

A phylogeographic investigation of the hybrid origin of a species of swordtail fish from Mexico

JULIA C. JONES,*† JUAN-ANTONIO PEREZ-SATO‡ and AXEL MEYER*

*Lehrstuhl für Zoologie und Evolutionsbiologie, Department of Biology, University of Konstanz, Universitätsstraße 10, 78457 Konstanz, Germany, †Zukunftskolleg, University of Konstanz, Konstanz, Germany, ‡Unidad de Abejas, Colegio de Postgraduados, Campus Cordoba, Carretera Federal Cordoba-Veracruz Km 348, Congregacion Manuel Leon, Amatlan de Los Reyes, Veracruz C.P 94946, Mexico

Abstract

Hybrid speciation may contribute significantly to generating biodiversity, but only a few well-documented examples for it exist so far that do not involve polyploidization as a mechanism. The swordtail fish, *Xiphophorus clemenciae*, shows common hallmarks of a hybrid origin and still overlaps in its current geographic distribution with its putative ancestral species (*Xiphophorus hellerii* and *Xiphophorus maculatus*). *Xiphophorus clemenciae* provides an ideal system for investigating the possible continued genetic interactions between a hybrid and its parental species. Here, we use microsatellite and mitochondrial markers to investigate the population structure of these species of swordtails and search for signs of recent hybridization. Individuals were sampled from 21 localities across the known range of *X. clemenciae* – the Isthmus of Tehuantepec (IT) Mexico, and several environmental parameters that might represent barriers to dispersal were recorded. The hybridization event that gave rise to *X. clemenciae* appears to be rather ancient, and a single origin is likely. We find negligible evidence for ongoing hybridization and introgression between the putative ancestral species, because they now occupy distinct ecological niches, and a common haplotype is shared by most populations of *X. clemenciae*. The population structure within these species shows an isolation-by-distance (IBD) pattern and genetic differentiation between most populations is significant and high. We infer that tectonic evolution in the Isthmus has greatly restricted gene flow between the southern and central IT populations of *X. clemenciae* and *X. hellerii* and provide preliminary information to aid in conservation management of this geographically restricted hybrid species, *X. clemenciae*.

Keywords: genetic differentiation, hybridization, microsatellite, mtDNA, niche analysis, *Xiphophorus*

Received 12 July 2011; revision received 7 February 2012; accepted 16 February 2012

Introduction

The role and relative importance of hybridization in speciation and evolution has been debated throughout the development of evolutionary theory. Traditionally, hybridization has been viewed as a rare event, and the potential for natural hybrids to form novel fit genotypes was downplayed (e.g. Fisher 1930; Mayr 1942, 1963), or as R.A. Fisher put it: ‘The grossest blunder in sexual

preference, which we can conceive of an animal making, would be to mate with a species different from its own...’ Generally, but not universally, the typical outcome of hybridization is seen as negative in that the hybrid offspring would be less fit than either parent and go extinct, or a reversal of speciation would occur leading to fewer species rather than the origin of new species (reviewed in Arnold 1997; Arnold & Meyer 2006). However, there was also early recognition of the potentially creative role of hybridization in evolution – particularly through polyploidization in plants (e.g. Anderson & Stebbins 1954; Lewontin & Birch 1966;

Correspondence: Axel Meyer, Fax: +49 7531 883018; E-mail: axel.meyer@uni-konstanz.de

Templeton 1981; reviewed in Arnold 1997; Arnold & Meyer 2006) – and today evidence continues to accumulate that suggests that hybridization is widespread, although there surely are taxon-specific differences (e.g. Scribner *et al.* 2001; Reyer 2008; Whitney *et al.* 2010). More recently, hybrid speciation was touted as a potential key player in the evolutionary history of plants in particular, but also some lineages of animals (Dowling & Secor 1997; Barton 2001; Mallet 2005, 2007; Soltis & Soltis 2009). In fishes, hybridization was documented quite early (Hubbs 1955), and there is increasing evidence for hybridization and even reverse speciation among a number of African cichlid species (Salzburger *et al.* 2002a; Koblmüller *et al.* 2004, 2007, 2010; Streelman *et al.* 2004; Schelly *et al.* 2006; Nyingi & Agnése 2007), and hybridization in cyprinid fishes as well (Scribner *et al.* 2001; Hayden *et al.* 2010). However, so far, there is little evidence from fishes for a creative role of hybridization in speciation (Meyer *et al.* 1994, 2006; but see Seehausen 2004; Nosil *et al.* 2009). It remains unclear how common hybrid origins are in various animal lineages, whether hybridization tends to consist of isolated events or alternatively persists over time, and how often hybrids might be fitter than their ancestral species and what portion of species arose by this type of mechanism (for a review see Arnold & Meyer 2006).

If natural hybridization and introgression – where genetic information passes from the gene pool of one species into the gene pool of another by hybridization – did play a major role in the evolutionary history of a particular lineage, we should be able to detect the genetic consequences in current species (Arnold 1992; Mallet 2005, 2007). An easily identified signature of past hybridization is incongruence between phylogenetic data sets based on different sets of genetic markers (Arnold 1992; Avise 1994, 2000; Seehausen 2004; Arnold & Meyer 2006). Different regions of the genome are expected to introgress at different rates due to the action of selection and drift, and because mtDNA is maternally inherited, the effective population size for mtDNA markers is smaller than for nuclear markers. This means different regions of a hybrid genome will carry a mosaic of markers from each parental species and would hence be expected to result in conflicting phylogenetic signals (Seehausen 2004; Arnold & Meyer 2006). Meyer *et al.* (1994, 2006) showed such discordance for *Xiphophorus clemenciae* (swordtail), from the family Poeciliidae, and suggested that this species might have a hybrid origin.

The Poeciliidae are a species-rich family of small freshwater fish from Central America. Poeciliids are a widely used model system for investigating a range of evolutionary and ecological questions such as the pre-existing bias hypothesis (e.g. Quattro & Vrijenhoek

1989; Basolo 1990b; Meyer *et al.* 1994; Scharl *et al.* 1994; Scharl 1995a) and sexual selection theory more generally (Wagner & Ryan 1987; Basolo 1990b, 1995a,b, 1997; Kirkpatrick & Ryan 1991; Houde 1997; Magurran 2005; Schlupp *et al.* 2010). Poeciliids have also featured in early work on conservation genetics (Vrijenhoek *et al.* 1985; Quattro & Vrijenhoek 1989; Quattro *et al.* 1996) and recent whole-genome research (Tripathi *et al.* 2009; Willing *et al.* 2010; Zhang *et al.* 2011; Shen *et al.* 2012).

However, knowledge of hybridization and its importance during evolution in this family, the population dynamics of different *Xiphophorus* (swordtails and platyfish) species and knowledge about the environmental features that permitted or even promoted hybridization is still wanting. Microgeographic population structure and population genetics have been investigated in *Xiphophorus hellerii* (Tatarenkov *et al.* 2010) and *X. cortezi* (Gutiérrez-Rodríguez *et al.* 2007), respectively. Hybridization has been suggested to occur in some species of the genus *Xiphophorus*. For example, in seven different streams, extensive hybridization has been found between the species *X. malinche* and *X. birchmanni*, with a distinct elevational gradient between the two species (Culumber *et al.* 2011).

Fish of the genus *Xiphophorus*, swordtails and platyfish, are interesting candidates for investigating the role of hybridization in evolution. Males of some species of *Xiphophorus*, swordtails, possess a swordlike elongated ventral caudal fin that – since Charles Darwin (Darwin 1871) – has been the subject of much behavioural research on the evolution of an exaggerated male trait through female preference (Darwin 1871; e.g. Basolo 1990a, 1995a,b; Rosenthal & Evans 1998; Rosenthal *et al.* 2002; Meyer *et al.* 2006) (see Fig. 1). Of specific interest is the established finding that females of some species of this genus, the platyfish, whose conspecific males do not possess a sword, prefer males of their own species with an artificial sword, over swordless ones (Basolo 1990a). Female mating preference for sworded males has been suggested to be a pre-existing ancient bias that may have promoted the evolution of the sword trait (Basolo 1995a,b, see also Meyer 1997) and possibly hybridization between female platyfish and male swordtails (Meyer *et al.* 2006).

There are several lines of evidence suggesting that at least one species of swordtail fish, *X. clemenciae*, arose through a hybridization event. First, in a mitochondrial phylogeny, the 'sworded' *X. clemenciae* was found to be nested among the 'swordless' southern platyfish, whereas in the nuclear phylogeny, it was grouped together with the other species of southern swordtails (Meyer *et al.* 1994, 2006). Incongruence between phylogenies based on molecular markers is recognized as



Fig. 1 Photographs of *Xiphophorus* (swordtails and platyfish) species examined in this study (scale bar = 1 cm). Swordtail males (e.g. *Xiphophorus hellerii* and *Xiphophorus clemenciae*) have an elongated ventral caudal fin of different relative length – those of *X. hellerii* are typically longer than those of *X. clemenciae* – whereas platyfish males do not. Here, we include a female sample of *Xiphophorus maculatus* only. (Note that dorsal fins were pinned for photographs, and a small fraction of the pin is visible).

evidence for past reticulate events (Avice 2004). The incongruent placement of *X. clemenciae* on mitochondrial and nuclear DNA-based phylogenetic trees suggests that this species inherited its mitochondrial DNA from a southern platyfish species (such as *Xiphophorus maculatus*) and its nuclear DNA from a southern swordtail species (most likely *X. hellerii*) (Meyer *et al.* 2006) (see Fig. 1).

Also, behavioural data suggest that such a hybridization event is plausible (Meyer *et al.* 2006). The putative parental species, *X. hellerii* and *X. maculatus*, hybridize under laboratory conditions, and as had been described before, female platyfish prefer males with artificially elongated caudal fins (swords) over nonsworded males (Basolo 1990a,b). We showed previously (Meyer *et al.* 2006) that hybrid offspring of extant representatives of the potential ancestral parents (*X. maculatus* and *X. hellerii*) are fertile and that the resulting hybrid males have

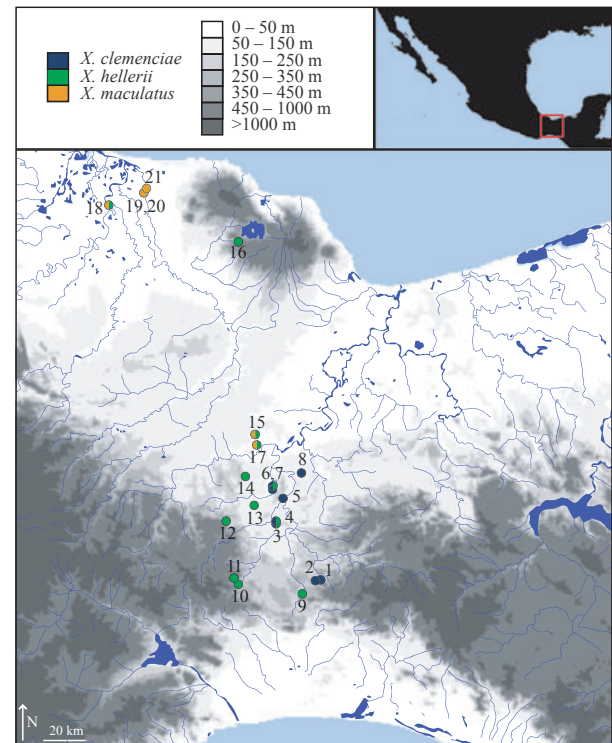


Fig. 2 Sample locations of the three *Xiphophorus* study species in the Isthmus of Tehuantepec, Mexico (sites are labelled as in Table 1 and Fig. 5). Sites are labelled by species and geographic location south to north. Sample locations are identified by a coloured circle, and the colour of the circle designates which species was found at each site (blue = *Xiphophorus clemenciae*, green = *Xiphophorus hellerii*, orange = *Xiphophorus maculatus*). The grey scale indicates the height above sea level.

swords of intermediate lengths that resemble those of *X. clemenciae*. Hybridizations and backcrosses in the laboratory between *X. hellerii* and *X. maculatus* (backcrosses = hybrid \times *X. hellerii*) produced hybrids that resembled *X. clemenciae* phenotypically and produced males with intermediate sword lengths (Meyer *et al.* 2006). In addition, when provided with a choice of mates, hybrid females prefer sworded *X. hellerii* males over nonsworded *X. maculatus* males (Meyer *et al.* 2006). These backcrosses of hybrid females with the southern swordtail parental lines would be expected to lead to a mostly southern swordtail nuclear background with a southern platy mitochondrial genome (Meyer *et al.* 2006).

Xiphophorus clemenciae has a restricted geographical distribution, being confined to small regions of the Isthmus of Tehuantepec (IT) in Mexico (Fig. 2). *Xiphophorus clemenciae* has been suggested to be a rare and threatened species with the most restricted distribution of all southern swordtail species (Alvarez 1959; Rosen 1960; Ramirez 1999). However, within its core range, *X. clemenciae* is quite abundant and widespread. Kallman *et al.*

(2004) readily found *X. clemenciae* in a particular area of the IT – the uplands of the Rio Coatzacoalcos system – but this species has never been found in the adjacent coastal plain. By contrast, the putative parental species, *X. hellerii* and *X. maculatus*, have the largest distribution of all *Xiphophorus* species extending from Veracruz, Mexico to Honduras and Guatemala, respectively, and both putative parental species overlap in their distribution with *X. clemenciae* (Kallman *et al.* 2004; Kallman & Kazianis 2006). We did not examine other potential maternal lineages, such as *X. milleri* (Meyer *et al.* 2006), as this platyfish has a very restricted distribution that does not overlap with that of *X. clemenciae* (Kallman & Kazianis 2006) or *X. hellerii* (M. Scharl personal communication). Given this distribution, it is rather improbable that this species had been or currently is hybridizing to produce *X. clemenciae*-like hybrids or introgressing with *X. clemenciae*.

To date, this group of fish provides one of the very few known examples, across all animal taxa, of speciation via hybridization (Mallet 2007). In this case, there are several lines of evidence that support the hypothesis that one species arose through an ancient hybridization event. All evidence available so far, including incongruent phylogenies, laboratory-reared fertile hybrid offspring, *X. maculatus*'s preference for sworded males and currently overlapping geographic distributions, suggests that ongoing hybridization and introgression between these species are plausible. We asked whether hybridization persists over time or whether hybridization ceased even when it was the initial mechanism by which this hybrid species arose. Here, we investigated the population structure of *X. clemenciae* and its putative parental species, *X. maculatus* and *X. hellerii*, and test for recent hybridization. We quantified environmental features that might influence the distribution of these species and their recent evolutionary history. Using both nuclear microsatellite variation and mitochondrial DNA sequence data, we determined the level of population differentiation within and between the putative parental and hybrid species and we compared a range of environmental parameters between their respective native habitats. Based on tectonic history in the IT, we reconstructed how geological history influenced the evolutionary history of these fishes and further made recommendations for conservation.

Methods

Sample collection

Samples were collected from the putative parental (*Xiphophorus hellerii* and *Xiphophorus maculatus*) and hybrid (*Xiphophorus clemenciae*) species during the dry

season (January–February) of 2009 from small streams and ponds in the IT, in southeastern Mexico. Specimens were collected at 21 locations; seven predominantly *X. hellerii* populations, seven predominantly *X. clemenciae* populations, five predominantly *X. maculatus* populations, one equally mixed *X. hellerii* – *X. clemenciae*, one equally mixed *X. hellerii* – *X. maculatus* (Table 1, Fig. 2, Table S1, Supporting information) covering most of the known current geographical distribution of the hybrid species. A total of 840 individuals, comprising juveniles, and adult males and females, were sampled (Table 1). *Xiphophorus hellerii* was found to live sympatrically with one of the other study species at a number of sites, although often in comparably small numbers. Fish were sampled with a consistent method across all sites; thus, sample sizes within and between sites can act as a proxy for abundance (Table 1). Fish were caught using a seine net (10 × 3 foot, with a 1/8-inch mesh) and photographed, and then fin clips were collected and stored in ethanol – after that, all individuals were released. All fish were identified as *X. hellerii*, *X. maculatus* or *X. clemenciae* based on the traits highlighted in Kallman *et al.* (2004) (Kallman, personal communication). Collection sites were chosen based on the most recent distribution and site descriptions (Kallman

Table 1 Specimens by sample site and species

Site	Species	N per species
1	<i>Xiphophorus clemenciae</i>	53
2	<i>X. clemenciae</i>	48
3	<i>X. clemenciae</i>	27
	<i>Xiphophorus hellerii</i>	3
4	<i>X. clemenciae</i>	36
	<i>X. hellerii</i>	6
5	<i>X. clemenciae</i>	17
6	<i>X. clemenciae</i>	55
7	<i>X. clemenciae</i>	44
	<i>X. hellerii</i>	21
8	<i>X. clemenciae</i>	42
9	<i>X. hellerii</i>	39
10	<i>X. hellerii</i>	33
11	<i>X. hellerii</i>	52
12	<i>X. hellerii</i>	26
13	<i>X. hellerii</i>	67
14	<i>X. hellerii</i>	28
15	<i>X. hellerii</i>	9
	<i>Xiphophorus maculatus</i>	10
16	<i>X. hellerii</i>	37
17	<i>X. maculatus</i>	32
	<i>X. hellerii</i>	2
18	<i>X. maculatus</i>	43
	<i>X. hellerii</i>	1
19	<i>X. maculatus</i>	39
20	<i>X. maculatus</i>	30
21	<i>X. maculatus</i>	40

et al. 2004; Kallman, personal communication) to cover the range of *X. clemenciae* and potential hybrid zones.

Microsatellite genotyping and analyses

Genomic DNA was extracted from fin clips using the Chelex[®] method (Walsh *et al.* 1991). Nine microsatellite loci (that were developed for *Xiphophorus* and selected for a previous study on *X. hellerii*) were amplified using standard PCR conditions: Msb080, Msd045, Msb069, Msd033, Msd036, Msd051, Msd060, Msd055, Msc045 (Tatarenkov *et al.* 2008). One locus (Msd060) could not be amplified in *X. clemenciae*; therefore, it was excluded. Fluorescently labelled fragments were analysed on an ABI3130XL (Applied Biosystems, CA, USA) with the internal size marker Liz 600 in GENEMAPPER version 4.0 (Applied Biosystems).

Genotyping efficiency for each population was verified by testing in MICRO-CHECKER (van Oosterhout *et al.* 2004) with 1000 randomizations and a 95% confidence interval. The potential influence of null alleles was low (Brookfield 1996) as no consistent differences were found across loci or populations using MICRO-CHECKER; therefore, all loci were retained. To assess the microsatellite diversity in all species, the allelic richness and observed (H_o) and expected (H_E) heterozygosity were calculated using FSTAT VERSION 2.9.3 (Goudet 1995) and GENETIX VERSION 4.05 (Belkhir *et al.* 2004), respectively. One-way ANOVA was used to compare allelic richness between species. Multilocus and locus-by-locus tests for Hardy–Weinberg equilibrium (HWE) were performed in GENEPOP VERSION 4.0 (Rousset 2008). Pairwise population genetic differentiation within and between species was estimated using Wright's F -statistics (F_{ST}) (Weir & Cockerham 1984) as calculated by ARLEQUIN VERSION 3.1 (Excoffier *et al.* 2005). Populations with less than ~15 samples were not included in our population differentiation analyses, except for populations of *X. hellerii* and *X. maculatus* from site 15 (with $n = 9$ and 10, respectively, Table 1). The latter populations were included in our analyses because this was the only study site where comparable numbers of both *X. hellerii* and *X. maculatus* were found and collected. An AMOVA was calculated in ARLEQUIN for each species to investigate the genetic variation within and between groups. Populations were grouped together by geographically proximal location; for example, *X. clemenciae* populations from sites 1 and 2 were grouped together, and populations from sites 3 and 4 were grouped together. Population differentiation within and between species and evidence for hybrid individuals were estimated using the program STRUCTURE that uses an individual-based Bayesian cluster approach (Pritchard *et al.* 2000). For all estimations, we used 2×10^6 steps of

the Markov chain and a burn-in period of 5×10^5 steps, and five iterations were used for each genetic cluster (K). The level of genetic differentiation in our pairwise comparisons was taken into account when nominating the upper bound of the different genetic clusters (K) for our population-level analyses: *X. maculatus* $K = 1-5$, *X. hellerii* $K = 1-8$, *X. clemenciae* $K = 1-6$, and the combined putative parental and hybrid species $K = 1-12$. The latter approach aims to avoid misclassification of genetically similar groups in STRUCTURE that can occur when inappropriately small numbers of clusters are designated (Kalinowski 2010). We used ΔK (Evanno *et al.* 2005) to estimate the most likely value for K in all population-level structure analyses. In a separate structure analysis, we limited K to 3 (i.e. the number of species in our study), and prior information on the species of each individual was included in the analysis. This analysis was also used to estimate the probability of having an immigrant ancestor for two generations (Gensback = 2). The latter analysis estimates the posterior probability that an individual is correctly assigned to the given population, and the probability that it is from, or has ancestry in, the other populations (here the other species). To test for a genetic pattern of isolation by distance (Wright 1943), for each individual species, the relationship between genetic differentiation (F_{ST}) and estimated waterway distances between collection sites (using digital maps of the sites in the program IMAGEJ) was evaluated with Mantel tests in ARLEQUIN (Excoffier *et al.* 2005).

Mitochondrial DNA sequencing and analyses

A subset of our population samples (~16 samples randomly selected per population) was sequenced for the complete mitochondrial control region using methods similar to those previously described (Meyer *et al.* 2006; using primers from Meyer *et al.* 1994; Lee *et al.* 1995; Salzburger *et al.* 2002b). At sites where the majority of samples collected were either *X. clemenciae* or *X. maculatus*, but also a small number of *X. hellerii* samples (1–2 samples per site, Table 1) were found, the latter were sequenced for additional comparisons. To determine the levels of current genetic diversity in our three study species, the number of mtDNA haplotypes, number of polymorphic sites and nucleotide diversity (π) were calculated for each population using DNASP VERSION 5.1 (Librado & Rozas 2009) and haplotype richness after rarefaction (HR) was calculated using Contrib-1.02 (Petit *et al.* 1998) and species means were compared with one-way ANOVA. To investigate the evolutionary relationships between and among species, a median-joining haplotype network was calculated in NETWORK VERSION 4.5.1.6 <http://www.fluxus-engineering.com/> using samples from all populations of all three species.

To examine intra- and interspecific population differentiation and the variation within and between geographically distinct species-specific populations, F_{ST} estimates and AMOVA were calculated in ARLEQUIN using a Tamura–Nei + gamma (unequal base frequencies, unequal substitution rates) corrected distance matrix as an approximation of the best model of evolution that was estimated to be TrN + I + G in MODELTEST VERSION 3.7 (Posada & Crandall 1998) in PAUP* VERSION 4.0 (Swofford 2003).

Niche analyses

For each collection location, we recorded a snapshot of a variety of environmental parameters including altitude, water temperature, pH, dissolved oxygen, conductivity, salinity, water flow, stream depth and bank width (Table S1, Supporting information). All parameters were recorded in the region of the stream or pond where the majority of fish samples were collected and recordings were made directly after sampling was completed. Water temperature, pH, dissolved oxygen and conductivity were measured using a Multi-350i hand-held multiparameter meter (WTW, GmbH, Germany). Water flow was measured using a MiniWater20 flow meter (Schiltknecht Messtechnik AG, Switzerland). For each of the parameters described above, an average of three measurements (upper, middle and lower thirds of the water body) was taken. Stream depth and bank width were measured using a metre stick and measuring tape, respectively. In addition, elevation and position were recorded using a hand-held global positioning system (GPS) unit (GARMIN, USA).

The niches of all three species were computed and compared using the ade4-package in R (R Development Core Team 2010). Nine site parameters (altitude, pH, water temperature, bank width, stream depth, stream velocity, conductivity and dissolved oxygen) were included in the analysis. The niches of the putative hybrid and parental species were calculated according to Doledec *et al.* (2000). Briefly, this analysis is a multivariate technique based on the outlying mean index (OMI), or marginality of a species, and can also be used to analyse species assemblages. The marginality or OMI describes the distance of a species' niche to the mean available habitat conditions of the study area. To transform the data into a duality diagram (Escoufier 1987), principal components analysis (PCA) was used. We performed the OMI analysis using R and tested for significance of the results with a random Monte–Carlo test, one sided with 1000 permutations. We compared the niches of the different species and ascertained which parameters potentially separate the niches using Pianka's niche overlap index (Pianka 1973; Giraudoux 2010).

This index was calculated for every species using the parameters listed above.

Results

Genetic diversity

Microsatellites. The levels of allelic richness and heterozygosity were higher on average for *Xiphophorus maculatus* populations than for *Xiphophorus hellerii* and *Xiphophorus clemenciae* populations and higher in *X. hellerii* than in *X. clemenciae* populations (Table 2). The mean allelic richness across populations for *X. maculatus* was significantly higher than in the other two species (post hoc, Bonferroni, $P \leq 0.001$), and the mean allelic richness for *X. hellerii* was higher than for *X. clemenciae* but not significantly different ($P = 0.574$). Interestingly, allelic richness in *X. clemenciae* was significantly negatively correlated with altitude ($Y = -0.014X + 7.485$, $R^2 = 0.741$, $P = 0.006$). Most populations of *X. hellerii* (all except one; Table 2) were in Hardy–Weinberg equilibrium for the loci sampled. By contrast, only one population of *X. maculatus* and only three of eight populations of *X. clemenciae* were in Hardy–Weinberg equilibrium. However, when these populations are examined locus by locus, only a few loci (one or two) are significantly different from what is expected under HWE, and it is not always the same loci; therefore, these results are not expected to reflect any important demographic effects.

mtDNA. A large number of individuals from most populations of *X. maculatus* and *X. clemenciae* share the most abundant haplotype for their respective species: *X. hellerii* shows two distinct major haplotypes (Fig. 3). The sampled populations of each species exhibit a range of low to moderate haplotype richness values, and nucleotide diversities were low in all populations (Table 2) (see also Gutiérrez-Rodríguez *et al.* 2007 for other *Xiphophorus*, and von der Heyden *et al.* 2010 for marine fish examples). The mean haplotype richness across populations was not significantly different between species (post hoc, Bonferroni, $P > 0.05$).

Genetic differentiation within and between species

Using eight nuclear microsatellite markers and mitochondrial control region sequences, we detected highly statistically significant levels of genetic differentiation between most *Xiphophorus* populations, both within and between species (Table 3). Co-occurring species groups sampled at the same geographic locations were highly significantly differentiated (e.g. site 7, Table 3, Fig. 2). In general, the F_{ST} values are extremely high, especially

Table 2 Summary of the level of genetic diversity in sampled populations of *Xiphophorus clemenciae*, *Xiphophorus hellerii* and *Xiphophorus maculatus* at both eight microsatellite loci and mtDNA control region.

Species		<i>X. clemenciae</i>							
Population		1	2	3	4	5	6	7	8
Microsatellites									
Mean allelic richness		4.599	3.720	6.264	6.860	6.728	6.948	6.247	5.156
H-W tests (<i>P</i>)		0.453	0.080	0.255	0.003	<0.001	<0.001	<0.001	0.505
<i>H_E</i>		0.535(±0.360)	0.508(±0.309)	0.633(±0.385)	0.672(±0.370)	0.726(±0.308)	0.705(±0.337)	0.649(±0.318)	0.619(±0.291)
<i>H_O</i>		0.536(±0.376)	0.479(±0.303)	0.609(±0.381)	0.633(±0.366)	0.691(±0.429)	0.725(±0.380)	0.641(±0.356)	0.619(±0.294)
MtDNA									
Number of haplotypes		1	2	1	1	3	4	3	2
No. of polymorphic sites		0	1	0	0	2	5	4	1
Haplotype richness (HR)		0	0.604	0	0	1.070	1.778	1.786	0.313
Nucleotide diversity (τ)		0	0.00033	0	0	0.00060	0.00287	0.00223	0.00016
Species		<i>X. hellerii</i>							
Population		9	10	11	12	13	14	15	16
Microsatellites									
Mean allelic richness		6.626	6.325	6.605	4.997	7.692	7.579	8.550	4.600
H-W tests (<i>P</i>)		0.443	0.996	0.540	0.041	0.352	0.919	0.532	0.917
<i>H_E</i>		0.746(±0.210)	0.654(±0.351)	0.671(±0.342)	0.668(±0.168)	0.811(±0.165)	0.782(±0.245)	0.838(±0.077)	0.628(±0.282)
<i>H_O</i>		0.742(±0.250)	0.693(±0.367)	0.650(±0.321)	0.658(±0.150)	0.807(±0.171)	0.821(±0.262)	0.787(±0.102)	0.631(±0.303)
MtDNA									
Number of haplotypes		2	2	1	1	1	2	3	1
No. of polymorphic sites		1	1	0	0	0	1	2	0
Haplotype richness (HR)		0.894	0.455	0	0	0	0.357	1.667	0
Nucleotide diversity (τ)		0.00058	0.00023	0	0	0	0.00018	0.00098	0
Species		<i>X. maculatus</i>							
Population		17	15	18	19	20	21	20	21
Microsatellites									
Mean allelic richness		7.979	7.306	10.970	11.743	11.989	11.637	11.989	11.637
H-W tests (<i>P</i>)		0.583	0.060	<0.001	0.006	<0.001	<0.001	<0.001	<0.001
<i>H_E</i>		0.853(±0.075)	0.776(±0.142)	0.927(±0.037)	0.942(±0.022)	0.941(±0.013)	0.938(±0.027)	0.941(±0.013)	0.938(±0.027)
<i>H_O</i>		0.874(±0.088)	0.793(±0.184)	0.922(±0.063)	0.893(±0.078)	0.868(±0.043)	0.917(±0.062)	0.868(±0.043)	0.917(±0.062)
MtDNA									
Number of haplotypes		1	1	3	2	3	4	3	4
No. of polymorphic sites		0	0	3	1	4	5	4	5
Haplotype richness (HR)		0	0	1.778	0.641	0.905	0.714	0.905	0.714
Nucleotide diversity (τ)		0	0	0.00116	0.00045	0.00115	0.00092	0.00115	0.00092

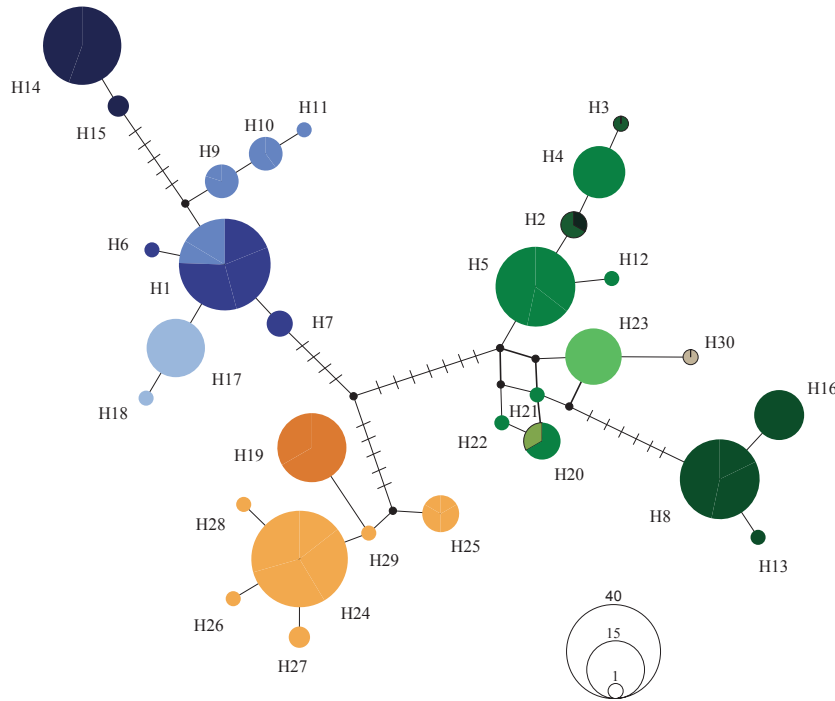


Fig. 3 A haplotype network of mtDNA sequences show that there are no shared haplotypes between species, even when found at the same sites. To enable clear comparisons, haplotype colours are designated by population assignment as found for the microsatellite data using *STRUCTURE* (see Fig. 5); note that there are also additional *Xiphophorus hellerii* populations in the haplotype analysis, see below. The southern most populations of *X. hellerii* (H8, H13 and H16) and *Xiphophorus clemenciae* (H14 and H15) show distinct haplotypes that differ from the rest of their respective species. (As noted in the text, *Xiphophorus hellerii* populations – where the haplotypes denoted H2 (site 3 and 4), H3 (site 4), H20 [site 17 (also site 15)], H30 (site 18) were found – consist of small numbers of *X. hellerii* samples found at the same sites as the other species (see also Table 1 and Fig. 2); these samples were included in the haplotype analysis to provide additional same site comparisons). Haplotype circle size is equivalent to haplotype frequency. Small black circles represent missing or inferred haplotypes. Connecting line distances are proportional to sequence divergence between haplotypes (2 mm = 1 bp difference) and short line strokes represent mutational steps between haplotypes.

those estimated using the mitochondrial sequences where some comparisons yielded values as high as 1.00 documenting a complete lack of gene flow between some populations both within and between species. However, some F_{ST} values, although very high, are not significant after Bonferroni correction (Table 3). As expected under isolation by distance (IBD, see below), geographically proximal populations of the same species, such as *X. clemenciae* populations three and four (Table 3, Fig. 2), were more genetically similar at both marker types.

Interestingly, geographically distant populations of all species have unique haplotypes, distinct from the rest of their respective species (Fig. 3). This suggests these populations are geographically isolated populations, albeit the sequence differences are less than those between species (i.e. average within genus distances of 207 marine fish 9.93%, Ward *et al.* 2005; for other teleostean fishes, 1.0–5.6% within species between drainages, Fajen & Breden 1992 5.089% between subspecies clades, Salzburger *et al.* 2003 5.29% between clades

within a marine species, Stefanni & Thorley 2003 0.11–1.3% within species (*X. cortezi*) between drainages, Gutiérrez-Rodríguez *et al.* 2007). Specifically, populations of *X. hellerii* and *X. clemenciae* sampled from the southern most region of the IT show the highest amount of within-species sequence divergence. The southern populations of *X. clemenciae* differ from populations in the central IT (closest haplotype) by 0.62% (Figs 2 and 3). The southern most populations of *X. hellerii* are slightly more distinct from their central IT counterparts differing by 0.74% sequence divergence (Figs 2 and 3).

Population differentiation significantly increased with distance within all examined species ($P < 0.01$, Mantel test, Fig. 4). Distance explained a high percentage of the variation in all species (*X. clemenciae*: $R^2 = 0.54$, *X. hellerii*: $R^2 = 0.40$ and *X. maculatus*: $R^2 = 0.72$). This differentiation is also reflected in the AMOVA results for all species using both marker types (Table 4). In other words, the percentage of variation among populations within geographic groups tends to be low when geographically proximal populations are grouped together

Table 3 Estimated pairwise F_{st} (microsatellite data, below diagonal, mtDNA data, above diagonal) values by species and geographic location.

Species	Population	<i>Xiphophorus clemenciae</i>										<i>Xiphophorus hellerii</i>										<i>Xiphophorus maculatus</i>									
		1	2	3	4	5	6	7	8	9	10	11	12	13	7	14	15	16	17	18	19	20	21								
<i>X. clemenciae</i>	1		0.084 ^{ns}	1.000	1.000	0.971	0.885	0.920	0.994	0.994	0.998	1.000	1.000	1.000	1.000 ^{ns}	0.998	0.994	1.000	1.000	1.000	0.985	0.980	0.979	0.983							
	2	0.230		0.977	0.980	0.954	0.861	0.893	0.982	0.991	0.994	0.996	0.995	0.995	0.993 ^{ns}	0.993	0.988	0.995	0.995	0.993	0.976	0.973	0.972	0.976							
	3	0.175	0.215		0.000 ^{ns}	0.028 ^{ns}	0.271 ^{ns}	0.433 ^{ns}	0.957	0.988	0.996 ^{ns}	1.000	1.000 ^{ns}	1.000 ^{ns}	1.000 ^{ns}	0.995	0.984 ^{ns}	1.000	1.000	1.000 ^{ns}	0.965 ^{ns}	0.960 ^{ns}	0.958	0.966							
	4	0.163	0.202	0.008 ^{ns}		0.066 ^{ns}	0.327 ^{ns}	0.498 ^{ns}	0.962	0.989	0.997	1.000	1.000	1.000 ^{ns}	1.000 ^{ns}	0.995	0.987	1.000	1.000	1.000 ^{ns}	0.971	0.966	0.963	0.970							
	5	0.196	0.206	0.098	0.080		-0.355 ^{ns}	0.489 ^{ns}	0.875	0.982	0.988	0.991	0.986	0.984	0.979	0.983	0.976	0.988	0.986	0.982	0.957	0.955	0.954	0.960							
	6	0.259	0.279	0.187	0.169	0.107		0.075 ^{ns}	0.811	0.972	0.975	0.980	0.968	0.961	0.944 ^{ns}	0.965	0.954 ^{ns}	0.973	0.968	0.952 ^{ns}	0.922 ^{ns}	0.931	0.933	0.941							
	7	0.304	0.318	0.228	0.210	0.152	0.026		0.712	0.962	0.965	0.970	0.951	0.943	0.924 ^{ns}	0.949	0.941	0.958	0.951	0.931	0.904	0.916	0.919	0.927							
<i>X. hellerii</i>	8	0.208	0.264	0.174	0.154	0.132	0.199	0.233		0.990	0.995	0.998	0.995	0.996	0.995	0.993	0.988	0.997	0.996	0.995	0.974	0.970	0.968	0.973							
	9	0.284	0.300	0.203	0.186	0.139	0.171	0.211	0.223		0.570 ^{ns}	0.641 ^{ns}	0.985	0.981	0.977 ^{ns}	0.980	0.953	0.977	0.993	0.991	0.975	0.973	0.972	0.975							
	10	0.310	0.320	0.229	0.210	0.172	0.175	0.219	0.251	0.093		0.016 ^{ns}	0.994	0.993	0.991 ^{ns}	0.989	0.961	0.992	0.997	0.996	0.977	0.974	0.972	0.976							
	11	0.296	0.310	0.218	0.203	0.164	0.173	0.214	0.239	0.088	0.001 ^{ns}		1.000	1.000	1.000 ^{ns}	0.995	0.974	1.000	1.000	1.000	0.984	0.979	0.977	0.981							
	12	0.346	0.399	0.309	0.292	0.267	0.281	0.309	0.326	0.235	0.289	0.282		1.000	1.000 ^{ns}	-0.977	0.951 ^{ns}	1.000	1.000	1.000	0.981 ^{ns}	0.976	0.974	0.978							
	13	0.276	0.313	0.237	0.222	0.188	0.201	0.219	0.251	0.171	0.223	0.218	0.157		0.000 ^{ns}	-0.026 ^{ns}	0.918 ^{ns}	1.000	1.000	1.000 ^{ns}	0.976	0.971	0.969	0.974							
	7	0.286	0.309	0.217	0.201	0.149	0.173	0.199	0.230	0.144	0.212	0.206	0.163	0.015 ^{ns}	0.097 ^{ns}		0.881 ^{ns}	1.000	1.000 ^{ns}	0.966 ^{ns}	0.962 ^{ns}	0.960 ^{ns}	0.967								
<i>X. maculatus</i>	14	0.320	0.345	0.257	0.240	0.197	0.213	0.229	0.274	0.192	0.247	0.241	0.215	0.039	0.031		0.911	0.978	0.997	0.996	0.976	0.972	0.970	0.974							
	15	0.278	0.320	0.206	0.178	0.138	0.166	0.198	0.223	0.115	0.187	0.183	0.157	0.044	0.027 ^{ns}	0.060		0.915	0.993	0.989 ^{ns}	0.964 ^{ns}	0.964	0.964	0.969 ^{ns}							
	16	0.394	0.414	0.335	0.315	0.290	0.294	0.326	0.333	0.265	0.331	0.321	0.306	0.234	0.219	0.259	0.177		0.915	1.000	0.989 ^{ns}	0.964 ^{ns}	0.964	0.964	0.980						
	17	0.291	0.297	0.229	0.217	0.170	0.201	0.225	0.244	0.177	0.218	0.213	0.204	0.150	0.120	0.150	0.103	0.243		1.000	1.000 ^{ns}	0.983	0.979	0.977	0.980						
	15	0.347	0.358	0.269	0.247	0.203	0.223	0.255	0.287	0.193	0.243	0.241	0.235	0.168	0.140	0.167	0.118	0.267	0.070		0.870 ^{ns}	-0.857	-0.851	-0.876							
	18	0.261	0.273	0.178	0.168	0.138	0.168	0.193	0.208	0.133	0.183	0.178	0.163	0.109	0.078	0.112	0.066	0.197	0.063	0.085		0.099 ^{ns}	0.084 ^{ns}	0.016 ^{ns}							
	19	0.253	0.261	0.174	0.164	0.127	0.157	0.184	0.199	0.118	0.167	0.163	0.155	0.098	0.069	0.101	0.055	0.185	0.056	0.077	0.012		-0.070 ^{ns}	-0.014 ^{ns}							
20	0.262	0.271	0.183	0.170	0.128	0.157	0.185	0.201	0.120	0.171	0.165	0.158	0.094	0.062	0.099	0.052	0.195	0.053	0.076	0.009	-0.003	-0.018 ^{ns}									
21	0.253	0.265	0.174	0.163	0.133	0.160	0.186	0.203	0.123	0.171	0.166	0.154	0.098	0.069	0.103	0.057	0.193	0.060	0.078	0.009 ^{ns}	0.000	-0.003									

Non-significant ($P > 0.05$) pairwise values designated as NS, all other values are significant ($P < 0.05$ after Bonferroni correction).

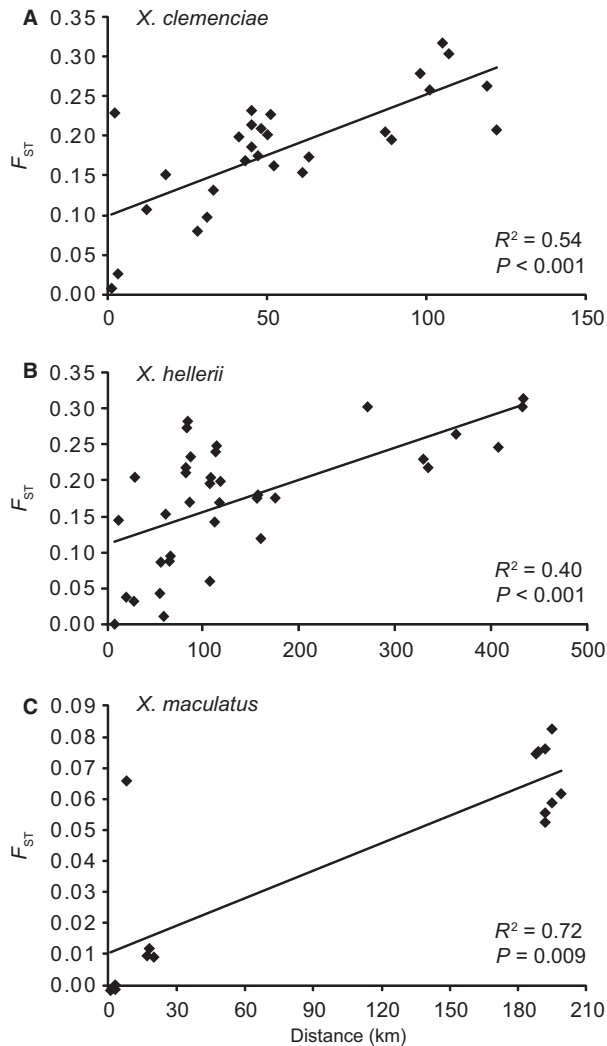


Fig. 4 Isolation by distance in *Xiphophorus clemenciae* (A), *Xiphophorus hellerii* (B) and *Xiphophorus maculatus* (C) populations in the Isthmus of Tehuantepec, Mexico.

(such as *X. clemenciae* populations five and six in the central IT, Fig. 2) compared to when geographically more disparate populations are grouped together (such as when more southern or northern IT populations are grouped with central IT populations, Fig. 2).

Structure analysis of microsatellite data

To evaluate the species and population differentiation at a broader scale, we used STRUCTURE (Pritchard *et al.* 2000) (microsatellite data). Using STRUCTURE, given the high level of genetic differentiation detected even within species in our pairwise comparisons (Table 2), we first estimated the population structure within species and then used the information generated on the most likely K values to conduct a STRUCTURE analysis incorporating both the putative hybrid and parental

species. The most likely genetic clusters for the individual species are two for *X. maculatus*, three for *X. hellerii* and four for *X. clemenciae*. Using the number of clusters detected in the individual species analyses, and with the indication of an increase in $\ln P(D)$ at $K = 9$ – the criterion in STRUCTURE used to detect the true K (the posterior probability of the data for a given K) – we estimate that the most likely number of genetic clusters among all species is nine (Fig. 5). This clustering provides no clear evidence of a single cluster being comprised of two different species. In addition, our estimations show the formation of single clusters for some geographically close populations of the same species, and also high differentiation between populations of the same species.

No evidence for current hybridization

To determine whether hybridization and introgression are ongoing between our study species, we used both mitochondrial haplotype network and STRUCTURE (microsatellite) analyses. Both analyses revealed similar differentiation within and between populations for each species and provided no strong evidence for admixture or hybridization between species (Figs 3 and 5). As an example of the similar differentiation detected between marker types, the southern most populations of *X. hellerii* (populations 9, 10 and 11) share a single main haplotype (H14, Fig. 3) and form a single cluster using STRUCTURE (Fig. 5). The differentiation shown in both analyses reflects the geographic distribution of the study populations (Fig. 2). As noted above, the most common *X. clemenciae* haplotype was found in most populations (five out of eight) from which *X. clemenciae* was sampled (H1, Fig. 3), and all of these populations occur in the central IT region (Fig. 2), suggesting a single origin for *X. clemenciae* in the central IT followed by moderate dispersal.

The haplotype network analysis strongly suggests a lack of current hybridization because no shared haplotypes between the putative hybrid and maternal species are found in the entire range of *X. clemenciae*. This also holds for all species even when they were collected at those sites where currently more than one species is found (Fig. 3) (note – *X. hellerii* populations where the haplotypes denoted H2 (site 3 and 4), H3 (site 4), H20 [site 17 (also site 15)], H30 (site 18) were found, consist of small numbers of *X. hellerii* samples found at the same sites as the other species (see also Table 1 and Fig. 2), these samples were included in the haplotype analysis to provide additional same site comparisons). Also all samples could be readily assigned to their species based on phenotypic characteristics (Kallman *et al.* 2004; Kallman, personal communication). We

Table 4 Hierarchical AMOVA of the genetic structure of populations and geographically distinct population groups of *Xiphophorus maculatus*, *Xiphophorus hellerii* and *Xiphophorus clemenciae*

Species	Locus	Source of variation	d.f.	Sum of squares	Variance components	Variance (%)	Fixation indices	P
<i>X. maculatus</i>	mtDNA control region	Among geographically distinct population groups	1	61.018	2.03419 Va	86.12	$F_{CT} = 0.8612$	0.067
		Among populations within groups	4	0.661	-0.01491 Vb	-0.63	$F_{SC} = -0.0455$	0.821
		Within populations	67	22.963	0.34273 Vc	14.51	$F_{ST} = 0.8549$	<0.001
		Total	72	84.642	2.36201			
	Microsatellites	Among geographically distinct population groups	1	27.964	0.16593 Va	4.44	$F_{CT} = 0.0444$	0.067
		Among populations within groups	4	25.494	0.04432 Vb	1.19	$F_{SC} = 0.0124$	<0.001
		Within populations	382	1346.034	3.52365 Vc	94.37	$F_{ST} = 0.0563$	<0.001
	Total	387	1399.492	3.73389				
<i>X. hellerii</i>	mtDNA control region	Among geographically distinct population groups	2	346.538	5.06919 Va	86.44	$F_{CT} = 0.8644$	<0.001
		Among populations within groups	6	46.318	0.70855 Vb	12.08	$F_{SC} = 0.8907$	<0.001
		Within populations	92	8.003	0.08699 Vc	1.48	$F_{ST} = 0.9852$	<0.001
		Total	100	400.859	5.86473			
	Microsatellites	Among geographically distinct population groups	2	268.639	0.5738 Va	15.20	$F_{CT} = 0.1520$	<0.001
		Among populations within groups	6	126.202	0.2836 Vb	7.51	$F_{SC} = 0.0885$	<0.001
		Within populations	613	1788.651	2.9178 Vc	77.29	$F_{ST} = 0.2271$	<0.001
	Total	621	2183.492	3.7753				
<i>X. clemenciae</i>	mtDNA control region	Among geographically distinct population groups	3	193.692	2.55975 Va	79.59	$F_{CT} = 0.7960$	0.029
		Among populations within groups	4	17.753	0.34803 Vb	10.82	$F_{SC} = 0.5302$	<0.001
		Within populations	88	27.134	0.30835 Vc	9.59	$F_{ST} = 0.9041$	<0.001
		Total	95	238.579	3.21613			
	Microsatellites	Among geographically distinct population groups	3	267.158	0.3954 Va	12.54	$F_{CT} = 0.1254$	0.01
		Among populations within groups	4	100.329	0.29312 Vb	9.30	$F_{SC} = 0.1063$	<0.001
		Within populations	636	1567.453	2.4645 Vc	78.16	$F_{ST} = 0.2183$	<0.001
	Total	643	1934.939					

conclude that current hybridization or introgression between the species does not occur.

Similar to the results of the mtDNA haplotype network analysis, the STRUCTURE analyses that would be able to identify nuclear genetic admixture between *X. clemenciae*, *X. hellerii* and *X. maculatus* showed strong evidence for a lack of hybrid genotypes (i.e. individuals for which Q values indicated a mixed genotype) (Fig. 5). Rather, most individuals were entirely representative of their respective species to which they were

assigned phenotypically. Only a single individual from *X. clemenciae* population four appears to be genotypically assigned to a mixture of both *X. maculatus* and *X. hellerii* (with consistent assignment probabilities across runs of approximately 32.6% and 46.4%, respectively, Fig. 5). In our species-restricted analysis, the same individual had a high probability of being from an entirely different group.

In addition, both our population-level (Fig. 5) and species-restricted analyses (data not shown) indicated a

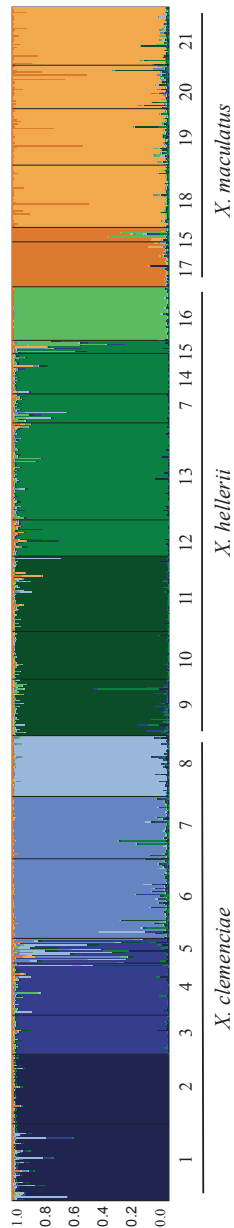


Fig. 5 Broad-scale assignment tests for the putative hybrid and parental *Xiphophorus* species using microsatellite markers in the software STRUCTURE. Tests were conducted for each individual species (data not shown) and for all species combined. Each vertical bar represents one individual, and the colour the proportion of its sampled genome that resembles a given genetic grouping. Populations of the same colour are most closely related compared to the other populations. Populations are ordered by geographic location (from south to north), and numbers below the figure indicate collection site (see Fig. 2). Within species, distinct genetic clusters are observed; however, geographically proximal populations tend to form single clusters. Populations of the different species do not form single genetic clusters even when found in the same geographic location (i.e. populations denoted by the same number). This analysis suggests there is negligible current introgression between the putative parental and hybrid species: see text.

small number of backcrossed individuals (two individuals from *X. clemenciae* population seven – assigned partially to a *X. hellerii* genotype – with assignment probabilities of approximately 17.8% and 26.8%, Fig. 5). Both analyses also assigned two individuals from the same population, but phenotypically *X. hellerii* (*X. hellerii* population seven), to a mixture of *X. hellerii* and *X. clemenciae* genotypes (with assignment probabilities of approximately 15.4% and 28.6%). Similarly, two phenotypically *X. maculatus* individuals (population 20 and 21) are assigned partially to an *X. hellerii* genotype for both analyses (with assignment probabilities of approximately 33.1% and 23.1%, Fig. 5). Individuals were noted as showing signs of backcrossing or introgression when they were identified in both analysis types as having mixed assignment with a contribution from, or probability of being from, an alternative genotype. However, in all cases, there are no phenotypic signs of mixing when photographs of each individual are examined for species characteristics. It should be noted that rare genotypes on occasion get poorly assigned by STRUCTURE and can lead to confusing or misleading cluster assignments – STRUCTURE is more effective at identifying large homogeneous genetic clusters (Kalinowski 2010).

Niche analyses

We utilized water parameter measurements along side additional environmental parameters to compare the ecological niches of *X. clemenciae* with those of its putative parental species. For our PCA, we chose seven factors to be included in the analysis. We used the first two orthogonal axes in the OMI analysis as they explained 97.92% of the total environmental variation (first axis, 85.85%; second axis, 12.07%). The niche analysis results (Fig. 6) reflect the geographical distribution of our study species across the IT. *X. maculatus*' niche is the farthest from the mean available habitat, while *X. hellerii*'s is closest. The two putative parental species have broad niches, while *X. clemenciae*, the putative hybrid, appears to have a much narrower niche with a smaller variance in habitat conditions compared to the variance of the average habitat (Fig. 6). Most of the observed niche variation of all three species is explained by three key parameters: dissolved oxygen, salinity and bank width (Fig. 7).

The niches of *X. maculatus* and *X. clemenciae* are significantly different from the mean available habitat (random Monte-Carlo-test, $P < 0.05$), while *X. hellerii*'s niche does not differ significantly from the mean available habitat ($P = 0.442$). Pianka's niche overlap index was used to gain an estimate of the percentage overlap of the niches of each species pair for each parameter

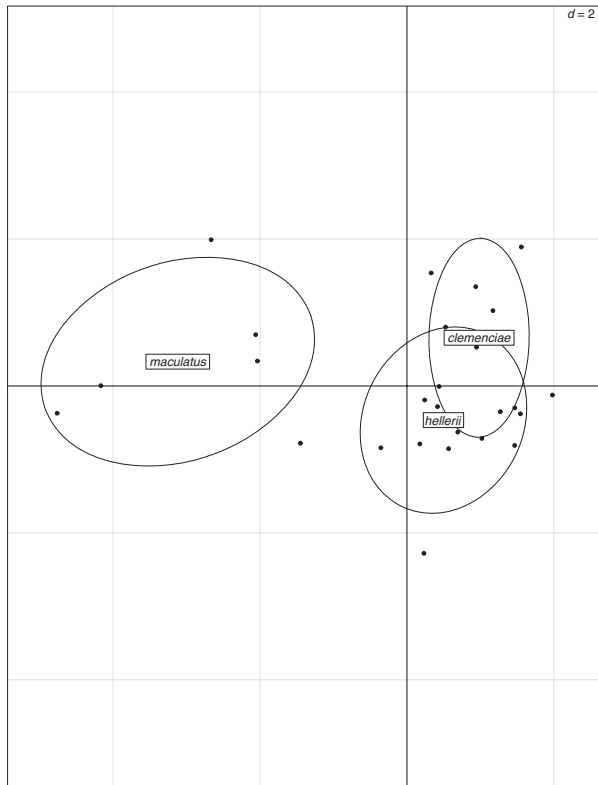


Fig. 6 Niches of *Xiphophorus clemenciae*, *Xiphophorus hellerii* and *Xiphophorus maculatus* in ecological space. The dots represent the different sample sites, and the origin of the axes is the mean available habitat. The circles represent the distribution of the species niches, and the centre of the circles represents the mean position of the species in ecological space. The niches of *X. hellerii* and *X. clemenciae* overlap and are close to the mean available habitat.

(Table 5). The niches of *X. clemenciae* and *X. hellerii* are most similar, while *X. clemenciae* and *X. maculatus* show the greatest differences and share 0% overlap for four of the nine parameters. The mean percentage overlap for all parameters combined for the different species comparisons are 25% overlap for *X. hellerii* and *X. clemenciae*, 16% for *X. hellerii* and *X. maculatus* and 14% for *X. maculatus* and *X. clemenciae*.

Discussion

We investigated the population genetic structure and ecological niches of 21 different populations of fish from three species of the genus *Xiphophorus*, in the IT, Mexico. We collected this information in an effort to determine the temporal nature of hybridization in an exemplar system. Specifically, we aimed to determine whether there is ongoing hybridization between *Xiphophorus hellerii*, *Xiphophorus maculatus* and *Xiphophorus clemenciae* because it has been suggested before that

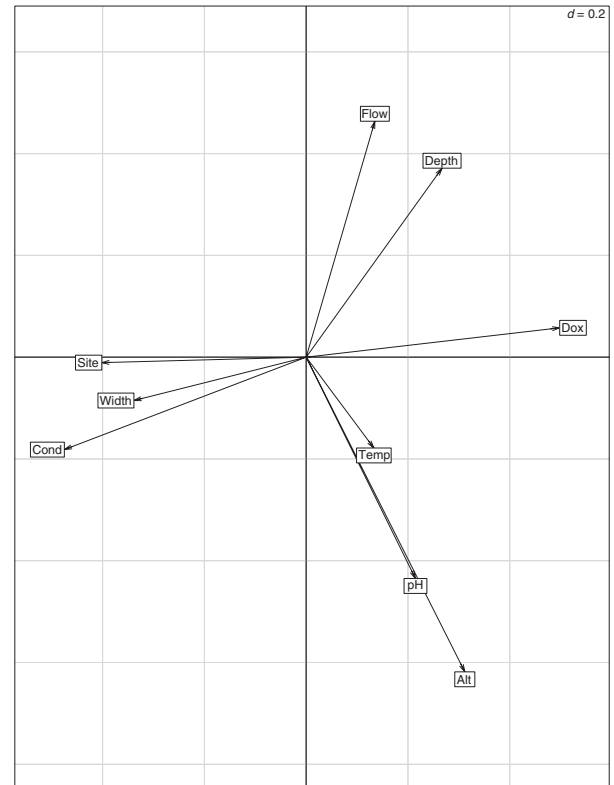


Fig. 7 Eigenvalues of the different parameters of the two axes of the niche analysis. The horizontal axis accounts for 87% of the variation and the vertical axis for 10.2%. Dissolved oxygen (dox), conductivity (cond) and bank width have the greatest influence on the species niches. The length and direction of the arrows indicate the relative loading of each parameter on the two principal components analysis axes.

Table 5 Pianka’s niche overlap index calculated for all pairwise species comparison. This index describes the percent overlap of two species niches for a single environmental parameter

Variable	<i>Xiphophorus hellerii</i> vs. <i>Xiphophorus maculatus</i>	<i>X. hellerii</i> vs. <i>Xiphophorus clemenciae</i>	<i>X. maculatus</i> vs. <i>X. clemenciae</i>
Altitude	0.020	0.128	0.000
Bank width	0.147	0.193	0.000
Conductivity	0.020	0.128	0.000
Dissolved oxygen	0.020	0.128	0.000
Flow velocity	0.782	0.808	0.778
pH	0.020	0.194	0.249
Stream depth	0.020	0.145	0.025
Temperature	0.276	0.249	0.100

X. clemenciae arose through hybridization between a maternal southern platyfish species, *X. maculatus*, and a paternal swordtail species, *X. hellerii*, and that female

preference, observed in the laboratory, has the potential to promote such hybridization. These laboratory results suggested that isolating mechanisms (IM) might not have evolved yet between the parental species and their hybrid species. We regarded this as a particularly interesting case where ecological characteristics and/or prezygotic IM might keep these species separate in nature, but do not apply in a laboratory setting. It is well known that the biological species concept faces operational challenges in the laboratory as many different taxa hybridize under such conditions but do not appear to do so in nature. Here, we set out to ask whether there is any current gene exchange detectable between this set of *Xiphophorus* fish species, where they occur in the same or different habitats.

Current evolutionary significant hybridization not detected

Both genetic marker types used in this study are fast evolving and therefore provide appropriate data for inferring recent population dynamics. Examining such genetic signatures of this fish system in an ecological and phylogenetic context allows us to obtain a more complete understanding of the initial hybridization process, subsequent differentiation and the evolution of prezygotic IM that now keep the two parental species apart from each other and their hybrid daughter species. Substantial genetic differentiation in nuclear markers and large divergence in mtDNA control region sequences between and within our putative parental (*X. hellerii* and *X. maculatus*) and hybrid (*X. clemenciae*) species, including co-occurring species groups, strongly suggest that there is virtually no currently ongoing hybridization between these species and that they are indeed good 'biological species'. Interestingly, now hundreds of thousands of generations after its origin, IM seem to have evolved both between the parental species and their daughter species. Using pairwise population analyses, STRUCTURE and haplotype network analyses, we did not detect significant levels of gene flow or shared haplotypes across species. These species are genetically distinct, and only few genetically mixed individuals (7 of 840) were found in our sample of 840 individuals throughout the entire distribution of *X. clemenciae*. Only the structure analyses detected this small number of putatively backcrossed and introgressed individuals. It is possible that these individuals actually reflect known limitations of STRUCTURE in designating rare genotypes (Kalinowski 2010). Alternatively, these rare individuals might indeed be due to current ongoing, but extremely infrequent, hybridization. If this alternative was true, the strong genetic clustering of the populations within and between species suggests this level of gene flow is

not enough to homogenize the gene pool of the different species. The main factors that could explain this outcome are hybridization is rare or very recent, and interestingly, selection may be acting against hybridization. At this point, we favour the first possibility given the extremely small number of mixed individuals detected, but do not entirely dismiss the possibility of rare and isolated hybridization and introgression.

Finding only negligible levels of current hybridization provides further support for a singular, much more ancient event having given rise to *X. clemenciae*, as suggested by the incongruence in the phylogenetic relationships using different genetic markers, and the divergence of those sequences (Meyer *et al.* 1994, 2006). Further, it is likely that *X. clemenciae* has a single origin in the central IT, followed by dispersal, as the most common haplotype of *X. clemenciae* is still found in most populations of this species. These *X. clemenciae* populations, which share the most common haplotype, all occur in the central IT within a 20-km radius and tend to show higher genetic diversity and allelic richness (e.g. populations 5, 6 and 7, Table 2, Fig. 2). Our extensive sampling of the entire geographic distribution of *X. clemenciae* suggests that recent hybridization is a rare event.

The pronounced genetic differentiation between the three species is also reflected in their distinct niches. Although some overlap is detected in Pianka's niche overlap index, we find large differences between the niches of all species pairs (Table 5). In particular, *X. maculatus* and *X. clemenciae* show no overlap in their estimated niche space (Fig. 6), suggesting that divergence in adaptations to certain environmental parameters, such as the level of dissolved oxygen, may play a role in preventing ongoing introgression between these species. *Xiphophorus clemenciae* tends to occur in well-oxygenated, higher-altitude streams, whereas *X. maculatus* tends to occur in low-lying pools and ponds with little water flow and low levels of dissolved oxygen (see also Kallman *et al.* 2004). However, we note that despite extensive sampling of the current known geographic distribution of *X. clemenciae*, we found no instances where *X. clemenciae* was sympatric with *X. maculatus*. The latter suggests that the genetic differentiation found between the hybrid species and its putative maternal parental lineage may also be maintained purely by geographic isolation. The putative parental species' niche spaces are also nonoverlapping. However, these two species do co-occur at three of six *X. maculatus* sites both in the central and in the northern IT (Fig. 2), and each site was typical for *X. maculatus* sampling locations – lower-lying stagnant ponds. *Xiphophorus hellerii* and *X. clemenciae* also differ in their typical niches but still do show some overlap (Fig. 6).

These species were found to co-occur in the central IT at three of eight *X. clemenciae* sampling locations (Fig. 2). This finding lends support to the hypothesis on the nature of the hybridization event (Meyer *et al.* 2006), where backcrosses of the initial hybrid females are thought to have occurred more often with southern swordtail males than swordless platyfish males. We found a few individuals in our large sample that are possible backcrosses between *X. clemenciae* and *X. hellerii*. Thus, *X. clemenciae* is not only genetically more similar to the swordtail lineage in five of six nuclear loci examined to date (Meyer *et al.* 2006) and may continue to backcross occasionally with *X. hellerii*, but this species also shares a more similar ecological niche with *X. hellerii*.

Dissolved oxygen, one of the three most important environmental variables distinguishing niche space among species, can act as a proxy for turbidity, as turbidity and dissolved oxygen tend to be negatively correlated in stagnant water bodies (Kramer 1987). Therefore, the known preference of female platyfish for sworded males may be tempered in their current natural habitat, even when sworded males do co-occur, by high water turbidity and a reduction in the ability of females to distinguish sworded vs. nonsworded males. In this kind of habitat, other mechanisms, such as chemical cues, may play an important role in sexual communication. Female *Xiphophorus* are known to show mating preferences for chemical cues produced by conspecific males (Crapon de Caprona & Ryan 1990; McLennan & Ryan 1997, 1999, Fisher & Rosenthal 2010) despite preferences for visual cues of heterospecific males (Ryan & Wagner 1987; Crapon de Caprona & Ryan 1990; Basolo 1995a,b; Hankison & Morris 2003). Such chemical cues may be one of the proximate mechanisms that act to maintain prezygotic isolating barriers. If visual cues would dominate in mate choice, such turbid habitats might promote hybridization, but we do not find evidence for this in our study.

Hybridization may have significantly contributed to the current species diversity among *Xiphophorus* fish more generally. Although examples of populations of poeciliid species with high frequencies of hybrids are rare, there are at least three well-characterized cases of hybridization in the genus *Xiphophorus* (reviewed in Rosenthal & García de León 2011). Using both molecular and morphological data, it has been shown that *X. birchmanni* and *X. cortezi*, and northern platyfish and *X. nezahualcoytl* hybridize in the wild. A third system, which has received considerable attention in the research community, the *X. birchmanni*-*X. malinche* species complex, may represent a case of incipient hybrid speciation, as hybrids in some populations are reproductively isolated from the parental species (Rosenthal

& García de León 2011, Culumber *et al.* 2011). Taking the latter case as an example, *Xiphophorus* hybrid systems appear to conform to expectations of recombinational speciation via hybridization where the ecological niches of hybrids are partially separated from the parental habitat (Coyne & Orr 2004). *Xiphophorus birchmanni*-*X. malinche* hybrids have been suggested to outperform or at least equal parental species at intermediate elevations and for a number of ecologically important traits, both hybrids and parentals covary with habitat (Culumber *et al.* 2011; Rosenthal & García de León 2011). In general, hybridization within the genus *Xiphophorus* appears to have occurred over different timescales and provides an ideal opportunity for examining the genomic and phenotypic signatures and ecological circumstances that contributed to past and present hybridization events and for comparing these types of events within a single genus. More work, such as further estimations of the phylogenetic relationships of these fish, is required to determine the overall prevalence of hybridization in this genus.

Genetic differentiation and phylogeography

This study revealed high levels of genetic differentiation between populations within species (*X. clemenciae*, *X. hellerii* and *X. maculatus*). Using different lines of evidence, we show that this differentiation is best explained by isolation by distance caused by restricted gene flow. All species show significant isolation by distance between populations, and both our STRUCTURE and haplotype network analyses assign individuals and populations to geographically proximal groupings. The different populations also contain a range of haplotype and nucleotide diversities and microsatellite allelic richness differs between populations. In particular, within *X. clemenciae* and *X. maculatus*, populations sharing the main haplotype and occurring at the lowest elevations detected for these species in this study (*X. clemenciae*: 40–50 m above sea level vs. 120–229 m, *X. maculatus*: 1–7 m vs. 26 and 39 m) tend to show higher levels of genetic diversity and allelic richness than the other populations. Lower genetic diversity and allelic richness at higher elevation populations suggest these populations have smaller effective population sizes. In addition, the latter result reiterates the finding of a lack of gene flow between populations and suggests that any gene flow that does occur is unidirectional, from high to low elevations. Similarly, in our AMOVA, we find a much higher percentage of variation between groups than within groups when geographically adjacent populations are grouped together (Table 4). These results are consistent with results obtained by Tatarenkov *et al.* (2010) in their microgeographic study of *X. hellerii* from Belize. They

found significant IBD both at the level of individual demes within a creek, and at larger spatial scales of up to 20 km between populations. In the current study, it follows that only geographically proximal populations show no significant genetic differentiation (Table 3, Fig. 2). At a microgeographic level (i.e. between creeks, similar or smaller waterway distances to Tatarenkov *et al.* 2010), close physical proximity, despite potential barriers such as separate streams with no permanent waterway connection, appears to facilitate gene flow between populations. For example, *X. clemenciae* populations three and four (Fig. 2) were found in geographically close (~1 km), but separate streams, and show no significant genetic differentiation at either marker type (Table 2). Therefore, seasonal flooding and/or migration between these sites via connecting waterways through stream capture might provide a mechanism for dispersing individuals to occupy new streams. Interestingly, Tatarenkov *et al.*'s (2010) results show significant genetic differentiation even between adjacent pools within creeks in contrast to our results and actually are surprising in the light of the documented among-pool movement of swordtails.

Freshwater fish in general are known to have high levels of genetic divergence (e.g. in small freshwater fish; Crispo *et al.* 2006; Tatarenkov *et al.* 2010), especially in comparison with marine fish (reviewed by Ward *et al.* 1994; DeWoody & Avise 2000). However, large differentiation, between the geographically distant central and southern *X. hellerii* populations (e.g. populations 12, 13, 14 vs. 9, 10, 11, Fig. 2), potentially provides the first indication of incipient speciation within *X. hellerii*. We note that the southern populations of *X. hellerii* are geographically the furthest from the type specimen of *X. hellerii* that was found north of the IT near Orizaba, Veracruz, Mexico (Heckel 1848). Further phylogenetic, ecological and morphological investigations are required to determine whether these different populations represent distinct species.

Here, the divergence between the southern IT populations of *X. hellerii* (9, 10, 11, Fig. 2) and *X. clemenciae* (1, 2, Fig. 2) and the central IT populations of the same species (Fig. 2) appears to coincide with regional tectonic events in the southern IT (Barrier *et al.* 1998). Using the fastest and slowest rates of known calibrated molecular clocks in teleost fish, from 0.044 to 0.004 changes/site/Myr (Schories *et al.* 2009), the separation of the central and southern *X. clemenciae* populations, and the central and southern *X. hellerii* populations would have occurred between 0.14 and 1.55 and 0.16 and 1.85 Ma, respectively. In a north to south cross-section of the IT, the elevation in the northern half is low and then increases regularly from the centre of the IT to a small coastal range (~200 m in

elevation, Fig. 2) that borders the Tehuantepec Coastal Plain (on the Pacific coast side). On the southern side of this range, a steep slope leads down to the plain (Barrier *et al.* 1998). The distinct tectonic episodes resulting in the present structure of the IT have taken place since the Late Miocene (6 Ma), and morphological evidence indicates that faults in the Coastal Plain and at its northern border are still active (Barrier *et al.* 1998). Therefore, a plausible scenario is that the southern populations of *X. hellerii* (9, 10, 11, Fig. 2) and *X. clemenciae* (1, 2, Fig. 2) became isolated from the central and northern IT populations when the coastal range was uplifted, within the last 6 Myr, and more specifically possibly <2 Ma as suggested by the molecular data. The large degree of divergence between these populations suggests that even seasonal flooding does not facilitate mixing of these populations. In the broader context of the Poeciliid family, the latter divergence in *Xiphophorus* populations occurred long after the main radiation of Poeciliid fishes from their South American area of origin, for example, via the Aves land bridge to Middle America (approximately 44 Ma) (Hrbek *et al.* 2007). Poeciliids are thought to have dispersed at least three times between South America and Middle America, with the second dispersal event resulting in the main poeciliid radiation of Central America and the Antilles. The latter clade, which includes *Xiphophorus*, is statistically well supported and is nested within a paraphyletic predominantly South American group (Hrbek *et al.* 2007). The early colonizing events preceded the final closing of the Isthmus of Panama 3.5–5 Ma. Following the closing, vicariant events, such as tectonic episodes in the IT, have continued to shape the subclades of the main Central American radiation.

Here, other genetic divergences observed within the study species, that is, within *X. maculatus*, in the IT reflect the geographic distance between populations and may have also been influenced by tectonic activity, although such associations do not appear to be as distinct. Interestingly, although the geographic distance between the central and northern *X. maculatus* populations is very large (15, 17 vs. 18, 19, 20, 21, Fig. 2), the sequence divergence between these populations is less than that found between the populations described above (Fig. 3). This may indicate a combination of two factors: first, *X. maculatus* appears to have a broad niche suggesting ease of movement between sites; second, the northern IT is downstream of the central IT, due to regional subsistence since the Pliocene (Barrier *et al.* 1998). We find a higher level of genetic diversity in the coastal *X. maculatus* populations (18, 19, 20, 21, Table 2), suggesting unidirectional gene flow from south to north, and possibly reflecting larger population sizes in the northern populations (Frankham 1996) or

higher dispersal between lowland habitats. During our field collections, we noticed that *X. maculatus* was more readily sampled from the northern coastal plain than from the central IT sites, again suggesting this species is more abundant on the coastal plain.

Conservation implications – *X. clemenciae*

Data from the present study can provide preliminary information for conservation management decisions through basic sampling successes, as a proxy for abundance, and by describing the genetic differentiation among populations and the pattern of geographic connectivity.

Our study lends further support to the description of *X. clemenciae*'s distribution proposed by Kallman *et al.* (2004), where these authors conclude that the species is widespread and common, but only in the uplands of the Rio Coatzacoalcos basin in the IT. We readily sampled *X. clemenciae* from many sites within this area, suggesting it is abundantly present in this restricted region. However, our sampling results also support the previously reported restricted distribution of this species, and we did not find any of these fish in the northern region of the IT including the northern coastal plain. Consequently, we recommend that this species is surveyed regularly, and possible anthropogenic effects are studied in the populations within this restricted distribution. Such effects can be noticed adjacent to villages and farms where waste water runoff and other human uses of the streams appear to occur regularly (JC Jones, personal observation). This may be particularly relevant to populations of this species occurring in the south of the IT, which are genetically very distinct from the other populations and may constitute distinct population units.

The process of identifying and then managing evolutionary significant units (ESUs; Moritz 1994) is not straightforward (Moritz *et al.* 2002; Mesquita *et al.* 2005), but phenotypic difference and historical isolation processes have been highlighted as important considerations (Moritz 1994, 2002). In recent studies (e.g. Salgueiro *et al.* 2003; Mesquita *et al.* 2005), mtDNA and microsatellite information have been used to identify ESUs and management units (MUs; Moritz 1994) (Mesquita *et al.* 2005; Angienda *et al.* 2010). Here, the use of both mtDNA and microsatellite data in addition to geographical distributions provides information on distinct population units of this geographically restricted species. If one considers the criteria of phylogeographic differentiation (including sequence divergence) and the observed significant F_{ST} values for both mtDNA and microsatellite data, two main ESUs are apparent, population one and two, and the central IT

populations (populations three-eight) (Fig. 2). *Xiphophorus clemenciae* also exhibits the lowest level of average microsatellite allelic richness across populations, although this is not reflected by the mtDNA haplotype and nucleotide diversity of this species (Table 2). The southern populations of *X. hellerii* also represent distinct population units.

Conclusion

We investigated the population dynamics and the possibility of ongoing hybridization and introgression between the putative hybrid and parental species, *X. clemenciae*, *Xiphophorus hellerii* and *Xiphophorus maculatus*, respectively. No strong evidence for ongoing hybridization or introgression between these species was discovered although partially overlapping distributions of these species would permit gene flow. Even where the species co-occur at the same locality, our molecular data show clear and consistent genetic differences between these species, and only few individuals appear to show evidence for hybridization. Data on ecological niches indicate a large separation between *X. maculatus* in particular and the other two species of *Xiphophorus*.

Populations of the three species studied within the IT are genetically highly differentiated, except for geographically proximal populations. The southern most populations show the highest levels of differentiation, and this may reflect a broader trend in population dynamics in fishes and other freshwater organisms occurring in this region. This suggests that tectonic activity in the region has created a barrier that has caused geographic isolation of some populations of *Xiphophorus* from the rest of their species, where speciation may be a possible outcome.

Acknowledgements

We are grateful to J. Shields for help with fieldwork and M. van Toor for assistance with laboratory work and the niche analyses. We thank SEMANART for permission to conduct our work in Mexico. The Colegio de Postgraduados, Campus Cordoba, provided helpful logistical support. S. Fan, H. Gunter, F. Henning, P. Hull, A. Kautt, H.-J. Lee, G. Machado-Schiaffino and M. Schartl provided valuable comments and discussions on the manuscript. This work was supported by grants of the Deutsche Forschungsgemeinschaft (DFG) (JCJ and AM) the Zukunftskolleg (JCJ) and the University of Konstanz (AM).

References

- Alvarez J (1959) Nuevas especies de *Xiphophorus* e *Hyporhamphus* procedentes del Rio Coatzacoalcos. *Ciencia*, **19**, 69–73.

- Anderson E, Stebbins GL (1954) Hybridization as an evolutionary stimulus. *Evolution*, **8**, 378–388.
- Angienda PO, Lee HJ, Elmer KR, Abila R, Waindi EN, Meyer A (2011) Genetic structure and gene flow in an endangered native tilapia fish (*Oreochromis esculentus*) compared to invasive Nile tilapia (*Oreochromis niloticus*) in Yala swamp, East Africa. *Conservation Genetics*, **12**, 243–255.
- Arnold ML (1992) Natural hybridization as an evolutionary process. *Annual Review of Ecology and Systematics*, **23**, 237–261.
- Arnold ML (1997) *Natural Hybridisation and Evolution*. Oxford University Press, Oxford, UK.
- Arnold ML, Meyer A (2006) Natural hybridization in primates: one evolutionary mechanism. *Zoology*, **109**, 261–276.
- Avise JC (1994) *Molecular Markers, Natural History and Evolution*. Chapman and Hall, New York City, New York.
- Avise JC (2000) *Phylogeography – The History and Formation of Species*. Harvard University Press, Cambridge, Massachusetts.
- Avise JC (2004) *Molecular Markers, Natural History, and Evolution*, 2nd edn. Sinauer Associates, Sunderland, Massachusetts.
- Barrier E, Velasquillo L, Chavez M, Gaulon R (1998) Neotectonic evolution of the Isthmus of Tehuantepec (southeastern Mexico). *Tectonophysics*, **287**, 77–96.
- Barton NH (2001) The role of hybridization in evolution. *Molecular Ecology*, **10**, 551–568.
- Basolo AL (1990a) Female preference for male sword length in the green swordtail *Xiphophorus hellerii* (Pisces: Poeciliidae). *Animal Behaviour*, **40**, 332–338.
- Basolo AL (1990b) Female preference predates the evolution of the sword in swordtail fish. *Science*, **250**, 808–810.
- Basolo AL (1995a) A further examination of a preexisting bias favoring a sword in the genus *Xiphophorus*. *Animal Behaviour*, **50**, 365–375.
- Basolo AL (1995b) Phylogenetic evidence for the role of a pre-existing bias in sexual selection. *Proceedings of the Royal Society of London B*, **259**, 307–311.
- Basolo AL (1997) Evolutionary change in a receiver bias: a comparison of female preference functions. *Proceedings of the Royal Society of London B*, **265**, 2223–2228.
- Belkhir P, Borsa P, Chikhi L, Raufaste N, Bonhomme F (2004) GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations, Laboratoire Génome, Populations, Interactions, Université de Montpellier II, Montpellier.
- Brookfield JFY (1996) A simple new method for estimating null allele frequency from heterozygote deficiency. *Molecular Ecology*, **5**, 453–456.
- Coyne JA, Orr HA (2004) *Speciation*. Sinauer, Sunderland.
- Crapon de Caprona M-D, Ryan MJ (1990) Conspecific mate recognition in swordtails, *Xiphophorus nigrensis* and *X. pygmaeus* (Poeciliidae): olfactory and visual cues. *Animal Behaviour*, **39**, 290–296.
- Crispo E, Bentzen P, Reznick DN, Kinnison MT, Hendry AP (2006) The relative influence of natural selection and geography on gene flow in guppies. *Molecular Ecology*, **15**, 49–62.
- Culumber ZW, Fisher HS, Tobler M *et al.* (2011) Replicated hybrid zones of *Xiphophorus* swordtails along an elevational gradient. *Molecular Ecology*, **20**, 342–356.
- Darwin CR (1871) *The Descent of Man, and Selection in Relation to Sex*. Murray, London.
- DeWoody JA, Avise JC (2000) Microsatellite variation in marine, freshwater and anadromous fishes compared with other animals. *Journal of Fish Biology*, **56**, 461–473.
- Doledec S, Chessel D, Gimaret-Carpentier C (2000) Niche separation in community analysis: a new method. *Ecology*, **81**, 2914–2927.
- Dowling TE, Secor C (1997) The role of hybridization and introgression in the diversification of animals. *Annual Review of Ecology and Systematics*, **28**, 593–619.
- Escoufier Y (1987) The duality diagram: a means of better practical applications. In: *Development in Numerical Ecology* (eds Legendre P, Legendre L), pp. 139–156. Springer Verlag, Berlin.
- Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, **14**, 2611–2620.
- Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0): an integrated software package for population genetic analysis. *Evolutionary Bioinformatics Online*, **1**, 47–50.
- Fajen A, Breden F (1992) Mitochondrial DNA sequence variation among natural populations of the Trinidad guppy. *Evolution*, **46**, 1457–1465.
- Fisher HS, Rosenthal (2010) Relative abundance of *Xiphophorus* fishes and its effect on sexual communication. *Ethology*, **116**, 32–38.
- Fisher RA (1930) *The Genetical Theory of Natural Selection*. Clarendon Press, Oxford.
- Frankham R (1996) Relationship of genetic variation to population size in wildlife. *Conservation Biology*, **10**, 1500–1508.
- Giraudoux P (2010). Pgirmiss: data analysis in ecology.
- Goudet J (1995) FSTAT (version 1.2): a computer program to calculate F-statistics. *Journal of Heredity*, **86**, 485–486.
- Gutiérrez-Rodríguez C, Morris MR, Dubois NS, de Queiroz K (2007) Genetic variation and phylogeography of the swordtail fish *Xiphophorus cortezi* (Cyprinodontiformes, Poeciliidae). *Molecular Phylogenetics and Evolution*, **43**, 111–123.
- Hankison SJ, Morris MR (2003) Avoiding a compromise between sexual selection and species recognition: female swordtail fish assess multiple species-specific cues. *Behavioral Ecology*, **14**, 282–287.
- Hayden B, Pulcini D, Kelly-Quinn M *et al.* (2010) Hybridisation between two cyprinid fishes in a novel habitat: genetics, morphology and life-history traits. *BMC Evolutionary Biology*, **10**, 169–180.
- Heckel JJ (1848) Eine neue Gattung von Poecilien mit rochenartigem Anklammerungs-Organ. Atlas zu der Reise im nördlichen Africa. Fische des Rothen Meeres. VI (pt1–5), 289–303.
- Houde AE (1997) *Sex, Color, and Mate Choice in Guppies*. Princeton University Press, Princeton, New Jersey.
- Hrbek T, Seckinger J, Meyer A (2007) A phylogenetic and biogeographic perspective on the evolution of poeciliid fishes. *Molecular Phylogenetics and Evolution*, **43**, 986–998.
- Hubbs CL (1955) Hybridisation between fish species in nature. *Systematic Zoology*, **4**, 1–20.
- Kalinowski ST (2010) The computer program STRUCTURE does not reliably identify the main genetic clusters within species: simulations and implications for human population structure. *Heredity*, **106**, 625–632.

- Kallman KD, Kazianis S (2006) The genus *Xiphophorus* in Mexico and Central America. *Zebrafish*, **3**, 271–285.
- Kallman KD, Walter RB, Morizot DC, Kazianis S (2004) Two new species of *Xiphophorus* (Poeciliidae) from the Isthmus of Tehuantepec, Oaxaca, Mexico, with a discussion of the distribution of the *X. clemenciae* clade. *American Museum Novitates*, **3441**, 1–34.
- Kirkpatrick M, Ryan MJ (1991) The evolution of mating preferences and the paradox of the lek. *Nature*, **350**, 33–38.
- Koblmüller S, Salzburger W, Sturmbauer C (2004) Evolutionary relationships in the sand-dwelling cichlid lineage of Lake Tanganyika suggest multiple colonization of rocky habitats and convergent origin of biparental mouthbrooding. *Journal of Molecular Evolution*, **60**, 277–289.
- Koblmüller S, Duftner N, Sefc KM *et al.* (2007) Reticulate phylogeny of gastropod-shell-breeding cichlids from Lake Tanganyika – the result of repeated introgressive hybridization. *BMC Evolutionary Biology*, **7**, 7–20.
- Koblmüller S, Egger B, Sturmbauer C, Sefc KM (2010) Rapid radiation, ancient incomplete lineage sorting and ancient hybridization in the endemic Lake Tanganyika cichlid tribe Tropheini. *Molecular Phylogenetics and Evolution*, **55**, 318–334.
- Kramer DL (1987) Dissolved oxygen and fish behavior. *Environmental Biology of Fishes*, **18**, 81–92.
- Lee WJ, Conroy J, Howell WH, Kocher TD (1995) Structure and evolution of teleost mitochondrial control regions. *Journal of Molecular Biology*, **40**, 1–13.
- Lewontin RC, Birch LC (1966) Hybridization as a source of variation for adaptation to new environments. *Evolution*, **20**, 315–336.
- Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, **25**, 1451–1452.
- Magurran AE (2005) *Evolutionary Ecology: The Trinidadian Guppy*. Oxford University Press, Oxford.
- Mallet J (2005) Hybridization as an invasion of the genome. *Trends in Ecology and Evolution*, **20**, 229–236.
- Mallet J (2007) Hybrid speciation. *Nature*, **446**, 279–283.
- Mayr E (1942) *Systematics and the Origin of Species*. Columbia University Press, New York City, New York.
- Mayr E (1963) *Animal Species and Evolution*. Belknap Press, Cambridge, Massachusetts.
- McLennan DA, Ryan MJ (1997) Responses to conspecific and heterospecific olfactory cues in the swordtail *Xiphophorus cortezi*. *Animal Behaviour*, **54**, 1077–1088.
- McLennan DA, Ryan MJ (1999) Interspecific recognition and discrimination based upon olfactory cues in northern swordtails. *Evolution*, **53**, 880–888.
- Mesquita N, Hänfling B, Carvalho GR, Coelho MM (2005) Phylogeography of the cyprinid *Squalius aradensis* and implications for conservation of the endemic freshwater fauna of southern Portugal. *Molecular Ecology*, **14**, 1939–1954.
- Meyer A (1997) The evolution of sexually selected traits in male swordtail fishes (*Xiphophorus*: Poeciliidae). *Heredity*, **79**, 329–337.
- Meyer A, Morrissey JM, Scharl M (1994) Recurrent origin of a sexually selected trait in *Xiphophorus* fishes inferred from molecular phylogeny. *Nature*, **386**, 539–542.
- Meyer A, Salzburger W, Scharl M (2006) Hybrid origin of a swordtail species (Teleostei: *Xiphophorus clemenciae*) driven by sexual selection. *Molecular Ecology*, **15**, 721–730.
- Moritz C (1994) Defining ‘evolutionary significant units’ for conservation. *TREE*, **9**, 373–375.
- Mortiz C (2002) Strategies to protect biological diversity and the evolutionary processes that sustain it. *Systematic Biology*, **51**, 238–254.
- Nosil P, Harmon LJ, Seehausen O (2009) Ecological explanations for (incomplete) speciation. *Trends in Ecology and Evolution*, **24**, 145–156.
- Nyingi DW, Agnèse J-F (2007) Recent introgressive hybridization revealed by exclusive mtDNA transfer from *Oreochromis leucostictus* (Trewavas, 1933) to *Oreochromis niloticus* (Linnaeus, 1758) in Lake Baringo, Kenya. *Journal of Fish Biology*, **70**, 148–154.
- van Oosterhout C, Hutchison WF, Wills DPM, Shipley P (2004) Micro-checker: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, **4**, 535–538.
- Petit RJ, El Mousadik A, Pons O (1998) Identifying populations for conservation on the basis of genetic markers. *Conservation Biology*, **12**, 844–855.
- Pianka ER (1973) The structure of lizard communities. *Annual Review of Ecology and Systematics*, **4**, 53–74.
- Posada D, Crandall K (1998) MODELTEST: testing the model of DNA substitution. *Bioinformatics*, **14**, 817–818.
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.
- Quattro JM, Vrijenhoek RC (1989) Fitness differences among remnant populations of the endangered sonoran topminnow. *Science*, **248**, 976–978.
- Quattro JM, Leberg PL, Douglas ME, Vrijenhoek RC (1996) Molecular evidence for a unique evolutionary lineage of endangered Sonoran desert fish (genus *Poeciliopsis*). *Conservation Biology*, **10**, 128–135.
- R Development Core Team. (2010) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Ramirez EM (1999) Taxonomia y zoogeografía de la ictiofauna dulceacuicola del Estado de Oaxaca, Mexico. Ph.D. thesis. Departamento de Ecología, Universidad de Barcelona.
- Reyer H-U (2008) Mating with the wrong species can be right. *Trends in Ecology and Evolution*, **23**, 289–292.
- Rosen DE (1960) Middle-American poeciliid fishes of the genus *Xiphophorus*. *Bulletin of the Florida State Museum*, **5**, 57–242.
- Rosenthal GG, Evans CS (1998) Female preference for swords in *Xiphophorus hellerii* reflects a bias for large apparent size. *Proceedings of the National Academy of Science USA*, **95**, 4431–4436.
- Rosenthal GG, García de León FJ (2011) Speciation and hybridization. In: *Ecology and Evolution of Poeciliid Fishes* (eds Schlupp I, Pilastro A, Evans J), University of Chicago Press, Chicago, Illinois.
- Rosenthal GG, Wagner WE Jr, Ryan MJ (2002) Secondary reduction of preference for the sword ornament in the pygmy swordtail *Xiphophorus nigrensis* (Pisces: Poeciliidae). *Animal Behaviour*, **63**, 37–45.
- Rousset F (2008) Genepop’007: a complete re-implementation of the genepop software for Windows and Linux. *Molecular Ecology Resources*, **8**, 103–106.

- Ryan MJ, Wagner WEJ (1987) Asymmetries in mating preferences between species: female swordtails prefer heterospecific males. *Science*, **236**, 595–597.
- Salgueiro P, Carvalho G, Collares-Pereira MJ, Coelho MM (2003) Microsatellite analysis of genetic population structure of the endangered cyprinid *Anaecypris hispanica*. Portugal: implications for conservation. *Biological Conservation*, **109**, 47–56.
- Salzburger W, Baric S, Sturmbauer C (2002a) Speciation via introgressive hybridization in East African cichlids? *Molecular Ecology*, **11**, 616–625.
- Salzburger W, Meyer A, Baric S, Verheyen E, Sturmbauer C (2002b) Phylogeny of the Lake Tanganyika cichlid species flock and its relationship to the central and east African haplochromine cichlid fish faunas. *Systematic Biology*, **51**, 113–135.
- Salzburger W, Brandstätter A, Gilles A *et al.* (2003) Phylogeography of the varione (*Leuciscus souffia*, Risso 1826) in Central Europe. *Molecular Ecology*, **12**, 2371–2386.
- Schartl M (1995a) Model systems for the study of genetic factors in cancer formation. *Naturwissenschaften*, **82**, 209–218.
- Schartl A, Dimitrijevic N, Schartl M (1994) Evolutionary origin and molecular biology of the melanoma-inducing oncogene of *Xiphophorus*. *Pigment Cell Research*, **7**, 428–432.
- Schelly R, Salzburger W, Kolbmüller S, Duftner N, Sturmbauer C (2006) Phylogenetic relationships of the lamprologine cichlid genus *Lepidolamprologus* (Teleostei: Perciformes) based on mitochondrial and nuclear sequences, suggesting introgressive hybridization. *Molecular Phylogenetics and Evolution*, **38**, 426–438.
- Schlupp I, Riesch R, Tobler M, Plath M, Parzefall J, Schartl M (2010) A novel, sexually selected trait in poeciliid fishes: female preference for mustache-like, rostral filaments in male *Poecilia sphenops*. *Behavioural Ecology and Sociobiology*, **64**, 1849–1855.
- Schories S, Meyer MK, Schartl M (2009) Description of *Poecilia (Acanthophaelus) obscura* n. sp., (Teleostei: Poeciliidae), a new guppy species from western Trinidad, with remarks on *P. wingei* and the status of the ‘‘Endler’s guppy’’. *Zootaxa*, **2266**, 35–50.
- Scribner KT, Page KS, Bartron M (2001) Hybridization in freshwater fishes: a review of case studies and cytonuclear methods of biological inference. *Reviews in Fish Biology and Fisheries*, **10**, 293–323.
- Seehausen O (2004) Hybridization and adaptive radiation. *Trends in Ecology and Evolution*, **19**, 198–207.
- Shen Y, Catchen J, Garcia T *et al.* (2012) Identification of transcriptome SNPs between *Xiphophorus* lines and species for assessing allele specific gene expression within F₁ interspecies hybrids. *Comparative Biochemistry and Physiology, Part C: Toxicology and Pharmacology*.
- Soltis PS, Soltis DE (2009) The role of hybridization in plant speciation. *Annual Review of Plant Biology*, **60**, 561–588.
- Stefanni S, Thorley JL (2003) Mitochondrial DNA phylogeography reveals the existence of an evolutionary significant unit of the sand goby *Pomatoschistus minutus* in the Adriatic (Eastern Mediterranean). *Molecular Phylogenetics and Evolution*, **28**, 601–609.
- Streelman JT, Gmyrek SL, Kidd MR *et al.* (2004) Hybridization and contemporary evolution in an introduced cichlid fish from Lake Malawi National Park. *Molecular Ecology*, **13**, 2471–2479.
- Swofford DL (2003) *PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods)*, Version 4. Sinauer Associates, Sunderland, Massachusetts.
- Tatarenkov A, Healey CIM, Grether GF, Avise JC (2008) Pronounced reproductive skew in a natural population of green swordtails, *Xiphophorus helleri*. *Molecular Ecology*, **17**, 4522–4534.
- Tatarenkov A, Healey CIM, Avise JC (2010) Microgeographic population structure of green swordtail fish: genetic differentiation despite abundant migration. *Molecular Ecology*, **19**, 257–268.
- Templeton AR (1981) Mechanisms of speciation – a population genetic approach. *Annual Review of Ecology and Systematics*, **12**, 23–48.
- Tripathi N, Hoffmann M, Willing E-M, Lanz C, Weigel D, Dreyer C (2009) Genetic linkage map of the guppy, *Poecilia reticulata*, and quantitative trait loci analysis of male size and colour variation. *Proceedings of the Royal Society of London B*, **276**, 2195–2208.
- von der Heyden S, Lipinski MR, Matthee A (2010) Remarkably low mtDNA control region diversity in an abundant demersal fish. *Molecular Phylogenetics and Evolution*, **55**, 1183–1188.
- Vrijenhoek RC, Douglas MC, Meffe GK (1985) Conservation genetics of endangered fish populations in Arizona. *Science*, **229**, 400–402.
- Wagner WE Jr, Ryan MJ (1987) Asymmetries in mating preferences between species: female swordtails prefer heterospecific males. *Science*, **236**, 595–597.
- Walsh PS, Metzger DA, Higuchi R (1991) Chelex® 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. *BioTechniques*, **10**, 506–513.
- Ward RD, Woodwark M, Skbinski DOF (1994) A comparison of genetic diversity levels in marine, freshwater and anadromous fishes. *Journal of Fish Biology*, **44**, 231–232.
- Ward RD, Zemlak TS, Innes BH, Last PR, Hebert PDN (2005) DNA barcoding Australia’s fish species. *Philosophical Transactions of the Royal Society B*, **360**, 1847–1857.
- Weir BS, Cockerham CC (1984) Estimating *F*-statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.
- Whitney KD, Ahern JR, Campbell LG, Albert LP, King MS (2010) Patterns of hybridization in plants. *Perspectives in Plant Ecology, Evolution and Systematics*, **12**, 175–182.
- Willing E-M, Bentzen P, Van Oosterhout C *et al.* (2010) Genome-wide single nucleotide polymorphisms reveal population history and adaptive divergence in wild guppies. *Molecular Ecology*, **19**, 968–984.
- Wright S (1943) Isolation by distance. *Genetics*, **28**, 114–138.
- Zhang Z, Wang Y, Wang S *et al.* (2011) Transcriptome analysis of female and male *Xiphophorus maculatus* Jp 163 A. *PLoS ONE*, **6**, e18379.

J.C.J. is a postdoctoral researcher focusing on hybridization and speciation in *Xiphophorus* fish. J.P. is a postdoctoral researcher who’s research primarily focuses on social insect biology. A.M. holds the chair in evolutionary biology at the University of Konstanz. He is an evolutionary biologist interested in genomic, developmental and morphological aspects of speciation.

Data accessibility

DNA sequences (all main MtDNA haplotypes): GenBank accessions JQ621886–JQ621898. Microsatellite data: DRYAD entry doi: 10.5061/dryad.524t4v16. Under the DRYAD entry, we provide a spreadsheet linking the mtDNA GenBank accession numbers to the microsatellite data and an alignment file for all main mtDNA haplotype sequences. Table S1 (Supporting information) provides all sampling site data.

Supporting information

Additional supporting information may be found in the online version of this article.

Table S1 GPS and site parameters recorded at each sampling site.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.