

# Cryptic species and parallel genetic structuring in Lethrinid fish: Implications for conservation and management in the southwest Indian Ocean

Amy J. E. Healey<sup>1</sup>  | Niall J. McKeown<sup>1</sup> | Amy L. Taylor<sup>2</sup> | Jim Provan<sup>1</sup> |  
Warwick Sauer<sup>3</sup> | Gavin Gouws<sup>4</sup> | Paul W. Shaw<sup>1,3</sup>

<sup>1</sup>Institute of Biological Environmental and Rural Sciences, Aberystwyth University, Aberystwyth, Dyfed, UK

<sup>2</sup>School of Biological Sciences, Royal Holloway University, Egham Hill, Egham, UK

<sup>3</sup>Department of Ichthyology and Fisheries Science, Rhodes University, Grahamstown, South Africa

<sup>4</sup>South African Institute for Aquatic Biodiversity, Grahamstown, South Africa

## Correspondence

Paul W. Shaw, Aberystwyth University, Aberystwyth, Dyfed, UK.  
Email: pws3@aber.ac.uk

## Funding information

Natural Environment Research Council

## Abstract

Analysis of genetic variation can provide insights into ecological and evolutionary diversification which, for commercially harvested species, can also be relevant to the implementation of spatial management strategies and sustainability. In comparison with other marine biodiversity hot spots, there has been less genetic research on the fauna of the southwest Indian Ocean (SWIO). This is epitomized by the lack of information for lethrinid fish, which support socioeconomically important fisheries in the region. This study combines comparative phylogeographic and population genetic analyses with ecological niche modeling to investigate historical and contemporary population dynamics of two species of emperor fish (*Lethrinus mahsena* and *Lethrinus harak*) across the SWIO. Both species shared similarly shallow phylogeographic patterns and modeled historical (LGM) habitat occupancies. For both species, allele frequency and kinship analyses of microsatellite variation revealed highly significant structure with no clear geographical pattern and nonrandom genetic relatedness among individuals within samples. The genetic patterns for both species indicate recurrent processes within the region that prevent genetic mixing, at least on timescales of interest to fishery managers, and the potential roles of recruitment variability and population isolation are discussed in light of biological and environmental information. This consistency in both historical and recurrent population processes indicates that the use of model species may be valuable in management initiatives with finite resources to predict population structure, at least in cases wherein biogeographic and ecological differences between taxa are minimized. Paradoxically, mtDNA sequencing and microsatellite analysis of samples from the Seychelles revealed a potential cryptic species occurring in sympatry with, and seemingly morphologically identical to, *L. mahsena*. BLAST results point to the likely misidentification of species and incongruence between voucher specimens, DNA barcodes, and taxonomy within the group, which highlights the utility and necessity of genetic approaches to characterize baseline biodiversity in the region before such model-based methods are employed.

## KEYWORDS

biodiversity hot spot, chaotic genetic patchiness, conservation, cryptic species, kinship, larval cohesion

## 1 | INTRODUCTION

Understanding the evolution of marine biodiversity requires knowledge of the patterns and processes underpinning genetic structuring of species. Combined demographic–genetic studies can reveal the spatial and temporal scale at which evolutionary forces are occurring (Waples, 1998) and, by extension, identify cryptic components of biodiversity and so allow optimization of spatial conservation strategies. Identification of spatial and temporal processes is particularly important for harvested taxa, as genetic diversity and adaptation are key factors underpinning the resilience/sustainability, and ultimately the evolutionary potential, of populations and species (Illes & Sinclair, 1982; Ruzzante et al., 2006; Ryman, Utter, & Laikre, 1995; Therkildsen et al., 2013).

The southwest Indian Ocean (SWIO), which forms a subdivision of the tropical Indo-Pacific, is renowned for its diverse marine habitats and resources, including an ichthyofaunal richness (Smith, Smith, & Heemstra, 2003) that supports significant artisanal, subsistence, and commercial fisheries across the region (Berg, Francis, & Souter, 2002). The complex oceanographic features of this region, characterized by the bifurcation of the South Equatorial Current into two continental currents and the series of eddies in the Mozambique Channel (Benny, 2002), are hypothesized to have contributed to this high biodiversity (Ridgway & Sampayo, 2005). Diverse patterns of genetic differentiation have been observed within species distributed across the SWIO, from high levels of gene flow (Chiang, Hsu, Wu, Chang, & Yang, 2008; Duda & Palumbi, 1999; Fratini, Ragionieri, & Cannicci, 2010; Muths, Grewe, Jean, & Bourjea, 2009; Muths, Le Couls, Evano, Grewe, & Bourjea, 2013; Ragionieri, Cannicci, Schubart, & Fratini, 2010; Silva, Mesquita, & Paula, 2010b) to more complex species-specific patterns of isolation, including disruption of gene flow across the Mozambique Channel (Gopal, Tolley, Groeneveld, & Matthee, 2006; Muths et al., 2011; Silva, Mesquita, & Paula, 2010a; Visram et al., 2010) and isolation of the northern (Ragioneiri et al. 2010; Visram et al., 2010) and/or southern (Fratini & Vannini, 2002; Lessios, Kessing, & Pearse, 2001; Muths, Tessier, & Bourjea, 2015; Muths et al., 2011) Mascarene Islands from continental populations. Some investigations have found signatures of historical vicariance associated with sea-level fluctuations during the Pleistocene (Ragionieri, Fratini, Vannini, & Schubart, 2009; Silva et al., 2010a). However, despite a growing body of ecological and evolutionary research the SWIO is still regarded as one of the least-studied tropical ecosystems (Beheregaray, 2008; Ridgway & Sampayo, 2005).

The complex array of genetic population structures resolved to date within the SWIO highlight the potential hazards of making region-wide predictions of population structure from single species. However, studies of ecologically similar and contemporaneously

co-distributed species represent analytical frameworks wherein variance between lineages, place, and time are minimized, and thus “general rules” informative to subsets of taxa may be resolved (Bird, Holland, Bowen, & Toonen, 2007, 2011; Bohonak, 1999; Dawson, 2012; Lester, Ruttenberg, Gaines, & Kinlan, 2007). In this study, we develop such an analytical framework by performing replicated genetic analysis of two lethrinid fish species: *Lethrinus harak*, the thumbprint emperor, and *Lethrinus mahsena*, the sky emperor. Both species are ecologically and biogeographically similar. The habitat of *L. harak*, the most abundant and commercially important lethrinid in the SWIO (Kulmiye, Ntiba, & Kisia, 2002), consists of shallow and protected coastal areas including reefs, mangroves, shallow lagoons, and sea-grass beds (Ebisawa & Ozawa, 2009; Kulmiye et al., 2002). *Lethrinus mahsena* occupies similar shallow habitats but has been observed predominantly in reefs or adjacent areas (Carpenter & Allen, 1989). Like the majority of lethrinids, both species are relatively long-lived (up to 15 years; Ebisawa & Ozawa, 2009) protogynous hermaphrodites (Ebisawa, 2006; Grandcourt, 2002), but many aspects of their reproductive biology are not known (Ebisawa, 2006; Kulmiye et al., 2002). Both species have pelagic larval stages with Nakamura et al. (2010) reporting a larval duration of around 29 days for *L. harak*.

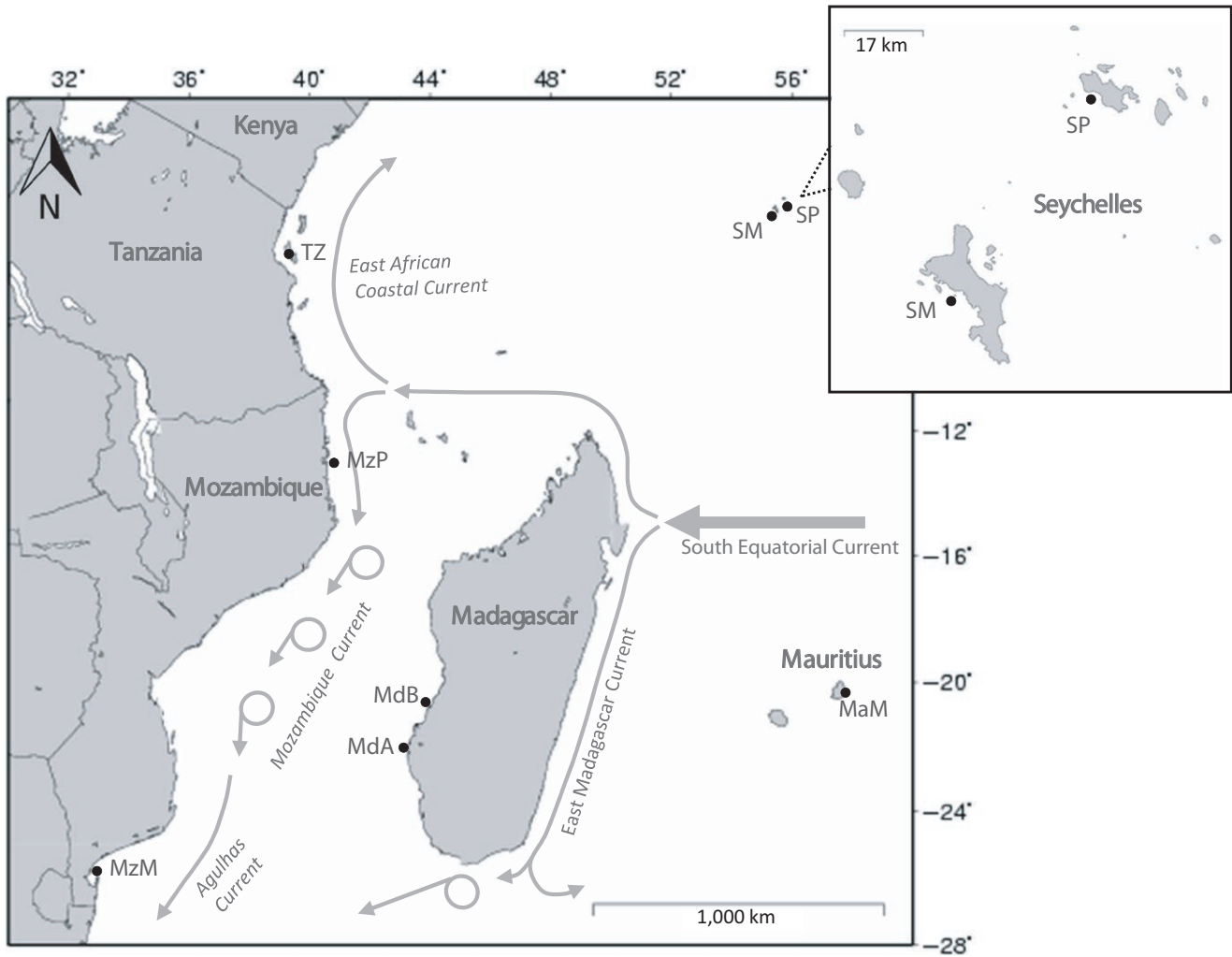
The aim was to test the prediction that both species would display similar patterns of historical (phylogeographic) and contemporary population genetic structure across the SWIO. As the Lethrinidae is one of the most important families of fish for artisanal, recreational, and subsistence fisheries (Carpenter & Allen, 1989; Gouws, 2012), the genetic patterns are interpreted in the context of recurrent recruitment processes relevant to stock sustainability. To this end, kinship analyses were employed to complement estimates of connectivity/isolation based on traditional population genetic (e.g.,  $F_{ST}$ ) methods. To provide an additional historical context, we combined genetic studies with ecological niche modeling (ENM), which has been shown to provide novel insights into the influence of historical biogeography on genetic variation in marine species (Glynn, Houghton, & Provan, 2015; Glynn et al., 2016).

## 2 | MATERIAL AND METHODS

### 2.1 | Sample collection and molecular analyses

Samples (fin clips fixed in 95% ethanol) were collected from commercial or subsistence catches landed in seven locations across the SWIO between 2009 and 2012 (Figure 1). Total DNA was extracted from all samples using a standard CTAB-chloroform/isoamyl alcohol method (Winnepenninckx, Backeljau, & Dewachter, 1993).

For both species, a fragment of the mtDNA cytochrome c oxidase subunit I (COI) gene was amplified by polymerase



**FIGURE 1** Sampling locations for *Lethrinus mahsena* and *Lethrinus harak* within the southwest Indian Ocean: Mauritius, Mahebourg (MaM); Seychelles, Praslin (SP); Seychelles, Mahe (SM); Tanzania, Zanzibar (TZ); Mozambique, Pemba (MzP); Madagascar, Belo sur Mer (MaB); Mozambique, Maputo (MzM)

chain reaction (PCR) with species-specific primers (*L. mahsena*: LmCOIF 5'-TGGTAGGAACAGCCCTAAGC-3' & LmCOIR 5'-AGAATTGGGTCCCCTCCTC-3'; *L. harak*: LhCOIF 5'-CGAACTTAGTCAGCCCGGA-3' & LhCOIR 5'-TGCTGATAGAGGATTGGTC-3') for a subset of individuals. PCRs were performed in a total volume of 20  $\mu$ l, containing 4  $\mu$ l template DNA, 2 mmol/L MgCl<sub>2</sub>, 0.5  $\mu$ mol/L forward primer and 0.5  $\mu$ mol/L of reverse primer, 0.2 mmol/L dNTP mix (20  $\mu$ mol/L each dATP, dCTP, dGTP, dTTP), 1 $\times$  reaction buffer [75 mmol/L Tris-HCl, 20 mmol/L (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>], and Taq polymerase (BioTaq, 5 U/ $\mu$ l). The PCR thermoprofile was: 180 s at 95°C, followed by 45 cycles of 30 s at 95°C, 45-s annealing at 55°C for both primer sets, and 60 s at 72°C, followed by a final 5-min extension at 72°C. PCR products were then purified using EXOSAPIT and sequenced from both directions on an Applied Biosystems 3500 platform using the respective PCR primers. Sequences were aligned using the Clustal W (Thompson, Higgins, & Gibson, 1994) program, available in BioEdit (Hall, 1999), and analyzed using BLAST. Sequences revealed the presence of a highly divergent clade among the Seychelles *L. mahsena* samples (described in Section 3). To investigate the

frequency of occurrence of this clade elsewhere, a diagnostic PCR-RFLP method was developed and used to genotype all *L. mahsena* samples (Appendix S1: Figure S1).

Nuclear genetic variation was assessed at nine dinucleotide (100RTE, 95ACRTE, 95TGRTE, 90RTE, 80RTE, 75RTE, 68RTE, 58RTE, BST2.33) and one trinucleotide (96RTE) microsatellite loci described by Van Herwerden, Benzie, Peplow, and Davies (2000); Van Herwerden, Benzie, and Davies (2003). PCRs were performed under the following conditions: 180 s at 95°C, followed by 35 cycles of 30-second denaturing at 95°C, 30-s annealing at 50°C, and 30 s at 72°C. In order to increase product strength and reduce nonspecific products, PCR protocols were altered for a subset of loci. Annealing temperature was increased to 60°C for 96RTE and 95ACRTE. For 75RTE, 68RTE, and 80RTE, the number of cycles was increased to 55 and annealing temperatures increased to 60°C, 55°C, and 54°C, respectively. Each 10  $\mu$ l reaction contained 3  $\mu$ l template DNA, 1 pmol of each primer, 5  $\mu$ l of 2xBioMix [1.5 mmol/L MgCl<sub>2</sub>] (Bioline, UK), and 1  $\mu$ l of ddH<sub>2</sub>O. Amplicons were separated on an Applied Biosystems 3500 with alleles inferred using the Peak Scanner software (Applied Biosystems).

To ensure robustness of genotypes, we employed a double genotyping as described by McKeown, Arkhipkin, and Shaw (2017).

## 2.2 | Statistical analysis of mtDNA sequences

After pruning sequences, a total of 486 bp of the mtDNA COI gene was aligned across 72 *L. mahsena* and 104 *L. harak* samples. All analyses of mtDNA sequences were conducted in Arlequin 3.5.1.2 (Excoffier & Lischer, 2010) unless otherwise stated. Phylogenetic relationships among sequences were inferred using maximum likelihood (ML) trees in MEGA v6.06 (Tamura, Stecher, Peterson, Filipiński, & Kumar, 2013) and Bayesian inference performed using MrBayes v3.2 (Ronquist & Huelsenbeck, 2003). For both species, the K2P + G + I substitution model, identified as optimal using Modeltest 3.7 (Posada & Crandall, 1998), was used. Maximum likelihood bootstrap values were calculated using 1000 bootstrap replicates, and Bayesian inference (BI) was calculated assuming unknown model parameters, and run over 5,000,000 generations, sampling the Markov chain every 1000 generations, with the first 15% of trees discarded as burn-in. Median-joining haplotype networks were constructed in Network 5.0 (Fluxus Technology). Percentage sequence divergence within and between species/clades and K2P distances were calculated using MEGA v6.06.

Genetic variation was assessed using calculations of the number of haplotypes ( $H$ ), in addition to indices of haplotype ( $h$ ) and nucleotide ( $\pi$ ) diversity (Nei, 1987) alongside their variances. Genetic differentiation among samples was further tested by global and pairwise  $\Phi_{ST}$  using pairwise haplotype distances (Weir & Cockerham, 1984), with associated  $p$  values estimated after 10,000 permutations.

## 2.3 | Statistical analysis of microsatellite data

The number of alleles (NA), allelic richness (AR), and observed ( $H_O$ ) and expected ( $H_E$ ) heterozygosities were calculated using GENALEX 6 (Peakall & Smouse, 2006). Deviations from Hardy–Weinberg (HW) expectations and linkage disequilibrium between pairs of loci were assessed using exact tests in Genepop v4.2 (Raymond & Rousset, 1995).

Genetic differentiation was quantified using global and pairwise  $F$ -statistics (Wright, 1978). To account for possible null allele effects, such indices were also calculated using the null alleles adjustment in FreeNA (Chapuis & Estoup, 2007). To test for signatures of isolation by distance (IBD), correlations between geographical distances (minimum sea distance in nautical miles) and linearly transformed genetic differences ( $F_{ST}/(1 - F_{ST})$ ) Slatkin, 1993) were tested using a MANTEL matrix correlation test in GENALEX. Genetic structure was also investigated using Bayesian “group assignment” “without admixture” methods implemented in BAPS 6.0 (Corander, Waldmann, Marttinen, & Sillanpää, 2004), for models of  $K = 1-6$  (10 independent runs per  $K$ ). Genetic relationships among samples were visualized using factorial correspondence analysis (FCA) in GENETIX (Belkhir, Borsa, Chikhi, Raufaste, & Bonhomme, 2004). Additionally, to assess nuclear differentiation between descendants of the two highly divergent mtDNA clades observed in *L. mahsena* (Seychelles), self-classification tests (i.e., assignment of nuclear genotypes to clade defined baseline groups)

were employed in GENECLASS 2 (Piry et al., 2004). The chord genetic distance (Cavalli-Sforza & Edwards, 1967) was used due to the deviations from Hardy–Weinberg equilibrium.

Mean pairwise relatedness within samples was calculated using the relatedness estimator  $r_{qg}$  of Queller and Goodnight (1989) in GENALEX (Peakall & Smouse, 2006) with associated 95% confidence intervals determined by 1,000 bootstraps. Permutation of genotypes among all samples (999 times) was used to calculate the upper and lower 95% confidence intervals for the expected range of  $r_{qg}$  under a panmictic model. The maximum likelihood method implemented in ML-RELATE (Kalinowski, Wagner, & Taper, 2006) was used to infer the relationships among pairs of individuals, specifically to categorize them as unrelated (U), half-sib (HS), full-sib (FS), or parent offspring (PO). Pairs of individuals were classified into three categories: (a) unrelated (classification as U only), (b) related (classification as any combination of HS, FS, PO but not U), or (c) ambiguous (classification as U as well as some related state).

## 2.4 | Estimation of Type I and Type II error rates

To assess differences in statistical power among the data sets for both species and marker types, Type I and Type II error rates were estimated using POWSIM (Ryman & Palm, 2006). Analyses were run for a combination of the average and smallest sample sizes for each species (adjusted for mtDNA following Larsson, Charlier, Laikre, & Ryman, 2009).

## 2.5 | Paleodistribution modeling

To evaluate suitable habitat ranges (and any geographical range changes) of *L. harak* and *L. mahsena* during the Last Glacial Maximum (LGM; ~21 KYA), paleodistribution modeling was conducted. Post-1950s species occurrence data were obtained from the Global Biodiversity Information Facility (GBIF; <http://www.gbif.org>), along with the Ocean Biogeographic Information System (<http://www.iobis.org>), resulting in a total of 335 occurrences of *L. harak* and 152 of *L. mahsena*. Contemporary bioclimatic data, specifically generated for marine environments, were collected at 5-min resolution from MARSPEC (Sbrocco & Barber, 2013; Table S1). Models were generated in MAXENT v3.3.3 (Phillips, Anderson, & Schapire, 2006) and cross-validation conducted using 10 replicate runs, under the default parameters. Model performance was assessed based upon the area under the receiver operating characteristic curve (AUC). Finally, the models were projected onto bioclimatic data reconstructed to represent the LGM (Sbrocco, 2014), using an ensemble of five different models (CNRN, ECBILTCLIO, FGOALS, Had-CM, and MIROC-322).

# 3 | RESULTS

## 3.1 | Cryptic genetic divergence in sympatry

Phylogenetic reconstruction of COI haplotypes among putative *L. mahsena* revealed the presence of two highly divergent clades within the SWIO, which were separated by 20 mutational steps

(Figure 2). Diagnostic PCR-RFLP analysis of all samples (Appendix S1) confirmed that clade A individuals occurred across all samples (Mauritius, Mahe Island Seychelles, Tanzania, Mozambique Pemba, Mozambique Maputo, Madagascar Belo sur Mer) excluding Praslin Island Seychelles, whereas clade B individuals were found in the Seychelles samples, where they co-occurred with, and were more frequent than, clade A fish.

BLAST searches of clade A individuals revealed a 99% sequence similarity to *L. mahsena* voucher specimens from Mozambique (JF493750.1, JF493751.1, JF493752.1); however, BLAST searches of clade B individuals also identified 99% sequence similarity to *L. mahsena* voucher specimens from India (EF609387.1, KM079305.1, KM079304.1, KJ920117.1); this was followed by 97% sequence similarity to all available *L. atkinsoni* voucher specimens from Western Australia, the Philippines, and Taiwan (KV944053.1, KP194639.1, KP194152.1, KP194307.1, KF009603.1, KC970391.1, EF609384.1). The average sequence divergence between clade A and clade B was 4.6% (k2p distance = 0.048, SE = 0.010; Table S4) with clade B less divergent from *L. atkinsoni* (2.2%; k2p distance = 0.021, SE = 0.006; Figure 3). Microsatellite genotypes revealed a clear differentiation between individuals partitioned according to their mtDNA clade in the FCA (Figure 4), with a corresponding high  $F_{ST}$  (0.01,  $p < .001$ ). This pattern was also supported by assignment tests which found complete self-classification to clade with the exception of a single individual from clade A, which had a similar probability of assignment to both clades (0.55 and 0.45 assignment probability to clade A and B groups, respectively).

### 3.2 | mtDNA COI sequence data

Overall haplotype and nucleotide diversities (Table 1) were similar for both *L. harak* ( $h = 0.52$ ,  $\pi = 0.0011$ ) and *L. mahsena* ( $h = 0.59$ ,

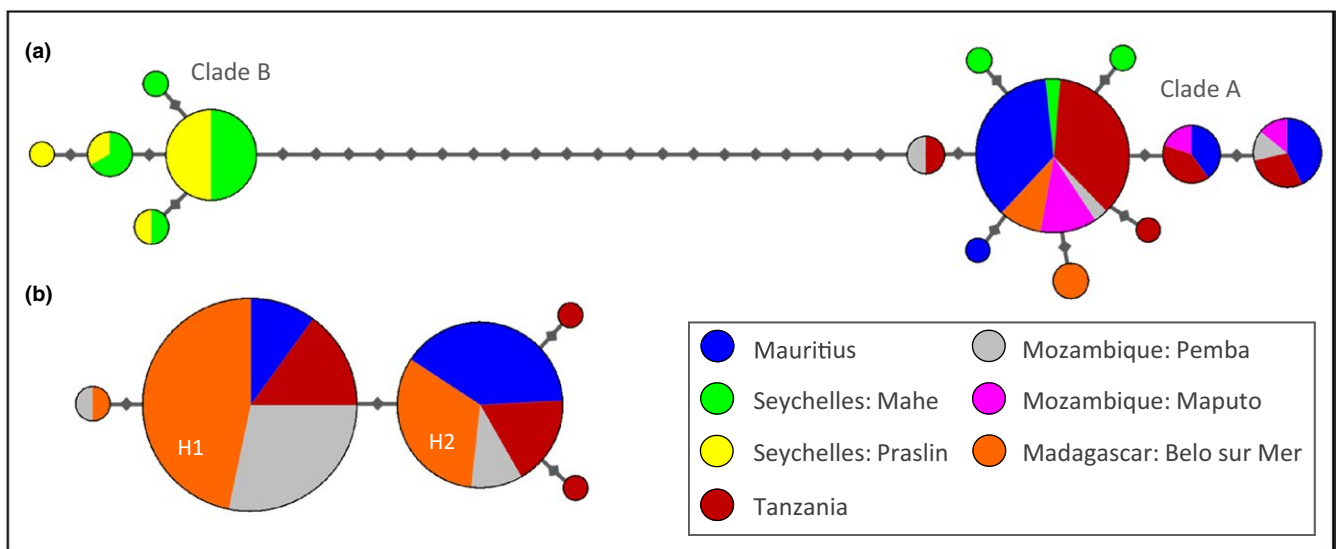
$\pi = 0.0018$ ). Network representation of *L. mahsena* (excluding clade B) revealed a single haplogroup dominated by a single common haplotype present across all SWIO localities. Similarly for *L. harak*, a single haplogroup was dominated by two common haplotypes (Figure 2): haplotype 2 which was dominant in Mauritius and haplotype 1 which was dominant across all other SWIO localities.

Although global values of genetic differentiation were high and significant for *L. harak* ( $\Phi_{ST} = 0.151$ ,  $p = .002$ ), pairwise tests revealed all significant values to be associated with two samples (Mauritius and Mozambique Pemba), with only the comparison between Mauritius and Mozambique Pemba remaining significant after Bonferroni correction (Table 2). Global differentiation was lower and not significant for *L. mahsena* ( $\Phi_{ST} = 0.024$ ,  $p = .277$ ), and although pairwise  $\Phi_{ST}$  was high in many cases, they were all nonsignificant (Table 2).

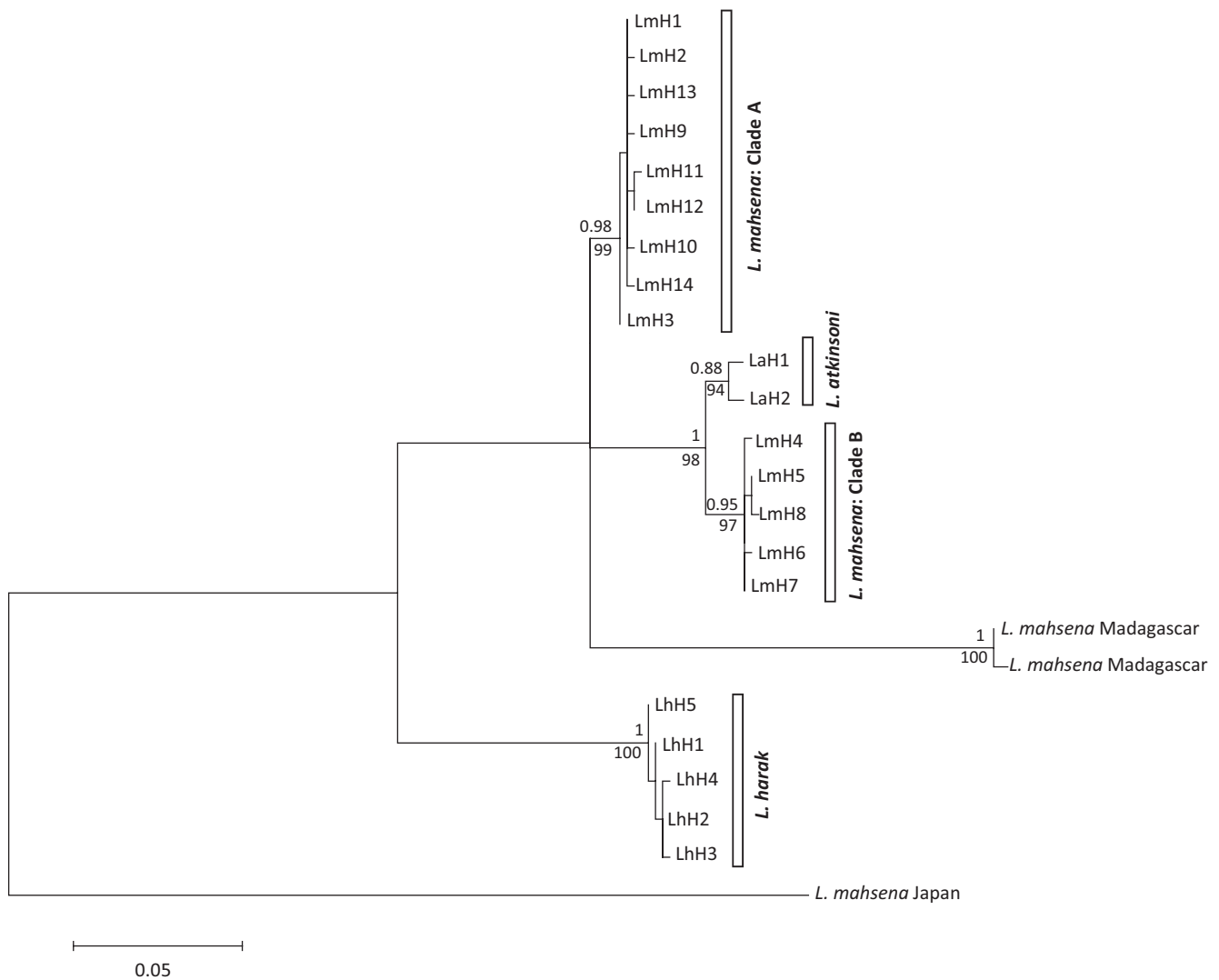
Owing to small sample sizes, demographic tests were conducted globally. Neither Tajima's  $D$  ( $D = -0.524$ ,  $p = .335$ ) nor Fu's  $F_s$  ( $F_s = -1.100$ ,  $p = .280$ ) tests detected global deviations from neutrality for *L. harak*. For *L. mahsena* Fu's  $F_s$  was significant ( $F_s = -4.617$ ,  $p = .002$ ), but Tajima's  $D$  was not ( $D = -1.330$ ,  $p = .070$ ). Mismatch analyses were compatible with expansions for *L. mahsena* ( $\tau = 1.078$ , 95% CI: 0–3.43,  $p = .690$ ) but not in *L. harak* ( $\tau = 0.723$ , 95% CI: 0.473–1.076,  $p = .010$ ).

### 3.3 | Microsatellite allele frequency data

A total of 247 individuals of *L. harak* from four sample locations and 72 *L. mahsena* individuals (excluding those belonging to clade B) from six sample locations were screened across nine and 10 microsatellite loci, respectively. Levels of intrasample genetic variability were similar in both species (Table 1). No significant linkage disequilibrium was indicated between any pair of loci in either species. POWSIM analysis indicated that microsatellite data conferred more power to detect



**FIGURE 2** Haplotype network of (a) *Lethrinus mahsena* and (b) *Lethrinus harak* based on 486 bp of COI mtDNA. Each small gray diamond represents a single mutation; branch lengths are proportional to the number of differences, and node size is proportional to the haplotype frequency



**FIGURE 3** Phylogenetic relationships within *Lethrinus* species using 486 bp COI mtDNA. Statistical support for nodes is given for both Bayesian analyses (posterior probabilities) above branches and ML analyses (bootstrap support) below branches. Branch tips are labeled as such: Lm = *Lethrinus mahsena*, La = *Lethrinus atkinsoni*, Lh = *Lethrinus harak*. Haplotype (H) numbers for *L. mahsena* and *L. harak* correspond to Tables S2 and S3, for *L. atkinsoni* H1 = GenBank sequences from Queensland, Australia (KP194639.1, KP194307.1). laH2 = GenBank sequences from the Philippines (KF009603.1, KC970391.1). *L. mahsena* Madagascar and *L. mahsena* Japan = sequences deposited as *L. mahsena* on GenBank (JQ350088.1, JQ350089.1, JF952782.1) that do not correspond phylogenetically to this classification

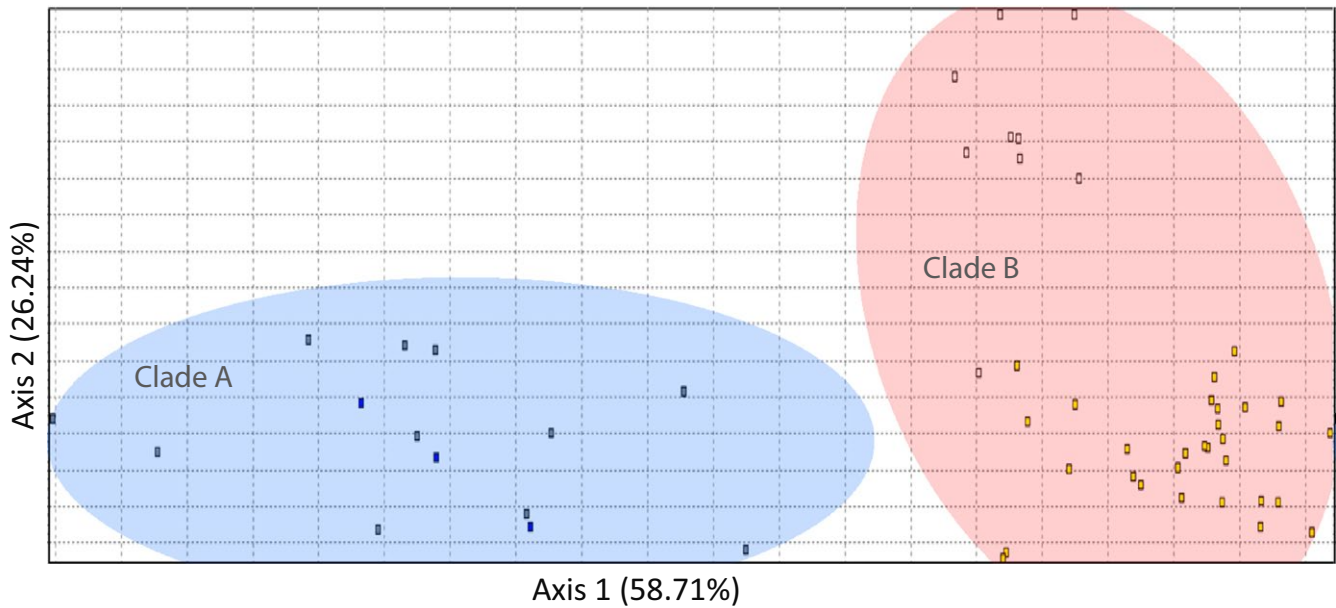
genetic differences than mtDNA, with a similar resolution for both species (Table S5), and with generally low Type I error probabilities.

For *L. mahsena*, eight of 60 locus by sample tests revealed significant deviations from HWE (Table S6), primarily due to heterozygote deficits spread relatively consistently across five loci (80RTE, BST2.33, 100RTE, 95TGRTE, and 68RTE). For *L. harak*, 12 of 36 tests demonstrated significant deviations from HWE (Table S7), again largely due to heterozygote deficits with locus 58RTE exhibiting a deficit of heterozygotes in all four samples.

Global estimates of differentiation (with null allele correction) were higher for *L. mahsena* ( $F_{ST} = 0.164$ ) than for *L. harak* ( $F_{ST} = 0.023$ ) but highly significant in both cases. Pairwise  $F_{ST}$  estimates were also high and significant between all samples for both *L. harak* and *L. mahsena* (Table 2). No significant IBD effects were detected for either species [*L.*

*mahsena*,  $R^2 = .071$ ,  $p = .340$ ; *L. harak*,  $R^2 = .434$ ,  $p = .090$ ]. BAPS analysis detected two genetically distinct groups within the *L. harak* data set, with Mauritius separated from the other SWIO samples (Figure S1); however, FCA divided samples into four groups corresponding to sampling sites (Figure 5). For *L. mahsena*, BAPS analysis identified three groups with samples from the Seychelles, Mozambique, and the rest of the SWIO each appearing to fall within distinct populations (Figure S1) supported by FCA (Figure 5).

Values of relatedness based on  $r_{sq}$  fell within the 95% confidence limits expected for panmictic populations for all samples of *L. harak* (Figure S2); however, mean values greater than expected under panmixia were identified in *L. mahsena* (Figure S3). Maximum likelihood inferences of relatedness among pairs of individuals (Figures S4 and S5) identified high levels of ambiguous kinship inference (i.e., individuals



**FIGURE 4** Factorial correspondence analysis showing multivariate relationships among microsatellite genotypes of *Lethrinus mahsena* samples from Mahe (SM) and Praslin (SP) Islands in the Seychelles. Samples were sorted by mitochondrial CO1 clade (clade A/clade B), classified by through PCR-RFLP analysis

**TABLE 1** (a) Mitochondrial genetic diversity levels, mtDNA neutrality tests, and a summary of mismatch distributions for 486 bp of *Lethrinus harak* and *Lethrinus mahsena* COI mtDNA: mtDNA sample size =  $N$ , number of haplotypes =  $H$ , number of private haplotypes = pHap, haplotype diversity =  $h$ , nucleotide diversity =  $\pi$ . (b) Microsatellite genetic diversity across nine (*L. harak*) and 10 (*L. mahsena*) loci in *L. harak* and *L. mahsena*, respectively: microsatellite sample size =  $N$ , number of alleles =  $N_A$ , expected heterozygosity =  $H_E$ , observed heterozygosity =  $H_O$ , and probability of deviation from Hardy–Weinberg expectations =  $p$ . For both mitochondrial and microsatellite data sets statistically significant estimates ( $p < .05$ ) are highlighted in bold and instances where tests could not be performed are indicated as such (–)

	(a) mtDNA (COI)					(b) Microsatellite				
	$N$	$H$	Phap	$h$	$\pi$	$N$	$N_A$	$H_O$	$H_E$	$p$
Mauritius										
<i>L. harak</i>	22	2	0	0.42 (0.09)	0.0008 (0.0009)	43	4.00	0.52	0.45	<b>.000</b>
<i>L. mahsena</i>	18	4	1	0.54 (0.12)	0.0017 (0.0014)	19	4.40	0.59	0.56	<b>.023</b>
Seychelles Mahe										
<i>L. mahsena</i>	3	3	2	1.00 (0.27)	0.0027 (0.0028)	12	3.60	0.58	0.50	<b>.044</b>
Tanzania										
<i>L. harak</i>	18	4	2	0.63 (0.07)	0.0015 (0.0013)	60	4.11	0.42	0.42	<b>.000</b>
<i>L. mahsena</i>	18	5	1	0.57 (0.13)	0.0016 (0.0014)	21	4.40	0.54	0.53	.412
Mozambique Pemba										
<i>L. harak</i>	22	3	0	0.39 (0.11)	0.0008 (0.0008)	44	3.89	0.41	0.42	<b>.000</b>
<i>L. mahsena</i>	3	3	0	1.00 (0.27)	0.0041 (0.0039)	7	3.10	0.46	0.45	.246
Mozambique Maputo										
<i>L. mahsena</i>	6	3	0	0.60 (0.22)	0.0018 (0.0020)	10	3.60	0.45	0.49	.053
Madagascar Belo sur Mer										
<i>L. harak</i>	42	3	0	0.47 (0.06)	0.0009 (0.0009)	100	4.78	0.42	0.42	<b>.000</b>
<i>L. mahsena</i>	5	2	1	0.60 (0.18)	0.0012 (0.0014)	5	2.30	0.53	0.40	.991
Total										
<i>L. harak</i>	104	5	2	0.52 (0.03)	0.0011 (0.0019)	247	4.19	0.44	0.42	<b>.000</b>
<i>L. mahsena</i>	53	9	7	0.59 (0.07)	0.0018 (0.0036)	72	3.57	0.52	0.49	<b>.025</b>

could not be unambiguously classified as related or unrelated for both *L. harak* (88.54%) and *L. mahsena* (79.56%) but did identify related dyads within all samples for both species except Madagascar, Belo sur Mer for *L. mahsena*.

### 3.4 | Paleodistribution modeling

Mean AUC values were 0.985 ( $SD = 0.003$ ) and 0.980 ( $SD = 0.016$ ) for *L. harak* and *L. mahsena*, respectively, indicating very good model performance. Present-day models for both *L. harak* and *L. mahsena* appear to accurately reflect the known contemporary ranges of both species, with suitable habitats aggregated in coastal environments across the Indian and western Pacific Oceans (Figure 6). Despite an overall reduction in suitable habitat during the LGM, this appears to take the form of a geographically homogeneous contraction across both species ranges rather than obvious vicariant fragmentation of present-day ranges (Figure 6).

## 4 | DISCUSSION

### 4.1 | Cryptic sympatric species

An important result was the detection of two highly divergent mtDNA clades (A and B) co-occurring within samples all described initially as *L. mahsena* from two sites in the Seychelles. Targeted PCR-RFLP analysis of all of our samples suggests that clade B was

confined to the Seychelles samples. Although both clade A and clade B exhibited high sequence similarity with *L. mahsena* voucher specimens from Mozambique and India, respectively, sequence divergence between the two clades (mean 4.4%) was in excess of even conservative thresholds of species-level divergence (Roux et al., 2016; Zmlak, Ward, Connell, Holmes, & Hebert, 2009). Microsatellite analysis also revealed a significant nuclear differentiation among members of the two clades, with biparentally restricted gene flow also apparent in results of assignment tests (high levels of assignment to clade). Clade B sequences were more similar to sequences of *L. atkinsoni*, a species described from west Pacific waters but not in the WIO. Sequence divergence between clade B and *L. atkinsoni* (mean = 2.2%; K2P distance = 0.021) is in the gray zone of the species divergence continuum (Roux et al., 2016), suggesting that clade B may represent a highly divergent *L. atkinsoni* clade lying well outside of the currently described species distribution, or that clade B represents an undescribed species most closely related to *L. atkinsoni*. Overall, the cytonuclear differentiation indicates the sympatric occurrence of a second species alongside *L. mahsena*, which is being indiscriminately harvested as *L. mahsena*. Furthermore, the data point to incongruence between current taxonomy and DNA barcodes. Specimens from India belonging to this potential cryptic species may have been misidentified as *L. mahsena* and used as voucher specimens. Further evidence of misidentification is apparent among sequences deposited as *L. mahsena* on GenBank: those from Madagascar (JQ350088.1, JQ350089.1)

**TABLE 2** Pairwise estimates of genetic differentiation ( $F_{ST}$ ) between *Lethrinus mahsena* samples (below diagonal—across 10 microsatellite loci corrected for null alleles (msat) and 486 bp of mtDNA COI) and *Lethrinus harak* samples (above diagonal—across nine microsatellite loci corrected for null alleles (msat) and 510 bp mtDNA COI)

	Mauritius	Seychelles Mahe	Tanzania	Mozambique Pemba	Mozambique Maputo	Madagascar Belo sur Mer
Mauritius						
COI			0.049	<b>0.407*</b>		0.256
msat			<b>0.039*</b>	<b>0.042*</b>		<b>0.029*</b>
Seychelles Mahe						
COI	0.152					
msat	<b>0.121*</b>					
Tanzania						
COI	-0.038	0.127		<b>0.127</b>		0.035
msat	<b>0.043*</b>	<b>0.241*</b>		<b>0.022*</b>		<b>0.012*</b>
Mozambique Pemba						
COI	-0.044	0.000	-0.043			0.002
msat	<b>0.118*</b>	<b>0.128*</b>	<b>0.239*</b>			<b>0.013*</b>
Mozambique Maputo						
COI	-0.116	0.100	-0.092	-0.154		
msat	<b>0.194*</b>	<b>0.191*</b>	<b>0.289*</b>	<b>0.075*</b>		
Madagascar Belo sur Mer						
COI	0.176	0.157	0.150	0.164	0.178	
msat	<b>0.065*</b>	<b>0.241*</b>	<b>0.088*</b>	<b>0.218*</b>	<b>0.267*</b>	

Statistically significant estimations ( $p < .05$ ) are indicated in bold, and those significant after Bonferroni correction are denoted with a \*.



have a greater similarity (99%) with *L. lentjan* voucher specimens from Iran, whereas supposed *L. mahsena* sequences from Japan (JF952782.1) exhibit greatest similarity (91%) to *L. rubrioperculatus*. In general, the meristic and morphological features of emperor fish are conservative (Carpenter & Allen, 1989) and distinct species are often difficult to identify easily by morphology alone (Sato, 1971; Smith, 1959). Failure to describe the full species diversity and/or misidentification of species may fundamentally compromise conservation strategies and tools. For example, the ENM performed here was based on reported occurrences of the species; however, based on the genetic results such occurrence data may be inaccurate. By extension, the modeled habitat patches may actually reflect distinct species ranges. The cryptic biocomplexity resolved here, as well as other cases of identification of cryptic lethriniid species (Borsa, Hsiao, Carpenter, & Chen, 2013), emphasizes the need for genetic studies on a wider geographical scale to fully resolve levels of intraspecific and interspecific diversification, as well as for the development of low-cost species identification and assignment techniques as resources for management such as the PCR-RFLP procedure developed here.

#### 4.2 | Population history of *Lethrinus harak* and *Lethrinus mahsena* across the SWIO

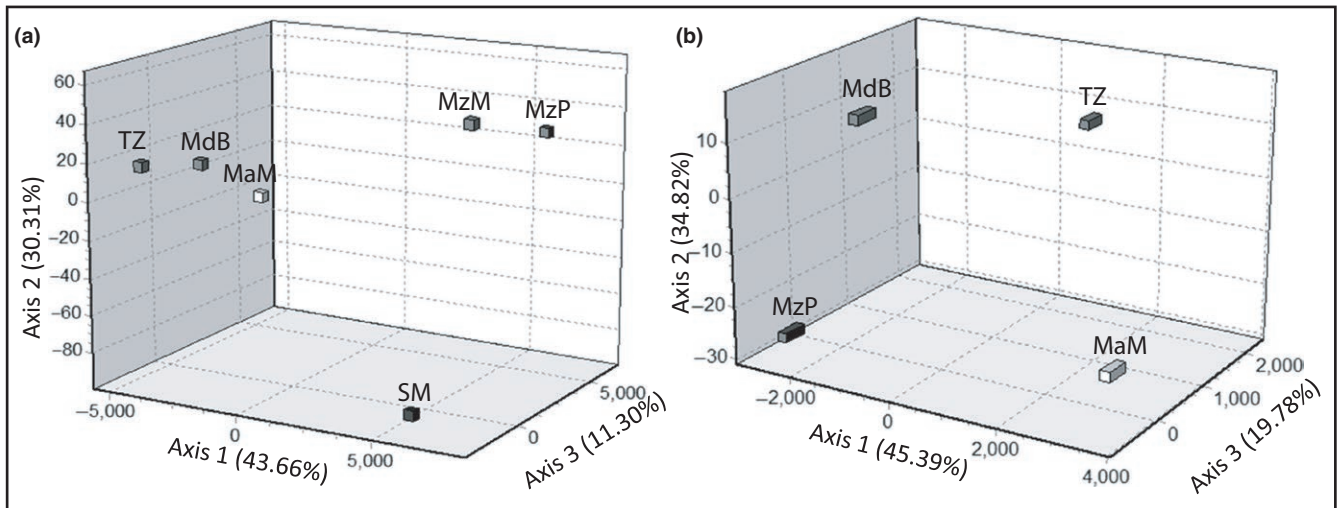
Hindcasting of ecological niche models (ENMs) suggested a marked reduction in suitable habitat for both *L. mahsena* and *L. harak* during the LGM that would be expected to have decreased their range-wide population sizes. Therefore, as for many other SWIO taxa including numerous reef-associated fish (Craig, Eble, Bowen, & Robertson, 2007; Visram et al. 2010) and crustaceans (Fratini et al., 2010; Gopal et al., 2006; Tolley, Groeneveld, Gopal, & Matthee, 2005), it is likely that eustatic sea-level fluctuations during the Pleistocene effected demographic changes in SWIO lethriniids. mtDNA phylogenies for both species conformed to classical “star”-shaped patterns (Slatkin & Hudson, 1991) expected under population expansion models. Demographic tests also provided some support for population expansion events for *L. mahsena* but not *L. harak*; however, the resolution of these tests may have been limited by a combination of low numbers of informative sites and differing sample sizes. The lack of genetic breaks within the mtDNA phylogenies for both species is compatible with a lack of prolonged vicariance. Similarly shallow phylogenies have been reported in many other SWIO species. Therefore, it would appear that for many species there has been little phylogeographic diversification within the SWIO (Hoareau, Boissin, & Berrebi, 2012; Muths et al., 2015) with the majority of cases of deep phylogeographic structure within the region due to colonization of allochthonous lineages (Ragionieri et al., 2009, 2010; Silva et al., 2010a).

#### 4.3 | Population structure of *L. harak* and *L. mahsena* across the SWIO

Although mtDNA  $\Phi_{ST}$  values were high for both species, pairwise comparisons between samples were largely nonsignificant. In contrast,

highly significant nuclear (microsatellite) differentiation was reported among samples for both species. Such a pattern is compatible with the greater statistical power of the nuclear data set as estimated by POWSIM. Similarly, Muths et al. (2011) detected a significant fine scale genetic structuring in the WIO for *Myripristis berndti*, while a previous mtDNA-based study indicated extensive connectivity (Craig et al., 2007). For both *L. harak* and *L. mahsena*, almost all pairwise tests ( $F_{ST}$  and exact) were significant, including comparisons between nearby samples, with no clear geographical pattern or significant IBD (there was also no geographical pattern to the mtDNA  $\Phi_{ST}$  values). The structuring for both species was also significant upon conservative correction for null alleles (as discussed by Shaw et al., 2010) and was not driven by single-locus effects (no loci identified as outliers).  $F_{ST}$  values were considerably higher for *L. mahsena* than for corresponding comparisons in *L. harak*. While POWSIM indicated a low Type I error rate for both data sets in qualitative (allele frequency) tests, the smaller sample sizes for *L. mahsena* may have inflated  $F_{ST}$  values (Willing, Dreyer, & Van Oosterhout, 2012). From a biological viewpoint, the considerably greater abundance of *L. harak* may also point to smaller effective population sizes for *L. mahsena* which could predispose the species to greater genetic drift effects and hence larger  $F_{ST}$  even if migration rates are similar (Whiteley, Spruell, & Allendorf, 2004). Further insight into the drivers of genetic differentiation was provided by the kinship analyses, which examine how alleles are shared among individuals and provide an independent test of the hypothesis that structure as quantified by  $F_{ST}$ , which focuses on population allele frequencies, is a result of connectivity (Christie, Johnson, Stallings, & Hixon, 2010; Iacchi et al., 2013). Firstly, for *L. mahsena* mean kinship values for all samples exceeded those predicted within a nonstructured system. Second, related individual pairs and/or potentially related dyads (i.e., dyads that could not be unambiguously described as unrelated) were found in all samples for both species. While the large number of dyads that could not be unanimously classified to a single relationship category (i.e., unrelated or related) highlights the resolution threshold of the data, the results for both allele frequency and allele-sharing analyses indicate that, even though sample sizes are small in some cases, the genetic heterogeneity cannot be dismissed as statistical noise but rather reflects some changes in the composition that may be a useful tool for better understanding recruitment dynamics and connectivity in these species (Knutsen et al., 2011; Selkoe, Gaines, Caselle, & Warner, 2006).

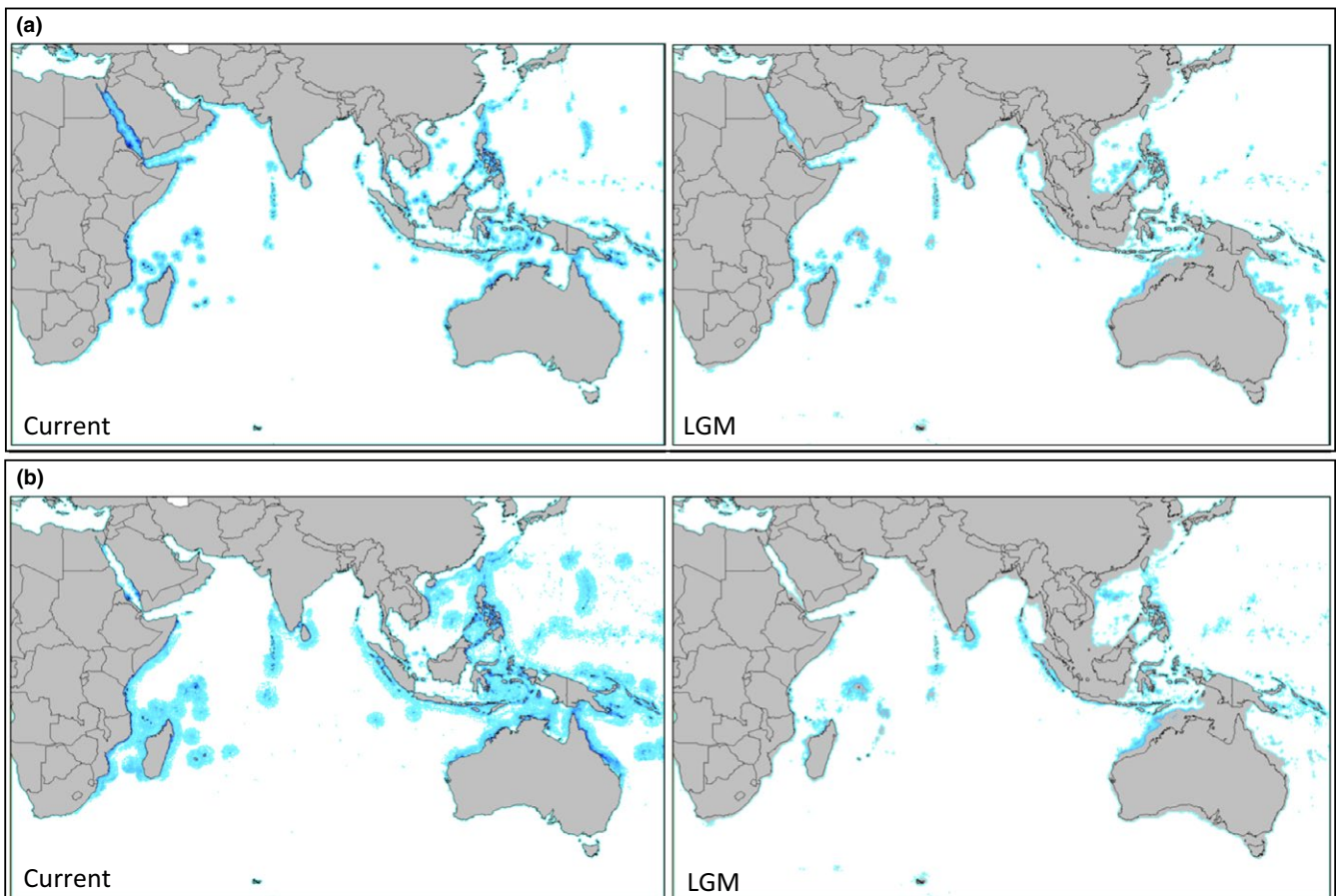
The geographically patchy genetic structuring and kinship patterns could be generated by three nonmutually exclusive processes: population isolation, large variances in individual reproductive success (sweepstakes recruitment), and limited mixing of larvae from genetically different sources (larval cohesion). Facets of the genetic structure exhibited some congruence with known biogeographic boundaries and patterns of population isolation inferred from other species. For *L. harak*, there was pronounced differentiation of the Mauritian population from the rest of the SWIO, supported by clustering and pairwise  $F_{ST} > 0.03$  in all comparisons. Although not resolved in the clustering analysis, the Mauritius *L. mahsena* sample was also highly differentiated from all other samples (pairwise  $F_{ST}$  values  $> 0.04$ ). Differentiation



**FIGURE 5** Factorial correspondence analysis showing multivariate relationships among microsatellite genotypes of SWIO samples of (a) *Lethrinus mahsena* and (b) *Lethrinus harak*

of Mascarene populations has been observed for several other reef-associated species (Muths et al., 2011, 2015) and linked to a combination of large geographical distances from other coasts (>1,000 km), and a barrier effect due to the landmass of Madagascar.

Genetic differentiation of the *L. mahsena* Seychelles sample was also supported by pairwise tests and clustering analysis. A combination of the South Equatorial Current and Equatorial Counter Current could serve to restrict connectivity between more western/southern locations as



**FIGURE 6** Results of species distribution models depicting predicted ranges of (a) *Lethrinus mahsena* and (b) *Lethrinus harak* during the present day and the Last Glacial Maximum (LGM, ~21KYA)

suggested by Muths et al. (2015). Such an isolating mechanism might also explain the seeming absence of clade B from other SWIO sites. The high fecundity and batch spawning behavior of both species could be conducive to sweepstakes effects, while local seascape features may contribute to recruitment heterogeneity (McKeown, Hauser, & Shaw, 2017). The Mozambique Channel has a dominant seasonal anticyclonic cell at the northern entrance, as well as a succession of mesoscale anticyclonic and cyclonic eddies along the Mozambique coast (Schouten, de Ruijter, Van Leeuwen, & Ridderinkhof, 2003; Swart, Lutjeharms, Ridderinkhof, & De Ruijter, 2010). Such features have been suggested to contribute to stochastic recruitment/genetic patchiness within the region in a number of species (Bourjea et al., 2007; Hoareau, Boissin, Paulay, & Bruggemann, 2013; Muths et al., 2015) and may contribute to the differentiation among samples from both species within the Mozambique Channel reported here. Even in the absence of genetically isolated source populations, larval cohesion (Selkoe et al., 2006) may enhance (Waples, 2002) and be effectively indistinguishable from sweepstake effects (Turner, Dowling, Marsh, Kesner, & Kelsen, 2007). Berry, England, Marriott, Burridge, and Newman (2012) reported evidence for cohesion among larval and juvenile *L. nebulosus* but suggested that this was subsequently dissipated by cumulative dispersal of juveniles with no nonrandom cohesion detected by 8 years of age, by which time *L. nebulosus* are reproductively mature. It is possible that the cohesion reported here is similarly transient, in which case the patterns would reflect harvesting of individuals that may not have attained reproductive maturity, a practice that can severely compromise sustainability.

Disentangling the exact roles of recruitment heterogeneity and restricted gene flow will require more sensitive genetic assays and would benefit from analysis of age-segregated samples (Burford, Carr, & Bernardi, 2011). However, the overall results indicate processes within the region that prevent genetic mixing, at least, on timescales of interest to fishery managers. These contrast with the broad-scale genetic homogeneity and subtle patchiness reported for *L. nebulosus* in Australian waters (Berry et al., 2012) and implicate regional seascape drivers within the SWIO. In these cases, a spatial “bet-hedging” approach is advised for marine resource management, including geographical dispersion of marine reserves if they are to be used (Larson & Julian, 1999). The ability to predict population structure across taxa applies directly to the implementation of MPAs. *L. harak* and *L. mahsena* reported similar historical modeled habitat occupancy and phylogeographic patterns, while resolved genetic patchiness points to similar recurrent population processes in both species. Therefore, while we agree that making sweeping predictions from alleged model organisms can be dangerous (Bird et al., 2007), this similarity suggests that the application of information from such models to subsets of taxa may be highly useful in light of finite resources for conservation. Paradoxically, the identification of a cryptic sympatric species assemblage with little (if any) morphological differentiation, that is potentially endemic to the Seychelles, highlights the utility and necessity of genetic approaches to characterize baseline biodiversity in the region before such model-based methods are employed.

## ACKNOWLEDGMENTS

The Institute of Biological Environmental and Rural Sciences at Aberystwyth University receives strategic funding from the Biotechnology and Biological Sciences Research Council. This research was funded through UK Natural Environment Research Council grants. We would like to thank Sophie Benbow, the staff at Blue Adventures, the Seychelles Fisheries Authority, Dr. Adriano Junior, and the Centre for Marine research in Pemba for helping us collect samples.

## CONFLICT OF INTEREST

None declared.

## AUTHOR CONTRIBUTIONS

P.W.S. and N.J.M. gained funding for and conceived the project with input from A.J.E.H. and A.T.; data were collected and analyzed by A.J.E.H. and A.T.; and J.P. conducted the ecological niche modeling. A.J.E.H. and N.J.M. led the writing of the manuscript, and all authors contributed to editing the manuscript.

## DATA ARCHIVING

All data will be uploaded on DRYAD.

## ORCID

Amy J. E. Healey  <http://orcid.org/0000-0001-6061-0044>

## REFERENCES

- Beheregaray, L. B. (2008). Twenty years of phylogeography: The state of the field and the challenges for the Southern Hemisphere. *Molecular Ecology*, 17, 3754–3774. <https://doi.org/10.1111/j.1365-294X.2008.03857.x>
- Belkhir, K., Borsa, P., Chikhi, L., Raufaste, N., & Bonhomme, F. (2004). *GENETIX 4.05, Population genetics software for Windows TM*. Montpellier, France: Université de Montpellier I.
- Benny, P. N. (2002). Variability of Western Indian Ocean currents. *Western Indian Ocean Journal of Marine Science*, 1, 81–90. <https://doi.org/hdl.handle.net/1834/29>
- Berg, H., Francis, J., & Souter, P. (2002). Support to marine research for sustainable management of marine and coastal resources in the Western Indian Ocean. *Ambio*, 31, 597–601. [https://doi.org/10.1639/0044-7447\(2002\)031\[0597:stmfrs\]2.0.co;2](https://doi.org/10.1639/0044-7447(2002)031[0597:stmfrs]2.0.co;2)
- Berry, O., England, P., Marriott, R. J., Burridge, C. P., & Newman, S. J. (2012). Understanding age-specific dispersal in fishes through hydrodynamic modelling, genetic simulations and microsatellite DNA analysis. *Molecular Ecology*, 21(9), 2145–2159. <https://doi.org/10.1111/j.1365-294X.2012.05520.x>
- Bird, C. E., Holland, B. S., Bowen, B. W., & Toonen, R. J. (2007). Contrasting phylogeography in three endemic Hawaiian limpets (*Cellana* spp.) with similar life histories. *Molecular Ecology*, 16, 3173–3186. <https://doi.org/10.1111/j-1365-294X.2007.03385.x>
- Bird, C. E., Holland, B. S., Bowen, B. W., & Toonen, R. J. (2011). Diversification of sympatric broadcast-spawning limpets (*Cellana* spp.) within the Hawaiian archipelago. *Molecular Ecology*, 20, 2128–2141. <https://doi.org/10.1111/j.1365-294X.2011.05081.x>

- Bohonak, A. J. (1999). Dispersal, gene flow, and population structure. *Quarterly Review of Biology*, 74, 21–45. <https://doi.org/10.1086/392950>
- Borsa, P., Hsiao, D. R., Carpenter, K. E., & Chen, W. J. (2013). Cranial morphometrics and mitochondrial DNA sequences distinguish cryptic species of the longface emperor (*Lethrinus olivaceus*), an emblematic fish of Indo-West Pacific coral reefs. *Comptes Rendus Biologies*, 336, 505–514. <https://doi.org/10.1016/j.crv.2013.09.004>
- Bourjea, J., Lapegue, S., Gagnevin, L., Broderick, D., Mortimer, J. A., Ciccione, S., ... Grizel, H. (2007). Phylogeography of the green turtle, *Chelonia mydas*, in the Southwest Indian Ocean. *Molecular Ecology*, 16(1), 175–186.
- Burford, M. O., Carr, M. H., & Bernardi, G. (2011). Age-structured genetic analysis reveals temporal and geographic variation within and between two cryptic rockfish species. *Marine Ecology Progress Series*, 442, 201–215. <https://doi.org/10.3354/meps09329>
- Carpenter, K. E., & Allen, G. R. (1989). *FAO species catalogue. Emperor fishes and large-eye breams of the world (family Lethrinidae)*, 125(9). Rome: Food and Agriculture Organization.
- Cavalli-Sforza, L. L., & Edwards, A. W. F. (1967). Phylogenetic analysis models and estimation procedures. *American Journal of Human Genetics*, 19, 233–257.
- Chapuis, M. P., & Estoup, A. (2007). Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution*, 24, 621–631. <https://doi.org/10.1093/molbev/msl191>
- Chiang, H. C., Hsu, C. C., Wu, G. C. C., Chang, S. K., & Yang, H. Y. (2008). Population structure of bigeye tuna (*Thunnus obesus*) in the Indian Ocean inferred from mitochondrial DNA. *Fisheries Research*, 90, 305–312. <https://doi.org/10.1016/j.fishres.2007.11.006>
- Christie, M. R., Johnson, D. W., Stallings, C. D., & Hixon, M. A. (2010). Self-recruitment and sweepstakes reproduction amid extensive gene flow in a coral-reef fish. *Molecular Ecology*, 19, 1042–1057. <https://doi.org/10.1111/j.1365-294X.2010.04524.x>
- Corander, J., Waldmann, P., Marttinen, P., & Sillanpää, M. J. (2004). BAPS 2: Enhanced possibilities for the analysis of genetic population structure. *Bioinformatics*, 20, 2363–2369. <https://doi.org/10.1093/bioinformatics/bth250>
- Craig, M. T., Eble, J. A., Bowen, B. W., & Robertson, D. R. (2007). High genetic connectivity across the Indian and Pacific Oceans in the reef fish *Myripristis berndti* (Holocentridae). *Marine Ecology Progress Series*, 334, 245–254. <https://doi.org/10.3354/meps334245>
- Dawson, M. N. (2012). Parallel phylogeographic structure in ecologically similar sympatric sister taxa. *Molecular Ecology*, 21, 987–1004. <https://doi.org/10.1111/j.1365-294X.2011.05417.x>
- Duda, T. F., & Palumbi, S. R. (1999). Population structure of the black tiger prawn, *Penaeus monodon*, among western Indian Ocean and western Pacific populations. *Marine Biology*, 134, 705–710. <https://doi.org/10.1007/s002270050586>
- Ebisawa, A. (2006). Reproductive and sexual characteristics in five *Lethrinus* species in waters off the Ryukyu Islands. *Ichthyological Research*, 53, 269–280. <https://doi.org/10.1007/s10228-006-0345-3>
- Ebisawa, A., & Ozawa, T. (2009). Life-history traits of eight *Lethrinus* species from two local populations in waters off the Ryukyu Islands. *Fisheries Science*, 75, 553–566. <https://doi.org/10.1007/s12562-009-0061-9>
- Excoffier, L., & Lischer, H. E. L. (2010). Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, 10, 564–567. <https://doi.org/10.1111/j.1755-0998.2010.02847.x>
- Fratini, S., Ragonieri, L., & Cannicci, S. (2010). Stock structure and demographic history of the Indo-West Pacific mud crab *Scylla serrata*. *Estuarine Coastal and Shelf Science*, 86, 51–61. <https://doi.org/10.1016/j.ecss.2009.10.009>
- Fratini, S., & Vannini, M. (2002). Genetic differentiation in the mud crab *Scylla serrata* (Decapoda: Portunidae) within the Indian Ocean. *Journal of Experimental Marine Biology and Ecology*, 272, 103–116. [https://doi.org/10.1016/s0022-0981\(02\)00052-7](https://doi.org/10.1016/s0022-0981(02)00052-7)
- Glynn, F., Houghton, J. D. R., Bastian, T., Doyle, T. K., Fuentes, V., Lilley, M. K. S., & Provan, J. (2016). High-resolution genetic analysis reveals extensive gene flow within the jellyfish *Pelagia noctiluca* (Scyphozoa) in the North Atlantic and Mediterranean Sea. *Biological Journal of the Linnean Society*, 117, 252–263. <https://doi.org/10.1111/bij.12654>
- Glynn, F., Houghton, J. D. R., & Provan, J. (2015). Population genetic analyses reveal distinct geographical blooms of the jellyfish *Rhizostoma octopus* (Scyphozoa). *Biological Journal of the Linnean Society*, 116, 582–592. <https://doi.org/10.1111/bij.12614>
- Gopal, K., Tolley, K. A., Groeneveld, J. C., & Matthee, C. A. (2006). Mitochondrial DNA variation in spiny lobster *Palinurus delagoae* suggests genetically structured populations in the southwestern Indian Ocean. *Marine Ecology Progress Series*, 319, 191–198. <https://doi.org/10.3354/meps319191>
- Gouws, G. (2012). Genetic stock assessment of the blue emperor (*Lethrinus nebulosus*) in the South West Indian Ocean. South West Indian Ocean Fisheries Project report. South African Institute for Aquatic Biodiversity.
- Grandcourt, E. M. (2002). Demographic characteristics of a selection of exploited reef fish from the Seychelles: Preliminary study. *Marine and Freshwater Research*, 53, 123–130. <https://doi.org/10.1071/mf01123>
- Hall, T. A. (1999). BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, 41, 95–98.
- Hoareau, T. B., Boissin, E., & Berrebi, P. (2012). Evolutionary history of a widespread Indo-Pacific goby: The role of Pleistocene sea-level changes on demographic contraction/expansion dynamics. *Molecular Phylogenetics and Evolution*, 62(1), 566–572. <https://doi.org/10.1016/j.ympev.2011.10.004>
- Hoareau, T. B., Boissin, E., Paulay, G., & Bruggemann, J. H. (2013). The southwestern Indian Ocean as a potential marine evolutionary hotspot: Perspectives from comparative phylogeography of reef brittle-stars. *Journal of Biogeography*, 40(11), 2167–2179. <https://doi.org/10.1111/jbi.12155>
- Iacchei, M., Ben-Horin, T., Selkoe, K. A., Bird, C. E., García-Rodríguez, F. J., & Toonen, R. J. (2013). Combined analyses of kinship and FST suggest potential drivers of chaotic genetic patchiness in high gene-flow populations. *Molecular Ecology*, 22(13), 3476–3494. <https://doi.org/10.1111/mec.12341>
- Iles, T. D., & Sinclair, M. (1982). Atlantic herring – stock discreteness and abundance. *Science*, 215(4533), 627–633. <https://doi.org/10.1126/science.215.4533.627>
- Kalinowski, S. T., Wagner, A. P., & Taper, M. L. (2006). ML-Relate: A computer program for maximum likelihood estimation of relatedness and relationship. *Molecular Ecology Notes*, 6(2), 576–579. <https://doi.org/10.1111/j.1471-8286.2006.01256.x>
- Knutsen, H., Olsen, E. M., Jorde, P. E., Espeland, S. H., André, C., & Stenseth, N. C. (2011). Are low but statistically significant levels of genetic differentiation in marine fishes 'biologically meaningful'? A case study of coastal Atlantic cod. *Molecular Ecology*, 20(4), 768–783. <https://doi.org/10.1111/j.1365-294X.2010.04979.x>
- Kulmiye, A. J., Ntiba, M. J., & Kisia, S. M. (2002). Some aspects of the reproductive biology of the thumbprint emperor, *Lethrinus harak* (Forsskal, 1775), in Kenyan Coastal Waters. *Western Indian Ocean Journal of Marine Science*, 1, 135–144.
- Larson, R. J., & Julian, R. M. (1999). Spatial and temporal genetic patchiness in marine populations and their implications for fisheries management. *California Cooperative Oceanic Fisheries Investigations Reports*, 40, 94–99.
- Larsson, L. C., Charlier, J., Laikre, L., & Ryman, N. (2009). Statistical power for detecting genetic divergence—organelle versus nuclear markers. *Conservation Genetics*, 10(5), 1255. <https://doi.org/10.1007/s10592-008-9693-z>

- Lessios, H. A., Kessing, B. D., & Pearse, J. S. (2001). Population structure and speciation in tropical seas: Global phylogeography of the sea urchin *Diadema*. *Evolution*, 55, 955–975. [https://doi.org/10.1554/0014-3820\(2001\)055\[0955:psasit\]2.0.co;2](https://doi.org/10.1554/0014-3820(2001)055[0955:psasit]2.0.co;2)
- Lester, S. E., Ruttenberg, B. I., Gaines, S. D., & Kinlan, B. P. (2007). The relationship between dispersal ability and geographic range size. *Ecology Letters*, 10, 745–758. <https://doi.org/10.1111/j.1461-0248.2007.01070.x>
- McKeown, N. J., Arkhipkin, A. I., & Shaw, P. W. (2017). Regional genetic population structure and fine scale genetic cohesion in the Southern blue whiting *Micromesistius australis*. *Fisheries Research*, 185, 176–184. <https://doi.org/10.1016/j.fishres.2016.09.006>
- McKeown, N. J., Hauser, L., & Shaw, P. W. (2017). Microsatellite genotyping of brown crab *Cancer pagurus* reveals fine scale selection and 'non-chaotic' genetic patchiness within a high gene flow system. *Marine Ecology Progress Series*, 566, 91–103. <https://doi.org/10.3354/meps12044>
- Muths, D., Grewe, P., Jean, C., & Bourjea, J. (2009). Genetic population structure of the Swordfish (*Xiphias gladius*) in the southwest Indian Ocean: Sex-biased differentiation, congruency between markers and its incidence in a way of stock assessment. *Fisheries Research*, 97, 263–269. <https://doi.org/10.1016/j.fishres.2009.03.004>
- Muths, D., Le Couls, S., Evano, H., Grewe, P., & Bourjea, J. (2013). Multi-genetic marker approach and spatio-temporal analysis suggest there is a single panmictic population of swordfish *Xiphias gladius* in the Indian Ocean. *PLoS ONE*, 8, <https://doi.org/10.1371/journal.pone.0063558>
- Muths, D., Tessier, E., & Bourjea, J. (2015). Genetic structure of the reef grouper *Epinephelus merra* in the West Indian Ocean appears congruent with biogeographic and oceanographic boundaries. *Marine Ecology-an Evolutionary Perspective*, 36, 447–461. <https://doi.org/10.1111/maec.12153>
- Muths, D., Tessier, E., Gouws, G., Craig, M., Mwale, M., Mwaluma, J., & Bourjea, J. (2011). Restricted dispersal of the reef fish *Myripristis berndti* at the scale of the SW Indian Ocean. *Marine Ecology Progress Series*, 443, 167–180. <https://doi.org/10.3354/meps09394>
- Nakamura, Y., Shibuno, T., Suzuki, N., Nakamori, J., Kanashiro, K., & Watanabe, Y. (2010). Interspecific variations in age and size at settlement of 8 emperor fishes (Lethrinidae) at the southern Ryukyu Islands, Japan. *Fisheries Science*, 76, 503–510. <https://doi.org/10.1007/s12562-010-0225-7>
- Nei, M. (1987). *Molecular evolutionary genetics*. New York, NY: Columbia University Press.
- Peakall, R., & Smouse, P. E. (2006). GENALEX 6: Genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, 6, 288–295. <https://doi.org/10.1111/j.1471-8286.2005.01155.x>
- Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, 190, 231–259. <https://doi.org/10.1016/j.ecolmodel.2005.03.026>
- 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. *Journal of Heredity*, 95, 536–539. <https://doi.org/10.1093/jhered/esh074>
- Posada, D., & Crandall, K. A. (1998). Modeltest: Testing the model of DNA substitution. *Bioinformatics (Oxford, England)*, 14(9), 817–818. <https://doi.org/10.1093/bioinformatics/14.9.817>
- Queller, D. C., & Goodnight, K. F. (1989). Estimating relatedness using genetic markers. *Evolution*, 43, 258–275. <https://doi.org/10.1111/j.1558-5646.1989.tb04226.x>
- Ragionieri, L., Cannicci, S., Schubart, C. D., & Fratini, S. (2010). Gene flow and demographic history of the mangrove crab *Neosarmatium meinerti*: A case study from the western Indian Ocean. *Estuarine Coastal and Shelf Science*, 86, 179–188. <https://doi.org/10.1016/j.eccs.2009.11.002>
- Ragionieri, L., Fratini, S., Vannini, M., & Schubart, C. D. (2009). Phylogenetic and morphometric differentiation reveal geographic radiation and pseudo-cryptic speciation in a mangrove crab from the Indo-West Pacific. *Molecular Phylogenetics and Evolution*, 52, 825–834. <https://doi.org/10.1016/j.ympev.2009.04.008>
- Raymond, M., & Rousset, F. (1995). GENEPOP (Version-1.2) – population-genetics software for exact tests and ecumenicism. *Journal of Heredity*, 86, 248–249. <https://doi.org/10.1093/oxfordjournals.jhered.a111573>
- Ridgway, T., & Sampayo, E. M. (2005). Population genetic status of the Western Indian Ocean: What do we know? *Western Indian Ocean Journal of Marine Science*, 4, 1–10.
- Ronquist, F., & Huelsenbeck, J. P. (2003). MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, 19(12), 1572–1574. <https://doi.org/10.1093/bioinformatics/btg180>
- Roux, C., Fraise, C., Romiguier, J., Anciaux, Y., Galtier, N., & Bierne, N. (2016). Shedding light on the grey zone of speciation along a continuum of genomic divergence. *PLoS Biology*, 14(12), e2000234.
- Ruzzante, D. E., Mariani, S., Bekkevold, D., Andre, C., Mosegaard, H., Clausen, L. A. W., & Carvalho, G. R. (2006). Biocomplexity in a highly migratory pelagic marine fish, Atlantic herring. *Proceedings of the Royal Society B-Biological Sciences*, 273(1593), 1459–1464. <https://doi.org/10.1098/rspb.2005.3463>
- Ryman, N., & Palm, S. (2006). POWSIM: A computer program for assessing statistical power when testing for genetic differentiation. *Molecular Ecology Notes*, 6, 600–602. <https://doi.org/10.1111/j.1365-294X.2006.01378.x>
- Ryman, N., Utter, F., & Laikre, L. (1995). Protection of intraspecific biodiversity of exploited fishes. *Reviews in Fish Biology and Fisheries*, 5, 417–446. <https://doi.org/10.1007/bf01103814>
- Sato, T. (1971). A revision of the Japanese sparoid fishes of the genus *Lethrinus*. *Journal of the Faculty of Science: University of Tokyo*, 4(12), 2.
- Sbrocco, E. J. (2014). Paleo-MARSPEC: Gridded ocean climate layers for the mid-Holocene and Last Glacial Maximum. *Ecology*, 95, 1710–1710. <https://doi.org/10.1890/14-0443.1>
- Sbrocco, E. J., & Barber, P. H. (2013). MARSPEC: Ocean climate layers for marine spatial ecology. *Ecology*, 94, 979–979. <https://doi.org/10.1890/12-1358.1>
- Schouten, M. W., de Ruijter, W. P., Van Leeuwen, P. J., & Ridderinkhof, H. (2003). Eddies and variability in the Mozambique Channel. *Deep Sea Research Part II: Topical Studies in Oceanography*, 50(12), 1987–2003. [https://doi.org/10.1016/S0967-0645\(03\)00042-0](https://doi.org/10.1016/S0967-0645(03)00042-0)
- Selkoe, K. A., Gaines, S. D., Caselle, J. E., & Warner, R. R. (2006). Current shifts and kin aggregation explain genetic patchiness in fish recruits. *Ecology*, 87(12), 3082–3094. [https://doi.org/10.1890/0012-9658\(2006\)87\[3082:CSAKAE\]2.0.CO;2](https://doi.org/10.1890/0012-9658(2006)87[3082:CSAKAE]2.0.CO;2)
- Shaw, P. W., Hendrickson, L., McKeown, N. J., Stonier, T., Naud, M. J., & Sauer, W. H. H. (2010). Discrete spawning aggregations of loliginid squid do not represent genetically distinct populations. *Marine Ecology Progress Series*, 408, 117–127. <https://doi.org/10.3354/meps08593>
- Silva, I. C., Mesquita, N., & Paula, J. (2010a). Genetic and morphological differentiation of the mangrove crab *Perisesarma guttatum* (Brachyura: Sesamidae) along an East African latitudinal gradient. *Biological Journal of the Linnean Society*, 99, 28–46.
- Silva, I. C., Mesquita, N., & Paula, J. (2010b). Lack of population structure in the fiddler crab *Uca annulipes* along an East African latitudinal gradient: Genetic and morphometric evidence. *Marine Biology*, 157, 1113–1126. <https://doi.org/10.1007/s00227-010-1393-9>
- Slatkin, M. (1993). Isolation by distance in equilibrium and nonequilibrium populations. *Evolution*, 47, 264–279. <https://doi.org/10.2307/2410134>
- Slatkin, M., & Hudson, R. R. (1991). Pairwise comparisons of mitochondrial-DNA sequences in stable and exponentially growing populations. *Genetics*, 129, 555–562.
- Smith, J. L. B. (1959). Fishes of the family Lethrinidae from the western Indian Ocean. *Ichthyological Bulletin Rhodes University*, 17, 285–295.

- Smith, J. L. B., Smith, M. M., & Heemstra, P. C. (2003). *Smiths' sea fishes*. Struik. Johannesburg, South Africa: Macmillan South Africa.
- Swart, N. C., Lutjeharms, J. R. E., Ridderinkhof, H., & De Ruijter, W. P. M. (2010). Observed characteristics of Mozambique Channel eddies. *Journal of Geophysical Research: Oceans*, 115(C9).
- Tamura, K., Stecher, G., Peterson, D., Filipski, A., & Kumar, S. (2013). MEGA6: Molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution*, 30(12), 2725–2729. <https://doi.org/10.1093/molbev/mst197>
- Therkildsen, N. O., Hemmer-Hansen, J., Hedeholm, R. B., Wisz, M. S., Pampoulie, C., Meldrup, D., ... Nielsen, E. E. (2013). Spatiotemporal SNP analysis reveals pronounced biocomplexity at the northern range margin of Atlantic cod *Gadus morhua*. *Evolutionary Applications*, 6(4), 690–705. <https://doi.org/10.1111/eva.12055>
- Thompson, J. D., Higgins, D. G., & Gibson, T. J. (1994). CLUSTAL-W – Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, 22, 4673–4680. <https://doi.org/10.1093/nar/22.22.4673>
- Tolley, K. A., Groeneveld, J. C., Gopal, K., & Matthee, C. A. (2005). Mitochondrial DNA panmixia in spiny lobster *Palinurus gilchristi* suggests a population expansion. *Marine Ecology Progress Series*, 297, 225–231. <https://doi.org/10.3354/meps297225>
- Turner, T. F., Dowling, T. E., Marsh, P. C., Kesner, B. R., & Kelsen, A. T. (2007). Effective size, census size, and genetic monitoring of the endangered razorback sucker, *Xyrauchen texanus*. *Conservation Genetics*, 8(2), 417–425. <https://doi.org/10.1007/s10592-006-9181-2>
- Van Herwerden, L., Benzie, J., & Davies, C. (2003). Microsatellite variation and population genetic structure of the red throat emperor on the Great Barrier Reef. *Journal of Fish Biology*, 62, 987–999. <https://doi.org/10.1046/j.1095-8649.2003.00075.x>
- Van Herwerden, L., Benzie, J., Peplow, L., & Davies, C. (2000). Microsatellite markers for coral trout (*Plectropomus laevis*) and red throat emperor (*Lethrinus miniatus*) and their utility in other species of reef fish. *Molecular Ecology*, 9, 1929–1931. <https://doi.org/10.1046/j.1365-294x.2000.01076-7.x>
- Visram, S., Yang, M. C., Pillay, R. M., Said, S., Henriksson, O., Grahn, M., & Chen, C. A. (2010). Genetic connectivity and historical demography of the blue barred parrotfish (*Scarus ghobban*) in the western Indian Ocean. *Marine Biology*, 157, 1475–1487. <https://doi.org/10.1007/s00227-010-1422-8>
- Waples, R. S. (1998). Separating the wheat from the chaff: Patterns of genetic differentiation in high gene flow species. *Journal of Heredity*, 89(5), 438–450. <https://doi.org/10.1093/jhered/89.5.438>
- Waples, R. S. (2002). Evaluating the effect of stage-specific survivorship on the N-e/N ratio. *Molecular Ecology*, 11, 1029–1037. <https://doi.org/10.1046/j.1365-294X.2002.01504.x>
- Weir, B. S., & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population-structure. *Evolution*, 38, 1358–1370. <https://doi.org/10.2307/2408641>
- Whiteley, A. R., Spruell, P., & Allendorf, F. W. (2004). Ecological and life history characteristics predict population genetic divergence of two salmonids in the same landscape. *Molecular Ecology*, 13(12), 3675–3688. <https://doi.org/10.1111/j.1365-294X.2004.02365.x>
- Willing, E. M., Dreyer, C., & Van Oosterhout, C. (2012). Estimates of genetic differentiation measured by  $F_{ST}$  do not necessarily require large sample sizes when using many SNP markers. *PLoS ONE*, 7(8), e42649. <https://doi.org/10.1371/journal.pone.0042649>
- Winnepenninckx, B., Backeljau, T., & Dewachter, R. (1993). Complete small ribosomal-subunit RNA sequence of the chiton *Acanthopleura japonica* (Lischke, 1873) (Mollusca, Polyplacophora). *Nucleic Acids Research*, 21, 1670–1670. <https://doi.org/10.1093/nar/21.7.1670>
- Wright, S. (1978). *Evolution and the genetics of populations. IV Variability Within and Among Populations*. Chicago, IL: University of Chicago Press.
- Zemlak, T. S., Ward, R. D., Connell, A. D., Holmes, B. H., & Hebert, P. D. N. (2009). DNA barcoding reveals overlooked marine fishes. *Molecular Ecology Resources*, 9, 237–242. <https://doi.org/10.1111/j.1755-0998.2009.02649.x>

#### SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

**How to cite this article:** Healey AJE, McKeown NJ, Taylor AL, et al. Cryptic species and parallel genetic structuring in Lethrinid fish: Implications for conservation and management in the southwest Indian Ocean. *Ecol Evol*. 2018;8:2182–2195. <https://doi.org/10.1002/ece3.3775>