Title

The pdm3 Locus Is a Hotspot for Recurrent Evolution of Female-Limited Color Dimorphism in Drosophila

Journal

Current Biology, 26(18)

ISSN

0960-9822

Authors

Yassin, Amir
Delaney, Emily K
Reddiex, Adam J
et al.

Publication Date

2016-09-01

DOI

10.1016/j.cub.2016.07.016

Peer reviewed

HHS Public Access

Author manuscript

Curr Biol. Author manuscript; available in PMC 2017 May 31.

Published in final edited form as:

Curr Biol. 2016 September 26; 26(18): 2412–2422. doi:10.1016/j.cub.2016.07.016.

The *pdm3* locus is a hotspot for recurrent evolution of female-limited color dimorphism in *Drosophila*

Amir Yassin[a,*], **Emily K. Delaney**[b,*], **Adam J. Reddiex**[c], **Thaddeus D. Seher**[b,d], **Héloïse Bastide**[a], **Nicholas C. Appleton**[c], **Justin B. Lack**[a], **Jean R. David**[e], **Stephen F. Chenoweth**[c], **John E. Pool**[a,+], and **Artyom Kopp**[b,+]

[a]Laboratory of Genetics, University of Wisconsin-Madison, 425-G Henry Mall, Madison, WI, 53705, USA

[b]Department of Evolution and Ecology, University of California-Davis, One Shields Avenue, Davis, CA, 95616, USA

[c]School of Biological Sciences, The University of Queensland, St. Lucia, Brisbane 4072, Australia

[d]Department of Quantitative and Systems Biology, University of California-Merced, 5200 North Lake Rd, Merced, CA, 95343, USA

[e]Laboratoire Evolution, Génomes, Comportement, Ecologie (EGCE), CNRS, IRD, Université Paris Sud, Université Paris-Saclay, Gif-sur-Yvette, France

Summary

Sex-limited polymorphisms are an intriguing form of sexual dimorphism that offer unique opportunities to reconstruct the evolutionary changes that decouple male and female traits encoded by a shared genome. We investigated the genetic basis of a Mendelian female-limited color dimorphism (FLCD) that segregates in natural populations of more than 20 species of the *Drosophila montium* subgroup. In these species, females have alternative abdominal color morphs, light and dark, whereas males have only one color morph in each species. A comprehensive molecular phylogeny of the *montium* subgroup supports multiple origins of FLCD. Despite this, we mapped FLCD to the same locus in four distantly related species – the transcription factor *POU domain motif 3* (*pdm3*), which acts as a repressor of abdominal pigmentation in *D. melanogaster*. In *D. serrata,* FLCD maps to a structural variant in the first intron of *pdm3*; however, this variant is not found in the three other species – *D. kikkawai*, *D. leontia,* and *D. burlai* – and sequence analysis strongly suggests the *pdm3* alleles responsible for FLCD originated

[+]Correspondence: John E. Pool, 425-G Henry Mall, Madison, WI, 53705, USA, +1-608-265-1036, jpool@wisc.edu, Artyom Kopp, One Shields Avenue, Davis, CA, 95616, USA, +1-530-752-8657, akopp@ucdavis.edu.

[*]Equal contributions.

Author contributions: A.Y. conducted phylogenetic analyses, introgression line analysis in *D. kikkawai*, *D. leontia* and *D. burlai*, population genetic analyses, and co-wrote the paper. E.K.D. conducted genetic and association mapping in *D. serrata*, population genetic analyses, functional tests in *D. melanogaster*, cuticle imaging, and co-wrote the paper. A.J.R. collected, inbred, and sequenced the association mapping population for *D. serrata*. T.S. did RNA-seq and produced and genotyped F_2 backcross and introgression mapping populations for *D. serrata*. H.B. performed genome alignment and CMH tests in *D. kikkawai*, *D. leontia* and *D. burlai*. N.C.A. phenotyped the *D. serrata* association mapping population. J.B.L. constructed genomic libraries in *D. kikkawai*, *D. leontia* and *D. burlai*. J.R.D. generated *D. kikkawai*, *D. leontia* and *D. burlai* light and dark lines and introgression lines and conducted complementation test between *D. kikkawai* and *D. leontia*. S.F.C. collected, phenotyped, and sequenced *D. serrata* lines and conducted the association analysis. J.E.P. co-directed the project and co-wrote the paper. A.K. co-directed the project, generated inbred parental and introgression mapping lines for *D. serrata* and *D. kikkawai,* made RNA-seq libraries, and co-wrote the paper.

independently at least three times. We propose that *cis*-regulatory changes in *pdm3* form sexually dimorphic and monomorphic alleles that segregate within species and are preserved, at least in one species, by structural variation. Surprisingly, *pdm3* has not been implicated in the evolution of sex-specific pigmentation outside the *montium* subgroup, suggesting that the genetic paths to sexual dimorphism may be constrained within a clade, but variable across clades.

Graphical abstract

Yassin et al. show that a female-limited polymorphism has evolved multiple times in the *Drosophila montium* species subgroup through independent changes in the same gene, *pdm3*. This strong genetic bias suggests that the evolution of sex-specific traits may follow relatively predictable paths constrained by genetic architecture.

Keywords

sex-limited polymorphism; sexual dimorphism; abdominal pigmentation; parallel evolution; *pdm3*

Introduction

Sexual dimorphism is associated with some of the most extravagant phenotypes found in nature [1]. Despite significant progress [2–6], the molecular mechanisms responsible for the origin of sex-specific traits remain poorly understood. Given the prevalence of sex-specific phenotypes, it is essential to take a broad comparative perspective. Does sexual dimorphism evolve via a limited set of genes and developmental pathways, or are there many ways to evolve a sexually dimorphic trait?

The genus *Drosophila* provides an excellent model for comparative analysis of sex-specific traits because it offers numerous transitions between sexual dimorphism and monomorphism. Many species in the *melanogaster* species group have sexually dimorphic pigmentation of posterior abdominal segments – usually dark in males and light in females [7–10]. The male-specific pigmentation is an evolutionarily derived state controlled by sex-specific expression of the *bric-à-brac* genes [7,11]. However, several species in this group have monomorphic pigmentation, indicating recurrent transitions to or reversions from sexual dimorphism. Lack of male-specific pigmentation has been traced to either loss of sexually dimorphic *bab* expression, or to changes in the *cis*-regulatory elements of pigment synthesis enzymes such as *yellow, tan,* and *ebony* [12–14]. These studies suggest that sex-specific abdominal pigmentation evolved through changes in the expression of a critical upstream regulator (*bab*), but can be further modified by mutations in downstream pigmentation genes.

In addition to male-specific pigmentation, some *Drosophila* species have female-limited color patterns, providing a rare opportunity to compare the genetic basis of evolutionary changes in male- and female-specific traits that are controlled by the same underlying gene network. Male pigmentation is seldom polymorphic within a species, and when variation exists, it usually distinguishes allopatric populations or subspecies [14]. In contrast, more than 20 species of the *D. montium* subgroup (a lineage within the *melanogaster* species

group, sometimes elevated to species group status [15]) have evolved female-limited color dimorphism (FLCD) in posterior abdominal segments [16,17] (Figure 1). In these species, females have discrete light and dark color morphs that segregate within natural populations, while males (with the exception of *D. truncata*) have only one abdominal color morph in each species, which can be either light or dark (Figure 2). FLCD, which is controlled by a single Mendelian locus (Table S1 and Supplemental Experimental Procedures), implies that sexually dimorphic and monomorphic genotypes segregate within a single species, making the *montium* subgroup an ideal model for reconstructing evolutionary transitions between sexually monomorphic and sex-specific genetic architecture.

FLCD has been described in only one species outside the *montium* subgroup, *D. erecta*, where it is controlled by *cis*-regulatory alleles of the X-linked melanin synthesis enzyme *tan* [18]. In contrast, FLCD in the *montium* subgroup is controlled by an unknown autosomal locus in all studied species [16], suggesting that this trait may have different genetic bases in the two lineages. It is unknown whether the numerous cases of FLCD in the *montium* subgroup are caused by mutations in the same gene, or perhaps even the same mutations. To elucidate this question, we undertook a comparative study of FLCD in several distantly related species. Although our phylogenetic analysis indicates multiple origins of FLCD in the *montium* subgroup, this polymorphism maps to the same gene, the POU domain transcription factor *pdm3*, in all cases. However, the sequence variants associated with FLCD are unique to each species except the two most closely related ones, supporting multiple origins of FLCD through independent changes in the same locus. This is the first time *pdm3* has been implicated in the evolution of color patterns. Intriguingly, an accompanying study of male-specific pigmentation in the closely related *ananassae* species subgroup [14] reveals a completely different, but equally reproducible genetic basis of evolutionary changes, suggesting that genetic biases that constrain the evolution of sex-specific traits may differ both between clades and between sexes.

Results

Phylogenetic relationships do not predict FLCD in the *montium* subgroup

To gain insight into the evolutionary history of FLCD in the *montium* subgroup, we inferred a molecular phylogeny of 44 species from one mitochondrial and three nuclear genes. The Bayesian tree with the greatest posterior probability identified six clades whose monophyly was supported by >95% posterior probability (clades I-VI in Figure 2A). We then evaluated phylogenetic signal, a measure of the tendency for species to share phenotypic traits because of their phylogenetic relationships, for male and female abdominal color using information gathered from the literature and our collections (Table S1 and Supplemental Experimental Procedures). Bayesian inference of color evolution in each sex produced contrasting results. While male ancestors of all clades could be significantly assigned to light or dark phenotypes indicating a strong signal ($\lambda = 0.82$), female ancestral states varied within all clades with a low signal ($\lambda = 0.33$) that was not significantly different from 0 (Bayes Factor = 1.58) (Figure 2B). This difference is consistent with the absence of evolutionary correlation between male and female pigmentation (Bayes Factor = 0.14). The rate of loss of dark pigmentation in females is higher than the rate of gain regardless of the male phenotype

(Figure 2C), and light monomorphic species are more common than dark monomorphic species.

FLCD in *D. serrata* maps to a structural variant in *pdm3*

To investigate the genetic basis of FLCD in *D. serrata,* we crossed light and dark strains and found that female pigmentation is controlled by a single autosomal Mendelian locus with a dominant dark allele, similar to other *montium* subgroup species. To identify the pigmentation locus, we implemented a three stage mapping process. First, we genotyped light and dark F_2 backcross females at 39 single nucleotide polymorphism (SNP) markers distributed across the genome. Two markers had complete association with female pigmentation, and permutation tests yielded a genome-wide-adjusted p-value of < 0.001 (Figure S1A). These markers were located on scaffolds 2 and 15 of the *D. serrata* genome assembly [19], corresponding to a single region on chromosome arm 2R in *D. melanogaster*. To further map the pigmentation locus, we introgressed the dominant dark allele into a light genomic background for 25 generations. We then examined F_{25} introgression and F_2 backcross genotypes for 29 SNP markers located within scaffolds 2 and 15 (Supplemental Experimental Procedures). This analysis localized *D. serrata* FLCD to a 708 kb interval on scaffold 2 (Figure S1B).

We then tested SNPs for association with female pigmentation after sequencing the genomes of 100 inbred *D. serrata* lines, which included 86 light, 12 dark, and 2 polymorphic lines (coded as “dark”). After removing three related lines, we tested 3,920,302 SNPs across the genome using the Cochran-Armitage Trend Test and used the Benjamini-Hochberg correction to control the false discovery rate. Five SNPs had a p-value $<10e^{-11}$, but one showed an exceptionally strong association with female pigmentation ($p = 10e^{-22}$). This top SNP occurred within the 708 kb interval on chromosome arm 2R identified by introgression mapping (Figure 3A). At this SNP, all pure-breeding dark lines were fixed for the minor allele and all the light lines were fixed for the major allele. This SNP and one of the other highly associated SNPs 480 bp away are both located in the first intron of the transcription factor *POU domain motif 3* gene (*pdm3*) (Figure 3B). *pdm3* is a member of the class VI POU domain protein family and has known roles in axon targeting and brain development [20,21], but was recently shown to affect abdominal pigmentation as well [10].

Closer inspection of the first intron in *pdm3* in the GWAS panel revealed a mapping bias between light and dark strains in the region between the *pdm3* intronic SNPs. All 86 light lines aligned well to the interval between these SNPs with an average coverage of 23× per line, but dark strains had an average coverage of 0× in this interval. To determine if structural variation could account for this mapping bias, we used reads from a ~5 kb region centered around this region to assemble haplotypes *de novo* for seven homozygous dark and seven homozygous light lines. Within this 5 kb region, we discovered a 786 bp haplotype common to all dark strains; this haplotype was 333 bases longer than the haplotype common to all light strains and to the reference genome. As expected if the light allele is recessive, no reads from light lines mapped to the reconstructed dark allele. We designed primers to amplify and sequence this region from light and dark *D. serrata* females (Figure 3C) and

corroborated the presence of two distinct alleles (Figure 3D). These results confirm the presence of a structural variant within the first intron of *pdm3*.

To study the potential origin of this structural variant, we used a dotplot to compare light and dark allele sequences. The two alleles did not align in either orientation (providing no support for an inversion), but there was weak evidence for four tandem repeats in the dark haplotype (Figure 3E). Using the MEME suite tools [22], we found three significant motifs with E-values < 0.05 in the dark allele that were absent in the light allele. In concordance with the dotplot, these three motifs were found in four tandem sets (Figure 3F), suggesting that duplications and modifications to these motifs might have led to the origin of this structural variant. We used BLAST to search for other instances of these motifs in the *D. serrata* genome; two motifs had top high-scoring segment pairs (E-values of $8e^{-9}$ and $9e^{-20}$) on scaffold 2 immediately proximal to the structural variant that were present in both light and dark flies. We therefore hypothesize that the dark structural variant may have formed via tandem duplications of a flanking precursor motif. Intriguingly, one of these repeated motifs has predicted binding sites for the HOX gene *Abdominal-B* (*Abd-B*) and the major effector of sex differentiation, *doublesex* (*dsx*). While this remains to be tested, co-regulation of *pdm3* by HOX and sexual differentiation genes is a plausible mechanism for establishing female-limited gene expression, as *Abd-B* and *dsx* have been shown to control sexually dimorphic pigmentation in *D. melanogaster* [7,9]. Overall, the divergence in length and sequence between the light and dark haplotypes most likely reduces homologous recombination between them, thereby preserving discrete alleles.

Genetic variation within and between pigmentation-associated haplotypes

The light and dark *D. serrata* haplotypes did not align within the structural variant, but aligned well in the flanking sequences (Figure S2). No excess of genetic divergence (D_{xy}) or nucleotide diversity (π) was observed between the two morphs outside the structural variant boundaries (Figure S3A). Our data therefore lack a signal of elevated divergence that might be expected under long term balancing selection; this is in contrast to FLCD in *D. erecta*, which shows excess polymorphism at the *tan* locus [18]. While it is possible that the high rates of recombination typical for *Drosophila* could erase the linked effects of balancing selection over time – in light of per-bp, per-generation recombination rates averaging roughly 1×10^{-8} [23] and approximately 10 generations per year, we expect on the order of one recombination event every 10 bp per million years – the dark allele may have originated recently. The dark alleles show modestly reduced genetic variation immediately around the structural variant, along with higher linkage disequilibrium (LD) (Figure S3B) and lower nucleotide divergence. Nucleotide diversity (π) was estimated to be 0.007 (95% CI 0.004 -0.009) for the dark allele and 0.015 (95% CI 0.011 - 0.018) for the light allele.

FLCD maps to the *pdm3* locus in three other *montium subgroup* species

We mapped the genetic basis of FLCD in three other species distantly related to *D. serrata*: *D. kikkawai*, *D. leontia*, and *D. burlai*. We performed crosses and found that the dark allele is dominant in *D. leontia*, as it is in *D. kikkawai* [24], *D. burlai* [16] and *D. serrata*. For *D. kikkawai*, we conducted two introgression experiments similar to those for *D. serrata,* performed between light and dark strains from South America (33 generations and from

India (10 generations). We then sequenced the genomes of 30 pooled dark introgression females, and compared allele frequencies of each pool to its corresponding light parental line by mapping reads to the *D. kikkawai* reference genome [25]. To test for SNPs that significantly deviated from the distribution in the parental lines in both experiments, we performed the Cochran-Mantel-Haenszel (CMH) test [26,27]. Two scaffolds, KB459527 and KB459631 of the *D. kikkawai* reference genome, contained an excess of significant SNPs. These scaffolds correspond to a syntenic block of ~460 kb of the *D. melanogaster* chromosome arm 2R, from base coordinates 8,173,480 to 8,631,633 (release 6.2), including the *pdm3* gene. However, the lowest CMH p-values within this peak were found in the upstream region of *pdm3*, rather than in its first intron (Figure 4A, Figure S4A). An outlier SNP was also found on chromosome arm 3L, falling in the intronic region of the gene *CTP synthase (CTPsyn)*. However, neither the introgression profile (only a single SNP), nor the known functions of this gene, nor the Mendelian nature of FLCD support a role of this outlier in the evolution of this trait.

D. leontia is closely related to *D. kikkawai* (Figure 2A), and reciprocal crosses between these two species produce fertile hybrid females but sterile males [28]. We generated *D. leontia* strains that were homozygous for dark and light pigmentation, and introgressed the dominant dark *D. leontia* allele into: (1) the light strain of *D. leontia* (intraspecific introgression) and (2) the light strain of *D. kikkawai* (interspecific introgression), each for 12 generations. Similar to *D. kikkawai*, the same two scaffolds were implicated, overlapping *pdm3* (Figure 4B, Figure S4B).

For *D. burlai*, we generated homozygous dark and light lines from a single polymorphic strain. We introgressed the dominant dark allele into the light background for 10 generations and compared allele frequencies between the light parental line and dark introgressed F_{10} females as well as between light and dark F_5 females with the CMH test. Here again, the most significant CMH peak in *D. burlai* was found in the *pdm3* region (Figure 4C, Figure S4C). Because FLCD is monogenic in *D. burlai*, lesser peaks on 2R and other arms may reflect incomplete synteny between *D. burlai* and the *D. kikkawai* genome that we mapped reads against (*i.e.* loci that are close to *pdm3* in *D. burlai* but not in *D. kikkawai*).

Multiple origins of FLCD alleles

FLCD in the *Drosophila montium* subgroup could be a trans-species polymorphism, caused by shared alleles inherited from a common ancestor or exchanged through gene flow. We tested whether the structural variant found in *D. serrata* was also present in the three other species by amplifying a ~1.5 kb fragment surrounding this variant. We did not detect any size or major sequence differences at this location between color strains in these species. We aligned ~5 kb sequences, obtained from reference genomes or Sanger sequencing, between all four species plus *D. birchii* and *D. bunnanda* (Figure S2), and inferred a phylogeny of this region (Figure 5A). The *D. serrata* dark haplotype grouped with other haplotypes from *D. serrata* and the closely related species (posterior probability 94%) and showed no clustering with dark haplotypes of *D. kikkawai*, *D. leontia,* or *D. burlai* (Figure 5A), suggesting that the structural variant perfectly associated with dark pigmentation in *D. serrata* is unique to this species.

We also looked for shared FLCD-associated SNPs between pairs of species. In the closely related *D. kikkawai* and *D. leontia*, the *pdm3* scaffold (KB459527) showed an excess of shared SNPs that consistently differentiate dark and light morphs in the two species. Consistent with a common genetic basis of FLCD between these two species, a complementation test between their light strains produced all light F_1 females. With the exception of a single SNP in the second intron of *pdm3*, most FLCD-associated SNPs shared between *D. kikkawai* and *D. leontia* occur in the intergenic region upstream of *pdm3*, and some fall in the coding region of the *LRP1* gene further upstream (Figure S4). RNAi knockdown of *LRP1* in *D. melanogaster* did not strongly alter pigmentation (Figure S5), suggesting that this gene is unlikely to be responsible for FLCD.

In contrast, we found no evidence of shared pigmentation-associated SNPs in any pairwise comparison between *D. serrata*, *D. burlai*, and the *D. kikkawai* / *D. leontia* clade. Neither the two SNPs associated with FLCD in *D. serrata*, nor the eleven SNPs that correlated with pigmentation in *D. kikkawai* and *D. leontia*, differed between color morphs in the other species (Figure 5B). *D. burlai* did not share FLCD-associated SNPs with any other species, either. We also checked for shared large indels that differentiate pigmentation morphs in multiple species, but found none. Examining called (shorter) indels, again we found no evidence of common structural variants between species at the *pdm3* locus (Figure S6). The lack of shared FLCD alleles among species is consistent with our finding of low phylogenetic signal for female pigmentation in the *montium* clade (Figure 2B), and suggests that FLCD is unlikely to be a *trans*-species polymorphism inherited from the *montium* subgroup ancestor.

pdm3 represses pigmentation in both sexes in *D. melanogaster*

Of all the genes in the introgressed intervals identified in our mapping experiments, only *pdm3* has been identified as a regulator of abdominal pigmentation in *D. melanogaster* [10,20]. A previous study, Rogers *et al.* [10] were unable to determine whether *pdm3* activates or represses pigmentation, as different RNAi hairpins appeared to have opposite effects on pigmentation. We replicated one of the Rogers et al. crosses and used an additional RNAi hairpin to knock down *pdm3* using *pnr*-GAL4 and *Abd-B*-GAL4 drivers [29], which drive expression along the dorsal midline and in the posterior abdominal segments, respectively. We observed increased pigmentation with all *pdm3* RNAi constructs and drivers (Figure 6 A-D, Figure S5). We also found that mutants homozygous for a *piggyBac* insertion in *pdm3* (PBac {WH}$pdm3^{f00828}$, [30]) had darker abdomens than flies with a precise excision of this transposon (Figure 6 E-H). Conversely, overexpression of *pdm3* in posterior abdominal segments resulted in decreased pigmentation (Figure 6 I-J). Together, these data suggest that *pdm3* acts as a repressor of abdominal pigmentation. Neither of the other genes that we tested in this interval (*Oaz* and *LRP1*) strongly affected pigmentation (Figure S5). Although *LRP1* showed a slight decrease in A6 pigmentation in females with the *Abd-B* driver, it did not show this change with the *pnr* driver, suggesting that the phenotype may be due to genetic background rather than *LRP1*. Thus, *pdm3* is most likely solely responsible for FLCD. It is possible that the *trans*-regulatory landscape of the abdominal epidermis has diverged between the *melanogaster* and *montium* subgroups – such a difference might explain why *pdm3* does not cause FLCD in the *melanogaster* subgroup

[18]. However, *pdm3* appears to be an important part of the pigmentation pathway in both lineages. Molecular studies of the regulatory interactions of *pdm3*, as well as genetic analyses of FLCD in additional *montium* subgroup species, may further illuminate the evolutionary constraint of this trait.

Discussion

pdm3 is a hotspot for female-specific evolutionary changes

To test whether the evolution of female-specific traits follows predictable developmental trajectories, we compared the genetic basis of a female-limited Mendelian dimorphism that has repeatedly evolved over ~19 million years (my) (Figure 2A) in the *Drosophila montium* subgroup. This trait consistently mapped to the same locus, *pdm3*, in four species whose last common ancestor lived ~13.1 my ago. Sequence analysis strongly suggests that *pdm3* alleles responsible for FLCD originated independently at least three times in different lineages. These findings indicate a strong genetic constraint in the evolution of FLCD.

The reasons for this constraint are not clear. *pdm3* functions in brain and olfactory neuron development [20,21] and was recently identified in a large RNAi screen for genes affecting abdominal pigmentation in *D. melanogaster* [10], but it has not been implicated in the evolution of pigmentation in any of the previously studied *Drosophila* species. This is surprising because color pattern variation in other species has been mapped both to transcription factors such as *bab* and *Abd-B* [14,31–36] and pigment synthesis enzymes such as *ebony*, *yellow*, and *tan* [13,14,18,33,35–38]. The Mendelian action and discrete alleles of *pdm3* in *D. serrata* suggest that this transcription factor could act as a binary switch that toggles the expression of multiple pigmentation enzymes to produce either very dark or very light pigmentation, with no intermediates. This property may favor genetic changes in *pdm3* during the evolution of starkly dimorphic, sex-limited phenotypes. *cis*-regulatory mutations that alter transcription factor expression have been observed in other species with sex-limited polymorphisms, for example, in the female-limited orange-blotch phenotype in Lake Malawi cichlids [2]. However, in the only known *Drosophila* species with FLCD outside of the *montium* subgroup – *D. erecta* of the *melanogaster* subgroup – female abdominal pigmentation is regulated by *cis*-regulatory changes in *tan*, not *pdm3* [18], suggesting that it is also possible to evolve discrete phenotypes through direct regulatory changes in a single enzyme.

Within the *montium* clade, *pdm3* could be a hotspot for FLCD evolution because all species share a similar gene regulatory architecture [39]. The last common ancestor of this clade most likely had dark males, but either light or polymorphic females (Figure 2B). If *pdm3* was ancestrally involved in generating sexually dimorphic pigmentation, that sex-specific regulatory architecture could have been co-opted repeatedly for the evolution of sex-limited polymorphism.

Our finding that causative *pdm3* alleles map to different parts of the locus suggest that *pdm3* may have several, partially redundant abdominal enhancers, and that the relative importance of these enhancers may change over evolutionary time as has been found for the *yellow* gene [40]. This raises the possibility that FLCD with distinct light and dark alleles evolved only

once in the last common ancestor of the *montium* subgroup, but the molecular basis of these alleles has subsequently diverged. However, we find no direct support for this scenario. Although phylogenetic analysis cannot recover the ancestral female states (clades IV-VI likelihoods: 27.3% dark, 30.6% light and 42.1% dimorphic), the overall transition rates for female pigmentation are very high (101.7 from dimorphic to light, and 15.6 from dimorphic to dark). Thus, it seems unlikely that FLCD would have been maintained over the full branch length connecting our four studied species. At a minimum, our data show that there are multiple ways to alter the regulation of *pdm3* to modulate sex-specific pigmentation, suggesting that strong constraint at the gene level masks greater flexibility at the level of *cis*-regulatory sequences.

Unique *pdm3* alleles associated with FLCD may reflect the fact that there are multiple ways to attain sex-specificity. Often, gains and losses of sexual dimorphism involve the fixation of sex-limited or sexually monomorphic alleles during species divergence [14]. For instance, gain and loss of Doublesex binding sites in the *desatF* gene is associated with evolutionary transitions between sexually dimorphic and sexually monomorphic cuticular hydrocarbon profiles [41]. In contrast, FLCD can be explained by long-term segregation of dimorphic and monomorphic alleles within a single species. We hypothesize that transitions between canonical sexual dimorphism and sex-limited polymorphism occur via changes in the sex-specificity of existing alleles. Different types of regulatory changes – inactivation of a female-specific allele in females, activation of a male-specific allele in females, etc. – can all result in FLCD as long as both ancestral and derived alleles are maintained in natural populations. This flexibility may explain both the frequent occurrence of FLCD in the *montium* subgroup, and the fact that it was arrived at through different sequence changes.

Discrete functional alleles may be maintained by structural variation

The divergent *pdm3* alleles associated with FLCD in *D. serrata* suggest a role for structural variation in maintaining alleles with widely disparate phenotypes within species. Chromosomal inversions have been found to preserve complex mimicry-associated alleles in *Heliconius* and *Papilio* butterflies [5,6,42,43], social morphs in fire ants [44] and male morphological/behavioral syndromes in ruffs [45] due to their ability to suppress recombination. However, we do not find evidence for inversions spanning the *pdm3* locus in any of the *montium* species. In *D. serrata*, we suspect that the structural polymorphism associated with FLCD contains *cis*-regulatory sequences that form light and dark alleles, and that these alleles are protected from recombination due to the lack of sequence homology between the light and dark haplotypes. If confirmed, this would be the first case of a fragment polymorphism (the existence of two alternative nonhomologous sequences) that maintains a sex-limited polymorphism.

Potential fitness consequences of light and dark *pdm3* alleles

Multiple abdominal color morphs may be maintained if they experience fitness trade-offs. Variation in abdominal pigmentation in flies has been associated with desiccation tolerance [46], UV resistance [47], and thermal adaptation [48]. However, environmental causes alone might not explain why FLCD is sex-specific. Alternatively, dark and light female morphs could reflect different reproductive strategies. Payant [17] found evidence for frequency-

dependent sexual selection on FLCD morphs in *D. erecta*: at some morph frequencies, male-resembling females had lower rates of mating, which may be costly in that species [49]. Thus far, there has been mixed support for balancing selection in several *montium* species [24,50–55]. Fitness experiments and population genetic analyses in multiple species with FLCD will be needed to address the adaptive significance of this trait.

Conclusion

Our study establishes the *Drosophila montium* subgroup as a prime model for investigating the molecular basis of sexual dimorphism and sex-limitation. Each species or lineage with an independent origin of FLCD can serve as a natural replication of the same evolutionary experiment. The segregation of sexually dimorphic and monomorphic alleles within species offers a rare glimpse at the earliest steps in the gain (or loss) of sex-specific phenotypes. Using unbiased mapping approaches, we have shown that FLCD has evolved independently at least three times in the *montium* clade through parallel non-coding changes in *pdm3*, presenting a remarkable example of genetic predictability. Interestingly, *pdm3* has not been implicated in the evolution of male-specific or sexually monomorphic color patterns in any other *Drosophila* species examined to date, suggesting that although the evolution of FLCD is genetically biased within the *montium* subgroup, the genetic architecture of sexually dimorphic pigmentation may differ between clades and possibly between sexes as well.

Experimental Procedures

Phylogenetic analysis

We reconstructed a molecular phylogeny for 44 *montium* subgroup species from concatenated sequences of three nuclear genes (*Adh*, *Amy1*, and *Amyrel*) and one mitochondrial (*COII*) gene using MrBayes [56]. We then estimated the phylogenetic signal, evolutionary correlation, and state transition rates for male and female abdominal pigmentation using BayesTraits [57] (see Supplemental Experimental Procedures).

Introgression mapping in *D. serrata, D. kikkawai, D. leontia* and *D. burlai*

We established homozygous light and dark strains and generated intraspecific introgressions for each species by backcrossing heterozygous dark females to males of the light parental strain for 10-33 generations. For the one interfertile species pair, *D. kikkawai* and *D. leontia*, we also introgressed the dark *D. leontia* allele into a light *D. kikkawai* strain. We mapped introgression breakpoints in *D. serrata* using 79 genotyping markers. For the other three species, all introgression and parental lines were sequenced. We extracted DNA from pools of 30 females per line using a phenol-chloroform extraction and sequenced on the Illumina HiSeq 2000 platform with paired end 100 bp reads. SNPs with significantly different allele frequencies were detected using the Cochran-Mantel-Haenszel (CMH) test implemented in PoPoolation2 [58].

Association mapping and detection of a structural variant in *D. serrata*

We mapped female pigmentation in a panel of 100 inbred lines. We extracted DNA from 30 virgin females/line using a phenol-chloroform extraction and sequenced each line with 100

bp paired end reads on the Illumina HiSeq platform to ~15× coverage. We aligned reads from each line to the *D. serrata* genome [19] using bwa mem [59] with soft clipping enabled and indels were realigned using the Genome Analysis Tool Kit [60,61]. Genotypes were inferred using the Joint Genotyper for Inbred Lines [62]. We used the Cochran-Armitage Trend Test in PLINK [63] to test for single-SNP associations. We detected missing reads between the top pigmentation-associated SNP and a neighboring associated SNP in the first *pdm3* intron, and reconstructed haplotypes from seven dark and seven light lines within this region using the Velvet assembler [64]. We verified these alleles using PCR and Sanger sequencing (see Supplemental Experimental Procedures). We estimated nucleotide diversity (π), genetic divergence (D_{xy}), and linkage disequilibrium (r^2) using the pegas package in R [65], MEGA [66], and PLINK. We constructed an allele tree with orthologous regions from light and dark *D. serrata*, *D. kikkawai*, *D. leontia*, *D. burlai, D. bunnanda,* and *D. birchii* using MrBayes.

Functional tests in *D. melanogaster*

We evaluated the role of *pdm3* in cuticular pigmentation using the GAL4/UAS system to either knock down *pdm3* (UAS-*pdm3*-RNAi) or ectopically express it (UAS-*pdm3*) in *D. melanogaster*. We knocked down *pdm3* and two nearby genes (*Oaz* and *LRP1*) using Harvard TRiP stocks and *pannier*-GAL4 and *Abdominal-B*-GAL4 drivers. We also drove increased UAS-*pdm3* expression using a stock that excludes the 9^{th} exon (which can be alternatively skipped in *D. melanogaster*) (gift from J. Carlson). In addition, we compared cuticular pigmentation between hypomorphic *pdm3* mutants (PBac {WH}$pdm3^{f008}$, formed by a *piggyBac* transposon insertion between exons 8 and 9) and strains with a precise excision of the *piggyBac* insertion. Cuticle mounts were prepared following Duncan [67].

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank J. Carlson for UAS-*pdm3* and precise *piggyBac* excision stocks, F. Karch for the *Abd-B-GAL4* stock, and the UC-San Diego *Drosophila* Species Stock Center, Bloomington *Drosophila* Stock Center, and TRiP at Harvard Medical School (NIH/NIGMS R01-GM084947) for providing other fly stocks. We also thank Nicole Bangart, Olga Barmina, Rachael Curtis, Ondrej Podlaha, and Romai Sebhatu for help with DNA extraction and library preparation, and Junhao Chen for assistance with cuticle preparation. This work was supported by the Richard Lounsbery Foundation (J.E.P. and J.R.D.), NIH grant R01 GM111797 (to J.E.P.), and NSF IOS-1256420 (to A.K).

References

1. Darwin, C. The Descent of Man. Appleton and Co; 1871.
2. Roberts RB, Ser JR, Kocher TD. Sexual conflict resolved by invasion of a novel sex determiner in Lake Malawi cichlid fishes. Science. 2009; 326:998–1001. [PubMed: 19797625]
3. Loehlin DW, Oliveira DCSG, Edwards R, Giebel JD, Clark ME, Cattani MV, Zande L, Verhulst EC, Beukeboom LW, Muñoz-Torres M, et al. Non-coding changes cause sex-specific wing size differences between closely related species of *Nasonia*. PLoS Genet. 2010; 6:e1000821. [PubMed: 20090834]
4. Tanaka K, Barmina O, Sanders LE, Arbeitman MN, Kopp A. Evolution of sex-specific traits through changes in HOX-dependent *doublesex* expression. PLoS Biol. 2011; 9:e1001131. [PubMed: 21886483]

5. Kunte K, Zhang W, Tenger-Trolander A, Palmer DH, Martin A, Reed RD, Mullen SP, Kronforst MR. *doublesex* is a mimicry supergene. Nature. 2014; 507:229–232. [PubMed: 24598547]
6. Nishikawa H, Iijima T, Kajitani R, Yamaguchi J, Ando T, Suzuki Y, Sugano S, Fujiyama A, Kosugi S, Hirakawa H, et al. A genetic mechanism for female-limited Batesian mimicry in *Papilio* butterfly. Nat Genet. 2015; 47:405–409. [PubMed: 25751626]
7. Kopp A, Duncan I, Carroll SB. Genetic control and evolution of sexually dimorphic characters in *Drosophila*. Nature. 2000; 408:553–559. [PubMed: 11117736]
8. Wittkopp PJ, Carroll SB, Kopp A. Evolution in black and white: genetic control of pigment patterns in *Drosophila*. Trends Genet. 2003; 19:495–504. [PubMed: 12957543]
9. Williams TM, Selegue JE, Werner T, Gompel N, Kopp A, Carroll SB. The regulation and evolution of a genetic switch controlling sexually dimorphic traits in *Drosophila*. Cell. 2008; 134:610–623. [PubMed: 18724934]
10. Rogers WA, Grover S, Stringer SJ, Parks J, Rebeiz M, Williams TM. A survey of the *trans*-regulatory landscape for *Drosophila melanogaster* abdominal pigmentation. Dev Biol. 2014; 385:417–432. [PubMed: 24269556]
11. Gompel N, Carroll SB. Genetic mechanisms and constraints governing the evolution of correlated traits in drosophilid flies. Nature. 2003; 424:931–935. [PubMed: 12931186]
12. Jeong S, Rebeiz M, Andolfatto P, Werner T, True J, Carroll SB. The evolution of gene regulation underlies a morphological difference between two *Drosophila* sister species. Cell. 2008; 132:783–793. [PubMed: 18329365]
13. Camino EM, Butts JC, Ordway A, Vellky JE, Rebeiz M, Williams TM. The evolutionary origination and diversification of a dimorphic gene regulatory network through parallel innovations in *cis* and *trans*. PLoS Genet. 2015; 11:e1005136. [PubMed: 25835988]
14. Signor SA, Liu Y, Rebeiz M, Kopp A. The genetic toolkit of convergent evolution: comparative genetics of male-specific pigmentation in multiple Drosophila species. This issue.
15. Yassin A. Phylogenetic classification of the Drosophilidae Rondani (Diptera): the role of morphology in the postgenomic era. Syst Entomol. 2013; 38:349–364.
16. Ohnishi S, Watanabe TK. Genetic analysis of color dimorphism in the *Drosophila montium* subgroup. Jpn J Genet. 1985; 60:355–358.
17. Payant V. Le polymorphisme de coloration abdominale de *Drosophila erecta* est-il gouverné par un gene sélectivement neutre? Genet Sel Evol. 1988; 20:37–50. [PubMed: 22879305]
18. Yassin A, Bastide H, Chung H, Veuille M, David JR, Pool JE. Ancient balancing selection at *tan* underlies female colour dimorphism in *Drosophila erecta*. Nat Commun. 2016; 7:10400. [PubMed: 26778363]
19. Chenoweth SF, Appleton NC, Allen SL, Rundle HD. Genomic evidence that sexual selection impedes adaptation to a novel environment. Curr Biol. 2015; 25:1860–1866. [PubMed: 26119752]
20. Tichy AL, Ray A, Carlson JR. A new *Drosophila* POU gene, *pdm3*, acts in odor receptor expression and axon targeting of olfactory neurons. J Neurosci. 2008; 28:7121–7129. [PubMed: 18614681]
21. Chen CK, Chen WY, Chien CT. The POU-domain protein Pdm3 regulates axonal targeting of R neurons in the *Drosophila* ellipsoid body. Dev Neurobiol. 2012; 72:1422–1432. [PubMed: 22190420]
22. Bailey TL, Elkan C. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Int Conf Intell Syst Mol Biol. 1994; 2:28–36.
23. Comeron JM, Ratnappan R, Bailin S. The many landscapes of recombination in *Drosophila melanogaster*. PLoS Genet. 2012; 8:e1002905. [PubMed: 23071443]
24. Freire-Maia N. Balanced polymorphism in *Drosophila montium*. Evolution. 1949; 3:98.
25. Chen ZX, Sturgill D, Qu J, Jiang H, Park S, Boley N, Suzuki AM, Fletcher AR, Plachetzki DC, FitzGerald PC, et al. Comparative validation of the *D. melanogaster* modENCODE transcriptome annotation. Genome Res. 2014; 24:1209–1223. [PubMed: 24985915]
26. Cochran WG. Some methods for strengthening the common $\chi 2$ tests. Biometrics. 1954; 10:417–451.

27. Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. J Natl Cancer Inst. 1959; 22:719–748. [PubMed: 13655060]
28. David JR, Lemeunier F, Tsacas L. Hybridization and genetic comparison of the subcosmopolitan species *Drosophila kikkawai* with its new sibling species *D. leontia* (Diptera, Drosophilidae). Egypt J Genet Cytol. 1978; 7:28–39.
29. Gligorov D, Sitnik JL, Maeda RK, Wolfner MF, Karch F. A novel function for the Hox gene *Abd-B* in the male accessory gland regulates the long-term female post-mating response in *Drosophila*. PLOS Genet. 2013; 9:e1003395. [PubMed: 23555301]
30. Thibault ST, Singer MA, Miyazaki WY, Milash B, Dompe NA, Singh CM, Buchholz R, Demsky M, Fawcett R, Francis-Lang HL, et al. A complementary transposon tool kit for *Drosophila melanogaster* using *P* and *piggyBac*. Nat Genet. 2004; 36:283–287. [PubMed: 14981521]
31. Kopp A, Graze RM, Xu S, Carroll SB, Nuzhdin SV. Quantitative trait loci responsible for variation in sexually dimorphic traits in *Drosophila melanogaster*. Genetics. 2003; 163:771–787. [PubMed: 12618413]
32. Bickel RD, Kopp A, Nuzhdin SV. Composite effects of polymorphisms near multiple regulatory elements create a major-effect QTL. PLoS Genet. 2011; 7:e1001275. [PubMed: 21249179]
33. Bastide H, Betancourt A, Nolte V, Tobler R, Stöbe P, Futschik A, Schlötterer C. A genome-wide, fine-scale map of natural pigmentation variation in *Drosophila melanogaster*. PLoS Genet. 2013; 9:e1003534. [PubMed: 23754958]
34. Rogers WA, Salomone JR, Tacy DJ, Camino EM, Davis KA, Rebeiz M, Williams TM. Recurrent modification of a conserved *cis*-regulatory element underlies fruit fly pigmentation diversity. PloS Genet. 2013; 9:e1003740. [PubMed: 24009528]
35. Dembeck LM, Huang W, Magwire MM, Lawrence F, Lyman RF, Mackay TFC. Genetic architecture of abdominal pigmentation in *Drosophila melanogaster*. PLoS Genet. 2015; 11:e1005163. [PubMed: 25933381]
36. Massey JH, Wittkopp PJ. The genetic basis of pigmentation differences within and between *Drosophila* species. Curr Topics Dev Biol. 2016; 119:27–61.
37. Rebeiz M, Pool JE, Kassner VA, Aquadro CF, Carroll SB. Stepwise modification of a modular enhancer underlies adaptation in a *Drosophila* population. Science. 2009; 326:1663–1667. [PubMed: 20019281]
38. Johnson WC, Ordway AJ, Watada M, Pruitt JN, Williams TM, Rebeiz M. Genetic changes to a transcriptional silencer element confers phenotypic diversity within and between *Drosophila* species. PLoS Genet. 2015; 11:e1005279. [PubMed: 26115430]
39. Conte GL, Arnegard ME, Peichel CL, Schluter D. The probability of genetic parallelism and convergence in natural populations. Proc R Soc B. 2012; 279:5039–5047.
40. Kalay G, Wittkopp PJ. Nomadic enhancers; tissue-specific *cis*-regulatory elements of *yellow* have divergent genomic positions among *Drosophila* species. PLOS Genet. 2010; 6:e1001222. [PubMed: 21151964]
41. Shirangi TR, Dufour HD, Williams TM, Carroll SB. Rapid evolution of sex pheromone-producing enzyme expression in *Drosophila*. PLOS Biol. 2009; 7:e1000168. [PubMed: 19652700]
42. Joron M, Papa R, Beltrán M, Chamberlain N, Mavárez J, Baxter S, Abanto M, Bermingham E, Humphray SJ, Rogers J, et al. A conserved supergene locus controls colour pattern diversity in *Heliconius* butterflies. PLoS Biol. 2006; 4:e303. [PubMed: 17002517]
43. Joron M, Frezal L, Jones RT, Chamberlain NL, Lee SF, Haag CR, Whibley A, Becuwe M, Baxter SW, Ferguson L, et al. Chromosomal rearrangements maintain a polymorphic supergene controlling butterfly mimicry. Nature. 2011; 477:203–206. [PubMed: 21841803]
44. Wang J, Wurm Y, Nipitwattanaphon M, Riba-Grognuz O, Huang YC, Shoemaker D, Keller L. A Y-like social chromosome causes alternative colony organization in fire ants. Nature. 2013; 493:664–668. [PubMed: 23334415]
45. Lamichhaney S, Fan G, Widemo F, Gunnarsson U, Thalmann DS, Hoeppner MP, Kerje S, Gustafson U, Shi C, Zhang H, et al. Structural genomic changes underlie alternative reproductive strategies in the ruff (*Philomachus pugnax*). Nat Genet. 2016; 48:84–88. [PubMed: 26569123]

46. Parkash R, Rajpurohit S, Ramniwas S. Changes in body melanisation and desiccation resistance in highland vs. lowland populations of D melanogaster. J Insect Physiol. 2008; 54:1050–1056. [PubMed: 18519137]
47. Bastide H, Yassin A, Johanning EJ, Pool JE. Pigmentation in *Drosophila melanogaster* reaches its maximum in Ethiopia and correlates most strongly with ultra-violet radiation in sub-Saharan Africa. BMC Evol Biol. 2014; 14:179. [PubMed: 25115161]
48. Gibert P, Moreteau B, Moreteau JC, Parkash R, David JR. Light body pigmentation in indian *Drosophila melanogaster*: A likely adaptation to a hot and arid climate. J Genet. 1998; 77:13–20.
49. Yassin A, Orgogozo V. Coevolution between male and female genitalia in the *Drosophila melanogaster* species subgroup. PLoS ONE. 2013; 8:e57158. [PubMed: 23451172]
50. Oshima C. Genetic analysis of the dimorphism of color pattern in *D. rufa*. Inf Serv. 1952; 26:116.
51. Freire-Maia N, Freire-Maia A. Segregational load in *Drosophila kikkawai*. III Natural populations. Genetics. 1964; 50:789–802. [PubMed: 14239768]
52. Lee TJ. Analysis of heterosis in populations of *D. auraria*. Drosoph Inf Serv. 1966; 41:120.
53. Gibert P, Moreteau B, David JR. Phenotypic plasticity of abdominal pigmentation in *Drosophila kikkawai*: multiple interactions between a major gene, sex, abdomen segment and growth temperature. Genetica. 1999; 105:165–176. [PubMed: 10568260]
54. Costa BEP, Rohde C, Valente VLS. Temperature, urbanization and body color polymorphism in South Brazilian populations of *Drosophila kikkawai* (Diptera, Drosophilidae). Iheringia Sér Zool. 2003; 93:381–393.
55. Parkash R, Sharma V, Kalra B. Sexual dimorphism for water balance mechanisms in montane populations of *Drosophila kikkawai*. Biol Lett. 2010; 6:570–574. [PubMed: 20106858]
56. Ronquist F, Teslenko M, Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. Syst Biol. 2012; 61:539–542. [PubMed: 22357727]
57. Pagel, MD., Meade, A. BayesTraits version 1.0 computer package. Software. 2007. avalaible from: http://www.evolution.reading.ac.uk/BayesTraits.html. 2007
58. Kofler R, Pandey RV, Schlötterer C. PoPoolation2: identifying differentiation between populations using sequencing of pooled DNA samples (Pool-Seq). Bioinformatics. 2011; 27:3435–3436. [PubMed: 22025480]
59. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. ArXiv Prepr ArXiv13033997. 2013
60. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, et al. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010; 20:1297–1303. [PubMed: 20644199]
61. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, Philippakis AA, Angel G, Rivas MA, Hanna M, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet. 2011; 43:491–498. [PubMed: 21478889]
62. Stone EA. Joint genotyping on the fly: Identifying variation among a sequenced panel of inbred lines. Genome Res. 2012; 22:966–974. [PubMed: 22367192]
63. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, Bakker PIW, Daly MJ, et al. PLINK: A tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007; 81:559–575. [PubMed: 17701901]
64. Zerbino DR, Birney E. Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. Genome Res. 2008; 18:821–829. [PubMed: 18349386]
65. Paradis E. pegas: an R package for population genetics with an integrated–modular approach. Bioinformatics. 2010; 26:419–20. [PubMed: 20080509]
66. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol. 2013; 30:2725–2729. [PubMed: 24132122]
67. Duncan IM. *Polycomblike*: A gene that appears to be required for the normal expression of the *Bithorax* and *Antennapedia* gene complexes of *Drosophila melanogaster*. Genetics. 1982; 102:49–70. [PubMed: 6813190]

68. Russo CAM, Mello B, Frazão A, Voloch CM. Phylogenetic analysis and a time tree for a large drosophilid data set (Diptera: Drosophilidae). Zool J Linn Soc. 2013; 169:765–775.

Highlights

* Female-limited color dimorphism exists in >20 *Drosophila montium* subgroup species.

* In four species, this trait maps near the same transcription factor gene, *pdm3*.

* The trait appears to have originated three independent times via the same locus.

* Repeated gene use indicates a bias in the evolution of female-specific dimorphism.

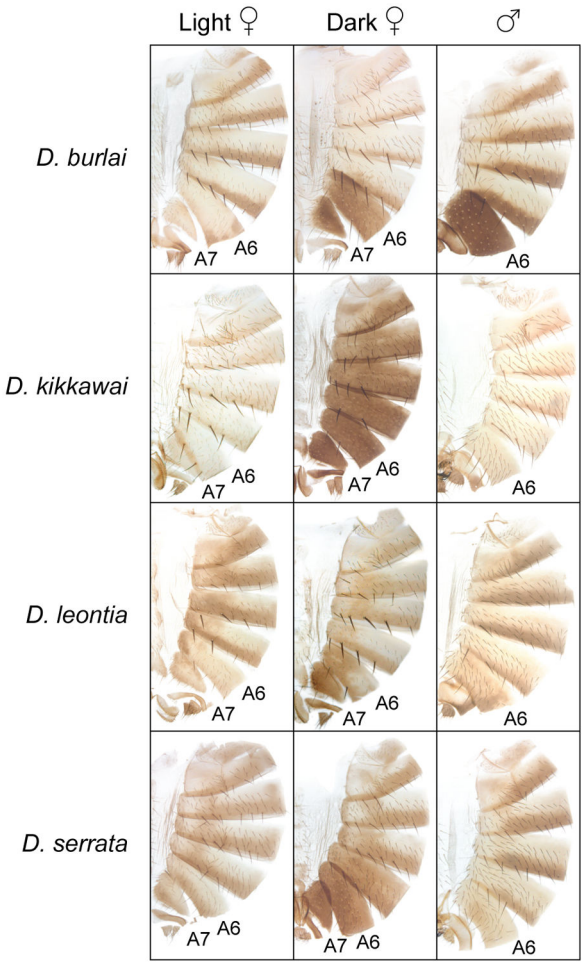

Figure 1.

(See also Table S1) Female-limited color dimorphism in four *Drosophila montium* subgroup species. Females have light or dark posterior abdominal segments (A6 and A7), but males of each species are invariant for pigmentation of their last abdominal segment (A6).

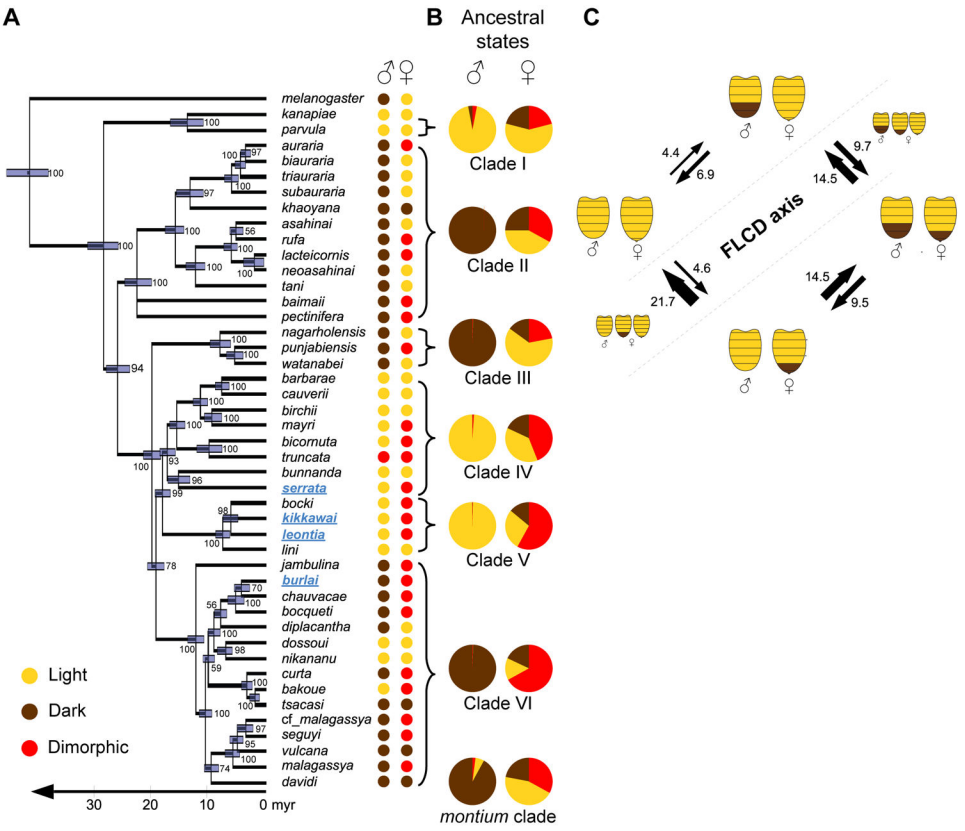

Figure 2.

(See also Table S1) Phylogenetic analysis of FLCD in the *Drosophila montium* subgroup. (A) Bayesian tree for 44 species inferred from four genes (*Adh*, *Amy1*, *Amyrel*, and *COII*). Male and female abdominal pigmentation is represented by colored dots (brown = dark morph, yellow = light morph, and red = dimorphic). Gray bars at nodes show the 95% confidence intervals for the estimated divergence times assuming *D. melanogaster* and *montium* species diverged 28 my ago [68]. The four species in our mapping analysis are underlined in blue. (B) Probabilities of ancestral pigmentation for the *montium* subgroup (bottom) and each major clade. (C) Inferred transition rates (the number of times a continuous-time Markov chain moves among states relative to branch lengths) between different male and female abdominal pigmentation phenotypes. A transition from one female state to another (dark → light or light → dark) requires a transitional state where light and dark alleles segregate within species so that the females are color-dimorphic until the new allele is fixed; this is represented by the FLCD axis. Arrow thickness reflects the frequency of transitions.

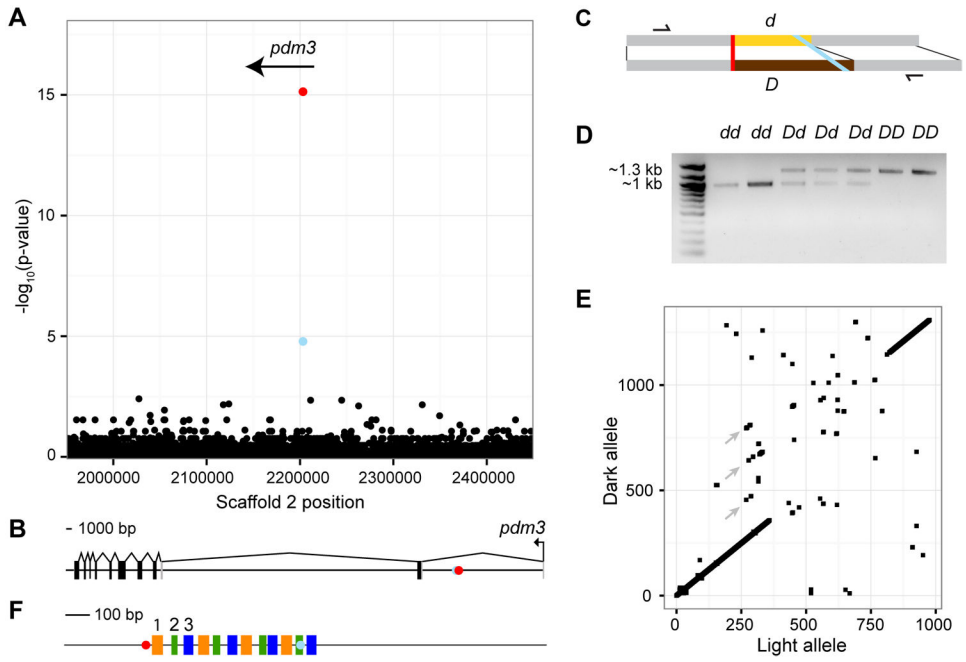

Figure 3.

(See also Figures S1, S2, S3) Pigmentation-associated variants in *D. serrata* map to the first intron of *pdm3*. (A) Association results for pigmentation in 97 inbred lines in the 708 kb interval identified through backcross and introgression mapping. One SNP (red dot) was significantly associated with female color and another SNP 480 bases away (blue dot) was weakly associated. (B) The top two SNPs fall within the first predicted intron of *pdm3*. Putative *pdm3* coding and non-coding exon boundaries in *D. serrata* were determined by reciprocal BLAST between the *D. serrata pdm3*-containing scaffold and *D. melanogaster* exons. Non-coding exons are represented with gray boxes and coding exons are represented with black boxes. *pdm3* is 69 kb long in *D. melanogaster* and is predicted to span approximately 75 kb in the *D. serrata* genome assembly (1000 bp scale bar for reference). (C) Cartoon of the structural fragment polymorphism between light (*d*) and Dark (*D*) alleles relative to the two FLCD-associated SNPs (red and blue lines). PCR amplification primers are shown. (D) PCR reveals two fragments perfectly correlated with genotype. From left to right, we PCR amplified the haplotypes from females from two light strains (*dd*: 681.3L and light FORS4), three heterozygotes (*Dd*: 681.3L (x2640) × 681.5D (x2642), dark FORS4 (x2640) × light FORS4 (x2642), and light FORS4 (x2640)x dark FORS4 (x2642)), and two dark strains (*DD*: 681.5D and dark FORS4). The light allele is shorter (453 bp) and the dark allele is longer (786 bp). (E) A dot plot of *de novo* assembled light (horizontal) and dark (vertical) haplotypes shows that flanking sequences are syntenic but the two haplotypes lack homology in the intervening region. Arrows show repeated sequence motifs in the dark haplotype. (F) A schematic representation of repeated sequence motifs found in the dark haplotype.

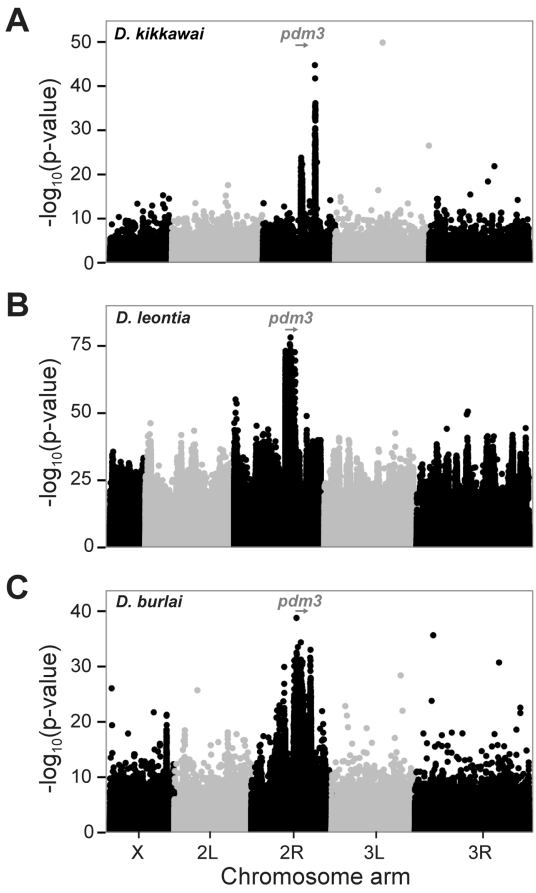

Figure 4.

(See also Figures S4, S6) FLCD in three other *montium* species maps to the *pdm3* locus. Manhattan plot for the Cochran-Mantel-Haenszel (CMH) test comparing SNP allele frequencies between introgression lines and their corresponding light parental strains for (A) *D. kikkawai* (two intraspecific introgression experiments), (B) *D. leontia* (one intraspecific introgression and one interspecific experiment), and (C) *D. burlai* (one intraspecific experiment). Scaffolds are arranged in descending order by size and colored according to the corresponding *D. melanogaster* chromosome arm, except for KB459631 and KB459527, which were concatenated according to their synteny.

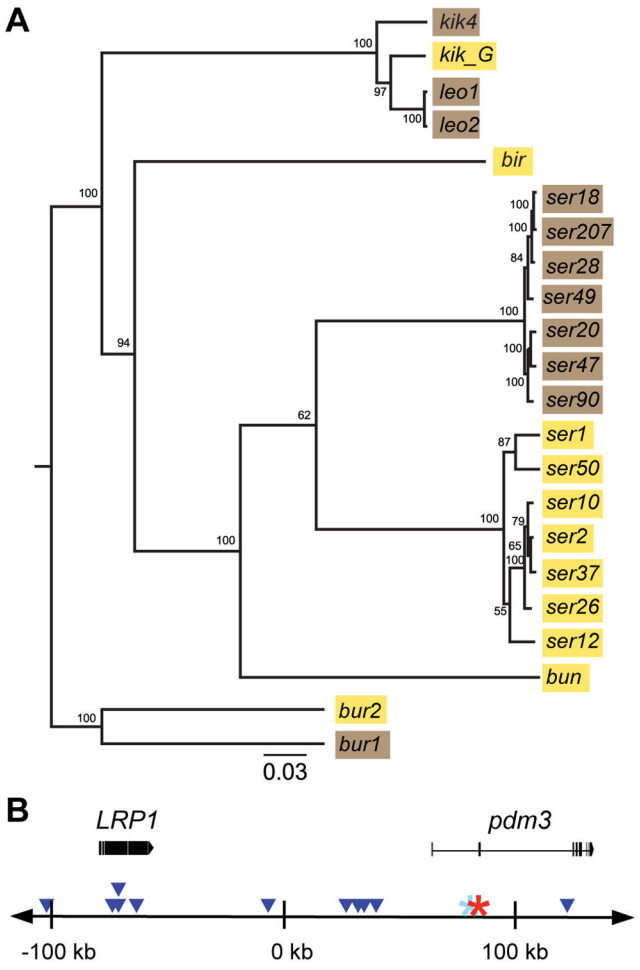

Figure 5.

(See also Figure S2) Different variants are associated with FLCD in four *montium* subgroup species. (A) Allele tree for a ~1 kb region surrounding the *D. serrata* structural variant in the first intron of *pdm3*. Sequences from light (shaded in yellow) and dark (shaded in brown) strains were amplified by PCR or extracted from reference genomes for *D. serrata* (*ser*), *D. kikkawai* (*kik*), *D. leontia* (*leo*), *D. burlai* (*bur*), *D. bunnanda* (*bun*), and *D. birchii* (*bir*). The *D. serrata* alleles form two clades by color morph, and do not group with sequences from other species. (B) FLCD-associated SNPs shared between *D. kikkawai* and *D. leontia* (red triangles) occur in the intergenic region upstream of *pdm3*, in contrast to the FLCD-associated SNPs in *D. serrata* that occur in the first intron of *pdm3* (blue asterisks). Positive and negative coordinates refer to scaffolds KB459527 and KB459631, respectively.

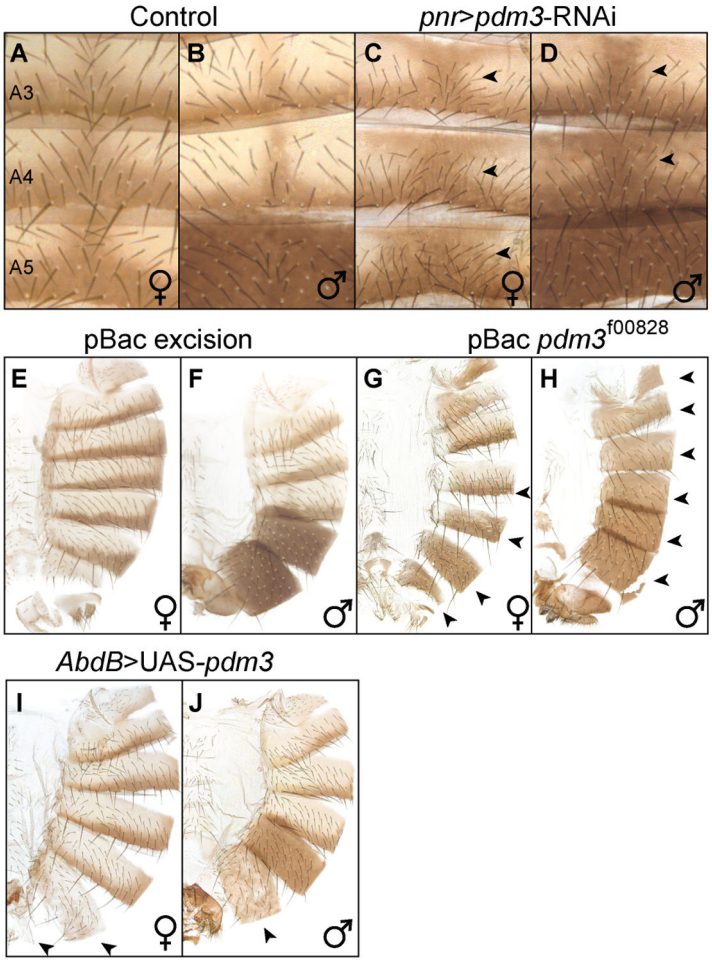

Figure 6.

(See also Figure S5) *pdm3* acts as a pigmentation repressor in *D. melanogaster*. (C,D) RNAi knockdown of *pdm3* (Bloomington *Drosophila* Stock Center #53887) using a *pnr*-GAL4 driver increased pigmentation along the dorsal midline (arrowheads) in females (n = 3) and males (n = 7) relative to control females (n = 3) and males (n = 4) lacking *pnr*-GAL4 (A,B). (G,H) Hypomorphic mutants where *pdm3* is disrupted by a *piggyBac* transposon insertion between exons 8 and 9 (PBac {WH}*pdm3*^{f00828}) have increased dark pigmentation throughout all abdominal segments (arrowheads) in females (n = 1) and males (n = 6) relative to females (n = 8) and males (n = 5) with a precise excision of the *piggyBac* transposon (E, F). (I,J) Ectopic expression of a short *pdm3* isoform (without exon 9) driven with *AbdB*-GAL4 and *tub*-GAL80^{ts} results in patchy, decreased pigmentation of posterior segments (arrowheads) in females (n = 3) and males (n = 6). Expression of a long *pdm3* isoform (including exon 9) produces similar phenotypes (not shown).