Science
AAAS

Intraspecific Polymorphism to Interspecific Divergence: Genetics of Pigmentation in Drosophila

Patricia J. Wittkopp, *et al.*

Science **326**, 540 (2009);
DOI: 10.1126/science.1176980

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published weekly, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005.

Intraspecific Polymorphism to Interspecific Divergence: Genetics of Pigmentation in *Drosophila*

Patricia J. Wittkopp,[1,2]* Emma E. Stewart,[1] Lisa L. Arnold,[2] Adam H. Neidert,[1] Belinda K. Haerum,[1] Elizabeth M. Thompson,[2] Saleh Akhras,[1] Gabriel Smith-Winberry,[1] Laura Shefner[1]

Genetic changes contributing to phenotypic differences within or between species have been identified for a handful of traits, but the relationship between alleles underlying intraspecific polymorphism and interspecific divergence is largely unknown. We found that noncoding changes in the *tan* gene, as well as changes linked to the *ebony* gene, contribute to pigmentation divergence between closely related *Drosophila* species. Moreover, we found that alleles linked to *tan* and *ebony* fixed in one *Drosophila* species also contribute to variation within another species, and that multiple genotypes underlie similar phenotypes even within the same population. These alleles appear to predate speciation, which suggests that standing genetic variation present in the common ancestor gave rise to both intraspecific polymorphism and interspecific divergence.

Similar phenotypes that vary within and between species may or may not be caused by similar genetic mechanisms. Quantitative trait mapping shows that loci contributing to polymorphism and divergence of a single character map to the same region of the genome approximately half of the time (table S1). These overlapping quantitative trait loci (QTLs) may or may not result from changes in the same genes, and most studies lack the power to distinguish between these possibilities. To determine whether the same genes (and potentially even the same alleles of these genes) contribute to phenotypic diversity within and between species, one must resolve intra- and interspecific QTLs to individual genes, localize functionally divergent sites within these genes, and then compare specific alleles within and between species.

***ebony* and *tan* QTLs contribute to pigmentation divergence.** To investigate the relationship between intraspecific polymorphism and interspecific divergence, we examined the genetic basis of pigmentation differences within and between a pair of closely related *Drosophila* species, *D. americana* and *D. novamexicana*. These two species are sister taxa within the *Drosophila* virilis species group that diverged about 300,000 to 500,000 years ago (*1*, *2*) (Fig. 1A). *D. novamexicana* has a derived light yellow body color, whereas other members of this group (including *D. americana*) retain an ancestral dark brown body color (*3*) (Fig. 1B). In the laboratory, these species can mate and produce fertile offspring. Genetic mapping showed that a region of the second chromosome containing the *ebony* gene contributes to pigmentation divergence between *D. novamexicana* and *D. americana* (*4*). This gene is required for pigmentation in *D. melanogaster* (*5*). Three other autosomal regions, as well as an unidentified region of the X chromosome, also contribute to pigmentation divergence, although none of these regions were linked to other pigmentation genes tested (i.e., *yellow*, *dopa-decarboxylase*, *optomotor blind*, and *bric-a-brac*).

Recently, the X-linked pigmentation gene *tan* was cloned in *D. melanogaster* (*6*). To test whether this gene might contribute to pigmentation differences between *D. americana* and *D. novamexicana*, we crossed *D. americana* females to *D. novamexicana* males, backcrossed F_1 hybrid females to *D. novamexicana* males, and scored 495 backcross progeny for body color (fig. S1). All of the lightest male offspring (n = 10) inherited the *D. novamexicana* allele of *tan*, whereas all of the darkest male offspring (n = 24) inherited the *D. americana* allele of this marker. These data show that sequences linked to *tan* contribute to pigmentation divergence ($P = 8 \times 10^{-9}$; Fisher's exact test). The previously described pigmentation QTL linked to *ebony* and the lack of a pigmentation QTL linked to *yellow* (*4*) were also reconfirmed in this population ($P = 3 \times 10^{-8}$ and 0.7, respectively; Fisher's exact test).

To determine the phenotypic effects of QTLs linked to *ebony* and *tan*, we created lines of *D. novamexicana* in which genomic regions containing these genes were replaced with orthologous sequences from *D. americana*. These genotypes were constructed by marker-assisted introgression, moving *ebony* and *tan* alleles from *D. americana* into *D. novamexicana*. F_1 hybrid females were backcrossed to *D. novamexicana* males, and a single female inheriting the *D. americana tan* (or *ebony*) allele was randomly selected and backcrossed to *D. novamexicana* males again. This process was repeated for 10 generations (fig. S2), with females carrying the *D. americana ebony* or *tan* allele selected randomly in each generation without regard to pigmentation. Introgressed *D. americana* sequences linked to either *tan* or *ebony* darkened pigmentation relative to wild-type *D. novamexicana* (Fig. 2, A to C), with sequences linked to *ebony* (Fig. 2C) causing darker pigmentation than sequences linked to *tan* (Fig. 2B). Digital quantification of pigmentation showed that, when combined, the introgressed *tan* and *ebony* regions recapitulated 87% of the pigmentation difference between species (Fig. 2, A, D, and E).

***ebony* and *tan* affect pigmentation development.** Studies of pigmentation in *D. melanogaster* suggest that *ebony* and *tan* may themselves be responsible for these interspecific QTLs. Loss-of-function mutations in *ebony* darken pigmentation (*5*), whereas loss-of-function mutations in *tan* lighten it (*7*). Biochemically, Ebony catalyzes the conversion of dopamine into *N*-β-alanyldopamine (NBAD), which is a precursor for (yellow) sclerotin, and Tan catalyzes the reverse reaction, converting NBAD back into dopamine, which is a precursor for (brown) melanin [reviewed in (*8*)] (Fig. 3A). Ectopic expression of Ebony induces yellow pigmentation (*9*) (Fig. 3D), whereas ectopic expression of Tan induces brown pigmentation (*6*) (Fig. 3E). Ectopic ex-

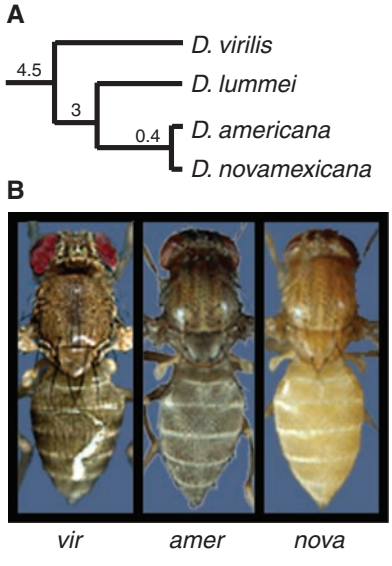


Fig. 1. *D. novamexicana* yellow body color is derived. (**A**) Phylogenetic relationships among members of the virilis phylad within the virilis group of *Drosophila* are shown with estimated divergence times (*1*) at each node (numbers denote millions of years ago). (**B**) Dorsal body pigmentation is shown for *D. virilis* (*vir*), *D. americana* (*amer*), and *D. novamexicana* (*nova*). *D. lummei* (not shown) has pigmentation similar to that of *D. virilis* and *D. americana* (*3*).

[1]Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, MI 48109, USA. [2]Department of Molecular, Cellular, and Developmental Biology, University of Michigan, Ann Arbor, MI 48109, USA.

*To whom correspondence should be addressed. E-mail: wittkopp@umich.edu

pression of both proteins simultaneously results in pigmentation intermediate to that caused by ectopic expression of either protein alone (Fig. 3F), showing that the balance between Ebony and Tan enzymatic activity affects pigmentation. Genetic and biochemical pathways controlling pigment synthesis are highly conserved among insects (*10*), which suggests that the *D. americana* and *D. novamexicana tan* and *ebony* genes function similarly to their *D. melanogaster* orthologs. Consistent with this prediction, the Ebony protein is more abundant in epidermal cells of the yellowish *D. novamexicana* during late pupal stages than in the darker *D. americana* (*4*).

Noncoding changes in *tan* contribute to pigmentation divergence. The above results are consistent with changes in *ebony* and *tan* contributing to pigmentation divergence, but they cannot be used to distinguish divergence affecting these genes from divergence affecting linked loci. This is particularly concerning for *ebony* because it is located in a part of the genome that is inverted between species (*4*, *11*). Inversions effectively suppress recombination, precluding genetic dissection of the region. Nonetheless, differences in Ebony protein expression between *D. americana* and *D. novamexicana* (*4*) strongly suggest that this gene is involved in pigmentation divergence.

Unlike *ebony*, *tan* is in a freely recombining region of the genome. This allowed us to use fine-scale genetic mapping to separate the effects of *tan* from neighboring genes and to determine whether *tan* contributes to the altered pigmentation observed in the *tan* introgression line (Fig. 2B). A 2.7-kb region of *tan* was identified that contributes to pigmentation divergence (fig. S3) and contains 57 single-nucleotide differences and 19 insertions or deletions (indels) (fig. S4). All of these changes affect noncoding sequences, and the region includes the entire first intron (fig. S3). Differences located 3′ of this region must also affect pigmentation, however, because the recombinant fly inheriting *D. americana tan* sequence only in this region was not as dark as flies inheriting *D. americana* sequence for the full *tan* gene (fig. S3). Within *tan*, this 3′ region includes many noncoding differences as well as two nonsynonymous differences that affect amino acids 190 and 267.

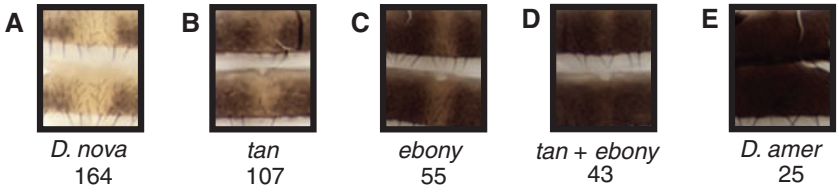


Fig. 2. QTLs linked to *tan* and *ebony* account for the majority of pigmentation divergence between species. Dorsal abdominal cuticle is shown from segments A4 and A5 of 7- to 10-day-old adult females. (**A** to **C**) Relative to *D. novamexicana* (A), introgression of alleles linked to *tan* (B) or *ebony* (C) led to darkened pigmentation. (**D** and **E**) Together, the introgressed regions produced even darker pigmentation (D), although these flies were not as dark as wild-type *D. americana* (E). Numbers indicate intensity of grayscale images, where 0 = black and 255 = white. Panels (B), (C), and (D) are all heterozygous for the introgressed region(s).

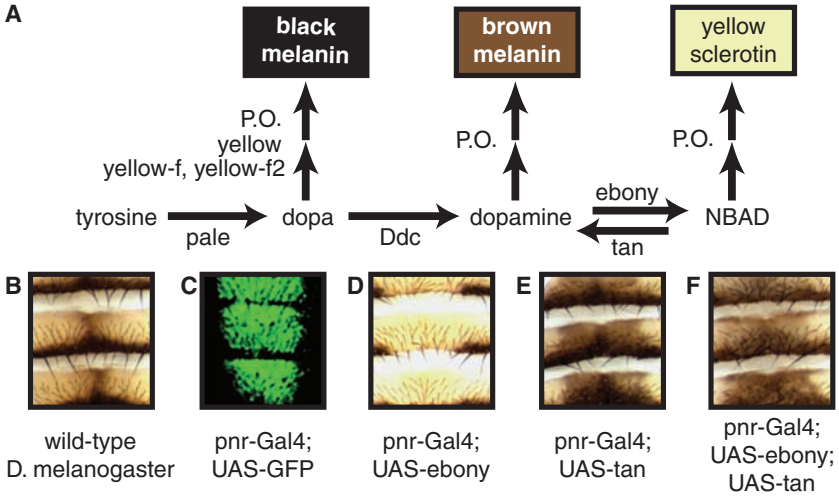


Fig. 3. Ebony and Tan have reciprocal effects on pigmentation development. (**A**) A simplified melanin biosynthesis pathway is shown. For a more complete pathway, see (*6*). The gene(s) controlling each enzymatic step are shown in italics. *P.O.* indicates genes encoding phenol oxidase proteins. Branches with two consecutive arrows include multiple enzymatic steps that are not well defined. (**B** to **F**) Dorsal abdominal cuticle (segments A3 to A5) is shown for various *D. melanogaster* genotypes. (B) Canton-S, a wild-type strain of *D. melanogaster*, shows the striped dorsal abdominal pigment pattern typical of this species. (C) Expression of UAS-GFP (green) shows that *pnr*-Gal4, the driver used to ectopically express Ebony and Tan in (D) to (F), activates gene expression in a stripe along the dorsal midline during late pupal development. (D) Ectopic expression of UAS-Ebony caused increased yellow pigmentation. (E) Ectopic expression of UAS-Tan caused increased brown pigmentation. (F) Simultaneous expression of both UAS-Tan and UAS-Ebony resulted in an intermediate phenotype. Cuticle is from 3- to 5-day-old females in all panels except (C), in which cuticle is from a female pupa just before eclosion (stage P15).

***tan* expression correlates with pigmentation differences.** Given the absence of coding changes in the 2.7-kb mapped region of *tan*, we expect that divergent sites in this region affect pigmentation by altering *tan* expression. Because of its darker pigmentation, we hypothesized that *D. americana* has higher levels of *tan* expression than *D. novamexicana*. In situ hybridization showed that *tan* is expressed throughout each dorsal abdominal segment ("tergite") in both species during the P14 and P15 pupal stages (*12*) when pigmentation develops (fig. S5, A and B). This expression pattern correlates with the distribution of pigments in adult *D. americana* and *D. novamexicana* tergites, and is distinct from the patterns of *tan* expression in *Drosophila* species with other pigment patterns (*13*). Differences in *tan* expression detected with in situ hybridization correlate with pigmentation divergence in these other species (*13*), yet we saw no obvious expression differences between *D. americana* and *D. novamexicana* during the same developmental stages with this technique (fig. S5, A and B).

To quantitatively compare levels of *tan* expression, we measured the relative abundance of *tan* transcripts in stage P14 and P15 pupae of each species with Pyrosequencing (*14*). We observed an average of 34% more *tan* transcripts in *D. americana* females than in *D. novamexicana* females ($n = 4$ samples, each containing six flies; $t = 3.7$, $P = 0.03$; t test) (fig. S5, C and D), consistent with the darker pigmentation of *D. americana*. To determine whether this expression difference results from cis-regulatory divergence of *tan*, we compared transcript abundance of *D. americana* and *D. novamexicana tan* alleles in F_1 hybrid females during the same pupal stages with Pyrosequencing (*14*). Surprisingly, no significant differences in allele-specific expression were observed ($n = 5$ samples, each containing six flies; $t = 0.72$, $P = 0.51$; t test) (fig. S5, E and F). Divergent expression levels may therefore be caused by differences in trans-regulatory factors and/or differences in the number of *tan*-expressing cells between species (*15*). The noncoding differences we identified by fine-scale genetic mapping may alter fine-scale temporal control (*16*) and/or posttranscriptional regulation (*17*) of *tan*. It is also possible that these noncoding differences may affect transcriptional regulation of a neighboring gene that is also

involved in pigmentation (*13*) or a cryptic, small, noncoding RNA encoded by the *tan* intron (*18*).

Phenotypic consequences of *tan* divergence revealed in transgenic flies. To determine whether evolutionary changes in the *tan* gene itself are sufficient to affect pigmentation, we inserted transgenes carrying the *D. americana* and *D. novamexicana tan* alleles into the *D. melanogaster* genome (*19*). Both transgenes were integrated at the same site, allowing us to compare pigmentation of flies whose genomes differed only for divergent sites within the transgenes. Both the *D. americana* and *D. novamexicana* alleles of *tan* rescued pigmentation in a *D. melanogaster tan* null mutant (Fig. 4, A to D), indicating that the transgenes were expressed in *D. melanogaster* and that Tan protein function is (at least largely) conserved. Flies carrying the *D. americana tan* allele had darker pigmentation than flies carrying the *D. novamexicana tan* allele (Fig. 4, C and D; $F = 26.94$, $P < 0.0001$ for abdominal segments A3 and A4, and $F = 6.51$, $P = 0.03$ for the darker A5 segment). This is consistent with the darker pigmentation of *D. americana* relative to *D. novamexicana*.

We also compared the phenotypic effects of *D. americana* and *D. novamexicana tan* alleles in *D. americana* and *D. novamexicana* themselves by randomly inserting both *tan* transgenes into the genomes of both species. Two independent insertions were recovered for each transgene in each species. In *D. americana*, we were unable to detect a difference in pigmentation between transformed and untransformed flies, presumably because of the already dark pigmentation of this species (see Fig. 1B). In *D. novamexicana*, however, transformant flies carrying the *D. americana tan* transgene (Fig. 4F) were visibly darker than flies carrying the *D. novamexicana tan* transgene (Fig. 4E).

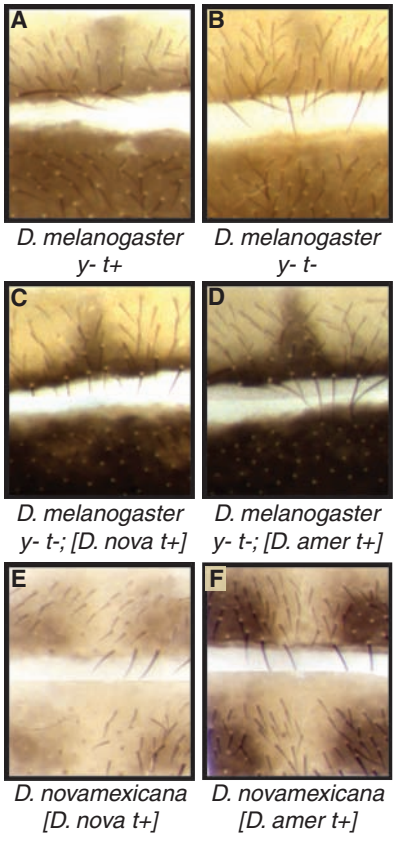


Fig. 4. The *D. americana* allele of *tan* causes darker pigmentation than the *D. novamexicana* allele of *tan*. Transgenes containing *tan* alleles from *D. americana* and *D. novamexicana* were transformed into *D. melanogaster*, *D. novamexicana*, and *D. americana*. In *D. melanogaster*, transgenes were crossed into a genetic background homozygous for null mutations in *yellow* (*y-*) and *tan* (*t-*). The *yellow* mutation was used to lighten pigmentation, making the effects of *tan* transgenes easier to see. (**A**) *D. melanogaster yellow* (*y-*) mutant, which is wild-type for *tan*. (**B**) *D. melanogaster yellow, tan* (*y-, t-*) double mutant. (**C**) *D. melanogaster yellow, tan* mutant carrying the *D. novamexicana tan* transgene (*D. nova t+*). (**D**) *D. melanogaster yellow, tan* mutant carrying the *D. americana tan* transgene (*D. amer t+*). (**E**) Wild-type *D. novamexicana* carrying the *D. novamexicana tan* transgene. (**F**) Wild-type *D. novamexicana* carrying the *D. americana tan* transgene.

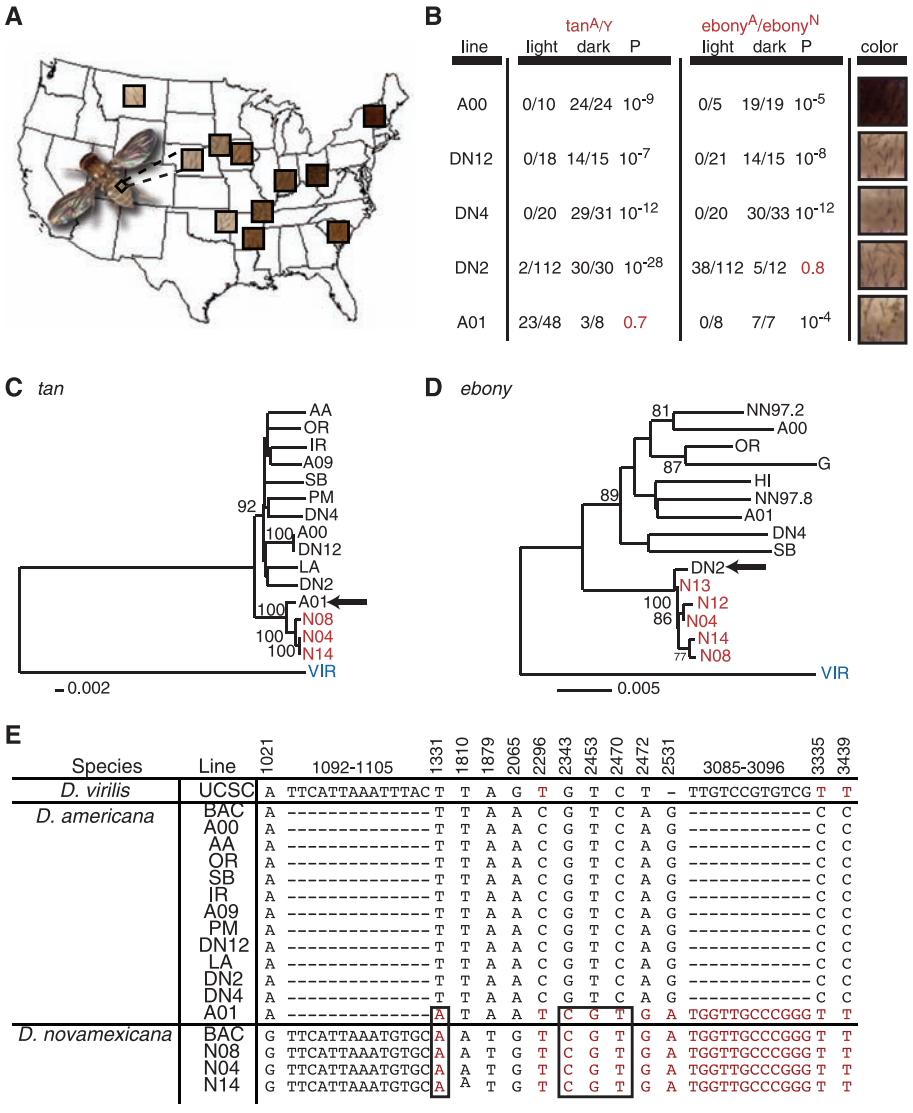


Fig. 5. *ebony* and *tan* QTLs also contribute to polymorphism. (**A**) Dorsal abdominal cuticle from *D. americana* isofemale lines (table S2) is shown. Eastern populations are darker than western populations. (**B**) Phenotypes, genotypes, and statistical significance for interspecific QTL mapping experiments. Dorsal abdominal pigmentation of each of isofemale line is shown in the color column; the middle columns show the proportion of male backcross progeny genotyped from the lightest (light) and darkest (dark) pigmentation classes with *D. americana tan* (tan^A) and *ebony* ($ebony^A$) alleles. *P* values are from 2×2 Fisher's exact tests of the genotype count data. (**C** and **D**) Neighbor-joining trees of *tan* [(C), 7625 base pairs] and *ebony* [(D), 1136 base pairs] from *D. americana* (black), *D. novamexicana* (red), and *D. virilis* (blue) are shown with bootstrap values >75% ($n = 1000$). Branch lengths are to scale. (**E**) Fixed differences within the 2.7-kb candidate region of *tan*. Sites 889 to 981 and 3521 to 3616 are exons 1 and 2, respectively. The *D. virilis* allele is from the 2005 assembly of the *D. virilis* genome sequence (*29*). Positions refer to alignment of GQ457336 through GQ457353. Alleles shared between *D. novamexicana* and A01 are red; the derived subset, relative to *D. virilis*, is boxed.

***ebony* and *tan* QTLs underlie variable pigmentation within *D. americana*.** Pigmentation of *D. americana* is always distinct from that of *D. novamexicana* (*3*), but the intensity of dark pigmentation varies within *D. americana*. This variation is geographically structured, with *D. americana* captured in the eastern United States visibly darker than those captured from the western part of the species range (*3*). These pigmentation differences remain visible after rearing flies under common environmental conditions in the laboratory (Fig. 5A), indicating a genetic basis for the pigmentation cline. Sequence variation within *D. americana* at putatively neutral loci shows no population structure (*2*, *20*–*24*), which suggests that this cline is due to local adaptation.

To determine whether sites linked to *ebony* and/or *tan* contribute to this intraspecific polymorphism, we used *D. novamexicana* alleles as a reference to compare the phenotypic effects of *ebony* and *tan* QTL alleles among lines of *D. americana*. Genetic mapping was performed using four isofemale lines of *D. americana* (A01 collected from Poplar, Montana; DN2, DN4, and DN12 collected from Duncan, Nebraska) with lighter pigmentation than the line of *D. americana* (A00) used previously. *D. americana* females from each strain were crossed to *D. novamexicana* males, F_1 hybrid females were backcrossed to *D. novamexicana* males, and molecular markers in *ebony* and *tan* were genotyped in flies from the lightest and darkest pigmentation classes (fig. S6). Despite their lighter pigmentation, the DN12 and DN4 lines of *D. americana* produced similar mapping results to A00: Both *ebony* and *tan* showed highly significant linkages to loci affecting pigmentation (Fig. 5B). Genotyping 101 males from the DN4 backcross population and fitting their genotypes and phenotypes to a linear model (*19*) showed that *ebony* (E) and *tan* (T) both had significant additive effects on pigmentation ($F_E = 132.98$, $P_E < 0.0001$; $F_T = 160.85$, $P_T < 0.0001$) with no significant epistatic interaction between them ($F = 3.02$, $P = 0.09$). These additive effects explained 76% of the pigmentation variance in the DN4 backcross population. For DN2, sites linked to *tan* contributed to pigmentation differences between species, but sites linked to *ebony* did not (Fig. 5B). The converse was true for A01: Sites linked to *ebony* contributed to pigmentation differences between species, whereas sites linked to *tan* did not (Fig. 5B).

Taken together, these data reveal three distinct genotypes among *D. americana* lines with a light-pigmentation phenotype. DN2 has alleles linked to *ebony* that appear to be functionally equivalent to those found in *D. novamexicana*; A01 has alleles linked to *tan* that appear to be functionally equivalent to those found in *D. novamexicana*; and DN4 and DN12 have alleles linked to *tan* and *ebony* that appear to be functionally distinct from those found in *D. novamexicana*. It remains to be seen whether the DN12 and DN4 alleles of these QTLs have the same effect on pigmentation as each other or as the alleles from the darker A00 line of *D. americana*. Two of the three *D. americana* lines (DN2 and DN4) were collected during the same year and DN12 the following year (table S2), which suggests that genetic heterogeneity for pigmentation exists within this local population. This heterogeneity may be caused by gene flow among populations and/or balancing selection within the population.

Shared pigmentation alleles contribute to polymorphism and divergence. The functional similarity observed for the A01 and *D. novamexicana* alleles linked to *tan*, as well as for the DN2 and *D. novamexicana* alleles linked to *ebony*, may result from shared ancestry (i.e., alleles that are identical by descent) or from convergent evolution. Sequences of *tan* and *ebony* from multiple lines of *D. americana* and *D. novamexicana* show that the functional similarity most likely reflects shared ancestry, as the A01 *tan* sequence is more similar to *D. novamexicana* alleles than to other *D. americana* alleles (Fig. 5C) and the DN2 *ebony* sequence is more similar to *D. novamexicana* alleles than to other *D. americana* alleles (Fig. 5D). The DN2 line of *D. americana* also has the same arrangement of the *ebony*-containing inversion ["In(2)b" in (*11*)] as *D. novamexicana* (fig. S7) (*25*), further suggesting that pigmentation alleles linked to *ebony* in DN2 and *D. novamexicana* have a common origin.

Sequence variation identifies candidate sites for divergent pigmentation. Sequence similarity between A01 and *D. novamexicana* was found to be highest beginning in the first intron of *tan* and extending 3′ of *tan* (fig. S8). Within the 2.7-kb region identified by fine-scale mapping (fig. S3), we observed 13 fixed single-nucleotide differences and two fixed indels between *D. americana* (excluding A01) and *D. novamexicana* (Fig. 5E). The A01 allele of *D. americana* contains the same sequence as *D. novamexicana* at nine of these 13 divergent sites and shares one of the two indels (Fig. 5E, red). Only four of the shared substitutions are derived changes relative to *D. virilis* (Fig. 5E, boxed). Because *D. virilis* has pigmentation similar to *D. americana* (Fig. 1B), we consider these four noncoding changes to be the best candidates for divergent function in this region. Derived changes outside of this region that are also unique to A01 and *D. novamexicana tan* may contribute to pigmentation divergence as well. The two amino acid differences between alleles used for fine-scale mapping are polymorphic and are thus unlikely to contribute to fixed differences between species.

A model of pigmentation evolution. Our data reveal the relationship between intraspecific polymorphism and interspecific divergence by showing that the same alleles contribute to pigmentation differences within and between species (fig. S9). These alleles may have been present in the common ancestor of *D. americana* and *D. novamexicana*, or they may have arisen in *D. novamexicana* and subsequently introgressed into *D. americana* after hybridization. Distinguishing between these two scenarios is notoriously difficult (*26*–*28*), yet haplotype sharing between *D. americana* and *D. novamexicana* has been postulated to be due to shared ancestral variation (*2*). Our data are consistent with this interpretation: Sequences of *D. americana* alleles that appear to have the same function as *D. novamexicana* alleles are basal to these *D. novamexicana* alleles in gene trees (Fig. 5, C and D); there is evidence of recombination (or gene conversion) within the *D. americana* A01 *tan* haplotype (fig. S5A), which argues against a recent introgression event; and *D. novamexicana* is thought to have evolved from a peripheral population of the common ancestor shared with *D. americana* (*2*). Therefore, we propose that light-pigmentation alleles segregating in this common ancestor became fixed in *D. novamexicana*, contributing to its yellow body color, and continue segregating in *D. americana*, contributing to clinal variation. Additional *D. novamexicana*–like alleles of *D. americana* are needed to further evaluate this model.

References and Notes

1. B. C. Caletka, B. F. McAllister, *Mol. Phylogenet. Evol.* **33**, 664 (2004).
2. R. Morales-Hojas, C. P. Vieira, J. Vieira, *Mol. Ecol.* **17**, 2910 (2008).
3. L. H. Throckmorton, in *The Genetics and Biology of* Drosophila, M. Ashburner, H. L. Carson, J. N. Thompson, Eds. (Academic Press, London, 1982), vol. 3b, pp. 227–296.
4. P. J. Wittkopp, B. L. Williams, J. E. Selegue, S. B. Carroll, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 1808 (2003).
5. C. B. Bridges, T. H. Morgan, *Publ. Carnegie Inst. 327* (1923).
6. J. R. True *et al.*, *PLoS Genet.* **1**, e63 (2005).
7. R. McEwen, *J. Exp. Zool.* **25**, 49 (1918).
8. P. J. Wittkopp, S. B. Carroll, A. Kopp, *Trends Genet.* **19**, 495 (2003).
9. P. J. Wittkopp, J. R. True, S. B. Carroll, *Development* **129**, 1849 (2002).
10. P. J. Wittkopp, P. Beldade, *Semin. Cell Dev. Biol.* **20**, 65 (2009).
11. T. C. Hsu, thesis, University of Texas, Austin (1951).
12. M. Ashburner, Drosophila*: A Laboratory Manual* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989).
13. S. Jeong *et al.*, *Cell* **132**, 783 (2008).
14. A. Ahmadian *et al.*, *Anal. Biochem.* **280**, 103 (2000).
15. P. J. Wittkopp, B. K. Haerum, A. G. Clark, *Nature* **430**, 85 (2004).
16. B. Cong, J. P. Liu, S. D. Tanksley, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 13606 (2002).
17. C. C. Laurie, L. F. Stam, *Genetics* **138**, 379 (1994).
18. A. Stark *et al.*, *Genome Res.* **17**, 1865 (2007).
19. See supporting material on *Science* Online.
20. X. Maside, A. W. Lee, B. Charlesworth, *Curr. Biol.* **14**, 150 (2004).
21. B. F. McAllister, *Genetics* **165**, 1317 (2003).
22. B. F. McAllister, A. L. Evans, *PLoS One* **1**, e112 (2006).
23. M. A. Schafer, L. Orsini, B. F. McAllister, C. Schlotterer, *Heredity* **97**, 291 (2006).
24. C. P. Vieira, P. A. Coelho, J. Vieira, *Genetics* **164**, 1459 (2003).
25. P. Mena, B. McAllister, personal communication.
26. J. H. Degnan, N. Rosenberg, *Trends Ecol. Evol.* **24**, 332 (2009).

27. M. T. Holder, J. A. Anderson, A. K. Holloway, *Syst. Biol.* **50**, 978 (2001).
28. M. A. Noor, J. L. Feder, *Nat. Rev. Genet.* **7**, 851 (2006).
29. *Drosophila* 12 Genomes Consortium, *Nature* **450**, 203 (2007).
30. We thank G. Kalay, X. Heng, M. Weisel, A. Ratnala, and E. Larimore for technical assistance; P. Mena and B. McAllister for sharing unpublished data (including the images shown in fig. S7) and lines of *D. americana*; C. Dick for use of the Mixer Mill, M. Rebeiz for sharing an unpublished in situ protocol; N. Rosenberg, G. Coop, and other colleagues for discussions of incomplete lineage sorting and introgression; and J. Gruber, J. Coolon, A. Cooley, G. Kalay, D. Yuan, B. McAllister, A. Kopp, B. Prud'homme, and N. Gompel for comments on the manuscript. Supported by NSF grant DEB140640485, the Margaret and Herman Sokol Endowment for Faculty and Graduate Student Research Projects, and the University of Michigan. P.J.W. is an Alfred P. Sloan Research Fellow. DNA sequences are available from GenBank (accession numbers GQ457336 to GQ457453).

Supporting Online Material
www.sciencemag.org/cgi/content/full/326/5952/540/DC1
Materials and Methods
Figures S1 to S9
Tables S1 to S3
References

28 May 2009; accepted 13 August 2009
10.1126/science.1176980

RNAi in Budding Yeast

Ines A. Drinnenberg,[1,2]* David E. Weinberg,[1,2,3]* Kathleen T. Xie,[1,2,3]* Jeffrey P. Mower,[4]† Kenneth H. Wolfe,[4] Gerald R. Fink,[1,3] David P. Bartel[1,2,3]‡

RNA interference (RNAi), a gene-silencing pathway triggered by double-stranded RNA, is conserved in diverse eukaryotic species but has been lost in the model budding yeast *Saccharomyces cerevisiae.* Here, we show that RNAi is present in other budding yeast species, including *Saccharomyces castellii* and *Candida albicans*. These species use noncanonical Dicer proteins to generate small interfering RNAs, which mostly correspond to transposable elements and Y′ subtelomeric repeats. In *S. castellii*, RNAi mutants are viable but have excess Y′ messenger RNA levels. In *S. cerevisiae*, introducing Dicer and Argonaute of *S. castellii* restores RNAi, and the reconstituted pathway silences endogenous retrotransposons. These results identify a previously unknown class of Dicer proteins, bring the tool of RNAi to the study of budding yeasts, and bring the tools of budding yeast to the study of RNAi.

RNA-silencing pathways contribute to transposon silencing, viral defense, DNA elimination, heterochromatin formation, and posttranscriptional repression of cellular genes (*1*, *2*). In the simplest form of silencing, known as RNA interference (RNAi), the ribonuclease III (RNaseIII) endonuclease Dicer successively cleaves double-stranded RNA (dsRNA) into small interfering RNAs (siRNAs), which are loaded into the effector protein Argonaute to guide the cleavage of target transcripts (*1*, *3*). RNAi arose in an early eukaryotic ancestor and appears to have been conserved throughout most of the fungal kingdom (*4*, *5*) (Fig. 1A). A prominent exception is *Saccharomyces cerevisiae*, a budding yeast that lacks recognizable homologs of Argonaute, Dicer, and RNA-dependent RNA polymerase (RdRP), which in some RNAi pathways produces dsRNA. Indeed, RNAi has been presumed lost in all budding yeasts. Despite this perceived loss, Argonaute genes are present in some budding yeasts (*6*, *7*), including *Saccharomyces castellii* and *Kluyveromyces polysporus* (both close relatives of *S. cerevisiae*) and *Candida albicans* [the most common yeast pathogen of humans (*8*)] (Fig. 1A). The presence of these genes in budding yeast has been enigmatic, because other RNAi genes, especially Dicer, have not been found in these species. A similar conundrum appears in prokaryotes, in which certain bacteria have Argonaute homologs yet lack the other genes associated with RNAi or related RNA-silencing pathways (*9*).

siRNAs in budding yeasts. To search for RNA silencing in budding yeast, we looked for short-guide RNAs, isolating 18- to 30-nucleotide (nt) RNAs from *S. castellii*, *K. polysporus*, and *C. albicans* and preparing sequencing libraries representing the subset of small RNAs with 5′-monophosphates and 3′-hydroxyls (*10*), which are the chemical features of Dicer products. The small RNAs of *S. castellii* and *K. polysporus* were most enriched in 23-RNAs beginning with U, and those of *C. albicans* were most enriched in 22-nt beginning with A or U (Fig. 1B). These biases were reminiscent of those observed for Argonaute-bound guide RNAs of animals, plants, and other fungi (*11*–*13*). Analogous RNAs were not found in *S. cerevisiae*, as expected for a species lacking RNAi (Fig. 1B).

Although some reads from the Argonaute-containing yeasts mapped to ribosomal RNA (rRNA) and transfer RNA (tRNA) and presumably represented degradation intermediates of abundant RNAs, many reads clustered at other types of genomic loci. The loci generating the most reads had sequence homology to repetitive elements, including long terminal repeat retrotransposons (Ty elements), LINE (long interspersed nuclear element)–like retrotransposons (Zorro elements), and subtelomeric repeats (Y′ elements) (Fig. 1C and table S1). Loci of *S. castellii* were also particularly enriched in long inverted repeats; these palindromic loci generated most of the reads with homology to Ty elements (Fig. 1, C and D). In *S. cerevisiae*, essentially all the reads appeared to represent degradation fragments of rRNA, tRNA, and mRNA.

The reads matching inverted repeats suggested origins from paired regions of transcripts that folded back on themselves to form hairpins (Fig. 1D). These inferred hairpins had 100- to 400-bp (base pair) stems, with loops ranging from 19 to >1600 nt. In regions of imperfect duplex, where reads could be mapped unambiguously, the small RNAs tended to match only one genomic strand, which further supported the idea that they originated from hairpin transcripts (Fig. 1D, bottom). Other reads did not map to inverted repeats and, instead, mapped uniquely to both genomic strands in a pattern suggesting that they originated from long bimolecular duplexes involving transcripts from both strands.

Most siRNAs of the fission yeast *Schizosaccharomyces pombe* correspond to the outer repeats of the centromeres and direct heterochromatin formation and maintenance (*14*). We therefore examined whether any of our sequenced small RNAs matched centromeres. Of the three Argonaute-containing species from which we sequenced (Fig. 1B), only *C. albicans* had annotated centromeres, and almost none (<0.001%) of our *C. albicans* reads matched these genomic loci. Also arguing against a function analogous to that in *S. pombe* is the lack in budding yeasts of recognizable orthologs of the histone H3 lysine 9 (H3K9) methyltransferase Clr4 and recognizable homologs of RdRP, Tas3, Chp1, and the heterochromatin protein HP1-like chromodomain protein Swi6—proteins all necessary for RNAi-dependent heterochromatin in *S. pombe* (*14*).

When mapped to the genome, the end of one 23-nt RNA was often next to the beginning of another 23-nt RNA, which suggested that endonuclease cleavage simultaneously generated the 3′ terminus of one small RNA and the 5′ terminus of the next. Consistent with this hypothesis, systematic analysis of the intervals spanning the mapped ends of all 23-nt RNA pairs revealed a clear phasing interval of 23 nt (Fig. 1E). Such phasing implied successive cleavage, beginning

[1]Whitehead Institute for Biomedical Research, 9 Cambridge Center, Cambridge, MA 02142, USA. [2]Howard Hughes Medical Institute, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [3]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [4]Smurfit Institute of Genetics, Trinity College Dublin, Dublin 2, Ireland.

*These authors contributed equally to this work.
†Present address: Center for Plant Science Innovation and Department of Agronomy and Horticulture, University of Nebraska, Lincoln, NE 68588, USA.
‡To whom correspondence should be addressed. E-mail: dbartel@wi.mit.edu

www.sciencemag.org/cgi/content/full/326/5952/540/DC1

Supporting Online Material for

Intraspecific Polymorphism to Interspecific Divergence: Genetics of Pigmentation in *Drosophila*

Patricia J. Wittkopp,* Emma E. Stewart, Lisa L. Arnold, Adam H. Neidert, Belinda K. Haerum, Elizabeth M. Thompson, Saleh Akhras, Gabriel Smith-Winberry, Laura Shefner

*To whom correspondence should be addressed. E-mail: wittkopp@umich.edu

Published 23 October 2009, *Science* **326**, 540 (2009)
DOI: 10.1126/science.1176980

This PDF file includes:

Materials and Methods:

Sequencing *D. americana* and *D. novamexicana tan* and *ebony*

Bacterial artificial chromosomes (BACs) containing *tan* and *ebony* were identified in *D. americana* and *D. novamexicana* genomic libraries by screening filters from the Arizona Genomics Institute. Radioactively labeled probes used for screening were produced with PCR products amplified by the following primers: *HMR*, located 5' of *tan* (5'-CATCTCGTCCAACTCCAGGT-3' and 5'-GCGCTATAAATATCAGCGTCA-3'); *CG7039* located 3' of *tan* (5'-CATTGCTGCACGGCTTTTAC-3' and 5'-CTCCACCAGCCATTTGATG-3'); *ETHR* located 5' of *ebony* (5'- GGCTGTCGCTGCTGTTATTT-3' and 5'- CCAAGCCGCAAATAAGTTTC-3'); and *CG5874* located 3' of *ebony*, (5'- GCCTGCACCTGCACCA-3' and 5'-CCACGCTAATTCCAACCAAC-3'). These primers were designed on the basis of sequence from the August 2005 *D. virilis* genome assembly (*S1*). We ordered six *D. americana* and six *D. novamexicana* clones positive for both genes flanking *tan* and five *D. americana* and ten *D. novamexicana* clones positive for both genes flanking *ebony* from the Arizona Genomics Institute. Each clone was evaluated by testing for amplification of PCR products from *tan*, *ebony*, and flanking genes. Ultimately, the DA_ABa0020L7 (*D. americana*) and DN_Ba0024C15 (*D. novamexicana*) clones containing *tan* as well as the DA_ABa0029H3 (*D. americana*) and DN_Ba0007J18 (*D. novamexicana*) clones containing *ebony* were selected for sequencing. With a combination of primers designed on the basis of the *D. virilis* sequence and primer walking, we sequenced ~14 kb from each *tan* BAC clone and ~4kb from each *ebony* BAC clone.

Fly strains, rearing, and imaging

The following lines of *D. melanogaster* were used for this work: *pnr*-Gal4 (G. Morata); *w*-;P[*w*+, UAS-Tan] (J. True); *w*-;P[*w*+, UAS-Ebony] (J. True); *yellow1*, *tan5* (J. True), and CantonS. A *w*-;*CyO/Sb*; *TM2/TM6* balancer line was used to construct genotypes described in Fig. 3. *D. americana* and *D. novamexicana* lines and sources are shown in Table S2. All flies were reared on standard yeast-glucose media, with *D. americana* and *D. novamexicana* reared at 20°C, and *D. melanogaster* reared at 25°C. Pigmentation of individual flies was documented by placing age-controlled adults in a solution of 10% glycerol in ethanol, storing at room temperature for 1-30 days, dissecting the dorsal abdominal cuticle, mounting the cuticle in Hoyer's solution, and imaging with a Scion 1394 (Frederick, MD) digital camera. All images presented within a figure (or within a panel for Fig. 5) are from age-matched flies with images captured at the same time, under the same lighting conditions. Colors were adjusted in Photoshop CS2 (Adobe, San Jose, CA) to best reproduce visual observations, with an

identical color transformation applied to all images shown within the same figure (or panel for Fig. 5). Pigmentation of a fly was quantified by measuring the intensity of grey-scale images with ImageJ (NIH, Bethesda, MD) in five to ten non-overlapping regions and averaging the median intensity from each region. Measurements range from 0 (black) to 255 (white).

Quantitative trait locus (QTL) mapping

All QTL mapping experiments were performed with interspecific backcross populations, as described in the main text. Within each population, progeny were visually classified into distinct pigmentation classes. All phenotypic scoring was performed by E.E.S. under controlled lighting conditions and pigmentation of each fly was verified by P.J.W prior to DNA extraction. Single fly DNA preparations were performed with the protocol described in (*S2*).

Genotypes of *tan*, *ebony*, and *yellow* were determined for backcross progeny with either DNA sequencing or amplified fragment length polymorphisms resolved with agarose gel electrophoresis. For *yellow,* genotypes were determined by directly sequencing a 632 bp band amplified with the following PCR primers: 5'-CTAAACATGCCTGAAAATCAATCACGGA-3' and 5'-CGTTGGTAAACGAAAGTCCAATTGG-3'. For *tan*, the primers 5'-CGAGTTTTTATTCCCACTGAATTAT-3' and 5'-GGGTTCGTCTTATCCACGAT-3' amplified a 99 bp band in *D. americana* and a 64 bp band in *D. novamexicana*. For *ebony*, the primers 5'-GTTGTGCCAAACTGAAAGATCC-3' and 5'-CACATTCACACTTTGTGCACTTG-3' amplified a 162 bp band in *D. novamexicana* and a 244 bp band in all *D. americana* lines except DN2, which amplified a 162 bp band identical to *D. novamexicana*. For the DN2 backcross population, *ebony* genotypes were determined by examining heterozygous bases in chromatograms from directly sequenced 162 bp PCR products.

To test for a significant association between genotypes and phenotypes, two-sided Fisher exact tests, evaluated with "fisher.test()" in R Cocoa GUI 1.12 (http://www.r-project.org/), were used to compare the number of *D. americana* alleles observed among the lightest and darkest flies. For the DN4 backcross population, 101 out of 127 flies were successfully genotyped for both *ebony* and *tan*. The proportion of pigmentation variance explained by *ebony* and *tan* genotypes in the DN4 backcross was determined by fitting the following model with PROC MIXED in SAS v.9.1 (Cary, NC) with Type III sums-of-squares: $Y_{ijk} = E_i + T_j + ET_{ij} + \varepsilon_{ijk}$, where Y is the pigmentation score of each fly, E is the *ebony* genotype (i = *D. americana* or *D. novamexicana*), T is the *tan* genotype (j = *D. americana* or *D. novamexicana*), ET is the interaction between *tan* and *ebony* genotypes, and ε is a random error term. E, T, and ET, were treated as fixed effects in the model. A second model, lacking the interaction term, was used to quantify the amount of variance explained by additive effects of *ebony* and *tan* genotypes.

Constructing introgression lines

D. americana alleles of *ebony* and *tan* from the DN12 line were introgressed into the N14 line of *D. novamexicana*. This line of *D. americana* was used after three failed attempts to introgress alleles from the line of *D. americana* (A00) used for the initial genetic mapping. Note that the *tan* sequences from A00 and DN12 are very similar (see Fig. 5C) and differ only by 2 bp in the 2.7 kb region identified by fine-scale genetic mapping (GQ457339 and GQ457347). DN12 and A00 also share the same arrangement of the In(2)b inversion containing *ebony* (*S3*).

As shown in Fig. S2, *D. americana* virgin females were crossed to *D. novamexicana* males, and F_1 hybrid virgin females were backcrossed to *D. novamexicana* males. Virgin females were collected from this first back-cross (BC1), and then we set up twenty matings, each containing one virgin female from BC1 and one *D. novamexicana* male. When third instar larvae were visible (two to three weeks after mating), the female parent was removed from each vial, DNA was extracted, and genotypes at *tan*

and *ebony* were determined with DNA sequencing and PCR-based genotyping as described above for QTL mapping. One vial containing larval progeny from a mother heterozygous for *tan* and/or *ebony* was randomly selected, and twenty virgin females were collected from this brood after eclosion. To begin the next backcross generation (BC2), twenty pair-matings were set-up by crossing each of these females to one *D. novamexicana* male. This process was repeated for ten generations.

On average, after ten generations of backcrossing, an introgressed region extends 10 cM to either side of the selected locus (i.e., *ebony* and *tan*) (*S4*). The precise breakpoints of introgressed *tan* and *ebony* regions remain unknown, although DNA sequencing showed that *D. americana* alleles remained at loci ~500 kb from *tan* in both directions. Polytene chromosomes squashes of the *tan* introgression line showed the loss of *D. americana* alleles for all chromosomal regions inverted between *D. americana* and *D. novamexicana*, and squashes of the *ebony* introgression lines showed the loss of inversions other than the *ebony* containing In(2)b. Importantly, mendelian inheritance of pigmentation was observed for both of the final introgression lines, showing that any remaining *D. americana* alleles not linked to *tan* or *ebony* do not visibly affect pigmentation.

Fine-scale genetic mapping

Virgin females heterozygous for the introgressed *D. americana tan* allele were crossed to *D. novamexicana* males. 5048 male offspring were visually scored for pigmentation by A.H.N. under constant light conditions, with each classified as either light or dark. Light flies had pigmentation most similar to wild-type *D. novamexicana* and dark flies had pigmentation most similar to the *tan* introgression line. DNA was extracted from each fly with the protocol described in (*S2*), except that flies were homogenized by placing a single fly into a well of a 96-well PCR plate, adding a single glass bead, and shaking on a Mixer Mill MM301 (Retsch, Inc., Haan, Germany) for 10 seconds at a frequency of 25 Hz. This shaking condition was found to homogenize the flies sufficiently for DNA extraction without completely destroying the abdominal cuticle.

Genotypes at molecular makers A, B, C, and D in Fig. S3 were determined with the following primer pairs: (A) 5'-TTATATCGCCGGGTATCAGC-3' and 5'-CGTCTGATGCTTTCTGACGA-3'; (B) 5'-CGAGTTTTTATTCCCACTGAATTAT-3' and 5'-GGGTTCGTCTTATCCACGAT-3'; (C) 5'-GGAGTCCATGTGGCCTAAGAAC -3' and 5'-GCCTTATCTTAATAGAAGTTTAATATGC-3'; and (D) 5'-TCGAACATGTTTGGCCTTGTCAC-3' and 5'-GTTTATAGCCAGCAGTTGCTG-3'.

PCR products from B, C, and D differed in length between the N14 line of *D. novamexicana* and the DN12 line of *D. americana* (i.e., the lines used for fine-scale genetic mapping). The PCR product amplified from locus A was cut with HaeIII, producing different sized fragments for the two alleles. Two flies were found that inherited the *D. americana* allele at one of these loci and the *D. novamexicana* allele at the other. The location of the recombination breakpoint in each of these flies was determined by re-sequencing the *tan* gene. The recombination breakpoint was located between positions 689 and 752 in F4 and between positions 3500 and 3658 in D1. These positions refer to the PopSet alignment of GQ457336-GQ457353. Flies were also genotyped for molecular markers outside of this region (i.e., A and D in Fig. S3) to estimate genetic distances between A and B as well as between C and D.

Analyzing *tan* mRNA expression

in situ hybridizations were performed as described in (*S5*). Briefly, *D. americana* (A00) and *D. novamexicana* (N14) pupae were collected one to four hours prior to eclosion and heated to 100°C for one minute. Dorsal abdominal cuticle was dissected, fixed in 4% paraformaldehyde, dehydrated, and stored at -20°C. After re-hydrating, cuticles were fixed again in 4% paraformaldehyde, treated with proteinase K, fixed a third time, and incubated overnight at 65°C with a Digoxigenin (DIG)-labeled RNA probe. Sense and anti-sense RNA probes were synthesized with T7 polymerase (Promega, Madison, WI) to transcribe PCR products containing ~300 bp of sequence from *D. novamexicana tan*

exon 8. *D. americana* and *D. novamexicana tan* alleles have identical sequence in this region. After washing, samples were incubated overnight with anti-DIG AP Fab fragments (Roche, Basel, Switzerland), washed again, and incubated with a solution containing NBT/BCIP (Promega, Madison, WI) for 20 minutes or until sufficient colorametric signal was obtained. Three to twenty cuticles of *D. americana* and *D. novamexicana* were processed in parallel, and a titration series of probe and antibody conditions was examined. This experiment was repeated more than ten times.

To quantify standing levels of *tan* mRNA, stage P14 and P15 pupae from *D. americana* (A00), *D. novamexicana* (N14) and F_1 hybrids were collected and stored at -80. Prior to freezing, pupae were dissected from their cases, and their heads were removed to eliminate *tan* transcripts associated with the visual system (*S6*). Four samples, each containing three *D. americana* and three *D. novamexicana* pupal bodies, as well as five samples, each containing six F_1 hybrid bodies, were homogenized and used for sequential RNA and genomic DNA extractions, as described in (*S7*). cDNA was synthesized from each RNA sample with a polyT primer and Superscript 2 reverse transcriptase (Invitrogen, Carlsbad, CA).

Pyrosequencing was performed as described in (*S8*) with the following pair of primers: 5'-GATGCTGAAGTCCAGCGTGTC-3' and 5'-BIO-CAGCCGCCAGTGACATCA-3', where "BIO" indicates the addition of a biotin molecule. A Pyrosequencing primer (5'-CGAGCACGATGTCCG-3') was used to analyze the sequence CAAYATG, in which the *D. americana* allele contains a thymine (T) and the *D. novamexicana* allele contains a cytosine (C) at the variable position. Pyrosequencing reactions were performed for each cDNA and genomic DNA sample, with a minimum average peak height of twelve for the conserved C, T, and G positions required for quality control.

The relative abundance of *D. americana* and *D. novamexicana tan* alleles was calculated as the ratio between the polymorphic T and C peaks (i.e., *D. americana*/*D. novamexicana* = T/C). These ratios were $\log_2$ transformed to make them normally distributed (*S8*), and the $\log_2$ genomic DNA ratio was subtracted from the corresponding $\log_2$ cDNA ratio to correct for any bias between alleles in PCR-amplification and/or nucleic acid extraction in mixed species pools (*S7*). We tested for expression differences between species and for allele-specific expression differences in F_1 hybrids using "t.test()" in R Cocoa GUI 1.12 (H_0: $\mu = 0$). The average percent difference in *tan* expression between species was calculated as $(1-2x) * 100$, where x = the mean value of $\log_2$(T/C) for the four biological replicate samples.

Transgene construction and transformation

piggyBac transgenes containing the *D. americana* and *D. novamexicana tan* alleles were constructed by recombineering with the general protocols and strains described at http://recombineering.ncifcrf.gov/. Targeting plasmids for *D. americana* and *D. novamexicana tan* were produced by amplifying ~500 bp "homology arms" from HMR and CG7039 with PCR primers, using PCR sewing to connect them with an Xho I site in between, and inserting them into the AscI site of a piggyBac plasmid (*S9*) with AscI sites flanking the homology arms that were introduced during the initial PCR amplification. The piggyBac plasmid, already containing a 3xP3-EGFP transformation marker (*S9*), was modified by inserting the attB sequence (*S10*) into the Xba I site located in the pUC18 backbone. Separate targeting vectors were made for *D. americana* and *D. novamexicana tan*, and the sequence of these constructs was confirmed by DNA sequencing.

Each 7 kb piggyBac targeting vector was linearized with Xho I, gel purified, treated with calf intestine alkaline phosphotase, and electroporated into SW102 cells carrying the appropriate BAC clone: DA_ABa0020L7 for *D. americana* and DN_Ba0024C15 for *D. novamexicana*. SW012 cells contain all of the genetic resources needed for recombineering (*S11*). Following induction of the recombinase and ampicillin selection for circularized piggyBac plasmids, individual colonies were screened by PCR, mini-prepped, and subject to diagnostic restriction digests. One positive clone from

each species was re-sequenced for the entire 14kb *tan* transgene. In both cases, recombineering was found to have produced an exact replica of the BAC sequence.

piggyBac plasmids carrying the *D. americana* and *D. novamexicana tan* transgenes were injected into *D. melanogaster white* mutants carrying the attP16 site on chromosome two (*S12*), and inserted into the genome with the phiC31 integrase (*S10*). Each piggyBac plasmid was also co-injected with a piggyBac transposase source (*S9*) into wild-type *D. novamexicana* (N14) and *white* mutant *D. americana* (provided by B. McAllister). Genetic Services, Inc (Sudbury, MA) performed all embryo injections and screening.

Sequence polymorphism discovery and analysis

DNA was extracted from males of each isofemale line listed in Table S2. *tan* was sequenced in twelve lines of *D. americana* and three lines of *D. novamexicana* with primers developed during the initial sequencing of *tan* alleles from BAC clones. Additional loci flanking *tan*, described in Fig. S8A, were also sequenced in multiple lines of *D. americana* and *D. novamexicana*. Three regions of *ebony* were amplified and sequenced from ten lines of *D. americana* and five lines of *D. novamexicana* (see Fig. 5D). These regions were concatenated for phylogenetic analysis. Primers used to collect these sequences are summarized in Table S3. In lieu of sequencing, AP-1 Gamma was genotyped with the primers 5'-TCGAACATGTTTGGCCTTGTCAC-3' and 5' TTTATAGCCAGCAGTTGCTG-3, which amplified a 158 bp product in *D. novamexicana*, and a 100bp product in *D. americana* (including the A01 allele).

Sequencing reactions were performed by the University of Michigan Sequencing Core Facility and raw sequence data was analyzed with Sequencher 4.6 (Gene Codes Corporation, Ann Arbor, MI) and Codon Code Aligner 2.0.6 (CodonCode Corporation, Dedham, MA). Sequences for each allele were aligned with MUSCLE (*S13*) and then manually inspected by P.J.W. Phylogentic trees were built with MEGA 4.0.2 (*S14*). DNAsp 4.90.1 (*S15*) and SITES (*S16*) were used to identify fixed differences and polymorphic sites.

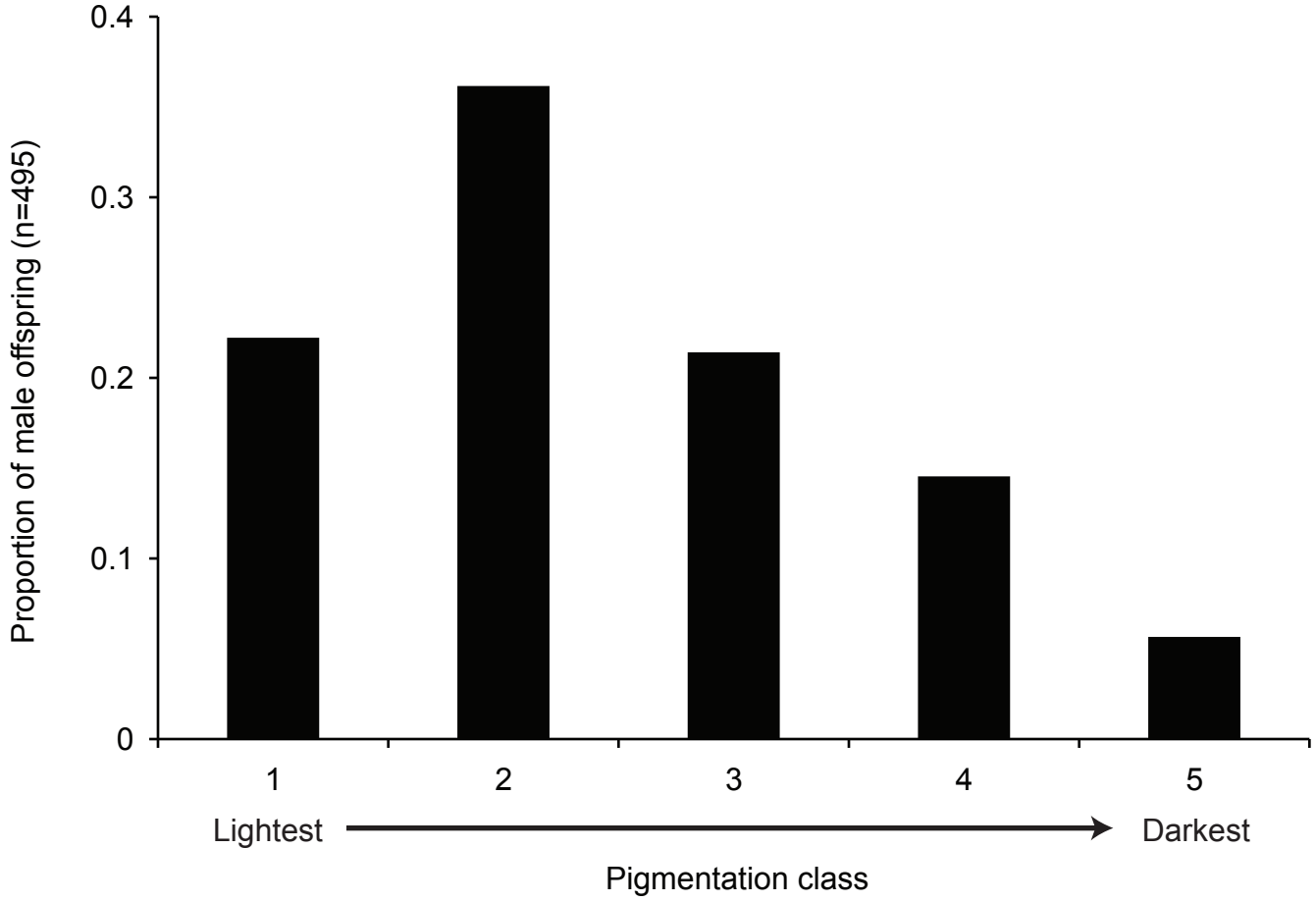


Fig. S1. Interspecific backcross progeny are distributed among five pigmentation classes. F_1 hybrid females, produced by crossing *D. americana* (A00) females to *D. novamexicana* (N14) males, were backcrossed to *D. novamexicana* (N14) males. Body color was scored by eye for 495 male progeny. Pigmentation phenotypes were not continuous, but rather fell into five distinct classes. The phenotypic distribution observed is similar to that reported in (*S17*). Light and dark flies genotyped came from classes 1 and 5, respectively.

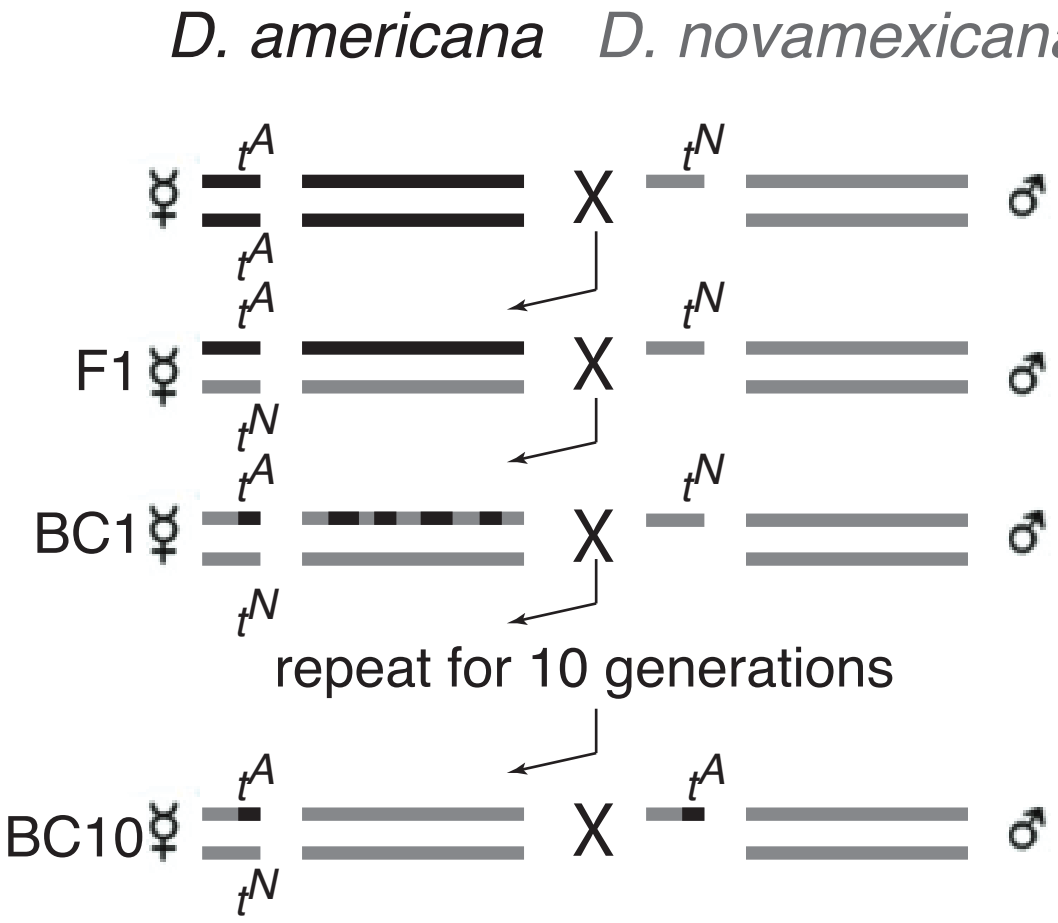


Figure S2. Crossing scheme used to introgress *D. americana* alleles linked to *tan* into *D. novamexicana*. Short bars represent X chromosomes and long bars represent all five autosomes for *D. americana* (black) or *D. novamexicana* (grey). Alleles of *tan* from *D. americana* (t^A) and *D. novamexicana* (t^N) are indicated. An analogous crossing scheme was used to introgress autosomal alleles linked to *ebony* from *D. americana* into *D. novamexicana*.

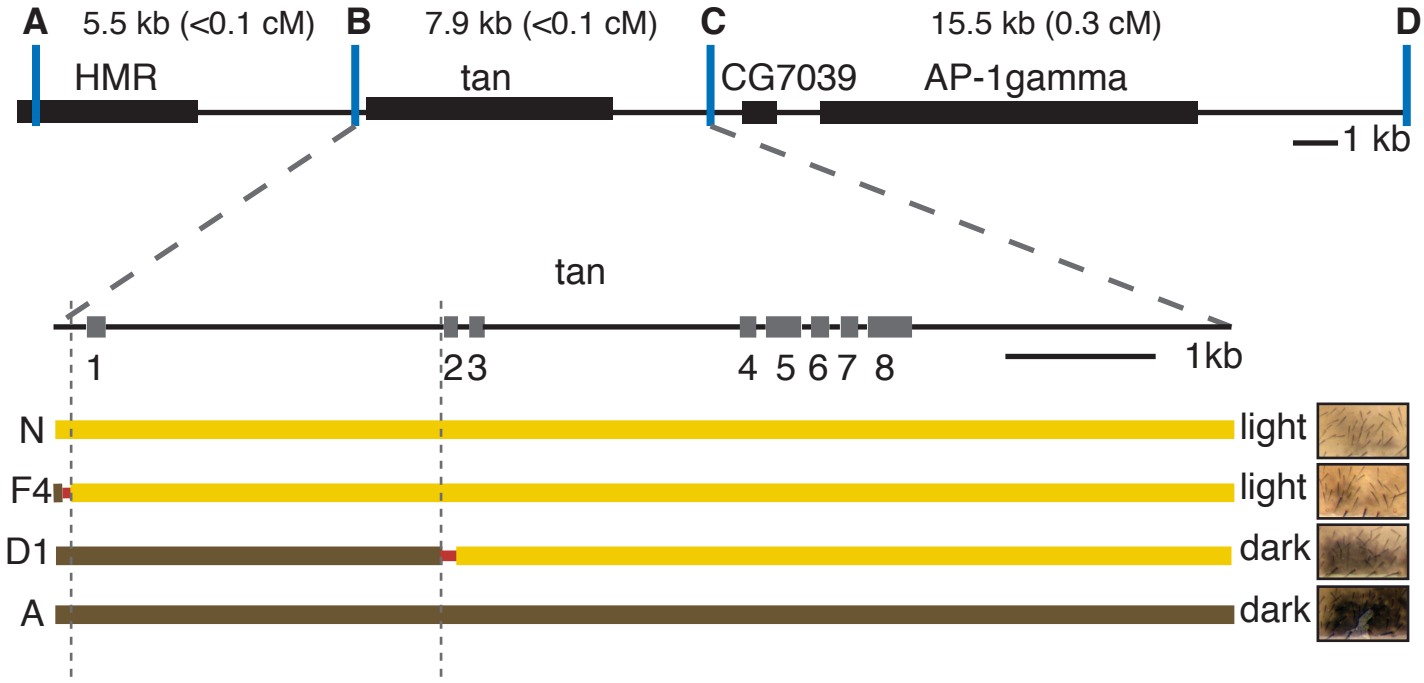


Figure S3. A 2.7 kb region of *tan* contributes to pigmentation divergence. Genomic structure of *tan* and flanking genes is shown to scale. Genetic distances indicated between molecular markers (labeled A-D) were determined empirically. Below, a more detailed schematic of *tan* is shown with vertical dotted lines delineating the 2.7 kb region containing sites inferred to affect pigmentation, and grey boxes representing exons. Recombination breakpoints occurred between positions 689 and 752 in F4 and between positions 3500 and 3658 in D1. Representations of recombinant genotypes, their corresponding phenotypic classifications (light or dark), and pictures of dorsal abdominal cuticle recovered after DNA extraction (19) are also shown. Yellow represents *D. novamexicana* sequence and brown represents *D. americana* sequence. Red bars show regions that do not differ between species. Genotypes A and N show flies with *D. americana* or *D. novamexicana* alleles at molecular markers B and C, respectively.

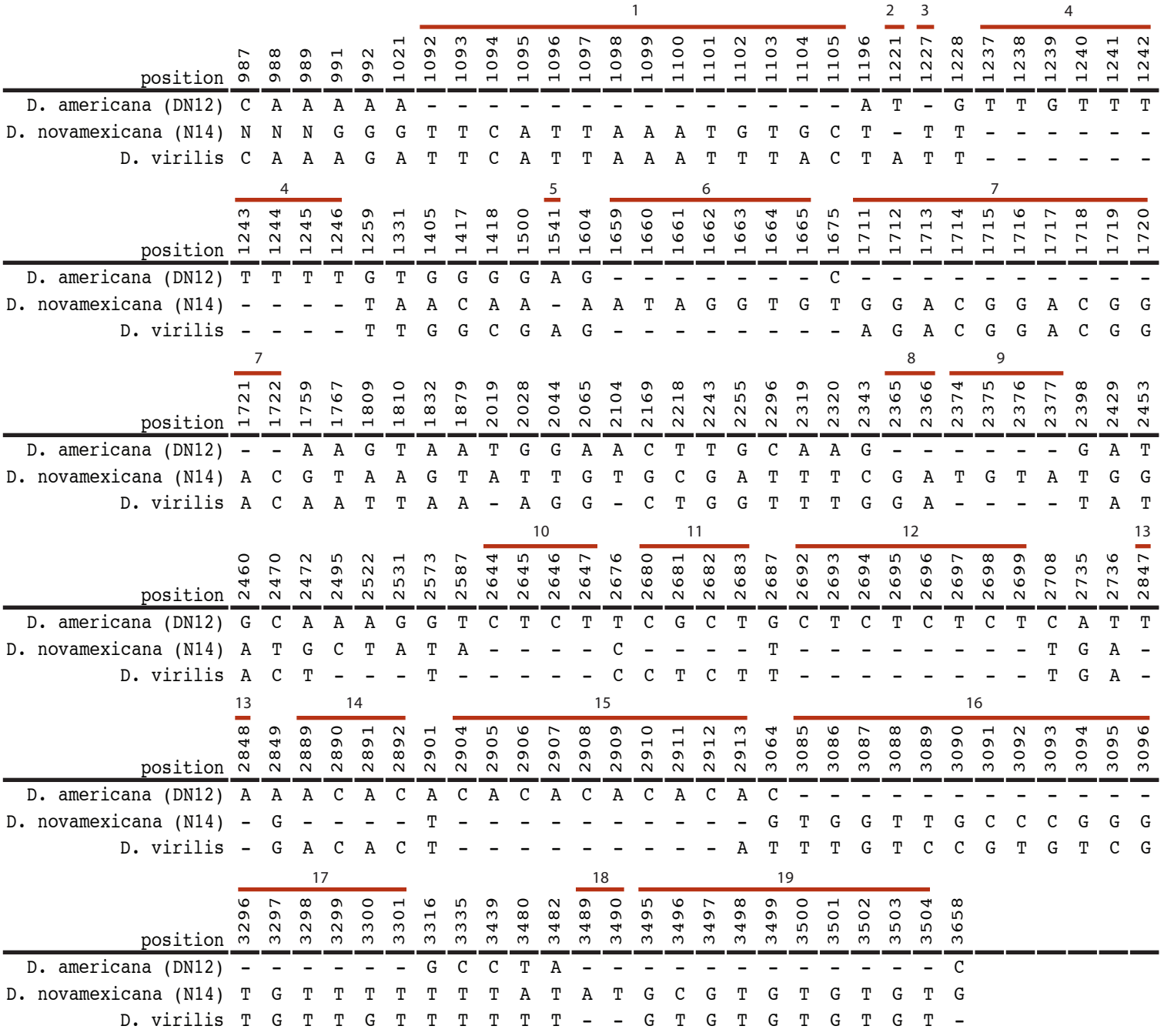


Fig. S4. Sequence alignment reveals candidate sites for functional divergence of *tan* alleles. Fine-scale genetic mapping identified a 2.7 kb region of *tan* that is functionally divergent between *D. americana* and *D. novamexicana* (Fig. S3). All divergent sites between the DN12 allele of *D. americana* and the N14 allele of *D. novamexicana* are shown. *D. virilis* alleles at these sites are also shown. Site positions refer to PopSet alignment of GQ457336-GQ457353. The 2.7 kb candidate region begins at position 752 and extends through position 3658 in this alignment. The red lines numbered 1 to 19 indicate indels.

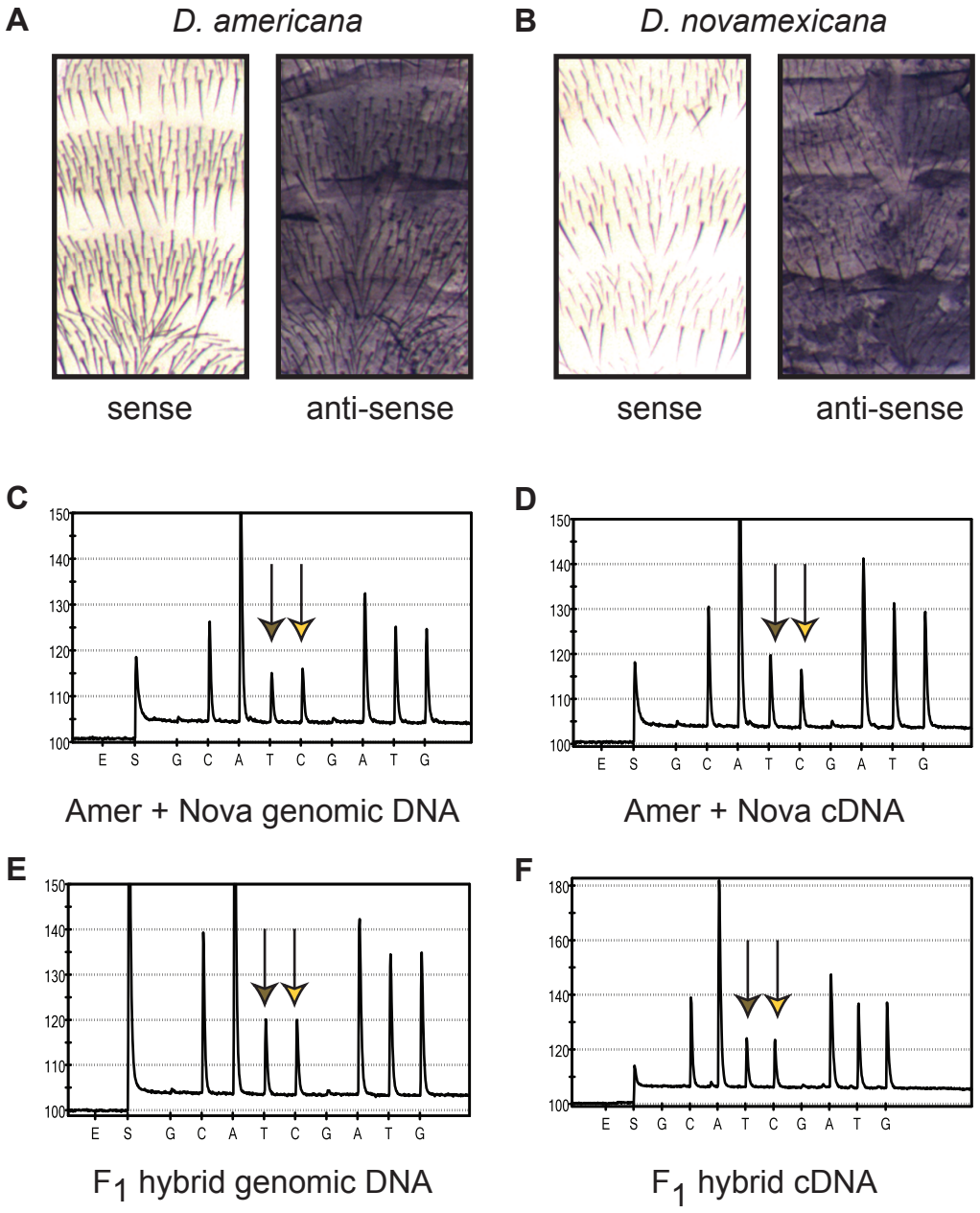


Fig. S5. *tan* mRNA is more abundant in *D. americana* than in *D. novamexicana*. *in situ* hybridization with an oligonucleotide probe complementary to *tan* mRNA (anti-sense) showed similar expression of *tan* in *D. americana* (A) and *D. novamexicana* (B). Control probes composed of sequences identical to the *tan* mRNA (sense) showed no hybridization signal in either species (A, B). The intensity of staining was variable from cuticle to cuticle and experiment to experiment, and no systematic differences in hybridization signal were apparent between species. To more quantitatively compare *tan* expression between species, we used Pyrosequencing (*S7*). Sample pyrograms are shown for (C) genomic DNA extracted from a pool of *D. americana* and *D. novamexicana* pupae, (D) cDNA synthesized from RNA extracted from the same pool of pupae, (E) F_1 hybrid genomic DNA, and (F) cDNA synthesized from F_1 hybrid pupal RNA. Letters below each pyrogram refer to the sequential addition of enzymes (E), substrate (S), and single nucleotides (A, G, C, or T) to the reaction. (C-F) *D. americana*-specific peaks are indicated by brown arrowheads and *D. novamexicana*-specific peaks are indicated by yellow arrowheads.

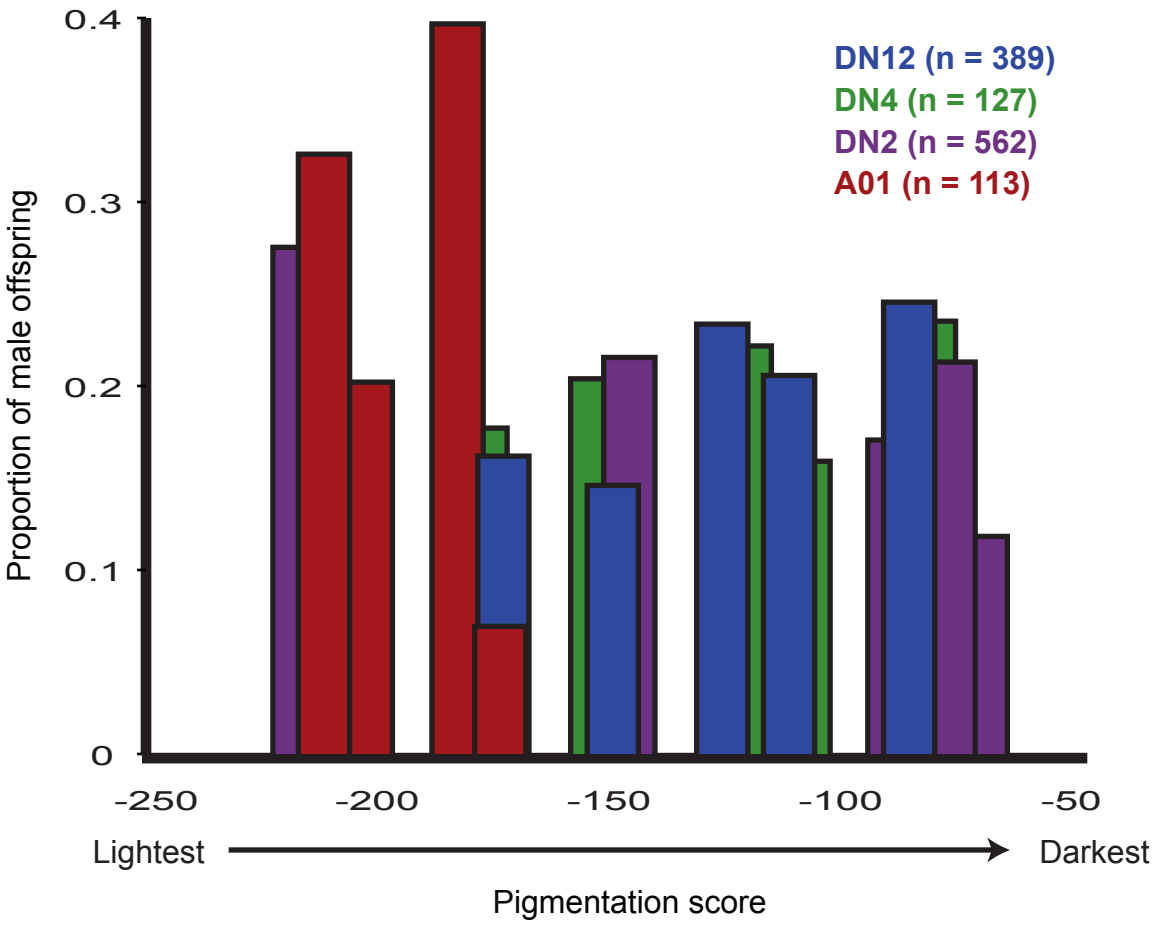


Fig. S6. Different phenotypic distributions were observed among backcross progeny from different strains of *D. americana*. Females from the DN12, DN4, DN2 and A01 strains of *D. americana* were each mated to *D. novamexicana* (N14) males; F_1 females were backcrossed to *D. novamexicana* (N14) males; and male progeny were classified by eye into as many distinguishable pigmentation classes as possible. A histogram of pigmentation phenotypes is shown, with the sample size (n) for each backcross population listed in the key. Pigmentation classes are plotted on the X-axis on the basis of the average pigmentation intensity of five randomly chosen flies from that class. DN12 (blue) and DN4 (green) backcrosses produced similar phenotypic distributions in terms of both the frequency and pigmentation of each phenotypic class. These distributions were distinct from that observed with A00 (Fig. S1). The DN2 backcross also produced a phenotypic distribution distinct from A00 (Fig. S1), but this population was also distinct (in both pigmentation phenotypes and frequency) from the DN12 and DN4 backcross distributions. Crosses with the A01 line of *D. americana* produced only four recognizable classes of backcross progeny, all of which were lighter than most backcross progeny produced with DN2, DN4, or DN12.

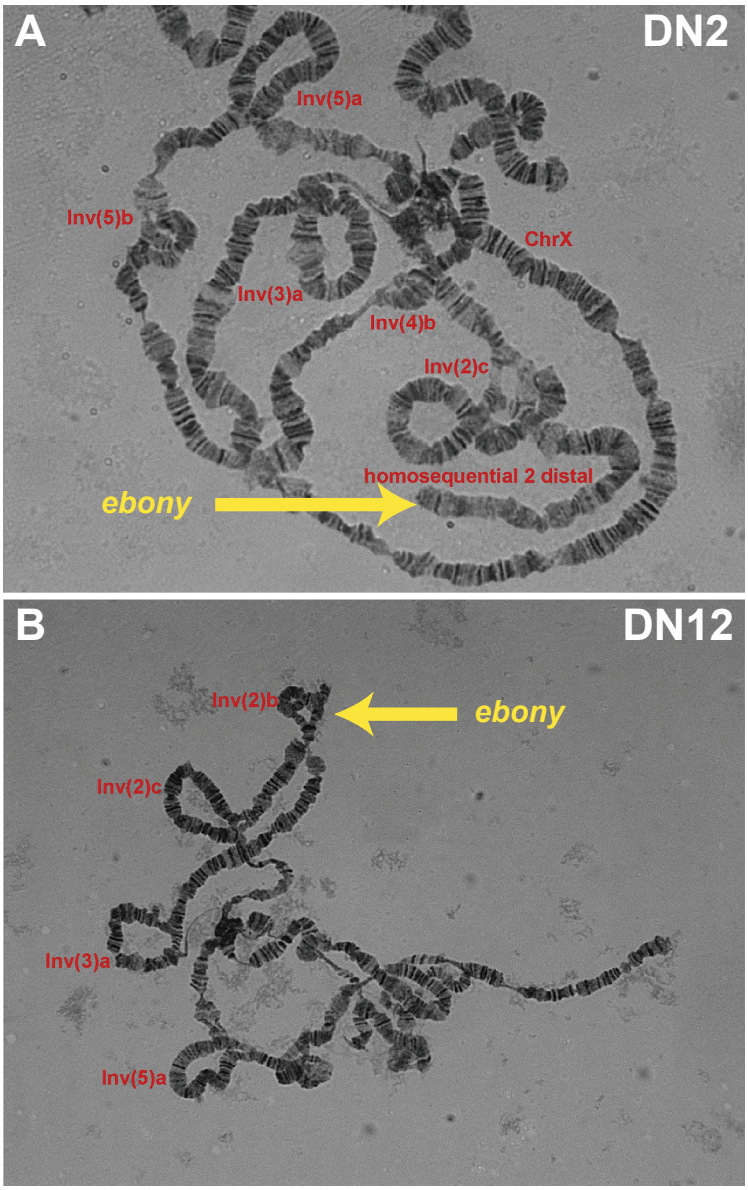
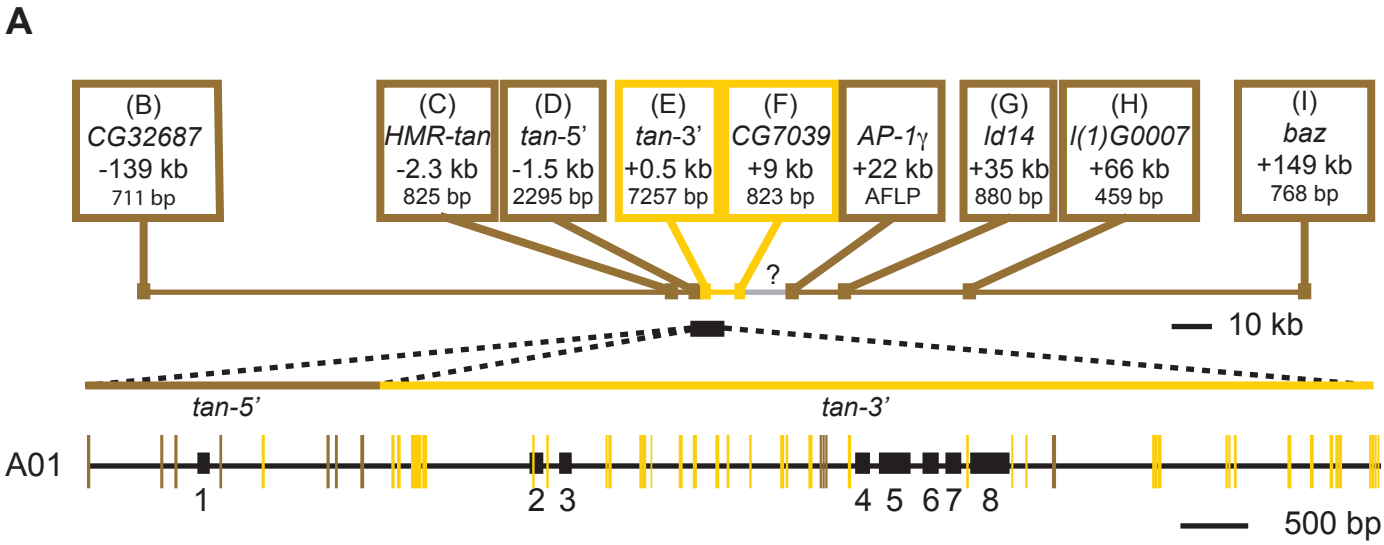
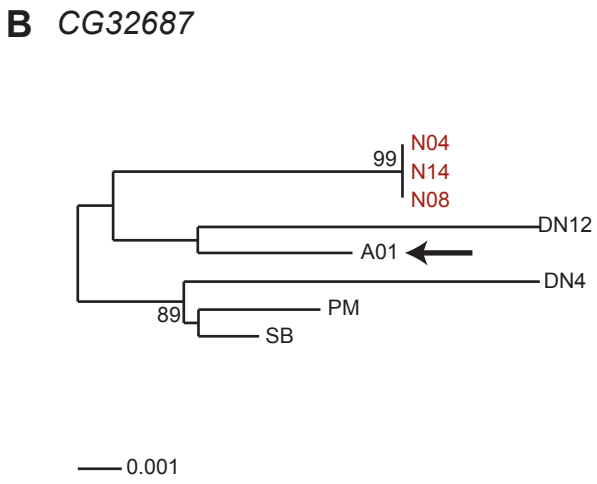
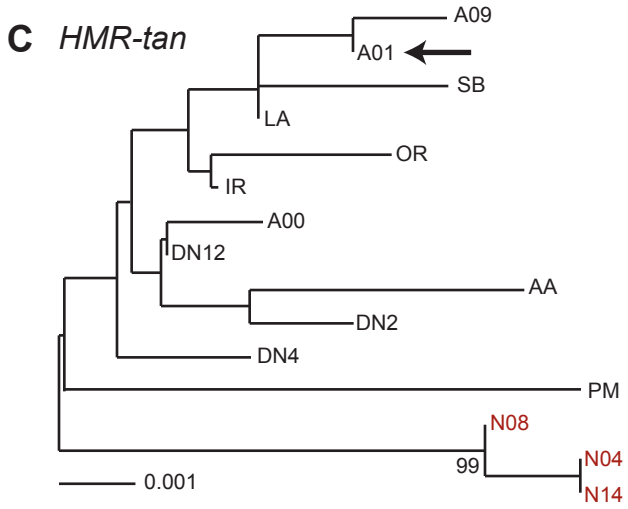


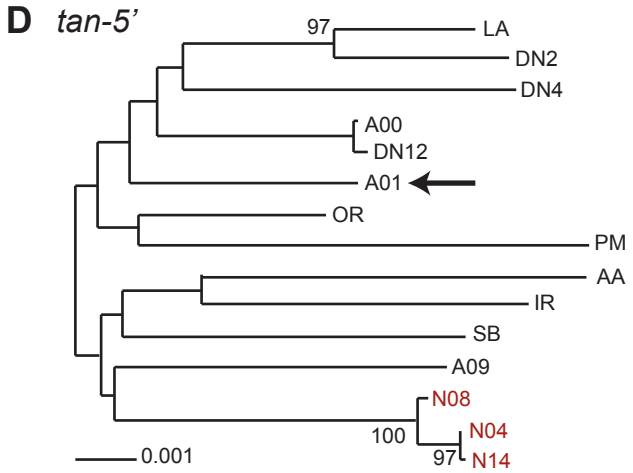
Fig. S7 The genomic region containing *ebony* is inverted between *D. novamexicana* and the DN12 line of *D. americana* , but not the DN2 line of *D. americana*. Polytene chromosomes from interspecific F_1 hybrid males show the location of inversions between *D. novamexicana* and two strains of *D. americana*: DN2 (A) and DN12 (B). The distal end of the 2nd chromosome, which contains *ebony*, is indicated with a yellow arrow in each panel. Note the presence of the *ebony*-containing In(2)b inversion in DN12 (B) but not DN2 (A). Other visible inversions are also indicated. This experiment was performed and analyzed by Paulina Mena and Bryant McAllister (University of Iowa); images are reproduced here with permission.

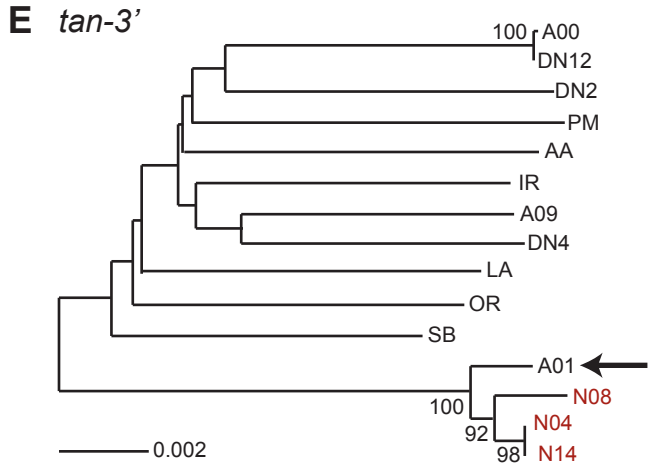
Fig. S8 (page 1)


A
(B)
CG32687
-139 kb
711 bp
(C)
HMR-tan
-2.3 kb
825 bp
(D)
tan-5'
-1.5 kb
2295 bp
(E)
tan-3'
+0.5 kb
7257 bp
(F)
CG7039
+9 kb
823 bp
AP-1γ
+22 kb
AFLP
(G)
ld14
+35 kb
880 bp
(H)
l(1)G0007
+66 kb
459 bp
(I)
baz
+149 kb
768 bp
?
10 kb
tan-5'
tan-3'
A01
1
2 3
4 5 6 7 8
500 bp


B CG32687
99
N04
N14
N08
DN12
A01
DN4
89
PM
SB
0.001


C HMR-tan
A09
A01
SB
LA
OR
IR
A00
DN12
AA
DN2
DN4
PM
N08
99
N04
N14
0.001


D tan-5'
97
LA
DN2
DN4
A00
DN12
A01
OR
PM
AA
IR
SB
A09
N08
100
N04
97
N14
0.001


E tan-3'
100
A00
DN12
DN2
PM
AA
IR
A09
DN4
LA
OR
SB
A01
100
N08
92
N04
98
N14
0.002

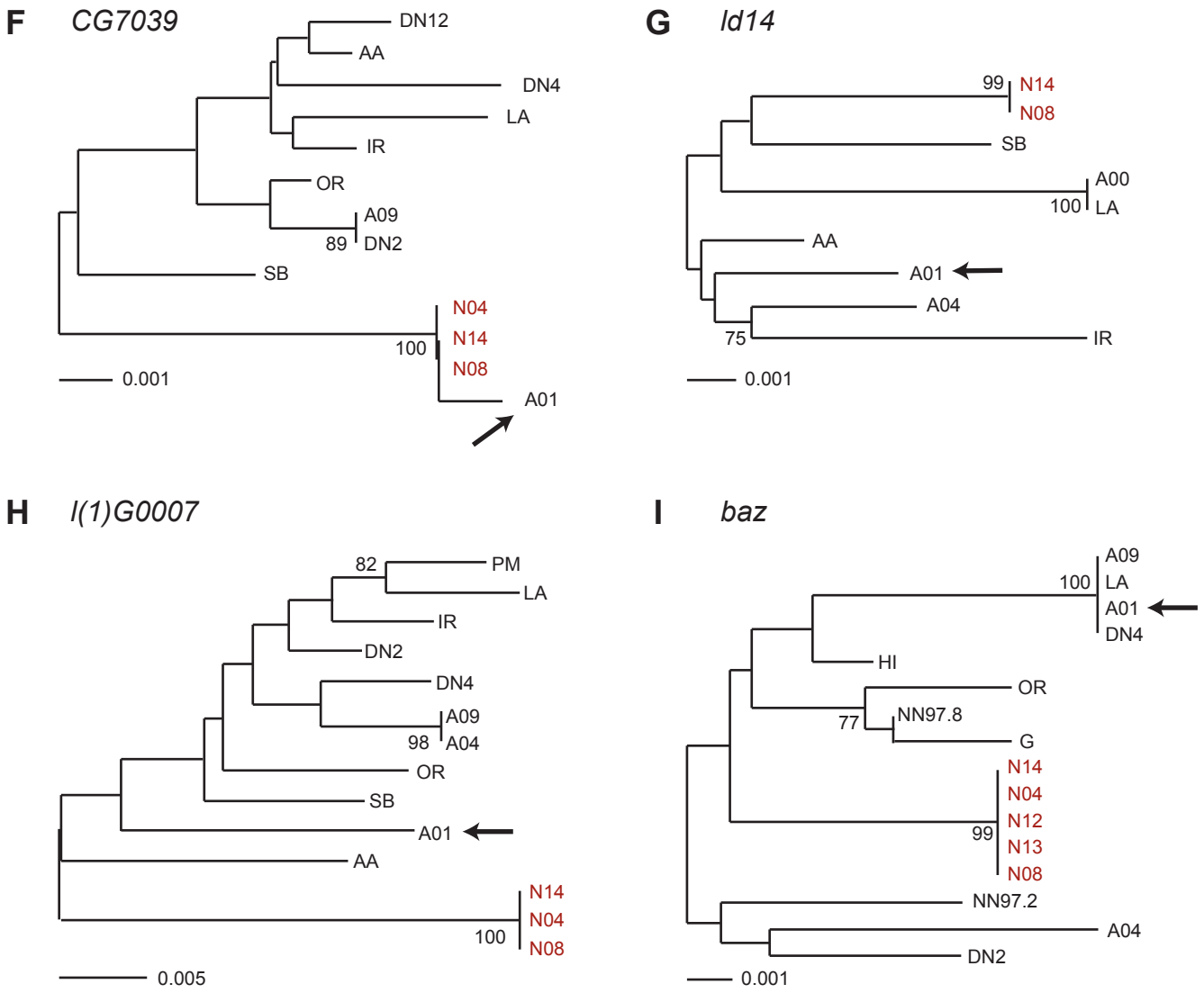


Fig. S8. Haplotype sharing between *D. novamexicana* and the A01 line of *D. americana* is limited to sequences within and near *tan*. To determine the extent of haplotype sharing, we sequenced multiple loci in and around *tan* (A) from multiple lines of *D. americana* and *D. novamexicana* and constructed a neighbor joining tree for each locus (B - I). Nodes appearing in >75% of 1000 replicate bootstrap trees are indicated by their exact frequency, and branch lengths are drawn to scale. Arrows indicate the location of A01 within each tree, and all *D. novamexicana* alleles are shown in red. (A) A summary of these neighbor joining trees is shown. Loci for which the A01 allele was most similar to other alleles of *D. americana* are shown in brown, and loci for which the A01 allele was most similar to *D. novamexicana* alleles are shown in yellow. The gene containing (or adjacent to) each locus is indicated with its position, in kilobases (kb), relative to the start of the *tan* gene (i.e., *D. virilis*, scaffold_12932: 1,852,445). The number of aligned basepairs (bp) analyzed for each locus (excluding gaps) is also indicated. The one exception is the *AP-1 Gamma* (*AP-1*) locus, which was genotyped with an amplified fragment length polymorphism. 840 bp of sequence (GQ457370-GQ457381) from the region shown in grey (indicated by the question mark) was also analyzed, but did not contain enough fixed differences to distinguish the *D. americana* and *D. novamexicana* alleles with bootstrap support greater than 75%. A larger schematic of the *tan* region, including exons 1 – 8 (black boxes), is also shown in panel A, with vertical lines indicating the positions of fixed differences between species GQ457336-GQ457353. Sequence of the A01 allele is indicated at each of these divergent sites, with brown lines representing *D. americana* alleles and yellow lines representing *D. novamexicana* alleles. On the basis of the pattern of haplotype sharing, this region was divided into two sections (*tan*-5' and *tan*-3') for phylogenetic analysis.

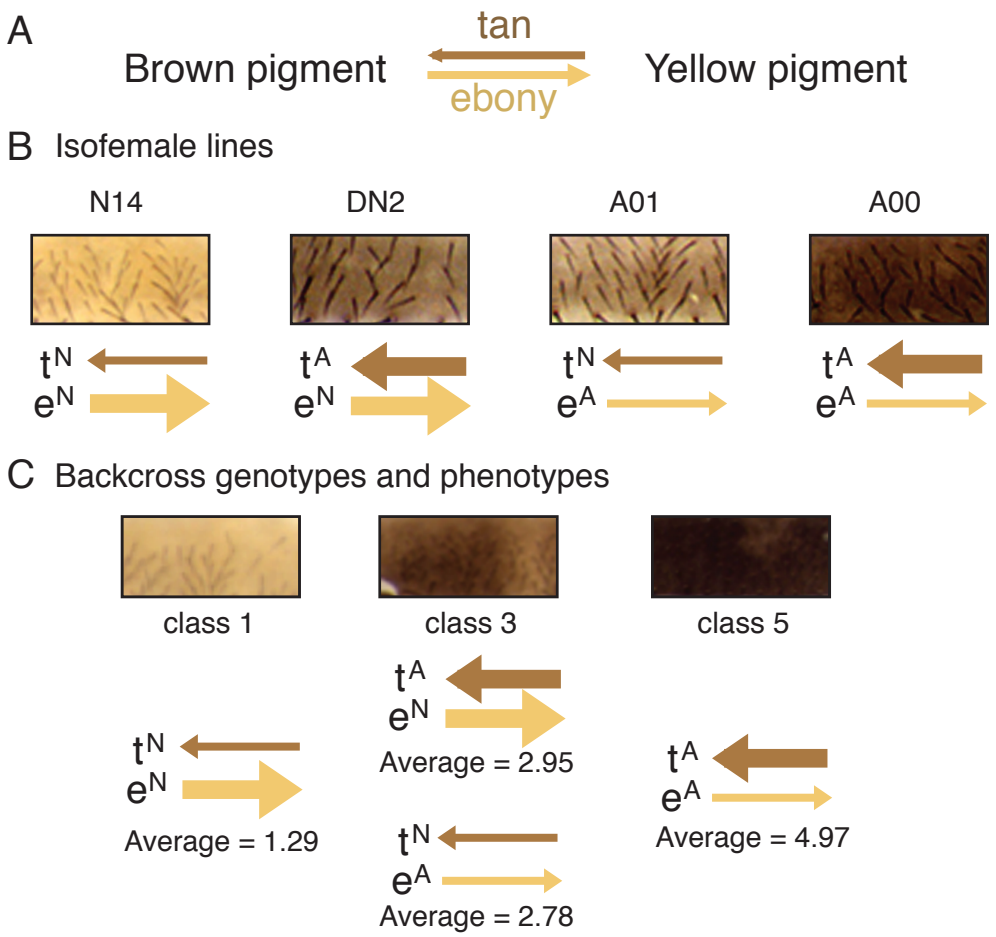


Fig. S9. A genetic model of pigmentation polymorphism and divergence. (A) *ebony* and *tan* control the production of yellow and brown pigments, respectively (Fig. 3A), and the relative expression of these proteins determines adult pigmentation (Fig. 3F). Tan is expressed at higher levels in *D. americana* (Fig. S3D), and Ebony is expressed at higher levels in *D. novamexicana* (S17). To reconcile the genotypes and phenotypes observed in this study, we propose a model of pigmentation evolution in which Tan and Ebony expression differences between species are caused by changes in the *tan* and *ebony* genes themselves. Specifically, we assume that the *D. americana tan* allele (t^A) causes greater Tan protein expression than the *D. novamexicana* allele (t^N), and that the *D. novamexicana ebony* allele (e^N) causes greater Ebony protein expression than the *D. americana* allele (e^A). For simplicity, we ignore the contribution of other genes, which in the DN4 backcross population explained up to 24% of the pigmentation variance. (B) *ebony* and *tan* genotypes are shown for isofemale lines of *D. americana* (DN2, A01, and A00) and *D. novamexicana* (N14). The thicker yellow arrow represents greater activity of e^N relative to e^A, and the thicker brown arrow represents greater activity of t^A relative to t^N. In N14, which carries t^N and e^N, our model predicts a net production of yellow pigment. In A00, which carries t^A and e^A, our model predicts a net production of brown pigment. In DN2 and A01, both of which carry one allele from each species, our model predicts pigmentation intermediate between A00 and N14. Dorsal abdominal cuticle from the A4 segment is shown for 7-10 day old males. (C) The four possible *tan* and *ebony* genotypes are shown along with the average pigmentation class for each genotype in the DN4 backcross population. Dorsal abdominal cuticle from the A4 segment of male flies categorized as class 1, class 3, and class 5 in the DN4 backcross are also shown. Flies inheriting e^N and t^N had light pigmentation, presumably because of excess *ebony* activity, while flies inheriting e^A and t^A had dark pigmentation, presumably because of excess *tan* activity. Flies inheriting the *D. americana* allele of only *ebony* or *tan* had intermediate pigmentation, presumably because of more balanced *tan* and *ebony* activities. The distribution of phenotypes in the DN4 backcross population is consistent with a simple two-locus Mendelian model assuming that class 1 contains $t^N;e^N$ flies, classes 2, 3, and 4 contain $t^N;e^A$ and $t^A;e^N$ flies, and class 5 contains $t^A;e^A$ flies ($\chi 2 = 4.7$, df = 2, P = 0.1).

Table S1. Prior studies comparing intra- and interspecific QTL

	Intraspecific			Interspecific			
Trait	**Species**	**# QTLs**	**Ref.**	**Species**	**# QTLs**	**Ref.**	**Overlap[1]**
sex comb tooth number	*Drosophila melanogaster*	2	(*S18*)	*D. simulans* and *D. mauritiana*	>=5	(*S19*)[2]	1
male courtship song	*Drosophila melanogaster*	3	(*S20*)	*D. simulans* and *D. sechellia*	6	(*S21*)	0
abdominal pigmentation	*Drosophila melanogaster*	3	(*S22*)	*D. melanogaster* and *D. willistoni*	N/A[3]	(*S23*)	1
various[4]	*Helianthus petiolaris*	28[5]	(*S24*)	*H. annuus* and *H. petiolaris*	72	(*S24*)	28
sex comb tooth number	*Drosophila simulans*	7	(*S25*)	*D. simulans* and *D. mauritiana*	2	(*S26*)	2
floral traits	*Mimulus guttatus*	16	(*S27*)	*M. guttatus* and *M. nasutus*	24	(*S28*)	11

[1] The number of QTLs that map to the same genomic region within and between species is shown.

[2] Interspecific QTLs were identified on each chromosome in this low-resolution study, providing a minimum estimate of 5 QTLs. True et al. (*S26*) refined this work, but only identified two QTLs, both on the third chromosome, and neither of which overlap the intraspecific QTL identified in (*S18*). The inconsistency between studies may also reflect intraspecific variation within *D. simulans* and/or *D. mauritiana*.

[3] Interspecific divergence of *bric-a-brac* (located within one of the intra-specific QTL) was demonstrated using transgenic assays.

[4] This study includes 40 different morphological, physiological, and life history traits.

[5] Intraspecific QTL were not mapped genome wide; rather, 72 interspecific QTL were directly tested for the presence of intraspecific QTL.

Table S2. Lines of *D. americana* and *D. novamexicana* used in this work

Species	Line	Full ID	Source	Collection Site	Collection Year
D. americana	A00	15010-0951.00	Drosophila Species Stock Center	Anderson, IN	unknown
D. americana	A01	15010-0951.01	Drosophila Species Stock Center	Poplar, MT	1947
D. americana	A04	15010-0951.04	Drosophila Species Stock Center	Lake Champlain, VT	1948
D. americana	A09	15010-0951.09	Drosophila Species Stock Center	Myrtle Beach, SC	1961
D. americana	AA	AA 99.6	B. McAllister (*S29*)	Augusta, AR	1999
D. americana	DN12	DN 01.2	B. McAllister (*S30*)	Duncan, NE	2001
D. americana	DN2	DN 00.2	B. McAllister (*S30*)	Duncan, NE	2000
D. americana	DN4	DN 00.4	B. McAllister (*S30*)	Duncan, NE	2000
D. americana	G	G96.21	B. McAllister (*S30*)	Gary, IN	1996
D. americana	IR	IR 03.10	B. McAllister (*S31*)	Iowa River, IA	2003
D. americana	LA	LA 99.48	B. McAllister (*S29*)	Lake Ashbaugh, AR	1999
D. americana	NN97.2	NN 97.2	B. McAllister (*S30*)	Niobrara, NE	1997
D. americana	NN97.8	NN97.8	B. McAllister (*S30*)	Niobrara, NE	1997
D. americana	OR	OR 01.52	B. McAllister (*S32*)	Ottawa, OH	2001
D. americana	PM	PM 99.28	B. McAllister (*S29*)	Puxico, MO	1999
D. americana	SB	SB 02.06	B. McAllister (*S31*)	Saulsbury, IA	2002
D. novamexicana	N04	15010-1031.04	Drosophila Species Stock Center	Moab, UT	1949
D. novamexicana	N08	15010-1031.08	Drosophila Species Stock Center	San Antonio, NM	1947
D. novamexicana	N12	15010-1031.12	Drosophila Species Stock Center	Antlers, CO	1949
D. novamexicana	N13	15010-1031.13	Drosophila Species Stock Center	Patagonia, AZ	1953
D. novamexicana	N14	15010-1031.14	Drosophila Species Stock Center	Moab, UT	1949

Table S3. Primers used for phylogenetic genetic analysis

Gene name	Primer name	Primer sequence	Accession # containing resulting sequence
tan	39	CGAACCGCAACTGATATTGA	GQ457336-GQ457353
tan	91	TAGTGAGTGCCACGTGTATAGAGAACG	
tan	204	AAGCTAGGCAAACGGCATGC	
tan	205	ACAATTTCGAATTCGATGAGC	
tan	226	GGCGCTCTTCAATGAGCCAAACAA	
tan	94	TCAGTTTGAATTCTGCCTTCAAGCGCT	
tan	40	GTTGTTGTTGTTGGGGGTTC	
tan	90	CGTTCTCTATACACGTGGCACTCACTA	
tan	204	AAGCTAGGCAAACGGCATGC	
tan	205	ACAATTTCGAATTCGATGAGC	
tan	191	CCCTTACCCACTTTCTATGG	
tan	229	TCACGACTGATCGACAGGGCAAAC	
tan	227	CCCGCGCACATAATTAACAAGCTG	
tan	89	GCGACTTTGGCCTTAGCTTC	
tan	69	GCCGAGGTGGAGTTCCA	
tan	81	GGTATTCAATCTTCGGCGTGCCAAA	
tan	71	CACACCTTTCAAAAGAT	
tan	85	CGCTGACCAGAATTTCAAATTTAATTGCC	
tan	87	GACCATTTAATGGTGCTCAAAATATGG	
tan	206	GGAATGCCTTTTACTGCACATAATG	
tan	230	CCGGCATAGCAGAGCGACATGAA	
tan	86	CCATATTTTGAGCACCATTAAATGGTC	
tan	228	GCCTTGACGGAGACGCTTATTCAT	
tan	41	TACTTTGTGGTTGCGCACAT	
tan	42	ATTTGTAAAGCAGGGGCAAC	
tan	180	CGCATACACTTGGACCAGGCC	
tan	231	TTGGCTTCGTTCTAACGGGCATCA	
tan	246	CATTAAGCTTTGTAGCTGACTATG	
tan	193	CACACACGCGAATTAGGCAAAAG	
ld14	215	GCGTGCAGTGTCTGTTAGCAG	GQ457425-GQ457433
ld14	216	AGGCCACGCCCACTAACTAAC	
l(1)G0007	217	GCGCGTCTCCGATGAGATGG	GQ457434-GQ457447
l(1)G0007	218	AATACCATCGGTCATGTACTTGATGAC	
CG32687	234	CCAGCGAAAGCACATGCAG	GQ457401-GQ457408
CG32687	235	GGATATCTGGAAGATGCAAAG	
baz	61	CGGTGGCATACATCAGAATG	GQ457409-GQ457424
baz	62	CGACTGCAAAAGCGTACAAA	
CG7039-AP1gamma	238	CAAATGGCTGGTGGAGGC	GQ457388-GQ457400
CG7039-AP1gamma	239	CCTGCTCATAATACGGATAC	

ebony	45	AATTACCCAACTGCGACTGG	GQ457354-GQ457369
ebony	46	CGCCCTCCATCTTCAGATAC	
ebony	130	CGCTCCCTGCTCATGTATCT	
ebony	131	GGCGACGTTCTTCTCAACCT	
ebony	e4f	CGTGTGCGCCCAATTAGC	
ebony	e4r	TGCTTAGATTTCACCTCATCAACAGAA	
HMR[1]	142	CATCTCGTCCAACTCCAGGT	GQ457448-GQ457453
HMR[1]	144	GCGCTATAAATATCAGCGTCA	
CG7039[1]	145	CATTGCTGCACGGCTTTTAC	GQ457382-GQ457387
CG7039[1]	146	CTCCACCAGCCATTTGATG	
AP1gamma[1]	213	GGAAATGTGCGAGAACAGTTCGG	GQ457370-GQ457381
AP1gamma[1]	214	GATTTTCACCTGCAGGAACGGATC	

[1] Phylogenetic analysis of these sequences was not included in Fig. S6 because they lacked either sufficient sequences or variable sites and/or were redundant with other loci for the purposes of this study. However, for completeness, they have been submitted to GenBank and are therefore included in this table.

Supplemental References

S1. D. G. Consortium *et al.*, *Nature* **450**, 203 (2007).
S2. G. B. Gloor *et al.*, *Genetics* **135**, 81 (1993).
S3. P. Mena, B. McAllister. Personal communication.
S4. D. S. Falconer, T. F. C. Mackay, *Introduction to Quantitative Genetics* (Longman Group Ltd., Essex, UK, ed. 4th, 1996).
S5. S. Jeong *et al.*, *Cell* **132**, 783 (2008).
S6. J. R. True *et al.*, *PLoS Genet.* **1**, e63 (2005).
S7. P. J. Wittkopp, B. K. Haerum, A. G. Clark, *Nature* **430**, 85 (2004).
S8. P. J. Wittkopp, B. K. Haerum, A. G. Clark, *Nat Genet* **40**, 346 (2008).
S9. C. Horn, E. A. Wimmer, *Dev. Genes Evol.* **210**, 630 (2000).
S10. A. C. Groth, M. Fish, R. Nusse, M. P. Calos, *Genetics* **166**, 1775 (2004).
S11. S. Warming, N. Costantino, D. L. Court, N. A. Jenkins, N. G. Copeland, *Nucleic Acids Res* **33**, e36 (2005).
S12. M. Markstein, C. Pitsouli, C. Villalta, S. E. Celniker, N. Perrimon, *Nat Genet* **40**, 476 (2008).
S13. R. C. Edgar, *Nucleic Acids Res* **32**, 1792 (2004).
S14. S. Kumar, K. Tamura, M. Nei, *Brief Bioinformatics* **5**, 150 (2004).
S15. J. Rozas, J. C. Sánchez-DelBarrio, X. Messeguer, R. Rozas, *Bioinformatics* **19**, 2496 (2003).
S16. J. Hey, J. Wakeley, *Genetics* **145**, 833 (1997).
S17. P. J. Wittkopp, B. L. Williams, J. E. Selegue, S. B. Carroll, *PNAS* **100**, 1808 (2003).
S18. S. V. Nuzhdin, S. G. Reiwitch, *Heredity* **84**, 97 (2000).
S19. J. A. Coyne, *Genet Res* **46**, 169 (1985).
S20. J. M. Gleason, S. V. Nuzhdin, M. G. Ritchie, *Heredity* **89**, 1 (2002).
S21. J. M. Gleason, M. G. Ritchie, *Genetics* **166**, 1303 (2004).
S22. A. Kopp, R. M. Graze, S. Xu, S. B. Carroll, S. V. Nuzhdin, *Genetics* **163**, 771 (2003).
S23. T. M. Williams *et al.*, *Cell* **134**, 610 (2008).
S24. C. Lexer, D. M. Rosenthal, O. Raymond, L. A. Donovan, L. Rieseberg, *Genetics* **169**, 2225 (2005).
S25. H. Tatsuta, T. Takano-Shimizu, *Genet Res* **87**, 93 (2006).
S26. J. R. True, J. Liu, L. F. Stam, Z. B. Zeng, C. C. Laurie, *Evolution* **51**, 816 (1997).
S27. M. C. Hall, C. J. Basten, J. H. Willis, *Genetics* **172**, 1829 (2006).
S28. L. Fishman, A. J. Kelly, J. H. Willis, *Evolution* **56**, 2138 (2002).
S29. B. F. McAllister, *Genome* **45**, 13 (2002).
S30. B. F. McAllister, *Genetics* **165**, 1317 (2003).
S31. B. F. McAllister, A. L. Evans, *PLoS ONE* **1**, e112 (2006).
S32. B. F. McAllister, *Dros. Inf. Serv.* **84**, 227 (2001).