

## ORIGINAL ARTICLE

Genetic variation tracks ecological segregation  
in Pacific island black fliesDA Joy<sup>1</sup>, DA Craig<sup>2</sup> and JE Conn<sup>3</sup><sup>1</sup>Department of Biology, University of Vermont, Burlington, VT, USA; <sup>2</sup>Department of Biological Sciences, University of Alberta, Edmonton, Alberta, Canada and <sup>3</sup>Griffin Laboratory, The Wadsworth Center, New York State Department of Health, Slingerlands, NY, USA

Geographic isolation is widely viewed as a key component of insular radiations on islands. However, strong ecological affinities may also reinforce isolation and promote genetic divergence. The black fly fauna in the Society Islands French Polynesia is notable for the number of closely related endemic species (31), and the morphological and habitat diversity of the larvae. Here, we measure ecological and morphological differences within and between two closely related species, *Simulium oviceps* and *Simulium dussertorum* and relate these differences to genetic distance. Phylogenetic analyses of a 920 bp fragment of the cyto-

chrome oxidase I (*COI*) gene revealed a well-supported, ecologically divergent *S. oviceps* clade (larvae found in rivers instead of cascades) that shows little morphological differentiation. For both *S. oviceps* and *S. dussertorum*, genetic distance among populations is related to larval habitat, with cascade populations showing greater isolation from each other than river populations. Our data support the hypothesis that larval ecological shifts have played a role in the radiation of this black fly fauna.

*Heredity* (2007) **99**, 452–459; doi:10.1038/sj.hdy.6801023; published online 4 July 2007

**Keywords:** insular radiation; habitat shift; mitochondrial DNA; simuliidae; island biogeography

## Introduction

Species radiations are a common feature of isolated islands, which provide natural laboratories for the study of speciation (Williamson, 1981). Inter-island colonization is widely viewed as a central component of island radiations (Carson and Templeton, 1984) although populations within islands may also become isolated from one another if dispersal is limited (Cowie, 1995; Shaw, 1996). Additionally, strong ecological affinities can reinforce isolation even in the absence of geographical distance (Kambyzellis *et al.*, 1995; Gillespie, 2004). Exploitation of novel resources and competition for shared resources are among the ecological causes of speciation expected to be especially important during a radiation because of the coexistence of many closely related species (Losos, 1994; Schluter, 1996).

*Simulium oviceps* and *Simulium dussertorum* are members of a larger radiation of simuliids within the Society Islands, French Polynesia (Craig *et al.*, 2001). The Society Islands, like the Hawaiian Islands, are a hot spot archipelago: a linear array of progressively younger islands formed by a submantle plume (Duncan and McDougall, 1976; Carson and Clague, 1995). Both *S. oviceps* and *S. dussertorum* are found on the youngest island, Tahiti, while *S. oviceps* is also found on Moorea (Figure 1), and an older island, Raiatea. The two species are sister taxa and are morphologically

similar, both showing a reduction in labral fan size and the number of fan rays, a specialized feeding adaptation in the larvae (Craig, 1977). Before the application of molecular techniques they were difficult to tell apart and Craig (1987) suggested, based on morphological variation, that *S. oviceps* might be a species complex, and in fact *S. dussertorum* was initially thought to be an environmental variant of *S. oviceps*, but was later described as a separate species (Craig, 1997).

A defining characteristic of black flies in French Polynesia is that larvae have colonized a diverse array of habitats but that most species have narrow habitat requirements – with the larvae of each species found in one or a few similar habitats (Craig *et al.*, 2001; Joy and Conn, 2001). In contrast, *S. oviceps* larvae have a relatively broad ecological niche: cascades and large rivers. The two habitats differ markedly from each other in terms of nutrient content and flow regime. Cascades in the Society Islands have a higher concentration of particulate matter, and a higher proportion of diatoms than rivers (Resh *et al.*, 1990), and diatoms are a high-quality food source for black fly larvae (Thompson, 1987). Unlike large rivers, water velocity can be highly variable among and even within individual cascades in the Society Islands (Craig *et al.*, 2001). Because of its seemingly broad habitat requirements, *S. oviceps* presents a unique opportunity to study the evolution of habitat specialization in the subgenus *Inseliellum*.

The objectives of this study were to reconstruct the molecular phylogeny of *S. oviceps* and *S. dussertorum* populations, to measure genetic and morphological variation across their ranges, and relate these differences to ecological and geographical isolation to gain insight into possible mechanisms underlying the Society Island black fly radiation. In addition, we discuss the feasibility

Correspondence: Dr DA Joy, NIH/NIAID/Laboratory of Malaria and Vector Research, 12735 Twinbrook Parkway, Room 3E-20B, Rockville, MD 20850, USA.

E-mail: djoy@niaid.nih.gov

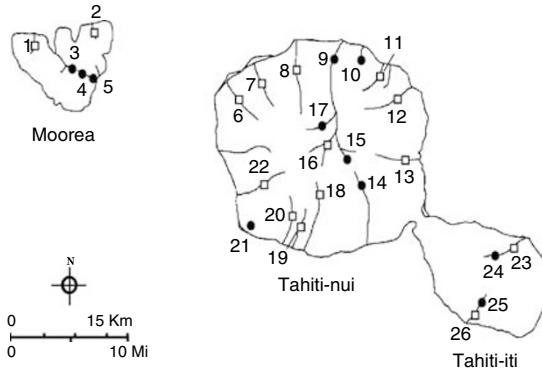
Received 18 May 2006; revised 1 May 2007; accepted 26 May 2007; published online 4 July 2007

of applying nested clade analysis (NCA) to a system with a known but highly complex geological history.

## Materials and methods

### Type specimens and sampling localities

The *Simulium (Inseliellum) oviceps* Edwards 1933 holotype is deposited at the Entomology Department, Natural



**Figure 1** Sampling locations for *S. oviceps* and *S. dussertorum* on Moorea and Tahiti. Corresponding location names are given in Table 1. Open square, river; closed circle, cascade.

History Museum, (Cromwell Road, London, SW7 5BD, England). The *Simulium (Inseliellum) dussertorum* Craig, 1997 holotype is deposited at the Department of Entomology, Bernice P Bishop Museum (1525 Bernice Street, Honolulu, Hawai'i, 96817–2704). A total of 339 *S. oviceps* and 94 *S. dussertorum* larvae were collected from 26 sites (15 rivers and 11 cascades) on Tahiti and Moorea (Figure 1; Table 1). *S. oviceps* was found on both islands while *S. dussertorum* was found only on Tahiti. The *S. dussertorum* type locality is Grottes de Maara (Craig, 1997). As only a single cytochrome oxidase I (COI) haplotype was collected from this location (haplotype A), we designated the clade containing this haplotype as *S. dussertorum*. *Simulium bogusium* and *Simulium proctorae* were chosen as outgroups because they are closely related to the two focal species (Joy and Conn, 2001) and are endemic to Raiatea, an older island.

### DNA extraction and amplification

Genomic DNA was isolated from frozen (−80°C) larvae following the protocol of Collins *et al.* (1987). Initial COI sequences were obtained with primers UEA5 and UEA10 (Lunt *et al.*, 1996) and were used to design *Simulium*-specific primers. These primers amplified a 920 bp fragment from the 5'-end of the COI gene. PCR parameters consisted of initial denaturing at 94°C

**Table 1** Collection sites, PCR-D haplotypes and sample sizes for PCR-D and sequence-based analyses

Locality name	Code	Location	Site type	Species	PCR-D Haplotypes	N PCR-D	N Sequence
Terahimaua	1	S17° 32' 25.8" W149° 49' 37.0"	r	ovi	D	4	4
Papeahi	2	S17° 29' 16.3" W149° 47' 56.7"	r	ovi	D	5	5
Belvedere	3	S17° 32' 25.8" W149° 49' 37.0"	c	ovi	E	10	4
Afareaitu, upper	4	S17° 32' 39.2" W149° 48' 48.2"	c	ovi	F,F3,M	15	5
Afareaitu, lower	5	S17° 32' 37.6" W149° 48' 29.7"	c	ovi	E,F	13	6
Tipaerui	6	S17° 33' 00.1" W149° 34' 37.1"	r	ovi	C,D	8	4
Fautaua	7	S17° 33' 22.0" W149° 32' 47.2"	r	dus	A	11	5
				ovi	C,D	17	4
				dus	A	3	1
Tuauru	8	S17° 31' 51.1" W149° 29' 20.8"	r	ovi	C,D,L2	19	4
Papanoo	9	S17° 33' 16.0" W 149° 25' 54.6"	c	ovi	G	13	4
				dus	A	7	5
Trois Cascades	10	S17° 32' 06.7" W149° 24' 05.9"	c	ovi	F1,G	13	3
				dus	A,B	10	5
Onoheha	11	S17° 32' 20.3" W149° 22' 14.5"	r	ovi	C,D,L1	22	4
Mahape	12	S17° 33' 55.9" W149° 19' 56.8"	r	ovi	C,D,L1	18	4
Utuofoai	13	S17° 39' 47.1" W149° 19' 02.0"	r	ovi	C,D,L1,L3	20	5
Lake Vaihiria	14	S17° 40' 15.6" W149° 25' 19.8"	c	ovi	F1,G,R	11	5
				dus	N,Q	7	5
Col Urufaa	15	S17° 39' 09.6" W149° 25' 50.6"	c	ovi	F1	3	1
				dus	A,N	17	5
Relais Marae	16	S17° 38' 36.1" W149° 26' 09.7"	r	ovi	C,F,K,K2,	16	6
				dus	N	2	2
Vaitemanu	17	S17° 37' 25.8" W149° 26' 06.9"	c	ovi	G	6	4
				dus	N,Q	11	5
Taharuu	18	S17° 44' 20.5" W149 28' 29.2"	r	ovi	C	20	4
Papeiti	19	S17° 45' 11.1" W149 29' 10.2"	r	ovi	C,D	20	4
Tamarua	20	S17° 45' 10.5" W149 30' 06.8"	r	ovi	C,D	17	4
				dus	A	1	1
Grottes de Maraa	21	S17° 44' 25.5" W149° 33' 29.7"	c	dus	A	15	6
Vaiatiu	22	S17° 41' 07.9" W149° 33' 50.6"	r	ovi	C,D,L1,L4	20	4
Vaitepiha river	23	S17° 46' 42.6" W149° 10' 40.0"	r	ovi	C	19	4
Vaitepiha cascade	24	S17° 46' 36.8" W149° 10' 40.6"	c	ovi	P	1	1
				dus	J	2	2
Tirahi cascade	25	S17° 50' 50.7" W149° 14' 12.3"	c	ovi	F1,P	12	5
				dus	H	8	5
Tirahi river	26	S17° 50' 49.2" W149° 14' 18.3"	r	ovi	C	17	4

Abbreviations: c, cascade; dus, *Simulium dussertorum*; ovi, *Simulium oviceps*; PCR-D, PCR-digest; r, river.

(3 min), five cycles at 94°C (1 min), 53°C (1 min) and 72°C (1.5 min), followed by 30 cycles at 94°C (1 min), and a ramp from 50 to 72°C (1°C/6s). Double-stranded PCR products were purified with polyethylene glycol. Reactions for automated sequencing were prepared using the ABI Prism Dye Terminator DNA Sequencing Kit.

### Sequence analyses

Sequences were compiled, edited and initially aligned using Sequencher 4.0 (Gene Codes Corp., Ann Arbor, MI, USA). Aligned sequences were imported into PAUP\*, v4.0b10 (Swofford, 2003). Phylogenetic relationships were reconstructed using maximum parsimony (MP) and maximum likelihood (ML) after redundant sequences were removed. We used MP to generate initial topologies for estimating ML model parameters and to evaluate models. One of 36 most-parsimonious trees (heuristic search, 20 random stepwise addition replicates, tree bisection-reconnection (TBR) branch swapping) was arbitrarily chosen to assess the fit of 56 models using likelihood ratio tests as implemented in MODELTEST 3.06 (Posada and Crandall, 1998). MP bootstrap support was estimated from 500 replicates (step-wise addition, TBR branch swapping). The ML tree was estimated using successive approximation (step-wise addition, TBR branch swapping), and 100 bootstrap replicates.

We estimated the time to most recent common ancestor (TMRCA) for *S. oviceps*, *S. dussertorum*, and subsets within *S. oviceps* from sequence data, using the Bayesian coalescent framework in BEAST v1.4 (Drummond and Rambaut, 2006). The TrN + I +  $\Gamma$  model (Tamura and Nei, 1993) was chosen based on likelihood ratio tests (Posada and Crandall, 1998). We used a mutation rate of  $2.0 \times 10^{-8}$  nucleotide substitutions/site year, based on estimates for COI in dipterans, specifically Hawaiian *Drosophila* (DeSalle *et al.*, 1987), *Drosophila melanogaster* and *Drosophila simulans* (Satta *et al.*, 1987), and in several other arthropod species (Brower, 1994). For each TMRCA estimate we ran two independent Markov chain Monte Carlo (MCMC) chains of 20 000 000 steps each, preceded by a burn-in of 2 000 000 steps. Convergence of the chains was checked using the program tracer (Rambaut and Drummond, 2004), and results from the two chains were found to be in agreement and were combined.

### PCR-digest screening and NCA

In addition to sequencing, we performed PCR digest (PCRD) on the 920 bp COI fragment. We screened 43 restriction enzymes using Sequencher 4.0 (Gene Codes Corp.), and found that a combination of three – *Hind*III, *Dde*I and *Acc*I – produced unique fragment patterns for 14 groups also identified by sequencing. DNA from up to 20 individuals (average = 16) from each of 26 collection sites was cut with these enzymes and new fragment patterns, when encountered, were sequenced, allowing us to identify four more restriction haplotypes. To verify that identical fragment patterns from different localities were the product of the same restriction sites, we sequenced at least one individual per haplotype per site ( $n = 149$ ). In doing so, we identified three additional haplotypes. We subsequently re-screened all localities with *Hpa*II, which in combination with the other three restriction enzymes produced unique restriction patterns

for the three new haplotypes. In all, four restriction enzymes allowed us to identify 21 restriction haplotypes (Supplementary Table 1). Because our method was not exhaustive, it is possible that we missed some haplotypes; however, restriction haplotypes largely reflected the pattern of sequence variation while allowing us to greatly increase our sample size.

We constructed a PCRD haplotype network based on the information in Supplementary Table 1. Statistical parsimony was used to determine the number of connections considered parsimonious at the 95% confidence level (Templeton *et al.*, 1992). The null hypothesis of no geographical association was tested by treating location as a categorical variable for nested exact permutational tests using GeoDis (Posada *et al.*, 2000). Geographical distance was incorporated as clade distance ( $D_c$ ) and nested clade distance ( $D_n$ ) that measure a clade's geographical range, and its distribution relative to closely related clades, respectively (Templeton *et al.*, 1995). Observed  $D_c$  and  $D_n$  values were compared to a distribution of randomly permuted values, which simulate the null hypothesis of no geographical association (Posada *et al.*, 2000). In cases where the null hypothesis was rejected three possibilities were considered: restricted gene flow, past fragmentation and range expansion. To distinguish among patterns, we used the inference key found at [http://zoology.byu.edu/crandall\\_lab/geodis.htm](http://zoology.byu.edu/crandall_lab/geodis.htm), which is a modification of the original key (Templeton *et al.*, 1995). Outgroup weights were also calculated for each PCRD haplotype using a heuristic procedure that takes into account haplotype connectedness (tip vs interior position) and frequency in the sample (Castelloe and Templeton, 1994).

### Population structure

Population structure was assessed by partitioning genetic diversity into hierarchical components using analysis of molecular variance (AMOVA) (Excoffier *et al.*, 1992). Populations were grouped according to larval habitat (river or cascade), and location with reference to putative geographic barriers to gene flow – the open ocean separating Moorea and Tahiti, and a narrow isthmus connecting Tahiti-nui and Tahiti-iti (Figure 1).

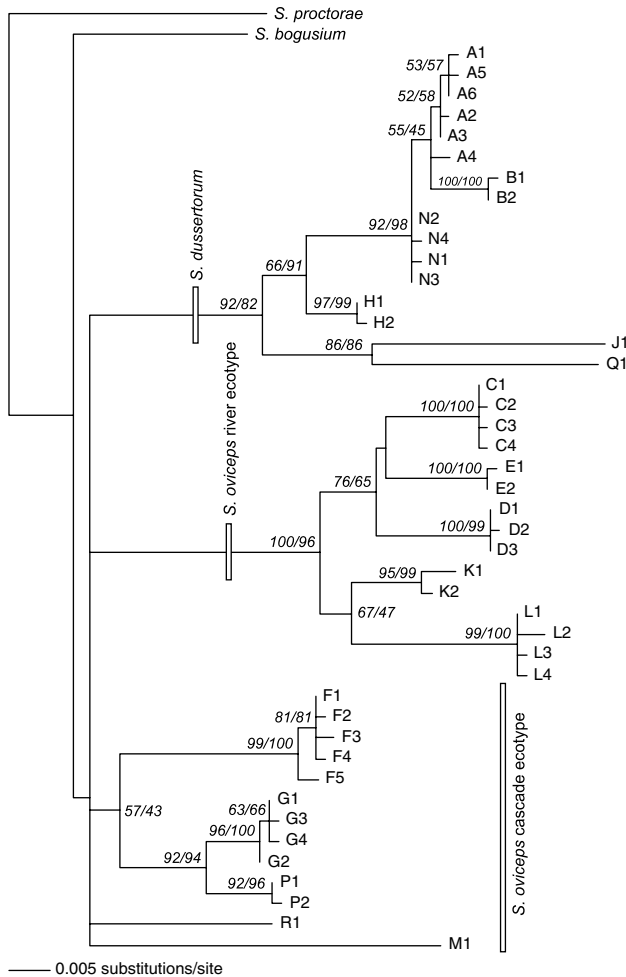
### Morphological variation

We measured head width, body length, length of the longest labral fan ray, and number of fan rays for 323 *S. oviceps* and 65 *S. dussertorum* final instar larvae. DNA from 87% of measured larvae ( $n = 338$ ) was successfully amplified and cut with restriction enzymes. We used analysis of variance to test for significant differences in larval morphology based on habitat (river vs cascade) and phylogeny (species and clades within *S. oviceps*).

## Results

### COI sequence phylogeny

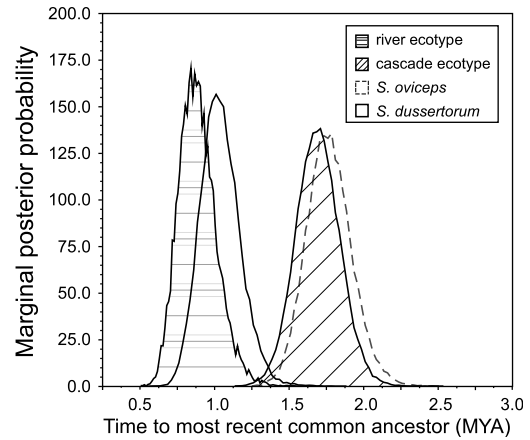
The TrN substitution model (Tamura and Nei, 1993) with a proportion of invariable sites and a  $\gamma$ -distribution of rate variation among sites (TrN + I +  $\Gamma$  model) was chosen based on likelihood ratio tests (Posada and Crandall, 1998). The ML and MP (not shown) trees



**Figure 2** ML tree estimated from *COI* sequences under the TrN+I+ $\Gamma$  model. The two outgroup species *S. bogusium* and *S. proctorae* are endemic to Raiatea. Sequence haplotype names are given at the tips of branches. ML followed by MP bootstrap proportions are shown above the branches (500 replicates). Branches with bootstrap values less than 50% are collapsed. *COI*, cytochrome oxidase I; ML, maximum likelihood; MP, maximum parsimony.

differed only in the arrangement of poorly supported branches within *S. dussertorum*. Neither ML nor MP support reciprocal monophyly between *S. oviceps* and *S. dussertorum* (Figure 2). Within *S. oviceps*, there is strong support for one clade, comprised of *S. oviceps* larvae collected from rivers (henceforth river ecotype, Figure 2) with the exception of haplotypes E1 and E2, collected from cascades on Moorea (Table 1). On the other hand, there is no support for reciprocal monophyly between this clade and cascade-dwelling *S. oviceps* (henceforth cascade ecotype, Figure 2).

The marginal posterior distribution of TMRCA for *S. oviceps*, *S. dussertorum* and the two ecotypes are shown in Figure 3. The TMRCA for *S. oviceps* (mean, 1.77 MY; 95% highest posterior density interval (HPD), 1.47–2.10 MY) and the cascade ecotype (mean, 1.70 MY; 95% HPD, 1.42–1.99 MY) showed considerable overlap and were the oldest (Figure 3). The river ecotype appears to have diverged more recently (mean, 0.89 MY; 95% HPD, 0.66–1.14 MY), with *S. dussertorum* intermediate between the



**Figure 3** Bayesian estimates of TMRCA for *S. oviceps*, *S. dussertorum*, *S. oviceps* river ecotype and *S. oviceps* cascade ecotype. ESS were 4284, 3454, 4085 and 3180 for *S. oviceps*, *S. dussertorum* and the river and cascade ecotypes, respectively. ESS, effective sample sizes; TMRCA, time to most recent common ancestor.

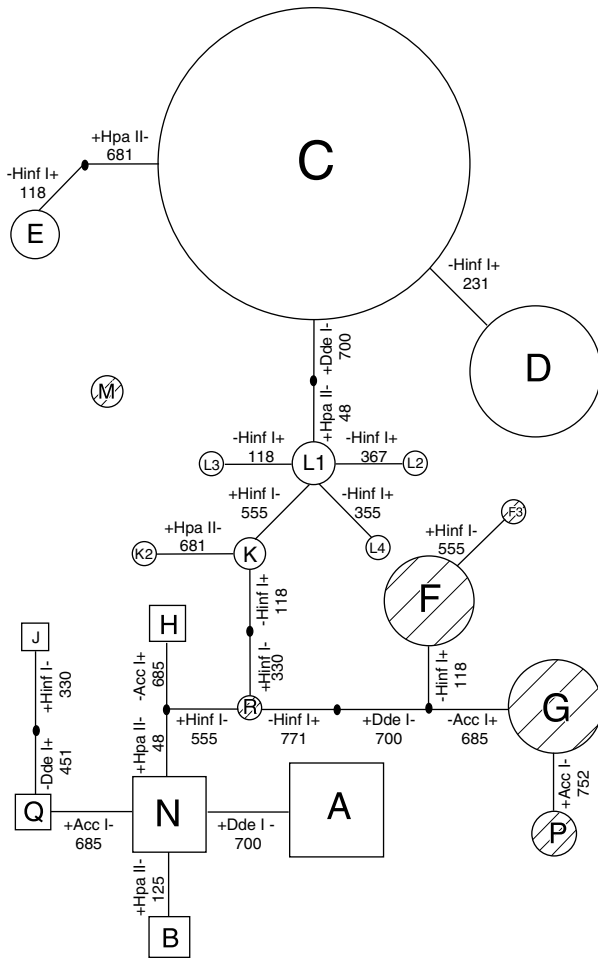
others but overlapping extensively with the river ecotype (mean, 1.05 MY; 95% HPD, 0.79–1.30 MY).

#### PCRD network and NCA

Restriction haplotypes corresponded well with sequence haplotypes, and divided *S. oviceps* into two groups, again reflecting larval habitat (Figure 4). Larvae were not distributed evenly among groups, with the *S. oviceps* river ecotype making up the majority of the collected larvae; almost one-third ( $n=161$ ) were haplotype C. Seven haplotypes were sampled only once. Two networks were supported by statistical parsimony, one consisting only of haplotype M. The two *S. oviceps* ecotypes and *S. dussertorum* were connected through haplotype R. Haplotype C had the highest outgroup weight (0.772; Supplementary Table 1) within the *S. oviceps* river ecotype and haplotype G had the highest outgroup weight (0.512) within the cascade ecotype, while within *S. dussertorum*, haplotype N had the highest outgroup weight (0.556).

Significant nonrandom associations between clades and geographical location were found for 3 out of 11 nested clades. In two other instances, geographical sampling was insufficient to make biological inferences. Within the *S. oviceps* river clade, haplotypes C and D had broadly overlapping ranges on Tahiti, while haplotype C was absent from Moorea (Figure 5a). An inference of contiguous range expansion from Tahiti to Moorea is supported by the distribution of these two haplotypes (Supplementary Figure 2). At the next highest nesting level, haplotype E, (clades 1–4 in Supplementary Figure 2) is found only in a few cascades on Moorea at the periphery of the nesting clade, which includes C, D and E (Figure 5b). The large discrepancy in range sizes and the location of haplotype E leads to an inference of past fragmentation (Supplementary Figure 2).

In *S. dussertorum*, the ranges of haplotypes A, B, and N overlap but differ in size (Figure 5c). The interior haplotype N was found only within the central crater of Tahiti, near the geographic center of the clade. Haplotype B was restricted to the Trois Cascades site.

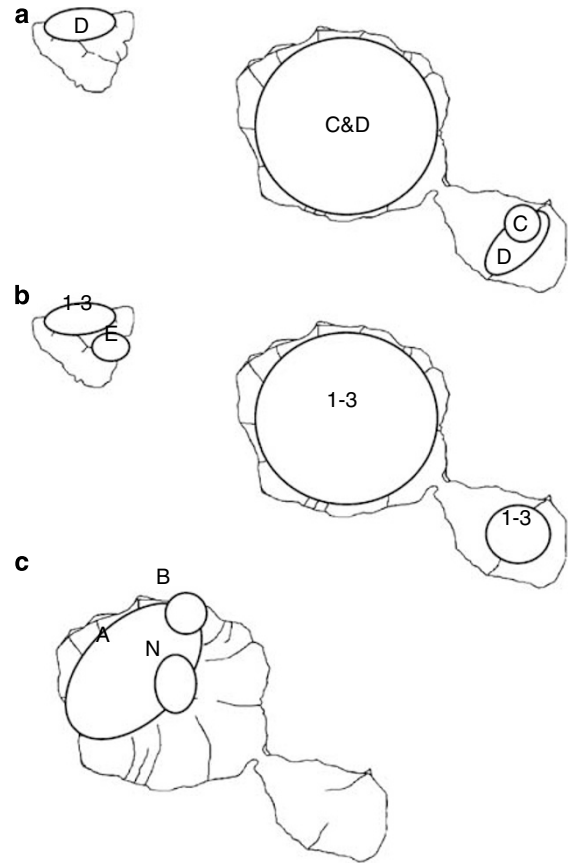


**Figure 4** Minimum spanning network for *S. oviceps* and *S. dussertorum* based on the restriction cut sites of a 920 bp *COI* fragment. Each branch represents the presence (+) or absence (–) of a restriction enzyme cut site with location given in bps. Black dots indicate intermediate haplotypes that were not sampled. Letters identify PCR/D haplotypes listed in Supplementary Table 1. White circle, *S. oviceps* river ecotype; cross-hatched circle, *S. oviceps* cascade ecotype; square, *S. dussertorum*. Symbol size is proportional to frequency of haplotype in the sample. *COI*, cytochrome oxidase I; PCR/D, PCR digest.

In contrast, the range of the tip haplotype A extended from within the central crater to the western and northern island periphery. These findings suggest range expansion outward from the central crater (Supplementary Figure 2).

#### Population structure

Most of the genetic variation within *S. oviceps* can be explained by differences in larval habitat while geographic barriers explained none of the variation (Table 2). In contrast, geographic barriers accounted for a large and significant percentage of the DNA variation in the *S. oviceps* cascade ecotype, although the amount of variation within populations was also substantial. For the river ecotype, genetic variation was split fairly and evenly between larval habitat and the within-populations component (Table 2), the former reflecting the fact that although E haplotypes fall within the river ecotype genetically, they were collected from cascades. In con-



**Figure 5** An overlay of nested clades on geography for clades in which the null hypothesis of no geographical structure was rejected. (a) *S. oviceps* river ecotype, range expansion; (b) *S. oviceps* river ecotype, past fragmentation and; (c) *S. dussertorum*, range expansion.

trast with the overall pattern for *S. oviceps* but in agreement with the *S. oviceps* cascade ecotype, the presence of geographical barriers accounted for more genetic variation in *S. dussertorum* than did larval habitat (Table 2).

#### Morphological variation

Mean head width and body length were significantly greater for larvae collected from cascades (head width:  $F_{1,336} = 138.57$ ,  $P < 0.0001$ ; body length:  $F_{1,334} = 95.94$ ,  $P < 0.0001$ ). Larvae from cascades also had significantly smaller labral fans than those from rivers (length of longest ray:  $F_{1,325} = 52.98$ ,  $P < 0.0001$ ; number of rays:  $F_{1,302} = 86.79$ ,  $P < 0.0001$ ).

Phylogenetic units (species and ecotype) also differed significantly from each other in overall size, with *S. oviceps* cascade ecotype larvae being the largest followed by *S. dussertorum* and finally *S. oviceps* river ecotype larvae (head width:  $F_{2,270} = 59.57$ ,  $P < 0.0001$ ; body length:  $F_{2,268} = 65.06$ ,  $P < 0.0001$ ); and differences among all three groups were significant (Tukey–Kramer honestly significant difference (HSD) *post hoc* test). This may simply reflect the effect of habitat. On the other hand, fan morphology differed significantly between species but not between *S. oviceps* ecotypes (length of longest ray:  $F_{2,264} = 81.62$ ,  $P < 0.0001$ ; number of rays:  $F_{2,243} = 173.66$ ,  $P < 0.0001$ ; Tukey–Kramer HSD *post hoc* test).

**Table 2** AMOVA for *S. oviceps* and *S. dussertorum* populations

Groups	Variance component	% Total variance	P-value
<i>Simulium oviceps</i>			
River, cascade	AG <sup>a</sup>	53.90	<0.0001
	AP/WG	18.15	<0.0001
	WP	27.95	<0.0001
Moorea, Tahiti-nui, Tahiti-iti	AG	2.29	0.2981
	AP/WG	59.57	<0.0001
	WP	38.14	<0.0001
<i>Cascade ecotype</i>			
Moorea, Tahiti-nui, Tahiti-iti	AG	48.51	0.0029
	AP/WG	13.30	0.0127
	WP	38.19	<0.0001
<i>River ecotype<sup>b</sup></i>			
River, cascade	AG	40.29	0.0049
	AP/WG	16.14	<0.0001
	WP	43.57	<0.0001
Moorea, Tahiti-nui, Tahiti-iti	AG	26.97	0.0020
	AP/WG	15.89	<0.0001
	WP	57.14	<0.0001
<i>Simulium dussertorum</i>			
River, cascade	AG	-8.28	0.8005
	AP/WG	77.01	<0.0001
	WP	31.27	<0.0001
Tahiti-nui, Tahiti-iti	AG	38.27	0.0450
	AP/WG	40.23	<0.0001
	WP	21.51	<0.0001

Abbreviations: AMOVA, analysis of molecular variance.

Geographic barriers refer to open ocean between Tahiti and Moorea and the narrow isthmus connecting Tahiti-nui with Tahiti-iti.

<sup>a</sup>AG is the among-groups variance component, AP/WG is the among populations/within groups variance component, WP is the within population variance component.

<sup>b</sup>Some *S. oviceps* river ecotype larvae (clade E) were collected from cascades that permitted partitioning based on larval habitat.

## Discussion

Our data illustrate the potential importance of larval ecology in the diversification of black flies in the Society Islands. We observed a pattern of genetic variation within *S. oviceps* not previously revealed by morphological analyses, and corresponding to larval aquatic habitat. Specifically, we found support for a genetically distinct river ecotype within *S. oviceps*. The two *S. oviceps* ecotypes are geographically sympatric, but are effectively allopatric because they occupy different stream types. The potential role of larval habitat in restricting gene flow is reinforced by the well-known vagility of black fly adults (Crosskey, 1990), making it surprising, in the absence of obvious barriers that subpopulation structure would exist on relatively small islands in close proximity to each other. Although *S. dussertorum* larvae are also found in rivers and cascades, ecological differences failed to explain genetic variation in this species, possibly because so few larvae were collected from rivers. An association between larval habitat shifts and cladogenesis has previously been proposed for Society Island black flies, based on the observation that species within the same clade tend to be found in similar stream types (Joy and Conn, 2001).

Despite strong support for a river ecotype clade, there is no such support for a monophyletic cascade ecotype within *S. oviceps*. As a result we are not able to distinguish between incipient and cryptic species status for the two *S. oviceps* ecotypes. Topological uncertainty

also exists regarding the reciprocal monophyly of *S. oviceps* and *S. dussertorum*. It is well known that incomplete lineage sorting can obscure evolutionary relationships among recently diverged taxa (Maddison, 1997). Craig (1987) speculated that *S. dussertorum* might be a variant within *S. oviceps*, and the morphological data presented here were equivocal. Labral fan size, which has been shown to be phenotypically plastic in response to current velocity (Zhang and Malmqvist, 1997), differed significantly between *S. oviceps* and *S. dussertorum* even within the same habitat. In contrast, fan size was conserved in *S. oviceps* across habitats. On the other hand, *S. dussertorum* was intermediate between the two *S. oviceps* ecotypes with respect to head width and body length.

*S. oviceps* and *S. dussertorum* are members of the *oviceps* group (Craig et al., 2001); a monophyletic clade consisting of at least nine species for which, with the exception of *S. oviceps*, cascades are the exclusive larval habitat (Craig et al., 2001; Joy and Conn, 2001). This strongly suggests that in the *oviceps* group, cascades are the ancestral stream type and that rivers were subsequently exploited; and our TMRCA estimates support this view. Large rivers on Tahiti and Moorea have relatively recent origins, their drainage basins having formed when the central calderas collapsed (Craig et al., 2001). The Papanoo River in the central cauldrea on Tahiti is estimated to have formed <0.87 MYA (Craig et al., 2001), which overlaps with the river ecotype age estimate (mean, 0.89 MY; 95% HPD, 0.66–1.14 MY).

The larvae of several other Society Island black fly species occupy rivers, although none have reduced feeding fans similar to *S. oviceps*. Schröder (1985) showed that *S. oviceps* in rivers had a high percentage of detritus particles in their gut while associated *Simulium tahitiense* larvae had a high percentage of unicellular algae in theirs, indicating that the two do not compete extensively for food. During a species radiation the invasion of novel habitats can provide new ecological opportunities and release from competition (Schluter, 1996). Here, we find *S. oviceps* exploiting resources in rivers by means of specialized feeding behavior (scraping) and associated morphology (reduced feeding fans) that presumably evolved in cascades. Thus, fine-scale partitioning of feeding niches may have facilitated the reinvasion of rivers.

Cascade populations exhibited higher levels of genetic sub-division than river populations. AMOVA indicated isolation between Tahiti-nui and Tahiti-iti for cascade but not river populations. The land bridge connecting the two is narrow and may restrict migration, but this does not explain the absence of sub-division among river populations. Nor can geographic distance between populations explain different levels of isolation for river and cascade populations. One possible explanation is that cascades may be intrinsically more isolated from each other than rivers due to their smaller size and patchy distribution. Of equal importance, small cascades in the Society Islands are often ephemeral, experiencing only intermittent, seasonal rainfall dependant flow (D Craig, personal communication). These conditions could lead to frequent local extinctions, thus augmenting isolation among cascade populations. Alternatively, given the younger geological age of rivers compared with cascades (Craig, 2003), and our estimated TMRCA

for the river ecotype, river populations may have had less time to diverge from each other.

The complex geological history of the Society Islands may have confounded estimates of relative haplotype age as well as NCA inferences. As hot spot islands age they erode and subside such that older islands are typically smaller than younger ones. Owing to the link between geological and hydrological profiles, the overall size of the catchment, as well as the diversity of stream environments, is expected to initially increase and later diminish as volcanic islands age and eventually subside (Haynes, 1990; Craig, 2003). Thus, we would expect local lineage extinction rates to be related to island age. The inability of NCA to detect localized haplotype extinction has been shown to lead to incorrect inferences (Masta et al., 2003). Given this and other difficulties in interpreting NCA results outlined by Knowles and Maddison (2002), we have treated NCA inferences not supported by other results as tentative. NCA and AMOVA both detected population structure with respect to the anomalous haplotype E. These larvae, collected from a cascade on Moorea are genetically within the *S. oviceps* river ecotype. On the other hand, the NCA inference of range expansion from Tahiti to Moorea for the *S. oviceps* river ecotype was not supported by other data. Additionally, *S. oviceps* haplotypes M and E are rare and endemic to cascades on Moorea, and outgroup weights for both are low due to their rarity and peripheral location in the PCRD network (Figure 3), suggesting recent origins. However, both had basal placements in the sequence phylogeny (Figure 2). The smaller size and greater age of Moorea (Figure 1) would predict a greater loss of suitable stream habitats overtime compared with Tahiti. This may in turn have distorted the haplotype network, such that older haplotypes on Moorea (possibly E and M), are currently less abundant than younger haplotypes on Tahiti.

Phenotypic plasticity in growth rate and labral fan morphology in response to food availability and current velocity has been widely documented for black fly larvae (Zhang and Malmqvist, 1997; Lucas and Hunter, 1999; Zhang, 2005). Genetic assimilation (Waddington, 1961), in which moderate levels of phenotypic plasticity promote establishment in novel habitats, and subsequent selection for extreme phenotypes leading to genetic differentiation and perhaps speciation, has been proposed as an important mechanism for black fly diversification (Zhang, 2005). We suggest that genetic assimilation may figure prominently in the radiation of black flies in the Society Islands by enhancing their ability to successfully colonize novel niches. Our finding that *S. oviceps* larvae from rivers are genetically distinct supports this view.

## Acknowledgements

We are indebted to the Institut Louis Malardé, Papeete and particularly Yves Séchan and Marc Faarua for providing DAJ and DAC with logistical support during fieldwork. The manuscript was improved by comments from CW Kilpatrick. Financial support was provided by the University of Vermont (UVM) College of Arts and Sciences Grant Seed Fund to JEC, a UVM Research and Scholarship Committee Grant to JEC, a National Science Foundation EPSCoR grant OSR-9350540 to UVM for both

DAJ and JEC, and a Canada Discovery Grant (No. 5753) to DAC.

## References

- Brower AVZ (1994). Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. *Proc Natl Acad Sci USA* **91**: 6491–6495.
- Carson HL, Clague DA (1995). Geology and biogeography of the Hawaiian Islands. In: Wagner WL, Funk VA (eds). *Hawaiian Biogeography: Evolution on a Hot Spot Archipelago*. Smithsonian Institution Press: Washington. pp 14–29.
- Carson HL, Templeton AR (1984). Genetic revolutions in relation to speciation phenomena: the founding of new populations. *Ann Rev Ecol Syst* **15**: 97–131.
- Castelloe J, Templeton AR (1994). Root probabilities for intraspecific gene trees under neutral coalescent theory. *Mol Phyl Evol* **3**: 102–113.
- Collins FH, Mendez MA, Rasmussen MO, Mehaffy PC, Besansky NJ, Finnerty V (1987). A ribosomal probe differentiates member species of the *Anopheles gambiae* complex. *Am J Trop Med Hyg* **37**: 37–41.
- Cowie RH (1995). Variation in species diversity and shell shape in Hawaiian landsnails: *in situ* speciation and ecological relationships. *Evolution* **49**: 1191–1202.
- Craig DA (1977). Mouthparts and feeding behaviour of Tahitian larval Simuliidae (Diptera: Nematocera). *Quaest Ent* **13**: 195–218.
- Craig DA (1987). A taxonomic account of the black flies (Diptera: Simuliidae) of the Society Islands- Tahiti, Moorea and Raiatea. *Quaest Ent* **23**: 372–429.
- Craig DA (1997). A taxonomic revision of the Pacific black fly subgenus *Inseliellum* (Diptera: Simuliidae). *Can J Zool* **75**: 855–904.
- Craig DA (2003). Geomorphology, development of running water habitats, and evolution of black flies on Polynesian islands. *BioScience* **53**: 1079–1093.
- Craig DA, Currie DC, Joy DA (2001). Geographical history of the central-western Pacific black fly subgenus *Inseliellum* (Diptera: Simuliidae: *Simulium*) based on a reconstructed phylogeny of the species, hot-spot archipelagos and hydrological considerations. *J Biogeogr* **28**: 1101–1127.
- Crosskey RW (1990). *The Natural History of Black Flies*. John Wiley and Sons: Chichester.
- DeSalle R, Freedman T, Prager EM, Wilson AC (1987). Tempo and mode of sequence evolution in mitochondrial DNA of Hawaiian *Drosophila*. *J Mol Evol* **26**: 157–164.
- Drummond AJ, Rambaut A (2006). BEAST v1.4. Available from <http://beast.bio.ed.ac.uk/>.
- Duncan RA, McDougall I (1976). Linear volcanism in French Polynesia. *J Volcanol Geotherm Res* **1**: 197–227.
- Excoffier L, Smouse P, Quattro J (1992). Analysis of molecular variance inferred from metric distance among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* **131**: 479–491.
- Gillespie R (2004). Community assembly through adaptive radiation in Hawaiian spiders. *Science* **303**: 356–359.
- Haynes A (1990). The number of freshwater gastropods on Pacific islands and the theory of island biogeography. *Malacologia* **31**: 237–248.
- Joy DA, Conn JE (2001). Molecular and morphological phylogeny of an insular radiation in Pacific black flies (*Simulium*). *Syst Biol* **50**: 18–38.
- Kambysellis MP, Ho K-F, Craddock EM, Piano F, Parisi M, Cohen J (1995). Patterns of ecological shifts in the diversification of Hawaiian *Drosophila* inferred from a molecular phylogeny. *Curr Biol* **5**: 1129–1139.
- Knowles LL, Maddison WP (2002). Statistical phylogeography. *Mol Ecol* **11**: 2623–2635.

- Losos JB (1994). Integrative approaches to evolutionary ecology: *Anolis* lizards as model systems. *Ann Rev Ecol Syst* **25**: 467–493.
- Lucas P, Hunter F (1999). Phenotypic plasticity in the labral fan of simuliid larvae (Diptera): effect of seston load on primary ray number. *Can J Zool* **77**: 1843–1849.
- Lunt DH, Zhang D-X, Szymura JM, Hewitt GM (1996). The insect cytochrome oxidase I gene: evolutionary patterns and conserved primers for phylogenetic studies. *Insect Mol Biol* **5**: 153–165.
- Maddison WP (1997). Gene trees in species trees. *Sys Biol* **46**: 523–536.
- Masta SE, Laurent NM, Routman EJ (2003). Population genetic structure of the toad *Bufo woodhousii*: an empirical assessment of the effects of haplotype extinction on nested clade analysis. *Mol Ecol* **12**: 1541–1554.
- Posada D, Crandall KA (1998). Modeltest: testing the model of DNA substitution. *Bioinformatics* **14**: 817–818.
- Posada D, Crandall KA, Templeton AR (2000). GeoDis: a program for the cladistic nested analysis of the geographical distribution of genetic haplotypes. *Mol Ecol* **9**: 487–488.
- Rambaut A, Drummond AJ (2004). *Tracer, Version 1.2*. University of Oxford: Oxford, UK.
- Resh VH, Barnes JR, Craig DA (1990). Distribution and ecology of benthic macroinvertebrates in the Opunohu River catchment, Moorea French Polynesia. *Ann Limnol* **26**: 195–214.
- Satta Y, Ishiwa H, Chigusa SI (1987). Analysis of nucleotide substitutions of mitochondrial DNAs in *Drosophila melanogaster* and its sibling species. *Mol Biol Evol* **4**: 638–650.
- Schluter D (1996). Ecological causes of adaptive radiation. *Am Nat* **148**: S40–S64.
- Schröder P (1985). Feeding biology of Tahitian blackfly larvae (Dipt Simuliidae). *Notes et Documents, Hygiène et Santé Publique Entomologie Médicale* **11**: 1–230.
- Shaw KL (1996). Sequential radiations and patterns of speciation in the Hawaiian cricket genus *Laupala* inferred from DNA sequences. *Evolution* **50**: 237–255.
- Swofford DL (2003). *PAUP\*: Phylogenetic Analysis Using Parsimony (\*and Other Methods)*, Version 4. Sinauer Associates: Sunderland, MA.
- Tamura K, Nei M (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzee. *Mol Biol Evol* **10**: 512–526.
- Templeton AR, Crandall KA, Sing CF (1992). A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics* **132**: 619–633.
- Templeton AR, Routman E, Phillips CA (1995). Separating population structure from population history: a cladistic analysis of the geographical distribution of mitochondrial DNA haplotypes in the tiger salamander, *Ambystoma tigrinum*. *Genetics* **140**: 767–782.
- Thompson BH (1987). The use of algae as food by larval Simuliidae (Diptera) of Newfoundland streams: II digestion of algae, and environmental factors affecting feeding rates and the degree of digestion. *Arch Hydrobiol Suppl* **76**: 443–457.
- Waddington C (1961). Genetic assimilation. *Adv Genet* **10**: 257–293.
- Williamson M (1981). *Island Populations*. Oxford University Press: Oxford.
- Zhang Y (2005). Balancing food availability and hydrodynamic constraints: phenotypic plasticity and growth in *Simulium noelleri* blackfly larvae. *Oecologia* **147**: 39–46.
- Zhang Y, Malmqvist B (1997). Phenotypic plasticity in a suspension-feeding insect *Simulium lundstromi* (Diptera: Simuliidae) in response to current velocity. *Oikos* **78**: 503–510.

Supplementary Information accompanies the paper on Heredity website (<http://www.nature.com/hdy>)