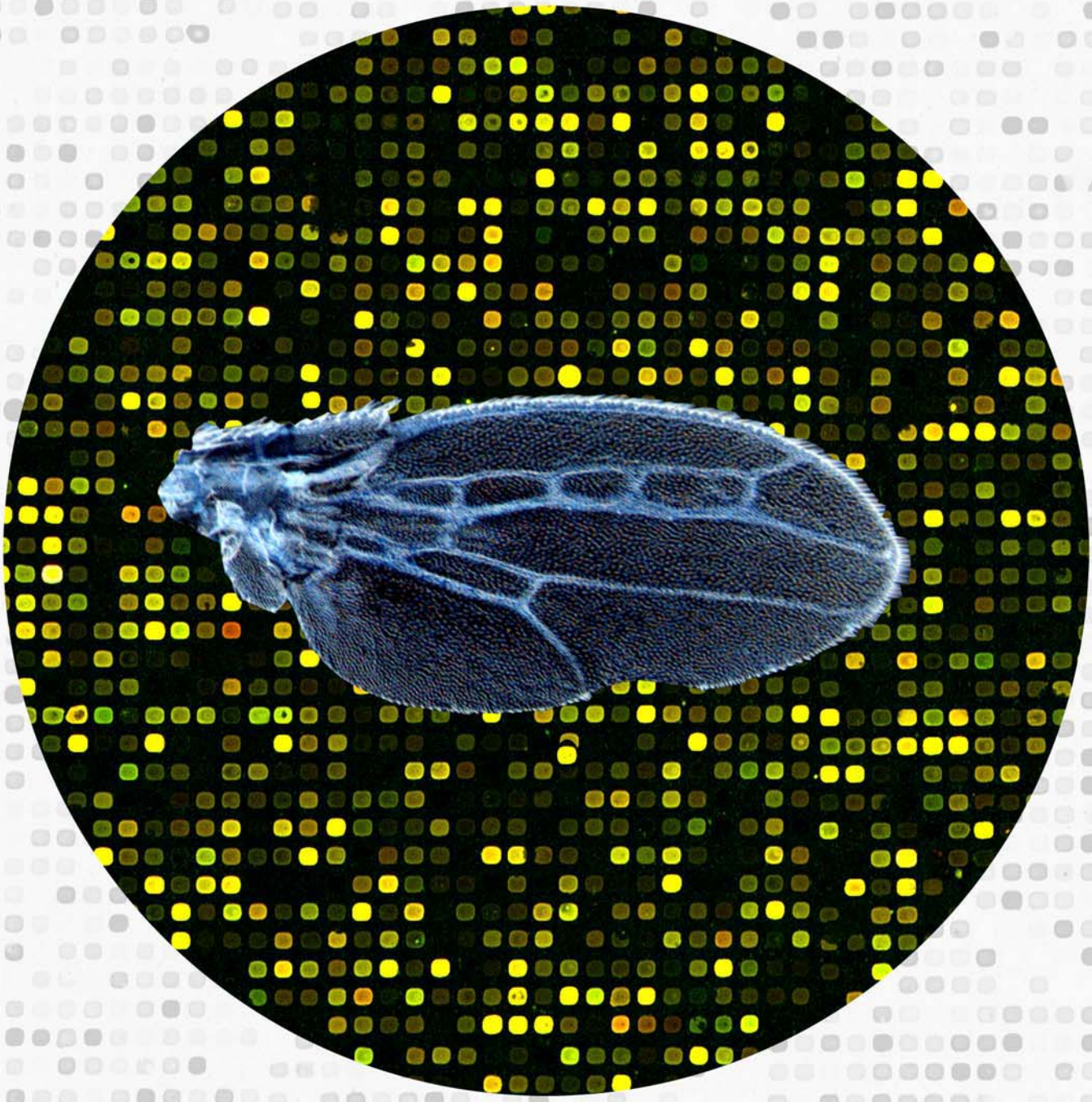
El transcriptoma d'*ash2*: dianes i funció



Sergi Beltran i Agulló

Departament de Genètica
Universitat de Barcelona

RESULTATS

Estructura dels Resultats

Els Resultats d'aquesta Tesi es presenten en dos articles, cadascun amb la seva Introducció, Material i mètodes, figures, taules i Resultats i Discussió particulars. Cada article ve precedit d'un resum en català, i a continuació del cos principal de cada article s'hi troben les dades, figures i taules suplementàries i/o de suport. Per tal de facilitar la visualització de les taules més llargues, aquestes només es troben en el disc compacte que s'adjunta.

ARTICLE 1

Títol

Transcriptional network controlled by the trithorax-group gene *ash2* in *Drosophila melanogaster*.

Autors (Any)

Beltran, S.[1], Blanco, E.[2], Serras, F.[1], Perez-Villamil, B.[3], Guigo, R.[2], Artavanis-Tsakonas, S.[3], and Corominas, M.[1] (2003)

[1]Departament de Genètica, Universitat de Barcelona. Diagonal 645, 08028 Barcelona
[2]Grup de Recerca en Informàtica Biomèdica, Institut Municipal d'Investigació Mèdica, Universitat Pompeu Fabra, Centre de Regulació Genòmica, Dr. Aiguader 80, 08003 Barcelona
[3]Massachusetts General Hospital Cancer Center, Harvard Medical School, Bl149, 13th Street, Charlestown, MA 02129

Referència

Proc Natl Acad Sci U S A **100**(6): 3293-3298.

Resum

El gen *absent, small, or homeotic discs 2* (*ash2*) és un membre del grup dels trithorax (*trxG*), reguladors positius de la transcripció dels gens homeòtics. Els al·lels mutants per *ash2* són letals larvals o pupals i presenten anomalies en els discs imaginals i en el cervell. Hem identificat un nou transcrit, més petit que el descrit fins al moment, pel gen *ash2*. L'al·lel utilitzat en aquest treball no presenta el transcrit més gran i és un mutant nul per la funció trxG d'ASH2. Per tal d'identificar gens diana d'ASH2, hem estudiat el patró d'expressió de larves mutants d'*ash2* mitjançant microarrays de cADN. Entre els gens que varien els seus nivells d'expressió se n'hi troben d'involucrats en múltiples processos, com la proliferació, l'adhesió i el cicle cel·lulars. Alguns d'aquests gens han estat validats mitjançant tècniques d'expressió alternatives o estudis funcionals. Aquest treball representa el primer anàlisi per microarrays d'un gen trxG.

Aportació personal al treball

La meva aportació a aquest article ha estat:

- seqüenciació de l'al·lel *ash2*^{l1}
- caracterització de la severitat dels al·lels *ash2*^{l1}, *ash2*^{112411} i *ash2*^{1} amb combinacions heteroal·lèliques i anàlisi d'expressió
- descripció del transcrit ash2.1 per 5' RACE
- anàlisi de dades dels microarrays de cADN
- comprovació de resultats per RT-PCR i anàlisi clonal (prèvia generació de la línia FRT82B- *ash2*^{l1})
- participació activa en l'el·laboració de les figures i del manuscrit

Transcriptional network controlled by the trithorax-group gene *ash2* in *Drosophila melanogaster*

Sergi Beltran*, Enrique Blanco†, Florenci Serras*, Beatriz Pérez-Villamil‡, Roderic Guigó†, Spyros Artavanis-Tsakonas‡, and Montserrat Corominas*§

*Departament de Genètica, Universitat de Barcelona, Diagonal 645, 08028 Barcelona, Spain; †Grup de Recerca en Informàtica Biomèdica, Institut Municipal d'Investigació Mèdica, Universitat Pompeu Fabra, Centre de Regulació Genòmica, Dr. Aiguader 80, 08003 Barcelona, Spain; and ‡Massachusetts General Hospital Cancer Center, Harvard Medical School, Bl 149, 13th Street, Charlestown, MA 02129

Communicated by Walter J. Gehring, University of Basel, Basel, Switzerland, January 10, 2003 (received for review July 20, 2002)

The transcription factor *absent, small,* or *homeotic discs 2* (*ash2*) gene is a member of the trithorax group of positive regulators of homeotic genes. Mutant alleles for *ash2* are larval/pupal lethals and display imaginal disc and brain abnormalities. The allele used in this study is a true mutant for the trithorax function and lacks the longest transcript present in wild-type flies. In an attempt to identify gene targets of *ash2*, we have performed an expression analysis by using cDNA microarrays. Genes involved in cell cycle, cell proliferation, and cell adhesion are among these targets, and some of them are validated by functional and expression studies. Even though trithorax proteins act by modulating chromatin structure at particular chromosomal locations, evidence of physical aggregation of *ash2*-regulated genes has not been found. This work represents the first microarray analysis, to our knowledge, of a trithorax-group gene.

The trithorax group (trx-G) of activators and the Polycomb group (Pc-G) of repressors maintain the correct expression of several key developmental regulators, including the homeotic genes. Pc-G mutants exhibit posterior transformations in embryos and adults caused by derepression of homeotic loci in flies (1) and vertebrates (2). In contrast, proteins of the trx-G are required for the maintenance of activation of homeotic loci (3). Pc-G and trx-G proteins function in distinct multiprotein complexes that are believed to control transcription by changing the structure of chromatin, organizing it into either a "closed" or an "open" conformation (ref. 4 and references therein). It is thought that Pc-G and trx-G regulate many targets in addition to homeotic genes, indicating that epigenetic maintenance of activated or repressed states might represent a fundamental developmental mechanism (5).

The *ash2* (*absent, small, or homeotic discs 2*) gene is a member of the trx-G discovered, together with *ash1*, in a screen for late larval/early pupal lethals that had imaginal discs abnormalities (6–9). The ASH2 protein has a proline-, glutamic acid-, serine-, and threonine-rich region sequence characteristic of short-lived proteins, a putative double zinc-finger domain, a bipartite nuclear localization signal, and a SPRY domain (10). Biochemical studies have shown that ASH1 and ASH2 are subunits of distinct protein complexes and that ASH2 elutes in fractions with an apparent native molecular mass of 500 kDa (11). More recently it has been reported that the *Saccharomyces cerevisiae* SET1 complex includes two putative ASH2 homologues as well as a protein (SET1) with high similarity to TRX. This complex methylates histone 3 lysine 4, reinforcing the notion that methylation is important for regulating the transcriptional accessibility of chromatin (12–14).

Mutations in *ash2* cause the homeotic transformations expected for genes in this group in addition to a variety of additional pattern formation defects. *ash2* mutant hemizygotes that are able to survive until eclosion include supernumerary legs, duplication of thoracic bristles, and transformation of campaniform sensilla to bristles (15). The line *l(3)112411* was isolated from a collection of *P-lacW* element insertional mutagenesis in the third chromosome (16) and corresponds to a new *ash2* allele. The few homozygous flies that reach the adult stage are sterile and display anomalous patterns of appendage differentiation. Clonal analysis in adult wings of homozygous cells for the stronger allele $ash2^{I1}$ reveals a role in vein–intervein patterning, because a reduction of intervein tissue and an increase of vein tissue are observed autonomously and nonautonomously in the clones (17). Moreover, a failure to form joints or fusion of several fragments leads to shortened legs when big clones are generated. Taken together, the pleiotropic phenotypes observed could not be explained only by changes in homeotic gene expression; therefore, more genes should be responding to the loss of *ash2* function.

In this work, we have applied cDNA microarray technology to analyze the transcription profile of $ash2^{I1}$ mutant larvae in comparison with WT, in an attempt to delineate the transcriptional consequences of lack of *ash2* function and to identify genes that may fulfill the criteria of *ash2* targets. Microarrays have been used to study a variety of biological processes, from differential gene expression in yeast sporulation (18) to human tumors (19). In the case of *Drosophila*, they were initially applied to analyze development during metamorphosis (20) and more recently for analyzing patterns of transcription under different situations or mutant conditions (21–26). The microarray analysis presented here represents the first approach, to our knowledge, to monitoring the genome wide-expression profile from a mutant of the trx-G. The regulated genes have been automatically classified and clustered according to the functional criteria in the Gene Ontology (GO) database (27), with the aim of finding a differential distribution among the regulated genes.

Materials and Methods

Canton-S and $ash2^{I1}/TM6C$ strains were maintained on standard medium and experiments performed at 25°C. Details of mitotic clone generation, 5′-rapid amplification of cDNA ends, Northern blot, and RT-PCR are published as *Supporting Materials and Methods* on the PNAS web site, www.pnas.org.

Microarray Analysis. One to three micrograms of poly(A) RNA from WT or mutant larvae were labeled by reverse transcription incorporation of Amino-allyl dUTP and coupling to cyanine dye (Cy3- or Cy5-NHS esters, Amersham Biosciences) and hybridized to cDNA microarrays constructed by using PCR products directly amplified from the DROSOPHILA gene collection 1.0 (www.fruitfly.org/dgc/index.html). GENEPIX 3.0 (Axon Instru-

Abbreviations: GO, Gene Ontology; SAM, significance analysis of microarrays; MF, molecular function; BP, biological process; CC, cellular component; trx-G, trithorax group; UBX, ultrabithorax; FLP–FRT, flipase–flipase recombination target.

§To whom correspondence should be addressed. E-mail: mcorominas@ub.edu.

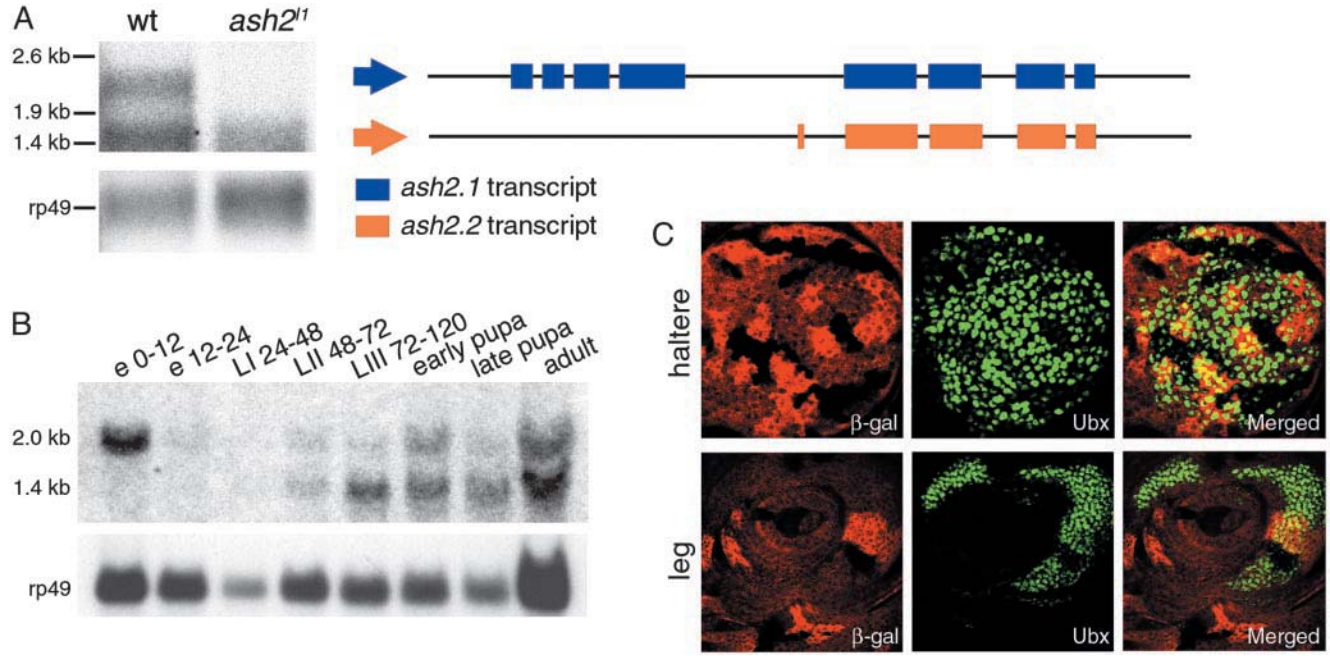


Fig. 1. (*A–C*) Molecular characterization of WT and mutant *ash2* mRNA. (*A*) Northern blot analysis of poly(A) RNA extracted from WT and $ash2^{I1}$ homozygous third instar larvae. Structure of the 2- (*ash2.1*) and 1.4-kb (*ash2.2*) transcripts after performing rapid amplification of cDNA ends is shown. (*B*) Developmental Northern blot of *ash2* expression. In *A* and *B*, *rp49* was used as a loading control. e, embryo; L, larval stages; numbering indicates hours after egg laying. (*C*) Down-regulation of UBX protein accumulation in $ash2^{I1}$ mutant clones generated in haltere and leg imaginal discs by FLP-FRT-induced mitotic recombination. (*Left*) Staining with anti-β-galactosidase antibody:WT cells (bright red), heterozygous cells (red), and homozygous $ash2^{I1}$ mutant cells (lack of red staining). (*Center*) Staining with anti-UBX antibody (green). (*Right*) Merged images.

ments, Foster City, CA), CONVERTDATA 3.33 (www.le.ac.uk/cmht/microarray_lab/microarray_softwares/microarray_softwares.htm) and Microsoft EXCEL were used for acquiring and mining data. One Class Response Significance Analysis of Microarrays (SAM) Analyses (28) were conducted with different fold-change thresholds. Automatic functional annotation of the regulated genes in our microarray experiments has been obtained by using the GO database (27). Complete information on the methods used and controls carried out is published as supporting information on the PNAS web site, which also contains a comprehensive description on how GO classification, chromosomal localization, and sequence analysis were performed. Additional data and detailed PERL scripts can be accessed at www.ub.es/epidd/arrays/index.htm.

Results and Discussion

The $ash2^{I1}$ Mutant Lacks the trx-G Function. In the work presented here, we have used the allele $ash2^{I1}$, obtained after excision of the *P-lacW* transposon present in line *l(3)12411* (17). It is lethal in early pupa, and homozygous larvae have reduced and abnormal imaginal discs and brain (data not shown and Fig. 5*A*). The molecular alterations present in the $ash2^{I1}$ allele are small changes (2-bp deletion and 5-bp insertion) in the fourth intron of the gene (17). Northern blot analysis of poly(A) RNA extracted from third instar larvae showed the presence of two transcripts (2 and 1.4 kb) with potential coding sequence in WT and only the small one in $ash2^{I1}$ mutant flies (Fig. 1*A*). The longer transcript would account for the already described ASH2 protein (15), and 5′-rapid amplification of cDNA ends results supported by *in silico* predictions from GENEID (29) and GENSCAN (30) showed that the 1.4-kb transcript, identical in both WT and $ash2^{I1}$, contains exons 5–8 present in the previously described *ash2* transcript plus a novel 62-bp exon containing 28-bp encoding for amino acids (Fig. 1*A*). If translated, the resulting protein would be 350 aa in length and would lack the proline-, glutamic acid-, serine-, and threonine-rich region sequence and the putative double zinc-finger domain, also found in other trx-G proteins. Developmental Northern blot of WT flies showed that both transcripts are present at all stages of development except in early embryos, where only the long maternal transcript was detected (Fig. 1*B*). Because the insertion (TTAGG) detected in the fourth intron of the $ash2^{I1}$ allele (17) creates a putative splicing acceptor site (31) that could generate a transcript containing a premature translation termination codon, it is tempting to speculate on a nonsense mediated decay of such RNA species (32). To confirm that our mutant behaves as a true trx-G mutant, we generated genetic mosaics in haltere and leg imaginal discs with the aid of the flipase–flipase recombination target (FLP-FRT) technique and analyzed the expression of ultrabithorax (UBX) by immunohistochemistry (see supporting information on the PNAS web site). The down-regulation of UBX accumulation in the homozygous $ash2^{I1}$ tissue (Fig. 1*C*) proves this mutant behaves as a trithorax mutant regarding homeotic function.

Transcription Profile of $ash2^{I1}$ Homozygous Mutant Larvae. To identify downstream genes of *ash2* function, we have compared the population of mRNA species isolated from homozygous $ash2^{I1}$ third instar larvae with that of stage-matched WT. Four completely independent cDNA microarray experiments were carried out with poly(A) RNA isolated from separate extractions. The microarrays were constructed by using the ESTs from the DROSOPHILA GENE COLLECTION 1.0, which contains about one-third of the *Drosophila* genes. This collection lacks some genes known to be regulated by *ash2* such as Ubx. A total of 5,139 cDNAs with a different FlyBase identifier (33) were printed; 4,163 of them passed the quality filters in at least two of the experiments (see supporting information on the PNAS web site) and could be used in a One Class Response SAM analysis (28). With a false discovery rate of ≈0.025 and a fold change threshold

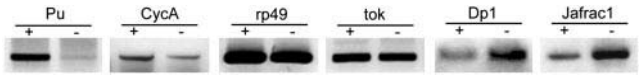


Fig. 2. Validation of the microarray analysis by RT-PCR of WT (+) and *ash2*^{11} (−) mRNA samples on selected genes. *Punch* (Pu) and *Cyclin A* (CycA) are down-regulated genes; ribosomal protein *rp49* and the endopeptidase *tolkin* (tok) are unaltered genes; *Dodeca-satellite-binding protein 1* (Dp1) and *thiredoxin peroxidase1* (Jafrac1) are up-regulated genes.

of 1.75, we identified 235 genes of the 4,163 (5.6%) in the SAM input whose expression levels change significantly in the mutant, pointing to *ash2* as a putative regulator of them (see Fig. 6, which is published as supporting information on the PNAS web site, and additional data at www.ub.es/epidd/arrays/index.htm). Before inputting data with SAM, the mean Pearson's correlation coefficient among all replicate experiments for the 235 regulated genes was 0.88, indicating a good reproducibility of ratios. One hundred forty of these genes were positively regulated and 95 negatively regulated. Down-regulated genes include *ash2* with a 1.76-fold change, a rather high value if we acknowledge the presence of the 1.4-kb transcript in the *ash2*^{11} mutant. We verified the differential expression levels of candidate genes by performing semiquantitative RT-PCR analysis on selected genes (Fig. 2).

Functional Classification of Regulated Genes. Functional annotation of the genes present in the SAM 1.75 output has been performed aiming to discover features characteristic to the sets of up- and down-regulated genes. We classified the cDNAs present in our chips according to GO database (27). GO provides a controlled vocabulary organized in a hierarchical way that can be applied to describe gene products from eukaryotic organisms regarding: molecular function (MF), biological process (BP), and cellular component (CC). GO terms were found for 55% of the 5,139 unique genes. From the 4,163 genes given to SAM as input, 56% had GO terms, and 62% of the 235 regulated genes also had them. The GO classifications of the SAM input set, the set of all genes in the chip, and the set of all genes in FlyBase (additional data at www.ub.es/epidd/arrays/index.htm) were very similar for the different categories studied (MF levels 1 and 2, BP levels 1 and 2, and CC levels 1 and 2), strongly suggesting that the SAM input set is a good representation of all known genes in *Drosophila*. To evaluate whether *ash2* was acting predominantly on a certain group of functionally related genes, we compared the GO classification of the regulated genes (up or down) with that of the whole set of genes given as input to SAM. Fig. 3 shows that the distribution of classes is quite different, both when comparing the regulated genes with the SAM input genes and when comparing up- with down-regulated genes (Fig. 3 and additional data at www.ub.es/epidd/arrays/index.htm), strongly suggesting

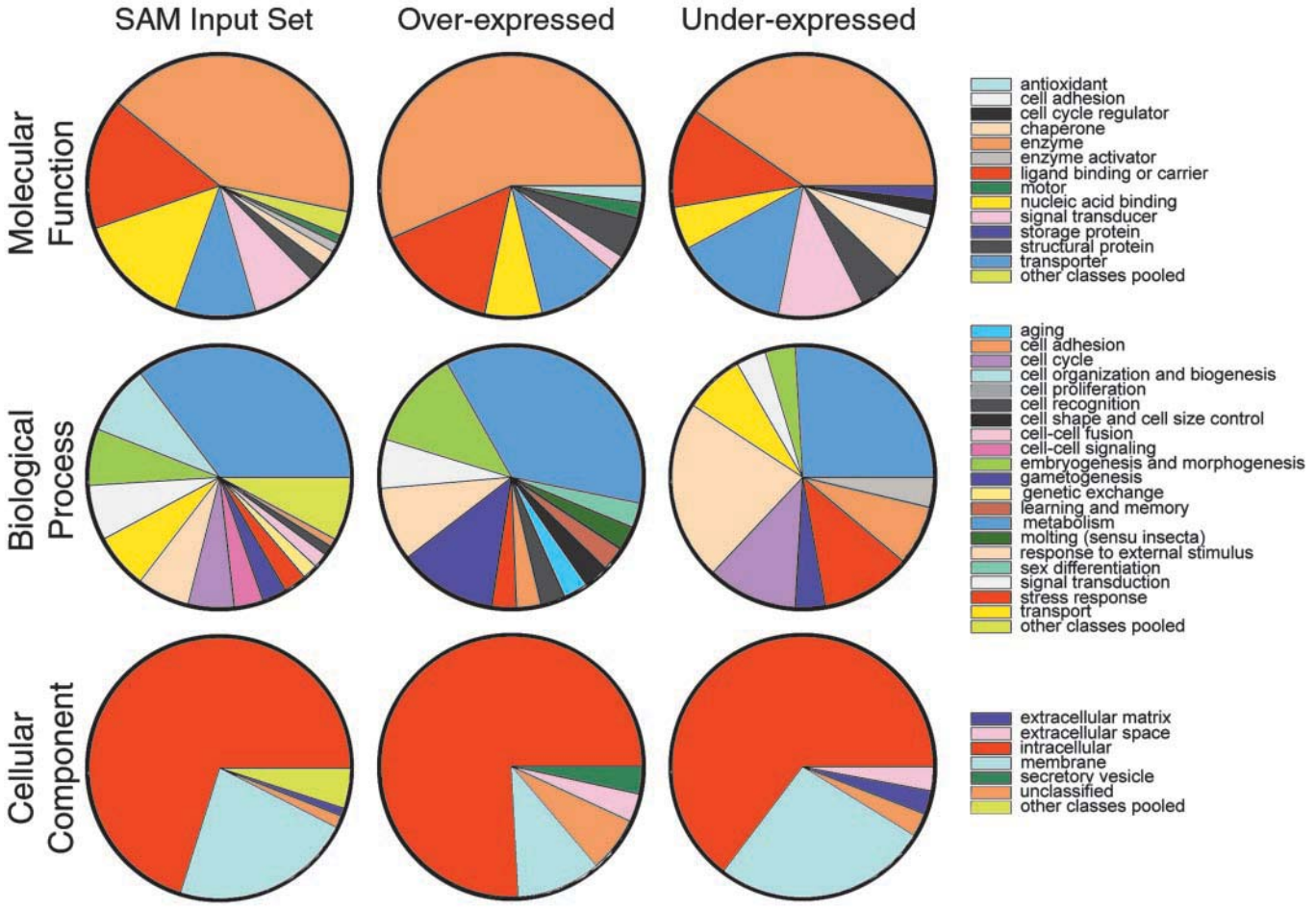


Fig. 3. Classification of regulated genes (up- and down-) according to GO (MF, level 1; CC, level 2; and BP, level 2) and comparison with the genes used for the study (SAM input set). To assess the statistical significance of differences between sets of genes, we have calculated pairwise Chi^2. Chi^2 *P* values are given in the following order for each of the classes: SAM input vs. whole fly, overexpressed vs. SAM input and underexpressed vs. SAM input. The *P* values are as follows: MF1, 0.820, 0.008, and 0.006; BP2, 0.845, 0.000, and 0.000; and CC2, 0.931, 0.002, and 0.384.

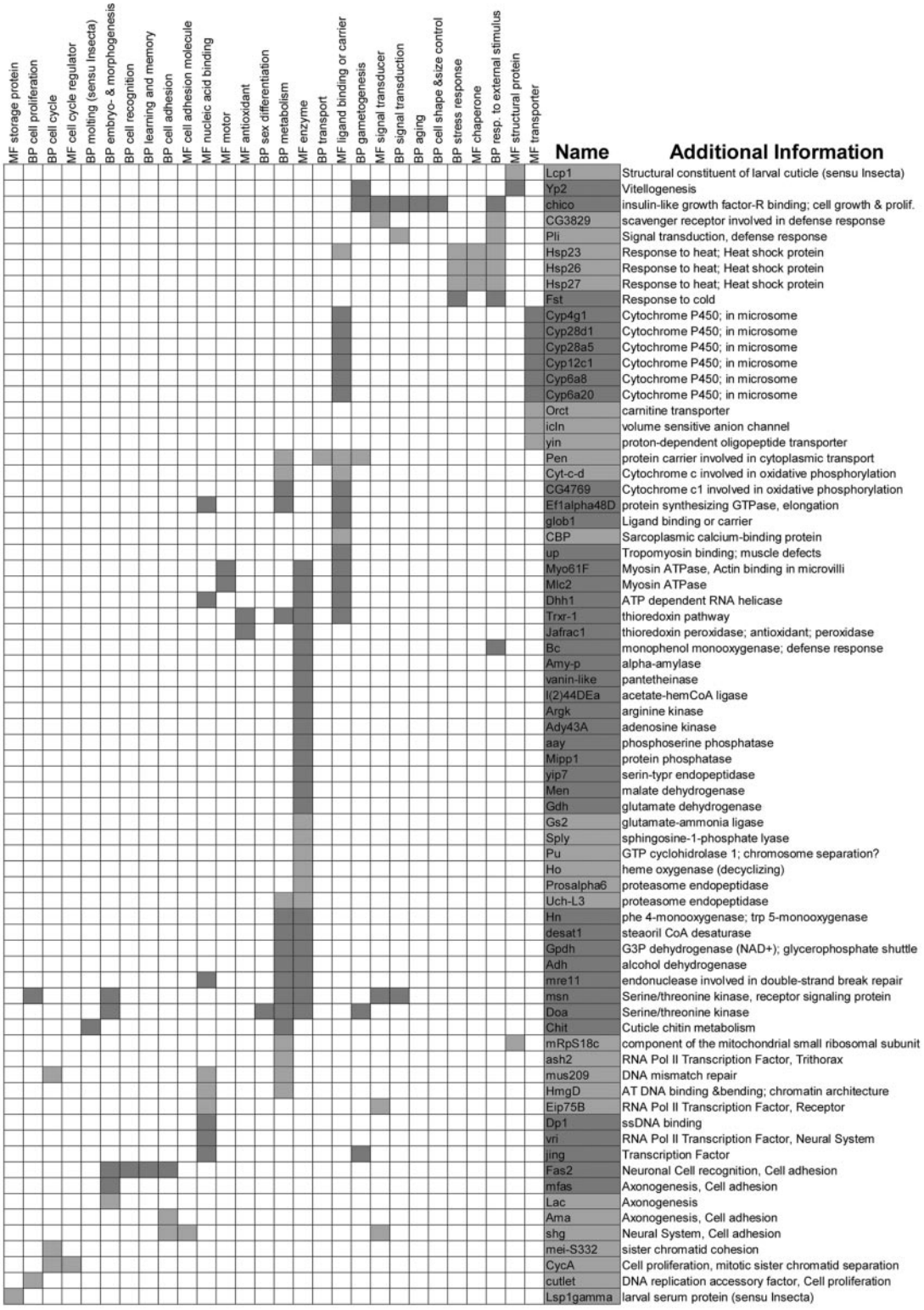


Fig. 4. Genes (rows) differentially expressed in the mutant *ash2*^{l1} clustered according to their MF level 1, and BP level 2 GO terms (columns). MF and BP descriptors are also clustered. A given gene has a GO term if the intersecting cell is dark gray (up-regulated) or light gray (down-regulated). Genes (CG nos. and BcDNAs) presenting only one GO term have been excluded. ''Additional Information'' has been extracted from FlyBase, interactive fly, and references herein. A POSTSCRIPT image with tree branches can be found at www.ub.es/epidd/arrays/index.htm.

that the results of our microarray experiment are indeed capturing underlying biological phenomena and are not purely artifactual.

In addition to distributing the SAM 1.75 set of genes among GO classes, we have hierarchically clustered them according to their GO terms (MF level 1 and BP level 2), so that genes having similar sets of GO terms are placed closer in Fig. 4. The position of the genes in this tabular representation is (broadly) associated

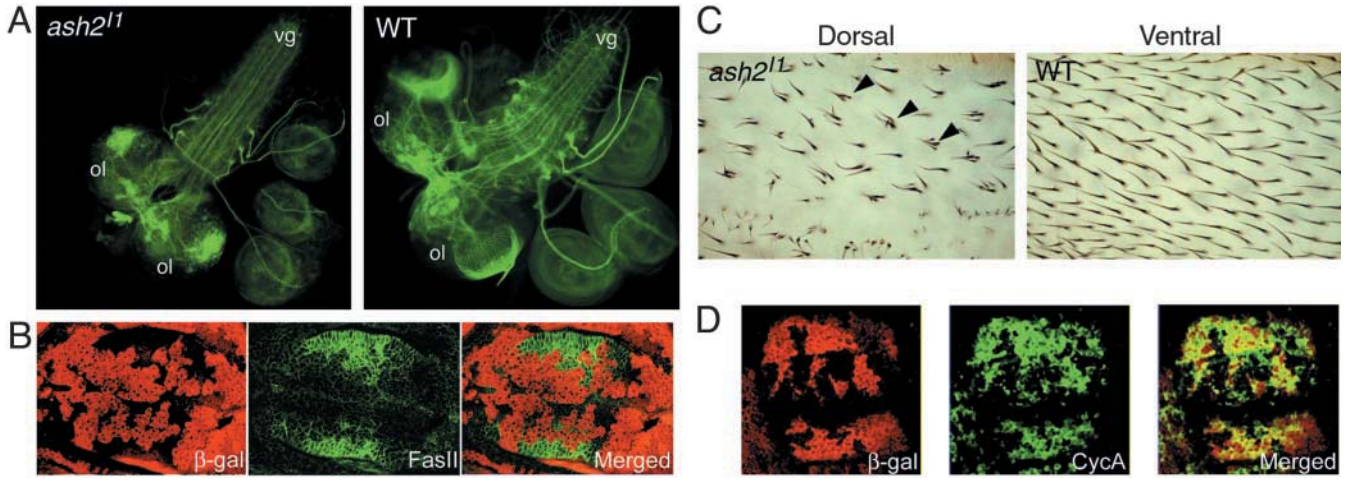


Fig. 5. *ash2* has a role in regulating cell adhesion, development of neural system, and cell cycle. (*A*) Third instar larval brains from $ash2^{l1}/ash2^{l1}$ and WT individuals stained with anti-FASII. ol, optic lob; vg, ventral ganglion. Note the disruption of the neural pattern in both the optic lobe and the comissures of the ventral ganglion. (*B*) FLP-FRT clones on *Minute* (M^-) background of $ash2^{l1}$ mutant cells from a wing imaginal disc tested for the up-regulated protein Fasciclin II (green). (*C*) Detail of a clone $ash2^{l1}/ash2^{l1}$ in the dorsal layer of cells of an adult wing (*Left*) and the same area (WT tissue) focused on the opposite wing side (*Right*). Note that in the WT condition each cell develops one hair, whereas in the dorsal mutant surface some cells develop multiple hairs (arrowheads). (*D*) Immunodetection of the down-regulated *CycA* (green) on M^- background FLP-FRT clones. (*B* and *D Left*) Staining with anti-β-galactosidase antibody in red (see Fig. 1 legend).

with its biological function. Indeed, the more GO terms share two of them (or two groups of them), the more similar they are supposed to be and the closer they will be to one another in the figure. For example, the down-regulated heat-shock proteins and the up-regulated gene *Frost* have been grouped together in Fig. 4 and are included in the classes MF chaperones and BP response to external stimulus, two of the classes that stand out in the set of down-regulated genes in Fig. 3. We believe methods like this of presenting microarray results will become more useful as more genes from *Drosophila* and other organisms are given GO terms, and as more microarray data from different eukaryotic organisms become available. By clustering genes from different sources using GO terms and microarray data, we will be able to find functional homologies between genes and/or pathways from different organisms even when there is no conservation whatsoever at the sequence level.

In Fig. 3, it is clear that regarding MF, nucleic acid-binding proteins are underrepresented in both sets of regulated genes and, whereas enzymes are overrepresented only in the set of up-regulated genes, so are chaperones in the down-regulated set. As to BP, GO classifications are quite characteristic of both up- and down-regulated genes. Genes related to cell organization and biogenesis are absent from the sets of regulated genes. The set of up-regulated genes is clearly enriched with genes involved in gametogenesis but lacks genes involved in cellular transport, whereas genes involved in responses to external stimulus, stress response, cell adhesion, and cell proliferation are overrepresented among down-regulated genes. Finally, concerning CC, regulated genes (both up- and down-) have an excess of genes acting in the extracellular space, whereas up-regulated genes show an excess of genes acting in the secretory vesicle and a deficiency of membrane proteins.

Because trx-G might be acting on chromatin structure, making DNA accessible to other proteins involved in transcriptional regulation, and some genes that have been found to be regulated by *ash2* are located adjacent or very close to each other on the genome (i.e., *Hsp23*, *Hsp26*, *Hsp27*, and *aay* or *Cyp6a8* and *Cyp6a20*), we mapped differentially expressed genes in the SAM 1.75 output set onto the *D. melanogaster* genome sequence with the aim of investigating possible colocalization of coregulated genes. To assess the degree of aggregation for coregulated genes, we computed the Moran's index (34) to measure the correlation between the localization of the genes along the chromosomes and their level of expression (see supporting information on the PNAS web site). We found no evidence of physical aggregation of coregulated genes, because the value of the Moran's index (0.01) was not significant after 1,000 Monte Carlo tests with randomly shuffled data. The same results have been obtained in whole-genome expression analysis of *snf/swi* mutants of *S. cerevisiae* (35). More recently, it has been found that chromosomal position of signature genes is random when studying the transcript profiles in aging and calorically restricted flies (36), although evidence for large domains of similarly expressed genes in the *Drosophila* genome has also been reported (37).

Finally, for 165 genes (72 down- and 93 up-regulated), complete mRNA exonic coordinates were annotated, and the corresponding full-length cDNA sequences were obtained. A number of sequence characteristics were computed on these genes (G+C content, number of exons, length of transcripts) and compared with those of the whole set of full-length cDNAs from *Drosophila* (4,002 sequences). As expected, no striking differences were found among up- and down-regulated genes and the whole set of fly genes (data not shown).

Confirmation of Target Genes. Genes involved in cell adhesion and/or development of the neural system (i.e., *FasII*, *mfas*, *Ama*, *Lac*, and *shg*) are two of the main classes regulated by *ash2* (Figs. 3 and 4). We have focused in the up-regulated gene *FasII* as an example of those and performed clonal analysis on a *Minute* background, to assess whether the behavior of the *FasII* transcript observed with this *ash2* mutant was also kept at the protein level in wing imaginal discs (Fig. 5*B*). Homozygous mutant cells show a clear up-regulation of FASII, mainly in the wing pouch area further away from the dorsoventral margin, where we have found FASII to be very slightly expressed in WT wing discs (data not shown). The up-regulation of *FasII* and other cell adhesion molecules like *mfas*, together with the up-regulation of the transcription factor *vri*, could explain some of the phenotypes previously found by clonal analysis (17), such as disruption of vein–intervein patterning, because it is known that preferential accumulation of specific adhesion molecules characterizes the final stages of vein differentiation (38). Furthermore, because

FASII is involved in the development of the neural system, we have compared its pattern of expression in *ash2*^{I1} mutant brains with that of WT and have found that they present a distorted phenotype in the optic lobes as well as fasciculation defects in the ventral ganglion (Fig. 5*A*), a process in which FASII plays a central role (39, 40). A gain-of-function screen done by Kraut *et al.* (41) identified *ash2* and *FasII* as genes involved in the development of the neural system, further supporting a relationship between them.

In a search for putative downstream genes for *Set1*, many genes involved in transcriptional regulation of growth and cell cycle control were found (12). According to our results, *ash2* also seems to act on genes that fall into these classes (see Figs. 3 and 4). For example, *Eip75B*, *vri*, *jing*, or *HmgD* could be classified in the first class, whereas *CyclinA* (*CycA*), *cutlet*, *Pu*, *mei-S332*, or *mus209* could do so in the second. We have confirmed, by clonal analysis in wing imaginal discs, that the protein product of the down-regulated gene *CycA* is present to a lesser extent in mutant clones (Fig. 5*D*). The down-regulation of *CycA* can play a role in the proliferation defects observed in mutant cells when clones are generated in the imaginal discs (17). Furthermore, in genetic mosaics, homozygous mutant *ash2*^{I1} cells show effects on both cell differentiation and cell size. Normal wing cells develop a single hair, whereas *ash2*^{I1} cells can develop either a single or multiple hairs. In addition, the spacing between *ash2*^{I1} mutant cells is increased compared with that of WT ones, suggesting that *ash2*^{I1} mutant cells are larger (Fig. 5*C*).

Preliminary promoter analysis of the regulated genes does not seem to reveal any defined and clear pattern (see supporting information on the PNAS web site). Although this needs further investigation, it may be difficult to find any consensus sequence, mainly because the regulated genes reflect all of the transcripts present at that moment and therefore a particular cellular status. Even though some authors (42, 43) have described *cis*-acting Trithorax Response Element sequences needed for some trx-G proteins to exert their function, it is not yet possible to discern between primary and secondary targets of ASH2 on the basis of this information. However, we can state that ASH2 function is required for a wide variety of biological processes during larval development.

We thank G. Hulburt for technical support, J. F. Abril for help with POSTSCRIPT, S. Vives for statistical advice, and J. Baguñà for helpful discussions. S.B. and E.B. are supported by fellowships from the Ministerio de Ciencia y Tecnología. Grants BMC2000-0766, GEN2001-4846-C05-02, and HI2001-0009 from Ministerio de Ciencia y Tecnología, Spain, have supported this work. We thank the Developmental Studies Hybridoma Bank, University of Iowa, for providing the anti-Fasciclin II antibody.

1. Simon, J. (1995) *Curr. Opin. Cell Biol.* **7,** 376–385.
2. Schumacher, A. & Magnuson, T. (1997) *Trends Genet.* **13,** 167–170.
3. Kennison, J. A. (1995) *Annu. Rev. Genet.* **29,** 289–303.
4. Mahmoudi, T. & Verrijzer, C. P. (2001) *Oncogene* **20,** 3055–3066.
5. Francis, N. J. & Kingston, R. E. (2001) *Nat. Rev. Mol. Cell Biol.* **2,** 409–421.
6. Shearn, A., Rice, T., Garen, A. & Gehring, W. (1971) *Proc. Natl. Acad. Sci. USA* **68,** 2695–2698.
7. Shearn, A., Hersperger, E. & Hersperger, G. (1987) *Roux's Arch. Dev. Biol.* **196,** 231–242.
8. Shearn, A. (1989) *Genetics* **121,** 517–525.
9. LaJeunesse, D. & Shearn, A. (1995) *Mech. Dev.* **53,** 123–139.
10. Ponting, C., Schultz, J. & Bork, P. (1997) *Trends Biochem. Sci.* **22,** 193–194.
11. Papoulas, O., Beek, S. J., Moseley, S. L., McCallum, C. M., Sarte, M., Shearn, A. & Tamkun, J. W. (1998) *Development (Cambridge, U.K.)* **125,** 3955–3966.
12. Nislow, C., Ray, E. & Pillus, L. (1997) *Mol. Biol. Cell* **8,** 2421–2436.
13. Roguev, A., Schaft, D., Shevchenko, A., Pijnappel, W. W., Wilm, M., Aasland, R. & Stewart, A. F. (2001) *EMBO J.* **20,** 7137–7148.
14. Nagy, P. L., Griesenbeck, J., Kornberg, R. D. & Cleary, M. L. (2002) *Proc. Natl. Acad. Sci. USA* **99,** 90–94.
15. Adamson, A. L. & Shearn, A. (1996) *Genetics* **144,** 621–633.
16. Deak, P., Omar, M. M., Saunders, R. D., Pal, M., Komonyi, O., Szidonya, J., Maroy, P., Zhang, Y., Ashburner, M., Benos, P., *et al.* (1997) *Genetics* **147,** 1697–1722.
17. Amorós, M., Corominas, M., Deak, P. & Serras, F. (2002) *Int. J. Dev. Biol.* **46,** 321–324.
18. Chu, S., DeRisi, J., Eisen, M., Mulholland, J., Botstein, D., Brown, P. O. & Herskowitz, I. (1998) *Science* **282,** 699–705.
19. Perou, C. M., Sorlie, T., Eisen, M. B., van de Rijn, M., Jeffrey, S. S., Rees, C. A., Pollack, J. R., Ross, D. T., Johnsen, H., Akslen, L. A., *et al.* (2000) *Nature* **406,** 747–752.
20. White, K. P., Rifkin, S. A., Hurban, P. & Hogness, D. S. (1999) *Science* **286,** 2179–2184.
21. Leemans, R., Egger, B., Loop, T., Kammermeier, L., He, H., Hartmann, B., Certa, U., Hirth, F. & Reichert, H. (2000) *Proc. Natl. Acad. Sci. USA* **97,** 12138–12143.
22. Furlong, E. E., Andersen, E. C., Null, B., White, K. P. & Scott, M. P. (2001) *Science* **293,** 1629–1633.
23. De Gregorio, E., Spellman, P. T., Rubin, G. M. & Lemaitre, B. (2001) *Proc. Natl. Acad. Sci. USA* **98,** 12590–12595.
24. Irving, P., Troxler, L., Heuer, T. S., Belvin, M., Kopczynski, C., Reichhart, J. M., Hoffmann, J. A. & Hetru, C. (2001) *Proc. Natl. Acad. Sci. USA* **98,** 15119–15124.
25. Churchill, G. A. & Oliver, B. (2001) *Nat. Genet.* **29,** 355–356.
26. Arbeitman, M. N., Furlong, E. E., Imam, F., Johnson, E., Null, B. H., Baker, B. S., Krasnow, M. A., Scott, M. P., Davis, R. W. & White, K. P. (2002) *Science* **297,** 2270–2275.
27. Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., *et al.* (2000) *Nat. Genet.* **25,** 25–29.
28. Tusher, V. G., Tibshirani, R. & Chu, G. (2001) *Proc. Natl. Acad. Sci. USA* **98,** 5116–5121.
29. Parra, G., Blanco, E. & Guigó, R. (2000) *Genome Res.* **10,** 511–515.
30. Burge, C. & Karlin, S. (1997) *J. Mol. Biol.* **268,** 78–94.
31. Pertea, M., Lin, X. & Salzberg, S. L. (2001) *Nucleic Acids Res.* **29,** 1185–1190.
32. Hentze, M. W. & Kulozik, A. E. (1999) *Cell* **96,** 307–310.
33. Anonymous (1999) *Nucleic Acids Res.* **27,** 85–88.
34. Moran, P. A. P. (1950) *Biometrika* **37,** 17–23.
35. Sudarsanam, P., Iyer, V. R., Brown, P. O. & Winston, F. (2000) *Proc. Natl. Acad. Sci. USA* **97,** 3364–3369.
36. Pletcher, S. D., Macdonald, S. J., Marguerie, R., Certa, U., Stearns, S. C., Goldstein, D. B. & Partridge, L. (2002) *Curr. Biol.* **12,** 712–723.
37. Spellman, P. T. & Rubin, G. M. (2002) *J. Biol.* **1,** 5.1–5.8.
38. De Celis, J. F. (1998) *Int. J. Dev. Biol.* **42,** 335–343.
39. Lin, D. M. & Goodman, C. S. (1994) *Neuron* **13,** 507–523.
40. Lin, D. M., Fetter, R. D., Kopczynski, C., Grenningloh, G. & Goodman, C. S. (1994) *Neuron* **13,** 1055–1069.
41. Kraut, R., Menon, K. & Zinn, K. (2001) *Curr. Biol.* **11,** 417–430.
42. Chang, Y. L., King, B. O., O'Connor, M., Mazo, A. & Huang, D. H. (1995) *Mol. Cell. Biol.* **15,** 6601–6612.
43. Tillib, S., Petruk, S., Sedkov, Y., Kuzin, A., Fujioka, M., Goto, T. & Mazo, A. (1999) *Mol. Cell. Biol.* **19,** 5189–5202.

Supporting Materials and Methods

Drosophila stocks

wt and mutant strains were maintained on standard medium and all experiments were performed at 25ºC. The WT stock used was Canton-S (Bloomington Stock Center, Indiana University, Bloomington). *ash2*^{l1}/*TM6C* mutant line was obtained from an excision of a *P-lacW* element inserted in the *ash2* gene (1), and homozygous *ash2*^{l1} larvae could be selected due to their *Tubby*$^{+}$ phenotype.

RNA analysis

Total RNA was extracted from larvae by using Trizol (GIBCO/BRL) and poly(A) by using the polyA-TRACT system (Promega). mRNA (0.5 µg) was used for the Northern blot analysis. For the developmental Northern blot, we used poly(A) RNA isolated from different stages (kindly provided by S. Campuzano, Consejo Superior de Investigaciones Científicas, Universidad Autónoma, Madrid). The *ash2* probe used for this analysis was a 659-bp PCR product (exons 5–8) obtained from the 3' end of the coding region, and a 300-bp fragment of the *rp49* gene was used as a loading control. All probes were labeled with [^{32}P]dCTP by random priming.

5'-RACE was carried out by using the SMART RACE cDNA Amplification Kit (CLONTECH) according to the manufacturer's instructions. The gene-specific primers were *ash2*- c1292 (5'-CGTGGCAGCTCCGTCGGGCA TCTCTTC-3') for the first 5'-RACE PCR and *ash2*-2R (5'-CTGACGAAATGG AGCATGTG-3') for the nested PCR. The resulting fragments of expected size were isolated from agarose gel by using the QIAquick Gel Extraction Kit (Qiagen, Chatsworth, CA) and automatically sequenced. *In silico* analysis was performed by using GENEID at

/www1.imim.es/software/geneid/index.html (28) and GENSCAN at

genes.mit.edu/GENSCAN.html (329).

Microarray construction

cDNA microarrays were constructed by using the ESTs from the *Drosophila* Gene Collection 1.0 (DGC1.0) set (www.fruitfly.org/DGC/index.htlm). cDNA clones were PCR amplified directly from the bacteria glycerol stocks in 96-well format by using specific primers from pOT2 and pBS vectors. All PCR reactions were analyzed by agarose gel electrophoresis, and DNA was ready to spot after purification of the PCR products by Ethanol precipitation. Concentration of DNA was around 1 µg/µl. Purified PCR products (1–7 kb long) were robotically spotted (Cartesian PS5500) on poly-lysine coated microscope slides (PL-25C, CEL Associates, Houston). Slides were UV cross-linked at 60 mJ in an UV Stratalinker (Stratagene) and made ready for prehybridization. To verify clone identity, some cDNAs were randomly chosen and sequenced by using an ABI Prism 377 DNA sequencer (Perkin–Elmer).

Probe preparation and hybridization

Total RNA and poly(A) were purified as described above. One to three micrograms of poly(A) RNA were labeled by reverse transcription incorporation of Amino-allyl dUTP and coupling to NHS–cyanine dye using the Atlas Glass fluorescent labeling kit (CLONTECH) following the manufacturer's specifications, with the exception of the probe purification that was performed by sodium acetate/EtOH precipitation. Ten microliters of Cy3- and Cy5-labeled probes resuspended in water were mixed with 5 µl of formamide, 5 µl of 20x SSC, 0.3 µl of 10% SDS. Probe was warmed at 90ºC for 5 min before being added to the slide. For the hybridization, Corning Hybridization Chamber (no. 2551) and Erie Scientific (Portsmouth, NH) cover slips (no. 22IX25-2-4635) were used, and slides were hybridized overnight at 42ºC, washed, and scanned by using a ScanArray 4000 (Perkin–Elmer) laser scanner.

Data acquisition and analysis

Fly Base gene number for all possible cDNAs was found by comparing the Berkeley Drosophila Genome Project Collection Table with "external-databases.txt (10-23-01)" from FLYBASE (4).

Data were acquired with GENEPIX PRO 3.0 (Axon Instruments, Foster City, CA) and filtered with the aid of CONVERT DATA 3.33 (www.le.ac.uk/cmht/microarray_lab/Microarray_Softwares/Microarray_Softwares.htm) default settings. For each experiment, the log (mF635-mB635) was plotted against the log (mF532-B532) for the filtered data, giving a mean slope for all experiments of 1.01 and a mean correlation coefficient (R^2) of 0.84; therefore, no dye bias correction was applied. Our four chips came from two different batches and did not have exactly the same genes. To assess the reproducibility of data between chips coming from the same batch and, therefore, having exactly the same set of genes, we calculated the Pearson's correlation coefficient between raw (before normalization of data) fluorescence intensities (at 532 and 635 nm) of all genes flagged as Good. The mean of these four correlation coefficients was 0.74. To estimate variability of the experimental procedure, we labeled two samples independently with Cy3 or Cy5 and hybridized a cDNA microarray. The correlation coefficient between raw fluorescent intensities F635-B635 and F532-B532 of the 3,819 spots flagged as okay was 0.97. The standardized ratios had a SD of 0.22, and 99.76% of the spots were below the 1.75-fold change, rendering this value as a good threshold to use in the One Class Response Significance Analysis of Microarrays (SAM) (5) conducted with a ~0.025 false discovery rate (FDR).

Gene ontology (GO) classification

Automatic functional annotation of the regulated genes in our microarray experiments has been obtained using the GO database (www.geneontology.org; ref. 6).

A number of scripts were written to parse the GO database and the association file (gene_association.fb from 10/10/2001) to obtain, in a recursive fashion, GO definitions for the *Drosophila* genes. Pie charts have been used to display GO classifications for different data sets (see Fig. 3 and http://www.ub.es/epidd/arrays/index.htm for details on the whole process).

To produce Fig. 4, a table was constructed where genes (rows) that had a given GO term (columns) presented a positive value in the intersecting cell. This table was used to perform an average linkage clustering with CLUSTER and TREEVIEW programs (7) that grouped together similar GO terms in one axis and genes with similar GO descriptions in the other. The *.cdt file was loaded to Microsoft EXCEL to modify color of cells according to their ratio and add additional information.

Chromosomal localization and sequence analysis

Genes from our microarray were mapped onto chromosome coordinates of the *Drosophila* genome. The National Center for Biotechnology Information distribution as of April 10, 2000, of the *Drosophila* genome was used. All annotated genes in the genomic scaffolds were mapped onto chromosomal coordinates, by matching the 120 bases at the 5' end of each gene against the assembled chromosomal sequences. We used BLASTN (8) to compare the gene to the genomic sequences.

For the set of genes from our microarray whose chromosomal coordinates were annotated, Moran's index (9) was computed to measure their aggregation in the *Drosophila* genome with respect to their level of expression in the microarray experiments. This index is a measure of correlation between the values of a spatial location attribute and a second attribute representing some specific property in a given dataset. Moran's index has been recently used to analyze the aggregation of genes in the human genome with respect to the observed mutational rates and GC content changes between human genes and their mouse orthologs (10). In our case, this index measures the correlation between the geographical localization of the genes along the chromosomes and their level of expression. As a correlation coefficient, the values for the index range from +1, meaning strong positive spatial correlation, to -1, indicating strong negative correlation, with 0 meaning a random pattern of distribution in the space. To assess the significance of the computed value for the index, 1,000 Monte Carlo

tests with randomly shuffled data were performed. The significance level is the proportion of times that the value of the index for a random permutation was larger than the computed value for the original data.

RT-PCR analysis

Total RNA (750 ng) and poly(dT)-24 (100 ng) were used for cDNA synthesis. The reaction was carried out in a final volume of 25 µl by using 5 units of Avian myeloblastosis virus–RT (Promega) and 200 units of Moloney murine leukemia virus RT (GIBCO). One microliter of the RT reaction was used for the PCR reaction, and all specific primers were designed to amplify a product of about 500 bp.

Generation of mitotic recombination clones

In imaginal discs, clones were generated by FLP-mediated mitotic recombination (11). Clones were induced in second instar larvae (50 and 60 h after egg laying, AEL) by 25-min heat shock at 37ºC. Mutant clones for *ash2*^{l1} were detected by the absence of β-galactosidase (β-gal) immunostaining in clones generated in the genotypes *hsFLP; FRT 82B arm lacZ/FRT82B ash2*^{l1} for twin-spot analysis and *hsFLP; FRT82B M(3)arm lacZ/FRT82B ash2*^{l1} for clonal analysis in a *Minute* background to generate large clones. After incubation in PBS containing 1% BSA (Sigma) and 0.1% TritonX-100 to block nonspecific binding, the larvae were immunostained on whole-mount preparations with the rabbit anti-CYCA (1:150), kindly provided by D. Glover (Cambridge University, Cambridge, U.K.), mouse anti-UBX (1:20), kindly provided by R. White (Cambridge University), mouse anti-FASII (1:4) from the University of Iowa and anti-β-gal (Cappel) antibodies. Incubation was done in blocking serum overnight at 4ºC. Secondary antibodies conjugated to Rhodamine red and FITC (Jackson ImmunoResearch) were used. The specimens were analyzed under a Leica confocal laser-scanning microscope. Mitotic recombination was also induced by x-rays from a Philipps x-ray source operating at a dose of 10 Gy (100 kV, 15 mA, 2-mm Al filter). We used *f36a, mwh f+87D M(3)w124/ash2*^{l1} to generate M^+ clones at 60-h AEL. Adult flies of appropriate genotype were dissected out and their wings mounted in lactic acid/ethanol (1:1) for microscopy.

References from Supporting Materials and Methods

1. Amorós, M., Corominas, M., Deak, P. & Serras, F. (2002) *Int. J. Dev. Biol.* **46,** 321–324.
2. Parra, G., Blanco, E. & Guigó, R. (2000) *Genome Res.* **10,** 511–515.
3. Burge, C. & Karlin, S. (1997) *J. Mol. Biol.* **268,** 78–94.
4. Anonymous (1999) *Nucleic Acids Res.* **27,** 85–88.
5. Tusher, V. G., Tibshirani, R. & Chu, G. (2001) *Proc. Natl. Acad. Sci. USA* **98,** 5116–5121.
6. Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., *et al.* (2000) *Nat. Genet.* **25,** 25–29.
7. Eisen, M. B., Spellman, P. T., Brown, P. O. & Botstein, D. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 14863–14868.
8. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997) *Nucleic Acids Res.* **25,** 3389–3402.
9. Moran, P. A. P. (1950) *Biometrika* **37,** 17–23.
10. Castresana, J. (2002) *Nucleic Acids Res.* **30,** 1751–1756.
11. Xu, T. & Rubin, G. M. (1993) *Development (Cambridge, U.K.)* **117,** 1223–237.

Additional data is available at www.ub.es/epidd/arrays/index.htm.

Supporting Figures to Main Text

Fig. 6. Down-regulated (green) and up-regulated (red) genes in the *ash2*^{l1} mutant identified after a One Class Response SAM analysis with a false discovery rate of ~0.025 and a fold-change threshold of 1.75. Numbers indicate the absolute fold-change of a given transcript when comparing *ash2*^{l1} to WT.

Additional Data can be found at http://www.ub.edu/epidd/arrays/pnas/index.htm

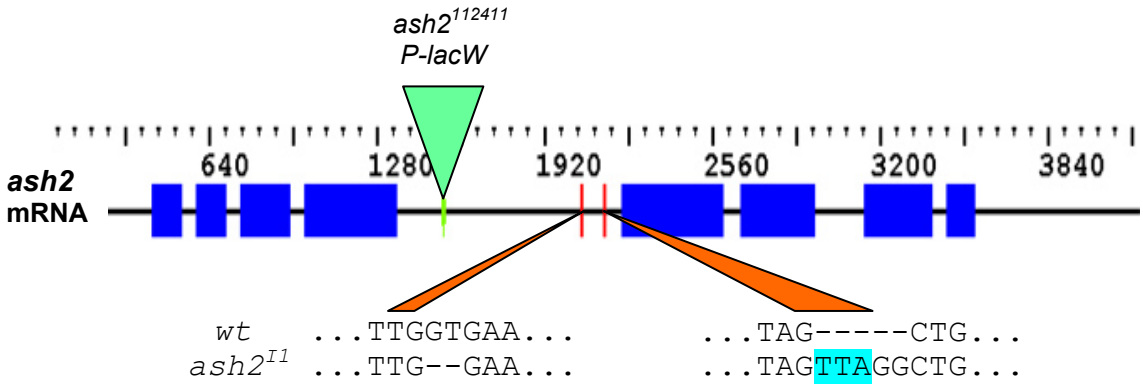
Additional Results and Discussion

Molecular and genetic analysis of *ash2*^{112411}, *ash2*^{1} and *ash2*^{l1}

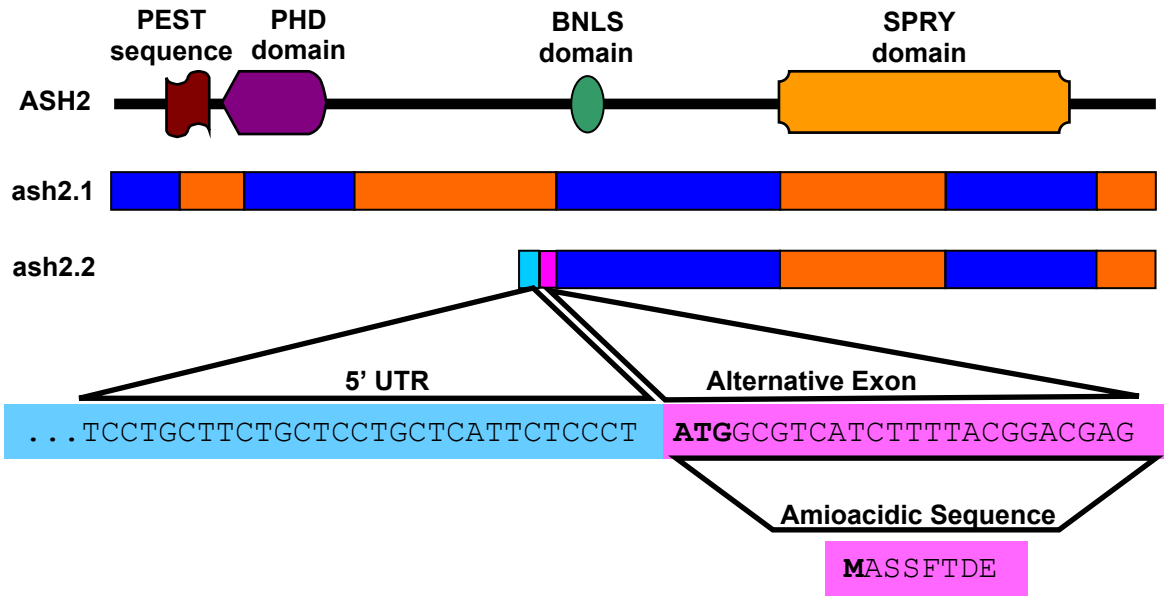
The *ash2* gene has at least two different transcripts and the lack of the ash2.1 transcript in the *ash2*^{l1} allele (Fig. 1) could be explained by a *nonsense mediated decay* due to the generation of an alternative acceptor splicing site (Supplementary Fig. 1 and Fig. 2). We have gone one step further and have compared the *ash2*^{l1} allele with two others (*ash2*^{112411} and *ash2*^{1}) by performing genetic analysis and RT-PCR. It has been previously reported that the *ash2*^{112411} allele is lethal at the pupal stage but some homozygous individuals reach adulthood. This allele fails to complement *ash2*^{l1}, but in the rare occasions in which a fly reaches the adult stage, their wings display extra cross-vein tissue as it does *ash2*^{112411}. Moreover, the *ash2*^{l1} mutant completely fails to complement the *ash2*^{1} allele (Amoros et al., 2002), an early pupal lethal with a 3 kb insertion on the 5' region of the gene (Adamson and Shearn, 1996). Although this data indicates that the *ash2*^{l1} and the *ash2*^{1} mutants are stronger than the *ash2*^{112411} allele, it is not conclusive and it does not provide information to elucidate if *ash2*^{l1} is a stronger mutant than *ash2*^{1} or vice versa. To address this issue, we have crossed *ash2*^{1} with *ash2*^{112411} to assess if the combination was viable and to analyze it's wing phenotypes if this was the case.

A high number of *ash2*^{112411}/*ash2*^{1} adults can be recovered and their wing phenotypes are stronger than those from *ash2*^{112411} homozygotes (Supplementary Fig. 3). The weaker *ash2*^{112411}/*ash2*^{1} and the average *ash2*^{112411}/*ash2*^{112411} wing phenotypes are very similar, and the strongest *ash2*^{112411}/*ash2*^{1} wing phenotype is never found in *ash2*^{112411} homozygotes. These wings are very small and display a severe reduction of intervein tissue and accumulation of vein tissue. Taken together, this data indicates that the *ash2*^{1} mutant is stronger than *ash2*^{112411} but weaker than *ash2*^{l1}.

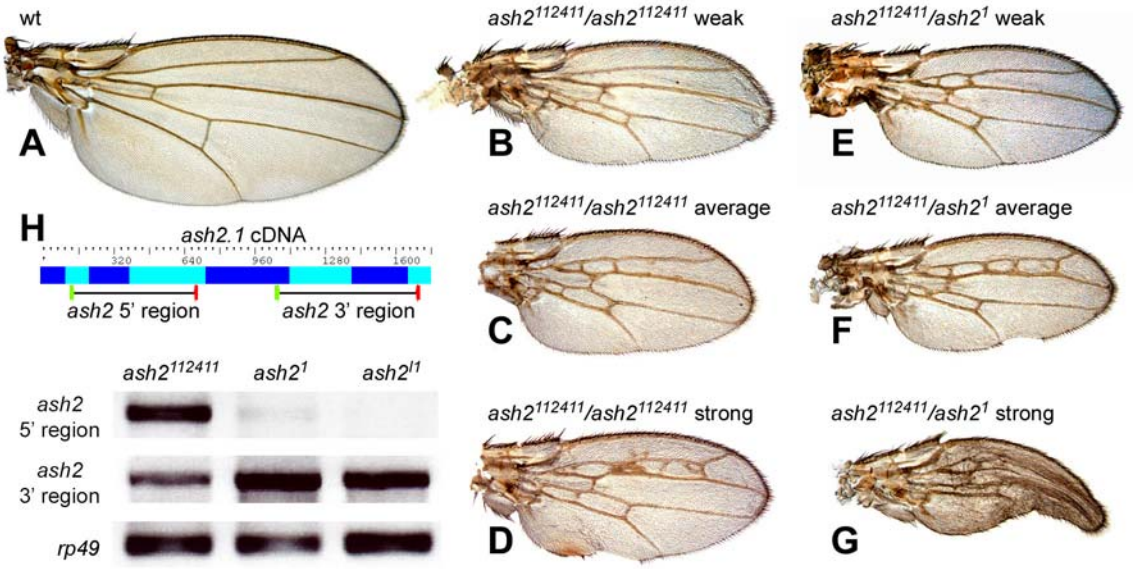
To get an insight at the molecular level, we have extracted total RNA from homozygous larvae of the three alleles to perform RT-PCRs of two distinct regions of the *ash2* cDNA (Supplementary Fig. 3). While the "5' region" amplification product only accounts for the ash2.1 transcript, the "3' region" encompasses both ash2.1 and ash2.2 transcripts. The *ash2*^{112411} mutant transcribes the "5' region" and, to a lesser extent, the "3' region". This could mean that although the transcription of ash2.1 is initiated normally, the *P-lacW* insertion detected in the biggest intron (Supplementary Fig. 1) would affect the transcription beyond this point. Since we cannot rule out that ash2.2 transcription is also affected by the presence of the nearby *P-lacW*, we cannot discern if the thinner "3' region" band is due to the presence of some ash2.1 transcript, to the ash2.2 transcript or both. Regarding *ash2*^{1} and *ash2*^{l1}, although the first has very little "5' region" product, the second has none. We conclude that this molecular data confirms the classification of the alleles obtained by genetic analyses, since it is likely that *ash2*^{112411} has reduced levels of ASH2.1 and/or ASH2.2, *ash2*^{1} has very little amounts of ASH2.1 and *ash2*^{l1} has none. This indicates that *ash2*^{l1} is a null mutant for the trxG function since, even if ash2.2 would be translated, ASH2 would lack the PEST sequence and the PHD domain, which is also found in the yeast and human ASH2 homologues (Roguev et al., 2001; Wang et al., 2001) and in other trxG and PcG proteins such as ASH1, Trx and PcL (Aasland et al., 1995; LaJeunesse and Shearn, 1995; Tripoulas et al., 1996).



Supplementary Fig. 1. The *ash2^{I1}* allele has a 2 bp deletion and a 5 bp insertion in comparison to wt or the *ash2^{112411}* allele. Dark blue boxes represent exons. In green is shown the location of the *P-lacW* insertion. Orange bars indicate the situation of the deletion and insertion identified in *ash2^{I1}*. In light blue is highlighted the first exon codon if the putative splicing site is used (score of 0.76 according to Splice View (Rogozin and Milanesi, 1997) at http://l25.itba.mi.cnr.it/~webgene/wwwspliceview.html.



Supplementary Fig. 2. Scheme of the ASH2 protein and the coding sequences of the ash2.1 and ash2.2 transcripts. Alternate blue and orange boxes indicate the different exons. In light blue is displayed the 5' UTR sequence preceding the alternative exon found by 5' RACE (pink). The sequence of this exon and the corresponding aminoacidic sequence are shown. Protein and transcript drawings are in scale: note that the protein encoded by the ash2.2 transcript would lack the PHD domain and the PEST sequence.



Supplementary Fig. 3. Comparative analysis of three *ash2* alleles. A) *wt* wing. B-D) Different wing phenotypes (weakest, average, strongest) from *ash2*^{112411} homozygotes. E-G) Different wing phenotypes (weakest, average, strongest) from the *ash2*^{112411}/*ash2*^{1} heteroal·lelic combination. Note that the weakest phenotype from *ash2*^{112411}/*ash2*^{1} is similar to the average one from *ash2*^{112411} homozygotes. All images were taken at the same magnification. H) Expression analysis of *ash2*. The scheme shows the exon structure (with alternating blue colors) of ash2.1 transcript and the two regions amplified by RT-PCR. The 5' region only amplifies the ash2.1 transcript, while the 3' region amplifies ash2.1 and ash2.2 (see Supplementary Fig. 2). Lower panel displays the RT-PCR product from each region of *ash2*^{112411}, *ash2*^{1} and *ash2*^{l1} homozygous larvae. Note that 5' region product is barely detectable in *ash2*^{1} homozygotes and not detectable in *ash2*^{l1} homozygotes. *rp49* was used as a control.

References for Additional Results and Discussion

Aasland, R., Gibson, T. J. and Stewart, A. F. (1995). The PHD finger: implications for chromatin-mediated transcriptional regulation. *Trends Biochem Sci* **20**, 56-9.

Adamson, A. L. and Shearn, A. (1996). Molecular genetic analysis of Drosophila ash2, a member of the trithorax group required for imaginal disc pattern formation. *Genetics* **144**, 621-33.

Amoros, M., Corominas, M., Deak, P. and Serras, F. (2002). The ash2 gene is involved in Drosophila wing development. *Int J Dev Biol* **46**, 321-4.

LaJeunesse, D. and Shearn, A. (1995). Trans-regulation of thoracic homeotic selector genes of the Antennapedia and bithorax complexes by the trithorax group genes: absent, small, and homeotic discs 1 and 2. *Mech Dev* **53**, 123-39.

Rogozin, I. B. and Milanesi, L. (1997). Analysis of donor splice sites in different eukaryotic organisms. *J Mol Evol* **45**, 50-9.

Roguev, A., Schaft, D., Shevchenko, A., Pijnappel, W. W., Wilm, M., Aasland, R. and Stewart, A. F. (2001). The Saccharomyces cerevisiae Set1 complex includes an Ash2 homologue and methylates histone 3 lysine 4. *Embo J* **20**, 7137-48.

Tripoulas, N., LaJeunesse, D., Gildea, J. and Shearn, A. (1996). The Drosophila ash1 gene product, which is localized at specific sites on polytene chromosomes, contains a SET domain and a PHD finger. *Genetics* **143**, 913-28.

Wang, J., Zhou, Y., Yin, B., Du, G., Huang, X., Li, G., Shen, Y., Yuan, J. and Qiang, B. (2001). ASH2L: alternative splicing and downregulation during induced megakaryocytic differentiation of multipotential leukemia cell lines. *J Mol Med* **79**, 399-405.

ARTICLE 2

Títol

Understanding phenotypes and protein interactions by functionally dissecting the *ash2* transcriptome

Autors

Sergi Beltran, Mireia Angulo, Miguel Pignatelli, Florenci Serras and Montserrat Corominas

Departament de Genètica, Universitat de Barcelona, Diagonal 645, 08028 Barcelona, Spain

Resum

El gen *absent, small or homeotic discs 2* (*ash2*) és un membre dels reguladors transcripcionals dels trxG que ha estat relacionat funcionalment amb els gens *absent, small or homeotic discs 1* (*ash1*) i *trithorax* (*trx*) mitjançant interaccions genètiques. ASH2 i ASH1 pertanyen a complexes multimèrics diferents de composició desconeguda i no es coneix amb exactitud de quina manera regulen la transcripció. L'examen minuciós dels patrons d'expressió de discs imaginals d'ala de mutants d'*ash2* i d'*ash1*, mostra que hi ha un alt grau de solapament (*overlap*) entre els dos, tant en les comparacions gen a gen com en els processos que es troben alterats. Sorprenentment, la comparació dels transcriptomes de larva de mutants d'*ash2* i de *trx* no mostren cap tipus de correlació. A més, ensenyem que la sub-expressió d'alguns d'aquests gens està directament correlacionada amb alguns dels fenotips observats en els mutants d'*ash2*. Per tal d'entendre millor el gen *ash2*, hem comparat els gens des-regulats en els seus mutants amb dades genòmiques provinents d'altres anàlisis *in vivo* i *in silico*. Una d'aquestes comparacions apunta cap a una putativa relació entre ASH2 i Sin3A, una proteïna associada a complexes desacetiladors d'histones. L'anàlisi en profunditat d'aquesta possibilitat ens ha portat a demostrar que tant ASH2 com Sin3A co-immunoprecipiten amb la proteïna Host Cell Factor (HCF), i que les tres co-localitzen en molts loci de cromosomes politènics. Aquestes evidències, juntament amb l'observació de que ASH2 co-localitza amb la tri-metilació de la H3K4, donen suport a un model en el qual ASH1 i ASH2 actuarien de forma seqüencial en la modificació de les histones per tal de mantenir els estats activats de transcripció.

Aportació personal al treball

La meva aportació en aquest article ha estat:

- participació en el disseny i la producció dels microarrays d'oligonucleòtids
- producció de controls *spike-in* a partir dels clons del TIGR
- optimització de protocols de laboratori i realització dels experiments de microarrays
- anàlisi de dades i comparació de mètodes de normalització
- comparació interal·lèlica i intergènica
- comparació entre plataformes
- anotació de resultats a partir de dades genòmiques
- comprovació de resultats per RT-PCR semi-quantitativa
- comprovació de resultats per anàlisi clonal
- estudi de la mort cel·lular amb anti-caspasa3
- generació de les construccions pAc5.1B-*ash2*-RC-V5-His i pAC5.1-*Sin3A*-HA-HIS
- participació activa en l'el·laboració de les figures i del manuscrit

Nota

Aquest article es troba actualment (Febrer de 2007) en revisió a la revista Genome Biology. Tan bon punt sigui publicat el treball serà accessible en aquesta web.