Molecular Ecology (2004) **13**, 1365–1375 doi: 10.1111/j.1365-294X.2004.02169.x

Contrasting population genetic patterns and evolutionary histories among sympatric Sonoran Desert cactophilic *Drosophila*

L. A. HURTADO, T. EREZ, S. CASTREZANA and T. A. MARKOW
Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ 85721, USA

Abstract

We studied population genetic differentiation in the sympatric Sonoran Desert cactophilic flies *Drosophila pachea*, *D. mettleri* and *D. nigrospiracula* across their continental and peninsular ranges. These flies show marked differences in ecology and behaviour including dispersal distances and host cactus specialization. Examination of a fragment of the mitochondrial cytochrome oxidase subunit I gene (*mtCOI*) reveals that the Sea of Cortez has constituted an effective dispersal barrier for *D. pachea*, leading to significant genetic differentiation between the continental and peninsular ranges of this species. No genetic differentiation was detected, however, within its continental and peninsular ranges. In contrast, our *mtCOI*-based results for *D. mettleri* and *D. nigrospiracula* are consistent with a previous allozyme-based study that showed no significant genetic differentiation between continental and peninsular ranges of these two species. For *D. mettleri*, we also found that the insular population from Santa Catalina Island, California, is genetically differentiated with respect to continental and peninsular localities. We discuss how differences in the genetic structure patterns of *D. pachea*, *D. mettleri* and *D. nigrospiracula* may correspond to differences in their dispersal abilities, host preferences and behaviour.

Keywords: cactophilic, *Drosophila*, mitochondrial COI, phylogeography, Sonoran Desert, Sea of Cortez, vicariance

Received 29 November 2003; revision received 4 February 2004; accepted 4 February 2004

Introduction

For most organisms there are insufficient ecological and biogeographical data to make predictions about population genetic patterns. One exceptional group is constituted by the species of cactophilic *Drosophila* endemic to the Sonoran Desert of North America. These species provide an important model system for studies of host plant associations (Fellows & Heed 1972; Mangan 1982; Fogleman & Danielson 2001), adaptations to extreme environments (Toolson *et al*. 1990; Stratman & Markow 1998; Gibbs *et al*. 2003), population genetics (Zouros 1973; Pfeiler & Markow 2001; Markow *et al*. 2002), and speciation (Zouros & d'Entremont 1980). While the species of Sonoran Desert *Drosophila* overlap almost completely in their distribution, they are phylogenetically unrelated (Remsen & O'Grady 2002), invaded the desert independently of each other (Heed 1978, 1982), and experience very different spatial and temporal patterns of resource availability (Breitmeyer & Markow 1998), owing to their specialization on different species of host cacti (Heed 1978, 1982). In addition, they inhabit an area where a well-described vicariant event occurred, making them ideally suited to studies of the interactions between geological and ecological factors in producing current patterns of population genetic differentiation.

Plate boundary expansion between the North American and Pacific plates resulted in the separation of the peninsula of Baja California from the mainland. This vicariant event occurred around 3–6 million years ago (Gastil *et al*. 1983; Lonsdale 1989; Helenes & Carreño 1999) and divided desert communities that were continuously distributed. As a consequence of this event, the Sea of Cortez or Gulf of California appeared as a barrier that has since separated continental and peninsular biotas. This vicariant event has been shown to be clearly imprinted in the evolutionary history of several lineages, including mammals, birds, amphibians, reptiles and cacti (Upton & Murphy 1997;

Correspondence: Luis A. Hurtado. Fax. 520 626 3522; E-mail: lhurtado@u.arizona.edu

Riddle 2000; Riddle *et al*. 2000a,b; Nason *et al*. 2002). Very little is known, however, about the consequences of this event on the evolution of arthropods, the most abundant and diverse faunal group of the region.

The cactophilic Sonoran Desert *Drosophila* species *D. mettleri* and *D. nigrospiracula* breed generally in association with necrotic patches of the same host, saguaro (*Carnegiea gigantea*) in Arizona and Sonora, and cardón (*Pachycereus pringlei*) in Baja California, although *D. mettleri* occasionally breeds in other host cacti as well (Heed 1982). However, marked differences in resource partitioning are observed between these two species because *D. nigrospiracula* breeds on the tissue of rotten saguaro or cardón, while *D. mettleri* breeds in soil soaked with the necrotic juices of the cactus. Necrotic patches of saguaro or cardón are found at low densities and are very scattered in space and time (Breitmeyer & Markow 1998); thus, these flies are expected to be good dispersers. This is consistent with field observations for *D. nigrospiracula*, for which dispersal distances of up to 2 km per day have been recorded (Markow & Castrezana 2000). The role of the Sea of Cortez as a dispersal barrier was examined for *D. mettleri* and *D. nigrospiracula* using a limited number of allozyme loci (Markow *et al*. 2002). No appreciable population genetic differentiation was found for these two *Drosophila* species across localities sampled within their Arizona–Sonora mainland range (Pfeiler & Markow 2001), or between mainland and peninsular ranges (Markow *et al*. 2002). These allozyme data suggest that neither the host shift between the mainland and the peninsula (i.e. saguaro vs. cardón), nor the barrier erected by the Sea of Cortez, contributed to genetic differentiation for these flies between the two geographical areas of the Sonoran Desert.

Another Sonoran Desert cactophilic fly, *D. pachea*, shows a very contrasting ecology compared to *D. nigrospiracula* and *D mettleri*. This species breeds only in the rotten arms of senita cactus (mainly *Lophocereus schottii*) across its mainland and Baja peninsular ranges (Mangan 1982). Necrotic patches of senita are encountered far more frequently than those of saguaro and cardón, the hosts for the two *Drosophila* species mentioned above. As predicted by the differences in host spatial abundance, dispersal distances in *D. pachea* are much shorter, only about 300 m per day (Markow & Castrezana 2000). Therefore, the potential for genetic differentiation with geographical distance should be higher in *D. pachea* compared to the other two species. However, genetic differentiation for this species has only been examined in its continental range and the results differ with the type of genetic marker examined. Allozyme studies showed no genetic structure across its continental range (Rockwood-Sluss *et al*. 1973; Pfeiler & Markow 2001). In contrast, a study based on an inversion polymorphism revealed a cline in the frequency of two chromosomal arrangements, suggesting local adaptation among continental populations (Ward *et al*. 1974). Because genetic comparisons between continental and peninsular populations of *D. pachea* have yet to be undertaken, it remains unknown whether the Sea of Cortez has constituted an effective historical dispersal barrier for this species and whether some local genetic structure is present within its peninsular range. Given the more limited dispersal ability of *D. pachea*, this species would be expected to exhibit greater differentiation across the Sea of Cortez and perhaps some local genetic differentiation within peninsular and continental ranges, compared to *D. nigrospiracula* and *D. mettleri*.

To address the above hypothesis, the genetic variation at the mitochondrial cytochrome oxidase subunit I gene (*mtCOI*) was examined across the entire range of *D. pachea*, and the following questions were asked: (i) Has the presence of the Sea of Cortez constituted a barrier for the dispersal of *D. pachea* between continental and peninsular ranges? (ii) Is genetic structure detected within peninsular or continental ranges of this species? Furthermore, *mtCOI* variation among localities was examined across the range of *D. mettleri* and *D. nigrospiracula*, and it was investigated whether this mitochondrial marker supports the conclusions of earlier allozyme work (Markow *et al*. 2002) that the Sea of Cortez is not an effective barrier for gene flow of these two species.

Materials and methods

Localities sampled

Genetic variation was examined in three cactophilic *Drosophila* found in the Sonoran Desert: *D. pachea*, *D. mettleri* and *D. nigrospiracula*. For *D. pachea* 203 individuals were sampled from 11 localities (four continental and seven peninsular) representing the entire distribution range of this species (Fig. 1, Table 1). For *D. mettleri* 117 individuals were sampled from eight localities (three continental, four peninsular, and Santa Catalina Island, California). For *D. nigrospiracula* 94 individuals were sampled from seven localities (three continental and four peninsular). The continental localities sampled for *D. mettleri* and *D. nigrospiracula* are separated by up to ~600 km. The peninsular localities sampled for these two species represent most of their range across the Baja peninsula.

DNA extraction, PCR amplification, and sequencing

The manufacturer's protocol for the DNEasy kit (Qiagen, Inc., Valencia, CA) was followed to extract total DNA from each individual. Whole flies were ground for DNA extraction. A 710-base-pair fragment of the *mtCOI* gene was amplified from all the individuals analysed. DNA primers and polymerase chain reaction (PCR) conditions are described in Folmer *et al*. (1994). The PCR products were sequenced in both directions using an ABI 3700 analyser. Sequences were

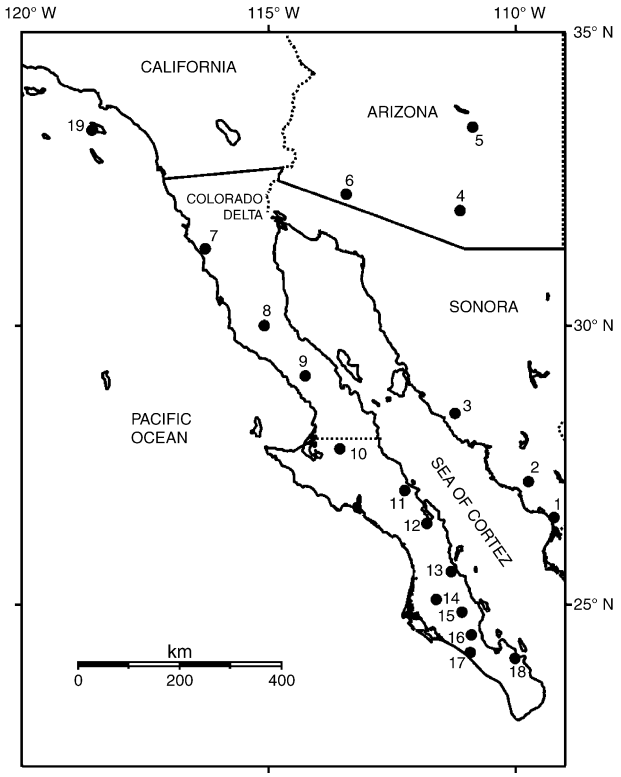

Fig. 1 Study area: mainland populations: 1 = Agiabampo (AG); 2 = Navojoa (NAV); 3 = Guaymas (GUY); 4 = Tucson (TC); 5 = Sierra Ancha (SA); 6 = Organ Pipe Cactus National Monument (OPN). Baja peninsula populations: 7 = Rancho Costa Rica (RCR); 8 = Cataviña (CAT); 9 = Chapala (CHA); 10 = Vizcaino Desert (VIZ); 11 = Vírgenes (VIR); 12 = Bahía Concepción (BACO); 13 = Nopolo (NOP); 14 = Ciudad Constitución (CICO); 15 = El Cién (CIEN); 16 = Puente Tevalle (PT); 17 = Punta Conejo (PUC); 18 = Ensenada de los Muertos (ENM). Insular population: 19 = Santa Catalina Island (SCI).

proofread and aligned with SEQUENCHER vs. 4.1 (Gene Codes Corp.). Assembled sequences were truncated to a fragment that contained only clear and readable nucleotides (661–663 base pairs, depending on the species). The *mtCOI* DNA sequences obtained for the three species were translated to amino acid sequences. No terminal codons or indels were found among the sequences, thus, we are confident to have amplified and sequenced *mtCOI* and not a pseudogene or another DNA fragment. Furthermore, a blast search of *mtCOI* sequences of the three species in the GenBank database revealed that they are most similar to *mtCOI* sequences reported for several species of *Drosophila*.

Population genetics and phylogenetic analyses

For each species the following statistics were calculated: the *mtCOI* haplotype diversity, denoted as *h* (equation 8.6 in Nei 1987); mean number of pairwise differences, denoted as π_1 (Tajima 1983); and nucleotide diversity, denoted as π_2 (equation 10.6 in Nei 1987); as implemented in ARLEQUIN 2000 (Schneider *et al*. 2000). Maximum parsimony and neighbour-joining analyses were performed using PAUP* vs. 4.0b8 (Swofford 1998). Maximum parsimony analyses were performed by heuristic searches with 50 stepwise random additions and tree bisection-reconnection (TBR) branch swapping. Bootstrap support values were calculated from a 50% majority-rule consensus tree based on 1000 bootstrap replicates. The maximum parsimony analyses assumed equally weighted characters and equally weighted substitution types. To correct for mutational saturation, the neighbour-joining analyses assumed Kimura two-parameter distances (K2P, Kimura 1980), although at the within-species level this is unlikely to present a problem. In addition, statistical parsimony haplotype networks were constructed (Templeton *et al*. 1992) using the TCS program (Clement *et al*. 2000). Only the statistical parsimony haplotype networks are reported here because they may provide better representations of gene genealogies at the population level (Templeton *et al*. 1992).

ARLEQUIN 2000 was also used to estimate genetic differentiation between populations (F_{ST}), the corresponding number of migrating individuals (*Nm*), and to conduct exact tests of population differentiation (Raymond & Rousset 1995). Values of F_{ST} and *Nm* were calculated following Hudson *et al*. (1992). *Nm*, or effective number of migrants, is the historical average number of individual migrants contributing to a population's gene pool each generation. It represents the number of migrating individuals necessary to produce the observed F_{ST} value in an island model population at equilibrium for gene flow and genetic drift (Wright 1951). For haploid data, such as mitochondrial DNA, *Nm* is calculated as $[(1/F_{ST}) - 1]/2$. Analysis of molecular variance (AMOVA, Excoffier *et al*. 1992) as implemented in ARLEQUIN 2000 was also conducted. AMOVA estimates the amount of genetic variation attributable to genetic differentiation among groups (Φ_{CT}), among localities within groups (Φ_{SC}), and among localities relative to the total sample (Φ_{ST}). These Φ-values are analogous to conventional *F*-values. For AMOVA of each species, localities were placed in two groups: continental vs. peninsular. A significant Φ_{CT} value (genetic differentiation between groups relative to the total) would suggest that the Sea of Cortez has constituted an effective dispersal barrier for continental and peninsular exchange of individuals.

Demographic analyses

It was examined whether the populations studied here were in mutation–drift equilibrium and whether they fitted a population-size stationary model or a range expansion model. To examine for evidence of range expansions in the populations of these species Tajima's *D* (Tajima 1989) and Fu's F_S (Fu 1996) tests were performed. Although these

Table 1 Localities and number of individuals collected

Localities	*D. pachea*	*D. mettleri*	*D. nigrospiracula*
Continental localities			
Agiabampo (AG)	9		
Navojoa (NAV)	16		
Guaymas (GUY)	16	20	19
Tucson (TC)		20	12
Organ Pipe Cactus National Monument (OPN)	20		
Sierra Ancha (SA)		5	10
Peninsular localities			
Rancho Costa Rica (RCR)	12		
Cataviña (CAT)	16		
Chapala (CHA)		17	
Vizcaino Desert (VIZ)	23		14
Vírgenes (VIR)	24		
Bahía Concepción (BACO)		20	17
Nopolo (NOP)			6
Ciudad Constitución (CICO)	19		
El Cién (CIEN)		13	
Puente Tevalle (PT)			16
Punta Conejo (PUC)	24	6	
Ensenada de los Muertos (ENM)	24		
Insular localities			
Santa Catalina Island, California (SCI)		16	
Total	203	117	94

tests were often used as tests of selective neutrality, they are also very sensitive to departures from population equilibrium. A significant Tajima's *D*-value may be the result of factors other than selective effects, such as population expansions, bottlenecks, or heterogeneity of mutation rates (Aris-Brosou & Excoffier 1996; Tajima 1996). Similarly, Fu's F_S test is very sensitive to population demographic expansion, which generally leads to large negative values (Fu 1997; Ray *et al*. 2003). ARLEQUIN 2000 was used to conduct these tests and to calculate the corresponding *P*-values. In addition, mismatch distributions, or the distributions of the number of pairwise differences among all DNA sequences, (Harpending 1994; Schneider & Excoffier 1999) were examined for each species. Ragged and erratic distributions are expected for stationary populations, while smooth distributions are expected for populations that experienced range expansions (Harpending 1994). We used ARLEQUIN 2000 to estimate parameters related to a population growth expansion, such as expansion time τ (where $\tau = 2ut$, where u is the mutation rate and t is the number of generations since the expansion), $\theta_0 = 2\ \mu N_0$, and $\theta_1 = 2\ \mu N_1$ (where N_0 and N_1 are the population sizes before and after the expansion). A null model of population range expansion is assumed if $\tau > 0$ and $\theta_1 > \theta_0$, or a null model of population stationarity if $\tau = 0$ or $\theta_1 = \theta_0$. The validity of the estimated demographic model is tested by obtaining the distribution of a test statistic, the sum of squared differences (SSD) between the observed and an estimated mismatch distribution obtained by a bootstrap approach. The *P*-value of the SSD statistic is computed as the proportion of simulated cases that show an SSD value larger than the original (Schneider & Excoffier 1999). A significant SSD value (*P*-values < 0.05) is taken as evidence for departure from a model of population range expansion (when $\tau > 0$ and $\theta_1 > \theta_0$), or from a model of static population size (when $\tau = 0$ or $\theta_1 = \theta_0$). The demographic tests were conducted separately for each locality and for pooled localities that were determined to be genetically homogeneous by the exact tests of population differentiation.

Results

Genetic variation and phylogenetic analyses

Large differences in nucleotide and haplotype genetic variation were observed among the three *Drosophila* species for the *mtCOI* gene fragment studied (Table 2). Genetic variation in *D. pachea* was the largest observed (π_1 = mean number of pairwise differences = 4.77), considerably lower in *D. mettleri* ($\pi_1 = 1.37$) and extremely low in *D. nigrospiracula* ($\pi_1 = 0.19$).

For *D. pachea*, there were 79 different haplotypes out of the 203 individuals examined (GenBank Accession numbers AY533710–AY533788). Of these, 23 were found exclusively among continental individuals, 49 exclusively among

Table 2 Haplotype diversity (h), mean number of pairwise differences (π_1), and nucleotide diversity (π_2) for *Drosophila pachea*, *D. mettleri* and *D. nigrospiracula*

Species	N	Haplotypes observed	Base pairs	Polymorphic sites	h	π_1	π_2
D. pachea	203	79	661	61	0.96 ± 0.00	4.77 ± 2.34	0.007 ± 0.004
D. mettleri	117	24	662	25	0.72 ± 0.04	1.37 ± 0.85	0.002 ± 0.001
D. nigrospiracula	94	10	663	9	0.18 ± 0.05	0.19 ± 0.24	0.0002 ± 0.000

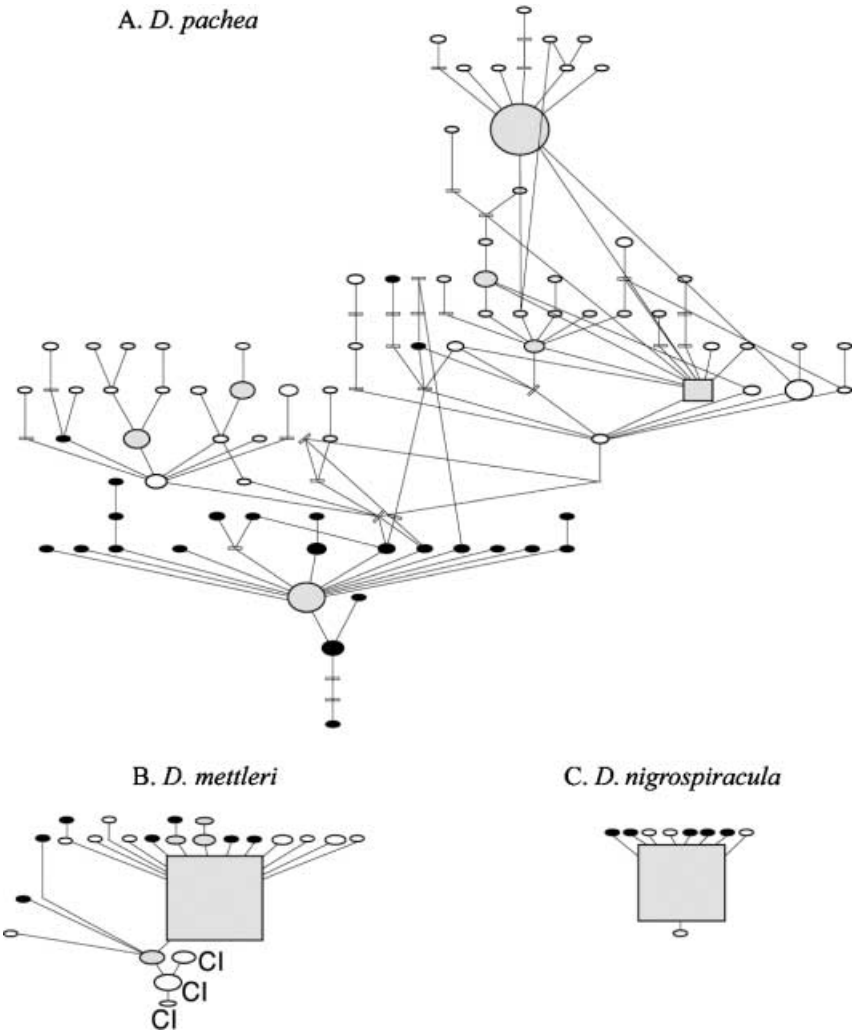

Fig. 2 Statistical parsimony haplotype networks of *Drosophila pachea*, *D. mettleri* and *D. nigrospiracula* haplotypes. Squares represent ancestral haplotypes as determined by the TCS program. Black ovals represent haplotypes observed only in continental individuals; white ovals represent haplotypes observed only in peninsular individuals; grey ovals and squares represent haplotypes observed in both continental and peninsular individuals; bars represent intermediate missing haplotypes. The three haplotypes of *D. mettleri* exclusive of Santa Catalina Island are indicated with the letters CI.

peninsular individuals, and only seven were shared between individuals of the two populations. More than 1000 most parsimonious trees were found for this species (not shown). The statistical parsimony haplotype network (Fig. 2A) and an unrooted neighbour-joining tree (not shown) for the 79 haplotypes suggest some evidence of continental vs. peninsular clade segregation. One region of the statistical parsimony haplotype network consists mainly of haplotypes exclusive to the continent (20 continental and one that was observed in 15 continental individuals and one peninsular individual). The rest of the network consists mainly of haplotypes exclusive to the peninsula (three continental, 49 peninsular, and six shared). This apparent continental/peninsular clade segregation may be suggestive of an ongoing lineage sorting process as a consequence of genetic differentiation between continental and peninsular ranges. However, neither the neighbour-joining bootstrap analysis nor the consensus tree of the most parsimonious trees support separation of these clades (not shown).

For *D. mettleri*, 24 different haplotypes were found out of

Table 3 Values of F_{ST} (above the diagonal) and *Nm* (below diagonal) for *Drosophila pachea*

	AG	NAV	GUY	OPN	RCR	CAT	VIZ	VIR	CICO	PUC	ENM
AG		0.00	0.00	0.00	***0.39***	***0.38***	***0.27***	***0.35***	***0.37***	***0.31***	***0.40***
NAV	*		0.02	0.00	***0.38***	***0.36***	***0.24***	***0.33***	***0.35***	***0.29***	***0.38***
GUY	*	23.8		0.00	***0.48***	***0.47***	***0.36***	***0.43***	***0.44***	***0.40***	***0.47***
OPN	*	*	*		***0.35***	***0.34***	***0.25***	***0.31***	***0.33***	***0.28***	***0.36***
RCR	*0.8*	*0.8*	*0.5*	*0.9*		**0.04**	0.09	0.04	0.01	0.03	0.02
CAT	*0.8*	*0.9*	*0.5*	*1.0*	12.9		0.02	0.00	0.00	0.00	0.00
VIZ	*1.4*	*1.6*	*0.9*	*1.5*	4.9	19.5		0.02	0.04	0.00	0.05
VIR	*0.9*	*1.0*	*0.7*	*1.1*	13.4	*	25.5		0.00	0.00	0.00
CICO	*0.9*	*0.9*	*0.6*	*1.0*	88.1	*	11.1	*		0.00	0.00
PUC	*1.1*	*1.2*	*0.8*	*1.3*	18.0	*	*	*	*		0.00
ENM	*0.8*	*0.8*	*0.6*	*0.9*	19.7	*	10.0	*	*	*	

Values in italic indicate continental vs. peninsular pairwise comparisons. Significant pairwise ($P < 0.05$) differences are indicated in bold.
* = N_m is undefined and approaches panmixia.
Mainland populations: AG = Agiabampo; NAV = Navojoa; GUY = Guaymas; OPN = Organ Pipe Cactus National Monument, Arizona.
Baja peninsular populations: RCR = Rancho Costa Rica; CAT = Cataviña; VIZ = Vizcaino Desert; VIR = Vírgenes; CICO = Ciudad Constitución; PUC = Punta Conejo; ENM = Ensenada de los Muertos.

Table 4 AMOVA analyses for *Drosophila pachea*, *D. mettleri* and *D. nigrospiracula* grouped into peninsular vs. continental groups

Source of variation	*D. pachea*	*D. mettleri*	*D. nigrospiracula*
Among groups	34.6%	–1.3%	0.9%
Among populations within groups	0.2%	5.6%	–0.8
Within populations	65.2%	95.7%	99.9%
Φ_{CT}	**0.345**	0.000	0.008
Φ_{SC}	0.003	0.055	0.000
Φ_{ST}	**0.348**	0.043	0.000

Φ_{CT} represents the amount of genetic variation attributable to genetic differentiation among groups; Φ_{SC} represents the amount of genetic variation attributable to genetic differentiation among localities within groups and Φ_{ST} represents the amount of genetic variation attributable to genetic differentiation among localities relative to the total sample.
Values in bold indicate significance at $P < 0.05$.

the 117 individuals examined (GenBank Accession numbers AY533789–AY533812). The most common haplotype was observed in 52.1% of the individuals sampled (29 continental and 32 peninsular). The same topology of the 24 different haplotypes was obtained with statistical parsimony (Fig. 2B), maximum parsimony, and neighbour joining (not shown). No evidence for a separation between continental and peninsular populations was observed from these topologies. However, a clear separation was observed between Santa Catalina Island haplotypes and continental/peninsular haplotypes. In Santa Catalina Island three different haplotypes were identified, none of which were observed in individuals from the continent or peninsula.

For *D. nigrospiracula*, only 10 haplotypes were found among the 94 individuals examined (GenBank Accession numbers AY533813–AY533822). The majority of the individuals (85) harboured the most common haplotype, and the other nine haplotypes, each differing from the common haplotype by only one nucleotide substitution, were observed in single individuals (five continental and four peninsular). The extremely low genetic variation found in this species was clearly depicted in the star-like topology of the unrooted neighbour-joining and maximum parsimony trees (not shown) and the statistical parsimony haplotype network (Fig. 2C).

Genetic structure

Genetic differentiation between continental and peninsular localities was clearly observed in *D. pachea*. All pairwise tests of genetic differentiation comparing continental vs. peninsular localities were significant (Table 3). In contrast, there was no evidence for genetic structure within continental or peninsular ranges. Except for one peninsular pairwise comparison, none of the tests of genetic differentiation among continental or among peninsular localities were significant. The AMOVA test that clustered *D. pachea* localities into continental vs. peninsular groups resulted in a significant Φ_{CT} value (i.e. genetic differentiation between groups; Table 4). Differences between continental and peninsular

Table 5 Values of F_{ST} (above the diagonal) and *Nm* (below) for *Drosophila mettleri*

	SA	TC	GUY	CHA	BACO	CIEN	PUC	SCI
SA		0.00	0.04	0.00	**0.06**	0.07	**0.15**	**0.76**
TC	*		0.02	0.00	**0.05**	0.00	**0.13**	**0.71**
GUY	12.8	27.7		0.01	**0.10**	0.05	0.00	**0.67**
CHA	*	*	43.9		0.03	0.00	**0.11**	**0.69**
BACO	7.5	8.7	4.6	17.3		0.05	**0.23**	**0.74**
CIEN	6.7	*	10.3	*	10.2		**0.29**	**0.81**
PUC	2.8	3.4	211.3	4.1	1.7	1.2		**0.63**
SCI	0.1	0.2	0.2	0.2	0.2	0.1	0.3	

Significant pairwise ($P < 0.05$) differences are indicated in bold. * = *Nm* is undefined and approaches panmixia.
Mainland populations: SA = Sierra Ancha; TC = Tucson; GUY = Guaymas. Baja peninsular populations: CHA = Chapala; BACO = Bahía Concepción; CIEN = El Cién; PUC = Punta Conejo. Insular population: SCI = Santa Catalina Island

Table 6 Demographic tests to detect range expansion among populations of Sonoran Desert *Drosophila*

Species/locality	*D*	*P*	F_S	*P*	τ	θ_0	θ_1	SSD	*P*
D. pachea									
Continental	–1.70	**0.02**	–11.63	**0.00**	4.43	0.42	8.44	0.01	**0.66**
Peninsular	–1.60	**0.02**	–12.95	**0.00**	5.30	0.43	10.45	0.01	**0.55**
D. mettleri									
Continental	–2.04	**0.00**	–8.92	**0.00**	1.45	0.00	1.93	0.00	**0.95**
Chapala	–2.20	**0.00**	–4.81	**0.00**	1.57	0.00	3.76	0.00	**0.82**
Bahía Concepción	–0.45	0.63	–0.32	0.36	1.11	0.00	1437.19	0.03	0.02
El cién	–1.47	0.06	–1.40	**0.02**	0.83	0.00	0.44	0.01	**0.38**
Punta Conejo	–0.18	0.41	–1.35	0.06	1.75	0.00	1891.25	0.04	**0.26**
Santa Catalina Island	0.09	0.62	0.04	0.39	0.82	0.00	1404.37	0.04	0.01
D. nigrospiracula									
Continental/peninsular	–2.26	**0.00**	–12.19	**0.00**	3.00	0.03	0.27	0.00	**0.36**

D = Tajima's *D* test; F_S = Fu's F_S tests; SSD = Sum of squared deviation mismatch method. Corresponding *P*-values (*P*) for each method are shown in the *P* column to the right of the column with the results for each particular method. *P*-values in bold indicate the test is consistent with a model of range expansion.

groups explained a large portion of the total genetic variance (34.6%). Differences among localities within each group (Φ_{SC}) were not significant, however, and explained only 0.23% of the total genetic variance. On the other hand, differences among localities relative to the total sample (Φ_{ST}) explained most of the genetic variance (65.2%).

For *D. mettleri* and *D. nigrospiracula*, no significant genetic differentiation was found between continental and peninsular ranges (Tables 4 and 5). The *D. mettleri* population from Santa Catalina Island, however, was significantly different from all the other *D. mettleri* localities examined (Table 5). Furthermore, the peninsular localities of Bahía Concepción and Punta Conejo were significantly different from most of the other localities. A complete lack of genetic structure was found among the localities examined for *D. nigrospiracula* (average $F_{ST} = 0.0$); however, the extremely low levels of *mtCOI* genetic diversity observed in this species may limit conclusions on gene flow based on this marker.

Demographic analyses

Demographic analyses (Table 6) showed evidence of range expansions in populations of all three species. Mismatch tests were consistent with a range expansion model in all cases in which significant *P*-values for Tajima's *D* and Fu's F_S tests were obtained. For *D. pachea*, there were significant Tajima's *D* and Fu's F_S *P*-values in the tests (i) pooling continental localities, and (ii) pooling peninsular localities. In both cases, the demographic parameters estimated by mismatch analyses corresponded to a null model of population range expansion ($\tau > 0$ and $\theta_1 > \theta_0$) that could not

be rejected (i.e. the SSD *P*-values were > 0.05). Likewise, mismatch analyses conducted for each locality of *D. pachea* separately (data not shown) resulted in a null model of population range expansion that was not rejected.

For *D. mettleri*, significant Tajima's *D* and Fu's F_S *P*-values were obtained in the tests where continental localities were pooled, and the mismatch analysis resulted in a null model of population range expansion that was not rejected. Mismatch analyses were also consistent with a model of population range expansion in tests where each continental locality was analysed separately (data not shown). Peninsular localities of this species were not pooled for demographic analyses since exact tests of population differentiation detected evidence for population structure in the peninsular region. Mismatch analyses for each peninsular locality of *D. mettleri* resulted in a null model of range expansion that was not rejected in most cases. In the case of the Santa Catalina Island *D. mettleri* population, the null model of range expansion was rejected and the *P*-values of Tajima's *D* and Fu's F_S were not significant. For *D. nigrospiracula*, significant Tajima's *D* and Fu's F_S *P*-values were obtained in the test that pooled all localities and the mismatch analysis resulted in a model of population range expansion that was not rejected.

Discussion

Despite occupying similar niches and having overlapping distributions, the three *Drosophila* species examined in this study exhibit very different patterns of genetic differentiation and evolutionary histories. On a broad scale, our data are consistent with predictions based upon resource distributions and dispersal differences observed among species, i.e. that *D. pachea*, because of its more limited dispersal relative to the other two species, would exhibit greater population structure. Significant genetic differentiation between *D. pachea* populations from the peninsula and the continent suggests that the Sea of Cortez has constituted an effective dispersal barrier for this species over evolutionary time. However, there was no evidence of any genetic structure either within the continental or within the peninsular ranges of this species. Our results for *D. mettleri* and *D. nigrospiracula* confirm those of earlier allozyme studies, concluding that the Sea of Cortez has not represented an effective dispersal barrier for gene flow, and suggesting further, that the host shift between the two sides of the Sea of Cortez (i.e. saguaro vs. cardón) has no impact on genetic differentiation in these two species.

Differences in dispersal potential and longevity may explain why the Sea of Cortez constitutes a barrier for gene flow for *D. pachea* but not for the other two species. Female *D. pachea* live only half as long as female *D. mettleri* and only one-third as long as *D. nigrospiracula* females (Markow and Marron unpublished data). A combination of a high dispersal and increased longevity of both sexes may allow *D. mettleri* and *D. nigrospiracula* to cross the Gulf of California more easily than *D. pachea*. The average distance between the Baja peninsula and the continent across the entire length (1100 km) of the Gulf of California is ~120 km. Dispersal across the Gulf may be facilitated by a chain of islands present in the mid-portion of the Gulf of California, the Midriff Islands, with oceanic gaps between ~2 and 16 km long. Alternatively, *D. pachea* may be sensitive to environmental clues associated with the sea and may avoid moving across the ocean. However, this species exhibits considerable gene flow across large distances within continental or peninsular ranges. Long-distance gene flow among the three species may be enhanced by the fact that adults of these species can feed on a range of cacti when dispersing to their own specific breeding sites (Fellows & Heed 1972; Markow *et al*. 2000), and their females exhibit frequent female re-mating (Markow 1982).

A previous study (Ward *et al*. 1974) showed a geographical gradient in the continental range of *D. pachea* for an inversion polymorphism in chromosome 7 (7+ and 7*A*) that exhibited a correlation with climate (precipitation and temperature). They claimed that individuals with inversions 7+, more common towards the southern range and rare at northern localities, are wet-and-warm adapted, but not hot adapted, and that the autumn should be the time of greatest selection pressure from seasonal weather (Ward *et al*. 1974). Given that these chromosomal polymorphisms are under apparent selection, they should not be used to infer historical gene flow (i.e. population genetic structure). Concordance between our mitochondrial results and previous allozyme-based results (Rockwood-Sluss *et al*. 1973; Pfeiler & Markow 2001), and the likely neutral or nearly neutral nature of our mitochondrial marker, allow us to assume that it portrays better the evolutionary history of gene flow for *D. pachea*.

Despite the obligate association of *D. pachea* with its senita host cactus (Heed & Kircher 1965), and thus the complete overlap in their distributions, patterns of population genetic differentiation of this fly differ markedly from those of the plant. For senita, population genetic differentiation exists not only between continental and peninsular ranges, but within the Baja Peninsula and within Sonora as well (Nason *et al*. 2002). This fly–host discrepancy probably only reflects the dispersal differences between the two taxa. Furthermore, there is no evidence that *D. pachea* has adapted to host differentiation at the species level. The senita species *L. gatesii* is restricted to a small area of the Magdalena plain west of La Paz, in southern Baja California, and the *D. pachea* sample collected from this area (i.e. Punta Conejo) did not differ significantly from any other peninsular samples taken from *L. schottii* var. *australis* (in the Cape region) and *L. schottii* var. *schottii* (in the rest of the peninsula).

Localized, though significant, genetic differentiation among some peninsular populations was reported previously using allozymes for *D. mettleri* and *D. nigrospiracula* (Markow *et al*. 2002). Using mtDNA this was also observed in two different peninsular populations of *D. mettleri*, suggesting that additional studies, at a finer geographic scale, may reveal additional population structure. Similarly, our study also identified the Santa Catalina Island population of *D. mettleri* as highly genetically differentiated from all continental and peninsular localities compared. This species uses a different host (*Opuntia*) on Santa Catalina Island compared to their Sonoran Desert conspecifics that use mainly cardón and saguaro, and evidence of incipient postzygotic isolation has been observed between individuals from this and other localities (Markow *et al*. 2002).

Striking differences among the three cactophilic fly species examined here were also observed with regard to intraspecific genetic diversity. Accordingly, *D. pachea* exhibits extremely high genetic diversity when compared to the other two species. High genetic diversity of *D. pachea* is consistent with a scenario of continental/peninsular vicariance, followed by local isolation at several times during glaciations, and subsequent mixing of isolated populations as climate conditions became warmer, as suggested for its senita host (Nason *et al*. 2002). In the case of *D. mettleri* and *D. nigrospiracula*, high levels of historical gene flow may have prevented substantial local genetic differentiation and could account for some of the reduced *mtCOI* genetic diversity observed, compared to *D. pachea*. However, *mtCOI* genetic diversity in *D. nigrospiracula* is extremely low, which we hesitate to explain only in terms of extensive historical gene flow. In the allozyme studies (Pfeiler & Markow 2001; Markow *et al*. 2002), *D. nigrospiracula* was reported to have a higher average observed heterozygosity ($H_O = 0.14$) than *D. mettleri* ($H_O = 0.10$). One scenario that may be consistent with these observations is that *D. nigrospiracula* experienced a recent bottleneck and a subsequent rapid range expansion in which mitochondrial genetic variability was lost after the bottleneck, but nuclear variation was retained. An alternative explanation is that a mitochondrial selective sweep reduced genetic variability for the *mtCOI* gene in *D. nigrospiracula*.

In conclusion, we found that emergence of the Sea of Cortez, as a consequence of the separation of Baja peninsula from the continent, is associated with significant genetic differentiation between continental and peninsular populations of *D. pachea*. The same pattern has been observed for senita, the host cactus of *D. pachea* (Nason *et al*. 2002), the Sonoran cactophilic fly *D. mojavensis* (Hocutt 2000; Zouros 1973), and several mammalian, avian, amphibian and reptilian taxa (Riddle *et al*. 2000b). However, there was no evidence for population genetic structure within either continental or peninsular regions of *D. pachea*, even though different species of its senita cactus host are present in the peninsular region. In contrast, neither the presence of the Sea of Cortez nor the host shift between continental and peninsular ranges (i.e. saguaro vs. cardón) appear to have reduced gene flow of *D. mettleri* and *D. nigrospiracula* to produce significant differentiation between continental and peninsular populations of these species. Given the contrasting genetic and evolutionary patterns found among the *Drosophila* species studied here, we are examining other arthropods that specialize on necrotic cacti but have longer life cycles and more limited dispersal, which may reveal higher levels of population genetic differentiation and a signature of vicariant or ecological events that have occurred in this dynamic region.

Finally, interspecific differences in genetic variability and degree of genetic differentiation provide a *priori* predictions concerning the degrees to which these three species will differ in their potential to undergo speciation. *Drosophila pachea*, for example, owing to its significant genetic differentiation across the Sea of Cortez, should be more closely examined for evidence of early stages of reproductive isolation between continental and peninsular populations. While an earlier study of sexual isolation between peninsular and mainland populations of this species (Markow *et al*. 1983) revealed no significant sexual isolation, postmating isolating mechanisms have yet to be examined for *D. pachea*. The genetic differentiation observed between *D. mettleri* from Santa Catalina Island and the rest of the range suggests that evidence of newly evolving isolating mechanisms may be found between these populations as well. *Drosophila nigrospiracula*, on the other hand, is more likely than the other two species to remain panmictic and undifferentiated.

Acknowledgements

We thank Mariana Mateos, Edward Pfeiler, Jason Wilder, William Heed and four anonymous reviewers for providing helpful advice on this study and the manuscript. This work was supported by NSF grants (DEB 95–10645 and DEB 00–75312) and the University of Arizona Center for Insect Science. We also thank the Santa Catalina Island Conservancy, especially Angela Aarhus and Frank Starkey, for collecting permits, logistical support and help collecting specimens in Santa Catalina Island.

References

Aris-Brosou S, Excoffier L (1996) The impact of population expansion and mutation rate heterogeneity on DNA sequence polymorphism. *Molecular Biology and Evolution*, **13**, 494–504.

Breitmeyer CM, Markow TA (1998) Resource availability and population size in cactophilic *Drosophila*. *Functional Ecology*, **12**, 14–21.

Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9**, 1657–1659.

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, **131**, 479–491.

Fellows DP, Heed WM (1972) Factors affecting host plant selection in desert-adapted cactiphilic *Drosophila*. *Ecology*, **53**, 850–858.

Fogleman JC, Danielson PB (2001) Chemical interactions in the cactus–microorganism–*Drosophila* model system of the Sonoran Desert. *American Zoologist*, **41**, 877–889.

Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R (1994) DNA primers for amplification of mitochondrial cytochrome *c* oxidase subunit I from metazoan invertebrates. *Molecular Marine Biology and Biotechnology*, **3**, 294–299.

Fu YX (1996) New statistical tests of neutrality for DNA samples from a population. *Genetics*, **143**, 557–570.

Fu YX (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, **147**, 915–925.

Gastil G, Minch J, Phillips RP (1983) The geology and ages of the islands. In: *Island Biogeography in the Sea of Cortez* (eds Case TJ, Cody ML), pp. 13–25. University of California Press, Berkeley, CA.

Gibbs AG, Perkins MC, Markow TA (2003) No place to hide: microclimates of sonoran desert *Drosophila*. *Journal of Thermal Biology*, **28**, 353–362.

Harpending HC (1994) Signature of ancient population-growth in a low-resolution mitochondrial-DNA mismatch distribution. *Human Biology*, **66**, 591–600.

Heed WB (1978) Ecology and genetics of Sonoran desert *Drosophila*. In: *Ecological Genetics: the Interface* (ed. Brussard PF), pp. 109–126. Springer, New York.

Heed WB (1982) The origin of *Drosophila* in the Sonoran Desert. In: *Ecological Genetics and Evolution: the Cactus–Yeast–Drosophila Model System* (eds Barker JSF, Starmer WT), pp. 65–80. Academic Press, Sydney.

Heed WB, Kircher HW (1965) A unique sterol in the ecology and nutrition of *Drosophila pachea*. *Science*, **149**, 758–761.

Helenes J, Carreño AL (1999) Neogene sedimentary evolution of Baja California in relation to regional tectonics. *Journal of South American Earth Sciences*, **12**, 589–605.

Hocutt GD (2000) Reinforcement of premating barriers to reproduction between *Drosophila arizonae* and *Drosophila mojavensis*. PhD Arizona State University, Tempe, AZ.

Hudson RR, Boos DD, Kaplan NL (1992) A statistical test for detecting geographic subdivision. *Molecular Biology and Evolution*, **9**, 138–151.

Kimura M (1980) A simple method for estimating evolutionary rates of base substitution through comparative studies of nucleotide sequences. *Journal of Molecular Evolution*, **16**, 111–120.

Lonsdale P (1989) Geology and tectonic history of the Gulf of California. In: *The Geology of North America*, **Vol. N**. The Eastern Pacific Ocean and Hawaii. (eds Winterer EL, Hussong DM, Decker RW), pp. 499–521. The Geological Society of America, Boulder, CO

Mangan RL (1982) Adaptation to competition in cactus breeding *Drosophila*. In: *Ecological Genetics and Evolution: the Cactus–Yeast–Drosophila Model System* (eds Barker JSF, Starmer WT). Academic Press, Sydney.

Markow TA (1982) Mating systems of cactophilic Drosophila. In: *Ecological Genetics and Evolution: the Cactus–Yeast–Drosophila Model System* (eds Barker JSF, Starmer WT), pp. 273–287. Academic Press, New York.

Markow TA, Castrezana S (2000) Dispersal in cactophilic *Drosophila*. *Oikos*, **89**, 378–386.

Markow TA, Fogleman JC, Heed WB (1983) Reproductive isolation in Sonoran Desert *Drosophila*. *Evolution*, **37**, 649–652.

Markow TA, Anwar S, Pfeiler E (2000) Stable isotope ratios of carbon and nitrogen in natural populations of *Drosophila* species and their hosts. *Functional Ecology*, **14**, 261–266.

Markow TA, Castrezana S, Pfeiler E (2002) Flies across the water: genetic differentiation and reproductive isolation in allopatric desert *Drosophila*. *Evolution*, **56**, 546–552.

Nason JD, Hamrick JL, Fleming TH (2002) Historical vicariance and postglacial colonization effects on the evolution of genetic structure in *Lophocereus*, a Sonoran Desert columnar cactus. *Evolution*, **11**, 2214–2226.

Nei M (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York.

Pfeiler E, Markow TA (2001) Ecology and population genetics of Sonoran Desert *Drosophila*. *Molecular Ecology*, **10**, 1787–1791.

Ray N, Currat M, Excoffier L (2003) Intra-deme molecular diversity in spatially expanding populations. *Molecular Biology and Evolution*, **20**, 76–86.

Raymond M, Rousset F (1995) An exact test for population differentiation. *Evolution*, **49**, 1280–1283.

Remsen J, O'Grady P (2002) Phylogeny of the Drosophilinae (Diptera: Drosophilidae), with comments on combined analysis and character support. *Molecular Phylogenetics and Evolution*, **24**, 243–264.

Riddle BR (2000) Comparative phylogeography of baileys' pocket mouse (*Chaetodipus baileyi*) and the *Peromyscus eremicus* species group: historical vicariance of the Baja California peninsular desert. *Molecular Phylogenetics and Evolution*, **17**, 161–172.

Riddle BR, Hafner DJ, Alexander LF (2000a) Phylogeography and systematics of the *Peromyscus eremicus* species group and the historical biogeography of North American warm regional deserts. *Molecular Phylogenetics and Evolution*, **17**, 145–160.

Riddle BR, Hafner DJ, Alexander LF, Jaeger JR (2000b) Cryptic vicariance in the historical assembly of a Baja California peninsular desert biota. *Proceedings of the National Academy of Sciences of the USA*, **97**, 14438–14443.

Rockwood-Sluss ES, Johnston JS, Heed WB (1973) Allozyme genotype–environment relationships. I. Variation in natural populations of *Drosophila pachea*. *Genetics*, **73**, 135–146.

Schneider S, Excoffier L (1999) Estimation of past demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: application to human mitochondrial DNA. *Genetics*, **152**, 1079–1089.

Schneider S, Roessli D, Excoffier L (2000) *ARLEQUIN, a Software Package for Population Genetics Data Analysis, Version 2000*. Genetics and Biometry Laboratory. Department of Anthropology. University of Geneva, Geneva.

Stratman R, Markow TA (1998) Resistance to thermal stress in desert *Drosophila*. *Functional Ecology*, **12**, 965–970.

Swofford DL (1998) *PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods)*. Sinauer, Sunderland, MA.

Tajima F (1983) Evolutionary relationships of DNA sequences in finite populations. *Genetics*, **105**, 437–460.

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

Tajima F (1996) The amount of DNA polymorphism maintained in a finite population when the neutral mutation rate varies among sites. *Genetics*, **143**, 1457–1465.

Templeton AR, Crandall KA, Sing CF (1992) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonulease mapping and DNA-sequence data. III. Cladogram Estimation. *Genetics*, **132**, 619–633.

Toolson EC, Markow TA, Jackson LL, Howard RW (1990) Epicuticular hydrocarbon composition of wild and laboratory-reared *Drosophila mojavensis* Patterson and Crow (Diptera, Drosophilidae). *Annals of the Entomological Society of America*, **83**, 1165–1176.

Upton DE, Murphy RW (1997) Phylogeny of the side-blotched lizards (Phrynosomatidae: *Uta*) based on mtDNA sequences: support for midpeninsular seaway in Baja California. *Molecular Phylogenetics and Evolution*, **8**, 104–113.

Ward BL, Starmer WT, Russell JS, Heed WB (1974) The correlation of climate and host plant morphology with a geographic gradient of an inversion polymorphism in *Drosophila pachea*. *Evolution*, **28**, 565–575.

Wright S (1951) The genetical structure of populations. *Annals of Eugenics*, **15**, 323–354.

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

Zouros E, d'Entremont CJ (1980) Sexual isolation among populations of *Drosophila mojavensis* — response to pressure from a related species. *Evolution*, **34**, 421–430.

Luis A. Hurtado is a postdoctoral fellow at T. A. Markow's laboratory, interested in the connection between geological and ecological processes on the evolution of invertebrate lineages. Tamar Erez and Sergio Castrezana are doctoral students on the Ecology and Evolutionary Biology Program at the University of Arizona. Tamar is interested in population genetics and evolution and Sergio in insect biology, including adaptive radiation, sexual selection and systematics. Therese Markow studies ecology and population biology of desert-adapted *Drosophila*, with special interests in speciation and the evolution of mating systems.
