Interspecies Comparison Reveals Evolution of Control Regions in the Nematode Sex-Determining Gene tra-2

Patricia E. Kuwabara

Medical Research Council Laboratory of Molecular Biology, Cambridge CB2 2QH, England

Manuscript received April 24, 1996
Accepted for publication June 22, 1996

ABSTRACT

The Caenorhabditis elegans sex-determining gene tra-2 promotes female development and expresses 4.7-, 1.9- and 1.8-kb mRNAs. The 4.7-kb mRNA encodes the major feminizing activity of the locus, a predicted membrane receptor that mediates cell-to-cell communication, named TRA-2A. The tra-2 gene was characterized from a close relative, C. briggsae. The Ch-tra-2 gene expresses only a 4.7-kb mRNA and alternatively spliced variants, which encode TRA-2A homologues. The Ch-TRA-2A and Ce-TRA-2A sequences are highly diverged, sharing only 43% identity, although their hydropathy profiles remain remarkably similar. Three potential regulatory sites of Ce-tra-2 activity were previously identified by analyzing tra-2(gf), tra-2(eg), and tra-2(mx) mutations. Two of these sites, the EG site and MX region, are conserved in Ch-tra-2. By contrast, the two direct repeat elements in the Ce-tra-2 3′ untranslated region, which are disrupted in tra-2(gf) mutants, are absent. Injection of Ch-tra-2 antisense RNA into C. briggsae mimics the Ce-tra-2 loss-of-function phenotype. Thus, antisense RNA permits studies of gene activity in nematodes that lack extensive genetics.

SEX determination is a fundamental part of the development of most animals. The importance of this process has led to the elucidation of the primary sex-determining signal for a number of model organisms. These studies reveal a conspicuous lack of evolutionary conservation in the mechanisms that control the decision to develop as a male or female. In mammals, male sex determination is induced by a dominant gene on the Y chromosome that encodes the testis-determining factor (TDF) (Sinclair et al. 1990). In mouse, the Sry gene, which encodes TDF, has been shown to be sufficient to direct male development in a chromosomally female mouse (Koopman et al. 1991). In contrast, both Drosophila and Caenorhabditis elegans share a primary sex-determining signal that relies on the ability to count chromosomes. It is the ratio of the number of X chromosomes to sets of autosomes, the X:A ratio, that dictates male or female development, not the absolute number of X chromosomes or the presence of a Y chromosome (reviewed in Hodgkin 1990; Parkhurst and Meneely 1994). No other similarities have been detected between the Drosophila and C. elegans sex determination pathways: the genes that transduce and respond to the X:A ratio appear to be completely unrelated in both sequence and function (for reviews, see Hodgkin 1990; Parkhurst and Meneely 1994).

The nematode C. elegans provides an excellent model for dissecting the mechanism of sex determination at both the genetic and molecular level and for asking questions regarding the evolution of sex-determining mechanisms. C. elegans naturally exists as morphologically distinct XX hermaphrodites or XO males. The hermaphrodite soma is virtually indistinguishable from the soma in females of closely related, yet exclusively female/male species (BaIRD et al. 1994). However, the germ line of the hermaphrodite first produces a limited number of sperm before switching to oogenesis for the remaining life of the hermaphrodite, thereby permitting self-fertility. A number of regulatory genes have been identified that control somatic sex determination. Loss-of-function mutations in any one of these genes transforms an animal from the sex initially specified by the X:A ratio to the opposite sex (reviewed in Hodgkin 1987b; Villeneuve and Meyer 1990; Kuwabara and Kimble 1992).

All of the known major regulatory genes that control sex determination in C. elegans have now been cloned. The basic mechanism underlying the somatic sex determination pathway appears to involve signal transduction. The predicted secreted HER-1 protein (Hunter and Wood 1992; Perry et al. 1993) and the membrane-spanning TRA-2A protein (Kuwabara et al. 1992) are postulated to mediate cell-to-cell signaling. In the absence of HER-1 activity, a predicted intracellular domain of TRA-2A is postulated to promote XX hermaphrodite development by negatively regulating the activity of the FEM proteins by protein-protein interaction (Kuwabara and Kimble 1995). Repression of the FEM proteins permits the master regulator in the somatic pathway of sex determination, a zinc-finger protein encoded by the tra-1 gene, to function as a transcriptional regulator that promotes XX hermaphrodite development (Hodgkin 1987a; Zarkower and Hodgkin 1992; Hodgkin 1993). In XO males, HER-1 is predicted to
repress TRA-2A activity. In turn, the FEM proteins are postulated to inhibit TRA-1 to allow male development.

The tra-2 gene is a large and complex locus that plays a key role in the regulatory cascade that controls sexual cell fate decisions. Three tra-2 transcripts of 4.7, 1.9, and 1.8 kb have been identified (Oskema and Kimble 1991); the 4.7-kb tra-2 mRNA is predicted to encode the major feminizing activity of the locus, a membrane-spanning protein of 170 kDa, named TRA-2A (Kuwabara et al. 1992; Kuwabara and Kimble 1995). In XO animals, it is predicted that TRA-2A is inactivated by a repressive ligand, HER-1 (Peery et al. 1993). A potential site of direct interaction between TRA-2A and HER-1 has been identified by the analysis of tra-2(eg) (for enhanced gain-of-function) mutations (Kuwabara 1996). These mutations encode a tra-2 activity that is insensitive to negative regulation by her-1, because XO tra-2(eg) mutants are transformed from male to hermaphrodite (Hodgkin and Albertson 1995; Kuwabara 1996).

The tra-2 gene regulates not only somatic sex determination, but also germ-line sex determination (Hodgkin and Brenner 1977; Doniach 1986; Schedl and Kimble 1988). Two classes of tra-2 mutations have been identified that disrupt germline sex determination but have little effect on somatic sex determination. These mutations, named tra-2(gf) and tra-2(mx) (for mixed character), transform self-fertile hermaphrodites to self-sterile females (spermless hermaphrodites) (Doniach 1986; Schedl and Kimble 1988). The tra-2(gf) mutations disrupt a direct repeat element (DRE) found in the tra-2 3′ untranslated region (UTR); the DREs are implicated in mediating translational regulation of tra-2 activity (Goodwin et al. 1993). The tra-2(mx) mutations encode missense changes in a discrete region of the carboxy terminal tra-2 coding region and collectively delineate a potential binding site for a repressor of tra-2 germline activity (P. Kuwabara, P. Oskema, J. Kimble, unpublished data). The existence of these two mutational classes suggest that in the hermaphrodite germ line, tra-2 feminizing activity must be negatively regulated to allow the onset of spermatogenesis. It follows that the acquisition of tra-2 germline-specific negative regulatory controls may have played an important role during the evolution of hermaphrodite/male nematode strains from exclusively female/male nematodes.

Given the importance of tra-2 in controlling C. elegans sex determination and the novelty of the signal transduction pathway that regulates sexual cell fate decisions, it was of interest to compare the C. elegans tra-2 gene sequence to its homologue in a closely related hermaphrodite/male strain, C. briggsae. C. briggsae and C. elegans are estimated to have diverged 20–50 million years ago (Emons et al. 1979; Heschl and Baillie 1990; Lee et al. 1992), but they remain extremely similar in development, morphology, and behavior. This phylogenetic comparison should reveal whether the function and regulation of tra-2 activity has been conserved in two different hermaphrodite species. More specifically, it was hoped that this comparison would also shed light on the notion that the acquisition of controls regulating tra-2 germ-line activity led to the evolution of hermaphrodite/male species, such as C. elegans and C. briggsae, from exclusively female/male nematodes. Previously, it was reported that the Cb-tra-2 gene had been cloned using a technique based on gene linkage (Kuwabara and Shah 1994); however, detailed sequence analysis was not presented. The complete Cb-tra-2 cDNA sequence and the characterization of the Cb-tra-2 gene and its products are described in this paper.

MATERIALS AND METHODS

Worm culture: Basic methods for culturing C. elegans were described by Brenner (1974). Genetic nomenclature follows the guidelines of Hodgkin (1995). Wild-type C. elegans refers to N2 var. Bristol and C. briggsae refers to the Gujarat isolate (Fodor et al. 1983). Mutations used in this paper are described in Hodgkin (1988). Worms were grown on petri dishes or in liquid as described (Sulston and Hodgkin 1988). To obtain synchronized populations, eggs were isolated by hypochlorite treatment of gravid hermaphrodites and growth arrested as L1 larvae by starvation in liquid medium. L1 larvae were released from growth arrest by the addition of Escherichia coli.

General manipulation of nucleic acids: General methods for the manipulation of nucleic acids are described in Sambrook (1989). Plasmid DNA was isolated either by Wizard (Promega) or QiaGen (Qiagen) column chromatography. DNA fragments were separated by electrophoresis on 1.0% agarose gels.

Nematode DNA was prepared as described by Emons and Yesner (1984). Nematode RNA was isolated from worms by the guanidinium thiocyanate protocol of Chirgwin et al. (1979). Poly(A)+ RNA was enriched using a single round of selection on oligo(dT) cellulose (Pharmacia) or QiaGen (Qiagen) column chromatography. Sequence-specific primers were used to complete the sequencing of first and second strands. All sequences obtained from PCR generated templates were verified by either sequencing an independently amplified Ch-tra-2 cDNA or by sequencing the corresponding region of genomic DNA. Partial Ch-tra-2 genomic sequence was also obtained by sequencing of lambda PK10, which was isolated from a C. briggsae genomic library (kindly provided by T. Snutch and D. Baillie) as described by Kuwabara and Shah (1994). cDNA clones corresponding to the 5' end of the Ch-tra-2 cDNA were synthesized by RT-PCR (Frohman et al. 1988).
Ob-tra-2 cDNA was primed using the oligonucleotide PK124CB and 2 μg of C. briggsae mixed-stage poly (A)+ mRNA as template. PCR was performed using this Ob-tra-2 cDNA as template and either SL1 or SL2 and PK124CB as primers. The full-length Ob-tra-2 cDNA sequence can be obtained from Gen-Bank (accession number U59879).

The clone pPK255 is a derivative of pPK223 that was constructed by ligating the insert from clone pPK223 into the NdeI-EcoRv site of the C. elegans heat shock promoter vector pPD49.83. It was necessary, however, to replace a 5-bp sequence of pPK255 with sequence from a RT-PCR-generated Ob-tra-2 cDNA because the pPK223 sequence carries a frame shift (see text, for details).

Computer analysis: DNA sequences were compiled and analyzed using the GCG Sequence Analysis Software Package (Devereux et al., 1984), available through the SEQNET facility at Daresbury, or the MacVector and AssemblyLIGN sequence analysis packages (Kodak). Sequence similarities were detected by BLAST analysis (Altschul et al. 1990; Gish and States 1993).

Northern blot analysis: Three micrograms of poly (A)+ mRNA per lane was electrophoresed through a 1.5% agarose gel containing formaldehyde and blotted to HybondN nylon membrane (Amersham). Northern blots were hybridized in 50% formamide, 5X Denhardt's, 5X SSC, 0.1% SDS, 100 pg/ml cDNA because the pPK223 sequence carries a frame shift (see text, for details).

Identification and characterization of Ob-tra-2 cDNA clones: Four positively hybridizing lambda phage clones were identified by screening a C. briggsae mixed-stage cDNA library with a Ob-tra-2 genomic DNA probe (Kuwabara and Shah 1994). All four clones shared a similar EcoRl restriction pattern (data not shown); therefore, only the lambda clone with the largest insert, 4.7 kb, lambda PK11, was analyzed in detail.

The original Ob-tra-2 cDNA clone, lambda PK11, contains the complete Ob-tra-2 coding sequence (see below), but does not contain a trans-splice leader sequence. To determine whether the Ob-tra-2 mRNA is trans-spliced, RT-PCR was used to clone additional Ob-tra-2 cDNAs that correspond specifically to the 5' end (see MATERIALS AND METHODS). Seven clones were obtained and each carried the SL2 trans-splice leader sequence (Huang and Hirsh 1989). The presence of SL2 at the 5' end of the Ob-tra-2 mRNA suggests that tra-2 is probably the downstream component of a polycistronic operon, as is Ce-tra-2 (Kuwabara et al. 1992; Spieth et al. 1993).

Identification of Ob-tra-2 RNA: To identify Ob-tra-2 mRNAs, a Northern blot containing poly(A)+ RNA prepared from mixed-stage C. briggsae hermaphrodites was hybridized with a radiolabeled lambda PK11 cDNA probe. For comparison, poly(A)+ RNA prepared from adult C. elegans hermaphrodites was loaded in an adjacent lane and hybridized with a full-length Ce-tra-2 cDNA probe (Kuwabara et al. 1992). Figure 1 (lane 2) shows that the Ob-tra-2 probe detects a single mRNA of 4.7 kb, which is the same size as the lambda PK11 insert. This indicates that the lambda PK11 Ob-tra-2 cDNA sequence specifically corresponds to a 4.7-kb Ob-tra-2 mRNA and that there are no additional small mRNAs. By contrast, a 4.7-kb and an additional 1.8-kb tra-2 mRNA are detected in a C. elegans poly(A)+ mRNA preparation (Figure 1, lane 1).

Analysis of the Ob-tra-2 cDNA sequence: The lambda PK11 Ob-tra-2 cDNA sequence was analyzed for open reading frames (ORFs). ORF 1 initiates at the first methionine codon (nucleotide 25) and contains three amino acids. ORF 1 encoded by the 4.7-kb Ce-tra-2 cDNA also consists of three codons. Moreover, the spacing between ORF 1 and ORF 2 is the same in both C. elegans and C. briggsae. The conservation of ORF 1 and the proximity of ORF 1 to ORF 2 in C. elegans and C. briggsae indicates that ORF 1 might regulate ORF 2.
translation (for example, see HINEBUSCH 1993). In *C. elegans*, ORF 2 encodes the primary feminizing activity of the *tra-2* locus, a predicted 1475-amino acid membrane protein with a molecular weight of 170 kDa, named TRA-2A (KUWABARA and KIMBLE 1995). In *C. briggsae*, ORF 2 is predicted to encode the Ce-TRA-2A homologue, a 1497-amino acid peptide with extended sequence similarity to Ce-TRA-2A. However, the lambda PK11 cDNA clone carries a frame-shift that prematurely terminates ORF 2 and, as discussed below, other *C. briggsae* cDNAs do not carry this frame-shift.

**C. elegans** mRNA is alternatively spliced: BLAST analysis established that Ce-TRA-2A is the most similar in sequence to the predicted *C. elegans* ORF 2 (ALTSCHUL et al. 1990; GISH and STATES 1993) and that lambda PK11 carries 2 frame-shifts. To determine the origin of this frame-shift, partial *C. elegans* genomic sequence was compared to the lambda PK11 *C. briggsae* cDNA sequence to identify potential sites of alternative splicing. Intervening sequences three and six (IVS 3 and IVS 6) were found to contain alternative splice acceptor sites, which are not present in the comparable regions of *C. briggsae* genomic sequence (Figure 2). In addition, the relative positions of the first six splicing events are conserved between *C. briggsae* and *C. elegans* (data not shown).

Comparison of *C. briggsae* genomic sequence to the sequences of lambda PK11 and the seven RT-PCR-generated 5' end *C. briggsae* cDNA clones verified that alternative splicing occurs at IVS 3 and IVS 6. IVS 3 (Figure 2A) contains two potential splice acceptor sites separated by five nucleotides (nt) (IVS 5A, B). Selection of IVS 5A by lambda PK11 and two of the seven 5' end *C. briggsae* cDNA clones produces a frame-shift; selection of IVS 5B by the remaining five clones allows the expression of a single protein consisting of 1497 amino acids with a predicted molecular weight of 172 kDa, which will henceforth be referred to as *C. briggsae* TRA-2A (Figure 3A).

Alternative splicing also occurs at IVS 6, where there are two splice acceptor sites separated by nine nt (Figure 2A). IVS 6A was selected by lambda PK11 and two of the RT-PCR clones, whereas IVS 6B was used by the remaining four clones. Neither of the IVS 6 splice variants has a frame-shift, although usage of IVS 6A leads to an in-frame insertion of three amino acids. One clone was identified that contains a frame-shift because it failed entirely to excise IVS 6 (IVS 6C). As summarized in Figure 2B, this analysis suggests that alternative splicing has the potential to generate a combination of active *C. briggsae* TRA-2 products, which may also contain in-frame amino acid insertions or deletions, and inactive products that are prematurely truncated.

**Comparison of *C. elegans* and *C. briggsae* regulatory elements in the *tra-2* gene and its gene products:** Potential regulatory elements and functional domains have been identified by the molecular analysis of three kinds of *tra-2* regulatory mutations: *tra-2*(eg) (HODGKIN and ALBERTSON 1995; KUWABARA 1996), *tra-2*(gf) (GOODWIN et al. 1993), and *tra-2*(mx) (P. KUWABARA, P. OKKEMA and J. KIMBLE, unpublished results) mutations. To determine whether the sequences associated with these mutations are conserved, regions of *C. briggsae* sequence that correspond to those affected by the *C. elegans* mutations were examined.

The *tra-2*(eg) mutations specifically affect XO animals, transforming them from males into hermaphrodites (HODGKIN and ALBERTSON 1995; KUWABARA 1996). It was proposed that the *tra-2*(egR177K) mutation identifies a potential site of direct interaction...
Evolution of \( \text{tra-2} \) Regulation

A

<table>
<thead>
<tr>
<th>IVS 3A</th>
<th>frame-shift</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAA gtagttatgaagggcttctttattagaaatctgttattcaaaatcggacagt</td>
<td>TTC AGC GTC TC 439</td>
</tr>
<tr>
<td>IVS 3B</td>
<td>in-frame</td>
</tr>
<tr>
<td>CAA gtagttatgaagggcttctttattagaaatctgttattcaaaatcggacagt</td>
<td>CGT GTC 434</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>IVS 6A</th>
<th>in-frame</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAG gtatatatcaattggctgattcgaaaatcacagaacagttaagagattgaag</td>
<td>TTT CAA CTC 878</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>IVS 6B</th>
<th>in-frame</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAG gtatatatcaattggctgattcgaaaatcacagaacagttaagagattgaag</td>
<td>TTT CAA CTC 878</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>IVS 6C</th>
<th>frame-shift (NO SPLICE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GTCACAG CAGATTAA GAGATTGAA GTTTCAA CTC 930</td>
<td></td>
</tr>
</tbody>
</table>

B

<table>
<thead>
<tr>
<th>SPlice Site Selection</th>
<th>IVS 3</th>
<th>IVS 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clone</td>
<td></td>
<td></td>
</tr>
<tr>
<td>lambda PK11</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>pPK158-1</td>
<td>A</td>
<td>C</td>
</tr>
<tr>
<td>pPK158-3</td>
<td>B</td>
<td>A</td>
</tr>
<tr>
<td>pPK158-5</td>
<td>B</td>
<td>B</td>
</tr>
<tr>
<td>pPK158-6</td>
<td>B</td>
<td>B</td>
</tr>
<tr>
<td>pPK158-8</td>
<td>B</td>
<td>A</td>
</tr>
<tr>
<td>pPK158-9</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>pPK158-10</td>
<td>B</td>
<td>A</td>
</tr>
</tbody>
</table>

Figure 2.—Comparison of genomic and cDNA sequences reveal that \( \text{Cb-tra-2} \) cDNAs are alternatively spliced. (A) Differences between the sequences of individual \( \text{Cb-tra-2} \) cDNAs are due to alternative splicing occurring at IVS 3 and IVS 6. Lowercase, intron sequences; uppercase, exon sequences; boldface, splice donor (gt) and acceptor sequences (ag). Coding sequence insertions are underlined. Numbering refers to the sequence of \( \text{Cb-tra-2} \) cDNA, lambda PK11. (B) Summary of IVS 3 and IVS 6 splice acceptor usage based on the sequence of eight \( \text{Cb-tra-2} \) cDNA clones.

between Ce-TRA-2A and HER-1 (Kuwabara 1996). Sequence comparisons indicate that the Ce-TRA-2A R177 residue is conserved in Cb-TRA-2A as R178. Furthermore, the Cb-TRA-2A R178 residue is embedded within the region most highly conserved between the two proteins (63% identity) (Figure 3A). These data support the idea that Cb-TRA-2A is regulated by a Cb-HER-1 homologue.

The \( \text{Ce-tra-2}(gf) \) and \( \text{Ce-tra-2}(mx) \) mutations primarily affect germline sex determination because they transform self-fertile hermaphrodites into females but have little effect on XO males (Doniach 1986; Schedl and Kimble 1988). It is proposed that these two sets of mutations identify two different control regions that negatively regulate \( \text{tra-2} \) germline activity to allow the onset of hermaphrodite spermatogenesis (Doniach 1986; Schedl and Kimble 1988). The \( \text{Ce-tra-2}(gf) \) mutations are caused by disruptions to a 28-nt DRE found in the \( \text{tra-2} \) 3' UTR (Figure 4, top), which is likely to have a role in translational regulation of \( \text{Ce-tra-2} \) (Kuwabara et al. 1992; Goodwin et al. 1993). No repeat elements larger than eight nt or sequences with extensive similarity to the \( \text{Ce-tra-2} \) DRE are detected in the 3' UTR of \( \text{Cb-tra-2} \) (Figure 4, bottom). By contrast, five \( \text{Ce-tra-2}(mx) \) mutations encode missense changes that affect a discrete cluster of amino acids in the carboxy terminal region of \( \text{Ce-tra-2A} \), named the MX site (P. Kuwabara, P. Okkema and J. Kimble, unpublished data); four of these residues are conserved in Cb-TRA-2A (Figure 3A). Therefore, only one of two potential germline controls of \( \text{tra-2} \) activity appears to be conserved in \( \text{C. briggsae} \).

\( \text{Cb-tra-2} \) antisense RNA phenocopies the \( \text{Ce-tra-2} \) loss-of-function phenotype: The syntenic conservation between the \( \text{Cb-tra-2} \) and \( \text{Ce-tra-2} \) gene regions combined with sequence analysis of the \( \text{Cb-tra-2} \) gene clearly indicate that the \( \text{C. briggsae} \) orthologue of \( \text{Ce-tra-2A} \) has been identified. However, \( \text{Cb-tra-2} \) mutants have not been identified, so it was not known whether the \( \text{Cb-tra-2} \) orthologue also functions to promote hermaphrodite development. In \( \text{C. elegans} \), \( \text{tra-2}(lf) \) mutations masculinize XX hermaphrodites. To provide evidence that \( \text{Cb-tra-2} \) functions in sexual cell fate decisions, experiments were performed to phenocopy the \( \text{Ce-tra-2}(lf) \) mutation by disrupting wild-type \( \text{Cb-tra-2} \) activity. This was achieved by injecting 1-kb fragments of \( \text{in vitro} \) synthesized antisense RNA corresponding to either 5' or 3' regions of \( \text{Cb-tra-2} \) mRNA into the germline syncytium.
of adult hermaphrodites (Guo and Kemphues 1995) and examining their progeny by Nomarski DIC optics for signs of somatic masculinization. As shown in Table 1, injection of 5' antisense RNA led to the formation of a single-lobed or intersex somatic gonad, absence of vulval induction, and perhaps most importantly, to the sexual transformation of the tail from hermaphrodite to male. Figure 5 shows such a masculinized animal, which has developed a male tail with fan, ray, and spicules. Generally, extensive tail masculinization was present only in animals with masculinized somatic gonads that failed to induce a vulva. Injection of 3' antisense RNA led to masculinization of the same tissues affected by 5' antisense, however, the extent of masculinization was comparatively less severe. Specifically, male tail structures such as a fan, ray, and spicules were not observed in the progeny of animals injected with 3' antisense RNA, although the

![Figure 3.—Comparisons of Ce-TRA-2A and Cb-TRA-2A amino acid sequences. To simplify the alignment and comparison of the Ce-TRA-2A and Cb-TRA-2A amino acid sequences using the GCG program Bestfit (Devereux et al. 1984) showed 43% identity and 64% similarity. Upper, Ce-TRA-2A sequence; lower, Cb-TRA-2A sequence. * +, Shadowed region, region of highest sequence conservation by BLAST analysis: 63% identity over 150 amino acids. Box, MX region. The positions of individual Ce-tra-2(mex) mutations will be reported elsewhere (P. Kawai, P. Okkema, and J. Kimble, unpublished data). A, site of lambda PKll frame-shift; circles, M, ORF 2 initiation site in Cb-tra-2 cDNA clone, lambda PKll. Note that ORF 3 may encode a TRA-2A variant with feminizing activity, because the sequences at its amino-terminus could substitute for the original signal sequence, which was truncated by the frame-shift. (B) Hydrophy profiles of Ce-TRA-2A and Cb-TRA-2B plotted according to the method of Kyte and Doolittle (1982).]
tail spike was sometimes found to be truncated (snub). Potential masculinization of the germ line was not assessed. The finding that antisense RNA corresponding to two different regions of \( \text{Cb-tra-2} \) disrupts sexual fate decisions in multiple tissues strongly argues that \( \text{Cb-tra-2} \) plays a major role in controlling sexual cell fates, as \( \text{Ce-tra-2} \) does in \( \text{C. elegans} \).

**DISCUSSION**

The \( \text{tra-2} \) gene plays a central role in regulating sexual cell fate decisions in \( \text{C. elegans} \); \( \text{Ce-TRA-2A} \), the major feminizing product of the \( \text{tra-2} \) locus, is a predicted novel membrane protein that mediates signal transduction (Kuwabara and Kimble 1995). One approach to learn more about the function and regulation of this novel protein was to characterize a homologue from \( \text{C. briggsae} \), because other proteins with significant sequence similarity to \( \text{TRA-2A} \) had not been identified in current databases. A composite \( \text{Cb-tra-2} \) cDNA sequence has now been obtained, and the extent of possible functional and regulatory conservation between the two genes and their products has been examined. The following discussion summarizes the results of this comparison and their implications for the control and evolution of nematode sex determination.

Preliminary analysis indicates that the overall gene organization of the \( \text{Cb-tra-2} \) and \( \text{Ce-tra-2} \) genes has been maintained. First, the genomic regions containing the \( \text{Cb-tra-2} \) and \( \text{Ce-tra-2} \) genes are syntenic (Kuwabara and Shah 1994). Second, both genes are trans-spliced to SL2, suggesting that they are likely to be downstream components of a polycistronic unit (Kuwabara et al. 1992; Speth et al. 1993). Third, both genes contain a three-codon ORF immediately upstream of the \( \text{TRA-2A} \) initiation codon. Finally, a limited sampling of six intron/exon boundaries indicates that their relative spacings are conserved.

A different picture emerges when the \( \text{Cb-tra-2} \) and \( \text{Ce-tra-2} \) gene products are compared. First, only a 4.7-kb mRNA is detected in RNA preparations from mixed-stage \( \text{C. briggsae} \) hermaphrodites, whereas a 4.7- and 1.8-kb \( \text{tra-2} \) mRNA are present in adult \( \text{C. elegans} \) hermaphrodites (Okkema and Kimble 1991). The role of the 1.8-kb \( \text{tra-2} \) mRNA may be partially dispensable for somatic sex determination (Kuwabara and Kimble 1995). The detection of only a single \( \text{Cb-tra-2} \) mRNA of 4.7 kb supports the previous report that the 4.7-kb \( \text{Ce-tra-2} \) mRNA provides the major feminizing activity of the \( \text{tra-2} \) locus (Kuwabara and Kimble 1995).

Second, several examples of alternatively spliced 4.7-kb \( \text{Cb-tra-2} \) cDNAs have been identified, but no alternatively spliced variants of the \( \text{Ce-tra-2} \) 4.7-kb mRNA have been detected (Kuwabara et al. 1992). The significance of alternative splicing in the \( \text{C. briggsae} \) sex determination pathway is open to speculation. In Drosophila, sex determination is controlled by a cascade of sex-specific alternative splicing events (reviewed in Parkhurst and Meneely 1994). These events lead to the formation of inactive or sex-specific products that regulate cell fate. In \( \text{C. briggsae} \), the lambda PK11 cDNA represents an alternatively spliced variant that encodes a severely truncated and probably inactive \( \text{Cb-TRA-2A} \). Therefore, alternative splicing may provide a mechanism for regulating \( \text{Cb-tra-2} \) activity. This is not to say that alternative
splicing of Cb-tra-2 is regulated by the same mechanism as Drosophila. In Drosophila, there is an initial sequence bias in splice site preference (reviewed in Parkhurst and Meneely 1994), which is not present in the Cb-tra-2 sequence; Cb-tra-2 splice acceptor choice might instead be random. Regardless of the mechanism underlying the control of alternative splicing in Cb-tra-2, alternative splicing has the potential to generate multiple Cb-TRA-2A proteins with in-frame insertions or deletions, which contribute to the sequence divergence between Cb-TRA-2A and Ce-TRA-2A. Conceivably, the net effect of alternative splicing is to downregulate tra-2 activity by producing a proportion of inactive products. In C. elegans, a similar downregulation of tra-2 activity might instead be achieved by translational control mediated through the DREs (Goodwin et al. 1993), which are not present in the Cb-tra-2 sequence (this paper).

Third, the comparison between Ce-TRA-2A and a composite Cb-TRA-2A sequence reveals that the two proteins are only 43% identical and 64% similar: the tra-2 gene is the most rapidly evolving of all genes so far compared in C. elegans and C. briggsae. The region most highly conserved between the two proteins is 65% identical over 150 amino acids; the remaining regions of similarity are essentially dispersed throughout the proteins. For comparison, two genes involved in the control of cell death ced-3 (Yuan et al. 1993) and ced-9 (Hengartner and Horvitz 1994) are 85% and 66% identical, when compared to their respective C. briggsae homologues. Despite the high degree of sequence divergence, the hydropathy profiles of Cb-TRA-2A and Ce-TRA-2A remain remarkably similar. It was previously noted that Ce-TRA-2A shows marginal sequence similarity and shares a similar hydropathy profile with the Drosophila Ptc protein (Kuwabara et al. 1992). Recently, it has been reported that the Drosophila Ptc and mouse Ptc protein are only 34% identical and 44% similar, yet they too have very similar hydropathy profiles (Goodrich et al. 1996). These two examples provide indications that proteins with similar topologies, as inferred from hydropathy predictions, might retain similar function and regulation, despite a lack of primary sequence conservation. The high level of divergence between Cb-TRA-2A and Ce-TRA-2A indicates that it will probably be difficult to identify TRA-2A-related proteins or other components of the signal transduction pathway, which may regulate processes other than sex determination, by primary sequence alone in evolutionarily more distant organisms.

To show that Cb-tra-2 plays a major role in controlling C. briggsae sex determination, antisense Cb-tra-2 RNA was used to disrupt wild-type Cb-tra-2 activity. This treatment led to the masculinization of most somatic tissues, which is a phenocopy of the Ce-tra-2(+/f) phenotype. In C. elegans, sexual cell fate decisions require tra-2 activity throughout larval development (Klass et al. 1976). These results indicate that antisense RNA provides a useful tool for phenocopying loss-of-function mutations, not only in the embryo (Guo and Kempfues 1995), but also at different stages throughout development. Furthermore, it is shown that the role of antisense RNA in eliciting mutant phenocopies is not dependent on the presence of a methyl-7-guanosine cap. More generally, these results indicate that antisense RNA may provide a tool for analyzing gene activity in

### TABLE 1

<table>
<thead>
<tr>
<th>Total progeny</th>
<th>Wild type</th>
<th>Dead eggs</th>
<th>Gonad and vulva</th>
<th>Masculinized phenotypes*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Snub Fan Ray Spicules</td>
</tr>
<tr>
<td>62</td>
<td>35</td>
<td>2</td>
<td>25</td>
<td>6 4 3 2</td>
</tr>
<tr>
<td>44</td>
<td>20</td>
<td>0</td>
<td>24</td>
<td>6 4 3 1</td>
</tr>
<tr>
<td>66</td>
<td>30</td>
<td>0</td>
<td>36</td>
<td>10 0 0 0</td>
</tr>
<tr>
<td>22</td>
<td>14</td>
<td>1</td>
<td>7</td>
<td>2 0 0 0</td>
</tr>
<tr>
<td>32</td>
<td>21</td>
<td>1</td>
<td>10</td>
<td>2 0 0 0</td>
</tr>
<tr>
<td>30</td>
<td>18</td>
<td>0</td>
<td>12</td>
<td>2 0 0 0</td>
</tr>
<tr>
<td>11</td>
<td>5</td>
<td>1</td>
<td>5</td>
<td>2 0 0 0</td>
</tr>
<tr>
<td>28</td>
<td>11</td>
<td>1</td>
<td>16</td>
<td>6 0 0 0</td>
</tr>
</tbody>
</table>

*First three lines represent the progeny from three individual C. briggsae hermaphrodites injected with 5' Cb-tra-2 antisense RNA. Last five lines represent the progeny from five individual C. briggsae hermaphrodites injected with 3' Cb-tra-2 antisense RNA. Uninjected C. briggsae hermaphrodites, n = 100, were never observed to produce masculinized progeny such as those described in Table 1.

*Total progeny refers only to the number of progeny produced by mothers up to 24–36 hour after antisense RNA injection and is obtained by adding columns 2–4. The number of animals with masculinized tail structures form a subset of column 4.

* Gonad refers to somatic gonad, not germ-line. Masculinized vulva refers to the absence of vulval induction or to the formation of a nonfunctional vulval structure.
other nematode species that do not have extensive genetics.

Given that *Cb-tra-2* plays a role in controlling sexual cell fate, then the sequences conserved between *Cb-tra-2* and *Ce-tra-2* are likely to shed light on those regions required for proper *tra-2* function or regulation. The region that is most highly conserved contains a potential HER-1 interaction site, which has been named the EG site in *C. elegans* (Kuwabara 1996). Conservation of this region suggests that *Cb-tra-2* activity is likely to be regulated by a yet to be identified *Cb-hpr-1* activity. A potential FEM interaction site is also predicted to reside within a carboxy terminal region of *Ce-TRA-2A* (Kuwabara and Kimble 1995). However, limited sequence conservation in this region presently precludes the identification of this site.

The *tra-2* gene also plays an important role in regulating sex determination in the germ line. It has been proposed that the *Ce-tra-2(gf)* and *Ce-tra-2(mx)* mutations identify two control regions that regulate *Ce-tra-2* germ-line activity to allow the onset of hermaphrodite spermatogenesis (Doniach 1986; Schedl and Kimble 1988). Obviously, the ability of an otherwise female animal to produce sperm was one of the key events during the evolution of hermaphrodite/male from exclusively female/male species. There is evidence that hermaphroditism may have evolved twice in the genus, because *C. briggsae* appears to be more closely related to a male/female species, *C. vulgařis*, than to *C. elegans* (Baird et al. 1992). Therefore, it was of considerable interest to determine whether the sites identified by *Ce-tra-2* germ-line feminizing mutations are both conserved in *C. briggsae*. This study showed that only the TRA-2A MX region is conserved (P. Kuwabara, P. Okkema and J. Kimble, unpublished data); the DREs in the *Ce-tra-2* 3′ UTR affected by the *tra-2(gf)* mutations (Goodwin et al. 1993) are absent in *Cb-tra-2*. It should be noted that the *tra-2(gf)* mutant phenotype is different from the *tra-2(mx)* mutant phenotype, because the strongest *tra-2(gf)* mutations feminize not only the germline, but also partially feminize the X0 soma (Doniach 1986; Schedl and Kimble 1988). Regulation of *Ce-tra-2* activity through the 3′ UTR may be a more recent acquisition that modulates *tra-2* activity in both germ line and soma.

The comparison of the *Cb-tra-2* and *Ce-tra-2* sequences indicate that the sex-determining genes are evolving more rapidly than any other nematode regulatory gene so far analyzed. This is further substantiated by the finding that the sequence of another sex-determining protein, TRA-1, is also rapidly diverging in *C. briggsae* (De Bono and Hodgkin 1996, this issue). Although there is a general lack of conservation in the mechanisms that control the decision to develop as a male or female, it is still surprising to find that sex-determining genes of two closely related hermaphroditic nematode species are diverging so quickly. The high rate of divergence could indicate that the rapidly evolving sequences have no functional importance or that change has been more actively selected. However, the changes in primary sequence cannot be entirely random, because the overall similarity in hydropathy profiles for the two TRA-2A proteins suggests that they have conserved topologies. This conservation in TRA-2A topology supports the prediction that the EG site, which may mediate cell-to-cell communication, is extracellular and the MX region is intracellular (Kuwabara et al. 1992). Potentially, the ability of the laboratory *C. briggsae* strain to survive at higher temperatures than *C. elegans* (Fo-
dor et al. 1983) also partially contributes to the sequence divergence between Cb-TRA-2A and Ce-TRA-2A; Cb-TRA-2A may be more thermostable than Ce-TRA-2A.

Interspecies comparison is a valuable tool for identifying sequences that have been conserved because of their role in regulating gene activity. This study identifies not only those sequences that may be crucial for tra-2 regulation and activity, but also suggests that the MX region may have played a key role in the evolution of hemphridote/male from female/male species.

The author thanks Mario de Bono and Jonathan Hodgkin for sharing results before publication, discussion, and critical review of the manuscript. Tim Schiedl and Allan Jones are thanked for discussion and experimental advice. I am also grateful to Annette Lenton and Brian Pashley for help with illustrations.

LITERATURE CITED


Perry, M. D., W. Li, C. Trent, B. Robertson, A. Fire et al., 1993 Molecular characterization of the her-1 gene suggests a direct role in cell signaling during Caenorhabditis elegans sex determination. Genes Dev. 7: 216–228.

Schiedl, T., and J. Kimble, 1988 fog-2, a germ-line-specific sex deter-
Evolution of *tra-2* Regulation


Communicating editor: R. K. Herman