The Drosophila Cystoblast Differentiation Factor, *benign gonial cell neoplasm*, Is Related to DExH-box Proteins and Interacts Genetically With *bag-of-marbles*

B. Ohlstein,*,[1] C. A. Lavoie,*,[1] O. Vef,[†] E. Gateff[†] and D. M. McKearin*

**Department of Molecular Biology, University of Texas Southwestern Medical Center, Dallas, Texas and [†]Department of Genetics, Johannes Gutenberg University, D55099 Mainz, Germany*

Manuscript received January 14, 2000
Accepted for publication April 21, 2000

ABSTRACT

Selection of asymmetric cell fates can involve both intrinsic and extrinsic factors. Previously we have identified the *bag-of-marbles* (*bam*) gene as an intrinsic factor for cystoblast fate in Drosophila germline cells and shown that it requires active product from the *benign gonial cell neoplasm* (*bgcn*) gene. Here we present the cloning and characterization of *bgcn*. The predicted Bgcn protein is related to the DExH-box family of RNA-dependent helicases but lacks critical residues for ATPase and helicase functions. Expression of the *bgcn* gene is extremely limited in ovaries but, significantly, *bgcn* mRNA is expressed in a very limited number of germline cells, including the stem cells. Also, mutations in *bgcn* dominantly enhance a *bam* mutant phenotype, further corroborating the interdependence of these two genes' functions. On the basis of known functions of DExH-box proteins, we propose that Bgcn and Bam may be involved in regulating translational events that are necessary for activation of the cystoblast differentiation program.

ASYMMETRIC stem cell divisions produce a new daughter stem cell and a second daughter that will undergo specialized differentiation (Morrison *et al.* 1997). The germline stem cells (GSCs) of Drosophila are an excellent model system for the study of stem cell biology (de Cuevas *et al.* 1997; Lin 1998). When oogenesis begins in Drosophila third instar larvae, GSC division produces another GSC and a cystoblast that will divide precisely four times with incomplete cytokinesis to form the syncytial germline cyst (Brown and King 1964; Wieschaus and Szabad 1979; Lin and Spradling 1993). Topics under active investigation include identifying factors that are essential to establish/maintain stem cells and determining the changes that take place in differentiating non-stem-cell daughters. Studies from several laboratories have shown that both extrinsic and intrinsic factors are necessary for GSC and cystoblast formation (McKearin 1997; Xie and Spradling 1998; King and Lin 1999).

Establishing and maintaining oogenic GSCs requires the transforming growth factor β (TGFβ)-like signal Decapentaplegic (Dpp) since mutations in Dpp receptors or signal transducing Smad transcription factors cause stem cell loss (Xie and Spradling 1998). Likewise, inactivating mutations of the *piwi* gene eliminate GSCs (Cox *et al.* 1998). Thus, *piwi*-dependent signaling and Dpp signaling are implicated in maintaining GSCs. In the absence of proper signals, GSCs apparently divide as cystoblasts, thus ablating a self-renewing population of stem cells for germline replenishment.

Intrinsic GSC maintenance factors include Pumilio (Pum) and perhaps Nanos (Nos; Forbes and Lehmann 1998) that were initially studied as key regulators of embryonic polarity (St Johnston and Nüsslein-Volhard 1992). Biochemical and molecular studies have shown that Pum and Nos are translational repressors of transcripts that carry a sequence-specific element known as the NRE (Curtis *et al.* 1997; Zamore *et al.* 1997; Wharton *et al.* 1998). Subsequent work, using strong or complete loss-of-function *nos* and *pum* alleles, established that proper germ cell cyst development required both genes (Wang *et al.* 1994; Lin and Spradling 1997; Forbes and Lehmann 1998). For example, GSCs carrying inactivating *pum* mutations divided symmetrically and produced only apparent cystoblasts; *nos* inactivation produced a range of phenotypes including the failure to establish GSCs and defective cyst formation. If Nos and Pum functions are the same in embryos and GSCs, we can infer that establishing and maintaining GSC identity requires translational repression of specific RNA targets. Presently, the germ cell targets of such regulation are unknown.

Studies focused on how asymmetric division produces the cystoblast have identified two intrinsic factors, *bag-of-marbles* (*bam*; McKearin and Spradling 1990) and *benign gonial cell neoplasm* (*bgcn*; Gateff 1982; Lavoie *et al.* 1999). Mutations in either gene produce tumorous egg chambers that are caused by symmetric GSC divisions that produce only more GSCs (McKearin and Spradling 1990; Gateff 1994; Lavoie *et al.* 1999). Misexpression of Bam is sufficient to eliminate GSCs, appar-

Corresponding author: Dennis McKearin, Department of Molecular Biology, U. T. Southwestern Medical Center, 5323 Harry Hines Blvd., Dallas, TX 75390-9148. E-mail: mckearin@utsw.swmed.edu

[1]These authors contributed equally to this article.

ently by converting them into cystoblasts, which is reminiscent of the phenotypes produced by abrogation of Dpp or *piwi*-dependent signaling (Ohlstein and McKearin 1997). Bam is a novel protein and is a component of the germ-cell-specific fusome organelle where it is required for proper formation of the endoplasmic reticulum (ER)-like fusome cisternae (McKearin and Ohlstein 1995; León and McKearin 1999). However, the low abundance of the Bam protein and its novelty has frustrated efforts to discover its biochemical function.

Previously, we have shown that Bgcn is an essential cystoblast differentiation factor that is required for Bam function and can regulate Bam localization (Lavoie *et al.* 1999). Mutations that inactivated the *bgcn* gene blocked the ability of misexpressed Bam to eliminate GSCs, implying that Bgcn was essential for Bam to exert its cystoblast-converting effect on wild-type GSCs. We also noted that Bam fusome localization was blocked by a strong *bgcn* allele. On the basis of these findings, we concluded that Bam and Bgcn action was interdependent (Lavoie *et al.* 1999). Furthermore, we postulated that $bgcn^{+}$ may act permissively in wild-type GSCs since Bam misexpression converted GSCs into apparent cystoblasts (Ohlstein and McKearin 1997). We have therefore cloned the *bgcn* gene to extend our studies of factors controlling cystoblast differentiation. Bgcn is a large protein related to the DExH-box ATP-dependent RNA helicase proteins (Aubourg *et al.* 1999; de la Cruz *et al.* 1999) although Bgcn lacks signature motifs essential for ATP binding and unwinding activity (Gorbalenya and Koonin 1993; de la Cruz *et al.* 1999). We speculate instead that Bgcn may share RNA-binding activity with proteins of the DExH family. Consistent with its genetic role as a highly specific regulator of cystoblast formation, *bgcn* transcripts are restricted to a small number of cells including GSCs. The activity of *bgcn* as a dominant enhancer of *bam* phenotype provides additional evidence that Bam and Bgcn function interdependently.

MATERIALS AND METHODS

Drosophila stocks and culture: Flies were maintained on standard molasses agar media as described in Ashburner (1989). Stocks carrying chromosomal deficiencies were obtained from the Drosophila Stock Centers at Bloomington and Umea.

Chromosomes carrying deficiencies of the 60A region were used to map the position of the *bgcn* gene. The distal breakpoint for *Df(2R)OV1* is in chromosomal locus 60A1 while the distal breakpoint of *Df(2R)b23* extends an additional 15 kb (Wharton *et al.* 1999). Appropriate matings were established to show that *Df(2R)OV1* complements *bgcn* while *Df(2R)b23* fails to complement.

Sequencing the *bam*[BW] allele and candidate genes from *bgcn* alleles: DNA corresponding to candidate genes was recovered from wild-type genomes and mutant *bgcn* alleles by PCR and the products were sequenced at the Department of Molecular Biology Sequencing Facility at U.T. Southwestern. The products from at least four parallel PCR reactions were pooled for sequencing to minimize the potential for mistaking PCR errors for *in vivo* mutations. When these experiments revealed mutations in relevant DNA fragments, the appropriate fragment was recovered again from multiple, parallel PCR reactions and sequenced to confirm the mutation.

Recovering *bgcn* candidates from cDNA and genomic DNA libraries: Genomic clones including the *bgcn* locus were obtained as cosmids from the European Genome Consortium. *wibg* cDNA clones were recovered from an ovarian cDNA library (Stroumbakis *et al.* 1994) and *bgcn* cDNA clones from a testis cDNA library (gift from T. Hazelrigg) by standard hybridization procedures. The longest *bgcn* cDNA clone started 19 nucleotides from a predicted initiating Met codon and extended 3886 nucleotides, stopping adjacent to a consensus *poly-A* addition signal. The sequence was extended at the 5′-end by RT-PCR to a position that corresponded to −92 nucleotides from the putative initiating Met. A termination codon in frame with the predicted *bgcn* open reading frame (ORF) lies at position −78; the next Met in frame with the longest ORF was selected as the most likely translational start codon and designated position +1.

***bgcn* alleles:** $bgcn^{QS2}$ and $bgcn^{QW34}$ were obtained from Drs. A. Mahowald and R. Steward; originally they were recovered from EMS screens by Dr. T. Schüpbach (Schüpbach and Wieschaus 1991). Both of these alleles carried the same transition mutation that produced a termination codon. For the purposes of correlating the *bgcn* gene with a particular segment of genomic DNA, we considered these alleles as nonindependent events.

The P[*lacW*] transposon (Bier *et al.* 1989) was mobilized and 19,000 second chromosome insertions were tested for new female sterile or lethal insertions that were not complemented by *Df(2R)b23*. One of these lines failed to complement the $bgcn^{1}$ allele and was designated $bgcn^{oe1}$.

The $bgcn^{1}$ allele was recovered from an EMS screen of *w*; *b* flies. Alleles $bgcn^{z2\text{-}1185}$, $bgcn^{z2\text{-}1748}$, and $bgcn^{z2\text{-}3112}$ were recovered from a mutagenesis screen of *cn bw* flies for male sterile mutations (B. Wakimoto, D. Lindsley, E. Koundakjian and C. Zuker, personal communication).

Germline transformation: Germline transformation was carried out essentially as described by Rubin and Spradling (1982). Three different *wibg* transgenes were constructed and used for germline transformation rescue experiments. Two transgenes utilized a full-length cDNA clone recovered from an ovarian cDNA library (Stroumbakis *et al.* 1994) fused to either a heat-shock 70-gene promoter (pCaSpeR; Thummel and Pirrotta 1991) or an *otu* gene promoter (pCOG; Robinson and Cooley 1997). The third *wibg* transgene was constructed by inserting a 4.2-kb *Eco*RI fragment of genomic DNA, recovered from cosmid 64H6 (Siden-Kiamos *et al.* 1990), into pCaSpeR. The *bgcn* transgene was constructed in cloning vector pCaSpeR-hs (Pirrotta 1988) using the full-length cDNA described above. Expression of the transgene was induced by 1 hr heat shocks at 37° repeated twice/day for 2 days. We tested the rescuing activity of the *bgcn* transgene by inducing expression in newly eclosed $P[w^{+}; hs\text{-}bgcn]/+; bgcn^{1}/bgcn^{1}$ animals and examining their ovaries on days 5–8 posteclosion.

Molecular biology: PCR, Northern blots, Southern blots, cloning, etc., were all carried out essentially as described in Sambrook *et al.* (1989). RNA *in situ* hybridization was performed as described in Christerson and McKearin (1994). Digoxigenin-labeled single-stranded DNA probes were prepared by cyclic reactions using a single antisense primer and double-stranded DNA template using reaction conditions described by the manufacturer (Boehringer Mannheim, Indianapolis).

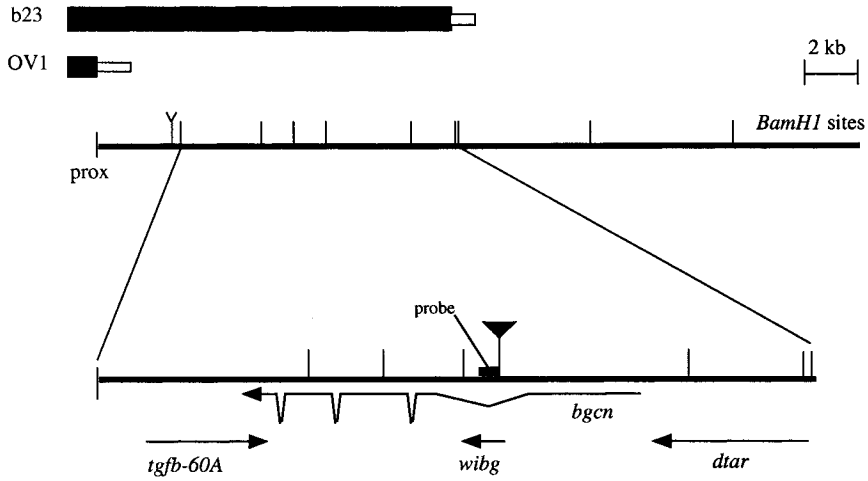

Figure 1.—*bgcn* maps to a 20-kbp region at chromosomal position 60A3. Deficiency chromosomes *Df(2R)b23* and *Df(2R)OV1* were used to map *bgcn* to a region of ~20 kbp. Position of the gene was further refined by a *bgcn* P-allele, which failed to complement female sterility when opposite *Df(2R)b23*. The position of the genomic fragment that was used as a molecular probe to recover cDNAs from the region around the *P* element inserted in $bgcn^{oe1}$ allele is indicated.

RESULTS

Mapping the *bgcn* locus: Previous meiotic mapping placed *bgcn* in the 60A region (Lindsley and Zimm 1992). We mapped the *bgcn* gene onto a 20-kb segment of the 60A1 region using combinations of chromosomal deficiencies placing the *bgcn* locus between positions shown on the map in Figure 1.

The P-allele $bgcn^{oe1}$ was used to refine the position of the *bgcn* gene. Chromosomal *in situ* hybridization showed that $bgcn^{oe1}$ flies (materials and methods) carried a *P*-element insertion at 60A. Homozygous and hemizygous $bgcn^{oe1}$ flies were male and female sterile and had tumorous gonads like canonical *bgcn* mutations. Finally, excision of the *P* element from $bgcn^{oe1}$ restored fertility to both males and females. Thus we concluded that the transposon in $bgcn^{oe1}$ was inserted in or very close to the *bgcn* gene.

A fragment of genomic DNA adjacent to the *P* element in $bgcn^{oe1}$ was recovered by plasmid rescue (Pirrotta 1986). When the sequence of the genomic fragment was aligned with the DNA sequence of the 60A1 region (Berkeley Drosophila Genome Project, unpublished result; http://www.fruitfly.org), we found that the transposon insertion site corresponded to the position on the genomic map (Figure 1) that fell within the 20-kb fragment between the *Df(2R)OV1* and *Df(2R)b23* breakpoints (materials and methods).

***bgcn* is a complex gene:** The genomic fragment recovered by plasmid rescue was used to identify cDNA clones that corresponded to a 1.0-kb transcript, which included the insertion site for the $bgcn^{oe1}$ transposon. However, several results suggested that this transcript did not belong to the *bgcn* gene. The most compelling data were that a 4.2-kb genomic fragment containing the coding sequence for the 1.0-kb transcript could not rescue *bgcn* mutant flies and the coding sequence was wild-type in four EMS-induced *bgcn* alleles. While this manuscript was in preparation, an article describing transcription units in the 60A region appeared (Lukacsovich *et al.* 1999). Comparison of the 1.0-kb transcript to mRNAs identified in that study showed that the 1.0-kb transcript corresponded to their anonymous mRNA, UD3. We have subsequently designated the 1.0-kb transcript *within bgcn* (*wibg*).

We concluded that the authentic *bgcn* gene must be near the $bgcn^{oe1}$ transposon insertion site and considered the possibility that a larger gene encompassed the region. Indeed, we found that probes derived from either side of the *wibg* gene recognized the same ~4100 nucleotide transcript of very low abundance in females and higher abundance in males (Figure 2). Thus we concluded that the *wibg* gene was located within the intron of a larger transcript that was a candidate for the *bgcn* gene.

cDNA clones for the larger transcript were recovered from a testis cDNA library and sequenced. The assembled transcript is ~3900 nucleotides in length and predicts a protein of 1215 amino acids with M_r 1.39×10^5. The intron-exon structure and features of the cDNA and predicted protein sequence are shown in Figures 1 and 3. As we had observed with genomic probes for mRNAs, the candidate *bgcn* transcript is abundant in samples of male *poly*(A^+) mRNA but exceedingly rare in female *poly*(A^+) samples (Figure 2).

We used the genomic DNA sequence (Berkeley Drosophila Genome Project, unpublished results; http://www.fruitfly.org) to design primers for sequencing the candidate *bgcn* ORF from EMS-induced *bgcn* alleles (Lavoie *et al.* 1999) since we expected to find mutations that would alter the predicted ORF. Four of five alleles sequenced contained point mutations that introduced a nonsense codon into the conceptual ORF while the fifth allele contained a G-to-A transition that altered a 3′-acceptor splice site and would produce a frameshift mutation in the mature transcript. The position of these mutations and affected amino acids are shown in Figure 3. All of these alleles show the same phenotype; males and females are viable but sterile with germline tumors in their gonads. We predict that this is the null pheno-

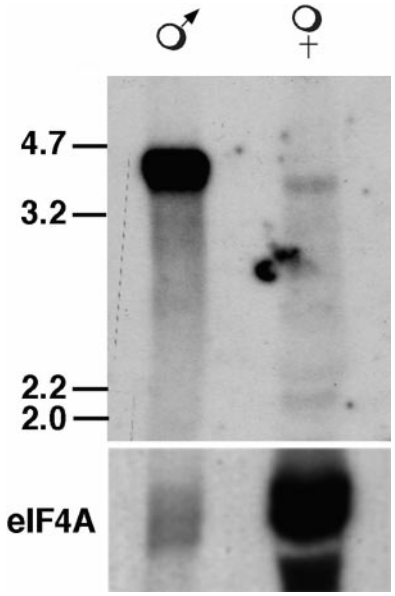

Figure 2.—*bgcn* mRNA is expressed in adult males and females. *Poly-A*$^+$ RNA was transferred from an agarose gel to nylon membrane and hybridized with a probe from the *bgcn* coding region. Expression of the 4-kb female transcript is very low compared with expression of the 4.2-kb male transcript and the female sample was therefore overloaded to improve the detection of the female transcript. Detection of mRNA encoding the ubiquitous translation factor eIF4A was used as a loading control (Cooley *et al.* 1992).

type since the mutation in allele *bgcn*$^{z2\text{-}1748}$ would cause translational termination after only 213 amino acids and almost certainly encodes a nonfunctional protein. The one-to-one correspondence between identical *bgcn* phenotypes and the occurrence of mutations in this ORF confirmed that the 4.1-kb transcript corresponded to *bgcn*.

Although the germ cell phenotypes of all *bgcn* alleles

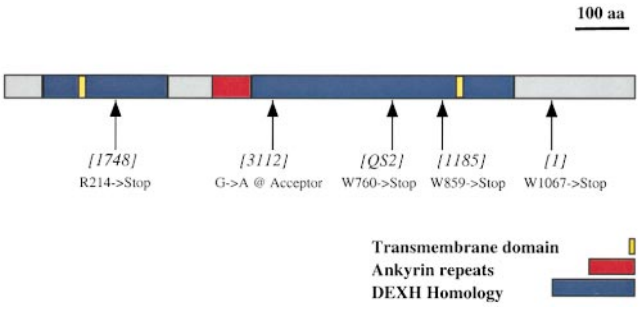

Figure 3.—*bgcn* mutants encode truncated protein products. DNA from *bgcn* mutant flies was amplified, sequenced, and compared to wild-type genomic sequence (materials and methods). Alleles *bgcn*^{QS2}, *bgcn*$^{z2\text{-}1185}$, and *bgcn*^{1} carry G-to-A transitions that introduce premature termination codons. Allele *bgcn*$^{z2\text{-}3112}$ carries a G-to-A mutation at the 3′-donor splice site and causes a predicted frameshift of one nucleotide. Allele *bgcn*$^{z2\text{-}1748}$ carries a C-to-T mutation that introduces a stop codon. The five alleles tested were isolated from three different mutagenic screens.

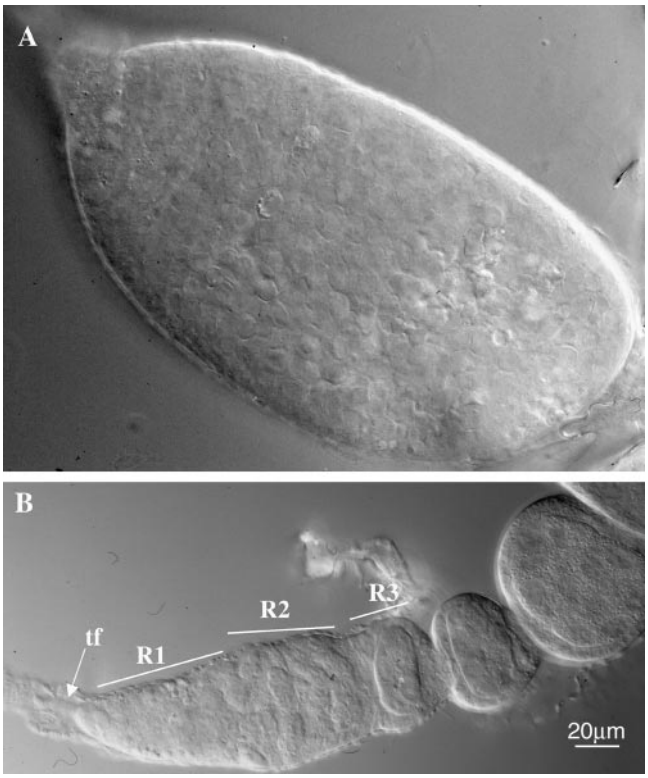

Figure 4.—Expression of *bgcn* cDNA from a heterologous promoter rescues the tumorous *bgcn*^{1} phenotype. (A) Ovaries in homozygous *bgcn* females are filled with tumorous egg chambers. (B) Flies transformed with a transgene containing *bgcn* cDNA under the control of the *hsp70* promoter were treated with heat shock for 2 days and ovaries dissected on the sixth day contained organized germaria (regions 1–3 as R1, R2, and R3) and maturing egg chambers. Bar, 20 μm.

were indistiguishable, immunolocalization experiments with Bam antisera revealed one significant difference. While Bam fusome localization was blocked in the *bgcn*^{1} allele (Lavoie *et al.* 1999), Bam protein associated with fusomes to varying degrees in the other four *bgcn* alleles. The molecular explanation for this allele-specific effect is currently under study.

Nonspecific expression of *bgcn* rescues the mutant phenotype: Using the cDNA clone recovered from the testis library, we constructed a heat-shock inducible *bgcn* transgene for germline rescue experiments. Ovaries of *bgcn* mutant females contain tumorous egg chambers such as those seen in Figure 4A. The ovaries of all P[*w*$^+$; *hs-bgcn*]/+; *bgcn*^{1}/*bgcn*^{1} females expressing transgenic *bgcn*$^+$ had maturing egg chambers and germaria organized into morphologically distinct regions 1, 2, and 3 (Figure 4B) that were formed as cysts assembled (King 1970). Ovaries in older animals also had egg chambers with 15 nurse cells and one yolk-accumulating oocyte (not shown). Females expressing the P[*w*$^+$; *hs-bgcn*] transgene, however, were not fertile. Late-stage egg chambers (stage 12 and later) contained small oocytes, misshapen follicle cell layers, and were not laid. To date,

we have also been unable to rescue spermatogenesis in *bgcn*^{1}/*bgcn*^{1} males with heat-induced expression from the P[w^+; *hs-bgcn*] transgene.

Since rescue of the oogenic tumorous phenotype verified that transgenic Bgcn was active in early germ cells, we could examine the effects of *bgcn* misexpression during early stages of the germ cell lineage. Previous experiments had demonstrated that *bam* and *bgcn* phenotypes were sufficiently similar to suggest a common function (Lavoie *et al.* 1999) and that misexpression of *bam* from a heat-shock promoter caused GSC ablation (Ohlstein and McKearin 1997). We therefore tested if ectopic *bgcn* expression also affected GSC maintenance. Induction of *bgcn* in either *bgcn*/*bgcn* or wild-type genotypes did not produce any of the phenotypes characteristic of GSC ablation. Even 8 days after the initial heat shocks, when P[hs-*bam*] transgenic animals have empty germaria (Ohlstein and McKearin 1997), germaria from P[*hs-bgcn*] females contained a normal array of developing cysts and GSCs could be identified by single, spherical fusomes (not shown) when reacted with anti-Hts antibodies (Zaccai and Lipshitz 1996).

Comparative analysis of the *bgcn* gene: The *bgcn* sequence was compared to sequences in the GenBank database using the BLAST algorithm available at NCBI (Altschul *et al.* 1997; http://www.ncbi.nlm.nih.gov/BLAST). The conceptual Bgcn protein showed a highly significant similarity to many members of the DExH family of RNA helicases (Figure 3); the top 65 scores were either DExH-box or DEAD-box proteins from diverse organisms. Members of the DExH/DEAD family contain seven highly conserved signature motifs (Figure 5, part 1; de la Cruz *et al.* 1999) that occur colinearly over a core region of ~675 amino acids that are more weakly conserved. Bgcn was ~25–30% identical to helicase proteins over the conserved core of 675 amino acids but, notably, showed little conservation in the hallmark helicase motifs.

Figure 5, part 1 shows the alignment of Bgcn and *Arabidopsis thaliana* HVT1 protein, the highest scoring BLAST match (~10^{-42}). *A. thaliana* HVT1 is a member of the DExH helicases based on sequence comparison although its specific function is unknown (Wei *et al.* 1997). The alignment shows that Bgcn does not match any of the helicase signature motifs (Walker motifs and domains 1a-VI in red in Figure 5, part 1). For example, the sequence from positions 184–188 (VILDD; Figure 5, part 2) in Bgcn matches the sequence and position of a Walker B site for Mg^{2+}-ATP binding commonly found in many RNA-dependent helicases (Walker *et al.* 1982; de la Cruz *et al.* 1999). However, Bgcn sequence at amino acids 171–178, which is the expected position of the Walker A site (phosphate-binding "P-loop"), lacks critical features of the ATP-binding P-loop (Saraste *et al.* 1990), making it very unlikely that Bgcn could bind any nucleotide triphosphate. In the DExH-box motif (DEIH in HVT1), Bgcn conserves the His residue at the fourth position but diverges in invariant D and E positions (Gorbalenya and Koonin 1993; de la Cruz 1999). In the case of the other signature helicase motifs, which are somewhat more flexible in sequence conservation, Bgcn does not have any previously recognized motif variant. Taken together, these divergences make it unlikely that Bgcn has either ATP-binding/hydrolysis or helicase activities (Gorbalenya and Koonin 1993).

Many residues outside of the canonical helicase family motifs are also conserved in helicases but they have not been associated with biochemical activities (Aubourg *et al.* 1999; de la Cruz *et al.* 1999). Figure 5, part 2 shows alignment of Bgcn fragments with the equivalent fragments from several diverse helicases. This method of comparison illustrated that Bgcn conservation was more extensive than would be concluded from simple pairwise alignments. It was especially striking that the sequences immediately surrounding several canonical motifs showed higher levels of conservation than within the signature motifs (Figure 5, part 2); the significance of conservation in these sequences is not known.

Bgcn alignment with HVT1 also includes sequences that are not typically part of the DExH/DEAD family core. Bgcn and HVT1 share limited sequence similarity in the N-terminal 150 amino acids and C-terminal 300 amino acids but these regions are not similar to any other proteins in the GenBank database. The N terminus of some DexH/DEAD-box proteins contains divergent versions of the double-stranded RNA-binding domain (DSR-BD; Aubourg *et al.* 1999) but the N-terminal sequences of Bgcn and HVT1 do not match the consensus for DSR-BD.

The region between residues 320 and 475 in Bgcn does not align with other helicase proteins except HVT1. The SMART algorithm, which recognizes common protein domains based on likely secondary structure (http://coot.embl-heidelberg.de/SMART; Schultz *et al.* 1998), predicts tandem ankyrin repeats at positions 407 to 436 and 440 to 472 in Bgcn and the first of these aligns with an ankyrin domain in HVT1 (Figure 5). Ankyrin repeats have been recognized in diverse proteins and are probably protein-protein interaction domains (Sedgewick and Smerdon 1999). The ankyrin repeat at amino acids 407–436 (AD1) is most closely related to one of the ankyrin domains found in black widow spider lactroinsectotoxin, a protein containing many tandemly repeated ankyrin domains (Kiyatkin *et al.* 1993). Despite the prediction of the SMART algorithm, AD2 (amino acids 440 to 472) is a divergent ankyrin repeat and does not show significant amino acid conservation with known ankyrin domains. Finally, the SMART algorithm also identified two possible transmembrane domains at positions 228 to 248 and 926 to 946 in the conceptual Bgcn sequence (http://coot.embl-heidelberg.de/SMART; Schultz *et al.* 1998).

***bgcn* mRNA pattern confirms predicted GSC expres-**

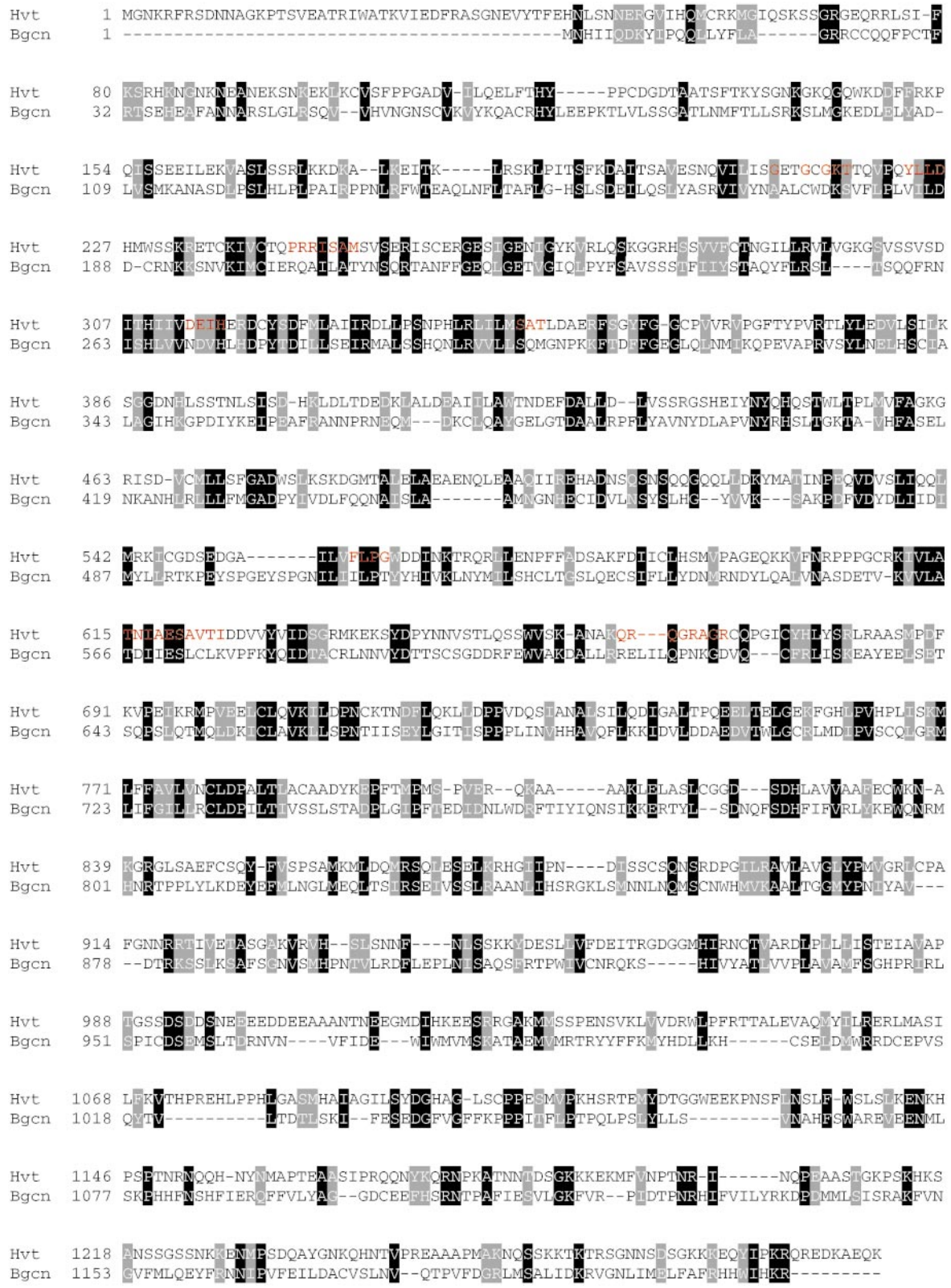

Figure 5.—The predicted Bgcn translation product is related to many members of the DExH family of RNA helicases. (A) Of all members of the DExH family, Bgcn shares the highest level of homology over its entire sequence with the HVT1 protein, a putative RNA helicase from *A. thaliana*. In order, the helicase consensus motifs are the following: Walker A [amino acids (aa) 210–217], Walker B (aa223–227), Ia (aa243–249), DExH-box (aa313–316), domain III (aa345–347), domain IV (aa556–559) domain V (aa615–624), and domain VI (aa672–679), which are shown in red type. (B) Analysis of Bgcn sequence against subdomains of family members from humans (Hhlx), Arabidopsis (Hvt), and Drosophila (Maleless) reveals that Bgcn lacks homology within domains shown to be necessary for RNA helicase activities (asterisks). The predicted Bgcn sequence has two putative ankyrin repeats. One of these is shown aligned with ankyrin domains from mouse Notch4 (Notch4), fly Cactus (Cact), black widow spider latroinsectotoxin (Toxin), and the Hvt helicase from Arabidopsis.

Figure 5.—*Continued.*

sion: We had previously observed that *bgcn* mutant GSCs were not eliminated by Bam misexpression (Lavoie *et al.* 1999). This suggested two conclusions about *bgcn*: (1) Bam requires $bgcn^{+}$ for full activity and (2) *bgcn* gene product could be expressed in wild-type germline stem cells and therefore might not be a limiting factor for cystoblast differentiation. We examined the distribution of *bgcn* expression by RNA *in situ* hybridization against wild-type ovaries (Figure 6). Like *bam* expression (McKearin and Spradling 1990), *bgcn* mRNA was detectable in very few cells (5–8 cells) at the anterior tip of the germarium; the small number of *bgcn*-positive cells explained the low signal on Northern blots. However, careful comparison of *bgcn* and *bam in situ* hybridizations revealed that the patterns were distinct; *bgcn*-positive germ cells included those immediately adjacent to the terminal filament. Thus it appeared that GSCs, in addition to cystoblasts and perhaps very young cystocytes, were positive for *bgcn* transcripts. This pattern can be contrasted with *bam* mRNA that is detectable in cystoblasts but not GSCs. The *bgcn* expression pattern is consistent with the prediction that *bgcn* is expressed in GSCs at levels that permit conversion to cystoblasts when Bam is misexpressed (Ohlstein and McKearin 1997). Also in contrast to *bam*, *bgcn* transcripts were not detectable in the remainder of the ovary.

***bgcn* is a dominant enhancer of *bam* phenotypes:** Previously, we have presented evidence that Bgcn and Bam might be interdependent cystoblast differentiation factors (Lavoie *et al.* 1999). Since *bgcn* RNA *in situ* hybridization results indicated that Bam and Bgcn were coexpressed in cystoblasts, we reasoned that *bam* and *bgcn* mutations might show synergistic effects in genetic tests of cystoblast differentiation. However, *bgcn*/+; *bam*/+ and wild-type flies were equally fertile. We next compared the effects of partial loss of *bgcn* activity in a genetic background with decreased bam^{+} activity. Recently we have characterized a weak *bam* allele, designated $bam^{z3\text{-}2884}$. Drs. B. Wakimoto and D. Lindsley first recognized this allele as a male sterile mutation that induced a tumorous gonial cell phenotype and we determined subsequently that it failed to complement the null $bam^{\Delta 86}$ allele (McKearin and Ohlstein 1995). The $bam^{z3\text{-}2884}/bam^{\Delta 86}$ genotype causes females to be weakly fertile with small ovaries that contain mostly wild-type egg chambers. Sequencing the $bam^{z3\text{-}2884}$ allele revealed a single C ⇒ T mutation that changed L255 ⇒ F. When the *bgcn* gene dosage was reduced by half in transallelic animals ($bgcn/+; bam^{z3\text{-}2884}/bam^{\Delta 86}$), females became sterile and produced only tumorous cysts similar to those found typically in $bam^{\Delta 86}/bam^{\Delta 86}$ flies (Figure 7). This interaction was tested for five *bgcn* alleles derived from four separate mutagenic screens and all behaved identically as dominant enhancers of the "sensitized" *bam* phenotype.

DISCUSSION

***bcgn* gene identification:** Transgenic rescue of the *bgcn* phenotype and DNA sequencing of mutant *bgcn* alleles allowed us to distinguish between the *bgcn* and *wibg* genes. However, while cyst formation and egg chamber

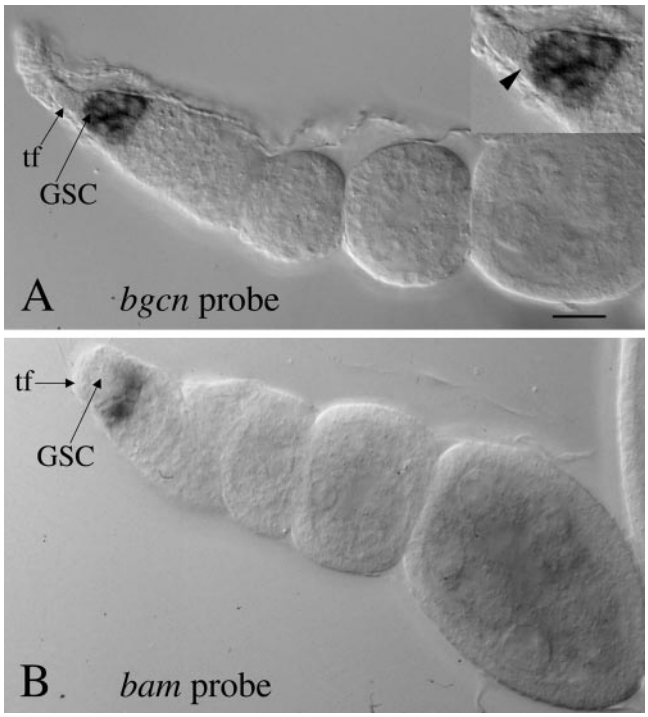

Figure 6.—*bgcn* mRNA is expressed within germline stem cells and cystoblasts at the tip of the germarium. (A) RNA *in situ* hybridization with a *bgcn* probe reveals expression in a few cells at the anterior end of the germarium including GSCs and cystoblasts. Cells immediately under the somatic cells at the base of the terminal filament (tf) and within a few cell diameters of the terminal filament cells are *bgcn* positive. The arrow labeled GSC denotes the presumed position of a germline stem cell. Note that *bgcn* RNA is undetectable in later stage egg chambers. The inset shows a higher magnification of the germarial tip to illustrate these points; the arrowhead indicates the base of the terminal filament. (B) This *in situ* hybridization with a probe for *bam* was performed to show the difference in the pattern of positive cells. The terminal filament (tf) is lying on top of the germarium tip. Note that cells immediately under the terminal filament base are *bam* negative (also see McKearin and Spradling 1990). Bar, 20 μm.

assembly was rescued by an inducible *bgcn* transgene, oogenesis was incomplete in rescued females. Possible explanations include that the P[hs-*bgcn*] transgene does not provide Bgcn product at all appropriate times for full oogenic rescue or misexpression of *bgcn* in somatic ovarian cells may interfere with proper oogenic progress. These same reasons, applied to spermatogenesis, could explain the failure of the P[hs-*bgcn*] transgene to rescue fertility in *bgcn* mutant males.

Ovarian *in situ* hybridization with *bgcn* revealed that the major site of *bgcn* accumulation was in a small number of cells at the most anterior tip of the germarium. Although the signal-to-noise ratio was reliable in these assays, we cannot exclude that *bgcn* is expressed at a low level elsewhere. The most significant aspect of *bgcn* expression was that, unlike *bam* mRNA, GSCs were positive for *bgcn* transcripts. At first, GSC expression appeared counterintuitive since *bgcn* is required for cystoblast, but not GSC, development (Gateff 1982; Lavoie *et al.* 1999). We had, however, considered it likely that *bgcn* would be expressed within GSCs based on the consequences of Bam misexpression. Since expression of the P[HS-Bam] transgene had ablated wild-type GSCs and apparently converted them to cystoblasts (Ohlstein and McKearin 1997), we expected that Bam accumulation would be limiting in wild-type GSCs while other cystoblast factors would be expressed in GSCs (Lavoie *et al.* 1999). From this perspective GSCs, which have no detectable *bam* mRNA (McKearin and Spradling 1990), are primed to become cystoblasts and lack only a higher expression level of the *bam* gene.

Northern blot analysis revealed that *bgcn* was expressed at very low levels in ovaries but at significantly higher abundance in testes. In addition, the female transcript was reproducibly smaller than the male. The very low abundance of *bgcn* mRNA in female *poly-A*$^+$ samples can be explained by the very restricted pattern of expression in ovaries although we do not yet know what factors account for the sexually dimorphic expression levels. Preliminary *in situ* hybridization suggested that *bgcn* was expressed throughout the testis but a more clear understanding of Bgcn expression will emerge when antibodies are available.

Bgcn is distantly related to DExH-box proteins: The predicted Bgcn sequence revealed two specific similarities; one to the superfamily of ATP-dependent RNA helicases and a second to ankyrin domains. RNA helicases are a very large family of proteins that are primarily involved in either pre-mRNA processing or in translational control (Aubourg *et al.* 1999; de la Cruz *et al.* 1999). Bgcn showed position alignment and sequence conservation with dozens of helicase family members. RNA helicases have been recognized and catalogued on the basis of seven conserved domains (Gorbalenya and Koonin 1993; de la Cruz *et al.* 1999); four motifs have been implicated in ATP binding and hydrolysis while two others have been implicated in nucleic acid unwinding. The final helicase motif, GRAGR, was implicated in RNA interaction in eIF4A (Pause *et al.* 1993) but was required for ATP hydrolysis and RNA unwinding, but not RNA interaction, in the NPH-II protein (Gross and Schuman 1996). Alignment of helicase family members illustrates that sequence conservation extends well beyond the short canonical motifs but no biochemical functions have been associated with conserved sequences outside of domains I–VII. A reasonable hypothesis is that some of the sequences conserved in DExH proteins are involved with RNA interactions especially since many DExH-box family members lack recognizable RNA-binding motifs (Nagai and Mattaj 1994; Zhang and Grosse 1997).

The degree of conservation predicts that Bgcn and helicases share some biochemical activities. Since Bgcn does not have the motifs required for ATP binding and

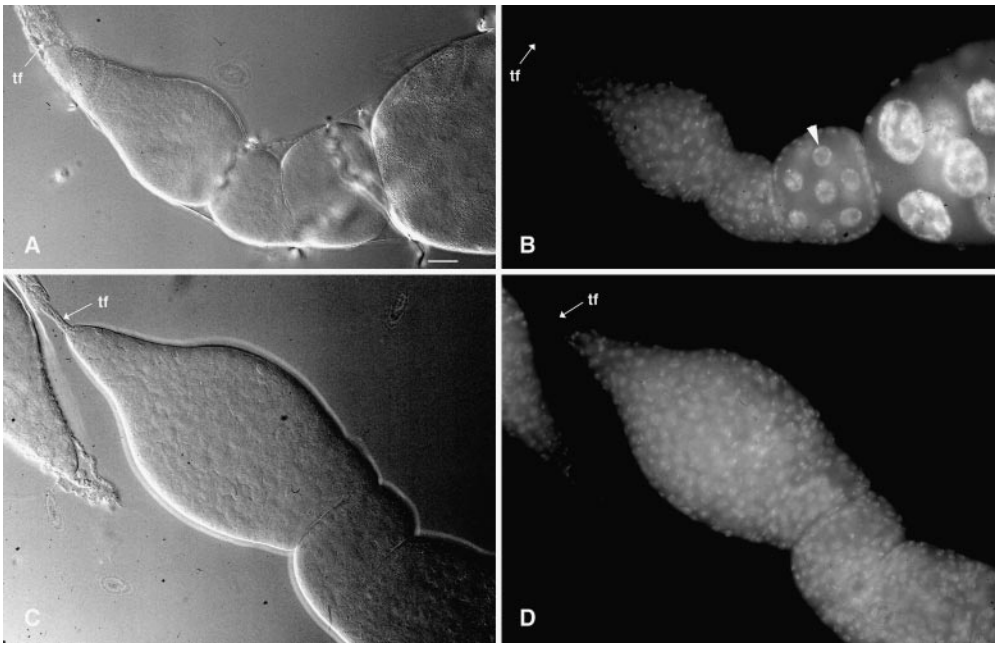

Figure 7.—*bgcn* interacts genetically with the cystoblast differentiation factor *bam*. (A) Ovaries dissected from $bam^{z3\text{-}2884}/bam^{\Delta 86}$ adults have wild-type egg chambers although many germaria lack morphologically distinct cysts reflecting the *bam* "sensitized" genetic background. Nevertheless, ovarioles contain maturing egg chambers, some of which are eventually fertilized and produce progeny. (B) Staining the ovariole shown in A with 4′,6-diamidino-2-phenylindole (DAPI) revealed egg chambers containing endoreduplicating nurse cells (arrowhead) and a monolayer of follicle cells surrounding each egg chamber. The germarium appears somewhat swollen and does not present the same degree of cyst organization seen in wild-type germaria. (C) Ovaries dissected from $bgcn^{1}/+$; $bam^{z3\text{-}2884}/bam^{\Delta 86}$ adults display germ cell hyperplasia that is indistinguishable from *bgcn* or *bam* homozygous mutant ovaries. DAPI staining these germaria (D) revealed that none of the $bgcn^{1}/+$; $bam^{z3\text{-}2884}/bam^{\Delta 86}$ germ cells contained the polyploid DNA complement that would be diagnostic of nurse cell formation. Bar, 20 μm.

helicase activity, we postulate that Bgcn shares the RNA interaction activity that characterizes the DExH proteins. Bgcn might represent an ancestral DExH protein that predates the acquisition of domains involved in ATP hydrolysis and RNA helicase catalysis. An alternative is that Bgcn represents a more modern branch that lost those domains involved in ATPase and helicase activity. Irrespective of the protein evolutionary implications of Bgcn and helicase similarities, we predict that Bgcn regulates post-transcriptional events.

The highest BLAST score of alignment was between Bgcn and the HVT1 protein of *A. thaliana*. HVT1 is clearly a member of the DEIH-box subfamily of helicases but differs from most other family members by having two ankyrin domains; Bgcn shares this feature with HVT1, having two ankyrin domains between amino acid positions 400 and 508. The Arabidopsis genome database contains at least one other helicase (NIH protein) that is similar to HVT1 protein and contains ankyrin domains (Isono *et al.* 1999). HVT1 transcripts are restricted to the plant's vascular and tapetum tissue (nutritive tissue that surrounds the microspores) and the gene is not essential (Wei *et al.* 1997) while NIH expression is widespread and the protein may be nuclear (Isono *et al.* 1999). The function of both proteins is unknown. They are, however, the only two DExH proteins in the GenBank database that also contain ankyrin repeats. Ankyrin repeats are commonly involved in protein interaction interfaces (Sedgewick and Smerdon 1999) and may play such a role in Bgcn interactions with partners.

***bgcn* and *bam* gene function are closely related and dosage sensitive:** In previous studies we demonstrated that $bgcn^{+}$ was necessary for proper Bam function and proposed that Bam and Bgcn may act together in a complex to accomplish cystoblast differentiation (Lavoie *et al.* 1999). More recently we discovered a weak *bam* allele that allowed limited female fertility and created flies that are sensitive to even small reductions in bam^{+} activity. We observed that these flies became sterile and produced tumorous, "*bam*-like" egg chambers when these "sensitized" *bam* females were made heterozygous for *bgcn*. The simplest interpretation for this observation is that decreasing *bgcn* dosage by half can effectively decrease bam^{+} activity and alter phenotype in the sensitized genetic background. This implies that Bam and Bgcn work together closely in the molecular pathway leading to cystoblast differentiation.

Is Bgcn a translational regulator? As a member of the DExH-box family, Bgcn may be an RNA interacting protein. To explain the genetic and molecular aspects of their expression, we have presented evidence (this article; Lavoie *et al.* 1999) that *bgcn* and *bam* functions are interdependent and suggest that Bam and Bgcn proteins may interact. This hypothesis predicts that Bgcn action would be cytoplasmic and its role as DExH-box protein would more likely be involved with translational control rather than splicing regulation. Studies of the key role that Pum plays in GSC maintenance implicate translational regulation in the transition between GSC and cystoblast fate (Lin and Spradling 1997; Forbes and Lehmann 1998). Perhaps a Bgcn-Bam protein complex acts as a translational regulator of cystoblast-promoting transcripts that would be translationally repressed in GSCs. We note that our previous

data suggest association between Bam and the fusome reticulum (McKearin and Ohlstein 1995; León and McKearin 1999), a structure resembling a germ cell modification of ER. A role for Bam in translational control could indicate that cystoblast activation depends on ER-associated translation.

The authors thank members of the S. Dinardo lab for sharing valuable reagents. We also thank B. Wakimoto and K. Wharton for sharing flies and unpublished information. Many thanks to E. Matunis, S. Wasserman, and L. Cooley for comments over the course of this work. Mary Kuhn provided expert technical assistance with genetic and cytological work for many of the experiments described; Erika Jost provided valuable technical assistance during the isolation of the *bgcn*^{oe1} allele. This work was supported by National Institutes of Health grant GM-45820 to D.M.

LITERATURE CITED

Altschul, S. F., T. L. Madden, A. A. Schäffer, J. Zhang, Z. Zhang *et al.*, 1997 Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25:** 3389–3402.

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

Aubourg, S., M. Kreis and A. Lecharny, 1999 The DEAD box RNA helicase family in *Arabidopsis thaliana*. Nucleic Acids Res. **27:** 628–636.

Bier, E., H. Vaessin, S. Shephard, K. Lee, K. McCall *et al.*, 1989 Searching for pattern and mutation in the *Drosophila* genome with a P-*lacZ* vector. Genes Dev. **3:** 1273–1287.

Brown, E. H., and R. C. King, 1964 Studies on the events resulting in the formation of an egg chamber in *Drosophila melanogaster*. Growth **28:** 41–81.

Christerson, L., and D. M. McKearin, 1994 *orb* is required for anteroposterior and dorsoventral patterning during *Drosophila* oogenesis. Genes Dev. **8:** 614–628.

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

Cox, D. N., A. Chao, J. Baker, L. Chang, D. Qiao *et al.*, 1998 A novel class of evolutionarily conserved genes defined by *piwi* are essential for stem cell renewal. Genes Dev. **12:** 3715–3727.

Curtis, D., K. Trieber, F. Tao, P. D. Zamore, J. R. Williamson *et al.*, 1997 A CCHC-metal binding domain in Nanos is essential for translational regulation. EMBO J. **16:** 834–843.

de Cuevas, M., M. A. Lilly and A. C. Spradling, 1997 Germline cyst formation in *Drosophila*. Annu. Rev. Genet. **31:** 405–428.

de la Cruz, J., D. Kressler and P. Linder, 1999 Unwinding RNA in Saccharomyces cerevisiae: DEAD-box proteins and related families. Trends Biochem. Sci. **24:** 192–198.

Forbes, A., and R. Lehmann, 1998 Nanos and Pumilio have critical roles in development and function of the *Drosophila* germline stem cells. Development **125:** 679–690.

Gateff, E., 1982 Gonial cell neoplasm of genetic origin affecting both sexes of *Drosophila melanogaster*, pp. 621–632 in *Embryonic Development, Part B: Cellular Aspects*, edited by R. Weber and M. M. Burger. Alan Liss, Inc., New York.

Gateff, E., 1994 Tumor suppressor and overgrowth suppressor genes of *Drosophila melanogaster:* developmental aspects. Int. J. Dev. Biol. **38:** 565–590.

Gorbalenya, A. E., and E. V. Koonin, 1993 Helicases: amino acid sequence comparisons and structure-function relationships. Curr. Opin. Struct. Biol. **3:** 419–429.

Gross, C. H., and S. Schuman, 1996 The QRxGRxGRxxxG motif of vaccinia virus DExH box RNA helicase NPH-II is required for ATP hydrolysis and RNA unwinding but not for RNA binding. J. Virol. **70:** 1706–1713.

Isono, K., H. Yamamoto, K. Sato and H. Kobayashi, 1999 An Arabidopsis cDNA encoding a DNA-binding protein that is highly similar to the DEAH family of RNA / DNA helicase genes. Nucleic Acids Res. **27:** 3728–3735.

King, F. J., and H. Lin, 1999 Somatic signaling mediated by fs(1)Yb is essential for germline stem cell maintenance during *Drosophila* oogenesis. Development **126:** 1833–1844.

King, R. C., 1970 *Ovarian Development in Drosophila melanogaster*. Academic Press, New York.

Kiyatkin, N., I. Dulubova and E. Grishin, 1993 Cloning and structural analysis of alpha-latroinsectotoxin cDNA. Abundance of ankyrin-like repeats. Eur. J. Biochem **213:** 121–127.

Lavoie, C., B. Ohlstein and D. M. McKearin, 1999 Localization and function of Bam protein requires the activity of the *benign gonial cell neoplasm* gene. Dev. Biol. **212:** 405–413.

León, A., and D. M. McKearin, 1999 Identification of TER94: a AAA ATPase protein, as a Bam-dependent component of the *Drosophila* fusome. Mol. Biol. Cell **10:** 3825–3834.

Lin, H., 1998 The self-renewing mechanism of stem cells in the germline. Curr. Opin. Cell Biol. **10:** 687–693.

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

Lin, H., and A. C. Spradling, 1997 A novel group of *pumilio* mutations affects the asymmetric division of germline stem cells in the *Drosophila* ovary. Development **124:** 2463–2476.

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

Lukacsovich, T., Z. Asztalos, N. Juni, W. Awano and D. Yamamoto, 1999 The *Drosophila melanogaster* 60A chromosomal division is extremely dense with functional genes: their sequences, genomic organization and expression. Genomics **57:** 43–56.

McKearin, D. M., 1997 The *Drosophila* fusome, organelle biogenesis and germ cell differentiation: If you build it, ... Bioessays **19:** 147–152.

McKearin, D. M., and B. Ohlstein, 1995 A role for the *Drosophila* Bag-of-marbles protein in the differentiation of cystoblasts from germline stem cells. Development **121:** 2937–2947.

McKearin, D. M., and A. C. Spradling, 1990 *bag-of-marbles*: a *Drosophila* gene required to initiate male and female gametogenesis. Genes Dev. **4:** 2242–2251.

Morrison, S. J., N. M. Shah and D. J. Anderson, 1997 Regulatory mechanisms in stem cell biology. Cell **88:** 287–298.

Nagai, K., and I. W. Mattaj, 1994 *RNA-Protein Interactions*. IRL Press, New York.

Ohlstein, B., and D. M. McKearin, 1997 Ectopic expression of the *Drosophila* Bam protein eliminates germline stem cells. Development **124:** 3651–3662.

Pause, A., N. Methot and N. Sonenberg, 1993 The HRIGRXXR region of the DEAD box RNA helicase eukaryotic translation initiation factor 4A is required for RNA binding and ATP hydrolysis. Mol. Cell. Biol. **13:** 6789–6798.

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

Pirrotta, V., 1988 Vectors for P-mediated transformation in *Drosophila*. Biotechnology **10:** 437–456.

Robinson, D. N., and L. Cooley, 1997 Examination of the function of two kelch proteins generated by stop codon suppression. Development **124:** 1405–1412.

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

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

Saraste, M., P. R. Sibbald and A. Wittinghofer, 1990 The P-loop: a common motif in ATP- and GTP-binding proteins. Trends Biochem. Sci. **15:** 430–434.

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

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

Sedgewick, S. G., and S. J. Smerdon, 1999 The ankyrin repeat: a diversity of interactions on a common structural framework. Trends Biochem. Sci. **24:** 311–316.

Siden-Kiamos, I., R. D. Saunders, L. Spanos, T. Majerus, J.

Treanear *et al.*, 1990 Towards a physical map of the *Drosophila melanogaster* genome: mapping of cosmid clones within defined genomic divisions. Nucleic Acids Res. **18:** 6261–6270.

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

Stroumbakis, N. D., Z. Li and P. P. Tolias, 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.

Walker, J. E., M. Saraste, M. J. Runswick and N. J. Gay, 1982 Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. EMBO J. **1:** 945–951.

Wang, C., L. K. Dickinson and R. Lehmann, 1994 Genetics of *nanos* localization in *Drosophila.* Dev. Dynamics **199:** 103–115.

Wei, W., D. Twell and K. Lindsey, 1997 A novel nucleic acid helicase gene identified by promoter trapping in *Arabidopsis.* Plant J. **11:** 1307–1314.

Wharton, R. P., J. Sonoda, T. Lee, M. Patterson and Y. Murata, 1998 The *pumilio* RNA-binding domain is also a translational repressor. Mol. Cell **1:** 863–872.

Wharton, K. A., J. M. Cook, S. Torres-Schumann, K. de Castro, E. Borod *et al.*, 1999 Genetic analysis of the Bone Morphogenetic Protein-related gene, *gbb*, identifies multiple requirements during *Drosophila* development. Genetics **152:** 629–640.

Wieschaus, E., and J. Szabad, 1979 The development and function of the female germ line in *Drosophila melanogaster*: a cell lineage study. Dev. Biol. **68:** 29–46.

Xie, T., and A. C. Spradling, 1998 *decapentaplegic* is essential for the maintenance and division of germline stem cells in the *Drosophila* ovary. Cell **94:** 251–260.

Zaccai, M., and H. D. Lipshitz, 1996 Differential distributions of two adducin-like protein isoforms in the *Drosophila* ovary and early embryo. Zygote **4:** 159–166.

Zamore, P., A. Williamson and R. Lehmann, 1997 The Pumilio protein binds RNA through a domain that defines a new class of RNA-binding proteins. RNA **3:** 1421–1433.

Zhang, S., and F. Grosse, 1997 Domain structure of human nuclear DNA helicase II (RNAse A). J. Biol. Chem. **272:** 11487–11494.

Communicating editor: T. Schüpbach