ARTICLE

The Caenorhabditis elegans orthologue of the human gene responsible for spinal muscular atrophy is a maternal product critical for germline maturation and embryonic viability

Irene Miguel-Aliaga, Emmanuel Culetto, Denise S. Walker, Howard A. Baylis, David B. Sattelle and Kay E. Davies

Department of Human Anatomy and Genetics, University of Oxford, South Parks Road, Oxford OX1 3QX, UK, Babraham Institute Laboratory of Molecular Signalling and Department of Zoology, University of Cambridge, Cambridge CB2 3EJ, UK

Received June 9, 1999; Revised and Accepted August 11, 1999

INTRODUCTION

Spinal muscular atrophy (SMA) is a common disorder characterized by loss of lower motor neurones of the spinal cord. The disease is caused by mutations in the survival motor neurone (SMN) gene. SMN is ubiquitously expressed and evolutionarily conserved, and its role in RNA processing has been well established. However, these properties do not explain the observed specificity of motor neurone death. To gain further insight into the function of SMN, we have isolated and characterized the Caenorhabditis elegans orthologue of the SMN gene (CeSMN). Here we show that CeSMN is transmitted maternally as a predominantly nuclear factor, which remains present in all the blastomeres throughout embryonic development and onwards into adulthood. In adult nematodes, a CeSMN–green fluorescent protein fusion protein is expressed in a number of cell types including the germline. Both disruption of the endogenous CeSMN function and overexpression of the gene result in a severe decrease in the number of progeny and in locomotive defects. In addition, its transient knockdown leads to sterility caused by a defect in germ cell maturation. The expression pattern and functional properties so far observed for CeSMN, together with its unusual behaviour in the germline, indicate that SMN may be involved in specific gene expression events at these very early developmental stages. We have also identified a deletion in the CeSMN promoter region in egl-32. This mutant may become a useful genetic tool with which to explore regulation of CeSMN and hence provide possible clues for novel therapeutic strategies for SMA.

DDBJ/EMBL/GenBank accession no. AF156887
tates cloning and genetic analysis (11). With respect to SMA, it offers an excellent system for the developmental study of the tissues involved, in that muscle cells and neurones make up ~45% of its total of 959 cells and the wiring diagram of its nervous system is extremely well characterized (9).

The existence of such a suitable system prompted us to study the nematode SMN gene in detail, in the hope that it would help us to overcome some of the problems that the current mouse model presents (the null mutant mouse is embryonic lethal) and would therefore shed new light on the aetiology of SMA. The results we present herein both characterize a new model system for the study of SMA and describe a novel key developmental function for SMN as a maternal factor and in the development of the germline.

RESULTS

Identification and characterization of the C. elegans SMN orthologue

A C. elegans genomic sequence (cosmid C41G7; GenBank accession no. Z81048) showing homology to the human SMN gene had been reported previously (12). We successfully amplified using RT–PCR and subsequently sequenced a cDNA fragment corresponding to a large part of the putative mRNA.

The ends of the transcript were then determined by RACE and RT–PCR using the SL-1 spliced leader (shaded box) at the trans-splice site (TTTCAG) 3 bp upstream of the initiator codon. The polyadenylation signal is underlined, 15 bp before the poly(A) tail. Many C. elegans genes have one of a limited group of tolerated mismatches to the consensus. TATAAA is the second most commonly used variant (53). A variation in one splice site (open box) was detected in one of the transcripts analysed. In this case, splicing occurred 3 bp downstream, giving rise to a slightly shorter cDNA (TTA instead of TCA GTA). This resulted in the substitution of SV for L in the amino acid sequence. (B) Putative regulatory region of the CeSMN gene. The TATAAAA box is located ~133 bp upstream of the START ATG, flanked by two SKN-1 sites (consensus sequence in boxes, usually associated with adjacent A/T-rich regions). The deletion found in egf-32 is underlined between the two arrows.

Figure 1. Sequence analysis of the CeSMN region. (A) The complete cDNA sequence of CeSMN. The CeSMN messenger is trans-spliced with the SL-1 spliced leader (shaded box) at the trans-splice site (TTTCAG) 3 bp upstream of the initiator codon. The polyadenylation signal is underlined, 15 bp before the poly(A) tail. Many C. elegans genes have one of a limited group of tolerated mismatches to the consensus. T

TATAAA is the second most commonly used variant (53). A variation in one splice site (open box) was detected in one of the transcripts analysed. In this case, splicing occurred 3 bp downstream, giving rise to a slightly shorter cDNA (TTA instead of TCA GTA). This resulted in the substitution of SV for L in the amino acid sequence. (B) Putative regulatory region of the CeSMN gene. The TATAAAA box is located ~133 bp upstream of the START ATG, flanked by two SKN-1 sites (consensus sequence in boxes, usually associated with adjacent A/T-rich regions). The deletion found in egf-32 is underlined between the two arrows.

Figure 1. Sequence analysis of the CeSMN region. (A) The complete cDNA sequence of CeSMN. The CeSMN messenger is trans-spliced with the SL-1 spliced leader (shaded box) at the trans-splice site (TTTCAG) 3 bp upstream of the initiator codon. The polyadenylation signal is underlined, 15 bp before the poly(A) tail. Many C. elegans genes have one of a limited group of tolerated mismatches to the consensus. TATAAA is the second most commonly used variant (53). A variation in one splice site (open box) was detected in one of the transcripts analysed. In this case, splicing occurred 3 bp downstream, giving rise to a slightly shorter cDNA (TTA instead of TCA GTA). This resulted in the substitution of SV for L in the amino acid sequence. (B) Putative regulatory region of the CeSMN gene. The TATAAAA box is located ~133 bp upstream of the START ATG, flanked by two SKN-1 sites (consensus sequence in boxes, usually associated with adjacent A/T-rich regions). The deletion found in egf-32 is underlined between the two arrows.
The nematode SMN protein exhibits a significant overall homology to its human counterpart (Fig. 2, smallest sum probability < 6.6e–29), the similarity being particularly strong for those regions which are most conserved amongst other SMN orthologues. This suggests an associated functional role, conserved throughout evolution. Interestingly, both the human and the nematode SMN protein sequences contain a single tudor domain. Such tudor domains have been reported previously in proteins with an RNA-binding function in development and/or a protein-binding role during RNA metabolism or transport (16).

The CeSMN mRNA is expressed uniformly throughout development

We used a full-length CeSMN cDNA clone as a probe on a northern blot. As can be seen in Figure 3A, a single CeSMN transcript is present from the embryonic stage. Its size confirms our RACE and RT–PCR results. To check whether its levels were developmentally regulated, we performed semiquantitative PCR on staged cDNA samples obtained at 2 h intervals from early L1 to adult. As shown in Figure 3B, the overall levels of CeSMN mRNA do not appear to vary significantly throughout larval development.

Figure 2. Evolutionary conservation of SMN orthologues. The multiple alignment was constructed using ClustalW (54,55) and displayed using the BOXHADDE server (K. Hofmann and M. Baron, unpublished data; http://www.ch.embnet.org/software/BOX_form.html ). Species: mus, Mus musculus; rat, Rattus norvegicus; dog, Canis familiaris; bos, Bos taurus; hum, Homo sapiens; zeb, Danio rerio; Cel, Caenorhabditis elegans. The overall similarity between the human and the nematode sequences is 36%. Homology is particularly significant at those regions that are thought to be functionally significant, i.e. amino acids 13–65 [SIP1 (6) and RNA (15,26) binding domain], amino acids 91–143 [tudor domain (16)] and amino acids 240–278 [binding to Sm proteins (6) and oligomerization (5) domain] in the human sequence.

Figure 3. The CeSMN messenger. (A) Northern analysis. A 500 bp CeSMN cDNA probe detects a single band of ~800 bp in both embryonic (e) and mixed-stages (ms) total RNA. (B) Semi-quantitative RT–PCR on staged cDNAs. The cDNA was amplified using two sets of primers, one set specific to CeSMN and the second to an internal control, ama-1, the levels of which remain constant throughout development (39). The number of PCR cycles was kept to a minimum to avoid saturation. As the larvae develop, there are no significant changes in the relative abundance of the CeSMN transcript with respect to the ama-1 control.
The CeSMN protein is expressed in all the nuclei of C. elegans embryos from the zygotic stage

In order to analyse the distribution of the CeSMN product, we generated a polyclonal anti-peptide antibody (amino acids 125–146). The antibody recognizes a single band of ~24 kDa on a western blot (data not shown). The CeSMN protein is expressed from the zygotic stage in both pronuclei and in all the blastomeres throughout embryonic development (Fig. 4). A strong homogeneous staining is present in all of the interphasic nuclei, although some diffuse fluorescence is also visible in the cytoplasm. In mitotic stages, the CeSMN protein disappears from the nucleoplasm, presumably being redistributed diffusely to the cytoplasm, and is re-incorporated into the newly formed nuclei (Fig. 4A). Since de novo zygotic transcription does not start until the 4-cell stage in C. elegans (17), the presence of CeSMN at such early stages indicates that the CeSMN protein and (perhaps) its mRNA are transmitted maternally at some point during germline development.

While analysing the embryos, we were able to observe the expression pattern of the endogenous CeSMN protein in newly hatched larvae (Fig. 4D). CeSMN is present in virtually all the cell nuclei of the young nematodes. This finding is in accord with the localization of the green fluorescent protein (GFP) fusion protein in adults described below.

The expression pattern of CeSMN in adults is heterogeneous

We investigated the spatial pattern of expression in adults by GFP tagging. We generated a CeSMN–GFP fusion construct containing the full-length CeSMN gene and ~3.5 kb of its upstream region. The transgenic adult nematodes showed intense fluorescence in the nuclei of many somatic cells including neurones, body wall and vulval muscle cells, hypodermal and gut cells, and the excretory cell (Fig. 5). Interestingly, the excretory cell canals and neuronal processes were clearly fluorescent, whereas no significant cytoplasmic signal was apparent for other cell types.

Furthermore, we also detected strong fluorescence in the nuclei of the germ cells (Fig. 5C), despite the fact that the C. elegans germline usually exhibits a striking ability to silence transgene arrays (18). Germline expression of CeSMN was confirmed subsequently by immunofluorescence (data not shown).

Curiously, although five independent stable transformed lines were established (as indicated by the inheritable rol-6 phenotype), the GFP signal was lost in all of them after one or two generations.

Down-regulation of CeSMN leads to severe effects in early development

We used RNA interference to decrease the endogenous levels of CeSMN product (Fig. 6). While the CeSMN dsRNA had no obvious effect on the injected individuals, a marked reduction in the number of F1 descendants was observed. This can be explained by the presence of a number of laid eggs arrested at a late stage of embryonic development. Reduced levels of maternal and/or early zygotic CeSMN may account for such a drastic effect on embryonic viability. The few surviving individuals showed a variety of deficiencies mainly involving neuronal, muscular and reproductive tissues. These included uncoordination, paralysis, lack of muscle tone, vulval abnormalities, molting defects and sterility. It is noteworthy that the less severely affected individuals, in which the knockdown allowed for nearly normal development in all other respects, were sterile. In these cases, the development of the somatic gonad was apparently normal, but the syncitial germ cells failed to proliferate and differentiate into oocytes. This reflects an inability to cope with even small changes in the level of CeSMN at this stage, this factor thus being absolutely necessary for the development and differentiation of the germline.
Increased levels of CeSMN also affect embryonic viability

In order to determine whether an excess of CeSMN product also had an effect, we overexpressed the full-length CeSMN ectopically in all the tissues of *C. elegans* adults. As shown in Table 1, the overexpression interfered strongly with the normal development of the progeny; the number of laid embryos was again drastically reduced in ~65% of the cases. Thus, the levels of CeSMN are particularly critical during very early embryogenesis, since both excess and defective expression are deleterious. Following heat shock, ~50% of the nematodes also showed variable degrees of locomotion defects. Abnormal pigmentation and a bag-of-worms phenotype (internal hatching of the eggs) were observed sporadically.

It is also worth noting that the extrachromosomal arrays containing CeSMN driven by the heat shock promoters appeared to be expressed efficiently in the germ cells, since the observed phenotype suggests that this cell type was the more severely affected in our experiments. The expression of this kind of array in the germline is usually poor (19).

The interaction properties of SMN are evolutionarily conserved

Defects in full-length SMN self-oligomerization have been shown in humans to correlate with the severity of SMA (5). The region of human SMN involved in this interaction (Fig. 2) is located at the C-terminus of the protein and is well conserved down to the nematode (12). Using a yeast two-hybrid system, we tested whether CeSMN is able to self-associate. Table 2 summarizes the results. The CeSMN protein interacts with itself. Its C-terminal domain is necessary and sufficient for this self-association to occur, as a HIS3*+*lacZ*+* phenotype was evident only in those cases in which both of the fusion proteins tested for interaction contained this domain.

If protein function is to be equivalent, one would expect not only self-association properties but also binding partners to be conserved. Several proteins have been identified in previous yeast two-hybrid screens using the human SMN as bait (6,20). One of them, fibrillarin, is present and extremely well conserved in *C. elegans* (75% identity). For this reason, we designed primers based on the predicted nematode sequence (cosmid T01C3; GenBank accession no. Z78413) and amplified the full-length fibrillarin from total *C. elegans* RNA by RT–PCR, thereby validating the GeneFinder prediction. When its ability to interact with CeSMN was assessed using a yeast two-hybrid system under the same conditions, co-expression also resulted in a HIS3*+*lacZ*+* phenotype. A shorter fibrillarin without the N-terminus, comprising the region which is homologous to the human protein (amino acids 111–352), retained the ability to associate.

The *C. elegans egl-32* mutant has a deletion in the CeSMN promoter region

Using ACEDB (http://wormsrv1.sanger.ac.uk/cgi-bin/ace/simple/worm), we identified several previously characterized genetic loci (*egl-32, egl-33* and *eat-16*) mapping close to *nob-3*, which we knew to be very near CeSMN. We isolated the CeSMN gene and flanking sequences for homozygous strains...
carrying alleles of these genes and determined their DNA sequence. We found that one of them, \textit{egl-32}, has a 12 bp deletion in the \textit{CeSMN} 5'-UTR immediately before the TATAAA box (Fig. 1B, nucleotides –141 to –152 upstream of the start ATG). This results in the loss of one of the two putative \textit{SKN-1} sites (TRANSFAC accession no. M00230) flanking the TATAAA box. \textit{egl-32} nematodes are recessive, egg-laying-defective mutants which exhibit abnormalities at the level of gonad formation and germline development (21). This, together with the presence of the deletion in a critical region of the \textit{CeSMN} promoter region, suggests that the anomalous transcription of \textit{CeSMN} may be the molecular basis for the \textit{egl-32} phenotype.

\textbf{DISCUSSION}

SMA is a frequent cause of infant mortality for which there is, to date, no effective treatment. The isolation of the gene responsible for the disease 4 years ago contributed to the development of accurate prenatal diagnosis methods (22,23). However, a number of SMA cases arise from \textit{de novo} mutational events in the \textit{SMN} region (24,25). Therefore, the development of postnatal therapeutic strategies is crucial. This depends on acquiring precise knowledge of how and why cell death of a particular subset of motor neurones occurs. It also relies on the existence of simple, suitable models with which to study the function and interactions of the \textit{SMN} gene; such models must provide an environment that recreates the complexity of the human situation to a certain extent. In this regard, \textit{C.elegans} is ideal, the evolutionary conservation of the \textit{SMN} gene lending justification to this approach.

\textbf{Sequence analysis: evolutionary considerations}

Comparison of the nematode and the human \textit{SMN} amino acid sequences allows the delineation of three conserved regions...
(Fig. 2). The N-terminus of both proteins is very well conserved. This region has been shown to be involved in RNA binding (26) and interaction with SMN-interacting protein 1 (SIP1) in humans (6). However, to date, there are no apparent SIP1 orthologues in C. elegans. As has been suggested previously (15), this may imply that the high degree of conservation is (at least partly) due to conservation of the RNA-binding capacity. In this regard, the finding that CeSMN is expressed in the germline and early embryos, together with the fact that the central domain of both human and nematode proteins shares homology to the Drosophila melanogaster tudor protein (27), suggest that this RNA-binding property might be involved in developmental processes other than snRNPs biogenesis and pre-mRNA splicing. This hypothesis is discussed below. The C-terminal end of CeSMN is also similar to its human counterpart. We have found in a yeast two-hybrid assay that CeSMN self-associates through this region. Therefore, this property may account (at least in part) for the evolutionary conservation of this domain and indicates that it may have an important functional role. In effect, defects in SMN oligomerization correlate with the severity of SMA (5). However, this region has also been implicated in the interaction with Sm proteins in humans (6). Sm proteins, components of snRNPs, are involved in a process that is intrinsic to eucaryotic cells and are therefore present in C. elegans. Thus, one would expect the similarity in this region to be a consequence also of the maintenance of this function, as the ability to bind fibrillarin additionally suggests. In this regard, SMN may have an even more important role in the RNA processing machinery in the nematode, since the spliceosomal snRNPs normally involved in cis-splicing are also required for trans-splicing. Trans-splicing is the process by which many of the nematode genes receive a 22 nucleotide leader sequence. The donors in the trans-splicing reaction, the SL RNAs, occur in the form of snRNPs (13), and deletions in the SL-1 gene complex are embryonic lethal (28).

Expression studies: functional implications

The CeSMN protein appears to be essentially ubiquitous, in common with its human orthologue (3). However, as C. elegans allows us to study this expression from a developmental point of view and at the level of the whole organism, we have been able to uncover some novel properties. The CeSMN–GFP fusion protein is expressed in the nuclei of most cell types. Nevertheless, the processes of neurones and the excretory cell show a strong cytoplasmic expression. Although the subcellular localization of GFP fusion proteins is not always conclusive, this difference in distribution may account for a specialized cytoplasmic function of the SMN product in neuronal processes, and may be achieved by interaction with a neurone-specific ligand (of proteic or ribonucleic nature). This would lend evolutionary support to the recently observed presence of the human SMN protein in dendrites (29).

Our data indicate that CeSMN is transmitted as a maternal factor during development. Its abundance in the germ cell nuclei and subsequently in both pronuclei and the zygote implies that it is synthesized by the germline and incorporated into the developing gametes. The germ cells may transfer CeSMN to the gonadal cytoplasmic core as a trophic factor, which will be transported to the nuclei of future oocytes, somehow contributing to their maturation or survival. Since the CeSMN protein is required at this stage (as shown by RNA interference and the overexpression experiments), it must be performing a function. While we cannot rule out the possibility that its presence is a mere consequence of its housekeeping activity [some previously identified maternal mRNAs belong to this category (17)], some unusual observations suggest an additional role. The maturation of the germline in those individuals

<table>
<thead>
<tr>
<th>GAL4-AD plasmid</th>
<th>GAL4-BD plasmid</th>
<th>Growth on SC –Leu, –Trp, –His</th>
<th>β-Galactosidase activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>pACT2</td>
<td>pAS2-1</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>pACT/SMN1–207</td>
<td>pAS2-1</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>pACT2/SMN1–50</td>
<td>pAS2-1</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>pACT2/SMN37–115</td>
<td>pAS2-1</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>pACT2/SMN157–207</td>
<td>pAS2-1</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>pACT2/Fib1–352</td>
<td>pAS2-1</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>pACT2/Fib111–352</td>
<td>pAS2-1</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>pACT2/SMN1–207</td>
<td>pAS2-1</td>
<td>+</td>
<td>++</td>
</tr>
<tr>
<td>pACT2/SMN1–207</td>
<td>pAS2-1/SMN1–207</td>
<td>+</td>
<td>++</td>
</tr>
<tr>
<td>pACT2/SMN1–207</td>
<td>pAS2-1/Fib1–352</td>
<td>+</td>
<td>++</td>
</tr>
<tr>
<td>pACT2/SMN1–207</td>
<td>pAS2-1/Fib111–352</td>
<td>+</td>
<td>++</td>
</tr>
</tbody>
</table>

Co-expression of the two CeSMN full-length fusion proteins (SMN1–207) resulted in a HIS3+ lacZ+ phenotype, indicating that CeSMN does indeed interact with itself. An interaction was also observed between full-length CeSMN (SMN1–207) and SMN157–207, while no such interaction was detected with SMN1–50, SMN37–115, or with the empty plasmid expressing GAL4-AD alone. The same assay was also positive when full-length CeSMN (SMN1–207) was tested against full-length or a shorter version of fibrillarin (Fib1–352 and Fib111–352, respectively).
less severely affected by the dsRNA injection was strongly disrupted, while the presumably low knockdown allowed for nearly normal development in all other respects. It could be argued that, since much of the metabolic activity of the late larva and adult worm is devoted to germline development, a partial loss of function is likely to affect the germline more severely than other tissues. However, the effect was absolute and highly penetrant. Both the nematode and the human genes are tudor-like proteins (16). The sequence similarity might be due to shared functional features, in that both tudor (30,31) and CeSMN are maternal factors essential for the successful development of the germline. In addition, both remain present during adulthood. In the case of tudor, it is also the protein that is involved in this function as a maternal factor in early developmental stages. Homeless, a better characterized tudor-like protein, is required for RNA localization in D. melanogaster oogenesis (32). CeSMN can bind RNA, and presumably small changes in the levels of the CeSMN protein (both overexpression and down-regulation) have drastic effects at this stage, which suggests a function involving direct interaction of the protein with a substrate. Thus, transport is a possibility in this regard. Curiously, an ncd-like kinesin gene lies ~500 bp upstream of the CeSMN gene in the opposite orientation. Both genes might therefore share common regulatory elements.

On the other hand, it has been shown recently in C. elegans that germ cell death is extensive, apoptotic and required to maintain germline homeostasis (33). It uses the same apoptotic execution machinery (ced-3, -4 and -9) as somatic cell death. However, this machinery is activated by a distinct pathway. It has been proposed that distinct CED-9 partners may be used to control germline and somatic cell death. In humans, SMN interacts with BCL-2, the orthologue of the nematode CED-9 (20). It is therefore possible that CeSMN acts as an anti-apoptotic factor in the development of the germline. Altered levels of CeSMN protein might disrupt germline homeostasis and trigger massive apoptosis. The RNA interference phenotype we have observed is consistent with this theory.

Finally, it is worth noting the unusual expression of extra-chromosomal arrays containing the CeSMN gene in the germ line. The C. elegans germline exhibits a striking ability to silence transgene arrays (18). This property is essential for the successful development of gametes, since mutations in the genes involved result in a mes (maternal-effect sterile) phenotype (34). One can speculate that CeSMN might escape or interfere with the silencing mechanism that is usually required for the maintenance of a functional germline in C. elegans.

**SMA and C. elegans**

Our experiments have established a good model organism for the study of SMA, describing aspects of the biology of CeSMN at the genetic, molecular and cellular levels. Both overexpression and RNA interference show interesting and useful phenotypes. The former suggests that SMN up-regulation might not be as innocuous or protective as previously thought. The down-regulation of the nematode orthologue is particularly informative, in that it is not absolute (as opposed to the embryonic lethal knockout mouse) and allows the observation of differently affected individuals. We are currently generating a CeSMN knockout nematode in collaboration with the C. elegans Gene Knockout project at the Sanger Centre, which, together with the overexpression lines, will provide us with a useful genetic tool to look for suppressor genes. If the levels of CeSMN prove to be affected in egl-32, we will have a suitable mutant for this kind of study, since it is both non-lethal and temperature-sensitive. egl-32 is suppressed by daf-3, a Smad transcription factor (35). This suggests the involvement of the transforming growth factor-β signalling pathway. In this context, the relevance of the two putative SKN-1 sites, as well as the relative involvement of CeSMN and/or the ncd-like kinesin, remain to be established. SKN-1 is a maternal factor required post-embryonically for the development of the intestine and to specify the fate of ventral stomatocytes in the early embryo (36). This is reminiscent once again of functions similar to those of Tudor-like proteins such as hfs, also involved in dorsoventral and anteroposterior axis formation. Further studies are in progress to elucidate the mechanisms underlying these novel findings and to enhance our understanding of their functional significance.

**MATERIALS AND METHODS**

**Wild-type and mutant strains**

Wild-type C. elegans strain Bristol N2 cells were maintained using standard methods (37). The mutant strain egl-32 (n155) was obtained from the Caenorhabditis Genetics Centre (University of Minnesota, MN) and grown under similar conditions at 15°C in order to avoid expression of the mutant phenotype.

**Sequence characterization**

RT–PCR was performed on total RNA isolated from mixed stage N2 C. elegans, following a previously described protocol (38). First, two oligonucleotides (SmmF 5′-ACG ATG TTT GGG ATG ATA CGG AGC-3′ and SmmR 5′-TGA TAT CCG CTC ATG TAC CAG CTC-3′) based on the predicted mRNA sequence were used to confirm the transcription of the putative orthologue. A single band (~500 bp was produced and its identity was confirmed by sequencing. The same conditions were used to amplify a single ~600 bp band using SmmR and an oligonucleotide to SL-1. The sequence of the 5′ end of the mRNA was then determined by cloning and sequencing this band. Whilst sequencing these clones, one cDNA with an alternative splice variant was identified (Fig. 1A). No specific product was identified with a second spliced leader SL-2. RACE experiments were performed using a Marathon RACE kit (Clontech, UK) to corroborate the RT–PCR result. The EST clone that subsequently became available (yk373d10; GenBank accession no. C33680) was also sequenced to confirm the rest of the predicted cDNA sequence.

**Northern analysis**

Total RNA from mixed stage wild-type N2 nematodes was isolated by Trizol extraction (Life Technologies, UK) following the manufacturer’s instructions. Samples (10–20 μg) of this preparation were run on a formaldehyde-containing 1% agarose gel at 4 V/cm and subsequently transferred onto a nylon membrane (Hybond-ND; Amersham, UK). The membrane was pre-hybridized for 2 h and hybridized overnight at 65°C in 0.5 M sodium phosphate pH 7, 1 mM EDTA,
7% SDS, 1% bovine serum albumin (BSA) with a 500 bp PCR-amplified CeSMN cDNA labelled by random priming (Rediprime; Amersham). The membrane was washed three times for 10 min in 2× SSC, 0.25% SDS, 25 mM sodium phosphate pH 7, at 65°C followed by one 15 min wash in 0.2× SSC, 0.25% SDS, 25 mM sodium phosphate pH 7, at 65°C.

**Semi-quantitative RT–PCR**

Semi-quantitative PCR was performed on a time course of cDNA samples obtained at 2 h intervals from early L1 to young adult (39). The cDNA was amplified using two sets of primers, one set specific to CeSMN (forward primer, 5′-ATG GCA AAA ATC TGG TCG AAA AGT G-3′ and reverse primer, 5′-TCC ATC TTC TCC TTT GAA TTT TC-3′) and the second to an internal control, ama-1, the levels of which remain constant throughout (39). In each case, oligonucleotides were chosen to span an intron. Cycling was performed under the following conditions: (i) 94°C 2 min x1; (ii) 94°C 30 s, 57°C 30 s, 72°C 1 min x35; and (iii) 72°C 7 min. The number of PCR cycles was kept to a minimum to avoid reaching the plateau phase of the reaction. PCRs were separated by electrophoresis in a 2% agarose gel.

**Immunofluorescence of embryos**

Antibody generation. A polyclonal anti-CeSMN antibody (Ab023) was raised against the CeSMN peptide NH2- DLQQTKKTT-STVNSVHAHNSKSTC-COOH (amino acids 125–146) by injecting it conjugated to PPD. Antibody generation, subsequent injections into rabbits and resulting antisera collections were performed by The Babraham Institute Microchemical Facility (Babraham Institute, Cambridge, UK). The antibody was subsequently affinity purified on a peptide antigen affinity column (Cymbus Biotechnology, UK).

**Immunofluorescence in embryos.** Immunofluorescence was performed as previously described (40). Briefly, C. elegans embryos were collected in hypochlorite solution, cracked on dry ice, fixed in methanol/acetone and incubated overnight with the primary antibody diluted 1/25. The anti-CeSMN polyclonal antibody was detected with a goat anti-rabbit antibody conjugated to Alexa 488 diluted 1/175 (Molecular Probes, Cambridge Bioscience, UK). The secondary antibody was cleaned with C. elegans acetone powder prior to use, according to a previously described protocol (41). Propidium iodide 0.2 µg/ml was used as a control of nuclear staining.

**GFP reporter construct**

The reporter gene construct was created using a PCR fusion approach. Primers based on the genomic information from the C41G7 cosmID sequence were designed to amplify a 4.5 kb genomic DNA fragment containing the full-length SMN gene and ~3.5 kb of its upstream region. This includes part of the putative upstream gene C41G7.2, which is transcribed in the opposite direction. The primer sequences (including additional HindIII–BamHI restriction sites) were as follows: forward primer, 5′-CCC AAG CTT GGT TTA GGA GCT AG-3′ and reverse primer, 5′-GGC GAA TTC ATG TAT GAG AGA TT-3′. The resulting fragment was directionally cloned into the pPD95.77 vector (42) in order to produce a CeSMN–GFP fusion protein under the control of the putative CeSMN promoter. Wild-type hermaphrodites were transformed according to a previously described protocol (43). Extrachromosomal transgenic lines were established by injecting the construct at a concentration of 50–80 ng/µl into C. elegans syncytial gonads together with the rol-6 [pRF4 (44)] dominant marker. Seventy individuals belonging to the progeny originated from eight independent injections were analysed. None of the transgenic F1 individuals studied gave rise to a transmitted line expressing the CeSMN–GFP fusion protein.

**RNA-mediated interference**

Double-stranded RNA used for the inhibition of CeSMN was synthesized from linearized template cloned into the pGEM-T vector (Promega, UK) using either SP6 or T7 RNA polymerase (Promega). The template used was a 500 bp fragment of the CeSMN cDNA obtained by RT–PCR (using the SmnF/R primers mentioned above). dsRNA was made according to a previously described protocol (45). Briefly, sense and anti-sense RNAs were produced from separate transcription reactions using ~1 µg of the linearized plasmid. Reaction products were treated with DNase, combined and extracted with phenol/ chloroform. Prior to ethanol precipitation, the mix was incubated for 10 min at 68°C followed by 30 min at 37°C to allow annealing to take place. The final pellet was resuspended in 10 µl of RNase-free TE and injected into the syncytial gonads of adult nematodes. Injection was via a glass micropipette under the control of an Eppendorf Transjector 524C, visualized using an Axiovert 35 microscope. Details of the procedure used were described earlier (46). The progeny of 20 individuals were analysed. The penetrance of the effect was ~90%.

**Overexpression experiments**

A full-length CeSMN cDNA was put under the control of a C. elegans heat-shock protein promoter to trigger a temperature-induced CeSMN overexpression. Two vectors [pPD49.78 and pPD49.83 (42)] containing the promoter of two distinct and complementarily expressed C. elegans heat-shock proteins fused to CeSMN were used and co-injected together with the rol-6 marker into the syncytial gonads of adult nematodes as described for the GFP fusion construct. These heat-shock promoters are expressed in many tissues (19), but show particularly strong expression in intestinal cells and neurones. The heat-shock was induced by transferring transgenic individuals to 31°C for 2 h, after which they were returned to 20°C, according to a previously described protocol (47). To control for heat-shock effect, similar temperature shifts were performed on wild-type animals and on transgenic lines expressing the GFP gene under the control of HSP promoters [pPD118.26 and pPD118.28 vectors (42)]. A set of duplicated plates was maintained at 20°C for phenotype comparison. Thirty individuals originated from two independent stable transgenic lines were analysed.

**Yeast two-hybrid analysis**

Oligonucleotides to the full-length CeSMN, fibrillarin and to the three more evolutionarily conserved domains of CeSMN were designed to contain additional linkers to facilitate
cloning. The amplification products were inserted in-frame into the NcoI site downstream of the GAL4-AD in pACT2 (48), while full-length CeSMN cDNA was also inserted downstream of the GAL4-BD in pAS2–1 (49). The fragments used were as follows: SMN1–207, full-length CeSMN, amino acids 1–207; SMN1–50, N-terminal domain of CeSMN, amino acids 1–50; SMN37–115, middle domain of CeSMN, amino acids 37–115; SMN157–207, C-terminal domain of CeSMN, amino acids 157–207; Fib1–352, full-length C. elegans fibrillarin, amino acids 1–352; Fib111–352, shorter C. elegans fibrillarin, amino acids 111–352.

Interactions were tested by co-transformation into the HIS3, lacZ reporter strain CG-1945 (Clontech). GAL4-AD and GAL4-BD fusions proteins were identified by assessing expression of HIS3 and lacZ reporters. Transformants were tested for their ability to grow in the absence of histidine by streaking onto SC–Leu, –Trp plates. Interaction between GAL4-AD and GAL4-BD fusion proteins could be identified by measurement of β-galactosidase activity. Single-worm PCR amplification of a 1.4 kb genomic fragment was performed using Expand High Fidelity Polymerase (Roche, UK) on both mutant and wild-type nematodes. Worms were lysed according to a previously described protocol (50) and transformants selected on SC–Leu, –Trp plates. Interaction between GAL4-AD and GAL4-BD fusion proteins was identified by assessing expression of HIS3 and lacZ reporters. Transformants were tested for their ability to grow in the absence of histidine by streaking onto SC–Leu, –Trp, –His. 50 mM 3-amino-1,2,4-triazole plates and assessing growth after 3–5 days. β-Galactosidase activity was tested using a liquid culture assay, using chlorophenol red-β-D-galactopyranoside (Calbiochem, UK) as a substrate, as described by Clontech (Yeast Protocols Handbook): overnight liquid cultures in selective SC medium were diluted 1:5 in YPAD and grown at 30°C, 250 r.p.m. to mid-log phase (A600 = 0.5–0.8). Aliquots of 1.5 ml were harvested in a microfuge, washed in 1 ml of buffer 1 (100 mM HEPES, 154 mM NaCl, 4.5 mM L-aspartate, 1% BSA, 0.05% Tween-20, pH 7.25–7.30) and resuspended in 300 μl of buffer 1. Samples (100 μl) were freeze–thawed rapidly three times, then mixed with 700 μl of buffer 2 (2.23 mM chlorophenol red-β-D-galactopyranoside in buffer 1) and incubated for 4 h at 37°C, before stopping the reaction by the addition of 0.5 ml of 3 mM ZnCl2 and centrifuging briefly to remove cell debris. β-Galactosidase activity was determined by measurement of A650 and results standardized by dividing by the A600 of the culture used. All yeast media were prepared as previously described (51).

**Mutant analysis**

Single-worm PCR amplification of a 1.4 kb genomic fragment was performed using Expand High Fidelity Taq Polymerase (Roche, UK) on both mutant and wild-type nematodes. Worms were lysed according to a previously described method (52), using Expand buffer instead of the lysis buffer suggested in the protocol. The primer sequences were as follows: forward, 5′-GAC GAT AAT GAG GTA ATC ATC CC-3′ and reverse, 5′-CCT GTA GCA TCA TGC TAG ATA CC-3′. The amplified fragments, containing all the genomic region coding for the full-length messenger plus ~400 additional upstream bases, were sub-cloned into pGEM-T vectors (Promega) and double sequenced. Two independent amplification products were analysed to verify the result.

**ACKNOWLEDGEMENTS**

We are grateful to Dr Julie Ahringer for advice, to Dr Andrew Fire for providing vectors, to Dr Iain L. Johnstone for his gift of staged cDNAs, to Yuji Kohara for supplying cDNA clones, and to Nick Gower, Lara Harrington and the SMA group (Oxford) for technical support. We thank Neville Huskisson of the Babraham Institute for help in designing and generating the polyclonal CeSMN anti-peptide antibody. We are indebted to Dr Kevin Talbot for taking an interest in C. elegans SMN and initiating this project, to Dr Chris P. Ponting and Dr Marcel van den Heuvel for helpful and fruitful discussion, and to Paul D. Jump for his priceless support. We also wish to thank the Caenorhabditis Genetics Center (which is funded by the National Center for Research Resources of the NIH) for strains. This work was funded by grants from the Muscular Dystrophy Group of the UK and the Muscular Dystrophy Association of the USA (K.E.D., I.M.-A.), the Medical Research Council of the UK (C.B.S., E.C.), the Biotechnology and Biological Sciences Research Council of the UK (D.B.S., E.C.) and the University of Oxford (I.M.-A. is a Mary Googder scholar).

REFERENCES


