doi: 10.1111/j.1420-9101.2008.01645.x

Dispersal and population genetic structure of *Telmatherina antoniae*, an endemic freshwater Sailfin silverside from Sulawesi, Indonesia

R. P. WALTER, G. D. HAFFNER & D. D. HEATH

Great Lakes Institute for Environmental Research, University of Windsor, Windsor, ON, Canada

Keywords:

dispersal;
Lake Matano;
microsatellites;
population structure;
Telmatherinidae.

Abstract

Population genetic structure in the presence of substantial dispersal provides a unique perspective on the evolution of reproductive isolation. We sampled *Telmatherina antoniae*, an endemic fish species, at 10 sites in Lake Matano, Indonesia. Significant genetic structure (F_{ST} = 0.03) was found, despite a migration rate of 10.2% and a mean dispersal distance of 13.6 km, estimated by genotype assignment. Neither dispersal distance nor direction differed from random expectations, indicative of no dispersal barrier in Lake Matano. However, Bayesian genotype cluster assignment identified a population structure consisting of four to six clusters that did not coincide with sample site distribution, but explained two to three times more genetic variance than sample site. The mechanism for continued isolation of those genetic clusters is unknown, but assortative mating and temporal isolation are obvious candidates. Our results resolve the apparent paradox of population genetic structure coupled with frequent dispersal, and highlight the importance of considering cryptic genetic structure.

Introduction

Dispersal, the movement of individuals or their gametes from their site of origin, is a principal factor in the maintenance of genetic connectivity among populations through gene flow. Constraints on dispersal restrict gene flow resulting in the isolation of groups of individuals, thereby initiating population divergence and ultimately driving allopatric speciation (Bossart & Prowell, 1998; Bohonak, 1999).

Limits to dispersal are the result of extrinsic (physical barriers) or intrinsic (behavioural and/or ecological) factors. In general, genetic differentiation is more pronounced among conspecific populations of freshwater compared with marine species (Bilton *et al.*, 2002; Wong *et al.*, 2004), and is largely attributed to geographic barriers (Ward & Elliot, 2001). Among lake-dwelling fishes population-level genetic structure ranges from lake-wide panmixia to highly divergent local populations depending on species-specific biology and lake-specific limnological characteristics (e.g. lake size, bathymetry and productivity). Within-species population structuring in a single lake has been documented for a number of fish species including cichlids (Shaw *et al.*, 2000; Abila *et al.*, 2004), sardines (Hauser *et al.*, 1998), walleye (Stepian & Faber, 1998; DuPont *et al.*, 2007), arctic charr and other salmonids (Hendry *et al.*, 1998, 1999; Power *et al.*, 2005). In those studies, spawning site fidelity, habitat partitioning and microallopatry have been postulated as driving population structure.

In a number of African cichlid species, habitat fidelity and microallopatry appear to be important predictors of population genetic structure. Habitat barriers including deep water, sandy areas and even habitat discontinuity have been identified as barriers to gene flow (Danley *et al.*, 2000; Rico & Turner, 2002). In the African Lake Tanganyika, high cichlid population differentiation was noted across barriers as well as along stretches of continuous shoreline (Rüber *et al.*, 2001; Taylor *et al.*, 2001; Duftner *et al.*, 2006). Given this influence of restricted dispersal and microallopatry in the strong population genetic divergence and speciation of African cichlids, similar patterns may exist for other adaptively radiating freshwater species flocks.

The telmatherinid species complex of the Malili Lakes, Sulawesi, Indonesia, have gained recent attention

Correspondence: Ryan P. Walter, Great Lakes Institute for Environmental Research, University of Windsor, 401 Sunset Ave, Windsor, ON, Canada N9B 3P4.
Tel.: +1 519 253 3000 ext. 429; fax: +1 519 971 3616;
e-mail: walterj@uwindsor.ca

© 2008 THE AUTHORS. *J. EVOL. BIOL.* **22** (2009) 314–323
JOURNAL COMPILATION © 2008 EUROPEAN SOCIETY FOR EVOLUTIONARY BIOLOGY

14209101, 2009, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/j.1420-9101.2008.01645.x by Univ Of California Santa Cruz - UCSC, Wiley Online Library on [17/10/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

through evidence of ecological segregation (Roy *et al.*, 2007a,b), hybridization (Herder *et al.*, 2006) and adaptive radiation (Roy *et al.*, 2004; Herder *et al.*, 2006) and have been compared with African cichlids as models for the study of adaptive divergence and speciation (Herder *et al.*, 2006; Roy *et al.*, 2007a,b). Within the species flock are forms confined to a single lake, including *Telmatherina antoniae* Kottelat, endemic to Lake Matano (Kottelat *et al.*, 1993). Little is known about population genetic structure in the Lake Matano *telmatherinids*; however, the conspicuous littoral distribution of all seven described species indicates that the deep waters toward the centre of the lake may serve as a barrier to dispersal, although other, perhaps less obvious, barriers may also exist. Thus, philopatric behaviour of individuals may result in marked genetic divergence (and ultimately speciation) similar to that noted in rock-dwelling cichlids (Rüber *et al.*, 2001; Taylor *et al.*, 2001; Rico & Turner, 2002; Duftner *et al.*, 2006; Sefc *et al.*, 2007).

Here, we quantify population structure and dispersal in *T. antoniae* using multilocus genotyping at polymorphic microsatellite loci coupled with a systematic sampling of the known species range (i.e. Lake Matano). Adult *T. antoniae* form loose aggregations of 10–100 individuals and exhibit a patchy distribution as conspicuous members of Lake Matano's littoral fish community. However, aggregations do not appear to be associated with a distinct microhabitat. Furthermore, there is a lack of continuous littoral habitat throughout Lake Matano with some sites separated by deep drop-offs greater than 100 m (Fig. 1). Two questions form the focus of this study: (1) Given the presumed littoral distribution of this species, is there genetic structure and/or dispersal among Lake Matano's *T. antoniae* populations? (2) Is there cryptic genetic structuring in the Lake Matano *T. antoniae* that does not conform to the spatial distribution of the fish in the littoral zone? As *T. antoniae* are endemic to Lake Matano, and are part of a species radiation, characterization of the factors shaping population genetic structure may provide important insights into mechanisms of sympatric speciation and local adaptation.

Materials and methods

Collection and genetic analyses

Lake Matano, Sulawesi, Indonesia, is the hydrological head of the Malili Lakes and is characterized by a narrow and steep littoral zone (Roy, 2006; Fig. 1). Telmatherinids were captured by beach seine at depths ranging 0–3 m from 10 littoral sites in Lake Matano (Fig. 1) in June 2006. Adult 'torpedo' morph individuals with round-shaped second dorsal and anal fins were identified in the field as *T. antoniae* (following Kottelat, 1991, Kottelat *et al.*, 1993; Roy, 2006) and were sampled nonlethally by removing a small section of the anal fin tissue, which was preserved in

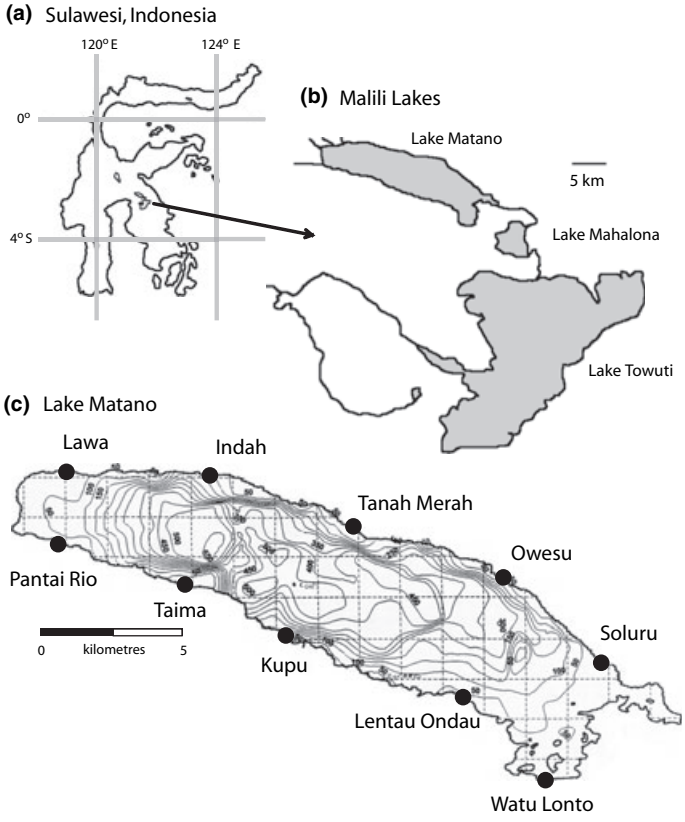

Fig. 1 Map of study area: (a) Sulawesi, Indonesia. (b) Malili Lakes in southern Sulawesi. (c) Lake Matano with sample sites indicated. Redrawn from Roy *et al.* (2007a).

14209101, 2009, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/j.1420-9101.2008.01645.x by Univ Of California Santa Cruz - UCSC, Wiley Online Library on [17/10/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

95% ethanol. The fish were released after a short recuperation period. DNA was recovered from tissue samples following the plate-based extraction method (Elphinstone *et al.*, 2003) and resuspended in 50 μL of Tris–EDTA buffer (10 mM Tris, 1.0 mM EDTA, pH 8.0).

To confirm our visual identification of *T. antoniae*, we genetically identified all fish based on mtDNA haplotype (Roy *et al.*, 2007a). We amplified a 1200-bp fragment of the cytochrome *b* gene using the forward primer GLUDG-5 (5′-TGACTTGAARAACCACCGTTG-3′) (Palumbi, 1992) and reverse primer CBtelm-R (5′-GTGGAGGAGGGG-TACGACTA-3′) (Roy *et al.*, 2007a) using the PCR conditions: initial denaturation at 95 °C for 2 min, followed by 35 cycles of 95 °C for 10 s, 60 °C for 10 s, 72 °C for 10 s, followed by a final extension at 72 °C for 1 min. Amplified *cyt b* fragments were then digested using PleI and FokI (New England Biolabs, MA, USA) to produce restriction fragment length polymorphism (RFLP) patterns diagnostic to *Telmatherina* mtDNA clades (following Roy, 2006). As introgressive hybridization has been documented in the adaptive radiation of Telmatherinid species complex (Herder *et al.*, 2006), any fish that did not conform to Roy's (2006) RFLP diagnostics were excluded from further analysis. Samples confirmed as *T. antoniae* were then genotyped at eight polymorphic microsatellite loci (Tan9, Tan10, Tan11, Tan12, Tan14, Tan17, Tan24 and Tan26) following Walter *et al.* (2007) using a LiCOR 4300 DNA analyser and scoring allele size using GENE IMAGIR 4.05 (Scanalytics Inc., Rockville, MD, USA) imaging software. We used MSA 4.0 (Dieringer & Schlötterer, 2002) to look for genotyping errors, MICROCHECKER (van Oosterhout *et al.*, 2004) to check for the presence of null alleles and GENEPOP 4.0 (Raymond & Rousset, 1995) test for linkage disequilibrium among loci.

Data analyses

Nonequilibrium- and equilibrium-based tests for genetic differentiation were performed. Pairwise Fisher's exact tests (10 000 dememorizations and 20 000 permutations) were performed to test for significant differences in allele frequency distributions among sampled sites using Tools For Population Genetic Analysis (TFPGA, 1.3; Raymond & Rousset, 1995; Miller, 1997). Exact tests for Hardy–Weinberg equilibrium (HWE) were also performed (20 000 permutations) in TFPGA. Pairwise F_{ST} (Weir & Cockerham, 1984) was calculated using MSA4.0 (Dieringer & Schlötterer, 2002). Significance values for Fisher's exact tests, HWE and pairwise F_{ST} calculations were Bonferroni corrected.

Population genetic structure was tested for adherence to an isolation-by-distance model via two methods: (1) 'littoral restriction' (assumes fishes avoid open water, with geographic distances calculated along the shortest littoral route between sites; and (2) 'open-water dispersal' using the shortest water distances between sites. Significance for IBD relationships were determined using Mantel's (1967) tests in TFPGA.

Partial Bayesian genotype assignments (Rannala & Mountain, 1997) using population exclusion methods in GENECLASS 2.0 (Piry *et al.*, 2004) were performed to detect first-generation migrants among sites using the likelihood ratio L-home/L-max (Paetkau *et al.*, 2004). Monte Carlo resampling using Paetkau *et al.*'s (2004) simulation algorithm was performed with 10 000 simulated individuals at an assignment threshold *P*-value of 0.05. Individuals identified as migrants were assigned to a site of origin on the basis of their −ln likelihood. Mean dispersal distances were calculated from the shortest straight line distances between sampled sites where migrants originated and sites at which they were caught. Individual dispersal distances were compared with expected pairwise distances between all sites using a Kolmogorov–Smirnov test. We also used the migrant data to assess patterns in dispersal direction by assigning migrants back to their putative sources, then mapped the shortest route distance onto a quadrant map of Lake Matano to determine whether fish dispersed with respect to north, south, east or west direction (roughly along the long and short axes of the lake).

The Bayesian clustering programs STRUCTURE and BAPS work well for inferring the number of genetic clusters even at low levels of differentiation (F_{ST} = 0.02–0.03, Latch *et al.*, 2006). Lake-level population genetic structure was assessed independent of the spatial sampling pattern employed using *a posteriori* Bayesian genotype clustering in STRUCTURE 2.1 (Pritchard *et al.*, 2000). The analyses were performed in three independent runs using the admixture model with 500 000 burn-in followed by 1 000 000 MCMC repetitions, with three iterations (*K* ranging from 1 to 10). *K* was determined using Δ*K* following Evanno *et al.* (2005). We also used BAPS 4.14 (Corander & Marttinen, 2006) to independently estimate the number of genetic clusters. Initial clustering of individuals was performed in three replicates for each possible *K* (1–10). We then performed an admixture analysis where the minimum population size prior to estimating admixture was 10, in 50 iterations, with admixture simulated from 200 reference individuals in 50 iterations. Hierarchical AMOVA was performed in ARLEQUIN 3.0 (Excoffier *et al.*, 2005) to determine the proportion of genetic variance explained for both the STRUCTURE and BAPS groups, excluding individuals who failed assignment ($Q < 0.5$ to any cluster), vs. the spatially defined groups. To determine the potential effects of migrants on population structure, the STRUCTURE, BAPS and IBD analyses were re-run with migrant fish (identified as above) removed from data sets.

Results

A total of 486 *T. antoniae* were genotyped at eight microsatellite loci with three to 25 alleles per locus and observed (H_O) and expected heterozygosities (H_E) ranging from 0.30 to 1.0 (Table 1). Exact tests revealed

14209101, 2009, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/j.1420-9101.2008.01645.x by Univ Of California Santa Cruz - UCSC, Wiley Online Library on [17/10/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table 1 Sample size (N), number of alleles (A), observed (H_O) and expected heterozygosities (H_E) for eight microsatellite loci in *Telmatherina antoniae* collected from 10 sites in Lake Matano, Indonesia.

	Sampling site									
Locus	Lawa	P Rio	Taima	Indah	Kupu	L Ondau	Owesu	T Merah	W Lonto	Soluru
Tan 9										
N	46	44	43	56	58	43	54	47	58	7
A	24	20	22	21	25	24	22	19	24	8
H_O	0.89	0.82	0.86	0.95	**0.88**	0.91	0.93	0.83	0.84	0.71
H_E	0.93	0.94	0.93	0.94	**0.94**	0.95	0.92	0.93	0.92	0.89
Tan 10										
N	49	52	47	57	53	46	56	50	56	8
A	5	4	3	4	3	4	4	6	4	4
H_O	**0.53**	0.40	0.36	0.49	0.47	0.30	0.55	0.36	0.39	0.38
H_E	**0.58**	0.37	0.51	0.47	0.59	0.69	0.61	0.56	0.47	0.71
Tan 11										
N	46	53	45	56	58	42	55	47	55	8
A	10	9	10	11	16	10	10	10	14	6
H_O	**0.41**	0.43	0.44	0.57	0.55	0.57	0.62	0.40	0.53	0.75
H_E	**0.67**	0.63	0.70	0.67	0.77	0.70	0.71	0.69	0.81	0.73
Tan 12										
N	50	53	47	57	58	45	57	49	58	8
A	13	9	14	14	9	16	8	14	15	9
H_O	**0.72**	0.64	0.72	0.54	0.55	**0.51**	**0.51**	0.73	0.60	0.75
H_E	**0.83**	0.78	0.72	0.73	0.79	**0.82**	**0.80**	0.84	0.79	0.91
Tan 14										
N	50	52	45	57	58	45	56	49	58	8
A	8	7	7	7	7	5	7	5	7	4
H_O	0.48	0.54	0.60	0.63	0.59	0.51	0.54	0.51	0.52	0.63
H_E	0.74	0.68	0.76	0.70	0.73	0.69	0.73	0.63	0.72	0.69
Tan 17										
N	51	50	46	57	58	45	57	50	59	8
A	25	19	22	25	19	21	23	22	20	11
H_O	0.90	0.92	0.91	0.86	0.98	0.89	0.84	0.88	0.90	1
H_E	0.93	0.93	0.92	0.92	0.93	0.94	0.93	0.93	0.92	0.95
Tan 24										
N	51	52	46	56	57	46	57	47	57	8
A	8	8	8	9	7	9	7	7	8	7
H_O	0.49	0.52	0.46	0.57	**0.40**	0.41	0.53	0.49	0.44	0.50
H_E	0.57	0.53	0.67	0.60	**0.58**	0.61	0.49	0.55	0.56	0.79
Tan 26										
N	51	53	47	57	58	46	57	49	58	8
A	23	22	22	25	20	24	23	21	22	12
H_O	0.96	0.89	0.94	0.95	1	0.89	0.89	0.88	0.97	1
H_E	0.95	0.95	0.95	0.94	0.94	0.94	0.95	0.94	0.94	0.94

Departures from HWE (following Bonferroni correction) are given in bold.

significant differences in allele frequency distributions among all pairs of sample sites after Bonferroni correction ($P < 0.001$). Ten of 90 tests revealed significant departures from HWE following Bonferroni correction, but none were consistent across sites or loci (Table 1). Pairwise F_{ST} estimates ranged from 0.009 to 0.061 (Table 2) with an overall significant global F_{ST} estimate of 0.029 ($P = 0.0001$). Pairwise comparisons that included the Kupu site produced the highest F_{ST} estimates (Table 2, Fig. 1).

Mantel tests for IBD relationships using genetic distance [$F_{ST}/(1 - F_{ST})$] vs. geographic distances (or log geographic distances) did not yield significant results under scenarios of straight line dispersal ($P = 0.231$), nor littoral restriction dispersal ($P = 0.181$). There was no difference in the results of IBD analyses following removal of migrant fish from the data set.

Migrant analysis and dispersal estimates

Detection of first-generation migrants using GENECLASS resulted in 50 individuals being classified as migrants, with three to nine migrants per site (Table 3). Estimated dispersal distances ranged from 3.5 to 28.8 km with a

14209101, 2009, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/j.1420-9101.2008.01645.x by Univ Of California Santa Cruz - UCSC, Wiley Online Library on [17/10/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table 2 Pairwise F_{ST} estimates for 10 sampled sites of *Telmatherina antoniae*.

Site	Lawa	P Rio	Taima	Indah	Kupu	L Ondau	Owesu	T Merah	W Lonto	Soluru
Lawa										
P Rio	**0.017***									
Taima	**0.038***	**0.034***								
Indah	0.010	**0.021***	**0.037***							
Kupu	**0.043***	**0.051***	**0.050***	**0.050***						
L Ondau	0.009	**0.034***	**0.030***	**0.015**	**0.036***					
Owesu	**0.035***	**0.028***	**0.012**	**0.044***	**0.040***	**0.030***				
T Merah	**0.029***	**0.025***	**0.014**	**0.031***	**0.039***	**0.018***	0.010			
W Lonto	**0.021***	**0.024***	**0.026**	**0.019***	**0.051***	0.013	**0.025***	**0.013**		
Soluru	0.037	**0.070***	**0.058**	**0.044***	**0.044**	0.018	**0.069***	0.038	0.036	

F_{ST} estimates significance levels < 0.05 are in bold, < 0.01 indicated by an asterisk (*) following Bonferroni correction.

Table 3 Summary of migrant analysis (exclusion analysis) using GENECLASS 2.0, mean dispersal estimates taken from migrant individuals assigned back to their site of origin.

	Source site										
	Lawa	P Rio	Taima	Indah	Kupu	L Ondau	Owesu	T Merah	W Lonto	Soluru	Mean distance (km)
Capture site											
Lawa			1	1		1		1	1		18.4
P Rio			1	1			1		1		16.0
Taima	1			1	4				1		7.5
Indah			1		1	1	1				9.0
Kupu	1					1		2		1	10.5
L Ondau	3	2	1					2	1		17.9
Owesu	1		1							1	15.8
T Merah		1				1	1		1		10.6
W Lonto	1		2	1	1	1					15.1
Soluru				1				1	1		13.0

mean estimated dispersal distance of 13.6 (SD 0.94) km (Table 3). Dispersal distances for *T. antoniae* were not significantly different from random expected open-water distances using a Kolmogorov–Smirnov test ($d_{\max\ 0.05,7,45} = 6$, $P > 0.05$, Fig. 2). No significant directional bias relative to random expectations was observed.

Lake-level genetic structuring

Individual-based genotype clustering using STRUCTURE returned similar −ln likelihood values for $K = 4$, 5 and 6; however, those clusters did not correspond to any of the sample sites or geographic structuring. We selected $K = 4$ as the estimate of the correct number of genetic clusters using both the ΔK criterion of Evanno *et al.* (2005) and the recommendations of Pritchard *et al.* (2000). We then determined the genetic cluster composition at each sample site, following removal of 65 individuals (13% study wide) identified as unassigned (e.g. these include 'admixed' individuals) by a $Q < 0.5$ to any given cluster. We also forced assignment of all individuals on the basis

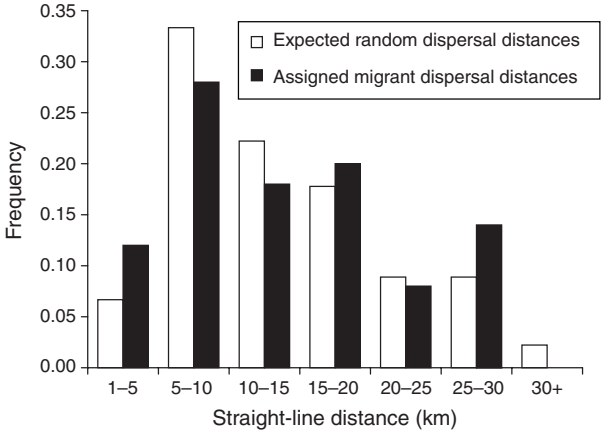

Fig. 2 Frequency histogram of open-water dispersal distances for individual migrant *Telmatherina antoniae* identified by genotype assignment with the comparable random expected distribution. The two distributions do not differ significantly (Kolmogorov–Smirnoff test ($P > 0.05$).

14209101, 2009, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/j.1420-9101.2008.01645.x by Univ Of California Santa Cruz - UCSC, Wiley Online Library on [17/10/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

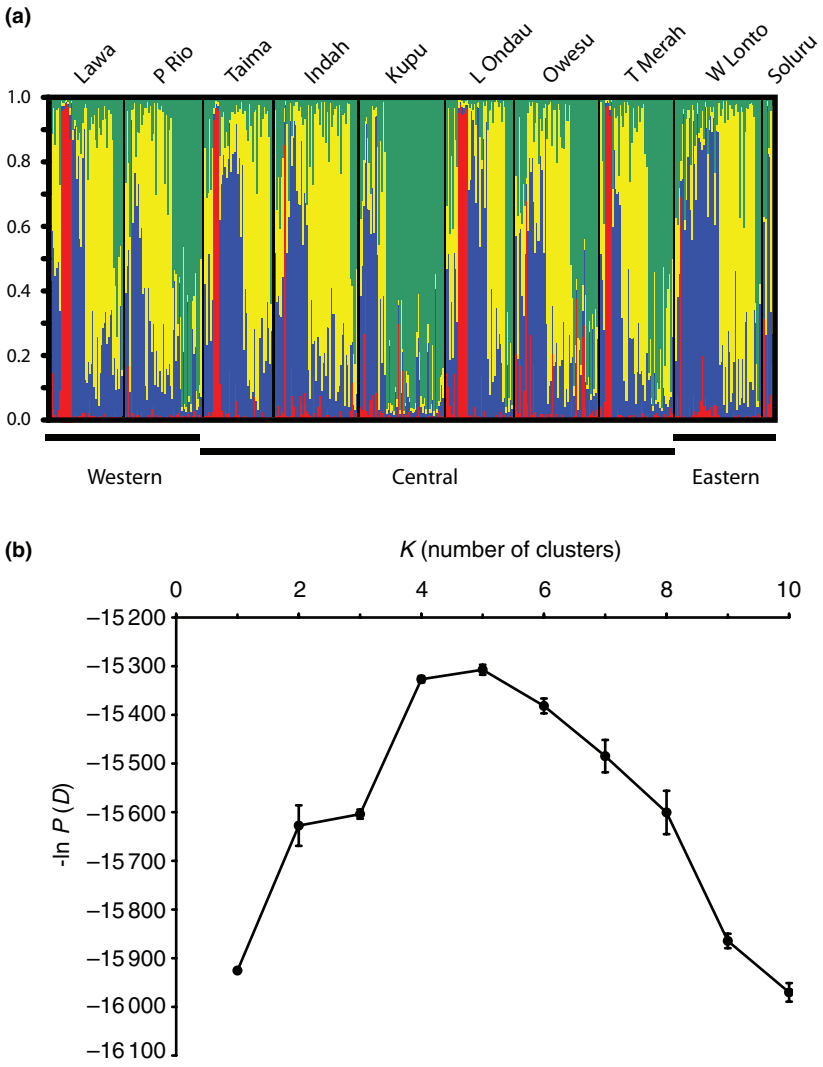

Fig. 3 (a) Summary of membership proportions for individuals from geographic sites in Bayesian clusters using STRUCTURE ($K = 4$). The column shading corresponds to each of the four clusters. (b) Negative log likelihoods [−ln P (D)] for microsatellite data fit to STRUCTURE models minimizing LD (k: 1–10), 500 000 burn-in, 500 000 iterations, with allele frequencies correlated and admixture allowed.

of their highest Q coefficient for a cluster. No gender- or size-based differences were found between clusters. Clustering using BAPS also failed to produce genetic structure corresponding to geographic structuring; with $K = 6$, of which the two additional groups primarily nested within STRUCTURE's groups 1 and 4. Both Bayesian programs grouped the same individuals into group 2, indicating high divergence of these individuals with respect to others in the data set. BAPS identified 11 individuals as significantly admixed (thus unassigned) using BAPS at $P < 0.05$; however, if the threshold is raised to $P < 0.10$, the number of significantly admixed individuals increases to 52, closer to the number identified by the STRUCTURE analysis. Significant differences in the proportion of each population cluster were apparent at each sampling site ($\chi^2_{27} = 134$, $P < 0.0001$) using both Bayesian clustering programs (Fig. 3, only STRUCTURE results shown).

From the hierarchical AMOVA, the genetic variance explained by the Bayesian clusters was two times greater using BAPS than geographic groupings (5.42%, $P < 0.001$; 2.32%, $P < 0.001$ respectively) and three times greater using STRUCTURE (7.30%, $P < 0.001$; 2.26%, $P < 0.001$ respectively). F_{ST} estimates between clusters were also larger (Table 4) in comparison with the geographic estimates (Table 2) with cluster 2 (both STRUCTURE and BAPS) as the most divergent. We reran the AMOVA using the data including forced assignment of individuals; the results were nearly identical with 7.30% of the variance explained by STRUCTURE's clusters, with only minor differences in spatial variance (2.92%).

Discussion

Among freshwater fish species, the degree of population substructure within a lake varies considerably. The

14209101, 2009, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/j.1420-9101.2008.01645.x by Univ Of California Santa Cruz - UCSC, Wiley Online Library on [17/10/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table 4 Mean and pairwise F_{ST} estimates for Bayesian clusters generated from STRUCTURE ($K = 4$) and BAPS ($K = 6$), all estimates are highly significant following Bonferroni correction ($P < 0.001$).

	1	2	3	4	5
STRUCTURE (F_{ST} = 0.102)					
2	0.120				
3	0.060	0.144			
4	0.031	0.181	0.079		
BAPS (F_{ST} = 0.087)					
2	0.157				
3	0.034	0.103			
4	0.111	0.158	0.083		
5	0.072	0.146	0.047	0.047	
6	0.011	0.152	0.019	0.107	0.070

majority of studies of within-lake genetic structure involving nontemperate species have focused on the cichlid species complexes of the African Rift lakes, with some species exhibiting little to no lake-wide structuring (Taylor & Verheyen, 2001; Koblmüller *et al.*, 2007) and others showing substantial genetic structure on small spatial scales (Taylor *et al.*, 2001; Rico & Turner, 2002; Duftner *et al.*, 2006; Sefc *et al.*, 2007). In the present study, we found significant spatial genetic structure in *T. antoniae* within a single lake, however; results of our migrant analysis and cluster-based assignments suggest that deep-water and/or habitat discontinuity does not represent barriers to dispersal. Adjacent sites lacking a continuous littoral habitat and separated by water exceeding 100 m in depth were among the least spatially differentiated (e.g. Tanah Merah vs. Owesu: $F_{ST} = 0.01$; Table 1). Mean estimated dispersal distances spanned approximately half the length of the lake, indicating a lack of strict allopatry and the opportunity for considerable gene flow. Furthermore, the Lake Matano *T. antoniae* do not conform to an isolation-by-distance model of population structure, making it unlikely that the observed genetic structure is due to distance-related reproductive isolation. We therefore conclude that, although constraints to gene flow exist for *T. antoniae,* on-going dispersal is occurring throughout Lake Matano.

The dispersal capacity of *T. antoniae* may be characteristic of Atheriniformes in general and the retention of such a trait would be particularly favoured in habitats where the cost of dispersal was low. Dispersal behaviour is risky given possibilities of failure to locate suitable habitat and mates, or *en route* mortality. The low productivity of Lake Matano (Haffner *et al.*, 2001) is characterized by an absence of pelagic predators, which would foster movement among sites thereby minimizing the predation risk of dispersal. However, it appears that some intrinsic factor(s) maintain significant genetic structuring despite such dispersal.

Local adaptation could account for the low but significant spatial structuring, as dispersing individuals (e.g. immigrants) may exhibit reduced fitness compared with residents (Hendry *et al.*, 2002). However, the low number of identified migrant individuals in this study is probably the result of reduced power in the assignment analyses because of low genetic differentiation among sites. Therefore, the true number of dispersing individuals is probably higher and cannot be definitively estimated due to spatial overlap of more than one population.

Multilocus genetic assignment methods such as those used in this study provide opportunities for defining populations with uncertain physical boundaries (Manel *et al.*, 2005; Rowe & Beebee, 2007). The genotype-based clustering explained a greater proportion of the genetic variance than the spatially based groups, indicating that the genetic clusters are present at all sites, spanning discontinuous habitat. Thus, the sampled fish from the 10 sites consist of differential proportions of individuals from the divergent and dispersive genetic clusters; however, the degree of isolation between the genetic clusters is only roughly estimated here. This type of population structure is analogous to a 'mixed stock fishery' where different mixtures of divergent populations (or 'stocks') result in genetic structure among samples. This type of genetic structure is common for a number of marine and aquatic organisms exhibiting philopatry, notably highly migratory sea turtles (Bowen & Karl, 2007), but has also been noted in landlocked salmonids (Potvin & Bernatchez, 2001). However, this observed among-sample site genetic divergence is problematic as it violates population genetic assumptions as specified by the Hardy–Weinberg principle.

First, the overlap of two or more distinct populations within a sample leads to a heterozygote deficit potentially creating a Wahlund effect, which may explain some departures from HWE among our data (Johnson & Black, 1984). As expected, assignment to the Bayesian clusters did slightly improve HWE, further supporting that these groupings may explain a higher proportion of genetic variance. Our estimation of *K* based on Pritchard *et al.* (2000) and Evanno *et al.* (2005), coupled with the assignment criteria of Latch *et al.* (2006) used here appears to be robust given concordance of STRUCTURE and BAPS results. Secondly, although the sympatry of these genetic clusters serves to explain the observed spatial genetic structure with high dispersal, it does not explain the persistence of these genetic clusters despite lake-wide dispersive ability. Thus, this apparent mixture of populations among sites may have resulted from two processes: (1) secondary contact following past allopatric divergence or (2) on-going reproductive isolation and genetic divergence.

The first possibility is less likely on two grounds. First, Lake Matano is considered an ancient lake (~4.5 Myr old) with no evidence to suggest that lake-level fluctuations have produced isolated refugia within Lake Matano at any time in its history. Furthermore, the estimated divergence times between *T. antoniae* and its

14209101, 2009, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/j.1420-9101.2008.01645.x by Univ Of California Santa Cruz - UCSC, Wiley Online Library on [17/10/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

sympatric sister mtDNA clade is 0.95–1.9 Ma (Roy *et al.*, 2007a). Thus, there is no evidence for past vicariance affecting Lake Matano *T. antoniae* populations, but it is a possibility. Our markers are based on rapidly evolving microsatellite DNA; therefore, gene flow facilitated by high dispersal would rapidly erode past allopatric divergence. Secondly, Roy *et al.* (2007a) also showed that the three mitochondrial telmatherinid clades probably diverged from a common ancestor within Lake Matano on the basis of ecological segregation, suggesting that any subsequent genetic structuring within *T. antoniae* may have occurred within Lake Matano as a result of similar processes.

The presence of sympatric but divergent genetic clusters may be the result of a number of intrinsic isolating mechanisms. For example, individuals may assortatively mate according to phenotype (e.g. colour and behavioural displays), as sexual selection has been implicated as an important factor driving sympatric genetic differentiation in other fishes (Seehausen, 2000). For *T. antoniae,* courtship and nuptial behaviours appear quite complex, with competition noted among males for access to females (Gray & McKinnon, 2006). Roy *et al.* (2007b) argued that colour was not an initiating factor in the earlier telmatherinid divergence, but did not assess colour as a potential structuring factor within specific telmatherinid clades. More recently, Herder *et al.* (2008) demonstrated a lack of colour-specific population differentiation among 'small' *T. antoniae* colour morphs using AFLP markers. However, for the closely related *Telmatherina sarasinorum* Gray *et al.* (2008) provide evidence for environment-contingent sexual selection favouring different colour morphs in various habitats, and it is possible that a similar pattern is operating for *T. antoniae*.

Herder *et al.* (2008) has further provided evidence for significant genetic differentiation among 'large' and 'small' morphtypes of *T. antoniae* and a lack of significant divergence between 'small' *T. antoniae* and the rare closely related but larger sized *Telmatherina prognatha*. It is not without possibility that immature 'large' *T. antoniae* and *T. prognatha* morphs may account for some of the genetic structure observed here. However, within our samples, spatially or temporally isolated reproductive aggregations could produce the observed genetic structuring. Genetic divergence at neutral microsatellite loci among populations of marine Atheriniformes with high dispersal capacity has classically been interpreted as a result of homing behaviour (Beheregaray & Sunnucks, 2001). However, little is known of the reproductive biology of *T. antoniae*, and the apparent lack of habitat specificity precludes the identification of natal sites for this species.

The pattern of 'divergence despite gene dispersal' observed in this study has also been identified in incipient species complexes where strong divergent selection drives isolation despite on-going gene flow (Rice & Hostert, 1993). Seemingly sympatric populations may in fact be isolated by extreme philopatry or micro-geographical structuring. This does not appear to be the case for *T. antoniae* where the lack of physical barriers to dispersal and the persistence of substantial genetic divergence lend support to the role of intrinsic mechanisms in the ongoing divergence among lineages in this ancient tropical lake.

Acknowledgments

We thank Sinyo Rio, Pak Dula, Peter Hehanussa and Andy Bramburger for assisting with field collections. We thank JL Shields, JD Hogan, CL Beneteau, CT Harvey and JR Ovenden for valuable comments on the manuscript. This research was supported by funding from the Natural Science and Engineering Council of Canada to DDH and GDH.

References

Abila, R., Barluenga, M., Engelken, J., Meyer, A. & Salzburger, W. 2004. Population-structure and genetic diversity in a haplochromine fish cichlid of a satellite lake of Lake Victoria. *Mol. Ecol.* **13**: 2589–2602.

Beheregaray, L.B. & Sunnucks, P. 2001. Fine-scale genetic structure, estuarine colonization and incipient speciation in the marine silverside fish *Odontesthes argentinensis*. *Mol. Ecol.* **10**: 2849–2866.

Bilton, D.T., Paula, J. & Bishop, J.D.D. 2002. Dispersal, genetic differentiation and speciation in estuarine organisms. *Estuar. Coast. Shelf Sci.* **55**: 937–952.

Bohonak, A.J. 1999. Dispersal, gene flow, and population structure. *Q. Rev. Biol.* **74**: 2145.

Bossart, J.L. & Prowell, D.P. 1998. Genetic estimates of population structure and gene flow: limitations, lessons and new directions. *Trends Ecol. Evol.* **13**: 202–206.

Bowen, B.W. & Karl, S.A. 2007. Population genetics and phylogeography of sea turtles. *Mol. Ecol.* **16**: 4886–4907.

Corander, J. & Marttinen, P. 2006. Bayesian identification of admixture events using multi-locus molecular markers. *Mol. Ecol.* **15**: 2833–2843.

Danley, P.D., Markert, J.A., Arnegard, M.E. & Kocher, T.D. 2000. Divergence with gene flow in rock-dwelling cichlids of Lake Malawi. *Evolution* **54**: 1725–1737.

Dieringer, D. & Schlötterer, C. 2002. Microsatellite Analyser (MSA): a platform independent analysis tool for large microsatellite data sets. *Mol. Ecol. Notes* **3**: 167–169.

Duftner, N., Sefc, K.M., Koblmüller, S., Nevado, B., Verheyen, E., Phiri, H. & Sturmbauer, C. 2006. Distinct population structure in a phenotypically homogeneous rock-dwelling cichlid fish from Lake Tanganyika. *Mol. Ecol.* **15**: 2381–2395.

DuPont, P.-P., Bourret, V. & Bernatchez, L. 2007. Interplay between ecological, behavioural and historical factors in shaping the genetic structure of sympatric walleye populations *Sander vitreus*. *Mol. Ecol.* **16**: 937–951.

Elphinstone, M.S., Hinten, G.N., Anderson, M.J. & Nock, C.J. 2003. An inexpensive and high-throughput procedure to extract and purify total genomic DNA for population studies. *Mol. Ecol. Notes* **3**: 317–320.

14209101, 2009, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/j.1420-9101.2008.01645.x by Univ Of California Santa Cruz - UCSC, Wiley Online Library on [17/10/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Evanno, G., Regnaut, S. & Goudet, J. 2005. Detecting the number of clusters of individuals using the software structure: a simulation study. *Mol. Ecol.* **14**: 2611–2620.

Excoffier, L., Laval, G. & Schneider, S. 2005. ARLEQUIN version 3.0: an integrated software package for population genetics data analysis. *Evol. Bioinform. Online* **1**: 47–50.

Gray, S.M. & McKinnon, J.S. 2006. A comparative description of mating behaviour in the endemic Telmatherinid fishes of Sulawesi's Malili Lakes. *Environ. Biol. Fishes* **75**: 471–482.

Gray, S.M., Dill, L.M., Tantu, F.Y., Loew, E.R., Herder, F. & McKinnon, J.S. 2008. Environment contingent sexual selection in a colour polymorphic fish. *Proc. R. Soc. B* **275**: 1785–1791.

Haffner, G.D., Hehanussa, P.E. & Hartoto, D. 2001. The biology and physical properties of large lakes of Indonesia. In: *The Great Lakes of the World: Food-Web Health and Integrity* (M. Manwar & R.E. Heck, eds), pp. 183–192. Backuys, Leiden.

Hauser, L., Carvalho, G.R. & Pitcher, T.J. 1998. Genetic population structure in the Lake Tanganyika sardine *Limnothrissa miodon*. *J. Fish Biol.* **53** (Suppl. A): 413–429.

Hendry, A.P., Hensleigh, J.E. & Reisenbichler, R.R. 1998. Incubation temperature, developmental biology and the divergence of sockeye salmon within Lake Washington. *Can. J. Fish. Aquat. Sci.* **55**: 1387–1394.

Hendry, A.P., Berg, O.K. & Quinn, T.P. 1999. Condition dependence and adaptation-by-time: breeding date, life history, and energy allocation in a population of salmon. *Oikos* **85**: 499–514.

Hendry, A.P., Taylor, E.B. & McPhail, J.D. 2002. Adaptive divergence and the balance between selection and gene flow: lake and stream stickleback in the Misty system. *Evolution* **56**: 119–1216.

Herder, F., Nolte, A.W., Pfaender, J., Schwarzer, J., Hadiaty, R.K. & Schliewen, U. 2006. Adaptive radiation and hybridization in Wallace's Dreamponds: evidence from sailfin silversides in the Malili Lakes of Sulawesi. *Proc. R. Soc. B* **273**: 2209–2217.

Herder, F., Pfaender, J. & Schliewen, U.K. 2008. Adaptive sympatric speciation of polychromatic ''roundfin'' sailfin silverside fish in Lake Matano. *Evolution* **62**: 2178–2195.

Johnson, M.S. & Black, R. 1984. The Wahlund effect and the geographical scale of variation in the intertidal limpet Siphonaria sp. *Mar. Biol.* **79**: 295–302.

Koblmüller, S., Sefc, K.M., Duftner, N., Warum, M. & Sturmbauer, C. 2007. Genetic population structure as an indirect measure of dispersal ability in a Lake Tanganyika cichlid. *Genetica* **130**: 121–131.

Kottelat, M. 1991. Sailfin silversides Pisces: Telmatherinidae of Lake Matano, Sulawesi, Indonesia, with descriptions of six new species. *Ichthyol. Explor. Freshw.* **1**: 321–344.

Kottelat, M., Whitten, A.J., Kartikasari, S.N. & Wirjoatmodjo, S. 1993. *Freshwater Fishes of Western Indonesia and Sulawesi*. Periplus, Jakarta.

Latch, E.K., Dharmarajan, G., Glaubitz, J.C. & Rhodes, O.E. Jr 2006. Relative performance of Bayesian clustering software for inferring population substructure and individual assignment at low levels of population differentiation. *Conserv. Genet.* **7**: 295–302.

Manel, S., Gaggiotti, O.E. & Waples, R.S. 2005. Assignment methods: matching biological questions with appropriate techniques. *Trends Ecol. Evol.* **20**: 136–142.

Mantel, N. 1967. The detection of disease clustering and a generalized regression approach. *Cancer Res.* **27**: 209–220.

Miller, M.. 1997. Tools for population genetic analysis (TFPGA) 1.3: a Windows program for analysis of allozyme and molecular population genetic data. Computer software distributed by the author: markperrymiller@gmail.com.

van Oosterhout, C., Hutchinson, W.F., Wills, D.P.M. & Shipley, P. 2004. MICROCHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Mol. Ecol. Notes* **4**: 535–538.

Paetkau, D., Slade, R., Burdens, M. & Estoup, A. 2004. Genetic assignment methods for the direct, real time estimation of migration rate: a simulation-based exploration of accuracy and power. *Mol. Ecol.* **13**: 55–65.

Palumbi, S.R. 1992. Marine speciation on a small planet. *Trends Ecol. Evol.* **7**: 114–118.

Piry, S., Alapetite, A., Cornuet, J.M., Paetkau, D., Baudouin, L. & Estoup, A. 2004. GENECLASS2: a software for genetic assignment and first-generation migrant detection. *J. Heredity* **95**: 536–539.

Potvin, C. & Bernatchez, L. 2001. Lacustrine spatial distribution of landlocked Atlantic salmon populations assessed across generations by multi-locus individual assignment and mixed-stock analyses. *Mol. Ecol.* **10**: 2375–2388.

Power, M., O'Connell, M.F. & Dempson, J.B. 2005. Ecological segregation within and among Arctic char morphotypes in Gander Lake, Newfoundland. *Environ. Biol. Fishes* **73**: 263–274.

Pritchard, J.K., Stephens, M. & Donnelly, P. 2000. Inference of population structure using multilocus genotype data. *Genetics* **155**: 945–959.

Rannala, B. & Mountain, J.L. 1997. Detecting immigration by using multilocus genotypes. *Proc. Natl Acad. Sci. Biol.* **94**: 9197–9201.

Raymond, M. & Rousset, F. 1995. GENEPOP version 1.2: population genetics software for exact tests and ecumenicism. *J. Heredity* **86**: 248–249.

Rice, W.R. & Hostert, E.E. 1993. Laboratory experiments on speciation: what have we learned in 40 years? *Evolution* **47**: 1637–1653.

Rico, C. & Turner, G.F. 2002. Extreme microallopatric divergence in a cichlid species from Lake Malawi. *Mol. Ecol.* **11**: 243–251.

Rowe, G. & Beebee, T.J.C. 2007. Defining population boundaries: use of three Bayesian approaches with microsatellite data from British natterjack toads *Bufo calamita*. *Mol. Ecol.* **16**: 785–796.

Roy, D. 2006. The evolutionary history and ecology of *Telmatherina* in Lake Matano: an example of adaptive radiation in an ancient lake. PhD Thesis, University of Windsor, Windsor, ON.

Roy, D., Docker, M.F., Hehanussa, P., Heath, D.D. & Haffner, G.D. 2004. Genetic and morphological data supporting the hypothesis of adaptive radiation in the endemic fish of Lake Matano. *J. Evol. Biol.* **17**: 1268–1276.

Roy, D., Paterson, G., Hamilton, P.B., Heath, D.D. & Haffner, G.D. 2007a. Resource-based adaptive divergence in the freshwater fish *Telmatherina* from Lake Matano, Indonesia. *Mol. Ecol.* **16**: 35–48.

Roy, D., Docker, M.F., Haffner, G.D. & Heath, D.D. 2007b. Body shape vs. colour associated initial divergence in the *Telmatherina* radiation in Lake Matano, Sulawesi, Indonesia. *J. Evol. Biol.* **20**: 1126–1137.

Rüber, L., Meyer, A., Sturmbauer, C. & Verheyen, E. 2001. Population structure in two sympatric species of the Lake

14209101, 2009, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/j.1420-9101.2008.01645.x by Univ Of California Santa Cruz - UCSC, Wiley Online Library on [17/10/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Tanganyika cichlid tribe Eretmodini: evidence for introgression. *Mol. Ecol.* **10**: 1207–1225.

Seehausen, O. 2000. Explosive speciation rates and unusual species richness in haplochromine cichlid fishes: effects of sexual selection. *Adv. Ecol. Res.* **31**: 237–274.

Sefc, K.M., Baric, S., Salzburger, W. & Sturmbauer, C. 2007. Species-specific population structure in rock-specialized cichlid species in Lake Tanganika, East Africa. *J. Mol. Evol.* **64**: 33–49.

Shaw, P.W., Turner, G.F., Idid, M.R., Robinson, R.L. & Carvalho, G.R. 2000. Genetic population structure indicates sympatric speciation of Lake Malawi pelagic cichlids. *Proc. R. Soc. B* **267**: 2273–2280.

Stepian, C.A. & Faber, J.E. 1998. Population genetic structure, phylogeography and spawning philopatry in walleye *Stizostedion vitreum* from mitochondrial DNA control region sequences. *Mol. Ecol.* **7**: 1757–1769.

Taylor, M.I. & Verheyen, E. 2001. Microsatellite data reveals weak population substructuring in *Copadichromis* sp. 'virginalis kajose', a demersal cichlid from Lake Malawi, Africa. *J. Fish Biol.* **59**: 593–604.

Taylor, M.I., Rüber, L. & Verheyen, R. 2001. Microsatellites reveal high levels of population substructuring in the species-poor Eretmodine cichlid lineage from Lake Tanganyika. *Proc. R. Soc. B* **268**: 803–808.

Walter, R.P., Ovenden, J.R. & Heath, D.D. 2007. Characterization of four tetranucleotide and six dinucleotide microsatellite markers for use in *Telmatherina antoniae* and related species. *Mol. Ecol. Notes* **7**: 651–653.

Ward, R.D. & Elliot, N.G. 2001. Genetic population structure of species in the South East Fishery of Australia. *Mar. Freshw. Res.* **52**: 563–573.

Weir, B.S. & Cockerham, C.C. 1984. Estimating *F*-statistics for the analysis of population structure. *Evolution* **38**: 1358–1370.

Wong, B.B.M., Keogh, J.S. & McGlashan, D.J. 2004. Current and historical patterns of drainage connectivity in eastern Australia inferred from population genetic structuring in a widespread freshwater fish *Pseudomugil signifer* (Pseudomugilidae). *Mol. Ecol.* **13**: 391–401.

Received 14 August 2008; revised 24 September 2008; accepted 11 October 2008

14209101, 2009, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/j.1420-9101.2008.01645.x by Univ Of California Santa Cruz - UCSC, Wiley Online Library on [17/10/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License