Molecular characterization of daughterless, a Drosophila sex determination gene with multiple roles in development

Claire Cronmiller, Paul Schedl, and Thomas W. Cline

Department of Biology, Princeton University, Princeton, New Jersey 08544-1003 USA

The daughterless (da) gene in Drosophila acts both maternally and zygotically to provide essential functions during development. Maternal da+ expression is required by embryos for the regulation of sex determination and dosage compensation. Zygotic da+ expression is required throughout development—early for the formation of the peripheral nervous system and perhaps for the proper functioning of genes in heterochromatic regions of the genome; during larval stages for growth and differentiation of the future adult epidermis; and in the somatic part of the gonad of adult females for proper egg membrane synthesis. Here we describe the cloning of da by the transposon tagging approach as well as some aspects of the molecular characterization of wild-type and mutant alleles. Despite the multiple developmental roles of da, the organization and expression of this gene appear relatively simple at this level of analysis. The gene codes for two transcripts, present in both sexes and at all stages of development. The nucleotide sequence of a nearly full-length cDNA predicts a protein product of 710 amino acids that shares sequence similarity with the His-Pro repeat of the Drosophila genes bicoid and paired. Two partial-loss-of-function da mutations (one of which is temperature sensitive) appear to be caused by DNA insertions in the 5'-untranslated region of the gene.

[Key Words: daughterless; sex determination; development; Drosophila; His-Pro repeat]

Received August 17, 1988; revised version accepted October 3, 1988.

In Drosophila, sexual phenotype and the overall level of X-linked gene expression are determined by the number of X chromosomes relative to the number of autosomes, a parameter called the X : A ratio. The X : A ratio accomplishes this by its action on the gene, Sex-lethal (Sxl) [1-19.2] (for recent reviews, see Baker et al. 1987; Lucchesi and Manning 1987; Cline 1988a). Higher values of this parameter [XXAA, X : A = 1.0] elicit female development and suppress transcriptional hyperactivation of the X chromosome by activating the female-specific functions of Sxl+. Male development and X chromosome hyperactivation occur at lower values [XYAA, X : A = 0.5] as a consequence of Sxl+ remaining inactive.

The effectiveness of this X/A balance in activating the feminizing functions of Sxl+ depends on the maternal expression of the gene daughterless (da; 2-41.5) (Cline 1980, 1983, 1984, 1988b). In the absence of da+ activity in the germ line of the mother, all progeny, regardless of their X/A balance, fail to activate Sxl+ and therefore develop as males (Cronmiller and Cline 1987). The name daughterless (Bell 1954) reflects the fact that daughters [X/A = 1] from mothers homozygous for the original spontaneous mutant allele, da1, die prior to the adult stage (Sandler 1972; Mange and Sandler 1973) as a consequence of upsets in X chromosome dosage compensation (Gergen 1987). In contrast, sons [X/A = 0.5] are unaffected by the lack of maternal germ line da+ activity (Cronmiller and Cline 1987) because they do not require (and indeed cannot tolerate) zygotic Sxl+ activity (Cline 1979; Salz et al. 1987). The da+ maternal product, in conjunction with components of the X/A signal such as sisterless-a and sisterless-b, interacts with Sxl+ only early in development (Cline 1976, 1980, 1988b; Sanchez and Nothiger 1983; Maine et al. 1985a). The effect of this early interaction persists as a consequence of the subsequent maintenance of the Sxl activity state (Cline 1984, 1985). Only the maternal expression of da+ is relevant to Sxl expression; the zygotic genotype with respect to da appears to have no effect on Sxl (Cline 1980; Cronmiller and Cline 1986).

Phenotypic analysis of mutant da alleles has demonstrated that the requirement for maternal germ line da+ activity in the control of progeny sex determination reflects only one of several important functions of this gene. Embryonic expression of da+ is required in both sexes for the formation of the peripheral nervous system (PNS) (Caudy et al. 1988) and for some process that is sensitive to the dose of genes in heterochromatic regions of the chromosomes (Mange and Sandler 1973; Sandler...
Molecular characterization of \textit{daughterless}

One new allele, \textit{da}^{pa}, was recovered in this screen. Phenotypic characterization of this allele shows it to be the weakest \textit{da} mutation yet identified. Homozygous \textit{da}^{pa} animals appear nearly wild type: Not only is their viability high (Table 1, Zygotic phenotype), but also the sex ratio among the progeny of these ‘daughterless’ females is normal (Table 1, Maternal phenotype). Nevertheless, \textit{da}^{pa} is defective clearly for both zygotic and maternal functions. First, \textit{da}^{pa} fails to complement the recessive lethality of the ethylmethane sulfonate (EMS)-induced null alleles, \textit{da}^{a} and \textit{da}^{3}. Second, \textit{da}^{1}/\textit{da}^{pa} heterozygotes display reduced viability (43\% versus 71\%); moreover, these heteroallelic females exhibit a female-lethal maternal effect, although one that is much less severe than that of \textit{da}^{1}/\textit{da}^{1} females (progeny sex ratio = 0.46 versus 0.00).

Two observations indicated that the insertion of a \textit{P} element is responsible for the mutant phenotype of \textit{da}^{pa}. First, in situ hybridization to polytene chromosomes identified \textit{P}-element sequences at 31E, consistent with the location of \textit{da} at 31B-F reported by Sandler (1977). Second, the \textit{da}^{pa} allele was found to be unstable under conditions that mobilize \textit{P} elements. When the allele was outcrossed under dysgenic conditions and selected for viability in \textit{trans} with a strong \textit{da} allele (\textit{da}^{a}), \textit{da}^{pa} was found to revert to wild type at a frequency of 1.5\%. Under the same dysgenic conditions, derivatives that were more extreme than the parental allele were generated at a frequency of 0.3\%. These more extreme alleles were identified by their failure to survive when hetero-

Table 1. Phenotypic characterization of \textit{da}^{pa}

| Maternal phenotype: sex-ratio maternal effect |
|---------------------------------------------|
| Genotype of mothers mated to +/- males* | Progeny sex ratio (female/male) | Number of male progeny recovered |
| \textit{da}^{a}/\textit{da}^{pa} | 1.03 | 120 |
| \textit{da}^{a}/\textit{da}^{1} | 0.46 | 292 |
| +/- \textit{da}^{1} | 0.85 | 368 |
| \textit{da}^{1}/\textit{da}^{a} | 0.00 | 1698* |

| Zygotic phenotype: recessive lethal effect |
|-------------------------------------------|
| Crossc Genotype | Viability of \textit{da} mutant flies relative to their \textit{da}^{+} siblings [%] | Number of \textit{da}^{+} siblings recovered |
| A \textit{da}^{a}/\textit{da}^{pa} | 71 | 166 |
| B \textit{da}^{a}/\textit{da}^{1} | 43 | 367 |
| C \textit{da}^{a}/\textit{da}^{2} | 0 | 229 |
| D \textit{da}^{a}/\textit{da}^{3} | 0 | 296 |
| E \textit{da}^{1}/\textit{da}^{a} | 0 | 540 |

* \textit{da}^{pa} = \textit{cl} \textit{da}^{a} \textit{b} \textit{pr} \textit{cn}; \textit{da}^{a}/\textit{da}^{1} and +/- \textit{da}^{1} mothers were obtained from cross B below.

* Data from Cronmiller and Cline (1986, Table 7).

- Full genotypes of crosses [at 25°C]: \textit{cl} \textit{da}^{a} \textit{b} \textit{pr} \textit{cn}/\textit{CyO}, \textit{da}^{a} \textit{pr} \textit{cn} females × the following males: [A] same as females; [B] \textit{cl} \textit{da}^{1} \textit{cn} \textit{bw}; [C] \textit{y} \textit{Y}; \textit{cl} \textit{da}^{a}/\textit{CyO}, \textit{da}^{a} \textit{pr} \textit{cn} females × the following males: [A] same as females; [B] \textit{cl} \textit{da}^{a} \textit{cn} \textit{bw}; [C] \textit{y} \textit{Y}; \textit{cl} \textit{da}^{a}/\textit{CyO}, \textit{da}^{a} \textit{pr} \textit{cn}.

* Data from Cronmiller and Cline (1987, Table 1).

1975; Pimpinelli et al. 1985]. Larval and/or pupal expression of \textit{da}^{+} is required for the proper growth and differentiation of cells that form the adult cuticle (Cronmiller and Cline 1987). Finally, expression of \textit{da}^{+} is required in the adult female in more than just her germ line: Somatic expression is required for proper egg membrane formation, and, hence, for survival of all progeny regardless of their sex (Cline 1976, 1980, Cronmiller and Cline 1987). The sex-ratio maternal effect for which the gene was named was discovered only because the first mutant allele of this gene, \textit{da}^{1}, provided sufficient zygotic activity, at least at lower growth temperatures, to allow homozygous mutant mothers to survive and produce offspring.

The involvement of \textit{da} in two processes as different as sex determination and neurogenesis is intriguing in light of recent studies on another class of developmentally important genes, the segmentation genes, whose functioning in the nervous system may be significantly different from their functioning in nonneuronal tissues. The fact that the regulatory relationships among even-skipped, \textit{fushi tarazu}, and \textit{engrailed} during neuronal development are very different from those during segmentation suggests that these genes may be interacting in different ways in the two processes (Doe et al. 1988). In the case of \textit{da}, molecular analysis should help elucidate the relationships among this gene’s multiple functions: Do they reflect a multitude of \textit{da} gene products, or does a single \textit{da} gene product participate in a variety of regulatory processes instead? Is the product of this regulatory gene related to the products of other regulatory genes in ways that might suggest its biochemical functions? To address these questions, we have (1) cloned \textit{da} through transposon tagging, (2) examined the transcription pattern for wild-type and partial-loss-of-function mutant alleles; and (3) determined the sequence of the wild-type coding region and of regions altered in two mutant \textit{da} alleles.

Results

Tagging \textit{da} with a \textit{P}-element transposon

As the first step in the transposon tagging strategy (Bingham et al. 1981), we used P-M hybrid dysgenesis to insert a \textit{P} element into \textit{da}. Because \textit{P}-element insertions often are associated with only partial-loss-of-function (hypomorphic) mutations (Tsubota and Sandler 1986), we exploited the recessive zygotic lethal aspect of the complex \textit{da} phenotype to serve as a sensitive assay for reductions in \textit{da} function: Hypomorphic alleles that are homozygous viable (such as the original allele, \textit{da}^{1}) are lethal when heterozygous with a \textit{da}^{−} deletion (Mange and Sandler 1973). In contrast, wild-type alleles fully complement \textit{da}^{−} deletions. Among the progeny of a dysgenic cross, 10,513 second chromosomes were tested for loss of the ability to complement a \textit{da}^{−} deletion for viability. Chromosomes satisfying this criterion were retested for their ability to complement various \textit{da} point mutants. This allowed us to distinguish new \textit{da} alleles from mutations in other genes missing from the tester deficiency, \textit{Df}(2L)\textit{der}-27 (Sandler 1977).

Cross 1. Genotype | Viability of \textit{da} mutant flies relative to their \textit{da}^{+} siblings [%] | Number of \textit{da}^{+} siblings recovered
|------------------|---------------------------------|------------------|

- Data from Cronmiller and Cline (1986, Table 7).

- Full genotypes of crosses [at 25°C]: \textit{cl} \textit{da}^{a} \textit{b} \textit{pr} \textit{cn}/\textit{CyO}, \textit{da}^{a} \textit{pr} \textit{cn} females × the following males: [A] same as females; [B] \textit{cl} \textit{da}^{1} \textit{cn} \textit{bw}; [C] \textit{y} \textit{Y}; \textit{cl} \textit{da}^{a}/\textit{CyO}, \textit{da}^{a} \textit{pr} \textit{cn} females × the following males: [A] same as females; [B] \textit{cl} \textit{da}^{1} \textit{cn} \textit{bw}; [C] \textit{y} \textit{Y}; \textit{cl} \textit{da}^{a}/\textit{CyO}, \textit{da}^{a} \textit{pr} \textit{cn}.

- Data from Cronmiller and Cline (1987, Table 1).
zygous with the hypomorphic da allele. Thus, mutations that increased da function as well as those that decreased its function occurred several orders of magnitude more frequently than the event that gave rise to da in the first place. This is the result expected if the mutant phenotype of da is caused by a P element that can be mobilized.

**Molecular cloning of DNA from the vicinity of da**

To clone DNA flanking the P element at da, we constructed a recombinant λ phage library from da DNA and isolated phage containing P-element sequences. In situ hybridization to polytene chromosomes lacking P elements revealed that one of these phage (CC1) hybridized to 31E (Fig. 1). Two restriction fragments from recombinant CC1, one from each side of the P element, were used then to isolate wild-type genomic recombinant phage from an Oregon R DNA library. In this way we isolated ~35 kb of DNA surrounding the P-element site of insertion at 31E. A composite restriction map of this DNA is shown in Figure 2 with the site of the P-element insertion designated as coordinate 0.

**DNA rearrangements associated with da mutant alleles and their derivatives**

To confirm that the P-element insertion in the cloned 31E region DNA indeed was responsible for the mutant phenotype of da, we examined the DNA sequence organization of the dysgenic derivatives of da mentioned above. In all cases examined, a change in the phenotype of da was accompanied by a change in this P element. Eleven independent wild-type revertants of da all appear to represent precise excisions of the P element (Fig. 3A): In contrast to da, they have a fragment that comigrates with the wild type. More complex rearrangements are associated with the three more extreme derivatives.

As indicated by the diagrams shown in Figure 3B, two appear to have lost some P-element and/or flanking sequences, while in the third the P element appears to have reversed its orientation.

To obtain additional evidence that this 31E DNA segment corresponds to the da locus, we looked for differences in the restriction fragment pattern of genomic DNA from three EMS-induced loss-of-function mutants (da2, da3, and da11831) and from one spontaneous partial-loss-of-function mutant, da4. The three EMS-induced alleles showed no evidence of sequence rearrangement at this level of analysis. In contrast, da4 was found to be associated with an insertion. Moreover, detailed restriction analysis placed the da4 insertion site between coordinates 0 and +0.9 (see Fig. 2), the same region into which the P transposon of da is inserted. The genomic Southern blot in Figure 3C illustrates the da4 insertion. The material inserted into da4 is a member of a moderately repetitive DNA family, but it is not homologous to the P-element family (data not shown).

**Identification of the da transcription unit**

To identify transcription units within the cloned DNA from the da region, we probed Northern blots of poly(A)+ RNA from 0- to 18-hr embryos, adult females, and adult males with seven fragments spanning 30 kb of the cloned region. The results of this Northern walk are summarized in Figure 4. Transcription units were found to be distributed throughout the region, and all seven fragments hybridized to RNA. All of these transcripts were present, although in varying abundance, in adults of both sexes and in embryos (data not shown). There was a large variety of transcripts in the region surrounding the P element and da4 insertion sites; hence, it was not immediately evident which of these RNAs might be da gene products.

Since the homozygous viable (hypomorphic) alleles, da and da4 are associated with DNA insertions, we thought that their da transcripts might be affected. For this reason we examined Northern blots of poly(A)+ RNA prepared from homozygous mutant da and da4 adult females and males. These blots were probed with fragments 3, 4, 5, and 6, as well as with the whole recombinant phage PS6 and PS10 (see Figs. 2 and 4). Of the transcripts detected by these probes, only the 3.2- and 3.4-kb RNAs were altered in these mutants. As shown in Figure 5A, the 3.2- and 3.4-kb transcripts are reduced substantially in da4, and may be absent in da4. Furthermore, the probes that label the 3.2- and 3.4-kb species in the wild-type strain hybridize to two, new, higher-molecular-weight RNA species in both mutants: In da4, the pair of mutant transcripts are ~3.5 and 3.7 kb, while those in da4 are even larger (3.8 and 4.0 kb). Because both insertion mutations disrupt the 3.2- and 3.4-kb transcripts, while no other RNAs from the cloned region appear to be affected, we conclude that the 3.2- and 3.4-kb RNAs are the transcription products of the da gene.

Figure 1. In situ hybridization of recombinant CC1 to wild-type polytene chromosomes. CC1 hybridizes to sequences at 31E (arrow), within the 'gooseneck' of chromosome 2L.
Developmental profile of da transcription

Genetic studies have indicated that wild-type da gene function is required for a number of very different developmental processes. Maternal expression of da in the germ line is required by female progeny very early in development for proper sex determination. In contrast, zygotic expression of da is required by both sexes—early, for development of the PNS, and later, for the growth and differentiation of cuticular tissues and for the proper functioning of the somatic component of the ovary. We wondered whether this functional complexity might be reflected in changes in the relative proportions of the two da transcripts during development, changes that could suggest different developmental roles for the two species.

The developmental profile of da gene expression is shown in Figure 5B. Both the 3.2- and 3.4-kb transcripts are present at every stage of development in approximately constant proportions. Embryos at 0–2.5 hr represent one potentially intriguing exception to this pattern. During this stage, the smaller RNA may be overrepresented relative to the pattern at all other stages, possibly suggesting an enrichment of this smaller message in maternally supplied RNA. Further analysis will be required to confirm this possibility.
da cDNAs: characterization and sequence

We isolated seven da cDNA recombinants from a 0- to 4-hr embryonic Drosophila cDNA library [Bopp et al. 1986]. Southern blot hybridization to genomic fragments suggested that all seven cDNAs represent RNAs originating from sequences between coordinates −0.1 and +6.5 [see Fig. 2]. This region encodes the 3.2- and 3.4-kb transcripts identified as da products by Northern blot analysis. When three of the cDNAs (1.0, 1.2, and 3.0 kb) were used to probe Northern blots, only the 3.2- and 3.4-kb wild-type transcripts were detected. Moreover, these same cDNAs also labeled the abnormal transcripts present in dal/da I flies (data not shown).

The complete nucleotide sequence of the 3.0 kb-cDNA (MN6) was determined on both strands and is presented in Figure 6. There is an ATG at position 226 of this cDNA which begins a 2130-base-long open reading frame [ORF]. [No other ORF exceeds several hundred bases.] The nucleotide sequence surrounding this ATG is a good match to the consensus sequence PUNNAUGG for translation initiation [Kozak 1984]. At the 5' end of MN6, there would be a 226-base untranslated leader. This cDNA, however, may be incomplete at its 5' end. Partial end sequencing of another, smaller cDNA suggests that the actual leader sequence may extend for another 50–100 bases. The precise location of the insertions present in da+ and da− was determined by sequencing [single-stranded] cloned genomic fragments containing the inserted elements. Both insertions are located in what would correspond to the leader sequence of MN6 [following bp 104 and 138, respectively; see arrows in Fig. 6].

The genomic sequence of the da+ region has been determined on one strand from −0.5 to +4.5 [see Fig. 4]. A comparison between the cDNA MN6 and this preliminary genomic sequence indicates that the mRNA represented by MN6 would be derived from a precursor with only a single, 1.5-kb intron that would be located between positions 213 and 214 of the cDNA [data not shown]. This intron is 13 bp upstream of the putative protein coding region, thus, the entire protein coding region of this cDNA appears to be colinear with the genomic sequence. With respect to the genomic map [see Fig. 2], transcription is directed from 0 toward +5, and the transcription unit would extend over a region that could be <5 kb in length [the size of the 3.0-kb cDNA MN6 plus a 1.5-kb intron plus the small additional amount of mRNA sequence not represented in MN6].

At the 3' end of the cDNA there is a 400-base stretch that contains numerous stop codons. Potential poly[A] addition sites [AATAAA] [Proudfoot and Brownlee 1976] are located at positions 2871 and 2926; however, this cDNA lacks a poly[A] tail, suggesting that the true 3' end of the message is missing from MN6.

The long ORF of cDNA MN6 would encode a protein of 710 amino acids with a molecular weight of ~74,000 and a pI of 6.72. The hydropathy plot of the predicted amino acid sequence shows no extended hydrophobic domains [data not shown]. Thus, the product would appear to be a soluble intracellular protein. There are seven potential glycosylation sites, most of which occur in the amino-terminal half of the protein [see Fig. 6].

A similarity search of this amino acid sequence against current sequence databanks [see Methods] revealed that the MN6 da protein contains an amino acid sequence motif that is similar to the His-Pro repeat present in the anterior-group maternal effect gene, bicoid [bcd] [Berleth et al. 1988], and in the segmentation gene, paired [prd] [Frigerio et al. 1986]. This histidine- and proline-rich region begins at residue 224 of the presumptive da protein and exhibits 59% identity with the
protein, far from its His-Pro repeat. It is not part of any previously described sequence motif.

Another similarity between the da and bcd predicted protein sequences is the presence of runs of repetitive glutamines called M (McGinnis et al. 1984) or opa repeats (Wharton et al. 1985), found in many other Drosophila proteins. In addition, da has two runs of glycine (five and seven adjacent) and a run of alanines (five) [see Fig. 6].

Discussion

We generated a P-element-induced mutation in the da locus of Drosophila melanogaster to clone this gene. Several lines of evidence make us confident that we have, in fact, isolated da. First, we have shown that the cloned P-element insertion is indeed responsible for the mutant phenotype of the new dysgenic allele, da<sup>pa</sup>. This P-element insertion invariably is altered in derivatives of da<sup>pa</sup> that have increased or decreased activity. Because essentially null derivatives of the da<sup>pa</sup> allele can be generated by rather modest changes in the inserted transposon and/or its flanking DNA, the P element seems to be near the da gene [the weak phenotype of da<sup>pa</sup> itself might have implied otherwise]. Second, the spontaneous partial-loss-of-function allele da<sup>1</sup> also is associated with an insertion of a middle repetitive element in this region. In fact, the da<sup>1</sup> insertion site is only 34 nucleotides away from that of the P element in da<sup>pa</sup>. Finally, both da<sup>pa</sup> and da<sup>1</sup> cause a similar, very specific disruption of the transcript pattern across the cloned region. Both insertions reduce or eliminate two overlapping mRNA species of 3.2 and 3.4 kb. In addition, both generate two new higher-molecular-weight RNA species that appear to correspond to the two wild-type transcripts.

Characterization of cDNAs from this region and analysis of the corresponding genomic region in both the wild-type and mutant strains place the da<sup>pa</sup> and da<sup>1</sup> insertions in the untranslated leader sequence of the gene. This finding could account for the high level of residual da function in these two mutant alleles: The aberrant transcripts that these alleles generate may serve as templates for the synthesis of wild-type protein product, albeit with reduced efficiency. The RNAs generated by the two mutants could result from initiation at the normal da promoter. Because the junctions of the da<sup>1</sup> and da<sup>pa</sup> insertions do not conform to the eukaryotic splice site consensus sequence, precise excision of the insertions during RNA processing would be unlikely. Alternatively, RNA synthesis might initiate at a site within the insertions. It should be noted that, in either case, the apparently wild-type transcripts found in da<sup>1</sup> would probably have an abnormal structure.

The molecular characterization of these two alleles generally is consistent with their mutant phenotypes: Both are weak mutants with reduced but nonzero levels of the maternal as well as the zygotic da functions. On the other hand, the discovery that the da<sup>1</sup> phenotype is apparently due to an insertion of DNA upstream of the

---

**Figure 5.** Transcriptional analysis of the da locus. (A) Northern blot hybridization of poly(A)<sup>+</sup> RNA from wild-type and mutant da adults and from wild-type 0- to 18-hr embryos (E) probed simultaneously with the 2.7-kb SalI–XhoI and 3.2-kb XhoI fragments from recombinant phage P510 (probes 5 and 6 in Fig. 4). The 1.2-kb transcript covered by one of the probes used here is not affected by da mutations [not shown]. Approximately 5 µg of RNA was loaded per lane. (B) Developmental profile of da gene expression. Northern blot hybridization of wild-type poly(A)<sup>+</sup> RNA from adult female and male flies, embryos, larvae, and pupae, probed with a 1.2-kb da cDNA. Embryonic stages are 0–2.5, 2.5–5, 5–10, 10–15, and 15–24 hr postoviposition. (L1) First larval instar, (L2) second larval instar. Approximately 5–10 µg of RNA was loaded per lane. RNA sizes are indicated in kilobases.

bcd His-Pro repeat with no gaps over a length of 27 residues [Fig. 7A]. If one includes conservative substitutions in the calculation [S-T, G-P], the similarity is 70%. The identity between da and the comparable region of the prd protein is 55% with three gaps introduced [Fig. 7A]. When all three sequences are considered together, there is 40% identity overall. A curious aspect of the da and bcd His-Pro region comparison is that of the five histidine mismatches, four involve leucine substitutions.

The region just upstream of the His-Pro repeat in da is similar to another region of the bcd sequence. Of 27 residues, 11 (40%) are identical between the two sequences [Fig. 7B]. This region is located in the middle of the bcd
Figure 6. Nucleotide and predicted amino acid sequence of the 3.0-kb da cDNA, MN6. Left column numbers refer to amino acids; right column numbers refer to nucleotides. The deduced translation product is shown for the longest open reading frame of the cDNA. The sites of the da* and da† insertion mutations are indicated by black arrowheads; the position of a 1.5-kb intron is marked by the open arrowhead. Potential glycosylation sites [NXS/T] are indicated by lines above each triplet amino acid signal. Shaded areas highlight regions found to share sequence similarity with bcd and prd (see text). Boxed amino acid sequence indicates runs of individual residues (>4 adjacent).
Amino acid identities shared by da, bcd, and prd might suggest that the role of the transcript might have. It is intriguing, however, that the protein-coding region would not lead one to expect that this mutant allele would be temperature sensitive (ts), yet all aspects of the phenotype of da1 have been shown to be ts; indeed, some aspects of the phenotype suggest that this mutant allele might generate a thermally unstable product (Cline 1976, 1980). Perhaps this result is not so surprising in view of the fact that there are so few cases in higher eukaryotes where the ts phenotype of a mutant allele can be ascribed unequivocally to the production of a thermolabile mutant protein product. It will be of interest to learn whether the ts phenotype of da1 might be related to the effects of temperature on the synthesis or functioning of the aberrant transcripts.

Both of the da* transcripts are present in both sexes of adults and at all stages of development. It is not yet clear from the data available whether these two transcripts have the same functions. The presence of a da transcript at all stages of development was anticipated from the previous genetic characterization of the locus which revealed essential zygotic functions during embryonic, larval, and adult stages. However, because only maternal germ line expression of da appears to be relevant to progeny sex determination, one might have expected a transcript that is involved exclusively in this function to be present only in adult females and young embryos. In this connection, it may be relevant that the proportion of the smaller (3.2 kb) RNA relative to the larger is highest in 0- to 2.5-hr embryos. Since little zygotic transcription is likely to have occurred in these very early embryos (Anderson and Lengyel 1979, Edgar and Schubiger 1986), this would imply that the smaller transcript is enriched among maternally-supplied messages. Whether this enrichment is significant with regard to the role of da in the sex determination process will have to be determined.

We do not know the precise structural relationship between the 3.2- and 3.4-kb da* transcripts, or which mRNA species the 3.0-kb cDNA, MN6, represents. A comparison of the MN6 sequence with our data on the genomic sequence indicates that the structure of this da transcript may be relatively simple: There is a single (1.5-kb) intron upstream of the protein-coding region. Analysis of additional cDNAs as well as RNA protection experiments will be required to determine whether the pattern of RNA processing that generates the other transcript is equally simple.

The relative simplicity of the da locus at the molecular level might seem surprising in view of the many developmental functions of this gene. It is worth noting, however, that so far the analysis of the five mutant alleles available has suggested little if any genetic complexity: These mutations appear to fit into a rather straightforward allelic series of increasing severity. Alleles isolated only on the basis of their maternal effect on progeny sex ratio have proven to be defective in their zygotic functions as well, and vice versa. Thus, while the da gene clearly is required for different developmental events in different tissues and at different stages, the results of our preliminary molecular analysis suggest that these various functions may be mediated by the same gene products.

What might those activities be? The best understood of all da functions is its participation in sex determination through the positive activation of Sxl in conjunction with the X/A signal very early in development. It is known that the transcript pattern of Sxl during the first few hours after fertilization is qualitatively different from that just a few hours later [the later pattern persists throughout the lifetime of the fly (Maine et al. 1985a, H. Salz, E. Maine, L. Keyes, M. Samuels, T. Cline, and P. Schell, in prep.). Moreover, the early expression of Sxl appears to be controlled by a promoter different from that used at later stages of development (Salz et al. 1987). It is also known that maternal mutations in da alter this early Sxl transcript pattern, although the precise nature of that alteration is not yet clear (Maine et al. 1985a and unpubl.). Thus, da could act as a transcription factor at Sxl early in development. Schupbach (1985) has suggested that the primary function of da might be to allow Sxl to be expressed at an unusually early stage in development when promoters are not active generally. Alternatively, the da product might not have a direct role in the transcription of Sxl per se, but might, instead, be involved in some other aspect of the process in which the X:A ratio is assessed.

The predicted amino acid sequence for the protein encoded by the da cDNA MN6 is not very informative with respect to the specific biochemical activity that a da protein might have. It is intriguing, however, that the only meaningful similarity we have discovered thus far for this protein sequence is to the predicted product of the bcd gene. bcd appears to encode an anteriorizing determinant whose graded distribution helps establish the spatial polarity of the Drosophila embryo. There are three regions of similarity between da and bcd, including the His-Pro repeat (also found in the segmentation gene prd) and the M repeat (polyglutamine), found in many Drosophila proteins. The functional significance of these regions is unknown. Unlike bcd, however, da has no hint of a homeo box, nor any clear indication of other structural motifs that might suggest a role in transcription or DNA binding. It is tempting to consider that the sequence similarity between da and

Figure 7. Comparison of similar amino acid domains among da, bcd, and prd. (A) Alignment of the His-Pro repeat found in da, bcd, and prd. (B) Comparison of a moderately conserved (40%) amino acid sequence between da and bcd. (Shaded boxes) Amino acid identities shared by da and either bcd or prd, or both.
The mutagenesis scheme that was carried out to generate Df(2L)J-der-27 b/CyO Df(2L)J-der-27 mutagenized chromosome of a lethal mutation within the mated to P-strain males (Hwb~v; Cline 1984). Individual female progeny from this mating were crossed to Hwb~p Df(2L)J-der-27 b/CyO tester males. These matings were scored for the absence of Cy+ progeny, indicating the presence of the mutagenized chromosome of a lethal mutation within the Df(2L)J-der-27 region.

Isolation of daPS

The mutagenesis scheme that was carried out to generate daPS was as follows: M-strain females [cl b pr cn/cl b pr cn] were mated to P-strain males [+/Cy(2L+2R)] derived from the Harwich P strain) at 25°C. The dysgenic hybrid F1 females [cl b l pr cn/Cy(2L+2R)] were then mated to Hwb~p cl b/CyO (M strain) males (Hwb~p; Cline 1984). Individual female progeny [cl b pr cn/CyO] from this mating were crossed to Hwb~p Df(2L)J-der-27 b/CyO tester males. These matings were scored for the absence of Cy+ progeny, indicating the presence of the mutagenized chromosome of a lethal mutation within the Df(2L)J-der-27 region.

Methods

Fly stocks

Flies were raised at 25°C on a standard Drosophila medium described in Cline [1978]. daP183 was obtained from L. Jan [see Caudy et al. 1988]. daP and daP are described in Cronmiller and Cline [1987]. Other mutations and chromosomes not listed in Lindsley and Grell (1968) or in Lindsley and Zimm [1987] are referenced in the text.

DNA sequencing and computer analysis

Genomic fragments and cDNAs were subcloned into Bluescript (Stratagene Cloning Systems, San Diego, California) vectors for sequencing. Nested deletions of these subclones were prepared by exonuclease III digestion [Henikoff 1984]. Both strands of cDNA MN6 were sequenced by the dideoxynucleotide chain-termination method (Sanger et al. 1977), using Sequenase and the Sequenase Sequencing Kit (U.S. Biochemicals). In some cases, specific synthetic oligonucleotide primers were used. Sequence assembly and analysis was accomplished by the computer programs DBUTIL, ANALYSEQ, and DIAGON (Staden 1980, 1982a,b, 1984). The NBRF Protein Identification Resource database (accessed through Bionet) and the translated GenBank library (provided by D. Welsh, Princeton University) were searched for sequence similarities using the algorithm of Lipman and Pearson [1985].

Acknowledgments

We are grateful to Dr. Markus Noll for the gift of the cDNA library and to Drs. Leslie Bell and John Yochem for help with the sequencing techniques and computer analysis. We also thank Dr. Iva Greenwald for helpful comments on the manuscript. This work was supported by grants from the National Institutes of Health and the March of Dimes to T.W.C. and P.S., and by an National Institutes of Health postdoctoral fellowship to C.C.

In situ hybridization and Southern blotting

In situ hybridizations to polytene chromosomes were done as described by Bingham et al. [1981]. Double-stranded DNA probes were labeled by nick-translation using [3H]TTP from ICN. Specific conditions for genomic Southern blots are described by Maine et al. [1985b].

Recombinant DNA library construction and screening

Whole genomic DNA was isolated from adult flies as described by Ish-Horowicz et al. [1979]. Recombinant libraries were constructed using genomic DNA isolated from daPS, daP [Oregon R], or daP flies and phage EMBL3 or EMBL4 vector DNA [Murray 1983]. Genomic DNA was digested with either MboI (partial: daP, daP) or BamHI (complete: daP) to produce 15- to 20-kb fragments, then ligated to BamHI-cut EMBL3 or EMBL4 vector DNA. Ligated DNA was in vitro-packaged into T7 phage particles as described in Maniatis et al. [1982]. The daP recombinant library was screened with 32P-labeled [ICN] nick-translated plasmid DNA containing an intact P element (p. 25.1: Bingham et al. 1981).

In situ hybridization and Southern blotting

In situ hybridizations to polytene chromosomes were done as described by Bingham et al. [1981]. Double-stranded DNA probes were labeled by nick-translation using [3H]TTP from ICN. Specific conditions for genomic Southern blots are described by Maine et al. [1985b].

Molecular analysis of more extreme derivatives of daP

The DNA rearrangements associated with the more extreme derivatives of daPS were analyzed by genomic Southern blotting. Blots containing EcoRI-digested genomic DNA from wild-type [cl b pr cn], daPS, and the three more extreme derivatives of daPS were probed with recombinant phage PS10 or with the 2.7-kb SalI–XhoI fragment from PS10 [probe 5 in Fig. 4] to ascertain the sizes of the insertions retained in the more extreme alleles. In addition, Southern blots containing EcoRI- and HindIII-restricted genomic DNA were probed with the 2.7-kb SalI–XhoI fragment mentioned above. Loss of the P element and/or flanking host sequences in a more extreme daPS derivative was signaled by a decreased size of the fragment containing the insertion, together with retention of the HindIII site of the P element.

The HindIII site of the P element of daPS was used as a marker (relative to the HindIII and EcoRI sites at -0.6 and +0.8, respectively) to determine the orientation of the P-element inserts. In this way, one more extreme daPS derivative was found to be an inversion of the P-element insertion with no associated loss of insert or host sequences. In the third more extreme daPS derivative, loss of the P-element HindIII site indicated deletion of some P-element sequences.

RNA isolation and Northern blotting

Poly[A]+ RNA was prepared by the method of O'Hare et al. [1983]. In general, 5–10 μg of poly[A]+ RNA was loaded per lane. RNA was size-fractionated on denaturing 0.7% formaldehyde–agarose gels, using a MOPS buffer [Maniatis et al. 1982]. RNA was transferred to nitrocellulose in 10× SSPE. Filters were prehybridized and hybridized using the same conditions employed for Southern blotting. 32P-Labeled probes were made either by nick-translation or by random primer extension [Feinberg and Vogelstein 1983]. The da RNAs were sized by probing Northern blot filters with probes specific for armadillo [3.2 kb; R. Riggleman, E. Wieschaus, and P. Schedl, in prep.].

Acknowledgments

We are grateful to Dr. Markus Noll for the gift of the cDNA library and to Drs. Leslie Bell and John Yochem for help with the sequencing techniques and computer analysis. We also thank Dr. Iva Greenwald for helpful comments on the manuscript. This work was supported by grants from the National Institutes of Health and the March of Dimes to T.W.C. and P.S., and by an National Institutes of Health postdoctoral fellowship to C.C.
Note

Sequence data described in this paper have been submitted to the EMBL/GenBank Data Libraries under accession number Y00221.

References

Anderson, K.V. and J.A. Lengyel. 1979. Rates of synthesis of major classes of RNA in Drosophila embryos. Dev. Biol. 70: 217–231.

Baker, B.S., R.N. Nagoshi, and K.C. Burtis. 1987. Molecular genetic aspects of sex determination in Drosophila. Bioessays 6: 66–70.

Bell, A.E. 1954. A gene in Drosophila melanogaster that produces all male progeny. Genetics 39: 958–959.

Berleth, T., M. Burri, G. Thoma, D. Bopp, S. Richstein, G. Frigerio, M. Noll, and C. Nüsslein-Volhard. 1988. The role of localization of bicoid RNA in organizing the anterior pattern of the Drosophila embryo. EMBO J. 7: 1749–1756.

Bingham, P.M., R. Levis, and G.M. Rubin. 1981. Cloning of DNA sequences from the white locus of D. melanogaster by a novel and general method. Cell 25: 693–704.

Bopp, D., M. Burri, S. Baumgartner, G. Frigerio, and M. Noll. 1986. Conservation of a large protein domain in the segmentation gene paired and in functionally related genes of Drosophila. Cell 47: 1033–1040.

Caudy, M., E.H. Grell, C. Dambly-Chaudiere, A. Ghysen, L.Y. Jan, and Y.N. Jan. 1988. The maternal sex determination gene daughterless has zygotic activity necessary for the formation of peripheral neurons in Drosophila. Genes Dev. 2: 843–852.

Cline, T.W. 1976. A sex-specific, temperature-sensitive maternal effect of the daughterless mutation of Drosophila melanogaster. Genetics 84: 723–742.

— 1978. Two closely linked mutations in Drosophila melanogaster that are lethal to opposite sexes and interact with daughterless. Genetics 90: 683–698.

— 1979. A male-specific lethal mutation in Drosophila melanogaster that transforms sex. Dev. Biol. 72: 266–275.

— 1980. Maternal and zygotic sex-specific gene interactions in Drosophila melanogaster. Genetics 96: 903–926.

— 1983. The interaction between daughterless and Sex-lethal in triploids: a lethal sex-transforming maternal effect linking sex determination and dosage compensation in Drosophila melanogaster. Dev. Biol. 95: 260–274.

— 1984. Autoregulatory functioning of a Drosophila gene product that establishes and maintains the sexually determined state. Genetics 107: 231–277.

— 1985. Primary events in the determination of sex in Drosophila melanogaster. In Origin and evolution of sex (ed. H.O. Halvarson and A. Monroy) pp. 301–327. Alan R. Liss, New York.

— 1988a. Exploring the role of the gene, Sex-lethal, in the genetic programing of Drosophila sexual dimorphism. In Evolutionary mechanisms in sex determination, CRC Uniscience series (ed. S.S. Wachtel). CRC Press, Boca Raton, Florida. [in press.]

— 1988b. Evidence that sisterless-a and sisterless-b are two of several discrete ‘numerator elements’ of the X/A sex determination signal in Drosophila that switch Sex between two alternative stable expression states. Genetics 119: 829–862.

Crommiller, C. and T.W. Cline. 1986. The relationship of relative gene dose to the complex phenotype of the daughterless locus in Drosophila. Dev. Genet. 7: 205–221.

— 1987. The Drosophila sex determination gene daughterless has different functions in the germline vs. the soma. Cell 48: 479–487.

Doe, C.Q., D. Smouse, and C.S. Goodman. 1988. Control of neuronal fate by the Drosophila segmentation gene, even-skipped. Nature 333: 376–378.

Edgar, B.A. and G. Schubiger. 1986. Parameters controlling transcriptional activation during early Drosophila development. Cell 44: 871–877.

Feinberg, A.P. and B. Vogelstein. 1983. A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. Biochemistry 132: 6–13.

Frigerio, G., M. Burri, D. Bopp, S. Baumgartner, and M. Noll. 1986. Structure of the segmentation gene paired and the Drosophila PRD gene set as part of a gene network. Cell 47: 735–746.

Gergen, J.P. 1987. Dosage compensation in Drosophila: evidence that daughterless and Sex-lethal control X chromosome activity at the blastoderm stage of embryogenesis. Genetics 117: 477–485.

Henikoff, S. 1984. Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. Gene 28: 351–359.

Ish-Horowicz, D., S.M. Pinchin, P. Schell, S. Artavanis-Tsakonas, and M.E. Mirault. 1979. Genetic and molecular analysis of the 87A7 and 87C1 heat-inducible loci of D. melanogaster. Cell 18: 1351–1358.

Kozak, M. 1984. Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs. Nucleic Acids Res. 12: 857–871.

Lindsley, D.L. and E.H. Grell. 1968. Genetic variations of Drosophila melanogaster. Carnegie Inst. Wash. Publ. 627.

Lindsley, D.L. and G. Zimm. 1987. The genome of Drosophila melanogaster, Part 3: rearrangements. Drosophila Inform. Serv. 65.

Lipman, D.J. and W.R. Pearson. 1985. Rapid and sensitive protein similarity searches. Science 227: 1435–1441.

Lucchesi, J.C. and J.E. Manning. 1987. Gene dosage compensation in Drosophila melanogaster. Adv. Genet. 24: 371–429.

Maine, E.M., H.K. Salz, P. Schell, and T.W. Cline. 1985a. Sex-lethal, a link between sex determination and sexual differentiation in Drosophila melanogaster. Cold Spring Harbor Symp. Quant. Biol. 50: 595–604.

Maine, E.M., H.K. Salz, T.W. Cline, and P. Schell. 1985b. The Sex-lethal gene of Drosophila: DNA alterations associated with sex-specific lethal mutations. Cell 43: 521–529.

Mange, A.P. and L. Sandler. 1973. A note on the maternal effect mutants daughterless and abnormal oocyte in Drosophila melanogaster. Genetics 73: 73–86.

Maniatis, T., E.R. Fritsch, and J.Sambrook. 1982. Molecular cloning: A laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

McGinnis, W., M.S. Levine, E. Hafen, A. Kuroiwa, and W.J. Murray, N.E. 1983. Phage lambda and molecular cloning. In Genetics of Drosophila: Antennapedia, bithorax and other complexes. Nature 308: 428–433.

Murray, N.E. 1983. Phage lambda and molecular cloning. In Lambda II (ed. R.W. Hendrix, J.W. Roberts, R.W. Stahl, and R.A. Weissberg), pp. 395–432. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

O'Hare, K., R. Levis, and G. Rubin. 1983. Transcription of the white locus in Drosophila melanogaster. Proc. Natl. Acad. Sci. 80: 6917–6921.

Pimpinelli, S., W. Sullivan, M. Prout, and L. Sandler. 1985. On biological functions mapping to the heterochromatin of Drosophila melanogaster. Genetics 109: 701–724.

Proudfoot, N. and G. Brownlee. 1976. 3' non-coding region se-
quences in eukaryotic messenger DNA. Nature 263: 211–213.

Salz, H.K., T.W. Cline, and P. Schedl. 1987. Functional changes associated with structural alterations induced by mobilization of a P element inserted in the Sex-lethal gene of Drosophila. Genetics 117: 221–231.

Sanchez, L. and R. Nothiger. 1983. Sex determination and dosage compensation in Drosophila melanogaster: production of male clones in XX females. EMBO J. 2: 485–491.

Sandler, L. 1972. On the genetic control of genes located in the sex-chromosome heterochromatin of Drosophila melanogaster. Genetics 70: 261–274.

———. 1975. Studies on the genetic control of heterochromatin in Drosophila melanogaster. Israel J. Med. Sci. 11: 1124–1134.

———. 1977. Evidence for a set of closely linked autosomal genes that interact with sex-chromosome heterochromatin in Drosophila melanogaster. Genetics 86: 567–582.

Sanger, F., S. Nicklen, and A.R. Coulson. 1977. DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. 74: 5463–5467.

Schupbach, T. 1985. Normal female germ cell differentiation requires the female X chromosome to autosome ratio and expression of Sex-lethal in Drosophila melanogaster. Genetics 109: 529–548.

Staden, R. 1980. A new computer method for the storage and manipulation of DNA gel reading data. Nucleic Acids Res. 8: 3673–3694.

———. 1982a. An interactive graphics program for comparing and aligning nucleic acid and amino acid sequences. Nucleic Acids Res. 10: 2951–2961.

———. 1982b. Automation of the computer handling of gel reading data produced by the shotgun method of DNA sequencing. Nucleic Acids Res. 10: 4731–4751.

———. 1984. Graphic methods to determine the function of nucleic acid sequences. Nucleic Acids Res. 12: 521–538.

Tsubota, S. and P. Schedl. 1986. Hybrid dysgenesis-induced revertants of insertions at the 5′ end of the rudimentary gene in Drosophila melanogaster: transposon-induced control mutations. Genetics 114: 165–182.

Wharton, K.A., B. Yedvobnick, V.G. Finnerty, and S. Artavanis-Tsakonas. 1985. opa: a novel family of transcribed repeats shared by the Notch locus and other developmentally regulated loci in D. melanogaster. Cell 40: 55–62.
Molecular characterization of daughterless, a Drosophila sex determination gene with multiple roles in development.

C Cronmiller, P Schedl and T W Cline

*Genes Dev.* 1988, 2:

Access the most recent version at doi:10.1101/gad.2.12a.1666