J Mol Evol (2002) 54:267–282
DOI: 10.1007/s0023901-0008-y

JOURNAL OF MOLECULAR EVOLUTION



# Evolution of the PP2C Family in *Caenorhabditis:* Rapid Divergence of the Sex-Determining Protein FEM-2

**Paul Stothard, Dave Hansen,* Dave Pilgrim**

Department of Biological Sciences, University of Alberta, Edmonton, Alberta T6G 2E9, Canada

Received: 11 April 2001 / Accepted: 6 August 2001

**Abstract.** To investigate the causes and functional significance of rapid sex-determining protein evolution we compared three *Caenorhabditis elegans* genes encoding members of the protein phosphatase 2C (PP2C) family with their orthologs from another *Caenorhabditis* species (strain CB5161). One of the genes encodes FEM-2, a sex-determining protein, while the others have no known sex-determining role. FEM-2's PP2C domain was found to be more diverged than the other PP2C domains, supporting the notion that sex-determining proteins are subjected to selective pressures that allow for or cause rapid divergence. Comparison of the positions of amino acid substitutions in FEM-2 with a solved three-dimensional structure suggests that the catalytic face of the protein is highly conserved among *C. elegans,* CB5161, and another closely related species *C. briggsae.* However, the non-conserved regions of FEM-2 cannot be said to lack functional importance, since *fem-2* transgenes from the other species were unable to rescue the germ-line defect caused by a *C. elegans fem-2* mutation. To test whether *fem-2* functions as a sex-determining gene in the other *Caenorhabditis* species we used RNA-mediated interference (RNAi). *fem-2* (*RNAi*) in *C. elegans* and *C. briggsae* caused germ-line feminization, but had no noticeable effect in CB5161. Thus the function of *fem-2* in CB5161 remains uncertain.

**Current address:* Department of Genetics, Washington University School of Medicine, St. Louis, MO 63110, USA
*Correspondence to:* D. Pilgrim; *email:* dave.pilgrim@ualberta.ca

**Key words:** *fem-2* gene — Sex determination — Protein phosphatase — PP2C — *C. briggsae* — CB5161 — Rapid evolution

## Introduction

Sexual dimorphism and the mechanisms used to achieve it have been studied extensively in the nematode *Caenorhabditis elegans,* which develops either as a self-fertile hermaphrodite or a male (for reviews see Cline and Meyer 1996; Hansen and Pilgrim 1999). *C. elegans* hermaphrodites are similar to females of related nematode species, and can be thought of as somatic females that are able to transiently produce sperm in their germ line. There are a number of obvious differences between the two sexes: males are smaller than hermaphrodites; males have a tail that is fan-like and specialized for mating while hermaphrodites have a tail that lacks a fan; hermaphrodites have a vulva through which sperm can enter and fertilized eggs exit while males lack a vulva; and hermaphrodites have a two-armed gonad while males have a single-armed gonad. Additional sexual dimorphisms are present in the nervous system, musculature, and intestine. In fact, almost every tissue in *C. elegans* shows some form of sexual specialization (Sulston and Horvitz 1977). The primary sex-determining signal is the ratio of X chromosomes to autosomes (Madl and Herman 1979). Animals with a single X chromosome (XO) normally develop as males, while those with two X chromosomes (XX) normally develop as hermaphrodites. Genes that interpret and transmit the pri-

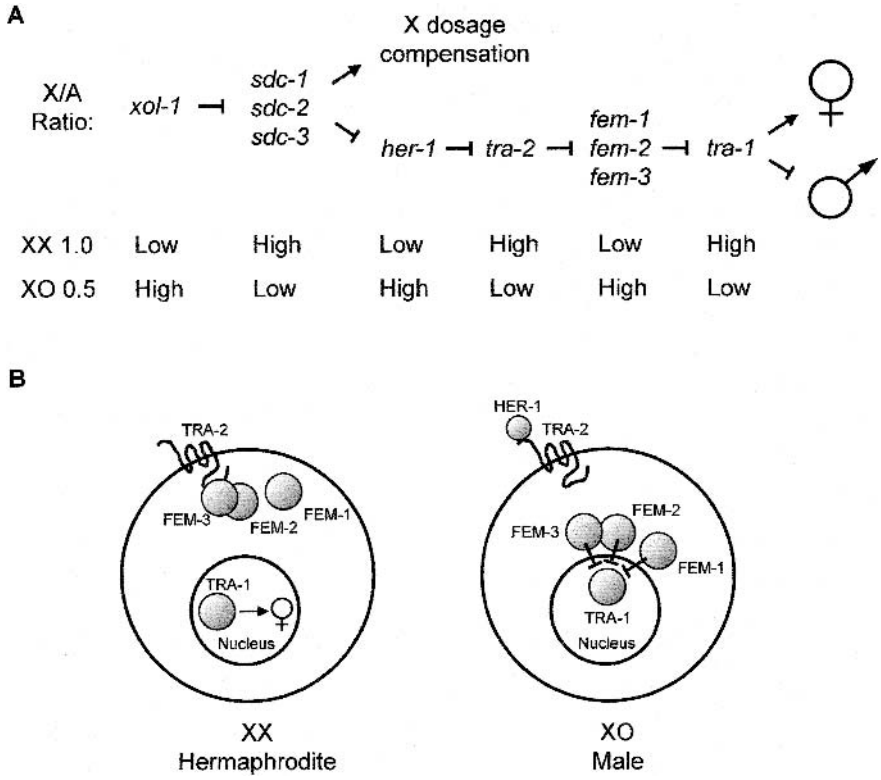


**Fig. 1.** **A:** The genetic pathway that regulates somatic sex determination in *C. elegans.* Arrows represent positive interactions, while barred lines indicate negative interactions. The ratio of the number of X chromosomes to the number of autosome sets (X/A ratio) is the initial signal. It establishes the sexual fate and dosage compensation modes via the *xol-1* and *sdc* genes, such that a ratio of 1.0 leads to female (hermaphrodite) development, while a ratio of 0.5 leads to male development. In XX animals *tra-2* negatively regulates the *fem* genes, allowing *tra-1* to promote female development. In XO animals *her-1* inhibits *tra-2,* allowing the *fems* to negatively regulate *tra-1.* In the absence of *tra-1* activity the male fate is established. **B:** The molecular model of somatic sex determination. In XX animals the TRA-2 transmembrane receptor binds to FEM-3 and interferes with the ability of the FEM proteins to inhibit the TRA-1 transcription factor. TRA-1 is then free to activate female-promoting genes and repress male-promoting genes. In XO animals the HER-1 secreted ligand is expressed, and it binds and inactivates TRA-2. This prevents TRA-2 from sequestering the FEMS, which are free to inactivate TRA-1. More detailed descriptions of these models and the supporting evidence is available in reviews by Cline and Meyer (1996) and Hansen and Pilgrim (1999).

mary signal have been isolated as mutations that transform the entire animal from one sex into the other. Genetic epistasis experiments have placed these genes into a regulatory hierarchy, in which the activity of each gene is controlled by the state of one or more upstream genes (Fig. 1). In somatic cells the most downstream gene is *tra-1;* if *tra-1* is active the soma is female, and if *tra-1* is inactive the soma is male. In the germ line the sex determination pathway is very similar, but there are additional genes involved that have no somatic role.

Comparisons of sex-determining genes between closely related species suggests that they evolve more rapidly than other genes. The orthologs of four *C. elegans* sex-determining genes have been isolated from the hermaphroditic species *C. briggsae* (de Bono and Hodgkin 1996; Kuwabara 1996; Hansen and Pilgrim 1998; Streit et at. 1999). In each case the predicted proteins have diverged more than the average level of divergence observed for 11 non-sex-determining proteins isolated from the same species (de Bono and Hodgkin 1996). Similar findings have been obtained from comparisons between mammalian species, and between *Drosophila* species (Whitfield et al. 1993; Tucker and Lundrigan 1993; O'Neil and Belote 1992; McAllister and McVean 2000). Often the poorly conserved regions of a protein are thought to lack functional importance, but in the case of sex-determining proteins it has been proposed that positive selection could be causing functionally important regions to diverge rapidly (Whitfield et al. 1993; Walthour and Schaeffer 1994). Positive selection has been demonstrated in the case of proteins that mediate sperm to egg binding in marine environments, where rapid divergence could facilitate reproductive isolation (Metz and Palumbi 1996; Swanson and Vacquier 1998). In the absence of positive selection proteins are said to evolve neutrally, with divergence occurring because it does not interfere with function. A recent study of the *transformer* sex-determining gene in *Drosophila* suggests that although it is evolving at a rapid pace, the changes are neutral in function (McAllister and McVean

2000). Why would one neutrally evolving protein diverge faster than another? One explanation is that the more quickly diverging protein contains more residues that are not important for function. Another is that the activity of the quickly diverging protein can vary more without having a detrimental effect on the fitness of the organism.

The sex-determining genes and pathways in more distant species appear to be mostly unrelated. Comparisons of the proteins involved in *C. elegans* and *Drosophila* sex determination have revealed almost no significant similarity, with each species relying on distinct biochemical mechanisms to propagate the sexual fate decision (for review see Cline and Meyer 1996). The somatic pathway in *C. elegans* involves negative regulatory interactions at the protein level, while the *Drosophila* somatic pathway involves a series of positive interactions at the level of RNA splicing. In the worm the somatic sex-determining genes all participate in the germ-line fate decision, while in the fly most somatic sex-determining genes have no germ-line role. To date the only molecular similarity involves a target of the *C. elegans tra-1* gene called *mab-3,* and the terminal sex-determining gene in *Drosophila, doublesex* (Raymond et al. 1998). From the limited information available, the mouse pathway lacks similarity to the worm or fly pathways, apart from a homolog of *mab-3/doublesex* that is required for testis differentiation (Raymond et al. 2000). It has been proposed that the most downstream components of the pathways are the most ancient, and that the upstream components have been assembled independently (Wilkins 1995).

To gain further insight into the evolution of sex-determining proteins and pathways we chose to study the evolution of the *C. elegans* sex-determining protein FEM-2. FEM-2 is particularly suited for this comparison for a variety of reasons. It is a member of the protein phosphatase type 2C (PP2C) family of serine/threonine phosphatases (Pilgrim et al. 1995; Chin-Sang and Spence 1996; Hansen and Pilgrim 1998), of which there are other members in *C. elegans* with no known role in sex determination. These FEM-2 paralogs can be used as internal controls to determine the rate and pattern of sequence divergence that is typical for PP2C proteins. The three-dimensional structure is known for a human PP2C (Das et al. 1996), allowing sequence conservation and structure/function information to be integrated. Also, the study of gene and gene pathway evolution in *Caenorhabditis* has practical advantages, as genes can be silenced using RNA-mediated interference, and transgenes from one species can be tested for function in another.

FEM-2's phosphatase activity plays a role in both somatic and germ-line sex determination in *C. elegans* (for review see Hansen and Pilgrim 1999). Animals mutant for *fem-2* develop as females, regardless of their karyotypic sex (Kimble et al. 1984; Hodgkin 1986). In male animals the *fem-2* gene product is required for both male somatic development and sperm production. In hermaphrodites *fem-2* is required for the short period of sperm production that takes place before oogenesis. Genetic evidence and molecular data indicate that in somatic cells FEM-2 is a component of a signal transduction pathway, linking a transmembrane receptor (encoded by the *tra-2* gene; Kuwabara et al. 1992) to a transcription factor (encoded by the *tra-1* gene; Zarkower and Hodgkin 1992). The germ-line role of FEM-2 is different from its somatic role (Hodgkin 1986). In the soma FEM-2 acts with FEM-1 and FEM-3 to inactivate (directly or indirectly) TRA-1. In the germ line FEM-2 and the other FEM proteins do not act via TRA-1, instead they are at the terminal position in the sex-determining pathway, and their downstream targets are not known.

The *fem-2* gene has been previously isolated from *C. elegans* and *C. briggsae* (Pilgrim et at. 1995; Hansen and Pilgrim 1998). Here we report the cloning and analysis of *fem-2* from a closely related male/female nematode species. We refer to this species using its strain name, CB5161, and the gene prefix 'Cs' for *Caenorhabditis* species. The *C. elegans* Genome Sequencing Consortium revealed that there are two other PP2C proteins in *C. elegans* that contain a PP2C domain similar in size and sequence to FEM-2's PP2C domain. We were able to isolate both of these FEM-2 paralogs from CB5161, and for further comparison we isolated PP2C sequences from a vertebrate, the zebrafish *Danio rerio.* Using a variety of techniques we investigate whether FEM-2 is subjected to different selective pressures than other PP2C proteins, and whether functionally important parts of the protein differ between species. Finally, we use RNA-mediated interference to ask whether *fem-2* has a sex-determining role in species other than *C. elegans.*

## Materials and Methods

*Nematode strains.* The following wild-type nematode strains were used in this study: *C. elegans* strain N2, *C. briggsae* strain AF16, and *Caenorhabditis* sp. strain CB5161. The Caenorhabditis Genetics Center and others originally referred to CB5161 as *C. remanei,* but it is now known to represent a distinct species (Thomas and Wilson 1991; Fitch et al. 1995).

*Oligonucleotides.*

| | |
|---|---|
| AP | 5′-GGCCACGCGTCGACTAGTACTTTTTTTTTTTTTTTT-3′ |
| BJM1 | 5′-G(GC)IGTIT(AT)(TC)GA(TC)GGICA(TC)G(GC)IGG-3′ |
| BJM2 | 5′-AC(AG)TCCCA(AGT)AIICC(AG)TC(GA)CAIGC-3′ |
| Cbrig3Xba | 5′-GGAATCTAGAAAGCAGTGGAATTCTATTC-3′ |
| Cbrig5Xho | 5′-CACCTCTCGAGATGTCCGGCCCACCACCAC-3′ |
| Cele3Xba | 5′-CCGATGTCACTGATCTAGAATAACTGCTT-3′ |
| Cele5Xho | 5′-GTGGTTTTAAAACTCGAGAAATACATGGA-3′ |
| Cspe3 | 5′-GGATGTCATATTGGATCGCG-3′ |
| Cspe3Nhe | 5′-CGCTAGCCTATTCATTGTCATCTTCTTCCTCTAGCTC-3′ |
| Cspe5Xho | 5′-CACCCTCTCGAGATGTCTGATTCGCTAAATC-3′ |
| CspefemFD | 5′-CCAGATTTTTGTGATCGGTGGT-3′ |
| CspefemFU | 5′-ACGAGTCGAAGAAGTCGGTGG-3′ |

| | |
|---|---|
| CspefemRD | 5′-GTGTTCAAATCAACAGCACAGCA-3′ |
| CspefemRU | 5′-TCTTTCGTCCAATGTTTCTAGAGCC-3′ |
| F25FI | 5′-GTCCACGATTGCTTCGCTATCT-3′ |
| F25RO | 5′-ACACATGCTTGATGTCACGTTG-3′ |
| LVSL1 | 5′-GGTTTAATTACCCAAGTTTGAG-3′ |
| T3becomesT7 | 5′-AGCTCGGTAATACGACTCACTATAGGGAAC-3′ |
| T7big | 5′-CCAGTGAATTGTAATACGACTCACTAT-3′ |
| UAP | 5′-CUACUACUACUAGGCCACGCGTCGACTAGTAC-3′ |

I = inosine and brackets surround possible bases present at a degenerate site.

*Isolation of PP2C homologs by PCR.* Degenerate oligonucleotide primers (BJM1 and BJM2) were designed based on conserved PP2C sequences detected through the alignment of FEM-2 with other known and predicted PP2Cs. The annealing sites of these primers correspond to portions of PP2C motifs two and eight, of the 11 conserved motifs detected in a detailed sequence analysis of the PP2C superfamily (Bork et al. 1996). To isolate PP2Cs from CB5161, PCR was performed using genomic DNA as the template. Products were separated on an agarose gel, and a band corresponding to approximately 650 bp was excised from the gel and the DNA was isolated using the Sephaglas BandPrep Kit (Pharmacia Biotech). The DNA was then cloned into pGEM-T (Promega) and six clones were examined by DNA sequencing. Two of the clones, which appeared to be identical, were found to be similar to *C. elegans fem-2* (pDP#PSr1 and pDP#PSr3), while another was similar to the predicted *C. elegans* PP2C gene T23F11.1 (pDP#PSr2). The remaining three sequences did not appear to be PP2C related and were ignored. The PCR and cloning procedures were repeated using the primers F25FI and F25RO, which are designed to anneal to the predicted *C. elegans* PP2C gene F25D1.1. Two clones were sequenced and one (pDP#PS3-1p.755) was found to contain the putative CB5161 F25D1.1 ortholog. To obtain PP2C sequences from zebrafish, total RNA was isolated from adult animals using RNAzol B as directed by the supplier (Tel-Test), and it was reverse transcribed by SuperScript II reverse transcriptase (Life Technologies) using oligo-dT as a primer. The resulting cDNA was amplified using the BJM1 and BJM2 primers, and the products were separated on an agarose gel. A band corresponding to approximately 500 bp in size was excised, cloned, and characterized as described above. Three different PP2C sequences were obtained, and based on their similarity to mammalian PP2Cs we refer to them as *pp2cα1, pp2cα2,* and *pp2cβ*.

*Isolation of the complete CB5161* fem-2 *cDNA.* Primers for RT-PCR were designed based on the partial genomic sequence of *Cs-fem-2* isolated as described above. To obtain the *Cs-fem-2* cDNA, total RNA was isolated from adult CB5161 worms as described (Hansen and Pilgrim 1998) and reverse transcribed by SuperScript II reverse transcriptase (Life Technologies) using the AP primer (Life Technologies). The resulting single-stranded cDNA served as the template in subsequent PCR reactions. A DNA copy of the 3′ end of the *Cs-fem-2* transcript was obtained by performing PCR using the UAP and CspefemFU primers. A nested primer (CspefemFD) was then used with UAP for further amplification, and the resulting product was cloned into pGEM-T (Promega) and sequenced. To obtain the 5′ end of the transcript, PCR was performed using the primer CspefemRD in conjunction with the LVSL1 primer, which is complementary to the spliced leader (SL1) that is present on the 5′ end of many *C. elegans* transcripts (Blumenthal and Steward 1997). The nested primer CspefemRU was then used with LVSL1 for further amplification, and the resulting product was cloned into pGEM-T and sequenced. A DNA copy of the middle portion of the *Cs-fem-2* transcript was obtained by performing RT-PCR with the degenerate primers BJM1 and BJM2. A product of approximately 500 bp was isolated from an agarose gel and cloned into pGEM-T, and four clones were sequenced. Two were found to correspond to *Cs-fem-2,* while the others matched Cs-T23F11.1. The cDNA sequences obtained using PCR overlap to give a complete cDNA that is identical to the genomic sequence (described later), apart from absent intronic sequences, a poly-A tail, an inferred SL1 sequence, and six single base differences between the genomic and cDNA sequence. We assume that these single base changes were introduced during PCR amplification or reverse transcription of the cDNA, and hence we consider the genomic sequence to be the correct sequence. These differences between the cDNA and genomic DNA do not affect the protein sequence of the PP2C domain segment used to assess the rate of FEM-2's evolution relative to other PP2Cs. Although no signs of alternative splicing were observed during the amplification and cloning of the cDNA fragments, we cannot exclude the possibility that *Cs-fem-2* is alternatively spliced.

*Northern Analysis.* Gel electrophoresis of RNA and Northern blotting were performed as described (Hansen and Pilgrim 1998). A 1.1% agarose gel was used to electrophorese five micrograms of total RNA, which was then transferred to GeneScreen Plus nylon membrane (New England Biolabs). A $^{32}$P-labeled DNA probe prepared from the insert of clone pDP#PSr1 was hybridized to the blot using Hybrisol II hybridization solution (Oncor) overnight at 65°C. The blot was washed twice with 2× SSC for 15 min each, once with 2× SSC, 0.1% SDS for 30 min, and once with 0.1× SSC, 0.1% SDS for 10 min.

*Obtaining a genomic clone of* Cs-fem-2. Southern blotting of genomic CB5161 DNA revealed a 6-kbp *Sst*I band that hybridizes to the insert from pDP#PSr1. A mini-library in pBluescript II SK- (Stratagene) was constructed from gel purified *Sst*I-digested genomic DNA 5 to 7 kbp in size, and the library was screened with the same probe. Partial sequence of a positive clone, pDP#DH150, revealed that it did not contain the entire coding region. A partially overlapping clone, pDP#DH160, was isolated from a similar library prepared from a size selected *Bam*HI/*Hin*dIII digest. A clone containing the entire *Cs-fem-2* coding sequence was constructed by combining portions of pDP#DH150 and pDP#DH160. Correct orientation of the insert in the resulting clone, pDP#DH161, was confirmed by sequencing.

*Construction of chimeric* fem-2 *genes.* Unique *Xba*I and *Xho*I sites were added to a wild-type genomic *Ce-fem-2* clone (pDP#DH11; Hansen and Pilgrim 1998) using in vitro mutagenesis (Leatherbarrow and Fersht 1986) and the primers Cele3Xba and Cele5Xho. The resulting plasmid (pDP#DH91) contains an *Xho*I cleavage site 11 bases upstream of the *Ce-fem-2* start codon, and an *Xba*I cleavage site six bases upstream of the stop codon. To prepare a Ce/Cs *fem-2* fusion, PCR was performed using pDP#DH161 as the template and the primers Cspe5Xho and Cspe3Nhe, which introduce an *Xho*I cleavage site five bases upstream of the *Cs-fem-2* start codon, and an *Nhe*I cleavage site one base downstream of the stop codon. The Cspe3Nhe primer also introduces a silent substitution to remove an existing *Xho*I site. The PCR product was digested with *Xho*I and *Nhe*I and ligated to the promoter-containing fragment of pDP#DH91, which was generated by an *Xho*I/*Xba*I digest. The resulting clone (pDP#PS3-1p.805) was sequenced to confirm that the expected changes were present. To generate a Ce/Cb *fem-2* fusion, PCR was performed using pDP#DH53 (Hansen and Pilgrim 1998) as the template and the primers Cbrig5Xho and Cbrig3Xba. These introduce an *Xho*I cleavage site five bases upstream of the *Cb-fem-2* start codon, and an *Xba*I cleavage site 14 bases downstream of the stop codon. The PCR product was digested with *Xho*I and *Xba*I and ligated to the promoter-containing fragment of pDP#DH91, which was generated using the same enzymes. The resulting clone (pDP#PS5-1p.805) was sequenced to confirm that the expected changes were present.

*Transformation of* fem-2 *animals.* The following plasmids were assayed for their ability to rescue *fem-2* mutants: pDP#DH11 (*Ce-fem-2;* Hansen and Pilgrim 1998), pDP#DH161 (*Cs-fem-2*), pDP#DH53 (*Cb-fem-2;* Hansen and Pilgrim 1998), pDP#PS3-1p.805 (*Cs-fem-2* with *Ce-fem-2* regulatory regions), and pDP#PS5-1p.805 (*Cb-fem-2* with *Ce-fem-2* regulatory regions). To test for somatic rescue, plasmids were injected with pRF4 (*rol-6*(*su1006dm*)) into the gonads of *C. elegans*

hermaphrodites of strain DP51 (*fem-2*(*e2105*) *unc-45*(*r450ts*)/*sC1*[*dpy-1* (*s2171*)]; *him-8*(*e1489*)) as described (Mello et al. 1991). Rolling Unc hermaphrodites were transferred to separate plates and allowed to self-fertilize at 20°C. Unc m-z- (i.e. lacking *Ce-fem-2* product from both mother and zygote) hermaphrodite (XX) and male (XO) rolling animals were raised at 20°C or 25°C and compared to control DP51 animals. To test for germ-line rescue, homozygous *fem-2*(*b245*) animals, which do not produce sperm when raised at 25°C, were injected with the same plasmids. Progeny from the injected animals were raised at 25°C, and rolling animals were examined for the presence of eggs.

*RNA-mediated interference.* Complete or near-complete *fem-2* cDNAs from *C. elegans, C. briggsae,* and CB5161 were cloned into pBluescript (Stratagene). PCR using the primers T7big and T3becomesT7 was performed on each construct to incorporate T7 promoter sequences into both strands of the cDNA. RNA was synthesized from the PCR products using the MEGAscript in vitro transcription kit (Ambion), and injected at a concentration of 1 mg/ml into the gonad or gut of *C. elegans, C. briggsae,* or CB5161 worms. Injected worms were transferred to separate plates after approximately 20 h, and their progeny were examined once they reached the adult stage. Injected CB5161 worms were always kept with five males. All worms were raised at 25°C.

*Sequence Analysis.* Sequences were aligned using CLUSTAL X with the default settings (Thompson et al. 1997). The multiple alignment showing the complete FEM-2 sequences with Ce-T23F11.1 and Ce-F25D1.1 was constructed by aligning the PP2C and the flanking regions separately. To estimate the number of synonymous ($K_s$) and nonsynonymous ($K_a$) substitutions per nucleotide site, orthologous PP2C domain protein sequences were aligned, and then the corresponding nucleotide sequence alignments were prepared manually. These were analyzed using K-Estimator 5.5 (Comeron 1999), which estimates $K_s$ and $K_a$ using the method described by Comeron (1995). The correction for multiple substitutions described by Kimura (1980) was used when applicable, otherwise the Jukes and Cantor (1969) correction was applied. A phylogenetic tree was generated from aligned PP2C domains using CLUSTAL X and the neighbor-joining method (Saitou and Nei 1987). Portions of the alignment containing gaps were excluded from the analysis, and the correction for multiple substitutions was performed. The results were analyzed using the bootstrap method (1000 replicates) to provide confidence levels for the tree topology. To color the three-dimensional structure of PP2Cα according to the sequence conservation of FEM-2, the Ce-FEM-2, Cb-FEM-2, and Cs-FEM-2 PP2C domain sequences were aligned with the PP2Cα sequence. Each residue in the solved structure was then assigned a color according to the level of conservation among the aligned residues in the FEM-2 proteins, using COMBOSA3D (Stothard 2001). The PP2Cα alignment procedure was repeated using the PP2C domains from Ce-FEM-2, Ce-F25D1.1, and Ce-T23F11.1, so that the structure could be colored to show sequence conservation among FEM-2 and its paralogs. The coloring procedures were repeated using the predicted structure of the Ce-FEM-2 PP2C domain, as modeled by SWISS-MODEL (Guex and Peitsch 1997). When using the predicted structure the comparison sequences were aligned with the PP2C domain of Ce-FEM-2 instead of PP2Cα. The same localization of conserved residues was obtained, and only the comparisons using the solved PP2Cα structure are presented.

## Results

### *Isolation of* Cs-fem-2

All previous studies of sex-determining gene evolution in *Caenorhabditis* have relied on low stringency hybridization or cloning by synteny to isolate sex-determining gene orthologs (de Bono and Hodgkin 1996; Kuwabara 1996; Hansen and Pilgrim 1998; Streit et al. 1999; Haag and Kimble 2000). In this study we designed degenerate PCR primers to anneal to two conserved PP2C motifs, and these were used to isolate a genomic fragment of the CB5161 homolog of *fem-2*. Additional primers were designed based on the sequence of the genomic fragment, and these were used to obtain the complete *Cs-fem-2* cDNA using 3′ RACE and RT-PCR. Cloning of the 5′ end of the transcript was facilitated by an SL1 spliced leader sequence, which is also present on *Ce-fem-2* and *Cb-fem-2* transcripts (Pilgrim et al. 1995; Hansen and Pilgrim 1998). The size of the complete cDNA is consistent with the size detected by Northern blotting (Fig. 2). Single product bands were obtained using RT-PCR, and Northern analysis shows a single band of hybridization. Thus there is no evidence for alternative splicing of *Cs-fem-2,* which is also the case for *Ce-fem-2* and *Cb-fem-2* (Pilgrim et al. 1995; Hansen and Pilgrim 1998). A complete genomic copy of *Cs-fem-2* was obtained from a CB5161 minilibrary probed with the sequence fragment obtained using PCR (Fig. 3A). Comparison of the genomic sequence to the cDNA sequence reveals the presence of seven introns, nearly twice as many as *Ce-fem-2* (Fig. 3B). Three of the introns are located in the same positions in all three *fem-2* genes, and only one of the introns in *Cs-fem-2* is not found in either *Cb-fem-2* or *Ce-fem-2* (Fig. 4). In contrast, *Cs-fem-2* does not share any intron positions with the *C. elegans* PP2C sequences T23F11.1 or F25D1.1 (Fig. 4).

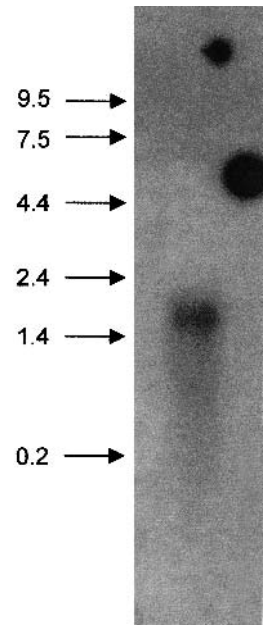


**Fig. 2.** Northern analysis shows that *Cs-fem-2* is expressed as a single transcript. Five micrograms of CB5161 total RNA was hybridized with a $^{32}$P-labelled probe generated from a genomic fragment of *Cs-fem-2.* The number indicate the relative locations of RNA size standards in kb. A single band of hybridization is seen, with an approximate size of 1.8 kb.

### *Comparison of Cs-FEM-2 to Ce-FEM-2 and Cb-FEM-2*

The protein predicted from the *Cs-fem-2* cDNA and genomic sequences consists of 483 amino acids, 34 amino acids longer than Ce-FEM-2, and 19 amino acids shorter than Cb-FEM-2. Cs-FEM-2 is 59% identical to Ce-FEM-2, which is similar to the conservation seen when Ce-

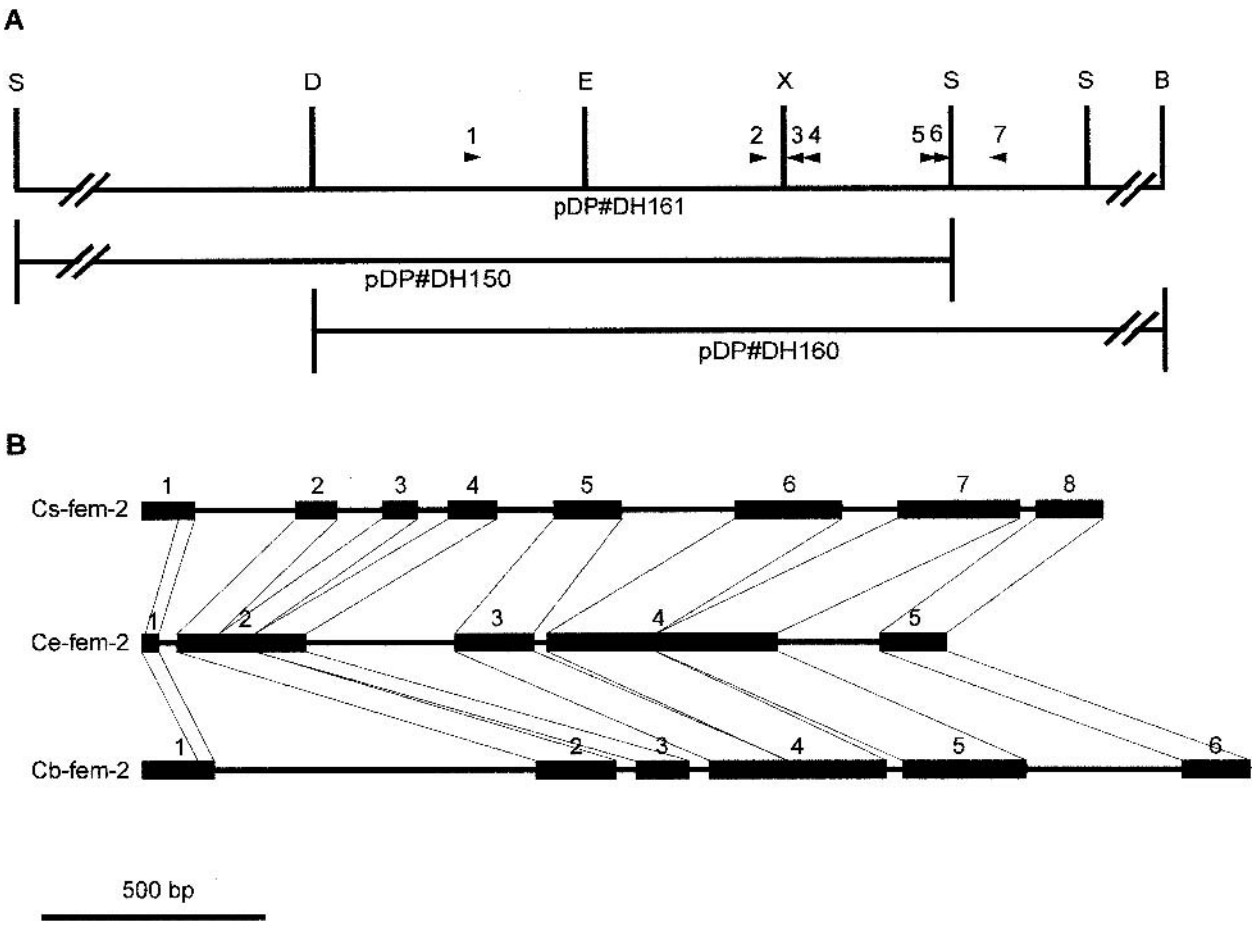


**Fig. 3.** **A:** Restriction map of the genomic region containing *Cs-fem-2*. Vertical lines represent restriction sites (B, *Bam*HI; D, *Hin*dIII; E, *Eco*RI; S, *Sst*I: and X, *Xba*I), and the numbered arrows show the approximate locations of the primers used for RT-PCR and 3′ RACE (1, Cspe3; 2, BJM1; 3, CspefemRU; 4, CspefemRD; 5, CspefemFU; 6, CspefemFD; and 7, BJM2). The upper horizontal line represents the DNA insert of clone pDP#DH161, while the lower lines represent clones pDP#DH150 and pDP#DH160. The hash marks represent portions of DNA not included in this figure. The distance between the first *Sst*I site and the *Hin*dIII site is ~5.7 kbp. The distance between the last *Sst*I site and the *Bam*HI site is ~3.5 kbp. **B:** Comparison of the genomic organization of *Cs-fem-2* with *Ce-fem-2* and *Cb-fem-2*. Boxes represent exons and intervening segments represent introns. The thinnest lines show the exon relationships between orthologs. The exons and introns of *Cs-fem-2* are aligned with the restriction map shown in **A.**

FEM-2 is compared with Cb-FEM-2, and when Cs-FEM-2 is compared to Cb-FEM-2. The residues that are mutated in two temperature-sensitive alleles of *fem-2* in *C. elegans* are conserved, as is a residue associated with functional mutations in two *Arabidopsis* PP2Cs (Fig. 4). Each protein also contains the seven metal binding and phosphate binding residues predicted by analysis of the three-dimensional structure of human PP2Cα (Das et al. 1996). In all three species FEM-2 contains a long amino terminal domain that is poorly conserved and absent from many other PP2Cs. There is also a group of acidic amino acids near the carboxy terminus that is not a general characteristic of PP2Cs (Bork et al. 1996).

### *FEM-2's PP2C Domain Evolves Faster Than the Same Domain in Other PP2Cs*

To compare the evolution of FEM-2 to the evolution of other PP2Cs, we isolated additional PP2C sequences from CB5161 and the zebrafish using PCR. The sequences were then aligned with their best matches in the GenBank database, as detected by performing BLAST searches (Altschul et al. 1997) (Fig. 5A). The PP2C domain of FEM-2 was found to be the most diverged of the PP2Cs we isolated, showing 68% identity between *C. elegans* and CB5161 (Table 1). The same domains in the *C. elegans* genes F25D1.1 and T23F11.1 were found to be 90% and 95% identical to their CB5161 orthologs, respectively. To better compare the rates of evolution of the PP2C domains, and to summarize the ortholog/paralog relationships, a phylogenetic tree was constructed using the PP2C domains we sequenced, along with the human orthologs of the zebrafish PP2C domains (Fig. 5B). The branch lengths of the tree, which serve as a better indication of evolutionary distance than simple percent differences in amino acid sequence, also indicate that the FEM-2 orthologs are more diverged than the other PP2C orthologs. The bootstrap values for the branches in the tree show that all three of the PP2C sequences from CB5161 can be unambiguously paired

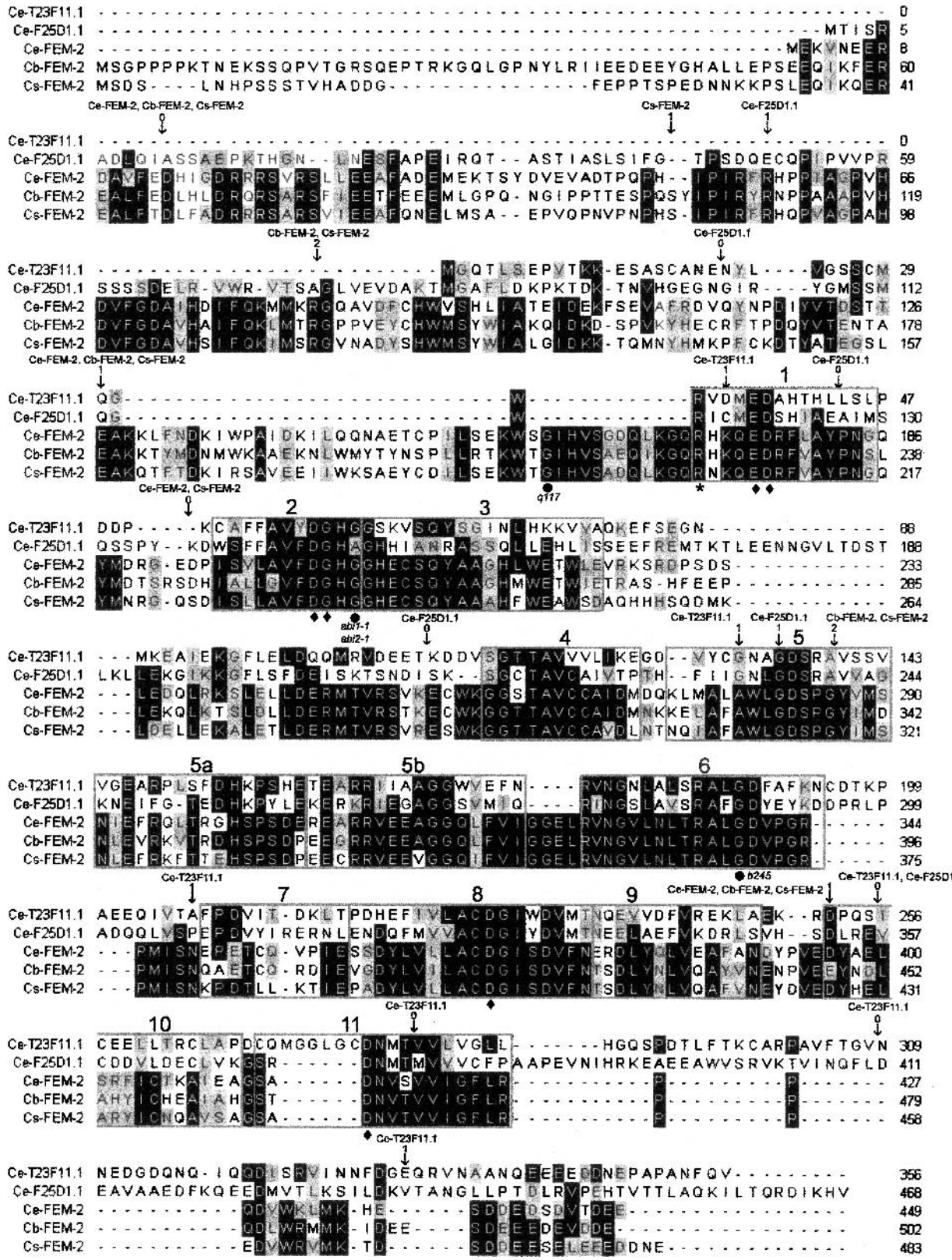


**Fig. 4.** Alignment of FEM-2 sequences with Ce-T23F11.1 and Ce-F25D1.1. Residues that are identical among three or more sequences are shown with a black background, while similar residues are shown with a gray background using the following similarity scheme: (I, L, V), (F, W, Y), (K, R, H), (D, E), (G, A, S), (P), (C), (T, N, Q, M). The PP2C motifs described by Bork et al. (1996) are surrounded by gray boxes. Black circles mark the locations of the lesions associated with two temperature-sensitive alleles of *Ce-fem-2, b245,* and *q117* (Pilgrim et al. 1995), and functionally mutant forms of the *Arabidopsis* PP2Cs ABI1 and ABI2 (Leung et al. 1994, 1997; Meyer et al. 1994). The black asterisk marks the residue thought to bind phosphate ions, and the black diamonds indicate residues thought to bind metal ions (Das et al. 1996). Arrows with numbers show the position and phase of introns (phase 0 introns are located between codons, phase 1 introns are located between the first and second base of a codon, and phase 2 introns are located between the second and third base of a codon). Each arrow is labeled with the names of the sequences that contain the intron. The GenBank accession numbers for the sequences are: Ce-T23F11.1, Z46343; Ce-F25D1.1, Z73973; Ce-FEM-2, U29515; Cb-FEM-2, AF054982; and Cs-FEM-2, AF177870.

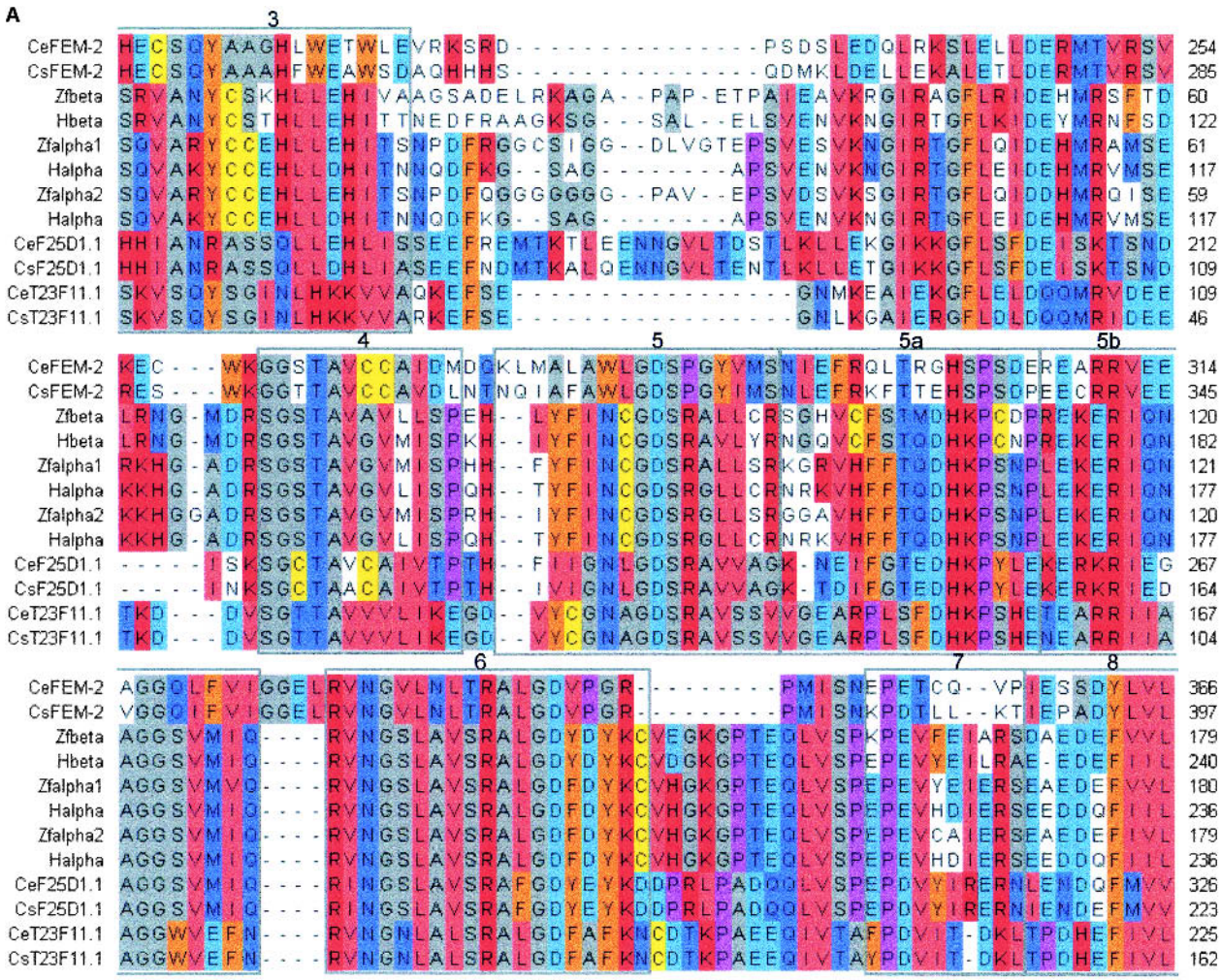


**Fig. 5. A** Legend appears on p. 271.

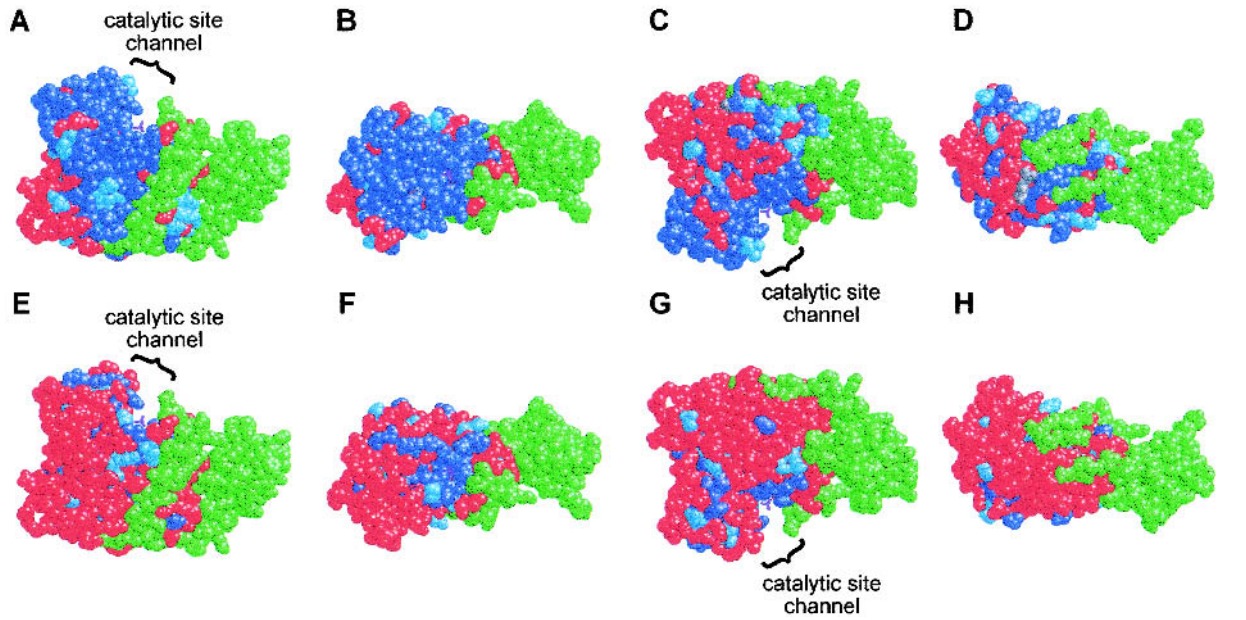


**Fig. 6.** Comparison of PP2C domain sequence conservation with the three-dimensional structure of human PP2Cα. **A:** View of human PP2Cα (Das et al. 1996) colored according to the conservation among the FEM-2 sequences. The phosphate and manganese ions are shown in purple. **B:** The same structure viewed onto the catalytic face. **C:** A view of the opposite side as that shown in **A. D:** A view of the opposite side as that shown in **B. E–H:** The same views as **A–D** but colored according to the conservation among Ce-FEM-2, Ce-F25D1.1, and Ce-T23F11.1. Residues outside of the PP2C domain are colored green, while residues within the PP2C domain that are identical or similar among all three sequences are colored dark blue and light blue, respectively. Residues that are not conserved among all three sequences are colored red. Portions of the PP2Cα sequence that align with gaps in the FEM-2 sequences are colored according to the surrounding sequence conservation (gray is used for ambiguous regions).

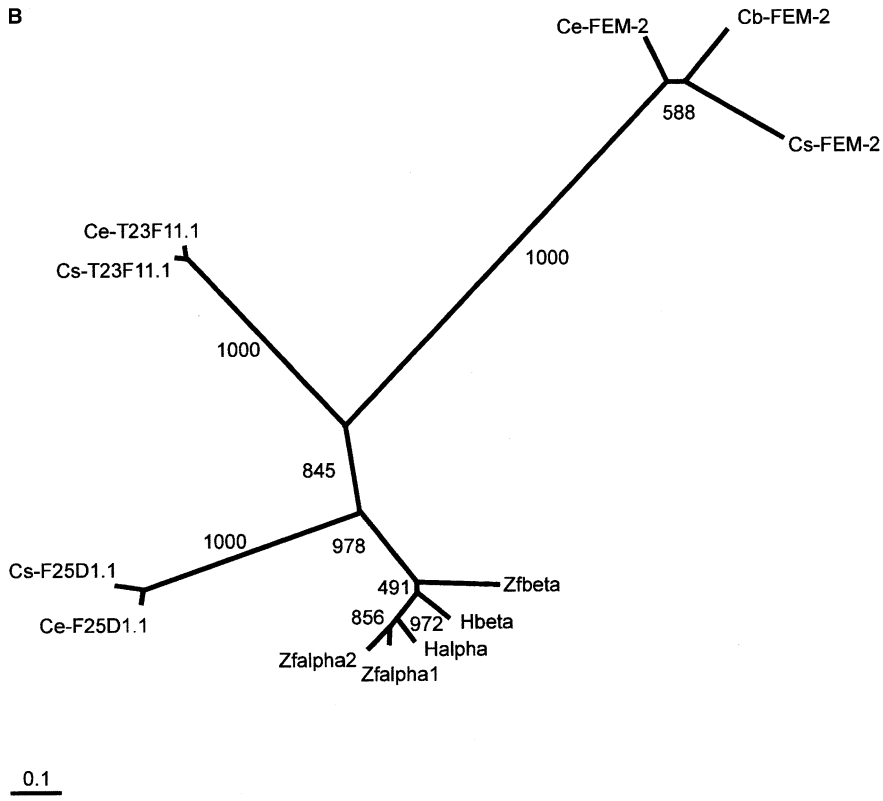


**Fig. 5.** **A:** Alignment of the PP2C domains isolated in this study (Cs-FEM-2, Cs-T23F11.1, Cs-F25D1.1, zebrafish Pp2cβ, zebrafish Pp2cα1, zebrafish Pp2cα2) with their *C. elegans* and human orthologs. Residues that are identical or similar between ortholog pairs are shown with colored backgrounds using the following similarity scheme: (I, L, V), (F, W, Y), (K, R, H), (D, E), (G, A, S), (P), (C), (T, N, Q, M). The PP2C motifs described by Bork et al. (1996) are surrounded by gray boxes. **B:** Phylogenetic analysis of the sequence segments shown in the alignment. Numbers at nodes indicate the bootstrap values (1000 replicates were performed). The GenBank accession numbers for the sequences are given in Table 1.

with their *C. elegans* orthologs. Thus the phylogenetic analysis of the predicted protein sequences, along with the conserved intron/exon boundaries, indicate that *Cs-fem-2* is the probable sequence ortholog of *Ce-fem-2.*

### *FEM-2 Has a Unique Catalytic Region That is Conserved Between* C. elegans, C. briggsae, *and CB5161*

When each non-FEM-2 partial PP2C domain sequence is aligned with its ortholog one of the most diverged regions is located between motifs three and four, while the most conserved region spans motif six (Fig. 5A). Although the motif six region is perfectly conserved within ortholog groups, between ortholog groups it differs at many sites, suggesting that it may convey specificity to the PP2C domain. The pattern of substitutions between Cs-FEM-2 and Ce-FEM-2 is similar to the pattern observed for the non-sex-determining PP2Cs. Within the FEM-2 group the region between motifs three and four has diverged, while motif six is completely conserved. The crystal structure of human PP2Cα (Das et al. 1996) indicates that the region between motifs three and four is located far from the catalytic site, while motif six is located in and around the catalytic channel. To better view the relationship between FEM-2's pattern of sequence conservation and the three-dimensional structure of the PP2C domain, the three-dimensional structure was colored according to the level of conservation among all three FEM-2 orthologs (Fig. 6A–D). The results suggest that there is a high concentration of conserved sequences on the catalytic face of the protein. When the same structure is colored according to the conservation among the *C. elegans* PP2C paralogs, it is apparent that the catalytic region differs between them (Fig. 6E–H).

### *Transgenic* Ce-fem-2, Cb-fem-2, *or* Cs-fem-2 *Can Rescue Somatic Feminization in* C. elegans fem-2 *Mutants*

To test whether the sequence divergence among the FEM-2 proteins affects regions of the protein important for masculinizing activity in the soma we introduced a variety of constructs into *C. elegans fem-2*(*null*) mutants. Wild-type XO animals normally develop as males, while XO m-z-*Ce-fem-2*(*null*) animals are feminized in both the germ line and soma. The degree of somatic feminization is temperature dependent. At 25°C the mutant XO worms are completely female, while at 20°C they have an intersexual soma that includes a partially formed vulva, a hermaphroditic double-armed gonad, and an incomplete male tail (Hodgkin 1986). XO m-z-*Ce-fem-2*(*null*) animals, either carrying a construct array or untransformed, were examined for somatic phenotype. At 20°C transgenic *Ce-fem-2, Cb-fem-2,* or *Cs-fem-2* was able to rescue the somatic feminization in *C. elegans fem-2* mutants, as were chimeric constructs consisting of the *Ce-fem-2* promoter and UTRs fused to the *Cb-fem-2*

**Table 1.** Comparison of CB5161 and zebrafish PP2C domains with their *C. elegans* and human orthologs

| Gene pair[a] | Protein % identity | Protein % similarity[b] | Number of synonymous substitutions per nucleotide site ($K_s$) | Number of nonsynonymous substitutions per nucleotide site ($K_a$) |
|---|---|---|---|---|
| *Cs-fem-2* and *Ce-fem-2* | 68 | 78 | 1.755 | 0.245 |
| Zebrafish *pp2c*β and human PP2Cβ | 73 | 83 | 1.022[c] | 0.178 |
| Zebrafish *pp2cα1* and human PP2Cα | 81 | 87 | 1.035[c] | 0.116 |
| Zebrafish *pp2cα2* and human PP2Cα | 80 | 85 | 1.266[c] | 0.122[c] |
| Cs-F25D1.1 and Ce-F25D1.1 | 90 | 94 | 1.641 | 0.048 |
| Cs-T23F11.1 and Ce-T23F11.1 | 95 | 98 | 0.958 | 0.027 |

[a] The GenBank accession numbers for the sequences are: *Ce-fem-2,* U29515; *Cs-fem-2,* AF177870; zebrafish *pp2c*β, AF177869; human PP2Cβ, AJ005801; zebrafish *pp2cα1,* AF177867; human PP2Cα, AF070670; zebrafish *pp2cα2,* AF177868; Ce-F25D1.1, Z73973; Cs-F25D1.1, AF268069; Ce-T23F11.1, Z46343; Cs-T23F11.1, AF177866.

[b] Similarity values were calculated using following similarity scheme: (I,L,V), (F,W,Y), (K,R,H), (D,E), (G,A,S), (P), (C), (T,N,Q,M).

[c] The correction for multiple substitutions described by Kimura (1980) was not applicable, so the Jukes and Cantor (1969) correction was applied.

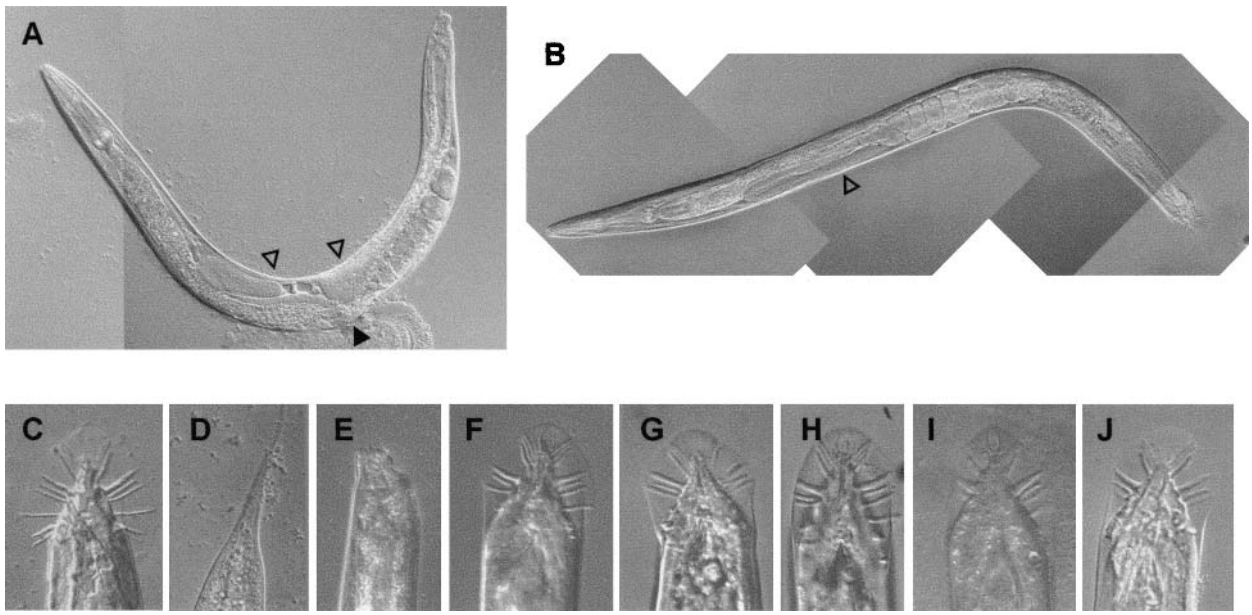


**Fig. 7.** Comparison of untransformed XO m-z-*Ce-fem-2*(*null*) worms with those carrying *fem-2* transgenic arrays. At 20°C untransformed mutant worms have a two-armed gonad, a vulva, and a slightly masculinized tail **(A),** while mutant worms carrying *fem-2* transgenes from *C. elegans, C. briggsae,* or CB5161 have a single-armed gonad, no vulva, and a more masculinized tail **(B).** The example shown in **B** is carrying a *Cb-fem-2* transgene fused to the *Ce-fem-2* regulatory regions. Open arrowheads point to gonad arms, and the closed arrowhead points to the vulva. **C–J:** Comparison of a wild-type male tail **(C)** and hermaphrodite tail **(D)** with mutant **(E)** and rescued **(F–J)** XO tails. The tails of males carrying a *Ce-fem-2* transgene **(F)** showed the same level of masculinization as those carrying a *fem-2* transgene from CB5161 **(G)** or *C. briggsae* **(H),** or a transgene consisting of the coding region of *Cs-fem-2* **(I)** or *Cb-fem-2* **(J)** fused to the *Ce-fem-2* flanking regions. The masculinization was incomplete, with missing rays and a reduced fan compared to the wild-type tail **(C).**

or *Cs-fem-2* coding regions. In each case, transformed XO animals were obtained that lacked a vulva, had a single-armed gonad, and a distinct male tail, while untransformed XO animals always had a vulva, a double-armed gonad, and a slightly masculinized tail (Fig. 7; Table 2). The masculinization of the tail was not complete in any of the transgenic animals, including those carrying a *Ce-fem-2* transgene. At 25°C many transgenic animals were obtained for each construct that lacked a vulva, had a single-armed gonad, and a distinct male tail, while control animals were always completely female (data not shown).

### *Transgenic* Ce-fem-2 *but Not* Cb-fem-2 *or* Cs-fem-2 *Can Rescue Germ-Line Feminization in* C. elegans fem-2 *Mutants*

Genetic evidence indicates that *fem-2* has distinct targets in the germ line and soma in *C. elegans* (Hodgkin 1986), raising the possibility that the FEM-2 protein from another species could act on the targets in one tissue but not the other. To test whether transgenic *fem-2* from *C. briggsae* or CB5161 can rescue germ-line feminization in *C. elegans* a variety of constructs were injected into *fem-2*(*b245*) animals, which are unable to produce sperm

**Table 2.** Attempted rescue of the somatic feminization of XO *C. elegans fem-2(null)* animals using *Ce-fem-2, Cb-fem-2,* or *Cs-fem-2* transgenic arrays

| Construct | % with male tail (n)[a] | % with single-armed gonad (n)[a] | % without vulval tissue (n)[a] |
|---|---|---|---|
| *Ce-fem-2* | 100 (17) | 100 (17) | 100 (17) |
| *Cb-fem-2* | 100 (34) | 100 (34) | 100 (34) |
| *Cs-fem-2* | 100 (30) | 100 (29) | 97 (30) |
| *Ce-fem-2: Cb-fem-2*[b] | 100 (22) | 100 (22) | 100 (22) |
| *Ce-fem-2: Cs-fem-2*[c] | 96 (25) | 96 (25) | 96 (25) |
| None[d] | 0 (37) | 0 (37) | 0 (37) |

[a] n, number of rolling XO animals examined. These animals were raised at 20°C. In the absence of a rescuing construct *fem-2(null)* m-z-XO animals are smaller than XX animals, and they have a slightly masculinized tail.
[b] The regions flanking the coding sequence in the rescuing *C. elegans* construct fused to the *C. briggsae fem-2* coding sequence.
[c] The regions flanking the coding sequence in the rescuing *C. elegans* construct fused to the CB5161 *fem-2* coding sequence.
[d] *fem-2(null)* m-z-XO progeny from non rolling hermaphrodites.

**Table 3.** Attempted rescue of the germ-line feminization of XX *C. elegans fem-2(b245)* animals using *Ce-fem-2, Cb-fem-2,* or *Cs-fem-2* transgenic arrays

| Construct | % rolling XX animals laying eggs (n)[a] |
|---|---|
| *Ce-fem-2* | 83 (12) |
| *Cb-fem-2* | 0 (11) |
| *Cs-fem-2* | 0 (10) |
| *Ce-fem-2: Cb-fem-2*[b] | 0 (15) |
| *Ce-fem-2: Cs-fem-2*[c] | 0 (12) |

[a] n, number of rolling XX animals examined. These animals were the F1 progeny of injected animals and were raised at 25°C.
[b] The regions flanking the coding sequence in the rescuing *C. elegans* construct fused to the *C. briggsae fem-2* coding sequence.
[c] The regions flanking the coding sequence in the rescuing *C. elegans* construct fused to the CB5161 *fem-2* coding sequence.

when raised at 25°C. The advantage of using the *b245* mutation is that animals raised at 20°C can be injected and moved to 25°C. The progeny of the injected animals can then be examined for germ-line rescue by their ability to reproduce at 25°C. Assaying the direct progeny of injected animals is important because transgenes in *C. elegans* show progressive germ-line silencing over generations (Kelly et al. 1997). Transgenic *Ce-fem-2* was able to restore sperm production in *fem-2*(*b245*) animals, as indicated by the presence of hermaphrodites with eggs after growth at the restrictive temperature. The brood size of the rescued hermaphrodites was variable (between 15 and 188), and the average brood size was 106. *Cb-fem-2* or *Cs-fem-2* were unable to rescue the germ-line feminization, as were chimeric constructs consisting of the *Ce-fem-2* promoter and UTRs fused to the *Cb-fem-2* or *Cs-fem-2* coding regions (Table 3).

### *RNA-Mediated Interference Against* fem-2 *Feminizes the Germ Line of* C. elegans *and* C. briggsae *but Not CB5161*

To address whether *Cb-fem-2* or *Cs-fem-2* are sex-determining genes in their respective species we used RNA-mediated interference (RNAi), a technique that has been used to silence genes in *C. elegans* (Fire et al. 1998), *C. briggsae* (Kuwabara 1996; Streit et al. 1999; Molin et al. 2000; Rudel and Kimble 2001), and *C. remanei* (Haag and Kimble 2000; Rudel and Kimble 2001). RNAi against *fem-2* in *C. elegans* caused germ-line feminization in males and hermaphrodites, but did not feminize the male soma (Fig. 8; Table 4). In *C. briggsae, fem-2*(*RNAi*) feminized the male germ line (Fig. 8; Table 4), but not the hermaphrodite germ line. We were unable to detect any reduction in the brood size of hermaphrodites arising from injected animals, suggesting that normal numbers of sperm were produced (data not shown). Although RNAi did not shift the sex ratio towards females, a few intersex animals were observed, which had a male tail, a single-armed gonad, and what appeared to be a hermaphrodite vulva. However, the role of *Cb-fem-2* in the soma remains uncertain because few animals were affected, and the tail and somatic gonad were not feminized. In CB5161, *fem-2*(*RNAi*) did not cause noticeable effects, with males having a normal germ line and soma. No other obvious defects were observed in any of the species. These results suggest that *fem-2* may not play a significant role in the germ line of CB5161 males or *C. briggsae* hermaphrodites, or that RNAi is less effective at inhibiting *fem-2* activity in these species.

## Discussion

Protein phosphatases play important roles in regulating a variety of cellular processes (Barford 1996; Pilgrim et al. 1995). The PP2C class of phosphatases is one of the two classes of enzymes that can dephosphorylate serine and threonine residues (Bork et al. 1996), and they are defined biochemically as having phosphatase activity that is $Mg^{2+}$-dependent or $Mn^{2+}$-dependent and okadaic acid resistant (Cohen 1989). Sequence analysis of the phosphatases meeting these criteria shows that they comprise a conserved protein superfamily (Bork et al. 1996). In addition to the FEM-2 sex-determining protein, there are three predicted PP2C genes in *C. elegans,* two of which contain a PP2C domain that is close in size to the one found in FEM-2. From a variety of sequence comparisons it appears that we have isolated the orthologs of the *C. elegans fem-2* gene from *C. briggsae* (Hansen and Pilgrim 1998) and CB5161. Cb-FEM-2 and Cs-FEM-2 are much more similar to *C. elegans* FEM-2 than to other *C. elegans* PP2Cs, they contain conserved sequence segments absent from other PP2Cs, and the positions of

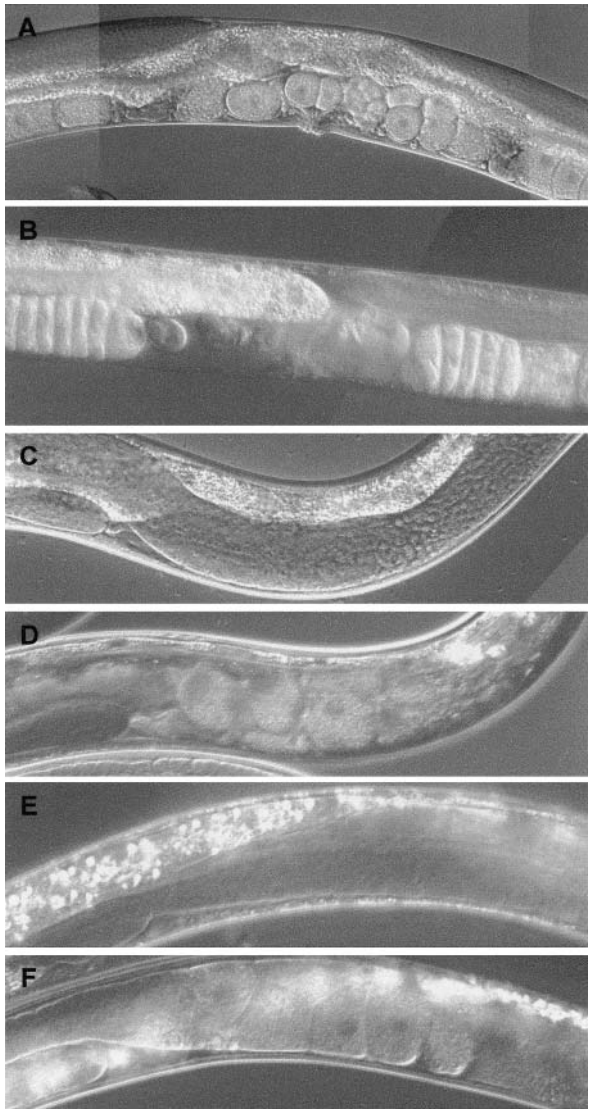


**Fig. 8.** Feminization of the germ line of *C. elegans* and *C. briggsae* animals using RNAi against *fem-2*. *C. elegans* worms arising from uninjected animals appeared to be normal hermaphrodites **(A)** or males **(C). B:** A typical feminized *C. elegans* XX worm containing stacked oocytes and no eggs. **D:** A typical feminized *C. elegans* XO worm containing oocyte-like cells in a single-armed gonad. *C. briggsae* XX worms arising from injected animals appeared to be identical to those arising from uninjected controls (not shown). **E:** *C. briggsae* XO animals arising from uninjected animals appeared to be normal males. **F:** A typical feminized *C. briggsae* XO worm containing oocyte-like cells in a single-armed gonad.

three introns are identical in all three *fem-2* sequences. The grouping of the additional CB5161 PP2C sequences with their *C. elegans* orthologs was straightforward, based on their high level of sequence identity. We cannot exclude the possibility that missing sequence information has led to incorrect ortholog assignments. Even with complete genome knowledge ortholog assignments remain uncertain because of the possibility of gene loss.

To assess how the observed substitutions in FEM-2 might affect the function of the protein we first examined the three-dimensional structure of human PP2Cα. In other studies the availability of x-ray crystallographic structures has made it possible to identify evolutionary changes that are likely to have important functional consequences. For example, comparisons of plant chitinase sequences with a chitinase three-dimensional structure revealed that numerous replacements occur in the active site cleft of the enzyme, consistent with coevolutionary interactions with pathogens (Bishop et al. 2000). In our analysis of FEM-2 we first looked at the residues predicted to bind metal and phosphate ions, and found that they were invariant. The conservation of the remaining residues appears to be related to their position in the PP2C domain, with those covering the catalytic face being conserved, and those remote from the catalytic site having diverged. The conservation of the catalytic face suggests that the protein dephosphorylates a substrate domain that is also similar among the three species we examined. Currently, the targets of FEM-2's phosphatase activity in *C. elegans* and the other nematode species are not known.

Although the diverged sequences in FEM-2's PP2C domain are concentrated away from the catalytic site, they could still play an important role in binding a targeting protein or substrate. Also, FEM-2 contains a long amino terminal domain that is absent from most other PP2C proteins and required for sex-determining function in *C. elegans,* but not for phosphatase activity (Hansen and Pilgrim 1998). We tested the *C. briggsae* and CB5161 *fem-2* orthologs for activity in *C. elegans,* to try to address more directly if the protein divergence affects function. *Cb-fem-2* has been previously shown to rescue the somatic feminization that occurs in XO animals carrying a mutant form of *fem-2* (Hansen and Pilgrim 1998). The *C. briggsae* orthologs of *her-1* and *tra-1* have also been introduced into *C. elegans,* and in both cases the genes were found to function in many of tissues regulated by the *C. elegans* gene (de Bono and Hodgkin 1996; Streit et al. 1999). In the case of *tra-1,* no rescue was observed in the somatic gonad or the germ line (de Bono and Hodgkin 1996). Although the *Cs-fem-2* and *Cb-fem-2* transgenes we tested could rescue the somatic feminization, they were unable to rescue the germ-line feminization caused by a *fem-2* mutation. This result suggests that the regions of FEM-2 that mediate an interaction with a germ-line substrate or targeting protein in *C. elegans* are not conserved. It seems unlikely that the absence of germ-line rescue reflects a lack of transgene expression, since a *Ce-fem-2* transgene was able to function robustly in the same assay. Even when the *fem-2* coding sequences from the other species were surrounded by the *Ce-fem-2* regulatory regions germ-line rescue was not seen.

Comparison of the *C. elegans* PP2Cs FEM-2, F25D1.1, and T23F11.1 with their orthologs in CB5161 shows that FEM-2's PP2C domain is the most diverged. A previous analysis of sequences suggests that FEM-2 is

**Table 4.** Effects of RNAi against *fem-2* in *C. elegans, C. briggsae,* and CB5161

| Species | Injected? | % XX feminized germ line (n)[a,b] | % intersex soma (n)[a,c] | % male soma, feminized germ line (n)[a,c] | % F1 males (n)[a,c] |
|---|---|---|---|---|---|
| *C. elegans* | Yes | 46 (347) | 0 (120) | 52 (19) | 37 (236) |
| | No | 0 (168) | 0 (88) | 0 (12) | 40 (375) |
| *C. briggsae* | Yes | 0 (171) | 8[d] (145) | 48 (33) | 26 (229) |
| | No | 0 (92) | 1[d] (105) | 4 (55) | 23 (267) |
| CB5161 | Yes | — | 0 (123) | 0 (123) | 47 (696) |
| | No | — | 0 (69) | 0 (69) | 46 (356) |

[a] n, number of animals examined.
[b] Hermaphrodites of *C. elegans* strain N2 or *C. briggsae* strain AF16 were injected and adult progeny were examined to see whether XX animals lost the ability to produce sperm.
[c] Unmated *C. elegans him-8(e1489)* hermaphrodites or *C. briggsae mih-3(s1290)* (a high incidence of males strain) hermaphrodites were injected. CB5161 females were crossed with males prior to and following injection. Adult progeny were examined for intersex somatic characteristics, XO germ-line feminization, and sex ratio.
[d] These intersex animals had a partially-formed vulva, a wild-type male tail, and a single-armed gonad.

also evolving rapidly between *C. briggsae* and *C. elegans* (Hansen and Pilgrim 1998). Why is FEM-2 evolving more rapidly than other proteins? In cases where positive selection is thought to cause sequence divergence, the active sites of proteins are often affected. For example, positive selection is believed to have caused a high level of amino acid divergence in the antigen-binding clefts of certain major-histocompatibility-complex molecules (Hughes et al. 1990), the active center regions of serine protease inhibitors (Hill and Hastie 1987), and the active site clefts of chitinases (Bishop et al. 2000). Our comparison of FEM-2's PP2C domain divergence to the three-dimensional structure of another PP2C domain suggests that the active site region of FEM-2 is conserved. Also, although FEM-2's PP2C domain contains more diverged residues than the other PP2C domains, the locations of many of the diverged residues are in regions that typically diverge between PP2C orthologs. Despite the consistency of the divergence pattern with neutral evolution we cannot rule out positive selection, particularly since the *fem-2* genes from the other species were unable to function in the *C. elegans* germ line. Tests commonly used to distinguish between positive selection and neutral evolution cannot be applied here because the high level of synonymous divergence between the sequences would mask evidence of positive selection, and polymorphisms within *C. elegans* are rare even in noncoding DNA (Kock et al. 2000). Indeed, we sequenced a diverged region of *fem-2*'s PP2C domain and a portion of another PP2C gene (including an intron) from a variety of geographic isolates and found no polymorphisms (data not shown). It is not clear why sex-determining genes would be subjected to positive selection. One hypothesis is that changes might facilitate reproductive isolation during the early stages of speciation, by causing hybrid sterility (Whitfield et al. 1993). Another explanation is that there could be selection for mutations that adjust the level or timing of gamete production towards an optimum (Hodgkin and Barnes 1991). Accelerated neutral evolution may occur because changes in sex-determining proteins are tolerated more than changes in other proteins. In *C. elegans* a variety of sex-determining gene mutations have a slight feminizing or masculinizing effect and yield fertile animals. Proteins that regulate other processes may not be as free to diverge, since changes in their activity may not simply skew the cell towards an alternate viable state. Thus equivalent mutations in two similar proteins, one that regulates sex determination and one that does not, could have dramatically different effects on the fitness of the organism.

The fact that a gene from a foreign species can rescue a mutant phenotype does not show that the gene regulates the same pathway in both species. Given the complexity of the sex-determining pathway, and its lack of conservation compared to pathways that regulate other developmental processes, it would not be surprising to identify orthologous genes that have a sex-determining role in one species but not in another. In *Drosophila* the *Sex-lethal* (*Sxl*) gene is a key regulator of sex determination, and clear orthologs have been isolated from *Chrysomya* (Muller-Holtkamp 1995), *Megaselia* (Sievert et al. 1997), *Musca* (Meise et al. 1998), and *Ceratitis* (Saccone et al. 1998). However, the alternative splicing of *Sxl* that determines sex in *Drosophila* is not observed in any of the other species, suggesting that its developmental role may not be conserved. The most downstream *Drosophila* gene is *doublesex,* and its ortholog has been isolated from *Megaselia* (Sievert et al. 1997) and *Bactrocera* (Shearman and Frommer 1998), and in both cases it shows the sex-specific splicing observed in *Drosophila.* Although the function of the orthologs remains to be tested in the other species, the results thus far are consistent with the model proposed by Wilkins (1995), in which sex-determining pathways are built from the bottom up. In *Caenorhabditis* the technique of RNA-mediated interference can be used to silence specific genes, and it has been used to test whether specific sex-determining genes have a conserved sex-determining role. RNAi has been performed against *Cb-tra-2* and

*Cb-her-1,* and in both cases the phenotype in *C. briggsae* was similar to the *C. elegans* phenotype (Kuwabara 1996; Streit et al. 1999). The *tra-2* gene was isolated from the closely related male/female species *C. remanei* and RNAi performed against *Cr-tra-2* showed that *tra-2* has a conserved sex-determining role (Haag and Kimble 2000). In this study we found that RNAi against *Ce-fem-2* can cause the germ-line feminization seen in animals carrying a *fem-2* mutation. When RNAi was performed against the *fem-2* ortholog in *C. briggsae,* the male germ line was feminized, but not the hermaphrodite germ line. There are a variety of explanations for why *fem-2*(*RNAi*) in *C. briggsae* appears to feminize the male germ line but not the hermaphrodite germ line. RNAi may act more effectively at the time or place that *fem-2* activity is required in the male, the male germ line may be more sensitive to a reduction in *fem-2* activity, or slight germ-line feminization may be easier to observe in male animals. A more intriguing possibility is that *fem-2* in *C. briggsae* regulates germ-line sex in males but not hermaphrodites. The absence of a *fem-2*(*RNAi*) phenotype in CB5161 cannot be interpreted at this time, because no cases of successful RNAi have been reported for this species. A better understanding of the role of *fem-2* in *C. briggsae* and CB5161 may require the isolation of mutations. With the availability of the genomic sequences of *Cb-fem-2* and *Cs-fem-2,* deletion mutants could eventually be obtained using the PCR-based selection schemes currently used in *C. elegans.*

A scenario in which the *fem-2* ortholog in a related species does not regulate some aspects of sex determination may still be consistent with Wilkins' model of the evolution of the *C. elegans* pathways, even in cases where the upstream components have a conserved sex-determining role. In *C. elegans,* the *fem-1* and *fem-3* genes act at the same level in the sex-determining pathway as *fem-2,* and the FEM-3 protein has been shown to interact directly with FEM-2, but the function of the interaction is unknown (Chin-Sang and Spence 1996). In the germ line two additional genes, *fog-1,* and *fog-3,* also act with the *fem* genes to promote the male fate (Barton and Kimble 1990; Ellis and Kimble 1995). Wilkins suggests that the levels of the pathway that consist of multiple genes, such as the *fem* level, could have grown in later steps, after the assembly of a core pathway. The *fem-3* gene is a good candidate for the most ancient *fem* gene, as it is the only one known to be directly regulated by *tra-2.* The *fem-1* and *fem-2* genes may have joined *fem-3* during subsequent optimization of the pathway. If the orthologs of *fem-1* and *fem-3* are isolated from *C. briggsae* or CB5161 it will be possible to test whether RNAi against these genes can cause feminization, and whether the FEM-2/FEM-3 interaction is maintained within species and between species. A recent and unexpected finding shows that in *C. elegans* a soluble version of TRA-2 binds to TRA-1 directly (Lum et al. 2000; Wang and Kimble 2001). Although this interaction is thought to play a minor role in *C. elegans* sex determination, it could have preceded the addition of the *fem* genes during the evolution of the pathway. Perhaps in some species it takes the place of one or more of the *fem* genes in the germ-line or soma. By isolating and testing the function of other sex-determining orthologs it should be possible to determine which components have been added most recently, and thus better understand how complex regulatory pathways evolve.

*Acknowledgements.* We thank Warren Gallin for critical reading of this manuscript, and Angela Manning for providing zebrafish RNA. Some nematode strains used in this work were provided by the Caenorhabditis Genetics Center, which is funded by the NIH National Center for Research Resources (NCRR). PS is supported by a Natural Sciences and Engineering Research Council of Canada post-graduate scholarship and an Alberta Heritage Foundation for Medical Research studentship. This work was supported by a grant from the Natural Sciences and Engineering Research Council of Canada to DP.

## References

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402

Barford D (1996) Molecular mechanisms of the protein serine/threonine phosphatases. Trends Biochem Sci 21:407–412

Barton MK, Kimble J (1990) *fog-1,* a regulatory gene required for specification of spermatogenesis in the germ line of *Caenorhabditis elegans.* Genetics 125:29–39

Bishop JG, Dean AM, Mitchell-Olds T (2000) Rapid evolution in plant chitinases: molecular targets of selection in plant-pathogen coevolution. Proc Natl Acad Sci USA 97:5322–5327

Blumenthal T, Steward K (1997) RNA processing and gene structure. In: Riddle D, Blumenthal T, Meyer B, Priess J (eds) *C. elegans* II. Cold Spring Harbor Laboratory Press, New York, pp. 117–145

Bork P, Brown NP, Hegyi H, Schultz J (1996) The protein phosphatase 2C (PP2C) superfamily: detection of bacterial homologues. Protein Sci 5:1421–1425

Chin-Sang ID, Spence AM (1996) *Caenorhabditis elegans* sex-determining protein FEM-2 is a protein phosphatase that promotes male development and interacts directly with FEM-3. Genes Dev 10:2314–2325

Cline TW, Meyer BJ (1996) Vive la difference: males vs. females in flies and worms. Annu Rev Genet 30:637–702

Cohen P (1989) The structure and regulation of protein phosphatases. Annu Rev Biochem 58:453–508

Comeron JM (1995) A method for estimating the numbers of synonymous and nonsynonymous substitutions per site. J Mol Evol 41: 1152–1159

Comeron JM (1999) K-Estimator: calculation of the number of nucleotide substitutions per site and the confidence intervals. Bioinformatics 15:763–764

Das AK, Helps NR, Cohen PT, Barford D (1996) Crystal structure of the protein serine/threonine phosphatase 2C at 2.0 A resolution. EMBO J 15:6798–6809

de Bono M, Hodgkin J (1996) Evolution of sex determination in *Caenorhabditis:* unusually high divergence of *tra-1* and its functional consequences. Genetics 144:587–589

Ellis RE, Kimble J (1995) The *fog-3* gene and regulation of cell fate in the germ line of *Caenorhabditis elegans.* Genetics 139:561–577

Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC (1998) Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans.* Nature 391:806–811

Fitch DH, Bugaj-Gaweda B, Emmons SW (1995) 18S ribosomal RNA gene phylogeny for some Rhabditidae related to *Caenorhabditis.* Mol Biol Evol 12:346–358

Guex N, Peitsch MC (1997) SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. Electrophoresis 18:2714–2723

Haag ES, Kimble J (2000) Regulatory elements required for development of *Caenorhabditis elegans* hermaphrodites are conserved in the *tra-2* homologue of *C. remanei,* a male/female sister species. Genetics 155:105–116

Hansen D, Pilgrim D (1998) Molecular evolution of a sex determination protein: FEM-2 (PP2C) in *Caenorhabditis.* Genetics 149: 1353–1362

Hansen D, Pilgrim D (1999) Sex and the single worm: sex determination in the nematode *Caenorhabditis elegans.* Mech Dev 83:3–15

Hill RE, Hastie ND (1987) Accelerated evolution in the reactive centre regions of serine protease inhibitors. Nature 326:96–99

Hodgkin J (1996) Sex determination in the nematode *C. elegans:* analysis of *tra-3* suppressors and characterization of *fem* genes. Genetics 114:15–52

Hodgkin J, Barnes TM (1991) More is not better: brood size and population growth in a self-fertilizing nematode. Proc R Soc Lond B Biol Sci 246:19–24

Hughes AL, Ota T, Nei M (1990) Positive Darwinian selection promotes charge profile diversity in the antigen-binding cleft of class I major-histocompatibility-complex molecules. Mol Biol Evol 7: 515–524

Jukes TH, Cantor CR (1969) Evolution of protein molecules. In: Munro HN (ed) Mammalian protein metabolism. Academic Press, New York, pp. 21–132

Kelly WG, Xu S, Montgomery MK, Fire A (1997) Distinct requirements for somatic and germline expression of a generally expressed *Caenorhabditis elegans* gene. Genetics 146:227–238

Kimble J, Edgar L, Hirsch D (1984) Specification of male development in *Caenorhabditis elegans:* The *fem* genes. Dev Biol 105:234–239

Kimura M (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J Mol Evol 16:111–120

Koch R, van Luenen HG, van der Horst M, Thijssen KL, Plasterk RH (2000) Single nucleotide polymorphisms in wild isolates of *Caenorhabditis elegans.* Genome Res 10:1690–1696

Kuwabara PE (1996) Interspecies comparison reveals evolution of control regions in the nematode sex-determining gene *tra-2.* Genetics 144:597–607

Kuwabara PE, Okkema PG, Kimble J (1992) *tra-2* encodes a membrane protein and may mediate cell communication in the *Caenorhabditis elegans* sex determination pathway. Mol Biol Cell 3: 461–473

Leatherbarrow RJ, Fersht AR (1986) Protein engineering [review]. Protein Eng 1:7–16

Leung J, Bouvier-Durand M, Morris PC, Guerrier D, Chefdor F, Giraudat J (1994) *Arabidopsis* ABA response gene *ABI1:* features of a calcium-modulated protein phosphatase. Science 264:1448–1452

Leung J, Merlot S, Giraudat J (1997) The *Arabidopsis ABSCISIC ACID-INSENSITIVE2* (*ABI2*) and *ABI1* genes encode homologous protein phosphatases 2C involved in abscisic acid signal transduction. Plant Cell 9:759–771

Lum DH, Kuwabara PE, Zarkower D, Spence AM (2000) Direct protein-protein interaction between the intracellular domain of TRA-2 and the transcription factor TRA-1A modulates feminizing activity in *C. elegans.* Genes Dev 14:3153–3165

Madl EM, Herman RK (1979) Polyploids and sex determination in *Caenorhabditis elegans.* Genetics 93:393–402

McAllister BF, McVean GA (2000) Neutral evolution of the sex-determining gene *transformer* in *Drosophila*. Genetics 154:1711–1720

Meise M, Hilfiker-Kleiner D, Dubendorfer A, Brunner C, Nothiger R, Bopp D (1998) *Sex-lethal,* the master sex-determining gene in *Drosophila,* is not sex-specifically regulated in *Musca domestics.* Development 125:1487–1494

Mello CC, Kramer JM, Stinchcomb D, Ambros V (1991) Efficient gene transfer in *C. elegans:* extrachromosomal maintenance and integration of transforming sequences. EMBO J 10:3959–3970

Metz EC, Palumbi SR (1996) Positive selection and sequence rearrangements generate extensive polymorphism in the gamete recognition protein binding. Mol Biol Evol 13:397–406

Meyer K, Leube MP, Grill E (1994) A protein phosphatase 2C involved in ABA signal transduction in *Arabidopsis thaliana.* Science 264: 1452–1455

Molin L, Mounsey A, Aslam S, Bauer P, Young J, James M, Sharma-Oates A, Hope IA (2000) Evolutionary conservation of redundancy between a diverged pair of forkhead transcription factor homologues. Development 127:4825–4835

Muller-Holtkamp F (1995) The *Sex-lethal* gene homologue in *Chrysomya rufifacies* is highly conserved in sequence and exon-intron organization. J Mol Evol 41:467–477

O'Neil MT, Belote JM (1992) Interspecific comparison of the *transformer* gene of *Drosophila* reveals an unusually high degree of evolutionary divergence. Genetics 131:113–128

Pilgrim D, McGregor A, Jackle P, Johnson T, Hansen D (1995) The *C. elegans* sex-determining gene *fem-2* encodes a putative protein phosphatase. Mol Biol Cell 6:1159–1171

Raymond CS, Murphy MW, O'Sullivan MG, Bardwell VJ, Zarkower D (2000) *Dmrt1,* a gene related to worm and fly sexual regulators, is required for mammalian testis differentiation. Genes Dev 14: 2587–2595

Raymond CS, Shamu CE, Shen MM, Seifert KJ, Hirsch B, Hodgkin J, Zarkower D (1998) Evidence for evolutionary conservation of sex-determining genes. Nature 391:691–695

Rudel D, Kimble J (2001) Conservation of *glp-1* regulation and function in nematodes. Genetics 157:639–654

Saccone G, Peluso I, Artiaco D, Giordano E, Bopp D, Polito LC (1998) The *Ceratitis capitata* homologue of the *Drosophila* sex-determining gene *Sex-lethal* is structurally conserved, but not sex-specifically regulated. Development 125:1495–1500

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425

Shearman DC, Frommer M (1999) The *Bactrocera tryoni* homologue of the *Drosophila melanogaster* sex-determination gene *doublesex.* Insect Mol Biol 7:355–366

Sievert V, Kuhn S, Traut W (1997) Expression of the sex-determining cascade genes *Sex-lethal* and *doublesex* in the phorid fly *Megaselia scalaris.* Genome 40:211–214

Stothard PM (2001) COMBOSA3D: combining sequence alignments with three-dimensional structures. Bioinformatics 17:198–199

Streit A, Li W, Robertson B, Schein J, Kamal IH, Marra M, Wood WB (1999) Homologs of the *Caenorhabditis elegans* masculinizing gene *her-1* in *C. briggsae* and the filarial parasite *Brugia malayi.* Genetics 152:1573–1584

Sulston J, Horvitz HR (1977) Post-embryonic cell lineages of the nematode *Caenorhabditis elegans.* Dev Biol 56:110–156

Swanson WJ, Vacquier VD (1998) Concerted evolution in an egg receptor for a rapidly evolving abalone sperm protein. Science 281: 710–712

Thomas WK, Wilson AC (1991) Mode and tempo of molecular evolution in the nematode *Caenorhabditis:* cytochrome oxidase II and calmodulin sequences. Genetics 128:269–279

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL X windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 25:4876–4882

Tucker PK, Lundrigan BL (1993) Rapid evolution of the sex determining locus in Old World mice and rats. Nature 364:715–717

Walthour CS, Schaeffer SW (1994) Molecular population genetics of sex determination genes: the *transformer* gene of *Drosophila melanogaster.* Genetics 136:1367–1372

Wang S, Kimble J (2001) The TRA-1 transcription factor binds TRA-2 to regulate sexual fates in *Caenorhabditis elegans.* EMBO J 20: 1363–1372

Whitfield LS, Lovell-Badge R, Goodfellow PN (1993) Rapid sequence evolution of the mammalian sex-determining gene *SRY.* Nature 364:713–715

Wilkins AS (1995) Moving up the hierarchy: a hypothesis on the evolution of a genetic sex determination pathway. BioEssays 17:71–77

Zarkower D, Hodgkin J (1992) Molecular analysis of the *C. elegans* sex-determining gene *tra-1:* a gene encoding two zinc finger proteins. Cell 70:237–249