

# Restricted dispersal of the reef fish *Myripristis berndti* at the scale of the SW Indian Ocean

D. Muths<sup>1,\*</sup>, E. Tessier<sup>2</sup>, G. Gouws<sup>3</sup>, M. Craig<sup>4</sup>, M. Mwale<sup>3</sup>, J. Mwaluma<sup>5</sup>,  
A. Mwandya<sup>6</sup>, J. Bourjea<sup>1</sup>

<sup>1</sup>Institut Français pour la Recherche et l'Exploitation de la Mer (IFREMER)—Delegation de la Réunion,  
Le Port, 97420 Réunion, France

<sup>2</sup>Réserve Naturelle Marine de La Réunion, La Saline les Bains, 97434 Réunion, France

<sup>3</sup>South African Institute for Aquatic Biodiversity, Grahamstown, 6140 South Africa

<sup>4</sup>University of Puerto Rico, Mayagüez, 00681 Puerto Rico

<sup>5</sup>Kenya Marine and Fisheries Research Institute, Mombasa, 080100 Kenya

<sup>6</sup>Sokoine University of Agriculture, Morogoro, 3000 Tanzania

**ABSTRACT:** The reef fish *Myripristis berndti* (Jordan & Everman 1903) is a pantropical species. A genetic analysis was conducted on 353 individuals from 10 localities distributed across the SW Indian Ocean (SWIO) in order to determine patterns of connectivity in the SWIO. Both the mtDNA sequences (711-bp cytochrome *b* sequences) and the microsatellites (8 newly developed loci) reveal spatial patterns of differentiation within the SWIO. There is, however, a discrepancy between the structure observed with each kind of marker. MtDNA revealed that 3 peripheral populations (NW Kenya, SE Reunion, and SW Europa) were isolated from the 7 more central populations, which form a more densely connected population network, while microsatellite data indicated a more restricted connectivity with significant differentiation between most pairs of localities. Higher genetic differences between Reunion and Europa were found, which might be explained by geography and isolation by distance pattern. In contrast, the genetic signature of Kenya—the most divergent locality identified by mtDNA basis but not with microsatellite—was probably the consequence of a particular colonisation history. These results indicate a much more restricted connectivity than previously thought for this species.

**KEY WORDS:** mtDNA · Microsatellite · Marine fish · SW Indian Ocean · Connectivity

—Resale or republication not permitted without written consent of the publisher—

## INTRODUCTION

A total of 22 species of the holocentrid fish genus *Myripristis* are recognized from the Indo-Pacific region (Randall & Greenfield 1996). *M. berndti* is the widest ranging of the 9 *Myripristis* species with a relatively wide distribution, occurring from the western Indian Ocean to the eastern Pacific (Randall & Greenfield 1996). Population genetic structuring of reef fish species with such a large geographical distribution could be expected as a consequence of many factors, including the patchy distribution of suitable habitat. In the Indian Ocean, reefs are often

separated by hundreds of km of open ocean or deep water, with 2 notably large expanses (>1200 km), one between the Seychelles and Chagos archipelagos and the other between Sri Lanka and Indonesia. Molecular tools provide an efficient approach to confirm within-species genetic homogeneity or, alternatively, to determine boundaries to effective dispersal within species. Recent mitochondrial DNA surveys of reef fish distributed throughout the Indo-Pacific region showed an equivocal pattern: many of the studied species showed no genetic structure at the pantropical scale (Horne et al. 2008, Gaither et al. 2010, Reece et al. 2010), while others showed the influence of the

\*Email: muthsdelphine@yahoo.fr

Indo-Pacific Barrier on the within-species genetic signature (Bernardi et al. 2002, Bay et al. 2004). In *M. berndti*, Craig et al. (2007) showed significant genetic structure between the Indian and Pacific Oceans, but no genetic structure was found at the scale of the central Pacific Ocean itself (even between populations >1000 km apart). As proposed by the authors, the pelagic larval phase of about 55 d (B. Victor unpubl. data), as well as a potential post-larval pelagic phase ('meeki' stage), is long enough to ensure homogeneity at this scale.

The present study focuses on this species in the SW Indian Ocean (SWIO)—defined here as the waters bounded by the eastern coast of Africa between Kenya and South Africa and as far east as 65° E longitude—and aims to determine if the high genetic connectivity encountered between Pacific populations also occurs in the smaller SWIO region. According to the partitioning of the world's oceans into provinces as defined by Longhurst (1998) on the basis of global hydrodynamics and water color data, the SWIO is divided into 4 main hydrogeographic areas (Fig. 1): the eastern Africa Coastal Province (EAFR), the NW Arabian Sea Upwelling Province (ARAB), the Indian Monsoon Gyre Province (MONS), and the Indian South Subtropical Gyre Province (ISSG). The last 2 are large oceanographical provinces, separated around 12° S, with the ISSG to the south and the MONS to the north of this parallel. The other 2 areas are mainly coastal areas, with the EAFR including the Mozambique Channel, East Africa coastline until 6° S and the entire Madagascan coastline. To the north of the 6° S longitude, the ARAB includes the adjacent coastlines of the Arabian Sea. Furthermore, these oceanographic provinces are in agreement with both tectonic plate boundaries and proposed biogeographic regions of the SWIO (Santini & Winterbottom 2002). The present study has such been conducted in an area that straddles these 4 regions (Fig. 1).

Several authors have emphasised the effects of the complex current system in this area (Schott et al. 2009), particularly in the Mozambique Channel, as mechanisms for the disruption of connectivity (Schouten et al. 2003, Swart et al. 2010). The westward flowing South Equatorial Current (SEC) splits at the east coast of Madagascar, near 17° S. The northward branch of the SEC splits at the northern tip of Madagascar (near 10° S) into the northward East African Coastal Current and a southward current that supplies a succession of eddies in the Mozambique Channel (consult Schott et al. (2009) for a diagrammatical presentation of this system). This hydrodynamic system occurring in the Mozambique

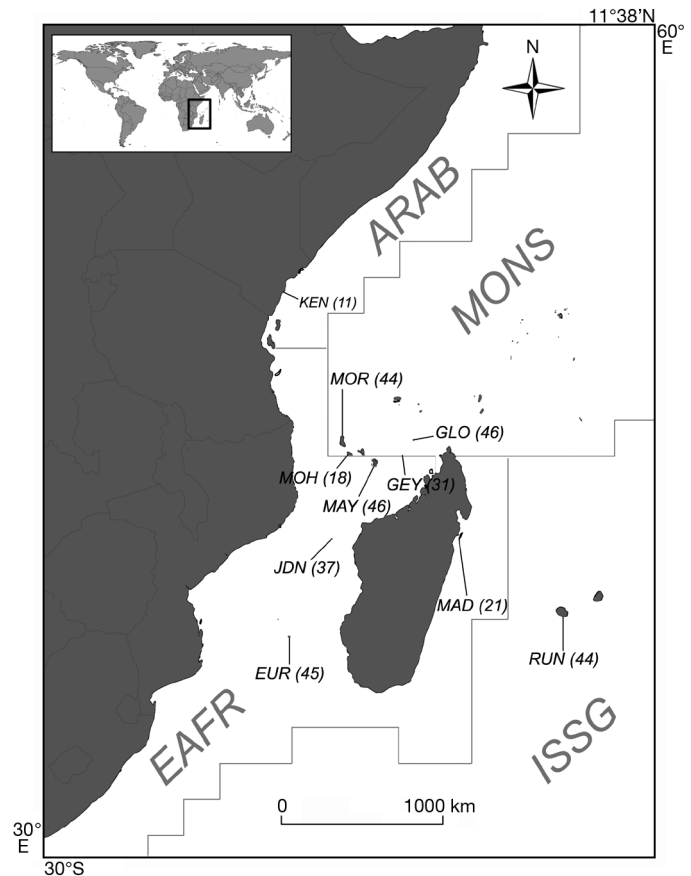


Fig. 1. *Myripristis berndti*. The 10 localities in the southwest Indian Ocean (SWIO) sampled in the present study, with sample sizes in brackets. Grey lines = biogeographical province boundaries (Longhurst 1998): Eastern Africa Coastal Province (EAFR), Indian Monsoon Gyre Province (MONS), Indian South Subtropical Gyre Province (ISSG), and Northwest Arabian Sea Upwelling Province (ARAB). See Table 1 for population abbreviations

Channel could greatly influence connectivity between populations, as was reported for the coral *Pocillopora verrucosa* (Ridgway et al. 2001) or the parrot fish *Scarus ghobban* (Visram et al. 2010). But, as recently emphasized (Ridgway & Sampayo 2005, Gaither et al. 2010), very few genetic studies are available in the SWIO, and it is impossible to draw general conclusions about patterns of connectivity of marine life in the SWIO.

Such knowledge on genetic connectivity is, however, crucial in terms of biodiversity protection, as it might significantly contribute to the design of marine protected areas (MPA) networks (Palumbi 2003). Even when MPAs are demographically self-sustaining, they need to be linked to other protected areas to promote recovery from local extinctions. In the SWIO, existing marine reserves appear largely ineffective and remain insufficient for the protection of

marine reef diversity (Mora et al. 2006). Examining the genetic structure of reef species in this area should allow an estimation of the level of larval exchange between populations or the identification of barriers to gene flow, and thus could help for the design of the MPAs network in the SWIO.

Here, we analysed mitochondrial cytochrome *b* sequences and 8 newly developed microsatellite loci from 343 individuals of the reef fish *Myripristis berndti*. The aim was to assess the level of connectivity of this reef fish between the 10 sites sampled in the SWIO. We also compared the mtDNA dataset with that obtained by Craig et al. (2007) from the same species in the Pacific Ocean.

## MATERIALS AND METHODS

### Sampling sites

Samples of *Myripristis berndti* were obtained by spearfishing or bought from fishermen at landing sites. From March 2007 to November 2010, we sampled 10 sites: Reunion Island, eastern coast of Madagascar (Sainte Marie Island), 7 islands in the Mozambique Channel (from south to north: Europa, Juan de Nova, 3 sites within the Comoros archipelago [Moroni, Mohéli, and Mayotte], Glorieuses, and the Geysers Bank), and 1 site on the Kenyan coast (Shimoni) (see Appendix 1 for dates and GPS coordinates). Locality codes and sample sizes are detailed in Table 1 and reported in Fig. 1. According to Longhurst (1998), these sites belong to 4 biogeographic areas: EAFR (Europa, Juan de Nova, Sainte

Marie Island), ISSG (Reunion), ARAB (Kenya), and MONS (Glorieuses), respectively. The 3 Comoros localities and those from Geysers are at the boundary between MONS and EAFR (see Fig. 1). Sampling at 3 other sites (Rodrigues, Mauritius, and Tanzania) was unsuccessful. Tissues were fixed in ethanol 95% and stored at  $-20^{\circ}\text{C}$ .

### Microsatellite development and genotyping

Total genomic DNA was isolated from individuals using Qiagen DNeasy spin columns and sent to GenoScreen, France ([www.genoscreen.fr](http://www.genoscreen.fr)). A total of 1  $\mu\text{g}$  was used for the development of microsatellite libraries through 454 GS-FLX Titanium pyrosequencing of enriched DNA libraries, as described in Malausa et al. (2011). Briefly, total DNA was enriched for AG, AC, AAC, AAG, AGG, ACG, ACAT, and ATCT repeat motifs and subsequently amplified. PCR products were purified, quantified, and GsFLX libraries were then constructed following manufacturer's protocols (Roche Diagnostics) and sequenced on a GsFLX-PTP. This technique identified 5306 potential markers. The bioinformatics program QDD (Megléczy et al. 2010) was used to filter for redundancy, resulting in a final set of sequences from which it was able to design primers. Tetra- and penta-repeats were favoured in order to reduce the presence of stutter bands and to increase the chance of accurate scoring. Finally, 70 primer sets were designed and tested for amplification. Primer sets were discarded if they failed to amplify, led to multiple fragments, or were revealed to be monomorphic. At the end, 8 microsatellite loci (see Appendix 1) were selected and used for population analysis.

PCR amplifications were performed in 20  $\mu\text{l}$  reactions containing 25 ng of template DNA, 1 $\times$  reaction buffer, 1.5 mM  $\text{MgCl}_2$ , 0.1 mM dNTP, 0.1  $\mu\text{M}$  fluorescent primer, 0.1  $\mu\text{M}$  primer, and 1U *Taq* polymerase. The PCR cycling consisted of an initial denaturation at  $95^{\circ}\text{C}$  for 5 min, followed by 35 cycles: denaturation at  $95^{\circ}\text{C}$  for 30 s, annealing at the appropriate temperature (see Appendix 1) for 30 s, and extension at  $72^{\circ}\text{C}$  for 1 min and a final extension at  $72^{\circ}\text{C}$  for 20 min. Each microsatellite amplification was diluted with  $\text{dH}_2\text{O}$  (1:20), mixed with Formamide and LIZ 500 size standard (Applied Biosystems). Fragments were separated using an Applied Biosystems 3100 DNA Analyzer. Alleles were scored using GeneMapper® ver. 4.0 (Applied Biosystems).

Table 1. *Myripristis berndti*. Microsatellite diversity indices for the 10 localities sampled in the southwest Indian Ocean (SWIO). n: number of samples analyzed;  $N_{\text{all}}$ : number of alleles per population;  $R_s$ : allelic richness;  $H_o$ : observed heterozygosity;  $F_{\text{IS}}$ : fixation index with associated significance, \* $p < 0.05$ ); and the % of 1 of the 2 clades identified with Structure analysis (called Clade S) (see Fig. 2)

Population	Abbr.	n	$N_{\text{all}}$	$R_s$	$H_o$	$F_{\text{IS}}$	% Clade S
Europa	EUR	45	4.75	3.43	0.416	0.177*	0.98 $\pm$ 0.01
Juan de Nova	JDN	37	5.25	3.60	0.410	0.219*	0.97 $\pm$ 0.03
Mayotte	MAY	46	4.75	3.73	0.409	0.191*	0.82 $\pm$ 0.13
Mohéli	MOH	18	5.12	3.68	0.306	0.359*	0.42 $\pm$ 0.39
Moroni	MOR	44	5.50	3.68	0.321	0.333*	0.29 $\pm$ 0.23
Geysers	GEY	31	4.50	3.82	0.400	0.301*	0.83 $\pm$ 0.11
Glorieuses	GLO	46	5.37	3.69	0.437	0.200*	0.88 $\pm$ 0.11
Kenya	KEN	11	3.62	3.62	0.441	0.155*	0.39 $\pm$ 0.17
Madagascar	MAD	21	4.00	3.30	0.353	0.280*	0.12 $\pm$ 0.16
Reunion	RUN	44	5.00	3.60	0.388	0.151*	0.39 $\pm$ 0.21
Totals		343	4.79	3.91	0.388	0.236*	0.61 $\pm$ 0.31

### Mitochondrial DNA sequencing

A 771 base pair (bp) fragment of the mitochondrial cytochrome *b* region was amplified by PCR using the primers CytbH 5'-GTG ACT TGA AAA ACC ACC GTT G-3' (Song et al. 1998) and CB13R 5'-TAT TCC GCC GAT TCA GGT AA-3' (Marko et al. 2004). Reactions were performed in 20  $\mu$ l volumes containing 1 $\times$  PCR buffer, 2 mM MgCl<sub>2</sub>, 20  $\mu$ M of each dNTP, 0.5  $\mu$ M of each primer, 0.5 U of Silverstar DNA *Taq* polymerase (Eurogentec), and 25 ng of genomic DNA. Cycling parameters were 94°C for 3 min; followed by 35 cycles of 94°C for 45 s, 54°C for 45 s, and 72°C for 45 s; and a final elongation at 72°C for 4 min. PCR products were purified and sequenced on an ABI 3100 sequencer (Macrogen). Samples were sequenced in forward and reverse directions. They were checked in BioEdit Sequence Alignment Editor (Hall 1999) and aligned using ClustalW (Thompson et al. 1994). Cytochrome *b* sequences of this species were available (GenBank accession number DQ868393-DQ969429; Craig et al. 2007) but were shorter; hence, all the present haplotypes were submitted to GenBank (accession number HQ687596-HQ687665 and HQ896824-HQ896828).

### Microsatellite analyses

Allele frequencies, the mean number of alleles ( $N_{\text{all}}$ ), observed ( $H_o$ ) heterozygosity (Nei 1987), deviations from Hardy-Weinberg equilibrium estimated by Wright's fixation index  $F_{\text{IS}}$  (Weir & Cockerham 1984) and exact tests were estimated per population with Arlequin 3.5 (Excoffier & Lischer 2010). Linkage disequilibrium was also investigated with this software. Micro-Checker ver. 2.2.3 (Van Oosterhout et al. 2004) was used to detect possible null alleles. Allelic richness ( $R_s$ ), which corrects the number of alleles according to the smallest sample size, was estimated using Fstat ver. 2.9.3.2 (Goudet 2001). Genetic differentiation between populations was estimated by calculating Wright's  $F_{\text{ST}}$  statistic (Weir & Cockerham 1984) with Arlequin ver. 3.5 (Excoffier & Lischer 2010). Significance levels for multiple tests were corrected in agreement with Narum (2006), using a sequential procedure from Benjamini-Yekutieli (2005). The unbiased estimator of divergence Jost's  $D$  (Jost 2008) was also calculated per pair of populations using Software for the Measurement Of Genetic Diversity (SMOGD) (Crawford 2010). Arlequin ver. 3.5 (Excoffier & Lischer 2010) was also used to perform AMOVA (analysis of molecular variance);

various locality groupings were made according to geographic proximity or biogeographic provinces. To test for patterns of isolation-by-distance, marine distances between localities were estimated with [www.geodistance.com](http://www.geodistance.com) and plotted against  $F_{\text{ST}}/(1 - F_{\text{ST}})$ , following the recommendations of Rousset & Raymond (1997). The significance of this relationship was tested by a Mantel test, performed in R (R Development Core Team 2010) using the spatial nonparametric covariance functions (nfc) package. A principal component analysis (PCA) on allelic frequencies was performed with PCA-gene 1.2 (Goudet 1995).

To determine if the samples belonged to 1 or more populations, data were also analysed using Structure ver. 2.3.2 (Pritchard et al. 2000), which uses an iterative computation process to infer the most likely number of populations ( $K$ ) represented in the total sample. For this analysis, an admixture model assuming independent allele frequencies was used and 20 replicates were run (each with a  $1.10^5$  generation burn-ins and  $5.10^5$  iterations) at  $K$  values from 1 to 10. The optimal value of  $K$  was selected based on Evanno et al. (2005) by choosing the number of clusters where the largest difference in log likelihoods ( $\Delta K$ ) was observed.

### Mitochondrial analyses

For each population, haplotype ( $h$ ) and nucleotide ( $\pi$ ) diversities as well as Tajima's  $D$  (1989) and Fu's  $F_s$  (Fu 1997) statistics were estimated using DNA Sequence Polymorphism ver. 5.0 (DNA SP 5.0) (Librado & Rozas 2009). These statistics test for departures from equilibrium between mutation and drift. A mismatch curve was examined with the same software package. Pairwise estimates of genetic differentiation ( $\Phi_{\text{st}}$ ) among localities were calculated with Arlequin ver. 3.5 (Excoffier & Lischer 2010). The unbiased estimator of divergence Jost's  $D$  (Jost 2008) was also calculated per pair of populations using SPADE (Species Prediction And Diversity Estimation; available at <http://chao.stat.nthu.edu.tw/softwareCE.html>). AMOVAs and tests for patterns of isolation-by-distance were conducted as for the microsatellite data. DnaSP ver. 5.0 (Librado & Rozas 2009) was also used to estimate the nearest-neighbour statistic,  $S_{\text{nn}}$  (Hudson 2000).  $S_{\text{nn}}$  is a measure of how often the 'nearest neighbours' (in sequence space) are from the same locality (in geographical space).  $S_{\text{nn}}$  varies from 0 to 1: under 0.5, it is assumed that populations are in panmixia, and values closer to 1 indicate that populations are differentiated. Median-joining networks

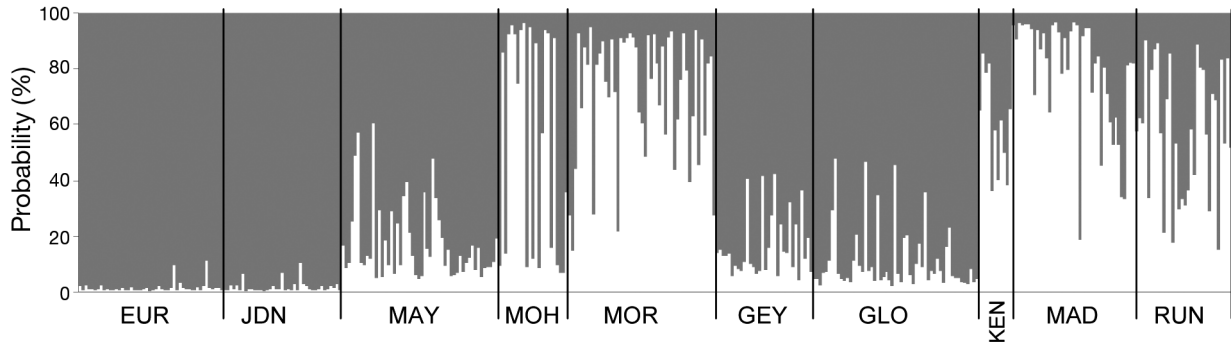


Fig. 2. *Myripristis berndti*. Clustering analysis made with Structure ver. 2.2 (Pritchard et al. 2000) based on the microsatellite data from individuals collected in the southwest Indian Ocean (See Table 1 for population abbreviations.). Posterior assignment probability is plotted for each individual represented by a vertical line, the 2 shades symbolizing the respective proportion of membership to each cluster ( $k = 2$ )

(Bandelt et al. 1999) were constructed using Network ver. 4.5 (available at [www.fluxus-technology.com/](http://www.fluxus-technology.com/)).

For comparison of the Indian and Pacific Oceans, we used the shorter sequences of Craig et al. (2007); only 164 sequences were taken into consideration (Table S1 in the supplement at [www.int-res.com/articles/suppl/m443p167\\_supp.pdf](http://www.int-res.com/articles/suppl/m443p167_supp.pdf); from all localities except for the Hawaii archipelago, where only the 2 largest samples of Midway and the Big Island of Hawaii were considered). We estimated parameters of demographic expansion:  $\tau$ , the time to the expansion and the mutation parameters  $\theta_0$  and  $\theta_1$  with Arlequin ver. 3.5 (Excoffier & Lischer 2010). The coalescent time can be deduced using the equation  $\tau = 2vt$ , where  $t$  is the age in generations and  $v$  is the mutation rate per generation. As done by Craig et al. (2007), an approximation of the population age could be estimated for *Myripristis berndti* with a generation time of 3 yr and a mutation rate of 2% per  $10^6$  yr.

## RESULTS

### Microsatellite analyses at the SWIO scale

The 8 loci screened for the 343 individuals were polymorphic, with a total of 75 alleles detected for the 10 localities. Allelic richness was of the same magnitude among the 10 localities, with a mean of  $3.91 \pm 0.15$  alleles per locus (Table 1). Four cases of significant linkage disequilibrium were observed, each for a single locality ( $p < 0.05$ ). The mean observed heterozygosity across populations was  $0.38 \pm 0.04$  (Table 1). Heterozygote deficiencies were highly significant in all localities ( $p < 0.001$ ) with  $F_{IS}$  values ranging from 0.151 for Reunion to 0.359 for Moheli (Table 1). However, heterozygote deficiency was locus-dependant, as only locus M-22 failed to meet Hardy-Weinberg

equilibrium for all populations. Indeed, this locus was characterized by presence of null alleles ( $p < 0.05$ ).

The clustering analysis conducted with Structure suggests that the highest likelihood of obtaining such data is to assume that 2 genetic groups exist (highest  $\Delta K$  for  $K = 2$ ). The posterior assignment probability plotted against locality of origin (Fig. 2; see also Table 1 for the proportion of individuals assigned to the first clade in each population) provided evidence of geographical subdivision. Thus the samples of Europa and Juan de Nova belonged nearly exclusively to the first of the 2 clusters ( $\geq 97\%$ ). Although a smaller proportion ( $\geq 80\%$ ), samples from Geyser, Glorieuses, and Mayotte were largely assigned to the first cluster, too. In contrast, Madagascan samples were assigned (88%) to the second cluster. The 4 remaining localities were also characterised by a large proportion assigned to the second cluster but show an important within-locality variability (Fig. 2, Table 1).

The principal component analysis (PCA; Fig. 3) performed with allelic frequencies confirmed this

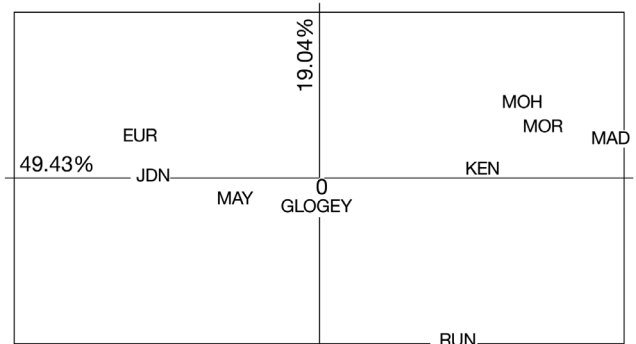


Fig. 3. *Myripristis berndti*. Principal component analysis plot based on allelic frequencies of microsatellite data of the 10 samples collected in the southwest Indian Ocean. PC1 and PC2 axis explain 49.43 and 19.04% of the variance, respectively. See Table 1 for abbreviations



Table 2. *Myripristis berndti*. Pairwise  $F_{ST}$  values of differentiation among the 10 populations (below diagonal) and Jost's  $D$  (above diagonal); significant  $F_{ST}$  values are indicated by \* $p < 0.05$ , and values still significant after Benjamini-Yekutieli correction for 45 comparisons are indicated in bold; \*\*\* $p < 0.011$ ). See Table 1 for population abbreviations

	EUR	JDN	MAY	MOH	MOR	GEY	GLO	KEN	MAD	RUN
EUR	–	0.004	0.003	0.003	0.006	0.013	0.013	0.024	0.018	0.025
JDN	0.013*	–	0.004	0.010	0.010	0.030	0.027	0.039	0.042	0.022
MAY	<b>0.021***</b>	<b>0.030***</b>	–	0.000	0.010	0.015	0.008	0.009	0.025	0.011
MOH	<b>0.076***</b>	<b>0.077***</b>	<b>0.049***</b>	–	0.006	0.010	0.009	0.016	0.018	0.018
MOR	<b>0.084***</b>	<b>0.077***</b>	<b>0.058***</b>	-0.000	–	0.010	0.016	0.018	0.015	0.016
GEY	<b>0.028***</b>	0.015	<b>0.022***</b>	<b>0.033***</b>	<b>0.029***</b>	–	0.000	0.005	0.005	0.026
GLO	<b>0.023***</b>	<b>0.022***</b>	<b>0.017***</b>	<b>0.035***</b>	<b>0.035***</b>	0.002	–	0.011	0.006	0.025
KEN	<b>0.074***</b>	<b>0.084***</b>	<b>0.044***</b>	0.015	0.023	0.032*	0.032*	–	0.018	0.012
MAD	<b>0.118***</b>	<b>0.119***</b>	<b>0.081***</b>	0.017	0.015*	<b>0.062***</b>	<b>0.065***</b>	0.043*	–	0.020
RUN	<b>0.092***</b>	<b>0.076***</b>	<b>0.051***</b>	<b>0.060***</b>	<b>0.050***</b>	<b>0.034***</b>	<b>0.037***</b>	<b>0.047***</b>	<b>0.059***</b>	–

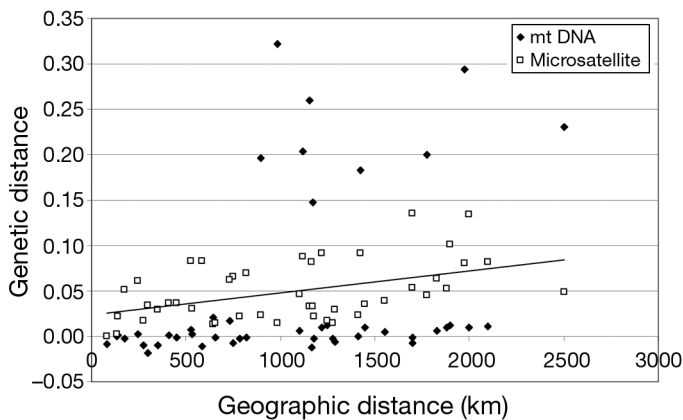


Fig. 4. *Myripristis berndti*. Isolation by distance graph showing corrected pairwise genetic distances,  $\Phi_{st} / (1 - \Phi_{st})$  for mitochondrial DNA (mtDNA) and  $F_{ST} / (1 - F_{ST})$  for microsatellite, plotted as a function of geographic distances. Line designates significant relationship between microsatellite and geographic distances (Mantel tests;  $p < 0.01$  for microsatellite;  $p = 0.11$  for mtDNA)

geographic structure, with the first axis extracting nearly 50% of the total variance by opposing Europa and Juan de Nova versus Madagascar. The second axis (extracting 19% of the total variance) differentiated the Reunion population from the others.

Pairwise multilocus  $F_{ST}$  values of microsatellite differentiation were very low (Table 2). However, of the 45 comparisons, 39 were significant ( $0.013 < F_{ST} < 0.119$ ;  $p < 0.05$ , even after Benjamini-Yekutieli correction). Values of Jost's  $D$  were also very low (Table 2). The highest values included some of the most distant populations (Madagascar versus Juan de Nova, Madagascar versus Europa, Reunion versus Europa). The smallest and non-significant values concerned the closest localities (Geysers/Glorieuses or

Mohéli/Moroni). Consequently, a significant relation exists between genetic and geographic distances ( $r = 0.46$ ,  $p < 0.01$ ; Fig. 4). In the same way, AMOVAs (Table 3) showed that some microsatellite variance could be significantly associated between geographic groups when localities are divided in 2 groups based on Structure results ( $\Phi_{sc} = 0.035$ ,  $p < 0.01$ ), but not when the grouping is done in agreement with Longhurst's (1998) biogeographic provinces ( $\Phi_{sc} = 0.010$ ,  $p = 0.15$ ).

#### Mitochondrial DNA analyses at the SWIO scale

The 771-bp sequences of the mitochondrial cytochrome *b* fragment revealed a high level of nucleotide and haplotype polymorphism in the samples of *Myripristis berndti* from the 10 localities. Sequence analysis indicated the occurrence of 61 polymorphic sites, from which 75 different haplotypes were identified.

Levels of genetic diversity within each population are provided in Table 4 (detailed haplotype frequencies are available in Table S2 in the supplement at [www.int-res.com/articles/suppl/m443p167\\_supp.pdf](http://www.int-res.com/articles/suppl/m443p167_supp.pdf)). The haplotype and nucleotide diversities were of the same magnitude for each population, with an average of  $0.88 (\pm 0.06)$  and  $0.0036 (\pm 0.0006)$ , respectively. The lowest diversity values were for the East Madagascar samples and the highest for Kenya. The reverse trend was observed for the proportion of private haplotypes (Table 4), which ranged from 0% in Kenya to 23.8% at Madagascar, with a mean of  $13.8\% (\pm 7.5)$ . The associated Tajima's  $D$  and Fu's  $F_s$  values for the sample overall were negative and significant ( $D = -2.08$ ,  $p < 0.05$  and  $F = -92.47$ ,  $p <$

Table 3. *Myripristis berndti*. Molecular variance (AMOVA) based on microsatellite and mtDNA data. The different groups are defined on the left of the table and based on (a) Longhurst's (1998) biogeographic provinces, (b)  $\Phi_{st}$  values, (c) Structure results. See Table 1 for site abbreviations. Pop. = population, % var. = % variation, F. ind. = Fixation index

Site	Groups (a) (b) (c)			Source of variation among	Microsatellite				mtDNA				
					df	% var	F. ind.	p	df	% var	F. ind.	p	
EUR	1	1	1										
JDN	1	2	1	Groups	3	1.1	0.010	0.15	3	-0.4	0.004	0.25	
GEY	2	2	1	Populations within groups	6	3.8	0.038	<0.001	6	3.0	0.025	<0.001	
GLO	2	2	1	Individuals within pop.	754	95.1	0.049	<0.001	333	97.4	0.029	<0.001	
MAY	1	2	1	Groups	3	1.1	0.011	0.17	3	4.2	0.042	0.21	
MOH	2	2	2	Populations within groups	6	4.0	0.040	<0.001	6	0.4	0.004	0.32	
MOR	2	2	2	Individuals within pop.	754	94.9	0.051	<0.001	333	95.4	0.046	<0.001	
KEN	3	3	2	Groups	1	3.6	0.035	0.003	1	-0.6	-0.006	0.62	
MAD	1	2	2	Populations within groups	8	2.6	0.026	<0.001	8	3.0	0.030	<0.001	
RUN	4	4	2	Individuals within pop.	754	93.8	0.061	<0.001	333	97.6	0.023	<0.001	

Table 4. *Myripristis berndti*. Cytochrome *b* diversity indices for the 10 populations sampled in the southwest Indian Ocean. See Table 1 for population abbreviations. (n: number of samples;  $N_{hap}$ : number of haplotypes per population; for Tajima's *D* and Fu's  $F_S$  values, significance: \*p < 0.05, \*\*p < 0.01, \*\*\*p < 0.001)

Population	n	$N_{hap}$	% of private haplotypes	Haplotype diversity ( <i>h</i> )	Nucleotide diversity ( $\pi$ )	Tajima's <i>D</i>	Fu's $F_S$
EUR	45	20	15.6	0.817	0.0031	-2.00**	-13.57***
JDN	37	18	18.9	0.907	0.0036	-1.81*	-10.07***
MAY	46	25	21.7	0.886	0.0036	-1.68*	-20.51***
MOH	18	11	16.7	0.817	0.0030	-2.26***	-5.88***
MOR	44	18	6.8	0.861	0.0039	-1.51	-8.11***
GEY	31	16	16.1	0.869	0.0042	-1.4	-7.13***
GLO	46	20	6.5	0.892	0.0032	-1.72*	-12.87***
KEN	11	5	0.0	0.933	0.0045	0.05	-0.31
MAD	21	13	23.8	0.709	0.0027	-0.81	-5.45***
RUN	44	21	11.4	0.864	0.0028	-2.02**	-17.01***
Total	343	75	13.8	0.876	0.0036	-2.08*	-92.47***

0.001); these negative values indicate a significant excess of low frequency variants. The Fu's  $F_S$  values were negative and significant for all populations, except Kenya. A mismatch distribution constructed using the 343 sequences shows 2 apparent peaks (Fig. 5). However, Harpending's raggedness index was low and not significant ( $r = 0.028$ ,  $p = 1.00$ ) indicating no departure between observed and expected unimodal distributions. All these elements indicate a recent population expansion for *Myripristis berndti* in the SWIO.

Pairwise values of genetic differentiation ( $\Phi_{st}$ ) are provided in Table 5. Of the 45 comparisons, 14 were significant ( $0.009 < \Phi_{st} < 0.20$ ;  $p < 0.05$ ). Of these significant values, 8 included the Kenyan samples (versus all the other samples, except Glorieuses) with a mean  $\Phi_{st} = 0.16$ . The 6 remaining significant differentiation values were 1 order of magnitude smaller than those induced by the Kenyan samples and

involve the Reunion and Europa samples. After a Benjamini-Yekutieli correction ( $p$ -limit value is thus 0.011), 6 comparisons still remained significantly different: 5 involved the Kenyan samples and the comparison among Europa and Juan de Nova ( $\Phi_{st} = 0.02$ ). Consistent with these results, values of Jost's *D* (Table 5) were very high between Kenya and other localities ( $0.58 < D < 0.83$ ). The differentiation between closed localities of Reunion and Madagascar on one side and of Europa and Juan de Nova on the other side was also confirmed. No significant relationship existed between mitochondrial and geographic distances ( $r = 0.35$ ,  $p = 0.12$ ). The nearest-neighbour statistic ( $S_{nn} = 0.13$ ) confirmed the non-significant association between sequence similarity and geographical location. The groupings examined in the AMOVAs (Table 3) failed to identify a hierarchical level where some mitochondrial variance could be significantly associated among groups, even in

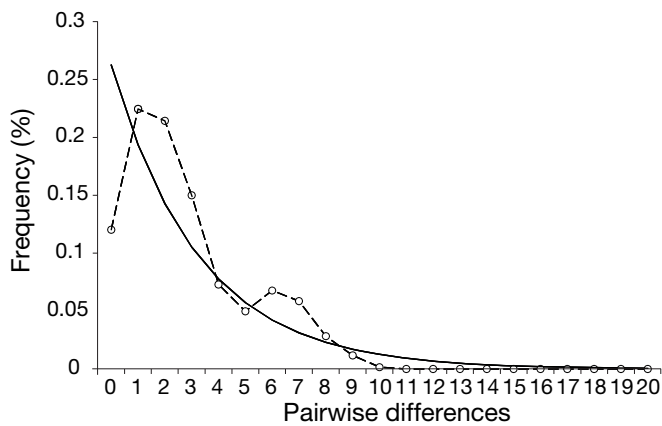


Fig. 5. *Myripristis berndti*. Mismatch distribution curves obtained from individuals collected in the southwest Indian Ocean. Solid line: expected curve under a constant population size. Dashed line: observed curve. Harpending's raggedness index  $r = 0.028$ ,  $p = 1.00$

the extreme case where the 3 most divergent populations (Kenya, Reunion, and Europa) are separated from the other localities.

A haplotype network constructed with the SWIO sequences is presented in Fig. 6a. The network revealed a 'star-like' pattern with the most common haplotypes at the centre, from which a crown of less common haplotypes radiated. However, one of the divergent branches of the crown revealed a more complex structure (composed of 16 interconnected

haplotypes, with the 2 most divergent being 7 mutations steps from the most common one). The most common haplotype (Hap\_1-HQ687596; Table S2), as well as most of the non-unique haplotypes, were well-represented across all 10 populations. This most common haplotype occurred in one-third of the SWIO individuals, with a frequency varying from 9% in Kenya to 44% in Mohéli (see Table S2 for detailed haplotype frequencies). The Kenyan sample, the only one without unique alleles, was mainly composed of a second shared haplotype (Hap\_19-HQ687614; see Table S2 and Fig. 6a).

### Mitochondrial DNA structure at the Indo-Pacific scale

Complementary analysis were conducted with the 624-bp in common among the 343 SWIO sequences and the 164 sequences obtained from 11 Pacific localities by Craig et al. (2007), for a total of 507 individuals (see Table S1). From this global dataset, 78 haplotypes were identified: 6 haplotypes shared between the SWIO and Pacific, 52 specific to the SWIO and 20 specific to the Pacific. The haplotype diversity was higher in SWIO ( $h = 0.84$ ) than in the Pacific Ocean ( $h = 0.61$ ). The parameter of demographic expansion ( $\tau$ ) and, as a consequence, the coalescent age, was higher in the Pacific Ocean than in the SWIO (see Table 6).

Table 5. *Myripristis berndti*. Pairwise  $\Phi_{st}$  values of differentiation among the 10 populations (below diagonal) and Jost's  $D$  (above diagonal); significant  $\Phi_{st}$  values are indicated as \* $p < 0.05$ , and values still significant after Benjamini-Yekutieli correction for 45 comparisons are indicated in **bold** \*\*\* $p < 0.011$ ). Populations labeled as indicated in Table 1

	EUR	JDN	MAY	MOH	MOR	GEY	GLO	KEN	MAD	RUN
EUR	–	0.125	–0.053	–0.016	0.047	–0.003	–0.036	0.778	0.050	0.063
JDN	<b>0.020***</b>	–	0.054	0.013	–0.082	–0.016	–0.011	0.731	–0.079	0.084
MAY	–0.003	0.002	–	–0.008	–0.045	0.013	–0.095	0.808	0.089	0.035
MOH	–0.012	0.007	–0.003	–	0.022	–0.080	0.005	0.746	0.049	–0.010
MOR	0.009*	–0.011	0.003	–0.009	–	–0.009	–0.066	0.663	–0.034	0.064
GEY	–0.006	–0.001	0.001	–0.019	–0.010	–	–0.002	0.584	–0.018	0.038
GLO	–0.000	–0.002	–0.010	0.001	–0.001	–0.000	–	0.833	–0.071	0.063
KEN	<b>0.185***</b>	<b>0.134***</b>	<b>0.145***</b>	<b>0.196***</b>	<b>0.141***</b>	<b>0.171***</b>	0.114	–	0.831	0.758
MAD	0.009	–0.007	0.006	0.012	–0.003	–0.007	–0.001	<b>0.143***</b>	–	0.136
RUN	0.012*	0.011*	–0.001	0.006	0.010	0.010*	0.005	<b>0.158***</b>	0.016*	–

Fig. 6. *Myripristis berndti*. Haplotype networks representing the evolutionary relationships between mitochondrial haplotypes. (a) Relationship among samples of the southwest Indian Ocean (present study;  $n = 343$ , 771 bp). (b) Network compiled using the data of the present study and those of Craig et al. (2007) ( $n = 164$ , 624 bp), for a total of 507 sequences. Length of connecting lines is proportional to the number of mutational steps between haplotypes (the smallest segment on the figure being for 1 mutation step). The size of circles is proportional to the number of individuals observed for a given haplotype. Colors of pie charts represent geographical origin of the haplotype, see Table 1 for abbreviations. The 2 main haplotypes are named on both networks



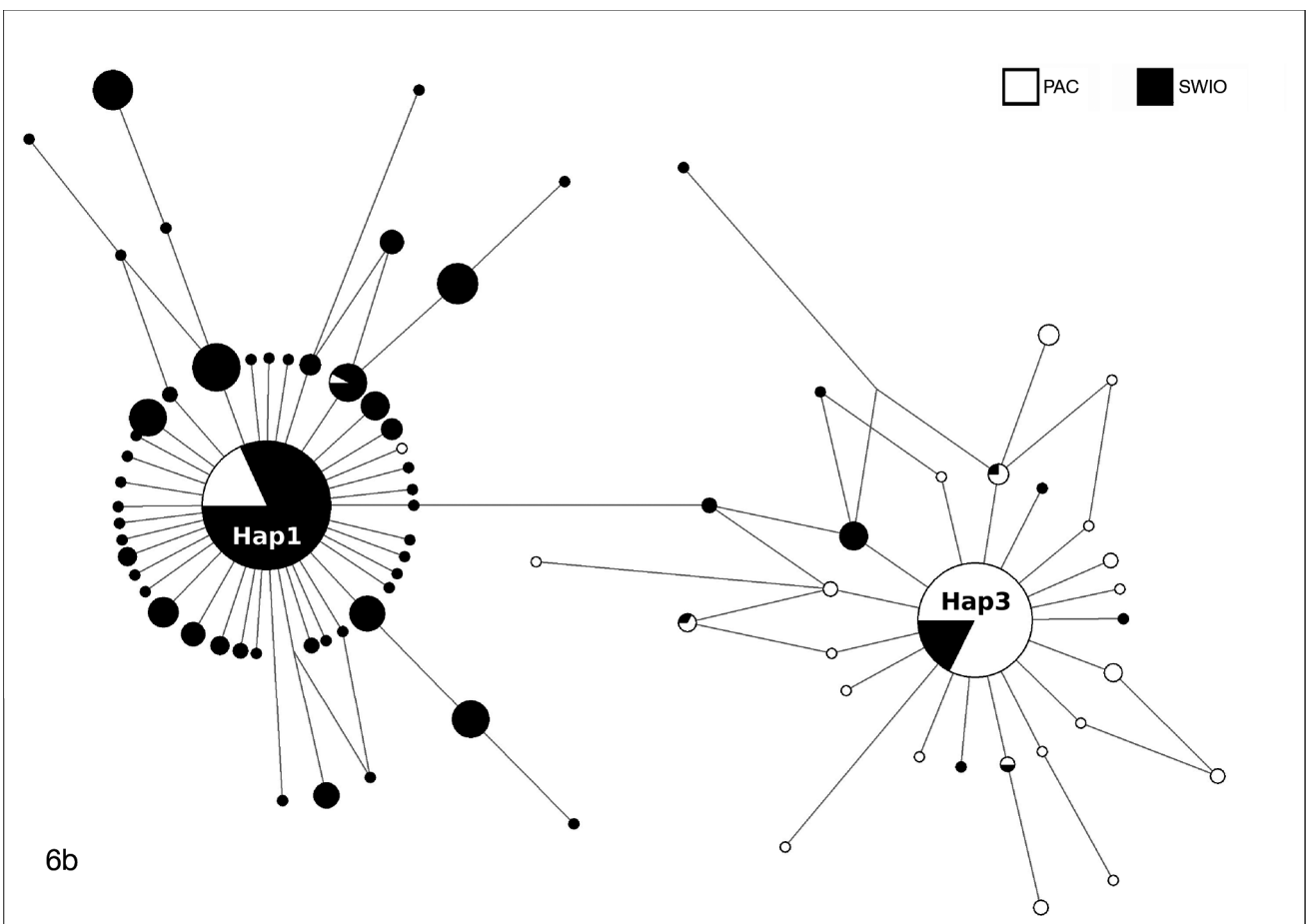
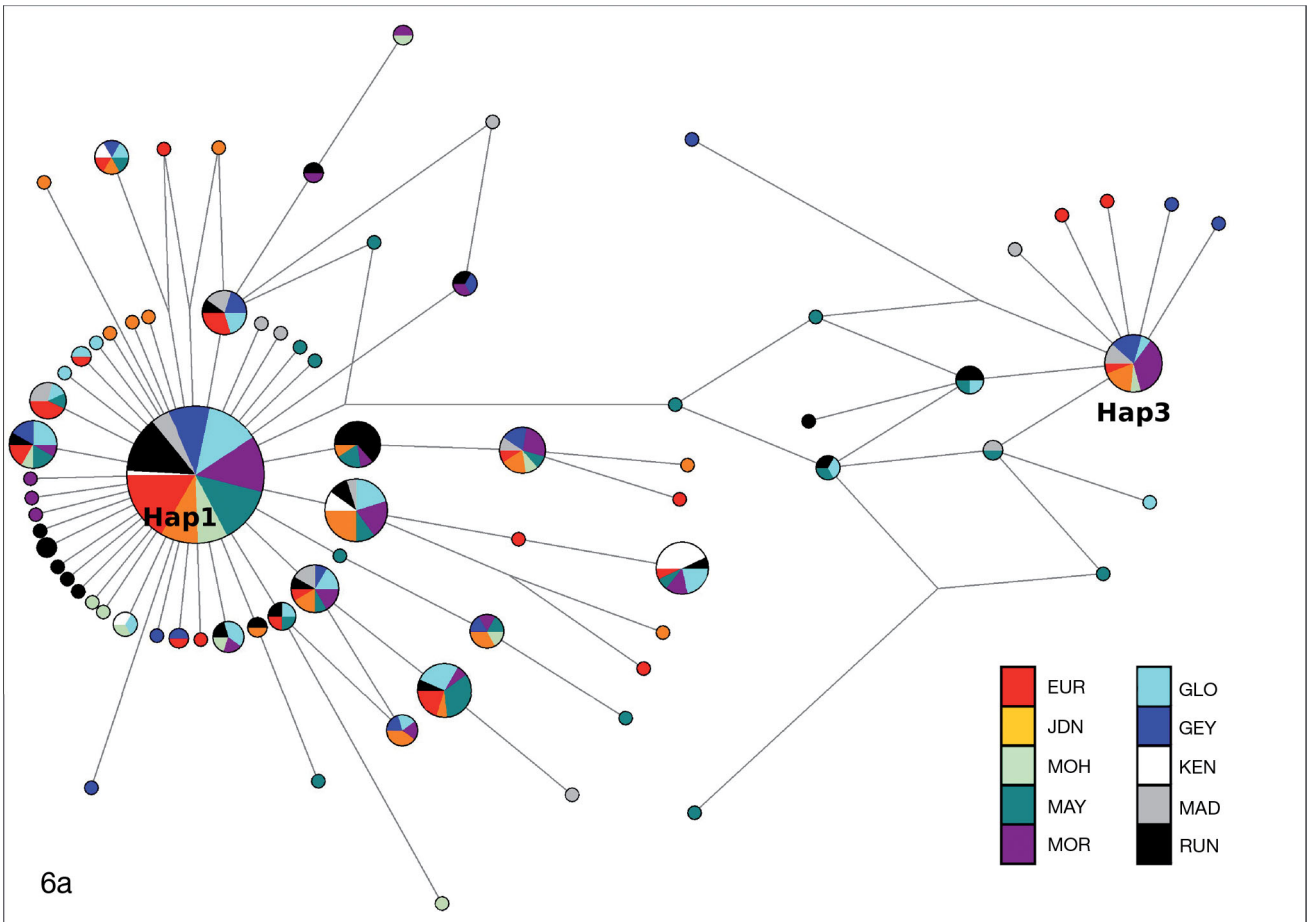


Table 6. *Myripristis berndti*. Cytochrome *b* diversity by oceans.  $\tau$ ,  $\theta_0$ , and  $\theta_1$  are the demographic parameters as reported by Arlequin ver. 3.5. The age of each population could be deduced from  $\tau$ . See Table 4 for further abbreviations

	n	$N_{hap}$	$h$	Tajima's $D$	Fu's $F$	$\tau$	$\theta_0$	$\theta_1$	Coalescence time (yr)
SWIO	343	58	0.838	-1.97***	-26.54***	0.768	0	$\infty$	57600
Pacific	164	25	0.626	-1.48*	-14.44***	5.801	1.25	1.25	435100

A second haplotype network was constructed with this dataset (Fig. 6b) and revealed a complex structure with 2 major haplotypes, diverging from one another by 5 mutational steps (divergence < 0.3%). The geographical distribution of these 2 major haplotypes and of related less-frequent haplotypes (see Fig. 6b) was clearly uneven with each genetic group being composed of >80% of samples from the same ocean. For example, the most common haplotype in the Pacific was absent from the Kenya, Reunion, and Mayotte samples. This strong geographic partition was confirmed by an AMOVA, with a significant variance associated among the oceanic groups ( $\Phi_{CT} = 0.18$ ;  $p < 0.001$ ). This value was even higher ( $\Phi_{CT} = 0.21$ ;  $p < 0.001$ ) when the 2 more eastern samples of the Pacific group (i.e. Panama and Clipperton samples) were excluded. These 2 populations were differentiated (as shown by  $\Phi_{st}$ ) from the other Pacific populations (see Craig et al. 2007 for further details). Despite being the most eastern population in the Pacific, their haplotypes frequencies were astonishingly more similar to those of the SWIO than to the other sites sampled in the Pacific Ocean (with the 2 major haplotypes in equal proportion).

## DISCUSSION

The present study aimed to evaluate genetic connectivity between *Myripristis berndti* sampled from 10 different localities of the SWIO. The mitochondrial study made on this species in the Hawaiian archipelago over a similar geographic scale (Craig et al. 2007) concluded that *M. berndti* is an extensive disperser. A study on another *Myripristis* species, *M. jacobus*, also showed no mitochondrial structure at the scale of the Atlantic Ocean (Bowen et al. 2006). The present analyses of genetic polymorphism showed that the connectivity of *M. berndti* in the SWIO is more restricted than could be expected on the basis of these 2 previous studies. Although the primary focus of the present study was the genetic connectivity within the SWIO for *M. berndti*, it also shed new light on the relationships between the Indian and Pacific Oceans.

## Signature of the Indo-Pacific barrier on the genetic structure of *Myripristis berndti*

MtDNA sequences confirmed the existence of significant genetic structure in the reef fish *Myripristis berndti* between the Indian and Pacific Oceans, with only 6 of the 79 haplotypes identified as being shared and, consequently, a high value of differentiation ( $\Phi_{CT} = 0.18$ ) among them. This oceanic differentiation, previously showed by Craig et al. (2007), is more evident now that the sampling size is consistent with >160 samples for both oceans. The role of the Indo-Pacific Barrier (i.e. the Indo-Australian archipelago) as a physical limitation to larval dispersal, enhanced during periods of lower sea levels, and its influence on genetic signature is widely recognized (Briggs 2005, Heads 2005). For example, during the Pleistocene, the sea level was up to 120 m lower than the current levels on at least 3 occasions (250 000, 150 000, and 17 000 yr before present). This caused nearly complete physical isolation of the 2 oceans, with Java, Sumatra, and Borneo connected by the exposed Sunda Shelf (Voris 2000). Such physical changes have led to the loss or at least a significant reduction of suitable habitat for marine species. This would have been accompanied by a decrease in population size that might have favoured the emergence of a different genetic lineage in each ocean (Hewitt 2000). However, the fact that the Panama and Clipperton samples are less differentiated from the SWIO populations than are the rest of samples from Pacific indicates that *M. berndti*'s phylogeographic history might be more complicated and that vicariance alone cannot explain the genetic pattern observed. Some parapatric divergence (Briggs 1999, Rocha & Bowen 2008) could be suspected. This could also be a consequence of an important drift effect in the central area or a signature of a once cosmopolitan species that is now extinct in the Atlantic basin. Further genome investigations and more intensive sampling in the eastern Pacific and the Indian Oceans are needed to (in)validate such hypotheses.

*Myripristis berndti* shows a similar pattern of recent demographic expansion in each ocean with significant negative Tajima's  $D$  and Fu's  $F_S$  values. These

expansions might be a consequence of the post-glacial sea level increase (Hewitt 2000). As the demographic parameter estimation is greater for the Pacific population, this should indicate a longest diversification time in the Pacific Ocean. This could be in agreement with results found in another widespread reef fish (Gaither et al. 2010). In contrast, the haplotype diversity is higher in the SWIO area ( $h = 0.84$  versus 0.62 in Pacific samples). Acquiring data on more Indo-Pacific species will allow investigators to deal on a potentially older diversification times in the SWIO.

### Restricted connectivity between localities of the SWIO

The analysis of mitochondrial sequences of *Myripristis berndti* revealed low and non-significant pairwise  $\Phi_{st}$  values between most of the localities. Together with a non-significant nearest-neighbour statistic and the lack of among-group mitochondrial variance as seen in the AMOVA, this provides evidence for high connectivity within the SWIO. However, this lack of mitochondrial structure is restricted to central localities as the 3 more peripheral ones (Kenya, Europa, and Reunion) showed significant levels of genetic structure. Moreover, all the microsatellite analyses show that connectivity between *M. berndti* populations is even lower than estimated using mitochondrial sequences. Indeed, the Structure analysis identifies 2 different genetic clusters; the  $F_{ST}$  values are low but significant between most localities, and some genetic variance could be associated among groups by AMOVA. Adding nuclear information has, thus, definitively confirmed the idea of restricted gene flow for *M. berndti* among localities of the SWIO.

Both genetic markers, thus, indicate restricted dispersal along the Mozambique Channel, mainly between Europa and the more northern localities, with Juan de Nova appearing intermediary, belonging either to the north or to the south depending of the marker of interest. Indeed, the Structure analysis indicates that Europa and Juan de Nova show the same microsatellite signature. Consequently, a low  $F_{ST}$  value is seen between these 2 localities, whereas the mitochondrial  $\Phi_{st}$  value between these localities is one of the largest and thus could indicate that Juan de Nova is more connected to the northern localities. This pattern is in agreement with what was previously observed in the green turtle *Chelonia mydas* (Bourjéa et al. 2007), for which genetic structure suggests that turtles nesting at the rookeries of the South

and those from the North Mozambique Channel belong to separate genetic stocks and that the Juan de Nova population comprises a mix of these 2 stocks. The authors suggested that this phylogeographic differentiation might persist along the Mozambique Channel owing to oceanic conditions that influence early passive stages in the green turtle life-cycle. In fact, the northern entrance of the channel is dominated by a large seasonal anticyclonic cell (Donguy & Piton 1991), whereas the central part is characterised by a succession of mesoscale anticyclonic and cyclonic eddies along the Mozambique coast (Schouten et al. 2003, Swart et al. 2010). This succession of gyres generates eddy-eddy interactions and strong dynamic gradients. Even if we considered the important differences that might exist in the behaviour of passive hatchling green turtle and fish larvae, it seems interesting to note that, for both species, the Europa population constitutes an independent gene pool from the northern populations of the Mozambique Channel and that the Juan de Nova population is an intermediate between both.

Microsatellite analysis indicates the important isolation of Reunion and Madagascar from the other localities, particularly when compared to the most distant localities of Europa and Juan de Nova, thus confirming the isolation-by-distance pattern already observed. The Structure analysis indicates a similar genetic signature between Reunion and Madagascar, but the PCA analysis clearly shows a unique allelic constitution for Reunion and the mitochondrial  $\Phi_{st}$  value between Reunion and Madagascar is one of the largest. Thus, the population of Madagascar appears to have a genetic signature intermediate to Reunion and the Mozambique Channel populations. In addition to the large geographic distances that separate Reunion from other localities that could create this genetic isolation, Reunion is the only locality sampled in the deep-ocean ISSG biogeographic area (Longhurst 1998). Both of these elements could explain the genetic differentiation observed. The fact that *Myripristis berndti* is rarely observed in Mauritius and Rodrigues (Fricke 1999, Heemstra et al. 2004) could also be indicative of the rarity of dispersal events of *M. berndti* that could connect the islands of the Mascarene Plateau and the other populations. Thus the genetic signature of the Reunion population might be the consequence of both past and present lack of connectivity with the other localities.

Even if less differentiated, the 5 northern localities of the Mozambique Channel are not in complete panmixia; the results of the Structure analysis tend to

indicate that the 2 nuclear gene pools co-occur in this area but not in the same proportions. Thus, localities separated by <200 km (for example, Mayotte and Mohéli) show significant genetic differences. This admixture could also partly explain the within-population genetic heterogeneity as revealed by high values of  $F_{IS}$  (for example, in Mohéli and Moroni).

Finally, it is important to highlight that some important discrepancies could appear in the structure given by each kind of genetic markers. The highest incoherence is for Kenyan locality, highly divergent on the basis of mitochondrial DNA but not with microsatellites. The Kenyan sample is indeed one of the least divergent population on the basis of microsatellites. The sample is not significantly different from Mohéli and Moroni localities in terms of  $F_{ST}$  and shows an admixture of the 2 genetic pools identified with Structure. But the mitochondrial differences between Kenya and the 9 other localities are the highest—of the same level of magnitude as the difference between the Indian and Pacific Oceans—despite an absence of private haplotypes and the highest haplotypic diversity. The small sample size from Kenya (only 11 fish) might partly exaggerate this result but the total absence of the major Pacific haplotype (only absent from the Kenyan locality) and the low frequency—or absence—of the main Indian haplotypes cannot be explained by sampling bias alone. The Kenyan population is also the only one that does not show a population expansion signature. It seems likely that the Kenyan locality had undergone a different history of colonisation, maybe a secondary event from a restricted SWIO stock with drift due to a founder effect that has created this particular pattern. Mitochondrial differentiation of the northern localities (Kenya and Seychelles versus Tanzania and Mauritius) has also been identified for *Scarus ghobban* (Visram et al. 2010). Further investigations in this area would confirm the historical role of the biogeographic ARAB/EAFR boundary (Longhurst 1998) as a potential barrier to gene flow.

In the context of MPA delineation for biodiversity protection, genetic studies constitute an indirect approach to determine effective dispersal and to delineate stock boundaries (Palumbi 2003). The present study thus provides crucial information about the restricted connectivity in the SWIO, even for a species supposed to be a great disperser. This is of primary interest in the design of the MPA network in the SWIO. It also highlights the importance of using both kind of genetic markers to discriminate the history of past colonisation from more recent connectivity. Finally, single-species studies are clearly insuffi-

cient as a tool for managers; an effort is now necessary to collect additional data in the SWIO, as was recently the case for the more studied Hawaiian archipelago with a comparison of genetic barriers identified in 39 species (Toonen et al. 2010).

*Acknowledgements.* We gratefully acknowledge the authorities of the different countries for allowing sample collections in their waters, particularly from marine parks: Terres australes et Antarctiques Françaises, Direction de l'Agriculture et de la Forêt de Mayotte, the Marine Parc of Mohéli, Kenya Marine and Fisheries Research Institute, and the Kenyan Department of Fisheries. We are very grateful to all people who helped us in the collection of samples: H. Evano, L. Le Rú, H. Sauvignet, R. Bochard, R. Le Goff, J. Wickel, F. Beudard, J. P. Quod, and T. Gazo. We are indebted to S. Ferreira and S. Duthoy from GenoScreen ([www.genoscreen.fr](http://www.genoscreen.fr)) for their work and their continuous help. A special thanks to L. Gagnevin and K. Vital from CIRAD-3P for their help at the genotyping platform. This work was mainly funded by European Funds—the European Cooperation Funds EU POCT-OI 2007-2013 (project CAMP) and the EU 7th Funding Program—Capacities, Research Potential (project RUNSea-Sciences), the French State (DIREN), La Réunion Council (Région Réunion), and a MASMA grant of the WIOMSA. Sample collection for specimen data from Craig et al (2007) was supported by NSF Grant OCE-0929031 (to B. W. Bowen). MTC was supported by NOAA National Marine Sanctuaries Program MOA Grant No. 2005-008/66882.

#### LITERATURE CITED

- Bandelt HJ, Forster P, Röhl A (1999) Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol* 16:37–48
- Bay LK, Choat JH, Herwerden L, Robertson DR (2004) High genetic diversities and complex genetic structure in an Indo-Pacific tropical reef fish (*Chlorurus sordidus*): evidence of an unstable evolutionary past? *Mar Biol* 144: 757–767
- Benjamini Y, Yekutieli Y (2005) False discovery rate controlling confidence intervals for selected parameters. *J Am Stat Assoc* 100:71–80
- Bernardi G, Holbrook SJ, Schmitt RJ, Crane NL, DeMartini E (2002) Species boundaries, populations and color morphs in the coral reef three-spot damselfish (*Dascyllus trimaculatus*) species complex. *Proc Biol Sci* 269:599–605
- Bourjea J, Lapegue S, Gagnevin L, Broderick D and others (2007) Phylogeography of the green turtle, *Chelonia mydas*, in the Southwest Indian Ocean. *Mol Ecol* 16: 175–186
- Bowen BW, Bass AL, Muss A, Carlin J, Robertson DR (2006) Phylogeography of two Atlantic squirrelfishes (Family Holocentridae): exploring links between pelagic larval duration and population connectivity. *Mar Biol* 149: 899–913
- Briggs J (1999) Modes of speciation: marine Indo-West Pacific. *Bull Mar Sci* 65:645–656
- Briggs J (2005) The Marine East Indies: diversity and speciation. *J Biogeogr* 32:1517–1522
- Craig T, Eble JA, Bowen BW, Robertson DR (2007) High genetic connectivity across the Indian and Pacific



- Oceans in the reef fish *Myripristis berndti* (Holocentridae). *Mar Ecol Prog Ser* 334:245–254
- Crawford NG (2010) SMOGD: software for the measurement of genetic diversity. *Mol Ecol Resour* 10:556–557
- Donguy J, Piton B (1991) The Mozambique Channel revisited. *Oceanol Acta* 14:549–558
- 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, Lischer H (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Resour* 10:564–567
- Fricke R (1999) Fishes of the Mascarene Islands (Réunion, Mauritius, Rodriguez): an annotated checklist, with descriptions of new species. Koeltz Scientific Books, Königstein
- Fu YX (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* 147:915–925
- Gaither MR, Toonen RJ, Robertson DR, Planes S, Bowen BW (2010) Genetic evaluation of marine biogeographical barriers: perspectives from two widespread Indo-Pacific snappers (*Lutjanus kasmira* and *Lutjanus fulvus*). *J Biogeogr* 37:133–147
- Goudet J (1995) PCA-GEN for Windows, ver. 1.2. Available at [www2.unil.ch/popgen/softwares/pcagen.htm](http://www2.unil.ch/popgen/softwares/pcagen.htm), Institute of Ecology, University of Lausanne
- Goudet J (2001) FSTAT, a program to estimate and test gene diversities and fixation indices (v 2.9.3.2), [www2.unil.ch/popgen/softwares/fstat.htm](http://www2.unil.ch/popgen/softwares/fstat.htm)
- Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acid Symposium Series* 41:95–98
- Heads M (2005) Towards a panbiogeography of the seas. *Biol J Linn Soc Lond* 84:675–723
- Heemstra E, Heemstra PC, Smale M, Hooper T, Pelicier D (2004) Preliminary checklist of coastal fishes from the Mauritian island of Rodrigues. *J Nat Hist* 38:3315–3344
- Hewitt G (2000) The genetic legacy of the Quaternary ice ages. *Nature* 405:907–913
- Horne JB, van Herwerden L, Choat JH, Robertson DR (2008) High population connectivity across the Indo-Pacific: congruent lack of phylogeographic structure in three reef fish congeners. *Mol Phylogenet Evol* 49:629–638
- Hudson RR (2000) A new statistic for detecting genetic differentiation. *Genetics* 155:2011–2014
- Jost L (2008)  $G_{ST}$  and its relatives do not measure differentiation. *Mol Ecol* 17:4015–4026
- Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25:1451–1452
- Longhurst A (1998) Ecological geography of the sea. Academic Press, San Diego, CA
- Malaua T, Gilles A, Megléc E, Blanquart H and others (2011) High-throughput microsatellite isolation through 454 GS-FLX Titanium pyrosequencing of enriched DNA libraries. *Mol Ecol Res* 11:638–644
- Marko PB, Lee SC, Rice AM, Gramling JM and others (2004) Mislabeling of a depleted reef fish. *Nature* 430:309–310
- Megléc E, Costedoat C, Dubut V, Gilles A, Malaua T, Pech N, Martin JF (2010) QDD: a user-friendly program to select microsatellite markers and design primers from large sequencing projects. *Bioinformatics* 26:403–404
- Mora C, Andrefouet S, Costello MJ, Kranenburg C and others (2006) Ecology: coral reefs and the global network of marine protected areas. *Science* 312:1750–1751
- Narum S (2006) Beyond Bonferroni: less conservative analyses for conservation genetics. *Conserv Genet* 7:783–787
- Nei M (1987) Molecular evolutionary genetics. Columbia University Press, New York, NY
- Palumbi SR (2003) Population genetics, demographic connectivity and the design of marine protected areas. *Ecol Appl* 13:146–158
- Pritchard J, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155:945–959
- R Development Core Team (2010) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna
- Randall JE, Greenfield D (1996) Revision of Indo-Pacific holocentrid fishes of the genus *Myripristis*, with descriptions of three new species. *Indo Pac Fish* 25:1–61
- Reece JS, Bowen BW, Joshi K, Goz V, Larson A (2010) Phylogeography of two moray eels indicates high dispersal throughout the Indo-Pacific. *J Hered* 101:391–402
- Ridgway T, Sampayo EM (2005) Population genetic status of the western Indian Ocean: what do we know? *West Ind Ocean J Mar Sci* 4:1–9
- Ridgway T, Hoegh-Guldberg O, Ayre D (2001) Panmixia in *Pocillopora verrucosa* from South Africa. *Mar Biol* 139:175–181
- Rocha LA, Bowen BW (2008) Speciation in coral-reef fishes. *J Fish Biol* 72:1101–1121
- Rousset F, Raymond M (1997) Statistical analyses of population genetic data: new tools, old concepts. *Trends Ecol Evol* 12:313–317
- Santini F, Winterbottom R (2002) Historical biogeography of Indo-western Pacific coral reef biota: is the Indonesian region a centre of origin? *J Biogeogr* 29:189–205
- Schott F, Xi S, McCreary P (2009) Indian Ocean circulation and climate variability. *Rev Geophys* 47:RG1002 doi: 10.1029/2007RG000245
- Schouten M, Ruijter W, Leeuwen P, Ridderinkhof H (2003) Eddies and variability in the Mozambique Channel. *Deep-Sea Res II* 50:1987–2003
- Song CB, Near TJ, Page LM (1998) Phylogenetic relations among percid fishes as inferred from mitochondrial cytochrome b DNA sequence data. *Mol Phylogenet Evol* 10:343–353
- Swart NC, Lutjeharms JRE, Ridderinkhof H, de Ruijter WPM (2010) Observed characteristics of Mozambique Channel eddies. *J Geophys Res* 115:C09006 doi: 10.1029/2009JC005875
- Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123:585–595
- Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22:4673–4680
- Toonen RJ, Andrews K, Baums I, Bird C and others (2011) Defining boundaries for applying ecosystem-based management: a multispecies case study of marine connectivity across the Hawaiian Archipelago. *J Mar Biol* 2011, doi:10.1155/2011/460173
- Van Oosterhout C, Hutchinson W, Wills D, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Mol Ecol*



- Notes 4:535–538
- Visram S, Yang MC, Pillay RM, Said S, Henriksson O, Grahn M, Chen CA (2010) Genetic connectivity and historical demography of the blue barred parrotfish (*Scarus ghobban*) in the western Indian Ocean. *Mar Biol* 157: 1475–1487
- Voris HK (2000) Maps of Pleistocene sea levels in southeast Asia: shorelines, river systems and time durations. *J Biogeogr* 27:1153–1167
- Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. *Evolution* 38: 1358–1370

Appendix 1. Characterization of the 8 microsatellite loci used for *Myripristis berndti* population survey and the amplification parameters

Locus	Primer sequence (5' to 3')	GenBank accession number	Repeat motif	Annealing temperature	Range size (bp)
Myr-22	CCACTTTGCTGCTGTAATGC CAGTTCCTAGCTGGCTCATGT	JN021280	TCTTA	60	126–176
Myr-25	TTCCCAAATAATTCTCAACTACCA CATAAAGGAAACGGACGGAA	JN021281	CTATT	60	169–219
Myr-28	AAATGGTTTGAGGGTTGGG AAGGCCTTGACCTTGAAAT	JN021282	TGAA	60	91–155
Myr-33	GGGCAACAGGGTAATAAGAGG GACATGTCGGCGATTAACCT	JN021283	TAGA	50	151–163
Myr-37	GCACTATGTTTCAGCTGTGGTT TCTGAATGCATTAAAGTGCCA	JN021284	GAGTT	60	227–262
Myr-44	TGGAGAAAGTTTAAAATGGACA GAGAATCCCACGCAGTTGTT	JN021285	CATA	60	126–154
Myr-55	CAGCAGCATCTGAGCACAAC CATGGGACAGTGCAGTCAAC	JN021286	AGAT	60	106–124
Myr-68	TGGAGAATGATCTGTGCCAA TGCCTGATGTTAGTTTGACAGG	JN021287	GATA	60	112–114

Editorial responsibility: Philippe Borsa,  
Montpellier, France

Submitted: February 28, 2011; Accepted: September 12, 2011  
Proofs received from author(s): December 9, 2011