Exosome-Mediated Recognition and Degradation of mRNAs Lacking a Termination Codon

Ambro van Hoof,^{1,2*} Pamela A. Frischmeyer,^{1,3} Harry C. Dietz,^{1,3}
Roy Parker^{1,2*}

One role of messenger RNA (mRNA) degradation is to maintain the fidelity of gene expression by degrading aberrant transcripts. Recent results show that mRNAs without translation termination codons are unstable in eukaryotic cells. We used yeast mutants to demonstrate that these "nonstop" mRNAs are degraded by the exosome in a 3'-to-5' direction. The degradation of nonstop transcripts requires the exosome-associated protein Ski7p. Ski7p is closely related to the translation elongation factor EF1A and the translation termination factor eRF3. This suggests that the recognition of nonstop mRNAs involves the binding of Ski7p to an empty aminoacyl-(RNA-binding) site (A site) on the ribosome, thereby bringing the exosome to a mRNA with a ribosome stalled near the 3' end. This system efficiently degrades mRNAs that are prematurely polyadenylated within the coding region and prevents their expression.

mRNA biogenesis is a multistep process with a certain frequency of errors, either due to inherent inaccuracies in transcription and processing or due to mutations in the DNA template. The cell has evolved mechanisms to rapidly degrade aberrant mRNAs, such as unspliced premRNAs, mRNAs with aberrantly long 3' untranslated regions (3'UTRs), and mRNAs with premature translation termination codons (1). Recently, it has been found that eukaryotic mRNAs that do not contain a termination codon are rapidly degraded (2). The rapid decay of these transcripts is referred to as nonstop mRNA decay and requires translation of the mRNA (2). However, degradation of a PGK1 mRNA, from which all in-frame termination codons have been removed (nonstop-PGK1), requires none of the enzymes involved in the major pathway for mRNA degradation, which occurs by deadenylation, decapping, and 5'-to-3' digestion (2-5). This suggests that nonstop mRNAs might be degraded by the exosome complex of 3'-to-5' exoribonucleases, the functions of which include 3'-to-5' degradation of mRNA in the cytoplasm, nuclear processing of ribosomal RNA and small nucleolar RNAs, and degradation of processing intermediates and stalled mRNAs in the nucleus (6-8).

To test whether the exosome functions in nonstop decay, we first examined nonstop decay in a ski4-1 strain of yeast. The ski4-1 allele encodes a point mutation in one of the core exosome subunits that specifically disrupts cytoplasmic 3'-to-5' degradation of mRNA without affecting any of the other known functions of the exosome (9). The ski4-1 mutation stabilizes the nonstop-PGK1 mRNA at least sixfold (Fig. 1, A and B). Exosome-mediated degradation of normal cellular mRNAs requires the exosome and two other factors (6, 9). One factor is a heterotrimeric helicase complex of Ski2p, Ski3p, and Ski8p (6, 10). As shown in Fig. 1, C through E, Ski2, -3, and -8 are all required for nonstop mRNA degradation. The second factor required for exosome-mediated mRNA decay is Ski7p, and deletion of SKI7 also caused stabilization of nonstop mRNAs (Fig. 1F). Because Ski2p and Ski7p localize to the cytoplasm (10, 11), we interpret these observations to indicate that nonstop mRNAs are degraded 3' to 5' by the cytoplasmic exosome.

Given that the major deadenylase (Ccr4p) is not required for nonstop decay (2) and that degradation occurs by the exosome, it is possible that the exosome both deadenylates and degrades nonstop mRNAs. This would be surprising because normal mRNAs cannot be deadenylated by the exosome (5). Alternatively, an unidentified nuclease may remove the polyadenylate [poly(A)] tail from nonstop mRNAs, followed by exosome-mediated decay.

To investigate whether the exosome degrades the poly(A) tail of nonstop transcripts, we performed transcriptional pulse-chase experiments. In these experiments, transcription of the reporter mRNA was induced briefly and was followed by transcriptional repression, which yielded a synchronous population of mRNA whose fate could be monitored. For comparison,

wild-type PGK1 mRNA was synthesized with a poly(A) tail of approximately 70 residues and was subsequently deadenylated slowly (Fig. 2A) (12). In contrast, nonstop-PGK1 transcripts disappeared rapidly without any detectable deadenylation intermediates (Fig. 2B). In addition, in a ski 7Δ strain, the nonstop mRNA persisted as a fully polyadenylated species for 8 to 10 min before disappearing (Fig. 2C). These data indicate that exosome function is required for rapid degradation of both the poly(A) tail and the body of the mRNA. Based on these observations, we suggest that nonstop mRNAs are rapidly degraded in a 3'-to-5' direction by the exosome, beginning at the 3' end of the poly(A) tail (13).

Two observations suggest a mechanism by which nonstop mRNAs are specifically recognized and targeted for destruction by the exosome. First, nonstop mRNA degradation requires that a translating ribosome reach at least the poly(A) tail, and most likely the 3' end of the mRNA (2, 14). The simplest interpretation of these data is that nonstop mRNAs are recognized when a ribosome reaches the 3' end of the mRNA. Such a recognition would be analogous to the recognition of ribosomes with an empty A site by a tRNA-mRNA hybrid (tmRNA) in prokaryotes (15, 16). Second, the COOH-terminal region of the Ski7 protein is closely related to the guanosine triphosphatases (GTPases) EF1A and eRF3, including similarity in the GTPase domain (17-19). EF1A and eRF3 are translation

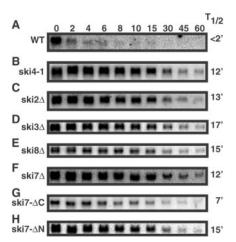


Fig. 1. Nonstop mRNA degradation requires exosome and cytoplasmic exosome cofactors. Nonstop-PGK1 mRNA stability was measured in wild-type (A), ski4-1 (B), ski2 Δ (C), ski3 Δ (**D**), ski8 Δ (**E**), ski7 Δ (**F**), ski7- Δ C (**G**), and ski7- ΔN (H) strains. Each strain contained a URA3 plasmid encoding the reporter gene and was grown to early- to mid-log phase at 30°C in media containing 2% galactose and lacking uracil. Transcription of the reporter gene was inhibited by replacing the media with media containing glucose (T = 0 min) and aliquots were taken thereafter. RNA was analyzed as described (9). The indicated half-lives are averages of at least two experiments and were calculated after correction for loading (9).

¹Howard Hughes Medical Institute, 4000 Jones Bridge Road, Chevy Chase, MD 20815, USA. ²University of Arizona, Department of Molecular and Cellular Biology, Tucson, AZ 85721, USA. ³Johns Hopkins University, Institute for Genetic Medicine, School of Medicine, Baltimore, MD 21205, USA.

^{*}To whom correspondence should be addressed. E-mail: ambro@u.arizona.edu (A.v.H.) or rrparker@u.arizona.edu (R.P.)

REPORTS

factors that interact with the A site of the ribosome when it contains a sense or nonsense codon, respectively. The interaction of Ski7p homologs with the ribosomal A site suggests that the homologous domain of Ski7p may function to distinguish nonstop from normal mRNAs by binding to the empty A site of ribosomes that have reached the 3' end of the mRNA. This hypothesis predicts that the COOH-terminal domain of Ski7p is specifically required for nonstop decay but may not be required for exosome-mediated degradation of normal mRNAs.

To determine the function of the Ski7p domains in exosome-mediated decay of nonstop and normal mRNAs, we generated yeast strains that express different deletion mutants of Ski7p (20). Two lines of evidence indicate that the NH₂-terminal nonconserved domain of Ski7p is necessary and sufficient for exosome-mediated degradation of normal mRNAs and that the

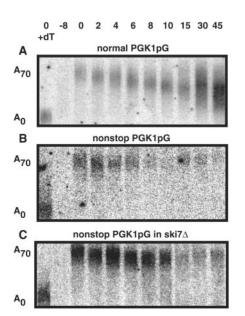


Fig. 2. Inactivation of the cytoplasmic exosome stabilizes polyadenylated nonstop mRNA. Normal (A) and nonstop (B and C) PGK1 mRNAs were analyzed by a transcriptional pulse-chase experiment in wild-type strains [(A) and (B)] and a ski7 deletion strain (C). A_o, mRNAs carrying a very short or no poly(A) tail. A₇₀, mRNAs carrying a poly(A) tail of approximately 70 residues. The strains in (B) and (C) had the nonstop PGK1 gene on a plasmid and were grown to early- to mid-log phase at 24°C in media containing 2% sucrose and lacking uracil. The strain shown in (A) was grown in 1% yeast, 2% peptone 2% sucrose media and carried the reporter integrated into the genome. However, similar results were obtained with a strain carrying PGK1 on a plasmid and grown in -URA media. We turned on transcription for 8 min by replacing the media with media containing 2% galactose. We then terminated transcription by adding 4% glucose (T = 0 min), and time points were taken. Forty micrograms of RNA isolated from each aliquot was cleaved with ribonuclease H (Promega) using oRP70 (CGGATAA-GAAAGCAACACCTGG) and analyzed by Northern blotting with a 6% polyacrylamide gel.

COOH-terminal GTPase domain does not play a role in exosome-mediated degradation of normal mRNAs. First, the NH₂-terminal domain, but not the COOH-terminal domain, is required for viability under conditions in which exosome-mediated decay is essential for viability (19). Second, the deletion of the NH₂-terminal part, but not the COOH-terminal part, of Ski7p causes a dramatic decrease in the rate of exosome-mediated decay of normal mRNAs (19).

Both *ski7* alleles stabilized the nonstop reporter transcript (Fig. 1, G and H), indicating that the COOH-terminal part of Ski7p functions in the nonstop mRNA degradation pathway. However, the COOH-terminal truncation of Ski7p has a smaller effect than either the NH₂-terminal deletion or complete deletion of SkI7. This suggests that other factors may to some extent be able to substitute for the COOH-terminal domain. Taken together, these results indicate that the NH₂-terminal part of Ski7p plays a central role in exosome-mediated mRNA decay and that the COOH-terminal domain plays a specific role in the degradation of nonstop mRNAs.

These results are consistent with the hypothesis that an interaction between the GTPase domain of Ski7p and the ribosome triggers exosome-mediated decay. One simple possibility is that Ski7p recruits the exosome to nonstop mRNAs. Consistent with this possibility, we observed that a large proportion of Ski7p copurified with two different subunits of the exosome (Ski4p or Rrp4p) (Fig. 3A) (21, 22). Ski7p remained in the unbound fraction in control purifications from strains with an untagged exosome (Fig. 3B). These results indicate that Ski7p physically associates with the exosome. This association is specific because neither Ski3p nor Lsm1p copurified with the exosome (22). In addition, the copurification of Ski7p with both Ski4p and Rrp4p is resistant to wash-

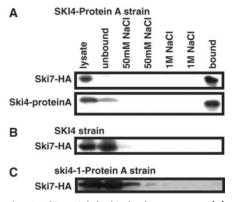


Fig. 3. Ski7p tightly binds the exosome. (A) Ski7p copurifies with protein A-tagged Ski4p. (B) Mock purification from a strain lacking the protein A tag. (C) Ski7p does not copurify with protein A-tagged ski4-1 mutant protein. Equal aliquots of each fraction of exosome purifications were analyzed by Western blotting using antibodies to protein A (Sigma) or to HA (Roche).

ing with 1 M NaCl (Fig. 3A) (22), suggesting a strong interaction between Ski7p and the exosome. The nuclear form of the exosome contains one additional subunit, Rrp6p (23, 24). Purification of protein A–tagged Rrp6 did not result in copurification of Ski7p (22), which is consistent with Ski7p being specific to the cytoplasmic exosome. Recently, Araki *et al.* (11) independently found that, when overexpressed, the NH₂-terminal part of Ski7p can coimmunoprecipitate with the exosome. The finding that Ski7p stably associates with the exosome through its NH₂-terminal suggests a mechanism to recruit the exosome to nonstop mRNAs recognized by the COOH-terminal of Ski7p.

To determine whether the interaction of Ski7p with the exosome is biologically relevant, we examined whether mutations in the exosome that disrupt all Ski7p-dependent functions of the exosome also disrupt Ski7-exosome interaction. Figure 3C shows that the ski4-1 mutation severely reduces the copurification of Ski7p with the exosome. One possibility is that the amino acid change in ski4-1 changes the binding site for Ski7p. This same ski4-1 mutation blocks exosome-mediated decay of both nonstop and normal mRNAs (Fig. 1B) (9). The observation that a mutation that prevents Ski7p from interacting with the exosome inhibits exosome-mediated mRNA decay indicates that the association of Ski7p with the exosome is important for the degradation of both normal and nonstop mRNAs.

One class of endogenous mRNAs subject to nonstop decay results from premature polyadenylation within the coding region (2). Another potential role for nonstop decay is to ensure the completeness of degradation for mRNAs that initiate 3'-to-5' decay while still being translated. In this case, as the exosome enters the coding region from the 3' end, it would encounter ribosomes coming from the 5' end. In both cases, the reason for the rapid degradation of nonstop mRNAs would be to prevent the production of truncated proteins. Similarly, translation of aberrant mRNAs containing premature termination codons has previously been shown to be deleterious to Caenorhabditis elegans (25). To test whether nonstop mRNAs can be translated into protein, we generated a nonstop allele of the HIS3 gene. Figure 4 shows that the nonstop his3 allele failed to complement a his3 deletion in a SKI+ strain. However, the nonstop his3 allele allowed rapid growth in the absence of added histidine when the strain was deleted for SKI2, SKI7, or SKI8. Even the COOH-terminal truncation of Ski7p, which specifically inhibits nonstop mRNA decay, allows for some growth in the absence of added histidine. These data suggest that the degradation of nonstop (his3) mRNA is effective in limiting the production of aberrant (His3p) protein, and in the absence of this

REPORTS

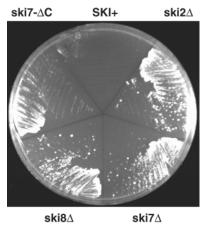


Fig. 4. Nonstop mRNA degradation is effective in limiting the translation of aberrant mRNAs. The HIS3 gene was amplified by polymerase chain reaction using oRP1075 (CGAGAGCTCAACA-CAGTCCTTTCCCGCAA) and oRP1077 (CGAG-GATCCACTTGCCACCTATCACC) and was cloned as a Sac I-Bam HI fragment into the CEN URA3 plasmid pRS416 (30). The nonstop his3 allele was created by deleting the first nucleotide of the termination codon (Quick-change kit, Stratagene). This creates an open reading frame that extends past the previously mapped polyadenylation sites (31). The nonstop his3 plasmid was transformed into strains that were $ura3\Delta$ and his 3Δ and were either SKI+, ski 2Δ , ski 7Δ . ski 8Δ , or ski7- Δ C. URA+ transformants were selected and streaked onto a plate lacking histidine. This plate is shown after a 2-day incubation at 30°C.

mRNA degradation pathway, protein products of nonstop mRNAs accumulate to functional levels.

In combination, these results define a mechanism of mRNA quality control that recognizes and degrades yeast mRNAs lacking translation codons, thereby preventing the production of truncated proteins. Because Ski protein homologs are present in the human genome (19, 26), we expect that the mechanism of nonstop decay is conserved. Transcripts that lack a termination codon are also recognized in prokaryotes (15, 16). It will be interesting to determine to what extent the prokaryotic and eukaryotic systems are similar and whether they are evolutionarily related.

References and Notes

- 1. P. Hilleren, R. Parker, RNA 5, 711 (1999).
- 2. P. A. Frischmeyer et al., Science 295, 2258 (2002).
- D. Muhlrad, C. J. Decker, R. Parker, Genes Dev. 8, 855 (1994).
- 4. C. Beelman et al., Nature 382, 577 (1996)
- 5. M. Tucker et al., Cell 104, 377 (2001).
- J. S. Jacobs Anderson, R. Parker, EMBO J. 17, 1497 (1998).
- 7. A. van Hoof, R. Parker, Cell 99, 347 (1999).
- 8. P. Mitchell, D. Tollervey, Nature Struct. Biol. 7, 843 (2000).
- A. van Hoof, R. R. Staples, R. E. Baker, R. Parker, Mol. Cell. Biol. 20, 8230 (2000).
- 10. J. Brown, X. Bai, A. W. Johnson, RNA 6, 449 (2000).
- 11. Y. Araki et al., EMBO J. 20, 4684 (2001).
- D. Muhlrad, C. J. Decker, R. Parker, Mol. Cell. Biol. 15, 2145 (1995).

- 13. Nonstop mRNAs are not detectably deadenylated, even when they are stabilized by deletion of SKI7 or several other exosome mutations. One possible explanation is that a stalled ribosome occupies the extreme 3' end of this mRNA and prevents exonucleases from digesting it. A corollary of this explanation is that in a wild-type strain, the exosome or associated proteins can dislodge a stalled ribosome at the 3' end of the mRNA or initiate 3'-to-5' decay of the mRNA in the presence of such a ribosome.
- 14. To test whether translation of nonstop PGK1 mRNA was required in cis, we introduced G₁₈ in its 5' UTR. This sequence forms a stable secondary structure and reduces translation by 4 orders of magnitude (27). This reduction in translation severely reduced exosome-mediated decay of the nonstop PGK1 mRNA (half-life = 14 min).
- A. W. Karzai, E. D. Roche, R. T. Sauer, *Nature Struct. Biol.* 7, 449 (2000).
- 16. A. Muto, C. Ushida, H. Himeno, *Trends Biochem. Sci.* 23, 25, (1998)
- L. Benard, K. Carroll, R. C. Valle, D. C. Masison, R. B. Wickner, J. Virol. 73, 2893 (1999).
- The SKI7 homology with translation factors is most evident in the GTPase domain, but multiple sequence alignment shows that the homology extends to the COOH-terminus of Ski7p, EF1A, and eRF3 (19).
- Supplemental data can be found on Science Online at www.sciencemag.org/
- 20. Alleles encoding either a COOH-terminal truncation or an NH₂-terminal deletion of Ski7p were integrated into the genome at the SkI7 locus and were expressed from the SkI7 promoter. The NH₂-terminal deletion removed amino acids 18 through 239, whereas the COOH-terminal truncation removed all amino acids from 265 to the COOH-terminal. The COOH-terminal truncation removes all of the translation factor homology.
- 21. Hemagglutinin (HA)-tagged Ski7p was generated as

- described (28) and introduced into strains that carried a protein A-tagged version of Rrp4p (29), Ski4p, or Rrp6p, which are subunits of the exosome. As a control, we used a similarly HA-tagged version of Ski3p, which is known not to copurify with the exosome (10). All five tagged proteins are expressed from their normal genomic locus and are functional (22). Ski3p, Ski4p, Ski7p, and Rrp6 are also expressed from their own promoters, whereas Rrp4p is expressed from the GAL10 promoter (29). Protein extracts were prepared by vortexing in the presence of glass beads and 50 mM Tris-HCl (pH 7.5), 50 mM NaCl, 2 mM MgCl₂, 1 mM β-mercaptoethanol, 0.1% NP40, and complete protease inhibitors EDTA free (Roche) and were incubated at 4°C for 1 hour with immunoglobulin G (IgG)-Sepharose beads. The beads were then washed twice with 40 volumes of the extraction buffer and twice with 40 volumes of the extraction buffer containing 1 M NaCl. The proteins bound to the IgG-Sepharose were recovered by boiling in sample buffer.
- 22. A. van Hoof, P. A. Frischmeyer, H. C. Dietz, R. Parker, data not shown.
- 23. C. Allmang et al., Genes Dev. 13, 2148 (1999).
- K. T. Burkard, J. S. Butler, Mol. Cell. Biol. 20, 604 (2000).
- 25. R. Pulak, P. Anderson, Genes Dev. 7, 1885 (1993).
- 26. C.-Y. Chen et al., Cell 107, 455 (2001).
- D. Muhlrad, C. J. Decker, R. Parker, Mol. Cell. Biol. 15, 2145 (1995).
- 28. M. S. Longtine et al., Yeast 14, 953 (1998).
- 29. P. Mitchell, E. Petfalski, D. Tollervey, *Genes Dev.* **10**, 502 (1996).
- 30. R. S. Sikorski, P. Hieter, Genetics 122, 19 (1989).
- 31. S. Mahadevan, T. R. Raghunand, S. Panicker, K. Struhl, *Gene* **190**, 69 (1997).
- We thank the members of the Parker laboratory for insightful comments and A. Sachs for antibodies to Lsm1p. Funded by the Howard Hughes Medical Institute.
 - 19 October 2001; accepted 31 January 2002

Brain to Plasma Amyloid-β Efflux: a Measure of Brain Amyloid Burden in a Mouse Model of Alzheimer's Disease

Ronald B. DeMattos, 1,2,3* Kelly R. Bales, 5* David J. Cummins, 5

Steven M. Paul, 5,6† David M. Holtzman 1,2,3,4†

The deposition of amyloid- β (A β) peptides into amyloid plaques precedes the cognitive dysfunction of Alzheimer's disease (AD) by years. Biomarkers indicative of brain amyloid burden could be useful for identifying individuals at high risk for developing AD. As in AD in humans, baseline plasma A β levels in a transgenic mouse model of AD did not correlate with brain amyloid burden. However, after peripheral administration of a monoclonal antibody to A β (m266), we observed a rapid increase in plasma A β and the magnitude of this increase was highly correlated with amyloid burden in the hippocampus and cortex. This method may be useful for quantifying brain amyloid burden in patients at risk for or those who have been diagnosed with AD.

Abundant evidence suggests that a key event in the pathogenesis of AD is the conversion of A β peptides from soluble to insoluble forms in the brain (1). This process is among the earliest pathological changes that characterizes AD, and is estimated to occur ~ 10 to 20 years before the appearance of the earliest

cognitive changes of the disease (2, 3). Whereas individuals with pre-clinical AD (i.e., cognitively normal individuals with plaque and tangle densities similar to those with AD) have no measurable neuronal loss in affected brain regions, individuals with even very mild cognitive impairment indica-

Supplementary web Figure 1. Ski7p alignment with several homologs. Homologs were identified by PSI-BLAST (http://www.ncbi.nlm.nih.gov/BLAST), and aligned using Clustalw 1.8 and box shade using the default settings. Residues in red are identical in 80% of the sequences, residues in blue are similar in 80% of the sequences. The consensus indicates residues similar in 80% of the sequences, with residues identical in all sequences capitalized. Genbank accession numbers and annotations are HsHBS1: NP_006611, *Homo Sapiens* HBS-like protein; MmHBS1: AAH10251, *Mus musculus* similar to HBS1; DmORF: AAF47584, *Drosophila melanogaster* hypothetical protein; ScEF1A: NP_009676, *Saccharomyces cerevisiae* EF1A; SpEF1A1: P50522, *Schizosaccharomyces pombe* EF1A; MtEF1A: NP_276188, *Methanothermobacter thermautotrophicus* EF1A; EcEFtu: 11514297,

Escherichia coli EFtu; TaEFtu: Q01698, Thermus aquaticus EFtu; BtEFtu: S62768, Bos taurus mitochondrial EFtu; CeORF1: T23393, Caenorhabditis elegans hypothetical protein; SpEF1A2: T40165, S. pombe EF1A; AtORF: T49975, Arabidopsis thaliana hypothetical protein; GieRF3: AAF74405, Giardia intestinalis eRF3; SceRF3: NP_010457 S. cerevisiae eRF3; CeORF2: T23102 C. elegans hypothetical protein; TveRF3: AAF74406, Trichomonas vaginalis eRF3; ScHBS1: P32769, S. cerevisiae Hbs1 protein; ScSKI7: NP_014719 S. cerevisiae Ski7 protein; SpORF: T39242, S. pombe hypothetical protein

Supplementary web Figure 2. The C-terminal domain of Ski7p is not required for exosome-mediated degradation of normal mRNAs. **A.** Ski7p consists of a nonconserved N-terminal part and a C-terminal part that is similar to GTPases that function in translation. Amino acids 18 through 239 or from 265 to the C-terminus were deleted from Ski7p. **B.** The N-terminal domain of Ski7p is required and sufficient for viability in a strain defective for decapping. Alleles of SKI7 were combined with a dcp1-2 mutation. The dcp1-2 strain degrades mRNA by decapping at 23°C, but is incapable of decapping at temperatures of 30°C and higher and therefore requires exosome-mediated mRNA decay for viability. The indicated strains were streaked and then incubated at 23°C or 33°C. **C.** The N-terminal domain of Ski7p is required and sufficient for exosome-mediated degradation of MFA2pG. Half-lives were determined after a one hour incubation at 37°C as previously described (A. van Hoof, R. R. Staples, R. E. Baker and R. Parker, *Mol. Cell. Biol.* **20**, 8230 (2000)).

```
258 KQLLNLVVIGHVDAGKSTLMGHMLYLLGNINKRTMHKYEQESKKAGK-ASFAYAWVLDETGEERERGVTMDVGMTKFETTT----KVITLMDAPGHKDFIPNMITG---AAQADVAVLVVDASRGEFEAGFE-TGGQTREHGLLVRSLG-----VTQLAVAVNKMDQVNWQ--QERFQEITGKL
         174 KQLLNLVVIGHVDAGKSTLMGHMLYLLGNVNKRTMHKYEQESKKAGK-ASFAYAWVLDETGEERERGVTMDVGMTKFETTT----KVITLMDAPGHKDFIPNMITG---AQADVAVLVVDASRGEFEAGFE-TGGQTREHGLLVRSLG-----VTQLAVAVNKMDQVNWQ--QERFQEITGKL
          236 KSHIHMIVIGHVDAGKSTLMGHLLVDTGNVSQRVMHKHEQESKKLGK-QSFMYAWVLDETGEERARGITMDVGQSRIETKT----KIVTLLDAPGHKDFIPNMISG---ATQADVALLVVDATRGEFESGFE-LGGQTREHAILVRSLG-----VNQLGVVINKLDTVGWS--QDRFTEIVTKL
DmORF
SCEE1A
           4 KSHINVVVIGHVDSGKSTTTGHLIYKCGGIDKRTIEKFEKEAAELGK-GSFKYAWVLDKLKAERERGITIDIALWKFETPK----YOVTVIDAPGHRDFIKNMITG---TSOADCAILIIAGGVGEFEAGIS-KDGOTREHALLAFTLG-----VROLIVAVNKMDSVKWD--ESRFOEIVKET
           4 KGHINVVVIGHVDSGKSTTTGHLIYKCGGIDKRTIEKFEKEATELGK-GSFKYAWVLDKLKAERERGITIDIALWKFETPK----YNVTVIDAPGHRDFIKNMITG---TSQADCAVLIIGGGTGEFEAGIS-KDGQTREHALRAYTLG-----VKQLIVAVNKMDTTGWS--QARFEEIVKET
SpEF1A1
            4 KEHMNLAFIGHVDHGKSTLVGHLLLOAG------AIAEOOLAE-GEDKFRFVMDRLSEERERGVTIDLAHAKFETDK----YEFTIVDCPGHRDFVKNMTTG---ASOADAAVLVVAVDDGVMP-------OTKEHVFLSRTLG-----INOLIVAINKMDLVNYD--EEKFNALKDEV
MtEF1A
            8 KPHVNVGTIGHVDHGKTTLTAAITTVLAKTYG------GARAFDQIDNAPEEKARGITINTSHVEYDTPT---RHYAHVDCPGHADYVKNMTTG--AAOMDGAILVVAATDGPMP-------OTREHILLGROVG-----VPYIIVFLNKCDMVDD---EELLELVEMEV
EcEFtu
           9 KPHVNVGTIGHVDHGKTTLTAALTYVAAAENPN------VEVKDYGDIDKAPEERARGITINTAHVEYETAK---RHYSHVDCPGHADYIKNMITG--AAOMDGAILVVSAADGPMP-----OTREHILLAROVG----VPYIVVFMNKVDMVDD---PELLDLVEMEV
TaEFtu
BtEFtu
           55 KPHVNVGTIGHVDHGKTTLTAAITKILAEGGG-----AKFKKYEEIDNAPEERARGITINAAHVEYSTAA---RHYAHTDCPGHADYVKNMITG--TAPLDGCILVVAANDGPMP-------QTREHLLLARQIG-----VEHVVVYVNKADAVQD--SEMVELVELEI
          496 KDLINLIVVGHVDAGKSTLMGHLLHDLEVVDSRTIDKFKHEAARNGK-ASFAYAWVLDETEEERERGVTMDIGRTSFETSH----RIVLLDAPGHKDFISNMITG---TSQADAAILVVNATTGEFETGFE-NGGOTKEHALLLRSLG-----VTQLIVAVNKLDTVDWS--QDRFDEIKNNL
CeORF1
          175 KPVVHLVVTGHVDSGKSTMLGRIMFELGEINSRSMQKLHNEAANSGK-GSFSYAWLLDTTEEERARGVTMDVASTTFESDK----KIYEIGDAPGHRDFISGMIAG---ASSADFAVLVVDSSQNNFERGFL-ENGQTREHAYLLRALG-----ISEIVVSVNKLDLMSWS--EDRFQEIKNIV
SpEF1A2
         374 LSOLNLAIVGHVDSGKSTLSGRLLHLLGRISOKOMHKYEKEAKLOGK-GSFAYAWALDESAEBRERGITMTVAVAYFNSKR----HHVVLLDSPGHKDFVPNMIAG---ATOADAAILVIDASVGAFEAGFDNLKGOTREHARVLRGFG-----VEOVIVAINKMDIVGYS--KERFDLIKOHV
AtORE
GieRF3
           13 RKNLNIVFIGHVDAGKSTISGHLVSDLGKLDKRQLEKLEQQAKALNR-ESWKYAFAMDTSEEEREKGKTVECARESFLTPNG---RRITIIDAPGHKGFVHNMISG---AAQADTAILVISARKGEFESGFE-RGGQTSEHALLAYVNG-----IKQIVCLINKMDDITVEYCKKRYDSIVSQL
          258 KDHVSLIFMGHVDAGKSTMGGNLLYLTGSVDKRTIEKYEREAKDAGR-QGWYLSWVMDTNKEERNDGKTIEVGKAYFETEK----RRYTILDAPGHKMYVSEMIGG---ASQADVGVLVISARKGEYETGFE-RGGQTREHALLAKTQG-----VNKMVVVVNKMDDPTVNWSKERYDQCVSNV
CeORF2
          107 KEHINMVFVGHVDAGKSTIGGQLMFLTGMVDKRTLEKYEREAKEKGR-ESWYLSWCMDTNDEEREKGKTVEVGRAYFETEK----RHFTILDAPGHKSFVPNMIVG---ANQADLAVLVISARRGEFETGFD-RGGOTREHSMLVKTAG-----VKHLVILVNKMDDPTVKWEEERFKEIEGKL
          160 KKHFNIVFIGHVDAGKSTLCGHVLYQAGCVDQRTIEQYQAESAKEGR-GSWYFSWVMDLSKEERSKGKTEEVGVAHFETAQ----NKYTILDAPGHRSYVPQMIGG---AVQADVAVLVISARNGEFEAGFE-NGGOTSEHLLIARTAG-----VREIIIVVNKMDDPTVKWSKERFDQIVTKF
TveRF3
          165 LPHLSFVVLGHVDAGKSTLMGRLLYDLNIVNQSQLRKLQRESETMGK-SSFKFAWIMDQTNEERERGVTVSICTSHFSTHR----ANFTIVDAPGHRDFVPNAIMG--ISQADMAILCVDCSTNAFESGFDLDG-QTKEHMLLASSLG-----IHNLIIAMNKMDNVDWS--QQRFEEIKSKL
ScHbs1
         265 PLNLTCLFLGDTNAGKSTLLGHLLYDLNEISMSSMRELOKKSSNLDPSSSNSFKVILDNTKTERENGFSMFKKVIOVENDLLPPSSTLTLIDTPGSIKYFNKETLNSILTFDPEVYVLVIDCNYDSWEKSLDGPNNOIYEILKVISYLNKNSACKKHLIILLNKADLISWD--KHRLEMIOSEL
         278 KPRTKLLLLGPPKSGKKTLLSRLFFQIGSFDPKTMQKCTVLN--AK--KESLSSVLKSTKTKWYDFETFSNSYSSTIIDFP--LGIFTTNASSRDNFLKHSSLF---QVMNTAIFTIDCLNPLEG-----LDGISSILQLMNGLS----ISSYMFAITKMDEIEWD--ENKFINLVNSI
             consensus
              . . . . . 10 . . . . 20 . . . . 30 . . . . 40 . . . . 50 . . . . 60 . . . . 70 . . . . 80 . . . . 90 . . . . . 100 . . . . 110 . . . . . 120 . . . . 130 . . . . 140 . . . . 150 . . . . 160 . . . . . 170 . . . . . 180 . . .
         HSHRS1
         MmHBS1
          S-NFIKKVGFNPKTVPFVPVSGFOGDNMIEPT-----TNMPWYOGWOKETKAGVVKGKTLLEAIDSIEP-PARPTDKPLRLPLODVYKI----GGIGTVPVGRVETGVIKPGMIVTFAPAG---------VTTEVKSVEMHHESLDAGLPGDNVGFNVKN------VSVKDIR
SpEF1A1
         MtEF1A
          ECEFtu
          R-DLLNOYEFPGDEVPVIRGSALLALEEMHKN-------PKTKRGENEWVDKIW-------ELLDAIDEYIPTPVRDVDKPFLMPVEDVFTI---TGRGTVATGRIERGKVKVGDEVEIVGLAP------ETRKTVVTGVEMHRKTLOEGIAGDNVGLLLRG------VSREEVE
TaEFtu
BtEFtu
         R-ELLTEFGYKGEETPIIVGSALCALEO--------KNIRTVVTGIEMFHKSLDRAEAGDNIGALVRG---------LKREDLR
          CeORF1
         SDFLIKMVGFKTSNVHFVPISAISGTNLIQKD----SS--DLYKWYKG-----PTLSALDQLVP-PEKPYRKPLRLSIDDVYR----SPRSVTVTGRVEAGNVQVNQVLYDVSSQ---------EDAYVKNVIRNSDPSSTWAVAGDTVTLQLAD------IEVNQLR
SpEF1A2
Atorr
          G-SFLQSCRFKDSSLTWIPLSAMENQNLVAAP------SD-NRLSSWYQG-------PCLLDAVDSVKS-PDRDVSKPLLMPICDAVRST---SQGQVSACGKLEAGAVRPGSKVMVMPSG--------DQGTIRSLERDSQACTIARAGDNVALALQG-----IDANQVM
GieRF3
          K-LYLENVGYASKNIFFLPISGFTGENLISTK------ELN-PKLSEWYSG-------PSFLDLLDELKV-PKRDTKSPLCACVSGHYKE-----NSAFIVVKVEQGKMSIGDTVLCLPENK------AFDILDIEGEQEEGSKLVQAKAGDNVRIKVKDD------IWEFLTE
         SCERF3
          CeORF2
          ScHbs1
         L-PYLVDIGFFENNINWVPISGFSGEGVVKIE------YT-DEVROWYNGPN------LMSTLENAAFKISKENEGINKDDPFLFSVLEIIPSK-KTSNDLALVSGKLESGSIOPGESLTIYPSEOSCIVD----KIOVGSOOGOSTNHEETDVAIKGDFVTLKLRK-------AYPEDIO
          NYVLKENFOWTDAEFOFIPCSGLLGSNLNKTENITKSKYKSEFDSINYVPEWYEG------PTFFSOLYLLVEHNMNKIETTLEEPFVGTILOSSVLOPIAEINYVSLKVLINSGY1OSGOTIEIHTOYEDFHYYGIVSRWKNSKOILETNTKNNISVGLNPDILEVLVKIHNTEDFTKKOFHIR
         Q-SFLKESCGIIEKSKFIPISGLKGTNLTSIS-------QEKLSQWYKS-------DTLGKIDKEADTNHGTWNFLLNLPLSLTISHITPLPENQSHIYCSIHSGMLQDSQKLYVGTGR-------LETQITGLSSDENPKGFNVAGDMIQAKIPT-------LPNLC
SpORF
consensus...l...f....fip.sq.q.l.....wy.....wy.....ll.id....r...p.l.v....qkie.G.i.q.l......a..qd.v..l.....m...i.
          200 - 200 - 210 - 220 - 230 - 240 - 250 - 260 - 270 - 280 - 290 - 300 - 310 - 320 - 330 - 340 - 350 - 360 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 - 370 
         VGCIFCGPKV----PIKACTRFRARILIFNIEI--PITKGFPVL-----LHYQTV-----SEPAVIKRLISVLNKSTGEVTKKKPKFLTKGQNALVELQTQRPIALELYKDFK---ELGRFMLRYGGSTIAAGVVTEIKE
HsHBS1
          MmHBS1
          VGCIISDPQT----PIPVTTRFQARIIVFNVKV--PITMGFPVL------LHHQSL-------IEPAVVCKLTASIHKSTGEVVKKKPRCLGNNSCALVELETGRPTCIERYADFK--ELGRVMLRVAGVTIAAGMVTKIR
          RGNVCGDAKND---PPKGCASFNATVIVLNHPG--0ISAGYSPV------LDCHTA-------HIACRFDELLEKNDRRSGKKLEDHPKFLKSGDAALVKFVPSKPMCVEAFSEYP---PLGRFAVRDMROTVAVGVIKSVDKTEK-AAKVTKAAOKAAKK
SCEF1A
         SpEF1A1
         Eceftu
          TaEFtu
          BtEFtn
          CeORF1
         PGDILSNYEN---PVRRVRSFVAEIQTFDIHG--PILSG-----S-----TLVLHLGRTVTSVSLKIVTVNNKRSRHIASRKRALVRISFLDGLFPLCLAEECP--ALGRFILRRSGDTVAAGIVKELC
SpEF1A2
          AGDVLCHPDFP---VSVATHLELMVLVLEGAT--PILLG-----SQLEFHVHHA------KEAATVVKLVAMLDPKTGQPTKKSPRCLTAKQSAMLEVSLQNPVCVETFSESR--ALGRVFLRSSGRTVAMGKVTRIIQDS
AtORF
         GieRF3
          PGFVLTSPKN---PIKSVTKFVAQIAIVELKS-IIAAGFS-----CVMHVHTA-----IEEVHIVKLLHKLEKGTNRKSKKPPAFAKKGMKVIAVLETEAPVCVETYQDYP--QLGRFTLRDQGTTIAIGKIVKIAE
          GGFIICSPDS----LAKTGRVFDAEVLVLEHRS--IIASGYSCVLHIQSAVEEVTVKVCLLRPEELGGCPIGGITNSGDWPRPQNRGLQYAHFCNCRTGVIATIDKKTGEKKRAKFVKQDEKCIMRLESPEPFVLEPFKEYP---YLGRFTLRDEGKTIAIGKVLKVVE
         SGSVICPVNAP----CDVAOKVIAKIRIVPSGPE-LITAG-----------YEAMCHIH----------TETVPVSVEKLRTLOIPGRELEKNPRFIKRGCLAEVILKFDHPICVEVAKDFP---OLGRFIIRKEGFTTIVGLVDKLPKATGK
TVERF3
         NGDLAASVDYS---SIHSACFVLELTTFDMNR--PLLPGTP------FILFIG------VKEQPARIKRLISFIDKGNTASKKKIRHLGSKQRAFVEIELIEVKRWIPLLTAHENDRLGRVVLRKDGRTIAAGKISEITQ
ScHbs1
          KGDIIIHSRKTN--TLSPNLPNTLKLLALRLIKLSIOTHALS-------DPVDLG-------SELLLYHNLTHNAVKLVKILGTNDISINPNOSLIVEVEIIEPDFALNVIDS----KYITNNIVLTSIDHKVIAVGRIACO
          consensus q.i....i.q....l..v.l..pi.le...r.lr..tia.q.v..i.
           \underbrace{\phantom{0} \dots 380 \dots 390 \dots 400 \dots 410 \dots 420 \dots 420 \dots 430 \dots 440 \dots 450 \dots 460 \dots 470 \dots 480 \dots 490 \dots 500 \dots 510 \dots 520 \dots 530 \dots 540 \dots 550 \dots 550 \dots 559 }_{\phantom{0} \dots 100 \dots 100
```

