Development 126, 5373-5385 (1999)
Printed in Great Britain © The Company of Biologists Limited 1999
DEV8642

spen encodes an RNP motif protein that interacts with Hox pathways to repress the development of head-like sclerites in the *Drosophila* trunk

Elizabeth L. Wiellette, Katherine W. Harding, Kimberly A. Mace, Matthew R. Ronshaugen, Fawn Y. Wang, and William McGinnis*

Department of Biology, University of California at San Diego, La Jolla, CA, 92093
*Author for correspondence (e-mail: mcginnis@biomail.ucsd.edu)

Accepted 21 September; published on WWW 9 November 1999

SUMMARY

***Drosophila* has eight Hox proteins, and they require factors acting in parallel to regulate different segmental morphologies. Here we find that the *Drosophila* gene *split ends (spen)*, has a homeotic mutant phenotype, and appears to encode such a parallel factor. Our results indicate that *spen* plays two important segment identity roles. One is to promote sclerite development in the head region, in parallel with Hox genes; the other is to cooperate with *Antennapedia* and *teashirt* to suppress head-like sclerite development in the thorax. Our results also indicate that without *spen* and *teashirt* functions, *Antennapedia* loses its ability to specify thoracic identity in the epidermis. *spen* transcripts encode extraordinarily large protein isoforms (approx. 5,500 amino acids), which are concentrated in embryonic nuclei. Both Spen protein isoforms and Spen-like proteins in other animals possess a clustered repeat of three RNP (or RRM) domains, as well as a conserved motif of 165 amino acids (SPOC domain) at their C-termini. Spen is the only known homeotic protein with RNP binding motifs, which indicates that splicing, transport, or other RNA regulatory steps are involved in the diversification of segmental morphology. Previous studies by Dickson and others (Dickson, B. J., Van Der Straten, A., Dominguez, M. and Hafen, E. (1996). *Genetics* 142, 163-171) identified *spen* as a gene that acts downstream of Raf to suppress Raf signaling in a manner similar to the ETS transcription factor Aop/Yan. This raises the intriguing possibility that the Spen RNP protein might integrate signals from both the Raf and Hox pathways.**

Key words: *split ends (spen)*, Hox, RNA binding, *Drosophila melanogaster*, *Antp*, Repressor

INTRODUCTION

The homeotic genes in *Drosophila* can be divided into three classes: (1) genes of the homeotic selector, or Hox class, which are located in the Antennapedia and Bithorax complexes, (2) genes of the Polycomb (Pc) and trithorax (trx) classes, which regulate the spatial and temporal location of Hox transcription, (3) genes of the Hox modulator class, such as *extradenticle*, which act in parallel with Hox genes to regulate downstream targets. *Hox* genes are activated in distinct metameres on the anterior-posterior body axis and morphologically diversify those serially arranged fields of cells (Lewis, 1978; McGinnis and Krumlauf, 1992). The Hox proteins do this by their differential regulation of batteries of downstream 'realizator' genes (Garcia-Bellido, 1977; Pradel and White, 1998). Despite much research, the basis of differential realizator gene expression by Hox proteins is illuminated by only a few facts of unknown generality (Chan et al., 1997; Biggin and McGinnis, 1997; Pradel and White, 1998; Li and McGinnis, 1999; Li et al., 1999), and the basis of morphological differentiation under Hox control is even more obscure.

The ability of Hox proteins to differentially regulate batteries of downstream genes and thereby activate different developmental pathways requires a variety of modulatory factors that contribute to the binding occupancy and/or activity of Hox proteins on different target sequences. Such factors include the Hox proteins themselves, which can phenotypically suppress the activity of other family members, usually the posterior Hox function dominating the more anterior (Gonzalez-Reyes et al., 1990). Recent work by Capovilla and Botas (1998) suggests that this phenotypic suppression involves a short range enhancer repression mechanism (Gray et al., 1999). A few other *Drosophila* modulatory factors are apparently also involved in the specificity of Hox output, as evidenced by their homeotic mutant phenotypes in a wild-type genetic background. These include the DNA binding proteins *extradenticle* (*exd*) (Peifer and Weischaus, 1990; Rauskolb et al., 1993), *teashirt* (*tsh*) (Fasano et al., 1991; Alexandre et al., 1996), *cap'n'collar B* (*cncB*) (Mohler et al., 1995; McGinnis et al., 1998) and *homothorax (hth)* (Rieckhof et al., 1997; Pai et al., 1998).

Exd and its mammalian homologs in the Pbx class are homeodomain proteins (Flegel et al., 1993; Rauskolb et al., 1993) which participate in the selection and regulation of many Hox targets by cooperative binding with Hox proteins and concerted activity with them once bound (van Dijk and Murre,

1994; Chan and Mann, 1996; Mann and Chan, 1996; Chan et al., 1996; Pinsonneault et al., 1997; Li et al., 1999). The ability of Exd protein to enter nuclei and interact with Hox proteins is dependent on Hth, another *Drosophila* homeodomain protein (Rieckhof et al., 1997; Kurant et al., 1998; Pai et al., 1998). This activity of Hth, in combination with its pattern of expression, generates cellular patterns of nuclei that either contain or lack Exd protein. In cells that accumulate Exd in nuclei, Hox proteins are capable of serving as transcriptional activators, but in cells that lack nuclear Exd, Hox proteins have only been observed to function as repressors (Pinsonneault et al., 1997; Li et al., 1999). Tsh is a zinc-finger DNA binding protein, and is required in combination with the trunk HOX proteins to repress head-specific developmental pathways and promote trunk identity (Fasano et al., 1991; Röder et al., 1992; Alexandre et al., 1996). For example, Tsh acts with Antp to repress the ectopic activation of the anterior-most *Hox* gene, *labial* (*lab*; Röder et al., 1992). Finally, the CncB protein, a bZIP class of transcription factor, functions as a Hox modulator by preventing the Deformed Hox protein from activating maxillary-specific targets in the mandibular segment (Mohler et al., 1995; McGinnis et al., 1998).

Using genetic modifier screens, we have identified additional factors which influence the function of the *Deformed (Dfd)* gene in *Drosophila*, focusing on those that act in parallel to *Dfd* and other Hox genes (Harding et al., 1995; Gellon et al., 1997; Florence and McGinnis, 1998). The normal function of *Dfd* is required for the morphogenesis of structures arising from the mandibular and maxillary head segments, dorsal ridge, optic lobe and central nervous system (McGinnis et al., 1990; Restifo and Merrill, 1994). Our modifier screens identified mutant alleles of both *exd* and *cncB*, but also identified a gene with a particularly interesting mutant phenotype in which the ventral regions of thoracic (and occasionally abdominal) segments are homeotically transformed to head-like skeletal structures. We previously mapped this gene to the tip of the left arm of chromosome 2, at 21B4-6, and called it *polycephalon* (Gellon et al., 1997). However, we recently found that our *polycephalon* alleles disrupt the previously named *split ends (spen)* gene (Kolodziej et al., 1995), which was originally mapped to the right arm of the 2nd chromosome. We will hereafter use the name *spen* to denote the gene. Mutant alleles in *spen* [called *E(Raf)2A*] also were isolated as enhancers of a dominant Raf-induced eye phenotype by Dickson et al. (1996), and yet other alleles of *spen* [called *E(E2F)2A* or *poc*) were isolated as enhancers of an eye phenotype induced by overexpression of dE2F and dDp (Staehling-Hampton et al., 1999).

The homeotic phenotypes in *spen* mutant embryos are not associated with changes in Hox gene expression, nor is the *spen* expression pattern changed in *Hox* mutants, therefore *spen* appears to act in parallel with Hox segmental identity functions. In addition, over-expression of the thoracic determinants *Antp* or *tsh* can suppress the ectopic sclerites in *spen* mutants, and thus we propose that *spen* acts in combination with Antp and Tsh pathways to regulate the distinction between head and trunk identity. The *spen* gene encodes very large protein isoforms (approx. 5,500 amino acids) with an arrangement of RNP motifs and conserved C-terminal regions that can be used to identify Spen orthologs and paralogs in *D. melanogaster*, *C. elegans* and *H. sapiens*. As members of the RNP motif family exert their diverse functions by sequence-specific binding to RNA (Burd and Dreyfuss, 1994), the Spen proteins may modulate Hox function through a regulatory influence on messenger RNAs produced from Hox downstream genes.

MATERIALS AND METHODS

Strains and genetics

Most mutations, aberrations and abbreviations are described in Lindsley and Zimm, (1992), or in FlyBase (http://flybase.bio.indiana.edu). Other alleles and lines include: $spen^{poc361}$ and $spen^{poc231}$ (Gellon et al., 1997), $spen^{E(Raf)2A16H1}$ and $spen^{E(Raf)2A1401}$ (Dickson et al., 1996; these are listed as *polycephalon* alleles in the current FlyBase), $spen^{E(CycE)D57}$ and $spen^{E(CycE)e9}$ alleles (Christian Lehner, personal communication), *Df(2L)TE21A*, Dll^5, cnc^{VL110}, lab^1, oc^2, ems^3, btd^1, $Antp^{25}$, $Antp^{hs.PHTA}$, $Scr^{C1}Antp^{NsRC3}$, Dfd^{16}, tsh^8, $tsh^{Scer/UAS.cGa}$, 69B-GAL4 (Brand and Perrimon, 1993; Chou et al., 1993), P{$ry^{+t7.2=neoFRT}$}40-A ('FRT-40A'), P{$w^{+mC=ovoD1\text{-}18}$}2L1P{$w^{+mC=ovoD1\text{-}18}$}2L2P{$ry^{+t7.2=neoFRT}$}40-A ('ovo^{D1}-FRT-2L'), P{$ry^{+7.2=hsFLP}$}*12;* CyO/noc^{Sco}. The *PZ1295* line [also known as *l(2)03350*)] was obtained from the Bloomington Stock Center. The *l(2)K06805*, *l(2)K07612*, *l(2)K07721*, *l(2)K08102*, *l(2)K10325* and *l(2)K13625* lines were obtained from the Berkeley Drosophila Genome Project (http://www.fruitfly.org). Embryonic stages are as described by Campos-Ortega and Hartenstein (1985).

Preparation of embryonic cuticles

Mutant chromosomes were outcrossed to eliminate balancer chromosomes from the stocks before mutant cuticles were collected. Embryos were collected for approximately 12 hours and aged for more than 24 hours before preparing cuticles as described by Wieschaus and Nüsslein-Volhard (1986). Overexpression of *tsh* in a *spen* mutant background was accomplished by crossing $spen^{poc361}$/CyO;UAS*tsh*/+ females to $spen^{poc361}$/+;69B-GAL4/+ males. Overexpression of *Antp* in a *spen* mutant background was accomplished by crossing $spen^{poc361}$/+;hs*Antp*/+ females to Df(2L)TE21A/+;hs*Antp*/+ males, collecting embryos for 1 hour intervals, aging at 25°C for 3 hours, and then heat shocking at 37°C for 20 minutes every 2 hours for a total of three heat treatments. Embryos were then aged at 25°C for at least an additional 22 hours.

Generation of maternal germline *spen/spen* clones

Maternal germline *spen* mutant clones were generated using the FLP-FRT system as detailed by Chou et al. (1993). $spen^{poc231}$ and $spen^{poc361}$ were recombined onto the FRT 40A chromosome. *spen* FRT 40A/CyO virgin females were then mated to hs-FLP12; ovo^{D1}-FRT-2L/CyO males; progeny of this cross were heat shocked for 1 hour at 37°C at 3, 4, 5 and 6 days after egg deposition. The heat-shocked female progeny carrying both the *spen* and ovo^{D1} chromosomes were mated to males heterozygous for the other *spen* allele (e.g. $spen^{poc361}/ovo^{D1}$ females were mated to $spen^{poc231}$/CyO,*wg* males).

Identification of P-element insertions and plasmid rescue

Genomic DNA from the *PZ1295* line corresponding to *l(2)03350* was prepared by standard protocols as described in Gellon et al. (1997). Rescue of the DNA surrounding the P-element insertion site was accomplished by digestion, ligation and transformation of the genomic DNA as described in Gellon et al. (1997).

Excision lines were generated from *PZ1295* by using the Δ2-3 strain to mobilize the element (Robertson et al., 1988). Genomic DNA was prepared from homozygous viable and homozygous lethal excision lines and mapped by Southern analysis using probes surrounding the region obtained by plasmid rescue. A similar analysis

was subsequently done to map the P insertion sites in the *l(2)K06805*, *l(2)K07612*, *l(2)K07721*, *l(2)K08102*, *l(2)K10325* and *l(2)K13625* alleles.

Isolation of genomic clones and cDNAs

We initiated a walk in the region surrounding the insertion point of PZ1295 using a 3.2 kb *Hin*dIII/*Xba*I genomic fragment obtained by plasmid rescue. Approximately 50 kb of overlapping λ clones were obtained from an EMBL-3 *iso-1* genomic library provided by J. Tamkun using standard protocols and nick translated probes. An additional 30 kb proximal to our walk was kindly provided by M. Noll (λ clones y17-3, y16-1, y15-1).

Poly(A)$^+$ RNA was prepared from 0- to 2-hour, 2- to 8-hour, 8 to 12-hour and 12- to 24-hour embryos, purified once over an oligo-dT column, electrophoresed in a formaldehyde gel and blotted according to standard procedures. Northern blots were hybridized with ^{32}P-labeled DNA probes.

cDNAs were isolated from an oligo-T primed 8- to 12-hour embryonic cDNA library provided by Nick Brown (Brown and Kafatos, 1988), and a randomly primed 0- to 18-hour embryonic λGT10 cDNA library (Clontech). Probes were labeled by nick translation. Hybridization and washes were done at high stringency following standard protocols. More than 100 overlapping partial cDNAs corresponding to the putative *spen* transcript were isolated using genomic and cDNA fragments as ^{32}P-labeled probes from the region shown in Fig. 2. Thirty cDNA clones were subcloned and mapped.

A short region (150 bp) within the 4.8 kb *Bam*HI-*Eco*RI fragment that spans most of the 3′ half of the large exon of the putative *spen* transcript sequence was not included in any of the cDNA clones. RT-PCR was used to determine that this sequence is included in *spen* transcripts. Poly(A)$^+$ RNA, isolated as described above, was used as a template for reverse transcriptase following the protocols included with the enzyme, Superscript II Reverse Transcriptase (GibcoBRL). The reaction was primed with a sequence-specific 3′ primer. PCR was performed using the 1st strand reaction as template and following the protocol included with the RT enzyme (GibcoBRL).

Sequencing and database searches

Sequencing was performed mainly at the Center for AIDS Research CORE Facility at UCSD. The BLAST algorithm (Altschul et al., 1997) was used to search protein and EST databases for sequences homologous to short fragments of protein sequence. The 3 RNP-type RNA binding domains were identified in this manner. Comparison of the Spen protein sequences (accession no. AF188205) to the ISREC PROSITE database identified five bipartite nuclear localization motifs. Searches of genomic and EST databases revealed cDNA and genomic sequences which are predicted to encode the protein domains shown in Fig. 4; from *D. melanogaster* (accession no. for short form – AAD38639), *Homo sapiens* Spen-like (accession no. for long form – AL034555, for short form – EST sequence AI621042, and genomic L13434), *H. sapiens* RNP2 from U1a protein (accession no. 2554638), *Caenorhabditis elegans* (accession no. for long form – CAA91320; for short form – AAC19192) and *Nicotiana tabacum* (accession no. U90212), *Xenopus laevis* (accession no. P20397).

RNA interference

RNAi was performed essentially as described by Kennerdell and Carthew (1998). A 972 bp DNA fragment from the region that encodes the 3′ UTR of the *spen* transcript (from 11 nucleotides to 983 nucleotides downstream of the Spen TAA stop codon) was transcribed with T7 RNA polymerase to generate RNA copies of both strands. The resulting double-stranded RNA was injected into the middle region of embryos less than one hour after egg deposition. Controls consisted of injected buffer at the same site and injection of *cap'n'collar B (cncB)* RNA. *cncB* loss-of-function mutants have an obvious homeotic phenotype in embryos (McGinnis et al., 1998) that is mimicked by injection of ds cncB-specific RNA (A.Veraksa and W. M., unpublished results). Injected embryos were aged and their cuticles prepared after removing the vitelline membranes by hand.

Mutation detection

Point mutations in *spen* open reading frames were detected by denaturing gel electrophoresis of 400-500 base pair fragments of genomic DNA isolated from *spen* heterozygotes and parental strains. Genomic DNA was purified using the QIAamp Tissue Kit (QIAGEN), and PCR was performed using primers designed to produce fragments with the melting behavior described by Myers et al. (1987). Melting point variations resulting from point mutations in the fragments were detected using a Denaturing Gel Electrophoresis System (BioRad). Fragments which showed melting point variation were sequenced on both strands to determine the specific nucleotide position that was changed in mutant chromosomes.

Antibodies and in situ hybridization

A GST-Spen fusion protein containing amino acids 3203-3714 from the common region of the SpenL protein was injected into mice to generate anti-Spen antiserum. The fusion protein was bacterially expressed and purified using glutathione-Sepharose (Pharmacia) and standard techniques. For staining with this antibody, embryos were collected, dechorionated and fixed for 15 minutes in 4% formaldehyde. The antiserum was used at a 1:100 dilution, and detected with FITC-labeled anti-mouse secondary antibody (Jackson Labs).

In situ hybridizations to detect transcript expression were performed using a variation of the protocol described by Tautz and Pfeifle (1989). DIG-labeled antisense probes were made from the following cDNA sequences: *Dfd* probe was generated from a 2.3 kb *Eco*RV-*Sma*I fragment isolated from pcDfd41 (Regulski et al., 1985), cloned into the *Eco*RV site of pBluescript, linearized with *Bam*HI and transcribed from the T3 promoter. *Scr* probe was generated from a 2.1 kb *Eco*RI-*Xba*I genomic fragment containing the homeobox (Kuroiwa et al., 1985), subcloned into pGEM2, linearized with *Eco*RI and transcribed from the T7 promoter. *Antp* probe was generated from a 2.2 kb *Bam*HI-*Eco*RI fragments from p903 (Hafen et al., 1983) subcloned into pBluescript, linearized with *Bam*HI and transcribed from the T7 promoter. *spen* antisense transcript probes were generated from a 5.2 kb cDNA fragment in clone pc1.65BH10-5.2R that included the RNP coding region subcloned into pBluescript, linearized with *Not*I and transcribed from the T3 promoter.

RESULTS

spen mutants develop head-like skeletal features in the trunk

We previously reported the isolation of six EMS-induced mutant alleles of *spen* based on its interaction with the Hox gene *Dfd* (*spen* was called *polycephalon* by Gellon et al., 1997). Some of these alleles behave as genetic nulls for the cuticular phenotype, i.e. *spen*^{361}/*spen*^{361} cuticles are indistinguishable from *spen*^{361}/*Df(2L)TE21A*(deletes the entire spen locus), and *spen*^{231} is slightly weaker than *spen*^{361}. Loss of zygotic *spen* function results in embryonic lethality that is associated with a loss of the anterior portion of the H-piece, and with a kinked median tooth (Fig. 1B,F, compare to Fig. 1A). The anterior parts of the H-piece are derived from the ventral maxillary segment, and are dependent on *Dfd* function (Merrill et al., 1987; Regulski et al., 1987). Head development is profoundly disrupted in maternal/zygotic *spen* mutants. These mutant embryos have non-involuted heads that are missing many head sclerites, including the base of the mouth

hooks, median tooth, anterior regions of the lateralgräten, and the dorsal bridge. The ventral arms and vertical plates are also strongly reduced (Fig. 1D,F). Many of these *spen*-dependent sclerites are also dependent on head Hox genes such as *labial (lab)*, and *Sex combs reduced (Scr)*, gap/homeotic genes such as *empty spiracles (ems)* and *orthodenticle (otd)*, and other head patterning genes (Merrill et al., 1989; Sato et al., 1985; Pattatucci et al., 1991; Dalton et al., 1989; Finkelstein and Perrimon, 1991). In the maternal/zygotic mutants, the trunk region of the embryos appears to be largely unaffected. The denticle belts and posterior structures such as filzkörper appear normally shaped, although these structures, as well as the remnants of head skeleton (and ecopic head-like sclerites, see below), are somewhat less pigmented/sclerotized than normal. The overall low level of sclerotization may be due to embryonic death in maternal/zygotic mutants before the cuticle is fully developed (Fig. 1D).

Approximately 50% of the embryos that are zygotic or maternal/zygotic *spen* mutants also develop sclerites in the thoracic segments. These sclerites appear most frequently in the second and third thoracic segments (T2 and T3), and overlap the anterior/posterior compartment boundary. The expressivity of this phenotype ranges widely. The weakest phenotypes exhibit a row of small chunks of ectopic sclerotic material, usually limited to T2 and T3 (Fig. 1B). Strong phenotypes show broad bands of sclerotic material in all thoracic segments, often accompanied by sclerotic patches in lateral regions of the abdominal segments (Fig. 1C). No matter how extensive, the sclerites do not form in the fields of denticle belts, nor do they change the overall size of a segment. In small patches, the ectopic material is variably sclerotized, scar-like, and not recognizable as any other embryonic structure. When manifest in broad patches, the sclerotic material is brown and striated, reminiscent of the brown, striated appearance of the ventral arms, vertical plates and dorsal arms of the head skeleton (Fig. 1E). We conclude that one role of *spen* is to suppress the production of head-type sclerotization from ventral thoracic cells, and it is needed to a lesser extent for this function in lateral abdominal cells.

This *spen*$^{-}$ homeotic phenotype is not specific to alleles isolated on the basis of *Dfd* interaction. In a screen for suppressors and enhancers of an activated Raf construct, Dickson et al. (1996) isolated three alleles of *E(Raf)2A*. These alleles do not complement the lethality of our *spen* alleles, and we find that the *E(Raf)2A* alleles also exhibit ectopic head-like sclerites, as do 11 additional alleles of *spen* that enhance the rough eye phenotype generated by ectopic expression of Cyclin E in the eye (Christian Lehner, personal communication).

Production of ectopic head-like sclerotic material in the trunk indicates that loss of *spen* function might result in de-repression of head patterning genes. We examined the cuticular phenotypes of embryos doubly mutant for *spen* and *Distal-less* (*Dll*), *cnc*, *lab*, *Dfd*, *buttonhead*, *otd* and *ems*, all of which have been shown to be required for determining head-specific pathways (Mohler et al., 1995; Cohen and Jurgens, 1990; Finkelstein and Perrimon, 1991). All of the double mutant combinations show the same degree of thoracic sclerotization as *spen* single mutants. We also examined the transcription patterns of *Dll*, *cnc*, *lab*, *Dfd*, *ems*, *otd* and *apontic* in *spen* mutant embryos

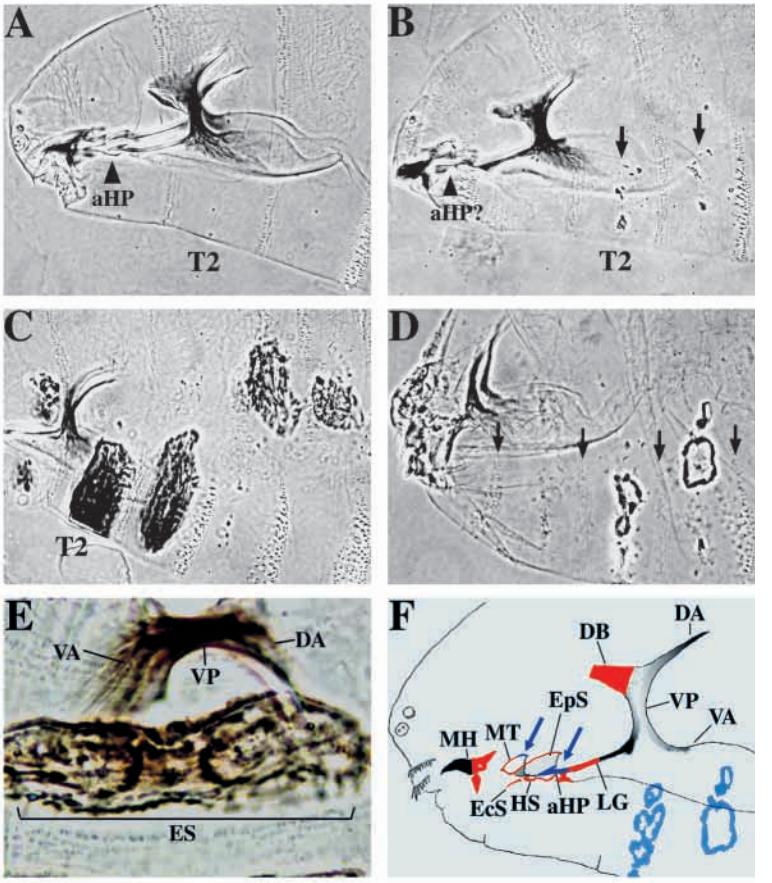

Fig. 1. *spen* is required for thorax and head development. (A) Wild-type cuticle showing normal head skeleton and thoracic segments. T2, second thoracic segment. Anterior is left and dorsal is up in this and all figures, except as noted. (B) *spen*^{poc361}/*Df(2L)TE21A* mutant phenotype with weak expressivity. The anterior H-piece is lost (arrowhead), and ectopic sclerites develop in the middle of T2 and T3 (arrows). (C) *spen*^{poc361}/*Df(2L)TE21A* mutant phenotype with strong expressivity. The ectopic sclerites cover more of the naked epidermis and are found in more segments, including T1, A1 and A2. Abdominal segments develop ectopic sclerites in lateral regions. (D) A maternal/zygotic *spen* mutant in which the germ line was homozygous for *spen*^{poc361}, half of these maternal mutant embryos then received the *spen*^{poc231} mutant allele. The maternal/zygotic mutant heads are severely disrupted, with posterior mouth hooks (MH), median tooth (MT), hypostomal (HS), ectostomal (EcS), and epistomal sclerites (EpS), dorsal bridge (DB) and lateralgräten (LG) either missing or disrupted (see F). The denticle belts (arrows) are slightly less sclerotized than normal. (E) In this close up of the mid thoracic cuticule of a *spen*^{poc361}/*Df(2L)TE21A* mutant, anterior is at the top. When present in large patches as seen here, the ectopic sclerites (ES) in *spen* mutants consist of brown, striated, thick sclerotic cuticle that is similar in color and texture to head skeleton. This can be seen by comparing the ectopic sclerites to the normal vertical plates (VP), ventral arms (VA) and dorsal arms (DA), which can be seen underlying the ectopic sclerites. (F) Schematic diagram of *spen* functional requirements in head and thorax. Blue structures are defective in zygotic *spen* mutants, while red components are only defective in embryos lacking both the maternal and zygotic *spen* function.

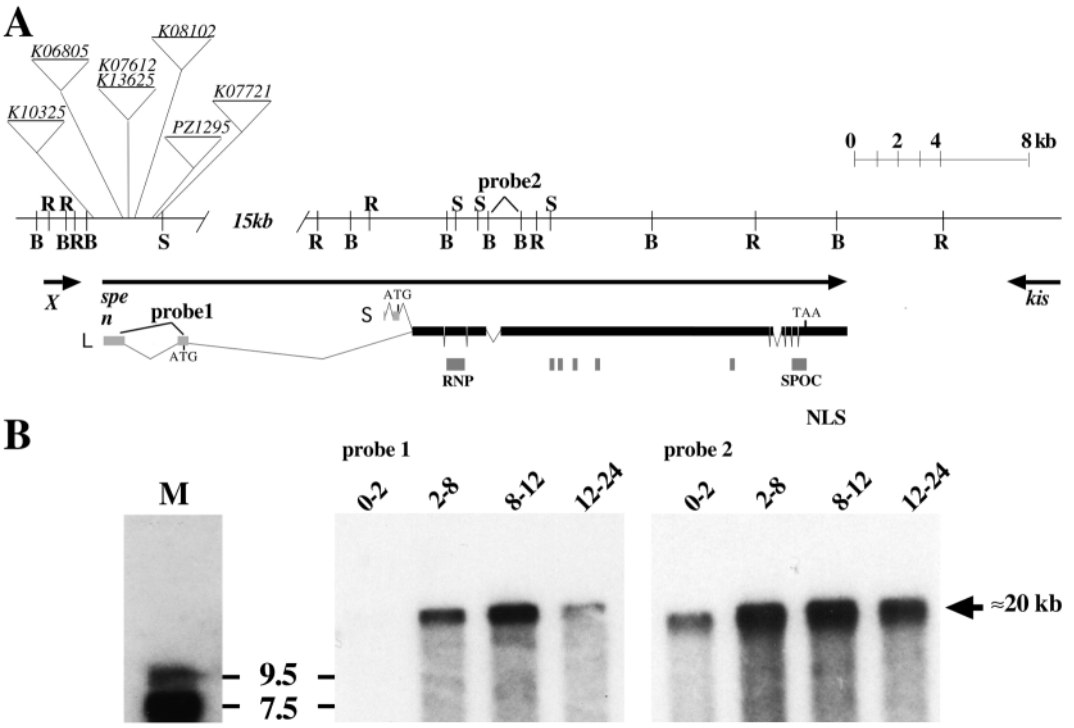

Fig. 2. (A) A molecular map of the *spen* locus. The top, thin line represents genomic DNA; the restriction sites are marked, *Bam*HI (B), *Eco*RI (R), *Sal*I (S). A size marker of 8 kb is shown in upper right. The location of P-inserts which fail to complement *spen* alleles are indicated with open triangles. The locations of transcripts within this region are shown, with arrows indicating extent and orientation of each. The intron/exon structures of the $spen^{L}$ and $spen^{S}$ transcription units are shown below the transcript arrow. Alternative 5′ *spen* exons are shown in light gray, and these include translation start codons and short ORFs that join the long common ORF. Common exons are shown in black. The stop codon for the predicted protein is indicated (TAA), and the positions of domains conserved with human and *C. elegans* homologs are shown below the sequence in dark gray (see text and Fig. 4 for descriptions of labels). (B) Northern blots of poly(A)$^{+}$ embryonic RNA from selected ages (0-2 hours, 2-8 hours, 8-12 hours, 12-24 hours) probed with two DNA fragments (shown above the cDNA and genomic structures by 1 and 2). Probe 1 was generated from $spen^{L}$-specific cDNA sequences, and hybridizes to a single band of approx. 20 kb (M, size marker lane). The $spen^{L}$ transcript is not detected in 0-2 hour embryos when only maternal transcripts are present. The $spen^{L}$ transcript levels are abundant at 2-8 hours and 8-12 hours of development, and less abundant at 12-24 hours. Probe 2 was generated from a genomic fragment in the common region, and detects a broader band of approx. 20 kb in 2-24 hour embryos. The common probe detects *spen* transcripts in 0- to 2-hour embryos, presumably due to maternal transcription from the $spen^{S}$ promoter.

and found them to be indistinguishable from wild type (data not shown). Therefore we conclude that *spen* does not repress ectopic head-like cuticle by repressing the expression or function of these known head-determining genes. Finally, ectopic expression of *Dfd* and *lab* in *spen* mutant embryos did not alter the degree or character of the ectopic sclerites in the thorax (data not shown).

Identification and sequence of *spen* transcripts

To understand how mutations in *spen* produce ectopic sclerites, we cloned the gene which encodes the *spen* function. All six of the EMS-induced *spen* mutant alleles recovered from the *Dfd* modifier screen (Gellon et al., 1997) fail to complement the lethality of the P-element insertion PZ1295 (*l(2)03350*), located within cytological interval 21B4-6. Precise excisions of the PZ1295 insertion are viable over all six *spen* alleles, whereas imprecise excision lines are still lethal over *spen* mutant chromosomes. These results suggest that PZ1295 interrupts a sequence crucial to the function of the *spen* locus. Genomic DNA adjacent to the site of PZ1295 insertion was rescued and overlapping λ clones spanning approximately 50 kb of surrounding genomic DNA were isolated. In addition, genomic clones covering the 3′ end of what turned out to be the *spen* transcription unit were kindly provided by Markus Noll. A structural map of the locus is shown in Fig. 2A. Northern analysis with genomic probes detected two transcripts from the region, one of 1.3 kb (*X* in Fig. 2A) and another of 20 kb (*spen* in Fig. 2A). The 3′ end of *kismet* (*kis*) is located 7.5 kb from the 3′ end of *spen* (Fig. 2A; Daubresse et al., 1999).

Five of the $spen^{-}$ P-element insertions map either in the first exon or the first intron of the long isoform of the 20 kb transcript, while one maps immediately 5′ of this transcript (Fig. 2A). Evidence that the 20 kb transcript corresponds to the *spen* locus is based on: (a) reversion of the mutant phenotype upon excision of the P element inserts; (b) identification of amino acid changes in two EMS-induced *spen* alleles, both isolated as enhancers of CycE (Christian Lehner, personal communication; see Fig. 4); and (c) the generation of $spen^{-}$ head defects and ectopic head-like sclerites upon injection of double stranded RNA specific to the 20 kb transcript sequence (data not shown).

Using genomic DNA fragments from these regions, approximately 200 cDNA clones were isolated from a poly-T primed, 8-12 hour embryonic cDNA library (Brown and Kafatos, 1988), and from a randomly primed, 0-18 hour embryonic cDNA library. Overlapping cDNA clones were mapped and sequenced, resulting in the exon/intron structure shown in Fig. 2A. The structure of the cDNA clones that we analyzed suggests that at least two alternative promoters initiate transcripts. A probe made from the most 5′ exons, which initiate the long ($spen^{L}$) transcript isoform does not detect maternal transcripts in the 0-2 hour RNA lane of a northern blot (Fig. 2B, probe 1), whereas a probe from the common region does (Fig. 2B, probe 2). However, RT-PCR

analysis (not shown) does detect *spen*^{L} and *spen*^{S} transcripts at all stages, including 0-2 hours, indicating that low levels of the *spen*^{L} isoform exist in maternal RNA. Both probes detect transcripts of approximately 20 kb at later stages of embryonic development, although the *spen*^{L} isoform is predicted to be approximately 1 kb larger based on the known cDNA structures. A third Spen protein isoform has been characterized by B. Kuang, S. Wu, Y. Shin, K. A. Mace, L. Luo and P. Kolodziej (personal communication), and is initiated from an alternative exon that maps upstream of the *spen*^{S} 5′ exon. The third isoform has a few unique amino acid residues at its N terminus that differ from two Spen protein isoforms reported here.

Spen proteins contain RNP motifs and a novel conserved C-terminal domain

The first ATG in the long open reading frame of the *spen*^{L} transcript (Fig. 2) begins a long open reading frame of 5533 codons, which is predicted to make a 597 kDa protein (Fig. 3). The first ATG in the *spen*^{S} transcript begins an open reading frame of 5476 codons. The SpenS protein isoform is essentially an N-terminal truncation of the SpenL protein with the addition of a Met Arg dipeptide, encoded in S exon 2, substituted for the first 59 aa residues of SpenL (Fig. 3). The N terminus of the SpenL isoform may contain unique functions due to its poly-Asn and poly-Gln sequences (Fig. 3). Three RNA-binding domains of the RNP type (amino acids 554 – 806 of L) are present in the 5474 aa common region of the predicted SpenL and SpenS proteins. Five bipartite nuclear localization sequences are scattered throughout the middle of the common sequence (positions 1872-1889, 1949-1966, 2135-2152, 2450-2467, 4556-4573 of SpenL). In addition, a region that is highly enriched in Glu (E) and Lys (K) residues is present from positions 1917-2053 of SpenL, and strings of glutamine (Q)

Fig. 3. The predicted amino acid sequences of Spen proteins. The transcripts that include the *spen*^{L}-specific exons are predicted to encode a longer protein (SpenL), the sequence of which is shown. Two putative translation start sites, α and β, are shown; α is the first initiation codon in the transcript, while β has a context closer to the consensus. The SpenS protein is predicted to contain only MetArg (MR) residues as unique N-terminal amino acids before joining the common sequence at residue 60. Protein domain predictions include three RNP domains within residues 553-806 (highlighted in black), five bipartite nuclear localization sequences (18-amino-acid blocks highlighted in black between residues 1800-4600), and a 165 residue Spen paralog and ortholog C-terminal domain, also highlighted in black, at the C terminus.

residues are found throughout the C-terminal 1/3rd of the Spen proteins (Fig. 4A).

Database searches combined with GENESCAN and FGENE predictions of exon structure reveal apparent structural orthologs of Spen in genome sequences from human and *C. elegans*. These Spen family members are defined by their large size (predicted to be at least 3300 aa for human and 2738 aa for *C. elegans*), by their possession of three RNP domains that are closely related to those in Spen isoforms, and by a lengthy sequence match to a 165 amino acid motif at the C terminus of the *D. melanogaster* Spen proteins (Fig. 4A). The sequence of the first RNP domain of the Spen family (RNPa in Fig. 4B) also is closely related to the first RNP domain of Nucleolin, a protein involved in ribosome biosynthesis (Ginisty et al., 1999). The second Spen family RNP domain (RNPb in Fig. 4B) is a good match to the canonical RNP domain (Fig. 4B, bottom line), and shares even more identities with an RNP domain in AC binding factor (ACBF), an RNP-motif protein isolated from tobacco. ACBF binds to DNA regulatory sequences of genes in the phenylpropanoid biosynthetic pathways of many plants (Séguin et al., 1997).

The C-terminal motif, which we call the SPOC domain (Spen Paralog and Ortholog C-terminal domain) is very similar among family members (57% identity between *D. melanogaster* Spen and *H. sapiens* Spen1; Fig. 4C), but matches no other sequence motifs with known biochemical function. The entire open reading frames of two EMS-induced *spen* alleles ($spen^{E(CycE++)D57}$ and $spen^{E(CycE++)e9}$) were sequenced. Single base substitutions in each of these alter the predicted amino acid sequence at highly conserved positions within the SPOC domain: codon position 5444 of $spen^{E(CycE++)e9}$, TGC – Cys to TAC – Tyr; and codon position 5480 of $spen^{E(CycE++)D57}$, GGC – Gly to GAC – Asp (Fig. 4C). With regard to the embryonic head and thoracic defects observed in $spen^{361}$ zygotic mutants, both of these alleles are hypomorphs. The order of allele strengths with respect to zygotic defects is $spen^{361}$ > $spen^{231}$ > $spen^{E(CycE++)e9}$ > $spen^{E(CycE++)D57}$.

Our sequence similarity searches also reveal a separate subfamily of short Spen-like proteins encoded in *D. melanogaster* and other animal genomes. The short Spen-like proteins have a similar arrangement of Spen-like RNP motifs and SPOC domains to those in the Spen orthologs, but the short Spen-like proteins are approximately one tenth the size of Spen (Fig. 4). All of the known short Spen-like proteins also encode an RGG motif, which is capable of destablizing RNA helices and often found in combination with RNP motifs (Siomi and Dreyfuss, 1997). The current sequence evidence suggests that a duplication event giving rise to the genes that encode the large and small Spen-like proteins occurred before the divergence of the deuterostome and protostome lineages in metazoan evolution (Aguinaldo et al., 1997).

Spen proteins are ubiquitously expressed and localized in nuclei

Antibodies were raised against a common fragment of the Spen proteins (amino acids 3203-3714) to assay the expression pattern and subcellular localization. Immunostaining of whole-mount embryos shows that Spen antigen is expressed in most or all cell types, and is concentrated in nuclei (Fig. 5A-C). Spen protein staining is first detectable prior to cellularization in stage 3 embryonic nuclei, and is present in all blastoderm cells, including the pole cells (Fig. 5A). The protein is expressed throughout the rest of embryogenesis, and is concentrated in nuclei. At later embryonic stages (9 through 14), nuclear Spen staining appears to be most abundant in CNS and epidermal cells (Fig. 5C), but most tissues exhibit nuclear staining at detectable levels.

spen cooperates with *Antp* and *tsh* to repress head-like sclerites in the thorax

Null mutations in *Antp* result in a transformation of T2 and T3 towards T1 in the embryonic body plan (Wakimoto and Kaufman, 1981; Martinez-Arias, 1986). In addition, *Antp* mutant embryos develop ectopic head-like sclerites in the dorsal thorax (between T1 and T2; Fig. 6A), similar in kind but not in position to the ectopic sclerite phenotype seen in *spen* mutants. To test whether *spen* and *Antp* function in an additive or synergistic manner in the repression of head-like sclerites in the thorax, we first studied $spen^{-}$; $Antp^{-}$ cuticle phenotypes. Embryos mutant for both *spen* and *Antp* have more sclerotic material in dorsal T2 than do *Antp* mutants alone (Fig. 6B). In addition, the ectopic head-like sclerites in the ventral thorax of $spen^{-}$; $Antp^{-}$ mutants are more sclerotized and extensive than in *spen* mutants alone (Fig. 6D, compare to 6E). The sclerotic material in $spen^{-}$; $Antp^{-}$ mutants frequently appears in two distinct bands, one in the center of the segment similar to the position in *spen* mutants, and at another position in the posterior of T1 and T2 (Fig. 6D, arrow). These posterior ectopic sclerites do not develop in T3 (Fig. 6D). The enhanced formation of head-like sclerites in $spen^{-}$; $Antp^{-}$ mutants suggests that *spen* and *Antp* function in a common or interacting pathway(s) in subregions of T1 and T2.

The synergistic effect of *Antp* and *spen* might be due to a regulatory effect of Antp on *spen* transcription pattern, or to Spen effects on *Antp* transcript pattern or translation. However, *Antp* transcript and protein expression patterns are unchanged in *spen* mutant embryos, and *spen* transcript expression is unchanged in *Antp* mutant embryos (data not shown). Therefore, *spen* and *Antp* appear to be acting in parallel, presumably due to direct or indirect regulation of common downstream genes.

If *spen* and *Antp* regulate common targets, then induction of high levels of exogenous *Antp* expression might result in suppression of the *spen* mutant phentoype. We therefore tested the ability of excess Antp protein to suppress the *spen* mutant phenotype. Overexpression of Antp under heat shock promoter control (*hsAntp*) causes a transformation of head regions to thoracic identity, but leaves T2 and T3 nearly unchanged (Fig. 6F; Gibson and Gehring, 1988). When Antp is overexpressed in a *spen* mutant background, the ectopic head-like sclerites are strongly suppressed (Fig. 6G). The number of *hsAntp*, $spen^{-}$ embryos which exhibit any detectable ectopic sclerites was less than half the expected number compared to $spen^{-}$ mutant siblings from the same cross, or compared to $spen^{-}$, *hsAntp* embryos that were not subjected to heat shock. In addition, the sclerites which do occasionally appear in heat shocked *hsAntp;* $spen^{-}$ embryos are smaller than those in their $spen^{-}$ siblings (Fig. 6G,H) The ability of excess *Antp* to suppress the $spen^{-}$ homeotic transformation indicates that the two genes interact to repress ectopic head-like sclerites.

In the head region, where *spen* is required for the

development of sclerites that also require *Dfd*, *Scr* and other head genes, it is also possible that *spen* might work in parallel to Hox pathways. This is the case for *Dfd* and *Scr*, since mutations in these genes had no effect on *spen* transcript expression pattern, and conversely, *spen* mutants had no effect on *Dfd* or *Scr* transcript or protein expression patterns (data not shown). We attempted to test whether the overexpression of Dfd could rescue the H-piece defect in *spen* mutant embryos, but the morphology of the head skeleton was so disrupted by heat shock induced ectopic Dfd protein (Kuziora and McGinnis, 1988) that it was not possible to conclude whether the anterior H-piece was restored.

Another *Drosophila* gene involved in distinguishing head from body is *teashirt* (*tsh*). Tsh protein is expressed only in the labial segment and trunk region of embryos, where it is required to repress head identity and to promote thoracic and abdominal segment identities (Fasano et al., 1991). *tsh* transcription levels in the thorax are maintained by *Antp* (Röder et al., 1992; McCormick et al., 1995), but a variety of genetic interaction tests have shown that Antp and Tsh have independent functions in repressing head development (Röder et al., 1992). Is *spen* integrated into the *Antp*; *tsh* pathways by regulation of the *tsh* or *Antp* expression patterns. Our experiments show that: (1) expression pattern of Tsh protein

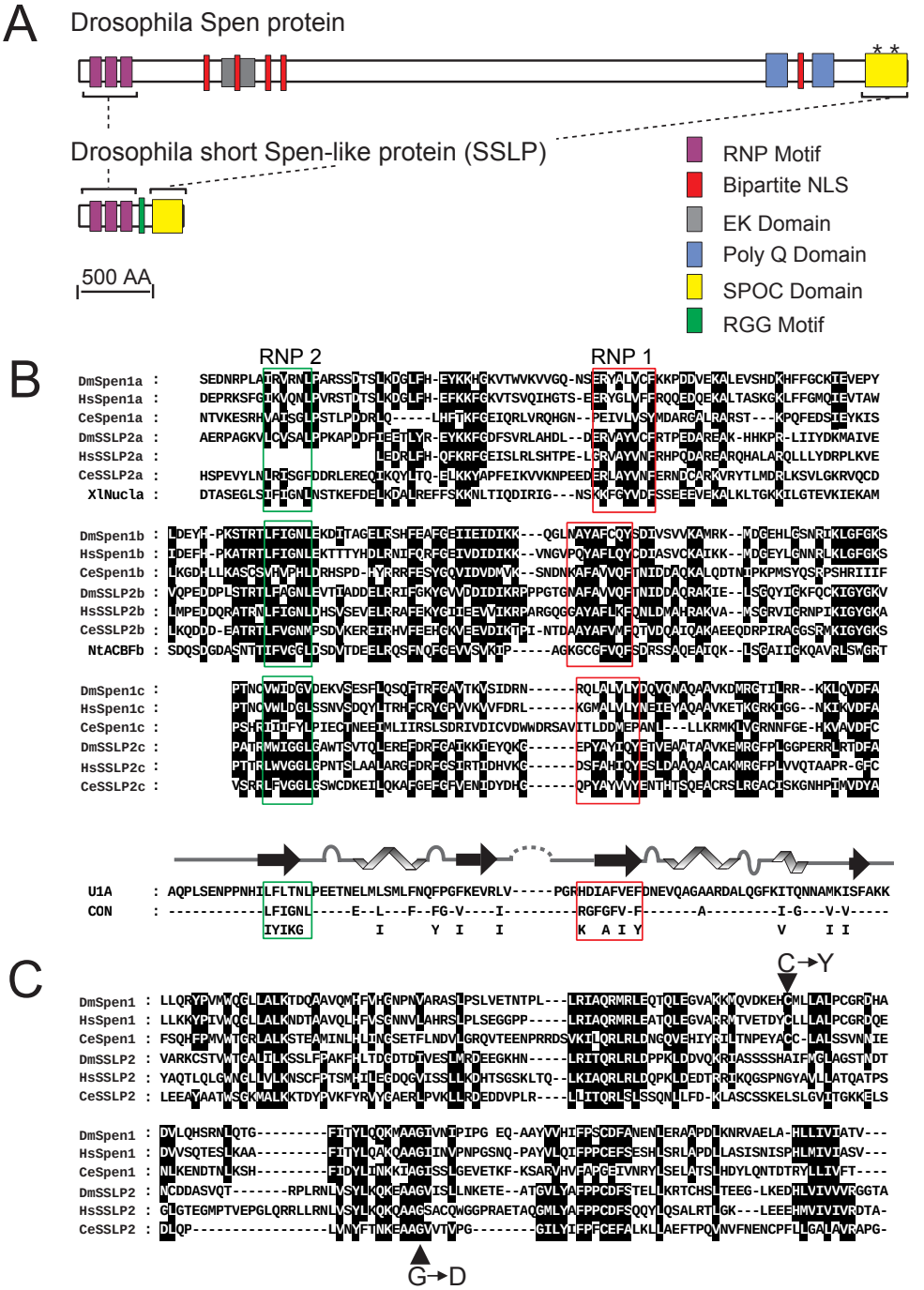

Fig. 4. Spen and the family of predicted Spen-like proteins. (A) The Spen orthologs from human and worm are all larger than 2738 amino acids, and each contain three RNP domains, bipartite nuclear localization sequences (NLS), and a novel, conserved domain, the Spen Paralog and Ortholog C-terminal domain (SPOC domain). Additional similarity among Spen orthologs consists of an acidic/basic domain (EK) and glutamine-rich regions (Poly Q). An additional group of proteins, the Short Spen Like Proteins (SSLPs) are encoded in fly, human and nematode genomes (Materials and Methods). These SSLPs are all smaller than 800 amino acids and possess Spen-like RNP domains, a SPOC domain, and an RGG motif. (B) Amino acid alignment of the three RNP domains from the six Spen family proteins, along with similar RNPs from other families. The first RNP (RNPa) is more similar to the first RNP of Nucleolin than to other RNP domains. The RNPb domain has more identity to an RNP domain of AC Binding Factor (ACBF) than to other RNP domains. ACBF is a tobacco protein which binds to double stranded, AC-rich DNA sequences (Séguin et al., 1997). The schematic structure of an RNP domain from human snRNP U1A is shown at the bottom to illustrate the RNP consensus and structure. The RNP2 (green) and RNP1 (red) sequences are boxed in all sequences. Hs, *Homo sapiens*; Dm, *D. melanogaster*; Ce, *C. elegans*; Xl, *Xenopus laevis*, Nt, *Nicotiana tabacum*. Accession numbers are provided in the Materials and Methods. (C). Amino acid alignment of the SPOC domains from fly, human and nematode. A nucleotide substitution (TGC to TAC) in spen$^{E(CycE)e9}$changes the conserved cysteine (C) to tyrosine (Y) at amino acid 5444, while a substitution (GGC to GAC) in spen$^{E(CycE)D57}$ changes the invariant glycine (G) to aspartate (D) at amino acid 5480.

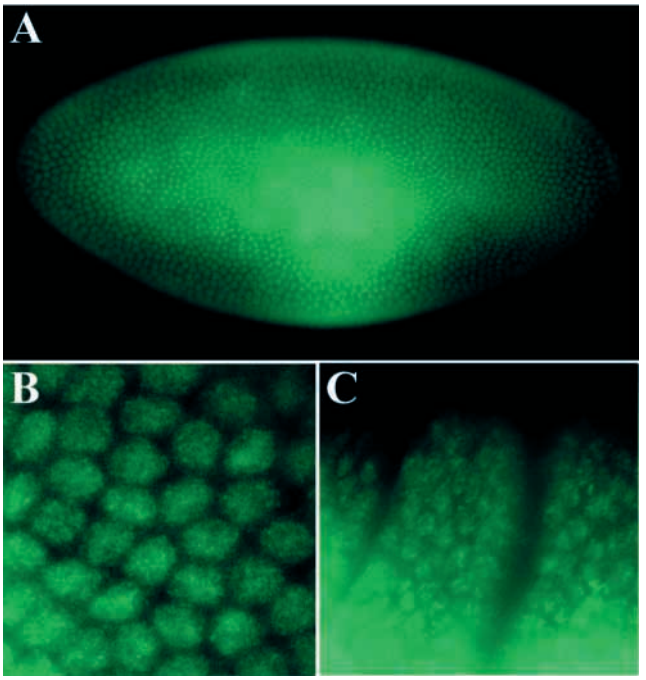

Fig. 5. Spen protein antigen is concentrated in nuclei of most embryonic cells. (A) Stage 5 embryo showing anti-Spen staining in all nuclei. The antibodies used were raised against the protein fragment 3203-3714, which is common to both SpenL and SpenS proteins. (B) Close-up of embryo in A. (C) Dorsolateral cells of the second and third thoracic segments in a stage 13 embryo; Spen protein staining is concentrated in nuclei.

(Röder et al., 1992) is unchanged in *spen* mutant embryos, (2) protein expression patterns of *Scr* or *Antp* in *spen*$^{-}$, *tsh*$^{-}$ double mutant embryos are unchanged from the pattern seen in *tsh* mutants alone (data not shown), and (3) the *spen* transcript pattern is normal in *tsh* mutants (data not shown). Therefore, Spen suppression of head-like sclerites is not exerted by a regulatory effect on the Tsh protein expression pattern, nor by Tsh effects on the *spen* transcript pattern, nor through combinatorial effects of *spen* and *tsh* on *Scr* or *Antp* protein abundance.

The phenotype of *tsh*$^{-}$, *spen*$^{-}$ mutant embryos suggests that the two genes act to promote thoracic development. In *tsh* mutant embryos, the T1 denticle belt is absent and although the remaining denticle belts appear to have the appropriate segmental identities, the denticles themselves are disorganized and smaller than in wild type (Fig. 7A; Fasano et al., 1991; Röder et al., 1992). In contrast, *tsh*$^{-}$, *spen*$^{-}$ double mutants completely lack denticle belts in the thorax (Fig. 7B). This may be due to the the death of cells in the denticle field in the thorax of the double mutants, or to the inability of Antp protein, still expressed in the remaining cells, to promote the development of thorax-specific structures.

As to whether *tsh* and *spen* collaborate in repressing head-like sclerites, we find that the *tsh*$^{-}$, *spen*$^{-}$ double mutants still have bits of sclerite in the 'thorax' of the double mutants, so this phenotype is not enhanced. However,we also examined the effects of Tsh overexpression on the ectopic head-like sclerites in *spen* mutants. In wild-type embryos, overexpression of Tsh protein throughout the embryo results in transformation of head regions toward thoracic identity (de Zulueta et al., 1994), as well as poorly differentiated denticle belts, especially in the thorax. In the thorax of *spen* mutant embryos that also overexpress Tsh, the ectopic ventral head-like sclerites are strongly suppressed (Fig. 7D). Taken together, these results suggest that *spen*, *tsh* and *Antp* function in a combinatorial manner to repress the development of head-like sclerites and promote the development of thoracic identity.

DISCUSSION

Spen function in Hox, Tsh, and Raf pathways

Encoding RNA binding domains, *spen* is a unique type of homeotic function. It is neither in the Hox class, nor is it required for the proper spatial regulation of Hox transcription, such as trx or Pc group homeotics. It falls into the growing class of functions that act in parallel to modulate the output of Hox pathways. The direction of spen's modulatory effect, either helping to promote or to repress head-type sclerotization, is apparently dictated by the type of Hox gene expressed in a certain region of the body.

Although expressed throughout the embryo, *spen* appears to be especially important to the morphogenesis of head cuticular structures, presumably due to modulation of its function on the anterior/posterior axis by specific Hox proteins and specific signaling pathways. The epistasis experiments reported here, along with previous results, indicate that in its role of promoting sclerite formation in the head, *spen* acts in parallel to head patterning genes such as *Dfd* and *ems* (Fig. 8). There are also other embryonic head structures that are missing or abnormal in both maternal/zygotic *spen* mutants and other head homeotic mutants such as *lab* and *Scr* (Merrill et al., 1989; Sato et al., 1985; Pattatucci et al., 1991). The homeotic function of *spen* is exerted in the trunk, where it represses head-like sclerite formation in combination with both the segment identity homeotic *Antp* and the regional homeotic gene *tsh* (Fig. 8).

The thorax-promoting *spen* function also genetically interacts with both *tsh* and *Antp*. In double mutants for *spen* and *tsh*, the thorax has few segmental specializations. This is an interesting result, for *Antp* expression is unaffected in the epidermal cells that remain in the thoracic region. Another result indicating a close genetic interaction between *tsh* and *spen* is that overexpression of Tsh protein can suppress the ectopic sclerite phenotype of *spen* mutants. The Spen and Tsh functions, both of which act in parallel to the Hox proteins, seem likely to exert this effect through by their collaboration (direct or indirect) on the regulation of Hox downstream target genes. Presumably this occurs by the Tsh DNA binding protein operating as a transcription factor (Alexandre et al., 1996), and by Spen acting at the level of RNA processing/transport.

Staehling-Hampton et al. (1999) recently discovered that in *Drosophila* eye cells, wild-type *spen* genetically antagonizes ectopically expressed E2F and Dp proteins, and genetically assists p21 in blocking S-phase entry. Thus at least in the eye, *spen* apparently has a role in suppressing progression of the cell cycle. Other evidence consistent with this derives from the study of interactions between *spen* and *Raf* mutations. spen$^{E(Raf)2A}$ alleles were found to enhance a dominant, constitutively active Raf phenotype in the R7 photoreceptors of the *Drosophila* eye (Dickson et al., 1996). Clones of *spen*

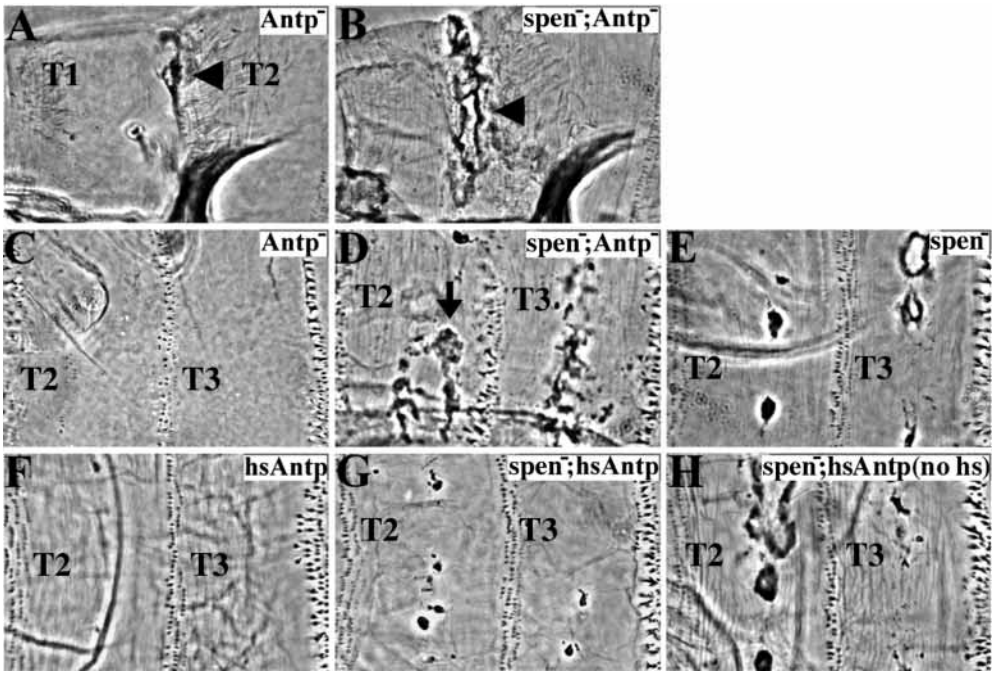

Fig. 6. *Antp* and *spen* cooperatively repress head-like sclerites in thoracic segments. (A) Dorsal head of an *Antp*⁻ embryo showing ectopic sclerotization (arrowhead). B) Dorsal head of a *spen*⁻; *Antp*⁻ embryo, a sibling of the embryo in A. The amount of dorsal ectopic sclerotization is increased (black arrowhead). (C-H) Ventral T2 and T3 embryonic cuticle. (C) *Antp*⁻ control embryo has sclerites in dorsal thoracic cuticle. There is no sclerotization in ventral regions of the thorax. (D) *spen*⁻; *Antp*⁻ embryos have more extensive sclerotization in the thorax than *spen*⁻ siblings, particularly in the posterior compartment (arrow). (E) *spen*⁻ control sibling of the embryo in D. (F) hs*Antp* control embryo treated with heat shock. (G) One of the rare *spen*⁻; hs*Antp* siblings of the embryo in F that shows bits of thoracic sclerotization. (H) *spen*⁻ ; hs*Antp* control embryo that was not treated with heat shock.

mutant cells in the eye paradoxically exhibited either loss of R7 cells, or supernumerary R7 and other photoreceptors (Dickson et al., 1996). Given the interactions and phenotypes they observed, Dickson et al. (1996) concluded that at least in R7 cells, *spen* acts downstream of Raf, as a constitutive repressor of Raf signaling, in a manner similar to the ETS transcription factor Aop/Yan (Nüsslein-Volhard et al., 1984; Lai and Rubin, 1992). It is interesting to note that the maternal/zygotic loss of function mutations reported here for *spen*, and those for null zygotic mutations of *aop/yan* (Rogge et al., 1995) are similar in that thoracic and abdominal features are relatively normal, while head skeletal structures are missing or malformed. So it is possible that Spen may act along with Aop/Yan in the nuclei of embryonic head cells as a repressor of Ras/Raf signaling.

Genetic interactions between Ras signaling and Hox function have been described previously; Boube et al. (1997) found that *proboscipedia* (*pb*) loss-of-function homeotic phenotypes in the *Drosophila* labium were enhanced in *Ras1*⁻*/Ras1*⁺ adults, and *Ultrabithorax (Ubx)* loss-of-function homeotic phenotypes were enhanced when haltere cells contained only one wild-type allele of Gap1 (a Ras inhibitor). In *C. elegans*, the ability of the *lin-39* Hox gene to specify vulval cell fates requires activation of the Ras pathway (Clandinin et al., 1997; Maloof and Kenyon, 1998; Eisenmann et al., 1998). Ras acts both in parallel to *lin-39* as a co-activator of vulval identity, and also amplifies the amount of lin-39 protein expression in vulval precursor cells. Another *C. elegans* Hox gene, *egl-5*, also requires Ras signaling to provide a neuroectoblast cell fate to the P12 cell, which is accomplished at least in part by Ras pathway activation of the transcription of *egl-5* (Jiang and Sternberg, 1998). Thus the impact of Ras/Raf signaling on Hox gene expression and Hox protein function may be commonly used in developmental genetic circuitry. In light of all the genetic interactions of Spen, it is possible that it has a general function in inhibiting cell division and acts in combination with Hox, E2F, and Ras/Raf signaling pathways to accomplish this role. How this putative general growth control function might be connected (if at all)

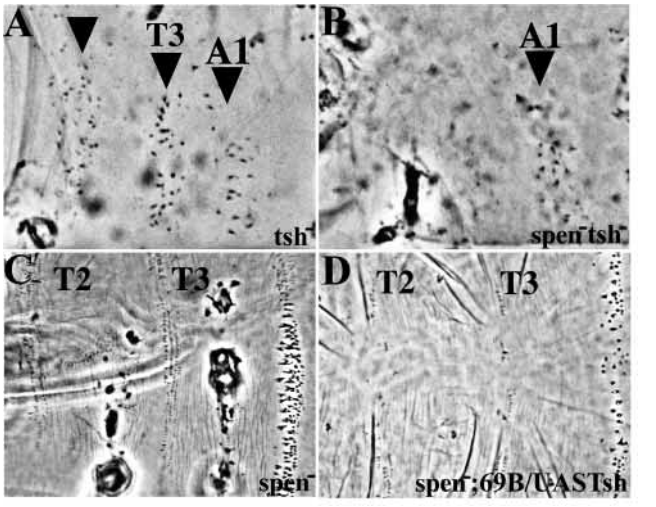

Fig. 7. *teashirt* (*tsh*) and *spen* cooperate in normal thoracic development. (A) Ventral cuticle of a *tsh*⁻ embryo. T1 denticles are absent. T2, T3 and first abdominal (A1) denticle belts are present (arrowheads), but have small denticles in disorganized patterns. The segments are also much narrower than wild type. (B) A comparable region from a *spen*⁻; *tsh*⁻ embryo. Some ectopic sclerotic material can be seen in the thoracic region, but the denticle belts are completely absent, which is not seen in *tsh*⁻ alone. The A1 denticle belt is still barely identifiable (arrowhead). (C) T2 and T3 of a control *spen*⁻ sibling of the embryo shown in (D). D) T2 and T3 segments of a *spen*⁻; UAS*tsh*/69B embryo, in which *tsh* was ubiquitously over-expressed under control of a 69B-GAL4 driver. This completely represses the ectopic sclerites observed in *spen* mutant embryos.

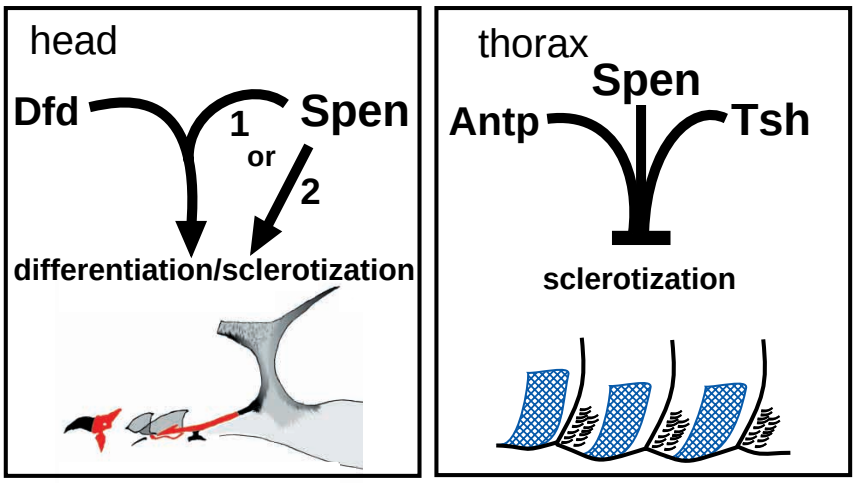

Fig. 8. A summary of Spen functions in the head and thorax. In the head, *spen* is required for production of specific head sclerites, which are also dependent on Hox gene function. Both *Dfd* and *spen* functions are required for the production of the base of the mouth hooks, anterior portion of the H-piece, lateralgräten and ectostomal sclerite (all shown in red). Thus, together *spen* and *Dfd* regulate the formation or sclerotization of head structures, either in a common pathway (1), or independently (2). In regions of ventral thoracic cuticle (region indicated in blue hatching), *spen* cooperates with *Antp* and *tsh* to repress head-like sclerites and promote normal thoracic development.

to the loss of sclerites in the head and the ectopic production of head-like sclerites in the trunk is unknown.

The extent to which Spen connects Ras/Raf pathways and Hox pathways is a fascinating but still largely unexplored area, as are the mechanisms by which Spen fits into these pathways. RNA-binding proteins have been postulated to function in signaling pathways by binding signal-responsive gene transcripts (Siomi and Dreyfuss, 1997). In this postulated mechanism, a transcript might be released for translation in response to activation of a signaling pathway, providing very rapid protein induction; or a different protein isoform might be generated by differential splicing in response to signaling pathways. However, these and other possible mechanisms are speculative until common target genes of Hox and Spen, or Spen and Raf, are identified.

Downstream targets for Hox and Spen

The ectopic head-like sclerites which develop in the thorax of *spen* mutant embryos are darker, more striated, and thicker than normal thoracic body wall cuticle. In addition, *spen* is required for the formation and sclerotization of head structures such as the H-piece and ventral arms. Thus a common theme of *spen* function in the epidermis is regulation of sclerotization. What are possible 'sclerite' target genes that might be activated by Dfd and Spen, and repressed by Antp and Spen?

At the molecular level, the formation of sclerites involves the incorporation of a complex array of substances in addition to chitin. Among these substances are cross-linking agents that generate darkly pigmented, thick, and hard sclerites (Poodry, 1980). Therefore, in thoracic cells, Spen and Antp are directly or indirectly repressing the expression or activities of the enzyme and transport pathways that synthesize and secrete these sclerotization-promoting substances. Genes in the pathways that synthesize some of these substances have been long studied in *Drosophila*, and include *Dopa decarboxylase* (*Ddc*) and *pale* (tyrosine hydroxylase; Wright, 1996; Jurgens et al., 1984; Neckameyer and White, 1993). These enzymes are required for the production of dopamine, which functions both as a neurotransmitter in the fly nervous system, and a precursor of cross-linking molecules required for the sclerotization of cuticle in the body wall. It will be interesting to investigate whether *Ddc*, *pale*, or other genes in the dopamine pathway are targets of the repressive functions of *spen* and *Antp* in thoracic cells.

Spen orthologs and Spen-like proteins

Spen and its homologs contain three clustered RNP-type RNA binding domains (Siomi and Dreyfuss, 1997; Handa et al., 1999). Other proteins which contain RNP domains include Sex lethal (Sxl), U1A snRNP, Nucleolin, and the circadian rhythm regulator Lark (Burd and Dreyfuss, 1994; Serin et al., 1997; Newby and Jackson, 1996). It therefore seems highly likely that Spen family protein functions will be exerted by RNA binding.

We thank T. Kaufman, J. Mahaffey, G.Morata, G. Struhl, S. Kerridge, Christian Lehner, Barry Dickson and Ernst Hafen for providing fly stocks, J. Tamkun and N. Brown for sharing their *Drosophila* libraries. Christian Lehner and Peter Kolodziej generously shared unpublished results. We are also grateful to M. Noll for providing genomic clones in the *spen* region, Eileen Furlong and Matt Scott for anti-Antp and anti-tsh antisera, and J. Tamkun for sharing the northern analysis of the *kis* region prior to publication. Both FlyBase and the Berkeley Drosophila Genome Project were of great value at many stages in the completion of these experiments. This work was supported by NIH grant HD 28315.

Note added in proof

The isolation of additional *spen* mutant alleles, along with the sequence of one of the Spen protein isoforms is also reported by Rebay, I., Chen, F., Hsaio, F., Kolodziej, P. A., Kuang, B. H., Laverty, T., Suh, C., Voas, M., Williams, A. and Rubin, G. M. (1999). A genetic screen for novel components of the Ras/MAPK signaling pathway that interact with the *yan* gene of Drosophila identifies *split ends*, a new RRM motif containing protein. *Genetics*, in press.

REFERENCES

Aguinaldo, A., Turbeville, J., Linford, L., Rivera, M., Garey, J., Raff, R. and Lake, J. (1997). Evidence for a clade of nematodes, arthropods and other moulting animals. *Nature* **387**, 489-493.

Alexandre, E., Graba, Y., Fasano, L., Gallet, A., Perrin, L., De Zulueta, P., Pradel, J., Kerridge, S. and Jacq, B. (1996). The Drosophila teashirt homeotic protein is a DNA-binding protein and *modulo*, a HOM-C regulated modifier of variegation, is a likely candidate for being a direct target gene. *Mech. Dev.* **59**, 191-204.

Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new

generation of protein database search programs. *Nucl. Acids Res.* **25**, 3389-3402.

Biggin, M. D. and McGinnis, W. (1997). Regulation of segmentation and segmental identity by *Drosophila* homeodomain proteins: the role of DNA binding in functional activity and specificity. *Development* **124**, 4425-4433.

Boube, M., Benassayag, C., Seroude, L. and Cribbs, D. L. (1997). Ras1-mediated modulation of Drosophila homeotic function in cell and segment identity. *Genetics* **146**, 619-628.

Brand, A. H. and Perrimon, N. (1993). Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* **118**, 401-415.

Brown, N. H. and Kafatos, F. C. (1988). Functional cDNA libraries from *Drosophila* embryos. *Mol. Biol.* **203**, 425-437.

Burd, C. G. and Dreyfuss, G. (1994). Conserved structures and diversity of RNA-binding proteins. *Science* **265**, 615-621.

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

Capovilla, M. and Botas, J. (1998). Functional dominance among Hox genes: repression dominates activation in the regulation of *Dpp*. *Development* **125**, 4949-4957.

Chan, S. K. and Mann, R. S. (1996). A structural model for a HOX-extradenticle-DNA complex accounts for the choice of the HOX protein in the heterodimer. *Proc. Natl. Acad. Sci. USA* **93**, 5223-5228.

Chan, S. K., Popperl, H., Krumlauf, R. and Mann, R. S. (1996). An extradenticle-induced conformational change in a HOX protein overcomes an inhibitory function of the conserved hexapeptide motif. *EMBO J.* **15**, 2476-2487.

Chan, S. K., Ryoo, H. D., Gould, A., Krumlauf, R. and Mann, R. S. (1997). Switching the in vivo specificity of a minimal Hox-responsive element. *Development* **124**, 2007-2014.

Chou, T. B., Noll, E. and Perrimon, N. (1993). Autosomal P[ovoD1] dominant female-sterile insertions in *Drosophila* and their use in generating germ-line chimeras. *Development* **119**, 1359-1369.

Clandinin, T. R., Katz, W. S. and Sternberg, P. W. (1997). *Caenorhabditis elegans* HOM-C genes regulate the response of vulval precursor cells to inductive signal. *Dev Biol.* **182**, 150-161.

Cohen, S. M. and Jurgens, G. (1990). Mediation of Drosophila head development by gap-like segmentation genes. *Nature* **346**, 482-485.

Dalton, D., Chadwick, R. and McGinnis, W. (1989). Expression and embryonic function of *empty spiracles*: a Drosophila homeo box gene with two patterning functions on the anterior-posterior axis of the embryo. *Genes Dev.* **3**, 1940-1956.

Daubresse, G., Deuring, R., Moore, L., Papoulas, O., Zakrajsek, I., Waldrip, W. R., Scott, M. P., Kennison, J. A. and Tamkun, J. W. (1999). The *Drosophila kismet* gene is related to chromatin-remodeling factors and is required for both segmentation and segment identity. *Development* **126**, 1175-1187.

de Zulueta, P., Alexandre, E., Jacq, B. and Kerridge, S. (1994). Homeotic complex and teashirt genes co-operate to establish trunk segmental identities in Drosophila. *Development* **120**, 2287-2296.

Dickson, B. J., Van Der Straten, A., Dominguez, M. and Hafen, E. (1996). Mutations modulating Raf signaling in *Drosophila* eye development. *Genetics* **142**, 163-171.

Eisenmann, D. M., Maloof, J. N., Simske, J. S., Kenyon, C. and Kim, S. K. (1998). The beta-catenin homolog BAR-1 and LET-60 Ras coordinately regulate the Hox gene *lin-39* during *Caenorhabditis elegans* vulval development. *Development* **125**, 3667-3680.

Fasano, L., Röder, L., Core, N., Alexandre, E., Vola, C., Jacq, B. and Kerridge, S. (1991). The gene *teashirt* is required for the development of *Drosophila* embryonic trunk segments and encodes a protein with widely spaced zinc finger motifs. *Cell* **64**, 63-79.

Finkelstein, R. and Perrimon, N. (1991). The molecular genetics of head development in *Drosophila*. *Development* **112**, 899-912.

Flegel, W. A., Singson, A. W., Margolis, J. S., Bang, A. G., Posakony, J. W. and Murre, C. (1993). *Dpbx*, a new homeobox gene closely related to the human proto-oncogene *pbx1*: molecular structure and developmental expression. *Mech. Dev.* **41**, 155-161.

Florence, B. and McGinnis, W. (1998). A genetic screen of the Drosophila X chromosome for mutations that modify Deformed function. *Genetics* **150**, 1497-1511.

FlyBase. (1999). The FlyBase database of the Drosophila Genome Projects and community literature. The FlyBase Consortium. *Nucleic Acids Res.* **27**, 85-88.

Garcia-Bellido, A. (1977). Homeotic and atavic mutations in insects. *Amer. Zool.* **17**, 613-629.

Gellon, G., Harding, K., McGinnis, N., Martin, M. M. and McGinnis, W. (1997). A genetic screen for modifiers of *Deformed* homeotic function identifies novel genes required for head development. *Development* **124**, 3321-3331.

Gibson, G. and Gehring, W. J. (1988). Head and thoracic transformation caused by ectopic expression of *Antennapedia* during *Drosophila* development. *Development* **102**, 657-675.

Ginisty, H., Sicard, H., Roger, B. and Bouvet, P. (1999). Structure and functions of nucleolin. *J. Cell Sci.* **112**, 761-772.

Gonzalez-Reyes, A., Urquia, N., Gehring, W. J., Struhl, G. and Morata, G. (1990). Are cross-regulatory interactions between homeotic genes functionally significant. *Nature* **344**, 78-80.

Gray, S., Szymanski, P. and Levine, M. (1994) Short-range repression permits multiple enhancers to function autonomously within a complex promoter. *Genes Dev.* **8,** 1829-1838.

Hafen, E., Levine, M., Garber, R. L. and Gehring, W. J. (1983). An improved in situ hybridization method for the detection of cellular RNAs in *Drosophila* tissue sections and its application for localizing transcripts of the homeotic Antennapedia gene complex. *EMBO J.* **2**, 617-623.

Handa, N., Nureki, O., Kurimoto, K., Kim, I., Sakamoto, H., Shimura, Y., Muto, Y. and Yokoyama, S. (1999). Structural basis for recognition of the tra mRNA precursor by the Sex-lethal protein. *Nature* **398**, 579-585.

Harding, K. W., Gellon, G., McGinnis, N. and McGinnis, W. (1995). A screen for *Dfd* modifier mutations in *Drosophila*. *Genetics* **140**, 1339-1352.

Jiang, L. I. and Sternberg, P. W. (1998). Interactions of EGF, Wnt and HOM-C genes specify the P12 neuroectoblast fate in *C. elegans*. *Development* **125**, 2337-2347.

Jurgens, G., Wieschaus, E., Nüsslein-Volhard, C. and Kluding, H. (1984). Mutations affecting the pattern of the larval cuticle in *Drosophila melanogaster*. II. Zygotic loci on the third chromosome. *Roux's Arch. Dev. Biol.* **193**, 283-295.

Kennerdell, J. R. and Carthew, R. W. (1998). Use of dsRNA-mediated genetic interference to demonstrate that frizzled and frizzled 2 act in the wingless pathway. *Cell* **95**, 1017-1026.

Kolodziej, P. A., Jan, L. Y. and Jan, Y. N. (1995). Mutations that affect the length, fasciculation, or ventral orientation of specific sensory axons in the *Drosophila* embryo. *Neuron* **15**, 273-286.

Kurant, E., Pai, C.-y., Sharf, R., Halachmi, N., Sun, Y. H. and Salzberg, A. (1998). *dorsotonals*/*homothorax*, the Drosophila homologue of *meis1*, interacts with *extradenticle* in patterning of the embryonic PNS. *Development* **125**, 1037-1048.

Kuroiwa, A., Kloter, U., Baumgartner, P. and Gehring, W. J. (1985). Cloning of the homeotic *Sex combs reduced* in *Drosophila* and in situ localization of its transcripts. *EMBO J.* **4**, 3757-3764.

Kuziora, M. A. and McGinnis, W. (1988). Autoregulation of a *Drosophila* homeotic selector gene. *Cell* **55**, 477-485.

Lai, Z. C. and Rubin, G. M. (1992). Negative control of the photoreceptor development in *Drosophila* by the product of the *yan* gene, an ETS domain protein. *Cell* **70**, 609-620.

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

Li, X. and McGinnis, W. (1999). Activity regulation of Hox proteins, a mechanism for altering functional specificity in development and evolution. *Proc. Natl. Acad. Sci. USA* **96**, 6802-6907.

Li, X., Murre, C. and McGinnis, W. (1999). Activity regulation of a Hox protein and a role for the homeodomain in inhibiting transcriptional activation. *EMBO J.* **18**, 198-211.

Lindsley, D. L. and Zimm, G. G. (1992). The Genome of *Drosophila melanogaster*. Academic Press, Inc., San Diego.

Maloof, J. N. and Kenyon, C. (1998). The Hox gene *lin-39* is required during *C. elegans* vulval induction to select the outcome of Ras signaling. *Development* **125**, 181-190.

Mann, R. S. and Chan, S. K. (1996). Extra specificity from extradenticle: the partnership between HOX and PBX/EXD homeodomain proteins. *Trends Genet.* **12**, 259-262.

Martinez-Arias, A. (1986). The *Antennapedia* gene is required and expressed in parasegments 4 and 5 of the *Drosophila* embryo. *EMBO J.* **5**, 135-141.

McCormick, A., Core, N., Kerridge, S. and Scott, M. (1995). Homeotic response elements are tightly linked to tissue-specific elements in a transcriptional enhancer of the *teashirt* gene. *Development* **121**, 2799-2812.

McGinnis, N., Ragnhildstveit, E., Veraksa, A. and McGinnis, W. (1998).

A cap 'n' collar protein isoform contains a selective Hox repressor function. *Development* **125**, 4553-4564.

McGinnis, W., Jack, T., Chadwick, R., Regulski, M., Bergson, C., McGinnis, N. and Kuziora, M. A. (1990). Establishment and maintenance of position-specific expression of the *Drosophila* homeotic selector gene *Deformed*. *Adv. Genet.* **27**, 363-402.

McGinnis, W. and Krumlauf, R. (1992). Homeobox genes and axial patterning. *Cell* **68**, 283-302.

Merrill, V. K. L., Diederich, R. J., Turner, F. R. and Kaufman, T. C. (1989). A genetic and developmental analysis of mutations in *labial*, a gene necessary for proper head formation in *Drosophila melanogaster*. *Dev. Biol.* **135**, 376-391.

Merrill, V. K. L., Turner, F. R. and Kaufman, T. C. (1987). A genetic and developmental analysis of mutations in the *Deformed* locus in *Drosophila melanogaster*. *Dev. Biol.* **122**, 379-395.

Mohler, J., J. W. Mahaffey, E. Deutsch, and Vani, K. (1995). Control of *Drosophila* head segment identity by the bZIP homeotic gene *cnc*. *Development* **121**, 237-247.

Myers, R. M., T. Maniatis, and Lerman, L. S. (1987). Detection and localization of single base changes by denaturing gradient gel electrophoresis. *Methods Enzymol.* **155**, 501-527.

Neckameyer, W. S. and White, K. (1993). Drosophila tyrosine hydroxylase is encoded by the pale locus. *J. Neurogenet.* **8**, 189-199.

Newby, L. M. and Jackson, F. R. (1996). Regulation of a specific circadian clock output pathway by lark, a putative RNA-binding protein with repressor activity. *J. Neurobiol.* **31**, 117-128.

Nüsslein-Volhard, C., Wieschaus, E. and Kluding, H. (1984). Mutations affecting the pattern of the larval cuticle in *Drosophila melanogaster*. I. Zygotic loci on the second chromosome. *Roux's Arch. Dev. Biol.* **193**, 267-282.

Pai, C.-Y., Kuo, T.-S., Jaw, T. J., Kurant, E., Chen, C.-T., Bessarab, D., Salzberg, A. and Sun, Y. H. (1998). The Homothorax homeoprotein activates the nuclear localization of another homeoprotein, Extradenticle, and suppresses eye development in Drosophila. *Genes Dev.* **12**, 435-446.

Pattatucci, A. M., Otteson, D. C. and Kaufman, T. C. (1991). A functional and structural analysis of the *Sex-combs reduced* locus of *Drosophila melanogaster*. *Genetics* **129**, 423-441.

Peifer, M. and Wieschaus, E. (1990). Mutations in the *Drosophila* gene *extradenticle* affect the way specific homeo domain proteins regulate segmental identity. *Genes Dev.* **4**, 1209-1223.

Pinsonneault, J., Florence, B., Vaessin, H. and McGinnis, W. (1997). A model for extradenticle function as a switch that changes Hox proteins from repressors to activators. *EMBO J.* **16**, 2032-2042.

Poodry, C. A. (1980). Epidermis, morphology and development. In *The Genetics and Biology of Drosophila* (ed. M. Ashburner and T. R. F. Wright), pp. 443-497. Academic Press, New York.

Pradel, J. and White, R. A. H. (1998). From selectors to realizators. *Int. J. Dev. Biol.* **42**, 417-421.

Rauskolb, C., Peifer, M. and Wieschaus, E. (1993). *extradenticle*, a regulator of homeotic gene activity, is a homolog of the homeobox-containing human proto-oncogene *pbx1*. *Cell* **74**, 1101-1112.

Regulski, M., McGinnis, N., Chadwick, R. and McGinnis, W. (1987). Developmental and molecular analysis of *Deformed*: A homeotic gene controlling *Drosophila* head development. *EMBO J.* **6**, 767-777.

Restifo, L. L. and Merrill, V. K. L. (1994). Two Drosophila regulatory genes, *Deformed* and the Broad-Complex, share common functions in development of adult CNS, head, and salivary glands. *Dev. Biol.* **162**, 465-485.

Rieckhof, G. E., Casares, F., Ryoo, H. D., Abu-Shaar, M. and Mann, R. S. (1997). Nuclear translocation of Extradenticle requires *homothorax*, which encodes an Extradenticle-related homeodomain protein. *Cell* **91**, 171-183.

Robertson, H. M., Preston, C. R., Phillis, R. W., Johnson-Schlitz, D. M., Benz, W. K. and Engels, W. R. (1988). A stable genomic source of P-element transposase in *Drosophila melanogaster*. *Genetics* **118**, 461-470.

Röder, V. Vola, C., and Kerridge, S. (1992). The role of the *teashirt* gene in trunk segmental identity in *Drosophila*. *Development* **115**, 1017-1033.

Rogge, R., Green, P. J., Urano, J., Horn-Saban, S., Mlodzik, M., Shilo, B. Z., Hartenstein, V. and Banerjee, U. (1995). The role of *yan* in mediating the choice between cell division and differentiation. *Development* **121**, 3947-3958.

Sato, T., Hayes, P. H. and Denell, R. E. (1985). Homeosis in *Drosophila*: Roles and spatial patterns of expression of the *Antennapedia* and *Sex combs reduced* loci in embryogenesis. *Dev. Biol.* **111**, 171-192.

Séguin, A., Laible, G., Leyva, A., Dixon, R. A. and Lamb, C. J. (1997). Characterization of a gene encoding a DNA-binding protein that interacts in vitro with vascular specific cis elements of the phenylalanine ammonia-lyase promoter. *Plant Mol. Biol.* **35**, 281-291.

Serin, G., Joseph, G., Ghisolfi, L., Bauzan, M., Erard, M., Amalric, F. and Bouvet, P. (1997). Two RNA-binding domains determine the RNA-binding specificity of nucleolin. *J. Biol. Chem.* **272**, 13109-13116.

Siomi, H. and Dreyfuss, G. (1997). RNA-binding proteins as regulators of gene expression. *Curr. Opin. Genet. Dev.* **7**, 345-353.

Staehling-Hampton, K., Ciampa, P. J., Brook, A. and Dyson, N. (1999) A genetic screen for modifiers of E2F in *Drosophila melanogaster*. *Genetics* **153**, 275-287.

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.

van Dijk, M. and Murre, C. (1994). extradenticle Raises the DNA Binding Specificity of Homeotic Selector Gene Products. *Cell* **78**, 617-624.

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

Wieschaus, E. and Nüsslein-Volhard, C. (1986). Looking at embryos. In *Drosophila, a Practical Approach* (ed. D. B. Roberts), Pp. 199-227. IRL Press: Oxford.

Wright, T. R. F. (1996). Phenotypic analysis of the *Dopa decarboxylase* gene cluster mutants in *Drosophila melanogaster*. *J. Hered.* **87**, 175-190.