Development 112, 119–129 (1991)
Printed in Great Britain © The Company of Biologists Limited 1991

Isolation and embryonic expression of an *abdominal-A-like* gene from the lepidopteran, *Manduca sexta*

LISA M. NAGY, RONALD BOOKER* and LYNN M. RIDDIFORD

Department of Zoology, University of Washington, Seattle, WA 98195, USA

*Present address Section of Neurobiology and Behavior, Cornell University, Ithaca, NY, 14853-2702, USA

Summary

Using sequence homology to the *Drosophila Antennapedia* gene, we isolated a homeobox-containing gene from the lepidopteran, *Manduca sexta*. Sequence analysis and *in situ* hybridizations to tissue sections suggest that the *Manduca* gene encodes a lepidopteran homologue of the *Drosophila Bithorax* complex gene *abdominal-A*. The predicted amino acid sequence of a 76 amino acid region that includes the homeobox and the regions immediately flanking it are identical between the *Manduca* and *Drosophila* genes. Northern blots reveal that the *Manduca abd-A* gene is expressed first in the early embryo and continues to be expressed throughout later embryonic and larval stages. *In situ* hybridizations show that the posterior half of the first abdominal segment marks the anterior border of the *Manduca abd-A* expression. This expression pattern demonstrates the conservation of parasegments as domains of gene activity in the lepidopteran embryo. The *Manduca abd-A* expression extends from the posterior half of the first abdominal segment through the tenth abdominal segment, a domain that is greater than that of the *Drosophila abd-A* expression, and reflects the difference in visible segment number between the two insects.

Key words: *abdominal-A*, homeobox-gene, *Manduca sexta*, segmentation.

Introduction

Modifications and specializations of segment character account for most of the enormous taxonomic diversity found in the insects. Underlying this diversity, however, is a striking similarity in the basic metameric body plan. Homeotic genes are involved in establishing segment identity (Morata and Lawrence, 1977) and are likely to play a crucial role in generating the diversity of segments, but what maintains the basic metameric plan is largely unknown.

Drosophila homeotic genes are clustered in the *Antennapedia* and *bithorax* gene complexes (Lewis, 1978; Mahaffey and Kaufman, 1987). The initial isolation and sequencing of these genes uncovered a region of sequence homology, the homeobox, found in nearly all the homeotic genes (McGinnis *et al.* 1984; Scott and Weiner, 1984). This homeobox region is able to bind to DNA, and these proteins act as transcriptional regulators (see reviews by Levine and Hoey, 1988; Hayashi and Scott, 1990). The *Drosophila* homeotic genes are first expressed at the cellular blastoderm stage as a result of the earlier action of the gap, pair-rule, and polarity genes (Akam, 1987; Ingham, 1988). At this time, each body segment is characterized by the expression of an unique set of homeotic genes, the products of which then are thought to control the subsequent expression of segment identity (Garcia-Bellido, 1977).

The common occurrence of homeotic mutations in other insects (for a review of early literature, see Ouweneel, 1976; Sibatani, 1980; Beeman, 1987; Booker and Truman, 1989) suggests that the fundamental aspects in the segmentation process that involve homeotic gene function will be conserved throughout the insects. In fact, clusters of homeotic alleles have been genetically identified in *Bombyx* (Tazima, 1964) and *Tribolium* (Beeman, 1987). The isolation and expression patterns of homeotic genes from other insects [*Apis* (Fleig *et al.* 1988; Walldorf *et al.* 1989), *Schistocerca* (Akam *et al.* 1988; Tear *et al.* 1990), and *Tribolium* (Brown *et al.* 1990)] as well as from other invertebrates and vertebrates (Gehring, 1987; Scott *et al.* 1989; Kessel and Gruss, 1990) indicates that at least some of the molecular components used to form and maintain repeated elements along a body axis are conserved throughout the eukaryotes.

Experimental and descriptive analyses of insect embryogenesis, however, suggest a divergence in mechanisms of segmentation within the insects. In particular, not all insects form their segments in the same way. Insects can be generally classified into three categories based on the germ anlage: short, intermediate and long (see Sander, 1976). Short and intermediate

germ insects are characterized by a progressive determination of the posterior segments, while long germ insects specify the identity of all the segments nearly simultaneously in the cellular blastoderm. In addition, the different types of germ band respond differently to experimental perturbations such as ligation, UV- and X-irradiation, and heat shock (see Kuhn, 1971; Seidel, 1971; and Sander, 1976 for summaries).

Even within these three categories, there is significant diversity of developmental pattern. For instance, the long germ anlage embryo of *Drosophila* spends approximately 4% of its developmental time in the cellular blastoderm stage (Campos-Ortega and Hartenstein, 1985), whereas that of the honeybee *Apis mellifera* spends 36 % of its developmental time in this stage and does not subsequently undergo germ band elongation (Fleig and Sander, 1986). These observations imply that, in spite of conserved components in the process of segmentation, the specific manner by which a particular insect constructs its overall body plan differs from group to group. These observed differences may result from simple differences in timing of events relative to one another, or they may in fact represent qualitatively different ways of making segments (Meinhardt, 1982, 1986; Sander, 1983, 1984).

The position of the lepidopterans within this system of classification remains uncertain (Sander, 1984). Studies using localized UV irradiation on both a primitive lepidopteran *Tineola* (Liischer, 1944) and the more derived *Bombyx* (Myohara and Kiguchi, 1990) suggest that segment determination may occur before or during cellular blastoderm as expected for a long germ insect. Yet recent studies with a fascicuhn-like antibody (Carr and Taghert, 1989) suggest that in another derived lepidopteran, *Manduca sexta,* there may be a progressive determination of segments as typical of intermediate and short germ insects.

To begin to understand novel mechanisms of segmentation and early lepidopteran embryogenesis, we have utilized sequence homology to isolate segmentation/homeotic genes from *Manduca sexta.* We report here the isolation of an homeobox-containing gene with significant sequence similarity to the *Drosophila* homeotic gene *abdominal-A* (Karch *et al.* 1990), and show its temporal and spatial pattern of expression during *Manduca* development.

Materials and methods

Genomic library screening

A *Manduca sexta* genomic library (Rebers *et al* 1987) was screened with a 570 bp *EcoRl-BamHl* fragment isolated from a clone (p904) of the *Antennapedw.* gene of *Drosophila* (Garber *et al* 1983). The probe contained the 180bp homeobox sequence plus 390 bp 3' to the homeobox. The library was screened under low stringency conditions exactly as described by McGinnis *et al.* (1984). Low stringency washes were carried out with 2xSSC (lxSSC: 0.15M NaCl, 0 015M sodium citrate), 0.1% SDS, for 4x15 min at room temperature, then 2x30 min at 55 °C.

Southern blot hybridizations

DNA was isolated by the method of McGinnis *et al.* (1983) from adult *Drosophila* and fifth instar larval epidermis of *Manduca.* Genomic DNA was electrophoresed on 10% agarose gels and transferred to Hybond N (Amersham). A 1.2 kb homeobox-containing *EcoKl* fragment was labeled with pPJdATP (New England Nuclear) by the method of random priming (Feinberg and Vogelstein, 1984). High-stringency hybridizations were earned out with the same buffer as used for the low-stringency hybridization (McGinnis *et al* 1984), but the hybridization temperature was raised to 42°C. Filters were washed 4x15min at room temperature in 2xSSC, 0.1%SDS, followed by two 30mm washes at 65°C with O.l xSSC, 0.1% SDS

RNA isolation and northern blot hybridization

Manduca adults were housed with tobacco plants under a 17L/7D photopenod at 21°C Egg collections were made from the plants beginning an hour before lights off (Sasaki and Riddiford, 1984) Eggs were stored at 26°C for the appropriate development time, washed with 50% bleach (concentrated bleach: 5.25% sodium hypochlonte), rinsed with water, then frozen in liquid nitrogen

RNA from embryos and larval tissues was extracted with guanidine isothiocyanate as described by Davis *et al* (1986). RNA was quantified by UV spectrophotometry, and the quality of the total RNA was detected by visualization with ethidium bromide in a 1% agarose gel Poly(A)+ RNA was selected by ohgo(dT) (Collaborative Research) chromatography (Aviv and Leder, 1972). RNA was separated on 0.66 M formaldehyde gels (Davis *et al.* 1986), which included 0.66 *ng* ml"^{1} ethidium bromide. The gels were photographed with a red filter under UV light, then transferred to Hybond N (Amersham). Blots were hybridized at 42°C, in 50% deionized formamide, 5xSSC, 50mM $NaPO_4$, pH6 5, 250/igmP^{1} sonicated, denatured herring sperm DNA, 250/igml"^{1} yeast tRNA, lxDenhardt's solution (Denhardt, 1966). The probe was a 216 bp *EcoRl-XhoI* fragment, which contained 133 bp of the homeobox and 92 bp 3' to the homeobox. Both the labeling of this fragment and the washing conditions were the same as for the high-stringency Southerns

Sequencing

The *Xhol* fragments indicated in Fig. 1 were subcloned into M13 vectors, MP18 or MP19 (Yanish-Perron *et al.* 1985), and sequenced using single-stranded templates by the dideoxy chain termination technique (Sanger *et al.* 1977), modified for the use of ^{35}S-dATP (Biggin *et al* 1983). Overlapping deletions for some of the clones were created by digesting M13 templates with T4 DNA polymerase (Dale *et al* 1985).

Analysis of the DNA sequence was performed using IBI/Pustell sequence analysis programs (International Biotechnologies, Inc.; Pustell and Kafatos, 1982)

In situ *hybridizations to tissue sections*

To generate RNA probes to both coding and non-coding strands, the 216 bp *EcoRl-Xhol* fragment was subcloned into Bluescnpt (Stratagene). This clone was linearized with either EcoRI or *Xhol* and labeled with ^{35}S-UTP (Amersham, 1300/iCimmor^{1}). RNA probes with specific activity greater than lxl^ctsmin"^{1}^"^{1} were made using T7 and T3 polymerases.

Staged *Manduca* embryos were rinsed briefly in bleach, then rinsed in water, and fixed in 4 % formalin in phosphate-buffered saline (PBS: 1 3 M NaCl, 0 007M Na_2HPO_4, 0 03M

NaH_2PO_4) for 1 h. At this point, the embryos were manually dechorionated and fixed for an additional 4–6 h. Dehydration, paraffin embedding, prehybridization and hybridizations were performed according to Angerer *et al* (1987). Sections were 6 μm thick. 30 μl of 1×10^5 disints min^{-1} μl^{-1} of probe were applied to each slide. Exposure times varied between 5 and 14 days. Sections were stained in 0.02 % toluidine blue, 0.02 M borate pH 9.4 after developing.

Results

Library screening and genomic southerns

Using a *Drosophila Antennapedia* probe (Garber *et al.* 1983) containing the 180 bp *Antennapedia* homeobox and 390 bp 3′ to the homeobox, we screened three genomic equivalents of a *Manduca* genomic library under low-stringency conditions. We isolated six clones; restriction enzyme mapping showed all of these clones to be overlapping or duplicates, and to encompass a 15 kb genomic region. The region of homology to the *Drosophila Antennapedia* probe was restricted to 225 bp between the *Xho*I sites at positions 39 and 264 in Fig. 1 (hereafter referred to as the *Xho*I fragment).

A Southern blot of *Manduca* genomic DNA probed under high-stringency conditions with the *Manduca* fragment containing the homeobox homology revealed one hybridizing band (Fig. 2), indicating that this clone represents a single copy gene in the *Manduca* genome. A rescreening of three genomic equivalents of the *Manduca* library with this same probe yielded an additional 14 homeobox-containing clones, which separated into 9 classes by restriction enzyme mapping. Preliminary sequence analysis of three of these classes

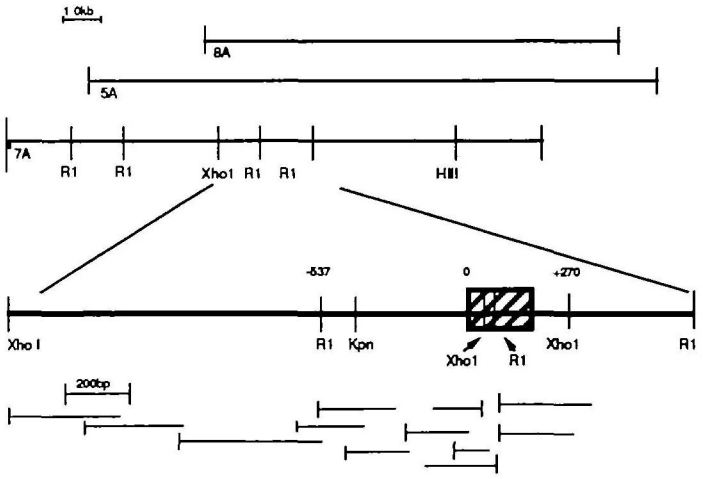

Fig. 1. Restriction map of the *Manduca* homeobox-containing genomic clones. The three overlapping clones that cover 15 kb of the *Manduca* genome are shown at the top. Beneath them and magnified is a 2.2 kb fragment containing the homeobox (shaded). We have sequenced between −1476 and +389, with position 0 marked as the first nucleotide in the homeobox. At the bottom are shown the fragments that were used in sequencing the region. HIII, *Hin*dIII; R1, *Eco*RI.

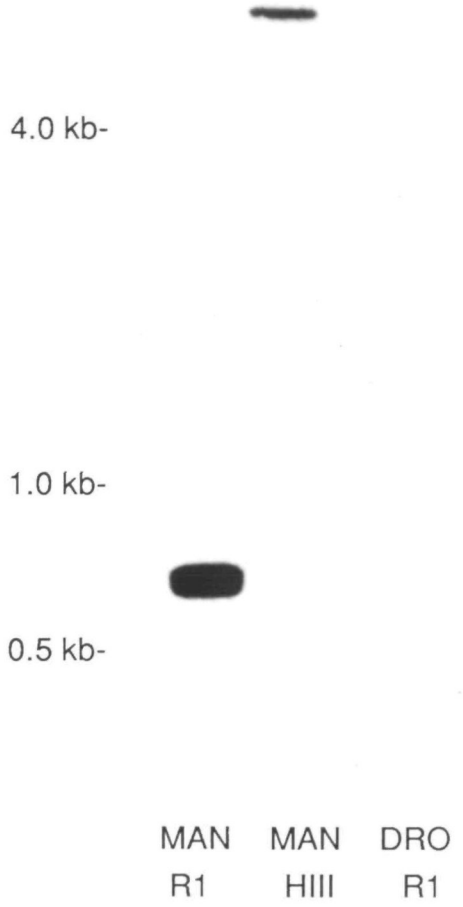

Fig. 2. Genomic Southern blots hybridized with *Manduca abd-A*. Migration distances of 0.5, 1 and 4 kb size standards are marked to the left of the blot. 10 μg *Manduca* genomic DNA cut with *Eco*RI (lane 1) or *Hin*dIII (lane 2), and 2.5 μg *Drosophila* DNA (lane 3) cut with *Eco*RI, hybridized to a *Manduca* homeobox-containing fragment under high stringency conditions.

indicate that they have significant sequence similarity to the *Drosophila* genes *Deformed, Antennapedia* and *Ultrabithorax* (Booker, in preparation).

Sequence analysis

Sequence analysis shows that the 225 bp *Xho*I fragment contains part of a homeobox that is 78 % similar at the nucleotide level and 92 % similar at the amino acid level to the *Drosophila Antennapedia* homeobox (McGinnis *et al.* 1984; Schneuwly *et al.* 1986; Stroeher *et al.* 1986), which was used as the initial hybridization probe. More importantly, the amino acid sequence of the *Manduca* homeobox (Fig. 3) is identical to that of the *Drosophila*

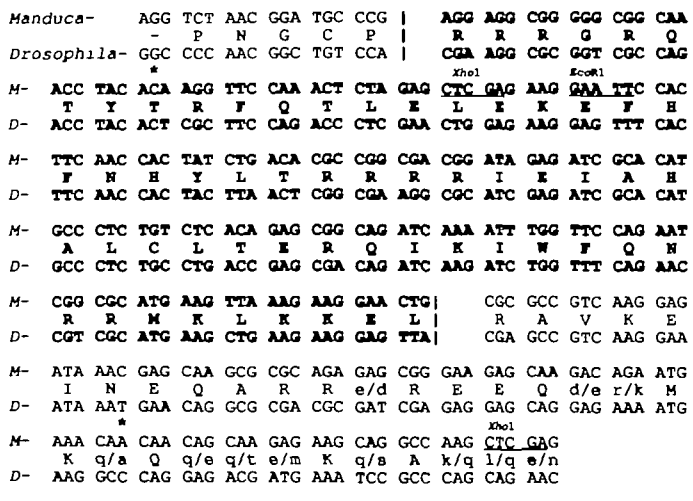

Fig. 3. Comparison of the sequence of the *Manduca* genomic clone to the cDNA sequence of the *Drosophila abd-A* homeobox-containing exon. The nucleotides encoding the sixty amino acids in the homeobox are in bold face; the *Manduca* sequence is marked M, *Drosophila*, D. The first amino acid shown is the first in the *Drosophila abd-A* homeobox-containing exon (amino acids 133–230 of the predicted *Drosophila abd-A* protein sequence are shown). Shared amino acid sequences of the *Manduca* homeobox region and the *Drosophila abd-A* sequence (Karch *et al* 1990) are in upper case letters between the two DNA sequences. The *Xho*I and *Eco*RI sites used to generate the hybridization probes from the *Manduca* clone are underlined The *Drosophila* splice sites are marked with an *.

abdominal-A homeobox (*abd-A*) (Karch *et al.* 1990), although the nucleotide similarity is only 78%.

The sequence identity goes beyond the homeobox as well. Upstream of the *Manduca* homeobox the identity extends an additional five amino acids, then ends at an intron–exon splice junction in the *Drosophila* abd-A gene (see Fig. 3). We have sequenced an additional 1.2 kb upstream and found AT-rich sequences with a low coding sequence probability (data not shown).

At the other end of the homeobox, the similarity of the predicted *Manduca* amino acid sequence extends another 21 amino acids beyond the end of the *Drosophila* homeobox-containing exon (18/21 amino acids are identical; Fig. 3). There is no consensus donor splice site that would correspond to the intron splice junction at the 3′ end of the *Drosophila abd-A* homeobox exon. The *Drosophila abd-A* gene encodes a glutamine repeat thirty amino acids downstream from the homeobox (27 of the next 32 amino acids are glutamine; Karch *et al.* 1990). This extended glutamine repeat does not appear in the *Manduca* sequence, although there is a region 21 amino acids downstream from the homeobox where 5 out of 7 predicted amino acids are glutamine. The sequence similarity between the *Manduca* and *Drosophila* genes does not resume within the next 128 bp of sequence (data not shown). The sequence similarities both within the homeobox, and extending in either direction, strongly suggest that we have isolated a *Manduca* homologue of the *Drosophila abd-A* gene.

Expression of the Manduca abd-A *gene*

To determine the developmental profile and the tissue specificity of the *Manduca abd-A* gene, we monitored the expression by both northern blots and *in situ* hybridizations to tissue sections. When the 216 bp clone was hybridized to a northern blot of either total RNA or poly(A)+, only one 4.7 kb band was detected (Fig. 4). Transcripts were not detected in the newly oviposited egg, but were visible in the extended germ band (24% development) and throughout embryogenesis, followed by a decline prior to hatching (88% development) (Fig. 4A). A band of the same size was detected during the third and fifth (final) larval instars (Fig. 4A,B). A northern blot containing total RNA from isolated tissues of 5th instar larvae, either during feeding or at the onset of wandering (the beginning of metamorphosis), showed that the 4.7 kb band was present in the epidermis, ventral nerve cord and muscle (data not shown), but not in the Malpighian tubules or fat body (Fig. 4B).

In situ *hybridizations to tissue sections*

To aid in understanding the tissue sections of *Manduca* embryos, we have included a brief description of *Manduca* development from the cellular blastoderm stage through katatrepsis (approximately 50% of development). A more detailed description of these events can be found in Dorn *et al.* (1987) and Dow *et al.* (1988). Events in *Manduca* embryogenesis will be referred to as the percentage of the developmental time to hatching; at 26°C embryogenesis takes approximately 92 h.

The early *Manduca* germ anlage is saddle-shaped and is two times wider (dorsal–ventral axis) than it is long (anterior–posterior axis). Gastrulation involves extensive shape change of the germ anlage. The ventral furrow invaginates first in the presumptive gnathal-thoracic region of the anlage, and proceeds over the length of the embryo in either direction (Fig. 5A,B). As the ventral furrow lengthens, the width of the entire germ anlage is greatly reduced. The germ band extends along the circumference of the egg, so that the head and tail nearly meet on the dorsal side of the egg (20% development), then retracts over the next 20% of development so that the head and tail are now at opposite poles of the egg. The movements of katatrepsis push the embryo across the interior of the egg, changing the flexure of the embryo from convex to concave in relation to the anterior–posterior axis (Fig. 6A: prekatatrepsis, Fig. 7A: postkatatrepsis).

The earliest developmental stage analyzed by *in situ* hybridization was 12% development (Fig. 5). By this stage, gastrulation has progressed through the anterior abdominal segments while the posterior abdominal segments remain expanded, as seen in the whole mount of the embryo in Fig. 5A,B. *In situ* hybridizations to frontal sections, using the anti-sense strand of the 216 bp fragment, showed abundant silver grains in the

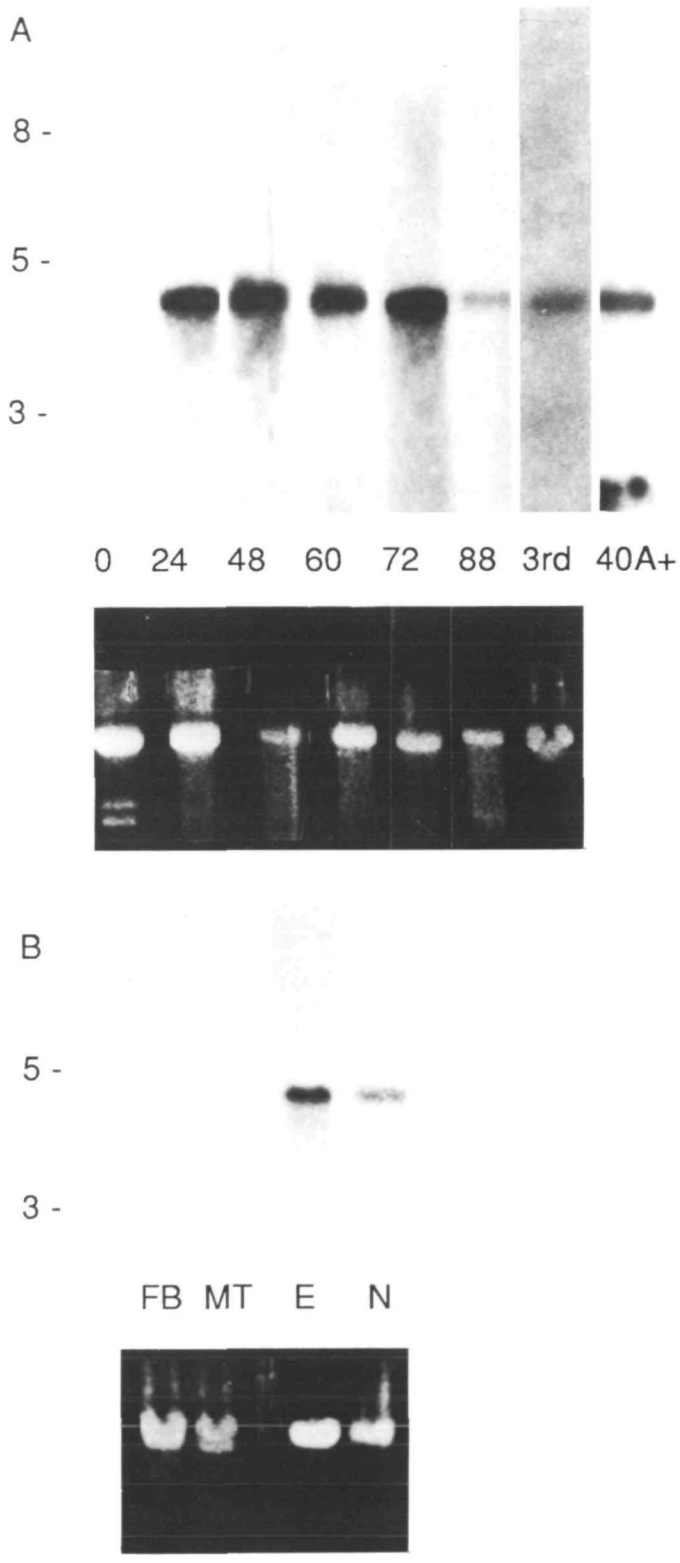

Fig. 4. Expression of *Manduca abd-A* during embryogenesis and late larval stages. (A) Northern blot of 15 μg total RNA from staged embryos (the numbers below the lanes indicate percentage of development completed) and from third instar larvae (3rd) The final lane contains 2 μg poly(A)$^+$ RNA from embryos that have completed 40% of development. (The spot in the lower portion of this lane is a hybridization artifact). The 3rd instar larval and poly(A)$^+$ RNA were run on separate gels (B) Northern blot of 15 μg total RNA from day 2 5th instar larval fat body (FB), wandering day 0 Malpighian tubules (MT), day 3 5th instar dorsal abdominal epidermis (E) and nerve cord (N). Beneath both blots A and B are photographs of the ethidium bromide staining of the RNA loaded in each lane (see Methods). The visible band is formed by the ribosomal RNA (the two lower bands in the 0 h lane are an unusual, yet consistent banding pattern that appears in early embryonic RNA). Transcript size is indicated in kilobases, based on the relative movement of ^{32}P-labeled DNA fragments. Blots were hybridized with the ^{32}P-labeled 216 bp anti-sense RNA probe. Autoradiographic exposures were 18 h

posterior 2/3 of the germ band (Fig. 5C,D,E,F). Signal is visible over both the presumptive ectoderm and mesoderm, but less abundant over the presumptive mesoderm (Fig. 5C,D). The exact limits of the silver grains with respect to segment number cannot be defined, but reconstruction of the serial sections confirms that signal is absent in the anterior portion of the gastrula, as well as in the most extreme posterior regions.

By 40% development, the embryo has completed germ-band elongation and retraction (Fig. 6A). Mouth parts and thoracic limb buds are substantially developed, and the posterior hindgut has extended one-third of the length of the embryo The nervous system is well developed, and the eighth, ninth and tenth abdominal ganglia are beginning to fuse. Hybridization signal from the *Manduca abd-A* transcripts formed an abrupt border on the midline of the first abdominal segment (Fig. 6C,D). No signal was detectable in the anterior portion of the first abdominal segment or any other more anterior segments (Fig. 6C,D), but was abundant over all the abdominal ganglia, as well as the abdominal ectodermal and mesodermal tissues through the tenth abdominal segment (Fig. 6E,F).

Tissue sections from 40% development (Fig. 6B), as well as from 12% and 50% development (data not shown), were all hybridized to the sense strand of the 216 bp fragment. In no case was signal above background detected.

By 50% development, the embryo has completed katatrepsis and begun dorsal closure (Fig. 7A). Note that the posterior-most ganglia in this stage embryo represents a fusion of eighth, ninth and tenth abdominal ganglia. The anterior boundary of expression remains the same at this stage. Transcripts were most abundant over all the abdominal ganglia posterior to the first abdominal segment, and present over the ectoderm and mesoderm of these same segments as well. Silver grains decreased in abundance in the ectoderm posterior to the last ganglia (Fig. 7B,C), and were absent over the most posterior ectodermal tissues (not visible in section shown).

Figs 7 D and E show a transverse section of the same stage embryo. Silver grains were abundant over the abdominal ganglia, epidermis and mesoderm, but were absent over the three paired Malpighian tubules and the

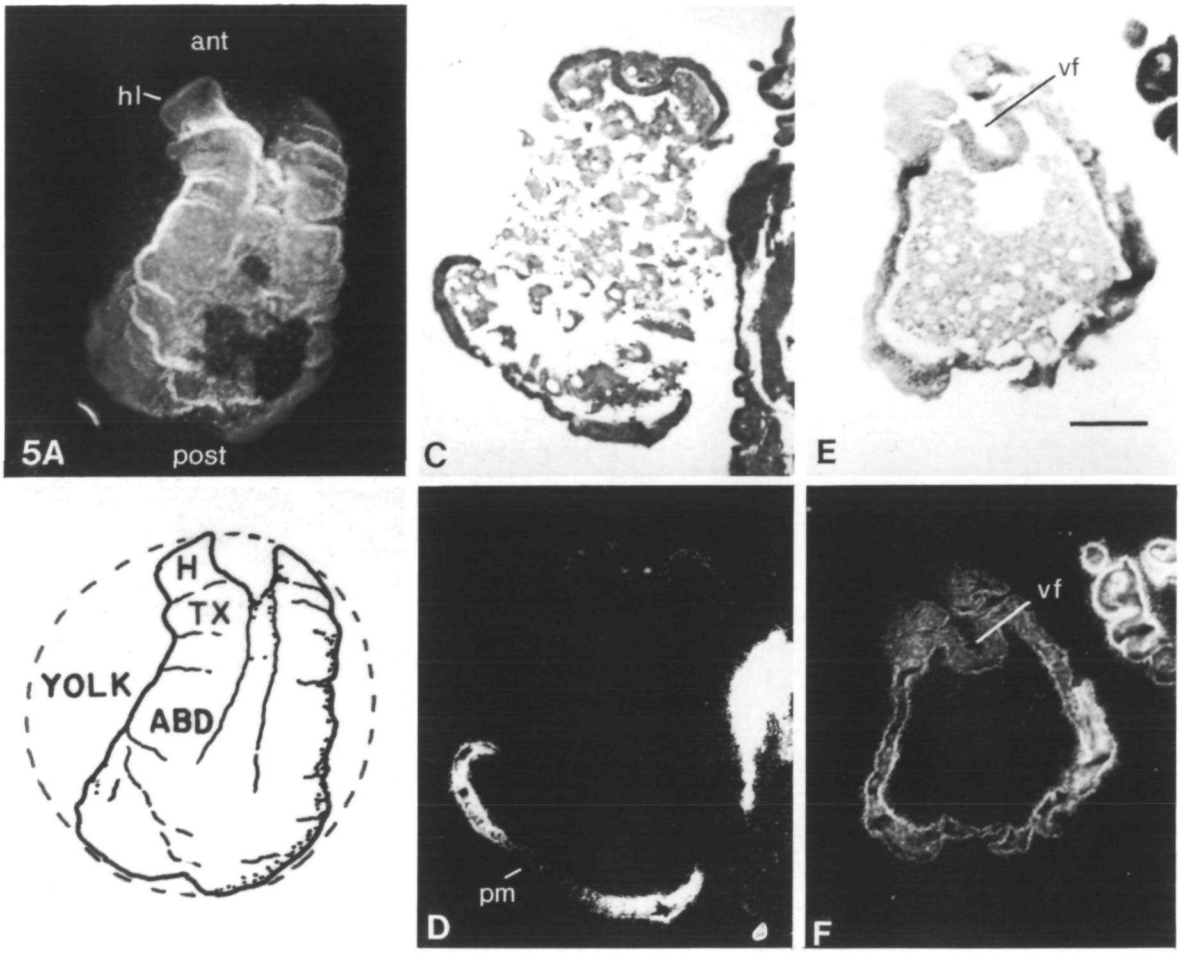

Fig. 5. Localization of the *Manduca abd-A* transcripts in tissue sections of embryos which have completed 12% of development (A) Whole mount preparation. (B) Diagrammatic representation of the same stage embryo, indicating the general location of head (H), thoracic (TX), and abdominal segments (ABD) (C) Frontal section, bright field. (D) Same section, dark-field. (E) Frontal section of the same embryo, more ventral to C. (F) Dark-field, same section as E. Anterior (ant) is up and posterior (post) is down in all frames. All sections are hybridized with the 216 bp anti-sense RNA probe. hl, head lobes; pm, presumptive mesoderm; vf, ventral furrow. Bar. 160 μm

hindgut. There were no grains visible above background in the head segments. This tissue specificity in the embryo matches that shown by the northern blot of RNA isolated from tissues from fifth instar and wandering stage larvae (Fig. 4B).

Discussion

We have isolated and partially sequenced a *Manduca* homeobox-containing gene. The predicted amino acid sequence of an 86 bp region of this gene is nearly identical to that of the homeobox-containing exon of the *Drosophila abd-A* gene (Karch *et al.* 1990). The sequence identity includes the homeobox, an additional five amino acids 5′ to the homeobox, and 18 out of 21 amino acids 3′ to the homeobox. The temporal expression of transcripts from the *Manduca* gene during both embryogenesis and postembryonic development mimics that from the *Drosophila* gene. The tissue-specific distribution of transcripts in the germ-band retracted embryo reflects the difference in visible segment number between the *Manduca* and *Drosophila* abdomen.

Sequence analysis

Genes that show a similar degree of identity to the *Drosophila abdominal-A* gene have now been isolated from four other insect species. *Drosophila melanogaster* (Karch *et al.* 1990), *Apis mellifera* (Walldorf *et al.* 1989), *Manduca sexta* (this paper), *Schistocerca gregaria* (Akam *et al.* 1988; Tear *et al.* 1990) and *Tribolium sp.* (Brown *et al.* 1990) all share a 72 amino acid stretch that extends in both directions beyond the homeobox. The conservation of this homeobox-containing domain in all 5 species (with the exception of position 22 in *Apis*) is remarkable. Treisman *et al.* (1989) and Hanes and Brent (1989) have shown that the asparagine in position 51 is critical for the DNA-binding specificity of the *Drosophila* homeobox domain. Yet it is clear that other regions of the domain are also important in forming the protein–DNA complex (Otting *et al.* 1990; Percival-

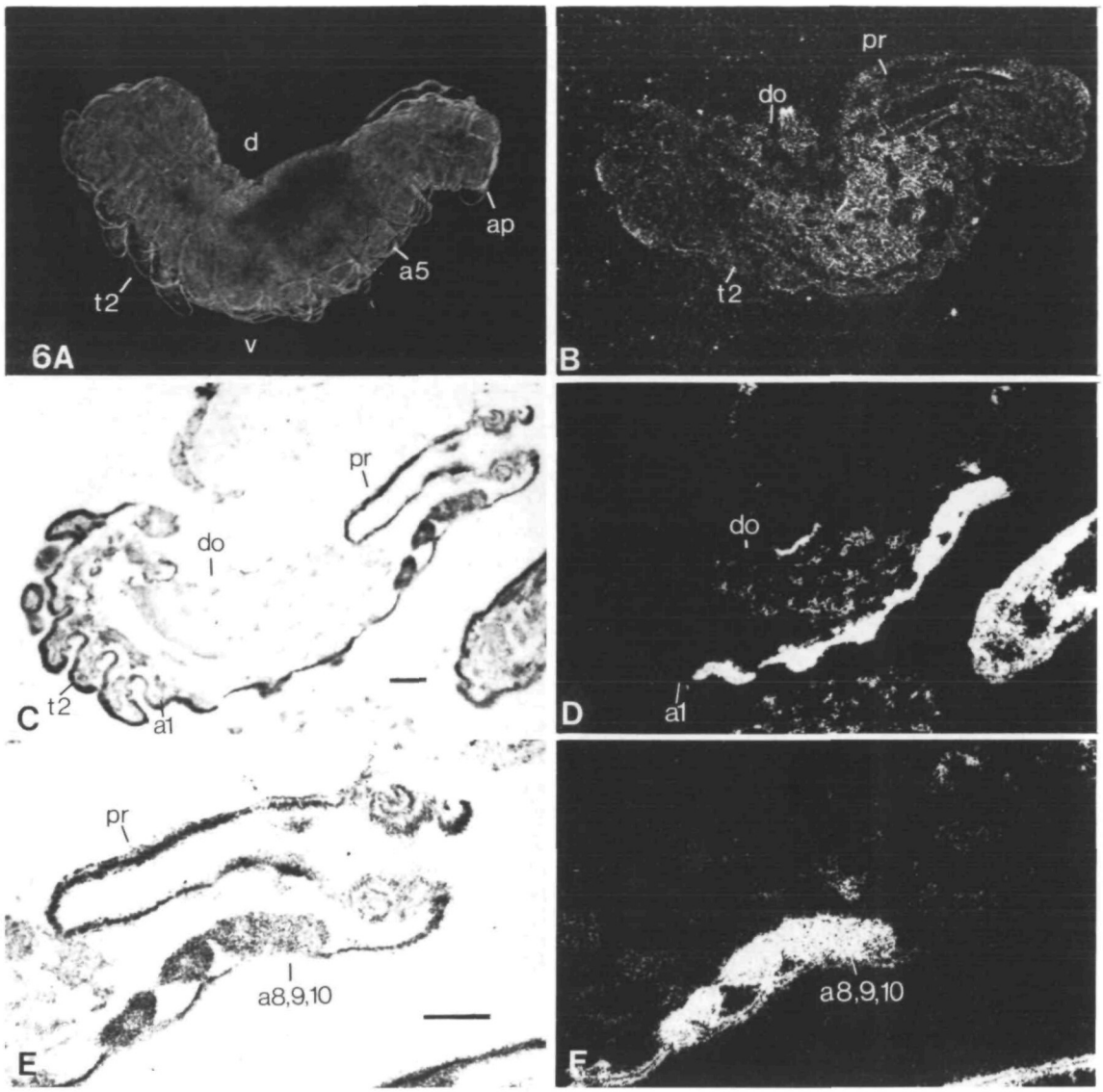

Fig. 6. Localization of the *Manduca abd-A* transcripts in tissue sections of embryos which have completed 40% development (A) Whole mount. (B) Sagittal section of a similarly staged embryo hybridized with the sense strand of the 216 bp homeobox fragment, dark field (C) Sagittal section, bright field (the thoracic ganglia are not visible in this section because the angle of sectioning was not exactly sagittal, but a few degrees off the axis in the direction of the transverse plane). (D) The same section as in C, dark field. (E) Higher magnification of the posterior ganglia of the same section. (F) Same section as in E, dark field. In these sections anterior is to the left. a1, first abdominal segment; a5, fifth abdominal segment; a8,9,10, fused ganglia of the eighth, ninth and tenth abdominal ganglia; ap, anal prolegs, d, dorsal; do, dorsal opening to the yolk; pr, proctodeum; t2, second thoracic segment; v, ventral. Sections in C,D,E,F are all hybridized with the 216 bp anti-sense RNA probe. Bar 90 μm

Smith *et al.* 1990) and in determining its target specificity (Hayashi and Scott, 1990). The resistance of the whole region to change-over evolution underscores its important functional role.

In addition to the amino acid conservation, the intron–exon splice junction 5′ to the homeobox also appears to be conserved among *Drosophila*, *Manduca*, *Schistocerca* and *Tribolium*. The *Manduca* gene appears to share this exon boundary by virtue of the conserved acceptor splice junction sequence (pyr,pyr, pyr,NAG) (Breathnach and Chambon, 1981) present at this site, and by the fact that the sequences immediately preceding this junction in the *Manduca* sequence are extremely AT-rich and the coding sequence probability decreases significantly. The *Apis abd-A* homologue does not appear to share this splice junction, as the sequence homology of the genomic clone continues an additional 17 amino acids 5′ to this junction without any disruption (Walldorf *et al.* 1989). The splice junction 3′ to the homeobox in *Drosophila* is not conserved in any of the other insects analyzed.

The sequence homology between the *Drosophila abd-A* cDNA and the *Manduca* genomic clones ends at a region that encodes a small glutamine repeat (5 amino

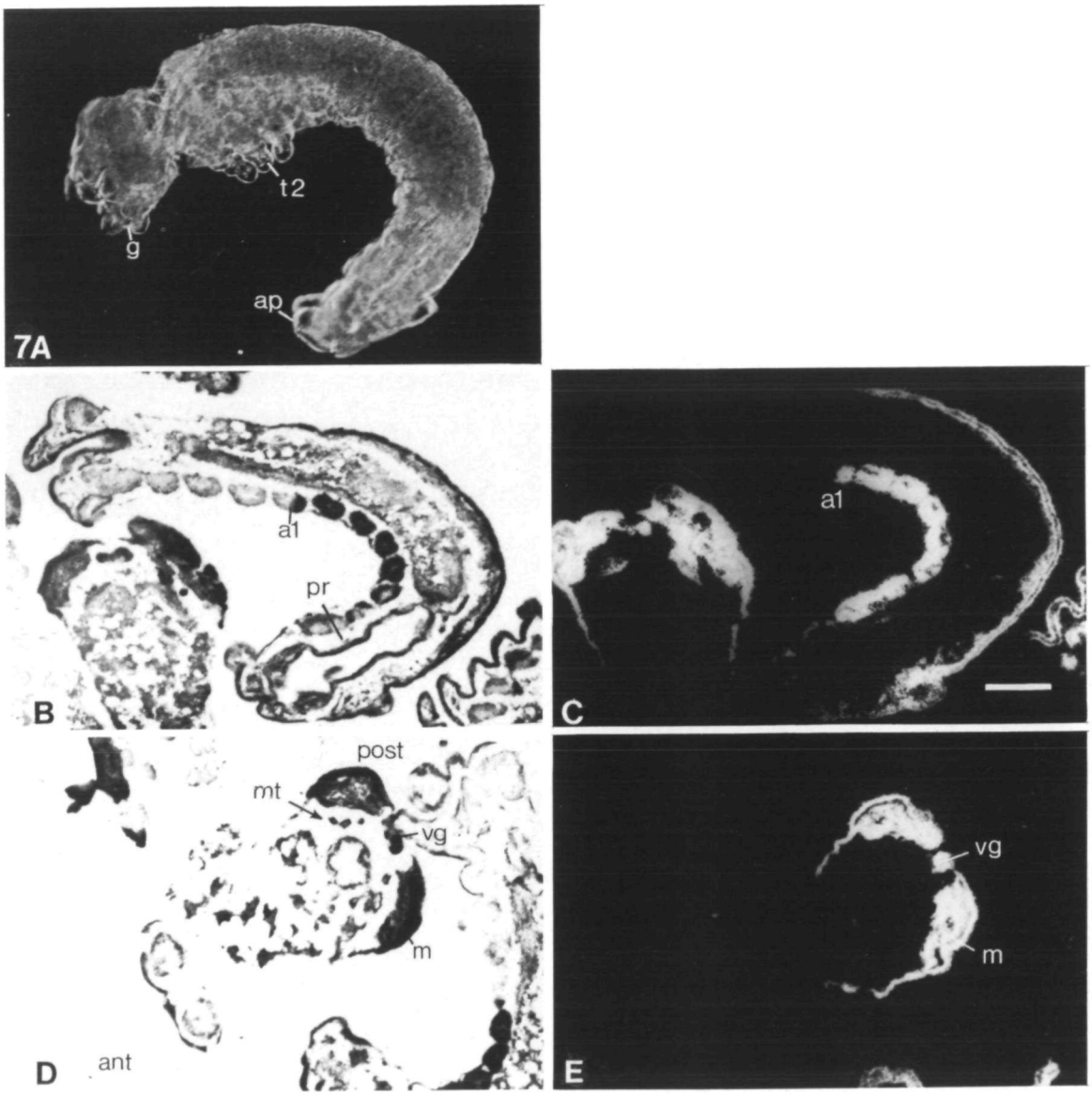

Fig. 7. Localization of transcripts from the *Manduca abd-A* gene in tissue sections of embryos which have completed 50% development. (A) Whole mount; (B) Sagittal section, bright field; (C) Same section as in B, dark field, (D) Transverse-section, bright field; and (E) same section as in D, dark field. (As the *Manduca* embryo normally lies curved in its egg chamber at this stage, a transverse section provides a simultaneous section through an abdominal segment (post) and a more anterior segment). a1, first abdominal segment, g, gnathal segments; m, mesoderm; mt, Malpighian tubules, vg, ventral ganglia Anterior (ant) is to the left in all frames. All sections are hybridized with the 216 bp anti-sense RNA probe. Bar· 160 μm

acids) in the *Manduca* sequence. This position is 8 amino acids upstream of a glutamine repeat in the translated *Drosophila* cDNA. Glutamine-rich regions are present in a number of known transcription factors and in most of the *Drosophila* homeobox-containing genes although their position in relation to the homeobox varies (Regulski *et al.* 1985; Laughon *et al.* 1985). There is evidence from mammalian transcriptional regulators that the presence of a glutamine-rich segment enhances transcriptional activation (Courey and Tjian, 1988), although the significance of the size or position of this segment is as yet unknown.

Preliminary reports indicate that multiple transcripts are produced from the *Drosophila abd-A* gene: 5.4 and 4.8 kb were reported by Rowe and Akam (1988), and 5.4 and 5.1 kb by Karch *et al.* (1990). Using a *Manduca abd-A* homeobox-containing probe, we detect only one 4.7 kb band on a northern blot containing either total or poly(A)+ RNA. This is consistent with the lack of conservation of splice junctions between the *Manduca* and *Drosophila* sequences. Further analysis of differential splicing in both the *Drosophila* and *Manduca* genes might reveal regulatory changes that have occurred in evolution.

Temporal expression and tissue specificity of abdominal-A *transcripts*

In *Drosophila*, homeotic genes are thought to have a role in maintenance of segment character throughout the life cycle of the fly (Garcia-Bellido and Lewis, 1976; Morata and Garcia-Bellido, 1976). The expression pattern of the *Manduca abd-A* gene indicates a similar requirement throughout the life cycle of *Manduca*: *abd-A* transcripts first appear in the maturing cellular blastoderm or early gastrula and expression is maintained throughout embryogenesis and larval life in both insects (Karch *et al.* 1990). The *Manduca abd-A* transcripts are also distributed amongst similar tissue types as the *Drosophila* abd-A protein (Karch *et al.* 1990): they are found in epidermal, mesodermal and neuronal tissues, but not in the Malpighian tubules.

The spatial distribution of transcripts in the late embryo reveals a striking similarity between *Manduca* and *Drosophila*. In both insects, the transcripts form an abrupt boundary in the middle of the first abdominal segment, the boundary that marks the beginning of parasegment (ps) 7 in *Drosophila*. The conservation of this domain of expression for a homeotic gene indicates the conservation of parasegments as domains of gene activity in *Manduca*. Whether the initial morphological metameric divisions in the *Manduca* embryo correspond to segmental or parasegmental iterations is not known.

In contrast, the posterior boundary of *abdominal-A* expression does not occur in the same segment in *Manduca* and *Drosophila*. The *Drosophila abd-A* protein is present from ps7 through the middle of ps13, indicating that its posterior boundary is not parasegmental (Karch *et al.* 1990; Macias *et al.* 1990). The posterior boundary to the expression of the transcript is less clearly resolved. *In situ* hybridization to germband-retracted embryos showed an abundant signal from ps7 to ps12 and a weaker signal in ps13 (Rowe and Akam, 1988). Harding *et al.* (1985) and Regulski *et al.* (1985) found an additional weak signal in ps14. In the *Manduca* embryo, transcripts appear equally abundant over ten abdominal segments (equivalent to ps7–15.5 in *Drosophila*), then are absent in the terminalia. Thus, as in *Drosophila*, the posterior boundary is not parasegmental, but it extends two segments further than in *Drosophila*. In the *Schistocerca* embryo, the *abd-A* protein also initially extends through the tenth abdominal segment, then later in development retracts to the ninth segment (Tear *et al.* 1990). The observed *abd-A* expression pattern correlates with the difference in the number of visible segments in the abdomen of these insects: *Manduca* and *Schistocerca* embryos have ten well-developed abdominal segments compared to the eight in *Drosophila*.

The conservation of parasegment 7 as the anterior boundary of *abd-A* expression and the divergence of the posterior boundary among these three insects suggests that these boundaries may be differently regulated in development and evolution. Although the specific regulation of the onset of transcription of *abd-A* is not well-characterized in *Drosophila*, *Abd-B* mutations affect the limits of *abd-A* expression so that late in development, *abd-A* appears in ps14 and 15 (Macias *et al.* 1990). Thus, *Abd-B* might be responsible for the late retraction of *abd-A* in *Schistocerca*. Other factors apparently act earlier to delimit the initial domain of *abd-A* expression in *Drosophila* (Macias *et al.* 1990) The variation in this domain among these three insects suggests that all of these factors may not be evolutionarily conserved.

Early pattern formation in the Manduca *embryo*

Lepidopterans are not easily classified into any of the categories of long, intermediate or short germ insects (see Sander, 1976). Segments appear with an obvious anterior–posterior progression, but it is not known when segmental fates are established: are segments specified in the cellular blastoderm, or are they progressively determined as segments appear during gastrulation? Luscher (1944) was able to cause disruption of individual segments by UV-irradiating transverse strips of the lepidopteran *Tineola* egg prior to the blastoderm stage. Similarly, Myohara and Kiguchi (1990) were able to construct a segment-specific fate map for *Bombyx mori* within 2 h of egg deposition (the cellular blastoderm is normally formed about 12 h). Although these results do not necessarily mean that determination has already occurred at the time of irradiation (see Sander, 1976 for discussion), they do show that the anlage for each abdominal segment can be mapped at the cellular blastoderm stage and does not undergo subsequent expansion Carr and Taghert (1989), using the onset of *Manduca* fasciculin II (TN1) as a marker, suggest that segments are progressively determined as gastrulation proceeds in *Manduca*. Our experiments show an abundance of transcripts in the presumptive abdominal region at 12 % of development, before this region has undergone the movements of gastrulation. This result suggests that, at least in relation to *abd-A* expression, the abdominal segments have some information as to their character before they are visibly distinguishable as abdominal segments.

We have shown that there is an extremely high degree of conservation of the amino acid sequences of the homeotic gene *abdominal-A*. Not only is the sequence conserved, but two of the principles of homeotic gene function – a parasegmental domain of gene expression and persistence of expression throughout the life of the insect – also appear to be conserved. The differences that we have detected appear in the regulation of the *abdominal-A* gene: its initial embryonic distribution is over a greater number of segments and, secondly, differential splicing of the mRNA is not detectable. It is not surprising to find the homeotic genes conserved in other insects but, nonetheless, the generation of morphologically diverse insects that respond differently to experimental perturbations remains a mystery. Our results suggest that the regulation of the homeotic gene products may begin to account for these differences.

We thank Frank Horodyski and John Rebers for their

undying patience and expert technical advice; M. Akam for pointing out the *abd-A* homology from unpublished sequences; R. Garber for the *Drosophila* probes, and G. Schubiger and B. Wakimoto for their comments on a draft of this manuscript. This study was supported by NIH AI12459 RB was supported by NIH NS-13079. LMN was supported by NIH Training Grant GM072780

References

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

Akam, M , Dawson, I and Tear, G. (1988) Homeotic genes and the control of segment diversity *Development* **104**, Sup 123–133

Angerer, L M , Stoler, M. H. and Angerer, R. C (1987) *In situ* hybridization with RNA probes. An annotated recipe In *In situ hybridization Applications to Neurobiology* (ed. K. L. Valentine, J H Eberwine and J D Barchas), pp 42–70 Oxford University Press, New York.

Aviv, H. and Leder, P (1972). Purification of biologically active globin messenger RNA by chromatography on oligothymidylic acid-cellulose *Proc natn. Acad Sci U S A* **69**, 1408–1412

Beeman, R W. (1987) A homeotic gene cluster in the red flour beetle *Nature* **327**, 247–249

Biggin, M D , Gibson, T. J. and Hong, G. F. (1983). Buffer gradient gels and ^{35}S label as an aid to rapid DNA sequence determination *Proc natn Acad Sci U S A* **80**, 3963–3965.

Booker, R and Truman, J W (1989) *Octopod*, a homeotic mutation of the moth *Manduca sexta*, influences the fate of identifiable pattern elements within the CNS *Development* **105**, 621–628

Breathnach, R. and Chambon, P. (1981) Organization and expression of eukaryotic split-genes coding for proteins *A. Rev Biochem* **50**, 349–383

Brown, S , Stewart, J , Beeman, R. and Denell, R (1990). Molecular genetics of *Tribolium* homeotic genes. Abstract, 31st *Ann Drosophila Res Conf* , March, 1990.

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

Carr, J N. and Taghert, P. H (1989) Pair-rule expression of a cell surface molecule during gastrulation of the moth embryo. *Development* **107**, 143–152

Courey, A J and Tjian, R (1988). Analysis of SP1 *in vivo* reveals multiple transcriptional domains, including a novel glutamine-rich activation motif. *Cell* **55**, 887–898.

Dale, R , McClure, B. and Houchins, J. (1985) A rapid single-stranded cloning strategy for producing a sequential series of overlapping clones for use in DNA sequencing Application to sequencing the corn mitochondrial 18S rDNA *Plasmid* **13**, 31–40

Davis, L G , Dibner, M. D. and Battey, J F (1986) *Basic Methods in Molecular Biology* Elsevier, N Y

Denhardt, D. (1966) A membrane-filter technique for the detection of complementary DNA *Biochem Biophys Res Commun* **23**, 641–651.

Dorn, A., Bishoff, S T and Gilbert, L I (1987) An incremental analysis of the embryonic development of the tobacco hornworm, *Manduca sexta Int J Inv Reprod. Dev* **11**, 137–158

Dow, R C , Carlson, S D and Goodman, W G (1988) A scanning electron microscope study of the developing embryo of *Manduca sexta* (L) (Lepidoptera·Sphingidae) *Int J Insect Morph. Embryol.* **17**, 231–242

Feinberg, A P and Vogelstein, R. (1984). A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity *Analyt Biochem* **137**, 266–267

Fleig, R. and Sander, K. (1986) The embryogenesis of the honeybee *Apis mellifera* L (Hymenoptera Apidae) a SEM study. *Int J Insect Morph Embryol.* **15**, 449–462

Fleig, R., Walldorf, U , Gehring, W J and Sander, K (1988) *In situ* localization of the transcripts of a homeobox gene in the honeybee *Apis mellifera* L (Hymenoptera). *Roux's Arch devl Biol* **197**, 269–274

Garber, R L , Kuroiwa, A and Gehring, W. J (1983) Genomic and cDNA clones of the homeotic locus *Antennapedia* in *Drosophila EMBO J* **2**, 2027–2036

Garcia-Bellido, A (1977) Homeotic and atavistic mutations in insects *Am Zool* **17**, 613–629

Garcia-Bellido, A and Lewis, E. B. (1976) Autonomous cellular differentiation of homeotic *Bithorax* mutants of *Drosophila. Devl Biol* **48**, 400–410

Gehring, W J (1987). Homeo boxes in the study of development *Science* **236**, 1245–1252

Hanes, S. D. and Brent, R (1989) DNA specificity of the bicoid activator protein is determined by homeodomain recognition helix residue 9. *Cell* **57**, 1275–1283

Harding, K , Wedeen, C , McGinnis, W and Levine, M (1985). Spatially regulated expression of homeotic genes in *Drosophila Science* **229**, 1236–1242

Hayashi, S and Scott, M P (1990) What determines the specificity of action of *Drosophila* homeodomain proteins? *Cell* **63**, 883–894

Ingham, P (1988) The molecular genetics of embryonic pattern formation in *Drosophila. Nature* **336**, 25–34.

Karch, F., Bender, W and Weiffenbach, B (1990) *AbdA* expression in *Drosophila* embryos. *Genes Dev* **4**, 1573–1587

Kessel, M. and Gruss, P (1990) Murine developmental control genes *Science* **249**, 375–379

Kuhn, A. (1971) *Lectures on Developmental Physiology* (translated by R. Milkman), Springer-Verlag, New York.

Laughon, A., Carroll, S B , Storfer, F A , Riley, P D. and Scott, M P (1985) Common properties of proteins encoded by the *Antennapedia* complex genes of *Drosophila melanogaster Cold Spring Harbor Symp. quant Biol* **50**, 253–262

Levine, M and Hoey, T (1988). Homeobox proteins as sequence specific transcription factors *Cell* **55**, 537–540

Lewis, E B (1978) A gene complex controlling segmentation in *Drosophila. Nature* **276**, 565–570

Lüscher, M (1944) Experimentale Untersuchungen uber die larvale und die imaginale Determination im Ei der Kleidermotte (*Tineola biselliella* Hum). *Revue suisse Zool.* **51**, 531–627

Macias, A , Casanova, J. and Morata, G (1990) Expression and regulation of the *abd-A* gene of *Drosophila. Development* **110**, 1197–1207

Mahaffey, J W. and Kaufman, T C (1987) The homeotic genes of the Antennapedia complex and the bithorax complex of *Drosophila* In *Developmental Genetics of Higher Organisms* (ed G. M Malcinski), pp. 329–360 New York Macmillan Publishing Co

McGinnis, W , Levine, M S , Hafen, E., Kuroiwas, A and Gehring, W J (1984). A conserved DNA sequence in homeotic genes of the *Drosophila* Antennapedia and *bithorax* complexes *Nature* **308**, 428–433

McGinnis, W , Shermoen, A. W and Beckendorf, S K (1983) A transposable element inserted just 5' to a *Drosophila* glue protein alters gene expression and chromatin structure *Cell* **34**, 75–84

Meinhardt, H (1982) *Models of Biological Pattern Formation.* Academic Press, London

Meinhardt, H (1986) Hierarchical inductions of cell states· a model for segmentation in *Drosophila. J. Cell Sci. Suppl* **4**, 357–381

Morata, G and Garcia-Bellido, A (1976) Developmental analysis of some mutants of the *bithorax* system of *Drosophila Wilhelm Roux's Arch devl Biol* **179**, 125–143.

Morata, G. and Lawrence, P A (1977) Homeotic genes, compartments and cell determination in *Drosophila Nature* **265**, 211–216

Myohara, M and Kiguchi, K (1990) Induction of localized cuticle defects by UV laser irradiation of early embryos in *Bombyx mori* Linne *Proc Arthropod Embryol Soc Jpn* (25) 13–14

Otting, G., Quian, Y Q , Billeter, M., Müller, M., Affolter,

M., Gehring, W. J. and Wiithrich, K. (1990). Protein-DNA contacts in the structure of a homeodomain-DNA complex determined by nuclear magnetic resonance spectroscopy in solution. *EMBO J.* 9, 3085-3092.

OUWENEEL, W. J. (1976). Developmental genetics of homoeosis. *Adv. Genet.* **18,** 179-238.

PERCIVAL-SMITH, A., MÜLLER, M., AFFOLTER, M. AND GEHRING, W. (1990). The interaction of DNA of wild-type and mutant *fushi tarazu* homeodomains. *EMBO J.* 9, 3967-3974.

PUSTELL, J. AND KAFATOS, F. C. (1982). A convenient and adaptable package of computer programs for DNA and protein sequence management, analysis, and homology determination. *Nucl. Acids Res.* 12, 643-645.

REBERS, J., HORODYSKl, F., HICE, R. AND RIDDIFORD, L. M. (1987). Genomic cloning of developmentally regulated larval cuticle genes from the tobacco hornworm, *Manduca sexta.* In *Molecular Entomology* (ed. J. H. Law), pp. 201-210. Alan R. Liss, New York.

REGULSKI, M., HARDING, K., KOSTRIKEN, R., KARCH, F., LEVINE, M. AND MCGINNIS, W. (1985). Homeo box genes of the *Antennapedia* and *Bithorax* complexes of *Drosophila. Cell* 43, 71-80.

ROWE, A. AND AKAM, M. (1988). The structure and expression of a hybrid homeotic gene. *EMBO J.* 7, 1107-1114.

SANDER, K. (1976). Specification of the basic body pattern in insect embryogenesis. *Adv. Ins. Phys.* 12, 125-238.

SANDER, K. (1983). The evolution of patterning mechanisms: gleanings from insect embryogenesis and spermatogenesis. In *Development and Evolution* (ed. B. P. Goodwin, N. Holder and C. C. Wylie), pp. 137-159. Cambridge University Press, Cambridge.

SANDER, K. (1984). Embryonic pattern formation in insects: Basic concepts and their experimental foundations. In *Pattern Formation. A Primer in Developmental Biology* (ed. G. M. Malacinski and S. V. Bryant), pp. 245-268. Macmillan Publishing Co., New York.

SANGER, F., NICKLEN, S. AND COULSON, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. natn. Acad. Sci. U.S.A.* 74, 5463-5467.

SASAKI, M. AND RIDDIFORD, L. M. (1984). Regulation of reproductive behavior and egg maturation in the tobacco hawk moth, *Manduca sexta. Phys. Ent.* 9, 315-327.

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. Ill (1989). The structure and function of the homeodomain. *BBA Rev. Cancer* **989,** 25-48.

SCOTT, M. P. AND WEINER, A. J. (1984). Structural relationships among genes that control development sequence homology between the *Antennapedia, Ultrabithorax,* and *fushi tarazu* loci of *Drosophila. Proc. natn. Acad. Sci. U.S.A.* 81, 4115-4119.

SEIDEL, F. (1971). The determination of the germ band in insects. In *Milestones in Developmental Physiology of Insects* (ed. D. Bodenstein), pp. 9-53. Appleton-Century-Crofts, New York.

SIBATANI, A. (1980). Wing homeosis in Lepidoptera: A survey. *Devi Biol.* 79, 1-18.

STROEHER, V. L., JORGENSEN, E. M. AND GARBER, R. L. (1986). Multiple transcripts from the *Antennapedia* gene of *Drosophila melanogaster. Molec. cell. Biol.* 6, 4667-4675.

TAZIMA, Y. (1964). *The Genetics of the Silkworm,* pp. 60-75. Logos Press, Nat'l Inst. of Gen., Japan.

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,** 915-925.

TREISMAN, J., GONCZY, P., VASHISHTHA, M., HARRIS, E. AND DESPLAN, C. (1989). A single amino acid can determine the DNA binding specificity of homeodomain proteins. *Cell* 59, 553-562.

WALLDORF, U., FLIEG, R. AND GEHRING, W. J. (1989). Comparison of homeobox-containing genes of the honeybee and *Drosophila. Proc. natn. Acad. Sci. U.S.A.* 86, 9971-9975.

YANISCH-PERRON, C, VIEIRA, J. AND MESSING, J. (1985). Improved M13 phage cloning vectors and host strains: Nucleotide sequences of the M13mpl8 and pUC19 vectors. *Gene* 33, 103-119.

(Accepted 6 February 1991)