To be cited as:
Liu S.-Y. V., Chang F.-T., Borsa P., Chen W.-J., Dai C.-F. (2014) Phylogeography of the humbug damselfish, *Dascyllus aruanus* (Linnaeus, 1758): evidence of Indo-Pacific vicariance and genetic differentiation of peripheral populations. *Biological Journal of the Linnean Society* 113, 931-942

Phylogeography of the humbug damselfish, *Dascyllus aruanus* (Linnaeus, 1758): evidence of Indo-Pacific vicariance and genetic differentiation of peripheral populations

SHANG-YIN VANSON LIU [1,2], FENG-TING CHANG [1], PHILIPPE BORSA [3], WEI-JEN CHEN [1], CHANG-FENG DAI [1*]

[1] Institute of Oceanography, National Taiwan University, Taipei, Taiwan
[2] Department of Ecology and Evolutionary Biology, University of California Los Angeles, U.S.A. [3] Institut de recherche pour le développement c/o Indonesian Biodiversity Research Center, Denpasar, Indonesia
* Author for correspondence: Institute of Oceanography, National Taiwan University, Taipei, Taiwan 10617, Republic of China.

The phylogeographic structure of coral-associated reef fishes may have been severely affected, more than species from deeper habitats, by habitat loss during periods of low sea level. The humbug damselfish, *Dascyllus aruanus*, is widely distributed across the Indo-West Pacific, and exclusively inhabits branching corals. We used mitochondrial cytochrome *b* sequence and seven microsatellite loci on *D. aruanus* samples (260 individuals) from 13 locations across the Indo-West Pacific to investigate its phylogeographic structure distribution-wide. A major genetic partition was found between the Indian and Pacific Ocean populations, which we interpret as the result of geographic isolation on either side of the Indo-Pacific barrier during glacial periods. The peripheral populations of the Red Sea and the Society Islands exhibited lower genetic diversity than, and substantial genetic differences with the other populations, suggesting relative isolation. Thus, vicariance on either side of the Indo-Pacific barrier and peripheral differentiation are thought to be the main drivers that have shaped the phylogeographic patterns presently observed in *D. aruanus*.

ADDITIONAL KEY WORDS: geographic barrier – reef fishes – isolation by distance – mitochondrial DNA – microsatellites.

INTRODUCTION

In the open ocean, geographic barriers shape a species' distribution by either physically restricting its dispersal range or limiting its dispersal success owing to a restriction of suitable habitat. In tropical marine provinces, a few such barriers have been intensively studied. These include the expanses of open ocean that separate the central and the eastern Pacific (the "East Pacific Barrier"), and the eastern and western

Atlantic (the "Mid Atlantic Barrier"), the Amazon–Orinocco outflow that separates the Brazilian and Caribbean provinces (the "Amazon Barrier"), and the cold upwelling area off southwest Africa that separates the tropical Atlantic and Indian oceans (the "Benguela Barrier") (Rocha, Craig & Bowen, 2007). The Indo-Pacific Barrier (Briggs, 1974) breaks the faunal composition of the Indo-West Pacific ensemble into Indian and Pacific provinces. During low sea-level episodes in the Quaternary, the Sunda shelf (i.e., the southeastern extension of the continental shelf of Southeast Asia, including the Malay Peninsula, Sumatra, Borneo, Java, Madura and Bali) repeatedly formed a broad land bridge protruding from Southeast Asia towards the Sahul shelf (i.e., the northwestern extension of the Australian continent, which includes most of the Timor Sea), thus restricting the exchange of marine fauna between the western Pacific and the Indian Ocean (Randall, 1998; Briggs, 1999; Voris, 2000).-This potentially caused genetic differentiation and, in some cases, eventually led to allopatric speciation on either side of the Sunda and Sahul shelves (e.g. Gaither *et al.*, 2010; Leray *et al.*, 2010; Gaither *et al.*, 2011a; Gaither & Rocha, 2013). The positive correlation between divergence time and the degree of geographic overlap in sister species of Indo-West Pacific reef fishes (Quenouille *et al.*, 2011) is further, indirect evidence for allopatric speciation.

Geographic isolation at the margin of the main distribution of a species ('peripheral isolation': Johannesson & André, 2006) is another factor thought to have promoted genetic differentiation and subsequent speciation in reef fishes (Rocha & Bowen, 2008, and references therein; Winters *et al.*, 2010). This process may be accelerated by higher genetic drift due to low effective population sizes (Johannesson & André, 2006; Nunes, Norris & Knowlton, 2009). Speciation by peripheral isolation has been documented at isolated oceanic islands at the margins of ocean basins where the endemism of marine fishes is highest (Allen, 2008). Unique genetic lineages that eventually characterize peripheral populations may be subsequently exported to adjacent, less peripheral regions (Drew & Barber, 2009; Bowen et al., 2013).

We aimed at understanding the mechanisms of geographic differentiation through the distribution-wide phylogeographic survey of an emblematic Indo-West Pacific coral-reef fish, the humbug damselfish, *Dascyllus aruanus*. The pelagic larval duration of *D. aruanus* is 16-24 days (Wellington & Victor, 1989), which is shorter than for most reef-fish species previously studied at similar geographic scales (e.g. Klanten, Choat & van Herwerden, 2007; Horne *et al.*, 2008; Gaither *et al.*, 2010; Winters *et al.*, 2010; Gaither *et al.*, 2011a). This species has been the focus of ecological and genetic investigations aimed at studying habitat preferences and home range (Sale, 1971a,b; Sale, 1972); at determining whether chemical cues trigger settlement in larvae (Sweatman, 1988); at assessing the degree of population genetic differentiation at the local scales (Planes, Bonhomme & Galzin, 1993; Fauvelot and Planes 2002; Raynal et al. 2014); and at estimating the genetic relatedness of individuals within a school (Buston *et al.*, 2009). No study has thus far addressed genetic variation at the scale of the Indo-West Pacific in *D. aruanus*. The mitochondrial phylogeny of species in the genus *Dascyllus* by McCafferty et al. (2002), which included 18 sequences of *D. aruanus* sampled across the Indo-West Pacific, suggested a split between Indian and Pacific haplotypes. This was recently confirmed from control region sequences sampled within the Coral Triangle by Raynal et al. (2014), who showed that the main phylogeographic break in *D. aruanus* is located at the eastern boundary of the Sunda Shelf.

Here, we studied samples collected across the entire distribution of *D. aruanus*, including peripheral populations from the western Indian Ocean and the eastern Central Pacific. Using microsatellite and mitochondrial genetic markers, we tested the effect of the Indo-Pacific barrier by confronting patterns of genetic differentiation in *D. aruanus* against S. Wright's isolation by distance model. We also addressed the possible effect of peripheral isolation on phylogeographic structure. We expect that the present study provide insights into the mechanisms that shape the phylogeographic structure of widely distributed reef fishes in the Indo-West Pacific.

MATERIALS AND METHODS

SAMPLE COLLECTION AND DNA EXTRACTION

In total, 260 *D. aruanus* were collected between 2007 and 2011 from 13 locations across the range of the species, from the Red Sea to the Society Islands (Table 1; Fig. 1). The geographical regions sampled included the Indian Ocean, the western Pacific, and two peripheral regions: the Red Sea and the Society Islands. The Red Sea, which is separated from the Indian Ocean by the 137-m deep Hanish Sill, is considered to be geographically isolated from the Indian Ocean because of the limited oceanic exchange between the two basins during periods of low sea level (Siddall *et al.*, 2003). The Society Islands archipelago is similarly considered as peripheral because of its location at the eastern extremity of the distribution of *D. aruanus*. The Society Islands are part of the eastern Polynesian ensemble which is separated from adjacent archipelagoes to the West (Samoa, Tonga) by a wide tract of ocean with sparse suitable reef habitat. Oceanographic isolation caused by the southerly flowing South Pacific Current (Bonjean & Lagerloef, 2002) may be an effective barrier between the Society Islands and other central West Pacific populations.

Specimens were collected either by scuba diving or by snorkeling using multiple fishing tools including hand-made darts, hand nets, clove oil, and rotenone. Tissue samples (fin clips, or pieces of muscle, or both) were preserved in 95% alcohol and stored at 4°C. Fish that survived capture and fin-clipping were released in the vicinity of the capture site. Individuals too small for fin clipping were anaesthetized with clove oil and preserved whole in 95% ethanol.

Genomic DNA was extracted from tissue fragments using commercial DNA extraction kits (GENO Marine Animals DNA Kit, MP Biomedicals, Singapore; Genomics BioSci & Tech, Taiwan). DNA extracts were diluted in TE buffer and stored at -20°C until amplification by polymerase chain reaction (PCR).

MITOCHONDRIAL DNA MARKER

The mitochondrial DNA cytochrome *b* gene was PCR-amplified using universal primers *GluDG-L* (5'-TGACCTGAARAACCAYCGTTG-3') and *H16460* (5'-CGAYCTTCGGATTACAAGACCG-3') (Palumbi, 1996). PCRs were run in 30 µ*L* reactions containing 10-40 ng template DNA, 3 µL 10X buffer, 0.2 m*M* dNTPs, 1.5 m*M* $MgCl_2$, 10 m*M* each primer, and 0.2 units of *Taq* polymerase (MDbio, Taipei). The thermocycling profile consisted of initial denaturation at 94°C for 2 min, followed by 41 cycles of denaturation at 94°C for 30 s, annealing at 57°C for 30 s, and extension at 72°C for 40 s, and a final extension at 72°C for 2 min. Purification of the PCR products and sequencing reactions were done by Genomics BioSci & Tech (Taiwan). Nucleotide sequences of both forward and reverse strands were determined using an ABI 3730XL automated sequencer (Applied Biosystems, Carlsbad CA, U.S.A.). Individual sequences were inspected visually and corrected based on the chromatograms. Sequences were assembled and edited using the SEQUENCHER version 4.2 software (Gene Code, Ann Arbor MI, U.S.A.). Consensus sequences were exported to the SE-AL program (Rambaut, 1996) to visually inspect all alignments.

MICROSATELLITE MARKERS

Studies of microsatellite variation are useful for evaluating ongoing or recent gene flow (Avise, 2004), at least among populations in recent contact or close proximity (Paetkau *et al.*, 1997). Seven microsatellite loci (*Da304*, *Da360*, *Da494*, *Da523*, *Da589*, *Da590*, and *Da593*) were selected among the 16 published for *D. aruanus* by Fauvelot *et al.* (2009) and used to genotype all individuals of our total sample. We selected

microsatellite loci which were under Hardy-Weinberg equilibrium and with higher levels of polymorphism in a test sample from Moorea, Society Islands (Fauvelot *et al.*, 2009). For each microsatellite marker, PCRs were run in 15 μ*L* reaction mixture containing 10-40 ng template DNA, 1.5 μL 10X buffer, 0.2 m*M* dNTPs, 1.5 m*M* $MgCl_2$, 10 m*M* each primer, and 0.2 unit *Taq* polymerase (MDbio, Taipei). The thermocycling profile consisted of initial denaturation at 94°C for 3 min, followed by 40 cycles of denaturation at 94°C for 30 s, annealing at the appropriate temperature (see Supporting information, Table S1) for 30 s, and extension at 72°C for 30 s, followed by a final extension step at 72°C for 5 min. Genotypes were determined by using an ABI 3730XL automated sequencer (Applied Biosystems) from the purified PCR products (by Genomics BioSci & Tech). Microsatellite fragments were identified and scored using the PEAK SCANNER software (Applied Biosystems).

ANALYSIS OF DATA

Mitochondrial-DNA sequence data were analyzed using ARLEQUIN 3.1 (Excoffier, Laval & Schneider, 2005). Haplotype diversity (*h*), nucleotide diversity (π), and their standard deviations were estimated for each population and for the total sample using Nei's (1987) estimators. Median-joining parsimony analysis was done using NETWORK (Bandelt, Forster & Röhl, 1999) on the nucleotide sequence matrix of 260 individual cytochrome *b* gene haplotypes, using the default settings of the program. In addition, phylogenetic analyses were done to tentatively root the network of mitochondrial haplotypes. Homologous sequences from 7 species in the genus *Dascyllus* [*D. melanurus* (GenBank no. AY208543), *D. flavicaudus* (AY208541), *D. trimaculatus* (AY208544), *D. reticulatus* (AY208545), *D. marginatus* (AY208542) and *D. carneus* (AY208540)] were used as outgroup. A maximum-likelihood (ML) phylogenetic tree was constructed using RAXML 7.2.7 (Stamatakis, 2006), based on the GTR+G model of molecular evolution. This model was selected according to jModelTest (Darriba *et al.*, 2012). A phylogenetic tree based on Bayesian inference (BI) was also constructed, using MRBAYES 3.1.2 (Huelsenbeck & Ronquist, 2001). The analysis was run for 1 million generations with an initial burn-in of 250,000 steps and the convergence diagnostic was applied and the stop probability set to 0·01. The sequential Bayesian stopping rule was used to find the optimal trade-off between reliability and computational effort (Boender & Rinnooy Kan, 1987). The ML and BI trees were viewed and edited in FIGTREE 1.3.1 (Rambaut, 2010). Genetic differentiation between populations was estimated, using θ, Weir & Cockerham's (1984) equivalent of S. Wright's *F*st. A hierarchical analysis of molecular variance (AMOVA) was conducted to quantify genetic differentiation between regions and among populations within regions. Mismatch distributions were computed (1000 replications) to determine whether the number of pairwise differences among DNA sequences reflects exponential population growth (Rogers & Harpending, 1992), which can be used to asses historical population expansion. Harpending's (1994) raggedness index (*r*) was computed to test whether the observed and modelled curves differed.

At microsatellite loci, allele frequencies, mean number of alleles (*Na*), observed (*Ho*) and expected (*He*) heterozygosities, and the inbreeding coefficient (f) were estimated using ARLEQUIN 3.1. An exact test on genotype counts was done to test for significant deviations from Hardy-Weinberg equilibrium, and pairwise estimates of Weir & Cockerham's (1984) θ, were obtained. MICRO-CHECKER 2.2.3 (van Oosterhout *et al.*, 2004) was used to check for the possible occurrence of null alleles and allelic dropout. Allelic richness (*Rs*) was estimated using FSTAT 2.9.3.2 (Goudet, 2001). Principal component analysis (PCA) was run using GENALEX 6.41 (Peakall & Smouse, 2006) to visualize overall geographic structure patterns. STRUCTURE 2.2 program (Pritchard, Stephens & Donnelly, 2000; Falush, Stephens & Pritchard, 2007) was used to infer population structure and assign individuals to clusters based on microsatellite genotype. The program was run using the admixture model that assumes that all individuals are potentially of mixed ancestry, and assigns each individual to a designated population (of *K* potential

populations) with a partial probability. Twenty independent runs, incorporating a burn-in of 1,000,000 Markov chain-Monte Carlo iterations followed by 1,000,000 replicates of data collection, were completed, using default settings, for $K = 2$ to 7. The optimal value of K was determined as the number of clusters where the largest difference in log likelihood (ΔK) was observed (Evanno, Regnaut & Goudet, 2005).

Geographic distance between sampling locations was calculated as the shortest ship distance using the PATH tool implemented in Google Earth (http://www.google.co.uk/intl/en_uk/earth/). The habitat is bi-dimensional and the correlation of genetic distance with geographic distance was analyzed according to the two-dimensional model of isolation by distance (Rousset, 1997). Pairwise genetic distance [$\hat{\theta} / (1-\hat{\theta})$] was plotted against the logarithm of geographic distance, and the correlation between the two distances was tested using the Mantel test implemented in IBDWS v. 3.23 (Jensen, Bohonak & Kelley, 2005). Under two-dimensional isolation by distance, a positive, linear relationship is predicted between genetic distance and the log of geographic distance (Rousset 1997).

RESULTS

GENETIC VARIABILITY

Sixty-three mitochondrial haplotypes (partial cytochrome *b* gene, 1058 bp) were scored in the total sample ($N = 260$). The haplotype sequences were deposited in GenBank under accession nos. KF754733-KF754795. Nucleotide diversity in *D. aruanus*, estimated from the total sample, was $\pi = 0.005 \pm 0.003$, while haplotype diversity was $h = 0.907 \pm 0.011$. The nucleotide diversity (π) ranged from 0.001 ± 0.001 to 0.005 ± 0.003 across populations, and haplotype diversity (h) ranged from 0.423 ± 0.119 to 1.000 ± 0.127 (Table 1). From all populations sampled in the Indian Ocean, the peripheral Red Sea population had the lowest level of genetic diversity. A similar observation was made for the Society Islands population in the Pacific Ocean.

The number of alleles (Na), expected heterozygosity (He), and observed heterozygosity (Ho) for microsatellite loci ranged from 6.4 to 24.7 (mean = 17.0), 0.895 to 0.943 (mean = 0.918), and 0.617 to 0.835 (mean = 0.716), respectively. Allelic richness (Rs) had the same magnitude across the distribution of *D. aruanus*, with a mean of 6.3 alleles per locus. There were no significant departures of genotypic proportions from Hardy-Weinberg equilibrium. Pairwise θ ranged from 0.105 to 0.356 ($P < 0.05$) across loci. No evidence of null alleles was detected at any locus (Table S1).

PHYLOGENETIC ANALYSES

The Median-Joining network of mitochondrial haplotypes (Fig. 2) was split into two main haplogroups separated by four mutational steps. One haplogroup exclusively included all haplotypes sampled in the Indian Ocean. The other haplogroup included all haplotypes from the Pacific Ocean. Most haplotypes of the Red Sea formed a lineage nested within the "Indian" haplogroup. Haplotypes from the Red Sea population were all private, and so were four out of six haplotypes from the Society Islands.

The ML and BI trees of haplotypes showed similar topologies (Supporting information, Fig. S1). The two separate haplogroups, here referred to as the Pacific and Indian Ocean haplogroups, were confirmed by phylogenetic analysis. Haplotypes from the Indian Ocean formed a sub-clade topologically distinct from the Pacific haplogroup.

POPULATION STRUCTURE

Pairwise θ estimates at the cytochrome *b* locus ranged from -0.080 to 0.858 (Supporting information, Table S2); pairwise multiple-locus θ estimates (averages over 7 microsatellite loci) ranged from -0.001 to 0.078 (Table S3). In comparisons between these two data sets, the values derived from microsatellite were

lower than from mitochondrial DNA.

Based on microsatellite markers, the number of population clusters was determined as $K = 3$ (Fig. 3A) for which the likelihood value [ln$P$ (3) = -1196.1] was highest. The three major groups included the populations of the Indian Ocean, the Society Islands, and all other Pacific populations. Although the Red Sea sample was clustered with the other Indian-Ocean samples, its detailed genetic composition differed from the latter (Fig. 3A). The PCA (Fig. 3B) showed a major partition between Indian and Pacific Ocean populations, and additional partition within the Pacific Ocean by singling out the Society Islands sample. The Red Sea population grouped with other Indian Ocean populations.

Pairwise genetic distances estimated from mitochondrial sequences were plotted against geographic distances (Fig. 4A). Within-ocean data points, excluding those comparisons involving either the Red Sea or Society Islands populations (black circles on Fig. 4A), all had an ordinate < 0.2 and were distributed in a linear fashion as expected under the bi-dimensional model of isolation by distance (Rousset, 1997). The trend of increasing genetic distance with increasing geographic distance was materialized by a regression line on Fig. 4A [Mantel's test: $P = 0.154$ (within-Indian Ocean comparisons) and $P = 0.115$ (within-Pacific Ocean comparisons); Fisher's combined probability (Sokal & Rohlf, 1995): $P = 0.087$]. Inter-oceanic data points all had an ordinate > 0.5 and were positioned clearly above the regression line. Intra-oceanic comparisons involving either the Red Sea population or the Society Islands population (light-grey circles on Fig. 4A) occupied an intermediate position on the graph. Likewise, pairwise genetic distances estimated from microsatellite genotypes were plotted against geographic distances (Fig. 4B). Within-ocean comparisons, excluding those comparisons involving either the Red Sea or Society Islands populations were similarly distributed in a linear fashion. A trend of increasing genetic distance with increasing geographic distance was similarly uncovered [Mantel's test: $P = 0.285$ (within-Indian Ocean comparisons) and $P = 0.154$ (within-Pacific Ocean comparisons); Fisher's combined probability (Sokal & Rohlf, 1995): $P = 0.181$]. Fig. 4B (microsatellites) was very similar to Fig. 4A (mitochondrial haplotypes), except that the range of genetic distances and the slope of the linear relationship were one order of magnitude lower. Thus, for both types of markers, a trend where genetic distance linearly increased with geographic distance was observed at the within-ocean scale, but not distribution-wide.

The AMOVA results (Table 2) indicated significant population structure when four groups were defined a priori (Red Sea, the Mozambique Channel, Society Islands, and the rest of Pacific populations) based on both mitochondrial and microsatellite data (respectively: Φct = 0.655, $P < 0.001$; and Φct = 0.030, $P < 0.001$). Significant structure was also present when considering the Indian Ocean vs. Pacific Ocean partition, as well as when comparing the Indian Ocean, the Pacific Ocean, and the Society Islands. For the mitochondrial data, most of the variance in the above three groupings resided among groups. Based on the microsatellite data, the total variance in the foregoing three groupings resided within populations.

POPULATION EXPANSION

Four groups were considered according to population structuring and used to do the mismatch distribution test: these were the Red Sea, Indian Ocean, Pacific Ocean, and Society Islands populations. A unimodal pattern was uncovered for the Indian Ocean population (Supporting information, Fig. S2B). In contrast, the mismatch distributions for the Red Sea, Pacific Ocean, and Society Islands populations were multimodal (Fig. S2A, C and D). The results of goodness of fit (Red Sea: $P = 0.295$; Pacific Ocean: $P = 0.201$) and Harpending's raggedness (Red Sea: P=0.526; Pacific Ocean: $P = 0.209$) tests showed that only the Red Sea and Pacific Ocean populations fit a model of population expansion.

DISCUSSION

Our results support a single major genetic break in *Dascyllus aruanus* on either side of the Indo-Pacific barrier. Furthermore, population structure analysis indicated substantial genetic differentiation between the populations of the Red Sea and the Western Indian Ocean as well as between the Society Islands population and other Pacific populations, indicating restricted gene flow at the margins of the geographical distribution of the humbug damselfish.

POPULATION STRUCTURE OF INDO-WEST PACIFIC REEF FISHES

Through the Pleistocene, episodes of low sea level at the apogee of glacial cycles eliminated local populations on the continental Sunda and Sahul shelves, thus facilitating genetic divergence between reef-fish populations from the Indian and Pacific Oceans. Meanwhile, land bridge formations restricted the oceanographic flow from the Pacific Ocean to the Indian Ocean, hampering the larval connectivity between them (Voris, 2000). Morphologically distinct sister-species pairs of reef fishes on either side of the Indo-Pacific barrier have been documented (Randall, 1998). Estimated divergence times for 17 fish species pairs across the Indo-Pacific barrier range from 0.3 and 6.6 Ma (Gaither & Rocha, 2013). Examples of marked Indo-West Pacific genetic partition at the infra-specific level in reef fishes include the bigscale soldierfish *Myripristis berndti* (Craig et al., 2007), the Coral trout *Plectropomus leopardus* (van Herwerden et al., 2009), the narrow-barred Spanish mackerel *Scomberomorus commerson* (Fauvelot and Borsa, 2011), the blue spotted grouper *Cephalopholis argus* (Gaither *et al.*, 2011a), and the honeycomb grouper *Epinephelus merra* (Muths,Teissier & Bourjea, 2014). This has been interpreted as an effect of vicariance on either side of the Indo-Pacific barrier, enhanced by various factors such as self- recruitment, philopatry, or relatively short larval duration. The level of distribution-wide genetic differentiation in these species (see Fauvelot & Borsa, 2011 and references therein) is comparable to that reported here for *D. aruanus*. In contrast, phylogeographic structure in *N. brevirostris* (Horne *et al.*, 2008) and *N. vlamingii* (Klanten, Choat & van Herwerden, 2007) suggests that the Indo-Pacific barrier has had an impact in the past but since the rise of sea level the initial signal of genetic isolation has been partly lost. Population structure of fishes with substantially higher dispersal ability than *D. aruanus*, e.g. *Lutjanus kasmira* (Gaither *et al.*, 2010), *Gymnothorax undulatus*, *G. flavimarginatus* (Reece *et al.*, 2010), *Naso unicornis* (Horne *et al.*, 2008), do not demonstrate a strong effect of the Indo-Pacific barrier.

Thus, species-specific differences in life history traits and in demography may result in different phylogeographical responses in shared environments. Throughout its range, *D. aruanus* shows close association with acroporid or pocilloporid corals; adults are sedentary (Sale, 1972; Holbrook, Forrester & Schmitt 2000). *D. aruanus* larvae have a pelagic life for 16–24 days prior to settlement on reefs (Wellington & Victor, 1989). They preferentially recruit near settled congeners by using chemical cues (Sweatman, 1985, 1988). This allows the pre-recruits to recognize their natal areas, thus potentially enhancing self-recruitment, hence hampering dispersal. These life history characteristics indicate that late *D. aruanus* larvae likely do not passively drift in surface currents, hence larvae may not be dispersed over distances as long as their pelagic duration would potentially allow.

PERIPHERAL DIFFERENTIATION IN THE RED SEA AND IN THE SOCIETY ISLANDS

Genetic diversity indices, including haplotype diversity and nucleotide diversity, were lower in the *D. aruanus* population from the Red Sea than in the other populations from the Indian Ocean, and a similar pattern was found in the Pacific Ocean where the Society Islands population harboured the lowest genetic diversity. In addition, all Red-Sea and the majority of Society-Islands haplotypes were private. The levels of genetic differentiation indicated low genetic exchange between these populations and the remaining populations in their respective ocean basins. To summarize, the lower levels of genetic diversity and

higher levels of genetic differentiation observed in these populations relative to the other populations from the same ocean, which we ascribe to increased genetic drift, fit the model of peripheral isolation.

Lower genetic diversity is commonly found in fish populations of geographically isolated peripheral environments such as Hawai'i (Eble *et al.*, 2011; Gaither *et al.*, 2011b). In this study, the Red Sea and the Society Islands populations were considered as peripheral because they are located at the boundaries of the range of *D. aruanus* and are geographically isolated from the main area of distribution. DiBattista *et al.* (2013) reported strong genetic differences between the Red Sea and Indian Ocean in five out of the seven reef fish species they studied, according with the hypothesis of repeated isolation during glacial episodes. A similar genetic break was here observed in *D. aruanus*, where no haplotype was shared between the Red Sea and the Indian Ocean populations. Similarly, several studies have shown reduced genetic diversity in French Polynesia's reef fishes (Gaither *et al.*, 2010, 2011a; Messmer *et al.*, 2012), a phenomenon assumed to be an effect of geographic isolation. This is also the case with *D. aruanus*.

GENETIC DIVERSITY AND POPULATION EXPANSION

The low nucleotide diversity and high haplotype diversity found in *D. aruanus* populations is commonly observed in other marine fishes where it has been interpreted as a signal of historical population expansion (Grant & Bowen, 1998). The results of the mis-match distribution analyses provide evidence that population expansion has occurred in the Red Sea and Pacific Ocean populations.

Population expansion after glacial periods has been previously inferred for several reef fishes (Horne *et al.*, 2008; Reece *et al.*, 2010). This phenomenon could result in shallow sequence divergence over wide areas (Liu *et al.*, 2012), as we found among Pacific haplotypes of *D. aruanus*. Moreover, the multiple salinity crises and sea level fluctuations in the Red Sea have affected the population structure and connectivity of reef fishes between the Indian Ocean and Red Sea (DiBattista *et al.*, 2013). In contrast, the Indian Ocean and Society Islands populations did not fit the expansion model, and showed no star-like network.

CONCLUSION

At the within-ocean scale, peripheral populations excluded, genetic differentiation between local *D. aruanus* populations slightly increased linearly with geographic distance, a property expected from S. Wright's isolation by distance model. Strong genetic differences were observed between populations from the Indian and Pacific oceans at both mitochondrial and microsatellite loci. This was interpreted as vicariance on either side of the Indo-Pacific barrier. Further analysis indicated that the Society Islands population is separated from the rest of the Pacific populations as a third evolutionary significant unit, presumably as a consequence of the geographic isolation of the Society Islands. Both the Red Sea and Society Islands populations have a genetic composition distinct from the other populations in the same ocean, which can be explained by genetic drift in a context of low gene flow due to their peripheral location and possible local adaptation. Overall, population structure of *D. aruanus* has been shaped by the combination of land barriers, isolation by distance, and by relative isolation of peripheral populations. These mechanisms affecting phylogeographic structure might be representative of a large proportion of Indo-West Pacific reef fishes.

ACKNOWLEDGMENTS

The sample from Raja Ampat was collected during expedition EWiN (LIPI, Indonesia) of RV *Baruna Jaya VIII* to West Papua, November 2007; the sample from New Caledonia was collected during the RESICOD workshop (IRD, New Caledonia), December 2010; the samples from Iles Eparses (Europa,

Juan de Nova,Glorieuses) were collected during the TAAF /INEE-sponsored PHYLIP expedition, April 2011. Field work was done under Surat Izin Penelitihan no. 5936/SU/KS/2007 issued by LIPI-P2O (West Papua), and under MLD/OD no. 19750 (West Papua), OD no. 84839 (New Caledonia), OD no. 95664 (Iles Eparses) and OD no. 106543 (Philippines, November 2011), issued by IRD's director general. Fish sampling in New Caledonia was in conformity with Article 7 of Deliberation no. 06-2009 (18 February 2009) concerning the collection and the exploitation of genetic resources in New Caledonia's Province Sud. Fish sampling at Iles Eparses was authorized under Arrêté no. 2011-32 of Terres australes et antarctiques françaises. Funded pro parte by FRB, CNRS-INEE, IRD, LIPI, and TAAF; the funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. We thank S. Arakaki, R. Beldade, B. Frédérich, J.-D. Durand, Husni, M. Kiflawi, M. Kulbicki, G. Mou-Tham, J. O'Donnell, K. Sakai, K.-N. Shen, M.-H. Tsai and J.-H. Yang for help in sampling in the Red Sea, Iles Eparses, Madagascar, the South China Sea, the Philippines, West Papua, New Caledonia, and the Society Islands. We would like to thank L. Sorenson and B. Frédérich for assisting in English proofreading and for comments on the manuscript.

REFERENCES

Allen GR. 2008. Conservation hotspots of biodiversity and endemism for Indo-Pacific coral reef fishes. *Aquatic Conservation: Marine and Freshwater Ecosystems* **18:** 541–556

Avise JC. 2004. Molecular Markers, Natural History, and Evolution (2nd edn.). Sinauer, Sunderland, MA.

Bandelt HJ, Forster P, Röhl A. 1999. Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution* **16:** 37–48.

Bonjean F, Lagerloef GSE. 2002. Diagnostic model and analysis of the surface currents in the tropical Pacific Ocean. *Journal of Physical Oceanography* **32:** 2938-2954.

Boender CGE, Rinnooy Kan AHG. 1987. Bayesian stopping rules for multistart global optimization methods. *Mathematical Programming* **37:** 59-80.

Bowen BW, Rocha LA, Toonen RJ, Karl SA, the ToBo Laboratory. 2013. The origins of tropical marine biodiversity. *Trends in Ecology and Evolution* **61:** 359-366.

Briggs JC. 1974. Marine zoogeography. New York: McGraw-Hill, 475 p.

Briggs JC. 1999. Extinction and replacement in the Indo-West Pacific Ocean. *Journal of Biogeography* **26:** 777–783.

Buston PM, Fauvelot C, Wong MYL, Planes S. 2009. Genetic relatedness in groups of the humbug damselfish *Dascyllusaruanus*: small, similar-sized individuals may be close kin. *Molecular Ecology* **18:** 4707–4715.

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.

Darriba D, Taboada GL, Doallo R, Posada D. 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods* **9:** 772.

DiBattista JD, Berumen ML, Gaither MR, Rocha LA, Eble JA, Choat JH, Craig MT, Skillings DJ, Bowen BW. 2013. After continents divide: comparative phylogeography of reef fishes from the Red Sea and Indian Ocean. *Journal of Biogeography* **40:** 1170-1181.

Drew J, Barber PH. 2009. Sequential cladogenesis of the reef fish *Pomacentrus moluccensis* (Pomacentridae) supports the peripheral origin of marine biodiversity in the Indo-Australian archipelago. *Molecular Phylogenetics and Evolution* **53:** 335–339.

Eble JA, Rocha LA, Craig MT, Bowen BW. 2011. Not all larvae stay close to home: insights into

marine population connectivity with a focus on the brown surgeonfish (*Acanthurus nigrofuscus*). *Journal of Marine Biology* **2011:** 518516.

Evanno G, Regnaut S, Goudet J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* **14:** 2611-2620.

Excoffier L, Laval G, Schneider S. 2005. Arlequin ver. 3.0: an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* **1:** 47−50

Falush D, Stephens M, Pritchard J. 2007. Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Molecular Ecology Resources* **7:** 574–578.

Fauvelot C, Borsa P. 2011. Patterns of genetic isolation in narrow-barred Spanish mackerel (*Scomberomorus commerson*) across the Indo-West Pacific. *Biological Journal of the Linnean Society* **104:** 886–902.

Fauvelot C, Planes S. 2002. Understanding origins of present-day genetic structure in marine fish: biologically or historically driven patterns? *Marine Biology* **141:** 773–788.

Fauvelot C, Smith-Keune C, Jerry DR, Buston PM, Planes S. 2009. Isolation and characterization of 16 microsatellite loci in the humbug damselfish, *Dascyllus aruanus* (family Pomacentridae). *Molecular Ecology Resources* **9:** 651–653.

Froese R, Pauly D. (eds.) 2012. FishBase. World Wide Web electronic publication (www.fishbase.org, version 07/2012).

Gaither MR, Bowen BW, Bordenave TR, Rocha LA, Newman SJ, Gomez JA, van Herwerden L, Craig MT. 2011a. Phylogeography of the reef fish *Cephalopholis argus* (Epinephelidae) indicates Pleistocene isolation across the Indo-Pacific Barrier with contemporary overlap in the Coral Triangle. *BMC Evolutionary Biology* **11:** 189.

Gaither MR, Jones SA, Kelley C, Newman SJ, Sorenson L, Bowen BW. 2011b. High connectivity in the deepwater snapper *Pristipomoides filamentosus* (Lutjanidae) across the Indo-Pacific with isolation of the Hawaiian archipelago. *PLoS ONE* **6:** e28913

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.

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: www2.unil.ch/popgen/softwares/fstat.htm.

Grant WS, Bowen BW. 1998. Shallow population histories in deep evolutionary lineages of marine fishes: insights from sardines and anchovies and lessons for conservation. *Journal of Heredity* **89:** 415-426.

Harpending RC 1994. Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. *Human Biology* **66:** 591-600.

Holbrook SJ, Forrester GE, Schmitt RJ. 2000. Spatial patterns in abundance of damselfish reflects availability of suitable habitat. *Oecologia* **122:** 109–120.

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. *Molecular Phylogenetics and Evolution* **49:** 629-638.

Huelsenbeck JP, Ronquist F. 2001. MRBAYES: Bayesian inference of phylogeny. *Bioinformatics* **17:** 754-755.

Jensen JL, Bohonak AJ, Kelley ST. 2005. Isolation by distance, web service. *BMC Genetics* **6:** 13.

Johannesson K, Andre C. 2006. Life on the margin: genetic isolation and diversity loss in a peripheral marine ecosystem, the Baltic Sea. *Molecular Ecology* **15:** 2013-2029.

Klanten OS, Choat JH, van Herwerden L. 2007. Extreme genetic diversity and temporal rather than spatial partitioning in a widely distributed coral reef fish. *Marine Biology* **150:** 659–670.

Leray M, Beldade R, Holbrook SJ, Schmitt RJ, Planes S, Bernardi G. 2010. Allopatric divergence and speciation in coral reef fish: the three-spot dascyllus, *Dascyllus trimaculatus*, species complex. *Evolution* **64:** 1218–1230.

Liu SYV, Dai CF, Allen GR, Erdmann MV. 2012. Phylogeography of the neon damselfish *Pomacentrus coelestis* indicates a cryptic species and different species origins in the West Pacific Ocean. *Marine Ecology Progress Series* **458:** 155-167

McCafferty S, Bermingham E, Quenouille B, Planes S, Hoelzer G, Asoh K. 2002. Historical biogeography and molecular systematics of the Indo-Pacific genus *Dascyllus* (Teleostei: Pomacentridae). *Molecular Ecology* **11:** 1377–1392.

Messmer V, Jones GP, Munday PL, Planes S. 2012. Concordance between genetic and species diversity in coral reef fishes across the Pacific Ocean biodiversity gradient. *Evolution* **66:** 3902-3917.

Muths D, Teissier E, Bourjea J. 2014. Genetic structure of the reef grouper *Epinephelus merra* in the West Indian Ocean appears congruent with biogeographic and oceanographic boundaries. *Marine Ecology* **2014:** 1-15.

Nei M. 1987. Molecular evolutionary genetics. Columbia University Press, New York, NY.

Nunes F, Norris RD, Knowlton N. 2009. Implications of isolation and low genetic diversity in peripheral populations of an amphi-Atlantic coral. *Molecular Ecology* **18:** 4283-4297.

Paetkau D, Waits IP, Clarkson PL, Craighead I, Strobeck C. 1997. An empirical evaluation of genetic distance statistics using microsatellite data from bear (Ursidae) populations. *Genetics* **147:** 1943–1957.

Palumbi SR. 1996. Nucleic acids II: the polymerase chain reaction. In: Hillis DM, Moritz C, Mable BK, eds. Molecular Systematics. Sunderland MA: Sinauer, 205-247.

Peakall R, Smouse PE. 2006. GenAlEx 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* **6:** 288-295.

Planes S, Bonhomme F, Galzin R. 1993. Genetic structure of *Dascyllus aruanus* populations in French Polynesia. *Marine Biology* **117:** 665-674.

Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics* **155:** 945-959.

Quenouille B, Hubert N, Bermingham E, Planes S. 2011. Speciation in tropical seas: allopatry followed by range change. *Molecular Phylogenetics and Evolution* **58:** 546–552.

Rambaut A. 1996. Se-Al: Sequence Alignment Editor ver. 2.0 (program distributed by the author, Department of Zoology, University of Oxford, Oxford. <http:// tree.bio.ed.ac.uk/software/seal/>).

Rambaut A. 2010. FigTree 1.3.1. Available from: http://tree.bio.ed.ac.uk/software/figtree/.

Randall JE. 1998. Zoogeography of shore fishes of the Indo-Pacific region. *Zoological Studies* **37:** 227–268.

Raynal JM, Crandall ED, Barber PH, Mahardika GN, Lagman MC, Carpenter KE. 2014. Basin isolation and oceanographic features influencing lineage divergence in the humbug damselfish (*Dascyllus aruanus*) in the Coral Triangle. *Bulletin of Marine Science* **90:** 53-532.

Reece JS, Bowen BW, Joshi K, Goz V, Larson AF. 2010. Phylogeography of two moray eels indicates high dispersal throughout the Indo-Pacific. *Journal of Heredity* **101:** 391–402.

Rocha LA, Bowen BW. 2008. Speciation in coral-reef fishes. *Journal of Fish Biology* **72:** 1101-1121.

Rocha LA, Craig MT, Bowen BW. 2007. Phylogeography and the conservation of coral reef fishes. *Coral Reefs* **26:** 501–512.

Rogers AR, Harpending H. 1992. Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution* **9:** 552-569.

Rousset F. 1997. Genetic differentiation and estimation of gene flow from *F*-statistics under isolation by

distance. *Genetics* **145:** 1219-1228.

Sale PF. 1971a. Apparent effect of prior experience on habitat preference exhibited by the reef fish, *Dascyllus aruanus* (Pisces, Pomacentridae). *Animal Behaviour* **19:** 251-256.

Sale PF. 1971b. Extremely limited home range in a coral reef fish, *Dascyllus aruanus* (Pisces, Pomacentridae). *Copeia* **1971:** 324-327.

Sale PF. 1972. Influence of corals in the dispersion of the pomacentrid fish, *Dascyllus aruanus*. *Ecology* **53:** 741-744.

Siddall M, Rohling EJ, Almogi-Labin A, Hemleben C, Meischner D, Schmelzer I, Smeed DA. 2003. Sea level fluctuations during the last glacial cycle. *Nature* **423:** 853–858.

Sokal RR, Rohlf FJ. 1995. Biometry, The Principles and Practice of Statistics in Biological Research, 3rd edn. Freeman and Co, New York.

Stamatakis A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22:** 2688-2690.

Sweatman HPA. 1985. The influence of adults of some coral reef fishes on larval recruitment. *Ecological Monographs* **55:** 469–485.

Sweatman HPA. 1988. Field evidence that settling coral reef fish larvae detect resident fishes using dissolved chemical cues. *Journal of Experimental Marine Biology and Ecology* **124:** 163–174.

van Herwerden L, Choat JH, Newman SJ, Leray M, Hillersøy G. 2009. Complex patterns of population structure and recruitment of *Plectropomus leopardus* (Pisces: Epinephelidae) in the Indo-West Pacific: implications for fisheries management. *Marine Biology* **156:** 1595-1607.

van Oosterhout C, Hutchinson W, Wills D, Shipley P. 2004. MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* **4:** 535-538.

Voris HK. 2000. Maps of Pleistocene sea levels in Southeast Asia: shorelines, river systems and time durations. *Journal of Biogeography* **27:** 1153–1167.

Weir BS, Cockerham CC. 1984. Estimating *F*-statistics for the analysis of population structure. *Evolution* **38:** 1358-1370.

Wellington GM, Victor BC. 1989. Planktonic larval duration of one hundred species of Pacific and Atlantic damselfishes (Pomacentridae). *Marine Biology* **101:** 557-567.

Winters KL, van Herwerden L, Choat JH, Robertson DR. 2010. Phylogeography of the Indo-Pacific parrotfish *Scarus psittacus*: isolation generates distinctive peripheral populations in two oceans. *Marine Biology* **157:** 1679–1691.

Table 1. Sampling locations and diversity indices based on individual cytochrome *b* gene sequences and multiple-locus microsatellite genotypes in 13 populations of *Dascyllus aruanus* from Indo-Pacific Ocean. *N*, sample size; *mtDNA*: mitochondrial DNA; *msat*: microsatellites; *nh*: number of haplotypes, *h*: haplotype diversity, π: nucleotide diversity; *Rs*: average allelic richness per locus

Location	Abbreviation	*N*	Diversity parameter			
			mt DNA			msat
			nh	*h* ± SD	π ± SD	*Rs*
Indian Ocean						
Gulf of Aqaba, Red Sea	RS	15	7	0.800 ± 0.083	0.002 ± 0.001	5.788
Europa, Mozambique Channel	EU	15	9	0.933 ± 0.040	0.003 ± 0.002	6.521
Juan de Nova, Mozambique Channel	JN	5	4	0.900 ± 0.161	0.003 ± 0.002	6.278
Toliara, Madagascar	MD	12	9	0.939 ± 0.058	0.002 ± 0.001	6.582
Glorieuses Is., Mozambique Channel	GL	5	5	1.000 ± 0.127	0.005 ± 0.003	5.587
Pacific Ocean						
Paracels Islands, South China Sea	PI	12	5	0.742 ± 0.116	0.002 ± 0.001	6.085
Dongsha, South China Sea	DS	16	8	0.825 ± 0.076	0.002 ± 0.001	6.602
Houbihu, Taiwan	TW	27	8	0.675 ± 0.091	0.002 ± 0.001	6.541
Caohagan, Lapu-Lapu, Philippines	LL	32	13	0.815 ± 0.061	0.002 ± 0.001	6.413
Sesoko Island, Japan	SK	25	7	0.767 ± 0.056	0.002 ± 0.002	6.396
Raja Ampat, West Papua	RA	33	10	0.763 ± 0.054	0.002 ± 0.001	6.814
SW Lagoon, New Caledonia	NC	38	13	0.852 ± 0.035	0.004 ± 0.002	6.473
Moorea Island, Society Islands	SI	25	5	0.423 ± 0.119	0.001 ± 0.001	5.871

Table 2. *Dascyllus aruanus*. Analysis of molecular variance (AMOVA) based on mitochondrial DNA haplotype sequences (*mtDNA*) and genotypes at 7 microsatellite loci (*microsatellites*), according to 4 distinct models for the geographic partition of populations. Φsc summarizes variation among populations within groups corresponding to geographical regions; Φct summarizes variation among geographical regions, and Φst summarizes the remaining variation within populations. In the 'No partition' model, Φst summarizes variation among populations without regard to hierarchical structure. *d.f.* degrees of freedom; *Indian Ocean samples* include RS, EU, JN, MD, and GL. *Pacific Ocean samples* include PI, DS, TW, LL, SK, RA, NC, and SI. See Table 1 for sample abbreviations

Partition model,	Genetic marker					
Source of variation	mt DNA			microsatellites		
	d.f.	% total variance	Φ statistic	d.f.	% total variance	Φstatistic
Indian vs Pacific:						
Among groups	1	65.2	Φct = 0.652 *	1	2.9	Φct = 0.029 *
Among populations within groups	11	7.4	Φsc = 0.212 *	11	2.1	Φsc = 0.022 *
Within populations	247	27.4	Φst = 0.726 *	507	95.0	Φst = 0.050 *
Indian vs Pacific vs SI:						
Among groups	2	64.16	Φct = 0.642 *	2	2.56	Φct = 0.026 *
Among populations within groups	10	3.55	Φsc = 0.099 *	10	1.89	Φsc = 0.019 *
Within populations	247	32.29	Φst = 0.677 *	507	95.55	Φst = 0.045 *
RS vs Indian vs Pacific vs SI:						
Among groups	3	65.5	Φct = 0.655 *	3	3.0	Φct = 0.030 *
Among populations within groups	9	2.0	Φsc = 0.058 *	9	1.6	Φsc = 0.017 *
Within populations	247	32.6	Φst = 0.675 *	507	95.4	Φst = 0.047 *
No partition						
Among populations	12	52.7	Φst = 0.527 *	12	3.2	Φst = 0.032 *

* $P < 0.05$; probability of observed value under null hypothesis ($\Phi=0$) was estimated from 1000 random permutations

Captions to figures

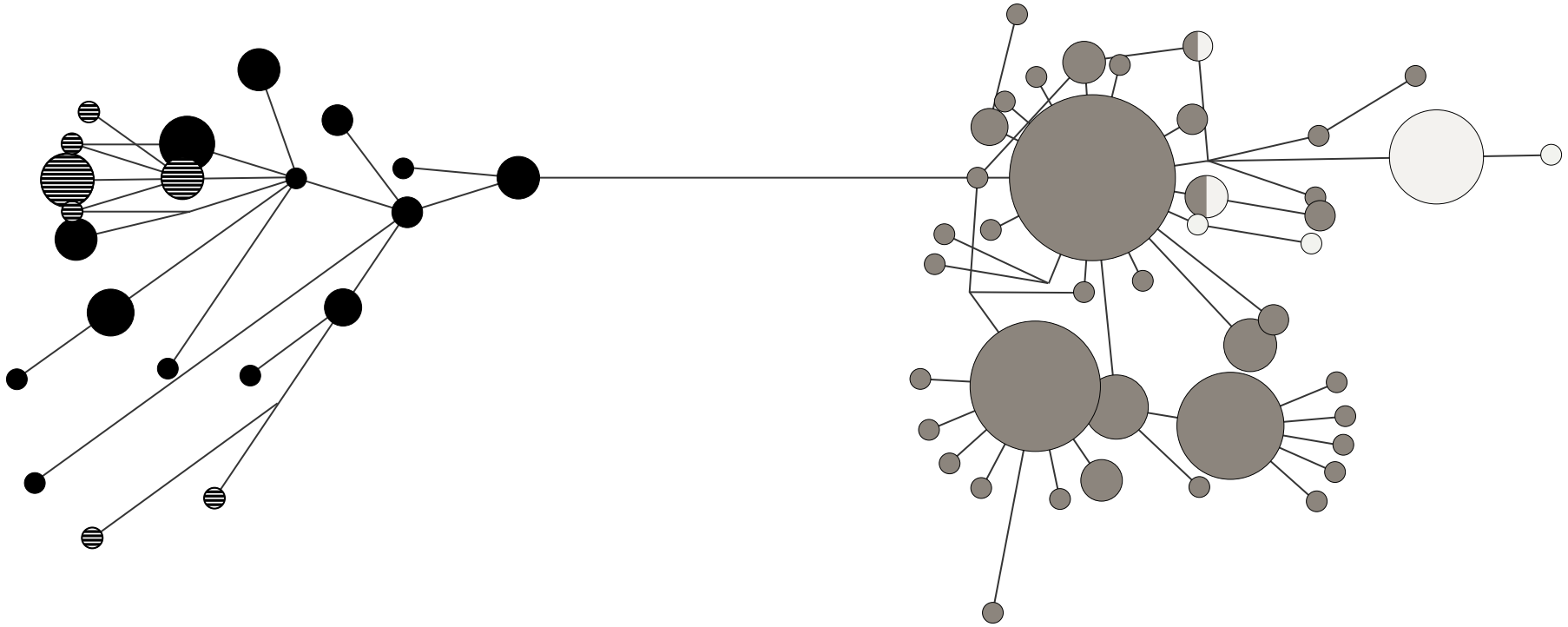
Figure 1. Sampling locations for humbug damselfish, *Dascyllus aruanus*. From West to East: *RS*, Gulf of Aqaba, Red Sea; *EU*, Europa Island, Mozambique Channel; *JN*, Juan de Nova Island, Mozambique Channel; *MD*, Toliara, Madagascar; *GL*, Glorieuses Islands, Mozambique Channel; *PI*, Paracels Islands, South China Sea; *DS*, Dongsha Island, South China Sea; *TW*, Houbihu, Taiwan; *LL*, Lapu-Lapu Island, Philippines; *SK*, Sesoko Island; *RA*, Raja Ampat, West Papua; *NC*, New Caledonia; *SI*, Moorea, Society Islands. Sampling details in Table 1; *grey-tone code* in accordance with the distinction of four groups of mitochondrial haplotypes (see Results). *Background-grey shading* indicates the geographic distribution of *D. aruanus* [from the point map presented in FishBase (Froese & Pauly, 2012); accessed 30 July 2012]; *dotted line* symbolizes the Indo-Pacific barrier (Rocha, Craig & Bowen, 2007).

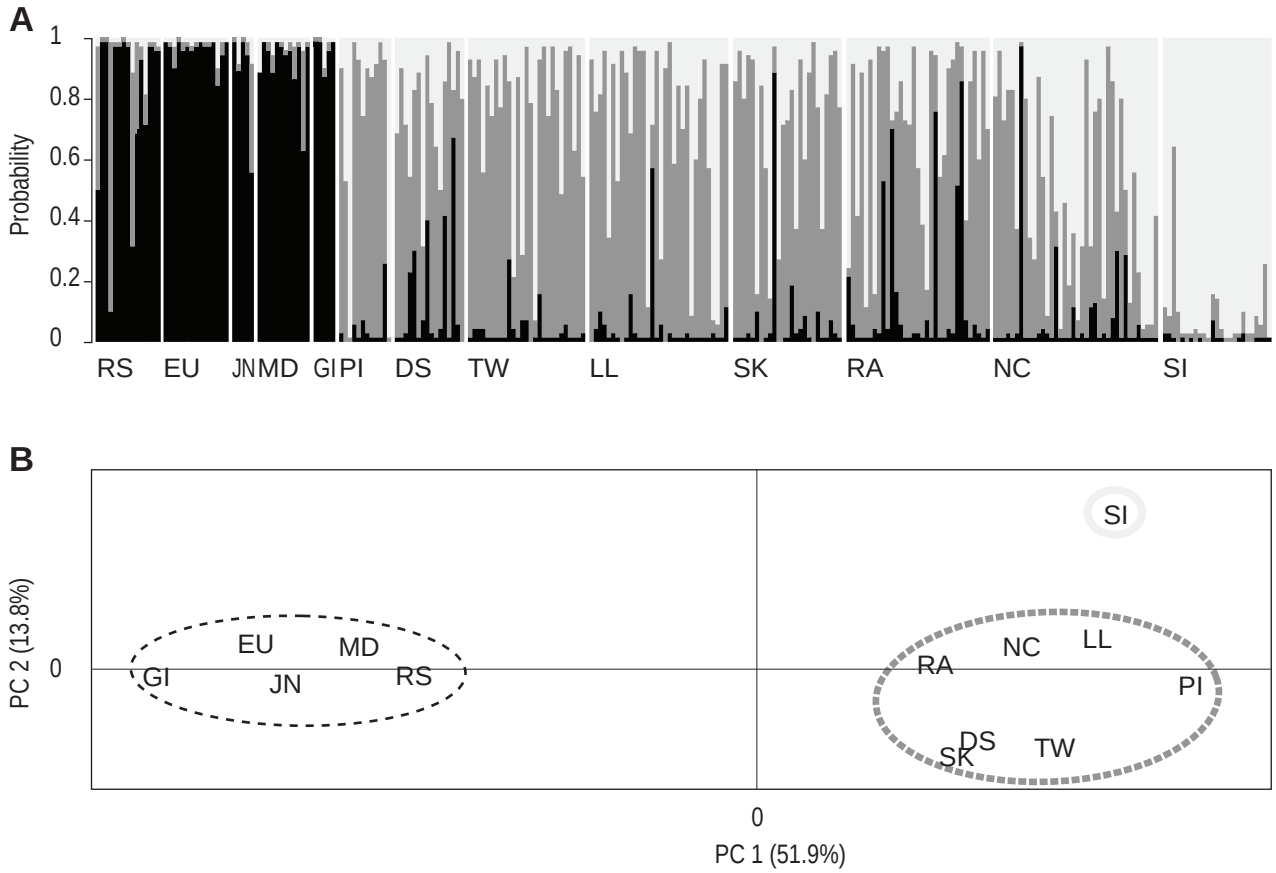
Figure 2. *Dascyllus aruanus*. Median-joining network for 260 cytochrome *b* gene sequences (1040 bp). Area of circle proportional to number of individuals with a given haplotype; *grey tone code* according to the region where the haplotype was sampled (see Fig. 1).

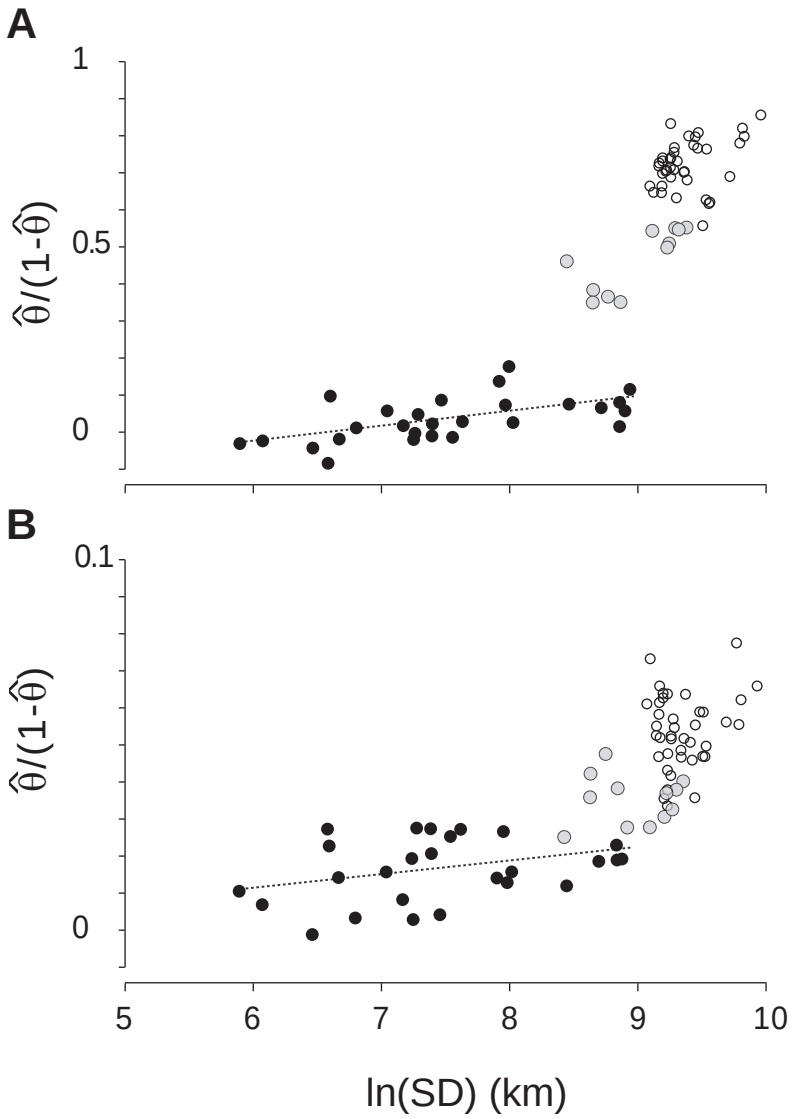
Figure 3. *Dascyllus aruanus*. Sample abbreviations as in Table 1. A. Outcome of Bayesian structure analysis on individuals defined by their genotype at seven microsatellite loci when K=3. B. Principal component analysis on samples defined by their allele frequencies at seven microsatellite loci. Samples represented by their abbreviations; ellipses delineate the three clusters identifed from Bayesian structure analysis, using the colour codes that characterize individuals in Fig. 3A.

Figure 4. Population genetic structure in *Dascyllus aruanus*. Pairwise estimates of genetic differentiation [$\hat{\theta}/(1-\hat{\theta})$; ordinates] plotted against the logarithm of ship distance [ln(SD); abscissa]. Pairwise θ were estimated from genotypic data according to Weir & Cockerham (1984). *Black circles*: intra-oceanic comparisons; *open circles*: inter-oceanic comparisons; *grey circles*: intra-oceanic comparisons involving samples from either the Red Sea (RS) or the Society Islands (SI); *dotted line*: regression of genetic distance against ship distance for intra-oceanic comparisons, to the exclusion of all comparisons involving samples RS or SI. A. Mitochondrial DNA data; regression line: $\hat{\theta}\cdot(1-\hat{\theta}) = 0.039\cdot\ln(SD)-0.257$ (Pearson's correlation coefficient, r=0.760). B. Microsatellite data; regression line: $\hat{\theta}\cdot(1-\hat{\theta}) = 0.004\cdot\ln(SD)-0.010$ (r=0.656).


RS
SK
TW
DS
PI
LL
RA
Pacific Ocean
Indian Ocean
GL
JN
EU
MD
NC
SI


Indian Ocean
Pacific Ocean


A
Probability
1
0.8
0.6
0.4
0.2
0
RS
EU
JN
MD
GI
PI
DS
TW
LL
SK
RA
NC
SI
B
PC 2 (13.8%)
0
EU
MD
GI
JN
RS
SI
NC
LL
RA
PI
DS
SK
TW
0
PC 1 (51.9%)


A
1
0.5
0
$\hat{\theta}/(1-\hat{\theta})$
B
0.1
0
$\hat{\theta}/(1-\hat{\theta})$
5
6
7
8
9
10
ln(SD) (km)

Supplementary material to "Phylogeography of the humbug damselfish, *Dascyllus aruanus*: evidence of an Indo-Pacific geographic barrier and genetic differentiation of peripheral populations"

by Shang-Yin Vanson Liu, Feng-Ting Chang, Philippe Borsa, Wei-Jen Chen, Chang-Feng Dai

Tables S1, S2, S3, S4 and Figures S1, S2 here appended.

Table S1. *Dascyllus aruanus*.Diversity indices based on microsatellite data for 13 samples collected across the Indo-West Pacific. *N*: sample size; *Na*: number of alleles; *Ho*: observed heterozygosity; *He*: expected heterozygosity; *f*: Weir & Cockerham's fixation index; *Rs*: allelic richness. Key to sample abbreviations in Table 1. ***Bold***: $P < 0.05$ (based on 1000 random permutations using ARLEQUIN 3.1)

Locus, Parameter	Sample												
	West Indian Ocean					Pacific Ocean							
	RS	EU	JN	MD	GL	PI	DS	TW	LL	SK	RA	NC	SI
(*N*)	(15)	(15)	(5)	(12)	(5)	(12)	(16)	(27)	(32)	(25)	(33)	(38)	(25)
Da304													
Na	13	14	8	21	9	14	19	21	19	21	27	21	14
Ho	0.667	0.933	1	0.560	1	1	0.875	0.815	0.719	0.560	0.758	0.711	0.720
He	0.887	0.924	0.956	0.925	0.978	0.928	0.960	0.934	0.888	0.925	0.946	0.886	0.910
Rs	6.655	7.361	8	8.320	9	7.722	8.392	7.590	6.620	7.497	8.046	6.507	6.869
f	0.255	- 0.010	- 0.053	0.224	- 0.026	- 0.069	0.091	0.129	0.193	0.399	0.202	0.200	0.212
Da593													
Na	17	18	8	18	6	13	19	19	30	18	23	24	19
Ho	0.800	0.933	1	0.720	0.800	0.583	0.875	0.778	0.844	0.720	0.879	0.816	0.800
He	0.954	0.959	0.956	0.927	0.889	0.873	0.944	0.923	0.958	0.927	0.943	0.925	0.927
Rs	8.160	8.397	8	7.129	6	7.457	7.979	7.398	8.402	7.403	7.899	7.423	7.479
F	0.166	0.027	- 0.053	0.091	0.111	0.382	0.075	0.160	0.121	0.226	0.069	0.119	0.139
Da494													
Na	15	21	7	23	7	14	23	30	33	23	31	29	21
Ho	0.643	0.667	0.600	0.682	0.600	0.750	0.800	0.667	0.844	0.682	0.625	0.774	0.760
He	0.907	0.977	0.911	0.964	0.867	0.946	0.984	0.977	0.955	0.964	0.962	0.952	0.904
Rs	7.332	8.995	7	7.828	7	7.317	9.276	9.041	8.431	8.537	8.508	8.270	7.276
F	0.299	0.325	0.368	0.139	0.333	0.178	0.192	0.322	0.118	0.298	0.354	0.190	0.162
Da360													
Na	13	11	6	15	5	10	12	15	16	15	20	15	9
Ho	0.600	0.733	0.600	0.640	0.400	0.750	0.933	0.815	0.844	0.640	0.758	0.676	0.760
He	0.897	0.892	0.844	0.863	0.844	0.783	0.828	0.836	0.869	0.863	0.902	0.837	0.807
Rs	6.743	6.449	6	6.966	5	5.661	5.765	6.119	6.345	6.207	6.963	6.083	5.212
F	0.339	0.183	0.314	0.087	0.556	0.043	- 0.133	0.026	0.030	0.262	0.162	0.195	0.059
Da590													
Na	11	15	9	18	7	13	12	18	27	18	21	28	18
Ho	0.600	0.733	1	0.560	0.800	0.667	0.800	0.667	0.781	0.560	0.813	0.784	0.760
He	0.837	0.954	0.978	0.944	0.933	0.906	0.897	0.923	0.933	0.944	0.946	0.964	0.937
Rs	5.716	8.101	9	6.454	7	7.831	6.626	7.348	7.849	7.845	7.929	8.555	7.652
f	0.290	0.238	- 0.026	0.150	0.158	0.304	0.111	0.282	0.165	0.412	0.143	0.189	0.192
Da523													
Na	12	12	7	18	5	10	18	19	20	18	20	24	12
Ho	0.733	0.800	0.600	0.640	0.400	0.750	0.875	0.778	0.813	0.640	0.758	0.838	0.880
He	0.924	0.917	0.911	0.853	0.844	0.928	0.942	0.905	0.917	0.853	0.941	0.908	0.886
Rs	7.155	7.080	7	8.859	5	6.226	7.888	7.035	7.285	6.181	7.79	7.373	6.343
f	0.212	0.132	0.368	0.151	0.556	0.147	0.073	0.143	0.116	0.254	0.197	0.078	0.007
Da589													
Na	11	14	7	24	6	13	17	24	28	24	27	32	18
Ho	0.467	0.467	0.600	0.520	0.200	0.833	0.688	0.440	0.625	0.520	0.688	0.541	0.400
He	0.908	0.938	0.933	0.956	0.911	0.938	0.956	0.962	0.952	0.956	0.959	0.954	0.936
Rs	6.750	7.636	7	7.559	6	7.457	8.192	8.491	8.241	8.281	8.395	8.401	7.606
f	0.495	0.511	0.385	0.205	0.800	0.106	0.287	0.548	0.347	0.461	0.287	0.437	0.578
All loci													
Na	13.1	15.0	7.4	19.6	6.4	12.4	17.1	20.9	24.7	19.6	24.1	24.7	15.9
Ho	0.644	0.752	0.771	0.617	0.600	0.762	0.835	0.708	0.781	0.617	0.754	0.734	0.726
He	0.902	0.937	0.927	0.919	0.895	0.900	0.930	0.923	0.925	0.919	0.943	0.918	0.901
Rs	5.788	6.521	6.278	6.582	5.587	6.085	6.602	6.541	6.413	6.396	6.814	6.473	5.871
f	0.293	0.203	0.185	0.150	0.356	0.159	0.105	0.236	0.157	0.333	0.203	0.203	0.198

Table S2. *Dascyllus aruanus*. Pairwise θ estimates between populations based on cytochrome *b* haplotype frequencies. Key to samples in Table 1.*: $P< 0.05$

Sample	Sample												
	RS	EU	JN	MD	Gl	PI	DS	TW	LL	SK	RA	NC	SI
RS		0.354*	0.369*	0.353*	0.386*	0.802*	0.798*	0.810*	0.777*	0.765*	0.769*	0.692*	0.858*
EU			-0.039	-0.028	0.000	0.702*	0.711*	0.733*	0.701*	0.683*	0.691*	0.619*	0.801*
JN				-0.080	-0.015	0.743*	0.743*	0.770*	0.728*	0.706*	0.717*	0.623*	0.835*
MD					0.053	0.734*	0.738*	0.757*	0.721*	0.703*	0.709*	0.629*	0.822*
GL						0.648*	0.666*	0.707*	0.666*	0.635*	0.648*	0.560*	0.781*
PI							0.102*	0.061	0.026	0.032	0.079	0.117*	0.556*
DS								-0.020	-0.006	0.021	0.180*	0.060	0.549*
TW									-0.016	0.015	0.142*	0.086	0.552*
LL										-0.010	0.090*	0.069	0.500*
SK											0.029	0.021	0.512*
RA												0.079	0.547*
NC													0.466*
SI													

Table S3. *Dascyllus aruanus*. Pairwise θ estimates between populations based on genotypes at seven microsatellite loci (above diagonal) and ship distances between locations (below diagonal, in km). Key to sample abbreviations in Table 1. *: $P< 0.05$

Sample	Sample												
	RS	EU	JN	MD	GL	PI	DS	TW	LL	SK	RA	NC	SI
RS	-	0.039	0.048*	0.043*	0.036	0.064*	0.046*	0.056*	0.051*	0.059*	0.036*	0.056*	0.066*
EU	6,986	-	-0.001	0.011	0.003	0.063*	0.042*	0.055*	0.052*	0.052*	0.038*	0.047*	0.062*
JN	6,341	638	-	0.028	0.015	0.066*	0.048*	0.052*	0.055*	0.047*	0.044*	0.050*	0.064*
MD	5,633	360	717	-	0.028	0.062*	0.034*	0.053*	0.053*	0.049*	0.036*	0.047*	0.056*
GL	5,619	1,407	784	1,444	-	0.073*	0.058*	0.064*	0.061*	0.057*	0.047*	0.059*	0.078*
PI	11,871	9,948	9,695	9,667	9,007	-	0.023*	0.016*	0.028*	0.028*	0.027*	0.028*	0.040*
DS	12,500	10,577	10,324	10,296	9,636	730	-	0.007	0.021*	0.009	0.013*	0.020*	0.038*
TW	12,804	10,881	10,626	10,600	9,940	1,138	431	-	0.020*	0.004	0.014*	0.019*	0.033*
LL	12,336	9,713	9,460	9,432	8,772	1,617	1,618	1,391	-	0.026*	0.005	0.019*	0.031*
SK	13,646	11,723	11,468	11,442	10,782	2,037	1,294	895	1,884	-	0.016*	0.023*	0.037*
RA	12,771	10,312	10,307	10,031	9,601	2,849	2,926	2,713	1,736	3,035	-	0.012*	0.028*
NC	16,340	13,865	13,940	13,584	13,234	7,528	7,212	6,955	6,010	6,920	4,679	-	0.025*
SI	20,787	18,312	10,307	18,031	17,681	11,604	11,026	10,667	10,039	10,203	8,980	4,587	-

Table S4. *Dascyllus aruanus*. Individual genotypes at 7 microsatellite loci. *xxx*, no data

Sample, individual	Locus													
	Da304		*Da593*		*Da494*		*Da360*		*Da590*		*Da523*		*Da589*	
Gulf of Aqaba, Red Sea														
RS1	143	143	155	172	146	183	212	233	163	163	153	170	201	201
RS2	121	143	151	159	146	146	208	221	148	157	139	155	191	201
RS3	120	141	161	166	152	152	224	224	157	157	128	132	191	195
RS4	154	160	155	159	144	144	208	208	157	165	133	149	197	197
RS5	121	143	139	161	146	192	210	210	161	161	164	164	191	201
RS6	144	161	151	174	143	143	212	240	152	157	133	155	196	196
RS7	121	133	173	179	146	211	210	210	161	171	128	133	191	207
RS8	141	141	149	160	146	202	206	211	161	161	139	139	196	196
RS9	160	167	162	162	148	202	208	240	163	163	144	161	187	187
RS10	164	164	159	159	187	194	202	214	157	167	160	170	197	197
RS11	141	141	166	170	xxx	xxx	202	208	161	173	128	128	191	201
RS12	143	156	162	166	152	175	212	229	153	163	133	133	202	202
RS13	141	141	153	157	146	146	208	208	150	161	155	160	197	209
RS14	128	141	153	153	154	216	212	215	161	161	132	139	199	209
RS15	143	161	155	180	175	196	210	210	150	161	132	160	193	193
Europa, Mozambique Channel														
Eu1	154	154	139	164	171	215	206	233	150	150	131	161	196	196
Eu2	133	139	151	161	152	162	233	240	138	152	128	135	197	197
Eu3	145	150	146	153	133	170	210	214	167	182	135	156	173	183
Eu4	121	133	153	153	173	206	210	210	148	157	142	162	193	193
Eu5	129	150	157	166	146	146	210	233	169	169	128	160	184	200
Eu6	150	179	135	159	133	209	212	212	141	186	128	135	185	199
Eu7	121	150	142	158	156	156	206	233	171	186	133	133	187	193
Eu8	121	143	151	178	183	183	215	234	171	171	131	144	193	193
Eu9	125	143	139	160	169	211	212	224	134	141	131	162	207	207
Eu10	129	179	149	164	141	141	206	232	148	148	137	155	175	211
Eu11	120	152	158	170	162	169	214	221	146	171	128	128	179	179
Eu12	121	179	149	157	162	198	212	212	134	159	131	142	175	175
Eu13	139	150	153	162	154	219	221	233	152	159	137	160	187	187
Eu14	121	147	142	172	217	217	206	214	138	147	137	144	175	197
Eu15	135	150	146	153	150	179	212	212	150	157	131	131	183	211
Juan de Nova, Mozambique Channel														
JN1	121	143	142	153	141	141	214	214	150	157	129	135	175	207
JN2	125	152	142	158	154	164	210	219	146	152	162	162	203	211
JN3	120	143	135	151	141	179	212	212	141	152	149	160	175	209
JN4	137	145	135	146	206	211	212	240	147	175	128	137	194	194
JN5	145	150	139	161	207	207	212	231	154	173	135	135	171	171
Toliara, Madagascar														
MD2	120	142	149	157	145	169	202	232	150	154	138	143	159	199
MD4	120	120	153	164	154	198	208	208	171	171	132	132	187	200
MD5	123	150	150	164	161	173	202	224	141	161	134	160	175	187
MD6	128	142	162	174	151	198	212	221	150	171	135	161	185	213
MD8	125	129	146	162	160	160	202	215	150	150	128	128	171	175
MD9	121	125	146	146	152	181	206	233	143	152	131	135	173	199
MD10	135	145	135	153	163	188	202	202	150	150	162	165	179	179
MD11	114	149	136	153	145	145	202	232	152	186	131	166	197	197
MD12	159	159	149	166	145	160	208	225	141	148	142	161	173	173
MD13	121	179	138	166	xxx	xxx	206	221	157	171	130	166	177	195
MD14	127	175	153	162	152	161	212	236	171	186	139	163	187	197
MD15	142	142	153	153	141	151	212	225	150	159	130	155	187	197
Glorieuses Is., Mozambique Channel														
LG1	128	145	151	166	179	202	212	212	152	152	133	133	177	193

LG2	121	125	153	160	152	152	212	233	150	155	131	131	195	195
LG3	141	179	147	147	154	169	208	214	167	171	131	149	188	188
LG4	129	154	151	162	150	190	210	210	150	171	133	144	197	197
LG5	121	150	151	160	152	152	214	214	161	204	135	135	198	198
Paracels Islands, South China Sea														
Xs1	139	152	168	168	190	235	214	225	171	194	133	133	183	191
Xs2	117	139	163	163	156	156	225	225	159	165	139	179	179	201
Xs3	140	173	157	157	156	156	217	227	161	206	168	168	189	195
Xs4	131	135	147	192	150	198	208	236	163	196	141	170	172	172
Xs5	131	152	163	176	162	185	225	236	175	175	135	139	195	209
Xs6	129	139	163	188	191	211	221	229	171	171	135	139	187	193
Xs7	125	141	166	176	188	192	225	225	163	163	139	147	183	207
Xs8	160	167	157	178	158	191	225	232	159	178	135	147	207	207
Xs9	131	139	132	172	167	185	225	225	165	173	139	164	207	245
Xs10	131	152	166	177	162	185	225	236	176	176	135	139	195	207
Xs11	126	143	161	161	156	202	215	225	163	190	128	139	201	233
Xs12	140	173	157	157	156	156	217	227	161	206	168	168	189	195
Dongsha, South China Sea														
Ds1	138	138	153	159	144	144	225	229	167	184	155	166	191	241
Ds2	129	139	159	176	166	208	208	225	163	163	129	147	187	187
Ds3	125	129	118	168	206	206	225	225	169	173	168	181	173	187
Ds4	130	140	153	153	175	175	208	225	152	169	139	179	179	193
Ds5	129	145	153	180	225	255	208	231	150	167	139	174	181	203
Ds6	169	169	128	159	160	167	208	225	159	169	135	141	191	219
Ds7	129	139	144	170	160	213	204	210	167	167	133	168	181	181
Ds8	109	127	146	159	141	196	216	225	150	163	141	172	173	191
Ds9	159	196	155	163	181	208	208	212	169	177	139	139	185	231
Ds10	138	152	155	155	139	204	xxx	xxx	163	171	135	179	219	219
Ds11	159	165	154	161	154	204	208	225	xxx	xxx	135	181	189	193
Ds12	143	154	166	172	156	200	225	234	163	186	141	141	175	175
Ds13	131	147	118	164	147	219	208	236	163	163	135	153	203	219
Ds14	125	139	155	159	141	183	223	233	150	169	149	162	179	179
Ds15	127	139	157	174	162	187	208	216	150	171	160	176	211	233
Ds16	140	160	155	161	xxx	xxx	225	229	159	182	135	139	215	221
Houbihu, Taiwan														
KTN1	128	140	118	170	143	143	212	225	169	184	147	147	183	191
KTN2	140	150	161	174	141	232	229	232	165	165	135	139	201	201
KTN3	128	152	164	176	181	196	214	225	163	163	160	166	201	201
KTN4	141	152	154	162	159	194	198	214	177	202	135	139	177	195
KTN5	128	140	162	170	206	206	225	236	159	165	135	139	xxx	xxx
KTN6	138	152	151	172	156	192	208	225	159	184	137	137	199	221
KTN7	132	139	170	170	xxx	xxx	208	225	169	184	135	162	233	257
KTN8	139	139	161	161	200	200	225	225	163	169	135	162	179	179
KTN9	128	128	155	155	198	198	225	231	169	192	133	133	183	219
Nu31	129	141	162	172	141	170	208	236	163	173	160	179	xxx	xxx
Nu32	129	159	118	153	xxx	xxx	225	225	177	177	128	135	189	189
Nu33	141	152	155	172	200	209	227	227	159	175	149	149	207	207
Nu34	124	124	157	157	xxx	xxx	219	225	141	141	139	157	219	219
Nu35	126	136	155	178	152	187	208	213	169	190	122	140	179	227
Nu36	141	141	132	164	158	158	220	225	179	179	139	157	215	231
Nu37	138	151	155	162	169	189	227	232	177	177	133	168	187	205
Nu38	128	150	155	155	152	217	225	231	171	194	135	178	173	173
Nu39	128	138	151	155	200	200	225	225	169	169	153	160	223	223
Nu310	139	142	163	170	150	161	219	232	163	163	135	147	235	257
Nu311	138	162	170	191	144	235	208	225	169	180	133	185	177	177
Nu312	141	152	168	182	177	207	225	225	150	173	135	135	167	167
Nu313	139	165	118	170	202	202	204	221	163	163	135	139	195	195
Nu314	114	120	155	168	179	207	208	225	159	202	135	149	183	183
Nu315	129	155	155	182	231	231	227	232	186	192	153	160	201	201

Nu316	138	138	157	168	141	156	225	229	163	175	139	139	211	211
Nu317	128	138	168	168	168	192	208	212	159	173	139	151	203	211
Nu318	140	149	155	172	194	202	221	227	163	173	139	149	199	211
Caohagan, Lapu-Lapu, Philippines														
Ph1	131	141	156	183	162	175	225	225	159	169	151	170	215	215
Ph2	139	157	145	157	187	209	206	227	163	163	142	149	191	191
Ph3	141	141	170	185	179	179	214	229	138	165	144	149	181	213
Ph4	141	141	157	161	162	171	225	236	127	152	151	162	201	233
Ph5	129	143	162	170	160	219	229	229	165	165	142	149	189	213
Ph6	129	147	153	161	158	158	208	208	127	161	124	139	183	203
Ph7	140	140	141	162	150	233	217	229	127	169	149	173	191	209
Ph8	141	163	148	168	148	200	216	225	165	165	142	151	182	207
Ph9	129	145	155	170	156	192	225	240	165	196	147	147	183	199
Ph10	137	141	118	166	143	156	214	231	171	182	149	160	211	211
Ph11	139	142	120	166	144	155	214	231	155	179	149	160	211	211
Ph12	129	160	168	189	156	215	208	227	165	165	170	177	201	201
Ph13	139	147	164	170	155	158	208	225	165	198	120	170	173	173
Ph14	131	139	153	159	181	181	217	225	165	165	120	135	179	207
Ph15	135	141	159	159	156	179	206	219	127	167	139	144	193	200
Ph16	129	141	161	161	171	200	208	229	150	155	146	168	183	183
Ph17	129	143	118	174	152	189	225	238	152	178	142	149	199	215
Ph18	129	145	155	170	156	192	225	240	163	196	147	147	183	183
Ph19	139	155	155	170	187	217	208	231	157	165	146	158	183	197
Ph20	152	160	159	159	154	209	227	230	165	175	142	158	187	249
Ph21	139	139	155	176	156	192	225	225	161	176	129	170	211	211
Ph22	127	141	118	172	156	198	208	232	177	177	149	149	183	191
Ph23	139	142	147	160	168	227	229	229	159	184	139	151	187	195
Ph24	141	141	155	168	150	172	225	230	171	194	120	126	195	251
Ph25	129	139	170	170	156	192	225	231	171	215	142	149	173	179
Ph26	139	152	163	163	147	179	225	238	148	165	147	147	205	243
Ph27	129	141	164	187	179	183	208	225	169	176	120	176	173	183
Ph28	139	159	126	168	156	206	214	225	127	154	149	149	219	243
Ph29	129	129	157	180	156	156	221	225	161	169	120	120	185	185
Ph30	173	173	155	159	183	183	214	225	154	173	135	149	217	217
Ph31	131	131	172	179	173	210	208	229	159	159	120	170	183	189
Ph32	139	139	151	164	192	215	214	225	127	152	149	168	173	173
Sesoko Island, Japan														
SK1	141	165	155	172	190	190	216	216	163	171	139	168	191	191
SK2	127	131	155	164	155	186	225	234	150	169	135	149	191	205
SK3	138	138	160	163	147	147	208	227	167	167	158	185	191	191
SK4	159	166	164	164	155	155	208	208	159	190	135	135	181	237
SK5	129	169	164	172	158	198	214	225	167	167	135	139	183	183
SK6	138	153	160	171	194	213	206	219	165	165	162	168	175	175
SK7	140	140	155	161	189	210	225	229	165	171	135	153	177	183
SK8	138	138	178	178	156	183	225	229	150	159	139	139	195	233
SK9	173	173	168	174	xxx	xxx	227	229	141	171	133	133	211	231
SK10	139	155	153	153	173	211	210	215	161	190	139	155	195	211
SK11	166	166	144	157	xxx	xxx	225	225	154	165	135	157	231	231
SK12	128	138	163	172	177	202	225	225	161	161	149	149	221	221
SK13	138	151	155	155	169	209	208	227	175	175	139	139	185	233
SK14	150	150	161	170	186	205	219	225	173	173	139	176	200	200
SK15	128	128	155	166	156	209	208	238	154	154	135	135	179	179
SK16	133	169	161	168	211	211	225	236	159	177	144	151	171	187
SK17	128	157	137	153	156	211	227	227	152	165	135	135	193	211
SK18	128	141	165	165	194	194	208	208	169	169	149	182	221	221
SK19	140	140	168	172	144	179	208	216	155	169	135	153	203	213
SK20	127	139	155	155	183	221	206	206	147	167	149	149	257	257
SK21	127	158	168	168	210	210	225	225	152	169	126	135	179	179
SK22	140	140	161	178	190	190	225	229	163	173	135	135	183	194

SK23	128	128	159	164	xxx	xxx	208	212	169	169	135	147	179	192
SK24	128	128	155	168	146	211	225	225	159	159	135	139	183	183
SK25	128	143	153	164	189	221	229	232	174	174	149	183	167	233
Raja Ampat, West Papua														
RA1	141	141	159	178	150	181	225	233	154	154	156	156	195	217
RA2	128	139	142	155	160	160	225	225	127	159	158	164	211	211
RA3	127	127	176	176	154	213	229	229	152	171	142	149	191	215
RA4	140	140	155	161	208	208	231	231	165	186	149	155	177	177
RA5	118	138	168	174	150	167	208	219	127	161	135	160	187	205
RA6	143	175	118	170	xxx	xxx	208	225	159	165	139	139	183	187
RA7	129	158	170	190	156	156	212	229	171	186	149	166	195	195
RA8	139	166	155	188	144	162	225	229	157	171	149	170	197	197
RA9	135	152	172	195	190	208	208	234	167	167	133	133	179	196
RA10	131	139	155	161	162	162	214	227	165	186	144	166	197	197
RA11	141	160	166	180	150	183	208	221	161	161	156	156	188	200
RA12	141	152	146	166	172	172	208	225	xxx	xxx	142	151	185	203
RA13	127	141	155	172	198	208	208	225	157	173	149	160	200	217
RA14	129	141	155	155	162	166	217	225	157	184	135	153	187	205
RA15	140	165	159	168	168	168	208	219	161	176	135	162	187	201
RA16	139	154	161	186	145	208	225	225	127	171	139	151	201	215
RA17	131	152	155	174	183	183	214	227	127	127	142	160	173	191
RA18	129	129	159	166	156	168	217	223	127	186	150	150	203	243
RA19	139	139	157	172	154	154	210	227	127	154	135	142	199	203
RA20	147	175	153	159	168	185	213	228	154	169	153	160	215	215
RA21	141	166	164	170	179	202	221	225	150	159	149	162	198	198
RA22	133	160	155	170	147	196	230	230	143	169	135	183	195	215
RA23	128	128	126	176	168	202	219	229	188	188	144	149	173	199
RA24	152	200	155	184	161	161	225	234	157	182	135	176	213	233
RA25	123	160	161	168	169	192	229	229	159	173	139	139	173	173
RA26	139	161	172	172	160	207	208	225	152	186	144	164	173	173
RA27	129	148	164	170	160	209	206	206	157	157	144	172	197	207
RA28	131	141	146	174	160	183	223	227	159	165	133	133	181	187
RA29	140	140	172	172	189	208	229	229	163	182	149	160	173	201
RA30	139	154	161	186	211	211	225	236	165	186	149	155	xxx	xxx
RA31	139	139	118	176	171	175	223	238	161	188	144	144	187	225
RA32	125	157	157	176	162	162	225	229	155	186	149	153	183	237
RA33	135	159	153	170	156	156	223	229	163	180	156	183	183	183
SW Lagoon, New Caledonia														
NC1	139	148	149	161	200	200	223	236	169	192	120	166	185	243
NC2	140	140	141	188	187	208	225	229	180	190	130	135	197	197
NC3	129	158	161	170	177	189	225	229	180	188	135	160	207	207
NC4	139	152	153	174	175	191	225	231	175	180	130	144	195	203
NC5	129	139	151	155	162	208	221	246	155	176	120	126	211	257
NC6	139	139	151	174	154	190	221	231	165	184	135	135	187	187
NC7	133	143	158	170	xxx	xxx	xxx	xxx	161	161	131	135	173	173
NC8	140	140	155	155	158	175	214	219	157	171	157	170	183	183
NC9	129	143	155	172	156	196	229	238	157	184	126	139	183	187
NC10	129	143	155	172	196	196	229	238	154	184	126	139	183	187
NC11	139	139	151	185	xxx	xxx	208	225	165	165	137	144	187	201
NC12	137	147	151	191	xxx	xxx	223	223	154	165	133	137	231	231
NC13	141	156	161	174	160	196	206	225	171	178	135	172	237	237
NC14	128	140	151	188	144	148	227	227	171	186	137	157	189	245
NC15	129	143	133	174	154	204	225	225	161	161	128	162	203	247
NC16	129	156	149	159	168	194	214	230	177	177	142	168	187	195
NC17	131	139	153	153	156	179	225	236	169	169	120	135	205	213
NC18	129	129	153	153	187	213	227	227	169	178	135	155	165	199
NC19	129	139	149	161	154	154	225	225	167	192	xxx	xxx	188	241
NC20	141	141	153	174	156	183	225	229	175	182	135	135	205	205
NC21	139	169	155	155	156	160	208	217	165	178	137	155	200	200

NC22	131	143	161	176	143	198	208	225	175	175	135	142	203	239
NC23	143	143	178	178	xxx	xxx	225	225	172	192	131	157	187	187
NC24	128	141	153	153	158	179	223	229	171	190	141	174	187	187
NC25	129	141	155	187	183	191	208	225	159	174	135	135	171	171
NC26	141	166	151	161	167	202	217	229	161	186	118	135	189	195
NC27	131	139	177	185	160	160	223	231	143	176	153	166	173	201
NC28	129	207	149	178	xxx	xxx	225	225	134	165	144	170	213	236
NC29	141	141	157	161	xxx	xxx	214	225	161	223	142	142	188	207
NC30	129	154	133	155	158	158	225	225	154	159	135	142	191	238
NC31	139	148	149	161	154	154	225	225	xxx	xxx	133	133	xxx	xxx
NC32	117	139	155	188	156	183	225	225	154	198	143	155	171	187
NC33	141	150	155	155	150	156	231	246	178	188	135	135	167	167
NC34	141	141	162	180	175	234	225	225	173	173	135	141	179	179
NC35	129	141	126	176	156	156	217	221	161	171	131	157	237	237
NC36	139	139	153	161	156	181	225	231	177	177	157	172	187	187
NC37	129	129	155	190	156	181	223	223	163	182	131	135	179	179
NC38	127	141	153	161	xxx	xxx	223	231	163	188	135	157	199	219
Society Islands														
Ra1	130	138	155	162	156	196	225	231	161	174	126	155	199	217
Ra2	141	141	171	171	158	167	219	225	173	182	126	155	195	213
Ra3	139	158	162	180	164	164	225	229	163	172	135	168	209	209
Ra4	131	143	153	170	156	161	217	231	159	167	126	143	215	215
Ra5	131	141	166	174	156	165	225	225	198	206	142	160	207	207
Ra6	130	130	155	162	163	179	219	225	171	171	126	135	179	179
FP1	141	141	155	164	156	160	219	227	167	177	142	155	197	215
FP2	141	141	153	168	156	164	225	229	150	173	126	137	211	211
FP3	139	139	161	161	190	217	225	225	178	178	130	142	211	211
FP4	140	140	176	194	150	156	219	225	165	165	120	155	203	247
FP5	116	142	164	190	156	156	219	225	161	167	126	157	175	175
FP6	138	157	155	170	163	163	217	232	161	178	126	137	199	199
FP7	130	138	161	174	156	179	227	230	148	154	120	120	175	187
FP8	130	140	163	171	156	189	225	231	161	178	126	135	183	187
FP9	142	156	154	162	163	188	219	225	167	167	137	142	179	179
FP10	131	139	162	162	161	202	217	225	171	182	126	126	195	195
FP11	116	139	155	174	167	167	217	219	171	171	143	143	195	195
FP12	140	144	168	183	156	164	212	232	161	171	142	155	215	215
FP13	140	156	161	161	171	183	225	225	167	177	120	137	211	211
FP14	130	138	155	176	154	154	225	227	161	204	120	126	175	175
FP15	116	138	172	172	150	156	219	219	154	167	126	155	205	215
FP16	116	140	162	190	164	183	225	230	154	206	139	143	183	187
FP17	139	139	149	155	156	156	219	219	163	173	142	168	179	179
FP18	141	156	162	174	183	188	217	217	165	165	120	135	201	213
FP19	133	141	155	162	169	184	227	230	173	198	120	143	173	199

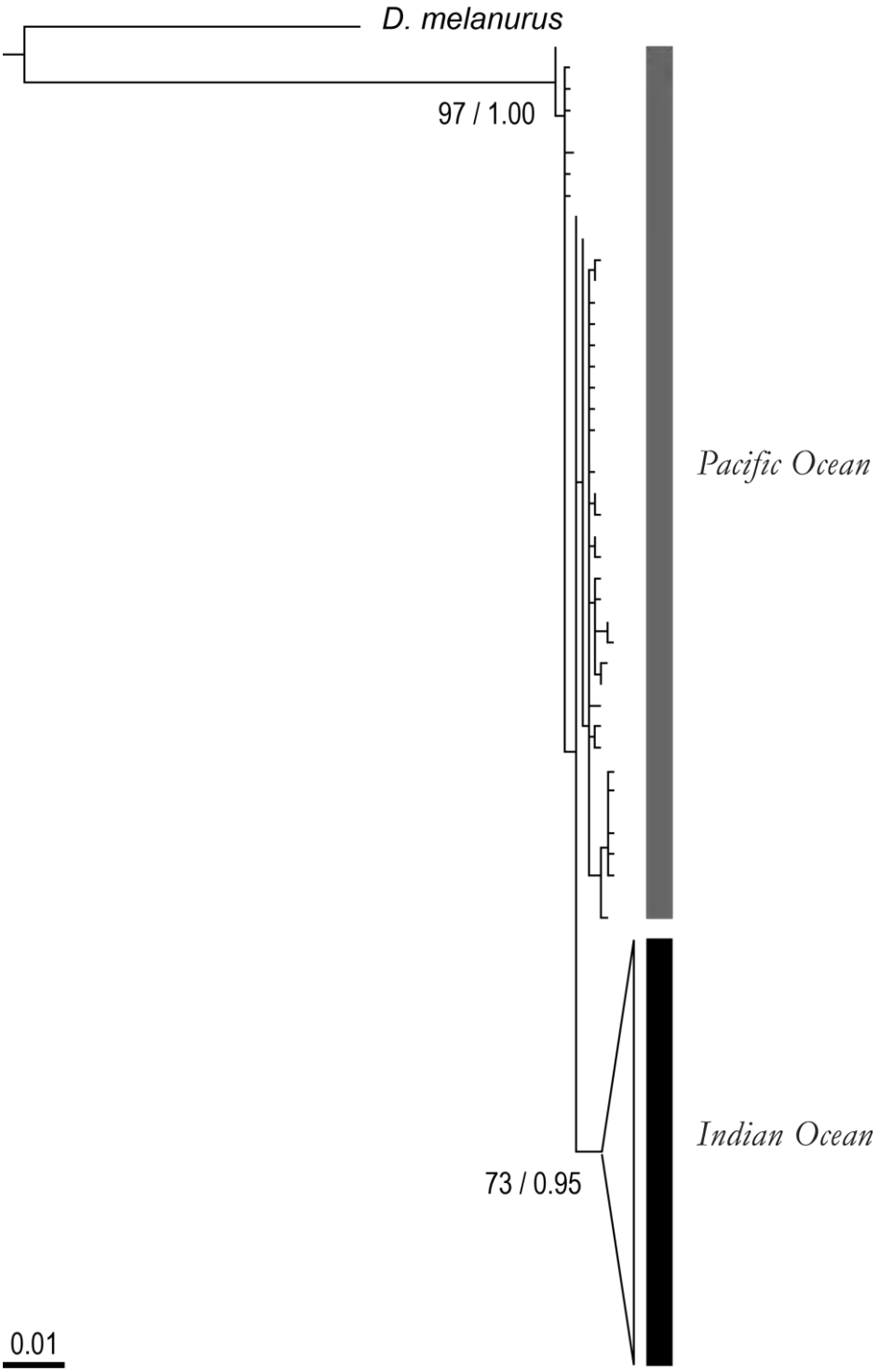


Figure S1. *Dascyllus aruanus*. Maximum likelihood phylogenetic tree based on mitochondrial (cytochrome *b* gene) nucleotide sequences. Numbers at nodes are bootstrap scores and concern the maximum likelihood phylogeny (before slash) and Bayesian probability values (after slash). *Black bars* indicates Pacific haplotypes; *gray bar* indicates Indian haplotypes.

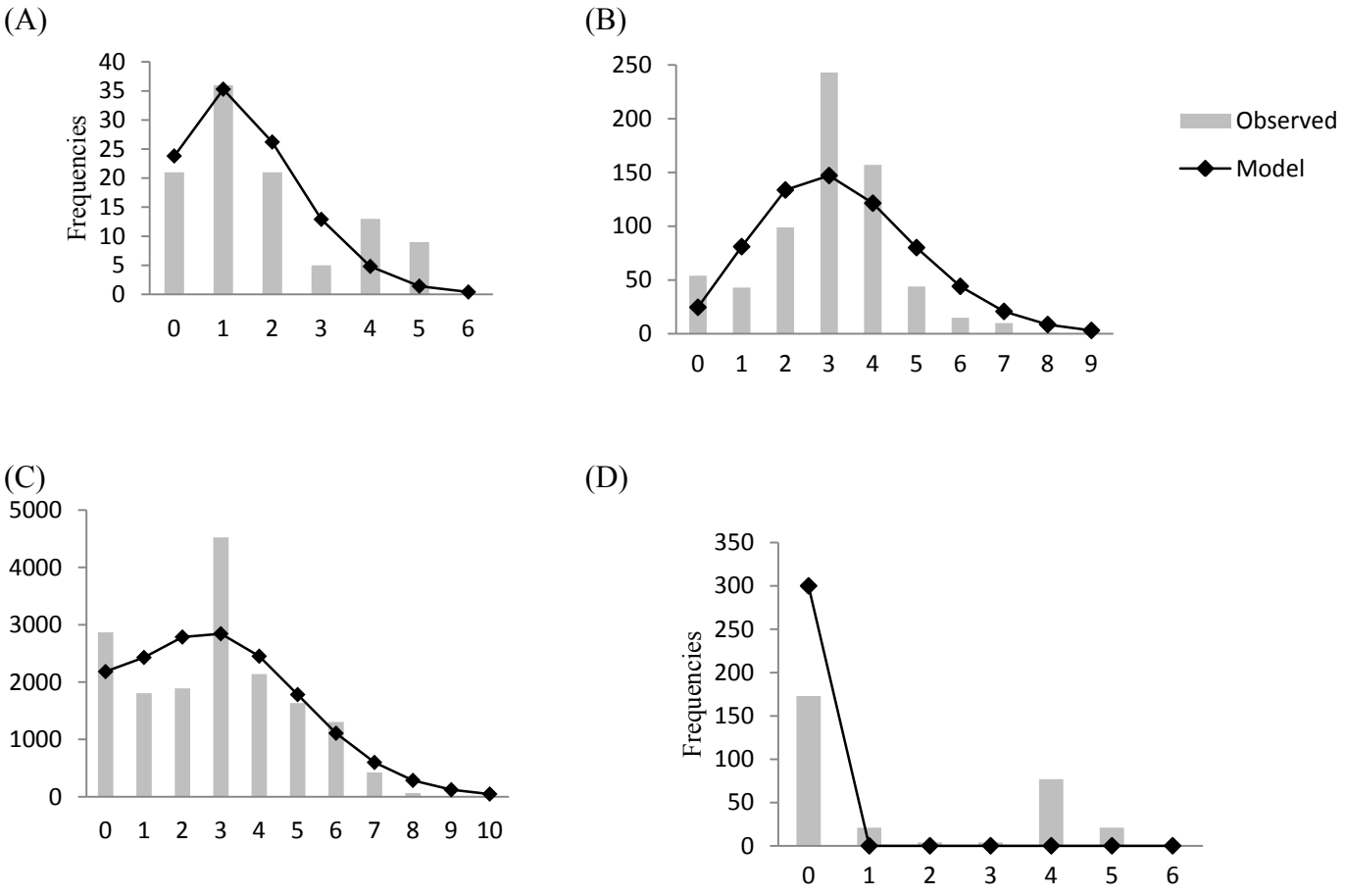


Fig. S2. *Dascyllus aruanus.* Mismatch distribution for the Red Sea population (A), the Indian Ocean population (B),the Pacific Ocean population (C), and the Society Islands population (D), based on cytochrome-*b* gene sequence data. *Black line* is the expected curve under constant population size. *Grey bars* indicate the observed frequency of haplotype pairs with the number of pairwise differences in abscissa.