Development 126, 873-881 (1999)
Printed in Great Britain © The Company of Biologists Limited 1999
DEV5283

Similarity of DNA binding and transcriptional regulation by *Caenorhabditis elegans* MAB-3 and *Drosophila melanogaster* DSX suggests conservation of sex determining mechanisms

Woelsung Yi[1] and David Zarkower[1,2,*]

[1]Biochemistry, Molecular Biology and Biophysics Graduate Program, and [2]Department of Genetics, Cell Biology and Development, University of Minnesota, Minneapolis, MN 55455, USA
*Author for correspondence (e-mail: zarkower@gene.med.umn.edu)

Accepted 17 December 1998; published on WWW 2 February 1999

SUMMARY

Although most animals occur in two sexes, the molecular pathways they employ to control sexual development vary considerably. The only known molecular similarity between phyla in sex determination is between two genes, *mab-3* from *C. elegans*, and *doublesex* (*dsx*) from *Drosophila*. Both genes contain a DNA binding motif called a DM domain and they regulate similar aspects of sexual development, including yolk protein synthesis and peripheral nervous system differentiation. Here we show that MAB-3, like the DSX proteins, is a direct regulator of yolk protein gene transcription. We show that despite containing different numbers of DM domains MAB-3 and DSX bind to similar DNA sequences. *mab-3* mutations deregulate vitellogenin synthesis at the level of transcription, resulting in expression in both sexes, and the vitellogenin genes have potential MAB-3 binding sites upstream of their transcriptional start sites. MAB-3 binds to a site in the *vit-2* promoter in vitro, and this site is required in vivo to prevent transcription of a *vit-2* reporter construct in males, suggesting that MAB-3 is a direct repressor of vitellogenin transcription. This is the first direct link between the sex determination regulatory pathway and sex-specific structural genes in *C. elegans*, and it suggests that nematodes and insects use at least some of the same mechanisms to control sexual development.

Key words: *Caenorhabditis elegans*, *Drosophila melanogaster*, Sex determination, Vitellogenin, Transcription, MAB-3, DSX

INTRODUCTION

Sexual differentiation is controlled by cascades of regulatory genes. In nematodes and insects, where they have been most extensively characterized, these can be divided into hierarchies composed of three classes of genes (reviewed in Cline and Meyer, 1996). The earliest acting class of genes are coordinately acting switch genes that control both sexual differentiation and dosage compensation, the process by which expression of genes on the sex chromosomes is equalized. Examples of such coordinate regulators include *sdc-2* in *Caenorhabditis elegans* and *sex-lethal* (*Sxl*) in *Drosophila melanogaster* (Cline, 1984; Nusbaum and Meyer, 1989). These coordinate regulators control the activity of a second class of switch genes that function exclusively in sex determination and affect all sexually dimorphic somatic features. Mutations in these genes, for example *tra-1* in *C. elegans* or *transformer* (*tra*) in *Drosophila*, can lead to full sex reversal (Sturtevant, 1945; Hodgkin, 1987). The first two classes of globally acting sex determination genes control the activity of a third class of genes, downstream regulators that direct more limited aspects of sexual development. Examples of downstream sexual regulators include *mab-3* in *C. elegans*, and *fruitless* (*fru*), *dissatisfaction* (*dsf*) and *doublesex* (*dsx*) in *Drosophila* (Baker and Ridge, 1980; Shen and Hodgkin, 1988; Ryner et al., 1996; Finley et al., 1997). Rather than causing complete sex-reversal, mutations in these downstream sexual regulators can cause incorrect sexual differentiation in specific tissues in one or both sexes.

Sex determination, unlike most major developmental processes, is notable for the lack of evolutionary conservation at the molecular level. For example, while alternative splicing plays a central role in *Drosophila* sex determination, there is no indication that this process regulates sex determination in *C. elegans* or mammals (Cline and Meyer, 1996; Schafer and Goodfellow, 1996). Indeed, the only molecular similarity in sex-determining genes that has been observed between phyla is between the *mab-3* gene of *C. elegans* and the *dsx* gene of *Drosophila* (Raymond et al., 1998). These two genes occupy analogous positions in their respective sex determination genetic pathways. Both are downstream regulators that are controlled by the cascade of globally acting switch genes and are required for a subset of sexually dimorphic features. Mammals may use similar genes in sex determination: a gene related to *dsx* and *mab-3* is expressed exclusively in human testis and maps to 9p24.3, a chromosomal region required in two copies for testis development (Raymond et al., 1998).

mab-3 and *dsx* perform similar roles in sexual

differentiation. *mab-3* functions in two male tissues. In the peripheral nervous system of the male tail it directs the development of sensory rays (V rays) that are required for mating, and in the male intestine it prevents expression of vitellogenins (yolk proteins) (Shen and Hodgkin, 1988). *dsx* is alternatively spliced to encode male- and female-specific proteins that act in each sex to drive development away from an intersexual state (Burtis and Baker, 1989). *dsx* controls a wider range of sex-specific features than *mab-3*, but these include sensory bristle formation in the peripheral nervous system and regulation of yolk protein expression in the fat body, features analogous to those controlled by *mab-3* (Baker and Ridge, 1980).

The proteins encoded by *mab-3* and *dsx* are structurally related, as both contain a DNA binding motif called the DM domain (Erdman and Burtis, 1993; Raymond et al., 1998). The DM domain chelates zinc and contains conserved cysteine and histidine residues, but it is otherwise distinct from classical 'zinc fingers' (Erdman et al., 1996). The DM domain was identified in DSX as the minimal DNA binding element of the protein, and it is required for *dsx* function in vivo (Erdman and Burtis, 1993). DSX has one DM domain, while MAB-3 has two. Missense mutations in either of the MAB-3 DM domains eliminate *mab-3* activity, demonstrating that both domains are essential in vivo (Shen and Hodgkin, 1988; Raymond et al., 1998). MAB-3 and DSX are also functionally similar: the protein encoded by the male splice form of the *dsx* mRNA, DSX^M, can replace MAB-3 in the nematode nervous system, restoring the formation of sensory rays to *mab-3* mutant males (Raymond et al., 1998).

The two sexes of *C. elegans* and *Drosophila* are highly dimorphic. For example, in *C. elegans* about 30% of hermaphrodite cells and 40% of male cells are overtly sexually specialized (reviewed in Hodgkin, 1988). From these differences it is evident that the sex determination regulatory pathways must control the sex-specific activity of many structural genes. However, the only structural genes known to be directly regulated by a sex determination pathway are the *Drosophila* yolk protein (*yp*) genes. Transcription of the *yp* genes in the fat body is regulated directly by *dsx* (reviewed by Bownes, 1994). DSX^F acts as an activator of *yp* transcription, while DSX^M represses *yp* transcription (Coschigano and Wensink, 1993; An and Wensink, 1995a,b; Li and Baker, 1998). The targets of *dsx* regulation in other tissues are unknown. Conserved positive regulatory elements, called VPE1 and VPE2 sequences, have been identified in the *C. elegans* vitellogenin promoters (MacMorris et al., 1994). However, the regulatory factors that act via these elements have not been identified, and the control of tissue-, stage- and sex-specific expression of the vitellogenin genes is poorly understood.

Here we investigate the conservation of sex-determining mechanisms by examining the DNA binding and transcriptional regulatory properties of MAB-3. We find that despite containing different numbers of DM domains (two versus one), MAB-3 and DSX bind to similar DNA sequences. The vitellogenin gene promoters in *C. elegans* contain potential MAB-3 binding sites upstream of the start site of transcription. We find that mutating the MAB-3 binding site in the *vit-2* promoter abolishes both binding of MAB-3 in vitro and sex-specific transcriptional regulation of a *vit-2* reporter gene in vivo, suggesting that MAB-3 acts as a direct repressor of vitellogenin transcription. Thus DSX in insects and MAB-3 in nematodes encode DM domain transcription factors that directly regulate expression of yolk protein genes by binding to similar DNA sites, suggesting that at least this aspect of sex determination has been conserved in evolution.

MATERIALS AND METHODS

C. elegans strains and culture

Culture and genetic manipulation of *C. elegans* were performed by standard methods as described previously (Sulston and Hodgkin, 1988). *mab-3* mutants were of phenotype *mab-3;him-5(e1490)*. The *mab-3* alleles used, *e1240* and *mu15*, are null alleles (Shen and Hodgkin, 1988; Raymond et al., 1998). The *him-5* mutation increases the frequency of X chromosome nondisjunction, and thus the proportion of males in a population (Hodgkin et al., 1977).

Plasmids

Plasmids for in vitro transcription are based on pT7NTagplink, which contains a T7 RNA polymerase promoter, the 5′ untranslated region of the human beta globin gene, and a c-myc epitope tag (Bardwell and Treisman, 1994). The MAB-3 expression plasmid pDZ130 contains the MAB-3 coding region as a *Eco*RI/*Xba*I fragment generated by PCR from *C. elegans* cDNA. The resulting protein has a c-myc epitope tag (MEQKLISEEDLNM) followed by the sequence AGSEF and the full MAB-3 coding region, minus the initiation methionine. The protein coding sequences in pDZ130 were confirmed by DNA sequencing. The DSX^F expression vector pDZ134 was made by inserting a *Eco*RI(filled)/*Sal*I fragment of the DSX^F cDNA (gift of K. Burtis) into *Bam*HI(filled)/*Xho*I cut pT7plink. The resulting protein contains the c-myc epitope tag followed by the sequence GSIPSEAGI and the full DSX^F coding sequence. MAB-3 and DSX proteins were generated by coupled transcription/translation in vitro with the TNT T7 coupled reticulocyte lysate system (Promega) according to the manufacturer's instructions.

The *vit-2*::GFP reporter pCR2 was made by inserting a 366 bp *Xba*I/*Sac*I(filled) fragment containing the 247 bp minimal *vit-2* promoter from pFPO-gfp-V3′(FUB1) (gift of M. MacMorris and T. Blumenthal) into *Xba*I/*Sma*I cut pPD95.70 (gift of A. Fire). The resulting plasmid has the minimal 247 bp *vit-2* promoter followed by the first 14 amino acids of *vit-2*, a synthetic intron and the GFP coding sequence. Mutant reporters M1 and M2 were generated from the wild-type *vit-2*::GFP reporter pCR2 using the GeneEditor™ mutagenesis system (Promega). Both reporters are highly expressed in hermaphrodites (not shown), confirming that *vit-2* promoter strength has not been impaired by the mutations. Mutagenic primers were the following: WY22, 5′ GGCTCTCACCGAATGGTTCAATTTGTTT-CTGAT 3′ (for *vit-2* M1), WY42, 5′ CAAAATTAATAGAC-AGGGCTCTCACCGAAGAGTTAAATTTGTTTCTGATAAGGGT-CACAAAGCG 3′ (for *vit-2* M2).

Oligonucleotides

Full sequences of double stranded oligonucleotides used for DNA binding experiments in Fig. 3 were as follows (one strand is shown):
Probe A: 5′ AATTCACAACTACAATGTTGCAATCAGCTAGCC 3′;
Probe B: 5′ AATTCTCTCACCGAATGTTGCAATTTGTTTCTC 3′;
Probe C: 5′ AATTCTCTCACACAATGTTGCAATTTGTTTCTC 3′;
Probe D: 5′ AATTCTCTCACCGAATGTTGCGATTTGTTTCTC 3′;
Probe E: 5′ AATTCTCTCACCGAATGGTTCAATTTGTTTCTC 3′;
Probe F: 5′ AATTCTCTCACCGAATATTGCAATTTGTTTCTC 3′.

Reporter gene analysis

Transgenic nematodes were generated by standard methods (Mello et al., 1991), by coinjecting the reporter plasmid and the dominant

transformation marker pRF4 (containing the mutant *rol-6* allele *su1006*) each at 100 ng/μl. Reporters were analyzed in lines containing heritable extrachromosomal DNA arrays, and each construct was assayed in at least five independent transgenic lines.

DNA binding site selection and gel-mobility shift assays

Binding site selection from random oligonucleotides was performed as described previously (Pollock and Treisman, 1990) except for the following changes. After in vitro translation of proteins, $ZnSO_4$ was added to 50 mM. For gel-mobility shift assays in Fig. 3, probes were made by annealing oligonucleotides and filling overhanging ends with Klenow DNA polymerase in the presence of [α-^{32}P]dATP, and unincorporated nucleotides were removed by chromatography through Sephadex G-50 resin. For other gel-mobility shift assays, probes were labeled with [α-^{32}P]dATP during PCR amplification as described (Pollock and Treisman, 1990). DSX^F was used in all experiments; DSX^F and DSX^M have identical DNA binding properties (Erdman and Burtis, 1993; Cho and Wensink, 1996, 1997; Erdman et al., 1996).

Sequence analysis of selected sites

Cloned oligonucleotides from site selections were sequenced using the Thermosequenase labeled terminator cycle sequencing kit (Amersham Life Science) according to the manufacturer's instructions. Sequences were aligned manually to derive consensus binding sequences. The potential MAB-3 binding sites in the vitellogenin promoters were identified using the MacVector program, allowing 4-bp mismatches. *C. briggsae vit* gene promoter sequences are described by Zucker-Aprison and Blumenthal (1989).

C. briggsae MAB-3

The *C. briggsae mab*-3 gene was identified by searching the *C. briggsae* DNA database with WU-BLAST 2.0 (http://www.isrec.isb-sib.ch/software/WUBLAST_form.html) using the *C. elegans mab*-3 sequence. The *C. briggsae* MAB-3 sequence was predicted by conceptual splicing of the genomic sequence with the splice sites used by the *C. elegans* gene, all of which are conserved. Database searches of the Wormpep database (http://www.sanger.ac.uk/Projects/C_elegans/blast_server.shtml) with the BLASTP program identified the following ten apparent DM domain-containing sequences in *C. elegans*: Y53C12B.5a (*mab-3*), F10C1.5, C34D1.1, C34D1.2, T22H9.4 (two DM domains), Y43F8A.G, F13G11.1, C27C12.6, Y67D8A_390.A and K08B12.2.

RESULTS

mab-3 prevents vitellogenin gene transcription in males

Vitellogenins are synthesized in the adult hermaphrodite intestine, exported into the body cavity, taken up by the ovary, and incorporated into maturing oocytes (Kimble and Sharrock, 1983). A short (247 bp) segment of the *vit-2* promoter is sufficient to recapitulate the tissue-, stage- and sex-specific expression of the intact *vit-2* gene, suggesting that the primary control of vitellogenin expression is by transcriptional regulation (MacMorris et al., 1992). Loss-of-function mutations in *mab-3* abrogate the sexual regulation of vitellogenin expression, causing the vitellogenins to be expressed abundantly in the adult male intestine (Shen and Hodgkin, 1988).

To determine whether the deregulation of vitellogenin expression in *mab-3* mutants results from a failure of transcriptional control, we tested the expression of a transgene containing 247 bp of *vit-2* promoter sequence fused to the green fluorescent protein (GFP) coding region. The *vit-2*::GFP reporter is expressed in wild-type hermaphrodites, but not in wild-type males (Fig. 1, top panels). In contrast, *mab-3* mutants express the *vit-2*::GFP reporter at high levels in both sexes (Fig. 1, bottom panels). Eliminating *mab-3* activity abolishes the sex-specificity of *vit* promoter activity, but stage-specificity and tissue-specificity are unaffected (Fig. 1, and data not shown). The misexpression of *vit-2*::GFP in *mab-3* mutant males suggests that *mab-3* regulates vitellogenin transcription, either directly or indirectly. To test whether the regulation is direct, we first investigated the DNA binding specificity of MAB-3.

MAB-3 and DSX bind to similar sites

We found previously that the male-specific isoform of DSX, DSX^M, can restore V rays to *mab-3* mutant males almost as well as MAB-3 (Raymond et al., 1998). This suggests that, despite containing different numbers of DM domains, DSX and MAB-3 can bind to similar DNA sequences in vivo and can regulate some of the same genes. To test this possibility more directly, we determined the preferred MAB-3 in vitro DNA binding site. As a control, we also determined the preferred binding site of DSX (Erdman et al., 1996).

To identify high-affinity DNA binding sites of MAB-3 and DSX, we used a selection method in which oligonucleotides containing functional binding sites are isolated from a large pool of random oligonucleotides by immunoprecipitation of protein/DNA complexes (Pollock and Treisman, 1990). The proteins used for selection were made by in vitro translation and contained a short amino-terminal c-myc sequence recognized by the 9E10 monoclonal antibody. Four rounds of DNA binding and immunoprecipitation with 9E10 were performed with each protein. Pools of oligonucleotides from the fourth round of selection were tested for binding in gel-mobility shift assays, and the shifted oligonucleotides were cloned and sequenced. DNA sequences of the selected oligonucleotides were compared to identify potential binding site preferences.

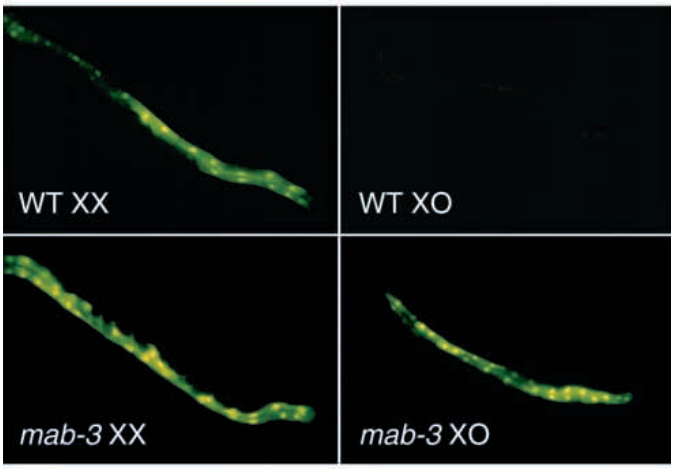

Fig. 1. *mab-3* prevents vitellogenin gene transcription in the XO intestine. Expression of a *vit-2*::GFP reporter gene containing 247 bp of *vit-2* promoter sequence fused to the GFP coding region in the intestine of wild-type adult XX hermaphrodites (upper left), wild-type XO males (upper right), *mab-3* mutant XX hermaphrodites (lower left) and *mab-3* mutant XO males (lower right). Anterior is to the left, and all exposures are equivalent.

A MAB-3 binding site consensus

base (%)	position 1	2	3	4	5	6	7	8	9	10	11	12	13
T	8	14	99	0	100	100	0	0	1	27	44	36	74
C	3	0	1	0	0	0	0	100	19	0	0	13	8
G	13	0	0	100	0	0	100	0	67	11	6	22	10
A	76	86	0	0	0	0	0	0	13	62	50	29	8
consensus:	A	A	T	G	T	T	G	C	G	A	T/A	N	T

B DSX binding site consensus

base (%)	position -6	-5	-4	-3	-2	-1	0	+1	+2	+3	+4	+5	+6
T	6	19	59	0	0	10	0	70	0	98	31	31	21
C	26	50	5	0	100	0	0	2	0	0	7	3	52
G	34	0	5	0	0	0	0	2	100	0	7	40	15
A	34	31	31	100	0	90	100	26	0	2	55	26	12
consensus:	nT	nG	T/A	A	C	A	A	T	G	T	A/T	nC	C

Fig. 2. Summary of DNA binding site selections. Selected oligonucleotides were sequenced and aligned to identify potential binding site sequences. The occurrence of each base at each position of the aligned sequences is shown (as a percentage). (A) Consensus table from 63 oligonucleotide sequences from the fourth round of selection with MAB-3, isolated from the shifted band in a gel-mobility shift assay. (B) Consensus table from 58 oligonucleotide sequences from the fourth round of selection with DSX^F isolated from the shifted band in a gel-mobility shift assay. T/A = T or A; nC = T, G or A (i.e. not C); N, no consensus.

The binding site consensus sequences derived from the selected oligonucleotides are summarized in Fig. 2. 63 of 65 sequenced oligonucleotides bound by MAB-3 contained a close match to the 13-nucleotide consensus AATGTTGCGA(T/A)NT, where N is any nucleotide. Most positions showed a very strong preference for one or two bases. The preferences at positions 3-8 are particularly strong, suggesting that these positions may be critical for high-affinity binding. All of 58 sequenced oligonucleotides bound by DSX contained a close match to the 13-nucleotide consensus (nT)(nG)(T/A)ACAATGT(A/T)(nC)C (nT = not T). The central nine base pairs form a punctuated palindrome around the central base pair, consistent with binding of DSX as a dimer to two half sites (Erdman et al., 1996). Our DSX site selection used full-length protein made by in vitro translation, but the consensus derived here is very similar to the 13-nucleotide consensus (G/A)NNAC(A/T)A(T/A)GTNN(C/T) reported previously for the bacterially expressed DM domain of DSX using an affinity chromatography method (Erdman et al., 1996). The most strongly preferred core nucleotides (ACAATGT) in the DSX site selected here and the previously reported site are identical. There are minor differences in flanking nucleotide preferences, possibly due to the use of full-length versus truncated protein or to differences in stringency between the two selection methods.

Although MAB-3 has two DM domains and DSX has only one, the DNA binding sites selected by the two proteins do bear some similarity. MAB-3 selected an asymmetrical site, while DSX prefers an inverted repeat (Erdman et al., 1996), but the sites selected by the two proteins are similar at eight base pairs (Fig. 3A). As a further test of the similarity of the MAB-3 and DSX binding sites, we assayed binding to several short double stranded oligonucleotides (Fig. 3B). These included a DSX binding site from the *yp1* promoter (probe A), a potential MAB-3 site from the *vit-2* promoter (probe B, see below), and several oligonucleotides based on the binding site consensus sequences described above (probes C-F).

A DSX binding site from the *Drosophila yp1* gene (DSX A site) (Coschigano and Wensink, 1993) contains both MAB-3 and DSX consensus binding sequences, and both proteins bind efficiently (Fig. 3C, lanes 1 and 2). The *C. elegans vit-2* promoter contains a potential MAB-3 binding site (see below), and this sequence is bound efficiently by MAB-3 (lane 3). However, the site in *vit-2* lacks an AC dinucleotide that is preferred by DSX, and is bound inefficiently by DSX (lane 4). As predicted by the consensus sequences, adding the AC dinucleotide results in efficient binding by both proteins (lanes 5 and 6). The potential MAB-3 site in the *vit-2* promoter differs at position 9 from an ideal MAB-3 consensus site (containing an A, rather than the preferred G), but optimizing this position to G has little effect (lane 7). This is consistent with the relatively weak preference for a G at that position (Fig. 2). At

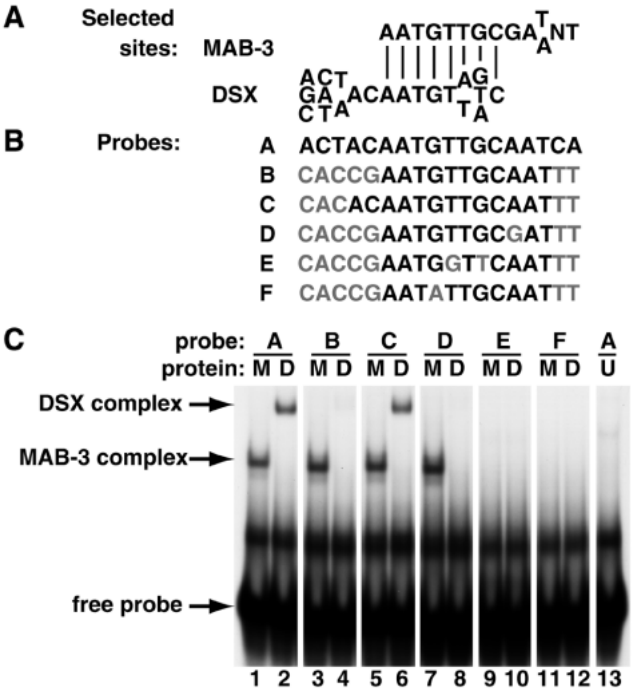

Fig. 3. MAB-3 and DSX bind similar DNA sequences. (A) Comparison of MAB-3 and DSX DNA consensus DNA binding sites. (B) Probe sequences used for gel-mobility shifts. The portion of the oligonucleotides containing potential MAB-3 and DSX binding sites is shown (the full 33-bp oligonucleotide sequences are in Materials and methods). Probe A is the DSX A site from *D. melanogaster yp1* promoter. Probe B is a potential MAB-3 binding site from *C. elegans vit-2* promoter. Probe C is the same site with a CG changed to AC to better match the DSX binding consensus. Probe D is the ideal MAB-3 binding site derived from site selection. Probes E and F are the same as probe B, but with changes in nucleotides that were absolutely preferred in the site selection. Mismatches relative to probe A are shown in gray. (C) Gel-mobility shift assay of oligonucleotides shown in B. Oligonucleotides were incubated with rabbit reticulocyte lysate containing MAB-3 (M), DSX (D) or unprogrammed lysate (U) and complexes were separated by native gel electrophoresis. The positions of free probe and shifted complexes containing MAB-3 and DSX are indicated.

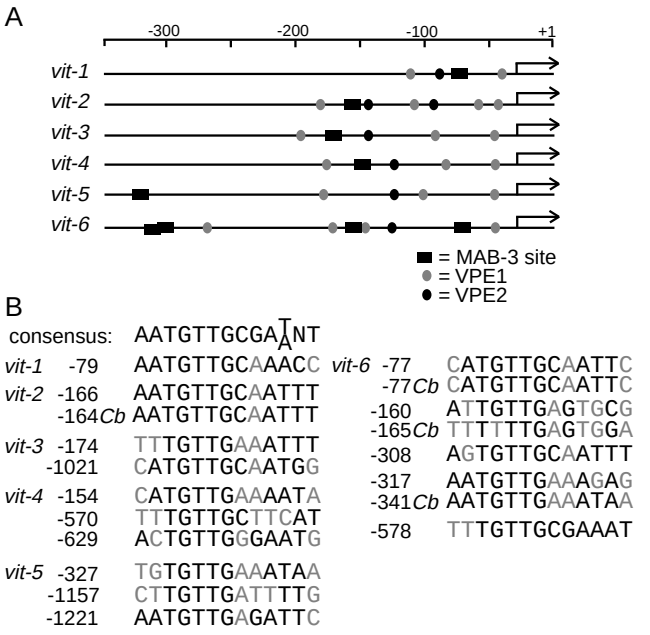

Fig. 4. (A) Vitellogenin gene promoters contain potential MAB-3 binding sites. The promoter transcription start sites are indicated by arrows and regulatory sequences by symbols. Potential regulatory elements are shown as follows: solid rectangles, MAB-3 sites; shaded ovals, VPE1 sequences; solid ovals, VPE2 sequences. VPE1 (TGTCAAT) and VPE2 (CTGATAA) elements shown are those that are conserved in position and orientation between *C. elegans* and *C. briggsae*, allowing a 1 bp mismatch (Zucker-Aprison and Blumenthal, 1989). MAB-3 binding sites with at least an 8/13 match to the consensus AATGTTGCGA(T/A)NT are shown. (B) MAB-3 binding sites in *vit* gene promoters. Binding sites identified in the *vit* gene promoters are shown. The position of the 5′ nucleotide in each site relative to start of transcription is indicated. Nucleotides that match the MAB-3 DNA binding site consensus are in black; mismatched positions are shaded. Sites that are conserved in *C. briggsae vit* genes are indicated (*Cb*). A *C. briggsae* sequence sufficient to compare with all *C. elegans* MAB-3 sites was not available for *vit-4*, *vit-5* and *vit-6*, and none was available for *vit-3*.

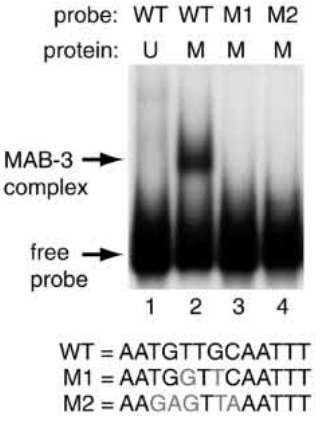

Fig. 5. MAB-3 binds the *vit-2* promoter in vitro. Gel-mobility shift assay with 247 bp *vit-2* promoter probe. Lanes 1 and 2: wild-type *vit-2* sequence. Lane 3: MAB-3 site mutated from AATGTTGCAATTT to AATGGTTCAATTT (M1). Lane 4: MAB-3 site mutated from AATGTTGCAATTT to AAGAGTTAAATTT (M2). Probe was incubated with rabbit reticulocyte lysate programmed with MAB-3 (M; lanes 2-4) or with unprogrammed lysate (U; lane 1). Positions of free probe and MAB-3/DNA complex are indicated by arrows. Sequences of MAB-3 binding sites are shown below. Mutated base pairs are indicated in gray.

other positions a particular nucleotide is absolutely preferred, and changing one or two of the nucleotides at those positions can severely reduce binding by MAB-3 (Fig. 3C, lanes 9 to 12). Taken together, these results confirm that the binding site selections provided optimal consensus DNA binding sites, as well as accurate indications of which positions are most important for efficient binding in vitro. They also confirm that MAB-3 and DSX can bind to very similar short DNA sequences.

vit gene promoters contain potential MAB-3 binding sites

To investigate the role of MAB-3 in regulating vitellogenin transcription, we searched for potential MAB-3 binding sites in the proximal promoters of the *C. elegans vit* genes. In all six *vit* genes, possible MAB-3 binding sites occur within the proximal 350 bp of the promoter (Fig. 4). Five of the *vit* genes have single proximal MAB-3 consensus binding sites, while *vit-6* has four sites. These sites range from 8/13 to 12/13 bp matches to the in vitro consensus DNA binding site (Materials and methods). Most of the sites do not contain a G residue at position 9, but as shown above this position is not critical for DNA binding (Fig. 3). The most proximal sites in *vit-3*, *vit-4* and *vit-5* are poorer matches to the consensus than those in *vit-1*, *vit-2* and *vit-6*. However, *vit-3*, *vit-4*, *vit-5* and *vit-6* contain additional sequences further upstream that more closely match the MAB-3 binding site consensus (Fig. 4B). Since MAB-3 is conserved between *C. elegans* and the related nematode *C. briggsae* (see below), functional MAB-3 binding sites might also be expected to be evolutionarily conserved between these two nematodes. Complete promoter sequences for the all of the *C. briggsae vit* genes are not available, so a full comparison is not currently possible. Nevertheless, the *C. briggsae vit-2* and *vit-6* homologs have potential MAB-3 binding sites in positions and orientations equivalent to those in *C. elegans* (Fig. 4B). In summary, searches for MAB-3 binding sites reveal that all of the *C. elegans vit* genes are candidates for direct transcriptional regulation by MAB-3.

MAB-3 is a direct transcriptional repressor of vitellogenin transcription

We next tested the functional importance of the MAB-3 binding site in the *vit-2* promoter. The *vit-2* promoter was chosen because it is the best characterized of the *vit* promoters and contains only one potential MAB-3 binding site near the transcriptional start site. MAB-3 binds the minimal 247 bp *vit-2* promoter in vitro in a gel-mobility shift assay (Fig. 5, lanes 1 and 2). To test whether MAB-3 binds the site identified by sequence analysis, we made two mutant promoters, *vit-2* M1 and *vit-2* M2, which have two and five nucleotide changes in the potential MAB-3 binding site, respectively. MAB-3 does not bind to either of the mutant promoters (Fig. 5, lanes 3 and 4), demonstrating that the site is required for MAB-3 binding in vitro.

To test the in vivo relevance of the MAB-3 binding site in

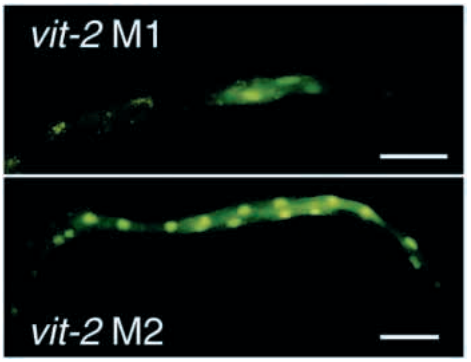

Fig. 6. Repression of *vit-2* transcription in vivo requires a MAB-3 binding site. Mutant *vit-2*::GFP reporters with the mutations shown in Fig. 5 were constructed and tested in *him-8* transgenic XO males. (Top) Adult male with extrachromosomal array containing the *vit-2* M1 reporter (AATGGTTCAATTT) expresses GFP in four posterior intestinal cells. (Bottom) Adult male with extrachromosomal array containing the *vit-2* M2 reporter (AAGAGTTAAATTT) expresses GFP in nearly all intestinal cells. Anterior is to left. Bars, 100 μm.

the *vit-2* promoter, we constructed *vit-2::*GFP reporters containing the two mutant *vit-2* promoters and assayed their expression in transgenic worms relative to the wild-type promoter. Wild-type males do not express the wild-type *vit-2* reporter (Fig. 1). In contrast, both of the mutant reporters are expressed in the wild-type male intestine (Fig. 6). The M1 reporter is expressed at relatively low levels in males, mainly in the posterior intestinal cells, but at high levels in hermaphrodites, suggesting that it may retain some MAB-3 binding in vivo (Fig. 6, top panel, and data not shown). The M2 reporter, with five nucleotide changes, is highly expressed in males throughout the intestine (Fig. 6, bottom panel). Neither mutation affects the tissue specificity or stage specificity of *vit-2* reporter gene expression (not shown). Disrupting the MAB-3 binding site in the minimal *vit-2* promoter results in a loss of sex specificity of *vit-2* transcription, the same effect as eliminating *mab-3* function. We conclude that MAB-3 is a direct transcriptional repressor of vitellogenins in the male intestine.

MAB-3 is more highly conserved than upstream sexual regulators

Sex determination pathways appear to evolve rapidly compared to other major developmental regulatory pathways, and there is some evidence that upstream sex-determining genes may evolve more rapidly than downstream genes (reviewed in Marin and Baker, 1998). The conservation between nematodes and insects of MAB-3 and DSX, not only in sequence, but also DNA binding and function, suggests that these two sex-determining genes are ancient (Raymond et al., 1998). If so,

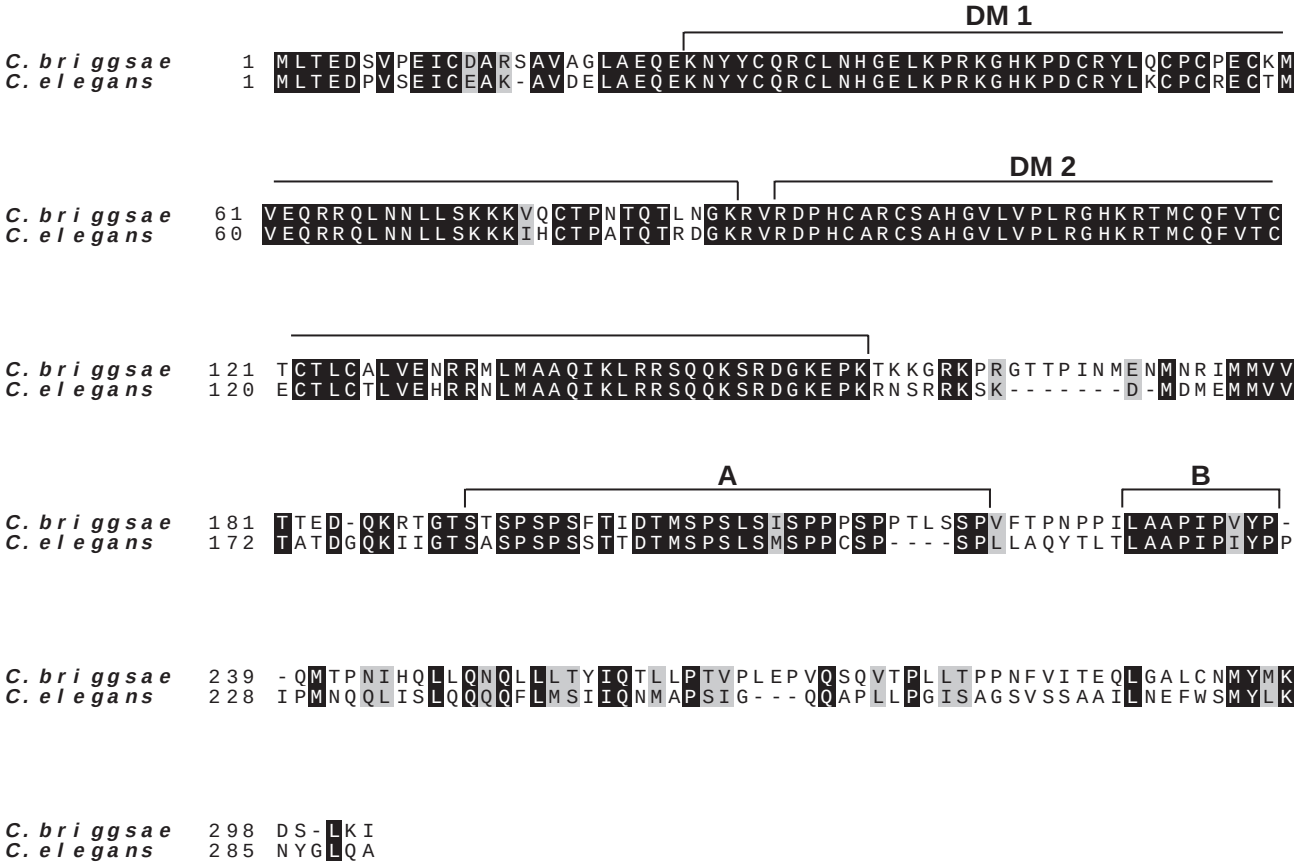

Fig. 7. Comparision of *C. elegans* and *C. briggsae* MAB-3 sequences. Sequences were aligned using the program Clustal (Higgins et al., 1996). Identical amino acids are shaded in black and similar amino acids are shaded in gray. Four highly conserved regions are labeled, including the two DM domains (DM1 and DM2), a proline- and serine-rich region ('A'), and another short highly conserved region ('B'). Protein sequences were derived from *C. elegans* cDNA and genomic DNA sequence and from *C. briggsae* genomic DNA sequence (Materials and methods).

then *mab-3* might be expected to be more highly conserved between nematode species than the upstream sex-determining genes. In addition, by comparing related MAB-3 genes it might be possible to identify conserved functional domains of the proteins.

We compared the *C. elegans* MAB-3 protein to that of the related nematode *C. briggsae*. Fig. 7 shows an alignment of the *C. elegans* protein sequence, *Ce*-MAB-3, predicted from the *mab-3* cDNA sequence, with the *C. briggsae* protein sequence, *Cb*-MAB-3, predicted from a genomic DNA sequence identified by database searching. *C. elegans* contains ten predicted DM domain-containing genes (Materials and methods), but only *mab-3* and one other gene (T22H9.4) have two DM domains, and there is little conservation between *mab-3* and the other *C. elegans* DM domain genes outside the DM domains (Raymond et al., 1998; W. Y. and C. Raymond, data not shown). From the presence of two highly similar DM domains, the similarity of other regions of the protein, and the apparent conservation of all splice sites between the two genes (data not shown), we conclude that the *C. briggsae* gene we identified is the homolog of *C. elegans mab-3*.

Overall the two predicted MAB-3 proteins are 64% identical. Several regions of the proteins are particularly similar. These include the two DM domains, which are 91% identical between Ce-MAB-3 and Cb-MAB-3, as well as two other regions, one of which is proline- and serine-rich (Fig. 7). These domains may have conserved functions such as mediating transcriptional repression or protein-protein interaction. The DSX proteins also contain a proline- and serine-rich region (Burtis and Baker, 1989), raising the possibility that these domains are a more widely conserved feature. MAB-3 is more similar between these two nematode species than are the upstream sexual regulators TRA-1 and TRA-2, which are 44% and 43% identical, respectively (de Bono and Hodgkin, 1996; Kuwabara, 1996). MAB-3 also is slightly more conserved between these nematode species than the upstream regulator FEM-2, which has a region that is 63% identical between species and an amino-terminal domain that is present only in the *C. briggsae* protein (Hansen and Pilgrim, 1998). Thus our data support the view that downstream sex-determining genes evolve more slowly than upstream genes.

DISCUSSION

Conservation of sex-determining mechanisms in worms and flies

Characterization of sex-determining genes from several species has revealed almost no molecular similarity between phyla, but the observation that *mab-3* and *dsx* encode proteins with a related motif, the DM domain, suggests that some evolutionary conservation may exist. Here we have described additional evidence that at least some features of sex determination in nematodes and insects are conserved. We have found that MAB-3 and DSX bind to similar DNA sequences, that *C. elegans vit* genes have potential MAB-3 binding sites in their promoter regions, and that a MAB-3 binding site is essential for male-specific repression of *vit-2* transcription. This strongly suggests that MAB-3 is a direct regulator of *vit-2* and possibly of other *vit* genes. Thus MAB-3 and DSX not only have related protein sequences, but they both directly control yolk protein gene transcription by binding to similar DNA sequences.

Although evidence is limited, upstream sexual regulators appear to diverge more rapidly than downstream regulators; for example, the upstream regulators *tra-1* and *tra-2* are among the most rapidly diverging nematode genes (de Bono and Hodgkin, 1996; Kuwabara, 1996). Similarly, in dipterans, the upstream regulator *Sxl* appears to have become involved in sex determination more recently than downstream genes like *dsx* (reviewed by Marin and Baker, 1998). The reasons for higher conservation of downstream genes are unclear, but among the proposed factors are the potentially higher pleiotropy of downstream genes due to a larger number of affected biochemical pathways (Waxman and Peck, 1998), and relatively recent recruitment of upstream genes to optimize regulation of more ancient downstream genes (Wilkins, 1995). *mab-3* and *dsx* both control sex-specific peripheral sensory organ formation, and ectopic DSX^M can at least partially replace MAB-3, restoring male sensory ray formation to a *mab-3* mutant (Raymond et al., 1998). Thus the two proteins can perform similar functions in vivo. Based on these results and the relatively high conservation of MAB-3 between nematode species, we speculate that the interface between sex-determining regulatory pathways and sex-specific structural genes may be more highly conserved than other aspects of sex determination.

DNA binding by MAB-3 and DSX

We have found that MAB-3 binds to DNA sequences similar to those bound by the DSX proteins of *Drosophila*. This result is striking for two reasons. First, while both proteins contain the DM domain, MAB-3 has two DM domains and DSX has only one. Second, while the yolk proteins of nematodes and fruitflies are functionally analogous, they are molecularly unrelated (Wahli, 1988; Spieth et al., 1991). It therefore appears from our results that the regulatory proteins have been conserved, but entirely different structural gene families have been selected. We speculate that the pleiotropy of MAB-3 and DSX, which need to regulate multiple aspects of sexual development, may constrain their evolution, while yolk proteins may be less constrained.

Consistent with the symmetrical nature of the DSX DNA binding site, DSX proteins have been shown to bind DNA as dimers. Two dimerization domains have been identified in DSX, and the proteins have been shown to dimerize both in vitro on DNA, and in vivo (An et al., 1996; Erdman et al., 1996). The MAB-3 DNA binding site is asymmetrical, suggesting that MAB-3 may bind DNA as a monomer, although this has not been shown. The similar sizes of the MAB-3 and DSX DNA binding sites are consistent with this view, as is the requirement for both DM domains for *mab-3* activity in vivo. Thus both DSX and MAB-3 might form complexes with DNA using two DM domains, with DSX doing so by forming dimers and MAB-3 binding as a monomer.

The role of *mab-3* in vitellogenin regulation

We have found that repression of *vit-2* transcription in the XO intestine requires both functional MAB-3 and an intact MAB-3 binding site in the *vit-2* promoter. We propose that MAB-3 is a direct transcriptional repressor of *vit-2*. What is the role of *mab-3* in regulating other *vit* genes? Since null mutations in

mab-3 deregulate all the vitellogenins, *mab-3* must play a role in regulating all six *vit* genes. The simplest model is that MAB-3 directly represses transcription of the *vit* genes. Supporting this possibility, we have found potential MAB-3 binding sites upstream of the transcriptional initiation sites of all six *vit* genes, suggesting that *mab-3* may regulate these genes directly.

The potential MAB-3 sites in the *vit* promoters are not perfect matches to the in vitro MAB-3 binding consensus or to the MAB-3 binding site in the *vit-2* promoter (Fig. 4), but the requirements in vivo are likely to be less stringent. In DNA binding assays in vitro, a single nucleotide change is sufficient to virtually eliminate binding of *mab-3* to an oligonucleotide (e.g. Fig. 3) and mutating two nucleotides in the *vit-2* promoter (probe E, Fig. 3C; *vit-2* M1, Fig. 5) completely eliminates binding. However, the *vit-2* M1 reporter is not fully deregulated in vivo, suggesting that MAB-3 is still capable of binding to the mutant promoter and partially repressing its transcription. One possibility is that DNA binding in the in vitro assay requires higher DNA/protein affinity than is required in vivo for regulation. A second possibility is that MAB-3 may bind cooperatively with other factors in vivo, and these factors may facilitate the efficient binding of MAB-3 to a less than optimal site. In either case it appears that the sequence requirements for regulation by MAB-3 in vivo are less strict than for MAB-3 DNA binding in vitro, and so MAB-3 might directly regulate all of the *vit* genes. Alternatively, it is possible that *mab-3* regulates some of the *vit* genes indirectly.

Sexual regulation of vitellogenin expression probably involves more than just repression of *vit* gene transcription by MAB-3 in males and repression of *mab-3* activity by *tra-1* in hermaphrodites. Genetic experiments suggest that *tra-1* promotes vitellogenin expression not only by negatively regulating *mab-3*, but also by a *mab-3*-independent mechanism (Shen and Hodgkin, 1988). If *tra-1* activity promoted *vit* gene expression solely by repressing *mab-3* activity, then animals deficient in *mab-3* would no longer require *tra-1* for high-level *vit* gene expression. This is not the case. Rather, while *mab-3(null)* XX animals express vitellogenins abundantly, *tra-1;mab-3(null)* XX animals express reduced levels of vitellogenins, with the amount of vitellogenin expression depending on the the *tra-1* allele tested. When the *tra-1* allele is the null allele *e1099*, *tra-1;mab-3(null)* XX animals accumulate reduced levels of vitellogenins relative to *mab-3(null)* XX animals (Shen and Hodgkin, 1988). These results suggest that *tra-1* can promote *vit* gene expression not only by repressing *mab-3* activity, but also through another mechanism that does not require *mab-3*. It is unclear whether this *mab-3*-independent activation of *vit* expression by *tra-1* is direct or indirect, but we have not found potential TRA-1 binding sites in any of the *vit* gene promoters, and TRA-1A does not bind the minimal *vit-2* promoter in vitro (data not shown). We speculate that one or more of the other nine *C. elegans* DM domain proteins might also have a role in regulating vitellogenin expression, possibly under the control of *tra-1*.

How conserved is sex determination?

From the experiments presented here and from previous work (Sievert et al., 1997; Raymond et al., 1998; Shearman and Frommer, 1998), the downstream genes *mab-3* and *dsx* do appear likely to be ancient sex-determining genes. At present it is unclear whether this is an isolated example or whether other downstream sexual regulators also are conserved in evolution. Intriguingly, mammals have DM domain-containing genes whose expression and chromosomal locations are consistent with a role in sex determination (Raymond et al., 1998; C. Raymond and D. Z., unpublished data), but this possibility has not been directly tested. Several other downstream regulators have been identified in *Drosophila* (Finley et al., 1997; Marin and Baker, 1998), but little is known in *C. elegans* about sexual regulators other than *mab-3* that act downstream of *tra-1*. Such genes must exist, since *mab-3* mutations only affect a fraction of sexually dimorphic cell lineages (Shen and Hodgkin, 1988; Hodgkin et al., 1989). It will be important to establish whether any other features of sexual dimorphism are controlled by conserved downstream regulatory genes.

We thank colleagues in the Developmental Biology Center of the University of Minnesota for helpful discussions, M. MacMorris, T. Blumenthal, R. Treisman, K. Burtis and A. Fire for clones and vectors and Christopher Raymond for clones and assistance. We thank Julie Ahringer, Vivian Bardwell, Jonathan Hodgkin and Ann Rougvie for critical reading of the manuscript. This work was supported by grants from the Minnesota Medical Foundation, University of Minnesota Graduate School and the NIH (GM53099).

REFERENCES

An, W., Cho, S., Ishii, H. and Wensink, P. C. (1996). Sex-specific and non-sex-specific oligomerization domains in both the doublesex transcription factors from *Drosophila melanogaster*. *Mol. Cell. Biol.* **16**, 3106-3111.

An, W. and Wensink, P. C. (1995a). Integrating sex- and tissue-specific regulation within a single *Drosophila* enhancer. *Genes Dev.* **9**, 256-266.

An, W. and Wensink, P. C. (1995b). Three protein binding sites form an enhancer that regulates sex- and fat body-specific transcription of *Drosophila yolk protein* genes. *EMBO J.* **14**, 1221-1230.

Baker, B. S. and Ridge, K. (1980). Sex and the single cell: on the action of major loci affecting sex determination in *Drosophila melanogaster*. *Genetics* **94**, 383-423.

Bardwell, V. J. and Treisman, R. (1994). The POZ domain: a conserved protein-protein interaction motif. *Genes Dev.* **8**, 1664-1677.

Bownes, M. (1994). The regulation of the *yolk protein* genes, a family of sex differentiation genes in *Drosophila melanogaster*. *BioEssays* **16**, 745-752.

Burtis, K. C. and Baker, B. S. (1989). *Drosophila* doublesex gene controls somatic sexual differentiation by producing alternatively spliced mRNAs encoding related sex-specific polypeptides. *Cell* **56**, 997-1010.

Cho, S. and Wensink, P. C. (1996). Purification and physical properties of the male and female doublesex proteins of *Drosophila*. *Proc. Natl. Acad. Sci. USA* **99**, 2043-2047.

Cho, S. and Wensink, P. C. (1997). DNA binding by the male and female doublesex proteins of *Drosophila melanogaster*. *J. Biol. Chem.* **272**, 3185-3189.

Cline, T. W. (1984). Autoregulatory functioning of a *Drosophila* gene product that establishes and maintains the sexually determined state. *Genetics* **107**, 231-237.

Cline, T. W. and Meyer, B. J. (1996). Vive la difference: males vs females in flies vs worms. *Annu. Rev. Genet.* **30**, 637-702.

Coschigano, K. T. and Wensink, P. C. (1993). Sex-specific transcriptional regulation by male and female doublesex proteins of *Drosophila*. *Genes Dev.* **7**, 42-54.

de Bono, M. and Hodgkin, J. (1996). Evolution of sex determination in *Caenorhabditis*: unusually high divergence of *tra-1* and its functional consequences. *Genetics* **144**, 587-595.

Erdman, S. E. and Burtis, K. C. (1993). The *Drosophila* doublesex proteins share a novel zinc finger related DNA binding domain. *EMBO J.* **12**, 527-535.

Erdman, S. E., Chen, H.-J. and Burtis, K. C. (1996). Functional and genetic characterization of the oligomerization and DNA binding properties of the *Drosophila* Doublesex proteins. *Genetics* **144**, 1639-1652.

Finley, K. D., Taylor, B. J., Milstein, M. and McKeown, M. (1997). *dissatisfaction*, a gene involved in sex-specific behavior and neural development of *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA* **94**, 913-918.

Hansen, D. and Pilgrim, D. (1998). Molecular evolution of a sex determination protein. FEM-2 (pp2c) in Caenorhabditis. *Genetics* **149**, 1353-62.

Higgins, D. G., Thompson, J. D. and Gibson, T. J. (1996). Using CLUSTAL for multiple sequence alignments. *Methods Enzymol.* **266**, 383-402.

Hodgkin, J. (1987). A genetic analysis of the sex-determining gene, *tra-1*, in the nematode *Caenorhabditis elegans*. *Genes Dev.* **1**, 731-745.

Hodgkin, J. (1988). Sexual dimorphism and sex determination. In *The Nematode* Caenorhabditis elegans (ed. W. B. Wood), pp. 243-279. Cold Spring Harbor, New York; Cold Spring Harbor Laboratory**.**

Hodgkin, J., Chisholm, A. D. and Shen, M. M. (1989). Major sex-determining genes and the control of sexual dimorphism in *Caenorhabditis elegans*. *Genome* **31**, 625-637.

Hodgkin, J., Horvitz, H. R. and Brenner, S. (1977). Nondisjunction mutants of the nematode *Caenorhabditis elegans*. *Genetics* **91**, 67-94.

Kimble, J. and Sharrock, W. J. (1983). Tissue-specific synthesis of yolk proteins in *Caenorhabditis elegans*. *Dev. Biol.* **96**, 189-196.

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

Li, H. and Baker, B. S. (1998). *hermaphrodite* and *doublesex* function both dependently and independently to control various aspects of sexual differentiation in *Drosophila*. *Development* **125**, 2641-2651.

MacMorris, M., Broverman, S., Greenspoon, S., Lea, K., Madej, C., Blumenthal, T. and Spieth, J. (1992). Regulation of vitellogenin gene expression in transgenic *Caenorhabditis elegans*: short sequences required for activation of the *vit-2* promoter. *Mol. Cell. Biol.* **12**, 1652-1662.

MacMorris, M., Speith, J., Madej, C., Lea, K. and Blumenthal, T. (1994). Analysis of the VPE sequences in the *Caenorhabditis elegans vit-2* promoter with extrachromasomal tandem array-containing transgenic strains. *Mol. Cell. Biol.* **14**, 484-491.

Marin, I. and Baker, B. S. (1998). The evolutionary dynamics of sex determination. *Science* **281**, 1990-1994.

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

Nusbaum, C. and Meyer, B. J. (1989). The *Caenorhabditis elegans* gene *sdc-2* controls sex determination and dosage compensation in XX animals. *Genetics* **122**, 579-593.

Pollock, R. and Treisman, R. (1990). A sensitive method for the determination of protein-DNA binding specificities. *Nucleic Acids Res.* **18**, 6197-6204.

Raymond, C. S., Shamu, C. E., Shen, M. M., Seifert, K. J., Hirsch, B., Hodgkin, J. and Zarkower, D. (1998). Evidence for evolutionary conservation of sex-determining genes. *Nature* **391**, 691-5.

Ryner, L. C., Goodwin, S. F., Castrillon, D. H., Anand, A., Villella, A., Baker, B. S., Hall, J. C., Taylor, B. J. and Wasserman, S. A. (1996). Control of male sexual behavior and sexual orientation in *Drosophila* by the *fruitless* gene. *Cell* **87**, 1079-1089.

Schafer, A. J. and Goodfellow, P. N. (1996). Sex determination in humans. *BioEssays* **18**, 955-963.

Shearman, D. C. and Frommer, M. (1998). The *Bactrocera tryoni* homologue of the *Drosophila melanogaster* sex-determination gene *doublesex*. *Insect Mol. Biol.* **7**, 355-366.

Shen, M. M. and Hodgkin, J. (1988). *mab-3*, a gene required for sex-specific yolk protein expression and a male-specific lineage in *C. elegans*. *Cell* **54**, 1019-1031.

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

Spieth, J., Nettleton, M., Zucker, A. E., Lea, K. and Blumenthal, T. (1991). Vitellogenin motifs conserved in nematodes and vertebrates. *J. Mol. Evol.* **32**, 429-38.

Sturtevant, A. H. (1945). A gene in *Drosophila melanogaster* that transforms females into males. *Genetics* **30**, 297-299.

Sulston, J. and Hodgkin, J. (1988). Methods. In *The Nematode* Caenorhabditis elegans (ed. W. W. Wood), pp. 587-606. Cold Spring Harbor, New York; Cold Spring Harbor Laboratory**.**

Wahli, W. (1988). Evolution and expression of vitellogenin genes. *Trends Genet.* **4**, 227-232.

Waxman, D. and Peck, J. R. (1998). Pleiotropy and the preservation of perfection. *Science* **279**, 1210-1213.

Wilkins, A. S. (1995). Moving up the hierarchy: a hypothesis on the evolution of a genetic sex determination pathway. *BioEssays* **17**, 71-77.

Zucker-Aprison, E. and Blumenthal, T. (1989). Potential regulatory elements of nematode vitellogenin genes revealed by interspecies sequence comparison. *J. Mol. Evol.* **28**, 487-496.