

Sensitivity of Allelic Divergence to Genomic Position: Lessons from the *Drosophila tan* Gene

Alisha V. John,* Lisa L. Sramkoski,* Elizabeth A. Walker,[†] Arielle M. Cooley,[†] and Patricia J. Wittkopp*^{†,1}

*Department of Molecular, Cellular, and Developmental Biology and [†]Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, Michigan 48109

ABSTRACT To identify genetic variants underlying changes in phenotypes within and between species, researchers often utilize transgenic animals to compare the function of alleles in different genetic backgrounds. In *Drosophila*, targeted integration mediated by the Φ C31 integrase allows activity of alternative alleles to be compared at the same genomic location. By using the same insertion site for each transgene, position effects are generally assumed to be controlled for because both alleles are surrounded by the same genomic context. Here, we test this assumption by comparing the activity of *tan* alleles from two *Drosophila* species, *D. americana* and *D. novamexicana*, at five different genomic locations in *D. melanogaster*. We found that the relative effects of these alleles varied among insertion sites, with no difference in activity observed between them at two sites. One of these sites simply silenced both transgenes, but the other allowed expression of both alleles that was sufficient to rescue a mutant phenotype yet failed to reveal the functional differences between the two alleles. These results suggest that more than one insertion site should be used when comparing the activity of transgenes because failing to do so could cause functional differences between alleles to go undetected.

KEYWORDS

allele-specific
transgenic
analysis
position effect
pigmentation
tan

Understanding the genetic basis of phenotypic change remains a pressing challenge for evolutionary biology. Addressing this challenge requires the identification of the genes contributing to phenotypic divergence as well as the specific changes within those genes that alter their function (Stern and Orgogozo 2008). Linkage mapping and genome-wide association studies (GWAS) are often used to identify regions of the genome associated with phenotypic divergence (Martin and Orgogozo 2013); however, these approaches must be supplemented with functional tests to demonstrate the phenotypic consequences of individual genes and sequence changes. This functional testing is often accomplished through transgenic analysis that evaluates the effects of a specific gene or region of a gene in different genetic backgrounds.

In *Drosophila*, the activity of divergent alleles is typically compared using transgenes inserted into the genome by transposon-mediated transformation (Wittkopp 2006). Most transposons used for this purpose (e.g., *P-elements*, *piggyBac*, and *Hermes*) insert a transgene semi-randomly into the genome (Engels 1996; Garza *et al.* 1991; Guimond *et al.* 2003; Handler and Harrell 1999; Smith *et al.* 1993; Spradling and Rubin 1983), which is not ideal because the genomic position of a gene can affect its activity, a phenomenon known as “position effect” (Sturtevant 1925; Wilson *et al.* 1990). These position effects can result from chromatin structure at the insertion site (Huisinga *et al.* 2016; Levis *et al.* 1985; Wilson *et al.* 1990) and/or interactions between the sequence of the transgene and the surrounding DNA that affect expression of the transgene (Venken and Bellen 2007; Wilson *et al.* 1990). The former generally affects the expression level of the transgene, whereas the latter can impact its expression level and/or spatiotemporal regulation. The extent of position effects has been hypothesized to be the product of two variables: (i) the strength of regulatory elements at the genomic location in which the transgene is inserted, and (ii) the susceptibility of the regulatory sequences in the transgene to altered activity (Wilson *et al.* 1990). The addition of insulator sequences flanking a transgene can reduce the effects of surrounding genomic context on its activity (Gdula *et al.* 1996; Kuhn and Geyer 2003; Silicheva *et al.* 2010).

Position effects are especially problematic when comparing activity among transgenes expected to vary in subtle ways. Targeted insertion of transgenes in *Drosophila*, most notably using the bacteriophage Φ C31

Copyright © 2016 John *et al.*

doi: 10.1534/g3.116.032029

Manuscript received June 6, 2016; accepted for publication July 19, 2016; published Early Online July 21, 2016.

This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Supplemental material is available online at www.g3journal.org/lookup/suppl/doi:10.1534/g3.116.032029/-/DC1

¹Corresponding author: Department of Ecology and Evolutionary Biology, 830 North University Avenue, University of Michigan, Ann Arbor, MI 48109. E-mail: wittkopp@umich.edu



Figure 1 Body color of *D. americana* and *D. novamexicana*. *D. novamexicana* (right) has evolved lighter body pigmentation since it diverged from the common ancestor shared with *D. americana* (left). *D. americana* has retained the darker body pigmentation shared by all other members of the *virilis* group.

integrase system (Groth *et al.* 2004), can help control for position effects by inserting each transgene of interest into the same genomic position of otherwise identical genomes (Venken and Bellen 2005). With large collections of “landing sites” (sequences that mediate integration of the transgene) for Φ C31-mediated transformation available (Bateman *et al.* 2006; Bischof *et al.* 2007; Venken *et al.* 2006), this method has become the standard for comparing the activity of related alleles in *Drosophila*. Typically, such a study compares a set of transgenic lines in which each transgene is integrated independently into a chosen landing site, with a single landing site used in most cases (*e.g.*, Cande *et al.* 2009; Duncan *et al.* 2010; Frankel *et al.* 2010; Haley *et al.* 2010; Joshi *et al.* 2010; Kalay and Wittkopp 2010; Perry *et al.* 2011; Rebeiz *et al.* 2011; Sayal *et al.* 2011). The use of a single landing site for such studies is justified by the assumption that all alleles compared will be affected similarly by the surrounding genomic context (Wimmer 2005). But is this true? Are sets of related transgenes influenced similarly by the surrounding DNA sequence?

Here, we test this assumption by examining the impact of position effects on a comparison of orthologous alleles that contribute to phenotypic divergence between a pair of closely related *Drosophila* species. Specifically, we compare the effects of *tan* alleles from *D. americana* and *D. novamexicana* integrated into the *D. melanogaster* genome at five different genomic locations. *D. americana* and *D. novamexicana* diverged ~400,000 years ago (Caletka and McAllister 2004; Morales-Hojas *et al.* 2008) and have evolved dramatic differences in adult pigmentation (Throckmorton 1982); *D. americana* has a brown body, whereas *D. novamexicana* has a yellow body (Figure 1). Prior work has shown that these differences in pigmentation are due in part to divergent sites located in the *tan* gene (Wittkopp *et al.* 2009). As described below, we found that position effects influenced whether or not a difference in activity could be detected between these two species-specific alleles of *tan*. Further analysis showed that the ability to detect a difference in activity was related to level of expression from the *tan* transgene at each site. These findings suggest that differences between transgenes should be assessed using multiple landing sites.

MATERIALS AND METHODS

Generation of transgenic flies

Previously constructed transgenes containing *D. americana* or *D. novamexicana tan* (Wittkopp *et al.* 2009) were injected into *D. melanogaster*

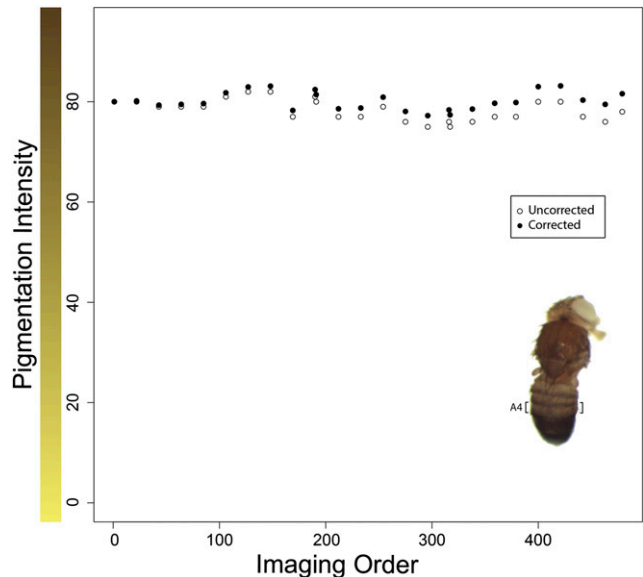


Figure 2 Measurements of pigmentation intensity in a control sample varied slightly during image collection. Raw median pigmentation intensity in tergite A4 (insert) is plotted against imaging order for the reference cuticle (open circles). All images were taken during in a single sitting without adjustment of lighting, focus, or other imaging parameters; the small ($\beta = -0.0075$), yet significant (p -value = 0.008), downward trend in pigmentation intensity as the imaging progressed, presumably as a result of changes in ambient lighting or other uncontrolled imaging variables. An imaging order correction was therefore applied to all measurements, as described in the *Materials and Methods* section. Corrected median pigmentation intensity values for the same images are also plotted against imaging order (closed circles) to show the effects of this correction.

using the Φ C31 integrase system. The transgenes contained all exonic and intronic sequences of *tan*, as well as 4.1 kb of sequence 5' of *tan* and 3.6 kb of sequence 3' of *tan* in a *piggyBac* vector (Horn and Wimmer 2000) containing an attB site used for Φ C31-mediated transformation and Pax6-EGFP, an eye-expressing green fluorescence marker used to detect successful integration (Wittkopp *et al.* 2009). In addition to noncoding and synonymous changes, these *D. americana* and *D. novamexicana tan* transgenes differ by two amino acids; however, these amino acid differences are not fixed between species and are thus unlikely to be responsible for the species-specific differences in pigmentation (Wittkopp *et al.* 2009). Each *D. melanogaster* host genotype carried a transgene on the X-chromosome using the *vasa cis*-regulatory sequences to express the Φ C31 integrase specifically in the germ-line and a single attP site located on the second (51C – BDSC #24482, 58A – BDSC #24484), third (86Fa – BDSC #24486, 86Fb – BDSC #24749), or fourth (102D – BDSC #24488) chromosome (Bischof *et al.* 2007). These lines were selected because they contain an eye-expressing red fluorescent protein (RFP) as a visible marker for the landing site; this is in contrast to the majority of strains containing attP landing sites that are commonly used, which use a copy of the *yellow* gene (which restores dark pigmentation in *yellow* mutant flies) as a visible marker for the landing site. GenetiVision (Houston, Texas) performed all DNA preparations and embryo injections according to their standard protocols (<http://www.genetivision.com/>). Transformant flies (expressing green fluorescent protein in their eyes) were used to establish lines homozygous for each transgene in a *D. melanogaster* background carrying loss-of-function mutations in the X-linked genes *tan*, *yellow*, and *white*. The

■ **Table 1 Analysis of variance (ANOVA) in pigmentation indicates that *tan* allelic identity, genomic location, and the interaction between allele and genomic location affect pigmentation intensity**

| Source of Variation | Degrees of Freedom | Sum of Squares | Mean Square | F value (F_0) | p-Value |
|----------------------------------|--------------------|--------------------|--------------------|--------------------|------------------------|
| 1) <i>tan</i> transgene identity | 2 | 3.08×10^4 | 1.54×10^4 | 1.54×10^2 | 2.76×10^{-46} |
| 2) Landing site | 4 | 5.74×10^4 | 1.44×10^4 | 1.43×10^2 | 1.89×10^{-67} |
| 3) Interaction between 1 and 2 | 4 | 2.54×10^3 | 6.34×10^2 | 6.33 | 6.73×10^{-5} |
| 4) Residuals | 290 | 2.90×10^4 | 90.5 | N/A | N/A |

N/A, not applicable.

mutant *D. melanogaster tan* allele allowed us to test for rescue of the *tan* mutant pigmentation phenotype by the heterologous *tan* alleles contained in the transgenes; the *yellow* mutant allele reduced the amount of black pigment present in these flies, providing a more sensitive assay for changes in abdominal pigmentation caused by the transgenes; and the *white* allele allowed for easier visualization of the eye-expressing fluorescent transformation marker.

Drosophila husbandry, collection, and abdominal cuticle dissection

For each line to be analyzed, virgin females were mated with males on standard yeast-glucose media at 20°. Upon formation of pupae, parents were removed and the offspring were allowed to continue development. Male offspring were collected 0–1 d posteclosion and aged to 7–8 d. Flies were stored in 10% glycerol in ethanol prior to dissection.

To harvest abdominal cuticles, 7–8-day-old males were removed individually from the 10% glycerol in ethanol solution and placed on a glass slide. Using a razor blade, the abdomen was separated from the rest of the body then cut along the lateral edge parallel with the anterior-posterior axis. The dorsal half of the abdomen was soaked overnight in a solution of phosphate buffered saline (PBS; 1.4 M NaCl, 27 mM KCl, 190 mM Na_2HPO_4 , and 18 mM KH_2PO_4 , adjusted to pH 7.4 with 1 M HCl). After soaking overnight, a single dorsal half of abdominal cuticle was removed from the PBS and placed on a glass slide, dorsal (cuticle) side down. Using forceps, the abdominal cuticle was cleared of any remaining debris. The cleaned cuticle was then mounted dorsal side up in polyvinyl alcohol (PVA) mounting media (BioQuip) on a clean glass slide, covered with a coverslip, and the coverslip was sealed with clear nail polish. This process was repeated for all genotypes analyzed, with 17–35 (mean = 27) flies analyzed for each genotype. To minimize effects of

any day-to-day differences in dissections, all genotypes were dissected during each dissection session.

Image collection and processing

Dissected abdominal cuticles were imaged in a single session using a Leica MZ6 microscope and Scion (CFW-1308C) camera operated via TWIN driver in Adobe Photoshop. Magnification was set to 3.2 with ring light illumination at ~75%. At the beginning of the imaging session, auto white balance (AWB) was used, resulting in a configuration of Gamma 0.605, Red Gain -1.4 db, Green Gain 5.4 db, and Blue Gain 8.9 db with Red Boost and Blue Boost active. These settings were not changed throughout the imaging session. Imaging was conducted at night to minimize changes in ambient lighting. Images were taken slide-by-slide (2 cuticles/slides, cuticles imaged individually) with samples arranged such that no more than two cuticles from the same line were imaged consecutively. A “reference” image of the same dissected cuticle was taken approximately every 10 slides to allow us to evaluate the consistency of the image collection, processing, and analysis pipeline during the multi-hour imaging session.

All images were compiled into a single document in Adobe Photoshop and the “Levels” function was used to adjust the color of all images simultaneously so that the images more closely matched the cuticle appearance visible by eye. This ensured that an identical color adjustment was applied to all photos.

Quantifying pigmentation

Using ImageJ (Rasband 1997–2016), the area of dorsal abdominal cuticle known as abdominal tergite 4, or “tergite A4,” (insert, Figure 2) was manually selected using the polygon tool, excluding any regions containing cracks, holes, or overlapping regions. Measurements of pixel intensity (area, mean, standard deviation, mode, min, max, and

■ **Table 2 Pairwise t-tests show which transgenes inserted at which insertion sites alter pigmentation relative to *D. melanogaster tan* mutants, as well as which landing sites show evidence of functional differences between the *D. americana* and *D. novamexicana tan* alleles**

| | | 51C | | 58A | | 86Fa | | 86Fb | | 102D |
|-------------------|----------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|
| <i>tan</i> mutant | | <i>D. amer</i> | <i>D. nova</i> | <i>D. amer</i> | <i>D. nova</i> | <i>D. amer</i> | <i>D. nova</i> | <i>D. amer</i> | <i>D. nova</i> | <i>D. amer</i> |
| 51C | <i>D. amer</i> | 4×10^{-9} | — | — | — | — | — | — | — | — |
| 51C | <i>D. nova</i> | 8×10^{-8} | 0.411 | — | — | — | — | — | — | — |
| 58A | <i>D. amer</i> | 5×10^{-18} | 3×10^{-10} | 2×10^{-11} | — | — | — | — | — | — |
| 58A | <i>D. nova</i> | 4×10^{-12} | 1×10^{-3} | 1×10^{-4} | 5×10^{-5} | — | — | — | — | — |
| 86Fa | <i>D. amer</i> | 4×10^{-24} | 3×10^{-14} | 1×10^{-15} | 0.317 | 2×10^{-7} | — | — | — | — |
| 86Fa | <i>D. nova</i> | 2×10^{-21} | 4×10^{-7} | 1×10^{-8} | 1×10^{-3} | 0.093 | 2×10^{-6} | — | — | — |
| 86Fb | <i>D. amer</i> | 3×10^{-23} | 1×10^{-15} | 9×10^{-17} | 0.012 | 2×10^{-9} | 0.064 | 1×10^{-8} | — | — |
| 86Fb | <i>D. nova</i> | 9×10^{-17} | 1×10^{-7} | 7×10^{-9} | 0.064 | 0.011 | 2×10^{-3} | 0.195 | 2×10^{-5} | — |
| 102D | <i>D. amer</i> | 0.125 | 1×10^{-4} | 1×10^{-3} | 9×10^{-14} | 3×10^{-8} | 4×10^{-15} | 8×10^{-11} | 9×10^{-17} | 1×10^{-11} |
| 102D | <i>D. nova</i> | 0.088 | 3×10^{-7} | 7×10^{-6} | 7×10^{-17} | 8×10^{-11} | 2×10^{-23} | 3×10^{-20} | 3×10^{-22} | 2×10^{-15} |

p-values adjusted by the Benjamini and Hochberg method from all possible pairwise t-tests using unpooled standard deviation are shown for each pair of genotypes compared. The *tan* mutant column shows results from comparisons between each transgenic genotype and the *tan* mutant (no transgene) control. Note that neither transgene darkened pigmentation relative to the *tan* mutant when inserted at landing site 102D. Highlighted boxes indicate comparisons between the *D. americana* and *D. novamexicana tan* alleles inserted at the same landing site. Significant differences ($p < 0.05$) in median pigmentation were observed for transgenes inserted at 58A, 86Fa, and 86Fb, but not 51C or 102D. *D. amer*, *D. americana*; *D. nova*, *D. novamexicana*.

■ **Table 3 Interaction between allelic identity and genomic location remains significant after excluding silenced transgenes**

| Source of Variation | Degrees of Freedom | Sum of Squares | Mean Square | F Value (F_0) | p-Value |
|----------------------------------|--------------------|--------------------|--------------------|--------------------|------------------------|
| 1) <i>tan</i> transgene identity | 2 | 3.53×10^4 | 1.77×10^4 | 1.66×10^2 | 6.39×10^{-46} |
| 2) Landing site | 3 | 2.41×10^4 | 8.05×10^3 | 75.5 | 2.00×10^{-34} |
| 3) Interaction between 1 and 2 | 3 | 1.54×10^3 | 5.13×10^2 | 4.82 | 2.83×10^{-3} |
| 4) Residuals | 240 | 2.56×10^4 | 1.07×10^2 | N/A | N/A |

Results from analysis of variance (ANOVA) after excluding flies with transgenes inserted at landing site 102D are shown. N/A, not applicable.

median) were taken for each selection. These results were compiled into a single Microsoft Excel spreadsheet where other identifying information was then added, including imaging order (ranging from 1 to 479), allele (no transgene control, *D. americana*, and *D. novamexicana*), and landing site (control, 51C, 58A, 86Fa, 86Fb, and 102D). Since ImageJ quantifies pigmentation (pixel intensity of a grayscale image) on a 0–255 scale (dark–light), we subtracted the reported pixel intensity from 255 so that darker cuticles had a higher pigmentation score. This file was then saved as a .csv file for statistical analysis in R.

Data analysis

Median pigmentation intensity of tergite A4 for each sample reported by ImageJ was analyzed using R v3.2.5 (R Development Core Team 2016). Median pigmentation was chosen for analysis instead of mean pigmentation intensity to minimize the impact of outlier (excessively white or black) pixels.

To test for systematic changes in imaging conditions that might have occurred during the imaging session, a linear regression was performed comparing median pigmentation values from the reference cuticle and the image order number. A small but significant regression coefficient ($\beta = -0.0075$, p-value = 0.008) was observed, so a correction for imaging order was applied to each median by subtracting (image order number * -0.0075) from the original median value. The differences in reference cuticle values before and after applying this correction are shown in Figure 2. Note that all analyses described below were also performed on data without this correction and produced the same pattern of statistically significant results (data not shown).

Median pigmentation intensity of tergite A4 for each sample reported by ImageJ was then fitted to the following model to test effects of landing site, allele, and the interaction between the two:

$$Y_{ijk} = site_i + allele_j + site \times allele_{ij} + e_{ijk}$$

Pairwise *t*-tests using unpooled standard deviations were then performed on the corrected pigmentation medians to identify which comparisons among *tan* alleles and/or landing sites were statistically significant. Statistical significance was assessed using p-values adjusted for multiple testing by the Benjamini and Hochberg method (Benjamini and Hochberg 1995) as implemented in the pairwise.t.test function in R.

Measuring relative expression of *tan* transgenes at each genomic location

To test for differences in the expression level of transgenes inserted at each genomic location in *D. melanogaster*, relative expression levels of the *D. americana tan* transgene were measured using pyrosequencing (Wittkopp 2011). Specifically, we measured the mRNA abundance of the *D. americana tan* allele inserted at each genomic location relative to the mRNA abundance of the *D. novamexicana tan* allele inserted at the 86Fa landing site. This *D. novamexicana* genotype was chosen as the internal reference point for measurements of *D. americana tan* expression because it caused an intermediate pigmentation phenotype, sug-

gesting it might also have an intermediate level of expression. The P14–P15 pupal stage was analyzed because *D. americana* and *D. novamexicana tan* have previously been shown to be most highly expressed during this time (Cooley *et al.* 2012). Pupal heads and wings were removed to avoid measuring *tan* expression in those tissues, focusing our measurements on expression in the thorax and abdomen where pigmentation phenotypes are most apparent.

For each landing site, both genomic DNA and total RNA were extracted from three replicate samples, each containing six dissected pupae expressing *D. americana tan* and six dissected pupae expressing *D. novamexicana tan*. cDNA was reverse transcribed from extracted RNA using a polyT primer for each sample. Both genomic DNA and cDNA were analyzed by pyrosequencing as described in Wittkopp (2011). PCR primers used to amplify the sequence analyzed (which was located in exon 7) were 5'-GATGCTGAAGTCCAGCGTGTC-3' and 5'-biotin-CAGCCGCCAGTGACATCA-3', and the primer used for pyrosequencing had the sequence 5'-CGAGCACGATGTCCG-3'. All measurements were then normalized to the relative expression of the *D. americana tan* transgene inserted at landing site 86Fa to compare expression among the *D. americana tan* transgenes at different landing sites.

Supplemental files

Supplemental Material, File S1 contains pigmentation measures from all individual cuticles. File S2 contains the code used to analyze the pigmentation measures. File S3 contains the raw and analyzed pyrosequencing data.

Data availability

The authors state that all data necessary for confirming the conclusions presented in the article are represented fully within the article.

RESULTS AND DISCUSSION

To test the assumption that position effects are negligible when comparing divergent alleles of the same gene at a single genomic location, we transformed *D. americana* and *D. novamexicana tan* alleles into five different genomic locations in *D. melanogaster* (51C, 58A, 86Fa, 86Fb, and 102D). Each of these transgenic lines was then crossed with *D. melanogaster yellow*, *white*, and *tan* mutants (see *Materials and Methods* for full genotype) to move the transgenes into genetic backgrounds lacking a functional copy of the *D. melanogaster tan* gene. Prior work has shown that the difference in body color seen between *D. americana* and *D. novamexicana tan* transgenes significantly increase abdominal pigmentation in a *D. melanogaster tan* mutant (Wittkopp *et al.* 2009). The transgenic *tan* allele from the more darkly pigmented *D. americana* was reported to increase pigmentation significantly more than the transgenic *tan* allele from the more lightly pigmented *D. novamexicana*, indicating that there is functional divergence between these species-specific alleles that affects pigmentation (Wittkopp *et al.* 2009).

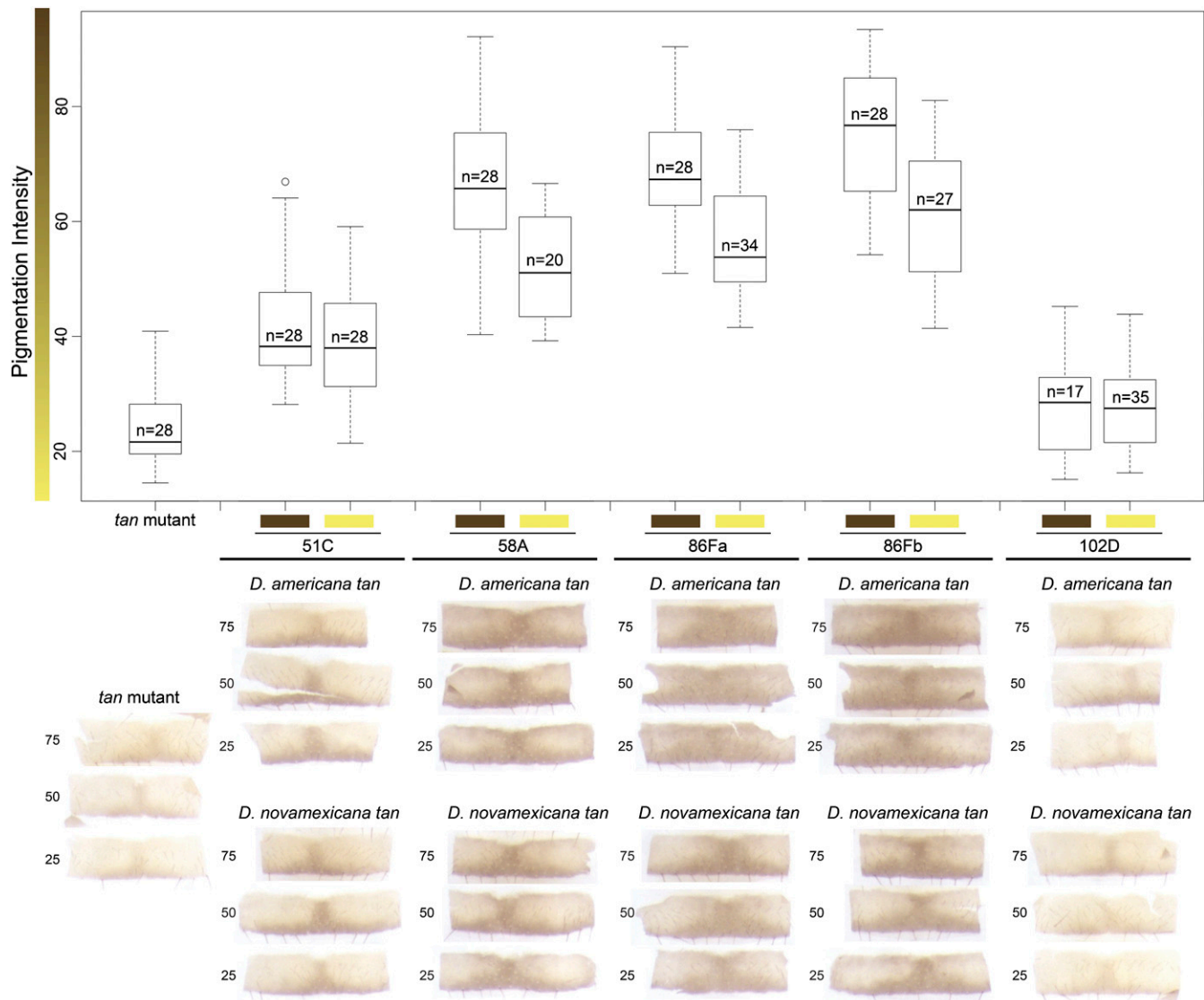


Figure 3 Genomic location can impact the relative difference in pigmentation caused by *D. americana* and *D. novamexicana* *tan* alleles. Box plots of pigmentation intensity summarize, for each genotype, the range of pigmentation phenotypes observed. The median (center line), first quartile (bottom of box), third quartile (top of box), and $\pm 1.5 \times$ the interquartile range (whiskers) are shown for each genotype examined. Yellow boxes along the x-axis represent the *D. novamexicana* allele and brown boxes represent the *D. americana* allele. Significant increases in pigmentation from the control were detected for all genomic locations except 102D (Table 2). Three of the other four landing sites (58A, 86Fa, and 86Fb) showed significant differences in pigmentation driven by the *D. americana* and *D. novamexicana* *tan* alleles, whereas the fourth landing site (51C) did not show a detectable difference in pigmentation between flies carrying the two species' alleles (Table 2). Representative images from the 25th percentile (first quartile), median, and 75th percentile (third quartile) are shown below the box plot for each genotype. The most striking differences between alleles are seen in the anterior regions outside the dorsal midline stripe.

To determine whether the insertion site of the *D. americana* and *D. novamexicana* *tan* transgenes affected their relative activity, we used an analysis of variance (ANOVA) to test for significant effects on pigmentation of allelic identity of the *tan* transgene (*D. americana* or *D. novamexicana*), genomic location of the landing site, and the interaction between the two. All three factors were found to be statistically significant predictors of pigmentation intensity (Table 1). In other words, pigmentation differences were detected between alleles and among landing sites, and the difference between alleles differed among landing sites. The significance of this interaction term is particularly interesting because it suggests that the effects of genomic context might differ between alleles, implying that the landing site used to compare the

function of *D. americana* and *D. novamexicana* alleles might alter the conclusions drawn about differences (or lack thereof) between these two alleles.

One way that the genomic context can affect a transgene is to simply silence it. To determine whether such silencing was contributing to the difference in allelic differences observed among insertion sites, we used *t*-tests to determine whether each transgene caused a statistically significant darkening of pigmentation in each transgenic line relative to the *D. melanogaster tan* mutant phenotype. Such a darkening would indicate that the transgene carried was being expressed at a level sufficient to restore at least some dark pigmentation in *D. melanogaster*. We found that the transgenic *tan* alleles from both *D. americana* and

■ **Table 4** Differences in pigmentation between flies carrying the *D. americana* and *D. novamexicana* *tan* transgenes are similar among the three landing sites that showed a significant difference between transgenes

| Source of Variation | Degrees of Freedom | Sum of Squares | Mean Square | F Value (F_0) | p-Value |
|----------------------------------|--------------------|--------------------|--------------------|--------------------|------------------------|
| 1) <i>tan</i> transgene identity | 2 | 4.58×10^4 | 2.29×10^4 | 2.13×10^2 | 6.11×10^{-49} |
| 2) Landing site | 2 | 1.85×10^3 | 9.27×10^2 | 8.66 | 2.52×10^{-4} |
| 3) Interaction between 1 and 2 | 2 | 27.8 | 13.9 | 0.13 | 0.88 |
| 4) Residuals | 186 | 1.88×10^4 | 1.07×10^2 | N/A | N/A |

Results from analysis of variance (ANOVA) after excluding flies with transgenes inserted at landing site 102D and 51C are shown. N/A, not applicable.

D. novamexicana failed to significantly alter pigmentation of the *D. melanogaster tan* mutant when inserted into the fourth chromosome at cytological position 102D (Table 2). This evidence of transgene silencing is consistent with prior studies showing that the fourth chromosome of *D. melanogaster* is highly heterochromatic (Riddle and Elgin 2006; Riddle *et al.* 2009) and can suppress the expression of transgenes (Salzler *et al.* 2013; Sun *et al.* 2000). Landing site 102D does not always silence transgenes, however; other transgenes inserted into the 102D landing site have been shown to be expressed during larval stages (Bischof *et al.* 2007; Barolo and Evans, personal communication). At each of the other four landing sites tested (all located on chromosome 2 or chromosome 3), both the *D. americana* and *D. novamexicana tan* transgenes caused a significant darkening of pigmentation relative to the *tan* mutant phenotype (Table 2), indicating that the transgenes were expressed and producing functional Tan protein. To determine whether the silencing of transgenes at landing site 102D was sufficient to explain the significant interaction observed between transgene identity and landing site in the initial ANOVA, we excluded flies with transgenes inserted into this site and repeated this ANOVA. We found that the two main effect terms (transgene identity and landing site) and the interaction term remained statistically significant (Table 3), indicating that the relative activity of the *D. americana* and *D. novamexicana tan* transgenes differed even among sites that allowed expression of both transgenes.

To further investigate this difference in relative transgene activity among insertion sites, we used a series of *t*-tests to compare the pigmentation phenotype caused by the *D. americana* and *D. novamexicana tan* alleles inserted at the same landing site. We found that the *D. americana tan* allele increased dark pigmentation of the *D. melanogaster tan* mutant significantly more than the *D. novamexicana tan* allele when inserted at three (58A, 86Fa, and 86Fb) of the four landing sites expressing the transgenes (Figure 3 and Table 2). The difference in activity between these two alleles was masked, however, when the transgenes were inserted into the landing site at 51C ($P = 0.411$, Figure 3 and Table 2). Excluding flies with transgenes at this landing site (51C) as well as flies with transgenes at the landing site that silenced the transgenes (102D) from the ANOVA described above resulted in a nonsignificant interaction between transgene allele and insertion site (Table 4), indicating that the relative effects of the *D. americana* and *D. novamexicana tan* transgenes on pigmentation were comparable at the 58A, 86Fa, and 86Fb landing sites.

Prior work has shown that position effects often alter expression levels of transgenes (*e.g.*, Markstein *et al.* 2008; Namciu *et al.* 1998; Wilson *et al.* 1990), thus we hypothesized that the different pigmentation phenotypes resulting from different insertion sites of the transgenes might be caused by differences in transgene expression among landing sites. To test this hypothesis, we used pyrosequencing to measure the relative expression of the *D. americana tan* transgene among landing sites (Figure 4). Genomic locations (58A, 86Fa, and 86Fb) that showed statistically significant differences in pigmentation caused by

the *D. americana* and *D. novamexicana tan* alleles had the highest levels of *D. americana tan* expression. The genomic location (51C) in which the *D. americana* and *D. novamexicana tan* alleles showed a significant increase in pigmentation relative to the *D. melanogaster tan* mutant, but no differences in pigmentation between flies carrying the *D. americana* and *D. novamexicana tan* alleles, had a lower level of *D. americana tan* expression. The genomic location (102D), in which neither the *D. americana* nor the *D. novamexicana tan* transgene increased pigmentation significantly relative to the *D. melanogaster tan* mutant, showed the lowest expression of *D. americana tan* among all five lines. These results confirm that different landing sites resulted in different levels of transgene expression and suggest that a threshold in transgene expression level must be reached before the different activities of the *D. americana* and *D. novamexicana tan* transgenes can be detected. We expect that this will be generally true when comparing activities of divergent alleles inserted into the same genomic location, but that the value of this threshold will likely differ depending on the strength of regulatory sequences in the transgenes, genomic context, and/or the relative difference in activity between alleles.

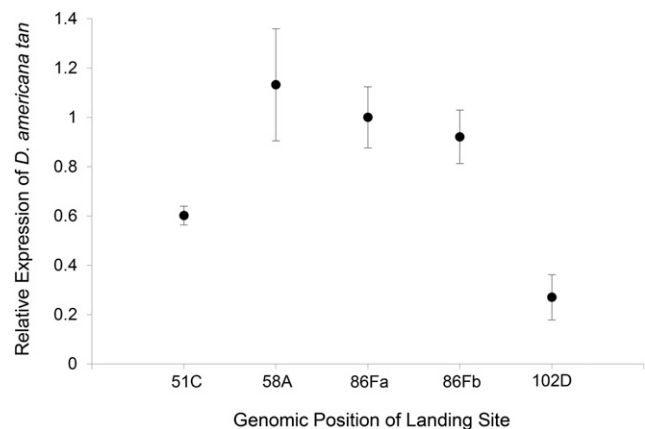


Figure 4 Genomic location impacts relative expression of the *D. americana tan* transgene in *D. melanogaster*. Expression of the *D. americana tan* transgene inserted at each of the five landing sites tested is shown relative to its expression when inserted in the 86Fa landing site. Circles indicate mean expression among replicate samples and the error bars show the 95% C.I. of the estimates. Note that the relative expression level of *D. americana tan* among landing sites correlates with the ability to detect differences in abdominal pigmentation (Figure 3). The *D. americana tan* transgene inserted at 58A, 86Fa, and 86Fb all showed similar expression as well as similar pigmentation phenotypes. The *D. americana tan* transgene inserted at 51C had a level of expression between these lines and the line with the transgene inserted at 102D, as well as pigmentation that was intermediate between these lines and 102D. The *D. americana tan* transgene inserted at 102D had the lowest transgene expression and failed to increase dark pigmentation relative to the *tan* mutant phenotype.

In summary, by comparing activities of divergent alleles of the same gene at five different genetic locations, we were able to test the assumption that position effects can be ignored as long as the two alleles compared are inserted into the same genomic location and the transgenes are expressed. We found this not to be true; *D. americana* and *D. novamexicana tan* transgenes inserted at landing site (51C) increased dark pigmentation relative to a *D. melanogaster tan* mutant, yet showed no significant difference in their relative activity. If we had only compared the effects of these *tan* alleles at the 51C landing site, we would have concluded that they had conserved functions. The lower level of transgene expression at this site relative to transgenes inserted at the three landing sites that allowed a functional difference between the *D. americana* and *D. novamexicana tan* alleles to be detected suggests that landing sites allowing the highest levels of transgene expression might provide the most power for detecting differences between alleles. We recommend that at least three genomic locations should be tested to search for allelic differences in activity. Although this increased production of transgenic lines would increase cost and workload, they would help prevent inaccurate conclusions from being drawn from transgenes affected by position effects.

ACKNOWLEDGMENTS

We would like to thank Scott Barolo and members of the Wittkopp lab for helpful discussions as well as Robert Dikeman for experimental assistance. This work was funded by the National Institutes of Health (R01-GM-089736).

LITERATURE CITED

- Bateman, J. R., A. M. Lee, and C. T. Wu, 2006 Site-specific transformation of *Drosophila* via phiC31 integrase-mediated cassette exchange. *Genetics* 173(2): 769–777.
- Benjamini, Y., and Y. Hochberg, 1995 Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. B* 57(1): 289–300.
- Bischof, J., R. K. Maeda, M. Hediger, F. Karch, and K. Basler, 2007 An optimized transgenesis system for *Drosophila* using germ-line-specific phiC31 integrases. *Proc. Natl. Acad. Sci. USA* 104(9): 3312–3317.
- Caletka, B. C., and B. F. McAllister, 2004 A genealogical view of chromosomal evolution and species delimitation in the *Drosophila virilis* species subgroup. *Mol. Phylogenet. Evol.* 33(3): 664–670.
- Cande, J., Y. Goltsev, and M. S. Levine, 2009 Conservation of enhancer location in divergent insects. *Proc. Natl. Acad. Sci. USA* 106(34): 14414–14419.
- Cooley, A. M., L. Shefner, W. N. McLaughlin, E. E. Stewart, and P. J. Wittkopp, 2012 The ontogeny of color: developmental origins of divergent pigmentation in *Drosophila americana* and *D. novamexicana*. *Evol. Dev.* 14(4): 317–325.
- Duncan, D., P. Kiefel, and I. Duncan, 2010 Control of the spineless antennal enhancer: direct repression of antennal target genes by Antennapedia. *Dev. Biol.* 347(1): 82–91.
- Engels, W. R., 1996 P elements in *Drosophila*. *Curr. Top. Microbiol. Immunol.* 204: 103–123.
- Frankel, N., G. K. Davis, D. Vargas, S. Wang, F. Payre *et al.*, 2010 Phenotypic robustness conferred by apparently redundant transcriptional enhancers. *Nature* 466(7305): 490–493.
- Garza, D., M. Medhora, A. Koga, and D. L. Hartl, 1991 Introduction of the transposable element mariner into the germline of *Drosophila melanogaster*. *Genetics* 128(2): 303–310.
- Gdula, D. A., T. I. Gerasimova, and V. G. Corces, 1996 Genetic and molecular analysis of the gypsy chromatin insulator of *Drosophila*. *Proc. Natl. Acad. Sci. USA* 93(18): 9378–9383.
- Groth, A. C., M. Fish, R. Nusse, and M. P. Calos, 2004 Construction of transgenic *Drosophila* by using the site-specific integrase from phage phiC31. *Genetics* 166(4): 1775–1782.
- Guimond, N., D. K. Bideshi, A. C. Pinkerton, P. W. Atkinson, and D. A. O’Brochta, 2003 Patterns of Hermes transposition in *Drosophila melanogaster*. *Mol. Genet. Genomics* 268(6): 779–790.
- Haley, B., B. Foy, and M. Levine, 2010 Vectors and parameters that enhance the efficacy of RNAi-mediated gene disruption in transgenic *Drosophila*. *Proc. Natl. Acad. Sci. USA* 107(25): 11435–11440.
- Handler, A. M., and R. A. Harrell, 2nd, 1999 Germline transformation of *Drosophila melanogaster* with the piggyBac transposon vector. *Insect Mol. Biol.* 8(4): 449–457.
- Horn, C., and E. A. Wimmer, 2000 A versatile vector set for animal transgenesis. *Dev. Genes Evol.* 210(12): 630–637.
- Huisinga, K. L., N. C. Riddle, W. Leung, S. Shimonovich, S. McDaniel *et al.*, 2016 Targeting of P-Element Reporters to Heterochromatic Domains by Transposable Element 1360 in *Drosophila melanogaster*. *Genetics* 202(2): 565–582.
- Joshi, R., L. Sun, and R. Mann, 2010 Dissecting the functional specificities of two Hox proteins. *Genes Dev.* 24(14): 1533–1545 10.1101/gad.1936910.
- Kalay, G., and P. J. Wittkopp, 2010 Nomadic enhancers: tissue-specific cis-regulatory elements of yellow have divergent genomic positions among *Drosophila* species. *PLoS Genet.* 6(11): e1001222 10.1371/journal.pgen.1001222.
- Kuhn, E. J., and P. K. Geyer, 2003 Genomic insulators: connecting properties to mechanism. *Curr. Opin. Cell Biol.* 15(3): 259–265.
- Levis, R., T. Hazelrigg, and G. M. Rubin, 1985 Effects of genomic position on the expression of transduced copies of the white gene of *Drosophila*. *Science* 229(4713): 558–561.
- Markstein, M., C. Pitsouli, C. Villalta, S. E. Celniker, and N. Perrimon, 2008 Exploiting position effects and the gypsy retrovirus insulator to engineer precisely expressed transgenes. *Nat. Genet.* 40(4): 476–483.
- Martin, A., and V. Orgogozo, 2013 The Loci of repeated evolution: a catalog of genetic hotspots of phenotypic variation. *Evolution* 67(5): 1235–1250.
- Morales-Hojas, R., C. P. Vieira, and J. Vieira, 2008 Inferring the evolutionary history of *Drosophila americana* and *Drosophila novamexicana* using a multilocus approach and the influence of chromosomal rearrangements in single gene analyses. *Mol. Ecol.* 17(12): 2910–2926.
- Namciu, S. J., K. B. Blochlinger, and R. E. Fournier, 1998 Human matrix attachment regions insulate transgene expression from chromosomal position effects in *Drosophila melanogaster*. *Mol. Cell. Biol.* 18(4): 2382–2391.
- Perry, M. W., A. N. Boettiger, and M. Levine, 2011 Multiple enhancers ensure precision of gap gene-expression patterns in the *Drosophila* embryo. *Proc. Natl. Acad. Sci. USA* 108(33): 13570–13575.
- Rasband, W. S., 1997–2016 ImageJ. U. S. National Institutes of Health, Bethesda, Maryland. Available at: <http://imagej.nih.gov/ij/>.
- R Development Core Team, 2016 R: A language and environment for statistical computing. Vienna, Austria: The R Foundation for Statistical Computing. Available at: <https://www.R-project.org/>.
- Rebeiz, M., N. Jikomes, V. A. Kassner, and S. B. Carroll, 2011 Evolutionary origin of a novel gene expression pattern through co-option of the latent activities of existing regulatory sequences. *Proc. Natl. Acad. Sci. USA* 108(25): 10036–10043.
- Riddle, N. C., and S. C. R. Elgin, 2006 The dot chromosome of *Drosophila*: Insights into chromatin states and their change over evolutionary time. *Chromosome Res.* 14(4): 405–416.
- Riddle, N. C., C. D. Shaffer, and S. C. R. Elgin, 2009 A lot about a little dot - lessons learned from *Drosophila melanogaster* chromosome 4. *Biochemistry and Cell Biology-Biochimie Et Biologie Cellulaire* 87(1): 229–241.
- Salzler, H. R., D. C. Tatomer, P. Y. Malek, S. L. McDaniel, A. N. Orlando *et al.*, 2013 A sequence in the *Drosophila* H3–H4 Promoter triggers histone locus body assembly and biosynthesis of replication-coupled histone mRNAs. *Dev. Cell* 24(6): 623–634.
- Sayal, R., S. M. Ryu, and D. N. Arnosti, 2011 Optimization of reporter gene architecture for quantitative measurements of gene expression in the *Drosophila* embryo. *Fly (Austin)* 5(1): 47–52.
- Silicheva, M., A. Golovnin, E. Pomerantseva, A. Parshikov, P. Georgiev *et al.*, 2010 *Drosophila* mini-white model system: new insights into positive

- position effects and the role of transcriptional terminators and gypsy insulator in transgene shielding. *Nucleic Acids Res.* 38(1): 39–47.
- Smith, D., J. Wohlgemuth, B. R. Calvi, I. Franklin, and W. M. Gelbart, 1993 hobo enhancer trapping mutagenesis in *Drosophila* reveals an insertion specificity different from P elements. *Genetics* 135(4): 1063–1076.
- Spradling, A. C., and G. M. Rubin, 1983 The effect of chromosomal position on the expression of the *Drosophila* xanthine dehydrogenase gene. *Cell* 34(1): 47–57.
- Stern, D. L., and V. Orgogozo, 2008 The loci of evolution: how predictable is genetic evolution? *Evolution* 62(9): 2155–2177.
- Sturtevant, A. H., 1925 The Effects of Unequal Crossing over at the Bar Locus in *Drosophila*. *Genetics* 10(2): 117–147.
- Sun, F. L., M. H. Cuaycong, C. A. Craig, L. L. Wallrath, J. Locke *et al.*, 2000 The fourth chromosome of *Drosophila melanogaster*: Interspersed euchromatic and heterochromatic domains. *Proc. Natl. Acad. Sci. USA* 97(10): 5340–5345.
- Throckmorton, L. H., 1982 The *virilis* species group, pp. 227–296 in *The Genetics and Biology of Drosophila* (Vol. 3b), edited by M. Ashburner, H. L. Carson, and J. N. Thompson, Jr. Academic Press, London.
- Venken, K. J., and H. J. Bellen, 2005 Emerging technologies for gene manipulation in *Drosophila melanogaster*. *Nat. Rev. Genet.* 6(3): 167–178.
- Venken, K. J., and H. J. Bellen, 2007 Transgenesis upgrades for *Drosophila melanogaster*. *Development* 134(20): 3571–3584.
- Venken, K. J., Y. He, R. A. Hoskins, and H. J. Bellen, 2006 P[acman]: a BAC transgenic platform for targeted insertion of large DNA fragments in *D. melanogaster*. *Science* 314(5806): 1747–1751.
- Wilson, C., H. J. Bellen, and W. J. Gehring, 1990 Position effects on eukaryotic gene expression. *Annu. Rev. Cell Biol.* 6: 679–714.
- Wimmer, E. A., 2005 Insect transgenesis by site-specific recombination. *Nat. Methods* 2(8): 580–582.
- Wittkopp, P. J., 2006 Evolution of cis-regulatory sequence and function in Diptera. *Heredity (Edinb)* 97(3): 139–147.
- Wittkopp, P. J., 2011 Using pyrosequencing to measure allele-specific mRNA abundance and infer the effects of cis- and trans-regulatory differences. *Methods Mol. Biol.* 772: 297–317 10.1007/978-1-61779-228-1_18.
- Wittkopp, P. J., E. E. Stewart, L. L. Arnold, A. H. Neidert, B. K. Haerum *et al.*, 2009 Intraspecific polymorphism to interspecific divergence: genetics of pigmentation in *Drosophila*. *Science* 326(5952): 540–544.

Communicating editor: H. K. Salz