Development 122, 3669-3676 (1996)
Printed in Great Britain © The Company of Biologists Limited 1996
DEV8342

The *Dichaete* gene of *Drosophila melanogaster* encodes a SOX-domain protein required for embryonic segmentation

Steven R. H. Russell*, Natalia Sanchez-Soriano, Charles R. Wright and Michael Ashburner

Department of Genetics, University of Cambridge, Downing Street, Cambridge, CB2 3EH, UK

*Author for correspondence

SUMMARY

We have cloned and characterised a member of the High Mobility Group superfamily of genes from *Drosophila, Sox70D,* which is closely related to the mammalian testis determining gene SRY. *Sox70D* corresponds to the dominant wing mutation *Dichaete.* Homozygous deletions of the *Sox70D* gene and recessive lethal *Dichaete* alleles have a variable embryonic segmentation phenotype. *Dichaete* is expressed in early embryos in a dynamic pattern reminiscent of gap and pair-rule genes and is required for the appropriate expression of the primary pair-rule genes *even skipped, hairy* and *runt.* The molecular nature of *Dichaete* and its expression pattern during early embryogenesis suggest that the gene plays a key role in early development; the variability in both the segmentation phenotype and the effects on pair-rule gene expression suggests that this role is to support the transcriptional regulation of key developmental genes rather than directly regulate any one of them.

Key words: *Sox70D, Dichaete*, pair-rule, gene regulation, segmentation, *Drosophila*

INTRODUCTION

The transcriptional regulation of eukaryotic genes during development is a dynamic process which requires the regulated assembly of multiprotein complexes at promoter and enhancer elements (Tjian and Maniatis, 1994). For example, early zygotic gene expression in the *Drosophila* embryo occurs in a syncytial environment and the enhancer elements that regulate the transcription of segmentation genes respond to gradients of transcription factors (Driever and Nüsslein-Volhard, 1988), resulting in the expression of segmentation genes in very precise temporal and spatial domains (Pankratz and Jackle, 1993 for review). Thus the regulation of segmentation gene expression provides an attractive system for the study of enhancer protein/DNA complexes and the elucidation of a molecular description of developmental gene regulation.

In *Drosophila* the genetic and subsequent molecular analysis of embryonic segmentation has identified a number of key transcriptional regulators that function in segmentation (Akam, 1987). The success of this approach relied upon the identification of mutations with unambiguous phenotypes which delete specific subsets of the body plan. These studies led to a model in which the embryo is progressively subdivided along the anterior-posterior axis by a hierarchy of regulatory genes (Nüsslein-Volhard and Wieschaus, 1980). Briefly, the products of maternal coordinate genes generate broad domains of zygotic gap gene expression. The products of the gap genes, along with the maternal coordinate genes, regulate the expression of the primary pair-rule genes in a characteristic seven-stripe pattern. The primary pair-rule and gap genes control the expression of secondary pair-rule genes and a combination of these factors generate the single segment expression of the segment polarity genes.

In the case of the primary pair-rule genes *even skipped (eve)* and *hairy (h)*, it has been shown that the seven-stripe expression of these genes is generated by transcriptional regulators which act at enhancer elements specific for each stripe (Goto et al., 1989; Riddihough and Ish-Horowicz, 1991). The regulatory elements for each of these genes are large and extend for several kb upstream of the transcriptional start sites. Here we describe a *Drosophila* gene, *Sox70D,* which encodes a member of the recently characterised family of SOX domain proteins, that is implicated in the regulation of pair-rule genes. The SOX domain is a sequence-specific DNA-binding domain found in those proteins of the High Mobility Group (HMG) superfamily, which are closely related to the mammalian sex determining factor SRY (Sinclair et al., 1990). Several genes containing this motif have been implicated in a variety of developmental processes from organisms as diverse as man and yeast (Laudet et al., 1993). One striking feature of HMG domain proteins is their ability to bend DNA upon binding (Ferrari et al., 1992; Giese et al., 1992). Moreover, two SOX domain proteins from mouse, SOX2 and LEF-1, have been shown to be unable to activate transcription on their own; they must act in concert with other enhancer binding proteins (Travis et al., 1991; Yuan et al., 1995). This has led to the suggestion that SOX domain proteins have an architectural role (Grosschedl et al., 1994). We show that *Sox70D* corresponds to the *Dichaete (D)* gene and that the striking features of null mutations in *D* are

the variability of their phenotypes and their variable effects on the expression of other segmentation genes. This suggests that the gene has a supporting role in regulating the expression of key developmental genes during segmentation.

MATERIALS AND METHODS

Drosophila stocks

Drosophila stocks and crosses were maintained on standard yeasted cornmeal-agar food at 25°C. Mutant nomenclature is as described (Lindsley and Zimm, 1992, FlyBase, 1996). In line with FlyBase nomenclature gene symbols are designated with lower case italics and their mRNA by upper case italics. D^1, D^3, D^4 and *Df(3L)fz-GS1a* are described in Lindsley and Zimm (1992). The remaining chromosomes have been generated in this laboratory and will be described elsewhere (S. R. H. R., Adelaide T. C. Carpenter and M. A., unpublished data). *Df(3L)D-5rv6*, *Df(3L)D-1rv16* and *Df(3L)fz-GS1a* all delete the *Sox70D* gene and two other complementation groups, *l(3)70Da* and *deviner;* neither of these other loci have segmentation defects and both fully complement all lethal *D* alleles.

Molecular biology

Genomic clones were isolated from a λGem11 library and cDNA clones from a λZapII prepupal cDNA library (a gift of P. Hurban). Northern blots were generated from poly(A)$^+$ mRNA separated on agarose-formaldehyde gels and probed with random primed *Sox70D* cDNA. All were carried out with minor modifications to standard techniques (Sambrook et al., 1989). A 4 kb *Hin*dIII genomic fragment encompassing the gene and a 1.8 kb cDNA clone corresponding to the *SOX70D* transcription unit were sequenced on both strands using a Sequenase-2 kit according to the manufacturer's instructions. In situ hybridisation to embryos was carried out using minor modifications to standard techniques (Tautz and Pfeifle, 1989). DNA probes for *ftz*, *h* and *run* were a gift from D. Ish-Horowicz and for *eve* a gift from A. Brandt. Grasshopper anti-EVE was a gift from N. Patel (Patel et al., 1992) and was used with minor modifications to the procedure of Rushton et al. (1995).

Transgenic flies

To construct transgenic flies containing the *Sox70D* gene, a full-length cDNA was cloned into the *Eco*RI site of the pCaSpeR-hs vector under the control of the *Hsp70* promoter. The construct was injected into *y w* embryos using standard techniques (Karess, 1985). In the experiments described, flies carrying three or four copies of the transgene, homozygous for inserts on both the X and second chromosomes, were used. For heat-shock experiments, embryos were collected from rapidly laying population cages for 30 minutes or 60 minutes and allowed to age for 2½ hours or 2 hours, respectively. The embryos were collected onto Nitex gauze and placed in a water bath at 36.5°C for 5 minutes or 30 minutes; they were then placed in a 25°C water bath for 15 minutes to recover. Embryos were dechorionated in 50% commercial bleach for 2 minutes and fixed for in situ hybridisation as before. Wild-type embryos were processed in parallel. The material from six consecutive collections was pooled and processed for in situ hybridisation, then divided and hybridised with appropriate probes.

Developmental biology

Embryonic staging was according to Campos-Ortega and Hartenstein (1985). For cuticle preparations, embryos were dechorionated in 50% commercial bleach, devitellinised in 1:1 heptane:methanol, washed with methanol and then 0.1% Triton X-100 and mounted in 3:1 Hoyer's:lactic acid medium. Embryos were cleared overnight at 65°C and viewed with dark-field optics.

RESULTS

Isolation of a *Drosophila* SOX box gene

We identified a *Drosophila* SOX box gene using a PCR product amplified from *Drosophila* genomic DNA with SRY specific primers (a gift of P. Koopman and R. Lovell-Badge). The PCR product was used to isolate clones from *Drosophila* genomic and cDNA libraries which correspond to a single copy gene that maps to 70D1-2 on the left arm of chromosome 3. Sequence analysis of cDNA and genomic clones reveals a single transcription unit of 1800 nucleotides without introns (a feature of some mammalian SOX genes (Colligon et al., 1996)) that would encode a protein of 382 amino acids. Within the predicted protein sequence, there is a 76 amino acid stretch with 88% identity (92% similarity) to the DNA-binding domain of SOX2 proteins from human, mouse and chicken (Colligon et al., 1996; Kamachi et al., 1995; Stevanovic et al., 1994) (Fig. 1). This extraordinary degree of sequence conservation suggests that the *Drosophila* protein may bind the same, or very similar, DNA sequences as SOX2. Outside of the DNA-binding domain there is no similarity to other proteins in the database, although there is a 30 amino acid stretch at the C-terminal end with limited similarity to a potential SOX2 activation domain (Kamachi et al., 1995).

SOX70D expression in the embryo is dynamic

The expression of the *SOX70D* transcript was characterised by northern blotting and whole-mount in situ hybridisation to embryos. A single 1800 nucleotide transcript, consistent with the length of the cDNA clone, is expressed at high levels during early embryogenesis and thereafter at very low levels through the remainder of the life cycle (Fig. 2). The temporal and spatial profile of expression in the early embryo is dynamic (Fig. 3). Zygotic expression is initiated late in stage 4 as a broad central domain which is rapidly followed by the appearance of an anterior domain. We do not detect transcripts prior to nuclear cycle 10, suggesting little or no maternal contribution of transcript. As cellularisation proceeds, the central domain splits and is resolved ventrally into seven stripes while dorsally it remains continuous. By the end of stage 5, the six anterior stripes and the dorsal expression have faded but the most posterior stripe remains strong. During stage 6, the posterior stripe follows the pole cells as the germ band extends and eventually fades as the pole cells are internalised at the amnioproctodeal invagination. Concomitantly, expression is initiated in a region of the neurectoderm which will give rise to the CNS. The neurectodermal expression is transiently observed as fourteen discrete stripes, which rapidly expand to become a continuum along the length of the germ band. Later in development expression is prominent in the nervous system (N. S. S., S. R. H. R., M. A. and Susana Romani, unpublished data). The early pattern of expression is reminiscent of many of the genes involved in segmentation, from an initial gap-like domain, through seven and then fourteen stripes; this suggests that the gene plays a role in segmentation.

Sox70D corresponds to the *Dichaete* gene

The dominant wing mutation *Dichaete (D)* maps to the 70D1-2 region (Bridges and Morgan, 1923) and was used as the starting point for our analysis. The dominant *D* phenotype corresponds to misexpression of *SOX70D* since it is ectopically

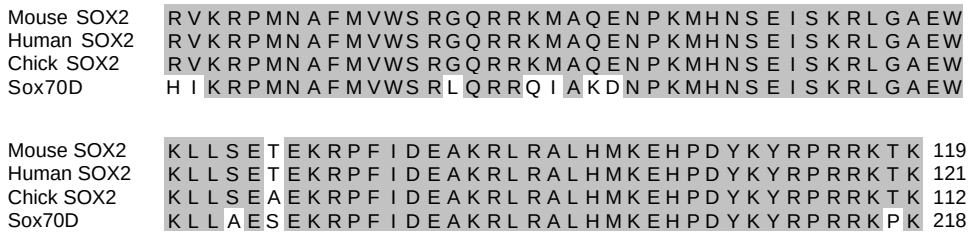

Fig. 1. Alignment of SOX box domains from SOX70D, mouse SOX2, human SOX2 and chicken SOX2. Shaded areas represent amino acid identity. Accession numbers: *Sox70D* X96419; Mouse SOX2 U31967; Human SOX2 Z31560; Chicken SOX2 D50603.

expressed in the wing imaginal discs of all dominant *D* alleles; moreover, dominant wing mutations, with phenotypes similar to those of *D* alleles, can result from ectopic expression of *SOX70D* induced in enhancer trapping experiments (S. R. H. R., Adelaide T. C. Carpenter and M. A., unpublished data). Three *D* alleles were available at the outset of this work; we have generated six new alleles by X-ray mutagenesis. Seven of the nine *D* alleles have an associated recessive lethal phenotype that maps to 70D1-2; the two that do not are the spontaneous $In(3L)D^{1}$ and the X-ray-induced $Df(3L)D^{5}$. One allele, D^{r8}, is recessive lethal but has no dominant phenotype. In five of the seven lethal alleles, early embryonic *SOX70D* expression is severely reduced (D^{3}, D^{9}) or altered (D^{6}, D^{r8}, D^{10}) (Fig. 4). In D^{6}, D^{r8} and D^{10}, the anterior domain never forms and the resolution of the central domain into seven stripes is abnormal and does not reach wild-type levels. Moreover, wild-type levels of expression are never observed. Later in development, *SOX70D* expression in the nervous system is also disrupted in *D* alleles (not shown). We have not detected disruptions of the early expression of *SOX70D* in the remaining two lethal alleles (D^{4}, D^{7}); however, both are chromosome translocations and half of the embryos are grossly aneuploid, which makes whole-mount preparations very difficult to analyse. All *D* alleles are chromosome aberrations and their relationship to the *Sox70D* gene has been determined by in situ hybridisation to polytene chromosomes and by Southern blotting. We have localised breakpoints in three *D* alleles close to the gene (Fig. 5); the remainder map more than 10 kb distal to the 3′ end of *Sox70D.* All lethal *D* alleles have breakpoints 3′ to the *Sox70D* transcription unit whereas the two viable alleles (D^{1}, D^{5}) have breakpoints 5′ (Fig. 5). D^{3}, a spontaneous partial revertant of D^{1} (Plunkett, 1926), shares the D^{1} breakpoint as well as an additional DNA lesion close to the 3′ end of *Sox70D*. In this allele over a deficiency for the region, *SOX70D* expression prior to embryonic stage 6 is barely detectable (Fig. 4B,C). Taken together, these data suggest that *D* corresponds to the *Sox70D* gene and that the recessive lethality associated with *D* alleles is a result of regulatory mutations that reduce *SOX70D* expression in the embryo. Here we focus on the lethal phenotype of *D* and show that the gene has a role in segmentation.

Dichaete mutations have a segmentation phenotype

The embryonic phenotype of lethal alleles was examined in cuticle preparations. In $In(3L)D^{3}/Df(3L)D\text{-}5rv6$ (D^{3}/Df) and also in D^{-} embryos generated from the overlapping deficiencies *Df(3L)D-1rv16/Df(3L)fz-GS1a* (not shown), we observe variable segmentation defects, which include deletions removing half of the segments as well as weaker partial deletions and also segment fusions; in all cases, even numbered metameres are more often affected (Fig. 6B-D). There are also variable defects in head development. In D^{-} embryos, 56% of mutants have a strong segmentation phenotype and 44% have intermediate or weak phenotypes (135/530 embryos examined from crosses between balanced deficiency stocks have a mutant phenotype). In D^{3}/Df, 51% of mutant embryos show strong segmentation phenotypes, the remainder being intermediate or weak (n=412 embryos, 111 with a mutant phenotype). D^{3}/Df is therefore close to the null condition for the gene as far as the segmentation defects are concerned and has been used in subsequent experiments as such. Other alleles, in which *SOX70D* expression is reduced but not absent, show weak segmentation defects in approximately half of the expected mutant embryos and most often these are segment fusions; here again even numbered metameres are most frequently affected (Fig. 6E,F). It is important to emphasise the variability in the phenotype; in both D^{3}/Df and the overlapping deletions, a range of phenotypes is observed, indicating that the gene may be acting as a supporting rather than a specific factor in segmentation.

Pair-rule gene expression is disrupted in *Dichaete* mutants

The effects of lethal *D* mutations on segmentation was further characterised by analysing the expression of a number of key genes in the segmentation pathway. Expression of the gap genes, *Krüppel*, *knirps* and *giant*, was normal in D^{3}/Df, indicating that D^{+} acts downstream or in parallel with these gap genes. With the pair-rule genes, however, we find strong, but

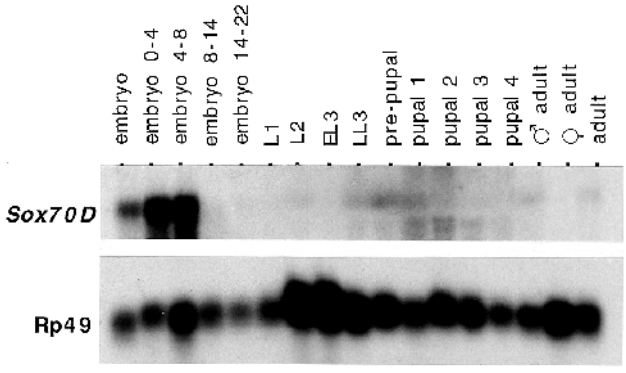

Fig. 2. Developmental northern blot probed with *Sox70D* cDNA. Each lane contains approximately 2 μg of polyA^{+} mRNA; stages are as indicated. Embryo lanes, hours after egg laying. L1, 1st instar larvae; L2, 2nd instar; EL3, early 3rd instar; Pupal, days of pupal development. The strong signal in early embryos is 1.8 kb in length; the additional lower molecular weight transcripts detected in pupae are not consistently detected and may represent cross hybridising genes.

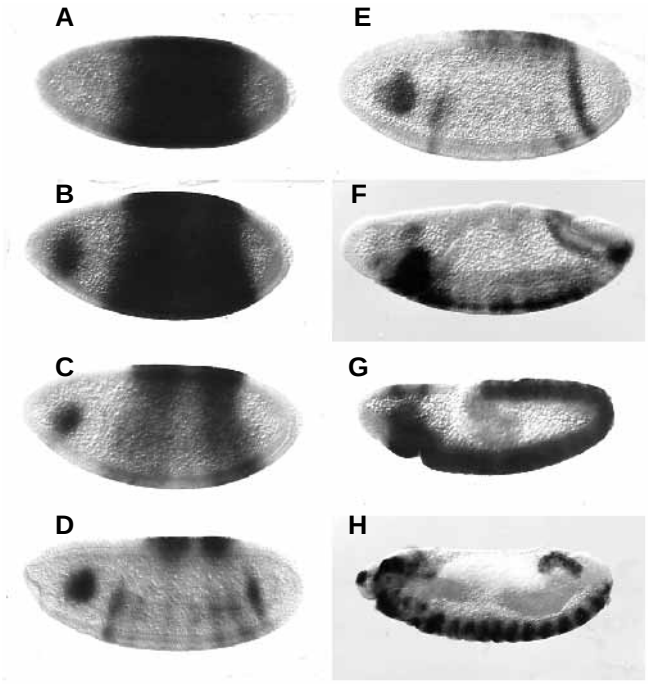

Fig. 3. Whole-mount in situ hybridisation to embryos with *Sox70D* cDNA probe. (A) Stage 4 embryo, the initial central domain. (B) Late stage 4, the development of the anterior domain. (C) Stage 5, the central domain begins to split. (D) Late stage 5, the central domain resolves into seven transient stripes. (E) Early stage 6, the anterior stripes are fading. (F) Late stage 6-early stage 7, neurectodermal expression is initiated as 14 ventral stripes, while the most posterior stripe remaining from stage 5 follows the invaginating pole cells. (G) Stage 10, neurectodermal expression is strong. (H) Stage 13, strong segmental expression in the nervous system. Embryos are oriented anterior to the left and dorsal up.

variable, effects on expression. We have focused on the so-called primary pair-rule genes *even-skipped (eve)*, *hairy (h)* and *runt (run)* since these are thought to be the earliest acting of this class (Howard and Ingham, 1986). In all three cases, we observe reductions in the levels of expression at syncytial blastoderm, a time when these genes are strongly expressed in seven-stripe domains in wild type (Fig. 7).

In the case of *EVE*, the weakest effects are reduced expression of stripe 5 while the strongest effects abolish most of the dorsal expression in stripes 2 through 6 and slightly disrupt the spacing of ventral stripes. Stripes 1 and 7 appear not to be affected. In the case of *HAIRY*, stripes 4, 5 and 6 are most affected with stripes 3 and 7 less so; again dorsal effects are more pronounced than ventral. With *RUN*, stripes 2, 5 and 6 are most frequently affected, 4 and 7 less so with 1 and 3 rarely affected. These patterns of pair-rule gene expression are never observed in wild-type embryos and do not resemble the temporal evolution of the stripe patterns. Again, we emphasise the variability of the effects of these genes; there is no single stripe in any of these phenotypes that is always affected. We also examined the expression of the *fushi tarazu (ftz)* gene in the same genotype. In this case, the defects appear more general; we find a reduction in the intensity of all seven stripes with some stripes more severely affected. Since the stripes of pair-rule gene expression generally occur in the correct anterior-posterior position in D^- embryos, the gene is unlikely to provide key positional infomation; it is more likely to be required in the maintainance or establishment of appropriate levels of pair-rule gene expression in the central region of the embryo.

Ectopic expression of *SOX70D* disrupts pair-rule gene expression

To support the hypothesis that loss or reduction of *SOX70D* expression is responsible for both the alterations in pair-rule gene expression and the segmentation defects observed in lethal *D* combinations, we constructed transgenic flies containing a full-length copy of the *Sox70D* cDNA regulated by the *Hsp70* promoter. None of the lines generated rescue the lethality of *D* alleles. We have used these transgenes to study the effects of ectopic *SOX70D* early in embryonic development in a wild-type background. These experiments assay direct effects of *SOX70D* on target genes, since the time between the heat shock and fixation is short and would not permit the action of a secondary factor induced by *SOX70D* (Manoukian and Krause, 1992). The heat-shock-induced *SOX70D* transcripts are very unstable; ectopic transcripts disappear within 15 minutes of a five minute heat shock (not shown). Two heat-shock regimes were used to assess the effects of ectopic *SOX70D*. In the first, a 5 minute heat shock of $2\frac{1}{2}$ to 3 hour old embryos was followed by 15 minutes recovery and then fixation. There is no consistent effect on the expression of *EVE* or *FTZ*. With *RUN* and *HAIRY* probes, however, there are reproducible alterations in expression (Fig. 8A-D). In both cases, there is precocious expression of normal wild-type

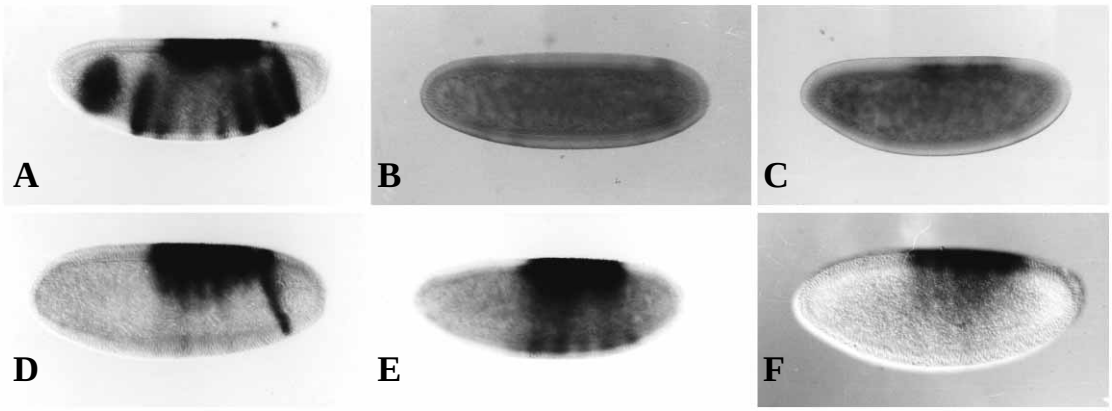

Fig. 4. *SOX70D* expression in hemizygous *D* alleles. (A) Wild type; (B,C) D^3/*D-5rv6*; (D) D^6/*D-5rv6*; (E) D^{r8}/*D-5rv6*; (F) D^{10}/*D-5rv6*. Embryos at stage 5/6 and oriented anterior to the left and dorsal up.

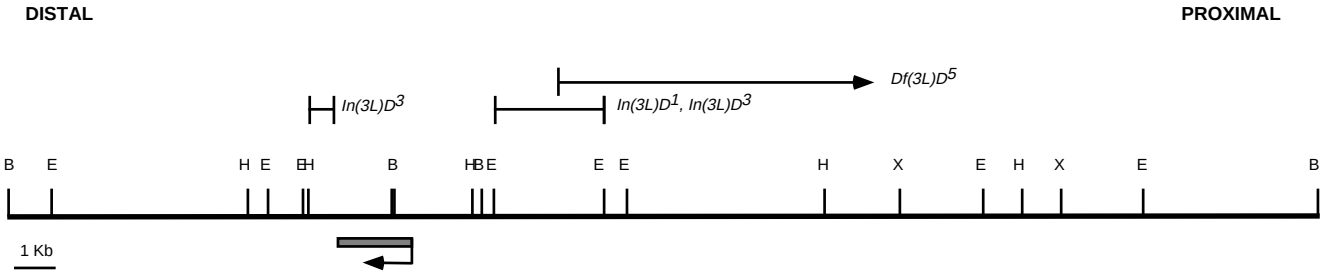

Fig. 5. A molecular map of the *Sox70D* region. Restriction sites are B, *Bam*HI; E, *Eco*RI; H, *Hin*dIII and X, *Xho*I. Below the map the position of the *Sox70D* transcription unit is indicated; the arrow shows the direction of transcription. Above the map are the locations of the *D* breakpoints which are in the region.

features. In the case of *RUN*, the transition from seven to fourteen stripes occurs earlier than in wild-type controls. The appearance of fourteen *RUN* stripes normally occurs during stage 6 (Klingler and Gergen, 1993); in the *Hs-Sox70D* embryos, this occurs during late stage 5. 65% of late stage 5/early stage 6 heat-shocked *Hs-Sox70D* embryos have initiated expression of *RUN* between the seven primary stripes (n=129) compared with 25% of heat-shocked wild type (n=83). With *HAIRY* almost all embryos have a patch of dorsal expression posterior to the seventh stripe and show a posterior expansion of the *HAIRY* anterior domain; 93% (n=112) of late stage 5/early stage 6 heat-shocked *Hs-Sox70D* embryos compared to 10% (n=71) of wild-type controls. These patches of expression normally appear during stage 7 (Hooper et al., 1989, Ingham et al., 1985). Thus short pulses of ectopic *SOX70D* early in development result in subtle, but reproducible, effects on the expression of a subset of pair-rule genes.

Since the effects of short heat shocks were subtle, the level of *SOX70D* was increased by giving a 30 minute heat shock to embryos 2 to 3 hours old. In these experiments, the longer egg collection allowed an analysis of the effects of ectopic *SOX70D* both before and after gastrulation. Here over 90% of embryos heat shocked before gastrulation exhibit severe disruptions in the expression patterns of all three primary pair-rule genes (n>100 in all cases; Fig. 8E-G). Specifically, we observe posterior expansion of the seventh stripe and repression of the central fourth stripe with *EVE*, *HAIRY* and *RUN*. In addition, there is a high level of ectopic expression of both *HAIRY* and *RUN* between the second and third stripes as well as other, more variable, patches of expression between stripes. These effects are never observed in heat-shocked wild-type controls and do not reflect any wild-type pattern of expression. In heat-shocked wild-type embryos, there are no obvious effects on *EVE*; with *HAIRY* 50% of embryos fail to resolve stripes 3 and 4 and with *RUN* 40% of embryos show a reduction in stripe 2.

We have also examined *FTZ* expression in *Hs-Sox70D* embryos. All of the stripes appear to be equal in size and intensity; however, the most posterior stripe is shifted posteriorly by 3-4 cells in over 90% of embryos. There are no obvious effects on wild-type embryos (Fig. 8H). This observation confirms that we are examining a direct effect of *SOX70D* since the effects on *FTZ* are subtle and are not those that would be expected if the aberrantly expressed primary pair-rule genes had time to affect *FTZ* expression.

In those embryos that had begun gastrulation before the heat shock, the expression of the pair-rule genes is more or less normal, indicating that *SOX70D* can only influence transcription of the pair-rule genes early in development (not shown). A proportion of embryos from the heat-shock experiments

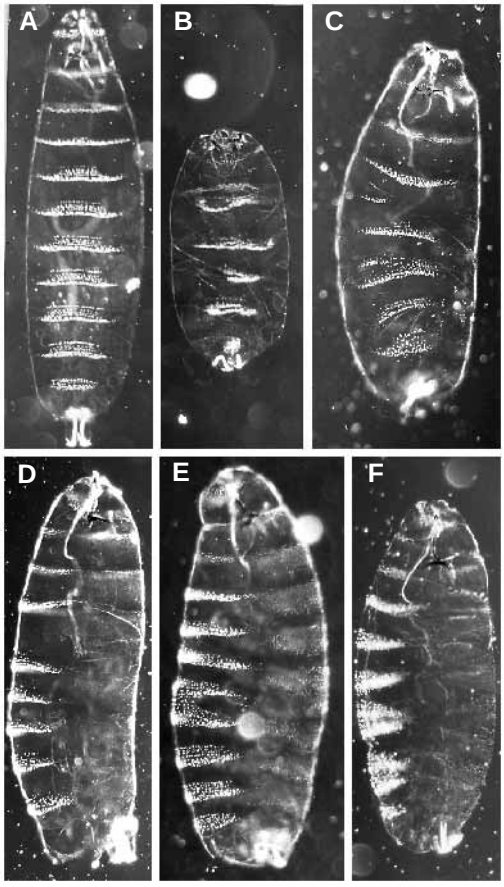

Fig. 6. Cuticle preparations from *D/TM3 x Df(3L)D-5rv6/TM3* crosses. (A) Wild-type cuticle. (B) Strong D^3/*Df(3L)D-5rv6* phenotype, showing severe segment deletions. (C) Intermediate D^3/*Df(3L)D-5rv6* phenotype showing partial deletion of segments and fusion of others. (D) Weak D^3/*Df(3L)D-5rv6* phenotype, naked cuticle in segments A2 and A8. (E) D^{r8}/*Df(3L)D-5rv6*, fusion between segments A4 and A5. (F) D^{10}/*Df(3L)D-5rv6*, intermediate phenotype showing segment fusions in A3 through A7.

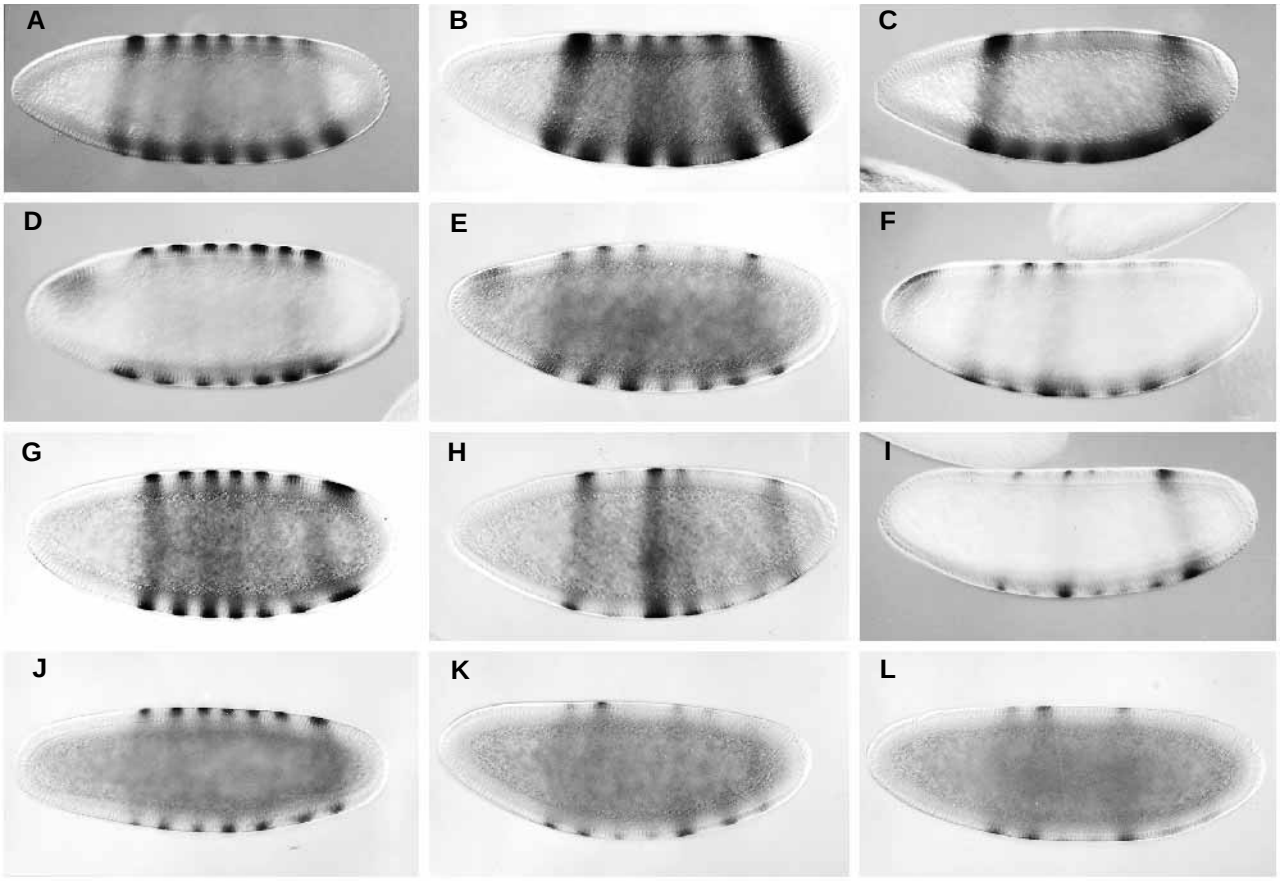

Fig. 7. Defects in pair-rule gene expression in strong *D* mutants visualised by in situ hybridisation. (A,D,G,J) Wild type; (B,C,E,F,H,I,K,L) D^3/*Df(3L)D-5rv6*. Hybridised with probes to (A-C) *eve*; (D-F) *h*; (G-I) *run* and (J-L) *ftz*. All embryos at late stage 4/stage 5. All embryos are oriented anterior to the left and dorsal up.

were allowed to develop and their cuticle phenotypes examined. After a 5 minute heat shock, there is a low level of weak segmentation defects. After a 30 minute heat shock, 65% of embryos show strong segmentation defects very similar to those of strong *D* alleles; this frequency approximates the proportion of embryos in the sample that had not initiated gastrulation at the time of the heat shock. Almost all of these have defects in the 6th abdominal segment as well as severe head defects (n=524) (Fig. 8I,J). In heat-shocked wild-type embryos processed in parallel, 35% have segmentation defects; however, these are more general than in *Hs-Sox70D* embryos and only 7% have head defects.

DISCUSSION

Sox70D corresponds to the *Dichaete* gene and is required for normal embryonic segmentation

We have cloned and characterised a gene encoding a SOX domain protein from *Drosophila* which is dynamically expressed in early embryos in a pattern resembling that of segmentation genes. By two criteria *Sox70D* corresponds to the *Dichaete* gene. (1) Its misexpression appears to cause the dominant wing phenotype since it is ectopically expressed in the wing discs of all eight dominant *D* alleles and dominant wing phenotypes can result from ectopic expression generated by mobilising a *Sox70D* cDNA containing P element in a wild-type background (S. R. H. R., Adelaide T. C. Carpenter and M. A., unpublished data). (2) Seven *D* alleles have an associated recessive lethal phenotype and, in at least five of these alleles, the early embryonic expression of *SOX70D* is aberrant. The data presented in this paper and discussed below support the hypothesis that the lethality of *D* alleles is due to regulatory mutations that disrupt *SOX70D* expression.

Lethal *D* alleles that eliminate or severly reduce *SOX70D* expression in the early embryo have severe segmentation defects. In alleles in which *SOX70D* expression is reduced but not eliminated segmentation defects are milder. Thus the severity of the segmentation defect correlates with the expression of *SOX70D*. Our data shows that the segmentation phenotype is due, in part, to repression of the primary pair-rule genes *eve*, *h* and *run* since the characteristic seven-stripe expression of these genes is altered in strong lethal *D* mutations. The expression pattern of *SOX70D* fits well with a factor required for the expression of pair-rule genes since it is strongly expressed at the same time, or shortly before, they are established in their seven-stripe domains. Furthermore, the stronger dorsal expression of *SOX70D* in wild type is consistent with the more pronounced dorsal effects on pair-rule gene expression observed in lethal *D* genotypes. The experiments do not allow us to ascertain whether alterations in the expression of a particular pair-rule gene are directly due to loss of *D* function or

due to secondary effects. For example, the general reduction in *FTZ* expression that we observe may be due to the prior disruption in primary pair-rule gene expression.

Dichaete directly regulates pair-rule genes

To address the possibility of direct effects of *D* on pair-rule gene expression, we have ectopically expressed the gene early in embryonic development and monitored changes in the expression of *EVE, FTZ, HAIRY* and *RUN*. High levels of *D* expression administered prior to gastrulation severely disrupt *EVE, HAIRY* and *RUN* expression in three ways. Firstly, there is an expansion of the most posterior stripe, sometimes reaching the posterior end of the embryo. Secondly there is expansion of stripe domains so that stripes are fused or patches of ectopic expression are visible between the normal stripes. These effects are the reciprocal of those observed in D^- embryos where pair-rule gene expression is reduced. The third effect of ectopic *D* expression is a repression of each of the primary pair-rule genes in the central domain of the embryo which encompasses stripe four. This is particularly apparent with *EVE*, where stripe four is eliminated in a high proportion of *Hs-Sox70D* embryos. Significantly, this corresponds to the region of the embryo that first loses wild-type *D* expression prior to its resolution into stripes (Fig. 3C). Double labelling of embryos with *SOX70D* and *EVE* confirm that *EVE* stripe four lies within this initial gap in *D* expression (N. S. S. and S. R. H. R., unpublished observations). We interpret the effects observed in the heat-shock experiments as reflecting the direct action of *D* for two reasons. Firstly, the time between the heat shocks and fixation is short and would not be expected to permit a secondary factor to be translated and then to modulate the expression of the pair-rule genes (Manoukian and Krause, 1992). Secondly, the effects of ectopic *D* on the expression of *FTZ* are not what would be expected if other regulators are being induced by *D*. If, for example, the aberrantly expressed EVE protein had time to affect the expression of *FTZ*, we would expect more dramatic effects on *FTZ* RNA (Manoukian and Krause, 1992; Yu and Pick, 1995). Taken together, the ectopic expression experiments and the loss-of-function studies indicate that *D* can act to both activate and repress a particular stripe domain and suggests that changing levels of D protein may be important in modulating pair-rule gene expression early in development.

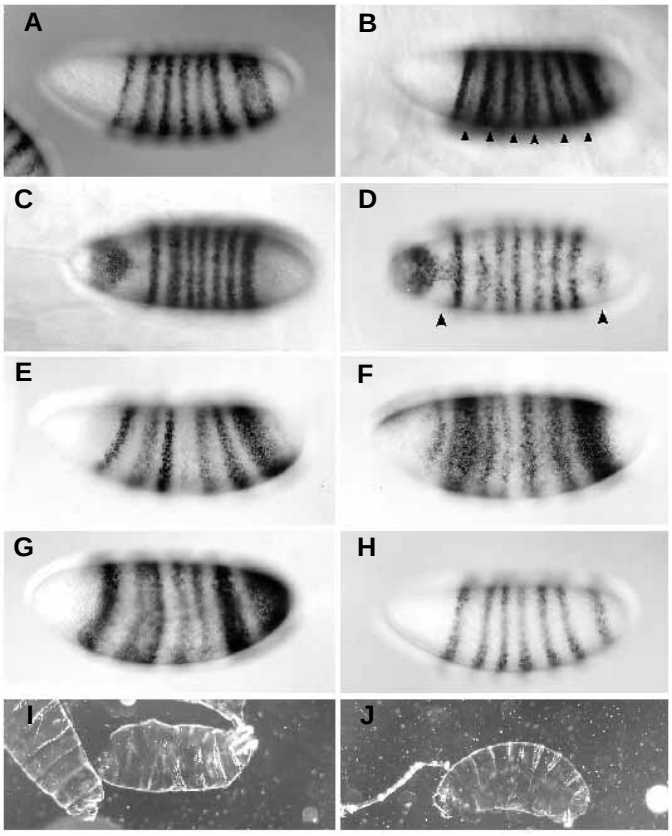

Fig. 8. Ectopic *SOX70D* expression induces ectopic expression of pair-rule genes. (A,C) Heat-shock wild type and (B,D,E-H) heat-shock *Hs-Sox70D* (probed with DNA for (A,B,G) *run*, (C,D,F) *h*, (H) *ftz* and (E) a monoclonal antibody against the EVE protein, after a 5 minute heat shock (A-D) or a 30 minute heat shock (E,H). Arrowheads point to ectopic patches of *run* and *h*. Embryos are oriented anterior to the left and dorsal up except for C and D which are dorsal views. The cuticles in I and J are from 30 minute *Hs-SOX70D* heat shocks.

We have also examined the effects of extremely short pulses of *D*. In this case, the effects are more subtle and appear to represent precocious expression of normal wild-type features. Thus there is expression of *RUN* in between the seven primary stripes and the appearance of anterior and posterior domains of *HAIRY*. It is possible that these effects are a reflection of wild-type *D* function slightly later in development since *D* is transiently expressed in fourteen stripes at the onset of gastrulation and at stage 7 in the areas that correspond to the patches of *HAIRY* expression induced by ectopic *D* (Fig. 3F). This supports the view that the level of D protein is important for regulating gene expression.

Dichaete may act as modulator of chromatin structure

A striking feature of the loss of *D* function is the variability observed at both the phenotypic and molecular levels; in overlapping deficiencies that remove *D* approximately half of the mutant embryos have strong segmentation defects and the remainder have weaker defects. Three possible explanations could account for the variable phenotype. Firstly, zygotic *D* function could be augmented by a maternal contribution; we do not detect maternal *D* transcripts but we cannot exclude maternal contribution of protein to the embryo. However, given that levels of *D* expression in the adult female are low, we do not expect a substantial maternal contrbution. Secondly, the variability might be due to partial complementation by the normal activity of another SOX gene. Although we cannot eliminate this possibility, we have cloned and partially characterised a further four members of the SOX family in *Drosophila* and these genes are not expressed as early in embryonic development as is *D* (S.R.H.R. unpublished observations). Thirdly, the variability may be intrinsic to the wild-type activity of *D*. This view implies that *D* is not absolutely required for pair-rule gene regulation and suggests that it has an accessory role in transcription.

The data that we have presented do not support models in which D acts as a general transcription factor at basal promoters. If the protein did act in this way we would expect more constant and dramatic effects on pair-rule gene expression in its absence. It is also unlikely that *D* provides key positional information, since stripes of pair-rule gene expression are expressed in approximately the correct anterior-posterior position in the absence of the gene. At present, we favour a model in which D

modulates chromatin structure to assist in generating appropriate levels of pair-rule gene expression in the domains established by both the gap genes and interactions between primary pair-rule genes. In this view, transcription factors act less efficiently at stripe-specific enhancers in the absence of D; conversely, ectopic expression of *D* allows transcriptional activation or repression of target genes by concentrations of transcription factors below those normally active. In the early syncytial embryo, where complex patterns of gene expression are rapidly generated by gradients of transcription factors, there may be a requirement for architectural components that regulate the formation of particular chromatin configurations. It is expected that favourable chromatin conformations would promote the interaction of regulatory molecules, bound to sequences distant from the promoter, with the basal transcription complex. The D protein is a candidate for such an activity. Since SOX box containing genes are known to bend DNA upon binding (Ferrari et al., 1992; Giese et al., 1992), the role of the gene may be mediated by this function. The primary pair-rule genes are distinguished by their very long 5′ regulatory sequences spanning several kb upstream of the transcription start site. We have examined the available regulatory sequences upstream of primary pair-rule genes for the presence of mouse SOX2 DNA-binding motifs (our preliminary data indicates that D binds to a similar sequence, Stefan Oehler and S. R. H. R, unpublished observations). We find a cluster of these elements located between the promoter and the stripe-specific enhancers in the *h* gene, a location consistent with a bending model.

The *Sox70D* gene has been independently isolated by P. A. Nambu and J. R. Nambu (1996, this issue of *Development*) as the *fish-hook* gene. These authors reach very similar conclusions to those described in this paper. We thank John Nambu for providing his manuscript prior to publication.

This work was supported by an MRC programme Grant to M. A., a Winston Churchill fellowship and Wellcome Trust studentship to C. R. W and a BBSRC fellowship to N. S. S. This work would not have been possible without the generous assistance of Adelaide Carpenter, whose genetic analysis of the region provided a basis for the characterisation of *Sox70D*. We are indebted to A. Brandt, J. F. de Celis, P. Goodfellow, D. Ish-Horowicz, R. Lovell-Badge, N. Patel, S. Romani and R. Rivera-Pomar for probes and helpful discussions.

REFERENCES

Akam, M. (1987). The molecular basis for metameric pattern in the *Drosophila* embryo. *Development* **101**, 1-22.

Bridges, C. B. and Morgan, T. H. (1923). The third-chromosome group of mutant characters of *Drosophila melanogaster*. *Publs Carnegie Instn.* **327**, 1-251.

Campos-Ortega, J. A. and Hartenstein, V. (1985). *The Embryonic Development of Drosophila melanogaster*. Berlin: Springer-Verlag.

Colligon, J., Sockanathan, S., Hacker, A., Cohen-Tannoudji, M., Norris, D., Rastan, S., Stevanovic, M., Goodfellow, P. N. and Lovell-Badge, R. (1996). A comparison of the properties of *Sox-3* with *Sry* and two related genes *Sox-1* and *Sox-2*. *Development* **122,** 509-520.

Driever, W. and Nüsslein-Volhard, C. (1988). The Bicoid protein determines position in the *Drosophila* embryo in a concentration-dependent manner. *Cell* **54**, 95-104.

Ferrari, S., Harley, V. R., Pontiggia, A., Goodfellow, P. N., Lovell-badge, R. and Bianchi, M. E. (1992). Sry, like HMG1, recognizes sharp angles in DNA. *EMBO. J.* **11**, 4497-4506.

FlyBase (1996). The *Drosophila* genetic database. *Nucleic Acids Research* **24**, 53-56.

Giese, K., Cox, J. and Grosschedl, R. (1992). The HMG domain of lymphoid enhancer factor-I bends DNA and facilitates assembly of functional nucleoprotein structures. *Cell* **69**, 185-195.

Goto, T., Macdonald, P. and Maniatis, T. (1989). Early and late periodic patterns of *even skipped* expression are controlled by distinct regulatory elements that respond to different spatial cues. *Cell* **57**, 413-422.

Grosschedl, R., Giese, K. and Pagel, J. (1994). HMG domain proteins – architectural elements in the assembly of nucleoprotein structures. *Trends Genet.* **10**, 94-100.

Hooper, K. L., Parkhurst, S. M. and Ish-Horowicz, D. (1989). Spatial control of Hairy protein expression during embryogenesis. *Development* **107**, 489-505.

Howard, K. R. and Ingham, P. W. (1986). Regulatory interactions between segmentation genes during *Drosophila* embryogenesis. *Heredity* **57**, 284

Ingham, P. W., Howard, K. R. and Ish-Horowicz, D. (1985). Transcription pattern of the *Drosophila* segmentation gene *hairy*. *Nature* **318**, 439--445.

Kamachi, Y., Sockanathan, S., Liu, Q. R., Breitman, M., Lovell-badge, R. and Kondoh, H. (1995). Involvement of Sox proteins in lens-specific activation of crystallin genes. *EMBO J.* **14**, 3510-3519.

Karess, R. E. (1995). P element mediated germ line transformation of *Drosophila.* In *DNA Cloning vol II.* (ed D. M. Glover), pp 121-142. Oxford: IRL Press.

Klingler, M. and Gergen, J. P. (1993). Regulation of *runt* transcription by *Drosophila* segmentation genes. *Mech. Dev.* **43**, 3-19.

Laudet, V., Stehelin, D. and Clevers, H. (1993). Ancestry and diversity of the HMG box superfamily. *Nucleic Acids Research* **21**, 2493-2501.

Lindsley, D. L. and Zimm, G. (1992). *The Genome of Drosophila melanogaster*. New York: Academic Press.

Manoukian, A. S. and Krause, H. M. (1992). Concentration-dependent activities of the Even-skipped protein in *Drosophila* embryos. *Genes Dev.* **6**, 1740-1751.

Nambu, P. A. and Nambu, J. R. (1996). The *Drosophila fish-hook* gene encodes a HMG domain protein essential for segmentation and CNS development. *Development* **122**, 3467-3475.

Nusslein-Volhard, C. and Wieschaus, E. (1980). Mutations affecting segment number and polarity in *Drosophila*. *Nature* **287**, 795-801.

Pankratz, M. J. and Jackle, H. (1993). Blastoderm segmentation. In *The Development of Drosophila melanogaster*. (ed. M. Bate and A. Martinez-Arias). pp 467-516 New York: Cold Spring Harbor Laboratory Press.

Patel, N. H., Ball, E. E. and Goodman, C. S. (1992). Changing role of *even-skipped* during the evolution of insect pattern formation. *Nature* **357**, 339-342.

Plunkett, C. R. (1926). The interaction of genetic and environmental factors in development. *J. Exp. Zool.* **46**, 181-244.

Riddihough, G. and Ish-Horowicz, D. (1991). Individual stripe regulatory elements in the *Drosophila hairy* promoter respond to maternal, gap, and pair-rule genes. *Genes Dev.* **5**, 840-854

Rushton, E., Drysdale, R. A., Abmayr, S. M., Michelson, A. M. and Bate, M. (1995). Mutations in a novel gene, *myoblast city*, provide evidence in support of the founder cell hypothesis for *Drosophila* muscle development. *Development* **121**, 1979-1988.

Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*. New York: Cold Spring Harbor Laboratory Press.

Sinclair, A. H., Berta, P., Palmer, M. S., Hawkins, J. R., Griffiths, B. L., Smith, M. J., Foster, J. W., Frishauf, A.-M., Lovell-Badge, R. and Goodfellow, P. N. (1990). A gene from the human sex-determining region encodes a protein with homology to a conserved DNA-binding motif. *Nature* **346**, 240-244.

Stevanovic, M., Zuffardi, O., Collignon, J., Lovell-Badge, R. and Goodfellow, P. N. (1994). The cDNA sequence and chromosomal location of the human *SOX2* gene. *Mamm. Genome* **5**, 640-642.

Tautz, D. and Pfeifle, C. (1989). A non-radioactive *in situ* hybridisation method for the localisation of specific RNAs in *Drosophila* embryos reveals translational control of the segmentation gene *hunchback*. *Chromosoma* **98**, 81-85.

Tjian, R. and Maniatis, T. (1994). Transcriptional activation: a complex puzzle with few easy pieces. *Cell* **77**, 5-8.

Travis, A., Amsterdam, A., Belanger, C. and Grosschedl, R. (1991). *LEF-1*, a gene encoding a lymphoid-specific protein with an HMG domain regulates T-cell receptor α enhancer function. *Genes Dev.* **5**, 880-894.

Yu, Y. and Pick, L. (1995). Nonperiodic cues generate seven *ftz* stripes in the *Drosophila* embryo. *Mech. Dev.* **50**, 163-175

Yuan, H., Corbi, N., Basilico, C. and Dailey, L. (1995). Developmental-specific activity of the FGF-4 enhancer requires the synergistic action of Sox2 amd Oct-3. *Genes Dev.* **9**, 2635-2645.

(Accepted 15 August 1996)