Development 125, 3645-3654 (1998)
Printed in Great Britain © The Company of Biologists Limited 1998
DEV7666

Regulation of the *Tribolium* homologues of *caudal* and *hunchback* in *Drosophila*: evidence for maternal gradient systems in a short germ embryo

Christian Wolff[1], Reinhard Schröder[1], Cordula Schulz[2], Diethard Tautz[1] and Martin Klingler[1,*]

[1]Zoologisches Institut der Universität München, Luisenstrasse 14, 80333 München, Germany
[2]Stanford University, School of Medicine, Department of Developmental Biology, Stanford CA94305, USA
*Author for correspondence (e-mail: klingler@zi.biologie.uni-muenchen.de)

Accepted 6 July; published on WWW 25 August 1998

SUMMARY

In short germ embryos, the germ rudiment forms at the posterior ventral side of the egg, while the anterior-dorsal region becomes the extra-embryonic serosa. It is difficult to see how an anterior gradient like that of *bicoid* in *Drosophila* could in these embryos be directly involved in patterning of the germ rudiment. Moreover, since it has not yet been possible to recover a *bicoid* homologue from any species outside the diptera, it has been speculated that the anterior *bicoid* gradient could be a late addition during insect evolution. We addressed this question by analysing the regulation of potential target genes of *bicoid* in the short germ embryo of *Tribolium castaneum*. We demonstrate that homologues of *caudal* and *hunchback* from *Tribolium* are regulated by *Drosophila bicoid*. In *Drosophila*, maternal *caudal* RNA is translationally repressed by *bicoid*. We find that *Tribolium caudal* RNA is also translationally repressed by *bicoid*, when it is transferred into *Drosophila* embryos under a maternal promoter. This strongly suggests that a functional *bicoid* homologue must exist in *Tribolium*. The second target gene, *hunchback*, is transcriptionally activated by *bicoid* in *Drosophila*. Transfer of the regulatory region of *Tribolium hunchback* into *Drosophila* also results in regulation by early maternal factors, including *bicoid*, but in a pattern that is more reminiscent of *Tribolium hunchback* expression, namely in two early blastoderm domains. Using enhancer mapping constructs and footprinting, we show that *caudal* activates the posterior of these domains via a specific promoter. Our experiments suggest that a major event in the evolutionary transition from short to long germ embryogenesis was the switch from activation of the *hunchback* gap domain by *caudal* to direct activation by *bicoid*. This regulatory switch can explain how this domain shifted from a posterior location in short germ embryos to its anterior position in long germ insects, and it also suggest how an anterior gradient can pattern the germ rudiment in short germ embryos, i.e. by regulating the expression of *caudal*.

Key words: *Tribolium castaneum*, *Drosophila melanogaster*, Evolution of development, Short germ embryogenesis, Maternal genes, Blastoderm fate map, *bicoid*, *caudal*, *hunchback*

INTRODUCTION

Pattern formation in the early *Drosophila* embryo is initiated by maternally provided gradient systems (St Johnston and Nüsslein-Volhard, 1992). One essential component of anterior-posterior axis specification is the interaction between *bicoid* (*bcd*) and *caudal* (*cad*). *bcd* RNA becomes localized during oogenesis at the anterior pole of the embryo. Protein translated from this localized RNA diffuses and forms an anterior-posterior gradient in the early embryo. The protein interacts with the homogeneously distributed *cad* RNA and represses its translation (Chan and Struhl, 1997; Driever and Nüsslein-Volhard, 1988; Dubnau and Struhl, 1996; Macdonald and Struhl, 1986; Mlodzik and Gehring, 1987; Rivera-Pomar et al., 1996). This translational regulation results in an opposing gradient for *cad* protein which has an essential function in activating the abdominal gap genes (Rivera-Pomar et al., 1995; Schulz and Tautz, 1995). The major role of *bcd* is, however, in the transcriptional activation of *hunchback* (*hb*) and other gap genes in a concentration-dependent manner. This ensures proper positioning of several *bcd*-dependent gap expression domains within the anterior half of the embryo (Driever and Nüsslein-Volhard, 1989; Struhl et al., 1989).

While the *bcd* gradient has served as a model system in understanding pattern formation in *Drosophila*, it is suspected not to have this role in more ancestral insects. The long-germ mode of development as found in *Drosophila* is probably an adaptation to its particularly rapid embryogenesis. The ancestral type of embryogenesis in insects and arthropods is the short germ type (Sander, 1983). In these embryos, the germ rudiment forms at the posterior ventral side of the egg. In extreme cases like the grasshopper, it may be restricted to only a few percent of the total egg length - which makes it difficult to imagine how an anteriorly localized *bcd* RNA could determine pattern formation at the posterior end of the egg. Moreover, classical experiments have only yielded evidence for a posteriorly localized organizing activity (Sander, 1976). Therefore, *bcd* could be considered a late addition during

insect evolution and its pivotal function during embryogenesis could be restricted to higher dipterans.

We are interested in early pattern formation of the flour beetle *Tribolium castaneum*. *Tribolium* is a typical example for short germ embryogenesis, representing the ancestral type of embryogenesis in insects (Tautz et al., 1994), albeit not in its extreme form like the grasshopper. In contrast to *Drosophila*, only cephalic and thoracic segments, but not abdominal segments, are determined during the blastoderm stage. Furthermore, the most anterior 20% of the *Tribolium* blastoderm cells form an extra-embryonic membrane, the serosa. This structure is not found in this form in higher Dipterans like *Drosophila*, but is again an ancestral feature of insect embryogenesis. Prior to gastrulation, most blastoderm cells move from anterior and dorsal positions towards the posterior ventral region where they form the embryo proper. This germ rudiment then continues to grow from its posterior end to form a germ band which eventually encompasses all abdominal segments.

So far, it has been impossible to recover a credible *bcd* homologue from *Tribolium*. Although *bcd* codes for a homeobox protein which would be expected to be conserved, the comparison of *bcd* homologues in higher dipterans already showed that the amino-acid replacement rate in this particular gene is exceptionally high (Schröder and Sander, 1993). As the evolutionary distance between beetles and *Drosophila* is more than twice the maximal distance among known *bcd* homologues within the diptera, there might well be too many changes in the homeobox to allow recovery of a *bcd* homologue by low stringency screens or redundant PCR primers. Therefore, the failure to identify *bcd* in *Tribolium* does not exclude the existence of a functional *bcd* homologue. To approach this problem from a different angle, we decided to look at known direct target genes of *bcd*, i.e. *cad* and *hb*, in this species (*Tc'cad* and *Tc'hb*). These two genes are much more conserved and could indeed be recovered from *Tribolium*. We used assays in *Drosophila* to show that most likely there is a functional *bcd* homologue in *Tribolium* after all. Moreover, our results suggest a molecular mechanism for the shift in position of the *hb* gap domain that occurred during evolution from short germ to long germ insects. Such repositioning of gap domains is responsible for one key difference between short and long germ embryos, i.e. their distinct blastoderm fate maps.

MATERIALS AND METHODS

In situ hybridization and immunohistochemistry

Pretreatment and fixation of embryos, as well as whole-mount in situ hybridization were done as described by Klingler and Gergen (1993); Tautz and Pfeifle (1989); Wolff et al. (1995). For visualization of the rather weakly expressed *Tc'hb-lacZ* transcripts, riboprobes yielding especially low background were used. For this end, in vitro transcription reactions (digoxigenin or fluorescein RNA labeling kit; Boehringer) were treated with DNAse I (Boehringer) and purified with Ultrafree-MC columns (100 000 NMWL, Millipore), without further precipitation and without hydrolyzation.

Simultaneous colorimetric detection of two different transcripts was done as published (Hauptmann and Gerster, 1994; Wolff et al., 1995). In addition, a novel double-fluorescence in situ hybridization procedure was developed involving HNPP/fast red (Boehringer) and tyramide signal amplification (NEN Life Science Products). Details of this protocol will be presented elsewhere but are available from the authors on request.

Tc'CAD protein was detected with an antibody described by Schulz et al. (1998). This antibody is specific for *Tribolium* and does not cross-react with the Dm'CAD protein in whole-mount stainings.

Mutant strains, generation of germ-line clones and transgenic lines

As wild-type strains, *Drosophila melanogaster* Oregon R and *Tribolium castaneum* San Bernardino were used. For genetic analysis, crossings were done with the *Drosophila* maternal-effect mutants bcd^{GB}, bcd^{E1}, $bcd^{2\text{-}13}$, bcd^{085} (Frohnhöfer and Nüsslein-Volhard, 1986), osk^{54}, osk^{150} (Lehmann and Nüsslein-Volhard, 1986) and tor^{4} (Schüpbach and Wieschaus, 1986).

Embryos lacking maternal *Dm'cad* function were obtained by generating cad^{2} germ-line clones with the FLP-DFS technique (Chou et al., 1993; Rivera-Pomar et al., 1995). Maternally and zygotically mutant embryos were generated by additional use of a cad^{3}/CyO *hb-lacZ* line (all lines kindly provided by R. Rivera-Pomar).

P-element mediated transformation was done as described by Rubin and Spradling (1982). Recipient strains were either ry^{506} for pCarnegie20 and pHZ50 based constructs or $y\ w^{67c23}$ for pCasper plasmids, respectively. For all transgenic lines generated, a minimum of two (but for most constructs, five) independent insertions were examined.

Plasmid constructs

To test translational regulation of the *Tc'cad* gene in *Drosophila*, a 0.9 kb *Xba*I/*Xho*I fragment from an almost full-length *Tc'cad* cDNA clone (Schulz et al., 1998) was cloned into a pBluescript plasmid (Stratagene) containing the maternal promoter of *Dm'hb* (Schulz and Tautz, 1994). Promoter plus *Tc'cad* gene were then transferred as a1.9 kb *Not*I/*Xho*I fragment into *Not*I and *Sal*I sites of a pCarnegie20 derivative (Rubin and Spradling, 1983), a gift from S. Cohen.

For the largest *Tc'hb-lacZ* construct (Fig. 4B, top), a genomic 5.2 kb *Xho*I/*Sac*II fragment (pBS-Tc'hb-XS5.2) was cloned into pBluescript using *Eco*RV and *Xho*I restriction sites (the genomic *Sac*II site was blunted). The whole region could then be transferred as a 5.2 kb *Bam*HI fragment into the pCasper-β-Gal vector (Thummel et al., 1988), resulting in an in-frame fusion with the *lacZ* gene within the first coding exon of *Tc'hb* (see also Fig. 4A,B). A similar construct was generated using pHZ50 (Hiromi and Gehring, 1987) as vector for transformation (the basal *hsp70* promoter of pHZ50 was removed).

To generate deletions within the *Tc'hb* regulatory region (2nd to 4th construct in Fig. 4B), a 5.2 kb *Xho*I/*Xba*I fragment from pBS-Tc'hb-XS5.2 was subcloned into pGem7Zf (Promega). Nested deletions within this fragment were then generated by restriction with *Cla*I, *Xho*I/*Sal*I or *Hin*dIII and subsequent religation. The remaining truncated *Tc'hb* upstream sequences were then transferred as 3.2 kb, 2.6 kb and 0.3 kb *Bam*HI fragments, respectively, into pCasper-β-Gal.

The construct containing the region distal of the P2 promoter but lacking P3 (5th construct in Fig. 4B), was generated by PCR using the High Fidelity Expand system (Boehringer). A pBluescript-specific upper primer (5′ AATACGACTCACTATAG) and a P2-5′-UTR-specific lower primer (5′ CCGACTGTTCGAAATTTTTGCAC) were used with the pBluescript subclone of *Tc'hb* (see above) as template. PCR products were digested with *Xho*I and cloned into *Xho*I/*Eco*RV sites of pBluescript. This region was then ligated as 3.0 kb *Bam*HI fragment into the respective site of pCasper-AUG-β-Gal (Thummel et al., 1988). The last construct which contained both promoters but lacked most DNA between P2 and P3 was generated by *Sal*I/*Hin*dIII digest of the pBluescript subclone of *Tc'hb* genomic region, and subsequent blunting and religation reactions. The resulting truncated region was again transferred as a 3.6 kb *Bam*HI fragment into pCasper-β-Gal.

In vitro footprinting

For in vitro binding assays, recombinant Dm'CAD protein was overexpressed in *E. coli BL21* (Stratagene) and crude extracts were prepared as described by Rivera-Pomar et al. (1995). Purification of the 6xHis-tagged protein was done under denaturing conditions with TALON resin (Clontech) according to the manufacturer's instructions. Purified protein was renatured through dialysis against Kadonaga buffer B (Kadonaga et al., 1987) with progressive reduction of urea concentrations.

DNaseI-footprinting and Maxam-Gilbert reactions were performed according to Dynan and Tijan (1983) with slight modifications (Rivera-Pomar et al., 1995), using components of the SureTrack footprinting kit (Pharmacia).

RESULTS

Regulation of *Tribolium caudal* by *Drosophila bicoid*

The earliest target gene of *bcd* during *Drosophila* embryogenesis is *cad*. *bcd* regulates the translation of *cad* by directly binding to the 3′-end of its mRNA (Chan and Struhl, 1997; Dubnau and Struhl, 1996; Rivera-Pomar et al., 1996). We have cloned a *cad* homologue from *Tribolium* and studied its expression pattern (Schulz et al., 1998). In the early blastoderm, *Tc'cad* mRNA is homogeneously distributed throughout the embryo. Prior to gastrulation, it becomes restricted to the posterior half of the egg, and as the germ band forms, *cad* mRNA gets confined to the growth zone. To detect a possible translational regulation of maternal *cad* mRNA, we prepared an antibody against the Tc'CAD protein. This antibody detects an initial homogeneous staining in all early blastoderm nuclei (Fig. 1A), as was observed for the maternal *Tc'hb* expression (Wolff et al., 1995). Subsequently, Tc'CAD protein withdraws from the anterior half of the blastoderm embryo, which could indicate translational regulation by an anteriorly localized factor (Fig. 1B). To determine whether this relates to a *bcd*-like activity, we transformed a construct into *Drosophila* bearing the *Tc'cad* gene under the control of a maternal promoter (see Methods). The transgene mRNA is equally distributed throughout the embryo (not shown). With the antibody specific for Tc'CAD, however, we find that Tc'CAD is excluded from the anterior quarter of the *Drosophila* blastoderm embryo (Fig. 1C). Therefore, *Tc'cad* in *Drosophila* is translationally repressed at the anterior pole like *Dm'cad*. To see if this regulation is dependent on *bcd*, we tested the *Tc'cad* construct in a *Drosophila* mutant background devoid of this maternal factor. Indeed we find that Tc'CAD protein is homogeneously distributed in *Drosophila* embryos lacking *bcd* activity (Fig. 1D). In principle, this effect of *bcd* on *cad* translation could be indirect, i.e. *bcd* might activate transcription of a factor that would regulate translation of *Tc'cad*. To exclude the possibility of zygotically expressed intermediate(s), we looked at Tc'CAD in unfertilized eggs (in a wild-type background). Also in this situation, a Tc'CAD protein gradient is formed (Fig. 1E). These findings strongly suggests that *bcd* regulates *Tc'cad* mRNA in the same way as it does *Dm'cad* RNA. Since *Drosophila* and *Tribolium* are separated by about 500 Myr of independent evolutionary history, the ability of *Tc'cad* mRNA to interact with BCD protein clearly demonstrates that the evolution of the *Tc'cad* sequence is functionally constrained and that a similar regulatory interaction must also occur in *Tribolium*. Therefore, this result strongly suggests that a *bcd*-like activity exists in the short germ embryo of *Tribolium*.

Regulation of *Tribolium hunchback* in *Drosophila*

The second well characterized target gene of *bcd* is *hb*. In this case, *bcd* acts as a transcriptional regulator on the *hb* promoter, thereby activating zygotic *hb* expression in the anterior half of the embryo. We have previously described *hb* expression in *Tribolium* (Wolff et al., 1995). There are two early zygotic expression domains of *Tc'hb*, one in the prospective serosa nuclei and one in the posterior half of the egg where the germ rudiment is about to form (see also left column in Fig. 2). From comparison of this pattern with that of pair-rule genes in *Tribolium* we have concluded that the posterior domain in *Tribolium* is the equivalent of the anterior domain in *Drosophila*, and that *Drosophila* (which has no serosa) lacks an equivalent for the anterior domain (Wolff et al., 1995). To see if the *bcd*-like activity described above may also serve as a transcriptional regulator of *Tc'hb*, we transformed a reporter gene construct with approximately 5 kb of *Tc'hb* upstream sequences into *Drosophila* (see Methods). The earliest expression of this *Tc'hb-lacZ* gene occurs in a posterior-to-anterior gradient, which is reminiscent of the CAD gradient in *Drosophila* (Fig. 2A, central column). Shortly afterwards, two separate domains become evident, one in the anterior region and one covering the posterior pole. These two domains are reminiscent of the early zygotic *hb* expression in *Tribolium* (Fig. 2B). Both domains later retract from the poles (Fig. 2C, D). At the end of the blastoderm stage, the posterior domain disappears (Fig. 2E). Later, in the extended germ band, we also find a neuronal expression pattern which represents a subset of neuroblasts that also express endogenous *Dm'hb* (Fig. 2F). Altogether, this tightly regulated expression pattern suggests that the regulatory regions of *Tc'hb* are faithfully recognized by regulatory factors in *Drosophila*. In particular, it appears that the anterior blastoderm domain of the reporter gene is the equivalent of the serosa domain in *Tribolium* and that its posterior blastoderm domain is the equivalent of the gap domain in *Tribolium*. Specifically, double staining for the endogenous *Dm'hb* mRNA and the RNA from the *Tc'hb-lacZ* reporter gene shows that both domains arise at the same time (Fig. 2A,B, right column), prior to the formation of the posterior domain of endogenous *Dm'hb* (Fig. 2C,D).

To see how these *Tc'hb* domains are regulated in *Drosophila*, we introduced the reporter construct into embryos of different maternally mutant backgrounds. In a hypomorphic *bcd* mutant background we find that the anterior domain disappears, while the posterior domain expands somewhat anteriorly (Fig. 3B). In amorphic *bcd* alleles, the posterior domain expands even more, but does not reach the anterior pole (Fig. 3C). The repression at the anterior pole in bcd^{-} is due to *torso* activity, because in a *torso* mutant background we find that both reporter gene domains are still visible, but do not recede from the poles (Fig. 3E). In a *oskar* mutant background (where embryos lack *nanos* activity), both domains remain unchanged (Fig. 3D). These results indicate that the anterior domain of *Tc'hb-lacZ* is activated by *bcd* (directly or indirectly). In contrast, the posterior domain of the reporter gene is repressed by *bcd* as well as by *torso*. Thus, this domain is regulated unlike any *Dm'hb* domain, i.e. it is neither activated by *bcd* (like the anterior domain of *Dm'hb*), nor is it

activated by the terminal system (like the posterior domain of *Dm'hb*; Margolis et al., 1995).

Drosophila caudal can activate *Tribolium hunchback*

In *Drosophila*, *cad* is known to act as a transcriptional activator of posteriorly expressed genes (Rivera-Pomar et al., 1995). To test the influence of *Dm'cad* on the posterior *Tc'hb-lacZ* domain in *Tribolium*, we generated embryos maternally mutant for *Dm'cad* (see Methods). In this genetic background, we find that the posterior domain of *Tc'hb-lacZ* is indeed lost, while the anterior domain is unaffected (Fig. 3F). To check whether CAD functions as a direct activator of *Tc'hb*, we constructed a series of reporter gene constructs with different *Tc'hb* upstream fragments. This allowed us to pinpoint a region around the P2 promoter (Wolff et al., 1995) as the region that mediates expression of the posterior *Tc'hb-lacZ* domain in *Drosophila* (Fig. 4). Using in vitro footprinting on this fragment, we find that Dm'CAD binds at multiple sites in this region. These sites are similar to the previously published CAD binding consensus (see legend for Fig. 4; Dearolf et al., 1989b; Rivera-Pomar et al., 1995). Thus, the high density of CAD binding sites in the fragment that mediates the CAD-dependent regulation strongly suggests that *cad* functions as a direct activator of *Tc'hb* in *Drosophila*. This finding can also explain why the earliest expression of the *Tc'hb-lacZ* construct in *Drosophila* occurs in a way that reflects the maternal *cad* gradient (Fig. 2A). Further, we note that the observed negative regulation of the posterior reporter gene domain by *bcd* (Fig. 3B,C) is likely to be indirect, since in *bcd* mutants the *Dm'cad* expression domain expands anteriorly, which must result in concomitant expansion of this *lacZ* domain.

To see whether *cad* could indeed mediate the activation of *hb* in *Tribolium*, we doubly stained *Tc'cad* and *Tc'hb* in early embryos. For this we employed a newly developed procedure for fluorescent double in situ hybridization suitable for confocal imaging (see Methods). Fig. 5 clearly shows that the earliest zygotic activation of *Tc'hb* in the posterior half of the embryo occurs at a time when *Tc'cad* is restricted to the same region, indicating that this domain could well be directly activated by CAD in *Tribolium* embryos.

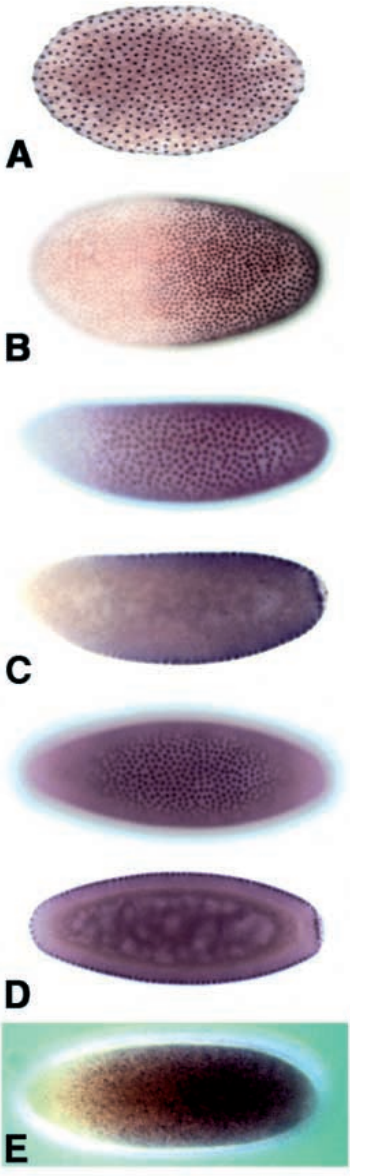

Fig. 1. Expression of Tc'CAD protein in *Tribolium* and *Drosophila*. (A,B) Distribution of Tc'CAD in early and late blastoderm embryos in *Tribolium*. (C) Expression of Tc'CAD protein in transgenic *Drosophila* embryos (two focal planes are shown). Maternally expressed *Tc'cad* mRNA is homogeneously distributed (not shown); anterior repression of *Tc'cad* translation therefore indicates posttranscriptional regulation. (D) Two focal planes of a embryo carrying the same construct, but in a bcd^{E1}/bcd^{GB} mutant background. In this situation, all blastoderm nuclei, including the anterior pole, express *Tc'CAD* protein. (E) A protein gradient is also formed in unfertilized *Drosophila* eggs (genotype as in C), indicating that this regulation is not mediated by zygotic target genes of *bcd*. As in the other figures, embryos are oriented dorsal up and anterior to left.

DISCUSSION

One way to investigate gene regulation in species not amenable to germline transformation is to transfer the gene in question into a model system like *Drosophila* to carry out functional experiments. This functional approach is especially important in cases where there is no obvious sequence conservation, as is the case for both, *hb* upstream DNA as well as for the *cad* 3′UTR. We have so far investigated the regulation of four *Tribolium* genes in *Drosophila*, and in all four cases we obtained meaningful expression patterns. As shown here, *Tc'cad* translation is properly regulated in *Drosophila*, and the *Tc'hb* promoter is regulated in a spatial pattern very similar to that in *Tribolium*. In two more attempts, we tested *Tc'hairy* (C. W., unpublished) and *Tc'tailless* (C. Eckert and R. S., unpublished) upstream DNA in *Drosophila*. Also in these cases, we obtained patterns reminiscent of the expression in *Tribolium*. Therefore, it appears that the molecular machinery in both species is sufficiently similar for *Drosophila* regulatory factors to properly recognize *Tribolium* sequences. Since it is often assumed that evolutionary changes are due to changes in gene regulation, this technique should help us to understand important aspects of the evolution of early patterning mechanisms.

Evidence for a *bcd* homologue in *Tribolium*

The data presented in Fig. 1 show that *Tc'cad* is posttranscriptionally regulated in *Drosophila* embryos, and that this regulation depends on *bcd*. Since a Tc'CAD protein gradient is also formed in unfertilized eggs, no zygotic target gene of *bcd* can act as an intermediate, which makes it very likely that *bcd* itself binds *Tc'cad* mRNA. This ability of *Tc'cad* to interact with *bcd* can most easily be explained by the assumption that *cad* in *Tribolium* is translationally regulated by a *bcd* homologue. The failure by us and others to isolate this homologue may indicate that *bcd* has acquired sequence changes in the homeobox which render detection by PCR and library screening more difficult.

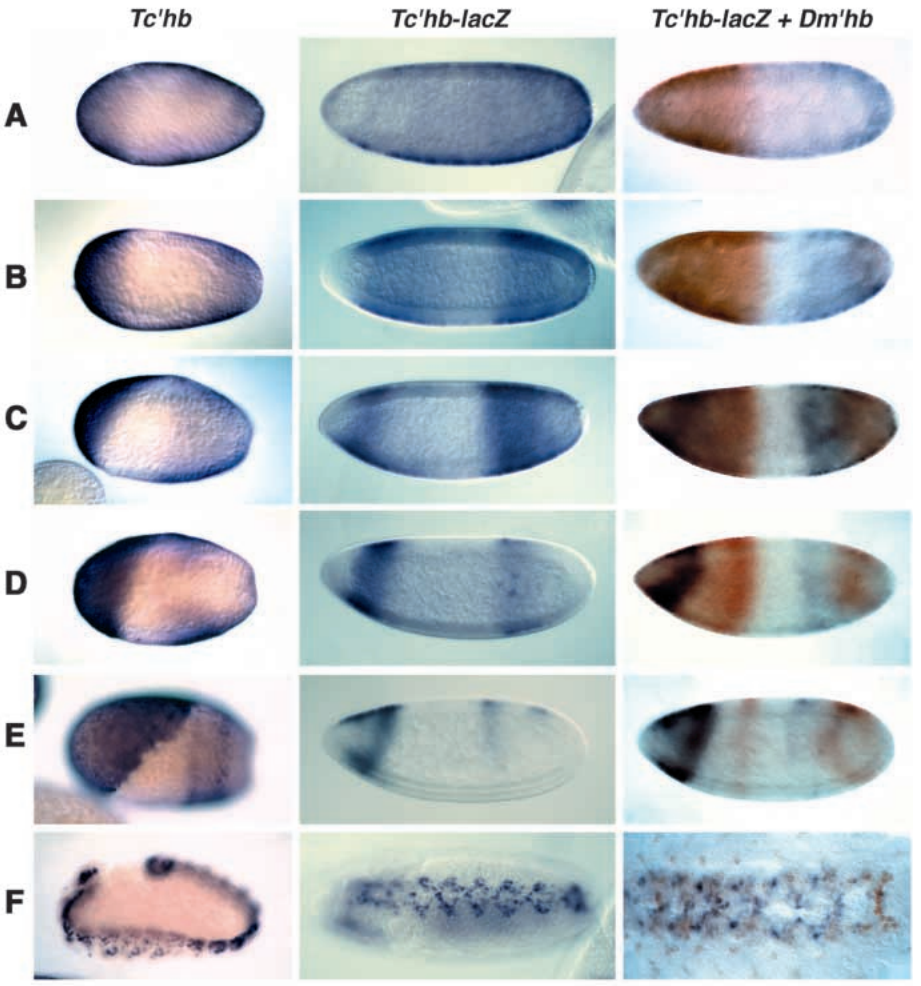

Fig. 2. RNA expression of the *Tc'hb-lacZ* reporter gene in *Drosophila* (central column). For comparison, *Tc'hb* expression in *Tribolium* is also shown (left column), and in the right column *Drosophila* embryos are pictured which are doubly stained for the reporter gene and endogenous *Dm'hb*. The reporter gene includes 5 kb of *Tc'hb* upstream DNA. A-F represent embryos of similar developmental stages across the columns. Early in the syncytial blastoderm, the reporter gene initiates near homogeneous expression (A) which resolves into two domains of which the posterior includes the posterior pole (B). During cellularization, these domains contract somewhat, and the posterior domain retracts from the pole similarly to *Tc'hb* (C,D). At the end of the cellular blastoderm, the posterior domain dissolves into two stripes and finally disappears (E). In the fully extended germ band, expression in the ventral nerve cord becomes apparent in *Drosophila* and *Tribolium* (F). The double in situ staining in the right column demonstrates that the reporter gene initiates expression coincidentally to the endogenous *Dm'hb* gap domain, and it shows that the reporter gene is expressed in a subset of CNS cells which also express the endogenous *Dm'hb* gene. The *Drosophila* embryos in F are ventral views.

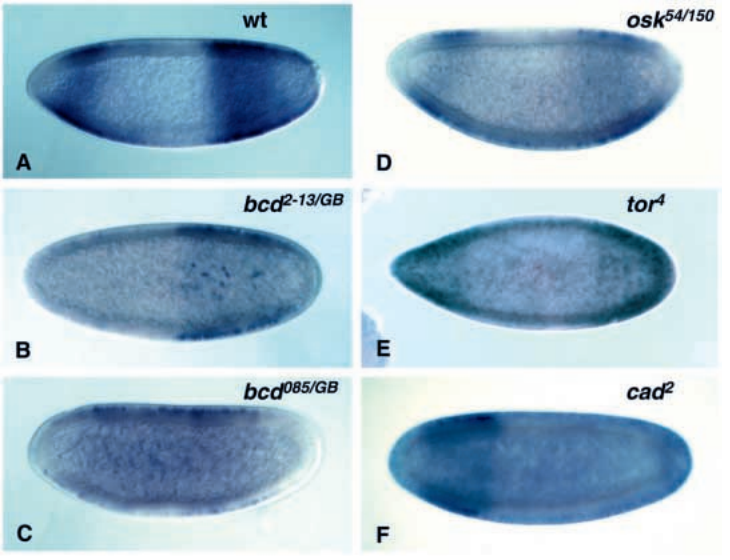

Fig. 3. Regulation of *Tc'hb-lacZ* by *Drosophila* segmentation genes. (A) Wild-type embryo. (B) In embryos of reduced *bicoid* activity ($bcd^{2\text{-}13}/bcd^{GB}$), the anterior domain is lost and the posterior domain expands anteriorly. (C) In a combination of strong *bicoid* alleles (bcd^{085}/bcd^{GB}), the posterior domain expands almost to the anterior pole. (D) In $oskar^{54}/oskar^{150}$ embryos, the size and position of the two domains is unaffected. (E) In $torso^{4}$, both the anterior and the posterior domains fail to retract from the poles. (F) The posterior domain of *Tc'hb-lacZ* is missing in embryos maternally mutant for *Dm'cad*. In B-F, the amount of transcript is reduced relative to A because only one copy of the transgene was present in mutant backgrounds.

Additional support for this interpretation comes from our finding that a *Tc'hb-lacZ* transgene in *Drosophila* is regulated by *bcd* as well. In this case, our promoter mapping experiments could not place this regulatory input in a defined fragment. However, we note that the upstream region surveyed shows multiple scattered *bcd* consensus binding sites. Such a distribution of *bcd* binding sites over a larger region also has been observed in *Musca domestica* (Bonneton et al., 1997).

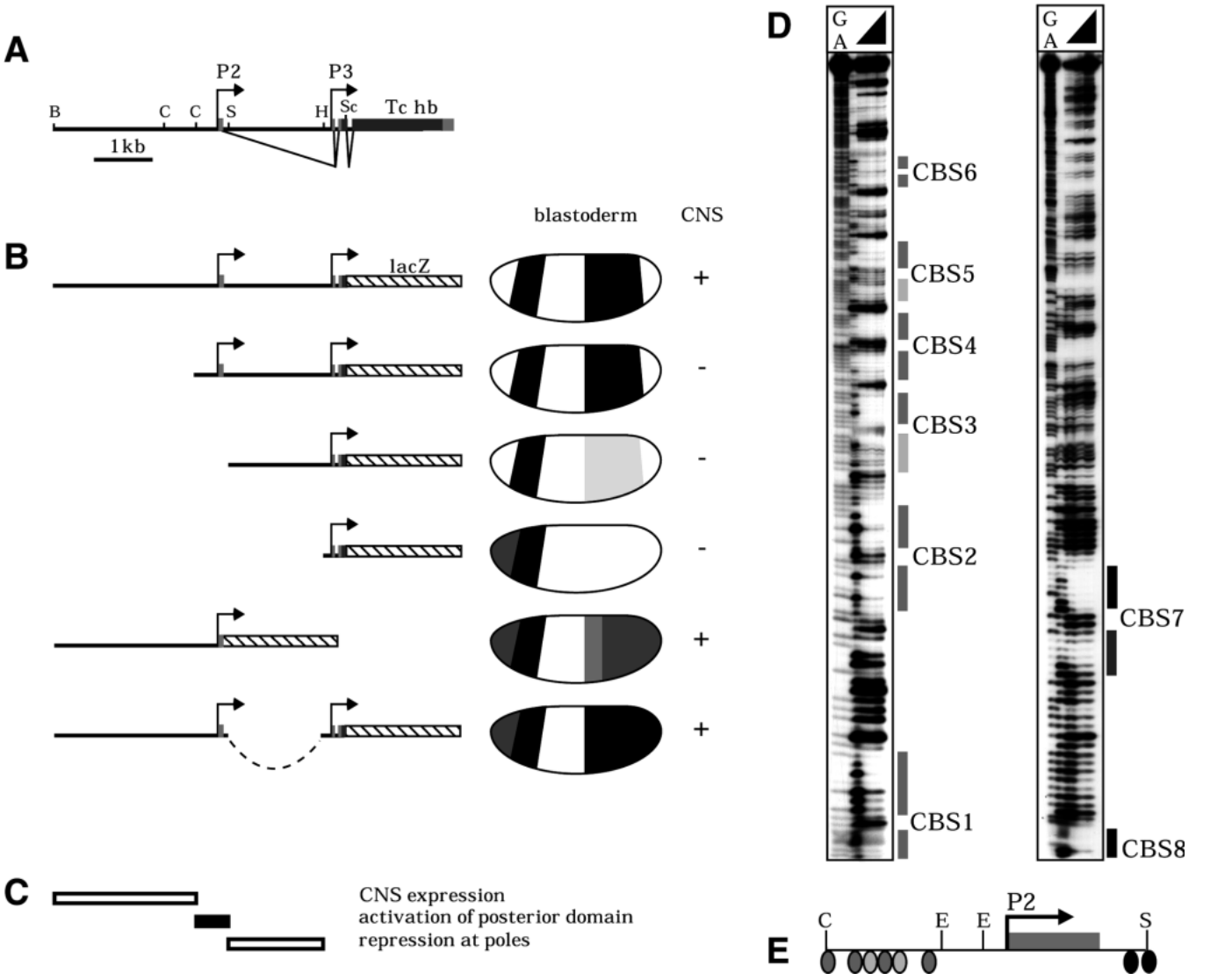

Fig. 4. Identification of a *Tc'hb* upstream fragment required for activation of the gap domain, and identification of CAD binding sites within this fragment. (A) Map of the *Tc'hb* locus. The region depicted includes two of the three promoters (P2, P3). The coding region (black) is interrupted by a short intron. Transcripts starting at P2 and P3 are spliced to the same coding sequences via introns in the untranslated regions (gray; Wolff et al., 1995). (B) Upstream fragments contained in different reporter gene constructs. *lacZ* mRNA expression from these construct is indicated to the right, i.e. expression in the blastoderm stage (embryo sketches) and presence/absence of CNS expression (+/–). (C) These constructs identify a region of 560 bp which is required for the *caudal*-dependent posterior domain of *Tc'hb-lacZ* (black box). In addition, a region required for repression near the poles is indicated, as well as a region required for expression in the CNS. The anterior domain is present in all constructs tested, indicating that sites mediating activation of this domain by *bicoid* may be dispersed throughout the region. Alternatively, the anterior domain in these lines could be due to sites in the transformation vector. Such artifactual anterior expression has occasionally been observed with pCasper vectors (Klingler et al., 1996). However, the anterior domain in our constructs arises much earlier during embryonic development than is the case with this artifactual expression. In addition, we have tested the largest construct in a different transformation vector, pHZ50 (Hiromi and Gehring, 1987), for which no anterior artifact has been described. Expression from this vector was identical to that in embryos carrying the corresponding Casper vector (not shown). (D) Footprinting reactions with bacterially expressed Dm'CAD (Rivera-Pomar et al., 1995). The *Cla*I-*Sal*I fragment highlighted in C was further cut with *Eco*RI and the resulting fragments labeled at the *Cla*I (left panel) and *Sal*I ends (right panel). In each panel, lane 1 is the Maxam-Gilbert reaction, lane 2 is CAD-free control, and lanes 3-5 contain 1, 5 and 10 μg of total protein per reaction, respectively. Protected sites are indicated at the side of each panel, each binding site consisting of two half sites (Dearolf et al., 1989a). Strength of protection is indicated by different gray shades. (E) Detail of the same *Tc'hb* upstream fragment. Ovals indicate the location of protected binding sites as identified in D. Six binding sites are located upstream of the P2 promoter, but the two strongest sites are situated downstream, in the intron (CBS8, CBS9). The sequences of these CAD binding sites (CBS1 to CBS8) are: TCGAAAN$_7$CATAAA, ATCTAAN$_6$TTTAAC, TTTATTN$_3$TTTATT, TTTGGAN$_3$TTTATG, ACTAAAN$_4$TTTACG, TTTCAGN$_4$TTTATT, TTTATTN$_5$TTTGCA and AGTAAAN$_4$CATAAA. Restriction sites in A and E are B, *Bam*HI; C, *Cla*I; E, *Eco*RI; H, *Hin*dIII; S, *Sal*I; Sc, *Sac*II.

Moreover, expression of the anterior domain of *Tc'hb-lacZ* is very similar to that of artificial constructs carrying low-affinity *bcd* binding sites (Ronchi et al., 1993). Yet, we cannot exclude the possibility of indirect action, i.e. zygotic genes depending on *bcd* could be responsible for the anterior expression of *Tc'hb-lacZ*. One potential intermediate factor, *orthodenticle*, is expressed in a pattern similar to the anterior domain of *Tc'hb-lacZ* (Finkelstein and Perrimon, 1990). However, this domain is still present in a otd^{YH} mutant background (data not shown). Nevertheless, other anteriorly expressed genes (not all of which are known to date) might still be involved in this regulation. Therefore, we regard the regulation of *Tc'hb-lacZ* by *bcd* as less persuasive evidence for the existence of a *bcd* homologue in *Tribolium* than the regulation of *Tc'cad*.

As an alternative interpretation one might consider that some other anteriorly localized factor could regulate *cad* in *Tribolium*. Intriguingly, a *cad* homologue in *Caenorhabditis* appears to be regulated by *mex-3*, which binds RNA with KH domains (Draper et al., 1996; Hunter and Kenyon, 1996). In *Drosophila*, *bcd* is so far the only known regulator of *cad* translation, and it is almost certain that *bcd* alone is responsible for the spatial aspect of the regulation. However, other factors may participate in this interaction, binding to the same 3′UTR sequences. Conceivably, such a 'cofactor' could form an anterior gradient in *Tribolium* and spatially regulate *Tc'cad*. This could explain conservation of the interacting mRNA sequences which then might still be recognized by *bcd* in *Drosophila*. If this model was true, it would be very interesting to see to what degree this factor would mimic, in *Tribolium*, the role of *bcd* in *Drosophila*, i.e how it would become localized (by a mechanism similar to *bcd*?), and how it might regulate anteriorly positioned expression domains like the anterior domain of *Tc'hb* (should this factor also be able, like *bcd*, to regulate translation as well as transcription of target genes?). We note that only the early expression of *Tc'hb* (Fig. 2B, left column) is likely to be under direct control of an anterior (presumably maternal) gradient in *Tribolium*. The later, sharply demarcated expression in the serosa cells (Fig. 2E, left column) is likely under zygotic control, possibly through interaction with *Tc'zen* which is also expressed in all serosa cells (Falciani et al., 1996).

That *Tribolium* has a functional *bcd* gene appears to us the more parsimonious explanation, however. If true, probably all holometabolous insects have *bcd*. It is then still an open question, if its range extends to hemimetabolous insects of the 'extreme short germ' type (Tautz et al., 1994), like *Schistocerca*, where the embryo forms in an area far from the anterior pole. If *bcd* indeed functions in these embryos, it would either have to become localized to a different position within the embryo, or it may form a protein gradient by some other mechanism. In any case, our evidence for the existence of *Tc'bcd* encourages new efforts to isolate the gene from *Tribolium*, and this then should provide sequence information to allow isolation of the gene from additional phylogenetic groups.

A switch in *hunchback* regulation

While our results provide evidence for a functional *bcd* homologue in *Tribolium*, they suggest a role somewhat different from that in *Drosophila*. An important function of *bcd* in *Drosophila* is the activation of the *hb* gap domain which is located anteriorly in this species. In *Tribolium*, the posterior blastoderm domain of *Tc'hb* must exert this segmentation function. However, an anteriorly positioned *bcd* gradient in *Tribolium* is not a good candidate for activating transcription in the posterior *Tc'hb* domain. The reporter gene data show that *Tc'hb* upstream sequences are capable of mediating expression in two domains in *Drosophila* embryos. The similarity of this pattern to that of *hb* in *Tribolium* suggests that the anterior reporter gene domain corresponds to the serosa domain of *Tc'hb*, while the posterior one represents the *Tc'hb* domain that carries the gap gene function. The anterior reporter gene domain is dependent on *Dm'bcd* activity, which is consistent with our interpretation that a *bcd* homologue is present also in *Tribolium*. The posterior domain of the reporter gene, however, depends on *Dm'cad*. Also this is consistent with the situation in *Tribolium* if we assume that this reporter gene domain corresponds to the gap domain of *Tc'hb*, because this posterior *Tc'hb* domain overlaps with the *Tc'cad* gradient in a way that suggests that it could be activated by *Tc'cad*. Therefore, it appears that a qualitative switch has occurred during evolution of the *hb* gene, since this domain is activated by *cad* in one species, and by *bcd* in the other (see below).

An alternative interpretation, namely that the posterior *Tc'hb-lacZ* domain was the equivalent of the posterior *Dm'hb* domain, would also imply that a regulatory switch has occurred in the evolution of *hb*, in this case a switch from activation by the terminal system in *Drosophila* (Margolis et al., 1995) to activation by *cad* in *Tribolium*. However, this alternative interpretation is not likely to be true, because the equivalent of the posterior *Dm'hb* domain in *Tribolium* arises at a much later stage, in the growing germ band (Wolff et al., 1995). We find it implausible that the regulatory environment in the growing germ band of *Tribolium* should be similar to that of the early syncitial blastoderm in *Drosophila*. Moreover, we have shown that the posterior domain of the reporter gene is negatively regulated by the terminal system (Fig. 3). However, the major executor of the terminal system in *Drosophila* – the *tailless* (*tll*) gene – is not expressed in the *Tribolium* germ band. Instead, *Tc'tll* is expressed in the blastoderm (R. S. unpublished) where it is well suited to repress the *Tc'hb* gap domain at the posterior pole. Thus, the observed inhibition of the reporter gene by the terminal system in *Drosophila* is consistent with the first, but not with this second interpretation.

The negative regulation of *Tc'hb-lacZ* by the terminal system in *Drosophila* is separable from the activating elements in the promoter mapping experiments, suggesting that it is mediated via specific regulatory elements (Fig. 4). Together with our observations about *Tc'tll*, this supports a blastoderm function for a terminal system in *Tribolium* as well. In the context of the *Tribolium* embryo, however, a terminal influence is only obvious for the posterior pole of the egg. Retraction of Tc'HB protein from the posterior pole is probably due to *nanos* activity at early stages, since maternal *hb* RNA at the posterior end of the embryo appears to be translationally regulated (Wolff et al., 1995). Our new data suggest that at subsequent stages the continued retraction of *Tc'hb* mRNA from the posterior pole (see Fig. 2) is due to transcriptional repression by a terminal activity. We have no evidence for an anterior terminal activity in *Tribolium* so far, as the serosa domain in that species always fully covers the anterior pole of the embryo. Thus, the repression of the serosa

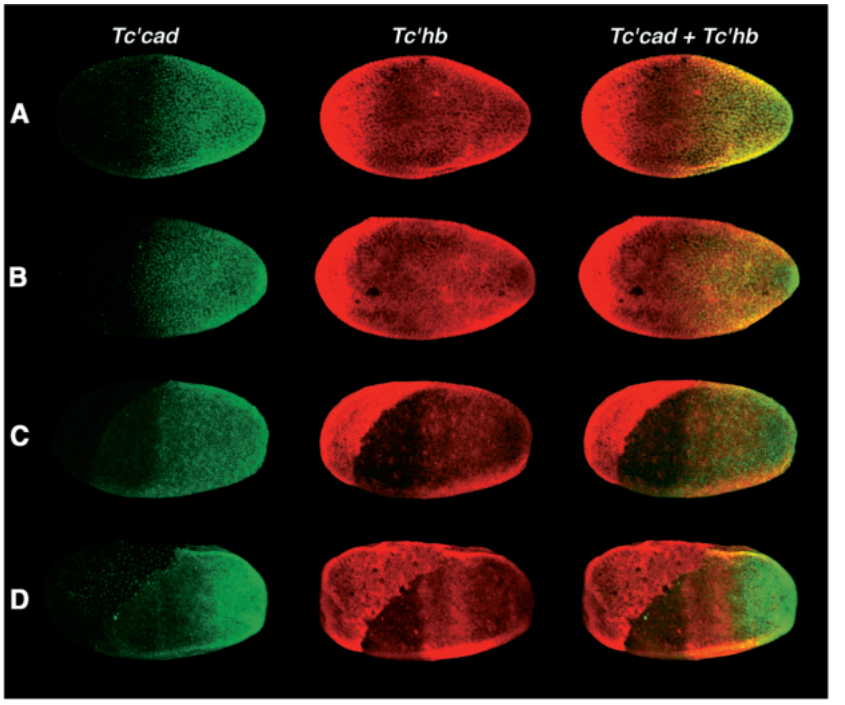

Fig. 5. Overlap of *Tc'cad* and *Tc'hb* mRNA expression domains in *Tribolium* embryos of successive developmental stages. (A,B) blastoderm stages; (C,D) formation of the germ rudiment. The anterior domain of *Tc'hb* demarcates the serosa primordium which expands during germ rudiment formation as embryonic cells move towards ventral and posterior regions. The posterior domain (which has a lower expression) arises within the *Tc'cad* domain. During germ rudiment formation, this domain weakens and resolves into two stripes, while the *Tc'cad* domain retracts towards the region of the prospective growth zone at the posterior pole (D).

domain from the anterior pole in *Drosophila* may reflect the symmetrical localization of *torso* activity in *Drosophila*. In short germ embryos, where no embryonic structures are formed at the anterior pole, terminal activity may not be present at the anterior pole.

A model for the evolution of *hunchback* regulation

We have previously argued that probably there is *nanos* activity in *Tribolium* (Wolff et al., 1995), and in this paper we provide evidence for anterior (*bcd*) and terminal (*torso*) systems in this species. Thus, all three of the anterior-posterior maternal systems known from *Drosophila* appear to be conserved in at least this short germ insect. In *Tribolium*, the gap domains of *hb* and *Krüppel* (*Kr*) occupy more posterior positions than in *Drosophila*, as abdominal segments are not represented in the blastoderm stage. The gap domain of *Dm'hb* is located in the anterior half of the embryo (Tautz et al., 1987), and *Dm'Kr* is expressed in the center of the embryo (Knipple et al., 1985). In *Tribolium*, the corresponding domains of *Tc'hb* and *Tc'Kr* occupy the posterior half of the embryo (Wolff et al., 1995) and the posterior pole (Sommer and Tautz, 1993), respectively. Since the maternal positional information in the *Tribolium* embryo appears to be similar to that in *Drosophila*, these changes must be due to differences in how these gap genes sense the positional information. In the case of *hb*, our results suggest which changes in the *hb* upstream region were necessary to achieve its change in expression. Assuming that the *Tribolium* situation represents the ancestral state (Tautz et al., 1994), it appears that the evolution of long germ development was accompanied by a loss of the *cad* regulatory input in parallel to the acquisition of (additional) *bcd* activating input.

We believe that the key to understanding this qualitative switch lies in the serosa expression domain of *Tc'hb*. Our reporter gene data suggest that this domain may already be

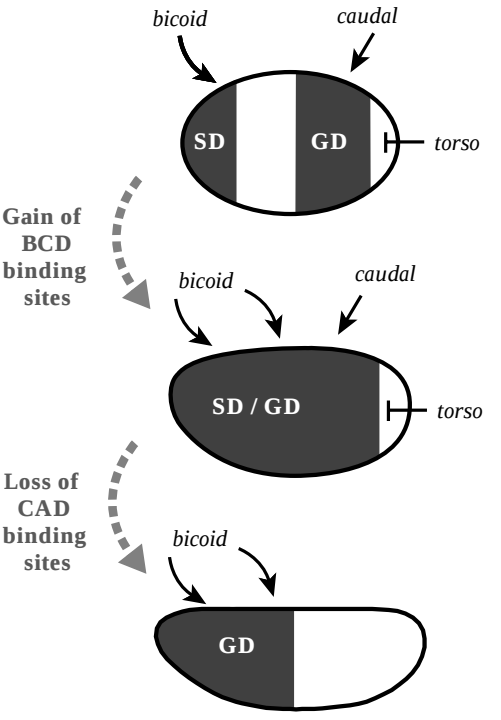

Fig. 6. Model for the evolution of the *hunchback* expression pattern. Top: *hunchback* expression in *Tribolium*, which represents the ancestral state. Regulation of the serosa domain by *bicoid* and regulation of the gap domain by *caudal* and *torso* is indicated. Center: hypothetical intermediate state. A large expression domain is formed by fusion of the two initial domains due to insertion of additional *bicoid* sites. Bottom: in *Drosophila*, the large gap domain is in the anterior half. This situation could be due to the loss of *caudal* activation, thereby shifting the posterior border of the domain towards the anterior.

activated by *bcd* in *Tribolium*. To explain the switch in the regulation of the gap domain, one can envision an intermediate state, where the serosa domain and the embryonic domain have fused into a single domain (Fig. 6). To achieve this, the evolution of a few additional *bcd* binding sites in the *hb* upstream region would have been sufficient. In this intermediate stage both *bcd* and *cad* would have acted as activators on the gap domain of *hb*. Subsequent loss of *cad* regulation could then have moved the posterior boundary of this combined domain towards the anterior. We note that the *Tc'hb* gene has three known promoters one of which appears to be specialized for mediating the *cad* regulation. In *Drosophila*, only two promoters are present, neither of which has a known responsiveness to *cad*. Thus, in all likelihood, the *cad* dependent promoter and its associated enhancer was lost. As we have found no other enhancer activity for later expression patterns of *hb* in the *cad* dependent fragment, the loss of this region could have been a single step. Intriguingly, a combined serosa and gap domain is still evident in the lower dipteran *Clogmia*. In this fly, *hb* is expressed in a large anterior domain, from which at later stages also the serosa is recruited (Rohr et al., personal communication). This mechanism – modification of the way how gap genes sense maternal positional information while this information itself remains constant – can explain how the blastoderm fate map changed during evolution of short germ insects to insects with long germ embryos. Moreover, it represents an intriguing example for the importance of regulatory adaptation during the evolution of developmental processes.

We thank G. Büttner for technical assistance, T. Zimmermann for help with confocal imaging, A. Ephrussi, C. Nüsslein-Volhard, R. Rivera, G. Struhl, as well as the Bloomington Stock Center for Drosophila strains, R. Rivera, S. Brown, C. Thummel, S. Cohen and N. Patel for DNA constructs and other reagents, and C. David and M. Kania for comments on earlier versions of the manuscript. We also thank the anonymous referees for their suggestions. This work was supported by grants from the Deutsche Forschungsgemeinschaft to D. T. and M. K.

REFERENCES

Bonneton, F., Shaw, P., Fazakerley, C., Shi, M. and Dover, G. (1997). Comparison of *bicoid*-dependent regulation of *hunchback* between *Musca domestica* and *Drosophila melanogaster*. *Mech. Dev.* **66,** 143-156.

Chan, S. K. and Struhl, G. (1997). Sequence-specific RNA binding by *bicoid*. *Nature* **388,** 634.

Chou, T. B., Noll, E. and Perrimon, N. (1993). Autosomal P[*ovo(D1)*] dominant-female-sterile insertions in *Drosophila* and their use in generating germ-line chimeras. *Development* **119,** 1359-1369.

Dearolf, C., Topol, J. and Parker, C. (1989a). Transcriptional control of *Drosophila fushi tarazu* zebra stripe expression. *Genes Dev.* **3,** 384-398.

Dearolf, C. R., Topol, J. and Parker, C. S. (1989b). The *caudal* gene product is a direct activator of *fushi-tarazu* transcription during *Drosophila* embryogenesis. *Nature* **341,** 340-343.

Draper, B. W., Mello, C. C., Bowerman, B., Hardin, J. and Priess, J. R. (1996). *Mex-3* is a KH domain protein that regulates blastomere identity in early *C. elegans* embryos. *Cell* **87,** 205-216.

Driever, W. and Nüsslein-Volhard, C. (1988). A gradient of *bicoid* protein in *Drosophila* embryos. *Cell* **54,** 83-93.

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

Dubnau, J. and Struhl, G. (1996). RNA recognition and translational regulation by a homeodomain protein. *Nature* **379,** 694-699.

Dynan, W. S. and Tijan, R. (1983). The promoter-specific transcription factor SP1 binds to upstream sequences in the SV40 early promoter. *Nature* **337,** 138-143.

Falciani, F., Hausdorf, B., Schröder, R., Akam, M., Tautz, D., Denell, R. and Brown, S. (1996). Class 3 *Hox* genes in insects and the origin of *zen*. *Proc. Natl. Acad. Sci. USA* **93,** 8479-8484.

Finkelstein, R. and Perrimon, N. (1990). The *orthodenticle* gene is regulated by *bicoid* and *torso* and specifies *Drosophila* head development. *Nature* **346,** 485-488.

Frohnhöfer, H. G. and Nüsslein-Volhard, C. (1986). Organization of anterior pattern in the *Drosophila* embryo by the maternal gene *bicoid*. *Nature* **324,** 120-125.

Hauptmann, G. and Gerster, T. (1994). Two-color whole-mount in situ hybridization to vertebrate and *Drosophila* embryos. *Trends Genet.* **10,** 266.

Hiromi, Y. and Gehring, W. J. (1987). Regulation and function of the *Drosophila* segmentation gene *fushi tarazu*. *Cell* **50,** 963-974.

Hunter, C. P. and Kenyon, C. (1996). Spatial and temporal controls target *pal-1* blastomere-specification activity to a single blastomere lineage in *C. elegans* embryos. *Cell* **87,** 217-226.

Kadonaga, J. T., Carner, K., Masiarz, F. and Tijan, R. (1987). Isolation of cDNA encoding transcription factor SP1 and functional analysis of the DNA binding domain. *Cell* **51,** 1079-1090.

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

Klingler, M., Soong, J., Butler, B. and Gergen, J. P. (1996). Disperse versus compact elements for the regulation of *runt* stripes in *Drosophila*. *Dev. Biol.* **177,** 73-84.

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 melanogaster* embryos. *Nature* **317,** 40-44.

Lehmann, R. and Nüsslein-Volhard, C. (1986). Abdominal segmentation, pole cell formation, and embryonic polarity require the localized activity of *oskar*, a maternal gene in *Drosophila*. *Cell* **47,** 141-152.

Macdonald, P. M. and Struhl, G. (1986). A molecular gradient in early *Drosophila* embryos and its role in specifying the body pattern. *Nature* **324,** 537-545.

Margolis, J. S., Borowsky, M. L., Steingrimsson, E., Shim, G. W., Lengyel, J. A. and Posakony, J. W. (1995). Posterior stripe expression of *hunchback* is driven from two promoters by a common enhancer element. *Development* **121,** 3067-3077.

Mlodzik, M. and Gehring, W. (1987). Expression of the *caudal* gene in the germ line of *Drosophila* – Formation of an RNA and protein gradient during early embryogenesis. *Cell* **48,** 465-478.

Rivera-Pomar, R., Lu, X. G., Perrimon, N., Taubert, H. and Jäckle, H. (1995). Activation of posterior gap gene expression in the *Drosophila* blastoderm. *Nature* **376,** 253-256.

Rivera-Pomar, R., Niessing, D., Schmidt-Ott, U., Gehring, W. J. and Jäckle, H. (1996). RNA binding and translational suppression by *bicoid*. *Nature* **379,** 746-749.

Ronchi, E., Treisman, J., Dostatni, N., Struhl, G. and Desplan, C. (1993). Down-regulation of the *Drosophila* morphogen *bicoid* by the *torso* receptor-mediated signal transduction cascade. *Cell* **74,** 347-355.

Rubin, G. M. and Spradling, A. C. (1982). Genetic transformation of *Drosophila* with transposable element vectors. *Science* **218,** 348-353.

Rubin, G. M. and Spradling, A. C. (1983). Vectors for P-element mediated gene transfer in *Drosophila*. *Nucl. Acids Res.* **11,** 6341-6351.

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

Sander, K. (1983). The evolution of patterning mechanisms: gleanings from insect embryogenesis and spermatogenesis. In *Development and Evolution: The sixth Symposium of the British Society for Developmental Biology* (ed. B. C. Goodwin, N. Holder, and C. C. Wylie), pp. 137-159. Cambridge University Press, Cambridge.

Schröder, R. and Sander, K. (1993). A comparison of transplantable *bicoid* activity and partial *bicoid* homeobox sequences in several *Drosophila* and blowfly species (Calliphoridae). *Roux's Arch. Dev. Biol.* **203,** 34-43.

Schulz, C., Schröder, R., Hausdorf, B., Wolff, C. and Tautz, D. (1998). A *caudal* homologue in the short germ band beetle *Tribolium* shows similarities to both the *Drosophila* and the vertebrate *caudal* expression patterns. *Dev. Genes Evol.* **in press**.

Schulz, C. and Tautz, D. (1994). Autonomous concentration-dependent

activation and repression of *Krüppel* by *hunchback* in the *Drosophila* embryo. *Development* **120,** 3043-3049.

Schulz, C. and Tautz, D. (1995). Zygotic *caudal* regulation by *hunchback* and its role in abdominal segment formation of the *Drosophila* embryo. *Development* **121,** 1023-1028.

Schüpbach, T. and Wieschaus, E. (1986). Maternal-effect mutations altering the anterior-posterior pattern of the *Drosophila* embryo. *Roux's Arch. Dev. Biol.* **195,** 302-317.

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

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

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

Tautz, D., Friedrich, M. and Schröder, R. (1994). Insect embryogenesis – what is ancestral and what is derived? *Development* **1994 Supplement,** 193-199.

Tautz, D., Lehmann, R., Schnuerch, 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.

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

Thummel, C. S., Boulet, A. M. and Lipschitz, H. D. (1988). Vectors for *Drosophila* P-element-mediated transformation and tissue culture transfection. *Gene* **74,** 445-456.

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.