Genetica **118:** 51–58, 2003.

Variability on the dot chromosome in the *Drosophila simulans* clade

Mohamed A.F. Noor[1] & Richard M. Kliman[2]
[1]*Department of Biological Sciences, Life Sciences Bldg., Louisiana State University, Baton Rouge, LA 70803, USA (Phone: +1-225-578-8556; Fax: +1-225-578-2597; E-mail: mnoor@lsu.edu);* [2]*Department of Biological Sciences, Kean University, Union, NJ 07083, USA; Current address: Department of Biological Sciences, Cedar Crest College, Allentown, PA 18104, USA*

Received 19 July 2002 Accepted 23 September 2002

Key words: Drosophila mauritiana, *Drosophila sechellia*, *Drosophila simulans*, introgression, phylogeny

Abstract

A recent study suggested that recent nuclear gene introgression between *Drosophila simulans* and *D. mauritiana* may have obscured efforts to estimate the phylogeny of the species of the *D. simulans* clade, which includes these two species and *D. sechellia*. Here, we report sequence variation of an intron of the *eyeless* gene in this species group. This gene should introgress freely between these species because it is not linked to any known barriers to gene exchange. We have also reevaluated levels of sequence divergence among species in this clade, noting differences between loci in regions of low recombination (as in all chromosome 4 loci) relative to other loci. Overall, none of the data analyzed were consistent with recent introgression exclusively between *D. simulans* and *D. mauritiana*.

Introduction

The *Drosophila simulans* species complex has been one of the most studied speciation model systems for the past twenty years (e.g., Coyne, 1984; Coyne, 1992; Zeng & Singh, 1993; Coyne, Crittenden & Mah, 1994; Palopoli & Wu, 1994; Price, 1997; Ting et al., 1998; Kliman et al., 2000). Its three species, *D. simulans*, *D. sechellia*, and *D. mauritiana*, are very closely related to each other and to *D. melanogaster*. However, their evolutionary relationships to each other have not been resolved, and every possible phylogenetic pairing between them has been suggested at some point based on available sequences or other kinds of data (see review in Kliman et al., 2000).

Recently, Ting, Tsaur and Wu (2000) performed a phylogenetic analysis of the *Odysseus* locus, which contributes to male hybrid sterility between *D. simulans* and *D. mauritiana*. Sequence data from the three species (and *D. melanogaster* as an outgroup) strongly suggested that *D. simulans* and *D. mauritiana* are more closely related to each other than either is to *D. sechellia*. The authors suggested that nuclear gene flow between *D. simulans* and *D. mauritiana* or lineage sorting may have obscured the species relationships in other sequence studies. However, *Odysseus* purportedly cannot introgress between these species since it would cause sterility in a foreign genetic background, so its analysis might be more likely to reconstruct the true 'species phylogeny.' Recent sequence data has suggested that *D. mauritiana* may have recently acquired a *D. simulans* mitochondrial haplotype (Ballard, 2000), and the abundance of this new haplotype suggests either selection favoring the introduced *D. simulans* sequence or fairly high levels of introgression.

High levels of nuclear gene flow exclusively between *D. simulans* and *D. mauritiana* should decrease divergence between these species without affecting their divergence from *D. sechellia*. It should also increase the number of synapomorphies, as derived states in one species can introgress and fix in the other. Evidence for this pattern of introgression should be readily apparent on the dot (fourth) chromosome. An estimated 120 genes across the genome

contribute to hybrid sterility between these species (Wu & Hollocher, 1998). However, the dot chromosome does not have any known hybrid sterility factors and can be made homozygous in a heterospecific genetic background with no apparent detrimental effects (Coyne & Berry, 1994). Hence, while linkage with hybrid sterility factors may impede introgression along much of the genome, the dot chromosome should introgress freely between species. Only one study has investigated sequence variability among these species on the dot chromosome (Hilton, Kliman & Hey, 1994). However, very few sequence differences between these species were observed, and none were phylogenetically informative, possibly because primarily coding sequences were studied.

This study addresses whether hybridization between *D. simulans* and *D. mauritiana* could have biased previous phylogenetic studies that used nuclear sequence data, as suggested by Ting, Tsaur and Wu (2000). We have sequenced 1161 bases from intron two of the *eyeless* gene on the dot chromosome in these three species, focusing especially on available lines of *D. mauritiana*. We have also reanalyzed much of the published sequence data from these taxa, testing the fit to a strict isolation model.

Materials and methods

We surveyed sequence variation in the second intron of the *eyeless* gene in 10 lines of *D. mauritiana*, seven lines of *D. simulans*, four lines of *D. sechellia*, and one line each of *D. melanogater* (OregonR) and *D. yakuba* (Tai 18). The *D. mauritiana* lines surveyed were: Bowling Green, 72, 75, 105, 197, 207 (from J. Coyne), MS9, MS11, MS34, and G52 (from B. Ballard). The *D. simulans* lines surveyed were C167.4; Florida City, FL; Winters, CA; Ottawa, Canada; Davis, CA; Valparaiso, IN; and Yaounde, Cameroon (from J. Coyne). The *D. sechellia* lines surveyed were 4, 15, 22, and 24 (from J. Coyne, collected on Cousin Island).

Single fly squish preparations were made from each isofemale line (Gloor & Engels, 1992). A 1300 bp region of the intron located between the second and third exons of the *eyeless* gene was amplified in two separate PCR reactions. The region proximal to exon 2 was amplified using primers 5′-ACTTACTACCACTTAACAGATGATGAATG-3′ and 5′-GGTGACCTGACAGAGAGTACTTAAC-3′. These primers amplify an approximately 600 bp fragment from *D. simulans* subgroup species. The second, overlapping amplification used primers 5′-GCGGAG TAGATTATAGGCATTCCTC-3′ and 5′-GAAAACCT CTGGCGAGCCC-3′. These primers amplify an approximately 700 bp fragment from *D. simulans* subgroup species. Standard 50 μl PCR reactions were prepared with 1.5 mM $MgCl_2$ and 1 μl of the single fly squish preparation. Annealing temperatures were 50°C for the first amplification and 56°C for the second. PCR was performed with 32 cycles of 1-min 94°C, 1-min annealing temperature, and 2-min 72°C.

PCR products were gel-extracted (Qiagen Gel Extraction kit) and eluted in 30 μl water. Sequencing was performed using the ABI BigDye Terminator mix, and sequences were run on the ABI 377 at the Louisiana State University Museum of Natural Science. Each PCR was sequenced in both directions, and we report approximately 1140 bp of this sequence. Sequences were aligned with ClustalW in BioEdit. SITES and WH (Hey & Wakeley, 1997) were used to characterize the variability within and between species, as well as to test the strict isolation model.

Results

Table 1 presents the variation present within *D. simulans*, *D. mauritiana*, and *D. sechellia* in the second intron of the *eyeless* gene. We noted a 592 bp insertion/deletion present in *D. melanogaster* but absent in all strains of *D. simulans*, *D. mauritiana*, and *D. sechellia* (not shown in Table 1). A 24 bp repetitive sequence was present in its place in the *D. simulans* clade species (see Table 1), and this sequence could not be aligned to any part of the *D. melanogaster* 592 bp sequence. To determine whether this large difference resulted from a deletion in the lineage leading to the *D. simulans* complex or an insertion in the lineage leading to *D. melanogaster*, we amplified and sequenced this region in *D. yakuba*. The longer *D. melanogaster* sequence was present in *D. yakuba*, suggesting that it was lost in the lineage leading to the *D. simulans* complex. Three other smaller indels (1–6 bp) also differentiated *D. melanogaster* from the three *D. simulans* complex species. In addition, there were 39–42 base differences that differentiated

Table 1. Sequence polymorphisms and differences among species in the *Drosophila melanogaster* species complex. The 592 bp deletion in the *D. simulans* clade species is not shown

		111111111	1111111111	1111112222	2222222222	2222222222	2222222333	3334455555	6666666788	9999999999	111111111
Position	2446688999	9111222222	2222333333	7788890002	2222233333	3333344444	4444558001	1493636799	0001888216	1114555668	000111111
	8491289012	3789012345	6789012345	4724612791	5678901234	5678901234	5678275015	7972675334	7890123824	1256024386	268001125
											342593544
Consensus	T*GCGAAAAT	G*********	**********	TGCTTGGACG	CACATCATCA	TCATCAGGTA	GCACCAAAAA	CACCTATCCT	****GACACT	CCCTGAATAT	CGGGCCTGG
MauBG	----------	----------	----------	----------	----------	----------	----------	-TA--C----	----***---	----------	---------
MauG52	----------	----------	----------	----------	----------	----------	----------	-TA--C----	----***---	----------	---------
MauMS9	----------	----------	----------	----------	----------	----------	----------	-TA--C----	----***---	----------	---------
MauMS11	----------	----------	----------	-----T----	----------	----------	----------	-TA--C----	----***---	----------	---------
MauMS34	----------	----------	----------	----------	----------	----------	----------	-TA--C----	----***---	----------	---------
Mau72	----------	----------	----------	----------	----------	----------	----------	-TA--C----	----***---	----------	---------
Mau75	----------	----------	----------	----------	----------	----------	----------	-TA--C----	----***---	----------	---------
Mau105	----------	----------	----------	----------	----------	----------	----------	-TA--C----	----***---	----------	---------
Mau197	----------	----------	----------	----------	----------	----------	----------	-TA--C----	----***---	----------	---------
Mau207	----------	----------	----------	----------	----------	----------	----------	-TA--C----	----***---	----------	---------
Sec4	----------	----------	----------	----------	----------	----------	----------	----C-----	----------	----------	--A------
Sec15	----------	----------	----------	----C-----	----------	----------	----------	----C-----	----------	----------	--A------
Sec22	----------	----------	----------	----------	----------	----------	----------	----C-----	----------	----------	--A------
Sec24	----------	----------	----------	----------	----------	----------	----------	----C-----	----------	----------	--A------
SimC167	----------	----------	----------	----------	----------	----------	----------	----------	----------	--T-------	---------
SimDavis	----------	----------	----------	---C------	----------	----------	----------	----------	----------	----------	---------
SimFlaCty	----------	----------	----------	---C------	----------	----------	----------	----------	----------	----------	---------
SimOttawa	----------	----------	----------	---C------	----------	----------	----------	----------	----------	----------	---------
SimValp	----------	----------	----------	---C------	----------	----------	----------	----------	----------	----------	---------
SimWinter	----------	----------	----------	---C------	----------	----------	----------	----------	----------	----------	---------
SimYaounde	----------	----------	----------	----------	----------	----------	----------	----------	ACCC-----C	----------	---------
Mel_OreR	CCATA*****	*TAAACTAAA	TATGAATACA	CTA-A-TGAT	**********	**********	****TGGCGG	T--T--GG**	-------TG-	TG-CTCGCTA	TA-AT*CCC

Table 2. Tests of the strict isolation model for *D. mauritiana* and *D. simulans*

Test	$\hat{\theta}_{sim}$	$\hat{\theta}_{mau}$	$\hat{\theta}_{anc}$	T	P_1	P_2
All 16 loci	87.734	57.034	223.951	0.315	0.4205	0.1458
All loci but *In(2L)t*	78.719	43.526	211.539	0.293	0.4148	0.1843
All loci but *In(2L)t* and *OdsH*	111.504	55.837	154.772	0.231	0.8720	0.1633

The estimated values of the four model parameters were derived by simulation using the WH program written by Jody Hey (Wang, Wakeley & Hey, 1997). The first test used all 16 loci; the second excluded *In(2L)t*, for which evidence of recent gene flow between *D. sechellia* and *D. simulans* has been reported (Kliman et al., 2000); the third excluded *In(2L)t* and *OdsH*, the latter hypothesized to be resistant to gene flow (Ting, Tsaur & Wu, 2000). P_1, proportion of times simulated WWH test statistic (Wang, Wakeley & Hey, 1997) exceeded that of the actual data. P_2, proportion of times simulated χ^2 test statistic (Kliman et al., 2000) exceeded that of the actual data.

the *D. melanogaster* sequence from the other three species.

Based on the proposal of Ting, Tsaur and Wu (2000) of gene flow between *D. simulans* and *D. mauritiana*, we predict that *D. simulans* and *D. mauritiana* should be more similar in sequence to each other than either is to *D. sechellia*. We did not observe any phylogenetically informative sites: there were no substitutions relative to *D. melanogaster* that were shared between any pair of *D. simulans* complex species (see Table 1). However, several character states were shared by *D. simulans* and *D. sechellia* that distinguished them from *D. mauritiana*, and all of these matched the *D. melanogaster* sequence. At least one polymorphic site was detected within each species, with the most in *D. simulans*, though no polymorphisms were shared between species.

We used two tests of the strict isolation model (i.e., no gene flow subsequent to initial isolation of populations) to contrast levels of interspecies divergence (here, fixed differences) with levels of shared polymorphism (Wang, Wakeley & Hey, 1997; Kliman et al., 2000). Both tests compare test statistic values to the distribution generated by coalescent simulations performed by the WH program (see Wang, Wakeley & Hey, 1997; Kliman et al., 2000 for details). Since *D. simulans* and *D. mauritiana* share polymorphism at several loci (Kliman et al., 2000), tests of the isolation model were performed on this species pair using sequences from 16 loci of all three members of the *D. simulans* species complex (14 loci described in Kliman et al. (2000), region A of *OdsH* (Ting, Tsaur & Wu, 2000) and our sequences of *eye*). If gene flow occurred recently between *D. simulans* and *D. mauritiana* at regions besides *OdsH*, as implied by Ting, Tsaur and Wu (2000), this test should reject a strict isolation model particularly when *OdsH* is excluded from the analyses. Neither test rejected a strict isolation model (see Table 2). To confirm this result, we also used WH to test for gene flow between *D. simulans* and *D. mauritiana* based on patterns of linkage disequilibrium (see Machado et al., 2002), but these were also nonsignificant. However, failure to reject the isolation model may be due to limited power.

Besides, to determine how the patterns of variation observed in regions of low recombination contrasted with those in other regions, we compared levels of average pairwise divergence among species pairs in 'hitchhiking' regions (*ase*, *ci*, and *eye*) to other loci (Table 3). Divergence at loci in these hitchhiking regions is generally lower than divergence at other loci among members of the *D. simulans* complex (see Figure 1). However, with the exception of noncoding divergence between *D. mauritiana* and *D. sechellia* ($F_{1,11} = 8.486$, $p = 0.014$, not significant after correction for multiple tests), this difference in divergence was not statistically significant. This trend does not appear to result from differential constraints among loci for two reasons. First, divergence among the two classes of loci is similar when members of the *D. simulans* complex are compared to the outgroup species *D. melanogaster*. Second, the pattern is observed for both coding and noncoding regions. Still, data from more loci are needed to provide sufficient statistical power before strong inferences can be made.

Our sequence for *D. melanogaster* OregonR was identical to that already in the databases (EMBL/GenBank Accession AJ131630). All other sequences have been submitted to these databases (Accession numbers AF491788–AF491807).

Table 3. Average pairwise divergence

Locus	L		sim versus mau		sim versus sec		mau versus sec		sim versus mel		mau versus mel		sec versus mel	
	cod	nc	cod	nc	cod	nc	cod	nc	cod	nc	cod	nc	cod	nc
Adh	570	139	0.00984	0.03117	0.00894	0.04892	0.01582	0.05876	0.01928	0.06447	0.02546	0.06192	0.01693	0.07309
est6	1472	57	0.02352	0.04265	0.02952	0.08369	0.02422	0.06465	0.05013	0.06411	0.04979	0.06250	0.05038	0.05982
Zw	1166	157	0.00699	0.02056	0.01208	0.02216	0.00706	0.02351	0.03443	0.04761	0.03450	0.05064	0.03970	0.05149
per	1682	196	0.01674	0.03421	0.01441	0.04618	0.02104	0.04111	0.02971	0.07953	0.03469	0.08700	0.03197	0.10061
yp2	1046	68	0.00383	0.00655	0.00288	0.02212	0.00511	0.04178	0.01999	0.11566	0.02219	0.12533	0.02124	0.13270
z	816	183	0.00824	0.01583	0.00726	0.03516	0.00728	0.03991	0.02834	0.08423	0.02654	0.08549	0.02722	0.08964
Sxl	–	297	–	0.00897	–	0.02125	–	0.02383	–	0.03624	–	0.03532	–	0.05649
w	–	226	–	0.02715	–	0.02900	–	0.02448	–	0.07035	–	0.06259	–	0.07597
janus	558	514	0.01105	0.02805	0.02040	0.04635	0.01986	0.04044	0.03594	0.06882	0.03546	0.06463	0.04213	0.07644
hb	–	291	–	0.00943	–	0.00770	–	0.01051	–	0.03237	–	0.03275	–	0.03526
mt:ND5	277	–	0.03209	–	0.03254	–	0.04377	–	0.03931	–	0.05054	–	0.05099	–
In(2L)t	–	722	–	0.02415	–	0.02607	–	0.03178	–	0.03819	–	0.04068	–	0.04149
ci	957	118	0.00552	0.00847	0.00429	0.00847	0.00436	0.00000	0.04613	0.08482	0.04723	0.07364	0.04602	0.07634
eye	–	1161	–	0.00362	–	0.00286	–	0.00472	–	0.03717	–	0.03918	–	0.03808
ase	1067	–	0.00244	–	0.00189	–	0.00435	–	0.02515	–	0.02669	–	0.02442	–
OdsH	770	–	0.04462	–	0.03990	–	0.04740	–	0.13479	–	0.13856	–	0.12323	–

Divergence was calculated as the average number of differences per base pair for all pairwise interspecies comparisons.
Abbreviations: L, approximate number of bases compared; cod, coding regions; nc, noncoding regions; sim, *D. simulans*; mau, *D. mauritiana*; sec, *D. sechellia*; mel, *D. melanogaster*.

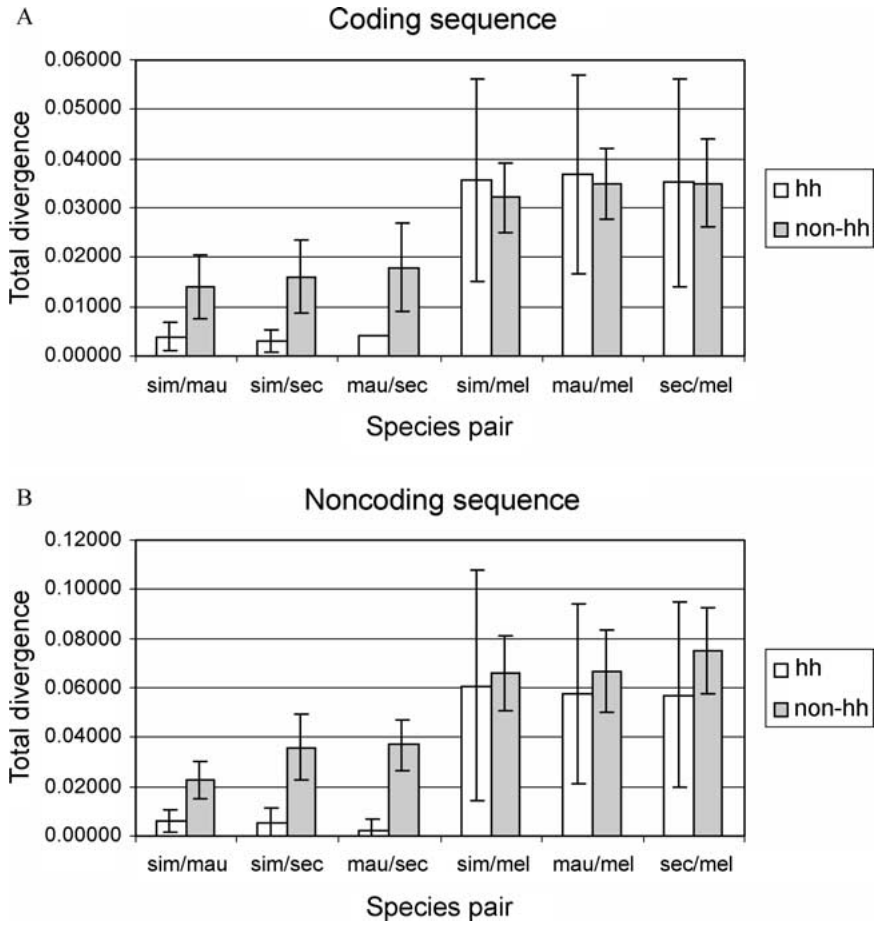

Figure 1. Divergence at 'hitchhiking' loci versus other loci. Shown are average divergences of 'hitchhiking' loci (*ase* and *ci* for coding regions, *ci* and *eye* for noncoding regions) versus all other loci. *OdsH* was excluded, since its divergence is much greater than that of other loci, presumably because of a history of adaptive evolution (Ting, Tsaur & Wu, 2000). Averages are unweighted by sequence length. 95% confidence intervals were calculated directly from the variance among divergence estimates for each class. Abbreviations: hh, hitchhiking loci; non-hh, non-hitchhiking loci; sim, *D. simulans*; mau, *D. mauritiana*; sec, *D. sechellia*; mel, *D. melanogaster*.

Discussion

Ting, Tsaur and Wu (2000) suggested that most analyses of the *D. simulans* clade species failed to identify almost any phylogenetically informative sites because of recent nuclear introgression between *D. simulans* and *D. mauritiana*. The mitochondrial genome of *D. simulans* appears to have invaded *D. mauritiana* recently (Ballard, 2000), and it is possible that nuclear gene flow occurred simultaneously. In contrast, phylogenetic analyses of loci associated with hybrid sterility could provide a more accurate depiction of the species phylogeny, since these loci cannot introgress due to their fitness consequences in heterospecific genetic backgrounds. According to this hypothesis, loci on the dot chromosome should be able to introgress freely between these species, since there is no obvious fitness consequence to make that chromosome homozygous in a heterospecific genetic background (Coyne & Berry, 1994). If this hypothesis is correct, then *D. mauritiana* and *D. simulans* sequences at the *eyeless* locus should be more similar to each other than sequences from either species are to *D. sechellia*, supporting a phylogeny in which the pair of hybridizing species appear to share the more recent common ancestor. Taken to an extreme, the hypothesis would predict no differences between *D. mauritiana* and *D. simulans* sequences at this locus. Regardless of the fact that such an inferred phylogeny may not reflect the natural history of isolation events, it would be fairly well supported.

The purpose of our study was to evaluate the suggestion of Ting, Tsaur and Wu (2000) that nuclear gene flow between *D. simulans* and *D. mauritiana* biased previous phylogenetic studies of this clade. In contrast to their prediction, we observed the greatest similarity in sequence to be between *D. simulans* and *D. sechellia*. Similar results were obtained in the study of Hilton, Kliman and Hey (1994) at another dot chromosome locus: greater differentiation was observed at *Cubitus interruptus* between *D. mauritiana* and *D. simulans* than between either and *D. sechellia*. Finally, using published sequence data from these three species along with the new data, all of our tests failed to reject a strict isolation model, irrespective of whether data from *OdsH* was included. Together, these results suggest that recent nuclear introgression between *D. simulans* and *D. mauritiana* is not the cause for the lack of informative sites for phylogenetic analyses of this clade.

In the course of these analyses, we observed reduced divergence among members of the *D. simulans* complex at loci in regions of low recombination ('hitchhiking loci') relative to other loci ('nonhitchhiking loci'). Consider that total divergence is the sum of (i) divergence since initial isolation and (ii) divergence prior to isolation, and that their expected ratio is $T_1{:}pN_e$ (where T_1 is the time in generations since isolation, p is ploidy, and N_e refers to the population size of the common ancestor). Low recombination should locally reduce N_e (Hill & Robertson, 1966; Maynard Smith & Haigh, 1974; Charlesworth, Morgan & Charlesworth, 1993; Comeron & Kreitman, 2002) and, therefore, the divergence prior to isolation. While such 'Hill–Robertson' effects would also reduce divergence from *D. melanogaster*, the effect would be proportionally smaller. However, unless N_e of the common ancestor of *D. simulans* complex species was on the order of T_1, it would be difficult to explain the apparently reduced divergence in hitchhiking loci. Also, less effective purifying selection in regions of reduced recombination may offset the divergence-reducing effect of a shallower gene tree.

A relaxation of the strict isolation model among all three species might also explain the pattern. One striking finding at the two chromosome 4 genes, *ci* and *eye*, is the absence of phylogenetically informative sites (i.e., sites with a character state in two in-group species that differs from that shared by the remaining in-group species and *D. melanogaster*). At a locus with historically low N_e, the number of fixed synapomorphies is nearly proportional to $T_2 - T_1$, where T_2 is the time back to the initial isolation of the ancestor of two in-group species from the remaining in-group species. The lack of fixed synapomorphies between any species pair at *ci* and *eye* is consistent with temporally close isolation events, but fixed synapomorphies at other loci conflict with this. Gene flow between a pair of species (but not among all three) would effectively decrease T_1, increasing the number of phylogenetically informative fixations. Gene flow among all three species would effectively decrease both T_2 and T_1; in fact, sufficient gene flow could make T_2 equal to the reduced T_1. The decreased divergence among members of the *D. simulans* complex at the chromosome 4 loci, coupled with the lack of phylogenetically informative nucleotide substitutions, could be consistent with recent gene flow among all three species, rather than between a particular pair of species.

As noted earlier, no loci contributing to hybrid fitness reduction among members of the *D. simulans* complex have been mapped to chromosome 4. Thus, if given the opportunity, gene flow would be relatively unimpeded at these loci. Emigration of *D. simulans* to both the Seychelles and Mauritius subsequent to the initial isolation events might explain the data. Such waves of migration of *D. simulans* from the mainland to the islands inhabited by *D. mauritiana* and *D. sechellia* could contribute to homogeneity among all three species, even if the latter did not directly exchange genes.

Ting, Tsaur and Wu (2000) note that *OdsH* provides strong phylogenetic signal placing *D. sechellia* as the outgroup to *D. simulans* and *D. mauritiana*. In fact, the latter share five fixed synapomorphies and one site at which the derived state is fixed in one species and polymorphic in the other. None of the 15 other loci supports this tree as well, and Ting, Tsaur and Wu (2000) note that the strong signal at *OdsH* may reflect the contribution of its interspecies variation to hybrid sterility. Interestingly, *D. mauritiana* and *D. sechellia* share three fixed synapomorphies at the *glucose-6-phosphate-dehydrogenase* (*Zw*) locus, and, relative to *OdsH*, this gene shows roughly one-third of the average divergence from *D. melanogaster*. A phylogeny estimated from sequence data of this locus would directly conflict with the one derived from *OdsH*.

Kliman et al. (2000) noted that variability at 14 loci across the genomes of these species was consistent with simple allopatric divergence without gene flow. Alternative phylogenies could arise from ancestral lineage sorting. Further, if there were population

structure in the ancestor, gene trees may be biased towards implying a particular species tree. Thus, it seems more likely that the unambiguous phylogeny recovered by Ting, Tsaur and Wu (2000) reflects the fast rate of adaptive evolution of the *Odysseus* gene rather than introgression at other loci. Alternatively, the 'clean' phylogeny may reflect the way the lineages assorted, with no subsequent loss of phylogenetic signal from gene flow.

Acknowledgements

We thank D. Ortiz-Barrientos for helpful comments on this manuscript and N. Crochet with assistance with the DNA sequencing. This work was supported by NIH grant GM58060 subcontracted through J. Hey at Rutgers University and NSF grant 9980797 to MAFN.

References

Ballard, J.W.O., 2000. When one is not enough: introgression of mitochondrial DNA in Drosophila. Mol. Biol. Evol. 17: 1126–1130.

Charlesworth, B., M.T. Morgan & D. Charlesworth, 1993. The effect of deleterious mutations on neutral molecular variation. Genetics 134: 1289–1303.

Comeron, J.M. & M. Kreitman, 2002. Population, evolutionary and genomic consequences of interference selection. Genetics 161: 389–410.

Coyne, J.A., 1984. Genetic basis of male sterility in hybrids between two closely related species of *Drosophila*. Proc. Natl. Acad. Sci. USA 81: 4444–4447.

Coyne, J.A., 1992. Genetics of sexual isolation in females of the *Drosophila simulans* species complex. Genet. Res. Camb. 60: 25–31.

Coyne, J.A. & A. Berry, 1994. Effects of the fourth chromosome on the sterility of hybrids between *Drosophila simulans* and its relatives. J. Hered. 85: 224–227.

Coyne, J.A., A.P. Crittenden & K. Mah, 1994. Genetics of a pheromonal difference contributing to reproductive isolation in *Drosophila*. Science 265: 1461–1464.

Gloor, G.B. & W.R. Engels, 1992. Single-fly DNA preps for PCR. Drosoph. Inf. Serv. 71: 148–149.

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

Hill, W.G. & A. Robertson, 1966. The effect of linkage on the limits to artificial selection. Genet. Res. 8: 269–294.

Hilton, H., R.M. Kliman & J. Hey, 1994. Using hitchhiking genes to study adaptation and divergence during speciation within the *Drosophila melanogaster* species complex. Evolution 48: 1900–1913.

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

Machado, C.A., R.M. Kliman, J.A. Markert & J. Hey, 2002. Inferring the history of speciation from multilocus sequence data: the case of *Drosophila pseudoobscura* and its close relatives. Mol. Biol. Evol. 19: 472–488.

Maynard Smith, J. & J. Haigh, 1974. The hitch-hiking effect of a favourable gene. Genet. Res. 23: 23–35.

Palopoli, M.F. & C.-I. Wu, 1994. Genetics of hybrid male sterility between Drosophila sibling species: a complex web of epistasis is revealed in interspecific studies. Genetics 138: 329–341.

Price, C.S.C., 1997. Conspecific sperm precedence in *Drosophila*. Nature 388: 663–666.

Ting, C.-T., S.-C. Tsaur & C.-I. Wu, 2000. The phylogeny of closely related species as revealed by the genealogy of a speciation gene, *Odysseus*. Proc. Natl. Acad. Sci. USA 97: 5313–5316.

Ting, C.-T., S.-C. Tsaur, M.-L. Wu & C.-I. Wu, 1998. A rapidly evolving homeobox at the site of a hybrid sterility gene. Science 282: 1501–1504.

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

Wu, C.-I. & H. Hollocher, 1998. Subtle is nature: the genetics of species differentiation and speciation, pp. 339–351 in Endless Forms: Species and Speciation, edited by D.J. Howard & S.H. Berlocher. Oxford University Press, New York.

Zeng, L.-W. & R.S. Singh, 1993. A combined classical genetic and high resolution two-dimensional electrophoretic approach to the assessment of the number of genes affecting hybrid male sterility in *Drosophila simulans* and *Drosophila sechellia*. Genetics 135: 135–147.