Development 120, 1561-1572 (1994)
Printed in Great Britain © The Company of Biologists Limited 1994

Dax, a locust Hox gene related to *fushi-tarazu* but showing no pair-rule expression

Rachel Dawes, Iain Dawson*, Francesco Falciani, Guy Tear† and Michael Akam

Wellcome/CRC Institute and Department of Genetics, Tennis Court Road, Cambridge, CB2 1QR, UK

*Present address: Department of Biology, Yale University School of Medicine, Boyer Centre for Molecular Medicine, 295 Congress Avenue, New Haven, Connecticut 06536 0812, USA
†Present address: Howard Hughes Medical Institute, Department of Molecular and Cell Biology, University of California, Berkeley, California 94720, USA

SUMMARY

We describe an unusual Antennapedia class homeobox gene from the grasshopper *Schistocerca gregaria* (Orthoptera, African Plague Locust). Its sequence is not sufficiently similar to that of any other insect Hom-Hox gene to identify it unambiguously, but short conserved elements suggest a relationship to the segmentation gene *fushi-tarazu*, (*ftz*). We term it *Sg Dax* (divergent Antennapedia class homeobox gene). Antibodies raised against the protein encoded by this gene reveal two phases of expression during embryogenesis. In the early embryo, it is a marker for the posterior part of the forming embryonic primordium, and subsequently for the posterior part of the growing germ band. In older embryos, it labels a subset of neural precursor cells in each trunk segment, very similar to that defined by the expression of *fushi tarazu* (*ftz*) in *Drosophila*. We suggest that *Schistocerca Dax* and *Drosophila ftz* are homologous members of a gene family whose members are diverging relatively rapidly, both in terms of sequence and role in early development.

Key words: homeobox, homeotic gene, *fushi tarazu*, short germ insect, grasshopper, locust

INTRODUCTION

Among the arthropods, homeotic genes of the Antennapedia (Antp) class are generally well conserved. Homeobox exons of these genes have been sequenced in a range of insects (Walldorf et al., 1989; Tear et al., 1990; Nagy et al., 1991; Eggleston, 1992; Ueno et al., 1992; Stuart et al., 1993), and in a single crustacean (Averof and Akam, 1993). In virtually all cases, these sequences can be identified as unique homologues (orthologues) of specific *Drosophila* genes, with near-identical homeodomains (>58/60 amino acids) and conserved flanking sequences. In the few cases where it has been tested, orthologues are expressed in similar spatial domains in different insects, implying a conserved role in development (Tear et al., 1990; Nagy et al., 1991; Stuart et al., 1993).

We report here the isolation of a gene from the Locust, *Schistocerca gregaria*, that does not fit this pattern. On the basis of homeodomain sequence, this gene is clearly of the Antp class (Scott et al., 1989), but resembles no particular *Drosophila* gene more closely than it does the Antp class consensus. It has a pattern of expression during early embryogenesis unlike that of any *Drosophila* Antp class gene, but it's expression during neurogenesis is very similar to that of the *Drosophila* gene *fushi tarazu* (Carroll and Scott, 1985; Doe et al., 1988a).

Fushi tarazu (*ftz*) is an exception among the set of *Drosophila* Antp class genes. Like all other Antp class genes, it is located in a Hom/Hox cluster, in this case between the homeotic genes *Sex combs reduced* (*Scr*) and *Antennapedia* (*Antp*) of the Antennapedia complex (Wakimoto and Kaufman, 1981). However, it does not contain a canonical YPWM motif characteristic of these Hox genes, and is not generally considered to be a homeotic gene. It is expressed in a pattern of stripes in the *Drosophila* blastoderm (Hafen et al., 1984; Carroll and Scott, 1985), where it is needed to define the pattern of segments (Wakimoto et al., 1984) and to establish the pattern of homeotic gene expression (Duncan, 1986; Ingham and Martinez-Arias, 1986). Later, in the nervous system, it is expressed in the same subset of neural cells in each segment; its role here appears to be the specification of cell type, not of segment differences (Carroll and Scott, 1985; Hiromi et al., 1985; Doe et al., 1988a).

Of these two processes in which the *ftz* gene is involved, neurogenesis has been well conserved in evolution, but segment formation differs markedly in different insect groups. The pattern of neural precursor cells, and the initial fate of their progeny, is very similar in insects as different as flies and locusts, allowing homologous cells to be identified unambiguously (Thomas et al., 1984). However, there is no simple way to equate the stages of early development in these insects prior to the formation of the segmented germ band. Flies (Diptera), are 'long germ' insects. They specify the complete pattern of segments at the syncytial blastoderm stage, virtually simultaneously. Grasshoppers and locusts (Orthoptera) are 'short germ' insects. They form a cellular embryonic primordium, which becomes segmented only as it grows. This short-germ mode of development is shared with the less specialised

members of other arthropod groups, and is therefore likely to be the primitive mode of embryogenesis for the phylum (Sander, 1976; Patel et al., 1994).

We suggest that the gene that we have isolated is a homologue of *ftz*, but that this gene family is diverging relatively rapidly, both in terms of sequence and role in early development. The expression of the *Schistocerca* gene suggests that it plays a role in the specification of the embryonic axis in this short germ insect.

MATERIALS AND METHODS

Locust embryos

Schistocerca gregaria (Forskål) eggs were kindly supplied by Cambridge University Zoology field station. They were collected overnight, laid in moist sand at 26°C, and then incubated in the laboratory at 28°C. All staging of locust embryos in terms of % development is based on that of Bentley et al. (1979).

Isolation of genomic clones

A *Schistocerca gregaria* partial *Sau* 3A library of genomic DNA (Tear et al., 1990) was screened with short fragments containing the *Drosophila Ubx, Antp* and *Dfd* homeoboxes (Dawson, 1989). Probes were hybridised in 42% deionised formamide, 5× SSPE, 5× Denhardts, 0.1% SDS with 250 μg/ml sheared, denatured salmon sperm DNA at 37°C. Washes were taken to a stringency of 65°C in 2× SSPE. Homeobox homologous fragments were subcloned and sequenced by standard techniques (Dawson, 1989). The Dax homeobox is contained on a 2.7 kb *Eco*RI/*Sal*I genomic fragment (subclone p2313RS2.7 derived from phage λ2313).

RACE PCR

Coding sequence 5′ of the homeobox exon of p2313RS2.7 was obtained by RACE PCR. First strand cDNA was made from 1 μg of poly(A)-enriched RNA isolated from embryos at 35% development, using 12.5 pmol of *Dax*-specific primer RED5 (see below for primer sequences) and Superscript™ RNase H⁻ Reverse Transcriptase (GIBCO BRL). First strand cDNA was size selected on Size Select-400 Spun Columns (Pharmacia) then tailed and used as substrate for RACE PCR, carried out as described by Frohman (1990), except that DMSO was not used.

For the first PCR, reactions containing 1/20th to 1/200th of the first strand were amplified with Adapter primer, dT_{18}-Adapter primer and *Dax*-specific primer RED6. Southern analysis of the first PCR reactions, using a *Dax* homeobox probe, identified positive bands. The largest of these products were gel purified and 1/50th of the removed gel slice was used as template for a second round of RACE PCR (using dT_{18}-Adapter primer and nested *Dax*-specific primer RED203; 40 cycles). A further round of Southern analysis, size selection and PCR (20 rounds) were carried out, using nested *Dax*-specific primer RED3. Bands were subcloned into the *Sma*I site of pBluescript II SK+ (Stratagene).

Three independent plasmids deriving from the final PCR reaction (pA1, pA2 and pA3) containing inserts of 1060 bp were sequenced using the Sequenase® Version 2.0 DNA Sequencing Kit (United States Biochemical). The full sequence of both strands was obtained by primer walking. The sequence of pA1, pA2 and pA3 showed differences at 9 positions over 1 kb. The consensus of these three sequences was taken to reduce PCR errors in the sequence.

Primers

RED5: 5′ ACTTCAGAATGGTAGTCAGGTGTGGCA 3′ (reverse)
RED6: 5′ GAGGGTCGGTCTCTGACGGGATTCGAGT 3′ (reverse)
RED203: 5′ GACCGCGAGACACTCCGGGTCGCGGACGCT 3′ (reverse)
RED3: 5′ GCCGCTCAGTCCATCGCCAA 3′ (reverse)
dT_{18}-Adapter primer: 5′ AGAGAACTAGTGTCGACGCGGC-$CGC(T)_{18}$ 3′ (forward)
Adapter primer: 5′ AGAGAACTAGTGTCGACGCGG 3′ (forward)

In situ hybridisation to embryo sections

Sg Dax probe was prepared by random priming the 2.7 kb *Eco*RI-*Sal*I fragment from p2313.RS2.7 in the presence of [^{35}S]dATP. Hybridization was carried out as described by Tear et al. (1990).

Production of fusion proteins

A 1.42 kb *Pst*I fragment from p2313RS2.7 (Dawson, 1989) contains the homeobox with 21 bp upstream and 1.2 kb downstream. This was subcloned into the *Pst*I site of pUR292 to produce an in frame fusion of the *Sg Dax* homeodomain coding region downstream of the *lacZ* gene to give plasmid pUR292.2313.

A second fusion construct was made by subcloning the 1.48 kb *Bam*HI-*Eco*RI homeobox containing fragment of pUR292.2313 into the *Bam*HI-*Eco*RI sites of pGEX3X (Smith and Johnson, 1988) (gift from S. Russell, Department of Genetics, Cambridge). The resulting plasmid, pGEX3X.2313, consists of the *Sg Dax* homeodomain coding region fused downstream of the glutathione S-transferase gene.

Fusion proteins from both these plasmids were produced in *E. coli* strain TG1 and isolated by SDS-polyacrylamide gel electrophoresis followed by electroelution as described by Tear et al. (1990). Protein was snap frozen in liquid nitrogen and stored at −70°C.

Antibody production

Two rabbits were injected with 200 μg of fusion protein produced from the pGEX3X.2313 construct in complete Freund's adjuvant. Five boosts of 200 μg protein in incomplete Freund's adjuvant were given to each rabbit at 2 week intervals. All injections were subcutaneous at multiple sites. Blood was taken from the rabbits after 5, 7, 9 and 11 weeks and sera collected and stored with 0.02% sodium azide at −20°C (Harlow and Lane, 1988).

Sera were tested for the ability to detect both the β-galactosidase and glutathione S-transferase fusion proteins on western blot strips made from protein extracts of *E. coli* expressing the fusion protein. Reaction to the fusion protein was seen in both rabbits after 5 weeks.

Purification of antisera

The sera were affinity purified against the second fusion protein (pUR292.2313) to select for antibodies that bind specifically to the *Sg Dax* fragment of the protein. 400 μg of fusion protein produced from plasmid pUR292.2313 was dialysed (room temperature, 3 hours) against 0.1 M Hepes-KOH, pH 7.5, bound to Affi-Gel 10 gel (Bio-Rad) and transferred to a column. Sera from the 11th week bleed of both rabbits were passed over the column and bound antibodies were eluted with 0.1 M glycine-HCl, pH 2.5 and neutralised with 0.1 volume 2 M Tris-HCl, pH 7.4. Bovine serum albumin (BSA) was added to 1% and the serum dialysed against PBS (overnight, 4°C) and stored at 4°C with 0.02% sodium azide.

Immunohistochemistry

Locust embryos were dissected from their eggs under locust embryo saline (150 mM NaCl, 3 mM KCl, 5 mM TES (Sigma), 2 mM $CaCl_2$, 1 mM $MgSO_4$) and fixed in 3.7% formaldehyde in PBS at 0°C for 40 minutes. Embryos were then washed in PTx (PBS+1% Triton X-100), blocked in PBTx (PTx+0.1% BSA) at 4°C for 2-4 hours and agitated at 4°C overnight in a 1:50 dilution of affinity-purified serum in PBTx. After the primary incubation, the serum was washed off with three changes of PBTx at room temperature over 3 hours and left in a 1:400 dilution of biotin-conjugated goat anti-rabbit antibody (Vector Laboratories, Peterborough, England) in PBTx, at 4°C on a rotating platform for 90 minutes. The embryos were then washed in three changes of PT (PBS+0.1% Tween 20) over 30 minutes at room temperature and incubated in avidin-biotin complex (ABC Elite, Vector

A

```
                             30                              60                              90                             120
CCAATACAGTACCAC ACTTACCAGGCCCAA AGGTACCACCAGCAA GAAGTCGGAGACCAA CAACAGGGGCATCAA CTGTCCAGTGTCACC CACCATCAACAAAAC AGGGCAATGTACCAA

                            150                             180                             210                             240
TCTCACGACTTGGGG ACCTATGAGAATCAC CACGCCCCACCAGCC TCAGACAGTCAGCAT CTGCCACAGACCCAA CAACCGAGACATCTT CCAGCTCCCTCTACG TATGCATCCCAGTTT
                        *
                            270                             300                             330                             360
CCGACGAGGCCACCC TCTGGTGGAAGAAAT GGGAACGCAAATAAC GCTGGCGCCTTCGAC ATCAGGTACGCCACC AACTTCTTTGCCCAA ATGCATTACGGCGCT TCCAGCAATGGTGAA
                                            *                                                                               *
                            390                             420                             450                             480
TGTGCGGCAACGCCA GTAGCTCCGTTTCCC CCTCTGGAGGACAGG ACTACGGAGAAGCTC ATAACGTAAATAACA ACTGTCCATCAACTG AAACTCCACCGAATG TTTATCCTCAATCAG
M  C  G  N  A   S  S  S  V  S   P  S  G  G  Q   D  Y  G  E  A   H  N  V  N  N   N  C  P  S  T   E  T  P  P  N   V  Y  P  Q  S

                            510                             540                             570                             600
CAGCGAGGGTCAGCA ATAAAGTGGACATAC CCGAGGCTCCAATCG GAATGGCGATGAAGA CATCAACATTATCAG AGAGATCAGTGAGCG GGAACGTTAGTCCAC CTCACCTTGTAGGAC
A  A  R  V  S   N  K  V  D  I   P  E  A  P  I   G  M  A  M  K   T  S  T  L  S   E  R  S  V  S   G  N  V  S  P   P  H  L  V  G

                            630                             660                             690                             720
AAGCATCATTTAGGG ATCACTTCCACAACA TGCAGCGTTTTGGAG AGGCCTTTTCGGATC TTCAAACTATCAGAA GAAACGACGATGGGC CCAGCAACTGCCACA TGAACTCACCACCAT
Q  A  S  F  R   D  H  F  H  N   M  Q  R  F  G   E  A  F  S  D   L  Q  T  I  R   R  N  D  D  G   P  S  N  C  H   M  N  S  P  P

                            750                             780                             810                             840
CGGCGGCGACTGCGA CTTTTGCGACGTCTT TCCCGAAGGTGCTAT ACAGTCCAGAGACGT TCCCAAAAGAAGAAA ATAATTCAAATCCCG AAGTCAGCTCTTACC CAGACGGCACGAATA
S  A  A  T  A   T  F  A  T  S   F  P  K  V  L   Y  S  P  E  T   F  P  K  E  E   N  N  S  N  P   E  V  S  S  Y   P  D  G  T  N

                            870                             900                             930                          ▽  960
CGACGTCAACCCCAT CGCCGCCTGTTCAAG ACCGTTGCCCGCCCT TACCAAGAGTGGAGA TGGATGTCCGACCGC AGTGCGCTTTTTTCT ACCCGTGGATGAAGA CACAATCCGGTGAGA
T  T  S  T  P   S  P  P  V  Q   D  R  C  P  P   L  P  R  V  E   M  D  V  R  P   Q  C  A  F  F   Y  P  W  M  K   T  Q  S  G  E

                            990                            1020                            1050                            1080
TTCCAGGACGGGAGA CAGCCAAACGCACCC GTCAGACGTACACTC GCTTCCAGACGCTGG AGCTGGAGAAGGAGT TCCACTACAACCGCT ACCTGACTCGCCGCC GACGAGTGGAGATCT
I  P  G  R  E   T  A  K  R  T   R  Q  T  Y  T   R  F  Q  T  L   E  L  E  K  E   F  H  Y  N  R   Y  L  T  R  R   R  R  V  E  I

                           1110                            1140                            1170                            1200
CGCAGGCGCTGGGCC TCACAGAGCGCCAGA TCAAGATCTGGTTCC AGAACCGGCGCATGA AGCTCAAGAAAGAGA ATGGTGGCACGCTAC CTGCCCCACTCGTGA GCTTAGGGCAGTCTC
S  Q  A  L  G   L  T  E  R  Q   I  K  I  W  F   Q  N  R  R  M   K  L  K  K  E   N  G  G  T  L   P  A  P  L  V   S  L  G  Q  S

                           1230                            1260                            1290                            1320
TGGCTGGGAGAGTCC CCATGCTCTCCATGG GGCTGCCCAGATCGT CGCCATCGTCTTCAC CTCTGAGGTCCTCGC CAGACGCTGGCTCTG ACGGTGTGGACTGAT GGTAAGACTTCAGTC
L  A  G  R  V   P  M  L  S  M   G  L  P  R  S   S  P  S  S  S   P  L  R  S  S   P  D  A  G  S   D  G  V  D  *

                           1350                            1380                            1410                            1440
TGTCCCAGTTGTGAC TCTGGAGTGGTCCCT GGTCACTGTGCCCAC ACTGGGGCCTCTGGC ATTGCAATCATCAAC ATCGTCTTCACCTCA GACATTGGTATGAAC TGTTAGCAAGTCCAC

                           1470
ACTATGTGGAACCCT GACGAAGGCCCTGTC GCGGAAGGACTCTTG AGTGTGCTCTGAT
```

B

```
                                                                                                  Match within
                                                  Homeodomain                                      homeodomain
                          1                                                         60

                 ▽
    Sg Dax :     GEIPGRET AKRTRQTYTRFQTLELEKEFHYNRYLTRRRRVEISQALGLTERQIKIWFQNRRMKLKKEN GGTLPAPLV

                          ▽
     Dm Antp:             R--G------Y----------F---------I--AH--C----------------W---- KTKGEPGSG      51/60
                     ▽
     Dm abdA:        PNGCP RR-G-----------------F-H-------I--AH--C--------------------L RAVKEINEQ      51/60
                      ▽
     Dm  Ubx:         TNGL RR-G------Y----------T-H-------I-MAH--C--------------------I QAIKELNEQ      48/60
                   ▽
     Dm  Scr:      TVNANGE T--Q-TS---Y----------F---------I--AH--C----------------W---H KMASMNIVP      48/60
                       ▽
     Dm  ftz:          CKD S---------Y----------F---I-----ID-AN--S-S--------------S--DR TLDSSPEHC      47/60

     Antp class
     consensus:           ....R..Y.R.QTLELEKEFH.N.Y.TRRR..........ERQIKIWFQNRRMK.KK..                 35/35
```

C

```
                                                                Homeodomain
                                    1                                                            60

                             ▽
    Sg Dax:    PQCAFFYPWMKTQSGEIPGRET  AKRTRQTYTRFQTLELEKEFHYNRYLTRRRRVEISQALGLTERQIKIWFQNRRMKLKKEN  GGTLPA
                             ▽
      Tc ftz:  EPP-N------AHGDSSAT  G  N---------Y----------F-K-------I--AES-R----------------A--DT  KF-EQS
                                 ▽
      Dm ftz:  -NG-GDFN-SHIEETLASDCKD  S---------Y----------F---I-----ID-AN--S-S--------------S--DR  TLDSSP
                             ▽
      Af Hx1:                XITIPMPG  Q---------KY--------L-------V--MD--SK-Q----------------A----  KNETNF
```

Fig. 1. (A) Sequence and conceptual translation of the *Sg Dax* gene. Sequence from 1 to 1003 bp was obtained by RACE PCR and from 955 to 1498 bp from genomic clone p2313RS2.7 (see Materials and Methods).The EMBL database accession number for this sequence is X73982. Amino acids of the homeodomain are double underlined and the FYPWMK peptide is single underlined. The genomic clone contains an intron, denoted by ∇, 24 bp upstream from the homeobox. In frame stop codons upstream and downstream of the presumed coding region are indicated by *. (B) Comparison of the amino acid sequence of the homeodomain and flanking regions of *Sg Dax* with the *Drosophila* Antp class proteins (Schneuwly et al., 1986; Kornfeld et al., 1989; Scott et al., 1989; Karch et al., 1990; Laughon and Scott, 1984; LeMotte et al., 1989). ∇ indicates the position of the 5′ splice acceptor in each of the genes. First and last residues of the homeodomain are numbered.
(C) Comparison of the amino acid sequence of the homeodomain and flanking regions of *Sg Dax* protein with the *ftz* proteins of *Tribolium* (Tc) and *Drosophila* (Dm) and the *Artemia Af Hx1* protein (Brown and Denell, personal communication; Laughon and Scott, 1984; Averof, 1993). Two spaces have been introduced into the *Tribolium* sequence for alignment. Only the sequence from the homeobox containing exon is available for *AfHx1*. ∇ denotes the position of an intron. First and last residues of the homeodomain are numbered. Note the conservation in the N terminus of the homeodomain of these proteins from position 2 to 10. Outside the homeodomain the sequence AFYPWMK is conserved between *Tc ftz* and *Sg Dax* proteins. In *Drosophila ftz* protein this region shows conservation of the A and W amino acids. The sequence of this region of *Af Hx1* is unknown.

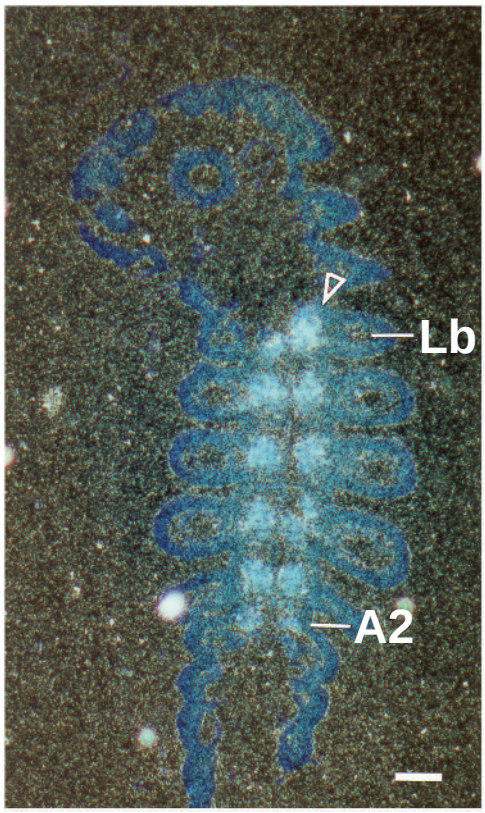


Fig. 2. In situ hybridisation of *Sg Dax* sequence to *Schistocerca* embryo (longitudinal section). Scale bar 100 μm. *Sg Dax* hybridises (arrowhead) to neuromeres in the centre of each segment (labial (Lb) to second abdominal (A2) visible in this section), but not to lateral epidermis. Embryo at 30% development. Dark-field optics.

Laboratories) for 30 minutes at room temperature. Excess ABC complex was washed off with five changes of PT in 30 minutes. They were stained in 0.25 mg/ml diaminobenzidine (Sigma), 0.1% hydrogen peroxide (from 30% solution, Sigma) and 0.03% nickel chloride in PT. Embryos were dehydrated in ethanol, cleared in xylene and mounted in DPX mountant (BDH Ltd., Poole, England).

Whole *Schistocerca* eggs were placed in 3.7% formaldehyde in PBS and pricked with a very fine needle (15 μm diameter) about 10 times through the chorion, avoiding the most posterior region where the embryo will develop. The eggs were left in formaldehyde overnight at 4°C. The following day, the chorion was peeled off the eggs, which were then washed in PBTx for 2 hours. Thereafter the whole eggs were treated in the same way as dissected embryos that have been fixed and blocked.

Drosophila embryos were dechorionated in bleach, fixed in 4% paraformaldehyde, devitellinised with methanol, washed in PBTx and then treated in the same way as the fixed *Schistocerca* embryos. The primary antibody was a mouse monoclonal anti-*ftz* antibody (Krause et al., 1988) kindly provided by David Ish-Horowicz and was used at a 1:1000 dilution. Biotinylated horse anti-mouse (Vector Laboratories) was used as the secondary antibody.

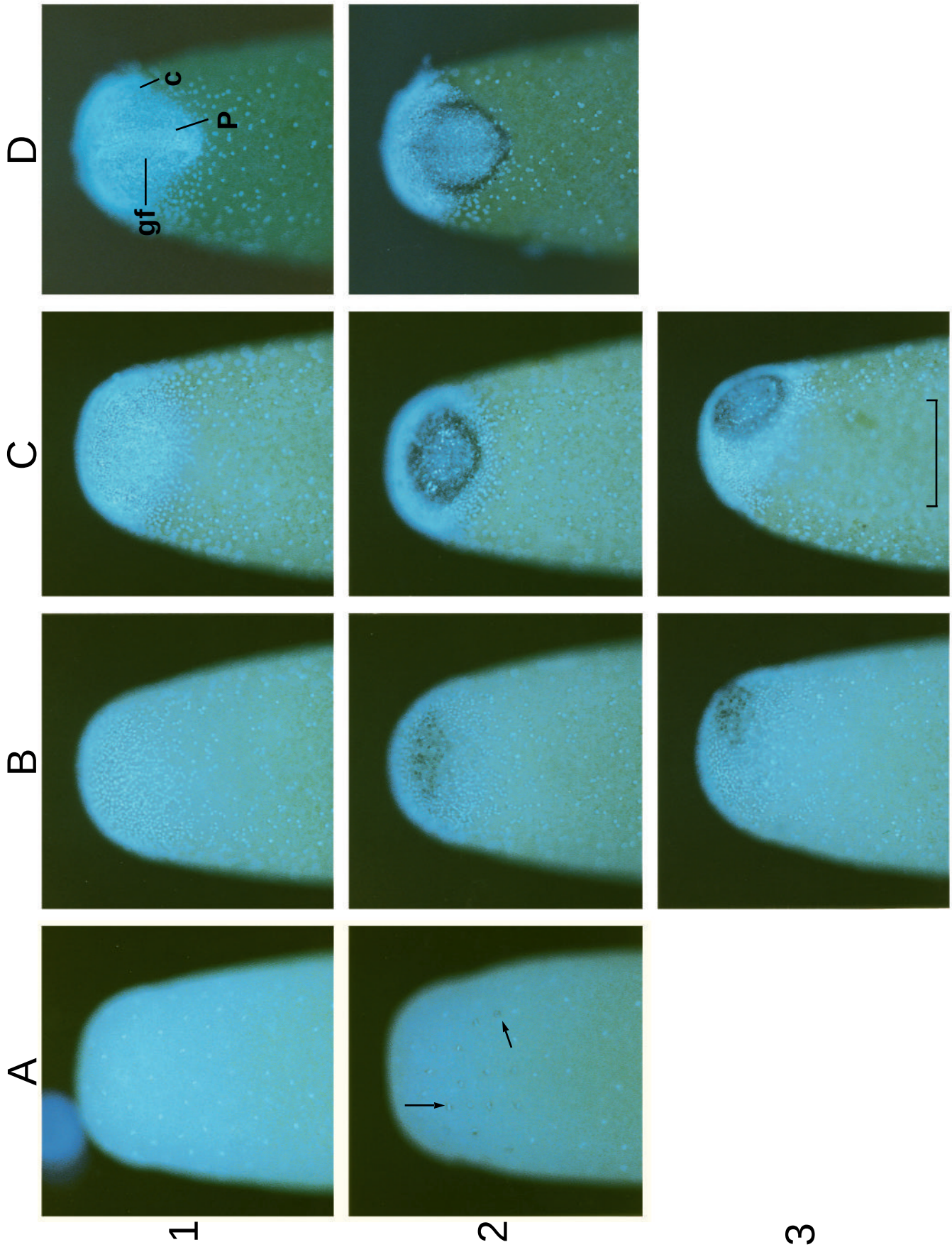
RESULTS

Library screens and characterisation of the *Dax* gene

Genomic clones carrying homeoboxes were isolated by further screens of the library described by Tear et al. (1990). Of four different Hom/Hox class homeoboxes sequenced, three were immediately identifiable as homologues of specific *Drosophila* genes (*Sex combs reduced* (*Scr*) (Akam et al., 1988; Dawson, 1989), *abdominal-A* (*abd-A*), (Tear et al., 1990) and *Abdominal-B* (*Abd-B*), (Kelsh et al., 1993)). The fourth homeobox was not immediately identifiable. It provides the focus for this study.

The complete coding sequence of the gene containing this homeobox was determined from genomic DNA spanning the homeodomain and carboxyl terminal region, and from RACE PCR clones covering the amino terminal portion of the protein. The gene encodes a protein of 314 amino acids, showing the characteristic structure of Hox/Hom genes, with the conserved YPWM motif lying 13 amino-acids upstream of the homeodomain, and separated from it by an intron (Fig. 1A). The homeodomain matches the Antp class consensus sequence (Scott et al., 1989; Burglin, 1994) at all 35 amino acid positions but differs from each specific *Drosophila* Antp class gene by 9-13 residues (Fig. 1B).

Fig. 3. *Sg Dax* expression in whole eggs during the first 15% of development. Whole eggs were stained with anti-Dax antibody and then counter stained with the fluorescent nuclear stain, Hoechst 33258. The fluorescence of the nuclear stain (blue) is quenched by the product of the antibody staining (black). Only the region of the egg that contains the developing embryo is shown (about 15% of the entire egg length). Eggs are oriented with the posterior pole at the top. For the top two rows, the view is of the concave, dorsal side of the egg. For the bottom row the view is lateral with dorsal (concave side) to the right and ventral (convex side) to the left. The eggs in the top row are unstained controls (no primary antibody is included in the antibody staining protocol). The bottom two rows are antibody stained for *Sg Dax*. The four columns depict sequential stages of development. Scale bar, 500 μm. (A) Early syncytial cleavage (less than 5% development): nuclei are undergoing rapid division and migrating from the posterior pole to the anterior and will eventually cover the entire egg. Weak stain (e.g. arrow) can be seen in the cytoplasm surrounding some of the nuclei in the most posterior region of the egg (A2). (B) Nuclear aggregation (about 7% development): nuclei have begun to concentrate at the posterior pole of the egg. These will form the embryo and the more sparsely distributed nuclei will form the serosa, an extraembryonic membrane. The exact boundary between embryonic and serosal nuclei at this stage is unclear. Nuclear staining is in a domain within the region of concentrated nuclei (B2). This domain is restricted to the dorsal surface of the egg (B3). The strength of staining of individual nuclei within this domain is variable. (C) Late germinal disc (about 12% development): nuclei have concentrated into a germinal disc which is skewed to the dorsal side of the egg. These embryonic nuclei are smaller than the sparsely distributed serosal nuclei, which will become polyploid. The exact boundary between embryonic and serosal nuclei is still not clear. Nuclear staining is in a near-circular domain within the part of the embryonic disc that lies on the dorsal side of the egg. The staining around the edge of the domain sometimes appears stronger than in the centre. (D) Early heart stage (about 15%): the embryo has differentiated a widened anterior cephalic region (c) and a narrower posterior region, the protocorm (p), that will form the trunk of the embryo. This gives the embryo a characteristic heart shape when flattened out. The cephalic region lies over the posterior pole of the egg and the protocorm lies over the dorsal side. The gastrulation furrow is visible (gf). The protocorm is stained from the anterior of the gastrulation furrow to the posterior edge of the embryo. Darkest staining is around the edge of the protocorm.


A
B
C
D
1
2
3
gf
P
c

Sequences flanking the homeodomain and the YPWM motif are frequently diagnostic for specific gene families within the Hox clusters (Bürglin, 1994). Residues flanking the homeodomain of this *Schistocerca* protein show no striking similarities to other Hox gene products, but the amino acids immediately flanking the YPWM motif are identical to those found in the *Tribolium ftz* gene (Fig. 1C, Brown and Denell, personal communication). Although this short sequence cannot be considered diagnostic, the extended FYPWMK motif is found in no other Hox/Hom protein. The remainder of the coding sequence shows only low levels of similarity with other sequences in the EMBL database.

We believe that this gene is a *ftz* homologue (see below), but as sequence alone does not identify it unambiguously, we refer to it as *Sg Dax* (divergent Antp-class homeobox gene).

To test whether this sequence is expressed in *Schistocerca*, we hybridised a probe spanning the putative homeobox exon with sections of *Schistocerca* embryos at 30% development. We observed labelling over the developing nervous system, but not over more lateral regions of the embryo (Fig. 2). Other Antp class genes are strongly expressed in the lateral epidermis at this stage (Tear, 1990; Tear et al., 1990), suggesting that the signal observed in this experiment is specific, despite the presence of a homeobox in the probe.

To examine the pattern of expression of *Sg Dax* in more detail, we raised antibodies to a fusion protein containing the homeodomain and carboxy terminal sequences of the putative protein. Two rabbits were immunised. Affinity-purified sera from both stained the same pattern of cells in the developing *Schistocerca* embryo. These sera reveal a pattern unlike that of the other Antp class genes that we have studied, but consistent with the results obtained by in situ hybridisation.

Expression of *Dax* in the early embryo

Figs 3 and 4 illustrate the early development of the *Schistocerca* embryo. The zygote nucleus is initially located near the posterior pole of the egg. After fertilisation, cleavage nuclei disperse to form a syncytial blastoderm, but only at the posterior pole do these nuclei coalesce to form a dense monolayer. This will become the embryonic primordium. Initially, this polar cap of nuclei is circular, but it rapidly extends along the dorsal side of the egg, forming first a heart shape, and then a 'tadpole', thereby revealing the anteroposterior axis of the embryo (Roonwal, 1936; Bentley et al., 1979). *Sg Dax* is expressed in posterior regions of this pri-

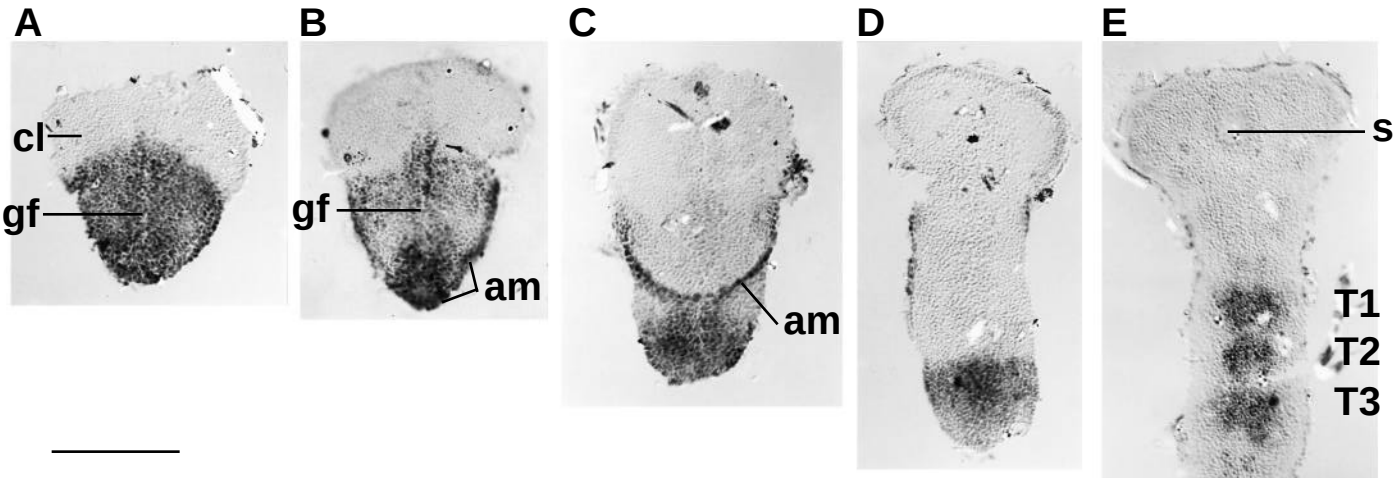


Fig. 4. *Sg Dax* protein expression in embryos from 15 to 20% development. Embryos were dissected from the serosa and underlying yolk and stained with anti-Dax antibody. All embryos are oriented with anterior upwards and viewed from the ventral side of the embryo (the side that apposes the egg chorion). (A) Heart stage (approximately 15% development): the protocorm stains uniformly from the anterior of the gastrulation furrow (gf) to the posterior edge of the embryo. The cephalic lobes (cl) are unstained. (B) Late heart stage: staining in the trunk of the embryo has now resolved into a diffuse anterior stripe and a posterior domain. The intervening cells are only weakly stained. Staining in the posterior still extends to the end of the embryo. At the anterior, staining extends further in the centre of the embryo than at the sides. It is unclear whether this central staining includes both mesoderm and ectoderm. Staining can also be seen in the amnion (am), which is attached at the edges of the trunk and has begun to grow over the ventral surface at the very posterior of the embryo. The gastrulation furrow (gf) is still visible. (C) Approximately 17% development: the trunk is longer and more uniform in width than in the earlier heart stages. The anterior domain of staining has virtually disappeared. The posterior domain still stains strongly. The leading edge of the amnion (am), which has now grown over the ventral surface of the posterior third of the trunk, also stains. (D) Approximately 18% development: the edges of the cephalic lobes are more distinct and the trunk is longer. The anterior domain of staining has disappeared completely. The posterior domain remains prominent but expression is fading in the most posterior cells of the embryo. The anterior limit of the posterior domain becomes fairly sharp at this stage. Staining in the centre of the embryo appears stronger than at the edges due to the staining of underlying mesoderm. Most of the amnion has been removed from this embryo. (E) Approximately 20% development: the stomodeum (s), the anterior invagination of the gut, is visible. The trunk has extended further and differentiated into a broader region that will form the gnathal and thoracic segments, and a narrower posterior region that will form the abdomen. At this stage, embryos show three engrailed stripes in the thorax but none in the abdomen (Patel et al., 1989). The posterior domain of expression is positioned about halfway along the prospective abdominal region. The most posterior cells in the abdomen do not stain. Within the three prospective thoracic segments (T1, T2, and T3), staining can be seen in the region of ectoderm from which the neuroblasts will delaminate. The amnion has been removed. Black and white dots are dirt that is unavoidably attracted to these sticky embryos! Scale bar, 200 μm.

mordium from its initial formation until about 23% of development.

While the nuclei are still condensing to form the embryonic disc, our antibody labels a crescent of nuclei near the posterior pole, on the dorsal side of the egg (Fig. 3B). We presume that these nuclei will form the posterior part of the embryonic axis, though this has not been demonstrated directly by fate mapping. It is not known precisely where the boundary between embryonic and extraembryonic tissue lies at this stage, or indeed whether such a boundary has yet been defined.

Once the embryonic primordium becomes visibly asymmetric, the region of nuclear staining can be seen to fill the posterior part of the presumptive germ band and to extend into the forming amnion. At 15% of development, the whole of the trunk is stained, from the anterior of the gastrulation groove backwards (Figs 3D, 4A). As far as we can tell this staining is both mesodermal and ectodermal. The head lobes are not stained. Shortly thereafter, staining in the anterior part of the growing trunk fades rapidly, transiently leaving a diffuse anterior stripe (Fig. 4B), but persisting only near the posterior tip (Fig. 4C,D). Between 18 and 20% development, cells at the very posterior tip of the embryo cease to stain, leaving only a subterminal domain of labelled cells in the presumptive abdomen. By 23%, this too has disappeared.

Segmentation does not become visible until the tadpole stage (17% development), when first stripes of Engrailed protein (Patel et al., 1989) and then segmentation of the mesoderm, define the three thoracic segments. Thereafter stripes of Engrailed protein appear sequentially down the abdomen and visible segmentation follows. In each region, *Dax* expression fades 2-3% of development (10-15 hours) before the first detection of Engrailed protein (compare Fig. 4 with Fig. 1 of Patel et al. (1989)). Thereafter, *Dax* expression is limited to the neurectoderm and its neural derivatives, as described below.

Expression in the central nervous system

At later developmental stages, our antibody stains a re-iterated pattern of cells within each neuromere of the *Schistocerca* nervous system (Fig. 5). When segmentation is completed, this staining extends from the posterior half of the first gnathal (mandibular) segment to the last abdominal segment (A11). Other tissues of the embryo are not stained, with the exception of a few isolated cells in the gnathal and thoracic appendages, and a single cell or pair of lateral cells in each abdominal segment (Fig. 5C). These may be peripheral neurons and/or glial cells.

The central nervous system of insects forms from the mid-ventral ectoderm of each segment. This region is termed the neurectoderm. Shortly after gastrulation, neuronal and glial precursor cells delaminate from the neurectoderm, moving dorsally and then enlarging. In *Drosophila*, this process occurs almost simultaneously in each trunk segment. In *Schistocerca*, segments develop sequentially, so that segments in a single embryo present a sequence of stages of neurogenesis (Fig. 5); (Bate, 1976; Bate and Grunewald, 1981; Doe and Goodman, 1985; Shepherd and Bate, 1990).

In each segment, the later phase of *Dax* expression proceeds in concert with neurogenesis. Thus staining in the neurectoderm is first seen in the thorax at about 20% embryogenesis (Fig. 4), but not until 30% in the posterior abdominal segments (Fig. 5) (Doe and Goodman, 1985). Blocks of about 40 cells stain on either side of the midline just as the first neuroblasts are beginning to delaminate. The number and intensity of stained nuclei varies considerably from embryo to embryo, reflecting dynamic or variable expression.

As the neuronal precursor cells delaminate, they form a stereotyped array that varies little from segment to segment, and even from one insect species to another (Thomas et al., 1984). These neuronal precursor cells consist of neuroblasts (NBs) and midline precursor cells (MPs). Neuroblasts divide to produce a fixed lineage of ganglion mother cells (GMCs) which themselves divide once to give rise to two neurons (Bate, 1976; Doe and Goodman, 1985). Midline precursor cells divide only once, giving rise to two neurons (Bate and Grunewald, 1981). The remaining cells of the neurectoderm give rise to non-neural support cells and contribute to the ventral epidermis (Doe and Goodman, 1985).

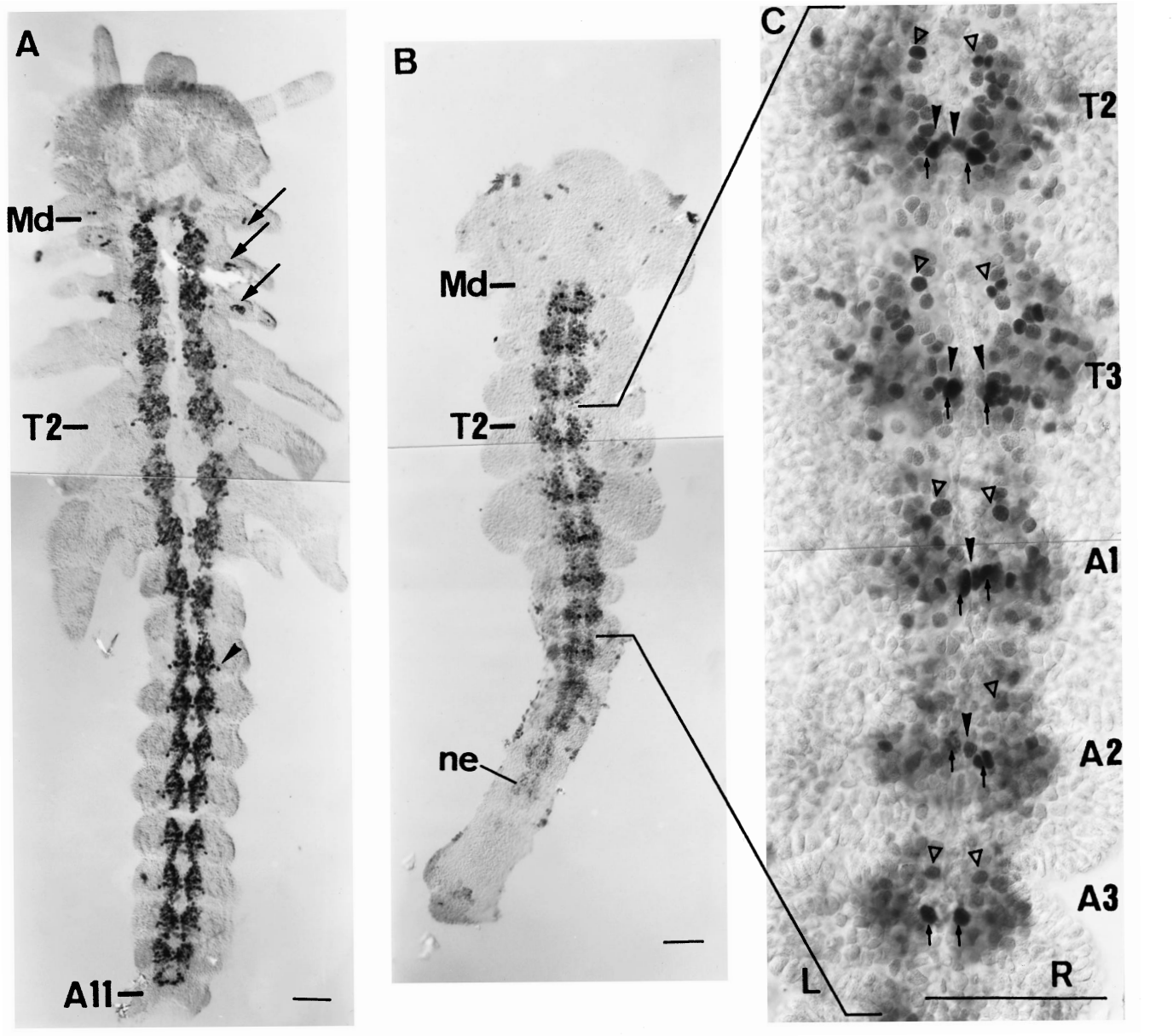
Shortly after the first generalised expression of *Dax* in the *Schistocerca* neurectoderm, staining appears in more dorsally located cells that have delaminated from the neurectoderm (Fig. 5B,C). Two large nuclei stain first, one on either side of the midline. The size and position of these cells identifies them as the midline precursor-2 cells (MP2) (Bate and Grunewald, 1981). In *Drosophila*, the MP2 cells are the first cells to express *ftz* in the neurectoderm (Doe et al. (1988a) and Fig. 6B).

About 1% of development after expression first appears in the MP2 nuclei, staining appears in a single nucleus lying between, and in the same dorsoventral plane, as the MP2 cells (Fig. 5C). From the size, position and subsequent behaviour of this cell, we identify it as the midline precursor-1 cell (MP1) (Bate and Grunewald, 1981). Shortly after staining, this cell divides, and its progeny move laterally until they lie over the progeny of the MP2 cells, which by this time have also divided.

By the time the MP1 cell has divided, several more nuclei in each hemisegment are strongly stained (Fig. 5C). One of these prominent cells is located among the anterior row of neuroblasts and their progeny, at the end of the row closest to the midline. It is too small to be a neuroblast. We identify it as a ganglion mother cell that is probably descended from neuroblast 1.1. It divides symmetrically about 2% of development later giving rise to two neurons which both stain. In *Drosophila*, the first ganglion mother cell produced by neuroblast 1.1 expresses *ftz* prominently, as do its daughters (Doe et al. (1988a) and Fig. 6C).

Subsequently, the pattern of expression becomes more complex. About 20% of the nuclei in each neuromere stain. These stained nuclei are concentrated in the anterior half of each neuromere. None of the labelled cells appear to be large enough to be neuroblasts; only midline precursor cells, ganglion mother cells and neurons stain. In all these respects, the pattern resembles that of *ftz* staining in the *Drosophila* CNS (Doe et al., 1988a; Fig. 6).

Even at these later stages some cells are still identifiable, especially in regions where the concentration of stained nuclei is lower. For example, less than 1% of development after the MP2 and MP1 progeny have formed bilateral clusters, a large nucleus between the two clusters begins to stain. This cell divides about 3% of development later. Its progeny remain on the midline until at least 35% development. This cell is likely to be the MP-3 cell (Bate and Grunewald, 1981; Goodman et al., 1981).



The description above applies to all trunk segments from the maxillary to the tenth abdominal segment. The pattern of staining is different in the mandibular segment. Only cells in the posterior half of this neuromere stain. Also, because MP2 cells are not present in the mandibular neuromere (Doe and Goodman, 1985), these cells are missing from the pattern. The eleventh abdominal (A11) neuromere is reduced in comparison with the other trunk segments (Doe and Goodman, 1985). *Dax* is expressed in the A11 neuromere, but we have not examined this pattern in detail.

Overall, the parallels between *Dax* and *ftz* expression in the nervous system are striking, and extend to both general and specific characteristics. Among the early, easily identifiable cells that express these genes, there is virtually complete overlap. The only clear difference is that *Dax* is expressed in MP3 cells as well as MP1 and MP2, whereas *Drosophila* has no homologous MP3 cell (Thomas et al., 1984). The anterior limit of both *ftz* and *Dax* expression lies in the mandibular neuromere. Posteriorly, the *Drosophila* CNS is compressed, making comparisons difficult, but *ftz* expression certainly extends into the fused terminal neuromere that represents the ninth and all more posterior segments of the Drosophila CNS.

DISCUSSION

A *ftz/Dax* gene family derived from Hox genes

Sg Dax is clearly a member of the Antp class of homeobox genes (Scott et al., 1989). In *Drosophila* the Antp class comprises four homeotic genes (*Antennapedia (Antp)*, *Sex combs reduced (Scr)*, *abdominal-A (abd-A)* and *Ultrabithorax* (*Ubx*)) and the pair rule segmentation gene *fushi tarazu* (*ftz*). Clear homologues for all four of these homeotic genes have now been identified in *Schistocerca*. These can easily be identified by sequence conservation in the homeobox exon. In each case this identification is consistent with a well conserved

Fig. 5. Neural expression of *Sg Dax* in embryos at 35% (A) and 25% (B,C) of development. Embryos were dissected from the serosa and underlying yolk and stained with anti-Sg Dax antibody. All embryos are oriented with anterior upwards. Scale bar, 100 μm. (A) 35% development: segmentation of the abdomen is complete. Staining in the central nervous system extends from the posterior of the mandibular (Md) neuromere up to and including the reduced neuromere of the last abdominal segment (A11). Groups of a few cells within the gnathal and thoracic appendages also stain (arrows). A single cell or pair of cells at the lateral edge of the neuromeres are also stained prominently (arrowhead). T2 marks the position of the second thoracic segment. (B) 25% development, showing the onset of the *Sg Dax* neural expression in the neuromeres of the abdomen. The first three abdominal segments are clearly delineated morphologically. The staining reveals a further five segments within the region of the abdomen that is not yet visibly segmented. Neurogenesis in the locust is not synchronous. The second and third thoracic segments (T2 and T3) are the oldest and segments decrease in age sequentially in both an anterior and posterior direction from T2 and T3 (Doe and Goodman, 1985). Staining is evident in the neural tissue in the centre of each segment and in the overlying ectoderm from which this derives (neurectoderm). In the youngest (most posterior) abdominal segments, staining is seen in the neurectoderm only (ne). In older (more anterior) abdominal segments, the delaminated cells of the developing nervous system are also stained. The anterior limit of staining is within the posterior of the first gnathal segment, the mandibular segment (Md). The cephalic lobes are unstained and the posterior domain seen at 20% of development (Fig. 4) has completely disappeared. (C) An enlarged view of the neuromeres of T2 to A3 from the embryo depicted in B. Because segmentation in the locust is not synchronous, the change in expression pattern seen by moving from posterior (younger) abdominal segments to more anterior (older) segments reflects the temporal progression of the pattern. In the youngest segment (A3) the midline precursor-2 (MP2) cells are prominent as the only strongly stained cells in the neuromere (arrows). Also distinguishable is a ganglion mother cell in the anterior of the neuromere, close to and on either side of the midline (empty arrowheads). From its position, this ganglion mother cell is probably derived from neuroblast 1.1 (Doe and Goodman, 1985). In A2 the MP2 cells have divided on the right (R) side of the midline giving rise to the vMP2 and dMP2 neurons (Bate and Grunewald, 1981). The other side of the neuromere is lagging slightly in development and the MP2 cell is undergoing cell division. The midline precursor-1 (MP1) cell, a single, unpaired cell lying on the midline between the MP2 neurons (Bate and Grunewald, 1981), is also staining lightly (solid arrowhead). In A1 the intensity of staining in MP1 and the presumptive NB1.1 ganglion mother cell is stronger than in the younger A2 neuromere. In T2 the MP1 cell has divided to give two MP1 neurons. On the right (R) side of the embryo, the presumptive NB1.1 ganglion mother cell has divided to give two neurons, the equivalent cell in the opposite half of the neuromere has not yet divided. Other ganglion mother cells and neurons are also staining. In the T3 neuromere the MP1 neurons have moved laterally to overlie the MP2 neurons on either side of the midline.

pattern of gene expression (Akam et al., 1988; Tear et al., 1990; Kelsh et al., 1993; Patel, 1993 and Kelsh et al., 1994). No clear homologue, however, has been identified for *ftz*.

Previous searches have failed to identify *ftz* homologues outside of the Diptera (Walldorf et al., 1989). This is probably because the sequences of *ftz* and its homologues in other insects have been less constrained than those of the canonical homeotic genes: even comparing species of *Drosophila*, *ftz* shows appreciable divergence within and flanking the homeobox (D. Maier, personal communication). However, the recent identification of a *ftz* gene from *Tribolium* (Brown and Denell, personal communication) enables us to recognize conserved features that link *Dax* with the *ftz* genes.

Tribolium and *Drosophila* are both endopterygotes, and therefore more closely related to one another than either is to Schistocerca (Kristensen, 1981, 1991). Even so, the *Tribolium ftz* gene is barely identifiable by sequence alone as a *ftz* homologue. Unambiguous recognition of this gene depends not only on sequence, but also on knowledge of its position in the *Tribolium* Hom-C, and on its expression pattern, which resembles that of *Drosophila*, both in the blastoderm and during neurogenesis (Brown and Denell, personal communication).

Three characteristics link *Schistocerca Dax* with the *ftz* genes. First, *Tribolium ftz* and *Schistocerca Dax* retain identical flanking amino acids around the YPWM motif (AFYPWMK). These residues are normally well conserved in Hox proteins of the same class. This motif has diverged almost beyond recognition in *Drosophila ftz* (Fig. 1C). Second, despite the overall divergence of their homeodomains, comparison of these three genes reveals one common feature: the homeodomains of all three are identical at the amino terminal end (residues 2-10), and are unique among Antp class genes in showing a T at residue 4 (Fig. 1C). This N-terminal region of the homeodomain is essential to specify functional differences between other Antp-class proteins, which all share identical sequences in helix three: swapping residues 1, 4, 6 and 7 alone is sufficient to switch the specificity of *Antp* to *Scr* and vice versa (Furukubo-Tokunaga et al., 1993; Zeng et al., 1993). Thus conservation of the *ftz/Dax* genes in this region is likely to be significant. Finally, *Dax* and *ftz* show a conserved pattern of expression in the same subset of cells in the developing nervous system. Thus we have no doubt that *Dax* is a homologue of *ftz*. However, because *Dax* shows no trace of the pair-rule pattern for which *ftz* is famous, and may well not play the same role in segmentation (see below), we retain a different name for the gene.

All three of the insect *ftz/Dax* genes map within or close to Hox clusters. The *Drosophila ftz* gene lies between *Scr* and *Antp*. Although its transcription orientation is reversed with respect to the flanking Hox genes, this appears to result from a recent inversion within one *Drosophila* subgenus; other *Drosophila* species carry *ftz* at the same location, but in the same orientation as flanking genes (Maier et al., 1993). The *Tribolium ftz* gene has been mapped by deletion to a region that includes *Scr*, *Antp*, *Ubx* and *abd-A* homologues (Stuart et al., 1991 and personal communication). We do not know the precise position of *Dax*, but it maps to the same chromosomal locus as *Scr* and *abd-A*, to within the limits of resolution of chromosome in situ hybridization (D. Ferrier and M. Akam, unpublished data).

The inclusion of the *Drosophila* segmentation gene *ftz* in the Hox clusters of *Drosophila* has always been somewhat enigmatic. Identifying the *Schistocerca Dax* and *Tribolium ftz* genes throws some light on this. It seems most likely that *ftz* originated as a typical Hox cluster gene, and has since lost, by sequence divergence, features that are strongly conserved in other Hox genes (e.g. the YPWM motif). In this context, it is interesting that the expression of *Dax* in *Schistocerca* displays a generic resemblance to that of a homeotic, Hom/Hox gene. First, it is expressed only in the trunk (including the mouthparts), not in the head. This is true in very early development,

but also during neurogenesis — it is striking that a relatively large fraction of cells in the ventral nerve cord express *Dax*, but none do so in the brain. Second, it shows a sharp anterior limit in its domain of expression. In the case of *Dax*, this limit is virtually at the anterior of the trunk — but in the nervous system, it is very clear that *Dax* is not expressed in the anterior part of the first gnathal neuromere (the equivalent of parasegment 0 in *Drosophila*). We note that both these characteristics are also shared by *ftz*, although the parallel has not previously been stressed.

Sg Dax, however, differs from the *Drosophila* homeotic genes in other important ways. In the later embryo expression is restricted (or virtually restricted) to the nervous system. Also, *Dax* shows no trace of down-regulation by other, posteriorly expressed homeotic genes — the protein is present in the same cells and at similar levels in all segments from the maxillary back. Finally, no *Drosophila* homeotic gene has an anterior limit of expression at the parasegment 0/1 boundary. Thus the expression pattern of *Dax* provides no evidence to suggest that it derives from a duplication of any known insect homeotic gene. It may represent a relatively old family of Hox genes that has, in arthropods, acquired functions rather different from those of the other homeotic genes.

Specific homologues of this *ftz/Dax* family may extend to the Crustacea. Among six Hox genes isolated from *Artemia* (Averof and Akam, 1993), five were immediately recognizable as orthologues of insect genes, but one was more divergent. This divergent sequence, *AfHx1*, shares with the *ftz/Dax* family conservation within the amino-terminus of the homeodomain (Fig. 1C). Tree-building based on sequence comparisons cluster the *AfHx1* homeobox specifically with *Dax* and the *ftz* genes, in preference to other insect Hox sequences (Averof, 1994).

Dax expression and early patterning

Whatever its origins, *Sg Dax* provides us with a molecular marker to define patterning at blastoderm stages in these short germ embryos. This is of particular interest, because the cellular context of early development differs markedly between *Drosophila* and *Schistocerca*.

In *Drosophila*, the nurse cells provision the egg with maternal information in the form of specific RNA and protein molecules that become localised to the poles of the egg. This information is used to define the anteroposterior axis of the developing embryo (St. Johnston and Nusslein-Volhard, 1992). In *Schistocerca*, no nurse cells are present in the egg tubes of the ovary (Mahowald, 1972). The only RNA in the egg is presumed to be provided by the oocyte nucleus. It is unclear to what extent maternal information is important for defining the axes of the embryo. The anterior half of the egg is unnecessary, for isolated posterior fragments of the egg can still develop a normally segmented embryo (Moloo, 1971; Sander, 1976). Any relevant maternal information must be located at the posterior tip of the egg.

The early expression of *Dax* shows that regional differences amongst the cleavage nuclei have already been established while the nuclei are still aggregating to form the embryonic primordium at the posterior pole. Without a fate map, it is not possible to say which parts of the prospective germ band this early *Dax* expression represents. It is clear, however, that by 15% development (heart stage), *Sg Dax* expression defines a distinction between head and trunk.

Shortly after the trunk begins to grow, Dax expression becomes restricted to just the posterior part, where segments

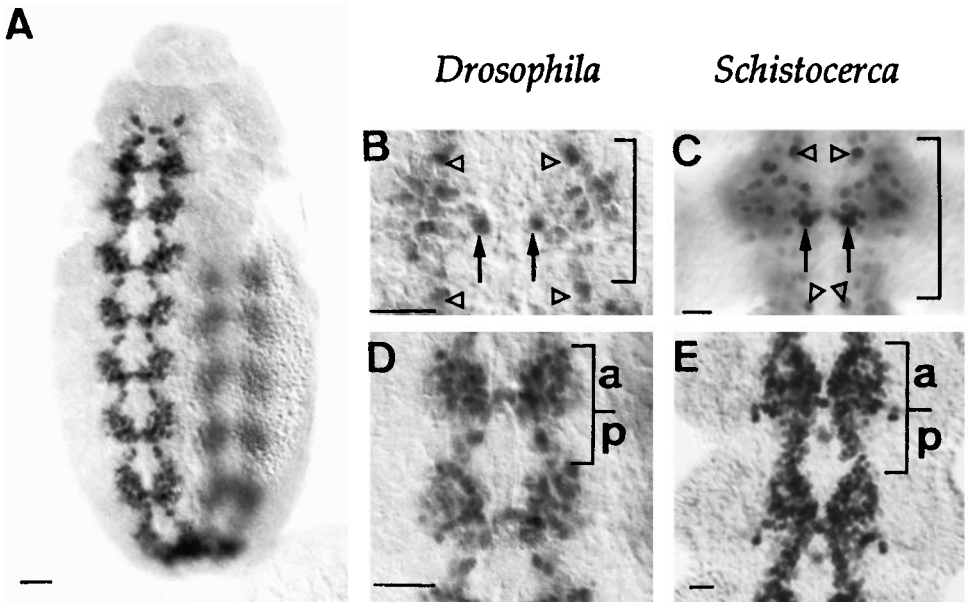


Fig. 6. Comparison of expression of *ftz* in the central nervous system of *Drosophila* with *Sg Dax* expression in the locust. Locust embryos were dissected from the serosa and underlying yolk and stained with anti-Sg Dax antibody. Drosophila embryos were stained as whole mounts with anti-ftz antibody (see Materials and Methods). Anterior is up. Scale bar, 20 μm. (A) *Drosophila* extended germ band just prior to shortening (late stage 11, (Campos-Ortega and Hartenstein, 1985)). Staining is seen in a segmentally reiterated pattern within the central nervous system. The view is ventral but turned slightly laterally. Staining in the abdominal segments that have extended onto the other side of the embryo is visible through the embryo. The expression in the central nervous system extends from the first gnathal (mandibular) segment to the reduced neuromere of the last (ninth) abdominal segment. Compare with Fig. 5A. (B,C) Early expression of *ftz* in a *Drosophila* neuromere (stage 10-11; Campos-Ortega and Hartenstein, 1985) and *Sg Dax* in a locust neuromere respectively. By this stage the majority of neuroblasts have delaminated from the neurectoderm and begun dividing, producing ganglion mother cells and neurons (Campos-Ortega and Hartenstein, 1985). In both species the pair of MP2 cells (arrows) and the more anterior ganglion mother cell that, from its position (Doe and Goodman, 1985), is probably derived from NB1.1 (empty arrowheads) are amongst the first strongly stained neuronal cells. Large bracket denotes a single neuromere. The presumptive NB1.1 ganglion mother cell from the segment below is also shown in both cases. (D,E) Expression of *ftz* and *Sg Dax* in abdominal neuromeres from the *Drosophila* embryo in Fig. 6A and locust embryo in Fig. 5C respectively. Note the overall similarity in expression pattern between the two insects including a greater number of expressing cells in the anterior of the neuromere (a) than in the posterior (p). Large bracket denotes a single neuromere.

have not yet formed. Later still, *Dax* expression is excluded from the posterior tip of the abdomen, but remains in a subterminal zone. Ablation experiments suggest that the extreme posterior abdominal segment is determined before segmentation of the abdomen is complete (Heinig, 1967). Thus there is some correlation between *Dax* expression, and regions that remain in the 'growth zone'.

This early pattern of *Dax* expression is strikingly different to that of *ftz*, in that it shows no trace of resolution into periodic stripes at double segment intervals.

Segmentation mechanisms in short germ insects

Ftz is one of a set of 'pair-rule' genes, all of which are required to elaborate segment pattern in the syncytial blastoderm of *Drosophila* (Akam, 1987; Ingham, 1988). When any one of these genes is mutant, embryos develop with alternate segments missing. One other gene from this pair-rule set has been characterised in *Schistocerca* — the homeobox gene *even-skipped (eve)*.

Like *ftz*, *Drosophila eve* plays several roles in development. In the blastoderm, *eve* is expressed in a pattern of stripes exactly complementary to those of *ftz* (Carroll and Scott, 1985; MacDonald et al., 1986) and serves to define the anterior boundaries of odd-numbered parasegments (Lawrence and Johnston, 1989). In later development, it is expressed in a subset of muscle precursors, and in a small subset of ganglion mother cells and neurons, where it is required for neural specification (MacDonald et al., 1986; Doe et al., 1988b).

Schistocerca eve is expressed in the same specific neural and muscle cells as its *Drosophila* counterpart but, like *Dax*, it shows no trace of 'pair-rule' expression during early development (Patel et al., 1992). It, like *Dax*, is expressed in the posterior part of the growing germ band from gastrulation onwards. Earlier *eve* expression, in the germinal disc, has not been described.

In *Drosophila*, the spatially patterned transcription of *eve* and *ftz* are central to the mechanism that generates positional signals in the blastoderm. Moreover, *eve* is directly involved in patterning the transcription of *ftz* (Carroll and Scott, 1986). The observation that neither *Sg eve* nor *Dax* show periodic stripes, implies that the molecular mechanisms generating segments in *Drosophila* are substantially different from those in *Schistocerca*. We discuss this question more fully elsewhere (Akam and Dawes, 1992).

The precise role of *eve* and *Dax* in the early *Schistocerca* embryo remains an open question. In vertebrates, *even-skipped* homologues are linked to the Hox gene cluster (Bastian and Gruss, 1990; Boncinelli et al., 1991). The vertebrate *eve* homologues are expressed in the posterior regions of early embryos, and have been implicated in the specification of 'posterior identities' (Ruiz i Altaba and Melton, 1989; Bastian and Gruss, 1990). The expression of *Dax* and *eve* at the heart stage onwards in *Schistocerca* also suggests a role in regionalisation of the embryo.

In *Tribolium*, pair-rule patterns of *eve* (Patel et al., 1994; Brown and Denell, personal communication), *hairy* (Sommer and Tautz, 1993) and *ftz* (Brown and Denell, personal communication) stripes appear sequentially along the length of the embryo, suggesting that a pair-rule patterning mechanism exists in beetles (Patel et al., 1994). However, a deletion of the Tribolium Hom-C (Stuart et al., 1991) spanning the *ftz* gene (Brown and Denell, personal communication) does not result in pair-rule segmentation defects. Thus in *Tribolium*, *ftz* appears to be regulated by the pair-rule patterning system, but not required for segmentation. No modern insect represents an actual intermediate in the evolution of insect developmental mechanisms, but in these comparisons we may be seeing intermediate stages in the process whereby *ftz* acquired a role in defining segment boundaries.

We thank Alfonso Martinez-Arias for enthusiatically supporting the initiation of this project, Michael Bate for patiently helping us to interpret the *Schistocerca* nervous system, Nipam Patel, Sue Brown, Rob Dennell and Dieter Maier for communicating their results prior to publication and David Micklem for his comments on the manuscript. This work was supported by grants from the Medical Research Council and the Wellcome Trust.

REFERENCES

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

Akam, M. and Dawes, R. (1992). More than one way to slice an egg. *Current Biology* **2**, 395-398.

Akam, M., Dawson, I. and Tear, G. (1988). Homeotic genes and the control of segment identity. *Development* **104 Supplement**, 123-133.

Averof, M. (1994). HOM/Hox Genes of a Crustacean; Evolutionary Implications. Ph.D. thesis, University of Cambridge.

Averof, M. and Akam, M. (1993). HOM/Hox genes of *Artemia*: implications for the origin of insect and crustacean body plans. *Current Biology* **3**, 73-78.

Bastian, H. and Gruss, P. (1990). A murine *even-skipped* homologue, *Evx 1*, is expressed during early embryogenesis and neurogenesis in a biphasic manner. *EMBO J.* **9**, 1839-1852.

Bate, C. M. (1976). Embryogenesis of an insect nervous system I: A map of the thoracic and abdominal neuroblasts in *Locusta migratoria*. *J.Embryol. Exp. Morph.* **35**, 107-123.

Bate, C. M. and Grunewald, E. B. (1981). Embryogenesis of an insect nervous system II: A second class of neuron precursor cells and the origin of the intersegmental connectives. *J. Embryol. Exp. Morph.* **61**, 317-330.

Bentley, D., Keshishian, H., Shankland, M. and Toroian-Raymond, A. (1979). Quantitative staging of embryonic development of the grasshopper *Schistocerca nitens*. *J. Embryol. Exp. Morph.* **54**, 47-74.

Boncinelli, E., Simeone, A., Acampora, D. and Mavilio, F. (1991). *HOX* gene activation by retinoic acid. *Trends Genet.* **7**, 329-334.

Burglin, T. R. (1994). The homeobox phylum: A comprehensive survey of homeodomain sequences and associated motifs. In *A Guidebook for Homeobox Genes*. (ed. D. Duboule), pp 27-71. Oxford: Oxford University Press.

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

Carroll, S. B. and Scott, M. P. (1985). Localization of the *fushi tarazu* protein during *Drosophila* embryogenesis. *Cell* **43**, 47-57.

Carroll, S. B. and Scott, M. P. (1986). Zygotically active genes that affect the spatial expression of the *fushi tarazu* segmentation gene during early *Drosophila* embryogenesis. *Cell* **45**, 113-126.

Dawson, I. (1989). Characterisation of homoeobox containing clones from the grasshopper *Schistocerca gregaria*. Ph. D thesis, University of Cambridge.

Doe, C. Q. and Goodman, C. S. (1985). Early events in insect neurogenesis. 1. Development and segmental differences in the pattern of neuronal precursor cells. *Dev. Biol.* **111**, 193-205.

Doe, C. Q., Hiromi, Y., Gehring, W. J. and Goodman, C. S. (1988a). Expression and function of the segmentation gene *fushi tarazu* during *Drosophila* neurogenesis. *Science* **239**, 170-175.

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

Duncan, I. (1986). Control of bithorax complex functions by the segmentation gene *fushi tarazu* of *Drosophila melanogaster*. *Cell* **47**, 297-309.

Eggleston, P. (1992). Identification of the *abdominal-A* homologue from *Aedes aegypti* and structural comparisons among related genes. *Nucleic Acids Research* **20**, 4095-4096.

Frohman, M. A. (1990). RACE: Rapid Amplification of cDNA Ends. In *PCR*

Protocols: A Guide to Methods and Applications. (ed. M. A. Innis, D. H. Gelfand, J. J. Sninsky and T. J. White), pp. 28-38. New York: Academic Press.

Furukubo-Tokunaga, K., Flister, S. and Gehring, W. J. (1993). Functional specificity of the Antennapedia homeodomain. *Proc. Natl. Acad. Sci. USA* **90**, 6360-6364.

Goodman, C. S., Bate, M. and Spitzer, N. C. (1981). Embryonic development of identified neurons: Origin and transformation of the H cell. *J. Neurosci.* **1**, 94-102.

Hafen, E., Kuroiwa, A. and Gehring, W. J. (1984). Spatial distribution of transcripts from the segmentation gene *fushi tarazu* during *Drosophila* embryonic development. *Cell* **37**, 833-841.

Harlow, E. and Lane, D. (1988). *Antibodies: A Laboratory Manual*. Cold Spring Harbor Laboratory.

Heinig, S. (1967). Die abänderung embryonaler differenzierungsprozesse durch totale röntgenbestrahlung in ei von Gryllus domesticus. *Zool. Jb. Anat. Bd.* **84**, 425-492.

Hiromi, Y., Kuroiwa, A. and Gehring, W. J. (1985). Control elements of the Drosophila segmentaion gene *fushi tarazu*. *Cell* **43**, 603-613.

Ingham, P. W. (1988). The molecular genetics of embryonic pattern formation in *Drosophila*. *Nature* **335**, 25-34.

Ingham, P. W. and Martinez-Arias, A. (1986). The correct activation of Antennapedia and bithorax complex genes requires the *fushi tarazu* gene. *Nature* **324**, 592-597.

Karch, F., Bender, W. and Weiffenbach, B. (1990). *abdominal-A* expression in *Drosophila* embryos. *Genes Dev.* **4**, 1573-1588.

Kelsh, R., Weinzierl, R. O. J., White, R. A. H. and Akam, M. (1994). Homeotic gene expression in the Locust *Schistocerca*: an antibody that detects conserved epitopes in Ultrabithorax and abdominal-A proteins. *Developmental Genetics*, **15**, 19-31.

Kelsh, R. N., Dawson, I. A. and Akam, M. (1993). An analysis of *Abdominal-B* expression in the locust *Schistocerca gregaria*. *Development* **117**, 293-305.

Kornfeld, K., Saint, R. B., Beachy, P. A., Harte, P. J., Peattie, D. A. and Hogness, D. S. (1989). Structure and expression of a family of *Ultrabithorax* mRNAs generated by alternative splicing and polyadenylation in *Drosophila*. *Genes Dev.* **3**, 243-258.

Krause, H. M., Klemenz, R. and Gehring, W. J. (1988). Expression, modification, and localization of the *fushi tarazu* protein in *Drosophila* embryos. *Genes Dev.* **2**, 1021-1036.

Kristensen, N. P. (1981). Phylogeny of insect orders. *Ann. Rev. Entomology* **26**, 135-157.

Kristensen, N. P. (1991). Phylogeny of extant hexapods. In *The Insects of Australia*. (ed. Division of Entomology CSIRO), pp 125-140. Melbourne University Press.

Laughon, A. and Scott, M. P. (1984). Sequence of a *Drosophila* segmentation gene: protein structure homology with DNA-binding proteins. *Nature* **310**, 25-31.

Lawrence, P. A. and Johnston, P. (1989). Pattern formation in the *Drosophila* embryo: allocation of cells to parasegments by *even-skipped* and *fushi tarazu*. *Development* **105**, 761-769.

LeMotte, P. K., Kuroiwa, A., Fessler, L. I. and Gehring, W J. (1989). The hoemotic gene *Sex Combs Reduced* of *Drosophila*: gene structure and embryonic expression. *EMBO J.* **8**, 219-229.

MacDonald, P. M., Ingham, P. and Struhl, G. (1986). Isolation, structure, and expression of *even-skipped*: a second pair-rule gene of *Drosophila* containing a homeobox. *Cell* **47**, 721-734.

Mahowald, A. P. (1972). Oogenesis. In *Developmental Systems: Insects.* (ed. S. J. Counce and C. H. Waddington), pp 1-47. London: Academic Press.

Maier, D., Sperlich, D. and Powell, J. R. (1993). Conservation and change of the developmentally crucial *fushi tarazu* gene in *Drosophila*. *J. Mol. Evol.* **36**, 315-326.

Moloo, S. K. (1971). The degree of determination of the early embryo of *Schistocerca gregaria* (Forskål). *J. Embryol. Exp. Morph.* **25**, 277-299.

Nagy, L. M., Booker, R. and Riddiford, L. M. (1991). Isolation and embryonic expression of an *abdominal-A*-like gene from the lepidopteran, *Manduca sexta*. *Development* **112**, 119-131.

Patel, N. H. (1993). Evolution of insect pattern formation: A molecular analysis of short germband segmentation. In *Evolutionary Conservation of Developmental Mechanisms*. (ed. A. Spradling), pp 85-110. New York: Wiley-Liss Inc.

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.

Patel, N. H., Condron, B. G. and Zinn, K. (1994). Pair-rule expression patterns of *even-skipped* are found in both short and long germ beetles. *Nature* **367**, 429-434.

Patel, N. H., Kornberg, T. B. and Goodman, C. S. (1989). Expression of *engrailed* during segmentation in grasshopper and crayfish. *Development* **107**, 201-213.

Roonwal, M. L. (1936). Studies on the embryology of the African migratory locust *Locusta migratoria migratoroides*. 1. Early development with a new theory of multiphased gastrulation among insects. *Phil. Trans. Roy. Soc. B* **226**, 391-421.

Ruiz i Altaba, A. and Melton, D. A. (1989). Involvement of the *Xenopus* homeobox gene *Xhox3* in pattern formation along the anterior-posterior axis. *Cell* **57**, 317-326.

Sander, K. (1976). Specification of the basic body pattern in insect embryogenesis. *Adv. Insect. Physiol* **12**, 125-238.

Schneuwly, S., Kuroiwa, A., Baumgartner, P. and Gehring, W. J. (1986). Structural organization and sequence of the homeotic gene *Antennapedia* of *Drosophila melanogaster*. *EMBO J.* **5**, 733-739.

Scott, M. P., Tamkun, J. W. and Hartzell, G. W. I. (1989). The structure and function of the homeodomain. *Biochim. Biophys. Acta* **989**, 25-48.

Shepherd, D. and Bate, C. M. (1990). Spatial and temporal patterns of neurogenesis in the embryo of the locust (*Schistocerca gregaria*). *Development* **108**, 83-96.

Smith, D. B. and Johnson, K. S. (1988). Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase. *Gene* **67**, 31-40.

Sommer, R. J. and Tautz, D. (1993). Involvement of an orthologue of the *Drosophila* pair-rule gene *hairy* in segment formation of the short germ band embryo of *Tribolium* (Coleoptera). *Nature* **361**, 448-450.

St. Johnston, D. and Nusslein-Volhard, C. (1992). The origin of pattern and polarity in the *Drosophila* embryo. *Cell* **68**, 201-220.

Stuart, J. J., Brown, S. J., Beeman, R. W. and Denell, R. E. (1991). A deficiency of the homeotic complex of the beetle *Tribolium*. *Nature* **350**, 72-74.

Stuart, J. J., Brown, S. J., Beeman, R. W. and Denell, R. E. (1993). The *Tribolium* homeotic gene *Abdominal* is homologous to *abdominal-A* of the *Drosophila* bithorax complex. *Development* **117**, 233-243.

Tear, G. (1990). A molecular analysis of embryonic development in the short-germ band insect *Schistocerca gregaria*. Ph. D thesis, University of Cambridge.

Tear, G., Akam, M. and Martinez-Arias, A. (1990). Isolation of an *abdominal-A* gene from the locust *Schistocerca gregaria* and its expression during early embryogenesis. *Development* **110**, 915-926.

Thomas, J. B., Bastiani, M. J., Bate, M. and Goodman, C. S. (1984). From grasshopper to *Drosophila*: a common plan for neuronal development. *Nature* **310**, 203-207.

Ueno, K., Hui, C.-C., Fukuta, M. and Suzuki, Y. (1992). Molecular analysis of the deletion mutants in the E homeotic complex of the silkworm *Bombyx mori*. *Development* **114**, 555-563.

Wakimoto, B. T. and Kaufman, T. C. (1981). Analysis of larval segmentation in lethal genotypes associated with the *Antennapedia* gene complex in *Drosophila melanogaster*. *Dev. Biol.* **81**, 51-64.

Wakimoto, B. T., Turner, F. R. and Kaufman, T. C. (1984). Defects in embryogenesis in mutants associated with the Antennapedia gene complex of *Drosophila melanogaster*. *Dev. Biol.* **102**, 147-172.

Walldorf, U., Fleig, R. and Gehring, W. J. (1989). Comparison of homeobox-containing genes of the honeybee and *Drosophila*. *Proc. Natl. Acad. Sci. USA* **86**, 9971-9975.

Zeng, W., Andrew, D. J., Mathies, L. D., Horner, M. A. and Scott, M. P. (1993). Ectopic expression and function of the *Antp* and *Scr* homeotic genes: the N terminus of the homeodomain is critical to functional specificity. *Development* **118**, 339-352.

(Accepted 18 March 1994)