Development 120, 3367-3377 (1994)
Printed in Great Britain © The Company of Biologists Limited 1994

diaphanous is required for cytokinesis in *Drosophila* and shares domains of similarity with the products of the *limb deformity* gene

Diego H. Castrillon and Steven A. Wasserman

Department of Biochemistry, University of Texas Southwestern Medical Center, Dallas, Texas 75235-9038, USA

SUMMARY

We show that the *Drosophila* gene *diaphanous* is required for cytokinesis. Males homozygous for the *dia*^{1} mutation are sterile due to a defect in cytokinesis in the germline. Females *trans*-heterozygous for *dia*^{1} and a deficiency are sterile and lay eggs with defective eggshells; failure of cytokinesis is observed in the follicle cell layer. Null alleles are lethal. Death occurs at the onset of pupation due to the absence of imaginal discs. Mitotic figures in larval neuroblasts were found to be polyploid, apparently due to a defect in cytokinesis. The predicted 123×10^{3} M_r protein contains two domains shared by the formin proteins, encoded by the *limb deformity* gene in the mouse. These formin homology domains, which we have termed FH1 and FH2, are also found in Bni1p, the product of a *Saccharomyces cerevisiae* gene required for normal cytokinesis in diploid yeast cells.

Key words: cytokinesis, mitosis, spermatogenesis, oogenesis, P element, *Drosophila*

INTRODUCTION

To complete mitosis successfully, a cell must carry out two separate processes: karyokinesis, or chromosome segregation, and cytokinesis, or the division of the entire cell. In animal cells, cytokinesis is believed to be mediated by the contractile ring, a transient cytoskeletal structure located midway between the spindle poles at the cell cortex. Actin and myosin, which are concentrated at the cleavage furrow in a variety of cells, are believed to generate the force leading to contraction of the ring and, ultimately, the separation into two daughter cells (Satterwhite and Pollard, 1992).

Despite increasing knowledge about mitosis and the cell cycle, many aspects of cytokinesis remain poorly understood. For example, it is not clear how actin and myosin are recruited to the contractile ring, nor how the contractile ring is attached to the cell membrane. The position of the contractile ring is believed to be determined by a signal originating at the mitotic spindle (Rappaport, 1986), but the nature of this signal is not known. Lastly, few of the structural and regulatory proteins involved in cytokinesis have been identified.

The isolation of mutations that disrupt cytokinesis in organisms such as yeast or *Drosophila* promises to be a useful approach for identifying genes required for cytokinesis. Indeed, mutations identified through genetic screens in *Drosophila* (Karess et al., 1991), or through a 'reverse genetic' approach in *S. cerevisiae* (Watts et al., 1987) and *Dictyostelium* (DeLozanne and Spudich, 1987), have demonstrated that myosin function is required for cytokinesis. More recently, the *peanut* locus has been shown to be required for cytokinesis in *Drosophila* and to encode a homolog of Cdc12 and related proteins required for cytokinesis in yeast (Neufeld and Rubin, 1994).

The *diaphanous* (*dia*) locus was identified in a P element screen for recessive male-sterile mutations (Castrillon et al., 1993). Here we present evidence that *dia* is generally required for cytokinesis. A failure in cytokinesis during spermatogenesis results in multinucleate spermatids. A defect in oogenesis in females is associated with a failure of cytokinesis in the somatically derived follicle cells, which surround each developing egg chamber. We further show that null *dia* alleles result in early pupal lethality and that the lethal phenotype is consistent with a defect in cytokinesis. Lastly, we have defined two evolutionarily conserved domains that are common to the protein products of *diaphanous*, the *limb deformity* locus of mouse and chicken, and *BNI1*, a yeast gene identified on the basis of its interaction with *CDC12*.

MATERIALS AND METHODS

Drosophila markers and manipulations

All crosses were performed at 25°C on yeasted cornmeal molasses agar. Genetic markers and balancers are described and referenced by Lindsley and Zimm (1992). Lethal *dia* mutations were balanced with *In(2LR)Gla, Bc*. RNA from germlineless flies was prepared from the progeny of females homozygous for the *tud*^{1} mutation.

Remobilization of P element to generate new alleles

The *dia*^{1} chromosome was brought together with a transposase source by crossing *dia*^{1}/*CyO* flies to flies carrying the *P[ry+, Δ2-3]* transposase source on the third chromosome (Robertson et al., 1987). In the next generation, the *P[ry+, Δ2-3]* chromosome was crossed out, and *ry*$^{-}$ derivatives of the *dia*^{1} chromosome were selected to establish lines.

Of the 226 *rosy*$^{-}$ lines generated, 17 were homozygous lethal. The results of several genetic tests argue that these events are lethal alleles

of the *dia* locus: (1) the cytologically visible deficiency *Df(2L)TW84*, which deletes the entire *dia* region, fails to complement any of these lethal mutations; (2) complementation tests indicate that these 17 lethal events are due to mutations in the same locus; (3) the lethal events fail to complement the original dia^1 allele and (4) a genomic fragment containing a single intact transcription unit can fully rescue the male-sterile and lethal phenotypes (discussed in more detail below).

Nucleic acid manipulations and analysis

Standard protocols were performed as described (Sambrook et al., 1989). For plasmid rescue, genomic DNA prepared from adults homozygous for dia^1 was digested to completion with *Xba*I and *Nhe*I, self-ligated and transformed into competent *E. coli* cells following a standard protocol (Ashburner, 1989) with selection for kanamycin resistance. The 140 bp fragment from one of the plasmid rescue clones was gel-purified, labelled and used to screen a *Drosophila* genomic library (EMBL3 vector) provided by J. Tamkun. The 3.6 kb *dia* cDNA was obtained by screening a 12-24 hour embryonic library (Brown and Kafatos, 1988). *fod* cDNAs were cloned from an adult testis library provided by T. Hazelrigg.

RNA amplification (RT-PCR) was performed as described in Kawasaki (1990). Synthetic oligonucleotides were synthesized to the 5′ end of the 3.6 kb cDNA and to the 140 bp fragment flanking the P element insertion site (see Fig. 7). The synthetic oligonucleotide complimentary to the genomic 140 bp fragment is 5′-ATGGTACCCAA-GAAAAAGTGTTCGGAGGG-3′. The first 8 bases were added to create an *Asp*718I/*Kpn*I site to facilitate subcloning. The sequence of the oligonucleotide complementary to the 5′ end of the cDNA is 5′-CTTCCGGGGATCCAGGGATTG-3′. This oligonucleotide was designed around a naturally occurring *Bam*HI site. 20 picomoles of each oligonucleotide were used per PCR reaction.

To avoid the complication of sequence polymorphisms that are common in non-isogenized *Drosophila* stocks, RT-PCR was performed on RNA from adult flies carrying *Df(2L)TW84*, which deletes the entire *dia* locus. 0.3 μg of total RNA from *Df(2L)TW84*/*CyO* adults or 0.3 μg of poly(A)$^+$ RNA from 8-16 hour embryos (Oregon-R) were used for cDNA synthesis preceding PCR. 75 picomoles of random hexanucleotides were used to prime cDNA synthesis by M-MLV reverse transcriptase. Amplification was carried out with 2.5 units of Taq polymerase in 35 cycles on a Perkin-Elmer-Cetus DNA thermal cycler using the following parameters: (1) raise temperature to 95°C, then hold 15 seconds; (2) lower to 55°C, then hold 30 seconds; (3) raise to 72°C, then hold 30 seconds. A total of four polymorphisms were identified among the *Df(2L)TW84*/*CyO* PCR products sequenced. Sequencing of an additional embryonic RT-PCR clone revealed no significant differences.

The 3.6 kb *dia* cDNA and various PCR products were subcloned into M13 mp18 and mp19 for sequencing by the dideoxy chain termination method. Sequences were assembled using Assemblylign (IBI/Kodak) and analyzed using the BLASTP sequence comparison program (Altschul et al., 1990). The BLOSUM62 matrix was used for scoring.

Predictions of coiled-coil structure from protein sequence were performed using the algorithm of Lupas et al. (1991). The mean score for globular proteins and coiled-coil sequences calculated by this method are 0.77 and 1.63, respectively.

Germline transformation and phenotypic rescue

pDC4 was constructed by ligating the 11 kb *Sal*I fragment from λ4138R to the w^+ P element vector pCaSpeR 4 (Pirrotta, 1988) digested with *Xho*I. P element-mediated germline transformation into a w^{1118} background was performed as described by Spradling (1986), except that a coinjected plasmid carrying *P[ry+, Δ2-3]* (Robertson et al., 1987) served as the transposase source. Genetic crosses were carried out to generate mutant flies carrying one copy of the *P*[w^+, pDC4] third chromosome. The second chromosomes of all *dia* alleles described in this paper are marked with *cn*, which allowed us to confirm that the rescued flies were mutant for *dia*.

Double-labelling of follicle cells

All steps were carried out at room temperature. Ovaries from 5-day-old females were dissected out and ovarioles were teased apart in PBS, fixed in PBT (PBT is PBS, 0.1% Tween-20) + 3.7% formaldehyde for 30 minutes, and washed 3× 5 minutes in PBT. To minimize non-specific propidium iodide staining, the ovaries were incubated overnight in PBS + 10 μg/ml RNAse (shorter incubations may be sufficient), and washed 3× 5 minutes in PBT. The ovaries were then stained in PBT + 25 μg/ml BODIPY-Concanavalin A (Molecular Probes) + 0.5 μg/ml propidium iodide (Sigma). Following three 10 minute washes, the ovaries were viewed immediately. Concanavalin A strongly binds to the epithelial sheath of the ovariole, which is closely apposed to the follicle cell layer; the follicle cells could be visualized by focusing just below this layer.

Cytological examination of dividing neuroblasts

Aceto-orcein squashes of the larval CNS were performed as described by Karess and Glover (1989).

RESULTS

Spermatogenesis phenotype

The male-sterile dia^1 allele was identified in a P element screen for mutations affecting spermatogenesis. Although both spermatocytes and spermatids are initially present in dia^1 testes, these cells degenerate and are not replenished. By 5 days after eclosion, most mutant testes are devoid of germinal contents. Female fertility is not significantly affected, and the viability of both males and females is normal. The locus maps to polytene interval 38E (Castrillon et al., 1993).

During spermatogenesis, a spermatogonium (the product of a stem cell division) undergoes four rounds of mitotic division to give rise to a cyst of 16 spermatocytes; meiosis then produces a cyst of 64 haploid spermatids. Wild-type chromosomal segregation and cytokinesis result in spermatids that each contain two major cytological structures of identical size and shape (Fig. 1A): a pale round nucleus (arrowhead), and an adjacent dark nebenkern (arrow). The nebenkern results from the fusion of all the mitochondria in a spermatid (see Fuller, 1993); its size thus serves as a marker for the amount of cytoplasm inherited by a spermatid.

The contents of testes from 50 newly eclosed dia^1 males were examined. dia^1 testes contain far fewer spermatids than normal, due to the reduced germinal content. The majority of spermatids were large and multinucleate (Fig. 1B). These abnormal spermatids contained either 2, 4 or 8 nuclei, with the size of the nebenkerne proportional to the number of nuclei. Of 159 unelongated spermatids identified, 51 (32%) contained one nucleus (phenotypically normal), 26 (16%) contained two nuclei, 81 (51%) contained four nuclei, and 1 (1%) contained 8 nuclei.

The presence of spermatids containing 2 or 4 nuclei within a common cytoplasm can be explained by a failure in cytokinesis in one or both of the meiotic divisions. Likewise, the rare spermatid containing 8 nuclei can be explained by a failure of cytokinesis in three consecutive cell divisions, the first being the mitotic division preceding meiosis. The nuclei in defective spermatids are almost always of wild-type size, indicating that

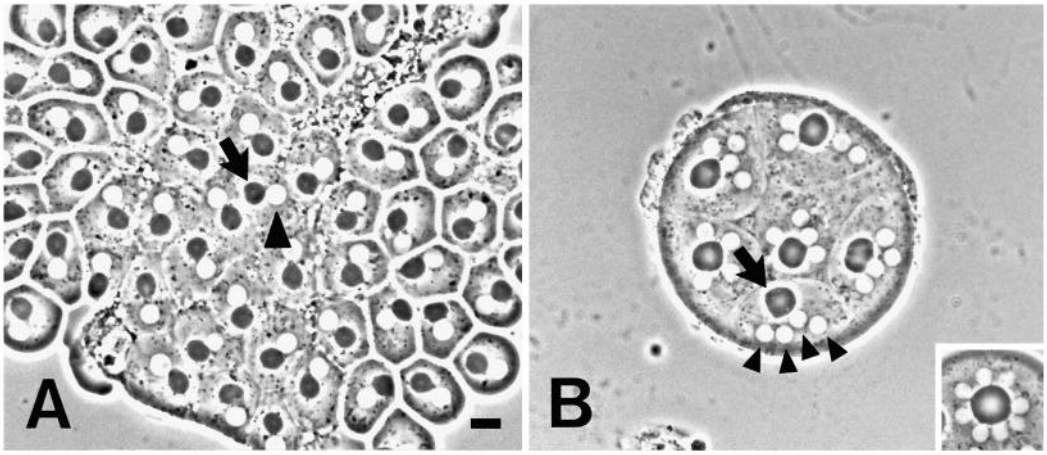

Fig. 1. Cytokinesis defect in *dia*^{1} testis. Photographs are of unfixed testis contents visualized by phase-contrast microscopy. Bar represents 10 μm for both A and B. (A) Part of a 64-cell cyst of wild-type spermatids. Each spermatid contains a single pale nucleus (arrowhead) and a single dark nebenkern (arrow). Although this cyst is intact, spermatid cysts typically rupture into smaller groups of cells due to the absence of a fixation step. (B) Group of six *dia*^{1} spermatids, each containing four nuclei (arrowheads) associated with a single large nebenkern (arrow). Inset: single *dia*^{1} spermatid containing eight nuclei.

chromosome segregation is normal in spite of the failure of cytokinesis. In contrast, in mutants that cause nondisjunction during meiosis, nuclei are of variable size (Gonzalez et al., 1989; Karess and Glover, 1989).

The finding that the initial failure of cytokinesis can occur at distinct points along the spermatogenesis pathway suggests why the germline in *dia*^{1} testes is eventually depleted. Since the 5-9 stem cells present in each testis continually divide to give rise to spermatogonia (Hardy et al., 1979), failure of cytokinesis during stem cell divisions should result in the permanent inactivation of these cells.

Oogenesis phenotype

In *trans* to a chromosomal deficiency or a null allele such as *dia*^{2}, the *dia*^{1} allele exhibits an oogenesis phenotype. Although the viability of such *trans*-heterozygous adults is normal and the male germline phenotype is similar to that of *dia*^{1} males, female fertility is dramatically decreased. The ovaries of *dia*^{1}/*dia*^{2} females are smaller than wild type. Egg chambers of all stages are present, the great majority of which contain 15 nurse cells and one oocyte. However, the eggs laid by *dia*^{1}/*dia*^{2} females are shorter than wild-type and have short, fused, or extra dorsal appendages (Fig. 2). Only 10% of the eggs hatch.

This eggshell phenotype suggested a defect in the somatically derived follicle cells, which surround each developing egg chamber and secrete the eggshell. Indeed, follicle cells in mutant ovaries have an abnormal appearance when viewed by differential interference contrast (DIC) microscopy (Fig. 3B). Their nuclei vary considerably in size (arrowheads) relative to wild-type controls (Fig. 3A). Some of the cells appear to contain two nuclei (arrows).

To visualize the follicle cell layer better, ovaries were double labelled with a fluorescent Concanavalin A derivative (to stain plasma membranes) and with propidium iodide (to stain nuclei) and examined by confocal microscopy. Individual follicle cells containing two nuclei were frequently observed in egg chambers from mutant mothers (Fig. 3D, arrows); such cells were not found in wild-type egg chambers (Fig. 3C). Therefore, it appears that cytokinesis fails to occur in some follicle cells. The abnormally large nuclei seen in Fig. 3B are likely due to nuclear fusion following the failure of cytokinesis.

Null mutations result in early pupal death and absence of imaginal discs

Lethal *dia* mutations, including null mutations, were generated by imprecise excision of the P element in the *dia*^{1} allele. Null alleles such as *dia*^{2} result in early pupal lethality and absence of imaginal discs. This phenotype is consistent with *dia* being an essential mitotic gene. Due to the presence of maternal gene products, an embryo with a null mutation in such a gene can develop into a larva (Gatti and Baker, 1989). However, since imaginal disc cells divide during the larval stages, by which time zygotic gene expression is required,

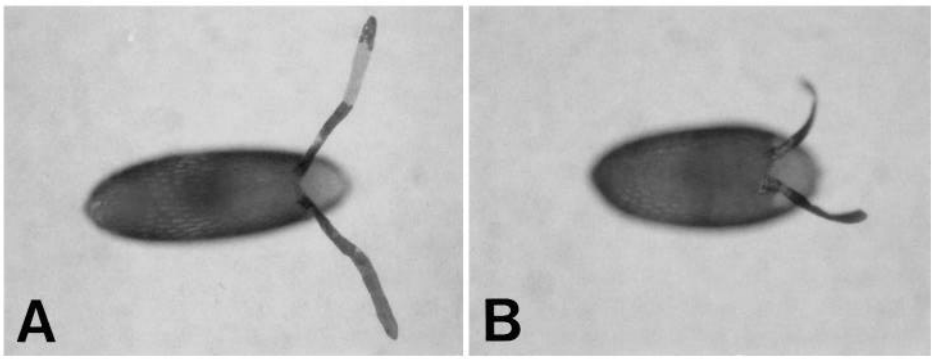

Fig. 2. Short egg phenotype. (A) Wild-type egg; (B) egg laid by *dia*^{1}/*dia*^{2} mother. Eggs from *dia*^{1}/*dia*^{2} females are significantly shorter and somewhat wider than wild type and have smaller dorsal appendages. This is an example of a mildly affected egg.

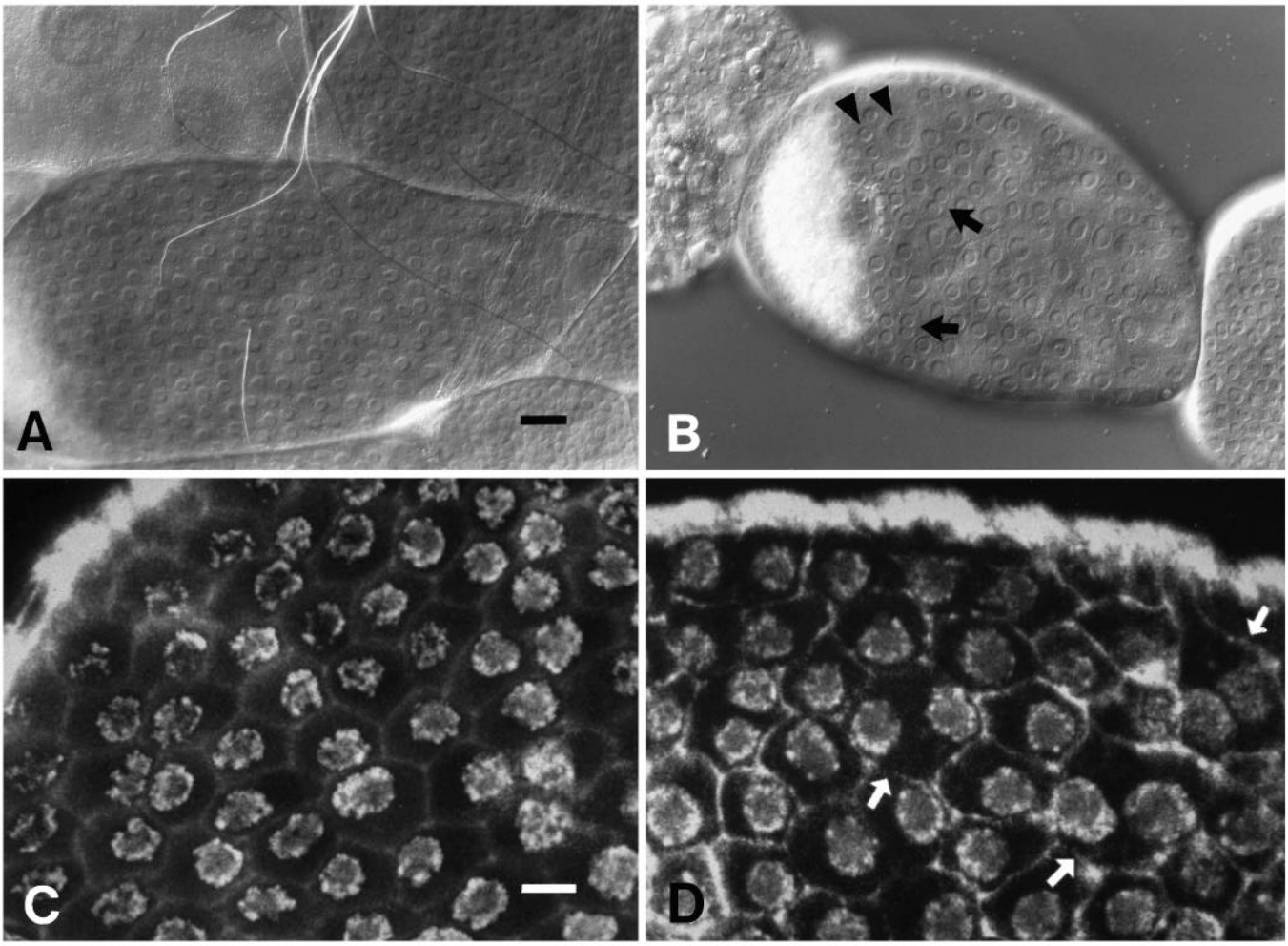

Fig. 3. Cytokinesis defect in oogenesis. Top two panels are of unfixed ovaries flattened under a coverslip and viewed with DIC; bar represents 20 μm. Bottom two panels are of fixed ovaries labelled with BODIPY-Concanavalin A and propidium iodide viewed by confocal microscopy; bar represents 5 μm. (A) Wild-type follicle cells. Nuclei are of equal size. The large egg chamber occupying the lower two-thirds of the field is at stage 8. (B) dia^1/dia^2 follicle cells, stage 8 egg chamber. Nuclei are of variable size (arrowheads). Some cells appear to contain more than one nucleus (arrows). The cytokinesis defect is apparent in earlier egg chambers, but is more easily observed in larger, more mature egg chambers. (C) Wild-type follicle cells. Each cell contains one nucleus. (D) dia^1/dia^2 follicle cells. Some cells contain two nuclei (arrows). Note that the nuclei within such a binucleate cell are of equal size.

such larvae will have defective or absent discs and will die at the onset of pupation.

Homozygotes for less severe alleles such as dia^3 also die as early pupae, but third instar larvae contain imaginal discs, albeit somewhat smaller than normal. Homozygotes for the weakest lethal allele, dia^9, have imaginal discs of normal appearance but die as late pupae or pharate adults. Very few (less than 1%) of dia^9 flies eclose. These flies are sickly and have a weak 'rough eye' phenotype (not shown), consistent with a mitotic defect that affects a small fraction of cells. Testes from these surviving adults are almost completely devoid of germinal content.

Polyploidy in dividing neuroblasts from the larval CNS

The larval central nervous system (brain and ventral ganglion) is a rich source of dividing cells (neuroblasts) and is the most suitable tissue for examining mitosis in larvae. Chromosome morphology and segregation can be examined in aceto-orcein-stained preparations of the larval CNS (Gatti et al., 1974). In wild type, a mitotic figure consists of 3 pairs of major chromosomes (Fig. 4A, arrowheads).

In larvae homozygous for a weak lethal allele, dia^9, only a small fraction of neuroblasts are polyploid and the number of mitotic figures is not affected (Table 1). Homozygotes for a stronger allele, dia^3, exhibit ploidy ranging from 2n, 4n and 8n to extreme hyperploidy (Fig. 4B,C). In addition, significantly fewer mitotic figures are present in dia^3 homozygotes than in wild type (Table 1). However, the fraction of mitotic figures in anaphase is the same as in wild type (see Table 1), whereas mutations that disrupt spindle function result in a decrease in the number of anaphases (Gatti and Baker, 1989). In addition, chromosome morphology is generally normal, although a small number of mitotic figures contain highly condensed chromosomes. Homozygotes for a null allele, dia^2, have very few mitotic figures (Table 1), and all are

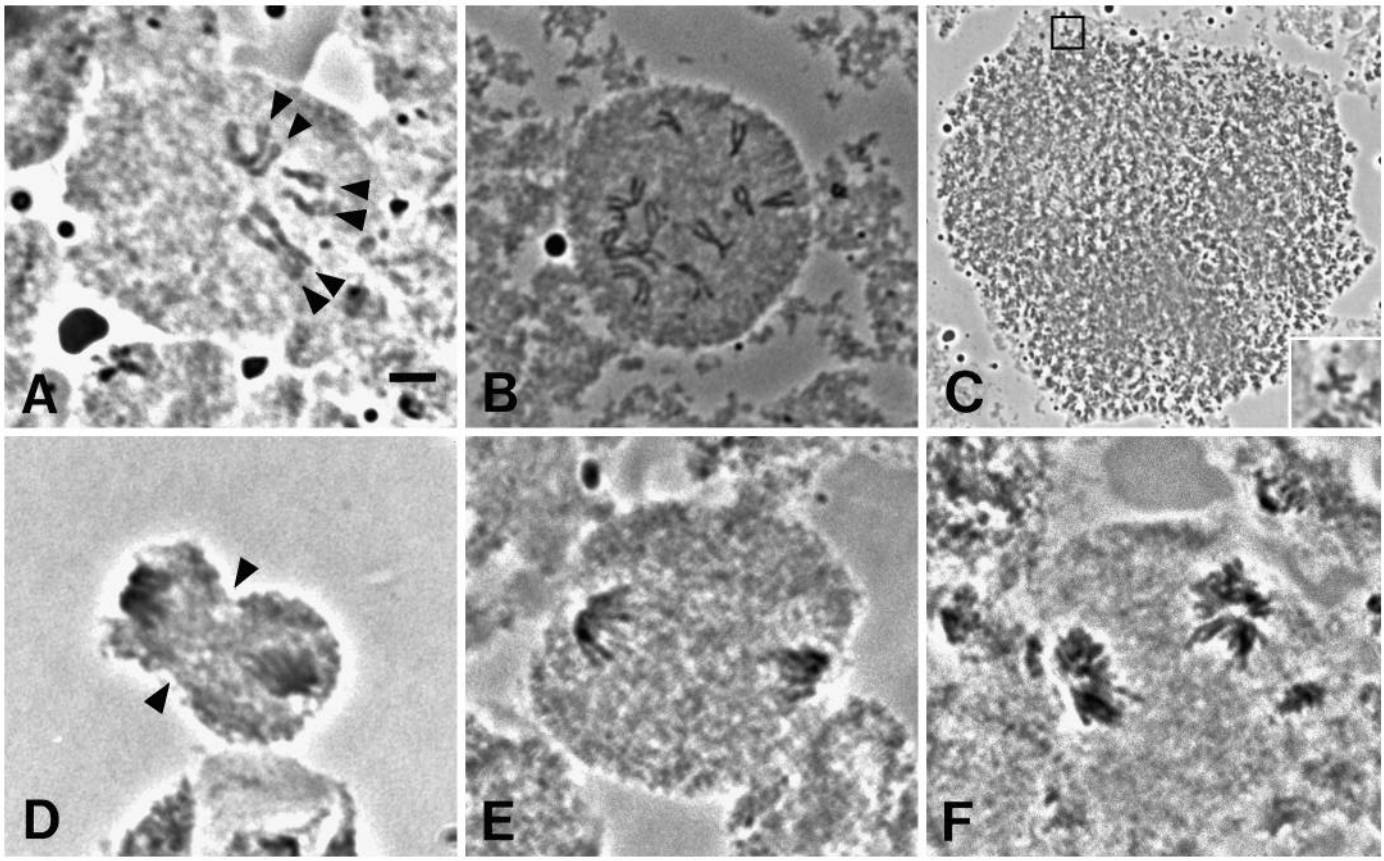

Fig. 4. Aceto-orcein stained mitotic chromosomes from larval neuroblasts. Cells in A and D are from dia^2/+ larvae; cells in B, C, E, and F are from dia^3/dia^3 larvae. A, D, E and F are at same magnification; bar represents 3 μm. For B, bar represents 5 μm, and for C, bar represents 18 μm. (A) dia^2/+. Wild-type mitotic figure consists of three pairs of major chromosomes and a pair of small fourth chromosomes which are not readily apparent in this micrograph. (B) dia^3/dia^3 tetraploid cell containing 12 major chromosomes. (C) dia^3/dia^3 hyperploid cell. This mitotic figure consists of hundreds of chromosomes packed together. Inset: magnification of boxed region, revealing individual chromosome. (D) dia^2/+. Wild-type anaphase figure. Cleavage furrow is evident (arrowheads). (E) dia^3/dia^3. Cleavage furrow is not evident in this mutant cell in anaphase. (F) dia^3/dia^3. Even though anaphase is clearly hyperploid, chromosomes are being segregated equally, indicating a functioning spindle. Again, cleavage furrow is absent.

Table 1. Quantitation of mitotic defects observed in *dia* mutants

Genotype	Mitotic figures per brain	Ploidy				% Anaphase
		%2n	%4n	%8n	%>8n	
dia^2/+	215 (32)	100	0	0	0	18
dia^9/dia^9	222 (26)	97	3	<1	0	17
dia^3/dia^3	76 (21)	11	28	17	44	20
dia^2/dia^2	5 (1)	0	0	0	100	ND

For each genotype, mitotic figures from five larvae were counted and averaged; standard deviation is shown in parentheses. Due to the high level of polyploidy, the exact number of chromosomes in most abnormal mitotic figures could not be determined. Brain size was comparable for all genotypes. ND=not determined.

enormously hyperploid, similar to the example shown in Fig. 4C.

The morphology of anaphase figures provides direct evidence that cytokinesis is defective in dia^3 cells. Cleavage furrows are sometimes evident in wild-type anaphase figures, especially in well-isolated cells (Fig. 4D). Cleavage furrows have not been observed, however, in any dia^3 anaphases and the cells appear completely round even when the chromosomes have finished migrating to opposite poles (Fig. 4E).

Despite the cytokinesis defect, chromosome segregation appears to be relatively normal in polyploid cells. In bipolar dia^3 anaphases, the spindles are well organized and the chromosomes are equally segregated, with no lagging chromosomes or other abnormalities. This is true even in anaphases that are clearly hyperploid (Fig. 4F). In more extremely hyperploid cells, anaphases are typically multipolar. Such multipolar spindles have also been observed in other mutants that produce hyperploid cells (Gatti and Baker, 1989; Karess et al., 1991).

Taken together, the male-sterile, female-sterile and lethal phenotypes associated with *dia* mutations demonstrate that *dia* is required for cytokinesis in both the soma and germline and in mitosis as well as meiosis.

Cloning and characterization of *dia* genomic region

A 140 bp fragment of DNA extending from the 5′ end of the

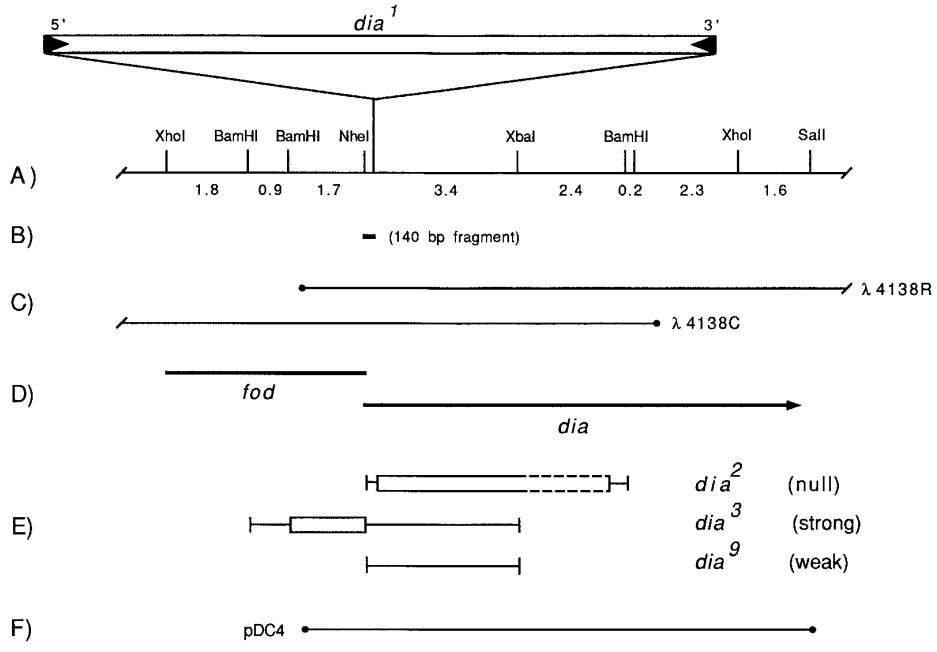

Fig. 5. Schematic diagram of the *diaphanous* region. (A) Restriction enzyme map of the *dia* region. Restriction fragment sizes are in kilobases. The *dia*^{1} allele is an insertion of a P element (drawn to scale) within the *dia* locus. The exact site of insertion was determined by sequencing the 140 bp fragment and both ends of the 3.4 kb *Nhe*I-*Xba*I fragment into which the P element inserted. (B) 140 bp genomic fragment obtained by plasmid rescue. (C) Genomic clones λ4138C and λ4138R. Only relevant portion of each clone is shown on this map. (D) Genomic extent of *fod* and *dia* transcription units. This was determined by northern analysis and hybridization of cloned cDNAs to the genomic clones. The direction of *dia* transcription (from left to right) was determined by hybridization of cDNA subfragments to the genomic clones. The direction of *fod* transcription was not determined. The intron/exon boundaries of *dia* are not known but, based on its size, the gene must contain at least one intron. (E) Extent of genomic deletions associated with lethal alleles. For each allele, a genomic probe which failed to hybridize (because of a complete deletion of that genomic DNA) is indicated by a box; a probe which detected a polymorphism relative to a progenitor strain (due to a deletion breakpoint within the DNA corresponding to that fragment) is marked by an error bar. For example, the deletion associated with *dia*^{9} could be no larger than 3.4 kb, but could be substantially smaller. For *dia*^{2}, the extent of DNA estimated to be deleted from the 2.4 kb *Xba*I-*Bam*HI fragment is indicated by dashed lines. (F) 11 kb genomic fragment from λ4138R used to make pDC4 construct. This fragment extends from one end of the genomic insert in λ4138R to a genomic *Sal*I site.

P element to a genomic *Nhe*I site was cloned by plasmid rescue (Fig. 5A, B). Overlapping genomic clones of the *diaphanous* region (Fig. 5C) were obtained by screening a genomic library with the 140 bp fragment.

Northern analysis revealed two transcription units near the P insertion site; both were found to be expressed throughout development (data not shown). cDNAs for both genes were obtained and mapped onto the genomic clones. As discussed below, we showed by germline transformation with a genomic fragment that one transcription unit represents *dia*; the other we have named *friend of diaphanous* (*fod*) (Fig. 5D). The P insertion disrupts the 5′ untranslated end of the *dia* gene, since both the 140 bp fragment and fragments to the right of the insertion site detect the *dia* but not the *fod* transcripts (Fig. 5).

Mapping of *dia* P element excision events

The extent of the genomic deletion associated with each lethal allele was determined by Southern analysis using DNA prepared from homozygous third instar larvae; the results for three representative alleles are shown in Fig. 5E. The 3.4 kb *Nhe*I-*Xba*I fragment is deleted in the *dia*^{2} allele. Moreover, the polymorphic band detected when *dia*^{2} DNA was probed with the 2.4 kb *Xba*I-*Bam*HI fragment was very faint, indicating that almost all of this genomic DNA is deleted. Thus, at least 5 kb of DNA from the *dia* locus are deleted by the *dia*^{2} mutation.

The *dia*^{2} allele, either when homozygous or in *trans* to *Df(2L)TW84*, results in early pupal lethality, complete absence of imaginal discs and extremely hyperploid neuroblasts. Although several alleles among the 17 lethal *dia* mutations recovered have a similar phenotype, none is more severe. Furthermore, the female-sterile phenotype of *dia*^{1} in *trans* to *dia*^{2} is as strong as that observed in *dia*^{1}/*Df(2L)TW84* females. On the basis of both the genetic and molecular data, we conclude that *dia*^{2} represents a null allele.

Northern analysis of *dia*

In wild-type adult males and females, two *dia* transcripts of 4.4 and 4.8 kb are observed in northern analysis (Fig. 6, lanes 1 and 2). In adults homozygous for *dia*^{1}, two *dia* transcripts are also observed, but they are larger than the corresponding wild-type transcripts (Fig. 6, lane 3). Reversion to the wild-type pattern is observed in a phenotypic revertant line in which remobilization of the P element resulted in its precise excision (Fig. 6, lane 4). Both *dia* transcripts are present in flies that lack a germline (Fig. 6, lane 5); therefore, the expression of neither *dia* transcript is limited to the germline. Of the two *fod* transcripts, the 3.5 kb species, which is specific to the male germline (Fig. 6, compare lanes 1, 2 and 5) is absent in *dia*^{1} flies (Fig. 6, lane 3); this is expected, however, since mature males of this genotype have no germline (Castrillon et al., 1993).

Germline transformation and phenotypic rescue

P element-mediated germline transformation allowed us to confirm that the disrupted gene is *dia*. An 11 kb fragment of genomic DNA was subcloned into the *Drosophila* germline transformation vector pCaSpeR 4 (Pirrotta, 1988) to yield the recombinant plasmid pDC4 (Fig. 5F). The 11 kb fragment encodes the entire *dia* transcription unit. A portion of the *fod* transcription unit also lies within the transgene, but is not

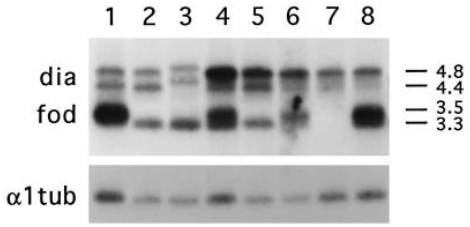

Fig. 6. Effect of *dia* mutations on *fod* and *dia* transcripts. Autoradiograph of northern blot probed with *dia* and *fod* cDNAs. Each lane was loaded with 5 μg of poly(A)$^+$ RNA. The *dia* and *fod* transcript sizes, in kilobases, are indicated to the right of the figure. An RNA loading control (the same blot reprobed with an α1-tubulin cDNA) is shown in the bottom panel. Lane 1, wild-type males; lane 2, wild-type females; lanes 3-8, RNA from mixture of adult males and females. Lane 3, dia^1/dia^1; lane 4, dia^1 revertant.; lane 5, germlineless progeny of tud^1/tud^1 females; lane 6, $dia^3/+$; $P[w^+$, pDC4]; lane 7, dia^3/dia^3 ; $P[w^+$, pDC4]; lane 8, dia^5/dia^5 ; $P[w^+$, pDC4]. Flies homozygous for the lethal dia^3 and dia^5 alleles survive only if they have inherited $P[w^+$, pDC4], which supplies *dia* function.

expressed at detectable levels (Fig. 6B, lane 7, compare to lanes 6 and 8). pDC4 plasmid DNA was injected into embryos to generate transgenic flies. One transformant line was identified that carried a copy of pDC4 ($P[w^+$, pDC4]) on the third chromosome. As shown in Fig. 6, lanes 6-8, this construct provided wild-type levels of the 4.8 kb *dia* transcript and somewhat lower levels of the 4.4 kb RNA. This transgene fully rescued the mutant phenotypes of all the alleles described in this paper: dia^1 males and dia^1/dia^2 females are rescued to wild-type levels of fertility, and dia^2, dia^3 and dia^9 flies are rescued to wild-type levels of viability and fertility.

Cloning of the 5′ end of the *dia* protein-coding sequence

The longest *dia* cDNA that we have identified (from an embryonic library) is 3.6 kb in length, significantly shorter than the *dia* transcripts detected by northern analysis. DNA sequencing revealed a single long open reading frame (ORF) extending to the 5′ end of the cDNA. Therefore, this cDNA appeared to be missing some protein-coding sequence at its 5′ end.

An RNA amplification strategy (RT-PCR) was devised to clone the missing 5′ cDNA sequence (see Experimental Procedures). RT-PCR reactions using the primers shown in Fig. 7 consistently resulted in a single product of 0.9 kb (data not shown). Two PCR products from each of two separate reactions were cloned and sequenced to resolve any Taq polymerase errors. Four sequence polymorphisms were identified; for each the correct base could be unambiguously determined, since the other sequences matched at these sites.

Analysis of dia protein sequence

The consensus PCR sequence contains a long ORF contiguous with the ORF from the 3.6 kb cDNA. The PCR and cDNA sequences were linked and the complete predicted amino acid sequence of the diaphanous protein was generated (Fig. 7). The nucleotides flanking the first start codon are a close match to the *Drosophila* translational start consensus sequence (Cavener, 1987). The next methionine in the ORF occurs 77 amino acids into the ORF.

No N-terminal signal peptide or obvious transmembrane domain is present in the 123×10^3 M_r protein. The protein does contain an unusual proline-rich domain near its middle. This proline-rich domain contains six repeats of 5 to 8 consecutive prolines separated by short stretches rich in the amino acids glycine, methionine, alanine or arginine, with a few interspersed prolines (Fig. 8C).

The diaphanous protein sequence was compared to protein sequences in available databases with the BLASTP program (Altschul et al., 1990). When the 'filter' directive was used to mask off segments of low compositional complexity, such as the proline-rich region, a region of about 130 amino acids was identified that was conserved in the yeast protein Bni1p, and in both mouse and chicken formins (Fig. 8B). Mutations in *BNI1* (bud neck involved) affect bud site selection and cytokinesis in diploid strains, and have a synthetic lethal interaction with *CDC12*, a gene required for cytokinesis in *S. cerevisiae* (Hartwell, 1971; H. Fares and J. Pringle, personal communication). In mice, the formin proteins are produced by the *limb deformity* gene; mutations in this gene result in limb malformations (Kleinebrecht et al., 1982; Woychick et al., 1985, 1990; Jackson-Grusby et al., 1992). In searching a six-frame translation of the DNA databases we found a fourth example of this conserved domain: a peptide encoded by a partial cDNA sequence from a rice expressed sequence tag (EST) (Fig. 8B).

The formins and Bni1p, like diaphanous, contain a large proline-rich domain located near the middle of the proteins (Fig. 8A,C). Although a number of proteins contain proline-rich domains (e.g. Mermod et al., 1989), those of the formins and Bni1p are the most similar to diaphanous in that they contain multiple stretches of consecutive prolines (Fig. 8C); proline-rich regions in other proteins usually consist of interspersed prolines or long stretches of polyproline. We have given the name FH1 (formin homology domain 1) to the proline-rich domains found in diaphanous, formin and Bni1p, and the name FH2 to the conserved 130 amino acid region in these proteins (Fig. 8). It is striking that not only the sequence composition of the FH1 and FH2 domains, but also the spacing between these domains, is conserved among diaphanous, Bni1p, and the formins (Fig. 8A).

Secondary structure analysis, using the algorithm devised by Stock and colleagues, revealed two regions of the diaphanous protein that are very likely to form coiled-coil domains (defined as a mean score >1.0; see Materials and Methods and Lupas et al., 1991). One of these, spanning amino acids 441 to 500 (mean score 2.1), ends just at the beginning of the FH1 domain. The second, spanning amino acids 863 to 1053 (mean score 1.3), begins 17 amino acids before the end of the FH2 domain, just after the region of highest sequence similarity. Thus it appears that coiled-coil domains flank the FH1 and FH2 domains of diaphanous. Moreover, sequence analysis also predicts the existence of coiled-coil domains at equivalent positions in mouse formin IV and Bni1p (H. Fares, personal communication).

DISCUSSION

Interpretation of mutant phenotypes

In this paper, we present the genetic and molecular characterization of the *diaphanous* locus. Mutations in *dia* affect a range

```
   1 CAAGAAAAAGTGTTCGGAGGGAGCGGTGTGCACAGCTTGTCAGCTTGTTTTTTGTACAGAGCGTCGCGCAATAAAACAAACAAAAAGACATAGTCCAAAAACAACAACCACAACAAAAGC  120
 121 CGTTTGATCGCGAGTTTGTGTGCGTGTGCGAGTGTGGCTCTATTTGTGTGAGAGTGGCAAGCCATTATTTTTCTGGTCTTATGACGTAAACCACCGAAAAAAATCAAAAGTGAGCAGGCG  240
 241 ACACAGATATCAAAACGAAGAATGTCTCGTCACGAGAAAACGAAATCCACGGGCGGCGGGCTCCTGGACAGTCTGTTCGGAAGACCCTCGAAGTCCAAGGGAGGAACCATCAGCAGTGGC  360
   1                     M  S  R  H  E  K  T  K  S  T  G  G  G  L  L  D  S  L  F  G  R  P  S  K  S  K  G  G  T  I  S  S  G   33

 361 ACCCTGGCCCATGGCGGACGACCCGTGTCCGCGGACAACTATGTGGTGCCGGGCGTGGAGGACTTTGAGCAGTACATCCAGCAGCTAAGCGTTGCGGAGCTGGATGCGAAGTTTCTGGAG  480
  34  T  L  A  H  G  G  R  P  V  S  A  D  N  Y  V  V  P  G  V  E  D  F  E  Q  Y  I  Q  Q  L  S  V  A  E  L  D  A  K  F  L  E   73

 481 ATCATCGAGGACATGAACATTCCGAAGGACAAGAGGGAGCCCCTGTTGGCCAAATCGAAGGAGGAGCGACAGAAGATGATTATGTGGCACTTGAAAGGTAAAAACTCACTGGAGCGTAGC  600
  74  I  I  E  D  M  N  I  P  K  D  K  R  E  P  L  L  A  K  S  K  E  E  R  Q  K  M  I  M  W  H  L  K  G  K  N  S  L  E  R  S  113

 601 GCCAACTCCCGCTTCGAGAAGCCCATAGACTATGTGGAATACCTGCAGAATGGGGAGCACAGCACGCACAAGGTGTACCAATGTGTGGAATCTCTGCGCGTGGCGCTCACCAGCAATCCG  720
 114  A  N  S  R  F  E  K  P  I  D  Y  V  E  Y  L  Q  N  G  E  H  S  T  H  K  V  Y  Q  C  V  E  S  L  R  V  A  L  T  S  N  P   153

 721 ATCTCGTGGATCAAGGAGTTTGGAGTGGCGGGCATCGGGACGATTGAGAAGCTGCTGGCCCGGTCAAAGAATAATGCCAGCTACGAGAAGATCGAGTTCGAGGCGATTCGGTGCCTGAAG  840
 154  I  S  W  I  K  E  F  G  V  A  G  I  G  T  I  E  K  L  L  A  R  S  K  N  N  A  S  Y  E  K  I  E  F  E  A  I  R  C  L  K   193

 841 GCGATCATGAACAACACATGGGGTCTGAACGTGGTGCTCAATCCGGATCAGCATAGTGTGGTGCTGCTGCTGGCGCAATCCCTGGATCCCCGGAAGCCGCAGACGATGTGTGAAGCCCTC  960
 194  A  I  M  N  N  T  W  G  L  N  V  V  L  N  P  D  Q  H  S  V  V  L  L  L  A  Q  S  L  D  P  R  K  P  Q  T  M  C  E  A  L   233

 961 AAGCTGCTGGCCTCGTTCTGCATTGTCTATGAGCGGAATGGCTACGAGAAGGTTCTCCGAGCCATAACCACTATTGCAGCCACATCCTTTAAAGCGAGCGAGCGCTTTCGACCCATAGTG 1080
 234  K  L  L  A  S  F  C  I  V  Y  E  R  N  G  Y  E  K  V  L  R  A  I  T  T  I  A  A  T  S  F  K  A  S  E  R  F  R  P  I  V   273

1081 GATGCCTTGTTTGCATCGGATCAGCAGGATCCCAAACGGGACTTGGCATGCCACAGCCTGATTTTTATTAATACTCTCACCAATACCCCCACGGATTTGAACTTCCGCCTGCACTTGCGA 1200
 274  D  A  L  F  A  S  D  Q  Q  D  P  K  R  D  L  A  C  H  S  L  I  F  I  N  T  L  T  N  T  P  T  D  L  N  F  R  L  H  L  R   313

1201 TGTGAGATTATGCGCATGGGTCTATACGATCGCCTAGATGAGTTCACCAAAATCGTGGAGGCCAGCAATAATGAGAACCTGCAGCAGCACTTCAAGATCTTCAACGAGATCCGCGAGGAT 1320
 314  C  E  I  M  R  M  G  L  Y  D  R  L  D  E  F  T  K  I  V  E  A  S  N  N  E  N  L  Q  Q  H  F  K  I  F  N  E  I  R  E  D   353

1321 GACTTCGAGGAGTTTGTGCAGCGCTTCGATAATGTCACCTTCAACATGGACGACGCCACCGATTGCTTCGATGTGCTGAAGAACCTGGTGACTGACACCACTTCCGAGCCCTACTTCCTG 1440
 354  D  F  E  E  F  V  Q  R  F  D  N  V  T  F  N  M  D  D  A  T  D  C  F  D  V  L  K  N  L  V  T  D  T  T  S  E  P  Y  F  L   393

1441 TCCATCCTGCAGCATTTGCTGTACATCAGGGATGACTTCTACTTCCGACCTGCCTATTATCAGCTGATTGAGGAGTGCATCTCACAAATCGTCTTCCACAAGGGTTACTGTGATCCGAAT 1560
 394  S  I  L  Q  H  L  L  Y  I  R  D  D  F  Y  F  R  P  A  Y  Y  Q  L  I  E  E  C  I  S  Q  I  V  F  H  K  G  Y  C  D  P  N   433

1561 TTCGAGAACCGGAACTTTAATATAGACACCTCGCTACTGCTGGACGACATTGTGGAGAAGGCCAAGGCCAAGGAGTCGAAGCGGTCGGAGGAGTACGAGAAGAAGATCGAGCAGCTGGAG 1680
 434  F  E  N  R  N  F  N  I  D  T  S  L  L  L  D  D  I  V  E  K  A  K  A  K  E  S  K  R  S  E  E  Y  E  K  K  I  E  Q  L  E   473

1681 AGTGCCAAGCAAGAGGCGGAAGCGAAGGCGGCTCATCTGGAGGAGAAAGTCAAGTTGATGGAGGCTAATGGTGTGGCGGCTCCGTCGCCCAATAAGCTACCCAAGGTAAACATACCCATG 1800
 474  S  A  K  Q  E  A  E  A  K  A  A  H  L  E  E  K  V  K  L  M  E  A  N  G  V  A  A  P  S  P  N  K  L  P  K  V  N  I  P  M   513

1801 CCACCGCCACCACCAGGAGGAGGAGGAGCACCTCCCCCACCGCCGCCACCTATGCCGGGACGAGCAGGTGGTGGACCTCCGCCTCCACCACCACCTCCCATGCCGGGAAGGGCAGGTGGA 1920
 514  P  P  P  P  P  G  G  G  G  A  P  P  P  P  P  P  P  P  M  P  G  R  A  G  G  G  P  P  P  P  P  P  P  P  M  P  G  R  A  G  G   553

1921 CCGCCACCACCTCCTCCACCGCCCGGAATGGGAGGCCCACCGCCTCCACCCATGCCTGGCATGATGCGCCCAGGAGGGGGACCTCCACCACCGCCCATGATGATGGGGCCCATGGTTCCC 2040
 554  P  P  P  P  P  P  P  P  G  M  G  G  P  P  P  P  P  P  M  P  G  M  M  R  P  G  G  G  P  P  P  P  P  P  M  M  M  G  P  M  V  P   593

2041 GTTCTGCCTCATGGCTTAAAACCGAAGAAGAAGTGGGACGTCAAGAATCCCATGAAGCGGGCCAACTGGAAAGCCATTGTCCCGGCCAAAATGTCCGACAAGGCATTCTGGGTCAAGTGC 2160
 594  V  L  P  H  G  L  K  P  K  K  K  W  D  V  K  N  P  M  K  R  A  N  W  K  A  I  V  P  A  K  M  S  D  K  A  F  W  V  K  C   633

2161 CAGGAGGATAAGCTGGCCCAGGATGACTTCCTCGCAGAACTGGCCGTGAAATTCTCTTCTAAACCTGTAAAGAAAGAACAGAAGGATGCGGTGGACAAGCCAACGACGCTGACAAAGAAA 2280
 634  Q  E  D  K  L  A  Q  D  D  F  L  A  E  L  A  V  K  F  S  S  K  P  V  K  K  E  Q  K  D  A  V  D  K  P  T  T  L  T  K  K   673

2281 AATGTCGATCTTCGTGTGCTCGACTCGAAGACTGCTCAGAACTTGGCTATTATGTTGGGAGGATCGCTGAAGCATCTGTCCTACGAACAGATCAAGATCTGCTTGCTGCGCTGCGACACC 2400
 674  N  V  D  L  R  V  L  D  S  K  T  A  Q  N  L  A  I  M  L  G  G  S  L  K  H  L  S  Y  E  Q  I  K  I  C  L  L  R  C  D  T   713

2401 GACATCCTGTCCTCCAATATCCTGCAGCAACTTATCCAGTACCTTCCGCCGCCAGAGCAACTCAAACGTTTGCAGGAGATCAAGGCAAAGGGCGAACCGCTACCGCCGATTGAACAGTTT 2520
 714  D  I  L  S  S  N  I  L  Q  Q  L  I  Q  Y  L  P  P  P  E  Q  L  K  R  L  Q  E  I  K  A  K  G  E  P  L  P  P  I  E  Q  F   753

2521 GCAGCCACAATAGGGGAAATTAAACGCCTTTCGCCGCGACTTCACAACCTGAACTTCAAGCTGACCTATGCGGACATGGTGCAGGATATCAAACCCGACATTGTGGCAGGAACGGCAGCA 2640
 754  A  A  T  I  G  E  I  K  R  L  S  P  R  L  H  N  L  N  F  K  L  T  Y  A  D  M  V  Q  D  I  K  P  D  I  V  A  G  T  A  A   793

2641 TGCGAAGAGATCCGGAATAGCAAAAAGTTCTCAAAGATCTTGGAGCTGATTCTGCTGCTTGGAAATTACATGAACTCGGGCTCCAAAAACGAGGCCGCCTTTGGCTTTGAGATCAGTTAT 2760
 794  C  E  E  I  R  N  S  K  K  F  S  K  I  L  E  L  I  L  L  L  G  N  Y  M  N  S  G  S  K  N  E  A  A  F  G  F  E  I  S  Y   833

2761 TTAACCAAACTGTCCAATACGAAGGATGCGGATAATAAGCAGACATTGCTGCACTACCTGGCTGACCTGGTGGAAAAGAAATTCCCAGATGCACTAAACTTCTACGACGATCTGTCGCAT 2880
 834  L  T  K  L  S  N  T  K  D  A  D  N  K  Q  T  L  L  H  Y  L  A  D  L  V  E  K  K  F  P  D  A  L  N  F  Y  D  D  L  S  H   873

2881 GTTAATAAAGCGTCGCGGGTCAACATGGATGCCATCCAAAAGGCCATGCGGCAAATGAATTCGGCGGTTAAGAACCTGGAAACTGATCTCCAGAACAACAAGGTGCCGCAGTGTGATGAT 3000
 874  V  N  K  A  S  R  V  N  M  D  A  I  Q  K  A  M  R  Q  M  N  S  A  V  K  N  L  E  T  D  L  Q  N  N  K  V  P  Q  C  D  D   913

3001 GACAAGTTTAGCGAGGTGATGGGCAAGTTTGCCGAGGAGTGCAGACAACAAGTGGACGTGCTGGGCAAAATGCAGCTGCAAATGGAGAAGCTGTACAAGGACCTCAGCGAGTACTATGCC 3120
 914  D  K  F  S  E  V  M  G  K  F  A  E  E  C  R  Q  Q  V  D  V  L  G  K  M  Q  L  Q  M  E  K  L  Y  K  D  L  S  E  Y  Y  A   953

3121 TTCGATCCCAGCAAATACACAATGGAGGAGTTCTTTGCGGACATCAAGACTTTCAAGGATGCCTTCCAAGCGGCCCACAACGACAATGTCCGGGTACGCGAGGAGCTGGAGAAGAAGCGT 3240
 954  F  D  P  S  K  Y  T  M  E  E  F  F  A  D  I  K  T  F  K  D  A  F  Q  A  A  H  N  D  N  V  R  V  R  E  E  L  E  K  K  R   993

3241 CGTCTGCAGGAGGCCCGAGAGCAGTCTGCTCGAGAGCAACAGGAGCGCCAACAGCGTAAGAAGGCAGTGGTTGACATGGATGCCCCGCAGACGCAGGAAGGCGTGATGGACAGTCTGTTG 3360
 994  R  L  Q  E  A  R  E  Q  S  A  R  E  Q  Q  E  R  Q  Q  R  K  K  A  V  V  D  M  D  A  P  Q  T  Q  E  G  V  M  D  S  L  L  1033

3361 GAGGCGCTGCAAACGGGCTCAGCCTTTGGCCAACGAAATCGACAGGCCCGGCGACAACGACCGGCGGGAGCGGAGCGGAGGGCACAGCTCAGCCGGAGTCGATCGCGGACGCGTGTCACC 3480
1034  E  A  L  Q  T  G  S  A  F  G  Q  R  N  R  Q  A  R  R  Q  R  P  A  G  A  E  R  R  A  Q  L  S  R  S  R  S  R  T  R  V  T  1073

3481 AACGGACAACTAATGACCCGCGAAATGATCCTCAACGAGGTTCTAGGCTCCGCGTAGAAGTAATGGAACGAGCAATTCCCCTCTGTACATATATAAATAGATATCTGTAGCGATATAAGG 3600
1074  N  G  Q  L  M  T  R  E  M  I  L  N  E  V  L  G  S  A  *                                                              1091

3601 GCCTGGGCTCATTAGCCGTGAGTCCGTGTGCGTCGCCTTGCTTGCATATTAACTCAATTGGCGTGTCCAAGAGATTTGACGAATTGGATGCCCAGCAGGAAGGACTGTATCAAGAGCCGC 3720
3721 AAAGTACTAATTAACTGGAACTAGAGTCACAGGTGGATTTTCCGAAATTATTAATTGCTGTAACGGAGTCCTTGGTTTATAGGATTGAGTTTTCCGTTTGAATTGATTGGAGTATTTATT 3840
3841 TAATTGATTGTTTAATAAATAACATTTTGTTTAGAATTTATAATATGCAAATAACACATACAAGATCACTTTTATTTAATTAAAATGATTGGCAGGCCTTTTAATTTCAGCTGTACATTT 3960
3961 TATTTTACTGCTATGACTGTCCCAATATTAAAGTACCAATTTTACAAGTCGAATATACTATAAAAAAACTTCTAGTTGAAACTAAATTCATGTCAAGAAATTAGCAATGCATATCGATGA 4080
4081 AAACTCCAACTTCGCTACCTATACGAAAAAACTAAATATATAGATATTTTTGATACGCTTTTACTTCTGAAATTATGTCGTTGTAAGTCTAAATACTTAAAAATGAAACTGAAAGCAACT 4200
4201 TTAAGATTTATCTGTATCTTGTGGTCTTTACATACTGTATTCCTATTTTACTGTGGCATGTTCCAATCATTTTGCTGTGCACTTCGTTTTCGAGGCACTCAGTAGAGCAGAGGCGAGATT 4320
4321 TATATCAAAGCAAAATTCCTTTTGTATTTAATTTGTTTTTAAGCGAAACTTTTTTGAACTTGTTTACCTTAAAATCCTATTTTAATCCGCGCTATACATAACATTGTTAATAAATGTCTA 4440
4441 TTTTTCGACAATGACGGTATATTGAAAATAAAAAAAAAAAA  4481
```

Fig. 7. Composite *dia* cDNA nucleotide sequence with predicted protein sequence (Genbank U11288). The synthetic oligonucleotides used for RT-PCR correspond to the underlined bases at 1-21 and 922-953. The 5′ end of the 3.6 kb cDNA corresponds to the ‘A’ at position 862.

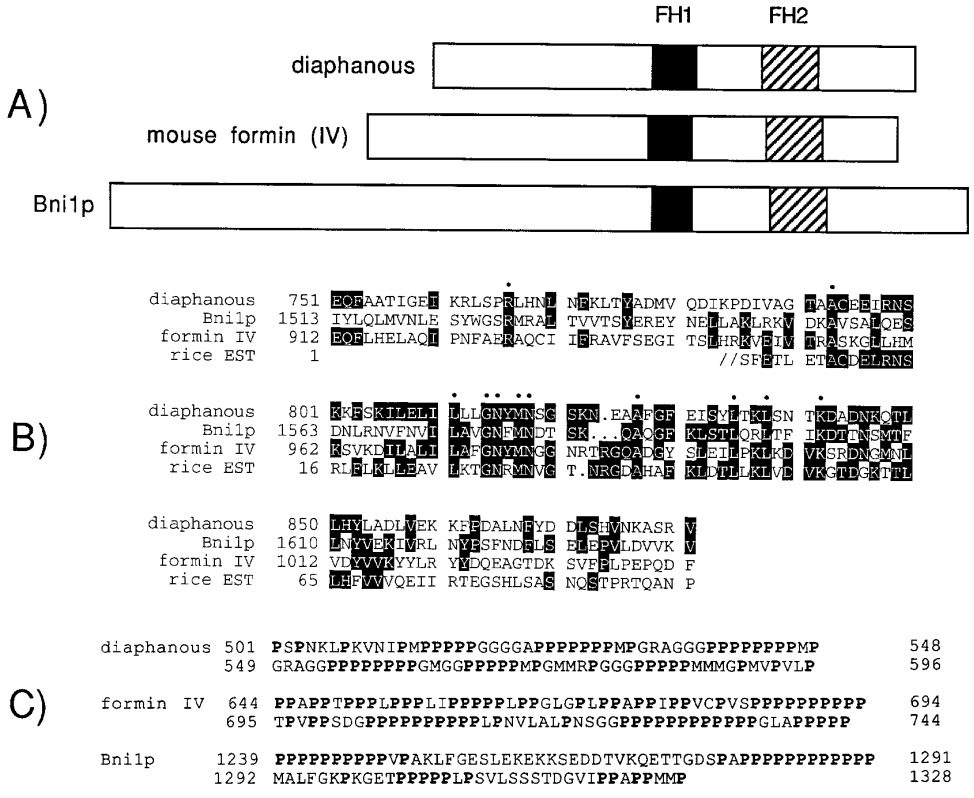

Fig. 8. FH1 and FH2 domains of diaphanous, mouse formin IV, and Bni1p. (A) Schematic representation of diaphanous (1091 amino acids), mouse formin IV (1206 amino acids) and yeast Bni1p (1953 amino acids) (Woychick et al., 1990; Genbank L31766). All three proteins contain a proline-rich domain, labelled FH1, and a second region of sequence similarity, labelled FH2. The FH2 domains are approximately 160 residues from the end of the FH1 domains. Three of the four mouse isoforms, as well as chicken formin, contain both of the regions of similarity (Trumpp et al., 1992). (B) Alignment of FH2 domains from diaphanous, mouse formin IV and Bni1p, and the translation product of a rice EST (Genbank D24760). Within a stretch of 68 amino acids beginning at residue 790, diaphanous shares 35% and 34% sequence identity with mouse formin IV and Bni1p, respectively. Dots indicate invariant residues; we note the highly conserved octapeptide motif: LxxGNxMN. Slash marks indicate end of EST sequence. (C) Sequence of the FH1 domains. The proline-rich FH1 domains of diaphanous, mouse formin and Bni1p each contain multiple stretches of consecutive prolines.

of cell types and result in diverse phenotypes, including male sterility, female sterility and lethality. Examination of mutant tissues suggests that a single underlying cellular defect, a failure of cytokinesis, is responsible for all of these phenotypes. Furthermore, analysis of null mutations indicates that *dia* is likely to be required for cytokinesis in all cells.

Sterile phenotypes

Failure of cytokinesis during meiosis is readily apparent in dia^1 spermatids. The presence of rare spermatids containing 8 nuclei indicates that cytokinesis can also fail in the mitotic divisions preceding meiosis. Finally, the fact that testes become devoid of germinal contents suggests that cytokinesis can also fail during stem cell divisions. Because the 5-9 germline stem cells present in each testis continually divide (Hardy et al., 1979), failure of cytokinesis in stem cells should block the initiation of spermatogenesis.

dia^1/dia^2 females lay short eggs with defective dorsal appendages. These defects are presumably secondary to the failure of cytokinesis observed in the follicle cell layer. Eggshell synthesis is a complex process requiring the precise migration and positioning of specialized follicle cells to lay down the various layers and structures of the eggshell (Spradling, 1993). The presence of abnormally large, binucleate follicle cells would likely interfere with follicle cell migration and with the construction of eggshell structures such as the dorsal appendages.

Both the male- and female-sterile phenotypes appear to be due to a failure in cytokinesis. However, the cells affected in spermatogenesis are of germinal origin whereas the cell type affected in oogenesis is of somatic origin. This specificity may be due to the fact that in females the somatic follicle cells undergo a greater number of mitotic divisions than germline cells. By the time follicle cell mitoses cease at stage 6 of oogenesis, an egg chamber consists of approximately 1000 follicle cells but only 16 germline cells. A clone of 16 germline cells arises through 5 rounds of mitosis, starting with a stem cell division. It is believed that a group of approximately 16 follicle cells originally envelop a single egg chamber (Spradling, 1993). Even if there were a direct progenitor for each of these 16 follicle cells, follicle cells must undergo at least two more rounds of mitotic divisions than the germline cells. We suggest that the requirement for these extra mitotic divisions, and the accompanying additional opportunities for the lack of *dia* function to have phenotypic consequences, is the cause of the apparent selectivity of the cytokinesis defect for follicle cells.

The failure of cytokinesis that we have observed in follicle cells has also been observed in *Drosophila* females lacking *peanut* function (Neufeld and Rubin, 1994). In addition, other female-sterile mutants that produce short eggs and short dorsal appendages have been described (Spradling, 1993). Since many of these lines are relatively uncharacterized, it is conceivable that a similar underlying defect is responsible for the phenotypes of some of these mutants.

Lethal phenotype

The lethal *dia* phenotypes are likely the result of a gradual dilution of maternal diaphanous gene product over many cell cycles. Presumably, when the amount of protein in a cell falls

below a certain critical point, cytokinesis fails in subsequent cell cycles. In flies homozygous for the null allele *dia*^{2}, the defect is manifested early, perhaps in the first few larval cell divisions, such that by the end of larval development, very few neuroblasts are present and all are highly hyperploid. In contrast, in partial-loss-of-function *dia*^{3} mutants, the cytokinesis defect apparently is manifested after a variable number of normal cell cycles. The number of cells and the degree of polyploidy is therefore intermediate between *dia*^{2} and wild type.

The survival of embryos homozygous for null *dia* mutations is likely due to a maternal contribution of *dia* mRNA or protein. Consistent with this hypothesis, *dia* transcripts are detectable in 0-3 hour embryos. Maternal *dia* gene product may be required for cytokinesis in postblastoderm mitoses and perhaps for cellularization of the syncytial blastoderm, a process similar in some respects to cytokinesis.

Role of *diaphanous*

We believe that *dia* is specifically required for cytokinesis; as such, the protein might be a component of the contractile ring or regulate its function. Other aspects of the cell cycle, including DNA replication, spindle assembly and function, and nuclear reformation appear to proceed normally in *dia* mutants. The chromosomes of salivary glands from larvae homozygous for a null allele reach wild-type levels of polytenization (unpublished data), demonstrating that *dia* is not required for DNA replication. Chromosome segregation occurs in mutant spermatids, follicle cells and neuroblasts.

In *dia* mutant cells that have failed to undergo cytokinesis, nuclear reformation appears to occur normally, although fusion of nuclei occurs in neuroblasts and possibly in follicle cells. In *Drosophila* and other species, fusion of nuclei is common in cells that become multinucleate due to a failure of cytokinesis. This phenomenon has been observed in the *Drosophila* cytokinesis mutant *spaghetti-squash*^{1} (*sqh*^{1}) (Karess et al., 1991; see also Gatti and Baker, 1989) and in a hamster cell line mutant defective for cytokinesis (Hatzfeld and Buttin, 1975). In *dia*^{1} spermatids, fusion of nuclei does not take place. This may be due to the fact that the spermatogenesis differentiation program, which includes the condensation of nuclei for packaging into spermheads, proceeds to some extent in these spermatids.

The partial loss-of-function *sqh*^{1} allele results in a mutant phenotype similar to *dia*^{3}: homozygotes die as early pupae, have small imaginal discs and exhibit a high frequency of polyploid neuroblasts in third instar larvae. In addition, cleavage furrows are missing in anaphase figures of *sqh*^{1} neuroblasts. Viable, sterile alleles of *sqh* have not been described. The *sqh* gene encodes the regulatory light chain of nonmuscle myosin, consistent with a requirement for myosin function during cytokinesis in *Drosophila* (Karess et al., 1991).

Significance of formin homology domains

Diaphanous, the formins and Bni1p contain two regions of sequence similarity, FH1 and FH2. Although no primary sequence similarity was detected in the N-terminal domains of these proteins, this portion of the mouse and chicken formins is much less conserved than the C-terminal domain (Trumpp et al., 1992; Jackson-Grusby et al., 1992). The fact that each of the proteins containing an FH2 domain also contains an FH1 domain, that the spacing between the two domains is also conserved and that the two domains in each protein are flanked by regions likely to form coiled-coil domains strongly suggests that the FH1 and FH2 domains function together.

Mutations in the mouse *limb deformity* gene, which encodes multiple formin isoforms, cause moderate to severe aplasia of both limbs and kidneys (Kleinebrecht et al., 1982; Woychick et al., 1985; 1990; Zeller et al., 1989). Within the limb bud, defects are observed in the apical epidermal ridge. However, complete loss-of-function mutations have not been identified, and little is known about the role of the formins at the cellular level. It remains possible that a defect in cytokinesis is responsible for the observed deformities.

A peptide representing a portion of the FH1 domains of the formins has been shown to bind in vitro to the SH3 domain of the Abl protein (Ren et al., 1993) and this sequence is a reasonable match with a recently defined consensus binding site for SH3 domains (Yu et al., 1994). However, the FH1 domains of diaphanous and Bni1p do not fit the defined consensus. Another possibility is that the FH1 domain serves as a hinge between two functionally separate domains. Hinge regions of proteins are also often rich in proline (e.g. Koenig and Kunkel, 1990), but tend to be smaller than the FH1 domain and have more evenly dispersed proline residues.

One intriguing possibility is that the FH1 domain binds to profilin, which is involved in the regulation of actin polymerization (Carlsson et al., 1977). Profilin binds with extremely high affinity to poly(L-proline), a feature that has been exploited to purify profilin from crude cell extracts in a single-step (Tanaka and Shibata, 1985). Profilin is therefore expected to interact in vivo with at least one as yet unidentified protein containing a proline-rich region. *Dictyostelium discoideum* cells mutated for both profilin genes form multinucleate cells due to a defect in cytokinesis (M. Schleicher, personal communication). Furthermore, male-sterile alleles of *chickadee*, which encodes the *Drosophila* homolog of profilin (Cooley et al., 1992) affect the proliferation of the germline (Castrillon et al., 1993). In this regard, it is notable that *dia* and the *Drosophila* profilin gene can both mutate to germlineless phenotypes.

The diaphanous protein might also interact directly with the product of the *peanut* gene (Neufeld and Rubin, 1994). Like *diaphanous*, *peanut* is required for cytokinesis in *Drosophila*. The peanut protein is similar throughout its length to the *S. cerevisiae* proteins Cdc3, Cdc10, Cdc11, and Cdc12, all of which are required for cytokinesis (Hartwell, 1971; Kim et al., 1991). Since *CDC12* has a synthetic lethal interaction with *BNI1* (H. Fares and J. Pringle, personal communication), there may be a direct interaction between the products of these two genes and, by analogy, between peanut and diaphanous. Whether or not this is the case, the observation that there are regions of similarity common to diaphanous, the formins and Bni1p indicates that these proteins are likely to share a common biochemical function.

We would like to give special thanks to Johnny Fares and John Pringle for pointing out the potential coiled-coil domains in diaphanous and for communication of unpublished results. We also thank Trudi Schüpbach, Kathy Matthews and the Bloomington *Drosophila* Stock Center for fly stocks; Tulle Hazelrigg, Nick Brown and John Tamkun for *Drosophila* libraries; Pierre Gönczy and Steve DiNardo for critical reading of the manuscript; and Teresa Gallardo for technical assistance. We gratefully acknowledge predoctoral

support for D. H. C. through Medical Scientist Training Grant GM-08014 and a predoctoral training grant from the National Institutes of Health. This work was supported by a grant to S. A. W. from the Texas Higher Education Coordinating Board.

REFERENCES

Altschul, S. F, Gish, W., Miller, W., Myers, E. W. and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* **215**, 403-410.

Ashburner, M. (1989). *Drosophila: a Laboratory Manual.* Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.

Brown, N. H. and Kafatos, F. C. (1988). Functional cDNA libraries from *Drosophila* embryos. *J. Mol. Biol.* **203**, 425-437.

Carlsson, L., Nyström, L.-E., Sundkvist, F., Markey, F. and Lindberg, U. (1977). Actin polymerization is influenced by profilin, a low molecular weight protein in non-muscle cells. *J. Mol. Biol.* **115**, 465-483.

Castrillon, D. H., Gönczy, P., Alexander, S., Rawson, R., Eberhart, C. G., Viswanathan, S., DiNardo, S. and Wasserman, S. A. (1993). Toward a molecular genetic analysis of spermatogenesis in *Drosophila melanogaster*: characterization of male-sterile mutants generated by single *P* element mutagenesis. *Genetics* **135**, 489-505.

Cavener, D. (1987). Comparison of the consensus sequences flanking translation start sites in *Drosophila* and vertebrates. *Nucl. Acids. Res.* **15**, 1353-1361.

Cooley, L., Verheyen, E. and Ayers, K. (1992). *chickadee* encodes a profilin required for intercellular cytoplasm transport during *Drosophila* oogenesis. *Cell* **69**, 173-184.

DeLozanne, A. and Spudich, J. A. (1987). Disruption of the Dyctiostelium myosin heavy-chain by homologous recombination. *Science* **236**: 1086-1091.

Fuller, M. (1993). Spermatogenesis. In *Development of Drosophila* (eds. A. Martinez-Arias and M. Bate), pp. 61-147. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.

Gatti, M., Tanzarella, C. and Olivieri, G. (1974). Analysis of the chromosome aberrations induced by X-rays in somatic cells of *Drosophila melanogaster*. *Genetics* **77**, 701-719.

Gatti, M. and Baker, B. S. (1989). Genes controlling essential cell-cycle functions in *Drosophila melanogaster*. *Genes Dev.* **3**, 438-453.

Gonzalez, C., Casal, J., and Ripoll, P. (1989). Relationship between chromosome content and nuclear diameter in early spermatids of *Drosophila melanogaster*. *Genet. Res.* **54**, 205-221.

Hardy, R. W., Tokuyasu, K. T., Lindsley, D. L. and Garavito, M. (1979). The germinal proliferation center of *Drosophila melanogaster*. *J. Ultrastruct. Res.* **69**, 180-190.

Hartwell, L. H. (1971). Genetic control of the cell division cycle in yeast. IV. Genes controlling bud emergence and cytokinesis. *Exp. Cell. Res.* **69**, 265-271.

Hatzfeld, J. and Buttin, G. (1975). Temperature-sensitive cell cycle mutants: a Chinese hamster cell line with a reversible block in cytokinesis. *Cell* **5**, 123-129.

Jackson-Grusby, L., Kuo, A. and Leder, P. (1992). A variant limb deformity transcript expressed in the embryonic mouse limb defines a novel formin. *Genes Dev.* **6**, 29-37.

Karess, R. E. and Glover, D. M. (1989). *rough deal*: A gene required for proper mitotic segregation in *Drosophila*. *J. Cell. Biol.* **109**, 2951-2961.

Karess, R. E., Chang, X., Edwards, K. A., Kulkarni, S., Aguilera, I. and Kiehart, D. P. (1991). The regulatory light chain of nonmuscle myosin is encoded by *spaghetti-squash*, a gene required for cytokinesis in *Drosophila*. *Cell* **65**, 1177-1189.

Kawasaki, E. S. (1990). Amplification of RNA. In *PCR protocols: A Guide to Methods and Applications* (eds. M.A. Innis, D.H. Gelfans and J.J. Sninsky), pp. 21-27. San Diego, California: Academic Press.

Kim, H. B., Haarer, B. K. and Pringle, J. R. (1991). Cellular morphogenesis in the *Saccharomyces cerevisiae* cell cycle: localization of the *CDC3* gene product and the timing of events at the budding site. *J. Cell Biol.* **112**, 535-544.

Kleinebrecht, J., Selow, J. and Winkler, W. (1982). The mouse mutant *limb-deformity* (*ld*). *Anat. Anz.* **152**, 313-324.

Koenig, M. and Kunkel, L. M. (1990). Detailed analysis of the repeat domain of dystrophin reveals four potential hinge segments that may confer flexibility. *J. Biol. Chem.* **265**, 4560-4566.

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

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

Mermod, N., O'Neill, E.A., Kelly, T. J. and Tjian, R. (1989). The proline-rich transcriptional activator of CTF/NF-I is distinct from the replication and DNA binding domain. *Cell* **54**, 741-753.

Neufeld, T. P. and Rubin, G. M. (1994). The *Drosophila peanut* gene is required for cytokinesis and encodes a protein similiar to yeast putative bud neck filament proteins. *Cell* **77**, 371-379.

Pirrota, V. (1988). Vectors for P-element transformation in *Drosophila*. In *Vectors: A Survey of Molecular Cloning Vectors and Their Uses* (eds. R.L. Rodriguez and D.T. Denhardt), pp.437-456. Boston, Massachussetts: Butterworth.

Rappaport, R. (1986). Establishment of the mechanism of cytokinesis in animal cells. *Int. Rev. Cytol.* **105**, 245-281.

Ren, R., Mayer, B. J., Cicchetti, P. and Baltimore, D. (1993). Identification of a ten-amino acid proline-rich SH3 binding site. *Science* **259**, 1157-1161.

Robertson, H.M., Preston, C. R., Phillis, R. W., Johnson-Schlitz, D. M., Benz, W. K. and Engels, W. R. (1987). A stable genomic source of P element transposase in *Drosophila melanogaster. Genetics* **118**, 461-470.

Sambrook, J. Fritsch, E. F. and Maniatis, T. (1989). Molecular cloning: A Laboratory Manual. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.

Satterwhite, L. and Pollard, T. D. (1992). Cytokinesis. *Curr. Opin Cell Biol.* **4**, 43-52.

Spradling, A. (1986). P element-mediated transformation. In *Drosophila: A Practical Approach* (ed. D.B. Roberts), pp. 175-197. Oxford: IRL Press.

Spradling, A. (1993). Oogenesis. In *Development of Drosophila* (eds. A. Martinez-Arias and M. Bate), pp. 1-70. New York: Cold Spring Harbor Laboratory Press.

Tanaka, M. and Shibata, H. (1985). Poly(L-proline)-binding proteins from chick embryo are a profilin and a profilactin. *Eur. J. Biochem.* **151**, 291-297.

Trumpp, A., Blundell, P.A., de la Pompa, J. L. and Zeller, R. (1992). The chicken *limb deformity* gene encodes nuclear proteins expressed in specific cell types during morphogenesis. *Genes Dev.* **6**, 14-28.

Watts, F. Z., Shiels, G. and Orr, E. (1987). The yeast *MYO1* gene encoding a myosin-like protein required for cell division. *EMBO J.* **6**, 3499-3505.

Woychick, R. P., Stewart, T. A., Davis, L. G., D'Eustachio, P. and Leder, P. (1985). An inherited limb deformity created by insertional mutagenesis in a transgenic mouse. *Nature* **318**, 36-40.

Woychick, R. P., Maas, R. L., Zeller, R., Vogt, T. F. and Leder, P. (1990). 'Formins': proteins deduced from the alternative transcripts of the *limb deformity* gene. *Nature* **346**, 850-852.

Yu, H., Chen, J. K., Feng, S., Dalgarno, D. C., Brauer, A. W and S. L. Schreiber. (1994). Structural basis for the binding of proline-rich peptides to SH3 domains. *Cell* **76**, 933-945.

Zeller, R., Jackson-Grusby, L. and Leder, P. (1989). The limb deformity gene is required for apical ectodermal ridge differentiation and anteroposterior limb pattern formation. *Genes Dev.* **3**, 1481-1492.

(Accepted 1 September 1994)