

A Single Gene Causes an Interspecific Difference in Pigmentation in *Drosophila*

Yasir H. Ahmed-Braimah*¹ and Andrea L. Sweigart[†]

*Department of Biology, University of Rochester, Rochester, New York 14627 and [†]Department of Genetics, University of Georgia, Athens, Georgia 30602

ABSTRACT The genetic basis of species differences remains understudied. Studies in insects have contributed significantly to our understanding of morphological evolution. Pigmentation traits in particular have received a great deal of attention and several genes in the insect pigmentation pathway have been implicated in inter- and intraspecific differences. Nonetheless, much remains unknown about many of the genes in this pathway and their potential role in understudied taxa. Here we genetically analyze the puparium color difference between members of the virilis group of *Drosophila*. The puparium of *Drosophila virilis* is black, while those of *D. americana*, *D. novamexicana*, and *D. lummei* are brown. We used a series of backcross hybrid populations between *D. americana* and *D. virilis* to map the genomic interval responsible for the difference between this species pair. First, we show that the pupal case color difference is caused by a single Mendelizing factor, which we ultimately map to an ~11-kb region on chromosome 5. The mapped interval includes only the first exon and regulatory region(s) of the dopamine *N*-acetyltransferase gene (*Dat*). This gene encodes an enzyme that is known to play a part in the insect pigmentation pathway. Second, we show that this gene is highly expressed at the onset of pupation in light brown taxa (*D. americana* and *D. novamexicana*) relative to *D. virilis*, but not in the dark brown *D. lummei*. Finally, we examine the role of *Dat* in *adult* pigmentation between *D. americana* (heavily melanized) and *D. novamexicana* (lightly melanized) and find no discernible effect of this gene in adults. Our results demonstrate that a single gene is entirely or almost entirely responsible for a morphological difference between species.

KEYWORDS morphological evolution; pigmentation; genetics of species differences

UNDERSTANDING the genetic basis of morphological differences between species is a central goal of evolutionary biology. A key question focuses on complexity: Do species differ by many genes of small effect or by a few genes of large effect (Haldane 1937)? The answer to this question has remained elusive; cases studied reveal examples of each scenario (Orr 2001). This is not surprising, however, as traits and species may differ in divergence rates and types of selection (or lack thereof). The question can best be stated as one of relative frequency: How often do major genes cause morphological differences between species (Orr 2001)?

Insect pigmentation traits have received much attention from geneticists (Wittkopp *et al.* 2003a; True 2003; Wittkopp

and Beldade 2009). These studies benefited from the fact that the pathway that determines insect cuticular pigment is both conserved and well understood (Wittkopp *et al.* 2002; Wittkopp *et al.* 2003a; Jeong *et al.* 2008; Williams *et al.* 2008; Wittkopp *et al.* 2009; Werner *et al.* 2010). Several genes in the pigmentation pathway have been implicated in pigmentation differences between and within species, particularly in *Drosophila* (Wittkopp *et al.* 2003b; Takahashi *et al.* 2007; Jeong *et al.* 2008; Telonis-Scott *et al.* 2011; Bastide *et al.* 2013). Furthermore, in cases where the molecular basis of a phenotypic difference has been determined, the results have usually shown that regulatory mutations cause differences in expression between the alternative alleles (Wittkopp *et al.* 2003b; Pool and Aquadro 2007; Jeong *et al.* 2008; Werner *et al.* 2010; Takahashi and Takano-Shimizu 2011; Arnoult *et al.* 2013). As the pathway is currently understood, the interplay between a handful of enzymes (*e.g.*, Ebony, Yellow, Tan), which act on tyrosine and its derivatives such as dopa and dopamine, and a set of regulatory proteins (*e.g.*, Bab) determine the amount and spatial distribution of cuticular pigment (Wittkopp *et al.* 2002;

Copyright © 2015 by the Genetics Society of America

doi: 10.1534/genetics.115.174920

Manuscript received January 27, 2015; accepted for publication March 11, 2015; published Early Online March 13, 2015.

Supporting information is available online at <http://www.genetics.org/lookup/suppl/doi:10.1534/genetics.115.174920/-/DC1>.

¹Corresponding author: Department of Biology, 340 Hutchinson Hall, University of Rochester, RC Box 270211, Rochester, NY 14627. E-mail: yasir.ahmed@rochester.edu

Wittkopp *et al.* 2003a). Variation in color pattern and pigment observed among insects, therefore, likely reflects evolutionary changes in the genes belonging to this pathway (Wittkopp *et al.* 2003a).

The virilis group of *Drosophila* provides a good model for analyzing the genetic basis of morphological differences among closely related species (Spicer 1991; Wittkopp *et al.* 2010; Fonseca *et al.* 2013). The virilis group contains two phylads, the montana phylad and the virilis phylad. The virilis phylad—the focus of this study—includes *Drosophila virilis*, *D. lummei*, *D. novamexicana*, and *D. americana*. Several studies have examined the phylogenetic history and reproductive incompatibilities among members of this group (Orr and Coyne 1989; Caletka and McAllister 2004; Sweigart 2010a,b; Morales-Hojas *et al.* 2011; Sagga and Civetta 2011; Ahmed-Braimah and McAllister 2012). The sequenced genome of *D. virilis*, coupled with the relative ease of producing backcross and advanced generation hybrids in species crosses, allows for easy genetic marker development and mapping of genomic regions underlying trait differences. This group, however, contains many segregating and fixed chromosomal inversions, which sometimes render fine-scale genetic mapping in parts of the genome difficult or impossible (Hsu 1952).

Species belonging to the virilis phylad differ in several morphological characters (Patterson *et al.* 1940; Spencer 1940; Stalker 1942). Perhaps best known, adults of *D. novamexicana* differ in color from their sister species. While all other members of the group are darkly pigmented, *D. novamexicana* has evolved a light brown color along its dorsal abdomen, head, and thorax. *D. novamexicana* also lacks pigment along the abdominal dorsal midline (Spicer 1991; Wittkopp *et al.* 2003b). Previous work has shown that *ebony* plays a major role in this pigment difference (Wittkopp *et al.* 2003b) and that *tan* also contributes substantially (Wittkopp *et al.* 2009).

Species belonging to the virilis phylad also differ in pupal case color (Supporting Information, Figure S1)(Stalker 1942). All members of the group (and including montana phylad species) have brown puparia. *D. virilis*, on the other hand, lacks brown pigment and has black puparia that are easily distinguished from the sister species early in pupal development. Although this trait appears fixed in *D. virilis*, the intensity of the black color varies between strains and is also affected by crowding and/or culture conditions (Spencer 1940; Stalker 1942). For example, uncrowded pupae appear darker than crowded ones, which appear tannish gray.

Previous genetic analyses of the pupal color difference between *D. americana* and *D. virilis* conducted during the early 1940s yielded mixed results. Early work by W. P. Spencer suggested a complex genetic basis of this trait, while his later work did not support this conclusion (Spencer 1940). Spencer attributed the discrepancies in his findings to the possibility that pupal color may behave differently in different strains. Interestingly, Spencer's original findings on this trait played an important role in shaping H. J. Muller's "multi-genic" view of morphological evolution (Muller 1940; Orr and Coyne 1992). Around the time that Spencer conducted his experiments, H. D.

Stalker performed an independent genetic analysis of the pupal color trait between *D. americana* and *D. virilis* and attributed the difference to a large effect on chromosome 5 (Muller element C) with supplementary genes for "brownness" carried on the *D. americana* chromosome 2–3 (Muller elements D–E) (Stalker 1942). His results differed from those of Patterson *et al.* (1940), who suggested that the difference is entirely explained by chromosomes 2–3. The latter authors based their conclusions on cytological data, but provided no detailed account of their findings.

Here we present a genetic analysis of the pupal case color difference between two species of the virilis group, *D. americana* and *D. virilis*. We first map this difference at the level of whole chromosomes and find that only chromosome 5 causes the difference. We ultimately identify a single genomic interval (~11 kb) on chromosome 5 that causes the pupal case color difference. This region contains only the first exon and regulatory region of a single gene, *GJ20215*. *GJ20215* is the homolog of the *D. melanogaster* dopamine *N*-acetyltransferase (*Dat*). *Dat* (also known as arylalkylamine *N*-acetyltransferase, or *aanAT*) is an enzyme known to act within the pigmentation pathway. In particular, *Dat* catalyzes the reaction from dopamine (DA) to *N*-acetyl dopamine (NADA). We then examine the expression differences between pupae of *D. americana* and *D. virilis*, their hybrids, as well as among members of the virilis phylad for *Dat*. Finally, we test the role of this gene in the adult pigmentation difference between *D. americana* and *D. novamexicana*. We conclude that reduced expression of *Dat* in early pupal development in *D. virilis* is the cause of dark pupae in this species.

Materials and Methods

Fly strains and husbandry

Flies were maintained at a constant temperature (22°) in a ~12-hr day/night cycle on standard cornmeal medium. The *D. virilis* strain used (15010–1051.31) was obtained from the University of California—San Diego (UCSD) *Drosophila* Species Stock Center (<http://stockcenter.ucsd.edu>) and carries two visible mutations on the fifth chromosome: Branched (*B*) and Scarlett (*st*). Both mutations behave recessively in a *D. americana* genetic background, although *B* is considered dominant within *D. virilis* strains. (*B* causes abnormal branching of wing veins, and *st* mutants have bright red eyes.) The *D. americana* strain (SB02.06) was collected by Dr. Bryant F. McAllister (University of Iowa) in 2002 near the Cedar River in Muscatine County, Iowa. While *D. americana* harbors a number of chromosomal inversions and two chromosomal fusions, the strain used here is known to differ from *D. virilis* in a small inversion on the telomeric end of chromosome 5 (In5a) and a large inversion encompassing the centromeric half of chromosome 2. Importantly it lacks the large In5b inversion that segregates in western populations of *D. americana* and that affects ~60% of the chromosome. In addition to the fixed fusion of chromosomes 2 and 3, the strain is also fixed for the clinally distributed fusion between chromosomes 4

and X, both of which carry fixed inversion differences. The *D. novamexicana* (15010–1031.04) and *D. lummei* (LM.08) strains were obtained from Dr. Bryant F. McAllister. For all crosses in this study, male and female flies were collected within 2 days of eclosion and reared separately until they reached sexual maturity and crossed at 12–14 days old.

Molecular genotyping

Up to 42 microsatellite markers and 6 single nucleotide polymorphism (SNP) markers were used to genotype recombinant individuals along chromosome 5 (Table S1; see Figure S3 for arrangement of genome scaffolds on chromosome 5). Tandem repeat regions were identified in the *D. virilis* genome (Release droVir2, <http://hgdownload.cse.ucsc.edu/downloads.html#droVir>) using Tandem Repeat Finder (Benson 1999), and primers flanking repeat regions were designed using Primer3 (Rozen and Skaletsky 2000). SNPs were identified by sequencing small orthologous fragments of *D. virilis* and *D. americana* and identifying base substitutions and/or insertions/deletions unique to either strain used in the study. Genomic DNA was obtained from each individual whole fly using the protocol of Gloor and Engels (1992). Microsatellite markers and genomic regions containing SNP and/or insertion/deletion markers were amplified and sized/sequenced on an ABI-3700 automated capillary sequencer (Applied Biosystems).

Whole-chromosome association backcross

The backcross population of pupae segregating for whole chromosomes was generated by backcrossing “AV” F1 males (A, *D. americana*; V, *D. virilis*; genotype of mother in crosses given first hereafter) to *D. virilis* females. As the *D. virilis* allele for puparium color is recessive (Stalker 1942), this cross produces a population of pupae that are either black (homozygous for *D. virilis* at the pupal case color locus, $n = 81$) or light brown (heterozygous at the pupal case color locus, $n = 84$). To infer the whole-chromosome genotypes of each individual, pupae were genotyped at a single microsatellite marker for each of the three autosomal linkage groups (chromosomes X-4:SSR21, chromosomes 2–3:SSR37, and chromosome 5:SSR146). Chromosome 6 (the small “dot” chromosome, Muller element F) was not surveyed.

Recombinant mapping of the pupal case color locus on chromosome 5

Recombinant backcross males [(AV)A] were generated by backcrossing AV F1 females to *D. americana* males. As all backcross progeny possessed the light brown pupal case color (caused by the dominant *D. americana* allele), the pupal case color alleles inherited by (AV)A males can be assessed only by crossing these males to *D. virilis* females; (AV)A males heterozygous at the pupal case color locus will sire light brown and black pupae in equal proportions whereas males homozygous for *D. americana* alleles will produce only light brown pupae. A total of 644 recombinant backcross males that sired progeny comprised the entire mapping population; however, only a subset of individuals was genotyped along the majority

of the recombining segment of the chromosome. This subset is composed of 46 individuals and was selected on the basis of the genotype at the visible marker, *B*, and whether that individual sired five or more progeny (only *D. americana* homozygotes at *B* were selected). In addition, a second subset of 24 individuals was selected on the basis of known recombination events in the vicinity of the pupal case locus from past genotyping efforts in the laboratory. The 70 individuals were genotyped at up to 42 microsatellite and SNP markers along a 14.5-Mb region that encompasses ~70% of the recombining portion of the chromosome.

Screening for recombinants in the candidate region

To screen for additional recombinants in the 27-kb candidate region, a large collection ($n \sim 30,000$) of recombinant backcross pupae was generated by backcrossing AV F1 females to *D. virilis* males. Pupae were separated by color (light brown or black) and allowed to eclose. The phenotype of eclosing flies at the two visible markers (*st* and *B*) was recorded and only flies that were recombinants between *st* and *B* (~30 cM) were frozen for molecular genotyping ($n \approx 12,000$). Recombinants between *st* and *B* were genotyped at two microsatellite markers (SSR169 and SSR173), which are ~32 kb apart and flank the pupal case color locus. A total of 16 recombinants were recovered between SSR169 and SSR173 and were subsequently genotyped at 10 additional markers (6 SNP, 4 microsatellite) spanning the 32-kb region.

Multiple sequence alignment of candidate region in *D. americana* and *D. virilis* strains

Publicly available genome sequences for two *D. americana* strains (H5 and W11, <http://cracs.fc.up.pt/~nf/dame/fasta/>; Fonseca *et al.* 2013) and two *D. virilis* strains (Str9:SRX496597, and Str160:SRX496709; Blumenstiel 2014) were used to compare the candidate region between the two species. To extract orthologous regions in the *D. americana* strains, the sequence of the ~11-kb candidate region from the *D. virilis* genome strain (droVir2) was Blasted (BLAST 2.2.28+) against the H5 and W11 genome assemblies. Matching scaffolds for each *D. americana* strain were concatenated and trimmed accordingly (Geneious R8). Sequence reads for the two *D. virilis* strains were obtained from the Sequence Read Archives (SRA) and mapped to the *D. virilis* genome (bwa v.0.7.9a). The consensus sequence from the mapped reads of the ~11-kb candidate region for each *D. virilis* strain was obtained through “pileup” of consensus bases and converted to FASTA sequences (Samtools v. 0.1.19). Finally, the ~11-kb region from all four strains and the *D. virilis* genome strain were aligned (MUSCLE v. 3.8.31).

Quantitative RT-PCR

The ortholog of *GJ20215* in *D. melanogaster* is known to have two isoforms, yet the annotated *GJ20215* consisted of only one isoform (hereafter isoform A, FlyBase v. 2014_04). To determine the gene structure of *GJ20215*, publicly available *D. virilis* mRNA short-sequence reads were obtained from the

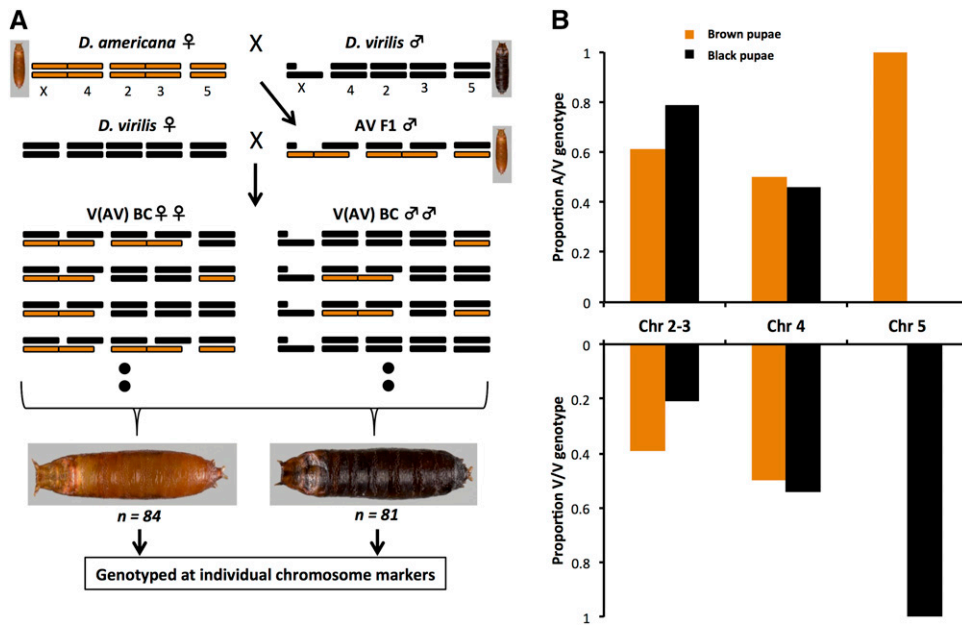


Figure 1 (A) Crossing scheme to generate backcross pupae segregating for whole chromosomes. *D. americana* (A) and *D. virilis* (V) chromosomes are shown in brown and black, respectively. Female genotypes are shown on the left and male genotypes on the right. For each hybrid the genotype notation is abbreviated and the maternal allele is indicated first. The pupal color phenotype for pure species and hybrids is shown next to the pure species and F1 karyotypes. Eight possible genotypes (four female and four male) are illustrated for backcross hybrids. (B) Associations between chromosomal genotypes and pupal color phenotype among whole-chromosome backcross hybrids. The proportion of the two alternative whole-chromosome genotypes for red/brown and black backcross pupae is plotted for each autosomal linkage group. The sum of AVV and VV genotypes for each phenotypic class within each chromosome should be 1.

modENCODE project (<http://www.modencode.org>). Reads were processed and mapped to the *D. virilis* genome using Tophat (Trapnell *et al.* 2009) and transcripts assembled using Cufflinks (Trapnell *et al.* 2010). The gene structure of *GJ20215* was found to contain a second isoform (isoform B) that utilizes an alternative, previously unknown first exon which lies ~12.5 kb upstream of the first exon of isoform A. Two different isoform-specific forward primers that overlap the exon–exon boundary, in addition to a shared reverse primer, were designed as described above. Primers were also designed for a reference gene, *RpS3*. Total RNA was extracted from each virilis group strain and AV F1 hybrids (~10 individual pupa/replicate) at three different stages during pupal development: (1) at the onset of cuticle hardening but prior to the development of any pigment, (2) after the development of light pigment (~6–8 hr after cuticle hardening), and (3) after full pigment development (~12–16 hr after cuticle hardening). Approximately 4 μ g of RNA from each sample was used to synthesize cDNA that was used in SYBR-green quantitative RT–PCR (qRT–PCR) reactions (Qiagen). Relative quantitation was performed using the $\Delta\Delta$ Ct method (Livak and Schmittgen 2001) with *RpS3* as the reference gene and *D. americana* as the control sample. Statistically significant differences in relative expression were assessed by performing a two-sample *t*-test (R v. 3.1.1).

***D. americana*/*D. novamexicana* adult pigmentation analysis and genotyping**

Sixth generation (F6) intercross hybrids between *D. americana* and *D. novamexicana* were produced by crossing *D. americana* females to *D. novamexicana* males *en masse* and subsequently allowing hybrids to intercross for five generations. F6 individuals were separated by sex soon after eclosion and aged for 14 days. A total of 188 flies (93 females, 95 males) were imaged

dorsally before being flash frozen for DNA extraction. Red–blue–green (RBG) measurements were obtained (ImajeJ) from three landmarks: the scutellum, the anterior portion of the thorax, and the dorsal abdominal cuticle. Only red measurements were used in the analyses as they showed the greatest difference. A microsatellite marker (for *tan*) and two sequenced fragments containing multiple SNP differences (for *ebony* and *GJ20215*) were developed to genotype F6 individuals as described above. The data were analyzed by performing a one-way analysis of variance (ANOVA) for each of the three markers separately where the genotype at a given marker is the independent variable and red mean value as the dependent variable (R v.3.1.1).

Results

Mapping to whole chromosomes

The *D. virilis* pupal case color allele(s) is completely recessive (Spencer 1940; Stalker 1942). Reciprocal F1 hybrid pupae between the two species are identical in pupal color to *D. americana*. In addition, backcross hybrids are either black (*D. virilis* homozygotes at pupal color locus) or brown (heterozygotes), with no observable gradation between these two classes. These observations suggest that the factor(s) causing the difference in pupal color in this species pair reside(s) on autosomes.

To map this phenotypic difference to whole autosomes, we generated a backcross population of pupae by crossing AV F1 hybrid males to *D. virilis* females (Figure 1A). As meiotic recombination does not occur in *Drosophila* males, whole chromosomes remain intact and species origin can be identified with single markers per chromosome.

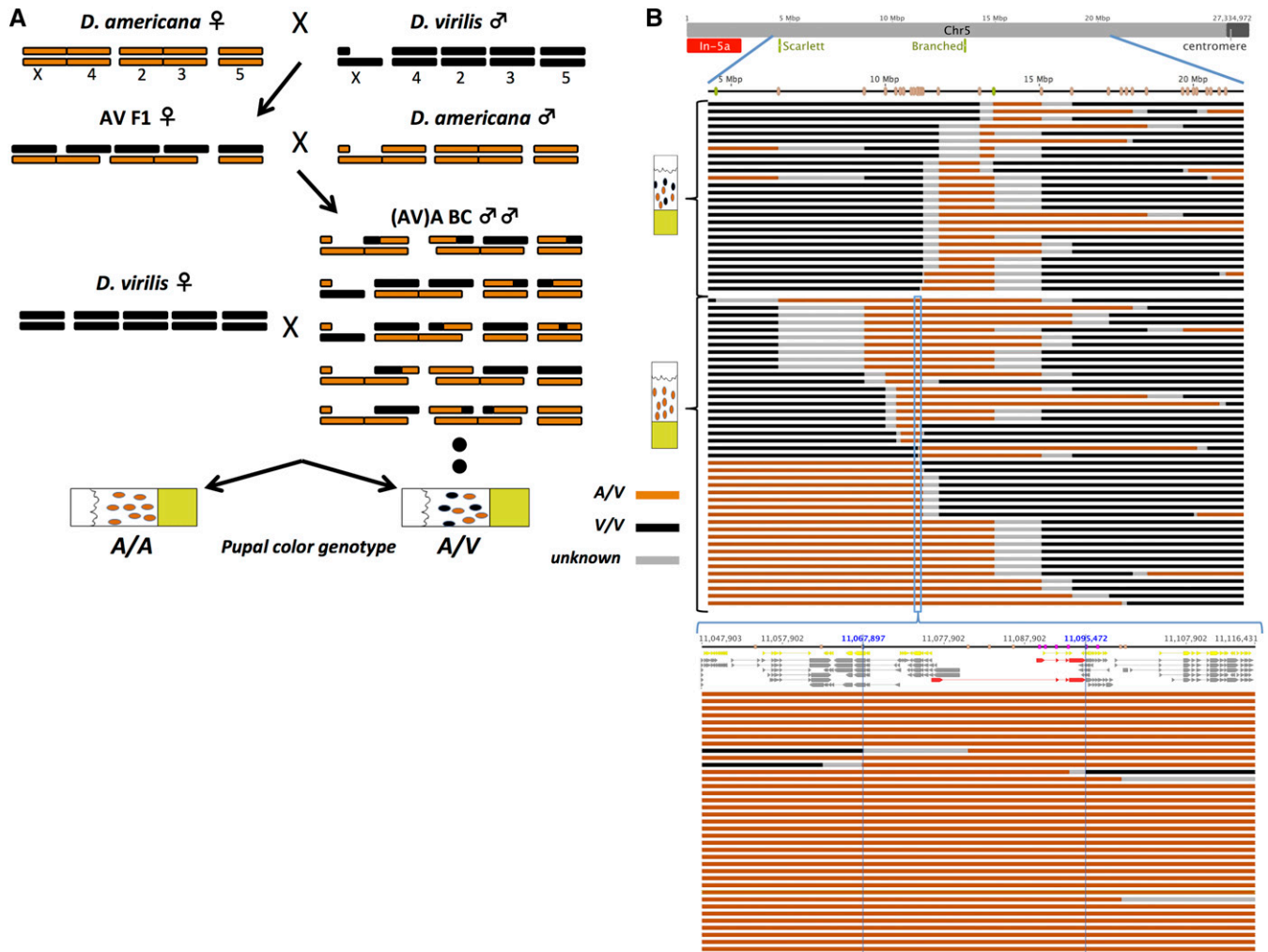


Figure 2 (A) Crossing scheme to generate recombinant backcross hybrid males. Color scheme and notation of male and female genotypes is the same as in Figure 1A. The two possible pupal color phenotypes and the corresponding genotypes at the pupal color locus for the recombinant backcross males are indicated at the bottom. (B) Genotypes of recombinant backcross males along chromosome 5. The gray line at the top represents the entire length of chromosome 5, with the location of the two visible mutations indicated in green and inversion 5a indicated in red. Blue lines magnify the molecularly genotyped region, where the location of microsatellite markers is indicated in orange. Each horizontal line below the chromosome represents a single recombinant backcross male, with regions along the chromosome color coded according to inferred genotypes using microsatellite markers; brown, A/A; black, A/V; gray, unknown. Recombination breakpoints reside within gray regions. Recombinant males are grouped according to pupal color phenotype (shown on the left). The pupal color locus resides in a black region (A/V) among recombinant males who sire both black and light brown pupae (top group), or in a light brown region (A/A) among recombinant males who sire light brown pupae only (bottom group). Bottom: close-up of a 68.5-kb candidate region among recombinant backcross males who sired light brown pupae only. Top: chromosomal coordinates and markers (orange, Microsatellite; purple, SNP) along with the gene annotations from Flybase (yellow) and those inferred by Cufflinks (gray). GJ20215 is shown in red (top, isoform A; bottom, isoform B).

Chromosome 5 (Muller element C) is perfectly associated with pupal case color (Figure 1B). Chromosomes 2–3 and 4 (Muller elements E–D and B), on the other hand, show no association with the phenotype. We also observe an excess of heterozygous genotypes for chromosome 2–3, but this excess is not correlated with segregation of black and brown pupae. The gene(s) that cause the pupal color difference between species thus reside on a single autosome: chromosome 5.

Fine mapping the pupal case color locus

D. americana harbors a number of fixed and segregating chromosomal inversions relative to *D. virilis* (Hsu 1952). Such

differences can often preclude finer recombination mapping. Fortunately, the *D. americana* strain used in this study is homo-sequential with *D. virilis* along ~85% of the euchromatic arm of chromosome 5. We can therefore fine map the locus (or loci) underlying the pupal color difference between *D. americana* and *D. virilis*. To do so, we performed two rounds of recombination mapping using two different recombinant backcross populations. We describe each below.

In the first round we used a recombinant backcross population that was created by backcrossing AV F1 females to *D. americana* males (Figure 2A). Because *Drosophila* females do recombine, backcross hybrids can carry recombinant

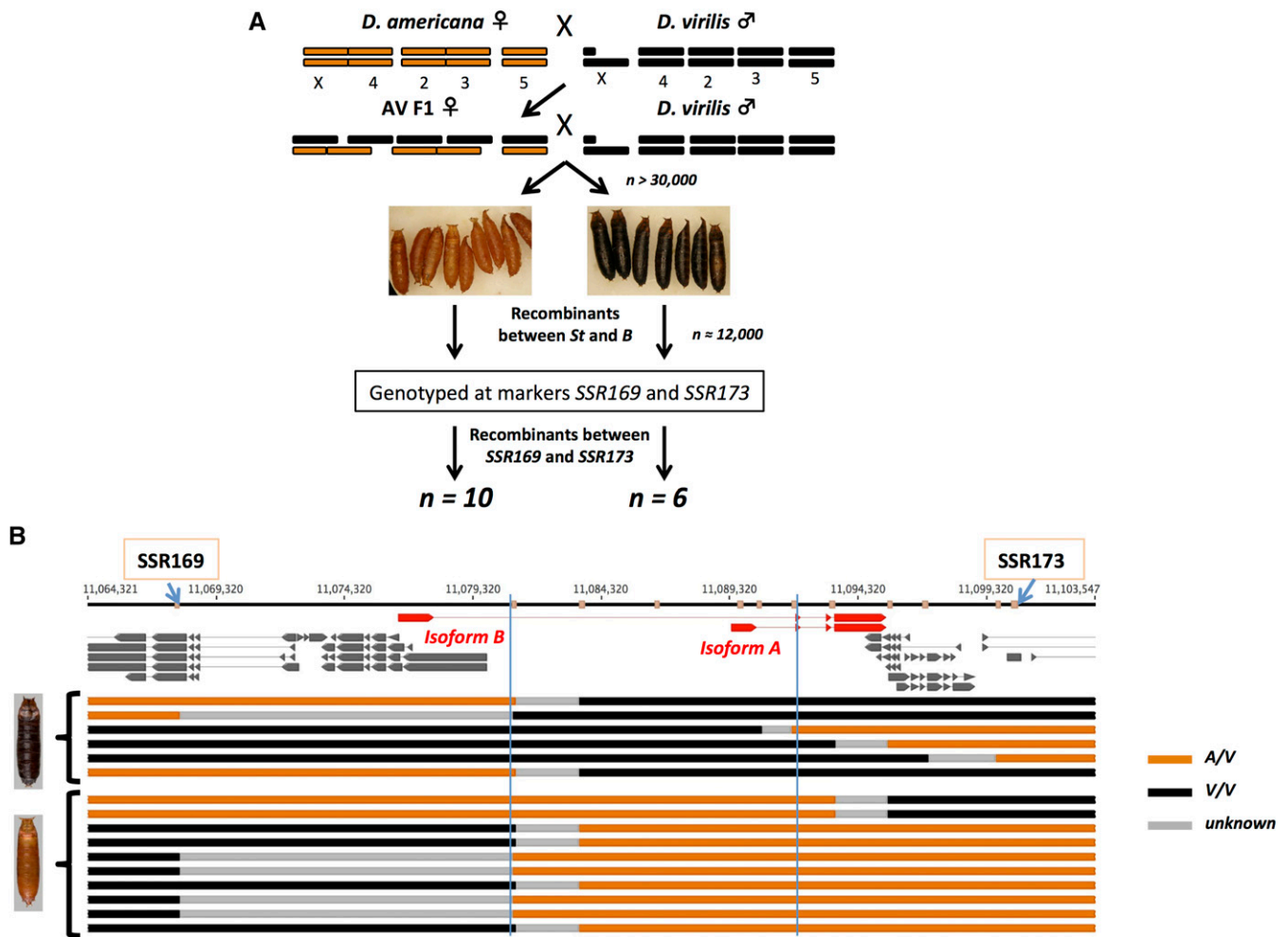


Figure 3 (A) Crossing scheme to enrich for recombinants in the candidate region. (B) Recombinant individuals recovered using the recombination enrichment strategy. Only Cufflinks assembled transcripts are shown and *GJ20215* is highlighted in red. The recombinants recovered are grouped by pupal color phenotype, where black indicates *D. virilis* homozygous regions (*V/V*) and light brown are heterozygous regions. The two vertical blue lines indicate the mapped candidate interval.

chromosomes. Backcross hybrids were either homozygous for *D. americana* alleles or heterozygous at any given locus, and therefore all had brown pupae. To infer their genotype at the putative pupal color locus we examined the pupal colors of the progeny sired when individual backcross males were crossed to *D. virilis* females. Backcross males that were heterozygous at the pupal color locus sired both black and brown pupae, while backcross males homozygous at the pupal color locus sired only brown pupae.

The *D. virilis* strain used in this study carries two visible markers on chromosome 5, *B* and *st*, both of which behave recessively in a *D. americana/D. virilis* F1 hybrid. *B* resides near the center of the chromosome (Figure 2B). We selected individuals from the recombinant backcross male population that were homozygous for *D. americana* alleles at the *B* locus ($n = 46$ such males), but carried both possible genotypes at the pupal color locus. These individuals were genotyped using microsatellite markers that spanned most of the length of the chromosome. This analysis revealed that the pupal

color locus resides in a 2.5-Mb region nearly midway between *B* and *st* (Figure S2). To refine our mapping here further, we studied an additional set of backcross males that was sparsely genotyped along chromosome 5 but carried a recombination breakpoint within the 2.5-Mb region ($n = 24$). The entire set of genotyped individuals is represented in Figure 2B. This mapping population lets us narrow the pupal case color locus to a genomic region that spans 27 kb. This region includes 4 genes: *GJ22136*, *GJ22137*, *GJ20214*, and *GJ20215* (*Dat*).

In the second round of mapping we generated another recombinant mapping population by backcrossing AV F1 females to *D. virilis* males and separated pupae based on color (Figure 3A). Performing the backcross in this way allows us to recover both color phenotypes among backcross individuals. Furthermore, we enriched for recombinants in the relevant region by genotyping only eclosed adults that had a recombination event between our two visible markers. These individuals were then genotyped at two microsatellite markers that

Table 1 Number of sequence differences within the mapped interval across four classes: fixed differences, polymorphic in *D. americana*, polymorphic in *D. virilis*, and shared polymorphisms found in the multiple sequence alignment of two *D. americana* and three *D. virilis* strains

| | Nucleotide substitution | Insertion/deletion | Total |
|------------------------------------|-------------------------|--------------------|-------|
| Fixed difference | 469 | 241 | 710 |
| Polymorphic in <i>D. americana</i> | 355 | 375 | 730 |
| Polymorphic in <i>D. virilis</i> | 130 | 92 | 222 |
| Shared polymorphism | 81 | 57 | 138 |
| Total differences | 1,035 | 765 | 1,800 |

Variants are classified separately as nucleotide substitutions and insertion/deletions.

flank the 27-kb candidate interval. We ultimately recovered 16 recombinants between the two microsatellite markers. The 16 recombinant individuals were subsequently genotyped along the 27-kb candidate region. Using this approach we narrowed the location of the relevant locus to an ~11-kb region (Figure 3B).

This region contains only the first intron, first exon, and upstream regulatory region of the annotated *Dat* isoform (isoform A). This region also includes the majority of the first intron of the unannotated isoform of this gene (isoform B). Divergence in pupal case color between *D. virilis* and *D. americana* is thus caused by sequence differences in this ~11-kb region.

Sequence alignment of mapped interval between and within *D. americana* and *D. virilis*

The mapped interval implicates *Dat* as causal and excludes adjacent genes. This interval mostly contains a sequence that is noncoding, but also includes the first exon of isoform A of *Dat*. To examine the extent of sequence divergence between the two species in this interval, we generated a multiple sequence alignment of the region using publicly available genome sequences from two strains of *D. americana* (strain H5 and strain W11; Fonseca *et al.* 2013) and two strains of *D. virilis* (strain 9 and strain 160; Blumenstiel 2014). By blasting the candidate region's sequence from the *D. virilis* genome to the two *D. americana* genomes, we recover two and three scaffolds from the H5 and W11 strains, respectively, that cover the ~11-kb region in *D. virilis*. The two H5 scaffolds (H5C.563 and H5C.564) cover all but 62 bp of the candidate region, whereas the three W11 scaffolds (W11C.1012, W11C.1013, and W11C.1014) cover all but ~200 bp. *D. virilis* strain 9 and strain 160 reads covered the entire candidate region.

We found a total of 1800 sequence differences in the candidate region between and within the strains sampled here (Table 1 and File S1). We classified base substitutions and insertion/deletions separately and partitioned them into fixed mutations, polymorphic in *D. americana*, polymorphic in *D. virilis*, and shared polymorphisms. The largest subset was nucleotide substitutions fixed between *D. americana* and *D. virilis* ($n = 469$; Table 1), while *D. americana* showed more within-species polymorphism than did *D. virilis*. The number of shared polymorphisms was also low ($n = 81$) but similar to polymorphism levels within *D. virilis*.

The sequence alignment reveals two nucleotide substitutions that are fixed between *D. americana* and *D. virilis* in the open reading frame of the first exon (Figure 4). One of these

mutations causes a nonsynonymous substitution (His → Arg) at the seventh codon, while the other is synonymous. Mutations elsewhere in the mapped interval are distributed roughly uniformly throughout the 5'-UTR of *Dat* and along the introns. Although the number of fixed nucleotide substitutions is nearly twice that of insertions/deletions, the latter affects ~60% of aligned variable sites (Figure 4 and File S1). These results show that many fixed differences exist between *D. americana* and *D. virilis* in the mapped interval, with all but two affecting noncoding sequences.

Relative expression of *Dat*

To examine the possible role of gene regulation in causing the two pupal color phenotypes, we measured relative expression levels of *Dat* using qRT-PCR at three phenotypically defined stages in *D. americana* and *D. virilis*, their hybrids, and in the other virilis phylad species. We defined first stage pupae ("pre-pupae"; *pre*) as those with hardened cuticles but no visible pigment, second stage pupae ("pupae"; *pup*) as slightly pigmented (faint brown; both species appear identical), and third stage pupae as fully pigmented ("post-pupae"; *post*) (Figure 5A).

First, we measured relative gene expression levels for both *Dat* isoforms in *D. americana* and *D. virilis* pupae (Figure 5B). We found that isoform A differs drastically in expression at all three stages ($P < 0.001$) (Table S2). Specifically, *D. americana* pupae express *Dat* isoform A >80-fold higher than *D. virilis* in all three stages. Expression of isoform B, on the other hand, differs modestly but significantly only in the first and third stages ($P = 0.0016$ and $P = 0.017$, respectively; Figure 5B, Table S2). Isoform B is on average ~1.5-fold and ~3-fold higher in *D. americana* at the first and third stages, respectively. These results suggest that isoform A is likely the isoform affected most by regulatory sequence divergence between *D. americana* and *D. virilis*.

Second, we examined expression levels of isoform A among F1 hybrids in all three stages of pupal development (Figure 5C). F1 hybrid pupae fully resemble pure *D. americana* pupae. Likewise, their expression of *Dat* isoform A did not significantly differ from that of *D. americana*, but was markedly reduced at the third stage (Table S2). These results suggest that expression levels earlier in pupal development likely determine this phenotypic difference.

As noted, the black pupal case color of *D. virilis* is unique in the virilis group: the sister species all feature brown pigment in their pupal cuticles. *D. novamexicana* pupae

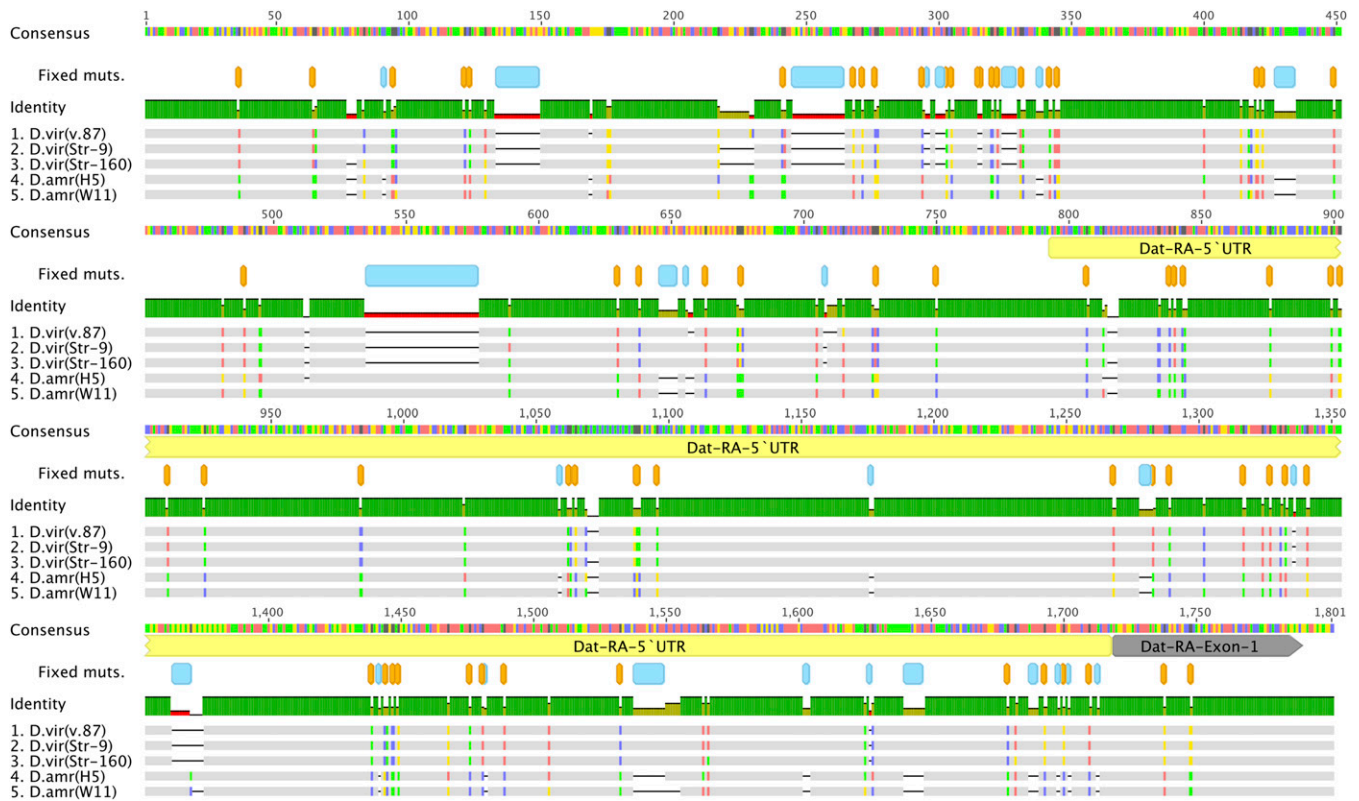


Figure 4 Multiple sequence alignment of a representative 1.8-kb region. This alignment shows a portion of the ~11-kb candidate region, which includes ~800 bp upstream of the transcription start site of GJ20215 isoform A, the 5'-untranslated region (yellow), and the first exon (gray). Fixed differences between *D. americana* and *D. virilis* are indicated in orange (nucleotide substitutions) and blue (insertion/deletions) (produced with Geneious R8).

are indistinguishable from *D. americana*. *D. lummei* pupae, on the other hand, appear intermediate between *D. virilis* and *D. americana*/*D. novamexicana*, with pupae containing both brown and black pigment (Figure 5A). We examined the expression of *Dat* isoform A in pupal samples from the three stages in all four virilis phylad species (Figure 5D). We found that expression levels in *D. novamexicana* samples were slightly higher than *D. americana* in the first two stages and slightly lower in the third stage, but this difference is not significant (Table S2). *D. lummei* expression levels, on the other hand, closely resemble *D. virilis* samples and are significantly lower than *D. americana* across all three stages ($P < 0.01$; Table S2). This suggests that the pupal color difference between *D. lummei* and *D. virilis* may have a genetic basis that differs from that identified here.

Taken together these observations are consistent with the brown pupal case color resulting from increased expression of *Dat* (isoform A) early in pupal development, whereas production of black pigment results from low expression.

The role of *Dat* in adult pigmentation

Previous work on the genetic basis of the difference in adult pigmentation between *D. americana* and *D. novamexicana* showed that it is largely caused by two genes, *ebony* and *tan*. Indeed these two genes together account for nearly 80% of the pigmentation difference (Wittkopp *et al.* 2003b;

Wittkopp *et al.* 2009). In the initial mapping study performed by Wittkopp *et al.* (2003b), a QTL region containing the *Ebony* locus showed the highest LOD score, with both *D. americana* and *D. novamexicana* alleles contributing to the abdominal pigmentation difference. Additional QTL were observed on chromosomes 3 and 5. No pigmentation candidate genes were known on these chromosomes. Given *Dat*'s mode of action described here and elsewhere (see Discussion), we hypothesized that it might play a role in the color difference between *D. americana* and *D. novamexicana* adults, contributing to the remaining ~20% difference between the two species.

We quantified the difference in pigmentation between adults of *D. americana* and *D. novamexicana* by measuring the amount of emitted red light from three landmarks: the scutellum, the lightly pigmented portion of the thorax, and the abdomen (Figure 6A). Pure species are clearly distinguishable in mean red values at all three landmarks, but the species differences are largest for abdominal coloration (Figure 6B). F1 hybrids tend to have intermediate mean red values, although with larger variance.

We tested the association between the mean red value at the three adult landmarks and genotype at three markers tightly linked to *ebony*, *tan*, and *Dat* in F6 intercross hybrids ($n = 188$; Table S3). The mean red value for genotypes at *ebony* increased significantly for the abdomen as *D. novamexicana*

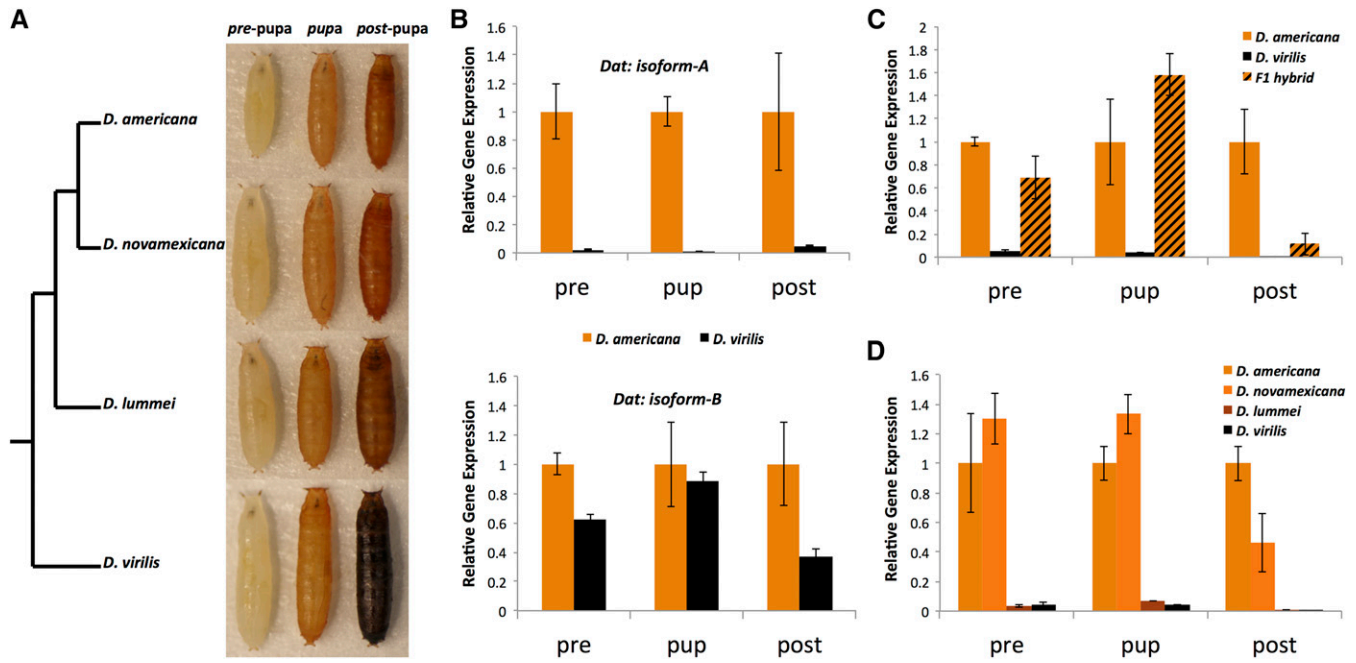


Figure 5 (A) Virilis group cladogram and pupae from the four virilis phylad members across the three developmentally defined stages used in RT-qPCR experiments. (B) Relative gene expression measures for the two isoforms in *D. americana* and *D. virilis*. Error bars represent standard error. All *D. americana* samples are normalized to 1 and considered the control sample. All comparisons are significant except for the “pup” sample of isoform B ($P < 0.05$). (C) Relative gene expression measures of GJ20215 isoform A in *D. americana*, *D. virilis*, and F1 hybrids. (D) Relative gene expression measures in all four species of the virilis group.

alleles are added ($P = 0.002$; Figure 6C and Table S3), but not the scutellum and thorax (Figure S4 and Table S3). A significant increase in abdominal mean red value was also observed with the addition of *D. novamexicana* alleles at *tan* ($P = 0.003$; Figure 6C and Table S3). In contrast, *Dat* showed no significant association with red values at any landmark ($P > 0.55$; Figure 6C and Table S3). We conclude that *Dat* plays no discernible role in the pigmentation difference between adults of *D. americana* and *D. novamexicana*.

Discussion

We have shown that a single gene entirely or almost entirely causes a morphological difference between two species of *Drosophila*. In particular, the difference between the black pupal case color of *D. virilis* and the brown pupal case color of *D. americana* maps to an ~11-kb region that includes only a single exon and part of the regulatory region of the gene, *GJ20215*. *GJ20215* is the homolog of the *D. melanogaster* *Dat*. *Dat* is highly conserved among insects and is known to play a role in the insect pigmentation pathway where it catalyzes the conversion of DA to NADA (Wittkopp *et al.* 2003a; Wittkopp and Beldade 2009).

Genes adjacent to *Dat* may be affected by regulatory changes in the mapped interval. However, we believe that they are unlikely to be involved in the pupal case color difference for two reasons. First, unlike *Dat*, the two flanking genes (*GJ22136* and *GJ22137*) have not been associated with melatonin biosynthesis in *Drosophila* or other insects.

Second, both flanking genes have orthologs in *D. melanogaster* (*DnaJ-60* and *CG4065*) that appear to show testis-biased expression (FlyBase, release FB2015_01), suggesting a role in male reproductive processes. Given our mapping resolution and *Dat*'s placement in the pigmentation pathway, *Dat* is the most likely candidate for causing the pupal case color difference.

The mapped interval contains many fixed differences between *D. americana* and *D. virilis*. Among all fixed differences between the two species (710 total), only one causes an amino acid substitution; this change resides in the first exon. The first exon also contains a silent substitution. All other fixed differences reside within introns, the 5'-UTR and the putative regulatory region(s). We have not identified the causal mutation(s) that explain the species difference. However, given the disproportionately large number of differences residing in noncoding regions, we examined differences in gene expression between the species.

The expression analysis shows that both isoforms of *Dat* show reduced abundance in *D. virilis* relative to *D. americana*. Across all stages surveyed, however, isoform A shows a more dramatic expression difference between the two species.

Dat's apparent mode of action in the species pair studied here resembles two recent findings in other insects. First, recent work found that loss of the *Dat* ortholog in the silkworm, *Bombyx Mori*, results in increased production of black pigment in larvae and adults (Dai *et al.* 2010; Zhan *et al.* 2010). Second, a recently developed dominant visible marker in insects uses a transgenic overexpression vector that expresses the *B. mori* ortholog of *Dat* (Osanai-Futahashi *et al.* 2012). This overexpression successfully reduces the

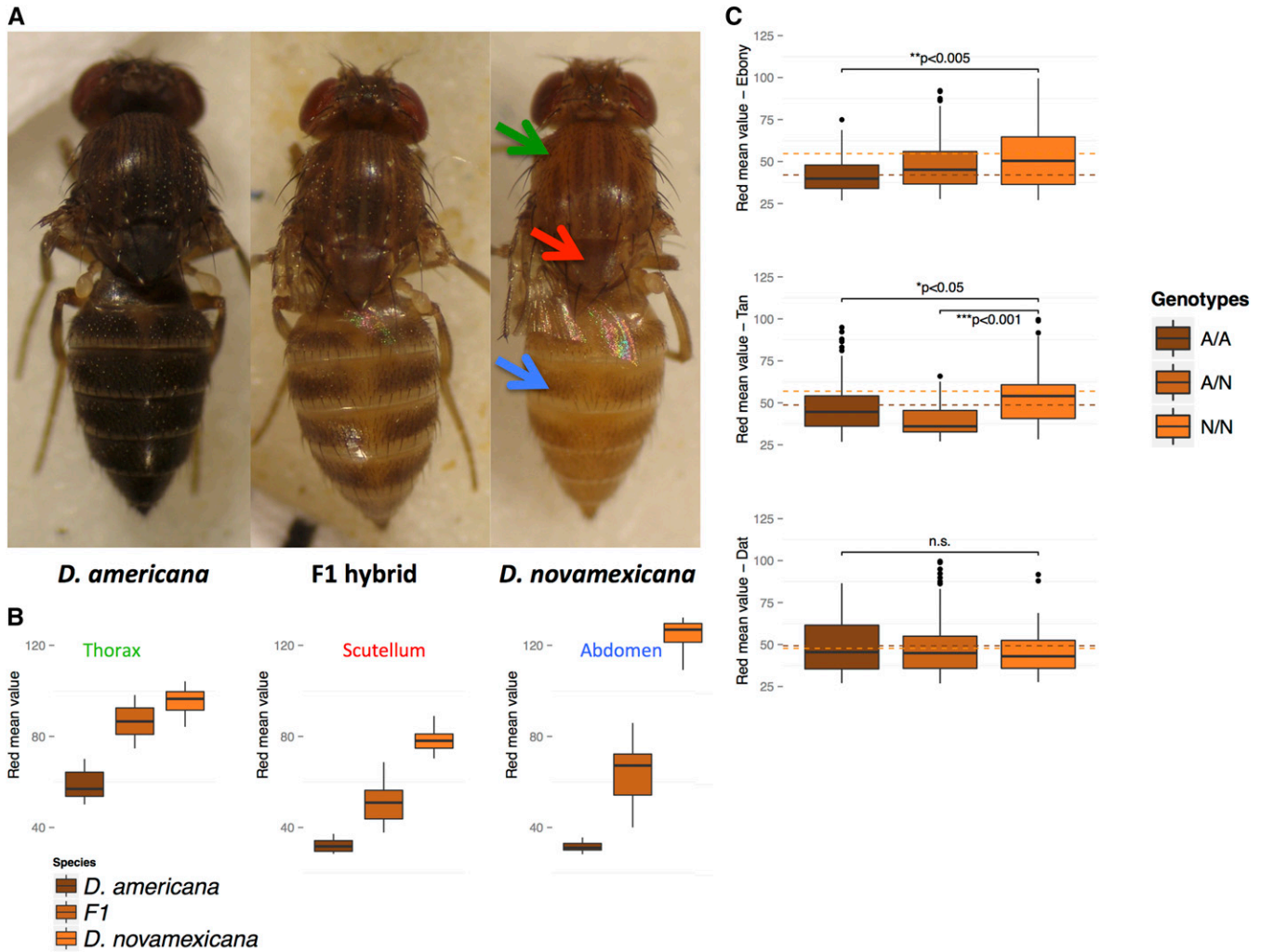


Figure 6 (A) Female *D. americana*, *D. novamexicana*, and F1 hybrids (14 days old). Arrows in the *D. novamexicana* image point to the adult landmarks used in quantifying emitted red values (arrow colors correspond to landmarks in B). (b) Distribution of mean emitted red values (represented by box plots) across a sample of *D. americana*, *D. novamexicana*, and F1 hybrids. (C) Box plot distribution of mean red values in the abdomen across genotypes for the three pigmentation genes surveyed. Red mean values are partitioned into the three genotypic classes recovered in the F6 population ($n = 188$). Dashed horizontal lines indicate the mean read value for A/A (brown) and N/N (orange) genotypes. Significant results from one-way ANOVA tests are indicated above each plot with asterisks and P -value, whereas the nonsignificant result is labeled n.s.

production of black melanin in three distantly related species (*B. mori*, *D. melanogaster*, and *Harmonia axyridis*).

These studies suggest that *Dat* has a conserved phenotypic effect in a wide range of insects and can behave as a “large-effect” gene that can render closely related species morphologically different if sufficiently divergent. In the *D. americana*/*D. virilis* pupal pigmentation difference described here, *Dat* appears to be the *only* gene responsible. Other genes in the pigmentation pathway have usually been implicated as partial or large contributors to pigmentation differences (e.g., Llopart *et al.* 2002, Wittkopp *et al.* 2003b, Jeong *et al.* 2008, but see Prud’homme *et al.* 2006), where additional unknown genes also play a role. Thus *Dat* is among a few other genes identified in *Drosophila* (e.g., Stern 1998, Sucena and Stern 2000, McGregor *et al.* 2007) that—by themselves—can cause striking differences in morphology between closely related species.

The adaptive significance of the pupal case color examined here, if any, is unknown. Pigmentation differences in adult insects have been attributed to a number of possible selective pressures, including desiccation resistance, ultraviolet protection, thermal regulation, crypsis, and sexual selection (Hollocher *et al.* 2000; Brisson *et al.* 2005; Pool and Aquadro 2007; Wittkopp *et al.* 2010; Clusella-Trullas and Terblanche 2011; Telonis-Scott *et al.* 2011; Matute and Harris 2013). Some of these—such as desiccation resistance (Terblanche and Kleynhans 2009) and crypsis (Hazel *et al.* 1998)—may also apply to puparia. Some insect larvae/pupae are also targets of endoparasitoids and pupal case characteristics may affect infection propensity (Fellowes *et al.* 1999). Selection on pigmentation phenotypes may also be indirect as genes in the pigmentation pathway are often pleiotropic, affecting a number of traits (True 2003; Wittkopp and Beldade 2009). While some of the habitat preferences exhibited by *virilis* group members

have been studied (Blight and Romano 1953), much remains to be understood regarding the ecological conditions that virilis group species experience across all life stages.

In summary, our study shows that *Dat* is a potentially important player in pigmentation in insects. More important, our study also shows that the genetic basis of morphological differences between species is sometimes simple.

Acknowledgments

We thank Bryant F. McAllister for providing the *D. americana*, *D. lummei*, and *D. novamexicana* stocks used in this study. We also thank Daniel McNabney, Robert Unckless, and two anonymous reviewers for helpful discussion and/or comments. We especially thank H. Allen Orr for guidance and support throughout the project. This work was supported by funds from a National Institutes of Health (NIH) Ruth L. Kirschstein National Research Service Award postdoctoral fellowship (GM078974) to A.L.S. and from an NIH grant (GM051932) to H. Allen Orr.

Literature Cited

- Ahmed-Braimah, Y. H., and B. F. McAllister, 2012 Rapid evolution of assortative fertilization between recently allopatric species of *Drosophila*. *Int. J. Evol. Biol.* 2012: 1–9.
- Arnoult, L., K. F. Y. Su, D. Manoel, C. Minervino, J. Magriña *et al.*, 2013 Emergence and diversification of fly pigmentation through evolution of a gene regulatory module. *Science* 339: 1423–1426.
- Bastide, H., A. Betancourt, V. Nolte, R. Tobler, P. Stöbe *et al.*, 2013 A genome-wide, fine-scale map of natural pigmentation variation in *Drosophila melanogaster*. *PLoS Genet.* 9: e1003534.
- Benson, G., 1999 Tandem repeats finder: a program to analyze DNA sequences. *Nuc. Ac. Res.* 27: 573–580.
- Blight, W. C., and A. Romano, 1953 Notes on a breeding site of *Drosophila americana* near St-Louis, Missouri. *Am. Nat.* 87: 111–112.
- Blumenstiel, J. P., 2014 Whole genome sequencing in *Drosophila virilis* identifies Polyphemus, a recently activated Tc1-like transposon with a possible role in hybrid dysgenesis. *Mob. DNA* 5: 6.
- Brisson, J. A., D. C. De Toni, I. Duncan, and A. R. Templeton, 2005 Abdominal pigmentation variation in *Drosophila polymorpha*: geographic in the trait, and underlying phylogeography. *Evolution* 59: 1046–1059.
- Caletka, B. C., and B. F. McAllister, 2004 A genealogical view of chromosomal evolution and species delimitation in the *Drosophila virilis* species subgroup. *Mol. Phyl. Evol.* 33: 664–670.
- Clusella-Trullas, S., and J. S. Terblanche, 2011 Local adaptation for body color in *Drosophila americana*: commentary on Wittkopp *et al.* *Heredity* 106: 904–905.
- Dai, F.-Y., L. Qiao, X.-L. Tong, C. Cao, P. Chen *et al.*, 2010 Mutations of an arylalkylamine-*N*-acetyltransferase, Bm-*i*AANAT, are responsible for silkworm melanism mutant. *J. Biol. Chem.* 285: 19553–19560.
- Fellowes, M. D. E., A. R. Kraaijeveld, and H. C. J. Godfray, 1999 The relative fitness of *Drosophila melanogaster* (Diptera, Drosophilidae) that have successfully defended themselves against the parasitoid *Asobara tabida* (Hymenoptera, Braconidae). *J. Evol. Biol.* 12: 123–128.
- Fonseca, N. A., R. Morales-Hojas, M. Reis, H. Rocha, C. P. Vieira *et al.*, 2013 *Drosophila americana* as a model species for comparative studies on the molecular basis of phenotypic variation. *Gen. Bio. Evo.* 5: 661–679.
- Gloor, G., and W. R. Engels, 1992 Single-fly DNA preps for PCR. *Dros. Info. Ser.* 71: 148–149.
- Haldane, J. B. S., 1937 The effect of variation of fitness. *Am. Nat.* 71: 337–349.
- Hazel, W., S. Ante, and B. Stringfellow, 1998 The evolution of environmentally-cued pupal colour in swallowtail butterflies: natural selection for pupation site and pupal colour. *Ecol. Ent.* 23: 41–44.
- Hollocher, H., J. L. Hatcher, and E. G. Dyreson, 2000 Evolution of abdominal pigmentation differences across species in the *Drosophila dunni* subgroup. *Evolution* 54: 2046–2056.
- Hsu, T. C., 1952 Chromosomal variation and evolution in the virilis group of *Drosophila*. *Univ. Texas Publ* 5204: 35–72.
- Jeong, S., M. Rebeiz, P. Andolfatto, T. Werner, J. True *et al.*, 2008 The evolution of gene regulation underlies a morphological difference between two *Drosophila* sister species. *Cell* 132: 783–793.
- Livak, K. J., and T. D. Schmittgen, 2001 Analysis of relative gene expression data using real-time quantitative PCR and the 2- $\Delta\Delta$ CT method. *Methods* 25: 402–408.
- Llopart, A., S. Elwyn, D. Lachaise, and J. A. Coyne, 2002 Genetics of a difference in pigmentation between *Drosophila yakuba* and *D. santomea*. *Evolution* 56: 2262–2277.
- Matute, D. R., and A. Harris, 2013 The influence of abdominal pigmentation on desiccation and ultraviolet resistance in two species of *Drosophila*. *Evolution* 67: 2451–2460.
- McGregor, A. P., V. Orgogozo, I. Delon, J. Zanet, D. G. Srinivasan *et al.*, 2007 Morphological evolution through multiple *cis*-regulatory mutations at a single gene. *Nature* 448: 587–590.
- Morales-Hojas, R., M. Reis, C. P. Vieira, and J. Vieira, 2011 Resolving the phylogenetic relationships and evolutionary history of the *Drosophila virilis* group using multilocus data. *Mol. Phyl. Evol.* 60: 249–258.
- Muller, H. J., 1940 Bearings of the *Drosophila* work on systematics, pp. 185–268 in *The New Systematics*, edited by Julian Huxley. Clarendon, Oxford.
- Orr, H. A., 2001 The genetics of species differences. *Trends Eco. Evol.* 16: 343–350.
- Orr, H. A., and J. A. Coyne, 1989 The genetics of postzygotic isolation in the *Drosophila virilis* group. *Genetics* 121: 527–537.
- Orr, H. A., and J. A. Coyne, 1992 The genetics of adaptation: a reassessment. *Am. Nat.* 140: 725–742.
- Osanaï-Futahashi, M., T. Ohde, J. Hirata, K. Uchino, R. Futahashi *et al.*, 2012 A visible dominant marker for insect transgenesis. *Nature Comm.* 3: 1295.
- Patterson, J. T., W. S. Stone, and A. B. Griffen, 1940 Evolution of the virilis group of *Drosophila*. *Univ. Texas Publ.* 4032: 218–250.
- Pool, J. E., and C. F. Aquadro, 2007 The genetic basis of adaptive pigmentation variation in *Drosophila melanogaster*. *Mol. Ecol.* 16: 2844–2851.
- Prud'homme, B., N. Gompel, A. Rokas, V. A. Kassner, T. M. Williams *et al.*, 2006 Repeated morphological evolution through *cis*-regulatory changes in a pleiotropic gene. *Nature* 440: 1050–1053.
- Rozen, S., and H. Skaletsky, 2000 Primer3 on the WWW for general users and for biologist programmers. *Methods Mol. Biol.* 132: 365–386.
- Sagga, N., and A. Civetta, 2011 Male-female interactions and the evolution of postmating prezygotic reproductive isolation among species of the virilis subgroup. *Int. J. Evol. Bio.* 2011: 1–11.
- Spencer, W. P., 1940 Subspecies, hybrids and speciation in *Drosophila hydei* and *Drosophila virilis*. *Am. Nat.* 74: 157–179.
- Spicer, G. S., 1991 The genetic basis of a species-specific character in the *Drosophila virilis* species group. *Genetics* 128: 331–337.

- Stalker, H. D., 1942 The inheritance of a subspecific character in the virilis complex of *Drosophila*. *Am. Nat.* 76: 426–431.
- Stern, D. L., 1998 A role of *Ultrabithorax* in morphological differences between *Drosophila* species. *Nature* 396: 463–466.
- Sucena, E., and D. L. Stern, 2000 Divergence of larval morphology between *Drosophila sechellia* and its sibling species caused by cis-regulatory evolution of *ovo/shaven-baby*. *Proc. Natl. Acad. Sci. USA* 97: 4530–4534.
- Sweigart, A. L., 2010a The genetics of postmating, prezygotic reproductive isolation between *Drosophila virilis* and *D. americana*. *Genetics* 184: 401–410.
- Sweigart, A. L., 2010b Simple Y-autosomal incompatibilities cause hybrid male sterility in reciprocal crosses between *Drosophila virilis* and *D. americana*. *Genetics* 184: 779–787.
- Takahashi, A., and T. Takano-Shimizu, 2011 Divergent enhancer haplotype of ebony on inversion In(3R)Payne associated with pigmentation variation in a tropical population of *Drosophila melanogaster*. *Mol. Ecol.* 20: 4277–4287.
- Takahashi, A., K. Takahashi, R. Ueda, and T. Takano-Shimizu, 2007 Natural variation of ebony gene controlling thoracic pigmentation in *Drosophila melanogaster*. *Genetics* 177: 1233–1237.
- Telonis-Scott, M., A. A. Hoffmann, and C. M. Sgrò, 2011 The molecular genetics of clinal variation: a case study of ebony and thoracic trident pigmentation in *Drosophila melanogaster* from eastern Australia. *Mol. Ecol.* 20: 2100–2110.
- Terblanche, J. S., and T. E. Kleyhans, 2009 Phenotypic plasticity of desiccation resistance in *Glossina puparia*: Are there ecotype constraints on acclimation responses? *J. Evol. Biol.* 22: 1636–1648.
- Trapnell, C., L. Pachter, and S. L. Salzberg, 2009 TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* 25: 1105–1111.
- Trapnell, C., B. A. Williams, G. Pertea, A. Mortazavi, G. Kwan *et al.*, 2010 Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* 28: 511–515.
- True, J. R., 2003 Insect melanism: the molecules matter. *Trends Ecol. Evol.* 18: 640–647.
- Werner, T., S. Koshikawa, T. M. Williams, and S. B. Carroll, 2010 Generation of a novel wing colour pattern by the Wingless morphogen. *Nature* 464: 1143–1148.
- Williams, T. M., J. E. Selegue, T. Werner, N. Gompel, A. Kopp *et al.*, 2008 The regulation and evolution of a genetic switch controlling sexually dimorphic traits in *Drosophila*. *Cell* 134: 610–623.
- Wittkopp, P. J., and P. Beldade, 2009 Development and evolution of insect pigmentation: Genetic mechanisms and the potential consequences of pleiotropy. *Sem. Cell Dev. Bio.* 20: 65–71.
- Wittkopp, P. J., K. Vaccaro, and S. B. Carroll, 2002 Evolution of yellow gene regulation and pigmentation in *Drosophila*. *Curr. Biol.* 12: 1547–1556.
- Wittkopp, P. J., S. B. Carroll, and A. Kopp, 2003a Evolution in black and white: genetic control of pigment patterns in *Drosophila*. *Trends Genet.* 19: 495–504.
- Wittkopp, P. J., B. L. Williams, J. E. Selegue, and S. B. Carroll, 2003b *Drosophila* pigmentation evolution: divergent genotypes underlying convergent phenotypes. *Proc. Natl. Acad. Sci. USA* 100: 1–6.
- 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: 540–544.
- Wittkopp, P. J., G. Smith-Winberry, L. L. Arnold, E. M. Thompson, A. M. Cooley *et al.*, 2010 Local adaptation for body color in *Drosophila americana*. *Heredity* 106: 592–602.
- Zhan, S., Q. Guo, M. Li, M. Li, J. Li *et al.*, 2010 Disruption of an *N*-acetyltransferase gene in the silkworm reveals a novel role in pigmentation. *Development* 137: 4083–4090.

Communicating editor: D. A. Barbash

GENETICS

Supporting Information

<http://www.genetics.org/lookup/suppl/doi:10.1534/genetics.115.174920/-/DC1>

A Single Gene Causes an Interspecific Difference in Pigmentation in *Drosophila*

Yasir H. Ahmed-Braimah and Andrea L. Sweigart



Figure S1: *D. americana* (left) and *D. virilis* (right) pupae at stage 3 (see Materials and Methods).

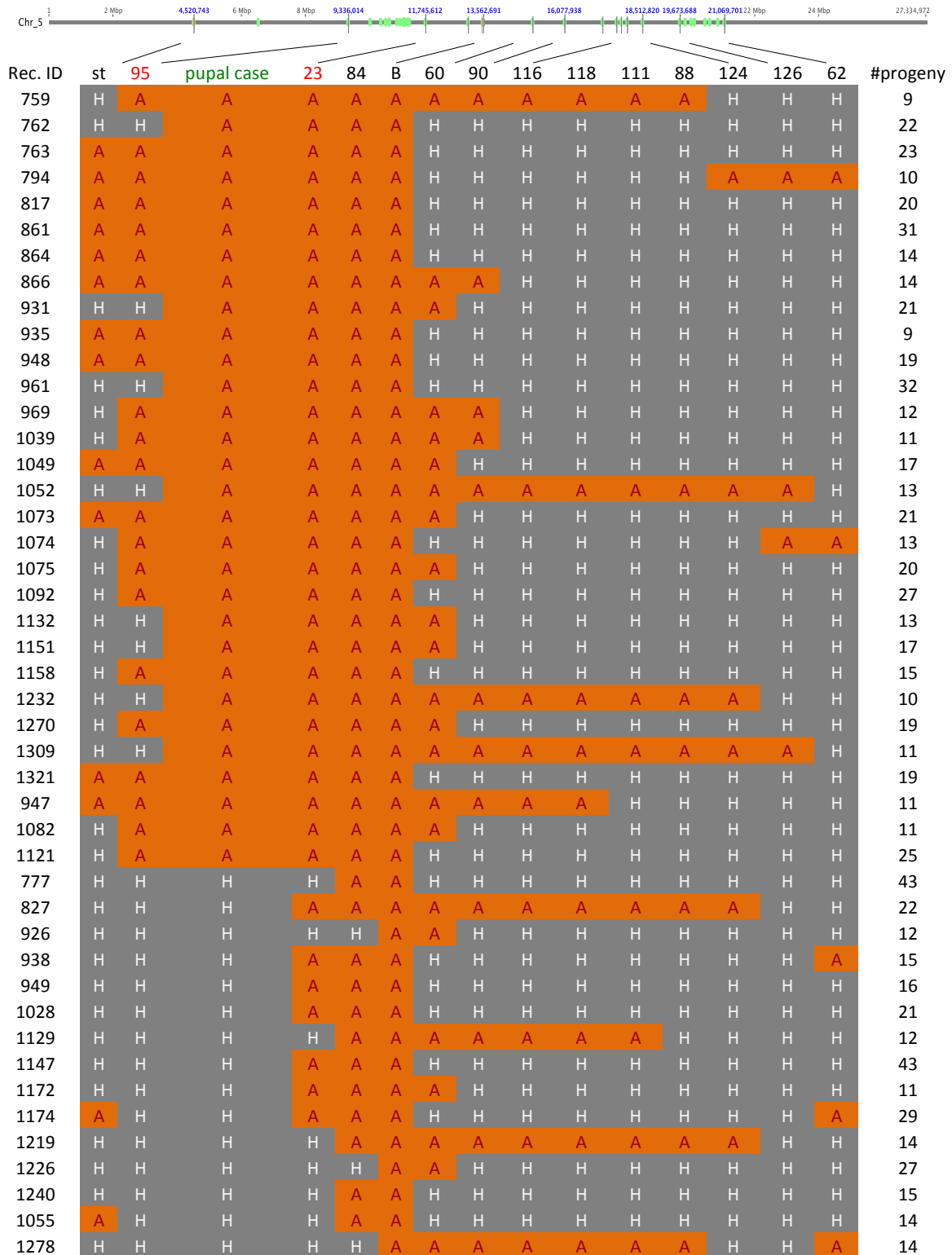


Figure S2: Genotype table of recombinant backcross males from the first round of genotyping (see Materials and Methods). Genotypes at each marker are indicated by letter and color (orange/A = *D. americana* homozygote, gray/H = heterozygote). Pupal case genotype is inferred from progeny pupal case colors and is indicated at the putative location (green) between markers SSR95 and SSR23 (red).

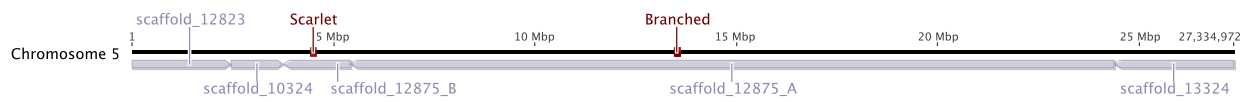


Figure S3: Correspondence of *D. virilis* genome scaffolds (release droVir2, UCSC) to chromosome 5 (Muller Element C; Schaffer *et al.* 2008).

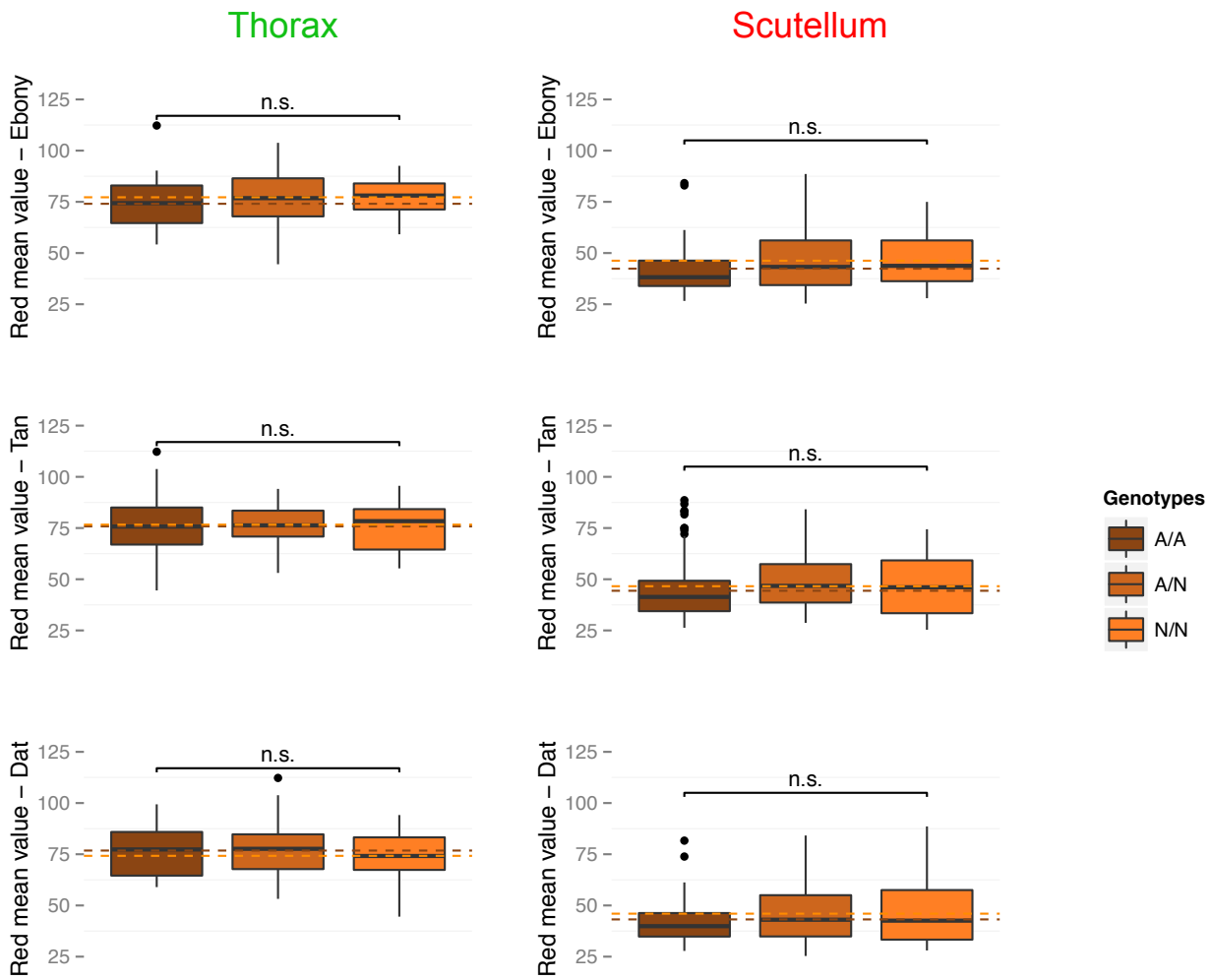


Figure S4: Boxplot distribution of mean red values in the thorax and scutellum across genotypes for the three pigmentation genes surveyed: Red mean values are partitioned into the three genotypic classes recovered in the F6 population (n=188). Dashed horizontal lines indicate the mean read value for A/A (brown) and N/N (orange) genotypes.

Table S1: List of genetic markers used in this study, their coordinates, and primer sequences.

| Name(type*) | Chr_5 coords. | scaffold_12875 coords. | Forward primer | Reverse primer |
|--------------------|-------------------|------------------------|-----------------------|-----------------------|
| St (VM) | 4519096-4520743 | 922609-924256 | N/A | N/A |
| SSR58 (MS) | 6545087-6545300 | 19509634-19509847 | TGCCTAGCATTTGGCACTTA | AAAAGAGCGTGGCAAAGAAA |
| SSR95 (MS) | 9335798-9336014 | 16718920-16719136 | TGTGCCTGCTGACAAAACAT | ACACTGCCTGCTTGCAATTA |
| SSR138 (MS) | 10017486-10017694 | 16037240-16037448 | TCGTACCAATGTTGTCAATGC | ACAGGAAGTGGGTGGAAGTG |
| SSR139 (MS) | 10361473-10361687 | 15693247-15693461 | CAAAGATGGAGGAAGCCTCA | CCAGCGCACACTCATAGATT |
| SSR157 (MS) | 10504504-10504720 | 15550214-15550430 | TTTCGTTTGCATTTCCCTGTG | GCGTTCGAACCTGAACAAAT |
| SSR155 (MS) | 10604771-10604976 | 15449958-15450163 | TGGGACAAATGTGCCAAATA | ATATGCTCGGTGCGGAGATTC |
| SSR154 (MS) | 10621121-10621371 | 15433563-15433813 | CGGGCAGAATCAACTCGTAG | CCGCAGACACTGAACCAAC |
| SSR150 (MS) | 10864161-10864552 | 15190382-15190773 | GCTATGTTTGTGCGAGTGA | ATTGCCAAAGCGCATAAATA |
| SSR149 (MS) | 10955634-10955876 | 15099058-15099300 | CAGTACATCGGGCAATTGTG | TGTGGGTGAAGTCAATGTTGT |
| SSR148 (MS) | 11054476-11054716 | 15000218-15000458 | ACGTGCCACATTGCTCATA | ACAACCTGTCGCGCACTAAA |
| SSR170 (MS) | 11062637-11062844 | 14992090-14992297 | TTTTTCGTTTTCTCCGTTTTT | ATCTAAGCAAGCGGGTTTCA |
| SSR169 (MS) | 11067738-11067901 | 14987033-14987196 | CATCAAGGAGCTGGCCTATC | ACTTCTCCGCGTACAACCTG |
| SSR168 (MS) | 11080856-11080998 | 14973936-14974078 | CTCCTGGCTGAAAAGCAAAC | CGGCAAAACAAATGACAATC |
| SSR165 (MS) | 11083472-11083698 | 14971236-14971462 | TTTGACCAGGGCATGACATA | AAGCACAAGCAAAGCACACAC |
| SSR167 (MS) | 11086448-11086593 | 14968341-14968486 | TTCGAATTTCCGCTTGATTT | TATAGCCGGCTCGTTGAAGT |
| pF10 (SNP) | 11089640-11089848 | 14965086-14965294 | AGCTTTAGCTTCCCTTGTC | TTGAAATTTGCGCGCTCTAC |
| pE10 (SNP) | 11090368-11090570 | 14964364-14964566 | CGCGATCGCAGCTATTCTT | TCCAGATGTGCGGTGAAGTGT |
| pD10 (SNP) | 11091748-11091976 | 14962958-14963186 | TCGACGACGACATATTGAGC | GTCTGGCTTTCTAAGAGCA |
| pC10 (SNP) | 11093207-11093456 | 14961478-14961727 | AGGGATCGGCCAATTATC | AAACCCCTCAAGGACAACCTG |
| pB10 (SNP) | 11095456-11095683 | 14959251-14959478 | TACGCAACGCCTATCTGAAA | TCAAGCTGTCGTGTAATCC |
| pA10 (SNP) | 11096841-11097053 | 14957881-14958093 | GATCGTCGCGACTTTCACTA | CGATGGGCAATAATATGGAG |
| SSR164 (MS) | 11099701-11099906 | 14955028-14955233 | TTGTGTTGTGCTGATGCTG | GCAAAGCAAAGCTGCGTA |
| SSR147 (MS) | 11129608-11129795 | 14925139-14925326 | TTAGGCGCTGCTACTCACAC | GCGTACAGCTGATAAGCACAG |
| SSR163 (MS) | 11179710-11179876 | 14875058-14875224 | TCAATGCACGAAAATTGGTC | GTTGCGGTAAAATGGATTCCG |
| SSR162 (MS) | 11205769-11205921 | 14849013-14849165 | AGGTAATGTTGGCCAATGC | TTCGAAAGGCATTTTGTGA |
| SSR146 (MS) | 11245420-11245584 | 14809350-14809514 | GTGACAATTGTTGCGGAGTG | CGCCGTCTGAAAAGAGAAAAC |
| SSR23 (MS) | 11745404-11745612 | 14309322-14309530 | AAACTGGCAGATGGGCATAG | CCACGATTTCAGAAGCACAA |
| SSR84 (MS) | 13073885-13074040 | 12980894-12981049 | CAGCATGGAGCATCTGTGTC | TGGAAGGGATGTCATGGACT |
| B (VM) | 13512463-13562691 | 12492265-12542347 | N/A | N/A |
| SSR60 (MS) | 15084664-15084836 | 10970098-10970270 | CAAAAGTGTTCCTTGATGG | GGGTTCTAGCCCCAAATAA |
| SSR90 (MS) | 16077689-16077938 | 9976996-9977245 | ACTTTGCCAAGCTGTGAAGG | GCGTCTCGTATGCTCTGCTA |
| SSR116 (MS) | 17269263-17269460 | 8785474-8785671 | CCCCATTGAAAGTTCATCCA | GTCAGGAGGCCACATTGTTT |
| SSR118 (MS) | 17695072-17695271 | 8359663-8359862 | GCCCAAAATCTTAGCCAAA | TGGCTTGGGTACTGTTTCT |
| SSR111 (MS) | 17849127-17849347 | 8205587-8205807 | TTTGATTGTTTCCCTCACTCG | TGTCATTGCTCTGGCAAAA |
| SSR88 (MS) | 18040111-18040269 | 8014665-8014823 | CCAAAAGGCAGGACCATAAA | TTGCGTAGACACCACAAGGT |
| SSR124 (MS) | 18512607-18512820 | 7542114-7542327 | CGCTTAAACGATCCAACGAT | GTTTATAACCGGTGCTCGAT |
| SSR126 (MS) | 19673479-19673688 | 6381246-6381455 | AATTGCCAAAGAAACCACCA | AGCGGCTCGTCTGCTGTA |
| SSR128 (MS) | 19839381-19839552 | 6215382-6215553 | ACGACTTGTGCTGATAGGC | CTGTTCTGGAACATGCAAGC |
| SSR129 (MS) | 20026180-20026357 | 6028577-6028754 | ACCATGAGCAGGCATCAGA | CCGGCAATCGTCTTTAACT |
| SSR130 (MS) | 20129136-20129321 | 5925613-5925798 | ACACAGGTCCCAACACAACA | CAGAAAATGATGTGCGTCTG |
| SSR133 (MS) | 20473430-20473645 | 5581289-5581504 | CTCGCAACTTGGCAGACATA | TGCCCAAATGTAAGTGTATC |
| SSR134 (MS) | 20601881-20602035 | 5452899-5453053 | GCACCGCCATACCCATAAT | ATTTGGCCACCCATTAGACC |
| SSR136 (MS) | 20858440-20858624 | 5196310-5196494 | ATGCCAAACAATGCTAAACA | TCCCGATGATCATTGTCTT |
| SSR62 (MS) | 21069407-21069701 | 4985233-4985527 | TGTTAGTTGGCAGCGCAAT | GATTATGCGTGTGCGAGTCG |

*VM: visible marker, MS: microsatellite marker, SNP: single nucleotide polymorphism.

Table S2: Statistical analysis of RT-qPCR results. Two sample *t*-tests (one-tailed) for each relevant comparison are shown. Significant *p*-values (<0.05) are indicated in red.

| Samples compared (isoform) | t-value | <i>p</i> -value | Associated figure |
|----------------------------|---------|-----------------|-------------------|
| Apré v. Vpre (A) | 17.54 | <0.00001 | 5b |
| Apup v. Vpup (A) | 35.19 | <0.00001 | 5b |
| Apos v. Vpos (A) | 5.69 | 0.00046 | 5b |
| Apré v. Vpre (B) | 4.63 | 0.00169 | 5b |
| Apup v. Vpup (B) | 0.31 | 0.76582 | 5b |
| Apos v. Vpos (B) | 2.99 | 0.01739 | 5b |
| Apré v. F1pre (A) | 1.49 | 0.13674 | 5c |
| Apup v. F1pup (A) | 0.27 | 0.40072 | 5c |
| Apos v. F1pos (A) | 2.21 | 0.07916 | 5c |
| Vpre v. F1pre (A) | 7.29 | 0.00915 | 5c |
| Vpup v. F1pup (A) | 25.46 | 0.00081 | 5c |
| Vpos v. F1pos (A) | 2.88 | 0.05133 | 5c |
| Apos v. F3Br (A) | 2.05 | 0.05457 | 5c |
| Vpos v. F3Bl (A) | 2.45 | 0.06719 | 5c |
| Apré v. Npre (A) | 1.67 | 0.11882 | 5d |
| Apup v. Npup (A) | 1.91 | 0.09801 | 5d |
| Apos v. Npos (A) | 1.89 | 0.10008 | 5d |
| Apré v. Lpre (A) | 7.92 | 0.00778 | 5d |
| Apup v. Lpup (A) | 29.88 | 0.00056 | 5d |
| Apos v. Lpos (A) | 32.38 | 0.00048 | 5d |
| Lpre v. Npre (A) | 13.38 | 0.00277 | 5d |
| Lpup v. Npup (A) | 36.6 | 0.00037 | 5d |
| Lpos v. Npos (A) | 8.02 | 0.00761 | 5d |
| Lpre v. Vpre (A) | 0.23 | 0.42108 | 5d |
| Lpup v. Vpup (A) | 2.72 | 0.05652 | 5d |
| Lpos v. Vpos (A) | 2.84 | 0.05235 | 5d |

Table S3: Statistical analysis of adult pigmentation between *D. americana* and *D. novamexicana*. One-way ANOVA was performed for genotype at each of the three genes and mean red value at three adult landmarks. Significant *p*-values (<0.05) are indicated in red.

| Tukey HSD adjusted <i>p</i>-value | | | | | | |
|------------------------------------------|-----------------|----------------|-------------------|------------|------------|------------|
| Gene | Landmark | F value | Pr (>F) | N-A | H-A | N-H |
| <i>ebony</i> | Abdomen | 4.969 | 0.00244 | 0.00221 | 0.0389 | 0.31829 |
| | Scutellum | 0.799 | 0.496 | 0.59866 | 0.46525 | 0.99959 |
| | Thorax | 0.558 | 0.644 | 0.61987 | 0.74585 | 0.96329 |
| <i>tan</i> | Abdomen | 4.85 | 0.00285 | 0.12649 | 0.07594 | 0.00205 |
| | Scutellum | 0.935 | 0.425 | 0.89283 | 0.4431 | 0.91896 |
| | Thorax | 0.177 | 0.912 | 0.99757 | 0.99912 | 0.99423 |
| <i>Dat</i> | Abdomen | 0.162 | 0.922 | 0.93054 | 0.98801 | 0.97699 |
| | Scutellum | 0.17 | 0.916 | 0.94041 | 0.89818 | 0.99981 |
| | Thorax | 0.692 | 0.558 | 0.66698 | 0.99866 | 0.55199 |

References:

Schaeffer, S. W. *et al.*, 2008 Polytene Chromosomal Maps of 11 *Drosophila* Species: The Order of Genomic Scaffolds Inferred From Genetic and Physical Maps. *Genetics* 179:1601-1655.

File S1

Nucleotide alignment in FASTA format.

Available for download at <http://www.genetics.org/lookup/suppl/doi:10.1534/genetics.115.174920/-/DC1>