MANY P-ELEMENT INSERTIONS AFFECT WING SHAPE IN DROSOPHILA MELANOGASTER

K. WEBER, N. JOHNSON, D. CHAMPLIN and A. PATTY

Department of Biological Sciences
University of Southern Maine
Portland, ME 04104-9300
Running Head:

"WING SHAPE GENES IN FLIES"

Key Words:
Development—morphology—wings—quantitative trait genes

Corresponding author:

K. WEBER
Department of Biological Sciences
University of Southern Maine
Box 9300
Portland, ME 04104-9300

Telephone: (207) - 780 - 4265
FAX: (207) - 228 - 8116
E-mail: keweber@usm.maine.edu
ABSTRACT

A screen of random, autosomal, homozygous-viable P-element insertions in *D. melanogaster* found small effects on wing shape in 11 of 50 lines. The effects were due to single insertions and remained stable and significant for over five years, in repeated, high-resolution measurements. All 11 insertions were within or near protein-coding transcription units, none of which were previously known to affect wing shape. Many sites in the genome can affect wing shape.

INTRODUCTION

For every quantitative trait, some set of genomic sites can yield mutational variation. The full distribution of potential effects is unknowable, being a function of all possible modifications; but any mutational screen can help to identify quantitative trait loci (QTLs). Here we report the provisional identification of 11 loci affecting wing shape in *D. melanogaster*, from effects produced with high frequency by random P-element insertions.

Studies of natural allelic variation affecting wing shape in *D. melanogaster* find evidence of numerous mainly additive effects (WEBER et al. 1999; 2001), which are largely independent of sex and body size (WEBER 1990; BIRDSALL et al. 2000; ZIMMERMAN et al. 2000). These effects are small and are sometimes concentrated in small regions of the wing (WEBER 1992). Wing shape is the third morphological trait in *Drosophila* to be mapped for natural QTLs (WEBER et al. 1999), after bristle number (LONG et al. 1995) and features of male genitalia (LIU et al. 1996).

Natural alleles can be quite different from new mutations. Their effects can arise from multiple sequence differences (HAENLIN et al. 1994; STAM and LAURIE 1996; ROBIN et al. 2002), unlike the new alleles that show up in mutational screens. Natural alleles may evolve piecemeal over long periods (KREITMAN and HUDSON 1991; PHILLIPS 1999). Moreover, the natural genetic variance for a quantitative trait can arise mainly from a few major loci (ROBERTSON 1967; MACKAY 2001a; 2001b), whereas mutational genetic variation for the same trait could have a broader base.

Common alleles affecting traits in nature may not always represent loci that would contribute to major evolutionary changes. Genes often affect particular traits through complex pleiotropic pathways, as the classical trait of bristle number in *Drosophila* has shown. For example, much of the standing genetic variance for bristle number appears to arise from genes with primary roles in the development of the peripheral nervous system (PNS) (NUZHDIN et al. 1999; NORGÅ et al. 2003). Similarly, one of the first quantitative trait genes to be identified was the *bobbed* (*bb*) locus, which codes for ribosomal RNA but also happens to affect bristle number (FRANKHAM et al. 1978). Despite these findings, one would think that natural selection could produce macroevolutionary changes in bristle number, without the inevitable derangement of either ribosome production or the PNS. Thus, during a major adaptive change, the common alleles first available to selection might sometimes present genetic conflicts, rather than a waiting reserve of evolutionary potential. New mutational effects should provide a more complete picture of the genetic basis of traits, and of their potential to evolve. A key question is the ultimate mutational target size for typical traits.

Most P-element insertion screens have focused on qualitative effects that can be scored in a few individuals. Other screens have studied quantitative effects, by measuring the increase in
phenotypic variance among P-element lines compared to controls, or by quantifying effects in individual insertion lines as deviations from the control mean. However, the creation and extraction of P-element insertion lines can cause various types of new genetic variation, and small effects of insertions are hard to separate from these and other residual variations in the genetic background. This study employs new tactics to reduce both genetic and environmental variation, to help ensure that insertion lines differ from controls only in the effect of the insertion.

The shape metrics employed in these studies are angular offsets (WEBER 1990). Angular offsets are sometimes inadvertently described as ratios (KLINGENBERG et al. 2001; HOULE et al. 2003), but ratios are not valid metrics of allometric shapes. Angular offsets resemble ratios in that each value is a function of two numbers (dimensions D1 and D2) with opposite effects on its magnitude, but unlike ratios they are size-independent in allometric forms. Angular offsets provide the most reductionistic shape metrics possible, based on the simplest aspect of shape, which is the allometric relation between two dimensions.

**MATERIALS AND METHODS**

**Creation of insertion lines:** The P-element construct P{lacW} contains a functional allele of white (w^{+mC}), providing a visible marker (BIER, et al. 1989). P{lacW} retains the left and right P-element terminal sequences required for transposition, but lacks the transposase gene. We crossed a line with a single P{lacW} inserted on the X chromosome, in a background marked by yellow (y) and w, to another M-cytotype line, also marked by y and w. Thereafter, red-eyed virgin flies, heterozygous for the insertion, were mated to their white-eyed brothers lacking the insertion, for 20 generations of single-pair matings. This created an inbred line in which this single P-element insertion was still segregating, in an otherwise nearly isogenic y w background. This line was the P-element source and the target genome in all mobilizations.

We mobilized P{lacW} in the standard way, by crossing red-eyed virgin heterozygotes, from the inbred stock just described, to males with dominantly marked and balanced second and third chromosomes, carrying the defective P-element "jumpstart" or “∆2-3” which produces transposase but cannot jump (ROBERTSON et al. 1988; BIER, et al. 1989). Sons containing both P-elements, and with red sectored eyes showing mobilization of P{lacW}, were backcrossed to virgin white-eyed cousins segregating from the inbred P-element source line. From this mating, new autosomal insertions of P{lacW} were picked up as heterozygous red-eyed males lacking their grandfather’s dominant markers. By this procedure, 60 random autosomal insertions were collected, each from a different vial to avoid duplicate insertions.

Each new heterozygous insertion male was crossed to one virgin white-eyed sib, also lacking dominant markers. Thereafter, each insertion was maintained in a segregating line, by single-pair matings of red-eyed insertion-heterozygote virgins to their white-eyed brothers, for ten more generations of inbreeding, leading up to the first screen.

**Screening for wing shape effects:** During the whole experiment each insertion was maintained continuously in its own inbred segregating line. Insertions were screened three times, and for each screen new, temporary, homozygous insertion-bearing and insertion-free lines were extracted from the segregating line and compared. Insertions were always made homozygous without using balancers, to avoid the chance of small effects entering the inbred lines from balancer stocks. Wherever possible, we used the slight differences in the depth of red
eye color to distinguish between insertion homozygotes and heterozygotes. Where eye colors were not distinguishable, we relied on the inevitability of success in large numbers of single pair matings between red-eyed flies, and the fact that eye colors would not segregate in fixed lines. Ten of the 60 insertions could not be made homozygous.

To eliminate effects of environmental differences among vials, flies of paired homozygous insertion and control lines were cultured together in vials containing five pregnant females from each line. From the progeny of each vial, 20 red-eyed males and 20 white-eyed males were compared. Paired samples of insertion and control were always measured by the same operator. Five vials were pooled per insertion, so that 100 insertion-homozygote males were compared to 100 insertion-free males. This measurement protocol was standard for all lines in all screens, except for line #41 in the third screen, when an extra comparison was done using pooled sample sizes of 300. Flies were cultured in plastic 75 ml vials on commercial potato-flake medium (Carolina Biological Supply, www.carolina.com) with yeast, at 26°.

**First screen:** Homozygous insertion and non-insertion control lines were both extracted from the same inbred segregating line, for each insertion that could be made homozygous. Wing shapes were compared between each homozygous insertion line and its paired control line. Insertions that showed significant differences were maintained by propagating the original lines, in which the insertions were still segregating. Insertions that failed the screen were abandoned.

**Second screen:** New homozygous insertion and non-insertion lines were extracted from each remaining segregating line, for the second time, at least 15 generations (depending on the line) after the completion of the first screen. During the interval between the first and second screens, the segregating lines were maintained by mass mating in single vial cultures. The second extraction was again done without balancers, as described above. Once again, flies of both homozygous types were cultured together in the same vials to eliminate effects of environmental differences between vials. Again, some insertions failed the screen and were abandoned.

**Third screen:** Insertions that showed the same pattern of effects in both the first and second screens were perpetuated in their segregating lines for about five more years after the second screen, on an approximately 2-week generation cycle. During this time, for about 50 consecutive generations, the lines were maintained by strict single pair mating of red-eyed virgins to their white-eyed brothers. Finally, new homozygous insertion and non-insertion lines were extracted for the third time. Again, homozygous lines were extracted without using balancers, and flies of both homozygous types were cultured together in the same vials for measurement.

**Angular offsets as a shape metric:** Vein intersection landmarks were digitized, and inter-landmark dimensions were used to quantify angular offsets for four wing-shape indexes, designated M, S, F and G. These are calculated from the four pairs of wing dimensions (D1 and D2) shown in Figure 1. The method is fully explained and illustrated in WEBER (1990). Briefly, to create each index, a curve was passed through a scatterplot of points (D1, D2) for wild-type flies, by regressing log θ on log r, where θ = arctan (D2/D1), and r = (D1² + D2²)⁰⁵. This yields a polar equation for each trait that expresses the mean allometric relation between dimensions D1 and D2 over all body sizes, in wild-type flies, and serves as a baseline. The trait value or angular offset of the point (D1, D2) for any individual wing is its rotation about the origin from the baseline, in radians. Selection on angular offset produces a rotation of this line, *i.e.*, a quantified change in allometry between D1 and D2 (WEBER 1990). The baselines for traits M, S, F and G were derived in WEBER (1990) and are still the same. This method gives a simple metric of shape that is independent of body size and most environmental influences.
Plasmid rescue and sequencing: A fragment of flanking genomic DNA to the right of each P{lacW} insertion site was retrieved by plasmid rescue (PIROTTA 1986). On its right end, P{lacW} includes a bacterial plasmid sequence with replication origin and ampicillin resistance, to the right of an EcoRI restriction site. Whole genomic DNA of male flies was digested with EcoRI, ligated, and used to transform E. coli. Clones surviving ampicillin were screened by electrophoresis. The flanking genomic DNA in typical clones was sequenced using the right end of P{lacW} as the primer: 5’-CGACGGGACCACCTTATGTTATTTCATCATG-3’ (BALLINGER and BENZER 1989; LINDSLEY and ZIMM 1992). Most clones were subsequently sequenced with a second primer starting 111 bases to the left of the right end, as a check on the exact insertion site: 5’-GGGTTAATCAACAATCATATCGCTGTCTCAC-3’.

Chromosome labeling: Polytene chromosomes from larval salivary glands were labeled using biotinylated DNA probes, including both a P-element sequence and rescued plasmid DNA, by a standard in situ protocol (cf LONG, et al. 1995). The treatments (separated by appropriate transitional baths) were 20 min 2XSSC at 65°C, 3 min 0.14M NaOH, air-drying, hybridization with biotinylated probe overnight, 20 min streptavidin/biotin treatment (Vectastain: Vector Laboratories, Inc., Burlingame, CA), 30 min in diaminobenzidine/H2O2 solution, and light Giemsa counterstaining.

RESULTS

Phenotypic screening of insertions: Each insertion was maintained in a segregating line, with inbreeding arranged to reduce genetic variation other than the insertion. To quantify insertion effects, paired homozygous lines with and without the insertion were extracted from each segregating line and compared. New homozygous lines were extracted from the same segregating lines and compared three times in consecutive screens, as inbreeding proceeded.

First screen: Of 60 original insertions, 10 could not be made homozygous, and were compared to controls as heterozygotes. In t-test comparisons (not corrected for the number of tests), we found differences in wing shape, in one or more of the four traits, which were significant at P≤0.05, in 29 of the 50 homozygous insertion lines, and in seven of the 10 still-heterozygous lines. Our criterion to retain a line in this initial screen was a difference with an individual P≤0.01, in at least one of the four shape traits (M, S, F, and G). Of the 50 homozygous lines, 17 passed this criterion. Of the 10 heterozygous lines, three passed. Only the 17 homozygous-viable insertions that passed were maintained for long-term study. No insertions had any obvious visible effects on the size, shape or venation of wings, except that one insertion (#36) sometimes showed a small gap at the distal end of vein L5.

Table 1 shows the grand means of line means, and the variances among line means, for all 50 homozygous-viable insertions, and for their controls. The difference between insertion and control grand means is not significant in t-tests for any trait (df=98). Thus the insertion of P{lacW} seems to act randomly with no consistent directional effect. This was also true in later comparisons involving only insertions that do show significant individual effects.

Model II ANOVAs, with line as a random, main effect and insertion (a random effect) nested within line, show a moderate variance component due to line for each trait (M: 19.2% of total variance, P<0.0001; S: 23.0%, P<0.0001; F: 8.9%, P=0.0006; G: 9.4%, P<0.0001) and a small, but significant variance component due to differences between insertion and non-insertion lines (M: 6.3% of total variance, P<0.0001; S: 4.2%, P<0.0001; F: 10.5%, P<0.0001; G: 3.7%, P<0.0001) for all four traits. Each insertion line sample and its paired control were measured by
the same operator, but many different operators measured the different lines. (The data set
includes 10,000 wings.) The among line variance, then, reflects both (1) the effect of the
mutagenic transposition process in 50 crosses on a common genetic background followed by
inbreeding in 50 separate segregating lines and (2) among and within operator error. Variance
due to insertions vs. controls reflects only insertion effects and within operator error.

Although the variance due to insertions is a significant component of total variance in these
ANOVAs, in all four traits, a direct comparison between the among line variances of insertion
line means and control line means is not significant in any of the four traits. As seen in Table 1,
the variances among the 50 lines with the insertion are larger in three cases out of four, but none
of the differences between variances are significant in terms of F-ratio (df=49, 49).

Second screen: At least 15 generations after the first screen, we extracted new homozygous
insertion and non-insertion lines from the segregating lines, for the 17 insertions remaining from
the first screen. These were tested again in the same way. Our criterion in the second screen
was not only a significant trait difference, but also consistency between the first and second
screens, in the pattern of values of the four traits. Two of the 17 lines (#6, #37) showed no
significant trait differences in the second screen. Three of the lines (#8, #50, #52) showed
significant differences, but in patterns that were not consistent with the first screen. The most
contradictory of these (#50) was immediately re-extracted and compared again, and showed no
significant differences in the third comparison. These five problematic lines were dropped
without further study, as either type I errors, unstable or multiple insertions, or lines with
persistent residual variation affecting the trait. The remaining 12 lines all still showed the same
approximate patterns of relative values in the four traits, and still showed significant trait
differences. Eleven of these were still significant at P≤0.01 in at least one trait. One (#47) was
only significant at P=0.036, but was retained for study.

Third screen: The 12 insertions that were retained after the second screen were maintained
in their original segregating lines for about 120 more generations in small vial cultures. During
this time, for about 50 consecutive generations, these lines were propagated by single pair
matings of red-eyed, heterozygous virgins to white-eyed, insertion-free brothers. Finally, new
homozygous insertion and non-insertion lines were extracted and compared a third time. In this
screen, all insertions showed patterns of differentiation consistent with previous measurements,
but only 10 of the 12 insertions passed initially, with P≤0.01 in at least one shape index. The
two insertions that failed this third screen (#41, #51) still showed some evidence of effects.
We immediately retested line #41 with sample sizes of N=300 for both insertions and
controls. This was the fourth test on #41, on the third pair of extracted homozygous lines. In
this test, line #41 showed P≤0.01 in one shape index, and the same pattern of differences as
before. Unfortunately, line #51 could not be retested because the fixed white-eyed control was
lost, and the segregating stock was contaminated by wild-type. This insertion survives only in a
fixed line.

Leaving aside the doubtful line #51, we conclude that 11 of these 12 insertions still show
effects of the P{lacW} insertion. Table 2 shows the mean effects of these 11 insertions,
calculated as the mean absolute difference between insertion and control lines for all three
assays, expressed as a fraction of the wild-type phenotypic standard deviation for each trait
(WEBER 1990). These differences are all too small to be reliably scored by eye.

Figure 2 shows the results of all three screens, for the 12 insertions in the third screen. Even
where individual trait differences are not significant, they tend to preserve a recognizable profile
across the four traits, though measured by different teams in each screen. Line #12 shows
significant deviations in opposite directions in index S, but otherwise retains the same overall pattern. Line #51 shows only P<0.05 in one trait in the third screen, and only weak consistency between screens.

**Gene identification and confirmation:** Flanking DNA on the right-hand (3’) side of each insertion was retrieved by plasmid rescue, for all 12 insertions that passed the second screen. Significant unique hits were obtained for all sequences with BLAST searches of the *D. melanogaster* genome (http://www.fruitfly.org/blast/). Polytenic chromosomes of each insertion line were twice labeled *in situ*, using as probes a P-element sequence and also the retrieved plasmid. Table 3 shows the approximate chromosome band locations assigned to each insertion by *in situ* labeling, and also the band locations assigned by subsequent BLAST searches using flanking DNA. In each case, the apparent *in situ* site was reasonably consistent with the computed or known band assignment according to Flybase (http://flybase.bio.indiana.edu). No secondary insertion sites were noted.

**Genes near the insertions:** Figure 3 shows each insertion in its genomic context of recognized transcription units. All insertions are shown at the same scale, including 30 kb of flanking DNA, except for line #18 (*heph*). In many cases one gene is the obvious candidate for the source of the effects, but not in all. Insertions may fall in a short interval between two genes, or in an intron containing other genes. In some cases the gene affected is uncertain because of the density of small genes near the insertion. The following summaries describe the genes most likely to be affected by each insertion, with reference to Figure 3. All information about gene functions, locations, exons and introns, alternative transcripts, etc., is based on the Flybase annotation under Release 3, if not attributed to other sources.

**Insertion #7:** The insertion site is 301 bp upstream of the 5’ end of the gene *hairy* (*h*). The closest other genes are 25.2 kb upstream and 42.4 kb downstream. *h* codes for a transcription factor of the basic helix-loop-helix type. It acts as a pair-rule gene in embryonic segmentation (INGHAM et al. 1985), and appears to be involved in patterning the nervous system (CARROLL and WHYTE 1989). It has widespread effects on sensory bristle patterning, and represses bristle formation by negative regulation of the Achaete-Scute Complex (ASC) (MOSCOSO DEL PRADO and GARCIA-BELLIDO 1984). *h* has been implicated as an important quantitative trait gene for sternopleural and abdominal bristle number (ROBIN et al. 2002; GURGANUS et al. 1999; LONG et al. 1995; SHRIMPTON and ROBERTSON 1988). Mutant alleles of *h* show effects on the placement of sensory bristles on the wing (RUSHLOW et al. 1989; THOMPSON and PRESTON 1992), and modulate the effects of *vestigial* (*vg*) and *Notch* (*N*) on the wing (ABU-ISSA and CAVICCHI, 1996). Aside from this, *h* has not previously been connected to wing morphology.

**Insertion #12:** The insertion is approximately 30 bp upstream of the *tribbles* (*trbl*) transcription start site. The closest other genes are 5.5kb upstream (*CG13248*) and 8kb downstream (*CG5571*). *trbl* codes for a cell-cycle regulatory protein, with an amino acid sequence that indicates protein serine/threonine kinase activity. This protein appears to act as a cell-cycle brake in G2-phase (MATA et al. 2000). Mutations of *trbl* have many effects on morphogenesis through their influence on the coordination of mitosis (SEHER and LEPTIN 2000; MATA et al. 2000). Overexpression of *trbl* in the posterior wing compartment causes cells inside the compartment to be fewer and larger than normal, without visibly changing the size of the compartment (MATA et al. 2000). Among genes known to interact with *trbl*, several show independent effects on the wing in some of their mutant alleles. These include *slow border cells* (*slbo*), a transcription factor (RØRTH et al. 2000); *string* (*stg*), another cell-cycle
control gene (VERHEYEN et al. 1996; MILAN et al. 1996a, 1996b; SALZBERG et al. 1997; TOBA et al. 1999); wee, a protein kinase that may be a cell-cycle control gene (PRICE et al. 2002); and Notch (N) (DE CELIS and BRAY 1997; LAWRENCE et al. 2000).

**Insertion #16:** This insertion falls in an intron of the gene gelded (gel), also designated as CG31605. Known alleles of gel include numerous recessive lethals (ROCH 1998), and one allele that causes male sterility by a recessive effect on spermatid development (CASTRILLON et al. 1993). This gene has no previously reported effects on the wing, nor any reported interactions with other genes that affect the wing. The complete transcribed region of gel is 26kb long, with many exons, at least nine known transcripts, and three transcription start sites. The insertion site falls in the central region of the gene, within the longest intron (16 kb in some transcripts), and 1.8kb upstream of the start site for the shortest transcript. The closest promoters of other genes are 18kb to the left of the insertion, at Macroglobulin complement-related (Mcr) and 16kb to the right at CG31756 (not included in Figure 3).

**Insertion #18:** The insertion is in an intron of the unusually long (142.3 kb) gene hephaestus (heph). This gene has at least 15 exons, 11 known transcripts and 5 different transcription start sites. Like insertion #16, this insertion occurs at an intermediate point in a long central intron (76 kb in most transcripts). Again, like insertion #16, the transcription start site of the shortest transcript falls in this same long intron, downstream (2,380 bp) of the insertion site. The location of the insertion in the middle of this long gene, rather far from other promoters, makes heph a strong candidate locus for the effects.

heph appears to be a spliceosome component. Its amino acid sequence contains an RNA polypyrimidine tract-binding region of four domains, already characterized in several vertebrates (DANSEREAU et al. 2002). In Drosophila, this gene has been implicated in the regulation of the Notch signaling pathway, appearing to suppress peripheral activity of Notch within the broad regions where Notch is first expressed during wing pattern formation. In homozygous clones, recessive lethal mutations in heph eliminate wing vein formation and also induce ectopic wing margin tissue, producing both effects by interaction of heph with Notch (DANSEREAU et al. 2002). heph also interacts with fringe (fng) (DANSEREAU et al. 2002), a gene that is expressed in dorsal cells of the wing (IRVINE and WIESCHAUS 1994), and that can independently induce ectopic wing vein material in hypomorphic alleles (CORREIA et al. 2003). NORGA et al. (2003) reported significant bristle number effects from a P-element insertion in heph.

**Insertion #24:** This insertion is located in the first intron of foxo, a long (30.4 kb) gene with one known transcript. The site of the insertion is 31.8 kb from the promoter of the next gene to the right (CG3153, not shown), and 13.1 kb from the promoter of the next gene to the left (CR31476, a tRNA gene). The exact location of the insertion is between the last two bases (A and G) of the intron. The right terminal base of P[lacW] would not reconstitute the AG signal. However, another AG site occurs six bases downstream, and with an eight-base target site duplication (O’HARE and RUBIN 1983), this and the original AG site may remain to the left of the insertion. Thus the insertion may be spliced out with the intron, or the intron may be spliced out leaving the insertion. Only the 5' UTR would be affected by this insertion, since the protein coding sequence begins after the second intron. This insertion could affect either transcription or translation rates of foxo, and might activate unknown alternative transcription start sites. foxo appears to have 9-14 exons, and numerous expressed sequence tags (ESTs) suggesting that variable start sites generate transcript diversity.

foxo codes for an apparent transcription factor, with a forkhead (or winged helix) DNA-
binding domain. In other organisms, proteins with this domain are involved in cell determination in early embryogenesis. Only one allele of foxo (wild-type) is known in Drosophila. Seven other transcription factors in Drosophila with forkhead domains are known. All are expressed in small, highly localized regions of embryos early in development (Häcker et al. 1992).

Insertion #25: This insertion involves two genes as the most likely targets—seven in absentia (sina) on the positive strand, and Rhodopsin 4 (Rh4), on the negative strand. Rh4 has a single long intron (8.8 kb), and sina is mainly contained within this intron, although its first transcription start site overlaps with the second exon of Rh4. The insertion site is near the middle of the 5' UTR of sina (base position 996 of 1902 bases), and 630 bp from the left end of the intron of Rh4. The closest flanking loci are all on the positive strand. Upstream, they include the nested genes CG9715 and CG32161, with their promoters 8 kb to the left of the insertion site. Downstream, they include two more genes which are both nested tightly together with sina within the same intron of Rh4. These are CG13030, with its promoter 5.3 kb to the right of the insertion, and CG13029.

Rh4 functions in the eye in phototransduction and is expressed exclusively in photoreceptor cells, so is unlikely to mediate any effect on the wing. sina codes for a protein with a single ring-type zinc finger domain, which is expressed in various tissues during all phases of development, particularly in sense organ development. Mutations in sina affect the formation of receptor cells in ommatidia, and also in sensory bristles. One mutant allele of sina causes outstretched wings (Lindsley and Zimm 1992), and another allele reduces the numbers of bristles along the anterior wing margin (Cartew and Rubin 1990). sina also interacts with many other genes, including at least six that are normally expressed in the wing and can show incidental effects on the wing. These include, for example, tramtrack (ttk) and phyllopod (phyl). ttk is expressed everywhere in the wing disk except in proneural clusters (Lehembre et al. 2000). When overexpressed it ablates almost all sensory bristle growth and severely reduces the size of the wing (Badenhorst et al. 2002). phyl and sina antagonize ttk, and overexpression of phyl in the wing causes ectopic bristle formation on the third wing vein (PI et al. 2001). The products of sina, ttk, and phyl appear to act together as a protein complex in both photoreceptor and sensory bristle differentiation (PI et al. 2001). musashi (msi) interacts with both sina and ttk in eye development (HirotA et al. 1999), and causes variation in the number of bristle support cells in sensory bristles, notably along the anterior wing margin (Nakamura et al. 1994).

GTPase-inactivating protein I (Gap1) is another gene of interest that interacts with sina. Gap1 is involved in differentiation of ommatidia and bristles and affects chromosome segregation, and Gap1 mutants also show effects on wing vein formation, including extra veinlets attached to the posterior crossvein parallel to the long veins (Gaul et al. 1992). Again, echinoid (ed) interacts with sina and also has direct effects on the differentiation of photoreceptor cells, and in one reported mutant genotype causes extra wing vein growth and enlarged wings (BaI et al. 2001). Finally, small wing (sl) has effects on the differentiation of ommatidia and also produces a short, blunted wing with extra wing vein material in two characteristic locations (Thackeray et al. 1998). A rather consistent pattern emerges for the genes in this group. They interact with sina, and, like sina, can affect the differentiation of both photoreceptors (specifically the R7 cell of the ommatidium) and also sensory bristles, including bristles on the notum and wing; and usually they also show some effects on wing shape, size, or venation.

Insertion #27: The insertion site is in the first intron of the gene sugarless (sgl). The intron is about 2.4 kb long, and the insertion is at the very beginning of the intron, between the fourth
and fifth bases. This location is close to the promoter region of the gene, because the first exon is short (406 bp). A short interval of 458 bp separates the 5’ end of sgl and the 3’ end of the next upstream gene (CG10064). The whole surrounding region is gene-dense, and the promoters of five other genes are located at ranges of 3.6 kb to 6.3 kb from the insertion site.

sgl encodes a UDP-glucose dehydrogenase. This enzyme is essential in the production of UDP-glucuronate, which is utilized in the biosynthesis of several glycosaminoglycans. Glycosaminoglycans have many structural and metabolic roles and are somehow involved in modulating the signaling of wingless (wg) (BINARI et al. 1997), a segment polarity gene, which interacts with sgl. Mutations in sgl produce a pattern of defects in embryonic cuticle that is similar to the effects of nonlethal mutations of wg (HAERRY et al. 1997). This phenotype in sgl mutants can be rescued by overexpression of wg (HÄCKER et al. 1997), and also by embryonic microinjection of exogenous heparan sulfate, a glycosaminoglycan (BINARI et al. 1997).

Insertion #36: This is the least conclusive case, in regard to the identity of the affected gene. The insertion is not located within any gene or promoter region, but is 3-4 kb from the transcription start sites of four different genes. Another complication of this case is a natural transposable element (Doc) that was discovered in our line, adjacent to the P{lacW} insertion.

The closest gene is stem cell tumor (stet), also known as rhomboid-2 (rho-2). The insertion is 118 bases to the left of the end of the 3’ UTR of transcript A (which begins at the second start site of stet), and 137 bases to the left of the end of transcript B (which begins at the first start site). The retrieved flanking sequence, to the right of P{lacW}, includes part of this 3’ UTR, and reveals the presence of an insertion of the transposable element Doc, just 22 bases inside the 3’ UTR of transcript A and three bases inside the 3’ UTR of transcript B. Thus, in our original inbred P{lacW} source line, stet most probably already carried this insertion of Doc inside the transcription unit, before P{lacW} was inserted just outside it. The retrieved flanking sequence includes ~500 bases matching the left end of Doc. Complete Doc sequences are around 5000 bases long (LINDSLEY and ZIMM 1992).

The product of stet belongs to the rhomboid-like group of proteins with seven transmembrane domains. The stet protein functions as a membrane-bound serine-peptidase in epidermal growth factor (EGF) signaling (KLÄMBT 2000). Mutations in stet affect male and female germline cells and cause defects in gonadal development (SCHULZ et al. 2002). Normal expression of stet appears to be limited to cells within male and female gonads at an early stage of differentiation. However, misexpression of stet in the wings of transgenic flies has direct effects on wing morphology, including thickening of veins and formation of ectopic vein material (GUICHARD et al. 2000; URBAN et al. 2002). Thus stet could be the source of wing shape effects, perhaps through effects on its second transcription start site. But the transcription start sites of two other genes—CG32319 and rob62A—are closer to the insertion.

The interpolation of a complete sequence of Doc to the right of P{lacW} would make CG32319 the closest transcription start site, and rob62A the next closest. Neither gene has a reported mutant phenotype but both have been classified by homology as to molecular function. CG32319 encodes a protein with N-acetyltransferase activity which is involved in acetylation of amino acids in proteins. rob62A encodes a dynein subunit protein with ATPase activity which should be involved in microtubule movement (GOLDSTEIN and GUNAWARDENA 2000).

Insertion #41: This insertion falls in a short interval of 230 bases, between two adjacent genes that both run left-to-right. The insertion is 140 bases downstream of the gene Srp54, and 90 bases upstream of the gene yippee interacting protein 2 (yip2), within the promoter region of
Aside from yip2, three other genes have their promoters within 3 kb of the insertion site, including Srp54, CG5899, and CG5885. yip2 codes for an acetyl-CoA C-acyltransferase which is found in the mitochondrion. Srp54 codes for a protein that includes an RNA-binding sequence. CG5899 codes for a DNA helicase. CG5885 codes for a signal sequence receptor component. No mutant effects have been described for any of these four loci.

Insertion #45: This insertion is inside a 2.1 kb intron, following the first exon of out at first (oaf), about 768 bp from the left end of the intron. oaf has three introns and four known transcripts, all with the same start site. The oaf gene codes for a protein of unknown functional type, involved in neurogenesis. It is transcribed in the embryonic central nervous system in segmental clusters, and in gonads of both sexes throughout development and adulthood (BERGSTROM et al. 1995). Some reported mutations in oaf are viable and have no obvious phenotypic effect; others are recessive lethal and affect the nervous system (BERGSTROM et al. 1995). oaf has no reported effects on the wing, but does appear to be expressed uniformly throughout the wing disk at low levels (MERLI et al. 1996).

Two other genes, SLY-1 homologous (Slh) and CG15393, are close to oaf but code in the opposite direction. The promoters of Slh and CG15393 are 1.4 kb and 2.0 kb from the insertion. Slh appears to be involved in membrane trafficking (LITTLETON 2000), and is not expressed in the wing (MERLI et al. 1996). CG15393, to the left of Slh, is a small gene of 121 amino acids with no recognized functional motif. The tight three-gene cluster of oaf, Slh, and CG15393 is isolated from neighboring genes by ~30 kb on the left and ~20 kb on the right.

The next gene upstream of oaf is decapentaplegic (dpp), which affects many aspects of development (SPENCER et al. 1982), and is strongly expressed in the wing. Although the coding regions of oaf and dpp are separated by 33.5 kb, most of this interval (30.0 kb) is occupied by an array of enhancers that control dpp (BLACKMAN et al. 1991; BERGSTROM et al. 1995). Perhaps insertion of the 10.7 kb P{lacW} in the oaf intron could have weak remote effects on the regulation of dpp, given the size of the region devoted to dpp transcriptional control, and the proximity of this region to the insertion site. A regulatory site for human $\alpha$-globin is also located in an intron of an adjacent gene, at a similar distance (VYAS et al. 1992).

Insertion #47: This insertion is in the Alhambra gene. This gene has four known transcripts. Two are long and include a 17 kb central intron. The other two are short and start inside this central intron, beginning with an exon that is spliced out of the long transcripts. The insertion site is near the middle of this intron, just 17 bases upstream of the common start site of the two short transcripts. About 5.6 kb upstream of the site of the insertion, and within the same long intron of Alhambra, is the small included gene Muscle LIM protein at 84B (Mlp84B), a gene that also runs right-to-left in the same direction as Alhambra. (The letters LIM stand for three homeodomain proteins with a shared motif.) The whole Alhambra transcription unit occupies 29.6 kb, so that the central insertion site is rather isolated from other genes.

Alhambra codes for a protein that is thought to incorporate two zinc ions in a domain resembling a plant homeodomain (PHD) finger, which is predicted to be involved in transcriptional regulation (BAHRI et al. 2001). Reported mutations in Alhambra have recessive effects on the larval nervous system and development rate. Mlp84B codes for a protein with a glucocorticoid receptor-like DNA-binding domain (again, a zinc-bound feature). Proteins with the LIM domain regulate cell growth and differentiation (DAWID et al., 1998). The product of Mlp84B belongs to a group of LIM proteins that regulate muscle differentiation, and the gene is expressed during differentiation at sites of muscle attachment (STRONACH et al. 1996; 1999).

Insertion #51: The insertion falls within the 5’ UTR of the gene Glycerol 3 phosphate
dehydrogenase (Gpdh), about 285 bp after the transcription start site, and about 145 bp before the first codon. Gpdh encodes an enzyme important in flight muscle metabolism (WOJTAS et al. 1997). Numerous recessive lethal and sublethal mutations of Gpdh have been reported, as well as a few mutations causing flightlessness (KOTARSKI et al. 1983). There are no reports of morphological effects. Insertion #51 may have had some variable effect on wing shape.

**DISCUSSION**

The use of paired insertions and controls, extracted from the same segregating, long-inbred stocks, with three repeated screens and multitrait comparisons, leads to high confidence in the detection of small genetic effects. Our protocol was not designed to confirm every minor effect that may have been present in the first screen, but only to select the more significant and stable effects. We found such effects in 11 of 50 random homozygous-viable insertions.

**These effects are strongly associated with the insertion of P(lacW):** No other cause would be likely to produce consistent trait profiles that segregate with the insertion in three independent extractions, separated by many generations of recombination and inbreeding. Initial variation was already low due to preliminary inbreeding of the source line. In the transposition generation, variation could arise by male recombination in mobilized hybrid males, or from mobilization-induced mutations. After transposition, much residual variation would have been eliminated by the first ten generations of single-pair matings between insertion-heterozygote virgins and their insertion-free brothers, before the first screen. Further recombination and inbreeding leading to the second and third screens would gradually homogenize and isogenize the genetic background even more. By never extracting insertions and controls except at the times of measurement, we avoided any chance for background divergence to accumulate. Balancers were never used in extractions so genetic variation could not arise from rare recombination with balancers. Some apparent effects in the first screen were not repeatable, due perhaps to unstable effects, multiple insertions or residual variation. Elimination of these problematic lines left 11 that were still stable and consistent, five years later in the third extraction and screen. During these five years, approximately 50 generations were single pair matings between heterozygous sisters and brothers lacking the insertion, increasing the likelihood that by the time of the third extraction, trait differences were caused only by the presence or absence of the insertion.

**The effects arise from gene disruption, not from P(lacW) itself:** In the region surrounding each insertion, nearby loci might be affected by interactions with specific parts of P(lacW), or simply by the insertion of 10.7 kb. But the shape changes are not a direct autonomous effect of P(lacW), because each insertion causes a unique pattern. For each trait the change was positive or negative at roughly equal frequencies among lines. The grand means of the 50 control and 50 insertion line means, from the first screen, are not significantly different in any of the four traits (Table 1). In their screen of the effects of P[ArB] insertions on bristle number, LYMAN et al. (1996) detected a significant mean directional effect that was due to the construct itself, relative to ry− controls, and this was attributed to rescue of ry− by ry+ in the insertion.

Although ANOVAs detected a significant contribution of insertion and control differences to total variance, F-ratios did not indicate significant increases in variance among insertion line means compared to controls. This is not surprising considering the relatively small number of
lines in our screen and the absence of large effects. In larger screens of P-element insertion lines, others have reported significant effects on variance in bristle number that were attributed mainly to a few insertions with large effects (MACKAY et al. 1992; LYMAN et al. 1996).

**No wing shape effects have been reported for these genes:** All the genes closest to or including each insertion have been discussed in at least one publication. Six have previously known effects on wing veins, wing bristles, or wing posture, but none were known to affect wing shape. Flybase lists about 1,900 loci (~14% of the genome) with known effects on wing morphology—including effects on size, shape, veins, bristles, etc. This whole list includes only four genes encountered in our screen (h, trbl, sgl, sina). Two other implicated genes not on the Flybase list are known to affect wing morphology in special circumstances, including heph (a recessive lethal affecting the wing only in homozygous clones (DANSEREAU et al. 2002)), and stet (affecting the wing only by ectopic expression (URBAN et al. 2002)). Thus some of these genes do have known effects on the wing, but none have been reported to affect wing shape. BUTLER et al. (2003) recently reported 56 genes with at least two-fold higher expression in the blade and hinge of wing imaginal disks, than in the body wall region. These are likely to include some that can affect wing shape, but include none of our 11.

**Diverse genetic pathways appear to affect wing shape:** The genes most likely affected by these insertions include some with prominent developmental roles, mainly outside the wing. These include four transcription factors (h, foxo, sina, Alhambra), two genes involved in signaling (sgl, stet), and one involved in cell cycle control (trbl). Also included are a putative spliceosome component (heph), a mitochondrial protein (yip2), and two genes of unknown types (gel, oaf). It is interesting that two genes previously implicated as quantitative trait genes for bristle number—hairy (ROBIN et al. 2002), and heph (NORGA et al. 2003)—now turn out to be potential genes for wing shape as well.

Mutations affecting wing morphogenesis have been reported in two other nucleus-encoded mitochondrial genes—colt (HARTENSTEIN et al. 1997) and Gart (TIONG and NASH 1990). Thus it is not implausible that a third such gene may affect wing shape, as suggested by our results for yip2. Although results for Gpdh were inconsistent, it is also not implausible that a basic metabolic enzyme could affect wing shape. For example, the rudimentary (r) locus codes for an enzyme involved in pyrimidine synthesis, and mutations at r cause variation in wing shape (FAUSTO-STERLING and HSIEH 1976).

**No effects are due to amino acid coding interruptions:** Seven insertions are in introns, four are in flanking regions, and two are in exons—with one in two categories at once (#25). The two in exons (sina and Gpdh) are both in the 5’ UTR of the first exon, where they could influence either transcription or translation. Of the four insertions in flanking regions, two are just upstream of a gene (h, trbl) in the promoter region, one is near the downstream end (stet), and one falls in a short interval between two genes, upstream of one (yip2) and downstream of the other (Srp54). Insertions near either end of a gene could influence transcription. Of the seven insertions in introns, one is in an intron of a gene (Alhambra) that contains another gene (Mlp84B). Another is in a gene (sina) that is inside an intron of another gene (Rh4). The remaining five insertions are all in introns without such complications (gel, heph, foxo, sgl, and oaf). Insertions in introns could affect transcription rates, alternative transcription start or stop sites, or the frequencies of different splicing patterns. Various P-element insertions have been reported to increase or decrease transcription rates, or to change the timing or the location of expression (reviewed in ENGELS 1989). Insertions might alter gene regulation in other ways. Alteration of the 5’ UTR might affect mRNA stability, or change the tertiary structure of mRNA.
in a way that affects gene expression (PARSCH et al. 1997). Insertions near promoters could activate secondary promoter sites.

**It is still not clear which genes produce the effects:** Eight insertions are inside genes and four are within a few hundred base pairs of one. In most cases one gene is an obvious primary candidate. However, the effect on wing shape could actually arise from some other nearby gene. Small regulatory influences can extend over local neighborhoods. For example, transgenes with transcription start sites embedded in large P-element vectors are not isolated from surrounding DNA, but show many position effects on the transgene, ranging from strong enhancement of weak promoters to complete suppression of strong promoters (HORN et al. 2003), and other effects such as new expression patterns within normally-expressing tissues (SUN et al. 1995). In this study, the distances from some of the insertion sites to other nearby loci are well within the distances of other cis regulatory sites with major effects, reported for various loci with complex regulation (BLACKMAN et al. 1991; DORSETT 1993; BACHMANN and KNUST 1998; BERMANN et al. 2002). MARONI (1994) found correlations in length among major functional gene regions, such as the 5’ and 3’ UTRs, the coding region, and the first intron, in Drosophila: therefore longer transcription units might also be influenced by more far-flung regions, even when they seem to be packed between other genes. Where genes are close together, major regulatory effects may be efficiently separated in various ways (VAZQUEZ and SCHEDL 2000; BELL et al. 2001; BURGESS-BEUSSE et al. 2002; LEVINE and TJIAN 2003). However, even if regulatory compartmentalization is qualitatively complete, minor changes in expression might occur as a kind of leakage through regulatory barriers. Figure 3 shows that some of these insertions are surrounded by many genes. The identity of the gene (or genes) causing the wing shape effect is especially uncertain in these cases.

**The percent of insertions with effects is unusually high:** Some recent screens have looked for genes with gain-of-function effects in various targets, by using a gene that normally turns on in the target to express Gal4, so as to drive overexpression or misexpression in a collection of other genes that carry insertions with a Gal4-driven transcriptional activator (RÖRTH 1996). RÖRTH et al. (1998) found various wing effects in 7% of 2,300 such insertion lines, driven by a Gal4 source covering most of the wing blade. ABDELILAH-SEYFRIED et al. (2000) found effects on sensory bristles in 5% of 2,293 insertion lines, driven by a Gal4 source in the scabrous gene. KRAUT et al. (2001) found effects on growth patterns of larval motor neurons in 5% of 2,293 lines, driven by a neuron-specific Gal4 source. PEÑA-RANGEL et al. (2002) found thorax modifications in 9% of 2,100 lines, driven by a thorax-specific Gal4 source. Finally, TSENG and HARIHAN (2002) found phenotypic effects on the eye in 2.3% of 2,296 insertion lines, driven by a Gal4 source in the eye disk.

Other screens have looked for direct effects of P-element insertions on specific quantitative traits. In a screen of 379 insertions, about 4% showed effects on avoidance of an odorant, benzaldehyde (ANHOLT et al. 1996). Two screens totaling 2,825 insertions were analyzed in NORGA et al. (2003), for P-element effects on bristle number. When insertion line means for abdominal and sternopleural bristle number were compared to the phenotypic distribution in control lines, more than 20% had phenotypes outside the 95% confidence limits, and 4-10% had phenotypes outside the 99.9% confidence limits.

A screen of metabolic effects of P-element insertions by CLARK et al. (1995) also yielded high percentages. In a screen of 263 random, single, autosomal, homozygous-viable P-element insertions, they found 153 with a significant difference from controls in one or more of 16 different traits, of which one was body weight, three were biochemical fractions, and 12 were
enzyme activities. Many insertions caused effects in more than one trait. From the published
data, the average number of lines affected per trait would have been about 24 out of 263, or
~9.1%. Since this is an average, the frequency of effects must have exceeded this in some traits.

We report positive results in 11 of our 50 lines, or 22%, supported by repeated high-
resolution tests of individual lines. Based on our small sample of lines, the percentage of all
random, autosomal, homozygous-viable insertions of P{lacW} that affect wing shape is, with
95% confidence, at least 11.5% (ROHLF and SOKAL 1981).

According to MIKLOS and RUBIN (1996), 65-75% of all genes in Drosophila have no
obvious loss-of-function phenotype. However, none of the small effects reported here are
obvious phenotypes. Perhaps most genes have small loss-of-function phenotypes, if the right
features are measured with sufficient precision.

**The genome must have a large potential for such variability:** P-elements do not insert
into DNA sites randomly, and perhaps are not even random with respect to traits (ENGELS
1989; SPRADLING et al. 1995). Still it seems clear from these results that the total number of
genes that could potentially affect wing shape must be much higher than the number that would
be segregating in any wild population. The latter number was estimated for one population
sample as at least 20, by QTL analysis of selected lines (WEBER et al. 2001), and at least 140,
by the Castle-Wright method (WEBER 1990). The number of possible effects must have some
inverse relation to their magnitude, but empirical studies are only beginning to quantify how
large the mutational target size of the genome is at different magnitudes of effect.

Transposable elements must generate similar variation in nature. P-elements can create
quantitative genetic variation (MACKAY et al. 1992; KEIGHTLEY et al., 1993), which is
selectable (MACKAY 1985; TORKAMANZEHI et al. 1992). In D. melanogaster, ~10% of the
genome consists of transposable elements, and their activity causes at least half of all
spontaneous mutations (FINNEGAN 1992). The long-term evolutionary importance of
transposable elements has been recognized in their ability to rearrange bits of the genome and to
generate small length variations (SHAPIRO 1999; MAKALOWSKI 2003), and partly depends
on how they are usually eliminated from the genome, which is still an issue (NUZHIDIN 1999).
Any other source of local variation in DNA length—such as imprecise repair of double strand
breaks (LIANG, et al. 1998)—could probably produce genetic variability in the same sites.
Recent detailed comparisons between whole genomes of related organisms show that genomes
are constantly undergoing localized reduction, expansion, and rearrangement (WATERSTON et
al. 2002). This may entail much minor regulatory variation.

These effects are all definitely microevolutionary: None of the effects are clearly visible
or qualitative except for the small occasional vein L5 gap noted in line #36. Yet we would
almost certainly have detected additional, even smaller effects if we had measured more wings
per line. We would probably also have found some large effects, as in other P-element screens
(CLARK et al. 1995; LYMAN et al. 1996), if we had looked at many lines. Our limited study
can show only an intermediate region of the distribution.

The primary motivation for this study was to explore the latent evolutionary potential of
shape traits in the wing. As P-element insertion screens are applied to more and more traits,
with increasing resolution, results show that very large portions of the genome can contribute
small effects. What is the evolutionary and adaptive significance of such findings?

If many genes can affect a trait, adaptive flexibility greatly increases. Selection will tend to
utilize the most suitable genes with the fewest pleiotropic complications. Moreover, smaller
effects allow more precise control of detail, like building with sand instead of stones. But
adaptive potential may not reside only in small effects that can accumulate at many loci. It may also depend importantly on particular small effects that have the potential to become large.

**Immediate effect is not ultimate adaptive potential:** A gene for a mitochondrial protein might affect wing shape, but would be unlikely to play a role in wing shape evolution. By contrast, a gene involved in genetic regulation or signaling might have the potential for a large role, even if its immediate mutational effects are small. The effect of an allele can increase, by selection of interacting alleles, or by incremental changes of the allele itself (STAM and LAURIE 1996). Such incremental changes could include multiple regulatory effects on timing, location, or rate of expression, as well as amino acid changes.

Because alleles can evolve, no analysis of quantitative trait genes, in a trait that is assumed to have been changed by selection, can with certainty establish the original distribution of selected effects by measuring the present distribution of fixed differences in effect, as is often assumed. Because alleles can evolve, a small mutational effect could be seen as the tip of an unknown wedge. Selected effects may start out small, but a bit part might become a starring role in a future adaptation, and could perhaps achieve that status by changes “insensibly small.”

**Acknowledgements:**

We thank two anonymous reviewers and Jeff Walker for comments. This project was supported by National Science Foundation grant DEB-9407005 to K. W. and National Institutes of Health grant R15GM65116-01 to D. C.
Table 1. Means and variances of homozygous control and insertion line means in first screen.

<table>
<thead>
<tr>
<th>Trait</th>
<th>50 Control Line Means</th>
<th>50 Insertion Line Means</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Grand Mean</td>
<td>Variance</td>
</tr>
<tr>
<td>M</td>
<td>+8.0 x 10^-3</td>
<td>2.4 x 10^-5</td>
</tr>
<tr>
<td>S</td>
<td>-2.1 x 10^-3</td>
<td>3.5 x 10^-5</td>
</tr>
<tr>
<td>F</td>
<td>-7.3 x 10^-3</td>
<td>1.6 x 10^-5</td>
</tr>
<tr>
<td>G</td>
<td>-1.9 x 10^-2</td>
<td>8.6 x 10^-6</td>
</tr>
</tbody>
</table>

Means are in radians of angular offset from wild-type baseline.
Table 2. Mean absolute trait differences, from three screens, in SD of wild flies.

<table>
<thead>
<tr>
<th>Insertion #</th>
<th>M</th>
<th>S</th>
<th>F</th>
<th>G</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>0.37*</td>
<td>0.11</td>
<td>0.50*</td>
<td>0.15</td>
</tr>
<tr>
<td>12</td>
<td>0.16</td>
<td>0.00</td>
<td>0.33*</td>
<td>0.52*</td>
</tr>
<tr>
<td>16</td>
<td>0.06</td>
<td>0.30*</td>
<td>0.37*</td>
<td>0.16</td>
</tr>
<tr>
<td>18</td>
<td>0.66*</td>
<td>1.02*</td>
<td>0.93*</td>
<td>0.15</td>
</tr>
<tr>
<td>24</td>
<td>0.24</td>
<td>0.15</td>
<td>0.60*</td>
<td>0.48*</td>
</tr>
<tr>
<td>25</td>
<td>0.74*</td>
<td>0.42*</td>
<td>0.30</td>
<td>0.10</td>
</tr>
<tr>
<td>27</td>
<td>0.21</td>
<td>0.15</td>
<td>0.22</td>
<td>0.39*</td>
</tr>
<tr>
<td>36</td>
<td>0.22</td>
<td>0.54*</td>
<td>0.15</td>
<td>0.05</td>
</tr>
<tr>
<td>41</td>
<td>0.18*</td>
<td>0.18</td>
<td>0.10</td>
<td>0.05</td>
</tr>
<tr>
<td>45</td>
<td>0.37*</td>
<td>0.35*</td>
<td>0.58*</td>
<td>0.14</td>
</tr>
<tr>
<td>47</td>
<td>0.21</td>
<td>0.01</td>
<td>0.23</td>
<td>0.23*</td>
</tr>
</tbody>
</table>

SD from WEBER (1990). * Difference had P < 0.01 in at least two screens.
Table 3. Insertion site identifications by *in situ* labeling and by BLAST search.

<table>
<thead>
<tr>
<th>Insertion</th>
<th>Gene(s)</th>
<th>Band: In situ/Flybase</th>
<th>DNA Site</th>
</tr>
</thead>
<tbody>
<tr>
<td>7 (+)</td>
<td>hairy (h) (+)</td>
<td>66D/66D10</td>
<td>301 bp upstream</td>
</tr>
<tr>
<td>12 (-)</td>
<td>tribbles (trbl) (-)</td>
<td>77B/77C1</td>
<td>31 bp upstream</td>
</tr>
<tr>
<td>16 (-)</td>
<td>CG31605, or gelled (gel) (+)</td>
<td>28D/28E3-E5</td>
<td>in 16 kb intron</td>
</tr>
<tr>
<td>18 (-)</td>
<td>hephaestus (heph) (-)</td>
<td>100F/100D3-E1</td>
<td>in 76 kb intron</td>
</tr>
<tr>
<td>24 (+)</td>
<td>foxo (+)</td>
<td>88A/88A6-A8</td>
<td>in short intron</td>
</tr>
<tr>
<td>25 (+)</td>
<td>seven in absentia (sina) (+) Rhodopsin 4 (Rh4) (-)</td>
<td>73C/73D2-D3</td>
<td>in exon (5’ UTR) in 8.8 kb intron</td>
</tr>
<tr>
<td>27 (+)</td>
<td>sugarless (sgl) (-)</td>
<td>65D/65D5</td>
<td>in 2.4 kb intron</td>
</tr>
<tr>
<td>36 (+)</td>
<td>stem cell tumor (stet) (-)</td>
<td>62A/62A2</td>
<td>downstream end a</td>
</tr>
<tr>
<td>41 (+)</td>
<td>yippee interacting protein 2 (yip2) (+) Srp54 (+)</td>
<td>30C/30E4</td>
<td>90 bp upstream 140 bp downstream</td>
</tr>
<tr>
<td>45 (+)</td>
<td>out at first (oaf) (+)</td>
<td>22F/22F3</td>
<td>in 2.1 kb intron</td>
</tr>
<tr>
<td>47 (+)</td>
<td>Alhambra (-)</td>
<td>84B/84B-C</td>
<td>in intron</td>
</tr>
<tr>
<td>51 (-)</td>
<td>Glycerol-3-phosph. dehyd. (Gpdh) (+)</td>
<td>26A/26A</td>
<td>in exon (5’ UTR)</td>
</tr>
</tbody>
</table>

Polarity of insertion or gene is indicated by + or –.

a Insertion #36: The 3’ UTR of *stet* contains an insertion of Doc, 140 bp to right of P{lacW}.
FIGURE 1.—The four size-independent shape indexes (M, S, F, and G) are defined by polar equations that express the allometric relation between paired dimensions (D1 and D2), in wild-type male flies.

FIGURE 2.—Individual differences between insertion and non-insertion homozygotes, in three separate extractions from the same inbred lines, for four shape indexes (M, S, F, and G), in males. Solid, dashed and dotted lines correspond to first, second and third extractions. Asterisks mark all differences with individual P<0.05. Insertion line numbers are shown to right. Line #41 was tested a fourth time on the third extraction (narrow solid line).

FIGURE 3.—P{lacW} insertions in genomic context. Each diagram shows an identical-sized region of 30 kb, except for No. 18 which includes 145 kb. Items above or below the central lines are in plus or minus orientation, respectively. Arrows show transcription start sites, including multiple start sites for some genes. Exons are thick, introns are thin. The diagrams do not show some very short introns. Gene symbols are given according to Flybase. Large open triangles show the location and orientation of P{lacW}, and are proportional to its length of 10.7 kb. The small filled triangle in No. 36 marks the site of an insertion of Doc, of unknown length.
LITERATURE CITED


ABU-ISSA, R. and S. CAVICCHI, 1996 Genetic interactions among *vestigial*, *hairy*, and *Notch* suggest a role of *vestigial* in the differentiation of epidermal and neural cells of the wing and haltere of *Drosophila melanogaster*. J. Neurogenet. **10**: 239-246


BERMAN, B. P., Y. NIBU, B. D. PFEIFFER, P. TOMANCAK, S. E. CELNIKER, et al., 2002 Exploiting transcription factor binding site clustering to identify cis-regulatory modules involved in pattern formation in the *Drosophila* genome. PNAS **99**: 757-762


BINARI, R. C., B. E. STAVELEY, W. A. JOHNSON, R. GODAVARTI, R. SASISEKHarAN,
et al., 1997  Genetic evidence that heparin-like glycosaminoglycans are involved in wingless signaling. Development 124: 2623-2632


CORREIA, T., V. PAPAYANNOPOULOS, V. PANIN, P. WORONOFF, J. JIANG, et al., 2003 Molecular genetic analysis of the glycosyltransferase Fringe in Drosophila. PNAS 100: 6404-6409


DE CELIS, J. F. and S. BRAY, 1997 Feed-back mechanisms affecting Notch activation at the
dorsoventral boundary in the Drosophila wing. Development **124**: 3241-3251

DORSETT, D., 1993 Distance-independent inactivation of an enhancer by the suppressor of Hairy-wing DNA-binding protein of Drosophila. Genetics **134**: 1135-1144


GURGANUS, M. C., S. V. NUZHDIN, J. W. LEIPS and T. F. C. MACKAY, 1999 High-resolution mapping of quantitative trait loci for sternopleural bristle number in Drosophila melanogaster. Genetics **152**: 1585-1604


HARTENSTEIN, K., P. SINHA, A. MISHRA, H. SCHENKEL, I. TÖRÖK et al., 1997 The congested-like tracheae gene of Drosophila melanogaster encodes a member of the mitochondrial carrier family required for gas-filling of the tracheal system and expansion of the wings after eclosion. Genetics 147: 1755-1768


LAWRENCE, N., T. KLEIN, K. BRENNAN and A. M. ARIAS, 2000 Structural requirements for Notch signaling with Delta and Serrate during the development and patterning of the wing disc of Drosophila. Development 127: 3185-3195


MAKALOWSKI, W., 2003 Not junk after all. Science 300: 1246-1247


specificity mediates the independent regulation of neighboring genes. Genes Dev. 10: 1260-1270


NUZHDIN, S. V., 1999 Sure facts, speculations, and open questions about the evolution of transposable element copy number. Genetica 107: 129-137


PHILLIPS, P. C., 1999 From complex traits to complex alleles. Trends Genet. 15: 6-8


RØRTH, P., K. SZABO and G. TEXIDO, 2000 The level of C/EBP protein is critical for cell migration during Drosophila oogenesis and is tightly controlled by regulated degradation. Molec. Cell 6: 23-30


SEHER, T. C. and M. LEPTIN, 2000 Tribbles, a cell-cycle brake that coordinates proliferation and morphogenesis during Drosophila gastrulation. Current Biology 10: 623-629

SHAPIRO, J. A., 1999 Transposable elements as the key to a 21st century view of evolution. Genetica 107: 171-179


STAM, L. F. and C. C. LAURIE, 1996 Molecular dissection of a major gene effect on a quantitative trait: the level of alcohol dehydrogenase expression in Drosophila melanogaster Genetics 144: 1559-1564


THACKERAY, J. R., P. C. W. GAINES, P. EBERT, and J. R. CARLSON, 1998 small wing encodes a phospholipase C-(g) that acts as a negative regulator of R& development in Drosophila. Development 125: 5033-5042


TORKAMANZEH, A., C. MORAN and F. W. NICHOLAS, 1992 P element transposition contributes substantial new variation for a quantitative trait in *Drosophila melanogaster*. Genetics **131**: 73-78

TSENG, A.-S. K., and I. K. HARIHARAN, 2002 An overexpression screen in *Drosophila* for genes that restrict growth or cell-cycle progression in the developing eye. Genetics **162**: 229-243

URBAN, S., J. R. LEE and M. FREEMAN, 2002 A family of rhomboid intramembrane proteases activates all *Drosophila* membrane-tethered EGF ligands. EMBO J. **21**: 4277-4286


WEBER, K. E., 1990 Selection on wing allometry in *Drosophila melanogaster*. Genetics **126**: 975-989

WEBER, K. E., 1992 How small are the smallest selectable domains of form? Genetics **130**: 345-353

WEBER, K. E., R. EISMAN, S. HIGGINS, L. MOREY, A. PATTY et al., 2001 An analysis of polygenes affecting wing shape on chromosome 2 in *Drosophila melanogaster*. Genetics **159**: 1045-1057


$\theta = 0.6263 \, r - 0.149$

$\theta = 0.5527 \, r - 0.253$

$\theta = 0.4048 \, r - 0.043$

$\theta = 0.7588 \, r - 0.135$