Development 122, 1859-1871 (1996)
Printed in Great Britain © The Company of Biologists Limited 1996
DEV5061

The *Drosophila pipsqueak* gene encodes a nuclear BTB-domain-containing protein required early in oogenesis

Heidi Horowitz and Celeste A. Berg

Department of Genetics, P.O. Box 357360, University of Washington, Seattle, Washington 98195-7360, USA

Author for correspondence (e-mail: berg@genetics.washington.edu)

SUMMARY

Mutations at the *pipsqueak* locus affect early patterning in the *Drosophila* egg and embryo. We have cloned *pipsqueak* and found that it is a large and complex gene, encoding multiple transcripts and protein isoforms. One protein, PsqA, is absent in all of the mutants that we have examined. We show that PsqA is a nuclear protein present in the germ cells and somatically derived follicle cells throughout oogenesis and that it is required prior to stage one of oogenesis. PsqA contains a BTB (POZ) domain at its amino terminus; additionally, we have identified an evolutionarily conserved motif of unknown function present four times in tandem at the C terminus of the protein. *PZ pipsqueak* mutants produce a putative fusion protein containing the *pipsqueak* BTB domain fused to sequences resident on the *PZ* element (H. Horowitz and C. Berg, 1995 *Genetics* 139, 327-335). We demonstrate here that expression of this fusion protein in wild-type flies has a dominant effect, resulting in infertility and eggshell defects. These dominant phenotypes are discussed in light of current theories on the role of the BTB domain in protein-protein interactions.

Key words: *pipsqueak,* psq, BTB, dominant-negative, *Drosophila*, oogenesis

INTRODUCTION

Proper development of the *Drosophila* early egg chamber is required to generate polarity in the egg and embryo. During the middle stages of oogenesis, RNAs and proteins required for germ cell determination and for the generation of embryonic abdominal segments are localized to the posterior pole of the oocyte (reviewed by St. Johnston, 1993). The localization of the germ cell determinant(s) and the mRNA encoding Nanos, the posterior morphogen, is accomplished by a common mechanism, which requires the activity of at least 9 maternal-effect genes. Mothers homozygous for mutations in any of these genes produce embryos that lack abdominal segments and pole cells (the germline precursors).

The *pipsqueak* (*psq*) gene has been shown to be required maternally for correct abdominal segmentation and pole cell formation in the embryo (Siegel et al., 1993). It has been suggested that the *psq* embryonic defect in posterior patterning might be due to insufficient expression of the maternal posterior group gene *vasa* in *psq* mutants (Siegel et al., 1993). In addition, we have shown that certain *psq* alleles also exhibit dorsal/ventral defects in eggshell structure (Horowitz and Berg, 1995).

Establishment of the dorsal/ventral axis is set into motion with the localization of *gurken* (*grk*) mRNA in the developing oocyte. The Grk protein serves as a signal to determine follicle cell fate at two distinct times during oogenesis (reviewed by Anderson, 1995). During mid-oogenesis, Grk signals to the follicle cells at the posterior of the egg chamber, inducing them to adopt a posterior fate (Roth et al., 1995; González-Reyes et al., 1995). Later, in stages 8-10 of oogenesis, Grk is localized as a cap above the oocyte nucleus (Roth et al., 1995) where it signals to adjacent follicle cells to determine their fate as dorsal. This signaling cascade culminates in the proper establishment of dorsal/ventral polarity in the embryo, and in the formation of the two dorsolateral respiratory appendages of the eggshell (reviewed by Schüpbach and Roth, 1994; Chasan and Anderson, 1993).

We report here the cloning and molecular characterization of the *psq* gene. Our analyses indicate that *psq* is a complex and large gene encoding multiple transcripts and protein isoforms. The PsqA isoform contains a BTB domain; these domains have been shown to mediate protein-protein interactions (Bardwell and Treisman, 1994; Chen et al., 1995). *psq* mutants that produce a truncated protein harboring this domain exhibit defects in the dorsal/ventral patterning of the eggshell and embryo.

MATERIALS AND METHODS

Fly stocks

P[lacZ; ry⁺] (PZ) insertional mutations at the *psq* locus (alleles psq^{2403}, psq^{0115}, psq^{8109} and psq^{0482}) were generated in a screen described in Karpen and Spradling (1992); *psq* is referred to as *zeppelin* in Spradling (1993). Excision lines psq^{RF9} and psq^{RF13} were created from allele psq^{2403} as described previously (Horowitz and Berg, 1995). The deficiency, *Df(2R)psq-lolaΔ18* (*Δ18)* was generated

by excising the two *PZ* elements resident in allele psq^{0482}; such an excision is likely to delete all of the DNA between the two *P* elements (Cooley et al., 1990). The *Δ18* deficiency fails to complement both *psq* and a nearby lethal (*lola,* E. Giniger, personal communication); Southern blot analysis (data not shown) demonstrates that the *Δ18* deficiency has lost DNA beginning at the *PZ* element within *psq* and extending through the length of our cloned DNA 5′ to the *psq* gene (see Fig. 6).

The EMS allele psq^{HK38} (Schüpbach and Wieschaus, 1991) was originally recovered as a second mutation on a *cappucino* mutant chromosome. Our studies employed a recombinant chromosome lacking the *cappucino* mutation.

The *P[lacZ, w⁺]* (*PW*) line psq^{fs1} and the excision allele $psq^{1\text{-}30}$ are described by Siegel et al. (1993). These authors refer to psq^{fs1} as psq^{P1} and $psq^{1\text{-}30}$ as $psq^{X1\text{-}30}$.

Transformant lines were produced by germline transformation carried out according to Spradling (1986).

Embryo and ovary preparations

Cuticles were prepared according to Wieschaus and Nüsslein-Volhard (1986). Ovaries and individual egg chambers were prepared according to Verheyen and Cooley (1994). In situ hybridizations were carried out as described previously (Gillespie and Berg, 1995). Whole-mount analyses using antibodies to the Psq protein were carried out as follows: ovaries from 2- to 3-day-old females were dissected in modified EBR and fixed in devitellinizing buffer for 10 minutes (Verheyen and Cooley, 1994). All washes were carried out for 5 minutes unless otherwise specified. Ovaries were rinsed three times in 1× PBS and subsequently extracted in 1% saponin in 1× PBS for one hour at room temperature. The ovaries were rinsed twice in 0.1% saponin in PBS and once in PBTB (1× PBS, 0.1% Triton-X-100, 0.2% BSA), then incubated for 30 minutes in 5% normal goat serum (NGS) in PBTB. Rabbit serum was diluted (1:500) into PBTB with 5% NGS and overnight incubation was carried out at 4°C. The tissue was rinsed four times for 15 minutes each in PBTB and incubated with a 1:100 dilution of BODIPY-conjugated goat anti-rabbit secondary antibody (Molecular Probes) in PBTB overnight at 4°C. The samples were subsequently rinsed four times for 15 minutes each in PBTB and mounted as described in Theurkauf et al. (1992). Fluorescence was observed on a Bio-Rad MRC-600 confocal microscope, utilizing the Bio-Rad COMOS program. Images were processed using the NIH Image program.

Cloning and characterization of *psq* DNA

Preparation of DNA and RNA, and northern and Southern blot analyses were described previously (Horowitz and Berg, 1995).

Genomic DNA flanking the *PZ* psq^{2403} insertion and the *PW* psq^{fs1} insertion were kindly provided by Haifan Lin and Vivian Siegel, respectively. Genomic DNA spanning the exons downsteam of the *PZ* insertions was isolated by probing a cosmid library (gift of Marc Champe, Genentech, San Francisco) with the 1.0 kb partial *psq* cDNA, 13-13, described previously (Horowitz and Berg, 1995). The 5.1 kb *psq*-1 cDNA, pHPT7-9, was isolated from an ovarian cDNA library (Stroumbakis et al., 1994) using a 300 bp 5′ UTR probe (see Fig. 6). Alternative cDNAs pHH62 and pHH63 (*psq*-2) were isolated by rescreening the library with the complete pHPT7-9 cDNA.

DNA sequencing was performed using the dideoxynucleotide chain termination method with the Sequenase kit (US Biochemical). Bases 615-624 (TATACTGCAG, Fig. 4A) were not present in pHPT7-9 but were found in both cDNA 13-13 and genomic DNA (TATACTGGAG). The most likely cause of this deletion is an error by the reverse transcriptase during the cDNA library construction: this region contains direct repeats and significant dyad symmetry.

Determination of genomic structure was performed by a combination of restriction analysis of cloned DNA, Southern blot analysis of whole fly DNA and DNA sequencing. For all introns, sequence determination of both the 5′ and 3′ junctions in the genomic DNA permitted unambiguous identification of the position of each intron relative to the cDNA (see Fig. 4A). We identified introns between the following pairs of nucleotides in the *psq-1* cDNA (numbering as in Fig. 4A): (614, 615), (883, 884), (1048,1049), **(2014, 2015)**, **(2153, 2154)**, **(2693, 2694)**, (3332, 3333), **(3435, 3436)**, and (3619, 3620). Nucleotide pairs presented above in bold indicate positions at which the introns were sequenced in their entirety. The presence of the intron at (883, 884) was inferred based on Southern analysis of genomic DNA and homology to consensus sequences; the precise position of this intron was determined by Zollman et al. (1994). The genomic structure that we have determined differs somewhat from that presented by Weber et al. (1995).

Sequence was compiled using the IntelliGenetics program. Database searches were carried out using the BLAST program from the NIH (Altschul et al. 1990).

The sequences of the three classes of embryonic *psq* transcripts identified by Weber et al. (1995) are distinct from the ovarian transcripts that we present here. The primary differences are as follows: The 5′ ends of all the embryonic transcripts differ from the ovarian transcripts *psq-1* and *psq-2*. The first 210 nucleotides of the E1A exon of Weber et al. (1995) are unique, but are followed by sequence starting at position 402 of the *psq-1* transcript (numbering as in Fig. 4A). Our genomic sequencing indicates that nucleotides 1-614 of *psq-1* are present in a single exon, suggesting that the 5′ end of E1A is derived from sequence 5′ to the *psq-1* transcript. The sequences of exons E1B and E1C of Weber et al. (1995) are not present in any of the ovarian transcripts that we have identified. The 3′ ends of the embryonic transcripts are extended by approximately 1 to 1.2 kb compared to the ovarian transcripts *psq-2* and *psq-1*, respectively.

The predicted proteins from all embryonic transcripts differ from those of the ovarian transcripts: the predicted embryonic proteins have 24 unique amino acids C-terminal to amino acid 1062 of the protein encoded by *psq-1* (see Fig. 4A). Comparison with our sequence shows that the embryonic transcripts lack the sequence from 3920 to 4096 of *psq-1*, suggesting that an additional splicing event occurs in the embryonic tissues, permitting translation to proceed somewhat further than predicted for the ovarian proteins.

Plasmids

Drosophila transformation constructs

P[hs-*psq*]

An *Xba*I fragment containing sequences from position 1 to 4619 of cDNA pHPT7-9 was cloned into the unique *Xba*I site of the transformation vector pCaSpeRhs (Thummel and Pirrotta, 1991), which uses the w^+ gene as a marker.

P[hs-psq-l(3)S12]

A 570 bp *Eco*RI-*Sal*I fragment isolated from a RT-PCR product of the *psq-l(3)S12* truncation mRNA (see Fig. 4A, Horowitz and Berg, 1995) was cloned into pCaSpeRhs. This fragment contains the entire coding region of the predicted fusion protein, as well as 47 bp of *psq* 5′ UTR and 29 bp of sequence 3′ to the predicted *l(3)S12* stop codon (Dutton and Chovnick, 1991).

Protein expression constructs

pHH14

A *Bam*HI fragment of *psq* extending from nucleotide 864 to 5162 of the pHPT7-9 cDNA (amino acids 45-1065) was cloned into the *Bam*HI and *Bgl*II sites of expression vector pQE-13 (Qiagen).

pHH70

A fragment extending from the *Nco*I site (position 2038 in Fig. 4A) to the 3′ end (*Not*I site) of the *psq*-2 cDNA pHH63 (see Results) was cloned into the *Nco*I and *Not*I sites of the expression vector, pET-28C (Novagen). Expression of pHH70 produces a protein of 69×10^3 M_r, derived entirely of *psq* sequence (amino acids 436-1065 in Fig. 4).

pHH69

This is identical to pHH70, except that the fragment was derived from cDNA pHH62 (see Results), and thus harbors a deletion of amino acids 719-736 (Fig. 4), producing a protein of 67×10^3 M_r.

Polyclonal antibody production

A 6xHis-Psq fusion protein generated from pHH14 was isolated from *E. coli* inclusion bodies. Protein purified by preparative SDS-polyacrylamide gel electrophoresis was used to immunize rabbits (R&R Rabbitry, Stanwood, WA). Antisera from two different rabbits (AS1 and AS2) showed slightly different immunoreactivity on westerns (see Results).

Western blot analysis

Drosophila ovary lysates were prepared by homogenizing ovaries in buffer (50 mM Tris, pH 7.5, 3 mM EDTA, 1% NP-40, 0.1% SDS, 1 mM N-ethylmaleimide, 2 µg/ml leupeptin, 200 KIU/ml Traysol, 100 µg/ml PMSF), and protein concentration was determined using the Bio-Rad Protein Assay reagent. Approximately 5 µg of protein were loaded per lane, subjected to SDS-polyacrylamide gel electrophoresis, and transfered to nitrocellulose. Membranes were probed with polyclonal antiserum at a 1:5000 dilution using the ECL Western blotting analysis system (Amersham).

Heat-shock experiments

Rescue of *psq*^{2403} mutant phenotype

The strain referred to as *X-3; psq*^{2403} has the genotype *w*^{1118} P[hs-*psq*]; *cn psq*^{2403} ; +/– *ry*^{506}. *w*^{1118} flies were used as a control. 3- to 5-day-old females (*n*=40-50) were aged with males for 2 days in embryo collection bottles at 25°C. Flies were subjected to a 1 hour heat pulse at 37°C in prewarmed vials, after which new males were provided. Timed collections on apple juice/agar plates were performed and hatch rates determined 48 hours postcollection. Fig. 7 shows combined data from two experiments; the curves were plotted by splining the data. For examination of heat-shock-induced Psq proteins, shocked flies were permitted to recover with new males at 25°C for 2 hours, after which ovaries were isolated and protein extracts prepared as described above.

Expression of the *Psq-l(3)S12* fusion protein in wild-type flies

Six strains containing insertions of P[hs-*psq-l(3)S12*] in a *w*^{1118} background were analyzed. *w*^{1118} was used as a control. Approximately 20 1- to 2-day-old female flies from each line were heat-shocked in prewarmed vials at 37°C for 1 hour, then transferred into fresh vials containing wild-type males and allowed to lay eggs for approximately 12 hours. Flies were then transferred into fresh vials for another 12 hour collection. Each day for 7 days the males were removed and the females heat-shocked as before. The number of eggs laid in each vial was counted immediately after the transfer of adults to fresh vials.

RESULTS

Analysis of *psq* mutants

Mutations at the *psq* locus cause a variety of defects in oogenesis and embryogenesis (see Table 1 for a summary of alleles). The phenotypes fall into three classes. The weak allele *psq*^{fs1} disrupts posterior segmentation and pole cell formation during embryogenesis (Siegel et al., 1993). The intermediate alleles *psq*^{HK38} and *psq*^{2403} exhibit dorsal/ventral defects in the eggshell and embryo in addition to the posterior axis defects described above (Figs 1, 2). Both of these intermediate alleles display a range of phenotypes; *psq*^{HK38} is significantly weaker than *psq*^{2403} (Figs 1, 2). Flies homozygous or transheterozygous for the strong alleles *psq*^{8109} and *psq*^{0115} have reduced viability, produce only a few, strongly dorsalized late stage eggs (Fig. 2), and have defects in early oogenesis. As expected, *psq*^{8109}/*psq*^{2403} transheterozygotes produce eggs with stronger dorsal/ventral eggshell defects than *psq*^{2403} homozygotes (Fig. 2). As is true for most mutants with eggshell defects, the number of embryos fertilized in all of the *psq* strains decreases with the increase in severity of the eggshell phenotypes.

Two other alleles do not fall readily into this classification scheme. The *PZ* insertion allele *psq*^{0482} is lethal, but acts as a weakly moderate allele in *trans* to all other alleles. *psq*^{0482} also fails to complement mutations in a nearby lethal locus (*lola*, E. Giniger, personal communication) that is a hotspot for *P* element insertion. In a complementation analysis between alleles of this lethal locus and all the *psq* alleles discussed here,

Table 1. *pipsqueak* alleles

Allele	Phenotype	Molecular nature
fs1	posterior group, grandchildless	*P[lacZ; w*$^+$*]* insertion (Siegel et al., 1993) into first intron of *psq*-1.
1-30	posterior group, grandchildless	Excision allele of *fs1* (Siegel et al., 1993). Deletes 5′-most exon and promoter for *psq*-1.
HK38	posterior group, weakly dorsalizes eggshell and embryo	EMS allele, unknown change (Schüpbach and Wieschaus, 1991).
2403	posterior group, moderately dorsalizes eggshell and embryo	*P[lacZ; ry*$^+$*]* insertion into largest *psq*-1 intron. Aberrant fusion protein created (Horowitz and Berg, 1995).
RF9	viable, fertile	Excision of *2403*. Removes all of *P* element except 3′ end (Horowitz and Berg, 1995).
RF13	viable, fertile	Excision of *2403*. Deletes entire *P* element and at least 40 kb surrounding insertion site (Horowitz and Berg, 1995).
8109	decreased viability, posterior group, strongly dorsalizes eggshell and embryo, early oogenesis defects	*P[lacZ; ry*$^+$*]* insertion into largest *psq*-1 intron. Aberrant fusion protein created. (Horowitz and Berg, 1995).
0115	decreased viability, posterior group, strongly dorsalizes eggshell and embryo, early oogenesis defects	*P[lacZ; ry*$^+$*]* insertion into largest *psq*-1 intron. Aberrant fusion protein created. (Horowitz and Berg, 1995).
0482	lethal	Two *P[lacZ; ry*$^+$*]* insertions: one in largest intron of *psq*-1 (Horowitz and Berg, 1995), second in 5′ end of *lola* (E. Giniger, personal communication).
Df(2R)psq-lola Δ18	lethal	Excision of both *P* elements in *0482*. Deletes DNA at *0482* site in largest intron of *psq*-1, extending through promoter for *psq*-1. Probable deletion of all DNA between *psq* and *lola*.

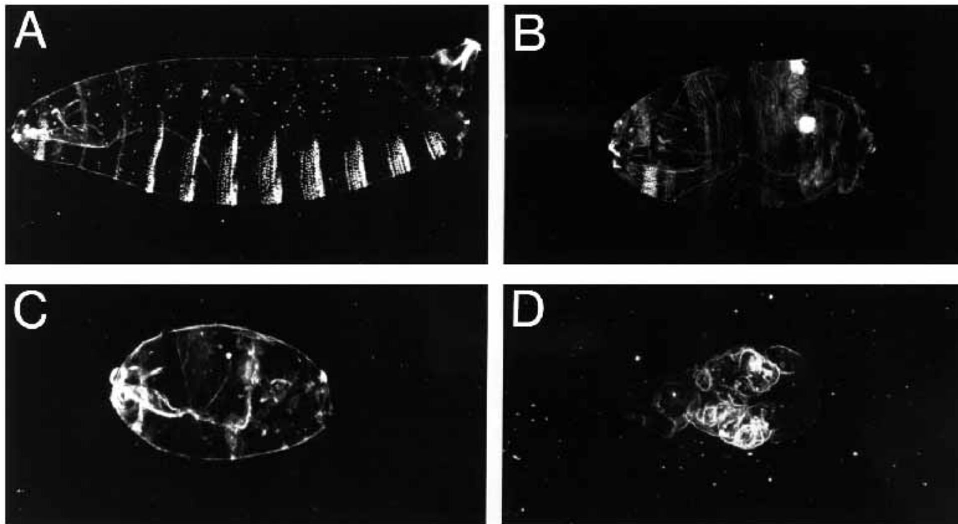


Fig. 1. Larval cuticle preparations from eggs laid by wild-type and *psq* mothers. Mutations in *psq* affect posterior segmentation. Some alleles also disrupt dorsal/ventral axis formation. (A) Canton S. (B) $psq^{1\text{-}30}/\Delta 18$ mothers produce embryos that lack posterior segments. Shown here is the most severe phenotype observed. (C) Cuticle from an egg laid by a psq^{HK38} mother in which posterior segments are absent, head defects are apparent and the dorsal hairs are slightly expanded into ventral regions. psq^{HK38} phenotypes range from weak posterior defects to strongly dorsalized, as in psq^{2403} mutants. (D) Cuticle from an egg laid by a psq^{2403} mother. This animal lacks posterior and head structures and is twisted.

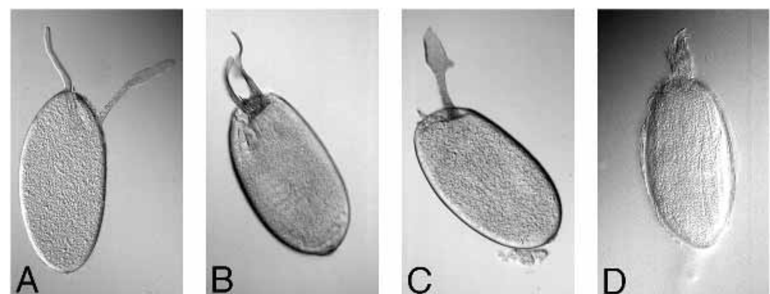
only psq^{0482} failed to complement mutations in both genes. Southern blot analysis demonstrates that psq^{0482} contains two *P* elements (data not shown), suggesting that both *psq* and the nearby hotspot locus contain *PZ* insertions.

A second allele that cannot be classified in our simplified groupings is the excision line $psq^{1\text{-}30}$ generated by mobilization of the *PW* element in psq^{fs1} (Siegel et al., 1993). As described below, this allele contains a deletion of the 5′ end of the *psq* gene. The $psq^{1\text{-}30}$ allele in homozygous flies produces defects early in oogenesis (Siegel et al., 1993), yet acts like a simple posterior group allele in *trans* to all other *psq* alleles. We have generated a larger deletion, $\Delta 18$ (see Materials and Methods) which encompasses and extends beyond deletion *1-30*. $psq^{1\text{-}30}/\Delta 18$ transheterozygotes produce embryos with a classic posterior group phenotype (Fig. 1); no defects in early oogenesis or in eggshell structures (Fig. 2) are observed. These *trans*heterozygotes are null for the *psq* 5.1 kb transcript (see below).

psq transcripts

To facilitate our understanding of the *psq* mutant phenotypes, we have conducted a molecular analysis of the gene. We previously reported the identification of a 5.1 kb *psq* transcript in wild-type females and the isolation of a 1.0 kb partial cDNA, 13-13, representing this message (Horowitz and Berg, 1995); we now call the 5.1 kb transcript *psq*-1. Using a 300 bp fragment from the 5′ UTR of *psq*-1 as a probe to northerns of RNA from wild-type adult males, females and isolated ovaries, two transcripts are revealed, the 5.1 kb *psq*-1 and a 6 kb message (Fig. 3). *psq*-1 is highly abundant in females but below the level of detection in males, while the 6 kb transcript appears to be present equally in both sexes. It should be noted that probes further 3′ in the *psq*-1 message detect additional transcripts by northern analysis (see below).

We have found that the *psq*-1 transcript is reduced in amount in females from all the *PZ* lines and is replaced by an abundant 1.6 kb transcript (Horowitz and Berg, 1995; and see *2403* and *0115/CyO* in Fig. 3A). In a previous study, we demonstrated that this product results from an



Genotype						n=
Canton S	100					379
Δ18/1-30	98.9	1.1				441
HK38	58.2	7.9	4.1	16.3	13.5	318
2403	6.1	24		65.2	4.4	342
2403/8109	1.4	4.5	6.7	80	6.7	371
8109/0015[a]		6.7	13.3	16	66.7	154
X-3;2403[b]	90.5	5.1		3.7	0.6	350

Fig. 2. Chorionic appendages of eggs produced by wild-type and *psq* females. (A) Canton S. (B) An egg from a psq^{2403} mother showing dorsal appendages that are fused at the base. (C) An egg from a psq^{2403} female showing a typical, single, spade-like dorsal appendage. (D) An egg, from a psq^{8109}/psq^{2403} transheterozygote showing a broader, thicker dorsal appendage characteristic of a more dorsalized phenotype. (E) Percent of eggs having wild-type or mutant eggshell structures in eggs laid by Canton S or *psq* females. (a) The values shown for psq^{8109}/psq^{0115} represent stage 14 egg chambers dissected from ovaries of transheterozygous flies, as these flies lay no eggs. (b) The *X-3;* psq^{2403} values were obtained from eggs laid by females kept for seven days at 24.5°C, without heat shock.

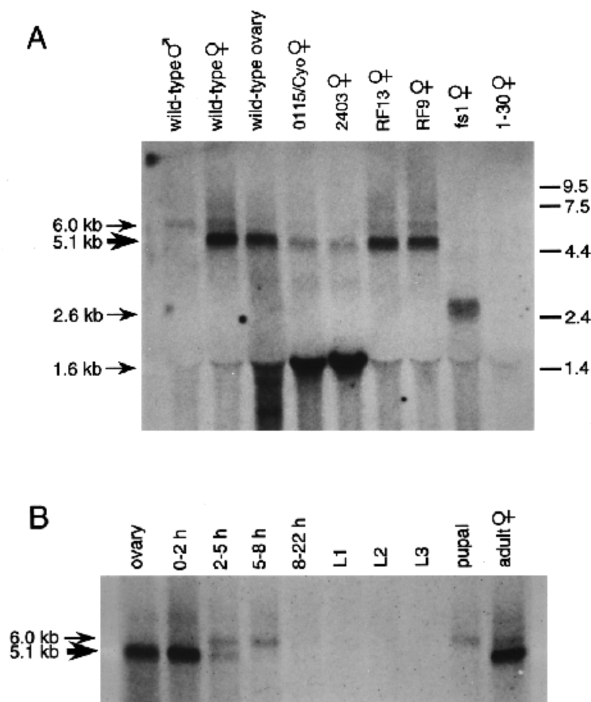


Fig. 3. *psq* encodes differentially regulated mRNAs. Northern blots in A and B contain total RNA and are probed with the 5′ UTR probe from cDNA 13-13 (see Fig. 6). Samples in A are from whole flies unless otherwise indicated; samples in B are from the indicated developmental stage. psq^{RF13} and psq^{RF9} are fertile revertants of the psq^{2403} allele. The 5.1 kb mRNA is *psq*-1. Position of RNA size markers are indicated to the right in A.

aberrant splicing event (Horowitz and Berg, 1995). The *PZ* elements in all four *PZ* mutant lines lie within a large intron of the *psq* gene. Sequences present within the partial *l(3)S12* gene residing on the *PZ* element provide a splice acceptor site and termination signal, causing the *psq*-initiated transcript to be spliced to *l(3)S12* sequence in the *PZ* element, then truncated. This truncated mRNA encodes a putative Psq-L(3)S12 fusion protein (Horowitz and Berg, 1995). Fertile revertants (such as psq^{RF13} and psq^{RF9}) generated by inducing excision of the *PZ* element in psq^{2403} delete part or all of the *l(3)S12* gene's splicing and termination signals and restore the wild-type *psq*-1 transcript (Fig. 3A, and Horowitz and Berg, 1995).

In the current study, we have extended our analysis of the *psq* gene by examining transcripts produced by other, non-*PZ* alleles of *psq*. In psq^{fs1} females, *psq*-1 is absent and is replaced by a transcript of approximately 2.5 kb (Fig. 3A). Sequence analysis reveals that the *PW* element of psq^{fs1} is inserted in a 3.2 kb intron in the 5′ UTR of the *psq*-1 message. The new 2.5 kb transcript observed in psq^{fs1} females most likely results from *psq*-initiated transcription that terminates within the *PW* element, as the transcript does not hybridize to cDNA sequences 3′ to the position of the *PW* insertion. RNA from the excision strain $psq^{1\text{-}30}$ produces no signal on the northern blot presented in Fig. 3A because the sequences hybridizing to the probe have been deleted from the genomic DNA. These results indicate that complete absence of the *psq*-1 transcript results in a weaker phenotype than observed for the *PZ psq* alleles, suggesting that the fusion protein produced in the *PZ* lines may be deleterious (see below). RNA from females harboring the EMS-induced allele, psq^{HK38} show no apparent change in size or abundance of *psq*-1 transcript (data not shown).

Developmental expression pattern of *psq-1*

The expression pattern of *psq*-1 on a developmental northern blot (Fig. 3B) conforms to the classic pattern observed for maternal mRNAs which are deposited into the developing oocyte. The message is abundant in ovaries and 0-2 hour embryos, but decreases to levels below detection in older embryos, larvae and pupae. Fig. 3B also demonstrates the expression of the 6 kb message described above, which is seen at varying levels throughout development.

We examined the expression of *psq* mRNA during oogenesis by in situ hybridization of the 5′ UTR *psq*-1 probe to whole ovaries. *psq* mRNA is expressed throughout oogenesis. Upon transport into the oocyte at stage 11, the message is uniformly distributed (data not shown).

Structure of the *psq* gene

The 5′ UTR *psq*-1 probe from cDNA 13-13 was used to screen an ovarian cDNA library (Stroumbakis et al., 1994). Sequence analysis of the longest clone, pHPT7-9, is presented in Fig. 4A.

The pHPT7-9 cDNA is 5162 nucleotides in length and contains an open reading frame encoding a protein of 1065 amino acids (M_r 115×10^3). The N terminus of the predicted Psq protein contains a BTB (or POZ) domain (underlined in Fig. 4A), a motif that has been shown to function in protein-protein interactions (Bardwell and Treisman, 1994; Chen et al., 1995). The BTB domain of *psq* has been cloned independently by virtue of its homology to other BTB domains in a study designed to isolate members of the BTB family of proteins (BTB-V in Zollman et al., 1994).

The Psq protein contains a number of repeated amino acid sequences, including multiple regions enriched for serine/glycine, proline or alanine residues, several opa repeats (Wharton et al., 1985), and a domain of histidine residues alternating with other amino acids (mostly glutamate or glycine). In addition, we have identified a 52-55 amino acid sequence, which we call the *psq* motif, that is repeated in tandem four times at the C terminus of the protein (Fig. 5). The *dTKR* gene of *Drosophila* (Haller et al., 1987) contains a single copy of the *psq* motif; the residues most highly conserved within the *psq* repeats are also conserved in TKR. A search of the sequence database revealed the presence of a variety of protein sequences from species as diverse as bacteria and humans with significant homology to the *psq* motif. The most highly conserved sequences are aligned in Fig. 5.

Alternative psq transcripts

Using a 1.2 kb fragment derived from the 3′ UTR of *psq*-1 as a probe to northern blots of RNA from wild-type females, we observed a heterogeneous band corresponding to an additional transcript(s) of approximately 4.4 kb in length. Of the mutant lines tested (psq^{2403}, psq^{HK38}, psq^{fs1} and $psq^{1\text{-}30}$), all retained the 4.4 kb transcript(s) (data not shown). The presence of the 4.4 kb transcript(s) in psq^{0115} or psq^{8109} homozygous flies was not tested due to the difficulty in obtaining suffficient flies from which to prepare RNA.

```
A
cacgagcactccgatgaacgacgagctgaaaaaataattattattttaaattccgttttgccagcgaaattcggtgcgtgcgcgaaaaagagtgcggaaaagcgcggttgagaaatttaa   120
tcggcccattatgaagcggtcaaagtgcggcgccaatgcaaatggttagccaaatggcaaaaaatatgtagcgatagcgtgccagcccagtaaaatcaaatagaatcaaatttaagtttc   240
ttgctttgcgaattatgaaaaaccaatcgcaatgcgtgtgtgtgtgtgtacgtgtggtgcgtgcgcgcgtgtgttagtgagcgagcgtagaagcagaaggcgtagcgagaaataaaaata   360
aaaaaccatacaataaaaagggccaacaaacgaaaaacagcagaagcaacaaaaacacacgacaagcggcagcaaaaacaacaacacaaaactggcggtaaatacgcatacagtacgata   480
caaaaaaataaactagaagcaacaacaaatacgcaacgtaacgggaactctttgacgtatttcctctattattatgcaaaaaggcaggcggcaaattgcagcagcagcgaaaatcaaata   600
taaatagcgaaaagtatactgcagtatacacacatttccgcagcgtgttgagcacgccacaagcataaacaaaagtgagcaacaggaattcagctaggattcggcgtgaaacgaatacga   720

aaaaccatacgaaATGGCAGCGGTTCGAGGACATCAGTATTTCAGTCTGCGCTGGAATAACTACCAGAACACGATGACGTCCGTGTTTCAGCAGCTGCGCGAGGATCTCTCCTTTGTGGA   840
             M  A  A  V  R  G  H  Q  Y  F  S  L  R  W  N  N  Y  Q  N  T  M  T  S  V  F  Q  Q  L  R  E  D  L  S  F  V  D   36

CGTCACTCTGTCCTGCGAGCATGGATCCCTCAAGGCGCACAAGGTCGTCCTGTCCGCCTGTTCGACATATTTCCAAAAACTACTGCTCGAGAATCCGTGCAAACATCCCACAATTATCCT   960
 V  T  L  S  C  E  H  G  S  L  K  A  H  K  V  V  L  S  A  C  S  T  Y  F  Q  K  L  L  L  E  N  P  C  K  H  P  T  I  I  L   76

TCCCGCCGACATTATCTTCACCGACCTGAAAACGATCATCGATTTTGTTTATCGTGGTGAAATCGACGTCACCGAATCGGAATTACAGGGCCTCTTGCGCACCGCCGAGCAGCTAAAGAT   1080
 P  A  D  I  I  F  T  D  L  K  T  I  I  D  F  V  Y  R  G  E  I  D  V  T  E  S  E  L  Q  G  L  L  R  T  A  E  Q  L  K  I   116

CAAGGGACTCTGCGAAACCGCCGAGAACGCGGATGACCTGAACGACGCGGCCACTGCAACGATCACAGTTTCAGAGAACATACAGCAGGCGGTCGTTGGCAATATAGTGAACGCCGCCAT   1200
 K  G  L  C  E  T  A  E  N  A  D  D  L  N  D  A  A  T  A  T  I  T  V  S  E  N  I  Q  Q  A  V  V  G  N  I  V  N  A  A  I   156

AGTTCCCGGAGCTCCGTCGCCCTCCTTGGAACAGCAGCAGCAACACCACCAACAGCAGCAACAGCAGCAGGCGCAGGAGCAGCACCAGCAACAGCAGGTGCATGCCCAGCAGCAACAGCA   1320
 V  P  G  A  P  S  P  S  L  E  Q  Q  Q  Q  H  H  Q  Q  Q  Q  Q  Q  Q  A  Q  E  Q  H  Q  Q  Q  Q  V  H  A  Q  Q  Q  Q  Q   196

GCAGCAGCAGCAAATAAACGCTGCCCTGCTTACGCAACACGGAGTGTCCAGCGGAAGTGTATCTCTTTCCGGGCAATTGCTGAGCTCCTCGGCTAGCGGCAGCAGCGGAAGTTTAGCCGG   1440
 Q  Q  Q  Q  I  N  A  A  L  L  T  Q  H  G  V  S  S  G  S  V  S  L  S  G  Q  L  L  S  S  S  A  S  G  S  S  G  S  L  A  G   236

CGGACAGCAGGCATCACAGACGCCATCCGGTTTGCAGCCGACGCCGCGCAAGTCTCGACTAAAGCGCTCCAAATCGCCAGATTTGTCATCGGGCGGAGGAGCAGGAAGCTCTGGAGGCTC   1560
 G  Q  Q  A  S  Q  T  P  S  G  L  Q  P  T  P  R  K  S  R  L  K  R  S  K  S  P  D  L  S  S  G  G  G  A  G  S  S  G  G  S   276

AAGTAGCGGCTCAACGCAGCAGCAACCGCAGCCGGCGCACCACCACCACCCGCAGACCATTCTCATCCAGGGCCAAAATCCGAATTCCATTGTGAGCCTTCAGCAGACGGCCGACGGCAA   1680
 S  S  G  S  T  Q  Q  Q  P  Q  P  A  H  H  H  H  P  Q  T  I  L  I  Q  G  Q  N  P  N  S  I  V  S  L  Q  Q  T  A  D  G  N   316

TTACATACCAGTTTCGGGTTCTGGCGACGATTCAGACGCCGAGGACCACGAGCACGAACACAACCACGGACACGGGCACGGGCACGGAGGACATACGCACGGACACGGTCATAGCCACGA   1800
 Y  I  P  V  S  G  S  G  D  D  S  D  A  E  D  H  E  H  E  H  N  H  G  H  G  H  G  H  G  G  H  T  H  G  H  G  H  S  H  E   356

GCACAGTCACGATCACGAGCACGAACATAAGCCGAACAAGATCTGCAAGACCGAGCACTCGGTGGCCTCGCCGGCGTCCAATTCATCCAGCGCCTCGAACAATGCAGCGGGCGTGAGCGG   1920
 H  S  H  D  H  E  H  E  H  K  P  N  K  I  C  K  T  E  H  S  V  A  S  P  A  S  N  S  S  S  A  S  N  N  A  A  G  V  S  G   396

CGTGACATCCACCGGCCAAGCCATCGTCACACAAATTGTAGTAGCGCGCGACGGAAAAGATACAAAAAATATGACTAGTTTAGGCATGGGCATGAACGGCGGCCTGTTGGGCGTGCCCAT   2040
 V  T  S  T  G  Q  A  I  V  T  Q  I  V  V  A  R  D  G  K  D  T  K  N  M  T  S  L  G  M  G  M  N  G  G  L  L  G  V  P  M   436
                                                                       *              *     *                          *

GGGATTCCTGGACTTCACACCCGAGCCACCAGCACCATCTGCCACCCCAGTCACCGTCACGGAGCACGTCGATCTGTCCTGCAACCCTAGCACGGACACGCGCGATCTATCAAATACCAC   2160
 G  F  L  D  F  T  P  E  P  P  A  P  S  A  T  P  V  T  V  T  E  H  V  D  L  S  C  N  P  S  T  D  T  R  D  L  S  N  T  T   476

CGAACCGCTCGATATAGACAATCATCTGGCTCAGCAGATCCATCGGCTGGATCAGTCCCCCATGCACTCGATATCGCACCATCATACAGGCGATGAGAGCAACTCGAATCTGGTGCAGCA   2280
 E  P  L  D  I  D  N  H  L  A  Q  Q  I  H  R  L  D  Q  S  P  M  H  S  I  S  H  H  H  T  G  D  E  S  N  S  N  L  V  Q  H   516

TATCAAGAGCGAAGTGATCGAGGCCAAGCACCTGGCTGCCCAGCATCATGCACTTAGCCAGGCCCAACAGCAACATGCCCATCATCAGGCCCACCAGCAGCATCAACAGCATCAGCAACA   2400
 I  K  S  E  V  I  E  A  K  H  L  A  A  Q  H  H  A  L  S  Q  A  Q  Q  Q  H  A  H  H  Q  A  H  Q  Q  H  Q  Q  H  Q  Q  Q   556

GCAGCATCAACAACAGCAGCAACATCTTCATGCCCAGCAGCTCTTGGCCCAGAGTCAGCTGCAGCAGCAACAGCAACAGCAACAGCAGCAGCAGCAACACCACCAGCAGCAGCAACAACA   2520
 Q  H  Q  Q  Q  Q  Q  H  L  H  A  Q  Q  L  L  A  Q  S  Q  L  Q  Q  Q  Q  Q  Q  Q  Q  Q  Q  Q  H  H  Q  Q  Q  Q  Q  Q   596

GGCAGCAGCAGCCGCCGCAGCAGCTGGTGTGCATGGCCAGCATGGGGGACATGTGACGCATGCCGACATCGGCGGTGCCACTGTCATGGAGATTGATCCAAGCCAGATTAAACACGAGCC   2640
 A  A  A  A  A  A  A  A  G  V  H  G  Q  H  G  G  H  V  T  H  A  D  I  G  G  A  T  V  M  E  I  D  P  S  Q  I  K  H  E  P   636

GGGCATGATAATAACGCCAGAGATTGTCAACATGATGTCCTCCGGGCATATGGACATGTATAACTCGGATACGAGTGAGGATTCGATGATGATCGCCAACGGTTCGCCGCACGATCAGAA   2760
 G  M  I  I  T  P  E  I  V  N  M  M  S  S  G  H  M  D  M  Y  N  S  D  T  S  E  D  S  M  M  I  A  N  G  S  P  H  D  Q  K   676

GGAACCGCACTATACGAATTTGGATCAGCAGCACGGTCTGGGTGGCAGCGTTTGCGGCCCGGGGCCCGGCGGTGCTGGTGGCGGTGGGGGCATGAGTGGTGGGGCTGGCTCTGGCTCTGG   2880
 E  P  H  Y  T  N  L  D  Q  Q  H  G  L  G  G  S  V  C  G  P  G  P  G  G  A  G  G  G  G  G  M  S  G  G  A  G  S  G  S  G   716

CGAGAAAGGTCAGTTTAATGGTCCCAAAGCTTGGACACAAGATGATATGAATTCAGCTTTAGATGCATTAAAGAACCAAAACATGAGCCTGACGAAAGCATCCGCCATTTATGGCATCCC   3000
 E  K {G  Q  F  N  G  P  K  A  W  T  Q  D  D  M  N  S  A  L} D  A  L  K  N  Q  N  M  S  L  T  K  A  S  A  I  Y  G  I  P   756

ATCGACCACCCTATGGCAACGTGCTCATCGCATGGGCATCGAAACGCCTAAGAAGGAGGGCGGCACCAAGTCGTGGAACGAGGATGCCCTGCAGAATGCCTTGGAGGCGCTGCGTTCGGG   3120
 S  T  T  L  W  Q  R  A  H  R  M  G  I  E  T  P  K  K  E  G  G  T  K  S  W  N  E  D  A  L  Q  N  A  L  E  A  L  R  S  G   796

TCAGATATCCGCGAACAAAGCGTCAAAGGCCTTCGGCATCCCATCCTCGACGCTCTACAAGATTGCACGGCGCGAAGGCATCCGCCTTGCGGCGCCCTTCAATGCCGCCCCGACCACGTG   3240
 Q  I  S  A  N  K  A  S  K  A  F  G  I  P  S  S  T  L  Y  K  I  A  R  R  E  G  I  R  L  A  A  P  F  N  A  A  P  T  T  W   836

GACGCCCGAGGATCTGGAGCGCGCCTTGGAGGCGATACGGGCGGGTAATACTTCGGTGCAGAAAGCCAGCGCTGAGTTTGGCATACCCACAGGAACACTTTACGGACGCTGCAAACGGA   3360
 T  P  E  D  L  E  R  A  L  E  A  I  R  A  G  N  T  S  V  Q  K  A  S  A  E  F  G  I  P  T  G  T  L  Y  G  R  C  K  R  E   876

GGGCATTGAGCTGTCGCGTTCGAATCCAACACCCTGGTCCGAGGATGCCATGAACGAGGCCCTGAACTCAGTGCGCGTTGGCCAAATGTCCATCAACCAGGCGGCCATTCACTATAACCT   3480
 G  I  E  L  S  R  S  N  P  T  P  W  S  E  D  A  M  N  E  A  L  N  S  V  R  V  G  Q  M  S  I  N  Q  A  A  I  H  Y  N  L   916

GCCCTACAGTTCGCTCTACGGCCGCTTCAAGCGCGGCAAATACGACGTGGTTGCCAACACGAGTGGCGTGGCGCTGTTGAACACCTCGGGCAACACCACCGGCAGCATTGAGATTATCGA   3600
 P  Y  S  S  L  Y  G  R  F  K  R  G  K  Y  D  V  V  A  N  T  S  G  V  A  L  L  N  T  S  G  N  T  T  G  S  I  E  I  I  E   956

GCACAGTCAGGAGAATTCGTTGCATATGTTACAGCAACAGTTCCCGTACAGTCCGTCGCCGCATCCGCCAACGCCGCAGCATCACAGCACGCCGCAGCACCACAGCTCGGCGCAGCAGGG   3720
 H  S  Q  E  N  S  L  H  M  L  Q  Q  Q  F  P  Y  S  P  S  P  H  P  P  T  P  Q  H  H  S  T  P  Q  H  H  S  S  A  Q  Q  G   996

TCCGCCGACGCCGCAGCACATGCAGCAGCACGTGGTGCACATGCAACAGCAGCAGCAGCAGCAGCAGCAACAGCAGTCGCCCCATCCGGGCCATCATCCGCAGCATATGCAGCAGGACGT   3840
 P  P  T  P  Q  H  M  Q  Q  H  V  V  H  M  Q  Q  Q  Q  Q  Q  Q  Q  Q  Q  S  P  H  P  G  H  H  P  Q  H  M  Q  Q  D  V  1036

GGTTACATCGAGCAGCCAGGTGGTGCACAGCCAACAGCAGCAGCAGCTTCAGCAGATCTATCAGCACCACGGGACGCCGGAGCGTAGTTGAgaatcactgcatgctgcgacattaatggg   3960
 V  T  S  S  S  Q  V  V  H  S  Q  Q  Q  Q  Q  L  Q  Q  I  Y  Q  H  H  G  T  P  E  R  S  •   1065

cagcaccggcggtggaggaggtggtggagttggtggagtagcagccagcactagcggcggcacggcggccagtagtacgccctttgccagcatttcggcgctggtcagcggaccgtccag   4080
tgccccaccaccgccggtaactgccagtggctcgtcggcagtcagtggcggcaccagcaagcgcaagaaagccaataagcaaaagcgttgagagcaggaggatgatcagccacacaggga   4200
gttggaaacaattgcgggagtggccaggcccgaggacatcaagtttagatttaattgcaagcactggaagcgcaagttgcggcagtttcatttgctattgtccaaacagtagtcagtcag   4320
tcagttaggaaggtacacacatatatgaatgacccacagacaaaaaattcccaagggtaggataaagcaatcacaaaaccactttcaactagatggacattaattacacaaccgagatg   4440
tgtataaaatgtaaatatataacgaaatatgctttatttacattatacggaacatagttatataggttttaggttcgcgaaagtcctttgttctagttaggtaaggtaaggtggagaaag   4560
gagaaagttaagaaaacatacctgaacgtaagagaatgatgaattttagatttgcctctagactcgaaagttaatggattcgttttgggctcgcagttttaagagcactttgttatcaa   4680
tggatataaacacgcgtatttatttggtctcgcgaccattctgagaaccttgaatcggcagttgattggaactagattgtttcccttcaatgtgcttatgtgctttgtt   4800
ataattttcgcttagtgtgagactgaaagtgcgcctaaaactgttgaaaaaccccatgcccaatattatagtgaagacgcataaaacacaacccatacatgaatcaaatgctttataacc   4920
atattgtttgaacagatagaacgaagtgggagaatacctttaacgtagaattaaaccaaattatcaaggttctagaacttaaatcattagctttgtaacgataaaagaaaggaaaacgagtt   5040
atttgaattttactggaggcgtcctgaattttggattgtatattgatttgatttcaactgaattaagaatatagccaataaataaatgtagaaatgaagtaattgcaaaaaaaaaaaaaa   5160
aa   5162

B
gtttttgtatgaatttaattccgtcacgtcgtctactacaatttgtcaactttaaagcgaaaaaacaaatcgaaatcaactaggcttggttccctggaataattaaccaacagtccttc
gatcaag...

...aaggaatcaacaaatggaatgtattttgatagaaccggaaacagaagcagctgaaattaagcgaatcaaactgcaaatgttgttaatactaagctcttaaataatgtcctttgaag
tcacattcaaacaaagatttaaaacaagcttgtaagaaaaagaaacc aataaa tacacacgtatatatatggatacaaaaaaaaaaaaaa
```

Fig. 4. Nucleotide and predicted amino acid sequence of *psq*. (A) The nucleotide and amino acid sequence of *psq*-1 (cDNA pHPT7-9). Single underline, BTB domain. Double underline, region of *psq* motif repeats. The position of amino acids deleted in the *psq*-1 alternatively spliced cDNA pHH62 (see text) are bracketed. Tailed arrow and arrowheads indicate position of introns. First intron, position of the *PW* insertion of *psq*^{fs1}. Tailed arrow, position of the large intron into which all four *PZ* elements are inserted. *, clustered methionines that could initiate translation in *psq*-2. Underlined italics, position of polyadenylation signal. (B) Sequence of the alternative 5′ (upper) and 3′ (lower) ends of *psq*-2. The 5′ end sequence joins to the sequence in A at the position indicated by the tailed arrow. The 3′ sequence continues from position 5145 of the sequence in A. The polyadenylation signal is underlined.

We used the entire *psq*-1 cDNA as a probe to rescreen the ovary cDNA library in search of alternative *psq* transcripts. One new cDNA, pHH63, contains an insert of approximately 4.4 kb and may correspond to the 4.4 kb transcript described above. We call the message encoded by pHH63 *psq*-2. *psq*-2 has 127 bp of unique sequence at the 5′ end (Fig. 4B) which is spliced to the fourth exon (position 1049, Fig. 4A) of the *psq*-1 sequence. The unique 5′ end of *psq*-2 is derived from within the large intron of *psq*-1 (see Fig. 6). Notably, the *psq*^{RF13} deletion line retains the sequences encoding the 5′ end of *psq*-2. The *psq*-2 transcript also differs from *psq*-1 by the presence of an additional 192 bp of sequence at the 3′ end (Fig. 4B). This novel 3′ sequence contains a consensus polyadenylation signal 29 bp 5′ to the poly(A) tail of the cDNA. Extensive restriction analysis suggests that, with the exception of the alternative 5′ end and the extended 3′ end, the sequence of this cDNA is identical to bases 1049 to 5146 of *psq*-1. The best candidates for translation initiation are four in-frame methionine codons clustered within a 17 amino acid stretch of sequence, positioned in the same frame as the conceptual protein encoded by *psq*-1 (asterisks, Fig. 4A). Translation initiated at one of these methionines is predicted to produce a protein ranging in length from 630 to 646 amino acids and which is identical to the carboxy-terminal two-thirds of the *psq*-1-encoded protein. Interestingly, the *psq*-2-encoded protein would completely lack the BTB domain found in the larger protein (see Fig. 6).

Another cDNA, pHH62, which contains an approximately 4 kb insert, was found to contain yet another transcript variant. This cDNA is distinguished from the *psq*-1 transcript by the presence of a 54 base-pair deletion extending from position 2889 to 2942. As the sequence flanking the ends of the deleted

Fig. 5. The *psq* motif is conserved. The four C-terminal *psq* motif repeats are aligned with homologous protein sequences. Numbers to the left indicate the position in the amino acid sequence of each protein. Homologous genes are: dTKR, *Drosophila* Tyrosine Kinase Related gene (Haller et al., 1987); lac F, lac permease of *Agrobacterium radiobacter* (Williams et al, 1992); T01C1.3, predicted protein from the *C. elegans* genome project. For a position to be shaded, at least three identical or similar amino acids must be present in the *Drosophila* sequences and no ambiguities in alignment can exist. Identities are indicated in black; similarities in grey.

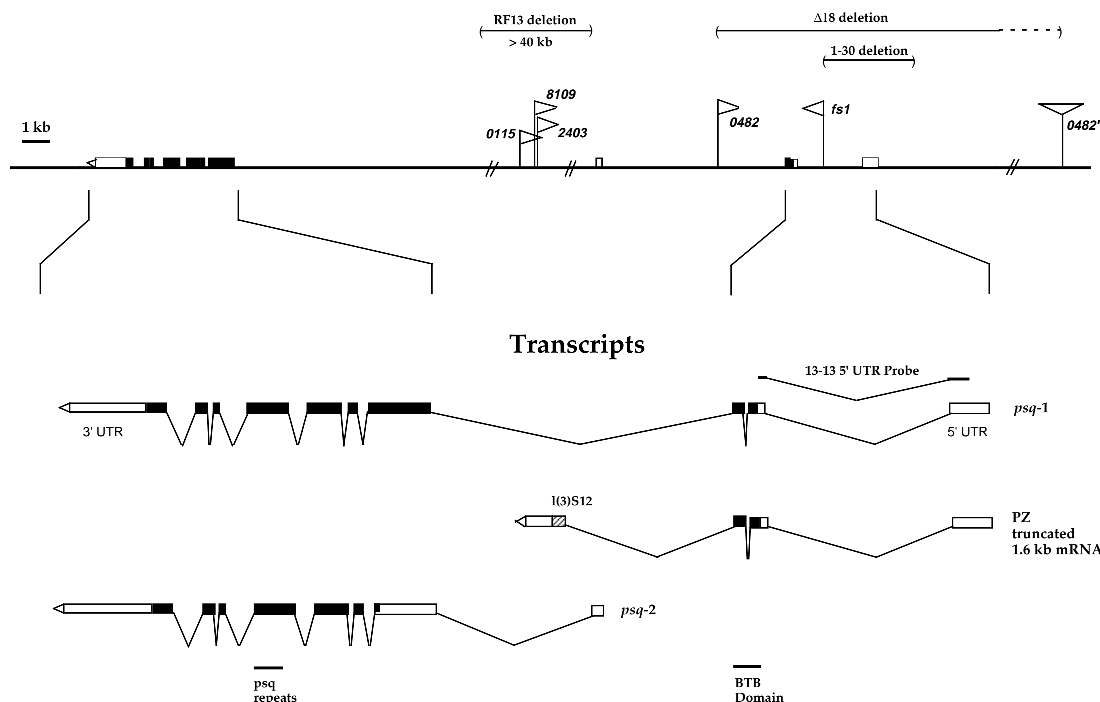


Fig. 6. *psq* gene structure. The map indicates the position of all *P* element insertions. Flags indicate the orientation of the element (wide part corresponding to the 5′ *P* element end). In the *psq*^{0482} line, two *P* elements are present; the orientation of 0482′ has not been determined. The extent of all deletions is indicated by parentheses. The boxes on the genomic DNA line show the relative position of exons for the *psq*-1 transcript. *psq* transcripts are shown in expanded form below the map. White boxes indicate 5′ and 3′ UTR sequences. Black boxes indicate predicted *psq* coding sequence. Hatched box indicates putative *l(3)S12* coding sequence fused to the BTB domain of *psq* in *PZ* mutants. Position of the BTB domain and four *psq* motif repeats are indicated, as is the source of probe used for the Northern blots in Fig. 3.

region provide excellent matches to the consensus splice donor and acceptor sequences (Mount et al., 1992), this deletion most likely represents an alternative splicing event. Such a splice perfectly deletes 18 amino acids but retains the open reading frame of the protein (Fig. 4A, bracketed amino acids). These deleted amino acids are located in the first of the four *psq* motif repeats described in Fig. 5. The 3′ end of pHH62 contains the same extended 3′ UTR found in *psq*-2. The 5′ end of the pHH62 cDNA begins within the coding region of the *psq*-1 sequence, at position 944, and thus it is unlikely that this represents the true 5′ end of this transcript. If the 5′ end of the message is the same as that for *psq*-1, a protein of 113 x 10^3 M_r, slightly smaller than that encoded by *psq*-1, should be produced.

It is clear that *psq* is a complex gene; there may be additional alternative ovarian transcripts for which we have not yet isolated cDNAs. We have chosen to focus on *psq*-1, which encodes a protein affected in all of the mutants that we have characterized (see below). Additional experiments reveal that mutation of *psq*-1 is the primary source of the defects observed in all of the mutants discussed here (see below).

The *psq*^{2403} mutation is rescued by expression of the cloned *psq*-1 mRNA

To verify that cDNA pHPT7-9 (*psq*-1) corresponds to the *psq* gene, we tested whether expression of the protein encoded by this cDNA could rescue *psq* mutants. We introduced a fragment of the cloned cDNA under the control of the hsp70 heat-shock promoter into various mutant lines and asked if expression of the transgene can restore fertility and/or wild-type eggshell phenotypes to the mutant females.

psq^{2403} females carrying an X-linked copy of the P[hs-*psq*] transgene (*X-3*; *psq*^{2403}) lay eggs at a frequency similar to that observed for the *w*^{1118} control. Up to 90% of the eggs laid in the absence of heat shock possess two dorsal appendages (see Fig. 2E), indicating that the low level expression of the hsp70 promoter in the absence of heat shock is sufficient to rescue the eggshell defect in this mutant line. These eggs (with the exception of a few rare escapers) do not hatch.

Heat-shock induction of the transgene in *psq*^{2403} females restores fertility. We conducted experiments to determine the oogenic stage at which *psq* is first required by subjecting *X-3*; *psq*^{2403} mutant females to heat shock and determining the hatching frequency of eggs laid at defined intervals following induction. A peak in the hatching frequency (as high as 70-80%) occurs approximately 80 hours post heat shock (Fig. 7). This timing corresponds to a requirement for the Psq protein prior to stage one of oogenesis (King, 1970; Mahowald and Kambysellis, 1980; Lin and Spradling, 1993). The lower levels of rescue observed prior to the peak at 80 hours post heat shock may reflect an additional requirement for *psq* function at multiple times throughout oogenesis.

Heat-shock induction of *psq*^{0115} females carrying the p[hs-*psq*] transgene results in the production of eggs with two dorsal appendages in over 60% of the eggs laid, but none of these eggs hatch (see below).

psq encodes multiple proteins, including a protein localized to follicle and nurse cell nuclei

In order to characterize the protein products of the *psq* gene, polyclonal antibodies were raised against bacterially expressed Psq protein. These antibodies, AS1 and AS2, recognize slightly different substrates on western blots of ovarian extracts.

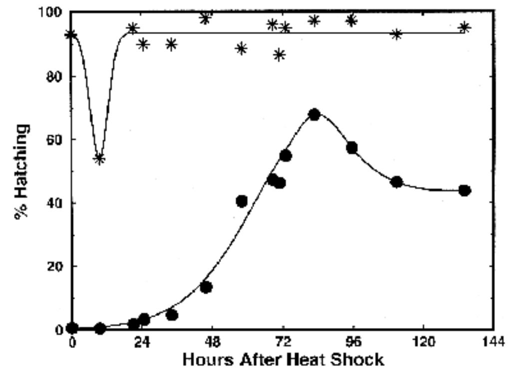
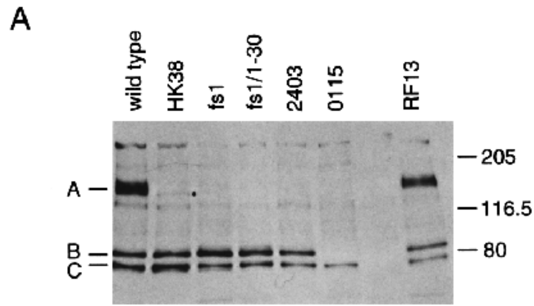


Fig. 7. Eggshell and fertility defects of *psq*^{2403} females are rescued by expression of the protein encoded by *psq*-1. The graph shows hatch rates of eggs laid by *X-3; psq*^{2403} females and *w*^{1118} controls at specified times after heat shock. (*) *w1118* (●) *X-3; psq*^{2403}.

On a western blot, immune serum AS1 recognizes at least two Psq protein species in wild-type ovaries (Fig. 8A). PsqA migrates as a broad band corresponding to a M_r of approximately 150×10^3. This species is found in protein extracts prepared from the ovaries of wild-type females (Fig. 8A), but is absent in mutants *psq*^{fs1}, *psq*^{fs1}/*psq*$^{1-30}$, *psq*^{2403} or *psq*^{0115} (Fig. 8A). Ovary extracts of the EMS-induced allele, *psq*^{HK38}, show a weak but reproducible signal at the fastest migrating position of the PsqA band (Fig. 8A). The PsqA band is restored in fertile revertants of *psq*^{2403} (Fig. 8A, RF13 lane) and in flies carrying the heat-shock *psq*-1 rescue construct after treatment at 37°C (Fig. 8B, *X-3; 2403* + lane). These findings demonstrate that protein PsqA is encoded by transcript(s) from cDNA pHPT7-9 (*psq*-1). The diffuse nature of the protein A signal could result from post-translational modification of the protein (e.g. glycosylation and/or cleavage). It is also possible that the band contains multiple protein species deriving from alternatively spliced messages. As described above, cDNA pHH62 is most likely a partial copy of its cognate message. If the true 5′ end of pHH62 extends to the same initiator methionine as that of the *psq*-1 mRNA, then the deletion of 18 amino acids of sequence observed in this cDNA would produce a slightly smaller protein. The observation that expression of the P[hs-*psq*] construct in *psq* mutant flies restores a broad band of PsqA is consistent with either modification of a single protein or the possibility that more than one species of mRNA, and consequently more than one species of protein, are produced from this construct.

Several additional protein species are recognized by the AS1 immune serum. One of these, PsqB, migrates with an apparent M_r of 80×10^3, and could be encoded by the *psq*-2 transcript. Expression in bacteria of a protein initiating at methionine 436 (Fig. 4A), one of the potential start sites in *psq-2*, produces a protein with the same relative mobility as PsqB (Fig. 8D, pHH70 lane). PsqB is absent in extracts prepared from ovaries of *psq*^{0115} and *psq*^{0115}/*psq*^{8109} females (Fig. 8A and data not shown), but is unaffected in any of the other mutants tested. Expression of the P[hs-*psq*] construct in *psq*^{0115} females produces the broad PsqA band, but does not restore the PsqB band (Fig. 8C, *X-3; 0115* + lane). The absence of PsqB in



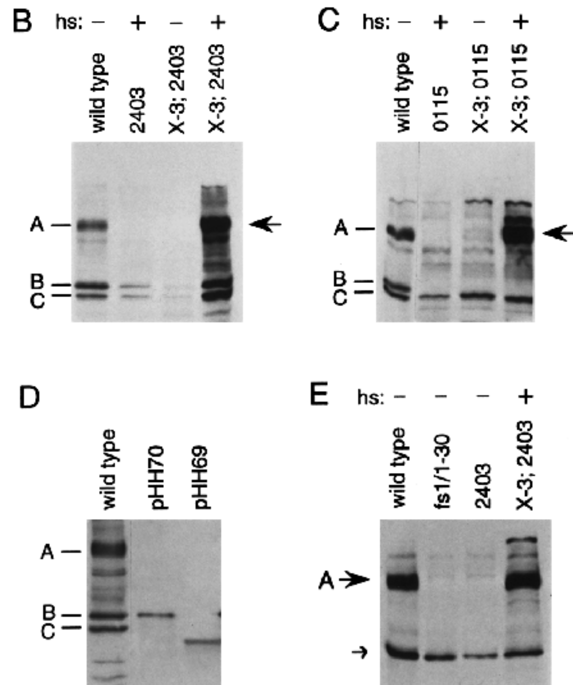


Fig. 8. *psq* proteins. Western blots of ovarian protein extracts from wild-type and *psq* mutant flies probed with AS1 (A-D) or AS2 (E) antisera. Lanes marked pHH70 and pHH69 in D contain bacterial extracts from the indicated expression clones (See Materials and Methods). The position of *psq* proteins A and B are indicated. It is not known whether the band labeled C is a protein encoded by the *psq* locus (see text). • (in A), faint but reproducible signal observed in extracts from psq^{HK38} females. The arrow in E indicates a protein species that migrates with the same mobility as B, but is either not B or not exclusively B based upon its presence in psq^{0115} extracts. Numbers to the right in A indicate positions of prestained MW markers ($M_r \times 10^3$).

psq^{0115} females expressing the P[hs-*psq*] construct may explain our observation that fertility is not restored in these flies.

On a western blot, the AS1 immune serum also recognizes a third protein, which is not affected in any of the *psq* mutants examined. AS2 immune serum does not recognize this protein (Fig. 8E); we do not know whether this protein is encoded by the *psq* locus.

AS2 immune serum recognizes PsqA and only one additional protein species (Fig. 8E). We used this antiserum to determine the localization of PsqA in whole ovaries by comparing the patterns observed in wild-type and psq^{2403} females. The results are shown in Fig. 9A-E. PsqA is abundant early in oogenesis, and is first detected at the anterior tip of the germarium, suggesting a role in the germline stem cells. PsqA is particularly abundant in the nuclei of the germ cells in region 2a, and is observed in the pro-follicular layer as well as the follicle cells of later stage egg chambers. The protein remains nuclear in both soma and germline throughout oogenesis. There is particularly strong staining in the anterior and posterior polar follicle cells of egg chambers during stages 5 and 6. A very distinct pattern is seen in the oocyte nucleus as two small and discrete spots (Fig. 9C). The staining observed in psq^{2403} ovaries is greatly reduced in comparison to the wild

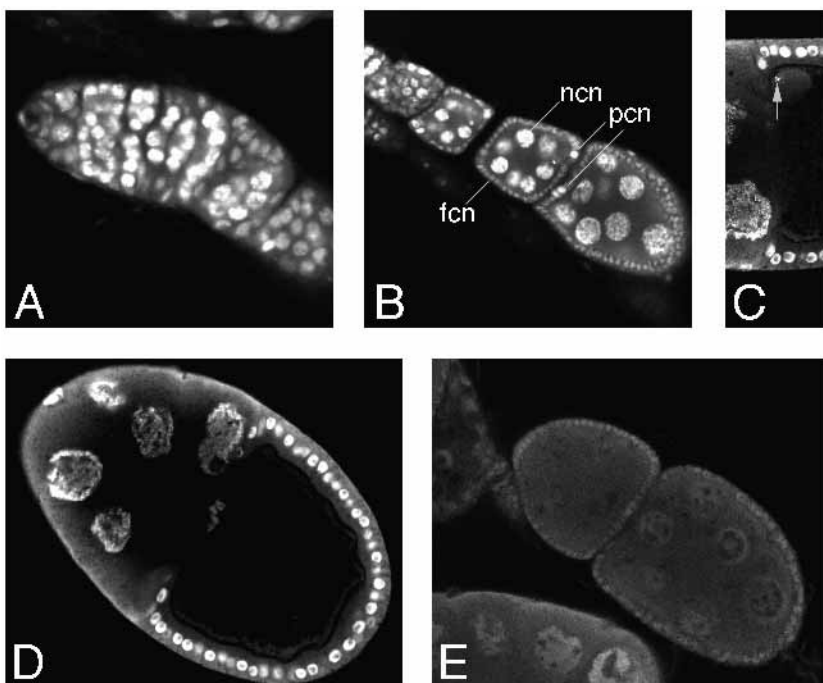


Fig. 9. Whole-mount immunocytochemistry of wild-type (A-D) and psq^{2403} (E) ovaries using antisera AS2. PsqA is localized to the nuclei of somatic and germline cells throughout oogenesis. (A) Germarium showing initial expression of PsqA in region 1, with intense fluorescence in the germ cells of region 2. The prefollicular somatic cells fluoresce weakly. (B) Ovariole showing fluorescence in follicle cell nuclei (fcn) and intense fluorescence in polar cell nuclei (pcn). Strong fluorescence continues to be observed in nurse cell nuclei (ncn). (C) Oocyte/nurse-cell border of a stage 10 egg chamber. Arrow indicates PsqA protein in the oocyte nucleus. This fluorescence frequently appears as two dots. (D) Stage 10 egg chamber showing fluorescence in nurse cell and follicle cell nuclei. (E) Previtellogenic egg chambers of a psq^{2403} ovary showing extremely weak fluorescence.

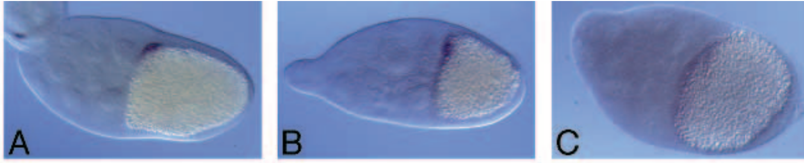


Fig. 10. In situ hybridization of wild-type (A) and psq^{PZ} (B-C) egg chambers using *grk* sequences as a probe. Lower levels of message in *PZ* egg chambers required that the color reaction proceed for twice the length of time as was utilized for wild-type egg chambers to obtain a photographable representation of the localization pattern. (A) In Canton S egg chambers, *grk* mRNA is localized in a tight cap above the oocyte nucleus. (B) In psq^{2403} egg chambers, *grk* message is localized to the dorsal region of the oocyte, but sometimes appears more diffuse than in wild type. (C) In approximately 25% of psq^{8109}/psq^{2403} egg chambers, *grk* message localization is abnormal.

type (Fig. 9E), indicating that the patterns described above predominantly reflect the distribution of PsqA.

grk localization is affected in psq mutants

The eggshell and cuticle defects that we observed with some *psq* mutant alleles suggested that dorsal/ventral patterning is disrupted in these mutants. To determine if this effect is mediated by the *gurken* signal transduction pathway, we examined the localization of *grk* mRNA in ovaries from *psq* mutants. In ovaries from wild-type females, *grk* mRNA is localized to the oocyte in early egg chambers and by stage 8 is seen as a tight cap on the dorsal side of the oocyte nucleus (Fig. 10A). In egg chambers from psq^{2403} homozygous females and psq^{2403}/psq^{8109} transheterozygous females, the amount of *grk* mRNA is reduced. In psq^{2403} mutants, *grk* mRNA is localized correctly in 90% (*n*=46) of stage 10 egg chambers (Fig. 10B). In the remaining 10% of egg chambers, *grk* mRNA is found as a diffuse band in the dorsal, anterior region of the oocyte, or fails to be localized at all. In psq^{2403}/psq^{8109} mutants, *grk* message is mislocalized in 24% (*n*=29) of stage 10 egg chambers. In these cases, message is either uniformly distributed throughout the oocyte or present in a ring at the anterior end (Fig. 10C). The combination of reduced levels and mislocalization of *grk* mRNA could lead to the unusual eggshell phenotypes that we observe in *PZ psq* mutants (Fig. 2). It is also possible that some contribution to the mutant phenotype is due to defects downstream of *grk* in the follicle cells.

Expression of a Psq-L(3)S12 fusion protein in wild-type flies causes sterility in females

Analysis of the mutant phenotypes associated with the *PZ* alleles suggests that the aberrant splicing of *psq* into the *l(3)S12* sequences in these lines creates a fusion protein with negative effects on oogenesis. To test this possibility, we expressed the predicted Psq-L(3)S12 fusion protein in transgenic flies. We generated a construct, P[hs-*psq-l(3)S12*], in which a mRNA encoding the fusion protein could be expressed by heat-shock induction from the hsp70 promoter. Six independent transgenic lines containing homozygous insertions of this construct were placed on a daily heat-shock regime and their egg-laying abilities compared to w^{1118} control animals. Three and a half days following the initial heat shock, all six transgenic lines demonstrate a dramatic drop in the number of eggs laid (Fig. 11, and data not shown), while control animals continue to lay eggs at a high rate. Dissection of the ovaries from these animals reveals a decreased number of egg chambers at all stages. In addition, a significant proportion of late stage egg chambers are short in length, with partially dorsalized eggshell structures. Thus, overexpression of the Psq-L(3)S12 fusion protein is sufficient to render flies sterile, phenocopying weak *PZ* alleles.

DISCUSSION

We report here the cloning and molecular characterization of the *Drosophila psq* gene, which encodes a novel nuclear protein involved in axis definition. Our results demonstrate that *psq* is a complex gene encoding multiple differentially spliced transcripts and several protein isoforms. In the course of submitting this manuscript, a characterization of the embryonic *psq* transcripts appeared in the literature (Weber et al., 1995). A sequence comparison of the ovarian and embryonic *psq* transcripts reveals that they have distinct 5′ and 3′ ends, and that the proteins encoded by transcripts from the two tissues differ slightly at their C termini (see Materials and Methods for details).

Protein structure

We have focused our efforts on the characterization of *psq*-1, the mRNA that encodes PsqA. This protein, which is drastically reduced or absent in all of the *psq* mutants that we have examined, has a number of interesting features. At the amino

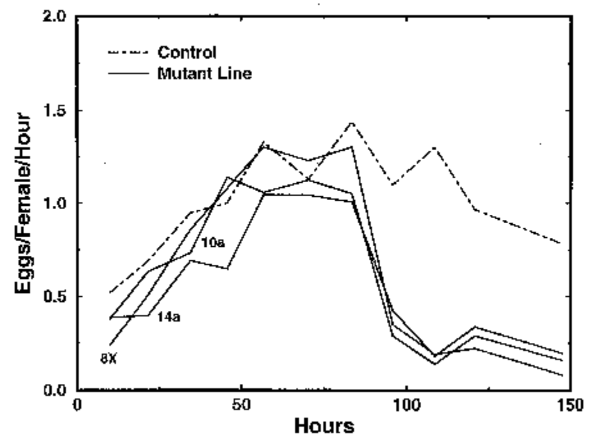


Fig. 11. Production of the PsqA-L(3)S12 fusion protein in wild-type females induces a dramatic decrease in egg-laying approximately 4 days following heat shock. Two samples of w^{1118} control animals and six independent insertion lines carrying the P[hs-*psq-l(3)S12*] transgene were placed on a daily heat-shock regimen. Number of laid eggs were counted at 12 hour intervals. Shown here are a weighted average of the two control samples (dashed line) and three of the six experimental lines (solid lines). Females laid more eggs at night than during the day, presumably due to circadian rhythms and to the effect of the daily heat shock.

terminus, PsqA contains a BTB domain (Godt et al. 1993; also referred to as a POZ domain by Bardwell and Treisman, 1994), a motif that has been shown to function in protein-protein interactions. Although BTB domains are often found near the N terminus of Cys_2-His_2 zinc finger proteins, PsqA does not appear to contain a zinc finger. Downstream of the BTB domain, PsqA contains 34 alternating histidine residues, $(HX)_n$, a motif that is present in a number of other *Drosophila* proteins, primarily transcription factors. It has been proposed that these histidine repeats could mediate protein-protein interactions by coordinating metal ions to form a 'histidine-metal zipper' between two proteins containing the repeats (Janknecht et al., 1991). The presence of two potential protein-protein interaction domains suggests that PsqA monomers may interact with each other or with heterologous protein species. Additionally, PsqA contains four tandem copies of a conserved sequence of unknown function at its carboxy terminus; we call this sequence the *psq* motif.

We have identified another *psq* transcript, *psq*-2, which encodes a second isoform of the Psq protein lacking both the BTB and $(HX)_n$ domains but retaining the *psq* motif repeats. This isoform is unlikely to be involved in the same type of protein-protein interactions predicted for PsqA. The finding of Bardwell and Treisman (1994) that a BTB-containing zinc finger protein was more competent to bind DNA when the BTB domain was removed raises the possibility that the isoform encoded by *psq*-2 could be an activated form of the Psq protein.

Dominant mutants

We previously hypothesized that the *PZ* fusion protein might act in a dominant-negative fashion, possibly by disrupting normal protein-protein interactions mediated by the BTB domain of the protein (Horowitz and Berg, 1995). This prediction was based on genetic analyses indicating that the production of a truncated protein is more harmful than production of no protein at all. Our observation that the psq^{2403} allele shows eggshell defects whereas the psq^{fs1} and $psq^{1\text{-}30}$ alleles do not, despite the fact that all three mutants display the same loss of PsqA using our antibodies, provides additional evidence that the fusion protein is deleterious.

Our experiments with P[hs-*psq-l(3)S12*] reveal that expression of an isolated BTB domain can behave in a dominant manner. Expression of this construct in wild-type flies leads to marked reduction in egg-laying. Furthermore, the presence of two copies of the transgene causes a more severe effect on egg production than a single copy (data not shown). These observations, along with the results of the rescue experiments (see below), suggest that the dominant nature of the *PZ* mutation is sensitive to the ratio of the *PZ* and wild-type forms of the Psq protein, a notion that we previously forwarded, based on the absence of eggshell defects in *PZ*/+ flies (Horowitz and Berg, 1995).

A model for Psq-L(3)S12 action

As no PsqA can be detected in any of the *psq PZ* alleles, the eggshell defects that we observe in these lines cannot be due to interaction of the *PZ* fusion protein Psq-L(3)S12 with PsqA. However, if PsqA normally interacts with a heterologous BTB-containing protein, the presence in the *PZ* mutants of the PsqA BTB domain without its associated carboxy-terminal sequences may result in inappropriate activation or inactivation of the heterologous protein.

Recent work by two groups lends support to this model for Psq-L(3)S12 action. Bardwell and Treisman (1994) and Chen et al. (1995) have demonstrated that the BTB domain can promote dimerization or multimerization of proteins in either homomeric or heteromeric complexes in vitro. Interestingly, the presence of the BTB domain is associated with a reduction in DNA binding by the associated zinc finger region in several zinc finger BTB proteins tested (Bardwell and Treisman, 1994). Of particular relevance to our work is the finding that coexpression of a protein fragment containing only the BTB domain along with the full-length BTB-containing protein relieves the BTB-mediated inhibition of DNA-binding in vitro (Bardwell and Treisman, 1994). Thus, interaction of the BTB domain present in the *psq PZ* fusion protein with that of a heterologous BTB-containing DNA-binding protein might inappropriately activate the binding activity of that protein.

In contrast, expression of a BTB domain without its associated C terminus (eg. a BTB domain fused to heterologous sequence) may inactivate proteins that interact with the mutant BTB protein. Several examples of human cancers associated with translocations that may involve the joining of a BTB domain from one protein to another protein have been identified (see, for example, Chen et al., 1993). This observation has led to the suggestion that transformation could arise by a dominant-negative mechanism involving sequestration of partner BTB domain proteins in inactive complexes (Bardwell and Treisman, 1994; Chen et al., 1995). Our results provide the first direct evidence supporting the hypothesis that an abnormal BTB-containing protein can exert a dominant effect in vivo.

The dominant effect of the fusion protein is apparent in the dorsal/ventral defects of the *PZ* mutants. Because proper localization of *grk* message is required for dorsal/ventral polarity (reviewed by Schüpbach and Roth, 1994), we examined *grk* mRNA localization in the *PZ* lines. We find that *PZ* mutants with more severe eggshell defects display a higher frequency of *grk* mRNA mislocalization. Interestingly, in wild-type egg chambers, we see an abundance of PsqA in polar cells at stages 6-7, a time when these cells signal to the oocyte. This signaling results in a repolarization of the microtubules in the oocyte, a process required for proper anterior/posterior and dorsal/ventral axis definition (Ruohola et al., 1991; Lane and Kalderon, 1994). It is possible that the mislocalization of *grk* mRNA in *psq PZ* mutants is due to disruption of the signaling between the polar follicle cells and the oocyte. In addition, these mutants may affect processes that occur in the dorsal follicle cells during later stages, which could contribute to the dorsal/ventral defects observed. These effects may be produced by interaction of the mutant *PZ* protein with factors normally associated with the wild-type PsqA. Alternatively, the *PZ* fusion protein may be interacting spuriously with proteins involved in these signaling cascades.

Rescue of *psq* mutants

Our heat-shock rescue experiments suggest that *psq* is required very early, prior to stage 1 in oogenesis. These results are consistent with the early oogenesis defects observed in the most severe *psq* mutants and with the abundant level of Psq protein detected in the 16-cell cysts in region 2 of the germarium. Interestingly, the rate of hatching of heat-shocked *X-3; psq*^{2403}

flies does not drop back to the pre-shock value, even after 160 hours post heat shock (data not shown). Rather, hatching appears to plateau at a value intermediate to the pre-shock and peak value of hatching observed. This may indicate that a permanent change is effected by expression of *psq* and that continued expression of the gene is not necessary to maintain fertility. Our finding that PsqA is present in the very tip of the germarium, and thus may be expressed in the stem cells, is consistent with this scenario. Alternatively, the PsqA protein produced upon heat shock may be quite stable, such that fertility drops off slowly after the heat-shock peak is achieved.

PsqA plays a role in regulating gene expression during oogenesis

Antibody staining of ovaries shows that PsqA is a nuclear protein, expressed in both the somatically derived follicle cells and in the germline. The apparent association of the PsqA protein with chromatin in the nurse cells, and the specific fluorescence in distinct dots in the oocyte nucleus, argue that PsqA protein is associated, directly or indirectly, with DNA. These observations, along with the results discussed below, suggest that PsqA plays a role in regulating gene expression in oogenesis.

Previous work (Siegel et al., 1993) has indicated that ovaries from *psq* mutant females have reduced levels of *vasa* mRNA and protein, implicating *psq* in the regulation of *vasa* expression. Expression of *vasa* is initially observed in the germarium; we find that PsqA is both abundant in the early germarium and required very early in oogenesis, correlating with the presence of Vasa at this time.

High levels of PsqA are observed in the posterior polar cells at stages 5-6, a time when these cells are purported to be involved in receiving and responding to a signal from the posteriorly localized Grk (reviewed by Anderson, 1995). It is possible that PsqA is involved in regulating gene expression as part of this complex signaling process.

The striking and specific pattern of PsqA localization in the oocyte nucleus at stages 6-10 (seen as two small bright spots of fluorescence in Fig. 9C) suggests that PsqA may bind, either directly or indirectly, to a specific site on the DNA in the oocyte nucleus. It is intriguing to consider the possibility that *psq* may be directly affecting the expression of a particular locus in the oocyte nucleus.

Finally, *psq* could also play a more general role in regulating gene expression by affecting chromatin structure. Our results indicate that PsqA protein is associated with the nurse cell chromatin; Siegel et al. (1993) found that *psq* mutants failed to undergo the normal decondensation of nurse cell DNA at stage 5. Interestingly, two other BTB-containing proteins, E(var)3-93D and GAGA, have been implicated in the modulation of chromatin structure, as well (Dorn et al., 1993; Croston et al., 1991). It is possible that the BTB-containing fusion protein produced in the *psq PZ* mutants may disrupt the protein-protein interactions of these or other BTB-containing modulators of chromatin structure, leading to some aspects of the mutant phenotypes we observe in these lines.

In conclusion, we have shown that *psq* encodes a nuclear protein which may be involved in the control of gene expression in oogenesis. The structure of PsqA and the nature of the *psq PZ* fusion protein suggest that these proteins interact directly with other proteins. These interactions could be involved in regulating the expression of *vasa*, or other genes required to establish polarity. Future experiments designed to identify the factors with which PsqA interacts should provide valuable insight into *psq*'s function.

We thank Haifan Lin, Vivian Siegel and Trudi Schüpbach for providing mutant strains and DNA. We acknowledge the expert technical assistance of Doreen Gillespie in whole-mount in situ hybridization and immunocytochemistry and express our gratitude to Philippa Webster and Alison Volpe for critical reading of the manuscript. We are especially indebted to members of the Berg lab for help in preparation of this manuscript and for insightful discussions: Doreen Gillespie, Philippa Webster, Kim Rittenhouse, Alison Volpe, Jon Schnorr, and Rebecca Botham, and to Michael Gelb and Mark Kot. This work was supported by the March of Dimes Basil O'Connor Research Program and NIH (GM45248).

The nucleotide sequence data reported in this paper are available from the GenBank Nucleotide Sequence Databases under the accession numbers U48358 (*psq-1*) and U48402 (*psq-2*).

REFERENCES

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

Anderson, K. (1995). One signal, two body axes. *Science* **269,** 489-490.

Bardwell, V. and Treisman, R. (1994). The POZ domain: A conserved protein-protein interaction motif. *Genes Dev.* **8**, 1664-1677.

Chasan, R. and Anderson, K. (1993). Maternal control of dorsal-ventral polarity and pattern in the embryo. In *The Development of* Drosophila melanogaster. (ed. M. Bate and A. Martinez-Arias), pp. 387-424. Cold Spring Harbor: Cold Spring Harbor Laboratory Press.

Chen, W., Zollman, S., Couderc, J.-L. and Laski, F. (1995). The BTB Domain of *bric a brac* mediates dimerization in vitro. *Mol. Cell. Biol.* **15**, 3424-3429.

Chen, Z., Brand, N., Chen, A., Chen, S., Tong, J., Wang, Z., Waxman, S. and Zelent, A. (1993). Fusion between a novel Krüppel-like zinc finger gene and the retinoic acid receptor-alpha locus due to a variant t(11;17) translocation associated with acute promyelocytic leukaemia. *EMBO J.* **12**, 1161-1167.

Cooley, L., Thompson, D. and Spradling, A. (1990). Constructing deletions with defined endpoints in Drosophila. *Proc. Natl. Acad. Sci. USA* **87**, 3170-3173.

Croston, G. E., Kerrigan, L. A., Lira, L. M., Marshak, D. R. and Kadonaga, J. T. (1991). Sequence-specific antirepression of histone H1-mediated inhibition of basal RNA Polymerase II transcription. *Science* **251,** 643-649.

Dorn, R., Krauss, V., Reuter, G. and Saumweber, H. (1993). The enhancer of position effect variegation of Drosophila, E(var)3-93D, codes for a chromatin protein containing a conserved domain common to several transcriptional regulators. *Proc. Natl. Acad. Sci. USA* **90**, 11376-11380.

Dutton, F. and Chovnick, A. (1991). The *l(3)S12* locus of *Drosophila melanogaster*: heterochromatic position effects and stage-specific misexpression of the gene in *P* element transposons. *Genetics* **128**, 103-118.

Gillespie, D. and Berg, C. (1995). *homeless* is required for RNA localization in *Drosophila* oogenesis and encodes a new member of the DE-H family of RNA-dependent ATPases. *Genes Dev*. **9,** 2495-2508.

Godt, D., Couderc, J.-L., Cramton, S. and Laski, F. (1993). Pattern formation in the limbs of *Drosophila*: *bric a brac* is expressed in both a gradient and a wave-like pattern and is required for specification and proper segmentation of the tarsus. *Development* **119**, 799-812.

González-Reyes, A., Elliott, H. and St. Johnston, D. (1995). Polarization of both major body axes in *Drosophila* by *gurken-torpedo* signalling. *Nature* **375,** 654-658.

Haller, J., Cote, S., Bronner, G. and Jäckle, H. (1987). Dorsal and neural expression of a tyrosine kinase-related *Drosophila* gene during embryonic development. *Genes Dev.* **1**, 862-867.

Horowitz, H. and Berg, C. (1995). Aberrant splicing and transcription termination caused by *P* element insertion into the intron of a *Drosophila* gene. *Genetics* **139**, 327-335.

Janknecht, R., Sander, C. and Pongs, O. (1991). $(HX)_n$ repeats: a pH-

controlled protein-protein interaction motif of eukaryotic transcription factors? *Fed. Eur. Bioch. Soc.* **295**, 1-2.

Karpen, G. and Spradling, A. (1992). Analysis of subtelomeric heterochromatin in the *Drosophila* minichromosome Dp1187 by single *P* element insertional mutagenesis. *Genetics* **132**, 737-753.

King, R. (1970). Ovarian development in *Drosophila melanogaster*. New York: Academic Press.

Lane, M. and Kalderon, D. (1994). RNA localization along the anteroposterior axis of the *Drosophila* oocyte requires PKA-mediated signal transduction to direct normal microtubule organization. *Genes Dev*. **8,** 2986-2995.

Lin, H. and Spradling, A. (1993). Germline stem cell division and egg chamber development in transplanted *Drosophila* germaria. *Dev. Biol.* **159,** 140-152.

Mahowald, A. and Kambysellis, M. (1980). Oogenesis. In *The Genetics and Biology of* Drosophila. vol. 2d (ed. M. Ashburner and T. Wright). pp. 141-224. New York: Academic Press.

Mount, S., Burks, C., Hertz, G., Stormo, G., White, O. and Fields, C. (1992). Splicing signals in *Drosophila*: intron size, information content, and consensus sequences. *Nucl. Acids Res.* **20,** 4255-4262.

Roth, S., Neuman-Silberberg, F., Barcelo, G. and Schüpbach, T. (1995). *cornichon* and the EGF receptor signaling process are necessary for both anterior-posterior and dorsal-ventral pattern formation in *Drosophila*. *Cell* **81,** 967-978.

Ruohola, H., Bermer, K., Baker, D., Swedlow, F., Jan, L. and Jan, Y. (1991). Role of neurogenic genes in establishment of follicle cell fate and oocyte polarity during oogenesis in *Drosophila*. *Cell* **66,** 433-449.

Schüpbach, T. and Roth, S. (1994). Dorsoventral patterning in *Drosophila* oogenesis. *Curr. Opin. Genet. Dev.* **4,** 502-507.

Schüpbach, T. and Wieschaus, E. (1991). Female sterile mutations on the second chromosome of *Drosophila melanogaster*. II. Mutations blocking oogenesis or altering egg morphology. *Genetics* **129**, 1119-1136.

Siegel, V., Jongens, T., Jan, L. and Jan Y. (1993). *pipsqueak*, an early acting member of the posterior group of genes affects *vasa* level and germ cell-somatic cell interaction in the developing egg chamber. *Development* **119**, 1187-1202.

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

Spradling, A. (1993). Developmental genetics of oogenesis. In *The Development of* Drosophila melanogaster. (ed. M. Bate and A. Martinez-Arias), pp. 1-69. Cold Spring Harbor: Cold Spring Harbor Laboratory Press.

St. Johnston, D. (1993). Pole plasm and the posterior group genes. In *The Development of* Drosophila melanogaster. (ed. M. Bate and A. Martinez-Arias), pp. 325-364. Cold Spring Harbor: Cold Spring Harbor Laboratory Press.

Stroumbakis, N., Li, Z. and Tolias, P. (1994). RNA- and single-stranded DNA-binding (SSB) proteins expressed during *Drosophila melanogaster* oogenesis: a homolog of bacterial and eukaryotic mitochondrial SSBs. *Gene* **143,** 171-177.

Theurkauf, W., Smiley, S., Wong, M. and Alberts, B. (1992). Reorganization of the cytoskeleton during *Drosophila* oogenesis: implications for axis specification and intercellular transport. *Development* **115,** 923-936.

Thummel, C. and Pirrotta, V. (1991). New pCaSpeR *P* element vectors. *Drosophila Information News* **2.**

Verheyen, E. and Cooley, L. (1994). Looking at oogenesis. In *Methods in Cell Biology.* vol. 44 (ed. L. Goldstein and E. Fyrberg), pp. 545-561. San Diego: Academic Press.

Weber, U., Siegel, V., and Mlodzik, M. (1995). *pipsqueak* encodes a novel nuclear protein required downstream of *seven-up* for the development of photoreceptors R3 and R4. *The EMBO J.* **14,** 6247-6257.

Wharton, K., Yedvobnick, B., Finnerty, V. and Artavanis-Tsakonas, S. (1985). opa: a novel family of transcribed repeats shared by the *Notch* locus and other developmentally regulated loci in *Drosophila melanogaster*. *Cell* **40,** 515-562.

Wieschaus, E. and Nüsslein-Volhard, C. (1986). Looking at embryos. In *Drosophila a Practical Approach* (ed. D. B. Roberts), pp. 199-227. Oxford: Information Printing Ltd.

Williams, S., Greenwood, J. and Jones, C. (1992). Molecular analysis of the *lac* operon encoding the binding-protein-dependent lactose transport system and β-galactosidase in *Agrobacterium radiobacter*. *Molecular Microbiology* **6**, 1755-1768.

Zollman, S., Godt, D., Privé, G., Couderc, J.-L. and Laski, F. (1994). The BTB domain, found primarily in zinc finger proteins, defines an evolutionarily conserved family that includes several developmentally regulated genes in *Drosophila*. *Proc. Natl. Acad. Sci USA* **91,** 10717-10721.

(Accepted 15 March1996)