Development 117, 585-596 (1993)
Printed in Great Britain © The Company of Biologists Limited 1993

Conservation of regulatory elements controlling *hairy* pair-rule stripe formation

James A. Langeland and Sean B. Carroll*

Howard Hughes Medical Institute, Laboratory Of Molecular Biology, University of Wisconsin, 1525 Linden Drive, Madison, Wiscosin 53706, USA

*Author for correspondence

SUMMARY

The *hairy (h)* gene is one of two pair-rule loci whose striped expression is directly regulated by combinations of gap proteins acting through discrete upstream regulatory fragments, which span several kilobases. We have undertaken a comparative study of the molecular biology of *h* pair-rule expression in order to identify conserved elements in this complex regulatory system, which should provide important clues concerning the mechanism of stripe formation. A molecular comparison of the *h* locus in *Drosophila virilis* and *Drosophila melanogaster* reveals a conserved overall arrangement of the upstream regulatory elements that control individual pair-rule stripes. We demonstrate that upstream fragments from *D. virilis* will direct the proper expression of stripes in *D. melanogaster*, indicating that these are true functional homologs of the stripe-producing *D. melanogaster* regulatory elements, and that the network of *trans*-acting proteins that act upon these regulatory elements is highly conserved. We also demonstrate that the spatial relationships between specific *h* stripes and selected gap proteins are highly conserved. We find several tracts of extensive nucleotide sequence conservation within homologous stripe-specific regulatory fragments, which have facilitated the identification of functional subelements within the *D. melanogaster* regulatory fragment for *h* stripe 5. Some of the conserved nulceotide tracts within this regulatory fragment contain consensus binding sites for potential *trans*-regulatory (gap and other) proteins, while many appear devoid of known binding sites. This comparative approach, coupled with the analysis of reporter gene expression in gap mutant embryos suggests that the Kr and gt proteins establish the anterior and posterior borders of *h* stripe 5, respectively, through spatial repression. Other, as yet unidentified, proteins are certain to play a role in stripe activation, presumably acting through other conserved sequence tracts.

Key words: *hairy*, sequence evolution, segmentation, *Drosophila*

INTRODUCTION

Investigations into the genetic control of early *Drosophila* development have identified a hierarchical network of genes that generate pattern along the anterior-posterior axis of the embryo (for review see Akam, 1987; Ingham, 1988). Typically, genes in this early network encode regulatory proteins, which activate and/or repress downstream genes, thus leading to a progressive refinement of the spatial organization of the embryo. Genes at the top of the hierarchy (*e.g. bicoid* (*bcd*) and gap genes) have relatively simple expression patterns (Driever and Nüsslein-Volhard, 1988; Knipple et al., 1985; Tautz, 1988; Pankratz et al., 1990; Pignoni et al., 1990; Kraut and Levine, 1991a; Eldon and Pirrotta, 1991); those acting further down (*e.g.* pair-rule and segment polarity genes) have more detailed patterns (Carroll and Scott, 1985; DiNardo et al., 1985) reflecting more complex regulatory input. While the regulatory interactions of these genes have been extensively analyzed by genetic means (Jäckle et al., 1986; Howard and Ingham, 1986; Carroll and Scott, 1986; DiNardo and O'Farrell, 1987; Frasch and Levine, 1987; Ingham and Gergen, 1988; Carroll and Vavra, 1989), recent work has increasingly focused on the dissection of the regulatory regions that control the patterns of expression of these genes and the localization of critical binding sites for proteins that genetic analyses indicate activate and/or repress their transcription (examples include Dearolf et al., 1989, Stanojevic et al., 1991; Hoch et al., 1992).

The pair-rule gene *hairy* (*h)* is an excellent example of a gene with a complex expression pattern, a large number of potential *trans*-regulating proteins, and a correspondingly complex regulatory region. *h* is expressed in seven transverse stripes encircling the embryo at cellular blastoderm (Carroll et al., 1988; Hooper et al., 1989) and is required for normal segmentation. *h* is also expressed later in the developing larva, forming detailed patterns in imaginal discs (Carroll and Whyte, 1989). Examinations of *h* pair-rule expression in mutant embryos have identified the gap genes *hunchback* (*hb*) *Krüppel (Kr), giant (gt)* and *knirps*

(kni) as critical regulators of this pattern (Ingham et al., 1986; Carroll and Vavra, 1989; Hooper et al., 1989), although which of the observed genetic interactions are direct and which are indirect remains unclear. These gap proteins are expressed regionally along the axis of the embryo in partially overlapping domains (Knipple et al., 1985; Tautz, 1988; Gaul and Jäckle, 1989; Pankratz et al., 1990; Pignoni et al., 1990; Kraut and Levine, 1991a; Eldon and Pirrotta, 1991); leading to the general model that individual *h* stripes are under the control of the specific combinations of *trans*-acting factors which are present at each position in the embryo (Pankratz et al., 1989; Howard and Struhl, 1990). However, it has not been shown that the precise relationships between domains of gap protein expression and *h* stripes are important for *h* pair-rule regulation. Each gap gene encodes a DNA-binding protein that binds well-defined recognition sequences (Stanojevic et al., 1989; Triesman and Desplan, 1989; Eldon and Pirrotta, 1991; Small et al., 1991; Hoch et al., 1992); three of these proteins (hb, Kr, and gt), along with the maternal determinant bcd, directly interact with sequences upstream of the pair-rule gene *even-skipped (eve)* to direct stripe formation (Small et al., 1991; Stanojevic et al., 1991). Additionally, both Kr and kni proteins have been shown to bind fragments of the *h* regulatory region (Pankratz et al., 1990). These results strongly suggest a direct role for these products in regulating the transcription of *h* stripes; however, in contrast to the case for *eve*, specific, functional binding sites have not yet been localized in *cis*-elements controlling *h* stripes.

The dissection of the region 5′ to *h*-coding sequences, as well as that of *eve*, has identified discrete regulatory elements spanning several kilobases of both genes which govern stripe formation (Goto et al., 1989; Pankratz et al.,1990; Howard and Struhl, 1990; Riddihough and Ish-Horowicz, 1991). When placed in front of the *lacZ* reporter gene and transformed into flies, specific upstream elements give rise to β-gal expression patterns, which coincide with particular subsets of native pair-rule stripes. These findings have simplified the study of pair-rule regulation by allowing both the genetic manipulations of expression arising from individual stripe elements (Pankratz et al., 1990; Small et al., 1991; Riddihough and Ish-Horowicz, 1991), as well as the biochemical analysis of interactions between *trans*-regulating DNA-binding proteins and stripe-specific promoter fragments (Pankratz et al., 1990; Small et al., 1991; Stanojevic et al., 1991.) However, the complexity of the *cis*-regulatory region and the intricate genetic interactions and expression patterns of potential regulatory proteins have remained a hinderance to a mechanistic understanding of the regulation of individual *h* stripes.

In this paper, we have undertaken a comparative approach of the structure, expression and regulation of the *h* pair-rule gene in order to identify conserved elements in this complex regulatory system, which provide potentially important clues concerning the mechanism of stripe formation. We have cloned the *h* locus from *Drosophila virilis*, and have sequenced and compared the overall molecular organization of the *h* locus in these species. We provide evidence for a high degree of conservation in the upstream organization of the *h* locus; both the gross arrangement of stripe regulatory elements and several tracts of nucleotide sequence within homologous stripe elements have been conserved. Interspecific comparisons have been used to determine conserved regulatory and coding regions of several genes in *Drosophila* (Kassis et al., 1989; Treier et al., 1989; Maier et al., 1990; Seeger and Kaufman, 1990; Williams et al., 1991; Michael et al., 1990). Conserved regulatory regions located by such comparisons have proven to correspond to functional *cis*-elements (Maier et al.,1990), and to contain binding sites for regulatory proteins (Kassis et al., 1989). Our finding that regulatory fragments from *D. virilis* will produce stripes in *D. melanogaster* indicates a highly conserved network of *trans*-acting proteins, in addition to a highly conserved *cis*-regulatory region. The comparative sequence alignment described here precisely delimits conserved, and therefore likely important regulatory sequences within upstream regions that govern stripe formation. We demonstrate that sequences uncovered in this manner allow us to identify functional subelements in previously defined stripe regulatory fragments and can provide new information regarding potential sites for regulatory proteins. We extend our comparative approach to include a comparison of patterns of regulatory gene expression, whereby we identify highly conserved spatial relationships between h protein stripes and the domains of select gap proteins. The analysis of reporter gene expression in gap mutant backgrounds confirms that the conserved spatial relationships represent important regulatory interactions and, augmented by the location of potential sites uncovered in the sequence alignment, suggests a model for the spatial regulation of *h* stripe 5.

MATERIALS AND METHODS

Cloning and mapping of the *D. virilis h* locus

Restriction fragments containing regulatory sequences for *h* stripes 1 and 5 (2.8 kb *Kpn*I-*Bam*HI fragment) as well as stripes 3 and 4 (1.4 kb *Cla*I-*Cla*I fragment) from *D. melanogaster* were subcloned from a parent plasmid containing ~12 kb of the *h* upstream region (provided by K. Howard) into Bluescript KSII+ vectors (Stratagene). These inserts, as well as the *D. melanogaster h* cDNA (Rushlow et al., 1989), were randomly labelled with α^{32}P-dATP (Feinberg and Vogelstein, 1983) and used to probe a λEMBL3 *D. virilis* genomic library (kindly provided by J. Tamkun). Roughly 2×10^5 plaques were screened and 12 positive plaques were picked, of which 4 were eventually purified. Purified DNA from these phages was restriction mapped and blots of mapping digests were probed with the labelled cDNA and various labelled *D. melanogaster* regulatory fragments to identify homologous fragments in *D. virilis*. All hybridizations were carried out as in Klessig and Berry (1983) at 42°C, except that the Southern hybridizations contained no dextran sulfate. Filters were washed at 37°C in three changes of 0.1× SSC, 0.1% SDS.

Interspecific transformants and in situ hybridizations

The *D. virilis* fragment (2.5 kb *Hin*dIII-*Sal*I), which hybridizes to the 2.8 kb *Kpn*I-*Bam*HI *D. melanogaster* fragment (controlling stripes 1 and 5), as well as the 11 kb *Xho*I-*Xho*I *D. virilis* fragment, which hybridizes to all the remaining *D. melanogaster* stripe-specific regulatory fragments, were cloned in their native orientation into the P-element transformation vector hsp-lacZ-

CASPER (Nelson and Laughon, 1992) and used to transform *D. melanogaster* flies as described by Spradling (1986). 0-4 hour embryos from lines homozygous for the resulting inserts were fixed and *lacZ* transcripts detected by hybridization with digoxigenin-labelled antisense *lacZ* RNA probes made with the Genius 4 kit (Boehringer-Mannheim) following the protocol of Jiang et al., (1991). This same procedure was used to generate and analyze transformant lines carrying various deletions of the *h* stripe 5 regulatory region. Various transformant lines were crossed into Kr^{-} and gt^{-} genetic backgrounds using the following stocks: Kr^{1}/SM1 and gt^{X11}/FM6.

Sequence comparison

Subcloned regulatory fragments from *D. melanogaster* and their homologs from *D. virilis* (described above) were sequenced on both strands using the Sequenase kit (US Biochemicals). A combination of nested deletions (created using the Erase-a-Base kit, Promega) and customized sequencing primers (UW Biotechnology Center) were employed to generate continuous sequence. Sequences from both species were aligned using the GAP program of the UWGCG software (Devereux et al., 1984). Optimal alignment was achieved using the following parameters within the GAP program: gap weight (2), gap length weight (0). Restriction fragments that do not interrupt conserved tracts of sequence were subcloned and used to transform flies as described above. Potential binding sites within conserved tracts were identified using the CONSENSUS and FITCONSENSUS programs of the UWGCG software and published binding sites for gt (Small et al., 1991; Capovilla et al., 1992), Kr (Stanojevic et al., 1989; Triesman and Desplan, 1989) and FTZ-F1 (Ueda et al., 1990).

Antibody production and embryo immunohistochemistry

Gap proteins were bacterially expressed (Sturdier et al., 1986) from full-length cDNAs cloned into inducible T7 expression vectors (generously provided by M. Levine and G. Struhl). Inclusion bodies containing the foreign protein were purified from whole lysates of induced cultures and used to immunize rabbits. The resulting sera were purified by affinity chromotography to generate specific polyclonal antisera for each of the gap proteins. 0-6 hour embryos from wild-type *D. melanogaster* and *D. virilis* stocks were fixed acccording to the procedures of Carroll and Scott (1985) and labelled for double immunoflourescence using the rabbit polyclonal gap antisera at 0.5 μg/ml and a monoclonal supernatant raised against the *h* protein (provided by J. Gates) at a dilution of 1:1. Biotinylated donkey anti-rabbit IgG, fluorescein-conjugated strepavidin and rhodamine-conjugated donkey anti-mouse IgG were used as secondary and tertiary reagents (Jackson ImmunoResearch) all at dilutions of 1:200 from the manufacturer's stocks. Images were taken on a Laser Scanning Confocal Microscope and processed according to the methods described in Paddock et al. (1993).

RESULTS

The molecular organization of the *h* locus from *D. virilis*: the order of stripe elements is conserved

We cloned the genomic region containing the homolog of the *h* locus from a λEMBL3 *D. virilis* genomic library (provided by J. Tamkun, see Materials and methods). Three DNA fragments from the *D. melanogaster* locus were used to screen the library and in subsequent rescreening steps. The specific probes are diagrammed in Fig. 1A and correspond to the *h* cDNA (Rushlow et al., 1989) and two upstream fragments, which drive the expression of stripes 3 and 4 (1.4 kb *Cla*I-*Cla*I) and stripes 1 and 5 (2.8 kb *Kpn*I-*Bam*HI; Pankratz et al., 1990, Howard and Struhl, 1990, Riddihough and Ish-Horowicz, 1991). Two plaques were purified that hybridized to all three probes and two that hybridized only to the upstream fragments; the upstream fragments provided stronger hybridization signals than the coding region. The restriction map for the genomic region contained in these overlapping phage is shown in Fig. 1B.

Southern blots from the cloned *D. virilis* genomic region were probed with labelled fragments from the *D. melanogaster h* locus in order to locate restriction fragments that are homologous to the known coding and regulatory fragments in *D. melanogaster*. The resulting comparison of the two loci is shown in Fig. 1A and B. The *h* cDNA hybridizes specifically to the restriction fragments corresponding to a 4.8 kb *Hin*dIII-*Eco*RI fragment. A portion of this fragment was sequenced starting from the internal *Sal*I site and the predicted translation product reveals near perfect amino acid sequence conservation with the amino end of the HLH domain of the *D. melanogaster h* protein (Rushlow et al., 1989), and matches perfectly with the sequence of the *D. virilis h* cDNA described by Wainwright and Ish-Horowicz (1992) (data not shown). Since this region contains the distinguishing sequence motifs for *h* and other related HLH proteins (Murre et al., 1991; Rushlow et al., 1989), we are confident that our clones represent the bona fide *h* homolog from *D. virilis*. Each of the defined stripe regulatory elements from the *D. melanogaster h* gene (Pankratz et al., 1990, Howard and Struhl, 1990, Riddihough and Ish-Horowicz, 1991) has a specific homolog in *D. virilis* (Fig. 1A and B). Although the *D. virilis h* locus appears to be ~20% larger than the *D. melanogaster* locus and contains extra upstream DNA that is not detected by any *D. melanogaster* fragments (arrows, Fig. 1B), the relative order of the stripe-specific regulatory sequences is the same in the two species. We also detect specific homologs for upstream fragments, which lie both 3′ and 5′ to the stripe-producing fragments, even though these have no known regulatory function.

Regulatory sequences from *D. virilis* produce stripes in *D. melanogaster*

The upstream fragments from *D. melanogaster* that we used to probe the *D. virilis h* locus have been shown to drive reporter gene expression corresponding to specific subsets of *h* stripes in the embryo (Pankratz et al., 1990; Howard and Struhl, 1990; Riddihough and Ish-Horowicz, 1991). In order to test the degree of functional conservation between regulatory fragments that we determined to be homologous by Southern hybridizations, we generated transformants in *D. melanogaster* using constructs containing *lacZ* under the control of upstream *h* fragments from *D. virilis* (see Materials and methods). We then compared patterns of *lacZ* transcript accumulation arising from *D. virilis* constructs with the known striped expression arising from our *D. melanogaster* constructs (provided by T. Orenic in this laboratory). A comparison of reporter gene expression arising from constructs containing the 2.8 kb *Kpn*I-*Bam*HI fragment from *D. melanogaster* and the 2.5 kb *Hin*dIII-*Sal*I fragment from *D. virilis* is shown in Fig. 2A and B (see

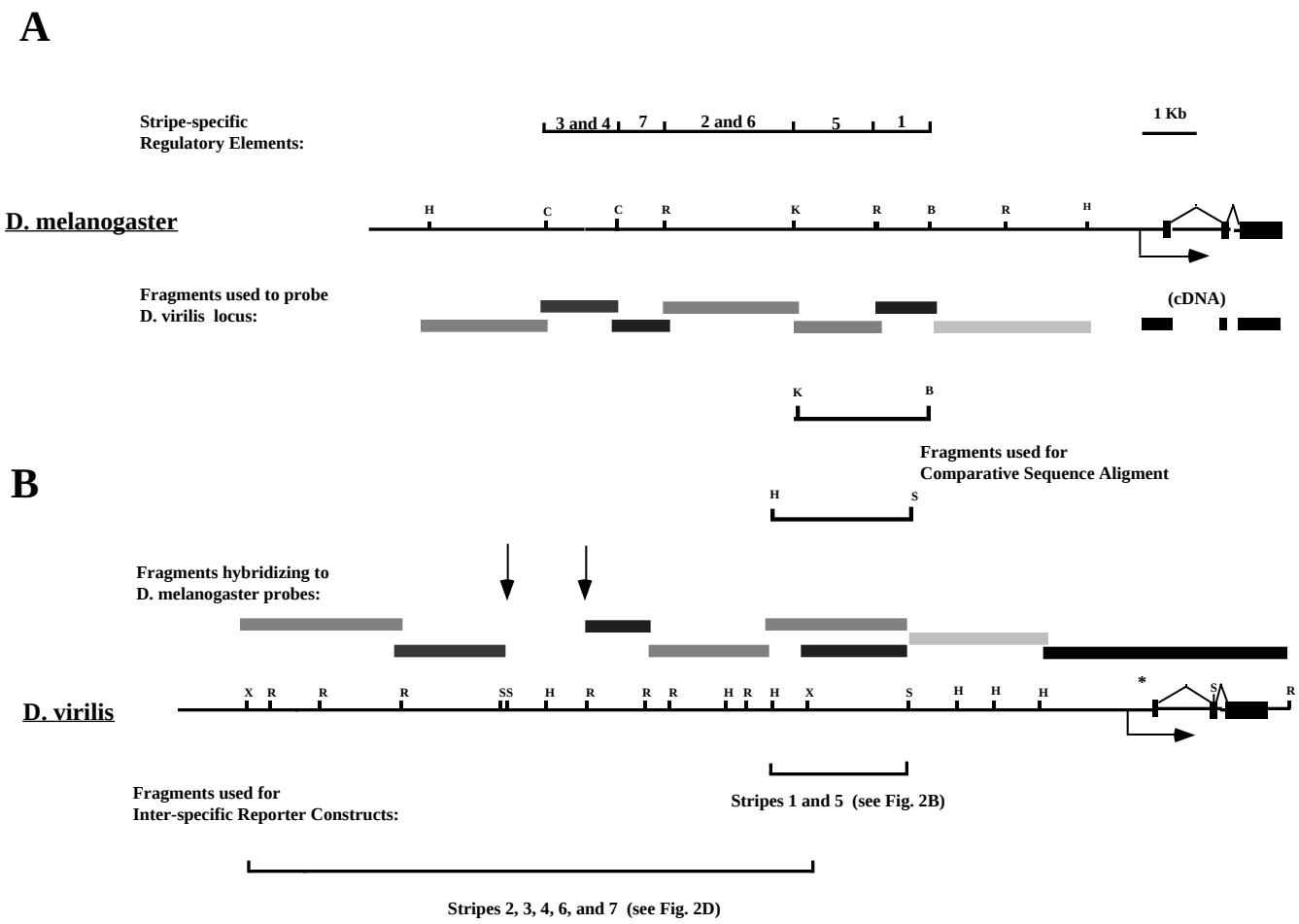

Fig. 1. The molecular organization of the *h* locus is highly conserved between *D. melanogaster* and *D. virilis*. A comparative alignment of the *h* locus in *D. melanogaster* and *D. virilis* indicating the locations of the coding region and selected stripe-specific enhancers in *D. melanogaster* (1A, Rushlow et al., 1989; Pankratz et al., 1990, Howard and Struhl, 1990, Riddihough and Ish-Horowicz, 1991) and their homologs in *D. virilis* (1B). The *h* locus was cloned from a *D. virilis* genomic library and Southern hybridizations on digests of the *D. virilis* locus using *D. melanogaster* coding and regulatory regions as probes (bars below *D. melanogaster* map) reveal that each of the defined stripe regulatory elements from *D. melanogaster* has a specific homolog in *D. virilis* (bars above *D. virilis* map). Conversely, only a small portion of the *D. virilis* locus fails to contain detectable homology with any *D. melanogaster* probe (arrows). Note that the linear order of these stripe elements is identical in the two species. Fragments producing stripes in interspecific reporter constructs are indicated below the *D. virilis* map; upstream fragments used in subsequent comparative sequence analysis are indicated between the two maps. *The location of the transcription and translation start sites, as well as the first and third exon splice sites for *D. virilis* and are extrapolated from Wainwright and Ish-Horowicz (1992). Restriction sites are as follows: B, *Bam*HI, C, *Cla*I, H, *Hin*dIII, R, *Eco*RI, S, *Sal*I, X, *Xho*I.

map in Fig. 1A and B for the locations of various fragments). Each fragment produces stripes at cellular blastoderm which correspond to *h* stripes 1 and 5. Although the posterior stripe arising from the *D. virilis* construct expresses rather weakly, it is spatially correct. Constructs containing the more distal region of the *D. virilis* locus (11 kb *Xho*I-*Xho*I fragment) drive expression corresponding to the remaining *h* stripes (stripes 2, 3, 4, 6 and 7, see Fig. 2D), matching the expression arising from the 7 kb *Hin*dIII-*Kpn*I *D. melanogaster* fragment (Fig. 2C). The striped expression arising from these interspecific reporter constructs demonstrates that the *D. virilis* fragments are true functional homologs of the *D. melanogaster* stripe elements and illustrates the highly conserved nature of both the *cis*-acting sequences, which produce pair-rule stripes, and the network of *trans*-acting factors, which act upon these sequences.

Homologous stripe regulatory elements contain large tracts of conserved sequence

The results from our hybridization and interspecific reporter constructs indicate that significant sequence homology exists between *h* regulatory fragments from *D. melanogaster* and *D. virilis*. In order to determine the extent and exact locations of this homology, we sequenced homologous stripe-producing regulatory fragments from each species. A comparative sequence alignment of the *D. melanogaster* regulatory fragment for h stripes 1 and 5 (2.8 kb *Kpn*I-*Bam*HI fragment, numbered from the transcription start site; Fig. 1) and its homolog from *D. virilis* (2.5 kb

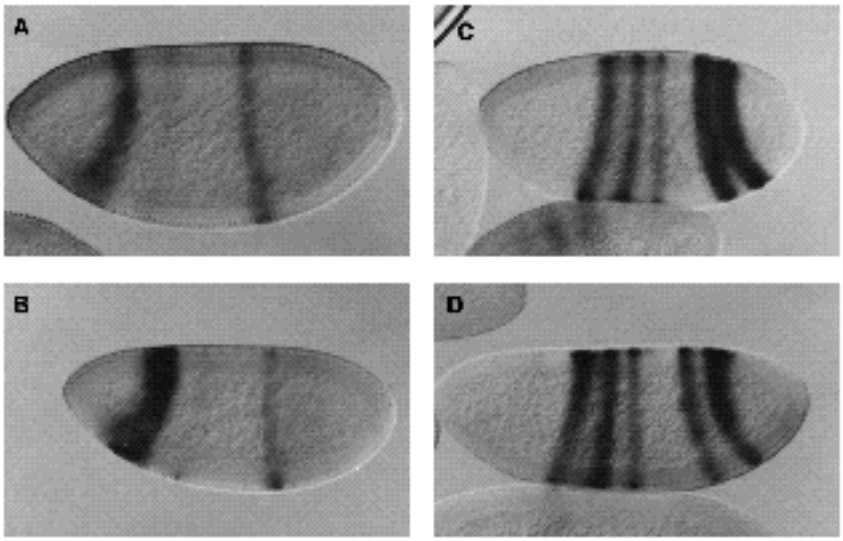

Fig. 2. Regulatory sequences from *D. virilis* produce stripes in *D. melanogaster*. Striped *lacZ* transcript accumulation in *D. melanogaster* embryos transformed with reporter constructs containing the 2.8 kb *Kpn*I-*Bam*HI fragment from *D. melanogaster* (A), the 2.5 kb *Sal*I-*Hin*dIII homolog of this fragment from *D. virilis* (B), the 7 kb *Hin*dIII-*Kpn*I fragment from *D. melanogaster* (C) and the 11 kb *Xho*I-*Xho*I fragment comprising most of the remaining *D. virilis h* upstream region (D, see map, Fig. 1). The embryos in A and B each produce stripes, which correspond to h protein stripes 1 and 5 indicating that the regulatory fragments are functionally homologous. The fifth stripe arising from the *D. virilis* fragment is weaker, but spatially correct. The embryos in C and D produce stripes corresponding to the remaining *h* stripes (stripes 2, 3, 4, 6 and7).

*Hin*dIII-*Sal*I fragment; Fig. 1) is shown in Fig. 3. This alignment reveals several tracts of perfect sequence conservation interspersed in divergent sequence. Conserved nucleotides are capitalized and regions that contain eight or more consecutive conserved nucleotides are shown in underlined bold type. Such tracts comprise 25% (695 out of 2832 nucleotides) of the *D. melanogaster* regulatory region for stripes 1 and 5, with the longest tract containing 38 continuous bases (*D. melanogaster* nucleotides −5339 to −5302, double underlined). We found a similar degree of conservation in regulatory fragments giving rise to stripes 3 and 4 (data not shown). As these conserved sequences are apparently highly constrained in evolution, they very likely contain the functional *cis*-elements which direct stripe formation.

Conserved sequences in functional regulatory subelements which produce *h* stripe 5

The alignment of *h* regulatory sequences for stripes 1 and 5 (Fig. 3) provides a molecular guide with which we have significantly narrowed the previously defined regulatory sequences for *h* stripe 5 (Pankratz et al., 1990; Howard and Struhl, 1990; Riddihough and Ish-Horowicz, 1991). We have constructed a deletion series through the regulatory region for stripe 5 which avoids the random splitting of large conserved tracts and used these to make *lacZ* reporter constructs. The *D. melanogaster* sequence of the 784 bp *Eag*I-*Bal*I fragment within the regulatory fragment for *h* stripes 1 and 5 is shown in Fig. 4A (see Fig. 3 for context). Nucleotides which are conserved in an alignment with the homologous *D. virilis* fragment are capitalized; large conserved tracts are shown in bold. A reporter construct containing the 569 bp *Eag*I-*Eco*RI fragment gives rise to a pattern of *lacZ* transcripts which corresponds to *h* stripe 5 (Fig. 4B). The posterior border of this stripe is not well-defined, very similar to the expression arising from the larger 1.7 kb *Kpn*I-*Eco*RI fragment described by Riddihough and Ish-Horowicz (1991, see Figs 1A and 3 for context) and suggesting that sequences required to establish the proper posterior border of stripe 5 lie proximal to the *Eco*RI breakpoint. Constructs containing only the 273 bp *Pvu*I-*Eco*RI fragment gives rise to no significant expression (data not shown) indicating that sequences that are neccessary for the activation of stripe 5 lie in the 302 bp *Eag*I-*Pvu*I fragment**.**

We examined conserved tracts within this more narrowly defined regulatory fragment for sequences that match published consensus binding sites for potential regulatory proteins. Both genetic evidence (Pankratz et al., 1990, Riddihough and Ish-Horowicz, 1991) and double-immunostaining of embryos (Howard and Struhl, 1991; this paper) indicate that *h* stripe 5 may be directly regulated by the gap proteins Kr, gt and kni. Using published binding sites for these proteins, we locate two potential binding sites for the gap protein Kr (Stanojevic et al., 1989; Triesman and Desplan, 1989) and one for the gap protein gt (Small et al., 1991; Capovilla et al., 1992) within conserved *h* stripe 5 regulatory sequence tracts. We did not find any reasonable matches for the kni-binding site (Hoch et al., 1992). We also find a well-conserved match for the binding site of the FTZ-F1 protein, which interacts with the stripe-producing zebra element of the pair-rule gene *ftz* and appears to be involved with the transcriptional activation of *ftz* (Ueda et al., 1990). There is a reasonable correlation between where we locate these putative sites and their potential biological function. The Kr recognition sequences could account for the relatively sharp anterior border of expression arising from our constructs, and the FTZ-F1 site is located within the fragment that we have shown to be required for stripe 5 activation. The potential gt site lies proximal to the *Eco*RI site and thus could partially account for the posterior expansion of expression in constructs lacking these sequences (Fig. 4B). Whether these are indeed bona fide binding sites is being tested in vitro by DNaseI footprinting.

The spatial relationships between gap proteins and *h* stripes are conserved in *D. melanogaster* and *D. virilis*

Genetic studies indicate that the gap proteins are critical components of the network of spatial regulators of *h* pair-rule stripe expression (Ingham et al., 1986; Howard, 1988; Carroll et al., 1988; Carroll and Vavra, 1989; Hooper et al.,

Fig. 3

1989). Since our reporter constructs suggest that these *trans*-acting elements should be conserved, we examined gap and h protein expression domains in each species to determine specifically which spatial relationships are conserved and therefore likely important. We produced antisera to each of the *D. melanogaster* gap proteins, and found

Fig. 3. Tracts of conserved sequence highlight stripe regulatory elements. Comparative alignment of the nucleotide sequence of the *D. melanogaster* regulatory fragment controlling *h* stripes 1 and 5 (2.8 kb *Kpn*I-*Bam*HI fragment) and its homolog from *D. virilis* (2.5 kb *Hin*dIII-*Sal*I fragment, see map Fig. 1). Conserved nucleotides are capitalized; regions that contain eight or more consecutive conserved nucleotides are shown in underlined bold type. This alignment reveals several tracts of near perfect sequence conservation interspersed in divergent sequence. 25% (695 out of 2832 nucleotides) of the *D. melanogaster* sequence is comprised of blocks of eight or more perfectly conserved contiguous bases. The longest such tract is 38 bp (*D. melanogaster* nucleotides −5339 to −5302, double underlined). We found a similar array of conserved tracts for regulatory fragments giving rise to *h* stripes 3 and 4 (data not shown). **D. melanogaster* sequence proximal to the *Eco*RI (R) site (nucleotide −5150) was determined by Rushlow et al. (1989)

that the affinity-purified anti-hb, anti-Kr and anti-gt antibodies cross-react well with the *D. virilis* proteins in whole-mount embryos. Using these antisera and a cross-reactive monoclonal antibody rasied against the *h* protein, we compared the expression domains of gap proteins and native *h* protein stripes in each of these species using double immunofluorescence imaging (see Materials and methods, anti-hb staining not shown). The relationships of Krüppel and giant protein domains to h stripes are conserved between *D. melanogaster* and *D. virilis* (Fig. 5), even at the fine level of detail observed at stripe borders. For example, h stripe 5 is flanked by Kr (Fig. 5A,B) to the anterior and gt (Fig. 5C,D) to the posterior in both species (indicated by arrows). Since these gap proteins (among others) are required for proper *h* pair-rule patterning, we conclude that these border relationships represent important regulatory interactions which have been conserved in evolution.

Genetic analysis indicates that *Kr* and *gt* establish the borders of *h* stripe 5

In order to test whether these conserved spatial relationships represent bona fide regulatory interactions, we examined *h* stripe 5 reporter gene expression in Kr⁻ and gt⁻ mutant embryos. Stripe 5 is clearly expanded posteriorly in gt⁻ embryos indicating that the gt protein is responsible for setting this border (compare Fig. 6A,C). Kr appears to have a dual effect on stripe 5 expression, depending upon the context and extent of regulatory DNA present. A 1.3 kb fragment clearly illustrates the role of Kr in repressing this element and setting the anterior border of stripe 5 expression (compare Fig. 6D,E), while a larger 2.7 kb fragment does not express stripe 5 in Kr⁻ background (Fig. 6B).

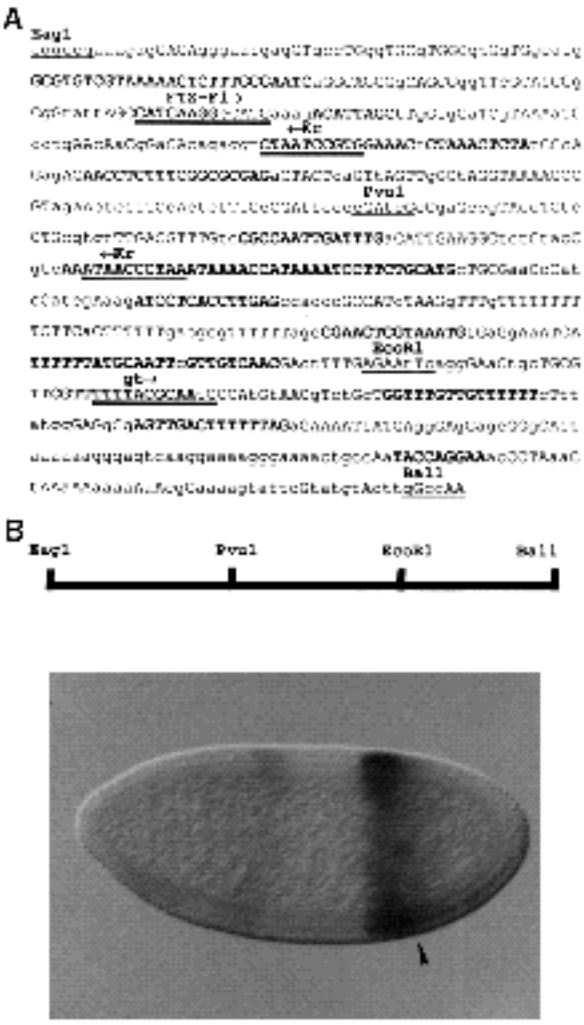

Fig. 4. Conserved regulatory sequences contain potential binding sites for regulatory proteins and guide the functional dissection of the *cis* region controlling *h* stripe 5. Sequence of a 784 bp *Eag*I-*Bal*I fragment, portions of which were used in various stripe 5 deletion constructs (A). Nucleotides that are conserved in an alignment with the homologous *D. virilis* fragment are capitalized, large tracts are shown in bold type. Sequences within conserved tracts that match published binding sites for the gap proteins Kr and gt, as well as the FTZ-F1 protein are indicated. (B) The pattern of *lacZ* transcript accumulation in a blastoderm embryo carrying a reporter construct under the control of the 569 bp *Eag*I-*Eco*RI fragment. This expression corresponds to *h* stripe 5. Note that this construct lacks the putative gt-binding site and that the stripe is expanded posteriorly. Constructs containing only the 273 bp *Pvu*I-*Eco*RI portion of this fragment do not give rise to any significant transcript accumulations, while constructs containing the 302 bp *Eag*I-*Pvu*I fragment produce only a feeble stripe 5 (data not shown).

DISCUSSION

We have compared *h* pair-rule stripe formation in *D. melanogaster* and *D. virilis* [estimated time of divergence 60 MYr, (Beverly and Wilson, 1984)] in order to identify conserved elements within the complex *cis*- and *trans*-regulatory network, which gives rise to striped expression. We have cloned the h locus from *D. virilis*, compared its gross molecular organization with that of *D. melanogaster*, and have examined the degree of functional conservation between homologous stripe elements by making reporter constructs using regulatory fragments from *D. virilis* transformed into *D. melanogaster*. Furthermore, we have determined and aligned the nucleotide sequences of homologous stripe regulatory elements, locating conserved and thus potentially critical regulatory sequences. We demonstrate that these conserved sequences contain potential binding

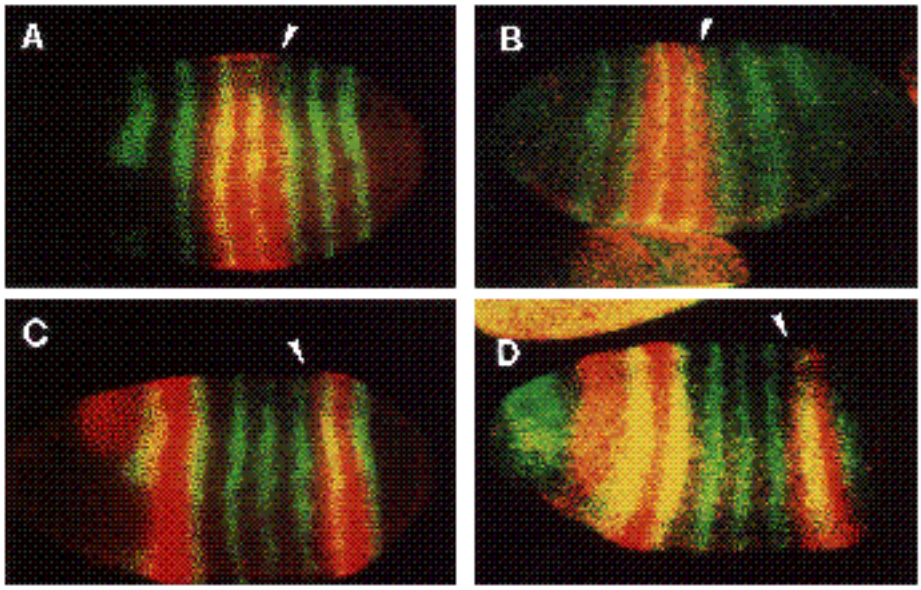

Fig. 5. Conserved relationships between *h* and gap protein expression domains. *D. melanogaster* (A,C) and *D. virilis* (B,D) blastoderm embryos labelled with antibodies against h (green) and either the gap protein Krüppel (Kr) (A,B) or the gap protein giant (gt) (C,D) (both red signals). Regions of overlap appear yellow. The relationships of these protein domains are identical in the two species, even at the fine level of detail observed at stripe borders. For example, h stripe 5 is flanked by Kr to the anterior and gt to the posterior in both species (indicated by arrows).

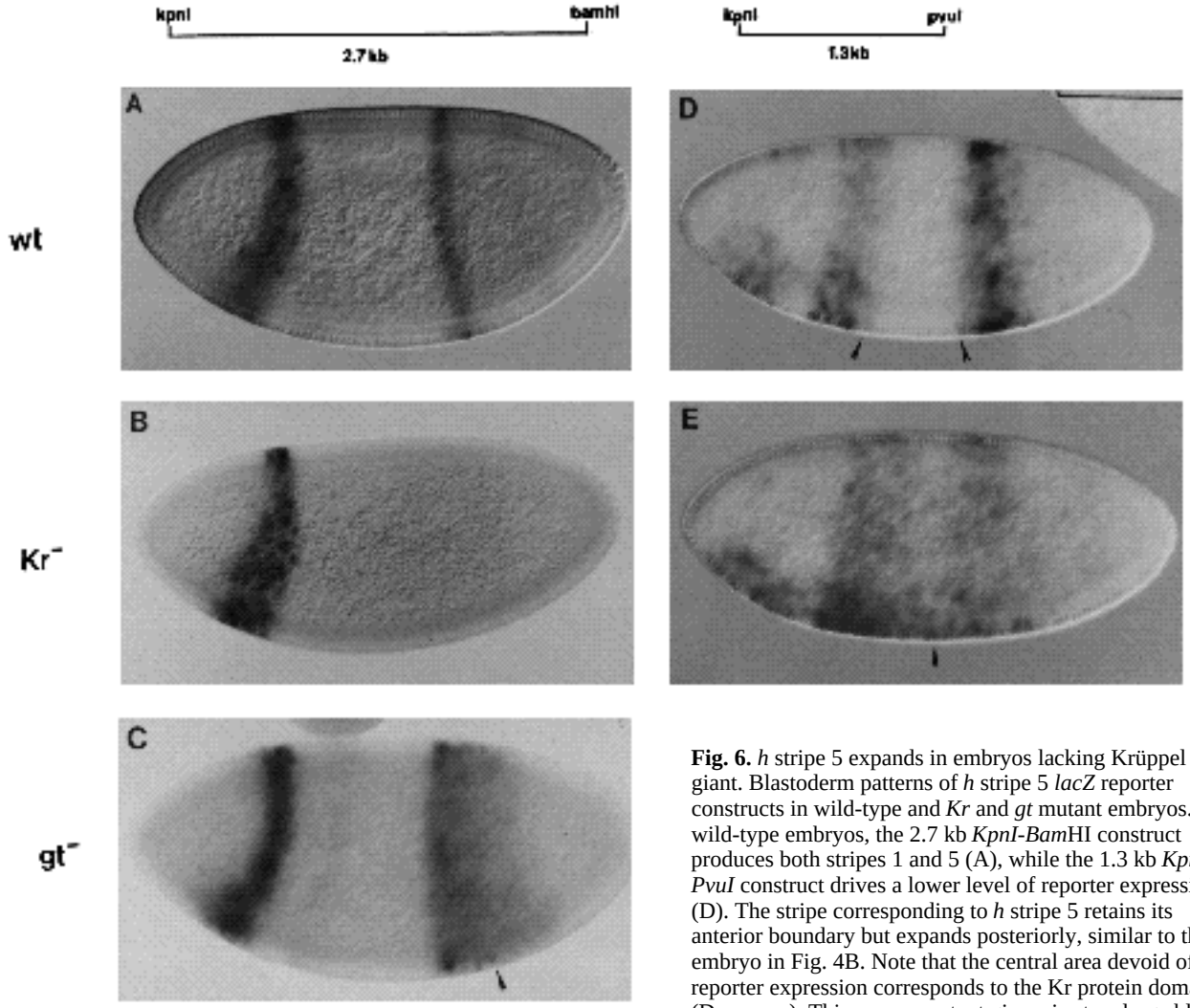

Fig. 6. *h* stripe 5 expands in embryos lacking Krüppel or giant. Blastoderm patterns of *h* stripe 5 *lacZ* reporter constructs in wild-type and *Kr* and *gt* mutant embryos. In wild-type embryos, the 2.7 kb *KpnI-Bam*HI construct produces both stripes 1 and 5 (A), while the 1.3 kb *KpnI-PvuI* construct drives a lower level of reporter expression (D). The stripe corresponding to *h* stripe 5 retains its anterior boundary but expands posteriorly, similar to the embryo in Fig. 4B. Note that the central area devoid of reporter expression corresponds to the Kr protein domain (D, arrows). This same constuct gives rise to a large block of expression in a Kr^- embryo (E), including expression across the wild-type Kr domain (arrow). The larger construct reveals a clearly derepressed stripe 5 in a gt^- embryo (C, arrow), yet does not produce a stripe 5 in a Kr^- embryo (B, see text).

sites for gap and other proteins and serve as a guide for the further dissection of regulatory elements. Finally, we have identified conserved relationships between h stripes and the domains of *trans*-regulatory (gap) proteins in each species and demonstrate that these relationships represent important regulatory interactions. This comparative approach both answers evolutionary questions about the degree of conservation of these elements and focuses genetic and molecular analyses by providing a model for the regulation of *h* stripe 5.

The organization of the *h* locus: conserved stripe elements

The dissection of the *h* regulatory region has revealed an organization that appears modular, with the various pair-rule stripes under the control of discrete regulatory fragments (Pankratz et al.,1990; Howard and Struhl, 1990). To date, *h* and the pair-rule gene *fushi-tarazu (ftz)* are the only zygotic *Drosophila* segmentation genes for which the entire region required for embryonic expression has been defined (Hiromi et al., 1985; Rushlow et al., 1989). An interspecific comparison of the the *ftz* locus has shown that all of the major *cis*-acting elements required for proper *ftz* expression have been highly conserved between *D. melanogaster* and *Drosophila hydei* (Maier et al. 1990). *h*, a primary pair-rule gene, has a much more complex upstream regulatory region than the secondary pair-rule gene *ftz* ; thus whether *h* would exhibit the same degree of regulatory conservation was not clear. In particular, since the various stripe-producing promoter fragments can seemingly produce stripes independent of one another, they could have possibly been rearranged during the radiation of *Drosophila*. Instead, we find not only that each of the defined elements in *D. melanogaster* has a specific counterpart in *D. virilis*, but also that the relative order of these elements has been fully conserved. This strongly suggests that these regulatory fragments are not strictly autonomous and that there is a functional constraint to their relative order. Indeed, attempts to separate the regulatory elements required for *h* stripes 3 and 4 have not been successful (Howard and Struhl, 1990). Furthermore, careful analysis of expression arising from stripe-specific reporter constructs in our lab ond others indicates that such expression does not always precisely correspond to perfectly formed individual h protein stripes; in some cases, stripes may be slightly misregulated and, in others, weak expression corresponding to a particular stripe may arise from more than one fragment of regulatory DNA (Riddihough and Ish-Horowicz, 1991). Taken together, these results indicate that there may be limited, but significant intermingling or even sharing of regulatory elements required for the formation of different *h* stripes. Such relationships would render significant rearrangements of the *h* regulatory region lethal, requiring the organizational conservation of this locus that we have described. The finding that homology exists outside of previously defined regulatory fragments suggests that these sequences also provide some critical regulatory function. While these sequences are not required to drive reporter gene expression corresponding to the seven *h* stripes, they have not been shown to be dispensible for wild-type function.

Our hybridization experiments indicate that significant sequence homology exists between *h* upstream restriction fragments in *D. virilis* and stripe-producing *h* regulatory fragments in *D. melanogaster*. We assayed the degree of functional conservation of these elements by generating transgenic lines carrying a *lacZ* reporter gene under the control of *D. virilis* upstream fragments. The patterns of expression arising from these constructs are virtually identical to those arising from their *D. melanogaster* counterparts, indicating that the *D. virilis* regulatory elements are correctly recognized by all *trans*-acting factors in *D. melanogaster* that are critical for *h* pair-rule patterning. Thus both the *cis*-acting regulatory sequences and the network of *trans*-acting factors that produce *h* pair-rule stripes have been very highly conserved throughout *Drosophila* evolution. While our assay is at the level of gene expression, similar studies have shown that the homologs of the maternal determinant *bcd* (Seeger and Kaufman, 1990) and the pair-rule gene *ftz* (Maier et al., 1990) from other *Drosophila* species are able to rescue *D. melanogaster* embryos mutant for those genes at least partially. Our results indicate that such a result should be possible with the *D. virilis h* homolog as well.

Regulatory sequence conservation: defining *cis*-elements controlling the expression of *h* stripe 5

The results from our hybridization and transformation experiments clearly indicate that significant, functional sequence homology exists between the *h* regulatory regions of *D. virilis* and *D. melanogaster*. Our sequence alignment for the regulatory region controlling *h* stripes 1 and 5 confirms this, revealing several tracts of perfect sequence conservation. Given the 60 MYr separating these species (Beverly and Wilson, 1984), nonfunctional DNA should be completely diverged; thus sequences that have been constrained from diverging very likely represent the *cis*-controls that provide for the functional conservation that we observe. This assumption is strongly supported by the finding that all of the functional regulatory elements of the pair-rule gene *ftz* have been highly conserved between *D. melanogaster* and *D. hydei*, while dispensible upstream fragments contain no detectable homology (Maier et al., 1990). Although we cannot conclude that all of the conserved *h* sequences are involved in embryonic stripe regulation, the extent of this sequence conservation indicates that the number of *trans*-interactions that are required to generate *h* stripes is potentially quite large. We note that these conserved sequences cannot be involved in the regulation of *h* expression in imaginal discs, since the *cis*-elements that control *h* disc expression map well outside of the region containing the embryonic stripe regulatory elements and no disc expression patterns are observed with the largest constructs described here (T. Orenic, L. Held, and S. Carroll, unpublished data).

Beyond illustrating the degree of divergence of such a complex regulatory region over a substantial period of evolutionary time, this comparative alignment of *h* regulatory sequences provides us with a molecular guide for the functional dissection of stripe-producing promoter fragments. Instead of relying on convenient restriction sites, which may randomly split conserved and likely functional sub-

elements, we are able to construct deletion series through this region, which retain the integrity of large conserved tracts. Using reporter constructs that contain these guided deletions, we have defined a 569 bp fragment, which produces *h* stripe 5 and have delimited sequences critical to the activation of this stripe to a 302 bp fragment. In addition to more precisely defining the *cis* elements that govern the formation of this stripe, these constructs also demonstrate that significant tracts of conserved sequence can be deleted without grossly disrupting spatially correct reporter gene stripe formation. However, many of these sequences may in fact be required for the precise patterning and level of native *h* expression and/or represent redundant *cis*-regulatory elements. While some of the conserved regulatory sequence tracts that we have uncovered may be constrained from diverging for other reasons, we are confident that at least a subset of the conserved tracts represent specific binding sites for critical *trans*-acting factors. In their interspecific comparison of intronic and upstream regions of the segment polarity gene *engrailed,* Kassis et al. (1989) demonstrate that certain conserved sequences do indeed contain binding sites for regulatory proteins. Within the regulatory fragment (569bp *Eag*I-*Eco*RI), which produces *h* stripe 5 stripe expression, some of the conserved tracts contain matches of binding site sequences for the gap proteins Kr (Stanojevic et al., 1989; Triesman and Desplan, 1989) and the transcriptional activator FTZ-F1 (Ueda et al., 1990). A reasonable match for the gt-binding site (Small et al., 1991; Capovilla et al., 1992) can be found just proximal to this fragment. Thus not only does the comparative sequence alignment provide a guide for defining minimal regulatory elements, but it can highlight potential binding sites for regulatory proteins. We stress that these are only putative sites at this point; a crucial test will be to determine whether the conserved sequences in the *h* regulatory region in fact correspond to in vivo functional binding sites for these regulatory proteins.

Spatial repression of *h* stripe 5 by the *Kr* and *gt* gap proteins

Several studies have indicated that the gap proteins are key spatial regulators of *h* pair-rule expression (Ingham et al., 1986; Howard, 1988; Carroll and Vavra, 1989; Hooper et al., 1989), and it is generally believed that each pair-rule stripe is the result of regulatory input from more than one gap protein (Pankratz et al., 1990; Riddihough and Ish-Horowicz, 1991). This has been conclusively shown to be the case for *eve* stripe 2. Stanojevic et al. (1991) demonstrated using reporter constructs containing mutagenized binding sites that this stripe is negatively regulated by *gt* on its anterior border and *Kr* on its posterior border, so that the limits of *eve* stripe 2 expression are determined by the domains of these gap proteins. Since the mechanisms that generate *h* pair-rule stripes are apparently highly conserved between *D. melanogaster* and *D. virilis*, and the Kr and gt proteins in particular have been implicated in the regulation of *h* stripe 5 (Pankratz et al., 1990; Riddihough and Ish-Horowicz, 1991), a comparison of the expression patterns of these regulatory genes is a logical extension of our comparative analysis of the upstream region of the *h* locus. By examining the relationships of *h* and gap protein expression in both *D. melanogaster* and *D. virilis*, we find that *h* stripe 5 is flanked by Kr to the anterior and gt to the posterior in both species, essentially a mirror image of *eve* stripe 2 (Small et al., 1991). The conserved spatial relationships between *h* and these gap proteins would not necessarily have been predicted given the numerous overlapping domains of gap protein in the blastoderm embryo (Howard and Struhl, 1990; Kraut and Levine, 1991a,b); that these spatial relationships have been perfectly conserved during evolution strongly indicates that they represent important regulatory interactions. Our analysis of reporter gene expression in mutant embryos confirms this, and we conclude that a situation analagous to *eve* stripe 2 is highly likely for the regulation of *h* stripe 5, with the Kr and gt proteins determining the limits of stripe expression via repression.

The case for giant setting the posterior border is relatively simple. The giant protein precisely flanks *h* stripe 5; when giant is removed, this stripe expands posteriorly. Deletions of stripe 5 regulatory sequences indicate that the site of action for this repression lies proximal to the *Eco*RI site, since stripe 5 is expanded posteriorly in constructs which lack these sequences. This model is supported by the identification of a conserved, potential gt-binding site proximal to the *Eco*RI site. The full role of Kr in the regulation of stripe 5 is less clear, although it appears to be responsible for the anterior border of stripe 5. As with gt, Kr precisely flanks *h* stripe 5; this relationship is highly conserved and consistent with a repressing role. A deletion fragment (1.3 kb *Kpn*I-*Pvu*I) which was guided by our comparative sequence alignment produces stripe 5 expression which is clearly derepressed when Kr is removed, indicating that Kr serves to establish the anterior border of stripe 5.

Interestingly, in the context of a larger regulatory fragment (2.7 kb *Kpn*I-*Bam*HI), stripe 5 fails to express significantly in a Kr^- background. This indicates that even though Kr ultimately limits stripe 5 expression via repression, there are sites proximal to the *Pvu*I breakpoint, which cause a requirement for Kr protein in order for this element to be activated. We cannot yet determine whether this is a direct requirement for Kr or an indirect effect of Kr on some other factor affecting stripe expression, perhaps another gap protein. We note that, prior to achieving their mature striped pattern, *h* transcripts are ubiquitous in the syncytial embryo at a time when gap protein domains are not well defined (Edgar et al., 1989). There may be critical *trans*-interactions required for early stripe activation, which are not represented at the time of stripe maturity. It is possible that the dual effect of Kr on stripe 5 constructs is the consequence of temporally separable programs of stripe activation and repression. Thus, while the repression of stripe borders at cellular blastoderm by gt and Kr appears certain, we do not yet understand how these elements are initially activated. A full understanding of *h* stripe 5 regulation will require the identification and disruption of specific protein-binding sites within *h cis*-regulatory regions.

Although *h* stripe 5 and *eve* stripe 2 have very similar negative controls, they will certainly prove to have quite different positive regulators. The primary activating proteins of *eve* stripe 2 transcription appear to be *bcd* and *hb*

(Small et al., 1991; Stanojevic et al., 1991); these proteins are not expressed in regions of the embryo that could account for the activation of *h* stripe 5. Our search for potential binding sites in the regulatory region for this stripe also turned up a well-conserved site for the FTZ-F1 protein. This protein binds to the zebra element of the pair-rule gene ftz and activates its expression (Ueda et al., 1990). Intriguingly, the potential FTZ-F1-binding site lies in the region that we have shown to be required for stripe 5 activation, suggesting FTZ-F1 may indeed have a role in the activation of *h* stripe 5, similar to its role in activating *ftz* expression. However, since this protein has not been shown to have a spatially restricted pattern in the blastoderm embryo, it does not likely play a region-specific role in regulating striped expression and may serve instead to provide a critical level of expression which is then spatially modulated by other factors. In addition to providing clues to the regulation of *h* stripe 5, the identification of this putative FTZ-F1 site demonstrates the general utility of a comparative approach in deciphering *cis*-controls, particularly in identifying potential interacting proteins for which no regulatory role has been identified through genetic means. Interestingly, there are many conserved *h* regulatory sequences for which we do not know of candidate binding proteins, indicating that more factors are likely to be uncovered.

While we have limited our comparison of the *h* regulatory region to two closely related Drosophilids, the identification of such highly conserved *cis*-elements may also serve as a tool with which to examine regulatory elements in other species. *h* expression is conserved in the housefly *Musca domestica* (Sommer and Tautz, 1991), thus the regulatory elements controlling pair-rule expression are probably highly conserved in all Diptera. However, whether the specific embryonic role and pattern of *h* expression have diverged significantly in organisms other than Diptera remains an open question. Indeed, the *eve* homolog from the grasshopper is apparently restricted in function to the nervous system, indicating that its pair-rule role in *Drosophila* is is rather recently derived from this more ancestral function (Patel et al., 1992). *h* homologs will almost certainly be cloned from other insects and there may be significant differences in the pattern of *h* expression in these species relative to *Drosophila*. Any differences in patterns of expression should be traceable to specific differences in *cis*-acting controls, thus allowing one to investigate how such a complex regulatory region has evolved. Toward this end, it will be of considerable interest to determine whether any of these highly conserved functional *h* *cis*-elements in *Drosophila* are present in more distantly related insects.

We are indebted to Kathy Vorwerk for her technical assistance in obtaining transformants, Teresa Orenic for the use of her *D. melanogaster* transformant lines, Steve Paddock for his expertise in confocal microscopy and image processing, Sherwin Attai for assistance with DNA sequencing and Julie Gates for her monoclonal antibody to *h*. We also thank Allen Laughon and Lisa Nagy for their critical reviews of this manuscript. J. A. L. is a Howard Hughes Medical Institute Predoctoral Fellow. This work was supported by NSF grant DCB-8801814 to S. B. C. and the Howard Hughes Medical Institute.

REFERENCES

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

Beverly, S. M. and Wilson, A. C. (1984). Molecular evolution in *Drosophila* and the higher Diptera II: A time scale for fly evolution. *J. Mol. Evol.* **21**, 1-13.

Capovilla, M., Eldon, E. and Pirrotta, V. (1992). The *giant* gene of *Drosophila* encodes a b-zip DNA-binding protein that regulates the expression of other segmentation gap genes. *Development* **114**, 99-112.

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

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

Carroll, S. B., Laughon, A. and Thalley, B. S. (1988). Expression, function and regulation of the *hairy* segementation gene during *Drosophila* embryogenesis. *Genes Dev.* **2**, 883-890.

Carroll, S. B. and Vavra, S. H. (1989). The zygotic control of pair-rule gene expression. II. Spatial repression by gap and pair-rule gene products. *Development* **107**, 673-683.

Carroll, S. B. and Whyte, J. S. (1989). The role of the *hairy* gene during *Drosophila* morphogenesis: Stripes in imaginal discs. *Genes Dev.* **3**, 905-916.

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

Devereux T., Haeberli, P. and Smithies, O. (1984). A comprehensive set of sequence analysis programs for the VAX. *Nuc. Acids Res.* **12**, 387-395.

DiNardo, S., Kuner, J. M., Theis, J. and O'Farrell, P. H. (1985). Development of embryonic pattern in *Drosophila melanogaster* as revealed by accumulation of the nuclear *engrailed* protein. *Cell* **43**, 59-69.

Dinardo, S. and O'Farrell, P. H. (1987). Establishment and refinement of segmental pattern in the *Drosophila* embryo: spatial control of *engrailed* expression by pair-rule genes. *Genes Dev.* **1**, 1212-1225.

Driever, W. and Nüsslein-Volhard, C. (1988). The bicoid protein determines position in the *Drosophila* embryo in a concentration-dependent manner. *Cell* **54**, 95-104.

Edgar, B. A., Odell, G. M. and Schubiger, G. (1989). A genetic switch, based on negative regulation, sharpens stripes in *Drosophila* embryos. *Dev. Genet.* **10**, 124-142.

Eldon, E. D. and Pirrotta, V. (1991). Interactions of the *Drosophila* gap gene *giant* with maternal and zygotic pattern-forming genes. *Development* **111**, 367-378.

Feinberg, A. P. and Vogelstein, B. (1983) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* **132**, 6-13.

Frasch, M. and Levine, M. (1987). Complementary patterns of *even-skipped* and *fushi-tarazu* expression involve their differential regulation by a common set of segmentation genes in *Drosophila*. *Genes Dev.* **1**, 981-995.

Gaul, U. and Jäckle, H. (1987). Pole region-dependent repression of the *Drosophila* gap gene *Krüppel* by maternal gene products. *Cell* **51**, 549-555.

Goto, T., Macdonald, P. and Maniatis, T. (1989). Early and late patterns of *even-skipped* expression are controlled by distinct regulatory elements that respond to different spatial cues. *Cell* **57**, 413-422.

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

Hoch, M., Gerwin, N., Taubert, H. and Jäckle, H. (1992). Competition of overlapping sites in the regulatory region of the *Drosophila* gene *Krüppel Science* **256**, 94-97.

Hooper, K., Parkhurst, S. and Ish-Horowicz, D. (1989). Spatial control of *hairy* protein expression during embryogenesis. *Development* **107**, 489-504.

Howard, K. (1988). The generation of periodic pattern during early *Drosophila* embryogenesis. *Development* **104 Supplement**, 35-50.

Howard, K. and Ingham, P. W. (1986). Regulatory interactions between the segmentation genes *fushi tarazu*, *hairy*, and *engrailed* in the *Drosophila* blastoderm. *Cell* **44**, 949-957.

Howard, K. R. and Struhl, G. (1990). Decoding positional information: regulation of the pair-rule gene *hairy*. *Development* **110**, 1223-1232.

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

Ingham, P. W., Ish-Horowicz, D. and Howard, K. R. (1986). Correlative changes in homeotic and segmentation gene expression in *Krüppel* embryos of *Drosophila*. *EMBO J.* **5**, 1527-1537.

Ingham, P. W. and Gergen, J. P. (1988). Interactions between the pair-rule genes *runt, hairy*, *even-skipped,* and *fushi tarazu* and the establishment of periodic pattern in the *Drosophila* embryo. *Development* **104 Supplement**, 51-60.

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

Jiang, J., Hoey, T. and Levine, M. (1991). Autoregulation of a segmentation gene in *Drosophila* : Combinatorial interaction of the *even-skipped* homeo box protein with a distal enhancer element. *Genes Dev.* **5**, 265-277.

Kassis, J. A., Desplan, C., Wright, D. K. and O'Farrell, P. (1989). Evolutionary conservation of homeodomain-binding sites and other sequences upstream and within the major transcription unit of the *Drosophila* segmentation gene *engrailed*. *Mol. Cell. Bio.* **9**, 4304-4311.

Klessig, D. and Berry, J. O. (1983). Improved filter hybridization method for detection of single copy sequences in large eukaryotic genomes. *Plant Mol. Biol. Rep.* **1**, 12-18.

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

Kraut, R. and Levine, M. (1991a). Spatial regulation of the gap gene *gaint* during *Drosophila* development. *Development* **111**, 601-609.

Kraut, R. and Levine, M. (1991b). Mutually repressive interactions between the gap genes *giant* and *Krüppel* define middle body regions of the *Drosophila* embryo. *Development* **111**, 611-622.

Maier, D., Preiss, A. and Powell, J. R. (1990). Regulation of the segmentation gene *fushi tarazu*has been functionally conserved in *Drosophila*. *EMBO J.* **9**, 3957-3966.

Michael, W. M., Bowtell, D. D. L. and Rubin, G. M. (1990) Comparison of the *sevenless* genes of *Drosophila virilis* and *Drosophila melanogaster*. *Proc. Nat. Acad. Sci.* **87**, 5351-5353.

Murre, C., McCaw, P. S. and Baltimore, D. (1991). A new DNA binding and dimerization motif in Immunoglobulin Enhancer Binding, daughterless, MyoD and myc proteins. *Cell* **56**, 777-783.

Nelson, H. B. and Laughon, A. (1992). Glial architecture and development in *Drosophila* embryos. *Development* (in press).

Paddock, S. W., Langeland, J. A., Devries, P. and Carroll, S. B. (1993). Three color immunofluorescence imaging of *Drosophila* embryos by laser scanning confocal microscopy. *Biotechniques* (in press).

Pankratz, M. J., Seifert, E., Gerwin, N., Billi, B., Nauber, U. and Jäckle, H. (1990). Gradients of *Krüppel* and *knirps* gene products direct pair-rule gene stripe patterning in the posterior region of the *Drosophila* embryo. *Cell* **61**, 309-317.

Patel, N. H., Ball, E. B. and Goodman, C. S. (1992). Changing role of *even-skipped* during the evolution of insect pattern formation. *Nature* **357**, 339-342.

Pignoni, F., Baldarelli, R. M., Steingrimsson, E., Diaz, R. J., Patapoutian, A., Merriam, J. R. and Lengyel, J. A. (1990). The *Drosophila* gene *tailless* is expressed at the embryonic termini and is a member of the steroid receptor superfamily. *Cell* **623**, 151-163.

Riddihough, G. and Ish-Horowicz, D. (1991). Individual stripe regulatory elements in the *Drosophilahairy* promoter respond to maternal, gap, and pair-rule genes. *Genes Dev.* **5**, 840-854.

Rushlow, C. A., Hogan, A., Pinchin, S. M., Howe, K. M., Lardelli, M. and Ish-Horowicz, D. (1989). The *Drosophilahairy* protein acts in both segmentation and bristle patterning and shows homology to *N-myc*. *EMBO J.* **8**, 3095-3103.

Seeger, M. A. and Kaufman, T. C. (1990). Molecular analysis of the *bicoid* gene from *Drosophila pseudoobscura* : Identification of conserved domains within coding and noncoding regions of the *bicoid* mRNA. *EMBO J.* **9**, 2977-2987.

Small, S., Kraut, R., Hoey, T., Warrior, R. and Levine, M. (1991). Transcriptional regulation of a pair-rule stripe in *Drosophila*.. *Genes Dev.* **5**, 827-839.

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

Spradling, A. C. (1986). P-element mediated transformation. In Drosophila, *a Practical Approach* (ed. D. B. Robert). Oxford: IRC Press Limited.

Stanojevi, D., Hoey, T. and Levine, M. (1989). Sequence-specific DNA binding activities of the gap proteins encoded by *hunchback* and *Krüppel* in *Drosophila*. *Nature* **341**, 331-335.

Stanojevic, D., Small, S. and Levine, M. (1991). Regulation of a segmentation stripe by overlapping activators and repressors in the *Drosophila* embryo. *Science* **254**, 1385-1387.

Sturdier, F. W., Rosenberg, A. H. and Dunn, J. J. (1986). Use of T7 RNA polymerase to direct the expression of cloned genes. *Meth. Enzymol.* **185**, 60-89.

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

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

Treisman, J. and Desplan, C. (1989). The products of the gap genes *hunchback* and *Krüppel* bind to the *hunchback* promoter. *Nature* **341**, 335-337.

Ueda, H., Sonoda, S., Brown, J. L., Scott, M. P. and Wu, C. (1990). A sequence-specific DNA-binding protein that activates *fushi-tarazu* segmentation gene expression. *Genes Dev.* **4**, 624-635.

Wainwright, S. M. and Ish-Horowicz, D. (1992). Point mutations in the *Drosophilahairy* gene demonstrate *in vivo* requirements for basic, helix-loop-helix, and WRPW domains. *Mol. Cell. Biol.* **12**, 2475-2483.

Williams, J. A., Bell, J. B. and Carroll, S. B. (1991). Control of *Drosophila* wing and haltere development by the nuclear *vestigial* gene product. *Genes Dev.* **5**, 2481-2495.

(Accepted 3 November 1992)