doi:10.1111/evo.12146

PUPARIATION SITE PREFERENCE WITHIN AND BETWEEN *DROSOPHILA* SIBLING SPECIES

Deniz F. Erezyilmaz[1,2,3] and David L. Stern[1,4]

[1]*Howard Hughes Medical Institute and Department of Ecology & Evolutionary Biology, Princeton University, Princeton, New Jersey 08544*

[2]*Current address: Department of Biochemistry and Cell Biology and Center for Developmental Genetics, Stony Brook University, Stony Brook, New York 11794*

[3]*E-mail: deniz.erezyilmaz@stonybrook.edu*

[4]*Current address: Janelia Farm and Howard Hughes Medical Institute, Ashburn, Virginia 20147*

Received July 1, 2012
Accepted April 15, 2013

Holometabolous insects pass through a sedentary pupal stage and often choose a location for pupation that is different from the site of larval feeding. We have characterized a difference in pupariation site choice within and between sibling species of *Drosophila*. We found that, in nature, *Drosophila sechellia* pupariate within their host fruit, *Morinda citrifolia*, and that they perform this behavior in laboratory assays. In contrast, in the laboratory, geographically diverse strains of *Drosophila simulans* vary in their pupariation site preference; *D. simulans* lines from the ancestral range in southeast Africa pupariate on fruit, or a fruit substitute, whereas populations from Europe or the New World select sites off of fruit. We explored the genetic basis for the evolved preference in puariation site preference by performing quantitative trait locus mapping within and between species. We found that the interspecific difference is controlled largely by loci on chromosomes X and II. In contrast, variation between two strains of *D. simulans* appears to be highly polygenic, with the majority of phenotypic effects due to loci on chromosome III. These data address the genetic basis of how new traits arise as species diverge and populations disperse.

KEY WORDS: **Behavioral evolution, Drosophila sechellia, Drosophila simulans, innate preference, pupation.**

A central goal of evolutionary genetics is to understand how new traits arise as populations diverge. Pupation site preference varies among metamorphosing insects, even among closely related species. Transplantation experiments of pupae from species of the butterfly genus *Papilo* have shown that selection of an appropriate site for pupation is essential to confer the benefits of cryptic coloration (Hazel et al. 1980; West and Hazel 1982). *Drosophila simulans* larvae from the Central Valley of Chile prefer rough surfaces and humid substrates, whereas *Drosophila busckii* larvae prefer smooth surfaces, and *Drosophila hydei* larvae prefer dry substrates and smooth surfaces (Godoy-Herrera and Silva-Cuadra 1998).

Drosophila larvae pass through three molts after hatching from eggs. At the end of the third larval stage, release of the steroid hormone ecdysone triggers wandering behavior (Riddiford 1993). During wandering, larvae search for a location to pass through metamorphic development, where they will remain for about 5 days. Larvae of Cyclorrhaphan flies, like *Drosophila*, use their final larval cuticle, called a puparium, as a pupal case. Several studies have examined environmental cues that underlie pupariation site choice by *Drosophila* larvae. Sokolowski (1985) showed that *D. melanogaster* larvae from wild lines that were derived from pupae found on fruit tended to wander shorter distances in laboratory assays than lines derived from pupae that were isolated on nonfruit substrates. Larvae that selected pupariation sites on fruit survived at a higher rate when the soil water content was less than 50%, but showed lower survivability when the water content was above 50% (Sokolowski 1985). Moreover, larvae selected sites on fruit when soil water content was low, suggesting that they chose sites for pupariation that minimized the probability of desiccating in dry conditions and of rotting in wet conditions (Sokolowski et al. 1986). Other studies have documented the effects of density

EVOLUTION
INTERNATIONAL JOURNAL OF ORGANIC EVOLUTION

(Ringo and Wood 1982), light (Rizki and Davis 1953; Schnebel and Grossfield 1992; Paranjpe et al. 2004), pH (Beltrami 2010), and the presence of con- or heterospecifics (Beltrami 2010) on pupariation site choice.

Here we examine the genetic basis for variation in pupariation site choice both within and between species of the *D. simulans* clade. The *D. simulans* clade emerged in tropical regions of East Africa and is believed to have diverged on islands in the Indian Ocean, generating three sibling species: *D. simulans*, *D. sechellia*, and *Drosophila mauritiana* (Lachaise et al. 1998; Kliman et al. 2000). *Drosophila mauritiana* is restricted to the island of Mauritius. *Drosophila sechellia* is endemic to the Seychelles Archipelago and current evidence suggests that its diet in the wild is limited to the toxic fruit, *Morinda citrifolia* (Louis and David 1986; R'Kha et al. 1991). *Drosophila simulans*, by contrast, is a cosmopolitan generalist. Analysis of *D. simulans* genetic variation suggests that the species evolved in east Africa or Madagascar (Begun and Aquadro 1995; Andolfatto 2000; Dean and Ballard 2004; Kopp et al. 2006; Begun et al. 2007), expanded to other continents during historic times (Keller 2007), and arrived in the western United States in the last 200 years (Sturtevant 1920; Keller 2007; Begun et al. 2007). The emigration of *D. simulans* from Africa to the New World is believed to have occurred due to its association with human activity, although the precise timing of the migration is unknown (Lachaise 1998).

We employed high-resolution genotyping methods (Salathia et al. 2007; Andolfatto et al. 2011) in quantitative trait locus (QTL) mapping and found that the genetic basis for pupariation site choice within species appears to differ from the genetic basis for differences between species.

Materials and Methods

SEYCHELLES FIELD EXPERIMENTS

We examined the distribution of *D. sechellia* on fruits of its host, *M. citrifolia*, at Anse Royale Beach on Mahe Island, Seychelles from 18 October 2009 to 25 October 2009. We studied fruits from three trees that were more than 200 ft from each other. We recorded the date that each fruit fell from the tree and placed fruits in $12'' \times 24'' \times 8''$ plastic boxes. The tops of boxes were covered with cotton fabric and sealed with rubber bands to prevent *Drosophila* from entering or leaving the box. The boxes were kept outdoors in the shade. The contents of fruit and boxes were examined carefully for larvae and pupae under $2\times$ magnification. Soil was screened for pupae by sifting through 1 tablespoon at a time under $2\times$ magnification. Fruit contents were examined by immersing pieces of fruit in 20% sucrose solution, which causes larvae to float to the surface.

LABORATORY PUPARIATION SITE CHOICE ASSAYS

Stocks of *D. simulans* and *D. sechellia* were reared on a standard corn meal–based fruit fly medium. Food containing larvae was immersed in a 20% sucrose solution to collect second instar (L2) larvae for experiments.

Vial assays

Thirty L2 larvae were placed in vials containing fly food medium, and these vials were cultured in constant light at 25°C and 35% relative humidity. The location of pupae (in the food or on the vial walls) was scored after 4 days.

Petri dish assays

At the onset of the wandering stage, third instar (L3) larvae stop eating and purge food from their digestive tract prior to metamorphosis. Identification of this event was simplified by raising larvae in cornmeal-based fly food containing 0.18 mg/mL bromophenol blue. The blue dye allows observation of food in the guts of the translucent larvae. Thirty wandering larvae with less than a full gut were selected in 20% sucrose and moved to 1 tablespoon of *Morinda* fruit (that was grown in Hawaii, or in greenhouses in North Carolina), cornmeal fly food, or 15 mg/mL agar in water plugs (2.2 cm in diameter $\times$ 0.75 cm high) that were placed in 150×15 mm petri dishes. We lined the walls of each petri dish with Insect-a-Slip (Sigma, St Louis, MO) to prevent larvae from leaving the petri dish. Petri dishes were placed in a room at 25°C, with constant light and 60–70% relative humidity overnight, and the location of puparia was recorded the next day. To account for handling errors and for differential survival, we used data from assays that had at least 25, but no more than 35 puparia, in each petri dish.

GENOTYPING

We used two techniques for generating high-density genetic marker information for the offspring from genetic crosses. First, we developed a microarray-based approach that exploited insertion–deletion differences between the genomes of *D. simulans* and *D. sechellia*. We used the microarray genotyping for the *D. sechellia* backcross experiment and describe this later. We later used a next-generation sequencing approach for genotyping, called Multiplexed Shotgun Genotyping (MSG; Andolfatto et al. 2011), and used this for all other mapping experiments.

MICROARRAY GENOTYPING

DNA extraction and labeling for microarrays

Genomic DNA was extracted from individual flies with the Gentra Systems' Puregene DNA Purification Kit (Qiagen, Valenica, CA). Labeled genomic DNA (gDNA) for microarray hybridization was prepared as follows: 50 ng of gDNA from each sample

was amplified with the Sigma WGA Kit (St. Louis, MO). We obtained sufficient amplification using half of the recommended quantity of reagents. The amplified product was labeled with the Agilent CGH Genomic DNA analysis kit Version 5.0 (Santa Clara, CA). We found that reactions using one fourth the recommended reagents produced satisfactory DNA labeling.

We employed a custom Python script (https://github.com/dstern/indel_probes) to identify 6975 insertion/deletion events of between 10 and 100 bp in length from the *D. simulans* and *D. sechellia* Mercator/MAVID genomic sequence alignment (DroSim_CAF1-DroSec_CAF1.tar.gz; 12 Genomes Consortium 2007). We excluded in/del events associated with di- and tri-nucleotide microsatellites, undetermined bases (Ns), or homopolymeric runs greater than 7 bp long. We also excluded sites containing less than 35% guanine/cytosine. We designed custom oligonucleotide Agilent 8× microarrays that contained 60 bp probes specific to each allele. Replicate pure samples of *D. simulans* and *D. sechellia* gDNA were competitively hybridized to measure the expected relative hybridization intensity of alleles from each species. We used these measured intensities to assess the likelihood that the hybridization signal at each probe in an experimental animal reflected homozygosity or heterozygosity at that locus (Fig. S1). We used a running-average smoothing algorithm to identify recombination breakpoints and called genotypes at approximately every 1 Mb. The microarray data are archived on the Gene Expression Omnibus, Bioproject PRJNA19076.

GENOTYPING BY NEXT GENERATION SEQUENCING

Multiplexed DNA libraries for Illumina sequencing were generated as described in Andolfatto et al. (2011), and the 384 barcodes that we used are provided in Appendix S1. We generated updated parental genomes by sequencing each of the parental strains in a single lane of an Illumina Genome Analyzer or HiSeq and mapped the reads (using the script msgUpdateParental.pl, part of the msg package) to an improved *D. simulans* genome assembly (Hu et al. 2012). For our *D. simulans* backcross progeny, we sequenced 96 individuals in each of two lanes of an Illumina Genome Analyzer and the reads were mapped to either of the parental genomes. For our intraspecific analysis, we multiplexed 384 and 279 individuals into two separate libraries that were sequenced on an Illumina HiSeq,

GENOTYPING OF MULTIPLEXED LIBRARIES

We used Multiplexed Shotgun Genotyping (Andolfatto et al. 2011; https://github.com/JaneliaSciComp/msg) to parse bar-coded data and generate whole-genome genotype probability distributions for individual flies. We used the python script pull_thin_tsv.py (https://github.com/dstern/pull_thin) to thin the dataset to only markers that flanked a recombination event in at least one individual, which we call informative markers. We used the .csv output files for analysis in r/qtl (Broman et al. 2003).

QTL analysis

To generate recombinant flies for QTL analysis, we crossed females from *D. simulans* w^3 (Orgogozo and Stern 2006), an EMS-induced mutant generated in a strain isolated from Nueva, California (*D. simulans*^{Nueva}; San Diego Species Stock Center No.: 14021-0251.0006), to males of a *D. sechellia* white mutant (*D. sechellia*^{w30}; San Diego Species Stock Center No.: 14021-0248.30). The F_1 females were backcrossed to either *D. sechellia*^{w30} males or *D. simulans*^{Nueva} males. For the *D. sechellia*^{w30} backcross, we genotyped 45 pupae found on food and 85 collected from the wall of the vial. For the *D. simulans*^{Nueva} backcross, we genotyped 63 pupae found on the food, and 92 pupae from the wall of the vial. For the intraspecific QTL study, we crossed *D. simulans*^{Nueva} to a strain isolated from Tsimbazaza, Madagascar (*D. simulans*^{Tsi}). We performed the parental cross in both directions; *D. simulans*^{Tsi} was the paternal grandmother of 175 F2 flies and *D. simulans*^{Nueva} was the paternal grandmother of 417 flies. We performed the F2 analysis with all 592 individuals together.

For QTL analysis, we treated pupariation site choice as a binary trait. Quantitative trait locus mapping was performed using R/QTL (Broman et al. 2003) and the R/QTL interface, J/qtl (Jackson Labs). For the *D. sechellia* backcross, we used sex as an additive and interactive covariate. For each of the single QTL models, we used the Haley–Knott algorithm. We used 1000 permutation replicates to assess the statistical significance of log of the odds (LOD) scores. Finally, we graphed the LOD scores, physical distances, marker locations, and significance thresholds in R. Our *D. sechellia* backcross QTL map was generated with 192 markers and the *D. simulans* map was produced with 4423 markers. Our intraspecific dataset included 6427 markers. For the two-dimensional two-QTL scans, we used reduced datasets of the *D. simulans* backcross (667 markers) and the F2 cross (480 markers) and 100 permutation replicates to assess the significance of LOD scores. The sequence reads have been archived in the NCBI sequence read archive, Bioproject PRJNA188380.

For our microarrays, the marker locations are based upon the published *D. simulans* genome assembly (12 Genomes Consortium 2007). For our sequence data, however, we used an improved genome assembly for mapping (Hu et al. 2012), which uses the *D. melanogaster* genome as a reference guide. Throughout the text, we translate the QTL locations of our sequence data into the reference *D. simulans* genome coordinates (12 Genomes Consortium 2007), although the Hu et al. (2012) coordinates are used in the QTL maps of Figures 4B and 6B.

We used the *powercalc* and *detectable* functions of the program, R/QTLDesign (Sen et al. 2007; Broman and Sen 2009) to determine the power of the QTL experiments.

FLY STRAINS

All fly strains used in this study are listed in Appendix S2.

Generation of EYFP-marked lines

We injected a piggyBac vector containing an EYFP gene expressed in the eyes. The two strains used here are derived from *D. simulans*^{w501} (San Diego Species Stock Number Strain No. 14021-0251.011), and are part of a larger project in the Stern lab to generate a high density of dominant markers in these species, and details of these strains will be published elsewhere.

Results

We asked, first, where wild *D. sechellia* pupariate in their native habitat. Mahe Island is the largest of the Seychelles Archipelago, and its interior is montane forest. We searched Mahe Island for *M. citrifolia* trees. We found a few isolated trees and saplings in the highland forests, but these rarely had mature fruit. In the two cases where we discovered ripe fruit below isolated trees, they did not contain larvae or pupae of *D. sechellia*. In contrast, the beach at Anse Royale contained a dense grove of mature trees. Here fruit fell on sandy soil. We found that larvae appeared in the fruit as soon as 1 day after fruit fall, and remained in the fruit until adult eclosion (Fig. 1). We placed ripe and rotting *Morinda* fruit in plastic boxes that included soil, leaves, and twigs. After 3 days we examined the contents for pupae, and found that most *D. sechellia* larvae pupariated in rotting *M. citrifolia* (Fig. 1A). We were unable to find pupae outside of the fruit either in the ground below *Morinda* trees or in the soil of our field experiments (Fig. 1A).

We asked next if *D. sechellia* larvae also pupariated preferentially in fruit or a fruit substitute using petri dishes in a laboratory assay. *Drosophila sechellia* larvae pupariated in *Morinda* fruit and agar significantly more often than they pupariated away from the medium (for *Morinda* and agar, $P = 0.001$, paired *t*-test; Fig. 1B, D). (The *Morinda* fruit we used was not obtained from *M. citrifolia* trees in the Seychelles, and may have different octanoic acid content than fruit from the Seychelles beaches.) *D. sechellia* larvae were equally likely to pupariate in cornmeal-based fly food or away from the food (Fig. 1D). In the absence of any fruit or fruit substitute, *D. sechellia* do not show any preference for the center of the dish (Fig. 1D). *Drosophila simulans* larvae rarely pupariated in the fly food (five trials) or in agar (six trials; $P < 0.001$ for larval medium or agar, paired *t*-test; Fig. 1C,D). These results agree with those of Welbergen and Sokolowski (1994), who showed that pupariation height of two *D. sechellia* lines cultured in vials was significantly lower

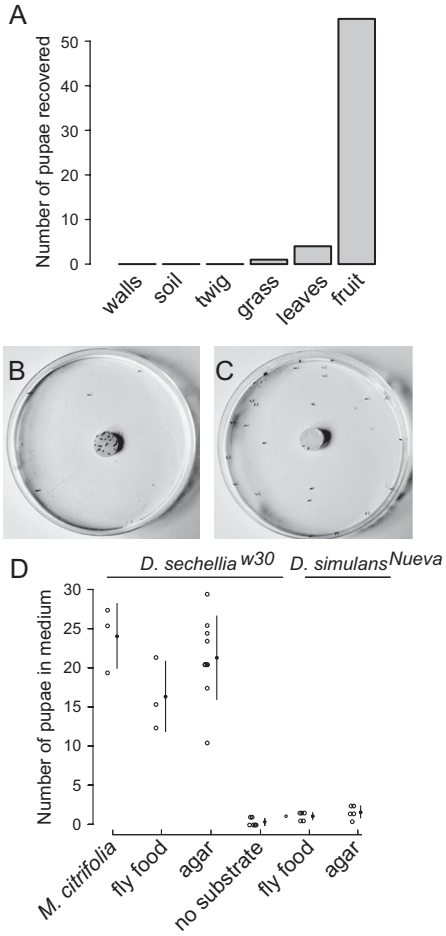

Figure 1. ***Drosophila sechellia*** **larvae pupariate most often in fruit. (A) Pupariation site selection in plastic boxes stocked with soil, fallen leaves, twigs, grass, and** ***Morinda*** **fruit. A majority of pupae were found in rotting fruit. (B,C) An example of a single test of pupariation site choice using 30 wandering larvae in petri dish assays. (B) A majority of** ***D. sechellia*** ^{w30} **pupariated on an agar disc. (C)** ***Drosophila simulans*** ^{Nueva} **larvae showed no preference for the agar disc. (D) A comparison of pupariation site choice between** ***D. sechellia*** ^{w30} **and** ***D. simulans*** ^{Nueva} **on a variety of substrates. For the "no substrate" control, we counted the number of pupae found in the center of the dish, where a substrate would otherwise be. Means are shown as filled small circles and error bars are standard deviations.**

than the average pupariation height of four *D. simulans* lines, although they detected considerable variation among *D. simulans* lines.

To determine if the difference in pupariation site choice that we observed between the *D. sechellia*^{w30} and *D. simulans*^{Nueva} strains is a fixed trait in each species, we tested strains collected from different locations in our pupariation site choice assay (Fig. 2). We found that most *D. sechellia* stocks had a strong tendency to pupariate on agar (mean $\pm$ SD $= 21.97 \pm 5.45$ in trials of 30 wandering larvae). Global populations of *D. simulans*,

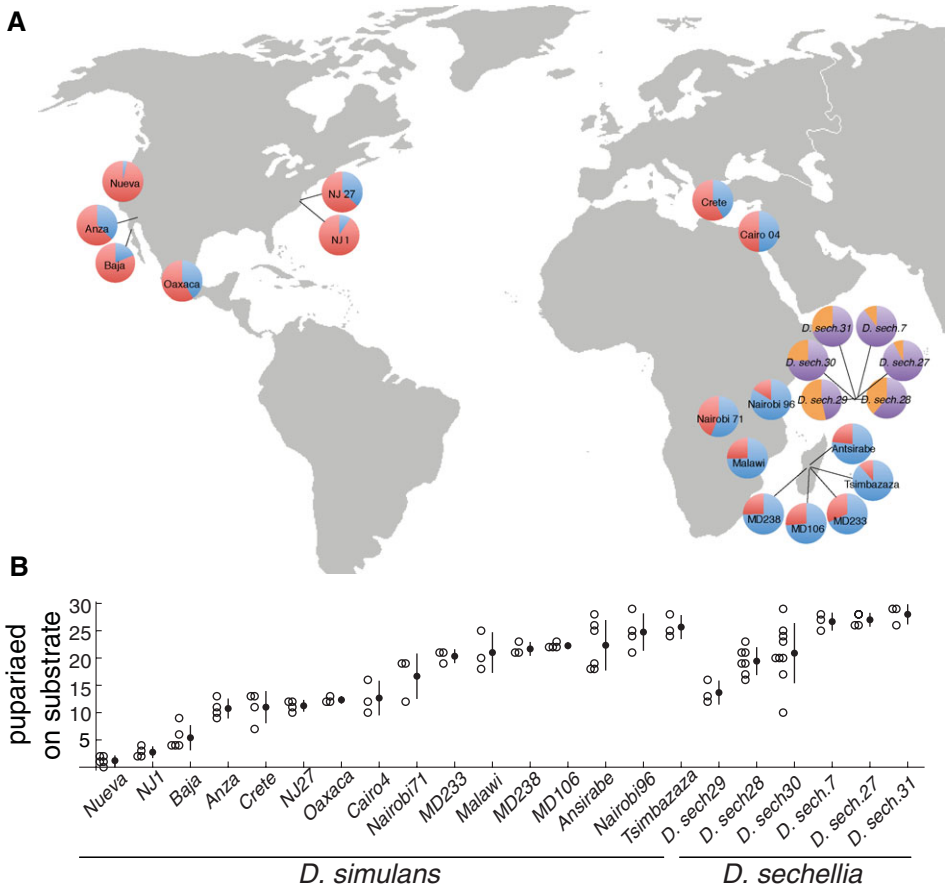

Figure 2. **Pupariation site preference of wandering larvae from six *Drosophila sechellia* lines and 16 *Drosophila simulans* lines tested in petri dish assays. (A) Global locations and pupariation site preferences for *D. sechellia* (purple and gold) and *D. simulans* (red and blue) pie charts. Purple and blue indicate the proportion of larvae that pupariated on agar. (B) A graph showing the number of pupae found on agar in trials of 30 wandering larvae (open circles). The means are indicated with closed circles, and the error bars show ± 1 SD. Nueva = a white mutant of 14021–0251.006; NJ = Princeton, New Jersey lines isolated by Josh Mast; Baja, from El Rosario, Baja California, Mexico (14021–0251.185), Anza Borrejo, CA, USA (14021–0251.268); Crete, Greece (14021–0251.181), Oaxaca, Mexico (4021–0251.180), Cairo4 from Cairo, Egypt; Nairobi71, Nairobi96, from Nairobi Kenya; MD from Madagascar; Lujeri, Malawi (14021–0251.261); Ansirabe, Madagascar (14021–0251.196), Tsimbazaza, from Ansirabe, Madagascar (NA); *D. sech.7* = 14021–0248.07; *D. sech.27* = 14021–0248.27; *D. sech.28* = 14021–0248.28; *D. sech.29* = 14021–0248.29; *D. sech.30* = 14021–0248.30; *D. sech.31* = 14021–0248.31. All of the *D. sechellia* stocks were collected on Cousin Island, Seychelles Republic, except for *D. sech.31*, which was isolated on Praslin Island.**

however, differed greatly in pupariation site choice. *Drosophila simulans* lines from Madagascar and southeast Africa were significantly different from lines that were collected outside of this region (22.42 ± 3.01 for the Madagascar lines, 11.02 ± 7.4 for the non-Madagascar lines in trials of 30 wandering lavae, $P < 0.0001$ *t*-test). Therefore, the intraspecific variation for pupariation site choice between southeast Africa and other *D. simulans* populations is as large as the average difference between species.

QTLS AFFECTING VARIATION IN PUPARIATION SITE CHOICE

The three species of the *D. simulans* clade can be crossed to produce fertile females, which permits a genetic approach to studying interspecific divergence. We examined the genetic basis of the evolved difference in pupariation site preference in interspecific hybrids between *D. simulans*^{Nueva} and *D. sechellia*^{30}. Hybrid F_1 larvae resulting from a cross between these strains pupariated on the larval substrate at a frequency that was intermediate between the pure species (Fig. 3). Larvae resulting from a backcross of F_1 female progeny to *D. simulans* males displayed pupariation site choices that did not differ significantly from the F_1 distribution (Fig. 3). In contrast, larvae resulting from the *D. sechellia*^{30} backcross pupariated in food at a frequency similar to the pure *D. sechellia*^{30} strain (Fig. 3). This result suggests that *D. sechellia* contains at least one dominant locus for pupariation site choice relative to *D. simulans*.

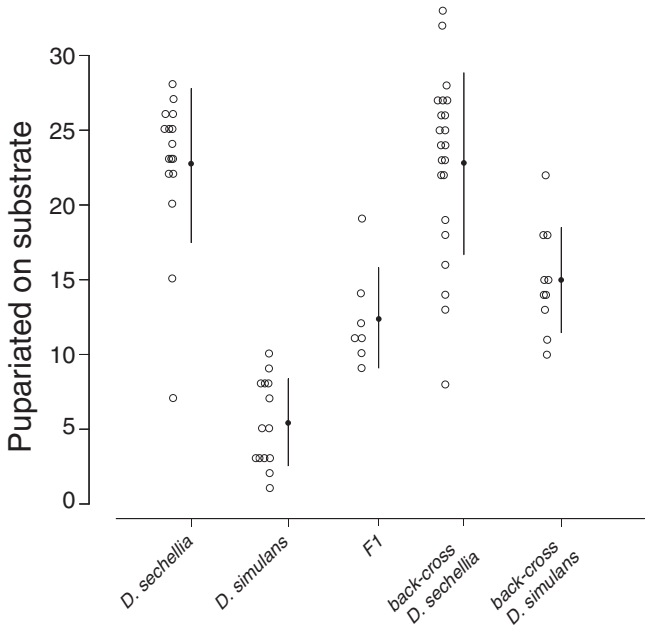

Figure 3. **Distributions of pupariation site preferences for larvae from the parental strains and crosses between *Drosophila simulans* and *Drosophila sechellia*. Data from individual trials in vials of 30 L2 larvae/trial for the parental species, F_1, and backcross progeny are shown. Open circles are the raw number of pupae on agar from each trial. Closed circles indicate the mean value, and the bars indicate ± 1 SD.**

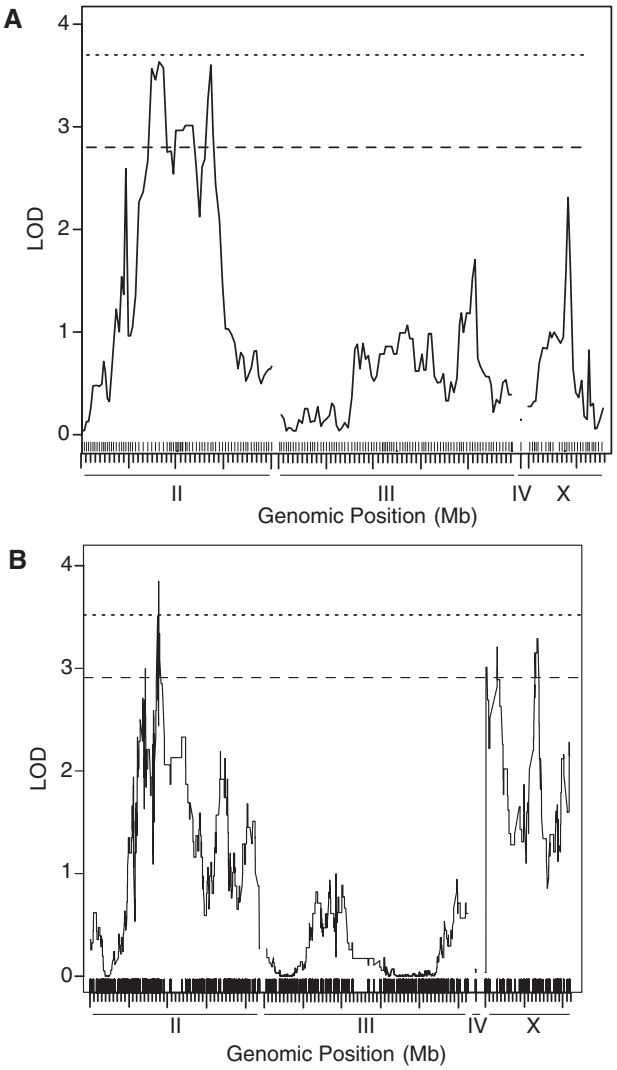

Figure 4. **Quantitative trait locus (QTL) maps for the two backcross populations. (A) QTL analysis reveals one broad significant region on chromosome II in the backcross to *Drosophila sechellia*. (B) The backcross to *Drosophila simulans* reveals significant regions on chromosomes X and II. The log of the odds (LOD) score is shown on the y-axis and genomic position on the x-axis. The dashed and dotted lines indicate LOD scores significant at $P = 0.05$ and 0.01 levels, respectively, determined by 1000 permutations. The upper rug of tick marks along the x-axes indicates marker locations, the lower tick marks are positioned every 10 Mbp along each chromosome.**

DROSOPHILA SECHELLIA BACKCROSS

In the *D. sechellia* backcross, we detected two QTL peaks on chromosome II, one at ~16,900,000 bp, and another at ~28,000,000 bp (Fig. 4A). We also observed a region on chromosome X at ~8,900,000 bp that was not significantly linked to pupariation site choice in a single QTL, additive model (LOD = 2.6, P = 0.15), but that had a large effect upon pupariation site in males (Figs. 5A, S4A). QTL II: ~16,900,000 bp shows an interaction with sex. Although female heterozygotes for this locus (*sec*/*sim*) were less likely to pupariate in food than were female homozygotes (*sec*/*sec*), male heterozygotes were four times more likely to remain in the food than were male homozygotes (Fig. S3A).

We tested for epistatic interactions using two-dimensional, 2-QTL scans (Broman et al. 2003) and found a significant interaction between regions 2: ~17,800,000 bp and X:~8,900,000 bp (LOD_f = 7.521, and $LOD_{f\text{-}v\text{-}I}$ = 4.5, LOD_i = 4.35; Figs. S2A, S3C). To determine if this interaction is distinct from the interaction between 2: ~17,800,000 bp and sex, we examined models that included both interactions separately or together. We found that although inclusion of either sex or QTL-X_{sec} as an interaction term substantially increased the likelihood of a model containing the two major loci of this map, a model including both terms showed the effect of sex is apparently due to the presence of *D. sechellia* DNA from X:~8,900,000 bp, which we call QTL-X_{sec} (Table 1). Next, we tested whether the two chromosome II peaks represent distinct QTLs. We found that a model including both peaks did not improve the likelihood score substantially over a model containing either marker alone (Table 1). We therefore treat this region on chromosome 2 as one QTL (QTL-II_{sec}; Fig. 4A).

Given our sample sizes and calculated genetic and environmental variances, a QTL must have an effect size of 0.37 or 0.34 (where 0 is pupariation in the periphery, and 1 is pupariation

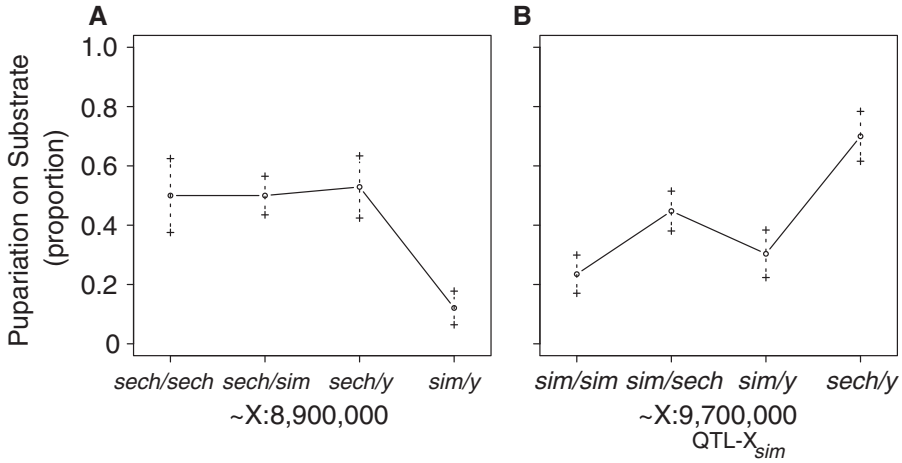

Figure 5. **Effect plots for the *Drosophila sechellia* and *Drosophila simulans* backcross populations at the X-linked quantitative trait locus (QTL). (A) The phenotype means for the four genotypes in the *D. sechellia* backcross at X:∼8,900,000 bp, and (B) for the four genotypes at X:∼9,700,000 in the *D. simulans* backcross. Error bars are ± 1 standard error. "sim" = the *D. simulans* allele of the QTL, "sec" is the *D. sechellia* allele.**

on the substrate) to be detected with a probability of 0.9 or 0.8, respectively. Therefore, only QTL with relatively large effects are detectable in our *D. sechellia* backcross, and it is possible that QTL-II is a false positive.

DROSOPHILA SIMULANS BACKCROSS

We next tested for the presence of QTL in pupariation site preference between species in the reciprocal backcross to *D. simulans*. We identified one QTL at 2: ∼17,400,000 bp (QTL-II$_{sim}$). We detected one significant peak on the X chromosome, at X:∼9,700,000 bp, and another significant region at the left end of the chromosome (Fig. 4B). A model containing both the QTL at the tip and the QTL at ∼9,700,000 in combination with QTL-II$_{sim}$ did not substantially improve the likelihood score over a model containing the QTL at X:∼9,700,000 in combination with QTL-II$_{sim}$ (Table 1). We therefore consider the locus at ∼9,700,000 to be the main QTL of the X chromosome, QTL-X$_{sim}$. QTL-X$_{sim}$ and QTL-II$_{sim}$ include the same genomic regions as the QTL identified in the back cross to *D. sechellia* (Figs. 4, S2). Like the effects of QTL-X$_{sec}$, the effects of QTL-X$_{sim}$ were expressed mainly in males. Males carrying the *D. sechellia* allele were 2.3 times as likely as males carrying the *D. simulans* allele to pupariate in food (Fig. 5B). We did not detect a significant interaction between QTL-II$_{sim}$ and sex (Fig. S3B). In contrast to the interchromosomal interactions in the *D. sechellia*[30] backcross, the effects of QTL-X$_{sim}$ and QTL-II$_{sim}$ are additive: LOD$_f$ = 8.1, LOD$_a$ = 6.7, and LOD$_{a\text{-}1}$ = 3.1 (Figs. S2B, S3B).

Our *D. simulans* backcross experiment had slightly greater power to detect QTL than did the *D. sechellia* backcross. Loci with effect sizes greater than 0.28 or 0.26 could be detected with a probability of 0.9 or 0.8, respectively, in the *D. simulans* backcross.

LINKAGE OF QTL-X TO INDEPENDENT MARKERS

Our interspecific QTL experiments suggest that a region near ∼10,000,000 on the X chromosome influences pupariation site choice (Figs. 4B, 5A, B), although these experiments were performed with relatively small sample sizes. We performed an independent test to determine if QTL-X is a false positive. We tested whether pupariation site choice segregated independently of two dominant markers that we generated in *D. simulans*. We used two strains carrying an EYFP marker on the X chromosome, one at ∼7, 100,000 bp (strain A-52.4) and one at ∼14,100,000 bp (strain A-19.1). We crossed females of these EYFP strains to *D. sechellia* males, and then backcrossed F_1 hybrid females to *D. sechellia* males. We tested 210 backcross larvae in seven trials for linkage of pupariation site choice to the A-19.1 EYFP marker, and 240 larvae in eight trials for linkage of pupariation site to the EYFP marker at A-52.4. We found that in both crosses, males carrying the EYFP markers were significantly underrepresented in larvae that pupariated in the agar (Fisher's exact test, *P*-values for A-52.4 and A-19.1 are 0.004 and 0.01, respectively; Table 2). These results confirm the existence of a QTL near ∼10,000,000 bp on the X chromosome that contributes to the difference in pupariation site choice between *D. simulans* and *D. sechellia*. The stronger linkage to A-52.4 than to A-19.1 suggests that the QTL is genetically closer to 7,100,000 bp than to 14,100,000 bp on the X chromosome.

QTL ANALYSIS OF PUPARIATION SITE PREFERENCES BETWEEN *D. SIMULANS* LINES

Our comparison of pupariation site choice between lines of *D. simulans* and *D. sechellia* showed that the difference between New World *D. simulans* and *D. sechellia* was as great as the

Table 1. Quantitative trait locus (QTL) models for the *Drosophila sechellia* and *Drosophila simulans* backcrosses.

Cross	Number of QTLs in model	LOD of model	Percent of explained	QTL locations	Percent of species difference	LOD drop one[1]
D. sechellia	2	5.4	31	X ∼ 8.9 Mbp	11.2	2.1
backcross				2 ∼ 28.0 Mbp	0.4	0.1
D. sechellia	2	5.3	30	X ∼ 8.9 Mbp	12.4	2.3
backcross				2 ∼ 17.8 Mbp	0.2	0.1
D. sechellia	3	5.5	31	X ∼ 8.9 Mbp	10.7	2.0
backcross				2 ∼ 28.0 Mbp	1.2	0.2
				2 ∼ 17.8 Mbp	0.9	0.2
D. sechellia	2 QTL,	9.3	49	X ∼ 8.9 Mbp	10.7	2.3
backcross	1 interaction			2 ∼ 17.8 Mbp	19.3	4.1
				sex	19.3	4.0
				2 ∼ 17.8 Mbp × sex	19.3	4.0
D. sechellia	2 QTL,	10.7	55	X ∼ 8.9 Mbp	37.3	7.7
backcross	1 interaction			2 ∼ 17.8 Mbp	25.1	5.4
				2 ∼ 17.8 Mbp	25.1	5.4
				× X ∼ 8.9 Mbp		
D. sechellia	2 QTL,	10.7	55	X ∼ 8.9 Mbp	16.5	3.6
backcross	2 interactions			2 ∼ 17.8 Mbp	25.1	5.4
				sex	0.0	0.0
				2 ∼ 17.8 Mbp	5.7	1.3
				× X ∼ 8.9 Mbp	0.0	0.0
				2 ∼ 17.8 Mbp		
				× sex		
D. simulans	2	7.0	32.8	X ∼ 9.7 Mbp	14.2	3.0
backcross				2 ∼ 17.4 Mbp	17.5	3.9
D. simulans	2	5.0	24.2	X ∼ 0.01 Mbp	5.8	1.3
backcross				2 ∼ 17.4 Mbp	13.0	2.8
D. simulans	3	7.5	35.1	X ∼ 0.01 Mbp	2.3	0.8
backcross				X ∼ 9.7 Mbp	8.1	1.9
				2 ∼ 17.4 Mbp	13.3	3.0

[1]Log likelihood ratios comparing the full model to a model with the specified QTL removed.

Table 2. Summary of results from the F1 backcross of *Drosophila simulans* lines A52-4 and A19.1, which each contain EYFP constructs on the X chromosome, into *Drosophila sechellia*^{w30}.

					Females			Males		
Strain	Cytological location	Physical location	Sample size[1]	Trials[2]	EYFP+[3]	EYFP−	P[4]	EYFP+	EYFP−	P[4]
A52-4	8C4	X:7,139,906	210	7	31:24	36:12	0.06	15:40	24:18	0.004
A19-1	17A10	X:14,106,580	240	8	51:20	44:8	0.1	27:32	27:10	0.01

[1]Total number of larvae tested.
[2]Number of groups of 30 L3 larvae tested in petri dish arenas.
[3]Proportion of larvae that pupariated in the substrate:periphery.
[4]Fisher's exact test.

differences between New World *D. simulans* lines and *D. simulans* lines from southeast Africa (Figs. 2, 6A). To determine if the same genomic regions that influence pupariation site choice between *D. simulans* and *D. sechellia* also influence differences between *D. simulans* populations, we performed an intercross between *D. simulans* strains showing divergent pupariation site choices and tested F2 larvae in our assay (Fig. 6). We identified multiple significant QTLs on chromosomes 2, 3, and X

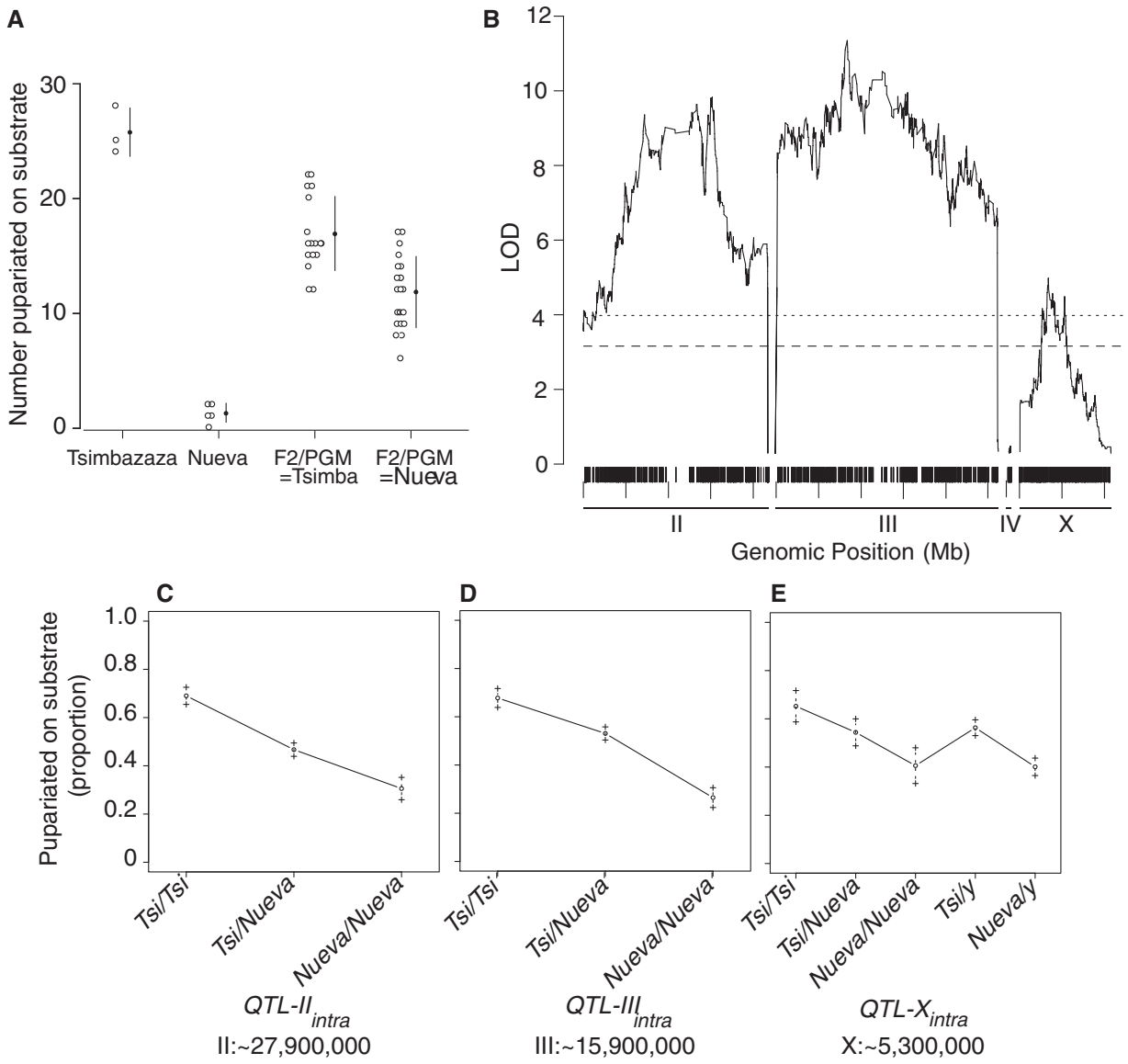

Figure 6. Intraspecific analysis of pupariation site preference. (A) Mean, SD, and individual values for tests of 30 wandering larvae of *Drosophila simulans*^{Nueva}, *D. simulans*^{Tsi}, and F2 progeny in the petri dish assay. (B) Quantitative trait locus (QTL) map of the F_2 progeny from the cross between *D. simulans*^{Nueva} and *D. simulans*^{Tsi} with log of the odds (LOD) scores plotted against physical map positions from Hu et al. (2012) genome. The dashed and dotted lines indicate the 5% and 1% significance levels, respectively. The upper "rug" of tick marks indicate the positions of markers, the lower ticks are positioned every 10 Mb along each chromosome. (C–E) Effect plots for the QTL peaks located at (C) II: ~27,900,000, (D) III: ~15,900,000, and (E) X: ~5,300,000 *D. simulans* genome coordinates (12 genomes consortium, 2007).

(Fig. 6B). The LOD scores at every location on the second and third chromosomes were significant at the 0.05 and 0.005 significance levels, respectively. On the X chromosome, only the markers between 6,000,000 bp and 14,000,000 bp were significantly linked to pupariation site choice (Fig. 6B). The LOD profiles on chromosomes 2 and 3 suggest that there are multiple QTLs on each of these chromosomes. Two-dimensional QTL scans did not reveal any significant epistatic interactions among the QTL (Fig. S2).

This experiment was powered to detect additive effect sizes as small as 0.104 and dominance effects as small as 0.148 with a probability of 0.8. Loci with dominance effect sizes as small as 0.164 and additive effect sizes as small as 0.115 could be detected with a probability of 0.9. Additive effects of magnitude larger than 0.115 were found throughout the second and third chromosomes, although dominance effects were always less than 0.1 for all chromosomes (Fig. S4C).

Table 3. Summary of tests for linkage between quantitative trait locus (QTL)-X and pupariation site preference in *Drosophila simulans* backcross progeny using ancestral and derived populations of *D. simulans,* and a *Drosophila sechellia* line marked with *vermilion* and *forked*.

			~X:8,400,000			~X:13,300,000		
Strain	Trials[1]	Sample size[2]	*vermilion*$^{+3}$	*vermilion*$^{-3}$	P^4	*forked*$^{+3}$	*forked*$^{-3}$	P^4
D. simulans. 185 (Baja, Mexico)	8	240	64:98	53:10	9.8e−10	69:90	44:18	0.0003
D. simulans (Oxnard, CA)	6	180	76:71	13:1	0.003	79:72	10:0	0.002
D. simulans. 268 (Anza Borrego Desert, CA)	8	240	79:91	31:7	0.0001	86:94	24:4	0.0002
D. simulans. 261 (Lujeri, Malawi)	8	240	110:70	19:1	0.002	112:68	17:3	0.05
D. simulans Tsimbazaza (Antsirabe, Madagascar)	6	180	124:35	9:0	0.2	126:35	7:0	0.35
D. simualns. 196 (Antsirabe, Madagascar)	7	210	96:23	26:0	0.01	106:23	20:0	0.04

[1]Number of groups of 30 L3 larvae tested in petri dish arenas, including females, which will not show the recessive markers.
[2]Total number of larvae tested.
[3]Proportion of larvae that pupariated in the substrate:periphery.
[4]Fisher's exact test.

LINKAGE OF QTL-X IN *D. SIMULANS* LINES FROM NORTH AMERICA AND SOUTHEAST AFRICA

We found that both inter- and intraspecific differences in pupariation site preference map, in part, to a location near 10,000,000 bp on the X chromosome (QTL-X). Together, our data suggest that a preference for the larval substrate at pupariation is ancestral in the *D. simulans* clade, but this preference is diminished, or a preference for off-fruit pupariation evolved, between populations of *D. simulans* from southeast Africa and the New World. To test if linkage to QTL-X differs among *D. simulans* populations, we crossed males of a strain of *D. sechellia* marked with both *vermillion* and *forked* (San Diego stock number 14021-0248.19) to females from three additional New World strains, and to three African strains, of *D. simulans*. The mutations *vermillion* and *forked* are located at genomic locations ~X:8,400,000 and ~X:13,300,000, respectively. The F_1 females were backcrossed to both parental strains, and the progeny were tested for pupariation site preference (Tables 3, 4). For the *D. simulans* backcrosses, only males will express the *vermilion* and *forked* markers. We found that, in both backcross directions, *vermillion* is linked to pupariation in agar in all but one of the crosses with *D. simulans* strains from different global locations. In addition, *forked* is linked to pupariation site choice in many of the crosses. In all but one case, *vermillion* is more closely linked to pupariation site choice than is *forked*. This observation is consistent with our mapping experiments using two EYFP markers, which positioned the QTL closer to 7,100,000 bp than to 14,100,000 bp on the X chromosome. Finally, the association between the *vermillion-forked* region and pupariation site choice is greater in California/Mexico strains than in the strains from southeast Africa (Tables 3, 4).

Discussion

ECOLOGICAL BASIS FOR PUPARIATION SITE CHOICE

Although the specialization of *D. sechellia* on the fruit of *Morinda citrifolia* is well documented, the behavior of *D. sechellia* larvae at the wandering stage in its natural environment has not been documented previously. We found that *D. sechellia* larvae pupariated on the rotting fruit of *M. citrifolia,* which falls on dry, sandy soil. Larvae that select pupariation sites in the sand may desiccate in these dry conditions. We suggest that the strong preference for host fruit pupariation in *D. sechellia*, at least for populations that live near the beach, may be related to the environment of the soil where the host fruit falls. The preference of *D. sechellia* for host fruit pupariation sites contrasts with field observations of *D. melanogaster*, which pupariates both on fruit and in the surrounding soil (Sokolowski 1985; Rodriguez et al. 1992).

Less is known of the pupariation site preferences of wild *D. simulans*. In a field study of resource utilization by *Drosophila* in grape residues at an Australian winery that was conducted over four vintage seasons, *D. melanogaster* larvae and pupae were found in grape residue during the early and late stages of fermentation, but *D. simulans* larvae and pupae were present in grape residue only during later stages. The early stages of grape fermentation are distinguished by greater water content, three-fold higher ethanol content, and a five-fold higher acetic acid

Table 4. Summary of tests for linkage between quantitative trait locus (QTL)-X and pupariation site preference in *Drosophila sechellia* backcross progeny using ancestral and derived populations of *Drosophila simulans,* and a *D. sechellia* line marked with *vermilion* and *forked*.

			∼X:8,400,000			∼X:13,300,000		
Strain	Trials[1]	Sample size[2]	*vermilion*$^{+3}$	*vermilion*$^{-3}$	P^4	*forked*$^{+3}$	*forked*$^{-3}$	P^4
D. simulans. 185 (Baja, Mexico)	14	420	74:89	114:40	2.3e−7	77:77	111:52	0.001
D. simulans (Oxnard, CA)	7	210	55:51	68:8	5.8e−8	69:40	54:19	0.15
D. simulans. 268 (Anza Borrego Desert, CA)	5	150	36:34	38:7	0.0003	40:28	34:13	0.17
D. simulans. 261 (Lujeri, Malawi)	5	150	61:20	34:3	0.05	58:15	37:8	0.8
D. simulans Tsimbazaza (Antsirabe, Madagascar)	10	300	80:46	96:7	4.2e−8	99:42	86:12	0.002
D. simualns. 196 (Antsirabe, Madagascar)	9	270	87:48	80:20	0.01	92:41	75:27	0.6

[1]Number of groups of 30 L3 larvae tested in petri dish arenas.
[2]Total number of larvae tested.
[3]Proportion of larvae that pupariated in the substrate:periphery.
[4]Fisher's exact test.

concentration (McKenzie and McKechnie 1979). In an analysis of pupariation site choice in Chilean orchards, Beltrami et al. (2010) mention that *D. simulans* pupae are found on and under the skin of fallen fruit as well as on the ground, although their study did not address the relative numbers on each substrate.

TWO MAJOR QTLS INFLUENCE THE DIFFERENCE IN PUPARIATION SITE PREFERENCE BETWEEN SPECIES

We observed QTLs for pupariation site preference on chromosomes 2 and X in both backcross directions. Although the two loci are additive in the *D. simulans* backcross, QTL-X_{sec} and QTL-II_{sec} interact substantially in the *D. sechellia* backcross. The effect of QTL-II_{sec} is entirely dependent upon the genotype at QTL-X_{sec} (Fig. S3). More than 85% of female larvae homozygous for the *D. sechellia* alleles at QTL-II_{sec} will remain in the larval substrate if QTL-X is homozygous for *D. sechellia* alleles, but only approximately 5% of males remain in the substrate if QTL-X contains a *D. simulans* allele. Although the peaks of the X-chromosome QTLs from the two backcross populations overlap, it is not possible to know whether the two QTLs reflect variation at the same locus.

Because our interspecific QTL analyses were performed with relatively small sample sizes that could result in an overestimation of QTL effects (Beavis 1998), we tested the association between the QTL-X region and pupariation site in an independent set of experiments. We used two strains of *D. simulans* that carry dominant markers at X:∼7,100,000 bp (line A52.4) and X:∼14,100,000 bp (line A19.1), which flank the peak of QTL-X_{sim}, to test whether pupariation site preference segregated with these markers. The results provide strong support for the existence of a QTL in this region that influences pupariation site preference (Table 2). These data also confirm that the majority of the QTL-X effects reflect differences in male behavior (Fig. 5), because the association of either EYFP marker with pupariation in the periphery was not significant in females (Table 2). In a separate set of tests with six different *D. simulans* strains, we asked whether a similar region of the *D. sechellia* X chromosome containing *vermilion*, at ∼X:8,400,000, and *forked*, located at ∼X:13,300,000, is associated with on-food pupariation in *D. simulans* and *D. sechellia* backcross progeny. We found linkage between *vermillion* and pupariation site preference in five out of six *D. simulans* backcrosses and in six out of six of the reciprocal *D. sechellia* backcrosses (Tables 3, 4). *forked* showed weaker linkage in these tests. These marker association tests, therefore, provide strong support for the existence of a locus on the X chromosome near 10 Mb that contributes to pupariation site choice. These positive association tests suggest that it may be possible to further refine the causal locus or loci underlying QTL-X with additional introgression experiments.

MULTIPLE QTL UNDERLIE THE DIFFERENCE IN PUPARIATION SITE CHOICE WITHIN POPULATIONS OF *D. SIMULANS*

Our intraspecific QTL analysis revealed a highly polygenic basis for variation within *D. simulans*. We detected strong QTL effects on chromosomes II, III, and X. The broad peaks on chromosomes II and III suggest that these chromosomes harbor multiple loci that

influence pupariation site choice. On the X chromosome, only the region between approximately 6 and 14 Mbp harbors a significant QTL.

Our intraspecific QTL results can be compared with a previous genetic analysis of pupariation height among *D. melanogaster* lines (Sokolowski and Bauer 1989). This study found that the majority of variation in pupariation distance from food is due to loci on chromosomes II and III, with most of the effect located on III and none on the X chromosome. Like the differences in pupariation site choice that we observed between *D. simulans* strains, the differences among *D. melanogaster* lines did not show significant epistasis. In contrast, QTL analysis of pupariation site using 76 recombinant inbred lines of *D. melanogaster* that were derived from separate strains revealed a single significant QTL on the tip of the right arm of chromosome 2 that does not show an interaction with sex (Riedl et al. 2007). This QTL does not overlap with any of the major QTL in our interspecific backcrosses, although it could correspond to a small peak on the right tip of choromosome 2 that is significant in our intraspecific cross. However, there are more differences than similarities between the QTL map of Riedl et al. (2007) and our own. It is difficult to know whether these differences are attributable to actual differences in *D. simulans* and *D. melanogaster* pupariation site loci, differences in phenotyping methods (binary vs. quantitative measures of pupariation height), or to the particular subsets of standing genetic variation that were sampled for the two studies.

It is not yet possible to determine whether QTL-X and QTL-II that we discovered in our interspecific QTL analyses are the same loci that we have discovered in our intraspecific QTL analysis of *D. simulans* strains. We are also unable to compare the effect sizes of QTLs between inter- and intraspecific crosses, because the number of progeny used for our interspecific analysis was relatively low, and our backcrosses had lower power to detect loci of minor effects than did the intraspecific cross. Loci on chromosome III, however, have the greatest influence upon differences in pupariation site preference between ancestral and New World strains of *D. simulans*. In contrast, loci on chromosome III did not have a detectable effect upon pupariation site preference in our interspecific analyses in either backcross direction. Therefore, the relative contributions of QTLs to intraspecific and interspecific variation differ, although the phenotypes in ancestral *D. simulans* populations and *D. sechellia* are indistinguishable.

All of our experiments were powered to detect mainly loci of large effect. Thus, it is possible that additional loci of small effect on pupariation site choice were not detected in our crosses. In addition, most of the QTL peaks were broad. There are several possible reasons for this. First, the interspecific experiments had relatively small sample sizes that would likely prevent high-resolution mapping of moderately complex traits. Second, *Drosophila* has a relatively low recombination rate limited to three chromosomes that contain most of the genome. This limits the resolution of QTL mapping in *Drosophila*, even with large sample sizes. Third, linked loci, combined with the low recombination rate, can generate wide QTL peaks.

EMERGENCE OF A SHIFT IN PUPARIATION SITE PREFERENCE IN THE *D. SIMULANS* CLADE

Studies of *D. mauritiana* pupariation site preference indicate that this species prefers to pupariate in fruit (Vandal et al. 2008), and the similarity in phenotype between *D. simulans* populations from southeast Africa, *D. sechellia*, and *D. mauritiana* suggests that pupariation on fruit is an ancestral trait for the *D. simulans* clade. Selection of sites off-fruit could have emerged in cosmopolitan populations as *D. simulans* emigrated from southeast Africa. Two possible scenarios could account for the shift in pupariation site preference in New World *D. simulans*. In the first scenario, a rare allele for pupariation off of fruit in the ancestral population would be enriched in a founder population, and become overrepresented after a population bottleneck. In a second scenario, populations outside of Africa would have experienced natural selection for pupariation off of fruit. For QTL-X, our association analysis shows a stronger correlation between pupariation site preference in populations from California and Mexico than in ancestral African populations. However, whether the allele underlying QTL-X became more prevalent through founder effects, or whether it emerged through selection can only be determined in future studies, once the identity of QTL-X is known.

GENETIC ARCHITECTURE OF PUPARIATION SITE CHOICE, WITHIN AND BETWEEN SPECIES

It is not clear how phenotypic differences arise and become fixed in populations. One way to address this issue is to identify the loci that underlie phenotypic differences that exist within a population, and ask whether the same loci underlie differences between species. Wittkopp et al. (2009) compared the roles of specific loci that underlie divergence in pigmentation between the species, *Drosophila novamexicana* and *Drosophila americana*, with loci that are responsible for variation in pigmentation of *D. americana*. Interspecific differences in this trait map to loci that contain the genes *tan* and *ebony*, which have established roles in pigmentation. Differences in pigmentation intensity among *D. americana* lines map to sites linked to *tan*, *ebony*, or both. In this case, intraspecific variation is a subset of the fixed differences that exist between species and fixation of shared ancestral variation may have led to the difference in pigmentation between species (Wittkopp et al. 2009).

In contrast to differences in pigmentation of *D. novamexicana*/*D. americana*, the phenotypic variation for pupariation site preference within *D. simulans* is as large as the difference between some strains of *D. simulans* and *D. sechellia*. Like Wittkopp et al.

(2009), we found extensive overlap between the QTL peaks for pupariation site choice within and between species. In contrast to Wittkopp et al. (2009), we find that more loci contribute to the intraspecific variation within *D. simulans* than cause the species difference, although we had greater power to detect effects in the intraspecific cross. The most striking difference, however, is that the third chromosome makes a large contribution to the intraspecific difference and no significant contribution to the interspecific difference.

ACKNOWLEDGMENTS

We thank J. Shapiro for input on the design and analysis of the in/del arrays; J. Buckles and D. Storton for technical help with the microarrays and next generation sequencing; R. Kuzmickas for help with Python coding; and N. Frankel, G. Kappes, R. Kuzmickas, and S. Picard for help with the *D. simulans* EYFP strains. We also thank the work of the anonymous reviewers, whose thoughtful comments contributed significantly to improving this manuscript. This work was supported by NIH F32 GM083546-02 to DFE. Our microarrays were processed and stored on PUMA-DB, which is supported by National Institutes of Health grant P50 GM071508. We also thank D. Matathiken and the Seychelles Bureau of Standards for permission to conduct research in the Seychelles. C. Jones (UNC-Chapel Hill), S. Nelson (University of Hawaii, Manoa), and S. Owens (Noni Biotech International) kindly provided us with the *M. citrifolia* fruit.

LITERATURE CITED

12 Genomes Consortium. 2007. Evolution of genes and genomes on the *Drosophila* phylogeny. Nature 450:203–218.

Andolfatto, P. 2000. Contrasting patterns of X-linked and autosomal nucleotide variation in *Drosophila melanogster* and *Drosophila simulans*. Mol. Biol. Evol. 18:279–290.

Andolfatto, P., D. Davison, D. Erezyilmaz, T. Hu, J. Mast, T. Sunayama-Morita, and D. L. Stern. 2011. Multiplexed shotgun genotyping for rapid and efficient genetic mapping. Genome Res. 4:610–617.

Beavis, W. D. 1998. QTL analyses: power, precision and accuracy. Pp. 141–161 *in* A. H. Paterson, ed. Molecular analysis of complex traits. CRC Press, Boca Raton, FL.

Begun, D. J., and C. F. Aquadro. 1995. Molecular variation at the vermillion locus in geographically diverse populations of *Drosophila melanogaster* and *D. simulans*. Genetics 140:1019–1032.

Begun, D. J., A. K. Holloway, K. Stevens, L. W. Hiller, Y. Poh, M. W. Hahn, P. M. Nista, C. D. Jones, A. D. Kern, C. N. Dewey, et al. 2007. Population genomics: whole-genome analysis of polymorphism and divergence in *Drosophila simulans*. PLoS Biol. 5:e310.

Beltrami, M., M. C. Medina-Munoz, D. Acre, and R. Godoy-Herrera. 2010. Drosophila pupation behavior in the wild. Ecol. Evol. 24:347–358.

Broman, K.W., and S. Sen. 2009. A guide to QTL mapping. Springer, New York.

Broman, K., H. Wu, S. Sen, and G. Churchill. 2003. R/QTL mapping in experimental crosses. Bioinformatics 19:8898–8890.

Dean, M. D., and J. W. O. Ballard. 2004. Linking phylogenetics with population genetics to reconstruct the geographic origin of a species. Mol. Phylogenet. Evol. 32:998–1009.

Godoy-Herrera, R., and J. L. Silva-Cuadra. 1998. The behavior of sympatric Chilean populations of *Drosophila* larvae during pupation. Genet. Mol. Biol. 21:1415–4757.

Hazel, W. N., R. Brandt, and T. Grantham. 1980. Genetic variability and phenotypic plasticity in pupal colour and its adaptive significance in the swallowtail butterfly Papilio polyxenes. Heredity 59:449–455.

Hu, T. T., M. B. Eisen, K. R. Thornton, and P. Andolfatto. 2012. A second generation assembly of the *Drosophila simulans* genome provides new insights into patterns of lineage-specific divergence. Genome Res. 23:89–98.

Keller, A. 2007. *Drosophila melanogaster's* history as a human commensal. Curr. Biol. 17:R77–R81.

Kliman, R. M., P. Andolfatto, J. A. Coyne, F. Depaulis, M. Kreitman, A. J. Berry, J. McCarter, J. Wakeley, and J. Hey. 2000. The population genetics of the origin and divergence of the *Drosophila simulans* complex species. Genetics 156:1913–1931.

Kopp, A., A. Frank, and J. Fu. 2006. Historical biogeography of *Drosophila simulans* based on Y-chromosomal sequences. Mol. Phylogenet. Evol. 38:355–362.

Lachaise, D., M.-L. Cariou, J. R. David, F. Lemeunier, L. Tsacas, and M. Ashburner. 1998. Historical biogeography of the *Drosophila melanogaster* species subgroup. Evol. Ecol. 22:159–225.

Louis, J., and J. R. David. 1986. Ecological specialization in the *Drosophila melanogaster* species subgroup: a case study of *D. sechellia*. Acta Oecol.7:215–229.

McKenzie, J. A., and S. W. McKechnie. 1979. A comparative study of resource utilization in natural populations of *Drosophila melanogaster*. Oecologia 40:299–309.

Orgogozo, V., and D. L. Stern. 2006. Thirty-three new mutations in *D. simulans*. Drosophila Inform. Service. 89:139–141.

Paranjpe, D. A., D. Anitha, V. K. Sharma, and A. Joshi. 2004. Circadian clocks and life-history related traits: is pupation height affected by circadian organization in *Drosophila melanogaster*. J. Genet. 83:73–77.

R'Kha, S., P. Capy, and J. R. David. 1991. Host-plant specialization in the *Drosophila melanogaster* species complex: a physiological, behavioral, and genetical analysis. Proc. Natl. Acad. Sci. 88:1853–1839.

Riddiford, L. M. 1993. Hormones and *Drosophila* development. Pp. 899–939 *in* M. Bate and A. Martinez-Arias, eds. The development of *Drosophila*. Cold Spring Harbor Laboratory Press, Plainview, NY.

Riedl, C. A., M. Riedl, T. F. Mackay, and M. B. Sokolowski. 2007. Genetic and behavioral analysis of natural variation in *Drosophila melanogaster* pupation position. Fly (Austin) 1:23–32.

Ringo, J., and D. Wood. 1982. Pupation site selection in *Drosophila simulans*. Behav. Genet. 13:17–27.

Rizki, T., and C. Davis. 1953. Light as an ecological determinant of interspecific competition between *Drosophila willistoni* and *Drosophila melanogaster*. Am. Nat. 87:389–392.

Rodriguez, L., M. B. Sokolowski, and J. S. Shore. 1992. Habitat selection by *Drosophila melanogaster* larvae. J. Evol. Biol. 5:61–70.

Salathia, N., H. N. Lee, T. A. Sangster, K. Morneau, C. R. Landry, K. Schellenburg, A. S. Behere, K. L. Gunderson, D. Cavalieri, G. Jander, et al. 2007. Indel arrays: an affordable alternative for genotyping. Plant J. 51:727–737.

Schnebel, E. M., and J. Grossfield. 1992. Temperature effects on pupation-height response in four *Drosophila* species group triads. J. Insect Physiol. 38:727–732.

Sen, S., J. M. Satagopan, K. W. Broman, and G. A. Churchill. 2007. R/qtlDesign: inbred line cross experimental design. Mammalian Genome 18:87–93.

Sokolowski, M. B. 1985. Genetics and Ecology of *Drosophila melanogaster* larval foraging and pupation behavior. J. Insect Physiol. 31:857–864.

Sokolowski, M. B., and S. J. Bauer. 1989. Genetic analyses of pupation distance in *Drosophila melanogaster*. Heredity 62:177–183.

Sokolowski, M. B., S. J. Bauer, V. Wai-Ping, L. Rodriguez, J. L. Wong, and C. F. Kent. 1986. Ecological genetics and behavior of *Drosophila melanogaster* larvae. Anim. Behav. 34:403–408.

Sturtevant, A. H. 1920. Genetic studies on *Drosophila simulans*. I. Introduction. Hybrids with *Drosophila melanogaster*. Genetics 5:488–500.

Vandal, N. B., G. S. Siddalingamurthy, and N. Shivanna. 2008. Larval pupation site preference on fruit in different species of *Drosophila*. Entomol. Res. 38:188–194.

Wittkopp, P. J., E. E. Stewart, A. H. Neidert, B. K. Haerum, L. L. Arnold, E. M. Thompson, G. Smith-Winberry, and L. Shefner 2009. Intraspecific polymorphism to interspecific divergence: genetics of pigmentation in *Drosophila*. Science 326:540–544.

Welbergen, P., and M. B. Sokolowski. 1994. Development time and pupation behavior in the *Drosophila melanogaster* subgroup (Diptera: Drosophilidae). J. Insect Behav. 7:263–277.

West, D. A., and W. N. Hazel. 1982. An experimental test of natural selection for pupation site in swallowtail butterflies. Evolution 36:152–159.

Associate Editor: P. Andolfatto

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Figure S1. We created custom Agilent high-density–CGH array (8 × 15K) e-array including 6975 pairs of probes (Agilent part number G4427A, Amadid #022089) *Drosophila sechellia* genomic regions are shaded, *Drosophila simulans* regions are clear.

Figure S2. Plot of interactions between loci.

Figure S3. Interactions between QTL-II and sex, and QTL-II and QTL-X for the *Drosophila sechellia* back-cross (A,C), and the *Drosophila simulans* back-cross (B,D).

Figure S4. Whole genome effect plots for (A) the *Drosophila sechellia* backcross, (B) the *Drosophila simulans* backcross, and (C) the F2 intercross.

Appendix S1. The sequence of 384 bar-coded adapters.

Appendix S2. The identity, stock number, and collection site of fly strains used in this study.