# Mutations Affecting the Development of the Peripheral Nervous System in Drosophila: A Molecular Screen for Novel Proteins

**Sergei N. Prokopenko,* Yuchun He,† Yue Lu† and Hugo J. Bellen*,†**

**Program in Developmental Biology, †Howard Hughes Medical Institute and Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, Texas 77030*

Manuscript received May 5, 2000
Accepted for publication July 27, 2000

ABSTRACT

In our quest for novel genes required for the development of the embryonic peripheral nervous system (PNS), we have performed three genetic screens using MAb 22C10 as a marker of terminally differentiated neurons. A total of 66 essential genes required for normal PNS development were identified, including 49 novel genes. To obtain information about the molecular nature of these genes, we decided to complement our genetic screens with a molecular screen. From transposon-tagged mutations identified on the basis of their phenotype in the PNS we selected 31 *P*-element strains representing 26 complementation groups on the second and third chromosomes to clone and sequence the corresponding genes. We used plasmid rescue to isolate and sequence 51 genomic fragments flanking the sites of these *P*-element insertions. Database searches using sequences derived from the ends of plasmid rescues allowed us to assign genes to one of four classes: (1) previously characterized genes (11), (2) first mutations in cloned genes (1), (3) *P*-element insertions in genes that were identified, but not characterized molecularly (1), and (4) novel genes (13). Here, we report the cloning, sequence, Northern analysis, and the embryonic expression pattern of candidate cDNAs for 10 genes: *astray*, *chrowded*, *dalmatian*, *gluon*, *hoi-polloi*, *melted*, *pebble*, *skittles*, *sticky ch1*, and *vegetable*. This study allows us to draw conclusions about the identity of proteins required for the development of the nervous system in Drosophila and provides an example of a molecular approach to characterize *en masse* transposon-tagged mutations identified in genetic screens.

THE peripheral nervous system (PNS) of Drosophila has been long used as an experimental paradigm to identify new genes and to further our understanding of the molecular mechanisms of neurogenesis (Modolell 1997; Dambly-Chaudiere and Vervoort 1998; Jan and Jan 1998). Most known genes that affect PNS development were isolated serendipitously, since they affect easily identifiable morphological markers, namely bristle number in adults (Jan and Jan 1993). These genes are often nonessential or correspond to partial loss- or gain-of-function mutations in essential genes. Other players that are required for PNS development are essential genes that were isolated because loss of one gene copy causes a visible, but often unrelated haploinsufficient phenotype (*e.g.*, *Notch*, *Delta*, and *Enhancer of split*; Lindsley and Zimm 1992). Later, these genes were shown to affect embryonic neurogenesis when homozygous, and the functional analysis that followed their initial characterization gradually integrated them into developmental pathways of neurogenesis (Jan and Jan 1993).

A subset of PNS genes that remained largely unidentified until the late 1980s corresponds to those essential genes that do not cause a haploinsufficient phenotype when mutated. These genes were identified in genetic screens designed to isolate mutations that cause aberrant development of the embryonic PNS (Salzberg *et al.* 1994; Kolodziej *et al.* 1995; Gao *et al.* 1999). Effects of mutations in these genes are typically pleiotropic and do not affect PNS development only. Two classical examples include the *daughterless* (Caudy *et al.* 1988) and *numb* (Uemura *et al.* 1989) genes. However, because screening by immunohistochemical staining of fixed whole-mount embryos with monoclonal antibodies is quite tedious (Jan and Jan 1993), no such systematic screens were performed prior to 1992.

We set out to screen for genes that are essential and affect PNS development in embryos using chemical agents (Salzberg *et al.* 1994) and *P* elements (Kania *et al.* 1995; Salzberg *et al.* 1997) as mutagens. Out of a total of 66 genes that affect PNS development, many genes were mapped and some were shown to be allelic to previously characterized genes on the basis of mapping information, similarity of phenotype, and complementation tests. However, numerous genes identified in *P*-element screens did not seem to correspond to known genes. The most direct approach to determine if these mutations correspond to novel genes and to establish the nature of mutations is to clone the genes adjacent to the *P*-element insertions. We selected 26 genes for cloning on the basis of several criteria. Combination of

*Corresponding author:* Sergei N. Prokopenko, The Salk Institute for Biological Studies, MNL-T, P.O. Box 85800, San Diego, CA 92186-5800. E-mail: prokopenko@salk.edu

**TABLE 1**

**Deficiencies and mutations used in this study**

| Deficiency | Cytology | | Reference |
|---|---|---|---|
| *Df(2L)s1402* | 30C1-2; 30F | | FlyBase (2000) |
| *Df(2R)H3C1* | 43F; 44D3-8 | | Zhang *et al.* (1996) |
| *Df(2R)H3D3* | 44D1-4; 44F4-5 | | Zhang *et al.* (1996) |
| *Df(2R)H3E1* | 44D1-4; 44F12 | | Zhang *et al.* (1996) |
| *Df(2R)Np3* | 44D2-E1; 45B8-C1 | | Konev *et al.* (1991) |
| *Df(2R)stil-A* | | | Pennetta and Pauli (1997) |
| Mutation | Cytological location | GenBank accession no.[a] | Reference |
| *EP(2)2116* | | AQ072916 | Rørth *et al.* (1998); E. J. Rehm and G. M. Rubin (unpublished results) |
| *EP(2)2257* | | AQ073437 | Rørth *et al.* (1998); E. J. Rehm and G. M. Rubin (unpublished results) |
| *EP(3)1019* | | AQ025152 | Rørth *et al.* (1998); E. J. Rehm and G. M. Rubin (unpublished results) |
| *EP(3)3110* | | AQ025288 | Rørth *et al.* (1998); E. J. Rehm and G. M. Rubin (unpublished results) |
| *EP(3)3377* | | AQ073956 | Rørth *et al.* (1998); E. J. Rehm and G. M. Rubin (unpublished results) |
| $spen^{1}$ | 21B2-4 | | Kolodziej *et al.* (1995) |
| $cyr^{k01021}$ | 29E5-6 | | Török *et al.* (1993); Kania *et al.* (1995) |
| *l(2)06825* | 29F1-2 | AQ025659 | Spradling *et al.* (1999) |
| $Pka\text{-}C1^{01272}$ | 30C1-2 | | Spradling *et al.* (1995) |
| *l(2)k09010* | 30D3-4 | AQ025879 | Török *et al.* (1993); Spradling *et al.* (1999) |
| $esg^{05729}$ | 35D1-2 | | Spradling *et al.* (1995) |
| $esg^{k08104}$ | 35D1-4 | | Török *et al.* (1993); Kania *et al.* (1995) |
| $CycE^{05206}$ | 35D3-4 | | Spradling *et al.* (1999) |
| $abof^{1}$ | [38B5-58E4] | | Kolodziej *et al.* (1995) |
| $gui^{1}$ | [38B5-58E4] | | Kolodziej *et al.* (1995) |
| $wwy^{1}$ | [38B5-58E4] | | Kolodziej *et al.* (1995) |
| $ptc^{4}$ | 44D2-5 | | Tearle and Nüsslein-Volhard (1987) |
| $ptc^{10}$ | 44D2-5 | | Tearle and Nüsslein-Volhard (1987) |
| $ptc^{11}$ | 44D2-5 | | Tearle and Nüsslein-Volhard (1987) |
| $ptc^{16}$ | 44D2-5 | | Tearle and Nüsslein-Volhard (1987) |
| *l(2)k04913* | 44F1-2 | AQ025750 | Török *et al.* (1993); Spradling *et al.* (1999) |
| $lola^{5D2}$ | 47A11-13 | | Giniger *et al.* (1994) |
| $lola^{A307}$ | 47A11-13 | | K. Madden and E. Giniger (personal communication) |
| $fra^{3}$ | 49A10-B2 | | Kolodziej *et al.* (1996) |
| $Sin3A^{ex4}$ | 49B2-3 | | Pennetta and Pauli (1998) |
| $Sin3A^{k05415}$ | 49B3-4 | | Török *et al.* (1993); Spradling *et al.* (1995); Pennetta and Pauli (1997, 1998) |
| $Sin3A^{k09715}$ | 49B3-4 | | Török *et al.* (1993); Spradling *et al.* (1995); Pennetta and Pauli (1998) |
| $Sin3A^{k11222}$ | 49B3-4 | | Török *et al.* (1993); Pennetta and Pauli (1998) |
| $Sin3A^{08269}$ | 49B3-6 | | Neufeld *et al.* (1998); Spradling *et al.* (1999) |
| $sca^{1}$ | 49C2-4 | | Lindsley and Zimm (1992) |
| *l(2)k07127* | 53C1-2 | AQ025818 | Török *et al.* (1993); Sprading *et al.* (1999) |
| *l(2)k03609* | 53C1-4 | AQ025726 | Török *et al.* (1993); Spradling *et al.* (1999) |
| $Sema\text{-}2A^{03021}$ | 53C6-10 | | Kolodkin *et al.* (1993); Spradling *et al.* (1999) |
| $Sema\text{-}2A^{k11240}$ | 53C6-10 | | Török *et al.* (1993); Spradling *et al.* (1995); Roch *et al.* (1998) |
| $mbl^{E27}$ | 54B1-2 | | Begemann *et al.* (1997) |
| $mbl^{E127}$ | 54B1-2 | | Begemann *et al.* (1997) |
| $mbl^{k05507}$ | 54B1-2 | | Török *et al.* (1993); Begemann *et al.* (1997) |
| $thr^{1}$ | 54F5-55A1 | | Nüsslein-Volhard *et al.* (1984) |
| $insc^{P8}$ | 57B | | Burchard *et al.* (1995) |

(*continued*)

**TABLE 1**

**(Continued)**

| Mutation | Cytological location | GenBank accession no.[a] | Reference |
|---|---|---|---|
| *insc*$^{P49}$ | 57B1 | | Kraut and Campos-Ortega (1996) |
| *l(2)05475* | 57B2-3 | G00585 | Spradling *et al.* (1999) |
| *insc*$^{8}$ | 57B4-5 | | Burchard *et al.* (1995) |
| *sktl*$^{\Delta 5}$ | 57B4-5 | | Hassan *et al.* (1995) |
| *sktl*$^{\Delta 15}$ | 57B4-5 | | Hassan *et al.* (1998) |
| *EP(3)3415* | 66A20-22[b] | AQ254711 | Rørth *et al.* (1998); E. J. Rehm and G. M. Rubin (unpublished results) |
| *l(3)09645* | 66B1-5[b] | AQ073344 | Spradling *et al.* (1999) |
| *puc*$^{1}$ | 84E10-13 | | Baker *et al.* (1991) |
| *puc*$^{A251.1F3}$ | 84E10-13 | | Martín-Blanco *et al.* (1998) |
| *Scm*$^{D1}$ | 85E1-2 | | Breen and Duncan (1986) |
| *dmt*$^{1184}$ | 85E5-11 | | Salzberg *et al.* (1994) |
| *dmt*$^{M572}$ | 85E5-11 | | Salzberg *et al.* (1994) |
| *EP(3)0359* | 86B[c] | AQ025068 | Rørth *et al.* (1998); E. J. Rehm and G. M. Rubin (unpublished results) |
| *l(3)06142* | 86B1-2 | AQ026212 | Spradling *et al.* (1999) |
| *l(3)j8B6* | 86B1-2 | AQ026332 | Spradling *et al.* (1999) |
| *stg*$^{\eta 80}$ | 99A | | Berg and Spradling (1991) |
| *stg*$^{01235}$ | 99A5-6 | G00587 | Spradling *et al.* (1999) |
| *stg*$^{4}$ | 99A5-7 | | Jürgens *et al.* (1984) |
| *veg*$^{Z322}$ | | | C. Russell and G. Tear (personal communication) |

[a] The GenBank accession number is given for those *P* and *EP* lines that were identified from BLAST searches as presumably allelic to *P*-element lines (Table 2) used in the screen.

[b] Prokopenko *et al.* (1999).

[c] This study.

plasmid rescue, sequencing, and cDNA cloning gave us molecular information about the mapping position of *P* elements, their physical locations in the genome, the nature of mutations, the possible identity of encoded proteins, etc. Here, we report on the nature of these mutations and their adjacent genes. We demonstrate that 11 mutations correspond to known genes and report the cloning, sequence, and analysis of expression in the embryo of 10 novel genes.

## MATERIALS AND METHODS

**Stocks:** All stocks were maintained on a standard corn meal/agar medium (Ashburner 1989) at room temperature. *P{lacZ,w*$^{+}$*}* *P*-element insertion lines used in this study derive from Istvan Kiss' collection of *P* elements on the second chromosome (Török *et al.* 1993) and from Peter Deák's collection on the third chromosome (Salzberg *et al.* 1997), and were shown to be associated with phenotypes in the embryonic PNS (Kania *et al.* 1995; Salzberg *et al.* 1997). *P*-element insertion lines used in the screen are listed in Table 2.

Deficiencies and mutations used for complementation tests are listed in Table 1. Deficiencies and mutations were obtained from the Bloomington Stock Center, the Berkeley Drosophila Genome Project, and individual laboratories. Genetic nomenclature, gene names, and cytology are according to Lindsley and Zimm (1992) and FlyBase (flybase.bio.indiana.edu; FlyBase Consortium 1999).

***In situ* hybridization to polytene chromosomes:** Digoxigenin-labeled DNA probes were prepared using the DIG DNA labeling kit (Roche Molecular Biochemicals). Pretreatment and hybridization to polytene chromosomes were essentially as described (Langer-Safer *et al.* 1982). Following hybridization, probes were detected using an anti-digoxigenin antibody conjugated to alkaline phosphatase (Fab fragments, 1:200; Roche Molecular Biochemicals) and 4-nitroblue tetrazolium chloride with 5-bromo-4-chloro-3-indolyl-phosphate (Roche Molecular Biochemicals). The chromosomes were counterstained with Giemsa (Sigma, St. Louis) and mounted in Permount mounting medium (Fisher Scientific, Pittsburgh, PA).

***In situ* hybridization to whole-mount embryos:** *In situ* hybridization to whole-mount Canton-S embryos was carried out as described (Tautz and Pfeifle 1989) using digoxigenin-labeled antisense riboprobes (DIG RNA labeling kit; Roche Molecular Biochemicals). To generate riboprobes by run-off transcription, the following combinations of restriction enzymes (to linearize the template plasmid DNA) and RNA polymerases were used: *aay* antisense probe (5B cDNA, *Xho*I, T3 polymerase), *aay* sense (5B cDNA, *Bam*HI, T7 polymerase), *dmt* antisense (16A cDNA, *Xho*I, T3 polymerase), *dmt* sense (16A cDNA, *Xba*I, T7 polymerase), *glu* antisense (glu11 cDNA, *Not*I, T3 polymerase), *glu* sense (glu11 cDNA, *Hin*dIII, T7 polymerase), *melt* antisense (8G cDNA, *Xba*I, T7 polymerase or HL03627 cDNA, *Not*I, T7 polymerase), *melt* sense (8G cDNA, *Eco*RV, T3 polymerase or HL03627 cDNA, *Xho*I, T3 polymerase), *stich1* antisense (GM05287 cDNA, *Xba*I, T7 polymerase), and *stich1* sense probe (GM05287 cDNA, *Xho*I, T3 polymerase).

**Molecular biology:** Genomic DNA isolation from Canton-S flies, poly(A)$^{+}$ RNA isolation from 0- to 20-hr-old Canton-S embryos, Southern and Northern analyses, and screening of cDNA libraries were performed according to standard protocols (Sambrook *et al.* 1989).

**Plasmid rescue:** Genomic sequences flanking the sites of *P{lacZ,w$^+$}* *P*-element insertions were isolated by plasmid rescue (Pirrotta 1986) using *Bam*HI, *Xba*I, and *Pst*I (for 5′ sequences) and *Eco*RI and *Sac*II (for 3′ sequences) restriction enzymes and Epicurian coli XL1-Blue supercompetent cells (Stratagene, La Jolla, CA). The typical number of transformant colonies with 3 μg of starting genomic DNA and one-third of a ligation reaction used for transformation ranged from 1 to 50.

Several tests were performed on each plasmid rescued genomic fragment to determine if they correspond to novel genes and if they can be used as probes to screen cDNA libraries to clone the corresponding genes. They were (1) checked molecularly by restriction analysis (Table 3), (2) checked cytologically by *in situ* hybridization to polytene chromosomes (data not shown), (3) analyzed by sequencing (Table 3), and (4) checked on a Southern of Canton-S genomic DNA for the absence of repetitive DNA (data not shown).

For each rescue, at least three colonies were checked by DNA miniprep and restriction analyses. In rare cases, when all three colonies exhibited different digestion patterns, three more colonies were analyzed. The lengths of isolated genomic fragments ranged from 150 bp to 15 kb (see Table 3). In some cases we found upon double digestion (using as a second enzyme *Xba*I for *Eco*RI and *Sac*II rescues and *Hin*dIII for *Bam*HI, *Pst*I, and *Xba*I rescues) that a plasmid did not carry a fragment corresponding to a *P*-element backbone (~2 kb for *Eco*RI and *Sac*II rescues and ~10 kb for all other rescues). The presence of a new band that had a size larger or smaller than expected suggested that there were rearrangements of genomic DNA associated with the *P*-element insertion.

Cytological location of each fragment was verified by *in situ* hybridization to polytene chromosomes. If a mapping position of a fragment did not correspond to the mapping position of a *P*-element line used for plasmid rescue, it was discarded.

Based on digestion pattern, a representative plasmid was chosen for sequencing. A single sequencing run was performed (see below). The sequences were used to perform BLAST searches against nucleotide and protein sequence databases. The sequence information from plasmid rescues also provided an independent verification of mapping positions of genomic fragments. If a mapping position of a genomic clone (cosmid, bacterial artificial chromosome, or P1) hit by plasmid rescue-derived sequence was different from a *P*-element mapping position, the plasmid rescue was excluded from further analysis. Genomic fragments listed in Table 3 have mapping positions identical to *P*-element lines from which they were derived. BLAST searches also allowed us to determine the origin of genomic fragments for multiple insertion lines (*e.g.*, *l(2)k00424*). The results of BLAST searches are presented in Table 3.

**cDNA cloning:** Plasmid rescue-derived genomic fragments (both 5′ and 3′, if available) were used to screen cDNA libraries. We used an adult head λEXLX M(−) cDNA library (Bruce A. Hamilton, personal communication) to isolate 31HC and 31HE clones and an embryonic (9–12 hr) λgt11 cDNA library (Zinn *et al.* 1988) to isolate all other cDNA clones (Table 4). For each set of probes, at least 400,000 plaques were screened.

cDNA clones derived from a λgt11 library were subcloned into pBluescript II KS(+) (clone glu11) or pBluescript II SK(+) (all other clones) (Stratagene). cDNA clones derived from a λEXLX library were converted into pEXLX plasmid clones by Cre-*loxP* automatic subcloning *in vivo* (Palazzolo *et al.* 1990) using BM25.8 Cre-expressing strain (Novagen).

For some genes, putative cDNA clones were identified through database searches among Drosophila expressed sequence tags (ESTs; Rubin *et al.* 2000a). These are cloned in pBluescript SK(+/−) (clones GM05287, HL03627, and LD13852; Table 4) or in pOT2a (clones GH03082, GH23250, GM14315, and LD47384).

**Sequencing:** To determine the terminal sequences of plasmid-rescued genomic fragments the following primers were used: P-ele-R (5′-CGACGGGACCACCTTATGTTATTTC-3′) for proximal ends of all rescues; 703 (5′-CGAAAAGTGCCAC CTGACGTC-3′) for distal ends of *Eco*RI and *Sac*II rescues; and 1706 (5′-GCCAGCAACGCAAGCTTCTAG-3′) for distal ends of *Bam*HI, *Xba*I, and *Pst*I rescues. To determine full-length sequence of cDNA clones, we used nested deletions generated with an ExoIII/mung bean nuclease deletion kit (Stratagene) in combination with primer walking. Dye primer and dye terminator sequencing (BigDye cycle sequencing ready reaction kits; PE Applied Biosystems, Foster City, CA) was carried out on an ABI Prism 377 DNA sequencer (PE Applied Biosystems). Nucleotide sequences were assembled using an AutoAssembler (PE Applied Biosystems). All sequences were annotated and deposited in GenBank prior to the end of 1999 (see Tables 3 and 4 for accession numbers).

**Biomolecular search and analysis tools:** Sequence similarity searches were performed using National Center for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov/BLAST/; Wheeler *et al.* 2000) and Berkeley Drosophila Genome Project (BDGP) (http://www.fruitfly.org/blast/) BLAST. Open reading frames (ORFs) were identified with NCBI ORF Finder (http://www.ncbi.nlm.nih.gov/gorf/) and codon preference was determined using SeqWeb (Wisconsin Package, Genetics Computer Group). To identify protein domains in predicted amino acid sequences, we used Motif (GenomeNet, Institute for Chemical Research, Kyoto University, Japan; http://www.motif.genome.ad.jp), SMART (EMBL, Heidelberg, Germany; http://smart.embl-heidelberg.de; Schultz *et al.* 1998, 2000), and ProfileScan (Swiss Institute for Experimental Cancer Research, Lausanne, Switzerland; http://www.isrec.isb-sib.ch/software/PFSCAN_form.html). Other biomolecular tools used were COILS (European Molecular Biology network—Swiss node, http://www.ch.embnet.org/software/COILS_form.html; Lupas *et al.* 1991) to predict coiled-coil domains, SignalP (Center for Biological Sequence Analysis, The Technical University of Denmark, Lyngby, Denmark; http://www.cbs.dtu.dk/services/SignalP/; Nielsen *et al.* 1997) to predict the presence and location of signal peptide cleavage sites, TMHMM (Center for Biological Sequence Analysis, The Technical University of Denmark, Lyngby, Denmark; http://www.cbs.dtu.dk/services/TMHMM-1.0/; Sonnhammer *et al.* 1998) to predict transmembrane helices in proteins, PSORT (University of Tokyo, Tokyo, Japan, http://psort.nibb.ac.jp; Nakai and Horton 1999) to predict subcellular localization sites of proteins, and PESTfind (Pasteur Institute, Paris, France; http://bioweb.pasteur.fr/seqanal/interfaces/pestfind.html; Rechsteiner and Rogers 1996) to identify PEST regions.

## RESULTS AND DISCUSSION

**Rationale of the molecular screen:** To identify novel proteins required for the development of the peripheral nervous system, we decided to clone the affected genes identified in our forward genetic screens (Kania *et al.* 1995; Salzberg *et al.* 1997). The basis of the molecular screen (Figure 1) is the observation that *P* elements often insert in the 5′ regions of genes (reviewed by Bellen 1999). Therefore, genomic sequences flanking the sites of *P*-element insertions often provide information about the identity of genes affected by *P* elements. Typically, flanking genomic DNA is isolated by plasmid rescue, the 5′- and 3′-ends of rescued fragments are

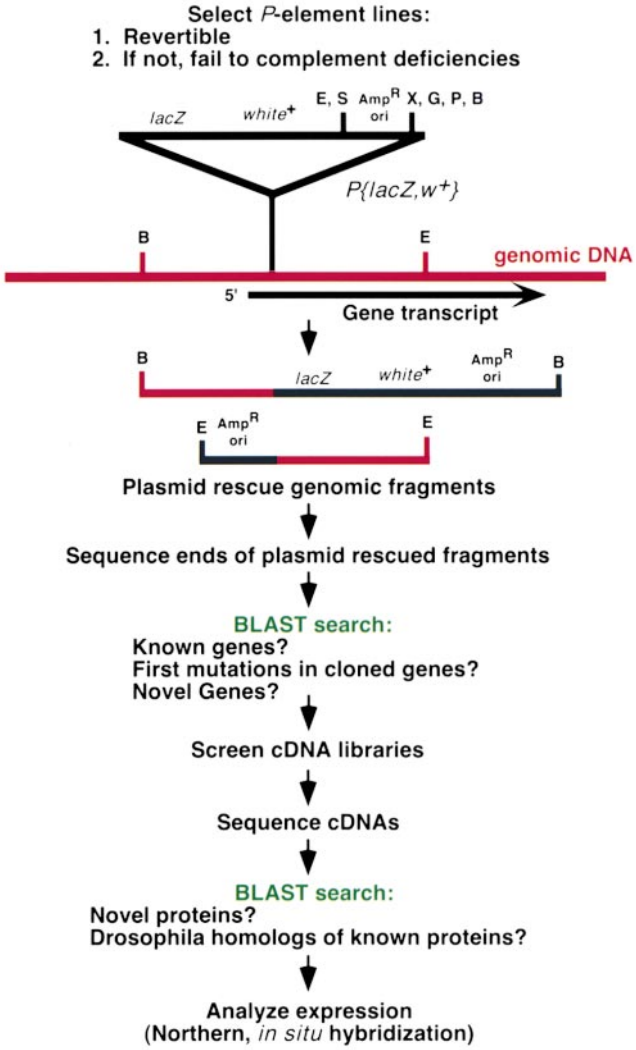


Figure 1.—Experimental design of the molecular screen. A general strategy for plasmid rescue using the *P{lacZ,w*$^{+}$*}* *P* element inserted in the 5′-end of a gene transcription unit is shown. The genomic DNA and rescued genomic fragments are shown in red. For details, see materials and methods and results. B, *Bam*HI; E, *Eco*RI; G, *Bgl*II; P, *Pst*I; S, *Sac*II; X, *Xba*I.

sequenced, and the sequences are used for BLAST searches. Three possible outcomes should be considered. First, *P* elements may affect known, previously identified genes. Indeed, *P*-element insertions are often partial loss-of-function mutations that cause mild phenotypes that are quite different from the phenotypes associated with severe loss-of-function alleles (*e.g.*, *emc*$^{S009426}$ and other *emc* alleles, *pbl*$^{S054203}$; Salzberg *et al.* 1997). Hence, relying on a similarity of phenotype with genes mapped to the region where the *P* element maps may not permit making an educated guess as to the identity of a gene. Thus, the molecular information obtained through plasmid rescue should greatly assist in the identification of affected genes. Second, BLAST searches may identify *P*-element insertions in genes that were cloned, but for which no mutations are available. Finally, *P* elements for which no significant matches were found in BLAST searches are good candidates for mutations affecting novel genes. Each presumably novel gene/mutation is then characterized using a combination of genetic (complementation tests), cytological (comparison of mapping positions of *P* elements, plasmid rescued genomic fragments, and candidate allelic genes), and molecular (positioning *P*-element insertions on the genomic sequence relative to neighboring genes, both known and predicted) approaches.

**Selection of *P*-element lines for the screen:** The screen is based on the assumption that the lethality caused by the insertion of a *P* element maps to the same molecular and cytological region as the *P* element itself. In other words, the *P*-element insertion itself, but not some other mutation on the chromosome, causes the observed phenotype. However, previous experience with *P*-element screens demonstrates that this is not always the case. *P*-element-encoded transposase is a mutagen (Beall and Rio 1997) that may cause chromosomal breaks resulting in inversions, deletions, and other rearrangements elsewhere on a chromosome. In addition, Δ2-3 transposase activity may result in the introduction of multiple *P* elements on a chromosome. Indeed, the screen efficiency defined as the percentage of raw lines that contain a single insertion causing its associated phenotype can be as low as 31% (*P*-element lines from Kiss' collection; Spradling *et al.* 1999). Such considerations precluded us from selecting many lines for the molecular screen (see below).

We, therefore, established the following criteria to select *P*-element lines for the screen. First, *P*-element lines have to be single insertion lines and revertible. Second, if the *P* elements are not revertible, they should fail to complement deficiencies that on the basis of their breakpoints should lack the affected gene. Alternatively, they should fail to complement other independently isolated alleles from the same complementation group that map to the same cytological position as the *P* element. Third, occasionally a line with multiple insertions can be used to clone the gene. However, this was done only if we were able to demonstrate that only one *P*-element insertion is responsible for the lethality and phenotype.

A total of 72 novel *P*-element mutations representing 44 complementation groups were identified in our genetic screens (Kania *et al.* 1995; Salzberg *et al.* 1997). All these genes are essential, since *P*-element insertions cause embryonic, larval, or pupal lethality. Some of the novel genes have been already characterized by others and, hence, are not discussed here—*barren* (*barr*, Bhat *et al.* 1996), *benchwarmer* (*bnch*; A. Kania and H. J. Bellen, unpublished results), *bunched* (*bun*; *short sighted*, *shs*; Treisman *et al.* 1995; Dobens *et al.* 1997, 2000), *gutfeeling* (*guf*, Salzberg *et al.* 1996; Ivanov *et al.* 1998), *homothorax* (*hth*; *dorsotonals*, *dtl*; Rieckhof *et al.* 1997; Pai *et al.*

**TABLE 2**

***P*-element insertion lines used in the screen**

| *P*-element strain | Cytology[a] | Gene[b] | Reversion[c] | MAb 22C10 phenotype in the PNS[c] | Complementation data | Reference on gene cloning |
|---|---|---|---|---|---|---|
| *l(2)k06921* | 21E2-3 | *S* (*fltr*) | Y | Loss of LCh, pathfinding defects | | Kolodkin *et al.* (1994) |
| *l(2)k01021* | 29E5-6 | *raw* (*cyr*) | N | Elongated, darkly stained neurons, thick axons, mild pathfinding defects | $raw^{JG-d}$ | Byars *et al.* (1999) |
| | | | | | $raw^{11F-d}$, $raw^{2418-d}$, $raw^{6825-d}$, *l(2)06825*− | This study |
| *l(2)k08801* | 29E5-6 | *raw* (*cyr*) | N[e] | | $cyr^{k01021}$−, *l(2)06825*− | |
| *l(2)k07104* | 30C1-2 | *hoip* | N | Mild organization and fasciculation defects | *Pka-C1*$^{01272}$+ | This study |
| *l(2)k05002* | 35D1-2, (44D5-6) | *CycE* (*fond*)[e] | N | Fused, elongated LCh5 | $CycE^{05206}$+, $esg^{05729}$+, $esg^{k08104}$+, $fond^{k02514}$−, *Df(2R)H3D3*+, *Df(2R)H3C1*+, *Df(2R)H3E1*+ | Richardson *et al.* (1993) |
| *l(2)k08104* | 35D1-4, 68C1-2 | *esg* | Y | Crowded, elongated LCh5 | $esg^{05729}$− | Whiteley *et al.* (1992) |
| *l(2)k02514* | 35D3-4 | *CycE* (*fond*)[e] | N | Fused, elongated LCh5 | $CycE^{05206-f}$, $esg^{05729}$+, $esg^{k08104}$+ | Richardson *et al.* (1993) |
| *l(2)k08819* | 36A12-14 | *glu* | N | Organization defects, irregularly shaped clusters | | This study |
| *l(2)k02507* | 44D5-6 | *ptc*(*rubr*)[e] | Y | Loss of neurons, organization and pathfinding defects | $ptc^{H84-g}$, $ptc^{4}$−, $ptc^{10}$−, $ptc^{11}$−, $ptc^{16}$+/−[h] | Hooper and Scott (1989); Nakano *et al.* (1989) |
| *l(2)k00424* | (30D1-2), 44F1-2 | *l(2)k00424* | N | Organization and fasciculation defects, abnormally positioned LCh5 | *l(2)k04913*−, *l(2)k09010*+, *Df(2R)Np3*−, *Df(2L)s1402*+ | |
| *l(2)k09901* | 47A1-47B14 | *btb* | Y | Stalled LCh5, pathfinding and connectivity defects | $lola^{5D2}$+, $lola^{A307}$+ | |
| *l(2)k15001* | 48A1-6 | *rack* | Y | Loss of LCh, elongated LCh | *EP(2)2116*+ | |
| *l(2)k15501* | 49B1-49C4 | *unch* (*Sin3A*?)[e] | N | Stalled LCh5, loss of neurons in the lateral cluster | $abof^{1}$+, $fra^{3}$+, $gui^{1}$+, $sca^{1}$+, $Sin3A^{08269}$+, $Sin3A^{k05415}$+, $Sin3A^{k09715}$+, $Sin3A^{k11222}$+, $Sin3A^{ex4}$+, $spen^{1}$+, $wwy^{1}$+, *Df(2R)stil-A*+ | Neufeld *et al.* (1998); Pennetta and Pauli (1998) |
| *l(2)k03402* | 53C1-4 | *veg* | N | Loss of neurons, pathfinding defects | $veg^{Z322}$−, *Sema-2a*$^{03021}$+, *Sema-2a*$^{k11240}$+, $wwy^{1}$+, *l(2)k03609*+, *l(2)k07127*+, *EP(2)2257*+ | |
| *l(2)k07202* | 53E1-2 | *veg* | Y[e] | | $veg^{Z322}$−, $abof^{1}$+, $gui^{1}$+, *Sema-2a*$^{k11240}$+, $spen^{1}$+ | This study |
| *l(2)k07228* | 53E1-2 | *veg* | Y | Severe loss of neurons, pathfinding defects | $veg^{Z322}$−, *Sema-2a*$^{k11240}$+, $wwy^{1}$+, *l(2)k03609*+, *l(2)k07127*+, *EP(2)2257*+ | |
| *l(2)k07103* | 54B1-2 | *mm* (*mbl*)[e] | Y | Fused LCh5 neurons, mild loss of neurons, abnormal neuronal morphology | $mbl^{E27-i}$, $mbl^{E127}$+, $mbl^{k05507}$+ | Begemann *et al.* (1997) |

(*continued*)

**TABLE 2**

**(Continued)**

| *P*-element strain | Cytology[a] | Gene[b] | Reversion[c] | MAb 22C10 phenotype in the PNS[c] | Complementation data | Reference on gene cloning |
|---|---|---|---|---|---|---|
| *l(2)k07805b* | 54F5-6 | *thr* (*anch*)[e] | N | Scattered neurons, dark rounded cells | $thr^{1}$− | D'Andrea *et al.* (1993); Philp *et al.* (1993) |
| *l(2)k07505* | 57B1-3 | *sktl* | Y | LCh5 stain faintly with MAb 22C10, elongated | $insc^{8}$−, $insc^{P8}$−, $insc^{P49}$−, *l(2)05475*−[j] | Knirr *et al.* (1997b); Hassan *et al.* (1998); This study |
| | | *insc* | | | | Kraut and Campos-Ortega (1996) |
| *l(2)k06908* | 59E1-2 | *chrw* | Y | Crowding and abnormal appearance of LCh5, thick axons | | This study |
| *l(3)S058701* | 61D | *emc* | Y | 22C10 expression in the gonads | | Ellis *et al.* (1990); Garrell and Modolell (1990) |
| *l(3)S144114* | 65E | *melt* | Y | Abnormal morphology and mild loss of neurons | *EP(3)1019*+, *EP(3)3110*+, *EP(3)3377*+ | This study |
| *l(3)S008320* | 66AB | *pbl* | Y | Severe pathfinding and fasciculation defects, loss of neurons, enlarged neurons | *l(3)09645*−, *EP(3)3415*− | Prokopenko *et al.* (1999); This study |
| *l(3)S054203* | 66AB | *pbl* | Y | Mild fasciculation defects | *l(3)09645*−, *EP(3)3415*− | |
| *l(3)S042314* | 67B1-10 | *aay* | Y | Axonal misrouting, intersegmental nerve crosses segment boundaries | | This study |
| *l(3)S023803* | 84F | *puc* (*hrt*)[e] | Y | 22C10 expression in the heart, defasciculated axonal bundles | $puc^{A251.1F3-}$[k], $puc^{1}$+[k] | Martín-Blanco *et al.* (1998) |
| *l(3)S114307* | 85E | *potp* (*Crc*)[e] | Y | Variable loss of neurons, decreased levels of 22C10 staining, disorganization and pathfinding defects | | Smith (1992) |
| *l(3)S048103* | 85E5–11 | *dmt* | Y | Occasional loss of neurons, pathway defects, and dark rounded cells | $Scm^{D1}$+[l] | This study |
| *l(3)S143702* | 86B[e] | *stich1* | Y[e] | Loss of neurons in v and v′ clusters. LCh5 stain very faintly with MAb 22C10 or do not stain at all | *l(3)06142*+, *l(3)j8B6*+, *EP(3)0359*− | This study |
| *l(3)S024108* | 92E8–14 | *bon* | Y | Extra 22C10-positive cells anterior to each dorsal cluster and posterior to v′ cluster | | R. B. Beckstead, S. N. Prokopenko and H. J. Bellen (unpublished results) |
| *l(3)S024532* | 99A | *stg* | Y | Loss of neurons, abnormal positioning of LCh5, severe disorganization, misrouting of axons | $stg^{4}$−, $stg^{01235}$−, $stg^{y80}$− | Edgar and O'Farrell (1989); Jimenez *et al.* (1990) |

Second and third chromosome strains are listed in the estimated physical order of their *P*-element insertions along the chromosomes.

[a] According to Kania *et al.* (1995), Spradling *et al.* (1995), and Salzberg *et al.* (1997). Mapping positions in parentheses denote *P*-element insertions complementing deficiencies.

[b] Gene name synonyms are in parentheses.

[c] According to Kania *et al.* (1995) and Salzberg *et al.* (1997).

[d] Byars *et al.* (1999).

[e] This study.

[f] Adult escapers are viable, but have rough eyes and wing venation defects.

[g] J. Lim, J. Chern and K.-W. Choi, personal communication.

[h] Partial complementation: 30% of expected transheterozygous progeny emerge. Escapers are viable and fertile.

[i] Adult escapers have wing blisters, wing venation defects, or unexpanded wings.

[j] *P*-element insertion allelic to *sktl*, since it fails to complement $sktl^{\Delta 5}$ and $sktl^{\Delta 15}$ alleles.

[k] Identical results were obtained for $hrt^{H246}$, $hrt^{S070806}$, and $hrt^{S135108}$.

[l] Complementation tests with $dmt^{1184}$ and $dmt^{M572}$.

1998), *pavarotti* (*pav*, Adams *et al.* 1998), *RpL30* (Kania *et al.* 1995), *sanpodo* (*spdo*; Dye *et al.* 1998; Skeath and Doe 1998), *schnurri* (*shn*; *quo vadis*, *quo*; Nüsslein-Volhard *et al.* 1984; Arora *et al.* 1995; Grieder *et al.* 1995; Staehling-Hampton *et al.* 1995), and *senseless* (*sens*, Nolo *et al.* 2000). Among the remaining mutations we found 31 *P*-element strains representing 26 complementation groups (Table 2) that satisfy the required criteria. This number includes four lines selected during an initial stage of the project that were subsequently shown to be allelic to known genes—*Star* (*S*, Kolodkin *et al.* 1994), *escargot* (*esg*, Whiteley *et al.* 1992), *extra macrochaetae* (*emc*, Ellis *et al.* 1990; Garrell and Modolell 1990), and *string* (*stg*, Edgar and O'Farrell 1989; Jimenez *et al.* 1990) (see Table 2 and Salzberg *et al.* 1997). We also selected three lines with multiple insertions—*l(2)k00424*, *l(2)k05002*, and *l(2)k08104*. Each of these lines carries two *P* elements; however, one of them is not responsible for the phenotype, since it either maps to another chromosome or complements deficiencies (Table 2).

The remaining lines described in Kania *et al.* (1995) and Salzberg *et al.* (1997) were not included in the screen, since they did not satisfy the established selection criteria. Some of them correspond to genes identified by a single nonrevertible *P*-element insertion, which complements deficiencies uncovering the region of the insertion [*e.g.*, *l(2)k05422*, *l(2)k06712*, *l(2)k08807*, *l(3)S136802*, $misn^{S081603}$]. In other cases, *P* elements map to cytological intervals with few or no deficiencies [*e.g.*, *l(3)S052908*]. In addition, absence of independent alleles generated in other *P*-element or chemical mutagenesis screens did not allow us to genetically map the genes and therefore we were unable to conclude if the *P*-element insertion is responsible for the phenotype.

**Characterization of *P*-element insertions using flanking genomic DNA sequences:** We used plasmid rescue to recover genomic DNA flanking the insertion sites of 30 *P*-element lines selected for the screen. For many lines (19 out of 30) we were able to recover DNA flanking both 5′- and 3′-ends of *P* elements (Table 3). Analysis of genomic sequences flanking *P*-element insertions provided several types of information. First, a significant sequence match found between a plasmid rescue and cDNA sequence of a known gene demonstrated that this gene is likely to be affected by the *P* element. Second, availability of genomic sequence information allowed us to physically associate *P* elements and their flanking genomic fragments with specific sites in the genome. Furthermore, we were able to precisely position the *P* elements relative to neighboring genes. This information allowed us to make predictions about the identity of genes affected by *P* elements, about allelic relationships between previously characterized mutations and those identified in our screens, and about other genes adjacent to *P* elements and linked molecularly to the gene of interest. Results of plasmid rescue experiments including molecular (analysis by gel electrophoresis), sequence (GenBank accession numbers for sequences of ends of plasmid rescued fragments), database (BLASTN and BLASTX search results), and genomic (prediction of *P*-element locations relative to neighboring genes) analyses are presented in Table 3.

**Four classes of genes identified in the screen:** Sequence information derived from plasmid rescues allowed us to assign all genes to one of four classes: (1) previously characterized genes (11 genes), (2) first mutations in cloned genes (1 gene), (3) *P*-element insertions in genes that were phenotypically characterized, but not identified (1 gene), and (4) novel genes (13 genes).

**Previously characterized genes:** Our initial analysis of mutations relied solely on the molecular information derived from genomic DNA flanking the sites of *P*-element insertions. Using this approach we identified 11 previously characterized genes. We provide both molecular and genetic evidence establishing new allelic relationships between 10 existing mutations (Tables 3 and 5). We found that *Cyclin E* (*CycE*, Richardson *et al.* 1993; Knoblich *et al.* 1994) is allelic to *fondue* (*fond*, Kania *et al.* 1995), *mindmelt* (*mm*, Kania *et al.* 1995) is allelic to *muscleblind* (*mbl*, Begemann *et al.* 1997; Artero *et al.* 1998), *patched* (*ptc*, Sturtevant 1948; Nüsslein-Volhard *et al.* 1984; Hooper and Scott 1989; Nakano *et al.* 1989) is allelic to *rubberneck* (*rubr*, Kania *et al.* 1995), *puckered* (*puc*, Ring and Martinez Arias 1993; Martín-Blanco *et al.* 1998) is allelic to *hearty* (*hrt*, Salzberg *et al.* 1994, 1997), and *three rows* (*thr*, Nüsslein-Volhard *et al.* 1984; D'Andrea *et al.* 1993; Philp *et al.* 1993) is allelic to *anarchist* (*anch*, Kania *et al.* 1995). *CycE*, *mbl*, and *thr* are directly affected by *P*-element insertions $fond^{k02514}$, $mm^{k07103}$, and $anch^{k07805b}$, respectively. In contrast, the $rubr^{k02507}$ *P* element is inserted in the first intron of *ptc* and the $hrt^{S023803}$ *P* element is inserted in the intron between exons 3 and 4 of *puc*. Since these introns do not contain any known or predicted genes (Adams *et al.* 2000), we considered the possibility that *ptc* is allelic to *rubr* and *puc* is allelic to *hrt*. In all five cases we were able to demonstrate by complementation tests that the mutations are indeed allelic (Table 2). For example, $fond^{k05002}$ insertion complements *CycE* mutation, but fails to complement another *P*-element allele, $fond^{k02514}$. $fond^{k02514}/CycE^{05206}$ flies are viable, but adult escapers have rough eyes and wing venation defects indicating that the two mutations are allelic. Similarly, $mm^{k07103}$ complements two hypomorphic alleles of *mbl* (Table 2), but fails to complement $mbl^{E27}$, a putative null allele affecting the coding sequence of the gene. There are few $mm^{k07103}/mbl^{E27}$ adult escapers that have wing blisters, wing venation defects, or unexpanded wings. Similarly, $hrt^{S023803}$ insertion complements $puc^{1}$, but fails to complement $puc^{A251.1F3}$ (Table 2), a *P*-element insertion that affects the same intron as the $hrt^{S023803}$ *P* element. In conclusion, results of complementation tests combined with the se-

**TABLE 3**

**Genomic sequences flanking *P*-element insertions**

| Gene | *P*-element strain | Plasmid rescue[a] | Size, kb[b] | GenBank accession no.[c] | NCBI BLAST search results (match length, *E* value)[d] | Location of *P*-element insertion |
|---|---|---|---|---|---|---|
| *l(2)k00424* | *l(2)k00424* | A21 (30D) | 1.4 | AF175915 | *FKBP59-bp1*,[e] Dm, 30D3-4 | 5′-UTR, 21 nt upstream of *FKBP59-bp1* AUG |
| | | | | (AF175916) | No significant similarities | |
| | | B21 (44F) | 4.2 | AF175917 | No significant similarities | |
| | | | | (AF175918) | No significant similarities | |
| *aay* | *l(3)S042314* | A5 | 4.9 | AF174664 | No significant similarities | 5′-UTR, 61 nt upstream of *aay* AUG |
| | | | | (AF174665) | No significant similarities | |
| *bon* | *l(3)S024108* | A3 | 0.15 (0.7) | AF174658 | No significant similarities | 5′-UTR, 686 nt upstream of *bon* AUG |
| | | B3 | 8.7 | AF174659 | No significant similarities | |
| | | | | (AF174660) | No significant similarities | |
| *btb* | *l(2)k09901* | B18-28 | 0.8 | AF175907 | *CG5380* DNA-directed RNA polymerase III, Dm (26, Xe-06) | 418 nt upstream of *CG5380* AUG |
| | | | | (AF175908) | *CG5380* DNA-directed RNA polymerase III, Dm (112, Xe-59)<br>RNA polymerase III (39 kD), Hs (106, Xe-32)<br>RNA polymerase III subunit, Sp (104, Xe-11) | |
| *chrw* | *l(2)k06908* | A9 | 1.6 | AF174673 | No significant similarities | ~294 nt upstream of the 5′-end of the LD47384 *chrw* cDNA |
| | | | | (AF174674) | No significant similarities | |
| *CycE* (*fond*) | *l(2)k02514* | A27 | 10.0 | AF175936 | No significant similarities | Intron 1 of the type II *Cyclin E* mRNA |
| | | | | (AF175937) | *l(2)35Df* ATP-dependent DNA helicase, Dm, 35D5-6 (122, Xe-62)<br>KIAA0052, Hs (106, Xe-33)<br>Putative helicase W08D2.7, Ce (114, Xe-31)<br>Mtr4p RNA helicase, Sc (89, Xe-26) | 6.7 kb upstream of the 5′-end of the GH07290 *l(2)35Df* cDNA |
| | | B27 | 5.0 | AF175938 | No significant similarities | |
| | | | | (AF175939) | *CycE* type II, Dm, 35D4-6 (32, Ne-08) | |
| *dmt* | *l(3)S048103* | A16 | 7.0 | AF174682 | No significant similarities | ~11 nt upstream of the 5′-end of the 16A *dmt* cDNA |
| | | | | (AF174683) | *hyperplastic disks* (*hyd*), Dm, 85E5-6 (352, Ne-176) | ~3.5 kb downstream of the 3′-end of *hyd* mRNA |
| | | E16 | 5.0 | (AF174687) | No significant similarities | |
| | | P16 | 2.5 | AF174685 | No significant similarities | ~165 nt downstream of the 3′-end of *CG9448* mRNA |
| | | | | (AF174686) | *CG9448* transcription factor, Dm, 85E5-6 (117, Xe-48) | |
| | | S16 | 1.6 | (AF174684) | No significant similarities | |
| *emc* | *l(3)S058701* | A7 | 2.5 | AF174666 | *emc*, Dm, 61D1-2 (355, N0.0) | Exon 1, 5′-UTR, 251 nt upstream of *emc* AUG |
| | | | | (AF174667) | No significant similarities | |
| | | B7 | 2.4 | AF174668 | *emc* exon 1, Dm, 61D1-2 (364, N0.0) | |
| *esg* | *l(2)k08104* | A13 | 3.0 | AF174681 | *esg*, Dm, 35D1 (489, N0.0) | 5′-UTR, ~180 nt upstream of *esg* AUG |
| *glu* | *l(2)k08819* | A24 | 6.5 | AF175924 | XCAP-C, X1 (112, Xe-30), CAP-C, Hs (112, Xe-29), Smc4p, Sc (112, Xe-28), cut3p, Sp (112, Xe-20) | Coding sequence of *glu* |
| | | B24 | 5.0 | AF175925 | XCAP-C, X1 (34, Xe-05), CAP-C, Hs (34, Xe-05) | |

(*continued*)

**TABLE 3**

**(Continued)**

| Gene | *P*-element strain | Plasmid rescue[a] | Size, kb[b] | GenBank accession no.[c] | NCBI BLAST search results (match length, *E* value)[d] | Location of *P*-element insertion |
|---|---|---|---|---|---|---|
| *hoip* | *l(2)k07104* | A22 | 4.5 | AF175919 | Non-histone chromosome protein 2 (NHP2)-like 1, Hs (88, Xe-21) | ~31 nt upstream of the 5′-end of the GH03082 *hoip* cDNA |
| | | | | | NHP2/RS6 family protein YEL026W homolog, Ce (87, Xe-20) | ~2.3 kb upstream of *Pka-C1* AUG |
| | | | | | SNU13 snRNP subunit homolog, Sp (87, Xe-19) | |
| | | | | | Ribosomal protein L7Ae-like, At (92, Xe-18) | |
| | | | | (AF175920) | No significant similarities | |
| | | B22 | 15.0 | AF175921 | *Pka-C1*, Dm, 30C5-7 (143, Ne-65) | |
| *melt* | *l(3)S144114* | A8 | 3.0 | AF174669 | No significant similarities | ~125 nt upstream of the 5′-end of *melt* cDNA |
| | | | | (AF174670) | No significant similarities | ~1 kb downstream of poly(A) signal of *Diazepam-binding inhibitor* (*Dbi*), 65E5-6 |
| | | B8 | 6.0 | AF174671 | No significant similarities | |
| | | | | (AF174672) | No significant similarities | |
| *mm* (*mbl*) | *l(2)k07103* | A33 | 5.5 | AF175949 | *muscleblind A* (*mbl*), Dm, 54B8-16 (507, N0.0) | ~57 nt upstream of the 5′-end of the *mbl A* mRNA |
| | | | | (AF175950) | No signficant similarities | |
| | | B33 | >12.0 | AF175951 | No significant similarities | |
| | | | | (AF175952) | No significant similarities | |
| *pbl* | *l(3)S008320* | A1 | 0.4 (0.8) | AF174650 | No significant similarities | Coding sequence of *pbl*[f] |
| | | B1 | 2.8 | AF174651 | No significant similarities | ~1.4 kb upstream of *CG8281* AUG |
| | | | | (AF174652) | *CG8281*, Dm, 66A18-20 (127, Xe-33) | ~2.5 kb downstream of the 3′-end of *CG8111* mRNA |
| | | | | | *CG8111*, Dm, 66A18-20 (31, Xe-12) | |
| | *l(3)S054203* | A6 | 0.8 | AF174653 | No significant similarities | 5′-UTR, 149 nt upstream of *pbl* AUG[f] |
| | | | | (AF174654) | No significant similarities | |
| *potp* (*Crc*) | *l(3)S114307* | A20-16 | 1.4 | AF175911 | *Calreticulin* (*Crc*), Dm, 86B-C (283, Ne-156) | 5′-UTR, 28 nt upstream of *Crc* AUG |
| | | | | (AF175912) | *Crc*, Dm, 86B-C (277, Ne-152) | |
| | | B20-36 | 1.3 | AF175913 | *Crc*, Dm, 86B-C (467, N0.0) | |
| | | | | (AF175914) | *Crc*, Dm, 86B-C (374, N0.0) | |
| *ptc* (*rubr*) | *l(2)k02507* | A17-5 | 0.22 (0.8) | AF175904 | No significant similarities | ~9.4 kb intron between exons 1 and 2 of *patched* (*ptc*), 44D2-5 |
| | | B17-23 | 5.4 | AF175905 | No significant similarities | |
| | | | | (AF175906) | No significant similarities | |
| *puc* (*hrt*) | *l(3)S023803* | B2 | 3.1 | AF174657 | No significant similarities | ~9.2 kb intron between exons 3 and 4 of *puckered* (*puc*), 84E10-13 |
| *rack* | *l(2)k15001* | A19-13 | 0.27 (1.0) | AF175909 | No significant similarities | |
| *raw* (*cyr*) | *l(2)k01021* | A26 | 2.2 | AF175930 | No significant similarities | ~5 kb upstream of *raw* AUG (GenBank accession no. AF186024) |
| | | | | (AF175931) | No significant similarities | Coding sequence of the alternatively spliced *raw* mRNA (GH23250 cDNA) |

(*continued*)

**TABLE 3**

**(Continued)**

| Gene | *P*-element strain | Plasmid rescue[a] | Size, kb[b] | GenBank accession no.[c] | NCBI BLAST search results (match length, *E* value)[d] | Location of *P*-element insertion |
|---|---|---|---|---|---|---|
| | *l(2)k08801* | A30 | 3.0 | AF175932 | No significant similarities | Intron of *raw*, 29F1-2 |
| | | | | (AF175933) | No significant similarities | |
| | | B30 | 9.5 | AF175934 | No significant similarities | |
| | | | | (AF175935) | No significant similarities | |
| *S* (*fltr*) | *l(2)k06921* | A10 | 12.0 | AF174675 | *Star* (*S*), Dm, 21E2 (265, Ne-130) | ~1060 nt upstream of *S* AUG |
| | | | | | *asteroid* (*ast*), Dm, 21E2 (117, Ne-54) | ~275 nt upstream of *ast* AUG |
| | | B10 | 3.0 | AF174676 | *asteroid* (*ast*), Dm, 21E2 (467, N0.0) | |
| *sktl* (*fam*) | *l(2)k07505* | B12 | 5.5 | AF174679 | No significant similarities | Intron of *sktl*[g] |
| | | | | (AF174680) | No significant similarities | |
| *stg* | *l(3)S024532* | A4 | 1.5 | AF174661 | *stg*, Dm, 99A5-7 (169, Ne-82) | ~440 nt upstream of *stg* AUG |
| | | | | (AF174662) | No significant similarities | |
| | | B4 | 1.2 | AF174663 | No significant similarities | |
| *stich1* | *l(3)S143702* | A32 | 1.3 | AF175945 | *Domina* (*Dom*), Dm, 86A2-4 (353, N0.0) | Intron between exons 1 and 2 of *Dom* |
| | | | | (AF175946) | *Dom*, Dm, 86A2-4 (180, Ne-88) | |
| | | B32 | 8.0 | AF175947 | No significant similarities | ~90 nt upstream of the 5′-end of the GM05287 cDNA |
| | | | | (AF175948) | No significant similarities | |
| *thr* (*anch*) | *l(2)k07805b* | A23 | 2.1 | AF175922 | *thr*, Dm, 54F5-55A1 (445, N0.0) | ~640 nt upstream of *thr* AUG |
| | | B23 | 12.0 | AF175923 | *thr*, Dm, 54F5-55A1 (181, Ne-66) | |
| *unch* | *l(2)k15501* | A25 | 1.3 | AF175926 | No significant similarities | ~0.8 kb upstream of the 5′-end of *amphiphysin* mRNA |
| | | | | (AF175927) | *amphiphysin*, Dm, 49B2-3 (71, Ne-31) | |
| | | B25 | 8.0 | AF175928 | *Sin3A*, Dm, 49B2-3 (358, Ne-176) | ~24 nt upstream of the 5′-end of *Sin3A* mRNA (GenBank accession no. AF024603) |
| | | | | (AF175929) | No significant similarities | |
| *veg* | *l(2)k03402* | A28 | 0.25 (0.6) | AF175940 | No significant similarities | |
| | | B28 | 9.0 | AF175941 | No significant similarities | |
| | | | | (AF175942) | No significant similarities | |
| | *l(2)k07202* | A31 | 4.0 | AF175943 | No significant similarities | ~309 nt upstream of the 5′-end of *veg* cDNA |
| | | | | (AF175944) | No significant similarities | |
| | *l(2)k07228* | B11 | 3.5 | AF174677 | No significant similarities | ~125 kb away from *veg* cDNA |
| | | | | (AF174678) | No significant similarities | |

[a] Plasmid-rescued genomic fragments are named according to the restriction endonuclease used for plasmid rescue: A, *Eco*RI; B, *Bam*HI; E, *Xba*I; P, *Pst*I; S, *Sac*II.

[b] Sizes in parentheses denote an apparent length of genomic fragments as analyzed by gel electrophoresis. Sequence analysis demonstrated that their actual length is shorter, probably due to rearrangement at the site of a *P*-element insertion.

[c] Accession numbers in parentheses correspond to the sequences of distal (relative to the site of a *P*-element insertion) ends of genomic fragments.

[d] BLAST search results are current as of March 2000. BLAST hits with genomic sequences, ESTs, and sequence-tagged sites are excluded from the table. Shown are the results of BLAST searches followed by species name and (for Drosophila proteins) mapping position. The function of predicted Drosophila proteins is cited according to Adams *et al.* (2000). Match length is in nucleotides or amino acids and the corresponding Expect (*E*) value is preceded by N (for BLASTN) or by X (for BLASTX searches), respectively. At, *Arabidopisis thaliana*; Ce, *C. elegans*; Dm, *Drosophila melanogaster*; Hs, *Homo sapiens*; Sc, *S. cerevisiae*; Sp, *S. pombe*; Xl, *Xenopus laevis*.

[e] Identified by visual analysis of the genomic sequence near the site of a *P*-element insertion.

[f] Prokopenko *et al.* (1999).

[g] Hassan *et al.* (1998).

quence data provide strong evidence that the mutations are correctly assigned as allelic to previously identified genes.

Previously, we and others have shown that the *l(2)k06921*, *l(2)k08104*, *l(3)S024532*, and *l(3)S058701* mutations are allelic to known genes on the basis of complementation data (Berkeley Drosophila Genome Project, unpublished results; Salzberg *et al.* 1997). Our plasmid rescue and sequence data confirm this allelism for the *l(2)k08104* and *l(3)S058701 P* elements, which are allelic to *escargot* (*esg*, Whiteley *et al.* 1992) and *extra macrochaetae* (*emc*, Ellis *et al.* 1990; Garrell and Modolell 1990), respectively, and are inserted in the 5′-untranslated regions (5′-UTRs) of the genes (Table 3). The *l(3)S024532 P* element is allelic to *string* (*stg*, Edgar and O'Farrell 1989; Jimenez *et al.* 1990) and is inserted in the *stg* gene (Tables 2 and 3). The *l(2)k06921 P*-element insertion was shown (Berkeley Drosophila Genome Project, unpublished results) to be allelic to *Star* (*S*; *floater*, *fltr*; Kolodkin *et al.* 1994) and is inserted 1060 nucleotides (nt) upstream of the *S* AUG on the (−) strand and 275 nt upstream of the *asteroid* (*ast*, Higson *et al.* 1993; Kotarski *et al.* 1998) AUG on the (+) strand (Table 3) and, therefore, may affect both genes.

*cyrano:* We and others (Byars *et al.* 1999) found that *cyrano* (*cyr*) is allelic to *raw* (*raw*, Nüsslein-Volhard *et al.* 1984). *raw* is a dorsal-open group gene required for the regulation of Jun N-terminal kinase (JNK) signaling during dorsal closure that encodes a novel protein of 989 amino acids (aa). We used genomic sequences derived from $cyr^{k01021}$ to independently identify through database searches the GH23250 cDNA (Rubin *et al.* 2000a) as a candidate clone for *raw/cyr* (Table 4). It encodes a new smaller isoform (805 aa) of the RAW protein (Figure 2) generated by alternative splicing and the use of an upstream initiation methionine (Byars *et al.* 1999). Interestingly, we were unable to align with each other any of the six plasmid rescue sequences derived from $cyr^{k01021}$ and $cyr^{k08801}$ *P* elements, suggesting that the two *P* elements may be far apart. Database analysis revealed that the distance between the two insertions is at least 22 kb. The $cyr^{k08801}$ *P* element is inserted in an intron of the *raw* gene, and $cyr^{k01021}$ is inserted ~5 kb upstream of the *raw* AUG (GenBank accession no. AF186024, Byars *et al.* 1999), but within the coding sequence of the alternatively spliced GH23250 cDNA.

*unchained:* The molecular and genetic analyses of *unchained* (*unch*, Kania *et al.* 1995) suggest that it is allelic to *Sin3A* (Neufeld *et al.* 1998; Pennetta and Pauli 1998). Mouse mSin3 is a transcriptional corepressor that forms a ternary complex with the Mad and Max basic helix-loop-helix leucine zipper proteins, recruits other corepressor proteins, and downregulates Myc target genes (Ayer *et al.* 1995; Schreiber-Agus and DePinho 1998). Drosophila embryos homozygous for *Sin3A* null mutation fail to hatch, but have no obvious defects in muscle or nervous system development (Pennetta and Pauli 1998). We found (Table 3 and data not shown) that the $unch^{k15501}$ *P* element is inserted ~24 nt upstream of the the 5′-end of the longest *Sin3A* mRNA (GenBank accession no. AF024603, Neufeld *et al.* 1998) or 131 nt upstream of the beginning of the alternatively spliced *Sin3A* mRNA (GenBank accession no. AJ007518, Pennetta and Pauli 1998). Furthermore, database searches using the B25 genomic fragment flanking $unch^{k15501}$ identified the LD13852 clone (Rubin *et al.* 2000a) as a putative *unch* cDNA (Table 4). It maps to the *unch* locus (Table 4) and is likely to be affected by the $unch^{k15501}$ insertion, which is located 33 nt upstream of its 5′-end (data not shown). The $unch^{k15501}$ insertion, although not revertible, fails to complement deficiencies and is therefore likely to be responsible for the *unch* phenotype (Kania *et al.* 1995). On the basis of these observations we concluded that *unch* may be allelic to *Sin3A*. However, complementation tests between $unch^{k15501}$ and several *Sin3A* alleles (including the $Sin3A^{ex4}$ null allele, Pennetta and Pauli 1998) showed that they complement each other (Table 2). Note that the $Sin3A^{08269}$, $Sin3A^{k05415}$, and $Sin3A^{k11222}$ *P* elements are inserted in the first intron of *unch* (Pennetta and Pauli 1998; data not shown), whereas the $unch^{k15501}$ *P* element is inserted ~5 kb upstream, in the very beginning of the *unch* transcription unit (see above). Therefore, the peculiar complementation data may be the result of intragenic complementation, and *unch* may indeed be allelic to *Sin3A*. Further phenotypic and genetic analyses will be required to resolve this matter.

The above issues are further complicated as we cannot exclude that the $unch^{k15501}$ insertion may affect the *amphiphysin* gene (Razzaq *et al.* 2000), which is transcribed from a complementary strand in the orientation opposite to *Sin3A*. The $unch^{k15501}$ *P* element is inserted ~0.8 kb upstream of the 5′-end of *amphiphysin* mRNA (Table 3) and may affect the *amphiphysin* promoter or enhancer elements.

**First mutations in cloned genes:** We identified only one *P* element affecting a gene for which there were no mutations available. *Calreticulin* (*Crc*) is the gene mutated by the $potp^{S114307}$ *P*-element insertion. The $potp^{S114307}$ *P* element affects the *Crc* gene (Table 3), since it is inserted in the 5′-UTR of *Crc* mRNA, 28 nt upstream of initiator methionine.

Vertebrate calreticulins are major $Ca^{2+}$-binding proteins of endoplasmic reticulum implicated in the regulation of intracellular $Ca^{2+}$ signaling, glycoprotein folding, integrin-mediated cell adhesion, and steroid-dependent gene expression (Coppolino *et al.* 1997; Krause and Michalak 1997). The Drosophila *Crc* gene (Smith 1992) was previously shown to map to 86B-C (Gamo *et al.* 1998) or 85E1-5 (Christodoulou *et al.* 1997). The $potp^{S114307}$ *P* element was mapped to 85E

**TABLE 4**

**cDNAs for novel genes**

| Gene (cytology) | cDNA | Mapping position | Length (kb) | Northern (kb) | ORF (aa)[a] | GenBank accession no. | BLASTP search results[b] (species, protein accession) | % ID/SIM[c] |
|---|---|---|---|---|---|---|---|---|
| *aay* (67B2-3) | 5B | 67B | 1.466 | 1.4, 1.7 | 270 | AF191498 | 3-Phosphoserine phosphatase | |
| | | | | | | | (Hs, NP_004568) | 51/61 |
| | | | | | | | (At, BAA33806) | 50/61 |
| | | | | | | | (Sm, AAC46897) | 46/54 |
| | | | | | | | (Ec, CAA26852) | 30/41 |
| *chrw* (59E2) | LD47384[d] | 59E1-2 | 1.632 | | 261 | AF209706 | Rab-related small G protein | |
| | | | | | | | RAB5 (Hs, NP_004574) | 42/54 |
| | | | | | | | RAB21 (Cf, P55745) | 40/56 |
| | | | | | | | rab5 (Rn, AAC26004) | 40/52 |
| | | | | | | | RAB25 (Mm, AAD39911) | 39/49 |
| | | | | | | | Drab5 (Dm, BAA87879) | 37/50 |
| *cyr* (29F1-2) | GH23250[d] | 29E | 3.130 | | 805 | AF208397 | RAW (Dm, AAF28462) | 99/99 |
| | | | | | | | C02C6.2 gene product (Ce, CAB01856) | 34/40[e] |
| *dmt* (85E5-6) | 16A | 85E | 3.163 | 3.3 | 857 | AF203478 | No significant similarities | |
| *glu* (36A7-8) | glu11 | 36A | 4.461 | 4.7 | 1409 | AF185287 | XCAP-C (X1, P50532) | 42/53 |
| | | | | | | | CAP-C (Hs, NP_005487) | 40/51 |
| | | | | | | | Smc4p (Sc, NP_013187) | 39/50 |
| | | | | | | | cut3p (Sp, CAB46756) | 37/47 |
| *hoip* (30C1-2) | GH03082[d,f] | 30C | 0.566 | 0.7 | 127 | AF208396 | Non-histone chromosome protein 2 (NHP2)-like 1 (Hs, NP_004999) | 79/88 |
| | | | | | | | NHP2/RS6 family protein YEL026W homolog (Ce, Q21568) | 74/87 |
| | | | | | | | Ribosomal protein L7Ae-like (At, CAB53753) | 70/81 |
| | | | | | | | Nhp2p homolog Snu13p (Sc, NP_010888) | 67/80 |
| | | | | | | | Nhp2p HMG-like nuclear protein (Sc, NP_010073) | 41/55 |
| *melt* (65E5-6) | 8C | 65E | 0.692 | 4.2 | 717+[g] | AF205831[g] | K10B4.3 gene product (Ce, AAB71008) | 33/44[h] |
| | 8G | 65E7-9 | 1.722 | 4.2 | | | | |
| | HL03627[d] | 65E | 2.000 | | | | | |
| *pbl* (6A18-20) | 1A | 66AB | 3.140 | 3.7, 4.0, 5.5, 7.0 | 853 | AF136492 | Ect2 oncoprotein (Mm, NP_031926) | 40/61 |
| | | | | | | | ECT2 (Hs) | 38/51[i] |
| | | | | | | | T19E10.1b gene product (Ce, CAB54311) | 32/42 |
| *sktl* (57B4-5) | 12C | 57B | 2.922 | 3.8 | 700 | AF071417 | PIP5K, phosphatidylinositol-4-phosphate 5-kinase | |
| | | | | | | | type I-β (Mm, BAA13031) | 60/69 |
| | | | | | | | type I-α (Hs, AAC50910) | 60/68 |
| | | | | | | | type II (Ce, AAB54130) | 59/68 |

(*continued*)

**TABLE 4**

**(Continued)**

| Gene (cytology) | cDNA | Mapping position | Length (kb) | Northern (kb) | ORF (aa)[a] | GenBank accession no. | BLASTP search results[b] (species, protein accession) | % ID/SIM[c] |
|---|---|---|---|---|---|---|---|---|
| *stich1* (86B) | GM05287[d] | 86B | 2.931 | 5.2 | 610+ | AF203477 | HEY (Dm, AAD46771) | 51/75[j] |
| | | | | | | | Hey1 (Mm, CAB51321) | 51/71[j] |
| | | | | | | | HRT2 (Mm, AAF14546) | 51/71[j] |
| | | | | | | | DEC1 (Hs, NP_003661) | 43/62[j] |
| | | | | | | | Stra13 (Mm, AAB64228) | 43/62[j] |
| | | | | | | | Hairy (Dm, P14003) | 42/58[j] |
| *unch* (49B1-49C4) | LD13852[d] | 49B1-5 | 7.3 | | | | Sin3A (Dm)[k] | |
| *veg* (53E1-2) | 31HC | 53E1-5 | 0.848 | | 449[g] | AF211892[g] | Unnamed protein (Hs, BAA91196) | 34/46 |
| | 31HE | 53E1-5 | 1.004 | 2.1 | | | F12F1.28 gene product (At, AAC17633) | 34/42 |
| | GM14315[d] | 53E1-5 | 1.251 | 2.1 | | | C18B11.05 predicted protein (Sp, Q09712) | 33/43 |
| | | | | | | | T09B4.1 gene product (Ce, AAB53004) | 31/43 |

[a] Length of predicted ORF followed by a plus sign indicates that the coding sequence is incomplete at the 3′-end.

[b] BLAST search results are current as of March 2000. In all cases a hit with the lowest Expect value is listed first followed by several other representative hits. At, *A. thaliana*; Ce, *C. elegans*; Cf, *Canis familiaris*; Dm, *D. melanogaster*; Ec, *Escherichia coli*; Hs, *H. sapiens*; Mm, *Mus musculus*; Rn, *Rattus norvegicus*; Sc, *S. cerevisiae*; Sm, *Schistosoma mansoni*; Sp, *S. pombe*; Xl, *X. laevis*.

[c] Percentage identity (ID) and similarity (SIM) was calculated for full-length amino acid sequences (unless otherwise indicated) using BestFit (Wisconsin Package, Genetics Computer Group).

[d] cDNA clone isolated by Berkeley Drosophila Genome Project/HHMI EST Project (Rubin *et al.* 2000a).

[e] Calculated for a 146-aa match between two proteins (in Cyrano, aa 158–303).

[f] Chimeric cDNA clone containing two cDNAs in a tail-to-tail orientation that correspond to *hoip* (5′ clone, mapped at 30C) and CYP6-like microsomal cytochrome P450 (GenBank accession no. AF083946; 3′ clone, mapped at 48F).

[g] Coding sequence was assembled from the sequences of three independently isolated cDNA clones.

[h] Calculated for a 148-aa match between two proteins (in Melted, aa 163–310).

[i] T. Miki, personal communication.

[j] Calculated for bHLH domains.

[k] BLASTN hits by sequences of the 5′- and 3′-ends of the LD13852 cDNA: AF024603, AF024604, AJ007518.

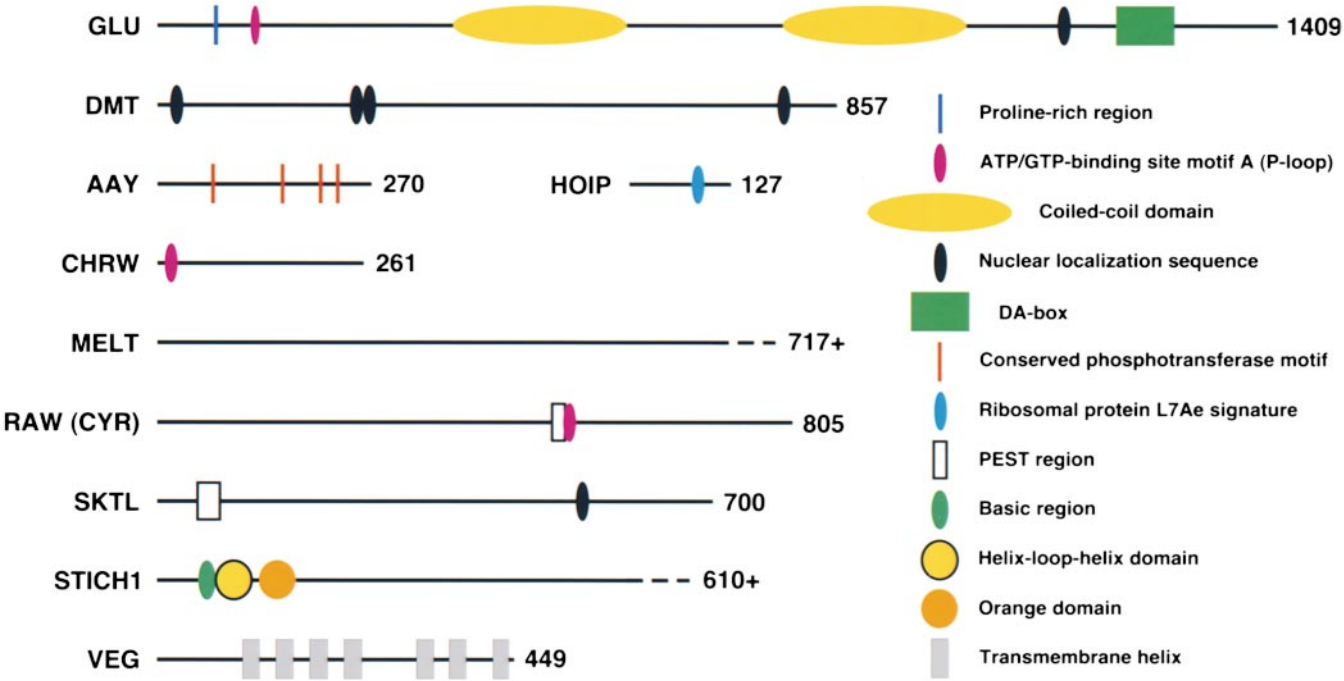


Figure 2.—Domain structure of novel proteins. Proteins are represented schematically with predicted functional domains and motifs shown in color. Abbreviated names of proteins are indicated on the left, and their lengths (in aa) are indicated on the right. Length of a protein followed by a plus sign indicates that the coding sequence of a cDNA clone is incomplete at the 3′-end. MELT protein does not have any known functional domains. Proteins and the respective domains are drawn to scale. For exact location of domains and motifs refer to respective GenBank accession numbers (Table 4).

(Salzberg *et al.* 1997). To resolve the ambiguity of the mapping of *Crc*, we identified through database searches the 1.5-kb LD07621 cDNA clone (Rubin *et al.* 2000a) corresponding to *Crc* and mapped it to 85E1 (data not shown). These observations, together with the ability to revert the lethality by a precise excision of $potp^{S114307}$ (Salzberg *et al.* 1997), indicate that it is indeed a mutation in the *Crc* gene.

A *Crc* mutation was reported to cause hypersensitivity of flies to diethylether anesthesia ($Crc^{eth\text{-}as311}$, Gamo *et al.* 1998). Our data show that *Crc* is an essential gene, since the $potp^{S114307}$ insertion is homozygous lethal and results in loss of neurons, decreased staining with the MAb 22C10 in the central nervous system (CNS), disorganization of the PNS, and pathfinding defects during embryonic development (Salzberg *et al.* 1997).

***P*-element insertions in genes that were identified, but not characterized molecularly:** This class of genes includes those *P*-element insertions that may serve as cloning tools for previously identified mutations that are not transposon tagged. The *pebble* (*pbl*) gene was identified in a chemical mutagenesis screen for mutations affecting the pattern of the embryonic cuticle (Jürgens *et al.* 1984) and is required for cytokinesis during postblastoderm mitoses (Hime and Saint 1992; Lehner 1992). We identified two alleles of *pbl* in our *P*-element screen (Salzberg *et al.* 1997), cloned the gene, and showed that it encodes a putative guanine nucleotide exchange factor for Rho1 G protein (Rho-GEF, Prokopenko *et al.* 1999).

**Novel genes:** We identified 13 novel genes and cloned or identified candidate cDNAs for 10 genes. The identity of the respective proteins, their domain structure, and RNA expression are described in the following sections. The information on cDNA clones including their names, lengths, mapping positions, GenBank accession numbers, and results of Northern analysis and sequence analysis is presented in Table 4 and summarized in Table 5. Cloning and functional characterization of *bonus* (*bon*) will be published elsewhere (R. B. Beckstead, S. N. Prokopenko and H. J. Bellen, unpublished results). The identity of three remaining genes remains unknown.

*l(2)k00424:* The *l(2)k00424* strain carries two *P*-element insertions that were mapped at 30D1-2 and 44F1-2. However, only one insertion (at 44F1-2) is responsible for the lethality and possibly the organizational defects observed in the dorsal cluster of neurons in the PNS (Kania *et al.* 1995) given the complementation data with deficiencies (Table 2). The genomic sequences flanking this insertion did not show any homologies in database searches. Hence, the *l(2)k00424* gene remains unidentified. The second genomic fragment isolated from this strain maps at 30D1-2. This *P* element is inserted in the 5′-UTR, 21 nt upstream of the initiator methionine of the *FKBP59-bp1* gene, which encodes the

**TABLE 5**

**Molecular screen: summary**

| Gene | Synonym | Cytology[a] | Protein identity | Known developmental or cellular roles |
|---|---|---|---|---|
| | | | I. Allelic genes | |
| *Cyclin E* (*CycE*) | *fondue* (*fond*) | 35D4-6 | $G_1$/S-specific cyclin E[b] | $G_1$/S transition during mitosis and endoreplication[b,c]<br>Coupling of growth to cell cycle[d] |
| *mindmelt* (*mm*) | *muscleblind* (*mbl*) | 54B8-16 | $Cys_3His$ zinc-finger protein[e] | Terminal differentiation of somatic muscles (organization of Z-bands and epidermal attachments)[f]<br>Photoreceptor differentiation[e] |
| *patched* (*ptc*) | *rubberneck* (*rubr*) | 44D2-5 | Integral membrane protein[g] | Segmental patterning[g]<br>Hedgehog receptor[h] |
| *puckered* (*puc*) | *hearty* (*hrt*) | 84E10-13 | JNK phosphatase[i] | Regulation of JNK activity during dorsal closure[i] |
| *raw* (*raw*) | *cyrano* (*cyr*) | 29F1-2 | Novel protein | Regulation of JNK signaling during dorsal closure[j] |
| *Star* (*S*) | *floater* (*fltr*) | 21E2 | Type II transmembrane protein[k] | Modulation of epidermal growth factor receptor signaling during photoreceptor,[k] embryonic nervous system,[l] and mesodermal[m] development |
| *three rows* (*thr*) | *anarchist* (*anch*) | 54F5-55A1 | Novel protein[n] | Chromatid segregation during mitosis[n] |
| *unchained* (*unch*) | *Sin3A* (?) | 49B1-49C4 | PAH motif protein[o] | Unknown[p] |
| | | | II. New gene-mutation associations | |
| *potpourri* (*potp*) | *Calreticulin* (*Crc*) | 85E1 | Calreticulin[q] | Sensitivity to ether anesthesia[r] |
| | | | III. *P*-element alleles of known genes | |
| *pebble* (*pbl*) | | 66A18-20 | RhoGEF, Ect2 (Mm, Hs) homolog[s] | Formation of a contractile ring and initiation of cytokinesis[s] |

(*continued*)

FK506-binding protein FKBP59, a member of the immunophilin family of proteins (Zaffran 2000).

*bumper-to-bumper:* The *bumper-to-bumper* (*btb*) gene was identified by a single revertible *P*-element insertion that leads to pathfinding and connectivity defects and affects dorsoventral migration of lateral chordotonal neurons (Kania *et al.* 1995). The *P*-element insertion in $btb^{k09901}$ may affect a predicted gene, *CG5380*, which encodes a DNA-directed RNA polymerase III. Genomic sequences flanking the 5′-end of the *P* element show homology to RNA polymerase III subunits from several species (Table 3). This *P*-element insertion is revertible (Kania *et al.* 1995) and is located 418 nt upstream of *CG5380* AUG. Both *P* element and a plasmid rescued genomic fragment were mapped to 47A-47B14.

*on-the-rack: on-the-rack* (*rack*) was identified by a single revertible *P*-element insertion that causes loss of LCh5 neurons and affects morphology of neurons in the lateral cluster (Kania *et al.* 1995). Genomic sequence flanking $rack^{k15001}$ *P*-element insertion did not give any significant matches in database searches (Table 3). Our attempts to clone *rack* were unsuccessful, because the $rack^{k15001}$ *P*-element insertion is associated with DNA rearrangements affecting the *P* element (Table 3).

*astray: astray* (*aay*) was identified by a single revertible *P*-element insertion ($aay^{S042314}$) that causes severe defects in the axonal trajectories in the embryonic PNS (Salzberg *et al.* 1997). We cloned the full-length *aay* cDNA, which encodes a 3-phosphoserine phosphatase (Table 4). L-3-phosphoserine phosphatase (PSPase) catalyzes the last rate-limiting step in the biosynthesis of serine—$Mg^{2+}$-dependent hydrolysis of L-phosphoserine to serine as well as an exchange reaction between L-serine and L-phosphoserine. The AAY protein is homologous to PSPases from mammals, plants, yeast, and bacteria (Table 4 and data not shown) and is most similar to human phosphoserine phosphatase (Collet *et al.* 1997). AAY and human PSPase belong to a new class of phosphotransferases (Collet *et al.* 1998) characterized by a conserved N-terminal DXDX(T/V) motif (aa 67–71 in AAY, Figure 2). The first aspartate in this motif is absolutely conserved among all proteins and has been implicated in covalent binding of phosphate and formation of a phosphoenzyme catalytic intermediate (Collet *et al.* 1999). The two proteins share two other highly conserved motifs (in AAY, aa 155–159 and 203–207, 225–232, Figure 2). Conserved residues within these motifs play an important role in catalysis, as demonstrated by site-directed mutagenesis of other phosphotransferases (P-type ATPases and human PSPase, Lingrel and Kuntzweiler 1994; Collet *et al.* 1999), and are likely to form a catalytic pocket, as shown by 3-D structure

**TABLE 5**

**(Continued)**

| Gene | Cytology[a] | Protein identity | Embryonic expression |
|---|---|---|---|
| | | IV. Novel genes | |
| *l(2)k00424* | 44F1-2 | Gene not cloned | |
| *astray* (*aay*) | 67B2-3 | 3-Phosphoserine phosphatase | Complex. Segmentally repeated expression in stripes and clusters of cells. Area surrounding the pole cells. Ring of large cells around the anterior gut |
| *bumper-to-bumper* (*btb*) | 47A1-47B14 | DNA-directed RNA polymerase III (?) | |
| *chrowded* (*chrw*) | 59E2 | Rab-related small G protein | |
| *dalmatian* (*dmt*) | 85E5-6 | Novel nuclear protein | Ubiquitous early, restricted to the CNS and the PNS late. Pole cells, gut, posterior spiracles |
| *gluon* (*glu*) | 36A7-8 | XCAP-C (X1)/Smc4p (Sc) homolog | Ubiquitous early. Expressed in dividing cells throughout embryogenesis—pole cells, neuroblasts in the CNS and the PNS |
| *hoi-polloi* (*hoip*) | 30C1-2 | YEL026W (Ce)/Snu13p (Sc) homolog | |
| *melted* (*melt*) | 65E5-6 | Novel protein | Segmentally repeated expression in stripes and domains. Ventral furrow. Brain, gut |
| *on-the-rack* (*rack*) | 48A1-6 | Gene not cloned | |
| *skittles* (*sktl*) | 57B4-5 | PIP5K | Dynamic, CNS, PNS, germ cells, dorsal vessel, gut, mesoderm[t] |
| *sticky ch1* (*stich1*) | 86B | Hairy-related bHLH protein | Complex. Anterior and posterior midgut primordia. Gut throughout emryogenesis. Segmentally repeated pattern. Amnioserosa, posterior spiracles, tracheal trees, head region |
| *vegetable* (*veg*) | 53E1-2 | Novel integral membrane protein | |

bHLH, basic helix-loop-helix; Ce, *C. elegans*; CNS, central nervous system; Hs, *H. sapiens*; Mm, *M. musculus*; PAH, paired amphipathic helix; PNS, peripheral nervous system; RhoGEF, guanine nucleotide exchange factor for Rho G proteins; Sc, *S. cerevisiae*; Xl, *X. laevis*.

[a] Cytological locations of genes with the exception of *l(2)k00424*, *btb*, *potp*, *rack*, *stich1*, *unch*, and *veg* are given according to FlyBase (2000).

[b] Richardson *et al.* (1993); Knoblich *et al.* (1994).

[c] Su and O'Farrell (1998).

[d] Prober and Edgar (2000).

[e] Begemann *et al.* (1997).

[f] Artero *et al.* (1998).

[g] Hooper and Scott (1989); Nakano *et al.* (1989).

[h] Chen and Struhl (1996, 1998); Marigo *et al.* (1996).

[i] Martín-Blanco *et al.* (1998).

[j] Byars *et al.* (1999).

[k] Kolodkin *et al.* (1994).

[l] Klämbt *et al.* (1991).

[m] Buff *et al.* (1998).

[n] D'Andrea *et al.* (1993); Philp *et al.* (1993).

[o] Neufeld *et al.* (1998); Penneta and Pauli (1998).

[p] Embryos homozygous for *Sin3A* null mutation fail to hatch, but have no obvious defects in muscle or nervous system development (Pennetta and Pauli 1998). *Sin3A* interacts genetically with *sina* and may be required for cell proliferation or survival, as suggested by mitotic analysis in the eye imaginal disk (Neufeld *et al.* 1998).

[q] Smith (1992).

[r] Gamo *et al.* (1998).

[s] Prokopenko *et al.* (1999).

[t] Knirr *et al.* (1997a); Hassan *et al.* (1998).

analysis of Pseudomonas haloacid dehalogenase (Li *et al.* 1998).

During embryonic development, *ASTRAY* is expressed in a complex pattern (Figure 3, A–E). During stage 5 (Figure 3A), *ASTRAY* is expressed in a highly specific pattern consisting of 7 stripes capped on the dorsal side by a longitudinal stripe. It is also expressed abundantly in the area surrounding the pole cells and

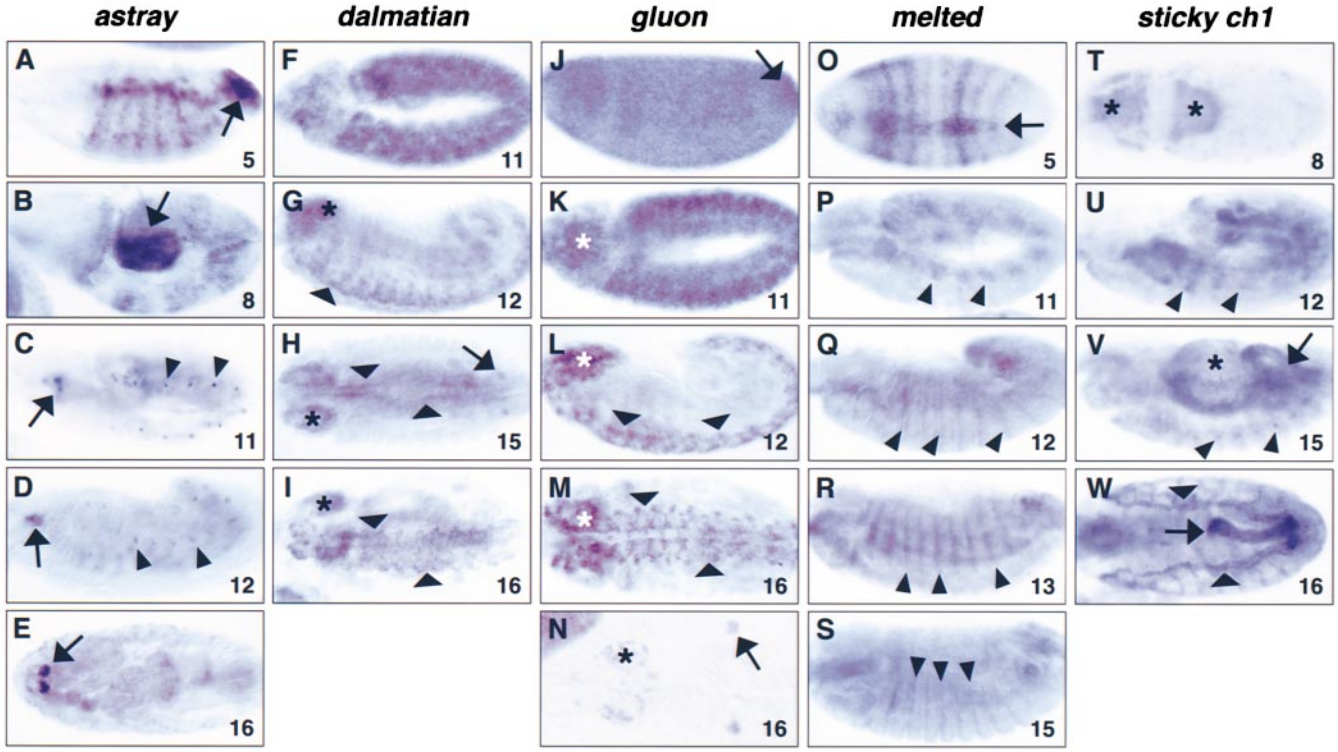


Figure 3.—Expression of novel genes during embryonic development. Wild-type embryos were hybridized with *aay* (A–E), *dmt* (F–I), *glu* (J–N), *melt* (O–S), and *stich1* (T–W) antisense RNA probes. Sense probes used in parallel *in situ* hybridization experiments gave either no staining or a low level of background staining (data not shown). Stages of embryonic development are indicated in each lower right corner. For details on generating RNA probes see materials and methods.

the invagination in which the pole cells migrate (Figure 3, A and B, arrows). The expression during germ band extension is characterized first by 7 broad stripes (Figure 3B) and later by 10 stripes that eventually fuse to form a peculiar pattern (data not shown). This expression then fades and gives rise to a pattern of segmentally repeated small clusters of cells (Figure 3C, arrowheads), a ring of large cells around the anterior gut (Figure 3, C–E, arrows), and low levels of expression in most of the gut in more mature embryos (Figure 3E).

How does a mutation in PSPase ($aay^{S042314}$ *P* element is inserted in the 5′-UTR of *aay*) lead to the axon guidance phenotype observed in the PNS? Serine is used not only as a building block for protein synthesis but also as a precursor of phospholipids (phosphatidylserine and sphingomyelin) and glycolipids. Loss of AAY may cause abnormalities in membrane biogenesis in specific cells that would affect transmembrane signaling in neuronal growth cones. Clearly, other alternative hypotheses are possible. It is interesting to note that l-serine does not cross the blood-brain barrier well (Smith *et al.* 1987) and that defects in the serine biosynthesis pathway leading to low serine levels in cerebrospinal fluid have been described in patients with Williams syndrome (3-phosphoserine phosphatase deficiency, Jaeken *et al.* 1997) and congenital encephalopathy (3-phosphoglycerate dehydrogenase deficiency, Jaeken *et al.* 1996).

*chrowded: chrowded* (*chrw*) was identified by a single revertible *P*-element insertion ($chrw^{k06908}$) that causes organizational and morphological defects in the PNS (Kania *et al.* 1995). We identified the LD47384 clone (Rubin *et al.* 2000a) as a candidate *chrw* cDNA. It encodes a Rab-related small G protein (Figure 2) that is identical to the ORF of the predicted gene *CG3870* (Rubin *et al.* 2000b). The $chrw^{k06908}$ *P* element does not directly affect the LD47384 cDNA, but is inserted in close proximity to its 5′-end (Table 3).

CHRW is distantly related to a number of Rab proteins from Drosophila, mammals, plants, and yeast (Table 4 and data not shown), but has a longer ORF (261 aa compared to 200–220 aa in most Rab proteins). Rab proteins compose a separate family within the Ras superfamily that consists of >30 members (in mammals, Rab1–25 and others) implicated in different aspects of intracellular vesicular trafficking (reviewed by Zerial and Huber 1995). Rab proteins that share at least 80% identity are included in the same subfamily (*e.g.*, Rab1, Rab2, *etc.*). In contrast, CHRW shares not more than 40–42% identity (51–54% similarity) with known Rab proteins, including several Rab proteins cloned in Drosophila (Sasamura *et al.* 1997; Satoh *et al.* 1997). Therefore, we propose that CHRW is a Rab-related protein. Alternatively, it may represent the first member of an unidentified family of Rab proteins. Like other small G

proteins, CHRW contains four conserved regions that form the GTP-binding and GTPase catalytic site and the C-terminal CCX cysteine motif as well as the Rab-specific effector region. It is possible that CHRW does not function as a Rab G protein, since it lacks two conserved amino acids present in Rab proteins—G41 in the effector region (A in CHRW) and A151 in region IV (T in CHRW). Further genetic and molecular analyses are necessary to demonstrate that CHRW is indeed a Rab-related protein and that the *chrw*$^{k06908}$ *P* element affects the *Rab* gene identified with the LD47384 cDNA.

*dalmatian:* We identified *dalmatian* (*dmt*) in our chemical mutagenesis screen as a mutation that leads to a loss of neurons, disorganization of the PNS, and formation in the ectoderm of small round cells that stain darkly with MAb 22C10 (Salzberg *et al.* 1994). We used a revertible *P*-element allele of *dmt*, *dmt*$^{S048103}$ (Salzberg *et al.* 1997), to clone the full-length cDNA (Table 4). This cDNA is likely to correspond to the *dmt* gene, since the *dmt*$^{S048103}$ *P* element is inserted 11 nt upstream of its 5′-end. *dmt* encodes a novel protein that is not related to any known or predicted proteins from other species. The only conspicuous feature of DMT is the presence of four nuclear localization sequences (Figure 2). In addition, the PSORT algorithm (Nakai and Horton 1999) predicts a 78% probability of DMT being a nuclear protein. Analysis of expression pattern by *in situ* hybridization revealed that at cellular blastoderm, *DALMATIAN* is expressed at low levels (data not shown). Expression levels increase during germ band extension and the gene is widely expressed at stage 11 (Figure 3F). From stage 12 (Figure 3G), expression becomes more restricted to the PNS and the CNS—ventral nerve cord (Figure 3, G and I, arrowheads) and brain (Figure 3, G–I, asterisks). In addition, *DALMATIAN* is expressed in the pole cells (data not shown), the gut (Figure 3H, arrowheads), and the posterior spiracles (Figure 3H, arrow). In summary, *dmt* is expressed in numerous tissues including the CNS throughout most of embryonic development and encodes a novel presumably nuclear protein of an unknown function.

*gluon: gluon* (*glu*) was identified by a single nonrevertible *P*-element insertion (*glu*$^{k08819}$) that fails to complement a deficiency uncovering the region where the transposon is inserted (Kania *et al.* 1995). The embryonic phenotype of *glu* is characterized by subtle organizational defects in ventral and lateral PNS resulting in irregularly shaped clusters. Sequencing of the genomic fragments flanking *glu*$^{k08819}$ revealed that the *P* element affects a gene related to SMC (structural maintenance of chromosomes) proteins (Table 3), which are required for chromosome condensation during mitosis (reviewed by Strunnikov 1998; Hirano 1999). The nearly full-length *glu* cDNA encodes a predicted ORF of 1409 amino acids related to members of a highly conserved family of SMC proteins (Table 4)—Xenopus XCAP-C (Hirano and Mitchison 1994), human CAP-C (Nishiwaki *et al.* 1999), *Saccharomyces cerevisiae* Smc4p (Koshland and Strunnikov 1996), and *Schizosaccharomyces pombe* cut3p (Saka *et al.* 1994). Therefore, GLU is a Drosophila XCAP-C/Smc4-like protein. Similar to other Smc4p-like proteins GLU contains an N-terminal nucleotide-binding motif, two central coiled-coil domains, and a C-terminal DA-box that has DNA-binding capability (Figure 2). The presence of a nuclear localization sequence and PSORT algorithm prediction (91% probability) suggest that GLU is a nuclear protein as demonstrated for XCAP-C (Hirano and Mitchison 1994) and cut3p (Saka *et al.* 1994; Sutani *et al.* 1999).

GLU is likely to be a component of the 13S condensin complex described in Xenopus (Hirano *et al.* 1997) and *S. pombe* (Sutani *et al.* 1999). This complex consists of two SMC subunits (in Xenopus, XCAP-E and XCAP-C) and three non-SMC subunits (in Xenopus, XCAP-D2, XCAP-G, and XCAP-H). 13S condensin as well as its individual components is absolutely required for mitotic chromosome condensation *in vitro* and *in vivo* (reviewed by Strunnikov 1998; Hirano 1999). Two other components of Drosophila condensin have been identified—Barren (XCAP-H homolog), which is required for sister chromatid segregation during mitosis (Bhat *et al.* 1996), and dSMC2 (XCAP-E homolog, Hirano 1999).

*glu* has a typical "mitotic" expression pattern similar to other genes implicated in cell cycle regulation or mitosis (*e.g.*, *stg*, *CycA*, and *barr*; Edgar and O'Farrell 1989; Lehner and O'Farrell 1989, 1990; Bhat *et al.* 1996). Prior to cellularization (Figure 3J), there is maternally provided *GLUON* RNA that is enriched at the posterior end of the embryo (Figure 3J, arrow), where the pole cells form during telophase of mitotic cycle 10 (stage 4). During germ band extension, there are high levels of *GLUON* in dividing neuroblasts in the PNS and the CNS (Figure 3K). *GLUON* expression in the brain (Figure 3, K–N, asterisks) and the ventral nerve cord (Figure 3, L and M, arrowheads) persists throughout embryogenesis. By the end of stage 16 (Figure 3N), when most of the embryonic cells have stopped dividing, *GLUON* RNA expression in most tissues is much lower than at earlier embryonic stages. However, *GLUON* continues to be expressed at elevated levels in tissues that will resume proliferation during larval development—in neuroblasts in the brain (Figure 3N, asterisk) and in gonads (Figure 3N, arrow). Hence, the expression pattern of *GLUON* is very similar to *BARREN*, which encodes another condensin subunit (Bhat *et al.* 1996), but its *in vivo* role remains to be determined.

*hoi-polloi: hoi-polloi* (*hoip*) was identified by a single nonrevertible *P*-element insertion (*hoip*$^{k07104}$) that fails to complement a deficiency uncovering the region where the transposon is inserted (Kania *et al.* 1995). *hoip*$^{k07104}$ causes subtle fasciculation and organization defects characterized by misplaced cells within few neuronal clusters. Database searches using genomic sequences flanking the site of *hoip*$^{k07104}$ insertion identified

HOIP as a member of a conserved family of YEL026W-like proteins (Table 3). We identified the GH03082 clone (Rubin *et al.* 2000a) as a candidate near full-length *hoip* cDNA (Table 4). It encodes a 127-aa protein related to human non-histone chromosome protein 2-like 1 (NHP2L1) protein (Saito *et al.* 1996), *Caenorhabditis elegans* YEL026W (GenBank accession no. Q21568), and *S. cerevisiae* Snu13p (Gottschalk *et al.* 1999). These proteins compose an evolutionary conserved family of YEL026W-like proteins found in phyla from plants to humans and distantly related to families of NHP2-like proteins and ribosomal L7Ae proteins (see Table 4 and Maiorano *et al.* 1999). We propose that HOIP is likely to be a functional homolog of these proteins, since it is very similar to human and *C. elegans* proteins (79 and 74% identity, respectively; Table 4). HOIP contains a central region corresponding to the ribosomal L7Ae signature (aa 72–89, Motif prediction, see Figure 2) that is highly conserved (83–100% identical) among family members. Human NHP2L1 was shown to bind directly to the 5′ stem-loop of U4 small nuclear RNA and is an essential component of a spliceosome (Gottschalk *et al.* 1999). On the basis of these observations we hypothesize that HOIP is an RNA-binding protein, component of a spliceosome, and is required for pre-RNA splicing.

*melted:* The *melted* (*melt*) gene was identified by a single revertible *P*-element insertion ($melt^{S144114}$) that results in abnormal morphology and mild loss of peripheral neurons (Salzberg *et al.* 1997). We used plasmid-rescued fragments to clone two partially overlapping *melt* cDNAs (8C and 8G, see Table 4). In addition, through database searches we identified the HL03627 clone (Rubin *et al.* 2000a) as a candidate *melt* cDNA. We used the full-length sequence of the three clones to assemble a 2.903-kb *melt* cDNA that is incomplete at the 3′-end. BLAST searches with the sequence of MELT gave no significant results, except a limited homology to a predicted protein from *C. elegans* (Table 4). In addition, MELT does not contain any functional domains or motifs. Early in embryogenesis (stage 5, Figure 3O), *MELTED* RNA is expressed in 8 or 9 stripes and in the invaginating ventral furrow (Figure 3O, arrow). During germ band extension, *MELTED* is expressed in discrete domains in each segment of the embryo (Figure 3P, arrowheads). Later, this pattern is refined to several rows of ectodermal cells in the anterior of each segment (Figure 3, Q–S, arrowheads). There are also low levels of expression in the brain and the gut (data not shown). In conclusion, MELT is a novel protein of unknown function, which, on the basis of its expression pattern, may be required for ectodermal patterning.

*skittles:* The *skittles* (*sktl*) cDNA was isolated in an attempt to clone the *fata morgana* (*fam*) gene. *fam* was identified by several *P*-element alleles that result in morphological defects of lateral and v′ chordotonal neurons in the PNS (Kania *et al.* 1995). We and others have shown (Knirr *et al.* 1997a; Hassan *et al.* 1998) that *sktl* is nested in the first intron of *inscuteable* (*insc*; *not enough muscles, nem*; Burchard *et al.* 1995; Kraut and Campos-Ortega 1996). *fam P*-element insertions (*e.g.*, $fam^{k07505}$) fail to complement independently generated alleles of *insc* (Table 2) and affect two genes (Knirr *et al.* 1997a; Hassan *et al.* 1998). Therefore, the name *fata morgana* does not refer to either of the two genes, but rather describes the composite phenotype caused by the loss of both. The isolated *sktl* cDNA (Table 4) encodes a putative phosphatidylinositol-4-phosphate 5-kinase that is longer than the published sequence of SKTL protein (Knirr *et al.* 1997b). SKTL contains a nuclear localization sequence (aa 577–593) and a PEST region (aa 49–78) and has a 78% probability to localize to the nucleus (PSORT prediction). Expression pattern and genetic and functional analyses of *sktl* have been published elsewhere (Knirr *et al.* 1997a; Hassan *et al.* 1998).

*sticky ch1:* The *sticky ch1* (*stich1*) gene was originally identified by a single EMS-induced allele ($stich1^{D233}$) as a mutation that affects morphology of neurons (Salzberg *et al.* 1994). Later, we identified an independently isolated *P*-element allele of *stich1*, $stich1^{S143702}$ (Salzberg *et al.* 1997), and showed that the $stich1^{S143702}$ insertion is revertible (this study, Table 2). We used one of the genomic sequences flanking the site of the $stich1^{S143702}$ insertion to identify through database searches the GM05287 clone (Rubin *et al.* 2000a) as a putative *stich1* cDNA. It encodes a predicted basic helix-loop-helix (bHLH) protein similar to Hairy and Enhancer of split-related transcriptional repressors (Table 4, reviewed in Fisher and Caudy 1998). Common functional domains shared among these proteins are the bHLH domain and the adjacent orange domain (Figure 2), which is thought to confer functional specificity among Hairy-related proteins (Dawson *et al.* 1995). The bHLH domain is most closely related to that of Drosophila HEY, mouse Hey1 (Leimeister *et al.* 1999), and mouse HRT2 (Nakagawa *et al.* 1999, see Table 4). The basic region is most closely related to that of rat SHARP-1 (Rossner *et al.* 1997), human DEC1 (Shen *et al.* 1997), and mouse Stra13 (Boudjelal *et al.* 1997), but the position of a conserved proline residue is shifted in STICH1 2 amino acids toward the N terminus, suggesting that it may have a different DNA-binding specificity. Since the GM05287 cDNA is incomplete at the 3′-end, we do not know if it contains a C-terminal WRPW motif found in all Hairy family proteins and required for interaction with Drosophila non-HLH corepressor protein Groucho (reviewed in Fisher and Caudy 1998). Finally, the predicted protein is much longer (at least 610 aa) than other Hairy and Enhancer of split-related proteins (250–400 aa) and, therefore, it may not be their functional homolog.

The gene has a complex expression pattern during embryonic development (Figure 3, T–W). During stage 8, the RNA is expressed in the anterior and posterior midgut primordia (Figure 3T, asterisks). Expression in

the gut continues throughout embryonic development (Figure 3U; hindgut in Figure 3, V and W, arrows). During germ band retraction, expression is initiated in many tissues in a prominent segmentally repeated pattern (Figure 3U, arrowheads). Later expression is quite ubiquitous, but has higher levels in segmentally repeated clusters of cells (Figure 3V, arrowheads). Expression is also found in cells of amnioserosa (Figure 3V, asterisk), in the head region (stage 16, Figure 3W), in posterior spiracles (Figure 3W), and in tracheal trees (Figure 3W, arrowheads).

Analysis of the recently released Drosophila genome sequence revealed that the A32 *Eco*RI and the B32 *Bam*HI genomic fragments (Table 3) are located at least 28 kb apart on the genomic sequence (data not shown). We propose two alternative hypotheses to explain this: (1) they may derive from two different *P* elements that map 28 kb apart on the $stich1^{S143702}$ chromosome or (2) the $stich1^{S143702}$ *P* element may be associated with a 28-kb deletion. We found that the $stich1^{S143702}$ *P* element is inserted ~90 nt upstream of the 5′-end of the GM05287 cDNA at 86B. However, the A32 plasmid rescue from $stich1^{S143702}$ maps within the first intron of the *Domina* (*Dom*) gene, which encodes a transcription factor and maps to 86A2-4 (GenBank accession nos. AJ243814 and AJ243916). This indicates that $stich1^{S143702}$ may indeed affect two genes—*GM05287* and *Domina*.

In addition, we identified the *EP(3)0359 P* element as allelic to *stich1*, since it fails to complement both $stich1^{S143702}$ and $stich1^{D233}$ (Table 2 and data not shown). Like $stich1^{S143702}$, this *EP* insertion is located in the first intron of *Dom*. The close proximity of $stich1^{S143702}$ and $stich1^{EP0359}$ (the two map ~85 bp apart), their failure to complement an independently generated allele of *stich1* ($stich1^{D233}$, data not shown), and the ability to revert the lethality of $stich1^{S143702}$ suggest that *stich1* may be allelic to *Dom*. However, without molecular information on the nature of mutations in $stich1^{EP0359}$ and $stich1^{D233}$ we cannot exclude the possibility that the $stich1^{S143702}$ *P* element in the intron of *Dom* is viable and not responsible for the phenotype. Alternatively, both *P* elements may contribute to the lethality and the PNS phenotype associated with the $stich1^{S143702}$ chromosome. We were unable to test these hypotheses, since deficiencies uncovering cytological region 86A-B are not available and there are no recorded *Dom* mutations. In conclusion, we present identification of a cDNA encoding a bHLH protein similar to a family of Hairy-related transcriptional repressors that may correspond to the *sticky ch1* gene.

*vegetable: vegetable* (*veg*) was identified by several *P*-element alleles that cause a severe loss of neurons and fasciculation defects in the PNS (Kania *et al.* 1995). *veg* *P*-element insertions map to 53C or 53E (Kania *et al.* 1995; Spradling *et al.* 1999), leaving the issue of where the *veg* gene maps unresolved. However, all *veg* alleles fail to complement each other (Kania *et al.* 1995). They also fail to complement (Table 2) an independently isolated EMS allele of *veg* (C. Russell and G. Tear, personal communication). To clone *veg*, we selected three lines—$veg^{k03402}$, $veg^{k07202}$, and $veg^{k07228}$. Two of them ($veg^{k07202}$ and $veg^{k07228}$) carry *P* elements at 53E1-2 that are revertible (Kania *et al.* 1995 and this study) and, therefore, are likely to affect the gene. Since plasmid-rescued fragments from $veg^{k03402}$ and $veg^{k07228}$ lines carry internal *P*-element-derived sequences (Table 3 and data not shown), we used the *Eco*RI plasmid rescue from the $veg^{k07202}$ line to clone the gene. We isolated two partially overlapping *veg* cDNAs (31HC and 31HE, see Table 4) that correspond partially to the sequence (Table 4) of the GM14315 cDNA (Rubin *et al.* 2000a). The assembled nearly full-length 1.809-kb *veg* cDNA encodes an ORF of 449 aa that is identical to a predicted *CG6657* gene product (Rubin *et al.* 2000b). VEG protein has a signal peptide (aa 1–25) and seven transmembrane helices (Figure 2) and is predicted to be a type 3a protein with the N terminus located inside the cell (membrane topology prediction using PSORT, Hartmann *et al.* 1989). VEG displays similarity to predicted proteins with several transmembrane helices from Arabidopsis, fission yeast, *C. elegans*, and human (Table 4). Since a *P* element in the $veg^{k07202}$ line is inserted 309 nt upstream of the 5′-end of *veg* cDNA, it remains to be demonstrated that this *P* element affects the *veg* gene. Unexpectedly, we found that $veg^{k07228}$ *P*-element plasmid rescue is ~125 kb away from $veg^{k07202}$ as well as from the cDNA we cloned, indicating that the $veg^{k07228}$ chromosome may carry two *P*-element insertions or a rearrangement. Since $veg^{k07202}$ and $veg^{k07228}$ map at 53E1-2 and since the two *P* elements define genetically the same complementation group and are revertible, we propose that the *veg* gene maps at 53E1-2. Unfortunately, lack of deficiencies uncovering the cytological divisions 53C-E did not allow us to test this hypothesis. In conclusion, we identified a cDNA that may correspond to the *veg* gene and that defines a novel family of predicted transmembrane proteins found in organisms from yeast to human.

**Candidate cDNAs, novel proteins, and PNS development:** In summary, we present molecular characterization of 26 *P*-element-tagged mutations and demonstrate that 11 mutations are allelic to previously characterized genes. We identify 13 genes as novel on the basis of genetic and molecular analyses and present cloning of 9 novel genes. At this point the cDNAs we identified should be considered candidate cDNA clones. In many cases *P* elements directly affect the cDNAs as they are inserted in the coding sequence or in the 5′-UTR (*aay*, *bon*, *glu*, *pbl*, and *raw/cyr*) or in close proximity (<100 nt) of the 5′-end of cDNAs (*dmt*, *hoip*, and *stich1*). This strongly suggests that the genes cloned correspond to the respective mutations. However, an mRNA most proximal to the site of a *P*-element insertion may not be the one or the only one affected by a *P* element. Known examples include intronic *P* elements that disrupt a regulatory element of a distal gene without affect-

ing the most proximal gene (*e.g.*, *dlt* and α-*Spec*, Bhat *et al.* 1999) and *P*-element insertions affecting two nested genes (*e.g.*, *insc* and *sktl* or *guf* and *SmD3*, Knirr *et al.* 1997a; Hassan *et al.* 1998; Ivanov *et al.* 1998). Further molecular, phenotypic, and genetic analyses including transgenic rescue of mutant phenotypes are required to demonstrate if the genes we cloned indeed correspond to the respective mutations.

Results presented here together with our earlier observations based on genetic analyses suggest that there are few genes that function solely in the PNS. Most mutations that affect PNS development are pleiotropic. Such mutations result in phenotypes in other tissues during embryogenesis and probably are required at later stages of development. Indeed, many of the genes identified initially as novel in our screens [*e.g.*, *bun* (*shs*), *CycE* (*fond*), *gcm*, *hth* (*dtl*), *mm* (*mbl*), *pbl*, *ptc* (*rubr*), *raw* (*cyr*), *S* (*fltr*), *shn* (*quo*), *thr* (*anch*), and *unch* (*Sin3A*)] have been independently isolated in other mutagenesis screens aimed at identifying mutations affecting other cellular or developmental processes. A summary of our current (Table 5) and previous (Salzberg *et al.* 1997) results demonstrates that mutations in genes required for cell cycle or cell division (*e.g.*, *CycA*, *CycE*, *glu*, *pav*, *pbl*, *stg*, and *thr*), dorsal closure (*e.g.*, *puc* and *raw*), patterning (*e.g.*, *ptc*), signal transduction (*e.g.*, *bun*, *S*, and *shn*), or cellular metabolism (*e.g.*, *aay*, *btb*, *guf*, *hoip*, *potp*, and *RpL30*) will also affect neuronal development. In conclusion, our results provide information on types of proteins required for the PNS development and may provide a framework for future studies of biochemical and genetic interactions within the molecular pathways operating during nervous system development.

We thank the Bloomington and Umeå Stock Centers, Istvan Kiss and the Szeged Stock Center, Todd Laverty and the Berkeley Drosophila Genome Project, Bill Chia, Ed Giniger, Corey S. Goodman, Bruce A. Hamilton, Bassem Hassan, Peter Kolodziej, Marek Mlodzik, Daniel Pauli, Giuseppa Pennetta, Claire Russell, Guy Tear, and Kai Zinn for sending us fly strains, antibodies, and cDNA libraries. We acknowledge Bassem Hassan for screening a cDNA library for *gluon* cDNA and Adi Salzberg for providing us with *Sac*II, *Pst*I, and *Xba*I plasmid rescues from the *l(3)S048103 P*-element line. We also thank Nancy Van Driessche for technical assistance; Kwang-Wook Choi (Baylor College of Medicine, Houston, TX), Toru Miki (National Cancer Institute, Bethesda, MD), and Guy Tear (King's College, London, Great Britain) for sharing with us unpublished data; and Anthony L. Lau and Adi Salzberg for helpful discussions. S.N.P. was a graduate student in the Program in Developmental Biology. H.J.B. is an Investigator of the Howard Hughes Medical Institute.

## LITERATURE CITED

Adams, M. D., S. E. Celniker, R. A. Holt, C. A. Evans, J. D. Gocayne *et al.*, 2000 The genome sequence of *Drosophila melanogaster*. Science **287:** 2185–2195.

Adams, R. R., A. A. Tavares, A. Salzberg, H. J. Bellen and D. M. Glover, 1998 *pavarotti* encodes a kinesin-like protein required to organize the central spindle and contractile ring for cytokinesis. Genes Dev. **12:** 1483–1494.

Arora, K., H. Dai, S. G. Kazuko, J. Jamal, M. B. O'Connor *et al.*, 1995 The Drosophila *schnurri* gene acts in the dpp/TGF beta signaling pathway and encodes a transcription factor homologous to the human MBP family. Cell **81:** 781–790.

Artero, R., A. Prokop, N. Paricio, G. Begemann, I. Pueyo *et al.*, 1998 The *muscleblind* gene participates in the organization of Z-bands and epidermal attachments of *Drosophila* muscles and is regulated by Dmef2. Dev. Biol. **195:** 131–143.

Ashburner, M., 1989 *Drosophila: A Laboratory Handbook*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Ayer, D. E., Q. A. Lawrence and R. N. Eisenman, 1995 Mad-Max transcriptional repression is mediated by ternary complex formation with mammalian homologs of yeast repressor Sin3. Cell **80:** 767–776.

Baker, B. S., G. Hoff, T. C. Kaufman, M. F. Wolfner and T. Hazelrigg, 1991 The *doublesex* locus of *Drosophila melanogaster* and its flanking regions: a cytogenetic analysis. Genetics **127:** 125–138.

Beall, E. L., and D. C. Rio, 1997 *Drosophila* P-element transposase is a novel site-specific endonuclease. Genes Dev. **11:** 2137–2151.

Begemann, G., N. Paricio, R. Artero, I. Kiss, M. Perez-Alonso *et al.*, 1997 *muscleblind*, a gene required for photoreceptor differentiation in *Drosophila*, encodes novel nuclear $Cys_3His$-type zinc-finger-containing proteins. Development **124:** 4321–4331.

Bellen, H. J., 1999 Ten years of enhancer detection: lessons from the fly. Plant Cell **11:** 2271–2282.

Berg, C. A., and A. C. Spradling, 1991 Studies on the rate and site-specificity of *P*-element transposition. Genetics **127:** 515–524.

Bhat, M. A., A. V. Philp, D. M. Glover and H. J. Bellen, 1996 Chromatid segregation at anaphase requires the *barren* product, a novel chromosome-associated protein that interacts with Topoisomerase II. Cell **87:** 1103–1114.

Bhat, M. A., S. Izaddoost, Y. Lu, K. O. Cho, K. W. Choi *et al.*, 1999 Discs Lost, a novel multi-PDZ domain protein, establishes and maintains epithelial polarity. Cell **96:** 833–845.

Boudjelal, M., R. Taneja, S. Matsubara, P. Bouillet, P. Dolle *et al.*, 1997 Overexpression of Stra13, a novel retinoic acid-inducible gene of the basic helix-loop-helix family, inhibits mesodermal and promotes neuronal differentiation of P19 cells. Genes Dev. **11:** 2052–2065.

Breen, T. R., and I. M. Duncan, 1986 Maternal expression of genes that regulate the Bithorax complex of *Drosophila melanogaster*. Dev. Biol. **118:** 442–456.

Buff, E., A. Carmena, S. Gisselbrecht, F. Jimenez and A. M. Michelson, 1998 Signalling by the *Drosophila* epidermal growth factor receptor is required for the specification and diversification of embryonic muscle progenitors. Development **125:** 2075–2086.

Burchard, S., A. Paululat, U. Hinz and R. Renkawitz-Pohl, 1995 The mutant *not enough muscles* (*nem*) reveals reduction of the *Drosophila* embryonic muscle pattern. J. Cell Sci. **108:** 1443–1454.

Byars, C. L., K. L. Bates and A. Letsou, 1999 The dorsal-open group gene *raw* is required for restricted DJNK signaling during dorsal closure. Development **126:** 4913–4923.

Caudy, M., H. Vassin, M. Brand, R. Tuma, L. Y. Jan *et al.*, 1988 *daughterless*, a Drosophila gene essential for both neurogenesis and sex determination, has sequence similarities to myc and the *achaete-scute* complex. Cell **55:** 1061–1067.

Chen, Y., and G. Struhl, 1996 Dual roles for Patched in sequestering and transducing Hedgehog. Cell **87:** 553–563.

Chen, Y., and G. Struhl, 1998 *In vivo* evidence that Patched and Smoothened constitute distinct binding and transducing components of a Hedgehog receptor complex. Development **125:** 4943–4948.

Christodoulou, S., A. E. Lockyer, J. M. Foster, J. D. Hoheisel and D. B. Roberts, 1997 Nucleotide sequence of a *Drosophila melanogaster* cDNA encoding a calnexin homologue. Gene **191:** 143–148.

Collet, J. F., I. Gerin, M. H. Rider, M. Veiga-da-Cunha and E. Van Schaftingen, 1997 Human L-3-phosphoserine phosphatase: sequence, expression and evidence for a phosphoenzyme intermediate. FEBS Lett. **408:** 281–284.

Collet, J. F., V. Stroobant, M. Pirard, G. Delpierre and E. Van Schaftingen, 1998 A new class of phosphotransferases phosphorylated on an aspartate residue in an amino-terminal D*X*D*X*(T/V) motif. J. Biol. Chem. **273:** 14107–14112.

Collet, J. F., V. Stroobant and E. Van Schaftingen, 1999 Mechanistic studies of phosphoserine phosphatase, an enzyme related to P-type ATPases. J. Biol. Chem. **274:** 33985–33990.

Coppolino, M. G., M. J. Woodside, N. Demaurex, S. Grinstein, R. St-Arnaud *et al.*, 1997 Calreticulin is essential for integrin-mediated calcium signalling and cell adhesion. Nature **386:** 843–847.

D'Andrea, R. J., R. Stratmann, C. F. Lehner, U. P. John and R. Saint, 1993 The *three rows* gene of *Drosophila melanogaster* encodes a novel protein that is required for chromosome disjunction during mitosis. Mol. Biol. Cell **4:** 1161–1174.

Dambly-Chaudiere, C., and M. Vervoort, 1998 The bHLH genes in neural development. Int. J. Dev. Biol. **42:** 269–273.

Dawson, S. R., D. L. Turner, H. Weintraub and S. M. Parkhurst, 1995 Specificity for the Hairy/Enhancer of split basic helix-loop-helix (bHLH) proteins maps outside the bHLH domain and suggests two separable modes of transcriptional repression. Mol. Cell. Biol. **15:** 6923–6931.

Dobens, L. L., T. Hsu, V. Twombly, W. M. Gelbart, L. A. Raftery *et al.*, 1997 The *Drosophila bunched* gene is a homologue of the growth factor stimulated mammalian TSC-22 sequence and is required during oogenesis. Mech. Dev. **65:** 197–208.

Dobens, L. L., J. S. Peterson, J. Treisman and L. A. Raftery, 2000 *Drosophila bunched* integrates opposing DPP and EGF signals to set the operculum boundary. Development **127:** 745–754.

Dye, C. A., J.-K. Lee, R. C. Atkinson, R. Brewster, P.-L. Han *et al.*, 1998 The *Drosophila sanpodo* gene controls sibling cell fate and encodes a tropomodulin homolog, an actin/tropomyosin-associated protein. Development **125:** 1845–1856.

Edgar, B. A., and P. H. O'Farrell, 1989 Genetic control of cell division patterns in the *Drosophila* embryo. Cell **57:** 177–187.

Ellis, H. M., D. R. Spann and J. W. Posakony, 1990 *extramacrochaetae*, a negative regulator of sensory organ development in Drosophila, defines a new class of helix-loop-helix proteins. Cell **61:** 27–38.

Fisher, A., and M. Caudy, 1998 The function of hairy-related bHLH repressor proteins in cell fate decisions. Bioessays **20:** 298–306.

FlyBase Consortium, 1999 The FlyBase database of the Drosophila genome projects and community literature. Nucleic Acids Res. **27:** 85–88.

Gamo, S., K. Dodo, H. Matakatsu and Y. Tanaka, 1998 Molecular genetical analysis of *Drosophila* ether sensitive mutants. Toxicol. Lett. **100/101:** 329–337.

Gao, F. B., J. E. Brenman, L. Y. Jan and Y. N. Jan, 1999 Genes regulating dendritic outgrowth, branching, and routing in *Drosophila*. Genes Dev. **13:** 2549–2561.

Garrell, J., and J. Modolell, 1990 The Drosophila *extramacrochaetae* locus, an antagonist of proneural genes that, like these genes, encodes a helix-loop-helix protein. Cell **61:** 39–48.

Giniger, E., K. Tietje, L. Y. Jan and Y. N. Jan, 1994 *lola* encodes a putative transcription factor required for growth and guidance in *Drosophila*. Development **120:** 1385–1398.

Gottschalk, A., G. Neubauer, J. Banroques, M. Mann, R. Luhrmann *et al.*, 1999 Identification by mass spectrometry and functional analysis of novel proteins of the yeast [U4/U6·U5] tri-snRNP. EMBO J. **18:** 4535–4548.

Grieder, N. C., D. Nellen, R. Burke, K. Basler and M. Affolter, 1995 *schnurri* is required for Drosophila *dpp* signaling and encodes a zinc finger protein similar to the mammalian transcription factor PRDII-BF1. Cell **81:** 791–800.

Hartmann, E., T. A. Rapoport and H. F. Lodish, 1989 Predicting the orientation of eukaryotic membrane-spanning proteins. Proc. Natl. Acad. Sci. USA **86:** 5786–5790.

Hassan, B. A., S. N. Prokopenko, S. D. Breuer, B. Zhang, A. Paululat *et al.*, 1998 *skittles*, a Drosophila phosphatidylinositol 4-phosphate 5-kinase, is required for cell viability, germline development and bristle morphology, but not for neurotransmitter release. Genetics **150:** 1527–1537.

Higson, T. S., J. E. Tessiatore, S. A. Bennett, R. C. Derk and M. A. Kotarski, 1993 The molecular organization of the *Star/asteroid* region, a region necessary for proper eye development in *Drosophila melanogaster*. Genome **36:** 356–366.

Hime, G., and R. Saint, 1992 Zygotic expression of the *pebble* locus is required for cytokinesis during the postblastoderm mitoses of *Drosophila*. Development **114:** 165–171.

Hirano, T., 1999 SMC-mediated chromosome mechanics: a conserved scheme from bacteria to vertebrates? Genes Dev. **13:** 11–19.

Hirano, T., and T. J. Mitchison, 1994 A heterodimeric coiled-coil protein required for mitotic chromosome condensation in vitro. Cell **79:** 449–458.

Hirano, T., R. Kobayashi and M. Hirano, 1997 Condensins, chromosome condensation protein complexes containing XCAP-C, XCAP-E and a *Xenopus* homolog of the *Drosophila* Barren protein. Cell **89:** 511–521.

Hooper, J. E., and M. P. Scott, 1989 The Drosophila *patched* gene encodes a putative membrane protein required for segmental patterning. Cell **59:** 751–765.

Ivanov, I. P., K. Simin, A. Letsou, J. F. Atkins and R. F. Gesteland, 1998 The *Drosophila* gene for antizyme requires ribosomal frameshifting for expression and contains an intronic gene for snRNP Sm D3 on the opposite strand. Mol. Cell. Biol. **18:** 1553–1561.

Jaeken, J., M. Detheux, L. Van Maldergem, M. Foulon, H. Carchon *et al.*, 1996 3-Phosphoglycerate dehydrogenase deficiency: an inborn error of serine biosynthesis. Arch. Dis. Child. **74:** 542–545.

Jaeken, J., M. Detheux, J. P. Fryns, J. F. Collet, P. Alliet *et al.*, 1997 Phosphoserine phosphatase deficiency in a patient with Williams syndrome. J. Med. Genet. **34:** 594–596.

Jan, L. Y., and Y. N. Jan, 1993 The peripheral nervous system, pp. 1207–1244 in *The Development of Drosophila melanogaster*, edited by M. Bate and A. Martinez Arias. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Jan, Y. N., and L. Y. Jan, 1998 Asymmetric cell division. Nature **392:** 775–778.

Jimenez, J., L. Alphey, P. Nurse and D. M. Glover, 1990 Complementation of fission yeast *cdc2*$^{ts}$ and *cdc25*$^{ts}$ mutants identifies two cell cycle genes in *Drosophila*: a *cdc2* homologue and *string*. EMBO J. **9:** 3565–3571.

Jürgens, G., E. Wieschaus, C. Nüsslein-Volhard and H. Kluding, 1984 Mutations affecting the pattern of the larval cuticle in *Drosophila melanogaster*. II. Zygotic loci on the third chromosome. Roux's Arch. Dev. Biol. **193:** 283–295.

Kania, A., A. Salzberg, M. Bhat, D. D'Evelyn, Y. He *et al.*, 1995 *P*-element mutations affecting embryonic peripheral nervous system development in *Drosophila melanogaster*. Genetics **139:** 1663–1678.

Klämbt, C., J. R. Jacobs and C. S. Goodman, 1991 The midline of the Drosophila central nervous system: a model for the genetic analysis of cell fate, cell migration, and growth cone guidance. Cell **64:** 801–815.

Knirr, S., S. Breuer, A. Paululat and R. Renkawitz-Pohl, 1997a Somatic mesoderm differentiation and the development of a subset of pericardial cells depend on the *not enough muscles* (*nem*) locus, which contains the *inscuteable* gene and the intron located gene, *skittles*. Mech. Dev. **67:** 69–81.

Knirr, S., A. Santel and R. Renkawitz-Pohl, 1997b Expression of the PI4P 5-kinase *Drosophila* homologue *skittles* in the germline suggests a role in spermatogenesis and oogenesis. Dev. Genes Evol. **207:** 127–130.

Knoblich, J. A., K. Sauer, L. Jones, H. Richardson, R. Saint *et al.*, 1994 Cyclin E controls S phase progression and its down-regulation during Drosophila embryogenesis is required for the arrest of cell proliferation. Cell **77:** 107–120.

Kolodkin, A. L., D. J. Matthes and C. S. Goodman, 1993 The *semaphorin* genes encode a family of transmembrane and secreted growth cone guidance molecules. Cell **75:** 1389–1399.

Kolodkin, A. L., A. T. Pickup, D. M. Lin, C. S. Goodman and U. Banerjee, 1994 Characterization of *Star* and its interactions with *sevenless* and *EGF receptor* during photoreceptor cell development in *Drosophila*. Development **120:** 1731–1745.

Kolodziej, P. A., L. Y. Jan and Y. N. Jan, 1995 Mutations that affect the length, fasciculation, or ventral orientation of specific sensory axons in the Drosophila embryo. Neuron **15:** 273–286.

Kolodziej, P. A., L. C. Timpe, K. J. Mitchell, S. R. Fried, C. S. Goodman *et al.*, 1996 *frazzled* encodes a Drosophila member of the DCC immunoglobulin subfamily and is required for CNS and motor axon guidance. Cell **87:** 197–204.

Konev, A. I., E. R. Varentsova and I. M. Khromykh, 1991 Cytogenetic analysis of the chromosomal region containing the radiosensitivity genes of Drosophila. Influence of pericentromeric heterochromatin on mutagenesis in the 44-45 region of chromosome 2. Genetika **27:** 667–675 [in Russian].

Koshland, D., and A. Strunnikov, 1996 Mitotic chromosome condensation. Annu. Rev. Cell Dev. Biol. **12:** 305–333.

Kotarski, M. A., D. A. Leonard, S. A. Bennett, C. P. Bishop, S. D. Wahn *et al.*, 1998 The *Drosophila* gene *asteroid* encodes a novel protein and displays dosage-sensitive interactions with *Star* and *Egfr*. Genome **41:** 295–302.

Krause, K. H., and M. Michalak, 1997 Calreticulin. Cell **88:** 439–443.

Kraut, R., and J. A. Campos-Ortega, 1996 *inscuteable*, a neural precursor gene of *Drosophila*, encodes a candidate for a cytoskeleton adaptor protein. Dev. Biol. **174:** 65–81.

Langer-Safer, P. R., M. Levine and D. C. Ward, 1982 Immunological method for mapping genes on Drosophila polytene chromosomes. Proc. Natl. Acad. Sci. USA **79:** 4381–4385.

Lehner, C. F., 1992 The *pebble* gene is required for cytokinesis in *Drosophila*. J. Cell Sci. **103:** 1021–1030.

Lehner, C. F., and P. H. O'Farrell, 1989 Expression and function of Drosophila cyclin A during embryonic cell cycle progression. Cell **56:** 957–968.

Lehner, C. F., and P. H. O'Farrell, 1990 The roles of Drosophila cyclins A and B in mitotic control. Cell **61:** 535–547.

Leimeister, C., A. Externbrink, B. Klamt and M. Gessler, 1999 *Hey* genes: a novel subfamily of *hairy*- and *Enhancer of split* related genes specifically expressed during mouse embryogenesis. Mech. Dev. **85:** 173–177.

Li, Y. F., Y. Hata, T. Fujii, T. Hisano, M. Nishihara *et al.*, 1998 Crystal structures of reaction intermediates of L-2-haloacid dehalogenase and implications for the reaction mechanism. J. Biol. Chem. **273:** 15035–15044.

Lindsley, D. L., and G. G. Zimm, 1992 *The Genome of Drosophila melanogaster*. Academic Press, San Diego.

Lingrel, J. B., and T. Kuntzweiler, 1994 Na+,K(+)-ATPase. J. Biol. Chem. **269:** 19659–19662.

Lupas, A., M. Van Dyke and J. Stock, 1991 Predicting coiled coils from protein sequences. Science **252:** 1162–1164.

Maiorano, D., L. J. Brimage, D. Leroy and S. E. Kearsey, 1999 Functional conservation and cell cycle localization of the Nhp2 core component of H + ACA snoRNPs in fission and budding yeasts. Exp. Cell Res. **252:** 165–174.

Marigo, V., R. A. Davey, Y. Zuo, J. M. Cunningham and C. J. Tabin, 1996 Biochemical evidence that Patched is the Hedgehog receptor. Nature **384:** 176–179.

Martín-Blanco, E., A. Gampel, J. Ring, K. Virdee, N. Kirov *et al.*, 1998 *puckered* encodes a phosphatase that mediates a feedback loop regulating JNK activity during dorsal closure in *Drosophila*. Genes Dev. **12:** 557–570.

Modolell, J., 1997 Patterning of the adult peripheral nervous system of *Drosophila*. Perspect. Dev. Neurobiol. **4:** 285–296.

Nakagawa, O., M. Nakagawa, J. A. Richardson, E. N. Olson and D. Srivastava, 1999 HRT1, HRT2, and HRT3: a new subclass of bHLH transcription factors marking specific cardiac, somitic, and pharyngeal arch segments. Dev. Biol. **216:** 72–84.

Nakai, K., and P. Horton, 1999 PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization. Trends Biochem. Sci. **24:** 34–35.

Nakano, Y., I. Guerrero, A. Hidalgo, A. Taylor, J. R. Whittle *et al.*, 1989 A protein with several possible membrane-spanning domains encoded by the *Drosophila* segment polarity gene *patched*. Nature **341:** 508–513.

Neufeld, T. P., A. H. Tang and G. M. Rubin, 1998 A genetic screen to identify components of the *sina* signaling pathway in Drosophila eye development. Genetics **148:** 277–286.

Nielsen, H., J. Engelbrecht, S. Brunak and G. von Heijne, 1997 Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Eng. **10:** 1–6.

Nishiwaki, T., Y. Daigo, T. Kawasoe, Y. Nagasawa, H. Ishiguro *et al.*, 1999 Isolation and characterization of a human cDNA homologous to the *Xenopus laevis* XCAP-C gene belonging to the structural maintenance of chromosomes (SMC) family. J. Hum. Genet. **44:** 197–202.

Nolo, R., L. A. Abbott and H. J. Bellen, 2000 Senseless, a Zn-finger transcription factor is necessary and sufficient for sensory organ development in *Drosophila*. Cell **102:** 349–362.

Nüsslein-Volhard, C., E. Wieschaus and H. Kluding, 1984 Mutations affecting the pattern of the larval cuticle in *Drosophila melanogaster*. I. Zygotic loci on the second chromosome. Roux's Arch. Dev. Biol. **193:** 267–282.

Pai, C. Y., T. S. Kuo, T. J. Jaw, E. Kurant, C. T. Chen *et al.*, 1998 The Homothorax protein activates the nuclear localization of another homeoprotein, Extradenticle, and suppresses eye development in *Drosophila*. Genes Dev. **12:** 435–446.

Palazzolo, M. J., B. A. Hamilton, D. Ding, C. H. Martin, D. A. Mead *et al.*, 1990 Phage lambda cDNA cloning vectors for subtractive hybridization, fusion-protein synthesis and Cre-*lox*P automatic plasmid subcloning. Gene **88:** 25–36.

Pennetta, G., and D. Pauli, 1997 *stand still*, a Drosophila gene involved in the female germline for proper survival, sex determination and differentiation. Genetics **145:** 975–987.

Pennetta, G., and D. Pauli, 1998 The *Drosophila Sin3* gene encodes a widely distributed transcription factor essential for embryonic viability. Dev. Genes Evol. **208:** 531–536.

Philp, A. V., J. M. Axton, R. D. Saunders and D. M. Glover, 1993 Mutations in the *Drosophila melanogaster* gene *three rows* permit aspects of mitosis to continue in the absence of chromatid segregation. J. Cell Sci. **106:** 87–98.

Pirrotta, V., 1986 Cloning Drosophila genes, pp. 83–110 in *Drosophila: A Practical Approach*, edited by D. B. Roberts. IRL Press, Oxford.

Prober, D. A., and B. A. Edgar, 2000 Ras1 promotes cellular growth in the *Drosophila* wing. Cell **100:** 435–446.

Prokopenko, S. N., A. Brumby, L. O'Keefe, L. Prior, Y. He *et al.*, 1999 A putative exchange factor for Rho1 GTPase is required for initiation of cytokinesis in *Drosophila*. Genes Dev. **13:** 2301–2314.

Razzaq, A., Y. Su, J. E. Mehren, K. Mizuguchi, A. P. Jackson *et al.*, 2000 Characterisation of the gene for *Drosophila* amphiphysin. Gene **241:** 167–174.

Rechsteiner, M., and S. W. Rogers, 1996 PEST sequences and regulation by proteolysis. Trends Biochem. Sci. **21:** 267–271.

Richardson, H. E., L. V. O'Keefe, S. I. Reed and R. Saint, 1993 A *Drosophila* G1-specific cyclin E homolog exhibits different modes of expression during embryogenesis. Development **119:** 673–690.

Rieckhof, G. E., F. Casares, H. D. Ryoo, M. Abu-Shaar and R. S. Mann, 1997 Nuclear translocation of Extradenticle requires Homothorax, which encodes an Extradenticle-related homeodomain protein. Cell **91:** 171–183.

Ring, J. M., and A. Martinez Arias, 1993 *puckered*, a gene involved in position-specific cell differentiation in the dorsal epidermis of the *Drosophila* larva. Dev. Suppl., 251–259.

Roch, F., F. Serras, F. J. Cifuentes, M. Corominas, B. Alsina *et al.*, 1998 Screening of larval/pupal *P*-element induced lethals on the second chromosome in *Drosophila melanogaster*: clonal analysis and morphology of imaginal discs. Mol. Gen. Genet. **257:** 103–112.

Rørth, P., K. Szabo, A. Bailey, T. Laverty, J. Rehm *et al.*, 1998 Systematic gain-of-function genetics in *Drosophila*. Development **125:** 1049–1057.

Rossner, M. J., J. Dorr, P. Gass, M. H. Schwab and K. A. Nave, 1997 SHARPs: mammalian *Enhancer-of-Split*- and *Hairy*-related proteins coupled to neuronal stimulation. Mol. Cell. Neurosci. **9:** 460–475.

Rubin, G. M., L. Hong, P. Brokstein, M. Evans-Holm, E. Frise *et al.*, 2000a A *Drosophila* complementary DNA resource. Science **287:** 2222–2224.

Rubin, G. M., M. D. Yandell, J. R. Wortman, G. L. Gabor Miklos, C. R. Nelson *et al.*, 2000b Comparative genomics of the eukaryotes. Science **287:** 2204–2215.

Saito, H., T. Fujiwara, S. Shin, K. Okui and Y. Nakamura, 1996 Cloning and mapping of a human novel cDNA (NHP2L1) that encodes a protein highly homologous to yeast nuclear protein NHP2. Cytogenet. Cell Genet. **72:** 191–193.

Saka, Y., T. Sutani, Y. Yamashita, S. Saitoh, M. Takeuchi *et al.*, 1994 Fission yeast cut3 and cut14, members of a ubiquitous protein family, are required for chromosome condensation and segregation in mitosis. EMBO J. **13:** 4938–4952.

Salzberg, A., D. D'Evelyn, K. L. Schulze, J.-K. Lee, D. Strumpf *et al.*, 1994 Mutations affecting the pattern of the PNS in Drosophila reveal novel aspects of neuronal development. Neuron **13:** 269–287.

Salzberg, A., K. Golden, R. Bodmer and H. J. Bellen, 1996 *gutfeeling*, a Drosophila gene encoding an antizyme-like protein, is required for late differentiation of neurons and muscles. Genetics **144:** 183–196.

Salzberg, A., S. N. Prokopenko, Y. He, P. Tsai, M. Pál *et al.*, 1997

*P*-element insertion alleles of essential genes on the third chromosome of *Drosophila melanogaster*: mutations affecting embryonic PNS development. Genetics **147:** 1723–1741.

Sambrook, J., E. F. Fritsch and T. Maniatis, 1989 *Molecular Cloning: A Laboratory Manual*, Ed. 2. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Sasamura, T., T. Kobayashi, S. Kojima, H. Qadota, Y. Ohya *et al.*, 1997 Molecular cloning and characterization of *Drosophila* genes encoding small GTPases of the rab and rho families. Mol. Gen. Genet. **254:** 486–494.

Satoh, A. K., F. Tokunaga and K. Ozaki, 1997 Rab proteins of *Drosophila melanogaster*: novel members of the Rab-protein family. FEBS Lett. **404:** 65–69.

Schreiber-Agus, N., and R. A. DePinho, 1998 Repression by the Mad(Mxi1)-Sin3 complex. Bioessays **20:** 808–818.

Schultz, J., F. Milpetz, P. Bork and C. P. Ponting, 1998 SMART, a simple modular architecture research tool: identification of signaling domains. Proc. Natl. Acad. Sci. USA **95:** 5857–5864.

Schultz, J., R. R. Copley, T. Doerks, C. P. Ponting and P. Bork, 2000 SMART: a web-based tool for the study of genetically mobile domains. Nucleic Acids Res. **28:** 231–234.

Shen, M., T. Kawamoto, W. Yan, K. Nakamasu, M. Tamagami *et al.*, 1997 Molecular characterization of the novel basic helix-loop-helix protein DEC1 expressed in differentiated human embryo chondrocytes. Biochem. Biophys. Res. Commun. **236:** 294–298.

Skeath, J. B., and C. Q. Doe, 1998 Sanpodo and Notch act in opposition to Numb to distinguish sibling neuron fates in the Drosophila CNS. Development **125:** 1857–1865.

Smith, M. J., 1992 Nucleotide sequence of a *Drosophila melanogaster* gene encoding a calreticulin homologue. DNA Seq. **3:** 247–250.

Smith, Q. R., S. Momma, M. Aoyagi and S. I. Rapoport, 1987 Kinetics of neutral amino acid transport across the blood-brain barrier. J. Neurochem. **49:** 1651–1658.

Sonnhammer, E. L. L., G. von Heijne and A. Krogh, 1998 A hidden Markov model for predicting transmembrane helices in protein sequences, pp. 175–182 in *Proceedings of the Sixth International Conference on Intelligent Systems for Molecular Biology*, edited by J. Glasgow. AAAI Press.

Spradling, A. C., D. Stern, I. Kiss, J. Roote, T. Laverty *et al.*, 1995 Gene disruptions using *P* transposable elements: an integral component of the *Drosophila* genome project. Proc. Natl. Acad. Sci. USA **92:** 10824–10830.

Spradling, A. C., D. Stern, A. Beaton, E. J. Rhem, T. Laverty *et al.*, 1999 The Berkeley Drosophila Genome Project gene disruption project: single *P*-element insertions mutating 25% of vital Drosophila genes. Genetics **153:** 135–177.

Staehling-Hampton, K., A. S. Laughon and F. M. Hoffmann, 1995 A *Drosophila* protein related to the human zinc finger transcription factor PRDII/MBPI/HIV-EP1 is required for *dpp* signaling. Development **121:** 3393–3403.

Strunnikov, A. V., 1998 SMC proteins and chromosome structure. Trends Cell Biol. **8:** 454–459.

Sturtevant, A. H., 1948 New mutants report. Dros. Inf. Serv. **22:** 55–56.

Su, T. T., and P. H. O'Farrell, 1998 Chromosome association of minichromosome maintenance proteins in *Drosophila* endoreplication cycles. J. Cell Biol. **140:** 451–460.

Sutani, T., T. Yuasa, T. Tomonaga, N. Dohmae, K. Takio *et al.*, 1999 Fission yeast condensin complex: essential roles of non-SMC subunits for condensation and Cdc2 phosphorylation of Cut3/SMC4. Genes Dev. **13:** 2271–2283.

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

Tearle, R. G., and C. Nüsslein-Volhard, 1987 Tübingen mutants and stock list. Dros. Inf. Serv. **66:** 209–269.

Török, T., G. Tick, M. Alvarado and I. Kiss, 1993 *P-lacW* insertional mutagenesis on the second chromosome of *Drosophila melanogaster*: isolation of lethals with different overgrowth phenotypes. Genetics **135:** 71–80.

Treisman, J. E., Z. C. Lai and G. M. Rubin, 1995 *shortsighted* acts in the *decapentaplegic* pathway in *Drosophila* eye development and has homology to a mouse TGF-beta-responsive gene. Development **121:** 2835–2845.

Uemura, T., S. Shepherd, L. Ackerman, L. Y. Jan and Y. N. Jan, 1989 *numb*, a gene required in determination of cell fate during sensory organ formation in *Drosophila* embryos. Cell **58:** 349–360.

Wheeler, D. L., C. Chappey, A. E. Lash, D. D. Leipe, T. L. Madden *et al.*, 2000 Database resources of the National Center for Biotechnology Information. Nucleic Acids Res. **28:** 10–14.

Whiteley, M., P. D. Noguchi, S. M. Sensabaugh, W. F. Odenwald and J. A. Kassis, 1992 The Drosophila gene *escargot* encodes a zinc finger motif found in *snail*-related genes. Mech. Dev. **36:** 117–127.

Zaffran, S., 2000 Molecular cloning and embryonic expression of *dFKBP59*, a novel *Drosophila* FK506-binding protein. Gene **246:** 103–109.

Zerial, M., and L. A. Huber (Editors), 1995 *Guidebook to the Small GTPases*. Oxford University Press, Oxford.

Zhang, N., J. Zhang, Y. Cheng and K. Howard, 1996 Identification and genetic analysis of *wunen*, a gene guiding *Drosophila melanogaster* germ cell migration. Genetics **143:** 1231–1241.

Zinn, K., L. McAllister and C. S. Goodman, 1988 Sequence analysis and neuronal expression of fasciclin I in grasshopper and Drosophila. Cell **58:** 577–587.

Communicating editor: T. F. C. Mackay