Multilocus Test for Introgression between the Cactophilic Species *Drosophila mojavensis* and *Drosophila arizonae*

Brian A. Counterman* **and Mohamed A. F. Noor**†

Levine Science Research Center, Developmental, Cell, and Molecular Biology Group, Box 91000, Duke University, Durham, North Carolina 27708

Submitted February 17, 2006; Accepted June 21, 2006; Electronically published September 29, 2006

Online enhancements: appendixes.

abstract: Information obtained from laboratory studies regarding the efficacy of barriers to gene flow (reproductive isolation) between species is often incomplete or misleading, so detailed genetic analyses are needed to determine whether hybridization and introgression occur in nature. Previous laboratory studies of the cactophilic species *Drosophila mojavensis* and *Drosophila arizonae* suggest that reproductive isolation is incomplete and that gene flow may occur in sympatry. We sampled 18 nuclear and one mitochondrial loci from multiple populations of *D. arizonae* and *D. mojavensis* to test for the signature of recent or historic gene flow between these two species. We located chromosomal regions that were inverted between these species and analyzed those regions independently of others. Statistical tests for introgression using all loci or only collinear loci failed to reject expectations of an isolation model. Further tests using average nucleotide differences between species and phylogenetic analyses also failed to find support for introgression between *D. mojavensis* and *D. arizonae*. Additional ecological and behavioral studies of these species in their natural habitats are required to explain why the signature of gene flow was not detected at the DNA sequence level in populations when laboratory studies suggest such gene flow should be possible.

Keywords: Drosophila mojavensis, *Drosophila arizonae*, introgression, chromosomal inversion, isolation model.

To understand the evolutionary forces operating on coexisting closely related species, one must determine whether natural hybridization is occurring or has occurred recently and whether such hybridization leads to interspecies introgression. Many putatively hybridizing insect species have been studied in part through laboratory investigations of inferred barriers to gene flow (reproductive isolating mechanisms), such as hybrid sterility and sexual mate preference. However, the correspondence between results of these laboratory studies and the patterns observed in nature is often mixed. For example, Llopart et al. (2005*a*) recently reported an anomalous hybrid zone between two *Drosophila* species, where the most abundant class of hybrid males was the one most difficult to produce in the laboratory and hybrid females were entirely absent for unknown reasons. Thus, it is essential that we also examine patterns of reproductive isolation using means that estimate evolutionary history, such as divergence population genetics (Kliman et al. 2000). Studies of population divergence in several species have sampled multiple loci across many individuals from many different populations to better understand their evolutionary history (e.g., Besansky et al. 2003; Machado and Hey 2003; Ramos-Onsins et al. 2004; Borge et al. 2005; Jennings and Edwards 2005; Won and Hey 2005; Bull et al. 2006; Dolman and Moritz 2006; Won et al. 2006).

Ideally, the connection between laboratory studies of reproductive isolation and evidence of interspecies gene flow from nature should be assessed in species for which abundant ecological information is available. Within the genus *Drosophila*, there have been extensive ecological studies of cactophilic species in the *Drosophila repleta* group, especially *Drosophila mojavensis* and its sibling species *Drosophila arizonae* (e.g., Etges et al. 1999). While *D. mojavensis* is endemic to the Sonoran and Mojave deserts and adjacent arid lands, the range of *D. arizonae* extends from Arizona and New Mexico in the United States to southern Mexico and Guatemala. In sympatry, these two species are known to share host cacti including cina (*Stenocereus alamosensis*) and pitaya agria (*Stenocereus gummosus*; Markow et al. 1983; Ruiz and Heed 1988).

Drosophila mojavensis and *D. arizonae* are estimated to have diverged approximately 2 million years ago (Mills et al. 1986; Matzkin and Eanes 2003). Nonetheless, several

* E-mail: brian.counterman@duke.edu.

† E-mail: noor@duke.edu.

Am. Nat. 2006. Vol. 168, pp. 682–696. © 2006 by The University of Chicago. 0003-0147/2006/16805-41647$15.00. All rights reserved.

laboratory studies of reproductive barriers between these species suggest isolation may be incomplete. Postzygotic isolation between these species appears incomplete; hybrid females are fertile, and hybrid males of crosses between some strains are also fertile (Ruiz et al. 1990; Reed and Markow 2004). Sexual isolation also appears incomplete between these species, but it has been found to vary among geographic locations and under different rearing conditions (Wasserman and Koepfer 1977; W. Etges, personal communication). Some postmating, prezygotic isolation also exists but is far from complete in the laboratory (Markow and Hocutt 1998; Knowles and Markow 2001). Finally, experimental sympatry in large population cages results in genetic introgression (Mettler and Nagle 1966). Hence, the potential for genetic introgression between these species exists.

There is some indirect evidence for hybridization between these species. A classic study found that species discrimination was stronger by strains derived from sympatry than those derived from allopatry, implying that there has been selection to prevent maladaptive hybridization (Wasserman and Koepfer 1977). Similarly, the frequency of hybrid male sterility is greater among strains derived from allopatry than among strains derived from sympatry (Ruiz et al. 1990; Reed and Markow 2004), which could result from homogenization following introgression. Such interspecies gene flow appears to have removed or prevented the evolution of alleles that confer hybrid sterility in the *Drosophila pseudoobscura* and *Drosophila persimilis* pair (Brown et al. 2004), yielding a pattern analogous to that observed among *D. mojavensis* populations. Finally, a recent study of sequence variation at three genes found the species to be "neither diagnosable nor monophyletic with respect to one another," potentially as a result of gene flow or incomplete lineage sorting (Oliveira et al. 2003, p. 223).

Despite the potential for gene flow and indirect evidence, no conclusive evidence for recent introgression between these species has been found. *Drosophila mojavensis* and *D. arizonae* differ by chromosomal inversions on their X, second, and third chromosomes, and despite several very-large-scale surveys of chromosomal inversions in *D. mojavensis*, *D. arizonae* arrangements have never been observed in *D. mojavensis* (Johnson 1980; Etges et al. 1999). However, chromosomal inversions may be generally less prone to introgression than collinear regions of the genome because they are completely linked to large segments of the genome that could confer specific adaptations or reproductive isolation between the two species (Noor et al. 2001; Rieseberg 2001; Navarro and Barton 2003*a*). That said, Reed et al. (2006) further failed to detect evidence for introgression between these species in a mitochondrial gene.

In this study, we survey polymorphism and divergence at 18 loci across inverted and collinear chromosomal regions of *D. mojavensis* and *D. arizonae*, using multiple strains derived from sympatric and allopatric populations. The recent availability of the *D. mojavensis* genome sequence and our linkage mapping efforts (see "Linkage Mapping"; Staten et al. 2004) have allowed us to identify whether these surveyed loci are within or outside regions inverted between the species, and we also directly test for differences in introgression between such regions. The observed patterns of polymorphism and divergence do not depart from the neutral expectations under the isolation model of speciation for any combination of loci. Despite evidence of incomplete reproductive isolation between these two species, we failed to detect any clear genetic signature of interspecies gene flow. We discuss the implications of these results in relation to the suggested evolution of reproductive isolation in the *D. mojavensis* species cluster.

Material and Methods

Population Sampling

Drosophila stocks were acquired from the Tucson *Drosophila* Species Stock Center (TSC; Arizona), through donation by W. J. Etges, and, in the case of some Santa Catalina Islands strains, from field collections that were submitted to the TSC. *Drosophila arizonae* flies were sampled from the following 12 strains that were originally collected from Tucson, Arizona (TU95 [Taz]); San Luis Potosi, Mexico (15081-1271.06); and 10 locations in Sonora, Mexico (15081-1271.00, Esperanza; 15081-1271.01, Rio Bavispe; 15081-1271.02, Caborca; 15081-1271.04, Navojoa; 15081-1271.108, Empalme; 15081-1271.16, Heroico Ejido Arturo; 15081-1271.18, Las Bocas; 15081-1271.23, La Pintada Canyon; A989, Ejido Porto Arturo; A990 (Daz), Las Bocas). We presume all of these strains were derived from populations sympatric with *Drosophila mojavensis.*

Drosophila mojavensis strains that were originally derived from populations sympatric to *D. arizonae* were sampled from the following locations: Los Mochis, Sinaloa, Mexico, 15081-1351.08; Santa Rosa Mountains, Arizona, 924 and A900; Rancho El Diamante, Sonora, 993 and A993; Las Bocas, Sonora, A990; El Fuerte, Sinaloa, A991; Organ Pipe National Monument, Arizona, OPNM2. These strains have been described as belonging to the subspecies *D. mojavensis baja* (Mettler 1963).

The following *D. mojavensis* strains were originally derived from populations allopatric to *D. arizonae.* The 15081-1352.10 strain was from Deubendorf Rapids, Grand Canyon, Arizona, and the following were from California: 15081-1352.00, Chocolate Mountains, Riverside County;

15081-1352.01, Anza Borrego Desert; 15081-1352.02, 15081-1352.14, and 15081-1352.22, Santa Catalina Island (SCI); 15081-1352.28, Cottonwood Beach, SCI; 15081-1352.27 and 15081-1352.29, Little Harbor, SCI. These strains have been described as belonging to the subspecies *D. mojavensis mojavensis*. The two subspecies of *D. mojavensis* have been described to be morphologically, karyotypically, and electrophoretically distinct (Mettler 1963; Zouros 1973). For some analyses, we also present SCI population samples alone because of recent interest in this population (Reed and Markow 2004). We did not survey strains of *D. mojavensis baja* from the California Baja peninsula because of sporadic reports of the presence of *D. arizonae* in the southern Cape region, which may confound appropriate designation of sympatry or allopatry required for our analyses (Fellows and Heed 1972; Ruiz et al. 1990). One line of *Drosophila navojoa*, 15081-1374.02 from Tomatlan, Jalisco, Mexico, was analyzed for some loci.

Loci Sampled

Eighteen loci across the X, second, and fourth chromosomes and one mitochondrial locus were sequenced successfully for five or more strains of *D. arizonae*, *D. mojavensis* from allopatry, and *D. mojavensis* from sympatry (for chromosomal locations, see fig. 1; table A1 in the online edition of the *American Naturalist*). Felsenstein (2006) suggests that sequencing single stretches for six to eight samples of as many loci as possible is optimal for estimating population genetic parameters. These particular loci were chosen for their proximity to previously identified microsatellite loci (Staten et al. 2004). We hoped that these loci would be predominantly in noncoding regions of the genome, but the recent genome sequence assembly suggests that three loci may contain coding region intervals. Sequencing of a single strain of *D. navojoa*, the closest known relative to *D. mojavensis* and *D. arizonae*, was attempted for all 19 loci; however, several X-linked loci could not be PCR amplified despite several attempts with numerous primers. DNA was extracted using a single-fly squish protocol (Gloor and Engels 1992). Primers were designed from either the original microsatellite clone sequences (Staten et al. 2004) or the *D. mojavensis* genome sequence (available from DroSpeGe at http://insects.eugenes.org/species/) for PCR and sequencing (table A1). Recent alignment and annotation of the *D. mojavensis* genome sequence allowed inference of putative coding sequences and determination of physical locations of our sampled loci. PCR amplification and sequencing were performed for all lines and loci listed in table A1, which contains GenBank accession numbers for all sequences used in analyses. PCR amplifications were performed using 50-μL reactions containing 0.5 μM of each primer, 5.0 μL

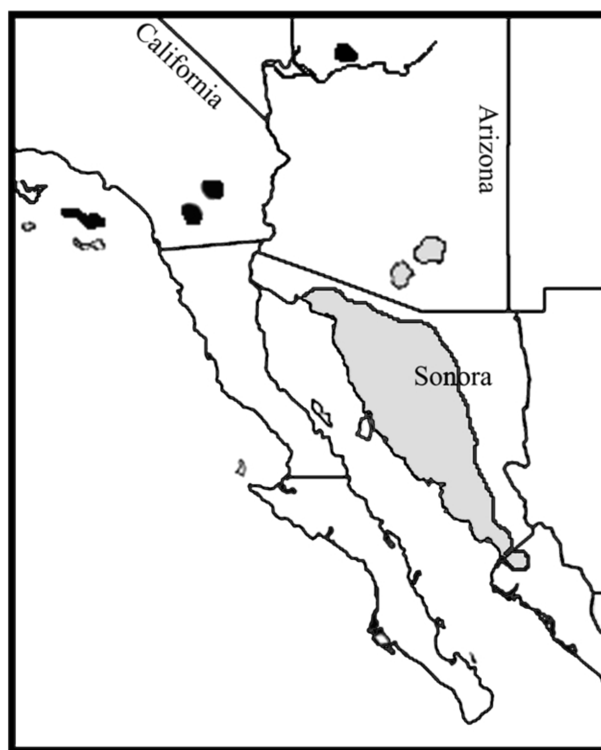

Figure 1: Geographical representation of the locations where *Drosophila mojavensis* and *Drosophila arizonae* strains were sampled. Black locations denote populations of *D. mojavensis* sampled where *D. arizonae* has been historically absent. Gray regions are locations where *D. mojavensis* and *D. arizonae* were sampled and occur in sympatry.

of 2 mM dNTPs, 10.0 μL of $\times 10$ PCR buffer (100 mM Tris, pH 8.3; 500 mM KCl; 15 mM $MgCl_2$), 0.5 μL of 50 mM $MgCl_2$, 1U Taq DNA polymerase, and 0.5 μL from a single-fly DNA preparation. Sizes of PCR products were confirmed by electrophoresis on 1% TBE agarose gels and purified with a Qiaquick gel extraction kit (Qiagen). Purified PCR products were sequenced in both directions with ABI Big Dye Terminator, version 3.1, on ABI 3700 DNA sequencers (Perkin-Elmer). Sequences were aligned using ClustalW (Thompson et al. 1994) via BioEdit (Hall 1994) and then edited manually.

Linkage Mapping

We produced a small-scale linkage map of the *D. arizonae* X chromosome and compared it with the physical and recombinational assemblies of *D. mojavensis* to identify the approximate location of the inversion breakpoints. Markers located inside the inverted region should have a reversed linkage order when compared with the physical assembly and linkage map of the X in *D. mojavensis* (Staten

et al. 2004). A female-parent backcross using *D. arizonae* strains 15081-1271.15 (collected from Safford, AZ, and acquired from the TSC) and 15081-1271.16 was reared, and 96 female progeny were genotyped for seven X-linked markers near inverted candidate loci. The Mapmaker software package (Lander et al. 1987) was used to estimate recombination distances between markers and to determine their appropriate order along the X chromosome.

Data Analysis

Basic polymorphism analyses were performed using SITES (Hey and Wakeley 1997). To determine whether the observed frequency spectrum of polymorphisms for each locus conformed to expectations under the neutral model, we estimated these test statistics: Tajima's (1989) *D*, Fu and Li's (1993) *D*, and Fay and Wu's (2000) *H*. Significant departures from neutral estimates for Tajima's *D* and Fu and Li's *D* were assessed using coalescent simulations in the SITES software package (distributed by J. Hey through http://lifesci.rutgers.edu/~heylab/). These simulated distributions of *D* values were compared with observed *D* values to determine whether each locus significantly departed from neutral expectations. One-sample sign tests were used to determine whether there was a significant excess of positive or negative *D* and *H* values across loci in each species/subspecies group.

Tests of neutrality for the three loci that contain protein coding regions (see table A1) were conducted using nonsynonymous and synonymous amino acid substitution rates estimated by DnaSP 4.0.1 (Rozas et al. 2003), according to Nei and Gojobori (1986). A ratio of nonsynonymous to synonymous substitutions (dN/dS) greater than 1 is considered to be a conservative test for detecting evidence of positive selection and lack of selective constraint on amino acid sequence divergence (Fay and Wu 2001). Another common method for detecting the signature of selection is to compare the relative number of amino acid substitutions within and between species for protein coding sequences. Under neutrality, the ratio of nonsynonymous and synonymous polymorphisms should be similar within species and between species (McDonald and Kreitman 1991). Tests for significant departures from these neutral expectations of amino acid substitution rates can be used to test for the signature of selection.

To test for the signature of gene flow between *D. mojavensis* and *D. arizonae*, we determined the fit of our data to an isolation model, which assumes no genic exchange between species, using the Wakeley-Hey (WH) program (Wakeley and Hey 1997). This model assumes that all substitutions are neutral and effective population sizes remain constant over time. Fixed, shared, and exclusive polymorphisms calculated using SITES during basic polymorphism analyses were used to estimate population parameters during the fit of the isolation model in WH. Fixed polymorphisms are nucleotide sites at which all strains of one taxon differ from all strains of the other taxon. Exclusive polymorphisms are those that are polymorphic in one taxon and invariant in the other. Shared polymorphisms are sites for which multiple nucleotides are found in both taxa. If speciation occurs by isolation without gene flow between two species, fixed differences would accumulate with time and shared polymorphisms would become increasingly rare. In an isolation model, any shared polymorphisms would be descended from variation in the ancestral species or result from recurrent parallel mutations between the two taxa. To determine whether our data deviate from the isolation model, we tested for an excess of shared polymorphisms between each species group.

Expected distributions of polymorphisms were constructed using WH. First, four population parameters—ancestral and two descendent population mutation rates (θ's) and time since separation (τ)—were estimated from observed numbers of polymorphisms. These parameter estimates were used in 10,000 coalescent simulations to create the expected distribution of polymorphisms. This method is outlined by Wang et al. (1997) and Wakeley and Hey (1997). Population recombination rates $4N_er$, where r is the recombination rate per generation, were estimated for loci with available data using Hudson's four-gamete test (Hudson et al. 1987). Both χ^2_{ss} and WH test statistics were used to test for a significant excess of shared polymorphisms between the two species, relative to the expected distribution from coalescent simulations (Kliman et al. 2000), but only χ^2_{ss} (the sum across loci of the χ^2 for shared polymorphisms) is presented, except when the tests statistics gave different results. Tests of significance were conducted in one direction (i.e., test for migration from species 1 to species 2); therefore, analysis was performed for all pairwise comparisons between *D. mojavensis* from all sampled populations, *D. mojavensis* from sympatry, and *D. arizonae.*

For the mitochondrial locus *COII*, which does not recombine, we were able to use the isolation with migration model to estimate the rate of migration for *D. arizonae* and sympatric *D. mojavensis.* Following the procedures described by Hey and Nielsen (2004), we used the program IM to determine the posterior probability density for migration, as well as several other population parameters. Using the value of θ (population mutation rate $= 4N_e\mu$) and m/μ (rate of migration per mutation event), with the highest posterior probabilities for each species, we were able to determine the population migration rate M for *D. arizonae* and sympatric *D. mojavensis* populations by solving for $M = [(\theta)(m/u)]/2 = [(4N_e\mu)(m/\mu)]/2 = 2N_em.$

The isolation-with-migration model implemented in IM assumes no intragenic recombination and may often lead to spurious results if intralocus recombination has occurred. For each nuclear locus, we found high estimates of intragenic recombination rates (see “Results and Discussion”) and therefore did not conduct analyses using the isolation-with-migration model. Similarly, simulated expectations of the isolation model are also affected by high intragenic recombination; however, in WH, estimates of population recombination rates can be provided to account for the nonindependence of segregating sites (see Kliman et al. 2000). Further, in the isolation model implemented by WH, high recombination rates between loci reduce the variance among observed polymorphisms and the distribution of the test statistics, thereby increasing confidence in WH’s tests of the isolation model. Consistent with recommendations from the authors of the WH software, we fit our data to the isolation model assuming no recombination and by providing estimates of population recombination rates available for each locus.

We also applied Feder et al.’s (2005) relative node depth (RND) test to the DNA sequence data for evidence of introgression. This test was done for each locus separately by dividing the average pairwise sequence difference (θ_π) between two populations by the average pairwise sequence difference between these populations and the outgroup *D. navojoa*. Assuming substitutions occur at a constant rate over time, RND is the amount of divergence between the two derived species relative to the total divergence from *D. navojoa*. If there has been recent introgression between *D. arizonae* and *D. mojavensis* at certain loci, RND will be reduced for those loci in the sympatric pair relative to the allopatric pair. Similarly, we predict that introgression should also reduce RND for loci within collinear regions relative to loci within inverted regions. For loci with available data from *D. navojoa*, we applied this test to all pairwise combinations between *D. arizonae*, sympatric *D. mojavensis*, and allopatric *D. mojavensis*. Rates of recombination should not affect estimates of divergence in our RND analyses.

Finally, a consensus Bayesian phylogenetic tree was constructed using sequences from 18 nuclear and one mitochondrial locus. Best fit models of substitution for each locus were obtained by hierarchical likelihood ratio tests implemented in Modeltest 3.7 (Posada and Crandall 1998). Using Mr. Bayes 3.1.2 (Ronquist and Huelsenbeck 2003), each run was initiated with a random tree and uniform priors. Four chains were run simultaneously for 5 million generations, and parameters including trees were sampled every 500 generations. The burn-in stage was determined by plotting a graph of the likelihood scores and estimated parameters. The first 2,000 trees were discarded, and the remaining 8,000 trees (post-burn-in) were used to generate a majority rule consensus tree using PAUP* 4.0b10 (Swofford 1998). To ensure that the search was not trapped in a local optimum, the Bayesian analysis was performed independently three times. For *COII* and all nuclear loci, separate Bayesian phylogenetic trees were also constructed using the same methods mentioned above.

Results and Discussion

Polymorphism and divergence data were collected for 10 X-linked, eight autosomal, and one mitochondrial locus from *Drosophila arizonae* and *Drosophila mojavensis* (table B1 in the online edition of the *American Naturalist*). Watterson’s theta (θ_w), a measure of within-species variation, is an estimate of $4N_e\mu$, where N_e is the effective population rate and μ is the mutation rate (Watterson 1975). Within-species variation was found to be significantly higher in *D. arizonae* than *D. mojavensis* when comparing the average θ_w for all loci from *D. arizonae* with *D. mojavensis* from allopatry and *D. mojavensis* from sympatry separately (Wilcoxon sign-rank test, $P = .0043$, relative to *D. mojavensis* from allopatry; $P = .0158$, relative to *D. mojavensis* from sympatry). However, when θ_w was estimated using both allopatric and sympatric populations of *D. mojavensis*, no significant difference was found for intraspecific variation between *D. mojavensis* and *D. arizonae* ($P = .939$; but see discussion of isolation between sympatric and allopatric *D. mojavensis* populations below).

Tests of the Isolation Model

Numbers of shared, fixed, and exclusive polymorphisms were counted and averaged for all loci in *D. arizonae* and *D. mojavensis* from sympatry and allopatry (table 1). Thirteen of the 19 loci surveyed possessed one or more putative fixed differences between sympatric populations of *D. mojavensis* and *D. arizonae*. Fixed differences were on average threefold more numerous than shared polymorphisms between *D. arizonae* and *D. mojavensis*. Numbers of polymorphisms and estimates of recombination rates for each locus were used to test for fit of the isolation model. Population recombination rates ($4N_er$; recombination/locus/generation) were high on average in the three taxa across loci (*D. arizonae* = 229.34; *D. mojavensis* from sympatry = 82.97; *D. mojavensis* from allopatry = 93.12), which increases the number of independent substitutions and reduces the variation in estimation of parameters used to simulate the expected χ^2_{ss} distribution. In addition to using these estimated recombination rates in our isolation model test, we also fit our data to the isolation model assuming no recombination; however, the results did not differ from those using available $4N_er$ estimates of recombination. When all 19 nuclear loci were analyzed

Table 1: Summary of nucleotide substitutions between *Drosophila* species

Locus	SS			*F*			*S* × 1			*S* × 2		
	D. arizonae–*D. mojavensis*		*D. arizonae*–*D. mojavensis*	*D. navojoa*–*D. arizonae*	*D. navojoa*–*D. mojavensis*		***D. arizonae***–*D. mojavensis*		***D. mojavensis***–*D. mojavensis*	*D. arizonae*–***D. mojavensis***		*D. mojavensis*–***D. mojavensis***
	(Allo)	(Sym)	(Allo-sym)		(Allo)	(Sym)	(**Allo**)	(**Sym**)	(**Allo**-sym)	(**Allo**)	(**Sym**)	(Allo-**sym**)
DMOJX010	0	0	0	4	4	0	7	7	3	3	2	2
DMOJX020	0	0	6	10	13	0	19	19	3	9	7	1
DMOJX030	0	0	1	1	2	0	4	4	2	3	3	2
DMOJX060	2	2	1	0	0	0	15	15	4	3	2	3
DMOJX090	0	0	0	7	10	3	11	11	5	5	1	1
DMOJX501	4	4	7	0	0	0	31	31	9	12	9	6
DMOJX504	1	0	2	5	9	1	23	24	17	18	7	5
DMOJX110	2	0	2	2	2	0	13	15	7	7	9	7
DMOJX120	0	0	0	0	0	0	14	14	8	8	3	3
DMOJX500	0	1	2	8	8	0	4	3	7	9	6	5
DMOJ2010	0	0	2	2	2	0	7	7	3	5	7	5
DMOJ2020	0	1	0	9	3	0	16	15	0	0	14	15
DMOJ2100	1	1	1	1	0	0	14	14	5	5	11	11
DMOJ2210	0	0	2	1	1	0	7	7	0	2	6	4
DMOJ4030	0	1	2	4	3	0	11	10	14	16	7	6
DMOJ4050	0	0	0	0	0	0	8	8	3	3	1	1
DMOJ4060	3	4	3	0	2	0	20	19	11	11	3	4
DMOJ4100	0	0	4	2	2	0	13	13	6	10	6	2
COII	5	6	3	0	0	0	18	17	14	12	4	7
Nuclear (average)	.72	.78	1.94	3.11	3.39	.22	13.17	13.11	5.94	7.17	5.78	4.61
Collinear (average)	1.00	1.30	2.00	2.80	3.60	.30	12.60	12.30	6.50	7.50	3.60	2.90
All loci (average)	.95	1.05	2.00	2.95	3.21	.21	13.42	13.32	6.37	7.42	5.68	4.74

Note: SS = number of shared polymorphisms; *F* = number of fixed substitution between species 1 and 2; *S* × 1 = number of exclusive polymorphisms in species 1 (bold type); *S* × 2 = number of exclusive polymorphisms in species 2 (bold type); Allo = allopatric; Sym = sympatric.

collectively, the data did not significantly depart from expectations of the isolation model between *D. arizonae* and sympatric *D. mojavensis* (χ^2_{ss} $P = .1420$, from *ariz* to *moj*; χ^2_{ss} $P = .1473$, from *moj* to *ariz* when estimates of recombination rate were provided) or between *D. arizonae* and all *D. mojavensis* populations (χ^2_{ss} $P = .1560$, from *ariz* to *moj*; χ^2_{ss} $P = .1603$, from *moj* to *ariz* when estimates of recombination rate were provided). When loci were analyzed independently, no nuclear loci had variation departing from expectations of the isolation model; however, the mitochondrial locus *COII* gave mixed results. Between *D. arizonae* and all *D. mojavensis* populations, *COII* did not fit the isolation model (χ^2_{ss} $P = .0513$, from *ariz* to *moj* and *moj* to *ariz*; WH statistic $P < .0001$ in either direction), but between *D. arizonae* and sympatric *D. mojavensis*, the χ^2_{ss} and WH test statistics gave conflicting results (χ^2_{ss} $P = .5046$, from *ariz* to *moj* and *moj* to *ariz*; WH statistic $P < .0001$ in either direction). Collectively, these results suggest there is no signature of gene flow between *D. arizonae* and *D. mojavensis* for nuclear loci, and from the limited data of the single mitochondrial locus *COII*, there may have been some historical mitochondrial gene flow between these species.

Additional unpublished sequence data from these species strongly support our conclusions regarding the lack of nuclear gene flow. Another study using 11 genes spread across the chromosomes of *D. arizonae* and *D. mojavensis* failed to reject an isolation model (C. Machado, personal communication). Hence, as in our study, Machado finds no evidence of nuclear introgression, although there is evidence of strong population structure in both species. However, Reed et al. (2006) examined mitochondrial *COI* sequences from 174 *D. mojavensis* and 102 *D. arizonae* strains and found no shared mitochrondial haplotypes and no evidence for recent introgression.

For our mitochondrial locus *COII*, we observed an abundance of shared polymorphisms and lack of fixed differences between *D. arizonae* and *D. mojavensis* (table 1). This pattern is consistent with a previous study of these species using *COII* and two nuclear loci, which concluded that *D. arizonae* and *D. mojavensis* were not monophyletic groups and either had diverged recently or still experienced gene flow occurring between sympatric populations (Oliveira et al. 2003). Also similar to Oliveira et al. (2003), we observed greater numbers of informative polymorphic sites for the mitochondrial locus than the average across all nuclear loci in our study (table B1). Since *COII* had the greatest number of shared polymorphisms of the loci studied, the estimated migration rate using *COII* provided an upper limit for migration between sympatric *D. arizonae* and *D. mojavensis*.

Because of the lack of recombination in the mitochondrial genome, we further tested for introgression at *COII*. Using the IM software, the estimated population migration rate, the rate at which genes come into the population per generation (*M*), was estimated to be 0.0334 per generation for *D. arizonae* and 0.0022 per generation for sympatric *D. mojavensis*, with neither one significantly different from 0. These very low estimates of migration rate suggest that if any mitochondrial gene flow did occur between *D. arizonae* and *D. mojavensis*, it was very limited.

As an aside, we also tested for evidence of gene flow between allopatric and sympatric *D. mojavensis*. Only two loci bore putatively fixed differences between these distinct populations/subspecies. When all 18 nuclear loci are analyzed collectively using available $4N_er$ estimates of recombination, the data significantly depart from expectations of the isolation model between allopatric and sympatric *D. mojavensis* populations using one test statistic but not the other (χ^2_{ss} $P = .0213$, from sympatric to allopatric; χ^2_{ss} $P = .0225$, from allopatric to sympatric; WH statistic $P > .10$ in either direction), suggesting possible recent or historical gene flow between them. However, when no recombination was assumed within or between loci, the observed data did not significantly violate expectations of the isolation model between these populations/ subspecies ($n = 18$, χ^2_{ss} $P > .1$ for each test). These populations/subspecies may differ in gene arrangements on the second chromosome, which is polymorphic in *D. mojavensis baja* but not in *D. mojavensis mojavensis* and may lead to an inflation of the average divergence between them. Therefore, we tested for fit of the isolation model using only X and fourth chromosome loci. Unexpectedly, after excluding second chromosome loci, we could not reject the isolation model between these populations/subspecies ($n = 14$, χ^2_{ss} $P > .1$ for each test) when either assuming no recombination or providing available $4N_er$ estimates. On balance, given these mixed (and mostly negative) results and the documented differentiation between these populations, it is premature to reject the isolation model between the two *D. mojavensis* subspecies/populations (see also our phylogenetic results in "Supporting Evidence for Lack of Interspecies Introgression: RND Test and Phylogenetic Tree").

Comparison of Inverted versus Collinear Regions

Drosophila mojavensis and *D. arizonae* differ by fixed chromosomal inversions on the second, third, and X chromosomes (Ruiz et al. 1990), and species are often unable to exchange genetic material located inside such inversions because their presence depresses homologous recombination rates (Noor et al. 2001; Rieseberg 2001; Navarro and Barton 2003*b*). If these rearrangements are old, then recent gene flow may not have been possible between the two species in the rearranged regions even if such gene

flow could have occurred in collinear regions. Therefore, a potentially more sensitive test for interspecies gene flow is to examine these parameters only at loci in collinear regions of the two species.

We compared physical and linkage maps from *D. mojavensis* and *D. arizonae* to identify which X-linked loci are in regions that are inverted versus collinear between the species. Using a *D. arizonae* backcross population, we estimated recombination distances between seven X-linked markers (fig. 2). When these were compared with marker locations from Staten et al.'s (2004) linkage map and physical genome assembly of *D. mojavensis*, we were able to narrow the location of the inversion breakpoint to a 2.2-megabase region between markers DMOJX501 and DMOJX100. Between *D. mojavensis* and *D. arizonae*, the order of markers on the X chromosome was reversed from DMOJX100 to DMOJX500. These markers are separated by more than 11.5 megabases in the *D. mojavensis* genome sequence assembly but appear completely linked within *D. arizonae*. We also observed the complete absence of recombination between these two markers in an interspecies backcross (B. A. Counterman and M. A. F. Noor, unpublished data), further supporting their proximity to the inverted region. This interspecies backcross also supported the lack of recombination between any of our second chromosome markers.

Using the information above, we identified which loci are in collinear regions of the X chromosome between these species and tested the isolation model for just these loci. This more sensitive WH test for the signature of gene flow using five X-linked and four fourth-chromosome collinear loci also failed to significantly depart from expectations of the isolation model for *D. arizonae* versus sympatric *D. mojavensis* ($n = 9$, $\chi^2_{ss} = 0.5546$, from *ariz* to *moj*; $\chi^2_{ss} = 0.6273$, from *moj* to *ariz*). We excluded the collinear markers DMOJX501 and DMOJX504 from this test because they are located near an inversion breakpoint and because the suppression of recombination associated with inversions may often extend well beyond the inversion breakpoints (Navarro et al. 1997; Andolfatto et al. 2001; Ortíz-Barrientos et al. 2006). In conclusion, all of our tests of the isolation model, using all loci and only collinear loci, were unable to detect the signature of nuclear introgression between *D. mojavensis* and *D. arizonae*.

Tests of Neutrality

Tajima's D values were calculated and averaged for all loci in each species (table B1). In *D. arizonae* and the sympatric and allopatric populations/subspecies of *D. mojavensis*, there was a significant excess of negative Tajima's D values across loci when tested via a sign test (one-sample sign test; $n = 19$, $P < .0001$; $n = 19$, $P = .0044$; and $n = 18$, $P = .0013$, respectively; see Ramos-Onsins et al. 2004 for another application of this method). Averages across all loci for Fay and Wu's H found a significant excess of negative estimates only in the allopatric *D. mojavensis* population. Similarly, Fu and Li's D found an excess of negative estimates only in *D. arizonae* populations. Tests for significant departures from neutrality of individual loci using these three test statistics revealed that estimates from only a small number loci violated neutral expectations (see table B1). The low number of loci significantly departing from neutrality for any of the three test statistics suggests there is little evidence that any single locus is under strong selection. When X-linked and autosomal loci were analyzed separately, there was more often an excess of negative test statistics on the X chromosome than on autosomes (results not shown). An excess of negative test statistic values across all loci suggests that a past demographic event such as a population expansion may have affected the polymorphism patterns across all the genome. However, incongruence between X-linked and autosomal loci suggests X-linked loci may have a unique history not shared with the rest of the genome. Collectively, there is no compelling evidence that the noncoding regions studied here violate the assumptions of the isolation model.

Using putative coding regions identified in the physical genome sequence assembly and annotation of *D. mojavensis*, we identified three of our loci as containing protein coding sequences (table A1). As expected, these loci were less variable relative to noncoding sequences (see tables 1, 2). Only one of these loci, DMOJ4050 in *D. arizonae*, displays a Tajima's D significantly less than expected under neutrality (table B1). Interestingly, DMOJ4050 contained a stop codon at the same position in two *D. arizonae* strains (15081-1271.16, A989), suggesting that functional constraints may be relaxed in these samples. All other summary statistics for the three loci did not significantly differ from simulated expectations in any species group. Rates of nonsynonymous (dN) and synonymous (dS) substitutions between *D. mojavensis* and *D. arizonae* were less than 2 and often 0 for the three loci. Further, the ratio of nonsynonymous to synonymous substitutions (dN/dS) was not greater than 1 for any pairwise comparison among all *D. arizonae* and *D. mojavensis* sequences. When comparing the numbers of nonsynonymous and synonymous substitutions within and between species, an excess number of nonsynonymous substitutions between species is the signature of positive selection (McDonald and Kreitman 1991). There was no significant difference between the estimates of within- and between-species variation in the McDonald-Kreitman test for *D. arizonae* and *D. mojavensis* from allopatry and sympatry, though there were only few substitutions in the putative coding regions. Collectively, there is no compelling evidence of selection acting

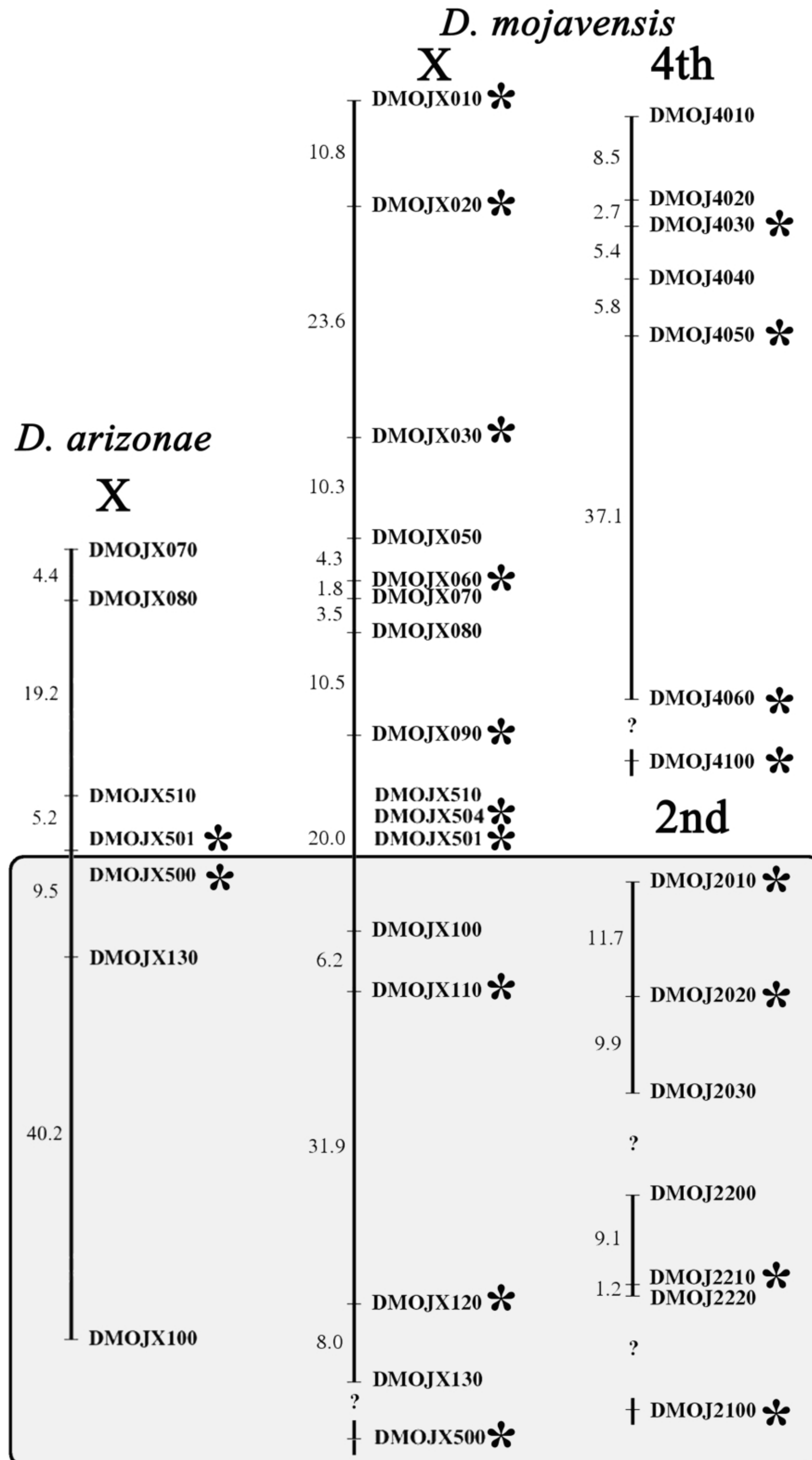

Figure 2: Linkage map of *Drosophila mojavensis* X, second, and fourth chromosomes and partial *Drosophila arizonae* X chromosome. Recombination distances are given in Kosambi centiMorgans for mapped loci; unmapped loci with unknown recombination distances are mapped using a *D. mojavensis* physical genome assembly. Chromosomal regions in shaded box are putative inverted regions between *D. mojavensis* and *D. arizonae*. The asterisks denote loci sampled in this study.

Table 2: Comparison of average group differences and relative node depth (RND) between *Drosophila* species groups

	Average group differences per base pair						RND		
	D. arizonae–D. mojavensis		*D. arizonae–D. mojavensis*	*D. navojoa–D. arizonae*	*D. navojoa–D. mojavensis*		*D. arizonae–D. mojavensis*		*D. mojavensis–D. mojavensis*
Locus	(Allo)	(Sym)	(Allo-sym)		(Allo)	(Sym)	(Allo)	(Sym)	(Allo-sym)
DMOJX010	.0104	.0103	.0018	.0362	.0394	.0416	.0686	.0660	.0108
DMOJX020	.0276	.0266	.0108	.1183	.0929	.0750	.0653	.0688	.0322
DMOJX030	.0101	.0120	.0043	.0833	.0870	.0889	.0296	.0349	.0121
DMOJX060	.0178	.0190	.0050	.0275	.0280	.0281	.1605	.1711	.0442
DMOJX090	.0460	.0567	.0163	.0996	.0801	.0862	.1280	.1527	.0489
DMOJX120	.0189	.0130	.0066	.0631	.0653	.0577	.0735	.0538	.0267
DMOJ2010	.0134	.0132	.0056	.0312	.0339	.0343	.1028	.1009	.0414
DMOJ2020	.0383	.0281	.0242	.0378	.0578	.0492	.2000	.1614	.1130
DMOJ2100	.0185	.0220	.0176	.0626	.0596	.0629	.0757	.0878	.0717
DMOJ2210	.0094	.0085	.0033	.0146	.0172	.0133	.1471	.1525	.0546
DMOJ4030	.0196	.0160	.0132	.0493	.0479	.0445	.1008	.0853	.0713
DMOJ4050	.0166	.0108	.0049	.0363	.0301	.0269	.1246	.0855	.0431
DMOJ4060	.0218	.0166	.0090	.0421	.0351	.0291	.1414	.1169	.0702
DMOJ4100	.0137	.0146	.0052	.0332	.0250	.0295	.1174	.1167	.0478
Nuclear	.0202	.0191	.0091	.0525	.0500	.0477	.1097	.1039	.0491
All loci	.0201	.0193	.0097	.0550	.0527	.0508	.1060	.1009	.0491
COII	.0196	.0221	.0181	.0904	.0908	.0946	.0541	.0597	.0488
Collinear	.0203	.0205	.0086	.0616	.0556	.0547	.0990	.0958	.0415
Inverted	.0197	.0170	NA	.0418	.0468	NA	.1198	.1113	NA

Note: For a discussion of *D* and *H* values, see "Data Analysis." NA = data required for estimation not available; Allo = allopatric; Sym = sympatric.

on the studied regions of the three putative protein coding loci sampled in *D. mojavensis* and *D. arizonae*.

Supporting Evidence for Lack of Interspecies Introgression: RND Test and Phylogenetic Tree

Relative node depth analysis was used to test for introgression between *D. mojavensis* and *D. arizonae* by looking for differences in RND values between allopatric and sympatric *D. arizonae*. *Drosophila navojoa* sequences were acquired for 11 of the nuclear loci and the mitochondrial locus *COII* and used to estimate RNDs among all three species groups. Consistent with our previous findings, average RNDs across all loci and for each individual locus were very similar between *D. arizonae* and allopatric and sympatric *D. mojavensis* (table 2). Additionally, between *D. arizonae* and sympatric *D. mojavensis*, there was no significant difference between RND estimates for inverted or collinear loci (table 2). Overall, RND analyses suggest there was no difference in divergence for allopatric and sympatric *D. mojavensis* from *D. arizonae* and that divergence was twofold higher between species than between geographically separated populations/subspecies within *D. mojavensis*.

Bayesian phylogenetic analyses of the 18 nuclear loci studied in *D. arizonae* and *D. mojavensis* revealed each to be monophyletic with respect to *D. navojoa* (fig. 3). Monophyly of *D. mojavensis* and *D. arizonae* suggests genetic divergence between these species has not been impeded by recent gene flow. Further, there does not appear to be greater differentiation in the phylogeny between allopatric *D. mojavensis* and *D. arizonae* than sympatric *D. mojavensis* and *D. arizonae*, also suggesting there has not been much recent gene flow between sympatric populations of the two species. Most noteworthy in *D. mojavensis* is the monophyletic grouping of all Santa Catalina Island strains; however, some variation does exist between every island strain sampled. This monophyly suggests Santa Catalina Island populations have had sufficient time (and/or small population size) to diverge from their closest mainland *D. mojavensis* relatives and can be considered isolated populations.

Results from phylogenetic analysis of *COII* did not resolve *D. arizonae* and *D. mojavensis* (allopatric or sympatric) into monophyletic groups (fig. 4). Additionally, neither allopatric nor sympatric populations/subspecies of *D. mojavensis* were found to be monophyletic. These results are consistent with the previous *COII* results of Oliveira et al. (2003) but discordant with the more recent study of *COI* sequence divergence in cactophilic *Drosophila* species by Reed et al. (2006). Even though there is a lack of phylogenetic resolution for *D. arizonae* and *D. moja-*

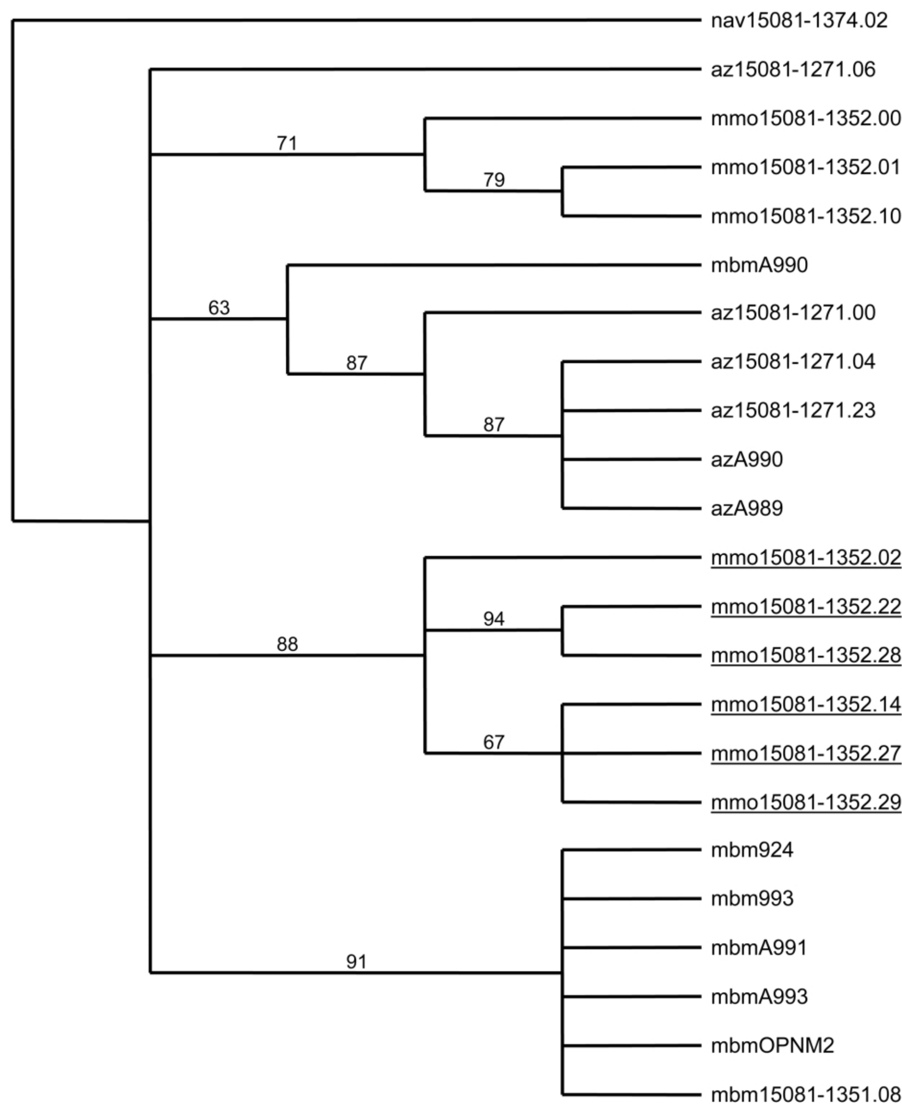

Figure 3: A 50% majority rule consensus phylogeny from 8,000 Bayesian trees using 18 nuclear loci. *az* = *Drosophila arizonae*; *mbm* = *Drosophila mojavensis*, sympatry; *mmo* = *D. mojavensis*, allopatry; *underline* = *D. mojavensis* from Santa Catalina Island, California; *nav* = *Drosophila navajoa*. Posterior probabilities are presented for their respective nodes.

vensis at *COII*, estimates of migration rate and RND suggest there has been limited (if any) gene flow between species at the *COII* loci. The inconsistencies observed between the nuclear loci and *COII* may result from an excess of retained ancestral polymorphism in mitochondria of *D. arizonae* and *D. mojavensis* or low-level historical mitochondrial gene exchange. This discrepancy is puzzling, however, because we would expect the effective population size of the mitochondrial gene to be smaller than that of the nuclear loci, and thus its coalescence time should be shorter. While this discrepancy points to the possibility for some low-level historical mitochondrial gene exchange, as seen in other organisms (i.e., Ruiz-Pesini et al. 2004; Roca et al. 2005), that conclusion is not supported by another large data set of mitochondrial sequences from these species (Reed et al. 2006).

Synopsis

No compelling evidence of introgression was found between *Drosophila mojavensis* and *Drosophila arizonae*, based on analysis of polymorphism patterns for 18 nuclear and one mitochondrial loci. A more sensitive test of gene flow was also conducted by fitting polymorphism patterns

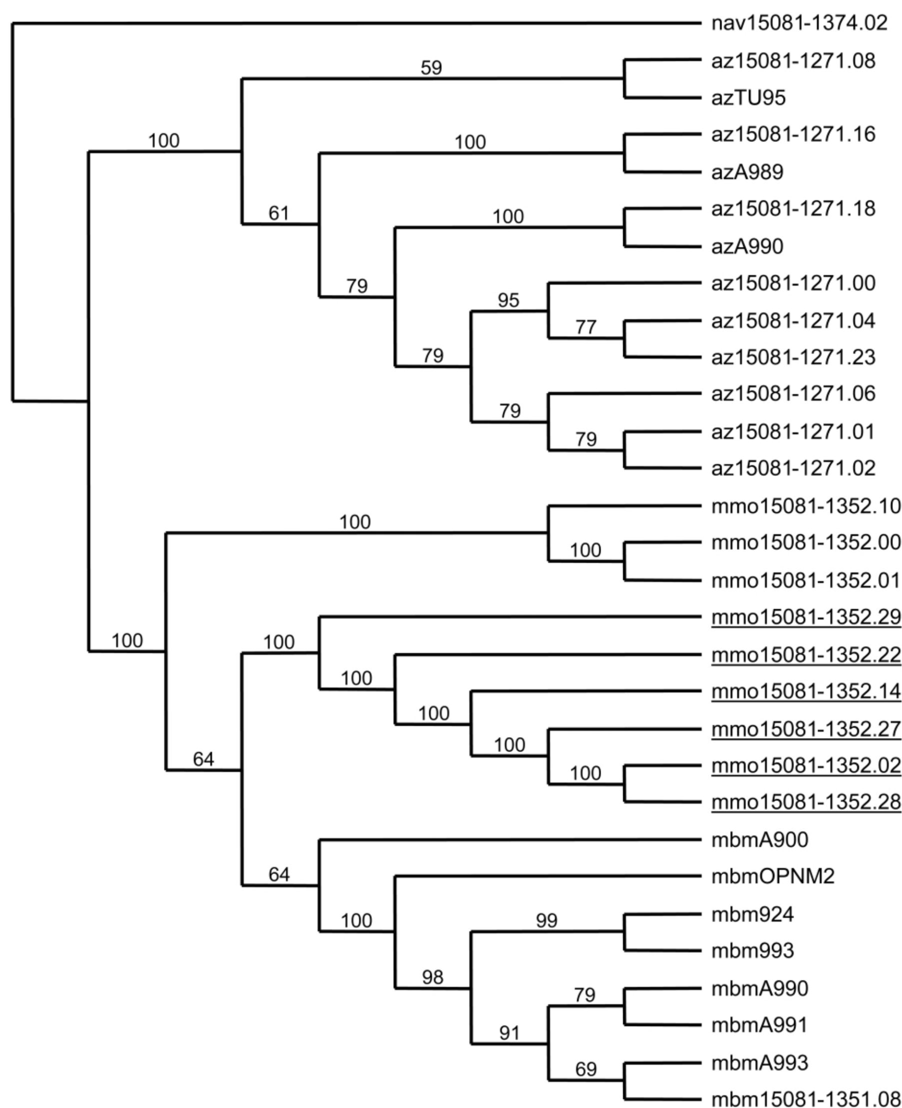

Figure 4: A 50% majority rule consensus phylogeny from 8,000 Bayesian trees using the *COII* locus. *az* = *Drosophila arizonae*; *mbm* = *Drosophila mojavensis*, sympatry; *mmo* = *D.* mojavensis *mojavensis*, allopatry; *underline* = *D. mojavensis* from Santa Catalina Island, California; *nav* = *Drosophila navajoa*. Posterior probabilities are presented for their respective nodes.

from only collinear loci to expectations of the isolation model, but this test also failed to reject an isolation model. Further analysis of relative node depths also suggests introgression is not occurring between theses two species. Results of phylogenetic analyses found *D. mojavensis* to be a monophyletic group distinct from *D. arizonae* and that the two described subspecies of *D. mojavensis* are also largely phylogenetically distinct. Support from divergence and phylogenetic analyses suggest the test for fit of an isolation model was robust to any departures from neutrality present in our data. No analyses provided any evidence for current or recent nuclear gene flow between *D. mojavensis* and *D. arizonae*, despite their sympatry in nature and incomplete reproductive isolation in laboratory studies.

When patterns of molecular evolution were collectively examined over all loci, our data departed from neutral expectations, but when analyzed individually, there was little evidence that any particular locus significantly departed from neutral expectations. A possible interpretation of these results is that the sampled populations experienced a demographic event, such as a population expansion, that affected all loci.

Collectively, this evidence strongly suggests interspecific

gene flow is not occurring in the genomic regions sampled in this study. These results were unexpected, considering that previous laboratory studies suggested reproductive isolation is incomplete between these two species. In general, meta-analyses have demonstrated reproductive isolation increasing as levels of neutral molecular divergence increase (e.g., Coyne and Orr 1997; Sasa et al. 1998; Presgraves 2002; Price and Bouvier 2002; Moyle et al. 2004). These studies and many more have documented greater molecular divergence between populations/species as barriers to gene flow increase. Correspondingly, several studies have detected the signature of hybridization or gene flow between sympatric populations of closely related species with incomplete reproductive isolation (e.g., see Mallet 2005). Despite incomplete reproductive isolation among sympatric populations of the cactophilic *Drosophila* species *D. mojavensis* and *D. arizonae*, we were unable to detect a signature of gene flow among the nuclear loci.

Although not as common, a few studies that have also found no signature of gene flow in natural populations despite evidence that reproductive isolation is incomplete. In an example mentioned previously, on the island of São Tomé, where *Drosophila santomea* and *Drosophila yakuba* come in contact along a hybrid zone, reproductive isolation appears incomplete and F_1 hybrids are often found, yet there is little evidence of gene flow among DNA sequences sampled across the genome. These two species hybridize along a contact zone (Llopart et al. 2005*a*, 2005*b*). A study of hydrenid beetles also found a discrepancy between the degree of reproductive isolation and genetic divergence (Urbanelli 2002). Among the *Ochthebius* complex, there are three species with geographically separated populations and populations with range overlap. In these species, those populations farthest from each other, which appeared to have the greatest physical barriers to gene flow, actually had intermediate levels of gene flow, while the sympatric populations that appeared to have the fewest barriers to gene flow were genetically distinct. These studies and ours stress the importance of using molecular markers to confirm suggested barriers to gene flow, such as physical barriers and sexual isolation. Identifying barriers to gene flow associated with increased molecular divergence will allow us to identify the actual mechanisms involved in the process of speciation.

For laboratory studies of reproductive isolation, it is important to recognize that laboratories cannot accommodate many unknown ecological factors that may influence interspecific reproductive frequency and success. These ecological factors may have pronounced affects on behavior and developmental patterns that distinguish these species. Without appropriate environmental cues in a homogenized laboratory environment, species-specific behavior and development patterns may also be homogenized. This could lead to the observation that some species successfully hybridize in the laboratory but not in their natural habitats. Inconsistency between studies of reproductive isolating mechanisms in the laboratory and our results from analysis of divergence population genetics suggest that the barriers to gene flow preventing introgression between *D. mojavensis* and *D. arizonae* have not been fully resolved and that estimates of isolation from laboratory studies have overestimated the potential of gene flow in the wild.

Acknowledgments

We thank A. Chang, W. J. Etges, N. Kandul, and anonymous reviewers for constructive comments on the manuscript and A. Chang and E. Gragg for technical assistance. This work was funded by a Sigma Xi grant-in-aid of research to B.A.C. and National Science Foundation grants 0211007, 0314552, and 0509780 to M.A.F.N.

Literature Cited

Andolfatto, P., F. Depaulis, and A. Navarro. 2001. Inversion polymorphisms and nucleotide variability in *Drosophila*. Genetical Research 77:1–8.

Besansky, N. J., J. Krzywinski, T. Lehmann, F. Simard, M. Kern, O. Mukabayire, D. Fontenille, et al. 2003. Semipermeable species boundaries between *Anopheles gambiae* and *Anopheles arabiensis*: evidence from multilocus DNA sequence variation. Proceedings of the National Academy of Sciences of the USA 100:10818–10823.

Borge, T., M. T. Webster, G. Andersson, and G.-P. Sætre. 2005. Contrasting patterns of polymorphism and divergence on the z chromosome and autosomes in two *Ficedula* flycatcher species. Genetics 171:1861–1873.

Brown, K. M., L. M. Burk, L. M. Henagan, and M. A. F. Noor. 2004. A test of the chromosomal rearrangement model of speciation in *Drosophila pseudoobscura*. Evolution 58:1856–1860.

Bull, V., M. Beltrán, C. D. Jiggins, W. O. McMillan, E. Bermingham, and J. Mallet. 2006. Polyphyly and gene flow between non-sibling *Heliconius* species. BMC Biology 4:11.

Coyne, J. A., and H. A. Orr. 1997. "Patterns of speciation in *Drosophila*" revisited. Evolution 51:295–303.

Dolman, G., and C. Moritz. 2006. A multilocus perspective on refugial isolation and divergence in rainforest skinks (*Carlia*). Evolution 60:573–582.

Etges, W. J., W. R. Johnson, G. A. Duncan, G. Huckins, and W. B. Heed. 1999. Ecological genetics of cactophilic *Drosophila*. Pages 164–214 *in* R. Robichaux, ed. Ecology of Sonoran Desert plants and plant communities. University of Arizona Press, Tucson.

Fay, J. C., and C.-I. Wu. 2000. Hitchhiking under positive Darwinian selection. Genetics 155:1405–1413.

———. 2001. The neutral theory in the genomic era. Current Opinion in Genetics & Development 11:642–646.

Feder, J. L., X. Xie, J. Rull, S. Velez, A. Forbes, B. Leung, H. Dambroski, et al. 2005. Mayr, Dobzhansky, and Bush and the complexities of sympatric speciation in *Rhagoletis*. Proceedings of the National Academy of Sciences of the USA 102:6573–6580.

Fellows, D. P., and W. B. Heed. 1972. Factors affecting host plant

selection in desert-adapted cactiphilic *Drosophila*. Ecology 53:850–858.

Felsenstein, J. 2006. Accuracy of coalescent likelihood estimates: do we need more sites, more sequences, or more loci? Molecular Biology and Evolution 23:691–700.

Fu, Y. X., and W. H. Li. 1993. Statistical tests of neutrality of mutations. Genetics 133:693–709.

Gloor, G. B., and W. R. Engels. 1992. Single-fly DNA preps for PCR. Drosophila Information Service 71:148–149.

Hall, T. A. 1994. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symposium Series 41:95–98.

Hey, J., and R. Nielsen. 2004. Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. Genetics 167:747–760.

Hey, J., and J. Wakeley. 1997. A coalescent estimator of the population recombination rate. Genetics 145:833–846.

Hudson, R. R., M. Kreitman, and M. Aguadé. 1987. A test of neutral molecular evolution based on nucleotide data. Genetics 116:153–159.

Jennings, W. B., and S. V. Edwards. 2005. Speciational history of Australian grass finches (*Poephila*) inferred from thirty gene trees. Evolution 59:2033–2047.

Johnson, W. R. 1980. Chromosomal polymorphism in natural populations of the desert adapted species, *Drosophila mojavensis*. PhD thesis. University of Arizona, Tucson.

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

Knowles, L. L., and T. A. Markow. 2001. Sexually antagonistic coevolution of a postmating-prezygotic reproductive character in desert *Drosophila*. Proceedings of the National Academy of Sciences of the USA 98:8692–8696.

Lander, E. S., P. Green, J. Abrahamson, A. Barlow, M. J. Daly, S. E. Lincoln, and L. Newburg. 1987. MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1:174–181.

Llopart, A., D. Lachaise, and J. A. Coyne. 2005*a*. An anomalous hybrid zone in *Drosophila*. Evolution 59:2602–2607.

———. 2005*b*. Multilocus analysis of introgression between two sympatric sister species of *Drosophila*: *Drosophila yakuba* and *D. santomea*. Genetics 171:197–210.

Machado, C. A., and J. Hey. 2003. The causes of phylogenetic conflict in a classic *Drosophila* species group. Proceedings of the Royal Society B: Biological Sciences 270:1193–1202.

Mallet, J. 2005. Hybridization as an invasion of the genome. Trends in Ecology & Evolution 20:229–237.

Markow, T. A., and G. D. Hocutt. 1998. Reproductive isolation in Sonoran desert *Drosophila* testing the limits of the rules. Pages 234–244 *in* D. J. Howard and S. H. Berlocher, eds. Endless forms: species and speciation. Oxford University Press, New York.

Markow, T. A., J. C. Fogleman, and W. B. Heed. 1983. Reproductive isolation in Sonoran Desert *Drosophila*. Evolution 37:649–652.

Matzkin, L. M., and W. F. Eanes. 2003. Sequence variation of alcohol dehydrogenase (Adh) paralogs in cactophilic *Drosophila*. Genetics 163:181–194.

McDonald, J. H., and M. Kreitman. 1991. Adaptive protein evolution at the Adh locus in *Drosophila*. Nature 351:652–654.

Mettler, L. E. 1963. *Drosophila mojavensis baja*, a new form in the *mulleri* complex. Drosophila Information Service 38:57–58.

Mettler, L. E., and J. J. Nagle. 1966. Corroboratory evidence for the concept of the sympatric origin of isolating mechanisms. Drosophila Information Service 41:76.

Mills, L. E., P. Batterham, J. Alegre, W. T. Starmer, and D. T. Sullivan. 1986. Molecular genetic characterization of a locus that contains duplicate Adh genes in *Drosophila mojavensis* and related species. Genetics 112:295–310.

Moyle, L. C., M. S. Olson, and T. P. 2004. Patterns of reproductive isolation in three angiosperm genera. Evolution 58:1195–1208.

Navarro, A., and N. H. Barton. 2003*a*. Accumulating postzygotic isolation in parapatry: a new twist on chromosomal speciation. Evolution 57:447–459.

———. 2003*b*. Chromosomal speciation and molecular divergence: accelerated evolution in rearranged chromosomes. Science 300: 321–324.

Navarro, A., E. Betran, A. Barbadilla, and A. Ruiz. 1997. Recombination and gene flux caused by gene conversion and crossing over in inversion heterokaryotypes. Genetics 146:695–709.

Nei, M., and T. Gojobori. 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Molecular Biology and Evolution 3:418–426.

Noor, M. A., K. L. Grams, L. A. Bertucci, and J. Reiland. 2001. Chromosomal inversions and the reproductive isolation of species. Proceedings of the National Academy of Sciences of the USA 98: 12084–12088.

Oliveira, D. C. S. G., E. H. Wintermute, W. J. Etges, W. B. Heed, P. M. O'Grady, and R. Desalle. 2003. Hierarchical structure in the *Drosophila mojavensis* cluster (Diptera: Drosophilidae). Hereditas 139:223–227.

Ortíz-Barrientos, D., A. S. Chang, and M. A. F. Noor. 2006. A recombinational portrait of the *Drosophila pseudoobscura* genome. Genetical Research 87:23–31.

Posada, D., and K. Crandall. 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14:817–818.

Presgraves, D. C. 2002. Patterns of postzygotic isolation in Lepidoptera. Evolution 56:1168–1183.

Price, T. D., and M. M. Bouvier. 2002. The evolution of F_1 postzygotic incompatibilities in birds. Evolution 56:2083–2089.

Ramos-Onsins, S. E., B. E. Stranger, T. Mitchell-Olds, and M. Aguade. 2004. Multilocus analysis of variation and speciation in the closely related species *Arabidopsis halleri* and *A. lyrata*. Genetics 166:373–388.

Reed, L. K., and T. A. Markow. 2004. Early events in speciation: polymorphism for hybrid male sterility in *Drosophila*. Proceedings of the National Academy of Sciences of the USA 101:9009–9012.

Reed, L. K., M. Nyboer, and T. A. Markow. 2006. Evolutionary relationships of *Drosophila mojavensis* geographic host races and their sister species *Drosophila arizonae*. Molecular Ecology, doi: 10.1111/j.1365-294X.2006.02941.x.

Rieseberg, L. H. 2001. Chromosomal rearrangements and speciation. Trends in Ecology & Evolution 16:351–358.

Roca, A. L., N. Georgiadis, and S. J. O'Brien. 2005. Cytonuclear genomic dissociation in African elephant species. Nature Genetics 37:96–100.

Ronquist, F., and J. P. Huelsenbeck. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572–1574.

Rozas, J., J. C. Sánchez-DelBarrio, X. Messeguer, and R. Rozas. 2003.

DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 19:2496–2497.

Ruiz, A., and W. B. Heed. 1988. Host-plant specificity in the cactophilic *Drosophila mulleri* species complex. Journal of Animal Ecology 57:237–249.

Ruiz, A., W. B. Heed, and M. Wasserman. 1990. Evolution of the *mojavensis* cluster of cactophilic *Drosophila* with descriptions of two new species. Journal of Heredity 81:30–42.

Ruiz-Pesini, E., D. Mishmar, M. Brandon, V. Procaccio, and D. C. Wallace. 2004. Effects of purifying and adaptive selection on regional variation in human mtDNA. Science 303:223–226.

Sasa, M. M., P. T. Chippindale, and N. A. Johnson. 1998. Patterns of postzygotic isolation in frogs. Evolution 52:1811–1820.

Staten, R., S. D. Schully, and M. A. F. Noor. 2004. A microsatellite linkage map of *Drosophila mojavensis*. BMC Genetics 5:5.

Swofford, D. L. 1998. PAUP*: phylogenetic analysis using parsimony (* and other methods). Sinauer, Sunderland, MA.

Tajima, F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123:585–595.

Thompson, J. D., D. G. Higgins, and T. J. Gibson. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. Nucleic Acids Research 22:4673–4680.

Urbanelli, S. 2002. Genetic divergence and reproductive isolation in the *Ochthebius* (Calobius) complex (Coleoptera: Hydraenidae). Heredity 88:333–341.

Wakeley, J., and J. Hey. 1997. Estimating ancestral population parameters. Genetics 145:847–855.

Wang, R. L., J. Wakeley, and J. Hey. 1997. Gene flow and natural selection in the origin of *Drosophila pseudoobscura* and close relatives. Genetics 147:1091–1106.

Wasserman, M., and H. R. Koepfer. 1977. Character displacement for sexual isolation between *Drosophila mojavensis* and *Drosophila arizonensis*. Evolution 31:812–823.

Watterson, G. A. 1975. On the number of segregating sites in genetical models without recombination. Theoretical Population Biology 7: 256–276.

Won, Y.-J., and J. Hey. 2005. Divergence population genetics of chimpanzees. Molecular Biology and Evolution 22:297–307.

Won, Y.-J., Y. Wang, A. Sivasundar, J. Raincrow, and J. Hey. 2006. Nuclear gene variation and molecular dating of the cichlid species flock of Lake Malawi. Molecular Biology and Evolution 23:828–837.

Zouros, E. 1973. Genic differentiation associated with the early stages of speciation in the *mulleri* subgroup of *Drosophila*. Evolution 27: 601–621.

Associate Editor: James D. Fry
Editor: Jonathan B. Losos