Mitch D. Price · Zhi-Chun Lai

The yan gene is highly conserved in Drosophila and its expression suggests a complex role throughout development

Abstract  Competence for cell fate determination and cellular differentiation is under tight control of regulatory genes. Yan, a nuclear target of receptor tyrosine kinase (RTK) signaling, is an E twenty six (ETS) DNA-binding protein that functions as a negative regulator of cell differentiation and proliferation in Drosophila. Most members of RTK signaling pathways are highly conserved through evolution, yet no yan orthologues have been identified to date in vertebrates. To investigate the degree of yan conservation during evolution, we have characterized a yan homologue from a sibling species of D. melanogaster, D. virilis. Our results show that the organization, primary structure and expression pattern of yan are highly conserved. Both genes span over 20 kb and contain four exons with introns at identical positions. The areas with highest amino acid similarity include the Pointed and ETS domain but there are other discrete regions with a high degree of similarity. Phylogenetic analysis reveals that yan’s closest relative is the human tel gene, a negative regulator of differentiation in hematopoietic precursors. In both species, Yan is dynamically expressed beginning as early as stage 4/5 and persisting throughout embryogenesis. In third instar larvae, Yan is expressed in and behind the morphogenetic furrow of the eye imaginal disc as well as in the laminar precursor cells of the brain. Ovarian follicle cells also contain Yan protein. Conservation of the structure and expression patterns of yan genes strongly suggests that regulatory mechanisms for their expression are also conserved in these two species.

Key words  Drosophila · Expression · Evolutionary conservation · Yan

Introduction

Inductive intercellular interactions play an important role in cell fate determination and pattern formation. The developing eye of Drosophila melanogaster provides an excellent system to study the molecular mechanisms of such signaling. The adult eye consists of approximately 800 repeating units known as ommatidia. Each ommatidium is comprised of 8 neuronal photoreceptor cells (R1–8), 4 cone cells, and 8 accessory cells. Eye development takes place in a monolayer epithelium known as the eye/antennal imaginal disc. Differentiation begins in the third instar larval stage and can be detected morphologically by the progression of an indentation, the morphogenetic furrow, from the posterior to the anterior of the disc. Differentiation of photoreceptor cells is a progressive phenomenon beginning with differentiation of the R8 cell. Following specification of R8, ommatidial assembly proceeds by the sequential recruitment of photoreceptors R2/R5, R3/R4, R1/R6, and finally, R7 (reviewed by Wolff and Ready 1993). This process of sequential recruitment requires inhibitory signals, which maintain precursor cells in an undifferentiated state until they receive an inductive signal.

Induction of R-cell differentiation is mediated by the Drosophila epidermal growth factor receptor homologue, DER (reviewed by Freeman 1997, 1998). In the case of the R7 cell, a receptor tyrosine kinase (RTK) encoded by sevenless is also required for proper specification. Several nuclear targets of RTK signaling have been identified including Pointed, Sina, Phyllopod (Phyl), D-Jun, Prospero and Yan (reviewed by Zipursky and Rubin 1994; Wassarman and Therrien 1997). The yan gene encodes a nuclear protein with an E twenty six (ETS) DNA-binding motif, a Pointed (PNT) domain, 8 mitogen-activated protein (MAP) kinase phosphorylation sites, and several putative PEST (Pro, Glu, Ser and Thr) signals. Yan functions as a negative regulator of photoreceptor cell differentiation since loss-of-function mutations in yan result in extra R cells in a cell autonomous manner (Lai and Rubin 1992; Tei et al. 1992). Recently, Yan has been shown...
to synergistically interact with another nuclear transcriptional repressor, tramtrack (ttk), to block neural differentiation in the developing eye (Lai et al. 1997). Inductive signaling is thought to allow expression of Phyl, which in turn associates with Sina to target Ttk for degradation (Li et al. 1997; Tang et al. 1997).

In addition to its function in the developing eye, yan plays critical roles during embryogenesis. The yan gene was independently identified in a screen for embryonic lethal mutations where it was named anterior open (aop; Nusslein-Volhard et al. 1984; Lindsley and Zimm 1992). Recent work has shown that Yan is a repressor of cell shape changes during dorsal closure, and that this repression is relieved following phosphorylation by Jun kinase (Hou et al 1997; Kockel et al. 1997; Riesgo-Escovar and Hafen 1997; reviewed by Noselli 1998). An additional function of yan in the ectoderm is in the differentiation of engrailed (en) expressing cells, where Yan must be inactivated to specify the denticle-secreting en cell type (O’Keefe et al. 1997). Analysis of Yan function earlier in embryogenesis revealed that it is a repressor of ventral cell fate in the ectoderm. In loss-of-function yan mutant embryos, expression of the ventral cell markers orthodenticle (otd), argos (aos), and tartan (trn) is expanded, whereas expression of activated Yan reduces expression of otd and aos (Gabay et al. 1996). Finally, yan is responsible for differentiation of the dorsal head ectoderm; a tissue giving rise to the visual system, medial brain, and head epidermis. Loss-of-function yan mutants display massive head defects caused by over proliferation and subsequent cell death in these precursors (Rogge et al. 1995).

Despite extensive genetic and biochemical data showing how Yan activity can be regulated by RTK signaling in D. melanogaster (reviewed by Wasserman and Theiren 1997; Schweitzer and Shilo 1997; Noselli 1998), there is relatively little data investigating the structural basis of Yan function. We were also troubled that no yan orthologues have been identified in vertebrates despite extensive conservation between other members of the RTK signaling cascade. To address these issues we took an evolutionary approach and report here the identification and structural characterization of the yan gene from D. virilis. Comparison of the Yan polypeptide sequence from D. virilis and D. melanogaster reveals a high degree of structural similarity, especially in specific subdomains of the protein. Phylogenetic analysis shows that yan’s only relative is the human gene tel. To further investigate the degree to which yan function is likely to be conserved, we compared the expression of Yan in D. melanogaster and D. virilis and again show that there is a high degree of conservation in the spatial and temporal pattern of Yan expression. This argues strongly that both the function as well as the mechanisms for regulating the expression and activity of Yan have been well conserved.

Materials and methods

Fly stocks

All stocks were maintained on standard yeast-agar medium at 18 or 25°C. In all cases Canton-S flies served as wild-type controls. D. virilis flies were obtained from either the Bloomington Stock Center or the Rubin Laboratory. The yan* allele, which has been previously described (Lai and Rubin 1992), contains a P-element insertion in the 5' portion of intron 1 and expresses β-galactosidase in all of the tissues and stages in which Yan protein is expressed.

Isolation of the D. virilis yan gene

A λEMBL3A D. virilis genomic DNA library (a gift of J. Tamkun) was screened at reduced stringency with two probes derived from D. melanogaster yan cDNA corresponding to the PNT and ETS domains. Restriction fragments from candidate phage DNA were subcloned into pBluescript KS+ (Stratagene) using standard conditions (Sambrook et al. 1989). Plasmid DNA was sequenced using dideoxynucleotide chain termination either by manual sequencing with T7 DNA polymerase (Sequenase version 2.0, United States Biochemical Corp) or automated sequencing using 3' dye-labeled dideoxynucleotide triphosphates in cycle sequencing reactions run on an ABI PRISM 377 DNA sequencer (Perkin-Elmer ABI, Foster City, Calif.) at the Nucleic Acid Facility of the Biotechnology Institute at The Pennsylvania State University. Sequence data was analyzed using the Lasergene software package (DNASTAR).

cDNA analysis

cDNA fragments spanning the intron/exon junctions were amplified using reverse transcription-polymerase chain reaction (RT-PCR). For sequences flanking intron 2, total RNA was isolated from 100 3rd instar larval imaginal disc complexes using the Promega total RNA isolation reagents. Reverse transcription and PCR were performed on 0.5 μg of RNA using the Access RT-PCR system (Promega) and the following primers: dv1373 (5'-CATGAGCCCAAATTTGACGCGC-3') and dv3912 (5'-CAGCTGCATTATTGCTCAGG-3'). Products from the first PCR amplification were diluted x100 and a second PCR reaction was performed using primers dv1373 and dv2810 (5'-TTGCCGAGAATTTGACGGCGG-3'). Products from these reactions were subcloned and sequenced.

Intron 3 boundaries were confirmed in a similar way, except amplifications were done using double stranded cDNA, synthesized from approximately 5 μg of whole third instar larval polyA+ RNA, and the Marathon cDNA Amplification Kit (Clontech). Primers for the first PCR reaction were dv1373 and dv3912. The second PCR reaction was performed with primers dv3368 (5'-CCGATCTGACCTGTAGACCCGCCA-3') and dv3912. A 480-bp product from this reaction was subcloned into pBluescript KS+ which had been digested with EcoRV and subsequently incubated for 2 h at 70°C with 2.5 mM EDTA in the presence of 5U Taq polymerase in a total volume of 200 μl.

A cDNA fragment containing a portion of exon 1 was amplified by rapid amplification of cDNA ends (RACE)-PCR from double stranded cDNA to which adaptors had been ligated (Marathon cDNA Amplification Kit, Clontech). Amplifications were performed with a primer in the adaptor (AP1) and primers dv628 (5'-GCAGCTGCATTATTGCTCAGG-3') or dv2810. Subsequent PCR reactions were performed on x100 dilutions of primary PCR products with nested primers in the adaptor (AP2) and either: dv672 (5'-GATCAAAGTCTGACCGTGGG-3'), dv628, or dv498 (5'-AGATGCTGG-GAT-TGAGCCG-3'). From this amplification regime specific products were identified, subcloned into pBluescript and sequenced.
Evolutionary analysis

The sequences of known ETS genes containing PNT domains were obtained from Genbank (http://ncbi.nlm.nih.gov) and a list of those used in this study is shown (Table 1). Amino acid sequences were aligned using the Clustal V program (Higgins et al. 1992). The MEGA program (Kumar et al. 1993) was used to construct phylogenetic trees by the Neighbor-Joining method (Saitou and Nei 1987). Pairwise evolutionary distances (number of substitutions/amino acid site) were calculated using the p-distance.

In situ hybridization

Embryos were collected on molasses agar plates smeared with live yeast paste. Fixation, pretreatment, and hybridization were carried out as described (Lehmann and Tautz 1994). Digoxigenin-labeled antisense RNA probes from the D. virilis or D. melanogaster Yan coding region were used for hybridization. Binding of the probe was detected using an alkaline phosphatase conjugated anti-DIG antibody (1:2000, Boehringer Mannheim).

Antibody staining

Throughout this study, a mouse monoclonal antibody against the C-terminal 367 amino acids of D. melanogaster Yan (Yan-Yan; a gift of I. Rebay and G. M. Rubin) was used. For early D. melanogaster embryos, a mouse monoclonal antibody raised against amino acids 69–382, which includes the Pointed domain, was used (Yan-A2; a gift of I. Rebay and G. M. Rubin). Embryos were fixed for whole-mount antibody staining as described (Patel 1994). Antibody binding was detected using a biotinylated goat anti-mouse antibody (1:200) and the Vectastain reagents for color development (Vector Laboratories). Embryos were mounted in methyl-salicylate and photographed on a Zeiss Axiophot compound microscope.

Tissues from third instar larvae including the brain, ventral ganglion, and imaginal discs, were immunostained with α-Yan antibodies diluted 1:5 as described (Kunes et al. 1993). Antibody staining was detected with Vectastain ABC color developing system (Vector Laboratories). Stained tissue was mounted in 80% glycerol and photographed.

Ovaries from mature D. melanogaster and D. virilis females were prepared as described (Verheyen and Cooley 1994). Antibody staining was performed using the same conditions as described for embryos except that α-Yan antibodies were used at 1:1 dilution. Stained ovaries were mounted in 50% glycerol and photographed.

Results

Primary structure of the yan gene is highly conserved in D. virilis

yan is an integral part of the Ras/MAP kinase signaling pathway in D. melanogaster. Although components of this pathway are well conserved in evolution, no yan orthologues have been identified to date. To begin investigating the structural and functional conservation of the yan gene, we sought to identify a yan homologue in D. virilis. D. virilis is separated from D. melanogaster by approximately 60 million years (Beverley and Wilson 1984), making it a good candidate in which to examine important functional domains of homologous genes on the basis of their evolutionary conservation. Using probes from the D. melanogaster yan gene we isolated a number of D. virilis yan genomic and cDNA clones. Structurally, the D. virilis yan gene is very similar to D. melanogaster (Fig. 1A). Both genes consist of four exons with introns at identical positions. The size of intron 3 is nearly identical, 64 bp in D. virilis and 62 bp in D. melanogaster. Intron 2 in D. virilis is larger, however, 931 bp compared to 66 bp in D. melanogaster. Comparing the sequence of D. melanogaster intron 2 with the second intron of D. virilis failed to identify any significant similarity, suggesting that there is little or no evolutionary constraint on sequences within the intron. Intron 1 in D. virilis is approximately 12 kb long as compared to the 17 kb intron 1 from D. melanogaster. To confirm the predicted intron/exon boundaries in D. virilis, sequences were obtained from RT-PCR amplified cDNA fragments. The sequences of these boundaries are shown (Fig. 1B), and in the case of intron 3 the entire intron sequence is shown. Nucleotide sequence similarity surrounding the splice junctions extends for several bases within the exon, but diverges rapidly within the introns.

The primary amino acid sequence of D. virilis Yan exhibits extensive similarity to the D. melanogaster protein (Fig. 1C). The most strikingly conserved regions are within the PNT and ETS domains, where the amino acid identity is 97 and 100% respectively. This underscores the importance of these domains with regard to the function of Yan. Identity within other regions of the protein is also high. The amino terminal 52 amino acids, which precede the PNT domain, exhibit 90% identity overall; whereas amino acid identity in the intervening region between PNT and ETS domains drops to 57%. The two proteins differ most at the carboxy terminus where the overall identity is 51%. Nevertheless there are still discrete regions within these domains that are well conserved. Most notably are two conserved sequences, FQFHP and P-DLS, that represent motifs important for the function of other proteins. The FQFHP sequence is the site of several gain-of-function mutations in the lin-1 gene from C. elegans (Jacobs et al. 1998). Lin-1 is an ETS gene that functions as a negative regulator of vulval cell fate (Beitel et al. 1995). Like Yan, Lin-1 activity is down-regulated by MAP kinase phosphorylation (Tan et al. 1998) suggesting that despite differences in their structure (Lin-1 has no PNT domain) they may be functionally related. P-DLS is the sequence recognized by the dCtBP transcriptional co-repressor (Nibu et al. 1998; Poortinga et al. 1998). Given the conservation of the dCtBP binding site, and Yan’s role as a transcriptional repressor, it is possible that the mechanism of Yan function involves association with the dCtBP co-repressor.

All eight consensus MAP kinase phosphorylation sites found in the Yan protein from D. melanogaster are also conserved in D. virilis. These sites are important for the inactivation of Yan following phosphorylation by MAP kinase in D. melanogaster (Rebay and Rubin 1995) and thus may play a similar role in D. virilis. There are also a number of putative PEST signals in Yan proteins from both species. These sequences are often found in rapidly degraded proteins, including those
found in signaling pathways (reviewed by Rechsteiner and Rogers 1996). The sequence from amino acid position 266–286 in *D. virilis* Yan represents a very strong consensus PEST site that is not present in *D. melanogaster* indicating there may be differences in Yan stability between these two species. The existence of multiple PEST sequences in Yan from evolutionarily distant species suggest that regulated proteolysis may play an important and conserved role in regulating Yan activity.

Within the Yan protein of both species there are multiple glutamine-rich repeats. These repeats are present in a number of developmentally important *Drosophila* genes as well as clinically relevant vertebrate genes.

---

Fig. 1A–C Structural comparison of *Drosophila melanogaster* and *D. virilis* yan genes. A Restriction map of the yan locus from both species. Restriction enzymes represented are BamHI (B), SalI (S), HindIII (H), and SacI (Ss). Exons are represented by boxes, with the coding portion solid. The 5' boundary of *D. virilis* exon 1 has not been determined, but its relative position is known from overlapping phage clones. B Sequence comparison of intron/exon boundaries. Exon sequences are contained within boxes and the nucleotides at the boundary are boldface. C Sequence of Yan protein from both species. The Pointed and ETS domains are indicated with arrows. MAP kinase phosphorylation sites are boxed, and the position of introns are marked with an arrowhead. The FQFHP and P-DLS sequences are underlined. Identical amino acids are marked * while conservative amino acid changes are designated with a
There is evidence to support a role for these regions in transcriptional regulation (Ali and Bienz 1991; Gerber et al. 1994), but little is known about the mechanism by which these repeats function.

Evolutionary relationship of ETS genes

Among the family of ETS genes, only a subset contain both ETS and PNT domains (Table 1; Klambt 1993). To assess the relationship of yan gene to other members of this group, we have compared the PNT domain of these genes. The amino acid alignment (Fig. 2A) exhibits several interesting points. There are 5 residues that are absolutely conserved among all 23 proteins tested in this analysis, suggesting that these residues are critical for the function of the PNT domain. It is interesting that these residues are not clustered in particular regions, but instead are evenly distributed within the PNT domain. An additional 6 residues, however, are conserved in all but yan and/or tel genes. If these positions are taken into ac-

![Fig. 2A](image-url) **Fig. 2A, B** Phylogenetic analysis of Pointed domains. A An alignment of Pointed domains from all ETS/Pointed proteins is shown. Positions that are identical in all 23 proteins are shaded with dark boxes and marked with a diamond while those that are conserved in all except Yan and/or Tel are shaded with light boxes and marked by a dot. B The phylogenetic tree showing the relationship of Pointed domains is shown. The numbers indicate the bootstrap confidence in each node. The three major branches representing evolutionary related groups are indicated by brackets.
count, there are 2 clusters of conserved sequences at the amino terminus and carboxy terminus of the PNT domain.

To further assess the relationship of PNT domains in this group, we generated a phylogenetic tree (Fig. 2B) using the Neighbor-Joining algorithm (Saitou and Nei 1987). This analysis indicates that the PNT domains group into three major branches. The first group consists of the human ets1, ets2 and Drosophila pnt, elg genes. The Drosophila elg gene may not be a bona fide member of this group as indicated by a bootstrap confidence level of less than 50%. The second major branch consists of vertebrate GABP, ERG and FLI genes. The final cluster in this tree contains only the yan gene from D. melanogaster and D. virilis and the human tel gene. To determine if the relationship of Yan and Tel extended beyond the PNT domain, we performed a similar analysis with the sequence of 87 known ETS domains and obtained comparable results (data not shown). The apparent segregation of yan and tel from the other ETS/PNT genes suggests that these two proteins may have arisen from a common lineage and despite relatively low sequence similarity might function in a similar manner.

The D. virilis yan gene exhibits conserved and complex expression during development

In order to determine if the conservation of Yan between the two species extended beyond the structural level, we analyzed the expression of Yan in D. virilis and D. melanogaster. Previously Gabay et al. (1996) examined yan expression and function in D. melanogaster embryos and found that yan mRNA is detectable by stage 5 in the neurogenic region and the dorsal ectoderm. During gastrulation, Gabay et al. reported that yan expression becomes restricted to the head region and the ectodermal layer, but is absent in several rows of cells along the ventral midline. Later in embryonic development, yan mRNA was detected in the tracheal pits and in cells close to the ventral midline in the epidermal layer.

To further define the temporal and spatial pattern of yan expression in both D. melanogaster and D. virilis we used antibodies raised against D. melanogaster Yan protein to examine expression in both species. In stage 5 D. virilis embryos, we detect Yan protein in all nuclei at the periphery of the cellular blastoderm (Fig. 3B). In D. melanogaster we also detect this pattern of Yan immuno-

### Table 1  Ets genes used for phylogenetic analysis with GenBank accession numbers

<table>
<thead>
<tr>
<th>Mammals</th>
<th>Birds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homo sapiens (Hs)</td>
<td>Gallus Gallus (Gg)</td>
</tr>
<tr>
<td>c-ets-1: X14798</td>
<td>c-ets: X10326</td>
</tr>
<tr>
<td>erg: M21535</td>
<td>c-ets-2: X07202</td>
</tr>
<tr>
<td>erg2: M17254</td>
<td>erg: X77159</td>
</tr>
<tr>
<td>ets-2: J04102</td>
<td>c-ets1: M21823</td>
</tr>
<tr>
<td>fli-1: M93255</td>
<td>ets: M22462</td>
</tr>
<tr>
<td>GABPa: Q06546</td>
<td>tel: U11732</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mus musculus (Mm)</th>
<th>Amphibians</th>
</tr>
</thead>
<tbody>
<tr>
<td>c-ets-1: X55787</td>
<td>Xenopus laevis (Xl)</td>
</tr>
<tr>
<td>c-ets-2A: M81683</td>
<td></td>
</tr>
<tr>
<td>c-ets-2B: X52635</td>
<td></td>
</tr>
<tr>
<td>fli-1: X59421</td>
<td></td>
</tr>
<tr>
<td>GABPa: M74515</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Rattus norvegicus (Rn)</th>
<th>Arthropods</th>
</tr>
</thead>
<tbody>
<tr>
<td>c-ets-1A: P18755</td>
<td>Drosophila melanogaster (Dm)</td>
</tr>
<tr>
<td>c-ets-1: X55787</td>
<td>elg: X68259</td>
</tr>
<tr>
<td>c-ets-2A: M81683</td>
<td>pnt: X69166</td>
</tr>
<tr>
<td>c-ets-2B: X52635</td>
<td>tel: M97694</td>
</tr>
<tr>
<td>fli-1: X59421</td>
<td>yan: AF045593</td>
</tr>
<tr>
<td>GABPa: M97694</td>
<td></td>
</tr>
<tr>
<td>fli-1: X59421</td>
<td></td>
</tr>
<tr>
<td>GABPa: M74515</td>
<td></td>
</tr>
</tbody>
</table>

Fig. 3A–L. Yan is dynamically expressed in early embryos. Yan expression in D. melanogaster (A, C, E, G, I, K) and D. virilis (B, D, F, H, J, L) is shown. In stage 5 embryos Yan immunoreactivity is detected in all peripheral nuclei (A, B). In situ hybridization with yan antisense RNA probes shows yan expression in a similar pattern in both species (C, D). In a ventral view of stage 10 embryos (E–F) Yan protein is detected in the ectoderm, except for several rows of cells along the midline. By stage 11 (G–J) expression is seen in the ventral-most cells in the ectoderm, with the exception of the midline (black arrowhead). A dorsolateral view of stage 11 embryos (I, J) shows that in addition to expression in the ventral ectoderm, Yan expression is concentrated in cells surrounding the tracheal pits (white arrowhead) and in a group of cells slightly posterior and beneath the tracheal pits. Expression of Yan in stage 13 embryos is limited to the ectoderm and tracheal system (K–L). In each case, anterior is to the left.
staining (Fig. 3A). Because this is not in agreement with the in situ hybridization result previously reported (Gabay et al. 1996) we examined the expression of yan mRNA in both species to corroborate our immunostaining result. In stage 5 embryos from both species we detect yan mRNA throughout the embryo (Fig. 3C–D). As expected, in situ hybridization with a sense RNA probe fails to generate any hybridization signal (data not shown). We also detect uniform expression of lacZ in the YanP enhancer trap line (data not shown). In some earlier stage embryos we could detect a gradient of expression similar to that described by Gabay et al. (1996), but we failed to detect this pattern in D. virilis (data not shown).

During germ band extension, the pattern of Yan immunostaining in the ectoderm reflects the in situ hybridization results reported by Gabay et al. (1996). That is, Yan protein is detected in the presumptive ectoderm but is excluded from the midline and several rows of ventral-most ectodermal cells (Fig. 3E–F). A closer examination of the Yan immunostaining pattern reveals repeating groups of cells extending into the dorsal ectoderm that have higher levels of Yan expression. In D. melanogaster embryos, yan is required for determination of cell fates in the ectoderm (Gabay et al. 1996; O’Keefe et al. 1997).

Given the similar pattern of expression in D. virilis, it is likely that Yan serves a conserved function in determination of ventral cell fate in this species. Yan expression is also seen in the procephalic region of the head in both D. melanogaster and D. virilis suggesting that the role of yan in formation of brain structures (Rogge et al. 1995; Dumstrei et al. 1998) has also been preserved.

The expression pattern in later germ band-extended embryos (stage 11) becomes much more complicated. Expression persists in the ventral ectoderm as before, but expands closer to the ventral midline (Fig. 3G–H). At this stage, expression can be further defined around the tracheal pits and in a group of cells located posterior and slightly beneath the tracheal pits (Fig. 3I–J). Additional expression in the presumptive mesoderm of both species is detected with a –Yan antibodies at this stage (data not shown). Expression in this tissue had not been previously reported and the role for yan in this tissue has not yet been ascribed, although over-expression of an activated form of Yan does cause defects in tissues derived from the mesoderm (Rebay and Rubin 1995; Shishido et al. 1997). Expression in the head region persists during this stage, but appears to become more concentrated around the maxillary, mandibular, and labial buds. By the end of germ band retraction, expression is prevalent in the epidermis and trachea (Fig. 3K–L) of both species.

Further functional conservation of Yan is suggested by the expression pattern in third instar larval eye/antennal discs where Yan protein is detected in most cells in the morphogenetic furrow and in the basal nuclei behind the furrow (Fig. 4D–F). In D. melanogaster this expression pattern is well established (Lai and Rubin 1992) and supports the idea that Yan plays a critical role in the control of cell proliferation and differentiation. In addition to the expression in the eye disc, Yan protein is found in the laminar precursor cells (LPCs) in the optic lobes of third instar larval brains from both species (Fig. 4B–C). To date, a role for Yan in larval brain development has...
the level of Yan protein could be affected by RTK signaling in these cells.

Discussion

The structure of Yan is highly conserved between D. melanogaster and D. virilis

To facilitate functional analysis of Yan and begin characterization of molecular mechanisms that specify yan expression in D. melanogaster, we have studied the yan homologue from a distantly related species, D. virilis. With approximately 60 million years separating these two species (Beverley and Wilson 1984), one would expect only the functionally important regions of a gene to be conserved. The D. virilis and D. melanogaster Yan proteins share an overall 72% identity, which is similar to other published inter-species comparisons (Heberlein and Rubin 1990; Michael et al. 1990; Pepling and Gergen 1995; Tillib et al. 1995). Homology in the subdomains of the proteins, however, is more telling with regard to functionally important regions. Namely, the PNT and ETS domains show a very high degree of identity clearly suggesting that these domains are indispensable for the function of Yan protein.

Areas of high homology outside the ETS and PNT domain are also indicative of functionally important sequences within the Yan protein. The MAP kinase phosphorylation sites, which are important for integration of RTK signaling (Rebay and Rubin 1995), are conserved in D. virilis. This suggests that phosphorylation by MAP kinase may represent a general mechanism for regulating Yan activity in other organisms. PEST sequences and glutamine rich repeats are much less well defined with respect to their functional importance. Certainly the role of PEST signals in promoting degradation of proteins has been well established (reviewed by Rechsteiner and Rogers 1996). The most highly conserved PEST site in both species overlaps the third MAP kinase phosphorylation site in the Yan protein. This is noteworthy because the first three MAP kinase phosphorylation sites are required for the inactivation and degradation of Yan in response to activated RTK signaling. Furthermore, it has been proposed that phosphorylation at or near PEST signals can be a conditional signal for selective degradation (reviewed by Rechsteiner and Rogers 1996). The potential role for glutamine rich repeats within the Yan protein has not been addressed, but their conservation in these evolutionarily divergent species does suggest that there is selection pressure for maintaining them in the molecule. Structurally, there is a fairly high degree of variation in these repeats between the two species. This may not affect the function of this domain, however, since the glutamine rich repeat region from the D. virilis elav gene can substitute for the D. melanogaster repeats despite differences in their primary structure (Yao and White 1991).

Our phylogenetic analysis of ETS genes underscores the importance of identifying yan homologues. In a com-

---

Fig. 5A–D Yan is expressed in ovaries from both species. Ovaries from D. melanogaster (A) and D. virilis (B) stained with anti-Yan antibodies reveal Yan protein in the nuclei of the follicle cells and at low levels in the nurse cells. C β-galactosidase activity in ovaries from yanP flies is detected in the follicle cells with higher levels at the ends of each egg chamber. D yanP expression in the border cells can be seen (arrow) in the middle focal plain of a late stage egg chamber. In all cases anterior is to the right.

not been demonstrated, but yan is required for development of the embryonic brain (Rogge et al. 1995; Dumstrei et al. 1998).

Finally, Yan expression is detected in follicle cells in the ovary from both D. melanogaster and D. virilis (Fig. 5A, B). We also observed Yan immunostaining in nurse cell nuclei, albeit less dramatic than follicle cell expression. The pattern of yan transcription, as detected in the yanP enhancer trap line (Fig. 5C–D), displays an interesting difference when compared with the antibody staining. Whereas Yan protein appears evenly distributed along the anterior posterior axis of each ovariole, there is an apparent gradient of β-galactosidase activity emanating from the ends of each egg chamber. That this pattern is not seen with α-Yan antibodies suggests that there may be only subtle differences in the overall transcription rate, but because of the difference in stability between Yan and β-galactosidase these subtleties are exaggerated in yanP ovaries. A similar situation has recently been shown for the fushi tarazu (ftz) gene in D. melanogaster where it had previously been thought that Ftz was expressed in a gradient within each stripe based on reporter gene expression. Upon quantitative analysis of Ftz protein, however, this was shown not to be the case (Lawrence and Pick 1998). It is also possible that an alternative isoform of Yan which is not recognized by the anti-Yan antibody is produced in follicle cells or that yanP transcription in these cells is influenced by an element outside of the yan locus. A second difference in the pattern of transcription and Yan protein distribution is that yanP expression is seen in the border cells at the anterior of the growing oocyte (Fig. 5D) whereas Yan protein is not detected in these cells. Ras signaling is important in these cells to mediate their migration to the proper position in the egg chamber (Lee et al. 1996); again,
parison of 87 known ETS sequences, Yan had no close relatives. The only gene segregating with yan in this analysis was the human tel gene. This in itself is very interesting given the role of Tel in hematopoiesis and leukemia. The tel locus was first identified as the site of chromosomal rearrangements associated with leukemias of myeloid and lymphoid origins (Golub et al. 1994; Stegmaier et al. 1996; Poirel et al. 1998). The normal function of tel appears to be in the establishment and maintenance of hematopoietic precursors within the bone marrow (Wang et al. 1998) which suggests that despite primary structural differences, yan and tel might be functionally related.

A similar relationship can be inferred between yan and the Caenorhabditis elegans lin-1 gene. Lin-1 is an ETS domain protein that acts as an inhibitor of vulval cell fate (Beitel et al. 1995). Like Yan, Lin-1 activity is down-regulated by MAP kinase phosphorylation (Jacobs et al. 1998; Tan et al. 1998). Despite very little overall structural similarity between Yan and Lin-1, they both contain a conserved FQFP motif in their C-termini. This motif is the site of several gain-of-function mutations in lin-1 (Jacobs et al. 1998) which raises the possibility that this conserved motif is indispensable for the function of these two proteins.

Transcriptional repression often involves the association of a transcription factor with a co-repressor protein. One such molecule, Groucho, is encoded at the Enhancer of split (E(spl)) locus in D. melanogaster. Groucho acts as a co-repressor with a number of transcription factor families including Hairy-related proteins, Runt-related proteins, Dorsal, Engrailed, and other members of the E(spl) complex (reviewed by Fisher and Caudy 1998). It is not known whether Yan requires a co-repressor protein to repress transcription, but conservation of the C-terminal P-DLS sequence suggests that this may be the case. P-DLS is the core sequence required for association of the dCtBP co-repressor with a number of transcriptional repressors of the Hairy/E(spl) basic helix-loop-helix family (Poortinga et al. 1998; Nibu et al. 1998). Although Yan is structurally different from this class of proteins, the existence and conservation of this motif argues that Yan may require co-repressor activity to function.

The expression of Yan suggests that its function is required in many tissues throughout development

Yan expression in a variety of tissues and stages throughout development, and the similarity in expression pattern between these two divergent species suggests that multiple and varied roles for Yan have been conserved in evolution. For instance, Yan expression in the presumptive ectoderm suggests that its role in controlling ventral cell fate has been conserved. Similarly, control of brain development requires that yan be expressed in the presumptive head during embryogenesis (Rogge et al. 1995; Dumstrei et al. 1998). There are, however, other developmental contexts where yan is expressed for which there are no known roles, namely, in the syncitial blastoderm and in the presumptive mesoderm. Although the yan loss-of-function phenotype in the mesoderm is not clear, over-expression of activated Yan causes a dramatic reduction in expression of mesodermal markers (Rebay and Rubin 1995; Shishido et al. 1997), indicating that Yan may play a role in regulating mesodermal differentiation. This activity in the mesoderm may be modulated through signaling from another receptor tyrosine kinase, the FGF receptor encoded by heartless (Shishido et al. 1997). Conservation of Yan expression in the mesoderm of both D. virilis and D. melanogaster suggests that its prospective role in this tissue may be widely conserved within invertebrates. Furthermore, the requirement for Yan function in different tissues and developmental stages underscores the importance of a strict regulatory hierarchy controlling its expression throughout development.

The role of Yan in regulating proliferation and differentiation in the developing eye is well established (Lai and Rubin 1992; O’Neill et al. 1994; Rebay and Rubin 1995; Rogge et al. 1995). In D. virilis, Yan undoubtedly serves the same function in the eye disc. It is less clear what the function of Yan is in the developing optic lobes. At this stage, laminar development is known to require signaling from incoming retinal axons (Huang and Kunes 1996, 1998) and activation of the Drosophila EGF receptor by the activating ligand Spitz (reviewed by Schweitzer and Shilo 1997). Yan may be involved in the integration of EGF receptor signaling in this tissue as in the developing eye. Another possibility is that Yan could play an antiproliferative role in these cells. LPCs are surrounded by two proliferative zones, the inner proliferative zone and outer proliferative zone, but they themselves only undergo limited proliferation (Richardson et al. 1995). In ways LPCs are analogous to the undifferentiated cells in the developing eye disc: that is they remain quiescent until an inductive signal stimulates their differentiation. Thus, the presence of Yan in the LPCs may point to a role very similar to that in the developing eye where Yan regulates both cell proliferation and differentiation.

Expression of Yan in ovaries is interesting for several reasons. The ovary represents a dynamic tissue in which RTK signaling is critical for proper differentiation of different cell types (Murphy et al. 1995; Sapir and Shilo 1998). In this context, models for yan function can be envisioned wherein Yan could serve to integrate these signaling events, thus coordinating cell fate decisions. Although the exact role of yan during oogenesis is not clear, there are no detectable maternal effects in yan mutants. Therefore, it is likely that Yan function in oogenesis involves integration of RTK signaling like in other tissues (Rogge et al. 1995; Gabay et al. 1996; O’Keefe et al. 1997; Hacohen et al. 1998).

In summary, the results presented here address two major issues with respect to the conservation, function, and regulation of the yan gene. First of all, the high degree of conservation in the Yan polypeptide sequence in-
icates that there are strong functional constraints on the majority of the protein. Certainly the ETS and PNT domains are critical for Yan function given their high degree of identity. Further functional importance can be inferred from preservation of the MAP kinase phosphorylation sites, PEST signals, poly-glutamate regions, and other putative functional motifs. The common expression pattern of Yan in these two species suggests that the mechanisms regulating yan expression are likely to be conserved as well. The similar expression pattern between the two species and the growing evidence linking Yan function to RTK signaling suggests that Yan may act as a general inhibitor of RTK signaling in multiple tissues throughout development.

Acknowledgements We thank J. Tamkun, I. Rebay, E.M. O’Neill, and G.M. Rubin and the Bloomington Stock Center for reagents which facilitated this work. Sequence analysis of the D. melanogaster yan gene was done by Z.-C. L. in G. M. Rubin’s laboratory at U.C. Berkeley. The initial molecular evolutionary analysis of ETS genes was performed by Kristi Balczerek and Sudhir Kumar. We are grateful to Drs. D. Nguyen, R. Ordway, and E. Siegfried for comments on this manuscript; also to members of the Lai laboratory for technical assistance and helpful discussions throughout this work. This work was supported by an NSF grant IBN9511201 to Z.-C.L.

References

Kumar S, Tamura K, Nei M (1993) MEGA: Molecular Evolutionary Genetic Analysis. Pennsylvania State University, University Park, PA


