
Genetic structure of the reef grouper *Epinephelus merra* in the West Indian Ocean appears congruent with biogeographic and oceanographic boundaries

Delphine Muths^{1,*}, Emmanuel Tessier^{2,3} and Jerome Bourjea¹

¹ IFREMER, Delegation de l'Océan Indien, Le Port, France

² Réserve Naturelle Marine de La Réunion, La Saline les Bains, Réunion, France

³ Agence des aires marines protégées, Antenne de Nouvelle-Calédonie, Nouméa, France

*: Corresponding author : Delphine Muths, email address : muthsdelphine@yahoo.fr

Abstract:

The reef fauna connectivity of the West Indian Ocean (WIO) is one of the least studied globally. Here we use genetic analyses of the grouper *Epinephelus merra* (Bloch 1793) to determine patterns of connectivity and to identify barriers to dispersal in this WIO marine area. Phylogeographic and population-level analyses were conducted on cytochrome *b* sequences and microsatellites (13 loci) from 557 individuals sampled in 15 localities distributed across the West Indian Ocean. Additional samples from the Pacific Ocean were used to benchmark the WIO population structure. The high level of divergence revealed between Indian and Pacific localities (of about 4.5% in sequences) might be the signature of the major tectonic and climatic changes operating at the Plio-Pleistocene transition, congruently with numerous examples of Indo-Pacific speciation. In comparison, the *E. merra* sequences from the Indian Ocean constitute a monophyletic clade with a low average genetic distance ($d < 0.5\%$). However both genetic markers indicated some structure within this ocean. The main structure revealed was the isolation of the Maldives from the WIO localities (a different group signature identified by clustering analysis, great values of differentiation). Both marker types reveal further significant structure within the WIO, mainly the isolation of the Mascarene Islands (significant AMOVA and isolation-by-distance patterns) and some patchy structure between the northernmost localities and within the Mozambique Channel. The WIO genetic structure of *E. merra* appeared congruent with main biogeographic boundaries and oceanographic currents.

Keywords: Cytochrome *b* ; marine connectivity ; microsatellite ; reef fish ; West Indian Ocean

1. Introduction

Biogeographic and oceanographic boundaries are known to rule gene flow disruption in many marine species (Aulsebrook 2000). The West Indian Ocean (WIO: defined here as the intertropical waters bounded by the eastern coast of Africa and by the Chagos Ridge on the west side at approximately 72° E) was first divided in three biogeographic areas on the basis of the endemism exhibited by the marine fauna and in correspondence to known tectonic plates (Santini and Winterbottom 2002). These three biogeographic regions are partly in agreement with the four WIO oceanographic provinces proposed by Longhurst (1998) on the basis of global hydrodynamics and water color data: the Indian Monsoon Gyre Province and the Indian South Subtropical Gyre Province (two large oceanographical provinces, separated around 12°S), and, the Eastern Africa Coastal Province and the Northwest Arabian Sea Upwelling Province (two coastal areas: the EAFR including the Mozambique Channel, East Africa coastline until 6°S and the entire Madagascan coastline; the ARAB including the adjacent coastlines of the Arabian Sea to the north of the 6°S longitude). More recently Spalding *et al.* (2007) subdivided these huge biogeographic and oceanographic areas in 11 smallest WIO ecoregions (see Figure 1), the boundaries of these 11 ecoregions still being discussed (Obura 2012). Despite such known biogeographic and oceanographic boundaries within the WIO, most genetic studies failed to identify any strong and congruent restriction to connectivity between localities (Ridgway *et al.* 2001; Ridgway and Sampayo 2005; Silva *et al.* 2010; Muths *et al.* 2012). Studies conducted on the parrot fish *Scarus ghobban* (Visram *et al.* 2010) and on the mangrove crab *Neosarmatium meinerti* (Ragionieri *et al.* 2010) however suggested the genetic isolation of Seychelles from sites on the coastline of Africa. The green turtle *Chelonia mydas* also showed some WIO structure, with individuals nesting in the south and in the north of the Mozambique Channel belonging to separate genetic stocks (Bourjea *et al.* 2007). These studies provide crucial information about marine connectivity restriction in the WIO, congruently with most recently defined biogeographic boundaries (Obura 2012; Spalding *et al.* 2007). But comparisons are difficult as most studies have only analyzed a few localities from a restricted geographical coverage. Moreover, conclusions in terms of population structure are weakened as these studies were based on results from a single genetic marker type, mostly mitochondrial DNA. Through finer, local-scale sampling and additional use of microsatellites, Muths *et al.* (2011) showed restricted connectivity in the WIO for *Myripristis berndti* while this species was previously assumed to be widely-dispersing on the basis of mtDNA (Craig *et al.* 2007). As recently emphasized (Ridgway and Sampayo 2005; Gaither *et al.* 2010), the genetic structure of marine species in the WIO remains globally one of the least studied of tropical waters: too few genetic studies are available, mostly from restricted areas, making it impossible to draw general conclusions about congruency across structure from different marine organisms and their causal mechanisms. Such knowledge is however fundamental for developing suitable marine conservation strategies (Almany *et al.* 2009).

Here, we present a genetic study on the grouper *Epinephelus merra* (Bloch 1793). *E. merra* is a widespread Indo-Pacific reef species, with a natural range from South Africa to the Central-East Pacific (Randall and Heemstra 1991; Craig *et al.* 2012), usually found in sheltered, shallow waters (usually less than 20 meters depth) around isolated coral heads and patches, inside lagoons or bays (Heemstra and Randall 1993). Because of its abundance in shallow water habitats, *E. merra* is important in artisanal fisheries and as a food resource for many local communities (Heemstra and Randall 1993) and provides a good model for the studying reef fish connectivity. The geographical isolation of Mascarene, Seychelles and Maldives islands from the east coast of Africa and the large distances between these archipelagos were supposed to be drivers of genetic divergence even for species with long larval stage duration as *E. merra*. Complex oceanography areas within the Mozambican channel may also act as barriers to dispersal as reported for other species

(Bourjea *et al.* 2007; Muths *et al.* 2011). By using both mitochondrial and microsatellite data and the most important, to date, reef fish sampling within the WIO, we aimed to better determine how the identified biogeographic and oceanographic boundaries determine the realized dispersal of *E. merra* and also how it is congruent with previous marine connectivity studies. Mitochondrial sequences (cytochrome *b*) and microsatellite genotypes (from 13 loci) were thus generated and analysed for 557 individuals of *E. merra*, sampled from 15 sites in the WIO to assess levels of connectivity for this reef fish. Some *E. merra* samples obtained in the Pacific Ocean were also added to the analyses to benchmark the structure observed in the WIO in the context and scale of the broader Indo-Pacific.

2. Methods

2.1. Sampling sites

Samples of *Epinephelus merra* were obtained by spearfishing or were purchased from fishermen. Fifteen sites of the WIO were sampled from March 2007 to November 2010: seven in the Mozambique Channel, three in the north of the studied area, four at the east of Madagascar and one site in the Maldives (Table 1 and Figure 1). Samples of *E. merra* were also collected in two sites of the Pacific Ocean (New Caledonia & Moruroa; see Table 1). The congener *E. hexagonatus* was collected to serve as an outgroup. Tissues were fixed in ethanol 95% and stored at -20° C.

2.2. Genetic analyses

Total genomic DNA was extracted using a DNAeasy Tissue Kit (Qiagen). An 809 base pair fragment of the mitochondrial cytochrome *b* gene region was amplified by PCR using the primers CB12F and CB13R (Marko *et al.* 2004) with an annealing temperature of 54° C. Reactions were performed in 20 μ l, containing 1X PCR buffer, 2 mM $MgCl_2$, 20 μ M of each dNTP, 0.5 μ M of each primer, 0.5 U of Silverstar DNA Polymerase Taq (Eurogentec) and 25 ng of genomic DNA. Sequences were generated on an ABI 3100 sequencer (Applied Biosystems Inc.) in both forward and reverse directions, sequence quality was checked in BIOEDIT (Hall 1999) and multi-alignment done using ClustalW (Thompson *et al.* 1994). Unique cytochrome *b* haplotypes of *Epinephelus* species were submitted to GenBank [JN255254-JN255345; JN545057-JN545096].

Thirteen species-specific microsatellite loci were amplified following the recommended conditions (Muths and Bourjea 2011). Simplex PCR were pooled in three multiplex for genotyping (Multiplex 1: EPI-04, EPI-08, EPI-12, EPI-27, EPI-53 ; Multiplex 2: EPI-05, EPI-07, EPI-35, EPI-50 ; Multiplex 3: EPI-13, EPI-16, EPI-48, EPI-49). Alleles were scored using a co-migrating size standard and identified using GENE MAPPER 4 on an ABI 3100 sequencer (Applied Biosystems Inc.). Full genotypes are available in Supplementary Material.

2.3. Data analyses

The analysis of sequences in jModeltest 0.1.1 (Posada and Crandall 1998; Guindon and Gascuel 2003) indicated that among the 88 candidate models of nucleotide substitution tested, the HKY + G model of nucleotide substitution best fits the data (*i.e.* the model with the smallest AIC score). The subsequent analyses were, therefore, run under these conditions. First, a maximum-likelihood tree was constructed in MEGA 5 (Tamura *et al.* 2011) using the *Epinephelus merra* sequences and rooted by the grouper *Epinephelus hexagonatus*. The statistical robustness of the analysis was estimated by bootstrapping with 1,000 replicates. Following this, analyses were centred on WIO samples only.

From the cytochrome *b* sequences, haplotype (*h*) and nucleotide (π) diversity and Fu's (Fu 1997) F_s -statistic were estimated for each locality in DNAsp 5.0 (Librado and Rozas 2009). Fu's F_s -statistic detects excess of rare alleles and as such could detect a departure from mutation-drift equilibrium. A mismatch distribution curve was constructed with the same software as well as estimation of Harpending's raggedness index. A significant value of Harpending's raggedness index ($P < 0.05$) indicated a departure between observed and expected unimodal distributions and taken as evidence for rejecting the sudden population expansion model. Pairwise values of genetic differentiation (Φ_{st}) were calculated with ARLEQUIN 3.5 (Excoffier and Lischer 2010). Significance levels for multiple tests were corrected using a sequential Benjamini-Yekutieli procedure (Benjamini and Yekutieli 2005). Jost's (Jost 2008) unbiased estimator of divergence (D) was also calculated for each pair of localities using SPADE (available at <http://chao.stat.nthu.edu.tw/softwareCE.html>). A median-joining network (Bandelt *et al.* 1999) was constructed using NETWORK 4.5 (available at <http://www.fluxus-technology.com/>).

For microsatellites, allele frequencies, the mean number of alleles per population (N_{all}), and the observed (H_o) and expected (H_e) heterozygosities (Nei 1987) were calculated in ARLEQUIN 3.5 (Excoffier and Lischer 2010). Allelic richness (R_s) – which corrects the number of alleles according to the smallest sample size - was estimated using Fstat 2.9.3.2 (Goudet 2001). Deviations from Hardy-Weinberg equilibrium were examined for each population, at each locus, by calculating Wright's fixation index F_{is} , as estimated by Weir and Cockerham (1984), and significance was assessed by exact tests using ARLEQUIN 3.5 (Excoffier and Lischer 2010). Genetic differentiation between populations was estimated by calculating Wright's F_{ST} statistic (Wright 1969) and the null hypothesis (the identity of allelic distributions across populations) tested using exact tests. Jost's (2008) D estimator was calculated for each pair of localities using SMOGD (Crawford 2009). Microsatellite data were analysed using STRUCTURE 2.3.2 (Pritchard *et al.* 2000) with an admixture model assuming independent allele frequencies; twenty replicates were run (each with 1.10^5 burn-in generations and 1.10^6 iterations) for each value of K from 1 to 15. The optimal value of K was selected according to Evanno *et al.* (Evanno *et al.* 2005), whereby the number of clusters showing the largest difference in log-likelihoods (ΔK) was selected, as implemented in STRUCTURE Harvester (Earl and VonHoldt 2012). CLUMPP (Jakobsson and Rosenberg 2007) was then used to summarize the individual cluster assignments for the 20 runs at the optimal K .

For both mitochondrial sequences and microsatellite data, the geographic partitioning of genetic structure was investigated using a number of techniques. First, we tested for patterns of isolation by distance. Marine distances were plotted against genetic distance to test for a linear relationship. Following the recommendations of Rousset and Raymond (1997), genetic distance used for the isolation by distance analysis were $\Phi_{st} / (1 - \Phi_{st})$ for mitochondrial data and $F_{ST} / (1 - F_{ST})$ for microsatellite data. The significance of this relationship was tested with Mantel tests, performed in R (R Development Core Team, 2010) using the *ncf* package (available at <http://cran.r-project.org/web/packages/ncf/>). ARLEQUIN 3.5 (Excoffier and Lischer 2010) was used to perform analysis of molecular variance (AMOVA) for both marker sets, independently, with *a priori* groupings based on geographical proximity, biogeographic ecoregions and oceanic currents (e.g. localities at the east of Madagascar *versus* localities at the west, which could be subdivided further as localities within the Mozambique Channel and northern localities; Figure 1 for biogeographic boundaries and currents and Table 2 for grouping details).

Asymmetric migration rates among ecoregions and effective population sizes were estimated using Markov chain Monte Carlo (MCMC) simulations of coalescent genealogies in migrate-n 3.6.4 (Beerli & Felsenstein 2001). For the simulations, 10 short chains of 50 000 steps were followed by three long chains of 500 000 steps. Chains were sampled every 100 steps following a burn-in of 10 000 steps, and default settings were used for the initial estimate of

the θ -value. The number of immigrants per generation (Nm) were then calculated by multiplying final estimates of θ and M .

3. Results

3.1. Indo-Pacific structure

The analysis of 809-bp long cytochrome *b* sequences revealed a high level of genetic structure within *Epinephelus merra*. As shown by the maximum-likelihood tree (Figure 2), the *E. merra* sequences constituted a monophyletic group, but an average sequence divergence of 4.5% separated the *E. merra* haplotypes found in the Indian Ocean (IO) from the fifteen individuals sampled in the Pacific Ocean (with 31 mutations between the least divergent haplotypes). This deep divergence contrasted greatly with the average distance of *E. merra* within each ocean region ($d < 0.5\%$). The pairwise estimate of genetic differentiation (Φ_{st}) between the two groups was 0.94 ($p < 0.001$). As a comparison, divergence between *E. merra* and *E. hexagonatus* (the outgroup) was about 10.6%.

Microsatellite amplification was initially tested on *E. merra* samples from the Pacific. There was complete failure for two of the 13 loci (loci EPI-16 and EPI-49). The pairwise estimate of genetic differentiation (F_{st}) between the two oceans based on these 11 remaining loci was 0.12 ($p < 0.001$).

3.2. Genetic diversity within the West Indian Ocean

The mitochondrial analysis was conducted using 475 sequences from 15 localities of the WIO. This revealed the occurrence of 68 polymorphic sites, from which 76 different haplotypes were identified. The mean haplotype diversity (h) and mean nucleotide diversity (π) for the overall sample were $0.779 (\pm 0.11)$ and $0.0045 (\pm 0.0008)$, respectively (diversities provided for each population in Table 3). The most common haplotype represented 41% of all sequences; it was well represented in 14 of the 15 localities, with frequencies varying from 20% in Tanzania to 66% in Mayotte but was wholly absent from Maldives (Table 3). The proportion of private haplotypes ranged from 0 % in Mayotte to 33 % in Kenya. Consequently, the haplotype diversity was the lowest in Mayotte ($h = 0.583$) and the highest in Kenya ($h = 0.944$).

A total of 557 individuals from the WIO were analysed for 13 polymorphic microsatellite loci. No loci were in linkage disequilibrium ($p < 0.001$) over the whole dataset, supporting the independent assortment of alleles at different loci. Allelic richness was comparable among the 15 localities, with values ranging from 5.28 for Mauritius to 6.50 for Maldives (Table 3), and with a mean of 5.95. The mean observed and expected heterozygosities across all populations were 0.559 and 0.732, respectively. Highly significant ($p < 0.001$) heterozygote deficiencies were observed at all localities, with F_{is} values ranging from 0.160 for Mayotte to 0.339 for Tanzania (Table 3). These heterozygote deficiencies were neither locus- nor population-dependent.

3.3. Demographic analyses and population structure within the West Indian Ocean

Fu's F_s value was negative and significant for the entire sample ($F_s = -5.15$, $p < 0.05$), but not for all individual localities (Table 3). The mismatch distribution (Figure 3) showed no significant deviation from what is expected under a sudden expansion model (Harpending's raggedness index $r = 0.05$, $p = 0.63$). Both implied that the sudden population expansion model could not be rejected.

The haplotype network constructed from the WIO sequences (Figure 4) revealed several common haplotypes, not equally-represented across all localities, from which crowns of rare haplotypes radiated. The distribution of mtDNA haplotypes among populations was plotted on Figure 5. As with the most common haplotype #5, the abundant haplotypes (#6 and #7) and all the rare haplotypes that radiated from these three main haplotypes were absent from Maldives (Figure 4 and 5). Conversely, haplotype #3 and all the rare haplotypes that radiated from it were virtually absent from localities to the east of Madagascar (only one haplotype #3 in Madagascar itself; Figure 4 and 5). Haplotypes #12 and #14 were absent from the Northern localities. Consequently, the AMOVAs indicated that significant genetic variation was associated with groupings done in agreement with geographic regions for both the mitochondrial and microsatellite data (Table 2). The optimal grouping (maximising variance among groups and with non-significant estimates of variance among populations within groups) was obtained when four geographic groups were considered: Maldives was considered alone, and in the WIO, localities from the north, to the east and west of Madagascar constituted three groups ($\Phi_{CT} = 0.217$ for mitochondrial data and $\Phi_{CT} = 0.008$ for microsatellites; $p < 0.001$). The grouping according to the 7 ecoregions induced higher variance among individuals within population. While the among-populations within-groups variance component was always higher than the among-groups variance component for the microsatellite data set, the optimal grouping induced an among-group variance component higher than the among-population variance component for the mtDNA dataset. Whatever the grouping considered, the highest part of the variance was always associated at the within-locality level ($p < 0.001$)

The clustering analysis conducted with STRUCTURE suggested that the highest likelihood of obtaining the present data was to consider that three genetic groups co-occurred ($\Delta K = 21.38, 40.42$ and 6.05 , respectively from $K = 2, 3$ and 4 , confirmed by a “plateau” in likelihood values after $K = 3$). However, only 25 of the 557 individuals (from which 8 were from Mauritius and 4 from Maldives, the 13 remaining ones being more equally dispatched in other localities) were assigned to one cluster with more than 90% probability and each locality had at least a small proportion of individuals assigned to each of the three clusters. The posterior assignment probability plotted per individuals and per locality (Figure 6) provided some little evidence of geographical subdivision: Maldives was predominantly assigned to clade 3 (60%) and the Mascarene Islands to clade 2 (57, 47 and 39% respectively for Mauritius, Rodrigues and Reunion).

The overall Φ_{ST} was 0.154 ($p < 0.001$). Of the 105 pairwise estimates of genetic differentiation (Φ_{st} ; Table 4), 30 were significant after correction. The highest values involved comparisons with the Maldives sample ($0.258 < \Phi_{st} < 0.727$; $p < 0.05$). Congruently, values of Jost’s D (Table 4) were highest for all the comparisons with Maldives [with the maximal value of no shared haplotypes ($D = 1$) obtained in comparisons with the three Mascarene localities: Mauritius, Rodrigues and Reunion]. Similarly but to a lesser extent, other cases of differentiation (significant Φ_{st} -values, confirmed by high values of D) involved the three Mascarene, with the highest values obtained in comparisons with Kenya and Tanzania.

Pairwise fixation F_{ST} and Jost’s D estimates are provided in Table 5 for microsatellite data. The overall F_{ST} was 0.019 ($p < 0.001$). Despite low values, 74 (56 after correction) of the 105 comparisons were significant ($0.009 < F_{ST} < 0.057$; $p < 0.05$). The highest F_{ST} and Jost’s D values were obtained for the Maldives and Mauritius samples, with similar values obtained for these two localities when compared to the Mozambique Channel or the northern localities.

Significant patterns of isolation-by-distance were identified for both the mitochondrial ($r = 0.58$, $p < 0.01$) and microsatellite ($r = 0.64$, $p = 0.01$) data. When Maldives was removed, the mitochondrial relationship became non-significant ($r = 0.19$, $p = 0.13$), but remained significant for microsatellite data ($r = 0.57$, $p = 0.001$).

Estimated gene flows between the seven ecoregions were congruent between the two markers and globally asymmetric (Table 6). The most exporting area is the Mozambique Channel, either Southern or Northern parts, with gene flow orientated towards all other localities by one to two orders of magnitude. Average donating flows of migrants per generation were 200 and 527 for EURJDN and NMOZ respectively while receiving flow is less than 10 migrants. Other localities contribute poorly to genetic exchanges.

4. Discussion

4.1. Indo-Pacific cryptic lineages of *E. merra*

Samples of *E. merra* from the Pacific Ocean, firstly considered to benchmark the structure in the WIO, appeared highly divergent from the Indian samples for both genetic markers (about 4.5% divergence in cytochrome *b* sequences, amplification failure at two of the 13 microsatellite loci for the Pacific samples). Pairwise differentiation values were highly significant ($\Phi_{st} = 0.94$ and $F_{ST} = 0.12$), twice as high between oceans as between the most divergent Indian Ocean samples and six times the overall value within the Indian Ocean. Both genetic markers thus suggest the existence of cryptic lineages in *E. merra*. Based on substitution rates estimated for another grouper species (1.9-2.0% per million years (Ma) (Craig *et al.* 2009)), the divergence between the two lineages identified in *E. merra* was estimated to date back to *circa* 2.3 Ma ago. While caution in interpretation might be required due to the small number of samples from Pacific Ocean and the use of non-specific molecular clock, the estimated date of divergence coincide approximately with the time of the Plio-Pleistocene transition (3 Ma ago (Zachos *et al.* 2001)). This geological period transition was the consequence of major tectonic and climatic changes (Gourlan *et al.* 2008) that have played a vicariant role on marine biodiversity (Cowman and Bellwood 2013) and might be one of the main mechanisms responsible for the emergence of *E. merra* lineages. The genetic signature of these Plio-Pleistocene events has been widely recognized in Indo-Pacific marine taxa (Benzie 1999; Reid *et al.* 2006; Gaither and Rocha 2013; Hoareau *et al.* 2013). Better understanding the history and the geographical location of the separation of the cryptic lineages in *Epinephelus merra* would clearly necessitate further sampling in the Pacific Ocean and around the Indo-Australian Archipelago but appeared interesting in the wider context of Indo-Pacific speciation.

4.2. Maldives isolation

Even though the average genetic distance between *E. merra* sequences from the Indian Ocean was low ($d < 0.5\%$), genetic heterogeneity was observed within this ocean. First of all was the isolation of Maldives. Both mitochondrial data and microsatellites pairwise estimates of fixation appeared highly significant when comparing Maldives to any WIO localities. The haplotype repartition within the WIO appeared geographically biased: one group of haplotype completely excluded from the Maldives (haplotypes #5, 6, 7) and a second group excluding the Mascarene Islands (see occurrence of haplotype #3, 11; see Figure 5 and 6). Consequently, the highest values of pairwise differentiation are obtained between Maldives and the Mascarene Islands for both genetic markers – with the extreme values of Jost's *D* (1) obtained for the mitochondrial data. The STRUCTURE analysis also assigned the Maldives samples and the rest of the WIO to different genetic clusters. The fact that the Maldives is the most distant locality (4000 km apart) of our study sites, separated from the others by the strong eastward Equatorial Counter Current, and the only locality in the Northern hemisphere, might be sufficient to explain the low degree of connectivity between this area and the WIO. The great genetic isolation between samples from Maldives and those from the WIO adds support to the status of Maldives being part of a different biogeographic province

than the WIO province (Spalding *et al.* 2007). Briggs and Bowen (2012) already suggested, on the basis of surveys of the fish fauna, that the Maldives Archipelago might be better included in the western Indo-Polynesian Province rather than in the WIO. The fact that cytochrome *b* sequences from Maldives clade with Indian sequences rather than Pacific sequences tend to mitigate this last proposal but all these data underline the peculiar status of Maldives.

4.3. Population structure within the West Indian Ocean

Both mitochondrial and microsatellite data underline the isolation of localities at the east of Madagascar from the other localities of the WIO (significant partition of molecular variance and highest values of differentiation), the three Mascarene localities being the most differentiated and samples from Madagascar that showed an „in-between“ genetic signature. Gene flow disruption between the Mascarene Islands and the rest of the WIO was previously observed for another reef fish species, *M. berndti* (Muths *et al.* 2011). These three islands are situated in the deep-ocean Indian South Subtropical Gyre Province (Longhurst 1998), an identified different oceanographic region. It is also worth noting that large geographic distances (> 1000km) and the landmass of Madagascar itself contribute to their physical isolation and separate these three Mascarene localities from other localities in WIO. These elements explain how such different genetic signature could be observed between the Mascarene Islands and the rest of the WIO. Moreover, the Mascarene are characterized by low haplotype diversity but a high private haplotype proportion, which was suspected for remote islands (Obura 2012), and confirm the status of Mascarene as a clearly independent ecoregion within the WIO. Finally, differentiation within Mascarene was also observed, with Mauritius samples showing a different genetic signature, a pattern already observed for the highly connected *Lutjanus kasmira* (Muths *et al.* 2012). Thus, regarding the Mascarene as an effective connected network (Crochelet *et al.* 2013) has to be taken with caution.

The other localities of the WIO appeared less differentiated among them, the Kenya locality being the rare one that induced significant differentiation pairwise values with mitochondrial data while Geyser and Glorieuses showed significant structure with microsatellites. A significant part of molecular variance was associated when grouping the northernmost localities (Kenya, Tanzania and Seychelles) against the Mozambique Channel localities, these localities being separated by the strong westward flowing South Equatorial Current (SEC). The northernmost localities (Kenya, Tanzania and Seychelles) showed more similarity with Maldives, but caution in interpretation might be applied as the northern localities displayed small sample sizes (only 9 and 10 samples for Kenya and Tanzania respectively). Within the Mozambique Channel, significant microsatellite structure was observed when separating Geyser and Glorieuses from the other Channel localities, similarly to the structure observed for *M. berndti* (Muths *et al.* 2011). Genetic structure along the Mozambique Channel was also reported for the green turtle *Chelonia mydas* (Bourjea *et al.* 2007) and for brittle-stars (Hoareau *et al.* 2013) without the emergence of a clear shared pattern. Local oceanic conditions [*i.e.* the northern entrance of the channel being dominated by a large seasonal anticyclonic cell (Donguy and Piton 1991), followed by a succession of mesoscale anticyclonic and cyclonic eddies along the Mozambique coast (Schouten *et al.* 2003; Swart *et al.* 2010)] create a complex and fluctuating system of mixing and retention that generates patchy pattern of connectivity depending of species life history traits. Further investigation might be necessary to better surround the connectivity patterns in this complex area and help to resolve the ambiguous definition of ecoregion boundaries in this area (Obura 2012; Spalding *et al.* 2007). Whatever, as shown by migration rates and despite a global westward oceanic inflow, localities from the Mozambique Channel appeared as the main source of genetic diversity for the whole WIO. It is consistent with high species richness of reef-building corals observed in the northern Mozambique Channel (Obura 2012) and could be viewed as an argument for their hypothesis of WIO hotspot of biodiversity, reinforcing the interest for this area.

Generally speaking, the present study provides information about the disruption to gene flow of *Epinephelus merra* in the WIO. Three main barriers to dispersal were identified: distance and open-waters isolating the Mascarene Islands, the South Equatorial Current in separating the northern sites (Tanzania, Kenya, Seychelles) and the complex oceanography that created patchy structure within the Mozambique Channel. The congruence between this structure and some of the biogeographic boundaries as identified for reef-building corals in Obura (2012) reinforce the idea of a dominant role played by present-day oceanography in shaping WIO population connectivity.

Acknowledgements

This work was funded by the European Union – through the European Cooperation Fund EU POCT-OI 2007-2013 (project CAMP) and the EU 7th Funding Program - Capacities, Research Potential (project RUNSeaSciences), the French Government (DEAL La Réunion), La Réunion Council (Région Réunion) and a MASMA grant from the WIOMSA.

We are hugely thankful to G. Gouws and G. Hugues for improving the English in this manuscript. We extend our gratitude to M. Mwale, J. Mwaluma and A. Mwandya for our collaboration under the MASMA project managed by G. Gouws. We acknowledge the authorities for allowing sample collection in their waters: Terres Australes et Antarctiques Françaises, Réserve Naturelle Marine de la Réunion, Direction de l'Agriculture et de la Forêt de Mayotte, Marine Park of Mohéli, Kenya Marine and Fisheries Research Institute and Kenyan Department of Fisheries, Seychelles Fishing Authority. We are very grateful to all people who helped us with the collection of samples: H. Evano, L. Le Rû, H. Sauvignet, C. Cadet, K. Pothin, R. Le Goff, R. Bochard and F. Beudard, Shoals Rodrigues and anonymous fishermen from the whole WIO. Samples from the Pacific Ocean were kindly provided by P. Borsa, those from Maldives by L. Vigliola and L. Bigot thanks to their project IFB (Institut Français de la Biodiversité). We extend special thanks to C. Poncet from the INRA-Gentiane platform and to L. Gagnevin and K. Vital from the CIRAD-3P for their help with genotyping and E. Wicker for kindly inviting D. Muths to the population genomics lesson. We thank E. Billard for her help with genetics software bugs. Finally, we thank the anonymous reviewers for their meaningful remarks on the manuscript.

References

- Almany G, Connolly S, Heath D, Hogan J, Jones G, McCook L, Mills M, Pressey R, Williamson D (2009) Connectivity, biodiversity conservation and the design of marine reserve networks for coral reefs. *Coral Reefs* **28**:339–351.
- Avice JC (2000) *Phylogeography: the history and formation of species*, Cambridge: Harvard University press : 447pp
- Bandelt HJ, Forster P, Röhl A. (1999) Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution* **16**:37–48.
- Beerli P, Felsenstein J (2001) Maximum likelihood estimation of a migration matrix and effective population sizes in subpopulations using a coalescent approach. *Proceedings of the National Academy of Sciences, USA* **98**: 4563–4568.

- Benjamini Y, Yekutieli Y (2005) False discovery rate controlling confidence intervals for selected parameters. *Journal of the American Statistical Association* **100**:71–80.
- Benzie JAH. (1999) Genetic Structure of Coral Reef Organisms: Ghosts of Dispersal Past. *American Zoology* **39**:131–145.
- Bourjea J, Lapegue S, Gagnevin L, Broderick D, Mortimer JA, Ciccione S, Roos D, Taquet C, Grizel H. (2007) Phylogeography of the green turtle, *Chelonia mydas*, in the Southwest Indian Ocean. *Molecular Ecology* **16**:175–186.
- Briggs JC, Bowen BW (2012) A realignment of marine biogeographic provinces with particular reference to fish distributions. *Journal of Biogeography* **39**:12–30.
- Cowman PF, Bellwood DR (2013) Vicariance across major marine biogeographic barriers: temporal concordance and the relative intensity of hard versus soft barriers. *Proceedings of the Royal Society B: Biological Sciences* **280**: 20131541.
- Craig MT, Eble JA, Bowen BW, Robertson DR (2007) High genetic connectivity across the Indian and Pacific Oceans in the reef fish *Myripristis berndti* (Holocentridae). *Marine Ecology Progress Series* **334**:245–254.
- Craig MT, Graham R, Torres R, Hyde JR, Freitas MO, Ferreira BP, Hostim-Silva M, Gerhardinger LC, Bertoincini AA, Robertson DR (2009) How many species of goliath grouper are there? Cryptic genetic divergence in a threatened marine fish and the resurrection of a geopolitical species. *Endangered Species Research* **7**:167–174.
- Craig MT, Mitcheson YJS, Heemstra PC (2012) *Groupers of the world: A field and market guide*. NISC (Pty) Ltd, Grahamstown, 2011, 356 pp
- Crawford NG (2009) SMOGD: software for the measurement of genetic diversity. *Molecular Ecology Resources* **10**:556–557.
- Crochelet E, Chabanet P, Pothin K, Lagabrielle E, Roberts J, Pennober G, Lecomte-Finiger R, Petit M (2013) Validation of a fish larvae dispersal model with otolith data in the Western Indian Ocean and implications for marine spatial planning in data-poor regions. *Ocean & Coastal Management* **86**: 13–21.
- Donguy J, Piton B (1991) The Mozambique Channel revisited. *Oceanologica Acta* **14**:549–558.
- Earl D, Von Holdt B (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetic Resources* **4**:359–361.
- Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software structure : a simulation study. *Molecular Ecology* **14**:2611–2620.
- Excoffier L, Lischer H (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.
- Fu YX (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* **147**:915–925.
- Gaither MR., Rocha LA (2013) Origins of species richness in the Indo-Malay-Philippine biodiversity hotspot: evidence for the centre of overlap hypothesis. *Journal of Biogeography* **40**: 1638–1648. doi: 10.1111/jbi.12126

- 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*). *Journal of Biogeography* **37**:133–147.
- Goudet J (2001) FSTAT, a program to estimate and test gene diversities and fixation indices (v 2.9.3.2). Available from: <http://www2.unil.ch/popgen/softwares/fstat.htm>
- Gourlan AT, Meynadier L, Allègre CJ (2008) Tectonically driven changes in the Indian Ocean circulation over the last 25 Ma: Neodymium isotope evidence. *Earth and Planetary Science Letters* **267**: 353-364
- Guindon S and Gascuel O (2003) A simple, fast and accurate method to estimate large phylogenies by maximum-likelihood. *Systematic Biology* **52**: 696-704
- 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
- Heemstra PC, Randall JE. (1993) *Groupers of the World (Family Serranidae, Subfamily Epinephelinae)*. An annotated and illustrated catalogue of the grouper, rockcod, hind, coral grouper and lyretail species known to date. FAO Fisheries Synopsis, No. **125**, 16: 382 pp
- Hoareau TB, Boissin E, Paulay G, Bruggeman JH (2013) The Southwestern Indian Ocean as a potential marine evolutionary hotspot: perspectives from comparative phylogeography of reef brittle-stars. *Journal of Biogeography* **40**: 2167–2179doi: 10.1111/jbi.12155
- Jakobsson M, Rosenberg N (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* **23**:1801–1806.
- Jost L (2008) GST and its relatives do not measure differentiation. *Molecular Ecology* **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*. San Diego, USA: Academic Press, 398 pp
- Marko PB, Lee SC, Rice AM, Gramling JM, Fitzhenry TM, McAlister JS, Harper GR, Moran AL (2004) Mislabeling of a depleted reef fish. *Nature* **430**:309–310.
- Muths D, Bourjea J (2011) Characterization of thirteen new polymorphic microsatellite markers from the honeycomb grouper *Epinephelus merra*. *Conservation Genetic Resources* **3**:629–631.
- Muths D, Gouws G, Mwale M, Tessier E, Bourjea J (2012) Genetic connectivity of the reef fish *Lutjanus kasmira* at the scale of the West Indian Ocean. *Canadian Journal of Fisheries and Aquatic Sciences* **69**: 842–853.
- Muths D, Tessier E, Gouws G, Craig M, Mwale M, Mwaluma J, Mwandya A, Bourjea J (2011) Restricted dispersal of the reef fish *Myripristis berndti* at the scale of the South-West Indian Ocean compared to known transoceanic species connectivity. *Marine Ecology Progress Series*. **443**: 167–180.
- Nei M (1987) *Molecular Evolutionary Genetics*. New York: Columbia University Press :512p

- Obura D (2012) The Diversity and Biogeography of Western Indian Ocean Reef-Building Corals. *PLoS ONE* **7**:e45013.
- Posada D, Crandall KA (1998) Model Test: testing the model of DNA substitution. *Bioinformatics* **14**:817 – 818.
- Pritchard J, Stephens M, Donnelly P (2000) Inference of Population Structure Using Multilocus Genotype Data. *Genetics* **155**:945–959.
- Ragionieri L, Cannicci S, Schubart C, 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.
- Randall JE, Heemstra PC (1991) Revision of Indo-Pacific groupers (Perciformes: Serranidae: Epinephelinae), with descriptions of five new species. *Indo-Pacific Fishes* **20**:1–332.
- Reid DG, Lal K, Mackenzie-Dodds J, Kaligis F, Littlewood DTJ, Williams ST (2006) Comparative phylogeography and species boundaries in *Echinolittorina* snails in the central Indo-West Pacific. *Journal of Biogeography* **33**:990–1006.
- Ridgway T, Hoegh-Guldberg O, Ayre D (2001) Panmixia in *Pocillopora verrucosa* from South Africa. *Marine Biology* **139**:175–181.
- Ridgway T, Sampayo EM (2005) Population Genetic Status of the Western Indian Ocean: What do we Know? *Western Indian Ocean Journal of Marine Science* **4**:1–9.
- Rousset F, Raymond M. (1997) Statistical analyses of population genetic data : new tools, old concepts. *Trends in Ecology & Evolution* **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? *Journal of Biogeography* **29**:189–205.
- Schouten M, Ruijter W, Leeuwen P, Ridderinkhof H (2003) Eddies and variability in the Mozambique Channel. *Deep-Sea Research II*:1987–2003.
- Silva I, Mesquita N, Paula J (2010) 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.
- Spalding M, Fox H, Allen GR, Davidson N, Ferdana Z, Finlayson M, Halpern BS, Jorge MA, Lombana A, Lourie SA, Martin KD, McManus E, Molnar J, Recchia CA, Robertson J (2007) Marine ecoregions of the world: a bioregionalization of coastal and shelf areas. *Bioscience* **57**:573–583.
- Swart NC, Lutjeharms JRE, Ridderinkhof H, de Ruijter WPM (2010) Observed characteristics of Mozambique Channel eddies. *Journal of Geophysical Research* **115**:C09006.
- Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Molecular Biology and Evolution*.
- Thompson J, Higgins D, Gibson T. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acid Research* **22**:4673–4680.

- 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. *Marine Biology* **157**:1475–1487.
- Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. *Evolution* **38**: 1358-1370
- Wright F (1969) *Volume 2: The theory of gene frequencies*. In: Evolution and the genetics of population. Chicago: Chicago Press: 512p.
- Zachos J, Pagani M, Sloan L, Thomas E, Billups K (2001) Trends, Rhythms, and Aberrations in Global Climate 65 Ma to Present. *Science* **292**:686–693.

Author contributions: M.D, E.T. and J.B. conceived the project; M.D and J.B. contributed to the sampling; M.D. collected the genetic data, conducted data analyses and drafted the manuscript, which was approved by co-authors.

Authors' information: Delphine Muths focused her work on population dynamics, connectivity, molecular ecology, phylogeography and speciation, both on temperate and intertropical marine fauna. J. Bourjea focuses his research on conservation biology of marine resources, using several scientific approaches at oceanic scales to understand marine connectivity. E. Tessier is broadly interested in marine conservation; he was involved in the present study as part of his manager position of a Marine Protected Area at Reunion Island.

Tables

Table 1. Characteristics of the fifteen localities sampled for *Epinephelus merra* in the Indian Ocean and the two localities in the Pacific Ocean.

Group	Population location	Label	GPS coordinates		N (mtDNA)	N (msat)	Sampling date
Mozambique Channel	Europa	EUR	22°24 S	40°23 E	40	48	May-2010
	Juan de Nova	JDN	17°03 S	42°47 E	41	48	May-2010
	Mayotte	MAY	12°52 S	45°15 E	9	9	March-08
	Moheli	MOH	12°24 S	43°41 E	46	48	October-09
	Moroni	MOR	11°47 S	43°14 E	46	48	October-09
	Geyser bank	GEY	12°21 S	46°26 E	32	36	May-2010
	Glorieuses	GLO	11°34 S	47°23 E	43	61	May-2009
North area	Tanzania	TAN	6°48 S	39°16 E	10	12	October-09
	Kenya	KEN	4°65 S	39°38 E	9	13	September-10
	Seychelles	SEY	4°35 S	55°27 E	23	25	March-09
East side Madagascar	Madagascar east	MAD	16°50 S	49°55 E	38	48	November-10
	Reunion	RUN	21°05 S	55°14 E	30	30	March-07
	Mauritius	MAU	20°16 S	57°51 E	31	45	October-10
	Rodrigues	ROD	19°71 S	63°42 E	44	46	October-10
	Maldives	MAL	5°24 N	73°15 E	33	40	June-09
Pacific ocean	New Caledonia	NCA	20°59 S	165°02 E	11	-	
	Moruroa	MOA	17°29 S	50°08 W	4	-	

N(mtDNA): number of sequences obtained

N(msat): number of fish genotyped

Table 2. Analysis of molecular variance (AMOVA) for *Epinephelus merra*. The different groupings of the fifteen localities tested are indicated in the grey part of the table, while the remainder of the table presents the corresponding results of the AMOVAs based on cytochrome *b* and microsatellites.

Site	Grouping pool:				cytochrome <i>b</i>				Microsatellite					
	(a)	(b)	(c)	(d)	Source of variation	d.f.	% variation	Fixation index	p	d.f.	% variation	Fixation index	p	
EUR	1	1	1	1	(a) Among groups	6	17.77	0.177	< 0.001	6	0.76	0.007	< 0.05	
JDN	1	1	1	1		Among populations within groups	8	-0.23	-0.002	0.72	8	1.29	0.013	< 0.001
MAY	2	1	1	1		Among individuals within pop.	460	82.47	0.175	< 0.001	1099	97.95	0.020	< 0.001
MOH	2	1	1	1	(b) Among groups	3	21.71	0.217	< 0.001	3	0.86	0.008	< 0.01	
MOR	2	1	1	1		Among populations within groups	11	0.07	0.001	0.54	11	1.36	0.013	< 0.001
GLO	2	1	1	1		Among individuals within pop.	460	78.22	0.217	< 0.001	1099	97.78	0.022	< 0.001
TAN	3	2	2	1	(c) Among groups	2	14.74	0.147	< 0.001	2	0.58	0.005	< 0.001	
KEN	3	2	2	1		Among populations within groups	12	5.17	0.060	< 0.001	12	1.55	0.015	< 0.001
SEY	4	2	2	1		Among individuals within pop.	460	80.09	0.199	< 0.001	1099	97.87	0.021	< 0.001
MAD	5	3	3	2	(d) Among groups	1	7.26	0.072	< 0.001	1	0.38	0.003	< 0.05	
RUN	6	3	3	2		Among populations within groups	12	0.29	0.003	0.33	12	1.41	0.014	< 0.001
MAU	6	3	3	2		Among individuals within pop.	428	92.46	0.075	< 0.001	1020	98.20	0.017	< 0.001
ROD	6	3	3	2										
MAL	7	4	2	0										

Table 3. Molecular diversity indices for the fifteen localities of *Epinephelus merra* sampled in the West Indian Ocean.

Locality	Cytochrome <i>b</i>							Microsatellite						
	<i>N</i>	<i>Nhap</i>	<i>h</i>	π	Main hap. proportion	Private hap. proportion	<i>Fu Fs</i>	<i>N</i>	<i>Nall</i>	<i>Rs</i>	Private all. proportion	<i>Ho</i>	<i>He</i>	<i>Fis</i>
EUR	40	15	0.841	0.0031	0.37	0.05	-6.41**	48	8.00	5.56	0.057	0.557	0.721	0.237***
JDN	41	10	0.695	0.0026	0.51	0.12	-2.10	48	7.92	5.53	0.009	0.562	0.729	0.239***
MAY	9	4	0.583	0.0013	0.66	0.00	-0.92	9	5.76	5.76	0.086	0.649	0.723	0.160***
MOH	46	16	0.794	0.0027	0.41	0.15	-8.16**	48	8.23	5.73	0.009	0.589	0.722	0.194***
MOR	46	13	0.671	0.0020	0.56	0.10	-7.69***	48	8.15	5.60	0.036	0.592	0.732	0.201***
GEY	32	12	0.866	0.0026	0.31	0.15	-4.75*	36	8.07	5.66	0.062	0.527	0.727	0.287***
GLO	43	19	0.863	0.0026	0.32	0.20	-13.4***	61	9.07	5.91	0.078	0.532	0.724	0.274***
TAN	10	7	0.933	0.0034	0.20	0.20	-2.29	12	6.53	5.90	0.020	0.480	0.687	0.339***
KEN	9	7	0.944	0.0039	0.22	0.33	-2.39*	13	6.30	5.67	0.026	0.656	0.703	0.107***
SEY	23	8	0.810	0.0026	0.34	0.17	-1.63	25	8.15	5.97	0.009	0.557	0.707	0.231***
MAD	38	10	0.652	0.0014	0.57	0.05	-5.07**	48	8.61	5.75	0.000	0.554	0.729	0.249***
RUN	30	11	0.770	0.0017	0.43	0.16	-1.31	30	7.38	5.47	0.009	0.546	0.698	0.234***
MAU	31	11	0.724	0.0019	0.51	0.19	-5.66**	45	7.69	5.28	0.020	0.483	0.691	0.310***
ROD	44	7	0.652	0.0013	0.54	0.02	-1.84	46	8.15	5.67	0.037	0.531	0.725	0.277***
MAL	33	11	0.632	0.0012	0.00	0.30	-8.47***	40	9.84	6.50	0.011	0.560	0.751	0.266***
TOTAL	475	76	0.779	0.0045	0.41	-	-5.15*	557	14.6	5.95	-	0.559	0.732	

N the number of individuals analyzed per marker

For cytochrome *b*: *Nhap* the number of haplotypes per locality, *h* the haplotype diversity, π the nucleotide diversity, the proportions of the most common (#5) and private haplotypes, *Fu*'s *Fs* value

For microsatellites: *Nall* the mean number of alleles per locus, *Rs* the allelic richness, the proportion of private alleles, *Ho* and *He* the observed and expected heterozygosities, *Fis* the fixation index.

Significance is noted as * for $p < 0.05$, ** for $p < 0.01$, *** for $p < 0.001$.

Table 4. Pairwise mtDNA Φ_{st} values of differentiation among the fifteen populations of *Epinephelus merra* (below diagonal) and Jost's D (above diagonal). Population labels are indicated in Table 1

	EUR	JDN	MAY	MOH	MOR	GEY	GLO	TAN	KEN	SEY	MAD	RUN	MAU	ROD	MAL
EUR	--	0.096	0.054	0.026	0.054	0.010	0.019	0.025	0.042	0.084	0.059	0.031	0.038	0.094	0.868
JDN	-0.004	--	-0.037	-0.022	0.009	0.053	0.038	0.133	0.098	0.022	0.041	0.100	0.051	0.086	0.604
MAY	0.009	-0.019	--	-0.001	-0.062	0.090	0.084	0.218	0.268	0.076	-0.080	0.003	-0.046	-0.031	0.828
MOH	-0.009	-0.018	-0.016	--	0.031	-0.042	-0.047	-0.037	-0.054	-0.023	0.058	0.082	0.067	0.109	0.579
MOR	0.034*	0.008	-0.05	0.011	--	0.130	0.103	0.267	0.238	0.118	-0.012	0.067	0.002	0.030	0.848
GEY	-0.002	-0.012	-0.032	-0.014	0.007	--	-0.107	-0.283	-0.105	-0.061	0.116	0.041	0.106	0.140	0.622
GLO	0.000	-0.012	-0.037	-0.012	0.001	-0.015	--	-0.238	-0.132	-0.073	0.126	0.060	0.099	0.137	0.608
TAN	-0.021	0.008	0.071	-0.00	0.097	0.004	0.010	--	-0.455	-0.221	0.262	0.056	0.226	0.232	0.442
KEN	0.020	0.065	0.169	0.058	0.191***	0.079*	0.083*	-0.048	--	-0.024	0.332	0.325	0.307	0.399	0.315
SEY	0.004	-0.012	-0.038	-0.009	0.007	-0.013	-0.01	0.005	0.079	--	0.139	0.033	0.107	0.107	0.527
MAD	0.089***	0.054*	-0.057	0.056*	0.003	0.034	0.036*	0.191**	0.306***	0.045	--	0.009	-0.026	-0.007	0.955
RUN	0.107*	0.081**	-0.040	0.081***	0.032*	0.051*	0.052*	0.185**	0.303***	0.047	-0.008	--	-0.026	-0.020	1.000
MAU	0.095***	0.070**	-0.046	0.070***	0.019	0.043*	0.044*	0.170***	0.281***	0.048*	-0.013	-0.015	--	-0.007	1.000
ROD	0.149***	0.107***	-0.016	0.113***	0.046**	0.088***	0.081***	0.262***	0.381***	0.080**	0.004	-0.008	0.001	--	1.000
MAL	0.404***	0.472***	0.687***	0.456***	0.593***	0.507***	0.492***	0.400***	0.258***	0.529***	0.690***	0.695***	0.675***	0.727***	--

Significant Φ_{st} values are indicated by * for $p < 0.05$ and values still significant after Benjamini-Yekutieli correction for 105 comparisons are indicated by bold *** - $p < 0.009$.

Table 5. Pairwise microsatellite F_{st} values of differentiation among the fifteen populations of *E. merra* (below diagonal) and Jost's D (above diagonal). Population labels are indicated in Table 1

	EUR	JDN	MAY	MOH	MOR	GEY	GLO	TAN	KEN	SEY	MAD	RUN	MAU	ROD	MAL
EUR	--	0.002	0.016	0.002	0.001	0.008	0.013	0.006	0.018	0.012	0.000	0.006	0.050	0.016	0.073
JDN	0.011 ***	--	0.013	0.004	0.000	0.018	0.010	0.000	0.016	0.009	0.007	0.006	0.048	0.009	0.048
MAY	0.021 *	0.024 *	--	0.010	0.006	0.014	0.016	0.017	0.020	0.011	0.011	0.013	0.071	0.028	0.065
MOH	0.005	0.010 ***	0.016	--	0.007	0.009	0.017	0.008	0.029	0.014	0.007	0.005	0.034	0.009	0.064
MOR	0.003	0.004	0.016 *	0.007 *	--	0.007	0.009	0.010	0.032	0.019	0.000	0.002	0.033	0.004	0.067
GEY	0.019 ***	0.013 ***	0.017	0.020 ***	0.011 ***	--	0.005	0.033	0.020	0.010	0.004	0.009	0.050	0.019	0.057
GLO	0.018 ***	0.012 ***	0.015	0.016 ***	0.010 ***	0.004	--	0.038	0.017	0.006	0.010	0.007	0.058	0.023	0.069
TAN	0.018	0.014	0.033	0.018 *	0.020 *	0.037 ***	0.031 ***	--	0.029	0.018	0.016	0.051	0.043	0.019	0.073
KEN	0.019 *	0.013	0.019	0.022 *	0.018 *	0.012	0.015	0.035 ***	--	0.012	0.021	0.019	0.076	0.017	0.065
SEY	0.023 ***	0.010	0.010	0.018 ***	0.012	0.017 *	0.007	0.024 *	0.013	--	0.014	0.003	0.048	0.024	0.035
MAD	0.002	0.007 *	0.014	0.007	0.002	0.013 ***	0.014 ***	0.020 *	0.014	0.018 ***	--	0.003	0.052	0.011	0.091
RUN	0.018 ***	0.008 *	0.020 *	0.014 ***	0.004	0.011 *	0.009 *	0.031 ***	0.013	0.007	0.008	--	0.026	0.004	0.051
MAU	0.043 ***	0.034 ***	0.046 ***	0.035 ***	0.026 ***	0.030 ***	0.031 ***	0.043 ***	0.047 ***	0.032 ***	0.032 ***	0.022 ***	--	0.012	0.099
ROD	0.022 ***	0.011 ***	0.020	0.013 ***	0.009 ***	0.017 ***	0.017 ***	0.021 *	0.024	0.016 *	0.012 ***	0.006	0.009 ***	--	0.087
MAL	0.040 ***	0.026 ***	0.038 ***	0.035 ***	0.036 ***	0.028 ***	0.031 ***	0.043 ***	0.028 ***	0.030 ***	0.041 ***	0.041 ***	0.057 ***	0.047 ***	--

Significant F_{st} values are indicated by * for $p < 0.05$ and values still significant after Benjamini-Yekutieli correction for 105 comparisons are indicated by bold *** - $p < 0.009$.

Table 6. Gene flow estimates (Nm, with N the effective female population size and m the migration rate) of *Epinephelus merra* across the West Indian Ocean obtained by migrate-n simulations. Donating populations are in columns (with estimates of Θ , the parameter representing the effective population size) and receiving areas are in rows. Locations were grouped according to the seven ecoregions (population labels are as in Table 1, except for NMOZ which includes the five localities in the Northern part of the Mozambique Channel and MASCA which includes the three Mascarene Islands).

		IMMIGRATION FROM:													
		Cytochrome <i>b</i>							Microsatellite						
		EURJDN	NMOZ	MAD	MASCA	TANKEN	SEY	MAL	EURJDN	NMOZ	MAD	MASCA	TANKEN	SEY	MAL
Θ		0.0025	0.0126	0.0012	0.0446	0.0006	0.0011	0.0020	0.0445	0.0984	0.0202	0.0182	0.0126	0.0077	0.0147
RECEIVING POPULATIONS:	EURJDN	-	31.00	0.79	0.09	8.40	< 1.10e-6	< 1.10e-6	-	24.51	2.09	1.23	0.87	0.15	1.24
	NMOZ	22.30	-	< 1.10e-6	0.31	0.07	< 1.10e-6	5.74	7.11	-	0.60	3.23	0.47	0.46	0.36
	MAD	183.0	589.0	-	15.0	< 1.10e-6	1.57	< 1.10e-6	7.43	19.78	-	3.70	0.90	0.48	0.27
	MASCA	91.60	< 1.10e-6	8.59	-	< 1.10e-6	0.56	< 1.10e-6	7.52	12.57	1.23	-	0.98	0.28	0.88
	TANKEN	3.13	1.56	< 1.10e-6	< 1.10e-6	-	0.21	0.68	8.29	13.94	1.32	0.89	-	0.19	0.57
	SEY	702.0	1930.0	< 1.10e-6	< 1.10e-6	< 1.10e-6	-	18.31	12.09	11.65	0.96	0.53	1.97	-	2.72
	MAL	< 1.10e-6	85.6	< 1.10e-6	< 1.10e-6	103.0	< 1.10e-6	-	4.11	13.88	0.53	2.70	0.75	0.18	-

Figure legends

Figure 1. Map showing the fifteen *E. merra* localities sampled for the study. The number of samples per locality analysed using the mtDNA marker are shown in brackets. The dashed blue lines correspond to the biogeographic boundaries between the 11 WIO ecoregions [as adjusted by Obura (2012) from the Marine Ecoregions Of the World (Spalding *et al.* 2007)].

Figure 2. Maximum-likelihood tree of cytochrome *b* sequences for the included *Epinephelus* species. Only the 10 most common haplotypes of *E. merra* in the WIO were plotted. White indicates Indian Ocean samples and black Pacific Ocean samples. Numbers at each node are the percentage bootstrap values of 1000 replicates (only indicated when greater than 90%).

Figure 3. Mismatch distribution for *E. merra* based on 475 cytochrome *b* sequences from the Indian Ocean. The black line represents the observed and the dashed grey line the simulated pairwise differences under a sudden demographic expansion model, respectively. Harpending's raggedness index $r = 0.05$, $p = 0.63$.

Figure 4. Haplotype network representing the evolutionary relationships between mitochondrial haplotypes identified among the *E. merra* samples of the WIO (N = 475). Connecting lines are proportional to the number of mutational steps between haplotypes (the smallest segment on the figure being one mutational step). The size of circles is proportional to number of individuals observed for a given haplotype. Colours of pie charts represent the geographical origin of the haplotype (in terms of four groups: Mozambique Channel, Northern area, East Madagascar and Maldives; see Table 1 for details). The most common haplotypes are numbered.

Figure 5. Map showing the distribution of cytochrome *b* haplotypes among populations of *Epinephelus merra*.

Figure 6. Results from the clustering assignment analysis based on the microsatellite data from *E. merra* collected in the WIO performed using STRUCTURE. The mean posterior assignment probabilities are plotted per individual (upper part of the Figure; one vertical bar per individual) and per locality (lower part of the Figure), with each of the three colors symbolizing each of a clade ($k = 3$).









