DOI: 10.1534/genetics.109.110189

Caenorhabditis elegans unc-82 Encodes a Serine/Threonine Kinase Important for Myosin Filament Organization in Muscle During Growth

Pamela E. Hoppe,*,†,[1] Johnnie Chau,*,[2,3] Kelly A. Flanagan,*,[2,4] April R. Reedy†,[2,5] and Lawrence A. Schriefer*

†*Department of Biological Sciences, Western Michigan University, Kalamazoo, Michigan 49008-5410 and* **Department of Genetics, Washington University School of Medicine, St. Louis, Missouri 63110*

Manuscript received September 23, 2009
Accepted for publication November 1, 2009

ABSTRACT

Mutations in the *unc-82* locus of *Caenorhabditis elegans* were previously identified by screening for disrupted muscle cytoskeleton in otherwise apparently normal mutagenized animals. Here we demonstrate that the locus encodes a serine/threonine kinase orthologous to human ARK5/SNARK (NUAK1/NUAK2) and related to the PAR-1 and SNF1/AMP-Activated kinase (AMPK) families. The predicted 1600-amino-acid polypeptide contains an N-terminal catalytic domain and noncomplex repetitive sequence in the remainder of the molecule. Phenotypic analyses indicate that *unc-82* is required for maintaining the organization of myosin filaments and internal components of the M-line during cell-shape changes. Mutants exhibit normal patterning of cytoskeletal elements during early embryogenesis. Defects in localization of thick filament and M-line components arise during embryonic elongation and become progressively more severe as development proceeds. The phenotype is independent of contractile activity, consistent with *unc-82* mutations preventing proper cytoskeletal reorganization during growth, rather than undermining structural integrity of the M-line. This is the first report establishing a role for the UNC-82/ARK5/SNARK kinases in normal development. We propose that activation of UNC-82 kinase during cell elongation regulates thick filament attachment or growth, perhaps through phosphorylation of myosin and paramyosin. We speculate that regulation of myosin is an ancestral characteristic of kinases in this region of the kinome.

THE contractile apparatus of striated muscle is a highly ordered cytoskeletal structure (Figure 1) composed of actin and myosin filaments, the filament anchoring structures, and a host of regulatory proteins. During *Caenorhabditis elegans* embryogenesis, the body-wall muscle cells polarize and assemble their cytoskeletons in response to contact with the epidermal cells, to which they attach through focal-adhesion-like structures. The epidermal cells respond in a similar fashion and assemble attachment structures and fibrous organelles at the sites of muscle-cell contact (reviewed in Moerman and Williams 2006). The coordination of the cytoskeletons of the two tissue types provides the physical attachment that transmits the force of muscle-cell contraction to the epidermis and its secreted cuticle and allows the worm to locomote through its environment. The patterning of the contractile apparatus occurs through integrin-mediated signaling at the plasma membrane where muscle cells contact the epidermis. The assembly of more interior (membrane-distal) components of the contractile apparatus follows and requires the membrane-proximal events (Hresko *et al.* 1994). Failure to assemble functional epidermal–muscle-cell contacts or failure to make contractile muscle cells prevents elongation of the embryo from an egg shape into a long tube. Many genes required for these early patterning events, as well as those essential for muscle contraction, have been identified by screening for embryonic lethal mutations that produce the Pat phenotype (**p**aralyzed, **a**rrested elongation at **t**wo-fold) (Williams and Waterston 1994).

However, proteins that act subsequent to the early patterning events or are not essential for contraction would not have been identified in the Pat screens. Mutations in the *unc-82* gene were isolated by screening apparently normal animals for muscle-cell disorganization using polarized light microscopy (Waterston *et al.* 1980). Animals homozygous for *unc-82* mutations exhibit patchy, bright birefringence rather than the

Supporting information is available online at http://www.genetics.org/cgi/content/full/genetics.109.110189/DC1.

[1]*Corresponding author:* Department of Biological Sciences, Western Michigan University, Wood Hall, 1903 W. Michigan Ave., Kalamazoo, MI 49008-5410. E-mail: pamela.hoppe@wmich.edu

[2]These authors contributed equally to this work.

[3]*Present address:* Department of Genetics, Case Western Reserve University, Cleveland, OH 44106-4955.

[4]*Present address:* Department of Cell Biology, The Scripps Research Institute, La Jolla, CA 92037.

[5]*Present address:* Université Claude Bernard Lyon I, Centre de Génétique Moléculaire et Cellulaire, 69622 Villeurbanne Cedex, France.

uniform bright bands of signal that mark the areas of organized myosin-containing thick filaments in wild-type worms. To define the mechanisms underlying filament organization within the contractile apparatus, we undertook molecular and phenotypic analyses of *unc-82* mutants. Our data suggest that UNC-82 is a kinase, orthologous to human ARK5 and SNARK, that is required specifically for myosin filament reorganization during cellular elongation in normal development.

MATERIALS AND METHODS

Nematode strains: We used the following nematode strains: CB1220 *unc-82(e1220)* IV; CB1323 *unc-82(e1323)* IV; RW3536 *unc-82(e1323) unc-24(e138)* IV; RW1350 *unc-44(e362) unc-82(e1323)/stDf7* IV; PZ51 *unc-54(s95)* I; *unc-82(e1323)* IV, PZ52 *unc-54(s95)* I; *unc-82(e1220)* IV; CB4856 Hawaiian; transgenic lines of *unc-82(e1323)* rescued by cosmid B0496, RW3918, RW3919, RW3920, RW3921; and transgenic line expressing UNC-82::GFP, PZ73 *unc-82(e1220)* IV; *phEx22*.

Antibody staining: Embryos were fixed with paraformaldehyde and methanol and stained using the methods of Hresko *et al.* (1994). Adults used in the *z*-series analysis (Figure 5, D–F) were processed using peroxide tube fixation (Duerr 2006). Images were collected over a depth of 3–4 μm at 0.2- to 0.3-μm intervals using a Leica DM5500 microscope and ImagePro 6.0 software. All other images show adults fragmented with a French press cell, extracted with detergent, fixed with methanol, and stained (Francis and Waterston 1985).

Time-lapse video recording: Embryos were mounted on slides and recorded using a modified version of the method described in Williams and Waterston (1994). Gravid adult hermaphrodites were cut in half with a razor blade in M9 buffer. The eggs in a small amount of buffer were transferred to a 2% agarose pad on a microscope slide and then covered and gently flattened with a Vaseline-lined coverslip. The developing embryos were videotaped overnight using Nomarski optics and a time-lapse VHS recorder.

Single nucleotide polymorphism mapping: Using the single nucleotide polymorphism (SNP) method of Jakubowski and Kornfeld (1999), N2 Bristol animals homozygous for the linked mutations *unc-82(e1323)* and *unc-24(e138)* were crossed to males of the Hawaiian strain CB4856, and single F_1 hermaphrodite cross progeny were picked to individual plates. F_2 recombinant animals that were homozygous for either *unc*-82 or *unc-24*, but not for both, were picked singly, and F_3 worms homozygous for the recombinant chromosome were isolated. SNPs within this interval were chosen from those identified and described by Wicks *et al.* (2001): F38A5-19715, D2024-25027, T12B3-3235, C48A7-11713, B0496-27999, F55G1-23412, and T09A12-14845. PCR fragments from 400 to 800 bp in length were amplifed from each recombinant strain and analyzed by restriction digest.

RNA interference: Double-stranded RNA (dsRNA) fragments ~500–1000 bp in length were made using the Megascript kit (Ambion). The dsRNA was injected into the gonads of L4 hermaphrodites at 1 μg/ml. Injected animals were maintained at 20° or 25°and transferred daily to fresh plates. Progeny were scored by polarized light microscopy. Primers used were aataatacgactcactatagggagaTACTCTAGCGGTGGAGAATT, aaatttaggtgacactatagaagagagCAGACTTCATCTCTTCCG, aataatacgactcactatagggagaACGGGCTGAAAGAGATGCTG, and aataatacgactcactatagggagaTCGAACTCCATTGCTTG. Lowercase letters denote residues that do not match *C. elegans* wild-type sequence.

DNA constructs and transgenic worm strains: Cosmid rescue was obtained by injecting *unc-82(e1323)* animals with a 210 ng/μl DNA cocktail containing pPHgfp1 (Hoppe and Waterston 2000), Bluescript, and B0496 in a 5:15:1 ratio. Transgenic lines were marked by GFP expression in the hypodermis and scored for rescue using polarized light microscopy.

A full-length UNC-82::GFP fusion construct was generated by recombination *in vivo* between a 17-kb PCR fragment and a plasmid encoding the C terminus of UNC-82 fused to GFP following the method of Yuan *et al.* (2000). Oligonucleotides GTCTCTGCTAAACAGCAATCG and GTTTGTGTACTTGTTGTGTGTG were used with cosmid template B0496 and the Expand Long Template PCR System (Roche) to amplify a genomic fragment beginning 2.6 kb upstream of the UNC-82 initiator methionine and terminating within intron 27. To fuse the C terminus of UNC-82 to GFP, primers gacacaagctTCGTTTCCGTCCAACTGCTCG and ccccggatccccATAAATATTTGGATCATCAT were used to amplify a 2-kb genomic fragment spanning exons 24–30, which was cut with *Hin*dIII and *Bam*HI and cloned into pPD95_67. Prior to injection, the plasmid was cut with *Hin*dIII, extracted with phenol/chloroform, and ethanol precipitated.

Sequencing alleles: Genomic DNA was isolated from worms homozygous for each *unc-82* mutant allele, and two overlapping PCR fragments spanning the locus were amplified using the Expand Long Template PCR system (Roche). A library was generated from each of the four PCR fragments using a protocol from the Washington University Genome Sequencing Center (http://genome.wustl.edu/tools/protocols/). Briefly, the PCR fragments were sonicated and treated with mung bean nuclease. The smaller fragments were ligated into pZERO-2 vector and transformed into DH10b cells. The sequence was assembled and analyzed using the Phred/Phrap program.

cDNA analysis: The following cDNA clones were obtained from Yuji Kohara for sequencing: yk76b5, yk159d1, yk286b1, yk360b4, yk896c08, yk1121c09, yk1232g01, yk1305a10, yk315d2, yk47c5, yk356a4, yk405f4, and yk8c7. Inserts from phage clones were amplified using the Expand Long Template PCR System (Roche) with vector primers GGTTTTCCCAGTCACGACGTTG and CAGGAAACAGCTATGACCATGATT. The PCR products were gel purified, and the DNA was recovered by using either a phenol extraction protocol or GenElute Minus EtBr Spin Columns (Sigma-Aldrich). DNA from plasmid clones was isolated using Wizard miniprep kits (Promega).

A cDNA fragment containing the 5′-end of the *unc-82* mRNA was generated from total adult *C. elegans* RNA (a gift from Ziva Misulovin) using the gene-specific primer exon6RevBam CTGTGGATCCAGACTTCATCTCTTCCG and SuperScript II RT (Invitrogen) for first-strand synthesis and primers exon6-RevBam and SL1HindDIII ACAGGAAGC TTCGGTTTAATTACCCAAGTTTGAG for PCR using Biolase DNA Polymerase (Bioline). The fragment was gel purified prior to DNA sequencing with ABI Big Dye Terminators version 3.1 and run on an ABI Prism 3100 Genetic Analyzer.

Sequence comparisons: The pairwise alignment of the genomic DNA sequences from *C. elegans* and *C. briggsae* was generated and annotated using BioEdit Sequence Alignment Editor. Homologous protein sequences were identified by BLAST and aligned using CLUSTALX. Repetitive elements in UNC-82 and its orthologs were detected using Radar (http://www.ebi.ac.uk/Radar/) and DotPlot.

RESULTS

***unc-82* is required for thick filament organization during embryonic elongation:** To further characterize the role of *unc-82* in muscle, the phenotype of *e1323*

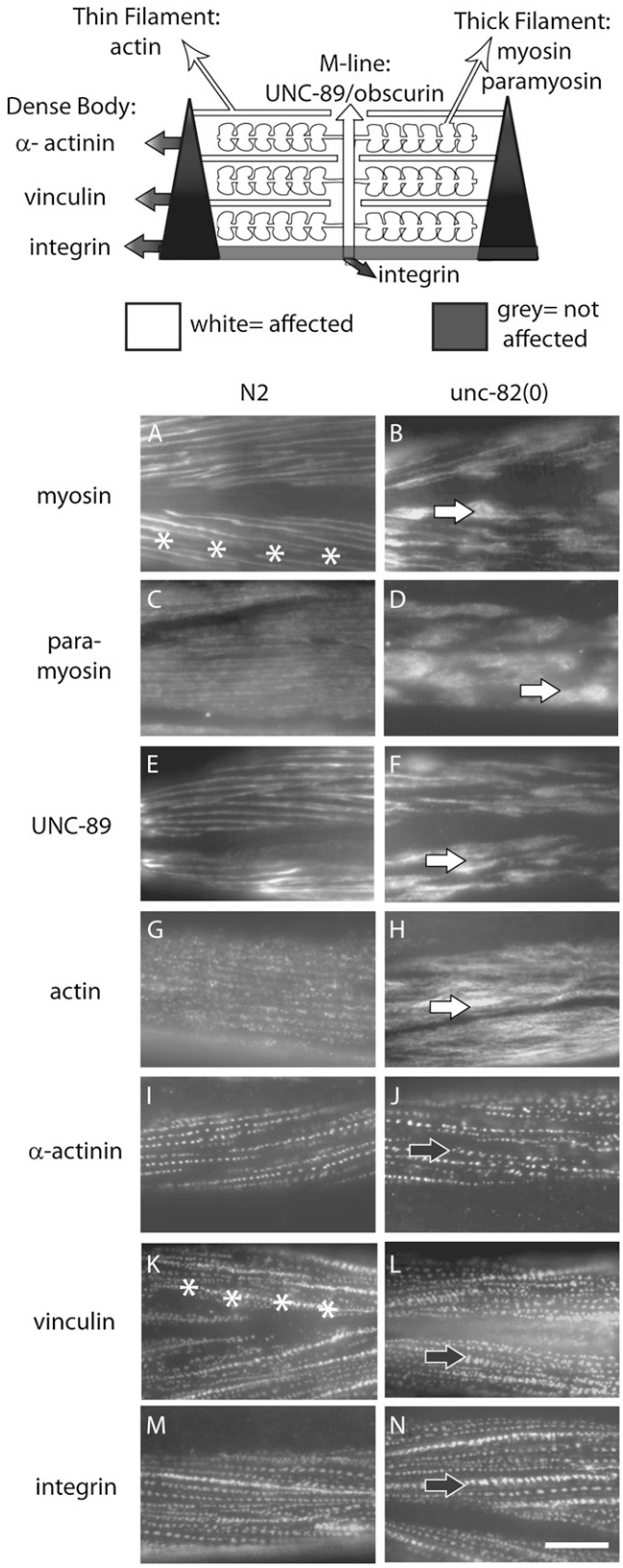


FIGURE 1.—*unc-82* mutants show dramatic defects in localization of thick-filament and M-line components, but normal patterning of membrane and dense-body proteins. A diagram of the sarcomere (top) is highlighted to indicate those components affected in *unc-82* mutants. Structures represented include the actin filaments anchored to the dense body (the Z-line analog) and myosin-containing thick filaments associated with the M-line. The components represented in white exhibit abnormal staining patterns in *unc-82* mutants; those represented in gray are relatively unaffected. (A–N) Adult fragments from wild-type (left column) and *unc-82* mutant worms (right column) were stained with antibodies specific for components of the contractile apparatus. Thick-filament proteins myosin A (A and B) and paramyosin (C and D) are grossly mislocalized in *unc-82* mutants, as is the M-line component UNC-89/obscurin (E and F). White arrows (B, D, F, and H) indicate abnormal accumulations of thick-filament and M-line proteins, and asterisks (A and K) mark a cell border. Actin staining (G and H) is mildly disrupted, but does not appear in large clumps. The distribution of α-actinin, vinculin, and integrin (I–N) (organized lines of puncta, solid arrows) is similar in mutant and wild type. Antibodies: myosin A, 5.6; paramyosin, 5.23; UNC-89, EU30; actin, C4; α-actinin, MH35; vinculin, MH24; integrin MH25. Bar, 10 μm.

mutant adults was examined using antibodies specific for various components of the contractile apparatus. The staining experiments revealed profound defects in the distribution of the thick filament proteins myosin and paramyosin (Figure 1, A–D), consistent with the structural defects observed by transmission electron microscopy (TEM; WATERSTON *et al.* 1980). In addition, the morphology of the thick-filament-attachment structures, assessed by staining for the M-line component UNC-89/obscurin (BENIAN *et al.* 1996; SMALL *et al.* 2004), was severely disrupted (Figure 1, E and F). In contrast, vinculin and α-actinin, which are components of the thin-filament-attachment structure called the dense body, are comparatively unaffected, as is the distribution of integrin, which is found in the muscle-cell membrane at the base of both M-lines and dense bodies (Figure 1, I–N). These results suggest that *unc-82* activity is required to organize internal proteins of the M-line and of the thick filament, but is not involved in overall patterning of the contractile apparatus, which occurs at the membrane. The distribution of actin staining is also notably altered in *unc-82* adults (Figure 1H), despite the relatively normal positioning of the structures anchoring the actin filaments. Given the physical interaction between actin and myosin, the actin phenotype is likely a secondary consequence of thick filament disorganization.

To distinguish between a role in the initial organization of the affected proteins and a role in the maintenance of organization during growth, wild-type and mutant embryos at various stages of development were examined to determine when abnormalites first appear in *unc-82* mutants. For all proteins examined, the antibody staining pattern in mutant embryos was indistinguishable from that of wild type up through the 1.5-fold stage (Figure 2, A–G), the time at which muscle contractions first occur. Furthermore, time-lapse video microscopy showed that the onset of muscle twitching (1.5-fold stage) and the progression to coordinated body movement (2-fold stage) is normal in *unc-82*

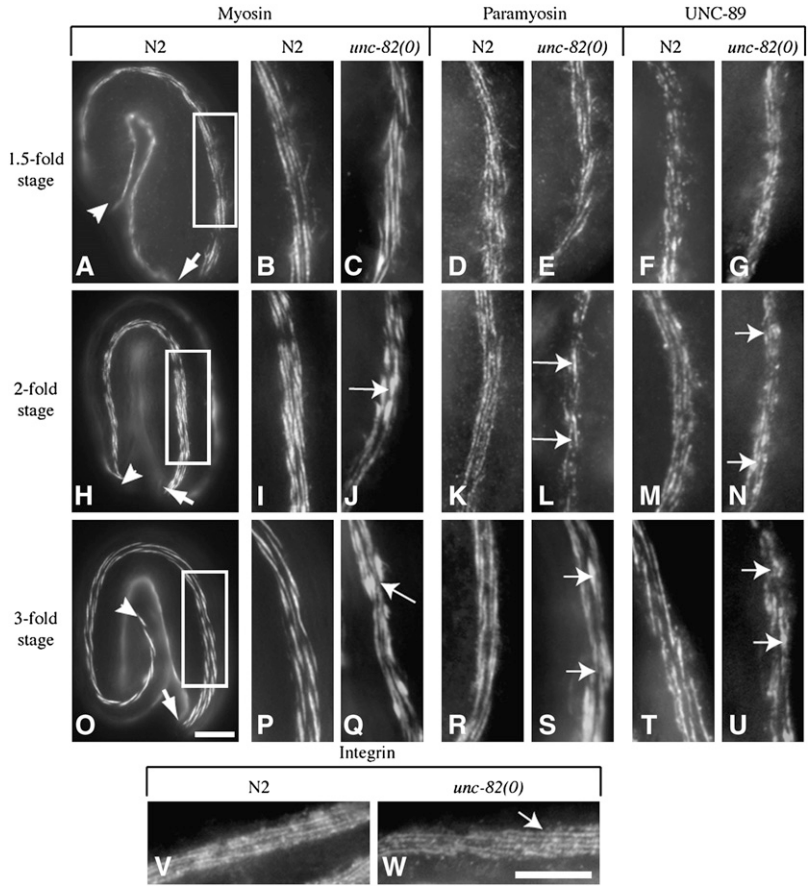


FIGURE 2.—Defects in localization of thick-filament and M-line proteins appear during embryonic elongation in *unc-82* mutants. Each row of micrographs shows a different stage in embryonic development. The first micrograph in each row shows a wild-type embryo at the stage and in the position and orientation of all embryos in that row. Arrowheads indicate the end of the elongating tail, and arrows mark the anterior tip of the head (A, H, and O). A single muscle quadrant is in focus. The boxed area (A, H, and O) indicates the portion of a dorsal muscle quadrant shown at higher magnification in the second column. Subsequent columns show an equivalent area of muscle from different animals. The staining pattern for each antibody is shown in wild-type and in *unc-82(e1323)* embryos. The *unc-82* embryos show normal localization of all proteins at the 1.5-fold stage (B–G). The staining patterns of thick-filament and M-line proteins become abnormal by the twofold stage of development (H–N), exhibiting aberrant blotches of protein (arrows in J, L, N, Q, S, U) that are larger and more numerous in 3-fold embryos (O–U). Integrin appears normal in *unc-82* at the 3-fold stage (V and W). Antibodies: myosin A, 5.6; paramyosin, 5.23; UNC-89, MH42; integrin, MH25. Bars, 10 μm.

mutant embryos. These data argue that *unc-82* is not required for the signaling events between the muscle and epidermal cells that establish the earliest pattern of the contractile apparatus.

Subsequent to the 1.5-fold stage, *C. elegans* embryos rapidly develop coordinated movement (WILLIAMS and WATERSTON 1994) and undergo body elongation. During this time, *unc-82* mutants begin to exhibit defects in localization of thick filament and M-line components. The first detectable defects, aberrant accumulations and gaps in myosin and paramyosin staining, are consistently seen at the 2-fold stage (Figure 2, H–N). By the 3-fold stage, animals exhibit more numerous and severe defects in thick-filament and M-line protein localization throughout the length of the muscle quadrants. The staining pattern of integrin, which marks the base of the M-lines and dense bodies, remains normal through at least the 3-fold stage (Figure 2W). These results suggest that *unc-82* activity is required to maintain proper thick filament and M-line organization during the stages when muscle cells are rapidly elongating and vigorously contracting.

To test whether the force of contraction is responsible for causing the observed cytoskeletal defects, contractile forces were diminished in an *unc-82* mutant background by constructing double-mutant strains homozygous for either *unc-82(e1323)* or *unc-82(e1220)* and *unc-54(s95)*. The *s95* allele, a point mutation in the motor domain of the major body-wall muscle myosin (DIBB *et al.* 1985), greatly reduces muscle contractility but does not alter sarcomere structure (MOERMAN *et al.* 1982). Examination of double-mutant animals by polarized light microscopy (Figure 3) revealed that reducing contraction altered the muscle-cell phenotype but did not restore wild-type structure. Like *unc-82* single mutants, double-mutant animals lacked organized striations and contained brightly birefringent material at the ends of the cells. However, the double mutants exhibited more uniform signal in the body of the muscle cell.

***unc-82* gene encodes a predicted serine/threonine kinase:** Previous three-factor mapping and deficiency complementation data had placed *unc-82* on chromosome IV to the right of *deb-1* and within nDf41 (WADDLE 1993). Using SNP mapping, we defined an ~400-kb region containing the *unc-82* gene (Figure 4A). Rescue of *unc-82(e1323)* was obtained with B0496, a cosmid within this interval, and subsequently with an ~18-kb PCR fragment containing genomic sequences that spanned the single predicted ORF called B0496.3 in Wormbase (http://www.wormbase.org/). The predicted protein product from this locus contains ~1600 amino acids. The single conserved motif is a predicted serine/threonine kinase catalytic domain near the N terminus.

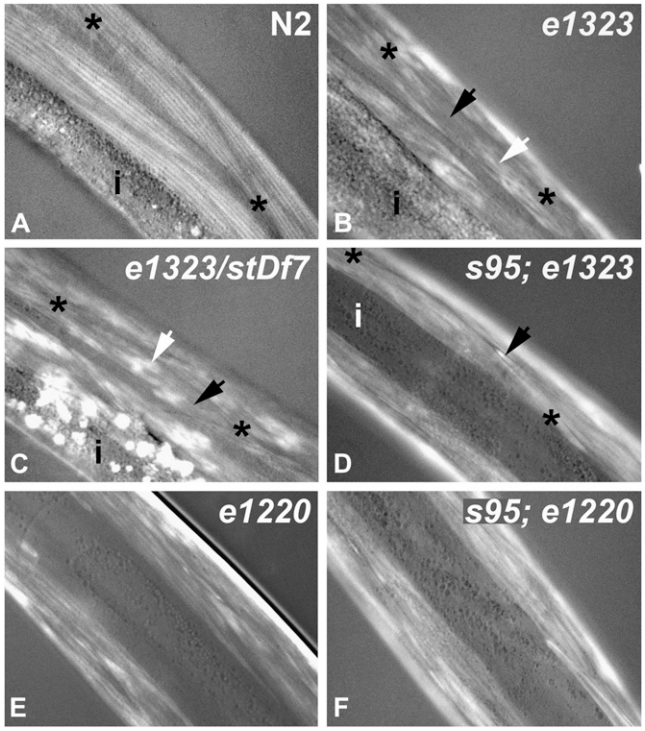


Figure 3.—Muscle-cell organization in young adult worms was assessed by polarized light microscopy. Asterisks mark ends of spindle-shaped muscle cells, and "i" marks the intestine. (A) In wild type, longitudinal bright birefringent bands mark the positions of highly ordered thick filaments. (B) Animals homozygous for the presumptive null allele *e1323* exhibit bright amorphous patches (white arrow) and regions of little signal (black arrow) within muscle cells. (C) The muscle-cell phenotype of *e1323/stDf7* hemizygotes is similar to that of the *e1323* homozgyote. There is an increase of birefringent material in the intestine. Compared to single mutants (B and E), double mutants homozygous for *unc-82* and the myosin mutation *unc-54(s95)* (D and F) exhibit relatively uniform signal in the muscle-cell body and brightly birefringent needles at the ends of cells (black arrow in D).

Several approaches were used to confirm the molecular identity of the *unc-82* gene. To test whether the ~18-kb rescuing fragment was likely to represent a single transcription unit, we used RNA interference to target sequences within either the 5′ or 3′ region of the predicted mRNA (Figure 4B). Injection of dsRNA derived from either region into wild-type worms phenocopied the cytoskeletal defects in *unc-82* mutants, suggesting that both targeted sequences were part of the *unc-82* gene. DNA sequencing revealed a molecular lesion within the predicted kinase domain in *unc-82* mutant strains. The *unc-82(e1220)* mutation is a missense allele that changes a glutamic acid to a lysine. This charge reversal occurs within the catalytic loop, which is involved in substrate binding (Hanks and Hunter 1995). The *unc-82(e1323)* mutation changes a glutamine to a stop codon, which terminates translation within the catalytic domain. Animals hemizygous for the *e1323* mutation over the chromosomal deficiency *stDf7*, which removes the *unc-82* gene, had a muscle phenotype similar to that of *e1323* homozygotes as young adult animals (Figure 3). However, older adult *e1323/stDf7* hemizygotes may contain larger, brighter patches of signal (not shown). Further, the hemizygotes are often sterile and have a distinct polarized light phenotype in the intestine (Figure 3).

The exon structure of the *unc-82* transcription unit was determined by sequencing several cDNA clones obtained from Yuji Kohara, as well as a 5′ PCR fragment. These data identified two previously undetected exons, including the true first exon, as well as a number of alternative splicing events. The *unc-82* gene contains 30 exons (Figure 4B), with the potential to produce a full-length mRNA containing 4803 coding bases. Alternatively spliced cDNAs (Figure 4C) suggest that several protein isoforms between 1300 and 1600 amino acids in length are generated. Because none of the cDNAs are full length, we do not know the precise exon content of any single isoform. Isolation of a single SL1-primed PCR product from cDNA using a reverse primer in exon 6 indicates that isoforms that contain exon 6 also contain the N-terminal kinase domain. No attempt was made to use reverse primers in other exons to identify alternative 5′-ends. Comparison of the *C. elegans* genomic sequence to that of *C. briggsae* revealed that the *unc-82* gene structure, including alternative exon borders, is conserved between the two species (see supporting information, Figure S1).

UNC-82 is located at or near the M-line: To determine the subcellular location of the UNC-82 protein *in vivo*, a full-length C-terminal GFP fusion construct (see materials and methods) was injected into *unc-82*mutant worms, and transgenic lines exhibiting improved motility and muscle-cell structure were isolated. Because the final exon is found in all known cDNAs (Figure 4C), we expected the inserted GFP tag to label all protein products from the locus. Expression of the fusion protein in *unc-82* mutants restored the wild-type pattern of UNC-89 (Figure 5A), demonstrating function *in vivo*. Further, UNC-82::GFP is localized near the M-line and present throughout the depth of the myofilament lattice (Figure 5, A–F). The distribution of UNC-82::GFP appears punctate by antibody staining in fragmented adults (Figure 5B), but as continuous lines when viewed by endogenous GFP fluorescence in living or fixed adults (Figure 5E). At the 1.5-fold stage, the antibody signal in transgenic embryos is punctate and concentrated in the region of the contractile apparatus (Figure 5, G–I). As development continues, the stain becomes organized into longitudinal lines of dots that roughly coincide with the striped myosin pattern (Figure 5, J and K).

Transgene expression is not limited to body-wall muscle. In embryos, anti-GFP produces signal in areas outside the muscle quadrants (Figure 5K). Interestingly, UNC-82::GFP is expressed in the pharynx but not

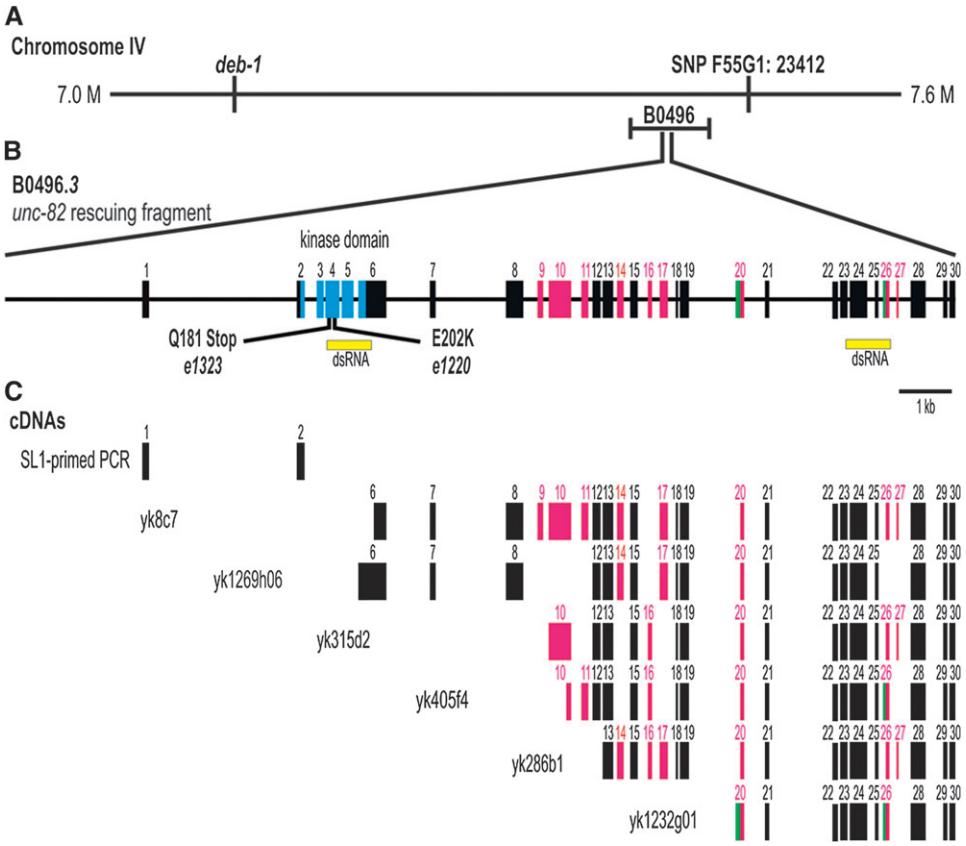


FIGURE 4.—The *unc-82* gene encodes a predicted serine/threonine kinase and produces multiple alternatively spliced messages. (A) Schematic of the 600-kb region containing *unc-82* as defined by SNP mapping, which placed *unc-82* to the left of marker F55G1: 23412. Within this interval, both (A) the cosmid B0496 and (B) an 18-kb PCR fragment containing B0496.3 rescue *unc-82(e1323)*. (B) The solid black line represents the introns and 2.6 kb of upstream sequence. Yellow rectangles indicate regions targeted in RNA interference experiments. Point mutations found in *e1323* and *e1220* are shown. (B and C) Rectangles represent exons: magenta denotes alternatively spliced exons, green marks exons with alternatively spliced borders, and blue shows the kinase domain. (C) The B0496.3 gene structure was determined by sequencing cDNAs and a 5′ PCR product. *unc-82* has at least nine alternatively included exons that can exist in numerous combinations. Bar, 1 kb.

detected in the regions occupied by thick filaments (Figure 5, L–N). No readily apparent defects in pharyngeal morphology or function were revealed by light microscopy, consistent with earlier observations (WATERSTON *et al.* 1980).

UNC-82 homologs are found in vertebrates and other invertebrates: BLAST searches using either the full-length UNC-82 protein sequence or only the kinase catalytic domain identified the same small set of high-scoring proteins, which includes the human proteins ARK5 and SNARK (NUAK1 and NUAK2), as well as anonymous proteins in many organisms. All proteins within the group contain a kinase domain near the N terminus, with no detectable conserved domains in the remainder of the protein. The sequences C-terminal to the kinase domain are noncomplex and repetitive and lack detectable homology among worms, flies, and humans. This 1225-amino-acid region of UNC-82 includes 33% charged residues (amino acids D, E, K, and R, or DEKR) and 12% serine. The C-terminal regions of the human orthologs have similar sequence composition: ARK5, 27% DEKR, 16% serine; SNARK, 26% DEKR, 12% serine. Comparison of the UNC-82 sequence to itself (see MATERIALS AND METHODS) revealed a variety of repeated elements ranging from 10 to 91 residues in length that are positioned throughout the C-terminal domain. None of these repeats match any repetitive elements detected in ARK5 and SNARK. The secondary structure prediction programs DSC (KING and STERNBERG 1996) and GGR (GARNIER *et al.* 1996) (http://workbench.sdsc.edu/) suggest that 20–34% of C-terminal domain residues form an α-helix and 9–17% form β-strands.

BLASTp identified the two highest-scoring *C. elegans* UNC-82 paralogs as the products of the *aak-2* and *par-1* genes. AAK-2 is a member of the Snf1/AMPK family, members of which have been studied for their roles in energy metabolism and the stress response. In contrast, the *par-1* gene was originally identified in screens for mutations affecting the anterior–posterior polarity of the earliest cell divisions in the *C. elegans* embryo (KEMPHUES *et al.* 1988). To examine the relationship of UNC-82 to Snf1/AMPK and PAR-1 kinases, a protein alignment of the catalytic domains of representatives from each of these families in *C. elegans*, Drosophila, and humans was constructed (Figure 6). Comparison of all sequences reveals 30% identity overall, whereas the percentage of identical residues is much higher among sequences within a given family. These data support the proposed orthology of the UNC-82/ARK5/SNARK group of kinases and suggest that this group diverged from the Snf1/AMPK and PAR-1 families prior to the divergence of worms, flies, and humans.

UNC-82 lies in a branch of the calcium/calmodulin-regulated kinases tree: To establish the evolutionary relationship of UNC-82 and its apparent orthologs to

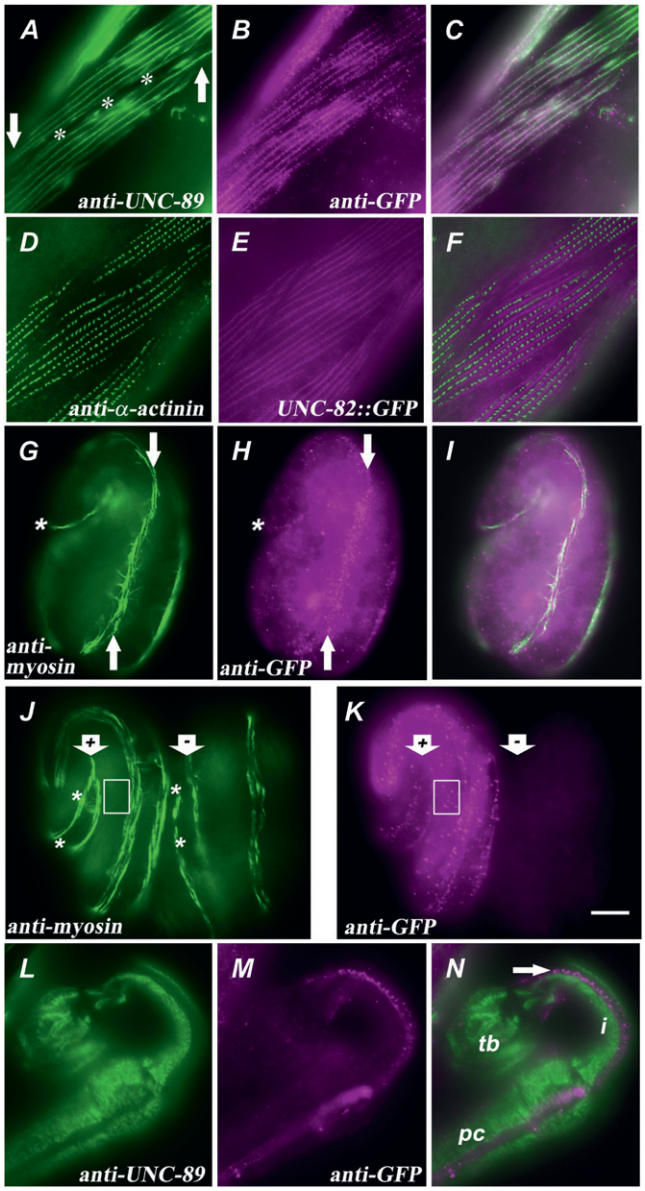


Figure 5.—A rescuing UNC-82::GFP fusion protein is localized near the M-line in body-wall muscle. (A–C) Portions of two body-wall muscle cells from an adult *unc-82(e1323)* homozygote expressing UNC-82::GFP are shown stained with antibodies against (A) UNC-89/obscurin and (B) GFP. The signals coincide in the merged image (C). The asterisks (A) mark the boundary between the two muscle cells, and the arrows indicate one end of each muscle cell. (D–F) One plane from a *z*-series in which UNC-82::GFP was detected throughout the depth of the myofilament lattice, as marked by α-actinin (Francis and Waterston 1985; see Figure 1). (G–I) A 1.5-fold embryo, positioned as in Figure 2A, exhibits punctate UNC-82::GFP concentrated near myosin in a single muscle quadrant (arrows). Asterisks mark the posterior tip of the embryo. (J and K) The *e1323* embryo on the left (arrow marked "+") expresses UNC-82::GFP, whereas the embryo on the right (arrow marked "−") has lost the unstable transgene. (J) The myosin staining (between asterisks) appears broken and patchy in the mutant without the transgene (−). Anti-GFP signal (K) outside the body-wall muscle quadrants (box in J and K) in only the rescued embryo (+) suggests fusion protein expression in nonmuscle cells. An adult pharynx (L–N) exhibits no overlap between the UNC-89 and UNC-82::GFP signals. (L) UNC-89 signal marks the regions of the single-sarcomere pharyngeal muscle cells that contain thick filaments. The UNC-82::GFP signal (M, arrow in N) appears localized to the marginal cells or to a region of the muscle cell that does not contain thick filaments. (N) The pharynx shown is bent back in the isthmus (i) such that the procorpus (pc) lies alongside the terminal bulb (tb). The antibodies used were anti-UNC-89, MH42; rabbit anti-GFP, Abcam; antimyosin A, 5.6; anti-α-actinin, MH35. Bar, 5 μm.

related kinases, a phylogenetic tree (Figure 7) was estimated. The analysis (see Figure S2 and materials and methods) included all available kinase sequences within a small branch of the calcium/calmodulin-regulated kinases (CAMK) in the human kinome tree (Manning *et al.* 2002; http://www.kinase.com) and all related kinases in worm, Drosophila, and yeast, which were identified by BLAST. Both parsimony and neighbor-joining methods produced a single unique tree; these trees are very similar to each other and consistent with the relevant branch of the published human tree. Both methods support the proposed orthology of the UNC-82/ARK5/SNARK group of proteins suggested by BLAST analysis. Similarly, the orthologs within the well-characterized kinase families, such as PAR-1 and Snf1, form groups that have strong statistical support (Figure 7).

In the parsimony tree (Figure 7), UNC-82 appears most closely related to the Kin1/Kin2 and PIG-1 groups, but these relationships are not statistically significant. Further, the grouping of UNC-82 with Kin1/Kin2 and PIG-1 does not occur in the tree constructed using neighbor-joining techniques (not shown). Therefore, the data do not allow us to discern whether UNC-82 is more closely related to any one of these conserved families: Kin1/Kin2, PIG-1, PAR-1, Snf1/AMPK, or SAD-1.

DISCUSSION

UNC-82 is a serine/threonine kinase: Our combined molecular and genetic approaches have unambiguously identified the *unc-82* transcription unit (Figure 4). The single conspicuous feature of the predicted ORF is a serine/threonine kinase catalytic domain. Both *unc-82* alleles contain a mutation in the kinase domain (Figure 6). The missense allele *e1220*, in which a widely conserved glutamic acid is replaced by a lysine, is a likely kinase-dead mutation. In a scanning alanine mutagenesis of a yeast kinase, mutation of the homologous glutamic acid reduced catalytic activity to 1.7% of wild type (Gibbs and Zoller 1991). This confirms the importance of the kinase domain for UNC-82 function and argues against the presumptive catalytic domain having a different, unknown function, as has been suggested by mutational analysis of the ILK kinase

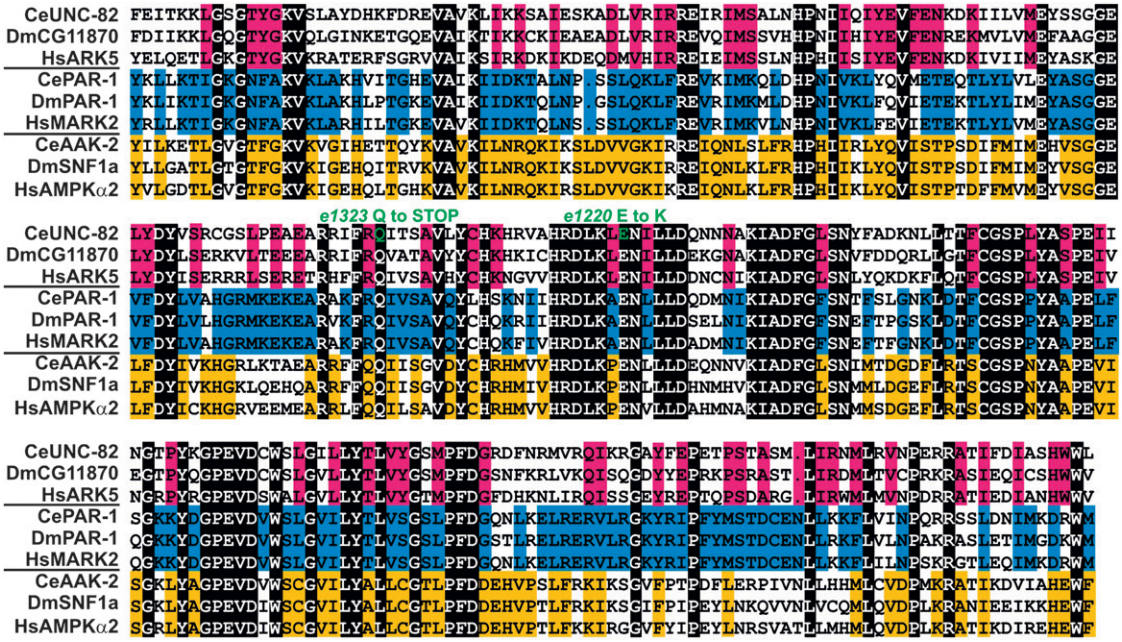


Figure 6.—Alignment of the 256-residue kinase domain protein sequences from representatives of the UNC-82, PAR-1, and Snf1 families from worm (Ce), fly (Dm), and human (Hs) reveals that sequence divergence between families has occurred throughout the catalytic domain. Amino acid identity among all included sequences, indicated by black backgrounds, is 30%. Sequence identity is higher within families: UNC-82, 56%; Snf1, 74%; PAR-1, 81%. Positions that are identical within the UNC-82 family (top three lines), but are not identical in all nine sequences, are indicated by a magenta background. Similarly, such family identities are in yellow backgrounds for the Snf1 sequences and in blue backgrounds for the PAR-1 group. The *unc-82* mutant alleles are represented in green.

(Mackinnon *et al.* 2002). The *e1323* allele contains a premature stop within the kinase domain and is therefore a candidate for a molecular and genetic null mutation. Consistent with this possibility, the polarized light phenotype of *e1323/stDf7* hemizygotes is similar to that of *e1323* homozygotes at the young adult stage (Figure 3). However, the muscle phenotype of *e1323/stDf7* animals becomes more severe in older adults, and the hemizygous animals exhibit additional phenotypes, such as sterility. These differences may reflect residual activity in the *e1323* allele. A more likely alternative is that changes in phenotype are due to the deficit of proteins, such as vinculin, that are encoded by other genes deleted in *stDf7*.

Sequences outside the kinase domain in UNC-82 and its orthologs are noncomplex and repetitive and show no homology among the worm, fly, and human orthologs. If the sequences outside the UNC-82 catalytic domain are responsible for subcellular localization of the protein, the lack of homology or identifiable motifs in these regions suggests two extreme possibilities: (1) The C-terminal sequences have different ancestral origins in worms, flies, and humans and are therefore likely to be functionally different; and (2) the C-terminal sequences have a common origin and function, such as binding to an M-line protein, but are rapidly evolving and/or diverged from related kinases at a very ancient branch. Distinguishing between these will require comparison of the subcellular localization of the orthologs in the different species and identification of the protein region required for correct localization.

Role of UNC-82 in thick filament and M-line organization: The first detectable defects in *unc-82* mutant embryos appear during body elongation (Figure 2). Therefore, loss of *unc-82* activity did not have a discernible effect on the early assembly events (Epstein *et al.* 1993; Hresko *et al.* 1994) such as protein localization to the muscle-cell membrane adjacent to the hypodermis, patterning of attachments at the membrane, or construction of functional contractile units. Vigorous contraction is not the primary cause of the M-line and thick filament defects in *unc-82* mutants, since greatly decreasing myosin activity using the *s95* allele did not rescue the muscle phenotype: double-mutant strains did not contain ordered bands of myosin filaments (Figure 3). However, the birefringent signal in the double-mutant lines appeared more uniform compared to that of the single mutants, suggesting that the more prominent patches of signal in the single mutants result from contractile forces acting on an already disorganized lattice. Previously, a myosin mutation was used to demonstrate that UNC-87/calponin is required to maintain sarcomere integrity during vigorous contraction (Goetinck and Waterston 1994).

Because thick filament disorganization in *unc-82* mutants is not dependent on contractile activity, we propose that *unc-82* is required for regulating some

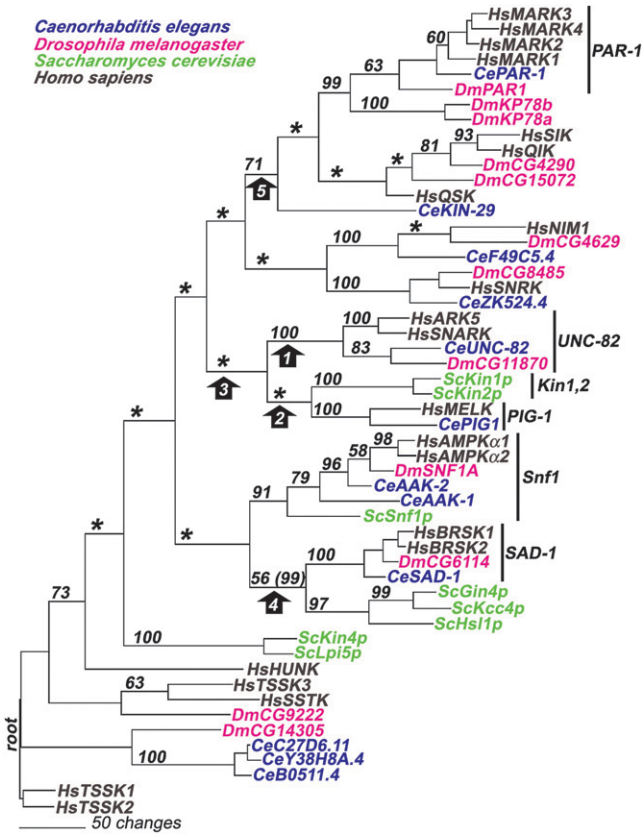


FIGURE 7.—The parsimony tree presents a model of the descent of UNC-82 and related kinases from a common ancestral sequence. The tree was rooted to the HsTSSK sequences as shown in the human kinome tree (MANNING *et al.* 2002; http://www.kinase.com). Bootstrap values (out of 100) are indicated on each branch; asterisks denote nodes that did not receive support >50. The proposed orthology of the UNC-82/ARK5/SNARK proteins is well supported (bootstrap value 100) for the node that separates this group from all other sequences (arrow 1). In contrast, neither the node connecting UNC-82 to the Kin1 and PIG-1 groups (arrow 2) nor the node that separates the UNC-82, Kin1, and PIG-1 groups from all other groups (arrow 3) is supported by bootstrap analysis. Two groupings that receive moderate bootstrap support in this parsimony analysis (arrows 4 and 5) receive higher support (in parentheses) using neighbor-joining techniques: the node grouping the SAD-1 family with the trio of yeast kinases Gin4/Kcc4/Hsl1 and the node grouping the PAR-1 family with HsSIK, HsQSK, and CeKIN-29.

aspect of thick filament organization during changes in muscle-cell length. During the embryonic stages in which defects appear, the number of muscle cells remains constant as the length of the animal grows from 1.5 to approximately four times egg length. As cells elongate, thick and thin filament attachment sites move farther apart (MOERMAN and WILLIAMS 2006) so that the interdigitated filaments (see Figure 1), which are oriented longitudinally in the cell, must become longer and slide past each other. The UNC-82 kinase may play a role in one of these processes. The normal patterning of integrin in *unc-82* mutants argues that, as cells elongate, M-lines are properly spaced at the membrane, but thick filament components and some M-line components become unevenly distributed within the cell (Figure 2). Since the central portion of the bipolar thick filament is positioned at the M-line (see Figure 1), as M-lines move apart, filament centers must also move. Possible roles for UNC-82 in filament translocation include proper attachment of thick filaments to the M-line or regulation of the ability of the interdigitated filaments to slide past each other. In these cases, the bright patches of staining in *unc-82* mutants may correspond to abnormally wide regions of filament lattice that reflect a failure in filament placement. Alternatively, it is possible that *unc-82* is required for the addition of protein components to the lengthening thick filament and that the abnormal accumulations of myosin and paramyosin represent unincorporated or misincorporated protein.

The localization of UNC-82::GFP near the M-line (Figure 5) suggests that this kinase targets proteins in the central portion of the bipolar thick filament or the M-line. The punctate localization of UNC-82 in embryos and some adult preparations (Figure 5, B and H) does not resemble the distribution of any known protein or structure. Because the endogenous GFP signal from UNC-82::GFP in living and fixed animals (Figure 5E) appears in unbroken lines, as in the anti-UNC-89 stain, the puncta may represent a portion of the UNC-82 protein pool that is more accessible to antibodies. The colocalization of UNC-82 and UNC-89/obscurin and the disorganization of UNC-89 in *unc-82* mutants (Figure 1) suggest that the two proteins may directly interact or at least be members of a single signaling pathway. UNC-89/obscurin is a giant protein that contains two potential kinase domains and many Ig domains, which are hypothesized to mediate thick filament attachment to the M-line (BENIAN *et al.* 1996; SMALL *et al.* 2004). It is possible that UNC-82 targets UNC-89/obscurin to regulate thick filament attachment or some other process that requires transmission of a signal from membrane-proximal components of the M-line to the thick filaments.

Myosin and paramyosin are the earliest affected proteins in *unc-82* mutants (Figure 2) and are therefore candidate targets of the enzyme as well. Previous ultrastructural analyses of single- and double-mutant adult worms revealed that *unc-82* mutations resulted in thick filaments that lacked paramyosin and aberrant filaments that were likely composed of paramyosin. This led to the suggestion that the *unc-82* gene product probably affects thick filament assembly through its actions on paramyosin (WATERSTON *et al.* 1980). Paramyosin is a coiled-coil protein homologous to the C-terminal two-thirds of the myosin heavy chain rod and may be considered a "headless myosin" (KAGAWA *et al.* 1989). The small N-terminal nonhelical "headpiece" region of paramyosin, but not the coiled-coil domain, is phosphorylated on serine residues by an endogenous

kinase (Schriefer and Waterston 1989). The nonhelical headpiece contains multiple copies of the proposed phosphorylation motif, S_S_A. The more acidic isoelectric species of paramyosin are absent from extracts of *unc-82* mutant worms (unpublished data cited in Schriefer and Waterston 1989), suggesting that the headpiece is a direct or indirect target of UNC-82.

The S_S_A motif is also present in multiple copies in the nonhelical C-terminal tailpieces of both myosin heavy chains expressed in body-wall muscle (Schriefer and Waterston 1989), suggesting that myosin and paramyosin may be targeted by the same kinase. Abnormal thick-filament structures found in *unc-82* mutants contain both myosin and paramyosin (Epstein *et al.* 1987). The possibility that myosin is a target of UNC-82 is supported by the observation that the myosin phenotype observed in *unc-82* mutant embryos is similar to that caused by removal of the phosphorylation motifs contained in the myosin nonhelical tailpiece: early patterning of myosin is normal, but aberrant patchy distribution appears as elongation proceeds (Hoppe *et al.* 2003). While the mechanisms guiding thick filament elongation and placement during growth are not well understood in striated muscle, many prior studies have examined the regulation of assembly and disassembly of nonmuscle and smooth muscle myosins by phosphorylation (Castellani and Cohen 1987; Castellani *et al.* 1988; reviews: Moussavi *et al.* 1993; Brzeska and Korn 1996; Redowicz 2001; Bosgraafa and van Haastert 2006). In *C. elegans* striated muscle, both myosin and paramyosin are phosphorylated in an assembly-dependent manner (Dey *et al.* 1992). UNC-82::GFP is present at or near the M-line and therefore distant from the ends of the thick filament where much of the subunit addition presumably occurs. Its location may imply that, if UNC-82 regulates these proteins, UNC-82 directly phosphorylates only a subpopulation of myosin and paramyosin molecules or that it serves as an intermediate in a kinase cascade that targets these proteins.

***unc-82* mutations identify a role for this conserved kinase family in normal development:** The UNC-82 orthologs ARK5 and SNARK (NUAK-1 and NUAK-2) (Figure 7) are named for their similarity to the Snf1/AMPK kinases, which are thought to be activated by increased levels of the nucleotide AMP during metabolic or other stress conditions. In humans, ARK5 (cDNA KIAA0537) is strongly expressed in both skeletal and cardiac muscle (http://www.kazusa.or.jp/huge/) and may therefore play a role similar to that of UNC-82 in developing striated muscle. Tests of the *in vivo* function of ARK5 in vertebrate muscle (Fisher *et al.* 2005; Niesler *et al.* 2007) have been limited to its putative role as an AMP-activated kinase.

The mechanism of UNC-82 activation during muscle development is unknown. Despite its similarity to AMPK, UNC-82 functions in cells of healthy developing embryos where it is unlikely to be exposed to the elevated levels of AMP or other activating signals important in the stress response. Further, although the UNC-82/ARK5/SNARK family lies within the so-called CAMK portion of the kinome tree (Manning *et al.* 2002; http://www.kinase.com), this designation is applied to a large region of the kinome, including many proteins whose specific mode of regulation has not been established experimentally.

If expression of UNC-82::GFP outside body-wall muscle (Figure 5) represents distribution of the native protein, it is likely that UNC-82 has unidentified functions in other tissue types. Similarly, ARK5 (http://www.kazusa.or.jp/huge/; Fisher *et al.* 2005) and SNARK (Lefebvre and Rosen 2005) are expressed in many nonmuscle tissues and organs. As putative AMP-activated enzymes, both have been studied for roles in the cellular stress response in tissue culture cells and cancer cell lines. ARK5 is induced in response to glucose starvation (Suzuki *et al.* 2003) and can confer resistance to apoptosis induced by stress (Suzuki *et al.* 2005) or apoptotic signals (Suzuki *et al.* 2004). The response of SNARK activity to different stresses varies greatly depending upon the cell type tested (Lefebvre and Rosen 2005). The role of UNC-82, ARK5, and SNARK in normal development or physiology in nonmuscle tissues is unknown. A role for the human kinases in the regulation of the cytoskeleton, such as that proposed for UNC-82 in muscle, might be implied by changes observed in tumor cells when kinase activity is altered. For example, increased SNARK activity is associated with increased motility and invasiveness (Legembre *et al.* 2004) and with cell detachment in culture (Suzuki *et al.* 2003). ARK5 activity is also correlated with metastatic invasiveness in tumors derived from several tissue types (Kusakai *et al.* 2004a,b; Suzuki *et al.* 2004, 2005).

UNC-82 lies in a branch of the kinome tree rich in cytoskeletal regulators: Our phylogenetic analyses support a close evolutionary relationship between UNC-82 and the AMPK, PAR-1, SAD-1, PIG-1, and Kin1/2 protein families. Several of the kinases within these families have been shown by mutation to play a role in normal development and in particular are involved in the regulation of the cytoskeleton and cell polarity. We propose that UNC-82 is an additional member of this superfamily of developmentally important kinases whose salient feature is the regulation of cellular organization rather than any presumed mode of activation.

In light of the disorganized myosin phenotype in *unc-82* mutants, it is intriguing to note that other kinases within the related kinase families interact with myosin in nonmuscle cells. The *C. elegans* nonmuscle myosin NMY-2 was identified through its interaction with the PAR-1 protein and is required for establishing embryonic polarity (Guo and Kemphues 1996). Drosophila SNF1/AMPK is required during embryogenesis for establish-

ment of epithelial cell polarity and acts through phosphorylation of the myosin regulatory light chain (Lee *et al.* 2007). The presence of various myosin-associated kinases in this region of the kinome suggests that myosin was a substrate of an ancestral kinase prior to the divergence of the UNC-82, AMPK, and PAR-1 groups.

We thank Mike Becker of the Genome Sequencing Center for instructions and assistance in library production and sequencing, Yuji Kohara for cDNA clones, Ziva Misulovin for RNA and expert advice, Guy Benian for antibodies and for suggesting that we analyze the protein sequence for internal repeats, and Xhirong Bao and Todd Barkman for help with our phylogenetic analysis. We thank Tim Schedl and his lab members for use of microscopes and scientific input and Doug Coulter for critical reading of the manuscript. Some strains were obtained from the *C. elegans* Genetic Center. This material is based upon work supported by the National Science Foundation under grant no. MCB 0454737.

LITERATURE CITED

Benian, G. M., T. L. Tinley, X. Tang and M. Borodovsky, 1996 The *Caenorhabditis elegans* gene *unc-89*, required for muscle M-line assembly, encodes a giant modular protein composed of Ig and signal transduction domains. J. Cell Biol. **132:** 835–848.

Bosgraafa, L., and P. J. M. van Haastert, 2006 The regulation of myosin II in Dictyostelium. Eur. J. Cell Biol. **85:** 969–979.

Brzeska, H., and E. D. Korn, 1996 Regulation of class I and class II myosins by heavy chain phosphorylation. J. Biol. Chem. **271:** 16983–16986.

Castellani, L., and C. Cohen, 1987 Rod phosphorylation favors folding in a catch muscle myosin. Proc. Natl. Acad. Sci. USA **84:** 4058–4062.

Castellani, L., B.W. Elliott, Jr. and C. Cohen, 1988 Phosphorylatable serine residues are located in a non-helical tailpiece of a catch muscle myosin. J. Muscle Res. Cell Motil. **9:** 533–540.

Dey, C. S., P. R. Deitiker and H. F. Epstein, 1992 Assembly-dependent phosphorylation of myosin and paramyosin of native thick filaments in *Caenorhabditis elegans*. Biochem. Biophys. Res. Commun. **14:** 1528–1532.

Dibb, N. J., D. M. Brown, J. Karn, D. G. Moerman, S. L. Bolten *et al.*, 1985 Sequence analysis of mutations that affect the synthesis, assembly and enzymatic activity of the *unc-54* myosin heavy chain of *Caenorhabditis elegans*. J. Mol. Biol. **183:** 543–551.

Duerr, J. S., 2006 Immunohistochemistry, in *WormBook*, edited by The *C. elegans* Research Community. *WormBook*, http://www.wormbook.org.

Epstein, H. F., I. Ortiz and G. C. Berliner, 1987 Assemblages of multiple thick filaments in nematode mutants. J. Muscle Res. Cell Motil. **8:** 527–536.

Epstein, H. F., D. L. Casey and I. Ortiz, 1993 Myosin and paramyosin of *Caenorhabditis elegans* embryos assemble into nascent structures distinct from thick filaments and multi-filament assemblages. J. Cell Biol. **122:** 845–858.

Fisher, J. S., J. S. Ju, P. J. Oppelt, J. L. Smith, A. Suzuki *et al.*, 2005 Muscle contractions, AICAR, and insulin cause phosphorylation of an AMPK-related kinase. Am. J. Physiol. Endocrinol. Metab. **289:** 986–992.

Francis, G. R., and R. H. Waterston, 1985 Muscle organization in *Caenorhabditis elegans*: localization of proteins implicated in thin filament attachment and I-band organization. J. Cell Biol. **101:** 1532–1549.

Garnier, J., J. F. Gibrat and B. Robson, 1996 GOR method for predicting protein secondary structure from amino acid sequence. Methods Enzymol. **266:** 540–553.

Gibbs, C. S., and M. J. Zoller, 1991 Rational scanning mutagenesis of a protein kinase identifies functional regions involved in catalysis and substrate interactions. J. Biol. Chem. **266:** 8923–8931.

Goetinck, S., and R. H. Waterston, 1994 The *Caenorhabditis elegans* UNC-87 protein is essential for maintenance, but not assembly, of bodywall muscle. J. Cell Biol. **127:** 71–78.

Guo, S., and K. J. Kemphues, 1996 A non-muscle myosin required for embryonic polarity in *Caenorhabditis elegans*. Nature **382:** 455–458.

Hanks, S. K., and T. Hunter, 1995 Protein kinases 6. The eukaryotic protein kinase superfamily: kinase catalytic domain structure and classification. FASEB J. **9:** 576–596.

Hoppe, P. E., and R. H. Waterston, 2000 A region of the myosin rod important for interaction with paramyosin in *Caenorhabditis elegans* striated muscle. Genetics **156:** 631–643.

Hoppe, P. E., R. C. Andrews and P. D. Parikh, 2003 Differential requirement for the nonhelical tailpiece and the C-terminus of the myosin rod in *C. elegans* muscle. Mol. Biol. Cell **14:** 1677–1690.

Hresko, M. C., B. D. Williams and R. H. Waterston, 1994 Assembly of body wall muscle and muscle cell attachment structures in *Caenorhabditis elegans*. J. Cell Biol. **124:** 491–506.

Jakubowski, J., and K. Kornfeld, 1999 A local, high-density, single-nucleotide polymorphism map used to clone *Caenorhabditis elegans cdf-1*. Genetics **153:** 743–752.

Kagawa, H., K. Gengyo, A. D. McLachlan, S. Brenner and J. Karn, 1989 Paramyosin gene (*unc-15*) of *Caenorhabditis elegans*: molecular cloning, nucleotide sequence and models for thick filament structure. J. Mol. Biol. **207:** 311–333.

Kemphues, K. J., J. R. Priess, D. G. Morton and N. S. Cheng, 1988 Identification of genes required for cytoplasmic localization in early *C. elegans* embryos. Cell **52:** 311–320.

King, R. D., and M. J. Sternberg, 1996 Identification and application of the concepts important for accurate and reliable protein secondary structure prediction. Protein Sci. **5:** 2298–2310.

Kusakai, G., A. Suzuki, T. Ogura, M. Kaminishi and H. Esumi, 2004a Strong association of ARK5 with tumor invasion and metastasis. J. Exp. Clin. Cancer Res. **23:** 263–268.

Kusakai, G., A. Suzuki, T. Ogura, S. Miyamoto, A. Ochiai *et al.*, 2004b ARK5 expression in colorectal cancer and its implications for tumor progression. Am. J. Pathol. **164:** 987–995.

Lee, J. H., H. Koh, M. Kim, Y. Kim, S. Y. Lee *et al.*, 2007 Energy-dependent regulation of cell structure by AMP-activated protein kinase. Nature **447:** 1017–1020.

Lefebvre, D. L., and C. F. Rosen, 2005 Regulation of SNARK activity in response to cellular stresses. Biochim. Biophys. Acta **1724:** 71–85.

Legembre, P., R. R. Schickel, B. C. Barnhart and M. E. Peter, 2004 Identification of SNF1/AMP kinase-related kinase as an NF-kappaB-regulated anti-apoptotic kinase involved in CD95-induced motility and invasiveness. J. Biol. Chem. **279:** 46742–46747.

Mackinnon, A. C., H. Qadota, K. R. Norman, D. G. Moerman and B. D. Williams, 2002 *C. elegans* PAT-4/ILK functions as an adaptor protein within integrin adhesion complexes. Curr. Biol. **12:** 787–797.

Manning, G., D. B. Whyte, R. Martinez, T. Hunter and S. Sudarsanam, 2002 The protein kinase complement of the human genome. Science **298:** 1912–1934.

Moerman, D. G., and B. D. Williams, 2006 Sarcomere assembly in *C. elegans* muscle, in *WormBook*, edited by The *C. elegans* Research Community. Wormbook, http://www.wormbook.org.

Moerman, D. G., S. Plurad, R. H. Waterston and D. L. Baillie, 1982 Mutations in the *unc-54* myosin heavy chain gene of *Caenorhabditis elegans* that alter contractility but not muscle structure. Cell **29:** 773–781.

Moussavi, R. S., C. A. Kelley and R. S. Adelstein, 1993 Phosphorylation of vertebrate nonmuscle and smooth muscle myosin heavy chains and light chains. Mol. Cell. Biochem. **128:** 219–227.

Niesler, C. U., K. H. Myburgh and F. Moore, 2007 The changing AMPK expression profile in differentiating mouse skeletal muscle myoblast cells helps confer increasing resistance to apoptosis. Exp. Physiol. **92:** 207–217.

Redowicz, M. J., 2001 Regulation of nonmuscle myosins by heavy chain phosphorylation. J. Muscle Res. Cell Motil. **22:** 163–173.

Schriefer, L., and R. H. Waterston, 1989 Phosphorylation of the N-terminal region of paramyosin in *Caenorhabditis elegans*. J. Mol. Biol. **207:** 451–454.

Small, T. M., K. M. Gernert, D. B. Flaherty, K. B. Mercer, M. Borodovsky *et al.*, 2004 Three new isoforms of *Caenorhabditis elegans* UNC-89 containing MLCK-like protein kinase domains. J. Mol. Biol. **342:** 91–108.

Suzuki, A., G. Kusakai, A. Kishimoto, J. Lu, T. Ogura *et al.*, 2003 Identification of a novel protein kinase mediating Akt survival signaling to the ATM protein. J. Biol. Chem. **278:** 48–53.

Suzuki, A., J. Lu, G. Kusakai, A. Kishimoto, T. Ogura *et al.*, 2004 ARK5 is a tumor invasion-associated factor downstream of Akt signaling. Mol. Cell. Biol. **24:** 3526–3535.

Suzuki, A., S. Iida, M. Kato-Uranishi, E. Tajima, F. Zhan *et al.*, 2005 ARK5 is transcriptionally regulated by the Large-MAF family and mediates IGF-1-induced cell invasion in multiple myeloma: ARK5 as a new molecular determinant of malignant multiple myeloma. Oncogene. **24:** 6936–6944.

Waddle, J. A., 1993 Molecular analysis of actin capping protein function in *Caenorhabditis elegans*. Ph.D. Thesis, Washington University, St. Louis.

Waterston, R. H., J. N. Thomson and S. Brenner, 1980 Mutants with altered muscle structure of *Caenorhabditis elegans*. Dev. Biol. **77:** 271–302.

Wicks, S. R., R. T. Yeh, W. R. Gish, R. H. Waterston and R. H. Plasterk, 2001 Rapid gene mapping in *Caenorhabditis elegans* using a high density polymorphism map. Nat. Genet. **28:** 160–164.

Williams, B. D., and R. H. Waterston, 1994 Genes critical for muscle development and function in *Caenorhabditis elegans* identified through lethal mutation. J. Cell Biol. **124:** 475–490.

Yuan, A., M. Dourado, A. Butler, N. Walton, A. Wei *et al.*, 2000 SLO-2, a K+ channel with an unusual Cl- dependence. Nat. Neurosci. **3:** 771–779.

Communicating editor: A. Villeneuve

GENETICS

Supporting Information
http://www.genetics.org/cgi/content/full/genetics.109.110189/DC1

C*aenorhabditi elegans unc*-82 Encodes a Serine/Threonine Kinase Important for Myosin Filament Organization in Muscle During Growth

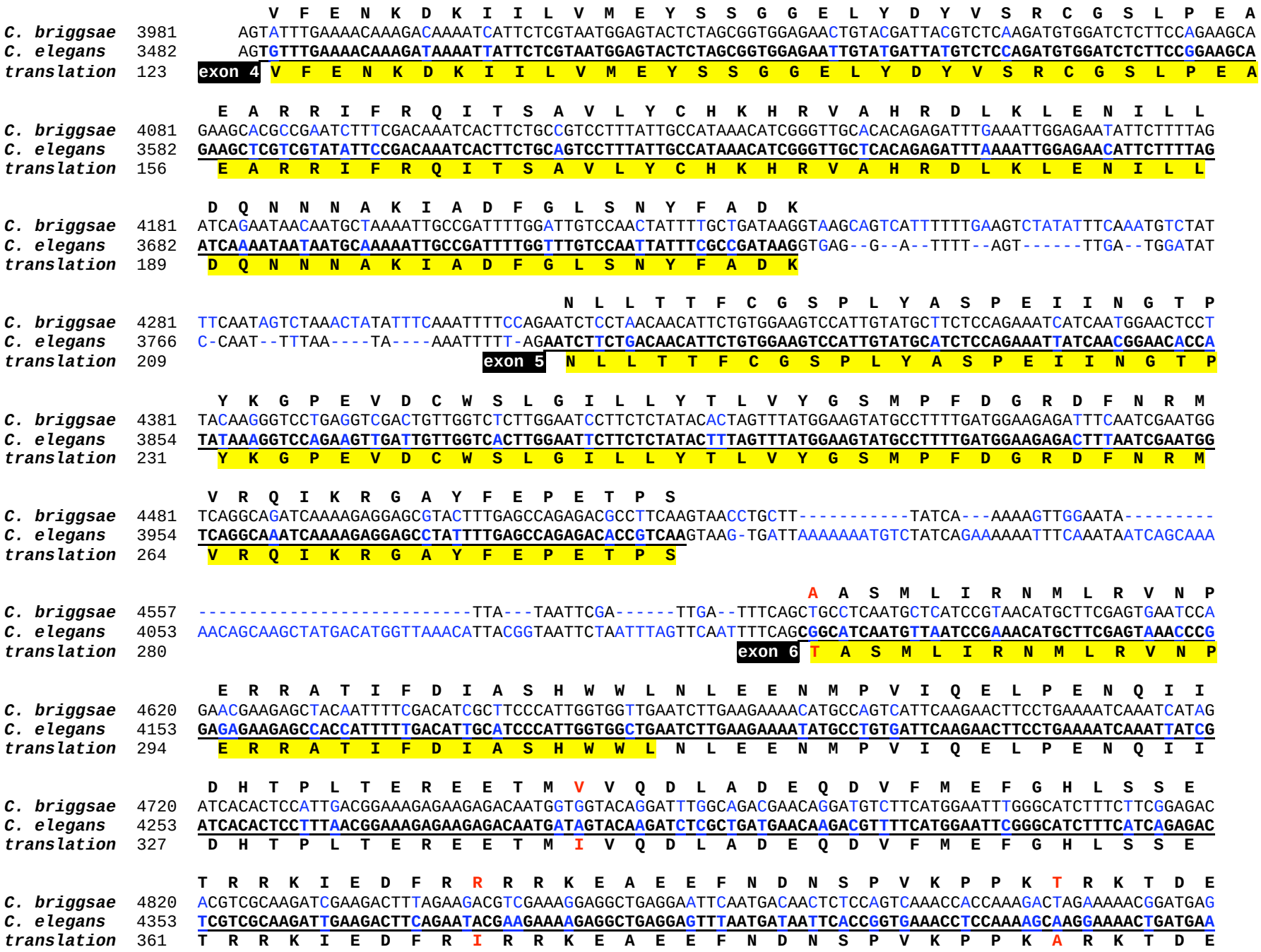
Pamela E. Hoppe, Johnnie Chau, Kelly A. Flanagan, April R. Reedy and Lawrence A. Schriefer

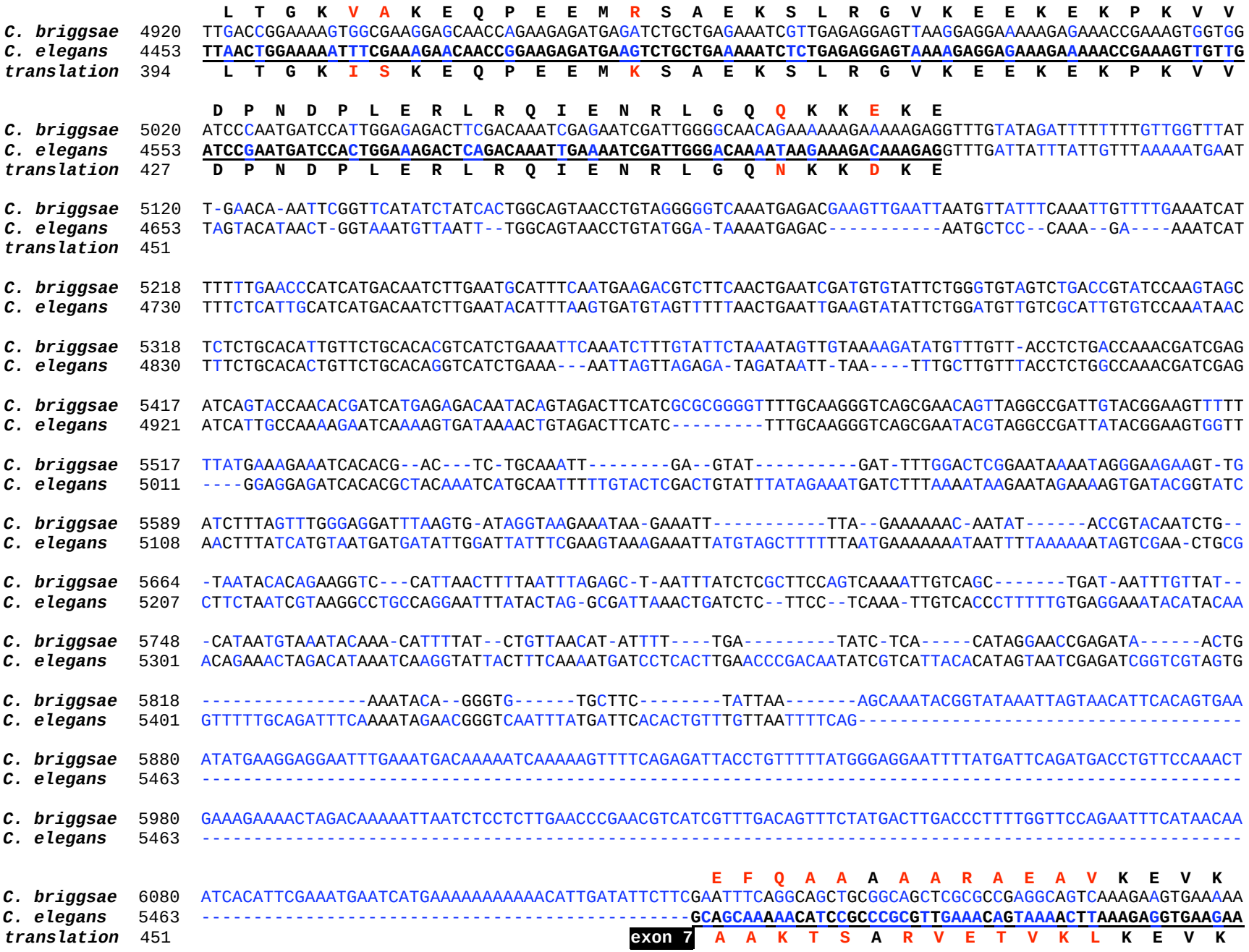
DOI: 10.1534/genetics.109.110189

```
                                M  P  V  G  N  S  G  S  R  S  G  R  L  A  R  I  S  Q  S  N  I  R  P  F  G  S  N  G  G  G
C. briggsae   1      GAGGACGAATGCCCGTCGGTAATTCGGGATCTCGATCGGGTCGTCTGGCACGTATCAGCCAATCCAATATTCGACCGTTTGGATCAAATGGCGGTGGAGC
C. elegans    1      GTGGACGAATGCCCGTCGGAAATTCGGGATCTCGATCGGGTCGATTGGCACGTATCAGTCAATCGAATATACGACCGTTTGGATCAAATGGCGGTGGAGC
translation   1        exon 1  M  P  V  G  N  S  G  S  R  S  G  R  L  A  R  I  S  Q  S  N  I  R  P  F  G  S  N  G  G  G

                        A  T  N  S  Y  G  P
C. briggsae   101    TACGAACAGTTATGGACCTGGTAAGTTGTTTCTTCGGAAGAAAAAATCCATTCTTTTTGTCTCTGTGTGTGAGATATAGAAGAAAGAGAGAGCAGGGTGA
C. elegans    101    TACGAACAGTTATGGACCCGGTTAGTAGTGTC---------AAAAA------------------------GAGATTTTGA-------GATGG--GGGTGA
translation   37     A  T  N  S  Y  G  P

C. briggsae   201    GAATTGAG--TGTAATAC---GAAACACTTAGTCCTCAAAAACTCAAAGGGACCCCATCCCCGAAACTCCTGTGCGGGAGAGATAA-ATCAATGACTGCG
C. elegans    159    GAAGTGAGAGTGTAACACCCCGAAACACTTAGTCCTCAAAAACTGAGAGAGA----------GAGA---------GAGAGAGAGAAGATCTA------CG

C. briggsae   295    AGAGCACAATA-AGAAGAAGAAGAAG-AA--GATACGAGTAGAAAAGACAGA-AGAAGAAGA-AGAAGACAGTTGTTCGTGTCCAGTGGTTTCCGCCCTC
C. elegans    234    AAAGTTCTAAAGAGAAAGAGATAAATCAATGGATACGAG-AGCAGCGGCAACGAGTAGAAAAGATAAAACAGTTGTTCGTGTCCAGTGGTTTCCGCCCTC

C. briggsae   389    AAAATGATTGTTAAAGAAGAGTTTGTAGGTGGGATGGAAGATGAAAGAGTACGAAGAATAGAAAAAGAAAATGACTTGGATATACAGGAAGCATATGGAA
C. elegans    333    AAAATCATTGTT------------GTA-------TAGAAAAAAATGAGTACGAAGAAT-GAAAATT----TGAGT--GAAGTATCGGA--CTTTTGGGA

C. briggsae   489    GA-GAAAGTTCAG-GTTAACTTGGTCA----GATGAAA--ACAACACAATTACTTCTGAAAATAGACCGCAGACAGGAATGCATTTTTAGTAAAGAAATG
C. elegans    405    GTTGAAAGATGAAAGTCAA-TTGATTATATAGATTAAGGTACATGACAATT-----TGAAAATTGA----AGAAAC-ATTGGAATTT--GCAGGGTTATA

C. briggsae   581    TAAACATCAACTGGATGACGTTTCTTAAATT--TTTTGGAACTTCTGTTTCGTAACTTTGAGAGCCTAAAAAATTAAATT--TGT-CA-ATTCACCATAT
C. elegans    492    TTATGTTGAATT---TGTC-TTTGAAAAATTAATTCTACAACTT-----TCG-ATCTATTTTAGG----AAAATGAATTTGTTATACATATTAATAATAT

C. briggsae   675    TTCCTCTAATAGAAGCTCCCATTCAGGGCGTCAACT-TTGA---GCTCAATTTTTTTGATTCCTGTGAAAGATATC--A--AAATTTCGTCAATTGATAA
C. elegans    578    TT-------TTGAATTT---ATTTAGTTT-TCAAATATTTACAAGCTCAATT------AGGCGTTTTAAAGTGAGCGGACCAAATATCG--AATT-ATA-

C. briggsae   767    AATATTGGT--GATATTATGATAAACATTTTATCAGTTTGCAGTTTTCTGATA-TCTTCCACAGGTATTTAGATAACTGAGCTCAAAGTTGACGGTTAGG
C. elegans    656    --TATTTCTAAGAAATTAAAAAAAAAATTTT--CAGTCAAAAAGTGTCAATTACTCTTTT-CA--TCTTTC-ATACCC-A--TCATAGTT-----TTGGA

C. briggsae   864    A-TGGGCTCCTTTGTAAACACGGTATTAGGAAGTCTCATGAAGTAGCGAGGCGTATGTTTTCAAAAAAATTACTGAAGTTTGTTACTTTTTTGGAA-GTT
C. elegans    741    AATCGACA-------AAACA--GTTTT-----------TTAA--A-CGA---------TTT-------ATAACTAAAGCTC---ACTTTTTTTTAAAGAT

C. briggsae   962    TGTT-ACTTTTTCGCATCAGAAGTATATGGCACCTCAAATTCTGT--AATTTTTCATTCA-TGTAGAATTTGAAATTTGTTTATTGGTATCTGGGTGACT
C. elegans    799    TAAGCAGTTTTTC--AT----AGTTT-TGGC---TGAAATTTTTTGAAAAATCTAAATCACTTTAGAGTGTTTTATTTATTTATT-----CAGCCTTTTT

C. briggsae   1058   GAAATT-GTGCACGTACCTACTTCTGAATTCTGAATTTCCAATCTGTTGACTATTGACCAAAAAGAAAAAAAAAATGACAAATCGTTCCACCTGAATCGG
C. elegans    884    CACATTAGCCCAACT-CC-ACCTCAAAGTAATAAGTTTACA--CTGTAAA-TATGGACGAAACATTTTCAAAA--TTA-------TTTCGGC-GATT---

C. briggsae   1157   ACTCCAGACCCACAATTCTCCCCTTTTTTCTTCTTGGACCCAAACTGAAAAAGGTCGTTTTGGTCATTGTTCTTTCTCAGTTTTCTTCTTCATTCCAGAC
C. elegans    965    --TCCAAAT----AATT-------TTTTAAATAATGAAACC---CCGTAAAA---CTATCTG--C------CTGTAAGAATTTTATGCATC--TC-----

C. briggsae   1257   ACATGTGACTCGTGACTCTAATTTGTGACATATA-TCTTGAAGTCGTTTGAAGAAGCCAACTTTTTTGCGCTTTTTCACCCATCCC--AGAGTCGTCATT
C. elegans    1032   ATATGTATAT---GAA---AAATTAT--CATATACTCTCAAAATACTCCGGAGTA-----CTATTT--CAATGTTTAATT--TCCCTTAAAAACAACAAT

C. briggsae   1354   TTTTCTCGAGTTGGGAGCAAGAACTCCCTGGAGAAGACGGTTGCTGAGGGAAAGGAAAAAAGGCCATT-TTA-T-TGTCAACTCTCTTGTGTTGCTCTCT
C. elegans    1115   TT----CAATTTTGGAATAAAAATTTCCTATTGAAAA-------------AAAT--AATAATTCCGTACTTACTATGTCA----TATTGTTTT-------
```

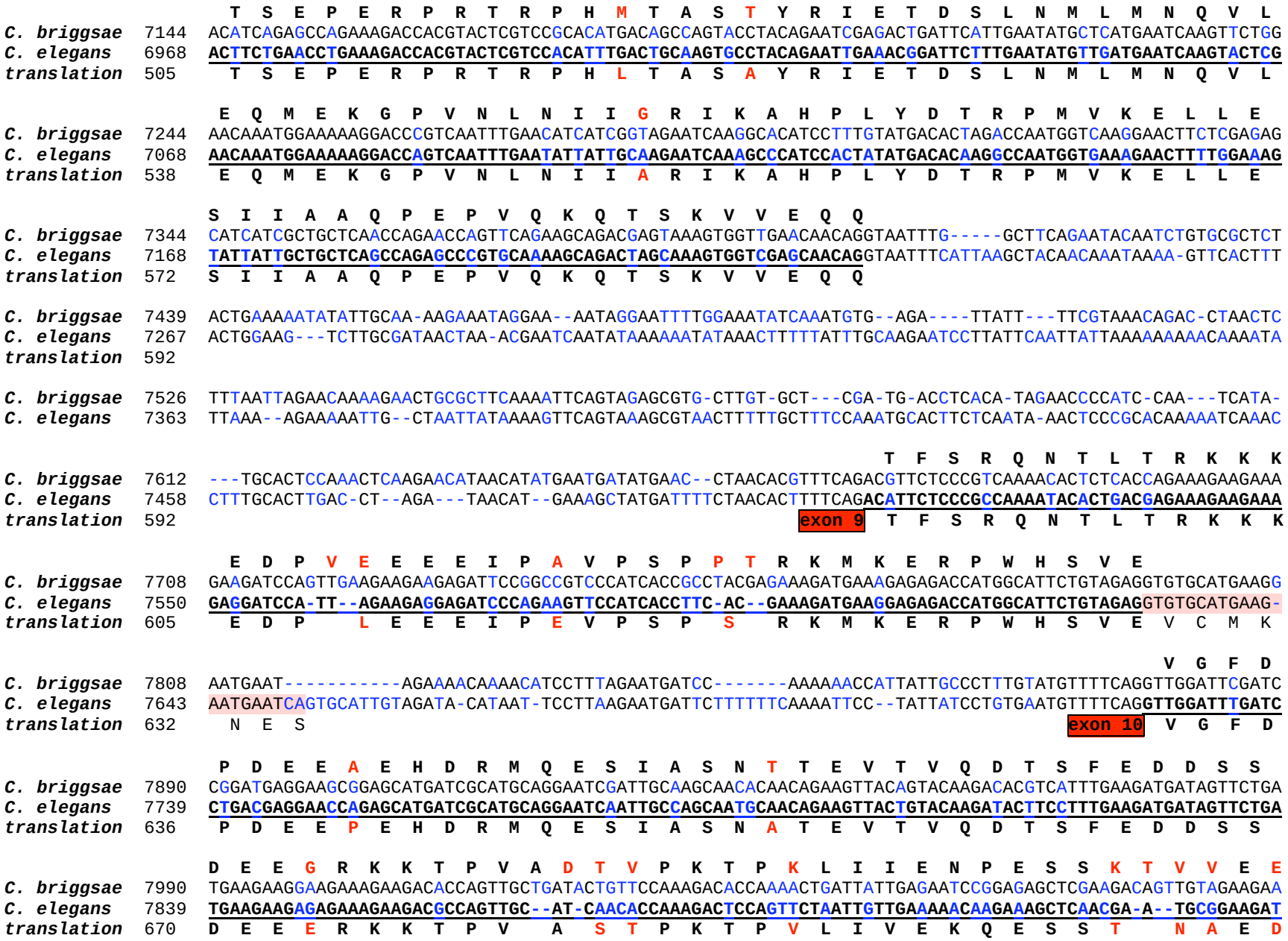
```
C. briggsae   1451  TCGCACTGCCAAAAATGATGAGAGATGCTCTTCTCTGGAGCAAAGAAGTTTCTTCTTCGGACCGACGACAAACTGACACCCACCCAAAAAACTGGGGAGG
C. elegans    1184  -------------AAGGTTG--------TATTTTTTG---------A-------CTTAGTACATATG-CAAACT------------------TGGTCATT
translation   37

C. briggsae   1551  AAAAAATGCGGAAACCGGAAAACCCAGAAGGCAAAAAAGTTTATGACTCAAGAAGAATTTAACAAAACTCCCAAAAACCGAGTCAACACCGTTATGAACT
C. elegans    1229  AAGAGTT----------------CAAGACTACAATAGAGACAATGACT------GAATT---CAAA--TCG----AATTT--TCAAAACGTTTAT----T

C. briggsae   1651  ATATCTCCTTTACTCTTTTCCGAACACTTCGTTCCTTCAAAAAGATTGCTCTAGGGTGACTAGCTCATCCATCACTTCATCCGTCATTTTTTGGAACGAA
C. elegans    1292  GG-TCGCCTGAAAAC----CC-AAC-CTT-GTTACCT-------ATT-------------TAG--------TCTCT------GTCATTTTTTTGCTCAAA

C. briggsae   1751  ATTTTCAGTATCTTTTCTTGCAACCACCAACTTTTTTTCTCGCTGACATGGAATGACTCCCAAGTGTTACCCTCGAAAAAAATCGAGACCTATATGGTTG
C. elegans    1350  AATT-CTGAAT------TTGAAATC------TGTTATTCT-----ACA--------------------ACCCACAATAAAAA----GAC--AAATCGTT-

C. briggsae   1851  GATAGAGCGTAATAAGTTATTTTAGATTCCAAGCGGAATAGTTGGTAAAATATTTCTTTTGGGAGTTATGGTCCAAAAAGTACCAAAAAATGCCTAAATT
C. elegans    1404  --------------------TTT-GACTC---------T--TTGCTCCC------CTTTT-----------TCAATACA-TACCAAAGAA----------

C. briggsae   1951  TTCAAAAAATCATAACTTTGTCAGTTTATGTCCGTTTGTTGAAATTTAAAAAACAGTATATAGACAATTAATTGAACACTTGAGAATAAATTTTTTAAAT
C. elegans    1444  ----AAAGGTCTT--CTCGGTCA-----------TTTGTT----TTT---------TCTGCTGCCTATT----------TT----------TTTTTACAT

C. briggsae   2051  TTTTTTCCTATTGATTGGATTATTGATTTTTTGAAAAAATCAAAAAATTTTTCAATATCTCAAAAATCCTAGGTTAACTTTTTTCTCATTATGGGGAAAAT
C. elegans    1495  GTGA--CCTAATTTGTGGCATAT--ATATCTT---AAAGTCG-------TTTGAAGA----AAAGGACC------AACTTTTTTTTCGTT----GCTCTT

C. briggsae   2151  CATGCTCACACAAGTTATGCCAAAAAATATCGATTTTTGGCACTTGGCTTCTTTGAATTCGTTGGCCGGTATGG-TGAGGG-GCATACTTGTTTTTGTGA
C. elegans    1567  CTTG---ACCCA--TCGCGGCG-------TC-ATTTTT-----TTGG------TGACTGGGT--GCTGGT--GGATGAGGGCGCGAACTCC---------

C. briggsae   2249  TTTTTACTTGCGTATGGTTTGATTGATTGATTTCCCCCGGGGTATTGATATGTTTCTATAATCAAAT-GT-GCTATATTT-TAATGGTCA-TAATGAGAA
C. elegans    1629  ------CTGGAGAAGAAGAAGACTG-TTGCTGACAAAAAAGCCATT-------TT--ATTGTCAACTTGTTGCTCTTTCTCTA-TAC-CAAAAATGA--A

C. briggsae   2345  GAAAAATTAACCTTAGGATTTTTGAGA-TATT-G-AAA-AATTTTTTGATTTTTTC--AAAAAATCAATATCCAATCAATAGGAAAAAAATTTAAAAAAA
C. elegans    1714  GAGAGAGTGAG---AG-ATCCTCCA-ACTCGTCGCAAAGAAGTTTCT--TTTT--CGGACAAACTGAC--ACCCAG-C--TACGCAAAAAAG----AGAA

C. briggsae   2439  TTTATTCTCAAGTGTTCAATC-AATTGTCTATCTACTGTTTTTTAAATTTC---AACAAACGGACA-TAAACTGGCAAA--GTT-ATGATTTTTTGAAAA
C. elegans    1792  G-----CGGAAGAG---AATGGAAAAGGCGGGC----GGG-----AA---CCGGAA--AACGGAGAGTTTCCTC-CAAAAAGTTTATGAGTC---GAAAA

C. briggsae   2531  TTTAGGCATTTTTTGGTA-CTTTTTGGACCATAACTCCCATA-AGA-AATTTTTTACCAACTATTCCGCTTGGAATCTAAAATAACTTTTCACTCTCTAT
C. elegans    1865  ---AAGAAAGACGGGTTAACGTCAT----CATAGAT----TAGAGACAATACTTGAC----TATT----------TCTAAAAT----TTTAATTTTCTGT

C. briggsae   2628  CC----AACCAT-ATAGGTCTAT-ATTTTT-TCCGA-GGTTAACACTTGGGAGTCATTCCATGTGAGTCACACATACAT-TTTAGAAAAA-TAAGTACCG
C. elegans    1937  CCTTGAAATTATTAGAGG-CAATGATTTTTGT-GGAAGGCTCCC-CTT-------ATTCCTT---A---AAACTTCC-TCGGTCGACAAGCTCA-T-CC-

C. briggsae   2718  AATCACAGT--GTTCAGAAAGGTCT-GATCTGAAATCGCTCCCCAGTG--AACG-TCAGATT-GAAGCTTCTTTTTTCTTCTTTTCGTTGGC--TAAGTA
C. elegans    2016  -ATCACC-TCCG-TCA--AACTT-TGGAACTAAAAT---TTC--AGTGTTTTTGGT-AGTTTTGCAGCCAATATTTT-TTCTTTA-----GCAATAGGAA

C. briggsae   2809  TTAAACGTACAACAAACACACGCCAAATCCACGAACACTTTGAAATTGCTAA-ATTTCATCTGGCCTAAAAGTTATCTGGGGAGCT-ATTTCTGACCAGA
C. elegans    2099  TTGAA-GTTTTTCA------C-CCAAA----CGAA---TTT----TCTCTTACA-GTCA---------AAATTT-----GGAAGCGCATGTTTTA---GA

C. briggsae   2907  CATTTTGGAACATTGTGCTGAGTATTCAAGAGTTTCGAATATATGAAAAGAAGTAGTGAACATTTTGGCATTTGGCATATGATGTCCCATTTTATGACCA
C. elegans    2161  --TTTTGTTAAA--------ACTATTCAC----TCCCAAGGT-TGA-------TCC--AATATTCTCTCAAAA--CAGATGGTGTGAAATTTTATAACCA
```

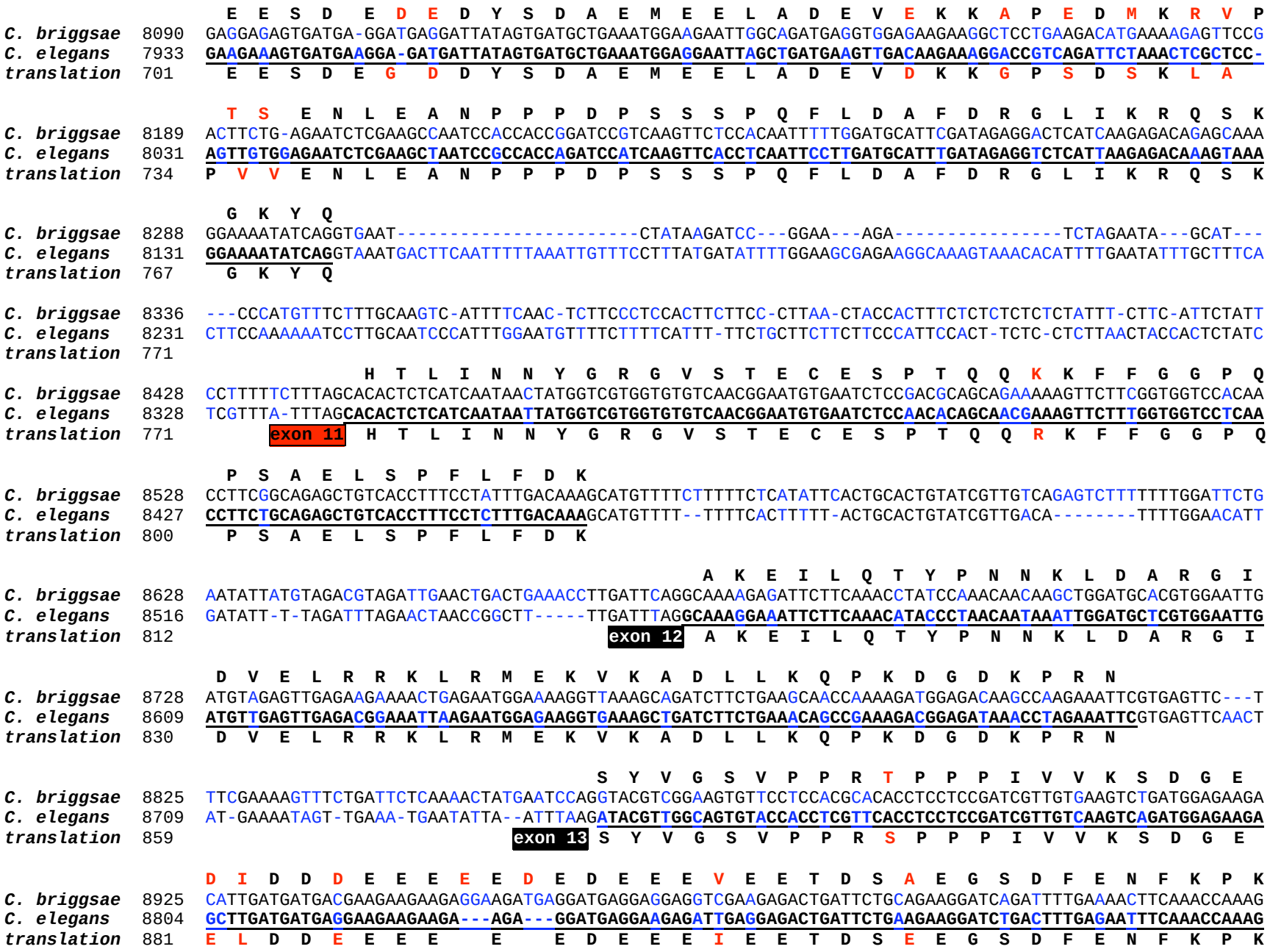
```
C. briggsae   3007  TATCTCGTCAGTTATTTTCACATTCT-CCCC-TCTCTGTTCTCAGCTGGTGCGTCAGTTTTGTCACTGAGCATTTTCACTCGTTTCAAAA-CCCATAAAC
C. elegans    2236  AATTTTGTCAGTTATATTCACATT-TGCTCAAT-TCTCCT-TCAGCTGAAGA---AGTTT----ACTTCTCGTTT-CTCCCGTT-CAAAAACCCATAAAC
translation   37

C. briggsae   3104  CGCAGCCCGTTTATTGGTCGT-GTCCTCAGATTCTTTCCTCC-TTCTCT-ATTGTTTTCTGATTTACTCGTCGTGTTCACTTTGGATCAGAGAGTCTTCT
C. elegans    2324  CGCCGCC-GTTTATCGTTC-TCTTGTTC---TTTTTACAAAGATTCTCTCA-TCTTTT-TGCT---CTTTTC-TATTCTC---GGATCA-----TTTTC-

C. briggsae   3201  AAATATTATCAGTCTTTCCATCAGAATGCCGTTGAGCTCATTTCTCTGTAACGCGGGTAAAATTGTCAAATGTGTTCGAAGAGCAATTTC-ATCAAATGA
C. elegans    2403  -AAT-TTCTCAGTGTGT---TG-GAATGACGTTAAGCTCATTTCTATGTAATGCGGGTAAACTTGTCAAATGCGTTCGAAGAGCAATTTCCATTCG-TGA

C. briggsae   3300  TCACG-AGCAGACTGATCGATGTATTGACTATTGCAACAATAATAGGCAATGCATTGTCACTCATGAGCCCACATGCTCCAAGTTCGTTTATGGTATAGT
C. elegans    2497  TCAAGCAGCCGAA-GATCAATGTATTGATGATTGTAATAATAAT---CATTGCGTCGTGACTCATGAACCAACATGCTCAAAATATATTTATGGTATTTT

C. briggsae   3399  T---CTGTTT----------------------------------------TGATCACG---AGAGA-----CT----------------------------
C. elegans    2593  TTTACTGTTTCAATTTCTTAATTTTAACAAATCAGCTCTCAAAAATATGATC-CGTTCAGAGAAAAAACTTTATGCGAAAAATTGTCAATAAATAAGAAA

C. briggsae   3421  AC--------------------ATGTT---TT---CAGTTTTCCAG----------------------------------------ATTGAT---------
C. elegans    2692  ACGGTGAAGTTTGAAAAGTGTATATGTTAAATTGCTCAGCTTTTGAGCTTTGAAAAAAAAATTTCACTAATTTTTGATAAATATTATTGATAGTACTGGG

C. briggsae   3446  --CAA---ACTTTTGG-----AAT-----CAATATTT--------TTCCTG----------------------------------------TTGA-----
C. elegans    2792  TCCATGTGACTTTTGGGACAAAATATGGACAAAATATGATGAGGCTTCATGATGGTAGTGGGATCCCATTATTCATAATTAAAGATAGAATTTGAATAAT

                                                                  G  T  H  P  M  S  G  A  E  P  C  S  P  T  K  K  E
C. briggsae   3478  --------CCT----------------AAA---ATTGTAATTTGTAGGTACTCATCCAATGTCCGGTGCAGAACCATGCAGTCCCACCAAAAAAGAGAA
C. elegans    2892  TATGGAACCCTTATTTTGATTTTCCGTTAAATCAATTTATTTTCCAGGAACCCATCCAATGTCCGGTGCTGAACCATGTAGCCCAACAAAAAAGGAAAA
translation   37                                              exon 2 G  T  H  P  M  S  G  A  E  P  C  S  P  T  K  K  E

                      K  H  R  F  E  I  T  K  K  L  G  S  G  T  Y  G  K  V  S  L  A  Y  D  H  K  F  D  R  E
C. briggsae   3551  GCATCGATTCGAGATCACAAAGAAGCTTGGCTCTGGAACATATGGAAAAGTGTCTCTGGCCTACGATCACAAATTCGATCGAGAGGTTTGTTCGTCTTCT
C. elegans    2992  GCATCGATTCGAAATCACAAAAAAACTTGGCTCCGGAACATACGGAAAAGTATCGTTGGCATATGATCACAAGTTCGATCGAGAGGTTCGTGTTCTTTTT
translation   55    K  H  R  F  E  I  T  K  K  L  G  S  G  T  Y  G  K  V  S  L  A  Y  D  H  K  F  D  R  E

C. briggsae   3651  ATCTTTA----------------------------------AAGACGTCAATCTACTTTTTATTGTCGTCGATTTTACGTCTCTAGGGATCGAATTGA
C. elegans    3092  TTTTATTTATTCGTCCATTTTTTAAACATTTGGCATATTCTCAAAGACGTCAATCTACTTTTTATTGTCGTCGATTTTACGTCTCTTGGGATCGAATA-A
translation   83

C. briggsae   3715  AATC--TTTTTTGTTGCTCGTTCTG--TGGGAACGAGCAAAAAAAAAATTCAATCAAGATTTGTCAGA-------AGTTTT-----CTCCAGGACC----
C. elegans    3191  GATCCATTTCTCGTTCTTCTTTTTGGGTGGGAACGAG----AAAAAAATTAAATAAA-ATCTGACAGAGAATGAAAGTTTTGAGATCTTGAAGATGTTTG

                                                  V  A  V  K  L  I  K  K  S  A  I  E  S  K  A  D  L  V  R  I  R
C. briggsae   3794  --AAAAATAG-----AAATCA--TG--TTTTGTAGGTTGCTGTTAAATTGATCAAAAAGAGCGCAATTGAGAGCAAAGCTGACTTGGTTCGGATTCGGAG
C. elegans    3286  ATAAAAATAGAAACCAAAATAAATTCAATTTT-AGGTTGCTGTCAAATTGATCAAAAAAGCGCAATTGAAGCAAAGCTGACTTGGTTCGAATTCGAAG
translation   83                                exon 3 V  A  V  K  L  I  K  K  S  A  I  E  S  K  A  D  L  V  R  I  R

                      R  E  I  R  I  M  S  A  L  N  H  P  N  I  I  Q  I  Y  E
C. briggsae   3884  GGAAATTCGTATTATGAGTGCATTAAACCATCCAAATATCATTCAGATTTATGAAGGTACTGCAATGAGGGATGTT-CTCAAAAAT-ACATTCGTTTTGC
C. elegans    3385  GGAAATTCGTATTATGAGTGCATTAAACCATCCAAATATTATTCAGATTTATGAAGGTACA-CAAT-ATAAATGAGACTATTAAATTATATAAGATTT-C
translation   105   R  E  I  R  I  M  S  A  L  N  H  P  N  I  I  Q  I  Y  E
```

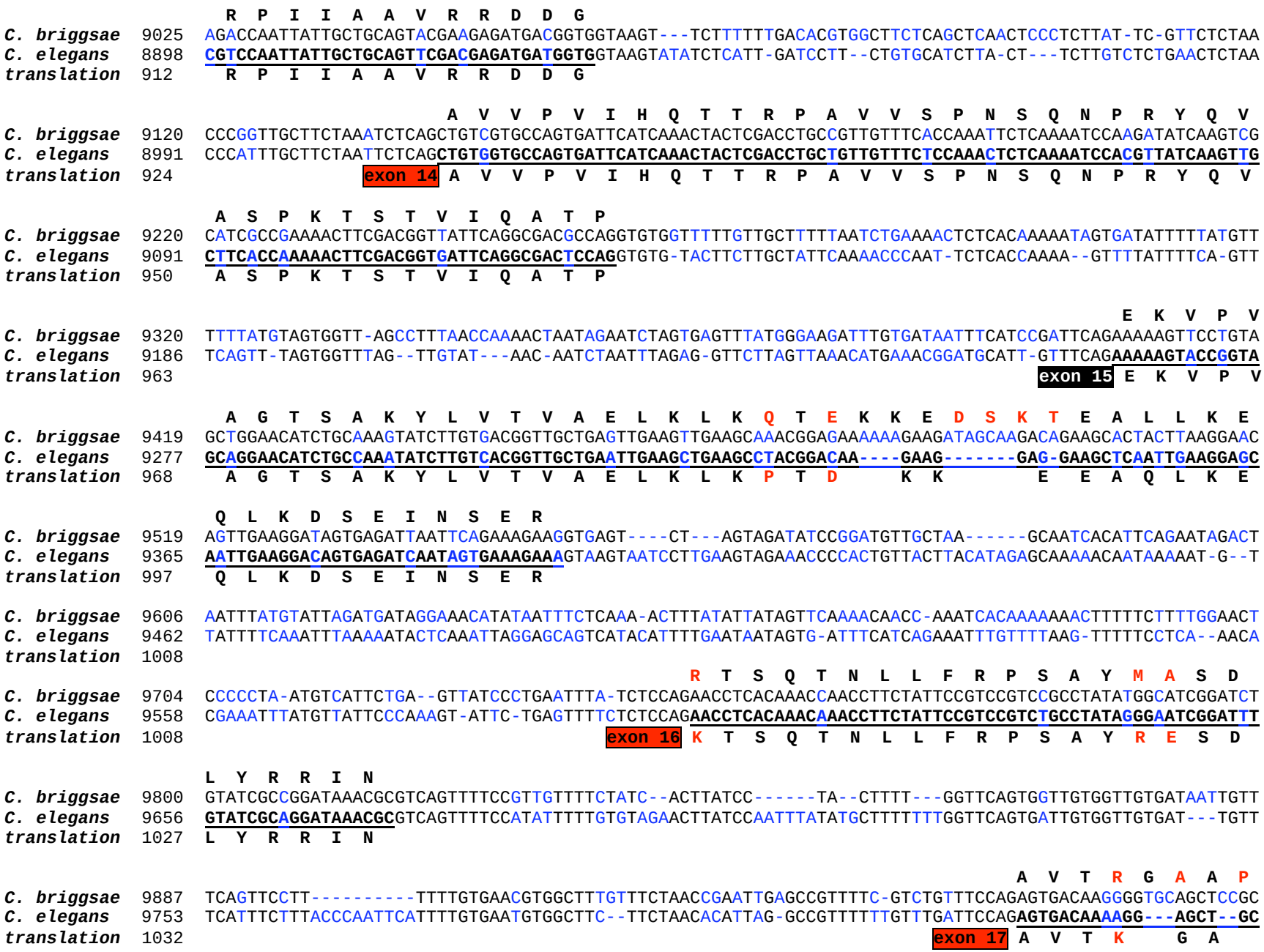

V F E N K D K I I L V M E Y S S G G E L Y D Y V S R C G S L P E A
C. briggsae 3981 AGTATTTGAAAACAAAGACAAAATCATTCTCGTAATGGAGTACTCTAGCGGTGGAGAACTGTACGATTACGTCTCAAGATGTGGATCTCTTCCAGAAGCA
C. elegans 3482 AGTGTTTGAAAACAAAGATAAAATTATTCTCGTAATGGAGTACTCTAGCGGTGGAGAATTGTATGATTATGTCTCCAGATGTGGATCTCTTCCGGAAGCA
translation 123 exon 4 V F E N K D K I I L V M E Y S S G G E L Y D Y V S R C G S L P E A
E A R R I F R Q I T S A V L Y C H K H R V A H R D L K L E N I L L
C. briggsae 4081 GAAGCACGCCGAATCTTTCGACAAATCACTTCTGCCGTCCTTTATTGCCATAAACATCGGGTTGCACACAGAGATTTGAAATTGGAGAATATTCTTTTAG
C. elegans 3582 GAAGCTCGTCGTATATTCCGACAAATCACTTCTGCAGTCCTTTATTGCCATAAACATCGGGTTGCTCACAGAGATTTAAAATTGGAGAACATTCTTTTAG
translation 156 E A R R I F R Q I T S A V L Y C H K H R V A H R D L K L E N I L L
D Q N N N A K I A D F G L S N Y F A D K
C. briggsae 4181 ATCAGAATAACAATGCTAAAATTGCCGATTTTGGATTGTCCAACTATTTTGCTGATAAGGTAAGCAGTCATTTTTTGAAGTCTATATTTCAAATGTCTAT
C. elegans 3682 ATCAAAATAATAATGCAAAAATTGCCGATTTTGGTTTGTCCAATTATTTCGCCGATAAGGTGAG--G--A--TTTT--AGT------TTGA--TGGATAT
translation 189 D Q N N N A K I A D F G L S N Y F A D K
N L L T T F C G S P L Y A S P E I I N G T P
C. briggsae 4281 TTCAATAGTCTAAACTATATTTCAAATTTTCCAGAATCTCCTAACAACATTCTGTGGAAGTCCATTGTATGCTTCTCCAGAAATCATCAATGGAACTCCT
C. elegans 3766 C-CAAT--TTTAA----TA----AAATTTTT-AGAATCTTCTGACAACATTCTGTGGAAGTCCATTGTATGCATCTCCAGAAATTATCAACGGAACACCA
translation 209 exon 5 N L L T T F C G S P L Y A S P E I I N G T P
Y K G P E V D C W S L G I L L Y T L V Y G S M P F D G R D F N R M
C. briggsae 4381 TACAAGGGTCCTGAGGTCGACTGTTGGTCTCTTGGAATCCTTCTCTATACACTAGTTTATGGAAGTATGCCTTTTGATGGAAGAGATTTCAATCGAATGG
C. elegans 3854 TATAAAGGTCCAGAAGTTGATTGTTGGTCACTTGGAATTCTTCTCTATACTTTAGTTTATGGAAGTATGCCTTTTGATGGAAGAGACTTTAATCGAATGG
translation 231 Y K G P E V D C W S L G I L L Y T L V Y G S M P F D G R D F N R M
V R Q I K R G A Y F E P E T P S
C. briggsae 4481 TCAGGCAGATCAAAAGAGGAGCGTACTTTGAGCCAGAGACGCCTTCAAGTAACCTGCTT----------TATCA---AAAAGTTGGAATA---------
C. elegans 3954 TCAGGCAAATCAAAAGAGGAGCCTATTTTGAGCCAGAGACACCGTCAAGTAAG-TGATTAAAAAAATGTCTATCAGAAAAATTTCAAATAATCAGCAAA
translation 264 V R Q I K R G A Y F E P E T P S
A A S M L I R N M L R V N P
C. briggsae 4557 ----------TTA---TAATTCGA------TTGA--TTTCAGCTGCCTCAATGCTCATCCGTAACATGCTTCGAGTGAATCCA
C. elegans 4053 AACAGCAAGCTATGACATGGTTAAACATTACGGTAATTCTAATTTAGTTCAATTTTCAGCGGCATCAATGTTAATCCGAAACATGCTTCGAGTAAACCCG
translation 280 exon 6 T A S M L I R N M L R V N P
E R R A T I F D I A S H W W L N L E E N M P V I Q E L P E N Q I I
C. briggsae 4620 GAACGAAGAGCTACAATTTTCGACATCGCTTCCCATTGGTGGTTGAATCTTGAAGAAAACATGCCAGTCATTCAAGAACTTCCTGAAAATCAAATCATAG
C. elegans 4153 GAGAGAAGAGCCACCATTTTTGACATTGCATCCCATTGGTGGCTGAATCTTGAAGAAAATATGCCTGTGATTCAAGAACTTCCTGAAAATCAAATTATCG
translation 294 E R R A T I F D I A S H W W L N L E E N M P V I Q E L P E N Q I I
D H T P L T E R E E T M V V Q D L A D E Q D V F M E F G H L S S E
C. briggsae 4720 ATCACACTCCATTGACGGAAAGAGAAGAGACAATGGTGGTACAGGATTTGGCAGACGAACAGGATGTCTTCATGGAATTTGGGCATCTTTCTTCGGAGAC
C. elegans 4253 ATCACACTCCTTTAACGGAAAGAGAAGAGACAATGATAGTACAAGATCTCGCTGATGAACAAGACGTTTTCATGGAATTCGGGCATCTTTCATCAGAGAC
translation 327 D H T P L T E R E E T M I V Q D L A D E Q D V F M E F G H L S S E
T R R K I E D F R R R R K E A E E F N D N S P V K P P K T R K T D E
C. briggsae 4820 ACGTCGCAAGATCGAAGACTTTAGAAGACGTCGAAAGGAGGCTGAGGAATTCAATGACAACTCTCCAGTCAAACCACCAAAGACTAGAAAAACGGATGAG
C. elegans 4353 TCGTCGCAAGATTGAAGACTTCAGAATACGAAGAAAAGAGGCTGAGGAGTTTAATGATAATTCACCGGTGAAACCTCCAAAAGCAAGGAAAACTGATGAA
translation 361 T R R K I E D F R I R R K E A E E F N D N S P V K P P K A R K T D E


L T G K V A K E Q P E E M R S A E K S L R G V K E E K E K P K V V
C. briggsae 4920 TTGACCGGAAAAGTGGCGAAGGAGCAACCAGAAGAGATGAGATCTGCTGAGAAATCGTTGAGAGGAGTTAAGGAGGAAAAAGAGAAACCGAAAGTGGTGG
C. elegans 4453 TTAACTGGAAAAATTTCGAAAGAACAACCGGAAGAGATGAAGTCTGCTGAAAAATCTCTGAGAGGAGTAAAGAGGAGAAAGAAAAACCGAAAGTTGTTG
translation 394 L T G K I S K E Q P E E M K S A E K S L R G V K E E K E K P K V V
D P N D P L E R L R Q I E N R L G Q Q K K E K E
C. briggsae 5020 ATCCCAATGATCCATTGGAGAGACTTCGACAAATCGAGAATCGATTGGGGCAACAGAAAAAAGAAAAGAGGTTTGTATAGATTTTTTTTGTTGGTTTAT
C. elegans 4553 ATCCGAATGATCCACTGGAAAGACTCAGACAAATTGAAAATCGATTGGGACAAAATAAGAAAGACAAAGAGGTTTGATTATTTATTGTTTAAAAATGAAT
translation 427 D P N D P L E R L R Q I E N R L G Q N K K D K E
C. briggsae 5120 T-GAACA-AATTCGGTTCATATCTATCACTGGCAGTAACCTGTAGGGGGTCAAATGAGACGAAGTTGAATTAATGTTATTTCAAATTGTTTTGAAATCAT
C. elegans 4653 TAGTACATAACT-GGTAAATGTTAATT--TGGCAGTAACCTGTATGGA-TAAAATGAGAC----------AATGCTCC--CAAA--GA----AAATCAT
translation 451
C. briggsae 5218 TTTTTGAACCCATCATGACAATCTTGAATGCATTTCAATGAAGACGTCTTCAACTGAATCGATGTGTATTCTGGGTGTAGTCTGACCGTATCCAAGTAGC
C. elegans 4730 TTTCTCATTGCATCATGACAATCTTGAATACATTTAAGTGATGTAGTTTTAACTGAATTGAAGTATATTCTGGATGTTGTCGCATTGTGTCCAAATAAC
C. briggsae 5318 TCTCTGCACATTGTTCTGCACACGTCATCTGAAATTCAAATCTTTGTATTCTAAATAGTTGTAAAGATATGTTTGTT-ACCTCTGACCAAACGATCGAG
C. elegans 4830 TTTCTGCACACTGTTCTGCACAGGTCATCTGAAA---AATTAGTTAGAGA-TAGATAATT-TAA----TTTGCTTGTTTACCTCTGGCCAAACGATCGAG
C. briggsae 5417 ATCAGTACCAACACGATCATGAGAGACAATACAGTAGACTTCATCGCGCGGGGGTTTTGCAAGGGTCAGCGAACAGTTAGGCCGATTGTACGGAAGTTTTT
C. elegans 4921 ATCATTGCCAAAAGAATCAAAAGTGATAAAACTGTAGACTTCATC---------TTTGCAAGGGTCAGCGAATACGTAGGCCGATTATACGGAAGTGGTT
C. briggsae 5517 TTATGAAAGAAATCACACG--AC---TC-TGCAAATT--------GA--GTAT----------GAT-TTTGGACTCGGAATAAAATAGGGAAGAAGT-TG
C. elegans 5011 ----GGAGGAGATCACACGCTACAAATCATGCAATTTTTGTACTCGACTGTATTTATAGAAATGATCTTTAAAATAAGAATAGAAAAGTGATACGGTATC
C. briggsae 5589 ATCTTTAGTTTGGGAGGATTTAAGTG-ATAGGTAAGAAATAA-GAAATT----------TTA--GAAAAAAC-AATAT------ACCGTACAATCTG--
C. elegans 5108 AACTTTATCATGTAATGATGATATTGGATTATTTCGAAGTAAAGAAATTATGTAGCTTTTTTAATGAAAAAAATAATTTTAAAAAATAGTCGAA-CTGCG
C. briggsae 5664 -TAATACACAGAAGGTC---CATTAACTTTTAATTTAGAGC-T-AATTTATCTCGCTTCCAGTCAAAATTGTCAGC-------TGAT-AATTTGTTAT--
C. elegans 5207 CTTCTAATCGTAAGGCCTGCCAGGAATTTATACTAG-GCGATTAAACTGATCTC--TTCC--TCAAA-TTGTCACCCTTTTTGTGAGGAAATACATACAA
C. briggsae 5748 -CATAATGTAAATACAAA-CATTTTAT--CTGTTAACAT-ATTTT----TGA---------TATC-TCA-----CATAGGAACCGAGATA------ACTG
C. elegans 5301 ACAGAAACTAGACATAAATCAAGGTATTACTTTCAAAATGATCCTCACTTGAACCCGACAATATCGTCATTACACATAGTAATCGAGATCGGTCGTAGTG
C. briggsae 5818 ----------------AAATACA--GGGTG------TGCTTC--------TATTAA-------AGCAAATACGGTATAAATTAGTAACATTCACAGTGAA
C. elegans 5401 GTTTTTGCAGATTTCAAAATAGAACGGGTCAATTTATGATTCACACTGTTTGTTAATTTTCAG----------
C. briggsae 5880 ATATGAAGGAGGAATTTGAAATGACAAAAATCAAAAAGTTTTCAGAGATTACCTGTTTTTATGGGAGGAATTTTATGATTCAGATGACCTGTTCCAAACT
C. elegans 5463 ----------
C. briggsae 5980 GAAAGAAAACTAGACAAAAATTAATCTCCTCTTGAACCCGAACGTCATCGTTTGACAGTTTCTATGACTTGACCCTTTTGGTTCCAGAATTTCATAACAA
C. elegans 5463 ----------
E F Q A A A A A R A E A V K E V K
C. briggsae 6080 ATCACATTCGAAATGAATCATGAAAAAAAAAAACATTGATATTCTTCGAATTTCAGGCAGCTGCGGCAGCTCGCGCCGAGGCAGTCAAAGAAGTGAAAAA
C. elegans 5463 ----------GCAGCAAAAACATCCGCCGCGTTGAAACAGTAAAACTTAAAGAGGTGAAGAA
translation 451 exon 7 A A K T S A R V E T V K L K E V K

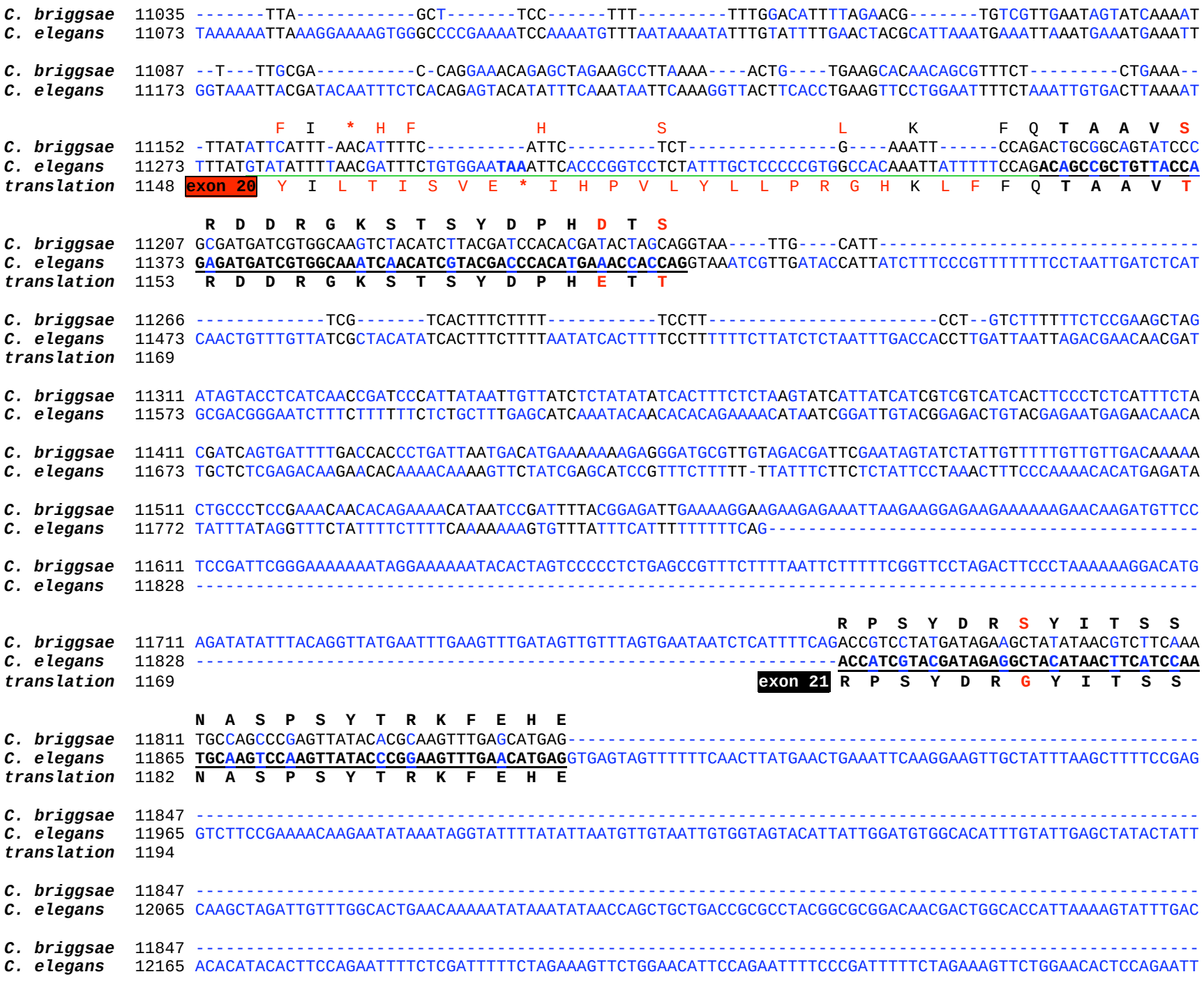
```
                      K  A  K  S  P  E  Q  Q  P  E  D  P  K  T  A  R
C. briggsae     6180  GGCAAAATCACCAGAGCAACAACCGGAGGACCCGAAGACAGCGAGAGGGTGAGT----------------------------------------------
C. elegans      5517  AGAAAAGTCTCCAGAGCAACAACCAGAAGATCCGAAAACTGCAAGAGGGTGAGTGGTTTTGATGTTTTTAAGAGAGGGATGACCGTTTAAGCAGGTGTTT
translation     469   K  E  K  S  P  E  Q  Q  P  E  D  P  K  T  A  R

C. briggsae     6233  ----------------------------------------------------------------------------------------------------
C. elegans      5617  TATTTGATTGGGGAATACGAAAAAGTCATGTAGTCTATGCGTGATACGGGAACTTTGAGCTCAAGTTGAAAATGTATGAGGCTTATTATTGAGTGAGATG
translation     484

C. briggsae     6233  ----------------------------------------------------------------------------------------------------
C. elegans      5717  TTAAACAATCCATTTTGCAAAAAAAAACTGAAGGACTATAATACTATAGTATATTCAAAAAAAACTCGCTCAAAAAAGACAGCAAAAATCTTGAAATTGA

C. briggsae     6233  ----------------------------------------------------------------------------------------TGGGGAA
C. elegans      5817  AAAGCAAAACAAGCAATAATTATTATTTGGCAGATTTTGGAAAAAATATTCTGACACCTAAACTTAGAAGAAATATCATCTTGACCACAATTTTGAACAA

C. briggsae     6241  AGACTTTGATGATTTAAAA----AG-TCACAAGTAACAGGT---GTTTTATT-TGAT-----TGA---TAAAA---AATATGTGGGTCA---TTCAAAGA
C. elegans      5917  ATTCTT-GAAGCACTAAAAGTAAAGGTTACAATTCCAAATTAAAGTTTTATTCTGATCCAAATCAATTTAAAATTAAACAAATTTTAAAATTTTCGATAA

C. briggsae     6318  AT-GATCACAAAA-----ACATTTA--TCTGAGTTCAG-CTGCCCT-CTGGTTAGAAACGGA-CA----AAAC-----------AACTTTGCAGATATGT
C. elegans      6016  GTTGCTTAGAAAATGTAGAAATAGAGATTTG-GTATATACTACATTACTT-TAAGAAACCAAACATTCGAAACCTTTTCTTTCAATTTTGT-GACTTAT

C. briggsae     6382  GTTGATCCC-----CCACTGGGAGCCGAATATTTTGTGA--CT--TGAA-GC-ACAAGTTACAC----------TAAT--------------TCTGGTG
C. elegans      6113  GGTGCTTACAAATACGACTG--AGCACAAAATTTTGAAATTCTGATGAATGCTACATATT-CACGTCAGATGTATAATAAGCATGAGATTCATCTATTCG

C. briggsae     6456  ATTTT-TAATGTTGCTATTA-ATTTTCTCCAGTTCAA----------TAA-CTTTGTTAA--T----TA--GTGAAGAAGTCA-GAAG--TTGTGC-CCT
C. elegans      6210  ATTTTCTAACAT--CTATTCGATTTTCTAAAAATAAAAGAATTGTTGTAAACTTTTTTAAACTAAAGTATCGTTTACAAATCAAGCAGCGTTGTTCTCCA

C. briggsae     6530  --TTATGAT-TATCAGA-GAA-TGTCAGACC--GAAA--AGA---GAAATGAA--ACT----GAAC-AGTTGGTGTT---TACAAAACTAAAAAT-T-CC
C. elegans      6308  CATTATGAGCTAGCTGTTGAACTATATGATCTTGAAAGAATACTTGAATTCAACGACTTCTAGAACGAACTCATGTCGGGTACACACTACTTTTGTACCG

C. briggsae     6607  GAAAACCGAAAAGAGGTGTCAAATAG--GGATTGCCGGT-TCGATATGAAAAA---ATAGGGGGAAT-GATG--CGA-GATTTGGGAATAAATCATAAAA
C. elegans      6408  AAAAATAAAAATATCTGGTACATTGTAGAATTTAAAGTATGGAACTGAACAATTGATTGTTTTTATAGTTGAACATCGATTTTGAAATAGGTG-TCAAAT

C. briggsae     6697  AGGTACATTGACTAG--AAAA--TGG-A-TGACAACCGCTAAAAA-AAGGGA---TTACCCGCGGGTGTT-TAGAC-CATTC-TCATGTCTTTTTCTCTC
C. elegans      6507  GATGGGTTGAAAGGGGAAAAGGTGGGAGTGAGAACCGATCATGCGAAGGCACATTTACTAGAAACAAAAATGGACTCAATCGTCCGCTAAATAGAGATTA

C. briggsae     6784  GC-GTTTTTATGATAAATCCGCCTTCA-GTCCGTTACACTTTGATCTTTTTTA-CGATGGCCTT--------TTTTCGG----G--AGAGCGTGCAATAG
C. elegans      6607  CAGTTTTTTCGAG---TCT-CTTTCATGTCCGATACACTTTGATCCCTTTTAACGATGGCCTTGTTCAAGTTTTTAAGTTTTGTAATAGCCAACAACAGA

C. briggsae     6867  TTGGAAACAAAATGTTGGGGGGA-AAAAGAGAAGGACAGCGTCCCTCCGCTCTGAAAAAAGGCTATTGCTATACGGATTTGGTGTCGGCTTTGAAAGAAA
C. elegans      6703  AGTGAATATTGTGTGTGGGAGAGAAAAG-GACGGACAGCG----------CAAAAAATAG-CTGTTGCTATACGGAAT--GTCT-GTTTTTGAGAGGAA-

C. briggsae     6966  AAAGGAGAAACTAACG-----GGCTA--------TTTT----TGGTTACTGGGAAACGGCAGCACGCAAACAAGCGGCTAGACTAAATGTTGAACGT---
C. elegans      6786  ---GAGAAACACACCTCAAAGGATAGAGACTTCTTTTCGGGTGGTTACCTGGATACGGGATTA-------AAGAGGAAA-ATGAACTTGTGAACATACGA

                                                                  G  T  S  K  P  A  D  S  R  A  P  S  F  V  P  V  K  D  K  P  E
C. briggsae     7045  --TTTCGTGGGCAAAAAACAAATCGGATTATTCTTTCAGAACATCAAAACCAGCGGACTCGCGTGCTCCTTCATTTGTTCCTGTCAAGGATAAACCAGAA
C. elegans      6876  CTTTTTGATA-CCTAAAACAA----GAA-ATT--TTCAGAACATCAAAACCAGCTGATTCTCGTGCTCCATCATTCGTTCCAGTTAAGGATCGTCCAGAA
translation     484                                  exon 8 G  T  S  K  P  A  D  S  R  A  P  S  F  V  P  V  K  D  R  P  E
```

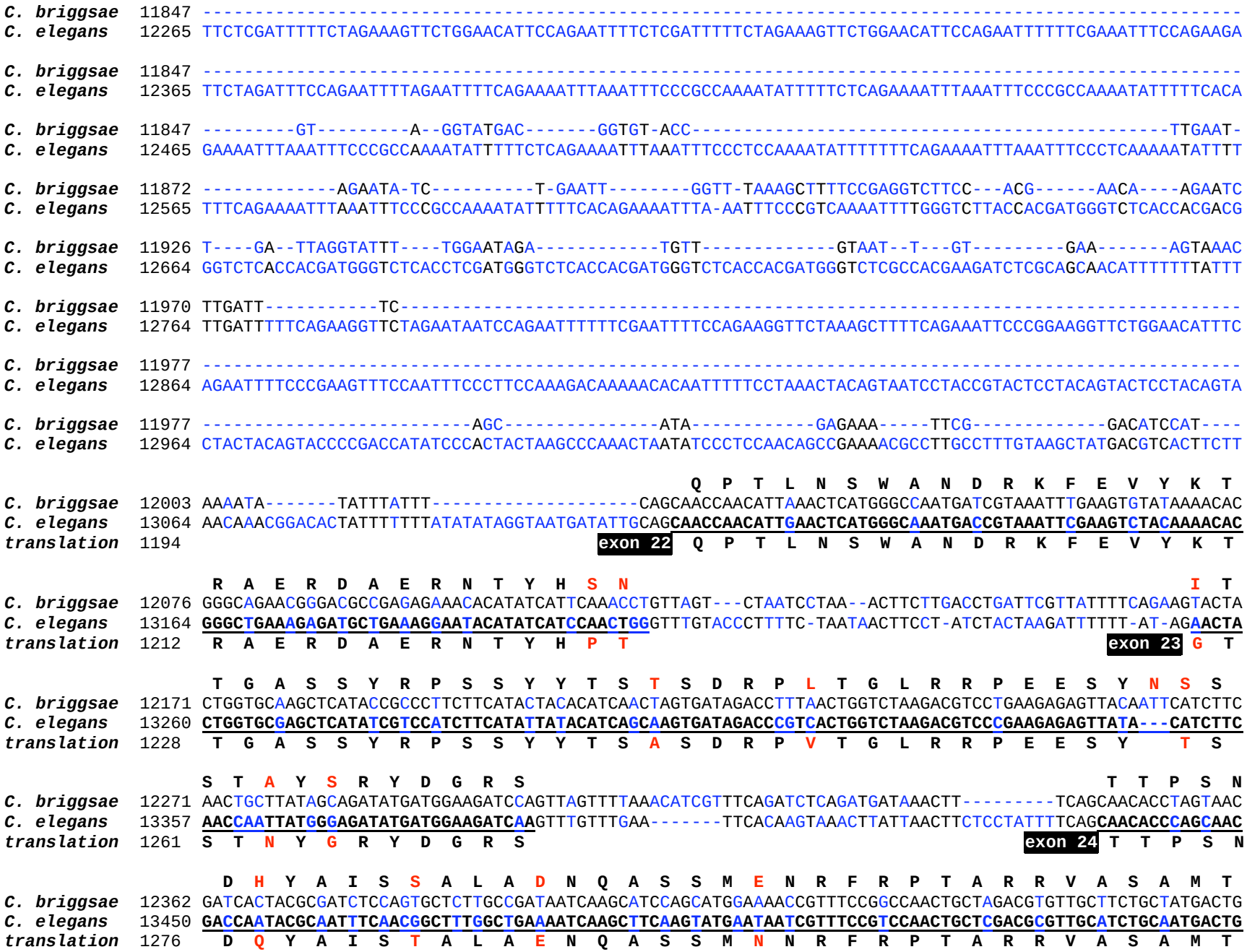

T S E P E R P R T R P H M T A S T Y R I E T D S L N M L M N Q V L
C. briggsae 7144 ACATCAGAGCCAGAAAGACCACGTACTCGTCCGCACATGACAGCCAGTACCTACAGAATCGAGACTGATTCATTGAATATGCTCATGAATCAAGTTCTGG
C. elegans 6968 ACTTCTGAACCTGAAAGACCACGTACTCGTCCACATTTGACTGCAAGTGCCTACAGAATTGAAACGGATTCTTTGAATATGTTGATGAATCAAGTACTCG
translation 505 T S E P E R P R T R P H L T A S A Y R I E T D S L N M L M N Q V L
E Q M E K G P V N L N I I G R I K A H P L Y D T R P M V K E L L E
C. briggsae 7244 AACAAATGGAAAAAGGACCCGTCAATTTGAACATCATCGGTAGAATCAAGGCACATCCTTTGTATGACACTAGACCAATGGTCAAGGAACTTCTCGAGAG
C. elegans 7068 AACAAATGGAAAAAGGACCAGTCAATTTGAATATTATTGCAAGAATCAAAGCCCATCCACTATATGACACAAGGCCAATGGTGAAAGAACTTTTGGAAAG
translation 538 E Q M E K G P V N L N I I A R I K A H P L Y D T R P M V K E L L E
S I I A A Q P E P V Q K Q T S K V V E Q Q
C. briggsae 7344 CATCATCGCTGCTCAACCAGAACCAGTTCAGAAGCAGACGAGTAAAGTGGTTGAACAACAGGTAATTTG-----GCTTCAGAATACAATCTGTGCGCTCT
C. elegans 7168 TATTATTGCTGCTCAGCCAGAGCCCGTGCAAAGCAGACTAGCAAAGTGGTCGAGCAACAGGTAATTTCATTAAGCTACAACAAATAAAA-GTTCACTTT
translation 572 S I I A A Q P E P V Q K Q T S K V V E Q Q
C. briggsae 7439 ACTGAAAAATATATTGCAA-AAGAAATAGGAA--AATAGGAATTTTGGAAATATCAAATGTG--AGA----TTATT---TTCGTAAACAGAC-CTAACTC
C. elegans 7267 ACTGGAAG---TCTTGCGATAACTAA-ACGAATCAATATAAAAAATATAAACTTTTTATTTGCAAGAATCCTTATTCAATTATTAAAAAAAAACAAAATA
translation 592
C. briggsae 7526 TTTAATTAGAACAAAAGAACTGCGCTTCAAAATTCAGTAGAGCGTG-CTTGT-GCT---CGA-TG-ACCTCACA-TAGAACCCCATC-CAA---TCATA-
C. elegans 7363 TTAAA--AGAAAAATTG--CTAATTATAAAAGTTCAGTAAAGCGTAACTTTTTGCTTTCCAAATGCACTTCTCAATA-AACTCCCGCACAAAAATCAAAC
T F S R Q N T L T R K K K
C. briggsae 7612 ---TGCACTCCAAACTCAAGAACATAACATATGAATGATATGAAC--CTAACACGTTTCAGACGTTCTCCCGTCAAAACACTCTCACCAGAAAGAAGAAA
C. elegans 7458 CTTTGCACTTGAC-CT--AGA---TAACAT--GAAAGCTATGATTTTCTAACACTTTTCAGACATTCTCCCGCCAAAATACACTGACGAGAAAGAAGAAA
translation 592 exon 9 T F S R Q N T L T R K K K
E D P V E E E E I P A V P S P P T R K M K E R P W H S V E
C. briggsae 7708 GAAGATCCAGTTGAAGAAGAAGAGATTCCGGCCGTCCCATCACCGCCTACGAGAAAGATGAAAGAGAGACCATGGCATTCTGTAGAGGTGTGCATGAAGG
C. elegans 7550 GAGGATCCA-TT--AGAAGAGGAGATCCCAGAAGTTCCATCACCTTC-AC--GAAAGATGAAGGAGAGACCATGGCATTCTGTAGAGGTGTGCATGAAG-
translation 605 E D P L E E E I P E V P S P S R K M K E R P W H S V E V C M K
V G F D
C. briggsae 7808 AATGAAT----------AGAAAACAAAACATCCTTTAGAATGATCC-------AAAAAACCATTATTGCCCTTTGTATGTTTTCAGGTTGGATTCGATC
C. elegans 7643 AATGAATCAGTGCATTGTAGATA-CATAAT-TCCTTAAGAATGATTCTTTTTTCAAAATTCC--TATTATCCTGTGAATGTTTTCAGGTTGGATTTGATC
translation 632 N E S exon 10 V G F D
P D E E A E H D R M Q E S I A S N T T E V T V Q D T S F E D D S S
C. briggsae 7890 CGGATGAGGAAGCGGAGCATGATCGCATGCAGGAATCGATTGCAAGCAACACAACAGAAGTTACAGTACAAGACACGTCATTTGAAGATGATAGTTCTGA
C. elegans 7739 CTGACGAGGAACCAGAGCATGATCGCATGCAGGAATCAATTGCCAGCAATGCAACAGAAGTTACTGTACAAGATACTTCCTTTGAAGATGATAGTTCTGA
translation 636 P D E E P E H D R M Q E S I A S N A T E V T V Q D T S F E D D S S
D E E G R K K T P V A D T V P K T P K L I I E N P E S S K T V V E E
C. briggsae 7990 TGAAGAAGGAAGAAAGAAGACACCAGTTGCTGATACTGTTCCAAAGACACCAAAACTGATTATTGAGAATCCGGAGAGCTCGAAGACAGTTGTAGAAGAA
C. elegans 7839 TGAAGAAGAGAGAAAGAAGACGCCAGTTGC--AT-CAACACCAAAGACTCCAGTTCTAATTGTTGAAAACAAGAAGCTCAACGA-A--TGCGGAAGAT
translation 670 D E E E R K K T P V A S T P K T P V L I V E K Q E S S T N A E D

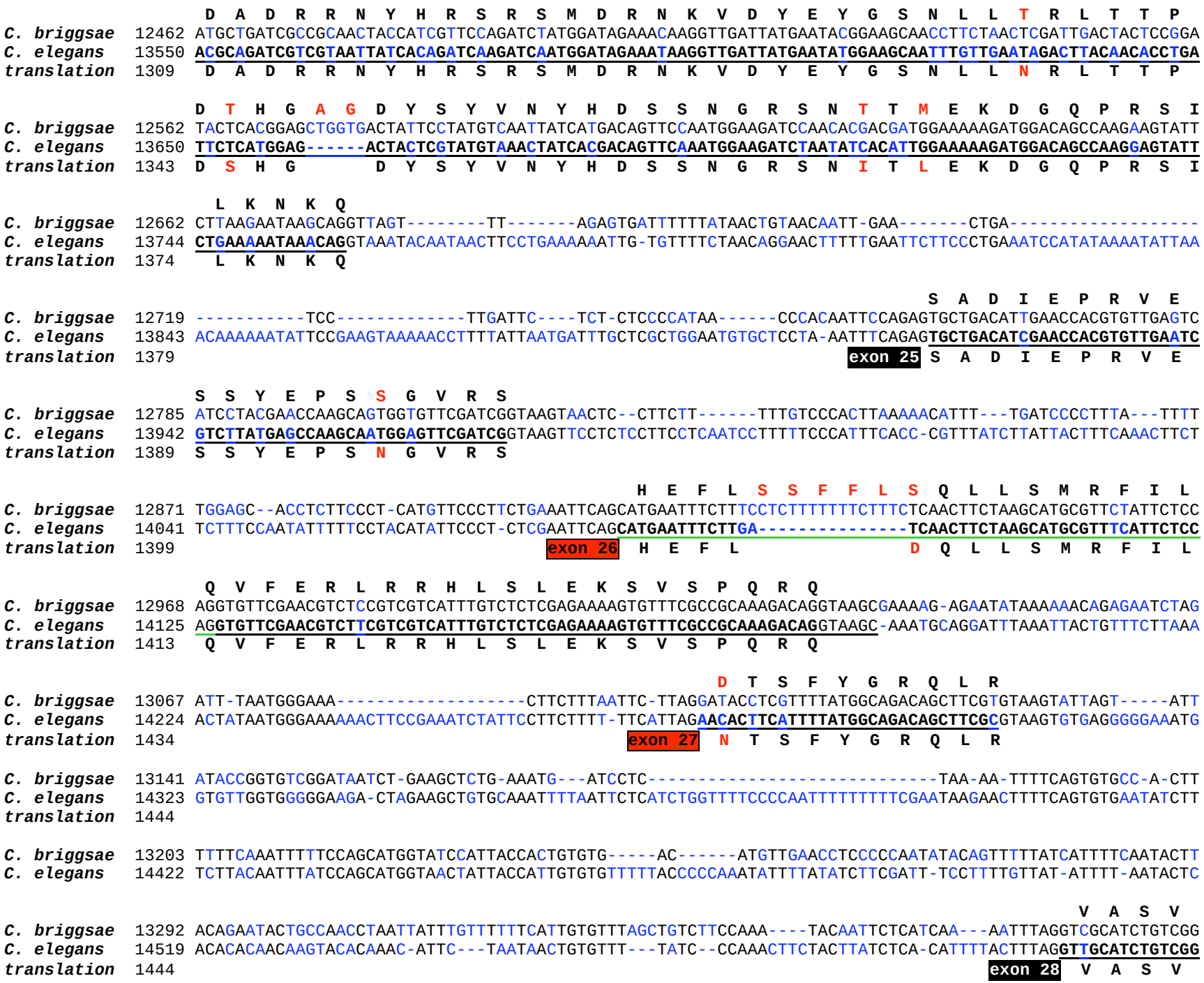

E E S D E D E D Y S D A E M E E L A D E V E K K A P E D M K R V P
C. briggsae 8090 GAGGAGAGTGATGA-GGATGAGGATTATAGTGATGCTGAAATGGAAGAATTGGCAGATGAGGTGGAGAAGAAGGCTCCTGAAGACATGAAAAGAGTTCCG
C. elegans 7933 GAAGAAAGTGATGAAGGA-GATGATTATAGTGATGCTGAAATGGAGGAATTAGCTGATGAAGTTGACAAGAAAGGACCGTCAGATTCTAAACTCGCTCC-
translation 701 E E S D E G D D Y S D A E M E E L A D E V D K K G P S D S K L A
T S E N L E A N P P P D P S S S P Q F L D A F D R G L I K R Q S K
C. briggsae 8189 ACTTCTG-AGAATCTCGAAGCCAATCCACCACCGGATCCGTCAAGTTCTCCACAATTTTTGGATGCATTCGATAGAGGACTCATCAAGAGACAGAGCAAA
C. elegans 8031 AGTTGTGGAGAATCTCGAAGCTAATCCGCCACCAGATCCATCAAGTTCACCTCAATTCCTTGATGCATTTGATAGAGGTCTCATTAAGAGACAAAGTAAA
translation 734 P V V E N L E A N P P P D P S S S P Q F L D A F D R G L I K R Q S K
G K Y Q
C. briggsae 8288 GGAAAATATCAGGTGAAT----------CTATAAGATCC---GGAA---AGA----------TCTAGAATA---GCAT---
C. elegans 8131 GGAAAATATCAGGTAAATGACTTCAATTTTTAAATTGTTTCCTTTATGATATTTTGGAAGCGAGAAGGCAAAGTAAACACATTTTGAATATTTGCTTTCA
translation 767 G K Y Q
C. briggsae 8336 ---CCCATGTTTCTTTGCAAGTC-ATTTTCAAC-TCTTCCCTCCACTTCTTCC-CTTAA-CTACCACTTTCTCTCTCTCTCTATTT-CTTC-ATTCTATT
C. elegans 8231 CTTCCAAAAAATCCTTGCAATCCCATTTGGAATGTTTTCTTTTCATTT-TTCTGCTTCTTCTTCCCATTCCACT-TCTC-CTCTTAACTACCACTCTATC
translation 771
H T L I N N Y G R G V S T E C E S P T Q Q K K F F G G P Q
C. briggsae 8428 CCTTTTTCTTTAGCACACTCTCATCAATAACTATGGTCGTGGTGTGTCAACGGAATGTGAATCTCCGACGCAGCAGAAAAAGTTCTTCGGTGGTCCACAA
C. elegans 8328 TCGTTTA-TTTAGCACACTCTCATCAATAATTATGGTCGTGGTGTGTCAACGGAATGTGAATCTCCAACACAGCAACGAAAGTTCTTTGGTGGTCCTCAA
translation 771 exon 11 H T L I N N Y G R G V S T E C E S P T Q Q R K F F G G P Q
P S A E L S P F L F D K
C. briggsae 8528 CCTTCGGCAGAGCTGTCACCTTTCCTATTTGACAAAGCATGTTTTCTTTTTCTCATATTCACTGCACTGTATCGTTGTCAGAGTCTTTTTTTGGATTCTG
C. elegans 8427 CCTTCTGCAGAGCTGTCACCTTTCCTCTTTGACAAAGCATGTTTT--TTTTCACTTTTT-ACTGCACTGTATCGTTGACA--------TTTTGGAACATT
translation 800 P S A E L S P F L F D K
A K E I L Q T Y P N N K L D A R G I
C. briggsae 8628 AATATTATGTAGACGTAGATTGAACTGACTGAAACCTTGATTCAGGCAAAAGAGATTCTTCAAACCTATCCAAACAACAAGCTGGATGCACGTGGAATTG
C. elegans 8516 GATATT-T-TAGATTTAGAACTAACCGGCTT-----TTGATTTAGGCAAAGGAAATTCTTCAAACATACCCTAACAATAAATTGGATGCTCGTGGAATTG
translation 812 exon 12 A K E I L Q T Y P N N K L D A R G I
D V E L R R K L R M E K V K A D L L K Q P K D G D K P R N
C. briggsae 8728 ATGTAGAGTTGAGAAGAAAACTGAGAATGGAAAAGGTTAAAGCAGATCTTCTGAAGCAACCAAAAGATGGAGACAAGCCAAGAAATTCGTGAGTTC---T
C. elegans 8609 ATGTTGAGTTGAGACGGAAATTAAGAATGGAGAAGGTGAAAGCTGATCTTCTGAAACAGCCGAAAGACGGAGATAAACCTAGAAATTCGTGAGTTCAACT
translation 830 D V E L R R K L R M E K V K A D L L K Q P K D G D K P R N
S Y V G S V P P R T P P P I V V K S D G E
C. briggsae 8825 TTCGAAAAGTTTCTGATTCTCAAAACTATGAATCCAGGTACGTCGGAAGTGTTCCTCCACGCACACCTCCTCCGATCGTTGTGAAGTCTGATGGAGAAGA
C. elegans 8709 AT-GAAAATAGT-TGAAA-TGAATATTA--ATTTAAGATACGTTGGCAGTGTACCACCTCGTTCACCTCCTCCGATCGTTGTCAAGTCAGATGGAGAAGA
translation 859 exon 13 S Y V G S V P P R S P P P I V V K S D G E
D I D D D E E E E E D E D E E E V E E T D S A E G S D F E N F K P K
C. briggsae 8925 CATTGATGATGACGAAGAAGAAGAGGAAGATGAGGATGAGGAGGAGGTCGAAGAGACTGATTCTGCAGAAGGATCAGATTTTGAAAACTTCAAACCAAAG
C. elegans 8804 GCTTGATGATGAGGAAGAAGAAGA---AGA---GGATGAGGAAGAGATTGAGGAGACTGATTCTGAAGAAGGATCTGACTTTGAGAATTTCAAACCAAAG
translation 881 E L D D E E E E E D E E E I E E T D S E E G S D F E N F K P K

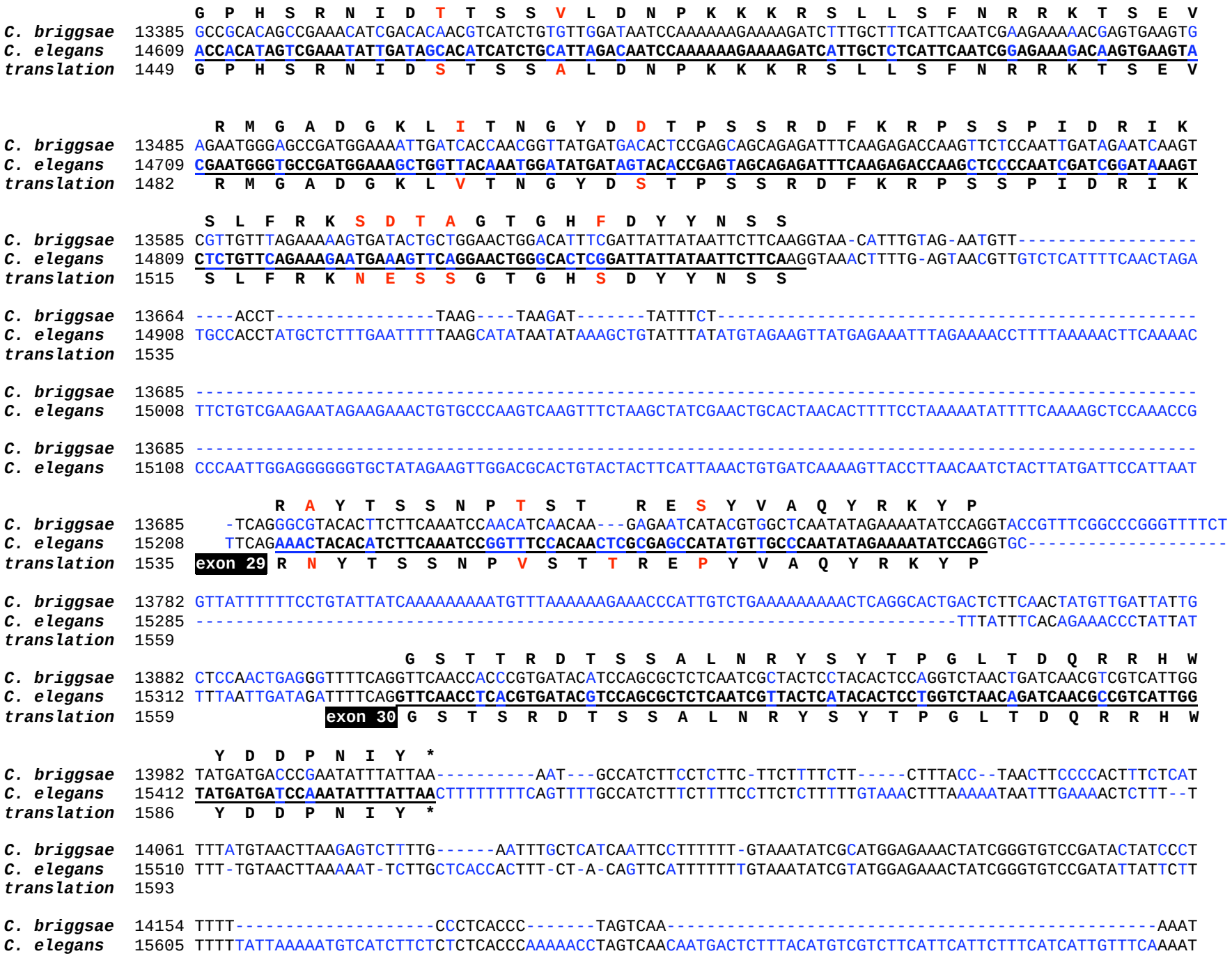

R P I I A A V R R D D G
C. briggsae 9025 AGACCAATTATTGCTGCAGTACGAAGAGATGACGGTGGTAAGT---TCTTTTTTGACACGTGGCTTCTCAGCTCAACTCCCTCTTAT-TC-GTTCTCTAA
C. elegans 8898 CGTCCAATTATTGCTGCAGTTCGACGAGATGATGGTGGTAAGTATATCTCATT-GATCCTT--CTGTGCATCTTA-CT---TCTTGTCTCTGAACTCTAA
translation 912 R P I I A A V R R D D G
A V V P V I H Q T T R P A V V S P N S Q N P R Y Q V
C. briggsae 9120 CCCGGTTGCTTCTAAATCTCAGCTGTCGTGCCAGTGATTCATCAAACTACTCGACCTGCCGTTGTTTCACCAAATTCTCAAAATCCAAGATATCAAGTCG
C. elegans 8991 CCCATTTGCTTCTAATTCTCAGCTGTGGTGCCAGTGATTCATCAAACTACTCGACCTGCTGTTGTTTCTCCAAACTCTCAAAATCCACGTTATCAAGTTG
translation 924 exon 14 A V V P V I H Q T T R P A V V S P N S Q N P R Y Q V
A S P K T S T V I Q A T P
C. briggsae 9220 CATCGCCGAAAACTTCGACGGTTATTCAGGCGACGCCAGGTGTGGTTTTTGTTGCTTTTTAATCTGAAAACTCTCACAAAAATAGTGATATTTTTATGTT
C. elegans 9091 CTTCACCAAAAACTTCGACGGTGATTCAGGCGACTCCAGGTGTG-TACTTCTTGCTATTCAAAACCCAAT-TCTCACCAAAA--GTTTTATTTTCA-GTT
translation 950 A S P K T S T V I Q A T P
E K V P V
C. briggsae 9320 TTTTATGTAGTGGTT-AGCCTTTAACCAAAACTAATAGAATCTAGTGAGTTTATGGGAAGATTTGTGATAATTTCATCCGATTCAGAAAAAGTTCCTGTA
C. elegans 9186 TCAGTT-TAGTGGTTTAG--TTGTAT---AAC-AATCTAATTTAGAG-GTTCTTAGTTAAACATGAAACGGATGCATT-GTTTCAGAAAAAGTACCGGTA
translation 963 exon 15 E K V P V
A G T S A K Y L V T V A E L K L K Q T E K K E D S K T E A L L K E
C. briggsae 9419 GCTGGAACATCTGCAAAGTATCTTGTGACGGTTGCTGAGTTGAAGTTGAAGCAAACGGAGAAAAAGAAGATAGCAAGACAGAAGCACTACTTAAGGAAC
C. elegans 9277 GCAGGAACATCTGCCAAATATCTTGTCACGGTTGCTGAATTGAAGCTGAAGCCTACGGACAA----GAAG-------GAG-GAAGCTCAATTGAAGGAGC
translation 968 A G T S A K Y L V T V A E L K L K P T D K K E E A Q L K E
Q L K D S E I N S E R
C. briggsae 9519 AGTTGAAGGATAGTGAGATTAATTCAGAAAGAAGGTGAGT----CT---AGTAGATATCCGGATGTTGCTAA------GCAATCACATTCAGAATAGACT
C. elegans 9365 AATTGAAGGACAGTGAGATCAATAGTGAAAGAAAGTAAGTAATCCTTGAAGTAGAAACCCCACTGTTACTTACATAGAGCAAAAACAATAAAAAT-G--T
translation 997 Q L K D S E I N S E R
C. briggsae 9606 AATTTATGTATTAGATGATAGGAAACATATAATTTCTCAAA-ACTTTATATTATAGTTCAAAACAACC-AAATCACAAAAAAACTTTTTCTTTTGGAACT
C. elegans 9462 TATTTTCAAATTTAAAAATACTCAAATTAGGAGCAGTCATACATTTTGAATAATAGTG-ATTTCATCAGAAATTTGTTTTAAG-TTTTTCCTCA--AACA
translation 1008
R T S Q T N L L F R P S A Y M A S D
C. briggsae 9704 CCCCCTA-ATGTCATTCTGA--GTTATCCCTGAATTTA-TCTCCAGAACCTCACAAACCAACCTTCTATTCCGTCCGTCCGCCTATATGGCATCGGATCT
C. elegans 9558 CGAAATTTATGTTATTCCCAAAGT-ATTC-TGAGTTTTCTCTCCAGAACCTCACAAACAAACCTTCTATTCCGTCCGTCTGCCTATAGGGAATCGGATTT
translation 1008 exon 16 K T S Q T N L L F R P S A Y R E S D
L Y R R I N
C. briggsae 9800 GTATCGCCGGATAAACGCGTCAGTTTTCCGTTGTTTTCTATC--ACTTATCC------TA--CTTTT---GGTTCAGTGGTTGTGGTTGTGATAATTGTT
C. elegans 9656 GTATCGCAGGATAAACGCGTCAGTTTTCCATATTTTTGTGTAGAACTTATCCAATTTATATGCTTTTTTTGGTTCAGTGATTGTGGTTGTGAT---TGTT
translation 1027 L Y R R I N
A V T R G A A P
C. briggsae 9887 TCAGTTCCTT----------TTTTGTGAACGTGGCTTTGTTTCTAACCGAATTGAGCCGTTTTC-GTCTGTTTCCAGAGTGACAAGGGGTGCAGCTCCGC
C. elegans 9753 TCATTTCTTTACCCAATTCATTTTGTGAATGTGGCTTC--TTCTAACACATTAG-GCCGTTTTTTGTTTGATTCCAGAGTGACAAAAGG---AGCT--GC
translation 1032 exon 17 A V T K G A

```
                    P  P  P  S  T  I  S  Q  E  S  D  G  K  D  T  H  E  S  A  S  A  Y  I  R  R  K  N  R  E  R  R  Q  R
C. briggsae  9976  CTCCACCATCGACAATTTCTCAGGAATCAGATGGAAAAGACACGCATGAAAGTGCTTCTGCGTATATAAGAAGGAAGAATCGAGAAAGGAGACAAAGGAA
C. elegans   9844  ----ACCGTCAACAATTTCTCAAGATTCCGATGGTAAAGACACGTTTGAGAGTGCTTCTGCATATATTCGACGCAAAAATCGAGAAAGACGTCAAAGAAA
translation  1038     A  P  S  T  I  S  Q  D  S  D  G  K  D  T  F  E  S  A  S  A  Y  I  R  R  K  N  R  E  R  R  Q  R

                 N  R  T  I  G  T  A  E  E  A  L
C. briggsae  10076 TCGGACTATTGGAACCGCCGAAGAAGCGCTACGGTTCAGTTTTTTTCTGTTTTCCTGGTTACTGACTCGAGAAAACTCAATTTATTCTCAGTTTCCTAA
C. elegans   9941  TCGGACGATTGGAACAGCAGAAGAAGCACTACGGTTCAGTTTTAATTAT-------TGATAACTGACTC---AAAATTACAT--ATTCTCAGTTTC-TAA
translation  1071  N  R  T  I  G  T  A  E  E  A  L

C. briggsae  10176 TTGAATGGATATTGATTATCATGATGATGCGTTCAGTTGTTTTTCTGATGCTCTAAACGCTTGAGTCAGGATTTCTTTGCGAAAATTTGACCATTCATTT
C. elegans   10028 TTGAATGCTT-TTG-TT--------GA------CAGTTTTGGTTCTGAT--TCTAAAAGCTCGAGTCAGGCTTTTTTTG---AAAT---AA-ATTCAAAT
translation  1081

                                                 R  A  L  E  R  P  S  V  D  P  Y  E
C. briggsae  10276 TTG-AAAAATTA-TC-GA---TTTTCAGAGCTCTGGAAAGACCATCTGTGGATCCATACGAAGGTACACCCCAAAAATTGAGAATTTTTCTTCA--ATGA
C. elegans   10103 TTGTAAAAATAAATCTGAAAATTTTCAGGGCTCTCGAGAGACCGTCTGTGGATCCATACGATGGTG-AGCTTAGAAAC--AAAATTATACAAAAGAATAA
translation  1081                        exon 18  R  A  L  E  R  P  S  V  D  P  Y  D

                          E  R  A  F  S  P  I  S  S  D  P  F  Y  S  H  H  S  S  G  V     A  A  P  S  A  L  G  Y  V
C. briggsae  10367 ---AACTTTTCAGAGCGGGCATTCTCTCCAATCTCTTCGGATCCATTCTACTCTCATCACTCTTCTGGAGTC-GCTGCTCCTTCGGCTCTTGGATACGTT
C. elegans   10200 TATAAATATTCAGATCGTGCATTCTCTCCTCTCTCGACTGATCCATTTTATTCTCATCACTCATCTGGT-TCAGCTGCTCCATCAGCACTTGGATATGTT
translation  1093        exon 19  D  R  A  F  S  P  L  S  T  D  P  F  Y  S  H  H  S  S  G  S  A  A  P  S  A  L  G  Y  V

                    R  P  R  Y  H  D  D  S  Y  R  T  H  R  D  S  D  Y  R  P  M  S  P  T  S  R  Y
C. briggsae  10464 CGTCCACGTTATCACGATGATAGCTACCGTACTCATCGTGATAGTGACTATCGTCCGATGTCCCCAACCAGTCGATATGTAAGTATAA-TAAAAGAAACA
C. elegans   10299 CGACCTCGTTATCATGATGATAGTTACCGTACTCATCGAGACAGTGATTATCGTCCAATGTCACCAACAAGTCGTTACGTAAGTATTTGTGGTAGAGAAT
translation  1122    R  P  R  Y  H  D  D  S  Y  R  T  H  R  D  S  D  Y  R  P  M  S  P  T  S  R  Y

C. briggsae  10563 GACAAAAAGGAGAGAATGATACCAATTCTCTCATCTAATTAGTCAAAATGAAGTAGTTCACACGAACAATCGGGTGA--CTGTCAAACATACACATCTGA
C. elegans   10399 GACAAAAACATAAAAACGATACCC-TTCTCTCATCTAATTAGTCGAAATGAAGTGGTTCAAATGAACAATTGCTTGCTTCTGGCTGTCTTATT--TCTGT
translation  1148

C. briggsae  10661 ATA---GAG-AGAAAACCCACACTTTGTATATGGCAGGAG-AAGAACACGCCTAAA----------GCGCGAACAAGCATTTGGAGGCGGAAAAAATGA
C. elegans   10496 ATCTTTGAGGAGAAAAG----ACTTTGTAT--GGCAAACGGAAGAAAACGCCTAAATGCTGCTGCTGGCGCGAACAGGCAGTTT-AGGCGAAAAAAA-GA

C. briggsae  10745 GAGAGAAAGGAATTGGGAGGAGAGAAGAGAGAGGAAGAAGAACACAGTGTACTTTGAATGA--------------GAGGAGTGGCCC----------C
C. elegans   10588 GAGAGAAA-----TGGGTGGAGA--AAGAGGGAGAATTGAGA---CAGTGTACTTTGAATGAATAAAGAAGAAAAAGAGGAGTAGAATAATAGAAAGAGC

C. briggsae  10820 CGT--CTTTTTAGTTTTGAACACTTTTGAGTATACCATTT------------------TCTA-ATTG----------TTTTC--------ATT-CACAT
C. elegans   10678 GGTTGCTTTTTAGTTTTAAATTCGCT--AGTATACCATTTATTCATGCTTATACACAATCAAGATTGGGTATGTTTACTTTTCTGCTTTTTATTGCAAAT

C. briggsae  10878 ---GCTT------------ATTCAC--AAT--------CAAGA---TTGGG--------TATGTTT---GTTTT-GAGAGTTA---GAAAAGCATT-G--
C. elegans   10776 AAGGGTTCTATAAATTTTTATTTTCCTAATTTAAGTTTCAAGAACATTATGCCTTCAATTATTTATATGGTTTTCGTGATTTTTCTGAAAA-CATTTGAT

C. briggsae  10932 -------TTTGG---------GTGTAAATAAACATTG------G---ATTAGT--------------G-----GTAGAAAGAACT----------A-GG
C. elegans   10875 CCATCAATTTGGAGCATATCTGTGCAGATATAGAGTGAGTGTCGTAAATTCGTAATTTTTGAAACACGTTTTTGTA-AAAGCACTGACTTATTTGAATGG

C. briggsae  10977 GTCGAATTG-----AAC--TGA------CAGG--------TTACGGT------A---GGTCT-----AAAATGTATCCCCAGATAAT-----AAGTTTG-
C. elegans   10974 GTCTACTTGGAAAAAACACTGATTTTTGCGGGGATCTCTTTTTCATTTATTTAATTTGTTCTTTGGAAAAATGAATTAT-AGATAATCATGAAAAATTGA
```


C. briggsae 11035 -------TTA------------GCT-------TCC------TTT---------TTTGGACATTTTAGAACG-------TGTCGTTGAATAGTATCAAAAT
C. elegans 11073 TAAAAAATTAAAGGAAAAGTGGGCCCCGAAAATCCAAAATGTTTAATAAAATATTTGTATTTTGAACTACGCATTAAATGAAATTAAATGAAATGAAATT
C. briggsae 11087 --T---TTGCGA----------C-CAGGAAACAGAGCTAGAAGCCTTAAAA----ACTG----TGAAGCACAACAGCGTTTCT---------CTGAAA--
C. elegans 11173 GGTAAATTACGATACAATTTCTCACAGAGTACATATTTCAAATAATTCAAAGGTTACTTCACCTGAAGTTCCTGGAATTTTCTAAATTGTGACTTAAAAT
F I * H F H S L K F Q T A A V S
C. briggsae 11152 -TTATATTCATTT-AACATTTTC----------ATTC---------TCT---------------G----AAATT------CCAGACTGCGGCAGTATCCC
C. elegans 11273 TTTATGTATATTTTAACGATTTCTGTGGAATAAATTCACCCGGTCCTCTATTTGCTCCCCCGTGGCCACAAATTATTTTTCCAGACAGCCGCTGTTACCA
translation 1148 exon 20 Y I L T I S V E * I H P V L Y L L P R G H K L F F Q T A A V T
R D D R G K S T S Y D P H D T S
C. briggsae 11207 GCGATGATCGTGGCAAGTCTACATCTTACGATCCACACGATACTAGCAGGTAA----TTG----CATT----------
C. elegans 11373 GAGATGATCGTGGCAAATCAACATCGTACGACCCACATGAAACCACCAGGTAAATCGTTGATACCATTATCTTTCCCGTTTTTTTCCTAATTGATCTCAT
translation 1153 R D D R G K S T S Y D P H E T T
C. briggsae 11266 ------------TCG-------TCACTTTCTTTT-----------TCCTT----------CCT--GTCTTTTTCTCCGAAGCTAG
C. elegans 11473 CAACTGTTTGTTATCGCTACATATCACTTTCTTTTAATATCACTTTTCCTTTTTTCTTATCTCTAATTTGACCACCTTGATTAATTAGACGAACAACGAT
translation 1169
C. briggsae 11311 ATAGTACCTCATCAACCGATCCCATTATAATTGTTATCTCTATATATCACTTTCTCTAAGTATCATTATCATCGTCGTCATCACTTCCCTCTCATTTCTA
C. elegans 11573 GCGACGGGAATCTTTCTTTTTTCTCTGCTTTGAGCATCAAATACAACACACAGAAAACATAATCGGATTGTACGGAGACTGTACGAGAATGAGAACAACA
C. briggsae 11411 CGATCAGTGATTTTGACCACCCTGATTAATGACATGAAAAAAGAGGGATGCGTTGTAGACGATTCGAATAGTATCTATTGTTTTTGTTGTTGACAAAAA
C. elegans 11673 TGCTCTCGAGACAAGAACACAAAACAAAAGTTCTATCGAGCATCCGTTTCTTTTT-TTATTTCTTCTCTATTCCTAAACTTTCCCAAAACACATGAGATA
C. briggsae 11511 CTGCCCTCCGAAACAACACAGAAAACATAATCCGATTTTACGGAGATTGAAAAGGAAGAAGAGAAATTAAGAAGGAGAAGAAAAAAGAACAAGATGTTCC
C. elegans 11772 TATTTATAGGTTTCTATTTTCTTTTCAAAAAAGTGTTTATTTCATTTTTTTTTCAG----------
C. briggsae 11611 TCCGATTCGGGAAAAAAATAGGAAAAAATACACTAGTCCCCCTCTGAGCCGTTTCTTTTAATTCTTTTTCGGTTCCTAGACTTCCCTAAAAAAGGACATG
C. elegans 11828 ----------
R P S Y D R S Y I T S S
C. briggsae 11711 AGATATATTTACAGGTTATGAATTTGAAGTTTGATAGTTGTTTAGTGAATAATCTCATTTTCAGACCGTCCTATGATAGAAGCTATATAACGTCTTCAAA
C. elegans 11828 ----------ACCATCGTACGATAGAGGCTACATAACTTCATCCAA
translation 1169 exon 21 R P S Y D R G Y I T S S
N A S P S Y T R K F E H E
C. briggsae 11811 TGCCAGCCCGAGTTATACACGCAAGTTTGAGCATGAG----------
C. elegans 11865 TGCAAGTCCAAGTTATACCCGGAAGTTTGAACATGAGGTGAGTAGTTTTTTCAACTTATGAACTGAAATTCAAGGAAGTTGCTATTTAAGCTTTTCCGAG
translation 1182 N A S P S Y T R K F E H E
C. briggsae 11847 ----------
C. elegans 11965 GTCTTCCGAAAACAAGAATATAAATAGGTATTTTATATTAATGTTGTAATTGTGGTAGTACATTATTGGATGTGGCACATTTGTATTGAGCTATACTATT
translation 1194
C. briggsae 11847 ----------
C. elegans 12065 CAAGCTAGATTGTTTGGCACTGAACAAAAATATAAATATAACCAGCTGCTGACCGCGCCTACGGCGCGGACAACGACTGGCACCATTAAAAGTATTTGAC
C. briggsae 11847 ----------
C. elegans 12165 ACACATACACTTCCAGAATTTTCTCGATTTTTCTAGAAAGTTCTGGAACATTCCAGAATTTTCCCGATTTTTCTAGAAAGTTCTGGAACACTCCAGAATT


C. briggsae 11847 ----------
C. elegans 12265 TTCTCGATTTTTCTAGAAAGTTCTGGAACATTCCAGAATTTTCTCGATTTTTCTAGAAAGTTCTGGAACATTCCAGAATTTTTTCGAAATTTCCAGAAGA
C. briggsae 11847 ----------
C. elegans 12365 TTCTAGATTTCCAGAATTTTAGAATTTTCAGAAAATTTAAATTTCCCGCCAAAATATTTTTCTCAGAAAATTTAAATTTCCCGCCAAAATATTTTTCACA
C. briggsae 11847 ---------GT---------A--GGTATGAC-------GGTGT-ACC----------TTGAAT-
C. elegans 12465 GAAAATTTAAATTTCCCGCCAAAATATTTTCTCAGAAAATTTAAATTTCCCTCAAAAATATTTTT
C. briggsae 11872 -------------AGAATA-TC----------T-GAATT--------GGTT-TAAAGCTTTTCCGAGGTCTTCC---ACG------AACA----AGAATC
C. elegans 12565 TTTCAGAAAATTTAAATTTCCCGCCAAAATATTTTCACAGAAAATTTA-AATTTCCCGTCAAAATTTTGGGTCTTACCACGATGGGTCTCACCACGACG
C. briggsae 11926 T----GA--TTAGGTATTT----TGGAATAGA------------TGTT-------------GTAAT--T---GT---------GAA-------AGTAAAC
C. elegans 12664 GGTCTCACCACGATGGGTCTCACCTCGATGGGTCTCACCACGATGGGTCTCACCACGATGGGTCTCGCCACGAAGATCTCGCAGCAACATTTTTTTATTT
C. briggsae 11970 TTGATT-----------TC----------
C. elegans 12764 TTGATTTTTCAGAAGGTTCTAGAATAATCCAGAATTTTTTCGAATTTTCCAGAAGGTTCTAAAGCTTTTCAGAAATTCCCGGAAGGTTCTGGAACATTTC
C. briggsae 11977 ----------
C. elegans 12864 AGAATTTTCCCGAAGTTTCCAATTTCCCTTCCAAAGACAAAAACACAATTTTTCCTAAACTACAGTAATCCTACCGTACTCCTACAGTACTCCTACAGTA
C. briggsae 11977 -------------------------AGC---------------ATA-----------GAGAAA-----TTCG-------------GACATCCAT----
C. elegans 12964 CTACTACAGTACCCCGACCATATCCCACTACTAAGCCCAAACTAATATCCCTCCAACAGCCGAAAACGCCTTGCCTTTGTAAGCTATGACGTCACTTCTT
Q P T L N S W A N D R K F E V Y K T
C. briggsae 12003 AAAATA-------TATTTATTT----------CAGCAACCAACATTAAACTCATGGGCCAATGATCGTAAATTTGAAGTGTATAAAACAC
C. elegans 13064 AACAAACGGACACTATTTTTTTATATATAGGTAATGATATTGCAGCAACCAACATTGAACTCATGGGCAAATGACCGTAAATTCGAAGTCTACAAAACAC
translation 1194 exon 22 Q P T L N S W A N D R K F E V Y K T
R A E R D A E R N T Y H S N I T
C. briggsae 12076 GGGCAGAACGGGACGCCGAGAGAAACACATATCATTCAAACCTGTTAGT---CTAATCCTAA--ACTTCTTGACCTGATTCGTTATTTTCAGAAGTACTA
C. elegans 13164 GGGCTGAAAGAGATGCTGAAAGGAATACATATCATCCAACTGGGTTTGTACCCTTTTC-TAATAACTTCCT-ATCTACTAAGATTTTTT-AT-AGAACTA
translation 1212 R A E R D A E R N T Y H P T exon 23 G T
T G A S S Y R P S S Y Y T S T S D R P L T G L R R P E E S Y N S S
C. briggsae 12171 CTGGTGCAAGCTCATACCGCCCTTCTTCATACTACACATCAACTAGTGATAGACCTTTAACTGGTCTAAGACGTCCTGAAGAGAGTTACAATTCATCTTC
C. elegans 13260 CTGGTGCGAGCTCATATCGTCCATCTTCATATTATACATCAGCAAGTGATAGACCCGTCACTGGTCTAAGACGTCCCGAAGAGAGTTATA---CATCTTC
translation 1228 T G A S S Y R P S S Y Y T S A S D R P V T G L R R P E E S Y T S
S T A Y S R Y D G R S T T P S N
C. briggsae 12271 AACTGCTTATAGCAGATATGATGGAAGATCCAGTTAGTTTTAAACATCGTTTCAGATCTCAGATGATAAACTT---------TCAGCAACACCTAGTAAC
C. elegans 13357 AACCAATTATGGGAGATATGATGGAAGATCAAGTTTGTTTGAA-------TTCACAAGTAAACTTATTAACTTCTCCTATTTTCAGCAACACCCAGCAAC
translation 1261 S T N Y G R Y D G R S exon 24 T T P S N
D H Y A I S S A L A D N Q A S S M E N R F R P T A R R V A S A M T
C. briggsae 12362 GATCACTACGCGATCTCCAGTGCTCTTGCCGATAATCAAGCATCCAGCATGGAAAACCGTTTCCGGCCAACTGCTAGACGTGTTGCTTCTGCTATGACTG
C. elegans 13450 GACCAATACGCAATTTCAACGGCTTTGGCTGAAAATCAAGCTTCAAGTATGAATAATCGTTTCCGTCCAACTGCTCGACGCGTTGCATCTGCAATGACTG
translation 1276 D Q Y A I S T A L A E N Q A S S M N N R F R P T A R R V A S A M T


D A D R R N Y H R S R S M D R N K V D Y E Y G S N L L T R L T T P
C. briggsae 12462 ATGCTGATCGCCGCAACTACCATCGTTCCAGATCTATGGATAGAAACAAGGTTGATTATGAATACGGAAGCAACCTTCTAACTCGATTGACTACTCCGGA
C. elegans 13550 ACGCAGATCGTCGTAATTATCACAGATCAAGATCAATGGATAGAAATAAGGTTGATTATGAATATGGAAGCAATTTGTTGAATAGACTTACAACACCTGA
translation 1309 D A D R R N Y H R S R S M D R N K V D Y E Y G S N L L N R L T T P
D T H G A G D Y S Y V N Y H D S S N G R S N T T M E K D G Q P R S I
C. briggsae 12562 TACTCACGGAGCTGGTGACTATTCCTATGTCAATTATCATGACAGTTCCAATGGAAGATCCAACACGACGATGGAAAAAGATGGACAGCCAAGAAGTATT
C. elegans 13650 TTCTCATGGAG------ACTACTCGTATGTAAACTATCACGACAGTTCAAATGGAAGATCTAATATCACATTGGAAAAAGATGGACAGCCAAGGAGTATT
translation 1343 D S H G D Y S Y V N Y H D S S N G R S N I T L E K D G Q P R S I
L K N K Q
C. briggsae 12662 CTTAAGAATAAGCAGGTTAGT--------TT-------AGAGTGATTTTTTATAACTGTAACAATT-GAA-------CTGA------------------
C. elegans 13744 CTGAAAAATAAACAGGTAAATACAATAACTTCCTGAAAAAATTG-TGTTTTCTAACAGGAACTTTTTGAATTCTTCCCTGAAATCCATATAAAATATTAA
translation 1374 L K N K Q
S A D I E P R V E
C. briggsae 12719 ----------TCC-------------TTGATTC----TCT-CTCCCCATAA------CCCACAATTCCAGAGTGCTGACATTGAACCACGTGTTGAGTC
C. elegans 13843 ACAAAAAATATTCCGAAGTAAAAACCTTTTATTAATGATTTGCTCGCTGGAATGTGCTCCTA-AATTTCAGAGTGCTGACATCGAACCACGTGTTGAATC
translation 1379 exon 25 S A D I E P R V E
S S Y E P S S G V R S
C. briggsae 12785 ATCCTACGAACCAAGCAGTGGTGTTCGATCGGTAAGTAACTC--CTTCTT------TTTGTCCCACTTAAAAACATTT---TGATCCCCTTTA---TTTT
C. elegans 13942 GTCTTATGAGCCAAGCAATGGAGTTCGATCGGTAAGTTCCTCTCCTTCCTCAATCCTTTTTCCCATTTCACC-CGTTTATCTTATTACTTTCAAACTTCT
translation 1389 S S Y E P S N G V R S
H E F L S S F F L S Q L L S M R F I L
C. briggsae 12871 TGGAGC--ACCTCTTCCCT-CATGTTCCCTTCTGAAATTCAGCATGAATTTCTTTCCTCTTTTTTTCTTTCTCAACTTCTAAGCATGCGTTCTATTCTCC
C. elegans 14041 TCTTTCCAATATTTTTCCTACATATTCCCT-CTCGAATTCAGCATGAATTTCTTGA----------TCAACTTCTAAGCATGCGTTTCATTCTCC
translation 1399 exon 26 H E F L D Q L L S M R F I L
Q V F E R L R R H L S L E K S V S P Q R Q
C. briggsae 12968 AGGTGTTCGAACGTCTCCGTCGTCATTTGTCTCTCGAGAAAAGTGTTTCGCCGCAAAGACAGGTAAGCGAAAAG-AGAATATAAAAAACAGAGAATCTAG
C. elegans 14125 AGGTGTTCGAACGTCTTCGTCGTCATTTGTCTCTCGAGAAAAGTGTTTCGCCGCAAAGACAGGTAAGC-AAATGCAGGATTTAAATTACTGTTTCTTAAA
translation 1413 Q V F E R L R R H L S L E K S V S P Q R Q
D T S F Y G R Q L R
C. briggsae 13067 ATT-TAATGGGAAA-------------------CTTCTTTAATTC-TTAGGATACCTCGTTTTATGGCAGACAGCTTCGTGTAAGTATTAGT-----ATT
C. elegans 14224 ACTATAATGGGAAAAAACTTCCGAAATCTATTCCTTCTTTT-TTCATTAGAACACTTCATTTTATGGCAGACAGCTTCGCGTAAGTGTGAGGGGGAAATG
translation 1434 exon 27 N T S F Y G R Q L R
C. briggsae 13141 ATACCGGTGTCGGATAATCT-GAAGCTCTG-AAATG---ATCCTC----------TAA-AA-TTTTCAGTGTGCC-A-CTT
C. elegans 14323 GTGTTGGTGGGGGAAGA-CTAGAAGCTGTGCAAATTTTAATTCTCATCTGGTTTTCCCCAATTTTTTTTTCGAATAAGAACTTTTCAGTGTGAATATCTT
translation 1444
C. briggsae 13203 TTTTCAAATTTTTCCAGCATGGTATCCATTACCACTGTGTG-----AC------ATGTTGAACCTCCCCCAATATACAGTTTTTATCATTTTCAATACTT
C. elegans 14422 TCTTACAATTTATCCAGCATGGTAACTATTACCATTGTGTGTTTTTACCCCCAAATATTTTATATCTTCGATT-TCCTTTTGTTAT-ATTTT-AATACTC
V A S V
C. briggsae 13292 ACAGAATACTGCCAACCTAATTATTTGTTTTTTCATTGTGTTTAGCTGTCTTCCAAA----TACAATTCTCATCAA---AATTTAGGTCGCATCTGTCGG
C. elegans 14519 ACACACAACAAGTACACAAAC-ATTC---TAATAACTGTGTTT---TATC--CCAAACTTCTACTTATCTCA-CATTTTACTTTAGGTTGCATCTGTCGG
translation 1444 exon 28 V A S V


G P H S R N I D T T S S V L D N P K K K R S L L S F N R R K T S E V
C. briggsae 13385 GCCGCACAGCCGAAACATCGACACAACGTCATCTGTGTTGGATAATCCAAAAAAGAAAAGATCTTTGCTTTCATTCAATCGAAGAAAAACGAGTGAAGTG
C. elegans 14609 ACCACATAGTCGAAATATTGATAGCACATCATCTGCATTAGACAATCCAAAAAAGAAAAGATCATTGCTCTCATTCAATCGGAGAAAGACAAGTGAAGTA
translation 1449 G P H S R N I D S T S S A L D N P K K K R S L L S F N R R K T S E V
R M G A D G K L I T N G Y D D T P S S R D F K R P S S P I D R I K
C. briggsae 13485 AGAATGGGAGCCGATGGAAAATTGATCACCAACGGTTATGATGACACTCCGAGCAGCAGAGATTTCAAGAGACCAAGTTCTCCAATTGATAGAATCAAGT
C. elegans 14709 CGAATGGGTGCCGATGGAAAGCTGGTTACAAATGGATATGATAGTACACCGAGTAGCAGAGATTTCAAGAGACCAAGCTCCCCAATCGATCGGATAAAGT
translation 1482 R M G A D G K L V T N G Y D S T P S S R D F K R P S S P I D R I K
S L F R K S D T A G T G H F D Y Y N S S
C. briggsae 13585 CGTTGTTTAGAAAAAGTGATACTGCTGGAACTGGACATTTCGATTATTATAATTCTTCAAGGTAA-CATTTGTAG-AATGTT----------
C. elegans 14809 CTCTGTTCAGAAAGAATGAAAGTTCAGGAACTGGGCACTCGGATTATTATAATTCTTCAAGGTAAACTTTTG-AGTAACGTTGTCTCATTTTCAACTAGA
translation 1515 S L F R K N E S S G T G H S D Y Y N S S
C. briggsae 13664 ----ACCT----------TAAG----TAAGAT-------TATTTCT----------
C. elegans 14908 TGCCACCTATGCTCTTTGAATTTTTAAGCATATAATATAAAGCTGTATTTATATGTAGAAGTTATGAGAAATTTAGAAAACCTTTTAAAAACTTCAAAAC
translation 1535
C. briggsae 13685 ----------
C. elegans 15008 TTCTGTCGAAGAATAGAAGAAACTGTGCCCAAGTCAAGTTTCTAAGCTATCGAACTGCACTAACACTTTTCCTAAAAATATTTTCAAAAGCTCCAAACCG
C. briggsae 13685 ----------
C. elegans 15108 CCCAATTGGAGGGGGGTGCTATAGAAGTTGGACGCACTGTACTACTTCATTAAACTGTGATCAAAAGTTACCTTAACAATCTACTTATGATTCCATTAAT
R A Y T S S N P T S T R E S Y V A Q Y R K Y P
C. briggsae 13685 -TCAGGGCGTACACTTCTTCAAATCCAACATCAACAA---GAGAATCATACGTGGCTCAATATAGAAAATATCCAGGTACCGTTTCGGCCCGGGTTTTCT
C. elegans 15208 TTCAGAAACTACACATCTTCAAATCCGGTTTCCACAACTCGCGAGCCATATGTTGCCCAATATAGAAAATATCCAGGTGC----------
translation 1535 exon 29 R N Y T S S N P V S T T R E P Y V A Q Y R K Y P
C. briggsae 13782 GTTATTTTTTCCTGTATTATCAAAAAAAAATGTTTAAAAAAGAAACCCATTGTCTGAAAAAAAAACTCAGGCACTGACTCTTCAACTATGTTGATTATTG
C. elegans 15285 ----------TTTATTTCACAGAAACCCTATTAT
translation 1559
G S T T R D T S S A L N R Y S Y T P G L T D Q R R H W
C. briggsae 13882 CTCCAACTGAGGGTTTTCAGGTTCAACCACCCGTGATACATCCAGCGCTCTCAATCGCTACTCCTACACTCCAGGTCTAACTGATCAACGTCGTCATTGG
C. elegans 15312 TTTAATTGATAGATTTTCAGGTTCAACCTCACGTGATACGTCCAGCGCTCTCAATCGTTACTCATACACTCCTGGTCTAACAGATCAACGCCGTCATTGG
translation 1559 exon 30 G S T S R D T S S A L N R Y S Y T P G L T D Q R R H W
Y D D P N I Y *
C. briggsae 13982 TATGATGACCCGAATATTTATTAA----------AAT---GCCATCTTCCTCTTC-TTCTTTTCTT-----CTTTACC--TAACTTCCCCACTTTCTCAT
C. elegans 15412 TATGATGATCCAAATATTTATTAACTTTTTTTTCAGTTTTGCCATCTTTCTTTTCCTTCTCTTTTTGTAAACTTTAAAAATAATTTGAAAACTCTTT--T
translation 1586 Y D D P N I Y *
C. briggsae 14061 TTTATGTAACTTAAGAGTCTTTTG------AATTTGCTCATCAATTCCTTTTTT-GTAAATATCGCATGGAGAAACTATCGGGTGTCCGATACTATCCCT
C. elegans 15510 TTT-TGTAACTTAAAAAT-TCTTGCTCACCACTTT-CT-A-CAGTTCATTTTTTTGTAAATATCGTATGGAGAAACTATCGGGTGTCCGATATTATTCTT
translation 1593
C. briggsae 14154 TTTT----------CCCTCACCC-------TAGTCAA----------AAAT
C. elegans 15605 TTTTTATTAAAAATGTCATCTTCTCTCTCACCCAAAAACCTAGTCAACAATGACTCTTTACATGTCGTCTTCATTCATTCTTTCATCATTGTTTCAAAAT

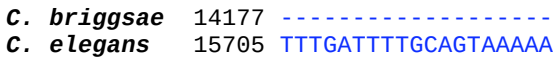


FIGURE S1.—The *unc-82* gene structure is highly conserved between *C. elegans* and *C. briggsae*. The pairwise alignment of *unc-82* genomic sequences are shown with nucleic acid identities in black, and mismatches and indels in blue. Nucleic acid residues in exons are in bold face and underlined. Each exon label, found at the 5' exon boundary, is color coded to match the splicing data presented in Figure 3: black indicates invariant exons, magenta indicates exons excluded from some isoforms, and green marks alternative exon donor or acceptor sites found in cDNA clones. The alternative splice acceptor site (green) for exon 26 is conserved in *C. briggsae*, and maintains the reading frame in both species. The alternative acceptor site for exon 20 is not present in *briggsae*, and introduces a stop codon in the *elegans* sequence. The translation of the *C. elegans* and predicted *C. briggsae* coding regions are shown below and above the DNA alignment, respectively. Amino acid identities are shown in black, whereas sites of amino acid substitution are colored gray. Comparison of the protein sequences within the 252-residue kinase domain, highlighted in yellow, revealed a single amino acid substitution. Overall, the coding regions are 82% identical at the nucleic acid level, and 92% identical at the protein level. Sequences highlighted in pink were predicted to be part of exons, but were not found in any cDNA clones. These include a region in intron 1 which was predicted by Genefinder (http://www.wormbase.org/) to be the first exon in the *unc-82* gene. However, this presumptive exon does not appear to be present in mRNA, based on our inability to amplify the sequence from first-strand cDNA or from a random-primed cDNA library (obtained from Bob Barstead). Further, this presumptive exon shows a lower degree of conservation than any confirmed *unc-82* exon when compared to the homologous sequences in *C. briggsae*. Inclusion of this presumptive exon as the second exon in the *briggsae* message would introduce an in-frame stop codon. Our cDNA analysis also identified the previously undetected exon #27.

Input data matrix

```
                    1111111111222222222233333333334444444444555555555566666666667777777777
Taxon/Node  12345678901234567890123456789012345678901234567890123456789012345678901234567890
-----------------------------------------------------------------------------------------------
HsMARK3     YRLL-KTIGKGNFAKVKLARHI----LTG------REVAIKIID---------------------------------
HsMARK1     YRLQ-KTIGKGNFAKVKLARHV----LTG------REVAVKIID---------------------------------
HsMARK2     YRLL-KTIGKGNFAKVKLARHI----LTG------KEVAVKIID---------------------------------
HsMARK4     YRLL-RTIGKGNFAKVKLARHI----LTG------REVAIKIID---------------------------------
DmPAR1      YKLI-KTIGKGNFAKVKLAKHL----PTG------KEVAIKIID---------------------------------
CePAR1      YKLL-KTIGKGNFAKVKLAKHV----ITG------HEVAIKIID---------------------------------
DmKP78b     YKII-KTLGKGNFAKVKLAIHL----PTG------REVAIKLID---------------------------------
DmKP78a     YKII-KTLGKGNFAKVKLAIHV----PTG------REVAIKVID---------------------------------
HsSIK       YDIE-RTLGKGNFAVVKLARHR----VTK------TQVAIKIID---------------------------------
HsQIK       YDIE-GTLGKGNFAVVKLGRHR----ITK------TEVAIKIID---------------------------------
DmCG4290    YDIE-RTIGKGNFAVVKLARHR----ITK------NEVAIKIID---------------------------------
HsQSK       YEID-RTIGKGNFAVVKRATHL----VTK------AKVAIKIID---------------------------------
DmCG15072   YELE-KTIGKGNFAVVKLATNI----VTK------TKVAIKIID---------------------------------
CeKIN29     YDVG-RAIGKGNFATVRIARHK----IAK------TKVAIKSID---------------------------------
HsARK5      YELQ-ETLGKGTYGKVKRATER----FSG------RVVAIKSIR---------------------------------
HsSNARK     YEFL-ETLGKGTYGKVKKARE-----SSG------RLVAIKSIR---------------------------------
CeUNC82     FEIT-KKLGSGTYGKVSLAYDH----KFD------REVAVKLIK---------------------------------
DmCG11870   FDII-KKLGQGTYGKVQLGINK----ETG------QEVAIKTIK---------------------------------
HsAMPKa1    YILG-DTLGVGTFGKVKVGKHE----LTG------HKVAVKILN---------------------------------
HsAMPKa2    YVLG-DTLGVGTFGKVKIGEHQ----LTG------HKVAVKILN---------------------------------
CeAAK2      YILK-ETLGVGTFGKVKVGIHE----TTQ------YKVAVKILN---------------------------------
DmSNF1A     YLLG-ATLGTGTFGKVKIGEHQ----ITR------VKVAVKILN---------------------------------
CeAAK1      FVIK-ETIGKGAFGAVKRGTHI----QTG------YDVAIKILN---------------------------------
ScSnf1p     YQIV-KTLGEGSFGKVKLAYHT----TTG------QKVALKIIN---------------------------------
HsBRSK1     YRLE-KTLGKGQTGLVKLGVHC----ITG------QKVAIKIVN---------------------------------
HsBRSK2     YRLE-KTLGKGQTGLVKLGVHC----VTC------QKVAIKIVN---------------------------------
DmCG6114    YRLE-KTLGKGQTGLVKLGVHC----VIG------KKVAIKIIN---------------------------------
CeSAD1      YKLE-KTLGKGQTGLVKTGTHC----ITG------RKVAIKIVN---------------------------------
ScGin4p     WKLG-ETLGLGSTGKVQLARNG----STG------QEAAVKVIS-KAVFNTGNVSGTSI-------------------
ScKcc4p     WKLG-ETLGFGSTGKVQLAQHE----RTG------HRTAVKVIS-KSIFNN---NGN---------------------
ScHsl1p     WKLG-KTLGKGSSGRVRLAKNM----ETG------QLAAIKIVPKKKAFVHCSNNGTVPNSYSSSMVTSNVSSPSIAS
ScKin4p     YIIG-STLGEGEFGKVKLGWTK--ASSSN---EVPKQVAIKLIR---------------------------------
ScLpi5p     YILG-STLGEGEFGKVKLGWPKNFSNSSNSTFDFPKQVAIKLIK---------------------------------
ScKin1p     WEFV-ETVGAGSMGKVKLAKHR----YTN------EVCAVKIVNRATKAFLHKEQMLPP---------PKNEQDVLER
ScKin2p     WEFL-ETVGAGSMGKVKLVKHR----QTK------EICVIKIVNRASKAYLHKQHSLPS---------PKNESEILER
HsMELK      YELH-ETIGTGGFAKVKLACHI----LTG------EMVAIKIMD---------------------------------
CePIG1      YAVH-DELGSGGFGKVRLATHL----LTN------QKVAIKIID---------------------------------
HsHUNK      YLIGSRKLGEGSFAKVREGLHV----LTG------EKVAIKVID---------------------------------
HsNIM1      YRIR-GEIGSGNFSQVKLGIHS----LTK------EKVAIKI-----------------------------------
DmCG4629    YRFC-GDIGRGNFSKVKLAVHQ----LTR------DKVAIKVVD---------------------------------
CeF49C5.4   YRLG-KELGAGNFSKVKLGVHQ----LTK------EKVAVKI-----------------------------------
DmCG8485    YDLE-ETLGSGHFAVVKLARHV----FTG------AKVAVKVVD---------------------------------
HsSNRK      YDLD-KTLGRGHFAVVKLARHV----FTG------EKVAVKVID---------------------------------
CeZK524.4   YDLE-KTIGQGHFAVVKLAKHV----FTG------EMVAVKIID---------------------------------
HsTSSK1     YLLG-INLGEGSYAKVKSA-YS-----ER----LKFNVAIKIIDRK-------------------------------
HsTSSK2     YIVG-INLGKGSYAKVKSA-YS-----ER----LKFNVAVKIIDRK-------------------------------
DmCG14305   YNVG-HKIGEGSYATVITAGYA-----DDH--GHGVHLACKIIDKA-------------------------------
HsTSSK3     YQLG-KTIGEGTYSKVKEA-FS-----KK----HQRKVAIKVIDKM-------------------------------
HsSSTK      YKLG-RTIGEGSYSKVKVA-TS-----KK----YKGTVAIKVVDRR-------------------------------
DmCG9222    IILG-KVIGTGNYAKVKIG-FS-----EE----YGKRVAVKIISKV-------------------------------
CeC27D6.11  LKFD-KILGSGSYSRVARATWG-----DK-----KLEVAAKVINITS-------------------------------
CeY38H8A.4  LKFD-KILGSGSYSRVARATWG-----DK-----KLEVAAKVINITP-------------------------------
```

```
CeB0511.4   LKFE-KVLGSGSYSRVARATFG-----EK-----KLEVAAKIINITP------------------------------

Input data matrix (continued):

                                 1111111111111111111111111111111111111111111111111111111111
            7888888888899999999990000000000011111111122222222223333333333444444444455555555
Taxon/Node  9012345678901234567890123456789012345678901234567890123456789012345678901234567
------------------------------------------------------------------------------------------
HsMARK3     --KTQLNPTSLQKLFREVRIMKILNHPNIVKLFEVIETEK-TLYLIMEYASGGEVFDYLVAHGR-MKEKEARSKFRQI
HsMARK1     --KTQLNPTSLQKLFREVRIMKILNHPNIVKLFEVIETEK-TLYLVMEYASGGEVFDYLVAHGR-MKEKEARAKFRQI
HsMARK2     --KTQLNSSSLQKLFREVRIMKVLNHPNIVKLFEVIETEK-TLYLVMEYASGGEVFDYLVAHGR-MKEKEARAKFRQI
HsMARK4     --KTQLNPSSLQKLFREVRIMKGLNHPNIVKLFEVIETEK-TLYLVMEYASAGEVFDYLVSHGR-MKEKEARAKFRQI
DmPAR1      --KTQLNPGSLQKLFREVRIMKMLDHPNIVKLFQVIETEK-TLYLIMEYASGGEVFDYLVLHGR-MKEKEARVKFRQI
CePAR1      --KTALNPSSLQKLFREVKIMKQLDHPNIVKLYQVMETEQ-TLYLVLEYASGGEVFDYLVAHGR-MKEKEARAKFRQI
DmKP78b     --KTALNTIARQKLYREVNIMKKLNHPNIVRLLQVIESER-TLYLVMEYVSGGELFNYLVKNGR-MRERDARVLFRQL
DmKP78a     --KTQLNTSARQKLYREVKIMKLLNHPNIVRLFQVIESER-TLYLVMEYASRGELFDHLVKNGR-MRERDARVIFRQL
HsSIK       --KTRLDSSNLEKIYREVQLMKLLNHPHIIKLYQVMETKD-MLYIVTEFAKNGEMFDYLTSNGH-LSENEARKKFWQI
HsQIK       --KSQLDAVNLEKIYREVQIMKMLDHPHIIKLYQVMETKS-MLYLVTEYAKNGEIFDYLANHGR-LNESEARRKFWQI
DmCG4290    --KSQLDQTNLQKVYREVEIMKRLKHPHIIKLYQVMETKN-MIYIVSEYASQGEIFDYIAKYGR-MSESAARFKFWQI
HsQSK       --KTQLDEENLKKIFREVQIMKMLCHPHIIRLYQVMETER-MIYLVTEYASGGEIFDHLVAHGR-MAEKEARRKFKQI
DmCG15072   --KTCLNEEYLNKTFREIAILKSLRHPHITRLYEVMESQS-MIYLVTEYAPNGEIFDHLVANGR-MKEPEAARVFTQL
CeKIN29     --VSALDRENLIKLEREVKIVKVIDHPHIVKSYEIMRVDN-MLYIVSEYCSSGELYETLIEKGR-VAENVARKWFSET
HsARK5      -KDKIKDEQDMVHIRREIEIMSSLNHPHIISIYEVFENKD-KIVIIMEYASKGELYDYISERRR-LSERETRHFFRQI
HsSNARK     -KDKIKDEQDLMHIRREIEIMSSLNHPHIIAIHEVFENSS-KIVIVMEYASRGDLYDYISERQQ-LSEREARHFFRQI
CeUNC82     -KSAIESKADLVRIRREIRIMSALNHPNIIQIYEVFENKD-KIILVMEYSSGGELYDYVSRCGS-LPEAEARRIFRQI
DmCG11870   -KCKIEAEADLVRIRREVQIMSSVHHPNIIHIYEVFENRE-KMVLVMEFAAGGELYDYLSERKV-LTEEEARRIFRQV
HsAMPKa1    -RQKIRSLDVVGKIRREIQNLKLFRHPHIIKLYQVISTPS-DIFMVMEYVSGGELFDYICKNGR-LDEKESRRLFQQI
HsAMPKa2    -RQKIRSLDVVGKIKREIQNLKLFRHPHIIKLYQVISTPT-DFFMVMEYVSGGELFDYICKHGR-VEEMEARRLFQQI
CeAAK2      -RQKIKSLDVVGKIRREIQNLSLFRHPHIIRLYQVISTPS-DIFMIMEHVSGGELFDYIVKHGR-LKTAEARRFFQQI
DmSNF1A     -RQKIKSLDVVGKIRREIQNLKLFRHPHIIKLYQVISTPS-DIFMIMEYVSGGELFDYIVKHGK-LQEHQARRFFQQI
CeAAK1      -RGRMKGLGTVNKTRNEIDNLQKLTHPHITRLFRVISTPS-DIFLVMELVSGGELFSYITRKGA-LPIRESRRYFQQI
ScSnf1p     -KKVLAKSDMQGRIEREISYLRLLRHPHIIKLYDVIKSKD-EIIMVIEYA-GNELFDYIVQRDK-MSEQEARRFFQQI
HsBRSK1     --REKLSESVLMKVEREIAILKLIEHPHVLKLHDVYENKK-YLYLVLEHVSGGELFDYLVKKGR-LTPKEARKFFRQI
HsBRSK2     --REKLSESVLMKVEREIAILKLIEHPHVLKLHDVYENKK-YLYLVLEHVSGGELFDYLVKKGR-LTPKEARKFFRQI
DmCG6114    --REKLSESVLMKVEREIAIMKLIDHPHVLGLSDVYENKK-YLYLILEHVSGGELFDYLVKKGR-LTPKEARKFFRQI
CeSAD1      --KEKLSESVLQKVEREIAIMKLIEHPHVLHLYDVYENKK-YLYLLLEHVSGGELFDYLVRKGR-LMSKEARKFFRQI
ScGin4p     -VGSTTPDALPYGIEREIIIMKLLNHPNVLRLYDVWETNT-DLYLVLEYAEKGELFNLLVERGP-LPEHEAIRFFRQI
ScKcc4p     -H-SNDDSVLPYNIEREIVIMKLLSHPNVLSLYDVWETNN-NLYLILEYAEKGELFNLLVDHGP-LPEREAINCFRQI
ScHsl1p     REHSNHSQTNPYGIEREIVIMKLISHTNVMALFEVWENKS-ELYLVLEYVDGGELFDYLVSKGK-LPEREAIHYFKQI
ScKin4p     RDTIKKDADKEIKIYREINALKHLTHPNIIYLEEVLQNSK-YIGIVLEFVSGGEFYKYIQRKRR-LKESSACRLFAQL
ScLpi5p     RDSISNDYRKEVKIYREINALKHLSHPNIVKLEEVLQNSR-YIGIVLEYACGGEFYKYIQKKRR-LKEMNACRLFSQL
ScKin1p     QKKLEKEISRDKRTIREASLGQILYHPHICRLFEMCTLSN-HFYMLFEYVSGGQLLDYIIQHGS-IREHQARKFARGI
ScKin2p     QKRLEKEIARDKRTVREASLGQILYHPHICRLFEMCTMSN-HFYMLFEYVSGGQLLDYIIHDGS-LKEHHARKFARGI
HsMELK      KNTLGSD---LPRIKTEIEALKNLRHQHICQLYHVLETAN-KIFMVLEYCPGGELFDYIISQDR-LSEEETRVVFRQI
CePIG1      KKQLGHD---LPRVQTEMDALRNLSHQNICRLYHYIETED-KFFIVMEYCSGGEMFDYIVRKER-LEESEARHFFRQL
HsHUNK      KKRAKKDTYVTKNLRREGQIQQMIRHPNITQLLDILETEN-SYYLVMELCPGGNLMHKIYEKKR-LEESEARRYIRQL
HsNIM1      LDKTKLDQKTQRLLSREISSMEKLHHPNIIRLYEVVETLS-KLHLVMEYAGGGELFGKISTEGK-LSEPESKLIFSQI
DmCG4629    LDRAGLDAKALRMLSSEIATLECVHHPNILRLFEVVETLG-RVYLVTEWIRGGELYNHITQGGP-LREIHAAPLLKQL
CeF49C5.4   MDKAKMDAKAQKLLSREIQAMEEMNHPNIVKLFEVVETLT-RVHLVIEYASGGELYTYVHERGK-LTEGDAKPLFAQI
DmCG8485    --KTKLDDVSKAHLFQEVRCMKLVQHPNVVRLYEVIDTQT-KLYLVLELGDGGDLYDYIMKHDSGLSEELARKYFRQI
HsSNRK      --KTKLDTLATGHLFQEVRCMKLVQHPNIVRLYEVIDTQT-KLYLILELGDGGDMFDYIMKHEEGLNEDLAKKYFAQI
CeZK524.4   --KTKMDEASTSQIMKEVRCMKLVQHANIVRLYEVLDTQT-KIFLILELGD-YDLHDFIIKHEKGVCESLAQQYFCQI
HsTSSK1     ---KAPADFLEKFLPREIEILAMLNHCSIIKTYEIFETSHGKVYIVMELAVQGDLLELIKTRGA-LHEDEARKKFHQL
HsTSSK2     ---KTPTDFVERFLPREMDILATVNHGSIIKTYEIFETSDGRIYIIMELGVQGDLLEFIKCQGA-LHEDVARKMFRQL
DmCG14305   ---KAPTDFVNKFFPRELEILTKIDHSNIIQIHSILQRGP-KIFIFMRYAENGDLLSHIKRSGP-IDEKQSKIWFFQM
HsTSSK3     ---GGPEEFIQRFLPRELQIVRTLDHKNIIQVYEMLESADGKICLVMELAEGGDVFDCVLNGGP-LPESRAKALFRQM
HsSSTK      ---RAPPDFVNKFLPRELSILRGVRHPHIVHVFEFIEVCNGKLYIVMEAAAT-DLLQAVQRNGR-IPGVQARDLFAQI
```

```
DmCG9222    ---KAPSEYTQKFLPREIEAVKGLHHENLITFYQSIETSH-RVYLIMQLAENGTLLDYVRERKF-LDEPQSRTLFKQL
CeC27D6.11  --TREKDDYIKKFLPREKEIVKLLKHDNICRLYEMISFPD-HIIFVTEFCAGGDLLRKMKDIKT-MNEDDAKFTFRQF
CeY38H8A.4  --TREKDDYIKKFLPREKEIVKLLKHDNVCRLYEMISFPD-HIIFVTEFCAGGDLLRKMKKIKT-MSEENAKFTFRQF
CeB0511.4   --TREKEDYIKKFLPREKEIVKLLKHDNICRLYEMISFQD-HIIFVNEYCAGGDLLRKMKDIVA-MKEEDAKFTFRQL
Input data matrix (continued):

            11111111111111111111111111111111111111111112222222222222222222222222222222222
            55566666666667777777777888888888999999999000000000011111111122222222233333
Taxon/Node  78901234567890123456789012345678901234567890123456789012345678901234567890123456789012345678901234
----------------------------------------------------------------------------------------------------
HsMARK3     VSAVQYCHQKRIVHRDLKAENLLLDADMN-IKIA---DFGFS-----NEFTVGG--KLDTFCGSPPYAAPELFQ-GKK
HsMARK1     VSAVQYCHQKYIVHRDLKAENLLLDGDMN-IKIA---DFGFS-----NEFTVGN--KLDTFCGSPPYAAPELFQ-GKK
HsMARK2     VSAVQYCHQKFIVHRDLKAENLLLDADMN-IKIA---DFGFS-----NEFTFGN--KLDTFCGSPPYAAPELFQ-GKK
HsMARK4     VSAVHYCHQKNIVHRDLKAENLLLDAEAN-IKIA---DFGFS-----NEFTLGS--KLDTFCGSPPYAAPELFQ-GKK
DmPAR1      VSAVQYCHQKRIIHRDLKAENLLLDSELN-IKIA---DFGFS-----NEFTPGS--KLDTFCGSPPYAAPELFQ-GKK
CePAR1      VSAVQYLHSKNIIHRDLKAENLLLDQDMN-IKIA---DFGFS-----NTFSLGN--KLDTFCGSPPYAAPELFS-GKK
DmKP78b     VSAIEYCHSKSIVHRDLKAENLLLDQQMK-LKIA---DFGFS-----TTFEPKA--PLETFCGSPPYAAPELFR-GKK
DmKP78a     VSAIQYCHSKFVVHRDLKAENLLLDQHMN-IKIA---DFGFG-----NTFDPNA--QLETFCGSPPYAAPELFM-GRK
HsSIK       LSAVEYCHDHHIVHRDLKTENLLLDGNMD-IKLAGTEDFGFG-----NFYKSGE--PLSTWCGSPPYAAPEVFE-GKE
HsQIK       LSAVDYCHGRKIVHRDLKAENLLLDNNMN-IKIA---DFGFG-----NFFKSGE--LLATWCGSPPYAAPEVFE-GQQ
DmCG4290    ISAVEYCHKKGIVHRDLKAENLLLDLNMN-IKIA---DFGFS-----NHFKPGE--LLATWCGSPPYAAPEVFE-GKQ
HsQSK       VTAVYFCHCRNIVHRDLKAENLLLDANLN-IKIA---DFGFS-----NLFTPGQ--LLKTWCGSPPYAAPELFE-GKE
DmCG15072   VSAVHYCHRRGVVHRDLKAENVLLDKDMN-IKLA---DFGFS-----NHYEEGA--TLKTWCGSPPYAAPEVFQ-GLE
CeKIN29     ASAVAYLHSQGIVHRDLKAENILLGKNSN-IKII---DFGFS-----NFQTGDQ--LLNTWCGSPPYAAPELLL-GNS
HsARK5      VSAVHYCHKNGVVHRDLKLENILLDDNCN-IKIA---DFGLS-----NLYQKDK--FLQTFCGSPLYASPEIVN-GRP
HsSNARK     VSAVHYCHQNRVVHRDLKLENILLDANGN-IKIA---DFGLS-----NLYHQGK--FLQTFCGSPLYASPEIVN-GKP
CeUNC82     TSAVLYCHKHRVAHRDLKLENILLDQNNN-AKIA---DFGLS-----NYFADKN--LLTTFCGSPLYASPEIIN-GTP
DmCG11870   ATAVYYCHKHKICHRDLKLENILLDEKGN-AKIA---DFGLS-----NVFDDQR--LLGTFCGSPLYASPEIVE-GTP
HsAMPKa1    LSGVDYCHRHMVVHRDLKPENVLLDAHMN-AKIA---DFGLS-----NMMSDGE--FLRTSCGSPNYAAPEVIS-GRL
HsAMPKa2    LSAVDYCHRHMVVHRDLKPENVLLDAHMN-AKIA---DFGLS-----NMMSDGE--FLRTSCGSPNYAAPEVIS-GRL
CeAAK2      ISGVDYCHRHMVVHRDLKPENLLLDEQNN-VKIA---DFGLS-----NIMTDGD--FLRTSCGSPNYAAPEVIS-GKL
DmSNF1A     ISGVDYCHRHMIVHRDLKPENLLLDHNMH-VKIA---DFGLS-----NMMLDGE--FLRTSCGSPNYAAPEVIS-GKL
CeAAK1      ISGVSYCHNHMIVHRDLKPENLLLDANKN-IKIA---DFGLS-----NYMTDGD--LLSTACGSPNYAAPELIS-NKL
ScSnf1p     ISAVEYCHRHKIVHRDLKPENLLLDEHLN-VKIA---DFGLS-----NIMTDGN--FLKTSCGSPNYAAPEVIS-GKL
HsBRSK1     VSALDFCHSYSICHRDLKPENLLLDEKNN-IRIA---DFGMA-----SLQVGDS--LLETSCGSPHYACPEVIK-GEK
HsBRSK2     ISALDFCHSHSICHRDLKPENLLLDEKNN-IRIA---DFGMA-----SLQVGDS--LLETSCGSPHYACPEVIR-GEK
DmCG6114    ISALDFCHSHSICHRDLKPENLLLDEKNN-IKIA---DFGMA-----SLQPAGS--MLETSCGSPHYACPEVIR-GEK
CeSAD1      ISALDFCHAHNICHRDLKPENLLLDERNN-IKVA---DFGMA-----SLQVEGS--MLETSCGSPHYACPEVIR-GEK
ScGin4p     IIGVSYCHALGIVHRDLKPENLLLDHKYN-IKIA---DFGMA-----ALETEGK--LLETSCGSPHYAAPEIVS-GIP
ScKcc4p     IIGISYCHALGIVHRDLKPENLLLDSFYN-IKIA---DFGMA-----ALQTDAD--LLETSCGSPHYAAPEIVS-GLP
ScHsl1p     VEGVSYCHSFNICHRDLKPENLLLDKKNRRIKIA---DFGMA-----ALELPNK--LLKTSCGSPHYASPEIVM-GRP
ScKin4p     ISGVNYMHYKGLVHRDLKLENLLLDKHEN-LVIT---DFGFVN----EFFEDNE--LMKTSCGSPCYAAPELVVSTKA
ScLpi5p     ISGVHYIHSKGLVHRDLKLENLLLDKNEN-LVIT---DFGFVN----EFCSRNE--LMKTSCGSPCYAAPELVISAEP
ScKin1p     ASALIYLHANNIVHRDLKIENIMISDSSE-IKII---DFGLSN-----IYDSRK--QLHTFCGSLYFAAPELLK-ANP
ScKin2p     ASALQYLHANNIVHRDLKIENIMISSSGE-IKII---DFGLSN-----IFDYRK--QLHTFCGSLYFAAPELLK-AQP
HsMELK      VSAVAYVHSQGYAHRDLKPENLLFDEYHK-LKLI---DFGLCAKP-KGNKDYH----LQTCCGSLAYAAPELIQ-GKS
CePIG1      VSAIAFVHSQGYAHRDLKPENLLLTEDLH-LKLI---DFGLCAKTEKGRIDKHN---LDTCCGSPAYAAPELIQ-GLQ
HsHUNK      ISAVEHLHRAGVVHRDLKIENLLLDEDNN-IKLI---DFGLSNCA--GILGYSD--PFSTQCGSPAYAAPELLA-RKK
HsNIM1      VSAVKHMHENQIIHRDLKAENVFYTSNTC-VKVG---DFGFS-----TVSKKGE--MLNTFCGSPPYAAPELFR-DEH
DmCG4629    LLAVKHMHSLGYVHRDIKAENVLLLSEDR-LKLA---DFGFS-----TQLINGANQKLDTFCGSPPYAAPELFS-DDH
CeF49C5.4   VSAVSHMHSRNIVHRDIKAENVMFSSPNT-VKLV---DFGFS-----CLVDREQ--MLRTFCGSPPYAAPELFQ-DTS
DmCG8485    LRAITYCHQLHVVHRDLKPENVVFFEKLGLVKLT---DFGFS-----NKFLPGQ--KLETFCGSLAYSAPEILL-GDS
HsSNRK      VHAISYCHKLHVVHRDLKPENVVFFEKQGLVKLT---DFGFS-----NKFQPGK--KLTTSCGSLAYSAPEILL-GDE
CeZK524.4   MTAIDYCHQLHVVHRDLKPENVVFFEKLGMVKLT---DFGFS-----NSYEPGE--QLNTSCGSLAYSAPEILL-GDS
HsTSSK1     SLAIKYCHDLDVVHRDLKCDNLLLDKD-FNIKLS---DFSFSKRC--LRDDSGRMALSKTFCGSPAYAAPEVLQ-GIP
HsTSSK2     SSAVKYCHDLDIVHRDLKCENLLLDKD-FNIKLS---DFGFSKRC--LRDSNGRIILSKTFCGSAAYAAPEVLQ-SIP
DmCG14305   SKALKYLHNLDIAHRDLKCENILLSKR-LNIKLA---DFGFARYC---RDDNGREMKSETYCGSAAYAAPEVVC-GRP
```

```
HsTSSK3     VEAIRYCHGCGVAHRDLKCENALLQG--FNLKLT---DFGFAKV-----LPKSHRELSQTFCGSTAYAAPEVLQ-GIP
HsSSTK      AGAVRYLHDHHLVHRDLKCENVLLSPDERRVKLT---DFGFGR------QAHGYPDLSTTYCGSAAYASPEVLL-GIP
DmCG9222    VSAVEYIHSKGVVHRDIKCENLLLDEN-WNLKLI---DFGFARKD--TRTSDNQVILSKTFCGSYAYASPEILK-GVA
CeC27D6.11  IAALMHLQSYNIVHRDLKCENIFLDKH-ENVKLG---DFGFSR-------ILKPGEKSGTFCGSRAYVAPEILR-GRE
CeY38H8A.4  IAALMHLQSYNIVHRDLKCENIFLDKY-ENVKLG---DFGFSR-------ILKPGEKSGTFCGSRAYVAPEILR-GRE
CeB0511.4   IAALTHLQSYNIVHRDLKCENIFMDKH-GNVKLG---DFGFSR-------ILKPGEKSGTFCGSRAYVAPEIFR-GRE
Input data matrix (continued):

            22222222222222222222222222222222222222222222222222222222222223333333333333
            33333444444444455555555566666666667777777777888888888999999999900000000111
Taxon/Node  56789012345678901234567890123456789012345678901234567890123456789012345678901
-------------------------------------------------------------------------------------------
HsMARK3     YDGPEVDVWSLGVILYTLVSGSLPFDGQ-------NLKELRERVLRGKYR-IPFY-----MSTDCENLLKRFLVLNPI
HsMARK1     YDGPEVDVWSLGVILYTLVSGSLPFDGQ-------NLKELRERVLRGKYR-IPFY-----MSTDCENLLKKLLVLNPI
HsMARK2     YDGPEVDVWSLGVILYTLVSGSLPFDGQ-------NLKELRERVLRGKYR-IPFY-----MSTDCENLLKKFLILNPS
HsMARK4     YDGPEVDIWSLGVILYTLVSGSLPFDGH-------NLKELRERVLRGKYR-VPFY-----MSTDCESILRRFLVLNPA
DmPAR1      YDGPEVDVWSLGVILYTLVSGSLPFDGS-------TLRELRERVLRGKYR-IPFY-----MSTDCENLLRKFLVLNPA
CePAR1      YDGPEVDVWSLGVILYTLVSGSLPFDGQ-------NLKELRERVLRGKYR-IPFY-----MSTDCENLLKKFLVINPQ
DmKP78b     YSGPEVDSWSLGVVLYTLVSGSLPFDGT-------NLKELRDRVLRGKYR-VPYY-----VSIECESLIRKFLVLNPT
DmKP78a     YAGPEVDAWSLGVVLYTLVSGSLPFDGG-------TLKELRERVLRGKYR-VPYY-----ISMDCENLMRKFLVLNPA
HsSIK       YEGPQLDIWSLGVVLYVLVCGSLPFDGP-------NLPTLRQRVLEGRFR-IPFF-----MSQDCESLIRRMLVVDPA
HsQIK       YEGPQLDIWSMGVVLYVLVCGALPFDGP-------TLPILRQRVLEGRFR-IPYF-----MSEDCEHLIRRMLVLDPS
DmCG4290    YTGPEIDIWSLGVVLYVLVCGALPFDGS-------TLQSLRDRVLSGRFR-IPFF-----MSSECEHLIRRMLVLEPT
HsQSK       YDGPKVDIWSLGVVLYVLVCGALPFDGS-------TLQNLRARVLSGKFR-IPFF-----MSTECEHLIRHMLVLDPN
DmCG15072   YDGPKSDIWSLGVVLYALVCGALPFDGK-------TILELKSRVVLGKFR-IPFF-----MSQECEQLIRNMLVVEPD
CeKIN29     YDGMKADIWSMGVLLYILVAGGFPFPSD-------SVNKLKRSVLSGLVK-IPYW-----VSVECADFIRKMLVLNPG
HsARK5      YRGPEVDSWALGVLLYTLVYGTMPFDGF-------DHKNLIRQISSGEYR-EPTQ------PSDARGLIRWMLMVNPD
HsSNARK     YTGPEVDSWSLGVLLYILVHGTMPFDGH-------DHKILVKQISNGAYR-EPPK------PSDACGLIRWLLMVNPT
CeUNC82     YKGPEVDCWSLGILLYTLVYGSMPFDGR-------DFNRMVRQIKRGAYF-EPET------PSTASMLIRNMLRVNPE
DmCG11870   YQGPEVDCWSLGVLLYTLVYGSMPFDGS-------NFKRLVKQISQGDYY-EPRK------PSRASTLIRDMLTVCPR
HsAMPKa1    YAGPEVDIWSSGVILYALLCGTLPFDDD-------HVPTLFKKICDGIFY-TPQY-----LNPSVISLLKHMLQVDPM
HsAMPKa2    YAGPEVDIWSCGVILYALLCGTLPFDDE-------HVPTLFKKIRGGVFY-IPEY-----LNRSVATLLMHMLQVDPL
CeAAK2      YAGPEVDVWSCGVILYALLCGTLPFDDE-------HVPSLFRKIKSGVFP-TPDF-----LERPIVNLLHHMLCVDPM
DmSNF1A     YAGPEVDIWSCGVILYALLCGTLPFDDE-------HVPTLFRKIKSGIFP-IPEY-----LNKQVVNLVCQMLQVDPL
CeAAK1      YVGPEVDLWSCGVILYAMLCGTLPFDDQ-------NVPTLFAKIKSGRYT-VPYS-----MEKQAADLISTMLQVDPV
ScSnf1p     YAGPEVDVWSCGVILYVMLCRRLPFDDE-------SIPVLFKNISNGVYT-LPKF-----LSPGAAGLIKRMLIVNPL
HsBRSK1     YDGRRADMWSCGVILFALLVGALPFDDD-------NLRQLLEKVKRGVFH-MPHF-----IPPDCQSLLRGMIEVEPE
HsBRSK2     YDGRKADVWSCGVILFALLVGALPFDDD-------NLRQLLEKVKRGVFH-MPHF-----IPPDCQSLLRGMSEVDAA
DmCG6114    YDGRKADVWSCGVILYALLVGALPFDDD-------NLRQLLEKVKRGVFH-IPHF-----VPPDCQSLLRGMIEVNPD
CeSAD1      YDGRKADVWSCGVILYALLVGALPFDDD-------NLRNLLEKVKRGVFH-IPHF-----VPADVQSLLRAMIEVDPG
ScGin4p     YQGFASDVWSCGVILFALLTGRLPFDEED-----GNIRTLLLKVQKGEFE-MPSD---DEISREAQDLIRKILTVDPE
ScKcc4p     YEGFASDVWSCGVILFALLTGRLPFDEEN-----GNVRDLLLKVQKGQFE-MPND---TEISRDAQDLIGKILVVDPR
ScHsl1p     YHGGPSDVWSCGIVLFALLTGHLPFNDD-------NIKKLLLKVQSGKYQ-MPSN-----LSSEARDLISKILVIDPE
ScKin4p     YEARKADVWSCGVILYAMLAGYLPWDDDHENPTGDDIARLYKYITQTPLK-FPEY-----ITPIPRDLLRRILVPNPR
ScLpi5p     YEARKADIWSCGVILYAILAGYLPWDDDPNNPEGSDIGRLYNYINSTPLK-FPDY-----ILPIPRDLLRRMLVSDPK
ScKin1p     YTGPEVDVWSFGVVLFVLVCGKVPFDDE-------NSSVLHEKIKQGKVE-YPQH-----LSIEVISLLSKMLVVDPK
ScKin2p     YTGPEVDIWSFGIVLYVLVCGKVPFDDE-------NSSILHEKIKKGKVD-YPSH-----LSIEVISLLTRMIVVDPL
HsMELK      YLGSEADVWSMGILLYVLMCGFLPFDDD-------NVMALYKKIMRGKYD-VPKW-----LSPSSILLLQQMLQVDPK
CePIG1      YKGNEADVWSMGILLYTLLVGALPFEDD-------NMQIMYKKIQSGCFY-EPEF-----LSPLSKQLLRAMLQVVPE
HsHUNK      Y-GPKIDVWSIGVNMYAMLTGTLPFTVEP-----FSLRALYQKMVDKEMNPLPTQ-----LSTGAISFLRSLLEPDPV
HsNIM1      YIGIYVDIWALGVLLYFMVTGTMPFRAE-------TVAKLKKSILEGTYS-VPPH-----VSEPCHRLIRGVLQQIPT
DmCG4629    YIGAPVDVWALGILLYFMVVGNMPFRAP-------TIPGLKAAILKGDYL-LPGQ-----LSLPCIRLIQRILIHIPA
CeF49C5.4   YAGELVDVWALGVLLFFMLIGVTPFKAE-------TVPDMKVLITAGKYQ-IPDY-----VSLLATELIKSMLKTDTG
DmCG8485    YDAPAVDIWSLGVILYMLVCGQAPFEKA-------NDSETLTMIMDCKYT-VPSH-----VSTDCRDLIASMLVRDPK
HsSNRK      YDAPAVDIWSLGVILFMLVCGQPPFQEA-------NDSETLTMIMDCKYT-VPSH-----VSKECKDLITRMLQRDPK
CeZK524.4   YDAPAVDVWSLGVILYMLVCGRLPFQEA-------NDSETLTKILDCKYS-IPDV-----LSDECRNLIQSMLVREPQ
HsTSSK1     YQPKVYDIWSLGVILYIMVCGSMPYDDS-------NIKKMLRIQKEHRVN-FPR--SK-HLTGECKDLIYHMLQPDVN
```

```
HsTSSK2     YQPKVYDIWSLGVILYIMVCGSMPYDDS-------DIRKMLRIQKEHRVD-FPR--SK-NLTCECKDLIYRMLQPDVS
DmCG14305   YDPKLADAWSLGVILFIMMNAKMPFDDS-------NLTKLLEDQRNRKFA-FRRKLQE-TISAQAKATVSVLLEPEAH
HsTSSK3     HDSKKGDVWSMGVVLYVMLCASLPFDDT-------DIPKMLWQQQ-KGVS-FPT--HL-SISADCQDLLKRLLEPDMI
HsSSTK      YDPKKYDVWSMGVVLYVMVTGCMPFDDS-------DIAGLPRRQK-RGVL-YPE--GL-ELSERCKALIAELLQFSPS
DmCG9222    YDPFMSDIWACGVVCYAMVFGRLPYDGS-------NVHILLKRIN-QSLV-FPK--SP-SASSECKHMIMHILAP-VK
CeC27D6.11  YSGNAVDVWSTGVILYIMLVGSMPFDDR-------NPTKMIERQLAHKIK-FPK--LC-TASVQSKALILEILQPHAP
CeY38H8A.4  YSGNAVDVWSTGVILYIMLVGTMPFDDR-------DPTRMIERQLAHKIK-FGK--TC-TASIHSKALILEILQPHAP
CeB0511.4   YSGNAVDVWSTGVILYIMLAGSMPFDDR-------DPRKMIERQLAHKIK-FPK--SC-TSSVFSKALVLEILQPHAP
Input data matrix (continued):

            333333333333333
            111111122222222
Taxon/Node  345678901234567
---------------------------
HsMARK3     KRGTLEQIMKDRWIN
HsMARK1     KRGSLEQIMKDRWMN
HsMARK2     KRGTLEQIMKDRWMN
HsMARK4     KRCTLEQIMKDKWIN
DmPAR1      KRASLETIMGDKWMN
CePAR1      RRSSLDNIMKDRWMN
DmKP78b     QRTSLSAVMADRWIN
DmKP78a     KRTSLSAVMSDKWIN
HsSIK       RRITIAQIRQHRWMR
HsQIK       KRLTIAQIKEHKWML
DmCG4290    RRYTIDQIKRHRWMC
HsQSK       KRLSMEQICKHKWMK
DmCG15072   RRYTIKQIIKHRWLS
CeKIN29     KRYTIQNVLQHRWMH
HsARK5      RRATIEDIANHWWVN
HsSNARK     RRATLEDVASHWWVN
CeUNC82     RRATIFDIASHWWLN
DmCG11870   KRASIEQICSHWWVN
HsAMPKa1    KRATIKDIREHEWFK
HsAMPKa2    KRATIKDIREHEWFK
CeAAK2      KRATIKDVIAHEWFQ
DmSNF1A     KRANIEEIKKHEWFQ
CeAAK1      KRADVKRIVNHSWFR
ScSnf1p     NRISIHEIMQDDWFK
HsBRSK1     KRLSLEQIQKHPWYL
HsBRSK2     RRLTLEHIQKHIWYI
DmCG6114    RRLTLAEINRHPWVT
CeSAD1      KRYSLADVFKHPWVS
ScGin4p     RRIKTRDILKHPLLQ
ScKcc4p     QRIKIRDILSHPLLP
ScHsl1p     KRITTQEILKHPLIK
ScKin4p     RRINLQTIKRHVWLK
ScLpi5p     KRINLKQIKKHEWLK
ScKin1p     RRATLKQVVEHHWMV
ScKin2p     RRATLKNVVEHPWMN
HsMELK      KRISMKNLLNHPWIM
CePIG1      RRISVKKLLEHDWLN
HsHUNK      KRPNIQQALANRWLN
HsNIM1      ERYGIDCIMNDEWMQ
DmCG4629    QRPTIDDMLNSQFVT
CeF49C5.4   QRADIDSVKKHFWMR
DmCG8485    KRATVEEIASSAWLK
HsSNRK      RRASLEEIENHPWLQ
```

```
CeZK524.4   KRASLEKIVSTSWVQ
HsTSSK1     RRLHIDEILSHCWMQ
HsTSSK2     QRLHIDEILSHSWLQ
DmCG14305   ARWNLREILNCAWLR
HsTSSK3     LRPSIEEVSWHPWLA
HsSSTK      ARPSAGQVARNCWLR
DmCG9222    IRYNIPQVKEDPWYS
CeC27D6.11  NRPTYKAICESEWLK
CeY38H8A.4  NRPTYKAICESEWLK
CeB0511.4   NRPTYKAICESEWLR
```

FIGURE S2.—The alignment of all sequences included in the phylogenetic analyses was generated by Clustal X. The following sites were excluded from the analysis when estimating trees and performing bootstrap analysis in both parsimony and neighbor-joining sites: 5, 23-26, 30-35, 45-80, 119, 143, 186, 191-193, 199-203, 211, 212, 231, 263-269, 285, 290-294.