The EMBO Journal vol.15 no.14 pp.3722–3731, 1996

The functional conservation of proteins in evolutionary alleles and the dominant role of enhancers in evolution

Lei Xue and Markus Noll[1]

Institute for Molecular Biology (II), University of Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland

[1]Corresponding author

***Drosophila paired*⁻ embryos can be rescued to viable adults by the evolutionary alleles *prd*–Gsb and *prd*–Pax3, which express the *Drosophila* Gooseberry and mouse Pax3 proteins under the control of the *paired cis*-regulatory region. As *prd*–Gsb uncovers a *prd* function involved in the proper abdominal segmentation of adults, evolutionary alleles, defined and constructed in this manner, may often be weak and thus serve to discover hitherto unknown functions of a gene. Our findings show that the Gooseberry and Pax3 proteins have conserved most or all functions of the related *Drosophila* Paired protein although their C-terminal halves appear unrelated in sequence but not in 3-D structure essential for function. It follows that the acquisition of new *cis*-regulatory regions rather than the divergence of the C-terminal coding regions has been the primary device for the functional diversification of the *Drosophila* genes *paired* and *gooseberry* and the mouse *Pax3* gene. The operation of this mechanism in insects as well as vertebrates suggests a major role in evolution.**

Keywords: *Drosophila*/enhancer/evolutionary alleles/*Pax* genes/protein function

Introduction

About 50 genes, the maternal coordinate genes, three classes of zygotic segmentation genes—the gap, pair-rule and segment-polarity genes—and the homeotic genes, interact in a hierarchical network to progressively define position along the antero-posterior axis with increasing precision in the developing *Drosophila* embryo (Nüsslein-Volhard and Wieschaus, 1980; Small and Levine, 1991; Peifer and Bejsovec, 1992; St Johnston and Nüsslein-Volhard, 1992). We have isolated and characterized three genes of this network that have arisen by independent duplications during evolution, the pair-rule gene *paired* (*prd*) and the segment-polarity genes *gooseberry* (*gsb*) and *gooseberry neuro* (*gsbn*) (Noll, 1993). The N-terminal halves of the transcription factors encoded by these genes share a paired-domain and a homeodomain, whose amino acid sequences have been highly conserved (~75% identity and 85% homology) (Bopp *et al.*, 1986), whereas their C-terminal portions exhibit no significant amino acid sequence similarity (Baumgartner *et al.*, 1987). Together with the other pair-rule genes, *prd* specifies position along the antero-posterior axis with a double-segment periodicity and activates several segment-polarity genes, including *gsb*, *wingless* (*wg*) and *engrailed* (*en*), which are expressed at a single-segment periodicity. After germ-band extension, *gsb* maintains the epidermal expression of *wg*, which represses cuticular denticle formation, through a *wg*–*gsb* autoregulatory loop (Li and Noll, 1993). In addition, *gsb* activates *gsbn* in a segmentally repeated pattern in the central nervous system (CNS) (Gutjahr *et al.*, 1993b), implying that *gsbn* might be required for proper segmentation of the CNS (Patel *et al.*, 1989).

Although these genes perform distinct developmental functions, their ubiquitous expression generates the same cuticular phenotypes (Li and Noll, 1994). Moreover, a *gsb*–Prd transgene, which expresses the Prd protein under the control of the *gsb cis*-regulatory region in the epidermis (note that proteins are abbreviated in roman type and initial capitals, genes in lower case italics), is able to rescue the cuticular phenotype of *gsb*⁻ embryos (Li and Noll, 1994). These surprising results suggest that the three related proteins Prd, Gsb and Gsbn might have retained the same functions despite their highly diverged C-terminal domains and point to the *cis*-regulatory rather than coding regions as being responsible for the essential difference between the distinct functions of the three genes. However, these experiments left unanswered the question of whether the Prd protein is able to substitute for all functions or only for the embryonic epidermal function of the Gsb protein because (i) expression of Prd was controlled by an incomplete *cis*-regulatory region of *gsb* and thus limited to the embryonic epidermis, and (ii) the only available *gsb*⁻ embryos were homozygous for the deficiency *Df(2R)IIX62*, which deletes additional genes that are required for survival beyond embryogenesis (Côté *et al.*, 1987).

Here, we have answered this question of the functional equivalence of Prd and Gsb. We show that the Gsb protein, when expressed under the control of the complete *prd cis*-regulatory region, is able to rescue *prd*⁻ embryos to viable adults, thus demonstrating that Gsb has retained all functions of Prd required for survival to the adult stage. In addition, we have extended these studies to include the mouse Pax3 or Splotch protein (Epstein *et al.*, 1991), a mouse homolog of Prd, Gsb and Gsbn and of the human Pax3 protein (Burri *et al.*, 1989) responsible for Waardenburg's syndrome type I (Baldwin *et al.*, 1992; Tassabehji *et al.*, 1992). The Pax3 protein again consists of the two highly conserved N-terminal domains and an apparently unrelated C-terminal half (Goulding *et al.*, 1991). We find that Pax3 protein expressed under the control of the *prd cis*-regulatory region is able to rescue the cuticular phenotype of *prd*⁻ embryos and thus is capable of substituting for the main functions of Prd protein, but that it is unable to support their development to adults. We conclude that the different functions of *prd*, *gsb* and *Pax3* are largely determined by their different

control regions and are modified by protein domains which are probably located within the C-terminal portion. The functional equivalence of the exchanged protein domains is striking and emphasizes the high degree of functional conservation of protein domains after gene duplication. An important general implication of this functional conservation of the proteins, which includes the largely unknown functions of their highly diverged C-terminal halves, is the conservation of their secondary and tertiary structures that is not necessarily reflected by their amino acid sequences. The validity of per cent amino acid identity as a measure of functional equivalence in homologous proteins has been questioned before (Yockey, 1992). Together these observations add further weight to our earlier suggestion that the acquisition of different *cis*-regulatory regions is the primary mechanism in the evolution of new functions (Li and Noll, 1994). Finally, we show that the construction of 'evolutionary' *prd* alleles has led to the discovery of *prd* functions not revealed by previously known *prd* alleles.

Results

Gsb and Pax3 proteins recognize Prd target site in vitro

Previous studies suggest that the establishment of *gsb* expression depends on the *gsb* early enhancer (GEE) and its activation by Prd protein (Li *et al.*, 1993). To test whether Prd is able to bind to GEE *in vitro*, a 35 bp fragment, GEE1, that exhibited a strong Prd footprint (X.Li and M.Noll, unpublished) was derived from GEE and used as probe in band-shift assays. As is evident from Figure 1A, partially purified Prd protein synthesized in bacteria forms a strong complex with GEE1 (lanes 2 and 3), whereas no band shift was detectable after its incubation with bacterial extracts devoid of Prd (lane 1). Prd binds specifically to GEE1 since the shifted band is not affected by the addition of an excess of Sp1 binding sites (lane 5) but is abolished by competition with unlabeled GEE1 (lane 4). Moreover, the formation of this complex is inhibited by Prd antiserum (lanes 6 and 7) but not by preimmune serum (lane 8), confirming its dependence on Prd protein. Similar to Prd, both Gsb (Figure 1B, lanes 1–7) and Pax3 proteins (Figure 1B, lanes 8–11) form a specific complex with GEE1, indicating that they are able to recognize *in vitro* the same target site as Prd.

Generation of prd–Gsb and prd–Pax3 transgenic flies

To examine whether and to what extent the Gsb or Pax3 proteins are able to substitute for the functions of the Prd protein, two DNA constructs, *prd*–Gsb and *prd*–Pax3, consisting of the *gsb* or *Pax3* coding region under the control of the complete *prd cis*-regulatory region (Figure 2A; Gutjahr *et al.*, 1994) were introduced into the *Drosophila* germline by P-element-mediated transformation. Three independent transgenic lines were obtained for *prd*–Gsb, whose transgenes, *prd*–Gsb-1, *prd*–Gsb-2 and *prd*–Gsb-3, are all located on the third chromosome, and one for *prd*–Pax3, whose transgene *prd*–Pax3-1 resides on the X-chromosome.

We have shown previously (Gutjahr *et al.*, 1993a) that in wild-type embryos Prd protein appears in a broad

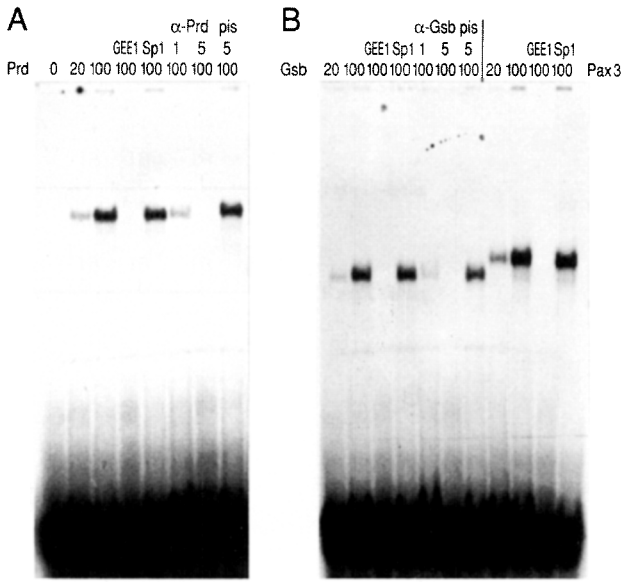

Fig. 1. Prd, Gsb and Pax3 proteins bind specifically to the Prd binding site GEE1. (**A**) Band-shift analysis of the oligonucleotide GEE1 with purified Prd protein. Band-shift assays of labeled GEE1 were performed with 1 μg protein from crude extracts of IPTG-induced *E.coli* BL21(DE3) cells carrying the empty expression vector pAR3040 (lane 1), and with 20 ng (lane 2) or 100 ng (lanes 3–8) of purified Prd protein expressed in the same cells. In the assays shown in lanes 4 and 5, a 400-fold molar excess of unlabeled GEE1 and Sp1 binding sites, respectively, was added to the labeled GEE1 prior to addition of Prd protein. In lanes 6–8, Prd protein was preincubated with 1 μl (lane 6) or 5 μl (lane 7) of rabbit anti-Prd antiserum (Gutjahr *et al.*, 1993a), or with 5 μl of preimmune serum (lane 8) before binding buffer was added. (**B**) Band-shift analysis of GEE1 with purified Gsb or Pax3 protein. Band-shift assays of labeled GEE1 were carried out with 20 ng (lanes 1 and 8) or 100 ng (remaining lanes) of purified Gsb (lanes 1–7) or Pax3 protein (lanes 8–11) expressed in *E.coli* BL21(DE3). In the assays shown in lanes 3, 4, 10 and 11, a 400-fold molar excess of unlabeled GEE1 (lanes 3 and 10) or Sp1 binding sites (lanes 4 and 11) was added to the labeled GEE1 prior to addition of Gsb or Pax3 protein. In lanes 5–7, Gsb protein was preincubated with 1 μl (lane 5) or 5 μl (lane 6) of rabbit anti-Gsb antiserum (Gutjahr *et al.*, 1993b), or with 5 μl of preimmune serum (lane 7) before binding buffer was added.

anterior stripe at the end of the last nuclear division (Figure 2B). By mid-cellularization, Prd protein is expressed in seven stripes with a double-segment periodicity characteristic of pair-rule genes (Figure 2C). Subsequently, Prd expression is converted into a pattern of 14 stripes with single-segment periodicity and an anterior dorsal patch (Figure 2D). As expected, under the control of the *prd cis*-regulatory region, the expression patterns of the transgenic Gsb protein (Figure 2E–G) and *Pax3* RNA (Figure 2H–J) are indistinguishable from those of the endogenous Prd protein (Figure 2B–D) although the three *prd*–Gsb lines express Gsb at slightly different levels (data not shown, but see below).

Rescue by prd–Gsb and prd–Pax3 of prd target gene expression

The early expression patterns of several segment-polarity genes, such as those of *gsb*, *wingless* (*wg*) and *engrailed* (*en*), are positively regulated by Prd (Bopp *et al.*, 1989; Ingham and Martinez Arias, 1992), as evident from the loss of their odd- or even-numbered stripes in *prd* mutant embryos (Figure 3D–F) as compared with wild type

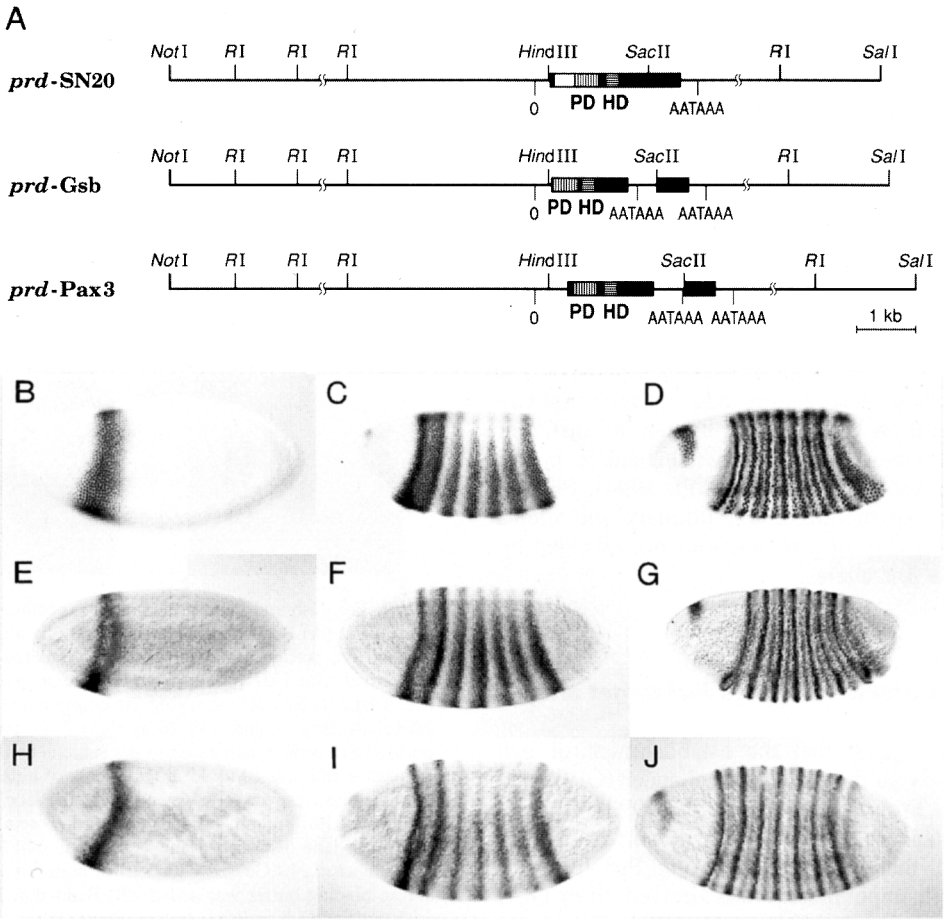

Fig. 2. Expression of Prd, Gsb and Pax3 proteins under the control of the complete *prd cis*-regulatory region. (**A**) Map of *prd*–SN20, *prd*–Gsb and *prd*–Pax3 transgenes. The *prd*–SN20 transgene is an 18.7 kb genomic fragment that includes the entire *prd* gene, consisting of its transcribed region as well as of adjacent 10.0 kb upstream and 5.9 kb downstream sequences. As a single copy, it is able to rescue prd^- embryos to wild-type adults (Gutjahr *et al.*, 1994). In *prd*–Gsb and *prd*–Pax3 transgenes, most of the *prd* coding region and the *prd* intron are replaced by *gsb*- and *Pax3* cDNAs, respectively, while the entire *prd cis*-regulatory region is retained. Coding regions are indicated as black boxes except for the paired-domains PD and the *prd*-type homeodomains HD which are hatched. The *prd* intron is indicated as an open box. Transcriptional starts are marked by 0 and poly(A) addition signals AATAAA are indicated. (**B–J**) Expression of Prd protein in wild-type embryos (B–D), of Gsb protein in homozygous transgenic *prd*–Gsb-2 embryos (E–G) and of *Pax3* mRNA in homozygous transgenic *prd*–Pax3-1 embryos (H–J) at the onset of cellularization (B, E and H), at mid-cellularization of syncytial blastoderm (C, F and I) and at the onset of gastrulation (D, G and J). Wild-type embryos were stained with anti-Prd antiserum and transgenic embryos, collected from homozygous *prd*–Gsb-2 or *prd*–Pax3-1 stocks, were stained with anti-Gsb antiserum or hybridized *in situ* with digoxigenin-labeled *Pax3* cDNA. Embryos are oriented with their anterior to the left and dorsal side up.

(Figure 3A–C). The *prd* mutant embryos used in these and all subsequent experiments were either homozygous or hemizygous, when combined with the deficiency *Df(2L)Prl*, for the $prd^{2.45}$ allele, which has a 1.1 kb insertion (Kilchherrr *et al.*, 1986) after amino acid 45 of the paired-domain (Frigerio *et al.*, 1986), and hence lacked any functional Prd protein. By the introduction of *prd*–Gsb or *prd*–Pax3 into such prd^- embryos, *gsb*, *wg* and *en* expression patterns are rescued partially by one copy (Figure 3G–I) and completely by two copies of the transgenes (Figure 3J–L), demonstrating that Gsb and Pax3 can substitute for Prd protein in the transcriptional regulation of these target genes.

Rescue by prd–Gsb and prd–Pax3 of the prd^- cuticular phenotype

Embryos that lack *prd* function exhibit the classical pair-rule phenotype in their cuticle (Nüsslein-Volhard and Wieschaus, 1980), in which segmental equivalents are deleted with a double-segment periodicity (Figure 4B) as compared with wild-type embryos (Figure 4A). This phenotype results from the failure to express several segment-polarity genes, like *gsb*, *wg* and *en*, on either side of every other parasegmental boundary (Ingham and Martinez Arias, 1992). Since *prd*–Gsb and *prd*–Pax3 transgenes are able to rescue the expression of these segment-polarity genes in *prd* mutants (Figure 3), they might also rescue the cuticular pair-rule phenotype. Indeed, *prd*–Gsb derived from all three independent lines and *prd*–Pax3 rescue the cuticular phenotype of prd^- embryos partially with one copy (Figure 4C and D) and completely with two copies (cuticles indistinguishable from those of wild-type embryos, cf. Figure 4A), demonstrating that Gsb and Pax3 are able to replace the function of Prd protein in the epidermis.

Two copies of prd–Gsb can rescue prd^- embryos to adult flies

When *prd*–Pax3 and *prd*–Gsb transgenes were tested for their ability to rescue prd^- embryos to adulthood, one

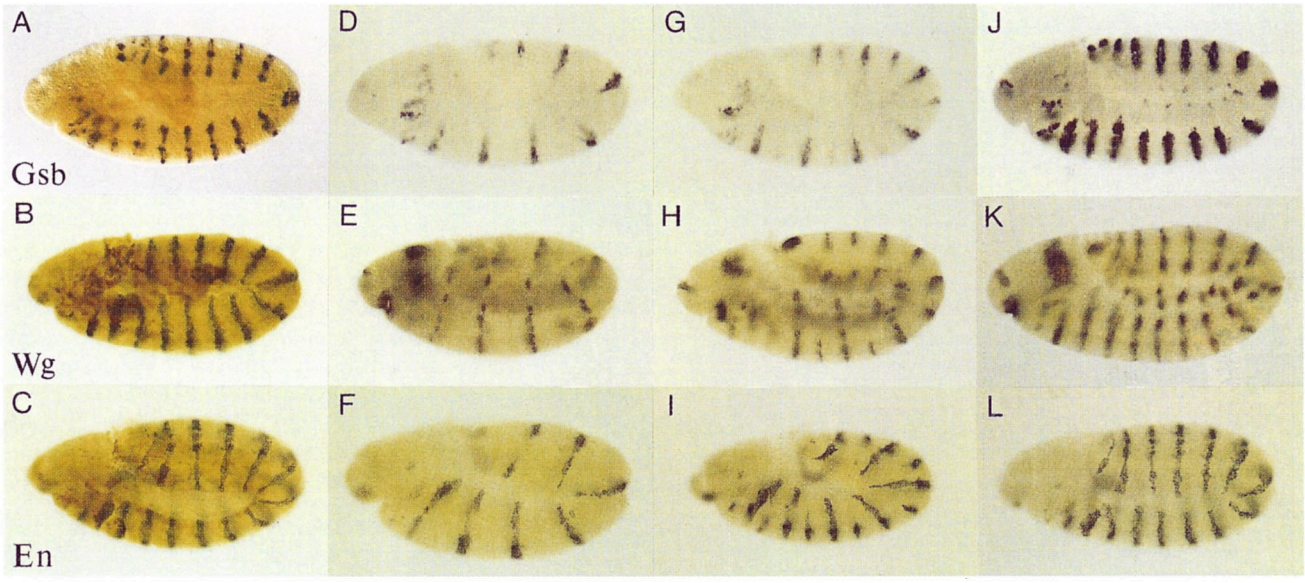

Fig. 3. Rescue of *gsb*, *wg* and *en* expression patterns in *prd*$^-$ embryos by *prd*–Gsb and *prd*–Pax3 transgenes. Expression of Gsb (**A**, **D**, **G** and **J**), Wg (**B**, **E**, **H** and **K**) and En protein (**C**, **F**, **I** and **L**) in wild-type (A–C), and homozygous *prd*$^{2.45}$ embryos carrying no (D–F), one (G–I) or two copies (J–L) of a *prd*–Pax3 transgene. Identical expression patterns were observed in homozygous *prd*$^{2.45}$ embryos carrying one or two copies of a *prd*–Gsb instead of the *prd*–Pax3 transgene (not shown). Embryos during the extended germ band stage are oriented with their anterior to the left and dorsal side up. Embryos were collected from *prd*–Pax3-1/X or Y; *prd*$^{2.45}$/*CyO*, *actin*–LacZ parents, whose *CyO* balancer chromosome expresses β-galactosidase under the control of the actin promoter, and double stained for β-galactosidase (brown) and Gsb, Wg or En protein (black), using monoclonal antibodies or rabbit antiserum against β-galactosidase and rabbit anti-Gsb or anti-Wg antiserum or anti-En monoclonal antibodies. One quarter of the embryos did not stain for β-galactosidase as expected. Of these embryos, one quarter express Gsb, Wg and En in a pair-rule pattern as expected for *prd*$^-$ embryos, half of the embryos display partially rescued and another quarter fully rescued expression patterns as in wild-type embryos.

copy was insufficient. However, all combinations of two *prd*–Gsb transgenes were able to rescue some of the *prd*$^-$ embryos to viable adults, except for two copies of *prd*–Gsb-1, which cannot be tested because its insertion is lethal, and two copies of *prd*–Gsb-3 (Table I). These results are explained by independent tests indicating that transgene expression is highest in the *prd*–Gsb-1 and lowest in the *prd*–Gsb-3 line. Hence, in order to substitute Gsb for Prd protein and rescue *prd*$^-$ embryos to adulthood, Gsb must be expressed above a threshold level which is higher than that for wild-type Prd protein. For one copy of a *prd* transgene with [*prd*–SN20 (Gutjahr *et al.*, 1994); Figure 2A] or without (*prd*–Prd) the *prd* intron is sufficient to rescue all *prd*$^-$ embryos to adulthood. Therefore, Gsb protein is able, although less efficiently, to substitute for all functions of Prd required for survival to the adult stage, as already shown for Gsb and Pax3 proteins in regard to the larval cuticular phenotype (Figure 4C and D).

One copy of prd–Pax3 enhances rescue to adulthood by prd–Gsb

In contrast to *prd*–Gsb, two copies of *prd*–Pax3-1 fail to rescue *prd*$^-$ embryos to adults (Table I) although they express *Pax3* mRNA at levels comparable with those of *prd*–Gsb-1 and suffice to rescue the *prd* cuticular phenotype. Neither did doubling the *prd*–Pax3 dosage make a difference as is evident from tests with one or two copies of six additional *prd*–Pax3 transgenic lines even though some of these lines express higher levels of *Pax3* mRNA than *prd*–Pax3-1 (data not shown). We conclude that *prd*–Pax3-1 fails to rescue *prd*$^-$ embryos to adulthood not because of insufficient levels of its expression but rather because the Pax3 protein is unable to substitute for all of the functions of Prd required to support development to the adult stage.

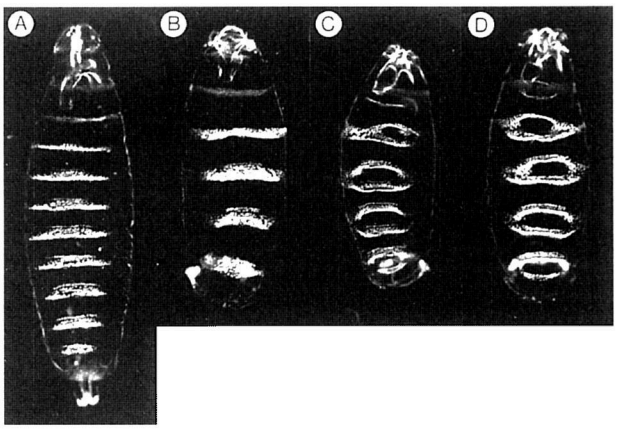

Fig. 4. Rescue of the cuticular pair-rule phenotype of *prd*$^-$ embryos by *prd*–Gsb or *prd*–Pax3 transgenes. Ventral views of cuticle preparations of wild-type (**A**) and *prd*$^{2.45}$/*prd*$^{2.45}$ embryos without (**B**) and with one copy of the *prd*–Gsb (**C**) or *prd*–Pax3 transgene (**D**) are shown under dark-field illumination (anterior is up). Embryos carrying two copies of the transgenes exhibit cuticles that are indistinguishable from wild type (not shown). To obtain *prd*$^-$ embryos carrying one or two copies of *prd*–Gsb, *prd*$^{2.45}$/*SM1*; *prd*–Gsb-2/*prd*–Gsb-2 flies were crossed with *prd*$^{2.45}$/*SM1* flies or *inter se*. To produce *prd*$^-$ embryos carrying one or two copies of *prd*–Pax3, *prd*–Pax3-1/*prd*–Pax3-1; *prd*$^{2.45}$/*SM1* virgins were crossed with *prd*$^{2.45}$/*SM1* or *prd*$^{2.45}$ *prd*–Pax3-2/*SM1* males. Embryos were collected and allowed to develop for 24 h at 25°C before cuticles were prepared as described (Wieschaus and Nüsslein-Volhard, 1986).

Surprisingly, combinations of one copy each of *prd*–Pax3-1 and *prd*–Gsb transgenes result in much higher rescue efficiencies than any combination of two *prd*–Gsb

Table I. Rescue of *prd*⁻ embryos to viable adults by two copies of *prd*–Gsb or *prd*–Pax3

	prd–Gsb-1	*prd*–Gsb-2 (%)	*prd*–Gsb-3 (%)	*prd*–Pax3-1 (%)
prd–Gsb-1	lethal	15 (40/260)	5 (6/123)	51 (62/121)
prd–Gsb-2		9 (50/576)	4 (33/779)	29 (40/139)
prd–Gsb-3			0 (0/987)	18 (11/61)
prd–Pax3-1				0 (0/885)

Percentages of rescued $prd^{2.45}$/*Df(2L)Prl* flies harboring combinations of two *prd*–Gsb or *prd*–Pax3-1 transgenes (actual numbers of rescued flies per total number of expected *prd* mutants are given in parentheses). Transgenic *prd*⁻ flies carrying two copies of *prd*–Gsb and/or *prd*–Pax3 were obtained as offspring from crosses between $prd^{2.45}$/*SM1* and *Df(2L)Prl/SM1* flies that were homozygous for *prd*–Gsb-2 or *prd*-Gsb-3 on the third chromosome, homozygous for *prd*–Pax3-1 on the first chromosome, or carried *prd*–Gsb-1, whose insertion is homozygous lethal, on one of the third chromosomes.

Table II. Rescue of *prd*⁻ embryos to viable adults by combinations of *prd*–Gsb-2 with *prd*–Pax3

prd–Pax3 line	Chromosome[a]	Rescued adults[b] (%)
prd–Pax3-1	X	29 (40/139)
prd–Pax3-2	2	23 (64/280)
prd–Pax3-3	2	17 (68/412)
prd–Pax3-4	2	7 (40/560)
prd–Pax3-5	3	29 (10/34)
prd–Pax3-6	3	30 (83/281)
prd–Pax3-7	3	11 (6/53)

[a]Chromosome carrying the *prd*–Pax3 transgene.
[b]Percentages of rescued $prd^{2.45}$/*Df(2L)Prl* flies harboring one copy each of *prd*–Pax3 and *prd*–Gsb-2 transgene (actual numbers of rescued flies per total number of expected *prd* mutants are given in parentheses).
The *prd*–Pax3 transgenes located on the second chromosome were recombined with $prd^{2.45}$ (*prd*–Pax3-2) or *Df(2L)Prl* (*prd*–Pax3-3 and *prd*–Pax3-4) and maintained as stocks over the *SM1* balancer chromosome while the other *prd*–Pax3 transgenes were kept as homozygous stocks in a transheterozygous *prd*⁻ background [$prd^{2.45}$ or *Df(2L)Prl/SM1*]. Transgenic $prd^{2.45}$/*Df(2L)Prl* offspring carrying one copy each of *prd*–Gsb-2 and *prd*–Pax3 were obtained from crosses of these with $prd^{2.45}$ or *Df(2L)Prl/SM1* stocks homozygous for *prd*–Gsb-2.

transgenes (Table I). For example, replacing *prd*–Gsb-2 in combinations with itself or the other two *prd*–Gsb transgenes by *prd*–Pax3-1 produces 3- to 4-fold higher rescue efficiencies. In general, if the expression of the *prd*–Gsb transgene that is replaced by *prd*–Pax3-1 is low, the enhancement of the rescue is high and vice versa. Since the replacement by *prd*–Pax3-1 always results in an enhanced rescue, it follows that Pax3 is better than Gsb protein in substituting for at least one of the functions of Prd protein required to support development to adulthood. Moreover, these results confirm our conclusion that the failure of two copies of *prd*–Pax3-1 to rescue *prd*⁻ embryos to the adult stage does not result from insufficient levels of Pax3 expression.

As is evident from Table II, all of six new *prd*–Pax3 transgenic lines, generated by mobilizing the P element carrying *prd*–Pax3-1, rescue *prd*⁻ embryos to adults when combined with one copy of *prd*–Gsb-2 with an efficiency that reaches a plateau at ~30% for several *prd*–Pax3 lines. Hence, in these lines, which include the original *prd*–Pax3-1 line, the level of Pax3 expression is optimal and not limiting. It follows that their failure to rescue *prd*⁻ embryos to adults in the absence of a *prd*–Gsb transgene results from the inability of the Pax3 protein to substitute for all functions of the Prd protein. Taken together, these experiments show that Gsb protein is able to perform all of the functions of Prd required for the rescue of *prd*⁻ embryos to adults, whereas Pax3 can substitute for some but not all of these functions. However, Pax3 is a better replacement than Gsb for at least one Prd function.

The rescued flies show, in addition to the wild-type phenotype (Figure 5A and D), a segmental phenotype exhibiting a variety of disorganized segments (Figure 5B and C). The yield of wild-type phenotypes among the rescued adults increases with the rescue ability of the different combinations of the two transgenes (data not shown, but see below).

Rescue of prd⁻ embryos depends on relative levels of Gsb and Pax3 protein

Since two copies of *prd*–Gsb or *prd*–Pax3 transgenes are able to rescue at best about half of the *prd*⁻ embryos to the adult stage (Tables I and II), additional copies of these transgenes might further increase the rescue efficiency and/or the fraction of wild-type phenotypes among the rescued adults. The results of such experiments, in which up to two copies each of *prd*–Gsb and *prd*–Pax3 transgenes were tested, are summarized in Table III. In agreement with our observation that Pax3 protein enhances the rescue activity of *prd*–Gsb (Tables I and II), replacing one of the two *prd*–Gsb-2 transgenes by *prd*–Pax3-2 or *prd*–Pax3-3 increases the rescue efficiency considerably while the percentage of wild-type phenotypes among the rescued adults is augmented 6- to 11-fold (Table III). If an additional copy of *prd*–Gsb-2 is added to embryos carrying one copy each of *prd*–Gsb-2 and *prd*–Pax3, the rescue efficiency and the fraction of wild-type phenotypes among the rescued adults are further enhanced several fold, indicating that the level of *prd*–Gsb-2 expression is limiting when only one copy is present. In contrast, if an additional copy of *prd*–Pax3 instead of *prd*–Gsb is added, only the fraction of normal rescued adults increases, whereas the rescue efficiency is reduced (Table III). Thus, Pax3 protein levels are too high for optimal rescue if two doses of a *prd*–Pax3 transgene and only one copy of *prd*–Gsb are present, demonstrating that the rescue efficiency depends not only on absolute but also on relative levels of Gsb and Pax3 proteins.

In other words, *prd*–Pax3 exerts a dominant negative effect if expressed at too high a level. This interpretation is confirmed by the finding that the rescue efficiency achieved by two copies of *prd*–Gsb-2 and one copy of *prd*–Pax3-2 is reduced 5-fold if one *prd*–Gsb-2 transgene is replaced by *prd*–Pax3-3, whereas the dominant negative effect exerted by the addition of a second *prd*–Pax3 transgene is largely buffered if two copies of *prd*–Gsb-2 remain present (Table III). Interestingly, the fraction of rescued adults displaying a wild-type phenotype is not subject to this dominant negative effect of *prd*–Pax3. On the contrary, doubling the dose of *prd*–Pax3 enhances this fraction to nearly wild-type levels (Table III). Therefore, the dominant negative effect exerted by Pax3 protein affects one or several functions of Prd required to support

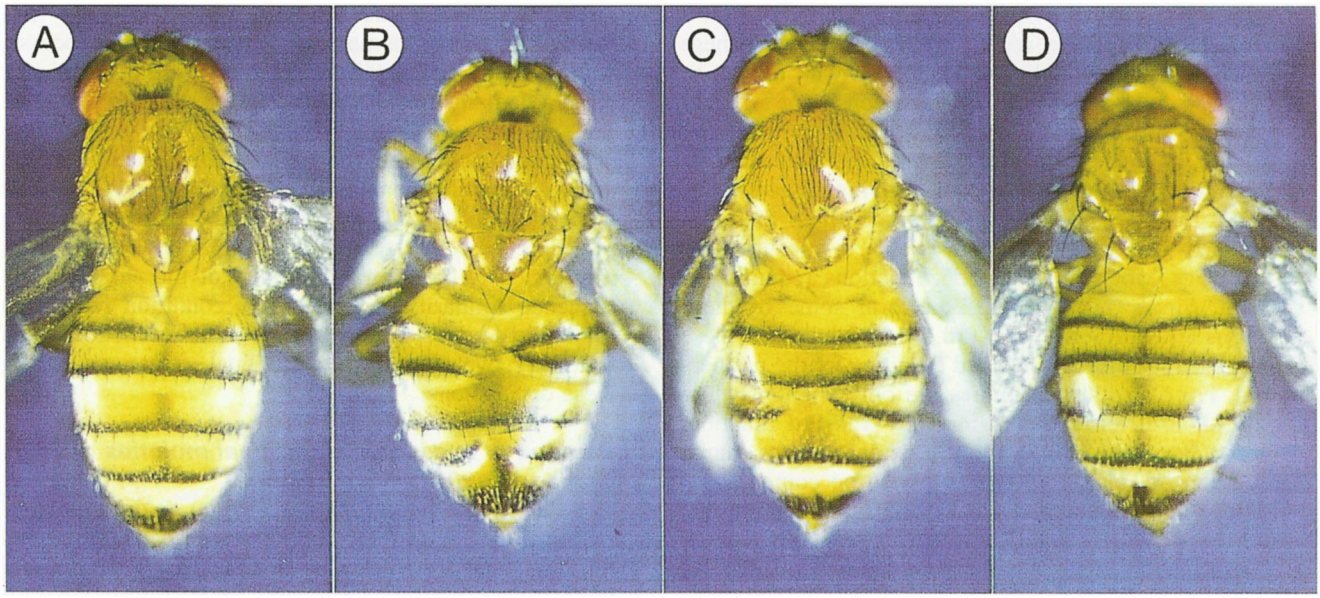

Fig. 5. Rescue of *prd*$^-$ embryos to adult flies by two copies of *prd*–Gsb or one copy each of *prd*–Gsb and *prd*–Pax3 transgenes; the adult segmental phenotype of *prd*. (**A**) Wild-type adult females. (**B–D**) *prd*$^{2.45}$/*Df(2L)Prl* adult females, rescued by one copy each of *prd*–Gsb-2 and *prd*–Pax3-1, displaying either a distorted (B and C) or wild-type (D) abdominal phenotype. Similar phenotypes are observed among *prd*$^-$ flies rescued by two copies of the *prd*–Gsb transgene. Rescued *prd*$^-$ flies were obtained from a cross of *prd*–Pax3-1/*prd*–Pax3-1; *prd*$^{2.45}$/*SM1* females with *Df(2L)Prl*/*SM1*; *prd*–Gsb-2/*prd*–Gsb-2 males.

Table III. Rescue of lethality and of adult segmental phenotype of *prd*$^-$ embryos by combinations of *prd*–Gsb-2 with *prd*–Pax3-2 and/or *prd*–Pax3-3 transgenes

	prd–Gsb-2		*prd*–Gsb-2/*prd*–Gsb-2	
	Rescued adults (%)[a]	Wild type (%)[b]	Rescued adults (%)[a]	Wild type (%)[b]
			9 (50/576)	2 (1/50)
prd–Pax3-2	23 (64/280)	22 (14/64)	61 (268/441)	74 (198/268)
prd–Pax3-3	17 (68/412)	13 (9/68)	55 (224/408)	65 (148/224)
prd–Pax3-2/*prd*–Pax3-3	12 (56/470)	86 (48/56)	53 (184/350)	94 (173/184)

[a]Percentages of rescued *prd*$^{2.45}$/*Df(2L)Prl* flies harboring combinations of *prd*–Gsb-2, *prd*–Pax3-2 and *prd*–Pax3-3 transgenes (actual numbers of rescued flies per total number of expected *prd* mutants are given in parentheses).
[b]Percentages of wild-type segmental phenotype among rescued adults (actual numbers of wild-type rescued adults per total number of rescued adults are given in parentheses).
Two to four transgenes were tested in combinations of *prd*–Gsb-2 with *prd*–Pax3-2 and *prd*–Pax3-3 for rescue of *prd*$^{2.45}$/*Df(2L)Prl* embryos to viable adults by corresponding crosses between *prd*–Gsb-2, *prd*–Pax3-2, *prd*–Pax3-3 stocks (see legend to Table II), and *prd*–Pax3-2 and *prd*–Pax3-3 stocks made homozygous for *prd*–Gsb-2, all in a *prd*$^{2.45}$ or *Df(2L)Prl/SM1* background.

development to the adult stage, but not those functions responsible for proper segmentation of the adult cuticle.

Discussion

In summary, our results demonstrate that (i) Gsb protein is able to substitute for all functions of Prd required to support development to adults but it cannot perform all of these functions with the same efficiency as Prd; (ii) Pax3 protein cannot substitute for all Prd functions but is able to replace some functions more efficiently than Gsb; (iii) Pax3 protein executes all of Prd's functions required for the development of wild-type larval cuticle although with lower efficiency than Prd; and (iv) *prd*–Pax3 is able to exert a dominant negative effect on *prd*–Gsb in the absence of wild-type Prd protein, since, unlike Gsb, Pax3 is unable to substitute for all functions of Prd. These findings may be explained by the following molecular model. We assume that Prd performs its various functions by binding to several different partners with different affinities. For wild-type Prd function, these DNA or protein binding affinities of the Prd protein for each of its partners must then be between certain threshold limits. If the affinity of a mutant Prd protein for one or several of its partners is not too far below the corresponding threshold values for wild-type function, it can be compensated by an increased dosage of the mutant allele enhancing the concentration of the mutant protein. On the other hand, if one of the mutant affinities is too low, this particular function of Prd cannot be compensated by additional dosages of the mutant allele.

Evolutionary alleles of prd

It is easy to see that our results (i)–(iv) may be explained by such a model if we consider, in a Gedankenexperiment, Gsb and Pax3 as leaky mutant proteins of Prd. Accordingly, *prd*–Gsb and *prd*–Pax3 are hypomorphic *prd* alleles, which we propose to call 'evolutionary' alleles of *prd*, as their

coding regions and that of *prd* have been derived from a common ancestral gene during evolution. Both of these mutant proteins have only few alterations with respect to Prd in their N-terminal halves, which consist of a highly conserved paired-domain and *prd*-type homeodomain, while their C-terminal halves have diverged to such an extent that their descent from a common precursor is difficult, if not impossible, to ascertain. It is therefore particularly astonishing that Gsb is able to substitute for all, and Pax3 for many, of Prd's functions required for survival. Since Gsb and Pax3 perform some of these functions less efficiently than Prd, we conclude that these proteins have somewhat reduced affinities for the partners of Prd. For Gsb, these affinities are only slightly below the threshold levels required for wild-type Prd function, while for Pax-3 the affinity for at least one partner crucial for survival to adulthood is too low to be compensated by increased dosage of *prd*–Pax3. By extrapolation of this argument, we expect that Gsb would be able to rescue all prd^- embryos to adults if *prd*–Gsb is present in sufficient numbers, whereas Pax3 would not. Formally, the *prd*–Gsb-2 transgene behaves as a leaky lethal allele of *prd* whose penetrance is ~90% since 9% of *prd*–Gsb-2 homozygotes survive to adulthood in a prd^- background. The *prd*–Gsb-2 allele is weaker than the weakest previously known *prd* allele, prd^2, which is homozygous lethal and displays a cuticular phenotype (Nüsslein-Volhard and Wieschaus, 1980) very similar to that of hemizygous *prd*–Gsb-2 (Figure 4C). Surprisingly, despite the numerous mutations in its coding region, the *prd*–Gsb-2 allele is the only known viable *prd* allele. This unusual fact uncovered the previously unknown involvement of *prd* in the proper abdominal segmentation of adults. Hence, our approach of constructing evolutionary alleles might provide a novel strategy for detecting protein functions not accessible by other methods.

General aspects of structure–function analysis by exchange of gene domains

The structure–function analysis of genes in relation to evolution has progressed through several stages. Experiments carried out by evolution itself demonstrate the generation of novel functions by combining similar protein coding regions with different control regions (Carroll, 1995). The same principle is observed with dramatic gain-of-function or neomorphic mutants like *Antp* in *Drosophila*. These arise spontaneously from novel combinations of *cis*-regulatory with protein coding domains, which result in the ectopic expression of the protein (Frischer *et al.*, 1986; Schneuwly *et al.*, 1987a,b). The same or similar dominant gain-of-function phenotypes have been generated in *Drosophila* by ectopic expression of the mouse or human homologs, suggesting that homologous proteins of different species may have conserved their functions during evolution (McGinnis *et al.*, 1990; Malicki *et al.*, 1990; Zhao *et al.*, 1993). The next stage examined the functional attributes of the various parts of homologous genes that have acquired novel functions within the same species as well as across phyla. Our previous experiments in *Drosophila* (Li and Noll, 1994) and elegant experiments in the mouse (Hanks *et al.*, 1995) are consistent with the essential role of control elements in determining a gene's function. While the mouse experiments emphasized the selection for overlapping gene functions prevalent in mammals, our experiments focussed on the functional significance of conserving dissimilar protein domains in genes with non-overlapping functions. The present work takes the analysis one step further by demonstrating the functional equivalence of highly diverged protein domains of duplicated genes that have evolved under conditions that ensured their independent selection. In contrast to earlier work, these experiments also seriously question the validity of the paradigm that sequence similarity is a necessary condition for protein folding and hence conservation of protein function. In the following, we discuss the individual aspects supporting these conclusions.

Paradox of conserved C-terminal moieties of Prd, Gsb and Pax3

The Prd protein needs to fulfill several distinct functions during development. In addition to its early embryonic function in segmentation (Nüsslein-Volhard and Wieschaus, 1980), we have shown that Prd functions are required for development to adulthood and for proper segmentation of the adult cuticle (Figure 5 and Table III), functions that have escaped detection by previously known *prd* alleles. It is possible that these requirements for Prd occur during embryogenesis as well since Prd protein begins to accumulate in a few specific cell clusters of the gnathal segments and in specific neurons of the central nervous system after its epidermal pattern of segmental stripes has disappeared (Gutjahr *et al.*, 1993a). Although these later embryonic functions of Prd are not known, they appear to be indispensable for survival as indicated by the observation that all prd^- embryos carrying two copies of *prd*–Pax3-1 die by the end of embryogenesis whereas 84% of prd^- embryos carrying two copies of *prd*–Gsb-2 survive to larval stages. Similarly, Prd functions for proper segmentation of the adult abdominal cuticle may be required during embryogenesis, for example for the proper activation of the cognate target genes in the abdominal histoblasts.

We have shown that both Gsb and Pax3 proteins are able to substitute for the Prd functions required for the proper segmentation of the larval and adult cuticle. However, while Gsb can provide all Prd functions necessary for survival, Pax3 cannot. Since the two N-terminal DNA binding domains, the paired-domain and the homeodomain, are highly conserved in Prd, Gsb and Pax3, it would make sense to postulate that those functions for which Pax3 is unable to substitute are associated with the C-terminal half of Prd. Quite unexpected, however, is our finding that Gsb and Pax3 are able to substitute for many Prd functions despite the lack of any apparent similarity in their C-terminal moieties. Clearly, only some functional equivalence not apparent in the amino acid sequence could have ensured the retention of the C-terminal portion of these molecules over hundreds of millions of years. We conclude that certain structural features, possibly of secondary and tertiary folding, must be responsible for the common functions observed. The validity of amino acid identity as a general measure of functional equivalence in homologous proteins has been criticized by Yockey (1992) who proposed to replace it by calculations of the mutual entropy, a more precise statistical measure that

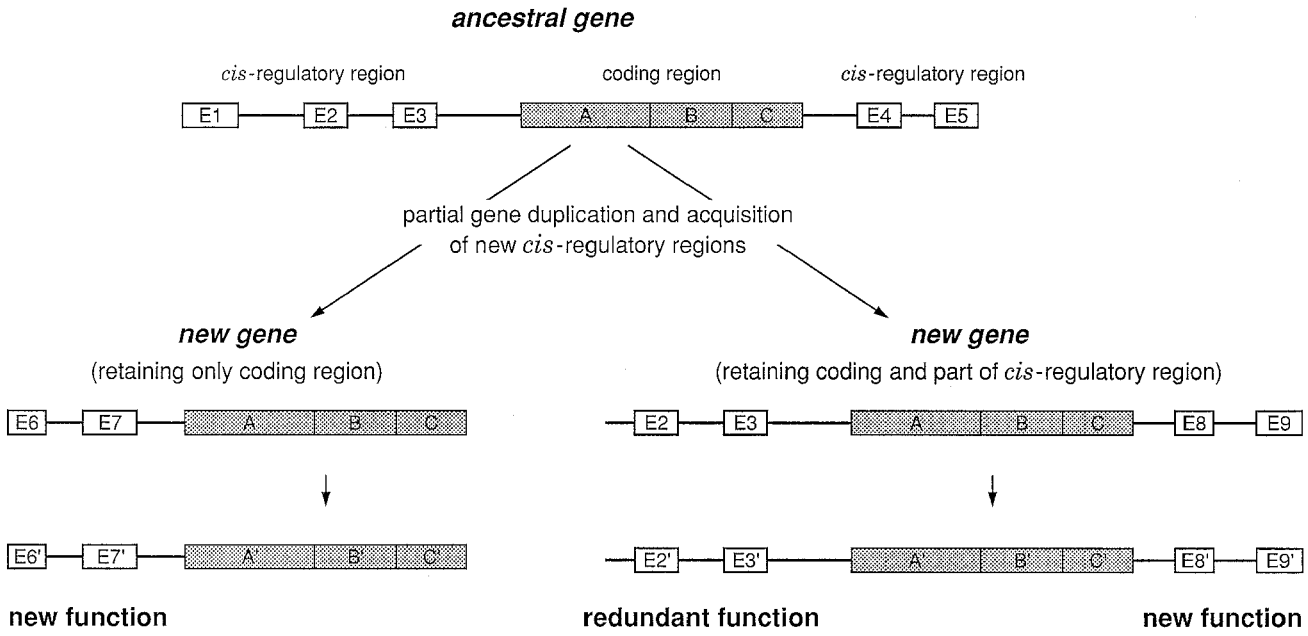

Fig. 6. Functional diversification of duplicated gene by acquisition of new *cis*-regulatory regions. The evolution of new functions by duplication of the coding region of an ancestral gene (consisting of the protein coding domains A, B and C, and the *cis*-regulatory regions E1–E5) and its combination with new *cis*-regulatory regions (E6–E9) is illustrated in new genes on the left and right. Redundancy in function may arise if some of the *cis*-regulatory regions of the ancestral gene are included in the duplication event, as shown for the new gene on the right. For discussion, see text.

takes into account the probability by which certain amino acids are replaced by others.

Driving evolution by the acquisition of new cis-regulatory regions

Our most astonishing finding is that *prd*–Gsb or combinations of *prd*–Gsb and *prd*–Pax3 transgenes can rescue *prd*$^-$ embryos to viable adults. It follows that the Prd, Gsb and Pax3 proteins are largely or entirely able to substitute for each other's functions although they have been separated for several hundred million years (Noll, 1993) and their C-terminal halves have diverged to a very large extent. Hence, the crucial differences in function between *prd*, *gsb* and presumably *Pax3* reside in their *cis*-regulatory rather than coding regions. These genes, therefore, serve as paradigm of an important evolutionary mechanism by which new functions evolve through the acquisition of new control regions, for example during gene duplication (Figure 6; Li and Noll, 1994). Furthermore, as proposed earlier (Li and Noll, 1994) and illustrated in Figure 6, gene duplications that include part of the control regions might explain the frequent observation of redundant functions of homologous genes (Hanks *et al.*, 1995). Because expression of the new gene overlaps in time and space with that of the ancestral gene, the functions regulated by the duplicated enhancers E2 and E3 become redundant while new functions evolve by the acquisition of new *cis*-regulatory elements (E8 and E9 in Figure 6). On the other hand, if gene duplication does not include any of the *cis*-regulatory regions of the ancestral gene, the new gene expresses only new functions that are regulated by the newly acquired *cis*-regulatory regions (E6 and E7 in Figure 6). Following the duplication event, protein coding or *cis*-regulatory domains mutate and thus optimize the interactions with their partners. As shown here, such subsequent changes of protein domains may result in considerable divergence of their primary structure although their functions are largely conserved.

Functional diversification of a gene may also occur by the acquisition of new *cis*-regulatory regions in the absence of its duplication as suggested recently by Carroll (1995). Similarly to the acquisition of new protein domains during evolution, the combination of new *cis*-regulatory elements with a gene may endow it with new functions. Accordingly, new genes evolve by the independent assortment of both protein domains and *cis*-regulatory 'domains', a feature that has been an explicit assumption of the gene network hypothesis (Frigerio *et al.*, 1986; Noll, 1993).

Materials and methods

Expression and purification of Prd, Gsb and Pax3 proteins

Prd, Gsb and Pax3 proteins were expressed in *Escherichia coli* BL21(DE3) cells by placing the *prd* cDNA c7340.6 (Frigerio *et al.*, 1986), the *gsb* cDNA BSH9c2 (Baumgartner *et al.*, 1987) and a *Pax3* cDNA (Goulding *et al.*, 1991), respectively, under the control of the T7 promoter in the pAR3040 vector (Studier and Moffat, 1986) and purified to ~50% purity as described (Gutjahr *et al.*, 1993a). The pARprd, pARgsb and pARPax3 expression plasmids were constructed by creating an *Nde*I site at the translation initiation codon of each cDNA (Li and Noll, 1994) and subsequent ligation of the *Nde*I–*Eco*RI cDNA fragment into *Nde*I/*Bam*HI-cut pAR3040.

Band-shift assays

For band-shift assays, proteins were preincubated with or without the indicated antiserum for 15 min at room temperature. The incubation mixture was then supplemented with binding buffer to a final volume of 20 μl of 20 mM HEPES pH 7.9, 50 mM KCl, 1 mM EDTA, 1 mM DTT, 4% Ficoll, 0.25 mg/ml BSA, 1 μg poly(dI–dC), 0.50 nM ^{32}P-labeled GEE1 (25 000 d.p.m.), incubated for another 15 min at room temperature and analyzed on a 4% polyacrylamide gel run in 0.25× TBE buffer at 10 V/cm for 3 h essentially as described (Gstaiger *et al.*, 1995). GEE1, which is derived from the *gsb* early enhancer GEE (Li *et al.*, 1993), is the following oligonucleotide (upper strand):

5′-GACCACAAGCGTCATAATTGCGGTGTGTCTGACGG-3′ (X.Li, unpublished). Its upper strand was labeled with [γ-^{32}P]ATP (3000 Ci/mmol) and T4 polynucleotide kinase and reannealed with its complementary strand according to standard procedures. The GC box oligonucleotide (36mer), including two tandem Sp1 binding site sequences from the ICP4 promoter of herpes simplex virus, was used as competitive Sp1 binding sites (Silke *et al.*, 1995).

Generation of transgenic prd–Gsb and prd–Pax3 flies

The *prd*–Gsb and *prd*–Pax3 transgenes were derived from *prd*–SN20-rescue (Gutjahr *et al.*, 1994) in two steps (cf. Figure 2A). First, the 1.7 kb *Hin*dIII–*Sac*II fragment, including 27 bp of *prd* leader sequence, most of the *prd* coding region and the *prd* intron (Frigerio *et al.*, 1986), was deleted to generate the *prd*-0 construct, by ligating the *Not*I–*Hin*dIII *prd*-upstream fragment and the *Sac*II–*Not*I *prd*-downstream and P-element vector fragment with a *Hin*dIII/*Sac*II adaptor containing a *Spe*I and *Kpn*I site (5′-AGCTTACTAGTGGTACCGC-3′). The *prd*–Gsb construct was then produced by cloning the *Xba*I–*Nhe*I fragment from pARgsb (see above in Materials and methods) in the proper orientation into the *Spe*I site of *prd*-0. The *prd*–Pax3 construct was prepared from a *Pax3* cDNA that included 297 bp of leader and the entire trailer sequence (Goulding *et al.*, 1991) and was cloned in the *Eco*RI site of pKSpL5, a derivative of Bluescript pKS$^+$ in which the *Kpn*I site is eliminated and the *Eco*RI–*Hin*dIII polylinker region replaced by the *Eco*RI–*Hin*dIII polylinker of the pUC19 vector. The *Xba*I fragment of this *Pax3* cDNA clone, which included the entire *Pax3* cDNA, was then cloned in the proper orientation into the *Spe*I site of *prd*-0 to generate the *prd*–Pax3 construct. The *prd*–Gsb and *prd*–Pax3 constructs were injected together with pUChspΔ2-3 P-element helper plasmid (D.Rio, unpublished) into ry^{506} embryos and ry^+ transformants were selected (Rubin and Spradling, 1982).

Immunostaining of embryos and in situ hybridization with DIG-labeled probe

Embryos were stained with anti-Prd or anti-Gsb antiserum (Gutjahr *et al.*, 1993a,b) or hybridized *in situ* with digoxigenin-labeled *Pax3* cDNA and stained essentially as described (Tautz and Pfeifle, 1989). Double-labeling of embryos for β-galactosidase and Gsb, Wg or En protein was performed according to Lawrence *et al.* (1987).

Acknowledgements

We are indebted to Xuelin Li for his advice and help in the initial stage of this work and Fritz Ochsenbein for his expert artwork. We thank P.Gruss for a *Pax3* cDNA, R.Smith and T.Kornberg for anti-En monoclonal antibodies, R.Nusse and M.van den Heuvel for anti-Wg antiserum, K.Basler for the *actin*–LacZ marked *CyO* chromosome stock and E.Hafen for the pUChspΔ2-3 P-element helper plasmid. We are obliged to Erich Frei and Hans Noll for stimulating discussions and thank them and Joy Alcedo for comments on the manuscript. This work has been supported by grants 31-26652.89 and 31-40874.94 from the Swiss National Science Foundation and by the Kanton Zürich.

References

Baldwin,C.T., Hoth,C.F., Amos,J.A., Da-Silva,E.O. and Milunsky,A. (1992) An exonic mutation in the *HuP2* paired domain gene causes Waardenburg's syndrome. *Nature*, **355**, 637–638.

Baumgartner,S., Bopp,D., Burri,M. and Noll,M. (1987) Structure of two genes at the *gooseberry* locus related to the *paired* gene and their spatial expression during *Drosophila* embryogenesis. *Genes Dev.*, **1**, 1247–1267.

Bopp,D., Burri,M., Baumgartner,S., Frigerio,G. and Noll,M. (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., Jamet,E., Baumgartner,S., Burri,M. and Noll,M. (1989) Isolation of two tissue-specific *Drosophila* paired box genes, *pox meso* and *pox neuro*. *EMBO J.*, **8**, 3447–3457.

Burri,M., Tromvoukis,Y., Bopp,D., Frigerio,G. and Noll,M. (1989) Conservation of the paired domain in metazoans and its structure in three isolated human genes. *EMBO J.*, **8**, 1183–1190.

Carroll,S.B. (1995) Homeotic genes and the evolution of arthropods and chordates. *Nature*, **376**, 479–485.

Côté,S., Preiss,A., Haller,J., Schuh,R., Kienlin,A., Seifert,E. and Jäckle,H. (1987) The *gooseberry-zipper* region of *Drosophila*: five genes encode different spatially restricted transcripts in the embryo. *EMBO J.*, **6**, 2793–2801.

Epstein,D.J., Vekemans,M. and Gros,P. (1991) *splotch* (*Sp^{2H}*), a mutation affecting development of the mouse neural tube, shows a deletion within the paired homeodomain of *Pax-3*. *Cell*, **67**, 767–774.

Frigerio,G., Burri,M., Bopp,D., Baumgartner,S. and Noll,M. (1986) Structure of the segmentation gene *paired* and the *Drosophila* PRD gene set as part of a gene network. *Cell*, **47**, 735–746.

Frischer,L.E., Hagen,F.S. and Garber,R.L. (1986) An inversion that disrupts the *Antennapedia* gene causes abnormal structure and localization of RNAs. *Cell*, **47**, 1017–1023.

Goulding,M.D., Chalepakis,G., Deutsch,U., Erselius,J.R. and Gruss,P. (1991) Pax-3, a novel murine DNA binding protein expressed during early neurogenesis. *EMBO J.*, **10**, 1135–1147.

Gstaiger,M., Knoepfel,L., Georgiev,O., Schaffner,W. and Hovens,C.M. (1995) A B-cell coactivator of octamer-binding transcription factors. *Nature*, **373**, 360–362.

Gutjahr,T., Frei,E. and Noll,M. (1993a) Complex regulation of early *paired* expression: initial activation by gap genes and pattern modulation by pair-rule genes. *Development*, **117**, 609–623.

Gutjahr,T., Patel,N.H., Li,X., Goodman,C.S. and Noll,M. (1993b) Analysis of the *gooseberry* locus in *Drosophila* embryos: *gooseberry* determines the cuticular pattern and activates *gooseberry neuro*. *Development*, **118**, 21–31.

Gutjahr,T., Vanario-Alonso,C.E., Pick,L. and Noll,M. (1994) Multiple regulatory elements direct the complex expression pattern of the *Drosophila* segmentation gene *paired*. *Mech. Dev.*, **48**, 119–128.

Hanks,M., Wurst,W., Anson-Cartwright,L., Auerbach,A.B. and Joyner, A.L. (1995) Rescue of the *En-1* mutant phenotype by replacement of *En-1* with *En-2*. *Science*, **269**, 679–682.

Ingham,P.W. and Martinez Arias,A. (1992) Boundaries and fields in early embryos. *Cell*, **68**, 221–235.

Kilchherr,F., Baumgartner,S., Bopp,D., Frei,E. and Noll,M. (1986) Isolation of the *paired* gene of *Drosophila* and its spatial expression during early embryogenesis. *Nature*, **321**, 493–499.

Lawrence,P.A., Johnston,P., Macdonald,P. and Struhl,G. (1987) Borders of parasegments in *Drosophila* embryos are delimited by the *fushi tarazu* and *even-skipped* genes. *Nature*, **328**, 440–442.

Li,X. and Noll,M. (1993) Role of the *gooseberry* gene in *Drosophila* embryos: maintenance of *wingless* expression by a *wingless–gooseberry* autoregulatory loop. *EMBO J.*, **12**, 4499–4509.

Li,X. and Noll,M. (1994) Evolution of distinct developmental functions of three *Drosophila* genes by acquisition of different cis-regulatory regions. *Nature*, **367**, 83–87.

Li,X., Gutjahr,T. and Noll,M. (1993) Separable regulatory elements mediate the establishment and maintenance of cell states by the *Drosophila* segment-polarity gene *gooseberry*. *EMBO J.*, **12**, 1427–1436.

McGinnis,N., Kuziora,M.A. and McGinnis,W. (1990) Human *Hox-4.2* and *Drosophila Deformed* encode similar regulatory specificities in *Drosophila* embryos and larvae. *Cell*, **63**, 969–976.

Malicki,J., Schughart,K. and McGinnis,W. (1990) Mouse *Hox-2.2* specifies thoracic segmental identity in *Drosophila* embryos and larvae. *Cell*, **63**, 961–967.

Noll,M. (1993) Evolution and role of *Pax* genes. *Curr. Opin. Genet. Dev.*, **3**, 595–605.

Nüsslein-Volhard,C. and Wieschaus,E. (1980) Mutations affecting segment number and polarity in *Drosophila*. *Nature*, **287**, 795–801.

Patel,N.H., Schafer,B., Goodman,C.S. and Holmgren,R. (1989) The role of segment polarity genes during *Drosophila* neurogenesis. *Genes Dev.*, **3**, 890–904.

Peifer,M. and Bejsovec,A. (1992) Knowing your neighbors: cell interactions determine intrasegmental patterning in *Drosophila*. *Trends Genet.*, **8**, 243–249.

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

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

Schneuwly,S., Klemenz,R. and Gehring,W.J. (1987a) Redesigning the body plan of *Drosophila* by ectopic expression of the homeotic gene *Antennapedia*. *Nature*, **325**, 816–818.

Schneuwly,S., Kuroiwa,A. and Gehring,W.J. (1987b) Molecular analysis of the dominant homeotic *Antennapedia* phenotype. *EMBO J.*, **6**, 201–206.

Silke,J., Rother,K.I., Georgiev,O., Schaffner,W. and Matsuo,K. (1995) Complex demethylation patterns at Sp1 binding sites in F9 embryonal carcinoma cells. *FEBS Lett.*, **370**, 170–174.

Small,S. and Levine,M. (1991) The initiation of pair-rule stripes in the *Drosophila* blastoderm. *Curr. Opin. Genet. Dev.*, **1**, 255–260.

Studier,F.W. and Moffat,B.A. (1986) Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.*, **198**, 113–130.

Tassabehji,M., Read,A.P., Newton,V.E., Harris,R., Balling,R., Gruss,P. and Strachan,T. (1992) Waardenburg's syndrome patients have mutations in the human homologue of the *Pax-3* paired box gene. *Nature*, **355**, 635–636.

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

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

Yockey,H.P. (1992) *Information Theory and Molecular Biology*. Cambridge University Press, Cambridge, UK.

Zhao,J.J., Lazzarini,R.A. and Pick,L. (1993) The mouse *Hox-1.3* gene is functionally equivalent to the *Drosophila Sex combs reduced* gene. *Genes Dev.*, **7**, 343–354.

Received on February 1, 1996; revised on February 15, 1996