Development 128, 3459-3472 (2001)
Printed in Great Britain © The Company of Biologists Limited 2001
DEV5933

Grasshopper *hunchback* expression reveals conserved and novel aspects of axis formation and segmentation

Nipam H. Patel[1,*], David C. Hayward[2], Sabbi Lall[1], Nicole R. Pirkl[1], Daniel DiPietro[1] and Eldon E. Ball[2]

[1]Department of Organismal Biology and Anatomy and Howard Hughes Medical Institute, University of Chicago, 5841 S. Maryland Ave., MC1028, Chicago, IL 60637, USA
[2]Molecular Genetics and Evolution Group, Research School of Biological Sciences, PO Box 475, Canberra, A.C.T. 2601, Australia
*Author for correspondence (e-mail: npatel@midway.uchicago.edu)

Accepted 25 June 2001

SUMMARY

While the expression patterns of segment polarity genes such as *engrailed* have been shown to be similar in *Drosophila melanogaster* and *Schistocerca americana* (grasshopper), the expression patterns of pair-rule genes such as *even-skipped* are not conserved between these species. This might suggest that the factors upstream of pair-rule gene expression are not conserved across insect species. We find that, despite this, many aspects of the expression of the *Drosophila* gap gene *hunchback* are shared with its orthologs in the grasshoppers *S. americana* and *L. migratoria*.

We have analyzed both mRNA and protein expression during development, and find that the grasshopper *hunchback* orthologs appear to have a conserved role in early axial patterning of the germ anlagen and in the specification of gnathal and thoracic primordia. In addition, distinct stepped expression levels of *hunchback* in the gnathal/thoracic domains suggest that grasshopper *hunchback* may act in a concentration-dependent fashion (as in *Drosophila*), although morphogenetic activity is not set up by diffusion to form a smooth gradient.

Axial patterning functions appear to be performed entirely by zygotic *hunchback*, a fundamental difference from *Drosophila* in which maternal and zygotic *hunchback* play redundant roles. In grasshoppers, maternal *hunchback* activity is provided uniformly to the embryo as protein and, we suggest, serves a distinct role in distinguishing embryonic from extra-embryonic cells along the anteroposterior axis from the outset of development – a distinction made in *Drosophila* along the dorsoventral axis later in development.

Later *hunchback* expression in the abdominal segments is conserved, as are patterns in the nervous system, and in both *Drosophila* and grasshopper, *hunchback* is expressed in a subset of extra-embryonic cells. Thus, while the expected domains of *hunchback* expression are conserved in *Schistocerca*, we have found surprising and fundamental differences in axial patterning, and have identified a previously unreported domain of expression in *Drosophila* that suggests conservation of a function in extra-embryonic patterning.

Key words: Hunchback, *Drosophila*, *Schistocerca*, *Locusta*, Segmentation, Axis formation, Grasshopper

INTRODUCTION

A great deal is known about the molecular mechanisms responsible for patterning and segmentation during *Drosophila melanogaster* embryogenesis, and this knowledge provides an excellent framework for examining both the conservation and the evolution of these processes in the arthropod phylum. Such comparative data not only provide a more accurate assessment of how these processes vary between insects, but also of how early developmental mechanisms evolve in all animals.

Genetic and molecular analyses have demonstrated that segmentation in *D. melanogaster* depends on progressive subdivision by a hierarchy of regulatory factors (Rivera-Pomar and Jäckle, 1996; St Johnston and Nüsslein-Volhard, 1992). Initially, maternal gradients of *bicoid* and *nanos* provide information that defines position along the length of the syncytial embryo. Translational repression by bicoid and nanos result, respectively, in gradients of caudal and hunchback (Hb) protein. The bicoid, Hb and caudal gradients lead to activation of zygotic gap gene transcription at distinct positions along the anteroposterior (AP) axis. Regulation by maternal and gap gene products then leads to the expression of pair-rule genes in stripes of a two-segment periodicity. The pair-rule genes thus define the initial reiterated pattern within the embryo. Segmental pattern is later maintained and refined by the expression, under pair-rule regulation, of the segment polarity genes.

Both the data summarized above, and the results of various embryonic manipulations, show that in *Drosophila* the entire body plan is established simultaneously across the length of the blastoderm embryo. This type of development (in which body regions are present at blastoderm in the same proportions as are found in the hatching larva) is known as long germ embryogenesis (reviewed by Sander, 1976). By contrast, short

germ embryos, such as those of the grasshoppers *Locusta* and *Schistocerca*, appear to pattern only the head region of the body before gastrulation, more posterior regions being patterned during a subsequent growth phase. Intermediate germ insects fall somewhere between these extremes. The long germ mode of development is seen only in the most phylogenetically derived insect orders, and thus ancestral insect development was probably closer to short or intermediate germ embryogenesis.

Significantly, features of short germ development suggest that the mechanisms of pattern formation in these embryos may be fundamentally different from those seen in *Drosophila*. For example, during formation of the grasshopper abdomen, segments are defined sequentially along the AP axis, suggesting temporal disparity in the response to positional information along this axis. In addition, cellularization of the grasshopper blastoderm occurs relatively early (Ho et al., 1997), so that positional information must be interpreted in a cellular environment: this differs from the fly segmentation paradigm, where a syncytial environment seems crucial to patterning.

Despite the very different embryogenesis of short germ insects, there is evidence that some molecular aspects of *Drosophila* segmentation are conserved. In particular the segment polarity genes *engrailed* and *wingless* are expressed in a conserved pattern at parasegment boundaries in many insects including the grasshopper, *Schistocerca*, and beetle, *Tribolium castaneum* (Brown et al., 1994b; Dearden and Akam, 2001; Nagy and Carroll, 1994; Patel et al., 1989a; Patel et al., 1989b). However, earlier events in *Drosophila* segmentation appear less conserved in short germ insects. Although genetic and molecular data suggest that pair-rule patterning is conserved in *Tribolium* (Brown et al., 1994a; Maderspacher et al., 1998; Patel et al., 1992; Patel et al., 1994; Schröder et al., 1999; Sommer and Tautz, 1993), such evidence has proven more elusive in the grasshopper. Recently, a homolog of *Drosophila paired* (*prd*) has been found to be expressed with two-segment periodicity during grasshopper segmentation (Davis et al., 2001). Given, however, that the grasshopper *eve* and *ftz* orthologs are not expressed in a pair-rule pattern, significant differences in the mechanisms of segmentation between *Drosophila* and the grasshopper may exist (Dawes et al., 1994; Patel et al., 1992). Reasoning that the source of these differences might lie upstream, we wondered whether the expression of regulators of *Drosophila* pair-rule genes might also differ significantly in the grasshopper.

In order to better understand the similarities and differences between long and short germ embryogenesis, we have tried to identify orthologs of the earliest factors involved in the *Drosophila* segmentation hierarchy, and we focus here on *hb*. The gap gene *hb* encodes a zinc-finger transcription factor that functions early in development, both as a maternal morphogen and a gap gene. During *Drosophila* oogenesis, the egg is loaded ubiquitously with maternal *hb* transcript (Tautz et al., 1987). Upon fertilization, maternal *hb* transcript is translated only anteriorly, owing to translational repression by a gradient of maternal nanos protein emanating from the posterior end of the embryo (Hülskamp et al., 1989; Irish et al., 1989; Wang and Lehmann, 1991). This translational repression results in the formation of a Hb protein gradient derived from the maternally supplied mRNA (Tautz, 1988). Previous studies in *Drosophila* reveal that this translational repression depends on the 3′ untranslated region (3′UTR) of the *hb* transcript. Here, a bipartite sequence (the nanos response element, NRE) is recognized by pumilio protein, which in turn recruits nanos (Murata and Wharton, 1995; Sonoda and Wharton, 1999; Wharton and Struhl, 1991). Together nanos and pumilio act to prevent *hb* translation in the posterior part of the embryo. The resulting Hb protein gradient is known to behave morphogenetically, regulating its own promoter, as well as other gap genes such as *Krüppel*, *knirps* and *giant* in a concentration-dependent manner (Hülskamp et al., 1990; Simpson-Brose et al., 1994; Struhl et al., 1992).

Zygotic expression of *Drosophila hb* begins at the blastoderm stage (Fig. 1A; Tautz et al., 1987). *Drosophila hb* mRNA is transcribed from two promoters, P1 and P2, resulting in transcripts of 3.2 and 2.9 kb, respectively (Schröder et al., 1988). The P2 promoter is activated by the anterior morphogen bicoid, while the 3.2 kb transcript is activated across parasegment 4 (PS4) by, among other factors, Hb itself (Schröder et al., 1988; Tautz, 1988). The zygotic Hb protein gradient (Fig. 1A) is similar to its maternal counterpart. Indeed the two gradients are almost redundant: maternal *hb* is dispensable for AP patterning, and zygotic *hb* activation by bicoid almost is (Hülskamp et al., 1989; Irish et al., 1989; Wimmer et al., 2000).

Anterior *hb* expression is crucial for the development of thoracic segments. In addition, a posterior domain of zygotic *hb*, under independent control, emerges and refines to a domain spanning PS13/14 of the embryo (Fig. 1A; see Tautz, 1988; Tautz et al., 1987). Loss of this expression accounts for the fusion of segments A7 and A8 in the *hb* mutant. Later in development, *hb* is expressed in specific mesodermal cells and in the nervous system. Within the nervous system, transient *hb* expression is seen in neuroblasts and in a sub-population of ganglion mother cells (GMCs) and neurons (Fig. 1B,C,E; see Kambadur et al., 1998; Tautz et al., 1987; Wolf and Schuh, 2000).

hb orthologs have been isolated from other dipterans including *D. virilis*, *Megaselia abdita* and the midge, *Clogmia albipunctata* (Rohr et al., 1999; Sommer et al., 1992; Stauber et al., 2000; Treier et al., 1989). As in *D. melanogaster*, an anterior domain of *hb* mRNA is observed for all three species, and may be generated from an initially ubiquitous pool whose maternally repressed transcripts are subsequently degraded. Gap-like expression domains are later observed in both the anterior (gnathal/thoracic) and posterior (abdominal) regions. Thus, the role of *hb* in anteroposterior patterning is apparently conserved within flies. Indeed, functional data for this has been collected from *M. abdita*, where RNA interference experiments yield a phenotype that resembles a hypomorphic *D. melanogaster hb* allele (Stauber et al., 2000).

hb orthologs have also been isolated from insects other than flies, and expression patterns have been described in lepidopterans (*Manduca sexta* and *Bombyx mori*) and the beetle, *Tribolium castaneum* (Kraft and Jäckle, 1994; Wolff et al., 1995; Xu et al., 1997). In all three cases, *hb* is expressed in the anterior region of the embryo. In the case of *Tribolium*, this expression domain is thought to be equivalent to the anterior gap domain of *Drosophila hb*. Also conserved is the abdominal expression domain (in the PS13/14 region) of *hb*.

In all insect species examined, neural expression of *hb* is conserved, suggesting that a neural function is ancestral. However, as the expression of the *eve* and *ftz* genes during segmentation is not conserved between grasshopper and *Drosophila*, and these genes lie below gap genes such as *hb* in the *Drosophila* segmentation hierarchy, it was unclear whether the role of *hb* in AP patterning would be conserved in more basal insects.

In order to answer this question, we have isolated full-length orthologs from two grasshopper species, *Locusta migratoria* and *Schistocerca americana*, the most basal insect *hb* homologs yet examined. Sequence analysis of these orthologs compared with *hb* genes from other phyla enables us to propose a novel ancestral structure for the Hb protein. Additionally, we have examined the expression pattern of *hb* in the grasshoppers and find that it is expressed both maternally and zygotically during early development. As in other insects, the anterior domain (gnathal/thoracic), abdominal domain and nervous system expression patterns are conserved in the grasshopper. In addition, as we detect at least two different levels of *hb* expression in the grasshopper gnathal/thoracic segments, we propose that the morphogenetic role of *hb* is conserved. We also find that grasshopper *hb* is present anteriorly but absent from the posterior growth zone at early stages, suggesting a conserved role in anterior patterning, despite the very short germ mode of embryogenesis of grasshoppers. Furthermore, we define *hb* expression domains in the grasshopper mesoderm and serosa. Examination of the *Drosophila hb* expression pattern shows that it is expressed extra-embryonically, suggesting that the gene may have played an ancestral role in serosal development. Finally, we note that maternally supplied Hb protein is spatially restricted within the newly laid *Schistocerca* egg, and propose that this initial expression domain distinguishes between the cells that will contribute to the embryo as opposed to extra-embryonic tissue.

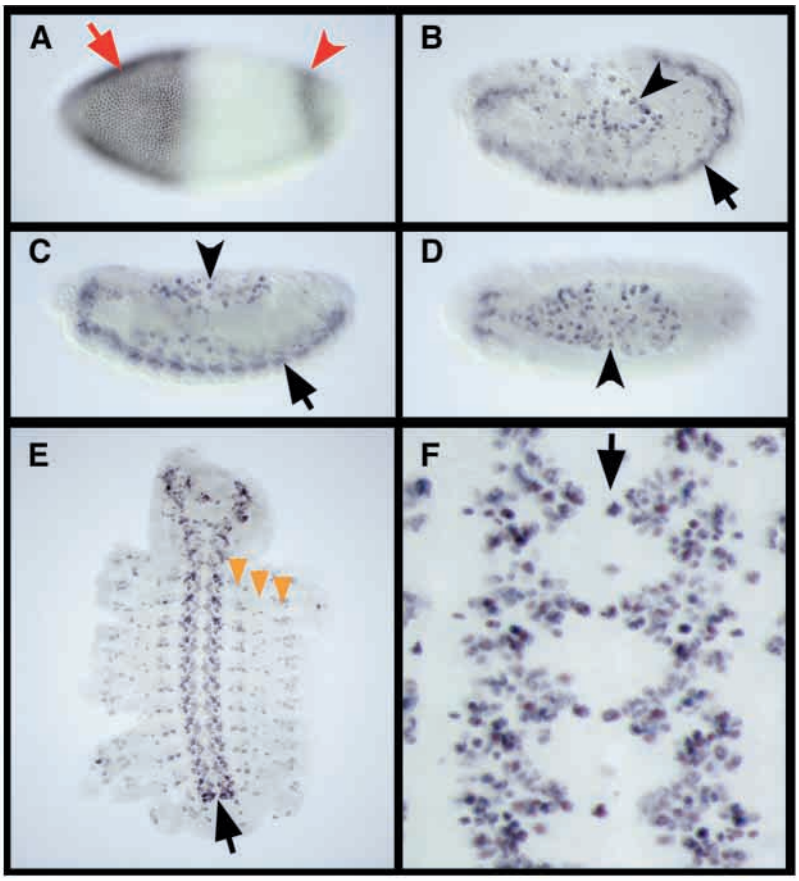

Fig. 1. Expression of *Drosophila hb* during embryogenesis. Immunostaining reveals the dynamic expression of Hb during *D. melanogaster* development, and provides a basis for interpretation of the grasshopper expression pattern. (A) At blastoderm stage, Hb is found in a large anterior domain (arrow) and a posterior domain around PS 13/14 (arrowhead). In a stage 12 germ band extended embryo (B), Hb is found in multiple tissues including the nervous system (arrow). In addition, it is expressed in extra-embryonic tissue and in nuclei around the yolk (arrowhead). Both domains are also seen in stage 13 germ band retracted embryos (C,D; arrow is neural expression, arrowhead is extra-embryonic expression), with the extra-embryonic domain shown in both a lateral (C) and dorsal view (D). *Drosophila hb* expression in a dissected stage 14 embryo (E) reveals a 'bow-tie'-like pattern in each neuromere (arrow points to CNS) as well as in lateral domains (orange arrowheads) of mesodermal and possibly PNS expression. (F) Close-up view showing *hb* expression in the stage 15 CNS (arrow indicates the midline).

MATERIALS AND METHODS

Embryo collections

Grasshopper eggs were collected from colonies maintained in Berkeley and Chicago, USA (*Schistocerca americana*) and Canberra, Australia (*Locusta migratoria*). To stage *S. americana* embryos from 0 to 10%, timed eggs were collected by taking newly laid eggpods and allowing them to develop at 32°C in a moist environment. At this temperature, *S. americana* embryos take approximately 20 days to hatch, and thus each day represents 5% of development. Embryos are described either by time after egg lay (0-10%=0-48 hours AEL) or staged (10-40%) using morphological criteria or staining for engrailed (Bentley et al., 1979; Patel et al., 1989a). We also used a set of staging criteria developed by Peter Dearden and Michael Akam for *Schistocerca gregaria*, to help set our own 0-10% timeline (*S. gregaria* and *S. americana* are not identical morphologically, but similar enough to make comparisons reasonable). The development of *L. migratoria* occurs at a different rate, so these embryos are staged by morphological comparison with *S. americana*.

Cloning of *hb*

The primers and conditions used for *hb* amplification are those described by Sommer and Tautz (Sommer and Tautz, 1992). mRNA was extracted from 30% *S. americana* embryos and used for first strand cDNA synthesis and then PCR. A multimerized copy of this PCR fragment was used to screen a *L. migratoria* genomic library, and a region of the genomic clone was used to isolate a cDNA from a *L. migratoria* ovarian cDNA library. Using the sequence of the original *S. americana* PCR fragment, non-degenerate primers were designed and used for 5′ and 3′ RACE

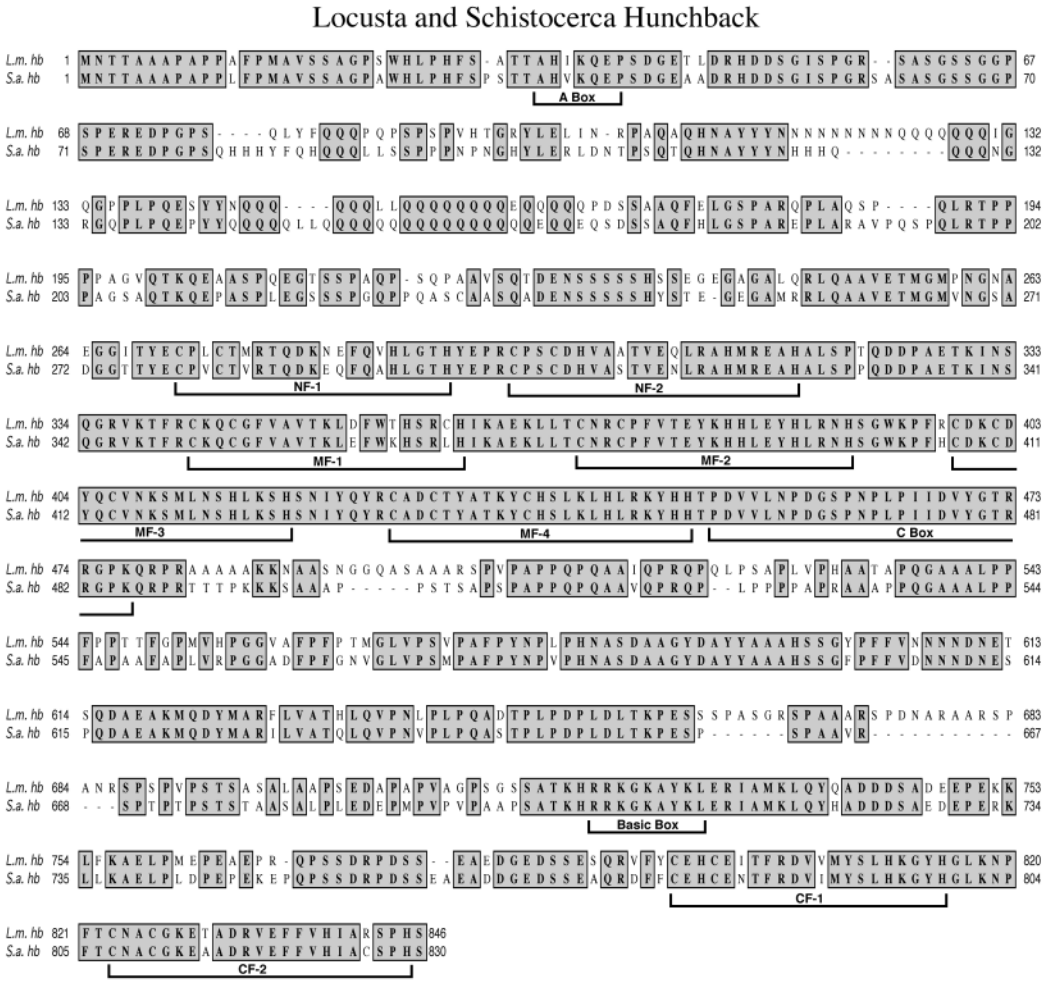

Fig. 2. Alignment of the *L. migratoria* and *S. americana hb* sequences. Clustal alignment of the predicted full-length amino acid sequences of the two grasshopper Hb proteins reveals 76% identity. Labeled brackets indicate conserved domains, including two zinc fingers towards the N terminus of the proteins (NF-1 and 2), four central zinc fingers (MF-1 to 4), two C terminal fingers (CF-1 and 2), and the conserved A, C and basic boxes. *L.m.*, *Locusta migratoria*, *S.a.*, *Schistocerca americana*.

(Gibco BRL) to obtain further cDNA sequence for *S. americana*. Genomic clones for *S. americana* were isolated using the GenomeWalker system (Clontech).

For *Schistocerca*, we have isolated cDNA sequence beginning 199 nucleotides 5′ of the start methionine and extending to a polyA tail 1243 nucleotides 3′ of the stop codon. Our genomic sequence extends 2740 nucleotides 3′ of the stop codon and contains all the cDNA sequence as well as complete introns with the exception of one intron between the N-2 and M-1 zinc fingers, which was hopped over using primers based on the cDNA sequence. The *Locusta* cDNA sequence begins 152 nucleotides 5′ of the start methionine and extends to a polyA tail 1215 nucleotides 3′ of the stop codon. The *Locusta* genomic sequence extends from a *Hin*dIII site in the last intron to a position 2529 nucleotides 3′ of the stop codon. The *S. americana* and *L. migratoria* cDNA and genomic sequences are deposited in GenBank (Accession Numbers, AY040605, AY040606 and AY040607, AY040537, AY040538).

In situ hybridization

For *S. americana* embryos, whole-mount in situ hybridization using a digoxigenin-labeled RNA probe corresponding to the *hb* 3′UTR was performed as previously described (Patel, 1996) with the following changes. Embryos of *S. americana* were fixed for 15-20 minutes in 3.7% formaldehyde after dissection in 1× PBS (pH=7.4) and were then dehydrated stepwise in methanol. Fixed embryos were not treated with xylene or proteinase K, and the hybridization was carried out overnight at 60°C in sodium dodecyl sulfate (SDS) hybridization solution (SDS-Hyb: 50% formamide, 5× saline sodium citrate (SSC) (pH 4.5), 0.1% Tween-20, 0.3% SDS, 50.0 μg/ml heparin and 100.0 μg/ml sonicated salmon sperm DNA). The embryos were then washed at 60°C in the following solutions of SDS-Hyb/ PTw (1× phosphate-buffered saline (PBS) containing 0.1% Tween-20): 2×20 minutes (80% SDS-Hyb/20% PTw), 2×20 minutes and 2×1 hour (50% SDS-Hyb/50% PTw), 2×20 minutes (20% SDS-Hyb/80% PTw), then 2×20 minutes (100% PTw). Finally, embryos were washed 2×20 minutes in 1×PBS with 0.1% bovine serum albumin (BSA) and 0.1% Triton X-100 (PBT) before adding sheep anti-digoxigenin-AP Fab fragments (Roche), diluted 1:3000 in PBT, and incubating at 4°C overnight. The next day embryos were washed 4×20 minutes and 2×60 minutes in PBT before performing the color reaction.

For *L. migratoria* in situ hybridization, embryos were dissected in PBS, fixed in 3.7% formaldehyde (Sigma), held in PTw at 4°C and hybridized within 24 hours. The probe used corresponds to a 930 bp *Xho*I/*Sal*I fragment from the *L. migratoria* cDNA clone hydrolyzed to an average length of approximately 250 nucleotides. A sense probe

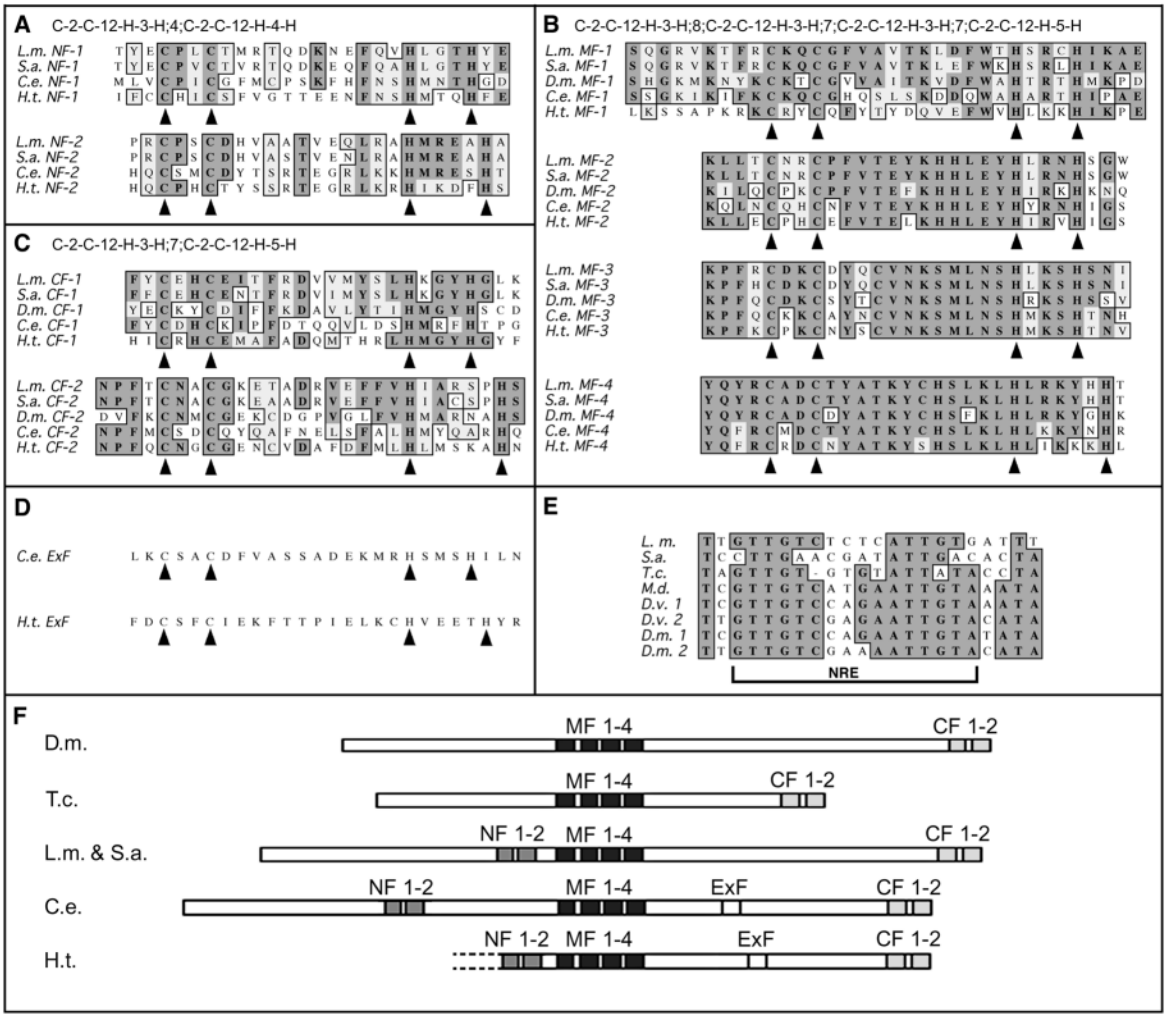

Fig. 3. Alignments of *hb* zinc fingers and putative NREs from different species. Alignment of orthologous zinc fingers from different species reveals that the various fingers are highly related across species in the position of structural residues. Alignments are shown of (A) the two NF fingers, (B) the four MF fingers and (C) the two CF metal-binding fingers. Black arrowheads indicate the structural residues of the putative metal binding fingers. The spacing between cysteine and histidine residues within and between each finger in each cluster (N, M and C) is shown at the top of each section of the figure. (D) Alignment of ExF, an additional putative zinc-finger identified in the *C. elegans* and *H. triserialis* Hb sequences. These fingers are in analogous positions in their respective proteins, but have little structural similarity to each other. (E) DNA nucleotide alignment of predicted nanos response element (NRE) sequences from the 3′UTR of *hb* transcripts of various species. *D.v.* and *D.m.* 1 and 2 indicate the two NRE sequences found in the 3′UTR of both of these *Drosophila* species. (F) Overall structure of Hb, with regard to zinc fingers. The putative *H. triserialis* Hb-coding sequence was reconstructed from the published genomic sequence (Savage and Shankland, 1996) by comparison with the grasshopper and *C. elegans hb* sequences, and manual assignment of splice junctions. This process yields a predicted leech Hb of similar structure to *C. elegans* Hb. As the overall structure illustrated for *L. migratoria* and *S. americana* is conserved across phyla, it may be the ancestral insect structure, with the NF fingers being lost in the lineage leading to *Tribolium* and *Drosophila*. Whether or not the ExF fingers are part of an ancestral protostome structure is less clear, as the leech and *C. elegans* ExF fingers lack noticeable similarity. Species used are *L.m.*, *L. migratoria*; *S.a.*, *S. americana*; *C.e.*, *C. elegans*; *H.t.*, *H. triserialis* (leech); *D.m.*, *D. melanogaster*; *T.c.*, *T. castaneum* (flour beetle); *M.d.*, *M. domestica* (house fly); and *D.v.*, *D. virilis*.

was used as a control. Techniques and solutions were as described in Kucharski et al. (Kucharski et al., 2000), except that hybridization was at 55°C.

Antibody production and staining

A TrpE/*L. migratoria* Hb fusion protein was constructed by cloning a fragment, beginning at the *Xho*I site within the M-1 finger and continuing through to the end of the coding region, into a pATH vector (Koerner et al., 1991). Rat antisera were produced, and affinity purified, as described previously (Patel et al., 1992). A mouse monoclonal antibody was produced by immunizing with the same TrpE/*L. migratoria* Hb fusion protein as above. ELISA-positive clones were then screened on *S. americana* embryos. Both the rat antisera and mouse monoclonal work equally well in *S. americana* and *L. migratoria*. We also produced a mouse monoclonal against *Drosophila* Hb by immunizing with a TrpE/*Drosophila* Hb fusion protein and screening clones on *Drosophila* embryos. Embryo staining followed the protocols outlined elsewhere (Patel, 1994). The monoclonal antibody to grasshopper Hb is designated PP 7C11, while that to *Drosophila* Hb is PP 1G10.

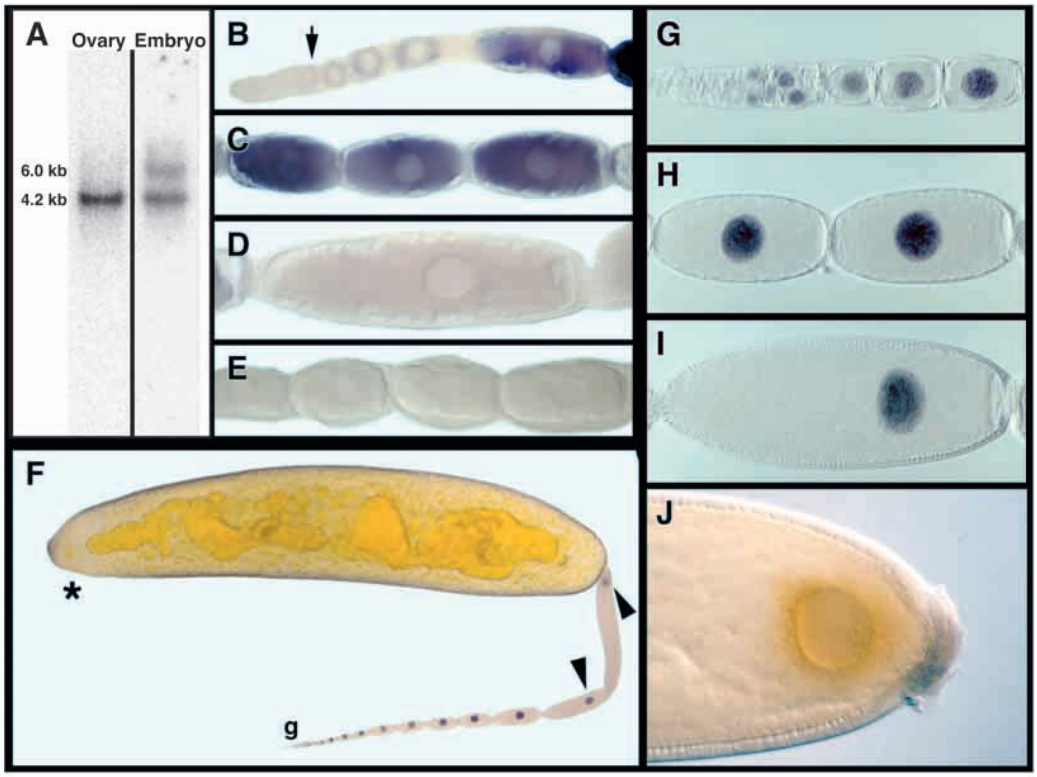

Fig. 4. Maternal *hb* expression in *S. americana*. Northern analysis, in situ hybridization and immunostaining reveal that *hb* is expressed during oogenesis. One transcript of 4.2 kb is expressed in the *L. migratoria* ovary (A), with an additional transcript of 6.0 kb appearing during embryogenesis. (B-E) In situ hybridization results using antisense (B-D) and sense (E) *hb* probes on *S. americana* ovarioles. Expression begins within the oocyte, as it leaves the germarium (B, arrow): there is no detectable expression in the germarium (left of arrow) or the somatic cells that surround the oocyte. By mid-oogenesis, *hb* transcript is abundant (C), but fades to background levels in older oocytes (D). (E) Control hybridization with a sense probe of oocytes at about the same stage as those seen in C. (F-J) Immunostaining of *S. americana* ovaries with monoclonal antibody PP 7C11. In F, the almost completed egg (*) is rotated relative to the rest of the ovariole and the presence of a vitelline membrane around it prevents the penetration of antibodies; (g) indicates the position of the germarium. Hb protein is expressed at high levels within developing oocytes and is localized to the germinal vesicle (oocyte nucleus). Translation begins around the time that transcript can first be detected (G), and continues through mid-oogenesis (H) until the germinal vesicle moves posteriorly in late oogenesis (arrowheads in F). As the nucleus moves posteriorly, Hb protein is still nuclear (I), but seems to exit the posteriorly located nucleus about the same time that the egg undergoes a phase of rapid expansion as it fills with yolk (J).

RESULTS

Grasshopper *hb* orthologs have a conserved structure

Using a degenerate PCR strategy, a small region of the *S. americana hb* sequence was isolated from 30% embryos. This *S. americana* fragment was then used to isolate genomic and cDNA clones of *hb* from two species of grasshoppers, *S. americana* and *L. migratoria* (see Materials and Methods). The predicted protein sequence encoded by both of these genes is shown in Fig. 2. The encoded proteins are approximately 76% identical at the amino acid level, and both contain eight zinc fingers (designated NF1-2, MF1-4 and CF1-2, see Fig. 2). The Hb proteins of *Caenorhabditis elegans* and *Helobdella triserialis* (leech) also contain zinc fingers that correspond to these eight (see Fig. 3A-C,F). Alignments of the grasshopper Hb zinc fingers with those of Hb proteins from *D. melanogaster*, *C. elegans* and *H. triserialis* are shown in Fig. 3.

The Hb protein of *D. melanogaster* contains six zinc fingers, four in the middle region of the protein and two in the C-terminal region (Fig. 3F). The four grasshopper Hb zinc fingers MF1-4 and CF1-2 correspond to the four central fingers and the two C-terminal fingers of *D. melanogaster*, respectively (see Fig. 3B, C and F). The two grasshopper fingers NF1 and NF2, however, do not have corresponding fingers in the *D. melanogaster* Hb protein, although these two fingers are present in *C. elegans* and *H. triserialis* Hb (see Fig. 3A,F). In addition, our analysis of the *hb* sequences of *C. elegans* and *H. triserialis* reveal that they also contain an additional zinc finger (ExF) located between the MF1-4 and CF1-2 (Fig. 3D,F). Beyond their relative position however, the *C. elegans* and *H. triserialis* ExFs bear little similarity to one another (see Fig. 3D). In addition to the zinc-finger regions, there are other conserved domains defined initially by comparative studies with other insect *hb* sequences, as well as functional analysis in *D. melanogaster* (Hülskamp et. al., 1994, Wolff at el., 1995). These domains are conserved in the grasshopper sequences and include the so called A box, Basic box and C box (see Fig. 2).

Another notably conserved region between grasshopper *hb* and its orthologs in other insects is a putative nanos response element (NRE) within the 3′ UTR (see Fig. 3E). *D. melanogaster* and *D. virilis hb* contain two NREs in their 3′ UTRs, while the 3′UTRs of *Musca domestica* (house fly) and *T. castaneum* (beetle) *hb* each contain a single NRE-like sequence. Like *Musca* and *Tribolium*, both *S. americana* and *L. migratoria* contain single NRE-like sequences of 540 and 460 nucleotides, respectively, after the stop codon (Fig. 3E).

Grasshopper *hb* is expressed during oogenesis

Northern analysis of *L. migratoria* tissue indicates that *hb* mRNA is expressed during both oogenesis and embryogenesis (Fig. 4A). As in *Drosophila*, a single transcript is detected during oogenesis, and a second transcript of different size is additionally expressed during embryogenesis (Fig. 4A). In *L.*

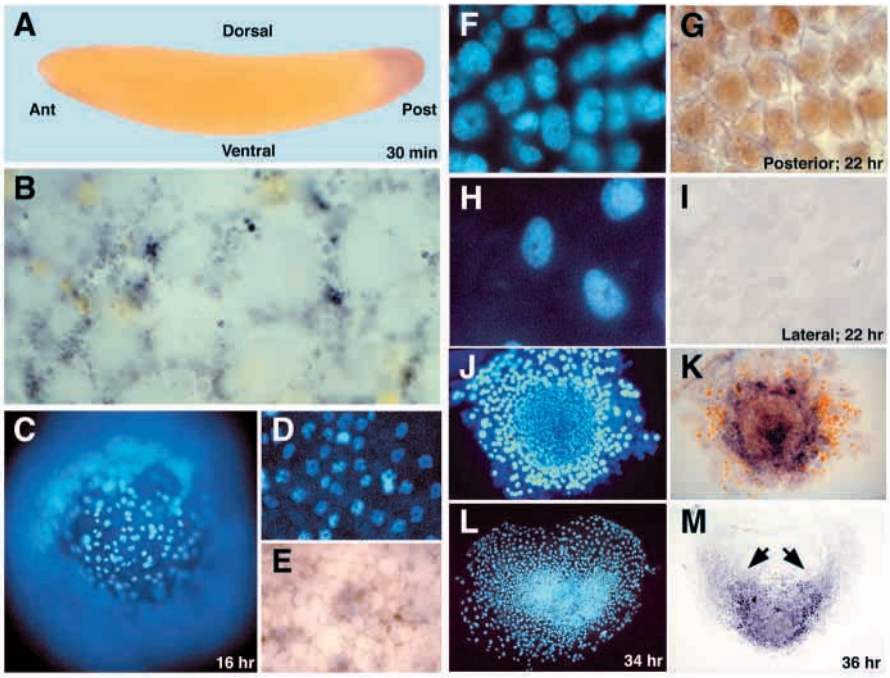

Fig. 5. *hb* expression during early embryogenesis. Nomarski views of immunostaining (A,B,E,G,I,M) compared with DAPI staining (blue, C,D,F,H,J,L) reveal *S. americana* (A-I,L,M) and *L. migratoria* (J,K) *hb* localization from egg lay to condensation of cells to form the blastodisc. After egg lay, Hb protein is found in the posterior 15% or so of the egg (A), in a granular network on the surface of the egg (B). As nuclei begin to clump at the posterior end of the egg (C,D), Hb protein is still evident cytoplasmically (E, same field as D), but then enters nuclei of the posterior (G) region of the egg (matching DAPI shown in F). Nuclei in more lateral regions of the same egg (outside of the posterior 15% region) do not contain Hb protein (I, matching DAPI in H). Grasshopper *hb* transcript is first detectable in the condensed blastodisc but not extra-embryonically (K, matching DAPI in J). At about 34 hours AEL, the disc starts to become asymmetric (L, this embryos is not stained for Hb), and a short while later when gastrulation has begun, all embryonic nuclei still express *hb* at some level. Expression is not seen in extra-embryonic nuclei (M, arrows point towards the head lobes of the embryo).

migratoria, the maternal transcript is 4.2 kb in length, and the zygotic transcripts are 4.2 kb and 6.0 kb in length. RT-PCR analysis reveals that *hb* is also transcribed both maternally and zygotically in *S. americana* (data not shown). The two *Drosophila hb* transcripts contain the same coding regions and 3′ UTRs, and differ only in their 5′ UTR regions. This is due to the sole use of the P1 promoter during oogenesis, and use of both P1 and P2 zygotically. Our analyses of cDNA clones and RT-PCR products also suggest that the grasshopper zygotic and maternal mRNAs do not differ in their coding regions or in their 3′ UTRs. This, therefore, resembles the situation in *Drosophila*, although we have yet to analyze maternal versus zygotic 5′ UTR regions to determine if the two different transcript sizes are due to differences in transcription start sites.

To study the expression pattern of *hb* during grasshopper development, we carried out whole-mount in situ hybridization and generated both polyclonal antisera and a monoclonal antibody for immunohistochemical localization. In our descriptions below, we will simply refer to the expression patterns seen as the ‘grasshopper pattern’, given that we obtained consistent results in both *L. migratoria* and *S. americana* from 12% of development onwards. The results presented for earlier stages are for *S. americana*.

Consistent with our northern hybridization and RT-PCR results, in situ analysis shows that *hb* mRNA accumulates in developing oocytes. Grasshopper *hb* transcript is detectable in oocytes as they leave the germarium (Fig. 4B). By mid-oogenesis *hb* mRNA levels have risen (Fig. 4C), but then drop and become undetectable during later oogenesis when the oocyte nucleus begins to move posteriorly (Fig. 4D). Grasshopper *hb* transcript is not detected in cells within the germarium or in any somatic cells of the ovary. One important difference between *Drosophila* and grasshoppers is that *Drosophila* possesses meroistic ovaries, in which each oocyte is part of a 16 cell cyst including 15 nurse cells (which transcribe maternal factors). Grasshoppers, however, have panoistic ovaries, in which the oocyte develops in the absence of nurse cells. This means that *hb* mRNA observed in developing grasshopper oocytes is likely to be transcribed by the oocyte nucleus.

Surprisingly, in addition to mRNA, Hb protein is also present in the oocytes of grasshoppers: this contrasts with *Drosophila* oocytes, which contain *hb* mRNA that is not translated until after fertilization (see Discussion). In grasshoppers, Hb protein is detected in oocytes shortly after they have exited the germarium, the protein being localized to the oocyte nucleus (the germinal vesicle, Fig. 4F,G). Initially the oocyte nucleus is positioned centrally within the developing oocyte, but as the oocyte matures and slowly enlarges, the germinal vesicle comes to lie at the extreme posterior end of the egg (Fig. 4F,H-J). At about this time the oocyte begins a phase of rapid enlargement and yolk accumulation, and Hb protein leaves the oocyte nucleus and enters the surrounding cytoplasm (Fig. 4J). Shortly afterwards, germinal vesicle breakdown occurs.

In freshly laid eggs (30 minutes after egg lay; AEL) we are unable to detect *hb* mRNA by in situ hybridization (data not shown). However, Hb protein (presumably maternally provided), is localized to small islands of cytoplasm in the posterior 15% of the egg length, with the boundary of the domain being slightly graded (Fig. 5A,B). At this time, embryonic nuclei (energids) are dividing and moving up towards the cortex of the egg, but none have yet reached the surface. During the next few hours, these energids reach the surface of the egg and rapidly cellularize as they do so (Fig. 5C-E; for cellularization timing in the related species *S. gregaria* see Ho et al. (Ho et al., 1997)). At this point, we observe that some of the Hb protein in the egg cytoplasm is now within the cytoplasm of newly formed cells (Fig. 5D,E). Consistent with maternal Hb protein localization to the

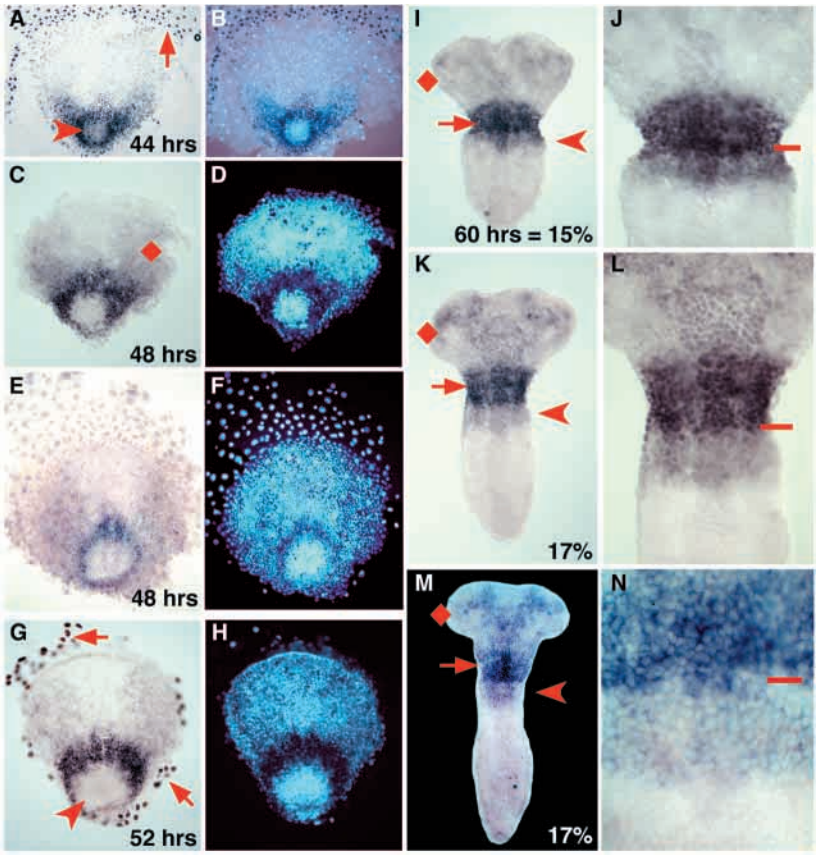

Fig. 6. *hb* expression in the developing germband. Immunostaining and paired DAPI images show the development of *S. americana hb* domains as the posterior growth zone forms and begins extension. (A,B) At 44 hours *hb* is expressed throughout the embryo, but starts to become less intense in more posterior regions of the embryo (arrowhead). In addition, strong Hb labeling is seen at this time in extra-embryonic nuclei of the serosa (arrow). By around 48 hours (C,D), Hb protein has disappeared from the posterior growth zone, but remains expressed at low levels in the headlobes (diamond). At this time the transcript pattern (E,F) matches that of the protein. (G,H) As posterior growth begins, *hb* is absent from the posterior growth zone (arrowhead), but is upregulated in an arc anterior to it. Strong expression continues in serosal nuclei (arrows). By about 15% (I, closer view in J), there are at least four levels of *hb* expression: weak in the headlobes (diamond), strong anteriorly (arrow), weaker posteriorly (arrowhead) in the gnathal/thoracic domain and absent in the more posterior part of the embryo. These expression domains are still evident at 17% (K-N), for both the protein (K, closer view in L) and the transcript (M, closer view in N) in *S. americana* (K,L) and *L. migratoria* (M,N). Red lines in J,L,N indicate the boundary between the strong and weak *hb* domains. Anterior is at the top in all cases.

posterior 15% or so of the egg, only those cells forming in this posterior region contain Hb protein. This maternally provided Hb protein appears to enter the nucleus, and at about the time that this occurs we begin to detect the first accumulation of zygotic *hb* mRNA in the cytoplasm of these same cells (data not shown). At ~22 hours AEL, the nuclei that are condensing at the posterior end of the egg to form the embryonic blastodisc contain Hb protein (Fig. 5F, G), while anteriorly located extra-embryonic nuclei do not (Fig. 5H,I). By 32 hours AEL, the cells at the posterior end have condensed, forming a defined circular blastodisc. At this stage, the egg is composed of two cell types: those that contain *hb* mRNA and protein, and are part of the blastodisc (embryonic), and those that are not part of the disc (extra-embryonic) and contain little or no *hb* mRNA or protein (mRNA pattern shown in Fig. 5J,K).

Expression of Hb in the developing blastodisc

In *Schistocerca*, the blastodisc at this point is oriented more or less perpendicular to the long axis of the egg. In *Locusta*, the disc is slightly offset down the side of the egg with the future embryonic tail pointing anteriorly with respect to the egg. By 34-36 hours AEL (Fig. 5L,M), the disc has become asymmetric, taking on a shallow 'V' shape that can be observed with DAPI staining (as cells are more densely packed embryonically than extra-embryonically), and the beginning of a gastrulation furrow is visible in the center of the embryo. The future head lobes are located at the top of the shallow 'V' and the future posterior region of the embryo at the base of the 'V'. At this stage, Hb protein is still expressed throughout the embryo (Fig. 5M).

Just after the embryo takes on this 'V' shape (~44 hours AEL), the pattern of *hb* expression undergoes two major alterations: expression of *hb* begins in extra-embryonic cells and expression is lost from within a circular domain at the posterior end of the embryo (Fig. 6A,B). At this stage, the extra-embryonic cells have a very flattened appearance and large nuclei relative to the embryonic cells, and express markers such as *zen*. The *zen* and later *hb*-positive extra-embryonic nuclei go on to form the serosa (Dearden et al., 2000). The amnion, an extra-embryonic membrane that comes to lie on top of the embryo, does not express *hb* when the tissue becomes morphologically obvious later in *Schistocerca* development (by the stage shown in Fig. 6I).

Interestingly, the loss of *hb* expression in the posterior circular domain is first seen at the level of protein expression (Fig. 6A), and at 44 hours AEL, *hb* mRNA is still seen throughout the embryo (data not shown). By about 48 hours, the discrepancy between the pattern of mRNA and protein disappears. Now both *hb* mRNA and protein are absent from the posterior region of the embryo (protein Fig. 6C,D; mRNA Fig. 6E,F), but present throughout the rest of the embryo and in the extra-embryonic cells surrounding the embryo. The *hb* expression domain is clearly modulated so that expression is relatively weak in the head lobes, but strong within an arc-shaped region just posterior to them (Fig. 6C-F). The posterior *hb*-negative region rapidly expands, apparently through a combination of cell proliferation and rearrangement. At ~52 hours, the posterior of the embryo starts to grow, and within the embryo, *hb* continues to be expressed weakly in the head lobes and strongly in an arc across the middle of the embryo

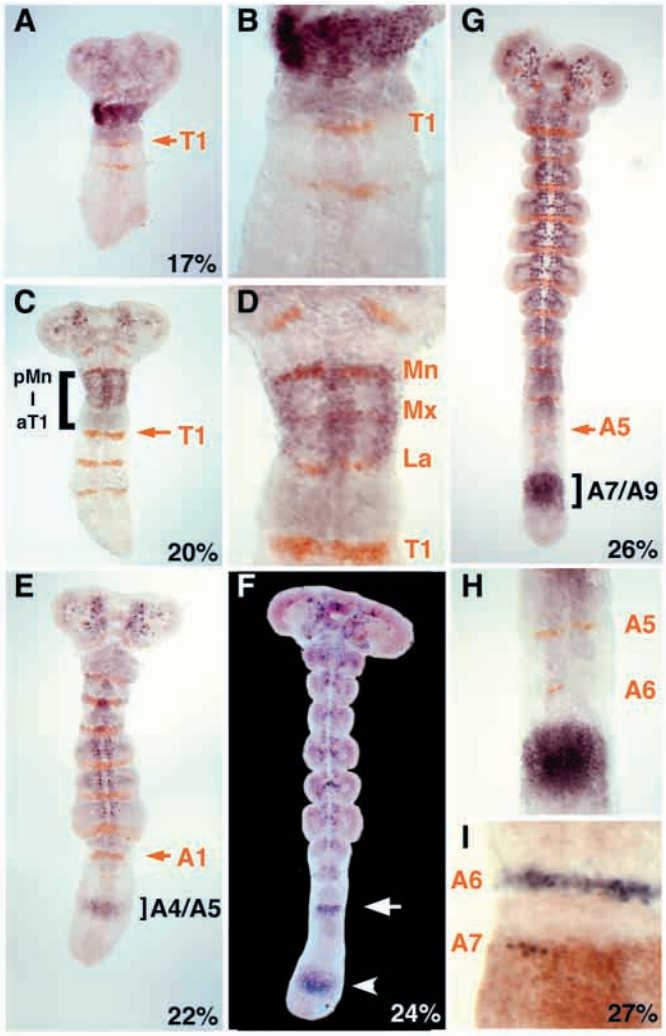

Fig. 7. *hb* expression during abdomen formation. Grasshopper Hb and engrailed double immunolabeling allow localization of *hb* domains during abdominal growth. At 17%, the posterior boundary of the *hb* weak domain (black) extends to the T1 engrailed stripe (brown, A, closer view in B). As the gnathal engrailed stripes appear (C, closer view in D), the anterior Hb domain (black, bracket) can be seen to extend from the posterior of Mn to anterior T1 (brown lettering indicates position of engrailed stripes for the various segments – mandibular (Mn), maxillary (Mx), labial (La) and the first thoracic segment (T1)). By 22%, a broad abdominal stripe appears (bracket in E). This is posterior to the A1 engrailed stripe, and is followed at 24% (F, in situ hybridization for *hb* in *L. migratoria*) with the appearance of a second abdominal domain (arrow indicates A4/A5 domain, arrowhead indicates the newly appearing A7-A9 domain). This second abdominal domain is also evident at the protein level at 26% (bracket in G, closer view in H), but the A4/A5 domain has faded away. The appearance of the engrailed A7 stripe (black, I) at 27%, allows placement of the anterior boundary of the second Hb domain at posterior A7. Immunostaining in *S. americana* (A-E,G-I) and in situ hybridization in *L. migratoria* (F), reveal similar expression domains.

(Fig. 6G,H). As the embryo grows (60 hours AEL=15% of development), this arc takes on the shape of a straight band of expression (Fig. 6I,J). This band is clearly not uniform, however, as there are two discrete sub-domains with different levels of Hb protein expression. This 'step' pattern at the posterior margin of the stripe is very noticeable at 17% of development (Fig. 6K,L). In situ analysis of *hb* mRNA shows that this step pattern of expression seen for the protein is also seen for *hb* mRNA (Fig. 6M,N).

The position of this *hb* domain was mapped at 17-20% of development by double labeling for both Hb and engrailed protein expression (Fig. 7A,B). At this stage, the engrailed stripes of T1 and T2 are just forming and the posterior boundary of *hb* expression is at the anterior edge of the T1 engrailed stripe. At 20% of development (Fig. 7C,D), the engrailed stripes of the gnathal segments have appeared, and reveal that the band of Hb protein extends from the posterior compartment of the mandibular (pMn) segment through to the anterior compartment of T1 (aT1), and thus spans parasegments 1-3. The more strongly stained subdomain spans pMn to pLa and the weaker subdomain is confined to the anterior compartment of T1 (Fig. 7C,D). This domain can be more precisely mapped using the pax3/7 homolog *pby1* (Davis et al., 2001).

Hb is found in two abdominal domains during the growth of the germ band

At 22% of development, a new domain of *hb* expression appears in the abdomen (Fig. 7E). This *hb* domain fades before engrailed stripes appear in this part of the abdomen, so we are not able to map its position precisely, but can place it roughly in the region of A4/A5. At about 24% of development, in situ hybridization shows that this A4/A5 *hb* domain is fading, and a new domain is appearing in a more posterior part of the abdomen, around A7/A9 (Fig. 7F). By 26% of development, the A4/A5 domain has disappeared, and the posterior (A7/A9) abdominal domain is intensifying (Fig. 7G). At this time, the A6 engrailed stripe is just appearing, and the second abdominal *hb* domain is clearly posterior to this. At 27% of development (Fig. 7I; engrailed in black and Hb in brown), the A7 engrailed stripe appears, and now we can see that this posterior *hb* domain begins at the anterior edge of this engrailed stripe, and thus has its anterior boundary in the posterior compartment of A7. This posterior *hb* expression domain fades before the next engrailed stripe forms, so we cannot precisely place its posterior boundary, but it is roughly within the anterior compartment of A9 (data not shown).

Grasshopper Hb is expressed in the mesoderm and nervous system

Several non-ectodermal patterns of *hb* expression are also visible beginning at about 20% of development. These include two phases of mesodermal expression. First, within the gap domains (Mn-T1, A4-A5 and A7-A9), *hb* is expressed in the underlying mesoderm at the same time that is expressed in the ectoderm. Second, *hb* is later expressed in the mesoderm of every segment at about the time that this germ layer condenses to take on a wedge shaped configuration within each segment. This occurs at roughly the same time that *engrailed* expression begins in the ectoderm of these segments (Figs 7G,H, 8E). Mesodermal condensation causes *hb* to appear to form a striped pattern with almost all of the mesodermal wedge located under the anterior compartment ectoderm (Fig. 8E). As the mesoderm migrates laterally, expression is eliminated from some

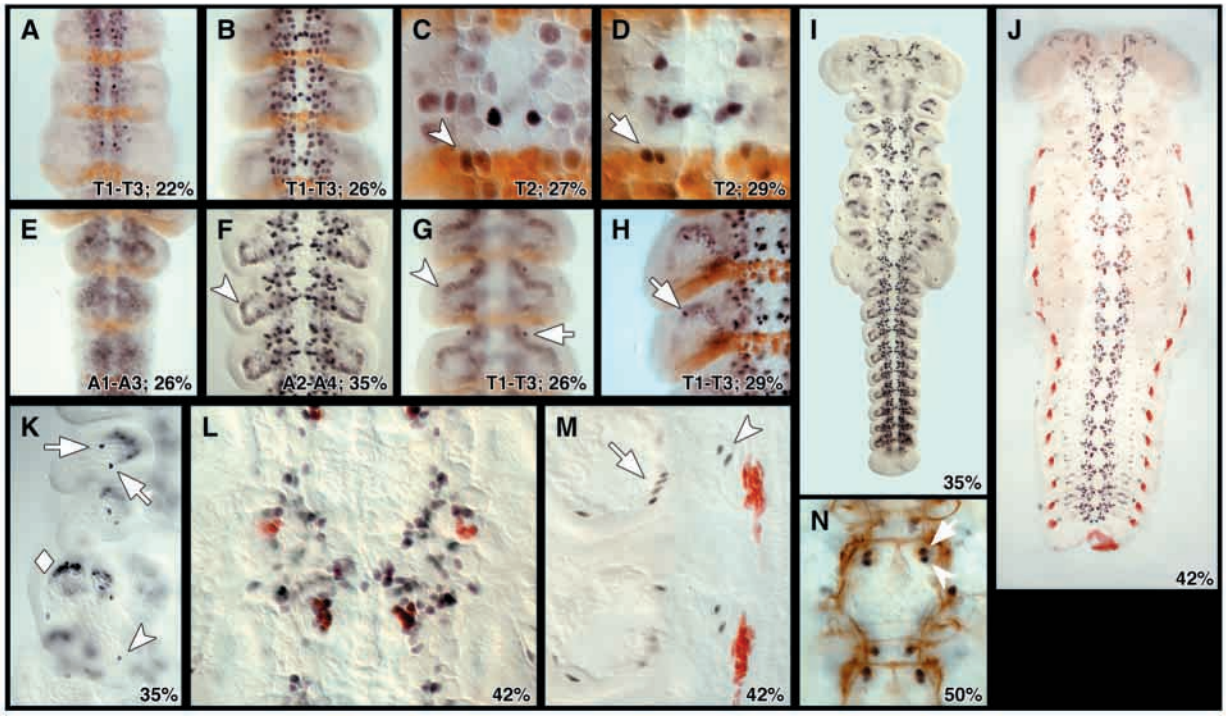

Fig. 8. Neural patterns of *hb* expression. Immunostaining for Hb (black) alone (in F, I, and K) or in combination with engrailed (brown in A-E, G, and H), or even-skipped (red in J, L, and M), or fasciclin II (brown in N). Grasshopper Hb protein is detected in all cells of the neuroepithelium before neuroblast delamination (A), but is then restricted to the newly delaminated neuroblasts (B). At 27%, NBs 6-1 and 6-2 (arrowhead in C) express both engrailed (brown) and Hb (black). At 29%, these same neuroblasts express only engrailed, but the initial GMCs that they have produced do express *hb* (arrow in D). Mesodermal expression in every segment is found as the mesoderm compacts and takes on a wedge shape (E). As the mesoderm moves out laterally, *hb* is expressed in the more distal cells in the abdomen (arrowhead in F), and in the more proximal cells at the limb bases of appendage-bearing segments (arrowhead in G). Expression is also seen in peripheral neurons of the leg (arrows in G,H). At 35% of development (I, closer view in K), expression can be seen in mesoderm at the base of the limbs (diamond), and in peripheral neurons of the gnathal segments (arrows in K indicate a few of these in the maxillary appendage) and several within each thoracic leg (arrowhead in K indicates the most distal Hb positive neuron in the T1 leg). By 42% (J, closer view in L, Hb in black, even-skipped in red), the pattern of Hb expression in the CNS resembles the pattern seen in *Drosophila* (see Fig. 1E,F). (M) A closer view of the lateral T1-T2 region from (J). Grasshopper Hb is expressed by several muscle fibers in the body wall (fibers 3 and 4, arrowhead) and at the base of the legs (arrow). The red staining in M shows eve expression in some pericardial tissue as well as muscle fibers 1 and 2. (N) At 50%, *hb* expression is still retained in some neurons, including aCC (arrow) and pCC (arrowhead).

mesodermal cells. In the abdomen, expression remains in the more distal cells within each body somite (Fig. 8F). In the gnathal and thoracic segments, expression is retained in the more proximal cells (Fig. 8G), which eventually form a ring of staining at the base of each appendage (Fig. 8H,I,K). At 42% of development, expression of *hb* can be seen in the nuclei of longitudinal muscles 3 and 4, and what appear to be muscles at the base of the limbs (Fig. 8J,M)

As in *Drosophila*, there is also dynamic expression of *hb* in the nervous system (Figs 1B,C,E,F, 8A-D). In grasshoppers, *hb* expression is seen throughout the neuroepithelium just before neuroblast delamination begins. Higher levels of Hb protein accumulate in the cells that will be the neuroblasts (Fig. 8A). Once delamination occurs, all neuroblasts appear to express *hb* at least transiently (Fig. 8B). As these neuroblasts begin to produce ganglion mother cell (GMC) progeny, expression is lost, but retained in the GMCs. This is clearly illustrated by following expression in neuroblasts 6-1 and 6-2. At 27% of development, NB 6-1 and 6-2 express both En and Hb proteins (Fig. 8C). By 29% of development, these neuroblasts express engrailed but not Hb, although the latter is detected in their initial GMC progeny (Fig. 8D). Thus, as in *Drosophila*, it appears that *hb* expression is transiently seen in all neuroblasts, but subsequently restricted to the first one or two of their GMC progeny. By 42% of development, Hb protein expression is seen in a number of neurons, with the overall 'bowtie-shaped' pattern of staining in each neuromere being quite similar to that seen in *Drosophila* at stage 15 (compare Fig. 8J,L with Fig. 1F). By 45% of development, Hb protein is found in a small set of neurons including aCC and pCC (Fig. 8N).

Expression of *hb* is also seen in a small number of peripheral neurons, most notably in specific neurons that form within the developing appendages. At 26% of development, the first *hb* positive neural precursors appear in the epithelium of the limbs (Fig. 8G). By 29%, these neurons have delaminated into the limb (Fig. 8H), and by 42% of development, several neurons are clearly visible within the gnathal and thoracic appendages.

DISCUSSION

In order to further understand the early patterning of

grasshopper embryos and make inferences about the ancestral role of *hunchback* (*hb*) in the insects, we have isolated and analyzed the expression pattern of the *S. americana* and *L. migratoria hb* orthologs. The comparison of the grasshopper *hb* sequences with previously isolated orthologs allows us to propose an ancestral structure for the Hb protein. Examination of the expression pattern of the grasshopper *hb* gene has raised questions about its regulation and function in these insects, as well as its ancestral role in arthropods.

Sequence alignments suggest a novel structure for an ancestral *hb* gene

Alignment of the Hb proteins highlights a highly stereotypical arrangement of putative zinc fingers that is apparently conserved across phyla (Fig. 3). Not only is the clustering pattern of the metal binding fingers conserved, but the spacing of structural residues within homologous fingers appears to be consistent across phyla (Fig. 3). From the evolutionary distribution of zinc fingers, we can infer that the common ancestor of the annelids, nematodes and arthropods possessed a *hb* gene encoding at least eight zinc fingers. Although the M fingers are known to be capable of binding DNA, the function of the N- and C-terminal zinc-finger domains remains unclear: they may, for example, be important for the recruitment of transcriptional co-factors. Whatever the case, the N terminal fingers are not found in the fly and the flour beetle (Fig. 3F), and may thus play some ancestral role that has been lost in the lineage leading to *Drosophila* and *Tribolium*. Other regions of conservation such as the A, B and basic domain are present in both *Drosophila* and grasshoppers (Fig. 2), and may play a conserved role in interactions with co-factors such as dMi-2 (Kehle et al., 1998).

Further evidence for a conserved role for the *hb* gene comes from the neural expression pattern in grasshopper. Grasshopper *hb* is transiently expressed in neuroblasts, and in a specific subset of GMCs and neurons, just as it is in *Drosophila* (Goodman and Doe, 1993; Kambadur et al., 1998; Tautz et al., 1987). The conservation of multiple protein domains across species, together with the highly conserved neural expression domains, shows that we have isolated true *hb* orthologs from *L. migratoria* and *S. americana*.

Although we have not identified a clear vertebrate *hb* ortholog in the databases, the family of Ikaros/Helios zinc-finger proteins have a similar structure to *hb* (Georgopoulos et al., 1992). In particular, they contain clusters of zinc fingers analogous to the M and C fingers in the grasshopper Hb protein. Interestingly, Ikaros feeds into gene regulation through chromatin remodeling complexes and interacts with the mouse ortholog of dMi-2 (Kim et al., 1999; Sun et al., 1996). This is reminiscent of the interaction of *Drosophila* Hb protein with dMi-2 and subsequent maintenance of Hox gene repression through the action of the polycomb group genes.

Regulation of *hb* function may occur by conserved and divergent mechanisms

The expression pattern of grasshopper *hb* allows us to make inferences about potential functions of the gene, as well as the transcriptional and post-transcriptional regulation of *hb* expression, both maternally and zygotically.

In grasshoppers, maternal *hb* is translated during oogenesis, and the protein initially accumulates in the oocyte nucleus (Fig. 4F-I). The nucleus is centrally placed during early stages of oogenesis, but moves posteriorly later in oocyte development (Fig. 4F). The Hb protein then appears to leave the nucleus, but remains in the posterior 15% of the egg (Figs 4J, 5A). This expression pattern suggests two novelties in our understanding of the regulation of grasshopper maternal *hb*. First, in contrast to *Drosophila* (see Tautz, 1988), *hb* transcript is translated during grasshopper oogenesis and maternally provided in the form of protein. This indicates that maternal *Drosophila hb* may be regulated by a translational control mechanism, so that the protein is not produced during oogenesis or very early embryogenesis (Wreden et al., 1997), and that Hb protein might play a novel function during grasshopper oogenesis or early embryogenesis. Second, spatial localization of maternal grasshopper Hb protein may be mediated by the oocyte nucleus.

This pattern of maternal Hb protein expression is particularly interesting because the position of the oocyte nucleus in short germ insects often correlates with the position where the germ anlagen (embryonic blastodisc) forms. Furthermore, irradiation of very early embryonic nuclei at what would normally be the position of the germ anlagen leads to the destruction of these nuclei. Adjacent nuclei, which would normally have taken on an extra-embryonic fate, enter the gap and go on to form a normal embryonic primordium (Patel, 2000). Such experiments imply that a non-nucleic acid factor is localized at the site occupied by the oocyte nucleus during oogenesis and that this factor plays an instructive role in determining the position at which the embryonic primordium will form. We propose that this factor may be maternally provided Hb protein. In this model, the position of the grasshopper oocyte nucleus at the end of oogenesis would specify the embryonic primordium by leaving Hb protein at a specific location. This maternally provided Hb protein is subsequently incorporated into cells that form in the posterior 15% of the egg. Relevant to this discussion is the observation that maternal grasshopper zen protein is initially found in the embryonic primordium, but becomes restricted to extra-embryonic cells (Dearden et al., 2000). Thus, early nuclei that express both zen and Hb proteins may be fated as embryonic, while those that express only zen may be fated to become extra-embryonic.

From about 44 to 96 hours AEL, *hb* expression in the early grasshopper embryo resembles the early fly pattern in that it is found throughout the anterior (head and anterior thorax) part of the embryo. Grasshopper *hb* is not expressed in the region of the embryo that will eventually contribute to the posterior thorax and abdomen. This suggests a conserved role for *hb* in setting up the initial embryonic axis. However, while grasshopper eggs inherit maternal Hb protein, we have been unable to detect maternal transcript in newly laid eggs. The *hb* expression seen during this 44-96 hour period would seem to be mainly derived from zygotically transcribed *hb*. This suggests that maternally supplied *hb* may not play a role in AP patterning, and that this role is played by zygotic *hb*. This contrasts with *Drosophila* where both maternal and zygotic *hb* participate in axis formation. In grasshoppers then, the importance of maternal *hb* may be to help distinguish between embryonic and extra-embryonic fate.

A zygotic patterning bias during axis formation outside of the Dipterans is also suggested by genetic evidence from

Hymenoptera. The existence of a protein carrying out a *hb* like role in A/P patterning of the wasp *Nasonia vitripennis* has been suggested by the close resemblance of the *headless* phenotype to that of *Drosophila hb* (Pultz et al., 1999). The *headless* phenotype, though zygotic, is stronger than that of the zygotic *Drosophila hb* mutant. As with our evidence from grasshopper, this suggests that although maternal AP patterning mechanisms exist in multiple insect orders, flies show an extreme bias towards maternal patterning.

At about 44 hours AEL (Fig. 5M), presence of *hb* transcript but lack of protein in cells of the posterior growth zone, implies translational control. The downregulation of *hb* in the posterior embryonic primordium may possibly be instigated by a factor such as nanos, as in *Drosophila*. Indeed, the sequences of the grasshopper *hb* 3′UTRs contain a putative nanos response element (NRE), the sequence required to recruit translational control proteins to the *Drosophila hb* transcript (Fig. 3E; Wharton and Struhl, 1991). By 17%, regulation probably occurs at the transcriptional level. Such regulation may either involve repression by a protein such as Krüppel, or the decreased concentration of a transcriptional activator (just as bicoid concentration regulates *Drosophila hb* transcription; Driever and Nüsslein-Volhard, 1989; Jäckle et al., 1986; Struhl et al., 1989). The posterior boundary of the weak Hb subdomain in T1 would not overlap with *Krüppel*, if the location of the latter matches its exact position in flies (PS4 to PS7; Harding and Levine, 1988; Knipple et al., 1985). Thus, it is possible that the anterior *Krüppel* boundary is set via repression by grasshopper *hb*. Alternatively, *Krüppel* may overlap with *hb* and be responsible for the weak *hb* subdomain in T1 region (see above). These scenarios are reminiscent of the cross-regulatory interactions between *Krüppel* and *hb* in *Drosophila*. Cloning of grasshopper *Krüppel* and analysis of its expression, as well as analysis of the *hb* promoter, will shed light on the transcriptional control of *hb*.

Expression patterns of grasshopper *hb* suggest a broadly conserved role in segmentation with some variation

The presence and positioning of the gnathal/thoracic domain of grasshopper *hb* suggests that the functions of this domain, which in *Drosophila* include control of gap, pair-rule and Hox genes, may be conserved in grasshopper. Anterior *hb* expression spans the gnathal segments, and seems well positioned for the control of possible head gap genes. The *hb* regulation of pair-rule stripes may also be conserved in the grasshopper although *eve*, for example, is not expressed in a pair-rule pattern in the grasshopper but rather in a posterior domain (Patel et al., 1992). This fact does not preclude *hb* regulation of the anterior *eve* border early in development, allowing for possible conservation of regulatory relationships in the face of clearly divergent spatial pattern. Grasshopper *pby1* (Davis et al., 2001), a *Pax 3/7* family member related to the *Drosophila* pair-rule gene *paired*, is expressed in a pair-rule pattern that may be regulated by *hb* expression, thus providing a pair-rule link between the gap patterning of *hb* and the segmental patterns of *engrailed* and *wingless*.

As in *Drosophila*, grasshopper *hb* may also be involved in Hox gene regulation, for example, by acting as a repressor to prevent anterior *Ubx* expression. However, the posterior boundary of grasshopper *hb* is shifted one segment anteriorly relative to *Drosophila*, while Ubx protein is expressed with the same anterior border in both insects (Kelsh et al., 1994). This highlights the absence of PS4 *hb* expression in grasshopper, and suggests the possibility of an additional level of *Ubx* regulation in this insect. It is also important to note that in *Tribolium*, *Ubx* transcripts appear more anteriorly than Ubx protein (Bennett et al., 1999). If this is also the case in the grasshopper, then the direct transcriptional control of *Ubx* by Hb protein may be conserved across insects, but with subsequent translational control setting the Ubx protein boundary in some species.

What were the ancestral roles of *hb* in development?

Some domains of grasshopper *hb* expression, such as the gnathal/thoracic, A7-A9 gap domains and CNS patterns, are presumably homologous to those observed in *Drosophila*, and thus probably ancestral. Other domains such as the A4-5 gap domain, and serosal, PNS and mesodermal expression domains are less obviously so.

The expression of *hb* in extra-embryonic primordia seems conserved in many species outside the higher flies, and it has even been suggested that this domain has been lost in *Drosophila* as a result of fusion and reduction of the amnion and serosa (Rohr et al., 1999). The reduction of extra-embryonic material and its movement into the DV axis of the egg maps phylogenetically to higher dipterans (the cyclorhapphan flies), and has been proposed to reflect profound changes in the specification of anterior tissue (Schmidt-Ott, 2000; Stauber et al., 1999). The cumulative data from insects suggests that serosal *hb* expression is ancestral, especially as there is extra-embryonic expression in *D. melanogaster* (Fig. 1C,D). Thus, it is formally possible that *hb* is involved in serosal patterning, whichever body axis this tissue or its evolutionary derivatives lie in.

The grasshopper *hb* mesodermal stripes at first sight seem unlike any domains observed in other species. The striped appearance of mesodermal *hb* is actually due to compaction of this tissue during morphological segmentation. Segmental stripes have also been observed in the trunk of *Musca domestica* and *Tribolium castaneum* (Sommer and Tautz, 1991; Wolff et al., 1995). If *hb* stripes in these insects are also found mesodermally, then mesodermal *hb* expression may be ancestral, the striped appearance being simply due to mesodermal condensation. Such early mesodermal expression has not been seen in *Drosophila*, but *hb* has recently been shown to be expressed later in specific mesodermal cells that play a role in tracheal guidance in *D. melanogaster* (Wolf and Schuh, 2000). Thus, while early *hb* expression in the mesoderm may differ between grasshoppers and flies, it may have a conserved expression in mesoderm later in development, and play a role in events such as tracheal guidance.

We have found that expression of *hb* in varying concentrations across the gnathal/T1 primordium is conserved through to basal insects such as the grasshopper. This suggests that *hb* may behave morphogenetically, despite lack of a prolonged syncytial stage in the grasshopper. Grasshopper *hb* does not, however, form a smooth gradient, as might arise by protein diffusion, but instead is expressed in a distinct stepped pattern that is seen at both the protein and mRNA level. This implies that morphogenetic activity may be set up or maintained by a fundamentally different mechanism from the

bicoid diffusion found in *Drosophila*. Orthologs of *hb* have been cloned from annelids, nematodes and a mollusk (Fay et al., 1999; Savage and Shankland, 1996; Sommer et al., 1992), and expression patterns have been described in *C. elegans* and the leech *H. triserialis* (Fay et al., 1999; Iwasa et al., 2000; Savage and Shankland, 1996). In neither organism is an early A/P gradient observed (although the *C. elegans* data are from a GFP construct and not from observation of endogenous mRNA or protein; Fay et al., 1999), suggesting that *hb* is not required for anteroposterior patterning in *C. elegans* and *H. triserialis*. These data suggest that *hb* may have been co-opted into a role in anterior patterning within the arthropods, or that such a role has been lost in the lineages leading to *C. elegans* and *H. triserialis*.

Concluding remarks

Cumulative evidence from insects seems to indicate that the role of *hb* as a gap gene is highly conserved. Thus, despite the fact that the expression of pair-rule orthologs may vary across insects, expression of an earlier component of the segmentation hierarchy is conserved. However, it is less certain that maternal *hb* is vital to patterning of the grasshopper anteroposterior axis. Instead, this role may be mediated by the regulation of zygotic *hb*, suggesting a bias toward zygotic axial patterning in more basal insects. Nevertheless, the maternal Hb protein product in grasshoppers may play a role in determining embryonic from extra-embryonic cells within the early egg.

We thank David Bentley for access to the Berkeley grasshopper colony, Bridget Lear for her unwavering commitment to keeping our Chicago grasshopper colony flourishing, and Greg Davis and Bridget Lear for helpful discussions and critical reading of the manuscript. We also thank Peter Dearden and Michael Akam for communication of unpublished results and for providing their timeline of early embryonic development in *S. gregaria*, Corey Goodman, David Bentley and Joyce West for help and support in the early stages of this project, Danny Brower for playing devil's advocate on *Locusta* development, and Kirsten Balding for help with sequencing. N. H. P. is an HHMI investigator, and S. L. is supported by an EMBO fellowship.

REFERENCES

Bennett, R. L., Brown, S. J. and Denell, R. E. (1999). Molecular and genetic analysis of the *Tribolium Ultrabithorax* orthologue, *Ultrathorax*. *Dev. Genes Evol.* **209**, 608-619.

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

Brown, S. J., Hilgenfeld, R. B. and Denell, R. E. (1994a). The beetle *Tribolium castaneum* has a *fushi tarazu* homolog expressed in stripes during segmentation. *Proc. Natl. Acad. Sci. USA* **91**, 12922-12926.

Brown, S. J., Patel, N. H. and Denell, R. E. (1994b). Embryonic expression of the single *Tribolium engrailed* homolog. *Dev. Genet.* **15**, 7-18.

Davis, G. K., Jaramillo, C. A. and Patel, N. H. (2001). Pax group III genes and the evolution of pair rule pattern. *Development* **128**, 3445-3458.

Dawes, R., Dawson, I., Falciani, F., Tear, G. and Akam, M. (1994). *Dax*, a locust Hox gene related to *fushi-tarazu* but showing no pair-rule expression. *Development* **120**, 1561-1572.

Dearden, P. and Akam, M. (2001). Early embryo patterning in the grasshopper, Schistocerca gregaria: wingless, dpp and caudal expression. *Development* **128**, 3435-3444.

Dearden, P., Grbic, M., Falciani, F. and Akam, M. (2000). Maternal expression and early zygotic regulation of the *Hox3/zen* gene in the grasshopper *Schistocerca gregaria*. *Evol. Dev.* **2**, 261-270.

Driever, W. and Nusslein-Volhard, C. (1989). The bicoid protein is a positive regulator of *hunchback* transcription in the early *Drosophila* embryo. *Nature* **337**, 138-143.

Fay, D. S., Stanley, H. M., Han, M. and Wood, W. B. (1999). A *Caenorhabditis elegans* homologue of *hunchback* is required for late stages of development but not early embryonic patterning. *Dev. Biol.* **205**, 240-253.

Georgopoulos, K., Moore, D. D. and Derfler, B. (1992). Ikaros, an early lymphoid-specific transcription factor and a putative mediator for T cell commitment. *Science* **258**, 808-812.

Goodman, C. S. and Doe, C. Q. (1993). Embryonic development of the *Drosophila* central nervous system. In *The Development of* Drosophila melanogaster. Vol. 2 (ed. M. Bate and A. Martinez Arias), pp. 1131-1206. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Harding, K. and Levine, M. (1988). Gap genes define the limits of Antennapedia and Bithorax gene expression during early development in *Drosophila*. *EMBO J.* **7**, 205-214.

Ho, K., Dunin-Borkowski, O. M. and Akam, M. (1997). Cellularization in locust embryos occurs before blastoderm formation. *Development* **124**, 2761-2768.

Hülskamp, M., Schröder, C., Pfeifle, C., Jäckle, H. and Tautz, D. (1989). Posterior segmentation of the *Drosophila* embryo in the absence of a maternal posterior organizer gene. *Nature* **338**, 629-632.

Hülskamp, M., Pfeifle, C. and Tautz, D. (1990). A morphogenetic gradient of hunchback protein organizes the expression of the gap genes Krüppel and knirps in the early *Drosophila* embryo. *Nature* **346**, 577-580.

Hülskamp, M., Lukowitz, W., Beermann, A., Glaser, G. and Tautz, D. (1994). Differential regulation of target genes by different alleles of the segmentation gene *hunchback* in *Drosophila*. *Genetics* **138**, 125-134.

Irish, V., Lehmann, R. and Akam, M. (1989). The *Drosophila* posterior-group gene nanos functions by repressing hunchback activity. *Nature* **338**, 646-648.

Iwasa, J. H., Suver, D. W. and Savage, R. M. (2000). The leech *hunchback* protein is expressed in the epithelium and CNS but not in the segmental precursor lineages. *Dev. Genes Evol.* **210**, 277-288.

Jäckle, H., Tautz, D., Schuh, R., Seifert, E. and Lehmann, R. (1986). Cross-regulatory interactions among the gap genes of *Drosophila*. *Nature* **324**, 668-670.

Kambadur, R., Koizumi, K., Stivers, C., Nagle, J., Poole, S. J. and Odenwald, W. F. (1998). Regulation of POU genes by castor and hunchback establishes layered compartments in the *Drosophila* CNS. *Genes Dev.* **12**, 246-260.

Kehle, J., Beuchle, D., Treuheit, S., Christen, B., Kennison, J. A., Bienz, M. and Muller, J. (1998). dMi-2, a hunchback-interacting protein that functions in polycomb repression. *Science* **282**, 1897-1900.

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

Kim, J., Sif, S., Jones, B., Jackson, A., Koipally, J., Heller, E., Winandy, S., Viel, A., Sawyer, A., Ikeda, T. et al. (1999). Ikaros DNA-binding proteins direct formation of chromatin remodeling complexes in lymphocytes. *Immunity* **10**, 345-355.

Knipple, D. C., Seifert, E., Rosenberg, U. B., Preiss, A. and Jäckle, H. (1985). Spatial and temporal patterns of Krüppel gene expression in early *Drosophila* embryos. *Nature* **317**, 40-44.

Koerner, T. J., Hill, J. E., Myers, A. M. and Tzagoloff, A. (1991). High-expression vectors with multiple cloning sites for construction of trpE fusion genes: pATH vectors. *Methods Enzymol.* **194**, 477-490.

Kraft, R. and Jäckle, H. (1994). *Drosophila* mode of metamerization in the embryogenesis of the lepidopteran insect *Manduca sexta*. *Proc. Natl. Acad. Sci. USA* **91**, 6634-6638.

Kucharski, R., Ball, E. E., Hayward, D. C. and Maleszka, R. (2000). Molecular cloning and expression analysis of a cDNA encoding a glutamate transporter in the honeybee brain. *Gene* **242**, 399-405.

Maderspacher, F., Bucher, G. and Klingler, M. (1998). Pair-rule and gap gene mutants in the flour beetle *Tribolium castaneum*. *Dev. Genes Evol.* **208**, 558-568.

Murata, Y. and Wharton, R. P. (1995). Binding of pumilio to maternal *hunchback* mRNA is required for posterior patterning in *Drosophila* embryos. *Cell* **80**, 747-756.

Nagy, L. M. and Carroll, S. (1994). Conservation of wingless patterning functions in the short-germ embryos of *Tribolium castaneum*. *Nature* **367**, 460-463.

Patel, N. H. (1994). Imaging neuronal subsets and other cell types in whole-mount *Drosophila* embryos and larvae using antibody probes. In *Methods in Cell Biology*. Vol. 44 (ed. L. S. B. Goldstein and E. A. Fyrberg), pp. 445-487. San Diego: Academic Press.

Patel, N. H. (1996). In situ hybridization to whole mount Drosophila embryos. In *A Laboratory Guide to RNA: Isolation, Analysis and Synthesis* (ed. P. A. Kreig), pp. 357-369. New York: Wiley, Liss.

Patel, N. H. (2000). It's a bug's life. *Proc. Natl. Acad. Sci. USA* **97**, 4442-4444.

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. (1989a). Expression of engrailed during segmentation in grasshopper and crayfish. *Development* **107**, 201-212.

Patel, N. H., Martin-Blanco, E., Coleman, K. G., Poole, S. J., Ellis, M. C., Kornberg, T. B. and Goodman, C. S. (1989b). Expression of engrailed proteins in arthropods, annelids, and chordates. *Cell* **58**, 955-968.

Pultz, M. A., Pitt, J. N. and Alto, N. M. (1999). Extensive zygotic control of the anteroposterior axis in the wasp *Nasonia vitripennis*. *Development* **126**, 701-710.

Rivera-Pomar, R. and Jäckle, H. (1996). From gradients to stripes in *Drosophila* embryogenesis: filling in the gaps. *Trends Genet.* **12**, 478-483.

Rohr, K. B., Tautz, D. and Sander, K. (1999). Segmentation gene expression in the mothmidge *Clogmia albipunctata* (Diptera, psychodidae) and other primitive dipterans. *Dev. Genes Evol.* **209**, 145-154.

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

Savage, R. M. and Shankland, M. (1996). Identification and characterization of a hunchback orthologue, Lzf2, and its expression during leech embryogenesis. *Dev. Biol.* **175**, 205-217.

Schmidt-Ott, U. (2000). The amnioserosa is an apomorphic character of cyclorrhaphan flies. *Dev. Genes Evol.* **210**, 373-376.

Schröder, C., Tautz, D., Seifert, E. and Jäckle, H. (1988). Differential regulation of the two transcripts from the *Drosophila* gap segmentation gene *hunchback*. *EMBO J.* **7**, 2881-2887.

Schröder, R., Jay, D. G. and Tautz, D. (1999). Elimination of EVE protein by CALI in the short germ band insect Tribolium suggests a conserved pair-rule function for even skipped. *Mech. Dev.* **80**, 191-195.

Simpson-Brose, M., Treisman, J. and Desplan, C. (1994). Synergy between the hunchback and bicoid morphogens is required for anterior patterning in Drosophila. *Cell* **78**, 855-865.

Sommer, R. and Tautz, D. (1991). Segmentation gene expression in the housefly Musca domestica. *Development* **113**, 419-430.

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.

Sommer, R. J., Retzlaff, M., Goerlich, K., Sander, K. and Tautz, D. (1992). Evolutionary conservation pattern of zinc-finger domains of *Drosophila* segmentation genes. *Proc. Natl. Acad. Sci. USA* **89**, 10782-10786.

Sonoda, J. and Wharton, R. P. (1999). Recruitment of Nanos to mRNA by Pumilio. *Genes Dev.* **13**, 2704-2712.

St Johnston, D. and Nüsslein-Volhard, C. (1992). The origin of pattern and polarity in the *Drosophila* embryo. *Cell* **68**, 201-219.

Stauber, M., Jäckle, H. and Schmidt-Ott, U. (1999). The anterior determinant *bicoid* of *Drosophila* is a derived Hox class 3 gene. *Proc. Natl. Acad. Sci. USA* **96**, 3786-3789.

Stauber, M., Taubert, H. and Schmidt-Ott, U. (2000). Function of *bicoid* and *hunchback* homologs in the basal cyclorrhaphan fly *Megaselia* (Phoridae). *Proc. Natl. Acad. Sci. USA* **97**, 10844-10849.

Struhl, G., Johnston, P. and Lawrence, P. A. (1992). Control of Drosophila body pattern by the hunchback morphogen gradient. *Cell* **69**, 237-249.

Struhl, G., Struhl, K. and Macdonald, P. M. (1989). The gradient morphogen bicoid is a concentration-dependent transcriptional activator. *Cell* **57**, 1259-1273.

Sun, L., Liu, A. and Georgopoulos, K. (1996). Zinc finger-mediated protein interactions modulate Ikaros activity, a molecular control of lymphocyte development. *EMBO J.* **15**, 5358-5369.

Tautz, D. (1988). Regulation of the *Drosophila* segmentation gene hunchback by two maternal morphogenetic centres. *Nature* **332**, 281-284.

Tautz, D., Lehmann, R., Schnürch, H., Schuh, R., Seifert, E., Kienlin, A., Jones, K. and Jäckle, H. (1987). Finger protein of novel structure encoded by *hunchback*, a second member of the gap class of *Drosophila* segmentation genes. *Nature* **327**, 383-389.

Treier, M., Pfeifle, C. and Tautz, D. (1989). Comparison of the gap segmentation gene *hunchback* between *Drosophila melanogaster* and *Drosophila virilis* reveals novel modes of evolutionary change. *EMBO J.* **8**, 1517-1525.

Wang, C. and Lehmann, R. (1991). Nanos is the localized posterior determinant in *Drosophila*. *Cell* **66**, 637-647.

Wharton, R. P. and Struhl, G. (1991). RNA regulatory elements mediate control of *Drosophila* body pattern by the posterior morphogen nanos. *Cell* **67**, 955-967.

Wimmer, E. A., Carleton, A., Harjes, P., Turner, T. and Desplan, C. (2000). Bicoid-independent formation of thoracic segments in *Drosophila*. *Science* **287**, 2476-2479.

Wolf, C. and Schuh, R. (2000). Single mesodermal cells guide outgrowth of ectodermal tubular structures in *Drosophila*. *Genes Dev* **14**, 2140-2145.

Wolff, C., Sommer, R., Schröder, R., Glaser, G. and Tautz, D. (1995). Conserved and divergent expression aspects of the *Drosophila* segmentation gene *hunchback* in the short germ band embryo of the flour beetle *Tribolium*. *Development* **121**, 4227-4236.

Wreden, C., Verrotti, A. C., Schisa, J. A., Lieberfarb, M. E. and Strickland, S. (1997). Nanos and pumilio establish embryonic polarity in *Drosophila* by promoting posterior deadenylation of hunchback mRNA. *Development* **124**, 3015-3023.

Xu, X., Xu, P. X., Amanai, K. and Suzuki, Y. (1997). Double-segment defining role of even-skipped homologs along the evolution of insect pattern formation. *Dev. Growth Differ.* **39**, 515-522.