A Screen for Modifiers of *decapentaplegic* Mutant Phenotypes Identifies *lilliputian*, the Only Member of the Fragile-X/Burkitt's Lymphoma Family of Transcription Factors in *Drosophila melanogaster*

Maureen A. Su,*,[1] Robert G. Wisotzkey[†] and Stuart J. Newfeld[‡]

**Department of Molecular and Cellular Biology, Harvard University, Cambridge, Massachusetts 02138, [†]Deltagen, Inc., Menlo Park, California 94025 and [‡]Department of Biology and Graduate Program in Molecular and Cellular Biology, Arizona State University, Tempe, Arizona 85287-1501*

Manuscript received July 26, 2000
Accepted for publication November 3, 2000

ABSTRACT

The *decapentaplegic* (*dpp*) gene directs numerous developmental events in *Drosophila melanogaster*. *dpp* encodes a member of the Transforming Growth Factor-β family of secreted signaling molecules. At this time, mechanisms of *dpp* signaling have not yet been fully described. Therefore we conducted a genetic screen for new *dpp* signaling pathway components. The screen exploited a transvection-dependent *dpp* phenotype: heldout wings. The screen generated 30 mutations that appear to disrupt transvection at *dpp*. One of the mutations is a translocation with a recessive lethal breakpoint in cytological region 23C1-2. Genetic analyses identified a number of mutations allelic to this breakpoint. The 23C1-2 complementation group includes several mutations in the newly discovered gene *lilliputian* (*lilli*). *lilli* mutations that disrupt the transvection-dependent *dpp* phenotype are also dominant maternal enhancers of recessive embryonic lethal alleles of *dpp* and *screw*. *lilli* zygotic mutant embryos exhibit a partially ventralized phenotype similar to *dpp* embryonic lethal mutations. Phylogenetic analyses revealed that *lilli* encodes the only Drosophila member of a family of transcription factors that includes the human genes causing Fragile-X mental retardation (FMR2) and Burkitt's Lymphoma (LAF4). Taken together, the genetic and phylogenetic data suggest that *lilli* may be an activator of *dpp* expression in embryonic dorsal-ventral patterning and wing development.

THE *decapentaplegic* (*dpp*) gene influences many developmental events in *Drosophila melanogaster*. These include dorsal-ventral patterning in the embryo, larval midgut morphogenesis, and formation of adult appendages (Gelbart 1989). *dpp* encodes a member of the highly conserved Transforming Growth Factor-β (TGF-β) family of secreted signaling molecules (Padgett *et al.* 1987). To understand how *dpp* directs developmental decisions in target cells, mechanisms of *dpp* activation and signal transduction must be fully described. Genetic screens have been successful in identifying components of Dpp's signal transduction pathway (Raftery *et al.* 1995; Sekelsky *et al.* 1995). These screens exploited recessive embryonic lethal *dpp* alleles to identify mutations that enhance this phenotype. *Mothers against dpp* (*Mad*) and *Medea* (*Med*) were identified in these screens. These genes, members of the Smad family, are also highly conserved across species (Newfeld *et al.* 1999). Smad family members play important roles in mouse development and act as tumor suppressor genes in several human cancers (Riggins *et al.* 1997; Waldrip *et al.* 1998).

Here we report a genetic screen for *dpp* signaling pathway components that exploits transvection effects at the *dpp* locus (Gelbart 1982). Transvection, or pairing-dependent intragenic complementation between two alleles of a gene, is seen at a number of loci (Lewis 1954). As a result of transvection, *trans*-heterozygous individuals of the genotype $dpp^{d\text{-}ho}/dpp^{hr4}$ display wild-type wings. The $dpp^{d\text{-}ho}$ mutation is a small deletion in the 3′ *cis*-regulatory region of *dpp*. $dpp^{d\text{-}ho}$ homozygous flies have wings that are held out laterally from the body axis (Spencer *et al.* 1982). The dpp^{hr4} mutation is a missense mutation in the protein-coding region of *dpp* (Wharton *et al.* 1996). When homozygous, the dpp^{hr4} allele is embryonic lethal. When $dpp^{d\text{-}ho}$ and dpp^{hr4} are paired, the wild-type regulatory region of the dpp^{hr4} allele appears to act *in trans* on the wild-type coding region of the $dpp^{d\text{-}ho}$ allele to generate viable adults with wild-type wings.

During transvection, the respective regions (regulatory and coding) must be in close physical proximity. A chromosomal rearrangement that physically moves a *dpp* allele to another part of the chromosome disrupts transvection (Gelbart 1982). Rather than having wild-type wings, $dpp^{d\text{-}ho}/dpp^{hr4}$ flies with chromosomal rearrangements have heldout wings. Analyses of polytene

Corresponding author: Stuart J. Newfeld, Department of Biology, Arizona State University, Tempe, AZ 85287-1501.
E-mail: newfeld@asu.edu

[1] *Present address:* Department of Pediatrics, University of California, San Francisco, CA 94143.

chromosomes from rearrangement genotypes showed asynapsis at the *dpp* locus. These rearrangements are referred to as normal *dpp* transvection-disruptors (normal DTDs). *Trans*-heterozygous $dpp^{d\text{-}ho}/dpp^{hr4}$ flies will also display a heldout phenotype if they contain a rearrangement with a breakpoint in a gene required for *dpp* function (*e.g.*, *Mad*; Sekelsky *et al.* 1995). This type of rearrangement, one that generates heldout phenotypes in *trans*-heterozygous flies without asynapsis at the *dpp* locus, is referred to as an exceptional DTD (Gelbart 1982).

To determine if a DTD is normal or exceptional, an unknown DTD is paired with a previously characterized normal DTD. If the unknown DTD is a normal DTD, *trans*-heterozygous flies will display wild-type wings. Two normal DTDs (even those with very different rearrangements) have the ability to arrange themselves in such a way that synapsis occurs at the *dpp* locus (Gelbart 1982). If the unknown DTD is an exceptional DTD, *trans*-heterozygous flies will display heldout wings. The presence of a normal DTD cannot suppress a heldout phenotype that is due to a mutation in a gene required for *dpp* function. Mutations that act as exceptional DTDs are therefore candidates for components of the *dpp* signaling pathway.

MATERIALS AND METHODS

Drosophila stocks: $dpp^{d\text{-}ho}$, dpp^{hr4}, dpp^{e87}, dpp^{hr56}, and dpp^{hr92} are described in St. Johnston *et al.* (1990). *DTD11*, *DTD24*, *Df(2L)DTD16xD42*, and *Df(2L)DTD51xD52* are described in Gelbart (1982). Med^{1}, scw^{E1}, and scw^{S12} are described in Raftery *et al.* (1995). *Df(2L)C28*, *Df(2L)C144*, *Df(2L)JS17*, *Df(2L)JS7*, *Df(2L)DTD62xH7*, Mad^{6}, Mad^{11}, and Mad^{12} are described in Sekelsky *et al.* (1995). sax^{1} and tkv^{8} are described in Brummel *et al.* (1994). *l(2)a16*, *l(2)k9*, *l(2)a4*, and *l(2)a6* were identified in a large screen (>5000 chromosomes) for lethal mutations over *Df(2L)JS17* described in Sekelsky (1993). gbb^{1} is described in Wharton *et al.* (1999). *l(2)00632* and *l(2)k05431* allelic to *lilli*, *l(2)01361* allelic to toucan, and *Df(3R)e-N19* are described in FlyBase (1999). $lilli^{s35}$ and $lilli^{xs407}$ are described in Neufeld *et al.* (1998) and Rebay *et al.* (2000).

Exceptional DTD screen: Homozygous $dpp^{d\text{-}ho}$ males were irradiated and crossed to dpp^{hr4}/CyO females. All G1 heldout progeny were isolated. These progeny carry DTDs (*). Single G1 heldout males were mated to females carrying a normal DTD (either *DTD11* or *DTD24*). If the G2 progeny was heldout, then the new DTD was an exceptional DTD. The $dpp^{d\text{-}ho}$ * chromosome was then balanced. Gravid G1 heldout females were placed alone in a vial and allowed to produce progeny. Heldout male progeny must be either $dpp^{d\text{-}ho}$ */$dpp^{d\text{-}ho}$ or $dpp^{d\text{-}ho}$ */dpp^{hr4}. These males were crossed to dpp^{hr4}/CyO females. Heldout progeny from this cross must bear the genotype $dpp^{d\text{-}ho}$ */dpp^{hr4}. These males were treated like G1 heldout males and crossed to females carrying a normal DTD. If the resulting progeny was heldout, then the new DTD was an exceptional DTD. The $dpp^{d\text{-}ho}$ * chromosome was then balanced. Wing angle measurements were performed as described (Gelbart 1982). Polytene chromosome squashes, cuticle preps, maternal enhancement, and stage of lethality tests were performed as described (Sekelsky *et al.* 1995).

Phylogenetic analysis of Lilli: Database searches for proteins similar to Lilli were conducted using the National Institutes of Health website: http://www.ncbi.nlm.nih.gov/BLAST. In addition to GenBank, we conducted extensive BLAST searches of the genome databases for *D. melanogaster* (Berkeley *Drosophila* Genome Project website: http://fruitfly.berkeley.edu) and *Caenorhabditis elegans* (Washington University Genome Sequence Center website: http://genome.wustl.edu/gsc). Proteins identified by these searches that showed strong similarity to Lilli (see the legend to Figure 5 for accession numbers) were aligned with MACAW (Schuler *et al.* 1991). The alignments were refined using CLUSTAL X (Jeanmougin *et al.* 1998) and then adjusted manually. Alignments not presented are available upon request. Protein motifs were identified in the alignments using the Kyoto University GenomeNet website: http://www.motif.genome.ad.jp. Pairwise calculations of amino acid identity and similarity were deduced from the alignments using the Baylor College of Medicine website: http://dot.imgen.bcm.tmc.edu:9331.

Phylogenetic trees were generated from the alignments using MEGA (Kumar *et al.* 1993). First, a Poisson correction distance (Nei and Kumar 2000) was calculated from each pairwise distance to account for multiple substitutions per site. Then the evolutionary divergence (the number of amino acid substitutions per site) between two sequences was calculated from the Poisson correction distance. Trees were then constructed on the basis of the corrected distance matrix using the neighbor-joining method (Saitou and Nei 1987). Trees not presented are available upon request. The degree of confidence for each branchpoint was obtained by the bootstrap method (1000 replications; Felsenstein 1985).

RESULTS

Exceptional DTD screen: A total of 44,000 $dpp^{hr4}/dpp^{d\text{-}ho}$ flies were screened (Figure 1) and 321 DTD mutations were isolated. Of these mutations, 30 were exceptional DTDs (Table 1). All exceptional DTDs were cytologically mapped. If an exceptional DTD chromosome appeared cytologically normal, the DTD mutation was mapped by recombination.

All exceptional DTDs were then tested for genetic interactions (enhancement of recessive embryonic lethality) with several classes of mutations affecting the Dpp signaling pathway. First, we tested for interactions with loss-of-function mutations in the Dpp receptors *saxophone* and *thickveins* (sax^{1} and tkv^{8}; Brummel *et al.* 1994) and in the Dpp signal transducers *Mad* and *Med* (Mad^{6}, Mad^{11}, Mad^{12}, *and* Med^{1}; Newfeld *et al.* 1997; Wisotzkey *et al.* 1998). All of these mutations are dominant maternal enhancers of *dpp* recessive embryonic lethal alleles. Second, we tested for interactions with *screw* (*scw*), a gene encoding a TGF-β family member that augments *dpp* signaling during dorsal-ventral patterning of the embryo (Nguyen *et al.* 1998). We used a gain-of-function mutation (scw^{E1}) that is a dominant zygotic enhancer of *dpp* recessive embryonic lethal alleles and a loss-of-function mutation (scw^{S12}) that does not interact with any *dpp* alleles (Raftery *et al.* 1995). Third, we tested for genetic interactions with a loss-of-function allele of *glass bottom boat*-60A (gbb^{1}). *gbb* encodes a TGF-β family member that cooperates with Dpp to specify positional information in imaginal disks (Khalsa

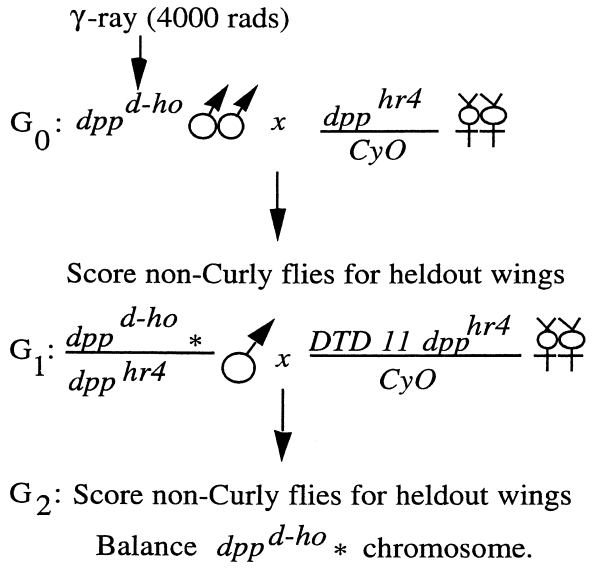

FIGURE 1.—Screen for exceptional DTDs. One version of the screen is shown in which the G1 heldout mutant is male and *DTD11* is used to test for exceptional DTDs. See MATERIALS AND METHODS for details.

et al. 1998). *gbb* is not involved in embryonic dorsal-ventral patterning (WHARTON *et al.* 1999). To date, no interactions between *gbb* mutations and *dpp* recessive embryonic lethal alleles have been reported.

Characterization of *DTD46.4*: *DTD46.4* is a recessive lethal strain obtained in our screen that has a T(2;3) 23C; 93F rearrangement (Figure 2A). To determine which translocation breakpoint results in the recessive lethality, *DTD46.4*-bearing flies were mated to flies with deletions spanning one of the two breakpoints. *DTD46.4* complemented *Df(3R)e-N19,* a deletion of 93B-94. *DTD-46.4* failed to complement *Df(2L)JS17,* a deletion spanning cytological region 23C-D that includes *Mad. Mad* is known to act as a *dpp* transvection disrupter (SEKELSKY *et al.* 1995), so we suspected that *DTD46.4* might be a new allele of *Mad.* To test this hypothesis we chose to further characterize *DTD46.4.*

Complementation tests were conducted with a number of deficiencies and other mutations in the 23C-D cytological region (Figure 2B). The *DTD46.4* chromosome failed to complement the deficiencies *Df(2L)C144*, *Df(2L)DTD52xD51*, and *Df(2L)JS17* and an EMS-induced loss-of-function mutation *l(2)a16.* These five strains are referred to as the 23C complementation group. However, the *DTD46.4* chromosome was viable over Mad^{6}, Mad^{11}, and Mad^{12} and the small deletion *Df(2L)C28* that uncovers *Mad.* These results place the recessive lethality of *DTD46.4* distal to *Mad* in 23C1-2. Polytene *in situ* hybridization studies utilizing a variety of probes demonstrated that the Drosophila Genome Project P1 clones DS00906 and DS07149 span the 23C1-2 breakpoint (data not shown).

We wanted to determine if the 23C1-2 breakpoint of *DTD46.4* was also responsible for disrupting the $dpp^{d\text{-}ho}/dpp^{hr4}$ transvection-dependent phenotype. We tested *Df(2L)C144* and *l(2)a16* for the ability to disrupt this phenotype. Forty-six percent of $dpp^{d\text{-}ho}$ *Df(2L) c144* $/dpp^{hr4}$ flies had heldout wings; of these flies, 47% were severely heldout. Twenty percent of $dpp^{d\text{-}ho}$ *l(2)a16*$/dpp^{hr4}$ flies had heldout wings; of these flies, 50% were severely heldout. These results are similar to those of *DTD46.4.* Twenty-six percent of $dpp^{d\text{-}ho}$ *DTD46.4*$/dpp^{hr4}$ flies had heldout wings; of these flies, 53% were severely heldout. We conclude that the site of *DTD46.4* recessive lethality in 23C1-2 is also the site that disrupts the $dpp^{d\text{-}ho}/dpp^{hr4}$ transvection-dependent phenotype.

During the course of this study we became aware of a new gene located in cytological region 23C1-2. This

TABLE 1

Summary of exceptional DTDs

DTD	Cytology	Comments[a]
42.0	*In(2L)27E-28A;35A*	
43.1	*T(2;4)*	Homozygous viable
43.3	*In(2L)27A;28A*	Homozygous viable
	T(2;3)28A;96B	
43.4	*T(2;3)58B;98F*	
43.5	*T(2;3)25C3-D2;76D*	
44.1	*T(2;3)34D;83E-F*	
44.2	Complex	
44.3	Normal	Homozygous viable
44.4	Normal	
44.5	Normal	Homozygous viable
45.1	Normal	
45.2	Complex	Pseudolinkage observed[b]
45.3	*T(2;3)22F1-3;79*	Pseudolinkage observed
45.4	*T(2;3)33D-F;86C*	Pseudolinkage observed
45.5	*T(2;3)23D1;62C*	
45.6	Three-break inversion on II	
	21A-F/54B-41A	
	60F-54B/29E-F-21F	
	29E-F-40	
45.7	Normal	
45.8	Normal	Homozygous viable
45.9	*T(2;3)55A;66A*	
45.10	Normal	
45.11	Normal	
45.12	*Dp(2;2)21A;24C-D*	Chromosome lost
46.2	*In(2;2)27F-28B;42A*	Homozygous viable
46.3	*In(2;2)21E-F;28D-F*	
46.4	*T(2;3)23C-D;93F*	
46.5	Normal	Homozygous viable
46.6	Normal	Pseudolinkage observed
46.7	*T(2;3)32D-E;82A-B*	
46.8	Normal	
46.9	*In(2;2)40F;59A*	

[a] All exceptional DTDs are recessive lethal unless otherwise indicated.

[b] No females with the exceptional DTD were seen possibly due to an undetected translocation between chromosomes II and Y.

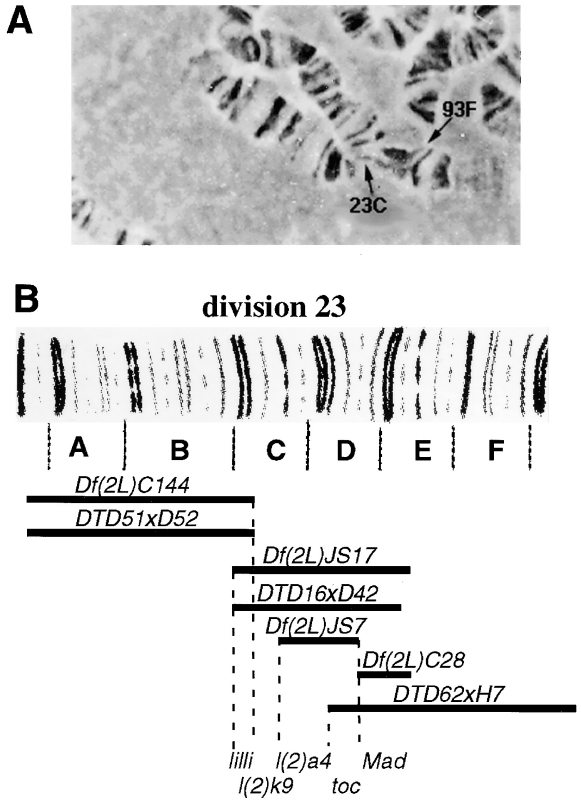

FIGURE 2.—Cytological and genetic mapping of *DTD46.4*. (A) Polytene chromosomes from larvae heterozygous for *DTD46.4* show a *T(2;3)23C;93F* rearrangement. The 23C and 93F cytological regions are indicated by arrows. (B) A schematic representation of cytological region 23C-D. The cytological locations of several deficiencies are shown. The endpoints of all deficiencies are approximate except that the distal breakpoints of *Df(2L)C28* and *Df(2L)JS17* have been cloned (SEKELSKY 1993). The cytological locations of five complementation groups ordered using the deficiencies are indicated by vertical dashed lines. The EMS-induced mutation *l(2)a16* and the *P*-element insertion lines *l(2)00632*, *l(2)k05431* were used to place *lilli* . The *P*-element insertion line *l(2)01361* was used to place *toucan* (*toc*). The EMS-induced mutations Mad^{6}, Mad^{11}, and Mad^{12} were used to place *Mad*. The EMS-induced mutations *l(2)k9* and *l(2)a4* represent complementation groups not currently assigned to a known gene.

gene, *lilliputian* (*lilli*), was identified in two screens for Ras/Mitogen-activated protein kinase (MAPK) signal transduction pathway components. In these screens, loss-of-function mutations in *lilli* were identified as suppressors of gain-of-function phenotypes of *seven in absentia* (SS2-1; NEUFELD *et al.* 1998) and as suppressors of gain-of-function phenotypes of *yan* (SY2-1; REBAY *et al.* 2000). Complementation tests showed that both *DTD46.4* and *l(2)a16* failed to complement either $lilli^{s35}$ (NEUFELD *et al.* 1998) or $lilli^{xs407}$ (REBAY *et al.* 2000). We conclude that members of our 23C1-2 complementation group are alleles of *lilli*. In addition, a screen for genes that interact with dRaf, another component of MAPK signaling pathways, identified a locus in 23C1-2 (DICKSON *et al.* 1996). Loss-of-function mutations in Su(Raf)2A suppress gain-of-function dRaf phenotypes. It seems likely that Su(Raf)2A mutations are also allelic to *DTD46.4* and *lilli*.

We tested four *lilli* alleles for dominant maternal enhancement of *dpp* recessive embryonic lethality. We excluded *Df(2L)JS17* because it uncovers *Mad*. We tested the *lilli* alleles with dpp^{e87}, dpp^{hr56}, dpp^{hr4}, and dpp^{hr92} (ST. JOHNSTON *et al.* 1990). No genetic interactions were detected with the weak alleles dpp^{e87} and dpp^{hr56} (data not shown). However, all *lilli* alleles tested showed significant dominant maternal enhancement of the strong alleles dpp^{hr4} (Figure 3A) and dpp^{hr92} (Figure 3B). Modest dominant zygotic enhancement of dpp^{hr4} was also detected (Figure 3A). Thus, *lilli* alleles that disrupt a *dpp* transvection-dependent phenotype are also dominant enhancers of *dpp* recessive embryonic lethality.

The same alleles of *lilli* were tested for genetic interactions with other genes that function in *dpp* signaling. *lilli* alleles did not enhance the recessive lethality of the loss-of-function mutations Mad^{12}, Med^{1}, sax^{1}, tkv^{8}, scw^{S12}, or gbb^{1}. However, *lilli* alleles showed dominant maternal enhancement of the recessive lethality of scw^{E1} (Figure 3C). scw^{E1} is a gain-of-function allele that is itself a dominant zygotic enhancer of *dpp* recessive embryonic lethality (RAFTERY *et al.* 1995).

We then determined the stage of lethality for the *lilli* loss-of-function mutation *l(2)a16*. We identified *lilli* mutant individuals [*l(2)a16/Df(2L)C144*] using the dominant visible marker *Black cells* (*Bc*; FLYBASE 1999). When *l(2)a16/In(2LR)Gla Bc* males were mated with *Df(2L)C144/In(2LR)Gla Bc* females, only *Bc* larvae were recovered (data not shown). *Bc* is not visible in first instar larvae, suggesting that *lilli* mutants were dying as embryos or as first instar larvae. Examination of *lilli* mutant embryos revealed a partially ventralized phenotype (Figure 4). This phenotype is also seen in zygotic mutant embryos of dpp^{hr56} (WHARTON *et al.* 1993) and scw^{E1} (RAFTERY *et al.* 1995). Several of the hallmarks of this phenotype are a herniated head, internalized filzkorper, and disorganized/expanded denticle bands. Embryos derived from germline clones of weak Su(Raf)2A mutations (*e.g.*, $Su(Raf)2A^{161H1}$) also show this partially ventralized phenotype (DICKSON *et al.* 1996).

Phylogenetic analysis of Lilli: The sequence of a full-length *lilli* cDNA has recently been identified (A. TANG, personal communication). A nearly identical protein of 1665 amino acids, except for an 8-amino-acid truncation at the N terminus, was predicted from genomic sequence by the Berkeley *Drosophila* Genome Project (GenBank accession no. AAF51180; ADAMS *et al.* 2000). BLAST searches using arbitrarily defined segments of the predicted Lilli protein identified very similar regions in four human proteins. These proteins belong to a multigene family called the FMR2/LAF4 family (GECZ *et al.* 1997; NILSON *et al.* 1997).

FMR2 was identified via mutations that result in Fragile-X mental retardation syndrome. Fragile X mental

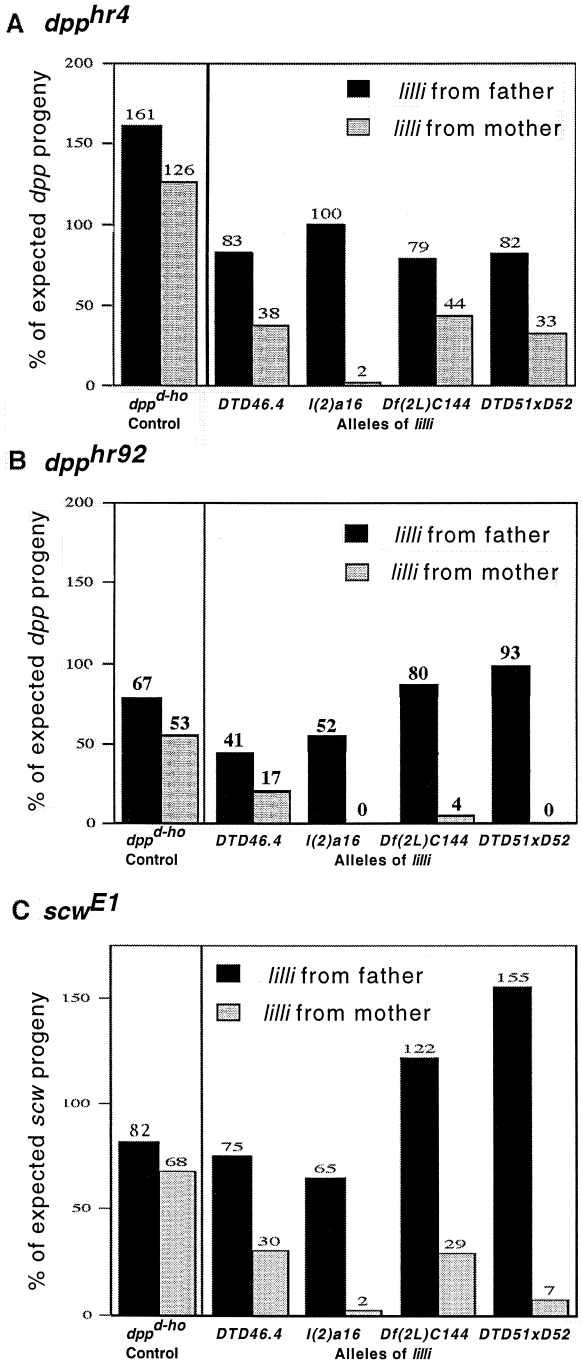

FIGURE 3.—*lilli* mutants are dominant maternal enhancers of *dpp* and *scw* recessive embryonic lethality. (A) dpp^{hr4}. (B) dpp^{hr92}. (C) scw^{E1}. Bars represent the percentage of expected progeny obtained from each mating. The actual value is shown. Solid bars indicate tests for zygotic enhancement of recessive lethality (matings where the father was heterozygous for *lilli*). Shaded bars indicate tests for maternal enhancement of recessive lethality (matings where the mother was heterozygous for *lilli*). In control crosses, $dpp^{d\text{-}ho}$ was utilized in place of the recessive embryonic lethal allele. At least 75 progeny were counted from each mating. In maternal enhancement crosses, adult escaper progeny with the *dpp* or *scw* mutant chromosomes had no defects in their eyes, wings, or legs and females were fertile.

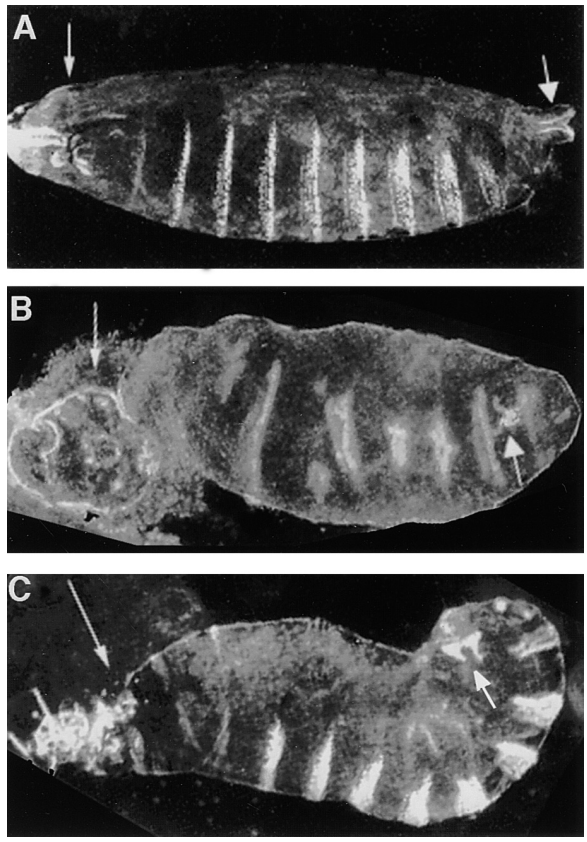

FIGURE 4.—*lilli* mutant embryos have a partially ventralized phenotype. (A) Wild-type embryo oriented anterior to the left and dorsal toward the top. The head skeleton at the anterior (thin arrow) and the filzkorper at the posterior (wide arrow) are noted. (B) *l(2)a16/Df(2L)C144* embryo. In this example, the head is dysmorphic (thin arrow) and the filzkorper are internalized (wide arrow). The ventral denticle bands are expanded toward the dorsal side and are disorganized. This embryo is similar to a dpp^{hr56} mutant embryo (WHARTON *et al.* 1993). (C) *l(2)a16/Df(2L)C144* embryo. In this example, the head skeleton is completely herniated (thin arrow), the embryo is bent into a U shape, and the filzkorper are internalized (wide arrow). This embryo is similar to a scw^{E1} mutant embryo (RAFTERY *et al.* 1995).

retardation syndrome is the most common form of inherited mental retardation in humans. FMR2 is highly expressed in the fetal brain (reviewed in JIN and WARREN 2000). LAF4 was identified via chromosomal translocations that result in Burkitt's lymphoma. Burkitt's lymphoma is associated with highly malignant tumors and is the most common form of childhood cancer. LAF4 is highly expressed in fetal lymphoid tissue, particularly in preB-cells (MA and STAUDT 1996). The other family members, AF4 and AF5, were identified via distinct chromosomal translocations that give rise to infant acute lymphoblastic leukemia (ALL). At this time, ALL is resistant to treatment and invariably fatal. AF5 is highly expressed in fetal heart, lung, and brain while AF4 is highly expressed in fetal heart, liver, and brain (LI *et al.* 1998; TAKI *et al.* 1999). These human proteins

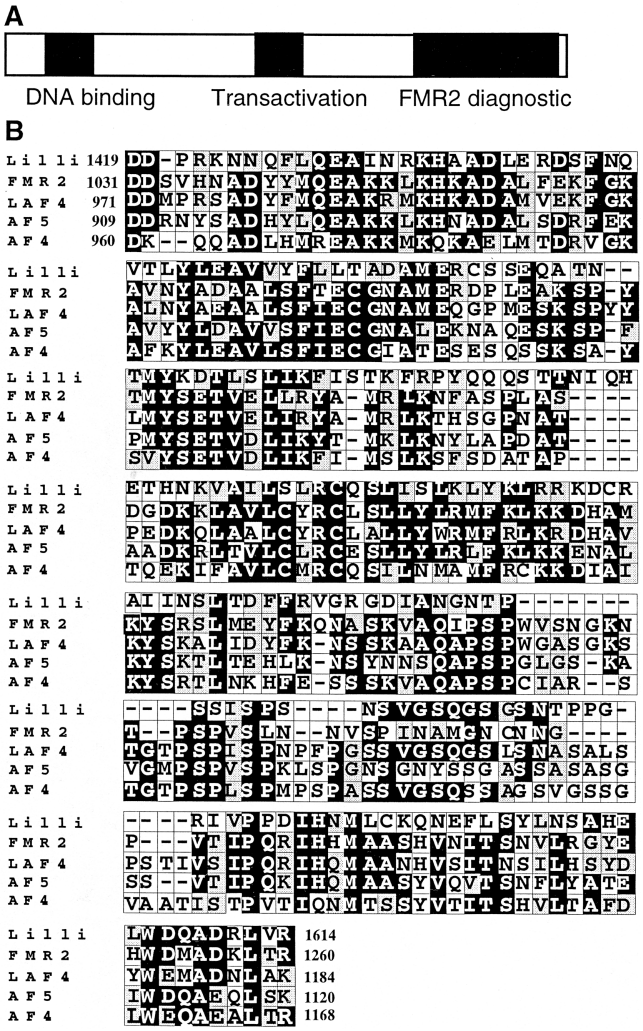

FIGURE 5.—Lilli is a member of the FMR2/LAF4 multigene family. (A) Schematic representation of an FMR2/LAF4 family member. The locations of three conserved domains are shown. (B) Amino acid alignment of the FMR2/LAF4 diagnostic domain of Lilli and the four human family members. Black boxes indicate identical amino acids at that position in at least three sequences. Shaded boxes indicate a similar amino acid at that position in at least three sequences. Gaps in the alignment minimize the number of mutations required to explain all differences between the sequences. Amino acid numbers for each sequence are indicated. Accession numbers are as follows: Lilli, AAF51180; FMR2, AAA99416, AF5–AAF18981, AF4–CAB69660, and LAF4–NP002276.

are nuclear proteins capable of DNA binding and transcriptional activation (LI *et al.* 1998).

Previous studies of this family identified three conserved domains (GECZ *et al.* 1997; Figure 5A). Near the N terminus there is a conserved domain that includes a high mobility group I (HMGI) DNA-binding motif. In the center there is a conserved transcriptional activation domain with no recognizable motif. At the C terminus there is a highly conserved domain diagnostic for the FMR2/LAF4 family with no recognizable motif and unknown function. BLAST searches showed that Lilli contains segments very similar to each of these domains in the proper location.

We conducted an exhaustive analysis of the *D. melanogaster* genome database using the conserved regions of Lilli and the four human FMR2/LAF4 family sequences. A total of 15 different domains were used as query sequences. We did not identify any additional proteins that contain all 3 conserved domains. Nor did we identify a group of consecutive (mis)predicted proteins that contain the 3 conserved domains in the proper order. We were not able to identify any additional proteins with obvious similarity to only the C-terminal domain diagnostic for the FMR2/LAF4 family. At this time, Lilli appears to be the only *D. melanogaster* member of this multigene family. We then conducted the same set of exhaustive searches using the *C. elegans* genome database. We did not identify any proteins with all three domains or any with convincing similarity to the C-terminal diagnostic domain.

An alignment of the C-terminal domain of Lilli with all of the human family members is shown in Figure 5B. This region of Lilli shows extensive amino acid similarity with all of the human proteins. However, the alignment gives the overall impression that the four human family members are more similar to each other than they are to Lilli. The degree of amino acid identity and similarity, calculated from pairwise comparisons between all five sequences for each of the conserved domains, is shown in Table 2. The comparisons show that there is a significant amount of amino acid similarity (>51%) between Lilli and each human protein in all domains. The human proteins show >63% similarity in all domains with most comparisons >72%.

Data derived from pairwise comparisons were used to construct phylogenetic trees for each domain. A composite tree was also constructed from an alignment consisting of all three domains (Figure 6). Only slight differences were noted between the individual domain trees and the composite tree. The similarity of the trees suggests that the tripartite structure of these proteins predates the divergence of arthropods and vertebrates. The composite tree shows that the human family members are indeed more similar to each other than they are to Lilli. This distinction is 100% supported by the bootstrap analysis. The composite tree contains two clusters of human sequences that are also strongly supported. Sequence clusters with bootstrap values >75% are considered biologically meaningful (NEWFELD *et al.* 1999).

DISCUSSION

Exceptional DTD screen: We conducted a genetic screen for new components of the *dpp* signaling pathway. The screen identified 30 exceptional DTDs. These mutations disrupt transvection at the *dpp* locus but are not associated with asynapsis at *dpp*. Mutations were not recovered in genes involved in *dpp* signaling that act as

TABLE 2

Pairwise amino acid comparisons

	Lilli	FMR2	AF5	LAF4	AF4
N-terminal domain (DNA binding)					
Lilli		24.2[a]	25.8	29.0	23.7
FMR2	52.4[b,c]		42.9	51.3	39.3
AF5	60.5	68.1		45.7	43.0
LAF4	54.8	75.6	69.8		39.0
AF4	51.7	63.2	68.5	65.2	
Central domain (transactivation)					
Lilli		31.8	24.5	31.4	19.0
MFR2	65.9		48.0	47.3	41.3
AF5	62.7	69.3		52.0	64.8
LAF4	64.7	75.7	83.0		51.9
AF4	64.8	72.0	86.7	83.0	
C-terminal domain (FMR2 diagnostic)					
Lilli		29.0	31.5	30.6	29.6
FMR2	61.4		53.5	60.5	42.3
AF5	61.1	79.8		50.9	45.8
LAF4	58.0	85.6	84.6		47.2
AF4	61.1	72.3	78.7	77.1	

[a] Percentage identity is indicated above the diagonal.
[b] Percentage similarity is indicated below the diagonal.
[c] Percentage similarity is the sum of the percentage of identical amino acids and the percentage of conservative amino acid substitutions. Conservative substitutions are defined biochemically (Smith and Smith 1990).

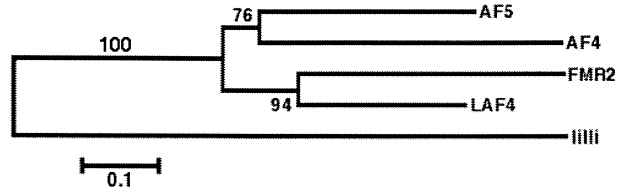

Figure 6.—Phylogenetic analysis of FMR2/LAF4 family members. Evolutionary relationships between FMR2/LAF4 family members, based on a composite alignment of the three conserved domains, are shown. The length of the alignment was 522 amino acids. The tree is unrooted. The numbers represent the relative incidence of that particular relationship (in *percentage*) during bootstrap resampling using 1000 replicates. Branch lengths are drawn to scale on the basis of the number of amino acid substitutions per site.

exceptional DTDs, such as *Mad* (Sekelsky *et al.* 1995), suggesting that our screen was not exhaustive.

To determine if any of the exceptional DTDs were associated with mutations in *dpp* signaling pathway components, we utilized three assays. These are the same tests used in the initial characterization of the Dpp signal transducers *Mad* and *Med* (Raftery *et al.* 1995; Sekelsky *et al.* 1995). First, we tested each DTD for genetic interactions with *dpp* alleles that were not part of the original screen. The original screen exploited *dpp*'s role in adult appendage formation. In this test we examined each DTD for dominant maternal enhancement of *dpp* alleles that disrupt embryonic dorsal-ventral patterning. Second, we tested each DTD for genetic interactions with other genes that participate in *dpp* signaling such as *sax*, *scw*, and *Mad*. Third, we looked for similarities between the mutant phenotype of a DTD, or another member of its complementation group, and *dpp* mutant phenotypes.

Characterization of *DTD46.4*: The first mutation we chose to characterize in detail was the 23C1-2 breakpoint of *DTD46.4*. This breakpoint is allelic to mutations in the newly discovered gene *lilli*. The results of our genetic tests suggest that *lilli* is a strong candidate for a new component of the *dpp* signaling pathway. First, *lilli* mutations enhance *dpp* heldout phenotypes and embryonic recessive lethality. The enhancement of *dpp* embryonic lethality by *lilli* mutations is not as strong as that of *Mad* or *Med* mutations (Raftery *et al.* 1995; Sekelsky *et al.* 1995). Mutations in *Mad* or *Med* enhance weak *dpp* alleles while *lilli* mutations do not. Second, *lilli* mutations enhance the recessive embryonic lethality of a gain-of-function allele of the TGF-β family member *scw*. *scw* augments *dpp* signaling in embryonic dorsal-ventral patterning. To date, tests for interactions between scw^{E1} and other *dpp* pathway components such as *Mad* or *Med* have not been reported. *lilli* mutations do not enhance the recessive lethality of mutations in genes that encode Dpp signal transduction proteins (*sax*, *tkv*, *Mad*, or *Med*). Third, *lilli* homozygous mutant embryos have dorsal-ventral patterning defects similar to zygotic mutant embryos of *dpp* and *scw*. Utilizing these genetic criteria, *lilli* has as strong a connection to *dpp* signaling as *Mad* and *Med*.

In addition to our screen, *lilli* mutations were identified in three other screens. In these screens, *lilli* mutations suppress dominant phenotypes generated by activated MAPK signaling pathways (Dickson *et al.* 1996; Neufeld *et al.* 1998; Rebay *et al.* 2000). MAPK signal transduction is initiated by transmembrane receptor tyrosine kinases. These receptors transmit the signal to transcription factors in the nucleus utilizing a cascade of tyrosine kinases. Alternatively, TGF-β family members such as Dpp bind to transmembrane receptor serine-threonine kinases. These receptors transmit the Dpp signal via a nonkinase mechanism of nuclear translocation by Mad and Med. The ability of *lilli* loss-of-function mutations to suppress MAPK signaling gain-of-function phenotypes and to enhance *dpp* loss-of-function phenotypes is very intriguing. To our knowledge, *lilli* is one of the first genes involved in MAPK and TGF-β signaling pathways in a developmental system.

lilli encodes a transcription factor (A. Tang, personal communication). This fact suggests one hypothesis for *lilli*'s role in MAPK signaling and another hypothesis for a role in Dpp signaling. For MAPK signaling, *lilli* may be a transcriptional effector of MAPK signal transduction pathways. This hypothesis fits the observation

that *lilli* loss-of-function mutations suppress MAPK signaling gain-of-function phenotypes (Dickson *et al.* 1996; Neufeld *et al.* 1998; Rebay *et al.* 2000). For Dpp signaling, *lilli* may be a maternally supplied transcriptional activator of *dpp* and/or *scw* during dorsal-ventral patterning. This hypothesis fits three of our observations. First, *lilli* loss-of-function mutations maternally enhance the recessive lethality of several dorsal-ventral patterning mutations. Second, *lilli* mutant phenotypes mimic the mutant phenotypes of dorsal-ventral patterning mutations. Third, *lilli* mutations do not enhance, either maternally or zygotically, the embryonic lethality of genes that encode Dpp signal transduction proteins. Alternatively, *lilli* could participate in a signaling pathway parallel to the Dpp pathway that is also required for the expression of Dpp target genes.

To test the hypothesis that *lilli* is a maternal activator of *dpp* and/or *scw* in dorsal-ventral patterning one would examine *dpp* and *scw* expression in embryos derived from *lilli* mutant germline clones. The prediction is that there would be reduced *dpp* and/or *scw* expression in these embryos during dorsal-ventral patterning. At this time, maternal activators of zygotic dorsal-ventral patterning genes such as *dpp* and *scw*, as opposed to well-known repressors such as Dorsal (Anderson 1998), are unknown. It is tempting to speculate that a maternal MAPK signal induces *lilli* to activate *dpp* in embryonic dorsal-ventral patterning.

Determining a role for *lilli* in *dpp* signaling in adult wings, where *lilli* mutations enhance the heldout phenotype, is more problematic. There is no *a priori* reason to believe that *lilli* plays the same role in *dpp* signaling during dorsal-ventral patterning and adult appendage formation but it seems a logical place to begin. Thus it is possible that *lilli* activates *dpp* expression in wing imaginal disks. This hypothesis fits a report of *dpp* transcriptional regulation by the heldout *cis*-regulatory region (Hepker *et al.* 1999). In this study, two consensus HMGI binding sites (A/TA/TCAAG; van de Wetering *et al.* 1991) are identified as dTcf binding sites in the heldout region. The expression of reporter genes containing the *dpp* heldout region was disrupted when these putative dTcf sites were mutagenized. In addition, dominant negative forms of dTcf expressed in wing imaginal disks eliminated *dpp* expression. As a result, the authors conclude that dTcf is required for *dpp* expression by the heldout *cis*-regulatory region. However, these data do not preclude the possibility that the HMGI binding sites are actually the target of another HMGI domain protein, such as Lilli. To determine which HMGI domain protein is actually responsible for *dpp* expression from heldout regulatory sequences, one would examine *dpp* expression in wing imaginal disks bearing dTcf or *lilli* somatic clones.

Phylogenetic analysis of Lilli: Lilli shares three conserved domains with the four human members of the FMR2/LAF4 multigene family of transcription factors. The human family members are all developmental genes with high levels of fetal tissue-specific expression. Mutations in these genes have devastating effects. Mutations in FMR2 lead to mental retardation and mutations in LAF4, AF4, and AF5 lead to treatment-resistant forms of childhood cancer. Our analyses revealed several interesting features of this newly expanded multigene family.

First, BLAST searches demonstrate that Lilli is unique among *D. melanogaster* genes. No other sequences with convincing similarity to any FMR2/LAF4 family member were found in the *D. melanogaster* genome. We found this surprising for a gene associated with *dpp* signaling. To date, all known components of *dpp* signaling pathways belong to large multigene families with several members in *D. melanogaster* (Newfeld *et al.* 1999). Second, pairwise amino acid comparisons suggested that the human genes in the FMR2/LAF4 family are more similar to each other than to Lilli. This suggestion is supported with a 100% bootstrap value by phylogenetic analysis. The basal branch of the tree separates Lilli from the human genes.

Taken together, these two observations strongly support the hypothesis that *lilli* is the *D. melanogaster* homolog of the human FMR2/LAF4 family members. We employ the strict evolutionary definition of homology (genes identical by descent from a common ancestor). In this case, we refer to the FMR2/LAF4 family progenitor in the common ancestor of arthropods and chordates. The absence of any FMR2/LAF4 family members in *C. elegans* suggests that the FMR2/LAF4 family progenitor arose after the split of nematodes and arthropods.

In addition, the appearance of two pairs of sequences for the four human FMR2/LAF4 family members in the phylogenetic tree is compatible with Ohno's (1970) hypothesis that two rounds of genome duplication have occurred in the vertebrate lineage. The original member of each pair of human sequences could have been generated during the first event and the second member of each pair by the second event. Alternatively, the four human sequences could have been generated by three independent gene duplication events. Additional phylogenetic data are needed to distinguish these hypotheses.

In summary, Lilli appears to function in both MAPK and Dpp signaling pathways, suggesting important roles in Drosophila development. Detailed studies of Lilli function in Drosophila will likely shed light on the wild-type function of human FMR2/LAF4 family members. For example, the functional conservation of *dpp* signaling pathway components suggests that human homologs of Lilli's transcriptional targets are likely to be targets of human FMR2/LAF4 family members. Given that mutations in these human genes lead to mental retardation or childhood cancer and that information on human developmental genes is difficult to gather directly, studies of Lilli are an important weapon in our efforts to combat these human syndromes.

We thank Amy Tang for providing the *lilli* cDNA sequence prior to publication and Sudhir Kumar for assistance in generating phylogenetic trees. This study was begun in Bill Gelbart's lab and we thank Bill and Jeff Sekelsky for assistance with cytology. We thank Mike O'Connor for valuable discussions. S.J.N. is supported by a Basil O'Connor Starter Scholar Research Award from the March of Dimes.

LITERATURE CITED

Adams, M., S. Celniker, R. Holt, C. Evans, J. Gocayne *et al.*, 2000 The genome sequence of *Drosophila melanogaster*. Science **287:** 2185–2194.

Anderson, K., 1998 Pinning down positional information: dorsal-ventral polarity in the Drosophila embryo. Cell **95:** 439–442.

Brummel, T., V. Twombly, G. Marques, J. Wrana, S. Newfeld *et al.*, 1994 Characterization and relationship of Dpp receptors encoded by the *saxophone* and *thickveins* genes in *Drosophila*. Cell **78:** 251–262.

Dickson, B., A. van der Straten, M. Dominguez and E. Hafen, 1996 Mutations modulating Raf signaling in Drosophila eye development. Genetics **142:** 163–171.

Felsenstein, J., 1985 Confidence limits on phylogenies: an approach using the bootstrap. Evolution **39:** 783–791.

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

Gecz, J., S. Bielby, G. Sutherland and J. Mulley, 1997 Gene structure and subcellular localization of FMR2, a member of a new family of putative transcription activators. Genomics **44:** 201–213.

Gelbart, W., 1982 Synapsis-dependent allelic complementation at the *decapentaplegic* gene complex in *Drosophila melanogaster*. Proc. Natl. Acad. Sci. USA **79:** 2636–2640.

Gelbart, W., 1989 The *decapentaplegic* gene: a TGF-β homolog controlling pattern formation in Drosophila. Development **107** (Suppl.): 65–74.

Hepker, J., R. Blackman and R. Holmgren, 1999 Cubitus interruptus is necessary but not sufficient for direct activation of a wing-specific *decapentaplegic* enhancer. Development **126:** 3669–3677.

Jeanmougin, F., J. Thompson, M. Gouy, D. Higgins and T. Gibson, 1998 Multiple sequence alignment with Clustal X. Trends Biochem. Sci. **23:** 403–405.

Jin, P., and S. Warren, 2000 Understanding the molecular basis of Fragile X syndrome. Hum. Mol. Genet. **9:** 901–908.

Khalsa, O., J. Yoon, S. Torres-Schumann and K. Wharton, 1998 TGF-β/BMP superfamily members, Gbb-60A and Dpp, cooperate to provide pattern information and establish cell identity in the Drosophila wing. Development **125:** 2723–2734.

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

Lewis, E., 1954 The theory and application of a new method of detecting chromosomal rearrangements in *Drosophila melanogaster*. Am. Nat. **88:** 225–239.

Li, Q., J. Frestedt and J. Kersey, 1998 AF4 encodes a ubiquitous protein that in both native and MLL-AF4 fusion types localizes to subnuclear compartments. Blood **92:** 3841–3847.

Ma, C., and L. Staudt, 1996 LAF4 encodes a lymphoid nuclear protein with transactivation potential that is homologous to AF4, the gene fused to MLL in t(4;11) leukemias. Blood **87:** 743–745.

Nei, M., and S. Kumar, 2000 *Molecular Evolution and Phylogenetics*. Oxford University Press, New York.

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

Newfeld, S., A. Mehra, M. Singer, J. Wrana, L. Attisano *et al.*, 1997 *Mad* participates in a DPP/TGF-β responsive serine-threonine kinase signal transduction cascade. Development **124:** 3167–3176.

Newfeld, S., R. Wisotzkey and S. Kumar, 1999 Molecular evolution of a developmental pathway: phylogenetic analyses of TGF-β family ligands, receptors and Smad signal transducers. Genetics **152:** 783–795.

Nguyen, M., S. Park, G. Marques and K. Arora, 1998 Interpretation of a BMP activity gradient in Drosophila embryos depends on synergystic signaling by two Type I receptors, SAX and TKV. Cell **95:** 495–506.

Nilson, I., M. Reichel, M. Ennas, R. Greim, C. Knorr *et al.*, 1997 Exon/intron structure of the human AF4 gene, a member of the AF4/LAF4/FMR2 gene family coding for a nuclear protein with structural alterations in acute leukaemia. Br. J. Haematol. **98:** 157–169.

Ohno, S., 1970 *Evolution by Gene Duplication*. Springer-Verlag, New York.

Padgett, R., R. St. Johnston and W. Gelbart, 1987 A transcript from a Drosophila pattern gene predicts a protein homologous to the TGF-β family. Nature **325:** 81–84.

Raftery, L., V. Twombly, K. Wharton and W. Gelbart, 1995 Genetic screens to identify elements of the *dpp* pathway in Drosophila. Genetics **139:** 241–254.

Rebay, I., F. Chen, F. Hsiao, P. Kolodziej, B. Kuang *et al.*, 2000 A genetic screen for novel components of the Ras/Mitogen-activated protein kinase signaling pathway that interact with the *yan* gene of Drosophila identifies *split ends*, a new RNA recognition motif-containing protein. Genetics **154:** 695–712.

Riggins, G., K. Kinzler, B. Vogelstein and S. Thiagalingam, 1997 Frequency of Smad gene mutations in human cancers. Cancer Res. **57:** 2578–2580.

Saitou, N., and M. Nei, 1987 The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. **4:** 406–425.

Schuler, G., S. Altschul and D. Lipman, 1991 A workbench for multiple alignment construction and analysis. Proteins **9:** 180–190.

Sekelsky, J., 1993 Mothers against dpp: a gene required for decapentaplegic function in Drosophila melanogaster. Ph.D. Thesis, Harvard University, Cambridge, MA.

Sekelsky, J., S. Newfeld, L. Raftery, E. Chartoff and W. Gelbart, 1995 Genetic characterization and cloning of *Mad*, a gene required for *dpp* function in Drosophila. Genetics **139:** 1347–1358.

Smith, R., and T. Smith, 1990 Automated generation of primary sequence patterns from sets of related protein sequences. Proc. Natl. Acad. Sci. USA **87:** 118–122.

Spencer, F., F. Hoffmann and W. Gelbart, 1982 *decapentaplegic*: a gene complex affecting morphogenesis in *Drosophila melanogaster*. Cell **28:** 451–461.

St. Johnston, R., F. Hoffmann, R. Blackman, D. Segal, R. Grimaila *et al.*, 1990 Molecular organization of the *dpp* gene in Drosophila. Genes Dev. **4:** 1114–1127.

Taki, T., H. Kano, M. Taniwaki, M. Sako, M. Yanagisawa *et al.*, 1999 AF5q31, a newly identified AF4 related gene, is fused to MLL in infant acute lymphoblastic leukemia with ins (5;11) (q31;q13q23). Proc. Natl. Acad. Sci. USA **96:** 14535–14540.

van de Wetering, M., M. Oosterwegel, D. Dooijes and H. Clevers, 1991 Identification and cloning of TCF1, a T lymphocyte-specific transcription factor containing a sequence specific HMG box. EMBO J. **10:** 123–132.

Waldrip, W., E. Bikoff, P. Hoodless, J. Wrana and E. Robertson, 1998 Smad2 signaling in extraembryonic tissues determines anterior-posterior polarity of the early mouse embryo. Cell **92:** 797–808.

Wharton, K., R. Ray and W. Gelbart, 1993 An activity gradient of *decapentaplegic* is necessary for the specification of dorsal pattern elements in the Drosophila embryo. Development **117:** 807–822.

Wharton, K., R. Ray, S. Findley, H. Duncan and W. Gelbart, 1996 Molecular lesions associated with alleles of *decapentaplegic* identify residues necessary for TGF-β/BMP cell signaling in *Drosophila melanogaster*. Genetics **142:** 493–505.

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

Wisotzkey, R., A. Mehra, D. Sutherland, L. Dobens, X. Liu *et al.*, 1998 *Medea* is a Drosophila Smad4 homolog that is differentially required to potentiate Dpp responses. Development **125:** 1433–1445.

Communicating editor: C.-I Wu