



# Phylogeography of the sergeants *Abudefduf sexfasciatus* and *A. vaigiensis* reveals complex introgression patterns between two widespread and sympatric Indo-West Pacific reef fishes

Joris A.M. Bertrand, Philippe Borsa, Wei-Jen Chen

## ► To cite this version:

Joris A.M. Bertrand, Philippe Borsa, Wei-Jen Chen. Phylogeography of the sergeants *Abudefduf sexfasciatus* and *A. vaigiensis* reveals complex introgression patterns between two widespread and sympatric Indo-West Pacific reef fishes. *Molecular Ecology*, 2017, 26, pp.2527 - 2542. 10.1111/mec.14044 . ird-01525350v3

HAL Id: ird-01525350

<https://ird.hal.science/ird-01525350v3>

Submitted on 3 Jun 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

To be cited as: Bertrand JAM, Borsa P, Chen W-J (2017) Phylogeography of the sergeants *Abudefduf sexfasciatus* and *A. vaigiensis* reveals complex introgression patterns between two widespread and sympatric Indo-West Pacific reef fishes. *Molecular Ecology* 26, 2527–2542.

## Phylogeography of the sergeants *Abudefduf sexfasciatus* and *A. vaigiensis* reveals complex introgression patterns between two widespread and sympatric Indo-West Pacific reef fishes

JORIS A.M. BERTRAND,\* PHILIPPE BORSA† and WEI-JEN CHEN\*

\*Institute of Oceanography, National Taiwan University, Taipei, Taiwan. †Institut de recherche pour le développement (IRD), UMR 250 "Ecologie marine tropicale des océans Pacifique et Indien", Nouméa, New Caledonia.

Correspondence: Joris Bertrand, University of Lausanne, Department of Ecology and Evolution, Quartier Sorge, Biophore, CH-1015 Lausanne, Switzerland; E-mail: jorisbertrand@gmail.com; Wei-Jen Chen; E-mail: wjchen.actinops@gmail.com

**Keywords:** coral reef, cryptic lineage, inter-specific gene flow, mito-nuclear discordance, multi-locus phylogeny, Pomacentridae.

Running title: Cryptic taxa and introgression in a reef fish

## Abstract

On evolutionary time scales, sea level oscillations lead to recurrent spatio-temporal variation in species distribution and population connectivity. In this situation, applying classical concepts of biogeography is challenging yet necessary to understand the mechanisms underlying biodiversity in highly diverse marine ecosystems such as coral reefs. We aimed at studying the outcomes of such complex biogeographical dynamics on reproductive isolation by sampling populations across a wide spatial range of a species-rich fish genus: the sergeants (Pomacentridae: *Abudefduf*). We generated a multi-locus data set that included ten species from 32 Indo-West Pacific localities. We observed a pattern of mito-nuclear discordance in two of the most common and widely distributed species: *Abudefduf sexfasciatus* and *A. vaigiensis*. The results showed three regional sub-lineages (Indian Ocean, Coral Triangle region, western Pacific) in *A. sexfasciatus* (0.6–1.5% divergence at locus *cytb*). The other species, *A. vaigiensis* consisted of three distinct lineages (A, B, and C) whose geographic ranges overlapped (9% divergence at locus *cytb*). Although *A. vaigiensis* lineage A and *A. sexfasciatus* were found to be distinct based on nuclear information, *A. vaigiensis* lineage A was nested within *A. sexfasciatus* in the mitochondrial gene tree. *Abudefduf sexfasciatus* from the Coral Triangle region and *A. vaigiensis* lineage A were not differentiated from each other at the mitochondrial locus. We then used coalescent-based simulation to characterize a spatially widespread, weak gene flow between the two species. These fishes may be considered as candidates of choice to investigate the complexity of the discrepancies between phenotypic and genetic evolution among sibling species.

## Introduction

Evolutionary processes that lead to reproductive isolation between populations and/or species can be temporally and spatially dynamic. For instance, closely related and previously allopatric species that come into contact may interbreed and form hybrid zones (Barton & Hewitt, 1985; 1989; Mallet, 2005; Price *et al.* 2008). Hybrid zones have been intensively studied in terrestrial systems (Harrison, 1993) but have been considered as rare in the sea until recently. However, the number of studies documenting marine hybrid zones has increased within the last decade (DiBattista *et al.* 2015; Montanari *et al.* 2012; 2014; 2016 and references therein for examples in fishes).

Characterizing spatial and temporal patterns of inter-specific gene flow in marine systems may have been more difficult than in terrestrial or riverine organisms because marine populations are generally more connected and are expected to experience higher levels of gene flow. In this context, inter-specific gene flow may give rise to complex patterns of hybridization (Fraisse *et al.* 2014; Le Moan *et al.* 2016), not only along the edges of species' ranges, but virtually everywhere their distributions overlap, sometimes resulting in mosaic-like patterns of hybridization (Hobbs *et al.* 2009; Hobbs & Allen, 2014). These particularities challenge the classical vision of the mechanisms underlying species formation and/or their integrity in spite of potential gene flow. Yet, it is of primary importance to investigate patterns of biodiversity in such marine ecosystems with complex and highly dynamic histories (see Bowen *et al.* 2013). Moreover, the role of hybridization in biodiversity is still debated, as it has been considered either as source of evolutionary novelty or as a homogenising process and a potential threat for biological diversity (Seehausen, 2004; Abbott *et al.* 2013).

The tropical Indo-West Pacific is the largest and most speciose marine biogeographic region (Crandall & Riginos, 2014). Out of the 3919 reef fish species that occur in the Indo-West Pacific, 69.1 % are widely distributed from the western Indian Ocean to the western Pacific (Allen, 2008; Briggs & Bowen, 2012; 2013). Marine populations of the western Pacific and the Indian ocean have undergone a succession of stages of geographic isolation on either side of the Indo-Pacific barrier, and subsequent secondary contact (Gaither & Rocha, 2013) due to cyclic sea-level oscillations (Voris, 2000; Siddall *et al.* 2003; Naish *et al.* 2009). This has considerably but also differently impacted the phylogeography of species in this region but also led to very different outcomes. While some species exhibit a strong Indian *vs.* Pacific phylogeographic structure, others show an apparent lack of differentiation as well as a range of intermediate situations (Gaither & Rocha, 2013; Bowen *et al.* 2016; Borsa *et al.* 2016). Even closely related congeneric species display strikingly different patterns of spatial genetic differentiation. Gaither *et al.* (2010) found a lack of genetic differentiation across the whole Indo-Pacific in the snapper *Lutjanus kasmira* but high levels of structure at all geographical scales in the closely related species: *L. fulvus*. These results are similar to the ones reported by DiBattista *et al.* (2012) between the undifferentiated *Chaetodon meyeri* and the structured populations of *C. ornatus*: two sister species of butterflyfishes. Horne and van Herwerden (2013) documented weak genetic differentiation at mitochondrial markers between *Naso hexacanthus* and *N. caesius* whereas these two sister species of unicornfishes were distinct at nuclear loci. This variability indicates that evolutionary and ecological dynamics in these organisms are species-specific. It is even possibly observable at the intra-specific level (Carpenter *et al.* 2011; Borsa *et al.* 2016).

Highly diverse and widely distributed animal groups such as the damselfishes (Pomacentridae) provide an opportunity to investigate the different factors involved in generating and maintaining spatial patterns of phenotypic and genetic diversity and differentiation in our oceans. This family comprises 399 valid species in the Indo-Pacific (Eschmeyer *et al.* 2016). Previous phylogenetic studies have documented the systematics of the Pomacentridae family (e.g. Quenouille *et al.* 2004; Cooper *et al.* 2009). Other biogeographic studies have investigated species boundaries at lower taxonomic levels in this family but only a few of them highlighted cases of cross-species hybridization (but see Litsios & Salamin, 2014 for a noticeable exception in clownfishes) suggesting that this phenomenon could be rare in Pomacentridae compared to other reef fishes from different families including Acanthuridae, Chaetodontidae, Labridae or Pomacanthidae (Montanari *et al.* 2016).

Within the Pomacentridae, the genus *Abudefduf* (sergeants) comprises Indo-West Pacific species that are generally widely distributed and locally abundant (Froese & Pauly 2015). The evolutionary history of this genus is complex, with the possible existence of cryptic species in *Abudefduf saxatilis* and *A. vaigiensis* that have not been confirmed (Quenouille *et al.* 2004; 2011) as well as blurred boundaries between *A. vaigiensis* and *A. sexfasciatus*. In a DNA barcoding survey based on the cytochrome oxidase c subunit 1 gene (COI), Hubert *et al.* (2012) found that out of 668 coral reef fish species sampled in both Indian and Pacific oceans (Madagascar, Réunion and French Polynesia), only three pairs did not show clear inter-specific boundaries, possibly because of mitochondrial introgression or because of the presence of different morphotypes in one of these species. One of these pairs was *A. sexfasciatus* (Lacépède, 1801) and *A. vaigiensis* (Quoy & Gaimard, 1825). A case of hybridization involving the Hawaiian endemic *A. abdominalis* and the recently introduced *A. vaigiensis* was also documented (Coleman *et al.* 2014). This study reported that hybridization is likely to be favoured by human activities, raised conservation issues for *A. abdominalis*, and demonstrated that the species barrier may quickly collapse after secondary contact in these fishes. Altogether, these results led us to consider the Indo-West Pacific species in the genus *Abudefduf* as candidates to investigate species formation in allopatry, secondary contact, hybridization and the mechanisms by which species integrity is maintained in parapatry or in sympatry, despite opportunity for inter-specific gene flow at broad geographic scales.

In this study, we first investigated the consistency between the phenotype-based taxonomy and the molecular systematics of species in the genus *Abudefduf*. By incorporating several individuals from different populations in each species, we conducted multi-locus phylogenetic analyses including 11 (out of 12) sergeant species occurring in the Indo-West Pacific. We tested for the existence of cryptic lineages in *A. vaigiensis*. Based on an extensive geographic sampling, we then evaluated patterns of intra- and inter-specific differentiation to look for genetic structure, identify geographic barriers to gene flow and highlight potential hybridization. Finally, we investigated the magnitude, direction and locus-dependence of inter-specific gene flow between *A. sexfasciatus* and *A. vaigiensis*, to propose an evolutionary scenario to explain patterns of diversity and differentiation in these two widely sympatric Indo-West Pacific species.

## Materials and methods

### *Collection of specimens*

We examined a total of 488 *Abudefduf* spp. specimens collected between 2006 and 2015 from 32 localities across the tropical Indo-West Pacific (Fig. 1A; sampling details in Supplementary Table S1). Fishes were caught using nets during snorkelling sessions, by angling, or bought on local markets. We used colour patterns (number and colour of lateral stripes, presence of bars or spot in the caudal region, presence of yellow coloration in the dorsal region and flanks) to unambiguously identify the specimens to species. For each specimen, a muscle-tissue or fin-clip sample was taken, then preserved in 95% ethanol and stored at -20°C in the Marine Biodiversity and Phylogenomics laboratory at the Institute of Oceanography, National Taiwan University (NTU). When possible, voucher specimens were photographed, then formaldehyde-fixed and deposited at the NTU Museum (NTUM), Taipei.

### *Molecular procedures*

Total genomic DNA was extracted from fin or muscle tissues using the automated *LabTurbo 48 Compact System* extractor (Taigene Biosciences Corp., Taipei) and *LGD 480-500* DNA-extraction kits (Taigene Biosciences Corp.) following the manufacturer's protocol. Amplification of the complete cytochrome *b* (*cyt b*) gene together with a short fragment of the adjacent tRNA-Thr was done by polymerase chain reaction (PCR). DNA fragments at two nuclear loci were also PCR-amplified: a partial sequence of an exon of the Recombination Activating gene 1 (*RAG1*) and a fragment of an intron of the Dystrophin gene (*Dyst*). Primer sequences and additional

details are provided in Supplementary Table S2. PCRs were done in 25 µL reaction volume, containing ~1-2 µL DNA template, 0.2 mM of each dNTP, 2 mM MgCl<sub>2</sub>, 0.4 µM of each primer and 0.5 U *GloTaq Flexi* DNA polymerase (Promega, Madison WI, U.S.A.) in 5 µL 5X manufacturer's buffer plus 12.9 µL sterile distilled water. The thermal profile of the PCR was as follows: initial denaturation at 95°C for 4 min, followed by 36 cycles of denaturation at 95°C for 40 s, locus-specific annealing temperature for 40 s, elongation at 72°C for a locus-specific time, and a final elongation at 72°C for 7 min. The PCR products were then checked on 1% agarose gel, purified using the *AMPure* magnetic bead cleanup protocol (Agencourt Bioscience Corp., Beverly MA, U.S.A.), and sequenced in both directions by Sanger's method on an *ABI 3730* analyzer (Applied Biosystems, Foster City CA, U.S.A) at the Centre of Biotechnology (National Taiwan University, Taipei).

#### *Phylogenetic inference*

We sequenced the amplified fragments at the mitochondrial locus *cyt b* on the full sample set and those at the nuclear loci, *RAG1* and *Dyst* on taxonomically and spatially representative subsets. Chromatograms were examined using the viewing and editing features of GENEIOUS R v. 6.1.8 (Kearse *et al.* 2012). For each gene, sequences were first automatically aligned using the GENEIOUS homemade alignment method with default values. The alignments were subsequently checked by eye and adjusted manually when necessary. Indels were treated as missing data. All the sequences were deposited in GenBank under accession numbers KU553479 to KU554420 (Supplementary Table S1). We built two matrices of sequences: a mitochondrial one (*cyt b*) and a nuclear one which consisted of the concatenated *RAG1* and *Dyst* sequences. The *RAG1* and *Dyst* markers were also considered separately. For single-gene tree reconstructions, we included all *Abudedefduf* sequences available from GenBank in order to verify their phylogenetic position.

For each alignment, we determined the best partition scheme and the corresponding models of nucleotide substitution that best fit the data using PARTITIONFINDER v. 1.1.1 (Lanfear *et al.* 2012). Coding sequences (*cyt b* and *RAG1*) were partitioned according to codon positions (first, second and third position). In PARTITIONFINDER, likelihood scores for all the models implemented in the phylogenetic reconstruction program MRBAYES v. 3.2.6 (Ronquist *et al.* 2012) were computed and we selected the best-fit models according to the Bayesian Information Criterion associated with the 'greedy search algorithm' option. We then conducted phylogenetic analyses based on maximum likelihood and Bayesian methods using the programs, RAxML v. 8 (Stamatakis, 2014) and MRBAYES, respectively. For Bayesian inference, two independent runs (each with four Markov chains) were made simultaneously for 10 millions of generations and sampled every 1000 generations. We then evaluated the convergence within and across runs. The first 25 % iterations were discarded as burn-in and the remainder was used to generate a 50% majority rule consensus tree. For Maximum Likelihood analyses, the ML search was run simultaneously with an automatic bootstrapping procedure ('autoMRE' criteria) on each gene partition and following the GTRCAT approximation. Computations were run online through the Cypress Science Gateway portal ([www.phylo.org](http://www.phylo.org); Miller *et al.* 2010).

#### *Genetic diversity, population differentiation and demography*

We first focused on mitochondrial sequences (*cyt b*) to compute parameters characterizing the genetic diversity of *A. sexfasciatus* and *A. vaigiensis* lineage A populations, including the number of polymorphic sites per sequence ( $S$ ), the nucleotide diversity ( $\pi$ ; Tajima, 1989), and the haplotype diversity ( $H_d$ ; Nei, 1987). To investigate spatial restriction in gene flow, genetic structure was assessed according to the AMOVA framework (Excoffier *et al.* 1992). We investigated the proportion of genetic variance explained by the differences among regions (considered as groups): Indian Ocean/Coral Triangle region/western Pacific ( $F_{CT}$ ), and among localities within regions ( $F_{SC}$ ) in each of these two species. The AMOVA has been primarily designed for the intra-specific level. However, the degree of inter-specific genetic differentiation at the mitochondrial locus between Coral Triangle region populations of *A. sexfasciatus* and *A. vaigiensis* lineage A was low compared to what it was at the intra-

specific level between regional subgroups of *A. sexfasciatus* (see Results). Therefore, we ran a second AMOVA to obtain a comparative quantification of the genetic differentiation between the two species (considered as groups,  $F_{CT}$ ), at the mitochondrial locus, in the Coral Triangle region localities where they cooccurred in our sampling (*i.e.* around Taiwan) and between these localities within each species ( $F_{SC}$ ).

Departure from the mutation/drift equilibrium was estimated with Tajima's (1993)  $D$  and Fu's (1997)  $F_s$ . These two indices were initially conceived to detect deviation from selective neutrality but have been also shown to be sensitive to modifications in demographical dynamics. Population genetics analyses were done in Arlequin v. 3.5 (Excoffier *et al.* 2005). To visualise with more detail the genetic differentiation between *A. sexfasciatus* and *A. vaigiensis* lineage A, we reconstructed minimum-spanning mitochondrial haplotype networks using PopART (University of Otago, <http://popart.otago.ac.nz>). We checked for shared haplotypes between these two species.

The analyses related to genetic diversity were done on each nuclear marker separately (*RAG1* and *Dyst*). The most likely haplotypes were first determined using the algorithm PHASE (Stephens & Donelly, 2001) implemented in DNAsP v. 5.10.1 (Librado & Rozas, 2009). To optimize haplotype inference without introducing false-positives, the PHASE thresholds were lowered to 0.6 as recommended in Garrick *et al.* (2010). For each of the two nuclear loci, we ran three replicate runs and changed the seed value each time.

#### *Inter-specific gene flow*

Information obtained from the above analyses led us to characterize the joint effect of time and gene flow on the divergence between the weakly differentiated populations of the Coral Triangle region of *A. sexfasciatus* and *A. vaigiensis* A. We used the Bayesian coalescent-based estimation implemented in IMa2 (Hey & Nielsen 2007; Hey 2010) to run this analysis on two populations. The first population consisted of the pooled samples of *A. vaigiensis* A, excluding the two genetically differentiated populations of the Maldives and Red Sea (see Results), while the other population consisted of the pooled samples of *A. sexfasciatus* from the Coral Triangle region (*i.e.* excluding the genetically differentiated Indian Ocean localities and Moruroa). The ‘isolation-with-migration’ model implemented in IMa2 assumes that the two populations (hereafter populations 1 and 2) of effective size  $\theta_1$  and  $\theta_2$  diverged from an ancestral population (of effective size  $\theta_A$ ) at time  $t$  and then exchanged genes at rates  $m_1$  (from population 1 to population 2) and  $m_2$  (from population 2 to population 1). The foregoing six parameters are scaled by mutation rate and therefore need to be converted to derive biologically realistic demographic estimates. Here,  $2.N.m$  corresponds to the number of effective migrants with  $N$  being the effective population size or the number of diploid individuals and  $m$  being the mutation rate (Wright 1931).

The IMa2 analysis was basically run on the three locus data set. Additional partitions (mitochondrial / nuclear) were also considered. For the mitochondrial gene, the mutation rate was set to  $1.3 \times 10^{-8}$  mutations per site per year (DiBattista *et al.* 2013) under the HKY model (Hasegawa *et al.* 1985) and an appropriate inheritance scalar of 0.25. For the nuclear genes, we first used IMgc (Woerner *et al.* 2007) to do a ‘four-gamete test’ in order to detect and exclude potentially recombining blocks within nuclear loci. The mutation rate was set to  $9.7 \times 10^{-10}$  for the autosomal exon *RAG1* (Lessios 2008) and  $7.5 \times 10^{-9}$  for the autosomal intron *Dyst*, following Eytan & Hellberg (2010), under the Infinite-sites model (Kimura 1969). We assumed an inheritance scalar of 1 for both nuclear loci. Preliminary tests were run to characterize prior ranges that encompassed the entire distribution of each parameter estimates for use in subsequent runs. We then ran additional replicate runs with refined priors distributions and different random number seeds. Parameter estimates were congruent across these runs. A last series of runs consisting of 25 to 100 million steps (sampled every 1000 steps) with 0.5 million burn-in steps was carried out to obtain the final estimates. Convergence was checked by visually inspecting the trend lines of all parameters. To assess whether the complete model with gene flow was better supported than the model without gene flow, we tested whether simpler demographic models (with no gene flow) fitted the data. To do so, we followed the nested model approach in the ‘load-tree mode’ available in IMa2 according to the procedure described in Hey & Nielsen (2007).

## Results

### Description of sequence datasets

A total of 942 sequences were generated in the course of this study (484 at locus *cytb*, 224 at locus *RAG1* and 234 at locus *Dyst*, see details in Supplementary Table S1, Supporting information). Once additional sequences retrieved from GenBank had been included, the mitochondrial data set consisted of 522 *cytb* gene sequences aligned over 1161 bp. The nuclear data set consisted of 246 sequences at locus *RAG1*, aligned over 1456 bp and 234 sequences at locus *Dyst* aligned over 1211 bp. The two nuclear sequence data sets were concatenated into a single matrix of 204 individuals x 2667 nucleotide sites.

### Mitochondrial gene-based phylogeny

At the mitochondrial locus, six of the species surveyed (*Abudedefduf bengalensis*, *A. saxatilis*, *A. septemfasciatus*, *A. sordidus*, *A. sparoides* and *A. taurus*) were monophyletic and presented low levels of intraspecific divergence (Fig. 2A). In contrast, *A. vaigiensis* was polyphyletic so we coined the three lineages sampled in this study A, B and C. *Abudedefduf sexfasciatus* was paraphyletic with respect to *A. vaigiensis* A (Figs 1B, 2 and S1, Supporting information). *Abudedefduf vaigiensis* A was the most abundant and widely distributed in our total *A. vaigiensis* samples; its representatives were spread within the clade that also contains the haplotypes from western Pacific *A. sexfasciatus* (Fig. 2A). *Abudedefduf vaigiensis* B consisted of two geographically separated sublineages with one containing the samples collected from Taiping Island in the South China Sea, from Little Liuqiu Island in the Taiwan strait, West Papua in the Coral Triangle and from Guam Island in the West Pacific and the other one consisting of a unique individual from the western Indian Ocean. *Abudedefduf vaigiensis* B was far less abundant than *A. vaigiensis* A with only six individuals of 217 *A. vaigiensis* sampled. The monophyletic *A. vaigiensis* C was recorded only from the Coral Triangle (with five individuals from the Bali Strait and six individuals from Caohagan Island in the Cebu Sea) (Fig. 1B). A fourth lineage within *A. vaigiensis* was represented by a single *cytb* gene sequence from Christmas Island (GenBank: AY208557; Quenouille *et al.* 2004).

The *A. sexfasciatus* samples included in the mitochondrial gene analysis can be subdivided into three phylogeographic or regional sublineages. One of them included all the individuals from the western Indian Ocean. A second one included all 12 individuals from Moruroa, one individual from New Caledonia, and one individual from West Papua. A third and the largest one, comprised the 124 remaining *A. sexfasciatus* from the Coral Triangle and the western Pacific Ocean, along with all 204 *A. vaigiensis* A individuals.

### Nuclear gene-based phylogeny

The nuclear gene-based phylogenetic tree also supported the monophyly of most *Abudedefduf* species surveyed and showed little evidence of intraspecific divergence in most of the species that were sampled from both the Indian and Pacific Oceans. An exception was *A. septemfasciatus* for which a phylogeographic break (Indian vs. western Pacific partition) was observed (Figs 2B and S1, Supporting information). Interestingly, twelve of the 20 phenotypically identified *A. vaigiensis* individuals from the Maldives fell into the *A. sexfasciatus* Indian Ocean clade at both nuclear loci. Three distinct lineages (A, B, C) were also observed in *A. vaigiensis* in the nuclear gene-based trees. In contrast with the mitochondrial tree, the samples from previously assigned *A. vaigiensis* A exclusive of the twelve Maldivian *A. vaigiensis* individuals mentioned above formed a monophyletic group which is completely separated from *A. sexfasciatus* (Fig. 2B). In *A. sexfasciatus*, nuclear markers confirmed the phylogeographic pattern as revealed from the results of mitochondrial gene analysis by dividing the samples into three regional sublineages: one from the western Indian Ocean, one from the Coral Triangle region and one from the western Pacific Ocean. There was only weak intraspecific variation within *A. vaigiensis* A across the whole Indo-West Pacific. The fast-evolving intronic marker (*Dyst*) provided better phylogenetic resolution than

the exonic one (*RAG1*). *RAG1* failed to fully resolve the terminal part of the tree, whereas *Dyst* did and even revealed an intraspecific structure with a distinct lineage including the Red Sea population in *A. vaigiensis* A (Fig. S1, Supporting information). The placement of *A. saxatilis* indicated that *A. sexfasciatus* and *A. vaigiensis* are not sister species which contrasts with the proximity we found between *A. sexfasciatus* and *A. vaigiensis* A in the mitochondrial gene tree.

#### *Genetic diversity, population genetic structure and demography*

*Abudeodus sexfasciatus* and *A. vaigiensis* lineage A displayed different patterns of intra-specific genetic diversity and differentiation at the mitochondrial locus. In *A. sexfasciatus*, we found a substantially reduced genetic diversity in the remote atoll of Moruroa where all 12 individuals sampled shared a unique *cytb* haplotype (Table 1). In both species, populations from several localities displayed significant negative Tajima's *D* and Fu's *F<sub>S</sub>* values. This trend was more obvious in *A. vaigiensis* lineage A than in *A. sexfasciatus*. It was observed, to a lesser extent, at locus *RAG1* but not at locus *Dyst* (Supplementary Table S3A; 3B).

The overall estimates of genetic differentiation at the mitochondrial locus (*cytb*) in *A. sexfasciatus* was  $F_{ST} = 0.786$  ( $P < 10^{-4}$ ). Most of the genetic differentiation occurred between regions: Indian Ocean, Coral Triangle region and western Pacific ( $F_{CT} = 0.669$ ;  $P = 0.03$ ). Genetic differentiation was also considerable between localities within regions ( $F_{SC} = 0.351$ ;  $P < 10^{-4}$ , Table 2). Within *A. vaigiensis* lineage A, there was only negligible differentiation among populations ( $F_{ST} = 0.020$ ,  $P = 0.77$ ;  $F_{SC} = -0.001$ ,  $P = 0.95$ ) even though it was significant between oceans with  $F_{CT} = 0.033$ ,  $P = 0.02$  (Table 2). Based on the AMOVA, the degree of genetic differentiation was nonsignificant ( $F_{CT} = 0.283$ ,  $P = 0.10$ ) between *A. sexfasciatus* and *A. vaigiensis* lineage A where the two morpho-species were found in sympatry ( $F_{SC} = -0.009$ ;  $P = 0.58$  and  $F_{ST} = 0.276$ ;  $P < 10^{-4}$ ) at this locus. This confirmed that genetic differentiation was lower between the two species in the Coral Triangle region than within *A. sexfasciatus* across regions.

Three mitochondrial haplotypes, of 137 in total, were shared by *A. sexfasciatus* and *A. vaigiensis* A individuals (Fig. 3A). One of these three haplotypes was common and it characterized several individuals from the two species and from different localities across the Indo-West Pacific. The two other haplotypes were less common in at least one of the two species. Overall, based on mitochondrial haplotypes, the differentiation between *A. sexfasciatus* and *A. vaigiensis* A was weak in the Coral Triangle region. The mitochondrial haplotype network confirmed that Indian Ocean and western Pacific haplotypes were relatively distinct from the Coral Triangle region ones (with 13 and 6 mutations, respectively; i.e. 1.4% and 0.6% of pairwise sequence divergence between groups). One individual of *A. sexfasciatus* from West Papua and one from New Caledonia had haplotypes similar to those found in Moruroa, whereas the other haplotypes present at these localities had affinities with those sampled in the Coral Triangle region. At the nuclear loci, several individuals from the Maldives formally identified as *A. vaigiensis* according to colour pattern possessed alleles characteristic of *A. sexfasciatus* (Fig. 3B, C). Elsewhere, the two species were clearly distinct at locus *Dyst* with some haplotypes from the Red Sea at an intermediate position between *A. vaigiensis* A and *A. sexfasciatus* (Fig. 3C).

#### *Coalescence and inter-specific gene flow*

Bayesian coalescent-based analyses highlighted contrasting patterns in the timing of divergence and gene flow between Coral Triangle region populations of *A. sexfasciatus* and *A. vaigiensis* A (Table 3; Figs 4 and S2, Supporting information). Splitting time ( $t$ ) between the two species based on both mitochondrial and nuclear information was found to be recent: 0.5 Myrs. Further examination suggested that  $t$  (as well as the times of the most recent common ancestor, tMRCA) were likely to be much younger at the mitochondrial locus than at nuclear loci (<0.2 vs.  $\geq 0.9$  Myrs, respectively). The effective migration rates between the two species ( $5 \times 10^{-7}$  ind./year, Table 3) were low but non-null. This result was corroborated by the model comparison procedure

according to which the complete model (allowing gene flow) better fit the data than the model without gene flow (Supplementary Table S4, Supporting information).

According to the shapes of the relative likelihood distributions, the overall effective migration rate could be marker-dependent. Interspecific gene flow was predominantly driven by mitochondrial gene flow, whereas nuclear gene flow was negligible (Figs 3 and S2, Supporting information). Based on effective migration rate values, we also noted that the estimates are consistent with a slight asymmetry in gene flow between the two species. Although gene flow should be considered as weak in both directions, it seems it preferentially occur from *A. vaigiensis* A to *A. sexfasciatus* (Table 3; Supplementary Table S4). Inferred effective population sizes were high in both species (on the order of 104 or 105 individuals), but the estimates differed according to the locus considered.

## Discussion

In this study, we reported mito-nuclear discordance in two sergeant fishes widely distributed in the Indo-West Pacific which we ascribed to interspecific gene flow. We relied upon a geographic sampling spread across the Indo-Pacific barrier and including peripheral localities (the Mozambique Channel, the Red Sea, Moruroa) and used three genetic markers assumed to evolve at different paces, to span different time frames of the evolutionary history of this system. The phylogenetic results enabled us to characterize the polyphyly of *A. vaigiensis* thus confirming its existence of cryptic diversity (see Quenouille *et al.* 2011). We drew the spatial distribution of our identified genetic lineages. Our results also showed contrasting spatial patterns of intraspecific differentiation between the phylogeographically structured *A. sexfasciatus* and the more homogeneous *A. vaigiensis* A. An examination of mito-nuclear patterns of variation then enabled us to characterize complete lineage sorting between the two species at the fast-evolving nuclear intronic locus (*Dyst*). This contrasted with a very likely incomplete lineage sorting at the nuclear exonic locus (*RAG1*) for which the whole terminal part of the phylogenetic tree we inferred was poorly resolved. We also found evidence of weak but nonnegligible interspecific gene flow at the mitochondrial (*cytb*) locus.

### *Systematics of the sergeant fishes*

Most species in the genus *Abudefduf* were monophyletic except *A. vaigiensis* (polyphyletic) and *A. sexfasciatus* (paraphyletic). We found that *A. vaigiensis* actually corresponds to at least three phenotypically similar but genetically divergent lineages (A, B and C; Fig. 2A,B). A previous study by Quenouille *et al.* (2011) reported three *A. vaigiensis* taxa (A, B and C), two of which (A and B, not C) match the ones characterized in this study. However, Quenouille *et al.* (2011) did not further discuss this result that was based on the distinctiveness of mitochondrial cytochrome b and ATP synthase 8 and 6 gene sequences obtained from two single individuals (sequenced and deposited in GenBank by Quenouille *et al.* 2004). In this study, we confirmed the existence of *A. vaigiensis* A and B taxa as we can unambiguously assign several individuals to each of them based on both mitochondrial and nuclear information. Cryptic diversity in reef fishes often concerns populations whose geographic distributions do not overlap, or weakly so (Taylor & Helberg 2005; Drew *et al.* 2010; Leray *et al.* 2010). In our study, this was the case of *A. sexfasciatus* which showed three regional sublineages (Indian Ocean / Coral Triangle region / western Pacific). The degree of genetic differentiation among these sublineages (up to 1.5% of divergence between the Indian Ocean sublineage and the two others at *cytb*) was sufficiently high to potentially assign them the rank of subspecies. In contrast, the three cryptic lineages uncovered in *A. vaigiensis* were substantially different from each other for both mitochondrial and nuclear genes but co-occurred widely (i.e. 8.8 to 9% of divergence between lineages A, B and C at *cytb*). *Abudefduf vaigiensis* A was the most frequently sampled and occurred throughout the Indo-West Pacific. *Abudefduf vaigiensis* B was geographically widespread too but seemed to be relatively uncommon. The *cytb* gene sequence of a single individual from Guam sequenced by Quenouille *et al.* (2004) (GenBank AY208561) clustered with *A. vaigiensis* B haplotypes (Fig. 2A). *Abudefduf*

*vaigiensis* C was here reported for the first time and is distinct from the cryptic taxon “C” reported in Quenouille et al. (2011) based on a single sample from Christmas Island that was sequenced at the cytochrome *b* and the ATP synthase 8 and 6 gene loci (GB Accessions: AY208557 and AY208418, respectively). The geographic distribution of *A. vaigiensis* C is possibly limited to the Coral Triangle. Our results thus revealed cryptic diversity in *A. vaigiensis* and indicated that the number of species was underestimated in the genus *Abudedefduf*. Altogether, these results on *A. sexfasciatus* and *A. vaigiensis* call for further investigations on the systematics and taxonomy of the genus, to be conducted with more samples and more nuclear gene markers.

#### *Phylogeography of A. sexfasciatus and A. vaigiensis lineage A*

Among the *Abudedefduf* species sampled in both the Indian Ocean and the Pacific Ocean, *A. vaigiensis* A showed nonsignificant population structure distribution-wide at the mitochondrial locus ( $F_{ST} = 0.020$ ,  $P = 0.78$ ). Genetic homogeneity across the entire Indo-West Pacific has been reported in other reef fishes (Klanten et al. 2007; Horne et al. 2008; Reece et al. 2010, 2011). Such broad scale homogeneity suggests high levels of contemporary intraspecific gene flow or insufficient time to observe divergence. In contrast, substantial levels of intraspecific genetic structure were observed among regional populations as well as within regions in *A. sexfasciatus*. The genetic discontinuities (Indian Ocean/Pacific Ocean) matched the well-known Indo-Pacific barrier (Gaither & Rocha 2013; Borsa et al. 2016). Geographic structure also reflected the isolation of peripheral insular populations, a pattern that has been documented for reef fishes from Hawaii (Eble et al. 2011; Gaither et al. 2014), the Society Islands (Liu et al. 2014) and the Marquesas (Planes & Fauvelot 2002; Gaither et al. 2010). Genetic diversity at the mitochondrial locus was extremely low in the Moruroa population of *A. sexfasciatus*, a pattern previously observed for populations from isolated peripheral habitats (Liu et al. 2014). The *A. sexfasciatus* population from Moruroa was monomorphic at the *cytb* locus while populations elsewhere had high gene and nucleotide diversities. Remote populations may be particularly sensitive to bottleneck events due to relatively low effective population sizes and/or lack of migrants from external genetic pool that could efficiently maintain genetic diversity.

We hypothesize that populations of *A. vaigiensis* A diverged from their most recent common ancestor more recently than that of *A. sexfasciatus*. Stronger negative Tajima’s *D* and Fu’s *F<sub>S</sub>* values in *A. vaigiensis* A support recent population demographic expansion in this species. However, in contrast with the overall lack of genetic structure at loci *cytb* and *RAG1*, the Red Sea population of *A. vaigiensis* A showed evidence of differentiation at the fast-evolving intronic locus *Dyst*, suggesting a more complex demographic history and a role of geographic isolation in the peripheral populations of this species too.

We documented contrasting mito-nuclear phylogenetic patterns between the two closely related (but not sister) species, *A. sexfasciatus* and *A. vaigiensis* A. These two species are genetically distinct based on the nuclear information but *A. vaigiensis* A and the populations of *A. sexfasciatus* from the Coral Triangle region were represented by the same mitochondrial lineage. We clearly identified distinct regional lineages in *A. sexfasciatus* but at the same time found shared mitochondrial haplotypes between the sublineage of *A. sexfasciatus* from the Coral Triangle region and *A. vaigiensis* A. This strongly supports the hypothesis of gene flow over that of incomplete lineage sorting.

#### *Proposed evolutionary scenario*

The comparative analysis of the patterns of genetic diversity and differentiation within and between *A. sexfasciatus* and *A. vaigiensis* lineage A led us to propose the following evolutionary scenario (Fig. 5). The split between *A. sexfasciatus* and *A. vaigiensis* A would be ancient enough to be unambiguously characterized at a relatively fast-evolving nuclear intronic locus such as *Dyst*. A lack of resolution associated with weak phylogenetic signal at locus *RAG1* suggests incomplete lineage sorting at this slower-evolving nuclear exon. At these nuclear loci, the estimates of splitting times and tMRCA are consistent with the existence of the last

common ancestor between the two species in the late Miocene (between 10 and 5 My B.P.; Litsios *et al.* 2012; Frédéric *et al.* 2013) or perhaps, even more recently according to our coalescent-based estimates.

In the Coral Triangle region, the non-significant level of divergence ( $F_{CT} = 0.283$ ,  $P = 0.10$ ) between *A. sexfasciatus* and *A. vaigiensis* lineage A at the mitochondrial locus (*cytb*) contrasted with the consensus pattern observed at nuclear loci. The estimated splitting time at locus *cytb* was at least about ten times younger than the one at nuclear genes. The signal that we detected as well as evidence of shared haplotypes is thus consistent with gene flow after secondary contact following at least one initial stage of divergence between the two species. This scenario has been previously invoked to explain the sympatry of sibling species in other reef fishes from various families [e.g. snappers (Lutjanidae): Gaither *et al.* 2010; wrasses (Labridae): Choat *et al.* 2012 or damselfishes (Pomacentridae): Quenouille *et al.* 2011] in which cross-species hybridization sometimes occurs (e.g. Hobbs *et al.* 2009; DiBattista *et al.* 2015). This does not rule out more local ecological and biological factors to explain the genetic introgression between *A. sexfasciatus* and *A. vaigiensis* lineage A in part of their range.

#### *Causes of inter-specific gene flow*

In sympatric and closely related reef fish species, traits associated with sexual selection, for example coloration in hamlets, *Hypoplectrus* spp. (Puebla *et al.* 2007) or pygmy angelfishes, *Centropyge* spp. (DiBattista *et al.* 2012; Bowen *et al.* 2013; and references therein) or natural selection (Rocha *et al.* 2005; Bowen *et al.* 2013 and references therein), generally act as species recognition signals and impede hybridization. In some situations however, divergence may be too recent to have led to complete pre- or postzygotic isolation. In the genus *Abudefduf*, Coleman *et al.* (2014) proposed that similarity in colour patterns might be one of the factors responsible for the failure of species recognition and for the subsequent hybridization between *A. abdominalis* and *A. vaigiensis* after a recent introduction of the former species to Hawaii. In our case, however, colour pattern differences are obvious between *A. sexfasciatus* and *A. vaigiensis* A, whereas they are not between lineages A, B and C of *A. vaigiensis*. Yet, there was no evidence of gene flow between the three cryptic lineages of *A. vaigiensis* (A, B and C), whereas we found evidence of it between *A. sexfasciatus* and *A. vaigiensis* A. Evidence of interspecific gene flow between *A. sexfasciatus* and *A. vaigiensis* A in part of their range thus suggests that factors other than coloration may have allowed hybridization.

*Abudefduf sexfasciatus* and *A. vaigiensis* A often co-occur in strict sympatry and seem to display similar ecological preferences (Aguilar-Medrano & Barber 2016; personal field observations). Although the niches of the two species broadly overlap, *A. vaigiensis* which is benthopelagic could be more generalist than *A. sexfasciatus* which is pelagic. Several ecological factors have been invoked as catalysts of hybridization and their putative effects have been discussed by Coleman *et al.* (2014) for the two *Abudefduf* species that occur in Hawaii (see also Yaakub *et al.* 2006; Montanari *et al.* 2012, 2014, 2016). Two species that form heterospecific social groups and have synchronized spawning seasons may experience external fertilization accidentally involving gametes of both species. However, this hypothesis is unlikely in damselfishes which lay demersal eggs (Gainsford *et al.* 2015) as do *A. sexfasciatus* and *A. vaigiensis*. Relative differences in population density may also favour hybridization. In the field, we often observed local differences in the ratio of *A. sexfasciatus* / *A. vaigiensis*, but the dominant species was not always the same depending on the locality. Our genetic results further suggest that interspecific gene flow predominantly concerns mitochondrial genes and could be asymmetric: it was slightly higher from *A. vaigiensis* A to *A. sexfasciatus*. As mitochondrial gene inheritance in fishes is assumed to occur through females, this result would be consistent with interspecific mating predominantly involving females of *A. vaigiensis* A and males of *A. sexfasciatus*. An accurate characterization of hybrids is still required to elucidate the biological mechanisms underpinning hybridization in this system, given that putative hybrids seem to be rare. We noted that most of the *A. vaigiensis* A individuals that shared mitochondrial haplotypes with *A. sexfasciatus* (but displayed nuclear characteristics of *A. vaigiensis*) were juveniles from particular insular localities. This could be consistent with hybrids having lower survival rate.

Only one Indian Ocean locality (Maldives) showed adult individuals with unambiguous *A. vaigiensis* phenotypes associated with *A. sexfasciatus* genotypes at both nuclear and mitochondrial loci. Deep introgression could explain such a situation.

### Conclusive remarks

Sergeants provide an opportunity to further investigate how hybridization and its evolutionary causes work at a trans-oceanic scale; in widely sympatric species and with apparent similar ecological niches. Between *A. sexfasciatus* and *A. vaigiensis* A, inter-specific gene flow occurs not only at spatially restricted hotspots but virtually across the whole western Pacific. Unlike discontinuous island-dwelling systems in which each sampling localities can be considered as a distinct replicate, this configuration could be particularly suitable to better understand what mechanisms prevent species from collapsing in the face of gene flow across unrivalled large scale continuums.

### Acknowledgments

Specimens from the western Indian Ocean were collected during the IRD/CNRS/TAAF-sponsored PHYLIP campaign on board RV *Marion Dufresne*, April 2011; those from West Papua were collected during expedition EWiN sponsored by LIPI (Indonesia), on board KR *Baruna Jaya VIII*, November 2007; those from New Caledonia were collected during the FRB-sponsored RESICOD project, December 2010; those from Moruroa were obtained from the IRD-sponsored ISOTOPE campaign on board RV *Alis*, June 2006; those from Papua New Guinea were collected during the joint "Our Planet Reviewed" PAPUA NIUGINI Biodiversity Expedition of Pro-Natura International, MNHN, IRD, and UPNG , November 2012, and during the MADEEP (April 2014) and Kavieng (September 2014) campaigns on board RV *Alis* under Taiwan France marine diversity exploration and evolution of deep-sea fauna (TFDeepEvo). We thank M. Adrim (LIPI, Jakarta), S. Bahri and A. Sembiring (IBRC, Denpasar), K. Conway (TAMU, College Station), J. DiBattista and J.-P. Hobbs (Curtin University, Perth), J.-D. Durand (IRD, Montpellier), H.-C. Ho (NMMBA, Pingtung), F. Giancarlo, (Proyek BioKor, Denpasar), M. Kulbicki (IRD, Perpignan), G. Mou-Tham and C. Paillon (IRD, Nouméa), T. Gurevich Raguso (Inter-university Institute for Marine Science, Eilat), K.-N. Shen (NTOU, Taipei), and T.H. Sinclair-Taylor (KAUST, Thuwal) for their participation in the collection or provision of specimens. We also thank the University of Milano-Bicocca Marine Research and High Education Centre in Magoodhoo, Maldives, the Ministry of Fisheries and Agriculture of the Republic of Maldives and the community of Maghoodhoo, Faafu Atoll, and the headquarters of Kenting National Park and Dongsha Atoll National Park of Taiwan. We are grateful to J. Yang and the Chen's Lab members for help in the laboratory and management of specimens. S. Lavoué and Y. Bourgeois provided helpful comments on earlier versions of the manuscript. This work was supported by research grants from the Ministry of Science and Technology of Taiwan (MOST 101-2611-M-002-016-MY3, MOST103-2811-M-002-109 and MOST 104-2611-M-002-002-MY3 to WJC). JAMB's salary was supported by National Taiwan University and MOST post-doctoral fellowships.

### References

- Abbott R, Albach D, Ansell S *et al.* (2013) Hybridization and speciation. *Journal of Evolutionary Biology*, **26**, 229–246.
- Aguilar-Medrano R, Barber PH (2016) Ecomorphological diversification in reef fish of the genus *Abudefduf* (Perciformes, Pomacentridae). *Zoomorphology*, **135**, 103-114.
- Allen GR (2008) Conservation hotspots of biodiversity and endemism for Indo-Pacific coral reef fishes. *Aquatic Conservation: Marine and Freshwater Ecosystems*, **18**, 541-556.
- Barton NH, Hewitt GM (1985) Analysis of hybrid zones. *Annual Review of Ecology and Systematics*, **16**, 113-148.

- Barton NH, Hewitt GM (1989) Adaptation, speciation and hybrid zones. *Nature*, **341**, 497-503.
- Borsig P, Durand J-D, Chen W-J, Hubert N, Muths D, Mou-Tham G, Kulbicki M (2016) Comparative phylogeography of the western Indian Ocean reef fauna. *Acta Oecologica*, **72**, 72-86.
- Bowen BW, Rocha LA, Toonen RJ, Karl SA, ToBo Lab (2013) The origins of tropical marine biodiversity. *Trends in Ecology and Evolution*, **28**, 359-366.
- Bowen BW, Shanker K, Yasuda N, Malay MC(M)D, von der Heyden S, Paulay G, Rocha LA, Selkoe KA, Barber PH, Williams ST, Lessios HA, Crandall ED, Bernardi G, Meyer CP, Carpenter KE, Toonen RJ (2014) Phylogeography unplugged: comparative surveys in the genomic era. *Bulletin of Marine Science*, **90**, 13-46.
- Briggs JC, Bowen BW (2012) A realignment of marine biogeographic provinces with particular reference to fish distributions. *Journal of Biogeography*, **39**, 12–30.
- Briggs JC, Bowen BW (2013) Marine shelf habitat: biogeography and evolution. *Journal of Biogeography*, **40**, 1023–1035.
- Carpenter KE, Barber PH, Crandall ED, Ablan-Lagman MCA, Ambariyanto, Mahardika GN, Manjaji-Matsumoto BM, Juinio-Meñez MA, Santos MD, Starger CJ, Toha AHA (2011) Comparative phylogeography of the Coral Triangle and implications for marine management. *Journal of Marine Biology*, **2011**, 14.
- Choat JH, Klanten OS, van Herwerden L, Robertson DR, Clements KD (2012) Patterns and processes in the evolutionary history of parrotfishes (Family Labridae). *Biological Journal of the Linnean Society*, **107**, 529-557.
- Crandall ED, Riginos C (2014) Magnificent dimensions, varied forms, and brilliant colors: the molecular ecology and evolution of the Indian and Pacific oceans. *Bulletin of Marine Science*, **90**, 1-11.
- Coleman RR, Gaither MR, Kimokeo B et al. (2014) Large-scale introduction of the Indo-Pacific damselfish *Abudefduf vaigiensis* into Hawai'i promotes genetic swamping of the endemic congener *A. abdominalis*. *Molecular Ecology*, **23**, 5552-5565.
- Cooper WJ, Smith LL, Westneat MW (2009) Exploring the radiation of a diverse reef fish family: Phylogenetics of the damselfishes (Pomacentridae), with new classifications based on molecular analyses of all genera. *Molecular Phylogenetics and Evolution*, **52**, 1-16.
- Cooper WJ, Albertson RC, Jacob RE, Westneat MW (2014) Re-description and reassignment of the damselfish *Abudefduf luridus* (Cuvier, 1830) using both traditional and geometric morphometric approaches. *Copeia*, **2014**, 473-480.
- DiBattista JD, Rocha LA, Craig MT, Feldheim KA, Bowen BW (2012) Phylogeography of two closely related Indo-Pacific butterflyfishes reveals divergent evolutionary histories and discordant results from mtDNA and microsatellites. *Journal of Heredity*, **103**, 617–629.
- 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.
- DiBattista JD, Rocha LA, Hobbs J-PA, He S, Priest MA, Sinclair-Taylor TH, Bowen BW, Berumen ML (2015) When biogeographical provinces collide: hybridization of reef fishes at the crossroads of marine biogeographical provinces in the Arabian Sea. *Journal of Biogeography*, **42**, 1601-1614.
- Drew JA, Allen GR, Erdmann MV (2010) Congruence between mitochondrial genes and color morphs in a coral reef fish: population variability in the Indo-Pacific damselfish *Chrysiptera rex* (Snyder, 1909). *Coral Reefs*, **29**, 439-444.
- Eble JA, Toonen RJ, Sorensen LL, Basch LV, Papastamatiou YP, Bowen BW (2011) Escaping paradise: larval export from Hawaii in an Indo-Pacific reef fish, the Yellow Tang (*Zebrasoma flavescens*). *Marine Ecology Progress Series*, **428**, 245-258.
- Eschmeyer WN, Fricke R, Van der Laan JD (2016) Catalog of Fishes: genera, species, references. <http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp>. Electronic version accessed 01 September 2016.
- Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes - Application to human mitochondrial-DNA restriction data. *Genetics*, **131**, 479-491.

- Excoffier L, Laval G, Schneider S (2005) ARLEQUIN (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, **1**, 47–50.
- Eytan RI, Hellberg ME (2010) Nuclear and mitochondrial sequence data reveal and conceal different demographic histories and population genetic processes in Caribbean reef fishes. *Evolution*, **64**, 3380-3397.
- Fraïsse C, Roux C, Welch JJ, Bierne N (2014) Gene-flow in a mosaic hybrid zone: is local introgression adaptive? *Genetics*, **197**, 939-951.
- Froese R, Pauly D (2015) FishBase, World Wide Web electronic publication, www.fishbase.org, version (10/2015).
- Frédéric B, Sorenson L, Santini F, Slater GJ, Alfaro ME (2013) Iterative ecological radiation and convergence during the evolutionary history of damselfishes (Pomacentridae). *American Naturalist*, **181**, 94-113.
- Fu YX (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, **147**, 915-925.
- Gainsford A, van Herwerden L, Jones GP (2015) Hierarchical behaviour, habitat use and species size differences shape evolutionary outcomes of hybridization in a coral reef fish. *Journal of Evolutionary Biology*, **28**, 205-222.
- 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.
- Gaither MR, Bowen BW, Bordenave T-R, Rocha LA, Newman SJ, Gomez JA, van Herwerden L, Craig MT (2011) 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, 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, Schultz JK, Bellwood DR, Pyle RL, DiBattista JD, Rocha LA, Bowen BW (2014) Evolution of pygmy angelfishes: recent divergences, introgression, and the usefulness of color in taxonomy. *Molecular Phylogenetics and Evolution*, **74**, 38-47.
- Garrick RC, Sunnucks, Dyer RJ (2010) Nuclear gene phylogeography using PHASE: dealing with unresolved genotypes, lost alleles, and systematic bias in parameter estimation. *BMC Evolutionary Biology*, **10**, 118.
- Hasegawa M, Kishino H, Yano T (1985) Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution*, **22**, 160–174.
- Hey J (2010) Isolation with migration models for more than two populations. *Molecular Biology and Evolution*, **27**, 905-920.
- Hey J, Nielsen R (2007) Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. *Proceedings of the National Academy of Sciences USA*, **104**, 2785-2790.
- Hobbs J-PA, Frisch AJ, Allen GR, van Herwerden L (2009) Marine hybrid hotspot at Indo-Pacific biogeographic border. *Biology letters*, **5**, 258-261.
- Hobbs J-PA, Allen GR (2014) Hybridisation among coral reef fishes at Christmas Island and the Cocos (Keeling) Islands. *Raffles Bulletin of Zoology*, **30**, 220-226.
- 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.
- Horne JB, van Herwerden L (2013) Long-term panmixia in a cosmopolitan Indo-Pacific coral reef fish and a nebulous genetic boundary with its broadly sympatric sister species. *Journal of Evolutionary Biology*, **26**, 783-799.
- Hubert N, Meyer CP, Bruggemann HJ, Guérin F, Komeno RJL, Espiau B, Causse R, Williams JT, Planes S (2012) Cryptic diversity in Indo-Pacific coral-reef fishes revealed by DNA-barcoding provides new support to the Centre-of-Overlap hypothesis. *PloS One*, **7**, e28987.
- Kearse M, Moir R, Wilson A et al. (2012) Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, **28**, 1647-1649.

- Kimura M (1969) The number of heterozygous nucleotide sites maintained in a finite population due to steady flux of mutations. *Genetics*, **61**, 893–903.
- Klanten SO, 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.
- Lanfear R, Calcott B, Ho SYW, Guindon S (2012) PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular Biology and Evolution*, **29**, 1695–1701.
- Le Moan A, Gagnaire P-A, Bonhomme F (2016) Parallel genetic divergence among coastal-marine ecotype pairs of European anchovy explained by differential introgression after secondary contact. *Molecular Ecology*, **25**, 3187–3202.
- 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.
- Lessios HA (2008) The Great American Schism: divergence of marine organisms after the rise of the Central American isthmus. *Annual Review of Ecology, Evolution, and Systematics*, **39**, 63–91.
- Lester SE, Ruttenberg BI (2005) The relationship between pelagic larval duration and range size in tropical reef fishes: a synthetic analysis. *Proceedings of the Royal Society of London B*, **272**, 585–591.
- Librado P, Rozas J (2009) DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, **25**, 1451–1452.
- Litsios G, Sims CA, Wüest RO et al. (2012) Mutualism with sea anemones triggered the adaptive radiation of clownfishes. *BMC Evolutionary Biology*, **12**, 212.
- Litsios G, Salamin N (2014) Hybridization and the diversification in the adaptive radiation of clownfishes. *BMC Evolutionary Biology*, **14**, 245.
- Liu S-YV, 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.
- Mallet J (2005) Hybridization as an invasion of the genome. *Trends in Ecology and Evolution*, **20**, 229–237.
- Miller M, Pfeiffer W, Schwartz T (2010) Gateway Computing Environments Workshop (GCE), 2010. Creating the CIPRES science gateway for inference of large phylogenetic trees, pp. 1–8.
- Montanari SR, van Herwerden L, Prachett MS, Hobbs J-PA, Fugedi A (2012) Reef fish hybridization: lessons learnt from butterflyfishes (genus *Chaetodon*). *Ecology and Evolution*, **2**, 310–329.
- Montanari SR, Hobbs J-PA, Pratchett MS, Bay LK, van Herwerden L (2016) The importance of ecological and behavioural data in studies of hybridisation among marine fishes. *Reviews in Fish Biology and Fisheries*, **26**, 181–198.
- Montanari SR, Hobbs J-PA, Pratchett MS, Bay LK, van Herwerden L (2014) Does genetic distance between parental species influence outcomes of hybridization among coral reef butterflyfishes? *Molecular Ecology*, **23**, 2757–2770.
- Naish T, Powell R, Levy R et al. (2009) Obliquity-paced Pliocene West Antarctic ice sheet oscillations. *Nature*, **458**, 322–328.
- Nei M (1987) *Molecular evolutionary genetics*. Columbia University Press, New York, NY.
- Planes S, Fauvelot C (2002) Isolation by distance and vicariance drive genetic structure of a coral reef fish in the Pacific Ocean. *Evolution*, **56**, 378–399.
- Price T (2008) *Speciation in birds*. Roberts & Company Publishers, Austin, TX.
- Puebla O, Bermingham E, Guichard F, Whiteman E (2007) Colour pattern as a single trait driving speciation in *Hypoplectrus* coral reef fishes. *Proceedings of the Royal Society B*, **274**, 1265–1271.
- Quenouille B, Bermingham E, Planes S (2004) Molecular systematics of the damselfishes (Teleostei: Pomacentridae): Bayesian phylogenetic analyses of mitochondrial and nuclear DNA sequences. *Molecular Phylogenetics and Evolution*, **31**, 66–88.
- 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.

- Reece JS, Bowen BW, Joshi K, Goz V, Larson A (2010) Phylogeography of two moray eels indicates high dispersal throughout the Indo-Pacific. *Journal of Heredity*, **101**, 391-402.
- Reece JS, Bowen BW, Smith DG, Larson A (2011) Comparative phylogeography of four Indo-Pacific moray eels species (Muraenidae) reveals comparable ocean-wide genetic connectivity despite five-fold differences in available adult habitat. *Marine Ecology Progress Series*, **437**, 269-277.
- Rocha LA, Robertson DR, Roman J, Bowen BW (2005) Ecological speciation in tropical reef fishes. *Proceedings of the Royal Society B*, **272**, 573-579.
- Ronquist FM, Teslenko P, van der Mark D *et al.* (2012) MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology*, **61**, 539-542.
- Seehausen O (2004) Hybridization and adaptive radiation. *Trends in Ecology and Evolution*, **19**, 198-207.
- Siddall M, Rohling EJ, Almogi-Labin A *et al.* (2003) Sea level fluctuations during the last glacial cycle. *Nature*, **423**, 853–858.
- Stamatakis A (2014) RAxML Version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, **30**, 1312-1313.
- Stephens M, Donnelly P (2001) A new statistical method for haplotype reconstruction from population data. *American Journal of Human Genetics*, **68**, 978-989.
- Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, **123**, 583-595.
- Taylor MS, Helberg ME (2005) Marine radiations at small geographic scales: speciation in Neotropical reef gobies (*Elacatinus*). *Evolution*, **59**, 374-385.
- Voris HK (2000) Maps of Pleistocene sea levels in Southeast Asia: shorelines, river systems and time durations. *Journal of Biogeography*, **27**, 1153–1167.
- Warren BH, Bermingham E, Bourgeois YXC *et al.* (2012) Hybridization and barriers to gene flow in an island bird radiation. *Evolution*, **66**, 1490-1505.
- Woerner AE, Cox MP, Hammer MF (2007) Recombination-filtered genomic datasets by information maximization. *Bioinformatics*, **23**, 1851-1853.
- Wright S (1931) Evolution in Mendelian populations. *Genetics*, **16**, 97-159.
- Yaakub S, Bellwood DR, van Herwerden L, Walsh FM (2006) Hybridization in coral reef fishes: Introgression and bi-directional gene exchange in *Thalassoma* (family Labridae). *Molecular Phylogenetics and Evolution*, **40**, 84-100.

## Authors' contributions

W.J.C. led the project including the search of funding support for completing this study. W.J.C. and P.B. contributed to the conception, design of the work, led the fieldwork and collected the samples. J.A.M.B. participated to fieldwork, generated most of the molecular data, did the analyses and wrote the manuscript. J.A.M.B., P.B. and W.J.C. interpreted the results. P.B. and W.J.C. revised successive versions of the manuscript.

## Data accessibility

The nucleotide sequences were deposited in GenBank, with accession numbers KU553479 to KU554420.

**Table 1** Genetic diversity and results of neutrality test at the mitochondrial locus for *A. sexfasciatus* and *A. vaigiensis* across sample localities and regions. Localities arranged from west to east. N sample size; S number of polymorphic sites;  $H_d$  haplotype diversity;  $\pi$  nucleotide diversity; SD standard deviation, estimated by bootstrap resampling. Values of Tajima's (1989) D and Fu's (1997)  $F_S$  were considered statistically significant (\*) when  $P < 0.05$  and  $P < 0.02$ , respectively.

Species or lineage, Region, Locality	N	S	$H_d \pm SD$	$\pi \pm SD$	D	$F_S$
<i>A. sexfasciatus</i>						
Indian Ocean						
Glorieuses	6	8	0.93 ± 0.12	0.0025 ± 0.0018	-1.07	-1.37
Mauritius	15	17	0.94 ± 0.04	0.0024 ± 0.0015	-1.86*	-4.43*
Coral Triangle Region						
Dongsha	49	23	0.86 ± 0.03	0.0016 ± 0.0010	-2.06*	-12.37*
Northern Taiwan	19	12	0.80 ± 0.08	0.0015 ± 0.0010	-1.85*	-4.25*
Southern Taiwan	14	6	0.60 ± 0.15	0.0008 ± 0.0007	-1.73*	-3.03*
Eastern Taiwan	13	5	0.88 ± 0.06	0.0016 ± 0.0011	-0.22	-2.59
Virac, Philippines	4	6	1.00 ± 0.18	0.0027 ± 0.0021	-0.31	-1.16
West Papua	3	11	1.00 ± 0.27	0.0063 ± 0.0051	0.00	0.81
Pacific Ocean						
New Caledonia	24	21	0.84 ± 0.06	0.0022 ± 0.0014	-2.14*	-5.01*
Moruroa	12	0	0.00 ± 0.00	0.0000 ± 0.0000	0.00	NA
<i>A. vaigiensis</i> lineage A						
Indian Ocean						
Eilat, Red Sea	8	13	0.96 ± 0.08	0.0033 ± 0.0021	-1.24	-2.60
Mozambique Channel <sup>a</sup>	6	9	0.93 ± 0.12	0.0030 ± 0.0021	-0.72	-1.01
Maldives	21	24	0.95 ± 0.04	0.0026 ± 0.0016	-2.08*	-14.28*
Coral Triangle Region						
Hainan	21	21	0.97 ± 0.03	0.0028 ± 0.0017	-1.68*	-11.29*
Taiping	8	13	0.96 ± 0.08	0.0038 ± 0.0024	-0.66	-2.24
Bali	17	22	0.97 ± 0.03	0.0035 ± 0.0021	-1.46	-8.14*
Dongsha	2	4	1.00 ± 0.50	0.0034 ± 0.0039	0.00	1.39
Little Liuqiu <sup>b</sup>	10	16	0.93 ± 0.08	0.0037 ± 0.0023	-1.08	-2.39
Southern Taiwan	34	33	0.96 ± 0.02	0.0038 ± 0.0021	-1.63*	-11.10*
Northern Taiwan	15	21	0.98 ± 0.03	0.0037 ± 0.0022	-1.37	-7.53*
Eastern Taiwan	25	25	0.96 ± 0.02	0.0035 ± 0.0020	-1.41	-8.87*
PNG <sup>c</sup>	33	31	0.97 ± 0.02	0.0037 ± 0.0021	-1.56*	-11.71*
<i>A. vaigiensis</i> lineage B						
(all localities)	6	44	0.93 ± 0.12	0.0128 ± 0.0077	-1.46*	1.48
<i>A. vaigiensis</i> lineage C						
(all localities)	11	10	0.89 ± 0.09	0.0017 ± 0.0012	-1.81*	-4.55*

<sup>a</sup>'Mozambique Channel' groups samples from Europa, Juan de Nova and Glorieuses Is.;

<sup>b</sup> Little Liuqiu is a small island off the southwestern coast of Taiwan;

<sup>c</sup>'PNG' groups samples from Madang, Kavieng and the Louisiade archipelago at the eastern extremity of New Guinea.

**Table 2** Summary of the AMOVA of *A. sexfasciatus* and *A. vaigiensis* lineage A. Sampling localities were nested into regions, at the mitochondrial locus. Percentage of total genetic variation as well as the corresponding fixation indices are given.

Species or lineage, Source of variation	Percentage of variation	Fixation index
<i>A. sexfasciatus</i>		
Among regions	66.9	$F_{CT} = 0.669^*$
Among populations within regions	11.6	$F_{SC} = 0.351^*$
Among populations across regions	21.4	$F_{ST} = 0.786^*$
<i>A. vaigiensis</i> lineage A		
Among regions	3.3	$F_{CT} = 0.033$
Among populations within regions	0	$F_{SC} = -0.001$
Among populations across regions	98.0	$F_{ST} = 0.020$

\*  $P < 0.05$

**Table 3** Demographic parameter estimates derived from the IMa2 analysis for *A. sexfasciatus* and *A. vaigiensis* lineage A from the Coral Triangle region. For each partition, the following parameters were estimated: mutation rate ( $\mu$ ); length of alignment ( $l$ ); effective population size [calculated as  $(\theta/4.\mu.l)$ , with  $\theta$  being the effective population size in the model] in *A. sexfasciatus* ( $N_1$ ), in *A. vaigiensis* lineage A ( $N_2$ ) and in the ancestral population ( $N_a$ ); splitting time ( $t$ ); time to most recent common ancestor [tMRCA, calculated as  $(t/\mu.l)$ ]; and relative migration rates from *A. sexfasciatus* to *A. vaigiensis* lineage A ( $m_1$ ) and reciprocally ( $m_2$ ) (with  $m = m.\mu.l$  being the mutation rate in the coalescent).

Marker	$\mu^*$	$l$	$N_1$	$N_2$	$N_a$	$t$	tMRCA	$m_1$	$m_2$
<i>Dyst</i>	7.50x10 <sup>-9</sup>	1040	54487	28846	176282	6198718	961538	2.34x10 <sup>-8</sup>	2.34x10 <sup>-8</sup>
<i>RAG1</i>	9.70x10 <sup>-10</sup>	985	562562	1295201	91580	889633	2616568	4.78x10 <sup>-9</sup>	4.78x10 <sup>-9</sup>
2 nuclear loci	2.70x10 <sup>-9</sup>	2025	89255	116718	514933	10518372	NA	2.73x10 <sup>-8</sup>	2.73x10 <sup>-8</sup>
<i>cytb</i>	1.30x10 <sup>-8</sup>	1161	1347976	2733055	66670	122772	163983	2.94x10 <sup>-7</sup>	3.55x10 <sup>-7</sup>
All loci	4.00x10 <sup>-9</sup>	3186	120753	197327	874719	514425	NA	4.90x10 <sup>-7</sup>	4.90x10 <sup>-7</sup>

\* When several loci were involved, the average mutation rate was calculated as the geometric mean of the individual mutation rates

## Captions to figures

**Figure 1** Map of the Indo-West Pacific, with sampling localities and sample sizes. **(A)** Sampling localities for *Abudefduf* spp. **(B)** Sampling localities for *A. vaigiensis* and *A. sexfasciatus*. Circle surface (red: *A. sexfasciatus*; blue: *A. vaigiensis*) proportional to sample size. Pie charts give the proportion of each lineage under *A. vaigiensis* (see inset) in a sample. Additional information provided in Table 1 and in Supplementary Table S1.

**Fig. 2** Phylogenetic trees of *Abudefduf* spp. Tree topology corresponds to the best ML tree. Node supports are indicated by bootstrap values (when >50%); white circles indicate nodes whose Bayesian probability was >90%. Sequences generated in the course of this study are in red for *Abudefduf sexfasciatus*, in blue for *Abudefduf vaigiensis*, with different tones for the cryptic groups (A, B or C), or in black (other *Abudefduf* species); sequences retrieved from GenBank are in grey. When nodes were collapsed for readability purpose, the number of sequences included in a cluster and their geographical origin are mentioned. (A) From mitochondrial sequences (*cytb* locus). Homologous sequences from *Oreochromis niloticus* (GenBank NC\_013663) and *Amphiprion ocellaris* (GenBank AP006017) were used as outgroup. *Similiaparma lurida* (previously *Abudefduf luridus*) is considered as a separate genus (Cooper *et al.* 2014). (B) From concatenated (*RAG1+Dyst*) nuclear sequences.

**Fig. 3** Minimum-spanning parsimony networks for *Abudefduf sexfasciatus* (in red) and *Abudefduf vaigiensis* A (in blue). Hatches depicts the number of mutations; circle surface is proportional to sample size; tones of red (A. *sexfasciatus*) differ according to geographic origin. (A) Haplotypes at the *cytb* locus. (B), haplotypes of the nuclear exon *RAG1*. (C) haplotypes of the nuclear intron *Dyst*.

**Fig. 4** *Abudefduf sexfasciatus* and *A. vaigiensis* lineage A. relative likelihood distributions for three demographic parameters estimated using IMa2. Row A concerns the nuclear data set (*RAG1+Dyst*); row B, the mitochondrial data set (*cytb*); and row C, the three-locus data set (*cytb+RAG1+Dyst*). *Full line*: splitting time ( $t$ ); *dashed line*: time to most recent common ancestor (tMRCA) in million years before present (BP); *black*: effective migration rate ( $m$ ) from *A. sexfasciatus* to *A. vaigiensis* lineage A; *grey*: effective migration rate from *A. vaigiensis* lineage A to *A. sexfasciatus*. Migration rates expressed in individuals per million year. Effective population size ( $N_e$ , in millions of individuals) for *A. sexfasciatus* (in black), *A. vaigiensis* lineage A (in grey) and for the ancestral population (black dashed line).

**Fig. 5** Schematic overview of the proposed evolutionary scenario. In the species tree depicting the split between Coral Triangle region populations of *A. sexfasciatus* and *A. vaigiensis* A, the mitochondrial locus (*cytb*) exhibit inter-specific gene flow. Among the nuclear loci, lineage sorting is complete at the intron *Dyst* but incomplete at the exon *RAG1*.

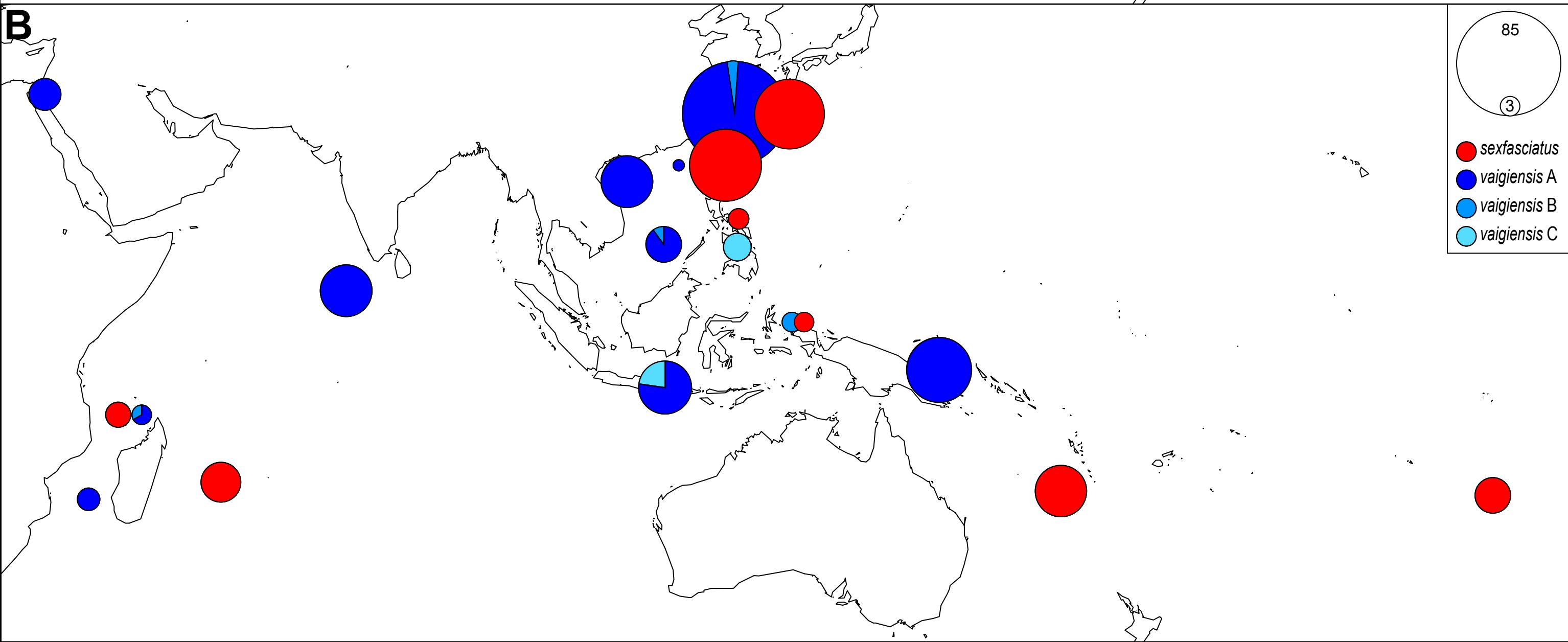
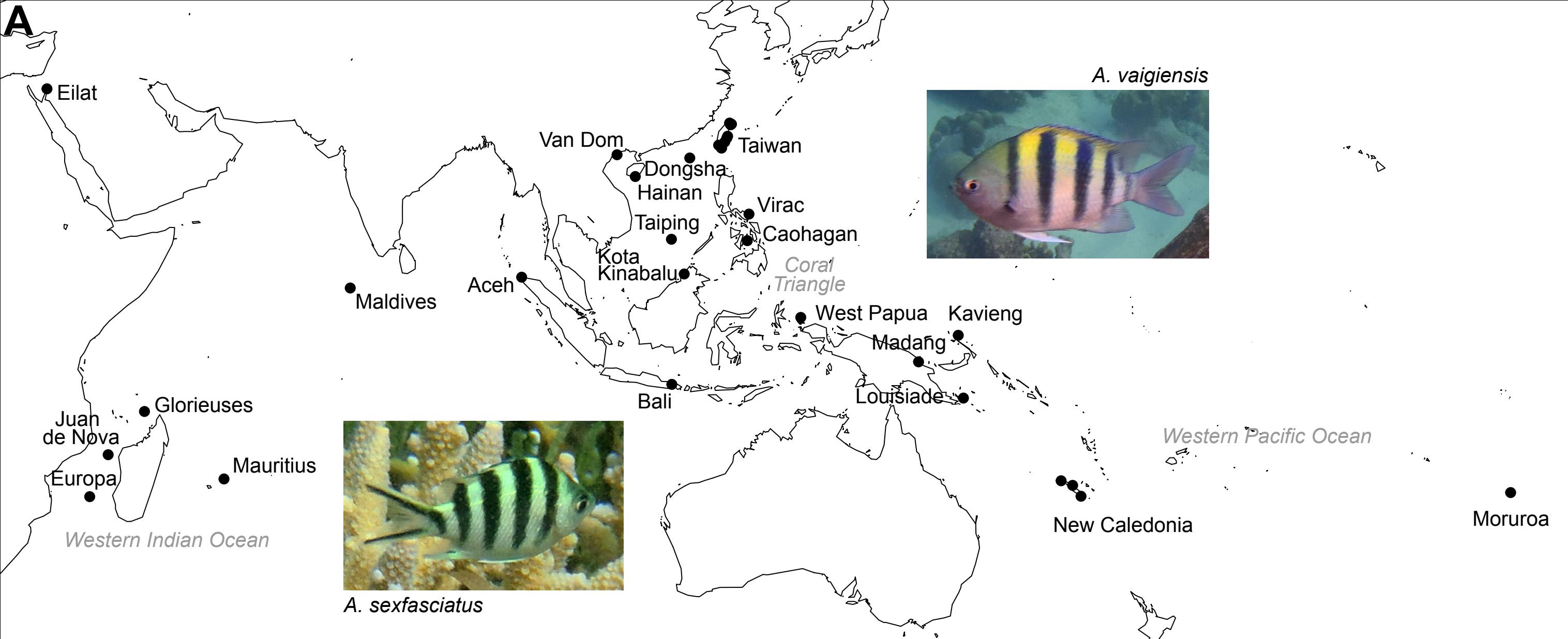
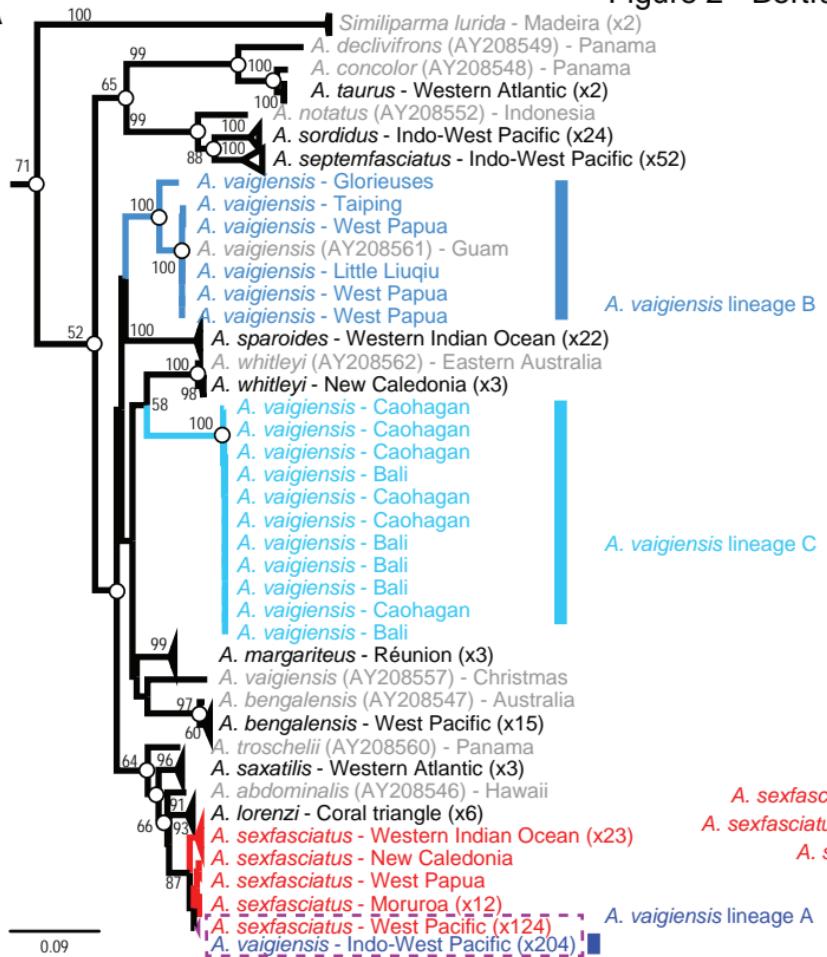


Figure 2 - Bertrand, Borsa, Chen

A



B

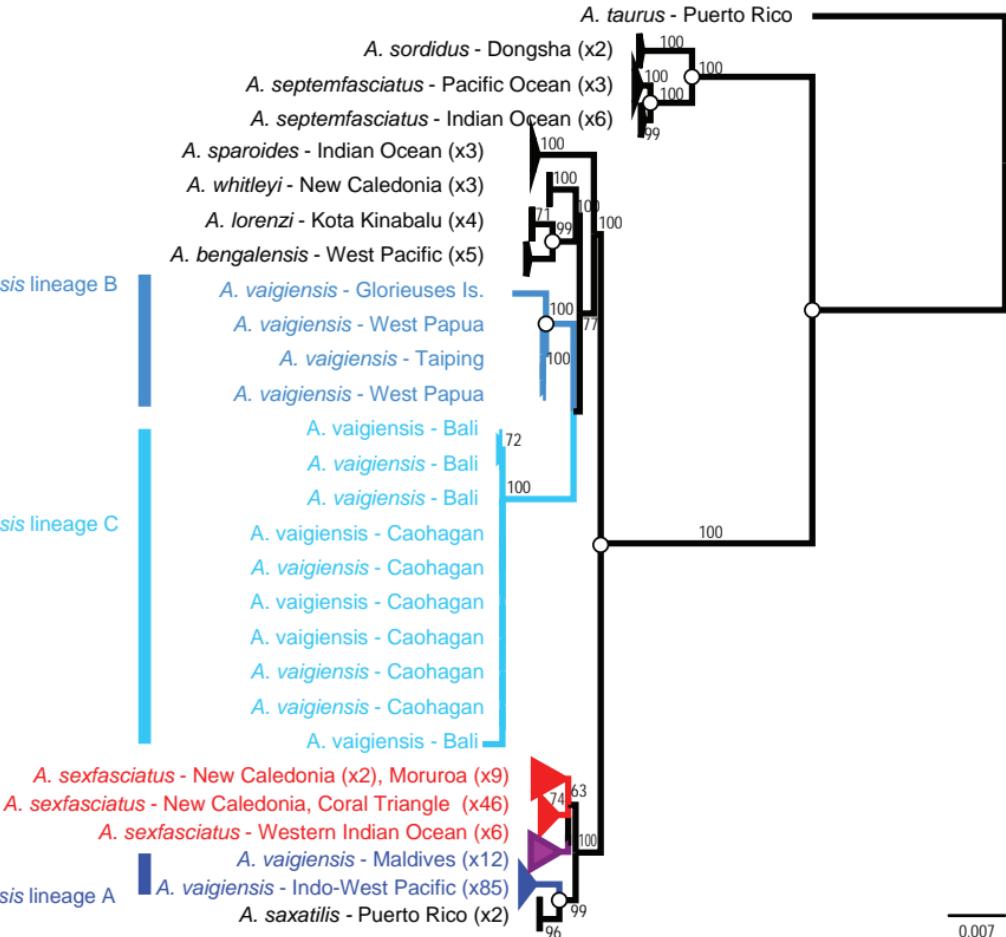


Figure 3 - Bertrand, Borsa, Chen

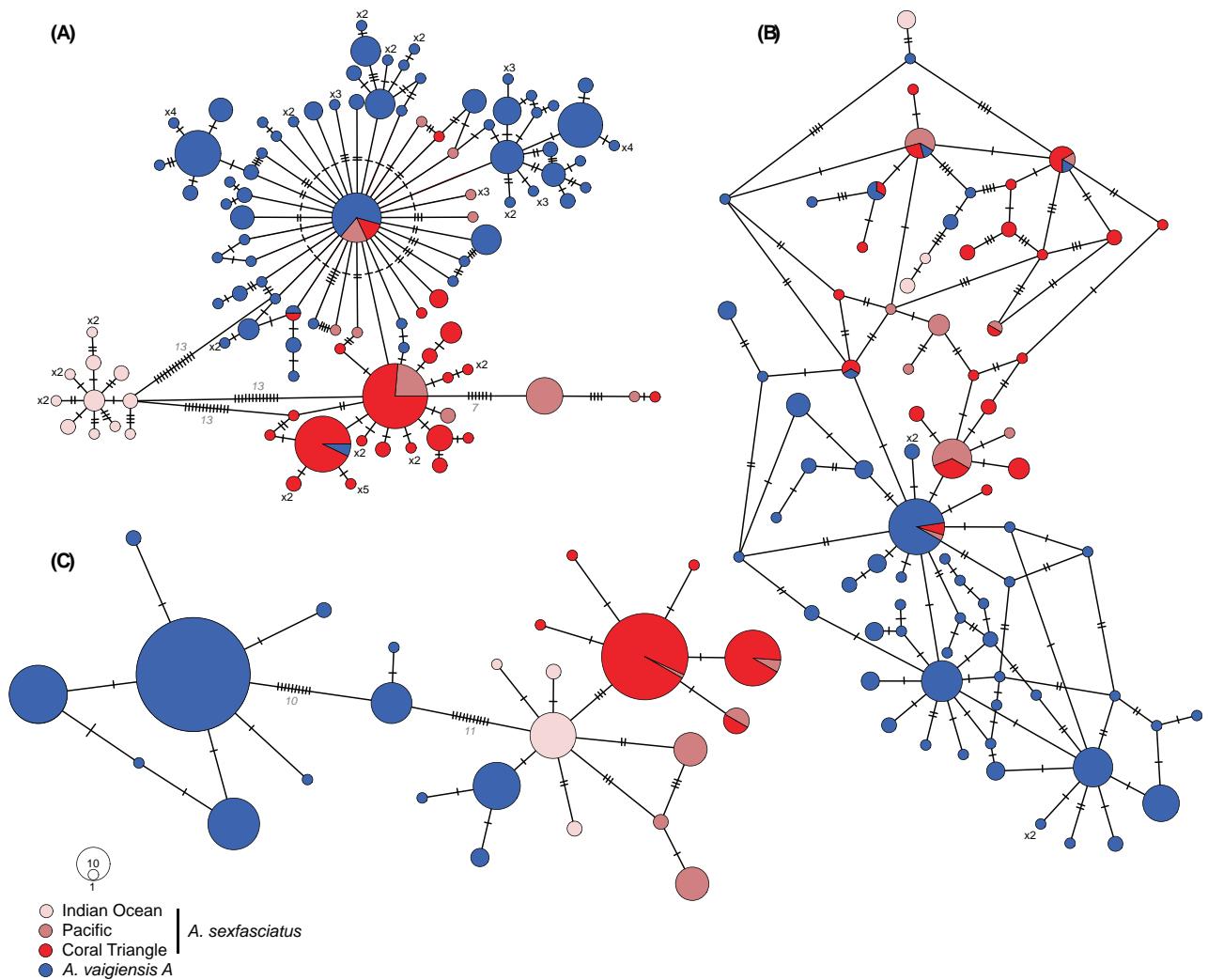
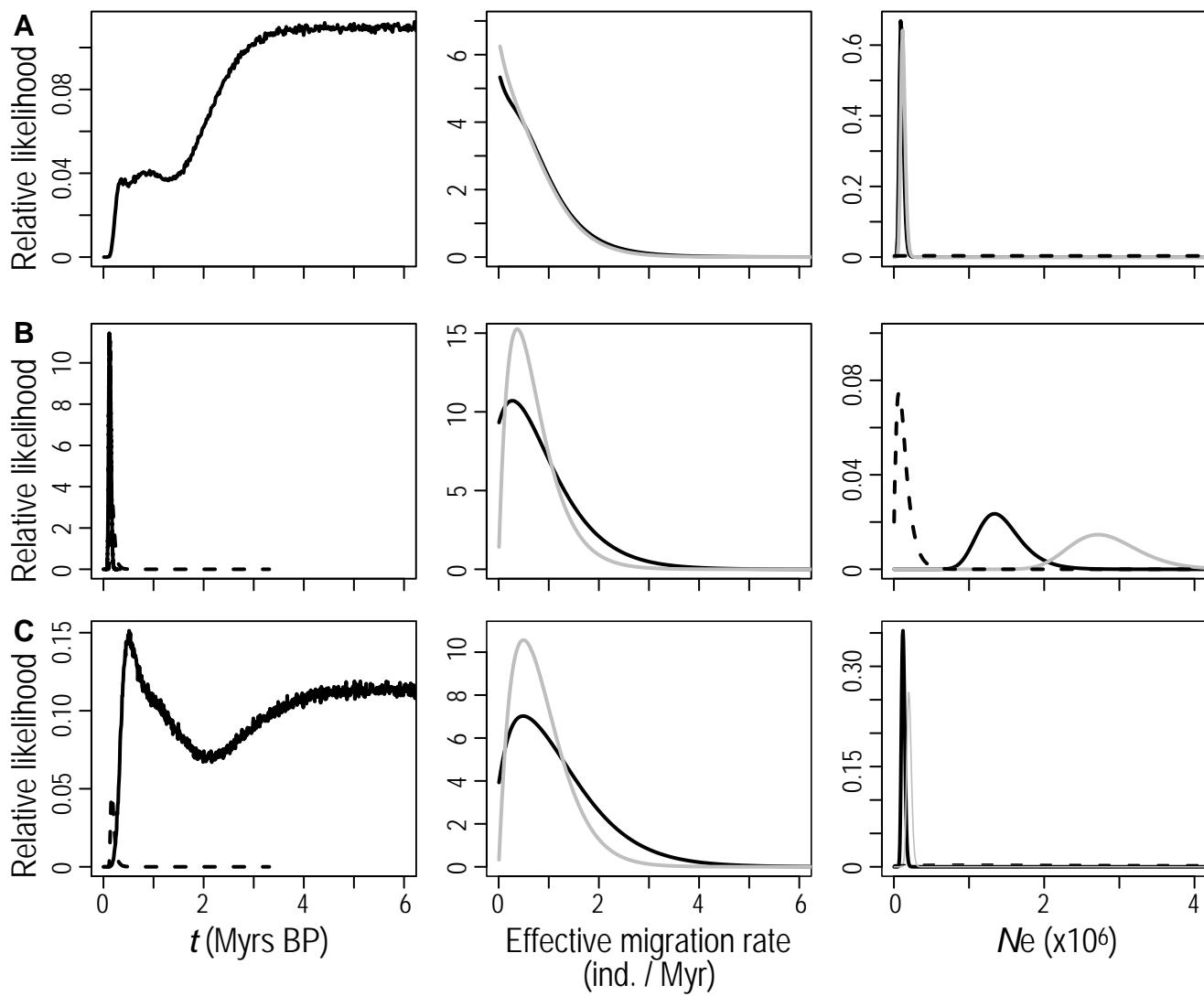
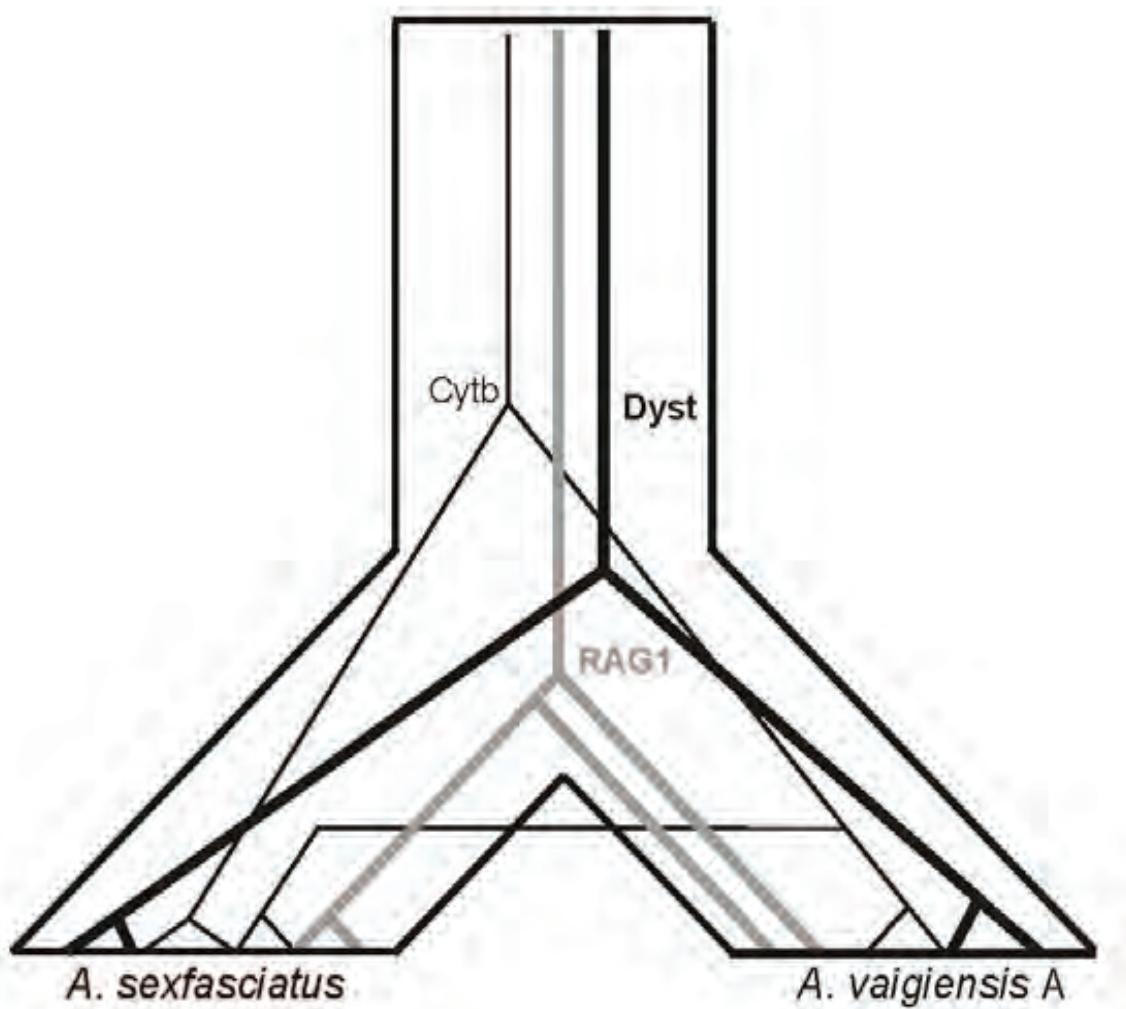


Figure 4 - Bertrand, Borsa, Chen





**Supplementary material for "Phylogeography of the sergeants *Abudefduf sexfasciatus* and *A. vaigiensis* reveals complex introgression patterns between two widespread and sympatric Indo-West Pacific reef fishes"**

Joris A. M. Bertrand, Philippe Borsa and Wei-Jen Chen

Supplementary Tables S1-S4 and Supplementary Figures S1 and S2 here appended.

**Supplementary Table S1** *Abudedefduf* spp. individuals sampled for the present study. Each row represents an individual identified to species, for which sampling location (region, locality and geographic coordinates), individual number, main collector and when available, GenBank accession number at three loci (*cytb*, *RAG1* and *Dyst*). N, total number of individuals sequenced at a locus

Individual no.	Species	Sampling details					GenBank no.		
		Region	Locality	Lat.	Long.	Coll.	<i>cytb</i>	<i>RAG1</i>	<i>Dyst</i>
WJC1514	<i>A. bengalensis</i>	Coral Triangle	East Taiwan	24.94	121.89	WJC	KU553479	KU554197	KU553963
WJC1515	<i>A. bengalensis</i>	Coral Triangle	East Taiwan	24.94	121.89	WJC	KU553480	KU554198	KU553964
WJC1728	<i>A. bengalensis</i>	Coral Triangle	Van Dom	21.13	107.45	WJC	KU553481	KU554199	KU553965
WJC4314	<i>A. bengalensis</i>	Coral Triangle	South Taiwan	120.71	22.01	WJC	KU553482	-	-
WJC4315	<i>A. bengalensis</i>	Coral Triangle	South Taiwan	120.71	22.01	WJC	KU553483	-	-
WJC4316	<i>A. bengalensis</i>	Coral Triangle	South Taiwan	120.71	22.01	WJC	KU553484	-	-
WJC5381	<i>A. bengalensis</i>	Coral Triangle	Kota Kinabalu	115.99	5.96	WJC	KU553485	KU554200	KU553966
WJC5382	<i>A. bengalensis</i>	Coral Triangle	Kota Kinabalu	115.99	5.96	WJC	KU553486	-	KU553967
WJC5383	<i>A. bengalensis</i>	Coral Triangle	Kota Kinabalu	115.99	5.96	WJC	KU553487	KU554201	KU553968
WJC5384	<i>A. bengalensis</i>	Coral Triangle	Kota Kinabalu	115.99	5.96	WJC	KU553488	-	KU553969
WJC5385	<i>A. bengalensis</i>	Coral Triangle	Kota Kinabalu	115.99	5.96	WJC	KU553489	-	-
WJC5386	<i>A. bengalensis</i>	Coral Triangle	Kota Kinabalu	115.99	5.96	WJC	KU553490	-	-
WJC5387	<i>A. bengalensis</i>	Coral Triangle	Kota Kinabalu	115.99	5.96	WJC	-	-	-
WJC5388	<i>A. bengalensis</i>	Coral Triangle	Kota Kinabalu	115.99	5.96	WJC	-	-	-
WJC5389	<i>A. bengalensis</i>	Coral Triangle	Kota Kinabalu	115.99	5.96	WJC	KU553491	-	-
WJC5390	<i>A. bengalensis</i>	Coral Triangle	Kota Kinabalu	115.99	5.96	WJC	KU553492	-	-
WJC5391	<i>A. bengalensis</i>	Coral Triangle	Kota Kinabalu	115.99	5.96	WJC	KU553493	-	-
WJC5375	<i>A. lorenzi</i>	Coral Triangle	Kota Kinabalu	115.99	5.96	WJC	KU553494	KU554202	KU553970
WJC5376	<i>A. lorenzi</i>	Coral Triangle	Kota Kinabalu	115.99	5.96	WJC	KU553495	KU554203	KU553971
WJC5377	<i>A. lorenzi</i>	Coral Triangle	Kota Kinabalu	115.99	5.96	WJC	KU553496	KU554204	KU553972
WJC5378	<i>A. lorenzi</i>	Coral Triangle	Kota Kinabalu	115.99	5.96	WJC	-	KU554205	KU553973
WJC5379	<i>A. lorenzi</i>	Coral Triangle	Kota Kinabalu	115.99	5.96	WJC	KU553497	-	-
WJC5380	<i>A. lorenzi</i>	Coral Triangle	Kota Kinabalu	115.99	5.96	WJC	KU553498	-	-
WJC67	<i>A. saxatilis</i>	Atlantic Ocean	Puerto Rico	18.48	-66.13	KC	KU553499	KU554206	KU553974
WJC84	<i>A. saxatilis</i>	Atlantic Ocean	Puerto Rico	18.48	-66.13	KC	KU553500	KU554207	KU553975
EPS159	<i>A. septemfasciatus</i>	Indian Ocean	Juan de Nova	-17.03	42.72	PB, WJC	KU553501	KU554208	KU553976
EPS160	<i>A. septemfasciatus</i>	Indian Ocean	Juan de Nova	-17.03	42.72	PB, WJC	KU553502	-	-
EPS214	<i>A. septemfasciatus</i>	Indian Ocean	Juan de Nova	-17.03	42.72	PB, WJC	KU553503	-	-
EPS215	<i>A. septemfasciatus</i>	Indian Ocean	Juan de Nova	-17.03	42.72	PB, WJC	KU553504	-	-
EPS216	<i>A. septemfasciatus</i>	Indian Ocean	Juan de Nova	-17.03	42.72	PB, WJC	KU553505	-	-
EPS217	<i>A. septemfasciatus</i>	Indian Ocean	Juan de Nova	-17.03	42.72	PB, WJC	KU553506	-	-
EPS218	<i>A. septemfasciatus</i>	Indian Ocean	Juan de Nova	-17.03	42.72	PB, WJC	KU553507	-	-
EPS221	<i>A. septemfasciatus</i>	Indian Ocean	Juan de Nova	-17.03	42.72	PB, WJC	KU553508	KU554209	KU553977
EPS222	<i>A. septemfasciatus</i>	Indian Ocean	Juan de Nova	-17.03	42.72	PB, WJC	KU553509	KU554210	KU553978
EPS223	<i>A. septemfasciatus</i>	Indian Ocean	Juan de Nova	-17.03	42.72	PB, WJC	KU553510	KU554211	KU553979
EPS224	<i>A. septemfasciatus</i>	Indian Ocean	Juan de Nova	-17.03	42.72	PB, WJC	KU553511	-	-
EPS225	<i>A. septemfasciatus</i>	Indian Ocean	Juan de Nova	-17.03	42.72	PB, WJC	KU553512	-	-
EPS226	<i>A. septemfasciatus</i>	Indian Ocean	Juan de Nova	-17.03	42.72	PB, WJC	KU553513	-	-
EPS227	<i>A. septemfasciatus</i>	Indian Ocean	Juan de Nova	-17.03	42.72	PB, WJC	KU553514	-	-
EPS228	<i>A. septemfasciatus</i>	Indian Ocean	Juan de Nova	-17.03	42.72	PB, WJC	KU553515	-	-
EPS229	<i>A. septemfasciatus</i>	Indian Ocean	Juan de Nova	-17.03	42.72	PB, WJC	KU553516	-	-
EPS232	<i>A. septemfasciatus</i>	Indian Ocean	Juan de Nova	-17.03	42.72	PB, WJC	KU553517	-	-
EPS234	<i>A. septemfasciatus</i>	Indian Ocean	Juan de Nova	-17.03	42.72	PB, WJC	KU553518	-	-
EPS235	<i>A. septemfasciatus</i>	Indian Ocean	Juan de Nova	-17.03	42.72	PB, WJC	KU553519	-	-
EPS236	<i>A. septemfasciatus</i>	Indian Ocean	Juan de Nova	-17.03	42.72	PB, WJC	KU553520	-	-
EPS238	<i>A. septemfasciatus</i>	Indian Ocean	Juan de Nova	-17.03	42.72	PB, WJC	KU553521	-	-
EPS239	<i>A. septemfasciatus</i>	Indian Ocean	Juan de Nova	-17.03	42.72	PB, WJC	KU553522	-	-
EPS240	<i>A. septemfasciatus</i>	Indian Ocean	Juan de Nova	-17.03	42.72	PB, WJC	KU553523	-	-
EPS241	<i>A. septemfasciatus</i>	Indian Ocean	Juan de Nova	-17.03	42.72	PB, WJC	KU553524	-	-
EPS285	<i>A. septemfasciatus</i>	Indian Ocean	Juan de Nova	-17.03	42.72	PB, WJC	KU553525	-	-
EPS286	<i>A. septemfasciatus</i>	Indian Ocean	Juan de Nova	-17.03	42.72	PB, WJC	KU553526	-	-
EPS287	<i>A. septemfasciatus</i>	Indian Ocean	Juan de Nova	-17.03	42.72	PB, WJC	KU553527	-	-
EPS445	<i>A. septemfasciatus</i>	Indian Ocean	Glorieuses	-11.53	47.33	PB, WJC	KU553528	-	-
EPS446	<i>A. septemfasciatus</i>	Indian Ocean	Glorieuses	-11.53	47.33	PB, WJC	KU553529	-	-
EPS447	<i>A. septemfasciatus</i>	Indian Ocean	Glorieuses	-11.53	47.33	PB, WJC	KU553530	-	-
EPS448	<i>A. septemfasciatus</i>	Indian Ocean	Glorieuses	-11.53	47.33	PB, WJC	KU553531	-	-
EPS449	<i>A. septemfasciatus</i>	Indian Ocean	Glorieuses	-11.53	47.33	PB, WJC	KU553532	-	-
WJC806	<i>A. septemfasciatus</i>	Coral Triangle	South Taiwan	22.35	120.37	WJC	KU553533	KU554212	KU553980
WJC807	<i>A. septemfasciatus</i>	Coral Triangle	South Taiwan	22.35	120.37	WJC	KU553534	KU554213	-
WJC1265	<i>A. septemfasciatus</i>	Pacific Ocean	New Caledonia	-20.93	165.40	WJC	KU553535	KU554214	KU553981
WJC1415	<i>A. septemfasciatus</i>	Coral Triangle	East Taiwan	23.24	121.41	WJC	KU553536	KU554215	KU553982
WJC1426	<i>A. septemfasciatus</i>	Coral Triangle	East Taiwan	23.48	121.51	WJC	KU553537	KU554216	-
WJC3447	<i>A. septemfasciatus</i>	Indian Ocean	Aceh	5.55	95.32	WJC	KU553538	KU554217	-
WJC3448	<i>A. septemfasciatus</i>	Indian Ocean	Aceh	5.55	95.32	WJC	KU553539	KU554218	KU553983
WJC3449	<i>A. septemfasciatus</i>	Indian Ocean	Aceh	5.55	95.32	WJC	KU553540	KU554219	KU553984
WJC3450	<i>A. septemfasciatus</i>	Indian Ocean	Aceh	5.55	95.32	WJC	KU553541	KU554220	-





WJC4206	<i>A. sexfasciatus</i>	Coral Triangle	East Taiwan	23.48	121.51	WJC	KU553748	KU554298	KU554065
WJC4207	<i>A. sexfasciatus</i>	Coral Triangle	East Taiwan	23.48	121.51	WJC	KU553749	KU554299	KU554066
WJC4208	<i>A. sexfasciatus</i>	Coral Triangle	East Taiwan	23.48	121.51	WJC	KU553750	KU554300	KU554067
WJC4209	<i>A. sexfasciatus</i>	Coral Triangle	East Taiwan	23.48	121.51	WJC	KU553751	KU554301	KU554068
WJC4210	<i>A. sexfasciatus</i>	Coral Triangle	East Taiwan	23.48	121.51	WJC	KU553752	KU554302	KU554069
WJC4211	<i>A. sexfasciatus</i>	Coral Triangle	East Taiwan	23.48	121.51	WJC	KU553753	KU554303	KU554070
WJC4212	<i>A. sexfasciatus</i>	Coral Triangle	East Taiwan	23.48	121.51	WJC	KU553754	KU554304	KU554071
WJC4317	<i>A. sexfasciatus</i>	Coral Triangle	South Taiwan	21.98	120.75	WJC	KU553755	-	-
WJC4320	<i>A. sexfasciatus</i>	Coral Triangle	South Taiwan	21.98	120.75	WJC	KU553756	-	-
WJC4323	<i>A. sexfasciatus</i>	Coral Triangle	South Taiwan	21.98	120.75	WJC	KU553757	-	-
WJC4324	<i>A. sexfasciatus</i>	Coral Triangle	South Taiwan	21.98	120.75	WJC	KU553758	-	-
WJC4366	<i>A. sexfasciatus</i>	Coral Triangle	North Taiwan	25.02	122.00	WJC	KU553759	-	-
WJC4367	<i>A. sexfasciatus</i>	Coral Triangle	North Taiwan	25.02	122.00	WJC	KU553760	-	-
WJC4796	<i>A. sexfasciatus</i>	Coral Triangle	South Taiwan	22.35	120.37	WJC	KU553761	-	-
WJC4797	<i>A. sexfasciatus</i>	Coral Triangle	South Taiwan	22.35	120.37	WJC	KU553762	-	-
EPS51	<i>A. sordidus</i>	Indian Ocean	Europa	-22.37	40.37	PB, WJC	KU553551	KU554224	-
EPS52	<i>A. sordidus</i>	Indian Ocean	Europa	-22.37	40.37	PB, WJC	KU553552	KU554225	-
EPS157	<i>A. sordidus</i>	Indian Ocean	Juan de Nova	-17.03	42.72	PB, WJC	KU553553	KU554226	-
EPS237	<i>A. sordidus</i>	Indian Ocean	Juan de Nova	-17.03	42.72	PB, WJC	KU553554	-	-
EPS283	<i>A. sordidus</i>	Indian Ocean	Juan de Nova	-17.03	42.72	PB, WJC	KU553555	KU554227	-
EPS284	<i>A. sordidus</i>	Indian Ocean	Juan de Nova	-17.03	42.72	PB, WJC	KU553556	KU554228	-
EPS442	<i>A. sordidus</i>	Indian Ocean	Glorieuses	-11.53	47.33	PB, WJC	KU553557	-	-
EPS443	<i>A. sordidus</i>	Indian Ocean	Glorieuses	-11.53	47.33	PB, WJC	KU553558	-	-
WJC1260	<i>A. sordidus</i>	Coral Triangle	Dongsha	20.72	116.70	WJC	KU553559	KU554229	KU553985
WJC1261	<i>A. sordidus</i>	Coral Triangle	Dongsha	20.72	116.70	WJC	KU553560	KU554230	KU553986
WJC1410	<i>A. sordidus</i>	Coral Triangle	East Taiwan	23.24	121.41	WJC	KU553561	-	-
WJC1411	<i>A. sordidus</i>	Coral Triangle	East Taiwan	23.24	121.41	WJC	KU553562	-	-
WJC1425	<i>A. sordidus</i>	Coral Triangle	East Taiwan	23.24	121.41	WJC	KU553563	KU554231	-
WJC4125	<i>A. sordidus</i>	Coral Triangle	Dongsha	20.72	116.70	WJC	KU553564	-	-
WC4155	<i>A. sordidus</i>	Coral Triangle	Dongsha	20.72	116.70	WJC	KU553565	-	-
WJC4368	<i>A. sordidus</i>	Coral Triangle	North Taiwan	25.02	122.00	WJC	KU553566	-	-
WJC4789	<i>A. sordidus</i>	Coral Triangle	South Taiwan	22.35	120.37	WJC	KU553567	-	-
WJC4960	<i>A. sordidus</i>	Coral Triangle	Taiping	10.38	114.37	WJC	KU553568	KU554232	-
EPS204	<i>A. sparoides</i>	Indian Ocean	Juan de Nova	-17.03	42.72	PB, WJC	KU553569	KU554233	KU553987
EPS255	<i>A. sparoides</i>	Indian Ocean	Juan de Nova	-17.03	42.72	PB, WJC	KU553570	KU554234	KU553988
EPS424	<i>A. sparoides</i>	Indian Ocean	Glorieuses	-11.53	47.33	PB, WJC	KU553571	KU554235	KU553989
EPS425	<i>A. sparoides</i>	Indian Ocean	Glorieuses	-11.53	47.33	PB, WJC	KU553572	KU554236	KU553990
EPS450	<i>A. sparoides</i>	Indian Ocean	Glorieuses	-11.53	47.33	PB, WJC	KU553573	-	-
EPS451	<i>A. sparoides</i>	Indian Ocean	Glorieuses	-11.53	47.33	PB, WJC	KU553574	-	-
EPS452	<i>A. sparoides</i>	Indian Ocean	Glorieuses	-11.53	47.33	PB, WJC	KU553575	-	-
EPS453	<i>A. sparoides</i>	Indian Ocean	Glorieuses	-11.53	47.33	PB, WJC	KU553576	-	-
EPS454	<i>A. sparoides</i>	Indian Ocean	Glorieuses	-11.53	47.33	PB, WJC	KU553577	-	-
EPS484	<i>A. sparoides</i>	Indian Ocean	Mauritius	-20.12	57.45	PB	KU553578	-	-
EPS485	<i>A. sparoides</i>	Indian Ocean	Mauritius	-20.12	57.45	PB	KU553579	-	-
EPS486	<i>A. sparoides</i>	Indian Ocean	Mauritius	-20.12	57.45	PB	KU553580	-	-
EPS487	<i>A. sparoides</i>	Indian Ocean	Mauritius	-20.12	57.45	PB	KU553581	-	-
EPS488	<i>A. sparoides</i>	Indian Ocean	Mauritius	-20.12	57.45	PB	KU553582	-	-
WJC85	<i>A. taurus</i>	Atlantic Ocean	Puerto Rico	-66.13	18.48	KC	KU553583	KU554237	KU553991
EPS15	<i>A. vaigiensis A</i>	Indian Ocean	Europa	-22.37	40.37	PB, WJC	KU553763	KU554305	KU554072
EPS16	<i>A. vaigiensis A</i>	Indian Ocean	Europa	-22.37	40.37	PB, WJC	KU553764	KU554306	KU554073
EPS17	<i>A. vaigiensis A</i>	Indian Ocean	Europa	-22.37	40.37	PB, WJC	KU553765	KU554307	KU554074
EPS18	<i>A. vaigiensis A</i>	Indian Ocean	Europa	-22.37	40.37	PB, WJC	KU553766	KU554308	KU554075
EPS342	<i>A. vaigiensis A</i>	Indian Ocean	Glorieuses	-11.53	47.33	PB, WJC	KU553767	KU554309	KU554076
EPS343	<i>A. vaigiensis A</i>	Indian Ocean	Glorieuses	-11.53	47.33	PB, WJC	KU553768	KU554310	KU554077
PNG24	<i>A. vaigiensis A</i>	Coral Triangle	PNG	-5.22	145.80	WJC	KU553769	-	-
PNG313	<i>A. vaigiensis A</i>	Coral Triangle	PNG	-5.22	145.80	WJC	KU553770	-	-
PNG610	<i>A. vaigiensis A</i>	Coral Triangle	PNG	-5.22	145.80	WJC	KU553771	KU554311	KU554078
PNG611	<i>A. vaigiensis A</i>	Coral Triangle	PNG	-5.22	145.80	WJC	KU553772	KU554312	KU554079
PNG612	<i>A. vaigiensis A</i>	Coral Triangle	PNG	-5.22	145.80	WJC	KU553773	KU554313	KU554080
PNG613	<i>A. vaigiensis A</i>	Coral Triangle	PNG	-5.22	145.80	WJC	KU553774	KU554314	KU554081
PNG614	<i>A. vaigiensis A</i>	Coral Triangle	PNG	-5.22	145.80	WJC	KU553775	KU554315	KU554082
PNG615	<i>A. vaigiensis A</i>	Coral Triangle	PNG	-5.22	145.80	WJC	KU553776	KU554316	KU554083
PNG616	<i>A. vaigiensis A</i>	Coral Triangle	PNG	-5.22	145.80	WJC	KU553777	KU554317	KU554084
PNG617	<i>A. vaigiensis A</i>	Coral Triangle	PNG	-5.22	145.80	WJC	KU553778	KU554318	KU554085
PNG618	<i>A. vaigiensis A</i>	Coral Triangle	PNG	-5.22	145.80	WJC	KU553779	KU554319	KU554086
PNG619	<i>A. vaigiensis A</i>	Coral Triangle	PNG	-5.22	145.80	WJC	KU553780	KU554320	KU554087
PNG620	<i>A. vaigiensis A</i>	Coral Triangle	PNG	-5.22	145.80	WJC	KU553781	-	-
PNG621	<i>A. vaigiensis A</i>	Coral Triangle	PNG	-5.22	145.80	WJC	KU553782	-	-
PNG622	<i>A. vaigiensis A</i>	Coral Triangle	PNG	-5.22	145.80	WJC	KU553783	-	-
PNG623	<i>A. vaigiensis A</i>	Coral Triangle	PNG	-5.22	145.80	WJC	KU553784	-	-
PNG624	<i>A. vaigiensis A</i>	Coral Triangle	PNG	-5.22	145.80	WJC	KU553785	-	-
PNG625	<i>A. vaigiensis A</i>	Coral Triangle	PNG	-5.22	145.80	WJC	KU553786	-	-
PNG626	<i>A. vaigiensis A</i>	Coral Triangle	PNG	-5.22	145.80	WJC	KU553787	-	-
PNG627	<i>A. vaigiensis A</i>	Coral Triangle	PNG	-5.22	145.80	WJC	KU553788	-	-
PNG628	<i>A. vaigiensis A</i>	Coral Triangle	PNG	-5.22	145.80	WJC	KU553789	-	-
PNG629	<i>A. vaigiensis A</i>	Coral Triangle	PNG	-5.22	145.80	WJC	KU553790	-	-
PNG630	<i>A. vaigiensis A</i>	Coral Triangle	PNG	-5.22	145.80	WJC	KU553791	-	-





WJC4853	<i>A. vaigiensis A</i>	Coral Triangle	Bali	-8.08	114.42	SB	KU553946	KU554398	KU554165
WJC4855	<i>A. vaigiensis A</i>	Coral Triangle	Bali	-8.08	114.42	SB	KU553947	KU554399	KU554166
WJC4856	<i>A. vaigiensis A</i>	Coral Triangle	Bali	-8.08	114.42	SB	KU553948	-	KU554167
WJC4857	<i>A. vaigiensis A</i>	Coral Triangle	Bali	-8.08	114.42	SB	KU553949	KU554400	KU554168
WJC4858	<i>A. vaigiensis A</i>	Coral Triangle	Bali	-8.08	114.42	SB	KU553950	KU554401	KU554169
WJC4859	<i>A. vaigiensis A</i>	Coral Triangle	Bali	-8.08	114.42	SB	KU553951	KU554402	KU554170
WJC4860	<i>A. vaigiensis A</i>	Coral Triangle	Bali	-8.08	114.42	SB	KU553952	KU554403	KU554171
WJC4861	<i>A. vaigiensis A</i>	Coral Triangle	Bali	-8.08	114.42	SB	KU553953	KU554404	KU554172
WJC4946	<i>A. vaigiensis A</i>	Coral Triangle	Taiping	10.38	114.37	WJC	KU553954	-	KU554173
WJC5227	<i>A. vaigiensis A</i>	Red Sea	Eilat	29.52	34.93	TGR	KU553955	-	KU554174
WJC5228	<i>A. vaigiensis A</i>	Red Sea	Eilat	29.52	34.93	TGR	KU553956	-	KU554175
WJC5229	<i>A. vaigiensis A</i>	Red Sea	Eilat	29.52	34.93	TGR	KU553957	-	KU554176
WJC5230	<i>A. vaigiensis A</i>	Red Sea	Eilat	29.52	34.93	TGR	KU553958	-	KU554177
WJC5231	<i>A. vaigiensis A</i>	Red Sea	Eilat	29.52	34.93	TGR	KU553959	-	KU554178
WJC5232	<i>A. vaigiensis A</i>	Red Sea	Eilat	29.52	34.93	TGR	KU553960	-	KU554179
WJC5233	<i>A. vaigiensis A</i>	Red Sea	Eilat	29.52	34.93	TGR	KU553961	-	KU554180
WJC5234	<i>A. vaigiensis A</i>	Red Sea	Eilat	29.52	34.93	TGR	KU553962	-	KU554181
EPS432	<i>A. vaigiensis B</i>	Indian Ocean	Glorieuses	-11.53	47.33	PB, WJC	KU553587	KU554405	KU554182
WC2961	<i>A. vaigiensis B</i>	Coral Triangle	Taiping	10.38	114.37	WJC	KU553588	KU554406	KU554183
WJC4193	<i>A. vaigiensis B</i>	Coral Triangle	West Papua	0.43	130.82	PB	KU553589	KU554407	-
WJC4194	<i>A. vaigiensis B</i>	Coral Triangle	West Papua	0.43	130.82	PB	KU553590	KU554408	KU554184
WJC4196	<i>A. vaigiensis B</i>	Coral Triangle	West Papua	0.43	130.82	PB	KU553591	KU554409	KU554185
WJC4791	<i>A. vaigiensis B</i>	Coral Triangle	South Taiwan	22.35	120.37	WJC	KU553592	KU554410	-
WJC1291	<i>A. vaigiensis C</i>	Coral Triangle	Caohagan	10.20	124.02	FG	KU553593	KU554411	KU554186
WJC1292	<i>A. vaigiensis C</i>	Coral Triangle	Caohagan	10.20	124.02	FG	KU553594	KU554412	KU554187
WJC1293	<i>A. vaigiensis C</i>	Coral Triangle	Caohagan	10.20	124.02	FG	KU553595	KU554413	KU554188
WJC1294	<i>A. vaigiensis C</i>	Coral Triangle	Caohagan	10.20	124.02	FG	KU553596	KU554414	KU554189
WJC1295	<i>A. vaigiensis C</i>	Coral Triangle	Caohagan	10.20	124.02	FG	KU553597	KU554415	KU554190
WJC1296	<i>A. vaigiensis C</i>	Coral Triangle	Caohagan	10.20	124.02	FG	KU553598	KU554416	KU554191
WJC4840	<i>A. vaigiensis C</i>	Coral Triangle	Bali	-8.08	114.42	SB	KU553599	KU554417	KU554192
WJC4846	<i>A. vaigiensis C</i>	Coral Triangle	Bali	-8.08	114.42	SB	KU553600	KU554418	KU554193
WJC4851	<i>A. vaigiensis C</i>	Coral Triangle	Bali	-8.08	114.42	SB	KU553601	KU554419	KU554194
WJC4852	<i>A. vaigiensis C</i>	Coral Triangle	Bali	-8.08	114.42	SB	KU553602	-	KU554195
WJC4854	<i>A. vaigiensis C</i>	Coral Triangle	Bali	-8.08	114.42	SB	KU553603	KU554420	KU554196
WJC1262	<i>A. whitleyi</i>	Pacific	New Caledonia	-22.32	166.47	PB	KU553584	KU554238	KU553992
WJC1263	<i>A. whitleyi</i>	Pacific	New Caledonia	-22.32	166.47	PB	KU553585	KU554239	KU553993
WJC1264	<i>A. whitleyi</i>	Pacific	New Caledonia	-22.32	166.47	PB	KU553586	KU554240	KU553994
<i>N</i>	-	-	-	-	-	-	484	224	234

Collectors' names abbreviations: SB, S. Bahri; PB, P. Borsa; WJC, W-J Chen; JDB, J. DiBattista; KC, K. Conway; FG, F. Giancarlo; HH, H.-C. Ho; MK, M. Kulwicki; TGR, T. Gurevich Raguso.

**Supplementary Table S2** Locus names, PCR primers used, amplification conditions, resulting sequence length and evolution models at the three loci scored in this study

Locus	PCR name (F/R)	Primer sequence	Source	Annealing temperature	Sequence length (bp)	Evolution model by codon position		
						1	2	3
Cytochrome <i>b</i>	14695-F / 15969-R	5'-AAGCCACCGTTGTTATTCAACTA-3' / 5'-ATCCTAGCTTGGGAGTTAGGG-3'	WJC, unpubl.	55°C	1161	K80+I+G	HKY+I	GTR+G
<i>RAG1</i>	2533-F / 4061-R	5'-CTGAGCTGCAGTCAGTACCATAAAGATGT-3' / 5'-AATACTTGGAGGTGTAGAGCCAGT-3'	López <i>et al.</i> , 2004 / Chen <i>et al.</i> , 2007	54°C	1456	K80+I	K80+I	HKY+G
<i>Dystrophin</i>	Dyst-MCE1-F / Dyst-MCE1-R	5'-TTACAGCAGACCAACAGTGAGAA-3' / 5'-GAAGTTAACACGTTGACCTGAG-3'	WJC, unpubl.	55°C	1211	HKY	-	-

#### References

- Chen W-J, Ruiz-Carús R, Ortí G (2007) Relationships among four genera of mojarras (Teleostei: Perciformes: Gerreidae) from the western Atlantic and their tentative placement among percomorph fishes. *Journal of Fish Biology*, 70 (sb), 202–218.  
 López JA, Chen W-J, Ortí G (2004) Esociform phylogeny. *Copeia*, 2004, 449-464.

**Supplementary Table S3A** Genetic diversity and results of neutrality test derived from *A. sexfasciatus* and *A. vaigiensis* lineage A at nuclear locus *RAG1* across sample localities and regions. N sample size (number of sequences); S number of polymorphic sites;  $H_d$  haplotype diversity;  $\pi$  nucleotide diversity; SD standard deviation, estimated by bootstrap resampling. Values of Tajima's (1989)  $D$  and Fu's (1997)  $F_S$  were considered statistically significant (\*) when  $P < 0.05$  and  $P < 0.02$ , respectively

Species, region, Locality	N	S	$H_d \pm SD$	$\pi \pm SD$	Tajima's $D$	Fu's $F_S$
<i>A. sexfasciatus</i> , Indian Ocean						
Glorieuses	6	11	$0.73 \pm 0.16$	$0.0045 \pm 0.0029$	2.15	3.64
Mauritius	NA	NA	NA	NA	NA	NA
<i>A. sexfasciatus</i> , Coral Triangle Region						
Coral Triangle	6	10	$1.00 \pm 0.96$	$0.0029 \pm 0.0022$	0.67	-2.21
North Taiwan	16	15	$0.96 \pm 0.04$	$0.0028 \pm 0.0016$	-0.39	-5.23*
South Taiwan	10	9	$0.93 \pm 0.08$	$0.0024 \pm 0.0015$	0.42	-3.05
East Taiwan	6	8	$1.00 \pm 0.10$	$0.0024 \pm 0.0017$	0.05	-2.90*
Dongsha	2	0	$0.00 \pm 0.00$	$0.0000 \pm 0.0000$	0	NA
<i>A. sexfasciatus</i> , Pacific Ocean						
New Caledonia	10	7	$0.53 \pm 0.18$	$0.0012 \pm 0.0009$	-1.31	0.25
Moruroa	14	7	$0.81 \pm 0.07$	$0.0017 \pm 0.0011$	0.34	-0.48
<i>A. vaigiensis</i> , Indian Ocean						
Mozambique Channel	12	10	$1.00 \pm 0.03$	$0.0022 \pm 0.0014$	-0.10	-10.80*
Maldives	18	18	$0.95 \pm 0.04$	$0.0035 \pm 0.0020$	-0.11	-4.51
<i>A. vaigiensis</i> , Coral Triangle Region						
PNG	26	16	$0.90 \pm 0.04$	$0.0020 \pm 0.0012$	-1.07	-3.96
North Taiwan	6	6	$0.93 \pm 0.12$	$0.0017 \pm 0.0013$	-0.35	-1.67
South Taiwan	20	8	$0.85 \pm 0.05$	$0.0011 \pm 0.0008$	-0.75	-3.36*
East Taiwan	22	17	$0.95 \pm 0.03$	$0.0021 \pm 0.0012$	-1.30	-6.02*
Hainan	6	9	$0.93 \pm 0.12$	$0.0025 \pm 0.0017$	-0.52	-0.94
Dongsha	2	1	$1.00 \pm 0.50$	$0.0007 \pm 0.0010$	0	0
Taiping	8	9	$0.93 \pm 0.08$	$0.0018 \pm 0.0012$	-1.22	-1.95
Bali	22	16	$0.98 \pm 0.02$	$0.0023 \pm 0.0014$	-0.86	-17.52*

**Supplementary Table S3B** Genetic diversity and results of neutrality test derived from *A. sexfasciatus* and *A. vaigiensis* lineage A at nuclear locus *Dyst* across sample localities and regions. N sample size (number of sequences); S number of polymorphic sites;  $H_d$  haplotype diversity;  $\pi$  nucleotide diversity; SD standard deviation, estimated by bootstrap resampling. Values of Tajima's (1989)  $D$  and Fu's (1997)  $F_S$  were considered statistically significant (\*) when  $P < 0.05$  and  $P < 0.02$ , respectively

Species, Region, Locality	N	S	$H_d \pm SD$	$\pi \pm SD$	Tajima's $D$	Fu's $F_S$
<i>A. sexfasciatus</i> ,						
Indian Ocean	24	5	$0.44 \pm 0.12$	$0.0008 \pm 0.0006$	-1.19	-1.50
Coral Triangle region						
Virac	8	1	$0.43 \pm 0.17$	$0.0004 \pm 0.0005$	0.33	0.54
North Taiwan	28	1	$0.49 \pm 0.05$	$0.0005 \pm 0.0005$	1.48	1.58
South Taiwan	16	2	$0.51 \pm 0.13$	$0.0005 \pm 0.0005$	-0.19	-0.18
East Taiwan	22	3	$0.54 \pm 0.09$	$0.0006 \pm 0.0005$	-0.71	-1.08
West Papua	6	1	$0.33 \pm 0.22$	$0.0003 \pm 0.0004$	-0.93	0
Dongsha	18	2	$0.45 \pm 0.12$	$0.0005 \pm 0.0005$	-0.44	-0.38
Pacific Ocean						
New Caledonia	8	92	$0.82 \pm 0.10$	$0.0383 \pm 0.0213$	0.71	9.54*
Moruroa	20	4	$0.53 \pm 0.04$	$0.0022 \pm 0.0014$	2.51*	5.56*
<i>A. vaigiensis</i> lineage A						
Indian Ocean						
Eilat (Red Sea)	16	1	$0.13 \pm 0.11$	$0.0001 \pm 0.0002$	-1.16	-0.70
Mozambique Channel	12	1	$0.41 \pm 0.13$	$0.0004 \pm 0.0004$	0.54	0.73
Maldives	38	36	$0.61 \pm 0.05$	$0.0155 \pm 0.0079$	2.35*	19.75*
Coral Triangle Region						
PNG	36	4	$0.67 \pm 0.07$	$0.0008 \pm 0.0007$	-0.33	-1.97
North Taiwan	8	1	$0.25 \pm 0.18$	$0.0002 \pm 0.0003$	-1.05	-0.18
South Taiwan	24	4	$0.65 \pm 0.08$	$0.0008 \pm 0.0006$	-0.73	-1.51
East Taiwan	24	2	$0.47 \pm 0.11$	$0.0005 \pm 0.0005$	-0.07	-0.02
Hainan	6	5	$0.53 \pm 0.17$	$0.0026 \pm 0.0019$	0	NA
Dongsha	4	1	$0.50 \pm 0.27$	$0.0005 \pm 0.0006$	-0.61	0.17
Taiping	16	6	$0.76 \pm 0.06$	$0.0025 \pm 0.0016$	-0.19	2.13
Bali	32	2	$0.55 \pm 0.08$	$0.0006 \pm 0.0005$	0.47	0.52

**Supplementary Table S4** Output from the nested-model comparison procedure of IMa2. Numbers in first column correspond to the following models: 1 full (allowing all the parameters to vary); 2: equal migration rates; 3: coalescent migration rate is zero from population 1 to 2 (*A. sexfasciatus* to *A. vaigiensis* lineage A); 4: coalescent migration rate is zero from population 2 to 1 (*A. vaigiensis* lineage A to *A. sexfasciatus*); and 5: coalescent migration rate is zero. Values in between brackets are fixed under the model considered. From left to right, the different columns report the log-likelihood of the model: log(P), the number of varying parameters in the model: #terms, the number of degrees of freedom: df, the log-likelihood ratio statistics: 2LLR, the effective sample size (ESS), the estimates for effective population sizes of the ancestral population ( $\theta_a$ ), population 1 (*A. sexfasciatus*,  $\theta_1$ ) and population 2 (*A. vaigiensis* lineage A,  $\theta_2$ ), the effective migration rates from population 1 to population 2 (from *A. sexfasciatus* to *A. vaigiensis* lineage A,  $m_{1>2}$ ) and from population 2 to population 1 (from *A. vaigiensis* lineage A to *A. sexfasciatus*,  $m_{2>1}$ ) and the statistical significance of the model

Model	log(P)	#terms	df	2LLR	ESS	$\theta_a$	$\theta_1$	$\theta_2$	$m_{1>2}$	$m_{2>1}$	P-value
1	-2,473	5	-	-	1.309	6.3572	9.2907	0.0117	0.00960	0.07270	NA
2	<b>-3,626</b>	4	1	<b>2.305</b>	<b>1678</b>	<b>6.3426</b>	<b>10.238</b>	<b>36.435</b>	<b>0.04660</b>	[0.04660]	<b>0.129</b>
3	<b>-4,084</b>	4	1*	<b>3.223</b>	<b>169.4</b>	<b>6.8252</b>	<b>9.7426</b>	<b>38.493</b>	[0.00000]	<b>0.075900</b>	0.036
4	<b>-15,54</b>	4	1*	<b>26.14</b>	<b>133.4</b>	<b>5.9223</b>	<b>10.356</b>	<b>46.515</b>	<b>0.08590</b>	[0.00000]	0.000
5	<b>-41,61</b>	3	2*	<b>78.27</b>	<b>1.000</b>	<b>4.6473</b>	<b>11.025</b>	<b>18.515</b>	[0.00000]	[0.00000]	0.000

\* test distribution of 2LLR is a mixture

**Supplementary Figure S1** Phylogenetic trees inferred for *Abudefduf* spp. based on nuclear loci (*RAG1* and *Dyst*). Tree topology corresponds to the best maximum likelihood (ML) tree. When  $>50$ , node supports are indicated by bootstrap values; white circles show nodes whose Bayesian posterior probability was  $> 90\%$ . *Abudefduf sexfasciatus* sequences generated in this study are in red, *A. vaigiensis* sequences are in different tones of blue depending on the lineage (A, B or C), and sequences from other *Abudefduf* species are in black; sequences retrieved from GenBank are in grey. When nodes were collapsed for readability purpose, the number of sequences included in each cluster as well as their geographical origin were reported. **(A)** ML tree of *RAG1* gene sequences. The homologous sequence in *Amphiprion ocellaris* (GenBank AY208631) was used as an outgroup. *Similiparma lurida* (previously *Abudefduf luridus*) is considered as a separate genus (Cooper *et al.*, 2014). The *A. whitneyi* sequence from Lizard Is. (GenBank FJ616626) highlighted in orange may have been misidentified. **(B)** ML tree of *Dyst* gene sequences.

**Supplementary Figure S2** *Abudefduf sexfasciatus* and *A. vaigiensis* lineage A, relative likelihood distributions at each of demographic parameters inferred from IMa2 analysis. Left column: splitting time ( $t$ , full line) and tMRCA (dashed line, when available) in million years before present (Myrs BP). Central column: effective migration rates ( $m$ ) from *A. sexfasciatus* to *A. vaigiensis* (in black) and from *A. vaigiensis* lineage A to *A. sexfasciatus* (in grey) in individuals per million year (ind./Myr). Right column: effective population size ( $N_e$ ) in millions of individuals (ind.) for *A. sexfasciatus* ( $N_1$ , in black), *A. vaigiensis* lineage A ( $N_2$ , in grey) and the ancestral population ( $N_A$ , dashed black line). **(A)** From *RAG1* gene sequences and **(B)** from *Dyst* gene sequences.

*Amphiprion ocellaris*

*Similiparma lurida* (FJ616619) - Canary

*A. taurus* (FJ616625) - Cap Verde

*A. whitleyi* (FJ616626) - Lizard

100  
*A. taurus* (WJC85) - Puerto-Rico

100  
*A. taurus* (AY208622) - Panama

89  
94  
A. notatus (FJ616620) - Calamian

96  
A. sordidus - Indo-West Pacific (x11)

87  
A. septemfasciatus - Indo-West Pacific (x18)

*A. margariteus* (AY208618) - Réunion

41  
*A. sparoides* - Western Indian Ocean (x5)

99  
61  
*A. whitleyi* - New Caledonia (x3)

96  
*A. bengalensis* - Coral Triangle (x6)

95  
*A. vaigiensis* (EPS 432) - Glorieuses

97  
95  
*A. vaigiensis* (WJC4791) - Little Liuqiu

*A. vaigiensis* (WJC 4196) - West Papua

*A. vaigiensis* (WJC4193) - West Papua

*A. vaigiensis* (WJC2961) - Taiping

*A. vaigiensis* (WJC4194) - West Papua

97  
*A. vaigiensis* (WJC1293) - Caohagan

*A. vaigiensis* (WJC1295) - Caohagan

*A. vaigiensis* (WJC1296) - Caohagan

*A. vaigiensis* (WJC1294) - Caohagan

*A. vaigiensis* (WJC4854) - Bali

*A. vaigiensis* (WJC1292) - Caohagan

*A. vaigiensis* (WJC1291) - Caohagan

*A. vaigiensis* (WJC4851) - Bali

*A. vaigiensis* (WJC4846) - Bali

*A. vaigiensis* (WJC4840) - Bali

*A. vaigiensis* (lineage B)

*A. vaigiensis* (lineage C)

100  
*A. lorenzi* - Kota Kinabalu (x4)

100  
*A. sexfasciatus* (x67) + *A. vaigiensis* (lineage A) (x100)  
+ *A. saxatilis* (x5) + *A. hoeftleri* (x1) + *A. troschelii* (x1)

0.007

**Relative likelihood**

