

Molecular genetics of *aristaleless*, a *prd*-type homeo box gene involved in the morphogenesis of proximal and distal pattern elements in a subset of appendages in *Drosophila*

Kay Schneitz,¹ Patrick Spielmann, and Markus Noll

Institute for Molecular Biology II, University of Zürich, CH-8057 Zürich, Switzerland

Viable *aristaleless* (*al*) mutations of *Drosophila* affect pattern elements at both ends of the proximodistal axis in a subset of adult appendages. The *al* gene has been cloned and identified by P-element-mediated germ-line transformation with a genomic DNA fragment, which rescues a lethal mutation of *al* as well as aspects of the adult *al* phenotype. The *al* gene contains a *prd*-type homeo domain and a Pro/Gln-rich domain and, hence, probably encodes a transcription factor. Its transcript distribution in third-instar imaginal discs closely corresponds to the anlagen of the tissues that later become visibly affected in adult *al* mutants. The striking similarity of a bimodal *al* expression in different imaginal discs indicates that *al* is under the control of a "prepattern," which is shared at least among antennal, leg, and wing discs. The *al* gene is also transcribed during embryogenesis. Apart from a function in the ontogeny of specific larval head and tail organs, its embryonic transcript pattern suggests a possible role in early imaginal disc development.

[Key Words: *aristaleless*; appendage morphogenesis; imaginal disc; pattern formation; *prd*-type homeo domain]

Received September 17, 1992; accepted October 26, 1992.

An overwhelming multiplicity of species is found in the insect world, as documented by the enormous variety of different sizes, puzzling shapes, and intricate patterns of their outer manifestations. A primary goal of developmental biology is the description of the ontogeny of such complex forms and patterns at the molecular level.

Two characteristics of the arthropod phylum are the tendency toward regional specification and the segmented body, with each segment bearing a pair of ventrally located and articulated appendages (Brusca and Brusca 1990). Thus, it is commonly accepted that insects have evolved from an annelid-like ancestor, consisting of an array of fairly similar segments, to their current specialized forms through the grouping of adjacent segments into three body parts (tagmosis) and a high degree of specification of single segments that is reflected, for example, by the variety in external morphology of insect appendages (Brusca and Brusca 1990 and references therein).

In *Drosophila*, the classical genetic studies on segment identity (Lewis 1963, 1978; García-Bellido 1975, 1977; Ouweneel 1976; Kaufman et al. 1980) and the pioneering

work on the generation of segments (Nüsslein-Volhard and Wieschaus 1980) have initiated a vast number of molecular studies that have provided basic insights into the gene networks and mechanisms regulating the establishment and specification of segments (for review, see Akam 1987; Ingham 1988). However, comparatively little is known about the nature of the processes generating a particular three-dimensional pattern element of a segment, like a leg or an antenna. Thus, although many genes are known that affect leg formation (Tokunaga and Gerhart 1976; Held et al. 1986), only recently detailed molecular studies, combined with genetic tools, have begun to analyze their function. Such studies showed that genes like the homeo box gene *Distal-less* (*Dll*), formerly called *Brista* (Sunkel and Whittle 1987), and *rotund* (*rn*) play an important role in establishing the pattern along the proximodistal axis of appendages (Kerridge and Thomas-Cavallin 1988; Agnel et al. 1989; Cohen and Jürgens 1989; Cohen et al. 1989; Cohen 1990).

In an attempt to understand the events of pattern formation and morphogenesis in *Drosophila* that follow determination and specification of segments, we focused our attention on the *aristaleless* (*al*) gene for several reasons. First, viable *al* mutants are known to exhibit characteristic and localized pattern disturbances in a subset of the appendages of the adult fly (Lindsley and Zimm

¹Present address: Department of Cellular and Developmental Biology, Harvard University, Cambridge, Massachusetts 02138 USA.

1992). Various *al* mutations reduce or remove the most distal part of the antennae and legs, the arista and claws (Fig. 1). Furthermore, *al* mutants display a shortened scutellum, as well as minor effects on wing venation. Second, *al* belongs to a particular subset of morphogenetic genes residing in the cytological interval 21C1,2; 21D1,2. This set includes *al*, *expanded* (*ex*), and *dachsous* (*ds*) (Stern and Bridges 1926; Lewis 1945; Lindsley and Grell 1968; Korochkina and Golubovsky 1978). Besides their close genetic linkage, these three genes interact with a locus, defined by a closely linked yet distinct, though still poorly characterized, "genetic regulator" function *R*, which fails to complement *al*, *ex*, and *ds* and thus behaves as an activator of the three genes (Golubovsky and Kulakov 1978; Korochkina and Golubovsky 1978). An additional property of these genes might lie in a common origin for their adult phenotypes. Phenotypic abnormalities associated with *ex* and *ds* mutations have been interpreted in terms of defects in growth behavior of the mutant imaginal discs during larval development (Waddington 1941, 1942, 1943, 1953). An altered growth pattern could also account for at least some of the phenotypic effects of *al* (Tokunaga and Stern 1969). Finally, both *al* and *ds* were shown to interact with a variety of homeotic genes (Villée 1945, 1946; Mglinets and Ivanov 1975; Mikuta and Mglinets 1978).

In a first step toward an understanding of the phenomena described above, the *al* gene has been cloned, identified, and characterized. In particular, its expression patterns are described in detail in whole-mount embryos and late third-instar imaginal discs.

Results

The *al* phenotype

A limited set of particular body parts are visibly affected in *al* flies (Lindsley and Zimm 1992; for a careful description of the wild-type adult structures, see Hodgkin and Bryant 1978). The most prominent effects are illustrated in Figure 1. Compared to a wild-type arista (Fig. 1A), the mutant arista is variably reduced in size (Fig. 1D), depending on the allelic combination. The other obviously affected part is the posterior mesonotum (Fig. 1B,E), the scutellum, which is considerably reduced in size (Tokunaga and Stern 1969). In addition, and again depending on the allelic combination, one may find supernumerary scutellar bristles (up to a total of eight; Fig. 1E) or the posterior scutellar bristles pointing anteriorly (Lindsley and Zimm 1992). In adults carrying strong viable alleles, occasionally a reduction or complete absence of the tarsal claws is observed (Lindsley and Zimm 1992), whereas the empodium and the pulvilli appear unaffected (Fig. 1C,F). The remaining mutant features include irregularities of the sternopleural bristles and of the first and second longitudinal wing vein (Lindsley and Zimm 1992; K. Schneitz, unpubl.). No cuticular phenotype was detected in embryos carrying lethal allelic combinations of *al* (see Discussion).

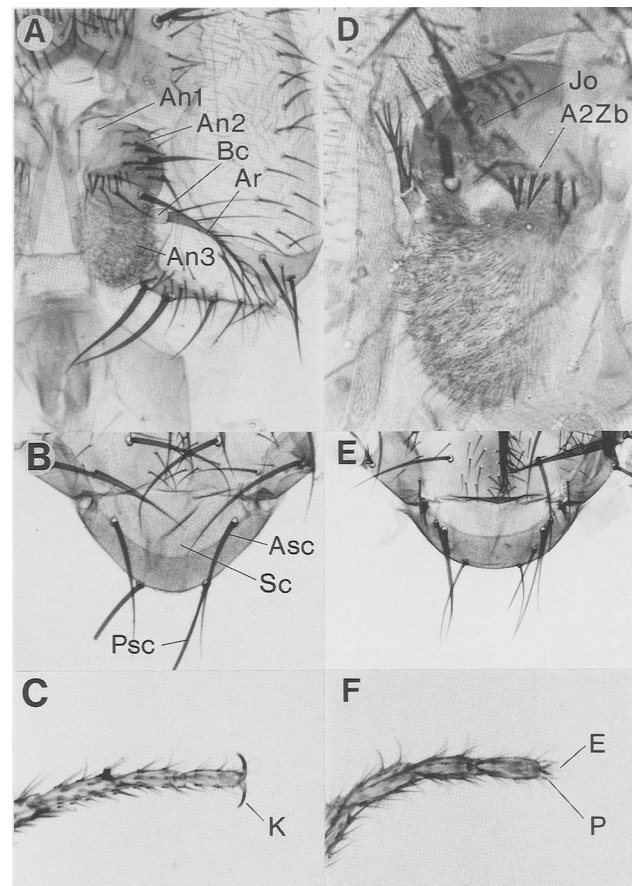
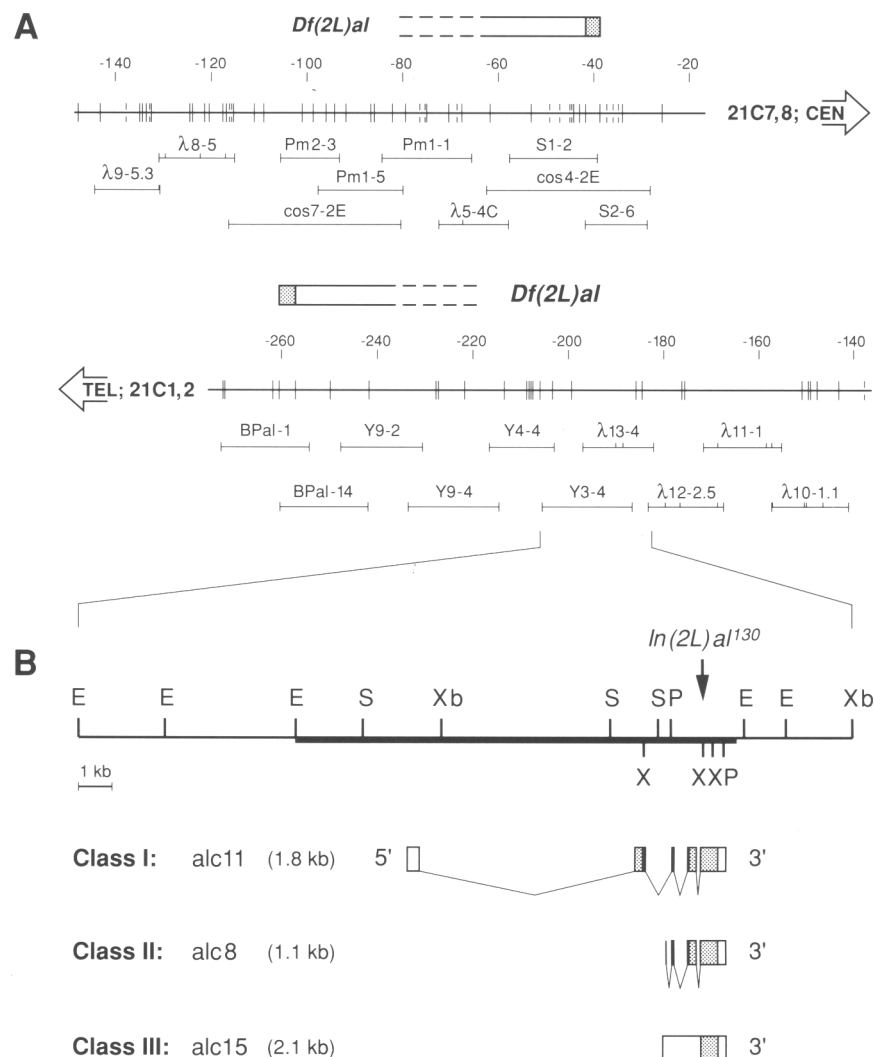


Figure 1. The *al* phenotype. Adult cuticle preparations illustrate corresponding wild-type (A–C) and *al* mutant (D–F) structures. The region of the antenna with the arista (A,D), the posterior mesonotum and scutellum (B,E), and the distal leg with the claw organ (C,F) are shown. (A2Zb) Zahnborsten of second antennal segment; (An1–An3) first to third antennal segments; (Ar) arista; (Bc) basal capsule; (Jo) Johnston's organ; (Asc, Psc) anterior and posterior scutellar bristle; (Sc) scutellum; (E) empodium; (K) unguis (claw); (P) pulvillus.

Cloning and identification of *al*

The *al* gene maps genetically at 0.0 map units and cytologically to 21C1,2 on the left arm of the second chromosome (Stern and Bridges 1926; Lewis 1945; Korochkina and Golubovsky 1978). To clone the *al* gene, a genomic walk was initiated from a previously isolated phage (containing position 0 of our walk and mapping cytologically to chromosomal band 21C7,8) and extended to cover the entire *Df(2L)al* deficiency (Fig. 2A; portions of the walk were generous gifts of Rick Jones and Corrado Caggese). The *al* gene was localized within the walk by in situ hybridization to *In(2L)al¹³⁰* polytene chromosomes and subsequent isolation from a genomic *In(2L)al¹³⁰* library of a region covering the proximal breakpoint of *In(2L)al¹³⁰* (21C1,2/21E1,2; E. Lewis pers. comm.) and, hence, containing *al* sequences proximal to the distal inversion breakpoint. The 0.29-kb *Xho*I fragment containing the distal breakpoint [Fig. 2B] was in-

Figure 2. Chromosomal walk uncovered by *Df(2L)al* and transcriptional organization of *al*. (A) Chromosomal walk of the 21C1,2/21C7,8 interval. Below is an *EcoRI* restriction map (scale in kb; broken lines indicate *EcoRI* sites of neighboring fragments whose order has not been determined) of the genomic region comprising the 220 kb deleted by *Df(2L)al*. Horizontal lines indicate the extent of phage or, in the case of cos4-2E and cos7-2E, cosmid inserts of a chromosomal walk. Clones λ 5-4C, λ 8-5 to λ 13-4 [derived from an isogenic II (*dp cn bw*) genomic library prepared in EMBL3; the small vertical lines represent internal *SalI* sites] and the cosmid clones cos4-2E and cos7-2E are part of a generous gift of Rick Jones (Southern Methodist University, Dallas, TX). All other inserts were obtained from a genomic library prepared in EMBL4 [Frischauf et al. 1983]. The cytological location of the chromosomal walk is indicated next to the open arrows, orienting the telomere of the left arm of the second chromosome to the left and the centromere to the right. The extent of the region deleted by *Df(2L)al* is represented by an open bar above its ends, with the stippled regions containing the deficiency breakpoints. Proximal and distal breakpoints lie within 3.0- to 4.0-kb *EcoRI* fragments. (B) Enlarged map of the *al* locus indicating its transcriptional organization. A partial restriction map of the *al* region is shown in which sites also found in isolated *al*-cDNAs are indicated below the line. The bold horizontal bar designates a 13.3-kb genomic fragment (proximal *EcoRI* fragment of Y3-4) used for rescue experiments of *al* mutants. The vertical arrow marks the distal breakpoint of the inversion *In(2L)al*¹³⁰ localizing the gene. Its precise position has been determined by cloning and sequencing (Fig. 3A). Below the map, the splicing pattern of three different classes of *al*-cDNAs is depicted. Only the longest cDNA representative of each class is shown. Class I consists of alc1, alc3, alc4, alc10-alc13, and alc17; class II consists only of alc8, class III consists of alc2, alc7, alc15, and alc18. Sequences with an open reading frame are shown as stippled; the *prd*-type homeo box is shown in black. The orientation of the *al* transcripts is indicated by the 5' and 3' ends. (E) *EcoRI*; (P) *PstI*; (S) *SalI*; (X) *XhoI*; (Xb) *XbaI*.



cluded in all cDNAs isolated from this region (see below). Therefore, the inversion interrupts a transcription unit that was later confirmed by genomic rescue experiments to code for the *al* transcript (see below). The precise location of the distal breakpoint was determined by sequencing (Fig. 3A).

Transcriptional organization of the *al* gene

Figure 2B depicts a schematic overview of the transcriptional organization of the *al* locus. On the basis of isolated cDNAs, at least three different groups of transcripts can be discriminated. Class I cDNAs are derived from an mRNA that encodes the longest open reading frame and includes a homeo domain. The longest cDNA representing this class is the ~1.8-kb alc11 [without poly(A) tail], which lacks ~0.4 kb of untranslated leader

sequence (Fig. 3A) as judged by Northern blot analysis (Fig. 5, below).

Only one member has been isolated from the second class, alc8, which deviates in its first 11 bp but shares the last three exons with class I cDNAs. These 11 bp are found in the genomic sequence of the second intron of class I primary transcripts (Fig. 3A). Hence, alc8 could have arisen from a splicing artifact or by a regular alternative splice event generating an altered amino terminus of the homeo domain. Consistent with either possibility is the fact that the 11-bp piece ends at a potential splice donor site (Fig. 3A). Although no genomic DNA sequence upstream of the 5' end of alc8 (within the second intron of class I transcripts) could encode for the remaining amino-terminal part of a homeo domain and no signal was detected upstream of the second exon after low stringency hybridization with the homeo domain of *prd*,

A

1 atttattgcactttgtgctggggtttattcaattacgcccaatcgaagaggcgagagagagggccatcgaggagcagagagcggtagagagcagctccgccattatttcgctccgtc 120
 121 gaatgctgctttcccataggccgaacccagaaagaggcgaggtcccgtccggccacctgcccgtccacctgcgacacctgcgacgatcgatcgacacgctccgacgcgctccc 240
 241 tctcgctctgctggcaactcgaacacactaaaaacactgggagggcagaaacacactctcgcaacattggaacaacacacactcaaccgggtgtgtgttcagaccacacagagttttcacct 360
 361 ggcgcaccaactgctcaccgtgtagttttgtcGCCGGCTCTGGAACAGCCCGGTCATCTCGTCGGTTCGGTTCGGATTCCGATTTCGAATAGTCGAGCTGGGGATATATTGTTGTTTCCGGG 480
 481 TTAAGTGCCTACGATTTCGGAAGTGTGGTATCTCGCGCTAGTACAAAGTTATACCCATCAAAGGTGTATCGCATCAAAGTGAACGCTTGGTACTCGGTATCTGGTGTGGAAGAA 600
 601 GTGGACAAAAGACCGTCGGCAAAAGTTACTGTTGGTGTGAGCTACAGGGCTTATTGGCCGGTTAAATAGGAAGAGAGATTTCAGAGGCCAACATTATCAAGACCATCCGTACAATA 720
 721 TCTGGCGATTAAAGGAGAGAAGGAAGgttagtctcagacatt... ca 6.3 kb ...ttatctctgtgtagCCGCTGTCTTCGAGATCAGACTAGATCGCCCTCTGGATTATG 7140
 7141 G I S E E I K L E E L P Q E A K L A H P D A V V L V D R A P G S S A A S A G A A 41
 GGCATATCGGAGAAATCAAGCTGGAGAACTACCGCAAGAAGCCAAGTAGCGCACCCGATGCCGTAGTGTGGTGTGATCGTCCCGAGGATCGTGGCCGCCCTCCGCCGCCGACG 7260
 7261 L T V S M S V S G G A P S G A S G A S G G T N S P V S D G N S D C E A D E Y A P 81
 CTGACGGTTTCGATGTGGTCTCGGGCGGAGCGCCAGTGGAGCAAGTGGGGCAGTGGCGGCACCAATAGCCCGTTTCGGATGGAAACAGCGACTCGCAAGCGGACGAGTACGCCCG 7380
 7381 K R K Q R R Y R T T F T S F Q L E E L E K A F S R T H Y P D V F T R 115
 AAGCGGAAGCAGCGGCTATCGCACAACTTACCAGCTTTCAGCTCAGGAGCTGGAGAAGGCATTCGCGGACCCATTATCCGACGTGTTTACGAGgtgagtaacatatgacgaa 7500
 7501 atgcgaataaggtcgaagcggcagaaaatgagtcgccaagagaaaagaatttgcaataatcgctcatcgaggcgagcaaaaggtggaaggtgattctgtgttttacgataccaagac 7620
 7621 gaatatgggaacgatcggtgataataataatgccataacgacgcgacccccctgagtgagtcgatcgatcttattcggagcagcggtttatacccatgggttttgctgggactcctccc 7740
 7741 gagctagccaaaatagcagcatctcgccctaagagactctaacgatcgcaattagcatgctgcaatgtggttgaagacagccacagtaattgccataaatttaagtcaggcactttt 7860
 7861 Sal I 5' end alc15 7980
 7981 gccatgctgctggcagttttcaattcattaggtcgcgattcatttaaatgcagcaatataaagaaaataaatcaattcgacaatgatctctgaataagcttcaattgctttggaatg 8100
 8101 agaagcgcagaggtgagctaaagtcagcgttttaacaacgatgacttttgcatttttgtaataaaattggttcgctaatacttttgatatagaagacagtttgacggcggttagtt 8220
 8221 Pst I 5' end alc2 5' end alc12 130
 aaatgctgcgcagctgctccctgcagttttccagctaaccaaggtttctcttcttcttcttctgcagAGAGGAGCTGGCGATGAAAATTGGATTACCGAGGCGCATTCAGgtgagca 8340
 8341 tttcgcaagagggaattcagggtcgatgctgcgcagattttgtggttaattcgctgccccagctgcaccaacaatgtgaaaattgaaaataatggagcatgtagctactaatga 8460
 8461 aatcgagacgggactagacgagacagcgagcagatccgcggtggagaaattcatttaattctgttttagtgccagtcattacatgcccggtcggtcatttttggtcggaagtcgag 8580
 8581 gdfctcttttaataactgttcatgctaaacaacattatttgctacatccatttccgtcgagacttcaactgctcgggactccccgagcaatcaggagccggttcttgggaaggactcc 8700
 8701 atattcagattgttatgttaaaaactaataaaggggcggttttcaataaacagcccatataattgatgctccgcatcgagcacatatgaacaggatggtcggaacttgatgcagata 8820
 8821 V W F Q N R R A K W R K Q E K V G P Q S H P 152
 cagtagcctgcttgagcacataactcataaagccattctcacttactttccagGTTGGTTCCTCAAAATCGCGCGCCCAAGTGGCGCAAGCAGGAGAGGTTGGCCCCAGTCGCATC 8940
 8941 Y N P Y L P G G A A T M Q T V V G A A L P P N P F T H L G F Q L R K P F D A Q H 192
 CGTATAATCCCTACCTGCCCGCGCGCGCCGCCACCATGCAGACGCTAGTTGGGGCGCCCTACCGCCCAATCCCTTCACTCACTGGGCTTCCAGTGCAGCAAGCCCTTCGACGCCAAC 9060
 9061 A A N L A A F R Y P H L S A A P M I P S G Y F N Q F Q R 220
 ATGCCGCAACCTTGCCGCTTCCGCTACCGCACCTCTCTGCGGCCCATGATTCCGTCGGGTACTTTAAACCAATCCAGAGgtgagtaatcagctctatcctgatttttaagaaaag 9180
 9181 tegttaaaaaattttataaataataaaccagatctatttatacgttatttgaaggttccattcttaatttatctgcaattttctgagGGCTCCTCCCATATGTTACCACATGGAA 230
 2301 A G M Y S P S S F Q S L L A N M T A V P R G P P L G K P P A L L V G S P D L H 270
 TGGCTGGTATGACTCTCCCTCGAGCTCATTTCAATCCTTGCTGGCAATATAGCGCGCTTCCGCGCGGCGCTCCACTTGGAAAACCTCTGCATCTAGTGGGATCTCCGATCTGC 9420
 9421 S P N H M L A S P P T S P A S G H A S Q H Q Q H P T A H P P P P Q A P P Q M P V 310
 ATTCGCCCAACCATGCTCGCTCACCACCCACATCGCTCGCTCGGGCAGCCTCCAGCATCAACAACATCAACCGCTCACCCCGCCACCTCAGGCGCTCTCAATGCCAG 9540
 9541 G V Q C A Q L S P C Q Q L V G I A L T Q Q A S S L S P T Q T S P V A L T L T L S H S P 350
 TTGGAGTCCAGCTGCTCAGCTTTCTCCCGAGCAGCTGGTTGGGATCGCTTGACCCAACAGGCTCGAGTCTATCGCCACCCAGACTTCCCGGTGGCCCTGACTTTGTGCACTCTC 9660
 9661 Q R Q L P P P S H Q A P R H L P E L Q R H P R T G A P P P L P H C V END 384
 CCCAACGACAACTGCCCTCCACCATCGCATCAGGCTCCGCGCCACCTCCAGAGGTGCAACGCCACCGAGGACAGGCGCACTCTCCATTGCGGCACTCGCTCTGAAGCGGAGAGCA 9780
 9781 CGAACTGAAACTGGAAGTGTTCGCTCAGAATGGACATGGAACAGATGTGGTCAGTAGTGTGGAGCAATTAAGGTTCAAGGATGCAATGGACGATACGAAGGTACGCGAGGATCCCC 9900
 9901 Pst I poly A (3' end alc6) 10020
 AAACCAAGTGCAAAAGATTTCCAATTCAGTGCAGTCTAACTAATAAGTTATTCAATTGTATTACCTAAGCTCCAATTTTAAATAAATCAATAAGTAAATTaaccatattgtt 10020
 10021 ttgggcaaaactgatattagccacaaataggagaaatagcttatttgtgaactaagataaataactgaataaagttcatgcaatcactagttcaaaaataaactacttttaccoc 10140
 10141 attttgagaagaccaagcacaccaactcttcatctctcattattgtctcttctgtaattccacaa 10206

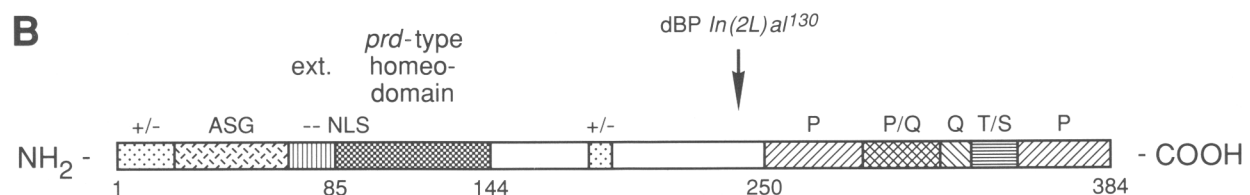
B

Figure 3. (See following page for legend.)

we cannot exclude that the 11 bp are part of a minixon, and additional exons encoding the missing portion of the homeo domain are located in the large intron or upstream of the 5' end of class I transcripts. Furthermore, it might be noteworthy that a potential translational start site is found 15 bp upstream of the beginning of *alc8*. Starting from this ATG, translation would result in a truncated homeo domain that is missing the amino-terminal half of the canonical homeo domain but contains the helix–turn–helix motif.

Class III transcripts exhibit a rather astonishing sequence. Four cDNAs of different sizes belong to this class and seem to be derived from an unspliced message overlapping the 3' portion (last three introns and exons) of class I transcripts. Its longest representative is *alc15*, starting 171 bp upstream from the 5' end of *alc8*. No reasonably long open reading frame could be derived from its genomic sequence upstream from the splice acceptor site of the third intron of class I cDNAs. We do not know whether there is any biological function associated with class II and class III transcripts. However, there are several ATGs in the open reading frame of the last exon of class I cDNAs, which is identical to that of class II and class III cDNAs (Fig. 3A). Hence, at present, we cannot eliminate the possibility that one or several of these ATGs is used in class II or class III transcripts to produce a small Pro- and Gln-rich protein (Fig. 3).

Rescue of the *al* lethality and features of the adult phenotype

In a first attempt of a P-element-mediated rescue of *al* mutants, a 13.3-kb *EcoRI* fragment derived from the right end of Y3-4 was used (Fig. 2). This fragment is able to rescue the lethality (Table 1) and some aspects of the adult *al* scutellar phenotype associated with the inversion *In(2L)al¹³⁰*. The rescued flies still exhibit reduced arista but are nearly wild type with regard to the scutellar phenotype (not shown).

Interestingly, flies homozygous for the inversion *In(2L)al¹³⁰* that carry only one copy of the rescue con-

struct show new additional features of the wing phenotype. These animals display gaps in the most proximal part of the second longitudinal vein (LII), a feature very reminiscent of the *telegraph* (*tg*) phenotype (Stern and Bridges 1926). Because the *tg* mutation is genetically closely linked to *al*, we suggest that these two mutations are allelic. In addition to the effect on the LII vein, we observe an altered wing shape in that the wing blade is bent downward at its posterior tip. In contrast to the mutant wing blade, the gap in LII is only rarely observed and only in flies harboring only one copy of the rescue construct (data not shown). The fact that the rescue construct fails to rescue some of the adult *al* phenotypic features in transgenic flies provides an internal control for the homozygous *In(2L)al¹³⁰* genotype of the rescued flies. In addition, the observed partial rescue of the scutellar phenotype proves that the *al* phenotype of the *In(2L)al¹³⁰* chromosome is caused by an effect on the transcription unit interrupted by the inversion and not by a mutation elsewhere on the chromosome. Homozygous *In(2L)al¹³⁰* flies without the genomic rescue construct have never been detected, suggesting that this *al* mutation is lethal.

DNA and protein sequence of *al*

The genomic DNA sequence of the five exons and the last three introns of the *al* transcription unit represented by class I cDNAs was determined on both strands, whereas only a small portion of the 6.3-kb intron was sequenced. In addition, the cDNA sequence was obtained from both strands of class I cDNAs while only the 5' ends of other cDNAs were sequenced. The result, including the conceptual translation of class I transcript sequences derived from the genomic sequence, is shown in Figure 3A. The longest class I cDNA, *alc11*, starts at nucleotide 391 and has a length of 1791 bp. As judged from the length of the major transcript detected by Northern blot analysis (Fig. 5, below), it is not full length but lacks ~0.4 kb at its 5' end. Four introns of 6.3 kb, 806 bp, 543 bp, and 126 bp are found. Transcripts correspond-

Figure 3. (A) DNA sequence of the *al* gene and corresponding amino acid sequence of the putative Al protein. The genomic DNA sequence between 390 bp upstream of the 5' end of the longest class I cDNA (*alc11*) and 199 bp downstream of the poly(A) addition site is shown with the exception of the 6.3-kb intron. Exon sequences found in class I cDNAs are represented by uppercase letters. The numbering of nucleotides (indicated in both margins) following the 6.3-kb intron is approximate, as only a small portion of it has been sequenced. Amino acid positions of the putative Al protein are numbered only in the right margin. The sequence corresponding to the *prd*-type homeo box, interrupted by two introns, is underlined twice and its amino-terminal extension is underlined once. They contain a potential nuclear localization signal labeled by a broken line. The 5' ends of *al*-cDNAs are indicated above the sequence. Asterisks mark the first 11 bp of the 5' end of *alc8*. The canonical poly(A) addition signal is underlined. Note that the site of poly(A) addition in *alc1* is 2–6 bp upstream of that found in *alc6* and that the sequence of *alc11* differs from the genomic sequence at two nucleotides that would result in an altered protein, namely at nucleotides 9372 (C → A) and 9575 (G → T). The distal breakpoint of the inversion [*dBPIIn(2L)al¹³⁰*] is marked by a vertical line and an arrow indicating the orientation of the inversion. Because only the proximal breakpoint of the inversion has been sequenced, the probability that the true breakpoint is *N* base pairs downstream of the one indicated is about $\frac{1}{2}^{2N}$. Restriction sites shown in Fig. 2B are denoted above the sequence. (B) Schematic representation of Al protein domains. The primary sequence of the Al protein suggests its composition of several domains: (+/-) charged domain; (ASG) domain rich in Ala, Ser, and Gly; (ext.) negatively charged (- -) amino-terminal extension of *prd*-type homeo domain; (NLS) nuclear localization signal; (P/Q) overlap of Pro- (P) and Gln-rich (Q) domains; (T/S) Thr/Ser-rich domain. Numbers refer to amino acid position in the primary sequence of the Al protein and mark domain boundaries. The arrow indicates the position of the distal breakpoint of *In(2L)al¹³⁰*.

Table 1. The *P*-element insertion *B38* carrying genomic *al* DNA rescues the lethality associated with the inversion *In(2L)al*¹³⁰

Genotypes of progeny ^a	Cross (inter se) [<i>w</i> ¹¹⁸ ; <i>In(2L)al</i> ¹³⁰ / <i>CyO</i> ; <i>B38</i> /+]			
	observed number of progeny	fraction of combinatorial events	calculated number of progeny per single combinatorial event	percentage of viability ^b
<i>In(2L)al</i> ¹³⁰ / <i>In(2L)al</i> ¹³⁰ ; +/+	0	1/16	0	0
<i>In(2L)al</i> ¹³⁰ / <i>In(2L)al</i> ¹³⁰ ; <i>B38</i> /+	39	2/16	19.5	15
<i>In(2L)al</i> ¹³⁰ / <i>In(2L)al</i> ¹³⁰ ; <i>B38</i> / <i>B38</i>	28	1/16	28	40 ^c
<i>In(2L)al</i> ¹³⁰ / <i>CyO</i> ; +/+	238	2/16	119	100
<i>In(2L)al</i> ¹³⁰ / <i>CyO</i> ; <i>B38</i> /+	530	4/16	132.5	100
<i>In(2L)al</i> ¹³⁰ / <i>CyO</i> ; <i>B38</i> / <i>B38</i>	139	2/16	69.5 [*]	55 ^e
Number of progeny scored	974			

^aThe lethal combinations involving the homozygous *CyO* chromosomes are not listed, and X chromosomes (carrying the *w*¹¹⁸ marker) are omitted.

^bNormalized to 126(d) with the exception of 40(c).

^cNormalized to 69.5(*) due to the reduced viability of the homozygous *B38* chromosome.

^dAverage of values from rows four and five.

^eNote the reduced viability of the homozygous *B38* chromosome.

ing to class I cDNAs consist of a leader (interrupted by the first intron) of at least 396 nucleotides but probably closer to ~0.8 kb, an open reading frame of 1152 nucleotides (beginning with an ATG at nucleotide 7138), and a trailer of 243 nucleotides.

The potential translational start site fits the consensus sequence well (Kozak 1986) and predicts a protein of 384 amino acids (Fig. 3B). Its amino-terminal part comprises a stretch of 22 amino acids rich in charged residues (8/22), followed by 45 amino acids, including 22% Ala (10/45), 22% Ser (10/45), and 18% Gly (8/45). A short stretch of acidic residues (5/10) is part of an extended homology of 18 amino acids frequently abutting the amino terminus of *paired* (*prd*)-type homeo domains (Bopp et al. 1986). Residues 82–89 might contain a nuclear localization signal (Silver 1991) overlapping with this extended homology and a *prd*-type homeo domain that begins at residue 85 (see below). Three of the nine residues encoded by amino acids 183–191 are charged and surrounded by long runs of uncharged amino acids. The carboxy-terminal portion of the protein, encoded entirely by the last exon, consists of two Pro-rich regions located at its ends (20/69 and 11/35). Between them, and partially overlapping with the first Pro-rich region, is a stretch of 41 amino acids rich in Gln (11/41), followed by 19 amino acids rich in Ser and Thr (9/19).

Sequence analysis of a genomic fragment containing the proximal breakpoint of the *al* inversion *In(2L)al*¹³⁰ shows that the inversion removes the region downstream of nucleotide 9332, encoding the Pro- and Gln-rich domain, from the *al* transcription unit.

The *al* gene encodes a homeo domain belonging to the *prd* class

A comparison of the *al* homeo domain with homeo domains found in a variety of metazoa reveals a striking

resemblance to the *prd* class of homeo domains (Fig. 4A; Bopp et al. 1986). Within this class, the *al* homeo domain is most closely related (75% identity) to the murine *Pax-7* (Jostes et al. 1991) and *Schistosoma mansoni smox-3* homeo domains (Webster and Mansour 1992), whereas it has diverged slightly more from the *gooseberry neuro* (*gsbn*, formerly designated *gsb-BSH4* or *gsb-p*), *Pax-3* (73%), and *prd* homeo domains (70%) (Bopp et al. 1986; Goulding et al. 1991). The lowest degree of similarity, found with the *Xenopus Mix.1* homeo domain (Rosa 1989), is still 60%.

With the exception of two residues, all *prd*-specific amino acids (Bopp et al. 1986) are present in the *al* homeo domain if conservative changes are neglected (Fig. 4A). One of the two nonconservative changes of *prd*-specific amino acids in the *al* homeo domain deserves special comment, namely the change of a Ser (so far typical for all *prd*-type homeo domains) into a Gln at position 50. This difference may be important because this particular residue was shown to be a major determinant of the homeo domain DNA-binding specificity in the case of *prd* (Treisman et al. 1989), as well as *bicoid* (*bcd*) (Hanes and Brent 1989; 1991). In the other nonconservative change, Ile replaces Thr at position 38. The significance of this alteration is unclear.

A subset of homeo domains belonging to the *prd* class, including that of *al*, *Splotch* (*Sp*), *Pax-7*, and the *Pax-6* homologs *Small eye* (*Sey*), *Aniridia* (*AN*), and *pax[zf-a]*, display an extended homology of 18 amino acids at their amino terminus (Fig. 4B), which was first described for *prd*, *gsb*, and *gsbn* as an additional characteristic trait of *prd*-type homeo domains (Bopp et al. 1986). Its most striking features are its high negative charge, which may amount to up to 7 acidic residues of 9 in the case of *gsb* (6 of 10 residues in *al*), and the basic tripeptide Lys-Arg-Lys at its carboxy-terminal end linked to the homeo domain. These amino-terminal extensions of *prd*-type ho-

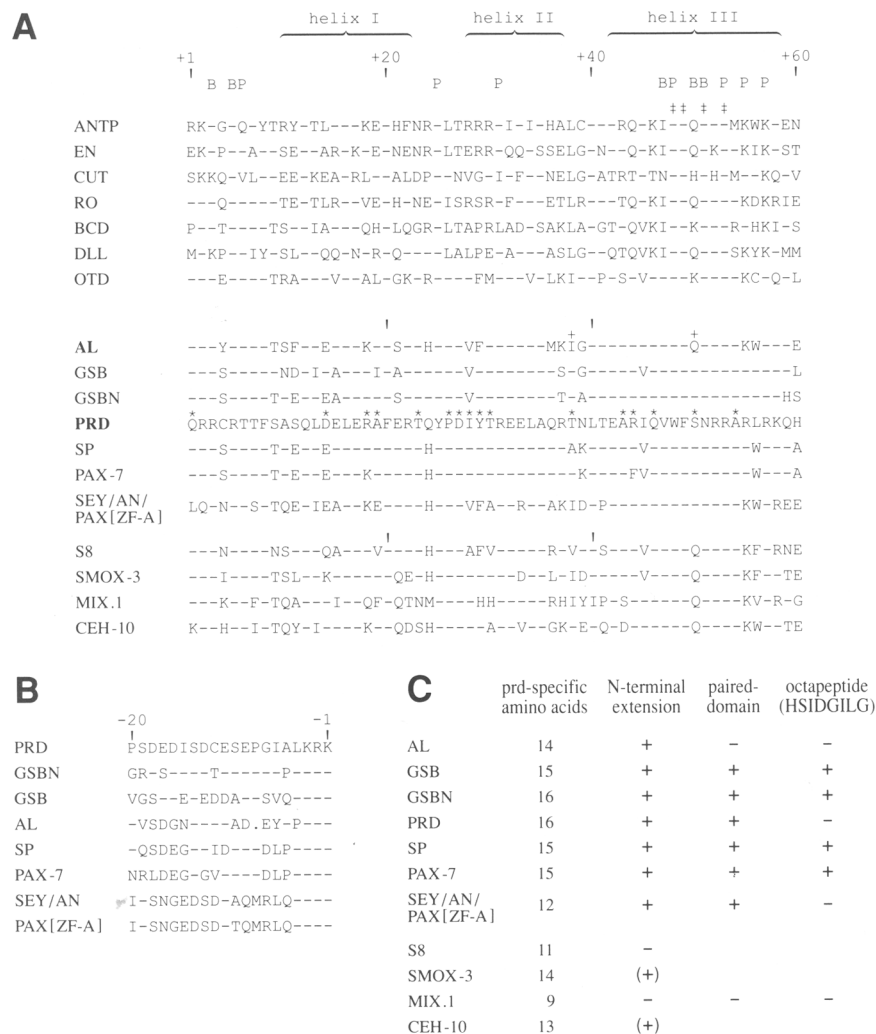


Figure 4. (A) Sequence alignment of various homeo domains. Sequences of *Drosophila* homeo domains are listed above; those of other organisms are listed below that of the Paired homeo domain (PRD). Amino acids identical to those of the Paired homeo domain are shown as dashes. The 16 residues that are conserved in *prd*-type homeo domains, yet differ from those of other homeo domain classes, are defined as *prd* specific (Bopp et al. 1986) and are marked by an asterisk above the Paired sequence; the two residues of the Al homeo domain differing from *prd*-specific amino acids are denoted by a + above its sequence. The residues important for DNA contacts with bases (B) or the phosphate backbone (P) (Kissinger et al. 1990) are marked. (+) The four conserved residues of helix III. (B) Amino-terminal extensions of *prd*-type homeo domains. The 20 residues immediately preceding several *prd*-type homeo domains are aligned. A gap of 1 amino acid (indicated by a dot) has been introduced into the extended homeo domain of *al* (AL). Sequences of homeo domains and their extensions were taken from the following references: *Antennapedia* (*Antp*), McGinnis et al. (1984), Scott and Weiner (1984); *engrailed* (*en*), Fjose et al. (1985), Poole et al. (1985); *cut*, Blochlinger et al. (1988); *rough* (*ro*), Saint et al. (1988), Tomlinson et al. (1988); *bicoid* (*bcd*), Frigerio et al. (1986); *Distal-less* (*Dll*), Cohen et al. (1989); *orthodenticle* (*otd*), Finkelstein et al. (1990); *gsb* (formerly *gsb-BSH9*) and *gsbn* (formerly *gsb-BSH4*; for new names see X. Li, T. Gutjahr, and M. Noll; T. Gutjahr, N.H. Patel, X. Li, C.S. Goodman, and

M. Noll; both in prep.), Bopp et al. 1986; *prd*, Frigerio et al. (1986); *Sp* (*Pax-3*), Goulding et al. (1991); *Pax-7*, Jostes et al. (1991); *Sei* (*Pax-6*), Walther and Gruss (1991); *AN*, Ton et al. (1991); *pax[zf-a]*, Krauss et al. (1991b); *S8*, Opstelten et al. (1991); *smox-3*, Webster and Mansour (1992); *Mix.1*, Rosa (1989); *ceh-10*, Hawkins and McGhee (1990). (C) Characteristic features of *prd*-type and *prd*-like homeo domains (for details, see Discussion).

meo domains exhibit >60% conservation (disregarding conservative changes) in comparison to the *prd* sequence, with the exception of the *Pax-6* extension which is, however, >60% homologous to that of *gsb*. The extended homeo domain homology has so far only been observed in association with *prd*-type homeo domains.

It should be noted that the helix-turn-helix motif of the *al* homeo box is interrupted by two introns, one in each of the two helices, at positions corresponding to amino acid residues 31 and 46 of the homeo domain. Remarkably, the location of the second intron is conserved in several homeo box genes, as, for example, in the mouse and *Xenopus* homologs of the *Drosophila* *even-skipped* gene, *Evx 1*, *Evx 2*, and *Xhox-3* (Ruiz i Altaba and Melton 1989; Bastian and Gruss 1990), the mouse gene *S8* (Kongsuwan et al. 1988; Opstelten et al. 1991), or the four *Caenorhabditis elegans* genes *ceh-4*, *ceh-8*, *ceh-14*, and *ceh-10* (Bürglin et al. 1989; Hawkins

and McGhee 1990). The *S8* homeo domain and *ceh-10* homeo domain belong to the *prd* class (Fig. 4A; see Discussion).

Temporal distribution of *al* transcripts

Several *al* transcripts of different sizes are revealed by Northern blot analysis (Fig. 5). The most prominent band corresponds to a 2.2-kb transcript that first appears in 4- to 8-hr-old embryos, peaks in 8- to 12-hr-old embryos, and remains detectable throughout development up to late larval stages. During the second half of embryogenesis, a much less abundant transcript of ~1.9 kb appears. In addition, two bands, which are barely detectable, appear at 3 and 4.5 kb in parallel to the main 2.2-kb transcript. Whereas the 3-kb band might be a precursor of the main transcript containing the second intron, the 1.9- and 4.5-kb bands cannot be explained in this manner

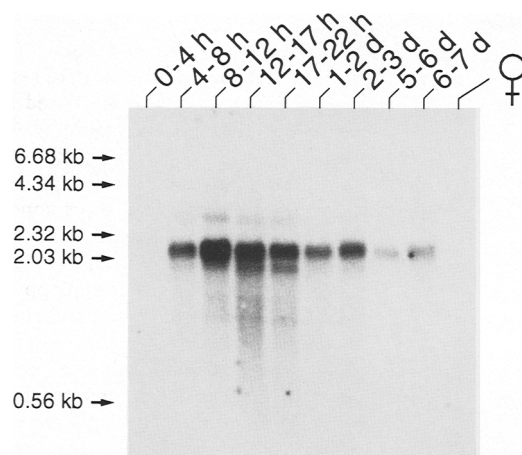


Figure 5. Developmental profile of *al* transcripts. Poly(A)⁺ RNA (5 µg/lane) of 0- to 4-, 4- to 8-, 8- to 12-, 12- to 17-, and 17- to 22-hr-old embryos of first instar (1–2 day), second instar (2–3 day), and late third-instar (5–6 day and 6–7 day) larvae and adult females were analyzed by Northern blot transfer and hybridization with a 1.7-kb genomic *Pst*I fragment (Fig. 3A) labeled with [α -³²P]dATP (3000 Ci/mmol) by nick translation. Exposure occurred for 2.5 days at -80°C with an intensifying screen. For size calibration, the positions of *Hind*III fragments of λ DNA run on the same gel are indicated at left.

and, hence, might represent transcripts processed differently and/or initiated from a different transcription start site. No transcripts were detectable in 0- to 4-hr-old embryos and in adult females.

Transcript pattern of *al* during embryogenesis

The temporal and spatial expression of *al* was examined by *in situ* hybridization to whole-mount embryos, using as probe a 1.63-kb DIG/dUTP-labeled fragment of a class I cDNA. The *al* gene is transcribed successively in the head, the germ band, and the gut in characteristic, bilaterally symmetrical patterns.

As shown in Figure 6A, the first signals of *al* transcripts are detected in the head of stage 10 embryos (late germ-band extension; Campos-Ortega and Hartenstein 1985). Two lateral spots are visible just behind the cephalic furrow in the maxillary and labial segment primordia. Shortly thereafter, a third, more anterior, spot appears (Fig. 6B) in what is probably the prospective antennal or perhaps mandibular segment (Turner and Mahowald 1979; Jürgens et al. 1986; Finkelstein and Perimon 1991). A complex dynamic transformation of the head expression pattern is subsequently observed until early stage 11 (Fig. 6C–E). Several weak spots appear in the antennal segment, and a single patch of cells begins to transcribe *al* in the mandibular segment. In addition, *al* transcripts emerge as three spots (regular three-spot icon) in the maxillary lobe while they are reduced in the labial segment (Fig. 6D,E). At the same time, a weak beads-on-a-string-like pattern of *al* staining is detected in the embryo along its border to the amnioserosa (Fig. 6E),

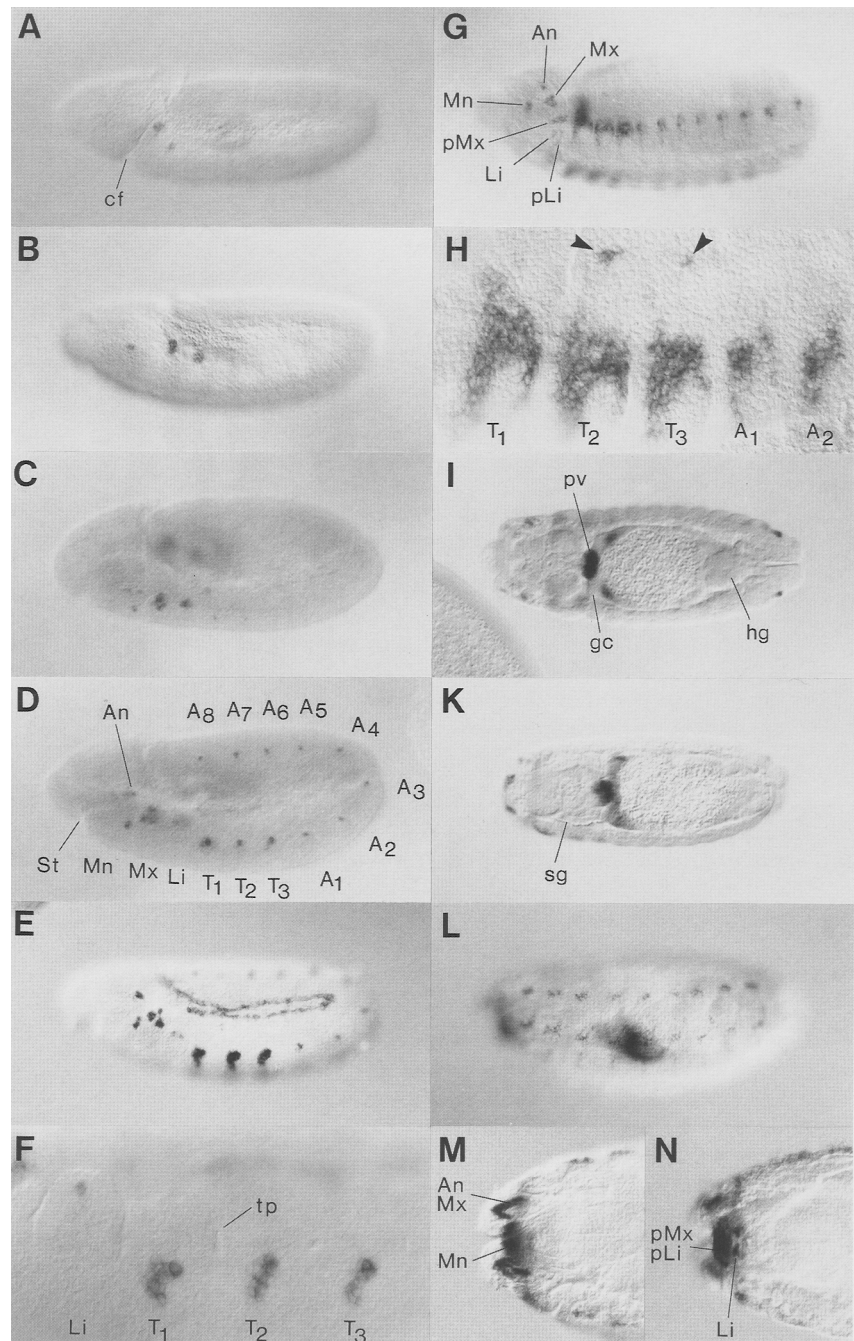
which persists until completion of dorsal closure. During head involution, a stripe of cells in the posterior labial lobe begins to express *al* (Fig. 6G, pLi), whereas *al* transcripts increase again in the more anterior patch of cells in this segment (Fig. 6G, Li). Similarly, an additional patch of cells transcribing *al* appears in the posterior maxillary lobe (Fig. 6G, pMx).

The head staining persists at least until late embryogenesis (Fig. 6M,N). In a more dorsal view, staining is observed in an internal structure belonging to the antenno-maxillary complex (Fig. 6M, An, Mx). This tissue consists of *al*-expressing cells derived from both the antennal and maxillary segment (Fig. 6G, An, Mx). More medially, structures at the very beginning of the atrium express *al* (Fig. 6M, Mn). This pattern, which can also be traced back to earlier stages (Fig. 6G,I,K), appears to be of mandibular origin and might consist of cells involved in mouth hook formation. In a more ventral view, the labium is stained (Fig. 6N, pMx, pLi). These cells are derived from the two patches of cells expressing *al* in the posterior labial (pLi) and posterior maxillary segment (pMx) at an earlier stage (Fig. 6G), which have approached each other during subsequent stages to form the labium. In addition, two small bilateral spots of *al* transcripts are visible (Fig. 6N, Li), which might correspond to cells of the labial sense organ (Jürgens et al. 1986) and apparently originate from those cells expressing *al* at stage 13 in the center of the labial lobe (Fig. 6G, Li).

In the epidermis, a segmentally repeated pattern of *al* transcripts, consisting of three thoracic and eight abdominal lateral patches, appears at the beginning of the extended germ-band stage (Fig. 6C–E). Subsequently, *al* transcripts accumulate to a much greater extent in the thoracic than in the abdominal regions (Fig. 6D,E). These epidermal cells expressing *al* are located immediately anterior to the parasegmental grooves (Martinez-Arias and Lawrence 1985), as suggested by their relative positions to the tracheal pits (Fig. 6F). However, shortly after the appearance of these epidermal clusters consisting solely of a few *al*-expressing cells, the shape of these clusters rapidly becomes more complex (Fig. 6F–H). Particularly in the thoracic segments, it resembles that of a horseshoe, partially surrounding a ventral cluster of cells that do not express *al* (Fig. 6H). Moreover, in the second and third thoracic segment, additional small clusters of *al*-expressing cells appear more dorsally (Fig. 6H, arrowheads). There is no obvious correlation of this pattern with morphological structures. Some of the *al*-expressing cells and the nonexpressing cells that they partially enclose are imaginal disc precursor cells [as judged by double-labeling experiments using *al* and *Dll* probes (K. Schneitz and M. Noll, in prep.)]. This pattern is stable up to the end of stage 15 when it begins to fade slowly. The eighth abdominal spot, conceivably labeling the anterior lateral sense organs (Jürgens 1987), seems to persist slightly longer (Fig. 6I).

At late stage 15 and during stage 16, two rows of segmentally repeated ventro- and dorsolateral spots are observed to express *al* in the epidermis of the thorax and the first seven abdominal segments (Fig. 6L). Expression

Figure 6. Patterns of *al* transcripts during wild-type embryogenesis. (A–G,L) Focus on the lateral surface (with slight ventral or dorsal shifts and an enlarged view in F) of whole-mount embryos oriented with their anterior to the left and dorsal side up. (H) A dissected and flattened lateral epidermis of a stage 13 embryo. (I,K) Slightly dorsally and ventrally located transverse optical sections. (M,N) Detailed ventral views of stained head structures with a more dorsal (M) and a more ventral (N) plane of focus in different embryos. Embryos are at stage 10 (A,B), stage 11 (C–F), early stage 13 (G,H), early stage 15 (I), late stage 15 or early stage 16 (K), stage 16 (L), and late stage 16 or early stage 17 (M,N); all stages are according to Campos-Ortega and Hartenstein (1985). A DIG-labeled 1.65-kb *Pst*I fragment from *alc1* (class I; containing all of the information for the open reading frame as well as for 311-bp leader and 171-bp trailer sequences) was hybridized to whole-mount embryos. Development of the stainings was for 6.5 hr except for the embryos shown in E and L, which were stained for 20 hr. (A1–A8) Abdominal segments; (An) antennal segment; (cf) cephalic furrow; (gc) gastric caecum; (hg) hindgut; (Li) labial segment; (Mn) mandibular segment; (Mx) maxillary segment; (pv) proventriculus; (sg) salivary gland; (st) stomodeum; (T1–T3) thoracic segments; (tp) tracheal pit.



of *al* in the dorsolateral cells appears to be newly induced, whereas the ventrolateral spots of *al* transcripts could be remnants of the dorsal-most portion of the thoracic horseshoe pattern and the ventrolateral abdominal pattern. Hence, this dorsal-most part of the horseshoe pattern might be regulated independently from the ventrally abutting “true” horseshoe pattern (Fig. 6H) and reflect a function of *al* distinct from its potential role in the early formation of imaginal discs. This idea is consistent with the finding that the dorsal-most and the horseshoe portions of the thoracic epidermal pattern are

differentially regulated by some segment polarity genes during stages 13 and 14 (K. Schneitz and M. Noll, in prep.). Similarly, the much smaller abdominal epidermal cell clusters expressing *al* are composed of differentially regulated dorsal and ventral parts. The double-row pattern disappears during late stage 16 and at the beginning of stage 17. We do not know what morphological structures are stained and what the specific function of *al* in these regions may be. By the end of embryogenesis, epidermal expression of *al* has disappeared from the trunk.

Finally, *al* begins to be expressed in the intestinal tract

at stage 13 [Fig. 6G; a detailed description of the gut pattern will be presented elsewhere (K. Schneitz and M. Noll, in prep.)]. Initially, a strongly stained ring is observed, which consists of cells forming the prospective proventriculus (Campos-Ortega and Hartenstein 1985; Reuter and Scott 1990). Its location coincides with the anterior end of the overlaying visceral mesoderm. During closure of the gut tube, a second, less intensely labeled ring is found posterior to the proventricular ring. During the following stages, 14 and 15, this posterior ring is visible in the endoderm of the anterior midgut (Fig. 6I) immediately posterior to the buds of the developing gastric caeca (Campos-Ortega and Hartenstein 1985; Reuter and Scott 1990). As development proceeds, this second ring expands anteriorly until it spans the buds of the gastric caeca, leaving a gap of cells that do not express *al* between the gastric caeca and the proventriculus (Fig. 6K). At stage 16, the first constriction marks the posterior end of *al* transcripts in the midgut (not shown). The proventricular pattern, but not the anterior midgut expression, disappears by the end of embryogenesis (not shown).

Transcript pattern of al in late third-instar imaginal discs

The distribution of *al* transcripts in imaginal discs from late third-instar larvae is shown in Figure 7. In leg discs (Fig. 7A), *al* is expressed in an anterior–lateral region corresponding to the prospective thorax (Schubiger 1968). Although not evident from this view of the disc, the lateral staining actually extends in a narrow stripe of cells much more posteriorly. In addition, twin spots are visible in the center of the disc, where the prospective claw organ has been mapped (Schubiger 1968). In a lateral view, the two spots appear as two short tubes.

In the eye–antennal disc (Fig. 7B), *al* expression is observed only in the antennal part of the disc. A laterally staining patch of cells is found in a region from which part of the head capsule originates (Haynie and Bryant 1986). This patch extends into a region of the folds, which produces lateral parts of the antenna. The center of the antennal disc, generating the arista, is marked by a single cluster of cells expressing *al*.

Finally, the wing disc pattern (Fig. 7C) also consists of two distinct regions. The first region generates at least part of the future scutellum (Bryant 1975), but presumably not the postscutellum, while the second region corresponds to part of the anterior compartment of the wing pouch. Here, staining is found on the dorsal as well as the ventral side of the future wing blade and includes the area of the wing margin that gives rise to the anterior triple row of bristles.

Discussion

The al protein contains a prd-type homeo domain yet no paired domain

The *al* locus codes for a protein containing a homeo domain as a potential DNA-binding domain. Furthermore,

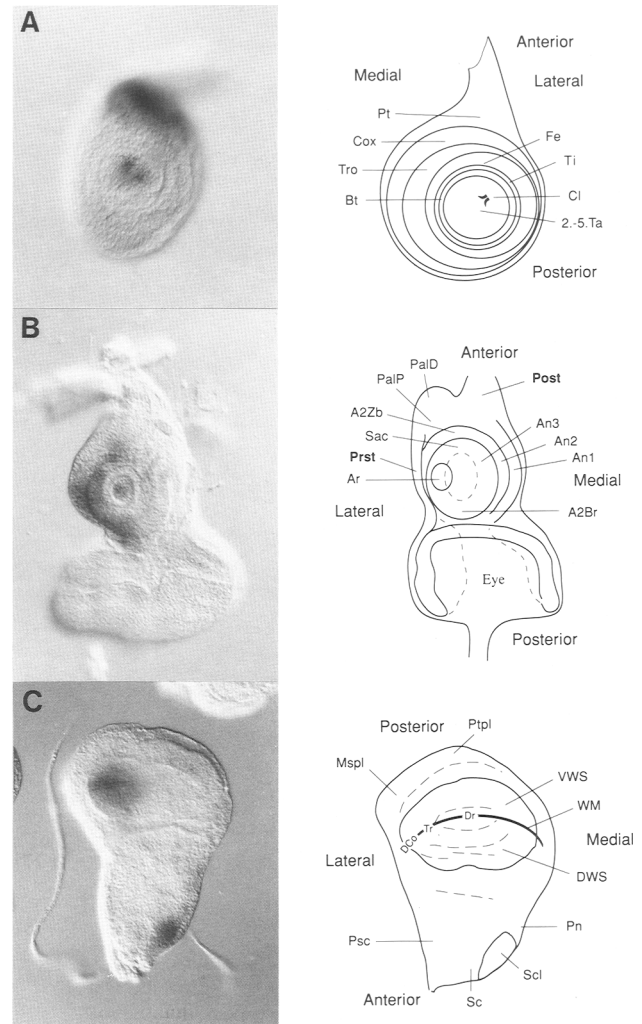


Figure 7. Distribution of *al* transcripts in whole-mount late third-instar imaginal discs. Panels show in situ hybridizations (with the same DIG-labeled cDNA fragment as described in the legend to Fig. 6; staining was developed for 2 hr) to a leg (A), eye–antennal (B), and wing disc (C). Corresponding fate maps of imaginal discs, drawn to the right of each in situ hybridization, were adopted from Schubiger (1968) (A), from Haynie and Bryant (1986) (B), and from Bryant (1975) (C). The fate maps further indicate the orientation of each disc with respect to the antero-posterior and dorsoventral axis of the larva. In the eye–antennal fate map, the areas denoted with bold letters were mapped onto the overlaying peripodial membrane. (Bt) Basitarsus; (Cl) claws; (Cox) coxa; (Fe) femur; (Pt) thorax; (Ti) tibia; (2-5.Ta) tarsal segments; (Tro) trochanter. (A2Br) Large bristles on An2; (A2Zb) Zahnborsten; (An1–An3) antennal segments; (Ar) arista; (PalD) distal maxillary palpus bristles; (PalP) proximal maxillary palpus sensilla trichodea; (Post) postoccipital sensilla trichodea; (Prst) proximal rostrum sensilla trichodea; (Sac) sacculus. (DCo) Distal costa; (Dr) double bristle row (distal wing margin); (DWS) dorsal wing surface; (MspI) mesopleura; (Pn) postnotum; (Psc) prescutum; (PtPl) pteropleura; (Sc) scutum; (Scl) scutellum; (Tr) triple bristle row (anterior wing margin); (VWS) ventral wing surface; (WM) wing margin.

the primary sequence of the *al* protein exhibits a short and two long stretches of charged amino acids in the amino-terminal half, as well as Pro-, Gln-, and Ser/Thr-rich regions in the carboxy-terminal half characteristic of transcriptional activator domains (for review, see Mitchell and Tjian 1989). These features of the putative *al* protein strongly suggest that it is a transcriptional regulator of target genes not yet identified.

The high degree of homology among the homeo domains of *al*, *prd*, and the two *gsb* genes (80% between those of *al* and *prd* if conservative changes are neglected) indicates that the *al* homeo domain belongs to the *prd* class [Bopp et al. 1986]. This class currently consists of the four *Drosophila* genes *prd*, *gsb*, *gsbn*, and *al*, the three mouse genes *Sp* (identical with *Pax-3*; Epstein et al. 1991; Goulding et al. 1991), *Sey* (identical with *Pax-6*; Hill et al. 1991; Walther and Gruss 1991), and *Pax-7* [Jostes et al. 1991], and the human and zebra fish homologs of *Sey*, *AN* (Ton et al. 1991), and *pax[zf-a]* [Krauss et al. 1991b] on the basis of the following criteria. First, their homeo domains are highly conserved in 16 amino acids characteristic of the *prd*-type homeo domains (Fig. 4A); that is, these residues differ from amino acids conserved in other classes of homeo domains [Bopp et al. 1986]. Second, their homeo domains display at their amino-terminal end an extended homology of 18 amino acids (Fig. 4A,B). Finally, their homeo domains are associated with a paired domain of the *prd* type [Bopp et al. 1986, 1989; Burri et al. 1989], with one exception. The *al* gene does not encode a paired domain and, hence, is not a member of the *PHox* gene set [Bopp et al. 1989]. Thus, just as the paired domain is not always associated with a *prd*-type homeo domain [Bopp et al. 1986, 1989], the *prd*-type homeo domain can also occur independently of a paired domain.

The *Sey/AN/pax[zf-a]* genes exhibit the lowest similarity with respect to the first two criteria (12/16 *prd*-specific amino acids, 6/18 amino acid amino-terminal extension conserved). However, the relatively low conservation of the amino-terminal extension is increased significantly if it is compared with that of the *gsb* (56%) rather than the *prd* homeo domain (33%). Because the *al* protein contains such an extended homology but no paired domain, it appears that this extension is not only associated with *prd*-type homeo domains in genes that also contain a paired domain. It is unclear at present whether or not all genes containing a *prd*-type homeo domain—as defined by containing at least 12 of the 16 *prd*-specific amino acids (Fig. 4A)—display this amino-terminal extension. Two additional genes, the *S. mansonii* gene *smox-3* and the *C. elegans* gene *ceh-10*, have *prd*-type homeo domains (Fig. 4A) that might also possess an amino-terminal extension (only part of their extensions are known, which, however, exhibit at least 70% identity), whereas the homeo domains of *S8* and *MIX.1* are clearly not extended (Fig. 4C). Therefore, it might be appropriate to classify *prd*-type homeo domains according to the first two criteria discussed above. On the other hand, homeo domains that exhibit a high degree of homology to that of the *prd* gene but display a

lower number of *prd*-specific amino acids (<12) and no amino-terminal extension might be qualified as *prd*-like rather than *prd*-type homeo domains.

Another domain, associated with a subset of genes containing a paired domain [Burri et al. 1989; Goulding et al. 1991; Jostes et al. 1991] but not necessarily also a *prd*-type homeo domain [Dressler et al. 1990; Plachov et al. 1990; Krauss et al. 1991a; Adams et al. 1992], is the octapeptide HSIDGILG [Burri et al. 1989]. In agreement with its exclusive association with paired domain genes, the *al* gene does not contain this octapeptide. Thus, the octapeptide and the amino-terminal extension have so far always been found associated with a paired domain and a *prd*-type homeo domain, respectively, or with both of these domains (Fig. 4C). An octapeptide with only 50% identity to HSIDGILG has also been found associated with other types of homeo domains [Allen et al. 1991]. The functions of the amino-terminal extension of *prd*-type homeo domains or of the octapeptide are still a matter of speculation. Their relatively small size and association with one or even two DNA binding domains [Bopp et al. 1989; Burri et al. 1989; Chalepakakis et al. 1991; Treisman et al. 1991; Adams et al. 1992], as well as the high negative charge of the homeo domain extension, seem to favor a role in the interaction with specific proteins rather than DNA sequences as, for example, in the formation of homodimers or heterodimers with other DNA-binding or adaptor proteins [Bopp et al. 1986]. It remains to be clarified whether the octapeptide or the amino-terminal extension of *prd*-type homeo domains constitute domains in the previously defined sense; that is, domains exhibiting independent assortment with other network-specific domains during evolution [Bopp et al. 1986; Frigerio et al. 1986].

What are the embryonic functions of *al*?

The embryonic expression of *al*, as revealed by class I transcripts, can be subdivided into distinct epidermal patterns of head, thorax, and abdomen and an endodermal pattern in the anterior intestinal tract. Here, we have studied the wild-type epidermal pattern of embryos and imaginal discs. A detailed analysis of the gut pattern will be published elsewhere (K. Schneitz and M. Noll, in prep.). Only two lethal alleles of *al* are available at present, the inversion *In(2L)al¹³⁰* and a large deficiency *Df(2L)al* [Lewis 1945]. The deficiency removes ~220 kb of the 21C1,2–21C7,8 interval (Fig. 2A) and several other genes, including *u-shaped* [Nüsslein-Volhard et al. 1984], which is responsible for the embryonic lethal phenotype of the deficiency and complements *al* alleles [K. Schneitz, unpubl.]. Essentially all homozygous or hemizygous *In(2L)al¹³⁰* embryos escape embryogenesis without apparent cuticular defects. However, as we have been able to detect *al* transcripts in homozygous *In(2L)al¹³⁰* embryos [K. Schneitz, unpubl.], this mutation is not a true null allele, which might explain our failure to detect an embryonic phenotype. In the absence of an embryonic phenotype, an assessment of possible embryonic functions of *al* relies on an interpretation of

al expression in primordial larval structures identified only on the basis of their position and morphology.

Embryonic head expression emerges first and can be subdivided into an early and a late pattern. The early expression, at late stage 10, consists of the spots in the antennal or mandibular segment, the maxillary, and the labial segment. It is unclear whether the early pattern is a prerequisite for the late pattern or whether it reflects an early and independent function of *al* in the head (see below). The late pattern can be traced from mid-stage 11 through subsequent stages. Several larval head structures express *al* at stage 17, including the vestigial appendages of the antennomaxillary complex and the labial sense organ (Cohen and Jürgens 1989), as well as other tissues like the labium and cells perhaps involved in mouth hook formation (Jürgens et al. 1986). Expression of *al* in these tissues suggests that it plays a role in the development of this subset of larval head structures and that they might be determined before or at mid-stage 11 of embryogenesis.

It is interesting to note the discrepancy of *al* expression between larval head primordia and their adult homologs. Whereas *al* expression in the antennomaxillary complex might label larval structures homologous to the arista stained later in the eye–antennal disc, no *al* transcripts were detected in adult primordia corresponding to the remaining larval head structures expressing *al*, that is, in the labial and clypeolabral discs. Hence, *al* is expressed in a more complex pattern in the late embryonic head than in its imaginal counterparts, a finding that might reflect a situation explained by the idea that during evolution of *Cyclorrhaphora* evolutionary pressure has acted mainly on the head part of the larva (Snodgrass 1953; Diederich et al. 1991). According to this speculative view, *al* thus might have acquired a secondary function during evolution of the larval head in *Cyclorrhaphora*.

Early role of *al* in imaginal disc development

In the thorax and abdomen, *al* expression appears in segmentally repeated small clusters of cells. Their position relative to the tracheal pits indicates a location at or close to the parasegmental boundaries (Martinez-Arias and Lawrence 1985). This conclusion is corroborated by the observation that *al* expression in the trunk depends on *wingless* (*wg*) and *engrailed* (*en*) (K. Schneitz and M. Noll, in prep.), which specify the parasegmental boundary (Ingham et al. 1991). The final thoracic pattern of *al* displays in each segment a characteristic horseshoe shape and partially surrounds a ventral cluster of unstained cells. On the basis of their location, these thoracic ventral cell clusters are probably the ventral disc cells expressing *Dll*. The homeo box gene *Dll* is crucial for the proximodistal axis formation during limb development of the larva as well as the adult, and its thoracic expression can be taken as marker for the embryonic origin of the limb portion of leg discs (Cohen 1990; Cohen et al. 1991). In this case, *al* expression, which abuts the *Dll*-expressing cells of the limb anlage dorsolaterally,

might mark the proximal or body wall anlage of the leg discs.

A similar pattern is observed in third-instar imaginal leg discs, where *al* is also expressed in the anlage of the thoracic body wall and thus flanks the anlage for the limb proper. Therefore, it is possible that *al* also plays a role in the early development of imaginal discs. If true, the more dorsal *al* spots in the second and third thoracic segments might hint at a similar function of *al* in wing and haltere disc development. Furthermore, the early head pattern could be interpreted in light of such a hypothesis as well. The early expression of *al* in head and thorax is also very similar to the *Dll* expression pattern (cf. Fig. 6E with Fig. 1f of Cohen 1990). Hence, *Dll* might activate the initial transcription of *al* and/or repress *al* in the thoracic unlabeled cluster of cells of the horseshoe pattern. The significance of the *al* pattern in abdominal segments A1–A7 is unclear, even though the localization of the staining might correlate with the ventral histoblast regions (Frayne and Sato 1991), indicating a possible role of *al* in histoblast development.

Bipartite pattern of *al* in antennal, wing, and leg discs

The primordia of all adult tissues that are phenotypically affected in *al* flies express *al*. These include the future arista in the eye–antennal disc, the claws and sternopleurum in the leg discs, as well as part of the wing blade and the scutellum in the wing disc. In addition, *al* expression has been observed in the prospective posterior head. However, the back of the head is not changed in mutant flies in any obvious manner (K. Schneitz, unpubl.), although effects in this part of the fly are rather difficult to discern so that minor alterations might have escaped our detection.

Formally, a common feature of *al* expression in leg, antennal, and wing discs is a bipartite pattern in which a proximal portion, belonging to the body proper (future head, ventral thorax, notum), and a very distal part of the appendage (future arista, claw, triple-row subportion of the wing blade) express *al*. However, it should be noted that in the case of the wing disc, the inner distal spot does not correspond to the geometrical center of the disc. The similar organization of the *al* pattern in these different discs might suggest its activation by a common system of positional information (Wolpert 1971), which is shared at least by leg, antennal, and wing discs. Such an idea is consistent with recent findings demonstrating that in the ontogeny of *Drosophila*, genes of the embryonic segmentation gene hierarchy or those belonging to the gene networks specifying the dorsoventral and proximodistal axes are used again in the development of imaginal discs for related tasks of positional specification (for review, see Wilkins and Gubb 1991).

In evolutionary terms, this result is not surprising with respect to leg and antennal discs. Morphology, classical embryology (e.g., Vogt 1947), and genetics (Roberts 1964; Postlethwait and Schneiderman 1971; Postlethwait and Girton 1974) point to a rather close, serial homology between these two appendages. More astonish-

ing is the similarity of these patterns to the *al* expression in the wing disc. The wing is not serially homologous to the leg, and its phylogenetic origin and primitive function are still obscure (see Brusca and Brusca 1990 and references therein). It may either represent a secondary outgrowth of the notum (paranotum) or, more likely, have originated from an epicoxal segment of the euarthropodan leg (Kukalová-Peck 1983). In either case, in view of the arguments mentioned above and taking *al* expression as marker, we consider it probable that a secondary adaption of the positional information system of leg and antenna has occurred in the wing during evolution of pterygote insects.

Morphogenetic role of *al*

We can only speculate about the morphogenetic role of *al*. Among several possibilities, we favor the view that *al* is a region-specific growth control gene involved in the regulation of localized growth patterns at various sites in the animal. Such speculation is supported by several independent lines of evidence. First, Tokunaga and Stern (1969) showed that the scutellar phenotype of *al* is based on a distorted growth pattern in the scutellum, an effect that could possibly also explain the reduction of the arista and claws. In addition, *al* interacts, as *ex* and *ds*, with the *R* locus (Golubovsky and Kulakov 1978; Korochkina and Golubovsky 1978). Interestingly, the effects of the *ex* and *ds* genes have also been explained by an altered growth pattern (Waddington 1941, 1942, 1943, 1953); and in the case of *ds*, an epistatic interaction with the *Gull* mutation has been found (Lindsley and Grell 1968). *Gull* is a dominant allele of the tumor suppressor gene *fat*, and recessive mutations in this locus cause massive overgrowth of the imaginal discs (Bryant et al. 1988; Mahoney et al. 1991).

The postulated morphogenetic function of *al* is less obvious with respect to the embryonic expression pattern. However, it might be interesting to note that its expression in the three thoracic segments corresponds roughly to the regions of enhanced BrdU incorporation that were taken as a sign for compensatory epidermal cell proliferation before the invagination of imaginal discs (Bate and Martinez-Arias 1991).

Materials and methods

General procedures

Standard procedures such as the construction and screening of a genomic library, chromosomal walking, in situ hybridization to salivary gland chromosomes, DNA sequencing by the dideoxynucleotide method, isolation of poly(A)⁺ RNA, or Northern blot analysis were carried out essentially as described (Maniatis et al. 1982; Frei et al. 1985; Kilchherr et al. 1986). A clone containing the proximal breakpoint of *In(2L)al¹³⁰* was isolated from a genomic DNA library prepared from *In(2L)al¹³⁰/Df(2L)S³* flies in EMBL4. Fine mapping of the distal inversion breakpoint located in the *al* transcript was achieved by DNA sequencing and comparison of the mutant DNA sequence with the corresponding wild-type sequence (Fig. 3A). DNA sequence

analysis and EMBL data bank searches were performed by the use of version 7.0 of the GCG sequence analysis software package (Devereux et al. 1984) on a VAX 9000-420/VMS 5.4-2.

Genomic rescue of *al*

The proximal 13.3-kb *EcoRI* fragment of the genomic clone Y3-4 (Fig. 2B) was cloned into the plasmid vector pW6 containing the mini-white gene as marker (Klemenz et al. 1987) and injected into *w¹¹¹⁸* embryos according to standard procedures (Rubin and Spradling 1982). Four independent transgenic lines were established, of which only one, B38, carried the rescue construct on the third rather than the second chromosome and was homozygous viable. Transgenic flies heterozygous and homozygous for the genomic *al* rescue construct were distinguishable by their different eye colors owing to the presence of one and two copies of the mini-white gene. To obtain the results shown in Table 1, it was important to perform crosses under uncrowded conditions.

Isolation of cDNA clones

A cDNA library of poly(A)⁺ RNA from 8- to 12-hr-old embryos was constructed using the Stratagene ZAP-cDNA synthesis kit (Stratagene, no. 200400), generating cDNAs subcloned between the *EcoRI* (5' end) and *XhoI* (3' end) site of the plasmid vector pBluescript SK(-). From a total of 1.3×10^6 phage clones, initially screened with a genomic 1.7-kb *PstI* fragment and subsequently with a 1.63-kb *PstI* fragment derived from *alcl1*, 18 independent *al*-cDNA clones were isolated.

In situ hybridization to whole-mount embryos and discs

In situ hybridization to embryos with digoxigenin-labeled probes was carried out essentially as described (Tautz and Pfeifle 1989). Modifications include the use of a high random primer concentration (2 mg/ml) for the preparation of the probe and the use of embryos for antibody preabsorption that were treated in the same way as those used for hybridization. Fixation steps were performed with formaldehyde (4003, Merck). [In case of a long incubation of embryos in the staining solution (>2 hr), it is recommended to exchange the staining solution once after 2 hr.] Late third-instar imaginal discs were treated essentially according to Phillips et al. (1990), except that an 8-min digestion with proteinase K at room temperature was included. Embryos and discs were mounted in either glycerol or D.P.X. (BDH Chemicals, England). For microscopy and photography, a Zeiss Axiophot with DIC optics and Kodak Tmax 100 films were used.

Drosophila strains and cuticle preparation

The following stocks were kindly supplied by the Bloomington Stock Center: *al b pr c sp*, *Df(2L)al cn/CyO*, and *In(2L)al¹³⁰*; *al ds/Df(2L)S³* [E. Lewis, unpubl. the inversion 21C1,2/21E1,2 was induced in a *ds* mutant background]. The *In(2L)al¹³⁰* chromosome was cleared of other lethals by two outcrosses with a well growing *Oregon-R* (Munich) strain. The *w¹¹¹⁸* stock was obtained from E. Hafen. Adult cuticle preparations were obtained by boiling the flies in 1 N KOH until the eye color faded and by subsequent dissection and mounting of the appropriate structures in D.P.X. (BDH Chemicals, England).

Acknowledgments

We are grateful to Rick Jones and Corrado Caggese for providing genomic clones, to Ed Lewis for comments on his unpublished *In(2L)al¹³⁰* stock, and to the Bloomington Stock Center and Ernst Hafen for fly strains. We thank Barbara and Steve Cohen, as well as Urs Kloter, for advice concerning the DIG in situ hybridization method, Maya Burri and Lorenz Borer for initial DNA sequencing and cDNA screenings, and Anna De Busturia and Xuelin Li for help with the adult cuticle preparations and the injection procedure. Our special thanks go to Fritz Ochsenbein for the art work. We are most grateful to our colleagues in the laboratory for stimulating discussions and to Pam Mitchell and David Arnosti for critical comments on the manuscript. K.S. is especially indebted to Ulrike Utans for continuous support. This work has been supported by Swiss National Science Foundation grants 3.348-0.86 and 31-26652.89 (to M.N.) and by the Kantons Zürich and Basel.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

Note added in proof

Sequence data described in this paper have been submitted to the EMBL/GenBank data libraries.

References

- Adams, B., P. Dörfler, A. Aguzzi, Z. Kozmik, P. Urbánek, I. Maurer-Fogy, and M. Busslinger. 1992. *Pax-5* encodes the transcription factor BSAP and is expressed in B lymphocytes, the developing CNS, and adult testis. *Genes & Dev.* **6**: 1589–1607.
- Agnel, M., S. Kerridge, C. Vola, and R. Griffin-Shea. 1989. Two transcripts from the *rotund* region of *Drosophila* show similar positional specificities in imaginal disc tissues. *Genes & Dev.* **3**: 85–95.
- Akam, M. 1987. The molecular basis for metameric pattern in the *Drosophila* embryo. *Development* **101**: 1–22.
- Allen, J.D., T. Lints, N.A. Jenkins, N.G. Copeland, A. Strasser, R.P. Harvey, and J.M. Adams. 1991. Novel murine homeo box gene on chromosome 1 expressed in specific hematopoietic lineages and during embryogenesis. *Genes & Dev.* **5**: 509–520.
- Bastian, H. and P. Gruss. 1990. A murine *even-skipped* homologue, *Evx 1*, is expressed during early embryogenesis and neurogenesis in a biphasic manner. *EMBO J.* **9**: 1839–1852.
- Bate, M. and A. Martinez Arias. 1991. The embryonic origin of imaginal discs in *Drosophila*. *Development* **112**: 755–761.
- Blochlinger, K., R. Bodmer, J. Jack, L.Y. Jan, and Y.N. Jan. 1988. Primary structure and expression of a product from *cut*, a locus involved in specifying sensory organ identity in *Drosophila*. *Nature* **333**: 629–635.
- Bopp, D., M. Burri, S. Baumgartner, G. Frigerio, and M. Noll. 1986. Conservation of a large protein domain in the segmentation gene *paired* and in functionally related genes of *Drosophila*. *Cell* **47**: 1033–1040.
- Bopp, D., E. Jamet, S. Baumgartner, M. Burri, and M. Noll. 1989. Isolation of two tissue-specific *Drosophila* paired box genes, *pox meso* and *pox neuro*. *EMBO J.* **8**: 3447–3457.
- Brusca, R.C. and G.J. Brusca. 1990. *Invertebrates*. Sinauer Associates, Inc., Sunderland, MA.
- Bryant, P.J. 1975. Pattern formation in the imaginal wing disc of *Drosophila melanogaster*: Fate map, regeneration and duplication. *J. Exp. Zool.* **193**: 49–78.
- Bryant, P.J., B. Huettner, L.I. Held Jr., J. Ryerse, and J. Szidonya. 1988. Mutations at the *fat* locus interfere with cell proliferation control and epithelial morphogenesis in *Drosophila*. *Dev. Biol.* **129**: 541–554.
- Bürglin, T.R., M. Finney, A. Coulson, and G. Ruvkun. 1989. *Caenorhabditis elegans* has scores of homoeobox-containing genes. *Nature* **341**: 239–243.
- Burri, M., Y. Tromvoukis, D. Bopp, G. Frigerio, and M. Noll. 1989. Conservation of the paired domain in metazoans and its structure in three isolated human genes. *EMBO J.* **8**: 1183–1190.
- Campos-Ortega, J.A. and V. Hartenstein. 1985. *The embryonic development of Drosophila melanogaster*. Springer-Verlag, New York.
- Chalepakakis, G., R. Fritsch, H. Fickenscher, U. Deutsch, M. Goulding, and P. Gruss. 1991. The molecular basis of the *undulated/Pax-1* mutation. *Cell* **66**: 873–884.
- Cohen, B., E.A. Wimmer, and S.M. Cohen. 1991. Early development of leg and wing primordia in the *Drosophila* embryo. *Mech. Dev.* **33**: 229–240.
- Cohen, S.M. 1990. Specification of limb development in the *Drosophila* embryo by positional cues from segmentation genes. *Nature* **343**: 173–177.
- Cohen, S.M. and G. Jürgens. 1989. Proximal-distal pattern formation in *Drosophila*: Cell autonomous requirement for *Distal-less* gene activity in limb development. *EMBO J.* **8**: 2045–2055.
- Cohen, S.M., G. Brönner, F. Küttner, G. Jürgens, and H. Jäckle. 1989. *Distal-less* encodes a homeodomain protein required for limb development in *Drosophila*. *Nature* **338**: 432–434.
- Devereux, J., P. Haeberli, and O. Smithies. 1984. A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* **12**: 387–395.
- Diederich, R.J., A.M. Pattatucci, and T.C. Kaufman. 1991. Developmental and evolutionary implications of *labial*, *Deformed* and *engrailed* expression in the *Drosophila* head. *Development* **113**: 273–281.
- Dressler, G.R., U. Deutsch, K. Chowdhury, H.O. Nornes, and P. Gruss. 1990. *Pax2*, a new murine paired-box-containing gene and its expression in the developing excretory system. *Development* **109**: 787–795.
- Epstein, D.J., M. Vekemans, and P. Gros. 1991. *splotch*, a mutation affecting development of the mouse neural tube, shows a deletion within the paired homeodomain of *Pax-3*. *Cell* **67**: 767–774.
- Finkelstein, R. and N. Perrimon. 1991. The molecular genetics of head development in *Drosophila melanogaster*. *Development* **112**: 899–912.
- Finkelstein, R., D. Smouse, T.M. Capaci, A.C. Spradling, and N. Perrimon. 1990. The *orthodenticle* gene encodes a novel homeo domain protein involved in the development of the *Drosophila* nervous system and ocellar visual structures. *Genes & Dev.* **4**: 1516–1527.
- Fjose, A., W. McGinnis, and W.J. Gehring. 1985. Isolation of a homeo box-containing gene from the *engrailed* region of *Drosophila* and the spatial distribution of its transcripts. *Nature* **313**: 284–289.
- Frayne, E.G. and T. Sato. 1991. The *Ultrabithorax* gene of *Drosophila* and the specification of abdominal histoblasts. *Dev. Biol.* **146**: 265–277.
- Frei, E., S. Baumgartner, J.-E. Edström, and M. Noll. 1985. Cloning of the *extra sex combs* gene of *Drosophila* and its identification by P-element-mediated gene transfer. *EMBO J.* **4**: 979–987.

- Frigerio, G., M. Burri, D. Bopp, S. Baumgartner, and M. Noll. 1986. Structure of the segmentation gene *paired* and the *Drosophila* PRD gene set as part of a gene network. *Cell* **47**: 735–746.
- Frischauf, A.-M., H. Lehrach, A. Poustka, and N. Murray. 1983. Lambda replacement vectors carrying polylinker sequences. *J. Mol. Biol.* **170**: 827–842.
- García-Bellido, A. 1975. Genetic control of wing disc development in *Drosophila*. *CIBA Symp.* **29**: 161–182.
- . 1977. Homoeotic and atavic mutations in insects. *Am. Zool.* **17**: 613–629.
- Golubovsky, M.D. and L.A. Kulakov. 1978. Regulator gene in *Drosophila*? *Dros. Inf. Serv.* **53**: 133–134.
- Goulding, M.D., G. Chalepakis, U. Deutsch, J.R. Erselius, and P. Gruss. 1991. Pax-3, a novel murine DNA binding protein expressed during early neurogenesis. *EMBO J.* **10**: 1135–1147.
- Hanes, S.D. and R. Brent. 1989. DNA specificity of the bicoid activator protein is determined by homeodomain recognition helix residue 9. *Cell* **57**: 1275–1283.
- . 1991. A genetic model for interaction of the homeodomain recognition helix with DNA. *Science* **251**: 426–430.
- Hawkins, N.C. and J.D. McGhee. 1990. Homeobox containing genes in the nematode *Caenorhabditis elegans*. *Nucleic Acids Res.* **18**: 6101–6106.
- Haynie, J.L. and P.J. Bryant. 1986. Development of the eye-antenna imaginal disc and morphogenesis of the adult head in *Drosophila melanogaster*. *J. Exp. Zool.* **237**: 293–308.
- Held, Jr., L.I., C.M. Duarte, and K. Derakhshanian. 1986. Extra tarsal joints and abnormal cuticular polarities in various mutants of *Drosophila melanogaster*. *Wilhelm Roux's Arch. Dev. Biol.* **195**: 145–157.
- Hill, R.E., J. Favor, B.L.M. Hogan, C.C.T. Ton, G.F. Saunders, I.M. Hanson, J. Prosser, T. Jordan, N.D. Hastie, and V. van Heyningen. 1991. Mouse *Small eye* results from mutations in a paired-like homeobox-containing gene. *Nature* **354**: 522–525.
- Hodgkin, N.M. and P.J. Bryant. 1978. Scanning electron microscopy of the adult of *Drosophila melanogaster*. In *The genetics and biology of Drosophila* (ed. M. Ashburner and T.R.F. Wright), Vol. 2c, pp. 337–358. Academic Press, London, England.
- Ingham, P.W. 1988. The molecular genetics of embryonic pattern formation in *Drosophila*. *Nature* **335**: 25–34.
- Ingham, P.W., A.M. Taylor, and Y. Nakano. 1991. Role of the *Drosophila* *patched* gene in positional signalling. *Nature* **353**: 184–187.
- Jostes, B., C. Walther, and P. Gruss. 1991. The murine paired box gene, *Pax7*, is expressed specifically during the development of the nervous and muscular system. *Mech. Dev.* **33**: 27–37.
- Jürgens, G. 1987. Segmental organization of the tail region in the embryo of *Drosophila melanogaster*. *Wilhelm Roux's Arch. Dev. Biol.* **196**: 141–157.
- Jürgens, G., R. Lehmann, M. Schardin, and C. Nüsslein-Volhard. 1986. Segmental organization of the head in the embryo of *Drosophila melanogaster*. *Wilhelm Roux's Arch. Dev. Biol.* **195**: 359–377.
- Kaufman, T.C., R. Lewis, and B. Wakimoto. 1980. Cytogenetic analysis of chromosome 3 in *Drosophila melanogaster*: The homoeotic gene complex in polytene chromosome interval 84A-B. *Genetics* **94**: 115–133.
- Kerridge, S. and M. Thomas-Cavallin. 1988. Appendage morphogenesis in *Drosophila*: A developmental study of the *rotund* (*rn*) gene. *Wilhelm Roux's Arch. Dev. Biol.* **197**: 19–26.
- Kilchherr, F., S. Baumgartner, D. Bopp, E. Frei, and M. Noll. 1986. Isolation of the *paired* gene of *Drosophila* and its spatial expression during early embryogenesis. *Nature* **321**: 493–499.
- Kissinger, C.R., B. Liu, E. Martin-Blanco, T.B. Kornberg, and C.O. Pabo. 1990. Crystal structure of an engrailed homeodomain-DNA complex at 2.8 Å resolution: A framework for understanding homeodomain-DNA interactions. *Cell* **63**: 579–590.
- Klemen, R., U. Weber, and W.J. Gehring. 1987. The *white* gene as a marker in a new P-element vector for gene transfer in *Drosophila*. *Nucleic Acids Res.* **15**: 3947–3959.
- Kongsuwan, K., E. Webb, P. Housiaux, and J.M. Adams. 1988. Expression of multiple homeobox genes within diverse mammalian haemopoietic lineages. *EMBO J.* **7**: 2131–2138.
- Korochkina, L.S. and M.D. Golubovsky. 1978. Cytogenetic analysis of induced mutations on the left end of the second chromosome of *D. melanogaster*. *Dros. Inf. Serv.* **53**: 197–200.
- Kozak, M. 1986. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell* **44**: 283–292.
- Krauss, S., T. Johansen, V. Korzh, and A. Fjose. 1991a. Expression of the zebrafish paired box gene *pax[zf-b]* during early neurogenesis. *Development* **113**: 1193–1206.
- Krauss, S., T. Johansen, V. Korzh, U. Moens, J.U. Ericson, and A. Fjose. 1991b. Zebrafish *pax[zf-a]*: A paired box-containing gene expressed in the neural tube. *EMBO J.* **10**: 3609–3619.
- Kukalová-Peck, J. 1983. Origin of the insect wing and wing articulation from the arthropodan leg. *Can. J. Zool.* **61**: 1618–1669.
- Lewis, E.B. 1945. The relation of repeats to position effect in *Drosophila melanogaster*. *Genetics* **30**: 137–166.
- . 1963. Genes and developmental pathways. *Am. Zool.* **3**: 33–56.
- . 1978. A gene complex controlling segmentation in *Drosophila*. *Nature* **276**: 565–570.
- Lindsley, D.L. and E.H. Grell. 1968. *Genetic variations of Drosophila melanogaster*. Carnegie Inst. Washington Publ. **627**.
- Lindsley, D.L. and G. Zimm. 1992. *The genome of Drosophila melanogaster*. Academic Press, New York.
- Mahoney, P.A., U. Weber, P. Onofrechuk, H. Biessmann, P.J. Bryant, and C.S. Goodman. 1991. The *fat* tumor suppressor gene in *Drosophila* encodes a novel member of the cadherin gene superfamily. *Cell* **67**: 853–868.
- Maniatis, T., E.F. Fritsch, and J. Sambrook. 1982. *Molecular cloning: A laboratory manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
- Martinez-Arias, A. and P.A. Lawrence. 1985. Parasegments and compartments in the *Drosophila* embryo. *Nature* **313**: 639–642.
- McGinnis, W., M.S. Levine, E. Hafen, A. Kuroiwa, and W.J. Gehring. 1984. A conserved DNA sequence in homoeotic genes of the *Drosophila* Antennapedia and bithorax complexes. *Nature* **308**: 428–433.
- Mglinets, V.A. and V.I. Ivanov. 1975. Investigation of the effects of homoeotic mutations. V. Interaction of the homoeotic *ss^{ad0a}* and the antennal *th* and *al* genes at different temperatures. *Genetika* **11**: 88–96.
- Mikuta, G.I. and V.A. Mglinets. 1978. Investigation of the interactions between the homeotic *Nasobemia* mutation and the *thread* and *aristaleless* mutations in *Drosophila melanogaster* at 17 and 29°C. *Genetika* **14**: 1578–1585.
- Mitchell, P.J. and R. Tjian. 1989. Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. *Science* **245**: 371–378.

- Nüsslein-Volhard, C. and E. Wieschaus. 1980. Mutations affecting segment number and polarity in *Drosophila*. *Nature* **287**: 795–801.
- Nüsslein-Volhard, C., E. Wieschaus, and H. Kluding. 1984. Mutations affecting the pattern of the larval cuticle in *Drosophila melanogaster*. I. Zygotic loci on the second chromosome. *Wilhelm Roux's Arch. Dev. Biol.* **193**: 267–282.
- Opstelten, D.-J.E., R. Vogels, B. Robert, E. Kalkhoven, F. Zwartkruis, L. de Laaf, O.H. Destrée, J. Deschamps, K.A. Lawson, and F. Meijlink. 1991. The mouse homeobox gene, *S8*, is expressed during embryogenesis predominantly in mesenchyme. *Mech. Dev.* **34**: 29–42.
- Ouweneel, W.J. 1976. Developmental genetics of homoeosis. *Adv. Genet.* **18**: 179–248.
- Phillips, R.G., I.J.H. Roberts, P.W. Ingham, and J.R.S. Whittle. 1990. The *Drosophila* segment polarity gene *patched* is involved in a position-signalling mechanism in imaginal discs. *Development* **110**: 105–114.
- Plachov, D., K. Chowdhury, C. Walther, D. Simon, J.-L. Guenet, and P. Gruss. 1990. *Pax8*, a murine paired box gene expressed in the developing excretory system and thyroid gland. *Development* **110**: 643–651.
- Poole, S.J., L.M. Kauvar, B. Drees, and T. Kornberg. 1985. The *engrailed* locus of *Drosophila*: Structural analysis of an embryonic transcript. *Cell* **40**: 37–43.
- Postlethwait, J.H. and J.R. Girton. 1974. Development in genetic mosaics of *aristopedia*, a homeotic mutant of *Drosophila melanogaster*. *Genetics* **76**: 767–774.
- Postlethwait, J.H. and H.A. Schneiderman. 1971. Pattern formation and determination in the antenna of the homeotic mutant *Antennapedia* of *Drosophila melanogaster*. *Dev. Biol.* **25**: 606–640.
- Reuter, R. and M.P. Scott. 1990. Expression and function of the homeotic genes *Antennapedia* and *Sex combs reduced* in the embryonic midgut of *Drosophila*. *Development* **110**: 289–303.
- Roberts, P. 1964. Mosaics involving *aristopedia*, a homeotic mutant of *Drosophila melanogaster*. *Genetics* **49**: 593–598.
- Rosa, M.F. 1989. *Mix.1*, a homeobox mRNA inducible by mesoderm inducers, is expressed mostly in the presumptive endodermal cells of *Xenopus* embryos. *Cell* **57**: 965–974.
- Rubin, G.M. and A.C. Spradling. 1982. Genetic transformation of *Drosophila* with transposable element vectors. *Science* **218**: 348–353.
- Ruiz i Altaba, A. and D.A. Melton. 1989. Bimodal and graded expression of the *Xenopus* homeobox gene *Xhox3* during embryonic development. *Development* **106**: 173–183.
- Saint, R., B. Kalionis, T.J. Lockett, and A. Elizur. 1988. Pattern formation in the developing eye of *Drosophila melanogaster* is regulated by the homeo-box gene, *rough*. *Nature* **334**: 151–154.
- Schubiger, G. 1968. Anlageplan, Determinationszustand und Transdeterminationsleistungen der männlichen Vorderbeinscheibe von *Drosophila melanogaster*. *Wilhelm Roux's Arch. Entwicklungsmech. Org.* **160**: 9–40.
- Scott, M.P. and A.J. Weiner. 1984. Structural relationships among genes that control development: Sequence homology between the *Antennapedia*, *Ultrabithorax*, and *fushi tarazu* loci of *Drosophila*. *Proc. Natl. Acad. Sci.* **81**: 4115–4119.
- Silver, P.A. 1991. How proteins enter the nucleus. *Cell* **64**: 489–497.
- Snodgrass, R.E. 1953. The metamorphosis of a fly's head. *Smithson. Misc. Collect.* **122**(3): 1–25.
- Stern, C. and C.B. Bridges. 1926. The mutants of the extreme left end of the second chromosome of *Drosophila melanogaster*. *Genetics* **11**: 503–530.
- Sunkel, C.E. and J.R.S. Whittle. 1987. *Brista*: A gene involved in the specification and differentiation of distal cephalic and thoracic structures in *Drosophila melanogaster*. *Wilhelm Roux's Arch. Dev. Biol.* **196**: 124–132.
- Tautz, D. and C. Pfeifle. 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.
- Tokunaga, C. and J.C. Gerhart. 1976. The effect of growth and joint formation on bristle pattern in *D. melanogaster*. *J. Exp. Zool.* **198**: 79–95.
- Tokunaga, C. and C. Stern. 1969. Determination of bristle direction in *Drosophila*. *Dev. Biol.* **20**: 411–425.
- Tomlinson, A., B.E. Kimmel, and G.M. Rubin. 1988. *rough*, a *Drosophila* homeobox gene required in photoreceptors R2 and R5 for inductive interactions in the developing eye. *Cell* **55**: 771–784.
- Ton, C.C.T., H. Hirvonen, H. Miwa, M.M. Weil, P. Monaghan, T. Jordan, V. van Heyningen, N.D. Hastie, H. Meijers-Heijboer, M. Drechsler, B. Royer-Pokora, F. Collins, A. Swaroop, L.C. Strong, and G.F. Saunders. 1991. Positional cloning and characterization of a paired box- and homeobox-containing gene from the *Aniridia* region. *Cell* **67**: 1059–1074.
- Treisman, J., P. Gönczy, M. Vashishtha, E. Harris, and C. Desplan. 1989. A single amino acid can determine the DNA binding specificity of homeodomain proteins. *Cell* **59**: 553–562.
- Treisman, J., E. Harris, and C. Desplan. 1991. The paired box encodes a second DNA-binding domain in the paired homeo domain protein. *Genes & Dev.* **5**: 594–604.
- Turner, F.R. and A.P. Mahowald. 1979. Scanning electron microscopy of *Drosophila melanogaster* embryogenesis. III. Formation of the head and caudal segments. *Dev. Biol.* **68**: 96–109.
- Villée, C.A. 1945. Developmental interactions of homeotic and growth rate genes in *Drosophila melanogaster*. *J. Morphol.* **77**: 105–118.
- . 1946. Phenogenetic studies of the homeotic mutants of *Drosophila melanogaster*. IV. Homeotic and “growth rate” genes. *Genetics* **31**: 428–437.
- Vogt, M. 1947. Beeinflussung der Antennendifferenzierung durch Colchicin bei der *Drosophilamutante Aristopedia*. *Experientia* **3**: 156–157.
- Waddington, C.H. 1941. The genetic control of wing development in *Drosophila*. *J. Genet.* **41**: 75–139.
- . 1942. Growth and determination in the development of *Drosophila*. *Nature* **149**: 264–265.
- . 1943. The development of some “leg genes” in *Drosophila*. *J. Genet.* **45**: 29–43.
- . 1953. The interactions of some morphogenetic genes in *Drosophila melanogaster*. *J. Genet.* **51**: 243–258.
- Walther, C. and P. Gruss. 1991. *Pax-6*, a murine paired box gene, is expressed in the developing CNS. *Development* **113**: 1435–1450.
- Webster, P.J. and T.E. Mansour. 1992. Conserved classes of homeodomains in *Schistosoma mansoni*, an early bilateral metazoan. *Mech. Dev.* **38**: 25–32.
- Wilkins, A.S. and D. Gubb. 1991. Pattern formation in the embryo and imaginal discs of *Drosophila*: What are the links? *Dev. Biol.* **145**: 1–12.
- Wolpert, L. 1971. Positional information and pattern formation. *Curr. Top. Dev. Biol.* **6**: 183–224.