Regulation of cellular plasticity in *Drosophila* imaginal disc cells by the Polycomb group, trithorax group and *lama* genes

Ansgar Klebes[1,*], Anne Sustar[2], Katherina Kechris[1], Hao Li[1], Gerold Schubiger[2] and Thomas B. Kornberg[1,†]

[1]Department of Biochemistry and Biophysics, University of California, San Francisco, CA 94143, USA
[2]Department of Biology, University of Washington, Seattle, WA 98195, USA
*Present address: Institut für Biologie, Genetik, Freie Universität Berlin, 14195 Berlin, Germany
†Author for correspondence (e-mail: tkornberg@biochem.ucsf.edu)

Accepted 2 June 2005

Development 132, 3753-3765
Published by The Company of Biologists 2005
doi:10.1242/dev.01927

Summary

***Drosophila* imaginal disc cells can switch fates by transdetermining from one determined state to another. We analyzed the expression profiles of cells induced by ectopic Wingless expression to transdetermine from leg to wing by dissecting transdetermined cells and hybridizing probes generated by linear RNA amplification to DNA microarrays. Changes in expression levels implicated a number of genes: *lamina ancestor*, CG12534 (a gene orthologous to mouse augmenter of liver regeneration), Notch pathway members, and the Polycomb and trithorax groups of chromatin regulators. Functional tests revealed that transdetermination was significantly affected in mutants for *lama* and seven different *PcG* and *trxG* genes. These results validate our methods for expression profiling as a way to analyze developmental programs, and show that modifications to chromatin structure are key to changes in cell fate. Our findings are likely to be relevant to the mechanisms that lead to disease when homologs of Wingless are expressed at abnormal levels and to the manifestation of pluripotency of stem cells.**

Key words: Transdetermination, Cellular plasticity, Imaginal disc, Polycomb group, Trithorax group, Expression profiling

Introduction

Transplantation experiments in many different systems are the basis for a robust operational definition of the concept of determination. Such experiments show that determined cells can differentiate structures that reflect their original fate even after many cell generations or after placement in ectopic locations. Among the many examples are adult stem cells that can, after prolonged culture, give rise to progeny that will differentiate structures consistent with their assigned fate. In *Drosophila*, the heritability of the determined state was established when imaginal disc cells that had been cultivated for long periods were implanted in larvae and their differentiated products examined after metamorphosis (Hadorn, 1963). Wing disc cells differentiated wing structures, leg disc cells differentiated leg structures, etc; the stability of these determined states in these cultures was high, but it was not absolute. Some disc cells changed their determined state, a phenomenon that is defined as transdetermination (Hadorn, 1965). Such changes show that some disc cells retain the capacity to adopt alternative states. The work reported here analyzes the transcriptional phenotype of transdetermined cells, with the goal of understanding the mechanisms that underlie their pluripotency.

Developmental plasticity has been observed and studied in many systems, but it has attracted increased interest recently because of the potential medical applications of human stem cells. In *Drosophila*, the phenotype of the classic homeotic mutants, which grow normal body parts in inappropriate locations, indicates that cells choose between developmental pathways and have the capacity to follow more than one. Changing the activity of these homeotic genes late in development results in similar mutant phenotypes, indicating that cells retain the capacity to change their determined state long after they adopt a commitment. Developmental plasticity in the context of losing or activating gene functions in individual cells has been extensively characterized in this system. The significance of transdetermination is founded in the likelihood that changes of determined state have an epigenetic basis, providing access to influences that establish or are needed to maintain determined states.

There are strong parallels between the processes of homeosis and transdetermination in *Drosophila* and similar processes in mammals. Homeotic transformations of vertebrates that are comparable to the homeotic mutations in *Drosophila* have been described. One is Barrett's esophagus, which results in the formation of intestinal tissue from the esophageal epithelium (Jankowski et al., 1999). Interestingly, ectopic activation of the Wnt signaling pathway [orthologous to fly Wingless (Wg)] in the lungs of mouse embryos results in the appearance of cells in the lung that have histological and molecular characteristics of intestinal epithelial cells (Okubo and Hogan, 2004). Hyperactivation of the Wnt signaling pathway also promotes cell fate switches in epidermal and hair follicle cells (Merrill et al., 2001; Niemann et al., 2002) and in the mammary gland and prostate (Bierie et al., 2003; Miyoshi et al., 2002).

When prothoracic (1st) leg discs are fragmented and cultivated in vivo, cells in a proximodorsal region known as the 'weak point' can switch fate and transdetermine. These 'weak

point' cells give rise to cuticular wing structures (Strub, 1977). The leg-to-wing switch is regulated, in part, by the expression of the *vestigial* (*vg*) gene, which encodes a transcriptional activator that is a key regulator of wing development (Kim et al., 1996; Williams et al., 1991). *vg* is not expressed during normal leg development, but it is expressed during normal wing development and in 'weak point' cells that transdetermine from leg to wing (Johnston and Schubiger, 1996; Maves and Schubiger, 1995). Activation of *vg* gene expression marks leg-to-wing transdetermination.

Sustained proliferation appears to be a prerequisite for fate change, and conditions that stimulate growth increase the frequency and enlarge the area of transdetermined tissue (Schubiger, 1973; Schweizer and Bodenstein, 1975; Tobler, 1966). As noted above, transdetermination was discovered when fragments of discs were allowed to grow for an extensive period of in vivo culture. More recently, ways to express Wg ectopically have been used to stimulate cell division and cell cycle changes in 'weak point' cells (Sustar and Schubiger, 2005), and have been shown to induce transdetermination very efficiently (Maves and Schubiger, 1998). In the work reported here, we designed experiments to characterize the genes involved in or responsible for transdetermination that was induced by ectopic Wg. We focused on leg-to-wing transdetermination because it is well characterized, it can be efficiently induced and it can be monitored by the expression of a real-time GFP reporter. These attributes make it possible to isolate transdetermining cells as a group distinct from dorsal leg cells, which regenerate, and ventral leg cells in the same disc, which do not regenerate; and, in this work, to directly define their expression profiles. Our analysis identified unique expression properties for each of these cell populations. It also identified a number of genes whose change in expression levels may be significant to understanding transdetermination and the factors that influence developmental plasticity. One is *lamina ancestor* (*lama*), whose expression correlates with undifferentiated cells and we show controls the area of transdetermination. Another has sequence similarity to the mammalian augmenter of liver regeneration (*Alr*; *Gfer* – Mouse Genome Informatics), which controls regenerative capacity in the liver and is upregulated in mammalian stem cells. We also found that fifteen regulators of chromatin structure [e.g. members of the Polycomb group (PcG) and trithorax group (trxG)] are differentially regulated in transdetermining cells and that mutants in seven of these genes have significant effects on transdetermination. These studies identify two types of functions that transdetermination requires – functions that promote an undifferentiated cell state and functions that re-set chromatin structure.

Materials and methods

Fly stocks and in situ hybridization

Transdetermination was induced in *y,w,hs-flp;* vgBE-*GAL4/UAS-GFP, Actin^{5C},FRT,y^{+},FRT,wg^{+}* larvae in which Wg was overexpressed in over 90% of the disc cells after heat shock (Maves and Schubiger, 1995). *lama^{410}* (Perez and Steller, 1996), *Su(z)2^{1}* (antimorph) (Adler et al., 1989), *Su(z)2$^{1.b7}$* (Soto et al., 1995), *E(z)^{61}* (both loss of function) (Jones and Gelbart, 1990), *Pc^{1}* (amorph) (Gindhart and Kaufman, 1995), *E(Pc)^{1}* (Lindsley and Zimm, 1992), *Scm^{D1}* (loss of function) (Bornemann et al., 1998), *osa^{2}* (hypomorph) (Gindhart and Kaufman, 1995), *brm^{2}* (Gindhart and Kaufman, 1995) and *E(Pc^{1})* were analyzed in the transdeterimination studies. Expression profiles were determined for male larvae that had been fed 1% bromphenol blue to stage wandering stage late third instar larvae. In situ hybridization was carried out as described previously (O'Neill and Bier, 1994) using full-length cDNAs (BDGP) or genomic fragments generated by PCR amplification using specific primer pairs (Incyte Genomics).

Microarray data is available at http://www.ncbi.nlm.nih.gov/projects/geo/index.cgi and can be accessed using the series ID number GSE2886.

Induction of transdetermination and culture of discs

Ectopic Wg expression was induced by heat shock either at 60 hours or 72 hours after egg deposition, and leg discs were isolated 3 days later. Wg overexpression delays metamorphosis by 1 day, so these larvae had not yet pupariated. For the regeneration experiments, the '3/4 fragment' of 2nd leg discs were isolated and cultured in the abdomen of adult female flies for 3-5 days.

Sample isolation and RNA amplification

Leg discs were dissected in PBS and either used as whole discs or were cut with tungsten needles to isolate the three cell populations: GFP-expressing TD, D_{Wg} and V_{Wg}, as indicated in Fig. 2A. Linear RNA amplification was performed essentially as described (Klebes et al., 2002). In brief, RNA was extracted from single imaginal discs or from pooled fragments from four to eight discs using the Mini RNA isolation kit (Zymo Research). Total RNA was used for a first round of reverse transcription and in vitro transcription using T7 RNA polymerase. Subsequently, the amplified RNA product was subjected to a 2nd round of reverse transcription and in vitro transcription yielding 10,000-fold amplification or more.

Production of microarrays, hybridization and composition of reference sample

Amplified RNA was indirectly labeled by a reverse transcription reaction in the presence of amino-allyl-modified dUTP (Klebes et al., 2002) (www.microarrays.org). Fluorescent dyes (Cy3 and Cy5, Amersham) were coupled to the modified nucleotides. Data was collected using a GenePix scanner 4000A (Axon Instruments). Microarrays were produced as described (Xu et al., 2003). In brief, 14,151 specific primer pairs (Incyte Genomics) were used for PCR amplification of 100-600 bp long fragments of annotated open reading frames. The common reference sample to which we compared control leg imaginal discs, Wg-induced whole leg imaginal discs and all experimental samples of the regeneration group was generated by pooling amplified RNA of male and female 1st, 2nd and 3rd leg imaginal discs in equal proportions. For about half of the experiments, dyes were reversed to avoid bias (see details in Table S1).

Data analysis

Scanned images were processed using GenePix software (Axon Instruments). Signal intensities were further processed as expression ratios ($\log_2$ transformed). Global median normalization was performed on NOMAD (www.microarrays.org). Genes with expression levels smaller than 350 as the sum of medians were not included. Cluster analysis was performed with Cluster software and visualized with Treeview (Eisen et al., 1998) with the filtering settings detailed in Fig. 3. The median expression ratios of replicate experiments was calculated in Excel (Microsoft). To determine the median expression, replicate experiments were filtered threefold: (1) data for a given gene had to be present in at least 60% of all replicate experiments; (2) the *P*-value (heteroscedastic, two-tailed student's *t*-test) had to be less than 0.05 (95% confidence level); and (3) the median expression level had to be greater than 1.85-fold (0.8, $\log_2$). To correct for multiple testing the comparisons of wild-type wing and leg discs and the Wg-induced TD, D_{Wg}, and V_{Wg} cells were subjected

to the significance analysis of microarrays software package (Tusher et al., 2001) and all genes from the median-based lists were determined to be significant.

Analysis of PcG and trxG expression levels was performed with a two-sided *t*-test using the limma package in the statistical software R (Ihaka and Gentleman, 1996). A 'moderated' t-statistic was calculated to account for small sample sizes and differences in variability of expression values between genes (Smyth, 2004). To correct for multiple testing, the *P* values from the *t*-test were adjusted by controlling the false discovery rate (Benjamini et al., 2001). Ratios within a 95% confidence level and a median ratio greater 0.25 ($\log_2$) were considered significant. Fifteen out of 32 (47%) of the genes satisfy the two conditions. Among all of the genes in these experiments (n=11.952), 14.4% satisfy these conditions (binomial test, P-value=1.140×10^{-5}).

Results

Experimental design

Strong overexpression of Wg produces ectopic growth in 1st leg discs. Previously, we have shown that this response to Wg overexpression is region specific (Maves and Schubiger, 1998; Sustar and Schubiger, 2005). Cells in the ventral part of the disc (V_{Wg} cells) produce only structures characteristic of this region – 'fate map structures'. Cells in the dorsal part respond in two ways. Dorsal cells in the 'weak point' (TD cells) transdetermine to wing; dorsal cells surrounding the transdetermined region (D_{Wg} cells) regenerate a new posterior compartment. These two dorsal populations can be distinguished in the disc with the use of a reporter transgene that expresses GFP under the control of the *vg* boundary enhancer (vgBE) (Kim et al., 1996; Williams et al., 1991). Leg-to-wing transdetermination depends upon *vg* expression driven by the vgBE and expression of the vgBE-GFP reporter clearly marks transdetermining leg cells (Maves and Schubiger, 1998). Seeking to identify genes expressed in TD cells, we overexpressed Wg just after the molt to the third instar, and 3 days later isolated leg discs prior to pupariation. Probes generated from mRNA by linear amplification were applied to DNA microarrays.

In order to distill the genes involved in transdetermination from the expression profiles of more than 14,000 genes of TD cells, we generated nine different categories of samples for control and comparative purposes (see Table S1): (1) wild-type wing (W); (2) leg (L) discs of 3rd instar larvae; (3-6) dorsal and ventral cells from both wild-type (D_{WT} and V_{WT}) and Wg-expressing (D_{Wg} and V_{Wg}) discs; (7) intact Wg-induced 1st leg discs; (8) micro-dissected TD cells from these discs (TD); and (9) fragments of 2nd leg discs that were cultivated to obtain a population of regenerating cells. Last, we prepared a 'reference sample' of wild-type leg discs (see Materials and methods). DNA arrays were hybridized to pairs of probes that had been generated from the same larva or, when this was impractical, to the common reference sample.

Expression profiles of wild-type wing and leg imaginal discs

Probes prepared from wing and leg discs were hybridized together to arrays and the median expression ratios were compared. Using a stringent filter setting (see Materials and methods), 67 wing-specific and 62 leg-specific transcripts were identified (Table 1, see Table S2 in the supplementary material). Genes in the wing cluster previously shown to be expressed most abundantly in wing discs included *collier*, *apterous* and *vg* (Diaz-Benjumea and Cohen, 1993; Vervoort et al., 1999; Williams et al., 1991). In situ hybridization with probes for transcripts from five genes in the wing list and four genes in the leg list detected expression in leg and wings discs that was consistent with the array experiments (Fig. 1). Interestingly, a significant proportion of the genes in the wing and leg lists encode transcription factors. Among the 20 genes with the highest levels of differential expression, seven (35%) in the wing list and nine (45%) in the leg list encode transcription factors. To put this result in context, ~5% (715/14,113) of the genes in the annotated database are predicted to encode transcription factors (Adams et al., 2000). If the wing- and leg-specific transcription factors have multiple targets, most of their targets apparently are not differentially expressed with comparable specificity.

To identify genes that are normally expressed in the dorsal cells of wild-type or Wg overexpressing 1st leg discs, we also determined the expression profiles of D_{WT} and D_{Wg} cells. We isolated D_{WT}, D_{Wg} V_{WT} and V_{Wg} cells by micro-dissection (the Wg-expressing discs were marked with vgBE-GFP and the D_{Wg} cells represented the non-GFP expressing D cells; Fig. 2A), and performed pairwise comparisons (see Table S1 in the supplementary material). Among genes that were expressed 1.8-fold more in wild-type discs, *wg* (Baker, 1988), *Wnt6* (Janson et al., 2001) and *Dfz3* (Sivasankaran et al., 2000) segregated to the V_{WT} list and *decapentaplegic* (*dpp*) segregated to the D_{WT} list (Raftery et al., 1991), as expected (see Table S3 in the supplementary material). Genes encoding transcription factors were prevalent among those whose expression differed most in D_{WT} cells (9/40, 22.5%).

Expression profiles of regenerating cells

Expression profiles characteristic of regenerating cells were obtained by analyzing disc fragments that had been cultured to promote proliferation. To minimize possible contamination with transdetermined cells, we did not use 1st leg discs, but instead examined the '3/4 fragment' of 2nd leg discs. This fragment regenerates missing parts, and in contrast to 1st leg discs, does not transdetermine (G.S., unpublished). The disc fragments were collected after 3-5 days of culture, a period that corresponds to the stage when transdetermined cells express the vgBE-GFP marker and the stage when we analyzed TD cells of 1st leg discs. Transcripts enriched in regenerating disc fragments encoded proteins involved in protein synthesis, cytoskeletal organization and energy metabolism (see Table S4 in the supplementary material), an array of functions that is consistent with the increased mitotic activity of proliferating cells. Seven genes out of the 130 (5.4%) that were upregulated at least 1.8-fold in the regenerating cells encode transcription factors.

We call attention to two genes in the regeneration cluster. *regucalcin* is expressed in hemocytes associated with wing but not in leg discs (Fig. 1E). Its expression has been detected at sites of wound healing (Vierstraete et al., 2004); its elevated expression in regenerating cells (sixfold) may suggest that hemocytes are recruited to sites of regeneration as well. *headcase* (*hdc*) was also upregulated in the regenerating fragments (7.3-fold). *hdc* is expressed in all imaginal discs during normal development and has been characterized as a

Table 1. Genes with transcripts enriched in wild-type wing and leg, and in transdetermining leg disc cells

CG number	Synonym	Fold	Function
Wing			
CG10197*	*knot/collier*	45.25	RNA polymerase II transcription factor activity
CG3830*	vesitigial†	40.39	Wing margin morphogenesis/transcription factor activity
CG10619*	*tailup*	16.37	RNA polymerase II transcription factor activity
CG5966	CG5966	15.01	Triacylglycerol lipase activity
CG8376*	*apterous*†	10.84	Zinc ion binding; RNA polymerase II transcription factor activity
CG4382	CG4382	8.98	Carboxylesterase activity
CG1897*	*Drop*	8.03	RNA polymerase II transcription factor activity
CG12843	*Tetraspanin 42Ei*	7.96	Tetraspanin
CG12287*	*POU domain protein 2*†	6.66	DNA binding; RNA polymerase II transcription factor activity
CG7160	CG7160	6.07	Not known
CG10570	CG10570	5.97	Not known
CG1803	*regucalcin*†	5.61	Anteroposterior axis specification; Ca^{2+}-mediated signaling
CG9023	*Drip*	5.60	Water transporter activity; carrier activity; water channel activity
CG2663	CG2663	5.47	Carrier activity; tocopherol binding
CG3132	*Ect3*†	5.37	β-Galactosidase activity
CG4914*	CG4914†	5.12	Transcription factor activity; trypsin activity; serine-type endopeptidase activity
CG9554	*eyes absent*	5.11	Hydrolase activity; protein tyrosine phosphatase
CG9427	CG9427	4.97	Not known
CG9623	*inflated*	4.92	Protein binding;receptor activity; cell-adhesion molecule binding
CG10501	*α methyl dopa resistant*	4.88	Carboxy-lyase activity; aromatic-L-amino-acid decarboxylase activity
Leg			
CG7807*	*AP-2*	102.18	RNA polymerase II transcription factor activity
CG5893*	*Dorsal*	42.28	DNA bending activity; transcription factor activity
CG11922*	*forkhead domain 96Cb*†	25.69	Transcription factor activity
CG6269*	*unc-4*†	22.61	Transcription factor activity
CG6570*	*ladybird late*	22.05	RNA polymerase II transcription factor activity
CG10382	*wrapper*	21.26	Axonogenesis; cell-cell adhesion; ectoderm development; gliogenesis; neurogenesis; signal transduction
CG3388*	*gooseberry*	18.69	RNA polymerase II transcription factor activity
CG6414	CG6414	15.27	Carboxylesterase activity
CG11354*	*Lim1*†	13.04	Transcription factor activity
CG18111	*Odorant-binding protein 99a*	12.18	Odorant binding
CG11921*	*forkhead domain 96Ca*	8.33	Transcription factor activity
CG4605	*Accessory gland-specific peptide 32CD*	7.52	Hormone activity
CG5888	CG5888	7.26	Transmembrane receptor activity
CG4501	*bubblegum*	6.02	Long-chain-fatty-acid-CoA ligase activity
CG6604*	H15	4.48	Transcription factor activity
CG10440	CG10440	3.79	
CG2056	CG2056	3.73	Trypsin activity;serine-type endopeptidase activity
CG17131	*SP71*	3.72	
CG9747	CG9747	3.63	Acyl-CoA delta11-desaturase activity
CG1004	*rhomboid*†	3.50	Serine-type peptidase activity; receptor binding

Continued on next page

negative regulator of terminal differentiation (Weaver and White, 1995). Suppression of terminal differentiation may be an essential step during regeneration. As described in the next section, analysis of the *lama* gene provides support for this suggestion.

Expression profiles of transdetermining cells

We used two methods to analyze the transcripts in TD cells. The first entailed direct comparisons of the expression ratios. We compared GFP-positive TD cells with D_{Wg} and V_{Wg} cells that we dissected from the same 1st leg discs. One-hundred and forty-three 'TD' genes whose expression is enriched in TD cells were identified (Table 1; see Table S5 in the supplementary material). Of these genes, 19 are also upregulated in either dorsal cells (D_{Wg} + D_{wt}) or regenerating cells. Fifteen genes are also upregulated in wing cells (see Table S5 in the supplementary material). The 109 genes in the TD set that are not characteristic of either dorsal, regenerating or wing disc cells are implicated in transdetermination.

We also analyzed the expression profiles of TD cells by hierarchical clustering (Eisen et al., 1998), a method that groups genes with expression levels that change in similar ways. Clustering analysis confirmed that TD cells have a distinct expression profile (see Fig. 3; see Tables S7 and S8 in the supplementary material). The threshold settings for the two methods of analysis were different, but the majority of TD genes (66%) were also grouped together by the clustering routine. Among the many genes identified by this analysis, we focus this description on the following.

CG14059

CG14059 is the gene in the transdetermination list whose expression differed most dramatically (26.7-fold as a median of eight TD-to-D_{Wg} and TD-to-V_{Wg} comparisons; Table 1). In

Table 1. Continued

CG number	Synonym	Fold	Function
Transdetermination			
CG14059	CG14059†	26.70	
CG5993	*unpaired* (*outstretched*)†	12.27	Receptor binding; morphogen activity;cytokine activity
CG3830*	*vestigial*†	10.04	Wing morphogenesis
CG6816	*Cyp18a1*	9.45	Electron transporter activity; oxidoreductase activity
CG15279	CG15279	7.42	Cation:amino acid symporter activity;neurotransmitter:sodium symporter
CG4746	*mab-2*	7.30	Encodes Mab-21, involved in cell fate determination
CG30445	CG3686	7.12	Aromatic-L-amino-acid decarboxylase activity; amino acid metabolism; transmission of nerve impulse
CG8394	CG8394	6.03	γ-Aminobutyric acid transporter activity
CG2198	*Amalgam*	5.22	Antigen binding
CG8404*	*Sox15*	5.12	Transcription factor activity
CG5518	*slamdance*	5.04	Membrane alanyl aminopeptidase activity
CG11822	*nAcRbeta-21C*	4.69	Nicotinic acetylcholine-activated cation-selective channel activity
CG4914*	CG4914†	4.64	Trypsin activity;serine-type endopeptidase activity
CG3359	*midline fasciclin*	4.31	Axonogenesis; cell-cell adhesion; ectoderm development; signal transduction
CG9307	CG9307	4.31	Chitinase activity; hydrolase activity; hydrolyzing N-glycosyl compounds
CG6906	CG6906	4.24	Carbonate dehydratase activity
CG4859	*Matrix metalloproteinase 1*	4.17	Metalloendopeptidase activity; structural molecule activity
CG7722	CG7722	4.10	Serine-type endopeptidase inhibitor activity
CG1897*	*Drop*	3.88	RNA polymerase II transcription factor activity

Nine independent pair-wise comparisons of wing discs to leg discs from single 3rd instar larvae were performed (see Table S1 in the supplementary material). The median of the expression ratios were ranked and subjected to a *t*-test (see Table S2 in the supplementary material). The name and function of the 20 genes with highest ratios (fold) and a confidence level greater than 95% are listed. To identify genes with elevated expression in TD cells, GFP-positive cells were micro-dissected from Wg-induced 1st leg imaginal discs and the remainder of the disc was cut approximately along the DV boundary (see Fig. 2A). This procedure generated three cell populations: TD, D_{Wg} and V_{Wg}. For each preparation, four to eight discs were dissected. The fold induction in TD cells was calculated as the median of three TD-to-D_{Wg} and five TD-to-V_{Wg} comparisons (see Table S5 in the supplementary material).

*Genes encoding transcription factors.

†Genes referred to in the text.

situ hybridization confirmed these differences, detecting CG14059 RNA only in the transdetermined region of Wg-expressing leg discs (Fig. 2B), but not in either wild-type leg or wing discs (not shown). CG14059 is predicted to encode a novel protein that shares 77% sequence identity with a conserved ortholog in *D. pseudoobscura*.

unpaired (*upd*; *outstretched* – FlyBase; CG5993)

upd expression is significantly upregulated in TD cells (12.3-fold; Table 1). It encodes a ligand that activates the JAK/STAT signaling cascade (Harrison et al., 1998). Two aspects are consistent with a role for *upd* in the plasticity of the TD cells. First, Upd regulation of the JAK/STAT pathway is essential for suppressing differentiation and for promoting self-renewal of stem cells in the *Drosophila* testis (reviewed by Hombria and Brown, 2002). Second, *upd* interacts genetically with *hdc* (Bach et al., 2003). Although we have not tested whether *upd* mutants affect transdetermination, the enhanced expression of *upd* suggests that the JAK/STAT pathway might play an important role.

apterous (*ap*; CG8376)

Ap is a LIM-homeodomain-containing protein whose function is essential to wing development (Cohen et al., 1992). It is expressed in normal leg discs in the presumptive cells of the 4th tarsal segment (Pueyo et al., 2000), and is expressed at higher levels in wing discs (10.8-fold; Table 1), where it functions as a selector gene in dorsal cells (Diaz-Benjumea and Cohen, 1993). *ap* expression was marginally enhanced in TD cells, and in our cluster analysis, *ap* segregated with the genes that were upregulated in the wing disc. Anti-Ap antibody did not detect protein in Wg-induced discs, but robust staining was observed in disc fragments that had been cultivated in vivo (Fig. 2E). Antibody staining was detected in the same region as vgBE-GFP expression, indicating that transdetermined cells express Ap. Although transdetermined cells induced by Wg expression or fragmentation share many genetic and cytological features (Johnston and Schubiger, 1996; Sustar and Schubiger, 2005), the differences in *ap* expression indicate that they are not identical.

lim1 (CG11354)

Lim1 is another LIM-homeodomain containing protein that is expressed in normal leg discs in a region that is distal to the *ap* expression domain. Lim1 functions in concert with Ap to specify distal leg development (Pueyo et al., 2000). Our array analysis showed that *lim1* expression is leg-specific (13-fold; Table 1) and is up-regulated in D_{Wg} cells (5.1-fold; not included in Table S3 because of the insufficient number of data points). D_{Wg} cells are adjacent to the TD cells, and expression of *lim1* in Wg-expressing discs mimics the adjacent expression domains of *ap* and *lim1* in wild-type leg, and suggests a functional interaction between the TD cells and the adjacent leg cells (Fig. 2F,G).

CG4914

CG4914 is predicted to encode a protein that may have serine protease, DNA-binding and/or transcription factor activities. Array hybridization showed a 4.6-fold enrichment in transcript levels in wing-to-leg comparisons (Table 1), and in situ hybridization corroborated the higher level of expression in

normal wing discs (Fig. 1D). Its designation as a marker of cells that switch to wing fate was validated by showing that its expression in *wg*-expressing leg discs was most abundant in the transdetermining region (Fig. 2C).

CG12534

Expression of CG12534 was enriched 2.3-fold in TD cells (see Table S5 in the supplementary material). Although expression of CG12534 was also enriched 3.8-fold in three experiments analyzing regenerating 2nd leg discs, it did not pass the filter settings for inclusion in the regeneration list (no hybridization signal was obtained in one of four replicate experiments). CG12534 has sequence homology with the mouse gene augmenter of liver regeneration (*Alr*). Both CG12534 and *Alr* encode conserved ERV1 domains (Lisowsky et al., 1995). ALR has been implicated as a growth factor that contributes to the regenerative capacity of mammalian liver (Hagiya et al., 1994) and pancreas (Adams et al., 1998), and expression of the *Alr* gene has been found to be common to mouse embryonic stem cells, neural stem cells and hematopoietic stem cells (Ramalho-Santos et al., 2002). CG12534 is the only gene we found that TD cells have in common with both regenerating cells and the three types of mouse stem cells. We examined CG12534 expression by in situ hybridization. In

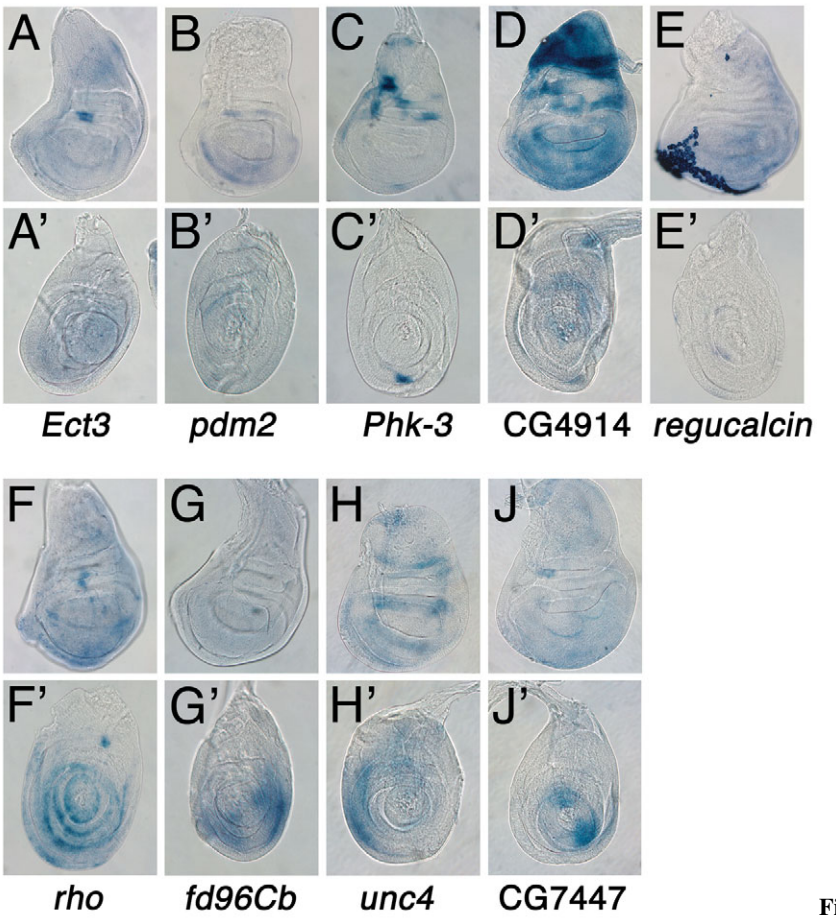

Fig. 1. Whole-mount in situ hybridization to imaginal discs. Wing and leg discs were hybridized with probes derived from genes predicted to have elevated expression in wing discs (A-J) or leg discs (A′-J′). (A) *Ect3* probe stained distinct cell clusters in the presumptive dorsal hinge region of the wing disc; (B) *pdm2* probe labeled cells surrounding the pouch with preferential expression in anterior cells; (C) Phk-3 probe labeled several domains in the hinge and notum region; (D) CG4914 expression is enriched throughout the wing disc with highest levels in the dorsal-most region. (E) *regucalcin* probe produced signal in a population of hemocytes that are associated with the wing imaginal disc, but not the leg imaginal discs (E′). Except for Phk-3, which is expressed in a distinct cell cluster in a ventral position of leg discs, these 'wing-specific' genes did not hybridize at detectable levels to leg discs (A′,B′,D′,E′). *rho*, *fd96CB*, *unc-4* and CG7447 generated more intense staining in leg discs (F′,G′,H′,J′). *rho* and *unc-4* are also detected in wing discs (F,H), whereas no signal above background could be detected for *fd96Cb* and CG7447 in wing discs (G,J). Dorsal is upwards for all discs; anterior is leftwards.

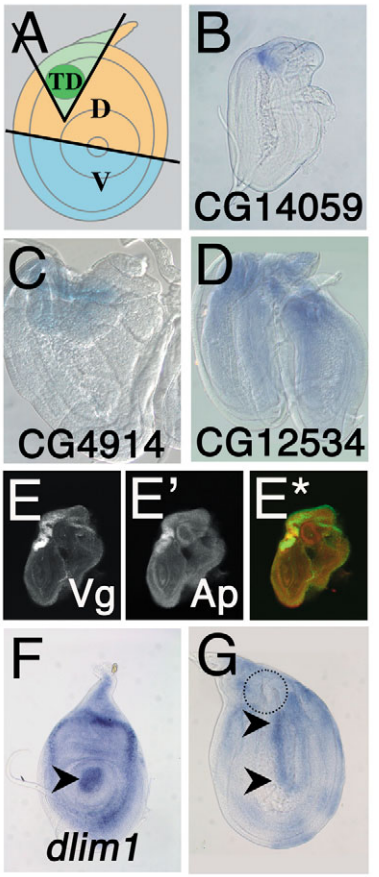

Fig. 2. Novel markers of transdetermining cells in Wg-expressing leg discs. (A) Schematic representation of a leg disc and the cuts performed to isolate transdetermining (TD, green), dorsal (D, D_{WT}, D_{Wg}, orange) and ventral cells (V, V_{WT}, V_{Wg}, blue). The area of the weak point is indicated by the dark-green circle. (B,C,D,F,G) In situ hybridization in *wg*-overexpressing leg discs. CG14059 (B), CG4914 (C) and CG12534 (D) are expressed in the area of the weak point. (E) A cultivated 3/4 fragment of a wild-type 1st leg disc. In the area of the weak point, Vg (E) and Ap protein (E′) are both expressed in the dorsal region (E*, merged image: Vg, green; Ap, red) that is similar to the expression domain of CG14059; the expression domain of CG12534 is broader. Expression of *lim1* RNA in normal (F) and a Wg-expressing leg disc (G). The distal expression domain of *lim1* expands in the Wg-expressing disc (arrowheads) but is absent from an adjacent proximodistal region (outlined). Wg expression causes overgrowth of the dorsal region.

wild-type wing and leg discs, no expression was detected (not shown). However, in leg discs that ectopically express Wg, we confirmed that CG12534 is expressed most abundantly in vgBE-GFP-expressing TD cells (Fig. 2D). We suggest that CG12534 may encode an evolutionarily conserved function in the regenerative process. Functional tests with knock-down mutants are in progress.

lama (CG10645)

lama transcripts were highly enriched in leg disc fragments containing TD cells (see Table S5 in the supplementary material). *lama* encodes a novel protein that is expressed by *Drosophila* neural and glial progenitors prior to, but not after, differentiation (Perez and Steller, 1996). A database search revealed that *lama* is evolutionarily conserved in mouse, human, *C. elegans* and Dictyostelium (see Table S6 in the supplementary material). Despite the high degree of conservation, loss-of-function *lama* mutants are viable, fertile and without apparent phenotype. In Wg-expressing discs, in situ hybridization detected *lama* transcripts in a broad dorsal domain that encompassed the region of vgBE-*lacZ* expression (Fig. 4G), a pattern that is consistent with the microarray hybridization results.

To investigate whether *lama* plays a role in transdetermination, we first examined *lama* expression during normal larval development. In situ hybridization revealed that *lama* is not expressed in late 3rd instar discs, but is expressed uniformly by early 3rd instar discs, in the imaginal ring of the proventriculus and in the salivary gland (Fig. 4). Functional tests were performed in the viable, null mutant $lama^{410}$ (Perez and Steller, 1996) by monitoring vgBE-*lacZ* expression in Wg-expressing mutant discs as an indicator of transdetermination. Compared with controls, the frequency of transdetermination was unchanged. However, the relative fraction of the leg disc that expressed *vg* decreased from 5% to 2% (Fig. 5). Expression of *lama* in early, but not late discs suggests that the role of *lama* in normal development may be to suppress pathways that promote differentiation, and the significant decrease in transdetermined region of *lama* mutant discs suggests that it may preserve the pluripotency of disc cells in Wg-expressing discs.

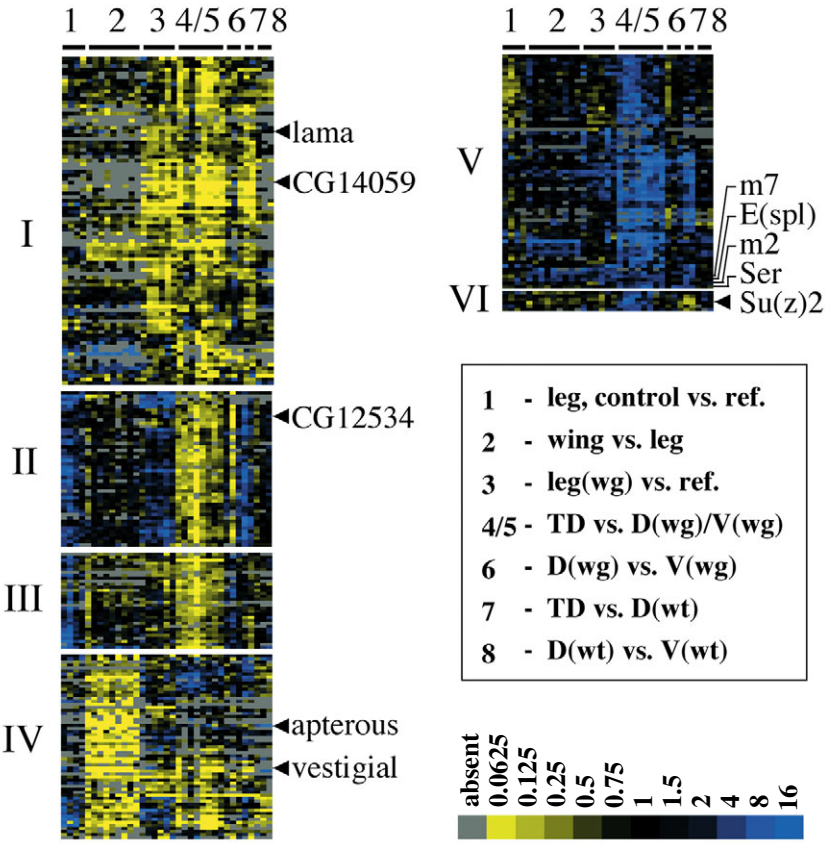

Fig. 3. Expression profiling analysis of transdetermining cells. Cluster analysis of eight different comparisons. Expression ratios are color coded as indicated. Each of 35 columns represents an array experiment grouped into categories as described in Table S1 (see supplementary material). Categories 4, 5 and 5 compare TD cells with D_{Wg}, V_{Wg} and D_{WT}, respectively. Prior to cluster analysis, the expression ratios were filtered at two levels: (1) spots were required to have intensities with a sum of medians greater than 350; and (2) ratios greater than 2 (1 $\log_2$) in at least five out of the 35 experiments. The calculation of self-organizing maps and hierarchical clustering produced a cluster with several sub-clusters, of which six are represented; labeled I-VI with I-IV including genes enriched in TD cells and V and VI containing genes depleted in TD cells. Genes that segregate to sub-cluster I had high levels of expression in TD cells (yellow) and include *lama* and CG14059. The many absent calls (grey) in the wing-to-leg comparison for genes of this sub-cluster suggest that these genes are not expressed in either wild-type wing or leg discs, but are enriched in Wg-induced leg discs undergoing transdetermination. Genes in subclusters II and III also had elevated expression in TD cells, but do not have elevated expression levels in wing discs. Subcluster II includes CG12534. Genes that segregated to subcluster IV show high expression in wing discs (yellow) and some of these genes are enriched in TD cells, indicating the realization of a wing developmental program in Wg-induced leg discs. This group includes *ap* and *vg*. Genes in subclusters V and VI are expressed at low levels in the TD cell preparation (blue). Components of the Notch signaling pathway and the PcG gene *Su(z)2* are included in this group. Gene names for all genes are listed in Tables S7 and S8 (see supplementary material). Some genes are represented by replicate spots on the microarray, causing multiple listings.

Notch

Levels of *Notch* expression did not change significantly. However, the cluster analysis identified four genes that encode proteins with roles in Notch signaling: *Enhancer of split*, *E(spl) region transcript m7*, *E(spl) region transcript m2* and *Serrate*, whose expression decreased; and two genes, *Notchless* and *bancal*, with increased levels in TD cells. As the vgBE enhancer includes a binding site for Suppressor of Hairy [Su(H)], a transcription factor that acts downstream of Notch and is required for *vg* expression (Couso et al., 1995), Notch signaling may contribute to activation of the vgBE enhancer in TD cells. There may be additional inputs from Notch signaling, as wing fate is among the many cell fate decisions that Notch signaling regulates (Kurata et al., 2000).

Polycomb Group and *trithorax Group* gene expression in transdetermination

The importance of chromatin structure to the transcriptional state of determined cells makes it reasonable to assume that re-programming cells to different fates entails reorganization of the Polycomb group (PcG) and trithorax group (trxG) protein complexes that bind to regulatory elements (Paro et al., 1998). Although altering the distribution of proteins that mediate chromatin states for transcriptional repression and activation need not involve changes in the levels of expression of the PcG and trxG proteins, we analyzed our array hybridization data to determine if they do. The PcG *Suppressor of zeste 2* [*Su(z)2*] gene had a median fold repression of 2.1 in eight TD to D_{Wg}/V_{Wg} comparisons (P=0.021), but our cut-off settings did not detect significant enrichment or repression of most of the other PcG or trxG protein genes with either clustering analysis or the method of ranking median ratios. As criteria for assigning biological significance to levels of change are purely subjective, we re-analyzed the transdetermination expression data to identify genes whose median ratio changes within a 95% confidence level. Fourteen percent of the genes satisfied these conditions [>0.25 (log2); binomial test, $P=1.140\times10^{-5}$]. Among these genes, 15/32 *PcG* and *trxG* genes (47%) had such statistically significant changes (Table 2). Identification of these 15 genes with differential expression suggests that transdetermination may be correlated with large-scale remodeling of chromatin structure.

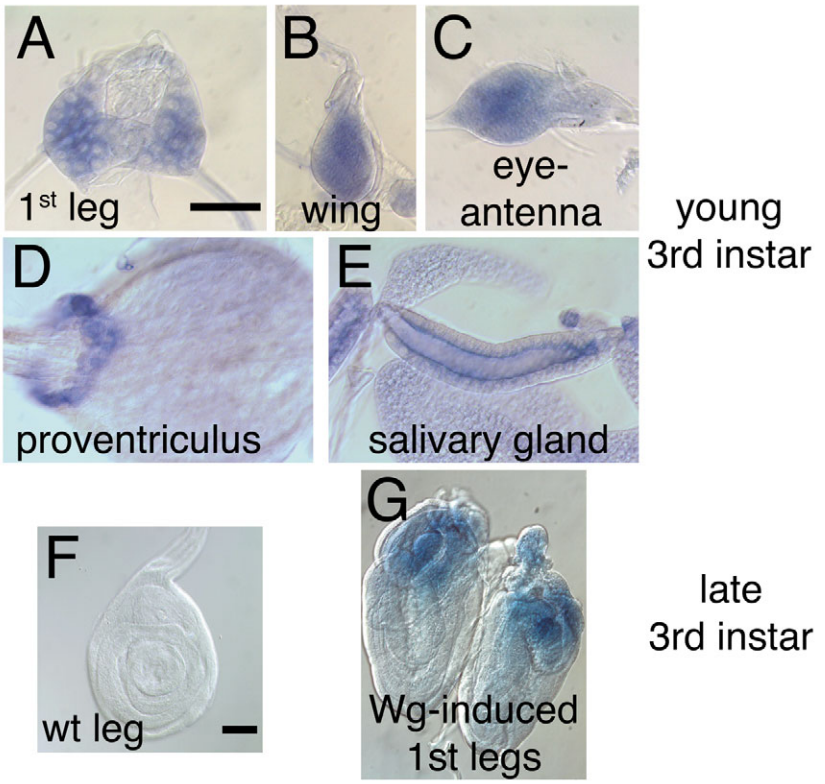

Fig. 4. *lama* expression in larval organs and transdetermined cells. In situ hybridization with a *lama* probe in young 3rd instar larval organs (A-E), a disc from a wandering 3rd instar (F) and in 1st leg discs from a late 3rd instar (G) induced to transdetermine by ectopic Wg expression. Scale bars: in A, 50 μm for A-E; in F, 50 μm for F,G.

To test if the small but statistically significant changes in the expression of PcG and trxG genes are indicative of a functional role in determination, we analyzed discs from wild-type, *Polycomb* (*Pc*), *Enhancer of Polycomb* [*E(Pc)*], *Sex comb on midleg* (*Scm*), *Enhancer of zeste* [*E(z)*], *Su(z)2*, *brahma* (*brm*) and *osa* (*osa*) larvae. We adjusted the level of Wg induction to reduce the frequency of transdetermination and we monitored both frequency of transdetermination and area of transdetermined cells. As shown in Fig. 5 (see also Table S11 in the supplementary material), the frequency of leg discs expressing *vg* increased significantly in *E(z)*, *Pc*, *E(Pc)*, *brm* and *osa* mutants, and the frequency of leg to wing transdetermination in adult cuticle increased in *Scm*, *E(z)*, *Pc*, *E(Pc)* and *osa* mutants. Remarkably, *Su(z)2* heterozygous discs had no *vg* expression, suggesting that the loss of Su(z)2 function limits *vg* expression.

Members of the PcG and trxG are known to act as heteromeric complexes by binding to cellular memory modules (CMMs). Our functional tests demonstrate that mutant alleles for members of both groups have the same functional consequence (they increase transdetermination frequency). Our findings are consistent with recent observations that the traditional view of PcG members as repressors and trxG factors as activators might be an oversimplification, and that a more complex interplay of a varying composition of PcG and trxG proteins takes place at individual CMMs (reviewed by Lund and van Lohuizen, 2004). Furthermore the opposing effects of Pc and Su(z)2 functions are consistent with the proposal that Su(z)2 is one of a subset of PcG genes that is required to activate as well as to suppress gene expression (Gildea et al., 2000). In addition to measuring the frequency of transdetermination, we also analyzed the relative area of *vg* expression in the various PcG and trxG heterozygous mutant discs. As shown in Fig. 5B (see also Table S11 in the supplementary material), the relative area decreased in *E(Pc)*, *brm* and *osa* mutant discs, despite the increased frequency of transdetermination in these mutants. We do not have evidence to explain these contrasting effects, but the roles of these seven PcG and trxG genes in transdetermination that these results identify support the proposition that the transcriptional state of determined cells is implemented through the controls imposed by the regulators of chromatin structure.

Regulation of genes involved in or responsive to transdetermination

As described above, we used the *vg* boundary enhancer (vgBE) that is activated in *wg*-expressing 1st leg discs to mark and identify transdetermined cells. This element is one of two identified and characterized enhancers that regulate *vg* expression in wing disc cells (Kim et al., 1996; Williams et al., 1991). It integrates regulatory input from the *wg* and *Notch* signaling pathways as well as from Vg itself (Kim et al., 1996; Klein and Arias, 1999). The Vg protein functions in concert

Table 2. Involvement of PcG and trxG genes and *lama* in transdetermination

Gene	Median	*P*
A trx Group		
Increased in TD cells		
CG4303, *Bap60*	–0.55	0.017
CG18740, *moira*	–0.52	0.027
CG8573, *Su(Hw)*	–0.28	0.018
Unchanged		
CG32346, *E(bx)*	–0.26	0.154
CG6677, *ash2*	–0.19	0.064
CG7803, *z*	–0.07	0.863
CG9343, *Trl*	–0.01	0.758
CG8887, *ash1*	0.05	0.840
CG32491, *mod (mdg4)*	0.19	0.017
CG8651, *trx*	0.22	0.218
CG1064, *SNF5 related*	0.29	0.100
CG8625, *ISWI*	0.51	0.097
Decreased		
CG7467, *eyelid, osa*	0.42	0.020
CG5942, *brm*	0.36	0.030
B Pc Group		
Increased in TD cells		
CG17743, *pho*	–0.65	0.045
CG14941, *esc*	–0.54	0.007
CG32443, *Pc*	–0.49	0.025
CG6502, *E(z)*	–0.33	0.035
Unchanged		
CG5738, *lola-like*	–0.42	0.443
CG2368, *psq*	–0.31	0.056
CG9696, *domino*	–0.29	0.171
CG8013, *Su(z)12*	–0.02	0.700
CG3886, *Psc*	–0.36	0.100
CG4236, *Caf1*	0.03	0.840
CG18412, *ph-p*	0.10	0.217
CG3895, *ph-d*	0.06	0.641
Decreased		
CG9495, *Scm*	0.46	0.020
CG5109, *Pcl*	0.48	0.033
CG8787, *Asx*	0.57	0.020
CG7776, *E(Pc)*	0.63	0.020
CG2714, *crm*	0.66	0.017
CG3905, *Su(z)2, Arp*	1.08	0.017

Based on the median expression ratios of eight replicate experiments (TD-to-D_{Wg} and TD-to-V_{Wg} comparisons) and *t*-test calculations, *trxG* (A) and *PcG* (B) genes with a statistically significant up- or downregulation were identified.

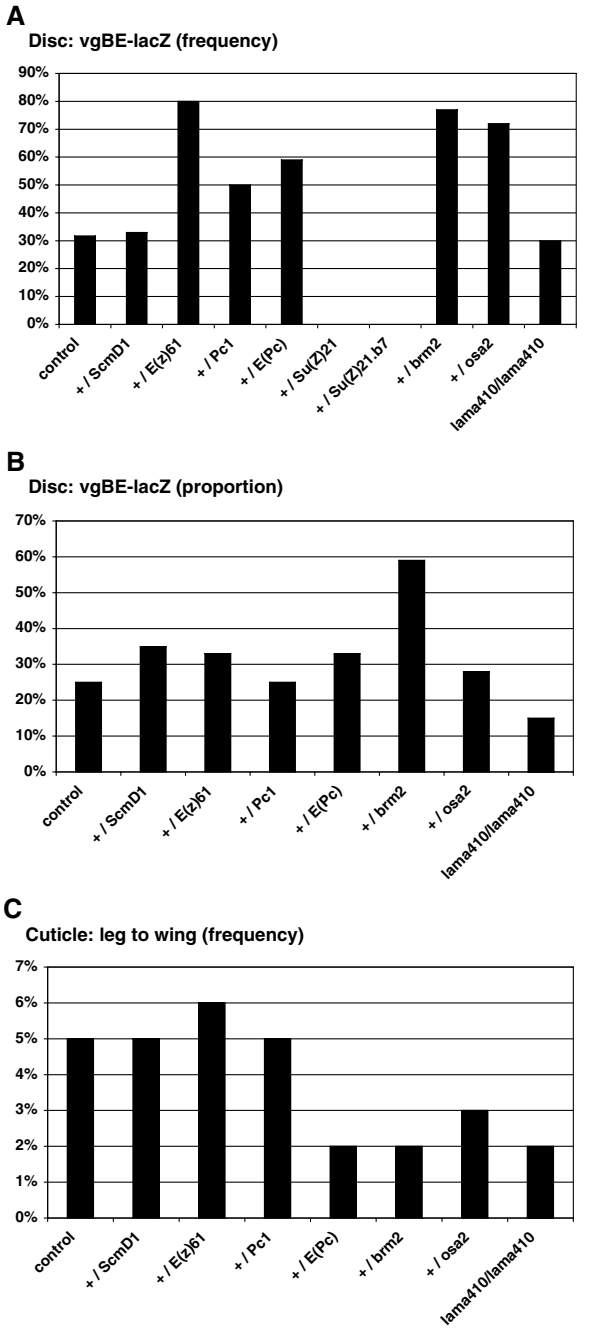

Fig. 5. Involvement of *PcG*, *trxG* and *lama* genes in transdetermination. The frequency of transdetermination was determined (A) by vgBE-*lacZ* expression and (B) by the occurrence of wing cuticle for control discs, *PcG* heterozygous mutant discs [*Scm*, *E*(*z*), *Pc*, *E*(*Pc*), and *Su*(*z*)*2*], *trxG* heterozygous mutant discs (*brm* and *osa*) and homozygous *lama* mutant discs. The frequency of wing cuticle could not be measured for *Su*(*z*)*2* as no adult survivors were recovered and it was not determined for *brm*. (C) Determinations of the relative area of transdetermination compared the area of vgBE-*lacZ* expressing cells with the total disc area (%). For details of the genetic background and *t*-test statistics, see the legend to Table S11 in the supplementary material.

with the Scalloped (Sd) protein (Halder et al., 1998; Paumard-Rigal et al., 1998; Simmonds et al., 1998), and we identified an element with similarity to the consensus Sd-binding site (Halder et al., 1998) within the vgBE. vgBE also has binding sites for Pangolin (Pan, which is homologous to LEF1/TCF1 and is a transcription factor that functions in Wg signal transduction) and Su(H), a transcription factor downstream of Notch (Kim et al., 1996; Klein and Arias, 1999; Williams et al., 1991). To gauge the significance of the three clustered binding sites in vgBE, we examined the *vg* genomic sequences of *D. yakuba* and *D. pseudoobscura*, and found Sd, Pan and Su(H) sites in putative *vgBE* sequences that are located in the same relative position as in the *D. melanogaster* genome (see Table S9 in the supplementary material).

We also analyzed the upstream regions of the genes whose transcripts were enriched in *D. melanogaster* TD cells, using a computational supervised search for vgBE-like motifs. Seventeen genes contain one or more clusters of Sd-, Pan- and Su(H)-binding site sequences and 45 had putative binding sites for at least one of these transcription factors (see Table S10 in

the supplementary material). Among the TD genes with binding sites are *vg*, CG14059 and CG12534. Verification that these sequences bind Sd, Pan and Su(H) is necessary, but the presence of these sequences is consistent with the expectation that many of the TD genes are directly regulated by the Vg, Wg and Notch pathways.

Cis-regulatory elements in the TD genes

We also conducted an unbiased query for putative cis-regulatory elements in the genes represented on our arrays. This analysis identified the sequence TATCGAYW at a statistically significant frequency in the 5 kb upstream genomic region of 77 genes. Interestingly, 21/77 (27%) of these genes were members of the group of genes whose expression was enriched ⩾twofold in TD cells (see Table S10 in the supplementary material). TATCGAYW closely resembles the DNA replication element (DRE), TATCGATA, which was previously found to be located upstream of many replication- and proliferation-related genes (Yamaguchi et al., 1995), and is a binding site for the DNA replication-related element binding factor DREF (Hirose et al., 1996). Among the genes in the TD group that have DRE sites are *vg*, CG14059, genes that encode the alpha DNA polymerase complex, primases, helicases, stress response proteins and *Brahma associated protein 60kD* (*Bap60*). *Bap60* is one of the *trx-G* genes that is upregulated in TD cells (Table 2).

Discussion

The determined states that direct cells to particular fates or lineages can be remarkably stable and can persist after many cell divisions in alien environments, but they are not immune to change. In *Drosophila*, three experimental systems have provided opportunities to investigate the mechanisms that lead to switches of determined states. These are: (1) the classic homeotic mutants; (2) the PcG and trxG mutants that affect the capacity of cells to maintain homeotic gene expression; and (3) transdetermination. During normal development, the homeotic genes are expressed in spatially restricted regions, and cells that lose (or gain) homeotic gene function presumably change the transcriptional profiles characteristic of the particular body part. In the work reported here, we used techniques of micro-dissection, RNA amplification and array hybridization to monitor the transcription profiles of cells in normal leg and wing imaginal discs, in leg disc cells that regenerate and in cells that transdetermine from leg to wing. Our results validate the idea that changing determined states involves global changes in gene expression. They also identify genes whose function may be unrelated to the specific fates of the cells we characterized, but instead may correlate with developmental plasticity.

The general issue these studies address is the molecular basis for changes in cell fate, a subject we tackled by analyzing transcriptional profiles to ask which genes instigate and elaborate such changes. The microarray experiments we performed yielded relative quantitative data for as many as 75% of the predicted transcription units in the *Drosophila* genome, and we showed using in situ hybridization probes for a small number of genes that the predictive value of the data is excellent. Genes identified in the arrays to be expressed in a wing-specific or leg-specific manner, or to be differentially regulated in TD cells had expression patterns consistent with the array results. Importantly, tests of several genes implicated by differential expression in TD cells indicated that these genes function to either promote or inhibit transdetermination. In part, the value of these studies is the demonstration that the methods we used yield high quality data from small amounts of tissue that can be isolated by hand with little effort, making it possible to carry out replicate experiments for many conditions and samples. In work that is yet unpublished, we used these same methods to analyze other developmental systems in *Drosophila* – for example, haltere discs transformed by various combinations of *bithorax* mutants, different larval tracheal metameres, various parts of the male reproductive apparatus, and anterior and posterior compartments of the wing disc (A.K. and T.B.K., unpublished). We suggest that these methods will be useful for studies of developmental programs in many other contexts as well. The method of analysis we used is not comprehensive, as the arrays do not query all the predicted protein coding sequences and the amplification technique we developed does not yield a complete representation of full-length transcripts. However, as we show in this work, the method constitutes a robust and high value screen that can be queried in various ways to identify candidate genes. The data we obtained is informative both about the general landscape of the transcriptional profiles and about individual genes.

Overlap between the transcriptional profiles in the wing and transdetermination lists (15 genes) and with genes in subcluster IV (Fig. 3) is extensive. The overlap is sufficient to indicate that the TD leg disc cells have changed to a wing-like program of development, but interestingly, not all wing-specific genes were activated in the TD cells. The reasons could be related to the incomplete inventory of wing structures produced (only ventral wing; G.S., unpublished) or to the altered state of the TD cells. During normal development, *vg* expression is activated in the embryo and continues through the 3rd instar. Although the regulatory sequences responsible for activation in the embryo have not been identified, in 2nd instar wing discs, *vg* expression is dependent upon the vgBE enhancer, and in 3rd instar wing discs expression is dependent upon the vgQE enhancer (Kim et al., 1996; Klein and Arias, 1999). Expression of *vg* in TD cells depends on activation by the vgBE enhancer (Maves and Schubiger, 1998), indicating that cells that respond to Wg-induction do not revert to an embryonic state. Recent studies of the cell cycle characteristics of TD cells support this conclusion (Sustar and Schubiger, 2005), but the role of the vgBE enhancer in TD cells and the incomplete inventory of 'wing-specific genes' in their expression profile probably indicates as well that at the time that we analyzed the TD cells, they were not equivalent to the cells of late 3rd instar wing discs.

Investigations into the molecular basis of transdetermination have led to a model in which inputs from the Wg, Dpp and Hh signaling pathways alter the chromatin state of key selector genes to activate the transdetermination pathway (Maves and Schubiger, 2003). Our analyses were limited to a period 2-3 days after the cells switched fate, because several cell doublings were necessary to produce sufficient numbers of marked TD cells. As a consequence, these studies did not analyze the initial stages. Despite this technical limitation, this study identified several genes that are interesting novel markers

of transdetermination (e.g. *ap*, CG12534, CG14059 and CG4914), as well as several genes that function in the transdetermination process (e.g. *lama* and the *PcG* genes). The results from our transcriptional profiling add significant detail to the general model proposed by Maves and Schubiger (Maves and Schubiger, 2003).

First, we report that ectopic *wg* expression results in statistically significant changes in the expression of 15 PcG and trxG genes. Moreover, although the magnitudes of these changes were very small for most of these genes, functional assays with seven of these genes revealed remarkably large effects on the metrics we used to monitor transdetermination – the fraction of discs with TD cells, the proportion of disc epithelium that TD cells represent, and the fraction of adult legs with wing cuticle. These effects strongly implicate PcG and trxG genes in the process of transdetermination and suggest that the changes in determined states manifested by transdetermination are either driven by or are enabled by changes in chromatin structure. This conclusion is consistent with the demonstrated roles of PcG and trxG genes in the self-renewing capacity of mouse hematopoietic stem cells (reviewed by Valk-Lingbeek et al., 2004), in Wg signaling and in the maintenance of determined states (Barker et al., 2001; Collins and Treisman, 2000; Petruk et al., 2001). Our results now show that the PcG and trxG functions are also crucial to pluripotency in imaginal disc cells, namely that pluripotency by 'weak point' cells is dependent upon precisely regulated levels of PcG and trxG proteins, and is exquisitely sensitive to reductions in gene dose.

Our data do not suggest how the *PcG* and *trxG* genes affect transdetermination, but several possible mechanisms deserve consideration. The recent study of Sustar and Schubiger (Sustar and Schubiger, 2005) reported that transdetermination correlates with an extension of the S phase of the cell cycle. Several proteins involved in cell cycle regulation physically associate with PcG and trxG proteins (Brumby et al., 2002; Trimarchi et al., 2001), and Brahma, one of the proteins that affects the metrics of transdetermination we measured, has been shown to dissociate from chromatin in late S-phase and to reassociate in G1. It is possible that changes in the S-phase of TD cells are a consequence of changes in PcG/trxG protein composition.

Another generic explanation is that transdetermination is dependent or sensitive to expression of specific targets of *PcG* and *trxG* genes. Among the 167 Pc/Trx response elements (PRE/TREs) predicted to exist in the *Drosophila* genome (Ringrose et al., 2003), one is in direct proximity to the *vg* gene. It is possible that upregulation of *vg* in TD cells is mediated through this element. Another factor may be the contribution of targets of Wg signaling, as Collins and Treisman reported that targets of Wg signaling are upregulated in *osa* and *brm* mutants (Collins and Treisman, 2000). These are among a number of likely possible targets, and identifying the sites at which the PcG and trxG proteins function will be necessary if we are to understand how transdetermination is regulated. Importantly, understanding the roles of such targets and establishing whether these roles are direct will be essential to rationalize how expression levels of individual *PcG* and *trxG* genes correlate with the effects of *PcG* and *trxG* mutants on transdetermination.

Second, the requirement for *lama* suggests that proliferation of TD cells involves functions that suppress differentiation. *lama* expression has been correlated with neural and glial progenitors prior to, but not after, differentiation (Perez and Steller, 1996), and we observed that *lama* is expressed in imaginal progenitor cells and in early but not late 3rd instar discs (Fig. 4). We found that *lama* expression is re-activated in leg cells that transdetermine. The upregulation of *unpaired* in TD cells may be relevant in this context, as the JAK/STAT pathway functions to suppress differentiation and to promote self-renewal of stem cells in the *Drosophila* testis (reviewed by Hombria and Brown, 2002). We suggest that it has a similar role in TD cells.

Third, a role for Notch is implied by the expression profiles of several Notch pathway genes. Notch may contribute directly to transdetermination through the activation of the vgBE enhancer [which has a binding site for Su(H)] and of similarly configured sequences that we found to be present in the regulatory regions of 45 other TD genes (see Table S10 in the supplementary material). It will be important to test whether Notch signaling is required to activate these co-expressed genes, and if it is, to learn what cell-cell interactions and 'community effects' regulate activation of the Notch pathway in TD cells.

Fourth, the upregulation in TD cells of many genes involved in growth and division, and the identification of a DRE sites in the regulatory region of many of these genes supports the observation that TD cells become re-programmed after passing through a novel proliferative state (Sustar and Schubiger, 2005), and suggests that this change is in part implemented through DRE-dependent regulation.

Two final comments: we are intrigued by the interesting correlation between transdetermination induced by Wg mis-expression and the role of Wg/Wnt signaling for stem cells. Wg/Wnt signaling functions as a mitogen and maintains both somatic and germline stem cells in the *Drosophila* ovary (Gonzalez-Reyes, 2003), and mammalian hematopoetic stem cells (Reya et al., 2003). Although the 'weak point' cells in the *Drosophila* leg disc might lack the self-renewing capacity that characterizes stem cells, they respond to Wg mis-expression by manifesting a latent potential for growth and transdetermination. It seems likely that many of the genes involved in regulating stem cells and in leading to disease states when the relevant regulatory networks lose their effectiveness are conserved.

We are also intrigued by the prevalence of transcription factors among the genes whose relative expression levels differed most in our tissue comparisons. It is commonly assumed that transcription factors function catalytically and that they greatly amplify the production of their targets, so the expectation was that the targets of tissue-specific transcription factors would have the highest degree of tissue-specific expression. In our studies, tissue-specific expression of 15 transcription factors among the 40 top-ranking genes in the wing and leg data sets (38%) is consistent with the large number of differentially expressed genes in these tissues, but these rankings suggest that the targets of these transcription factors are expressed at lower relative levels than the transcription factors that regulate their expression. One possible explanation is that the targets are expressed in both wing and leg disc cells, but the transcription factors that regulate them are not. This would imply that the importance of

position-specific regulation lies with the regulator, not the level of expression of the target. Another possibility is that these transcription factors do not act catalytically to amplify the levels of their targets, or do so very inefficiently and require a high concentration of transcription factor to regulate the production of a small number of transcripts. Further analysis will be required to distinguish between these or other explanations, but we note that the prevalence of transcription factors in such data sets is neither unique to wing-leg comparisons (A.K. and T.B.K., unpublished) nor universal.

We thank Miguel Ramalho-Santos for help comparing expression profiles, Songmei Liu and Ulla Cinque for technical assistance, and members of the Joe DeRisi laboratory for their help and advice. This work was supported by grants from the Herbert Boyer Fund to A.K., by a postdoctoral fellowship to K.K. from the Sloan Foundation/DOE, by a grant from the Ellison Medical Foundation to T.B.K., and by grants from the NIH to H.L., G.S. and T.B.K.

Supplementary material

Supplementary material for this article is available at
http://dev.biologists.org/cgi/content/full/132/16/3753/DC1

References

Adams, G. A., Maestri, M., Squiers, E. C., Alfrey, E. J., Starzl, T. E. and Dafoe, D. C. (1998). Augmenter of liver regeneration enhances the success rate of fetal pancreas transplantation in rodents. *Transplantation* **65**, 32-36.

Adams, M. D., Celniker, S. E., Holt, R. A., Evans, C. A., Gocayne, J. D., Amanatides, P. G., Scherer, S. E., Li, P. W., Hoskins, R. A., Galle, R. F. et al. (2000). The genome sequence of *Drosophila* melanogaster. *Science* **287**, 2185-2195.

Adler, P. N., Charlton, J. and Brunk, B. (1989). Genetic interactions of the suppressor 2 of zeste region genes. *Dev. Genet.* **10**, 249-260.

Bach, E. A., Vincent, S., Zeidler, M. P. and Perrimon, N. (2003). A sensitized genetic screen to identify novel regulators and components of the Drosophila janus kinase/signal transducer and activator of transcription pathway. *Genetics* **165**, 1149-1166.

Baker, N. E. (1988). Transcription of the segment-polarity gene wingless in the imaginal discs of Drosophila, and the phenotype of a pupal-lethal wg mutation. *Development* **102**, 489-497.

Barker, N., Hurlstone, A., Musisi, H., Miles, A., Bienz, M. and Clevers, H. (2001). The chromatin remodelling factor Brg-1 interacts with beta-catenin to promote target gene activation. *EMBO J.* **20**, 4935-4943.

Benjamini, Y., Drai, D., Elmer, G., Kafkafi, N. and Golani, I. (2001). Controlling the false discovery rate in behavior genetics research. *Behav. Brain Res.* **125**, 279-284.

Berman, B. P., Nibu, Y., Pfeiffer, B. D., Tomancak, P., Celniker, S. E., Levine, M., Rubin, G. M. and Eisen, M. B. (2002). Exploiting transcription factor binding site clustering to identify cis-regulatory modules involved in pattern formation in the Drosophila genome. *Proc. Natl. Acad. Sci. USA* **99**, 757-762.

Bierie, B., Nozawa, M., Renou, J. P., Shillingford, J. M., Morgan, F., Oka, T., Taketo, M. M., Cardiff, R. D., Miyoshi, K., Wagner, K. U. et al. (2003). Activation of beta-catenin in prostate epithelium induces hyperplasias and squamous transdifferentiation. *Oncogene* **22**, 3875-3887.

Bornemann, D., Miller, E. and Simon, J. (1998). Expression and properties of wild-type and mutant forms of the Drosophila sex comb on midleg (SCM) repressor protein. *Genetics* **150**, 675-686.

Brumby, A. M., Zraly, C. B., Horsfield, J. A., Secombe, J., Saint, R., Dingwall, A. K. and Richardson, H. (2002). Drosophila cyclin E interacts with components of the Brahma complex. *EMBO J.* **21**, 3377-3389.

Bussemaker, H. J., Li, H. and Siggia, E. D. (2001). Regulatory element detection using correlation with expression. *Nat. Genet.* **27**, 167-171.

Cohen, B., McGuffin, M. E., Pfeifle, C., Segal, D. and Cohen, S. M. (1992). apterous, a gene required for imaginal disc development in Drosophila encodes a member of the LIM family of developmental regulatory proteins. *Genes Dev.* **6**, 715-729.

Collins, R. T. and Treisman, J. E. (2000). Osa-containing Brahma chromatin remodeling complexes are required for the repression of wingless target genes. *Genes Dev.* **14**, 3140-3152.

Corpet, F. (1988). Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res.* **16**, 10881-10890.

Couso, J. P., Knust, E. and Martinez Arias, A. (1995). Serrate and wingless cooperate to induce vestigial gene expression and wing formation in Drosophila. *Curr. Biol.* **5**, 1437-1448.

Diaz-Benjumea, F. J. and Cohen, S. M. (1993). Interaction between dorsal and ventral cells in the imaginal disc directs wing development in Drosophila. *Cell* **75**, 741-752.

Eisen, M. B., Spellman, P. T., Brown, P. O. and Botstein, D. (1998). Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. USA* **95**, 14863-14868.

Frith, M. C., Hansen, U. and Weng, Z. (2001). Detection of cis-element clusters in higher eukaryotic DNA. *Bioinformatics* **10**, 878-889.

Gildea, J. J., Lopez, R. and Shearn, A. (2000). A screen for new trithorax group genes identified little imaginal discs, the Drosophila melanogaster homologue of human retinoblastoma binding protein 2. *Genetics* **156**, 645-663.

Gindhart, J. G., Jr and Kaufman, T. C. (1995). Identification of Polycomb and trithorax group responsive elements in the regulatory region of the Drosophila homeotic gene Sex combs reduced. *Genetics* **139**, 797-814.

Gonzalez-Reyes, A. (2003). Stem cells, niches and cadherins: a view from Drosophila. *J. Cell Sci.* **116**, 949-954.

Hadorn, E. (1963). Differenzierungsleistungen wiederholt fragmentierter Teilstücke männlicher Genitalscheiben von Drosophila melanogaster nach Dauerkultur in vivo. *Dev. Biol.* **7**, 617-629.

Hadorn, E. (1965). Problems of determination and transdetermination. *Brookhaven Symp. Biol.* **18**, 148-161.

Hagiya, M., Francavilla, A., Polimeno, L., Ihara, I., Sakai, H., Seki, T., Shimonishi, M., Porter, K. A. and Starzl, T. E. (1994). Cloning and sequence analysis of the rat augmenter of liver regeneration (ALR) gene: expression of biologically active recombinant ALR and demonstration of tissue distribution. *Proc. Natl. Acad. Sci. USA* **91**, 8142-8146.

Halder, G., Polaczyk, P., Kraus, M. E., Hudson, A., Kim, J., Laughon, A. and Carroll, S. (1998). The Vestigial and Scalloped proteins act together to directly regulate wing-specific gene expression in Drosophila. *Genes Dev.* **12**, 3900-3909.

Harrison, D. A., McCoon, P. E., Binari, R., Gilman, M. and Perrimon, N. (1998). Drosophila unpaired encodes a secreted protein that activates the JAK signaling pathway. *Genes Dev.* **12**, 3252-3263.

Hirose, F., Yamaguchi, M., Kuroda, K., Omori, A., Hachiya, T., Ikeda, M., Nishimoto, Y. and Matsukage, A. (1996). Isolation and characterization of cDNA for DREF, a promoter-activating factor for Drosophila DNA replication-related genes. *J. Biol. Chem.* **271**, 3930-3937.

Hombria, J. C. and Brown, S. (2002). The fertile field of Drosophila Jak/STAT signalling. *Curr. Biol.* **12**, R569-R575.

Ihaka, R. and Gentleman, R. R. (1996). A language for data analysis and graphics. *J. Comp. Graphic. Stat.* **5**, 2999-2344.

Jankowski, J. A., Wright, N. A., Meltzer, S. J., Triadafilopoulos, G., Geboes, K., Casson, A. G., Kerr, D. and Young, L. S. (1999). Molecular evolution of the metaplasia-dysplasia-adenocarcinoma sequence in the esophagus. *Am. J. Pathol.* **154**, 965-973.

Janson, K., Cohen, E. D. and Wilder, E. L. (2001). Expression of DWnt6, DWnt10, and DFz4 during Drosophila development. *Mech. Dev.* **103**, 117-120.

Johnston, L. A. and Schubiger, G. (1996). Ectopic expression of wingless in imaginal discs interferes with decapentaplegic expression and alters cell determination. *Development* **122**, 3519-3529.

Jones, R. S. and Gelbart, W. M. (1990). Genetic analysis of the enhancer of zeste locus and its role in gene regulation in Drosophila melanogaster. *Genetics* **126**, 185-199.

Kim, J., Sebring, A., Esch, J. J., Kraus, M. E., Vorwerk, K., Magee, J. and Carroll, S. B. (1996). Integration of positional signals and regulation of wing formation and identity by Drosophila vestigial gene. *Nature* **382**, 133-138.

Klebes, A., Biehs, B., Cifuentes, F. and Kornberg, T. B. (2002). Expression profiling of Drosophila imaginal discs. *Genome Biol.* **3**, R0038.

Klein, T. and Arias, A. M. (1999). The vestigial gene product provides a molecular context for the interpretation of signals during the development of the wing in Drosophila. *Development* **126**, 913-925.

Kurata, S., Go, M. J., Artavanis-Tsakonas, S. and Gehring, W. J. (2000). Notch signaling and the determination of appendage identity. *Proc. Natl. Acad. Sci. USA* **97**, 2117-2122.

Lindsley, D. L. and Zimm, G. G. (1992). *The Genome of Drosophila melanogaster*. San Diego, CA: Academic Press.
Lisowsky, T., Weinstat-Saslow, D. L., Barton, N., Reeders, S. T. and Schneider, M. C. (1995). A new human gene located in the PKD1 region of chromosome 16 is a functional homologue to ERV1 of yeast. *Genomics* **29**, 690-697.
Lund, A. H. and van Lohuizen, M. (2004). Polycomb complexes and silencing mechanisms. *Curr. Opin. Cell Biol.* **16**, 239-246.
Maves, L. and Schubiger, G. (1995). Wingless induces transdetermination in developing Drosophila imaginal discs. *Development* **121**, 1263-1272.
Maves, L. and Schubiger, G. (1998). A molecular basis for transdetermination in Drosophila imaginal discs: interactions between wingless and decapentaplegic signaling. *Development* **125**, 115-124.
Maves, L. and Schubiger, G. (2003). Transdetermination in Drosophila imaginal discs: a model for understanding pluripotency and selector gene maintenance. *Curr. Opin. Genet. Dev.* **13**, 472-479.
Merrill, B. J., Gat, U., DasGupta, R. and Fuchs, E. (2001). Tcf3 and Lef1 regulate lineage differentiation of multipotent stem cells in skin. *Genes Dev.* **15**, 1688-1705.
Miyoshi, K., Shillingford, J. M., Le Provost, F., Gounari, F., Bronson, R., von Boehmer, H., Taketo, M. M., Cardiff, R. D., Hennighausen, L. and Khazaie, K. (2002). Activation of beta–catenin signaling in differentiated mammary secretory cells induces transdifferentiation into epidermis and squamous metaplasias. *Proc. Natl. Acad. Sci. USA* **99**, 219-224.
Niemann, C., Owens, D. M., Hulsken, J., Birchmeier, W. and Watt, F. M. (2002). Expression of DeltaNLef1 in mouse epidermis results in differentiation of hair follicles into squamous epidermal cysts and formation of skin tumours. *Development* **129**, 95-109.
O'Neill, J. W. and Bier, E. (1994). Double-label in situ hybridization using biotin and digoxigenin-tagged RNA probes. *Biotechniques* **17**, 870, 874-875.
Okubo, T. and Hogan, B. L. (2004). Hyperactive Wnt signaling changes the developmental potential of embryonic lung endoderm. *J. Biol.* **3**, 11.
Paro, R., Strutt, H. and Cavalli, G. (1998). Heritable chromatin states induced by the Polycomb and trithorax group genes. *Novartis Found. Symp.* **214**, 51-61; discussion 61-66, 104-113.
Paumard-Rigal, S., Zider, A., Vaudin, P. and Silber, J. (1998). Specific interactions between vestigial and scalloped are required to promote wing tissue proliferation in Drosophila melanogaster. *Dev. Genes Evol.* **208**, 440-446.
Perez, S. E. and Steller, H. (1996). Molecular and genetic analyses of lama, an evolutionarily conserved gene expressed in the precursors of the Drosophila first optic ganglion. *Mech. Dev.* **59**, 11-27.
Petruk, S., Sedkov, Y., Smith, S., Tillib, S., Kraevski, V., Nakamura, T., Canaani, E., Croce, C. M. and Mazo, A. (2001). Trithorax and dCBP acting in a complex to maintain expression of a homeotic gene. *Science* **294**, 1331-1334.
Pueyo, J. I., Galindo, M. I., Bishop, S. A. and Couso, J. P. (2000). Proximal-distal leg development in Drosophila requires the apterous gene and the Lim1 homologue dlim1. *Development* **127**, 5391-5402.
Raftery, L. A., Sanicola, M., Blackman, R. K. and Gelbart, W. M. (1991). The relationship of decapentaplegic and engrailed expression in Drosophila imaginal disks: do these genes mark the anterior-posterior compartment boundary? *Development* **113**, 27-33.
Ramalho-Santos, M., Yoon, S., Matsuzaki, Y., Mulligan, R. C. and Melton, D. A. (2002). 'Stemness': transcriptional profiling of embryonic and adult stem cells. *Science* **298**, 597-600.
Reya, T., Duncan, A. W., Ailles, L., Domen, J., Scherer, D. C., Willert, K., Hintz, L., Nusse, R. and Weissman, I. L. (2003). A role for Wnt signalling in self-renewal of haematopoietic stem cells. *Nature* **423**, 409-414.
Ringrose, L., Rehmsmeier, M., Dura, J. M. and Paro, R. (2003). Genome-wide prediction of Polycomb/Trithorax response elements in Drosophila melanogaster. *Dev. Cell* **5**, 759-771.
Schubiger, G. (1973). Regeneration of Drosophila melanogaster male leg disc fragments in sugar fed female hosts. *Experientia* **29**, 631-632.
Schweizer, P. and Bodenstein, D. (1975). Aging and its relation to cell growth and differentiation in Drosophila imaginal discs: developmental response to growth restricting conditions. *Proc. Natl. Acad. Sci. USA* **72**, 4674-4678.
Simmonds, A. J., Liu, X., Soanes, K. H., Krause, H. M., Irvine, K. D. and Bell, J. B. (1998). Molecular interactions between Vestigial and Scalloped promote wing formation in Drosophila. *Genes Dev.* **12**, 3815-3820.
Sivasankaran, R., Calleja, M., Morata, G. and Basler, K. (2000). The Wingless target gene Dfz3 encodes a new member of the Drosophila Frizzled family. *Mech. Dev.* **91**, 427-431.
Smyth, G. K. (2004). Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. *Statistical Applications in Genetics and Molecular Biology* http://www.bepress.com/sagmb/vol3/iss1/art3
Soto, M. C., Chou, T. B. and Bender, W. (1995). Comparison of germline mosaics of genes in the Polycomb group of Drosophila melanogaster. *Genetics* **140**, 231-243.
Strub, S. (1977). Localization of the cells capable of transdetermiantion in a specific region of the male foreleg disc of *Drosophila*. *Wilhelm Roux's Arch. Dev. Biol.* **182**, 69-74.
Sustar, A. and Schubiger, G. (2005). A transient cell cycle shift in Drosophila imaginal disc cells precedes multipotency. *Cell* **120**, 383-393.
Tobler, H. (1966). Zellspezifische Determination und Beziehung zwischen Proliferation und Transdetermination in Bein- und Flu-gelprimordien von Drosophila melanogaster. *J. Embryol. Exp. Morphol.* **16**, 609-633.
Trimarchi, J. M., Fairchild, B., Wen, J. and Lees, J. A. (2001). The E2F6 transcription factor is a component of the mammalian Bmi1-containing polycomb complex. *Proc. Natl. Acad. Sci. USA* **98**, 1519-1524.
Tusher, V. G., Tibshirani, R. and Chu, G. (2001). Significance analysis of microarrays applied to the ionizing radiation response. *Proc. Natl. Acad. Sci. USA* **98**, 5116-5121.
Valk-Lingbeek, M. E., Bruggeman, S. W. and van Lohuizen, M. (2004). Stem cells and cancer; the polycomb connection. *Cell* **118**, 409-418.
van de Wetering, M., Cavallo, R., Dooijes, D., van Beest, M., van, Es, J., Loureiro, J., Ypma, A., Hursh, D., Jones, T., Bejsovec, A. et al. (1997). Armadillo coactivates transcription driven by the product of the Drosophila segment polarity gene dTCF. *Cell* **88**, 789-799.
Vervoort, M., Crozatier, M., Valle, D. and Vincent, A. (1999). The COE transcription factor Collier is a mediator of short-range Hedgehog-induced patterning of the Drosophila wing. *Curr. Biol.* **9**, 632-639.
Vierstraete, E., Verleyen, P., Baggerman, G., D'Hertog, W., Van den Bergh, G., Arckens, L., De Loof, A. and Schoofs, L. (2004). A proteomic approach for the analysis of instantly released wound and immune proteins in Drosophila melanogaster hemolymph. *Proc. Natl. Acad. Sci. USA* **101**, 470-475.
Weaver, T. A. and White, R. A. (1995). headcase, an imaginal specific gene required for adult morphogenesis in Drosophila melanogaster. *Development* **121**, 4149-4160.
Williams, J. A., Bell, J. B. and Carroll, S. B. (1991). Control of Drosophila wing and haltere development by the nuclear vestigial gene product. *Genes Dev.* **5**, 2481-2495.
Xu, E. Y., Lee, D. F., Klebes, A., Turek, P. J., Kornberg, T. B. and Reijo Pera, R. A. (2003). Human BOULE gene rescues meiotic defects in infertile flies. *Hum. Mol. Genet.* **12**, 169-175.
Yamaguchi, M., Hayashi, Y., Nishimoto, Y., Hirose, F. and Matsukage, A. (1995). A nucleotide sequence essential for the function of DRE, a common promoter element for Drosophila DNa replication-related genes. *J. Biol. Chem.* **270**, 15808-15814.