Evolution and role of *Pax* genes

Markus Noll

University of Zürich, Zürich, Switzerland

Pax genes encode a class of highly conserved transcription factors containing a paired-domain. These factors play important roles in *Drosophila* and vertebrate development, for example, in segmentation and neurogenesis. Their developmental roles are assessed in terms of their participation in conserved gene networks and mechanisms that establish positional information.

Current Opinion in Genetics and Development 1993, **3**:595–605

Discovery of the paired-box

Genes containing a paired-box, the *Pax* genes, were first found and isolated in *Drosophila* [1] in an attempt to test the gene network concept that had been postulated on the basis of evolutionary considerations [2]. Evolution from simple to complex systems proceeds by the formation of hypercycles [3] in which two reactions are coupled to the advantage of both. This idea implies that new genes that arise by duplication of old genes or by recombination of different domains must interact with the components from which they were derived. Participants in such interactions include genes and their RNA and protein products. The salient point is that successful new combinations are created by interactions between new and old components in all possible combinations (DNA–protein, RNA–protein and protein–protein), some of which will be discussed in this review. As a result, gene networks with integrated functions evolve [1,2,4,5]. According to their origin, all genes belonging to such networks consist of combinations of a relatively small number of protein-coding domains or *cis*-regulatory elements. It further follows that selection pressure also acts to conserve the integrity of the network as a whole. For this reason, networks of analogous function are expected to be conserved in widely different organisms such as *Drosophila* and man.

To test this gene network hypothesis, it was important to devise an experimental approach for finding the postulated network-specific domains. The prediction was that any member of a particular gene network shares one or several domains with other members. In principle, therefore, it should be possible to find all genes and domains belonging to a network. The *paired* (*prd*) gene was selected because it is a member of an interesting set of genes determining segmentation in *Drosophila* [6]. The transcribed portion of the *prd* gene was subdivided arbitrarily into three segments of about equal length that were then used to search for homologous sequences. Thus far, this search has led to the isolation of a number of new genes. Those which have been identified — *bicoid* (*bcd*), *gooseberry* (*gsb*), *gooseberry neuro* (*gsbn*), *pox neuro* (*poxn*) and *pox meso* (*poxm*) — belong to the same network as *prd* [1,2,4]. As expected, new domains were discovered that corresponded to the sequence homologies of the newly isolated genes, namely the PRD or His-Pro repeat (present in *bcd*), the (extended) *prd*-type homeodomain (found in *gsb* and *gsbn*), and the paired-domain (in *gsb*, *gsbn*, *poxn* and *poxm*).

Developmental significance of *Pax* genes

As predicted by the gene network hypothesis [1], *Pax* genes have been isolated from a number of other organisms. Of particular interest from a developmental as well as an evolutionary point of view is their presence in vertebrates (including mouse [7–14,15•,16], man [5,15•,17,18] and zebrafish [19–21]). In man, these genes are of considerable importance because they were found to be associated with severe developmental defects. Much light will be thrown on the etiology of these human diseases by the study of their murine homologs. Thus, *Splotch* (*Sp*) mice, mutant for *Pax-3* [22], may serve as a model for the human disorder Waardenburg's syndrome type I (WSI), which evokes in patients who are heterozygous for the *Pax-3* homolog *HuP2* [5,23•,24•] sensorineural deafness, dystopia canthorum, and pigment abnormalities of hair and iris. Heterozygous *Sp* mice exhibit a phenotype similar to that of WSI patients, while homozygous *Sp* mutants display more severe defects such as exencephaly, spina bifida, and defects in neural crest cell derivatives. Similarly, *Aniridia* (*AN*) patients, heterozygous mutants of the human *Pax-6* gene [17], are char-

Abbreviations

AN—*Aniridia*; **ANT-C**—*Antennapedia* complex; **AS-C**—*achaete-scute* complex; ***bcd***—*bicoid*; **BSAP**—B-cell lineage-specific activator protein; **BX-C**—*bithorax* complex; **CMV**—cytomegalovirus; **CNS**–central nervous system; ***da***—*daughterless*; ***en***—*engrailed*; **m-es**—mono-innervated external sensory; **p-es**— poly-innervated external sensory; ***eve***–*even-skipped*; ***gsb***–*gooseberry*; ***gsbn***—*gooseberry neuro*; **HLH**—helix-loop-helix; **Hox**—homeobox; **Pax**—paired-box; ***prd***—*paired*; **PNS**—peripheral nervous system; ***poxm***—*pox meso*; ***poxn***—*pox neuro*; ***Sd***—Danforth's short tail; ***Sey***—*Small eye*; **SMC**—sensory mother cell; ***Sp***—*Splotch*; **Tg**—thyroglobulin; **TPO**—thyroperoxidase; ***un***—*undulated*; ***wg***—*wingless*; **WSI**—Waardenburg's syndrome type I.

 ISSN 0959-437X

acterized by the complete or partial absence of the iris, whereas homozygous *Small eye* (*Sey*; homolog of human *Pax-6*) mutant mice completely lack eye and nasal structures [25].

In the following, I will discuss mainly the role of the five known *Drosophila* paired-box genes in development, their regulatory relationship within the network of genes specifying positional information along the anteroposterior axis of the embryo, and their relationship to the *Pax* genes and their networks in vertebrates (for a recent review on *Pax* genes, see [26•]).

The paired-domain is a DNA-binding domain

The paired-domain is a large, highly conserved sequence of about 130 amino acids. All known paired-domains of *Drosophila* and vertebrates are located close to the amino terminus of their proteins. Secondary structure analysis of all paired-domains strongly predicts an amphipathic α-helix (residues 23–31) and a helix-turn-helix motif (residues 80–105) in which the first helix is again amphipathic [4,5,14].

The association of the paired-domain with the DNA-binding homeodomain in *prd*, *gsb* and *gsbn* implies that the paired-domain has a gene regulatory function and that paired-box genes, in general, encode transcription factors. This conclusion is consistent with the known gene regulatory function of these genes and the nuclear localization of paired-domain proteins that contain no homeodomain [4]. Thus, the paired-domain could constitute either a DNA-binding domain or a domain specifically interacting with certain proteins; alternatively, and most probably, it could be endowed with both functions. Support for a DNA-binding function was first derived from *in vitro* binding studies of a truncated prd protein consisting essentially only of its paired-domain [27]. This paired-domain binds specifically to the downstream half of the e5 site of the *Drosophila even-skipped* (*eve*) promoter. Deletion of the third or destruction of the second α-helix does not affect binding to this site. It is, however, abolished by helix-breaking mutations in the first α-helix or by a point mutation of residue 15 [27], which gives rise to the *undulated* (*un*) mutation of murine *Pax-1* [28]. A sequence similar to the e5 site has been determined by mutational analysis to be recognized *in vitro* by murine Pax-1 protein, which contains a paired-domain but no homeodomain [11]. This sequence is also bound, although at strongly reduced affinity, by the amino-terminal portion of Pax-1 in which the paired-domain has been truncated at the end of helix 2, indicating that the paired-domain is responsible for the DNA binding. A drawback of both studies, however, is that they deal with DNA sequences for which a function *in vivo* as target sites for prd or Pax-1 protein is improbable because *eve* expression remains unaffected in *prd* mutants [29] or after ectopic expression of *prd* [30].

Recently, probable targets of two vertebrate *Pax* gene products have been identified. The Pax-8 protein activates thyroid-specific transcription of the TPO (thyroperoxidase) and probably of the Tg (thyroglobulin) gene by binding to homologous sequences in the promoters of both genes [31•]. This sequence is also bound efficiently by a truncated Pax-8 protein consisting only of its paired-domain. The B-cell lineage-specific activator protein (BSAP), the product of the human *Pax-5* gene [15•], binds *in vivo* to a defined site in the promoter of the CD19 gene to activate its B-lymphoid-specific expression [32•]. In this case, an intact paired-domain is both necessary and sufficient for *in vitro* binding of Pax-5 to its CD19 target site [15•]. Interestingly, this Pax-5 binding site, which is completely different from any of the e5-related sites, is also bound by Pax-1 protein, although with a much lower affinity. In contrast, Pax-5 binds to the e5-related site, to which Pax-1 binds with the highest affinity, with a twofold higher affinity than Pax-1 and only with a twofold lower affinity than to the CD19 promoter site used *in vivo* [15•]. Therefore, while paired-domains are clearly involved in DNA binding, the relevance of binding specificities determined solely *in vitro* are uncertain because these might depend *in vivo* on additional unknown parameters such as the interaction of the paired-domain with other transcription factors.

Role of *paired* and *gooseberry* in segmentation

The anlagen that give rise to the metameric pattern observed at later stages in *Drosophila* embryos are determined by the time of cellular blastoderm. Thus, segmentation is initiated very early in *Drosophila* development by a gene regulatory cascade consisting of the classes of gap, pair-rule, segment-polarity and homeotic genes that progressively refine positional information in a combinatorial manner along the anteroposterior axis of the early embryo (for reviews, see [33,34•]). The metameric organization of the embryo is most conspicuous in the repetitive patterns elaborated by the epidermis, but is also evident in the ectodermally derived central nervous system (CNS) and peripheral nervous system (PNS) as well as in the mesoderm. While many of the segmentation and homeotic genes are re-deployed to specify position during neurogenesis, specification of position within the mesoderm appears to be controlled by additional classes of genes.

The three most closely related *Drosophila* paired-box genes, *prd*, *gsb* and *gsbn*, are segmentation genes of the pair-rule and segment-polarity class. Their products are highly homologous in their amino-terminal halves (85%), consisting essentially of a paired-domain and an extended *prd*-type homeodomain [1], but exhibit little or no homology in their carboxy-terminal halves [35]. As all pair-rule genes, *prd* integrates the aperiodic positional information expressed by gap genes into a pattern of stripes exhibiting a double-segment periodicity [36•]. Interestingly, this initial pair-rule pattern of *prd* expression is only transient and is converted, by the end of cellularization at blastoderm, to a pattern displaying a single-segment repeat that is characteristic for segment-polarity genes [36•,37]. This transition in *prd* expression results from the ability of *prd* to integrate positional information of a double-segment periodicity supplied by the other pair-rule genes into a single-segment periodicity [38], in a process similar to the initial activation of segment-polarity genes [34•]. Thus, the expression of *prd* exhibits both pair-rule and segment-polarity gene characteristics.

The transition from a pair-rule to a segment-polarity type expression of *prd* may reflect the evolution from short to long germband insects. One might imagine that duplication of the *prd/gsb* ancestral gene delegated its segment-polarity function to the *gsb* gene, while the *prd* gene acquired the new or more recent pair-rule function. It is not clear, however, whether *prd* still retains some of the original segment-polarity function.

The initial activation of the segment-polarity genes *engrailed* (*en*), *wingless* (*wg*) [34•], and *gsb* [4] has been shown to depend on *prd* at least in every other stripe (Fig. 1a). These genes are probably direct targets of the prd protein. Indeed, *cis*-regulatory target sites have recently been identified for *gsb* that respond to the prd protein *in vivo* [39•] and interact with it *in vitro* (X Li, M Noll, unpublished data). It is not known whether these sites, which show no similarity to the e5 sequence, interact with the paired-domain and/or the homeodomain of prd.

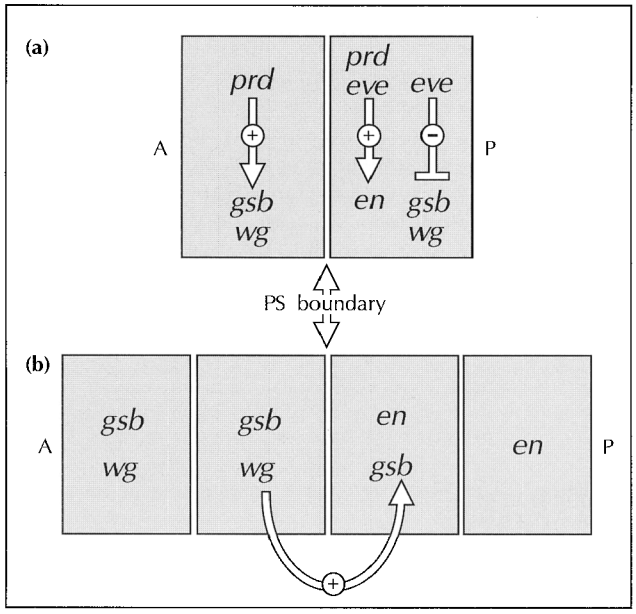

Fig. 1. Activation by prd of the odd-numbered gsb, wg and en stripes. **(a)** The initial activation of *gsb*, *wg* and *en* at cellular blastoderm and **(b)** after cell division during gastrulation is illustrated schematically in cells anterior (A) and posterior (P) to the parasegmental (PS) boundary. **(a)** The activation (+) by prd of odd-numbered gsb and wg stripes is counteracted by another pair-rule protein, eve, repressing (−) *gsb* and *wg* in posterior cells. The *en* gene is activated only in posterior cells as its activation depends on both *prd* and *eve*, which is not active in anterior cells. The even-numbered gsb, wg, and en stripes (not shown) are similarly established by the combinatorial action of pair-rule gene products [34•] (X Li, M Noll, unpublished data). **(b)** After cell division, *gsb* is activated by the wg signal also in cells posterior to the PS boundary [39•].

Although the prd protein is strictly required for the activation of odd-numbered en, gsb and wg stripes, it does not activate them in congruent stripes. Thus, the wg stripes are activated anterior and the en stripes posterior to the parasegmental boundary that they establish, while the gsb stripes coincide anteriorly with those of wg and overlap posteriorly with the anterior portion of en stripes (Fig. 1). The precise locations of the odd-numbered en, wg and gsb stripes depend on the differential role of other pair-rule genes as, for example, on the *eve* product that activates *en* but represses *wg* [34•] and *gsb* (Fig. 1a) (X Li, M Noll, unpublished data). The maintenance of segment-polarity gene expression also depends on interactions among these genes. Thus, the posterior boundary of odd-numbered en stripes is determined by prd protein [30,36•], while the posterior limit of gsb stripes initially coincides with and depends on the anterior border of eve stripes (Fig. 1a) but is later shifted posteriorly due to the activation by wg [39•] expressed in the anterior neighbouring cells (Fig. 1b).

The initial activation of the *gsb* gene by a combination of pair-rule gene products generates a striped segment-polarity pattern with single-segment periodicity (Fig. 1a). During germ band extension, *gsb* expression begins to depend on the extracellular product of the *wg* gene (Fig. 1b). A signal transduction cascade, initiated by the wg signal and probably involving several segment-polarity gene products [39•], takes over the activation of *gsb* from prd and other pair-rule proteins. During this transition, *gsb* expression is restricted to the ventrolateral neurogenic region [40•]. In combination with some 15 other segment-polarity genes, *gsb* serves to maintain and elaborate positional information along the anteroposterior axis within each primordial segment during embryogenesis (for a recent review, see [41•]). In this prepatterning process of the epidermis that determines the larval cuticular pattern, the fundamental role of *gsb* is to activate *wg*, thereby ensuring the propagation of the wg signal (X Li, M Noll, unpublished data).

Role of *gooseberry* in epidermal and of *gooseberry neuro* in CNS pattern formation

A surprising outcome of our search for genes sharing putative domains with the *prd* gene was that two genes, *gsb* and *gsbn*, rather than one were found [1] at the genetically defined segment-polarity locus of *gsb* [6]. As they are spaced by less than 10 kb and both encode highly homologous (85%) paired-domains and extended *prd*-type homeodomains, the possibility was considered that inactivation of both genes is necessary to generate the *gsb* cuticular phenotype, particularly because all known *gsb* mutants were deletions eliminating the activity of both genes. This supposition appears to be incorrect as only one of the two genes, *gsb*, suffices to rescue the *gsb* cuticular phenotype of these mutants [40•]. However, it remains to be demonstrated whether the sole lack of *gsb* function generates a *gsb* cuticular phenotype. In other experiments, *gsb* was shown to activate *gsbn* in neuroblasts. In contrast to *gsb*, *gsbn* continues to be expressed in the CNS in what appears to be the progeny of *gsb*-expressing neuroblasts [40•]. It seems probable, therefore, that *gsbn* rather than *gsb* is mainly responsible for the *gsb* CNS phenotype whose most prominent feature is the failure of proper formation of posterior commissures [42]. These differential roles of *gsb* and *gsbn* probably evolved from a single gene that was responsible for both functions (X Li, M Noll, unpublished data).

Role of *pox neuro* in neurogenesis

The two paired-box genes *poxn* and *poxm* contain no homeobox. Their segmentally repeated expression patterns depend on the activities of the segmentation genes and are restricted to single germ layers [4]. The *poxn* gene is first expressed in specific sensory mother cells (SMCs) of the PNS and in neuroblasts of the CNS during neurogenesis [4,43•]. Two SMCs and their progeny express *poxn* in each hemisegment of the embryonic thorax and abdomen. These give rise to the poly-innervated external sensory (p-es) organs of the peripheral nervous system. Indeed, failure of *poxn* expression converts p-es into mono-innervated external sensory (m-es) organs, while ectopic expression of *poxn* generates supernumerary p-es organs [43•]. Thus, *poxn* determines p-es versus m-es organ development and acts as a switch between these different fates of es SMCs. This function of *poxn* is not limited to the larval PNS but is, at least in some imaginal structures, extended to the adult PNS [43•,44•]. Hence, similar to the homeobox gene *cut* that specifies es versus chordotonal organ development [45•], *poxn* specifies the identity of SMCs in combination with other genes and thus is a 'neuroblast identity' gene [46•]. Extending this concept of *poxn* function from the PNS to the CNS where *poxn* is also first expressed in two specific neuroblasts per hemisegment [43•], one might speculate that *poxn*, in combination with other neuroblast identity genes, determines the fate of specific neuroblasts and their progeny.

Similarity of *pox meso* and *Pax-1* expression patterns

Similar to *gsb*, *poxm* expression depends completely on the prd protein in odd-numbered bands [4]. Poxm protein first appears during germ band extension in the posterior half of each segment in the somatic mesoderm (Fig. 2a). At subsequent stages, *poxm* continues to be expressed in a segmentally repeated subpopulation of cells of the somatic mesoderm until it is expressed in specific muscles of the larval body wall (W Fu, E Jamet, M Noll, unpublished data). The phenotype of *poxm* mutants and thus the role *poxm* might play in myogenesis is not yet known.

It is striking that the *Pax* gene encoding the paired-domain most closely related to that of *poxm* (86% identity), *un* or *Pax-1*, is also expressed in a segmentally repeated pattern of mesodermal cells (Fig. 2b). Both *poxm* in *Drosophila* and *un* in the mouse are expressed only after mesoderm formation, but there are important differences in how the periodic patterns arise. The periodicity of *poxm* expression is based on its direct or indirect activation by pair-rule proteins whereas *Pax-1* is initially expressed in a non-periodic pattern in sclerotome cells during somite differentiation [7]. Its expression in a segmental repeat presumably occurs by migration of the *un*-expressing sclerotome cells into the perichordal region and their subsequent condensations along the rostrocaudal axis where they form the anlagen of the intervertebral

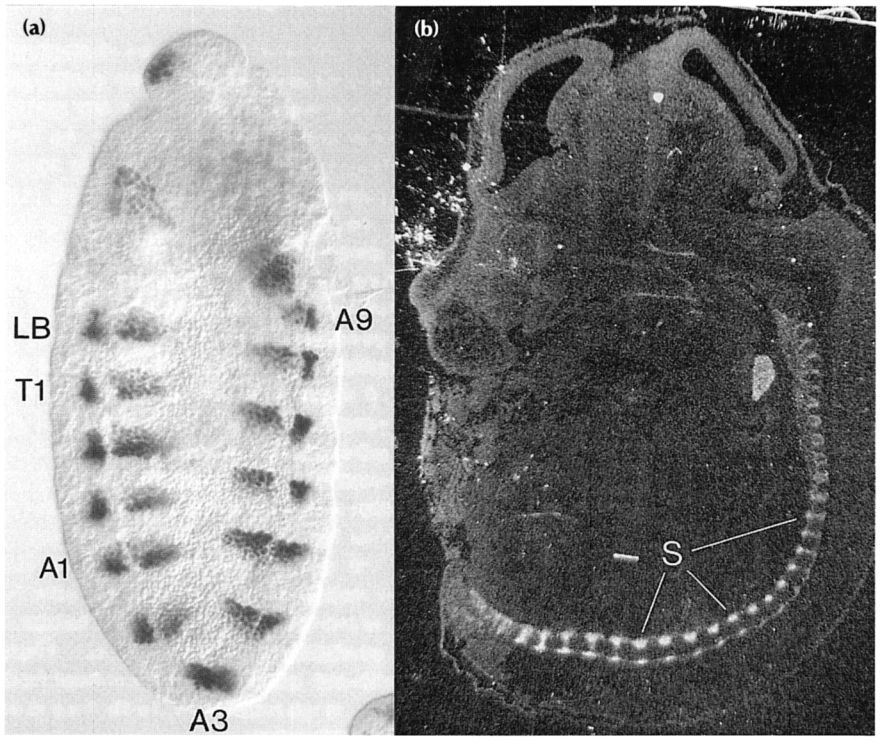

Fig. 2. Segmentally repeated expression of *poxm* and *Pax-1* (*un*) in the embryonic mesoderm of *Drosophila* and mouse. **(a)** Expression of *poxm* is visualized by an anti-*poxm* antiserum in a whole-mount *Drosophila* embryo during germ band extension (stage 10). At this stage, the segment primordia extend around the posterior pole of the embryo (A3) with the most posterior segments located at the anterior-most dorsal position of the germ band. A bilaterally symmetric expression pattern with a single segment periodicity is observed in the somatic mesoderm, including the labial head segment (LB), the three thoracic segments, T1–T3, and the abdominal segments, A1–A9 (photograph, courtesy of W Fu and X Li). **(b)** Expression of *Pax-1* is revealed in a near-midsagittal section of a 14 days *post coitum* mouse embryo by *in situ* hybridization with a ^{35}S-labeled *Pax-1*-specific RNA probe and subsequent photography under dark-field illumination. At this stage, *Pax-1* expression exhibits a somitic repeat (S) in the intervertebral disk anlagen (photomicrograph from [7], with permission).

disks [7]. The periodicity of *un* expression thus follows that of the somites and, therefore, is probably indirectly induced by processes that generate the somitic repeat.

Classes of paired-domains

Paired-domains fall into different classes according to several criteria: class-specific amino acids, degree of conservation, sequence organization (location of introns), and association with other domains such as the homeodomain or the HSIDGILG (one-letter amino acid code) octapeptide [4,5]. Three of the five *Drosophila* paired-box genes, *prd*, *gsb* and *gsbn*, belong to the class of *PHox* genes characterized by a highly conserved class of paired-domains as well as a *prd*-type homeodomain [1]. In contrast, *poxn* and *poxm* encode separate classes of paired-domains and no homeodomain [4]. This establishes the paired-domain as an independent domain, contrary to the POU-specific domain, which until now has always been found to be associated with a homeodomain [47]. The octapeptide has only been found in association with a paired-domain or a homeodomain (Table 1). Interestingly, in all cases analyzed, the octapeptide is encoded by the same exon as the paired-domain. Its function is unknown.

Additional classes of paired-domains, associated with or without a *prd*-type homeodomain, have been found in vertebrates [14,48], raising the obvious question of whether these classes of paired-domains also exist in *Drosophila* or whether they have evolved more recently. The division of paired-domains into different classes follows naturally from their evolutionary history as outlined below.

Evolution of *Pax* genes

Considering the degree of homology among the known paired-domains of vertebrates and insects [14], it is clear that at the time of the separation of deuterostomes from protostomes at least two and perhaps as many as four different ancestral paired-box genes existed (Fig. 3). One of them encoded a paired-domain associated with an octapeptide and evolved into *poxm* and *Pax-1*. Another contained a paired-domain, an octapeptide, and a homeodomain and gave rise to *gsb*, *gsbn* and *prd* in *Drosophila* and to *Pax-3* and *Pax-7* in *Mus*. As the paired-domain of *poxn* is more closely related to that of *Pax-2*, *Pax-5* or *Pax-8* than to any other known paired-domain, these genes probably also had a common ancestral gene when protostomes and deuterostomes separated. In agreement with this conclusion, the *Pax-2/-5/-8/poxn* line must have separated from the *Pax-1/poxm* or the *Pax-3/-7/prd/gsb/gsbn* line very early because all three lines have diverged to the same extent from each other. *Pax-4* and *Pax-6* are most closely related to each other [13] and have clearly diverged less from *Pax-2/-5/-8/poxn* than from the other two lines, indicating that they separated from the *Pax-2/-5/-8/poxn* line. At least one additional paired-box gene encoding both a paired-domain and a *prd*-type homeodomain exists in *Drosophila* (cited in [27]). As an extensive previous search for additional *Drosophila* paired-box genes was negative (M Noll, unpublished data), the paired-domain of this newly discovered, still unpublished *Drosophila* paired-box gene has probably diverged considerably from the five known *Drosophila* paired-boxes used as probes in this search. Hence, it might well have the same origin as *Pax-4* and *Pax-6* of the mouse. Therefore, it is probable that *Pax-4/-6* began to diverge from *Pax-2/-5/-8/poxn* also before separation of proto- and deuterostomes, and thus four *Pax* genes existed at this time. Assuming that this last separation occurred around the time of coelomate divergence into protostomes and deuterostomes, nearly 700 million years ago, we can roughly extrapolate the time of appearance of the first paired-boxes to around 850 million years ago (Fig. 3).

The evolutionary trees constructed on the basis of the degree of paired-domain conservation (Fig. 3) are consistent with the acquisition and loss of other domains and their conservation during evolution. For example, the *prd*-type homeodomain, which had been combined with

Table 1. Octapeptides of paired-domain and homeodomain proteins.

Protein	Octapeptide	Distance
Gsb	H S I D G I L G	5
Gsbn	Y T - N - - - -	9
Pax-7	- - - - - - - -	20
Pax-3	- - - - - - - S	22
Poxn	Y - - E D L - K	225
Pax-2	Y - - N - - - -	40
Pax-8	Y - - N - L - -	42
Pax-5	Y - - S - - - -	34
Poxm	- - V S S - - A	138
Pax-1	- - V S N - - -	63
En	F - - S N - - S	270
Inv	F - - - N - - K	183
H2.0	F - V - R - - -	214
Hlx	F G - - R - - S	92

The octapeptide, first found to be conserved between the human *PAX-3* (*HuP2*), *PAX-7* (*HuP1*) and the *Drosophila gsb* and *gsbn* genes [5], is listed for known *Drosophila* [43•] (M Noll, unpublished data) and mouse *Pax* genes [9–12,15•] in the three upper groups. No octapeptide is present in *prd* [5], *Pax-4* and *Pax-6* [26•]. The bottom group includes four octapeptides found to be encoded by *Drosophila* (*en*, *invected* (*inv*) and *H2.0*) and mouse (*Hlx*) homeobox genes that contain no paired-box [70]. These groups suggest different classes of octapeptides reflecting their evolutionary history (Fig. 3), with the exception of the Gsbn octapeptide, which is closer to the octapeptides of the Pax-2 group. The distance of the octapeptides (in amino acids) from the carboxy-terminal end of the paired-domains (upper three groups) or from the amino-terminal end of the homeodomains (bottom group) is indicated. For all known cases, the octapeptide is encoded by the same exon as the carboxy-terminal portion of the paired-domain but by an exon upstream from that or those encoding the homeodomain.

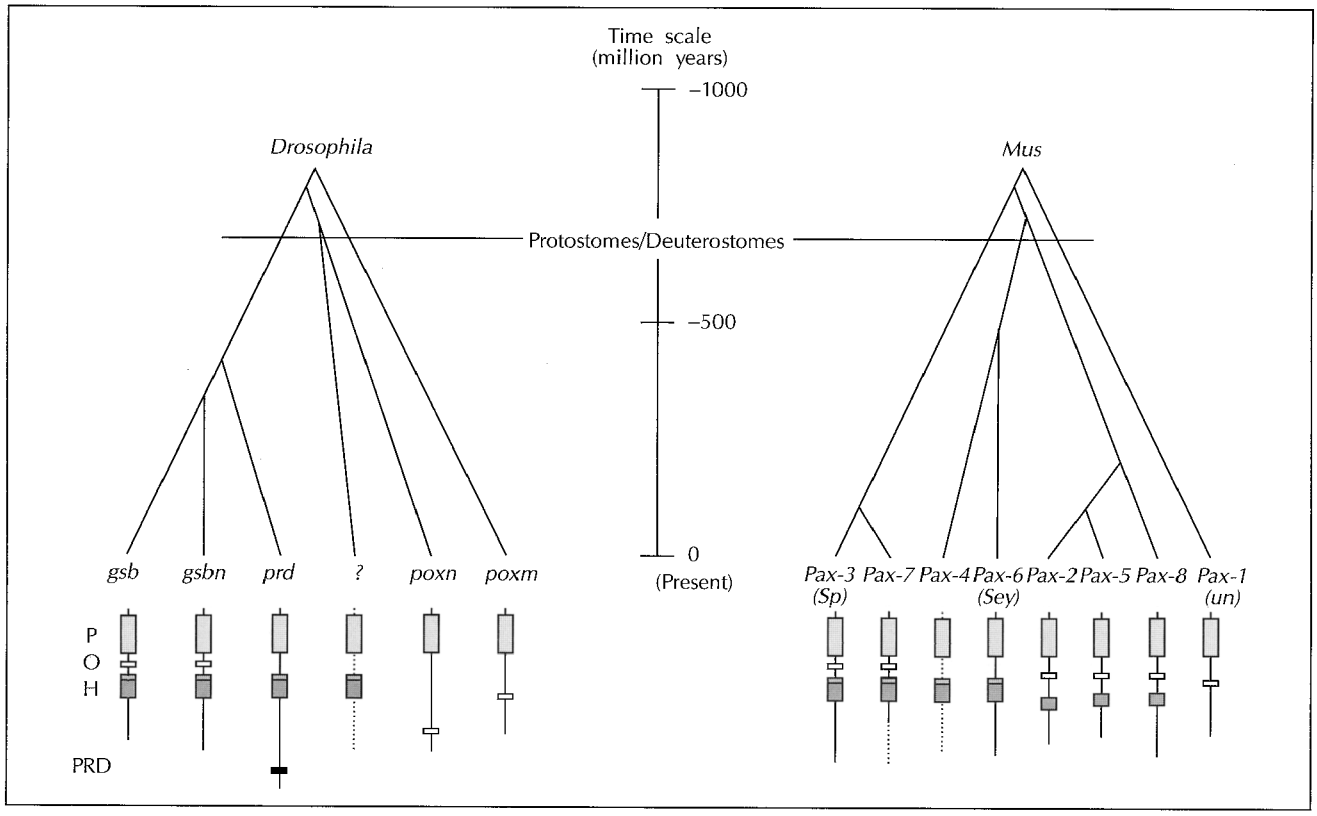

Fig. 3. Evolutionary trees of *Drosophila* and mouse paired-box genes. The evolutionary trees are constructed from the known amino acid differences between paired-domains of *Drosophila* and mouse (*Mus*) [14]. After the separation of protostomes and deuterostomes, the trees leading to the present *Drosophila* and *Mus* paired-box genes have evolved separately, originating from four common ancestral genes. In man, homologs of *Pax-3* (*HuP2*), *Pax-7* (*HuP1*), *Pax-6* (*AN*), *Pax-5*, *Pax-8* and *Pax-1* (*HuP48*) have been isolated [5,15•,17,18]. Below the trees, the composition of each paired-domain protein from various domains is indicated schematically: P, paired-domain; O, octapeptide; H, *prd*-type homeodomain; PRD, His-Pro repeat. Note that *Pax-2/-5/-8* have only retained the amino-terminal half of the *prd*-type homeodomain without its amino-terminal extension [13,15•,20]. (For detailed explanation, see text.)

a paired-domain and octapeptide (Table 1), was partially lost in *Pax-2/-5/-8* [13,15•,20], while it was retained in *Pax-4/6* which, however, failed to conserve the octapeptide [26•]. As argued above, a *Pax-4/-6* line might also exist in protostomes. The evolutionary tree of mouse *Pax* genes is also roughly consistent with respect to its time scale. It predicts that the duplications that produced *Pax-2*, *Pax-5* and *Pax-8* occurred after the separation of the lines leading to zebrafish and mammals about 400 million years ago. Indeed, in zebrafish only one *Pax-2/-5/-8* homolog, *pax[zfb]*, has been found that is expressed in regions of the developing brain, eye, ear, neural tube and kidney [20] that correspond to the sum of the regions in which all three mouse genes are expressed [8,9,15•,16,48–51]. Hence, in zebrafish, *pax[zfb]* might be the only gene derived from the *Pax-2/-5/-8* ancestral gene [16].

One might wonder why so few *Pax* genes have evolved, as compared for example to *Hox* genes, even though they have been in existence for a long time. The paired-domain is highly specialized and of a relatively large size that might have restricted its versatility during evolution. Indicative of such an extreme specialization of paired-domains during evolution is the high conservation not only of the paired-domain but of the entire protein in homologs of vertebrates. For example, only one amino acid change is observed in the *Pax-6* paired-domains of zebrafish [19,21] and mouse [13], while their proteins exhibit 96% identity over a length of 422 amino acids. In contrast, the enhanced versatility of the smaller homeodomain or helix-loop-helix (HLH) domain favoured the evolution of large *Hox* or HLH gene clusters by repeated duplications. These gene clusters probably evolved integrated properties or functions, such as providing positional information for segment identity [52•] or proneural cell clusters [45•], that allowed their repeated use during evolution by duplications of the entire cluster. The integration of genes into clusters — *bithorax* complex (BX-C), *Antennapedia* complex (ANT-C), and *achaete-scute* complex (AS-C) — and their conservation is reflected by and depends on their coordinate regulation. This property and the interaction of one gene with the product of the duplicated gene or the specific interactions among their products, for example, as heterodimers, again illustrate the principle of hypercycles. Such higher integrations of *Pax* genes into clusters apparently have not evolved (with the exception of the mini-cluster of *gsb* and *gsbn*) nor is there any evidence that paired-domain proteins are capable of heterodimer formation [15•].

Conservation of *Pax* gene networks during evolution

For species that have evolved separately for as long a time as *Drosophila* and mouse or man, we expect only fundamental mechanisms of development to be conserved yet little resemblance between detailed morphogenetic processes that evolved later. Indeed, the high degree of conservation of the paired-domain and octapeptide (Table 1) in *poxm* and *Pax-1* must reflect a stringent conservation of their interactions with other genes or gene products and thus a considerable conservation of the networks to which these genes belong. Consistent with this view is the observation that both genes are expressed exclusively in mesodermal cells soon after this germ layer has been established (Fig. 2). Although their segmentally repeated expression patterns appear to arise in widely different ways, it is conceivable that their underlying molecular mechanisms are similar and that the segmental periodicity is generated in both cases by a periodic prepattern of combinations of active regulatory genes. The existence of such a prepattern in the mouse is evident from the appearance of somites preceding *Pax-1* activation.

The activation of *poxm* clearly depends on *prd* [4] and perhaps also involves *gsb* that is expressed in the mesoderm as well [35,40•]. Hence, if similar interactions have been conserved among their *Pax* gene homologs, we might expect *Pax-3* or *Pax-7* to activate *Pax-1*. Consistent with such regulatory interactions, both *Pax-3* and *Pax-7* are expressed in the mesoderm [10,12]; however, *Pax-7* can be excluded because its expression coincides with rather than precedes that of *Pax-1* and occurs in dermomyotome but not in sclerotome cells, to which the somitic expression of *Pax-1* is restricted [7]. A better candidate for the activation of *Pax-1* is *Pax-3*, which is expressed in somites before their differentiation into sclerotome and dermomyotome cells. It is thus possible that similar interactions among *Pax* genes lead to the activation of *Pax-1* and *poxm*.

An additional analogy might exist between these two networks, namely, if *Pax-3*, whose expression precedes that of *Pax-7* in neuroblasts of the alar plate, would activate *Pax-7* as *gsb* activates *gsbn* in neuroblasts during neurogenesis. This similarity, however, would be based on convergent evolutionary processes because the duplication of both the *gsb/gsbn* and of the *Pax-3/-7* ancestors occurred after the protostome–deuterostome separation (Fig. 3). In fact, such an evolutionary process in which a hypercycle evolves by one of the duplicated regulatory genes regulating the other might occur rather frequently. Examples in *Drosophila* are both of the duplications that produced the *prd*, *gsb* and *gsbn* genes (Fig. 3) and the HLH gene cluster of the AS-C [45]. Another such duplication may have generated *Pax-2* and *Pax-8*, which are both expressed in the same regions of the developing neural tube, brain and kidney. In all of these developing organs *Pax-2* expression always precedes that of *Pax-8* [8,9,48,49].

Another property that appears to be conserved between *Pax-3/-7* and *prd/gsb/gsbn* is their expression in the developing nervous system. Similarly, *poxn* and *Pax-2/-5/-8*, which have originated from a common ancestral gene, are expressed predominantly in the CNS. In fact, all *Pax* genes in *Drosophila* and vertebrates except *Pax-1* and *poxm* are expressed in the developing nervous system. Evidently, at least the specificity for germ layers has been largely conserved in the expression of *Pax* genes derived from the same ancestral coelomate genes.

Strikingly similar expression patterns of *Drosophila* and vertebrate *Pax* genes observed at later developmental stages, such as the expression of *poxm* in somatic muscles (W Fu, E Jamet, M Noll, unpublished data) and of *Pax-7* in the skeletal musculature [10], are more difficult to interpret. They would have to be the result of convergent evolution as the *Pax-3/-7* line separated from the *Pax-1/poxm* line before the divergence of proto- and deuterostomes. This convergence may have been driven by the putative regulatory linkage of *Pax-3/-7* and *Pax-1* that ensured their expression in various compartments of the somites. The selection of which compartment, the dermomyotome or the sclerotome lineage, the expression of the *Pax-3/-7* and *Pax-1* ancestral genes was allocated to might have been accidental when the first vertebrates appeared about 500 million years ago.

The *gooseberry–wingless* autoregulatory loop: an ancient control mechanism of morphogenetic processes

In *Drosophila*, other genes are known to interact with some of the paired-box genes. Particularly, the segment-polarity gene *gsb* has been shown to be activated by the *wg*-encoded signal and itself serves to activate the *wg* gene (Fig. 4a). This type of autoregulatory loop could link the expression of these genes in neighbouring cells. As such a mechanism might, in principle, serve to propagate their expression over many cells, it could be a fundamental process of pattern formation.

That this mechanism already existed before the separation of proto- and deuterostomes is suggested by comparing recent experiments on the role of *pax[zf-b]* in zebrafish with that of *gsb* and *wg* in *Drosophila*. Injection of an antiserum specific for pax[zf-b] protein into fertilized zebrafish eggs resulted in failure of proper formation of the midbrain–hindbrain boundary and subsequent cell death in the region of the posterior midbrain and cerebellum [53•]. Before the time of furrow formation between mid- and hindbrain, expression of three genes normally expressed in this region [20,50,51,54,55] is strongly reduced or abolished [53•]. These genes are *pax[zf-b]* itself and the zebrafish homologs of *wg* (*wnt-1*) and *en* (*eng-2*). It is significant that these results observed in zebrafish can be explained by the same regulatory interactions as those of the *Drosophila* genes *gsb*, *wg* and *en* (Fig. 4). According to the regulatory interactions indicated in Fig. 4b, inactivation of the pax[zf-b] protein by injection of the corresponding antiserum interrupts the *pax[zf-b]-wnt-1* autoregulatory loop and thereby the expression of *pax[zf-b]* and *wnt-1*. Assuming that zebrafish *eng-2* depends on *wnt-1* as *en* expression depends on

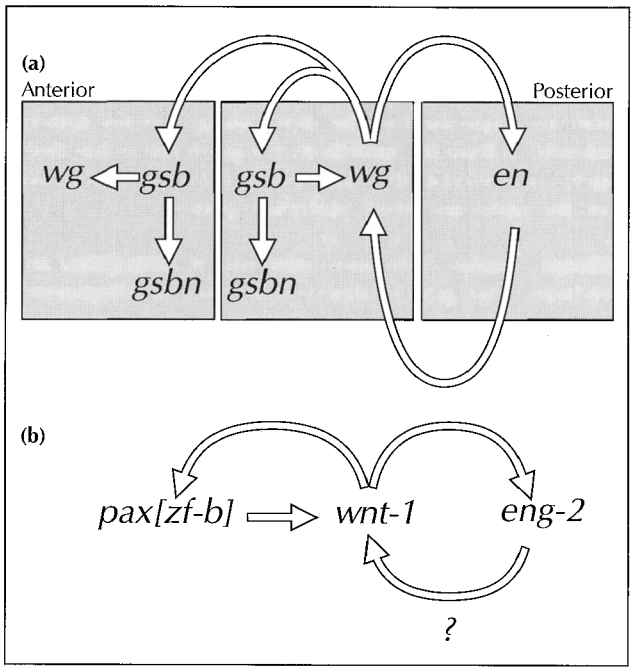

Fig. 4. The *gsb–wg* autoregulatory loop, a paradigm of a new pattern-forming process. **(a)** Interactions among the segment-polarity genes *gsb*, *wg*, and *en* in neighbouring cells along the anteroposterior axis. Arrows indicate activation by signal transduction or by a (hypothetical) direct interaction of a gene's product with its target. **(b)** Hypothesized *pax[zf-b]–wnt-1* autoregulatory loop in zebrafish homologous to that of *gsb* and *wg* in *Drosophila*. On the basis of recent observations [53•], regulatory interactions between *pax[zf-b]*, *wnt-1*, and *eng-2* are suggested that are homologous to those observed for the homologous *Drosophila* genes *gsb*, *wg*, and *en*. The question mark refers to an activating pathway drawn in analogy to the situation in *Drosophila*, but for which no evidence exists in zebrafish. (For further explanation, see text).

wg in *Drosophila*, inactivation of the *pax[zfb]* product further results in the failure of *eng-2* activation in agreement with the observations [53•].

In the mouse, where similar expression patterns of *Pax-5*, *Wnt-1*, and *En-1* occur at the midbrain–hindbrain boundary [15•,16,56,57], the same mechanism is likely to be at work. In this case, it is known that homozygous *Wnt-1* mutants show a complete loss of the midbrain and the adjacent cerebellum of the hindbrain [58–60] preceded by the corresponding failure of *En* expression in this region [61•]. *Pax-5* expression has not been examined in these mutants, but is expected to be eliminated in this region as well, according to the model shown in Fig. 4b. The homology between the *Drosophila* and zebrafish gene networks is not complete. The *gsb* gene encodes both a paired-domain and a homeodomain [1], while *pax[zfb]* [20] as well as mouse *Pax-5* [15•] have lost the carboxy-terminal half of the *prd*-type homeodomain in their ancestral gene. However, the mechanism depicted in Fig. 4a is general. In principle, it requires only a specific transcription factor (e.g. gsb) that activates a gene producing an extracellular signal (e.g. wg), which in turn activates, by signal transduction, the gene encoding the transcription factor. In particular, such a mechanism need not necessarily be restricted to *Pax* genes. For example, a similar mechanism might occur during axis formation by the Spemann's organizer in *Xenopus*, where injection of RNA derived from the homeobox gene *goosecoid* [62] or from the *Wnt* family member *Xwnt-8* [63,64] results in the ectopic formation of a second axis.

In addition to zebrafish *pax[zfb]* or mouse *Pax-5*, a number of *Pax* genes might use this mechanism in inductive processes during development. This autoregulatory process might be combined with the duplication of a *Pax* gene in which one of the duplicated genes becomes dependent on the other, as in the case of *gsb* and *gsbn* (Fig. 4a). Observations of *Pax* gene expression suggest that similar mechanisms operate in the mouse, particularly during inductive events [48]. An instructive example might be the induction of mesenchymal condensations around the ureter during nephrogenesis [65•]. In this case, *Pax-2* is expressed in the branching ureteric buds. Subsequently, the surrounding mesenchymal cells condense around the ureter and begin to express both *Pax-2* and *Pax-8*. Expression of *Pax-8* remains restricted to the mesenchymal condensations, while *Pax-2* is also expressed in the ureter [8,9]. The induction of *Pax-8* expression might be explained by the activation by *Pax-2* of a gene producing an extracellular signal, for example, Wnt, that in the mesenchymal cells activates *Pax-2* whose product in turn activates *Pax-8*. *Pax-2* expression is repressed during maturation of the mesenchyme-derived renal tubules and glomeruli. Consistent with a determinative role of *Pax-2* in early nephrogenesis is the recent demonstration that in transgenic mouse embryos, where *Pax-2* is expressed under the human cytomegalovirus (CMV) immediate early gene promoter, persistent *Pax-2* expression prevents terminal differentiation of kidneys [66•]. This situation is also found in the undifferentiated renal epithelium of human Wilms' tumours [67] that likewise continue to express *Pax-8* [18]. Moreover, *Pax-8* maps to the same region as Danforth's short tail (*Sd*) mutations [9,18], which are characterized by the reduction or absence of kidneys.

Analogous examples of regulatory networks of *Pax* genes might include the induction of *Pax-6* in the developing lens and cornea by the preceding *Pax-6* expression in the optic vesicle and optic cup [13,19,21], or the induction of *Pax-2* expression in the otic vesicle [49]. As demonstrated recently for chick embryos [68•], a similar mechanism might establish the polarity of the developing spinal cord [69] reflected in the differential expression of several *Pax* genes along the dorsoventral axis of the neural tube [10,12,13,26•]. This polarity is induced by the notochord and the floor plate of the neural tube [68•,69]. In this case, expression of *Pax-6* depends on a signal emanating from the floor plate whereas another signal activates *Pax-3* in cells of the dorsal neural folds [68•]. Expression of *Pax-3* and *Pax-6* might then be propagated in opposite directions along the dorsoventral axis of the developing spinal cord [68•] by autoregulatory loops as illustrated above for *gsb* and *wg* (Fig. 4a). Although the boundary between *Pax-3* and *Pax-6* expression is located initially in the dorsal half of the spinal cord, it is subsequently shifted ventrally until an apparent equilibrium is reached midway between the roof plate and floor plate [68•]. This shift is presumably achieved by the repression of *Pax-6* by *Pax-3* and *vice versa*, an explanation that would agree well with the ob-

served altered patterns of *Pax-3* and *Pax-6* expression after removal of pieces of the notochord or after implantations of ectopic notochord or floor plate fragments [68•].

Although some of these ideas might appear rather speculative, the availability of *Pax* gene mutants will undoubtedly soon test their heuristic value experimentally.

Conclusions

Pax genes encode transcription factors in which the DNA-binding paired-domain is frequently associated with a second DNA-binding domain, the *prd*-type homeodomain. The discovery of the paired-domain was based on evolutionary considerations summarized by the gene network hypothesis. A mechanistic explanation for this hypothesis is derived from Eigen's principle of hypercycles [3].

Paired-domains fall into separate classes that are interpreted in terms of their evolutionary history. Probably at least four paired-box genes existed at the time of protostome–deuterostome separation.

Some of the interactions among *Pax* genes are probably very old and might have been conserved in *Drosophila* and vertebrates. In addition, interactions of *Pax* genes or their products with other genes or their products seem to be conserved, suggesting that considerable portions of these gene networks have been conserved. In particular, the *gsb–wg* and *pax[zf-b]–wnt-1* autoregulatory loops appear to be conserved in *Drosophila* and zebrafish and might represent paradigms of a fundamental mechanism of pattern formation which, in *Drosophila*, involves other segment-polarity genes. It will be exciting to see whether their zebrafish or mouse homologs participate in the same autoregulatory loops forming a conserved gene network that has been used repeatedly during evolution to regulate morphogenetic events.

Acknowledgements

I thank Hans Noll for stimulating discussions and criticism of a first draft of this manuscript, and Leslie Pick, Xuelin Li, Daniel Bopp, Bill McGinnis, and Hans Noll for helpful suggestions and comments on a later version. I am also indebted to Fritz Ochsenbein for artwork and photography. Particularly, I would like to express my gratitude to my former and present colleagues who contributed to and shared the excitement of working on *Drosophila* and human *Pax* genes with me.

References and recommended reading

Papers of particular interest, published within the annual period of review, are highlighted as:

- • of special interest
- •• of outstanding interest

1. BOPP D, BURRI M, BAUMGARTNER S, FRIGERIO G, NOLL M: **Conservation of a Large Protein Domain in the Segmentation Gene *paired* and in Functionally Related Genes of *Drosophila*.** *Cell* 1986, 47:1033–1040.

2. FRIGERIO G, BURRI M, BOPP D, BAUMGARTNER S, NOLL M: **Structure of the Segmentation Gene *paired* and the *Drosophila* PRD Gene Set as Part of a Gene Network.** *Cell* 1986, 47:735–746.

3. EIGEN M: **Selforganization of Matter and the Evolution of Biological Macromolecules.** *Naturwiss* 1971, 58:465–523.

4. BOPP D, JAMET E, BAUMGARTNER S, BURRI M, NOLL M: **Isolation of Two Tissue-Specific *Drosophila* Paired Box Genes, *pox meso* and *pox neuro*.** *EMBO J* 1989, 8:3447–3457.

5. BURRI M, TROMVOUKIS Y, BOPP D, FRIGERIO G, NOLL M: **Conservation of the Paired Domain in Metazoans and its Structure in Three Isolated Human Genes.** *EMBO J* 1989, 8:1183–1190.

6. NÜSSLEIN–VOLHARD C, WIESCHAUS E: **Mutations Affecting Segment Number and Polarity in *Drosophila*.** *Nature* 1980, 287:795–801.

7. DEUTSCH U, DRESSLER GR, GRUSS P: ***Pax 1*, a Member of a Paired Box Homologous Murine Gene Family, is Expressed in Segmented Structures During Development.** *Cell* 1988, 53:617–625.

8. DRESSLER GR, DEUTSCH U, CHOWDHURY K, NORNES HO, GRUSS P: ***Pax2*, a New Murine Paired-Box-Containing Gene and its Expression in the Developing Excretory System.** *Development* 1990, 109:787–795.

9. PLACHOV D, CHOWDHURY K, WALTHER C, SIMON D, GUENET J-L, GRUSS P: ***Pax8*, a Murine Paired Box Gene Expressed in the Developing Excretory System and Thyroid Gland.** *Development* 1990, 110:643–651.

10. JOSTES B, WALTHER C, GRUSS P: **The Murine Paired Box Gene, *Pax7*, is Expressed Specifically During the Development of the Nervous and Muscular System.** *Mech Dev* 1990, 33:27–37.

11. CHALEPAKIS G, FRITSCH R, FICKENSCHER H, DEUTSCH U, GOULDING M, GRUSS P: **The Molecular Basis of the *undulated/Pax-1* Mutation.** *Cell* 1991, 66:873–884.

12. GOULDING MD, CHALEPAKIS G, DEUTSCH U, ERSELIUS JR, GRUSS P: **Pax-3, a Novel Murine DNA Binding Protein Expressed During Early Neurogenesis.** *EMBO J* 1991, 10:1135–1147.

13. WALTHER C, GRUSS P: ***Pax-6*, a Murine Paired Box Gene, is Expressed in the Developing CNS.** *Development* 1991, 113:1435–1450.

14. WALTHER C, GUENET J-L, SIMON D, DEUTSCH U, JOSTES B, GOULDING MD, PLACHOV D, BALLING R, GRUSS P: **Pax: A Murine Multigene Family of Paired Box-Containing Genes.** *Genomics* 1991, 11:424–434.

15. • ADAMS B, DÖRFLER P, AGUZZI A, KOZMIK Z, URBANEK P, MAURER–FOGY I, BUSSLINGER M: ***Pax-5* Encodes the Transcription Factor BSAP and is Expressed in B Lymphocytes, the Developing CNS, and Adult Testis.** *Genes Dev* 1992, 6:1589–1607.

The isolation of a *Pax* gene based entirely on the transcription factor activity of its product, BSAP, is described. It is shown that the paired-domain of BSAP is both necessary and sufficient for its binding *in vitro* to a DNA site shown in [32•] to be its target *in vivo*.

16. ASANO M, GRUSS P: ***Pax-5* is Expressed at the Midbrain-Hindbrain Boundary During Mouse Development.** *Mech Dev* 1992, 39:29–39.

17. TON CCT, HIRVONEN H, MIWA H, WEIL MM, MONAGHAN P, JORDAN T, VAN HEYNINGEN V, HASTIE ND, MEIJERS–HEIJBOER H, *ET AL.*: **Positional Cloning and Characterization of a Paired Box- and Homeobox-Containing Gene from the Aniridia Region.** *Cell* 1991, 67:1059–1074.

18. POLEEV A, FICKENSCHER H, MUNDLOS S, WINTERPACHT A, ZABEL B, FIDLER A, GRUSS P, PLACHOV D: ***PAX 8*, a Human Paired Box Gene: Isolation and Expression in Developing Thyroid, Kidney and Wilms' Tumors.** *Development* 1992, 116:611–623.

19. KRAUSS S, JOHANSEN T, KORZH V, MOENS U, ERICSON JU, FJOSE A: **Zebrafish *pax[zf-a]*: A Paired Box-Containing Gene Expressed in the Neural Tube.** *EMBO J* 1991, 10:3609–3619.

20. KRAUSS S, JOHANSEN T, KORZH V, FJOSE A: **Expression of the Zebrafish Paired Box Gene *pax[zf-b]* During Early Neurogenesis.** *Development* 1991, 113:1193–1206.

21. PÜSCHEL AW, GRUSS P, WESTERFIELD M: **Sequence and Expression Pattern of *pax-6* are Highly Conserved Between Zebrafish and Mice.** *Development* 1992, 114:643–651.

22. EPSTEIN DJ, VEKEMANS M, GROS P: ***splotch* (*Sp^{2H}*), a Mutation Affecting Development of the Mouse Neural Tube, Shows a Deletion within the Paired Homeodomain of *Pax-3*.** *Cell* 1991, 67:767–774.

23. • BALDWIN CT, HOTH CF, AMOS JA, DA-SILVA EO, MILUNSKY A: **An Exonic Mutation in the *HuP2* Paired Domain Gene Causes Waardenburg's Syndrome.** *Nature* 1992, 355:637–638.
See [24•].

24. • TASSABEHJI M, READ AP, NEWTON VE, HARRIS R, BALLING R, GRUSS P, STRACHAN T: **Waardenburg's Syndrome Patients Have Mutations in the Human Homologue of the *Pax-3* Paired Box Gene.** *Nature* 1992, 355:635–636.
This paper and [23•] demonstrate that mutant alleles of the human *PAX-3* (*HuP2*) gene are autosomal dominant and cause WSI. As their phenotype is similar to that of mouse *Pax-3* or *Sp* mutants [22], *Sp* mice may serve as a model system for the study of this human disease.

25. HILL RE, FAVOR J, HOGAN BLM, TON CCT, SAUNDERS GF, HANSON IM, PROSSER J, JORDAN T, HASTIE ND, VAN HEYNINGEN V: **Mouse *Small eye* Results from Mutations in a Paired-Like Homeobox-Containing Gene.** *Nature* 1991, 354:522–525.

26. • GRUSS P, WALTHER C: **Pax in Development.** *Cell* 1992, 69:719–722.
This review and [14,48] provide a comprehensive description of vertebrate *Pax* genes and discuss their developmental roles with an emphasis on mouse development.

27. TREISMAN J, HARRIS E, DESPLAN C: **The Paired Box Encodes a Second DNA-Binding Domain in the Paired Homeo Domain Protein.** *Genes Dev* 1991, 5:594–604.

28. BALLING R, DEUTSCH U, GRUSS P: ***undulated*, a Mutation Affecting the Development of the Mouse Skeleton, Has a Point Mutation in the Paired Box of *Pax 1*.** *Cell* 1988, 55:531–535.

29. FRASCH M, LEVINE M: **Complementary Patterns of *even-skipped* and *fushi tarazu* Expression Involve Their Differential Regulation by a Common Set of Segmentation Genes in *Drosophila*.** *Genes Dev* 1987, 1:981–995.

30. MORRISSEY D, ASKEW D, RAJ L, WEIR M: **Functional Dissection of the *paired* Segmentation Gene in *Drosophila* Embryos.** *Genes Dev* 1991, 5:1684–1696.

31. • ZANNINI M, FRANCIS-LANG H, PLACHOV D, DI LAURO R: **Pax-8, a Paired Domain-Containing Protein, Binds to a Sequence Overlapping the Recognition Site of a Homeodomain and Activates Transcription from Two Thyroid-Specific Promoters.** *Mol Cell Biol* 1992, 12:4230–4241.
It is suggested that the thyroid-specific expression of the TPO and Tg genes depends on their activation by Pax-8 protein. The evidence consists of a clear correlation of *in vitro* binding of Pax-8 protein or its paired-domain to promoter sequences of TPO (and Tg), with the ability of these sequences to activate a reporter gene in cotransfections with *Pax-8*.

32. • KOZMIK Z, WANG S, DÖRFLER P, ADAMS B, BUSSLINGER M: **The Promoter of the CD19 Gene is a Target for the B-Cell-Specific Transcription Factor BSAP.** *Mol Cell Biol* 1992, 12:2662–2672.
This paper identifies the CD19 gene, which encodes a cell surface protein involved in the control of proliferation and differentiation during B-cell maturation, and a defined sequence of its promoter as target of the *Pax-5* protein BSAP *in vivo*.

33. SMALL S, LEVINE M: **The Initiation of Pair-Rule Stripes in the *Drosophila* Blastoderm.** *Curr Opin Genet Dev* 1991, 1:255–260.

34. • INGHAM PW, MARTINEZ ARIAS A: **Boundaries and Fields in Early Embryos.** *Cell* 1992, 68:221–235.
This review discusses developmental fields, their polarities and boundaries in the context of the molecular mechanisms that determine segmentation in *Drosophila*.

35. BAUMGARTNER S, BOPP D, BURRI M, NOLL M: **Structure of Two Genes at the *gooseberry* Locus Related to the *paired* Gene and their Spatial Expression During *Drosophila* Embryogenesis.** *Genes Dev* 1987, 1:1247–1267.

36. • GUTJAHR T, FREI E, NOLL M: **Complex Regulation of Early *paired* Expression: Initial Activation by Gap Genes and Pattern Modulation by Pair-Rule Genes.** *Development* 1993, 117:609–623.
Initial changes in the establishment of *prd* expression in gap mutants as compared to wild-type embryos are explained by a model in which *prd* is initially activated by the products of gap genes rather than primary pair-rule genes.

37. KILCHHERR F, BAUMGARTNER S, BOPP D, FREI E, NOLL M: **Isolation of the *paired* Gene of *Drosophila* and its Spatial Expression During Early Embryogenesis.** *Nature* 1986, 321:493–499.

38. BAUMGARTNER S, NOLL M: **Network of Interactions Among Pair-Rule Genes Regulating *paired* Expression During Primordial Segmentation of *Drosophila*.** *Mech Dev* 1990, 33:1–18.

39. • LI X, GUTJAHR T, NOLL M: **Separable Regulatory Elements Mediate the Establishment and Maintenance of Cell States by the *Drosophila* Segment-Polarity Gene *gooseberry*.** *EMBO J* 1993, 12:1427–1436.
This work identifies the *cis*-regulatory elements responsible for the initial activation of *gsb* by pair-rule proteins, such as prd, and its maintenance by the wg signal.

40. • GUTJAHR T, PATEL NH, LI X, GOODMAN CS, NOLL M: **Analysis of the *gooseberry* Locus in *Drosophila* Embryos: *gooseberry* Determines the Cuticular Pattern and Activates *gooseberry neuro*.** *Development* 1993, 118:21–31.
It is shown by P-element-mediated transformation that one of the two *gooseberry* genes, *gsb*, is sufficient to rescue the cuticular phenotype of mutants in which both *gooseberry* genes are inactive. The other *gooseberry* gene, *gsbn*, is mainly expressed in the CNS, and its activation is shown to depend on the gooseberry protein.

41. • PEIFER M, BEJSOVEC A: **Knowing Your Neighbors: Cell Interactions Determine Intrasegmental Patterning in *Drosophila*.** *Trends Genet* 1992, 8:243–249.
This review discusses current knowledge and models of how segment-polarity genes establish and maintain positional information within segments.

42. PATEL NH, SCHAFER B, GOODMAN CS, HOLMGREN R: **The Role of Segment Polarity Genes During *Drosophila* Neurogenesis.** *Genes Dev* 1989, 3:890–904.

43. • DAMBLY-CHAUDIERE C, JAMET E, BURRI M, BOPP D, BASLER K, HAFEN E, DUMONT N, SPIELMANN P, GHYSEN A, NOLL M: **The Paired Box Gene *pox neuro*: A Determinant of Poly-Innervated Sense Organs in *Drosophila*.** *Cell* 1992, 69:159–172.
It is shown that *poxn* activity specifies p-es versus m-es organs in sensory mother cells, the stem cells of the PNS. As *poxn* specifies the fate of sensory mother cells in combination with other genes, such as *cut* [45•], it qualifies as a 'neuroblast identity' gene, as defined by Doe [46•], and may also determine the fate of neuroblasts in the CNS by a similar mechanism.

44. • NOTTEBOHM E, DAMBLY-CHAUDIERE C, GHYSEN A: **Connectivity of Chemosensory Neurons is Controlled by the Gene *poxn* in *Drosophila*.** *Nature* 1992, 359:829–832.
This paper demonstrates that ectopic *poxn* expression in early pupae transforms mono-innervated mechanosensory bristles into poly-innervated chemosensory bristles in adult legs. The underlying neurons have established the axonal connections and elicit the behavioural response characterisitic for chemosensory neurons, demonstrating that *poxn* also specifies their axonal pathways and connections.

45. • JAN YN, JAN LY: **Neuronal Specification.** *Curr Opin Genet Dev* 1992, 2:608–613.
This review discusses the genes and mechanisms involved in the specification of neuronal precursor cells in *Drosophila* embryos.

46. • DOE CQ: **The Generation of Neuronal Diversity in the *Drosophila* Embryonic Central Nervous System.** In *Determi-*

nants of Neuronal Identity. Edited by Shakland M, Macagno E. New York: Academic Press; 1992:119–154.
This review introduces the concept of 'neuroblast identity' genes that specify neuronal diversity in a combinatorial manner.

47. ROSENFELD MG: **POU-Domain Transcription Factors: Pou-erful Developmental Regulators.** *Genes Dev* 1991, **5**:897–907.

48. DEUTSCH U, GRUSS P: **Murine Paired Domain Proteins as Regulatory Factors of Embryonic Development.** *Semin Dev Biol* 1991, **2**:413–424.

49. NORNES HO, DRESSLER GR, KNAPIK EW, DEUTSCH U, GRUSS P: **Spatially and Temporally Restricted Expression of *Pax2* During Murine Neurogenesis.** *Development* 1990, **109**:797–809.

50. KRAUSS S, JOHANSEN T, KORZH V, FJOSE A: **Expression Pattern of Zebrafish *pax* Genes Suggests a Role in Early Brain Regionalization.** *Nature* 1991, **353**:267–270.

51. PÜSCHEL AW, WESTERFIELD M, DRESSLER GR: **Comparative Analysis of Pax-2 Protein Distributions During Neurulation in Mice and Zebrafish.** *Mech Dev* 1992, **38**:197–208.

52. • MCGINNIS W, KRUMLAUF R: **Homeobox Genes and Axial Patterning.** *Cell* 1992, **68**:283–302.
The functions of *Drosophila* and vertebrate homeobox gene clusters in providing positional information along the anteroposterior axis during embryonic development are comprehensively discussed.

53. • KRAUSS S, MADEN M, HOLDER N, WILSON SW: **Zebrafish *pax[b]* is Involved in the Formation of the Midbrain–Hindbrain Boundary.** *Nature* 1992, **360**:87–89.
Injection of anti-pax[zf-b] antiserum into fertilized zebrafish eggs is shown to down-regulate expression of *wnt-1*, *pax[zf-b]* and *eng-2* in the posterior midbrain, which subsequently fails to develop. As *prd* and *gsb* are required for proper expression of *wg* and *en* in *Drosophila* embryos, it is suggested that comparable networks of interacting factors exist in *Drosophila* and vertebrates.

54. MOLVEN A, NJØLSTAD PR, FJOSE A: **Genomic Structure and Restricted Neural Expression of the Zebrafish *wnt-1* (*int-1*) Gene.** *EMBO J* 1991, **10**:799–807.

55. PATEL NH, MARTIN-BLANCO E, COLEMAN KG, POOLE SJ, ELLIS MC, KORNBERG TB, GOODMAN CS: **Expression of *engrailed* Proteins in Arthropods, Annelids, and Chordates.** *Cell* 1989, **58**:955–968.

56. WILKINSON DG, BAILES JA, MCMAHON AP: **Expression of the Proto-Oncogene *int-1* is Restricted to Specific Neural Cells in the Developing Mouse Embryo.** *Cell* 1987, **50**:79–88.

57. DAVIS CA, JOYNER AL: **Expression Patterns of the Homeo Box-Containing Genes *En-1* and *En-2* and the Proto-Oncogene *int-1* Diverge During Mouse Development.** *Genes Dev* 1988, **2**:1736–1744.

58. MCMAHON AP, BRADLEY A: **The *Wnt-1* (*int-1*) Proto-Oncogene is Required for the Development of a Large Region of the Mouse Brain.** *Cell* 1990, **62**:1073–1085.

59. THOMAS KR, CAPECCHI MR: **Targeted Disruption of the Murine *int-1* Proto-Oncogene Resulting in Severe Abnormalities in Midbrain and Cerebellar Development.** *Nature* 1990, **346**:847–850.

60. THOMAS KR, MUSCI TS, NEUMANN PE, CAPECCHI MR: ***Swaying* is a Mutant Allele of the Proto-Oncogene *Wnt-1*.** *Cell* 1991, **67**:969–976.

61. • MCMAHON AP, JOYNER AL, BRADLEY A, MCMAHON JA: **The Midbrain-Hindbrain Phenotype of *Wnt-1*$^-$/*Wnt-1*$^-$ Mice Results from Stepwise Deletion of *engrailed*-Expressing Cells by 9.5 Days Postcoitum.** *Cell* 1992, **69**:581–595.
Homozygous *Wnt-1*$^-$ mice fail to express *En-1* (and *En-2*) in the entire midbrain–hindbrain domain and subsequently are unable to form the midbrain and adjacent cerebellum. These results and those of [53•] suggest that the regulatory interactions between *Wnt-1* and *En* in vertebrates are similar to those of their homologs in *Drosophila*.

62. CHO KWY, BLUMBERG B, STEINBEISSER H, DE ROBERTIS EM: **Molecular Nature of Spemann's Organizer: The Role of the *Xenopus* Homeobox Gene *goosecoid*.** *Cell* 1991, **67**:1111–1120.

63. SOKOL S, CHRISTIAN JL, MOON RT, MELTON DA: **Injected *Wnt* RNA Induces a Complete Body Axis in *Xenopus* Embryos.** *Cell* 1991, **67**:741–752.

64. SMITH WC, HARLAND RM: **Injected *Xwnt-8* RNA Acts Early in *Xenopus* Embryos to Promote Formation of a Vegetal Dorsalizing Center.** *Cell* 1991, **67**:753–765.

65. • VAN HEYNINGEN V, HASTIE ND: **Wilms' Tumour: Reconciling Genetics and Biology.** *Trends Genet* 1992, **8**:16–21.
A review on the genetic and developmental aspects of Wilms' tumour.

66. • DRESSLER GR, WILKINSON JE, ROTHENPIELER UW, PATTERSON LT, WILLIAMS-SIMONS L, WESTPHAL H: **Deregulation of *Pax-2* Expression in Transgenic Mice Generates Severe Kidney Abnormalities.** *Nature* 1993, **362**:65–67.
Ectopic expression of *Pax-2* in non-terminally differentiated renal epithelial cells of CMV–*Pax-2* transgenic mouse embryos leads to abnormal kidney development, probably by restricting the differentiation potential of these cells.

67. DRESSLER GR, DOUGLASS EC: ***Pax-2* is a DNA-Binding Protein Expressed in Embryonic Kidney and Wilms' Tumor.** *Proc Natl Acad Sci USA* 1992, **89**:1179–1183.

68. • GOULDING MD, LUMSDEN A, GRUSS P: **Signals from the Notochord and Floor Plate Regulate the Region-Specific Expression of Two *Pax* Genes in the Developing Spinal Cord.** *Development* 1993, **117**:1001–1016.
Removal of the notochord or implantation of a supernumerary notochord or floor plate elicit dramatic changes in the dorsoventral expression patterns of *Pax-3* and *Pax-6* in the neural tube of chick embryos. The observation that these changes in expression patterns occur rapidly in response to the implanted notochord or floor plate raises the possibility that these genes might mediate the subsequent changes in cell differentiation [69].

69. YAMADA T, PLACZEK M, TANAKA H, DODD J, JESSELL TM: **Control of Cell Patterning in the Developing Nervous System: Polarizing Activity of the Floor Plate and Notochord.** *Cell* 1991, **64**:635–647.

70. ALLEN JD, LINTS T, JENKINS NA, COPELAND NG, STRASSER A, HARVEY RP, ADAMS JM: **Novel Murine Homeo Box Gene on Chromosome 1 Expressed in Specific Hematopoietic Lineages and During Embryogenesis.** *Genes Dev* 1991, **5**:509–520.

M Noll, Institute for Molecular Biology II, University of Zürich, Winterthurerstr. 190, CH-8057 Zürich, Switzerland.