Quantitative Trait Loci Affecting The Difference in Pigmentation
Between *Drosophila yakuba* and *D. santomea*

MARY ANNA CARBONE,* 1, ANA LLOPART,†1,2 MATTHEW deANGELES,†,3 JERRY A. COYNE† AND TRUDY F. C. MACKAY*

*Department of Genetics, North Carolina State University, Raleigh, North Carolina 27695

†Department of Ecology and Evolution, University of Chicago, Chicago, Illinois 60637

1 The first two authors contributed equally to this work

2 Present address: Department of Biological Sciences, The University of Iowa, Iowa City, IA 52242

3 Present address: Pediatric Intensive Care Unit, Lucile Packard Children’s Hospital

725 Welch Road, Palo Alto, CA 94304
Running head: QTL for Drosophila pigmentation

Keywords: Speciation, species differences, composite interval mapping

Corresponding author: Mary Anna Carbone
Department of Genetics
Campus Box 7614
North Carolina State University
Raleigh, NC 27695
Tel: 919-515-5789
Fax: 919-515-3355
E-mail: macarbon@unity.ncsu.edu
ABSTRACT
Using quantitative trait locus (QTL) mapping, we studied the genetic basis of the difference in pigmentation between two sister species of Drosophila: D. yakuba, which, like other members of the D. melanogaster subgroup, shows heavy black pigmentation on the abdomen of males and females, and D. santomea, an endemic to the African island of São Tomé, which has virtually no pigmentation. Here we mapped four QTLs with large effects on this interspecific difference in pigmentation: two on the X chromosome, and one each on the second and third chromosomes. The same four QTLs were detected in male hybrids in the backcrosses to both D. santomea and D. yakuba, and in the female D. yakuba backcross hybrids. All four QTLs exhibited strong epistatic interactions in male backcross hybrids, but only one pair of QTLs interacted in females from the backcross to D. yabuka. All QTLs from each species affected pigmentation in the same direction, consistent with adaptive evolution driven by directional natural selection. The regions delimited by the QTLs included many positional candidate loci in the pigmentation pathway, including genes affecting catecholamine biosynthesis, melanization of the cuticle, and many additional pleiotropic effects.
INTRODUCTION

During the Modern Synthesis, the dominant view of the genetics of species differences was that of Ronald Fisher (1930), who believed that such differences were almost invariably due to the accumulation of many genes, each of small phenotypic effect. Tests of this proposition, however, were limited by the lack of genetic markers in most crossable but differentiated species, although some data suggested that species differences could occasionally be due to genes of large effect (Orr and Coyne 1992).

Recently, however, the advent of molecular techniques has improved our ability to study the genetics of species differences. Two innovations have been crucial. First, quantitative trait locus (QTL) mapping enables us to localize genes responsible for species differences by determining their association with molecular markers at known sites. Second, molecular techniques such as germline transformation enable us to determine directly whether a candidate gene affects a species difference. Orr (2001) describes these innovations and the results of recent genetical studies using them. Although most data derive from a small number of organisms (Orr’s study describes only 13 analyses, six from Drosophila and four from the monkeyflower genus Mimulus), the results show that while differences in traits between species can be polygenic, genes of large effect are involved more frequently than previously suspected.

Here we describe a genetic analysis of a striking character difference — the degree of abdominal pigmentation — between two sister species of Drosophila: D. yakuba and D. santomea.

D. yakuba is widely distributed in open habitats across sub-Saharan Africa and the islands near the continent (including Madagascar). In contrast, D. santomea is endemic to the 860 km² volcanic island of São Tomé, in the Gulf of Benin 255 km west of the coast of
Gabon (LACHAISE et al. 2000). On the mountain Pico do São Tomé, D. yakuba occurs at elevations below 1450 m, while D. santomea occupies the mist forests at elevations between 1153 and 2024 m. (D. yakuba is also widespread throughout lowland São Tomé, probably as a result of a secondary invasion after the initial common ancestor evolved into D. santomea.) Between about 1100 and 1450 m in elevation, the ranges of the two species overlap, with the ratio of D. yakuba/D. santomea shifting from 2:1 to 1:20 as one moves upward through this zone. The species show substantial sexual isolation when tested in the laboratory (LACHAISE et al. 2000), and, using morphological criteria, one finds a low frequency (ca. 1%) of hybrids in the zone of overlap.

Molecular evidence puts the divergence between D. yakuba and D. santomea at about 400,000 years ago (LLOPART et al. 2002). In interspecific crosses, F₁ male hybrids are sterile but female hybrids are fertile, and thus can be crossed to either parental species (LACHAISE et al. 2000; CARIOU et al. 2001). This fertility permits genetic analysis using backcross individuals.

The diagnostic differences between the species include male genital morphology and sex-comb tooth number (LACHAISE et al. 2000; COYNE et al. 2004), but the most striking difference involves abdominal pigmentation. Among the nine species in the D. melanogaster subgroup, eight of them, including D. yakuba, have similar patterns of dark pigmentation: males possess thin black stripes along the posterior portions of tergites 2, 3, and 4, while tergites 5-7 are completely black. Females of these species have stripes along the posterior portions of all tergites, and tergites 5-7 show substantial but not complete black pigmentation. In contrast, D. santomea males show virtually no pigmentation, and females show only very light striping on the posterior parts of tergites 2-5, with no pigmentation on other tergites.
(Photographs of these differences are given in Figures 1 and 2 of LACHAISE et al. 2000 and Fig. 1 of LLOPART et al. 2002). Given the dark pigmentation in all but one species in the subgroup, including the outgroup species D. orena and D. erecta, it is nearly certain that the absence of dark pigmentation in D. santomea is a novel derived trait.

The adaptive significance of this pigmentation difference, if any, is unknown. Although these species show strong sexual isolation (COYNE et al. 2002), this does not diminish when flies are tested in the dark, suggesting that the pigmentation difference is not a cue for mate discrimination (LLOPART et al. 2002). Moreover, the relationship between pigmentation and temperature is opposite to that expected from other studies of Drosophila: individuals within a species or closely related species living in colder conditions are almost invariably darker (e.g., DAVID et al. 1985; GIBERT et al. 1998), yet D. santomea, which lives at higher altitudes than D. yakuba, is lighter.

In previous genetic analyses using three morphological markers and eight molecular markers, we determined that at least three genes were involved in the pigmentation difference between D. santomea and D. yakuba females, and five genes between males. In each case, the genes resided on all three major chromosomes, with the X chromosome having a particularly strong effect in males (LLOPART et al. 2002).

In this study we extend and refine our previous analysis, using a more accurate method of measuring pigmentation as well as a QTL analysis employing 32 molecular markers, which enables us to map “pigmentation genes” more accurately. Our goals are to determine the number of genes involved in this morphological difference, their chromosomal locations, whether the same genetic regions affect the pigmentation difference in both males and females, and whether QTLs from a given species tend to affect the character in the same direction, implying that the
species difference evolved by natural selection (Orr 1998). Finally, the mapping of QTLs to fairly restricted regions of the genome may eventually allow us to identify specific loci involved in this species difference.

MATERIALS AND METHODS

**Drosophila strains:** All flies were maintained in 8-dram vials containing standard cornmeal-agar-Karo media on a 12h:12h light:dark cycle at 24°C. One isofemale strain was used from each species. The *D. yakuba* Taï 18 strain was derived from a female collected by D. LACHAISE in 1983 in the Taï rainforest on the border between Liberia and the Ivory Coast. The *D. santomea* STO.4 stock was derived from a female collected in March 1998 in the Obo Natural Reserve on São Tomé Island. These two strains were used (and further described) in our previous work on the genetics of pigmentation in these species (LLOPART et al. 2002). The strains are homosequential in chromosome banding pattern except for the right arm of the second chromosome: the *D. yakuba* Taï 18 strain is polymorphic for inversion 2*Rn*, which covers about 40% of the right arm of chromosome 2 (LEMEUNIER and ASHBURNER 1976).

**Crosses:** Backcross (BC) hybrids were produced by crossing four-day-old virgin *D. yakuba* Taï 18 females to virgin *D. santomea* STO.4 males, and then backcrossing virgin F₁ females to males from both species. Genetically, female BC hybrids to *D. yakuba* are either homozygous *D. yakuba* or heterozygous *D. yakuba/santomea*, and have mitochondrial DNA from *D. yakuba*. Similarly, female BC hybrids to *D. santomea* are either homozygous *D. santomea* or heterozygous *D. yakuba/santomea*, and have mitochondrial DNA from *D. yakuba*. Male BC hybrids have the same autosomal and mitochondrial genotypes as females, but the X-linked loci are either pure *D. santomea* or *D. yakuba*, with the Y chromosome from the parental
male used in the backcross. We scored pigmentation in about 50 males and 50 females from each of the pure species and the reciprocal F1 hybrids, and between 73 and 544 males and females from each of the two backcrosses (Table 1). To improve the precision of QTL mapping, we selected backcross individuals with extreme and intermediate pigmentation phenotypes for subsequent pigmentation scoring and genotyping. For each of the first three BC genotypes listed in Table 1, we selected flies of each sex from a sample of about 12,500 individuals (about 50 bottles, each containing about 250 flies of each sex). One-third of the total individuals chosen were judged by eye to have very dark pigmentation, one-third to have very light pigmentation, and the remaining third were chosen randomly from individuals with intermediate phenotypes. (Equal numbers of all three classes were chosen from each bottle inspected until we had accumulated about 500 flies of each sex in each backcross. Thus for each sex we selected about 4% of total individuals inspected; this stringent selection facilitates the precision of mapping). However, females from the backcross of F1 individuals to D. santomea showed little variation in pigmentation, and so all of these were chosen randomly.

**Pigmentation scores:** All scoring of pigmentation was done on 4-day-old virgin flies. We scored only the three posterior tergites of each fly (segments 5, 6, and 7) by examining the fly under a dissecting microscope. A pigmentation score was assigned based on both the percentage of the tergite that was pigmented and the degree of pigmentation. First, the proportion of the tergite covered by black pigment was estimated to the nearest 0.05 (5%). Then, the relative degree of pigmentation was measured within the pigmented area. Using color standards, we assessed the degree of pigmentation within the pigmented area using a five-point scale ranging from 1 (very light pigmentation, slightly darker than background color) to 5 (dark, shiny black), with intermediate numbers representing intermediate degrees of pigmentation.
Unpigmented areas were given a score of 0. We limited ourselves to assigning only three shades of black to each tergite. The percentage of the area of each tergite covered by each shade of pigmentation was then multiplied by the intensity of pigmentation, and these areas were summed. This gives each tergite a minimum possible pigmentation score of 0 (no area pigmented) to 5 (tergite completely covered with very dark pigmentation \(1.0 \times 5\)). These areas were summed for all three tergites, yielding a minimum possible pigmentation score for a given fly of 0 and a maximum possible score of 15. As shown in Table 1, this procedure discriminates well between the pigmentation of these species.

**Molecular markers:** We identified single nucleotide polymorphisms (SNPs) and insertion/deletion variants (indels) that discriminated between the *D. yakuba* Taï18 and the *D. santomea* STO.4 strains for 41 nuclear regions: \(y, per, sog, v, sn, Rux, f, bnb, Hex-A, AnnX, su(f), l(2)gl, Radl, RpL27A, Gart, salr, vkg, Rep4, Adh, His3, barr, Sara, Hex-C, Ngp, Kr, Lsp1-\(\gamma\), RpL14, dib, sfl, Sod, Est6, Ssl1, hb, Xdh (ry), Rpn5, AP-50, Mlc1, ymp, janB, krz, ci\). Newly reported sequences were deposited in GenBank under accession numbers DQ068949 (Gart_T18) and DQ068950 (Gart_sto4); otherwise see Llopard et al. (2005) for details. To determine nucleotide sites differentially fixed between *D. yakuba* Taï18 and *D. santomea* STO.4, we tested 20 individuals (10 per strain) for each region by direct sequencing of PCR products obtained from single fly DNA extractions (Ashburner 1989) using *D. melanogaster* or *D. yakuba* primers. We cleaned PCR products using the Wizard Magnesil PCR clean-up system (Promega Corp. Madison, WI), and sequenced them directly with an ABI PRISM® 3100 Genetic Analyzer (Applied Biosystems Inc. Foster City, CA). We edited the sequences with the Sequencher 3.0 software (Gene Codes Corp. Ann Harbor, MI), and aligned them using the ClustalX program (Thompson et al. 1997).
In total, we sequenced approximately 17.5 kb in each of the 20 flies tested. We detected 263 nucleotide differences fixed between Taï18 and STO.4, that is, 70% of the total nucleotide variation. Among these fixed differences we selected 32 that affect a restriction endonuclease site to be used as markers in the genotyping procedure. Table 2 lists the 32 markers, their relative order within the *D. yakuba* chromosomes, and the conditions for genotyping.

We inferred the relative order of markers within each chromosome in *D. yakuba/D. santomea* from the *D. yakuba* genome project (http://www.genome.wustl.edu/projects/yakuba/; version 040407).

**Marker genotypes:** All BC individuals from the pigmentation assays were stored at –80°C in 0.5 ml eppendorf tubes. Genomic DNA was extracted from each BC individual using the Puregene (Gentra Systems; Minneapolis, MN) single-fly DNA extraction protocol. The sample consisted of 506 BC *D. santomea* males, 73 BC *D. santomea* females, 537 BC *D. yakuba* males, and 526 BC *D. yakuba* females. The genotypes of the 1642 BC hybrids were determined for all 32 markers (*i.e.* 52,544 genotypes).

The 32 molecular markers were designed using sequence data from the parental strains of *D. santomea* and *D. yakuba*. The aligned sequences were used to develop PCR primers using Primer3 (ROZEN and SKALESKY 2000) and restriction enzyme (RE) digestions. Genotyping was performed using Restriction Fragment Length Polymorphism (RFLP) analysis by PCR amplification from genomic DNA using RedTaq DNA polymerase (Sigma; St. Louis, MO) followed by RE digestion (see Table 2 for primers, RE, and conditions). All RE’s were purchased from New England Biolabs (Beverly, MA) and primers were purchased from MWG Biotech (High Point, NC). The digested PCR products were run on a 3% agarose gel stained with ethidium-bromide, imaged with the Bio-Rad ChemiDoc System PC RS-170 using Quantity
One (version 4.2.1) software, and manually genotyped by assigning a “0” (homozygous \( D.\) santomea), “1” (\( D.\) yakuba/\( D.\) santomea heterozygote) or “2” (homozygous \( D.\) yakuba) to each marker genotype. A recombination map based on the 1642 BC hybrids (Table 3) was constructed using the Haldane mapping function.

**QTL mapping:** QTLs affecting variation in pigmentation between \( D.\) yakuba and \( D.\) santomea were mapped in each BC population using composite interval mapping (CIM; ZENG 1994) and implemented using QTL Cartographer software (BASTEN et al. 1999). CIM tests whether an interval between two markers contains a QTL affecting the trait while simultaneously controlling for the effect of QTL located outside the interval using multiple regression on marker co-factors. Marker co-factors were chosen by forward selection – backward elimination stepwise regression. The likelihood ratio (LR) test statistic is \(-2\ln(L_0/L_1)\), where \( L_0/L_1 \) is the ratio of the likelihood under the null hypothesis (i.e., there is no QTL in the test interval) to the alternative hypothesis (there is a QTL in the test interval). LR test statistics were computed every 2 cM with marker co-factors 10 cM or more from the test location. We used permutation analysis to determine appropriate significance thresholds that take into account the multiple tests performed and correlations among markers. We permuted trait and marker data 1000 times, and recorded the maximum LR statistic across all intervals for each permutation. LR statistics calculated from the original data that exceed the 50\(^{th}\) greatest LR statistic from the permuted data are significant at the experiment-wise 5\% level under the null hypothesis (CHURCHILL and DOERGE, 1994; DOERGE and CHURCHILL, 1996). The approximate boundaries of regions containing QTLs were determined by taking 2 LOD intervals (9.22 LR) surrounding the point of greatest significance and interpolating the cytological location of the interval based on the observed amount of recombination between flanking markers.
We estimated the effects of each QTL as the difference between the appropriate homozygous genotypes and heterozygous *D. santomea*/*D. yakuba* genotypes at the peak LR, scaled by the phenotypic standard deviation. The effects in females and autosomes of both sexes are thus estimates of \( a - d \) in the cross to *D. yakuba* and \(-a - d\) in the cross to *D. santomea*, where \(-a\) and \(a\) are, respectively, the genotypic values in *D. santomea* and *D. yakuba*, and \(d\) is the heterozygous effect (FALCONER and MACKAY 1996). The effects of *X*-linked QTLs in males are estimates of \(2a\).

We evaluated pairwise epistatic interactions between all possible marker pairs by running ANOVA models to account for the main effects of all significant markers and one pairwise interaction between markers (DILDA and MACKAY 2002). The ANOVAs were performed with the PROC GLM procedure using SAS 8.02 software (Cary, NC). Interactions with \(P\)-values less than 0.0001 are significant based on a Bonferroni correction for 496 tests per BC population. We estimated the effects of significant two-locus interactions from the least-squares means of the four marker locus classes as \([\overline{x}_{11} + \overline{x}_{22}] - (\overline{x}_{12} + \overline{x}_{21})\], where the first subscript is 1 if the marker has a homozygous genotype for either parental species and 2 if the marker has a heterozygous genotype, and the second subscript takes on the same values for the other marker in the interaction. Standard errors of the interaction effects were estimated as described by DILDA and MACKAY (2002).

**Candidate genes:** Cytological bands in *D. melanogaster* of the markers that define the interval under the QTL peak were obtained using Flybase (DRYSDALE and CROSBY, 2005). Using the known cytological positions of these markers in *D. melanogaster*, we determined the corresponding positions in *D. yakuba* (LEMEUNIER and ASHBURNER, 1976; ASHBURNER, 1989). This defined an interval both in *D. melanogaster* and in *D. yakuba* that allowed us to search for
candidate genes within that interval. We obtained a complete list of candidate genes involved in pigmentation in *D. melanogaster* from Flybase (Drysdale and Crosby, 2005). Candidate genes were identified based on the markers that delimit each QTL, and the presence of pigmentation genes between these two markers (Table 6).

**RESULTS**

**Pigmentation of pure species and F₁ hybrids:** Table 1 gives the mean pigmentation scores, standard errors, and sample sizes for the pure species, the reciprocal F₁ hybrids, and the backcross hybrids used for genotyping. The difference between the pure species is substantial: the mean pigmentation score of *D. yakuba* males and females is 14.22 and 9.85, respectively, and for *D. santomea* males and females 0.63 and 1.02, respectively. As seen in our previous analysis, (Llopart et al. 2002) reciprocal F₁ hybrid males show a large effect of the X chromosome on pigmentation: these males have pigmentation scores fairly close to those of males from the species of the maternal parent. The difference in pigmentation scores between the two classes of F₁ males is highly significant (*t* = 27.1, 100 d.f., *P* < 0.001). The relative effect of the X chromosome in male pigmentation can be judged as the percentage of the total difference between males of the two species explained by the difference between the reciprocal F₁ males; this effect is about 60%. This effect is much larger than the relative size of this chromosome (constituting roughly 21% of the haploid genome [Table 3]) and suggests either that the X chromosome carries a disproportionate number of genes affecting pigmentation, or that individual X-linked genes have disproportionately large effects. (The QTL analysis below shows that the second explanation is most likely to be correct.)
In contrast to males, the F1 females do not differ significantly in pigmentation scores (Table 1; $t = 1.55$, 99 d.f., $P = 0.12$). There is thus no evidence for a maternal or mitochondrial effect affecting pigmentation of these females, who are identical in nuclear genotype. The mean score of all F1 females (4.45) is slightly lighter than the average score of females for the two species (5.43), showing a small amount of dominance for the $D.\ santomea$ phenotype.

**QTLs affecting variation in pigmentation in BC hybrids:** We mapped four QTLs with large effects on pigmentation (Table 4, Figure 1). The same four QTLs were detected in male hybrids in the backcrosses to both $D.\ santomea$ and $D.\ yakuba$: two QTLs were on the X chromosome (between markers 1-2 and markers 6-10), one QTL was on the second chromosome (between markers 15-18) and one QTL was on the third chromosome (between markers 23-26). The magnitude of the QTL effects ranged from 0.49 – 1.42 phenotypic standard deviations in the backcross to $D.\ santomea$, and accounted for 67% of the total phenotypic variation. Similarly, the QTL effects ranged from 0.62 – 1.62 phenotypic standard deviation in the backcross to $D.\ yakuba$, and accounted for 58% of the total phenotypic variation. In both backcrosses, the sum of the estimated QTL effects equaled or exceeded that expected from the difference between the parental genotypes. In the backcross to $D.\ santomea$, the expected difference in pigmentation is $-11.29$ (i.e. the difference in pigmentation between $D.\ santomea$ males (0.63) and (Y X S) F1 males (11.92)); whereas the sum of the QTL effects was $-11.64$. In the backcross to $D.\ yakuba$, the expected difference in pigmentation is 10.5 (i.e. the difference in pigmentation between $D.\ yakuba$ males (14.22) and (S X Y) F1 males (3.72)); whereas the sum of the QTL effects is 16.15. Thus, it is likely that we have detected all of the QTLs affecting variation in pigmentation in this hybridization, and that our selective genotyping protocol led to over-estimation of effects.
LYNCH and WALSH 1998). It is also possible that estimates of main effects have been biased by epistatic interactions (see below).

Four QTLs in the same positions affected variation in pigmentation in the female \textit{D. yakuba} BC hybrids. The magnitude of the QTL effects ranged from 0.31 – 0.83 phenotypic standard deviations, and accounted for 63\% of the total phenotypic variance. The sum of the estimated QTL effects slightly exceeded that expected from the difference between the parental genotypes in this backcross, again suggesting that we have detected all of the QTL affecting variation in pigmentation in this hybridization. The expected difference in pigmentation is 5.39 (i.e. the difference in pigmentation between \textit{D. yakuba} females (9.85) and F1 females (4.46 on average)); whereas the sum of the QTL effects is 5.62. We observed only a single \textit{X} chromosome QTL (between markers 6-10) in the \textit{D. santomea} BC females, accounting for 43\% of the total phenotypic variance. This could be attributable to a lack of power to detect QTLs in this cross, since only 73 flies were assessed for genotype – phenotype associations, compared to over 500 individuals in each of the other crosses. Indeed, this QTL only accounted for 40\% of the expected difference in pigmentation (−3.44; i.e. the difference in pigmentation between \textit{D. santomea} females (1.02) and F1 females (4.46 on average)).

The QTL effects were largely additive within loci. In the BC to \textit{D. yakuba}, the effects of the two \textit{X}-chromosome QTLs in females \((a − d)\) were approximately half that of the effects in males \((2a)\), consistent with \(d = 0\). In addition, the effects of the chromosome 3 QTL in males from the backcrosses to \textit{D. yakuba} and \textit{D. santomea} were equal and opposite, as expected if \(d = 0\). The second chromosome QTL had a larger effect in the BC to \textit{D. yakuba} than to \textit{D. santomea}, suggesting partial dominance of the \textit{D. santomea} genotype. Since dominance of \textit{D. santomea}
QTLs reduces the power to detect QTLs in the backcross to *D. santomea*, this could also account for our failure to detect this QTL in females from this backcross.

**Epistatic interactions:** We assessed all possible epistatic interactions between pairs of markers within each cross and sex (Table 5, Figures 2 – 4). We observed significant epistasis (after correcting for multiple tests) in males from both backcrosses between markers in regions encompassed by the QTLs, but not between QTL regions and regions without main effects, or between two regions with no main effects on pigmentation (Figure 2). The significant interactions were between the two *X*-chromosome QTLs; the *X*-chromosome QTL between markers 6-10 and the chromosome 2 QTL; the *X*-chromosome QTL between markers 6-10 and the chromosome 3 QTL; and the chromosome 2 and chromosome 3 QTLs (Figure 2, Table 5). The nature of these interactions is illustrated in Figure 3, where the effect of a Y/S substitution at the second locus is shown in the form of reaction norms, conditional on the genotype of the first locus (where Y denotes a *D. yakuba* allele and S denotes a *D. santomea* allele at the QTL). In the absence of epistasis, the effect of the substitution at the second locus would be independent of the genotype of the first, and the reaction norms would be parallel.

In the backcross to *D. santomea*, we expect the hemizygous Y or heterozygous SY genotype at the second locus to be more pigmented than the hemizygous S or homozygous SS genotype at this locus. However, for all the interacting markers in this backcross, this is only true if the genotype at the first locus is Y (or SY). Either there is no difference between the genotypes at the second locus if the first is S (SS), or, for the case of the interaction between the second *X*-chromosome QTL and the chromosome 2 QTL, the SS genotype at the chromosome 2 QTL is actually more pigmented than the SY genotype at this QTL when the *X*-chromosome QTL is S (Figure 3). In other words, the effect of a Y-S substitution in an otherwise S background is
smaller than the effect of an S-Y substitution at each QTL in the Y background. Equivalently, the
effect of Y-S substitutions at two interacting loci in the homozygous S background is greater
than additive, and the effect of S-Y substitutions at two interacting loci in the heterozygous SY
background is less than additive. This is illustrated in Figure 4, where the sum of the effects of
substituting single Y alleles at each QTL in the S background would yield a predicted
pigmentation score of 4.09 for the YYYYY haplotype, whereas the observed score is 9.43.

The epistatic interactions are more complicated in the backcross to D. yakuba. Here we
expect the hemizygous Y or homozygous YY genotype at the second locus to be more
pigmented than the hemizygous S or heterozygous SY genotype at this locus. In the interaction
between the two X-chromosome QTLs (marker 1 × marker 9), this is true if marker 1 is Y. On
the other hand, in the interactions with the chromosome 3 QTL (marker 8 × marker 25 and
marker 18 × marker 25), this is true if marker 8 is S and marker 25 is SY (Figure 3). However,
the interaction between the second X-chromosome QTL and the chromosome 2 QTL (marker 8 ×
marker 18) is in the opposite direction to that expected: the SY genotype at marker 18 is actually
more pigmented than the YY genotype at this marker, but only if marker 8 is Y. Overall, the
effect of Y-S substitutions at two interacting loci in the heterozygous SY background is less than
additive, and the effect of S-Y substitutions at two interacting loci in the homozygous YY
background is greater than additive. Figure 4 shows that the sum of the effects of substituting
single Y alleles at each QTL would yield a predicted pigmentation score of 15.15 for YYYYY
haplotype, whereas the observed score is 13.43.

A single epistatic interaction was observed in females from the backcross to D. yakuba,
between the second X chromosome QTL (markers 6-10) and the chromosome 2 QTL (data not
shown). The direction of the epistatic effects between these QTLs is the same as in males from
this hybridization. The effect of Y-S substitutions at the two interacting loci in the heterozygous SY background is less than additive, and the effect of S-Y substitutions at two interacting loci in the homozygous YY background is greater than additive. Figure 4 shows that the sum of the effects of substituting single Y alleles at each QTL would yield a predicted pigmentation score of 11.17 for YYYY haplotype, whereas the observed score is 9.49.

DISCUSSION

We have mapped at least four QTLs with large effects associated with the variation in pigmentation between \textit{D. yakuba} and \textit{D. santomea}. The QTLs mapped to the same locations in both males and females in the backcrosses to \textit{D. yakuba} and in males in the backcross to \textit{D. santomea}: two QTLs mapped to the X chromosome, and one each to the second and third chromosomes. Thus, the loss of pigmentation in \textit{D. santomea} involved evolutionary changes in several genes, which probably affected both sexes. (Although only a single QTL was detected in females in the backcross to \textit{D. santomea}, it mapped to the same location as one of the X chromosome QTLs, and the small sample size of this population and narrow range of pigmentation conspire to reduce the power to detect QTLs with small effects.)

This study not only largely expands and refines the earlier results of LLOPART \textit{et al.} (2002), in which pigmentation differences were assessed in backcross hybrids using eight molecular markers, but also provides the first accurate chromosomal locations of genetic factors associated with these differences. In LLOPART \textit{et al.} (2002), the QTL of largest effect was also associated with \textit{AnnX} at the base of the X chromosome, with a second QTL with smaller effect at the tip of the X chromosome associated with \textit{y}. The locations of the autosomal QTLs are also concordant between the two studies. The QTL with the smallest effect detected by LLOPART \textit{et al.}
al. (2002) was associated with the marker at *bric-à-brac 1 (bab1)* at the tip of 3L, which is only marginally significant in the backcross to *D. yakuba* females in this study (LR = 14 between the *Lsp1γ* and *dib* markers). One possible explanation for this small discrepancy is that the *bab1* marker is in linkage disequilibrium with the chromosome 3 QTL mapped in this study. This, however, is not likely because the map distance between the major QTL detected on chromosome 3 and the *bab1* region is greater than 100 cM. It is possible that the discrepancy could be due to the fact that the methods used to score abdominal pigmentation in both studies, although correlated, are different.

The results presented here raise the interesting possibility that the genetic basis of pigmentation differences between *D. yakuba* and *D. santomea* is fairly simple. We infer that we have detected all major QTLs accounting for variation in pigmentation in these backcross hybrids (with the exception of females in the backcross to *D. santomea*), since the sum of the QTL effects equals or exceeds that expected from the difference between parental strain means. Further high-resolution mapping is required to determine whether single genes or multiple closely linked loci are responsible for the large QTL effects. Nevertheless, all QTLs from the same species affected pigmentation in the same direction, suggesting that the species difference might have arisen by natural selection (ORR 1998). We were not able to formally test this hypothesis, since a minimum of six QTLs are required to reject the null hypothesis (ORR 1998). However, sequencing of the relevant loci may show, by the ratio of coding versus non-coding substitutions, whether selection was involved in their divergence.

Only two other studies have investigated the genetic basis of pigmentation differences between closely related species of Drosophila. HOLLOCHER *et al.* (2000b) studied two Caribbean species in the *Drosophila cardini* group having extremely different pigmentation patterns: *D.*
arawakana (a light-colored sexually dimorphic species) and D. nigrodunni (the darkest sexually monomorphic species of the D. dunni subgroup). Using quantitative measures of abdominal pigmentation in F₁ hybrids and backcross flies, the authors conclude that, at least for the posterior segment of the abdomen (“area 3” in HOLLOCHER et al. 2000a, fig. 2), which is roughly equivalent to the area scored in our analysis, there are paternal and maternal effects, with no particular effect of the X chromosome. The second study mapped QTLs for the difference in pigmentation between D. americana and D. novamexicana (WITTKOPP et al. 2003) using 23 molecular markers. Five genes (y, e, Ddc, omb and bab) previously implicated in the development and evolution of abdominal pigmentation were used as markers. The authors indicate that this species difference is polygenic with no significant effect of the X chromosome but with significant effects of three of the five autosomes. There is little genetic commonality between the results reported by these two studies and our results. Of course, unless there is a very limited number of genes that could be potentially responsible for differences in pigmentation, one does not expect the genetic architecture to be shared among distantly related species.

Our observations of epistatic interactions between QTLs with main effects on pigmentation are consistent with genes corresponding to the QTLs that are in the same pathway(s). In the absence of high resolution mapping, however, we can only speculate about what candidate genes might correspond to the QTLs. An obvious candidate for the QTL at the tip of the X chromosome is yellow (y) itself, and complementation tests to D. santomea using a y mutation in D. yakuba are consistent with a very small contribution of mutations at the y locus in the pigmentation difference between these species (LLOPART et al. 2002). In addition, two
enhancers of $y$ are located in the region embraced by the QTL at the base of the X chromosome; these could contribute to the interactions between the two X chromosome QTLs.

Several candidate genes affecting body pigmentation have been identified by mutagenesis in *D. melanogaster* (DRYSDALE and CROSBY, 2005), and co-localize to the regions containing QTLs affecting pigmentation differences between *D. yakuba* and *D. santomea* (Table 6). Catecholamines are required for proper melanization and sclerotization of the Drosophila cuticle (WRIGHT 1987; WALTER *et al.* 1996). The *Ddc* gene cluster on chromosome 2 (genetically defined by *Df(2L)TW130; 37B9-C1.2;D1-2*) contains at least 18 functionally related genes involved in the catecholamine pathway, including *Catsup, Ddc, Dox-A2, amd*, and *l(2)37Ca* (STATHAKIS *et al.* 1995). Mutations in 11 of the loci in this complex (including *Ddc* and *amd*) produce melanotic psuedotumors, indicating abnormal catecholamine metabolism (WRIGHT 1996), and mutations in 14 of the loci affect the formation, sclerotization or melanization of the cuticle (WRIGHT 1996). The *Ddc* cluster co-localizes with the QTL on chromosome 2. *Pu* encodes GTP cyclohydrolase, the rate limiting step in the synthesis of tetrahydobiopterin, the co-factor required for the phosphorylation of tyrosine hydroxylase, which is in turn the rate limiting step in the synthesis of dopamine (STATHAKIS *et al.* 1999). *Pu* also co-localizes with the QTL on chromosome 2. The *silver* (*svr*) gene, which encodes proteins that are members of the carboxypeptidase family (SETTLE *et al.* 1995), co-localizes with the QTL at the tip of the X chromosome. Mutations in *svr* affect pigmentation, wing shape and catecholamine pools (WRIGHT 1987). *tan* (*t*) is an excellent candidate gene corresponding to the QTL at the base of the X chromosome. *t* is probably the structural gene for beta-alanyldopamine hydrolase activity; *t* mutants have reduced dopamine levels (WRIGHT 1987).
Additional candidate genes in the pigmentation pathway include *optomotor-blind* (*omb*), *black* (*b*), *Cysteine proteinase-1* (*Cp1*), and *Black cells* (*Bc*) (Wright 1987; Wittkopp et al. 2002; 2003). The developmental gene, *omb*, co-localizes with the QTL at the tip of the X chromosome. *omb* encodes a T-box transcription factor that is necessary for patterning the pigment band in each adult abdominal segment of *Drosophila melanogaster* (Kopp and Duncan 1997; 2002; Kopp et al. 1997). A recent study by Brisson et al. (2004) examined patterns of nucleotide variation at the *omb* locus in *D. polymorpha*, a species highly polymorphic for abdominal pigmentation. Two classes of haplotypes that appear to be under balancing selection were associated with variation in abdominal pigmentation in this species.

*b*, *Cp1* and *Bc* all co-localize with the chromosome 2 QTL. *b* encodes a product involved in beta-alanine biosynthesis; *b* mutants are heavily pigmented. *Cysteine proteinase 1* (*Cp1*) encodes a product with cathepsin L activity; deletion studies of *Cp1* have shown complete female sterility and reduced pigmentation in abdominal segments 1-5 (Gray et al. 1998). *Bc* encodes a tyrosinase, which catalyzes the *de novo* synthesis of melanin from tyrosine (Wittkopp et al. 2003).

Conspicuously absent from the list of strong potential candidate genes are *bab1* and *bab2*, two closely linked genes at the tip of chromosome 3L that are thought to be repress male-specific abdominal pigmentation in females (Kopp et al. 2000) and contribute significantly to variation of abdominal pigmentation in females of *D. melanogaster* (Kopp et al. 2003). The expression of *bab* is correlated with pigmentation across a diverse range of Drosophila species, such that species in which neither sex is pigmented express exhibit similar expression of Bab in males and females, but species in which abdominal tergites of males are more pigmented than females have female-specific Bab expression (Kopp et al. 2000). Thus, it was possible *a priori*
that overexpression of Bab in *D. santomea* could have resulted in loss of pigmentation in both sexes. This is not the case, however, since none of the QTLs map in the vicinity of *bab*. Further, Bab2 protein is expressed in a dimorphic *melanogaster*-like pattern in *D. santomea* (GOMPEL and CARROLL 2003), which is inconsistent with mutations at *bab* affecting the difference in pigmentation between *D. santomea* and *D. yakuba*.

While it is plausible that the loss of pigmentation in *D. santomea* was driven by natural selection, it is also possible that selection acted on pleiotropic effects of genes affecting pigmentation, and not pigmentation itself. All of the candidate genes listed in Table 3 have highly pleiotropic effects on traits related to fitness, including reproduction and immune response. For example, *Ddc* catalyzes the final step in the biosynthesis of the neurotransmitters dopamine and serotonin. Dopamine is required in Drosophila for normal development (NECKAMEYER 1996); ovarian maturation, fecundity and sexual receptivity in females (NECKAMEYER 1996; 1998a); learning (TEMPEL et al. 1984; NECKAMEYER 1998b); locomotion (PENDLETON et al. 2002) and aggressive behavior (BAIER et al. 2002). Serotonin also regulates or modulates a variety of behaviors in many animal species, including aggression, feeding, learning, locomotion, sleep and mood (BLENAU and BAUMANN 2001). Further speculation about the nature of the pleiotropic effects (and sex-specific epistatic effects) of genes affecting variation in pigmentation between these species must await the positional cloning of these genes.

**Acknowledgements:** We thank BETHUEL MGUMBA and ERIC GROSSMAN for technical help as well as AMANDA J. MOEHRING and TED J. MORGAN for helpful discussions. This work was funded by NIH research grants to J. A. C. (GM 58260) and T. F. C. M. (GM45344, GM 58260).
ASHBURNER, M., 1989  *Drosophila*: A laboratory manual. Cold Spring Harbor Press, Cold Spring Harbor, N. Y.

BASTEN, C. J., B. S. WEIR and Z-B. ZENG, 1999  QTL Cartographer, Version 1.13. Department of Statistics, North Carolina State University, Raleigh, NC.

BAIER, A., B. WITTEK and B. BREMBS, 2002  *Drosophila* as a new model organism for the neurobiology of aggression? J. Exp. Biol. **205**: 1233-1240.

BLENNU, W., and A. BAUMANN, 2001  Molecular and pharmacological properties of insect bioamine receptors: Lessons from *Drosophila melanogaster* and *Apis mellifera*. Arch. Insect Biochem. Phys. **48**: 13-38.

BRISON, J. A., A. R. TEMPLETON and I. DUNCAN, 2004  Population genetics of the developmental gene *optomotor-blind*(omb) in *Drosophila polymorpha*: evidence for a role in abdominal pigmentation variation. Genetics **168**: 1999-2010.

CARIOU, M. L., J. F. SILVAIN, V. DAUBIN, J. L. DALAGE and D. LACHAISE, 2001  Divergence between *Drosophila santomea* and allopatric or sympatric populations of *D. yakuba* using paralogous amylase genes and migration scenarios along the Cameroon volcanic line. Mol. Ecol. **10**: 649-660.

CHURCHILL, G. A., and R. W. DOERGE, 1994  Empirical threshold values for quantitative trait mapping. Genetics **138**: 963-971.

COYNE, J. A., S. ELWYN, S. Y. KIM and A. LLOPART, 2004  Genetic studies of two sister species in the *Drosophila melanogaster* subgroup, *D. yakuba* and *D. santomea*. Genet. Res. **84**: 11-26.
COYNE, J. A., S. Y. KIM, A. S. CHANG, D. LACHAISE and S. ELWYN, 2002  Sexual isolation between two species with overlapping ranges: Drosophila santomea and Drosophila yakuba. Evolution **56**: 2424-2434.

DAVID, J. R., P. CAPY, V. PAYANT and S. TSAKAS, 1985  Thoracic trident pigmentation in Drosophila melanogaster: differentiation of geographical populations. Genet. Sel. Evol. **17**: 211-224.

DILDA, C. L., and T. F. C. MACKAY, 2002  The genetic architecture of Drosophila sensory bristle number. Genetics **162**: 1655-1674.

DOERGE, R. W., and G. A. CHURCHILL, 1996  Permutation tests for multiple loci affecting a quantitative character. Genetics **142**: 285-294.

DRYSDALE R.A., CROSBY M.A. and The FlyBase Consortium, 2005 FlyBase: genes and gene models. Nucleic Acids Research **33**:D390-D395.  http://flybase.org/

FALCONER, D. S., and T. F. C. MACKAY, 1996  *Introduction to Quantitative Genetics*, Ed. 4. Longman, London.

FISHER, R. A., 1930  *The Genetical Theory of Natural Selection: A complete variorum edition*. Oxford University Press, Oxford.

GRAY, Y. H. M., J. A. SVED, C. R. PRESTON and W. R. ENGELS, 1998  Structure and associated mutational effects of the cysteine proteinase (CP1) gene of Drosophila melanogaster. Ins. Mol. Biol. **7**: 291-293.

GIBERT, P., B. MORETEAU, J. C. MORETEAU, R. PARKASH and J. R. DAVID, 1998  Light body pigmentation in Indian Drosophila melanogaster: a likely adaptation to a hot and arid climate. J. Genetics **77**: 13-20.

GOMPEL, N. and S. B. CARROLL, 2003  Genetic mechanisms and constraints governing the
evolution of correlated traits in drosophilid flies. Nature **424**: 931-935.

HOLLOCHER, H., J. L. HATCHER, AND E. G. DYRESON, 2000a. Evolution of abdominal pigmentation differences across species in the *Drosophila dunni* subgroup. Evolution **54**:2046-2056.

HOLLOCHER, H., J. L. HATCHER, AND E. G. DYRESON, 2000b. Genetic and developmental analysis of abdominal pigmentation differences across species in the *Drosophila dunni* subgroup. Evolution **54**:2057-2071.

KOPP, A., R. K. BLACKMAN and I. DUNCAN, 1999 Wingless, decapentaplegic and EGF receptor signaling pathways interact to specify dorso-ventral pattern in the adult abdomen of *Drosophila*. Development **126**: 3495-507.

KOPP, A., and I. DUNCAN, 1997 Control of cell fate and polarity in the adult abdominal segments of *Drosophila* by *optomotor-blind*. Development **124**: 3715-3726.

KOPP, A., and I. DUNCAN, 2002 Anteroposterior patterning in adult abdominal segments of *Drosophila*. Dev. Biol. **242**: 15-30.

KOPP, A., I. DUNCAN, D. GODT and S. B. CARROLL, 2000 Genetic control and evolution of sexually dimorphic characters. Nature **408**: 553-559.

KOPP, A., R. M. GRAZE, S. XU, S. B. CARROLL and S. V. NUZHDIN 2003 Quantitative trait loci responsible for variation in sexually dimorphic traits in *Drosophila melanogaster*. Genetics **163**: 771-787.

LACHAISE, D., M. HARRY, M. SOLIGNAC, F. LEMEUNIER, V. BENASSI, *et al.*, 2000 Evolutionary novelties in islands: *Drosophila santomea*, a new melanogaster sister species from Sao Tome. Proc. Roy. Soc. Lond. B **267**: 1487-95.

LEMEUNIER, F., and M. ASHBURNER, 1976 Relationship within the *melanogaster* subgroup
of the genus *Drosophila* (Sophophora). II. Phylogenetic relationships between six species based upon polytene chromosome banding sequences. Proc. Roy. Soc. Lond. B 193: 275-294.

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.

Llopart, A., S. Elwyn and J. A. Coyne, 2002 Pigmentation and mate choice in *Drosophila*. Nature 419: 360.

Llopart, A., D. Lachaise and J. A. Coyne, 2005 Multilocus analysis of introgression between two sympatric sister species of Drosophila: *D. yakuba* and *D. santomea*. Genetics, in press.

Lynch, M., and B. Walsh, 1998 Genetics and analysis of quantitative traits. Sinauer Associates, Sunderland, MA.

Neckameyer, W., 1996 Multiple roles for dopamine in *Drosophila* development. Dev. Biol. 176: 209-219.

Neckameyer, W., 1998a Dopamine modulates female sexual receptivity in *Drosophila melanogaster*. J. Neurogenet. 12: 101-114.

Neckameyer, W. 1998b Dopamine and mushroom bodies in *Drosophila*: experience-dependent and –independent aspects of sexual behavior. Learn. Mem. 5: 157-165.

Orr, H. A., 1998 Testing natural selection vs. genetic drift in phenotypic evolution using quantitative trait locus data. Genetics 149: 2099-2014.

Orr, H. A., 2001 The genetics of species differences. Trends Ecol. Evol. 16: 343-350.

Orr, H. A., and J. A. Coyne, 1992 The genetics of adaptation: a reassessment. Amer. Natur. 140: 725-742.
PENDLETON, R. G., A. RASHEED, T. SARDINA, T. TULLY and R. HILLMAN, 2002  Effects of tyrosine hydroxylase mutants on locomotor activity in *Drosophila*: a study in functional genomics. Behav. Genet. **32**: 89-94.

ROZEN, S., and H. J. SKALETSKY, 2000  Primer3 on the WWW for general users and for biologist programmers, pp. 365-386 in *Bioinformatics Methods and Protocols: Methods in Molecular Biology*, edited by S. KRAWETZ and S. MISENER. Humana Press, Totowa, N. J. (http://fokker.wi.mit.edu/primer3/).

SETTLE, S. H. Jr., M. M. GREEN and K. C. BURTIS, 1995  The silver gene of *Drosophila melanogaster* encodes multiple carboxypeptidases similar to mammalian prohormone-processing enzymes. Proc Natl. Acad. Sci. USA **92**: 9470-9474.

STATHAKIS, D. G., D. Y. BURTON, W. E. MCIivor, S. KRISHNAKUMAR, T. R. F. WRIGHT, *et al.*, 1999  The Catecholamines up (Catsup) protein of *Drosophila melanogaster* functions as a negative regulator of tyrosine hydroxylase activity. Genetics **153**: 361-382.

STATHAKIS, D. G., E. S. PENTZ, M. E. FREEMAN, J. KULLMAN, G. R. HANKINS, *et al.*, 1995  The genetic and molecular organization of the *Dopa decarboxylase* gene cluster of *Drosophila melanogaster*. Genetics **141**: 629-655.

TEMPEL, B. L., M. S. LIVINGSTONE and W. G. QUINN, 1984  Mutations in the *dopa decarboxylase* gene affect learning in *Drosophila*. Proc. Natl. Acad. Sci. USA **81**: 3577-3581.

THOMPSON, J. D., T. J. GIBSON, F. PLEWNIAK, F. JEANMOUGIN and D. G. HIGGINS, 1997 The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res **25**: 4876-4882.
WALTER, M. F., L. L. ZEINEH, B. C. BLACK, W. E. McIVOR, T. R. F. WRIGHT, et al., 1996  Catecholamine metabolism and *in vitro* induction of premature cuticle melanization in wild type and pigmentation mutants of *Drosophila melanogaster*. Arch. Insect Biochem. Physiol. **31**: 219-233.

WITTKOPP, P. J., S. B. CARROLL and A. KOPP, 2003  Evolution in black and white: genetic control of pigment patterns in *Drosophila*. Trends Genet. **19**: 495-504.

WITTKOPP, P. J., J. R. TRUE and S. B. CARROLL, 2002  Reciprocal functions of the *Drosophila* yellow and ebony proteins in the development and evolution of pigment patterns. Development **129**: 1849-1858.

WRIGHT, T. R. F., 1987  The genetics of biogenic amine metabolism, sclerotization, and melanization in *Drosophila melanogaster*. Adv. Genet. **24**: 127-222.

WRIGHT, T. R. F., 1996  The Wilhelmine E. Key 1992 Invitational lecture. Phenotypic analysis of the *Dopa decarboxylase* gene cluster mutants in *Drosophila melanogaster*. J. Hered. **87**: 175-190.

ZENG, Z.-B., 1994  Precision mapping of quantitative trait loci. Genetics **136**: 1457-1468.
## TABLE 1

Pigmentation scores of pure *D. yakuba*, *D. santomea*,

$F_1$ hybrids from the reciprocal crosses, and backcross individuals.

| Genotype             | Sex | Mean score (SE) | N   |
|----------------------|-----|-----------------|-----|
| *D. yakuba* Taï 18   | M   | 14.22 (0.30)    | 51  |
|                      | F   | 9.85 (0.15)     | 53  |
| *D. santomea* STO.4 | M   | 0.63 (0.05)     | 50  |
|                      | F   | 1.02 (0.06)     | 56  |
| $F_1$ (Y X S)        | M   | 11.92 (0.21)    | 51  |
|                      | F   | 4.62 (0.18)     | 50  |
| $F_1$ (S X Y)        | M   | 3.72 (0.21)     | 51  |
|                      | F   | 4.29 (0.12)     | 51  |
| Backcross (F1 X Y)   | M   | 7.26 (0.18)     | 544 |
|                      | F   | 7.42 (0.09)     | 544 |
| Backcross (F1 X S)   | M   | 4.50 (0.16)     | 517 |
|                      | F   | 2.05 (0.13)     | 73  |

“Y” = *D. yakuba*, “S” = *D. santomea*. (Taï 18 and STO.4 strains were used in all crosses). In all crosses the genotype of the female parent is given first. All $F_1$ females were produced by crossing *D. yakuba* females to *D. santomea* males.
| Marker | Cytological Location | Primer Sequence (5’-3’) | Type | PCR Ta (°C) | Restriction Endonuclease |
|--------|---------------------|-------------------------|------|-------------|--------------------------|
| y      | 1A5                 | CGCTGCCTGTTTGTTTTATT    | S    | 55          | AvaII                    |
|        |                     | GCGAATGGTTCAAAAGAATAATTTTC |
| per    | 3B1-2               | TTTCCAGTTCTCCGAATCAGC   | S    | 55          | BbvI                     |
|        |                     | CCTTAGGGGTGAGCCACTCT    |
| sog    | 13E1                | GCTGGGCCTACAACATTGAAA   | S    | 57          | XhoI                     |
|        |                     | CTCGGTGCCACATTCAC       |
| v      | 9F11                | AGACTCCCTTCCTGCTTTTC    | S    | 55          | SspI                     |
|        |                     | TGAGAGCTCCAGTTCCGACT    |
| rux    | 5D2                 | CATTGCTCTCCGTTTCTC     | S    | 55          | HpyCH4IV                  |
|        |                     | GTGCTTGTAGCGCTGTGTC     |
| Gene  | Primer | Sequence | Size | Enzyme |
|-------|--------|----------|------|--------|
| *f*   | 15F4--7| CTCGCGAATGGCAGCAGAT S 55 | HpaII |
|       |        | AATGTACGTCGCCTGGAT |      |
| *bnb* | 17D6   | TTCCTTTCTCTCTTCTCTTTGA S 55 | PvuII |
|       |        | CCGAGAAGAAGTCCATCGAG |      |
| *Hex-A* | 8E10 | GGTACCCAGCTCTTCGATCA S 57 | HhaI |
|       |        | GGCAATGGCATCCTTTAGAA |      |
| *AnnX* | 19C1 | AAACCAGAGAGCTGCTTTCA S 55 | Taq1α |
|       |        | ATTCTCCTTGCGACGTCTTG |      |
| *su(f)* | 20E | TGGTGGGCAAAAGTCAAAT ID 57 | NcoI |
|       |        | AAAATCTTAGCCGCCTGGAC |      |
| *l(2)gl* | 21A5 | TGACGTCGCTGAAGTCTTCTTG S 54 | MseI |
|       |        | GATGGGCACGCCCTGGAC |      |
| *Rad1* | 23A1 | ATGAATGTGCTGTCCGAGTG S 55 | HpyCH4IV |
|       |        | GTTCGTGGAACACCTTCGAT |      |
| *RpL27A* | 24F3 | ATCAAGCGGAAGAAGACCAG S 55 | NcoI |
|       |        | GACCTTGGCCGAAGTAACCAG |      |
| Gene | Location | Sequence | Restriction | Identifier |
|------|----------|----------|-------------|------------|
| salr | 32E4-F1  | AGCTGACTGATCCCAACCAG | ScrF1       |            |
|      |          | GATGATGCCGTGGAGAACT   |             |            |
| Rep4 | 34B4     | TCACGGGAGTACGAACCCA   | Taq1^α      |            |
|      |          | GATGATGCCGTTGGAGAACT  |             |            |
| His3 | 39D3-E1  | TTTCAGGACCACAAACCA    | MfeI        |            |
|      |          | CCGTTGCCCTTATAAACA    |             |            |
| barr | 38B1-2   | GCAGTGCAGGATGAAGATCA  | HaeIII      |            |
|      |          | TTGGAGTCCAACCTCAGA    |             |            |
|      |          | TTGGAGTCCAACCTCAGA    |             |            |
| Sara | 57E6     | CGACCACAAACCTGAATT   | HphI        |            |
|      |          | CATGTTATCCGCGACCATA   |             |            |
|      |          | CATGTTATCCGCGACCATA   |             |            |
| Kr   | 60F5     | ACCAGCCATGATGGAGATT   | MlyI        |            |
|      |          | CTACAGAGCTGGCTCCATCC  |             |            |
|      |          | CTACAGAGCTGGCTCCATCC  |             |            |
| Lsp1γ| 61A6     | CAAAACCACCCACAAGCAG   | AluI        |            |
|      |          | CCTTTGTACTCTTCTCGTACATGAT |         |            |
| dib  | 64A5     | AGTCCTTTTCTCCCCAGGAA  | NruI        |            |
|      |          | ATTGGGCCTGGCTGAGTT    |             |            |
| Gene  | Chromosome | Sequence 1 | Type | Length | Enzyme |
|-------|------------|------------|------|--------|--------|
| sfl   | 65B-4      | GGGTAATCCCTGTGACGATG TTCCGATGGAAAGAAGTCCA | S    | 52     | NsiI   |
| Est-6 | 69A1       | TCCTGCCTACGCTTTGTCT AAAAGTAGTCGTCGCCATGC | ID   | 52     | MseI   |
| Ssl1  | 80B2       | GGTGCCCAAGTAGTGTTGAGT GACGCACATTTTCGAGATCA | S    | 52     | BsrI   |
| ry    | 87D9       | CGCTTTGAAGCAAAAAATCCA GAAGAACAAGCTCACCACCA | S    | 52     | SacI   |
| Rpn5  | 83C4       | TACCGAGGGCAAGATTACG TGCTGATCTTTCTGGCAATG | S    | 57     | MseI   |
| AP-50 | 94A15-16   | AGTGCAAGTTCCGGCATCAA GAATGGCAGCGAAATGTCTT | S    | 57     | HaeIII |
| Mlc1  | 98A14-15   | TGCAAAACAGAGTTCCGTCAG | S    | 52     | Tsp5091 |
| ymp   | 96E        | CCTCGAGACCCGCAGTATG CACCTCGCAGTTCTGATTGA | S    | 53     | HaeIII |
| Marker | Chromosome | Forward Primer | Marker Type | Reverse Primer | Restriction Enzyme |
|--------|------------|----------------|-------------|----------------|-------------------|
| janB   | 99D3       | CATGGCTTCACGAAATACGG | S | 57 | SalI |
|        |            | CTTACCTGGAGGTGCCATA |         |      |      |
| kraz   | 100E3      | CGCATGTGGTCAAATAAAATCG | S | 55 | MseI |
|        |            | TTTTTGGGATAACCCATTATTCA |         |      |      |
| ci     | 102A1-3    | AGCCCTTGCGATGAAGACTC | S | 50 | HpaI |
|        |            | TGGTAGGTCTGCTACGTCC |         |      |      |

Cytological locations are given on the basis of *D. melanogaster* cytology (LEMEUNIER and ASHBURNER 1976). The order of the markers in the first column reflects their relative positions in *D. yakuba/D. santomea* chromosomes inferred from the *D. yakuba* genome project (http://www.genome.wustl.edu/projects/yakuba/).

The marker type is S, SNP; ID, insertion/deletion. The PCR protocol for all markers is 1 cycle 94 °C, 2 minutes; 35 cycles 94 °C, 30 seconds; T<sub>A</sub>, 30 seconds; 72 °C, 30 seconds; 1 cycle 72 °C, 4 minutes; where the annealing temperature, T<sub>A</sub>, is listed. PCR products were digested with a restriction endonuclease, run on a 3% agarose gel stained with ethidium bromide, imaged with the Bio-Rad Chemi Doc System PC RS-170 using Quantity One (version 4.2.1) software, and manually genotype
TABLE 3

Molecular markers and map positions

| Marker Number | Marker Name | Cytological Location | D. santomea r | Genetic distance (cM) | D. yakuba r | Genetic distance (cM) |
|---------------|-------------|----------------------|---------------|-----------------------|-------------|-----------------------|
| 1             | y           | 1A5                  | 0.0432        | 0.0                   | 0.0245      | 0.0                   |
| 2             | per         | 3B1-2                | 0.1572        | 4.5                   | 0.1468      | 2.5                   |
| 3             | sog         | 13E1                 | 0.1883        | 23.4                  | 0.1364      | 19.9                  |
| 4             | v           | 9F11                 | 0.3005        | 47.0                  | 0.2690      | 35.8                  |
| 5             | rux         | 5D2                  | 0.0570        | 93.0                  | 0.0913      | 74.4                  |
| 6             | f           | 15F4-7               | 0.0777        | 99.0                  | 0.1110      | 84.5                  |
| 7             | bnb         | 17D6                 | 0.0363        | 107.5                 | 0.0329      | 97.2                  |
| 8             | Hex-A       | 8E10                 | 0.0760        | 111.2                 | 0.0593      | 100.6                 |
| 9             | AnnX        | 19C1                 | 0.0501        | 119.5                 | 0.0254      | 106.9                 |
| 10            | su(f)       | 20E                  | 0.0000        | 124.7                 | 0.0000      | 109.5                 |
| Chromosome 2  |  |  |  |  |
|--------------|---------------|--------|--------|--------|
| 11           | l(2)gl        | 21A5   | 0.0639 | 0.0    | 0.0865 | 0.0 |
| 12           | Rad1          | 23A1   | 0.1054 | 6.8    | 0.0922 | 9.5 |
| 13           | Rpl27A        | 24F3   | 0.1364 | 18.7   | 0.1665 | 19.7|
| 14           | salr          | 32E4-F1| 0.1002 | 34.6   | 0.1001 | 39.9|
| 15           | Rep4          | 34B4   | 0.2073 | 45.8   | 0.2023 | 51.2|
| 16           | His3          | 39D3-E1| 0.0345 | 72.6   | 0.0132 | 77.1|
| 17           | barr          | 38B1-2 | 0.1140 | 76.1   | 0.1477 | 78.4|
| 18           | Sara          | 57E6   | 0.2694 | 89.1   | 0.2653 | 96.0|
| 19           | Kr            | 60F5   | 0.0000 | 127.8  | 0.0000 | 133.8|

| Chromosome 3  |  |  |  |  |
|--------------|---------------|--------|--------|--------|
| 20           | Lsp1γ         | 61A6   | 0.2729 | 0.0    | 0.1966 | 0.0 |
| 21           | Dib           | 64A5   | 0.2712 | 39.5   | 0.1345 | 25.0|
| 22           | sfl           | 65B3-4 | 0.2314 | 78.5   | 0.2700 | 40.6|
| 23           | Est-6         | 69A1   | 0.3057 | 109.6  | 0.2493 | 79.5|
| 24           | Ssl1          | 80B2   | 0.2211 | 156.9  | 0.1637 | 114.0|
| 25           | ry            | 87D9   | 0.2159 | 186.1  | 0.1797 | 133.8|
|   | Marker | Chromosome | cM | cM | Recombination Rate |
|---|--------|------------|----|----|--------------------|
| 26| Rpn5   | 83C4       | 0.0967 | 214.3 | 0.1176               |
| 27| AP-50  | 94A15-16   | 0.2297 | 225.1 | 0.1844               |
| 29| Mlc1   | 98A14-15   | 0.1036 | 255.8 | 0.1072               |
| 30| ymp    | 96E        | 0.2107 | 267.4 | 0.1110               |
| 28| janB   | 99D3       | 0.2055 | 294.8 | 0.1326               |
| 31| krz    | 100E3      | 0.0000 | 321.3 | 0.0000               |

**Chromosome 4**

|   | Marker | Chromosome | cM | cM | Recombination Rate |
|---|--------|------------|----|----|--------------------|
| 32| ci     | 102A1-3    | 0.0000 | N/A | 0.0000               |

$r$ is the recombination rate between two adjacent markers. The genetic distance $d$ was inferred from $r$ using the Haldane map function, $d = -1/2 \ln(1 - 2r)$. Cytological locations are given on the basis of *D. melanogaster* cytology (LEMEUNIER and ASHBURNER 1976).
**TABLE 4**

QTLs affecting variation in pigmentation between *D. yakuba* and *D. santomea*

| Backcross Population | Sex | QTL     | Peak LR | LR | Effect (SE) | Effect/σ_p | R^2 |
|----------------------|-----|---------|---------|----|-------------|------------|-----|
| F1 females           | F   | 1A5-13E1| 1A5     | 26.71 | 0.68 (0.18) | 0.31       | 0.0169 |
| × *D. yakuba* males  |     | 15F4-20E| 8E10    | 87.08 | 1.44 (0.56) | 0.66       | 0.0584 |
|                      |     | 34B4-57E6| 34B4   | 441.74 | 1.81 (0.76) | 0.83       | 0.4382 |
|                      |     | 69A1-83C4| 80B2  | 130.57 | 1.69 (0.69) | 0.78       | 0.1135 |
| M                    | M   | 1A5-13E1| 3B1-2  | 65.13 | 2.60 (0.65) | 0.62       | 0.0327 |
|                      |     | 15F4-20E| 17D6   | 525.94 | 6.75 (2.24) | 1.62       | 0.3186 |
|                      |     | 34B4-57E6| 34B4   | 193.19 | 3.29 (0.70) | 0.79       | 0.1060 |
|                      |     | 69A1-83C4| 80B2  | 216.23 | 3.51 (0.91) | 0.84       | 0.1227 |
| Strain          | Interval       | Location | Sex | LOD Support Interval | Peak Location | LR     | Standard Error |
|-----------------|---------------|----------|-----|----------------------|---------------|--------|----------------|
| F1 females      | F 15F4-20E 19C1 | 53.485   | 53.485 | -1.39 (0.68)         | -1.27         | 0.4280 |
| × D. santomea males | M 1A5-13E1 1A5 | 67.88    | 67.88 | -2.16 (0.45)         | -0.61         | 0.0409 |
|                 | F 15F4-20E 17D6 | 450.81   | 450.81 | -5.03 (1.84)         | -1.42         | 0.3964 |
|                 | 34B4-57E6 34B4 | 199.18   | 199.18 | -1.73 (0.61)         | -0.49         | 0.1397 |
|                 | 69A1-83C4 80B2 | 133.46   | 133.46 | -2.72 (0.76)         | -0.77         | 0.0966 |

a. QTL regions are estimated from 2 LOD support intervals ($P \leq 0.05$). The peak is the cytological location with the highest likelihood ratio (LR). Cytological locations are given on the basis of *D. melanogaster* cytology (LEMEUNIER and ASHBURNER 1976).

b. Effects were estimated from the least-squares means of the two marker locus classes as: $[\bar{x}_1 - \bar{x}_2]$, where the subscript is 1 if the marker has a homozygous genotype and 2 if the marker has a heterozygous or hemizygous genotype. The standard error is listed in brackets (SE).

c. Effect divided by the phenotypic standard deviation. See footnote “a” for the calculation of the effect.

d. $R^2$ is the proportion of variance explained by the QTL and is estimated by: $R^2 = (s_0^2 - s_1^2)/s^2$, where $s^2$ is the variance of the trait, $s_0^2$ is the sample variance of the residuals, and $s_1^2$ is the variance of the residuals (BASTEN et al. 1999).
TABLE 5

Epistatic effects of QTLs affecting variation in pigmentation between *D. yakuba* and *D. santomea*

| Backcross Population                  | Sex | Markers       | Effect (SE)a | Effect/σ_p b | P-value |
|--------------------------------------|-----|---------------|--------------|--------------|---------|
| F1 females × *D. yakuba* males       | M   | 1A5 × 19C1    | 1.38 (0.30)  | 0.33         | <0.0001 |
|                                      |     | 8E10 × 57E6   | −1.54 (0.32) | −0.37        | <0.0001 |
|                                      |     | 8E10 × 87D9   | −2.50 (0.33) | −0.60        | <0.0001 |
|                                      |     | 57E6 × 87D9   | −1.33 (0.17) | −0.32        | 0.0001  |
|                                      | F   | 8E10 × 39D3-E1| −1.04 (0.25) | −0.48        | <0.0001 |

| F1 females × *D. santomea* males     | M   | 1A5 × 19C1    | 1.29 (0.30)  | 0.36         | 0.0001  |
|                                      |     | 8E10 × 57E6   | 2.59 (0.58)  | 0.73         | <0.0001 |
|                                      |     | 8E10 × 87D9   | 2.17 (0.47)  | 0.61         | <0.0001 |
|                                      |     | 57E6 × 87D9   | 1.64 (0.36)  | 0.46         | <0.0001 |

a. See text for details. The standard error (SE) is given in brackets.

b. The QTL effect divided by the phenotypic standard deviation.
| Symbol   | Gene Name          | Cytological Location | Functiona                  |
|----------|--------------------|----------------------|----------------------------|
| cin      | cinnamon           | 1A1                  | Mo-molybdopterin cofactor biosynthesis |
| y        | yellow             | 1A5                  | cuticle pigmentation        |
| a(1)HM26 | abnormal abdomen   | 1B1-5C2              | mutants affect abdominal tergite & sternite |
| mk       | murky              | 1B1-5C2              | oogenesis                   |
| svr      | silver             | 1B5--7               | cuticle biosynthesis        |
| su(b)    | suppressor of black| <1B8                 | interacts genetically with black |
| su(s)    | suppressor of sable| 1B13                 | transcriptional repressor activity |
| dor      | deep orange        | 2B5                  | pteridine biosynthesis      |
| A        | Abnormal abdomen   | 3A5                  | mutations affect the abdominal tergite |
| l(1)3B2  | lethal(1)3B2       | 3B2                  | mutations affect the tergite |
| omb      | optomotor-blind    | 4C3-4                | patterning the pigment band |
| lac      | lacquered          | 4C6                  | mutations are body color defective |
| amb      | amber              | 4C6-D1               | mutations are recessive body color defective |
| Gene | Description | Chromosome | Notes |
|------|-------------|------------|-------|
| pt   | platinum    | 7F1        | mutations are body color defective |
| t    | tan         | 8A1-B8     | beta-alanyl-dopamine hydrolase activity |
| s    | sable       | 11F1-12A1  | encodes a product involved in pigmentation |
| e(y)l| enhancer of yellow-1 | 16E1 | interacts genetically with y, w, z, ct and sc |
| e(y)3| enhancer of yellow-3 | 18C-D | interacts genetically with y, w, z, ct and sc |
| mel  | melanized   | 19B3-C3    | mutations are recessive body color defective |
| mal  | maroon-like | 19D1       | Mo-molybdopterin cofactor biosynthesis |
| mel1 | melanized-like | 19E1+    | mutations affect the abdominal tergite |
| vao  | varied outspread | 19E7 | mutations are body color defective |
| su(f)| suppressor of forked | 20E | encodes a product with putative poly(A) binding |
| b    | black       | 34D5       | glutamate decarboxylase activity |
| yellow-c | yellow-c | 35B8     | cuticle pigmentation |
| Catsup | Catecholamines up | 37B11 | regulation of catecholamine metabolism |
| Dox-A2| Diphenol oxidase A2 | 37B12 | endopeptidase activity |
| Ddc  | Dopa decarboxylase | 37C1 | dopamine biosynthesis; pigmentation patterning |
| amd  | α methyl dopa-resistant | 37C1 | involved in cuticle biosynthesis |
| l(2)37Ca | lethal(2)37Ca | 37C5 | interacts genetically with Ddc |
tyr1  tyrosine-1   38A6-C1  mutations are body color defective
pr  purple   38B3  involved in pteridine biosynthesis
Bkd  Blackoid   43E18-52D7  mutations are body color defective
dkb  dark bubbly   <49D7  mutations are body color defective
Cp1  Cysteine proteinase-1   50C18-20  mutations affect the abdominal segment 1-5
U  Upturned   53A  mutations are body color defective
Bc  Black cells   54F6  involved in melanization defense response
Pu  Punch   57C7-8  tetrahydrobiopterin biosynthesis
D  Dichaeite   70D3  mutations affect the abdominal segment 3-7
db  dark body   73C1-D2  mutations are body color defective
Crn  Crown   <77B3  mutations are dominant body color defective
kkv  krotzkopf verkehrt   83A1  chitin synthase activity

a. In cases where the function is unclear, mutant phenotypes are listed. All information, including cytological locations, was retrieved from the FlyBase website (http://www.flybase.org) (DRYSDALE and CROSBY, 2005).
**FIGURE CAPTIONS**

**FIGURE 1.** – QTLs affecting variation in pigmentation between *D. yakuba* and *D. santomea*. (A) F1 females [from *D. yakuba* (males) × *D. santomea* (females)] backcrossed to *D. yakuba* males. (B) F1 females [from *D. yakuba* (females) × *D. santomea* (males)] backcrossed to *D. santomea* males. Molecular makers are indicated as closed triangles on the X-axis. Plots are likelihood ratio (LR) test statistics for pigmentation differences between males (teal) and females (magenta) as determined by composite interval mapping. Significance thresholds for each cross were determined by permutation and are approximately LR = 10 for each cross, denoted by the dashed horizontal line.

**FIGURE 2.** – Pairwise epistasis between markers. The significance of all pairwise interactions between markers for males from the backcross to *D. santomea* is indicated above the diagonal, and for interactions for males from the backcross to *D. yakuba* below the diagonal. ■ $P < 0.0001$ (Bonferroni correction); □ $0.0001 < P < 0.001$; ▶ $0.001 < P < 0.01$; ◀ $0.01 < P < 0.05$.

**FIGURE 3.** – Significant epistatic interactions between QTLs, depicted as reaction norms. Values shown are the mean pigmentation scores (Y-axis) for a particular genotype in the background of another genotype. (A) Backcrosses to *D. santomea* males. (B) Backcrosses to *D. yakuba* males. Epistatic interactions are shown for the significant markers common to both backcross populations, as shown in Fig. 2 (M1 × M9, M8 × M18, M8 × M25, M18 × M25). See Table 3 for marker definitions. Marker genotypes are indicated as *D. santomea* (A) or *D. yakuba* (B) homozygote (●); *D. santomea/D. yakuba* heterozygote (▲).
**FIGURE 4.** – Mean pigmentation scores (Y-axis) of the eight marker haplotypes derived from the four QTLs with large effects on pigmentation. The letters denote the genotype of the QTL allele (S = *D. santomea*, Y = *D. yakuba*), and the order of the letters indicates the genotype for the QTL at the tip of the X chromosome, the QTL at the base of the X chromosome, the second chromosome QTL, and the third chromosome QTL, respectively. Haplotypes refer to the markers at the peak LR in the respective QTL analyses. See Table 3 for marker descriptions. Haplotypes for male backcross hybrids are hemizygous *D. yakuba* and/or *D. santomea* for the two X chromosome QTLs. The X chromosome QTLs in female backcross hybrids and all autosomal QTLs are heterozygous or homozygous; the marker genotype of the non-recurrent parent is indicated. (A) Males from the backcross to *D. santomea*. Haplotypes are for markers M1, M7, M15 and M24. (B) Males from the backcross to *D. yakuba*. Haplotypes are for markers M2, M7, M15 and M24. (C) Females from the backcross to *D. yakuba*. Haplotypes are for markers M1, M8, M15 and M24. The bar-graphs are color-coded (increasing ratio of black to yellow) to indicate increasing number of *D. yakuba* alleles and pigmentation scores.
A

B

Position (cM)

Likelihood Ratio

Likelihood Ratio
