

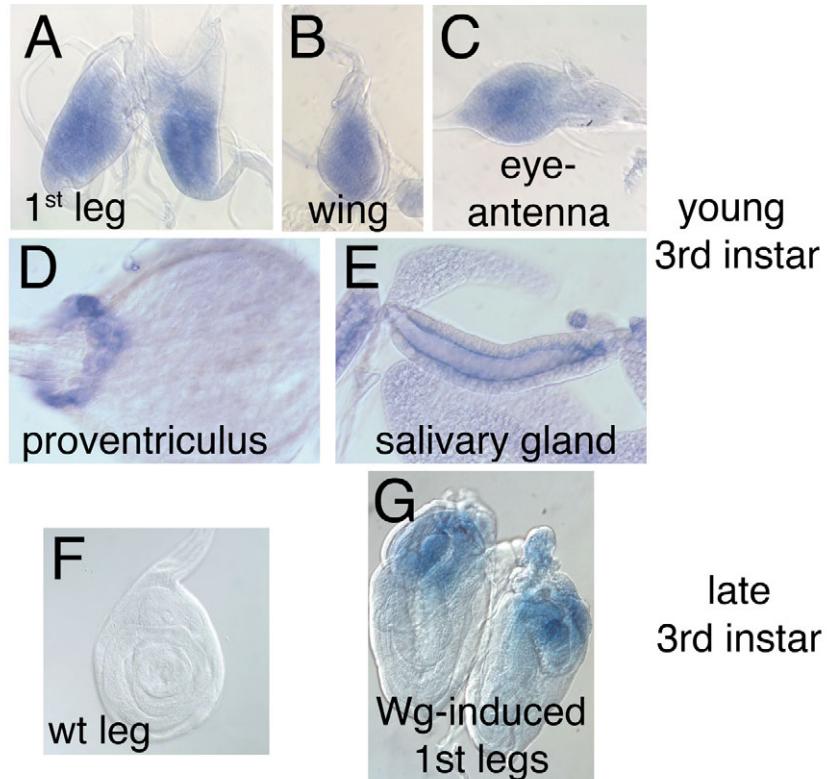
Regulation of cellular plasticity in *Drosophila* imaginal disc cells by the Polycomb group, trithorax group and *lama* genes

Ansgar Klebes, Anne Sustar, Katherina Kechris, Hao Li, Gerold Schubiger and Thomas B. Kornberg *Development* **132**, 3753-3765.

The wrong image was published in panel A of Fig. 4.

The correct figure is printed below.

The authors apologise to readers for this mistake.



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

Ansgar Klebes^{1,*}, Anne Sustar², Katherina Kechris¹, Hao Li¹, Gerold Schubiger² and Thomas B. Kornberg^{1,†}

¹Department of Biochemistry and Biophysics, University of California, San Francisco, CA 94143, USA

²Department 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). *lama⁴¹⁰* (Perez and Steller, 1996), *Su(z)^{2l}* (antimorph) (Adler et al., 1989), *Su(z)^{2l,b7}* (Soto et al., 1995), *E(z)^{6l}* (both loss of function) (Jones and Gelbart, 1990), *Pc^l* (amorph) (Gindhart and Kaufman, 1995), *E(Pc)^l* (Lindsley and Zimm, 1992), *Scm^{D1}* (loss of function) (Bornemann et al., 1998), *osa²* (hypomorph) (Gindhart and

Kaufman, 1995), *brm²* (Gindhart and Kaufman, 1995) and *E(Pc^l)* were analyzed in the transdetermination studies. Expression profiles were determined for male larvae that had been fed 1% bromophenol 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 P_cG 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*	<i>knot/collie</i>	45.25	RNA polymerase II transcription factor activity
CG3830*	<i>vestigial</i> [†]	40.39	Wing margin morphogenesis/transcription factor activity
CG10619*	<i>tailup</i>	16.37	RNA polymerase II transcription factor activity
CG5966	CG5966	15.01	Triacylglycerol lipase activity
CG8376*	<i>apterous</i> [†]	10.84	Zinc ion binding; RNA polymerase II transcription factor activity
CG4382	CG4382	8.98	Carboxylesterase activity
CG1897*	<i>Drop</i>	8.03	RNA polymerase II transcription factor activity
CG12843	<i>Tetraspanin 42Ei</i>	7.96	Tetraspanin
CG12287*	<i>POU domain protein 2</i> [†]	6.66	DNA binding; RNA polymerase II transcription factor activity
CG7160	CG7160	6.07	Not known
CG10570	CG10570	5.97	Not known
CG1803	<i>regucalcin</i> [†]	5.61	Anteroposterior axis specification; Ca ²⁺ -mediated signaling
CG9023	<i>Drip</i>	5.60	Water transporter activity; carrier activity; water channel activity
CG2663	CG2663	5.47	Carrier activity; tocopherol binding
CG3132	<i>Ect3</i> [†]	5.37	β-Galactosidase activity
CG4914*	CG4914 [†]	5.12	Transcription factor activity; trypsin activity; serine-type endopeptidase activity
CG9554	<i>eyes absent</i>	5.11	Hydrolase activity; protein tyrosine phosphatase
CG9427	CG9427	4.97	Not known
CG9623	<i>inflated</i>	4.92	Protein binding; receptor activity; cell-adhesion molecule binding
CG10501	<i>α methyl dopa resistant</i>	4.88	Carboxy-lyase activity; aromatic-L-amino-acid decarboxylase activity
Leg			
CG7807*	<i>AP-2</i>	102.18	RNA polymerase II transcription factor activity
CG5893*	<i>Dorsal</i>	42.28	DNA bending activity; transcription factor activity
CG11922*	<i>forkhead domain 96Cb</i> [†]	25.69	Transcription factor activity
CG6269*	<i>unc-4</i> [†]	22.61	Transcription factor activity
CG6570*	<i>ladybird late</i>	22.05	RNA polymerase II transcription factor activity
CG10382	<i>wrapper</i>	21.26	Axonogenesis; cell-cell adhesion; ectoderm development; gliogenesis; neurogenesis; signal transduction
CG3388*	<i>gooseberry</i>	18.69	RNA polymerase II transcription factor activity
CG6414	CG6414	15.27	Carboxylesterase activity
CG11354*	<i>Lim1</i> [†]	13.04	Transcription factor activity
CG18111	<i>Odorant-binding protein 99a</i>	12.18	Odorant binding
CG11921*	<i>forkhead domain 96Ca</i>	8.33	Transcription factor activity
CG4605	<i>Accessory gland-specific peptide 32CD</i>	7.52	Hormone activity
CG5888	CG5888	7.26	Transmembrane receptor activity
CG4501	<i>bubblegum</i>	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	<i>SP71</i>	3.72	
CG9747	CG9747	3.63	Acyl-CoA delta11-desaturase activity
CG1004	<i>rhomboid</i> [†]	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	<i>unpaired (outstretched)</i> [†]	12.27	Receptor binding; morphogen activity; cytokine activity
CG3830*	<i>vestigial</i> [†]	10.04	Wing morphogenesis
CG6816	<i>Cyp18a1</i>	9.45	Electron transporter activity; oxidoreductase activity
CG15279	CG15279	7.42	Cation:amino acid symporter activity; neurotransmitter:sodium symporter
CG4746	<i>mab-2</i>	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	<i>Amalgam</i>	5.22	Antigen binding
CG8404*	<i>Sox15</i>	5.12	Transcription factor activity
CG5518	<i>slamdance</i>	5.04	Membrane alanyl aminopeptidase activity
CG11822	<i>nAcRbeta-21C</i>	4.69	Nicotinic acetylcholine-activated cation-selective channel activity
CG4914*	CG4914 [†]	4.64	Trypsin activity; serine-type endopeptidase activity
CG3359	<i>midline fasciclin</i>	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	<i>Matrix metalloproteinase 1</i>	4.17	Metalloendopeptidase activity; structural molecule activity
CG7722	CG7722	4.10	Serine-type endopeptidase inhibitor activity
CG1897*	<i>Drop</i>	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 *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).

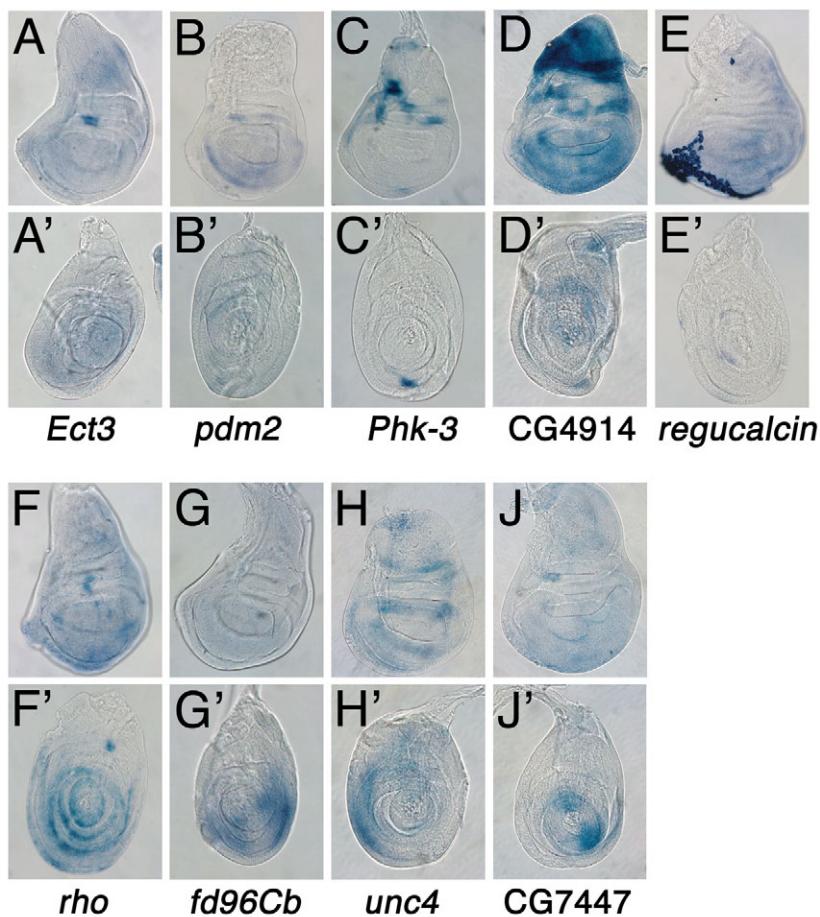


Fig. 1. Whole-mount in situ hybridization to imaginal discs. Wing and leg discs were hybridized with probes derived from genes predicted to have elevated expression in wing discs (A–J) or leg discs (A'–J'). (A) *Ect3* probe stained distinct cell clusters in the presumptive dorsal hinge region of the 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.

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

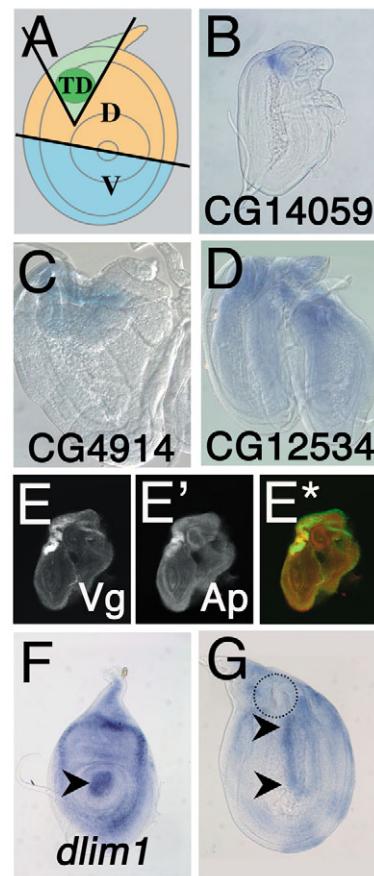


Fig. 2. Novel markers of transdetermining cells in *Wg*-expressing leg discs. (A) Schematic representation of a leg disc and the cuts performed to isolate transdetermining (TD, green), dorsal (D, D_{WT} , D_{Wg} , orange) and ventral cells (V, V_{WT} , V_{Wg} , blue). The area of the weak point is indicated 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.

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

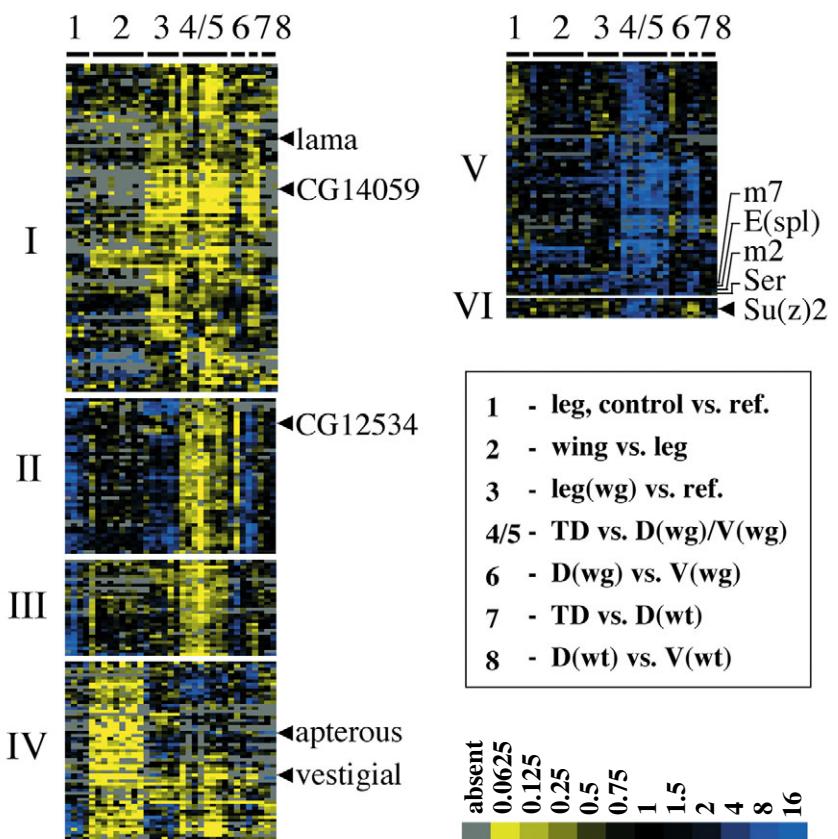


Fig. 3. Expression profiling analysis of transdetermining cells. Cluster analysis of eight different comparisons. Expression ratios are color coded as indicated. Each of 35 columns represents an array experiment grouped into categories as described in Table S1 (see supplementary material). Categories 4, 5 and 6 compare TD cells with Dwg, Vwg and Dwt, 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.

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, $P=1.140 \times 10^{-5}$]. 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.

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

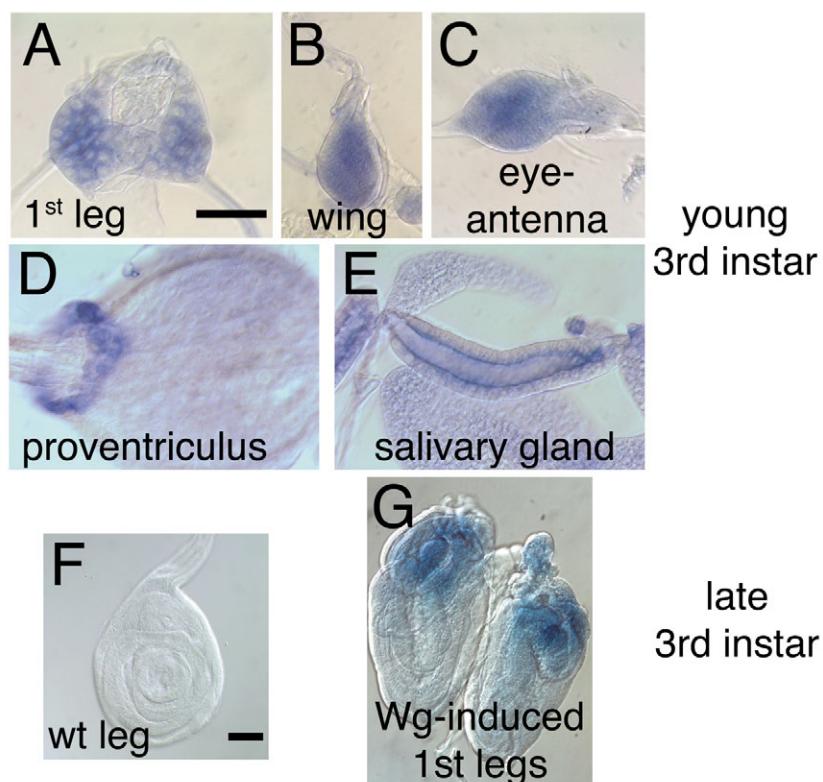


Fig. 4. *lama* expression in larval organs and transdetermined cells. In situ hybridization with a *lama* probe in young 3rd instar larval organs (A-E), a disc from a wandering 3rd instar (F) and in 1st leg discs from a late 3rd instar (G) induced to transdetermine by ectopic Wg expression. Scale bars: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

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 P_cG and tr_xG genes and *lama* in transdetermination

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

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

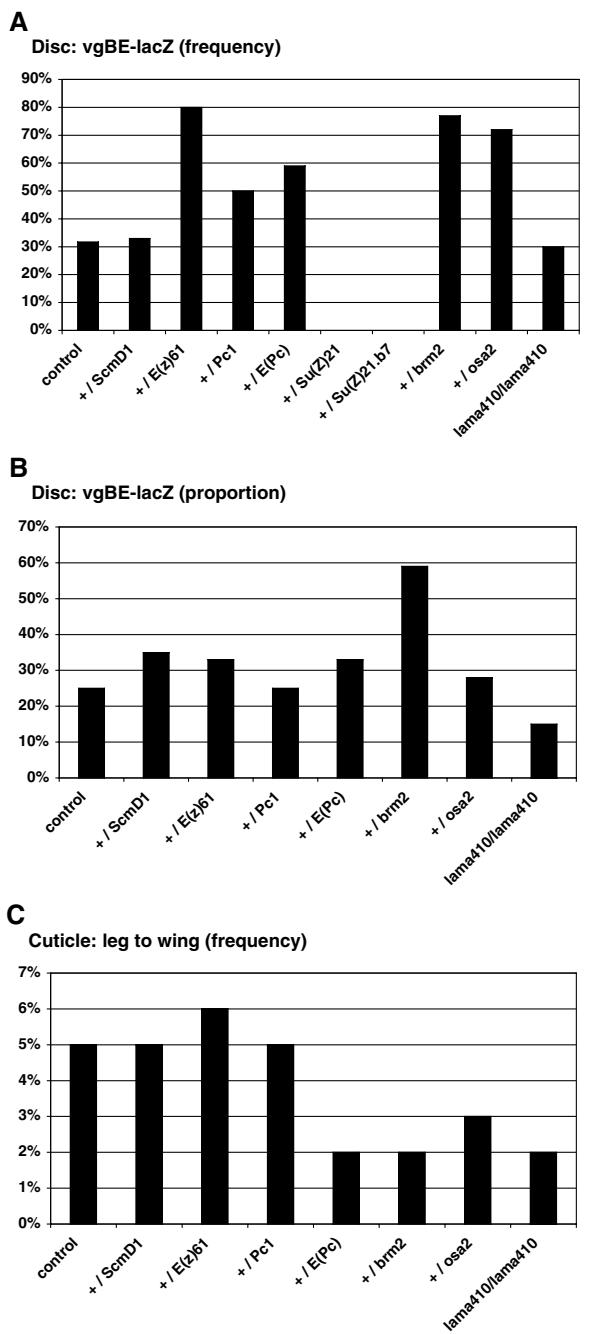


Fig. 5. Involvement of P_cG, tr_xG and *lama* genes in transdetermination. The frequency of transdetermination was determined (A) by vgBE-lacZ expression and (B) by the occurrence of wing cuticle for control discs, P_cG heterozygous mutant discs [Scm, *E(z)*, *Pc*, *E(Pc)*, and *Su(z)2*], tr_xG heterozygous 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 to 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.

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 \geq 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 P_cG 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 microdissection, 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 metamer, 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 hematopoietic 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., Celtniker, 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., Celtniker, 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 Sigga, E. D. (2001). Regulatory element detection using correlation with expression. *Nat. Genet.* **27**, 167-171.
- Cohen, B., McGuffin, M. E., Pfeiffe, 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 wiederholter fragmentierter Teilstücke männlicher Genitalscheiben von *Drosophila melanogaster* nach Dauerkultriv 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., Shimomishi, 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., Hulskens, J., Birchmeier, W. and Watt, F. M.** (2002). Expression of Delta/NLef1 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 transdetermination 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 Berghe, 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.

Table S1. Nine categories of microarray experiments

Category #	Experiment	Array ID	Cy3 Channel	Cy5 Channel
1	wildtype prothoracic leg disc (wt 1st leg) vs. reference sample	5-069 5-131 5-085 AK93	reference sample reference sample reference sample wt 1st leg	wt 1st leg wt 1st leg wt 1st leg reference sample
2	wing disc vs. leg disc of the same animal (all wildtype)	5-251 5-256 6-59 5-249 5-255 6-62 5-248 5-254 6-64	wing disc prothoracic leg disc wing disc mesothoracic leg disc wing disc mesothoracic leg disc wing disc metathoracic leg disc wing disc	prothoracic leg disc wing disc prothoracic leg disc wing disc mesothoracic leg disc wing disc metathoracic leg disc wing disc metathoracic leg disc
3	wg-induced prothoracic leg disc (wg+ 1st leg disc) vs. reference sample	6-069 7-067 7-159 7-049 7-065 7-078	reference sample reference sample wg+ 1st leg disc wg+ 1st leg disc wg+ 1st leg disc reference sample	wg+ 1st leg disc wg+ 1st leg disc reference sample reference sample reference sample wg+ 1st leg disc

4	microdissected TD cells vs. surrounding Dwg cells of the same wg-induced prothoracic leg discs	7-041 7-210 7-047	TD cells Dwg cells Dwg cells	Dwg cells
5	microdissected TD cells vs. Vwg cells of the same wg-induced prothoracic leg discs	7-044 7-204 7-035 7-050 7-226	Vwg cells Vwg cells TD cells TD cells TD cells	TD cells TD cells Vwg cells Vwg cells Vwg cells
6	Dwg cells vs. Vwg cells of the same wg-induced prothoracic leg discs	7-046 7-048 7-218	Dwg cells Vwg cells Vwg cells	Vwg cells Dwg cells Dwg cells
7	TD cells vs. D cells of wildtype prothoracic leg discs	7-017 7-008	wildtype D cells wildtype D cells	TD cells TD cells
8	D cells of wildtype prothoracic leg discs vs. V cells of the same discs	7-018 7-020 7-059	wildtype D cells wildtype V cells wildtype D cells	wildtype V cells wildtype D cells wildtype V cells
9	3/4 fragment of 2nd leg imaginal disc after cultivation vs. reference sample	5-121 5-129 5-189 5-122	3/4 fragment 2nd leg disc 3/4 fragment 2nd leg disc reference sample reference sample	reference sample reference sample 3/4 fragment 2nd leg disc 3/4 fragment 2nd leg disc

Table S10. Computational search for regulatory elements

A. Supervised search for vgBE-like elements

CG# from TD list	Synonym	# Clusters	# Different Motifs	Function
CG10121	/	1	3	/
CG10336	CG10336	1	2	/
CG11822	nAcRbeta-21C	1	2	nicotinic acetylcholine-activated cation-selective channel activity; GO:0004889; ; GO:0004889; ; GO:0004889
CG1210	Pk61C	1	3	protein serine/threonine kinase activity; GO:0004674; receptor signalling protein serine/threonine kinase activity; GO:0004702; EC:2.7.1.-; ; GO:0004674; protein kinase activity; GO:000
CG12534	similar to ALR	1	3	flavin-linked sulphydryl oxidase activity; GO:0016971; ; GO:0016971
CG12840	Tsp42Ei	1	2	receptor signalling protein activity; GO:0005057; ; GO:0005057
CG13780	Pvf2	1	1	vascular endothelial growth factor receptor binding; GO:0005172; ; GO:0005172
CG14059	CG14059	1	2	/
CG14879	CG14879	1	2	/
CG15561	CG15561	1	2	/
CG17108	CG17108	1	3	acetyl-CoA carboxylase activity; GO:0003989 ; EC:6.4.1.2.; ; GO:0003989 ; EC:6.4.1.2
CG17914	yellow-b	1	2	/
CG18290	Act87E	1	1	structural constituent of cytoskeleton; GO:0005200; ; GO:0005200
CG1874	CG1874	1	1	/
CG1994	I(1)G0020	1	2	N-acetyltransferase activity; GO:0008080 ; EC:2.3.1.-; ; GO:0008080 ; EC:2.3.1.-
CG2028	CkIalpha	1	2	casein kinase I activity; GO:0004681 ; EC:2.7.1.-; ; GO:0004681 ; EC:2.7.1.-; ; GO:0004681 ; EC:2.7.1.-; receptor signalling protein serine/threonine kinase activity; GO:0004702 ; EC:2.7
CG3198	CG3198	1	2	nuclear mRNA splicing, via spliceosome; ; GO:000398
CG3200	Reg-2	1	3	wide-spectrum protease inhibitor activity; GO_0017114; ; GO:0017114
CG3568	CG3568	1	3	/
CG3830	vg	1	2	wing margin morphogenesis; ; GO:0008587; wing morphogenesis; ; GO:0007476
CG4766	CG4766	1	2	/
CG5355	CG5355	1	2	prolyl oligopeptidase activity; ; GO:0004287 ; EC:3.4.21.26; ; GO:0004287 ; EC:3.4.21.26
CG5518	sda	1	3	membrane alanyl aminopeptidase activity; GO:0004179 ; EC:3.4.11.2; ; GO:0004179 ; EC:3.4.11.2
CG5836	SF1	1	1	pre-mRNA splicing factor activity; GO:0008248; ; GO:0008248; transcription cofactor activity; GO:0003712; ; GO:0003712
CG5840	CG5840	1	1	pyrrolidine-5-carboxylate reductase activity; GO:0004735 ; EC:1.5.1.2; ; GO:0004735 ; EC:1.5.1.2
CG5913	CG5913	1	2	/
CG6044	CG6044	1	3	/
CG6743	Nup170	1	3	transporter activity; ; GO:0005215; ; GO:0005215
CG7539	Edg91	1	1	structural constituent of pupal cuticle (sensu Insecta); GO:0008011; ; GO:0008011
CG7583	CtBP	1	2	protein C-terminus binding; GO:0008022; ; GO:0008022; transcription corepressor activity; GO:0003714; ; GO:0003714
CG7804	CG7804	1	1	RNA binding; ; GO:0003723; ; GO:0003723; transcription regulator activity; ; GO:0030528; ; GO:0030528

CG8171	dup	1	3	DNA binding; Mutants affect: S-phase, embryo, imaginal disc
CG8394	CG8394	1	2	gamma-aminobutyric acid transporter activity ; GO:0005331 ; ; GO:0005331
CG8545	CG8545	1	2	nucleic acid binding ; GO:0003676 ; ; GO:0003676
CG8805	wun2	1	3	phosphatidate phosphotransferase activity ; GO:0008195 ; EC:3.1.3.4 ; ; GO:0008195 ; EC:3.1.3.4 ; ; GO:0008195 ; EC:3.1.3.4
CG9023	Drip	1	2	water transporter activity ; GO:0005372 ; ; GO:0005372; carrier activity ; GO:0005386 ; ; GO:0005386; water channel activity ; GO:0015250 ; ; GO:0005372 ; ; GO:0015250
CG11450	net	2	1,3	RNA polymerase II transcription factor activity; interacts with rho and EGFR, wing vein patterning
CG15279	CG15279	2	3,2	cation/amino acid symporter activity ; GO:0005416; neurotransmitter:sodium symporter activity ; GO:0005328 ; ; GO:0005328
CG17337	CG17337	2	1,2	/
CG17843	CG17843	2	3,1	flavin-linked sulfhydryl oxidase activity; GO:0016971; glucosidase activity; GO:0015926 ; EC:3.2.1.- ; GO:0015926 ; EC:3.2.1.- ; GO:0016971
CG2863	Nle	2	1,3	Notch signaling pathway ; GO:0007219
CG3036	CG3036	2	1,1	sodium: phosphate symporter activity ; GO:0005436 ; ; GO:0005436
CG4591	Tsp86D	2	3,2	receptor binding; GO:0005102 ; ; GO:0005102
CG5114	CG5114	3	3,3,1	/
CG6751	CG6751	2	2,1	general RNA polymerase II transcription factor activity; GO:0016251 ; ; GO:0016251

B. Unsupervised search for putative regulatory elements

GENE	Count of TATCGATA	Ratio in TD	TATCGATA Positions upstream	Function
CG10223	1	-1.334	-37	Top2; DNA binding; ; GO:0003677; RNA binding; ; GO:0003723; ATPase activity; ; GO:0003723; topoisomerase (ATP-hydrolyzing) activity; ; GO:0013918; EC:5.99.1.3.; ; GO:0
CG1512	1	-1.103	-20	cul-2; proteolysis and peptidolysis; ; GO:0006508; regulation of cell cycle; ; GO:0000074
CG15218	1	-1.164	-486	CycK; transcription regulator activity; ; GO:003058; kinase activator activity; ; GO:0019209; cyclin-dependent protein kinase activity; ; GO:0016538; EC:2.7.1.-.; ; GO:0030528; ; ; GO:0019
CG15484	1	-1.5095	-125	CG15484
CG10642	1	-1.122	-171	Kip64D; structural constituent of cytoskeleton; ; GO:0005200; microtubule motor activity; ; GO:0003777; ; ; GO:0005200; ; ; GO:0003777
CG11255	1	-1.057	-574	/
CG5553	1	-1.04345	-45	DNAPrim; nucleic acid binding; GO:0003676; DNA primase activity; GO:0003896; EC:2.7.7.-.; ; GO:0003896; EC:2.7.7.-.; ; GO:0003896; EC:2.7.7.-.; ; GO:0003
CG5837	2	-1.558	-438	/
CG6143	2	-1.064	-140	Hem; receptor binding; GO:0005102; ; GO:0005102
CG14685	1	-1.318	-96	/
CG1483	1	-1.245	-83	CG1485
CG1646	2	-1.121	-38	Map205; microtubule binding; ; GO:0008017; ; GO:0008017; ; GO:0008017
CG18572	1	-1.124	-299	CG1646; pre-mRNA splicing factor activity; GO:0008248; ; GO:0008248
CG4303	1	-1.062	-217	/
CG4954	1	-1.745	-220	Bap60; chromatin binding; ; GO:0003682; general RNA polymerase II transcription factor activity; ; GO:0016251; ; ; GO:0003743; ; GO:0003743
				eIF3-S8; translation initiation factor activity; ; GO:0003743; ; ; GO:0003743; ; GO:0003743

CG9253	1	-1.388	-47		CG9253; nucleic acid binding; ; GO:0003676; ; ATP-dependent RNA helicase activity; ; GO:0004004; ; GO:0004004
CG6945	1	-1.4435	-13		CG6945
CG9128	1	-1.138	-269		Sac1; polyphosphoinositide phosphatase activity; ; GO:0017120; ; GO:0017120
CG1053	1	-1.183	-67		CG1053
CG1190	1	-1.279	-44		CG1190; attachment of GPI anchor to protein; GO:0016255
CG5902	2	-1.093	-80	-24	CG5902; molecular function unknown; GO:0005554; ; GO:0005554
CG1242	1	-1.614	-240		Hsp83; ATPase activity, coupled; ; GO:0042623; ; GO:0042623
CG7425	1	-1.325	-140		/
CG7769	1	-1.289	-209		DDB1; damaged DNA binding; GO:0003684; ; GO:0003684
CG6801	1	-1.149	-104		/
CG8188	2	-1.139	-100	.92	CG8188; ubiquitin conjugating enzyme activity; ; GO:0004840; ligase activity; ; GO:0016874 ; EC:6.-.-.-; ; GO:0016874 ; GO:0016874
CG6120	1	-1.03835	-131		GO:0016840
CG11259	2	-1.89	-2570	-56	Tsp96F; receptor binding; GO:0005102; ; GO:0005102
CG12301	2	-1.118	-4265	-42	/
CG2135	2	-1.011	-4767	-363	CG12301; CG2135; beta-glucuronidase activity; ; GO:0004566 ; EC:3.2.1.31; ; GO:0004566 ; EC:3.2.1.31
CG5874	2	-1.176	-4605	-116	CG5874; RNA binding; GO:0003723; ; GO:0003723
CG1810	3	-1.1	-4789	-560	mRNA capping-enzyme; protein tyrosine/serine/threonine phosphatase activity; ; GO:0008138 ; EC:3.1.3.-; mRNA guanylyltransfer activity; ; GO:004484 ; EC:2.7.7.50; ; GO:0008138 ; EC:3.1.3.-; ; GO:0004484 ; E
CG13900	3	-1.039	-623	-61	CG13900; poly(A) binding; ; GO:0008143; ; damaged DNA binding; ; GO:0003684; ; pre-mRNA splicing factor GO:008248; ; GO:008143; ; GO:008248
CG2165	2	-1.14	-1262	-18	CG2165; calcium-transporting ATPase activity; ; GO:0005388 ; EC:3.6.3.8; ; GO:0005388 ; EC:3.6.3.8
CG10087	2	-1.286	-1848	-390	BcDNA:GH0645J4/ RNA binding; ; GO:0003723; asparagine-tRNA ligase activity; ; GO:0004816 ; EC:6.1.1.22.; ; GO:0003723; ; RNA binding; ; GO:0003723; ; GO:0004816 ; EC:6.1.1.22.; ; GO:0003723; ; GO:0004816 ; EC:6.1.1.22.
CG13691	1	-1.2	-1996	-1771	CG13691
CG9915	1	-1.123	-1771		CG9915
CG6204	1	-1.2847	-956		CG6204; nucleic acid binding; GO:0003676; ; DNA helicase activity; ; GO:0003678; RNA helicase activity; GO:0003678
CG2666	2	-1.045	-962	-950	EC:2.7.7.-; helicase activity; GO:0004386; ; GO:0003724; ; EC:2.7.7
CG9191	1	-1.87	-1818		kv; chitin synthase activity; GO:0004100 ; EC:2.4.1.16; ; GO:0004100 ; EC:2.4.1.16; ; GO:0004100 ; EC:2.4.1.16; ; GO:0004100
CG6767	1	-1.2725	-1580		Kip61F; motor activity; ; GO:0003774; structural constituent of cytoskeleton; ; GO:0005200; microtubule motor activity; ; G
CG12325	1	-1.425	-1987		/
CG1341	1	-1.243	-1380		CG12325; cell cycle; ; GO:0007049; cytokinesis; ; GO:000910; mitosis; ; GO:0007067
CG1874	1	-1.0517	-1899		Rpt1; endopeptidase activity; ; GO:004175; ATPase activity; ; GO:0016887; ; GO:0016887
CG8502	1	-1.7881	-1887		CG1874
CG7033	2	-1.056	-2794	-1196	CG8502; structural constituent of larval cuticle (sensu Insecta); ; GO:0008010; ; GO:0008010
CG10636	1	-1.456	-4476		CG7033; ATPase activity, coupled; GO:0042623; ; GO:0042623
CG10697	1	-1.3805	-3555		TepIV; protease inhibitor activity; ; GO:0030414; ; GO:0030414
CG10954	1	-1.5	-3186		/
					Arc-P34; actin binding; ; GO:0003779; ; GO:0003779; structural constituent of cytoskeleton; ; GO:0005200

CG11822	1	-1.026	-2561			
CG5001	1	-1.0995	-2235			nAChBeta-21C; nicotinic acetyl/choline-activated cation-selective channel activity; ; GO:0004889; ; ; GO:00044889; ; ; GO:000
CG5776	1	-1.241	-3231			CG5001; defense response; ; GO:0006952; protein folding; ; GO:0006457; response to heat; ; GO:0009408; response to str
CG7294	1	-1.116	-2600			GO:0006950
CG7296	1	-1.307	-4881			CG5776; ATPase activity; ; GO:0016887; ; ; GO:0016887
CG9188	2	-1.26	-4220	-3581		CG7294
CG9246	1	-1.112	-4643			CG7296; molecular_function unknown; ; GO:0005554
CG11073	2	-1.451	-3243	-2348		CG9246; molecular_function unknown; ; GO:0005554
CG3830	1	-3.682	-4023			CG11073
CG6530	1	-1.477	-3818			vg; wing margin morphogenesis; ; GO:0008587; wing morphogenesis; ; GO:0007476
CG8414	1	-1.569	-3875			mth13; G-protein coupled receptor activity; ; GO:0004930; ; ; GO:0004930
CG8811	1	-1.146	-4240			CG8414
CG14059	1	-2.0365	-4434			muskelin; cell adhesion; ; GO:0007155; cytoskeleton organization and biogenesis; ; GO:0007010
CG17947	1	-2.06	-2135			/
CG1898	1	-1.228	-3431			alpha-Cat; actin binding; GO:0003779; cytoskeletal protein binding; GO:0008092; ; GO:0008092; ; cadherin b
CG3424	1	-1.11605	-4385			GO:0045296; structural constituent of cytoskeleton; ; GO:0005200; ; GO:0
CG6650	1	-1.081	-2348			HBS1; translation release factor activity; ; GO:003747; translation elongation factor activity; ; GO:0003746; ; GO:0003747
CG6897	1	-1.132	-4729			GO:0003747
CG11990	1	-1.058	-4250			/
CG2512	1	-1.289	-3269			CG6650
CG5913	1	-1.15885	-4977			CG6897
CG6378	2	-1.511	-4572	-3750		CG11990
CG6962	1	-1.352	-3502			alphaTub84D; GTP binding; GO:0005525; ; GO:0005525; structural constituent of cytoskeleton; ; GO:0005200; ; tubulin binding
CG8064	1	-1.2695	-4227			GO:0005200; ; GO:0015631
CG8165	1	-1.508	-2928			CG5913
CG9373	1	-1.002	-2753			Bm-40-SPARC; structural_molecule activity; GO:0005198; growth factor activity; GO:0008083; ; GO:0005198; ; GO:0008083
CG31998	1	-1.01353	-4331			CG9373; RNA binding; GO:0003723; ; GO:0003723
CG11750	1	-1.26	-3436			CG11572; CG11578
						CG11750

Table S11. *PcG* and *lama* genes function in transdetermination

Genotype	Disc: vgBE- <i>lacZ</i> (frequency)	Disc: vgBE- <i>lacZ</i> (proportion)*	Cuticle: leg to wing (frequency)
<i>y^{flp}; vgBE-LacZ/+ ; Act>>wg/+</i>	33/104 (31.7%)	<i>n</i> =23 (5%)	17/69 (25%)
<i>y^{flp}; vgBE-LacZ/+ ; Act>>wg, Scm^{+/+}/Scm^{D1}</i> (heterozygous loss of function)	28/85 (33%) <i>P</i> =0.99	<i>n</i> =28 (5%) <i>P</i> =0.17	14/40 (35%) <i>P</i> =0.25
<i>y^{flp}; vgBE-LacZ/+ ; Act>>wg, E(z)^{+/+}/E(z)⁶¹</i> (heterozygous loss of function)	41/51 (80%) <i>P</i> ≤0.001	<i>n</i> =41 (6%) <i>P</i> =0.10	14/42 (33%) <i>P</i> =0.32
<i>y^{flp}; vgBE-LacZ/+ ; Act>>wg, P^c^{+/+}/P^c^l</i> (heterozygous amorph)	29/58 (50%) <i>P</i> ≤0.05	<i>n</i> =29 (5%) <i>P</i> =0.14	9/36 (25%) <i>P</i> =0.97
<i>y^{flp}; vgBE-LacZ/+ ; Act>>wg, E(P^c)^{+/+}/E(P^c)^l</i> (heterozygous)	22/37 (59%) <i>P</i> ≤0.01	<i>n</i> =22 (2%) <i>P</i> <0.05	7/21 (33%) <i>P</i> =0.4
<i>y^{flp}; vgBE-LacZ, Su(z)2^{+/+}/Su(Z)2^l; Act>>wg/+</i> (heterozygous antimorph)	0/24 (0%) <i>P</i> ≤0.01		Ø [†]
<i>y^{flp}; vgBE-LacZ, Su(z)2^{+/+}/Su(Z)2^{l,b7}; Act>>wg/+</i> (heterozygous loss of function)	0/27 (0%) <i>P</i> ≤0.01		Ø [†]
<i>y^{flp}; vgBE-LacZ/+ ; Act>>wg, brm^{+/+}/brm²</i> (heterozygous amorph)	23/30 (77%) <i>P</i> ≤0.001	<i>n</i> =23 (2%) <i>P</i> ≤0.0001	24/41 (59%) <i>P</i> ≤0.001
<i>y^{flp}; vgBE-LacZ/+ ; Act>>wg, osa^{+/+}/osa²</i> (heterozygous hypomorph)	13/18 (72%) <i>P</i> ≤0.01	<i>n</i> =13 (3%) <i>P</i> <0.05	11/40 (28%) <i>P</i> =0.74
<i>y^{flp}; vgBE-LacZ/Act>>wg ; lama⁴¹⁰/lama⁴¹⁰</i> (homozygous loss of function)	20/67 (30%) <i>P</i> =0.71	<i>n</i> =20 (2%) <i>P</i> <0.01	5/35 (15%) <i>P</i> =0.2

*Percent of total median area.

[†]No adult survivors.

Table S2**Genes preferentially expressed in wildtype wing and leg discs**

CG Number	Synonym	Median (log2)	p-value
Wing			
CG10197	kn	-5.50	1.84E-03
CG3830	vg	-5.34	5.23E-05
CG10619	tup	-4.03	2.65E-05
CG5966	CG5966	-3.91	3.85E-06
CG8376	ap	-3.44	1.14E-08
CG4382	CG4382	-3.17	3.02E-04
CG1897	msh, Dr	-3.01	2.39E-07
CG12843	Tsp42Ei	-2.99	3.31E-04
CG12287	pdm2	-2.74	7.38E-07
CG7160	CG7160	-2.60	1.17E-03
CG10570	CG10570	-2.58	9.35E-03
CG1803	regucalcin	-2.49	2.74E-03
CG9023	Drip	-2.49	4.01E-06
CG2663	CG2663	-2.45	1.84E-03
CG3132	Ect3	-2.42	4.68E-04
CG4914	CG4914	-2.36	9.54E-07
CG9554	eya	-2.35	5.96E-05
CG9427	CG9427	-2.31	1.03E-06
CG9623	if	-2.30	9.48E-08
CG10501	amd	-2.29	3.45E-03
CG1698	CG1698	-2.17	1.60E-03
CG8501	CG8501	-2.17	1.10E-03
CG7539	Edg91	-2.13	3.93E-02
CG7924	CG7924	-2.11	4.13E-05
CG9008	CG9008	-2.08	5.72E-09
CG4766	CG4766	-2.07	4.26E-05
CG9358	Phk-3	-2.05	1.78E-05
CG5397	CG5397	-1.99	1.95E-02
CG1058	rpk	-1.98	1.45E-05
CG5518	sda	-1.89	5.45E-04
CG9307	CG9307	-1.87	1.01E-04
CG5187	Doc2	-1.72	7.42E-06
CG8404	Sox15	-1.69	8.62E-07
CG10704	toe	-1.67	1.36E-04
CG5133	Doc1	-1.60	5.40E-04
CG1089	alpha-Est5	-1.59	3.61E-02

CG5093	Doc3	-1.48	1.10E-05
CG17348	drl	-1.43	1.08E-06
CG16885	CG16885	-1.35	4.52E-05
CG10962	CG10962	-1.31	3.73E-02
CG3385	nvy	-1.29	3.36E-03
CG17044	yellow-e2	-1.29	2.32E-04
CG6044	CG6044	-1.26	3.11E-07
CG6680	CG6680	-1.25	5.45E-03
CG12023	GV1	-1.25	3.19E-05
CG9812	CG9812	-1.22	8.27E-05
CG15064	Him	-1.21	4.51E-03
CG8502	CG8502	-1.18	6.62E-05
CG4746	mab-2	-1.17	6.96E-04
CG18657	NetA	-1.15	2.77E-06
CG7052	TepII	-1.13	3.75E-05
CG30069	CG13017	-1.12	6.40E-03
CG2762	ush	-1.10	3.35E-05
CG11905	CG11905	-1.10	1.12E-04
CG7820	CAH1	-1.09	3.16E-05
CG4319	rpr	-1.07	2.52E-05
CG7811	b	-1.06	5.95E-03
CG5392	CG5392	-1.04	5.66E-04
CG11835	CG11835	-1.02	4.24E-04
CG16959	CG16959	-0.94	5.13E-06
CG17559	dnt	-0.94	9.93E-07
CG9552	rost	-0.94	2.39E-05
CG9266	CG9266	-0.88	6.57E-04
CG15093	CG15093	-0.88	8.04E-05
CG17941	ds	-0.88	2.33E-05
CG17032	CG17032	-0.86	1.66E-05
CG1794	Mmp2	-0.86	7.39E-04

Leg

CG7807	AP-2	6.68	7.37E-05
CG5893	D	5.40	2.87E-05
CG11922	fd96Cb	4.68	3.33E-04
CG6269	unc-4	4.50	3.14E-06
CG6570	lbl	4.46	2.80E-08
CG10382	wrapper	4.41	1.97E-05
CG3388	gsb	4.22	2.34E-05
CG6414	CG6414	3.93	2.11E-04
CG11354	Lim1	3.71	5.33E-04
CG18111	Obp99a	3.61	1.84E-06

CG11921	fd96Ca	3.06	3.36E-04
CG4605	Acp32CD	2.91	1.47E-04
CG5888	CG5888	2.86	3.82E-03
CG4501	bgm	2.59	4.07E-06
CG6604	H15	2.16	1.79E-03
CG10440	CG10440	1.92	8.55E-07
CG2056	CG2056	1.90	3.73E-09
CG17131	SP71	1.89	8.01E-04
CG9747	CG9747	1.86	9.47E-04
CG1004	rho	1.81	5.24E-08
CG33207	pxb	1.77	5.96E-07
CG13023	CG13023	1.74	1.59E-05
CG14191	CG14191	1.73	1.40E-02
CG3479	osp	1.68	2.00E-08
CG17943	comm	1.54	1.06E-05
CG10546	Cralbp	1.47	6.34E-04
CG2360	Ccp84Aa	1.47	1.07E-02
CG10108	phyl	1.45	3.32E-03
CG9196	CG9196	1.44	8.56E-07
CG2341	Ccp84Ad	1.43	2.30E-03
CG10241	Cyp6a17	1.43	9.21E-05
CG4559	Idgf3	1.28	2.74E-02
CG8333	HLHmgamma	1.25	1.49E-05
CG3827	sc	1.23	3.35E-06
CG9801	CG9801	1.23	2.28E-02
CG3851	odd	1.21	5.33E-06
CG15085	edl	1.17	4.28E-05
CG7088	bnb	1.16	2.23E-06
CG31869	CG17130	1.13	3.43E-04
CG11121	so	1.07	7.51E-06
CG1916	Wnt2	1.07	7.19E-08
CG13082	CG13082	1.02	5.50E-05
CG3649	CG3649	1.02	1.63E-05
CG9336	CG9336	1.02	4.82E-04
CG7194	CG7194	1.02	1.13E-06
CG17211	CG17211	0.98	1.40E-03
CG6234	CG6234	0.97	2.56E-02
CG32150	CG12712	0.97	3.33E-02
CG4969	Wnt6	0.97	1.17E-05
CG8363	Paps	0.96	2.60E-05
CG10433	CG10433	0.95	2.09E-02
CG15757	CG15757	0.93	3.41E-04
CG7447	CG7447	0.92	1.25E-02
CG7231	CG7231	0.91	8.99E-06
CG7294	CG7294	0.90	2.06E-05

CG2062	Cyp4e1	0.89	4.14E-04
CG3396	Ocho	0.89	1.62E-03
CG1342	CG1342	0.89	5.07E-07
CG9674	CG9674	0.88	8.64E-05
CG11797	Obp56a	0.88	5.77E-04
CG11136	CG11136	0.87	2.02E-02
CG3629	DII	0.86	7.56E-05

Table S3

Genes preferentially expressed in dorsal or ventral cell populations of wildtype and *Wg* over-expressing leg discs

D(wt)	CG#	Synonym	median (log2)	p-value	Function
CG13067	CG13067		-3.23	4.66E-02	
CG8827	Ance		-2.97	5.27E-03	zinc ion binding; GO:0008241 ; peptidyl-dipeptidase activity; GO:0008241 ; EC:3.4.15.-; ; GO:0004246 ; EC:3.4.15.1.; ; GO:000424
CG4766	CG4766		-2.68	3.20E-02	
CG18111	Obp99a		-2.63	1.90E-02	odorant binding; GO:0005549
CG5133	Doc1		-2.36	1.22E-02	transcription factor activity; GO:0003700; ; GO:0003700
CG16820	CG16820		-2.07	1.93E-02	
CG10625	CG10625		-1.83	5.22E-03	
CG15101	Jheh1		-1.78	7.53E-03	juvenile hormone epoxide hydrolase activity; GO:0008096 ; EC:3.3.2.-; ; GO:0008096 ; EC:3.3.2.3.; ; GO:0004301 ; EC:3.3.2.3.; ; GO:000
CG3935	al		-1.77	1.44E-02	specific RNA polymerase II transcription factor activity; GO:0003704; transcription factor activity; ; GO:0003704
CG5160	CG5160		-1.73	3.67E-02	GTPase activity; GO:0003924; ; GO:0003924
CG8333	HLHmgamma		-1.62	1.54E-02	DNA binding; GO:0003677; ; transcription factor activity; GO:0003700; specific transcriptional repressor activity; GO:0016566; ; GO:0003700; ; GO:0016566
CG1650	unpg		-1.59	4.25E-02	RNA polymerase II transcription factor activity; ; GO:0003702; ; GO:0003702
CG6906	CG6906		-1.50	2.84E-02	carbonate dehydratase activity; ; GO:0004089 ; EC:4.2.1.1; ; GO:0004089 ; EC:4.2.1.1
CG15611	CG15611		-1.49	2.10E-02	
CG17914	yellow-b		-1.49	4.94E-02	
CG1698	CG1698		-1.42	4.71E-02	potassium/amino acid transporter activity; ; GO:0017032; cation transporter activity; ; GO:0008324; ; GO:0017032
CG5093	Doc3		-1.41	3.79E-03	transcription factor activity; ; GO:0003700; ; GO:0003700
CG17348	drl		-1.39	2.70E-02	transmembrane receptor protein tyrosine kinase activity; ; GO:0004714 ; EC:2.7.1.-; ; GO:0004714 ; EC:2.7.1.-; ; GO:0004713 ; EC:2.7.1.112; ; GO:0004713 ;
CG17045	yellow-e3		-1.37	2.11E-03	
CG3244	CG3244		-1.36	1.91E-02	
CG9656	grn		-1.34	3.11E-02	RNA polymerase II transcription factor activity; GO:0003702; general RNA polymerase II transcription factor activity; GO:0016251; ; GO:0003702; ; GO:0016251; ; GO:0016251
CG4914	CG4914		-1.29	2.54E-03	trypsin activity; ; GO:0004295 ; EC:3.4.21.4; ; GO:0004252 ; EC:3.4.21.-
CG10119	LamC		-1.24	2.28E-02	structural constituent of cytoskeleton; ; GO:0005200; ; GO:0005200
CG12287	pdm2		-1.11	4.50E-02	DNA binding; ; GO:0003677; ; GO:0003677; specific RNA polymerase II transcription factor activity; ; GO:0003704; ; GO:0003704
CG9885	dpp		-1.10	1.12E-02	signal transducer activity; ; GO:0004871; morphogen activity; ; GO:0016015; ; GO:0005160; ; GO:0005160; ; GO:0016015; ; GO:0004713
CG4746	mab-2		-1.09	3.95E-02	transforming growth factor beta receptor binding; ; GO:0005160; ; GO:0005160; ; GO:0016015; ;
CG10200	CG10200		-1.07	2.84E-03	encodes Mab-21, involved in cell fate determination

CG12840	Tsp42E1	-1.07	2.57E-02	receptor signaling protein activity; ; GO:0005057; ; GO:0005057
CG8483	CG8483	-1.06	2.01E-02	defense response; GO:0006952
CG9593	CG9593	-1.03	7.28E-03	receptor binding; GO:0005102; ; GO:0005102
CG6868	tld	-1.02	1.24E-02	procollagen C-endopeptidase activity; GO:0017026 ; EC:3.4.24.19; metalloendopeptidase activity; GO:0004222 ; EC:3.4.24.-; ; GO:0017026 ; EC:3.4.24.19; ; GO:0004222 ; EC:3.4.24.-
CG5912	arr	-1.00	3.23E-02	low-density lipoprotein receptor activity; ; GO:0005041; ; GO:0005041; ; GO:0005041; ; GO:0005041; ; GO:0005041
CG2125	ci	-0.97	2.15E-02	specific RNA polymerase II transcription factor activity; ; GO:0003704; transcription factor activity; ; GO:0003700; transcriptional activator activity; ; GO:0016563; transcriptional represso
CG17559	dnt	-0.96	1.06E-02	protein-tyrosine kinase activity; ; GO:0004713 ; EC:2.7.1.112; ; GO:0004713 ; EC:2.7.1.112; transmembrane receptor protein tyrosine kinase activity; ; GO:0004714 ; EC:2.7.1.-; ; GO:0004714
CG14843	CG14843	-0.96	1.32E-02	/
CG17058	Peritrophin-A	-0.95	2.24E-02	chitin binding; ; GO:0008061; ; GO:0008061; structural constituent of peritrophic membrane (sensu Insecta); ; GO:0016490; ; ; GO:0016490
CG7722	CG7722	-0.92	3.61E-03	serine-type endopeptidase inhibitor activity; ; GO:0004867; ; GO:0004867
CG7860	CG7860	-0.90	4.58E-02	asparagine activity; ; GO:0004067 ; EC:3.5.1.1; ; GO:0004067 ; EC:3.5.1.1
CG14946	CG14946	-0.86	2.20E-03	oxidoreductase activity; ; GO:0016491 ; EC:1.-.-.-; ; GO:0016491 ; EC:1.-.-.-
CG15905	CG15905	-0.86	2.21E-02	/

D(wg) CG#	Synonym	median (log2)	p-value	Function
CG7539	Edg91	-3.03	3.18E-02	structural constituent of pupal cuticle (sensu Insecta); GO:0008011; ; GO:0008011
CG7892	nmo	-1.65	4.14E-02	protein serine/threonine kinase activity; ; GO:0004674; receptor signaling protein serine/threonine kinase activity; ; GO:0004702 ; EC:2.7.1.-; ; GO:0004674; ; ; GO:0004674; ; G
CG14598	CG14598	-1.63	4.54E-03	/
CG6906	CG6906	-1.37	1.23E-02	carbonate dehydratase activity; ; GO:0004089 ; EC:4.2.1.1; ; GO:0004089 ; EC:4.2.1.1
CG8846	Thor	-1.24	3.11E-03	eukaryotic initiation factor 4E binding; ; GO:0008190; ; ; GO:0008190; ; ; GO:0008190
CG1956	R	-1.23	4.42E-02	GTP binding; ; GO:0005525; GDP binding; ; GO:0019003; GTPase activity; ; GO:0003924; ; ; GO:0019003; ; ; GO:0003924
CG9425	CG9425	-1.16	3.63E-02	nucleic acid binding; ; GO:0003676; ; DNA helicase activity; ; GO:0003678; helicase activity; ; GO:0004386; ; ; GO:0004386; ; ; GO:0003678
CG13679	CG13679	-1.11	1.24E-02	/
CG1691	Imp	-1.10	1.13E-02	mRNA binding; GO:0003729; ; GO:0003729
CG15629	CG15629	-1.05	3.19E-02	oxidoreductase activity; ; GO:0016491 ; EC:1.-.-.-; ; GO:0016491 ; EC:1.-.-.-
CG3935	al	-1.02	3.68E-02	specific RNA polymerase II transcription factor activity; ; GO:0003704; transcription factor activity; ; GO:0003700; ; ; GO:0003700; ; ; GO:0003704
CG9192	CG9192	-1.01	2.01E-02	/
CG14191	CG14191	-0.88	2.83E-02	/

V(wt) CG#	Synonym	median (log2)	p-value	Function
CG4889	wg	4.80	1.88E-02	frizzled-2 binding; ; GO:0005110; ; ; GO:0005110; receptor binding; ; GO:0005102; Notch binding; ; GO:0005112; signal transducer activity; ; GO:0004871; morphogen activity; ; GO:00160

CG4969	Wnt6	3.83	4.55E-03	receptor binding; ; GO:0005102; ; GO:0004871; ; GO:0005102; signal transducer activity; ; GO:0004871; ;
CG6570	lbl	3.30	9.89E-03	specific RNA polymerase II transcription factor activity; GO:0003704; ; GO:0003704
CG16785	fz3	2.62	5.25E-03	Wnt-protein binding; ; GO:0017147; transmembrane receptor activity; ; GO:0004888; Wnt receptor activity; ; GO:0042813; G-protein coupled receptor activity; ; GO:0004930; ; GO:0004888; ; ; GO:0004
CG13857	CG13857	2.36	4.06E-02	/
CG4559	Idgf3	2.33	2.08E-02	imaginal disc growth factor activity; ; GO:0008084; NOT chitinase activity; ; GO:0004568 ; EC:3.2.1.14; hydrolase activity, hydrolyzing N-glycosyl compounds; ; GO:0016799 ; EC:3.2.2.-; ; GO:0
CG7777	CG7777	2.28	2.86E-03	water transporter activity; ; GO:0005372; carrier activity; ; GO:0005385; binding; ; GO:0005488; ; ; GO:0005372
CG8376	ap	2.18	3.36E-02	zinc ion binding; ; GO:0008270; specific RNA polymerase II transcription factor activity; ; GO:0003704; ; ; GO:0008270
CG3388	gsb	2.16	5.58E-03	specific RNA polymerase II transcription factor activity; ; GO:0003704; ; ; GO:0003704
CG11922	fd96cb	2.15	9.71E-03	transcription factor activity; GO:0003700; ; GO:0003700
CG11921	fd96Ca	1.92	1.72E-02	transcription factor activity; GO:0003700; ; GO:0003700
CG14933	CG14933	1.80	7.11E-04	/
CG1897	msh, Dr	1.75	8.12E-03	specific RNA polymerase II transcription factor activity; GO:0003704; ; GO:0003704
CG9008	CG9008	1.63	2.36E-02	isomerase activity; ; GO:0016853 ; EC:5.-.-.-; racemase and epimerase activity; ; GO:0016854 ; EC:5.1.-.-
CG9747	CG9747	1.47	1.06E-03	-; ; GO:0016853 ; EC:5.-.-.-; ; GO:0016854 ; EC:5.1.-.-
CG6604	H15	1.44	1.63E-02	transcription factor activity; ; GO:0003700; ; GO:0003700
CG8967	otk	1.37	2.82E-03	transmembrane receptor protein tyrosine kinase activity; ; GO:0004714 ; EC:2.7.1.-; receptor activity; ; GO:0004872; ; GO:0004872; semaphorin receptor binding; ; GO:0030215; protein-tyrosine
CG14598	CG14598	1.33	7.87E-03	/
CG11121	so	1.18	9.97E-03	transcription factor activity; ; GO:0003700; ; RNA polymerase II transcription factor activity; ; GO:0003702
CG15209	CG15209	1.15	2.63E-02	/
CG9338	CG9338	1.12	4.47E-02	/
CG2056	CG2056	1.11	2.49E-02	trypsin activity; ; GO:0004295 ; EC:3.4.21.4; serine-type endopeptidase activity; GO:0004252 ; EC:3.4.21.-; ; GO:0004295 ; EC:3.4.21.4; ; GO:0004252 ; EC:3.4.21.-
CG1743	Gs2	1.09	4.24E-02	glutamate:ammonia ligase activity; ; GO:0004356 ; EC:6.3.1.2; ; ; GO:0004356 ; EC:6.3.1.2; ; ;
CG3629	DII	1.04	4.53E-02	specific RNA polymerase II transcription factor activity; ; GO:0003704; ; ; GO:0003704
CG4637	hh	1.00	3.45E-03	smoothened binding; ; GO:0005119; ; patched binding; ; GO:0005113; ; GO:0005119; ; GO:0005113; signal transducer activity; ; GO:0004871; morphogen activity; ; GO:0016015; cysteine-type endopeptid
CG8964	CG8964	0.99	1.28E-02	non-membrane spanning protein tyrosine kinase activity; ; GO:0004715 ; EC:2.7.1.-; receptor activity; ; GO:0004872; ; transmembrane receptor prote
CG9015	en	0.99	5.83E-03	transcription regulator activity; ; GO:0030528; transcriptional repressor activity; ; GO:0016564; specific
CG6206	CG6206	0.99	2.70E-02	GO:0004872; ; hydrolyase activity, hydrolyzing N-glycosyl compounds; ; GO:0016799 ; EC:3.2.2.-; ; ; GO:0016799 ; EC:3.2.2.-; alpha-mannosidase activity; ; GO:0004559 ; EC:3.2.1.2
CG9653	brk	0.96	4.18E-02	RNA polymerase II transcription factor activity; GO:0003702; ; GO:0003704; ; ; GO:0003704; ; GO:0
CG8927	CG8927	0.91	1.20E-02	/
CG6173	kal-1	0.89	3.93E-02	cell adhesion; ; GO:0007155; transmission of nerve impulse; GO:0019226
CG9614	pip	0.87	1.13E-02	heparin-sulfate 2-sulfotransferase activity; GO:0004394 ; EC:2.8.2.-; ; GO:0004394 ; EC:2.8.2.-; ; GO:0008

Table S4**Genes preferentially expressed in regenerating cells**

CG #	Synonym	median (\log_2) p-value	Molecular Function
CG4250	CG4250	-6.03	3.00E-04
CG13053	CG13053	-4.40	4.80E-03
CG10112	CG10112	-4.39	6.00E-05
CG8846	Thor, Phas1	-3.95	5.00E-04
CG15212	CG15212	-3.77	1.50E-03
CG1148	Osi2	-3.30	6.90E-03
CG9812	CG9812	-3.08	1.00E-04
CG15532	hd_c	-2.87	1.39E-02
CG10131	CG10131	-2.72	2.48E-02
CG3752	CG3752	-2.62	8.70E-03
CG1803	regucalcin	-2.59	6.00E-03
CG3699	CG3699	-2.41	8.80E-03
CG5575	ken	-2.37	2.00E-03
CG4822	CG4822	-2.26	3.36E-02
CG4145	Cg25C	-2.24	1.61E-02
CG1102	CG1102	-2.12	1.77E-02
CG2150	CG2150	-2.01	5.00E-02
CG1762	betaInt-nu	-1.94	3.97E-02
CG8788	CG8788	-1.92	9.70E-03
ND3	ND3	-1.89	3.42E-02
CG6899	Ptp4E	-1.81	2.40E-03
CG17245	plexB	-1.81	4.00E-04
CG11784	CG11784	-1.79	2.60E-03

CG1368	CG1368	-1.78	4.10E-03	structural constituent of chorion (sensu Insecta) ; GO:0005213 ; ; EC:2.5.1.18	
CG17299	SNF4Agamma	-1.77	1.28E-02	receptor signalling protein serine/threonine kinase activity; GO:0004702 ; EC:2.7.1.-; protein serine/threonine kinase activity; GO:004674 ; GO:0004672 ; EC:2.7.1.-; AMP- /	GO:0005213
CG4686	CG4686	-1.75	2.50E-03	/	
CG10211	CG10211	-1.72	1.09E-02	peroxidase activity; ; GO:0004601 ; EC:1.11.1.7; ; GO:0004601 ; EC:1.11.1.7	
CG8245	CG8245	-1.71	1.27E-02	/	
CG4807	ab, abrupt	-1.70	3.60E-03	specific RNA polymerase II transcription factor activity; ; GO:0003704; ; GO:0003704.	GO:0003704.
CG9358	Phk-3	-1.70	2.19E-02	diacylglycerol binding; ; GO:0019992; ; GO:0019992; protein serine/threonine kinase activity; ; GO:004674; ; GO:0004674; protein kinase activity; ; GO:0004672 ; EC:2.7.1.37; carrier activity	
CG8066	CG8066	-1.68	2.50E-02	cysteine protease inhibitor activity; GO:004869; ; GO:004869	
CG3629	DII	-1.67	7.60E-03	specific RNA polymerase II transcription factor activity; ; GO:0003704; ; GO:0003704.	GO:0003704.
CG8727	cyc	-1.65	3.45E-02	specific RNA polymerase II transcription factor activity; GO:0003704; ; RNA polymerase II transcription factor activity; GO:0003702; ; GO:0003702; ; GO:0003704	
CG12822	CG12822	-1.64	1.32E-02	/	
CG9689	CG9689	-1.61	8.20E-03	/	
CG18507	CG18507	-1.58	1.05E-02	/	
CG5677	CG5677	-1.56	2.64E-02	signal peptidase activity; GO:0009003; ; GO:0009003	
CG7267	CG7267	-1.55	1.05E-02	/	
CG7724	CG7724	-1.55	2.45E-02	oxidoreductase activity, acting on CH-OH group of donors; ; GO:0016614 ; EC:1.1.-.-; ; GO:0016614 ; EC:1.1.-.-	
CG15611	CG15611	-1.52	9.60E-03	/	
CG6191	CG6191	-1.52	1.96E-02	/	
CG5518	sda	-1.52	3.00E-04	membrane aranyl aminopeptidase activity; GO:0004179 ; EC:3.4.11.2; ; GO:0004179 ; EC:3.4.11.2	
CG3831	CG3831	-1.49	2.13E-02	/	
CG6357	CG6357	-1.47	1.20E-03	cathepsin L activity; ; GO:0004217 ; EC:3.4.22.15; ; GO:0004217 ; EC:3.4.22.15	
CG32549 = new ID in v4, CG6247	CG32549 = new ID in v4, CG6247	-1.47	1.42E-02	CG3549, 5'-nucleotidase activity' GO:0008253 ; EC:3.1.3.5; ; GO:0008253 ; EC:3.1.3.5	
CyP4e1	Cyp4e1	-1.44	1.71E-02	electron transporter activity; ; GO:0005489 ; ; GO:0005489; oxidoreductase activity; ; GO:0016491 ; EC:1.-.-.-; ; GO:0016491 ; EC:1.-.-.-; ; GO:0008253 ; GO:0005489	
CG4778	CG4778	-1.43	8.30E-03	structural constituent of peritrophic membrane (sensu Insecta); ; GO:0016490 ; ; GO:0016490	
CG4382	CG4382	-1.42	9.90E-03	carboxylesterase activity; ; GO:0004091 ; EC:3.1.1.1; ; GO:0004091 ; EC:3.1.1.1	
CG8790	CG8790	-1.42	4.40E-03	carrier activity; GO:005386; dicarboxylic acid transporter activity;	
CG13044	CG13044	-1.41	4.00E-04	/	
CG3704	CG3704	-1.41	2.48E-02	purine nucleotide binding; ; GO:0017076; ; GO:0017076	
CG10126	CG10126	-1.40	2.61E-02	calcium ion binding; GO:0005509; ; GO:0005509	
CG10916	CG10916	-1.39	4.90E-03	/	

CG13430	CG13430	-1.38	1.81E-02	trypsin activity; ; GO:0004295 ; EC:3.4.21.4; ; GO:0004295 ; EC:3.4.21.4; ;
CG9338	CG9338	-1.36	4.30E-02	/
CG17035	GXIVsPLA2	-1.35	4.90E-02	phospholipase A2 activity; ; GO:0004623 ; EC:3.1.1.4; ; GO:0004623 ;
CG7860	CG7860	-1.34	9.40E-03	EC:3.1.1.4 asparaginase activity; ; GO:0004067 ; EC:3.5.1.1; ; GO:0004067 ; EC:3.5.1.1
CG14812	CG14812	-1.33	1.97E-02	/
CG6667	dI	-1.32	2.85E-02	DNA binding; ; GO:0003677; ; GO:0003677; transcription factor activity; ; GO:0003677; transcriptional activator activity; ; GO:0016563; RNA polymerase II transcription factor activity; ; GO:000
CG12164	CG12164	-1.31	4.16E-02	/
CG1572	CG1572	-1.29	2.25E-02	/
CG11143	Inos	-1.27	4.60E-03	Inositol-3-phosphate synthase activity; ; GO:0004512 ; EC:5.1.4; ;
CG4592	CG4592	-1.26	2.44E-02	GO:0004512 ; EC:5.1.4; ; GO:0004512 ; EC:4.2.1.-; ; GO:0016836 ; EC:4.2.1.-;
CG9134	CG9134	-1.25	1.52E-02	GO:00044165 ; EC:5.3.3.8 dodecenoyl-CoA delta-isomerase activity; ; GO:00044165 ; EC:5.3.3.8; ;
CG10119	LamC	-1.24	4.87E-02	sugar binding; ; GO:0005529; ; GO:0005529
CG8289	CG8289	-1.24	1.90E-03	/
CG3712	mRpl33	-1.23	4.54E-02	structural constituent of cytoskeleton; ; GO:0005200; ; GO:0005200
CG4380	usp	-1.19	4.10E-02	structural constituent of ribosome; ; GO:0003735; ; ; GO:0003735
CG4096	CG4096	-1.18	3.12E-02	DNA binding; ; GO:0003677; juvenile hormone binding; ; GO:0005500; ; ;
CG10071	Rpl29	-1.18	2.99E-02	GO:0003677; ecdysteroid hormone receptor activity; ; GO:0004884; ;
CG12023	GW1	-1.18	2.88E-02	GO:0004884; ligand-dependent nuclear receptor activity;
CG1516	CG1516	-1.18	2.33E-02	metalloendopeptidase activity; ; GO:0004222 ; EC:3.4.24.-; ; ; GO:0004222 ;
CG11077	CG11077	-1.16	1.20E-03	EC:3.4.24.- structural constituent of ribosome; ; GO:0003735; ; ; GO:0003735; ; ;
CG7224	CG7224	-1.16	4.58E-02	GO:0003735; ; ; GO:0003735
ND4	ND4	-1.14	3.60E-02	/
COX3	COX3	-1.14	4.90E-02	NADH dehydrogenase subunit 4, mitochondrial NADH-ubiquinone oxidoreductase chain 4, abbreviated as mt:ND4
CG7291	NPC2	-1.12	4.50E-02	mitochondrial Cytochrome c oxidase subunit III , abbreviated as mt:CoII
CG1402	CG1402	-1.12	1.25E-02	receptor binding; ; GO:0005102; ; GO:0005102
CG9535	CG9535	-1.11	4.58E-02	carbonate hydratase activity; GO:0004089 ; EC:4.2.1.1; ; GO:0004089 ;
CG13660	CG31370 = new ID in v4, CG13660	-1.10	1.66E-02	EC:4.2.1.1 UDP-N-acetylglucosamine diphosphorylase activity; ; GO:0003977 ;
CG11207	feo	-1.10	3.53E-02	EC:2.7.7.23; ; GO:0003977 ; EC:2.7.7.23 Cg31370
CG9704	Nrt	-1.07	2.10E-03	receptor binding; GO:0005102; ; GO:0005102 cell adhesion; ; GO:0007155; ectoderm development; ; GO:0007398;
CG12340	CG12340	-1.07	4.89E-02	neurogenesis; ; GO:0007399; amine receptor activity; ; GO:0004989 ;
CG11141	CG11141	-1.06	1.15E-02	octopamine receptor activity; ; GO:0004989 ; ; GO:0008227; ; ; GO:0004989

CG17033	CG17033	-1.06	6.40E-03	/
CG6999	CG6999	-1.06	4.90E-02	/
CG1827	CG1827	-1.05	3.70E-03	4-(beta-N-acetylglucosaminy)-L-asparaginase activity; ; GO:0003948 ; EC:3.5.1.26 ; ; GO:0003948 ; EC:3.5.1.26
CG11562	CG11562	-1.04	1.23E-02	/
CG4860	CG4860	-1.04	1.00E-03	acyl-CoA dehydrogenase activity; GO:0003995 ; EC:1.3.99.3 ; ; GO:0003995 ; EC:1.3.99.3
CG12238	((1)G0084	-1.03	3.00E-02	/
CG3937	cher	-1.02	6.50E-03	actin binding; GO:0003779 ; GO:0003779 ; structural constituent of cytoskeleton; GO:0005200 ; ; GO:0003779
CG18657	NetA	-1.02	4.66E-02	structural molecule activity; ; GO:0005198 ; ; GO:0005198
CG6575	glec	-1.02	1.50E-03	carbohydrate binding; GO:0030246 ; ; GO:0030246
CG6073	CG6073	-1.01	1.75E-02	/
CG8980	NiPP1	-1.00	4.69E-02	RNA binding; ; GO:0003723 ; ; GO:0004864 ; ; GO:0004864; type 1
CG15802	not found in v4, CG15802	-1.00	4.60E-03	RNA polymerase II transcription factor; ; GO:0003702
CG5753	stau	-0.99	4.20E-03	RNA binding; ; GO:0003723; double-stranded RNA binding; ; GO:0003725; ; ;
CG5571	CG5571	-0.99	6.10E-03	GO:0003723; microtubule binding; ; GO:0008017; ; ; GO:0008017; ; ;
CG17059	CG17059	-0.99	1.09E-02	GO:0003725; mRNA 3'-UTR bind
CG4260	alpha-Adaptin	-0.99	4.70E-03	/
CG12874	CG12874	-0.99	3.69E-02	asymmetric cytokinesis; ; GO:0008356; neurotransmitter secretion; ;
CG6695	CG6695	-0.98	1.20E-03	GO:0003723; ; GO:0003723
CG5887	desat1	-0.98	3.75E-02	RNA binding; GO:0004768 ; EC:1.14.19.1 ; ; GO:0004768 ; EC:1.14.19.1 ; ; GO:0004768 ; EC:1.14.19.1
CG11025	isopeptidase-T-3	-0.97	1.69E-02	ubiquitin thiolesterase activity; ; GO:0004221 ; EC:3.1.2.15 ; ; GO:0004221 ; EC:3.1.2.15
CG8430	G0t1	-0.96	1.87E-02	aspartate transaminase activity; ; GO:0004069 ; EC:2.6.1.1 ; ; GO:0004069 ; EC:2.6.1.1
CG12840	Tsp42E1	-0.96	3.80E-02	stearoyl-CoA 9-desaturase activity; GO:0004768 ; EC:1.14.19.1
CG10863	CG10863	-0.96	2.75E-02	receptor signaling protein activity; ; GO:0005057; ; ; GO:0005057
CG6528	not found in v4, CG6528	-0.95	6.00E-04	aldehyde reductase activity; ; GO:0004032 ; EC:1.1.1.21 ; ; GO:0004032 ;
CG3523	CG3523	-0.95	4.83E-02	#N/A
CG10800	Rca1	-0.94	7.20E-03	fatty-acid synthase activity; ; GO:0004312 ; EC:2.3.1.85
CG4698	Wnt4	-0.94	8.30E-03	eye-antennal disc metamorphosis; ; GO:0007455; regulation of mitosis; ;
CG4118	nxf2	-0.93	1.58E-02	GO:0005102; ; ; GO:0005102; ; ; GO:0004871
CG6117	Pka-C3	-0.93	1.36E-02	mRNA-nucleus export; ; GO:0006406, transporter activity; ; GO:0005215; ; ;
				receptor signalling protein serine/threonine kinase activity; ; GO:004702; ; ;
				EC:2.7.1.-; protein serine/threonine kinase activity; ; GO:0004674; cAMP-

				dependent protein kinase activity; ; GO:00
CG6673	CG6673	-0.93	2.24E-02	glutathione transferase activity; ; GO:004364 ; EC:2.5.1.18
CG6494	h	-0.93	2.93E-02	specific transcriptional repressor activity; ; GO:0016566; general transcriptional repressor activity; ; GO:0016565; specific RNA polymerase II transcription factor activity; ; GO:003704; ;
CG2948	rev7	-0.91	2.19E-02	ephrin receptor binding; GO:0046875; ; GO:0046875
CG12299	CG12299	-0.91	1.20E-03	transcription regulator activity; ; GO:0030528; ; GO:0030528
CG5576	imd	-0.91	2.42E-02	antibacterial humoral response (sensu Protostomia); ; GO:0006961; ; GO:0006963; antimicrobial humoral response (sensu Protostomia); ; GO:0006960; defense response; ; GO:0006952; defense response to Gram-negative bacteria; ; GO:0042742; humoral immune response; ; GO:0006959; immune response; ; GO:0006955; response to bacteria; ; GO:0009617; signal transduction; ; GO:0007165
CG9914	CG9914	-0.90	4.18E-02	structural molecule activity; ; GO:0005198; oxidoreductase activity; ;
CG6454	CG6454	-0.90	4.91E-02	/ ; GO:0016491 ; EC:1.-.-.-; ; GO:0005198; ; GO:0005198
CG1618	comt	-0.89	8.90E-03	ATPase activity; ; GO:0016887; ; ; GO:0016887
CG3359	mf1as	-0.89	1.10E-03	axonogenesis; GO:0007409; cell-cell adhesion; GO:0016337; ectoderm development; GO:0007398; signal transduction; GO:0007165
CG17725	Pepck	-0.89	1.65E-02	phosphoenolpyruvate carboxykinase (GTP) activity; ; GO:0004613 ; EC:4.1.1.32; ; GO:0004613 ; EC:4.1.1.32
CG17967	DebA	-0.88	1.37E-02	/
CG18343	CG18343	-0.88	4.40E-03	/
CG4210	CG4210	-0.87	1.29E-02	/
CG9438	Cyp6a2	-0.87	1.59E-02	electron transporter activity; ; GO:0005489; ; GO:0005489; ; GO:0005489; ; GO:0005489
CG9091	Rpl37a	-0.87	2.80E-02	nucleic acid binding; ; GO:0003676; structural constituent of ribosome; ; GO:0003735; ; GO:0003676; ; GO:0003735
CG5025	Sps2	-0.85	4.96E-02	selenide, water dikinase activity; ; GO:0004756 ; EC:2.7.9.3.; ; GO:0004756 ; EC:2.7.9.3.; ; GO:0004756 ; GO:0017076; ; ; GO:0017076; ; GO:0004756 ; EC: ; GO:0004756 ; EC: ; GO:0005083
CG1093	plx	-0.85	4.28E-02	small GTPase regulatory/interacting protein activity; ; GO:0005083; ; GO:0005083

Table S5. Genes preferentially expressed in TD cells**Overlap of TD genes with other lists**

TD & D(wg)	TD & D(wt)	TD & REG	TD & wing
CG14191	CG4766	CG10863	CG10570
CG15629	CG6906	CG12840	CG1698
CG6906	CG15611	CG15611	CG1794
CG7539	CG17914	CG1572	CG1897
CG1598	CG3359	CG3830	CG4319
CG4914	CG5518	CG4746	CG4766
CG985, dpp		CG4914	
CG4746		CG4914	
CG1840		CG5518	
CG6868		CG6044	
CG7722		CG7539	
TD & V(wg)	TD & V(wt)	CG8404	CG8404
CG6530	CG1597	CG9023	CG9307
CG2918			
CG4027			
CG54567			

Genes enriched in TD cells

confidence level >95%, median >0.85 (log2)

CG #	Synonym	Median (log2)	p-value	Function
CG14059	CG14059	-4.74	2.01E-04	/
CG5993	upd, os	-3.62	1.36E-04	JAK-STAT cascade; GO:0007259; blastoderm segmentation; GO:0007350; border cell migration (sensu Insecta); GO:0007298; defense response; GO:0006952; embryonic development (sensu Insecta); GO:0001700; foregut morphogenesis; GO:0042078; hindgut morphogenesis; GO:0042079; germ-line stem cell division; GO:0030707; primary sex determination; GO:0007538; sex determination; GO:0007530; sex determination; ovarian follicle cell development (sensu Insecta); GO:0007540; somatic stem cell division; GO:0048103; stem cell division; GO:0017145
CG3830	vg	-3.33	4.46E-04	microtubule-based movement; GO:0007018; wing margin morphogenesis; GO:0008587; wing morphogenesis; GO:0007476
CG6816	Cyp.8a1	-3.24	5.53E-03	electron transport; GO:0006118; steroid biosynthesis; GO:0006604
CG15279	CG15279	-2.89	5.43E-04	amino acid metabolism; GO:0006520; amino acid transport; GO:0006865; cation transport; GO:0006612; extracellular transport; GO:0006858; neurotransmitter transport; GO:0006836
CG4746	mab-2	-2.83	1.67E-03	/
CG3045	CG3686	-2.59	2.41E-04	amino acid metabolism; GO:0006520; transmission of nerve impulse; GO:0019226
CG8394	CG8394	-2.58	8.17E-05	amino acid transport; GO:0006520; amino acid transport; GO:0006865
CG2198	Ama	-2.38	1.05E-06	cell adhesion; GO:0007155; cell-cell adhesion; GO:0016337; signal transduction; GO:0007165
CG6840	Sox15	-2.36	6.67E-07	ectoderm development; GO:0007398; neurogenesis; GO:0007399; regulation of transcription from Pol II promoter; GO:0045449; regulation of transcription from Pol II promoter; GO:0006418
CG5158	nAchbeta-21C	-2.33	8.94E-04	tRNA aminacylation for protein translation; GO:0006527; tRNA aminacylation for protein translation; GO:0006508
CG11822		-2.23	3.37E-04	ion transport; GO:0006811; transport; GO:0006810
CG4914	CG4914	-2.21	1.42E-05	proteolysis and peptidolysis; GO:0006508; regulation of transcription, DNA-dependent; GO:0006355
CG3359	mfafs	-2.11	1.83E-03	axonogenesis; GO:0007409; cell-cell adhesion; GO:0016337; ectoderm development; GO:0007398; muscle development; GO:0007517; nervous system development; GO:0007417; dorsal/ventral pattern formation; GO:0009953; dorsal/ventral pattern formation; GO:0007165
CG9307		-2.11	8.53E-04	chitin metabolism; GO:0006050; polysaccharide metabolism; GO:0006357; ventral cord development; GO:0007419; wing morphogenesis; GO:0007476
CG6906		-2.08	2.31E-03	one-carbon compound metabolism; GO:0006730
CG4859	Mmp1	-2.06	2.90E-05	anti-apoptosis; GO:0006916; autophagic cell death; GO:0035001; larval development (sensu Insecta); GO:0002168; proteolysis and peptidolysis; GO:0006508; salivary gland cell death; GO:0035071; tracheal system development (sensu Insecta); GO:0007424
CG7722	wun2	-2.04	2.35E-05	brain development; GO:0007420; central nervous system development; GO:0007430; cell adhesion; GO:0007398; muscle development; GO:0007389; pattern specification; GO:0007389; pattern specification from Po II promoter; GO:0006337; ventral cord development; GO:0007419; wing morphogenesis; GO:0007476
CG10570		-1.89	4.23E-03	G-protein coupled receptor protein signaling pathway; GO:0007186; dephosphorylation; GO:0016311; germ cell migration; GO:0008554; germ cell programmed cell death; GO:0035233; lipid metabolism; GO:0006629
CG7447	CG7447	-1.85	1.12E-04	/
CG5739	Edg91	-1.83	2.83E-03	/
CG17914	yellow-b	-1.82	3.70E-03	/
CG10121	SP1173	-1.79	9.58E-05	/
CG15484	CG15484	-1.76	1.82E-04	/
CG9023	Drip	-1.74	1.42E-03	cell homeostasis; GO:0019725; water homeostasis; GO:0030104; water transport; GO:0006833
CG15629		-1.71	2.91E-04	metabolism; GO:0008152; visual perception; GO:0007601
CG5840	CG5840	-1.71	1.20E-04	amino acid biosynthesis; GO:0008652; proline biosynthesis; GO:00066561
CG12840	Tsp42El	-1.68	4.23E-03	ectoderm development; GO:0007398; neurogenesis; GO:0007399; transmission of nerve impulse; GO:0019226
CG15611	CG15611	-1.65	7.08E-04	/
CG3983		-1.62	7.65E-04	cell surface receptor linked signal transduction; GO:0035234; germ cell repulsion; GO:00035233; lipid metabolism; GO:0007242;

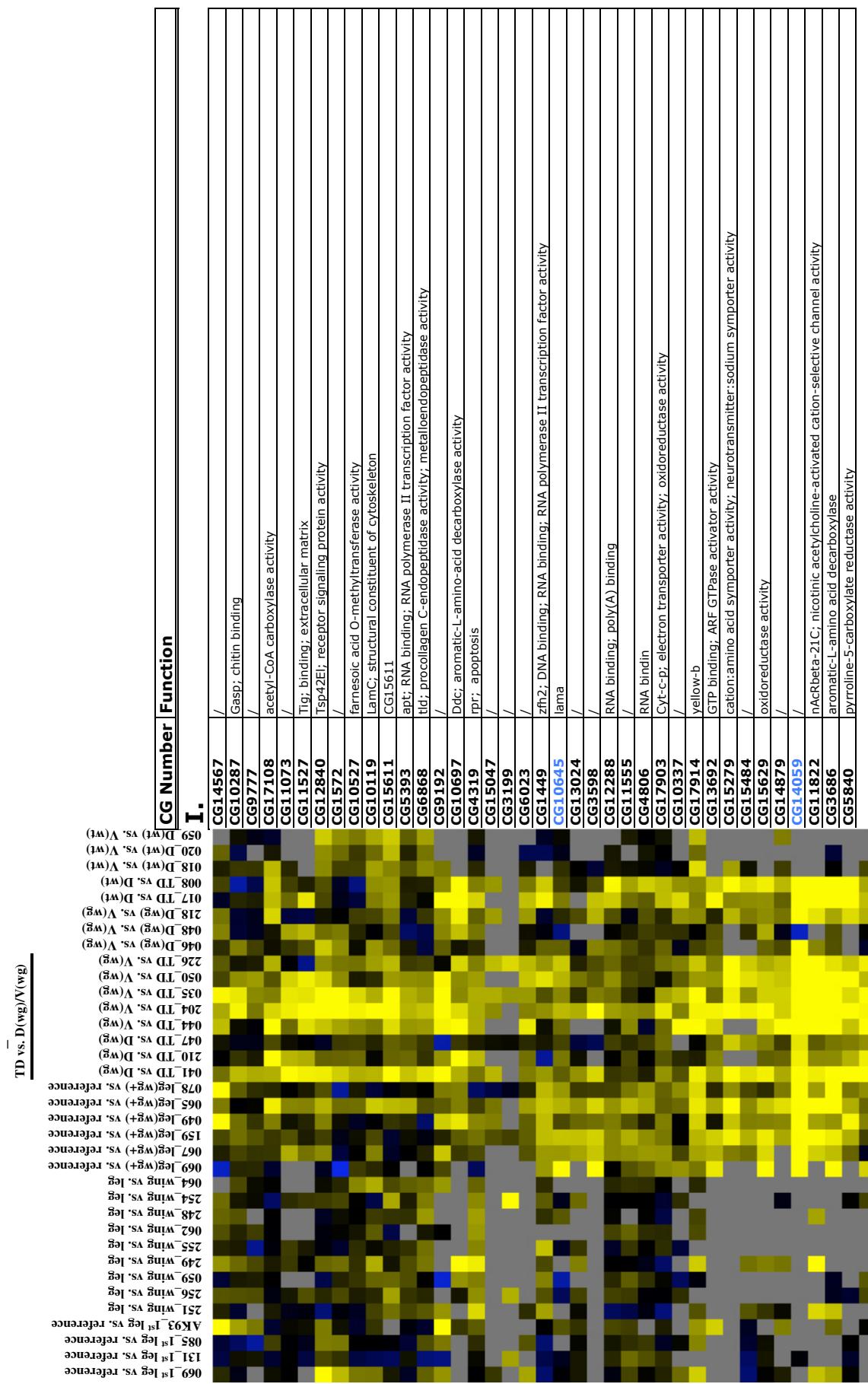
			nucleobase, nucleoside, nucleotide and nucleic acid metabolism; GO:0006139; nucleobase, nucleoside, nucleotide and nucleic acid transport; GO:0015931;
			protein metabolism; GO:0019538; regulation of translation; GO:0006445; signal transduction; GO:0006810
CG11073	CG11073	-1.60	6.09E-03 /
CG10697	Ddc	-1.57	1.10E-03 catecholamine metabolism; GO:0006584; courtship behavior; GO:0007619; cuticle biosynthesis; GO:0042335; dopamine biosynthesis from tyrosine; GO:0006761; melanin biosynthesis; GO:0042438; pigmentation; GO:0048066; serotonin biosynthesis from tryptophan; GO:0006587
CG14879	CG14879	-1.56	2.45E-04 /
CG4766	CG4766	-1.54	2.14E-03 /
CG13097	CG13097	-1.48	1.90E-03 nucleobase, nucleoside, nucleotide and nucleic acid metabolism; GO:0006139; rRNA metabolism; GO:0016072
CG5874	CG5874	-1.46	3.67E-04 negative regulation of transcription from Pol II promoter; mitotic; GO:0007070
CG3200	Reg2	-1.46	1.03E-03 metabolism; GO:0008152
CG11259	MICAL-like	-1.45	4.16E-03 cytoskeleton organization and biogenesis; GO:0007010
CG7367	CG7367	-1.43	3.50E-03 lipid metabolism; GO:0006629
CG6044	CG6044	-1.41	2.61E-04 /
CG6724	CG6724	-1.39	7.40E-05 /
CG11708	CG11708	-1.39	1.32E-03 transport; GO:0006810
CG7182	CG7182	-1.33	4.57E-04 defense response; GO:0006952; protein folding; GO:0006457; response to heat; GO:0009408; response to stress; GO:0006950
CG66962	CG66962	-1.33	5.88E-04 /
CG1698	CG1698	-1.33	5.88E-04 amino acid metabolism; GO:0006520; amino acid transport; GO:0006865; cation transport; GO:0006812; extracellular transport; GO:0006858; neurotransmitter transport; GO:0006821; sodium ion transport; GO:0006814
CG33188	CG33188	-1.33	3.13E-04 perception of sound; GO:0006865; cation transport; GO:0006812; chloride transport; GO:0006861; sodium ion transport; GO:0006814
CG4567	CG4567	-1.32	3.96E-03 transcriptional elongation; GO:0006865; cation transport; GO:0006915; apoptosis; GO:00068632; embryonic development (sensu Insecta); GO:0001700; induction of apoptosis; ecdisone-mediated induction of apoptosis by ionic changes; GO:0086627; larval midgut cell death; GO:0035096; negative regulation of protein biosynthesis; GO:0006917; induction of apoptosis and peptidofusins; GO:0030162; response to DNA damage stimulus; GO:0017448; programmed cell death; GO:002501; protein ubiquitination; GO:0016567; regulation of proteolysis and peptidofusins; GO:0035071
CG4357	CG4357	-1.29	9.90E-04 amino acid transport; GO:0006865; cation transport; GO:0006915; apoptosis; GO:00035075; central nervous system metamorphosis; GO:0035193; NOT nurse cell apoptosis; GO:0045476; apoptosis; GO:0006915; apoptotic program; GO:00068632; embryonic development (sensu Insecta); GO:0001700; induction of apoptosis; ecdisone-mediated induction of apoptosis by ionic changes; GO:0086627; larval midgut cell death; GO:0035096; negative regulation of protein biosynthesis; GO:0006917; induction of apoptosis and peptidofusins; GO:0030162; response to DNA damage stimulus; GO:0017448; programmed cell death; GO:002501; protein ubiquitination; GO:0016567; regulation of proteolysis and peptidofusins; GO:0035071
CG14479	CG14479	-1.29	1.14E-03 /
CG4319	rpr	-1.27	1.76E-04 DNA replication; GO:0006260; translational initiation; GO:0006413
CG5836	SF1	-1.26	3.67E-04 nuclear mRNA splicing, via spliceosome; GO:0000398
CG2512	alphaTub84D	-1.24	8.02E-06 cell motility; GO:0006928; cellular physiological process; GO:0046785; microtubule-based movement; GO:0007018; microtubule-based process; GO:0007017; mitosis; GO:0007057
CG10161	eIF-3d66	-1.23	6.05E-03 protein biosynthesis; GO:0006412; translational initiation; GO:0006413
CG1108	DNAPol-alpha50	-1.23	7.18E-05 DNA replication; GO:0006260; DNA replication, synthesis of RNA primer; GO:0006269
CG6728	CG6728	-1.23	3.00E-03 phosphate metabolism; GO:0006796; phosphate transport; GO:0006817
CG12554	CG12554	-1.23	6.45E-04 electron transport; GO:0006118; oxidative phosphorylation; GO:0006119
			actin cytoskeleton organization and biogenesis; GO:0003036; anti-apoptosis; GO:0006242; oogenesis (sensu Insecta); GO:0006242; actin filament-based movement; GO:0006246; intracellular signaling cascade; GO:0006286; pathway; GO:0006286; positive regulation of cell size; GO:0045793; potassium ion transport; GO:0006813; protein amino acid phosphorylation; GO:0006468; regulation of cell growth; GO:0030307; positive regulation of cell size; GO:0006620; sex differentiation; GO:0007548; spermatogenesis; GO:0007283
CG1210	Pk61C	-1.22	2.77E-03 GO:0001558; regulation of organ size; GO:0006620
CG1091	CG1091	-1.22	3.21E-03 protein biosynthesis; GO:0006412
CG3691	Pof	-1.21	4.87E-03 /
CG6014	CG6014	-1.21	1.01E-03 /
CG5355	CG5355	-1.21	1.98E-03 proteolysis and peptidolysis; GO:0006508
CG17124	CG17124	-1.20	5.00E-06 /
CG17843	CG17843	-1.19	3.45E-03 electron transport; GO:0006118
CG17337	CG17337	-1.19	4.61E-04 /
CG13849	Nop56	-1.18	1.02E-03 rRNA metabolism; GO:0016072
CG49778	Mcm7	-1.17	5.32E-04 DNA replication initiation; GO:0006270; pre-replicative complex formation and maintenance; GO:0006267
CG10363	TepIV	-1.17	9.13E-03 antibacterial humoral response (sensu Protostomia); GO:0006961; phosphotriehydroxyacetate-dependent sugar phosphotransferase system; GO:0009401
CG1572	CG1572	-1.16	1.05E-03 /
CG7583	CtBP	-1.16	3.75E-03 L-serine biosynthesis; GO:0006564; negative regulation of transcription from Pol II promoter; GO:0000122
CG1450	CG1450	-1.15	6.91E-03 regulation of transcription; GO:0045449
CG18410	CG18410	-1.14	1.13E-03 /
CG13813	CG13813	-1.14	1.09E-03 /
CG2863	Nle	-1.12	7.95E-04 G-protein coupled receptor protein signaling pathway; GO:0007186; Notch signalling pathway; GO:0007219
CG6375	pit	-1.12	4.89E-04 nucleobase, nucleoside, nucleotide and nucleic acid metabolism; GO:0006139
CG6897	CG6897	-1.12	1.57E-05 /
CG6868		-1.12	amnioserosa formation; GO:0007378; embryonic pattern specification; GO:0009680; positive regulation of BMP signalling pathway; GO:0030513; proteolysis and peptidofusins; GO:0006508; regulation of transforming growth factor beta receptor signalling pathway; GO:0017015; terminal region determination; GO:0007362; torso signalling pathway; GO:0008393; zygotic determination of dorsal/ventral axis; GO:0007352
CG5393	apt	-1.11	2.37E-04 tracheal branching (sensu Insecta); GO:0007424
CG11525	CycG	-1.10	6.53E-04 DNA repair; GO:0006281
CG3568		-1.07	
CG6739		-1.07	4.84E-05 /
CG13692	CG13692	-1.06	4.95E-03 development; GO:0007275
CG6623		-1.06	1.90E-04 protein amino acid ADP-ribosylation; GO:0006471
CG11190	CG11190	-1.06	9.42E-04 attachment of GPI anchor to protein; GO:0016255

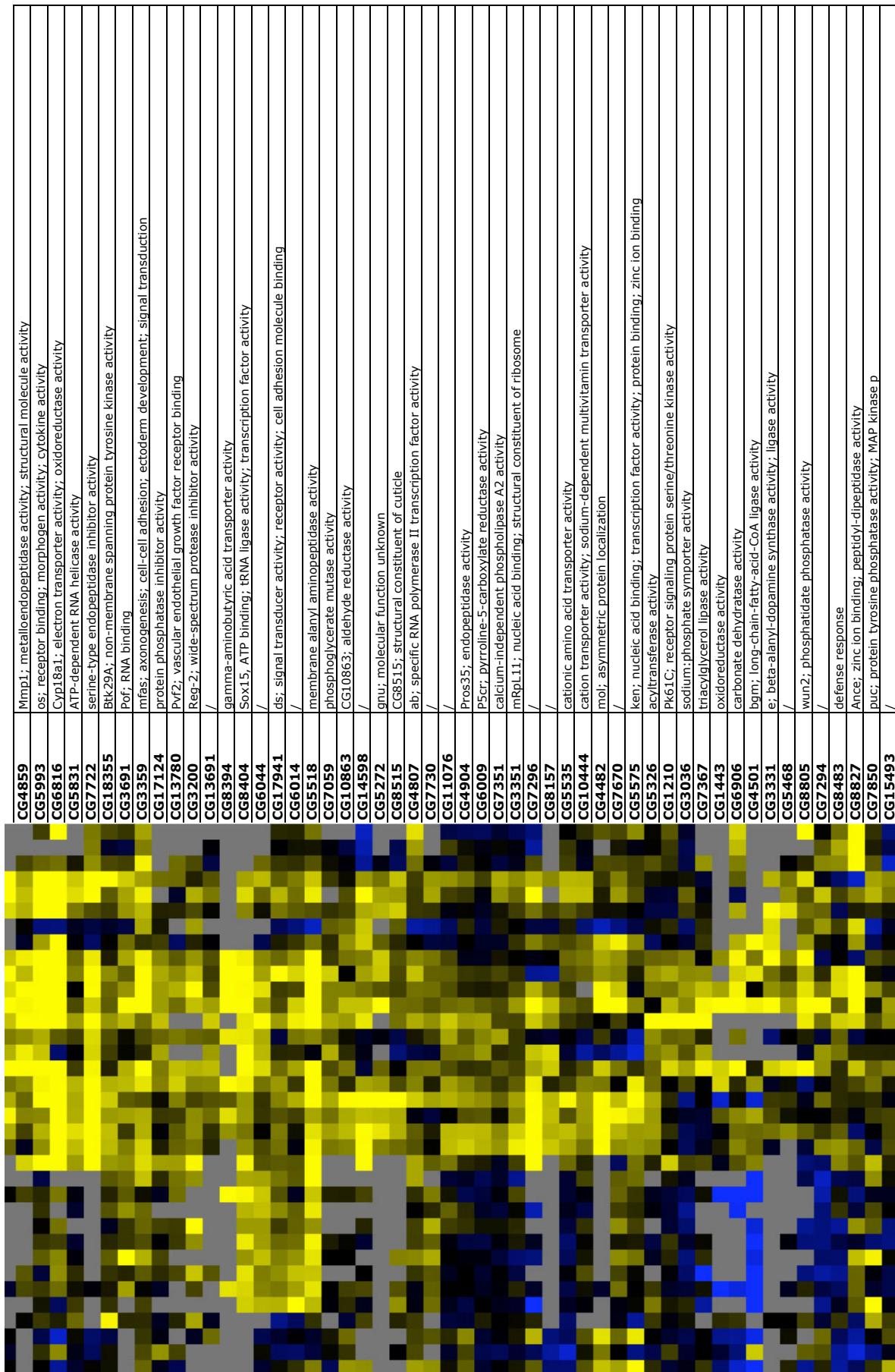
G1794	Mmp2	-1.05	5.54E-04	anti-apoptosis; GO:0006916; oogenesis (sensu Insecta); GO:0009933; proteolysis and peptidolysis; GO:0006508
G14984	C14984	-1.05	8.72E-03	/
G12358	Paipl2	-1.05	2.03E-03	negative regulation of translation; GO:0016478
G3254	pgant2	-1.05	1.30E-03	oligosaccharide biosynthesis; GO:0009312; polysaccharide metabolism; GO:0005976; protein amino acid glycosylation; GO:0006486
G5913	C5913	-1.04	2.11E-04	/
G74798	Act798	-1.03	4.75E-03	cytoskeleton organization and biogenesis; GO:0007010
G10354	C10354	-1.03	1.01E-03	DNA metabolism; GO:0006259; nucleobase, nucleoside, nucleotide and nucleic acid metabolism; GO:0006139
G18535	C18535	-1.03	7.57E-04	metabolism; GO:0008152
G3036	C3036	-1.02	6.28E-04	carbohydrate metabolism; GO:0005975; carbohydrate transport; GO:0008643; cation transport; GO:0006812; extracellular transport; GO:0006858; phosphate metabolism; GO:0006796; phosphate transport; GO:0006817
G15561	C15561	-1.02	2.33E-03	/
G1874	C1874	-1.01	3.24E-03	/
G5114	C5114	-1.00	1.68E-03	/
G7564	C7564	-1.00	7.09E-03	nuclear mRNA splicing, via spliceosome; GO:0000398
G11990	C11990	-1.00	8.99E-05	/
G10336	C10336	-0.97	1.04E-03	/
G64027	Act5C	-0.97	3.91E-03	cytoskeleton organization and biogenesis; GO:0007010; sperm individualization; GO:0007291
G5014	Vap-33.1	-0.97	8.06E-03	intracellular protein transport; GO:00056886; neuromuscular junction development; GO:0007528; neurotransmitter secretion; GO:0007269; synaptic transmission; GO:0007268; synaptic vesicle priming; GO:0016082; transmission of nerve impulse; GO:0019226; vesicle-mediated transport; GO:0016192
G13386	Itbp	-0.97	8.36E-04	neuropeptide signaling pathway; GO:0007218
G8545	C8545	-0.97	2.26E-03	RNA metabolism; GO:0016072; rRNA processing; GO:0006364
264591	Tsp86D	-0.96	9.46E-07	cell-cell adhesion; GO:0013337; signal transduction; GO:0007165
G68171	dup	-0.96	1.22E-03	DNA replication; GO:0006260; DNA replication checkpoint; GO:0000076; DNA-dependent DNA replication; GO:0006261; chorion gene amplification; GO:0007307; eggshell formation (sensu Insecta); GO:0077304
G1483	Map-205	-0.96	4.67E-03	microtubule-based process; GO:0007017
G13780	Pvt2	-0.95	7.70E-04	hemocyte cell migration (sensu Arthropoda); GO:0035099; hemocyte proliferation (sensu Arthropoda); GO:0035172
G10527	C10527	-0.95	2.46E-03	/
G10337	C10337	-0.95	7.78E-03	/
G7839	C7839	-0.95	2.87E-03	regulation of transcription; GO:0004549; regulation of transcription from Pol II promoter; GO:00065357
G7059	C7059	-0.95	5.91E-04	glycolysis; GO:0006096; microtubule-based movement; GO:0007018
G13690	C13690	-0.95	3.34E-04	RNA catabolism; GO:0006401
[G19994]	[G19994]	-0.94	6.63E-03	biological process unknown; GO:0000004
G10545	Lama	-0.94	2.49E-04	/
G7842	C7842	-0.94	1.23E-06	fatty acid biosynthesis; GO:0006633
G18290	Act87E	-0.93	3.71E-03	cytoskeleton organization and biogenesis; GO:0007010
G1677	C1677	-0.93	3.87E-03	/
G13691	C13691	-0.93	8.90E-04	/
G10722	C10722	-0.93	1.01E-03	/
G6751	G6751	-0.93	4.03E-03	general RNA polymerase II transcription factor activity; GO:0006251; ; GO:0016251; defense response; GO:000952; development; GO:0007275; extracellular transport; GO:0006858; response to toxin; GO:0006636
G67899	[G103659]	-0.93	7.08E-04	defense response; GO:0006281; regulation of transcription from Pol II promoter; GO:00065357
G7804	C7804	-0.92	2.95E-03	nuclear mRNA splicing, via spliceosome; GO:0000398; regulation of transcription from Pol II promoter; GO:0006858; response to toxin; GO:0006636
G1677	C1677	-0.92	3.87E-03	/
G13691	C13691	-0.92	8.90E-04	/
G10722	C10722	-0.92	1.01E-03	/
G6751	G6751	-0.93	4.03E-03	general RNA polymerase II transcription factor activity; GO:0006251; ; GO:0016251; defense response; GO:000952; extracellular transport; GO:0006858; response to toxin; GO:0006636
G67627	C67627	-0.92	6.61E-03	defense response; GO:0006281; regulation of transcription from Pol II promoter; GO:00065357
G62918	C62918	-0.91	1.07E-03	/
G5353	C5353	-0.91	7.27E-04	/
G9373	C9373	-0.91	4.95E-03	Bmp signaling pathway; GO:0006533; lipid metabolism; GO:0006629
G59923	DNApol-alpha7.3	-0.90	8.09E-03	aminoacid formation; GO:0007378; anterior/posterior axis specification; GO:0009948; anterior/posterior pattern formation, imaginal disc; GO:0007448; branch cell fate determination (sensu Insecta); GO:0006845; cell fate specification; GO:0001709; cell fate specification; GO:00064843; dorsal closure; GO:0007391; dorsal closure, leading edge cell fate determination; GO:0007393; dorsal/ventral pattern formation, imaginal disc; GO:0007450; actoderm cell fate specification; GO:0007315; ectoderm development; GO:0007398; embryonic morphogenesis; GO:0009705; foregut morphogenesis; GO:0007440; genital disc sexually dimorphic development; GO:0007323; germ-line stem cell division; GO:0042078; germ-line stem cell maintenance; GO:0030718; heart development; GO:0007446; imaginal disc development; GO:0007447; larval development (sensu Insecta); GO:0002168; leg disc proximal/distal pattern formation; GO:0007479; negative regulation of cell proliferation; GO:0006468; ovarian follicle cell development (sensu Insecta); GO:0001710; negative regulation of salivary gland development; GO:0045475; nurse cell apoptosis; GO:0007447; negative regulation of cell proliferation; GO:0009933; ovarian follicle cell development (sensu Insecta); GO:00030707; pigmentation; GO:0048066; progression of morphogenetic furrow (sensu Endopterygota); GO:0007458; regulation of cell proliferation; GO:0042127; regulation of organ size; GO:0046620; regulation of tracheal tube diameter; GO:0035158; sensory organ development; GO:0007442; stem cell division; GO:0017145; stem cell maintenance; GO:0007427; tracheal branch fusion; GO:0035147; tracheal cell fate determination (sensu Insecta); GO:0007425; tracheal cell migration; GO:0042079; wing disc anterior/posterior pattern formation; GO:0007410; wing disc proximal/distal pattern formation; GO:0007413; wing vein specification; GO:0008386; wing vein specification; GO:0007474; zygotic determination of anterior/posterior axis; embryo; GO:0007352
G9885	dpp	-0.88	3.83E-03	G-protein coupled receptor protein signaling pathway; GO:0007186; determination of adult life span; GO:0008340; response to stress; GO:0006950
G65330	math13	-0.88	9.78E-03	/

CG14191	CG14191	-0.88	1.31E-03	/
CG5452	dnk	-0.88	3.23E-03	TMP biosynthesis; GO:0006230; nucleoside diphosphate phosphorylation; GO:0006165; phosphorylation; GO:0016310; pyrimidine base metabolism; GO:0006206
CG3198	CG3198	-0.88	5.35E-03	nuclear mRNA splicing, via spliceosome; GO:0000398
CG17060	Rab10	-0.87	1.64E-03	endocytosis; GO:0006897; intracellular protein transport; GO:0006886; regulation of exocytosis; GO:0017157; regulation of transcription, DNA-dependent;
CG5018	CG5018	-0.87	4.72E-04	GO:0006355; small GTPase mediated signal transduction; GO:0007264; two-component signal transduction system (phosphorelay); GO:000160
CG9143	CG9143	-0.86	2.34E-03	
CG33129	CG6087	-0.86	5.22E-03	nucleobase, nucleoside, nucleotide and nucleic acid metabolism; GO:0006139
CG10863	CG10863	-0.86	7.11E-03	
CG5553	DNAPrim	-0.85	2.74E-03	DNA replication; GO:0006260; DNA replication, synthesis of RNA primer; GO:0006269; S phase of mitotic cell cycle; GO:000084; compound eye morphogenesis (sensu Endopterygota); GO:001745

Table S6

Table S7 Genes upregulated in TD cells





III.

CG16753	/	Act57B; structural constituent of cytoskeleton; glucuronosyltransferase activity
CG10067	CG10067	cabut; transcriptional activator activity
CG4427	/	
CG17494	/	nuclear mRNA splicing, via spliceosome
CG3198	/	
CG15561	/	
CG12358	Pain2; protein binding	
CG12534	CG12534	flavin-linked sulphydryl oxidase activity
CG18535	/	
CG874	RNA binding	
CG1898	HBS1; translation release factor activity; translation elongation factor activity	
CG12581	/	
CG6116	molecular function unknown	
CG3039	ogre; qab-4 junction forming channel activity	
CG11525	CycG; cyclin-dependent protein kinase regulator activity	
CG2093	/	
CG1091	/	
CG13097	RNA binding; nucleic acid binding	
CG14479	/	
CG3983	receptor binding; hormone activity	
CG10161	eIF-3p66; translation initiation factor activity; GO:0003743; ; GO:0003743; ; GO:0003743;	
CG11990	/	
CG14804	intracellular protein transport; lysosome organization and biogenesis	
CG11259	MICAL-like; actin binding; structural constituent of cytoskeleton	
CG4006	Akt1; receptor signaling protein serine/threonine kinase activity	
CG5353	threonine-tRNA ligase	
CG5836	SF1; pre-mRNA splicing factor activity; transcription cofactor activity	
CG1602	Sip5; RNA binding; pre-mRNA splicing factor activity	
CG7564	nuclear mRNA splicing, via spliceosome	
CG7788	Ice; caspase activity	
CG6962	/	
CG5355	prolyl oligo-GOPetidase activity	
CG16750	protein binding; receptor binding; perception of sound	
CG4994	Mpcp; carrier activity; phosphate transporter activity	
CG9277	beta tub56D; GTP binding; structural constituent of cytoskeleton	
CG6767	nucleotide kinase activity; ribose-phosphate diphosphokinase activity	
CG9915	/	
CG17187	/	
CG3183	geminin; DNA binding; regulation of cell cycle	
CG4528	sni; RNA binding; pre-mRNA splicing factor activity	
CG2028	Ck1alpha; casein kinase I activity; receptor signaling_protein serine/threonine kinase activity	
CG5014	Vap-33-1; intracellular protein transport; synaptic transmission	
CG7583	CBP; protein C-terminal binding; transcription corepressor activity	
CG5605	erR1; translation termination factor activity; translation release factor activity	
CG4863	Rpl3; nucleic acid binding; structural constituent of ribosome	
CG5130	zinc ion transporter activity	

III.

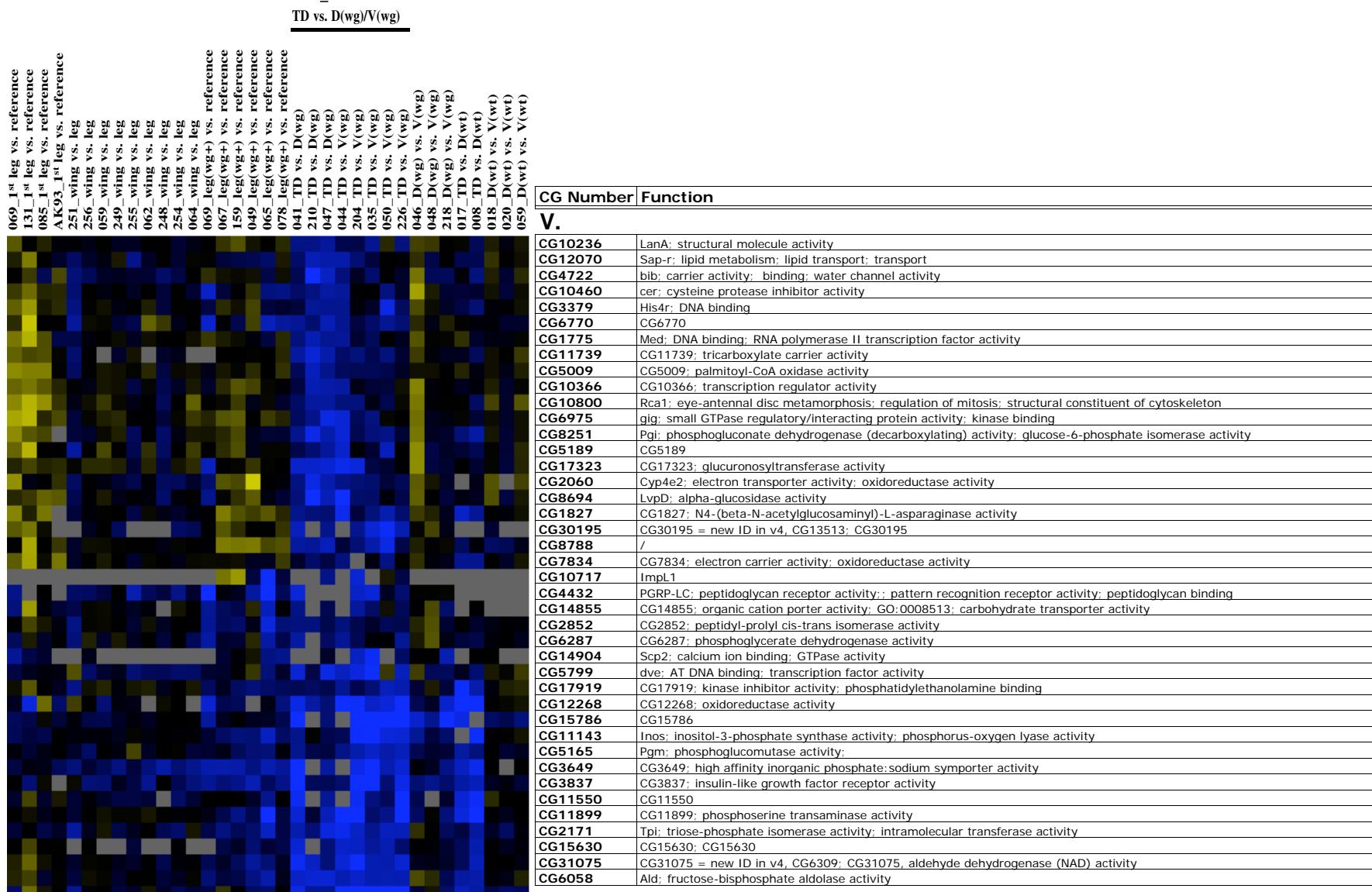
CG11450	net; RNA polymerase II transcription factor activity
CG2982	/
CG2198	Ama; antigen binding
CG12018	nucleic acid binding; delta DNA polymerase activity
CG2863	Nle; Notch signaling pathway
CG3568	/
CG13813	/
CG5913	/
CG17337	/
CG18410	/
CG17843	flavin-linked sulfhydryl oxidase activity; glucosidase activity
CG7108	DNApol-alpha50; nucleic acid binding; primase activity; alpha DNA polymerase activity
CG10121	SP1173
CG4978	Mcm7; chromatin binding; DNA replication origin binding; DNA helicase activity
CG5371	RnrL; ribonucleoside-diphosphate reductase activity
CG8171	dup; DNA binding; DNA replication checkpoint
CG7628	phosphate transporter activity
CG6897	/
CG13849	Nod56; RNA binding
CG6375	nucleic acid binding; RNA helicase activity; ATP-dependent RNA helicase activity
CG7182	defense response; protein folding; response to heat; response to stress
CG13425	bl; RNA binding; transcription factor binding; appendage morphogenesis; cell fate commitment; proliferation
CG7878	nucleic acid binding; helicase activity; ATP-dependent RNA helicase activity
CG11803	ATP binding; ATPase activity, coupled to transmembrane movement of substances; transporter activity
CG9253	nucleic acid binding; ATP-dependent RNA helicase activity
CG10778	acyltransferase activity; transferase activity
CG5383	PSR
CG18600	/
CG6983	Hrb98DE; RNA binding
CG6724	/
CG3593	r-; orotate phosphoribosyltransferase activity; orotidine-5'-phosphate decarboxylase activity
CG7447	/

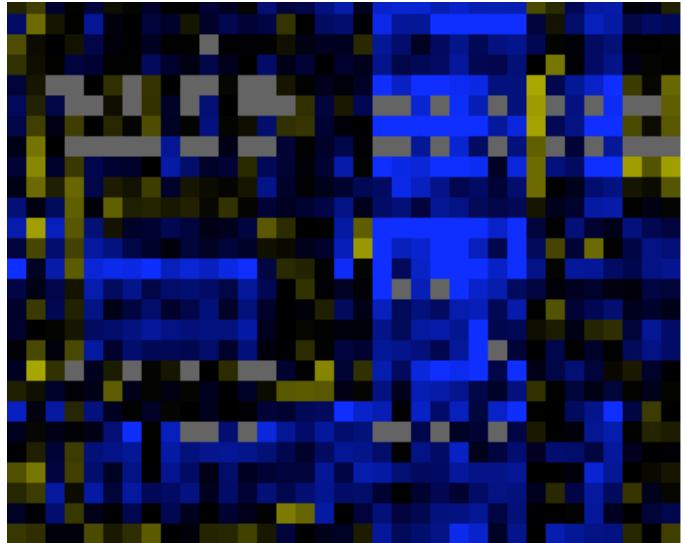
IV.

CG1910	molecular function unknown
CG15353	/
CG14489	olfr186-M
CG15093	3-hydroxyisobutyrate dehydrogenase activity
CG14846	/
CG17032	/
CG12023	/
CG12287	pdm2; DNA binding; specific RNA polymerase II transcription factor activity; :CG15487:CG15486
CG7924	/
CG2762	ush; transcription factor binding; ligand-dependent nuclear receptor activity; transcription factor activity
CG3385	nvv; transcription factor activity
CG9552	rost; mesoderm development; myoblast fusion
CG2444	/
CG7860	asparaginase activity
CG17738	/
CG17345	amid; carboxy-lyase activity; aromatic-L-amino-acid decarboxylase activity
CG12843	Tsp4E1; Tetraspanin
CG13250	/
CG9812	/

	CG2663	carrier activity; tocopherol binding
	CG10197	ktl; collier; transcription regulator activity; specific RNA polymerase II transcription factor activity
	CG10619	tup; specific RNA polymerase II transcription factor activity
	CG5966	tricylglycerol lipase activity
	CG68376	abp; zinc ion binding; specific RNA polymerase II transcription factor activity
	CG1148	Ost2; transcription regulator activity
	CG1803	regulation; anterior/posterior axis specification; calcium-mediated signaling
	CG10704	toe1; transcription factor activity
	CG3132	Ect3; beta-galactosidase activity
	CG9427	/
	CG7820	CAH1; carbonic anhydrase activity
	CG1058	rpk; sodium channel activity; amiloride-sensitive sodium channel activity
	CG9023	Drip; water transporter activity; carrier activity
	CG10601	mirr; protein binding; transcription factor activity
	CG6680	serine-/tyr-endopeptidase inhibitor activity
	CG9623	Ifi; protein binding; receptor activity; cell adhesion molecule binding
	CG10570	/
	CG3254	pgant2; polypeptide N-acetylgalactosaminyltransferase activity
	CG38330	vqj; wing margin morphogenesis; wing morphogenesis
	CG1897	mstn; specific RNA polymerase II transcription factor activity
	CG8216	/
	CG69358	Pmk-3; diacylglycerol binding; protein serine/threonine kinase activity; carrier activity
	CG18589	TepII; wide-spectrum protease inhibitor activity
	CG11370	/
	CG10208	/
	CG6597	sterol O-acyltransferase activity; hydrolase activity, acting on ester bonds
	CG69372	trypsin activity; serine-type peptidase activity
	CG7539	Edg91; structural constituent of pupal cuticle
	CG10625	/
	CG1698	potassium:amino acid transporter activity; cation transporter activity
	CG11905	/
	CG68301	/
	CG7766	/
	CG1794	Mmp2; metalloendopeptidase activity; structural molecule activity
	CG6746	mab-2; encodes Mab-21, involved in cell fate determination
	CG85902	structural constituent of larval cuticle
	CG64914	SP39; trypsin activity; serine-type endopeptidase activity
	CG9307	chitinase activity; hydrolase activity; hydrolyzing N-glycosyl compounds
	CG7811	b ₂ ; aspartate 1-decarboxylase activity; glutamate decarboxylase activity
	CG8587	Cyp301a1; electron transporter activity; oxidoreductase activity
	CG16884	/
	CG8012	/

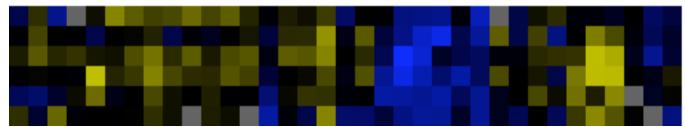
Table S8 Genes downregulated in TD cells





CG9914	CG9914; structural molecule activity; ; GO:0005198; oxidoreductase activity
CG5390	CG5390; NOT serine-type endopeptidase activity
CG9027	CG9027; superoxide dismutase activity
CG10663	CG10663; trypsin activity; serine-type endopeptidase activity
CG3743	MTF-1; transcription factor activity; specific RNA polymerase II transcription factor activity
CG17052	CG17052; structural constituent of peritrophic membrane (sensu Insecta);
CG15676	CG15676; chaperone binding
CG9656	grn; RNA polymerase II transcription factor activity; general RNA polymerase II transcription factor activity
CG5273	CG5273
CG6871	Cat; heme binding; antioxidant activity; catalase activity; peroxidase activity
CG16733	CG16733; aryl sulfotransferase activity; retinol dehydratase activity
CG6206	CG6206; hydrolase activity, hydrolyzing N-glycosyl compound; alpha-mannosidase activity
CG2056	CG2056; trypsin activity; GO:0004295 ; EC:3.4.21.4; serine-type endopeptidase activity
CG7231	CG7231
CG7675	CG7675; oxidoreductase activity, acting on CH-OH group of donors
CG1916	Wnt2; receptor bindin; signal transducer activity
CG9614	pip; heparin-sulfate 2-sulfotransferase activity; sulfotransferase activity
CG6117	Pka-C3; receptor signaling protein serine/threonine kinase activity; cAMP-dependent protein kinase activity
CG9355	dy; structural constituent of cuticle (sensu Insecta)
CG4463	Hsp23; actin binding;
CG32150	CG32150 = new ID in v4, CG15714; CG32150
CG15085	edl; protein binding; Ras signaling pathway;
CG8361	HLHm7; specific RNA polymerase II transcription factor activity; transcription factor activity
CG8365	E(spl); DNA binding; transcription factor activity
CG6104	m2; Notch signaling pathway
CG6127	Ser; receptor binding; Notch binding; signal transducer activity; epidermal growth factor receptor binding

VI.



CG4316	Sb; serine-type endopeptidase activity; trypsin activity
CG32940	BcDNA: GH12504/ /
CG3376	CG3376; sphin; GOmyelin phosphodiesterase activity
CG3905	Su(2)2; DNA binding; transcription regulator activity
CG8681	clumsy; kainate selective glutamate receptor activity; ligand-gated ion channel activity
CG5096	CG5096; receptor activity

Table S9
vgBE-like elements in *D. melanogaster*, *D. yakuba* and *D. pseudoobscura*