JOURNAL OF Evolutionary Biology

doi: 10.1111/jeb.12391

Hybridization occurs between *Drosophila simulans* and *D. sechellia* in the Seychelles archipelago

D. R. MATUTE* & J. F. AYROLES†‡

*Department of Human Genetics, University of Chicago, Chicago, IL, USA
†Department of Molecular Biology and Genetics, Cornell University, Ithaca, NY, USA
‡Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA, USA

Keywords:

adaptation;
Drosophila;
reproductive isolation;
speciation.

Abstract

Drosophila simulans and *D. sechellia* are sister species that serve as a model to study the evolution of reproductive isolation. While *D. simulans* is a human commensal that has spread all over the world, *D. sechellia* is restricted to the Seychelles archipelago and is found to breed exclusively on the toxic fruit of *Morinda citrifolia*. We surveyed the relative frequency of males from these two species in a variety of substrates found on five islands of the Seychelles archipelago. We sampled different fruits and found that putative *D. simulans* can be found in a variety of substrates, including, surprisingly, *M. citrifolia*. Putative *D. sechellia* was found preferentially on *M. citrifolia* fruits, but a small proportion was found in other substrates. Our survey also shows the existence of putative hybrid males in areas where *D. simulans* is present in Seychelles. The results from this field survey support the hypothesis of current interbreeding between these species in the central islands of Seychelles and open the possibility for fine measurements of admixture between these two *Drosophila* species to be made.

Introduction

Because gene flow is a homogenizing force that can swamp genetic and phenotypic differences accumulated between groups of individuals, reproductive isolation must be a key feature in the persistence of species (Dobzhansky, 1937; Coyne & Orr, 2004). The biological processes that limit gene flow between groups of organisms, termed 'reproductive isolating mechanisms', are the central focus of speciation research.

In cases where the nascent or fully formed species come into secondary contact and exchange genetic material, the effectiveness of reproductive isolation mechanisms is tested, with three possible outcomes. First, reproductive isolation can be absolute, such that no hybrids are formed. In this case, speciation has been fully completed. The second possible outcome is that hybridization occurs, and reproductive isolation is not strong enough to avoid the homogenizing effect of interspecific gene flow. In this case, it is likely the two populations will eventually fuse into a single species. The third outcome is that hybrids are formed, but the two species nonetheless persist. This might occur in two different ways. First, if hybridization is a rare event (the number of migrants per generation is much lower than effective population size), the two species will not fuse (Bank *et al.*, 2012). If hybridization is a common process and the hybrids show a reduced fitness (either intrinsic or extrinsic), then the species might persist and prezygotic isolation might become stronger over time as a by-product of selection against maladaptive hybridization (Servedio & Noor, 2003; Hopkins, 2013). No matter how frequently hybrids are produced, if they are fertile, they allow for the migration of alleles from one species into the other (introgression) through hybridization followed by backcrossing to a parental species. Gene flow through fertile hybrids is thought to occur only in relatively recent instances of speciation (e.g. Song *et al.*, 2011; Pardo-Diaz *et al.*, 2012 reviewed in Arnold, 2006; Arnold & Martin, 2009; Hedrick, 2013). Studying secondary contact in natural conditions thus provides a unique opportunity to understand how reproductive isolation evolves.

Correspondence: Daniel R. Matute, University of Chicago, 920 E 58th Suite 411, Chicago, IL 60637, USA.
Tel.: +1 773 834 0525; fax: +1 773 834 0505;
e-mail: dmatute@uchicago.edu
Sequences GenBank ID: KJ425848 - KJ426017, KJ426058 - KJ426257.

Studies of gene flow during the build up of reproductive isolation in *Drosophila* have been limited by the very few known hybrid zones, that is, areas of hybridization between two or more species. Many pairs of *Drosophila* species have overlapping ranges (reviewed in Coyne & Orr, 1989 and Yukilevich, 2012), but few have been proposed to currently hybridize. *D. recens* and *D. subquinaria* coexist across southern Canada and there is evidence that they hybridized in the past (Jaenike *et al.*, 2006); *D. metzii* and *D. pellewae* produce hybrids in their secondary contact area in Northern Colombia and Panama (Bedichek-Pipkin, 1968).*D. subobscura* has recently invaded the island of Madeira (Pinto *et al.*, 1997), where its sister species, *D. madereinsis* is endemic. Nuclear polymorphism scans show the existence of recent introgression between these two species (Herrig *et al.*, 2013), but no natural hybrids have been found yet. Finally, *D. yakuba* and *D. santomea* hybridize in the midlands of the volcanic island of São Tomé and produce a high proportion of hybrids in natural conditions (Lachaise *et al.*, 2000; Llopart *et al.*, 2005a,b; Matute, 2010); they have been shown to exchange genes (Llopart *et al.*, 2005a,b; Bachtrog *et al.*, 2006; Llopart *et al.*, 2014). As this certainly constitutes an underestimate of the actual number of hybrid zones, the identification of more areas of secondary contact between species of *Drosophila* and their potential for admixture have much to offer studies of speciation.

The *simulans* subcomplex is composed of three sister species (*D. simulans, D. mauritiana* and *D. sechellia*) all of which diverged approximately 0.2 million years ago (based on 1–2% divergence in synonymous sites, Coyne & Kreitman, 1986; McBride, 2007; Garrigan *et al.*, 2012; Kliman *et al.*, 2000; McDermott & Kliman, 2008). *D. simulans* is a human commensal that is thought to have originated in Madagascar and currently has a worldwide range, including the Seychelles archipelago in the Indian Oceans. *D. sechellia,* an endemic species to the Seychelles (Tsacas & Bächli, 1981; Lachaise & Silvain, 2004; David *et al.*, 2007) and *D. mautitiana*, an endemic to the Mauritius islands (Kliman *et al.*, 2000; McDermott & Kliman, 2008; Garrigan *et al.*, 2012), are sister species of *D. simulans*. The restricted distribution of *D. sechellia*, and the presumed colonization of the Seychelles archipelago by *D. simulans* with the recent settlement of humans (XVII century) offers the opportunity to pursue a variety of questions related to adaptation, reproductive isolation and the interface of the two processes.

First, there are clear ecological differences between species that may function as barriers to gene flow. *D. sechellia* specializes in breeding in *Morinda citrifolia* (Rubiaceae, commonly known as Noni, Legal *et al.*, 1994; Farine *et al.*, 1996; Amlou *et al.*, 1998) and is thought to be restricted to breeding and ovipositing only in this fruit (Louis & David, 1986; R'Kha *et al.*, 1991; Jones, 2005; Erezyilmaz & Stern, 2013). The *M. citrifolia* fruit is noxious and repulsive to most *Drosophila* species, including *D. simulans* (Higa & Fuyama, 1993; Farine *et al.*, 1996). *D. sechellia*, in contrast, is attracted to this fruit and can tolerate its toxic compounds, hexanoic and octanoic acids.

A second tier of reproductive isolation between *D. simulans* and *D. sechellia* involves behavioural prezygotic mechanisms. In nonchoice experiments, *D. sechellia* females discriminate strongly against *D. simulans* males (Cobb *et al.*, 1989; Coyne, 1996; Gleason *et al.*, 2005). *D. simulans* females also discriminate against *D. sechellia* males, but the magnitude of the isolation is lesser in this direction of the cross (Coyne, 1996; Tomaru *et al.*, 2004). Despite these and other known mechanisms of reproductive isolation (Cabot *et al.*, 1994; Price, 1997), interspecific crosses do produce hybrids: the two reciprocal crosses produce F_1 hybrid females that are fertile and F_1 hybrid males that show complete hybrid sterility due to defects in spermatogenesis. Any current gene flow between these two species could happen only through the hybrid females.

Although speciation could have occurred strictly in allopatry (Kliman *et al.*, 2000), *D. simulans* and *D. sechellia* have been proposed as a case of divergence in the presence of gene flow (Garrigan *et al.*, 2012). In *D. sechellia*, two different haplotypes segregate in a small region of chromosome 2L (cytological position 34A). One of these haplotypes resembles one of the haplotypes present in *D. simulans* (Solignac *et al.*, 1986; Kliman *et al.*, 2000), whereas the other one is an allele found only in *D. sechellia*, suggestive of either recent gene flow or of retention of ancestral polymorphism. Genome resequencing of individual isofemale lines has suggested the existence of introgression between *D. sechellia* and *D. simulans* in the last 5000 years, strengthening the case for divergence with gene flow (Garrigan *et al.*, 2012; Brand *et al.*, 2013).

Even though this pair of species has been extensively studied in the laboratory, no field survey has explored whether there is currently hybridization between the species in the wild. Given that recent studies have found evidence of introgression between the two species, we aimed to establish whether such hybridization is ongoing. We collected individuals from the *simulans* subcomplex on five islands of the Seychelles archipelago and identified individuals as *D. simulans*, *D. sechellia* or interspecific hybrids. We document, to our knowledge for the first time, the presence of hybrids between the two species in the Seychelles islands. The results from this survey lend strong support to the hypothesis of current interbreeding between these species and open the possibility of introgression studies between them.

Materials and methods

Sampling

The Seychelles archipelago is an enclave of 115 islands in the Indian Ocean that lie off the coast of East Africa,

north-east of Madagascar. We sampled five islands in the central and north part of the archipelago: Mahé, Praslin, Marianne and La Digue are relicts of the Mauritia microcontinent and are granitic in origin (Torsvik *et al.*, 2013), whereas Denis is the most northern island of the group and is a volcanic sand cay. All samples were collected in January 2012.

Females from the three species of the *simulans* subcomplex are largely indistinguishable to the untrained eye (the two species however show differences in wing morphology, Orgogozo & Stern, 2009). For that reason, we restricted our taxonomic survey to males, because we could use phenotypic traits to identify individuals from the two species. We collected flies that were either breeding or standing on the surfaces of mangoes, papayas, figs and *M. citrifolia* ripe fruits laying on the forest ground. In addition, we netted and collected flies breeding on vegetable debris in the main urban centres of Victoria, the largest and capital city of Seychelles. Males and females collected under the netting protocol were anaesthetized using FlyNap (Carolina Biological Supply Company, Burlington, NC, USA). All males were fixed in RNAlater (Invitrogen, Carlsbad, CA, USA). The collection sites are listed in Table S1.

Laboratory-produced individuals: crosses

Our approach to identify field-collected individuals involved the use of three phenotypic traits. Male genital morphology is one of the fastest evolving traits within the *melanogaster* group of species and one of the classical traits used to identify species (Markow & O'Grady, 2005). The posterior lobe of the male genital arch shows significant differences in both size and shape among closely related species of the *Drosophila melanogaster* species subgroup. *D. sechellia* has small and elongated genital arches, whereas *D. simulans* has spheroid large genital arches (Macdonald & Goldstein, 1999). The hybrids show intermediate genital morphology and can be distinguished from the parental species (Fig. 1, Coyne & Kreitman, 1986). We note that the hybrid males from the two reciprocal crosses cannot be distinguished (Macdonald & Goldstein, 1999). We also used two additional male traits that also show clear differences between species: the number of teeth in the sex combs and the length of the foreleg tibia (one of the segments of the proximal legs). *D. simulans* sex combs show an average number of teeth that ranges between nine and 11 (Among line SD = 0.69), whereas *D. sechellia* ranges between 12 and 13 (Among line SD = 0.81; Coyne & Kreitman, 1986; Macdonald & Goldstein, 1999). Tibial length in *D. simulans* is 0.464 mm on average (Among line SD = 0.017) and in *D. sechellia* 0.481 mm (Among line SD = 0.013). The F_1 hybrid males show both intermediate number of teeth (mean = 11.8, SD = 0.8, Macdonald & Goldstein, 1999) and intermediate tibial lengths (mean = 0.470 mm, SD = 0.07 mm, Macdonald & Goldstein, 1999).

We compared the field-collected individuals with laboratory-produced crosses. We used two isofemale lines that had been previously sequenced: *D. simulans* w^{501} (Hu *et al.*, 2012) and *D. sechellia* Robertson 3C (Tsacas & Bächli, 1981; Clark *et al.*, 2007). *D. simulans* was kept on corn meal, whereas *D. sechellia* was kept on ripe *Morinda* fruit. We produced laboratory-reared individuals (henceforth referred to as laboratory crosses) for the two pure species, the two reciprocal F_1s (*D. sechellia*. se*D. simulans* and ♀*D. simulans*. si*D. sechellia*) and the four possible backcrosses (each of the two types of F_1 females crossed to the males of the parental species). To make backcrossed individuals, hybrid females were crossed to males from the lines that were used to produce them. These individuals were raised and scored in the same manner as above-mentioned crosses. We collected 100 males per genotype and anaesthetized them until dissection.

Phenotypic scoring

We followed similar procedures to score traits in laboratory-produced and field-captured individuals. For laboratory-produced individuals, we anaesthetized 20 males from each of the intraspecific and interspecific crosses. For field-captured individuals, we extracted them from RNAlater and lightly washed them with distilled water. Briefly, genital arches were cut and mounted on a 50% glycerol solution. The mounted genitalia were photographed and the area of the genital lobes was calculated

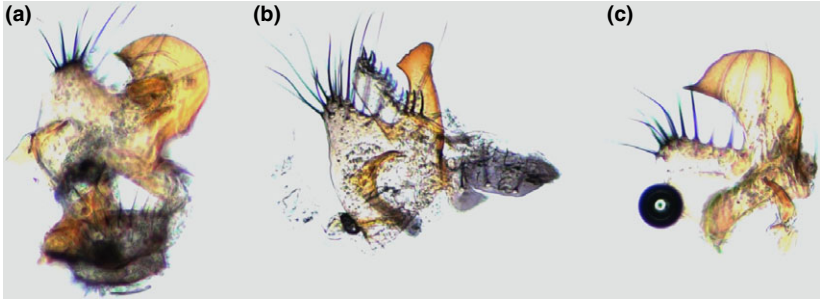

Fig. 1 *Drosophila simulans* (a) and *D. sechellia* (b) males have characteristic shapes in the posterior lobes of the male genitalia. F_1 males (c) show intermediate and distinct morphology. All pictures were taken within one hour of the dissection.

with ImageJ (Schneider *et al.*, 2012). Forelegs were cut from the adult fly and mounted on two-sided tape. We counted the number of teeth in the sex combs and measured the length of the tibia. Phenotypic distributions for these three traits are shown in Fig. S1.

Mahalanobis distance calculation

We aimed to assign each field-collected individual to one of eight groups: *D. simulans*, *D. sechellia*, F_1 hybrids (two directions) or backcrossed individuals (four directions). To date, the few attempts to detect *Drosophila* hybrid individuals from field collections have exclusively relied on the use of single traits and have not used a multivariate framework. To assign each collected individual to a genotypic class, we calculated the Mahalanobis distance of each individual to the phenotypic distribution of each of the eight genotypes in laboratory-produced crosses. This distance is based on both the mean and variance of the predictor variables (i.e. area of the posterior genital lobe, number of teeth in the sex combs and length of the tibia), plus the covariance matrix of all the variables. The Mahalanobis distance (Mahalanobis, 1936) between an individual and the average of the laboratory-produced individuals for a given genotype (i.e. either *D. sechellia*, *D. simulans*, each of the two F_1 genotypes, or the four backcrosses) was calculated as:

$$MD_{ij} = (F_i - \mu_{LB})^T \times S_{LB}{}^{-1} \times (F_i - \mu_{LB})$$

where the super-index T denotes matrix transpose, S denotes the covariance matrix of a given dataset (each of the genotypes of the laboratory-produced individuals), F_i is the vector of phenotypic observations in a field-collected individual, i, and μ_{LB} is the vector of average phenotypic observations in the laboratory-produced individuals for each genotype. As the distance was not weighted, all the three traits contributed equally to the assignment.

We assessed whether the phenotypic distributions from the eight phenotypes could be used to assign individuals to each of the eight genotypes by determining the group to which that observation showed the lowest distance. We assigned each of the 100 individuals from the eight genotypes to a genotypic class using a Mahalanobis distance. This assessment provided a measurement of the false-assignment rate that the use of phenotypic distributions and the Mahalanobis distance. We found an extremely high level of misassignment when the eight distributions were used (Table S2). Of the 200 pure species, 57 were correctly identified. Of the 600 hybrid individuals, 185 were correctly identified. Given this low rate of success, we resorted to an alternative approach and restricted our assignment to three different genotypic categories: the two parental species and F_1 hybrids. For this classification, we did not use the individuals produced in backcrosses. We calculated the Mahalanobis distance of each individual to the pure species and F_1 matrices. The second approach correctly identified 377 of the 600 hybrid individuals (F_1s + backcrosses), all the 100 *D. sechellia* individuals and 95 of the 100 *D. simulans* individuals. On the other hand, as in this approach we chose not to include the distributions of the backcrosses, all the backcrossed individuals got assigned to either the parental species or the F_1 hybrids (Table S3). We chose the latter approach to classify field-collected individuals for two reasons. First, the proportion of individuals categorized as the wrong genotype was overall higher in the first approach (~70% of misidentified individuals in the first approach vs. 28.5% of misidentified individuals in the second approach). The second reason was the nature of the misassignment. While the first approach spuriously inflates the amount of assigned backcrosses, and thus of hybridization, the second underestimates the amount of hybridization by assigning backcrossed individuals mostly to the pure species category (Table S3). For the purposes of this report, we found it preferable to underestimate the amount of hybridization in field collections. This assignment, however, does not preclude the possibility of advanced intercrosses on our sample, which we simply have no power to detect.

As using individual inbred lines is not representative of the levels of phenotypic variance within a species, we made a large sample of outbred pure species and F_1 hybrids to compare to the field collections. We used 10 *D. simulans* from cosmopolitan collections and 10 isofemale lines from *D. sechellia* collected in Denis island (Table S4). We made all the possible crosses between these isofemale lines. Denis is the northernmost island of the Seychelles archipelago. No collections of *D. simulans* from this island have previously been reported. We scored 20 individuals per cross which led to having 4000 laboratory-produced individuals for each pure species and 8000 for the F_1 group (4000 for each direction). The mean phenotypic values of the genital lobe area, anterior tibial length, and the number of teeth in the sex combs for the intraspecific and crosses are shown in Figs S2–S4 and show little within-species variation in the three characters. The two-dimensional distribution of each pair of traits in these laboratory-produced individuals is shown in Fig. S5.

Field-collected individuals: species identification

Even though *D. simulans* and *D. sechellia* are not the only *Drosophila* species in the Seychelles islands, we focused our analyses to these two parental species and the hybrids between them. To determine whether an individual was *D. simulans*, *D. sechellia* or a hybrid, we characterized the genital morphology, the number of sex combs, and the anterior tibial length in each individual. For each field-collected individual, we calculated the Mahalanobis distance to each laboratory-produced group (two pure spe-

cies and F_1s pooled together, each group with more than 4000 observations) and assign them to one of the three genotypic groups. These distances are shown in Table S5. The distance of each individual to each genotypic class could not be converted to *P*-values due to the non-normality of their distribution and we assign individuals to the group to which they showed the lowest distance.

Niche specificity

Drosophila sechellia has been described as a specialist of *M. citrifolia* fruits (R'Kha *et al.*, 1991). *D. simulans*, on the other hand, is a cosmopolitan generalist highly sensitive to the toxic compounds produced by the *M. citrifolia* fruit (Jones, 2005; Hungate *et al.*, 2013). To determine whether these two species genuinely have two discrete substrates or if there is any overlap in the substrates used by the two, we determined which substrates were used by the two parental species and the hybrids in the field (if any).

Molecular typing

We used two different molecular markers to assess ancestry of the putative hybrid individuals. First, we used a *Y*-chromosome marker that shows fixed differences between *D. sechellia* and *D. simulans*: *kl-5* is a 500 bp *Y*-chromosome linked marker that has five substitutions fixed between all lines of *D. simulans* and *D. sechellia* surveyed to date (Kopp, 2006; Tao *et al.*, 2007; Larracuente & Clark, 2013). We used previously described sequences (AJ874917.1-AJ874931.1) to design a pair of primers to produce an amplicon of 400 base pairs that includes the five substitutions. We used a second marker, Cytochrome Oxidase I (*COI*), a mitochondrial marker used, which also distinguishes the two species, with no known shared polymorphism between them. We used published primers to amplify 560 bp from *COI* (HM631513.1- HM631531.1, Nunes *et al.*, 2010). This fragment showed nineteen substitutions between *D. simulans* and *D. sechellia*. These two markers allowed us to confirm whether the putative hybrid individuals are indeed hybrids and to establish the genotype of the father and the mitochondrial lineage of the mother. We extracted DNA from single male flies using the Qiagen DNeasy Blood and Tissue Kit and performed PCR amplifications using approximately 25 ng of total DNA. PCR products were directly sequenced. Sequences were edited with 4Peaks 1.7.1 (http://nucleobytes.com/index.php/4peaks) and aligned using the ClustalX program (Larkin *et al.*, 2007). All the individual alleles were deposited in Genbank (KJ425848–KJ426017; KJ426058–KJ426257).

Results

Our goal was to determine the relative frequency of *D. sechellia*, *D. simulans* and any potential hybrids between these two species on different islands of the Seychelles archipelago. Our approach consisted of netting individuals feeding (or resting) on specific ripe fruits on the ground. We collected 14 285 drosophilid males. We restricted all our analyses to males of the *melanogaster* group that were identified by their characteristic genital morphology (i.e. external male genitalia includes genital arch, pair of anal plates and dentate claspers, Bock, 1980; Lemeunier *et al.*, 1986; Markow & O'Grady, 2005).

The most abundant class belonged to *Zaprionus* sp. (38% of the total collected individuals). Only 17% of the collected males belonged to the *melanogaster* subgroup, of which the *simulans* subcomplex is a part (Fig. S6). We characterized what species were found breeding on ripe fruits in the Seychelles (including the fruits of *M. citrifolia*). First, we found that it is possible to capture *D. melanogaster* males on *M. citrifolia* fruits but only in La Digue and Denis Island and only rarely (of all male drosophilids collected in La Digue and Denis island, 0.34% and 0.51%, respectively, were *D. melanogaster*). Unlike previous collections (David & Capy, 1982), we did not find *D. melanogaster* in Victoria. Approximately, 17% of the males that we collected belonged to the *simulans* subcomplex. We focused on the identification of these males and described whether they appeared to be *D. simulans*, *D. sechellia* or hybrids between these two species.

D. simulans and *D. sechellia* coexist and hybridize in the central islands of Seychelles

We used the number of teeth in the sex combs, the length of the tibia and the area of the posterior lobe of the male genital arches could help us distinguish between *D. simulans*, *D. sechellia* and F_1 hybrids (see Methods and Table S2). Using phenotypic traits of 4000 individuals from each pure species and 8000 F_1 hybrid individuals (4000 from each direction), we assigned each field-collected individual to the group to which it showed the lower Mahalanobis distance (Fig. 2, Figs S2–S5). F_1 hybrids produced by the two directions of the cross could not be distinguished from one another (Figs S2–S5). Using this approach, we identified 2160 individuals as *D. sechellia*, 215 as *D. simulans*, and 78 as F_1 hybrids. Results for the 2453 field-collected individuals are shown in Table S5.

We found *D. sechellia* males on the five sampled islands. This species was, by far, the most abundant of the *simulans* subcomplex on all islands sampled, accounting for 88% of the males collected from the *simulans* subcomplex overall. In Denis and Marianne, *D. sechellia* was the only species collected from the *simulans* subcomplex.

Surprisingly, we also found apparent *D. simulans* males on three of the five islands: Mahé, Praslin and La Digue. These islands lie at the centre of the Seychelles

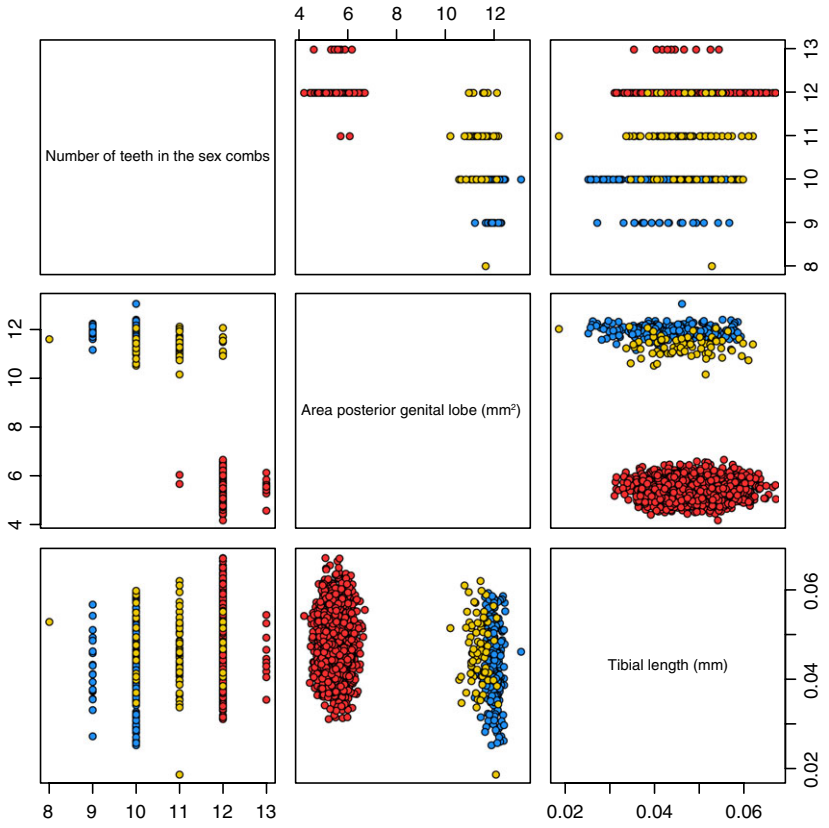

Fig. 2 Scatter plots of pairwise combinations of the three phenotypic traits assessed in this study for 2453 field-collected individuals. Red: putative *Drosophila sechellia* (n = 2160); Blue: putative *D. simulans* (n = 215); Yellow: putative F_1 hybrids (n = 78).

archipelago and are the main hubs of human activity. This geographical distribution suggest that *D. simulans* has followed humans to the Seychelles islands as a commensal species, as has occurred in other geographical locations (Lachaise *et al.*, 1986). This implies that its presence is likely relatively recent, as humans are thought to have settled in the Seychelles in the last 500 years (Cheke, 2010). On all the islands where the two species coexist, *D. simulans* was four-fold (or less) frequent than *D. sechellia*.

Wild-caught hybrid males have a *D. sechellia* father and a mother with a *D. simulans* mitochondria

Strikingly, on the three islands where we found *D. simulans*, we also found F_1 hybrid males between *D. simulans* and *D. sechellia* at lower frequencies than pure *D. simulans* males (see Fig. 3 for the geographic distribution of the two species and the F_1 hybrids). Within the *simulans* subcomplex, the F_1 hybrids were by far the rarest genotype (approximately 3% of the collections from the *simulans* subcomplex, and ~0.5% of the total Drosophilids collected). These results reveal that hybrids between *D. simulans* and *D. sechellia* are presently being produced in nature.

To complement the morphological characterization, we typed *D. simulans*, *D. sechellia* and hybrids with molecular markers to confirm hybrid genotypes in individuals assigned as F_1s based on phenotypes, and to determine the relative contribution of each type of hybrid cross to the F_1s we collected (i.e. ♀ *D. simulans* × ♂ *D. sechellia* vs. ♀ *D. sechellia* × ♂ *D. simulans*). To this end, we used two molecular markers: one paternally inherited (a *Y*-linked marker, *kl-2*) and one maternally inherited (a mitochondrial marker, *COI*). Although these two markers are well-established species-specific markers in molecular taxonomy, we sequenced these two loci in the 10 isofemale lines of each species used to study morphological traits in order to verify the existence of putative fixed differences between the species. All the 10 *D. simulans* isofemale lines were collected in mainland Africa. We collected the *D. sechellia* lines in a single geographical location, Denis island (Table S5). We found four fixed differences in *kl-2* and 22 in *COI* between *D. simulans* and *D. sechellia*. These results confirm that these two loci should be effective at identifying the origin of the mitochondrial genome and *Y*-chromosome in these two species.

We then typed 50 of our field-collected *D. sechellia* and 50 field-collected *D. simulans* from our Seychelles collection. Molecular typing did not identify any new variants compared with the sample of isofemale lines that had been sequenced in previous efforts (Nunes *et al.*, 2010) or the group of ten isofemale lines we

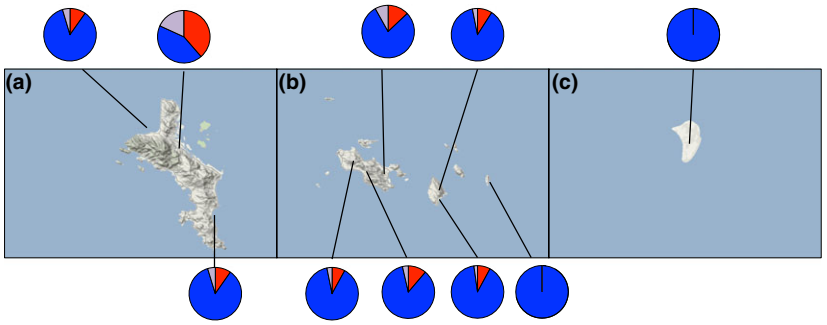

Fig. 3 Geographic distribution of the *Drosophila simulans*, *D. sechellia* and their F_1 hybrids between these two species in the Seychelles archipelago. We sampled five different islands: (a) Mahé, (b) Praslin, La Digue, and Marianne (c) Denis. All figures were created with the RgoogleMaps library in R.

sequenced (see above). For *D. sechellia*, we found no intraspecific polymorphism at either locus. For *D. simulans*, we found no intraspecific variation in *kl-2*. For *COI*, 49 of the 50 sequenced samples have a mitochondrial haplotype that corresponds with a known *D. simulans* haplotype; the remaining sample has a mitochondrial haplotype matching that of *D. mauritiana*. Taking this sample into account, the number of fixed differences between species that could be used for *D. simulans*/*D. sechellia* identification was reduced (*kl-2*: four fixed differences; *CO1*: 9 fixed differences).

We next used these markers to identify the mitochondrial type and the *Y*-chromosome from the field-captured individuals that we identified morphologically as F_1 hybrids. Our results indicate that all the individuals that were categorized as F_1 hybrids were indeed of hybrid origin ($n = 78$), thus validating our approach of hybrid identification using morphological measurements. Strikingly, all the putative F_1 hybrid individuals collected in the Seychelles carried the mitochondrial genotype of *D. simulans* and the *Y*-chromosome of *D. sechellia*. We cannot, however, discern whether these individuals are truly F_1 hybrids, backcrossed individuals (e.g. from the crosses ([♀ (♀*D. simulans*. s*D. sechellia*) × ♂*D. sechellia*]), or further intercrosses. In any case, these results indicate that there is ongoing hybridization between *D. simulans* and *D. sechellia*, and that such hybridization occurs exclusively in the direction of females that carry the *D. simulans* mitochondria and males with the *D. sechellia Y*-chromosome. This direction of the cross is most likely to occur as predicted by the strength of reproductive isolation in laboratory crosses (Lachaise *et al.*, 1986; Coyne, 1992; Gleason *et al.*, 2009).

Drosophila simulans, *D. sechellia* and their F_1 hybrids can be found on a variety of substrates

Drosophila simulans is known to be a generalist species, whereas *D. sechellia* is one of the best documented examples of the evolution of niche specialization in *Drosophila*. We tallied up the substrates on which each species was collected and calculated which species was more likely to be found on each of the four different substrates sampled (i.e. mango, papaya, figs, and *M. citrifolia*). Figure 4 shows the proportion of males we identified as *D. simulans*, *D. sechellia* and hybrids found in each substrate.

We found that, as expected, males that we classified as *D. simulans* based on genital, anterior tibial, and sex combs morphology could be found in a variety of substrates. Unexpectedly, however, these males were also found on *M. citrifolia*, a substrate previously thought to be toxic to *D. simulans*. Even more surprisingly, the *majority* of the males with *D. simulans* morphology were found on *M. citrifolia* fruits (52% of *D. simulans* males). The remainder of the *D. simulans* males was found at similar frequencies breeding on other fruits. The relative frequency of *D. simulans* found on different fruits split by island is shown in Table S6.

This survey also revealed the presence of F_1 hybrid males on *M. citrifolia* fruits. A substantial proportion of

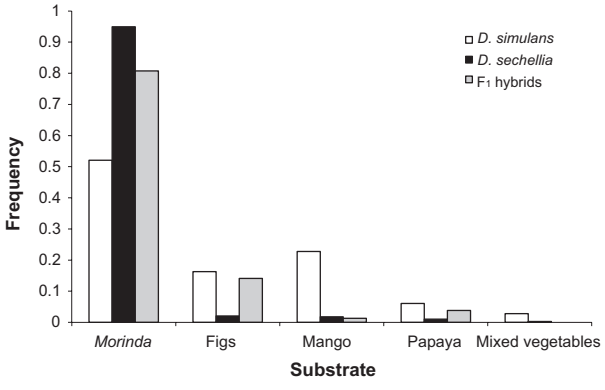

Fig. 4 Proportion of drosophilid males found in each substrate with emphasis on putative *Drosophila simulans*, putative *D. sechellia* and their hybrids. We show the frequency of the three genotypic classes and when individuals from all islands are pooled together.

the males with hybrid morphology were found on *M. citrifolia* fruits (~80%) over any other type of fruit. This result holds for all the three islands where F_1 males were found; the majority of them were found in *M. citrifolia* fruits regardless of what other substrates were available on the island. Together these results suggest that *M. citrifolia* is not an exclusive niche to *D. sechellia*, and that some *D. simulans* and interspecific hybrids are found in the *M. citrifolia* fruit.

Lastly, even though the majority of putative *D. sechellia* males were found on *M. citrifolia* fruits (~94%), a few males were netted on other substrates indicating that *D. sechellia* can be found on other fruits besides *M. citrifolia*. This observation was true for the five islands where we found *D. sechellia*. In Marianne island (a relatively small sample, 14 males), *D. sechellia* was found more frequently on figs than in *M. citrifolia* in this island (even though *M. citrifolia* is abundant in Marianne island). Figure S6 shows the breakdown of found species on each island.

Discussion

The creative role of hybridization in speciation remains hotly debated (recently reviewed in Abbott *et al.*, 2013). On the one hand, gene flow between species could swamp their divergence and ultimately lead to the collapse of the species boundary. On the other hand, hybridization could propel the evolution of new species. The best understood mechanism through which this happens is reinforcement, by which selection against maladaptive hybridization – specifically, costs to females who waste gametes on interspecific matings – leads to the enhancement of prezygotic isolation (Dobzhansky, 1937; Servedio & Noor, 2003; Hopkins, 2013). A more contentious mechanism by which hybridization could advance the process of speciation is via hybrid speciation, when two parental species give rise to a sexually reproducing hybrid lineage that is reproductively isolated from either parental species and is ecologically and genetically stable over time (Arnold, 2006). Speciation by hybridization occurs in plants where changes in ploidy are relatively common, but the frequency of this mode of speciation in animals is unclear (Coyne & Orr, 2004; Arnold, 2006; Brower, 2013).

Hybridization is actively occurring in the Seychelles between *D. simulans* and *D. sechellia*. Hybrids were found on all three islands where *D. simulans* was found. This does not constitute a hybrid zone in the classical sense, with a smooth continuum in the relative abundance of the two parental species and a transition area of overlap in a geographical window where hybrids are found. Future field studies will be needed to address the question of the geographical structure of hybridization, and whether there is a stable hybrid zone (in terms of time and geographical location) between *D. simulans* and *D. sechellia*.

Our approach likely underestimates the extent of hybridization, for two main reasons. First, we only identified individuals with clear hybrid morphology (likely F_1 hybrids) and likely missed backcrosses and advanced intercrosses between the two species that resemble the pure species in the phenotypic traits considered (Macdonald & Goldstein, 1999; see above). Second, we may have incorrectly assigned individuals to species if the genes underlying the phenotypes used for classification can introgress between species. Our molecular typing has similar limitations. First, it is possible that the putative F_1s are backcrossed individuals (or even further intercrosses). For example, we will correctly identify as hybrids individuals from the backcross [♀ (♀ *D. simulans* × ♂ *D. sechellia*) × ♂ *D. sechellia*] because they have a *D. simulans* mitochondria and a *D. sechellia Y*-chromosome. However, individuals from the backcross between the same mother and the other parental species [♀ (♀ *D. simulans* × ♂ *D. sechellia*) × ♂ *D. simulans*] will have both mitochondrial and *Y*-chromosome elements from *D. simulans* and will not be identified as a hybrid. Both the morphology-based identification and the molecular typing underestimate the number of hybrid males collected, so these results provide a conservative estimate of current hybridization between *D. simulans* and *D. sechellia*. Notably, the hybrids collected in the field come from the direction predicted by the strength of reproductive isolation in lab hybridizations: *D. simulans* females accept *D. sechellia* males much more easily than *D. sechellia* females accept *D. simulans* males, indicating that hybridization is more likely to occur through *D. simulans* females (Coyne & Kreitman, 1986; Lachaise *et al.*, 2004). Our ongoing work uses a genomic approach to estimate the levels of admixture between these two species in the Seychelles islands.

Oddly, we found putative *D. simulans* on *M. citrifolia* on every island where the species was collected. Most *D. simulans* lines have been reported to be sensitive to the octanoic and hexanoic acids in this toxic fruit and show a modest response to selection for resistance, indicating low levels of genetic variance for this trait (Colson, 2004). Equally puzzling is the fact that putative *D. sechellia* can be collected on a variety of fruits, since it had been previously proposed that *D. sechellia* is an obligate specialist of *M. citrifolia*. The fact that these males were found on other fruits does not prove that *D. sechellia* breeds in other fruits but it strongly suggests that *D. sechellia* males at least wander and explore other fruits. There are three possible explanations for the presence of flies of one species in the niche of the other. The first one is that some of the individuals classified as one species carry genetic elements from the other that allow them to explore (or even breed) in the other niche. The second one is that the flies only breed in their known niche, but males explore a wide variety of resources. A third possibility is that a small fraction

of individuals have independently evolved the ability to breed on other fruits. Regardless of the reason, as males of both species, putative *D. simulans* and putative *D. sechellia*, can be found in each other's niche, the possibility of hybridization is high, especially if this behaviour applies to females as well.

Our results also suggest that the distribution of *D. simulans* and *D. sechellia* in the Seychelles archipelago has changed in the last 25 years. Several past collections have revealed the presence of *D. simulans* on the central and larger islands of the Seychelles archipelago, namely Mahé and Silhouette (Lachaise *et al.*, 1988; Mercot *et al.*, 1995; Poinsot *et al.*, 2000). *D. sechellia* was restricted to the smaller, outer islands (Lachaise *et al.*, 2004; Legrand *et al.*, 2009). Additionally, *D. simulans* was usually found at higher elevations than *D. sechellia* (Lachaise *et al.*, 2004; Capy *et al.* 2004). Our report shows that currently *D. simulans* is widespread in three of the five sampled islands and is not restricted to the high elevations of Mahé. *D. sechellia*, on the other hand, was abundant on all the sampled islands. These results suggest that the ranges of the two species might have expanded in the last 25 years thus leading to secondary contact.

The importance of finding hybrid individuals in nature is two-fold. First, *D. simulans* and *D. sechellia* have been instrumental in the identification of alleles involved in hybrid sterility (Cabot *et al.*, 1994; Hollocher & Wu, 1996), meiotic drive (Johnson & Wu, 1992; Dermitzakis *et al.*, 2000; Tao *et al.*, 2007; Bayes & Malik, 2009), and in the mapping of loci involved in ecological divergence (Jones, 2005; Matsuo *et al.*, 2007; McBride, 2007; Dworkin & Jones, 2009). They have also been studied for the possible role that introgression can have during the speciation process and how gene exchange can proceed during early species divergence (Palopoli *et al.*, 1996; Garrigan *et al.*, 2012; Brand *et al.*, 2013). Nonetheless, this report is the first instance describing that *D. simulans* and *D. sechellia* hybridize in their natural environment, suggesting that reproductive isolation might not be complete in some locations. These results coupled with analyses of limited genetic data indicate that *D. simulans* and *D. sechellia* might currently be exchanging genes (R'Kha *et al.*, 1991; Garrigan *et al.*, 2012) but that such gene exchange might be might not be uniform across the genome (Fang *et al.*, 2012).

The second important contribution of natural hybrids pertains to the possibility of using these hybrids to dissect the genetic basis of interspecific phenotypic differences (Rieseberg & Buerkle, 2002). Currently, most of the mapping strategies in *Drosophila* rely on the generation of backcrosses and thus depend on the generation of recombination breakpoints in laboratory populations. The existence of hybrids in the field could facilitate genetic mapping by allowing the application of alternative mapping techniques, such as admixture mapping (Winkler *et al.*, 2010).

In conclusion, we have added a case of naturally occurring hybrids to the short list of known hybrid zones in *Drosophila*. We have also illuminated the natural history of the hybridizing species, *D. simulans* and *D. sechellia*, and discovered novel behaviours for these organisms that have been keenly observed in the laboratory. This new hybrid zone and the behavioural traits that exist within it are ripe for future study. Classical studies have used hybrid zones to estimate the number of genes involved in hybrid breakdown, to identify what segments of the genome can cross the species boundary, and to study the natural forces that affect hybrids and pure species alike (Gay *et al.*, 2008; Sweigart, 2009; Guedj & Guillot, 2011). With the advent of genomic data, it is possible to identify tracts of admixture, genomic regions that have been exchanged, and ultimately the genetic basis of interspecific differences. The existence of natural hybrids between *D. simulans* and *D. sechellia* opens the possibility of these kinds of studies.

Acknowledgments

We would like to acknowledge the Seychelles Bureau of Standard for their assistance in obtaining the collection permits, especially to Elvis Nicette and Ronley Fanchette. We are also grateful for Pierre-André Adam's support at the Island Conservation Society for his assistance in obtaining the collection permits. We would like to thank K. Mason, R. M. Mason, M. Ultsch-Unrath and James Chang-Tave for their generosity and for granting us permission to collect in Denis Island. We would like to thank K. L. Gordon, M. F. Przeworski, D. L. Hartl, A. G. Clark, N. Johnson and the Przeworski lab for scientific discussions and comments at every stage of the manuscript. D.R.M is funded by a Chicago fellowship. J.F.A is funded by the Harvard Society of Fellows.

References

Abbott, R., Albach, D., Ansell, S., Arntzen, J.W., Baird, S.J.E., Bierne, N. *et al.* 2013. Hybridization and speciation. *J. Evol. Biol.* **26**: 229–246.

Amlou, M., Moreteau, B. & David, J.R. 1998. Genetic analysis of *Drosophila sechellia* specialization: oviposition behavior toward the major aliphatic acids of its host plant. *Behav. Genet.* **28**: 455–464.

Arnold, M.L. 2006. *Evolution through Genetic Exchange*. Oxford University Press, Oxford; New York.

Arnold, M.L. & Martin, N.H. 2009. Adaptation by introgression. *J. Biol.* **8**: 82.

Bachtrog, D., Thornton, K., Clark, A. & Andolfatto, P. 2006. Extensive introgression of mitochondrial DNA relative to nuclear genes in the *Drosophila yakuba* species group. *Evolution* **60**: 292–302.

Bank, C., Bürger, R. & Hermisson, J. 2012. The Limits to parapatric speciation: Dobzhansky-Muller incompatibilities in a continent-Island model. *Genetics* **191**: 845–863.

Bayes, J.J. & Malik, H.S. 2009. Altered heterochromatin binding by a hybrid sterility protein in *Drosophila* sibling species. *Science* **326**: 1538–1541.

Bedichek-Pipkin, S. 1968. Introgression between closely related species of *Drosophila* in Panama. *Evolution* **22**: 140–156.

Bock, I.R. 1980. Current status of the *Drosophila melanogaster* species group (*Diptera*). *Syst. Entomol.* **5**: 341–356.

Brand, C.L., Kingan, S.B., Wu, L. & Garrigan, D. 2013. A selective sweep across species boundaries in *Drosophila*. *Mol. Biol. Evol.* **30**: 2177–2186.

Brower, A.V. 2013. Introgression of wing pattern alleles and speciation via homoploid hybridization in *Heliconius* butterflies: a review of evidence from the genome. *Proc. R. Soc. B* **280**: 20122302.

Cabot, E.L., Davis, A.W., Johnson, N.A. & Wu, C.-I. 1994. Genetics of reproductive isolation in the *Drosophila simulans* clade: complex epistasis underlying hybrid male sterility. *Genetics* **137**: 175–189.

Capy, P. & Gibert, P. 2004. Drosophila melanogaster, Drosophila simulans: so similar yet so different. In: *Drosophila melanogaster, Drosophila simulans: So Similar, So Different* (P. Capy, P. Gibert & I. Boussy, eds), pp. 5–16. Springer, Dordrecht, Netherlands.

Cheke, A. 2010. The timing of arrival of humans and their commensal animals on Western Indian Ocean oceanic islands. *Phelsuma* **18**: 38–69.

Clark, A.G., Eisen, M.B., Smith, D.R., Bergman, C.M., Oliver, B., Markow, T.A. *et al.* 2007. Evolution of genes and genomes on the *Drosophila* phylogeny. *Nature* **450**: 203–218.

Cobb, M., Burnet, B., Blizard, R. & Jallon, J.-M. 1989. Courtship in *Drosophila sechellia*: its structure, functional aspects, and relationship to those of other members of the *Drosophila melanogaster* species subgroup. *J. Insect Behav.* **2**: 63–89.

Colson, I. 2004. *Drosophila simulans*' response to laboratory selection for tolerance to a toxic food source used by its sister species *D. sechellia*. *Evol. Ecol.* **18**: 15–28.

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

Coyne, J.A. 1996. Genetics of a difference in male cuticular hydrocarbons between two sibling species, *Drosophila simulans* and *D. sechellia*. *Genetics* **143**: 1689–1698.

Coyne, J.A. & Kreitman, M. 1986. Evolutionary genetics of two sibling species, *Drosophila simulans* and *D. sechellia*. *Evolution* **40**: 673–691.

Coyne, J.A. & Orr, H.A. 1989. Patterns of speciation in *Drosophila*. *Evolution* **43**: 362–381.

Coyne, J.A. & Orr, H.A. 2004. *Speciation*. Sinauer Associates, Sunderland, MA. 545 pp.

David, J.R. & Capy, P. 1982. Genetics and origin of a *Drosophila melanogaster* population recently introduced to the Seychelles. *Genetics Res.* **40**: 295–303.

David, J.R., Lemeunier, F., Tsacas, L. & Yassin, A. 2007. The historical discovery of the nine species in the *Drosophila melanogaster* species subgroup. *Genetics* **177**: 1969–1973.

Dermitzakis, E.T., Masly, J.P., Waldrip, H.M. & Clark, A.G. 2000. Non-Mendelian segregation of sex chromosomes in heterospecific *Drosophila* males. *Genetics* **154**: 687–694.

Dobzhansky, T.G. 1937. *Genetics and the Origin of Species*. Columbia University Press, New York.

Dworkin, I. & Jones, C.D. 2009. Genetic changes accompanying the evolution of host specialization in *Drosophila sechellia*. *Genetics* **181**: 721–736.

Erezyilmaz, D.F. & Stern, D.L. 2013. Pupariation site preference within and between *Drosophila* sibling species. *Evolution* **67**: 2714–2727.

Fang, S., Yukilevich, R., Chen, Y., Turissini, D.A., Zeng, K., Boussy, I.A. *et al.* 2012. Incompatibility and competitive exclusion of genomic segments between sibling *Drosophila* species. *PLoS Genet.* **8**: e1002795.

Farine, J.-P., Legal, L., Moreteau, B. & Le Quere, J.-L. 1996. Volatile components of ripe Fruits of *Morinda citrifolia* and their effects on *Drosophila*. *Phytochemistry* **41**: 433–438.

Garrigan, D., Kingan, S.B., Geneva, A.J., Andolfatto, P., Clark, A.G., Thornton, K. *et al.* 2012. Genome sequencing reveals complex speciation in the *Drosophila simulans subcomplex*. *Genome Res.* **22**: 1499–1511.

Gay, L., Crochet, P.A., Bell, D.A. & Lenormand, T. 2008. Comparing clines on molecular and phenotypic traits in hybrid zones: a window on tension zone models. *Evolution* **62**: 2789–2806.

Gleason, J.M., Jallon, J.-M., Rouault, J.-D. & Ritchie, M.G. 2005. Quantitative trait loci for cuticular hydrocarbons associated with sexual isolation between *Drosophila simulans* and *D. sechellia*. *Genetics* **171**: 1789–1798.

Gleason, J.M., James, R.A., Wicker-Thomas, C. & Ritchie, M.G. 2009. Identification of quantitative trait loci function through analysis of multiple cuticular hydrocarbons differing between *Drosophila simulans* and *Drosophila sechellia* females. *Heredity* **103**: 416–424.

Guedj, B. & Guillot, G. 2011. Estimating the location and shape of hybrid zones. *Mol. Ecol. Resour.* **11**: 1119–1123.

Hedrick, P.W. 2013. Adaptive introgression in animals: examples and comparison to new mutation and standing variation as sources of adaptive variation. *Mol. Ecol.* **22**: 4606–4618.

Herrig, D.K., Modrick, A.J., Brud, E. & Llopart, A. 2013. Introgression in the *Drosophila subobscura—D. madeirensis* sister species: evidence of gene flow in nuclear genes despite mitochondrial differentiation. *Evolution* **68**: 705–719.

Higa, I. & Fuyama, Y. 1993. Genetics of food preference in *Drosophila sechellia* I. Responses to food attractants. *Genetica* **88**: 129–136.

Hollocher, H. & Wu, C.-I. 1996. The genetics of reproductive isolation in the *Drosophila simulans* subcomplex: *X* vs. autosomal effects and male vs. female effects. *Genetics* **143**: 1243–1255.

Hopkins, R. 2013. Reinforcement in plants. *New Phytol.* **197**: 1095–1103.

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

Hungate, E.A., Earley, E.J., Boussy, I.A., Turissini, D.A., Ting, C.T., Moran, J.R. *et al.* 2013. A locus in *Drosophila sechellia* affecting tolerance of a host plant toxin. *Genetics* **195**: 1063–1075.

Jaenike, J., Dyer, K.A., Cornish, C. & Minhas, M.S. 2006. Asymmetrical reinforcement and *Wolbachia* infection in *Drosophila*. *PLoS Biol.* **4**: e325.

Johnson, N.A. & Wu, C.-I. 1992. An empirical test of the meiotic drive models of hybrid sterility: sex-ratio data from hybrids between *Drosophila simulans* and *Drosophila sechellia*. *Genetics* **130**: 507–511.

Jones, C.D. 2005. The genetics of adaptation in *Drosophila sechellia*. *Genetica* **123**: 137–145.

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

Kopp, A. 2006. Basal relationships in the *Drosophila melanogaster* species group. *Mol. Phylogenet. Evol.* **39**: 787–798.

Lachaise, D. & Silvain, J.F. 2004. How two Afrotropical endemics made two cosmopolitan human commensals: the *Drosophila melanogaster-D. simulans* palaeogeographic riddle. In *Drosophila melanogaster*, *Drosophila simulans*: so similar, so different. *Genetica* **20**: 17–39.

Lachaise, D., David, J.R., Lemeunier, F., Tsacas, L. & Ashburner, M. 1986. The reproductive relationships of *Drosophila sechellia* with *D. mauritiana*, *D. simulans*, and *D. melanogaster* from the Afrotropical region. *Evolution* **40**: 262–271.

Lachaise, D., Cariou, M.L., David, J.R., Lemeunier, F., Tsacas, L. & Ashburner, M. 1988. Historical biogeography of the *Drosophila melanogaster* species subgroup. *Evol. Ecol.* **22**: 159–225.

Lachaise, D., Harry, M., Solignac, M., Lemeunier, F., Bénassi, V. & Cariou, M.L. 2000. Evolutionary novelties in islands: *Drosophila santomea*, a new *melanogaster* sister species from São Tomé. *Proc. R. Soc. B* **267**: 1487–1495.

Lachaise, D., Capy, P., Cariou, M.L., Joly, D.O., Lemeunier, F. & David, J.R. 2004. Nine relatives from one African ancestor: population biology and evolution of the *Drosophila melanogaster* subgroup species. In: *The Evolution of Population Biology* (R.S. Singh & M.K. Uyenoyama, eds), pp. 315–343. Cambridge University Press, Cambridge, U.K.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H. *et al.* 2007. Clustal W and Clustal X version 2.0. *Bioinformatics* **23**: 2947–2948.

Larracuente, A.M. & Clark, A.G. 2013. Surprising differences in the variability of *Y*-chromosomes in African and cosmopolitan populations of *Drosophila melanogaster*. *Genetics* **193**: 201–214.

Legal, L., Chappe, B. & Jallon, J.-M. 1994. Molecular basis of *Morinda citrifolia* (L.): toxicity on *Drosophila*. *J. Chem. Ecol.* **20**: 1931–1943.

Legrand, D., Tenaillon, M.I., Matyot, P., Gerlach, J., Lachaise, D. & Cariou, M.L. 2009. Species-wide genetic variation and demographic history of *Drosophila sechellia*, a species lacking population structure. *Genetics* **182**: 1197–1206.

Lemeunier, F., David, J.R., Tsacas, L. & Ashburner, M. 1986. The melanogaster species group. In: *The Genetics and Biology of Drosophila*. (M. Ashburner, H.L. Carson & J.N. Thompson Jr, eds), pp. 147–256. New York Academic Press, New York.

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

Llopart, A., Lachaise, D. & Coyne, J.A. 2005b. Multilocus analysis of introgression between two sympatric sister species of *Drosophila*: *Drosophila yakuba* and *D. santomea*. *Genetics* **171**: 197–210.

Llopart, A., Herrig, D., Brud, E. & Stecklein, Z. 2014. Sequential adaptive introgression of the mitochondrial genome in *Drosophila yakuba* and *D. santomea*. *Mol. Ecol.* **23**: 1124–1236.

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

Macdonald, S.J. & Goldstein, D.B. 1999. A quantitative genetic analysis of male sexual traits distinguishing the sibling species *Drosophila simulans* and *D. sechellia*. *Genetics* **153**: 1683–1699.

Mahalanobis, P.C. 1936. On the generalized distance in statistics. *Proc. Nat. Inst. Sci. India* **2**: 49–55.

Markow, T.A. & O'Grady, P. 2005. *Drosophila: A Guide to Species Identification and Use*. Academic Press (Elsevier), London.

Matsuo, T., Sugaya, S., Yasukawa, J., Aigaki, T. & Fuyama, Y. 2007. Odorant-binding proteins *OBP57d* and *OBP57e* affect taste perception and host-plant preference in *Drosophila sechellia*. *PLoS Biol.* **5**: e118.

Matute, D.R. 2010. Reinforcement of gametic isolation in *Drosophila*. *PLoS Biol.* **8**: e1000341.

McBride, C.S. 2007. Rapid evolution of smell and taste receptor genes during host specialization in *Drosophila sechellia*. *PNAS* **104**: 4996–5001.

McDermott, S.R. & Kliman, R.M. 2008. Estimation of isolation times of the island species in the *Drosophila simulans* complex from multilocus DNA sequence data. *PLoS ONE* **3**: e2442.

Mercot, H., Llorente, B., Jacques, M., Atlan, A. & Montchamp-Moreau, C. 1995. Variability within the Seychelles cytoplasmic incompatibility system in *Drosophila simulans*. *Genetics* **141**: 1015.

Nunes, M.D., Orozco-Ter Wengel, P., Kreissl, M. & Schlötterer, C. 2010. Multiple hybridization events between *Drosophila simulans* and *D. mauritiana* are supported by mtDNA introgression. *Mol. Ecol.* **19**: 4695–4707.

Orgogozo, V. & Stern, D.L. 2009. How different are recently diverged species?: more than 150 phenotypic differences have been reported for the *Drosophila melanogaster* species subgroup. *Fly* **3**: 117.

Palopoli, M.F., Davis, A.W. & Wu, C.-I. 1996. Discord between the phylogenies inferred from molecular vs. functional data: uneven rates of functional evolution or low levels of gene flow? *Genetics* **144**: 1321–1328.

Pardo-Diaz, C., Salazar, C., Baxter, S.W., Merot, C., Figueiredo-Ready, W., Joron, M. *et al.* 2012. Adaptive introgression across species boundaries in *Heliconius* butterflies. *PLoS Genet.* **8**: e1002752.

Pinto, F.M., Brehm, A., Hernandez, M., Larruga, J.M., González, A.M. & Cabrera, V.M. 1997. Population genetic structure and colonization sequence of *Drosophila subobscura* in the Canaries and Madeira Atlantic islands as inferred by autosomal, sex-linked and mtDNA traits. *J. Hered.* **88**: 108–114.

Poinsot, D., Montchamp-Moreau, C. & Merçot, H. 2000. *Wolbachia* segregation rate in *Drosophila simulans* naturally bi-infected cytoplasmic lineages. *Heredity* **85**: 191–198.

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

Rieseberg, L.H. & Buerkle, C.A. 2002. Genetic mapping in hybrid zones. *Am. Nat.* **159**: S36–S50.

R'Kha, S., Capy, P. & David, J.R. 1991. Host-plant specialization in the *Drosophila melanogaster* species complex: a physiological, behavioral, and genetical analysis. *PNAS* **88**: 1835–1839.

Schneider, C.A., Rasband, W.S. & Eliceiri, K.W. 2012. NIH image to ImageJ: 25 years of image analysis. *Nat. Methods* **9**: 671–675.

Servedio, M.R. & Noor, M.A. 2003. The role of reinforcement in speciation: theory and data. *Annu. Rev. Ecol. Evol. Syst.* **34**: 339–364.

Solignac, M., Monnerot, M. & Mounolou, J.C. 1986. Mitochondrial DNA evolution in the *melanogaster* species subgroup of *Drosophila*. *J. Mol. Evol.* **23**: 31–40.

Song, Y., Endepols, S., Klemann, N., Richter, D., Matuschka, F.R., Shih, C.H. *et al.* 2011. Adaptive introgression of anticoagulant rodent poison resistance by hybridization between old world mice. *Curr. Biol.* **21**: 1296–1301.

Sweigart, A. 2009. Sculpin hybrid zones: natural laboratories for the early stages of speciation. *Mol. Ecol.* **18**: 2547–2548.

Tao, Y., Araripe, L., Kingan, S.B., Ke, Y., Xiao, H. & Hartl, D.L. 2007. A sex-ratio meiotic drive system in *Drosophila simulans*. II: an X-linked distorter. *PLoS Biol.* **5**: e293.

Tomaru, M., Yamada, H. & Oguma, Y. 2004. Female mate recognition and sexual isolation depending on courtship song in *Drosophila sechellia* and its siblings. *Genes Genet. Syst.* **79**: 145–150.

Torsvik, T.H., Amundsen, H., Hartz, E.H., Corfu, F., Kusznir, N., Gaina, C. *et al.* 2013. A Precambrian microcontinent in the Indian Ocean. *Nat. Geosci.* **6**: 223–227.

Tsacas, L. & Bächli, G. 1981. *Drosophila sechellia*, N. Sp., huitième espèce du sous-groupe *melanogaster* des iles Seychelles (*Diptera*, *Drosophilidae*). *Rev. Fr. Entomol.* **3**: 146–150.

Winkler, C.A., Nelson, G.W. & Smith, M.W. 2010. Admixture mapping comes of age. *Annu. Rev. Genomics Hum. Genet.* **11**: 65–89.

Yukilevich, R. 2012. Asymmetrical patterns of speciation uniquely support reinforcement in *Drosophila*. *Evolution* **66**: 1430–1446.

Supporting information

Additional Supporting Information may be found in the online version of this article:

Figure S1 Backcrossed individuals show a high degree of phenotypic variance in their morphological traits.

Figure S2 The posterior genital lobe area shows no heterogeneity in *D. simulans* or *D. sechellia* intraspecific crosses. The identity of each cross was removed from the *x*-axis for clarity.

Figure S3 The number of teeth in sex combs shows no heterogeneity in *D. simulans* or *D. sechellia* intraspecific crosses.

Figure S4 Anterior tibial length shows no heterogeneity in *D. simulans* or *D. sechellia* intraspecific crosses.

Figure S5 Scatter plots of pairwise combinations of the three phenotypic traits assessed in this study for 16 000 laboratory-produced individuals.

Figure S6 Proportion of males from three broadly defined taxonomic groups.

Table S1 Geographical location of the ten collection sites in five different islands.

Table S2 *D. simulans*, *D. sechellia*, and the F_1s between these two species differ significantly in their phenotypic values in three different phenotypic traits.

Table S3 Mahalanobis distances (D^2) for each laboratory-produced individual when the assignment was carried out with eight genotypic groups.

Table S4 Mahalanobis distances (D^2) for each laboratory-produced individual when the assignment was performed with three genotypic groups.

Table S5 Isofemale lines of *D. simulans* and *D. sechellia* used for the morphological characterization and laboratory crosses.

Table S6 Mahalanobis distances (D^2) for the 2453 field-collected individuals to the three phenotypic classes (*D. simulans*, *D. sechellia* and F_1 hybrids).

Data deposited at Dryad: doi:10.5061/dryad.1c88s

Received 16 December 2013; revised 12 February 2014; accepted 20 March 2014