VRKGHVAKLMATE YOD IEGG FEDAED YTWS SLLNEGLVEY IDA
sp|Q6FLD5|RPB2 CANGA
                                                                                                                           LIAMOPEDLEP
                                                                                                                                                     710
                                                                                                                                                     712
tr|A0A0L8VHA5|A0A0L8VHA5_9SACH
                                                    VRKGHIAKLMATE YOD IEGG FEDVEE YTWS SLLNEGLVEY IDA
                                                                                                                          LIAMOPEDLEP
tr|A0A0L8RB33|A0A0L8RB33_SACEU
                                                    VRKGHVAKLMTTE YQD IEGG FED VEE FTWS SLLNEGLVEY IDA
                                                                                                                          LIAMOPEDLEP
                                                                                                                                                     712
tr|G0VJ71|G0VJ71_NAUCC
tr|G8ZM49|G8ZM49 TORDC
                                                    VRKGHIAKLMATE YOD IEGG FED AEE YTWT SLLNEGLVEY IDA
VRKGHIGKLMATE YOD IEGG FED SED YTWS SLLNEGLVEY IDA
                                                                                                                          ILISMOPEDLEP
ILISMOPEDLEP
                                                                                                                                                     712
                                                                                                                                                     710
tr|A0A1Q3A090|A0A1Q3A090_ZYGRO
                                                    IRKGHISKLMATEYQDIEGGFEDSEEYTWTSLLNDGLVEYIDA
                                                                                                                          LISMOPEDLEP
                                                                                                                                                     712
tr|A0A0N7IS35|A0A0N7IS35 9SACH
                                                    IRKGHVGKLMATE YOD IEGG FED SEE YTWT SLLNEGLVEY IDA
                                                                                                                     EETILISMQ PEDLEP
                                                                                                                                                     710
                                                                                                                    EEETILISMOPEDLEP
tr|A0A212MG88|A0A212MG88_ZYGBA
                                                    IRKGHVAKLMATE YOD IEGG FED SEE YTWT SLLNEGLVEY IDA
                                                                                                                                                     711
                                                    IRKGHVAKLMATEYODIEGGFEDSELYIWISLLNEGLVEYIDA EEEET LISMOPEDLEP
tr|A0A1S7HHE1|A0A1S7HHE1 9SACH
                                                                                                                                                     711
tr|S6ESB4|S6ESB4_ZYGB2
                                                                                                                                                     711
tr|B6K5Q5|B6K5Q5_SCHJY
                                                    IRKEHIQQLLEDRD---RFDINAEQRFGWSSLIASGLIEYLDA
                                                                                                                    EEETVMIAMSPDDLEL
                                                                                                                                                     696
                                                    IRKEHIQQLIEDKD---RYDIDPEQRFGWTALVSSGLIEYLDA
sp|Q02061|RPB2 SCHPO
                                                                                                                    EEETVMIAMS PEDLEA
                                                                                                                                                     696
tr|S9R8U4|S9R8U4 SCHOY
                                                    IRKEHVQQLIEDRD---RFDIDPEQKFGWTALVSSGLIEYLDA EEEE TVMIAMT PEDLEA
                                                                                                                                                     696
                                                    tr|S9W8C6|S9W8C6_SCHCR
```

```
------VAYCSTYTHCEIHPSM<mark>ILG</mark>VCASIIPFPDHNQSPRN
spiP30876|RPB2 HUMAN
                                                                                                                                    722
                                                  -----VAYCSTYTHCEIHPSM<mark>ILG</mark>VCASIIPFPDHNOSPRN
tr|G3V8Y5|G3V8Y5 RAT
                                                                                                                                    722
                                                  -----VAYCST YTHCETHPSMTLGVCAS IIP FPDHNQS PRN
-----VAYCST YTHCETHPSMTLGVCAS IIP FPDHNQS PRN
-----VAYCST YTHCETHPSMTLGVCAS IIP FPDHNQS PRN
tr|A0A250Y753|A0A250Y753_CASCN
tr|A0A1U7R4C7|A0A1U7R4C7_MESAU
                                                                                                                                    722
                                                                                                                                    722
tr|A0A286XIQ9|A0A286XIQ9_CAVPO
                                                      ------PAYCSTYTHCEIHPSM<mark>ILG</mark>VCASIIPFPDHNQSPRN
                                                                                                                                    722
                                                       -----VAYCSTYTHCEIHPSM<mark>ILG</mark>
                                                                                                     VCASIIPFPDHNOSPRN
VCASIIPFPDHNOSPRN
tr|I3M351|I3M351 ICTTR
                                                                                                                                    722
tr|G7P5R6|G7P5R6_MACFA
tr|H2QPI8|H2QPI8 PANTR
                                                                  ·-----VAYCSTYTHCEIHPSM<mark>IL</mark>
                                                                                                                                    722
                                                  -----VAYCSTYTHCEIHPSM<mark>IL</mark>
                                                                                                    <mark>G</mark>VCASIIPFPDHNQSPRN
GVCASIIPFPDHNQSPRN
                                                                                                                                    722
                                                  -----VAYCSTYTHCEIHPSMIL
tr|A0A1U7V0T5|A0A1U7V0T5_TARSY
                                                                                                                                    722
                                                                                                    GVCASIIPFPDHNQSPRN
tr|A0A1S2ZSL2|A0A1S2ZSL2_ERIEU
                                                                                                                                    722
tr|A0A0D9QYL1|A0A0D9QYL1_CHLSB
tr|A0A2K5ZNR7|A0A2K5ZNR7_MANLE
                                              KE-----VAYCSTYTHCEIHPSM<mark>IL</mark>
                                                                                                    GVCASIIPFPDHNQSPRN
                                                                                                                                    715
                                              KE-----VAYCSTYTHCEIHPSM<mark>ILG</mark>
                                                                                                     VCASIIPFPDHNQSPRN
                                              KE-----VA YCST YTHCEIH PSM<mark>ILG</mark>
KE-----VA YCST YTHCEIH PSM<mark>ILG</mark>
                                                                                                     VCASIIPFPDHNOSPRN
VCASIIPFPDHNOSPRN
tr|A0A2I2ZIU3|A0A2I2ZIU3 GORGO
                                                                                                                                    722
tr|A0A1D5QGA5|A0A1D5QGA5_MACMU
                                                                                                                                    722
                                              KE-----VAYCSTYTHCEIHPSM
tr|A0A2J8S2N1|A0A2J8S2N1_PONAB
                                                                                                      VCASIIPFPDHNQSPRN
                                                                                                                                    722
                                              KE-----VAYCSTYTHCEIHPSMIL
KE-----VAYCSTYTHCEIHPSMIL
tr|A0A2K5K5J5|A0A2K5K5J5_COLAP
tr|A0A2J8PEW7|A0A2J8PEW7_PANTR
                                                                                                    GVCASIIPFPDHNQSPRN
GVCASIIPFPDHNQSPRN
GVCASIIPFPDHNQSPRN
                                                                                                                                    722
                                                                                                                                    715
tr|A0A2K5CY83|A0A2K5CY83 AOTNA
                                              KE----VAYCSTYTHCEIHPSM<mark>II</mark>
                                                                                                                                    722
                                                  -----VAYCSTYTHCEIHPSM<mark>I</mark>
tr|A0A096NEY4|A0A096NEY4 PAPAN
                                                                                                     VCASIIPFPDHNOSPRN
                                                                                                                                    722
tr|C9J2Y9|C9J2Y9_HUMAN
tr|G8BY61|G8BY61_TETPH
                                              KE-----VAYCSTYTHCEIHPSM<mark>T</mark>
                                                                                                     VCASIIPFPDHNQSPRN
                                                                                                                                    715
                                              VSDEPEAI-PDVETDTAKRIR--AVHHATTFTHCEIHPSM<mark>I</mark>
                                                                                                      VAASIIPFPDHNQSPRN
                                                                                                                                    767
                                              ATEEANA--AINEMDPARRIK--AVQHATTFTHCEIHPSM
tr|A0A1X7QYA1|A0A1X7QYA1_9SACH
                                                                                                      VAAS IIP FPDH NQS PRN
                                                                                                                                    764
                                              VIEEEN PI---DDMDPAKRIK--ASQNATTFTHCEIH PSM<mark>I</mark>
tr|J7RV95|J7RV95_KAZNA
                                                                                                     VAASIIP FPDHNQS PRN
                                                                                                                                    763
                                              MGEEEETQ-NDTAMDPAKRIK--ATQNATTFTHCEIHPSMT
TAVEQDI--PKENVDLAKRIK--VTHHATTFTHCEIHPSMT
                                                                                                     VAASIIP FPDHNQS PRN
VAASIIP FPDHNQS PRN
tr|H2AVJ8|H2AVJ8 KAZAF
                                                                                                                                    765
sp|Q6FLD5|RPB2_CANGA
                                                                                                                                    766
tr|A0A0L8VHA5|A0A0L8VHA5 9SACH
                                              AEANEE --- NDLDVDPAKRIR--VSHHATT FTHCEIH PSM
                                                                                                     VAASIIPFPDHNQSPRN
                                                                                                                                    767
                                              TEGNEE --- NGLGVDHAKRIR--VTHHATT FTHCEIH PSM
tr|A0A0L8RB33|A0A0L8RB33_SACEU
                                                                                                     VAAS IIP FPDHNQS PRN
                                                                                                                                    767
                                              PMENEE---VIDDMDPAKRIR--ATQHATTFTHCEIHPSM
                                                                                                     VAASIIPFPDHNQSPRN
VAASVIPFPDHNQSPRN
tr|G0VJ71|G0VJ71 NAUCC
                                                                                                                                    767
                                              VQEAHE --- FNEDDDLAKRIR--ATQHATT FTHCEIH PSM
VQENTF---NEDDDDLARRIK--ATHHATT FTHCEVH PSM
tr|G8ZM49|G8ZM49_TORDC
                                                                                                                                    765
                                                                                                     VAAS IIP FPDHNOS PRN
VAAS IIP FPDHNOS PRN
tr|A0A1Q3A090|A0A1Q3A090_ZYGRO
                                                                                                                                    767
                                              VOENSY---AEEDDDLAKRIK-ATOHATTFTHCEIHPSM
VOENPY---SEEEDDLARRIK-ATHHATTFTHCEIHPSM
VOENPY---SEEEDDLARRIK-ATHHATTFTHCEIHPSM
tr|A0A0N7IS35|A0A0N7IS35_9SACH
                                                                                                                                    765
tr|A0A212MG88|A0A212MG88 ZYGBA
                                                                                                     VAASVTPFPDHNOSPRN
                                                                                                                                    766
tr|A0A1S7HHE1|A0A1S7HHE1_9SACH
                                                                                                     VAASVIPFPDHNOSPRN
                                                                                                                                    766
                                              VQENPY---SEEEDDLARRIK--ATHHATTFTHCEIHPSM
tr|S6ESB4|S6ESB4 ZYGB2
                                                                                                     VAASVIPFPDHNQSPRN
                                                                                                                                    766
                                              SRQANAGYEMQEELDPAKRVKPAPNPHVHAWTHCEIHPAM
                                                                                                     ILASIIPFPDHNQSPRN
tr|B6K5Q5|B6K5Q5_SCHJY
                                                                                                                                    756
sp|Q02061|RPB2_SCHPO
                                              SRQMQAGYEVKEELDPAQRVKPAPNPHVHAWTHCEIHPAM<mark>ILG</mark>ILASIIPFPDHNQSPRN
                                              SRQMNAGYEVKEELDPAQRVKPAPNPHVHTYTHCEIHPAM<mark>ILG</mark>ILASIIPFPDHNQSPRN
SRQMNAGYEVKEELDPAQRVKPAPNPHVHAYTHCEIHPAM<mark>ILG</mark>ILASIIPFPDHNQSPRN
tr|S9R8U4|S9R8U4_SCHOY
                                                                                                                                    756
tr|S9W8C6|S9W8C6 SCHCR
                                                                                                                                    756
                                                                                   ::****:*
                                                                                                        **:*******
 sp | P30876 | RPB2_HUMAN
                                               TYQSAMGKQAMGVYITNFHVRMDTLAHVLYYPQKPLVTTRSMEYLRFRELPAGINSIVAI
                                                                                                                                  782
 tr|G3V8Y5|G3V8Y5 RAT
                                              TYOS AMGKO AMGVYITNFHVRMDTLAHVLYYPOKPLVTTRSMEYLRFRELPAGINS IVAI
                                                                                                                                  782
 tr|A0A250Y753|A0A250Y753_CASCN
                                              TYQSAMGKQAMGVYITNFHVRMDTLAHVLYYPQKPLVTTRSMEYLRFRELPAGINSIVAI
                                                                                                                                  782
 tr|A0A1U7R4C7|A0A1U7R4C7_MESAU
                                              TYQSAMGKQAMGVYITNFHVRMDTLAHVLYYPQKPLVTTRSMEYLRFRELPAGINSIVAI
                                                                                                                                  782
                                              TYQSAMGKQAMGVYITNFHVRMDTLAHVLYYPQKPLVTTRSMEYLRFRELPAGINSIVAI
TYQSAMGKQAMGVYITNFHVRMDTLAHVLYYPQKPLVTTRSMEYLRFRELPAGINSIVAI
 tr|A0A286XIQ9|A0A286XIQ9 CAVPO
                                                                                                                                  782
 tr||I3M351||I3M351 | ICTTR
                                                                                                                                  782
 tr|G7P5R6|G7P5R6 MACFA
                                              TYQSAMGKQAMGVYITNFHVRMDTLAHVLYYPQKPLVTTRSMEYLRFRELPAGINSIVAI
                                                                                                                                   782
 tr|H2QPI8|H2QPI8_PANTR
                                              TYQSAMGKQAMGVYITNFHVFMDTLAHVLYYPQKPLVTTRSMEYLRFRELPAGINSIVAI
TYQSAMGKQAMGVYITNFHVRMDTLAHVLYYPQKPLVTTRSMEYLRFRELPAGINSIVAI
                                                                                                                                  782
 tr|A0A1U7V0T5|A0A1U7V0T5 TARSY
                                                                                                                                  782
 tr|A0A1S2ZSL2|A0A1S2ZSL2 ERIEU
                                              TYQSAMGKQAMGVYITNFHVRMDTLAHVLYYPQKPLVTTRSMEYLRFRELPAGINSIVAI
                                                                                                                                  782
 tr|A0A0D9QYL1|A0A0D9QYL1_CHLSB
                                               TYQSAMGKQAMGVYITNFHVRMDTLAHVLYYPQKPLVTTRSMEYLRFRELPAGINSIVAI
                                                                                                                                  775
 tr|A0A2K5ZNR7|A0A2K5ZNR7_MANLE
                                              {\tt TYQSAMGKQAMGVYITNFHVRMDTLAHVLYYPQKPLVTTRSMEYLRFRELPAGINSIVAI}
                                                                                                                                  775
 tr|A0A2I2ZIU3|A0A2I2ZIU3 GORGO
                                              TYOSAMGKOAMGVYITNFHVRMDTLAHVLYYPOKPLVTTRSMEYLRFRELPAGINSIVAI
                                                                                                                                  782
 tr|A0A1D5QGA5|A0A1D5QGA5_MACMU
                                               TYQSAMGKQAMGVYITNFHVRMDTLAHVLYYPQKPLVTTRSMEYLRFRELPAGINSIVAI
 tr|A0A2J8S2N1|A0A2J8S2N1_PONAB
                                               TYQSAMGKQAMGVYITNFHVRMDTLAHVLYYPQKPLVTTRSMEYLRFRELPAGINSIVAI
                                                                                                                                  782
 tr|A0A2K5K5J5|A0A2K5K5J5_COLAP
tr|A0A2J8PEW7|A0A2J8PEW7 PANTR
                                              TYQSAMGKQAMGVYITNFHVPMDTLAHVLYYPQKPLVTTRSMEYLRFRELPAGINSIVAI
TYQSAMGKQAMGVYITNFHVPMDTLAHVLYYPQKPLVTTRSMEYLRFRELPAGINSIVAI
                                                                                                                                  782
                                                                                                                                   775
 tr|A0A2K5CY83|A0A2K5CY83 AOTNA
                                               TYQSAMGKQAMGVYITNFHVRMDTLAHVLYYPQKPLVTTRSMEYLRFRELPAGINSIVAI
                                                                                                                                  782
 tr|A0A096NEY4|A0A096NEY4_PAPAN
tr|C9J2Y9|C9J2Y9_HUMAN
tr|G8BY61|G8BY61_TETPH
                                              TYQSAMGKQAMGVYITNFHVRMDTLAHVLYYPQKPLVTTRSMEYLRFRELPAGINSIVAI
TYQSAMGKQAMGVYITNFHVRMDTLAHVLYYPQKPLVTTRSMEYLRFRELPAGINSIVAI
                                                                                                                                  782
                                                                                                                                  775
                                               TYOSAMGKOAMGVFLTNYNVRMDTMANILYYPOKPLGTTRAMEYLKFRELPAGONAIVAI
                                                                                                                                  827
                                              TYQSAMGKQAMGVFLTNYNVRMDTMAN ILYYPQKPLGTTRSMEYLK FRELPAGQNA IVAI
TYQSAMGKQAMGVFLTNYNVRMDTMAN ILYYPQKPLGTTRAMEYLK FRELPAGQNA IVAI
 tr|A0A1X7QYA1|A0A1X7QYA1_9SACH
                                                                                                                                  824
 tr|J7RV95|J7RV95_KAZNA
tr|H2AVJ8|H2AVJ8 KAZAF
                                                                                                                                  823
                                              TYQSAMGKQAMGVFLTNFNVRMDTMANILYYPQKPLGTTRAMEYLKFRELPAGQNAIVAI
                                                                                                                                  825
                                              TYQSAMGKQAMGVFLTNYNFRMDTMANILYYPQKPLGTTRAMEYLKFRELPAGQNAIVAI
 sp | Q6FLD5 | RPB2_CANGA
 tr|A0A0L8VHA5|A0A0L8VHA5 9SACH
                                              TYOSAMGKOAMGVFLTNYNVRMDTMANILYYPOKPLGTTRAMEYLKFRELPAGONAIVAI
                                                                                                                                  827
 tr|A0A0L8RB33|A0A0L8RB33_SACEU
                                              TYQSAMGKQAMGVFLTNYNVRMDTMANILYYPQKPLGTTRAMEYLKFRELPAGQNAIVAI
                                                                                                                                  827
 tr|G0VJ71|G0VJ71_NAUCC
                                              TYQSAMGKQAMGVFLTNYNVRMDTMANILYYPQKPLGTTRSMEYLKFRELPAGQNAIVAI
                                                                                                                                  827
tr|G8ZM49|G8ZM49_TORDC
tr|A0A1Q3A090|A0A1Q3A090_ZYGRO
tr|A0A0N7IS35|A0A0N7IS35_9SACH
                                              TYQSAMGKQAMGVFLTNYNVPMDTMANILYYPQKPLGTTRAMEYLKFRELPAGQNAIVAI
TYQAAMGKQAMGVFLTNYNVPMDTMANILYYPQKPLGTTRAMEYLKFRELPAGQNAIVAI
                                                                                                                                  825
                                                                                                                                  827
                                               TYQSAMGKQAMGVFLTNYNVRMDTMANILYYPQKPLGTTRAMEYLKFRELPAGQNAIVAI
                                                                                                                                  825
 tr|A0A212MG88|A0A212MG88 ZYGBA
                                              TYQSAMGKQAMGVFLTNYNVPMDTMAN ILYYPQKPLGTTRAMEYLKFRELPAGQNA IVAI
TYOSAMGKOAMGVFLTNYNVPMDTMAN ILYYPOKPLGTTRAMEYLKFRELPAGONA IVAI
                                                                                                                                  826
 tr|A0A1S7HHE1|A0A1S7HHE1 9SACH
                                                                                                                                  826
                                               TYOSAMGKQAMGVFLTNYNVRMDTMANILYYPOKPLGTTRAMEYLKFRELPAGONAIVAI
 tr|S6ESB4|S6ESB4_ZYGB2
 tr|B6K5Q5|B6K5Q5_SCHJY
                                              TYQS AMGKQAMGIYLTNYQVRMDTMANILYYPQKPLATTRSMEYLKFRELPAGQNAIVAI
                                                                                                                                  816
 sp|Q02061|RPB2_SCHPO
                                              TYQSAMGKQAMGVYLTNYQVRMDTMANILYYPQKPLATTRSMEYLKFRELPAGQNAIVAI
                                                                                                                                  816
 tr|S9R8U4|S9R8U4 SCHOY
                                              TYQSAMGKQAMGVYLTNYQVRMDTMANILYYPQKPLATTRSMEYLKFRELPAGQNAIVAI
                                                                                                                                  816
                                              TYQSAMGKQAMGVYLTNYQVRMDTMANILYYPQKPLATTRSMEYLKFRELPAGQNAIVAI
 tr|S9W8C6|S9W8C6 SCHCR
```

****-*--****** ***-***-**** *-***



```
IVDQVMVTLNQEGYKFCKIRVRSVRIPQI
                                                                                            GDKFASRHGOKGTCGIQYRDEDMPFTCEGIT
sp|P30876|RPB2 HUMAN
                                                                                                                                                  962
                                                   IVDQVMVTLNQEGYKFCKIRVRSVRIPQ
IVDQVMVTLNQEGYKFCKIRVRSVRIPQ
                                                                                               KFASRHGOKGTCGIOYROEDMPFTCEGIT
KFASRHGOKGTCGIOYROEDMPFTCEGIT
tr|G3V8Y5|G3V8Y5 RAT
                                                                                                                                                  962
tr|A0A250Y753|A0A250Y753 CASCN
                                                                                                                                                  962
tr|A0A1U7R4C7|A0A1U7R4C7_MESAU
                                                   IVDQVMVTLNQEGYKFCKIRVRSVRIPQ
                                                                                               KFASRHGQKGTCGIQYRQEDMPFTCEGIT
                                                                                                                                                  962
tr|A0A286XIQ9|A0A286XIQ9 CAVPO
                                                   IVDQVMVTLNQEGYKFCKIRVRSVRIPQ
                                                                                               KFASRHGQKGTCGIQYRQEDMPFTCEGIT
                                                                                                                                                  962
                                                   IVDOVMVTLNQEGYKFCKIRVRSVRIPQ
tr|I3M351|I3M351_ICTTR
tr|G7P5R6|G7P5R6_MACFA
tr|H2QPI8|H2QPI8_PANTR
                                                                                               KFASRHGOKGTCGIØYRDEDMPFTCEGIT
KFASRHGOKGTCGIØYRDEDMPFTCEGIT
                                                                                                                                                  962
                                                   IVDQVMVTLNQEGYKFCKIRVRSVRIPQ
                                                                                                                                                  962
                                                                                               KFASRHGOKGTCGIOYROEDMPFTCEGIT
                                                   IVDQVMVTLNQEGYKFCKIRVRSVRIPQ
                                                                                                                                                  962
tr|A0A1U7V0T5|A0A1U7V0T5_TARSY
                                                   IVDQVMVTLNQEGYKFCKIRVRSVRIPQ
                                                                                               KFASRHGQKGTCGIQYRQEDMPFTCEGIT
                                                                                                                                                  962
                                                   IVDQVMVTLNQEGYKFCKIRVRSVRIPQ
                                                                                               KFASRHGOKGTCGIØYRDEDMPFTCEGIT
KFASRHGOKGTCGIØYRDEDMPFTCEGIT
tr|A0A1S2ZSL2|A0A1S2ZSL2_ERIEU
                                                                                                                                                  962
tr|A0A0D9QYL1|A0A0D9QYL1 CHLSB
                                                   IVDQVMVTLNQEGYKFCKIRVRSVRIPQ
                                                                                                                                                  955
                                                                                               KFASRHGOKGTCGIOYROEDMPFTCEGIT
tr|A0A2K5ZNR7|A0A2K5ZNR7_MANLE
tr|A0A2I2ZIU3|A0A2I2ZIU3_GORGO
                                                   IVDOVMVTLNOEGYKFCKIRVRSVRIPO
                                                                                                                                                  955
                                                                                               KFASRHGOKGTCGIOYRDEDMPFTCEGIT
                                                   IVDQVMVTLNQEGYKFCKIRVRSVRIPQ
                                                                                                                                                  962
                                                                                                KFASRHGOKGTCGI OYR DEDMPFTCEGIT
tr|A0A1D5QGA5|A0A1D5QGA5 MACMU
                                                   IVDQVMVTLNQEGYKFCKIRVRSVRIPQ
                                                                                                                                                  962
tr|A0A2J8S2N1|A0A2J8S2N1_PONAB
                                                   IVDQVMVTLNQEGYKFCKIRVRSVRIPQ
                                                                                                KFASRHGOKGTCGIOYRDEDMPFTCEGIT
                                                                                                                                                  962
                                                                                               KFASRHGQKGTCGIQYRQEDMPFTCEGIT
tr|A0A2K5K5J5|A0A2K5K5J5_COLAP
                                                   IVDQVMVTLNQEGYKFCKIRVRSVRIPQ
                                                                                                                                                  962
                                                                                               KFASRHGOKGTCGIQYRQEDMPFTCEGIT
                                                   IVDOVMVTLNOEGYKFCKIRVRSVRI PO
tr|A0A2J8PEW7|A0A2J8PEW7_PANTR
tr|A0A2K5CY83|A0A2K5CY83_AOTNA
                                                                                                                                                  955
                                                                                               KFASRHGOKGTCGIOYROEDMPFTCEGIT
                                                   IVDQVMVTLNQEGYKFCKIRVRSVRIPQ
                                                                                                                                                  962
                                                   IVDQVMVTLNQEGYKFCKIRVRSVRIPQ
                                                                                               KFASRHGOKGTCGIOYRDEDMPFTCEGIT
tr|A0A096NEY4|A0A096NEY4_PAPAN
                                                                                                                                                  962
tr|C9J2Y9|C9J2Y9_HUMAN
                                                   IVDQVMVTLNQEGYKFCKIRVRSVRIPQ
                                                                                               KFASRHGQKGTCGIQYRQEDMPFTCEGIT
                                                                                                                                                  955
                                                                                               kfasrhgokgtigitygredmpftaegiv
tr|G8BY61|G8BY61_TETPH
                                                   IVDQVLVTTNQDGLKFVKVRVRTTKVPQ
                                                                                                                                                  1007
                                                                                               KFASRHGOKGTIGITYRREDMPFTAEGIV
KFASRHGOKGTIGITYSREDMPFTAEGIV
KFASRHGOKGTIGITYRREDMPFTADGIV
tr|A0A1X7QYA1|A0A1X7QYA1 9SACH
                                                   IIDQVLVTTNQDGLKFVKVRVRTTKVPQ
                                                                                                                                                  1004
tr|J7RV95|J7RV95_KAZNA
tr|H2AVJ8|H2AVJ8_KAZAF
                                                   IIDOVLVTTNODGLKFVKVRVRTTKVPO
                                                                                                                                                  1003
                                                   IVDQVLITTNQDGLKFVKVRVRTTKVPQ
                                                                                                                                                  1005
sp|Q6FLD5|RPB2_CANGA
                                                   IVDQVLITTNQDGLKFVKVRVRTTKVPQ
                                                                                               kfasrhgqkgtigityrredmpftaegiv
                                                                                                                                                  1006
                                                                                               KFASRHGOKGTIGITYRREDMPFTAEGIV
KFASRHGOKGTIGITYRREDMPFTAEGIV
KFASRHGOKGTIGITYRREDMPFTAEGIV
                                                   IVDQVLVTTNQDGLKFVKVRVRTTKIPQ
IVDQVLVTTNQDGLKFVKVRVRTTKVXQ
tr|A0A0L8VHA5|A0A0L8VHA5_9SACH
                                                                                                                                                  1007
tr|A0A0L8RB33|A0A0L8RB33 SACEU
                                                                                                                                                  1007
tr|G0VJ71|G0VJ71 NAUCC
                                                   IVDQVLITTNQDGLKFVKVRVRTTKVPQ
                                                                                                                                                  1007
tr|G8ZM49|G8ZM49_TORDC
                                                   IVDQVLITTNQDGLKFVKVRVRTTKVPQ
                                                                                               kfasrhgokgtigitygredmpftaegiv
                                                                                                                                                  1005
tr|A0A1Q3A090|A0A1Q3A090_ZYGRO
                                                   IVDQILITTNQDGLKFVKVRVRTTKIPQ
                                                                                               kfasrhgqkgtigitygredmpftaegiv
                                                                                                                                                  1007
tr|A0A0N7IS35|A0A0N7IS35_9SACH
tr|A0A212MG88|A0A212MG88_ZYGBA
                                                   IVDQILITTNQDGLKFVKVRVRTTKI PQ
                                                                                               KFASRHGOKGTIGITYGREDMPFTSEGIV
KFASRHGOKGTIGITYSREDMPFTAEGIV
                                                                                                                                                  1005
                                                   IVDQILITTNQDGLKFVKVRVRTTKIPQ
                                                                                                                                                  1006
tr|A0A1S7HHE1|A0A1S7HHE1 9SACH
                                                   IVDQILITTNQDGLKFVKVRVRTTKIPQ
                                                                                               KFASRHGQKGTIGITYSREDMPFTAEGIV
                                                                                                                                                  1006
                                                   IVDQILITTNQDGLKFVKVRVRTTKI PQ
                                                                                               KFASRHGQKGTIGITYSREDMPFTAEGIV
tr|S6ESB4|S6ESB4_ZYGB2
                                                                                                                                                  1006
                                                                                              AFASAHGUNGTIGHTISE DWFFIAQGIV
KFASRHGOKGTIGMTYRHEDMPFSAQGIV
KFASRHGOKGTIGMTYRHEDMPFSAQGIV
KFASRHGOKGTIGMTYRHEDMPFSAQGIV
KFASRHGOKGTIGMTYRHEDMPFSAQGIV
                                                   IVDQVMVTTNQEGLKFVKVRMRSTRI PQ
tr|B6K5Q5|B6K5Q5_SCHJY
                                                                                                                                                  996
sp|Q02061|RPB2_SCHPO
tr|S9R8U4|S9R8U4 SCHOY
                                                   IVDQVMVTTNQEGLKFVKVRMRSTRIPQI
IVDQVLVTTNQEGLKFVKVRMRSTRVPQI
                                                                                                                                                  996
                                                                                                                                                  996
                                                   IVDQVLVTTNQEGLKFVKVRMRSTRVPQ<mark>I</mark>
tr|S9W8C6|S9W8C6 SCHCR
                                                                                                                                                  996
                                                   *:**:::* **:* ** *:*:*:::: *
sp|P30876|RPB2 HUMAN
                                                  PDIIIN PHAIPSRMTIGHLIECLQGKVSANKGE IGDATPFNDAVNVQKISNLLSD
                                                                                                                                                1022
tr|G3V8Y5|G3V8Y5 RAT
                                                  PDIIINPHAIPSRMTIGHLIECLQGKVSANKGE GDATPFNDAVNVQKISNLLSD
                                                                                                                                                1022
tr|A0A250Y753|A0A250Y753_CASCN
tr|A0A1U7R4C7|A0A1U7R4C7_MESAU
                                                  PDIIIN PHAIPSRMTIGHLIECLOGKVSANKGE IGDATPFNDAVNVOKISNLLSD
PDIIN PHAIPSRMTIGHLIECLOGKVSANKGE IGDATPFNDAVNVOKISNLLSD
                                                                                                                                                1022
                                                                                                                                                1022
tr|A0A286XIQ9|A0A286XIQ9_CAVPO
                                                  PDIIIN PHAIPSRMTIGHLIECLQGKVSANKGE IGDATPFNDAVNVQKISNLLSD
                                                                                                                                                 1022
tr|I3M351|I3M351_ICTTR
                                                  PDIIIN PHAIPSRMTIGHLIECLQGKVSANKGE GDATPFNDAVNVQKISNLLSD<mark>YGY</mark>HL
                                                                                                                                                1022
                                                  PDIIIN PHAIPSRMTIGHLIECLOGKVSANKGE IGDATPFNDAVNVOKISNLLSD
PDIIN PHAIPSRMTIGHLIECLOGKVSANKGE IGDATPFNDAVNVOKISNLLSD
tr|G7P5R6|G7P5R6_MACFA
tr|H2QPI8|H2QPI8_PANTR
                                                                                                                                                1022
                                                                                                                                                1022
tr|A0A1U7V0T5|A0A1U7V0T5_TARSY
                                                  PDIIIN PHAIPSRMTIGHLIECLQGKVSANKGE IGDATPFNDAVNVQKISNLLSD<mark>YGY</mark>HL
                                                                                                                                                1022
tr|A0A1S2ZSL2|A0A1S2ZSL2_ERIEU
                                                  PDIIINPHAIPSRMTIGHLIECLOGKVSANKGE IGDATPFNDAVNVOKISNLLSD
                                                                                                                                                1022
                                                  PDIIIN PHAIPSRMIIGHLIECLOGKVSANKGE IGDATPFNDAVNVOKISNLLSDYG YHL
PDIIIN PHAIPSRMIIGHLIECLOGKVSANKGE IGDATPFNDAVNVOKISNLLSDYG YHL
tr|A0A0D9QYL1|A0A0D9QYL1_CHLSB
tr|A0A2K5ZNR7|A0A2K5ZNR7_MANLE
                                                                                                                                                1015
                                                                                                                                                1015
                                                  PDIIIN PHAIPSRMTIGHLIECLOGKVSANKGE GDATPFNDAVNVOKISNLLSD<mark>YGY</mark>HL
PDIIIN PHAIPSRMTIGHLIECLOGKVSANKGE GDATPFNDAVNVOKISNLLSD<mark>YGY</mark>HL
PDIIIN PHAIPSRMTIGHLIECLOGKVSANKGE GDATPFNDAVNVOKISNLLSD<mark>YGY</mark>HL
tr|A0A2I2ZIU3|A0A2I2ZIU3 GORGO
                                                                                                                                                1022
tr|A0A1D50GA5|A0A1D50GA5 MACMU
                                                                                                                                                1022
tr|A0A2J8S2N1|A0A2J8S2N1 PONAB
                                                                                                                                                1022
tr|A0A2K5K5J5|A0A2K5K5J5|COLAP
tr|A0A2J8PEW7|A0A2J8PEW7_PANTR
tr|A0A2K5CY83|A0A2K5CY83_AOTNA
                                                  PDIIIN PHA IPSRMTIGHLIECL QGKVSANKGE IGDA TPFNDAVNVQKISNLLS DV
PDIIIN PHA IPSRMTIGHLIECL QGKVSANKGE IGDA TPFNDAVNVQKISNLLS DV
PDIIIN PHA IPSRMTIGHLIECL QGKVSANKGE IGDA TPFNDAVNVQKISNLLS DV
                                                                                                                                                1015
                                                                                                                                                1022
tr|A0A096NEY4|A0A096NEY4_PAPAN
                                                  PDIIINPHAIPSRMTIGHLIECLQGKVSANKGE GDATPFNDAVNVQKISNLLSD
                                                                                                                                                1022
tr|C9J2Y9|C9J2Y9_HUMAN
tr|G8BY61|G8BY61_TETPH
                                                  PDIIIN PHAIPSRMTIGHLIECLOGKVSANKGE GDATPFNDAVNVOKISNLLSD
PDLIIN PHAIPSRMTVAHLIECLLSKVAALSGNEGDASPFTD-ITVEGISKLLRE
                                                                                                                                                1015
                                                                                                                                                1066
tr|A0A1X7QYA1|A0A1X7QYA1_9SACH
                                                  PDLIINPHAIPSRMTVAHLIECLLSKVAALSGNEGDASPFTD-ITVEGISTLLRE
                                                                                                                                                 1063
                                                  PDLIIN PHA IPSRMTVAHLIEC. LSKVAALSGNIGDA SPFTD-ITVEGISKLIRE<mark>HG</mark>YHS
PDLIIN PHA IPSRMTVAHLIEC. LSKVAALSGNIGDA SPFTD-ITVEGISKLIREHGYHS
PDLIIN PHA IPSRMTVAHLIEC. LSKVAALSGNIGDA SPFTD-ITVEGISKLIRE<mark>HGY</mark>OS
tr|J7RV95|J7RV95_KAZNA
                                                                                                                                                1062
tr|H2AVJ8|H2AVJ8 KAZAF
                                                                                                                                                1064
sp|Q6FLD5|RPB2_CANGA
                                                                                                                                                1065
                                                  PDLIIN PHAIPSRMTVAHLIECLISKVAALSGNEGDASPFTD-ITVEGISKLIRE
tr|A0A0L8VHA5|A0A0L8VHA5_9SACH
                                                                                                                                                1066
                                                  PDLIINPHAIPSRMTVAHLIECLLSKVAALSCNEGDASPFTD-TTVEGISKLLRENGYDS
PDLIINPHAIPSRMTVAHLIECLLSKVAALSCNEGDASPFTD-TTVEGISKLLRENGYDS
tr|A0A0L8RB33|A0A0L8RB33 SACEU
                                                                                                                                                1066
tr|G0VJ71|G0VJ71_NAUCC
                                                                                                                                                1066
tr|G8ZM49|G8ZM49_TORDC
tr|A0A1Q3A090|A0A1Q3A090_ZYGRO
tr|A0A0N7IS35|A0A0N7IS35_9SACH
                                                  PDLIIN PHA IPSRMTVAHLIECLLSKVAALSGNEGDASPFTD-ITVE GISKLLREH
PDLIIN PHA IPSRMTVAHLIECLLSKVAALSGNEGDASPFTD-ITVE GISKLLREH
PDLIIN PHA IPSRMTVAHLIECLLSKVAALSGNEGDASPFTD-ITVE GISKLLREH
                                                                                                                                                1064
                                                                                                                                                1066
                                                                                                                                                1064
                                                  PDLIINPHAIPSRMTVAHLIECLLSKVAALSGNIGDASPFTD-ITVEGISKLLRE
PDLIINPHAIPSRMTVAHLIECLLSKVAALSGNIGDASPFTD-ITVEGISKLLREH
PDLIINPHAIPSRMTVAHLIECLLSKVAALSGNIGDASPFTD-ITVEGISKLLREH
tr|A0A212MG88|A0A212MG88_ZYGBA
                                                                                                                                                 1065
tr|A0A1S7HHE1|A0A1S7HHE1_9SACH
                                                                                                                                                1065
tr|S6ESB4|S6ESB4 ZYGB2
                                                                                                                                                1065
tr|B6K5Q5|B6K5Q5_SCHJY
                                                  PDIIINPHAIPSRMTVAHLVECOLSKVSALSGLEGDATPFTE-VTVE AVSKLLRS
                                                                                                                                                1055
                                                  sp|Q02061|RPB2_SCHPO
tr|S9R8U4|S9R8U4 SCHOY
                                                                                                                                                1055
                                                                                                                                                1055
tr|S9W8C6|S9W8C6 SCHCR
```

```
sp|P30876|RPB2 HUMAN
                                                          RGNEVLYNGFTGRKITSOIFIGPTYYORLKHMVDDKIHSRARGPIOILNROPMEGRSRDG
                                                                                                                                                                      1082
tr|G3V8Y5|G3V8Y5 RAT
                                                                                    SQIFIGPTYYQRLKHMVDDKIHSRARGPIQILNRQPMEGRSRDG
                                                          RGN EVL YNG FTGRKI
                                                                                                                                                                      1082
tr|A0A250Y753|A0A250Y753_CASCN
tr|A0A1U7R4C7|A0A1U7R4C7_MESAU
                                                                                    SQIFIGPTYYORIKHMYDDKTHSRARGPIQILNROPMEGRSRDG
SQIFIGPTYYORIKHMYDDKIHSRARGPIQILNROPMEGRSRDG
                                                                                                                                                                       1082
                                                               EVLYNG FTGRKI
                                                          RGNEVLYNGFTGRKIT
                                                                                                                                                                      1082
tr|A0A286XIQ9|A0A286XIQ9_CAVPO
                                                          RGN EVL YNG FTGRKIT
                                                                                    SQIFIGPTYYQRLKHMVDDKIHSRARGPIQILNRQPMEGRSRDG
                                                                                                                                                                      1082
tr||I3M351||I3M351 | ICTTR
                                                          RGNEVT, YNG FTGRKTT
                                                                                    BOTETGPTYYORLKHMVDDKTHSRARGPTOTLNROPMEGRSRDG
                                                                                                                                                                      1082
tr|G7P5R6|G7P5R6_MACFA
tr|H2QPI8|H2QPI8 PANTR
                                                                                    SQIFIGPTYYQRLKHMVDDKIHSRARGPIQILNRQPMEGRSRD
                                                                                    SOIFIGPTYYORLKHMVDDKIHSRARGPIOILNROPMEGRSRDG
                                                          RGN EVL YNG FTGRKII
                                                                                                                                                                      1082
tr|A0A1U7V0T5|A0A1U7V0T5_TARSY
                                                                                    SQIFIGPTYYQRLKHMVDDKIHSRARGPIQILNRQPMEGRSRD
                                                                                                                                                                      1082
                                                          RGFEVLYNGFTGRKITSOIFIGPTYYQRLKHMVDDKIHSRARGPIQIINRQPMEGRSRDG
RGFEVLYNGFTGRKITSOIFIGPTYYQRLKHMVDDKIHSRARGPIQIINRQPMEGRSRDG
RGFEVLYNGFTGRKITSOIFIGPTYYQRLKHMVDDKIHSRARGPIQIINRQPMEGRSRDG
RGFEVLYNGFTGRKITSQIFIGPTYYQRLKHMVDDKIHSRARGPIQIINRQPMEGRSRDG
tr|A0A1S2ZSL2|A0A1S2ZSL2_ERIEU
                                                                                                                                                                      1082
tr|A0A0D9QYL1|A0A0D9QYL1_CHLSB
tr|A0A2K5ZNR7|A0A2K5ZNR7_MANLE
tr|A0A2I2ZIU3|A0A2I2ZIU3_GORGO
                                                                                                                                                                       1075
                                                                                                                                                                      1075
                                                                                                                                                                      1082
tr|A0A1D5OGA5|A0A1D5QGA5 MACMU
                                                          RGWEVLYNG FTGRKIT SQIFIGPTYYORLKHMVDDKIHSRARG PIQILNROPMEGRS RDG
RGWEVLYNG FTGRKIT SQIFIGPTYYORLKHMVDDKIHSRARG PIQILNROPMEGRS RDG
                                                                                                                                                                      1082
tr|A0A2J8S2N1|A0A2J8S2N1 PONAB
                                                                                                                                                                      1082
tr|A0A2K5K5J5|A0A2K5K5J5_COLAP
tr|A0A2J8PEW7|A0A2J8PEW7_PANTR
tr|A0A2K5CY83|A0A2K5CY83_AOTNA
                                                          RGN EVL YNG FTGR KIT BOIF IGP TYYORLKHMVDDKI HSRARG PIQI LNR OPME GRS RDG
RGN EVL YNG FTGR KIT BOIF IGP TYYORLKHMVDDKI HSRARG PIQI LNR OPME GRS RDG
                                                                                                                                                                      1082
                                                               EVLYNGFTGRKIT SQIFIGPTYYQRLKHMVDDKIHSRARGPIQILNRQPMEGRSRDG
                                                                                                                                                                      1082
tr|A0A096NEY4|A0A096NEY4 PAPAN
                                                          RGN EVL YNG FTGR KIT BOIFIGPTYY ORLKHMVDDKI HSR ARG PIOI LNR OPMEGRS RDG
                                                                                                                                                                      1082
tr|C9J2Y9|C9J2Y9_HUMAN
tr|G8BY61|G8BY61 TETPH
                                                          RGMEVLYNGFTGRKIT SOIFIGPTYYORLKHMVDDKIHSRARGPIOILNROPMEGRSRDG
RGFEVMYNGHTGKKIMAQIFFGPTYYORLRHMVDDKIHARARGPMOVLTROPVEGRSRDG
                                                                                                                                                                       1075
                                                                                                                                                                      1126
                                                          RGTEVMINGHTGKKUMAQIFFGFTTYQRIRHMYDDKIHARARGFMQVLIKQFVEGRSKDG
RGTEVMYNGHTGKKUMAQIFFGFTYYQRIRHMYDDKIHARARGFMQVLTRQPVEGRSRDG
RGTEVMYNGHTGKKUMAQIFFGFTYYQRIRHMYDDKIHARARGFMQVLTRQPVEGRSRDG
RGTEVMYNGHTGKKUMAQIFFGFTYYQRIRHMYDDKIHARARGFMQVLTRQPVEGRSRDG
RGTEVMYNGHTGKKUMAQIFFGFTYYQRIRHMYDDKIHARARGFMQVLTRQPVEGRSRDG
tr|A0A1X7QYA1|A0A1X7QYA1_9SACH
tr|J7RV95|J7RV95_KAZNA
tr|H2AVJ8|H2AVJ8_KAZAF
                                                                                                                                                                      1122
                                                                                                                                                                      1124
sp|Q6FLD5|RPB2 CANGA
                                                                                                                                                                      1125
tr|A0A0L8VHA5|A0A0L8VHA5_9SACH
                                                         RGF EVMYNGHTGKKIM AQIF FGPTYY QRLRHMVDDKI HARARG PMQVLTR QFVE GRS RDG
RGF EVMYNGHTGKKIM AQIF FGPTYY QRLRHMVDDKI HARARG PMQVLTR QFVE GRS RDG
RGF EVMYNGHTGKKIM AQIF FGPTYY QRLRHMVDDKI HARARG PMQVLTR QFVE GRS RDG
RGF EVMYNGHTGKKIM AQIF FGPTYY QRLRHMVDDKI HARARG PMQVLTR QFVE GRS RDG
                                                                                                                                                                      1126
tr|A0A0L8RB33|A0A0L8RB33 SACEU
                                                                                                                                                                      1126
tr|G0VJ71|G0VJ71_NAUCC
tr|G8ZM49|G8ZM49 TORDC
                                                                                                                                                                       1126
                                                                                                                                                                       1124
tr|A0A1Q3A090|A0A1Q3A090_ZYGRO
tr|A0A0N7IS35|A0A0N7IS35_9SACH
tr|A0A212MG88|A0A212MG88_ZYGBA
                                                          RGFEVMYNGHTGKKIM BOIFFGPTYYQRLRHMVDDKIHARARGPMQVLTRQPVEGRSRDG
RGFEVMYNGHTGKKIMBQIFFGPTYYQRLRHMVDDKIHARARGPMQVLTRQPVEGRSRDG
                                                                                                                                                                      1126
                                                                                                                                                                      1124
                                                         1125
tr|A0A1S7HHE1|A0A1S7HHE1 9SACH
                                                                                                                                                                      1125
tr|S6ESB4|S6ESB4_ZYGB2
tr|B6K5Q5|B6K5Q5_SCHJY
sp|Q02061|RPB2_SCHPO
tr|S9R8U4|S9R8U4_SCHOY
                                                                                                                                                                      1125
                                                                                                                                                                      1115
                                                                                                                                                                      1115
tr|S9W8C6|S9W8C6 SCHCR
                                                           GLR FGEMER DCQI AHG AAQF LRERLFEASD PYQVHVCNL
GLR FGEMER DCQI AHG AAQF LRERLFEASD PYQVHVCNL
sp|P30876|RPB2 HUMAN
                                                                                                                              GIMA-IANTRIHITE
                                                                                                                                                                         1141
trlG3V8Y5IG3V8Y5 RAT
                                                                                                                               MA-IANTRTHTYE
                                                                                                                                                                         1141
tr|A0A250Y753|A0A250Y753_CASCN
tr|A0A1U7R4C7|A0A1U7R4C7_MESAU
                                                           GLR FGEMER DCOI AHGAAOF LRERLFEASD PYOVHV
                                                                                                                                MA-TANTRTHT
                                                                                                                                                                         1141
                                                           GLR FGEMER DCQI AHG AQF LRE RLF EASD PYQVHV
GLR FGEMER DCQI AHG AQF LRE RLF EASD PYQVHV
                                                                                                                                MA-IANTRTHTY
                                                                                                                                                         kdcR
                                                                                                                                                                         1141
tr|A0A286XIQ9|A0A286XIQ9 CAVPO
                                                                                                                                MA-IANTRTHTY
                                                                                                                                                                         1141
                                                          GLR FGEMERDCQL AHGAAQF LRERLF FEASD PYQVHV
tr|I3M351|I3M351 ICTTR
                                                                                                                                MA-IANTRTHTY
                                                                                                                                                                         1141
tr|G7P5R6|G7P5R6_MACFA
                                                                                                                                MA-IANTRTHT
                                                                                                                                                                         1141
tr|H2QPI8|H2QPI8_PANTR
tr|A0A1U7V0T5|A0A1U7V0T5_TARSY
                                                                                                                                MA-IANTRTHT
                                                                                                                                                                         1141
                                                                                                                                MA-IANTRTHTY
                                                                                                                                                                         1141
tr|A0A1S2ZSL2|A0A1S2ZSL2 ERIEU
                                                           GLR FGEMER DCQI AHGAAQFLRERLFEASD PYQVHV
                                                                                                                                MA-IANTRTHT
                                                                                                                                                           ¢R
                                                                                                                                                                         1141
                                                           GLR FGEMER DCQI AHGAAQF LRERLFEASD PYQVHV
tr|A0A0D9QYL1|A0A0D9QYL1 CHLSB
                                                                                                                                MA-IANTRTHTY
                                                                                                                                                         ≀d<mark>¢</mark>R
                                                                                                                                                                         1134
                                                          GLR FGEMERDCQI AHGAAQF LRERLFEASD PYQVHV
tr|A0A2K5ZNR7|A0A2K5ZNR7_MANLE
                                                                                                                                MA-IANTRIHTY
                                                                                                                                                                         1134
tr|A0A2I2ZIU3|A0A2I2ZIU3_GORGO
                                                                                                                                MA-IANTRTHT
                                                                                                                                                                         1141
tr|A0A1D5QGA5|A0A1D5QGA5_MACMU
                                                                                                                                MA-TANTRTHTY
                                                                                                                                                                         1141
tr|A0A2J8S2N1|A0A2J8S2N1 PONAB
                                                                                                                                MA-TANTRTHTY
                                                                                                                                                            tR
                                                                                                                                                                         1141
                                                                                                                                                        RGCR
RGCR
tr|A0A2K5K5J5|A0A2K5K5J5_COLAP
tr|A0A2J8PEW7|A0A2J8PEW7_PANTR
tr|A0A2K5CY83|A0A2K5CY83_AOTNA
                                                           GLR FGEMER DCQI AHGAAQF LRERLFEASD PYQVHV
                                                                                                                                MA-IANTRTHTY
                                                                                                                                                                         1141
                                                           GLR FGEMER DCOI AHGAAOF LRERLFEASD PYOVH
                                                                                                                                MA-IANTRTHT
                                                                                                                                                                         1134
                                                           GLR FGEMER DCQI AHGAAQF LRERLFEASD PYQVHV
                                                                                                                                MA-IANTRTHT
                                                                                                                                                                         1141
tr|A0A096NEY4|A0A096NEY4 PAPAN
                                                           GLR FGEMER DCQI AHG AAQF LRERLFEASD PYQVHV
                                                                                                                                MA-IANTRTHTY
                                                                                                                                                                         1141
                                                           GLR FGEMER DCQI AHG AAQF LRERLFE ASD PYQVHV
GLR FGEMER DCMI AHG AAAF LKE RLMEA SDAFRVHV
GLR FGEMER DCMI AHG AAAF LKE RLMEA SDAFRVHI
tr|C9J2Y9|C9J2Y9_HUMAN
                                                                                                                                MA-IANTRTHT
                                                                                                                                                                         1134
tr|G8BY61|G8BY61_TETPH
tr|A0A1X70YA1|A0A1X70YA1 9SACH
                                                                                                                                                        KGCE
                                                                                                                                                                         1186
                                                                                                                                MTVIAKLNHNQI
                                                                                                                                MSVIAKLSHNO
                                                                                                                                                                         1183
tr|J7RV95|J7RV95 KAZNA
                                                           GLR FGEMER DCMI AHGAAAF LKERLMEASDSFRVHI
                                                                                                                                MSVIAKLSHNO
                                                                                                                                                        KGCN
                                                                                                                                                                         1182
tr|H2AVJ8|H2AVJ8_KAZAF
                                                           GLR FGEMER DCMI AHGAAAF LKE RLMEASDAFRVHI
                                                                                                                                MSVIAKLSHNQI
                                                                                                                                                           ¢N
                                                                                                                                                                         1184
sp | Q6FLD5 | RPB2_CANGA
                                                           GLR FGEMERDCMI AHGAAAF LKE RLMEASDAFRVH:
                                                                                                                                MSVIAKLNHNQI
                                                                                                                                                            D
                                                                                                                                                                         1185
                                                           GLR FGEMER DCMI AHGAASF LKE RIMEA SDAFR VHI
GLR FGEMER DCMI AHGAASF LKE RIMEA SDAFR VHI
GLR FGEMER DCMI AHGAAAF LKE RIMEA SDAFR VHI
tr|A0A0L8VHA5|A0A0L8VHA5_9SACH
                                                                                                                                MTVIAKLNHNO
                                                                                                                                                                         1186
                                                                                                                               MTVIAKLNHNO
tr|A0A0L8RB33|A0A0L8RB33 SACEU
                                                                                                                                                                         1186
tr|G0VJ71|G0VJ71_NAUCC
tr|G8ZM49|G8ZM49 TORDC
                                                                                                                                MSVIAKLNHNO
                                                                                                                                                                         1186
                                                           GLR FGEMER DCMI AHGAAAF LKE RLMEASDAFRVHI
                                                                                                                                MSVIAKLNHNO
                                                                                                                                                                         1184
tr|A0A1Q3A090|A0A1Q3A090_ZYGRO
tr|A0A0N7IS35|A0A0N7IS35_9SACH
                                                           GLR FGEMER DCMI AHGAAAF LKERLMEASDAFRVHI
                                                                                                                                MSVVAKLNHNQ
                                                                                                                                                                         1186
                                                           GLR FGEMER DCMI AHGAASFLKERLMEASDAFRVH
                                                                                                                                MSVVAKLNHNQI
                                                                                                                                                            D
                                                                                                                                                                         1184
                                                          GLR FGEMERDCMI AHGAASF LKE RLMEASDAFR VHI
GLR FGEMERDCQI SHGÇSSV LRE RLF DQSDAYR VIV
tr|A0A212MG88|A0A212MG88 ZYGBA
                                                                                                                                MSVVAKLNHNQ<mark>FE</mark>
                                                                                                                                                                         1185
                                                                                                                                                                         1185
tr|A0A1S7HHE1|A0A1S7HHE1_9SACH
                                                                                                                                MSVVAKLNHNO
                                                                                                                                                            D
tr|S6ESB4|S6ESB4_ZYGB2
tr|B6K5Q5|B6K5Q5_SCHJY
sp|Q02061|RPB2_SCHPO
                                                                                                                                MSVVAKLNHNO
                                                                                                                                                                         1185
                                                                                                                                IA-IASFKKNS
                                                                                                                                                            Q
                                                                                                                                                                         1174
                                                           GLR FGEMER DCQI SHGCSSVLRERLFDCSDAYRVIV
                                                                                                                                IA-IASYKKDSY
                                                                                                                                                                         1174
tr|S9R8U4|S9R8U4 SCHOY
                                                           GLR FGEMER DCQI SHG CSSVMRERLFD CSDAYRVVV
                                                                                                                                VA-IASYKKNSY
                                                                                                                                                                         1174
tr|S9W8C6|S9W8C6_SCHCR
                                                           GLRFGEMERDCQISHGCSSVMRERLFDCSDAYRVVV
                                                                                                                             GI.VA-IASYKKNS
                                                                                                                                                                         1174
```

****** *:**

::***::.

:::*.

sp P30876 RPB2_HUMAN	NKTQISLVRMPYACKLLFQELMSMSIAPRMMSV	1174
tr G3V8Y5 G3V8Y5_RAT	NKTQISLVRMPYACKLLFQELMSMSIAPRMMSV	1174
tr A0A250Y753 A0A250Y753_CASCN	NKTQISLVRMPYACKLLFQELMSMSIAPRMMSV	1174
tr A0A1U7R4C7 A0A1U7R4C7_MESAU	NKTQISLVRMPYACKLLFQELMSMSIAPRMMSV	1174
tr A0A286XIQ9 A0A286XIQ9_CAVPO	NKTQISLVRMPYACKLLFQELMSMSIAPRMMSV	1174
tr I3M351 I3M351_ICTTR	NKTQISLVRMPYACKLLFQELMSMSIAPRMMSV	1174
tr G7P5R6 G7P5R6_MACFA	NKTQISLVRMPYACKLLFQELMSMSIAPRMMSV	1174
tr H2QPI8 H2QPI8_PANTR	NKTQISLVRMPYACKLLFQELMSMSIAPRMMSV	1174
tr A0A1U7V0T5 A0A1U7V0T5_TARSY	NKTQISLVRMPYACKLLFQELMSMSIAPRMMSV	1174
tr A0A1S2ZSL2 A0A1S2ZSL2_ERIEU	NKTQISLVRMPYACKLLFQELMSMSIAPRMMSV	1174
tr A0A0D9QYL1 A0A0D9QYL1_CHLSB	NKTQISLVRMPYACKLLFQELMSMSIAPRMMSV	1167
tr A0A2K5ZNR7 A0A2K5ZNR7_MANLE	NKTQISLVRMPYACKLLFQELMSMSIAPRMMSV	1167
tr A0A2I2ZIU3 A0A2I2ZIU3_GORGO	NKTQISLVRMPYACKLLFQELMSMSIAPRMMSV	1174
tr A0A1D5QGA5 A0A1D5QGA5_MACMU	NKTQISLVRMPYACKLLFQELMSMSIAPRMMSV	1174
tr A0A2J8S2N1 A0A2J8S2N1_PONAB	NKTQISLVRMPYACKLLFQELMSMSIAPRMMSV	1174
tr A0A2K5K5J5 A0A2K5K5J5_COLAP	NKTQISLVRMPYACKLLFQELMSMSIAPRMMSV	1174
tr A0A2J8PEW7 A0A2J8PEW7_PANTR	NKTQISLVRMPYACKLLFQELMSMSIAPRMMSV	1167
tr A0A2K5CY83 A0A2K5CY83_AOTNA	NKTQISLVRMPYACKLLFQELMSMSIAPRMMSV	1174
tr A0A096NEY4 A0A096NEY4 PAPAN	NKTQISLVRMPYACKLLFQELMSMSIAPRMMSV	1174
tr C9J2Y9 C9J2Y9_HUMAN	NKTQISLVRMPYACKLLFQELMSMSIAPRMMSV	1167
tr G8BY61 G8BY61_TETPH	NKIDIYQIR PYAAKLLFQELMAMNITPRLYTDRSKNF	1224
tr A0A1X7QYA1 A0A1X7QYA1 9SACH	NKIDIYQIQIPYAAKLLFQELMAMNITPRLYTDRSRDF	1221
tr J7RV95 J7RV95_KAZNA	NKIDIYQIR PYAAKLLFQELMAMNITPRLYTDRSRDF	1220
tr H2AVJ8 H2AVJ8_KAZAF	NKI DIYQIH IPYAAKLLFQELMAMNI TPRLFTDRSRDF	1222
sp Q6FLD5 RPB2_CANGA	NKIDIYQIH PYAAKLLFQELMAMNITPRLYTDRSRDF	1223
sp P08518 RPB2_YEAST	NKIDIYQIHIPYAAKLLFQELMAMNITPRLYTDRSRDF	1224
tr A0A0L8VHA5 A0A0L8VHA5_9SACH	NKIDIYQIHİPYAAKLLFQELMAMNITPRLYTDRSRDF	1224
tr A0A0L8RB33 A0A0L8RB33_SACEU	NKIDIYQIH PYAAKLLFQELMAMNITPRLYTDRSRDF	1224
tr G0VJ71 G0VJ71_NAUCC	NKI DIYQIH IPYAAKLLFQELMAMNI TPRLYTERSRDF	1224
tr G8ZM49 G8ZM49_TORDC	NKIDIYQIN PYAAKLLFQELMAMNITPRLYTDRSKDF	1222
tr A0A1Q3A090 A0A1Q3A090_ZYGRO	NKIDIYQIH PYAAKLLFQELMAMNITPRLYTDRSKDF	1224
tr A0A0N7IS35 A0A0N7IS35_9SACH	NKIDIFQIH PYAAKLLFQELMAMNITPRLYTDRSKDF	1222
tr A0A212MG88 A0A212MG88_ZYGBA	NKI DIYQIH IPYAAKLLFQELMAMNI TPRLYTDRSKDF	1223
tr A0A1S7HHE1 A0A1S7HHE1_9SACH	NKI DIYQIH IPYAAKLLFQELMAMNI TPRLYTDRSKDF	1223
tr S6ESB4 S6ESB4_ZYGB2	NKIDIYQIH PYAAKLLFQELMAMNITPRLYTDRSKDF	1223
tr B6K5Q5 B6K5Q5_SCHJY	NRTRFSQIYLPYAAKLLFQELMSMNIAPRLFTKSHH	1210
sp Q02061 RPB2_SCHPO	NRTRFSQVYLPYAAKLLFQELMSMNIAPRLFTKNHK	1210
tr S9R8U4 S9R8U4 SCHOY		
	NRTRFSQIYLPYAAKLLFQELMSMNIAPRLFTKNHKI-	1211
tr S9W8C6 S9W8C6_SCHCR	NRTRFSQIYLPYAAKLLFQELMSMNIAPRLFTKNHKI- NRTRFSQVYLPYAAKLLFQELMSMNIAPRLFTKNHKN- *: : : *** *** *** *** ***************	1211 1211

Fig. 3 MSA of the Rpb2, initiation subunits of eukaryotic RNAPs II

P30876|RPB2_HUMAN, Homo sapiens G3V8Y5_RAT, Rattus norvegicus A0A250Y753 CASCN, Castor Canadensis A0A1U7R4C7_MESAU, Mesocricetus auratus G7P5R6_MACFA, Macaca fascicularis A0A286XIQ9_CAVPO, Caviaporcellus I3M351_ICTTR, Ictidomys tridecemlineatus H2QPI8_PANTR, Pan troglodytes A0A1U7V0T5 TARSY, Tarsiussyrichta A0A1S2ZSL2 ERIEU, Erinaceus europaeus A0A0D9QYL1_CHLSB, Chlorocebussabaeus A0A2K5K5J5_COLAP, Colobus angolensis palliates A0A2I2ZIU3_GORGO, Gorilla gorillagorilla A0A1D5QGA5_MACMU, Macaca mulatta A0A2J8PEW7_PANTR, Pan troglodytes A0A2J8S2N1_PONAB, Pongo abelii A0A2K5CY83_AOTNA, Aotus nancymaae A0A2K5ZNR7_MANLE, Mandrillus leucophaeus C9J2Y9_HUMAN, Homo sapiens A0A096NEY4 PAPAN, Papio Anubis A0A1X7QYA1_9SACH, Kazachstania saulgeensis G8BY61 TETPH, Tetrapisisporaphaffii J7RV95_KAZNA, Kazachstaniana ganishii H2AVJ8_KAZAF, Kazachstania africana Q6FLD5_CANGA, Candida glabrata P08518 Yeast, Saccharomyces cerevisiae A0A0L8VHA5_9SACH, Saccharomyces boulardii A0A0L8RB33_SACEU, Saccharomyces eubayanus G0VJ71 NAUCC, Naumovozyma castellii G8ZM49| TORDC, Torulasporadel brueckii A0A1Q3A090_ZYGRO, Zygosaccharomyces rouxii A0A0N7IS35_9SACH, Zygosaccharomyces kombuchaensis A0A212MG88_ZYGBA, *Zygosaccharomyces bailii* A0A1S7HHE1_9SACH, *Zygosaccharomyces parabailii* S6ESB4_ZYGB2, Zygosaccharomyces bailii (strain CLIB 213) B6K5Q5_SCHJY, Schizosaccharomyces japonicas Q02061_SCHPO, Schizosaccharomyces pombe S9R8U4_SCHOY, Schizosaccharomyces octosporus S9W8C6 SCHCR, Schizosaccharomyces cryophilus

conservations strongly suggest that the DNA, SSU and MSU RNA polymerases use the same set of amino acids for template, substrate binding and catalysis establishing a structure-function relationship among the DNA and RNAPs. Interestingly the catalytic region harbouring the Zn binding motif is found very close to the Nterminal region. This is in sharp contrast to the equivalent region in the β' subunit in all eubacteria (Table 1). The Zn binding motif was originally identified by X-ray crystallographic analysis in the equivalent β ' subunit of the thermophilic bacterium, Thermus aquaticus [36]. The Zn binding 3 invariant Cs were located in the catalytic region and suggested in possible proofreading activity during elongation. A completely conserved R found upstream from the catalytic R is located at -6th position which is implicated in NTP selection. It is interesting to note a completely conserved R found upstream from the catalytic R is missing in eubacteria (Table 1) [29, 34]. The immediate downstream amino acid from catalytic K in DNA polymerases is usually a G or A [29], but in viral RNA polymerases it is a K or R, in MSU β ' subunits, it is a D and in all β ' subunits it is an S and it is an S/T in eukaryotes suggesting a possible role in NTP selection. In

SSU RNAPs, an invariant hydroxyl amino acid located very close to the YG pair is shown to involve in NTP selection by SDM experiments [37]. In addition to the template binding FG pair, there are 1 YG pair, 4 LG pairs and 3 I/VLG pairs in this subunit. A C- terminal conservation -SPDDSDEEN- (where the penultimate E is replaced with functionally equivalent D in some species) is seen in all higher forms of life and such sequence was conspicuously absent in the lower eukaryotes like yeasts and also not found in the initiation subunits, Rpb2. The -DXD- is a metal binding motif commonly found in glycosyl transferases and interestingly similar motif -**D**XDXT- in metal-dependent phosphatases (DxDXE in C-terminals of Rpb1 elongation subunits) where it plays a role for Ser phosphate removal from the CTD at the end of the cycle [38]. Interestingly, this motif invariably ends in N preceded by D/E in all the cases in Rpb1 (Fig 4). As expected the CTD is filled with the "heptapeptide repeats" (YSPTSPS) in all species and its role is discussed elsewhere (Not shown in the Figure) and results provide evidence for RNA exit in the vicinity of the carboxyl-terminal repeat domain, coupling synthesis to RNA processing by enzymes bound to this domain [28].

CLUSTAL O (1.2.4) MSA of the Rpb1, elongation subunits of MSU RNAPs- II 4

```
-MHGGAPSGDSACPLRTIKRVQFGILSPDEMKRMSVTEGGIKYPETT--EGGRPKLGGLM
tr|A0A1U8DYN0|A0A1U8DYN0 ALLSI
                                                                                                                      57
                                         --MHGA PSG DSAC PLRTIKR VOFGVI GPDE LKRMSVT EGG IKY SETT -- EGGRPKLGGLI
tr|A0A1L8H4P4|A0A1L8H4P4 XENLA
                                                                                                                      56
                                         -MHGGA PSGDSAC PLRTIKR VOFGIL SPDE MKRMSVTEGGIKY PETT -- EGGRPKLGGL
MHGGGP PSGDSAC PLRTIKR VOFGVLSPDE LKRMSVTEGGIKY PETT -- EGGRPKLGGL
tr|H9GLG5|H9GLG5_ANOCA
                                                                                                                      57
tr|H2R1J6|H2R1J6 PANTR
                                                                                                                      58
                                         MHGGGPPSGDSACPLRTIKRVQFGVLSPDELKRMSVTEGGIKYPETT--EGGRPKLGGL
tr|G1MCZ1|G1MCZ1_AILME
                                                                                                                      58
tr|008847|008847 MOUSE
                                         MHGGGPPSGDSACPLRTIKRVQFGVLSPDELKRMSVTEGGIKYPETT--EGGRPKLGGL
                                                                                                                      58
tr|S7PWZ6|S7PWZ6 MYOBR
                                         MHGGGPPSGDSACPLRTIKRVQFGVLSPDELKRMSVTEGGIKYPETT--EGGRPKLGGLM
tr|D4A5A6|D4A5A6_RAT
                                         MHGGGPPSGDSACPLRTIKRVQFGVLSPDELKRMSVTEGGIKYPETT--EGGRPKL
                                                                                                                      58
                                         MHGGGPPSGDSACPLRTIKRVQFGVLSPDELKRMSVTEGGIKYPETT--EGGRPKLGGL
sp|P08775|RPB1 MOUSE
                                                                                                                      58
                                         MHGGGPPSGDSACPLRTIKRVOFGVLSPDELKRMSVTEGGIKYPETT--EGGRPKLGGL
MHGGGPPSGDSACPLRTIKRVQFGVLSPDELKRMSVTEGGIKYPETT--EGGRPKLGGL
sp | P24928 | RPB1 HUMAN
                                                                                                                      58
tr|A0A1S3EWL2|A0A1S3EWL2 DIPOR
                                                                                                                      58
                                         MHGGGPPSGDSACPLRTIKRVQFGVLSPDELKRMSVTEGGIKYPETT--EGGRPKLGGL
                                                                                                                      58
sp|P11414|RPB1 CRIGR
                                         MHGGGPPSGDSACPLRTIKRVQFGVLSPDELKRMSVTEGGIKYPETT--EGGRPKLGGLM
tr10355591035559 CRIGR
                                                                                                                      58
tr|A0A2I3M9H2|A0A2I3M9H2 PAPAN
                                         MHGGGPPSGDSACPLRTIKRVQFGVLSPDELKRMSVTEGGIKYPETT--EGGRPKLGGL
                                                                                                                      58
tr|F7HB40|F7HB40 MACMU
                                         MHGGGPPSGDSACPLRTIKRVQFGVLSPDELKRMSVTEGGIKYPETT--EGGRPKL
                                                                                                         GGL
tr|A0A2K6RYW9|A0A2K6RYW9_SAIBB
                                         MHGGGPPSGDSACPLRTIKRVQFGVLSPDELKRMSVTEGGIKYPETT--EGGRPKLGGL
tr|W5N8Z6|W5N8Z6 LEPOC
                                         --mhgppsgdsacplrlikrvqfgiispdelkrmsvteggikypett--eggrpklggl
                                                                                                                      56
                                         --MHGPPSGDSACPLRTIKRVOFGILSPDELKRMSVTEGGIKYPETT--EGGRPKLGGLM
tr|I3JRW6|I3JRW6 ORENI
                                                                                                                      56
                                         --MHGPPSSDSACPLRLIKRVOFGVLSPDELKRMSVTEGGIKYPETT--EGGRPKLGGLM
tr|A0A0R4IMS9|A0A0R4IMS9_DANRE
                                                                                                                      56
                                         --MHGPPSGDSACPLRTIKRVQFGILSPDELKRMSVTEGGIKYPETT--EGGRPKLGGL
tr|A0A1A7X327|A0A1A7X327 9TELE
                                                                                                                      56
                                         --MHGPPSGDSACPLRTIKRVQFGIISPDELKRMSVTEGGIKYPETT--EGGRPKLGGL
tr|A0A1A8UKD7|A0A1A8UKD7_NOTFU
                                                                                                                      56
                                         --MHGPPSGDSACPLRTIKRVQFGILSPDELKRMSVTEGGIKYPETT--EGGRPKLGGLM
tr|A0A1A8ER05|A0A1A8ER05 9TELE
                                                                                                                      56
tr|A0A1A8DQ60|A0A1A8DQ60_9TELE
                                         --MHGPPSGDSACPLRTIKRVQFGILSPDELKRMSVTEGGIKYPETT--EGGRPKLGGLM
                                         --MHGP PSGDSAC PLRTIKR VOFGIL SPDE LKRMSVT EGGIKY PETT -- EGGRPKL GGL
--MHGP PSGDSAC PLRLIKR VOFGIL SPDE LKRMSVT EGGIKY PETT -- EGGRPKL GGL
tr|A0A1A8NSR8|A0A1A8NSR8_9TELE
tr|A0A1W4YLM7|A0A1W4YLM7_9TELE
                                                                                                                      56
tr|A0A1M8A6L7|A0A1M8A6L7 MALS4
                                         MMGH--QFAPSVAPVRPVKEVQFGILSPEFIRALSVCK--IEFPEVKDDATGKYKVGGLS
                                                                                                                      56
                                         --MS--QFPYSSAPLRSVKE VQFQLLSPEE IRAISVVK--IEYPEIMDESRQRPREGGL
--MS--QFPYSSAPLRSVKE VQFGLLSPEE IRAISVVK--IEYPEIMDESRQRPREGGL
tr|A0A1B2J8C6|A0A1B2J8C6 PICPA
                                                                                                                      54
tr|F2QW17|F2QW17 KOMPC
                                                                                                                      54
tr|A3GID7|A3GID7 PICST
                                         -MSR--QFPYSSAPLRSVKEVQFGLLSPEEVRAISVAK--IEYPETMDQTTKTPREGGLI
                                                                                                                      55
tr|A0A1D8PUA6|A0A1D8PUA6 CANAL
                                         -msr--qfpyssaplrsvkevqfgllspeevraisvak--ieypetmdqttkrpreggli
                                                                                                                      55
                                         -MSR--TFPFSNAPLRSVKEVQFGLLSPEEVRAISVAK--IEYPETMDQATKRPREGGLM
tr|G8BEH9|G8BEH9 CANPC
                                                    * .*:* :*. ****:: .*:* : ** : *.: *
```

```
DPROGVIERIGR
DPROGVIERIGR
DPROGVIERIGR
DPROGVIERIGR
                                                                    OTCA-GNMTE
OTCA-GNMTE
OTCA-GNMTE
OTCA-GNMTE
                                                                                           <mark>sh</mark>ielakpve<mark>hv</mark>¢fl¢ktmkilr<mark>cv</mark>q<mark>f</mark>fc<mark>$</mark>kl
                                                                                                           GFL GKTMKLLRCVC FFC
tr|A0A1U8DYN0|A0A1U8DYN0 ALLSI
                                                                                    PGH
tr|A0A1L8H4P4|A0A1L8H4P4 XENLA
                                                                                             VELAKPVY
                                                                                    PGHI
                                                                                                                                                 115
                                                                                    PGH
tr|H9GLG5|H9GLG5_ANOCA
                                                                                           HIELAKPVE
                                                                                                                                   SKT.
                                                                                                                                                 116
                                                                                            H IELAKPVF<mark>HV</mark>GFLVKTMKVLR<mark>CVCFF</mark>
H IELAKPVFH<mark>V</mark>GFLVKTMKVLRCVCFF
H IELAKPVFHVGFLVKTMKVLRCVCFF
tr|H2R1J6|H2R1J6 PANTR
                                                                                     PGHP
                                                                                                                                                 117
                                                  DPROGVIERTGR
DPROGVIERTGR
tr|G1MCZ1|G1MCZ1 AILME
                                                                         A-GNMT
                                                                                     PGHI
                                                                                                                                    SKL.
                                                                                                                                                 117
tr10088471008847 MOUSE
                                                                         A-GNMT
                                                                                     PGH
                                                                                                                                   SKT.
                                                                                                                                                 117
                                                                                           H IELAKPVFHVGFLVKIMKVLRCVGFF
H IELAKPVFHVGFLVKIMKVLRCVGFF
H IELAKPVFHVGFLVKIMKVLRCVGFF
                                                  DPROGVIERIGRO
DPROGVIERIGRO
DPROGVIERIGRO
DPROGVIERIGRO
DPROGVIERIGRO
DPROGVIERIGRO
DPROGVIERIGRO
DPROGVIERIGRO
DPROGVIERIGRO
DPROGVIERIGRO
DPROGVIERIGRO
DPROGVIERIGRO
DPROGVIERIGRO
DPROGVIERIGRO
DPROGVIERIGRO
DPROGVIERIGRO
DPROGVIERIGRO
DPROGVIERIGRO
DPROGVIERIGRO
DPROGVIERIGRO
DPROGVIERIGRO
DPROGVIERIGRO
DPROGVIERIGRO
DPROGVIERIGRO
DPROGVIERIGRO
DPROGVIERIGRO
DPROGVIERIGRO
DPROGVIERIGRO
tr|S7PWZ6|S7PWZ6_MYOBR
                                                                         A-GNMT
                                                                                    PGHF
                                                                                                                                                 117
                                                                         A-GNMT
tr|D4A5A6|D4A5A6 RAT
                                                                                     PGHI
                                                                                                                                                 117
                                                                                                           GFLVKTMKVLR<mark>CV</mark>CFF
                                                                         A-GNMT
                                                                                     PGH
                                                                                           GH IELAKPVE
sp|P08775|RPB1 MOUSE
                                                                                                                                   SKL
                                                                                                                                                 117
                                                                                                          VGFLVKTMKVLRCVCFF
VGFLVKTMKVLRCVCFF
sp|P24928|RPB1 HUMAN
                                                                         A-GNMT
                                                                                    PGH
                                                                                            HIELAKPVF
                                                                                                                                                 117
tr|A0A1S3EWL2|A0A1S3EWL2 DIPOR
                                                                         A-GNMT
                                                                                     PGHI
                                                                                            IELAKPVF
                                                                                                                                    SKL.
                                                                                                                                                 117
                                                                    CQTCA-GNMTE
                                                                                           <mark>sh</mark>ielakpyf<mark>hy</mark>gflyktmkylr<mark>cy</mark>d<mark>ffcs</mark>kl
sp|P11414|RPB1 CRIGR
                                                                                    PGHI
                                                                                                                                                 117
                                                                                           <mark>GH</mark> IELAKPV F<mark>HV</mark> GFLYKTMKVLRC<mark>V</mark> CEF
GH IELAKPV FHV GFLYKTMKVLRCV CEF
                                                                                                                                   SKT.
tr|035559|035559 CRIGR
                                                                                    PGH
                                                                                                                                                 117
                                                                    COTCA-GNMTE
COTCA-GNMTE
COTCA-GNMTE
tr|A0A2I3M9H2|A0A2I3M9H2 PAPAN
                                                                                     PGHI
                                                                                                                                                 117
                                                                                            HIELAKPVF<mark>HV</mark>GFLVKTMKVLR<mark>CV</mark>CFFCSKL
tr|F7HB40|F7HB40 MACMU
                                                                                    PGH
                                                                                                                                                 117
                                                                                            H IELAKPV FHVGFL VKIMKVLRCVCEF
E IELAKPV FHVGFL VKIMKVLRCVCEF
                                                                                    PGH
tr|A0A2K6RYW9|A0A2K6RYW9 SAIBB
                                                                                                                                   SKL
                                                                                                                                                 117
                                                                    COTCAAGNMTE
COTCAAGNMTE
COTCA-GNMTE
COTCA-GNMTE
COTCA-GNMTE
COTCA-GNMTE
tr|W5N8Z6|W5N8Z6_LEPOC
                                                                                     PGHI
                                                                                                                                   SKT.
                                                                                                                                                 116
                                                                                           <mark>gh</mark>ielakpvf<mark>hv</mark>¢fi$kimkvlr<mark>cvcf</mark>fc$kl
tr|I3JRW6|I3JRW6 ORENI
                                                                                    PGHF
                                                                                                                                                 115
                                                                                           H IELAKPVFHVGFI TKIMKVLRCVCFF
H IELAKPVFHVGFI SKIMKILRCVCFF
tr|A0A0R4IMS9|A0A0R4IMS9_DANRE
tr|A0A1A7X327|A0A1A7X327_9TELE
                                                                                    PGH
                                                                                                                                   SKT.
                                                                                                                                                 115
                                                                                    PGH
                                                                                                                                                 115
                                                                                    GGHTGH IELAKPYFTVGFISKIMKILRCVGFESKL
CPGHTGH IELAKPYFTVGFISKIMKILRCVGFFCSKL
CPGHTGH IELAKPYFTVGFISKIMKILRCVGFFCSKL
CPGHTGH IELAKPYFTVGFISKIMKILRCVGFFCSKL
CPGHTGH IELAKPYFTVGFISKIMKILRCVGFFCSKL
CPGHTGH IELAKPYFTVGFIKKIMKILRCVGFFCSKL
CPGHTGH IDLARPYFTVGFLGKVKKLLECVGVHCGKL
tr|A0A1A8UKD7|A0A1A8UKD7 NOTFU
                                                                                                                                                 115
                                                                     QT(
tr|A0A1A8ER05|A0A1A8ER05_9TELE
                                                                         A-GNMT
                                                                                                                                                 115
tr|A0A1A8DQ60|A0A1A8DQ60_9TELE
                                                                         A-GNMT
                                                                                                                                                 115
tr|A0A1A8NSR8|A0A1A8NSR8_9TELE
                                                                         A-GNMT
                                                                                                                                                 115
                                                               RTGR<mark>C</mark>QTCA-GNMTE
RNYKCQTOGE-GQAE
tr|A0A1W4YLM7|A0A1W4YLM7 9TELE
                                                   DPROGVIER
                                                                                                                                                 115
tr|A0A1M8A6L7|A0A1M8A6L7 MALS4
                                                   DPRLGTID
                                                                                                                                                 115
                                                  PGH<mark>TGH</mark>MELAKPVF<mark>HIG</mark>FIPKIKKVCE<mark>CIOMNCG</mark>KL
PGHTGHMELAKPVFHIGFIPKIKKVCE<mark>CIOMNCG</mark>KL
tr|A0A1B2J8C6|A0A1B2J8C6 PICPA
                                                                                                                                                 113
tr|F2QW17|F2QW17 KOMPC
                                                                                                                                                 113
                                                                                    CPGH<mark>TGH</mark> IELAKPVF<mark>HI</mark> GFIAKIKKVCE C<mark>V</mark>CMHCGKL
CPGHTGH IELAKPVFHIGFIAKIKKACE SVCMHCGKT.
tr|A3GID7|A3GID7 PICST
                                                                                                                                                 114
                                                                                    CPGH<mark>FGH</mark> IELAKPVF<mark>HI</mark> GFI AKIKKACE S<mark>VOMHCG</mark>KL
CPGHFGH IELAKPVFHI GFI AKIKKVCE SIOMHCGKL
tr|A0A1D8PUA6|A0A1D8PUA6_CANAL
                                                                                                                                                 114
tr|G8BEH9|G8BEH9 CANPC
                                                                                                                                                 114
                                                                                            *::**:**:
tr|A0A1U8DYN0|A0A1U8DYN0_ALLSI
                                                  LVDSNNPKIKD-ILGKSKGQPKKRLTHVYDL<mark>CH</mark>GKNICHGGEEMDNKFGVEQTEGDEDLT
                                                                                                                                                 175
tr|A0A1L8H4P4|A0A1L8H4P4 XENLA
                                                   LVDANNPKIKD-ILIKSKGQPKKRLTHVYEL
                                                                                                 GKNI
                                                                                                         GGEEMDNKFGVEQTEGDEDIQ
                                                                                                                                                 174
                                                                                                        EGGEEMDNK FGVEQPEGDEDIT
tr|H9GLG5|H9GLG5 ANOCA
                                                  LVDSNN PKT KD-T LAKSKGO PKKRLT HVYD L
                                                                                                 GKNI
                                                                                                                                                 175
tr|H2R1J6|H2R1J6 PANTR
                                                   LVDSNNPKIKD-ILAKSKGQPKKRLTHVYDL
                                                                                                 GKNI <mark>CE</mark>GGEEMDNKFGVEQPEGDEDLT
                                                                                                                                                 176
tr|G1MCZ1|G1MCZ1 AILME
                                                   LVDSNNPKIKD-ILAKSKGQPKKRLTHVYDL
                                                                                                 GKNT
                                                                                                         GGEEMDNKFGVEQPEGDEDLT
                                                                                                                                                 176
                                                                                                         EGGEEMDNKFGVEQPEGDEDLT
tr|008847|008847 MOUSE
                                                  LVDSNNPKIKD-ILAKSKGOPKKRLTHVYDL
                                                                                                KGKNI (
                                                                                                                                                 176
                                                                                                KGKNI <mark>CE</mark>GGEEMDNK FGVEQPEGDEDLT
tr|S7PWZ6|S7PWZ6 MYOBR
                                                  LVDSNN PKI KD-I LAKSKGQ PKKRLTHVYDL
                                                                                                                                                 176
triD4A5A6iD4A5A6 RAT
                                                  I.VD SNN PKT KD-T I.AK SKGO PKKRT.THVYD I.
                                                                                                 GKNT
                                                                                                         GGEEMDNK FGVEOPEGDEDI.T
                                                                                                                                                 176
sp|P08775|RPB1 MOUSE
                                                  LVDSNNPKIKD-ILAKSKGQPKKRLTHVYDL
                                                                                                 GKNI
                                                                                                         EGGEEMDNKFGVEQPEGDEDLT
                                                                                                                                                 176
sp|P24928|RPB1 HUMAN
                                                  LVDSNN PKI KD-I LAKSKGQ PKKRLT HVYDL
                                                                                                         EGGEEMDNKFGVEQPEGDEDLT
                                                                                                                                                 176
tr|A0A1S3EWL2|A0A1S3EWL2_DIPOR
                                                                                                GKNI
                                                                                                        EGGEEMDNKFGVEOPEGDEDLT
                                                  LVDSNNPKIKD-ILAKSKGQPKKRLTHVYDL
                                                                                                                                                 176
                                                   LVDSNNPKIKD-ILAKSKGQPKKRLTHVYDL
                                                                                                 GKNI
                                                                                                         GGEEMDNKFGVEQPEGDEDLT
sp|P11414|RPB1 CRIGR
                                                                                                                                                 176
tr|035559|035559 CRIGR
                                                  LVDSNN PKI KD-I LAKSKGO PKKRLT HVYDL
                                                                                                        EGGEEMDNKFGVEOPEGDEDLT
                                                                                                GKNI
                                                                                                                                                 176
tr|A0A2I3M9H2|A0A2I3M9H2 PAPAN
                                                  LVDSNNPKIKD-ILSKSKGQPKKRLTHVYDL
                                                                                                KGKNI (
                                                                                                        EGGEEMDNKFGVEQPEGDEDLT
                                                                                                                                                 176
tr|F7HB40|F7HB40 MACMU
                                                  LVDSNN PKI KD-ILSKSKGQ PKKRLTHVYDL
                                                                                                 GKNI
                                                                                                         GGEEMDNKFGVEQPEGDEDLT
                                                                                                                                                 176
                                                                                                         EGGEEMDNKFGVEQPEGDEDLT
tr|A0A2K6RYW9|A0A2K6RYW9_SAIBB
                                                  LVDSNNPKIKD-ILAKSKGOPKKRLTHVYDL
                                                                                                KGKNT (
                                                                                                                                                 176
tr|W5N8Z6|W5N8Z6 LEPOC
                                                   LVDSNNPKIKD-ILGKSKGQPRKRLTHVYDL
                                                                                                 GKNI
                                                                                                         EGGEEMDNKFGVEONETEEDIT
                                                                                                                                                 175
tr|I3JRW6|I3JRW6 ORENI
                                                   LVDSNN PKI KD-I LAKSKGQ PRKRLTHVYEL
                                                                                                 GKNI
                                                                                                         EGGEEMDNKFGMEQQETEEDLT
                                                                                                                                                 174
tr|A0A0R4IMS9|A0A0R4IMS9 DANRE
                                                   LVDANNPKIKD-ILTKSKGQPRKRLTHVYDL
                                                                                                 GKNI
                                                                                                         EGGEEMDNKFGVEQPES-EDIT
                                                                                                                                                 173
tr|A0A1A7X327|A0A1A7X327 9TELE
                                                  LVDANN PKI KE-I LVKSKGQ PRKRLTHVYEL
                                                                                                KGKNI<mark>CE</mark>GGEEMDNKFGMEQQETEEDIT
                                                                                                                                                 174
                                                   LVDANNPKIKE-ILVKSKGQPRKRLTHVYEL
tr|A0A1A8UKD7|A0A1A8UKD7_NOTFU
                                                                                                 GKNT
                                                                                                         EGGEEMDNKFGMEOOETEEDIT
                                                                                                                                                 174
tr|A0A1A8ER05|A0A1A8ER05 9TELE
                                                   LVDANNPKIKE-ILVKSKGQPRKRLTHVYEL
                                                                                                 GKNI
                                                                                                         EGGEEMDNKFGMEQQETEEDIT
                                                                                                                                                 174
tr|A0A1A8DQ60|A0A1A8DQ60 9TELE
                                                   LVDANNPKIKE-ILVKSKGQPRKRLTHVYEL
                                                                                                GKNT
                                                                                                         EGGEEMDNKFGMEQQETEEDIT
                                                                                                                                                 174
                                                                                                KGKNI
tr|A0A1A8NSR8|A0A1A8NSR8 9TELE
                                                   LVDANNPKIKE-ILVKSKGQPRKRLTHVYEL
                                                                                                        EGGEEMDNKFGMEOOETEEDIT
                                                                                                                                                 174
tr|A0A1W4YLM7|A0A1W4YLM7_9TELE
tr|A0A1M8A6L7|A0A1M8A6L7_MALS4
                                                   LVDSNNPKIKD-ILAKSKGQPRKRLTHVYDL
                                                                                                 GKNI
                                                                                                         EGGEEMDNKFGMEQAETEEDIT
                                                                                                                                                 174
                                                                                                KITI CHADEGKDE----DELGDTTQ
                                                   KADPISDPVFKSLLQSTRANRKRRFQRVWEY
                                                                                                                                                 169
tr|A0A1B2J8C6|A0A1B2J8C6_PICPA
                                                   LLDETNPTMAQAI --- RIRDPKKRFNAVWQLC
                                                                                                TKMV
                                                                                                        EADAPVDE-----YSEQ
                                                                                                                                                 160
                                                   LLDETNPTMAQAI --- RIRDPKKRFNAVWQLCKTKMVCEADAPVDE-----YSEO
tr|F2QW17|F2QW17 KOMPC
                                                                                                                                                 160
                                                   LLDENN PAMAQAI --- KIRD PKKRFNAVWQLCHAKMVCET DII EEG-----ATE-
tr|A3GID7|A3GID7 PICST
                                                                                                                                                 160
                                                  LLDETN PAMAQAI ---KIRD PKKRFNAVWS LCHTKMV CET DNNEDE-----MTD-
tr|A0A1D8PUA6|A0A1D8PUA6 CANAL
                                                                                                                                                 160
                                                  LLDESNPLMAQAI---KIRDPKKRFNAVWSICKSKMVCRTATSEEE-------MND-
* . : . : :: *: ::
tr|G8BEH9|G8BEH9 CANPC
                                                                                                                                                 160
```

```
KEKGHGGCCRYQPRIRRVGLELYAEWKHVNE-----DSQEKKILLSPERVHEIFKRIA
tr|A0A1U8DYN0|A0A1U8DYN0 ALLSI
                                                                                                                                       228
                                                          GRYQPRIRRTGLELYAEWKHVNE------DSQEKKILLSPERVHEIFKRIT
tr|A0A1L8H4P4|A0A1L8H4P4 XENLA
                                               KEKGHGG
                                                                                                                                       227
                                               KEKGHGGCGRYOPRIRRSGLELYAEWKHVNE-----DSOEKKILLSPERVHEIFKRIS
tr|H9GLG5|H9GLG5 ANOCA
                                                                                                                                       228
                                               KEKGHGGCCRYOPRIRRSGLELYAEWKHVNE------DSQEKKILLSPERVHEIFKRIS
tr|H2R1J6|H2R1J6_PANTR
                                                                                                                                       229
                                               KEKGHGGCGRYQPRIRRSGLELYAEWKHVNE------DSQEKKILLSPERVHEIFKRIS
tr|G1MCZ1|G1MCZ1 AILME
                                                                                                                                       229
tr|008847|008847 MOUSE
                                               KEKGHGGCGRYQPRIRRSGLELYAEWKHVNE-----DSQEKKILLSPERVHEIFKRIS
                                                                                                                                       229
                                               KEKGHGGGGRYOPRIRRSGLELYAEWKHVNE-----DSOEKKILLSPERVHE IFKRIS
KEKGHGGGGRYOPRIRRSGLELYAEWKHVNE-----DSOEKKILLSPERVHE IFKRIS
tr|S7PWZ6|S7PWZ6 MYOBR
                                                                                                                                       229
tr|D4A5A6|D4A5A6_RAT
                                                                                                                                       229
sp|P08775|RPB1_MOUSE
                                               KEKGHGGCGRYQPRIRRSGLELYAEWKHVNE------DSQEKKILLSPERVHEIFKRIS
                                                                                                                                       229
                                               KEKGIGGCGRYOPRIRRSGLELYAEWKHVNE-----DSQEKKILLSPERVHEIFKRIS
KEKGIGGCGRYOPRIRRSGLELYAEWKHVNE-----DSQEKKILLSPERVHEIFKRIS
sp|P24928|RPB1 HUMAN
                                                                                                                                       229
tr|A0A1S3EWL2|A0A1S3EWL2 DIPOR
                                                                                                                                       229
                                               KEKGHGGCCRYQPRIRRSGLELYAEWKHVNE------DSQEKKILLSPERVHEIFKRIS
sp|P11414|RPB1 CRIGR
                                                                                                                                       229
tr|035559|035559 CRIGR
                                               KEKGHGGCGRYQPRIRRSGLELYAEWKHVNE------DSQEKKILLSPERVHEIFKRIS
                                                                                                                                       229
                                               KEKGHGGCGRYOPRIRRSGLELYAEWKHVNE------DSOEKKILLSPERVHEIFKRIS
tr|A0A2I3M9H2|A0A2I3M9H2 PAPAN
                                                                                                                                       229
                                               KEKGHGGCGRYOPRIRRSGLELYAEWKHVNE------DSQEKKILLSPERVHEIFKRIS
tr|F7HB40|F7HB40 MACMU
                                                                                                                                       229
                                               KEKGIGGCCRYOPRIRRSGLELYAEWKHVNE-----DSOEKKILLSPERVHE IFKRIS
KEKGIGGCCRYOPRIRRSGLELYAEWKHVNE-----DSOEKKILLSPERVHE IFKRIS
tr|A0A2K6RYW9|A0A2K6RYW9 SAIBB
                                                                                                                                       229
tr|W5N8Z6|W5N8Z6 LEPOC
                                                                                                                                       228
                                               KEKGIGGCGRYOPRIKRSGLELYAEWKHVNE-----DSQEKKILLSPERVHE IFKRIS
KEKGIGGCGRYOPRIRRSGLELYAEWKHVNE-----DSQEKKILLSPERVHE IFKRIA
tr|I3JRW6|I3JRW6 ORENI
                                                                                                                                       227
tr|A0A0R4IMS9|A0A0R4IMS9_DANRE
                                                                                                                                       226
                                               KEKGHGG<mark>CG</mark>RYQPRIRRSGLELYAEWKHVNE------DSQEKKILLSPERVHEIFKRIS
tr|A0A1A7X327|A0A1A7X327 9TELE
                                                                                                                                       227
                                               KEKGHGG<mark>CG</mark>RYQPRIRRSGLELYAEWKHVNE------DSQEKKILLSPERVHEIFKRIS
tr|A0A1A8UKD7|A0A1A8UKD7 NOTFU
                                                                                                                                       227
                                               KEKGHGGCGRYOPRIRRSGLELYAEWKHVNE------DSOEKKILLSPERVHEIFKRIS
tr|A0A1A8ER05|A0A1A8ER05 9TELE
                                                                                                                                       227
tr|A0A1A8DQ60|A0A1A8DQ60_9TELE
                                               KEKGHGGCGRYOPRIRRSGLELYAEWKHVNE------DSQEKKILLSPERVHEIFKRIS
                                                                                                                                       227
tr|A0A1A8NSR8|A0A1A8NSR8 9TELE
                                               KEKGHGGCGRYQPRIRRSGLELYAEWKHVNE-----DSQEKKILLSPERVHEIFKRIS
                                                                                                                                       227
                                               KEKGHGGGGRYOPLIRRSGLELYAEWKHYNE-----DSOEKKILLSPERVHEIFKRIS
OKIGGGGGRFOPAIRKEALKLFSVWKQSKDEDEDSGGMAQSEKRPLPASEVHTILKKIT
tr|A0A1W4YLM7|A0A1W4YLM7_9TELE
                                                                                                                                       227
tr|A0A1M8A6L7|A0A1M8A6L7 MALS4
                                                                                                                                       229
tr|A0A1B2J8C6|A0A1B2J8C6 PICPA
                                               KVVSRGG<mark>CG</mark>NTQPVVRKDGMKLWGTWKKSGFSDR----DAQPERKLLTPGEILNVFRHIS
                                                                                                                                       216
                                               KVVSRGGCONTOPVVRKDGMKLWGTWKKSGFSDR----DAOPERKLLTPGEILNVFKHIS
-TTTRGGCONTOPTIRRDGLKLWGTWRHNKNFEE----NEQPERRLLTPSEILNVLKHIS
tr|F2QW17|F2QW17_KOMPC
tr|A3GID7|A3GID7_PICST
                                                                                                                                       216
                                                                                                                                       215
tr|A0A1D8PUA6|A0A1D8PUA6 CANAL
                                               -QPSkGGCCHPQPTIRRDGLKLWGTWKQNKNYDD----NDQPERRLLTPSEILNVFKHIS
                                                                                                                                       215
                                                -HNVRGG<mark>CG</mark>HTQPTIRRDGLKLWGTWKHNKNFEE----NDQPERRLLTPSEILNVFKHIS
tr|G8BEH9|G8BEH9 CANPC
                                                                                                                                       215
                                                           . ** ::: .::* . *::
                                                                                                   . .
                                               DDECLV<mark>LG</mark>NDPKFARPEWMVCTVLPVPHLSVRPAVVMQGSARNQDDLTHKLADIVKINNQ
tr|A0A1U8DYN0|A0A1U8DYN0 ALLSI
                                                                                                                                        288
tr|A0A1L8H4P4|A0A1L8H4P4 XENLA
                                               DEECFI<mark>LG</mark>MDPRFARPEWLIITVLPVPHLCVRPAVVMQGSARNQDDLTHKLADIVKINNQ
                                                                                                                                        287
                                               DEECFILGMDPMFAREEWMICTVLPVPELSVRPAVVMQGSARNODDLTHKIADIVKINNO
DEECFVLGMEPRYARPEWMIVTVLPVPELSVRPAVVMQGSARNODDLTHKIADIVKINNO
DEECFVLGMEPRYARPEWMIVTVLPVPELSVRPAVVMQGSARNODDLTHKIADIVKINNO
tr|H9GLG5|H9GLG5 ANOCA
                                                                                                                                        288
tr|H2R1J6|H2R1J6 PANTR
                                                                                                                                        289
tr|G1MCZ1|G1MCZ1_AILME
                                                                                                                                        289
                                               DEECTVLGMEPRYARPEMMITTVLEVPHLSVRPAVVMQGSARNODDLTHKLADIVKINNQ
DEECTVLGMEPRYARPEMMITTVLEVPHLSVRPAVVMQGSARNODDLTHKLADIVKINNQ
DEECTVLGMEPRYARPEMMITTVLPVPHLSVRPAVVMQGSARNODDLTHKLADIVKINNQ
DEECTVLGMEPRYARPEMMITTVLPVPHLSVRPAVVMQGSARNODDLTHKLADIVKINNQ
tr|008847|008847 MOUSE
                                                                                                                                        289
tr|S7PWZ6|S7PWZ6 MYOBR
                                                                                                                                        289
tr|D4A5A6|D4A5A6 RAT
                                                                                                                                        289
sp|P08775|RPB1 MOUSE
                                                                                                                                        289
                                               DEECFV<mark>LG</mark>MEPRYARPEWMIVTVLPVPHLSVRPAVVMQGSARNQDDLTHKLADIVKINNQ
sp|P24928|RPB1 HUMAN
                                                                                                                                        289
                                               DEECFVLGMEPRYARPEWMIVTVLPVPELSVRPAVVMQGSARNODDLTHKLADIVKINNO
DEECFVLGMEPRYARPEWMIVTVLPVPELSVRPAVVMQGSARNODDLTHKLADIVKINNO
DEECFVLGMEPRYARPEWMIVTVLPVPELSVRPAVVMQGSARNODDLTHKLADIVKINNO
tr|A0A1S3EWL2|A0A1S3EWL2 DIPOR
                                                                                                                                        289
sp|P11414|RPB1 CRIGR
                                                                                                                                        289
tr|035559|035559 CRIGR
                                                                                                                                        289
                                               DEECTVICHEPRYARPEWMIVTVLPVPELSVRPAVVMQGSARNODDLTHKLADIVKINNO
DEECTVICHEPRYARPEWMIVTVLPVPELSVRPAVVMQGSARNODDLTHKLADIVKINNO
tr|A0A2I3M9H2|A0A2I3M9H2 PAPAN
                                                                                                                                        289
tr|F7HB40|F7HB40 MACMU
                                                                                                                                        289
                                               DEECTVLGHEPRVARPEWMIVTVLPVPHLSVRPAVVMQGSARNODDLTHKIADIVKINNO
DEEDIILGHDPKFARPEWMIVTVLPVPHLAVRPAVVMQGSARNODDLTHKIADIVKINNO
tr|A0A2K6RYW9|A0A2K6RYW9 SAIBB
                                                                                                                                        289
tr|W5N8Z6|W5N8Z6 LEPOC
                                                                                                                                        288
                                               DEEDIILGADPKFARPEWMIVTVLPVPHLAVRPAVVMQGSARNODDLTHKLADIVKINNO
DEEDIILGADPKFARPEWMIVTVLPVPHLAVRPAVVMQGSARNODDLTHKLADIVKINNO
DEEDIILGADPKFARPEWMIVTVLPVPHLAVRPAVVMQGSARNODDLTHKLADIVKINNO
tr|I3JRW6|I3JRW6 ORENI
                                                                                                                                        287
tr|A0A0R4IMS9|A0A0R4IMS9_DANRE
                                                                                                                                        286
tr|A0A1A7X327|A0A1A7X327 9TELE
                                                                                                                                        287
                                               DEE DII <mark>LG</mark>MDPKFARPEWMIVTVLPVPHLAVRPAVVMQGSARN QDDLTHKLADIVKINNQ
tr|A0A1A8UKD7|A0A1A8UKD7_NOTFU
                                                                                                                                        287
                                                        gudpkfarpewmivtvlpvpelavrpavvmogsarnoddlthkladivkinno
gudpkfarpewmivtvlpvpelavrpavvmogsarnoddlthkladivkinno
tr|A0A1A8ER05|A0A1A8ER05_9TELE
tr|A0A1A8DQ60|A0A1A8DQ60_9TELE
                                               DEEDII
                                                                                                                                        287
                                               DEEDII
                                                                                                                                        287
tr|A0A1A8NSR8|A0A1A8NSR8_9TELE
tr|A0A1W4YLM7|A0A1W4YLM7_9TELE
                                               DEEDIILGADPKFARPEWMIVTVLPVPELAVRPAVVMQGSARNODDLITKLADIVKINNO
DEEDIILGADPKFARPEWMIVTVLPVPELAVRPAVVMQGSARNODDLITKLADIVKINNO
                                                                                                                                        287
                                                                                                                                        287
                                               PEDVVTLGLSEDFAQPDWMVLTVLPVPHPQVRPGVTEFGSGMGQDDLTYKLADIIKASAN
tr|A0A1M8A6L7|A0A1M8A6L7_MALS4
                                                                                                                                        289
tr|A0A1B2J8C6|A0A1B2J8C6 PICPA
                                               PEDCFR<mark>LGFNEDYARPEWMILTVLPVPHPQVRPSIAMDETTQGQDDLTHKLBDILKANI</mark>N
                                                                                                                                        276
tr|F2QW17|F2QW17_KOMPC
                                               PEDCFR<mark>LGFNEDYARPEWMILTVLPVPHPQVRPSIAMDETTQGQDDLTHKLBDILKANIN</mark>
                                                                                                                                        276
                                               tr|A3GID7|A3GID7 PICST
                                                                                                                                        275
tr|A0A1D8PUA6|A0A1D8PUA6_CANAL
                                                                                                                                        275
tr|G8BEH9|G8BEH9 CANPC
                                                                                                                                        275
```

```
LRRNEQNGAAAHVIAEDVELLQFHVATMVDNELPGLPRAMQKSGRPLKSLKORLKGKEGR
tr|A0A1U8DYN0|A0A1U8DYN0_ALLSI
                                                                                                                                  348
tr|A0A1L8H4P4|A0A1L8H4P4 XENLA
                                             LRRNEQNGAAAHVIAE DVKLLQFHVA TMVDNEI PGLPRAMQK$GRPLKSVKDKLKGKEGR
                                                                                                                                  347
tr|H9GLG5|H9GLG5 ANOCA
                                             LRRNEQNGAAAHVIAE DVKLLQFHVATMVDNEL PGLPRAMQKSGRPLKSLKDRLKGKEGR
                                                                                                                                  348
                                             LRRNEQNGAAAHVIAE DVILLQFHVA TMVDNEL PGLPRAMQKSGRPLKSLKDRLKGKEGR
tr|H2R1J6|H2R1J6 PANTR
                                                                                                                                  349
tr|G1MCZ1|G1MCZ1_AILME
                                             LRRNEQNGAAAHVIAEDVKLLQFHVATMVDNELPGLPRAMQK$GRPLKSLKORLKGKEGR
                                                                                                                                  349
tr|008847|008847 MOUSE
                                             LRRNEQNGAAAHVIAE DVKLLQFHVA TMVDNEL PGLPRAMQK$GRPLKSLKDKLKGKEGR
                                                                                                                                  349
tr|S7PWZ6|S7PWZ6_MYOBR
                                             LRRNEQNGAAAHVIAE DVILLQFHVA TMVDNEL PGLPRAMQKSGRPLKSLKDRLKGKEGR
                                                                                                                                  349
tr|D4A5A6|D4A5A6 RAT
                                             LRRNEQNGAAAHVIAE DVKLLQFHVA TMVDNEL PGLPRAMQK$GRPLKSLKDRLKGKEGR
                                                                                                                                  349
sp|P08775|RPB1 MOUSE
                                             LRRNEQNGAAAHVIAE DVKLLQFHVA TMVDNEL PGLPRAMQKSGRPLKSLKDRLKGKEGR
                                                                                                                                  349
sp | P2 4928 | RPB1_HUMAN
                                             lrrneqngaaahviaedvkllqfhvatmvdnelpglpramqksgrplkslkprlkgkegr
                                                                                                                                  349
tr|A0A1S3EWL2|A0A1S3EWL2 DIPOR
                                             LRRNEQNGA AAHV I AE DVIKLLQFHVA TMVDNEL PGLPRAMQK SGRPLKSLKDRLKGKEGR
                                                                                                                                  349
sp|P11414|RPB1 CRIGR
                                             LRRNEQNGAAAHVIAE DVKLLQFHVATMVDNEL PGLPRAMQKSGRPLKSLKQKLKGKEGR
                                                                                                                                  349
                                             lrrneongaaahviae dvillofhva tmydnel pglp ramoks grplkslidtlikgkegr
tr|035559|035559 CRIGR
                                                                                                                                  349
                                             LRRNEQNGA AAHV IAE DVILLQFHVA TMVDNEL PGLPRAMQKSGRPLKSLKORLKGKEGR
tr|A0A2I3M9H2|A0A2I3M9H2_PAPAN
                                                                                                                                  349
tr|F7HB40|F7HB40 MACMU
                                             LRRNEQNGAAAHVIAE DVKLLQFHVATMVDNEL PGLPRAMQKSGRPLKSLKDRLKGKEGR
                                                                                                                                  349
tr|A0A2K6RYW9|A0A2K6RYW9_SAIBB
                                             LRRNEQNGAAAHVIAE DVKLLQFHVA TMVDNEL PGLPRAMQK$GRPLKSLKDKLKGKEGR
                                                                                                                                  349
tr|W5N8Z6|W5N8Z6 LEPOC
                                             LRRNEQSGAAAHVIAEDVKLLQFHVATMVDNELPGLPRAMQK$GRPLKSIKDKLKGKEGR
                                                                                                                                  348
tr|I3JRW6|I3JRW6 ORENI
                                             LRRNEQSGAAAHVIAE DVKLLQFHVATMVDNELPGLPRAMQKSGRPLKSIKDRLKGKEGR
                                                                                                                                  347
                                             LKRNEQSGAAAHVIAEDVKLLOFHVATMVDNELPGLPRAMOKSGRPLKSIKORLKGKEGR
tr|A0A0R4IMS9|A0A0R4IMS9 DANRE
                                                                                                                                  346
                                             LKRNEQSGAAAHVIAEDVKLLQFHVATMVDNELPGLPRAMQK$GRPLKSIKDKLKGKEGR
tr|A0A1A7X327|A0A1A7X327_9TELE
                                                                                                                                  347
tr|A0A1A8UKD7|A0A1A8UKD7 NOTFU
                                             LKRNEQSGAAAHVIAEDVKLLQFHVATMVDNELPGLPRAMQK$GRPLKSLKDRLKGKEGR
                                                                                                                                  347
tr|A0A1A8ER05|A0A1A8ER05 9TELE
                                             LKRNEQSGAAAHVIAEDVKLLQFHVATMVDNELPGLPRAMQKSGRPLKSIKDRLKGKEGR
                                                                                                                                  347
tr|A0A1A8DQ60|A0A1A8DQ60_9TELE
                                             LKRNEQSGAAAHVIAEDVKLLQFHVAIMVDNELPGLPRAMQK$GRPLKSLKORLKGKEGR
                                                                                                                                  347
                                             LKRNEQSGAAAHVIAEDVKLLQFHVATMVDNELPGLPRAMQK$GRPLKSLKDKLKGKEGR
tr|A0A1A8NSR8|A0A1A8NSR8_9TELE
                                                                                                                                  347
                                             LRRNEQSGAAAHVIAEDVILLQFHVATMVDNELPGLPRAMQKSGRPLKSIKORLKGKEGR
tr|A0A1W4YLM7|A0A1W4YLM7_9TELE
                                                                                                                                  347
tr|A0A1M8A6L7|A0A1M8A6L7 MALS4
                                             LRRMEQEGA PAHI LND FADLLQYHTA TYMDNDI AGLPQSLQS$GRPVKAI RARLKGKEGR
                                                                                                                                  349
                                             VOKLEMDGS POHI INEVEOLLOFHVA TYMDNDI AGOP QALOKSGRPVKA I RARLKGKEGR
VOKLEMDGS POHI INEVEOLLOFHVA TYMDNDI AGOP QALOKSGRPVKA I RARLKGKEGR
tr|A0A1B2J8C6|A0A1B2J8C6 PICPA
                                                                                                                                  336
tr|F2QW17|F2QW17_KOMPC
                                                                                                                                  336
tr|A3GID7|A3GID7 PICST
                                             VQRLEMDGSPQHVISEFEALLQFHVATYMDNDIAGQPQALQKTGRPIKSIRARLKGKEGR
                                                                                                                                  335
tr|A0A1D8PUA6|A0A1D8PUA6_CANAL
                                             VQRLET DGS PQHVISE FEALLQFHVATYMDNDI AGQPQALQKTGRPIKSI RARLKGKEGR
                                                                                                                                  335
                                             VORLEIDGS POHVISE FEALLOFH VATYMDNDIAGOP OALOKT GRPIKSIRARLKGKEGR
tr|G8BEH9|G8BEH9 CANPC
                                                                                                                                  335
                                                                     ***:*
                                                                             ** :**:: * *:::*.: ***: *:::
                                                  * .*: * : :
                                             VRGNIMGKRVDFSARTVITT DPNLS DOVGVPR SIAANMT FAE IVTP FNI DRLQELVRRG
VRGNIMGKRVDFSARTVITT DPNLS DOVGVPR SIAANMS FPE IVTP FNI DRLQELVRRG
tr|A0A1U8DYN0|A0A1U8DYN0 ALLSI
                                                                                                                                  408
tr|A0A1L8H4P4|A0A1L8H4P4 XENLA
                                                                                                                                  407
                                              VRGNLMGKRVDFSARTVITPDPNL$1DQVGVPRSIAANMTFAEIVTPFNIDRLQELVRRG
tr|H9GLG5|H9GLG5 ANOCA
                                                                                                                                  408
                                             VRGNLMGKRVDFSARIVIII ÞPNLSI DOVGVPR SIAANMT FAE IVTP FNI DRLOELVRRG
VRGNLMGKRVDFSARIVIII ÞPNLSI DOVGVPR SIAANMT FAE IVTP FNI DRLOELVRRG
tr|H2R1J6|H2R1J6 PANTR
                                                                                                                                  409
tr|G1MCZ1|G1MCZ1 AILME
                                                                                                                                  409
                                              VRGNIMGKRVDFSARTVITP DPNL$ DQVGVPH SIAANMT FAE IVTP FNI DRLQELVRRG
tr|008847|008847_MOUSE
                                                                                                                                  409
                                             VRGNIMGKRVDFSARIVITE DPNIS I DQVGVPR SIAANMT FAE IVTP FNI DRIQELVRRG
VRGNIMGKRVDFSARIVITE DPNIS I DQVGVPR SIAANMT FAE IVTP FNI DRIQELVRRG
tr|S7PWZ6|S7PWZ6 MYOBR
                                                                                                                                  409
tr|D4A5A6|D4A5A6 RAT
                                                                                                                                  409
                                              VRGNIMGKRVDFSARTVITP PPNL$ DQVGVPR SIAANMTFAE IVTP FNI DRLQELVRRG
sp|P08775|RPB1 MOUSE
                                                                                                                                  409
                                             VRGNIMGKRVDFSARTVITT DPNLSI DOVGVPR SIAANMI FAE IVTP FNI DRLQELVRRG
VRGNIMGKRVDFSARTVITT DPNLSI DOVGVPR SIAANMI FAE IVTP FNI DRLQELVRRG
VRGNIMGKRVDFSARTVITT DPNLSI DOVGVPR SIAANMI FAE IVTP FNI DRLQELVRRG
VRGNIMGKRVDFSARTVITT DPNLSI DOVGVPR SIAANMI FAE IVTP FNI DRLQELVRRG
sp|P24928|RPB1 HUMAN
                                                                                                                                  409
tr|A0A1S3EWL2|A0A1S3EWL2 DIPOR
                                                                                                                                  409
sp|P11414|RPB1 CRIGR
                                                                                                                                  409
tr|035559|035559 CRIGR
                                                                                                                                  409
                                              vrgnlmgkrvdfsartvitt ppnlsidqvgvprsiaanmtfaeivtpfnidrlqelvrrg
tr|A0A2I3M9H2|A0A2I3M9H2 PAPAN
                                             VRGNLMGKRVDFSARTVITP DPNLST DOVGVPH SIAANMT FAE IVTP FNI DRLOELVRRG
VRGNLMGKRVDFSARTVITT DPNLST DOVGVPH SIAANMT FAE IVTP FNI DRLOELVRRG
tr|F7HB40|F7HB40 MACMU
                                                                                                                                  409
tr|A0A2K6RYW9|A0A2K6RYW9 SAIBB
                                                                                                                                  409
                                             VRGNIMGKRVDFSARTVITF PPNLOI DOVGVPR SIAANMT FAE IVTF FNI DRLOELVRRG
VRGNIMGKRVDFSARTVITF DPNLOI DOVGVPR SIAANMT FPE IVTP FNI DRLOELVRRG
VRGNIMGKRVDFSARTVITF DPNLOI DOVGVPR SIAANMT FPE IVTP FNI DRLOELVRRG
tr|W5N8Z6|W5N8Z6 LEPOC
                                                                                                                                  407
tr|I3JRW6|I3JRW6 ORENI
tr|A0A0R4IMS9|A0A0R4IMS9 DANRE
                                                                                                                                  406
                                             VRGNIMGKRVDFSARTVITT DPNLOI DQVGVPR SIAANMT FPE IVTP FNI DRLQELVRRG
VRGNLMGKRVDFSARTVITT DPNLOI DQVGVPR SIAANMT FPE IVTP FNI DRLQELVRRG
VRGNLMGKRVDFSARTVITT DPNLOI DQVGVPR SIAANMT FPE IVTP FNI DRLQELVRRG
tr|A0A1A7X327|A0A1A7X327_9TELE
                                                                                                                                  407
tr|A0A1A8UKD7|A0A1A8UKD7 NOTFU
                                                                                                                                  407
tr|A0A1A8ER05|A0A1A8ER05 9TELE
                                                                                                                                  407
                                              vrgnlmgkrvdfsartvitt ppnlotdovgvpfsiaanmtfpeivtpfnidrloelvrrg
tr|A0A1A8DQ60|A0A1A8DQ60 9TELE
                                                                                                                                  407
                                             VRGNIMGKRVDFSARTVITT DPNILD DOVGVPR SIAANMT FPE IVTP FNI DRIQELVRRG
VRGNIMGKRVDFSARTVITT DPNILD DOVGVPR SIAANMT FPE IVTP FNI DRIQELVRRG
LRGNIMGKRVDFSARTVITG DPNIL I DOVGVPR SIARNITYPERVTPYNRAYISDIVRNG
tr|A0A1A8NSR8|A0A1A8NSR8_9TELE
tr|A0A1W4YLM7|A0A1W4YLM7 9TELE
                                                                                                                                  407
                                                                                                                                  407
tr|A0A1M8A6L7|A0A1M8A6L7 MALS4
                                                                                                                                  409
tr|A0A1B2J8C6|A0A1B2J8C6_PICPA
                                             LRGNLMGKRVDFSARTVISCOPNLEIDQVGVPISIAKTLSYPETVTQYNIHRLTEYVRNG
                                                                                                                                  396
                                             tr|F2QW17|F2QW17_KOMPC
                                                                                                                                  396
tr|A3GID7|A3GID7 PICST
                                                                                                                                  395
tr|A0A1D8PUA6|A0A1D8PUA6 CANAL
                                                                                                                                  395
tr|G8BEH9|G8BEH9_CANPC
                                                                                                                                  395
```

```
NSQYPGAKYIIRDNGDRIDLRFHPKPSDLHLQIGYKVERHMCDGDIVIFNRQPTLHKMSM
tr|A0A1U8DYN0|A0A1U8DYN0 ALLSI
                                                                                                                   468
                                        NSQYPGAKYIIRDNGDRIDLRTHPKPSDLHLQIGYKVERHMCDGDIVIFNRQPTLHKMSM
tr|A0A1L8H4P4|A0A1L8H4P4 XENLA
                                                                                                                   467
                                        NSQYPGAKYIIRDNGDRIDLRTHPKPSDLHLQTGYKVERHMCDGDIVIFNRQPTLHKMSM
tr|H9GLG5|H9GLG5_ANOCA
                                                                                                                   468
tr|H2R1J6|H2R1J6 PANTR
                                        NSQYPGAKYIIRDNGDRIDLRFHPKPSDLHLQTGYKVERHMCDGDIVIFNRQPTLHKMSM
                                                                                                                   469
                                        NSQYPGAKYIIRDNGDRIDLRFH PKP SDLHLQTGYKVERHMCDGDIVIFNRQPTLHKMSM
tr|G1MCZ1|G1MCZ1 AILME
                                                                                                                   469
                                        NSQYPGAKYIIRDNGDRIDLRFHPKPSDLHLOTGYKVERHMCDGDIVIFNROPTLHKMSM
NSQYPGAKYIIRDNGDRIDLRFHPKPSDLHLOTGYKVERHMCDGDIVIFNROPTLHKMSM
tr|008847|008847_MOUSE
tr|S7PWZ6|S7PWZ6_MYOBR
                                                                                                                   469
                                                                                                                   469
                                        NSQYPGAKYIIRDNGDRIDLRFHPKPSDLHLQTGYKVERHMCDGDIVIFNRQPTLHKMSM
tr|D4A5A6|D4A5A6 RAT
                                                                                                                   469
sp|P08775|RPB1 MOUSE
                                        NSQYPGAKYIIRDNGDRIDLRFHPKPSDLHLQTGYKVERHMCDGDIVIFNRQPTLHKMSM
                                                                                                                   469
                                        NSQYPGAKYIIRDNGDRIDLRFHPKPSDLHLQTGYKVERHMCDGDIVIFNRQPTLHKMSM
sp | P24928 | RPB1 HUMAN
                                                                                                                   469
                                        NSQYPGAKYIIRDNGDRIDLRFHPKPSDLHLQTGYKVERHMCDGDIVIFNRQPTLHKMSM
tr|A0A1S3EWL2|A0A1S3EWL2_DIPOR
                                                                                                                   469
                                        NSQYPGAKYIIRDNGDRIDLRFHPKPSDLHLQTGYKVERHMCDGDIVIFNRQPTLHKMSM
sp|P11414|RPB1 CRIGR
                                                                                                                   469
tr|035559|035559 CRIGR
                                        NSQYPGAKYIIRDNGDRIDLRFHPKPSDLHLQTGYKVERHMCDGDIVIFNRQPTLHKMSM
                                                                                                                   469
                                        NSQYPGAKYIIRDNGDRIDLRFHPKPSDLHLOTGYKVERHMCDGDIVIFNRQPTLHKMSM
NSQYPGAKYIIRDNGDRIDLRFHPKPSDLHLOTGYKVERHMCDGDIVIFNRQPTLHKMSM
NSQYPGAKYIIRDNGDRIDLRFHPKPSDLHLOTGYKVERHMCDGDIVIFNRQPTLHKMSM
tr|A0A2I3M9H2|A0A2I3M9H2 PAPAN
                                                                                                                   469
tr|F7HB40|F7HB40 MACMU
                                                                                                                   469
tr|A0A2K6RYW9|A0A2K6RYW9_SAIBB
                                                                                                                   469
tr|W5N8Z6|W5N8Z6_LEPOC
                                        NSQYPGAKYIIRDNGDRIDLRTHPKPSDLHLQTGYKVERHMCDGDIVIFNRQPTLHKMSM
                                                                                                                   468
tr|I3JRW6|I3JRW6 ORENI
                                        NSQYPGAKYIIRDNGDRIDLRFH PKP SDLHLQIGYKVERHMCDGDIVIFNRQPTLHKMSM
                                                                                                                   467
                                        NSQYPGAKYIIRDNGDRIDLRFHPKPSDLHLOIGYKVERHMCDGDIIVFNROPTLHKMSM
NSQYPGAKYIIRDNGDRIDLRFHPKPSDLHLOIGYKVERHMCEGDIVIFNROPTLHKMSM
tr|A0A0R4IMS9|A0A0R4IMS9_DANRE
tr|A0A1A7X327|A0A1A7X327_9TELE
                                                                                                                   466
                                                                                                                   467
tr|A0A1A8UKD7|A0A1A8UKD7 NOTFU
                                        NSQYPGAKYIIRDNGDRIDLRFHPKPSDLHLQTGYKVERHMCDGDIVIFNRQPTLHKMSM
                                                                                                                   467
                                        NSQYPGAKYIIRDNGDRIDLRFHPKPSDLHLQIGYKVERHMCDGDIVIFNRQPTLHKMSM
tr|A0A1A8ER05|A0A1A8ER05_9TELE
                                                                                                                   467
                                        NSQYPGAKYIIRDNGDRIDLRFHPKPSDLHLQIGYKVERHMCDGDIVIFNRQPTLHKMSM
tr|A0A1A8DQ60|A0A1A8DQ60 9TELE
                                                                                                                   467
                                        NSQYPGAKYIIRDNGDRIDLRFHPKPSDLHLOIGYKVERHMCDGDIVIFNRQPTLHKMSM
NSQYPGAKYIIRDNGDRIDLRFHPKPSDLHLOIGYKVERHMCDGDIIIFNRQPTLHKMSM
tr|A0A1A8NSR8|A0A1A8NSR8_9TELE
tr|A0A1W4YLM7|A0A1W4YLM7_9TELE
                                                                                                                   467
                                                                                                                   467
                                        ENEYPGARYVIED TGERIDLKYNR-RGDIALQAGWIVERHLKDGDYVLFNRQPSLHKMSM
tr|A0A1M8A6L7|A0A1M8A6L7 MALS4
                                                                                                                   468
                                        ENEHPGAKYVIRDNGDRIDLRYHKRAGDIVLQYGWKVERHLMDDDPVLFNRQPSLHKMSM
tr|A0A1B2J8C6|A0A1B2J8C6 PICPA
                                                                                                                   456
tr|F2QW17|F2QW17_KOMPC
                                         nehpgakyvirdngdridlr‡hkragdivlq¥gwkverhlmdddpvlfnrqpslhkmsm
                                                                                                                   456
                                         NEHPGAKYVIRD TGDRIDLRYNKRAGDIALQYGWKVERHLMDNDPVLFNRQPSLHKMSM
tr|A3GID7|A3GID7 PICST
                                                                                                                   455
                                        ENEHPGAKYVIRD TGDRIDLRYNKRAGDIALQYGWKVERHLMD DDPVLFNRQPSLHKMSM
tr|A0A1D8PUA6|A0A1D8PUA6_CANAL
                                                                                                                   455
                                        tr|G8BEH9|G8BEH9_CANPC
                                                                                                                   455
                                       MGHRVRIL WSTFRLNLSVTTPYNAD FDGDEM LHLPQSLETRAEIQELAMVPRNIVTPQ
tr|A0A1U8DYN0|A0A1U8DYN0_ALLSI
                                                                                                                 528
tr|A0A1L8H4P4|A0A1L8H4P4 XENLA
                                       MGHRVRILDWSTFRLNLSVTTPYNADFDGDEM LHLPQSLETRAEIQELAMVPRMIVTPQ
                                                                                                                 527
                                       MGHRVRILPWSTFRLNLSVTTPYNADFDGDEMNLHLPQSLETRAEIQEIAMVPRMIVTPQ
tr|H9GLG5|H9GLG5 ANOCA
                                                                                                                 528
tr|H2R1J6|H2R1J6 PANTR
                                       MGHRVRIL WSTFRLNLSVTTPYNAD FDGDEM
                                                                           NLHLPQSLETRAEIQELAMVPRMIVTPQ
                                                                                                                 529
                                       MGHRVRILDWSTFRLNLSVTTPYNADFDGDEMILHLPQSLETRAEIQEIAMVPRMIVTPQ
tr|G1MCZ1|G1MCZ1 AILME
                                                                                                                 529
                                       MGHRVRILPWSTFRLNLSVTTPYNADFDGDEM LHLPQSLETRAEIQEIAMVPRNIVTPQ
tr|008847|008847_MOUSE
                                                                                                                 529
tr|S7PWZ6|S7PWZ6 MYOBR
                                       MGHRVRIL WSTFRLNLSVTTPYNAD FDGDEM
                                                                           NLHLPQSLETRAEIQELAMVPRNIVTPQ
                                                                                                                 529
tr|D4A5A6|D4A5A6 RAT
                                       MGHRVRILPWSTFRLNLSVTTPYNADFDGDEMNLHLPQSLETRAEIQEIAMVPRMIVTPQ
                                                                                                                 529
sp|P08775|RPB1_MOUSE
                                       MGHRVRILDWSTFRLNLSVTTPYNADFDGDEM
                                                                           NLHLPOSLETRAEIQELAMVPRHIVTPO
                                                                                                                 529
                                       MGHRVRIL WSTFRLNLSVTTPYNAD FDGDEM LHLPQSLETRAEIQEIAMVPR MIVTPO
sp | P24928 | RPB1 HUMAN
                                                                                                                 529
tr|A0A1S3EWL2|A0A1S3EWL2 DIPOR
                                       MGHRVRIL WSTFRLNLSVTTPYN<mark>AD FDGDEMM</mark>LHLPQSLETRAEIQEIAMVPRNIVTPQ
                                                                                                                 529
sp|P11414|RPB1_CRIGR
                                       MGHRVRIL WSTFRLNLSVTTPYNAD FDGDEM LHLPQSLETRAEIQEIAMVPR NIVTPQ
                                                                                                                 529
                                       MGHRVRIL WSTFRLNLSVTTPYNAD FDGDEM LHLPQSLETRAEIQELAMVPRNIVTPQ
tr|035559|035559 CRIGR
                                                                                                                 529
tr|A0A2I3M9H2|A0A2I3M9H2_PAPAN
                                       MGHRVRILPWSTFRLNLSVTTPYNADFDGDEM
                                                                           NLHLPQSLETRAEIQELAMVPRHIVTPQ
                                                                                                                 529
                                       MGHRVRIL WSTFRLNLSVTTPYNADFDGDEMNLHLPQSLETRAEIQELAMVPRNIVTPQ
tr|F7HB40|F7HB40_MACMU
                                                                                                                 529
tr|A0A2K6RYW9|A0A2K6RYW9 SAIBB
                                       MGHRVRIL WSTFRLNLSVTTPYNADFDGDEM LHLPQSLETRAEIQELAMVPRHIVTPQ
                                                                                                                 529
tr|W5N8Z6|W5N8Z6 LEPOC
                                       MGHRVRIL WSTFRLNLSVTTPYNAD FDGDEM LHLPQSLETRAEIQELAMVPRHIVTPQ
                                                                                                                 528
                                       MGHRVRILDWSTFRLNLSVTTPYNADFDGDEMNLHLPQSLETRAEIQEIAMVPRMIVTPQ
tr|I3JRW6|I3JRW6 ORENI
                                                                                                                 527
                                       MGHRVRILPWSTFRLNLSVTTPYNADFDGDEM
                                                                           NLHLPOSLETRAEIOELAMVPRNIVTPO
tr|A0A0R4IMS9|A0A0R4IMS9_DANRE
                                                                                                                 526
                                                                           LHLPQSLETRAEIQELAMVPRHIVTPQ
tr|A0A1A7X327|A0A1A7X327_9TELE
                                       MGHRVRIL WSTFRLNLSVTTPYNAD FDGDEM
                                                                                                                 527
                                                                           LHLPOSLETRAEIQELAMVPRHIVTPO
tr|A0A1A8UKD7|A0A1A8UKD7_NOTFU
                                       MGHRVRIL WSTFRLNLSVTTPYNAD FDGDEM
                                                                                                                 527
tr|A0A1A8ER05|A0A1A8ER05 9TELE
                                       MGHRVRILDWSTFRLNLSVTTPYNADFDGDEM
                                                                           LHLPQSLETRAEIQELAMVPRMIVTPQ
                                                                                                                 527
                                       MGHRVRILDWSTFRLNLSVTTPYNADFDGDEM
tr|A0A1A8DQ60|A0A1A8DQ60_9TELE
                                                                           NLHLPQSLETRAEIQELAMVPRNIVTPQ
                                                                                                                 527
tr|A0A1A8NSR8|A0A1A8NSR8_9TELE
tr|A0A1W4YLM7|A0A1W4YLM7_9TELE
                                       MGHRVRIL WSTFRLNLSVTTPYNAD FDGDEM
                                                                           NLHLPQSLETRAEIQELAMVPRHIVTPQ
NLHLPQSLETRAEIQELAMVPRHIVTPQ
                                                                                                                 527
                                       MGHRVRIL WSTFRLNLSVTTPYNAD FDGDEM
                                                                                                                 527
                                       MAHRVKLMIYSTFRLNLSVTPPYN<mark>ADFDGDEMI</mark>LHVPQSEEARAELAQIAWVPRQIVSPQ
tr|A0A1M8A6L7|A0A1M8A6L7_MALS4
                                                                                                                 528
tr|A0A1B2J8C6|A0A1B2J8C6_PICPA
                                       MAHRVKVMPYSTFRLNLSVTSPYNADFDGDEM LHVPQSEETRAELSQLCAVPLQIVSPQ
                                                                                                                 516
                                                                           LHVPQSEETRAELSQLCAVPLQIVSPQ
LHVPQSPETRAELSEICAVPLQIVSPQ
                                       MAHRVKVMP YSTFRLNLSVT SPYN<mark>AD FDGDEM</mark>
tr|F2QW17|F2QW17_KOMPC
tr|A3GID7|A3GID7_PICST
                                                                                                                 516
                                       MAHRVKVM#YSTFRLNLSVTSPYNADFDGDEM
                                                                                                                 515
tr|A0A1D8PUA6|A0A1D8PUA6_CANAL
                                       MAHRVRVMPYSTFRLNLSVTSPYNADFDGDEM LHVPQSPETRSELSQICAVPLQIVSPQ
                                                                                                                 515
                                       MAHRVKVMPYSTFRLNLSVTSPYNADFDGDEM LHVPQSPETRSELSQICAVPLQIVSPQ
tr|G8BEH9|G8BEH9 CANPC
                                                                                                                 515
```

```
SNR PVMGIVQDTLTAVRKFT KRDVFLERGEVMNLLMFLSTWDGKVPQPAILKPR PLWTGK
tr|A0A1U8DYN0|A0A1U8DYN0 ALLSI
                                                                                                                                                          588
tr|A0A1L8H4P4|A0A1L8H4P4_XENLA
                                                      SNR PVMGIVQDTLTAVRKFTKRDVFLERGEVMNLLMFLSTWDGKVPQ PAILKPR PLWTGK
                                                                                                                                                          587
tr|H9GLG5|H9GLG5 ANOCA
                                                      SNR PVMGIVQDTLTAVRKFT KRDVFLERGEVMNLLMFLSTWDGKVPQ PAILKPR PLWTGK
                                                                                                                                                          588
tr|H2R1J6|H2R1J6 PANTR
                                                      SNR PVMGIVQDTLTAVRKFTKRDVFLERGEVMNLLMFLSTWDGKVPQPAILKPRPLWTGK
                                                                                                                                                          589
                                                      SNR PVMGIVQDTLITAVRKFTKRDVFLERGEVMNLLMFLSTWDGKVPQ PAVLKPR PLWTGK
tr|G1MCZ1|G1MCZ1 AILME
                                                     SNR PVMGIVODTLTAVRKFT KROVFLERGE VMNLLMFLSTWOG KVPOPAILKPR PLWTGK
SNR PVMGIVODTLTAVRKFT KROVFLERGE VMNLLMFLSTWOG KVPOPAILKPR PLWTGK
tr|008847|008847 MOUSE
                                                                                                                                                          589
tr|S7PWZ6|S7PWZ6 MYOBR
                                                                                                                                                          589
                                                      SNR PVMGIVQDTLTAVRKFT KRDVFLERGEVMNLLMFLST WDG KVPQ PAILKPR PLWTGK
                                                                                                                                                          589
tr|D4A5A6|D4A5A6 RAT
                                                      SNR PVMGIVQDTLITAVRKFTKRDVFLERGEVMNLLMFLSTWDGKVPQ PAILKPR PLWTGK
sp|P08775|RPB1_MOUSE
                                                                                                                                                          589
sp|P24928|RPB1 HUMAN
                                                      SNR PVMGIVQDTLTAVRKFTKRDVFLERGEVMNLLMFLSTWDGKVPQPAILKPRPLWTGK
                                                                                                                                                          589
tr|A0A1S3EWL2|A0A1S3EWL2 DIPOR
                                                      SNR PVMGIVQDTLTAVRKFTKRDVFLERGEVMNLLMFLSTWDGKVPQPAILKPR PLWTGK
                                                                                                                                                          589
sp|P11414|RPB1_CRIGR
                                                      SNR PVMGIVQDTLTAVRKFTKRDVFLERGEVMNLLMFLSTWDGKVPQPAILKPR PLWTGK
                                                     SNR PVMGIVODTL TAVRKFT KRDVFLERGEVMNLLMFLSTWDG KVPOPAILKPR PLWTGK
SNR PVMGIVODTL TAVRKFT KRDVFLERGEVMNLLMFLSTWDG KVPOPAILKPR PLWTGK
tr|035559|035559 CRIGR
                                                                                                                                                          589
tr|A0A2I3M9H2|A0A2I3M9H2_PAPAN
                                                                                                                                                          589
tr|F7HB40|F7HB40 MACMU
                                                      SNR PVMGIVQDTLTAVRKFTKRDVFLERGEVMNLLMFLSTWDGKVPQPAILKPRPLWTGK
                                                                                                                                                          589
tr|A0A2K6RYW9|A0A2K6RYW9_SAIBB
                                                      SNR PVMGIVQDTLTAVRKFTKRDVFLERGEVMNLLMFLSTWDGKVPQPAILKPRPLWTGK
                                                                                                                                                          589
tr|W5N8Z6|W5N8Z6 LEPOC
                                                      SNR PVMGIVQDTLTAVRKFTKRDVFLDRGEVMNLLMFLSTWDGKVPQPAILKPRPLWTGK
                                                                                                                                                          588
                                                      SNR PVMGIVQDTLTAVRKFTKRDVFLERGEVMNLLMFLSTWDGKMPQPAILKPRPLWTGK
tr|I3JRW6|I3JRW6 ORENI
                                                                                                                                                          587
                                                      SNR PVMGIVODTLITAVRKFT KRDVFLERGEVMNLLMFLSTWDGKVPQ PAI LKPR PLWTGK
tr|A0A0R4IMS9|A0A0R4IMS9_DANRE
                                                                                                                                                          586
tr|A0A1A7X327|A0A1A7X327_9TELE
tr|A0A1A8UKD7|A0A1A8UKD7_NOTFU
                                                     SNR PVMGIVODTLTAVRKFT KROVFLERGE VMNLLMFLST WOG KMPO PAILKPR PLWTGK
SNR PVMGIVODTLTAVRKFT KROVFLERGE VMNLLMFLST WOG KMPO PAILKPR PLWTGK
                                                                                                                                                          587
                                                                                                                                                          587
                                                      SNR PVMGIVQDTLTAVRKFTKRDVFLERGEVMNLLMFLSTWDGKMPQPAILKPRPLWTGK
tr|A0A1A8ER05|A0A1A8ER05_9TELE
                                                                                                                                                          587
tr|A0A1A8DQ60|A0A1A8DQ60 9TELE
                                                      SNR PVMGIVQDTLITAVRKFTKRDVFLERGEVMNLLMFLSTWDGKMPQPAILKPR PLWTGK
                                                                                                                                                          587
tr|A0A1A8NSR8|A0A1A8NSR8_9TELE
tr|A0A1W4YLM7|A0A1W4YLM7_9TELE
                                                      SNR PVMGIVQDTLTAVRKFTKRDVFLERGEVMNLLMFLSTWDGKMPQ PAILKPR PLWTGK
                                                                                                                                                          587
                                                      SNR PVMGIVQDTLTAVRKFTKRDVFLERGEVMNLLMFLSTWDGKVPQPAILKPRPLWSGK
                                                                                                                                                          587
tr|A0A1M8A6L7|A0A1M8A6L7_MALS4
                                                      ANKPVMGIVQDTLCGIRKFTLRDCLLDFDQVQNVLMWLTIWDGIVPQPCILKPKPYWSGK
                                                                                                                                                          588
                                                     SNKPVMGIVODTLCGVRMTLRDTFIEYEQVMNMLFWVPSWDGVVPQPAIMKPKPLWTGK
SNKPVMGIVODTLCGVRMTLRDTFIEYEQVMNMLFWVPSWDGVVPQPAILKPKPLWTGK
SNKPVMGIVODTLCGIRMTLRDNFIDYDQVMNMLYWIPNWDGVIPPPAIAKPKPLWTGK
tr|A0A1B2J8C6|A0A1B2J8C6_PICPA
                                                                                                                                                          576
tr|F2QW17|F2QW17_KOMPC
                                                                                                                                                          576
tr|A3GID7|A3GID7_PICST
                                                                                                                                                          575
                                                      SNKPVMGIVQDTLCGIRKMTLRDIFIEYDQVMNMCYWIPNWDGVIPPPAVVKPKQLWTGK
tr|A0A1D8PUA6|A0A1D8PUA6 CANAL
                                                                                                                                                          575
                                                     tr|G8BEH9|G8BEH9 CANPC
tr|A0A1U8DYN0|A0A1U8DYN0 ALLSI
                                                     QVFSLIIPGHINCVRTHSTHPDDEDSGPYKHISPGDTKVIVENGELIMGILCKKSLGTSA
tr|A0A1L8H4P4|A0A1L8H4P4 XENLA
                                                      OT FST. TVPGHTNCTR THSTHPDDEDSGPYKHT SPGDTKVTVENGELVMGTLCKKSLGTSA
                                                                                                                                                          647
tr|H9GLG5|H9GLG5 ANOCA
                                                      QIFSLIIPGHINCIRTHSTHPDDEDSGPYKHISPGDTKVIVENGELIMGILCKKS<mark>LGT</mark>SA
                                                                                                                                                          648
tr|H2R1J6|H2R1J6 PANTR
                                                     QIFSLIIPGHINCIRTHSTHPDDEDSGPYKHISPGDTKVVVENGELIMGILCKKSLGTSA
                                                                                                                                                          649
                                                     QIFSLIVPGHINCIRTHSTHPDDEDSGPYKHISPGDTKVVVENGELIMGILCKKSLG
tr|G1MCZ1|G1MCZ1_AILME
                                                                                                                                                          649
tr|008847|008847 MOUSE
                                                      QIFSLIIPGHINCIRTHSTHPDDEDSGPYKHISPGDTKVVVENGELIMGILCKKSLG
                                                                                                                                                          649
tr|S7PWZ6|S7PWZ6 MYOBR
                                                     QIFSLIIPGHINCIRTHSTHPDDEDSGPYKHISPGDTKVVVENGELIMGILCKKS<mark>LGT</mark>SA
                                                                                                                                                          649
                                                     QIFSLIIPGHINCIRTHSTHPDDEDSGPYKHISPGDTKVVVENGELIMGILCKKSLG
tr|D4A5A6|D4A5A6 RAT
                                                                                                                                                          649
sp|P08775|RPB1 MOUSE
                                                      QIFSLIIPGHINCIRTHSTHPDDEDSGPYKHISPGDTKVVVENGELIMGILCKKSLGTSA
                                                                                                                                                          649
sp|P24928|RPB1 HUMAN
                                                     QIFSLIIPGHINCIRTHSTHPDDEDSGPYKHISPGDTKVVVENGELIMGILCKKSLGTSA
                                                                                                                                                          649
                                                      QIFSLIIPGHINCIRTHSTHPDDEDSGPYKHISPGDTKVVVENGELIMGILCKKSLG
tr|A0A1S3EWL2|A0A1S3EWL2 DIPOR
                                                                                                                                                          649
sp|P11414|RPB1_CRIGR
                                                      QIFSLIIPGHINCIRTHSTHPDDEDSGPYKHISPGDTKVVVENGELIMGILCKKS<mark>LG</mark>
                                                                                                                                                          649
tr|035559|035559 CRIGR
                                                     QIFSLIIPGHINCIRTHSTHPDDEDSGPYKHISPGDTKVVVENGELIMGILCKKS<mark>LGT</mark>SA
                                                                                                                                                          649
tr|A0A2I3M9H2|A0A2I3M9H2_PAPAN
                                                      QIFSLIIPGHINCIRTHSTHPDDEDSGPYKHISPGDTKVVVENGELIMGILCKKSLG
                                                                                                                                                          649
tr|F7HB40|F7HB40_MACMU
                                                      QIFSLIIPGHINCIRTHSTHPDDEDSGPYKHISPGDTKVVVENGELIMGILCKKSLG
                                                                                                                                                          649
tr|A0A2K6RYW9|A0A2K6RYW9 SAIBB
                                                      OTESTITE CHINCIP THAT HE PODED SCHOKK SECTION AND A SECTION AS A SECTI
                                                                                                                                                          649
tr|W5N8Z6|W5N8Z6 LEPOC
                                                      QVFSLIIPGHINAIRTHSTHPDEEDSGPYKHISPGDTKVIVENGELIMGILCKKS<mark>LGT</mark>SA
                                                                                                                                                          648
tr|I3JRW6|I3JRW6 ORENI
                                                      QIFSLIIPGHINVIRTHSTHPDDEDSGPYKHISPGDTKVIVENGELIMGILCKKS<mark>LG</mark>
                                                                                                                                                          647
tr|A0A0R4IMS9|A0A0R4IMS9 DANRE
                                                      OIFSLIIPGHINAIRTHSTHPDEEDSGPYKNISPGDTKVIVENGELIMGILCKKSLGTSA
                                                                                                                                                          646
tr|A0A1A7X327|A0A1A7X327_9TELE
                                                      QVFSLIIPGHINVIRTHSTHPDDEDSGPYKHISPGDTKVIVENGELIMGILCKKSLG
                                                                                                                                                          647
tr|A0A1A8UKD7|A0A1A8UKD7 NOTFU
                                                      QVFSLIIPGHINVIRTHSTHPDDEDSGPYKHISPGDTKVIVENGELIMGILCKKSLG
                                                                                                                                                          647
tr|A0A1A8ER05|A0A1A8ER05 9TELE
                                                                                                                                                          647
                                                      QVFSLIIPGHINVIRTHSTHPDDEDSGPYKHISPGDTKVIVENGELIMGILCKKSLGTSA
tr|A0A1A8DQ60|A0A1A8DQ60 9TELE
                                                      QVFSLIIPGHINVIRTHSTHPDDEDSGPYKHISPGDTKVIVENGELIMGILCKKS<mark>LGT</mark>SA
                                                                                                                                                          647
tr|A0A1A8NSR8|A0A1A8NSR8 9TELE
                                                      QVFSLIIPGHINVIRTHSTHPDDEDSGPYKHISPGDTKVIVENGELIMGILCKKSLG
                                                                                                                                                          647
tr|A0A1W4YLM7|A0A1W4YLM7 9TELE
                                                      QVFSLIIPGHINAIRTHSTHPDDEDSGPYKHISPGDTKVIVENGELIMGILCKKS<mark>LGT</mark>SA
                                                                                                                                                          647
                                                      QLLSMCIPKGINVFLG-----DAKAAANNFLKDDGVHIENGEIMYGVINKKV<mark>VG$</mark>SA
tr|A0A1M8A6L7|A0A1M8A6L7 MALS4
tr|A0A1B2J8C6|A0A1B2J8C6_PICPA
                                                      QLLSIAIPSGIHLQRT-----DGGN-SLLSPKDNQMLIVDGNVMFGVVDKKTVG$GG
                                                                                                                                                          627
                                                      QLLSIAIPSGIHLQRT-----DGGN-SLLSPKDNCMLIVDGKVMFGVVDKKTVGSGG
tr|F2QW17|F2QW17 KOMPC
                                                                                                                                                          627
                                                      QLLSMAIPKGIHLQRF-----DGGK-DLLSPKDTGMLIVDGEIMFGVVDKKTVGATG
tr|A3GID7|A3GID7 PICST
                                                                                                                                                          626
tr|A0A1D8PUA6|A0A1D8PUA6_CANAL
                                                     QLLSLAIPKGIHLQRF-----DGGR-DLLSPKDTGMLIVDGEIMFGVVDKKT<mark>VGA</mark>TG
                                                                                                                                                          626
                                                     QMLSMAIPKGIHIQRF-----DGGR-DLLSPKDTQMLIVDGEIMFGVVDKKT<mark>VGA</mark>TG
tr|G8BEH9|G8BEH9 CANPC
                                                      *::*: :* *:
                                                                                                        * : : :*::: *:: ** <mark>:*:</mark>
```

```
GSLVHISYLEMGHDVTRLXXXXXXXXXXPQTPLALHFLVXKRGRGHT1GIGDSIADAKTYQ
tr|A0A1U8DYN0|A0A1U8DYN0 ALLSI
                                                                                                            708
                                     GSLVHISYLEMGHDTTRLFYSNIQTVIN-----N---WLLIEGHTIGIGDSIADAKTYO
tr|A0A1L8H4P4|A0A1L8H4P4 XENLA
                                                                                                            698
                                     GSLVHISYLEMGHDVTRLFYSNIOTVIN-----N---WLLIEGHTIGIGDSIADAKTYO
tr|H9GLG5|H9GLG5 ANOCA
                                                                                                            699
tr|H2R1J6|H2R1J6 PANTR
                                     GSLVHISYLEMGHDITRLFYSNIQTVIN-----N--WLLIEGHTIGIGDSIADSKTYQ
                                                                                                            700
                                     GSLVHISYLEMGHDVTRLFYSNIQTVIN-----N---WLLIEGHTIGIGDSIADSKTYQ
tr|G1MCZ1|G1MCZ1 AILME
                                     GSLVHISYLEMGHDITRLFYSNIQTVIN-----N---WLLIEGHTIGIGDSIADSKTYQ
tr|008847|008847 MOUSE
                                                                                                            700
                                     GSLVHISYLEMGHDITRLFYSNIQTVIN-----N---WLLIEGHTIGIGDSIADSKTYQ
tr|S7PWZ6|S7PWZ6 MYOBR
                                                                                                            700
                                     GSLVHISYLEMGHDVTRLFYSNIQTVIN-----N---WLLIEGHT GIGDSIADSKTYQ
tr|D4A5A6|D4A5A6_RAT
                                                                                                            700
sp|P08775|RPB1 MOUSE
                                     GSLVHISYLEMGHDITRLFYSNIQTVIN-----N---WLLIEGHTIGIGDSIADSKTYQ
                                                                                                            700
                                     GSLVHISYLEMGHDITRLFYSNIQTVIN-----N---WLLIEGHTIGIGDSIADSKTYQ
sp | P24928 | RPB1 HUMAN
                                                                                                            700
tr|A0A1S3EWL2|A0A1S3EWL2 DIPOR
                                     GSLVHISYLEMGHDITRLFYSNIQTVIN----N--WLLIEGHTIGIGDSIADSKTYQ
                                                                                                            700
                                     GSLVHISYLEMGHDITRLFYSNIQTVIN-----N---WLLIEGHTIGIGDSIADSKTYO
sp|P11414|RPB1 CRIGR
                                                                                                            700
                                     GSLVHISYLEMGHDITRLFYSNIQTVIN-----N---WLLIEGHTIGIGDSIADSKTYQ
tr|035559|035559 CRIGR
                                                                                                            700
                                     GSLVHISYLEMGHDITRLFYSNIQTVIN----N--WLLIEGHT GIGDSIADSKTYQ
tr|A0A2I3M9H2|A0A2I3M9H2 PAPAN
                                                                                                            700
tr|F7HB40|F7HB40 MACMU
                                     GSLVHISYLEMGHDITRLFYSNIQTVIN-----N---WLLIEGHTIGIGDSIADSKTYQ
                                                                                                            700
                                     GSLVHISYLEMGHDITRLFYSNIQTVIN-----N---WLLIEGHTIGIGDSIADSKTYQ
tr|A0A2K6RYW9|A0A2K6RYW9 SAIBB
                                                                                                            700
                                     GSLVHISYLEMGHDVTRLFYSNIQTVVNN------WLLIEGHSIGIGDSIADAKTYL
tr|W5N8Z6|W5N8Z6_LEPOC
                                                                                                            699
tr|I3JRW6|I3JRW6 ORENI
                                     GSLVHISYLEMGHDITRLFYSNIQTVVNN------WLLIEGHSIGIGDSIADAKTYL
                                                                                                            698
                                     GSLVHISYLEMGHDITRLFYSNIQTVVNN------WLLIEGHSIGIGDSIADKATYQ
tr|A0A0R4IMS9|A0A0R4IMS9 DANRE
                                     GSLVHISYLEMGHDITRLFYSNIQTVVNN------WLLIEGHSIGIGDSIADAKTYL
tr|A0A1A7X327|A0A1A7X327 9TELE
                                     GSLVHISYLEMGHDITRLFYSNIOTVVNN------WLLIEGHSIGIGDSIADAKTYL
tr|A0A1A8UKD7|A0A1A8UKD7_NOTFU
                                                                                                            698
tr|A0A1A8ER05|A0A1A8ER05 9TELE
                                     GSLVHISYLEMGHDITRLFYSNIQTVVNN------WLLIEGHS GIGDSIADAKTYL
                                                                                                            698
                                     GSLVHISYLEMGHDITRLFYSNIQTVVNN------WLLIEGHS GIGDSIADAKTYL
tr|A0A1A8DQ60|A0A1A8DQ60_9TELE
                                                                                                            698
                                     GSLVHISYLEMGHDITRLFYSNIQTVVNN------WLLIEGHS GIGDSIADAKTYL
tr|A0A1A8NSR8|A0A1A8NSR8 9TELE
                                                                                                            698
                                     GSLVHISYLEMGHDVTRLFYSNIQTVVNN------WLLIEGHSIGIGDSIADAKTYL
tr|A0A1W4YLM7|A0A1W4YLM7 9TELE
                                                                                                            698
                                     GGLIHIIFRERGPVVCRDFFSGVORLVNF------WLLHNGFSIGIGDTVADKATTA
tr|A0A1M8A6L7|A0A1M8A6L7 MALS4
                                                                                                            691
                                     GGLIHTVMREKGPKICAELFGNIQKVVNY------WLLHNGFSIGIGDAIADASTMK
tr|A0A1B2J8C6|A0A1B2J8C6 PICPA
                                                                                                            678
                                     GGLIHTVMREKGPKICAELFGNIQKVVNY------WLLHNGFSIGIGDAIADASTMK
tr|F2QW17|F2QW17 KOMPC
                                     GGLIHTVMREKGPRVCAQLFSSIQKVTNY------WLLHNGFSIGIGDTIADVSTMK
tr|A3GID7|A3GID7 PICST
                                                                                                            677
                                     GGLIHTVMREKGPKVCAELFSSIQKVVNY------WLLHNGFSIGIGDTIADAQTMR
tr|A0A1D8PUA6|A0A1D8PUA6 CANAL
                                                                                                            677
                                     GGLIHTVMREKGPOVCAQLFSSIQKVVNF------WLLHNGFS GIGDTIADASTMK
tr|G8BEH9|G8BEH9_CANPC
                                                                                                            677
                                                                                       ****::**
                                      DIONTIKKAKODVIEVIEKAHNNELEPTPGNTLROTFENOVNRILNDARDKTGSSAOKSL
tr|A0A1U8DYN0|A0A1U8DYN0 ALLSI
                                                                                                             768
                                      DIONTIKKAKODVIEVIEKAHNNELEPTPGNILROTFENOVNRILNDARDKTGSSAOKSL
DIONTIKKAKODVIEVIEKAHNNELEPTPGNILROTFENOVNRILNDARDKTGSSAOKSL
tr|A0A1L8H4P4|A0A1L8H4P4 XENLA
                                                                                                             758
tr|H9GLG5|H9GLG5 ANOCA
                                                                                                             759
                                      DIQNTIKKAKQDVIEVIEKAHNNELEPTPGNTLRQTFENQVNRILNDARDKTGSSAQKSL
tr|H2R1J6|H2R1J6_PANTR
                                                                                                             760
tr|G1MCZ1|G1MCZ1 AILME
                                      DIONTIKKAKODVIEVIEKAHNNELEPTPGNTLROTFENOVNRILNDARDKTGSSAOKSL
                                                                                                             760
tr|008847|008847_MOUSE
                                      DIQNTIKKAKQDVIEVIEKAHNNELEPTPGNTLRQTFENQVNRILNDARDKTGSSAQKSL
                                                                                                             760
tr|S7PWZ6|S7PWZ6 MYOBR
                                      DIQNTIKKAKQDVIEVIEKAHNNELEPTPGNTLRQTFENQVNRILNDARDKTGSSAQKSL
                                                                                                             760
                                      DIQNTIKKAKQDVIEVIEKAHNNELEPTPGNTLRQTFENQVNRILNDARDKTGSSAQKSL
DIQNTIKKAKQDVIEVIEKAHNNELEPTPGNTLRQTFENQVNRILNDARDKTGSSAQKSL
tr|D4A5A6|D4A5A6 RAT
                                                                                                             760
sp | P08775 | RPB1 MOUSE
                                                                                                             760
                                      DIQNTIKKAKQDVIEVIEKAHNNELEPTPGNTLRQTFENQVNRILNDARDKTGSSAQKSL
sp | P24928 | RPB1_HUMAN
                                                                                                             760
tr|A0A1S3EWL2|A0A1S3EWL2 DIPOR
                                      DIQNTIKKAKQDVIEVIEKAHNNELEPTPGNTLRQTFENQVNRILNDARDKTGSSAQKSL
                                                                                                             760
sp|P11414|RPB1_CRIGR
                                      DIONTIKKAKODVIEVIEKAHNNELEPTPGNTLROTFENOVNRILNDARDKTGSSAOKSL
                                                                                                             760
tr|035559|035559 CRIGR
                                      DIQNTIKKAKQDVIEVIEKAHNNELEPTPGNTLRQTFENQVNRILNDARDKTGSSAQKSL
                                                                                                             760
                                      DIONTIKKAKODVIEVIEKAHNNELEPTEGNTIROT FENOVARILADARDKTGSSAOKSL
DIONTIKKAKODVIEVIEKAHNNELEPTEGNTIROT FENOVARILADARDKTGSSAOKSL
tr|A0A2I3M9H2|A0A2I3M9H2 PAPAN
                                                                                                             760
tr|F7HB40|F7HB40 MACMU
                                                                                                             760
                                      DIQNTIKKAKQDVIEVIEKAHNNELEPTPGNTLRQTFENQVNRILNDARDKTGSSAQKSL
tr|A0A2K6RYW9|A0A2K6RYW9 SAIBB
                                                                                                             760
tr|W5N8Z6|W5N8Z6 LEPOC
                                      DIQNTIKKAKQDVIEVIEKAHNNELEPTPGNTLRQTFENQVNRILNDARDKTGSSAQKSL
                                                                                                             759
                                      DIONTIKKAKODVIEVIEKAHNNELEPTPGNTLROTFENOVNRILNDARDKTGSSAOKSL
tr|I3JRW6|I3JRW6 ORENI
                                                                                                             758
tr|A0A0R4IMS9|A0A0R4IMS9 DANRE
                                      DIQNTIKKAKQDVIEVIEKAHNNELEPTPGNTLRQTFENQVNRILNDARDKTGSSAQKSL
                                                                                                             757
                                      DIONTIKKAKODVIEVIEKAHNNELEPTPGNILROTFENOVNRILNDARDKTGSSAOKSL
DIONTIKKAKODVIEVIEKAHNNELEPTPGNILROTFENOVNRILNDARDKTGSSAOKSL
tr|A0A1A7X327|A0A1A7X327_9TELE
tr|A0A1A8UKD7|A0A1A8UKD7_NOTFU
                                                                                                             758
                                                                                                             758
                                      DIONTIKKAKODVIEVIEKAHNNELEPTPGNTLROTFENOVNRILNDARDKTGSSAOKSL
tr|A0A1A8ER05|A0A1A8ER05 9TELE
                                                                                                             758
                                      DIQNTIKKAKQDVIEVIEKAHNNELEPTPGNTLRQTFENQVNRILNDARDKTGSSAQKSL
tr|A0A1A8DQ60|A0A1A8DQ60 9TELE
tr|A0A1A8NSR8|A0A1A8NSR8 9TELE
                                      DIONTIKKAKODVIEVIEKAHNNELEPTPGNTLROTFENOVNRILNDARDKTGSSAOKSL
                                                                                                             758
tr|AOA1W4YLM7|AOA1W4YLM7 9TELE
                                      DIONTIKKAKODVIEVIEKAHNNELEPTPGNTLROTFENOVNRILNDARDKTGSSAOKSL
                                                                                                             758
                                      NINOTIAKAKADVMDLIQAARHDWLKADPGMTLRESFEANVNRILNKARDDVGSHAEONL
tr|A0A1M8A6L7|A0A1M8A6L7 MALS4
                                                                                                             751
tr|A0A1B2J8C6|A0A1B2J8C6 PICPA
                                      EITHAISSAKEQVQEIIYKAQHNELELKPGMTLRESFEGEVSRTLNDARD$AGRSAEMNL
                                                                                                             738
                                      EITHAISSAKEQVQEIIYKAQHNELELKPGMTLRESFEGEVSRTLNDARDSAGRSAEMNL
tr|F2QW17|F2QW17_KOMPC
                                                                                                             738
tr|A3GID7|A3GID7_PICST
                                      DITSTISEAKIKVQEIILDAQSNKLEPEPGMTLRESFEHNVSRVLNQARDTAGRSAEMNL
                                      DVNKTIQEAKQKVQEIIIDAQHNKLEPEPGMTLRESFEHNVSRVLNQARDTAGRSAEMSL
tr|A0A1D8PUA6|A0A1D8PUA6 CANAL
                                                                                                             737
                                      DITKTIQEAKQKVQEIILDAQHNKLDPEPGMTLRESFEHNVSRVLNQARDTAGRSAEMSL
tr|G8BEH9|G8BEH9 CANPC
                                                                                                             737
                                                                     * ***::**
                                      :: :* .** .* ::
                                                          *:: *
```

```
SEYNNFKSMVVSGAKGSKINISQVIAVVGQQUVEGKRIPFGFKHRTLPHFIKDDYFPESR
SEYNNFKSMVVSGAKGSKINISQVIAVVGQQUVEGKRIPFGFKHRTLPHFIKDDYFPESR
SEYNNFKSMVVSGAKGSKINISQVIAVVGQQUVEGKRIPFGFKHRTLPHFIKDDYFPESR
SEYNNFKSMVVSGAKGSKINISQVIAVVGQQUVEGKRIPFGFKHRTLPHFIKDDYFPESR
tr|A0A1U8DYN0|A0A1U8DYN0 ALLSI
                                                                                                                                         828
tr|A0A1L8H4P4|A0A1L8H4P4 XENLA
                                                                                                                                         818
tr|H9GLG5|H9GLG5 ANOCA
                                                                                                                                         819
tr|H2R1J6|H2R1J6 PANTR
                                                SEYNNFKSMVVSGAKGSKINISQVIAVVGQQUVEGKRIPFGFKHRTLPHFIKDDYSPESR
SEYNNFKSMVVSGAKGSKINISQVIAVVGQQUVEGKRIPFGFKHRTLPHFIKDDYSPESR
                                                                                                                                         820
tr|G1MCZ1|G1MCZ1_AILME
tr|008847|008847 MOUSE
                                                                                                                                         820
tr|S7PWZ6|S7PWZ6 MYOBR
                                                SEYNNFKSMVVSGAKGSKINISQVIAVVGQQQVVEGKRIPFGFKHRTLPHF1KDDYGPESR
                                                                                                                                         820
                                               SEYNNFKSMVVSGAKGSKINISQVIAVVGQQNVEGKRIPFGFKHRTLPHFIKDDYPESR
SEYNNFKSMVVSGAKGSKINISQVIAVVGQQNVEGKRIPFGFKHRTLPHFIKDDYPESR
SEYNNFKSMVVSGAKGSKINISQVIAVVGQQNVEGKRIPFGFKHRTLPHFIKDDYPESR
SEYNNFKSMVVSGAKGSKINISQVIAVVGQQNVEGKRIPFGFKHRTLPHFIKDDYPESR
SEYNNFKSMVVSGAKGSKINISQVIAVVGQQNVEGKRIPFGFKHRTLPHFIKDDYPESR
SEYNNFKSMVVSGAKGSKINISQVIAVVGQQNVEGKRIPFGFKHRTLPHFIKDDYPESR
SEYNNFKSMVVSGAKGSKINISQVIAVVGQQNVEGKRIPFGFKHRTLPHFIKDDYPESR
SEYNNFKSMVVSGAKGSKINISQVIAVVGQQNVEGKRIPFGFKHRTLPHFIKDDYPESR
SEYNNFKSMVVSGAKGSKINISQVIAVVGQQNVEGKRIPFGFKHRTLPHFIKDDYPESR
SEYNNFKSMVVSGAKGSKINISQVIAVVGQQNVEGKRIPFGFKHRTLPHFIKDDYPESR
SEYNNFKSMVVSGAKGSKINISQVIAVVGQQNVEGKRIPFGFKHRTLPHFIKDDYPESR
SEYNNFKSMVVSGAKGSKINISQVIAVVGQQNVEGKRIPFGFKHRTLPHFIKDDYPESR
                                                                                                                                         820
triD4A5A6|D4A5A6 RAT
sp|P08775|RPB1 MOUSE
                                                                                                                                         820
sp | P24928 | RPB1 HUMAN
                                                                                                                                         820
tr|A0A1S3EWL2|A0A1S3EWL2 DIPOR
                                                                                                                                         820
sp|P11414|RPB1 CRIGR
                                                                                                                                         820
tr|035559|035559_CRIGR
                                                                                                                                         820
tr|A0A2I3M9H2|A0A2I3M9H2 PAPAN
                                                                                                                                         819
tr|F7HB40|F7HB40 MACMU
                                                                                                                                         820
tr|A0A2K6RYW9|A0A2K6RYW9 SAIBB
                                                                                                                                         820
tr|W5N8Z6|W5N8Z6 LEPOC
                                                                                                                                         819
                                               SEYNNFKSMVVAGSKGSKINISQVIAVVGQQUVEGKRIPFGFKHRILPHFIKDDYSPESR
SEYNNFKSMVVAGSKGSKINISQVIAVVGQQUVEGKRIPFGFKHRILPHFIKDDYSPESR
SEYNNFKSMVVAGSKGSKINISQVIAVVGQQUVEGKRIPFGFKHRILPHFIKDDYSPESR
tr||I3JRW6||I3JRW6 ORENI
                                                                                                                                         818
tr|A0A0R4IMS9|A0A0R4IMS9 DANRE
                                                                                                                                         817
tr|A0A1A7X327|A0A1A7X327 9TELE
                                                                                                                                         818
                                                SEYNNFKSMVVAGSKGSKINISQVIAVVGQOVVEGKRIPFGFKHRILPHFIKDDYSPLSR
SEYNNFKSMVVAGSKGSKINISQVIAVVGQOVVEGKRIPFGFKHRILPHFIKDDYSPLSR
SEYNNFKSMVVAGSKGSKINISQVIAVVGQOVVEGKRIPFGFKHRILPHFIKDDYSPLSR
tr|A0A1A8UKD7|A0A1A8UKD7 NOTFU
                                                                                                                                         818
tr|A0A1A8ER05|A0A1A8ER05 9TELE
                                                                                                                                         818
tr|A0A1A8DQ60|A0A1A8DQ60 9TELE
                                                                                                                                         818
                                                SEYNNFKSMVVAGSKGSKINISQVIAVVGQUIVEGKRIPFGFKHRTLPHFIKDDYSPESR
SEYNNFKSMVVAGSKGSKINISQVIAVVGQUIVEGKRIPFGFKHRTLPHFIKDDYSPESR
PDWNNVKQMVIAGSKGSFINISQMSACVGQQSVEGKRIPFGFRHRSLPHFIKDDFIPESR
tr|A0A1A8NSR8|A0A1A8NSR8 9TELE
                                                                                                                                         818
tr|A0A1W4YLM7|A0A1W4YLM7 9TELE
                                                                                                                                         818
tr|A0A1M8A6L7|A0A1M8A6L7 MALS4
                                                                                                                                         811
                                                KDLNNVKOMVSAGSKGSFINIAOMSACVGOONVEGKRIAFGFADRSLPHFTKDDFSPESK
tr|A0A1B2J8C6|A0A1B2J8C6_PICPA
                                                                                                                                         798
                                                KDLNNVKOMVSAGSKGSFINIAOMSACVGOMVEGKRIAFGFADRSLPHFTKDDFSPESK
KDLNNVKOMVVSGSKGSFINISOMSACVGOM VEGKRIPFGFSDRTLPHFTKDDYSPESK
KDLNNVKOMVTSGSKGSFINISOMSACVGOM VEGKRIPFGFGDRSLPHFTKDDYSPESK
tr|F2QW17|F2QW17 KOMPC
                                                                                                                                         798
tr|A3GID7|A3GID7 PICST
                                                                                                                                         797
tr|A0A1D8PUA6|A0A1D8PUA6_CANAL
                                                                                                                                         797
                                                KDLNNVKOMVTSGSKGSFINISOMSACVGOOTVEGKRIPFGFALRSLPHFTKDDYSPESK
tr|G8BEH9|G8BEH9 CANPC
                                                                                                                                         797
                                               GEVENSYLAGLTPTEFFFHAMSGREGLIDTAVKTAETGYIORRLIKSMESVMVKYDATVR
tr|A0A1U8DYN0|A0A1U8DYN0 ALLSI
                                                                                                                                        888
tr|A0A1L8H4P4|A0A1L8H4P4 XENLA
                                                GFVENSYLAGLTPTEFFFHAMGGREGLIDTAVKTAETGYIQRRLIKSMESVMVKYDATVR
                                                                                                                                        878
                                                GFVENSYLAGLTPTEFFFHAMGGREGLIDTAVKTAETGYIQRRLIKSMESVMVKYDATVR
tr|H9GLG5|H9GLG5 ANOCA
                                                                                                                                        879
                                                GFVENSYLAGLTPTEFFFHAMGGREGLIDTAVKTAETGYIORRLIKSMESVMVKYDATVR
tr|H2R1J6|H2R1J6 PANTR
                                                                                                                                        880
tr|G1MCZ1|G1MCZ1 AILME
                                                GFVENSYLAGLTPTEFFFHAMGGREGLIDTAVKTAETGYIQRRLIKSMESVMVKYDATVR
                                                                                                                                        880
tr|008847|008847 MOUSE
                                                GFVENSYLAGLTPTEFFFHAMGGREGLIDTAVKTAETGYIQRRLIKSMESVMVKYDATVR
                                                                                                                                        880
                                                GFVENSYLAGLTPTEFFFHAMGGREGLIDTAVKTAETGYIQRRLIKSMESVMVKYDATVR
tr|S7PWZ6|S7PWZ6 MYOBR
                                                                                                                                        880
                                                GFVENSYLAGLTPTEFFFHAMGGREGLIDTAVKTAETGYIQRRLIKSMESVMVKYDATVR
tr|D4A5A6|D4A5A6 RAT
                                                GFVENSYLAGLTPTEFFFHAMGGREGLIDTAVKTAETGYIQRRLIKSMESVMVKYDATVR
sp | P08775 | RPB1 MOUSE
                                                                                                                                        880
                                                GFVENSYLAGLTPTEFFFHAMSGREGLIDTAVKTAETGYIQRRLIKSMESVMVKYDATVR
sp|P24928|RPB1 HUMAN
                                                                                                                                        880
tr|AOA1S3EWL2|AOA1S3EWL2 DIPOR
                                                GFVENSYLAGLTPTEFFFHAMGGREGLIDTAVKTAETGYIQRRLIKSMESVMVKYDATVR
                                                                                                                                        880
                                                GFVENSYLAGLTPTEFFFHAMGGREGLIDTAVKTAETGYIQRRLIKSMESVMVKYDATVR
sp|P11414|RPB1_CRIGR
                                                                                                                                        880
tr|035559|035559_CRIGR
                                                GFVENSYLACITPTEFFFHAMGGREGLIDTAVKTAETGYIQRRLIKSMESVMVKYDATVR
                                                                                                                                        880
tr|A0A2I3M9H2|A0A2I3M9H2 PAPAN
                                                GFVENSYLAGLTPTEFFFHAMGGREGLIDTAVKTAETGYIQRRLIKSMESVMVKYDATVR
                                                                                                                                        879
tr|F7HB40|F7HB40 MACMU
                                                GFVENSYLAGLTPTEFFFHAMGGREGLIDTAVKTAETGYIQRRLIKSMESVMVKYDATVR
                                                                                                                                        880
tr|A0A2K6RYW9|A0A2K6RYW9 SAIBB
                                                GFVENSYLAGLTPTEFFFHAMGGREGLIDTAVKTAETGYIQRRLIKSMESVMVKYDATVR
                                                                                                                                        880
                                                GFVENSYLAGLTPTEFFFHAMGGREGLIDTAVKTAETGYIORRLIKSMESVMVKYDATVR
tr|W5N8Z6|W5N8Z6 LEPOC
                                                                                                                                        879
tr|I3JRW6|I3JRW6 ORENI
                                                GFVENSYLAGLTPTEFFFHAMGGREGLIDTAVKTAETGYIQRRLIKSMESVMVKYDGTVR
                                                                                                                                        878
tr|A0A0R4IMS9|A0A0R4IMS9 DANRE
                                                GFVENSYLACLTPTEFFFHAMGGREGLIDTAVKTAETGYIORRLIKSMESVMVKYDATVR
                                                                                                                                        877
                                                GFVENSYLACLTPTEFFFHAMGGREGLIDTAVKTAETGYIQRRLIKSMESVMVKYDATVR
tr|A0A1A7X327|A0A1A7X327_9TELE
                                                                                                                                        878
tr|A0A1A8UKD7|A0A1A8UKD7 NOTFU
                                                gfvensylagltptefffhamggreglidtavktaetgyiqrrliksmesvmvkydatvr
                                                                                                                                        878
tr|A0A1A8ER05|A0A1A8ER05 9TELE
                                                GFVENSYLAGLTPTEFFFHAMGGREGLIDTAVKTAETGYIQRRLIKSMESVMVKYDATVR
                                                                                                                                        878
                                                GFVENSYLAGLTPTEFFFHAMSGREGLIDTAVKTAETGYIQRRLIKSMESVMVKYDATVR
tr|A0A1A8DQ60|A0A1A8DQ60 9TELE
                                                                                                                                        878
                                                GFVENSYLAGLTPTEFFFHAMGGREGLIDTAVKTAETGYIQRRLIKSMESVMVKYDATVR
tr|A0A1A8NSR8|A0A1A8NSR8 9TELE
                                                                                                                                        878
                                                gfvensylagliptefffhamggreglidtavktaetgyiqrrliksmesvmvkydatvr
tr|A0A1W4YLM7|A0A1W4YLM7_9TELE
                                                                                                                                        878
tr|A0A1M8A6L7|A0A1M8A6L7_MALS4
                                                GFVENSYLRGLTAREFFFHAMAGREGLIDTAVKTAETGYIQRRLVKALEDVTICYDGTVR
                                                                                                                                        871
                                                GFVENSYLRGLTPOEFFFHAMAGREGLIDTAVKTAETGYIORRLVKALEDIMVHYDGTTR
tr|A0A1B2J8C6|A0A1B2J8C6 PICPA
                                                                                                                                        858
                                                GFVENSYLRGLTPGEFFFHAMAGREGLIDTAVKTAETGYIQRRLVKALEDIMVHYDGTTR
tr|F2QW17|F2QW17 KOMPC
                                                                                                                                        858
tr|A3GID7|A3GID7 PICST
                                               GFVENSYLRGLTPOEFFFHAMAGREGLIDTAVKTAETGYIORRLVKALEDIMVHYDGTTR
                                                                                                                                        857
tr|A0A1D8PUA6|A0A1D8PUA6_CANAL
                                               GFVENSYLRGLTPOEFFFHAMAGREGLIDTAVKTAETGYIQRRLVKALEDIMVHYDGTTR
                                                                                                                                        857
                                               GFVENSYLRGLTPOEFFFHAMAGREGLIDTAVKTAETGYIORRLVKALEDIMVHYDGTTR
tr|G8BEH9|G8BEH9 CANPC
                                                                                                                                        857
                                                                             *****************
```

```
tr|A0A1U8DYN0|A0A1U8DYN0 ALLSI
                                     NSINQVVQLTYGEDGLAGESVEFQNLATLKPSNKAFEKKFKFDYANERALRRTLQEEMVK
                                                                                                          948
tr|A0A1L8H4P4|A0A1L8H4P4 XENLA
                                     NS INQVVQLEYGEDGLAGEGVEFQNLATLKPSNKAFEKKFKFDYANERALRRTLQEDVVK
                                                                                                          938
tr|H9GLG5|H9GLG5 ANOCA
                                     NSINOVVOLRYGEDGLAGESVEFONLATLKPSNKAFEKKFKFDYTNERALRRTLGEEMVK
                                                                                                          939
                                     NSINQVVQLEYGEDGLAGESVEFONLATLKPSNKAFEKKFRFDYTNERALRRTLOEDLVK
tr|H2R1J6|H2R1J6_PANTR
                                                                                                          940
tr|G1MCZ1|G1MCZ1 AILME
                                     NSINQVVQLTYGEDGLAGESVEFQNLATLKPSNKAFEKKFRFDYTNERALRRTLQEDLVK
                                                                                                          940
tr|008847|008847 MOUSE
                                     NSINOVVOLHYGEDGLAGESVEFONLATLKPSNKAFEKKFRFDYTNERALRRTLOEDLVK
                                                                                                          940
tr|S7PWZ6|S7PWZ6 MYOBR
                                     NSINQVVQLRYGEDGLAGESVEFQNLATLKPSNKAFEKKFRFDYTNERALRRTLQEDLVK
                                                                                                          940
                                     NSINOVVOLRYGEDGLAGESVEFONLATLKPSNKAFEKKFRFDYTNERALRRTLOEDLVK
tr|D4A5A6|D4A5A6 RAT
                                                                                                          940
sp|P08775|RPB1 MOUSE
                                     NSINQVVQLEYGEDGLAGESVEFQNLATLKPSNKAFEKKFRFDYTNERALRRTLQEDLVK
                                                                                                          940
                                     NSINQVVQLHYGEDGLAGESVEFQNLATLKPSNKAFEKKFRFDYTNERALRRTLQEDLVK
sp | P24928 | RPB1 HUMAN
                                                                                                          940
                                     NSINQVVQLRYGEDGLAGESVEFQNLATLKPSNKAFEKKFRFDYTNERALRRTLQEDLVK
tr|A0A1S3EWL2|A0A1S3EWL2 DIPOR
                                                                                                          940
                                     NSINQVVQLTYGEDGLAGESVEFQNLATLKPSNKAFEKKFRFDYTNERALRRTLQEDLVK
sp | P11414 | RPB1 CRIGR
                                                                                                          940
tr|035559|035559 CRIGR
                                     NSINOVVOLHYGEDGLAGESVEFONLATLKPSNKAFEKKFRFDYTNERALRRTLOEDLVK
                                                                                                          940
                                     NSINQVVQLEYGEDGLAGESVEFQNLATLKPSNKAFEKKFRFDYTNERALRRTLQEDLVK
tr|A0A2I3M9H2|A0A2I3M9H2 PAPAN
                                     NSINQVVOLEYGEDGLAGESVEFONLATLKPSNKAFEKKFRFDYTNERALRRTLQEDLVK
NSINQVVOLEYGEDGLAGESVEFONLATLKPSNKAFEKKFRFDYTNERALRRTLQEDLVK
tr|F7HB40|F7HB40 MACMU
                                                                                                          940
tr|A0A2K6RYW9|A0A2K6RYW9 SAIBB
                                                                                                          940
tr|W5N8Z6|W5N8Z6 LEPOC
                                     NSINQVVQLHYGEDGLAGEAVEFQNMATLKPSNKAFEKKFKFDYTNERALRRTLQEDVVK
                                                                                                          939
tr|I3JRW6|I3JRW6 ORENI
                                     NSINQVVQLEYGEDGLAGENVEFQNLATVKPSHKAFEKKFKFDCTNERALRRTLQEDVVK
                                                                                                          938
tr|A0A0R4IMS9|A0A0R4IMS9 DANRE
                                     NSINOVVOLRYGEDGLAGEAVEFONMATLKPSHKAFEKKFKFDYANERALRRTLOEDVVK
                                                                                                          937
tr|A0A1A7X327|A0A1A7X327_9TELE
                                     NSINQVVQLEYGEDGLAGENVEFQNLATLKPSHKAFEKKFKFDCTNERALRRILQEDVVK
                                                                                                          938
tr|A0A1A8UKD7|A0A1A8UKD7 NOTFU
                                     NSINQVVQLRYGEDGLAGENVEFQNLATLKPSHKAFEKKFKFDCTNERALRRILQEDVVK
                                                                                                          938
tr|A0A1A8ER05|A0A1A8ER05 9TELE
                                     NSINOVVOLEYGEDGLAGENVEFONLATLKPSHKAFEKKFKFDCTNERALRRILQEDVVK
                                     NSINOVVOLTYGEDGLAGENVEFONLATLKPSHKAFEKKFKFDCTNERALRRILQEDVVK
tr|A0A1A8DQ60|A0A1A8DQ60 9TELE
                                                                                                          938
                                     NSINQVVQLEYGEDGLAGENVEFQNLATLKPSHKAFEKKFKFDCTNERALRRILQEDVVK
tr|A0A1A8NSR8|A0A1A8NSR8 9TELE
                                                                                                          938
tr|A0A1W4YLM7|A0A1W4YLM7 9TELE
                                     NSINQVVQLEYGEDGLAGEAVEFQNMATLKPSNKAFEKKFKFDYTNERALRRTLQEDVVK
                                                                                                          938
tr|A0A1M8A6L7|A0A1M8A6L7 MALS4
                                     NSTNNVIEFAYGEDGIDGAMVERQKLITHGLNDKEFRRRFKVDLSHGGF--KKGTLRAGL
                                                                                                          929
                                     NSLGDIIQFIYGEDGLDGTQVERQTIDTIPGSDKAFHKRYYVDLMDEKNSIKADVIEYAA
tr|A0A1B2J8C6|A0A1B2J8C6 PICPA
                                                                                                          918
tr|F2QW17|F2QW17 KOMPC
                                     NSLGDIIQFIYGEDGLDGTQVERQTIDTIPGSDKAFHKRYYVDLMDEKNSIKPDVIEYAA
                                                                                                          918
tr|A3GID7|A3GID7 PICST
                                     NSLGDIIQFVYGEDGIDGTQVEKQSVDTIPGSNDSFERRFRIDVLDSSKSIPESLLESGK
                                                                                                          917
tr|A0A1D8PUA6|A0A1D8PUA6 CANAL
                                     NSLGDIIQFTYGEDGIDGTQVEKQSVDTIPGSDESFERRYKIDVLDTENVISESLLESGK
                                                                                                          917
tr|G8BEH9|G8BEH9_CANPC
                                     NSLGDIIQFTYGEDGIDATQVEKQSVDTIPGSDASFERRYRIDLLEKNGSISESLLESGK
                                                                                                          917
                                     ** .::::: ****:
                                                       . :* *.: * .. *.::: .*
 tr|A0A1U8DYN0|A0A1U8DYN0_ALLSI
                                      DILSNAHIQNELERE FEKMREDREVLRV-IFPTGDSKVVLPCNLLRMINNAQKIFHVNAR
                                                                                                          1007
 tr|A0A1L8H4P4|A0A1L8H4P4 XENLA
                                      EILSDAHVQNELEKEFEKMKEDREVLRV-IFPTGDSKVVLPCNLQRMINNAQKIFHINTR
 tr|H9GLG5|H9GLG5 ANOCA
                                      DILSNAHIQNELERE FEKMKEDREVLRV-IFPTGDSKVVLPCNLLRMIWNAOKI FHINTR
                                                                                                          998
 tr|H2R1J6|H2R1J6 PANTR
                                      DVLSNAHIQNELERE FERMREDREVLRV-IFPTGDSKVVLPCNLLRMIWNAQKIFHINPR
                                                                                                          999
 tr|G1MCZ1|G1MCZ1_AILME
                                      DVLSNAHIQNELEREFERMREDREVLRV-IFPTGDSKVVLPCNLLRMIWNAQKIFHINPR
                                                                                                          999
 tr|008847|008847 MOUSE
                                      DVLSNAHIQNELEREFERMREDREVLRV-IFPTGDSKVVLPCNLLRMIWNAQKIFHINPR
                                                                                                          999
 tr|S7PWZ6|S7PWZ6 MYOBR
                                      DVLSNAHIONELERE FERMREDREVLRV-IFPTGDSKVVLPCNLLRMIWNAOKI FHINPR
                                                                                                          999
 tr|D4A5A6|D4A5A6 RAT
                                      DVI.SNAHTONET.EREFERMREDREVI.RV-TFPTGDSKVVI.PCNI.IRMTWNAGKTFHTNPR
                                                                                                          999
 sp|P08775|RPB1_MOUSE
                                      DVLSNAHIQNELEREFERMREDREVLRV-IFPTGDSKVVLPCNLLRMIWNAQKIFHINPR
                                                                                                          999
 sp|P24928|RPB1 HUMAN
                                      DVLSNAHIQNELERE FERMREDREVLRV-IFPTGDSKVVLPCNLLRMINNAQKIFHINPR
 tr|AOA1S3EWL2|AOA1S3EWL2 DIPOR
                                      DVLSNAHIQNELEREFERMREDREVLRV-IFPTGDSKVVLPCNLLRMIWNAQKIFHINPR
                                                                                                          999
                                      DVLSNAHIQNELERE FERMREDREVLRV-IFPTGDSKVVLPCNLLRMIWNAQKIFHINPR
 sp | P11414 | RPB1 CRIGR
                                                                                                          999
 tr|035559|035559 CRIGR
                                      DVLSNAHIQNELEREFERMREDREVLRV-IFPTGDSKVVLPCNLLRMIWNAQKIFHINPR
                                                                                                          999
 tr|A0A2I3M9H2|A0A2I3M9H2_PAPAN
                                      DVLSNAHIQNELEREFERMREDREVLRV-IFPTGDSKVVLPCNLLRMIWNAQKIFHINPR
                                                                                                          998
 tr|F7HB40|F7HB40 MACMU
                                      DVLSNAHIQNELERE FERMREDREVLRV-IFPTGDSKVVLPCNLLRMIWNAQKIFHINPR
 tr|A0A2K6RYW9|A0A2K6RYW9 SAIBB
                                      DVLSNAHIQNELERE FERMREDREVLRV-IFPTGDSKVVLPCNLLRMIWNAQKIFHINPR
                                                                                                          999
 tr|W5N8Z6|W5N8Z6 LEPOC
                                      EVLTNAHVQGELERE FEKMREDREILRA-IFPTGDSKVVLPCNLARMIWNAOKIFRINTR
                                                                                                          998
 tr|I3JRW6|I3JRW6 ORENI
                                      DVMTNAHVQGTLEREFDKMKEDREILRA-IFPTGDSKVVLPCNLARMIWNAQKIFRINPR
                                                                                                          997
 tr|A0A0R4IMS9|A0A0R4IMS9 DANRE
                                      DVMTNAHVQSSLEREFEKMREDREILRA-IFPTGDSKVVLPCNLARMIWNAQKIFRINPR
                                                                                                          996
 tr|A0A1A7X327|A0A1A7X327 9TELE
                                      DVQTNALVQSTLEREFEKMKEDREILRA-IFPTGDSKVVLPCNLARMIWNAQKIFRINPR
                                                                                                          997
                                      DVOTNAHVOSVLEREFEKMKEDREILRA-IFPTGDSKVVLPCNLARMIWNAOKIFRINPR
 tr|A0A1A8UKD7|A0A1A8UKD7 NOTFU
                                                                                                          997
 tr|A0A1A8ER05|A0A1A8ER05 9TELE
                                      DVQTNAHVQSVLEREFEKMKEDREILRA-IFPTGDSKVVLPCNLARMIWNAQKIFRINPR
                                                                                                          997
 tr|A0A1A8DQ60|A0A1A8DQ60 9TELE
                                      DVQTNAHVQSVLEREFEKMKEDREILRA-IFPTGDSKVVLPCNLARMIWNAQKIFRINPR
                                                                                                          997
 tr|A0A1A8NSR8|A0A1A8NSR8 9TELE
                                      DVQTNAHVQSVLEREFEKMKEDREILRA-IFPTGDSKVVLPCNLARMINNAQKIFRINPR
                                                                                                          997
 tr|A0A1W4YLM7|A0A1W4YLM7 9TELE
                                      DVMTNAHVOSALEHE FEKMREDREILRA-I FPTGDSKVVLPCNLARMIWNAOKI FRINPR
                                                                                                          997
                                      G-DWSPELEOLLDEEFEQLERDRMLLRTEIFRTDRVDTYLPLNIARLVLNAQQIFHIDPR
 tr|A0A1M8A6L7|A0A1M8A6L7 MALS4
                                                                                                          988
 tr|A0A1B2J8C6|A0A1B2J8C6 PICPA
                                      DIIGDVELQKELNSEYEQLVNDRKFLREIVFVNGDHNWPLFVNLRRIIONAQQIFHLDRA
 tr|F2QW17|F2QW17_KOMPC
tr|A3GID7|A3GID7 PICST
                                      DILGDVELQKELNSEYEQLVSDRKFLREIVFVNGDHNWPLPVNLRRIIQNAQQIFHLDRA
                                                                                                          978
                                      EIKGDVKLQKVLDEEYKQLLDDRKYLREVCFPNGDFSWPLPVNLRRIIDNAQQIFHNGRY
                                                                                                          977
 tr|A0A1D8PUA6|A0A1D8PUA6_CANAL
                                      EIRGDVQLQKILDEEYNQLLKDRKYLREVCFPNGDFSWPLPVNLRRIIQNAQQIFHNGRY
                                                                                                          977
 tr|G8BEH9|G8BEH9 CANPC
                                      EIQGDVKLQKLLDEEYNQLLKDRRYLRDVCFPNGDFSWPLPVNLRRIIQNAQQIFHNGRY
                                                                                                          977
                                          . .. *: *:.:: ** ** * . . ** *: *:: ***:
```

```
tr|A0A1U8DYN0|A0A1U8DYN0_ALLSI
                                         LSAQAFDWL<mark>LG</mark>EIESKFNQATAHPGEMGGALAAQS<mark>LGE</mark>PATQMTLNTFHYAGVSAKNVTI
                                                                                                                     1127
                                                    <mark>g</mark>eteskfnoatahpgemygalaaos<mark>lge</mark>patomtintfhyagysaknyti
tr|A0A1L8H4P4|A0A1L8H4P4 XENLA
                                         LSGEAFDWL
                                                                                                                     1117
                                                    GEIESKFNOATAHPGEMVGALAAOS<mark>LGE</mark>PATOMTLNTFHYAGVSAKNVTI
tr|H9GLG5|H9GLG5 ANOCA
                                         LSGEAFDWL
                                                                                                                     1118
                                        LSGEAFDWLIGEIESKFNQA AHPGEMUGALAAQS<mark>LG</mark>PATQMTLNTFHYAGVSAKNVT
LSGEAFDWLIGEIESKFNQA AHPGEMUGALAAQS<mark>LGE</mark>PATQMTLNTFHYAGVSAKNVT
tr|H2R1J6|H2R1J6 PANTR
                                                                                                                     1119
tr|G1MCZ1|G1MCZ1_AILME
                                                                                                                     1119
tr|008847|008847_MOUSE
                                         LSGEAFDWL<mark>LG</mark>EIESKFNQATAHPGEMVGALAAQSI<mark>LGE</mark>PATQMTLNTFHYAGVSAKNVTI
                                                                                                                     1119
                                                    <mark>lg</mark>e iesk fnoa tahpgemygalaaosi<mark>lge</mark>patomtint fhyagysaknyti
tr|S7PWZ6|S7PWZ6 MYOBR
                                         LSGEAFDWL
                                                                                                                     1119
                                         LSGEAFDWLLGEIESKFNQATAHPGEMVGALAAQSLGEPATQMTLNTFHYAGVSAKNVT
tr|D4A5A6|D4A5A6 RAT
                                                                                                                     1119
                                         LSGEAFDWLLGEIESKFNQATAHPGEMVGALAAOSLGEPATOMTLNTFHYAGVSAKNVT
sp|P08775|RPB1 MOUSE
                                                                                                                     1119
                                                    <mark>lg</mark>e i esk fnoa tah pgemvga laags<mark>lge</mark> patomt int fhyagvsaknvti
sp|P24928|RPB1 HUMAN
                                         LSCEAFDWI
                                                                                                                     1119
tr|A0A1S3EWL2|A0A1S3EWL2 DIPOR
                                         LSGEAFDWL<mark>LG</mark>EIESKFNQATAHPGEMVGALAAQS<mark>LGE</mark>PATQMTLNTFHYAGVSAKNVTI
                                                                                                                     1119
sp|P11414|RPB1 CRIGR
                                         LSGEAFDWL<mark>LG</mark>EIESKFNQATAHPGEMVGALAAQSI<mark>GE</mark>PATQMTLNTFHYAGVSAKNVTI
                                                                                                                     1119
                                         LSGEAFDWL<mark>LG</mark>EIESKFNQATAHPGEMVGALAAQS<mark>LGE</mark>PATQMTLNTFHYAGVSAKNVTI
tr|035559|035559 CRIGR
                                                                                                                     1119
                                                    <mark>lge</mark>ieskfnoa tahpgemygalaags <mark>ge</mark>patomtint fhyagysaknyt.
Lgeieskfnoa tahpgemygalaags gepatomtint fhyagysaknyt.
tr|A0A2I3M9H2|A0A2I3M9H2_PAPAN
                                         LSGEAFDWLI
                                                                                                                     1118
tr|F7HB40|F7HB40 MACMU
                                         LSGEAFDWL
                                                                                                                     1119
                                         LSGEAFOWL<mark>LG</mark>EIESKFNQATAHPGEMVGALAAQSI<mark>LGE</mark>PATQMTLNTFHYAGVSAKNVTI
tr|A0A2K6RYW9|A0A2K6RYW9 SAIBB
                                                                                                                     1119
tr|W5N8Z6|W5N8Z6 LEPOC
                                         LSTEAYEWL<mark>LG</mark>EIETKFNQS AHPGEMVGALAAQSI<mark>GE</mark>PATQMTLNTFHYAGVSAKNVTI
                                                                                                                     1118
tr|I3JRW6|I3JRW6 ORENI
                                         LSTEAFDWL<mark>LG</mark>EIETKFNQS AHPGEMVGALAAQSI<mark>LGE</mark>PATQMTLNTFHYAGVSAKNVTI
                                                                                                                     1117
                                                    <mark>g</mark>e ietkfnos iahpgemygalaaos <mark>lge</mark>patomtint fhyagysaknyti
tr|A0A0R4IMS9|A0A0R4IMS9 DANRE
                                         LSTEAFDWL
                                                                                                                     1116
                                                    GEIETKFNQS VHPGEMVGALAAQS<mark>LGE</mark>PATQMTLNTFHYAGVSAKNVT
tr|A0A1A7X327|A0A1A7X327 9TELE
                                         LSTEAFDWL
                                                                                                                     1117
tr|A0A1A8UKD7|A0A1A8UKD7 NOTFU
                                                    GEIETKFNOS VHPGEMVGALAAQS GEPATOMTLNTFHYAGVSAKNVT
                                         LSMEAFDWL
                                                                                                                     1117
tr|A0A1A8ER05|A0A1A8ER05 9TELE
                                         LSMEAFDWL
                                                    <mark>g</mark>e i etkfnos tvhpgemvgalaaosl<mark>ge</mark>patomtint fhyagvsaknvti
                                                                                                                     1117
tr|A0A1A8DQ60|A0A1A8DQ60 9TELE
                                         LSMEAFDWL<mark>LG</mark>EIETKFNQSTVHPGEMVGALAAQST<mark>LGE</mark>PATQMTLNTFHYAGVSAKNVTI
                                                                                                                     1117
tr|A0A1A8NSR8|A0A1A8NSR8 9TELE
                                         LSMEAFDWL<mark>LG</mark>EIETKFNQSTVHPGEMVGALAAQSTGEPATQMTLNTFHYAGVSAKNVTI
                                                                                                                     1117
                                         LSTEAFEWL<mark>LC</mark>EIETKFNQATVHPGEMVGALAAQSI<mark>.GE</mark>PATQMTLNTFHYAGVSAKNVTI
tr|A0A1W4YLM7|A0A1W4YLM7 9TELE
                                                                                                                     1117
                                         LSREAWEWILGEIEGOFARSVAOPGEMOGILAAOSIGEPATOMTINT FHYAGVSSKNVTI
tr|A0A1M8A6L7|A0A1M8A6L7 MALS4
                                                                                                                     1108
                                         LNRDAFEWV<mark>LG</mark>TIEAQFQRSIVHPGEMVGVIAAQSIGEPATQMTLNTFHYAGVSSKNVTI
tr|A0A1B2J8C6|A0A1B2J8C6 PICPA
                                                                                                                     1098
tr|F2QW17|F2QW17 KOMPC
                                         LNRDAFEWV<mark>LG</mark>TIEAQFQRSLVHPGEMVGVIAAQSIGEPATQMTLNTFHYAGVSSKNVTI
                                                                                                                     1098
tr|A3GID7|A3GID7 PICST
                                         LNRSSFEWV<mark>VG</mark>EIETQFQKSTVHPGEMVGVIAAQSTGEPATQMTLNTFHYAGVSSKNVTI
                                                                                                                     1097
tr|A0A1D8PUA6|A0A1D8PUA6 CANAL
                                         LNRASFEWV
                                                    <mark>g</mark>eietofokstvhpgemvgvvaaost
                                                                                   EPATQMTLNT FHYAGVS SKNVTI
                                                                                                                     1097
                                         LNRSSFDWVLGEIETOFOKSTVHPGEMVGVVAAOSTGEPATOMTINTFHYAGVSSKNVT
tr|G8BEH9|G8BEH9 CANPC
                                                                                                                     1097
                                                       ** :* :::::****
                                         *. :::*:
                                         GVPRLKELINISKKEKTPSLTVFLLGQSARDAERAKDILCRLEHTTLRKVTANTALYYDP
tr|A0A1U8DYN0|A0A1U8DYN0 ALLSI
                                                                                                                     1187
                                         GVPRLKELINISKKEKTPSLTVFLLGQSARDAERAKDILCRLEHTTLRKVTANTALYYDF
tr|A0A1L8H4P4|A0A1L8H4P4 XENLA
                                                                                                                     1177
                                         GVPRLKELINISKKEKTPSLTVFLLGQSARDAERAKDILCRLEHTTLRKVTANTAIYYDP
                                                                                                                     1178
tr|H9GLG5|H9GLG5 ANOCA
tr|H2R1J6|H2R1J6 PANTR
                                         GVPRLKELINISKKEKTPSLTVFLLGQSARDAERAKDILCRLEHTTLRKVTANTALYYDP
                                                                                                                     1179
tr|G1MCZ1|G1MCZ1 AILME
                                         GVPRLKELINISKKEKTPSLTVFLLGQSARDAERAKDILCRLEHTTLRKVTANTAIYYDP
                                                                                                                     1179
tr|008847|008847 MOUSE
                                         GVPRLKELINISKKEKTPSLTVFLLGQSARDAERAKDILCRLEHTTLRKVTANTALYYDP
                                                                                                                     1179
tr|S7PWZ6|S7PWZ6 MYOBR
                                         GVPRLKELINESKKEKTPSLTVFLLGOSARDAERAKDILCRLEHTTLRKVTANTALYYDP
                                                                                                                     1179
tr|D4A5A6|D4A5A6 RAT
                                         GVPRLKELINISKKEKTPSLTVFLLGQSARDAERAKDILCRLEHTTLRKVTANTAIYYDP
                                                                                                                     1179
sp|P08775|RPB1 MOUSE
                                         GVPRLKELINISKKIKTPSLTVFLLGQSARDAERAKDILCRLEHTTLRKVTANTALYYDP
                                                                                                                     1179
sp | P24928 | RPB1 HUMAN
                                         GVPRLKELINISKKEKTPSLTVFLLGQSARDAERAKDILCRLEHTTLRKVTANTAIYYDP
                                                                                                                     1179
                                         GVPRLKELINISKKEKTPSLTVFLLGQSARDAERAKDILCRLEHTTLRKVTANTALYYDP
tr|A0A1S3EWL2|A0A1S3EWL2_DIPOR
                                                                                                                     1179
sp|P11414|RPB1 CRIGR
                                         GVPRLKELINESKKEKTPSLTVFLLGQSARDAERAKDILCRLEHTTLRKVTANTALYYDP
                                                                                                                     1179
tr|035559|035559 CRIGR
                                         GVPRLKELINISKKEKTPSLTVFLLGQSARDAERAKDILCRLEHTTLRKVTANTALYYDP
                                                                                                                     1179
                                         GVPRLKELINISKKEKTPSLTVFLLGQSARDAERAKDILCRLEHTTLRKVTANTALYYDP
tr|A0A2I3M9H2|A0A2I3M9H2_PAPAN
                                                                                                                     1178
tr|F7HB40|F7HB40 MACMU
                                         GVPRLKELINESKKEKTPSLTVFLLGQSARDAERAKDILCRLEHTTLRKVTANTALYYDP
                                                                                                                     1179
tr|A0A2K6RYW9|A0A2K6RYW9 SAIBB
                                         GVPRLKELINISKKEKTPSLTVFLLGQSARDAERAKDILCRLEHTTLRKVTANTALYYDP
                                                                                                                     1179
                                         GVPRLKELINISKRUKTPSLTVFLLGQAARDAERAKDILCRLEHTTLRKVTANTALYYDP
tr|W5N8Z6|W5N8Z6 LEPOC
                                                                                                                     1178
tr|I3JRW6|I3JRW6 ORENI
                                         GVPRLKELINISKREKTPSLTVFLLGQAARDAERAKDILCRLEHTTLRKVTANTALYYDP
                                                                                                                     1177
tr|A0A0R4IMS9|A0A0R4IMS9 DANRE
                                         GVPRLKELINISKREKTPSLTVFLLGQAARDAERAKDILCRLEHTTLRKVTANTALYYDF
                                                                                                                     1176
                                         GVPRLKELINISKREKTPSLTVFLLGQAARDAERAKDILCRLEHTTLRKVTANTATYYDP
tr|A0A1A7X327|A0A1A7X327 9TELE
                                                                                                                     1177
                                         GVPRLKELINISKREKTPSLTVFLLGQAARDAERAKDILCRLEHTTLRKVTANTAIYYDP
tr|A0A1A8UKD7|A0A1A8UKD7 NOTFU
                                                                                                                     1177
tr|A0A1A8ER05|A0A1A8ER05 9TELE
                                         GVPRLKELINISKRIKTPSLTVFLLGQAARDAERAKDILCRLEHTTLRKVTANTAIYYDF
                                                                                                                     1177
                                         GVPRLKELINISKRUKTPSLTVFLLGQAARDAERAKDILCRLEHTTLRKVTANTALYYDP
tr|A0A1A8DQ60|A0A1A8DQ60 9TELE
                                                                                                                     1177
tr|A0A1A8NSR8|A0A1A8NSR8 9TELE
                                         GVPRLKELINISKREKTPSLTVFLLGQAARDAERAKDILCRLEHTTLRKVTANTALYYDP
                                                                                                                     1177
                                         GVPRLKELINISKREKTPSLTVFLLGQAARDAERAKDILCRLEHTTLRKVTANTALYYDP
tr|AOA1W4YLM7|AOA1W4YLM7 9TELE
                                                                                                                     1177
                                         GVPRLKEIINCAENIKTPSVTVYLHPKYSASSESAKVIQTALAYTTLQTVTSAVEVFYDP
tr|A0A1M8A6L7|A0A1M8A6L7_MALS4
                                                                                                                     1168
tr|A0A1B2J8C6|A0A1B2J8C6 PICPA
                                         GVPRLKEILNVAKNIKTPALTVYLDREIALDIEKAKVIQSSIEYTTLKNVTSATEIYYDF
                                                                                                                     1158
tr|F2QW17|F2QW17_KOMPC
                                         GVPRLKEILNVAKNIKTPALTVYLDREIALDIEKAKVIQSSIEYTTLKNVTSATELYYDP
                                                                                                                     1158
tr|A3GID7|A3GID7 PICST
                                         GVPRLKEILN/VAKNIKTPALTVYLDPALSDDIEKAKVVQSAIEHTSLKNVTSSTELYYDP
                                                                                                                     1157
                                         GVPRLKEILNVAKNIKTPAMTVFLDPEVASDIEKAKIVQSAIEHTTLKNVTSSTEIYYDF
tr|A0A1D8PUA6|A0A1D8PUA6 CANAL
                                                                                                                     1157
                                         GVPRLKEILMVAKNIKTPALTVYLDPEVASDIELAKIVQSSIEHTTLKNVTSSTETYYDP
tr|G8BEH9|G8BEH9 CANPC
                                                                                                                     1157
                                          *******::* ::. | ***::**:*
                                                                        : . * ** : : :*:*:.**
```

```
SPQSTVVAEDQEWVNVYYEMPDFDV----SRISPWLLRVELDRKHMTDRKLTMEQIAEKI
NPQNTVVAEDQEWVNVYYEMPDFDV----TRISPWLLRVELDRKHMTDRKLTMEQIAEKI
NPQNTVVAEDQEWVNVYYEMPDFDV----TRISPWLLRVELDRKHMTDRKLTMEQIAEKI
NPQSTVVAEDQEWVNVYYEMPDFDV----ARISPWLLRVELDRKHMTDRKLTMEQIAEKI
tr|A0A1U8DYN0|A0A1U8DYN0 ALLSI
                                                                                                                       1243
tr|A0A1L8H4P4|A0A1L8H4P4_XENLA
                                                                                                                       1233
tr|H9GLG5|H9GLG5 ANOCA
                                                                                                                       1234
tr|H2R1J6|H2R1J6 PANTR
                                                                                                                       1235
                                         NPQSTVVAEDQEWVNVYYEMPDFDV----ARTSPWLLRVELDRKHMTDRKLTMEQIAEKI
tr|G1MCZ1|G1MCZ1 AILME
                                                                                                                       1235
                                         NPQSTVVAEDQEWVNVYYEMPDFDV----ARI SPWLLRVELDRKHNTDRKLINEQIAEKI
NPQSTVVAEDQEWVNVYYEMPDFDV----ARI SPWLLRVELDRKHNTDRKLINEQIAEKI
tr|008847|008847 MOUSE
                                                                                                                       1235
tr|S7PWZ6|S7PWZ6 MYOBR
                                                                                                                       1235
                                         NPQSTVVAEDQEWVNVYYEMPDFDV----ARISPWLLRVELDRKHMTDRKLTMEQIAEKI
tr|D4A5A6|D4A5A6 RAT
                                                                                                                       1235
                                         NPQSTVVAEDQEWVNVYYEMPDFDV----ARI SPWLLRVELDRKHNTDRKLIMEQIAEKI
NPQSTVVAEDQEWVNVYYEMPDFDV----ARI SPWLLRVELDRKHNTDRKLIMEQIAEKI
sp|P08775|RPB1 MOUSE
                                                                                                                       1235
sp|P24928|RPB1 HUMAN
                                                                                                                       1235
                                         NPQSTVVAEDQEWVNVYYEMPDFDV----ARISPWLLRVELDRKHMTDRKLTMEQIAEKI
tr|A0A1S3EWL2|A0A1S3EWL2 DIPOR
                                                                                                                       1235
                                         NPQSTVVAEDQEWVNVYYEMPDFDV----ARI SPWLLRVELDRKHMTDRKLTMEQIAEKI
NPQSTVVAEDQEWVNVYYEMPDFDV----ARI SPWLLRVELDRKHMTDRKLTMEQIAEKI
sp|P11414|RPB1 CRIGR
                                                                                                                       1235
tr|035559|035559 CRIGR
                                                                                                                       1235
                                         NPQSTVVAEDQEWVNVYYEMPDFDV----ARISPWLLRVELDRKHMTDRKLTMEQIAEKI
tr|A0A2I3M9H2|A0A2I3M9H2_PAPAN
                                                                                                                       1234
                                         NPQSTVVAEDQEWVNVYYEMPDFDV----ARI SPWLLRVELDRKHMTDRKLTMEQIAEKI
tr|F7HB40|F7HB40 MACMU
                                                                                                                       1235
                                         NPOSTVVAEDOEWVNVYYEMPDFDV----ARI SPWLLRVELDRKHNTDRKLIMEQIAEKI
NPONTVVAEDOEWVNVYYEMPDFDV----TRI SPWLLRIELDRKHNTDRKLIMEQIAEKI
tr|A0A2K6RYW9|A0A2K6RYW9 SAIBB
                                                                                                                       1235
tr|W5N8Z6|W5N8Z6 LEPOC
                                                                                                                       1234
                                         NPQNTVVAEDQEWVNVYYEMPDFDV----TRISPWLLRIELDRKHMTDRKLTMEQIAEKI
tr|I3JRW6|I3JRW6 ORENI
                                                                                                                       1233
                                         NPONTVVTEDQEWVNVYYEMPDFDV----TRISPWLLRIELDRKHMTDRKLTMEQIAEKI
NPONTVVAEDQEWVNVYYEMPDFDV----TRISPWLLRIELDRKHMTDRKLTMEQIAEKI
tr|A0A0R4IMS9|A0A0R4IMS9 DANRE
                                                                                                                       1232
tr|A0A1A7X327|A0A1A7X327_9TELE
                                                                                                                       1233
                                         NPQNTVVAEDQEWVNVYYEMPDFDV----TRISPWLLRIELDRKHMTDRKLIMEQIAEKI
tr|A0A1A8UKD7|A0A1A8UKD7 NOTFU
                                                                                                                       1233
                                         NPONTVVAEDQEWVNVYYEMPDFDV----TRISPWLLRIELDRKHNTDRKLIMEQIAEKI
NPONTVVAEDQEWVNVYYEMPDFDV----TRISPWLLRIELDRKHNTDRKLIMEQIAEKI
tr|A0A1A8ER05|A0A1A8ER05_9TELE
                                                                                                                       1233
tr|A0A1A8DQ60|A0A1A8DQ60 9TELE
                                                                                                                       1233
                                         NPQNTVVAEDQEWVNVYYEMPDFDV----TRISPWLLRIELDRKHMTDRKLTMEQIAEKI
tr|A0A1A8NSR8|A0A1A8NSR8 9TELE
                                                                                                                       1233
                                         NPONTVVTEDOEWVNVYYEMPDFDV----TRISPWLLRIELDRKHNTDRKLTMEQIAEKI
DPSSTVIPEDRDFVDAFFAIPDEEVEASLERGSPWLLRIVLDRAGMLDKNLTMAEVASKI
tr|A0A1W4YLM7|A0A1W4YLM7_9TELE
                                                                                                                       1233
tr|A0A1M8A6L7|A0A1M8A6L7 MALS4
                                                                                                                       1228
                                                  PEDERTIOLHESILDERAROSED
                                         DPTSTVIEEDFDTVEAYFSIPDEKVEETIDKOSPWLLRIELDRARMLDKOLTMNOVADKI
DPTSTVIEEDFDTVEAYFSIPDEKVEETIDKOSPWLLRIELDRARMLDKOLTMNOVADKI
DPRTTVIEEDYDTVEAYFSIPDEKVEESIEKOSPWLLRIELDRAKMLDKOLTMAQVAEKI
tr|A0A1B2J8C6|A0A1B2J8C6 PICPA
                                                                                                                       1218
tr|F2QW17|F2QW17 KOMPC
                                                                                                                       1218
tr|A3GID7|A3GID7 PICST
                                                                                                                       1217
tr|A0A1D8PUA6|A0A1D8PUA6 CANAL
                                         DPRTTVIEDDYDTVEAYFAIPDQKVEESIDKQSPWLLRIELDRAKMLDKQLTMAQVAEKI
                                                                                                                       1217
                                         DPRTTVIEEDYDTVEAYFAIPDQKVEESIEKQSPWLLRIELDRAKMLDKQLTMAQVAEKI
tr|G8BEH9|G8BEH9 CANPC
                                                                                                                       1217
                                                                               ......
                                                                                       ***
                                                                                             4 4<u>-</u> 444
                                         .* .**: :* : :: .: : * ..
                                                                            : |
tr|A0A1U8DYN0|A0A1U8DYN0 ALLSI
                                          MTLQGIEQISKVYMHLPQTDNKKKIIITEDGEFKALQEWILETDGVSLMRVLSEKDVDPV
tr|A0A1L8H4P4|A0A1L8H4P4 XENLA
                                          MTLQGIEQISKVYMHLPQTDNKKKIIITEDGEFKALQEWILETDGVSLMRVLSEKDVDPV
                                                                                                                       1353
                                          MTLQGIEQISKVYMHLPQTDNKKKIIITEDGEFKALQEWILETDGVSLMRVLSEKDVDPV
tr|H9GLG5|H9GLG5 ANOCA
                                                                                                                       1354
tr|H2R1J6|H2R1J6 PANTR
                                          MTLQGIEQISKVYMHLPQTDNKKKIIITEDGEFKALQEW LETDGVSIMRVLSEKDVDPV
                                                                                                                       1355
tr|G1MCZ1|G1MCZ1 AILME
                                          MTLQGIEQISKVYMHLPQTDNKKKIIITEDGEFKALQEWILETDGVSLMRVLSEKDVDPV
                                                                                                                       1355
                                          MTLQGIEQISKVYMHLPQTDNKKKIIITEDGEFKALQEWILETDGVSLMRVLSEKDVDPV
tr|008847|008847 MOUSE
                                                                                                                       1355
                                          MTLQGIEQISKVYMHLPQTDNKKKIIITEDGEFKALQEWILETDGVSLMRVLSEKDVDPV
tr|S7PWZ6|S7PWZ6 MYOBR
                                                                                                                       1355
                                          MTLQGIEQISKVYMHLPQTDNKKKIIITEDGEFKALQEWILETDG/SLMRVLSEKDVDPV
tr|D4A5A6|D4A5A6 RAT
                                                                                                                       1355
sp|P08775|RPB1 MOUSE
                                          MTLQGIEQISKVYMHLPQTDNKKKIIITEDGEFKALQEWILETDGVSLMRVLSEKDVDPV
                                                                                                                       1355
sp|P24928|RPB1 HUMAN
                                          MTLQGIEQISKVYMHLPQTDNKKKIIITEDGEFKALQEWILETDGVSLMRVLSEKDVDFV
                                                                                                                       1355
tr|A0A1S3EWL2|A0A1S3EWL2 DIPOR
                                          MTLQGIEQISKVYMHLPQTDNKKKIIITEDGEFKALQEWILETDGVSIMRVLSEKDVDPV
                                                                                                                       1355
                                          MTLQGIEQISKVYMHLPQTDNKKKIIITEDGEFKALQEWILETDGVSIMRVLSEKDVDFV
                                                                                                                       1355
sp|P11414|RPB1 CRIGR
tr|035559|035559 CRIGR
                                          MTLQGIEQISKVYMHLPQTDNKKKIIITEDGEFKALQEWILETDGVSLMRVLSEKDVDFV
                                                                                                                       1355
                                          MTLQGIEQISKVYMHLPQTDNKKKIIITEDGEFKALQEWILETDGVSLMRVLSEKDVDFV
tr|A0A2I3M9H2|A0A2I3M9H2 PAPAN
                                                                                                                       1354
tr|F7HB40|F7HB40 MACMU
                                          MTLQGIEQISKVYMHLPQTDNKKKIIITEDGEFKALQEWILETDGVSLMRVLSEKDVDPV
                                                                                                                       1355
tr|A0A2K6RYW9|A0A2K6RYW9 SAIBB
                                          MTLQGIEQISKVYMHLPQTDNKKKIIITEDGEFKALQEWILETDGVSLMRVLSEKDVDFV
                                                                                                                       1355
tr|W5N8Z6|W5N8Z6_LEPOC
                                          MTLQGIEQISKVYMHLPQTDNKKKIIITEDGEFKALQEWILETDGVGLMRVLSEKDVDPV
                                                                                                                       1354
tr||I3JRW6||I3JRW6 ORENI
                                          MTLQGIEQISKVYMHLPQTDNKKKIIITEDGEFKALQEWILETDGVSLMRVLSEKDVDFV
                                                                                                                       1353
                                          MTLQGIEQISKVYMHLPQTDNKKKIIITEDGEFKALQEWILETDGVSIMRVLSEKDVDPV
tr|A0A0R4IMS9|A0A0R4IMS9 DANRE
                                                                                                                       1352
tr|A0A1A7X327|A0A1A7X327_9TELE
                                          MTLQGIEQISKVYMHLPQTDNKKKIIITEDGEFKALQEWILETDGVSLMRVLSEKDVDPV
                                                                                                                       1353
tr|A0A1A8UKD7|A0A1A8UKD7 NOTFU
                                          MTLQGIEQISKVYMHLPQTDNKKKIIITEDGEFKALQEWILETDGVSIMRVLSEKDVDFV
                                                                                                                       1353
tr|A0A1A8ER05|A0A1A8ER05 9TELE
                                          MTLQGIEQISKVYMHLPQTDNKKKIIITEDGEFKALQEWILETDGVSIMRVLSEKDVDPV
                                                                                                                       1353
                                          MTLQGIEQISKVYMHLPQTDNKKKIIITEDGEFKALQEWILETDGVSLMRVLSEKDVDPV
tr|A0A1A8DQ60|A0A1A8DQ60 9TELE
                                                                                                                       1353
                                          MTLQGIEQISKVYMHLPQTDNKKKIIITEDGEFKALQEWILETDGVSIMRVLSEKDVDPV
tr|A0A1A8NSR8|A0A1A8NSR8 9TELE
                                                                                                                       1353
tr|A0A1W4YLM7|A0A1W4YLM7 9TELE
                                          MTLQGIEQISKVYMHLPQTDNKKKIIITEDGEFKALQEWILETDGVGIMRVLSEKDVDPV
                                                                                                                       1353
tr|A0A1M8A6L7|A0A1M8A6L7 MALS4
                                          IALKGVPGISKVFIVKQD--KSSRRFDPQTGEWDTIKEYVLETDGTNLKDVLAVDGVDVS
                                                                                                                       1339
                                          IALRGIPGISKVYMVKHK--VS---VPDESGEYKNEELWALETDGINLAEVMAVPGVDSS
tr|A0A1B2J8C6|A0A1B2J8C6 PICPA
                                                                                                                       1328
                                          IALRGIPGISKVYMVKHK--VS---VPDESGEYKNEELWALETDGINLAEVMAVPGVDSS
tr|F2QW17|F2QW17_KOMPC
                                                                                                                       1328
                                          ISLRGIPGITRVFMMQHK--VN---TPDATGEFKQGKEWVLETDGVNLADVMAVPGVDSS
tr|A3GID7|A3GID7 PICST
                                                                                                                       1327
                                          ISLRGIPGITRVFMMQHK--VS---HPDETGEFKQGKEWVLETDGVNLADVMAVPGVDST
tr|A0A1D8PUA6|A0A1D8PUA6 CANAL
                                                                                                                       1327
                                          ISLRGIPGITRVFMMOHK--VS---KPDETGEFKOGKEWVLETDGVNLADVMAVPGVDSA
tr|G8BEH9|G8BEH9 CANPC
                                                                                                                       1327
                                                                                         .. :: .:*
                                          ::*:*: * :* :
```

```
RTTSNDIVEIFTVLGIEAVRKALERELYHVISTDGSYVNYRHLALLCDTMTSRGHLMAIT
tr|A0A1U8DYN0|A0A1U8DYN0_ALLSI
                                                                                                                       1423
                                         RTTSNDIVEIFTVLGIRVRKALERELYHVISFDGSYVNYRHLALLCDIMTCRGHLMAIT
RTTSNDIVEIFTVLGIRVRKALERELYHVISFDGSYVNYRHLALLCDIMTCRGHLMAIT
tr|A0A1L8H4P4|A0A1L8H4P4_XENLA
                                                                                                                       1413
tr | H9GLG5 | H9GLG5_ANOCA
                                                                                                                       1414
tr|H2R1J6|H2R1J6_PANTR
                                          RTTSNDIVEIFTVLGIENVRKALERELYHVISTDGSYVNYRHLALLCDTMTCRGHLMAIT
                                                                                                                       1415
tr|G1MCZ1|G1MCZ1_AILME
tr|O08847|O08847_MOUSE
                                          RTTSNDIVEIFTVLGIEAVRKALERELYHVISFDGSYVNYRHLALLCDTMTCRGHLMAIT
                                                                                                                       1415
                                          RTTSNDIVEIFTVLGIENVRKALERELYHVISTDGSYVNYRHLALLCDTMTCRGHLMAIT
                                                                                                                       1415
tr|S7PWZ6|S7PWZ6 MYOBR
                                          RTTSNDIVEIFTVLGIEAVRKALERELYHVISTDGSYVNYRHLALLCDTMTCRGHLMAIT
                                                                                                                       1415
                                                            <mark>iea</mark>vrkalerelyhvis#dgsyvnyrhlallcd#mtcrghlmait
tr|D4A5A6|D4A5A6_RAT
                                          RTTSNDIVEIFTVLO
                                                                                                                        1415
                                          RTTSNDIVEIFTV
sp|P08775|RPB1 MOUSE
                                                            <mark>iea</mark>vrkalerelyhvistogsyvnyrhlallcotmtcrghlmait
                                                                                                                       1415
sp | P24928 | RPB1 HUMAN
                                                            EAVRKALERELYHVISFDGSYVNYRHLALLCDTMTCRGHLMAIT
                                          RTTSNDIVEIFTVL
                                                                                                                       1415
tr|A0A1S3EWL2|A0A1S3EWL2_DIPOR
                                          RTTSNDIVEIFTV
                                                            <mark>iea</mark>vrkalerelyhvisedgsyvnyrhlallodimtcrghlmait
                                                                                                                       1415
sp|P11414|RPB1 CRIGR
                                                            EAVRKALERELYHVISFDGSYVNYRHLALLCDIMICRGHLMAIT
                                          RTTSNDIVEIFTVL
                                                                                                                       1415
tr|035559|035559 CRIGR
                                          RTTSNDIVEIFTV
                                                            ILAVRKALERELYHVIS PDGSYVNYRHLALLCDIMTCRGHLMAIT
                                                                                                                       1415
tr|A0A2I3M9H2|A0A2I3M9H2_PAPAN
                                                            EAVRKALERELYHVISEDGSYVNYRHLALLCDTMTCRGHLMAIT
                                          RTTSNDIVEIFTVL
                                                                                                                        1414
tr|F7HB40|F7HB40_MACMU
                                          RTTSNDIVEIFTV
                                                            <mark>iea</mark>vrkalerelyhvis#dgsyvnyrhlallcd#mtcrghlmait
                                                                                                                       1415
tr|A0A2K6RYW9|A0A2K6RYW9_SAIBB
                                          RTTSNDIVEIFTVL
                                                            EAVRKALERELYHVISEDGSYVNYRHLALLCDIMICRGHLMAIT
                                                                                                                       1415
                                                           I PAVRKALERELYHVIS TOGSYVNYRHLALLCOTMTCRGHLMAIT
tr|W5N8Z6|W5N8Z6_LEPOC
                                          RTTSNDIVEIFTVLO
                                                                                                                        1414
tr|I3JRW6|I3JRW6 ORENI
                                          RTTSNDIVEIFTVLGIENVRKALERELYHVISTDGSYVNYRHLALLCDTMTCRGHLMAIT
                                                                                                                       1413
tr|A0A0R4IMS9|A0A0R4IMS9 DANRE
                                          RTTSNDIVEIFTVLGIEAVRKALERELYHVISTDGSYVNYRHLALLCDTMTCRGHLMAIT
                                                                                                                       1412
tr|A0A1A7X327|A0A1A7X327 9TELE
                                          RTTSNDIVEIFTVL
                                                            <mark>ika</mark>vrkalerelyhvisfdgsyvnyrhlallcdtmtcrghlmait
                                                                                                                       1413
tr|A0A1A8UKD7|A0A1A8UKD7 NOTFU
                                          RTTSNDIVEIFTVLGIEAVRKALERELYHVISTDGSYVNYRHLALLCDTMTCRGHLMAIT
                                                                                                                        1413
tr|A0A1A8ER05|A0A1A8ER05 9TELE
                                          RTTSNDIVEIFTVL
                                                            <mark>iea</mark>vrkalerelyhvistogsyvnyrhlallcotmtcrghlmait
                                                                                                                       1413
                                                           ILAVRKALERELYHVISTOGSYVNYRHLALLCOTMTCRGHLMAIT
tr|A0A1A8DQ60|A0A1A8DQ60 9TELE
                                          RTTSNDIVEIFTVL
                                                                                                                       1413
tr|A0A1A8NSR8|A0A1A8NSR8 9TELE
                                                            IEAVRKALERELYHVISEDGSYVNYRHLALLCDIMICRGHLMAIT
                                                                                                                        1413
tr|A0A1W4YLM7|A0A1W4YLM7 9TELE
                                          RTTSNDIVEIFTVL
                                                           CIEAVRKALERELYHVISFDGSYVNYRHLALLCDIMICRGHLMAIT
                                                                                                                       1413
                                                            ieaargslikevrnviefdgsyvnyrhlallud mtsqgtlmait
tr|A0A1M8A6L7|A0A1M8A6L7 MALS4
                                          RTISNNCVEVERVE
                                                                                                                       1399
                                          RTYSNS FVEILSV<mark>LGIEA</mark>TRSSLYKEILNVIA FDGSYVNYRHMALLVDVMTSRGYLMAIT
tr|A0A1B2J8C6|A0A1B2J8C6 PICPA
                                                                                                                       1388
tr|F2QW17|F2QW17_KOMPC
tr|A3GID7|A3GID7_PICST
                                          RTYSNSFVEILSVLGIEATRSSLYKEILNVIAFDGSYVNYRHMALLVDVMTSRGYLMAIT
                                                                                                                       1388
                                          RTYSNNFIEILSV<mark>IGIE</mark>ATRAALFKEILNVLSFDGSYVNYRHMALLVDVMTSRGHLMAIT
                                                                                                                        1387
                                          RTYSNDFIEVLSVLGIEATRSSLYKEILNVIAFDGSYVNYRHMALLVDVMTSRGHLMAIT
tr|A0A1D8PUA6|A0A1D8PUA6 CANAL
                                                                                                                       1387
                                         tr|G8BEH9|G8BEH9 CANPC
                                                                                                                       1387
                                          RHGVNRODTGFLMKCSFEETVDVIMEAAAHGESDFMKGVSENIMLGQLAFAGTGCFDLLL
tr|A0A1U8DYN0|A0A1U8DYN0 ALLSI
                                                                                                                        1483
                                          RHGVNRQDTGFLMK<mark>CS</mark>FEETVDVIMEAAAHGESDPMKGVSENIMLGQLAPAGTGCFDLLL
tr|A0A1L8H4P4|A0A1L8H4P4 XENLA
                                                                                                                        1473
                                         RHG NRODTGFLMKCS FEETVDVIMEAAAHGESDPMKGVSENIMLGOLAPAGTGCFDLLL
RHG NRODTGFLMKCS FEETVDVIMEAAAHGESDPMKGVSENIMLGOLAPAGTGCFDLLL
RHG NRODTGFLMKCS FEETVDVIMEAAAHGESDPMKGVSENIMLGOLAPAGTGCFDLLL
                                                                                                                        1474
tr|H9GLG5|H9GLG5 ANOCA
tr|H2R1J6|H2R1J6_PANTR
                                                                                                                        1475
tr | G1MCZ1 | G1MCZ1 AILME
                                                                                                                        1475
                                                           S FEETVDVIMEAAAHGESDPMKGVSENIMLGQLAPAGTGCFDLLL
S FEETVDVIMEAAAHGESDPMKGVSENIMLGQLAPAGTGCFDLLL
                                          RHGVNRQDTGPLMK
tr|008847|008847 MOUSE
                                                                                                                        1475
                                          RHGVNRQDTGPLMK
tr|S7PWZ6|S7PWZ6 MYOBR
                                                                                                                        1475
                                          RHGVNRQDTGPLMKCSFEETVDVIMEAAAHGESDPMKGVSENIMLGQLAPAGTGCFDLLL
tr|D4A5A6|D4A5A6_RAT
                                                                                                                        1475
                                         RHGVNRODTGFLMKCSFEETVDVIMEAAAHGESDPMKGVSENIMLGQLAPAGTGCFDLLL
RHGVNRODTGFLMKCSFEETVDVIMEAAAHGESDPMKGVSENIMLGQLAPAGTGCFDLLL
RHGVNRODTGFLMKCSFEETVDVIMEAAAHGESDPMKGVSENIMLGQLAPAGTGCFDLLL
sp|P08775|RPB1 MOUSE
                                                                                                                        1475
sp|P24928|RPB1 HUMAN
                                                                                                                        1475
tr|A0A1S3EWL2|A0A1S3EWL2 DIPOR
                                                                                                                        1475
                                                            S FEETVDVIMEAAAHGESDPMKGVSENIMLGQLAPAGTGCFDLLL
S FEETVDVIMEAAAHGESDPMKGVSENIMLGQLAPAGTGCFDLLL
                                          RHGVNRQDTGPLMK
sp|P11414|RPB1 CRIGR
                                                                                                                        1475
                                          RHGVNRQDTGPLMK
tr|035559|035559 CRIGR
                                                                                                                        1475
tr|A0A2I3M9H2|A0A2I3M9H2 PAPAN
                                          RHGVNRQDTGPLMKC
                                                            SFEETVDVIMEAAAHGESDPMKGVSENIMLGQLAPAGTGCFDLLL
                                                                                                                        1474
                                                           S FEETVDVIMEAAAHGESDPMKGVSENIMIGQIAPAGTGCFDLLL
S FEETVDVIMEAAAHGESDPMKGVSENIMIGQIAPAGTGCFDLLL
                                          RHGVNRQDTGPLMK
tr|F7HB40|F7HB40 MACMU
                                                                                                                        1475
                                          RHGVNRQDTGPLMKC
tr|A0A2K6RYW9|A0A2K6RYW9 SAIBB
                                                                                                                        1475
tr|W5N8Z6|W5N8Z6 LEPOC
                                          RHGVNRQDTGPLMKC
                                                            SFEETVOVIMEAAAHGESDPMKGVSENIMLGQLAPAGTGCFOLLL
                                                                                                                        1474
                                                           S FEETVOVIMEASSIGECDPMKGVSENIMIGQIAPASTGCFDLIL
S FEETVOVIMEASSIGESDPMKGVSENIMIGQIAPASTGCFDLIL
S FEETVOVIMEASSIGECDPMKGVSENIMIGQIAPASTGCFDLIL
                                          RHGINRODIGPLMK
tr|I3JRW6|I3JRW6 ORENI
                                                                                                                        1473
                                          RHGINRODIGPLMK
tr|A0A0R4IMS9|A0A0R4IMS9 DANRE
                                                                                                                        1472
tr|A0A1A7X327|A0A1A7X327 9TELE
                                          RHGINRODIGPLMK
                                                                                                                        1473
tr|A0A1A8UKD7|A0A1A8UKD7 NOTFU
                                          RHGINRODIGPLMK
                                                            FEETVDVIMEA$SHGECDPMKGVSENIMLGQLAPAGTGCFDLLL
                                                                                                                        1473
                                          RHGINRODTGPLMK
                                                            FEETVDVIMEA$SHGECDPMKGVSENIMLGQLAPAGTGCFDLLL
tr|A0A1A8ER05|A0A1A8ER05 9TELE
                                                                                                                        1473
tr|A0A1A8DQ60|A0A1A8DQ60 9TELE
                                          RHGINRQDIGPLMK
                                                            FEETVOVIMEA$SHGECDPMKGVSENIMLGQLAPAGTGCFDLLL
                                                                                                                        1473
                                                            FEETVDVIMEASSHGECDPMKGVSENIMIGOLAPASTGCFDLLL
FEETVDVIMEASSHGECDPMKGVSENIMIGOLAPASTGCFDLLL
tr|A0A1A8NSR8|A0A1A8NSR8 9TELE
                                          RHGINRODTGPLMK
                                                                                                                        1473
                                          RHGINRODIGPLMK
tr|A0A1W4YLM7|A0A1W4YLM7 9TELE
                                                                                                                        1473
tr|A0A1M8A6L7|A0A1M8A6L7 MALS4
                                          RHGINRTSQGALMR
                                                            FEETVEILMEAASMGDMDDCKGVGONILLGOMAPMGTGAFDLNL
                                                                                                                        1459
                                                            S FEETVEILFEAGAAAELDDCRGVSENVMLGQLAPMGTGAFDVMI
tr|A0A1B2J8C6|A0A1B2J8C6 PICPA
                                          RHGINRADIGALMR
                                                                                                                        1448
tr|F2QW17|F2QW17 KOMPC
                                          RHGINRADIGALMR
                                                            SFEETVEILFEAGAAAELDDCRGVSENVMLGQLAPMGTGAFDVMI
                                                                                                                        1448
                                                            FEETVEILLEAGASAELDDCRGISENVMLGQMAPLGTGAFDVML
FEETVEILLDAAAAAELDDCKGISENVMLGQMAPLGTGSFDLMV
                                          RHGINRSDIGALMR
tr|A3GID7|A3GID7 PICST
                                                                                                                        1447
                                          RHGINRSDIGALMR
tr|A0A1D8PUA6|A0A1D8PUA6 CANAL
                                                                                                                        1447
                                          RHGINRSDIGALMR
                                                            SFEETVEILLDAAAAAELDDCKGISENVMLGQMAPLGTGAFDVMV
tr|G8BEH9|G8BEH9 CANPC
                                                                                                                        1447
                                                             *****:*:* : .: * :*:.:**:**
```

```
SPNYSPTSPSYS---PTSPSYSPTTPSY--SPTSPSYSPTTPSYSPTT------
tr|A0A1U8DYN0|A0A1U8DYN0 ALLSI
                                    SPSYSPTSPNYS---PTSPSYSPTSPSY---------
                                                                                                      1668
tr|A0A1L8H4P4|A0A1L8H4P4 XENLA
                                    SPSYSPTSPNYS---PTSPSYSPTSPSY--SPTSPSYSPTSPSYSPTSPS-----
tr|H9GLG5|H9GLG5_ANOCA
                                                                                                       1669
                                    SPSYSPTSPNYS---PTSPSYSPTSPSY---SPTSPSYSPTSPSYSPTSPS------
tr|H2R1J6|H2R1J6 PANTR
                                                                                                       1660
                                    SPSYSPTSPNYS---PTSPSYSPTSPSY--SPTSPSYSPTSPSYSPTSPS------
tr|G1MCZ1|G1MCZ1 AILME
                                                                                                       1673
                                    SPSYSPTSPNYS---PTSPSYSPTSPSY--SPTSPSYSPTSPSYSPTSPS------
tr|008847|008847_MOUSE
                                                                                                       1670
                                    SPSYSPTSPNYS---PTSPSYSPTSPSY--SPTSPSYSPTSPSYSPTSPS------
tr|S7PWZ6|S7PWZ6 MYOBR
                                                                                                       1670
tr|D4A5A6|D4A5A6 RAT
                                    SPSYSPTSPNYS---PTSPSYSPTSPSY--SPTSPSYSPTSPSYSPTSPS------
                                                                                                       1670
                                    SPSYSPTSPNYS---PTSPSYSPTSPSY--SPTSPSYSPTSPSYSPTSPS-----
sp|P08775|RPB1 MOUSE
                                                                                                       1670
                                    SPSYSPTSPNYS---PTSPSYSPTSPSY--SPTSPSYSPTSPSYSPTSPS------
                                                                                                      1670
sp | P24928 | RPB1 HUMAN
tr|A0A1S3EWL2|A0A1S3EWL2 DIPOR
                                    SPSYSPTSPNYS---PTSPSYSPTSPSY--SPTSPSYSPTSPSYSPTSPS-----
                                                                                                       1670
                                    SPSYSPTSPNYS---PTSPSYSPTSPSY--SPTSPSYSPTSPSYSPTSPS------
sp|P11414|RPB1 CRIGR
                                                                                                       1670
                                    SPSYSPTSPNYS---PTSPSYSPTSPSY--SPTSPSYSPTSPSYSPTSPS-----
tr|035559|035559 CRIGR
                                                                                                       1670
                                    SPSYSPTSPNYS---PTSPSYSPTSPSY--SPTSPSYSPTSPSYSPTSPS-----
tr|A0A2I3M9H2|A0A2I3M9H2_PAPAN
tr|F7HB40|F7HB40 MACMU
                                    SPSYSPTSPNYS---PTSPSYSPTSPSY--SPTSPSYSPTSPSYSPTSPS-----
                                    SPSYSPTSPNYS---PTSPSYSPTSPSY--SPTSPSYSPTSPSYSPTSPS-----
tr|A0A2K6RYW9|A0A2K6RYW9 SAIBB
                                    LSIGG-----G-----G------G------
tr|W5N8Z6|W5N8Z6_LEPOC
                                                                                                       1662
tr||I3JRW6||I3JRW6 ORENI
                                    SPSYSPTSPNYS---PTSPSYSPTSPSY--SPTSPSYSPTSPSYSPTSPS-----
                                    SPSYSPTSPNYS---PTSPSYSPTSPSY--SPTSPSYSPTSPSYSPTSPS-----
tr|A0A0R4IMS9|A0A0R4IMS9 DANRE
                                                                                                       1667
                                    SPSYSPTSPNYS---PTSPSYSPTSPSY--SPTSPSYSPTSPSYSPTSPS-----
tr|A0A1A7X327|A0A1A7X327 9TELE
                                                                                                       1668
                                    SPSYSPTSPNYS---PTSPSYSPTSPSY--SPTSPSYSPTSPSYSPTSPS-----
tr|A0A1A8UKD7|A0A1A8UKD7 NOTFU
                                                                                                       1668
                                    SPSYSPTSPNYS---PTSPSYSPTSPSY--SPTSPSYSPTSPSYSPTSPS-----
tr|A0A1A8ER05|A0A1A8ER05 9TELE
                                                                                                       1668
tr|A0A1A8DQ60|A0A1A8DQ60_9TELE
tr|A0A1A8NSR8|A0A1A8NSR8|9TELE
                                    SPSYSPTSPNYS---PTSPSYSPTSPSY--SPTSPSYSPTSPSYSPTSPS--------
SPSYSPTSPNYS---PTSPSYSPTSPSY--SPTSPSYSPTSPSYSPTSPS------
                                                                                                       1668
                                                                                                       1668
                                    SPSYSPTSPNYS---PTSPSYSPTSPSY--SPTSPSYSPTSPSYSPTSPS--------
SPAMGAMSPWVGAAGATSPAYSPTSPRIFAEATSPAYSPTSPSYSPSSPMIGATHVRSAY
tr|A0A1W4YLM7|A0A1W4YLM7_9TELE
                                                                                                       1668
tr|A0A1M8A6L7|A0A1M8A6L7 MALS4
                                                                                                      1633
                                    SPGYSP----TSPSYSPTSPSY--SPTSPSYSPTSPSYSPTSPS-----
tr|A0A1B2J8C6|A0A1B2J8C6 PICPA
                                                                                                      1594
                                    SPGYSP------SPSYSPTSPSY--SPTSPSYSPTSPSSYSPTSPS--------
SPSYSP-------SPSYSPTSPAY--SPTSPSYSPTSPSYSPTSPS------
tr|F2QW17|F2QW17_KOMPC
tr|A3GID7|A3GID7 PICST
                                                                                                      1594
                                                                                                      1598
tr|A0A1D8PUA6|A0A1D8PUA6 CANAL
                                    SPAYSP-----TSPSYAPTSPAY--SPTSPSYAPTSPAYSPTSPA-----
                                                                                                      1598
                                    SPSYSP-----TSPSYSPTSPSY--SPTSPSYSPTSPSYSPTSPS-----
tr|G8BEH9|G8BEH9_CANPC
                                                                                                      1602
                                                    **:*:**:*
                                   PTSPSYSPTSPSYSPTSPNYTPTSP
tr|A0A1U8DYN0|A0A1U8DYN0 ALLSI
                                                                                                    1754
                                   PTSPSYSPTSPSYSPTSPSYSPTSPSYSPTSPNYSPSSPNYTPTSPSYSPTSPSYSPTSP
                                                                                                    1779
tr|AOA1L8H4P4|AOA1L8H4P4 XENLA
                                                                                                    1780
tr|H9GLG5|H9GLG5 ANOCA
                                   PTSPSYSPTSPSYSPTSPSYSPTSPNYSPTSPNYTPTSPSYSPTSPSYSPTSPNYTPTSP
tr|H2R1J6|H2R1J6 PANTR
                                   PTSPSYSPTSPSYSPTSPSYSPTSPNYSPTSPNYTPTSPSYSPTSPSYSPTSPNYTPTSP
                                                                                                    1771
tr|G1MCZ1|G1MCZ1 AILME
                                   PTSPSYSPTSPSYSPTSPSYSPTSPNYSPTSPNYTPTSPSYSPTSPSYSPTSPNYTPTSP
                                                                                                    1784
tr|008847|008847 MOUSE
                                  PTSPSYSPTSPSYSPTSPSYSPTSPNYSPTSPNYTPTSPSYSPTSPSYSPTSPNYTPTSP
                                                                                                    1781
tr|S7PWZ6|S7PWZ6 MYOBR
                                  PTSPSYSPTSPSYSPTSPSYSPTSPNYSPTSPNYTPTSPSYSPTSPSYSPTSPNYTPTSP
                                                                                                    1781
                                  PTSPSYSPTSPSYSPTSPNYSPTSPNYTPTSPSYSPTSPSYSPTSPNYTPTSP
tr|D4A5A6|D4A5A6 RAT
                                                                                                    1781
sp|P08775|RPB1 MOUSE
                                   PTSPSYSPTSPSYSPTSPNYSPTSPNYTPTSPSYSPTSPSYSPTSPNYTPTSP
                                                                                                    1781
sp|P24928|RPB1 HUMAN
                                   PTSPSYSPTSPSYSPTSPSYSPTSPNYSPTSPNYTPTSPSYSPTSPSYSPTSPNYTPTSP
                                                                                                    1781
tr|AOA1S3EWL2|AOA1S3EWL2 DIPOR PTSPSYSPTSPSYSTTSPSYSPTSPNYTPTSPSYSPTSPSYSPTSPNYTPTSP
                                                                                                    1781
sp|P11414|RPB1 CRIGR
                                  PTSPSYSPTSPSYSPTSPSYSPTSPNYSPTSPNYTPTSPSYSPTSPSYSPTSPNYTPTSP
                                                                                                    1781
tr|035559|035559 CRIGR
                                   PTSPSYSPTSPSYSPTSPSYSPTSPNYSPTSPNYTPTSPSYSPTSPSYSPTSPNYTPTSP
                                                                                                    1781
tr|A0A2I3M9H2|A0A2I3M9H2_PAPAN
                                   PTSPSYSPTSPSYSPTSPNYSPTSPNYTPTSPSYSPTSPSYSPTSPNYTPTSP
                                                                                                    1742
                                                                                                    1743
tr|F7HB40|F7HB40_MACMU
                                   PTSPSYSPTSPSYSPTSPSYSPTSPNYSPTSPNYTPTSPSYSPTSPSYSPTSPNYTPTSP
tr|A0A2K6RYW9|A0A2K6RYW9 SAIBB
                                   PTSPSYSPTSPSYSPTSPSYSPTSPNYSPTSPNYTPTSPSYSPTSPSYSPTSPNYTPTSP
                                                                                                    1743
tr|W5N8Z6|W5N8Z6 LEPOC
                                   PTSPSYSPTSPSYSPTSPSYSPTSPSYSPTSPNYTPTSPSYSPTSPSYSPTSPNYTPTSP
                                                                                                    1770
                                   PTSPSYSPTSPSYSPTSPSYSPTSPSYSPTSPNYTPTSPSYSPTSPSYSPTSPSYSPTSP
tr|I3JRW6|I3JRW6 ORENI
                                                                                                    1779
tr|A0A0R4IMS9|A0A0R4IMS9 DANRE
                                   PTSPSYSPTSPSYSPTSPSYSPTSPSYSPTSPSYSPTSPSYSPTSPSYSPTSP
                                                                                                    1778
tr|A0A1A7X327|A0A1A7X327 9TELE
                                   PTSPSYSPTSPSYSPTSPSYSPTSPSYSPTSPNYTPTSPSYSPTSPSYSPTSPSYSPTSP
                                                                                                    1779
tr|A0A1A8UKD7|A0A1A8UKD7 NOTFU
                                                                                                    1779
                                   PTSPSYSXTSPSYSPTSPSYSPTSPSYSPTSPSYSPTSPSYSPTSPSYSPTSP
tr|A0A1A8ER05|A0A1A8ER05 9TELE
                                   PTSPSYSPTSPXXXXTSPSYSPTSPSYSPTSPNYTPTSPSYSPTSPSYSPTSPSYSPTSP
                                                                                                    1779
tr|A0A1A8DQ60|A0A1A8DQ60 9TELE
                                   PTSPSYSPTSPSYSPTSPSYSPTSPSYSPTSPNYTPTSPSYSPTSPSYSPTSPSYSPTSP
                                                                                                    1772
tr|A0A1A8NSR8|A0A1A8NSR8 9TELE
                                   PTSPSYSPTSPSYSPTSPSYSPTSPSYSPTSPSYSPTSPSYSPTSPSYSPTSP
                                                                                                    1772
tr|AOA1W4YLM7|AOA1W4YLM7 9TELE
                                   PTSPSYSPTSPSYSPTSPSYSPTSPSYSPTSPNYTPTSPSYSPTSPSYSPTSPSYSPTSP
                                                                                                    1779
tr|A0A1M8A6L7|A0A1M8A6L7 MALS4
                                   PTSPQYSPTSPQYSPTSPQYSPTSPQYSPTSPQYSPTSPQYSPTSPQYSPTSP
                                                                                                    1753
sp|P04050|RPB1 YEAST
                                   PTSPSYSPTS
                                                                                                    1687
tr|A0A1B2J8C6|A0A1B2J8C6 PICPA
                                   PTSPQYSPTSPQYSPTSPQYSPTSPQYSPTSPQYSPTSPQYSPTSPQYSPTSPQ
                                                                                                    1698
tr|F2QW17|F2QW17_KOMPC
                                   PTSPQYSPTSPQYSPTSPQYSPTSPQYSPTSPQYSPTSPQYSPTSPQYSPTSPQYSPTSP
                                                                                                    1698
tr|A3GID7|A3GID7 PICST
                                   ----SYSPTSPSYSPTSPSYSPTSPSYSPTSPSYSPTSPSYSPTSPSYSPTSPOYSPTSP
                                                                                                    1695
tr|A0A1D8PUA6|A0A1D8PUA6 CANAL
                                   PTSPSYSPTSPAYSPTSPSYSPTSPSYSPTSPQYSPTSPSYSPTSPQYSPTSPSYSPTSP
                                                                                                    1688
tr|G8BEH9|G8BEH9 CANPC
                                   PTSPSYSPTSPSYSPTSPSYSPTSPSYSPTSPSYSPTSPSYSPTSPSYSPTSPSYSPTSP
                                                                                                    1706
                                       . . . . . . . .
```

```
tr|A0A1U8DYN0|A0A1U8DYN0_ALLSI
                                                  NYSPTSPSYSPTSPSYSPTSPSYSPSSPRYTPQSPTYTPSSPSYSPSSPSYSPTSPKYTP
                                                                                                                                               1814
tr|A0A1L8H4P4|A0A1L8H4P4_XENLA
                                                  SYSPTSPNYSPTSPSYSPTSPSYSPSSPRYTPQSPSYTPSSPSYSPSSPSYSPTSPKYTP
                                                                                                                                               1839
tr|H9GLG5|H9GLG5 ANOCA
                                                  NYSPTSPSYSPTSPSYSPTSPSYSPSSPRYTPOSPTYTPSSPSYSPSSPSYSPTSPKYTP
                                                                                                                                               1840
tr|H2R1J6|H2R1J6_PANTR
                                                  NYSPTSPSYSPTSPSYSPTSPSYSPSSPRYTPOSPTYTPSSPSYSPSSPSYSPTSPKYTP
                                                                                                                                               1831
tr|G1MCZ1|G1MCZ1_AILME
tr|008847|008847_MOUSE
                                                  NYSPTSPSYSPTSPSYSPTSPSYSPSSPRYTPQSPTYTPSSPSYSPSSPSYSPTSPKYTP
                                                                                                                                               1844
                                                  NYSPTSPSYSPTSPSYSPTSPSYSPSSPRYTPQSPTYTPSSPSYSPSSPSYSPTSPKYTP
tr|S7PWZ6|S7PWZ6_MYOBR
                                                  NYSPTSPSYSPTSPSYSPTSPSYSPSSPRYTPQSPTYTPSSPSYSPSSPSYSPTSPKYTP
                                                                                                                                               1841
triD4A5A6|D4A5A6 RAT
                                                  NYSPTSPSYSPTSPSYSPTSPSYSPSSPSYTPOSPTYTPSSPSYSPSSPSYSPTSPKYTP
                                                                                                                                               1841
sp|P08775|RPB1 MOUSE
                                                  NYSPTSPSYSPTSPSYSPTSPSYSPSSPRYTPQSPTYTPSSPSYSPSSPSYSPTSPKYTP
                                                                                                                                               1841
sp | P24928 | RPB1_HUMAN
                                                  NYSPTSPSYSPTSPSYSPTSPSYSPSSPRYTPQSPTYTPSSPSYSPSSPSYSPASPKYTP
                                                                                                                                               1841
tr|A0A1S3EWL2|A0A1S3EWL2 DIPOR
                                                  NYSPTSPSYSPTSPSYSPTSPSYSPSSPRYTPQSPTYTPSSPSYSPSSPSYSPTSPKYTP
                                                                                                                                               1841
sp|P11414|RPB1 CRIGR
                                                  MYSPTSPSYSPTSPSYSPTSPSYSPSSPRYTPQSPTYTPSSPSYSPSSPSYSPTSPKYTP
                                                                                                                                               1841
                                                  NYSPTSPSYSPTSPSYSPTSPSYSPSSPRYTPQSPTYTPSSPSYSPSSPSYSPTSPKYTP
tr|035559|035559 CRIGR
                                                                                                                                               1841
tr|A0A2I3M9H2|A0A2I3M9H2 PAPAN
                                                  NYSPTSPSYSPTSPSYSPTSPSYSPSSPRYTPQSPTYTPSSPSYSPSSPSYSPTSPKYTP
                                                                                                                                               1802
tr|F7HB40|F7HB40_MACMU
                                                  NYSPTSPSYSPTSPSYSPTSPSYSPSSPRYTPQSPTYTPSSPSYSPSSPSYSPTSPKYTP
                                                                                                                                               1803
tr|A0A2K6RYW9|A0A2K6RYW9 SAIBB
                                                  NYSPTSPSYSPTSPSYSPTSPSYSPSSPRYTPOSPTYTPSSPSYSPSSPSYSPTSPKYTP
                                                                                                                                               1803
                                                  NYSPTSPSYSPTSPSYSPTSPSYSSTSPRYTPOSPTYTPSSPSYSPSSPSYSPTSPKYTP
tr|W5N8Z6|W5N8Z6_LEPOC
tr|I3JRW6|I3JRW6 ORENI
                                                                                                                                               1830
                                                  NYTPTSPNYSPTSPSYSPTSPSYSPSSPRYTPQSPTYTPSSPSYSPSSPSYSPTSPKYTP
tr|A0A0R4IMS9|A0A0R4IMS9_DANRE
                                                  NYTPTSPNYSPTSPSYSPTSPSYSPSSPRYTPQSPTYTPSSPSYSPSSPSYSPTSPKYTP
                                                                                                                                               1838
tr|A0A1A7X327|A0A1A7X327_9TELE
tr|A0A1A8UKD7|A0A1A8UKD7 NOTFU
                                                  NYTPTSPNYSPTSPSYSPTSPSYSPSSPRYTPQSPTYTPSSPSYSPSSPSYSPTSPKYTP
NYTPTSPNYSPTSPSYSPTSPSYSPSSPRYTPQSPTYTPSSPSYSPSSPSYSPTSPKYTP
                                                                                                                                               1839
                                                                                                                                               1839
tr|A0A1A8ER05|A0A1A8ER05_9TELE
                                                  NYTPTSPNYSPTSPSYSPTSPSYSPSSPRYTPQSPTYTPSSPSYSPSSPSYSPTSPKYTP
                                                                                                                                               1839
                                                  NYTPTSPNYSPTSPSYSPTSPSYSPSSPRYTPQSPTYTPSSPSYSPSSPSYSPTSPKYTP
tr|A0A1A8DQ60|A0A1A8DQ60_9TELE
tr|A0A1A8NSR8|A0A1A8NSR8_9TELE
                                                  NYTPTSPNYSPTSPSYSPTSPSYSPSSPRYTPOSPTYTPSSPSYSPSSPSYSPTSPKYTP
                                                                                                                                               1832
tr|AOA1W4YLM7|AOA1W4YLM7 9TELE
                                                  NYTPTSPNYSPTSPSYSPTSPSYSPSSPSYTPOSPTYTPSSPSYSPSYSPSYSPTSPKYTP
                                                                                                                                               1839
tr|AOA1M8A6L7|AOA1M8A6L7 MALS4
                                                  QYSPTSPQYSPTSPQYSPTSPQYSPGGAGGAAGAASPRRSRMT---SKPTWQR--
                                                                                                                                               1803
tr|A0A1B2J8C6|A0A1B2J8C6_PICPA
                                                  QYSPTSPQYSPTSPQYSPTSPQYSPASPQYSPSRHSPNGESKE---GE------
                                                                                                                                               1743
tr|F2QW17|F2QW17_KOMPC
tr|A3GID7|A3GID7 PICST
                                                  QYSPTSPQYSPTSPQYSPTSPQYSPASPQYSPSRHSPNGESKE---GE-------
QYSPTSPSYSPTSPQYSPTSPQYSPGSPEYSPNSPKTEDKKNE---D--------
                                                                                                                                               1743
                                                                                                                                               1739
tr|A0A1D8PUA6|A0A1D8PUA6_CANAL
                                                  QYSPTSPQYSPTSPQYSPTSPQYSPGSPGYEPEPPKKDEN------
tr|G8BEH9|G8BEH9 CANPC
                                                  TYSPTSPQYSPTSPQYSPTSPSYSPGSPGYNPESPKKEEK------
        tr|A0A1U8DYN0|A0A1U8DYN0_ALLSI
tr|A0A1L8H4P4|A0A1L8H4P4_XENLA
tr|H9GLG5|H9GLG5_ANOCA
tr|H2R1J6|H2R1J6_PANTR
tr|G1MC21|G1MC21_AILME
tr|O08847|O08847_MOUSE
tr|S7PWZ6|S7PWZ6_MYOBR
tr|D4A5A6|D4A5A6_PAT
                                                                                                                                   1926
                                                                                               SPI
                                                                                                                                   1968
                                                                                                              EEN
EEN
EEN
                                                                                                                                   1960
1973
1966
                                                                                               SPI
       tr|O08847|O08847|MOUSE

tr|S7PWZ6|S7PWZ6_MYOBR

tr|D4A5A6|D4A5A6_RAT

sp|P08775|RPB1_MOUSE

sp|P24928|RPB1_HUMAN

tr|A0A153EWL2|A0A153EWL2_DIPOR

sp|P11414|RPB1_CRIGR

tr|O35559|O35559_CRIGR

tr|A0A213M9H2|A0A213M9H2_PAPAN

tr|F7HB40|F7HB40_MACMU

tr|F7HB40|F7HB40_MACMU

tr|A0A2K6RYW9|A0A2K6RYW9_SAIBB

tr|WSN826|WSN826_LEPOC

tr|I3JRW6|I3JRW6_ORENI

tr|A0A1A8UKD7|A0A0R4IMS9_DANRE

tr|A0A1A8UKD7|A0A1A8UKD7_NOTFU

tr|A0A1A8UKD7|A0A1A8UKD7_NOTFU

tr|A0A1A8UKD7|A0A1A8UKD7_NOTFU

tr|A0A1A8NSR8|A0A1A8NSR8_9TELE

tr|A0A1A8NSR8|A0A1A8NSR8_9TELE

tr|A0A1M8NSR8|A0A1A8NSR8_9TELE

tr|A0A1M8NSR8|A0A1A8NSR8_9TELE

tr|A0A1M8NSR8|A0A1A8NSR8_9TELE

tr|A0A1M8NSR8|A0A1A8NSR8_9TELE

tr|A0A1M8NSR8|A0A1A8NSR8_9TELE

tr|A0A1M8NSR8|A0A1A8NSR8_9TELE

tr|A0A1M8NSR8|A0A1A8NSR8_9TELE
                                                                                                              CEN
CEN
CEN
                                                                                               SPI
                                                                                                                                   1970
                                                                                                                                   1970
1970
1970
1970
1970
1970
                                                                                               SPI
                                                                                               SPI
                                                                                                              EEN
EEN
EEN
EEN
                                                                                               SPI
                                                                                               SPI
                                                                                               SPI
                                                                                                                                    1931
1932
                                                                                               SPI
                                                                                                                                    1959
1966
                                                                                                              EENN---
ESDEENN
ESDEENN
                                                                                               SPI
                                                                                                                                    1965
                                                                                               SPI
                                                                                                              esdeenn
Esdeenn
                                                                                                                                    1962
1962
                                                                                                              DD1
                                                                                                                                    1968
                                                                                                                                    1803
        tr|F2QW17|F2QW17_KOMPC
tr|A3GID7|A3GID7=PICST
tr|A0A1D8PUA6|A0A1D8PUA6_CANAL
                                                                                                                                    1743
1739
         tr|G8BEH9|G8BEH9_CANPC
```

Fig. 4 MSA of the Rpb1, the elongation subunits of eukaryotic RNAP II

A0A1U8DYN0 _ALLSI Alligator sinensis
H9GLG5 _ANOCA Anolis carolinensis
G1MCZ1 _AILME Ailuropoda melanoleuca
S7PWZ6 _MYOBR Myotis brandtii
P08775 _MOUSE Mus musculus
A0A1S3EWL2 _DIPOR Dipodomys ordii
O35559 _CRIGR Cricetulus griseus
F7HB40| _MACMU Macaca mulatta
W5N8Z6 _LEPOC Lepisosteus oculatus
A0A0R4IMS9 _DANRE Danio rerio
A0A1A8UKD7 _NOTFU Nothobranchius furzeri
A0A1A8DQ60 _9TELE Nothobranchius kadleci
A0A1W4YLM7 _9TELE Scleropages formosus
P04050|RPB1_YEAST Saccharomyces cerevisiae
F2QW17 _KOMPC Komagataella phaffii
A0A1D8PUA6 _CANAL Candida albicans

A0A1L8H4P4 _XENLA Xenopus laevis H2R1J6 PANTR Pan troglodytes O08847 MOUSE Mus musculus D4A5A6 RAT Rattus norvegicus P24928|RPB1 HUMAN Homo sapiens P11414 CRIGR Cricetulus griseus A0A2I3M9H2 PAPAN Papio Anubis A0A2K6RYW9 SAIBB Saimiri boliviensis 13JRW6 ORENI Oreochromis niloticus A0A1A7X327 _9TELE Aphyosemion striatum A0A1A8ER05 9TELE Nothobranchius korthausae A0A1A8NSR8 9TELE Nothobranchius rachovii A0A1M8A6L7 _MALS4 Malassezia sympodialis A0A1B2J8C6 PICPA Komagataella pastoris A3GID7 PICST Scheffersomyces stipitis G8BEH9 CANPC Candida parapsilosis

7. MIX AND MATCH ANALYSIS OF EUBACTERIAL AND EUKARYOTIC INITIATION AND ELONGATION SUBUNITS

For this analysis a small number of initiation and elongation subunits of eubacterial and eukaryotic MSU RNAPs were subjected to MSA with ClustalW Omega programme and presented in Figs. 5 and 6. For eubacteria, *E. coli* MSU RNAP and for eukaryotes, *S. cerevisiae* MSU RNAP II sequences were used as standards. Only highly conserved regions are shown in the figures.

7.1 Mix and Match Analysis of the *E. coli*(β) and *S. cerevisiae* (Rpb2) Initiation Subunits

Fig. 5 shows the mix and match analysis of the initiation subunits from pro- and eukaryotic MSU RNAPs. The eukaryotic sequences are shown in red. The active site regions are highlighted in yellow the representative sequences are highlighted in yellow. There are a good number of sequences aligning in both. Significant among them are -FI/VINGS/TEK/RVL/II/VA/SQ- (~200), IETPE (~500) preceding this sequence a WG in

eukaryotic and YG in prokaryotic sequences, -ASI/LIPF-, the metal ion binding regions GYNQ/FEDS- (~800), -LDD/ED/SGL/I- (~850); -GDKF/MAS/GRHGXKG- (~1000), HLVDDKI/ MHAR (~1050), RFGEME (~1100). The catalytic regions are matching but among themselves only and there is no complete consensus among them, i.e., they are located at different regions. However, the metal binding motifs are aligning in both the initiation subunits suggesting the metal binding regions are preserved as a domain and did not diverge much during evolution. The prokaryotic active site region is placed around 540-555/1342 amino acids whereas eukaryotic active site region is placed around 851-866/1224 amino acids. The notable difference in the eukaryotic template binding pair sis KG in lower eukaryotes like human and in lower eukaryotes it uses YG like yeasts. The eukaryotes follow two very similar ending sequences at the C- terminal -ACKLLFQELMSMSIAPRMMSVend as (~ 1150) and AAKLLFQELMAMNITPRLYT-(~1200) (deviating amino acids are shown in red) the prokaryotes follow an altogether different consensus end sequence as -SFNVLLKEIRSL- (~1280). Thus, the eukaryotic and prokaryotic ending sequences are different and their significance is not clear now.

Table 4 Amino acids around the catalytic amino acid K/R and the YG/FG pair in DNA polymerases, DNA dependent SSU and MSU RNAPs

Polymerase Type	Catalytic Region
SSU RNAP family	
Viral SSU RNA pol (T7, T3, SP6)	-TR ⁻⁴ VTKR ¹ SVMTLAY ⁸ GS-
Mitochondrial SSU RNA pol (Yeast)	-TR ⁻⁴ KVVKQ ¹ TVMTNVY ⁸ GV
Chloroplast SSU pol (ARATH)	-DR⁴KLVKQ¹TVMTSVY ⁸ GV-
DNA polymerase family	
E. coli DNA pol I	-QR ⁻⁴ RSAKA ¹ INFGLIY ⁸ GM-
MSU RNAP family (Initiation subunits	
E. coli MSU RNAP β subunit	- ⁵³⁹ TR ⁻⁸ ERAGFEVRD ¹ VHPTHY ⁷ G ⁸ RV ⁵⁵⁸ -
S. cerevisiae MSU RNAP II Rpb2 subunit	.851FR-5SLFFRS1YMDQEKKY9GMSI870-
Human MSU RNAP II Rpb2 subunit	.806FR-5SVFYRS1YKEQESKK9GFDQ825-
MSU RNAP family (Elongation subun	its)
E. coli MSU RNAP β' subunit	- ⁸³³ NSV ⁻ ⁶ DAVKVRS ¹ VVSC ⁵ DTDFGVC ¹² AHC ¹⁵ Y ¹⁶ G ¹⁷ RDL ⁸⁶¹ -
S.cerevisiaeMSURNAPIIRpb1 subunit	-55DPR-6LGSIDRN1LKC4QTC7QEGMNEC14PGHF18G19HI84-
HumanMSURNAPIIRpb1subunit	-59DPR-6QGVIERT1GRC4QTC7AGNMTEC14PGHF18G19HI88-

NB: MSU RNA polymerases which use R in the catalytic site is shown in bold. The invariant R, at -6/-7 is not found in β' subunits of eubacteria; The eukaryotic elongation subunits use an FG instead of a YG

CLUSTAL O (1.2.4): Mix and Match analysis of eukaryotic Rpb2 and eubacterial $\boldsymbol{\beta}$ initiation subunits 5

sp P30876 RPB2 HUMAN		
	LLNGLT DRDLCELNEC PLDPGGY FIINGSE KVL IAQE KMATNT VYVF -AKKD-SKYAYTG	219
tr G3V8Y5 G3V8Y5 RAT	LLNGLTDRDLCELNECPLDPGGYFIINGSEKVLIAQEKMATNTVYVF-AKKD-SKYAYTG	219
tr A0A250Y753 A0A250Y753 CASCN	LLNGLTDRDLCELNEC PLDPGGY FIINGSE KVLIAOE KMATNTVYVF-AKKD-SKYAYTG	219
tr A0A1U7R4C7 A0A1U7R4C7_MESAU	LLNGLTDRDLCELNECPLDPGGYFIINGSEKVLIAQEKMATNTVYVF-AKKD-SKYAYTG	219
tr A0A286XIQ9 A0A286XIQ9_CAVPO	LLNGLTDRDLCELNECPLDPGGYFIINGSEKVLIAQEKMATNTVYVF-AKKD-SKYAYTG	219
tr I3M351 I3M351 ICTTR	LLNGLTDRDLCELNECPLDPGGYFIINGSEKVLIAQEKMATNTVYVF-AKKD-SKYAYTG	219
tr G7P5R6 G7P5R6 MACFA	LLNGLTDRDLCELNEC PLDPGGY FIINGSEKVLIAOEKMATNTVYVF-AKKD-SKYAYTG	219
tr H2QPI8 H2QPI8_PANTR	LLNGLTDRDLCELNECPLDPGGYFIINGSEKVLIAQEKMATNTVYVF-AKKD-SKYAYTG	219
tr A0A1U7V0T5 A0A1U7V0T5_TARSY	LLNGLTDRDLCELNECPLDPGGYFIINGSEKVLIAQEKMATNTVYVF-AKKD-SKYAYTG	219
tr A0A1S2ZSL2 A0A1S2ZSL2 ERIEU	LLNGLTDRDLCELNECPLDPGGYFIINGSEKVLIAQEKMATNTVYVF-AKKD-SKYAYTG	219
tr A0A0D9QYL1 A0A0D9QYL1 CHLSB	LLNGLTDRDLCELNEC PLDPGGY FIINGSEKVLIAOEKMATNTVYVF-AKKD-SKYAYTG	212
tr A0A2K5ZNR7 A0A2K5ZNR7_MANLE	LLNGLTDRDLCELNECPLDPGGYFIINGSEKVLIAQEKMATNTVYVF-AKKD-SKYAYTG	212
tr A0A2I2ZIU3 A0A2I2ZIU3_GORGO	LLNGLTDRDLCELNECPLDPGGY FIINGSEKVLIAQEKMATNTVYVF-AKKD-SKYAYTG	219
tr A0A1D5QGA5 A0A1D5QGA5 MACMU	LLNGLT DRDLCELNEC PLDPGGY FIINGSE KVL IAQE KMA TNT VYVF -AKKD-SKYA YTG	219
tr A0A2J8S2N1 A0A2J8S2N1 PONAB	LLNGLTDRDLCELNECPLDPGGY FIINGSEKVLIAQEKMATNTVYVF-AKKD-SKYAYTG	219
tr A0A2K5K5J5 A0A2K5K5J5_COLAP	LLNGLTDRDLCELNECPLDPGGYFIINGSEKVLIAQEKMATNTVYVF-AKKD-SKYAYTG	219
tr A0A2J8PEW7 A0A2J8PEW7_PANTR	LLNGLTDRDLCELNECPLDPGGYFIINGSEKVLIAQEKMATNTVYVF-AKKD-SKYAYTG	212
tr A0A2K5CY83 A0A2K5CY83_AOTNA	LLNGLT DRD LCELNEC PLDPGGY FIINGSE KVL IAQE KMA TNT VYVF - AKKD-SKYA YTG	219
tr A0A096NEY4 A0A096NEY4 PAPAN	LLNGLTDRDLCELNECPLDPGGYFIINGSEKVLIAQEKMATNTVYVF-AKKD-SKYAYTG	219
tr C9J2Y9 C9J2Y9 HUMAN	LLNGLTDRDLCELNECPLDPGGYFIINGSEKVLIAQEKMATNTVYVF-AKKD-SKYAYTG	212
· · · · · · · · · · · · · · · · · · ·		
tr G8BY61 G8BY61_TETPH	YLSDATESDLYKLKECPFDMGGYFIINGSEKVLIAQERSAGNIVQVF-KWAAPSPISHVA	236
tr A0A1X7QYA1 A0A1X7QYA1 9SACH	YLSDATESDLYKLKECPFDMGGY FIINGSEKVLIAQERSAGNIVQVF-KHAAPSPISHIA	234
tr J7RV95 J7RV95 KAZNA	YLSDATESDLYKLKECPFDMGGYFIINGSEKVLIAQERSAGNIVQVF-KKAAPSPISHVA	234
tr H2AVJ8 H2AVJ8 KAZAF	YLSDATESDLYKLKECPFDMGGYFIINGSEKVLIAOERSAGNIVQVF-KWAAPSPISHVA	234
sp Q6FLD5 RPB2_CANGA	YLSDATESDLYKLKEC PFDMGGY FIINGSEKVLIAQERSAGNIVQVF-KMAAPSPISHVA	236
sp P08518 RPB2_YEAST	YLSEATESDLYKLKECPFDMGGYFIINGSEKVLIAQERSAGNIVQVF-KMAAPSPISHVA	238
tr A0A0L8VHA5 A0A0L8VHA5 9SACH	YLSEATESDLYKLKEC PFDMGGY FIINGSEKVLIAQERSAGNIVQVF-KKAAPS PISHVA	238
tr A0A0L8RB33 A0A0L8RB33 SACEU	YLSEATESDLYKLKECPFDMGGYFIINGSEKVLIAQERSAGNIVQVF-KWAAPSPISHVA	238
tr G0VJ71 G0VJ71_NAUCC	YLSDATESDLYKLKECPFDMGGYFIINGSEKVLIAQERSAGNIVQVF-KMAAPSPISHVA	238
tr G8ZM49 G8ZM49_TORDC	YLSDATELDLYKLKECPFDMGGYFIINGSEKVLIAQERSAGNIVQVF-KKAAPSPISHVA	236
tr A0A1Q3A090 A0A1Q3A090 ZYGRO	YLSDATESDLYKLKECPFDMGGYFIINGSEKVLIAQERSAGNIVQVF-KMAAPSPISHVA	238
tr A0A0N7IS35 A0A0N7IS35 9SACH	YLSDATESDLYKLKECPFDMGGYFIINGSEKVLIAQERSAGNIVQVF-KMAAPSPISHVA	236
	YLSDATESDLYKLKECPFDMGGY FIINGSEKVLIAQERSAGNIVQVF-KWAAPSPISHVA	237
tr A0A212MG88 A0A212MG88_ZYGBA		
tr A0A1S7HHE1 A0A1S7HHE1_9SACH	YLSDATESDLYKLKECPFDMGGYFIINGSEKVLIAQERSAGNIVQVF-KMAAPSPISHVA	237
tr S6ESB4 S6ESB4 ZYGB2	YLSDATESDLYKLKECPFDMGGY FIINGSEKVLIAQERSAGNIVQVF-KHAAPSPISHVA	237
tr B6K5Q5 B6K5Q5 SCHJY	ILNGVSDSELYDLNECPYDQGGYFIINGSEKVIIAQERSAANIVQVF-RKAAPSPIAYVA	223
sp Q02061 RPB2 SCHPO	ILNGVSDSELYDLNECPYDQGGYFIINGSEKVIIAQERSAANIVQVF-KKAAPSPIAYVA	223
tr S9R8U4 S9R8U4_SCHOY	ILNGVSDAELYDLNECPYDQGGYFIINGSEKVIIAQERSAANIVQVF-RKAAPSPIALVA	223
tr S9W8C6 S9W8C6_SCHCR	ILNGVSDAELYDLNECPYDQGGYFIINGSEKVIIAQERSAANIVQVF-RWAAPSPVAYVA	223
tr S9W8C6 S9W8C6_SCHCR	ILNGVSDAELYDLNECPYDQGGYFIINGSEKVIIAQERSAANIVQVF-RHAAPSFVAYVA	223
-		
spiQ8RQE9 RPOB_THET8	ILNGVS DAE LYDLNEC PYDQGGY #IINGSEKVI IAQE RSAANI VQVF-RHAAPS PVA YVADGS #IINGADRVI VSQI #RS PGVYFT P DPAR PGRYIA	223 160
sp Q8RQE9 RPOB_THET8	dgs tiingadrvivsqihrs pgvyftp dparpgryia	160
sp Q8RQE9 RPOB_THET8 ASR51304.1	dgsfiingadrvivsqihrspgvyftpdparpgryia ngt vingtervivsqmhrspgvifdhdrgkthssgkylf	160 184
sp Q8RQE9 RPOB_THET8	DGS FIINGADRVIVSQIHRS PGVYFTP DPARPGRYIANGT FVINGTERVIVSQMHRS PGVLFDH DRQKTHS SGKYLF	160
sp Q8RQE9 RPOB_THET8 ASR51304.1 OXR47929.1	DGS FIINGADRVIVSQIHRS PGVYFTP DPARPGRYIANGT FVINGTERVIVSQMHRS PGVLFDH DRQKTHS SGKYLF	160 184 176
sp Q8RQE9 RPOB_THET8 ASR51304.1 OXR47929.1 WP_093971860.1	DGS FIINGADRVIVSQIHRS PGVYFTP DPARPGRYIANGT FVINGTERVIVSQHRS PGVLFDH DRGKTHS SGKYLFTGS FVINGTERVIVSQLHRS PGVFFEH DRGKTHS SGKLLFTGS FVINGTERVIVSQLHRS PGVFFEH DRGKTHS SGKLLF	160 184 176 176
sp Q8RQE9 RPOB_THET8 ASR51304.1 OXR47929.1 WP_093971860.1 sp Q2NWR6 RPOB_SODGM	DGS FIINGADRVIVSQIHRS PGVYFTP DPARPGRYIANGT FVINGTERVIVSQMHRS PGVLFDH DRQKTHS SGKYLF	160 184 176 176 172
sp Q8RQE9 RPOB_THET8 ASR51304.1 OXR47929.1 WP_093971860.1 sp Q2NWR6 RPOB_SODGM	DGS FIINGADRVIVSQIHRS PGVYFTP DPARPGRYIANGT FVINGTERVIVSQHRS PGVLFDH DRGKTHS SGKYLFTGS FVINGTERVIVSQLHRS PGVFFEH DRGKTHS SGKLLFTGS FVINGTERVIVSQLHRS PGVFFEH DRGKTHS SGKLLF	160 184 176 176
sp Q8RQE9 RPOB_THET8 ASR51304.1 OXR47929.1 WP_093971860.1 sp Q2NWR6 RPOB_SODGM sp B4EYU9 RPOB_PROMH	DGS FIINGADRVIVSQI HRS PGVYFTP DPARPGR YIANGT FVINGTERVIVSQHRS PGVLFDH DRGKTHS SGKYLFTGS FVINGTERVIVSQLHRS PGVFFEH DRGKTHS SGKLLFTGS FVINGTERVIVSQLHRS PGVFFEH DRGKTHS SGKLLFNGT FVINGTERVIVSQLHRS PGVFFDS DKGKTHS SGKVLY	160 184 176 176 172 172
sp Q8RQE9 RPOB_THET8 ASR51304.1 OXR47929.1 WP_093971860.1 sp Q2NWR6 RPOB_SODGM sp B4EYU9 RPOB_PROMH sp A7FNI3 RPOB_YERP3	DGS FIINGADRVIVSQI HRS PGVYFTP DPARPGRYIANGT FVINGTERVIVSQM HRS PGVLFDH DRG KTHS SGKYLFTGS FVINGTERVIVSQL HRS PGVFFEH DRG KTHS SGKLLFTGS FVINGTERVIVSQL HRS PGVFFEH DRG KTHS SGKLLFNGT FVINGTERVIVSQL HRS PGVFFDS DKG KTHS SGKVLYNGT FVINGTERVIVSQL HRS PGVFFDS DKG KTHS SGKVLYNGT FVINGTERVIVSQL HRS PGVFFDS DKG KTHS SGKVLY	160 184 176 176 172 172
sp Q8RQE9 RPOB_THET8 ASR51304.1 OXR47929.1 WP_093971860.1 sp Q2NWR6 RPOB_SODGM sp B4EYU9 RPOB_PROMH	DGS FIINGADRVIVSQI HRS PGVYFTP DPARPGR YIANGT FVINGTERVIVSQHRS PGVLFDH DRGKTHS SGKYLFTGS FVINGTERVIVSQLHRS PGVFFEH DRGKTHS SGKLLFTGS FVINGTERVIVSQLHRS PGVFFEH DRGKTHS SGKLLFNGT FVINGTERVIVSQLHRS PGVFFDS DKGKTHS SGKVLY	160 184 176 176 172 172
sp Q8RQE9 RPOB_THET8 ASR51304.1 OXR47929.1 WP_093971860.1 sp Q2NWR6 RPOB_SODGM sp B4EYU9 RPOB_PROMH sp A7FNI3 RPOB_YERP3 sp Q1C1U1 RPOB_YERPA	DGS FIINGADRVIVSQI HRS PGVYFTP DPARPGRYIANGT FVINGTERVIVSQM HRS PGVLFDH DRG KTHS SGKYLFTGS FVINGTERVIVSQL HRS PGV FFEH DRG KTHS SGKLLFTGS FVINGTERVIVSQL HRS PGV FFEH DRG KTHS SGKLLFNGT FVINGTERVIVSQL HRS PGV FFDS DKG KTHS SGKVLYNGT FVINGTERVIVSQL HRS PGV FFDS DKG KTHS SGKVLYNGT FVINGTERVIVSQL HRS PGV FFDS DKG KTHS SGKVLYNGT FVINGTERVIVSQL HRS PGV FFDS DKG KTHS SGKVLY	160 184 176 176 172 172 172
sp Q8RQE9 RPOB_THET8 ASR51304.1 OXR47929.1 WP_093971860.1 sp Q2NWR6 RPOB_SODGM sp B4EYU9 RPOB_PROMH sp A7FNI3 RPOB_YERP3 sp Q1C1U1 RPOB_YERPA sp B2K113 RPOB_YERPB	DGS FIINGADRVIVSQI HRS PGVYFTP DPARPGRYIANGT FVINGTERVIVSQM HRS PGVLFDH DRG KTHS SGKYLFTGS FVINGTERVIVSQL HRS PGVFFEH DRG KTHS SGKLLFTGS FVINGTERVIVSQL HRS PGVFFEH DRG KTHS SGKLLFNGT FVINGTERVIVSQL HRS PGVFFDS DKG KTHS SGKVLYNGT FVINGTERVIVSQL HRS PGVFFDS DKG KTHS SGKVLYNGT FVINGTERVIVSQL HRS PGVFFDS DKG KTHS SGKVLY	160 184 176 176 172 172 172 172 172
sp Q8RQE9 RPOB_THET8 ASR51304.1 OXR47929.1 WP_093971860.1 sp Q2NWR6 RPOB_SODGM sp B4EYU9 RPOB_PROMH sp A7FNI3 RPOB_YERP3 sp Q1C1U1 RPOB_YERPA	DGS FIINGADRVIVSQI HRS PGVYFTP DPARPGRYIANGT FVINGTERVIVSQM HRS PGVLFDH DRG KTHS SGKYLFTGS FVINGTERVIVSQL HRS PGV FFEH DRG KTHS SGKLLFTGS FVINGTERVIVSQL HRS PGV FFEH DRG KTHS SGKLLFNGT FVINGTERVIVSQL HRS PGV FFDS DKG KTHS SGKVLYNGT FVINGTERVIVSQL HRS PGV FFDS DKG KTHS SGKVLYNGT FVINGTERVIVSQL HRS PGV FFDS DKG KTHS SGKVLYNGT FVINGTERVIVSQL HRS PGV FFDS DKG KTHS SGKVLY	160 184 176 176 172 172 172
sp Q8RQE9 RPOB_THET8 ASR51304.1 OXR47929.1 WP_093971860.1 sp Q2NWR6 RPOB_SODGM sp B4EYU9 RPOB_PROMH sp A7FNI3 RPOB_YERP3 sp Q1C1U1 RPOB_YERPA sp B2K113 RPOB_YERPB sp A8G8E7 RPOB_SERP5		160 184 176 176 172 172 172 172 172 172
sp Q8RQE9 RPOB_THET8 ASR51304.1 OXR47929.1 WP_093971860.1 sp Q2NWR6 RPOB_SODGM sp B4EYU9 RPOB_PROMH sp A7FNI3 RPOB_YERP3 sp Q1C1U1 RPOB_YERPA sp B2K113 RPOB_YERPB sp A8G8E7 RPOB_SERP5 sp Q6DAN0 RPOB_PECAS		160 184 176 176 172 172 172 172 172 172 172
sp Q8RQE9 RPOB_THET8 ASR51304.1 OXR47929.1 WP_093971860.1 sp Q2NWR6 RPOB_SODGM sp B4EYU9 RPOB_PROMH sp A7FNI3 RPOB_YERP3 sp Q1C1U1 RPOB_YERPA sp B2K113 RPOB_YERPB sp A8G8E7 RPOB_SERP5		160 184 176 176 172 172 172 172 172 172
sp Q8RQE9 RPOB_THET8 ASR51304.1 OXR47929.1 WP_093971860.1 sp Q2NWR6 RPOB_SODGM sp B4EYU9 RPOB_PROMH sp A7FNI3 RPOB_YERP3 sp Q1C1U1 RPOB_YERPA sp B2K113 RPOB_YERPB sp A8G8E7 RPOB_SERP5 sp Q6DAN0 RPOB_PECAS sp C6DHR5 RPOB_PECCP		160 184 176 176 172 172 172 172 172 172 172
sp Q8RQE9 RPOB_THET8 ASR51304.1 OXR47929.1 WP_093971860.1 sp Q2NWR6 RPOB_SODGM sp B4EYU9 RPOB_PROMH sp A7FNI3 RPOB_YERP3 sp Q1C1U1 RPOB_YERPA sp B2K113 RPOB_YERPB sp A8G8E7 RPOB_SERP5 sp Q6DAN0 RPOB_PECAS sp C6DHR5 RPOB_PECCP sp Q7N9A4 RPOB_PHOLL		160 184 176 176 172 172 172 172 172 172 172 172 172
sp Q8RQE9 RPOB_THET8 ASR51304.1 OXR47929.1 WP_093971860.1 sp Q2NWR6 RPOB_SODGM sp B4EYU9 RPOB_PROMH sp A7FNI3 RPOB_YERP3 sp Q1C1U1 RPOB_YERPA sp B2K113 RPOB_YERPB sp A8G8E7 RPOB_SERP5 sp Q6DAN0 RPOB_PECAS sp C6DHR5 RPOB_PECCP sp Q7N9A4 RPOB_PHOLL sp C5BHE3 RPOB_EDWI9		160 184 176 176 172 172 172 172 172 172 172
sp Q8RQE9 RPOB_THET8 ASR51304.1 OXR47929.1 WP_093971860.1 sp Q2NWR6 RPOB_SODGM sp B4EYU9 RPOB_PROMH sp A7FNI3 RPOB_YERP3 sp Q1C1U1 RPOB_YERPA sp B2K113 RPOB_YERPB sp A8G8E7 RPOB_SERP5 sp Q6DAN0 RPOB_PECAS sp C6DHR5 RPOB_PECCP sp Q7N9A4 RPOB_PHOLL sp C5BHE3 RPOB_EDWI9		160 184 176 176 172 172 172 172 172 172 172 172 172 172
sp Q8RQE9 RPOB_THET8 ASR51304.1 OXR47929.1 WP_093971860.1 sp Q2NWR6 RPOB_SODGM sp B4EYU9 RPOB_PROMH sp A7FN13 RPOB_YERP3 sp Q1C1U1 RPOB_YERPA sp B2K113 RPOB_YERPB sp A8G8E7 RPOB_SERP5 sp Q6DAN0 RPOB_PECAS sp C6DHR5 RPOB_PECAS sp C6DHR5 RPOB_PECCP sp Q7N9A4 RPOB_PHOLL sp C5BHE3 RPOB_EDWI9 sp A7MQQ9 RPOB_CROS8		160 184 176 176 172 172 172 172 172 172 172 172 172 172
sp Q8RQE9 RPOB_THET8 ASR51304.1 OXR47929.1 WP_093971860.1 sp Q2NWR6 RPOB_SODGM sp B4EYU9 RPOB_PROMH sp A7FN13 RPOB_YERP3 sp Q1C1U1 RPOB_YERPA sp B2K113 RPOB_YERPB sp A8G8E7 RPOB_SERP5 sp Q6DAN0 RPOB_PECAS sp C6DHR5 RPOB_PECCP sp Q7N9A4 RPOB_PECCP sp Q7N9A4 RPOB_PHOLL sp C5BHE3 RPOB_EDWI9 sp A7MQQ9 RPOB_CROS8 sp B5XYF5 RPOB_KLEP3		160 184 176 176 172 172 172 172 172 172 172 172 172 172
sp Q8RQE9 RPOB_THET8 ASR51304.1 OXR47929.1 WP_093971860.1 sp Q2NWR6 RPOB_SODGM sp B4EYU9 RPOB_PROMH sp A7FN13 RPOB_YERP3 sp Q1C1U1 RPOB_YERPA sp B2K113 RPOB_YERPB sp A8G8E7 RPOB_SERP5 sp Q6DAN0 RPOB_PECAS sp C6DHR5 RPOB_PECAS sp C6DHR5 RPOB_PECCP sp Q7N9A4 RPOB_PHOLL sp C5BHE3 RPOB_EDWI9 sp A7MQQ9 RPOB_CROS8		160 184 176 176 172 172 172 172 172 172 172 172 172 172
sp Q8RQE9 RPOB_THET8 ASR51304.1 OXR47929.1 WP_093971860.1 sp Q2NWR6 RPOB_SODGM sp B4EYY9 RPOB_PROMH sp A7FN13 RPOB_YERP3 sp Q1C1U1 RPOB_YERPA sp B2K113 RPOB_YERPB sp A8G8E7 RPOB_SERP5 sp Q6DAN0 RPOB_PECAS sp C6DHR5 RPOB_PECCP sp Q7N9A4 RPOB_PECCP sp Q7N9A4 RPOB_PHOLL sp C5BHE3 RPOB_EDWI9 sp A7MQQ9 RPOB_CROS8 sp B5XYF5 RPOB_KLEP3 sp POA8V2 RPOB_ECOLI		160 184 176 176 172 172 172 172 172 172 172 172 172 172
sp Q8RQE9 RPOB_THET8 ASR51304.1 OXR47929.1 WP_093971860.1 sp Q2NWR6 RPOB_SODGM sp B4EYU9 RPOB_PROMH sp A7FN13 RPOB_YERP3 sp Q1C1U1 RPOB_YERPA sp B2K113 RPOB_YERPB sp A8G8E7 RPOB_SERP5 sp Q6DAN0 RPOB_PECAS sp C6DHR5 RPOB_PECCP sp Q7N9A4 RPOB_PHOLL sp C5BHE3 RPOB_EDWI9 sp A7MQQ9 RPOB_CROS8 sp B5XYF5 RPOB_KLEP3 sp POA8V2 RPOB_ECOLI sp C5AOS7 RPOB_ECOLI sp C5AOS7 RPOB_ECOLI		160 184 176 176 172 172 172 172 172 172 172 172 172 172
sp Q8RQE9 RPOB_THET8 ASR51304.1 OXR47929.1 WP_093971860.1 sp Q2NWR6 RPOB_SODGM sp B4EYY9 RPOB_PROMH sp A7FN13 RPOB_YERP3 sp Q1C1U1 RPOB_YERPA sp B2K113 RPOB_YERPB sp A8G8E7 RPOB_SERP5 sp Q6DAN0 RPOB_PECAS sp C6DHR5 RPOB_PECCP sp Q7N9A4 RPOB_PECCP sp Q7N9A4 RPOB_PHOLL sp C5BHE3 RPOB_EDWI9 sp A7MQQ9 RPOB_CROS8 sp B5XYF5 RPOB_KLEP3 sp POA8V2 RPOB_ECOLI		160 184 176 176 172 172 172 172 172 172 172 172 172 172
sp Q8RQE9 RPOB_THET8 ASR51304.1 OXR47929.1 WP_093971860.1 sp Q2NWR6 RPOB_SODGM sp B4EYU9 RPOB_PROMH sp A7FN13 RPOB_YERP3 sp Q1C1U1 RPOB_YERPA sp B2K113 RPOB_YERPB sp A8G8E7 RPOB_SERP5 sp Q6DAN0 RPOB_PECAS sp C6DHR5 RPOB_PECCP sp Q7N9A4 RPOB_PHOLL sp C5BHE3 RPOB_EDWI9 sp A7MQQ9 RPOB_CROS8 sp B5XYF5 RPOB_KLEP3 sp POA8V2 RPOB_ECOLI sp C5AOS7 RPOB_ECOLI sp C5AOS7 RPOB_ECOLI sp C5AOS7 RPOB_ECOLI sp C5AOS7 RPOB_ECOLU		160 184 176 176 172 172 172 172 172 172 172 172 172 172
sp Q8RQE9 RPOB_THET8 ASR51304.1 OXR47929.1 WP_093971860.1 sp Q2NWR6 RPOB_SODGM sp B4EYU9 RPOB_PROMH sp A7FNI3 RPOB_YERP3 sp Q1C1U1 RPOB_YERPA sp B2K113 RPOB_YERPB sp A8G8F7 RPOB_SERP5 sp Q6DAN0 RPOB_PECAS sp C6DHR5 RPOB_PECCP sp Q7N9A4 RPOB_PHOLL sp C5BHE3 RPOB_EDWIP9 sp A7MQQ9 RPOB_EDWIP9 sp A7MQQ9 RPOB_CROS8 sp B5XYF5 RPOB_KLEP3 sp POA8V2 RPOB_ECOLI sp C5AOS7 RPOB_ECOLI sp C5AOS7 RPOB_ECOLI sp C5AOS7 RPOB_SHIBS sp Q32AF9 RPOB_SHIDS		160 184 176 176 172 172 172 172 172 172 172 172 172 172
sp Q8RQE9 RPOB_THET8 ASR51304.1 OXR47929.1 WP_093971860.1 sp Q2NWR6 RPOB_SODGM sp B4EYU9 RPOB_PROMH sp A7FNI3 RPOB_YERP3 sp Q1C1U1 RPOB_YERPA sp B2K113 RPOB_YERPB sp A8G8E7 RPOB_SERP5 sp Q6DANO RPOB_PECAS sp C6DHR5 RPOB_PECCP sp Q7N9A4 RPOB_PHOLL sp C5BHE3 RPOB_EDW19 sp A7MQQ9 RPOB_EDW19 sp A7MQQ9 RPOB_CROS8 sp B5XYF5 RPOB_KLEP3 sp POA8V2 RPOB_ECOLI sp C5AOS7 RPOB_ECOLI sp C5AOS7 RPOB_ECOLI sp C5AOS7 RPOB_SHIBS sp Q32AF9 RPOB_SHIDS sp A8AKT9 RPOB_CITK8		160 184 176 176 172 172 172 172 172 172 172 172 172 172
sp Q8RQE9 RPOB_THET8 ASR51304.1 OXR47929.1 WP_093971860.1 sp Q2NWR6 RPOB_SODGM sp B4EYU9 RPOB_PROMH sp A7FNI3 RPOB_YERP3 sp Q1C1U1 RPOB_YERPA sp B2K113 RPOB_YERPB sp A8G8E7 RPOB_SERP5 sp Q6DANO RPOB_PECAS sp C6DHR5 RPOB_PECCP sp Q7N9A4 RPOB_PHOLL sp C5BHE3 RPOB_EDW19 sp A7MQQ9 RPOB_EDW19 sp A7MQQ9 RPOB_CROS8 sp B5XYF5 RPOB_KLEP3 sp POA8V2 RPOB_ECOLI sp C5AOS7 RPOB_ECOLI sp C5AOS7 RPOB_ECOLI sp C5AOS7 RPOB_SHIBS sp Q32AF9 RPOB_SHIDS sp A8AKT9 RPOB_CITK8		160 184 176 176 172 172 172 172 172 172 172 172
sp Q8RQE9 RPOB_THET8 ASR51304.1 OXR47929.1 WP_093971860.1 sp Q2NWR6 RPOB_SODGM sp B4EYU9 RPOB_PROMH sp A7FNI3 RPOB_YERP3 sp Q1C1U1 RPOB_YERPA sp B2K113 RPOB_YERPB sp A8G8E7 RPOB_SERP5 sp Q6DANO RPOB_PECAS sp C6DHR5 RPOB_PECCP sp Q7N9A4 RPOB_PHOLL sp C5BHE3 RPOB_EDW19 sp A7MQQ9 RPOB_CROS8 sp C6DHR5 RPOB_CROS8 sp B5XYF5 RPOB_KLEP3 sp POA8V2 RPOB_ECOLI sp C5AOS7 RPOB_ECOLI sp C5AOS7 RPOB_ECOLI sp C3AOS7 RPOB_SHIBS sp Q32AF9 RPOB_SHIDS sp A8AKT9 RPOB_CITK8 sp B5RFK1 RPOB_SALG2		160 184 176 176 172 172 172 172 172 172 172 172
Sp Q8RQE9 RPOB_THET8 ASR51304.1 OXR47929.1 WP_093971860.1 sp Q2NWR6 RPOB_SODGM sp B4EYU9 RPOB_PROMH sp A7FN13 RPOB_YERP3 sp Q1C1U1 RPOB_YERPA sp B2K113 RPOB_YERPA sp B2K113 RPOB_YERPB sp A8G8E7 RPOB_SERP5 sp Q6DAN0 RPOB_PECAS sp C6DHR5 RPOB_PECCP sp Q7N9A4 RPOB_PHOLL sp C5BHE3 RPOB_EDW19 sp A7MQQ9 RPOB_CROS8 sp B5XYF5 RPOB_KLEP3 sp POA8V2 RPOB_ECOLI sp C5AOS7 RPOB_ECOBW sp Q31U10 RPOB_SHIBS sp Q32AF9 RPOB_SHIDS sp A3AKT9 RPOB_CITK8 sp B5RFK1 RPOB_SALG2 sp B5RFK1 RPOB_SALG2 sp B5RFK1 RPOB_SALG2 sp B5RFK1 RPOB_SALG2 sp B5BJQ3 RPOB_SALPK		160 184 176 176 172 172 172 172 172 172 172 172
Sp Q8RQE9 RPOB_THET8 ASR51304.1 OXR47929.1 WP_093971860.1 sp Q2NWR6 RPOB_SODGM sp B4EYU9 RPOB_PROMH sp A7FN13 RPOB_YERP3 sp Q1C1U1 RPOB_YERPA sp B2K113 RPOB_YERPA sp B2K113 RPOB_YERPB sp A8G8E7 RPOB_SERP5 sp Q6DAN0 RPOB_PECAS sp C6DHR5 RPOB_PECCP sp Q7N9A4 RPOB_PHOLL sp C5BHE3 RPOB_EDW19 sp A7MQQ9 RPOB_CROS8 sp B5XYF5 RPOB_KLEP3 sp POA8V2 RPOB_ECOLI sp C5AOS7 RPOB_ECOBW sp Q31U10 RPOB_SHIBS sp Q32AF9 RPOB_SHIDS sp A3AKT9 RPOB_CITK8 sp B5RFK1 RPOB_SALG2 sp B5RFK1 RPOB_SALG2 sp B5RFK1 RPOB_SALG2 sp B5RFK1 RPOB_SALG2 sp B5BJQ3 RPOB_SALPK		160 184 176 176 172 172 172 172 172 172 172 172
Sp Q8RQE9 RPOB_THET8 ASR51304.1 OXR47929.1 WP_093971860.1 sp Q2NWR6 RPOB_SODGM sp B4EYU9 RPOB_PROMH sp A7FNI3 RPOB_YERP3 sp Q1C1U1 RPOB_YERP4 sp B2K113 RPOB_YERP5 sp A8G8E7 RPOB_SERP5 sp Q6DAN0 RPOB_PECAS sp C6DHR5 RPOB_PECCP sp Q7N9A4 RPOB_PHOLL sp C5BHE3 RPOB_EDW19 sp A7MQQ9 RPOB_CROS8 sp B5XYF5 RPOB_KLEP3 sp POA8V2 RPOB_ECOLI sp C5AOS7 RPOB_ECOBW sp Q31U10 RPOB_SHIDS sp A8AKT9 RPOB_SHIDS sp A8AKT9 RPOB_CITK8 sp B5RFK1 RPOB_SALG2 sp B5BJQ3 RPOB_SALPK sp B4T0Y9 RPOB_SALNS		160 184 176 176 172 172 172 172 172 172 172 172 172 172
Sp Q8RQE9 RPOB_THET8 ASR51304.1 OXR47929.1 WP_093971860.1 sp Q2NWR6 RPOB_SODGM sp B4EYU9 RPOB_PROMH sp A7FN13 RPOB_YERP3 sp Q1C1U1 RPOB_YERPA sp B2K113 RPOB_YERPA sp B2K113 RPOB_YERPB sp A8G8E7 RPOB_SERP5 sp Q6DAN0 RPOB_PECAS sp C6DHR5 RPOB_PECCP sp Q7N9A4 RPOB_PHOLL sp C5BHE3 RPOB_EDW19 sp A7MQQ9 RPOB_CROS8 sp B5XYF5 RPOB_KLEP3 sp POA8V2 RPOB_ECOLI sp C5AOS7 RPOB_ECOBW sp Q31U10 RPOB_SHIBS sp Q32AF9 RPOB_SHIDS sp A3AKT9 RPOB_CITK8 sp B5RFK1 RPOB_SALG2 sp B5RFK1 RPOB_SALG2 sp B5RFK1 RPOB_SALG2 sp B5RFK1 RPOB_SALG2 sp B5BJQ3 RPOB_SALPK		160 184 176 176 172 172 172 172 172 172 172 172
Sp Q8RQE9 RPOB_THET8 ASR51304.1 OXR47929.1 WP_093971860.1 sp Q2NWR6 RPOB_SODGM sp B4EYU9 RPOB_PROMH sp A7FNI3 RPOB_YERP3 sp Q1C1U1 RPOB_YERP4 sp B2K113 RPOB_YERP5 sp A8G8E7 RPOB_SERP5 sp Q6DAN0 RPOB_PECAS sp C6DHR5 RPOB_PECCP sp Q7N9A4 RPOB_PHOLL sp C5BHE3 RPOB_EDW19 sp A7MQQ9 RPOB_CROS8 sp B5XYF5 RPOB_KLEP3 sp POA8V2 RPOB_ECOLI sp C5AOS7 RPOB_ECOBW sp Q31U10 RPOB_SHIDS sp A8AKT9 RPOB_SHIDS sp A8AKT9 RPOB_CITK8 sp B5RFK1 RPOB_SALG2 sp B5BJQ3 RPOB_SALPK sp B4T0Y9 RPOB_SALNS		160 184 176 176 172 172 172 172 172 172 172 172 172 172

```
YSLATGNWGDQKKAHQARAGVSQVLNRLTFASTLSHLRRLNSP----IGR
sp|P30876|RPB2 HUMAN
                                                                                                            501
tr|G3V8Y5|G3V8Y5 RAT
                                      YSLATGNWGDQKKAHQARAGVSQVLNRLTFASTLSHLRRLNSP----IGF
                                                                                             LAKPE
                                                                                                            501
tr|A0A250Y753|A0A250Y753 CASCN
                                      YSLATGNWGDQKKAHQARAGVSQVLNRLTFASTLSHLRRLNSP-
                                                                                              LAKPE
                                                                                                            501
tr|A0A1U7R4C7|A0A1U7R4C7_MESAU
                                      YSLATGNWGDQKKAHQARAGVSQVLNRLTFASTLSHLRRLNSP-
                                                                                             LAKP
                                                                                                            501
                                      YSLATGNWGDQKKAHQARAGVSQVLNRLTFASTLSHLRRLNSP-
tr|A0A286XIQ9|A0A286XIQ9 CAVPO
                                                                                              LAKP
                                                                                                            501
tr||I3M351||I3M351 | ICTTR
                                      YSLATGNWGDQKKAHQARAGVSQVLNRLTFASTLSHLRRLNSP-
                                                                                             LAKP
                                                                                                            501
tr|G7P5R6|G7P5R6 MACFA
                                      YSLATGNWGDQKKAHQARAGVSQVLNRLTFASTLSHLRRLNSP
                                                                                             LAKP
                                                                                                            501
                                      YSLATGNWGDQKKAHQARAGVSQVLNRLTFASTLSHLRRLNSP-
                                                                                             LAKP
tr|H2QPI8|H2QPI8 PANTR
                                                                                                            501
                                      YSLATGNWGDQKKAHQARAGVSQVLNRLTFASTLSHLRRLNSP-
                                                                                             LAKPI
tr|A0A1U7V0T5|A0A1U7V0T5_TARSY
                                                                                                            501
tr|A0A1S2ZSL2|A0A1S2ZSL2_ERIEU
                                      YSLATGNWGDQKKAHQARAGVSQVLNRLTFASTLSHLRRLNSP
                                                                                             LAKPI
                                                                                                            501
 tr|A0A0D9QYL1|A0A0D9QYL1 CHLSB
                                      YSLATGNWGDQKKAHQARAGVSQVLNRLTFASTLSHLRRLNSP-
                                                                                              LAKP
                                                                                                            494
                                                                                             LAKPI
tr|A0A2K5ZNR7|A0A2K5ZNR7_MANLE
                                      YSLATGNWGDQKKAHQARAGVSQVLNRLTFASTLSHLRRLNSP-
                                                                                                            494
tr|A0A2I2ZIU3|A0A2I2ZIU3 GORGO
                                      YSLATGNWGDQKKAHQARAGVSQVLNRLTFASTLSHLRRLNSP
                                                                                           GKLAKPR
                                                                                                            501
                                      YSLATGNWGDQKKAHQARAGVSQVLNRLTFASTLSHLRRLNSP-
tr|A0A1D5QGA5|A0A1D5QGA5 MACMU
                                                                                                            501
tr|A0A2J8S2N1|A0A2J8S2N1 PONAB
                                      YSLATGNWGDQKKAHQARAGVSQVLNRLTFASTLSHLRRLNSP-
                                                                                                            501
tr|A0A2K5K5J5|A0A2K5K5J5 COLAP
                                     YSLATGNWGDQKKAHQARAGVSQVLNRLTFASTLSHLRRLNSP----IGRDCKLAKPR
                                                                                                           501
tr|A0A2J8PEW7|A0A2J8PEW7_PANTR
                                     YSLATGNWGDQKKAHQARAGVSQVLNRLTFASTLSHLRRLNSP----IG
                                                                                            KLAKP
                                                                                                           494
tr|A0A2K5CY83|A0A2K5CY83 AOTNA
                                     YSLATGNWGDQKKAHQARAGVSQVLNRLTFASTLSHLRRLNSP----IG
                                                                                            KLAKP
                                                                                                           501
tr|A0A096NEY4|A0A096NEY4 PAPAN
                                     YSLATGNWGDOKKAHOARAGVSOVLNRLTFASTLSHLRRLNSP--
                                                                                           KLAKP
                                                                                                           501
                                      YSLATGNWGDQKKAHQARAGVSQVLNRLTFASTLSHLRRLNSP--
tr|C9J2Y9|C9J2Y9_HUMAN
                                                                                           KT.AKE
                                                                                                           494
tr|G8BY61|G8BY61 TETPH
                                      YALATGNWGEQKKAMTSRAGVSQVLNRYTYSSTLSHLRRTNTP-
                                                                                           KLAKE
                                                                                                           512
                                     YALATGNWGEQKKAMSSRAGVSQVLNRYTYSSTLSHLRRTNTP----IG
                                                                                            KLAKP
tr|A0A1X7QYA1|A0A1X7QYA1 9SACH
                                                                                                           510
tr|J7RV95|J7RV95 KAZNA
                                     YALATGNWGEQKKAMSSRAGVSQVLNRYTYSSTLSHLRRTNTP--
                                                                                           KLAKP
                                                                                                           510
tr|H2AVJ8|H2AVJ8 KAZAF
                                      YALATGNWGEQKKAMSSRAGVSQVLNRYTYSSTLSHLRRTNTP-
                                                                                           KLAKP
                                                                                                           510
sp | Q6FLD5 | RPB2_CANGA
                                     YALATGNWGEQKKAMSSRAGVSQVLNRYTYSSTLSHLRRTNTP-
                                                                                            KLAKE
                                                                                                           512
sp|P08518|RPB2_YEAST
                                      YALATGNWGEQKKAMSSRAGVSQVLNRYTYSSTLSHLRRTNTP-
                                                                                           KLAKE
                                                                                                           514
                                     YALATGNWGEQKKAMSSRAGVSQVLNRYTYSSTLSHLRRTNTP----IG
tr|A0A0L8VHA5|A0A0L8VHA5 9SACH
                                                                                           KLAKP
                                                                                                           514
                                                                                                           514
tr|A0A0L8RB33|A0A0L8RB33 SACEU
                                     YALATGNWGEQKKAMSSRAGVSQVLNRYTYSSTLSHLRRTNTP--
                                                                                            KLAKP
tr|G0VJ71|G0VJ71 NAUCC
                                     YALATGNWGEQKKAMSSRAGVSQVLNRYTYSSTLSHLRRTNTP-
                                                                                           KLAKP
                                                                                                           514
tr|G8ZM49|G8ZM49 TORDC
                                      YALATGNWGEQKKAMSSRAGVSQVLNRYTYSSTLSHLRRTNTP-
                                                                                           KLAKP
                                                                                                           512
                                      YALATGNWGEQKKAMSSRAGVSQVLNRYTYSSTLSHLRRTNTP-
tr|A0A1Q3A090|A0A1Q3A090 ZYGRO
                                                                                           KLAKP
                                                                                                           514
tr|A0A0N7IS35|A0A0N7IS35 9SACH
                                     YALATGNWGEQKKAMSSRAGVSQVLNRYTYSSTLSHLRRTNTP----IG
                                                                                           KT.AKE
                                                                                                           512
tr|A0A212MG88|A0A212MG88 ZYGBA
                                      YALATGNWGEQKKAMSSRAGVSQVLNRYTYSSTLSHLRRTNTP-
                                                                                           KLAKP
                                                                                                           513
                                                                                     --IG
tr|A0A1S7HHE1|A0A1S7HHE1 9SACH
                                      YALATGNWGEQKKAMSSRAGVSQVLNRYTYSSTLSHLRRTNTP--
                                                                                          DGKLAKP
                                                                                                           513
tr|S6ESB4|S6ESB4 ZYGB2
                                     YALATGNWGEQKKAMSSRAGVSQVLNRYTYSSTLSHLRRTNTP----IG
                                                                                           KLAKP
                                                                                                           513
tr|B6K5Q5|B6K5Q5 SCHJY
                                     YSLATGNWGDQKRGLANRVGVSQVLNRYTFASTLSHLRRTNTP-
                                                                                          DGKLAKP
                                                                                                           500
                                     YSLATGNWGDQKRSMVNRVGVSQVLNRYTFASTLSHLRRTNTP--
                                                                                          DGKLAKP
sp|Q02061|RPB2_SCHPO
                                                                                                           500
tr|S9R8U4|S9R8U4 SCHOY
                                     YSLATGNWGDQKRSMMNRVGVSQVLNRYTFASTLSHLRRTNTP----IG
                                                                                          DCKLAKE
                                                                                                           500
tr|S9W8C6|S9W8C6 SCHCR
                                     YSLATGNWGDQKRSMMNRVGVSQVLNRYTFASTLSHLRRTNTP----IG
                                                                                          DOKTAKP
                                                                                                           500x8
sp|Q8RQE9|RPOB THET8
                                     -----EFFSRSQLSQFKDETNPLSSLRHKRRISALGPGGLTRERAGFDVRDV
                                                                                                           430x8
                                         -----effgssqlsqfmdqtnplaevthkrrvsalgpgglt<mark>rer</mark>agfevr<mark>dv</mark>
ASR51304.1
                                                                                                           564
OXR47929.1
                                         -----EFFGSSQLSQFMDQTNPLSEITHKRRVSALGPGGLT<mark>RER</mark>AGFEV<mark>RDV</mark>
                                                                                                           556
WP 093971860.1
                                         -----EFFGSSQLSQFMDQTNPLSEITHKRRVSALGPGGLT<mark>RER</mark>AGFEV<mark>RDV</mark>
                                                                                                           556
sp|Q2NWR6|RPOB SODGM
                                          -----EFFGSSQLSQFMDQNNPLSEITHKRRISALGPGGLTRERAGFEVR
                                                                                                           550
                                          -----EFFGSSQLSQFMDQNNPLSEITHKRRISALGPGGLT<mark>RER</mark>AGFEVR<mark>DV</mark>
sp | B4EYU9 | RPOB PROMH
                                                                                                           550
sp|A7FNI3|RPOB_YERP3
                                            -----EFFGSSQLSQFMDQNNPLSEITHKRRISALGPGGLT<mark>RER</mark>AGFEVR<mark>D</mark>V
                                                                                                           550
                                     -----EFFGSSQLSQFMDQNNPLSEITHKRRISALGPGGLTRERAGFEVRDV
sp | Q1C1U1 | RPOB YERPA
                                                                                                           550
                                     -----EFFGSSQLSQFMDQNNPLSEITHKRRISALGPGGLTRERAGFEVRDV
sp|B2K113|RPOB YERPB
                                                                                                           550
sp | A8G8E7 | RPOB SERP5
                                         -----EFFGSSQLSQFMDQNNPLSEITHKRRISALGPGGLT<mark>RER</mark>AGFEVR<mark>DV</mark>
                                                                                                           550
                                     -----EFFGSSQLSQFMDQNNPLSEITHKRRISALGPGGLTRERAGFEVRDV
sp | Q6DANO | RPOB PECAS
                                                                                                           550
sp | C6DHR5 | RPOB PECCP
                                     -----EFFGSSQLSQFMDQNNPLSEITHKRRISALGPGGLTRERAGFEVRDV
                                                                                                           550
sp | Q7N9A4 | RPOB PHOLL
                                     -----EFFGSSQLSQFMDQNNPLSEITHKRRISALGPGGLTRERAGFEVR
                                                                                                           550
sp | C5BHE3 | RPOB EDWI9
                                     -----EFFGSSOLSOFMDONNPLSEITHKRRISALGPGGLTRERAGFEVRD
                                                                                                           550
                                     -----EFFGSSQLSQFMDQNNPLSEITHKRRISALGPGGLTRERAGFEVRDV
sp|A7MQQ9|RPOB CROS8
                                                                                                           550
sp|B5XYF5|RPOB KLEP3
                                     -----EFFGSSQLSQFMDQNNPLSEITHKRRISALGPGGLTRERAGFEVRDV
                                                                                                           550
sp|P0A8V2|RPOB ECOLI
                                          -----EFFGSSQLSQFMDQNNPLSEITHKRRISALGPGGLTRERAGFEVRI
                                                                                                           550
                                     -----EFFGSSQLSQFMDQNNPLSEITHKRRISALGPGGLTRERAGFEVRDV
sp|C5A0S7|RPOB ECOBW
                                                                                                           550
sp|Q31U10|RPOB SHIBS
                                     -----EFFGSSQLSQFMDQNNPLSEITHKRRISALGPGGLT<mark>RER</mark>AGFEVR<mark>DV</mark>
                                                                                                           550
sp | Q32AF9 | RPOB SHIDS
                                     -----EFFGSSQLSQFMDQNNPLSEITHKRRISALGPGGLTRERAGFEVRDV
                                                                                                           550
                                     -----EFFGSSQLSQFMDQNNPLSEITHKRRISALGPGGLTRERAGFEVEDV
sp|A8AKT9|RPOB CITK8
                                                                                                           550
sp|B5RFK1|RPOB_SALG2
                                     -----EFFGSSQLSQFMDQNNPLSEITHKRRISALGPGGLTRERAGFEVRDV
                                                                                                           550
                                         -----EFFGSSQLSQFMDQNNPLSEITHKRRISALGPGGLT<mark>RER</mark>AGFEV<mark>RDV</mark>
sp|B5BJQ3|RPOB SALPK
                                                                                                           550
                                     -----EFFGSSQLSQFMDQNNPLSEITHKRRISALGPGGLTRERAGFEVRDV
sp|B4T0Y9|RPOB_SALNS
                                                                                                           550
sp|P06173|RPOB SALTY
                                     -----EFFGSSQLSQFMDQNNPLSEITHKRRISALGPGGLTRERAGFEVRDV
                                                                                                           550
                                                          :**. :. . : : * ** .:
```

```
sp|P30876|RPB2 HUMAN
                                   HNT
                                                  GHAVGLVKNLALMAYISVGSQPSPILEFLEEWSMENLEEISPAAI
                                                                                                     561
tr|G3V8Y5|G3V8Y5 RAT
                                   HNT
                                                   GHAVGLVKNLALMAYISVGSQPSPILEFLEEWSMENLEEISPAAI
                                                                                                     561
tr|A0A250Y753|A0A250Y753 CASCN
                                   HNT
                                                  GHAVGLVKNLALMAYISVGSQPSPILEFLEEWSMENLEEISPAAI
                                              ETPE
                                                                                                     561
tr|A0A1U7R4C7|A0A1U7R4C7 MESAU
                                              AETPE
                                                  GHAVGLVKNLALMAYISVGSQPSPILEFLEEWSMENLEEISPAAI
                                   HNT
                                                                                                     561
tr|A0A286XIQ9|A0A286XIQ9 CAVPO
                                             AETPEGHAVGLVKNLALMAYISVGSOPSPILEFLEEWSMENLEEISPAAI
                                   HNT
tr|I3M351|I3M351 ICTTR
                                   HNT
                                             AETPEGHAVGLVKNLALMAYISVGSQPSPILEFLEEWSMENLEEISPAAI
                                                                                                     561
tr|G7P5R6|G7P5R6 MACFA
                                   HNT
                                                  GHAVGLVKNLALMAYISVGSQPSPILEFLEEWSMENLEEISPAAI
                                                                                                     561
tr|H2QPI8|H2QPI8 PANTR
                                             AETPEGHAVGLVKNLALMAYISVGSQPSPILEFLEEWSMENLEEISPAAI
                                   HNT
                                             AETPEGHAVGLVKNLALMAYISVGSQPSPILEFLEEWSMENLEEISPAAI
tr|A0A1U7V0T5|A0A1U7V0T5 TARSY
                                                                                                     561
tr|A0A1S2ZSL2|A0A1S2ZSL2 ERIEU
                                   HNT
                                                  GHAVGLVKNLALMAYISVGSQPSPILEFLEEWSMENLEEISPAAI
                                                                                                     561
tr|A0A0D9QYL1|A0A0D9QYL1 CHLSB
                                   HNT
                                             AETPEGHAVGLVKNLALMAYISVGSQPSPILEFLEEWSMENLEEISPAAI
                                                                                                     554
tr|A0A2K5ZNR7|A0A2K5ZNR7 MANLE
                                   HNT
                                             AETPEGHAVGLVKNLALMAYISVGSQPSPILEFLEEWSMENLEEISPAAI
                                                                                                     554
tr|A0A2I2ZIU3|A0A2I2ZIU3 GORGO
                                   HNT
                                                  GHAVGLVKNLALMAYISVGSQPSPILEFLEEWSMENLEEISPAAI
                                                                                                     561
tr|A0A1D5QGA5|A0A1D5QGA5 MACMU
                                   HNT
                                              AETPEGHAVGLVKNLALMAYISVGSQPSPILEFLEEWSMENLEEISPAAI
                                                                                                     561
tr|A0A2J8S2N1|A0A2J8S2N1 PONAB
                                   HNT
                                             AETPEGHAVGLVKNLALMAYISVGSQPSPILEFLEEWSMENLEEISPAAI
                                                                                                     561
tr|A0A2K5K5J5|A0A2K5K5J5 COLAP
                                   HNT
                                                  GHAVGLVKNLALMAYISVGSQPSPILEFLEEWSMENLEEISPAAI
                                                                                                     561
tr|A0A2J8PEW7|A0A2J8PEW7 PANTR
                                   HNT
                                              AETPEGHAVGLVKNLALMAYISVGSOPSPILEFLEEWSMENLEEISPAAI
                                                                                                     554
                                             AETPEGHAVGLVKNLALMAYISVGSQPSPILEFLEEWSMENLEEISPAAI
tr|A0A2K5CY83|A0A2K5CY83 AOTNA
                                   HNT
                                                                                                     561
tr|A0A096NEY4|A0A096NEY4 PAPAN
                                                  GHAVGLVKNLALMAYISVGSQPSPILEFLEEWSMENLEEISPAAI
                                   HNT
                                              AETPE
                                                                                                     561
tr|C9J2Y9|C9J2Y9 HUMAN
                                   HNT
                                              AETPEGHAVGLVKNLALMAYISVGSOPSPILEFLEEWSMENLEEISPAAI
                                                                                                     554
                                             AETPEGQACGLVKNLSLMSCISVGTDPMPIITFLSEWGMEPLEDYVPHQS
tr|G8BY61|G8BY61 TETPH
                                   HNT
                                                                                                     572
                                              AETPEGQACGLVKNLSLMSCISVGTDPTPIITFLSEWGMEPLEDYVPHQS
tr|A0A1X7QYA1|A0A1X7QYA1 9SACH
                                   HNT
                                                                                                     570
                                                  GQACGLVKNLSLMSSISVGTDPMPIITFLSEWGLEPLEDYVPHQS
tr|J7RV95|J7RV95 KAZNA
                                   HNT
                                              AETPE
                                                                                                     570
tr|H2AVJ8|H2AVJ8 KAZAF
                                   HNT
                                                  GQACGLVKNLSLMSCISVGTDPMPIITFLSEWGMEPLEDYIPHQS
                                                                                                     570
sp|Q6FLD5|RPB2 CANGA
                                   HNT
                                              AETPE
                                                  GQACGLVKNLSLMSCISVGADPMPIITFLSEWGMEPLEDYVPHQS
                                                                                                     572
sp|P08518|RPB2_YEAST
                                                   GQACGLVKNLSLMSCISVGTDPMPIITFLSEWGMEPLEDYVPHQS
                                   HNT
                                                                                                     574
tr|A0A0L8VHA5|A0A0L8VHA5 9SACH
                                                   GQACGLVKNLSLMSCISVGTDPMPIITFLSEWGMEPLEDYVPHQS
                                   HNT
                                             AETPE
                                                                                                     574
tr|A0A0L8RB33|A0A0L8RB33 SACEU
                                   HNT
                                                  GOACGLVKNLSLMSCISVGTDPMPIITFLSEWGMEPLEDYVPHQS
                                                                                                     574
                                                  GQACGLVKNLSLMSCISVGTDPMPIITFLSEWGMEPLEDYVPHQS
tr|G0VJ71|G0VJ71 NAUCC
                                   HNT
                                              ETPE
                                                                                                     574
tr|G8ZM49|G8ZM49 TORDC
                                                  GQACGLVKNLSLMSCISVGTDPMPIITFLSEWGMEPLEDYVPHQS
                                   HNT
                                             AETPE
                                                                                                     572
                                             AETPEGQACGLVKNLSLMSCISVGTDPMPIITFLSEWGMEPLEDYIPHQS
tr|A0A1Q3A090|A0A1Q3A090 ZYGRO
                                   HNT
                                         LVC
                                                                                                     574
tr|A0A0N7IS35|A0A0N7IS35 9SACH
                                   HNT
                                                  GQACGLVKNLSLMSCISVGTDPMPIITFLSEWGMEPLEDYVPHQS
                                                                                                     572
                                             AETPEGQACGLVKNLSLMSCISVGTDPMPIITFLSEWGMEPLEDYVPHQS
tr|A0A212MG88|A0A212MG88 ZYGBA
tr|A0A1S7HHE1|A0A1S7HHE1 9SACH
                                   HNTHWGLVCPAETPEGOACGLVKNLSLMSCISVGTDPMPIITFLSEWGMEPLEDYVPHQS
                                          LVCPAETPECQACGLVKNLSIMSCISVGTDAMPIITFLSEWGMEPLEDYVPHOS
tr|S6ESB4|S6ESB4_ZYGB2
                                   HNTH
                                                                                                     573
tr|B6K5Q5|B6K5Q5 SCHJY
                                   HNTH
                                         GI/IVCP/
                                              ETPE
                                                   QACGLVKNLSLMSYVSVGSPSAPIIEFLEEWGMESLEDYNPSAS
                                                                                                     560
                                         GNVCPAETPEGOACGLVKNLSIMSYVSVGSPSAPIIEFLEEWGLETLEDYNPSAS
sp|Q02061|RPB2 SCHPO
                                   HNTH
                                                                                                     560
                                         GUVCPAET PEGQACGLVKNLALMSYVSVGSPAAPIIEFLEEWGLESLEDFNPSAS
tr|S9R8U4|S9R8U4 SCHOY
                                   HNTH
                                                                                                     560
                                              ETPEGQACGLVKNLALMSYVSVGSPAAPIIEFLEEWGLESLEDYNPSAS
tr|S9W8C6|S9W8C6 SCHCR
                                         MVCP
                                                                                                     560
                                   HRTHYGRICPVETPEGANIGLITSLAAYARVD-----ELGFI-----
                                                                                                     4677
sp|Q8RQE9|RPOB THET8
                                   HPTHYGRICPIETPEGPNIGLINSLASFSRVN-----KYGFI-----
ASR51304.1
                                   HPTHYGRVCPIETPEGPNIGLINSMALYARLN-----EYGFL-----
OXR47929.1
                                   HPTHYGRVCPIETPEGPNIGLINSMALYARLN-----EYGFL-----
WP 093971860.1
                                   HPTHYGRVCPIETPEGPNIGLINSLSVYAQTN-----EYGFL-----
sp|Q2NWR6|RPOB SODGM
                                   HPTHYGRVCPIETPEGPNIGLINSLSVYAQTN-----EYGFL-----
sp | B4EYU9 | RPOB PROMH
                                                                                                     587
                                   HPTHYGRVCPIETPEGPNIGLINSLSVYAQTN-----EYGFL-----
sp | A7FNI3 | RPOB YERP3
                                                                                                     587
                                   HPTHYGRVCPIETPEGPNIGLINSLSVYAOTN-----EYGFL-----
sp | Q1C1U1 | RPOB YERPA
                                   HPTHYGRVCPIETPEGPNIGLINSLSVYAQTN-----EYGFL-----
sp|B2K113|RPOB YERPB
                                   HPTHYGRVCPIETPEGPNIGLINSLSVYAQTN-----EYGFL-----
sp | A8G8E7 | RPOB SERP5
                                   HPTHYGRVCPIETPEGPNIGLINSLSVYAQTN-----EYGFL-----
sp | Q6DANO | RPOB PECAS
                                   HPTHYGRVCPIETPEGPNIGLINSLSVYAQTN-----EYGFL-------
sp | C6DHR5 | RPOB PECCP
sp | Q7N9A4 | RPOB PHOLL
                                                                                                     587
                                   HPTHYGRVCPIETPECPNIGLINSLSVYAQTN-----EYGFL-----
sp|C5BHE3|RPOB EDWI9
                                                                                                     587
                                   HPTHYGRVCPIETPEGPNIGLINSLSVYAQTN-----EYGFL-----
sp | A7MQQ9 | RPOB CROS8
                                   HPTHYGRVCPIETPEGPNIGLINSLSVYAQTN-----EYGFL-----
sp | B5XYF5 | RPOB KLEP3
                                                                                                     587
                                   HPTHYGRVCPIETPEGPNIGLINSLSVYAQTN-----EYGFL-------
HPTHYGRVCPIETPEGPNIGLINSLSVYAQTN-----EYGFL------
sp|P0A8V2|RPOB ECOLI
                                                                                                     587
sp | C5A0S7 | RPOB ECOBW
                                                                                                     587
                                   HPTHYGRVCPIETPEGPNIGLINSLSVYAQTN-----EYGFL-----
sp|Q31U10|RPOB SHIBS
                                                                                                     587
                                   HPTHYGRVCPIETPEGPNIGLINSLSVYAQTN-----EYGFL-----
sp | Q32AF9 | RPOB SHIDS
                                                                                                     587
                                   HPTHYGRVCPIETPEGPNIGLINSLSVYAQTN-----EYGFL-----
sp|A8AKT9|RPOB_CITK8
                                                                                                     587
                                   HPTHYGRVCPIETPEGPNIGLINSLSVYAQTN-----EYGFL----
sp|B5RFK1|RPOB SALG2
                                                                                                     587
                                   HPTHYGRVCPIETPEGPNIGLINSLSVYAQTN-----EYGFL-----
sp|B5BJQ3|RPOB_SALPK
                                                                                                     587
                                   HPTHYGRVCPIETPEGPNIGLINSLSVYAQTN-----EYGFL-----
sp | B4T0Y9 | RPOB SALNS
                                                                                                     587
                                   HPTHYGRVCPIETPEGPNIGLINSLSVYAQTN-----EYGFL-----
sp | P06173 | RPOB SALTY
                                   * * <mark>:*</mark> :** *****
                                                      **:..:: : .
```

```
sp|P30876|RPB2 HUMAN
                                    ILGVCASII PFPDHNOSPRNTYOSAMGKOAMGVYITNFHVRMDTLAH-------
                                                                                                       749
                                    ILGVCASII PFPDHNQSPRNTYQSAMGKQAMGVYIT NFHVRMDTLAH------
tr|G3V8Y5|G3V8Y5 RAT
                                                                                                       749
tr|A0A250Y753|A0A250Y753 CASCN
                                    ILGVCASII PFPDHNQSPRNTYQSAMGKQAMGVYITNFHVRMDTLAH--------
                                                DHNQSPRNTYQSAMGKQAMGVYITNFHVRMDTLAH------
tr|A0A1U7R4C7|A0A1U7R4C7 MESAU
                                    ILGVCASIIPFE
                                    ILGVCASII PFPDHNOSPRNTYOSAMGKOAMGVYITNFHVRMDTLAH------
tr|A0A286XIQ9|A0A286XIQ9 CAVPO
                                                                                                       749
                                    ILGVCASII PFPOHNQSPRNTYQSAMGKQAMGVYITNFHVRMDTLAH------ILGVCASII PFPOHNQSPRNTYQSAMGKQAMGVYITNFHVRMDTLAH------
tr|13M351|13M351_ICTTR
tr|G7P5R6|G7P5R6 MACFA
                                    ILGVCASII PFPDHNQSPRNTYQSAMGKQAMGVYITNFHVRMDTLAH------
tr|H2QPI8|H2QPI8 PANTR
                                    tr|A0A1U7V0T5|A0A1U7V0T5 TARSY
tr|A0A1S2ZSL2|A0A1S2ZSL2 ERIEU
                                    ILGVCASII PFPDHNQSPRNTYQSAMGKQAMGVYITNFHVRMDTLAH------
tr|A0A0D9QYL1|A0A0D9QYL1 CHLSB
                                    tr|A0A2K5ZNR7|A0A2K5ZNR7 MANLE
tr|A0A2I2ZIU3|A0A2I2ZIU3 GORGO
tr|A0A1D5QGA5|A0A1D5QGA5_MACMU
                                    ILGVCASII PFPDHNQSPRNTYQSAMGKQAMGVYITNFHVRMDTLAH-------
                                                                                                       749
                                    ILGVCASII PFPDHNQSPRNTYQSAMGKQAMGVYITNFHVRMDTLAH------
tr|A0A2J8S2N1|A0A2J8S2N1 PONAB
                                    ILGVCASIIPFPDHNQSPRNTYQSAMGKQAMGVYITNFHVRMDTLAH------
tr|A0A2K5K5J5|A0A2K5K5J5 COLAP
                                                DHNQSPRNTYQSAMGKQAMGVYITNFHVRMDTLAH------
tr|A0A2J8PEW7|A0A2J8PEW7 PANTR
                                    ILGVCASIIPFP
                                    ILGVCASII PFPDHNQSPRNTYQSAMGKQAMGVYITNFHVRMDTLAH------
tr|A0A2K5CY83|A0A2K5CY83 AOTNA
                                                                                                       749
tr | A0A096NEY4 | A0A096NEY4 PAPAN
                                    ILGVCASII PFPDHNQSPRNTYQSAMGKQAMGVYITNFHVRMDTLAH------
                                    ILGVCASIIPFPDHNQSPRNTYQSAMGKQAMGVYITNFHVRMDTLAH-------
tr|C9J2Y9|C9J2Y9 HUMAN
                                                                                                       742
                                    ILGVAASII PFPDHNQSPRNTYQSAMGKQAMGVFLTNYNVRMDTMAN-------
tr|G8BY61|G8BY61 TETPH
                                    ILGVAASIIPFP
                                                DHNQSPRNTYQSAMGKQAMGVFLTNYNVRMDTMAN------
tr|A0A1X7QYA1|A0A1X7QYA1 95ACH
                                    ILGVAASII PEPDHNQSPRNTYQSAMGKQAMGVFLTNYNVRMDTMAN------
tr|J7RV95|J7RV95 KAZNA
                                    ILGVAASII PFPDHNQSPRNTYQSAMGKQAMGVFLTNFNVRMDTMAN------
tr|H2AVJ8|H2AVJ8 KAZAF
                                    sp|Q6FLD5|RPB2 CANGA
sp | POS518 | RPB2 YEAST
tr|A0A0L8VHA5|A0A0L8VHA5 9SACH
                                    ILGVAASII PFPDHNQSPRNTYQSAMGKQAMGVFLTNYNVRMDTMAN------
                                                                                                       794
                                   tr|A0A0L8RB33|A0A0L8RB33 SACEU
                                                                                                       794
tr|G0VJ71|G0VJ71 NAUCC
                                    ILGVAASVIPFPDHNQSPRNTYQSAMGKQAMGVFLTNYNVRMDTMAN------
tr|G8ZM49|G8ZM49 TORDC
                                    ILGV AASII PFPDHNQSPRNTYQAAMGKQAMGVFLTNYNVRMDTMAN-----
tr|A0A1Q3A090|A0A1Q3A090 ZYGRO
tr|A0A0N7IS35|A0A0N7IS35_9SACH
                                    ILGVAASII PFPDHNQSPRNTYQSAMGKQAMGVFLTNYNVRMDTMAN------
                                    ILGVAASVI PFPDHNQSPRNTYQSAMGKQAMGVFLTNYNVRMDTMAN------
tr|A0A212MG88|A0A212MG88 ZYGBA
                                    ILGVAASVI PFPDHNQSPRNTYQSAMGKQAMGVFLTNYNVRMDTMAN-----
tr|A0A137HHE1|A0A137HHE1 9SACH
                                    ILGVAASVI PFPDHNQSPRNTYQSAMGKQAMGVFLTNYNVRMDTMAN------
tr|S6ESB4|S6ESB4 ZYGB2
                                    ILGILASII PFPDHNQSPRNTYQSAMGKQAMGIYLTNYQVRMDTMAN------
tr|B6K5O5|B6K5O5|SCHJY
                                    ILGILASII PFPDHNQSPRNTYQSAMGKQAMGVYLTNYQVRMDTMAN------
sp|Q02061|RPB2 SCHP0
                                           IIPFPDHNQSPRNTYQSAMGKQAMGVYLTNYQVRMDTMAN------
tr|39R8U4|39R8U4 3CHOY
                                    ILGILAS
                                    ILGILASII PFPDHNQSPRNTYQSAMGKQAMGVYLTNYQVRMDTMAN------
tr|39W8C6|39W8C6 3CHCR
                                    VFSVNTNLI PFLEHDDANRA LMGSNMQTQAV PLIRA QAPVVMTGLEE RVVRD SLAALY AE
sp|Q8RQE9|RPOB THET8
                                                                                                       598
                                    LVSVAASLI PFL<mark>ENDDA</mark>NRA LMGSNMQRQAV PLVQA EAPFV GTGMEE TVARD SGAAIA AK
ASR51304.1
                                                                                                       733
OXR47929.1
                                    IVSVAASLI PFL<mark>EHD DA</mark>NRA LMGANMQRQAV PCLRPEKTLVGTGIER TVAVD SGTTVQ AL
WP 093971860.1
                                    IVSVAASLI PFL<mark>EHD DA</mark>NRA LMGANMQRQAV PCLRP EKTLV GTGIER TVAVD SGTTVQ AL
                                    VVSVGASLI PFLEHDDANRA LMGANMOROAV PTLCT DKPLVGTGMER AVAVD SGVTAV AK
sp Q2NWR6|RPOB SODGM
                                                                                                       719
                                    VVSVGASLI PFL<mark>EHDDA</mark>NRA LMGANMQRQAV PTLRG DKPLV GTGMER AVAVD SGVTAV AK
sp|B4EYU9|RPOB PROMH
                                                                                                       719
                                    IVSVGASLI PFLEHDDANRA LMGANMORQAV PTLRA DKPLVGTGMER AVAVD SGVTSV AK
sp|A7FNI3|RPOB YERP3
                                                                                                       719
sp |Q1C1U1 |RPOB YERPA
                                    IVSVGASLI PFL<mark>EHD DA</mark>NRA LMGANMORQAV PTLRA DKPLVGTGMER AVAVD SGVTSV AK
                                   IVSV GAS LI PEL<mark>EHDDA</mark>NRA LMGANMORQAV PTLRA DKPLVGTGMER AVAVD SGVTSV AK
VVSV GAS LI PEL<mark>EHDDA</mark>NRA LMGANMORQAV PTLRA DKPLVGTGMER AVAVD SGVTAV AR
sp|B2K113|RPOB YERPB
                                                                                                       719
sp|A8G8E7|RPOB SERP5
                                                                                                       719
sp | Q6DAN 0 | RPOB PECAS
                                   IVSVGASLI PFL<mark>EHD DA</mark>NRA LMGANMQRQAV PTLRA DKPLVGTGMER AVAVD SGVTAV AK
                                                                                                       719
sp | C6DHR5 | RPOB PECCP
                                   VVSVGASLI PFLEHD DANRA LMGANMORQAV PTLRA DKPLVGTGMER AVAVD SGVTAV AK
                                                                                                       719
sp|Q7N9A4|RPOB PHOLL
                                   VVSVGASLI PFLEHDDANRA LMGANMOROAV PTLRA DKPLVGTGMER AVAVD SGVTSV AK
                                                                                                       719
sp|C5BHE3|RPOB EDWI9
                                   VVSVGASLI PFIEHDDANRA LMGANMOROAV PTLRA DKPLVGTGMER AVAVD SGVTSV AK
                                                                                                       719
sp|A7MQQ9|RPOB CROSS
                                   VVSVGASLI PFIEHD DANRA LMGANMOROAV PTLRA DKPLVGTGMER AVAVD SGVTAV AK
                                   VVSVGASLI PFLEHD DANRA LMGANMORQAV PTLRA DKPLV GTGMER AVAVD SGVTAV AK
sp|B5XYF5|RPOB KLEP3
                                                                                                       719
sp|P0A8V2|RPOB ECOLI
                                   VVSV GASLI PFLEHD DANRA LMGANMOROAV PTLRA DKPLV GTGMER AVAVD SGVTAV AK
                                   VVSVGASLI PFLEHDDANRA LMGANMOROAV PTLRA DKPLV GTGMER AVAVD SGVTAV AK
sp|C5A037|RPOB ECOBW
                                                                                                       719
                                   VVSV GASLI PELEHD DANRA LMGANMOROAV PTLRA DKPLV GTGMER AVAVD SGVTAVAK
sp|Q31U10|RPOB SHIBS
                                                                                                       719
sp|Q32AF9|RPOB SHIDS
                                   VVSV GASLI PFLEHD DANRA LMGANMOROAV PTLRA DKPLV GTGMER AVAVD SGVTAV AK
                                                                                                       719
                                   VVSV GASLI PFILEHD DANRA LMGANMORQAV PTLRA DKPLV GTGMER AVAVD SGVTAV AK
sp|A8AKT9|RPOB CITK8
                                   VVSVGASLI PFLEHDDANRA LMGANMOROAV PTLRA DKPLVGTGMER AVAVD SGVTAV AK
sp|B5RFK1|RPOB SALG2
                                                                                                       719
                                   VVSVGASLI PFLEHDDANRA LMGANMORQAV PTLRA DKPLVGTGMER AVAVD SGVTAV AK
sp|B5BJQ3|RPOB SALPK
                                                                                                       719
sp|B4T0Y9|RPOB SALNS
                                   VVSVGASLI PFIEHDDANRA LMGANMQRQAV PTLRA DKPLVGTGMER AVAVD SGVTAV AK
                                                                                                      719
                                   VVSVGASLI PFLEHD DANRA LMGANMORQAV PTLRA DKPLVGTGMER AVAVD SGVTAV AK
sp|P06173|RPOB SALTY
                                                                                                      719
```

```
----SMEYLRFRELPAGINSIVAIASYTGYNQEDSVIMNRSAVDRGFFRSVFY
sp/P30876/RPB2 HUMAN
                                                                                                                        817
                                          ----SMEYLRFRELPAGINSIVAIASYTOMOEDSVIMNRSAVDRGFF
tr|G3V8Y5|G3V8Y5 RAT
tr|A0A250Y753|A0A250Y753 CASCN
                                                                                                             KEO
                                                                                                                        817
                                          ----SMEYLRFRELPAGINSIVAIASYTGYNOEDSVIMNRSAVDRGFFF
tr|A0A1U7R4C7|A0A1U7R4C7 MESAU
                                                                                                             KEQ
                                                                                                                        817
tr|A0A286XIQ9|A0A286XIQ9_CAVPO
                                          ----SMEYLRFRELPAGINSIVAIASYTGYNQEDSVIMNRSAVDRGFFR
                                                                                                             KEQ
                                                                                                                        817
                                          ----SMEYLRFRELPAGINSIVAIASYT<mark>OYMQEDS</mark>VIMNRSAVDRGFF
----SMEYLRFRELPAGINSIVAIASYT<mark>OYMQEDS</mark>VIMNRSAVDRGFF
tr|I3M351|I3M351 ICTTR
                                                                                                             KEO
                                                                                                                        817
tr|G7P5R6|G7P5R6 MACFA
                                                                                                             KE0
                                                                                                                        817
                                          ----SMEYLRFRELPAGINSIVAIASYTGYNQEDSVIMNRSAVDRGFFF
tr|H2QPI8|H2QPI8 PANTR
                                                                                                             KEQ
                                                                                                                        817
tr|A0A1U7V0T5|A0A1U7V0T5_TARSY
                                          ----SMEYLRFRELPAGINSIVAIASYTGYNQEDSVIMNRSAVDRGFF
                                                                                                             KEO
                                                                                                                        817
                                          ----SMEYLRFRELPAGINSIVAIASYT<mark>CYNQEDS</mark>VIMNRSAVDRGFFF
----SMEYLRFRELPAGINSIVAIASYT<mark>CYNQEDS</mark>VIMNRSAVDRGFFF
tr|A0A1S2ZSL2|A0A1S2ZSL2_ERIEU
tr|A0A0D9QYL1|A0A0D9QYL1_CHLSB
                                                                                                             KE<sub>0</sub>
                                                                                                             KEO
                                                                                                                        810
                                          ----SMEYLRFRELPAGINSIVAIASYTGYNQEDSVIMNRSAVDRGFFF
tr|A0A2K5ZNR7|A0A2K5ZNR7 MANLE
                                                                                                             KEQ
                                                                                                                        810
tr|A0A2I2ZIU3|A0A2I2ZIU3_GORGO
                                          ----SMEYLRFRELPAGINSIVAIASYTGYNQEDSVIMNRSAVDRGFFF
                                                                                                             KEO
                                                                                                                        817
                                           ----SMEYLRFRELPAGINSIVAIASYT<mark>GYNQEDSV</mark>IMNRSAVDRGFF
                                                                                                             KEQ
tr|A0A1D5QGA5|A0A1D5QGA5 MACMU
                                                                                                                        817
                                          ----SMEYLRFRELPAGINSIVAIASYT<mark>GYNQEDS</mark>VIMNRSAVDRGFFF
tr|A0A2J8S2N1|A0A2J8S2N1_PONAB
                                                                                                             KE0
                                                                                                                        817
                                          ----SMEYLRFRELPAGINSIVAIASYTGYNQEDSVIMNRSAVDRGFF
tr|A0A2K5K5J5|A0A2K5K5J5 COLAP
                                                                                                             KEQ
                                                                                                                        817
tr|A0A2J8PEW7|A0A2J8PEW7_PANTR
                                          ----SMEYLRFRELPAGINSIVAIASYTGYNQEDSVIMNRSAVDRGFFF
                                                                                                             KE0
                                                                                                                        810
                                           ----SMEYLRFRELPAGINSIVAIASYT<mark>GYNQEDS</mark>VIMNRSAVDRGFFF
tr|A0A2K5CY83|A0A2K5CY83 AOTNA
                                                                                                             KEQ
                                                                                                                        817
                                          -----SMEYLRFRELPAGINSIVAIASYT<mark>GYNOEDS</mark>VIMNRSAVDRGFFF
tr|A0A096NEY4|A0A096NEY4 PAPAN
                                                                                                             KE0
                                                                                                                        817
tr|C9J2Y9|C9J2Y9 HUMAN
                                           ----SMEYLRFRELPAGINSIVAIASYT<mark>GYNQEDS</mark>VIMNRSAVDRGFFF
                                                                                                             KEQ
                                                                                                                        810
                                                                                                             MDQ
                                          ----AMEYLKFRELPAGQNAIVAIACYS<mark>GYNQEDSM</mark>IMNQSSIDRGLFR
tr|G8BY61|G8BY61_TETPH
                                                                                                                        862
tr|A0A1X7QYA1|A0A1X7QYA1 9SACH
                                           ----SMEYLKFRELPAGQNAIVAIACYS
                                                                            GYNQEDSMIMNQSSIDRGLFF
                                                                                                             MDO
                                                                                                                        859
                                                                            <mark>GYNQEDS</mark>MIMNQSSIDRGLFR
GYNQEDSMIMNQSSIDRGLFR
tr|J7RV95|J7RV95 KAZNA
                                          ----AMEYLKFRELPAGQNAIVAIACYS
                                                                                                                        858
tr|H2AVJ8|H2AVJ8_KAZAF
                                           ----AMEYLKFRELPAGQNAIVAIACYS
                                                                                                                        860
sp|Q6FLD5|RPB2 CANGA
                                           ----AMEYLKFRELPAGQNAIVAIACYS<mark>GYNQEDS</mark>MIMNQSSIDRGLF<mark>RSL</mark>FFI
                                                                                                             MDO
                                                                                                                        861
sp|P08518|RPB2_YEAST
                                               -AMEYLKFRELPAGQNAIVAIACYS
                                                                            GYNQEDSMIMNQSSIDRGLFRSLFFRS
tr|A0A0L8VHA5|A0A0L8VHA5_9SACH
                                                                            <mark>GYNQEDSI</mark>/IMNQSSIDRGLF<mark>RSI</mark>LFFI
                                          ----AMEYLKFRELPAGQNAIVAIACYS
                                                                                                                        862
tr|A0A0L8RB33|A0A0L8RB33 SACEU
                                          ----AMEYLKFRELPAGQNAIVAIACYS
                                                                            GYNQEDSMIMNQSSIDRGLFR
                                                                                                                        862
                                          ----SMEYLKFRELPAGQNAIVAIACYSGYNQEDSMIMNQSSIDRGLFR
                                                                                                             MDQ
tr|G0VJ71|G0VJ71 NAUCC
                                                                                                                        862
                                                                            GYNQEDSMIMNQSSIDRGLFF
                                                                                                             MDQ
tr|G8ZM49|G8ZM49 TORDC
                                           ----AMEYLKFRELPAGQNAIVAIACYS
                                                                                                                        860
                                          ----AMEYLKFRELPAGQNAIVAIACYS<mark>GYNQEDG</mark>MIMNQSSIDRGLF<mark>R</mark>
tr|A0A1Q3A090|A0A1Q3A090_ZYGRO
                                                                                                                        862
tr|A0A0N7IS35|A0A0N7IS35 9SACH
                                           ----AMEYLKFRELPAGQNAIVAIACYS<mark>GYNQEDS</mark>NIMNQSSIDRGLF<mark>RSL</mark>FF
                                                                                                                        860
tr|A0A212MG88|A0A212MG88_ZYGBA
                                          ----AMEYLKFRELPAGONAIVAIACYSGYNOEDSMIMNOSSIDRGLFR
                                                                                                             MDQ
                                                                                                                        861
                                                                            GYNQEDSMIMNQSSIDRGLFF
                                                                                                             MDQ
tr|A0A1S7HHE1|A0A1S7HHE1 9SACH
                                           ----AMEYLKFRELPAGQNAIVAIACYS
                                                                                                                        861
                                          ----AMEYLKFRELPAGQNAIVAIACYSGYNQEDSMIMNQSSIDRGLFR
tr|S6ESB4|S6ESB4_ZYGB2
                                                                                                             MDO
                                                                                                                        861
                                          ----SMEYLKFRELPAGQNAIVAILCYSGYNQEDSIIMNQSSIDRGLFR
tr|B6K5Q5|B6K5Q5 SCHJY
                                                                                                             MDQ
                                                                                                                        851
                                          ----SMEYLKFRELPAGQNAIVAILCYSGYNQEDSIIMNQASIDRGLFR
sp|Q02061|RPB2 SCHPO
                                                                                                             TDQ
                                                                                                                        851
tr|S9R8U4|S9R8U4 SCHOY
                                          ----SMEYLKFRELPAGQNAIVAILCYS
                                                                            GYNQEDSIIMNQASIDRGLF
                                                                                                             TDO
                                                                                                                        851
                                          ----SMEYLKFRELPAGQNAIVAILCYSGYNQEDSIIMNQASIDRGLFRS
tr|S9W8C6|S9W8C6 SCHCR
                                                                                                                        851x6
                                          GDLLADGPASENGFLALGQNVLVAIMPFDGYNFEDATVISEELLKRDFYTSIHIERYEIE
sp|Q8RQE9|RPOB THET8
                                                                                                                          711
                                          GDIIADGPSTDLGELALGRNALVAFMPWN<mark>GYNYEDS</mark>ILISERIVKDDVFTSIHIEEFEVM
ASR51304.1
                                                                                                                          852
                                          GDVLADGASTDLGELALGONMLIAFMPWNGYNFEDSILISEKVVADDRYTSIHTEELTVV
GDVLADGASTDLGELALGONMLIAFMPWNGYNFEDSILISEKVVADDRYTSIHIEELTVV
OXR47929.1
                                                                                                                          845
WP 093971860.1
                                                                                                                          845
sp|Q2NWR6|RPOB SODGM
                                          GDVLADGPSTDLGELALGQNMRIAFMPWNGYNFEDSMLVSERVVQEDRFTTIHIQELACV
                                                                                                                          839
sp|B4EYU9|RPOB PROMH
                                          GDVLADGPSTDLGELALGQNMRVAFMPWN<mark>GYNYEDS</mark>ILVSERVVQEDRFTTIHIQELACV
                                                                                                                          839
                                          GDVLADGPSTDLGELALGQNMRVAFMPWNGYNFEDSILVSERVVQEDRFTTIHIQELACV
GDVLADGPSTDLGELALGQNMRVAFMPWNGYNFEDSILVSERVVQEDRFTTIHIQELACV
sp|A7FNI3|RPOB YERP3
                                                                                                                          839
sp|Q1C1U1|RPOB_YERPA
                                                                                                                          839
sp|B2K113|RPOB YERPB
                                          GDVLADGPSTDLGELALGQNMRVAFMPWNGYNFEDSILVSERVVQEDRFTTIHIQELACV
                                                                                                                          839
                                          GDVLADGPSTDLGELALGQNMRVAFMPWNGYNFEDS ILVSERVVQEDRFTTIHIQELACV
sp|A8G8E7|RPOB SERP5
                                                                                                                          839
                                          GDVLADGPSTDLGELALGQNMRVAFMPWNGYNFEDSILVSERVVQEDRFTTIHIQELACV
sp|Q6DAN0|RPOB PECAS
                                                                                                                          839
                                          GDVLADGPSTDLGELALGQNMRVAFMPWNGYNFEDSILVSERVVQEDRFTTIHIQELACV
sp | C6DHR5 | RPOB PECCP
                                                                                                                          839
                                          GDVLADGPSTDLGELALGQNMRVAFMPWNGYNFEDSILVSERVVQEDRFTTIHIQELACV
sp|Q7N9A4|RPOB PHOLL
                                                                                                                          839
                                          GDVLADGPSTDLGELALGQNMRVAFMPWNGYNFEDS
sp|C5BHE3|RPOB_EDWI9
                                                                                                                          839
                                          GDVLADGPSTDLGELALGQNMRVAFMPWNGYNFEDSILVSERVVQEDRFTTIHIQELACV
GDVLADGPSTDLGELALGQNMRVAFMPWNGYNFEDSILVSERVVQEDRFTTIHIQELACV
sp|A7MQQ9|RPOB CROS8
                                                                                                                          839
sp|B5XYF5|RPOB KLEP3
                                                                                                                          839
sp|P0A8V2|RPOB ECOLI
                                           GDVLADGPSTDLGELALGONMRVAFMPWNGYNFEDSILVSERVVQEDRFTTIHIQELACV
                                                                                                                          839
                                          GDVLADGPSTDLGELALGQNMRVAFMPWNGYNFEDSILVSERVVQEDRFTTIHIQELACV
sp|C5A0S7|RPOB ECOBW
                                                                                                                          839
                                           GDVLADGPSTDLGELALGQNMRVAFMPWNGYNFEDSILVSERVVQEDRFTTIHIQELACV
sp|Q31U10|RPOB_SHIBS
                                                                                                                          839
                                          GDVLADGPSTDLGELALGQNMRVAFMPWNGYNFEDSILVSERVVQEDRFTTIHIQELACV
sp|Q32AF9|RPOB SHIDS
                                                                                                                          839
                                           GDVLADGPSTDLGELALGQNMRVAFMPWNGYNFEDSILVSERVVQEDRFTTIHIQELACV
sp|A8AKT9|RPOB CITK8
                                                                                                                          839
                                          GDVLADGPSTDLGELALGQNMRVAFMPWNGYNFEDSILVSERVVQEDRFTTIHIQELACV
GDVLADGPSTDLGELALGQNMRVAFMPWNGYNFEDSILVSERVVQEDRFTTIHIQELACV
GDVLADGPSTDLGELALGQNMRVAFMPWNGYNFEDSILVSERVVQEDRFTTIHIQELACV
sp|B5RFK1|RPOB_SALG2
                                                                                                                          839
                                                                                                                          839
sp|B5BJQ3|RPOB SALPK
sp|B4T0Y9|RPOB SALNS
                                                                                                                          839
sp|P06173|RPOB SALTY
                                          GDVLADGPSTDLGELALGQNMRVAFMPWN<mark>GYNFEDS</mark>ILVSERVVQEDRFTTIHIQELACV
                                                                                                                          839
                                                           . . . . . . .
                                                                              *** **::::. : . : ::. .
```

```
sp|P30876|RPB2 HUMAN
                                                   <mark>G</mark>FDQEEVFEKPTRETCQGMRHAIYDK<mark>LDDDGL</mark>IAPGVRVSGDDVIIGKTVTLPENE
                                                                                                                                    877x9
                                                    FDQEEVFEKPTRETCQGMRHAIYDK<mark>LDDDGL</mark>IAPGVRVSGDDVIIGKTVTLPENE
tr|G3V8Y5|G3V8Y5 RAT
                                              ESE
                                                                                                                                    877
tr|A0A250Y753|A0A250Y753 CASCN
                                              ESK
                                                    <mark>g</mark>fdqeevfekptretcqgmrhaiydk<mark>ldddgl</mark>iapgvrvsgddviigktvtlpene
                                                                                                                                    877
                                                    <mark>g</mark>fd<u>oe evfekptretcogmrhaiydk<mark>loddgl</mark>iapgvrvsgddviigktvtlpene</u>
tr|A0A1U7R4C7|A0A1U7R4C7 MESAU
                                              ESE
                                                                                                                                    877
                                                    FDQEEVFEKPTRETCQGMRHAIYDK<mark>LDDDGL</mark>IAPGVRVSGDDVIIGKTVTLPENE
tr|A0A286XIQ9|A0A286XIQ9 CAVPO
                                              ESE
                                                                                                                                    877
                                                    <mark>F</mark>DQEEVFEKPTRETCQGMRHAIYDK<mark>LDDDGL</mark>IAPGVRVSGDDVIIGKTVTLPENE
tr|I3M351|I3M351 ICTTR
                                                                                                                                    877
                                              ESK
tr|G7P5R6|G7P5R6 MACFA
                                                    <mark>g</mark>fdqeevfekptretcqgmrhaiydk<mark>loddgl</mark>iapgvrvsgddviigktvtlpene
                                                                                                                                    877
tr|H2QPI8|H2QPI8_PANTR
                                              ESK
                                                    <mark>g</mark>fdqeevfekptretcqgmrhaiydk<mark>ldddgl</mark>iapgvrvsgddviigktvtlpene
                                                                                                                                    877
                                                    FDQEEVFEKPTRETCQGMRHAIYDK<mark>LDDDGL</mark>IAPGVRVSGDDVIIGKTVTLPENE
tr|A0A1U7V0T5|A0A1U7V0T5 TARSY
                                              ESF
                                                                                                                                    877
tr|A0A1S2ZSL2|A0A1S2ZSL2 ERIEU
                                                    <mark>g</mark>fdqe evfe kptretcqgmrha iydk<mark>loddgl</mark>i apgvrvsgddvi i gktvtlpene
                                                                                                                                    877
                                              ESK
                                                    <mark>g</mark>fd<u>oe</u>evfekptretcogmrhaiydk<mark>iddogl</mark>iapgvrvsgd<del>dviig</del>ktvtlpene
<mark>g</mark>fdoeevfekptretcogmrhaiydk<mark>iddogl</mark>iapgvrvsgddviigktvtlpene
tr|A0A0D9QYL1|A0A0D9QYL1 CHLSB
                                                                                                                                    870
tr|A0A2K5ZNR7|A0A2K5ZNR7_MANLE
                                              ESK
                                                                                                                                    870
                                                     FDQE EVFEKPT RETCQGMRHA I YDK<mark>LDDDGL</mark>I APGVRV SGI
tr|A0A2I2ZIU3|A0A2I2ZIU3 GORGO
                                              ESK
                                                                                                          DVIIGKTVTLPENE
                                                                                                                                    877
tr|A0A1D5QGA5|A0A1D5QGA5 MACMU
                                                    FDOE EVFEKPTRETCOGMRHA IYDK LODDGLI APGVRVSGDDVII GKTVTLPENE
                                                                                                                                    877
                                              ESK
tr|A0A2J8S2N1|A0A2J8S2N1_PONAB
                                                    <mark>g</mark>fdqeevfekptretcqgmrhaiydk<mark>loddgl</mark>iapgvrvsgddviigktvtlpene
                                                                                                                                    877
tr|A0A2K5K5J5|A0A2K5K5J5_COLAP
tr|A0A2J8PEW7|A0A2J8PEW7_PANTR
                                                     FDQEEVFEKPTRETCQGMRHAIYDK<mark>LDODGL</mark>IAPGVRVSGDDVIIGKTVTLPENE
FDQEEVFEKPTRETCQGMRHAIYDK<mark>LDODGL</mark>IAPGVRVSGDDVIIGKTVTLPENE
                                              ESK
                                                                                                                                    877
                                              ESK
                                                                                                                                    870
                                                    FDOE EVFEKPTRETCOGMRHA IYDK<mark>LDDDGL</mark>I APGVRVSGDDVI I GKTVTLP ENE
tr|A0A2K5CY83|A0A2K5CY83 AOTNA
                                              ESK
                                                                                                                                    877
                                                     FDQEEVFEKPTRETCQGMRHAIYDK<mark>LDDDGL</mark>IAPGVRVSGI
                                                                                                         DVIIGKTVTLPENE
tr|A0A096NEY4|A0A096NEY4 PAPAN
                                              ESK
                                                                                                                                    877
                                                    GFDQEEVFEKPTRETCQGMRHAIYDK<mark>IDDDGL</mark>IAPGVRVSGDDVIIGKTVTLPENE
G<mark>d</mark>SITETFEKPQRTNTLRMKHGSYDK<mark>IDDDGL</mark>ISPGVRVSGEDIIIGKTTPISPDE
tr|C9J2Y9|C9J2Y9 HUMAN
                                              ESK.
                                                                                                                                    870
tr|G8BY61|G8BY61_TETPH
                                              EKK
                                                                                                                                    922
                                                   <mark>/G</mark>MSITETFEKPQRTNTLRMKHGTYDK<mark>LDDDGL</mark>IAPGVRVSGEDVIIGKTTPISPDE
tr|A0A1X7QYA1|A0A1X7QYA1 9SACH
                                                                                                                                    919
                                                    <mark>g</mark>asitetfekportntlrmkhgtydk<mark>lodogl</mark>iapgvrvsgebiiigkttpispde
tr|J7RV95|J7RV95_KAZNA
                                              EKK'
                                                                                                                                    918
                                                    <mark>G</mark>MSITETFEKPHRTNTLRMKHGTYDK<mark>LDDDGL</mark>IAPGVRVSGE
<mark>G</mark>MSITETFEKPQRTNTLRMKHGTYDK<mark>LDEDGL</mark>IAPGVRVSGE
                                                                                                         DIIIGKTTPISPDE
DIIIGKTTPIAPDE
tr|H2AVJ8|H2AVJ8 KAZAF
                                                                                                                                    920
sp/Q6FLD5/RPB2 CANGA
                                                                                                                                    921
                                              EKKYGMSITETFEKPORTNTLRMKHGTYDKLDDDGLIAPGVRVSGEDVIIGKTTPISPDE
sp|P08518|RPB2 YEAST
                                                                                                                                    922
                                              EKK YGUSITETFEKPQRTNTLRMKHGTYDK DDDGL APGVRVSGEDVIIGKTTPISPDE
tr|A0A0L8VHA5|A0A0L8VHA5 9SACH
                                                                                                                                    922
tr|A0A0L8RB33|A0A0L8RB33 SACEU
                                                     MSITETFEKPORTNTLRMKHGTYDK<mark>LDDDGL</mark>IAPGVRVSGE
                                                                                                         DVIIGKTTPISPDE
                                                                                                                                    922
                                                    <mark>e</mark>nsitetfekportntlrmkhgtydk<mark>loddgl</mark>iapgvrvsgediiigkttpispde
tr|G0VJ71|G0VJ71 NAUCC
                                              EKK
                                                                                                                                    922
tr|G8ZM49|G8ZM49 TORDO
                                              EKK
                                                    <mark>G</mark>MSITETFEKPQRTNTLRMKHGTYDK<mark>LDDDGL</mark>IAPGVRVSGR
                                                                                                         DIIIGKTTPISPDE
                                                                                                                                    920
                                                    G<mark>MSITETFEKPQRTNTLRMKHGTYDK<mark>LDDDGL</mark>IAPGVRVSGEDIIIGKTTPISPDE</mark>
tr|A0A1Q3A090|A0A1Q3A090 ZYGRO
                                              EKK
                                                                                                                                    922
                                                    <mark>S</mark>ASITETFE KPQRTNT LRMKHGTYDK <mark>LDD DGL</mark>I APGVRV SGE
tr|A0A0N7IS35|A0A0N7IS35_9SACH
                                              EKK
                                                                                                          DVIIGKTTPISPDE
                                                                                                                                    920
tr|A0A212MG88|A0A212MG88 ZYGBA
                                              EKK
                                                    <mark>G</mark>MSITETFEKPQRTNTLRMKHGTYDK<mark>LDDDGL</mark>IAPGVRVSGR
                                                                                                         DIIIGKTTPISPDE
                                                                                                                                    921
                                                    <mark>g</mark>isitetfekportntlrmkhgtydk<mark>loddgl</mark>tapgvrvsgf
tr|A0A1S7HHE1|A0A1S7HHE1 9SACH
                                              EKK
                                                                                                         DIIIGKTTPISPDE
                                                                                                                                    921
                                                    <mark>g</mark>isitetfekpqrtntlrmkhgtydk<mark>i.dddgi</mark>lapgvrvsgebiiigkttpispde
tr|S6ESB4|S6ESB4 ZYGB2
                                              EKK
                                                                                                                                    921
                                                    GNTVMEEFERPTRSTTLRMKHGTYDK<mark>LEDDGL</mark>IAPGTRVSGDDIIIGKTAPVPPDH
tr|B6K5Q5|B6K5Q5 SCHJY
                                              EKK
                                                                                                                                    911
sp|Q02061|RPB2 SCHPO
                                                   <mark>.G</mark>MTVMEEFERPVRSTTLRMKHGTYDK<mark>LEDDGL</mark>IAPGTRVSGEDIIIGKTAPIPLDH
                                                                                                                                    911
                                              EKK
tr|S9R8U4|S9R8U4_SCHOY
                                                    <mark>G</mark>MTVMEEFERPARSTTLRMKHGTYDK<mark>LEDDGL</mark>IAPGTRVSGDDIIIGKTAPIPVDN
                                              FKK
                                                                                                                                    911
tr|S9W8C6|S9W8C6 SCHCR
                                                    HTVMEEFERPARSTTLRMKHGTYDK<mark>LEDDGL</mark>IAPGTRVSGDDIIIGKTAPIPLDN
                                                                                                                                    911x9
                                             ARDTKLGPERI----TRDIPHLSEAALRDLDEEGVVRIGAEVKPGDILVGRTSFKGESE
sp|Q8RQE9|RPOB THET8
                                                                                                                                   766
                                             ARDTKLGPEDI----TRDIPNVGEEALRN LDEAGIVYIGAEVHPGDILVGKITPKGESP
                                                                                                                                   907
ASR51304.1
OXR47929.1
                                             ARDTKLGAEEI----TRDISNLPEIQLNRL
                                                                                      <mark>DDSGI</mark>VHIGAEVRADDVLVGKVT PKGETQ
                                                                                                                                   900
WP 093971860.1
                                             ARDTKLGAEEI----TRDISNLPEIQLNR
                                                                                      <mark>DDSGI</mark>VHIGAEVRADDVLVGKVTPKGETQ
                                                                                                                                   900
                                             SRDTKLGPEEI----TADI PNVGEAALSK<mark>LDE SGI</mark>VYIGAEVTGGD ILVGKVT PKGETQ
sp|Q2NWR6|RPOB SODGM
                                                                                                                                   894
                                             SRDTKLGPEEI----TADI PNVGEAALSKLDE SGIVYIGAEVKGGDILVCKVT PKGETO
sp|B4EYU9|RPOB_PROMH
                                                                                                                                   894
                                             SRDTKLGPEEI----TADI PNVGEAALSK<mark>LDESGI</mark>VYIGAEVTGGDILVCKVT PKGETQ
sp|A7FNI3|RPOB YERP3
                                                                                                                                   894
                                             SRDTKLGPEEI----TADI PNVGEAALSK<mark>LDE SGI</mark>VYIGAEVTGGD ILV<mark>G</mark>KVT PKGETQ
sp|Q1C1U1|RPOB YERPA
                                                                                                                                   894
                                             SRDTKLGPEEI----TADI PNVGEAALSKLDESGIVYIGAEVTGGDILVGKVT PKGETQ
sp|B2K113|RPOB YERPB
                                                                                                                                   894
                                             SRDTKLGPEEI----TADI PNVGEAALSK<mark>LDE SGI</mark>VYIGAEVTGGDILVGKVT PKGETQ
sp|A8G8E7|RPOB SERP5
                                                                                                                                   894
                                             SRDTKLGPEEI----TADI PNVGEAALSKLDE SGIVYIGAEVTGGD ILVGKVT PKGETQ
sp | Q6DANO | RPOB_PECAS
                                                                                                                                   894
                                             SRDTKLGPEEI----TADI PNVGEAALSKLDESGIVYIGAEVTGGDILVGKVT PKGETQ
sp|C6DHR5|RPOB PECCP
                                                                                                                                   894
                                             SRDTKLGPEEI----TADI PNVGEAALSK<mark>LDE SGI</mark>VYIGAEVTGGD ILVGKVT PKGETQ
sp|Q7N9A4|RPOB PHOLL
                                                                                                                                   894
                                             SRDTKLGPEEI----TADI PNVGEAALSK<mark>LDESGI</mark>VYIGAEVKGGDILVGKVT PKGETQ
sp|C5BHE3|RPOB_EDWI9
                                                                                                                                   894
                                             SRDTKLGPEEI----TADI PNVGEAALSKLDE SGIVYIGAEVTGGDILVGKVT PKGETQ
sp|A7MQQ9|RPOB CROS8
                                                                                                                                   894
sp|B5XYF5|RPOB_KLEP3
                                             SRDTKLGPEEI----TADI PNVGEAALSK<mark>LDE SGI</mark>VYIGAEVTGGD ILVGKVT PKGETQ
                                                                                                                                   894
sp|P0A8V2|RPOB ECOLI
                                              SRDTKLGPEEI----TADI PNVGEAALSKLDESGIVYIGAEVTGGD ILVGKVT PKGETQ
                                                                                                                                   894
                                             SRDTKLGPEEI----TADI PNVGEAALSK<mark>LDE SGI</mark>VYIGAEVTGGDILVGKVT PKGETQ
sp|C5A0S7|RPOB ECOBW
                                                                                                                                   894
sp|Q31U10|RPOB_SHIBS
                                             SRDTKLGPEEI----TADI PNVGEAALSK<mark>LDE SGI</mark>VYIGAEVTGGD ILV<mark>C</mark>KVT PKGETQ
                                                                                                                                   894
sp|Q32AF9|RPOB SHIDS
                                             SRDTKLGPEEI----TADI PNVGEAALSKLDE SGIVYIGAEVTGGDILVGKVT PKGETQ
                                                                                                                                   894
                                             SRDTKLGPEEI----TADI PNVGEAALSK<mark>LDE SGI</mark>VYIGAEVTGGDILVGKVT PKGETQ
sp|A8AKT9|RPOB_CITK8
                                                                                                                                   894
                                             SRDTKLGPEEI----TADI PNVGEAALSK<mark>LDE SGI</mark>VYIGAEVTGGDILVGKVT PKGETQ
                                                                                                                                   894
sp|B5RFK1|RPOB_SALG2
                                             SRDTKLGPEEI----TADI PNVGEAALSKLDESGIVYIGAEVTGGDILVGKVT PKGETO
sp|B5BJQ3|RPOB_SALPK
                                                                                                                                   894
sp|B4T0Y9|RPOB SALNS
                                             SRDTKLGPEEI----TADI PNVGEAALSK<mark>LDE SGI</mark>VYIGAEVTGGD ILVGKVT PKGETQ
                                                                                                                                   894
                                             SRDTKLGPEEI----TADI PNVGEAALSKLDE SGIVYIGAEVTGGD ILVGKVT PKGETQ
sp|P06173|RPOB_SALTY
                                                                                     :: *:
                                                                          : .
                                                                                              *..*
```

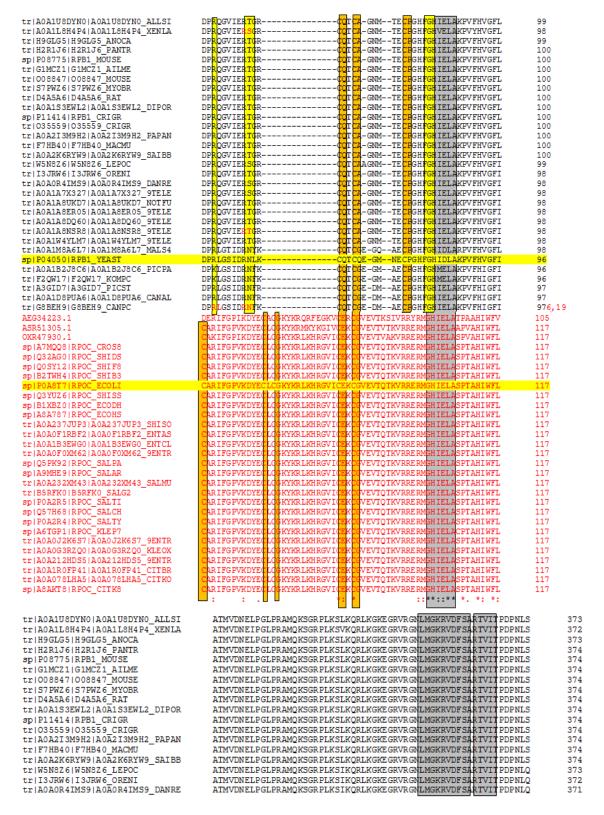
sp P30876 RPB2 HUMAN	IGDKFASRHGQKGTCGIQYRQEDMP#TCEGIT#DIIINFHAIPSRMTIGHLIECLQGKVS	990
tr G3V8Y5 G3V8Y5_RAT	IGDKFASRHGQKGTCGIQYRQEDMPHTCEGITHDIIINHHAIPSRMTIGHIIECLQGKVS	990
tr A0A250Y753 A0A250Y753_CASCN	IGDKFASRHGQKGTCGIQYRQEDMPFTCEGITFDIIINPHAIPSRMTIGHLIECLQGKVS	990
tr A0A1U7R4C7 A0A1U7R4C7_MESAU	IGDKFASRHGQKGTCGIQYRQEDMPHTCEGITHDIIINFHAIPSRMTIGHIIECLQGKVS	990
tr A0A286XIQ9 A0A286XIQ9_CAVP0 tr I3M351 I3M351 ICTTR	IGDKFASRHGQKGTCGIQYFQEDMPFTCEGITEDIIINFHAIPSRMTIGHLIECLQGKVS IGDKFASRHGQKGTCGIQYFQEDMPFTCEGITEDIIINFHAIPSRMTIGHLIECLQGKVS	990 990
tr G7P5R6 G7P5R6 MACFA	IGDKFASRHGQKGTCGIQYRQEDMFTCEGITEDIIINFHAIPSRMTIGHLIECLQGKVS	990
tr H2QPI8 H2QPI8 PANTR	IGDKFASRHGQKGTCGIQYRQEDMPFTCEGITFDIIINFHAIPSRMTIGHIIECLQGKVS	990
tr A0A1U7V0T5 A0A1U7V0T5 TARSY	IGDKFASRHGQKGTCGIQYRQEDMP#TCEGIT#DIIINPHAIPSRMTIGHIIECLQGKVS	990
tr A0A1S2ZSL2 A0A1S2ZSL2_ERIEU	IGDKFASRHGQKGTCGIQYRQEDMPFTCEGITEDIIINPHAIPSRMTIGHIIECLQGKVS	990
tr A0A0D9QYL1 A0A0D9QYL1_CHLSB	IGDKFASRHGQKGTCGIQYRQEDMPFTCEGITFDIIINPHAIPSRMTIGHIIECLQGKVS	983
tr AOA2K5ZNR7 AOA2K5ZNR7_MANLE	IGDKFASRHGQKGTCGIQYFQEDMPFTCEGITEDIIINFHAIPSRMTIGHLIECLQGKVS IGDKFASRHGQKGTCGIQYFQEDMPFTCEGITEDIIINFHAIPSRMTIGHLIECLQGKVS	983 990
tr A0A2I2ZIU3 A0A2I2ZIU3_GORG0 tr A0A1D5QGA5 A0A1D5QGA5_MACMU	IGDKFASRHGQKGTCGIQYRQEDMPHTCEGITHDIIINPHAIPSRMTIGHLIECLQGKVS	990
tr A0A2J8S2N1 A0A2J8S2N1 PONAB	IGDKFASRHGOKGTCGIOYROEDMPHTCEGITHDIIINPHAIPSRMTIGHIIECLOGKVS	990
tr A0A2K5K5J5 A0A2K5K5J5 COLAP	IGDKFASRHGQKGTCGIQYRQEDMPFTCEGITEDIIINPHAIPSRMTIGHIIECLQGKVS	990
_		
tr A0A2J8PEW7 A0A2J8PEW7 PANTR	IGDKFASRHGQKGTCGIQYRQEDMP#TCEGITPDIIINPHAIPSRMTIGHL!ECLQGKVS	983
tr A0A2K5CY83 A0A2K5CY83 AOTNA	IGDKFASRHGQKGTCGIQYRQEDMPFTCEGITPDIIINPHAIPSRMTIGHLIECLQGKVS	990
tr A0A096NEY4 A0A096NEY4_PAPAN	IGDKFASRHGQKGTCGIQYRQEDMP#TCEGITPDIIINPHAIPSRMTIGHLIECLQGKVS	990
tr C9J2Y9 C9J2Y9_HUMAN	IGDKFASRHGQKGTCGIQYRCEDMPFTCEGITPDIIINPHAIPSRMTIGHLIECLQGKVS	983
tr G8BY61 G8BY61_TETPH	IGDKFASRHGQKGTIGITYGREDMPETAEGIVPDLIINPHAIPSRMTVAHLIECLLSKVA	1035
tr A0A1X7QYA1 A0A1X7QYA1_9SACH	IGDKFASRHGQKGTIGITYRREDMPHTAEGIVPDLIINPHAIPSRMTVAHLIECLLSKVA	1032
tr J7RV95 J7RV95_KAZNA tr H2AVJ8 H2AVJ8 KAZAF	ISDKFASRHGQKGTIGITYSREDMP#TAEGIVPDLIINPHAIPSRMTVAHLIECLLSKVA ISDKFASRHGQKGTIGITYRREDMP#TADGIVPDLIINPHAIPSRMTVAHLIECLLSKVA	1031 1033
sp Q6FLD5 RPB2 CANGA	IGDKFASRHGQKGTIGITYRREDMPTAEGIVPDLIINPHAIPSRMTVAHLIECLLSKVA	1033
sp P08518 RPB2 YEAST	IGDKFASRHGQKGTIGITYRREDMPFTAEGIVPDLIINPHAIPSRMTVAHLIECLLSKVA	1035
tr A0A0L8VHA5 A0A0L8VHA5 9SACH	IGDKFASRHGQKGTIGITYRREDMP#TAEGIVPDLIINPHAIPSRMTVAHLIECLLSKVA	1035
tr A0A0L8RB33 A0A0L8RB33_SACEU	IGDKFASRHGQKGTIGITYRREDMPFTAEGIVPDLIINPHAIPSRMTVAHLIECLLSKVA	1035
tr G0VJ71 G0VJ71_NAUCC	IGDKFASRHGQKGTIGITYRREDMP#TAEGIVPDLIINPHAIPSRMTVAHLIECLLSKVA	1035
tr G8ZM49 G8ZM49_TORDC	IGDKFASRHGQKGTIGITYGREDMPETAEGIVPDLIINPHAIPSRMTVAHLIECLLSKVA	1033
tr A0A1Q3A090 A0A1Q3A090_ZYGRO	ISDKFASRHGQKGTIGITYGREDMP#TAEGIVPDLIINPHAIPSRMTVAHLIECLLSKVA ISDKFASRHGQKGTIGITYGREDMP#TSEGIVPDLIINPHAIPSRMTVAHLIECLLSKVA	1035 1033
tr A0A0N7IS35 A0A0N7IS35_9SACH tr A0A212MG88 A0A212MG88_ZYGBA	IGDKFASRHGQKGTIGITYSREDMPHTAEGIVPDLIINPHAIPSRMTVAHLIECLLSKVA	1033
tr A0A1S7HHE1 A0A1S7HHE1 9SACH	IGDKFASRHGOKGTIGITYSREDMP#TAEGIVPDLIINPHAIPSRMTVAHLIECLLSKVA	1034
tr S6ESB4 S6ESB4 ZYGB2	IGDKFASRHGQKGTIGITYSREDMPFTAEGIVPDLIINPHAIPSRMTVAHLIECLLSKVA	1034
tr B6K5Q5 B6K5Q5_SCHJY	IGDKFASRHGQKGTIGMTYRHEDMPFSAQGIVPDIIINPHAIPSRMTVAHLVECQLSKVS	1024
sp Q02061 RPB2_SCHPO	IGDKFASRHGQKGTIGMTYRHEDMPFSAQGIVPDIIINPHAIPSRMTVAHLVECQLSKVS	1024
tr S9R8U4 S9R8U4_SCHOY	IGDKFASRHGQKGTIGMTYRHEDMP#SAQGVVPDVIINPHAIPSRMTVAHL1ECQLSKVS	1024
tr S9W8C6 S9W8C6_SCHCR	IGDKFASRHGQKGTIGMTYRHEDMPTSAQGIVPDVIINPHAIPSRMTVAHLTECQLSKVS	1024
eningDorgiddor Thrts	VGDKLANRHGNKGVVAKILPVEDMPHLPDGTPVDVILNPLGVPSRMNLGQILETHLGLAG	894
sp Q8RQE9 RPOB_THET8 ASR51304.1	PGDKMAGRHGNKGIISRILPIEDMPHLEDGTHVDFVLNPLGVPSRMNVGQIFETHLGWAA	1147
OXR47929.1	PGDKMAGRHGNKGVVSRITPVEDMPHMADGTPADIVLNPLGVPSRMNVGQVLEVHLGWAA	1140
WP 093971860.1	PGDKMAGRHGNKGVVSRITPVEDMPHMADGTPADIVLNPLGVPSRMNVGQVLEVHLGWAA	1140
sp Q2NWR6 RPOB SODGM	pgdkmagrhgnkgviskinpiedmpydengvpydivlnplgvpsrmnigqilethlgmaa	1121
sp B4EYU9 RPOB_PROMH	PGDKMAGRHGNKGVISKINPIEDMPYDENGNPYDLVLNPL¢VPSRMNIGQILETHLGMAA	1121
sp A7FNI3 RPOB_YERP3	PGDKMAGRHGNKGVISKINPIEDMPYDENGTPVDIVLNPL¢VPSRMNIGQILETHLGMAA	1121
sp Q1C1U1 RPOB_YERPA	PGDKMAGRHGNKGVISKINPIEDMPYDENGTPYDIVLNPLGVPSRMNIGQILETHLGMAA	1121
sp B2K113 RPOB_YERPB	PGDKMAGRHGNKGVISKINPIEDMPYDENGTPYDIVLNPLGVPSRMNIGQILETHLGMAA	1121
sp A8G8E7 RPOB_SERP5 sp Q6DAN0 RPOB_PECAS	PGDKMAGRHGNKGVISKINPIEDMPYDENGTPVDIVLNPLGVPSRMNIGQILETHLGMAA PGDKMAGRHGNKGVISKINPIEDMPYDENGTPVDIVLNPLGVPSRMNIGQILETHLGMAA	1121 1121
sp C6DHR5 RPOB PECCP	PGDKMAGRHGNKGVISKINFIEDMPYDENGTPVDIVLNPLGVPSRMNIGQILETHLGMAA	1121
sp Q7N9A4 RPOB PHOLL	PSDKMAGRHGNKGVISKINPIEDMPYDENGTPYDIVLNPLGVPSRMNIGQILETHLGMAA	1121
sp C5BHE3 RPOB EDWI9	PGDKMAGRHGNKGVISKINPIEDMPYDENGTPVDIVLNPLGVPSRMNIGQILETHLGMAA	1121
sp A7MQQ9 RPOB_CROS8	pgdkmagrhgnkgviskinpiedmpydengtpydivlnpl¢vpsrmnig@ilethlgmaa	1121
sp B5XYF5 RPOB_KLEP3	PGDKMAGRHGNKGVISKINPIEDMPHDANGTPVDIVLNPL¢VPSRMNIGQILETHLGMAA	1121
sp P0A8V2 RP0B_ECOLI	PGDKMAGRHGNKGVISKINPIEDMPYDENGTPVDIVLNPLGVPSRMNIGQILETHLGMAA	1121
sp C5A0S7 RPOB_ECOBW	PGDKMAGRHGNKGVISKINPIEDMPYDENGTPYDIVLNPLGVPSRMNIGQILETHLGMAA	1121
sp Q31U10 RPOB_SHIBS	PGDKMAGRHGNKGVISKINPIEDMPYDENGTPYDIVLNPLGVPSRMNIGQIIETHLGMAA	1121
sp Q32AF9 RPOB_SHIDS sp A8AKT9 RPOB_CITK8	PGDKMAGRHGNKGVISKINPIEDMPYDENGTPVDIVLNPLGVPSRMNIGQILETHLGMAA PGDKMAGRHGNKGVISKINPIEDMPYDENGTPVDIVLNPLGVPSRMNIGQILETHLGMAA	1121 1121
sp B5RFK1 RPOB_C11K6	PGDKMAGRHGNKGVISKINPIEDMPYDENGTPVDIVLNPLGVPSRMNIGOILETHLGMAA	1121
sp B5BJQ3 RPOB_SALPK	PGDKMAGRHGNKGVISKINFIEDMFYDENGTFYDIVLNFLGVFSRMNIGQILETHLGMAA	1121
sp B4T0Y9 RPOB SALNS	PSDKMAGRHGNKGVISKINPIEDMPYDENGTPVDIVLNPLGVPSRMNIGQILETHLGMAA	1121
sp P06173 RPOB_SALTY	PGDKMAGRHGNKGVISKINPIEDMPYDENGTPVDIVLNPLGVPSRMNIGQILETHLGMAA	1121
-	***:*.***:** .	

sp P30876 RPB2 HUMAN	RKITSQIFIGPTYYORLKHUVDDKIHSRARGPIQILNROPMEGRSRDGGURFGEMERDCQ	1094
tr G3V8Y5 G3V8Y5_RAT	RKITSQIFIGPTYYQRLKHUVDDKIHSRARGPIQILNRQPMEGRSRDGGIRFGEMERDCQ	1094
tr A0A250Y753 A0A250Y753_CASCN	RKITSQIFIGPTYYQRLKHIVDDKIHSRARGPIQILNRQPMEGRSRDGGIRFGEMERDCQ	1094
tr A0A1U7R4C7 A0A1U7R4C7_MESAU tr A0A286XIQ9 A0A286XIQ9_CAVPO	RKITSQIFIGPTYYQRLKHUVDDKIHSRARGPIQILNRQPNEGRSRDGGIRFGEMERDCQ RKITSQIFIGPTYYQRLKHUVDDKIHSRARGPIQILNRQPNEGRSRDGGIRFGEMERDCQ	1094 1094
tr I3M351 I3M351 ICTTR	RKITSQIFIGFIIIQRLKHNVDDKIHSRARGFIQILNRQFMEGRSRDGGIRFGEMERDCQ RKITSQIFIGFTYYQRLKHNVDDKIHSRARGFIQILNRQFMEGRSRDGGIRFGEMERDCQ	1094
tr G7P5R6 G7P5R6 MACFA	RKITSQIFIGPTYYQRLKHNVDDKIHSRARGPIQILNRQPMEGRSRDGGIRFGEMERDCQ	1094
tr H2QPI8 H2QPI8_PANTR	RKITSQIFIGPTYYQRLKHUVDDKIHSRARGPIQILNRQPMEGRSRDGGIRFGEMERDCQ	1094
tr A0A1U7V0T5 A0A1U7V0T5_TARSY	RKITSQIFIGPTYYQRLKHMVDDKIHSRARGPIQILNRQPMEGRSRDGGIRFGEMERDCQ	1094
tr A0A1S2ZSL2 A0A1S2ZSL2_ERIEU	RKITSQIFIGPTYYQRLKHUVDDKIHSRARGPIQILNRQPMEGRSRDGGIRFGEMERDCQ RKITSQIFIGPTYYORLKHUVDDKIHSRARGPIQILNROPMEGRSRDGGIRFGEMERDCQ	1094
tr A0A0D9QYL1 A0A0D9QYL1_CHLSB tr A0A2K5ZNR7 A0A2K5ZNR7_MANLE	RKITSQIFIGFTITQRLKHNVDDKIHSRARGFIQILNRQPNEGRSRDGGIRFGEMERDCQ RKITSQIFIGPTYYQRLKHNVDDKIHSRARGPIQILNRQPNEGRSRDGGIRFGEMERDCQ	1087 1087
tr A0A2I2ZIU3 A0A2I2ZIU3 GORGO	RKITSQIFIGPTYYQRLKHUVDDKIHSRARGPIQILNRQPMEGRSRDGGURFGEMERDCQ	1094
tr A0A1D5QGA5 A0A1D5QGA5_MACMU	RKITSQIFIGPTYYQRLKH <mark>UVDDKIHSRA</mark> RGPIQILNRQPMEGRSRDGGIRFGEMERDCQ	1094
tr A0A2J8S2N1 A0A2J8S2N1_PONAB	RKITSQIFIGPTYYQRLKHMVDDKIHSRARGPIQILNRQPMEGRSRDGGIRFGEMERDCQ	1094
tr A0A2K5K5J5 A0A2K5K5J5_COLAP tr A0A2J8PEW7 A0A2J8PEW7_PANTR	RKITSQIFIGPTYYQRLKHUVDDKIHSRARGPIQILNRQPMEGRSRDGGIRFGEMERDCQ RKITSQIFIGPTYYQRLKHUVDDKIHSRARGPIQILNRQPMEGRSRDGGIRFGEMERDCQ	1094 1087
tr A0A2K5CY83 A0A2K5CY83 AOTNA	RKITSOIFIGFTYYORLKHWVDDKIHSRARGPIOILNROPNEGRSRDGGIRFGEMERDCO	1007
tr A0A096NEY4 A0A096NEY4 PAPAN	RKITSQIFIGPTYYQRLKHMVDDKIHSRARGPIQILNRQPMEGRSRDGGIRFGEMERDCQ	1094
tr C9J2Y9 C9J2Y9_HUMAN	RKITSQIFIGPTYYQRLKHMVDDKIHSRARGPIQILNRQPMEGRSRDGGIRFGEMERDCQ	1087
tr G8BY61 G8BY61_TETPH	KKLMAQIFFGPTYYQRLRHMVDDKIHARARGPMQVLTRQPVEGRSRDGGIRFGEMERDCM	1138
tr A0A1X7QYA1 A0A1X7QYA1_9SACH tr J7RV95 J7RV95 KAZNA	KKLMAQIFFGPTYYQRLRHIVDDKIHARARGPMQVLTRQPVEGRSRDGGIRFGEMERDCM KKLMAQIFFGPTYYORLRHIVDDKIHARARGPMOVLTROPVEGRSRDGGIRFGEMERDCM	1135 1134
tr H2AVJ8 H2AVJ8 KAZAF	KKLMAQIFFGFTIYQRLRHAVDDKIHARARGPMQVLTRQPVEGRSRDGGIRFGEMERDCM KKLMAQIFFGPTYYQRLRHAVDDKIHARARGPMQVLTRQPVEGRSRDGGIRFGEMERDCM	1134
sp Q6FLD5 RPB2 CANGA	KKLMAQIFFGPTYYQRLRHNVDDKIHARARGPMQVLTRQPVEGRSRDGGIRFGEMERDCM	1137
sp P08518 RPB2_YEAST	KKLMAQIFFGPTYYQRLRHMVDDKIHARARGPMQVLTRQPVEGRSRDGGLRFGEMERDCM	1138
tr A0A0L8VHA5 A0A0L8VHA5_9SACH	KKLMAQIFFGPTYYQRLRHIVDDKIHARARGPMQVLTRQPVEGRSRDGGIRFGEMERDCM	1138
tr A0A0L8RB33 A0A0L8RB33_SACEU tr G0VJ71 G0VJ71 NAUCC	KKLMAQIFFGPTYYQRLRHIVDDKIHARARGPMQVLTRQPVEGRSRDGGIRFGEMERDCM KKLMAQIFFGPTYYQRLRHIVDDKIHARARGPMQVLTRQPVEGRSRDGGIRFGEMERDCM	1138 1138
tr G8ZM49 G8ZM49 TORDC	KKLMAQIFFGFTITQKLKHIVDDKIHARARGFMQVLTRQFVEGRSRDGGIRFGEMERDCM KKLMAQIFFGFTYYORLRHIVDDKIHARARGFMOVLTRQFVEGRSRDGGIRFGEMERDCM	1136
tr A0A1Q3A090 A0A1Q3A090 ZYGRO	KKLMSQIFFGPTYYQRLRHNVDDKIHARARGPMQVLTRQPVEGRSRDGGIRFGEMERDCM	1138
tr A0A0N7IS35 A0A0N7IS35_9SACH	KKLMSQIFFGPTYYQRLRH <mark>M</mark> VDDKIHARARGPMQVLTRQPVEGRSRDGGIRFGEMERDCM	1136
tr A0A212MG88 A0A212MG88_ZYGBA	KKLMSQIFFGPTYYQRLRHMVDDKIHARARGPMQVLTRQPVEGRSRDGGIRFGEMERDCM	1137
triãoàis7HHE112oàis7HHF1 9SACH	KKLMSOT FFGPTYYORI RHWYDDKTHARA G GPMOVI. TROPY FGRSR DGGIR FGENFR DCM	1137
tr A0A1S7HHE1 A0A1S7HHE1_9SACH tr S6ESB4 S6ESB4_ZYGB2	KKLMSQIFFGPTYYQRLRHHVDDKIHARARGPMQVLTRQPVEGRSRDGGLRFGEMERDCM KKLMSQIFFGPTYYQRLRHHVDDKIHARARGPMQVLTRQPVEGRSRDGGLRFGEMERDCM	1137 1137
tr S6ESB4 S6ESB4_ZYGB2 tr B6K5Q5 B6K5Q5_SCHJY		
tr S6ESB4 S6ESB4 ZYGB2 tr B6K5Q5 B6K5Q5 SCHJY sp Q02061 RPB2 SCHPO	KKLMSQIFFGPTYYQRLRHMVDDKIHARARGPMQVLTRQFVEGRSRDGGIRFGEMERDCM RKLVAQVFLGPTYYQRLKHLVDDKIHARARGPVQILTRQFVEGRSRDGGIRFGEMERDCQ RKLVAQVFLGPTYYQRLKHLVDDKIHARARGPVQILTRQFVEGRSRDGGIRFGEMERDCQ	1137 1127 1127
tr S6ESB4 S6ESB4 ZYGB2 tr B6K5Q5 B6K5Q5_SCHJY sp Q02061 RPB2 SCHP0 tr S9R8U4 S9R8U4_SCHOY	KKLMSQIFFGPTYYQRLRHMVDDKIHARARGPMQVLTRQFVEGRSRDGGIRFGEMERDCM RKLVAQVFLGPTYYQRLKHLVDDKIHARARGPVQILTRQFVEGRSRDGGIRFGEMERDCQ RKLVAQVFLGPTYYQRLKHLVDDKIHARARGPVQILTRQFVEGRSRDGGIRFGEMERDCQ RKLVSQVFLGPTYYQRLKHLVDDKIHARARGPVQILTRQFVEGRSRDGGIRFGEMERDCQ	1137 1127 1127 1127
tr S6ESB4 S6ESB4 ZYGB2 tr B6K5Q5 B6K5Q5_SCHJY sp Q02061 RPB2 SCHP0 tr S9R8U4 S9R8U4 SCHOY tr S9W8C6 S9W8C6_SCHCR	KKLMSQIFFGPTYYQRLRH MVDDKIHARA RGPMQVLI RQFV EGRSRDGGI RFGEMERDCM RKLVAQVFLGPTYYQRLKH LVDDKIHARA RGFVQILIRQFV EGRSRDGGI RFGEMERDCQ RKLVAQVFLGPTYYQRLKH LVDDKIHARA RGFVQILIRQFV EGRSRDGGI RFGEMERDCQ RKLVSQVFLGPTYYQRLKH LVDDKIHARA RGFVQILIRQFV EGRSRDGGI RFGEMERDCQ RKLVSQVFLGPTYYQRLKH LVDDKIHARA RGFVQILIRQFV EGRSRDGGI RFGEMERDCQ	1137 1127 1127 1127 1127
tr S6ESB4 S6ESB4 ZYGB2 tr B6K5Q5 B6K5Q5_SCHJY sp Q02061 RPB2 SCHP0 tr S9R8U4 S9R8U4_SCHOY	KKLMSQIFFGPTYYQRLRHMVDDKIHARARGPMQVLTRQFVEGRSRDGGIRFGEMERDCM RKLVAQVFLGPTYYQRLKHLVDDKIHARARGPVQILTRQFVEGRSRDGGIRFGEMERDCQ RKLVAQVFLGPTYYQRLKHLVDDKIHARARGPVQILTRQFVEGRSRDGGIRFGEMERDCQ RKLVSQVFLGPTYYQRLKHLVDDKIHARARGPVQILTRQFVEGRSRDGGIRFGEMERDCQ	1137 1127 1127 1127
tr S6ESB4 S6ESB4 ZYGB2 tr B6K5Q5 B6K5Q5 SCHJY sp Q02061 RPB2 SCHPO tr S9R8U4 S9R8U4 SCHOY tr S9W8C6 S9W8C6 SCHCR sp Q8RQE9 RPOB THET8	KKLMSQIFFGPTYYQRLRH WDDKIHARA RGPMQVL RQFW EGRSRDGGI RFGEMERDCM RKLVAQVFLGPTYYQRLKH LVDDKIHARA RGFVQILTRQFV EGRSRDGGI RFGEMERDCQ RKLVAQVFLGPTYYQRLKH LVDDKIHARA RGFVQILTRQFV EGRSRDGGI RFGEMERDCQ RKLVSQVFLGPTYYQRLKH LVDDKIHARA RGFVQILTRQFV EGRSRDGGI RFGEMERDCQ RKLVSQVFLGPTYYQRLKH LVDDKIHARA RGFVQILTRQFV EGRSRDGGI RFGEMERDCQ EPIEGPIVVGQMFIMKLYHWEDKMHARSTGFYSLITQQPLEGKAQFGGQRFGEMEVWAL DKFDRKVIVGYIYMLKIHH LVDDKHARSTGFYSLVTQQPLGGKAQFGGQRFGEMEVWAL EQFERPVIVGYMHYLKIHH LVDDKMHARSTGFYSLVTQQPLGGKAQFGGQRFGEMEVWAL	1137 1127 1127 1127 1127 1040
tr S6ESB4 S6ESB4_ZYGB2 tr B6K5Q5 B6K5Q5_SCHJY sp Q02061 RPB2_SCHPO tr S9R8U4 S9R8U4_SCHOY tr S9R8C6 S9W8C6_SCHCR sp Q8RQE9 RPOB_THET8 ASR51304.1 OXR47929.1 WP_093971860.1	KKLMSQIFFGPTYYQRLRH VVDDKIHARA RGPMQVL RQFV EGRSRDGGI RFGEMERDCM RKLVAQVFLGPTYYQRLKH LVDDKIHARA RGFVQILTRQFV EGRSRDGGI RFGEMERDCQ RKLVAQVFLGPTYYQRLKH LVDDKIHARA RGFVQILTRQFV EGRSRDGGI RFGEMERDCQ RKLVSQVFLGPTYYQRLKH LVDDKIHARA RGFVQILTRQFV EGRSRDGGI RFGEMERDCQ RKLVSQVFLGPTYYQRLKH LVDDKIHARA RGFVQILTRQFV EGRSRDGGI RFGEMERDCQ EFIEGPIVVGQMFIMKLYH MVEDKHARA RGFVQILTRQFV EGRSRDGGI RFGEMEVWAL DKFDRKVTVGYIYMKLHH LVDDKIHARS IGFYSLVTQQPLGGKAQFGGQ RFGEMEVWAL EQFERPVTVGYMHYLKLHH LVDDKMHARSTGPYSLVTQQPLGGKAQFGGQ RFGEMEVWAL EQFERPVTVGYMHYLKLHH LVDDKMHARSTGPYSLVTQQPLGGKAQFGGQ RFGEMEVWAL	1137 1127 1127 1127 1127 1040 1310 1306
tr S6ESB4 S6ESB4_ZYGB2 tr B6K5Q5 B6K5Q5_SCHJY sp Q02061 RPB2_SCHPO tr S9R8U4 S9R8U4_SCHOY tr S9R8C6 S9W8C6_SCHCR sp Q8E6 SPW8C6 SCHCR sp Q8E8 RPOB_THET8 ASR51304.1 OXR47929.1 WP_093971860.1 sp Q2NWR6 RPOB_SODGM	KKLMSQIFFGPTYYQRLRH NVDDKIHARA RGPMQVL RQFV EGRSRDGGI RFGEMERDCM RKLVAQVFLGPTYYQRLKH LVDDKIHARA RGFVQILTRQFV EGRSRDGGI RFGEMERDCQ RKLVAQVFLGPTYYQRLKH LVDDKIHARA RGFVQILTRQFV EGRSRDGGI RFGEMERDCQ RKLVSQVFLGPTYYQRLKH LVDDKIHARA RGFVQILTRQFV EGRSRDGGI RFGEMERDCQ RKLVSQVFLGPTYYQRLKH LVDDKIHARA RGFVQILTRQFV EGRSRDGGI RFGEMERDCQ EFIEGPIVVGQMFIMKLYHNVEDKIHARA RGFVQILTRQFV EGRSRDGGI RFGEMEVWAL DKFDRKVTVGYIYMLKIHH LVDDKIHARA SIGPYSLVTQQPLGGKAQFGGQ RFGEMEVWAL EQFERPVTVGYMHYLKIHH LVDDKMHARSTGPYSLVTQQPLGGKAQFGGQ RFGEMEVWAL EQFERPVTVGYMHYLKIHH LVDDKMHARSTGPYSLVTQQPLGGKAQFGGQ RFGEMEVWAL EQFERQVTVGYMYMLKLHH LVDDKMHARSTGPYSLVTQQPLGGKAQFGGQ RFGEMEVWAL	1137 1127 1127 1127 1127 1040 1310 1306 1306
tr S6ESB4 S6ESB4_ZYGB2 tr B6K5Q5 B6K5Q5_SCHJY sp Q02061 RPB2_SCHPO tr S9R8U4 S9R8U4_SCHOY tr S9W8C6 S9W8C6_SCHCR sp Q8RQE9 RPOB_THET8 ASR51304.1 OXR47929.1 WP_093971860.1 sp Q2NWR6 RPOB_SODGM sp B4EYU9 RPOB_PROMH	KKLMSQIFFGPTYYQRLKH DVDDKIHARA RGPMQVL RQFV EGRSRDGGI RFGEMERDCM RKLVAQVFLGPTYYQRLKH LVDDKIHARA RGFVQILTRQFV EGRSRDGGI RFGEMERDCQ RKLVAQVFLGPTYYQRLKH LVDDKIHARA RGFVQILTRQFV EGRSRDGGI RFGEMERDCQ RKLVSQVFLGPTYYQRLKH LVDDKIHARA RGFVQILTRQFV EGRSRDGGI RFGEMERDCQ RKLVSQVFLGPTYYQRLKH LVDDKHHARA RGFVQILTRQFV EGRSRDGGI RFGEMERDCQ EPIEGPIVVGQMFIMKLYHMVEDKHHARA TGFYSLIT QQPLGGKAQFGGQ RFGEMEVWAL DKFDRKVTVGYIYMLKIHH LVDDKHHARA TGFYSLVTQQPLGGKAQFGGG RFGEMEVWAL EQFERPVTVGYMHYLKIHH LVDDKMHARA TGFYSLVTQQPLGGKAQFGGG RFGEMEVWAL EQFERPVTVGYMYMLKIHH LVDDKMHARA TGFYSLVTQQPLGGKAQFGGG RFGEMEVWAL EQFERQVTVGYMYMLKIHH LVDDKMHARA TGSYSLVTQQPLGGKAQFGGG RFGEMEVWAL EQFERQVTVGYMYMLKIHH LVDDKMHARA TGSYSLVTQQPLGGKAQFGGG RFGEMEVWAL EQFERPVTVGYMYMLKINH LVDDKMHARA TGSYSLVTQQPLGGKAQFGGG RFGEMEVWAL	1137 1127 1127 1127 1127 1040 1310 1306 1306 1278 1278
tr S6ESB4 S6ESB4_ZYGB2 tr B6K5Q5 B6K5Q5_SCHJY sp Q02061 RPB2_SCHPO tr S9R8U4 S9R8U4_SCHOY tr S9R8C6 S9W8C6_SCHCR sp Q8E6 SPW8C6 SCHCR sp Q8E8 RPOB_THET8 ASR51304.1 OXR47929.1 WP_093971860.1 sp Q2NWR6 RPOB_SODGM	KKLMSQIFFGPTYYQRLRH NVDDKIHARA RGPMQVL RQFV EGRSRDGGI RFGEMERDCM RKLVAQVFLGPTYYQRLKH LVDDKIHARA RGFVQILTRQFV EGRSRDGGI RFGEMERDCQ RKLVAQVFLGPTYYQRLKH LVDDKIHARA RGFVQILTRQFV EGRSRDGGI RFGEMERDCQ RKLVSQVFLGPTYYQRLKH LVDDKIHARA RGFVQILTRQFV EGRSRDGGI RFGEMERDCQ RKLVSQVFLGPTYYQRLKH LVDDKIHARA RGFVQILTRQFV EGRSRDGGI RFGEMERDCQ EFIEGPIVVGQMFIMKLYHNVEDKIHARA RGFVQILTRQFV EGRSRDGGI RFGEMEVWAL DKFDRKVTVGYIYMLKIHH LVDDKIHARA SIGPYSLVTQQPLGGKAQFGGQ RFGEMEVWAL EQFERPVTVGYMHYLKIHH LVDDKMHARSTGPYSLVTQQPLGGKAQFGGQ RFGEMEVWAL EQFERPVTVGYMHYLKIHH LVDDKMHARSTGPYSLVTQQPLGGKAQFGGQ RFGEMEVWAL EQFERQVTVGYMYMLKLHH LVDDKMHARSTGPYSLVTQQPLGGKAQFGGQ RFGEMEVWAL	1137 1127 1127 1127 1127 1040 1310 1306 1306
tr S6ESB4 S6ESB4_ZYGB2 tr B6KSQ5 B6KSQ5_SCHJY sp Q02061 RPB2_SCHPO tr S9R8U4 S9R8U4 SCHOY tr S9R8U6 S9W8C6_SCHCR sp Q8RQE9 RPOB_THET8 ASR51304.1 OXR47929.1 WP_093971860.1 sp Q2WWR6 RPOB_SODGM sp B4EYU9 RPOB_PROMH sp B4FYU9 RPOB_PROMH sp A7FN13 RPOB_YERP3 sp Q1C1U1 RPOB_YERPA sp B2K113 RPOB_YERPB	KKLMSQIFFGPTYYQRLRH VVDDKIHARA RGPMQVL RQFV EGRSRDGGI RFGEMERDCM RKLVAQVFLGPTYYQRLKH LVDDKIHARA RGFVQILTRQFV EGRSRDGGI RFGEMERDCQ RKLVAQVFLGPTYYQRLKH LVDDKIHARA RGFVQILTRQFV EGRSRDGGI RFGEMERDCQ RKLVSQVFLGPTYYQRLKH LVDDKIHARA RGFVQILTRQFV EGRSRDGGI RFGEMERDCQ RKLVSQVFLGPTYYQRLKH LVDDKIHARA RGFVQILTRQFV EGRSRDGGI RFGEMERDCQ RKLVSQVFLGPTYYQRLKH LVDDKIHARA RGFVQILTRQFV EGRSRDGGI RFGEMERDCQ EFIEGFIVVGQMFIMKLYHHVEDKMHARSTGFYSLVTQQPLGGKAQFGGQ RFGEMEVWAL DKFDRKVTVGYIYMLKLHH LVDDKMHARSTGFYSLVTQQPLGGKAQFGGQ RFGEMEVWAL EQFERPVTVGYMHYLKLHH LVDDKMHARSTGFYSLVTQQPLGGKAQFGGQ RFGEMEVWAL EQFERQVTVGYMYMLKLNH LVDDKMHARSTGSYSLVTQQPLGGKAQFGG RFGEMEVWAL	1137 1127 1127 1127 1127 1040 1310 1306 1306 1278 1278 1278 1278
tr S6ESB4 S6ESB4_ZYGB2 tr B6KSQ5 B6KSQ3_SCHJY sp Q02061 RPB2_SCHPO tr S9R8U4 S9R8U4 SCHOY tr S9R8U4 S9R8U4 SCHOY tr S9R8U6 S9W8C6 SCHCR sp Q8RQE9 RPOB_THET8 ASR51304.1 OXR47929.1 WP_093971860.1 sp Q2NWR6 RPOB_SODGM sp B4EYU9 RPOB_PROMH sp A7FN13 RPOB_YERP3 sp Q1C1U1 RPOB_YERPA sp B2K113 RPOB_YERPB sp A8G8E7 RPOB_SERP5	KKLMSQIFFGPTYYQRLKH DVDDKIHARA RGPMQVL RQFV EGRSRDGGI RFGEMERDCM RKLVAQVFLGPTYYQRLKH LVDDKIHARA RGFVQILTRQFV EGRSRDGGI RFGEMERDCQ RKLVSQVFLGPTYYQRLKH LVDDKIHARA RGFVQILTRQFV EGRSRDGGI RFGEMERDCQ RKLVSQVFLGPTYYQRLKH LVDDKIHARA RGFVQILTRQFV EGRSRDGGI RFGEMERDCQ RKLVSQVFLGPTYYQRLKH LVDDKIHARA RGFVQILTRQFV EGRSRDGGI RFGEMERDCQ EFIEGPIVVGQMFIMKLYHMVEDKMHARS TGFYSLITQQPLGGKAQFGGQ RFGEMEVWAL DKFDRKVTVGYIYYMKLHHLVDDKHHARS TGFYSLITQQPLGGKAQFGGQ RFGEMEVWAL EQFERPVTVGYMHYLKLHHLVDDKMHARS TGFYSLVTQQPLGGKAQFGGQ RFGEMEVWAL EQFERPVTVGYMYMLKLHHLVDDKMHARS TGFYSLVTQQPLGGKAQFGGQ RFGEMEVWAL EQFERVTVGYMYMLKLHHLVDDKMHARS TGSYSLVTQQPLGGKAQFGGQ RFGEMEVWAL EQFERVTVGYMYMLKLHHLVDDKMHARS TGSYSLVTQQPLGGKAQFGGQ RFGEMEVWAL EQFERVVTVGYMYMLKLHHLVDDKMHARS TGSYSLVTQQPLGGKAQFGGQ RFGEMEVWAL EQFERQVTVGYMYMLKLHLVDDKMHARS TGSYSLVTQQPLGGKAQFGG RFGEMEVWAL EQFERQVTVGYMYMLKLHHLVDDKMHARS TGSYSLVTQQPLGGKAQFGG RFGEMEVWAL	1137 1127 1127 1127 1127 1040 1310 1306 1306 1278 1278 1278 1278 1278
tr S6ESB4 S6ESB4 ZYGB2 tr B6K5Q5 B6K5Q5 SCHJY sp Q02061 RPB2 SCHPO tr S9R8U4 S9R8U4 SCHOY tr S9R8C4 S9R8U4 SCHOY tr S9R8C6 S9R8C6 SCHCR sp Q8RQE9 RPOB THET8 ASR51304.1 OXR47929.1 WP_093971860.1 sp Q2NRR6 RPOB SODGM sp B4EYU9 RPOB PROMH sp A7FN13 RPOB YERP3 sp Q1C1U1 RPOB YERPA sp B2K113 RPOB YERPA sp B4S8E7 RPOB SERP5 sp Q6DAN0 RPOB PECAS	KKLMSQIFFGPTYYQRLKH DVDDKIHARA RGPMQVL RQFV EGRSRDGGI RFGEMERDCM RKLVAQVFLGPTYYQRLKH DVDDKIHARA RGPVQILTRQPV EGRSRDGGI RFGEMERDCQ RKLVSQVFLGPTYYQRLKH DVDDKIHARA RGPVQILTRQPV EGRSRDGGI RFGEMERDCQ RKLVSQVFLGPTYYQRLKH DVDDKIHARA RGPVQILTRQPV EGRSRDGGI RFGEMERDCQ RKLVSQVFLGPTYYQRLKH DVDDKIHARA RGPVQILTRQPV EGRSRDGGI RFGEMERDCQ EPIEGPIVVGQMFIMKLYHMVEDKMHARSTGPYSLITQQPLGGKAQFGGQRFGEMEVWAL DKFDRKVTVGYIYYMLKLHH LVDDKTHARSIGPYSLITQQPLGGKAQFGGQRFGEMEVWAL EQFERPVTVGYMHYLKLHH LVDDKMHARSTGPYSLVTQQPLGGKAQFGGQRFGEMEVWAL EQFERPVTVGYMYMLKLHH LVDDKMHARSTGSYSLVTQQPLGGKAQFGGQRFGEMEVWAL EQFERQVTVGYMYMLKLHH LVDDKMHARSTGSYSLVTQQPLGGKAQFGGQRFGEMEVWAL EQFERQVTVGYMYMLKLHH LVDDKMHARSTGSYSLVTQQPLGGKAQFGGQRFGEMEVWAL EQFERQVTVGYMYMLKLHH LVDDKMHARSTGSYSLVTQQPLGGKAQFGGQRFGEMEVWAL EQFERQVTVGYMYMLKLHH LVDDKMHARSTGSYSLVTQQPLGGKAQFGGQRFGEMEVWAL EQFERQVTVGYMYMLKLHL LVDDKMHARSTGSYSLVTQQPLGGKAQFGGQRFGEMEVWAL EQFERQVTVGYMYMLKLHL LVDDKMHARSTGSYSLVTQQPLGGKAQFGGGRFGEMEVWAL EQFERQVTVGYMYMLKLHL LVDDKMHARSTGSYSLVTQQPLGGKAQFGGGRFGEMEVWAL EQFERQVTVGYMYMLKLHHLVDDKMHARSTGSYSLVTQQPLGGKAQFGGGRFGEMEVWAL EQFERQVTVGYMYMLKLHHLVDDKMHARSTGSYSLVTQQPLGGKAQFGGGRFGEMEVWAL EQFERQVTVGYMYMLKLHHLVDDKMHARSTGSYSLVTQQPLGGKAQFGGGRFGEMEVWAL EQFERQVTVGYMYMLKLHHLVDDKMHARSTGSYSLVTQQPLGGKAQFGGGRFGEMEVWAL EQFERQVTVGYMYMLKLHHLVDDKMHARSTGSYSLVTQQPLGGKAQFGGGRFGEMEVWAL	1137 1127 1127 1127 1127 1040 1310 1306 1306 1278 1278 1278 1278 1278
tr S6ESB4 S6ESB4 ZYGB2 tr B6K5Q5 B6K5Q5 SCHJY sp Q02061 RPB2 SCHPO tr S9R8U4 S9R8U4 SCHOY tr S9R8C4 S9R8U4 SCHOY tr S9R8C6 S9R8C6 SCHCR sp Q8R8C99 RPOB THET8 ASR51304.1 OXR47929.1 WP_093971860.1 sp Q2NWR6 RPOB SODGM sp B4EYU9 RPOB PROMH sp A7FNI3 RPOB YERP3 sp Q1C1U1 RPOB YERPA sp B2K113 RPOB YERPB sp A8G8E7 RPOB SERP5 sp Q6DANO RPOB PECAS sp C6DHR5 RPOB PECCP	KKLMSQIFFGPTYYQRLKH DVDDKIHARA RGPMQVL RQFV EGRSRDGGI RFGEMERDCM RKLVAQVFLGPTYYQRLKH DVDDKIHARA RGPVQILTRQPV EGRSRDGGI RFGEMERDCQ RKLVAQVFLGPTYYQRLKH DVDDKIHARA RGPVQILTRQPV EGRSRDGGI RFGEMERDCQ RKLVSQVFLGPTYYQRLKH DVDDKIHARA RGPVQILTRQPV EGRSRDGGI RFGEMERDCQ EPIEGPIVGQMFIMKLYHIVEDKHMARSTGPYSLITQQPLGGKAQFGGQ RFGEMEVWAL DKFDRKVTVGYIYMLKLHH LVDDKHHARSTGPYSLVTQQPLGGKAQFGGQ RFGEMEVWAL EQFERPVTVGYMHYLKLHH LVDDKMHARSTGPYSLVTQQPLGGKAQFGGQ RFGEMEVWAL EQFERPVTVGYMYMLKLHH LVDDKMHARSTGFYSLVTQQPLGGKAQFGGQ RFGEMEVWAL EQFERQVTVGYMYMLKLHH LVDDKMHARSTGSYSLVTQQPLGGKAQFGGQ RFGEMEVWAL EQFERQVTVGYMYMLKLHH LVDDKMHARSTGSYSLVTQQPLGGKAQFGGGRFGEMEVWAL EXFERQVTVGYMYMLKLHH LVDDKMHARSTGSYSLVTQQPLGGKAQFGGGRFGEMEVWAL	1137 1127 1127 1127 1127 1040 1310 1306 1306 1278 1278 1278 1278 1278 1278
tr S6ESB4 S6ESB4 ZYGB2 tr B6K5Q5 B6K5Q5 SCHJY sp Q02061 RPB2 SCHPO tr S9R8U4 S9R8U4 SCHOY tr S9R8C4 S9R8U4 SCHOY tr S9R8C6 S9R8C6 SCHCR sp Q8RQE9 RPOB THET8 ASR51304.1 OXR47929.1 WP_093971860.1 sp Q2NRR6 RPOB SODGM sp B4EYU9 RPOB PROMH sp A7FN13 RPOB YERP3 sp Q1C1U1 RPOB YERPA sp B2K113 RPOB YERPA sp B4S8E7 RPOB SERP5 sp Q6DAN0 RPOB PECAS	KKLMSQIFFGPTYYQRLKH DVDDKIHARA RGPMQVL RQFV EGRSRDGGI RFGEMERDCM RKLVAQVFLGPTYYQRLKH DVDDKIHARA RGPVQILTRQPV EGRSRDGGI RFGEMERDCQ RKLVSQVFLGPTYYQRLKH DVDDKIHARA RGPVQILTRQPV EGRSRDGGI RFGEMERDCQ RKLVSQVFLGPTYYQRLKH DVDDKIHARA RGPVQILTRQPV EGRSRDGGI RFGEMERDCQ RKLVSQVFLGPTYYQRLKH DVDDKIHARA RGPVQILTRQPV EGRSRDGGI RFGEMERDCQ EPIEGPIVVGQMFIMKLYHMVEDKMHARSTGPYSLITQQPLGGKAQFGGQRFGEMEVWAL DKFDRKVTVGYIYYMLKLHH LVDDKTHARSIGPYSLITQQPLGGKAQFGGQRFGEMEVWAL EQFERPVTVGYMHYLKLHH LVDDKMHARSTGPYSLVTQQPLGGKAQFGGQRFGEMEVWAL EQFERPVTVGYMYMLKLHH LVDDKMHARSTGSYSLVTQQPLGGKAQFGGQRFGEMEVWAL EQFERQVTVGYMYMLKLHH LVDDKMHARSTGSYSLVTQQPLGGKAQFGGQRFGEMEVWAL EQFERQVTVGYMYMLKLHH LVDDKMHARSTGSYSLVTQQPLGGKAQFGGQRFGEMEVWAL EQFERQVTVGYMYMLKLHH LVDDKMHARSTGSYSLVTQQPLGGKAQFGGQRFGEMEVWAL EQFERQVTVGYMYMLKLHH LVDDKMHARSTGSYSLVTQQPLGGKAQFGGQRFGEMEVWAL EQFERQVTVGYMYMLKLHL LVDDKMHARSTGSYSLVTQQPLGGKAQFGGQRFGEMEVWAL EQFERQVTVGYMYMLKLHL LVDDKMHARSTGSYSLVTQQPLGGKAQFGGGRFGEMEVWAL EQFERQVTVGYMYMLKLHL LVDDKMHARSTGSYSLVTQQPLGGKAQFGGGRFGEMEVWAL EQFERQVTVGYMYMLKLHHLVDDKMHARSTGSYSLVTQQPLGGKAQFGGGRFGEMEVWAL EQFERQVTVGYMYMLKLHHLVDDKMHARSTGSYSLVTQQPLGGKAQFGGGRFGEMEVWAL EQFERQVTVGYMYMLKLHHLVDDKMHARSTGSYSLVTQQPLGGKAQFGGGRFGEMEVWAL EQFERQVTVGYMYMLKLHHLVDDKMHARSTGSYSLVTQQPLGGKAQFGGGRFGEMEVWAL EQFERQVTVGYMYMLKLHHLVDDKMHARSTGSYSLVTQQPLGGKAQFGGGRFGEMEVWAL	1137 1127 1127 1127 1127 1040 1310 1306 1306 1278 1278 1278 1278 1278
tr S6ESB4 S6ESB4 ZYGB2 tr B6KSQ5 B6KSQ5 SCHJY sp Q02061 RPB2 SCHPO tr S9R8V4 S9R8U4 SCHOY tr S9R8V6 S9R8C6 SCHCR sp Q8RQE9 RPOB THET8 ASR51304.1 OXR47929.1 WP 093971860.1 sp Q2NWR6 RPOB SODGM sp B4EYU9 RPOB PROMH sp A7FN13 RPOB YERP3 sp Q1C1U1 RPOB YERPA sp B2K113 RPOB YERPA sp B2K113 RPOB YERPA sp A8G8E7 RPOB SERP5 sp Q6DANO RPOB PECCP sp Q7N9A4 RPOB PHOLL sp C5BHE3 RPOB EDW19 sp A7MQ09 RPOB CROS8	RKLWAQVFLGPTYYQRLKH LVDDKIHARA RGPWQVL RQFV GRSRDGGI RFGEMERDCM RKLVAQVFLGPTYYQRLKH LVDDKIHARA RGFVQILTRQFV GRSRDGGI RFGEMERDCQ RKLVAQVFLGPTYYQRLKH LVDDKIHARA RGFVQILTRQFV GRSRDGGI RFGEMERDCQ RKLVSQVFLGPTYYQRLKH LVDDKIHARA RGFVQILTRQFV GRSRDGGI RFGEMERDCQ RKLVSQVFLGPTYYQRLKH LVDDKIHARA RGFVQILTRQFV GRSRDGGI RFGEMERDCQ RKLVSQVFLGPTYYQRLKH LVDDKIHARA RGFVQILTRQFV GRSRDGGI RFGEMEVDCQ EFIEGPIVVGQMFIMKLYHWEDKHHARSTGFYSLVTQQPLGKAQFGGQ RFGEMEVWAL DKFDRKVTVGYIYMLKLHH LVDDKMHARSTGFYSLVTQQPLGGKAQFGGQ RFGEMEVWAL EQFERPVTVGYMYMLKLHH LVDDKMHARSTGFYSLVTQQPLGGKAQFGGQ RFGEMEVWAL EQFERPVTVGYMYMLKLHH LVDDKMHARSTGSYSLVTQQPLGGKAQFGGQ RFGEMEVWAL EQFERPVTVGYMYMLKLHH LVDDKMHARSTGSYSLVTQQPLGGKAQFGGQ RFGEMEVWAL EQFERQVTVGYMYMLKLHH VDDKMHARSTGSYSLVTQQPLGGKAQFGGQ RFGEMEVWAL EQFERQVTVGYMYMLKLHHLVDDKMHARSTGSYS	1137 1127 1127 1127 1127 1040 1310 1306 1306 1278 1278 1278 1278 1278 1278 1278 1278
tr S6ESB4 S6ESB4 ZYGB2 tr B6KSQ5 B6KSQ3 SCHJY sp Q02061 RPB2 SCHPO tr S9R8U4 S9R8U4 SCHOY tr S9R8U6 S9R8U6 SCHCR sp Q8RQE9 RPOB THET8 ASR51304.1 OXR47929.1 WP 093971860.1 sp Q2RWR6 RPOB SODGM sp B4EYU9 RPOB PROMH sp A7FN13 RPOB YERP3 sp Q1C1U1 RPOB YERPA sp B2K113 RPOB YERPB sp A8G8E7 RPOB SERP5 sp Q6DANO RPOB PECCP sp Q7N9A4 RPOB PHOLL sp C5BHE3 RPOB PHOLL sp C5BHE3 RPOB EDW19 sp A7MQ09 RPOB CROS8 sp B5XYF5 RPOB CROS8 sp B5XYF5 RPOB KLEP3	RKLWAQVFLGPTYYQRLKH LVDDKIHARA RGPWQVL RQFV EGRSRDGGI RFGEMERDCM RKLVAQVFLGPTYYQRLKH LVDDKIHARA RGFVQILTRQFV EGRSRDGGI RFGEMERDCQ RKLVAQVFLGPTYYQRLKH LVDDKIHARA RGFVQILTRQFV EGRSRDGGI RFGEMERDCQ RKLVSQVFLGPTYYQRLKH LVDDKIHARA RGFVQILTRQFV EGRSRDGGI RFGEMERDCQ RKLVSQVFLGPTYYQRLKH LVDDKIHARA RGFVQILTRQFV EGRSRDGGI RFGEMERDCQ RKLVSQVFLGPTYYQRLKH LVDDKHHARA RGFVQILTRQFV EGRSRDGGI RFGEMERDCQ RFGERFVIVGYMFINKLYHVE DKWHARS TGFYSLVTQQPLGGKAQFGGQ RFGEMEVWAL DKFDRKVIVGYIYMLKLHH LVDDKHHARS TGFYSLVTQQPLGGKAQFGGQ RFGEMEVWAL EQFERPVIVGYMYMLKLHH LVDDKMHARS TGFYSLVTQQPLGGKAQFGGQ RFGEMEVWAL EQFERQVIVGYMYMLKLHH LVDDKMHARS TGFYSLVTQQPLGGKAQFGG RFGEMEVWAL EQFERQVIVGYMYMLKLHH LVDDKMHARS TGSYSLVTQQPLGGKAQFGG RFGEMEVWAL EQFERPVIVGYMYMLKLHH LVDDKMHARS TGSYSLVTQQPLGGKAQFGG RFGEMEVWAL EQFERPVIVGYMYMLK	1137 1127 1127 1127 1127 1040 1310 1306 1306 1278 1278 1278 1278 1278 1278 1278 1278
tr S6ESB4 S6ESB4 ZYGB2 tr B6KSQ5 B6KSQ5 SCHJY sp Q02061 RPB2 SCHPO tr S9R8U4 S9R8U4 SCHOY tr S9R8U6 S9R8U4 SCHOY tr S9R8U6 S9R8U6 SCHCR sp Q8RQE9 RPOB THET8 ASR51304.1 OXR47929.1 WP 093971860.1 sp Q2WWR6 RPOB SODGM sp B4EYU9 RPOB PROMH sp A7FN13 RPOB YERPA sp B2K113 RPOB YERPA sp B2K113 RPOB YERPA sp B2K113 RPOB PECAS sp C6DHR5 RPOB PECCP sp Q7N9A4 RPOB PECCP sp Q7N9A4 RPOB PHOLL sp C5BHE3 RPOB EDWI9 sp A7MQQ9 RPOB CROS8 sp B5XYF5 RPOB KLEP3 sp P0A8V2 RPOB ECOLI	RKLWAQVFLGPTYYQRLKH LVDDKIHARA RGPWQVLT RQPV GRSRDGGI RFGEMERDCM RKLVAQVFLGPTYYQRLKH LVDDKIHARA RGPVQILT RQPV GRSRDGGI RFGEMERDCQ RKLVAQVFLGPTYYQRLKH LVDDKIHARA RGPVQILT RQPV GRSRDGGI RFGEMERDCQ RKLVSQVFLGPTYYQRLKH LVDDKIHARA RGPVQILT RQPV GRSRDGGI RFGEMERDCQ RKLVSQVFLGPTYYQRLKH LVDDKIHARA RGPVQILT RQPV GRSRDGGI RFGEMERDCQ RKLVSQVFLGPTYYQRLKH LVDDKHHARA RGPVQILT RQPV GRSRDGGI RFGEMERDCQ RFGEPVIVGQMFIMKLYHMVEDKHARA RGPVQILT RQPV GRSRDGGI RFGEMEVWAL DKFDRKVIVGYI YMLKLHH LVDDKHHARS TGPYSLVT QQPL GGKAQFGGQ RFGEMEVWAL EQFERPVIVGYMYMLKLHH LVDDKMHARS TGPYSLVT QQPL GGKAQFGGQ RFGEMEVWAL EQFERQVIVGYMYMLKLHH LVDDKMHARS TGSYSLVT QQPL GGKAQFGG RFGEMEVWAL EQFERPVIVGYMYMLKLHH VDDKMHARS TGSY	1137 1127 1127 1127 1127 1127 1040 1310 1306 1306 1278 1278 1278 1278 1278 1278 1278 1278
tr S6ESB4 S6ESB4 ZYGB2 tr B6K5Q5 B6KSQ5 SCHJY sp Q02061 RPB2 SCHPO tr S9R8U4 S9R8U4 SCHOY tr S9R8U6 S9W8U6 SCHCR sp Q8RQE9 RPOB THET8 ASR51304.1 OXR47929.1 WP 093971860.1 sp Q2NWR6 RPOB SODGM sp B4EYU9 RPOB PROMH sp A7FN13 RPOB YERP3 sp Q1C1U1 RPOB YERPA sp B2K113 RPOB YERPB sp A8G8E7 RPOB SERPS sp Q6DANO RPOB PECCP sp Q7M9A4 RPOB PHOLL sp C5BH3 RPOB PHOLL sp C5BH3 RPOB CROS8 sp B5XYF5 RPOB KLEP3 sp P0A8V2 RPOB ECOLI sp C5AOS7 RPOB ECOLI sp C5AOS7 RPOB ECOLI	RKLWAQVFLGPTYYQRLKH LVDDKIHARA RGPWQVL RQPV GRSRDGGI RFGEMERDCM RKLVAQVFLGPTYYQRLKH LVDDKIHARA RGPVQILTRQPV GRSRDGGI RFGEMERDCQ RKLVAQVFLGPTYYQRLKH LVDDKIHARA RGPVQILTRQPV GRSRDGGI RFGEMERDCQ RKLVSQVFLGPTYYQRLKH LVDDKIHARA RGPVQILTRQPV GRSRDGGI RFGEMERDCQ RKLVSQVFLGPTYYQRLKH LVDDKIHARA RGPVQILTRQPV GRSRDGGI RFGEMERDCQ RKLVSQVFLGPTYYQRLKH LVDDKIHARA RGPVQILTRQPV GRSRDGGI RFGEMERDCQ EPIEGPIVVGQMFIMKLYH VEDKMHARSTGPYSLITQQPLGGKAQFGGQ RFGEMEVWAL DKFDRKVTVGYIYYMLKLHH LVDDKHHARSTGPYSLITQQPLGGKAQFGGQ RFGEMEVWAL EQFERPVTVGYMYMLKLHH LVDDKMHARSTGPYSLVTQQPLGGKAQFGGQ RFGEMEVWAL EQFERQVTVGYMYMLKLHH LVDDKMHARSTGSYSLVTQQPLGGKAQFGGQ RFGEMEVWAL EQFERQVTVGYMYMLKLHH LVDDKMHARSTGSYSLVTQQPLGGKAQFGG RFGEMEVWAL EQFERQVTVGYMYMLKLHH LVDDKMHARSTGSYSLVTQQPLGGKAQFGG RFGEMEVWAL EQFERQVTVGYMYMLKLHH LVDDKMHARSTGSYSLVTQQPLGGKAQFGGQ RFGEMEVWAL EQFERPVTVGYMYMLKLHH VDDKMHARSTGSYSLVTQQPLGGKAQFGGG RFGEMEVWAL EQFERPVTVGYMYMLKLHHLVDDKMHARSTGSYSLVTQQPLGGKAQFGGG RFGEMEVWAL EQFERPVTVGYMYMLKLHHLVDDKMHARSTGSYSLVTQQPLGGKAQFGGG RFGEMEVWAL EQFERPVTVGYMYMLKLHHLVDDKMHARSTGSYSLVTQQPLGGKAQFGGG RFGEMEVWAL EQFERPVTVGYMYMLKLHHLVDDKMHARSTGSYSLVTQQPLGGKAQFGGG RFGEMEVWAL EQFERPVTVGYMYMLKLHHLVDDKMHARSTGSYSLVTQQPLGGKAQFGGGRFGEMEVWAL EQFERPVTVGYMYMLKLHHLVDDKMHARSTGSY	1137 1127 1127 1127 1127 1127 1040 1310 1306 1306 1278 1278 1278 1278 1278 1278 1278 1278
tr S6ESB4 S6ESB4 ZYGB2 tr B6KSQ5 B6KSQ5 SCHJY sp Q02061 RPB2 SCHPO tr S9R8U4 S9R8U4 SCHOY tr S9R8U6 S9R8U4 SCHOY tr S9R8U6 S9R8U6 SCHCR sp Q8RQE9 RPOB THET8 ASR51304.1 OXR47929.1 WP 093971860.1 sp Q2WWR6 RPOB SODGM sp B4EYU9 RPOB PROMH sp A7FN13 RPOB YERPA sp B2K113 RPOB YERPA sp B2K113 RPOB YERPA sp B2K113 RPOB PECAS sp C6DHR5 RPOB PECCP sp Q7N9A4 RPOB PECCP sp Q7N9A4 RPOB PHOLL sp C5BHE3 RPOB EDWI9 sp A7MQQ9 RPOB CROS8 sp B5XYF5 RPOB KLEP3 sp P0A8V2 RPOB ECOLI	RKLWAQVFLGPTYYQRLKH LVDDKIHARA RGPWQVLT RQPV GRSRDGGI RFGEMERDCM RKLVAQVFLGPTYYQRLKH LVDDKIHARA RGPVQILT RQPV GRSRDGGI RFGEMERDCQ RKLVAQVFLGPTYYQRLKH LVDDKIHARA RGPVQILT RQPV GRSRDGGI RFGEMERDCQ RKLVSQVFLGPTYYQRLKH LVDDKIHARA RGPVQILT RQPV GRSRDGGI RFGEMERDCQ RKLVSQVFLGPTYYQRLKH LVDDKIHARA RGPVQILT RQPV GRSRDGGI RFGEMERDCQ RKLVSQVFLGPTYYQRLKH LVDDKHHARA RGPVQILT RQPV GRSRDGGI RFGEMERDCQ RFGEPVIVGQMFIMKLYHMVEDKHARA RGPVQILT RQPV GRSRDGGI RFGEMEVWAL DKFDRKVIVGYI YMLKLHH LVDDKHHARS TGPYSLVT QQPL GGKAQFGGQ RFGEMEVWAL EQFERPVIVGYMYMLKLHH LVDDKMHARS TGPYSLVT QQPL GGKAQFGGQ RFGEMEVWAL EQFERQVIVGYMYMLKLHH LVDDKMHARS TGSYSLVT QQPL GGKAQFGG RFGEMEVWAL EQFERPVIVGYMYMLKLHH VDDKMHARS TGSY	1137 1127 1127 1127 1127 1127 1040 1310 1306 1306 1278 1278 1278 1278 1278 1278 1278 1278
tr S6ESB4 S6ESB4 ZYGB2 tr B6KSQ5 B6KSQ5 SCHJY sp Q02061 RPB2 SCHPO tr S9R8V6 S9R8V6 SCHCR sp Q8RQE9 RPOB THET8 ASR51304.1 OXR47929.1 WP O93971860.1 sp Q2NWR6 RPOB SODGM sp B4EYU9 RPOB PROMH sp A7FN13 RPOB YERP3 sp Q1C1U1 RPOB YERPA sp B2K113 RPOB YERPB sp A8G8E7 RPOB SERP5 sp Q6DANO RPOB PECCP sp Q7N9A4 RPOB PECCP sp Q7N9A4 RPOB PMOLL sp C5BHE3 RPOB EDW19 sp A7MQQ9 RPOB CROS8 sp B5XYF5 RPOB ECOLI sp C5AOS7 RPOB ECOLI sp C31U10 RPOB SHIBS sp Q32AF9 RPOB SHIBS sp Q32AF9 RPOB SHIDS sp A8AKT9 RPOB CITK8	RKLWAQVFLGPTYYQRLKH LVDDKIHARA RGPWQVL RQFV GRSRDGGI RFGEMERDCW RKLVAQVFLGPTYYQRLKH LVDDKIHARA RGFVQILT RQFV GRSRDGGI RFGEMERDCQ RKLVAQVFLGPTYYQRLKH LVDDKIHARA RGFVQILT RQFV GRSRDGGI RFGEMERDCQ RKLVSQVFLGPTYYQRLKH LVDDKIHARA RGFVQILT RQFV GRSRDGGI RFGEMERDCQ RKLVSQVFLGPTYYQRLKH LVDDKIHARA RGFVQILT RQFV GRSRDGGI RFGEMERDCQ EXTRACTION OF THE RESEARCH REPROPERTY OF THE RESEARCH RES	1137 1127 1127 1127 1127 1040 1310 1306 1306 1278 1278 1278 1278 1278 1278 1278 1278
tr S6ESB4 S6ESB4 ZYGB2 tr B6KSQ5 B6KSQ5 SCHJY sp Q02061 RPB2 SCHPO tr S9R8U4 S9R8U4 SCHCR sp Q8RQE9 RPOB THET8 ASR51304.1 OXR47929.1 WP 093971860.1 sp Q2RWR6 RPOB SODGM sp B4EYU9 RPOB PROMH sp A7FN13 RPOB YERP3 sp Q1C1U1 RPOB YERPA sp B2K113 RPOB YERPB sp A8G8E7 RPOB SERP5 sp Q6DANO RPOB PECCP sp Q7N9A4 RPOB PHOLL sp C5BHE3 RPOB PEDW19 sp A7MQ09 RPOB CROSS sp B5XYF5 RPOB ECOBW sp Q31U10 RPOB SHIBS sp Q32AF9 RROB SHIDS sp A8KT9 RPOB CITKS sp A8KT9 RPOB CITKS sp A8KT9 RPOB CITKS sp B5RFK1 RPOB SALG2	KKLMSQIFFGPTYYQRLKH. VVDDKIHARA RGPMQVLL RQPV GRSRDGGI RFGEMERDCM RKLVAQVFLGPTYYQRLKH. LVDDKIHARA RGPVQILT RQPV GRSRDGGI RFGEMERDCQ RKLVAQVFLGPTYYQRLKH. LVDDKIHARA RGPVQILT RQPV GRSRDGGI RFGEMERDCQ RKLVSQVFLGPTYYQRLKH. LVDDKIHARA RGPVQILT RQPV GRSRDGGI RFGEMERDCQ RKLVSQVFLGPTYYQRLKH. LVDDKIHARA RGPVQILT RQPV GRSRDGGI RFGEMERDCQ RKLVSQVFLGPTYYQRLKH. LVDDKIHARA RGPVQILT RQPV GRSRDGGI RFGEMERDCQ EPIEGPIVVQQMFIMKLYHVEDKMHARSTGPYSLVT QQPLGKAQFGGQ RFGEMEVWAL DKFDRKVTVGYIYMLKLHH. LVDDKMHARSTGPYSLVT QQPLGGKAQFGGQ RFGEMEVWAL EQFERPVTVGYMYMLKLHH. LVDDKMHARSTGFYSLVT QQPLGGKAQFGG RFGEMEVWAL EQFERQVTVGYMYMLKLHH. LVDDKMHARSTGSYSLVT QQPLGGKAQFGG RFGEMEVWAL EQFERPVTVGYMYMLKLHH. LVDDKMHARSTGSYSLVT QQPLGGKAQFGG RFGEMEVWAL EQFERPVTVGYMYMLKLHHLVDDKMHARSTGSYSLVT QQPLGGKAQFGG RFGEMEVWAL EQ	1137 1127 1127 1127 1127 1040 1310 1306 1306 1278 1278 1278 1278 1278 1278 1278 1278
tr S6ESB4 S6ESB4 ZYGB2 tr B6KSQ5 B6KSQ3 SCHJY sp Q02061 RPB2 SCHPO tr S9R8U4 S9R8U4 SCHOY tr S9R8U4 S9R8U4 SCHOY tr S9R8U6 S9W8C6 SCHCR sp Q8RQE9 RPOB THET8 ASR51304.1 OXR47929.1 WP 093971860.1 sp Q2WWR6 RPOB SODGM sp B4EYU9 RPOB PROMH sp A7FN13 RPOB YERPA sp B2K113 RPOB YERPA sp B2K113 RPOB YERPA sp B2K113 RPOB PECAS sp C6DHR5 RPOB PECCP sp Q7N9A4 RPOB PHOLL sp C5BHE3 RPOB PHOLL sp C5BHE3 RPOB EDWI9 sp A7MQ09 RPOB CROS8 sp B5XYF5 RPOB KLEP3 sp Q32AF9 RPOB SHIBS sp Q32AF9 RPOB SHIDS sp A8AKT9 RPOB CITKS sp B5RFK1 RPOB SALG2 sp B5BJQ3 RPOB SALFK	KKLMSQIFFGPTYYQRLKH. VVDDKIHARA RGPMQVLI RQPV GRSRDGGI RFGEMERDCM RKLVAQVFLGPTYYQRLKH LVDDKIHARA RGPVQILI RQPV GRSRDGGI RFGEMERDCQ RKLVSQVFLGPTYYQRLKH LVDDKIHARA RGPVQILI RQPV GRSRDGGI RFGEMEVDCQ EPIEGP IVVGQMFIMKLYHIVEDKMHARS TGPYSLIT QQPLGGKAQFGGQ RFGEMEVWAL EQFERPVTVGYMYMLKLHH LVDDKMHARS TGPYSLVT QQPLGGKAQFGG RFGEMEVWAL EQFERQVTVGYMYMLKLHH LVDDKMHARS TGSYSLVT QQPLGGKAQFGG RFGEMEVWAL EQFERPVTVGYMYMLKLHH LVDDKMHARS TGSYSLVT QQPLGGKAQFGG	1137 1127 1127 1127 1127 1127 1040 1310 1306 1306 1278 1278 1278 1278 1278 1278 1278 1278
tr S6ESB4 S6ESB4 ZYGB2 tr B6KSQ5 B6KSQ5 SCHJY sp Q02061 RPB2 SCHPO tr S9R8U4 S9R8U4 SCHOY tr S9R8U4 S9R8U4 SCHOY tr S9R8U6 S9W8C6 SCHCR sp Q8RQE9 RPOB THET8 ASR51304.1 OXR47929.1 WP 093971860.1 sp Q2NWR6 RPOB SODGM sp B4EYU9 RPOB PROMH sp A7FN13 RPOB YERPA sp B2K113 RPOB YERPA sp B2K113 RPOB YERPA sp B2K113 RPOB PECAS sp C6DHR5 RPOB PECCP sp Q7N9A4 RPOB PHOLL sp C5BHE3 RPOB EMILE sp C5BHE3 RPOB EMILE sp C5AOS7 RPOB ECOLI sp C5AOS7 RPOB ECOLI sp Q31U10 RPOB SHIBS sp Q32AF9 RPOB SHIDS sp A8AKT9 RPOB SALG2 sp B5BFKK1 RROB SALG2 sp B5BFJQ3 RPOB SALPK sp B4TOY9 RPOB SALNS	KKLMSQIFFGPTYYQRLKH LVDDKIHARA RGPMQVLI RQPV GRSRDGGI RFGEMERDCM RKLVAQVFLGPTYYQRLKH LVDDKIHARA RGPVQILTRQPV GRSRDGGI RFGEMERDCQ RKLVSQVFLGPTYYQRLKH LVDDKHHARA RGPVQILTRQPV GRSRDGGI RFGEMERDCQ RKLVSQVFLGPTYYQRLKH LVDDKHHARA RGPVQILTRQPV GRSRDGGI RFGEMERDCQ RKLVSQVFLGPTYYQRLKH LVDDKHHARA RGPVQILTRQPV GRSRDGGI RFGEMEVWAL DKFDRKVTVGYIYMKLHHLVDDKHHARA RGPVSLVTQQPLGGKAQFGGQ RFGEMEVWAL EQFERPVTVGYMMYLKLHHLVDDKMHARA RGPVSLVTQQPLGGKAQFGGQ RFGEMEVWAL EQFERQVTVGYMYMLKLHHLVDDKMHARA RGSYSLVTQQPLGGKAQFGGQ RFGEMEVWAL EQFERPVTVGYMYMLKLHHLVDDKMHARA RGSYSLVTQQPLGGKAQFGGQ RFGEMEVWAL EQFERPVTVGYMYMLKLHLLVDDKMHARA RGSYSLVTQQPLGGKAQFGGQ RFGEMEVWAL EQFERPVTVGYMYMLKLHLLVDDKMHARA RGSYSLVTQQPLGGKAQFGGQ RFGEMEVWAL EQFERPVTVGYMYMLKLHHLVDDKMHARA RGSYSLVTQQPLGGKAQFGGQ RFGEMEVWAL EQFERPVTVGYMYMLKLHHLVDDKMHARA RGSYSLVTQQPLGGKAQFGGQ RFGEMEVWAL EQFERPVTVGYMYMLKLHHLVD	1137 1127 1127 1127 1127 1127 1040 1310 1306 1306 1278 1278 1278 1278 1278 1278 1278 1278
tr S6ESB4 S6ESB4 ZYGB2 tr B6KSQ5 B6KSQ3 SCHJY sp Q02061 RPB2 SCHPO tr S9R8U4 S9R8U4 SCHOY tr S9R8U4 S9R8U4 SCHOY tr S9R8U6 S9W8C6 SCHCR sp Q8RQE9 RPOB THET8 ASR51304.1 OXR47929.1 WP 093971860.1 sp Q2WWR6 RPOB SODGM sp B4EYU9 RPOB PROMH sp A7FN13 RPOB YERPA sp B2K113 RPOB YERPA sp B2K113 RPOB YERPA sp B2K113 RPOB PECAS sp C6DHR5 RPOB PECCP sp Q7N9A4 RPOB PHOLL sp C5BHE3 RPOB PHOLL sp C5BHE3 RPOB EDWI9 sp A7MQ09 RPOB CROS8 sp B5XYF5 RPOB KLEP3 sp Q32AF9 RPOB SHIBS sp Q32AF9 RPOB SHIDS sp A8AKT9 RPOB CITKS sp B5RFK1 RPOB SALG2 sp B5BJQ3 RPOB SALFK	KKLMSQIFFGPTYYQRLKH. VVDDKIHARA RGPMQVLI RQPV GRSRDGGI RFGEMERDCM RKLVAQVFLGPTYYQRLKH LVDDKIHARA RGPVQILI RQPV GRSRDGGI RFGEMERDCQ RKLVSQVFLGPTYYQRLKH LVDDKIHARA RGPVQILI RQPV GRSRDGGI RFGEMEVDCQ EPIEGP IVVGQMFIMKLYHIVEDKMHARS TGPYSLIT QQPLGGKAQFGGQ RFGEMEVWAL EQFERPVTVGYMYMLKLHH LVDDKMHARS TGPYSLVT QQPLGGKAQFGG RFGEMEVWAL EQFERQVTVGYMYMLKLHH LVDDKMHARS TGSYSLVT QQPLGGKAQFGG RFGEMEVWAL EQFERPVTVGYMYMLKLHH LVDDKMHARS TGSYSLVT QQPLGGKAQFGG	1137 1127 1127 1127 1127 1127 1040 1310 1306 1306 1278 1278 1278 1278 1278 1278 1278 1278

sp P30876 RPB2_HUMAN	ACKLLFQELMSMS IAPRMMSV	1174
tr G3V8Y5 G3V8Y5_RAT	ACKLLFQELMSMS IAPRMMSV	1174
tr A0A250Y753 A0A250Y753 CASCN	ACKLLFQELMSMS IAPRMMSV	1174
tr A0A1U7R4C7 A0A1U7R4C7 MESAU	ACKLLFQELMSMS IAPRMMSV	1174
tr A0A286XIQ9 A0A286XIQ9 CAVPO	ACKLLFOELMSMSIAPRMMSV	1174
tr I3M351 I3M351 ICTTR	ACKLLFQELMSMSIAPRMMSV	1174
tr G7P5R6 G7P5R6 MACFA	ACKLLFQELMSMSIAPRMMSV	1174
tr H2OPI8 H2OPI8 PANTR	ACKLLFQELMSMS IAPRMMSV	1174
· · · · · · · · · · · · · · · · · · ·		
tr A0A1U7V0T5 A0A1U7V0T5_TARSY	ACKLLFQELMSMS IAPRMMSV	1174
tr A0A1S2ZSL2 A0A1S2ZSL2_ERIEU	ACKLLFQELMSMS IAPRMMSV	1174
tr A0A0D9QYL1 A0A0D9QYL1_CHLSB	ACKLLFQELMSMS IAPRMMSV	1167
tr A0A2K5ZNR7 A0A2K5ZNR7_MANLE	ACKLLFQELMSMSIAPRMMSV	1167
tr A0A2I2ZIU3 A0A2I2ZIU3_GORGO	ACKLLFQELMSMSIAPRMMSV	1174
tr A0A1D5QGA5 A0A1D5QGA5 MACMU	ACKLLFQELMSMSIAPRMMSV	1174
tr A0A2J8S2N1 A0A2J8S2N1 PONAB	ACKLLFQELMSMSIAPRMMSV	1174
tr A0A2K5K5J5 A0A2K5K5J5 COLAP	ACKLLFQELMSMS IAPRMMSV	1174
tr A0A2J8PEW7 A0A2J8PEW7 PANTR	ACKLLFQELMSMS IAPRMMSV	1167
tr A0A2K5CY83 A0A2K5CY83 AOTNA	ACKLLFQELMSMS IAPRMMSV	1174
	ACKLLFQELMSMSIAPRMMSV	1174
tr A0A096NEY4 A0A096NEY4_PAPAN		
tr C9J2Y9 C9J2Y9_HUMAN	ACKLLFQELMSMSIAPRMMSV	1167
tr G8BY61 G8BY61_TETPH	AAKLLFQELMAMNITPRLYTDRSKNF	1224
tr A0A1X7QYA1 A0A1X7QYA1_9SACH	AAKLLFQELMAMNITPRLYTDRSRDF	1221
tr J7RV95 J7RV95_KAZNA	AAKLLFQELMAMNITPRLYTDRSRDF	1220
tr H2AVJ8 H2AVJ8 KAZAF	AAKLLFQELMAMNITPRLFTDRSRDF	1222
sp Q6FLD5 RPB2 CANGA	AAKLLFQELMAMNITPRLYTDRSRDF	1223
sp P08518 RPB2 YEAST	AAKLLFQELMAMNITPRLYTDRSRDF	1224
tr A0A0L8VHA5 A0A0L8VHA5 9SACH	AAKLLFOELMAMNITPRLYTDRSRDF	1224
	AAKLLFQELMAMNITPRLYTDRSRDF	1224
tr A0A0L8RB33 A0A0L8RB33_SACEU		
tr G0VJ71 G0VJ71_NAUCC	AAKLLFQELMAMNITPRLYTERSRDF	1224
tr G8ZM49 G8ZM49_TORDC	AAKLLFQELMAMNITPRLYTDRSKDF	1222
tr A0A1Q3A090 A0A1Q3A090_ZYGRO	AAKLLFQELMAMNITPRLYTDRSKDF	1224
tr A0A0N7IS35 A0A0N7IS35_9SACH	AAKLLFQELMAMNITPRLYTDRSKDF	1222
tr A0A212MG88 A0A212MG88 ZYGBA	AAKLLFQELMAMNITPRLYTDRSKDF	1223
tr A0A1S7HHE1 A0A1S7HHE1 9SACH	AAKLLFQELMAMNITPRLYTDRSKDF	1223
tr S6ESB4 S6ESB4 ZYGB2	AAKLLFQELMAMNITPRLYTDRSKDF	1223
tr B6K5Q5 B6K5Q5 SCHJY	AAKLLFQELMSMNIAPRLFTKSHH	1210
sp Q02061 RPB2 SCHP0	AAKLLFQELMSMN IAPRLFTKNHK	1210
	AAKLLFQELMSMNIAPRLFTKNHKI	1211
tr S9R8U4 S9R8U4_SCHOY		
tr S9W8C6 S9W8C6_SCHCR	AAKLLFQELMSMNIAPRLFTKNHKN	1211
LOGDOTO LE DOD. MUTANO		
sp Q8RQE9 RPOB_THET8	SFRVLVKELQALALDVQTLDEKDNPVDIFEGLASKR	1119
ASR51304.1	SFNVLVKEMRSLGLNVELNSIDALPDPDEIAEAAE	1388
OXR47929.1	SFNVLVKEIRSLSLDMDLERN	1370
WP_093971860.1	SFNVLVKEIRSLSLDMIDLERN	1370
sp Q2NWR6 RPOB_SODGM	SFNVLLKEIRSLGINTELEED	1342
sp B4EYU9 RPOB_PROMH	SFNVLLKEIRSLGINTELEDE	1342
sp A7FNI3 RPOB_YERP3	SFNVLLKEIRSLGINTELEEE	1342
sp Q1C1U1 RPOB YERPA	SFNVLLKEIRSLGINTELEEE	1342
sp B2K113 RPOB YERPB	SFNVLLKEIRSLGIN ELEEE	1342
sp A8G8E7 RPOB_SERP5	SENVILKETESIGINTELEGE	1342
sp Q6DAN0 RPOB PECAS	SFNVLLKEIRSLGIN ELEEK	1342
sp/C6DHR5/RPOB_PECCP	SFNVLLKEIRSLGINIELEEE	
- <u>-</u>	SFNVLLKEIRSLGINIELEGE	1342
sp Q7N9A4 RPOB_PHOLL	SENVEL METROLOGINI ELEGE	1342
sp C5BHE3 RPOB_EDWI9	SFNVLLKEIRSLGINIELEDE	1342
sp A7MQQ9 RPOB_CROS8	SFNVLLKEIRSLGINTELEDE	1342
sp B5XYF5 RPOB_KLEP3	SFNVLLKEIRSLGIN ELEDE	1342
sp POA8V2 RPOB_ECOLI	SFNVLLKEIRSLGINIELEDE	1342
sp C5A0S7 RPOB_ECOBW	SFNVLLKEIRSLGINTELEDE	1342
sp Q31U10 RPOB_SHIBS	SFNVLLKEIRSLGINTELEDE	1342
sp Q32AF9 RPOB SHIDS	SFNVLLKEIRSLGINTELEDE	1342
sp A8AKT9 RPOB CITK8	SENVLLKEIRSLGINTELEDE	1342
sp B5RFK1 RPOB SALG2	SFNVLLKEIRSLGIN ELEDE	1342
sp B5BJQ3 RPOB SALPK	SFNVLLKEIRSLGIN ELEDE	1342
sp B4T0Y9 RPOB_SALNS	SFNVLLKEIRSLGIN ELEDE	1342
sp P06173 RPOB_SALTY	SFNVLLKEIRSLGINIELEDE	1342
PPIEGGI/SIKEOD_SHLII		1342
	: .:*.:*: :: :	

Fig. 5 Mix and match analysis of the initiation subunits β of *E. coli* and Rpb2 of yeast RNAP II For figure legends, refer to Figs. 3 and 4

CLUSTAL O (1.2.4): MSA of eubacterial β' and eukaryotic Rpb1 subunits 6



```
ATMVDNELPGLPRAMQKSGRPLKSIKQRLKGKEGRVRCNLMGKRVDFSARTVITPDPNLQ
tr|A0A1A7X327|A0A1A7X327_9TELE
tr|A0A1A8UKD7|A0A1A8UKD7_NOTFU
                                                                                                                                                      372
                                                     ATMVDNELPGLPRAMQKSGRPLKSLKQRLKGKEGRVRCNLMGKRVDFSARTVIT PDPNLQ
                                                                                                                                                      372
                                                    ATMYDNELPGLPRAMQKSGRPLKSIKQRLKGKEGRVRGNLMGKRVDFSARTVITPDPNLQ
ATMYDNELPGLPRAMQKSGRPLKSIKQRLKGKEGRVRGNLMGKRVDFSARTVITPDPNLQ
ATMYDNELPGLPRAMQKSGRPLKSIKQRLKGKEGRVRGNLMGKRVDFSARTVITPDPNLQ
tr|A0A1A8ER05|A0A1A8ER05_9TELE
                                                                                                                                                      372
tr|A0A1A8DQ60|A0A1A8DQ60_9TELE
tr|A0A1A8NSR8|A0A1A8NSR8_9TELE
tr|A0A1W4YLM7|A0A1W4YLM7_9TELE
                                                                                                                                                      372
                                                    ATMVDNELPGLPRAMQKSGRPLKSIKQRLKGKEGRVRGNLMGKRVDFSARTVITPDPNLQ
                                                                                                                                                      372
                                                    ATYMDNDIAGLPQSLQSSGRPVKAIRARLKGKEGRLRCNLMGKRVDFSARTVITGDPNLE
tr|A0A1M8A6L7|A0A1M8A6L7 MALS4
                                                                                                                                                      374
sp|P04050|RPB1_YEAST
tr|A0A1B2J8C6|A0A1B2J8C6_PICPA
tr|F2QW17|F2QW17_KOMPC
tr|A3GID7|A3GID7_PICST
                                                    ATYMDNDIAGOPOALOKSGR PVKAIRARLKGKE GRLR (NLMGKRVDF SAR TVIS SDPNLE
ATYMDNDIAGOPOALOKSGR PVKAIRARLKGKE GRLR (NLMGKRVDF SAR TVIS SDPNLE
ATYMDNDIAGOPOALOKTGR PIKSIRARLKGKE GRLR (NLMGKRVDF SAR TVIS SDPNLD
                                                                                                                                                      361
                                                                                                                                                      361
                                                                                                                                                      360
                                                    ATYMDNDIAGOPOALOKTGRPIKSIRARLKGKEGRIRGNLMGKRVDFSARTVISGDPNLD
ATYMDNDIAGOPOALOKTGRPIKSIRARLKGKEGRIRGNLMGKRVDFSARTVISGDPNLD
tr|A0A1D8PUA6|A0A1D8PUA6_CANAL
                                                                                                                                                      360
tr|G8BEH9|G8BEH9_CANPC
                                                                                                                                                      360
AEG34223.1
                                                     DALLDNGRRGAPVTNPGSDRPLRSLTDILSGKQGRFR
                                                                                                                                                      638
ASR51305.1
                                                     DALFDNGRRGRIITG-ANKRPLKSLSDMLKGKQGRFR
                                                                                                                                                      363
OXR47930.1
                                                    DSLLDNGRRGKAMTG-ANKROLKSLADMIKGKSGRFR(
                                                                                                                                      POLK
                                                                                                                                                      362
sp|A7MQQ8|RPOC_CROS8
sp|Q32AG0|RPOC_SHIDS
sp|Q0SY12|RPOC_SHIF8
                                                    DALLDNGRRGRAITG-SNKRPLKSLADMIKGKOGRFR
                                                                                                                                     GPYLR
                                                                                                                                                      362
                                                     DALLDNGRRGRAITG-SNKRPLKSLADMIKGKQGRFR
                                                    DALLDNGRRGRAITG-SNKRPLKSLADMIKGKOGRFR
                                                                                                                                      PYT.R
                                                                                                                                                      362
                                                     DALLDNGRRGRAITG-SNKRPLKSLADMIKGKQGRFRONLLGKRVDYSGRSVIT
sp|B2TWH4|RPOC SHIB3
                                                                                                                                                      362
sp|P0A8T7|RPOC_ECOLI
sp|Q3YUZ6|RPOC_SHISS
sp|B1XBZ0|RPOC_ECODH
sp|A8A787|RPOC_ECOHS
                                                                                                                                                      362
                                                    DALLDNGRRGRAITG-SNKRPLKSLADMIKGKOGRFRONLLGKRVDYSGR$VIT
                                                                                                                                     GPYLR
                                                    DALLDNGRRGRAITG-SNKRPLKSLADMIKGKQGRFR
                                                                                                                                     SPYLR
                                                                                                                                                      362
                                                     DALLDNGRRGRAITG-SNKRPLKSLADMIKGKQGRFR
                                                                                                                                                      362
tr|A0A237JUP3|A0A237JUP3|SHISO
tr|A0A0F1RBF2|A0A0F1RBF2_ENTAS
tr|A0A1B3EWG0|A0A1B3EWG0_ENTCL
                                                    DALLDNGRRGRAITG-SNKRPLKSLADMIKGKOGRFRDALLDNGRRGRAITG-SNKRPLKSLADMIKGKOGRFR
                                                                                                                                      DVI.D
                                                                                                                                                      362
                                                                                                                                                      362
                                                                                                                                      PYLR
                                                     DALLDNGRRGRAITG-SNKRPLKSLADMIKGKQGRFR
                                                                                                                                      PYLR
tr|A0A0F0XM62|A0A0F0XM62_9ENTR
                                                    DALLDNGRRGRAITG-SNKRPLKSLADMIKGKOGRFR
                                                                                                                                      PYLR
                                                                                                                                                      362
sp|Q5FK92|RFOC_SALFA
sp|A9MHE9|RFOC_SALAR
tr|A0A232XM43|A0A232XM43_SALMU
                                                    DALLDNGRRGRAITG-SNKRPLKSLADMIKGKQGRFR
                                                                                                                                      PYLR
                                                                                                                                                      362
                                                     DALLDNGRRGRAITG-SNKRPLKSLADMIKGKQGRFR
                                                                                                                                                      362
                                                    DALLDNGRRGRAITG-SNKRPLKSLADMIKGKOGRFR
                                                                                                                                      PYT.R
                                                                                                                                                      362
tr|B5RFK0|B5RFK0 SALG2
                                                    DALLDNGRRGRAITG-SNKRPLKSLADMIKGKQGRFR
                                                                                                                                     SPYLR
                                                                                                                                                      362
tr|BSRrKO|BSRrKO_SALTI
sp|POA2R5|RPOC_SALTI
sp|Q57H68|RPOC_SALCH
sp|POA2R4|RPOC_SALTY
sp|A6TGP1|RPOC_KLEP7
                                                     DALLDNGRRGRAITG-SNKRPLKSLADMIKGKQGRFR
                                                                                                                                     SPYLR
                                                                                                                                                      362
                                                    DALLDNGRRGRAITG-SNKRPLKSLADMIKGKOGRFRDALLDNGRRGRAITG-SNKRPLKSLADMIKGKOGRFR
                                                                                                                                      PYLR
                                                                                                                                                      362
                                                                                                                                     GPYLR
                                                                                                                                                      362
                                                     DALLDNGRRGRAITG-SNKRPLKSLADMIKGKQGRFR
                                                                                                                                      PYLR
                                                                                                                                                      362
tr|A0A0J2K6S7|A0A0J2K6S7_9ENTR
                                                    DALLDNGRRGRAITG-SNKRPLKSLADMIKGKQGRFR
                                                                                                                                      PYLR
                                                                                                                                                      362
tr|A0A0G3RZQ0|A0A0G3RZQ0_KLEOX
tr|A0A212HDS5|A0A212HDS5_9ENTR
                                                     DALLDNGRRGRAITG-SNKRPLKSLADMIKGKOGRFR
                                                                                                                                     SPYLR
                                                                                                                                                      362
                                                     DALLDNGRRGRAITG-SNKRPLKSLADMIKGKOGRFR
                                                                                                                                                      362
tr|A0A1R0FP41|A0A1R0FP41_CITBR
tr|A0A078LHA5|A0A078LHA5_CITKO
                                                    DALLDNGRRGRAITG-SNKRPLKSLADMIKGKQGRFR
                                                                                                                                      PYLR
                                                                                                                                                      362
                                                    DALLDNGRRGRAITG-SNKRPLKSLADMIKGKOGRFR
                                                                                                                                     GPYLR
                                                                                                                                                      362
sp|A8AKT8|RPOC CITK8
                                                     DALLDNGRRGRAITG-SNKRPLKSLADMIKGKQGRFR
                                                                                                                                     GPYLR
                                                                                                                                                      362
                                                                            . * ::::
                                                                                            :.**.**.*
tr|A0A1U8DYN0|A0A1U8DYN0_ALLSI
                                                     PSDLHLQIGYKVERHMCDGDIVI NRQPTLHKMSMMGHRVRILPWSTFRLNLSVTTPYN
                                                                                                                                                      493
tr|A0A1L8H4P4|A0A1L8H4P4_XENLA
                                                     PSDLHLQIGYKVERHMCDGDIVI FNR QPTLHKM SMMGHRVRIL PWST FRLNLSVTTPYN
                                                                                                                                                      492
                                                     PSDLHLQIGYKVERHMCDGDIVI TNROPTLHKM SMMGHRVRIL PWST FRLNLSVTT PYNA
PSDLHLQTGYKVERHMCDGDIVI FNROPTLHKM SMMGHRVRIL PWST FRLNLSVTT PYNA
PSDLHLQTGYKVERHMCDGDIVI FNROPTLHKM SMMGHRVRIL PWST FRLNLSVTT PYNA
tr|H9GLG5|H9GLG5_ANOCA
                                                                                                                                                      493
triH2R1J6|H2R1J6 PANTR
                                                                                                                                                      494
sp|P08775|RPB1_MOUSE
                                                                                                                                                      494
tr|G1MCZ1|G1MCZ1_AILME
                                                     PSDLHLQTGYKVERHMCDGDIVIFNRQPTLHKM SMMGHRVRILPWSTFRLNLSVTTPYNA
                                                                                                                                                      494
tr|008847|008847_MOUSE
tr|S7PWZ6|S7PWZ6_MYOBR
                                                     PSDLHLOTGYKVERHMCDGDIVI NROPTLHKMSMMGHRVRILPWSTFRLNLSVTTPYNA
PSDLHLOTGYKVERHMCDGDIVI NROPTLHKMSMMGHRVRILPWSTFRLNLSVTTPYNA
                                                                                                                                                      494
                                                                                                                                                      494
tr|D4A5A6|D4A5A6_RAT
                                                     PSDLHLQTGYKVERHMCDGDIVIFNRQPTLHKM SMMGHRVRILPWSTFRLNLSVTTPYNA
                                                                                                                                                      494
tr|A0A1S3EWL2|A0A1S3EWL2_DIPOR
                                                     PSDLHLOTGYKVERHMCDGDIVI NRQPTLHKMSMMGHRVRILPWSTFRLNLSVTTPYNA
PSDLHLOTGYKVERHMCDGDIVI NRQPTLHKMSMMGHRVRILPWSTFRLNLSVTTPYNA
                                                                                                                                                      494
sp|P11414|RPB1 CRIGR
                                                                                                                                                      494
 tr|035559|035559_CRIGR
                                                     PSDLHLQTGYKVERHMCDGDIVI FNRQPTLHKMSMMGHRVRIL PWST FRLNLSVTTPYNA
                                                                                                                                                      494
                                                     PSDLHLOTGYKVERHMCDGDIVIENROPTLHKMSMMGHRVRILPWSTFRLNLSVTTPYNA
PSDLHLOTGYKVERHMCDGDIVIENROPTLHKMSMMGHRVRILPWSTFRLNLSVTTPYNA
tr|A0A2I3M9H2|A0A2I3M9H2_PAPAN
                                                                                                                                                      494
tr|F7HB40|F7HB40 MACMU
                                                                                                                                                      494
tr|A0A2K6RYW9|A0A2K6RYW9_SAIBB
                                                     PSDLHLQTGYKVERHMCDGDIVI FNRQPTLHKMSMMGHRVRILPWSTFRLNLSVTTPYNA
                                                                                                                                                      494
tr|W5N8Z6|W5N8Z6_LEPOC
tr|I3JRW6|I3JRW6 ORENI
                                                     PSDLHLOIGYKVERHMCDGDIVIENROPTLHKMSMMGHRVRILPWSTFRLNLSVTTPYNA
PSDLHLOIGYKVERHMCDGDIVIENROPTLHKMSMMGHRVRILPWSTFRLNLSVTTPYNA
                                                                                                                                                      493
                                                                                                                                                      492
tr|A0A0R4IMS9|A0A0R4IMS9_DANRE
                                                     PSDLHLQIGYKVERHMCDGDIIV NRQPTLHKMSMMGHRVRILPWSTFRLNLSVTTPYNA
                                                                                                                                                      491
tr|A0A1A7X327|A0A1A7X327_9TELE
tr|A0A1A8UKD7|A0A1A8UKD7 NOTFU
                                                     PSDLHLOIGYKVERHMCEGDIVI NROPTLHKMSMMGHRVRILPWSTFRLNLSVTTPYNA
PSDLHLOIGYKVERHMCDGDIVI NROPTLHKMSMMGHRVRILPWSTFRLNLSVTTPYNA
                                                                                                                                                      492
                                                                                                                                                      492
tr|A0A1ASENCS|A0A1ASERO5 9TELE

tr|A0A1ASENO5|A0A1ASERO5 9TELE

tr|A0A1ASENOSR8|A0A1ASENOSR8 9TELE

tr|A0A1ASENOSR8|A0A1ASENOSR8 9TELE

tr|A0A1MSA6L7|A0A1MSA6L7 MALS4
                                                     PSDLHLQIGYKVERHMCDGDIVI FNRQPTLHKMSMMGHRVRILPWSTFRLNLSVTTPYNA
                                                                                                                                                      492
                                                     PSDLHLQIGYKVERHMCDGDIVI NRQPTLHKMSMMGHRVRILPWSTFRLNLSVTTPYNA
PSDLHLQIGYKVERHMCDGDIVI NRQPTLHKMSMMGHRVRILPWSTFRLNLSVTTPYNA
                                                                                                                                                      492
                                                                                                                                                      492
                                                     PSDLHLQIGYKVERHMCDGDIII NRQPTLHKMSMMGHRVRILPWSTFRLNLSVTTPYNA
RGDIALQAGWIVERHLKDGDYVLENRQPSLHKMSMMAHRVKLMDYSTFRLNLSVTPPYNA
                                                                                                                                                      492
                                                                                                                                                      493
sp/P04050/RPB1 YEAST
tr|A0A1B2J8C6|A0A1B2J8C6_PICPA
                                                     AGD IVLQYGWKVERHLMDDD PVL FNRQPSLHKM SMMAHRVKVM PYST FRLNLSVT SPYNA
                                                                                                                                                      481
tr|F2QW17|F2QW17_KOMPC
tr|A3GID7|A3GID7 PICST
                                                    AGDIVLOYGWKVERHLMDDDPVLINROPSLHKMSMMAHRVKVMPYSTFRLNLSVTSPYNA
AGDIALOYGWKVERHLMDNDPVLINROPSLHKMSMMAHRVKVMPYSTFRLNLSVTSPYNA
                                                                                                                                                      481
                                                                                                                                                      480
tr|A0A1D8PUA6|A0A1D8PUA6_CANAL
                                                     AGDIALQYGWKVERHLMDDDPVL#NRQPSLHKM$MMAHRVRVMPYSTFRLNLSVT$PYN
                                                                                                                                                      480
triG8BEH9IG8BEH9 CANPC
                                                     AGDIALOYGWKVERHLMDDDFVL#NROPSLHKM$MMAHRVKVMFYSTFRLNLSVT$FYNA
                                                                                                                                                      480
AEG34223.1
                                                     ORD I-KDEVWDALEEV IHGKVVL
                                                                                                      IOA FOPVLVEGOS IOLHPL
                                                                                                                                                      738
                                                            -KEVWDILDEVIREHPVM
ASR51305.1
                                                                                                                                                      460
OXR47930.1
                                                     OE----PVVWDILEEVIREHPVMI
                                                                                                      TOA FERVILLEGRA TOT HELD
                                                                                                                                                      459
sp|A7MQQ8|RPOC_CROS8
sp|Q32AG0|RPOC_SHIDS
sp|Q0SY12|RPOC_SHIF8
sp|B2TWH4|RPOC_SHIF8
                                                     EE----AVVWDILDEVIREHPVLINRAPTLHRI
                                                                                                     SIQAFEPVLIEGKA IQLHPLV
                                                                                                                                                      459
                                                     EE----AVVWDILDEVIREHPVLINRAPTLHRI
                                                                                                     GIQAFEPVLIEGKAIQLHPLV
                                                                                                                                                      459
                                                     EE---- AVVWDTI.DEV TREH PVT.
                                                                                                       OAFEPVLIEGKAIOLHPL
                                                                                                                                                      459
                                                     EE----AVVWDILDEVIREHPVLINRAPTLHRLGIQAFEPVLIEGKAIQLHPLV
                                                                                                                                                      459
 sp|Q3YUZ6|RPOC_SHISS
                                                     EE----AVVWDILDEVIREH PVLINRAPTLHRLGIQA FEPVLIEGKA IQLHPLVCAA YNA
```

```
sp|B1XBZ0|RPOC ECODH
                                          EE----AVVWDILDEVIREHPVLINRAPTLHRLGIQAFEPVLIEGKAIQLHPLVCAAYNA
                                                                                                                      459
sp|A8A787|RPOC_ECOHS
tr|A0A237JUP3|A0A237JUP3 SHISO
                                          EE----AVVWDILDEVIREHPV
                                                                                 GIQAFEPVLIEGKAIQLHPLVCAAYNA
                                                                                                                      459
                                         EE----AVVWDILDEVIREHP
                                                                                 JOAFEPVLIEGKAIOLHPLVCAAYNA
                                                                     LNRAPTLHR
                                                                                                                      459
                                             ---AVVWDILDEVIREHP
tr|A0A0F1RBF2|A0A0F1RBF2 ENTAS
                                                                    LNRAPTLHR
                                                                                 IQAFEPVLIEGKAIQLHPLVCAAYNA
                                                                                                                      459
tr|A0A1B3EWG0|A0A1B3EWG0_ENTCL
                                          EE----AVVWDILDEVIREHPV
                                                                    LNRAPTLHR
                                                                                 :IOAFEPVLIEGKAIOLHPLVCAAYNA
                                                                                                                      459
tr|A0A0F0XM62|A0A0F0XM62_9ENTR
                                          EE----AVVWDILDEVIREHPY
                                                                                SIQAFEPVLIEGKAIQLHPLVCAAYNA
                                                                    LNRAPTLHRI
                                                                                                                      459
sp|Q5PK92|RPOC_SALPA
                                          EE----AVVWDILDEVIREHPV
                                                                    LNRAPTLHRI
                                                                                GIQAFEPVLIEGKAIQLHPLVCAAYNA
                                                                                                                      459
sp|A9MHE9|RPOC_SALAR
                                          EE----AVVWDILDEVIREHPV
                                                                    LNRAPTLHR
                                                                                GIOAFEPVLIEGKAIOLHPLVCAAYNA
                                                                                                                      459
tr|A0A232XM43|A0A232XM43 SALMU
                                         EE----AVVWDILDEVIREHPY
                                                                    LNRAPTLHR
                                                                                 :IOAFEPVLIEGKAIOLHPLVCAAYNA
                                                                                                                      459
tr|B5RFK0|B5RFK0 SALG2
                                         EE----AVVWDILDEVIREHPV
                                                                                GIQAFEPVLIEGKAIQLHPLVCAAYNA
                                                                   LLNRAPTLHR
                                                                                                                      459
sp|P0A2R5|RPOC_SALTI
                                          EE----AVVWDILDEVIREHPV
                                                                    LINRAPTLHR
                                                                                 GIQAFEPVLIEGKAIQLHPLVCAAYNA
                                                                                                                      459
sp|Q57H68|RPOC_SALCH
sp|P0A2R4|RPOC_SALTY
                                          EE----AVVWDILDEVIREHPV
                                                                   LLNRAPTLHRI
                                                                                GIQAFEPVLIEGKAIQLHPLVCAAYNA
                                          EE----AVVWDILDEVIREHPV
                                                                    LNRAPTLHR
                                                                                SIQAFEPVLIEGKAIQLHPLVCAAYNA
                                                                                                                      459
sp|A6TGP1|RPOC KLEP7
                                          EE----AVVWDILDEVIREHPV
                                                                    LNRAPTLHRI
                                                                                SIQAFEPVLIEGKAIQLHPLVCAAYNA
                                                                                                                      459
tr|A0A0J2K6S7|A0A0J2K6S7_9ENTR
                                          VQHAQTVAU''LLUMAAA----AA
                                                                   LLNRAPTLHRI
                                                                                GIOAFEPVLIEGKAIOLHPLVCAAYNA
                                                                                                                      459
tr|A0A0G3RZQ0|A0A0G3RZQ0_KLEOX
tr|A0A212HDS5|A0A212HDS5_9ENTR
tr|A0A1R0FP41|A0A1R0FP41_CITBR
                                          EE----AVVWDILDEVIREHPV
                                                                                GIOAFEPVLIEGKAIOLHPLVCAAYNA
                                                                    LNRAPTLHRI
                                                                                                                      459
                                          EE----AVVWDILDEVIREHPV
                                                                   LLNRAPTLHRI
                                                                               GIQAFEPVLIEGKAIQLHPLVCAAYNA
                                                                                                                      459
                                          EE----AVVWDILDEVIREHPV
                                                                   LLNRAPTLHRI
                                                                                GIQAFEPVLIEGKAIQLHPLVCAAYNA
                                                                                                                      459
                                                                               GIQAFEPVLIEGKAIQLHPLVCAAYNA
tr|A0A078LHA5|A0A078LHA5_CITKO
                                          EE----AVVWDILDEVIREHPVLLNRAPTLHRI
sp|A8AKT8|RPOC CITK8
                                          EE----AVVWDILDEVIREHPVLLNRAPTLHRI
                                                                               GIQAFEPVLIEGKAIQLHPLVCAAYNA
                                                               . :::** *:**::
                                                                                        ::
tr|A0A1U8DYN0|A0A1U8DYN0 ALLSI
                                                  NLHLPQSLETRAEIQELAMVPRMIVTPQSNRPVMGIVQDTLTAVRKFTKRDVF
                                                                                                                      553
tr|A0A1L8H4P4|A0A1L8H4P4 XENLA
                                                  NLHLPQSLETRAEIQELAMVPRMIVTPQSNRPVMGIVQDTLTAVRKFTKRDVF
                                                                                                                      552
tr|H9GLG5|H9GLG5 ANOCA
                                                  NLHLPQSLETRAEIQELAMVPRMIVTPQSNRPVMGIVQDTLTAVRKFTKRDVF
tr|H2R1J6|H2R1J6_PANTR
                                                  NLHLPQSLETRAEIQELAMVPRMIVTPQSNRPVMGIVQDTLTAVRKFTKRDVF
                                                                                                                      554
sp|P08775|RPB1 MOUSE
                                                  NLHLPQSLETRAEIQELAMVPRMIVTPQSNRPVMGIVQDTLTAVRKFTKRDVF
                                                                                                                      554
tr|G1MCZ1|G1MCZ1_AILME
tr|O08847|O08847_MOUSE
                                                  NLHLPOSLETRAEIQELAMVPRMIVTPOSNRPVMGIVODTLTAVRKFTKRDVF
                                                                                                                      554
                                                  NLHLPQSLETRAEIQELAMVPRMIVTPQSNRPVMGIVQDTLTAVRKFTKRDVF
                                                                                                                      554
tr|S7PWZ6|S7PWZ6 MYOBR
                                                  NLHLPQSLETRAEIQELAMVPRMIVTPQSNRPVMGIVQDTLTAVRKFTKRDVF
tr|D4A5A6|D4A5A6 RAT
                                                 NLHLPQSLETRAEIQELAMVPRMIVTPQSNRPVMGIVQDTLTAVRKFTKRDVF
                                                                                                                      554
tr|A0A1S3EWL2|A0A1S3EWL2 DIPOR
                                                  NLHLPQSLETRAEIQELAMVPRMIVTPQSNRPVMGIVQDTLTAVRKFTKRDVF
                                                                                                                      554
sp|P11414|RPB1_CRIGR
tr|035559|035559 CRIGR
                                                  NLHLPQSLETRAEIQELAMVPRMIVTPQSNRPVMGIVQDTLTAVRKFTKRDVF
                                                                                                                      554
                                                  NLHLPQSLETRAEIQELAMVPRMIVTPQSNRPVMGIVQDTLTAVRKFTKRDVF
                                                                                                                      554
tr|A0A2I3M9H2|A0A2I3M9H2 PAPAN
                                                  NLHLPQSLETRAEIQELAMVPRMIVTPQSNRPVMGIVQDTLTAVRKFTKRDVF
                                                                                                                      554
tr|F7HB40|F7HB40 MACMU
                                                  NLHLPQSLETRAEIQELAMVPRMIVTPQSNRPVMGIVQDTLTAVRKFTKRDVF
                                                                                                                      554
tr|A0A2K6RYW9|A0A2K6RYW9_SAIBB
                                                  NLHLPQSLETRAEIQELAMVPRMIVTPQSNRPVMGIVQDTLTAVRKFTKRDVF
tr|W5N8Z6|W5N8Z6_LEPOC
                                                  NLHLPQSLETRAEIQELAMVPRMIVTPQSNRPVMGIVQDTLTAVRKFTKRDVF
                                                                                                                      553
tr||I3JRW6||I3JRW6 ORENI
                                                  NLHLPQSLETRAEIQELAMVPRMIVTPQSNRPVMGIVODTLTAVRKFTKRDVF
                                                                                                                      552
tr|A0A0R4IMS9|A0A0R4IMS9 DANRE
                                                  NLHLPOSLETRAEIOELAMVPRMIVTPOSNRPVMGIVODTLTAVRKFTKRDVF
                                                                                                                      551
tr|A0A1A7X327|A0A1A7X327_9TELE
tr|A0A1A8UKD7|A0A1A8UKD7_NOTFU
                                                  NLHLPQSLETRAEIQELAMVPRMIVTPQSNRPVMGIVQDTLTAVRKFTKRDVF
                                                                                                                      552
                                                  NLHLPQSLETRAEIQELAMVPRMIVTPQSNRPVMGIVQDTLTAVRKFTKRDVF
tr|A0A1A8ER05|A0A1A8ER05 9TELE
                                                  NLHLPQSLETRAEIQELAMVPRMIVTPQSNRPVMGIVQDTLTAVRKFTKRDVF
                                                                                                                      552
tr|A0A1A8DQ60|A0A1A8DQ60 9TELE
                                                  NLHLPQSLETRAEIQELAMVPRMIVTPQSNRPVMGIVQDTLTAVRKFTKRDVF
                                                                                                                      552
tr|A0A1A8NSR8|A0A1A8NSR8_9TELE
tr|A0A1W4YLM7|A0A1W4YLM7_9TELE
                                                  NLHLPQSLETRAEIQELAMVPRMIVTPQSNRPVMGIVQDTLTAVRKFTKRDVF
                                                                                                                      552
                                                  NLHLPQSLETRAEIQELAMVPRMIVTPQSNRPVMGIVQDTLTAVRKFTKRDVF
                                                                                                                      552
tr|A0A1M8A6L7|A0A1M8A6L7 MALS4
                                                  NLHVPQSEEARAELAQIAWVPRQIVSPQANKPVMGIVQDTLCGIRKFTLRDCL
                                                                                                                      553
sp|P04050|RPB1 YEAST
                                                   LHVPQSEETRAELSQLCAVPLQIVSPQSNKPCMGIVQDTLCGIRKLTLRDTF
                                                                                                                      540
tr|A0A1B2J8C6|A0A1B2J8C6 PICPA
                                                 NLHVPQSEETRAELSQLCAVPLQIVSPQSNKPVMGIVQDTLCGVRKMTLRDTF
tr|F2QW17|F2QW17_KOMPC
tr|A3GID7|A3GID7_PICST
                                                  NLHVPQSEETRAELSQLCAVPLQIVSPQSNKPVMGIVQDTLCGVRKMTLRDTF
                                                  NLHVPOSPETRAELSEICAVPLOIVSPOSNKPVMGIVODTLCGIRKMTLRDNF
                                                                                                                      540
tr|A0A1D8PUA6|A0A1D8PUA6 CANAL
                                                  NLHVPOSPETRSELSOICAVPLOIVSPOSNKPVMGIVODTLCGIRKMTLRDIF
                                                                                                                      540
tr|G8BEH9|G8BEH9_CANPC
                                                  NLHVPQSPETRSELSQICAVPLQIVSPQSNKPVMGIVQDTLCGIRKMTLRDIF
                                                                                                                      540
                                                  AVHVPLSSFAQAEARIQMLSAHNLLSPASGEPLAKPSRDIILGLYYITQVR-K
AEG34223.1
                                                                                                                       797
                                                   VHVPLSLEAQLEARVLMMSTNNILSPANGKPIIVPSQDMVLGLYYLSMDR-E
OXR47930.1
                                           <mark>FDGDQM</mark>AVHVPLSLEAQLEARTLMLASNNVLFPANGEPSIVPSQDIVLGLYYTTRER-I
                                                                                                                      518
sp|A7MQQ8|RPOC_CROS8
sp|Q32Ag0|RPOC_SHIDS
sp|Q0SY12|RPOC_SHIF8
sp|B2TWH4|RPOC_SHIB3
                                           FDGDOMAVHVPLTLEAQLEARALMMSTNNILSPANGEPIIVPSQDVVLGLYYMTRDK-I
                                                                                                                      518
                                                   VHVPLTLEAQLEARALMMSTNNILSPANGEPIIVPSQDVVLGLYYMTRDC-
                                                                                                                       518
                                           FDGDQMAVHVPLTLEAQLEARALMMSTNNILSPANGEPIIVPSQDVVLGLYYMTRDC-V
                                                                                                                       518
                                           FDGDQMAVHVPLTLEAQLEARALMMSTNNILSPANGEPIIVPSQDVVLGLYYMTRDC-V
                                                                                                                      518
sp | POAST7 | RPOC_ECOLI
sp | Q3YUZ6 | RPOC_SHISS
sp | B1XBZ0 | RPOC_ECODH
sp | A8A787 | RPOC_ECOHS
                                          <mark>DFDGDQMA</mark>VHVPLTLEAQLEARALMMSTNNILSPANGEPIIVPSQDVVLGLYYMTRDC-V
                                           FDGDQMAVHVPLTLEAQLEARALMMSTNNILSPANGEPIIVPSQDVVLGLYYMTRDC-V
                                                                                                                      518
                                                   VHVPLTLEAQLEARALMMSTNNILSPANGEPIIVPSQDVVLGLYYMTRDC-V
                                                                                                                      518
tr|A0A237JUP3|A0A237JUP3 SHISO
                                                   VHVPLTLEAQLEARALMMSTNNILSPANGEPIIVPSQDVVLGLYYMTRDC-V
                                                                                                                       518
                                                   .VHVPLTLEAQLEARALMMSTNNILSPANGEPIIVPSQDVVLGLYYMTRDC-V
.VHVPLTLEAQLEARALMMSTNNILSPANGEPIIVPSQDVVLGLYYMTRDC-V
tr|A0A0F1RBF2|A0A0F1RBF2_ENTAS
tr|A0A1B3EWG0|A0A1B3EWG0 ENTCL
                                                                                                                      518
tr|A0A0F0XM62|A0A0F0XM62 9ENTR
                                                   VHVPLTLEAQLEARALMMSTNNILSPANGEPIIVPSQDVVLGLYYMTRDC-V
                                                                                                                       518
sp|Q5PK92|RPOC_SALPA
sp|A9MHE9|RPOC_SALAR
tr|A0A232XM43|A0A232XM43_SALMU
                                                   VHVPLTLEAQLEARALMMSTNNILSPANGEPIIVPSQDVVLGLYYMTRDC-V
                                                   .
VHVPLTLEAQLEARALMMSTNNILSPANGEPIIVPSQDVVLGLYYMTRDC-V
.VHVPLTLEAQLEARALMMSTNNILSPANGEPIIVPSQDVVLGLYYMTRDC-V
                                                                                                                      518
                                                                                                                      518
tr|B5RFK0|B5RFK0 SALG2
                                                   VHVPLTLEAQLEARALMMSTNNILSPANGEPIIVPSQDVVLGLYYMTRDC-V
                                                                                                                       518
sp|P0A2R5|RPOC_SALTI
sp|Q57H68|RPOC_SALCH
sp|P0A2R4|RPOC_SALTY
                                           FDGDQMAVHVPLTLEAQLEARALMMSTNNILSPANGEPIIVPSQDVVLGLYYMTRDC-V
FDGDQMAVHVPLTLEAQLEARALMMSTNNILSPANGEPIIVPSQDVVLGLYYMTRDC-V
                                                                                                                       518
                                                                                                                      518
                                                   VHVPLTLEAQLEARALMMSTNNILSPANGEPIIVPSQDVVLGLYYMTRDC-V
                                                                                                                       518
sp | A6TGP1 | RPOC KLEP7
                                           FDGDQMAVHVPLTLEAQLEARALMMSTNNILSPANGEPIIVPSQDVVLGLYYMTRDS-V
                                           FOGDOM AVHVPT.TT.EAGT.EARAT.MMSTNNTT.SPANGEPTTVPSGDVVT.GT.YYMTRDC-V
tr|A0A0J2K6S7|A0A0J2K6S7 9ENTR
                                                                                                                      518
tr|A0A0G3RZQ0|A0A0G3RZQ0 KLEOX
tr|A0A212HDS5|A0A212HDS5_9ENTR
                                                   VHVPLTLEAQLEARALMMSTNNILSPANGEPIIVPSQDVVLGLYYMTRDC-V
                                                                                                                      518
                                                   .VHVPLTLEAQLEARALMMSTNNILSPANGEPIIVPSQDVVLGLYYMTRDC-V
tr|A0A1R0FP41|A0A1R0FP41_CITBR
                                           FDGDQMAVHVPLTLEAQLEARALMMSTNNILSPANGEPIIVPSQDVVLGLYYMTRDC-V
                                                                                                                      518
tr|A0A078LHA5|A0A078LHA5_CITKO
                                          DFDGDOMAVHVPLTLEAOLEARALMMSTNNILSPANGEPIIVPSODVVLGLYYMTRDC-V
                                                                                                                      518
sp|A8AKT8|RPOC CITK8
                                           FDGDQMAVHVPLTLEAQLEARALMMSTNNILSPANGEPIIVPSQDVVLGLYYMTRDC-V
                                                                            :: * ..*
```

```
tr|A0A1U8DYN0|A0A1U8DYN0 ALLSI
                                          REGLIDTAVKTAETGYIQRRLIKSMESVMVKYDATVRNSINQVVQLRYGEDGLAGESVEF
                                                                                                                       911
tr|A0A1L8H4P4|A0A1L8H4P4 XENLA
                                          REGLIDTAVKTAETGYIQRRLIKSMESVMVKYDATVRNSINQVVQLRYGEDGLAGEGVEF
                                                                                                                       901
tr|H9GLG5|H9GLG5_ANOCA
tr|H2R1J6|H2R1J6 PANTR
                                          REGLIDTAVKTAETGYIORRLIKSMESVMVKYDATVRNSINQVVQLRYGEDGLAGESVEF
REGLIDTAVKTAETGYIORRLIKSMESVMVKYDATVRNSINQVVQLRYGEDGLAGESVEF
                                                                                                                       902
                                                                                                                       903
                                          REGLIDTAVKTAETGYIQRRLIKSMESVMVKYDATVRNSINQVVQLRYGEDGLAGESVEF
sp|P08775|RPB1 MOUSE
                                                                                                                       903
tr|G1MCZ1|G1MCZ1 AILME
                                          REGLIDTAVKTAETGYIQRRLIKSMESVMVKYDATVRNSINQVVQLRYGEDGLAGESVEF
                                                                                                                       903
tr|008847|008847_MOUSE
tr|S7PWZ6|S7PWZ6_MYOBR
                                          REGLIDTAVKTAETGYIQRRLIKSMESVMVKYDATVRNSINQVVQLRYGEDGLAGESVEF
                                                                                                                       903
                                          REGLIDTAVKTAETGYIQRRLIKSMESVMVKYDATVRNSINQVVQLRYGEDGLAGESVEF
                                                                                                                       903
                                          REGLIDTAVKTAETGYIORRLIKSMESVMVKYDATVRNSINOVVOLRYGEDGLAGESVEF
tr|D4A5A6|D4A5A6 RAT
                                                                                                                       903
tr|A0A1S3EWL2|A0A1S3EWL2 DIPOR
                                          REGLIDTAVKTAETGYIORRLIKSMESVMVKYDATVRNSINQVVQLRYGEDGLAGESVEF
                                                                                                                       903
sp|P11414|RPB1 CRIGR
                                          REGLI<mark>DTAVKTAETGYI</mark>QRRLIKSMESVMVKYDATVRNSINQVVQLRYGEDGLAGESVEF
                                                                                                                       903
tr|035559|035559 CRIGR
                                          REGLIDTAVKTAETGYIQRRLIKSMESVMVKYDATVRNSINQVVQLRYGEDGLAGESVEF
                                                                                                                       903
tr|A0A2I3M9H2|A0A2I3M9H2_PAPAN
                                          regliptavktaetgyi<mark>o</mark>rrliksmesvmvkydatvrnsinovvolrygedglagesvef
                                                                                                                       902
                                          REGLIDTAVKTAETGYIORRLIKSMESVMVKYDATVRNSINOVVOLRYGEDGLAGESVEF
tr|F7HB40|F7HB40 MACMU
                                                                                                                       903
                                          REGLIDTAVKTAETGYIORRLIKSMESVMVKYDATVRNSINOVVOLRYGEDGLAGESVEF
trlA0A2K6RYW9LA0A2K6RYW9 SATBB
                                                                                                                       903
tr|W5N8Z6|W5N8Z6 LEPOC
                                          REGLIDTAVKTAETGYIORRLIKSMESVMVKYDATVRNSINOVVOLRYGEDGLAGEAVEF
                                                                                                                       902
tr|I3JRW6|I3JRW6 ORENI
                                          REGLIDTAVKTAETGYIQRRLIKSMESVMVKYDGTVRNSINQVVQLRYGEDGLAGENVEF
tr|A0A0R4IMS9|A0A0R4IMS9_DANRE
                                          REGLIDTAVKTAETGYIQRRLIKSMESVMVKYDATVRNSINQVVQLRYGEDGLAGEAVEF
                                                                                                                       900
tr|A0A1A7X327|A0A1A7X327_9TELE
                                          REGLIDTAVKTAETGYIORRLIKSMESVMVKYDATVRNSINOVVOLRYGEDGLAGENVEF
                                                                                                                       901
tr|A0A1A8UKD7|A0A1A8UKD7_NOTFU
tr|A0A1A8ER05|A0A1A8ER05_9TELE
                                          REGLIDTAVKTAETGYIQRRLIKSMESVMVKYDATVRNSINQVVQLRYGEDGLAGENVEF
                                                                                                                       901
                                          REGLIDTAVKTAETGYIORRLIKSMESVMVKYDATVRNSINOVVOLRYGEDGLAGENVEF
                                                                                                                       901
                                          REGLIDTAVKTAETGYIQRRLIKSMESVMVKYDATVRNSINQVVQLRYGEDGLAGENVEF
tr|A0A1A8DQ60|A0A1A8DQ60 9TELE
tr|A0A1A8NSR8|A0A1A8NSR8 9TELE
                                          REGLIDTAVKTAETGYIQRRLIKSMESVMVKYDATVRNSINQVVQLRYGEDGLAGENVEF
                                                                                                                       901
tr|A0A1W4YLM7|A0A1W4YLM7 9TELE
                                          regli<mark>dtavktaetgyio</mark>rrliksmesvmvkydatvrnsinovvolrygedglageavef
                                                                                                                       901
tr|A0A1M8A6L7|A0A1M8A6L7_MALS4
sp|P04050|RPB1_YEAST
                                          REGLIDTAVKTAETGYIQRRLVKALEDVTICYDGTVRNSTNNVIEFAYGEDGIDGAMVER
                                                                                                                       894
                                                               RRLVKALEDIMVHYDNTTRNSLGNVIQFIYGEDGMDAAHIEK
                                                                                                                       880
tr|A0A1B2J8C6|A0A1B2J8C6 PICPA
                                          REGLIDTAVKTAETGYIORRLVKALEDIMVHYDGTTRNSLGDIIQFLYGEDGLDGTQVER
                                                                                                                       881
tr|F2QW17|F2QW17_KOMPC
tr|A3GID7|A3GID7_PICST
                                          reglidtavktaetgyiqrrlvkaledimvhydgttrnslgdiiqflygedgldgtqver
                                          REGLIDTAVKTAETGYIQRRLVKALEDIMVHYDGTTRNSLGDIIQFVYGEDGIDGTQVEK
                                                                                                                       880
                                          REGLIDTAVKTAETGYIDRRLVKALEDIMVHYDGTTRNSLGDIIOFIYGEDGIDGTOVEK
tr|A0A1D8PUA6|A0A1D8PUA6 CANAL
                                                                                                                       880
tr|G8BEH9|G8BEH9_CANPC
                                          REGLIDTAVKTAETGYIORRLVKALEDIMVHYDGTTRNSLGDIIOFIYGEDGIDATOVEK
                                                                                                                       880
                                                              TRKLVDVTHEIVVREADCGTTNYISV-PLFQPDEVT------
TRRLVDVSQDCVIVQEDCGTENALEMRAIVQGGSVI------
AEG34223.1
                                          RKGG
                                                                                                                       1129
ASR51305.1
OXR47930.1
                                                 FALKTANSGYLTRRLVDVTQDLVITETDCGTTSGYTMKALVEGGEVI-----
                                          RKGL.
sp|A7MQQ8|RPOC_CROS8
sp|Q32AG0|RPOC_SHIDS
sp|Q0SY12|RPOC_SHIF8
                                          REGIA
                                                              RRLVDVAQDLVVTEDDCGTLEGITMTPVIEGGDVK-----
                                                                                                                       832
                                                              TRRLVDVAQDLVVTEDDCGTHEGIMMTPVIEGGDVK-----
TRRLVDVAQDLVVTEDDCGTHEGIMMTPVIEGGDVK-----
                                          RKGT.
                                                                                                                       832
                                          RKGL
                                                                                                                       832
sp | B2TWH4 | RPOC SHIB3
                                               DTALKTANSGYLTRRLVDVAQDLVVTEDDCGTHEGIMMTPVIEGGDVK-----
                                                                                                                       832
                                          RKGLADTALKTANSGYLTRRLVDVAQDLVVTEDDCGTHEGIMMTFVIEGGDVK------
RKGLADTALKTANSGYLTRRLVDVAQDLVVTEDDCGTHEGIMMTFVIEGGDVK------
RKGLADTALKTANSGYLTRRLVDVAQDLVVTEDDCGTHEGIMMTFVIEGGDVK------
RKGLADTALKTANSGYLTRRLVDVAQDLVVTEDDCGTHEGIMMTFVIEGGDVK------
sp|Q3YUZ6|RPOC_SHISS
sp|B1XBZ0|RPOC_ECODH
sp|A8A787|RPOC_ECOHS
                                                                                                                       832
                                                                                                                       832
tr|A0A237JUP3|A0A237JUP3 SHISO
                                                               RRLVDVAQDLVVTEDDCGTHEGIMMTPVIEGGDVK-----
                                          RKGL.
                                                                                                                       832
tr|A0A0F1RBF2|A0A0F1RBF2_ENTAS
tr|A0A1B3EWG0|A0A1B3EWG0_ENTCL
                                          RKGL
                                                              RRLVDVAQDLVVTEDDCGTLEGITMTPVIEGGDVK-----
                                                                                                                       832
                                          RKGL
                                                               RRLVDVAQDLVVTEDDCGTLEGITMTPVIEGGDVK-----
                                                                                                                       832
tr|A0A0F0XM62|A0A0F0XM62 9ENTR
                                          RKGL
                                                 TALKTANSGY.
                                                              TRRLVDVAQDLVVTEDDCGTLEGITMTPVIEGGDVK-----
                                                                                                                       832
sp|Q5FK92|RPOC_SALPA
sp|A9MHE9|RPOC_SALAR
tr|A0A232XM43|A0A232XM43_SALMU
                                          RKGLADTALKTANSGY
                                                              RRLVDVAODLVVTEDDCGTHEGILMTPVIEGGDVK-----
                                                                                                                       832
                                                DTALKTANSGY
                                          RKGL
                                                               RRLVDVAODLVVTEDDCGTHEGILMTPVIEGGDVK-----
                                                                                                                       832
                                          RKGLADTALKTANSGY
                                                              RRLVDVAQDLVVTEDDCGTHEGILMTPVIEGGDVK-----
                                                                                                                       832
tr|B5RFK0|B5RFK0 SALG2
                                                               RRLVDVAQDLVVTEDDCGTHEGILMTPVIEGGDVK-----
sp|P0A2R5|RPOC_SALTI
                                          RKGLA
                                                DTALKTANSGY:
                                                              TRRLVDVAQDLVVTEDDCGTHEGILMTPVIEGGDVK-----
sp|Q57H68|RPOC_SALCH
sp|P0A2R4|RPOC_SALTY
sp|A6TGP1|RPOC_KLEP7
                                          RKGT.A
                                                              FRRLVDVAODLVVTEDDCGTHEGILMTPVIEGGDVK-----
                                                                                                                       832
                                          RKGT.Z
                                                 TALKTANSGY.
                                                              RRLVDVAQDLVVTEDDCGTHEGILMTPVIEGGDVK-----
                                                                                                                       832
                                                              TRRLVDVAQDLVVTEDDCGTLEGITMTPVIEGGDVK-----
                                          RKGLADTALKTANSGY
                                                                                                                       832
tr|A0A0J2K6S7|A0A0J2K6S7 9ENTR
                                          RKGL
                                                              RRLVDVAQDLVVTEDDCGTLEGITMTPVIEGGDVK-----
                                                                                                                       832
tr|A0A0G3RZQ0|A0A0G3RZQ0 KLEOX
tr|A0A212HDS5|A0A212HDS5 9ENTR
tr|A0A1R0FP41|A0A1R0FP41_CITBR
                                                              RRLVDVAQDLVVTEDDCGTLEGITMTPVIEGGDVK-----
                                          RKGL
                                          RKGLA
                                                              RRLVDVAQDLVVTEDDCGTLEGITMTPVIEGGDVK-----
                                                               RRLVDVAODLVVTEDDCGTLEGITMTPVIEGGDVK-----
                                          REGIA
                                                                                                                       832
                                                              TRRLVDVAQDLVVTEDDCGTHEGILMTPVIEGGDVK-----
tr|A0A078LHA5|A0A078LHA5_CITKO
                                          RKGLADTALKTANSGY
                                                                                                                       832
sp|A8AKT8|RPOC CITK8
                                          RKGL
                                                DTALKTANSGYLTRRLVDVAQDLVVTEDDCGTHEGILMTPVIEGGDVK-----
                                                                                                                       832
                                                              *:*:. .. :
                                                                                          . .
                                         VNGDDPLSRQAQENATLLFNIHLRSTLCARRMIEEFRLSAQAFDWLLGEIESKFNQAIAH
tr|A0A1U8DYN0|A0A1U8DYN0 ALLSI
                                                                                                                       1090
tr|A0A1L8H4P4|A0A1L8H4P4 XENLA
                                         VNGDDPLSRQAQENATLLFNIHLRSTLSSRRMIEEFRLSGEAFDWLLGEIESKFNQAIAH
                                                                                                                       1080
tr|H9GLG5|H9GLG5_ANOCA
tr|H2R1J6|H2R1J6_PANTR
                                          VNGDDPLSKQAQENATLLFNIHLRSTLCSRRMIEEFRLSGEAFDWLLGEIESKFNQAIAH
                                                                                                                       1081
                                         VNGDDPLSRQAQENATLLFNIHLRSTLCSRRMAEEFRLSGEAFDWLLGEIESKFNQAIAH
                                                                                                                       1082
sp|P08775|RPB1 MOUSE
                                         VNGDDPLSROAOENATLLFNIHLRSTLCSRRMAEEFRLSGEAFDWLLGEIESKFNOAIAH
                                                                                                                       1082
tr|G1MCZ1|G1MCZ1 AILME
                                         VNGDDPLSROAOENATI.I.FNTHI.RSTI.CSRRMAEFFRI.SGEAFDWI.I.GETESKFNOATAH
                                                                                                                       1082
tr|008847|008847 MOUSE
                                         VNGDDPLSRQAQENATLLFNIHLRSTLCSRRMAEEFRLSGEAFDWLLGEIESKFNQAIAH
                                                                                                                       1082
                                          VNGDDPLSRQAQENATLLFNIHLRSTLCSRRMAEEFRLSGEAFDWLLGEIESKFNQAIAH
tr|S7PWZ6|S7PWZ6 MYOBR
tr|D4A5A6|D4A5A6 RAT
                                         VNGDDPLSRQAQENATLLFNIHLRSTLCSRRMAEEFRLSGEAFDWLLGEIESKFNQAIAH
                                                                                                                       1082
tr|A0A1S3EWL2|A0A1S3EWL2_DIPOR
                                         VNGDDPLSRQAQENATLLFNIHLRSTLCSRRMAEEFRLSGEAFDWLLGEIESKFNQAIAH
                                                                                                                       1082
sp|P11414|RPB1_CRIGR
tr|035559|035559 CRIGR
                                         VNGDDPLSROAOENATLLFNIHLRSTLCSRRMAEEFRLSGEAFDWLLGEIESKFNOAIAH
                                                                                                                       1082
                                          VNGDDPLSRQAQENATLLFNIHLRSTLCSRRMAEEFRLSGEAFDWLLGEIESKFNQAIAH
                                                                                                                       1082
                                         VNGDDPLSRQAQENATLLFNIHLRSTLCSRRMAEEFRLSGEAFDWLLGEIESKFNQAIAH
tr|A0A2I3M9H2|A0A2I3M9H2 PAPAN
                                                                                                                       1081
tr|F7HB40|F7HB40 MACMU
                                         VNGDDPLSRQAQENATLLFNIHLRSTLCSRRMAEEFRLSGEAFDWLLGEIESKFNQAIAH
                                                                                                                       1082
tr|A0A2K6RYW9|A0A2K6RYW9 SAIBB
                                         VNGDDPLSRQAQENATLLFNIHLRSTLCSRRMAEEFRLSGEAFDWLLGEIESKFNQAIAH
                                                                                                                       1082
tr|W5N8Z6|W5N8Z6_LEPOC
tr|I3JRW6|I3JRW6 ORENI
                                         VNGDDPLSRQAQENATLLFNIHLRSTLCSRRMTEEFRLSTEAYEWLLGEIETKFNQSIAH
                                                                                                                       1081
                                         VNGDDPLSRQAQQNATLLFNIHLRSTLCSKRMTEEFRLSTEAFDWLLGEIETKFNQSIAH
                                                                                                                       1080
tr|A0A0R4IMS9|A0A0R4IMS9 DANRE
                                          VNGDDQLSRQAQENATLLFNIHLRSTLCSKRMTEEFRLSTEAFDWLLGEIETKFNQSIAH
                                                                                                                       1079
                                         VNGDDPLSRQAQENATLLFNIHLRSTLCSKRMTEEFRLSTEAFDWLLGEIETKFNQSIVH
tr|A0A1A7X327|A0A1A7X327_9TELE
                                                                                                                       1080
tr|A0A1A8UKD7|A0A1A8UKD7_NOTFU
tr|A0A1A8ER05|A0A1A8ER05_9TELE
                                         VNGDDPLSRQAQENATLLFNIHLRSTLCSRRMTEEFRLSMEAFDWLLGEIETKFNQSIVH
                                                                                                                       1080
                                         VNGDDPLSROAOENATLLFNIHLRSTLCSRRMTEEFRLSMEAFDWLLGEIETKFNOSIVH
                                                                                                                       1080
```

```
tr|A0A1A8DQ60|A0A1A8DQ60_9TELE
                                              VNGDDPLSROAGENATILENTHLRSTLCSRRMTEEFRLSMEAFDWLLGETETKENGSTVH
                                                                                                                                   1080
tr|A0A1A8NSR8|A0A1A8NSR8_9TELE
tr|A0A1W4YLM7|A0A1W4YLM7_9TELE
                                              VNGDDPLSRQAQENATLLFNIHLRSTLCSRRMTEEFRLSMEAFDWLLGEIETKFNQSIVH
                                                                                                                                   1080
                                              VNGDDPLSRQAQENATLLFNIHLRSTLCSRRMTEEFRLSTEAFEWLLGEIETKFNQAIVH
                                                                                                                                   1080
tr|A0A1M8A6L7|A0A1M8A6L7_MALS4
                                              \tt IRGNDPISRSMQENATLLFK \cite{thm} RSFLCTKQVIEVHHLSREAWEWILGEIEGQFARSVAQ
                                                                                                                                   1071
tr|A0A1B2J8C6|A0A1B2J8C6 PICPA
                                              LRGENELIKEAQQNATSLFQC<mark>L</mark>VRARLATRRILEEFRLNRDAFEWVLGTIEAQFQRSLVH
                                                                                                                                    1061
tr|F20W17|F20W17_KOMPC
tr|A3GID7|A3GID7_PICST
tr|A0A1D8PUA6|A0A1D8PUA6_CANAL
                                              LRGENELIKEAOONATSLFOO
                                                                          VRARIATRRILEEFRINRDAFEWVLGTIEAOFORSIVH
                                                                                                                                   1061
                                              VRGDTELVKEAOANATLLFOO
                                                                          VRSRLASRRVIEEFKLNRSSFEWVVGEIETOFOKSIVH
                                                                                                                                    1060
                                              VRGDTPLVKEAQENATLLFQ
                                                                          LRSRLAARRVIEEFKLNRASFEWVMGEIETQFQKSIVH
                                                                                                                                    1060
tr|G8BEH9|G8BEH9 CANPC
                                              VRGKTKI AKEAOENATLI FO
                                                                          VRSRIAARRVIEEFKINRSSEDWVLGEIETOFOKSIVH
                                                                                                                                   1060
AEG34223.1
                                              -GE IO----EVP
                                                                                                       DLS----MARPVS
                                                                                                                                   1216
OXR47930.1
                                              -LGVD-
                                                      ----EVK
                                                                                                       RDLG-----RGSLVN
                                                                                                                                   910
sp|A7MQQ8|RPOC_CROS8
sp|Q32AG0|RPOC_SHIDS
sp|Q0SY12|RPOC_SHIF8
sp|B2TWH4|RPOC_SHIB3
                                              -NSVD----SVK
                                                                                                       RDLA----RGHIIN
                                                                                                                                   910
                                                                                                                                    910
                                              -NSVD----
                                                                                                       RDLA----RGHIIN
                                                           -AVK
                                                                                                                                   910
                                              -NSVD----
                                                           -AVK
                                                                                                      RDI.A----RGHIIN
                                                                                                                                   910
sp|Q3YUZ6|RPOC_SHISS
sp|B1XBZ0|RPOC_ECODH
sp|A8A787|RPOC_ECOHS
                                                           -AVKVI
                                                                                                                  RGHIIN
                                               -NSVD
                                                                                                                                    910
                                              -NSVD----AVK
                                                                                                       DLA----RGHIIN
                                                                                                                                   910
                                              -NSVD-
                                                                                                                   RGHIIN
                                                                                                                                   910
                                                                                                       DT.A--
tr|A0A237JUP3|A0A237JUP3_SHISO
                                              -NSVD-
                                                           -AVK
                                                                                                       DT.A----RGHTTN
                                                                                                                                   910
tr|A0A0F1RBF2|A0A0F1RBF2_ENTAS
tr|A0A1B3EWG0|A0A1B3EWG0_ENTCL
                                              -NSVD--
                                                           -SVK
                                                                                                       DLA----RGHIIN
                                                                                                                                   910
                                              -NSVD-
                                                                                                                   RGHIIN
                                                                                                                                   910
tr|A0A0F0XM62|A0A0F0XM62_9ENTR
                                              -NSVD-
                                                                                                       RDLA----RGHIIN
sp|Q5PK92|RPOC_SALPA
sp|A9MHE9|RPOC_SALAR
                                              -NSVD--
                                                           -AVK
                                                                                                       RDLA----RGHIIN
                                                                                                                                   910
                                              -NSVD-
                                                                                                                   RGHIIN
                                                                                                       RDLA----
                                                                                                                                   910
tr|A0A232XM43|A0A232XM43 SALMU
                                                                                                       DLA----RGHIIN
                                              -NSVD-
                                                            -AVK
                                                                                                                                    910
tr|B5RFK0|B5RFK0_SALG2
                                              -NSVD-
                                                           -217K
                                                                                                       PDT.A----RGHITN
                                                                                                                                   910
sp|P0A2R5|RPOC_SALTI
sp|Q57H68|RPOC_SALCH
                                              -NSVD-
                                                                                                       RDLA-----RGHIIN
                                                           -AVK
                                                                                                                                   910
                                              -NSVD--
                                                                                                       RDLA----RGHIIN
                                                                                                                                    910
sp|P0A2R4|RPOC_SALTY
sp|A6TGP1|RPOC_KLEP7
                                              -NSVD-
                                                           _277K
                                                                                                       PDT. A ---- RGH T TN
                                                                                                                                   910
                                              -NSVD-
                                                                                                       RDLA-----RGHLIN
                                                           -SVK
                                                                                                                                   910
tr|A0A0J2K6S7|A0A0J2K6S7_9ENTR
tr|A0A0G3RZQ0|A0A0G3RZQ0_KLEOX
                                                                                                                                   910
                                              -NSVD-
                                                           -SVR
                                                                                                       PDT. A ---- PGH T TN
                                                                                                                                   910
tr|A0A212HDS5|A0A212HDS5_9ENTR
tr|A0A1R0FP41|A0A1R0FP41_CITBR
                                              -NSVD-
                                                      ----SVK
                                                                                                       RDLA-----RGHIIN
                                                                                                                                   910
                                              -NSVD
                                                                                                       RDLA-----RGHIIN
                                                                                                                                   910
tr|A0A078LHA5|A0A078LHA5_CITKO
sp|A8AKT8|RPOC_CITK8
                                                                                                       RDLA----RGHIIN
                                                                                                      RDLA-----RGHIIN
                                              -NSVD----AVK
                                                                                                                                   910x16
                                             tr|A0A1U8DYN0|A0A1U8DYN0 ALLSI
                                                                                                                                    1142
tr|AOA1L8H4P4|AOA1L8H4P4_XENLA
                                                                                                                                    1132
tr|H9GLG5|H9GLG5_ANOCA
tr|H2R1J6|H2R1J6_PANTR
                                                                                                                                    1133
                                                                                                                                    1134
sp|P08775|RPB1 MOUSE
                                                                                                                                    1134
tr|G1MCZ1|G1MCZ1_AILME
tr|O08847|O08847_MOUSE
                                                                                                                                    1134
                                                                                                                                    1134
tr|S7PWZ6|S7PWZ6 MYOBR
                                                                                                                                    1134
tr|D4A5A6|D4A5A6_RAT
tr|A0A1S3EWL2|A0A1S3EWL2_DIPOR
                                                                                                                                    1134
                                                                                                                                    1134
sp|P11414|RPB1 CRIGR
                                                                                                                                    1134
tr|035559|035559_CRIGR
tr|A0A2I3M9H2|A0A2I3M9H2_PAPAN
                                                                                                                                    1134
                                                                                                                                    1133
tr|F7HB40|F7HB40 MACMU
                                                                                                                                    1134
tr|A0A2K6RYW9|A0A2K6RYW9_SAIBB
                                                                                                                                    1134
tr|W5N8Z6|W5N8Z6_LEPOC
                                                                                                                                    1133
tr||I3JRW6||I3JRW6 ORENI
                                                                                                                                    1132
tr|A0A0R4IMS9|A0A0R4IMS9_DANRE
                                                                                                                                    1131
tr|A0A1A7X327|A0A1A7X327_9TELE
tr|A0A1A8UKD7|A0A1A8UKD7 NOTFU
                                                                                                                                    1132
                                                                                                                                    1132
tr|A0A1A8ER05|A0A1A8ER05_9TELE
                                                                                                                                    1132
tr|A0A1A8DQ60|A0A1A8DQ60_9TELE
                                                                                                                                    1132
tr|AOA1A8NSR8|AOA1A8NSR8_9TELE
                                                                                                                                    1132
tr|A0A1W4YLM7|A0A1W4YLM7_9TELE
                                                                                                                                     1132
tr|A0A1M8A6L7|A0A1M8A6L7_MALS4
                                                                                                                                    1123
sp|P04050|RPB1 YEAST
                                              PGEMVGVI AAQSIGE PATQMTIN TEH TAGVSSKNVTIGVPR------LKEILNVAKNI
PGEMVGVI AAQSIGE PATQMTIN TEH TAGVSSKNVTIGVPR------LKEILNVAKNI
PGEMVGVI AAQSIGE PATQMTIN TEH TAGVSSKNVTIGVPR------LKEILNVAKNI
tr|A0A1B2J8C6|A0A1B2J8C6_PICPA
tr|F2QW17|F2QW17_KOMPC
tr|A3GID7|A3GID7 PICST
                                                                                                                                    1113
                                                                                                                                    1112
tr|A0A1D8PUA6|A0A1D8PUA6_CANAL
                                              PGEMVGVVAAQSIGE PATQMTINTFH AGVSSKNVTLGVPR-----LKEILNVAKNI
PGEMVGVWAAOSIGE PATOMTINTFH AGVSSKNVTLGVPR-----LKEILNVAKNI
tr|G8BEH9|G8BEH9 CANPC
                                                                                                                                    1112
AEG34223.1
                                               IGEAVG
                                                                                                                                    1248
                                                                                   GAAQLNE-TSNLEAVADGTLQYRDI-PTIVNK
                                                                                   GAASRAAMASSVETKAAGTVSFGVSMRYVTNA
OXR47930.1
                                              RGEAVG
                                                                                                                                    970
sp|A7MQQ8|RPOC_CROS8
sp|Q32AG0|RPOC_SHIDS
sp|Q0SY12|RPOC_SHIF8
                                              KGEAIG
                                                                                  GAASRAAAESSIQVKNKGSIRLSNA-KSVVNS
                                                                                                                                    969
                                                                                  GAASRAAAESSIQVKNKGSIKLSNV-KSVVNS
                                              KGEAIG
                                                                                  GAASRAAAESSIOVKNKGSIKLSNV-KSVVNS
                                                                                                                                    969
sp|B2TWH4|RPOC SHIB3
                                                                                  GGAASRAAAESSIQVKNKGSIKLSNV-KSVVNS
                                              KGEAIG
                                                                                                                                     969
sp|POA8T7|RPOC_ECOLI
sp|Q3YUZ6|RPOC_SHISS
                                                                                  <mark>GGAASRAAAESSIQVKNKGSIKLSNV-KSVVNS</mark>
GGAASRAAAESSIQVKNKGSIKLSNV-KSVVNS
                                              KGEAIGVIAAQSIGEPGTQLTMR
                                                                                                                                     969
sp | B1XBZ0 | RPOC ECODH
                                              KGEAIGVIAAQSIGEPGTQLT
                                                                                  GAASRAAAESSIQVKNKGSIKLSNV-KSVVNS
                                                                                                                                    969
sp|A8A787|RPOC_ECOHS
tr|A0A237JUP3|A0A237JUP3 SHISO
                                                                                   GAASRAAAESSIQVKNKGSIKLSNV-KSVVNS
                                                                                     AASRAAAESSIQVKNKGSIKLSNV-KSVVNS
                                              KGEATG
                                                                                                                                    969
```

```
GAASRAAAESSIQVKNKGSIKLSNA-KSVVNS
tr|A0A0F1RBF2|A0A0F1RBF2 ENTAS
                                            KGEAIG
                                                     TAAQSIGEPGTQLT
                                                                                                                               969
tr|A0A1B3EWG0|A0A1B3EWG0 ENTCL
                                            KGEAIG
                                                     IAAQSIGEPGTQLT
                                                                                GAASRAAAESSIQVKNKGSIKLSNA-KSVVNS
                                                                                                                               969
tr|A0A0F0XM62|A0A0F0XM62_9ENTR
                                             KGEAIG
                                                                               GAASRAAAESSIQVKNKGSIKLSNA-KSVVNS
                                                                                                                               969
sp|Q5PK92|RPOC_SALPA
                                             KGEAI
                                                                                GAASRAAAESSIQVKNKGSIKLSNV-KSVVNS
                                                                                                                               969
sp|A9MHE9|RPOC_SALAR
tr|A0A232XM43|A0A232XM43_SALMU
                                            KGEAIG
                                                                                GAASRAAAESSIQVKNKGSIKLSNV-KSVVNS
                                                                                                                               969
                                                                                GAASRAAAESSIOVKNKGSIKLSNV-KSVVNS
                                            KGEAIG
                                                                                                                               969
tr|B5RFK0|B5RFK0 SALG2
                                            KGEAIG
                                                                                GAASRAAAESSIQVKNKGSIKLSNV-KSVVNS
                                                                                                                               969
                                                     IAAQSIGEPGTQ
sp|P0A2R5|RPOC_SALTI
sp|Q57H68|RPOC_SALCH
                                                     /IAAQSIGEPGTQI
                                            KGEAIGV
                                                                               GAASRAAAESSIQVKNKGSIKLSNV-KSVVNS
                                                                                                                               969
                                                                                GAASRAAAESSIQVKNKGSIKLSNV-KSVVNS
                                                                                                                               969
                                            KGEAIGVIAAQSIGEPGTQ
sp | P0A2R4 | RPOC_SALTY
                                                                               GAASRAAAESSIQVKNKGSIKLSNV-KSVVNS
                                                                                                                               969
sp|A6TGP1|RPOC_KLEP7
tr|A0A0J2K6S7|A0A0J2K6S7_9ENTR
tr|A0A0G3RZQ0|A0A0G3RZQ0_KLEOX
                                            KGEAIGVIAAQSIGEPGT
KGEAIGVIAAQSIGEPGT
                                                                               GAASRAAAESSIQVKNKGSIKLSNA-KSVVNS
                                                                                                                               969
                                                                                GAASRAAAESSIQVKNKGSIRLSNA-KSVVNS
                                                                                                                               969
                                            KGEAIGVIAAOSIGEPGTOLT
                                                                               GAASRAAAESSIOVKNKGSIRLSNA-KSVVNS
                                                                                                                               969
tr|A0A212HDS5|A0A212HDS5_9ENTR
tr|A0A1R0FP41|A0A1R0FP41_CITBR
                                                                               GAASRAAAESSIOVKNKGSIRLSNA-KSVVNS
                                            KGEAIG
                                                     TAAOSIGEPGTO1
                                                                                                                               969
                                                                                GAASRAAAESSIQVKNKGSIRLSNA-KSVVNS
                                            KGEAIG
                                                                                                                               969
tr|A0A078LHA5|A0A078LHA5_CITKO
                                             KGEAIGVIAAQSIGEPGTQLT
                                                                               GAASRAAAESSIQVKNKGSIRLSNA-KSVVNS
sp|A8AKT8|RPOC_CITK8
                                                                               GAASRAAAESSIQVKNKGSIRLSNA-KSVVNS
                                            KGEAIG
                                                                          FH
                                                                                                                               969
tr|A0A1U8DYN0|A0A1U8DYN0 ALLSI
                                            --WSPS---VGSGMTPGAAGFSPSAASDA-SGFSPGYSPAWSPTPGSPGSPGPSSPY--WSPS---IGSGMTPGAAGFSPSAASDA-SGFSPGYSPAWSPTPGSPGSPGPSSPY-
                                                                                                                              1589
tr|A0A1L8H4P4|A0A1L8H4P4 XENLA
                                                                                                                              1579
tr | H9GLG5 | H9GLG5_ANOCA
                                            --WSPS---VGSGMTPGAAGFSPSAASDA-
                                                                                 SGLSPGYSPAWSPTPGSPGSPGPSSPY
                                                                                                                              1580
                                                                   -----LTYA-SGFSPGYSPAWSPTPGSPGSPGPSSPY
tr|H2R1J6|H2R1J6 PANTR
                                                                                                                              1571
sp|P08775|RPB1_MOUSE
                                            --WSPSV---GSGMTPGAAGFSPSAASDA-SGFSPGYSPAWSPTPGSPGSPGPSSPY
                                                                                                                              1581
                                            --WSPSVDITGSGMTPGAAGFSPSAASDA-SGFSPGYSPAWSPTPGSPGSPGPSSPY--WSPSV---GSGMTPGAAGFSPSAASDA-SGFSPGYSPAWSPTPGSPGSPGPSSPY--WSPSV---GSGMTPGAAGFSPSAASDA-SGFSPGYSPAWSPTPGSPGSPGPSSPY-
tr|G1MCZ1|G1MCZ1_AILME
                                                                                                                              1584
tr|008847|008847 MOUSE
                                                                                                                              1581
tr|S7PWZ6|S7PWZ6 MYOBR
                                                                                                                              1581
tr|D4A5A6|D4A5A6_RAT
                                            --WSPSV---GSGMTPGAAGFSPSAASDA-SGFSPGYSPAWSPTPGSPGSPGPSSPY
                                                                                                                              1581
                                            --WSPSV---GSGMTPGAAGFSPSAASDA-SGFSPGYSPAWSPTPGSPGSPGPSSPY--WSPSV---GSGMTPGAAGFSPSAASDA-SGFSPGYSPAWSPTPGSPGSPGPSSPY-
tr|A0A1S3EWL2|A0A1S3EWL2_DIPOR
                                                                                                                              1581
sp|P11414|RPB1_CRIGR
                                                                                                                              1581
                                            --WSPSV---GSGMTPGAAGFSPSAASDA-GGFSPGYSPAWSPTPGSPGSPGPSSPY
tr|035559|035559 CRIGR
                                                                                                                              1581
                                            -----GMT PGAAGFS PSAAS DA-SGF SPGY SPAWS PT PGS PGS PGPS SPY-----GMT PGAAGFS PSAAS DA-SGF SPGY SPAWS PT PGS PGS PGPS SPY-
tr|A0A2I3M9H2|A0A2I3M9H2_PAPAN
                                                                                                                              1542
tr|F7HB40|F7HB40 MACMU
                                                                                                                              1543
tr|A0A2K6RYW9|A0A2K6RYW9 SAIBB
                                            -----GMTPGAAGFSPSAASDA-SGFSPGYSPAWSPTPGSPGSPGPSSPY
                                                                                                                              1543
                                            --WSPS---VGSGMTPGAAGFSPSAASDA-SGFSPGYSPAWSPTPGSPGSPGPASPY--WSPS---VGSGMTPGGAGFSPSAASDA-SGFSPGYSPAWSPTPGSPGSPGPASPY
tr|W5N8Z6|W5N8Z6_LEPOC
                                                                                                                              1580
tr|I3JRW6|I3JRW6_ORENI
                                                                                                                              1579
                                            --WSPS---VGSGMTPGAAGFSPSAASDA-SGFSPGYSPAWSPTPGSPGSPGPASPY-
--WSPS---VGSGMTPGAAGFSPSAASDA-SGFSPGYSPAWSPTPGSPGSPGPASPY-
--WSPS---VGSGMTPGAAGFSPSAASDA-SGFSPGYSPAWSPTPGSPGSPGPASPY-
tr|A0A0R4IMS9|A0A0R4IMS9_DANRE
                                                                                                                              1578
tr|A0A1A7X327|A0A1A7X327_9TELE
tr|A0A1A8UKD7|A0A1A8UKD7_NOTFU
                                                                                                                              1579
                                                                                                                              1579
                                            --WSPS---VGSGMTPGAAGFSPSAASDA-SGFSPGYSPAWSPTPGSPGSPGPASPY
tr|A0A1A8ER05|A0A1A8ER05 9TELE
                                                                                                                              1579
                                            --WSPS---VGSGMTPGAAGFSPSAASDA-SGFSPGYSPAWSPTPGSPGSPGPASPY--WSPS--VGSGMTPGAAGFSPSAASDA-SGFSPGYSPAWSPTPGSPGSPGPASPY
tr|A0A1A8DQ60|A0A1A8DQ60_9TELE
                                                                                                                              1579
tr|A0A1A8NSR8|A0A1A8NSR8_9TELE
                                                                                                                              1579
tr|AOA1W4YLM7|AOA1W4YLM7 9TELE
                                            --WSPS---VGSGMTPGAAGFSPSAASDA-SGFSPGYSPAWSPTPGSPGSPGPASPY-
                                                                                                                              1579
                                            tr|A0A1M8A6L7|A0A1M8A6L7 MALS4
                                                                                                                              1542
sp|P04050|RPB1 YEAST
                                                                                                                              1523
tr|A0A1B2J8C6|A0A1B2J8C6 PICPA
                                                                                                                              1527
tr|F2QW17|F2QW17_KOMPC
tr|A3GID7|A3GID7_PICST
                                            ---DEF----NHDDVADVMFSPMAETGSGDDRSGGL---TEYAGIQSPYQP-----
                                                                                                                              1527
                                            ---DD-----KIQFEEGAGFSPIHTAQVQ-DVSGGL----TSYGGQPTSPSATSPFSYG
                                                                                                                              1529
                                            ---DE-----NIDIDAGAGFSPIHIAQMNEGNIGGL----TSYGGQPTSPAATSPFSYG
tr|A0A1D8PUA6|A0A1D8PUA6_CANAL
                                                                                                                              1527
                                            ---DD-----RIQVDESAGFSPVHQAPTAEGMNGGL----TSYGGQPTSPSATSPFSYS
tr|G8BEH9|G8BEH9_CANPC
                                                                                                                              1526
AEG34223.1
                                           GKQA-----
                                                                                                                             1524
ASR51305.1
                                                                                                                             1403
OXR47930.1
                                                                                                                             1416
sp|A7MQQ8|RPOC_CROS8
                                                                                                                             1407
sp|Q32AG0|RPOC_SHIDS
sp|Q0SY12|RPOC_SHIF8
                                                                                                                             1407
                                                                                                                             1407
sp|B2TWH4|RPOC_SHIB3
                                                                                                                             1407
sp|Q3YUZ6|RPOC_SHISS
sp|B1XBZ0|RPOC_ECODH
                                                                                                                             1407
                                                                                                                             1407
sp | A8A787 | RPOC_ECOHS
                                                                                                                             1407
tr|A0A237JUP3|A0A237JUP3_SHISO
                                                                                                                             1407
tr|A0A0F1RBF2|A0A0F1RBF2_ENTAS
                                                                                                                             1407
tr|A0A1B3EWG0|A0A1B3EWG0_ENTCL
                                                                                                                             1407
tr|A0A0F0XM62|A0A0F0XM62_9ENTR
                                                                                                                             1407
sp|Q5PK92|RPOC_SALPA
sp|A9MHE9|RPOC_SALAR
                                                                                                                             1407
                                                                                                                             1407
tr|A0A232XM43|A0A232XM43_SALMU
                                                                                                                             1407
tr|B5RFK0|B5RFK0 SALG2
                                                                                                                             1407
sp|P0A2R5|RPOC_SALTI
                                                                                                                             1407
sp|Q57H68|RPOC_SALCH
sp|P0A2R4|RPOC_SALTY
sp|A6TGP1|RPOC_KLEP7
                                                                                                                             1407
                                                                                                                             1407
                                                                                                                             1407
tr|A0A0J2K6S7|A0A0J2K6S7 9ENTR
                                                                                                                             1407
tr|A0A0G3RZQ0|A0A0G3RZQ0_KLEOX
tr|A0A212HDS5|A0A212HDS5_9ENTR
                                                                                                                             1407
                                                                                                                             1407
tr|A0A1R0FP41|A0A1R0FP41_CITBR
                                                                                                                             1407
tr|A0A078LHA5|A0A078LHA5 CITKO
                                                                                                                             1407
sp|A8AKT8|RPOC CITK8
                                                                                                                             1407
```

tr A0A1U8DYN0 A0A1U8DYN0 ALLSI		1926
tr A0A1L8H4P4 A0A1L8H4P4 XENLA	PDOSDEDN	1968
tr H9GLG5 H9GLG5 ANOCA	PDDSDEEN	1969
tr H2R1J6 H2R1J6 PANTR	PDDSDEEN	1960
splP08775 RPB1 MOUSE	PDDSDEEN	1970
tr G1MCZ1 G1MCZ1 AILME	PDOSDEEN	1973
tr 008847 008847 MOUSE	DEEN	1966
tr S7PWZ6 S7PWZ6 MYOBR	PDOSDEEN	1970
tr D4A5A6 D4A5A6 RAT	PEDSDEEN	1970
tr A0A1S3EWL2 A0A1S3EWL2 DIPOR	PDDSDEEN	1970
sp P11414 RPB1 CRIGR	PDDSDEEN	1970
tr 035559 035559 CRIGR	PDOSDEEN	1970
tr A0A2I3M9H2 A0A2I3M9H2 PAPAN	PDOSDEEN	1931
tr F7HB40 F7HB40 MACMU	PDOSDEEN	1932
tr A0A2K6RYW9 A0A2K6RYW9 SAIBB	PDDSDEEN	1932
tr W5N8Z6 W5N8Z6 LEPOC	PDDSDEDN	1959
tr I3JRW6 I3JRW6 ORENI	PDDSDEENN	1966
tr A0A0R4IMS9 A0A0R4IMS9_DANRE	PDDSDEENN	1965
tr A0A1A7X327 A0A1A7X327 9TELE	PDDSDE SDEENN	1969
tr A0A1A8UKD7 A0A1A8UKD7 NOTFU	PDDSDE SDEENN	1969
tr A0A1A8ER05 A0A1A8ER05_9TELE	PDDSDE SDEENN	1969
tr A0A1A8DQ60 A0A1A8DQ60 9TELE	PDDSDE SDEENN	1962
tr A0A1A8NSR8 A0A1A8NSR8 9TELE	PDDSDE SDEENN	1962
tr A0A1W4YLM7 A0A1W4YLM7_9TELE	PDDSDDDN	1968
tr A0A1M8A6L7 A0A1M8A6L7 MALS4		1803
sp P04050 RPB1 YEAST		1733
tr A0A1B2J8C6 A0A1B2J8C6_PICPA		1743
tr F2QW17 F2QW17_KOMPC		1743
tr A3GID7 A3GID7_PICST		1739
tr A0A1D8PUA6 A0A1D8PUA6_CANAL		1728
tr G8BEH9 G8BEH9_CANPC		1746

Fig. 6 Mix and Match analysis of the elongation subunits β' of *E. coli* and Rpb1 of yeast RNAP II. For figure legends, refer to Figs. 3 and 4

Fig. 6 shows the mix and match analysis of the elongation subunits from pro- and eukaryotes MSU RNAPs. The active site regions are highlighted in yellow and the representative sequences are highlighted in yellow. In eukaryotic elongation subunits, the catalytic region is brought to the N-terminal region whereas it is found very close to the C-terminal region. There are 4 Cs at the N-terminal region in prokaryotic elongation subunits (out of which 2 are aligning in both) suggesting a possible additional Zn binding region. There are a very few motifs aligning in both. Significant among them are -GHIELA-, -NLM/LGKRVDF/YS-, possible metal binding region -DFDGDE/QM-, -DTAV/LKTAE/NT/SGYI/L-, -L/V/IAAQSIGEPA/ GTQM/LTL/MXTFH-. The metal binding region -NADFDGD- is aligning in both and placed in the middle of the enzyme. The C-terminal ends in eukaryotes usually end in -PDDSDE/DE/DN-(except in yeasts and Alligator) and -GSDNE/Din prokaryotes, suggesting a possible metal binding site (DxD) at the ends. It is interesting to note that the prokaryotic active site is placed at the C-terminal region (~900 amino acids) whereas the eukaryotic active site is placed at the N-terminal region (~90 amino acids) with their built-in Zn binding motifs.

8. ACTIVE SITE ANALYSIS OF THE S. cerevisiae MSU RNAP- II

8.1 Catalytic Region

It has been found that almost all DNA polymerases and SSU RNAPs use an invariant K for catalysis, i.e., in the initial proton transfer reactions [34,29]. However, in all MSU RNAPs analyzed, no K was found at the expected distance from the template binding YG/FG pair but an equivalent invariant R (Table 4). However, a detailed analysis has shown that all prokaryotic DNA polymerases II also use an invariant R in catalysis with similar distance conservations [34, 4] instead of a usual K; interestingly, an enzyme also possesses primase activity and along with associated 3'—5' exonuclease activity.

Table 4 shows the invariant template binding YG pair with its catalytic R in the initiation (β and Rpb2) and in the elongation subunits (β ' and Rpb1) of eubacterial and eukaryotic MSU RNAPs, respectively (Figs. 3-7). The catalytic R in both in the initiation and elongation subunits of the MSU RNAPs will serve as a proton abstractor for initiating the catalysis as explained by Palanivelu

catalytic R in the initiation and elongation subunits of eubacterial and eukaryotic MSU RNAPs.

It is interesting to note that the YG pair appears to be specific for polymerases using DNA as the prokaryotic template (including the eukaryotic MSU RNAPs, data not shown) as it is not reported in RNA dependent RNAPs where they use RNA as the template [35]. In fact, Kotsyuk et al. [37] have shown that the DNA dependent T7 RNAP requires both the strands for activity and uses two YG pairs and there was no activity when single-stranded DNA was used as the substrate. However, the eukaryotic initiation subunit Rpb2 uses a functional equivalent FG and the initiation subunit from higher eukaryotes including human uses a completely different pair, KG but followed by an F. The elongation subunits in all eukaryotes use a functionally equivalent FG pair. However, it is interesting to note the catalytic amino acid is R and is completely conserved in all MSU RNAPs II. Three invariant Cs between catalytic R and YG/FG pair in the elongation subunits (β' and Rpb1) are highly conserved in both pro- and eukaryotes but the three Cs are placed at a different distances as 5+12+15 and 4+7+14 with an 11 amino acids span in both the cases (Table 4). These completely conserved Cs are implicated in Zn binding in the bacterial elongation subunits by crystallographic analysis [36].

of NTP 8.2 Mechanism and **dNTP** Discrimination in Eukaryotic MSU **RNA Polymerases**

NTP and dNTP discrimination in nucleic acid polymerases is usually achieved by base pairing, base stacking, hydrogen bonding and also by specific interaction(s) of completely conserved amino acid(s) in the nucleotide interacting domains. A completely conserved R at -5 in the Rpb2 and -6 in Rpb1 is known to participate in the nucleotide discrimination and select only NTPs in the polymerization site (Table 4). Similar absolute conservation is seen in SSU RNAPs from viruses, mitochondria and chloroplasts and DNA polymerases too. A conspicuous absence of an invariant R at the expected distance at -6 in the elongation subunit β' of eubacterial MSU RNAPs (Table 4) raises the question of how they discriminate the NTPs from dNTPs. Modelling of the substrate NTP bound to the T. thermophilus RNAP active site suggests that N⁴⁵⁸ (numbering

[4]. Table 4 shows the invariant YG/FG pair and its from E. coli RNA polymerase) within a highly conserved sequence motif 458NADFDGD464 that includes the catalytic Asp triad (D460, D462, D464) could mediate specific recognition of the [39,36]. ribose atom Functional in vitro analysis demonstrated that the substitutions of the corresponding β' N^{458} residue led to the loss of discrimination between NTP and dNTP substrates as well as to defects in RNA chain extension [39]. It is interesting to note that substitution of the corresponding amino acid in yeast Rpb1 (N479Y) is lethal in the same sequence motif ⁴⁷⁸YNADFDGD in eukaryotes [40]. The absolutely conserved N in both the cases could discriminate the NTP from dNTP by recognizing the 2'-OH of the ribose and suggested that the crucial N could interact with both the 2'-OH as well 3'-OH [40].

> The highly conserved S/T adjacent (N in S. cerevisiae) to the catalytic R in the initiation and elongation subunits likely recognizes the 2'-OH in the NTPs and possibly makes a hydrogen bond and discriminate dNTPs, which lack a 2'-OH (the yeast elongation subunit, Rpb1 could also use the invariant N⁴⁷⁹ to make the necessary H bond with the 2'-OH a well as 3'-OH as shown in elongation subunit β' of eubacterial MSU RNAPs [4]. Kaplan et al. [41] have shown by SDM experiments that the completely conserved His1085 both in prokaryotes (β') and eukaryotes Rpb1 might involve in NTP selection and substitution of Ala or Phe resulted in inviability. This H1085 in the trigger loop (amino acids from 1060 to 1101) was shown to make a contact with the β phosphate either through H-bond or salt bridge and the L1081 was located at the 3' end of the RNA [42]. Mix and match analysis also shows and that particular His is completely conserved in both prokaryotes and eukaryotes in highly conserved stretch (Fig. 6). Another amino acid Q1078 in the same block is also shown to be essential and replacing Q1078 in Rpb1 to either N or S is lethal in yeast [41], consistent with a key role of these residues in NTP/dNTP discrimination.

9. THE INITIATION AND ELONGATION SUBUNITS WORK IN TANDEM DURING TRANSCRIPTION **CYCLES EUKARYOTIC MSU RNAP II**

A close similarity is observed between the prokaryotic and eukaryotic transcription cycles. For example, it has been shown in prokaryotes (E. coli) that the initiation of transcription by the β subunit is not smooth and it makes many aborted transcripts of sizes 2-7 nts before the elongation step is taken over by the β' subunit [43]. It was found that RNA/DNA hybrids of less than 8-bp display markedly less stability than those that are 8 bp or longer. In fact, the E. coli MSU RNAPs use these short RNA/DNA hybrids of 8 bp or longer for the ternary elongation complex (TEC) formation and further processivity Furthermore, Zaychikov et al. [45] have shown that an ~17 bp region of the DNA called the 'transcription bubble' was melted to expose the template strand for transcription in E. coli. MSA analysis agreed with these findings with the 7 amino acid gap between the catalytic R and the template binding YG pair in the initiation subunit β and 17 amino acid gap ('transcription bubble') between the catalytic R and the template binding YG pair in the elongation subunit β ' (Table 5) [4].

A similar observation of abortive initiations in eukaryotic Rpb2 was also reported by Pal and Luse [32]. In addition to, as in prokaryotes, in eukaryotes also a 9 bp DNA-RNA stable hybrid is formed which extends from the active centre at nearly right angles to the entering DNA during the elongation cycle [16,33]. Fiedler and Timmers [33] results further support the MSA findings where the transition from abortive to

productive elongation cycle occurred once the RNAP register +10 nts (Table 5). These observations are further supported by Luse who have shown that the promoter clearance was complete with ~10 nts [46]. Pal et al on analyzing the 'transcription bubble' have found that regardless of promoter spacing, the upstream edge of the 'transcription bubble' formed 20 bp from TATA which is in close agreement with MSA analysis data [47]. A similar finding was reported by Giardina and Lis [48]. Furthermore, Holstege et al analyzed the transcription initiation by the yeast RNAP II in a highly efficient in vitro transcription system composed of essentially homogeneous protein preparations and found that the downstream part of the 'transcription bubble' expands in a continuous motion, but the initially opened region on the non-template strand reclosed abruptly when transcription reached 11, which was accompanied by a switch from abortive initiations to productive mRNA synthesis (elongation) [49]. Similar findings were also reported by Barnes et al by analyzing the crystal structure of the transcribing RNAP II complex. The transcription bubble was unwound ~18-25 bases and transcripts of 10 nts or more resulted in promoter escape with stabilization of a mature bubble [50].

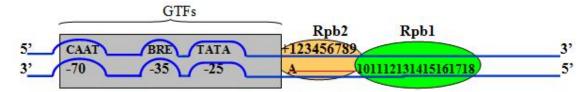


Fig. 7 A consensus model showing pre-initiation, initiation and elongation events by GTFs, Rpb2 and Rpb1 subunits of MSU RNAP- II of *S. cerevisiae*

NB: Thick lines indicate DNA strands and the thin line indicates the mRNA transcript A is shown as the first nucleotide at the +1 TSS GTFs, General Transcription Factors (TFII-D, -A, -B, -F, -E, -H) BRE, TFII-B Recognition Element (G/C-G/C-G/A-C-G-C-C)

Table 5. Metal-binding sites in prokaryotic and eukaryotic MSU RNAPs

Subunit (Organism)	Metal-biding sites	Method and Reference
β eubacteria (E. coli)	-671LEHDDA/ -809GYNFEDS* - (Mg ²⁺)	MSA (This communication)
β' eubacteria (<i>E. coli</i>)	- ⁴⁵⁸ YNADFDGDQM - (Mg ²⁺)	X-ray crystallographic data
- ⁸⁸³ RS	S ¹ VVSC ⁵ DTDFGVC ¹² AHC ¹⁵ Y ¹⁶ GR ⁹⁰¹ -(Zn ²⁺)*	[36]
Rpb2 eukaryote (Sc)	-893LDDDG897832GYNQED*S838-(Mg ²⁺)	MSA (This communication)
	-478YNAD*FD*GDEM ⁴⁸⁷ - (Mg ²⁺)	MSA (This communication)
- ⁵⁶ PR- ⁶ LGSIDRN ¹ LKC ⁴ QTC ⁷ QEGMNEC ¹⁴ PGHF ¹⁸ GH ⁸³ -(Zn ²⁺) ,,		

Possible metal binding sites arrived at by MSA and SDM

NB: The β' and RpB1 elongation subunits of eubacteria and eukaryotes contain both the Mg²⁺ and Zn²⁺ binding sites. In both the cases, the Zn²⁺ binding site is built in the catalytic region with the 3 invariant Cs which coordinates the Zn atom. Sc, S. cerevisiae

In contrast to prokaryotes, in eukaryotes, it has been found that an A is inserted (and not a usual G, as in prokaryotes) in the initiation site as it invariably uses a G for capping enzyme which is also associated with the RNAP II.

MSA analysis agrees with these findings with the 9 amino acid gap between the catalytic R and the template binding YG pair in the initiation subunit Rpb2 and 19 amino acid gap ('transcription bubble') between the catalytic R and the template binding FG pair in the elongation subunit Rpb1. Furthermore, Gnatt et al found that the contacts to the downstream and upstream parts of the hybrid are made by Rpb1 and Rpb2, respectively, which further supports the tandem arrangement of the initiation and elongation subunits as proposed in this model [16].

Fig. 7 shows a consensus model of the yeast MSU RNAP II subunits, from promoter recognition to initiation and elongation events during the transcription process. This has been confirmed that the RNAP II undergoes abortive initiations until it reached a position beyond +9 (i.e., the distance between the template binding FG pair and the catalytic R), at which stage the RNAP II was released from its promoter contacts and an elongation complex (TEC) is formed and a 20 bp "transcription bubble" formed from the TATA box [47]. A modular structure is proposed for the initiation and elongation subunits' function in eukaryotes also as suggested by in prokaryotes [51].

10. METAL-BINDING SITES

The metal-binding sites of S. cerevisiae MSU RNAP II is arrived at from the data obtained by the MSA (this work) and X-ray crystallographic analyses [36] and SDM experiments [39] on eubacterial MSU RNAPs. The eukarvotic Rpb2 and Rpb1 subunits show many possible metal binding sites (highlighted in green). A Mg²⁺ ion binding site is found in the mRNA initiation subunit, Rpb2, from all eukaryotes. Similarly, a Mg²⁺ and a Zn²⁺ binding motifs are also found in the elongation subunits, Rpb1, of all eukaryotic RNAPs II. Crystallographic analysis of the T. aquaticus RNAP by Zhang et al have shown that the Mg atom is chelated at an absolutely conserved -NADFDGD- motif in the elongation subunit and surprisingly the same invariant motif is found by MSA in all eukaryotic elongation subunits as well [36]. Moreover, this is one of the regions that align in both the pro- and eukaryotes by mix and match analysis (Fig. 6). Interestingly, substitution of these Ds by A

(D→A) gave rise to a dominant lethal phenotype and showed no detectable enzyme activity [45]. In addition to that, a Zn binding motif is also recognized in the eubacterial and eukaryotic elongation subunits with three conserved Cs. However, their distance arrangements between the three Cs are found to be different (Table 5). Another distinguishing feature between the eukaryotic and eubacterial elongation subunits is the catalytic and Zn binding regions, i.e., in E. coli β' subunit, the amino acids 888, 895 and 898 are shown to be involved in Zn binding (Table 5) but in eukaryotic elongation subunits it is placed at the very beginning of the N-terminal and covers the amino acids region 57-82 (Table 5). The Zn binding motif that is located in Rpb1 subunits may play a similar role in the Zn mediated proof-reading function as proposed for DNA polymerases and MSU RNAPs from eubacteria [4,29].

11. MECHANISM OF ACTION OF THE MSU RNAP II of S. cerevisiae

A minimal number of steps involved in the catalytic cycle of RNAPs consist of NTP selection, Watson-Crick base pairing with the complementary nucleotide to the template, catalysis, pyrophosphate release and translocation. As the catalytic regions are found to be similar in both the initiation and elongation subunits, the polymerization mechanism could be also similar in both the initiation and elongation reactions. Figs. 8 and 9 describe the reactions involved in the initiation and elongation cycles during transcription in *S. cerevisiae*. (all participating amino acids are not shown in the figures).

11.1 Mechanism of Initiation by Rpb2 Subunit of the MSU RNAP II from *S. cerevisiae*

Step 1: Enzyme and the NTP at the Entry Site: Template binding pair (Tyr-Gly) and nucleotide discrimination by the invariant amino acids, Watson-Crick base pairing of the incoming NTP with the template DNA. The catalytic site amino acid Arg in positioned for proton abstraction. ATP is the initiating nucleotide.

Step 2: Proton abstraction and nucleophilic attack of the α -phosphate of NTP: Electronic transition at the active site for proton abstraction by the active site Arg and an electrophilic and nucleophilic attack of the incoming NTPs on the 3'-OH.

Step 3: Phosphodiester bond formation: Proton abstraction by the active site amino acid Arg with simultaneous formation of $3'\rightarrow 5'$ phosphodiester bond with the incoming NTPs.

Step 4: Inorganic pyrophosphate formation and translocation of the enzyme to next nucleotide: Proton transfer from the active site amino acid Arg, formation of inorganic pyrophosphate, active site restoration and translocation of the enzyme to the next nucleotide (Fig. 8).

11.2 Mechanism of elongation by Rpb1 subunit of the MSU RNAP II from S. cerevisiae

Step1: Enzyme and the NTP at the Entry Site: Template binding by the template binding amino acids (Phe-Gly) and nucleotide discrimination by the invariant amino acids, Watson-Crick base pairing of the incoming NTP with the template

DNA. The catalytic site amino acid Arg in positioned for proton abstraction.

Step 2: Proton abstraction and nucleophilic attack of α -phosphate of NTP: Electronic transition at the active site for proton abstraction by the active site Arg and an electrophilic and nucleophilic attack of the incoming NTPs on the 3'-OH.

Step 3: Phosphodiester bond formation: Proton abstraction by the active site amino acid Arg with simultaneous formation of $3'\rightarrow 5'$ phosphodiester bond with the incoming NTPs.

Step 4: Inorganic pyrophosphate formation and translocation of the enzyme to next nucleotide: Proton transfer from the active site amino acid Arg, formation of inorganic pyrophosphate, active site restoration and translocation of the enzyme to the next nucleotide (Fig. 9).

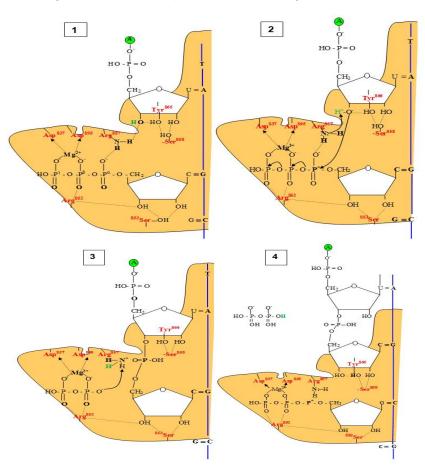


Fig. 8 Steps (1-4) proposed for the polymerization reactions during initiation of transcription by the yeast initiation subunit Rpb2 of the MSU RANP II

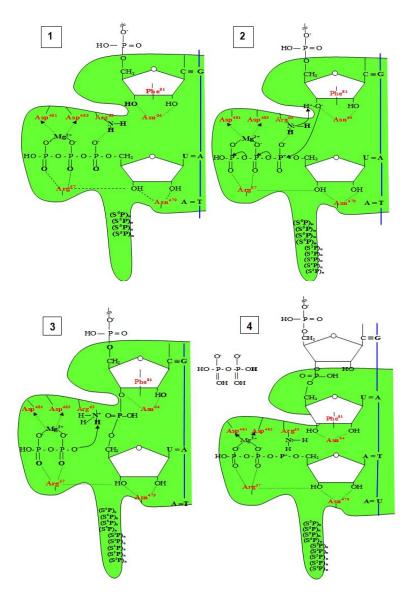


Fig. 9 Steps (1-4) proposed for the polymerization reactions during elongation of transcription by the yeast elongation subunit Rpb1 of the MSU RNAP II

12. PROOF-READING MECHANISM AND TRANSCRIPTION TERMINATION IN S. cerevisiae MSU RNAP II

Maintaining high fidelity during transcription is essential for the accurate transfer of genetic information from DNA to RNA as the first step in gene expression. The proof-reading mechanism is well established in DNA polymerases [34]. However, it is poorly understood in RNAPs. Transcription accuracy is relatively high, as RNAPs generally misincorporate one wrong nucleotide/~100000 bases. As RNAPs are also Zn metalloenzymes the Zn-mediated deletion of

the misincorporated NTP could be a possible mechanism as explained for DNA polymerases and eubacterial MSU RNAPs as both use DNA template to produce a complementary DNA/RNA strand [34,4,29,30]. There is no distinct proofreading domain or separate proof-reading subunit as reported in DNA polymerases. However, in eubacterial and eukarvotic elongation subunits of the MSU RNAPs a highly conserved Zn binding motif built into the catalytic region itself might involve in the proof-reading function [4]. Presence of two metal ions at the active site is supported by a 2.8 Å difference Fourier map, one persistently bound and the

other possibly exchangeable during RNA synthesis [28]. Unlike the DNA polymerases, the RNAPs could move forward (when correct NTP is bound) and backtrack (when wrong NTP is bound and secondary structures) [42]. When the enzyme stalls due to secondary structures and/or misincorporation, the proof-reading begins with fraying of the misincorporated nucleotide away from the DNA template and subsequent backtracking of RNAP by one position followed by nucleolytic cleavage of a dinucleotide that contains the misincorporated nucleotide [52]. Sydow and Cramer [52] have also proposed that the exo- or endonucleolytic cleavage occurs at the same active site that is used for polymerization, which is further confirmed by Xray crystallographic analysis of the E. coli and T. thermophilus MSU RNAPs suggesting a possible Zn mediated cleavage [36]. The complete conservation of the active sites and Zn binding motif in all prokaryotic and eukaryotic elongation subunits, found by MSA analysis also further proves this point. The proof-reading mechanism may not be that much important during initiation events by the Rpb1 subunits, as they transcribe only very small regions, i.e., only ~ 10 nts and again such transcripts are well within the 5' untranslated regions of the RNA; so any error it makes may not be deleterious. Furthermore, a Zn binding motif is not found in the active site region of the initiation subunits. However, proofreading is very important for the elongation subunit, viz., Rpb1 to produce an error-free transcript, as they are going to be translated into active proteins. Therefore, when the enzyme stalls at the wrong nucleotide, it could backtrack and make an endonucleolytic cleavage or simply remove the wrong nucleotide by exonuclease action [53, 34]. Unlike in DNAdependent RNA polymerases, where the polymerization and proofreading functions are separated, in RNAPs, synthesis and degradation are performed by the same active centre [54]. In support of this, Liu et al have shown that following the initiation, RNAP II alone was capable of RNA transcript elongation and of proofreading [55]. Thus, the fidelity in the transcription is ensured by two mechanisms in MSU RNAPs, viz. by the selection of correct NTPs at the entry site and an efficient Zn mediated proof-reading mechanism.

The enzyme could use the same strategy for transcription terminations (intrinsic types) at TTSs. That is, when the enzyme stalls due to the stem-loop structures commonly found at the

TTSs, it could backtrack and make an endo- or exonucleolytic cleavage and release the RNA transcript, using the same active site by Zn mediated reaction.

13. OTHER CONSERVED REGIONS AND AMINO ACID(s) IN THE MSU RNAP II

The above analysis to find out the conserved motifs, active sites and invariant amino acids, with respect to the substrate binding and catalytic region, form only a small region of the MSU RNAPs. Apart from these, there are a large number of single amino acid conservations like Ys, Ws, Cs, Ps and Gs, diads, triads and long stretches of conserved amino acids in all these polymerases (Figs. 4 – The highly conserved stretches of amino acids may represent the different conserved structural motifs such as loops, rudders, jaws, bridge helix, exit/entry points for DNA/RNA, etc. Thus, they could form small functional domains. A good number of completely conserved Ps in these polymerases is implicated in making the necessary bents on the enzyme's structure during unique folding, substrate and template binding. The completely conserved Cs other than the catalytic region may play a role in the formation of disulphide bridges to make the enzyme more compact and stable.

14. CONCLUSIONS

Unlike reported for the DNA polymerases and SSU RNA polymerases, the MSU enzymes of eukaryotic MSU RNAPs are similar to the prokaryotic ones and use an R as the catalytic amino acid. However, they maintain a different distance conservation in the initiation and elongation subunits. An invariant Zn²⁺ binding motif found in the Rpb1 elongation subunits, is proposed to participate in the proof-reading function. Therefore, it is clear from this communication that even though the transcription machinery of eukaryotes is much more complex than that of prokarvotes, the general principals of transcription and its regulation are found to be highly preserved. Hence, it is proposed that the MSU RNAP II of eukaryotes may also follow very polymerization and proof-reading mechanisms like their counterparts in eubacteria. MSA data and the available experimental data show that both the eubacterial and eukaryotic MSU RNAPs would have evolved from a common ancestor.

ACKNOWLEDGEMENTS

The author wishes to thank Dr. H. Shakila, Professor & Head, Department of Molecular Microbiology, School of Biotechnology, Madurai Kamaraj University, Madurai for useful suggestions on the manuscript.

COMPETING INTERESTS

The author has declared that no competing interests exist.

REFERENCES

- Anikin M, Molodtsov V, Temiakov D, McAllister WT. Transcript slippage and recoding. In: Atkins JF, Gesteland RF, Bujnicki JM. (eds). Recoding: Expansion of decoding rules enriches gene expression. 24th edn. Springer, New York. 2010;409– 432.
- Sahin U, Kariko K, Türeci Ö. mRNAbased therapeutics — Developing a new class of drugs. Nat Rev Drug Discov. 2014;13:759–780.
- Conry RM, LoBuglio AF, Wright M, Sumerel L, Pike MJ. Characterization of a messenger RNA polynucleotide vaccine vector. Cancer Res. 1995;55:1397–1400.
- Palanivelu P. Multi-subunit RNA Polymerases of Bacteria An insight into their active sites and catalytic mechanism. Indian J Sci Technol. 2018;11:1-37.
- Roeder RG, Rutter, WJ. Multiple forms of DNA-dependent RNA polymerase in eukaryotic organisms. Nature. 1969;224: 234–237.
- Werner F, Grohmann D. Evolution of multisubunit RNA polymerases in the three domains of life. Nat Rev Microbiol. 2011;9: 85–98.
- Domecq C, Trinhl V, Langelier MF, Archambault J, Coulombe B. Inhibitors of multisubunit RNA polymerases as tools to study transcriptional mechanisms in prokaryotes and eukaryotes. Curr Chem Biol. 2008;2:20–31.
- 8. Ma C, Yang X, Lewis PJ. Bacterial transcription as a target for antibacterial drug development. Microbiol Mol Biol Rev. 2016;80:139–60.
- Ream TS, Haag JR, Pikaard, CS. Plant multisubunit RNA polymerases IV and V: in Murakami, KS, Trakselis, MA (eds.), Nucleic Acid Polymerases, Nucleic Acids

- and Molecular Biology. Springer-Verlag Berlin Heidelberg. 2014;30. DOI: 10.1007/978-3-642-39796-7 13
- Lane WJ, Darst SA. Molecular evolution of multisubunit RNA polymerases: Sequence analysis. J Mol Biol. 2010;395:671–85.
- Sweetser D, Nonet M, Young RA. Prokaryotic and eukaryotic RNA polymerases have homologous core subunits. Proc Natl Acad Sci. USA. 1987; 84:1192–1196.
- Minakhin L, Bhagat S, Brunning A, Campbell EA, Darst SA, Ebright RH, Severinov K. Bacterial RNA polymerase subunit omega and eukaryotic RNA polymerase subunit RPB6 are sequence, structural, and functional homologs and promote RNA polymerase assembly. Proc Natl Acad Sci USA. 2001;98:892-897.
- Nonet M, Sweetser D, Young RA. Functional redundancy and structural polymorphism in the large subunit of RNA polymerase II. Cell. 1987;50:909-915.
- Todone F, Weinzierl R, Brick P, Onesti S. Crystal structure of RPB5, a universal eukaryotic RNA polymerase subunit and transcription factor interaction target, Proc Natl Acad Sci. USA. 2000;97:6306-6310.
- Bushnell DA, Kornberg RD. Complete, 12subunit RNA polymerase II at 4.1-Å resolution: Implications for the initiation of transcription. Proc Natl Acad. Sci. USA. 2003;100:6969–6973.
- Gnatt AL, Cramer P, Fu J, Bushnell DA, Kornberg RD. Structural basis of transcription: An RNA polymerase ii elongation complex at 3.3Å resolution; 2001.
 - Available:www.sciencexpress.org DOI: 10.1126/science.1059495
- 17. Young RA. RNA Polymerase II. Ann Rev Biochem. 2003;60:689–715.
- Hahn S. Structure and mechanism of the RNA polymerase II transcription machinery, Nat. Str. Biol. Mol. Biol. 2004; 11:394-403.
- West ML, Corden JL. Construction and analysis of yeast RNA polymerase II CTD deletion and substitution mutations, Genetics. 1995;140:1223-1233.
- Sylvain E, Shona M. Cracking the RNA polymerase II CTD code. Trends Genet., 2008;24:280–288
- Egloff S, O'Reilly D, Chapman RD, Taylor A, Tanzhaus K, Pitts L, Eick D, Murphy S. Serine 7 of the RNA polymerase II CTD is specifically required for snRNA gene

- expression. Science. 2007;318:1777–1779.
- 22. Phatnani HP, Greenleaf AL. Phosphorylation and functions of the RNA polymerase II CTD. Genes Dev. 2006;20:2922-2936.
- McCracken S, Fong N, Rosonina E, Yanku lov K, Brothers G, Siderovski D, Hesse A Foster S, Shuman S, Bentley DL. 5'-Capping enzymes are targeted to premRNA by binding to the phosphorylated carboxy-terminal domain of RNA polymerase II. Genes Dev. 1997;11:3306-3318.
- Fong N, Bentley DL. Capping, splicing, and 3' processing are independently stimulated by RNA polymerase II: different functions for different segments of the CTD. Genes Dev. 2001;15:1783–1795.
- 25. Kolodziej PA, Woychik, N, Liao SM, Young RA. RNA polymerase II Subunit composition, stoichiometry, and phosphorylation. Mol Cell Biol. 1990;10:1915-1920.
- Sheffer A, Varon M, Choder M. Rpb7 can interact with RNA polymerase II and support transcription during some stresses independently of Rpb4. Mol. Cell. Biol. 1999;19:2672–2680.
- Woychik NA, Young RA. Genes encoding transcription factor IIIA and the RNA polymerase common subunit RPB6 are divergently transcribed in *Saccharomyces* cerevisiae. Proc. Natl. Acad. Sci. USA. 1992;89:3999-4003.
- 28. Cramer P, Bushnell DA, Kornberg RD. Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. Science. 2001;292:1863-76.
- 29. Palanivelu P. Single subunit RNA Polymerases An insight into their active sites and mechanism of action, Biotech J Int. 2017;20:1-35.
- Palanivelu P. Active sites of the multisubunit RNA polymerases of Eubacteria and chloroplasts are very similar in Structure and Function. Indian J Sci Technol. 2019;12:1-32.
- Cramer P. Multisubunit RNA polymerases. Curr Opin Struct Biol. 2002;12:89–97.
- 32. Pal M, Luse DS. The initiation–elongation transition: Lateral mobility of RNA in RNA polymerase II complexes is greatly reduced at +8/+9 and absent by +23. EMBO J. 1997;16:7468–7480.
- 33. Fiedler U, Timmers HTM. Analysiis of the open region of RNA polymerase II

- transcription complexes in the early phase of elongation. Nucleic Acids Res. 2001;29: 2706-2714.
- Palanivelu P. DNA polymerases An insight into their active sites and mechanism of action, Int. J. Biochem. Res. Rev. 2013;3:205-247.
- 35. Tunitskaya VL, Kochetkov SN. Structural and functional analysis of bacteriophage T7 RNA polymerase. Biochemistry (Moscow). 2002;67:1124–35.
- Zhang G, Campbell EA, Minakhin L, Richter C, Severinov K, Darst SA. Crystal structure of *Thermus aquaticus* core RNA polymerase at 3.3 A resolution. Cell. 1999; 98:811-824.
- Kostyuk SM, Dragan DL, Lyakhov VO, Rechinsky VL, Tunitskaya BK. Chernov SN, Kochetkov E. Mutants of T7 RNA polymerase that are able to synthesize both RNA and DNA. FEBS Lett. 1995; 369:165–168.
- Hausmann S, Shuman S. Characterization of the CTD Phosphatase Fcp1 from Fission Yeast: Preferential dephosphorylation of serine 2 versus serine 5. J Biol Chem. 2002;277:21213-21220.
- Svetlov V, Vassylyev DG, Artsimovitch I. Discrimination against deoxyribonucleotide substrates by bacterial RNA polymerase. J Biol Chem. 2004;279:38087-90.
- Trinh V, Langelier MF, Archambault J, Coulombe B. Structural perspective on mutations affecting the function of multisubunit RNA polymerases. Microbiol Mol Biol Rev. 2006;70:12–36.
- Kaplan CD, Larsson KM, Kornberg RD. The RNA polymerase II trigger loop functions in substrate selection and is directly targeted by alpha-amanitin. Mol Cell. 2008;30:547–556.
- 42. Wang D, Bushnell D, Westover K, Kaplan C, Kornberg RD. Structural basis of transcription: Role of the trigger loop in substrate specificity and catalysis. Cell. 2006;127:941–954.
- Campbell EA, Korzheva N, Mustaev A, Murakami K, Nair S, Goldfarb A, Darst SA. Structural mechanism for rifampicin inhibition of bacterial RNA polymerase. Cell. 2001;104:901–912.
- 44. Kireeva ML, Komissarova N, Waugh DS, Kashlev M. The 8-nucleotide-long RNA: DNA hybrid is a primary stability determinant of the RNAP II elongation

- complex. J Biol Chem. 2000;275:6530-6536.
- Zaychikov E, Denissova L, Meier T, Gotte M, Heumann H. Influence of Mg²⁺ and temperature on formation of the transcripttion bubble. J Biol Chem. 1997;272:2259–67.
- 46. Luse DS. Promoter clearance by RNA polymerase II. Biochim Biophys Acta. 2013;1829:63–68.
- Pal M, Ponticelli AS, Luse DS. The role of the transcription bubble and TFIIB in promoter clearance by RNA polymerase II. Mol. Cell. 2005;19:101-110.
- 48. Giardina C, Lis JT. DNA melting on yeast RNA polymerase II promoters. Science. 1993;261:759-762.
- 49. Holstege FCP, Fiedler U, Timmers HTM. Three transitions in the RNA polymerase II transcription complex during initiation. EMBOJ. 1997;16:7468–7480.
- 50. Barnes CO, Calero M, Malik I, Graham BW, Spahr H, Lin G, Cohens A, et al. Crystal structure of a transcribing RNA polymerase II complex reveals a complete

- transcription bubble. Mol Cell. 2015;59: 258-269.
- Severinov K, Mustaev A, Kukarin A, Muzzin O, Bass I, Darst SA, Goldfarb A. Structural modules of the large subunits of RNA polymerase. Introducing archaebacterial and chloroplast split sites in the beta and beta'subunits of Escherichia coli RNA polymerase. J Biol Chem. 1996;271: 27969–27974.
- 52. Sydov JH, Cramer P. RNA polymerase fidelity and transcriptional proofreading. Curr Opin Struct Biol. 2009;19:732-9.
- 53. Nudler E. RNA polymerase active center: The molecular engine of transcription. Ann Rev Biochem. 2009;78:335–361.
- Sosunov V, Sosunova E, Mustaev A, Bass I, Nikiforov V, Goldfarb A. Unified twometal mechanism of RNA synthesis and degradation by RNA polymerase. EMBO J. 2003;22:2234–44.
- Liu X, Bushnell DA, Kornberg RD. RNA polymerase II transcription: Structure and mechanism. Biochim Biophys Acta. 2013; 1829:2-8.

© 2019 Palanivelu; This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Peer-review history:
The peer review history for this paper can be accessed here:
http://www.sdiarticle3.com/review-history/49940