RESEARCH ARTICLE

Contrasting genetic diversity and structure between endemic and widespread damselfishes are related to differing adaptive strategies

Vanessa Robitzch[1,2] | Pablo Saenz-Agudelo[2] | Tilman J. Alpermann[3] |
Bruno Frédérich[4] | Michael L. Berumen[1]

[1]Red Sea Research Centre, Division of Biological and Environmental Science and Engineering, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia

[2]Instituto de Ciencias Ambientales y Evolutivas, Facultad de Ciencias, Universidad Austral de Chile, Valdivia, Chile

[3]Senckenberg Research Institute and Natural History Museum Frankfurt, Frankfurt, Germany

[4]Laboratory of Functional and Evolutionary Morphology, FOCUS, University of Liège, Liège, Belgium

Correspondence
Vanessa Robitzch, Instituto de Ciencias Ambientales y Evolutivas, Universidad Austral de Chile, Campus Isla Teja, Casilla 567, Valdivia, Chile.
Email: vanessa.robitzchsierra@kaust.edu.sa

Funding information
Fondo Nacional de Desarrollo Científico y Tecnológico; King Abdullah University of Science and Technology; LOEWE-Hessen; Naturalis Biodiversity Center, Netherlands; Michel Claereboudt (Sultan Qaboos University); The Ministry of Environment and Climate Affairs

Handling Editor: Sergio Floeter

Abstract

Aim: Several marine biogeographical provinces meet at the Arabian Peninsula. Where and how these junctions affect species is poorly understood. We herein aimed to identify the barriers to dispersal and how these shape fish populations, leading to differing biogeographies despite shared habitat and co-ancestry.

Taxon: *Dascyllus marginatus* (endemic) and *Dascyllus abudafur* (widespread).

Location: Coral reefs from the Red Sea (RS), Djibouti, Yemen, Oman, and Madagascar.

Methods: We tested potential barriers to gene flow using RADseq-derived SNPs and identified whether population genetic differences on each side of these barriers were neutral or selective to relate this to the biogeography of the species. Seven locations (ranging over 5100 km) were sampled for the endemic and six (ranging over 7400 km) for the widespread species, taking 20 individuals per location, with two exceptions.

Results: *Dascyllus marginatus* populations (comprising 5648 SNPs) had an order of magnitude higher genetic differentiation compared to *D. abudafur* (comprising 10,667 SNPs), as well as several outlier loci that were absent in *D. abudafur* despite equal sampling locations. In both species, the RS and Djibouti specimens formed one genetic cluster separated from all other locations. Although ranging from the RS to Madagascar, *D. abudafur* was absent in Yemen and Oman.

Main Conclusions: Stronger genetic structure at smaller geographical scales and outlier loci in the endemic species seem associated with faster adaptation to environmental differences and selective pressure. Genetic differentiation in the widespread species is neutral and only occurs at large geographical distances. Restrictive transitions (between the Gulf of Aqaba and the RS or the RS and the Gulf of Aden) do not hinder gene flow in either species, and the environmental shift within the RS (at 22°N/20°N) only affected the endemic species. The genetic break in the Gulf of Aden likely reflects historical colonization processes and not contemporary environmental regimes.

KEYWORDS
coral reefs, *Dascyllus*, ddRAD, Djibouti, gene flow, Madagascar, Oman, Red Sea, SNPs, Yemen

1 | INTRODUCTION

Questions about today's species distribution and the origin of endemism (i.e. restricted geographical distributions) remain vividly debated among biogeographers (Mittermeier et al., 2005; Myers et al., 2000; Roberts et al., 2002). The distributions we see today are most likely the result of a combination of processes such as past migration events, local extinctions and/or range expansions of species that could have once been endemic. One particularly synthetic hypothesis proposed is the 'Gene pool—Ecological niche interaction hypothesis' (Stebbins, 1980), which suggests that a given species' ecological niche and its gene pool are the main determinants of its range size. Several factors shaping the ecological niche of a species, such as competition, adaptation, selective pressure or its genetic diversity, will influence a species' dispersal and, therefore, distribution. If endemism has been relatively well studied on terrestrial environments (Frankham, 1997; Stebbins, 1942), far less is understood about the patterns and processes that shape the distributions of species in the oceans. Tropical marine ecosystems, and especially coral reefs, are among the most diverse ecosystems on the planet. Explaining the actual biogeography of coral reef species and why some regions will favour and maintain a certain rate of endemism (Hughes et al., 2002) has become a central question in ecology and evolution. Many coral reef fish species differ in their geographical ranges, despite sharing habitats along their distributions. Herein, the environment, bathymetry, and oceanographic and geographical settings of these habitats are crucial for gene flow and for the maintenance or diversification of populations in space and time (Duputié & Massol, 2013; Heinz et al., 2009).

Our study targeted four different types of isolation processes that have been identified in the Western Indo-Pacific (after Saenz-Agudelo, DiBattista, et al., 2015; Spalding et al., 2007) and that have shaped connectivity and the distribution of coral reef biota across provinces and ecoregions. We focused on two congeneric and ecologically similar coral-dwelling damselfishes (Pomacentridae): *Dascyllus abudafur*, Forsskål 1775, widespread throughout the Western Indo-Pacific, and *Dascyllus marginatus*, Rüppell 1829, endemic to the Red Sea and Gulf of Aden. By comparing their population genetic composition, we assessed where and to which extent potential barriers to their dispersal lead to isolation and ultimately differences in biogeography and connectivity among populations.

In marine environments, (1) narrow connections or landmasses between oceanographic locations can be significant barriers for gene flow between regions leading to vicariance due to isolation-by-barrier (IBB; Rahel, 2007). For coral reef fishes dependent on a certain type of habitats, this barrier can also be posed by vast expanses of uninhabitable deep pelagic waters. (2) Populations can also differentiate genetically without the presence of an obvious barrier when isolation is a function of the distance between the populations, known as isolation-by-distance (IBD; reviewed in Hellberg, 2007; Jones et al., 2009; Selkoe & Toonen, 2011). In coral reef fishes, IBD is differently influenced by species-specific

mobility, range restrictive characteristics in adults and the dispersal capability of larvae. Both forementioned types of isolation lead to neutral differentiation of populations and eventually to the evolution of new species. Oppositely, (3) adaptation can also lead to genetic differentiation, generally driven by environmental differences (e.g. imposed by fluxes in temperature or water chemistry). This type of differentiation is called isolation-by-environment (IBE; Cooke et al., 2012; Zellmer et al., 2012). Lastly, (4) oceanographic features such as current regimes can act on neutral differentiation, for example, by altering the rate and/or direction of exchange and dispersal of larvae among locations with equal geographical distance. Such processes are categorized as isolation-by-oceanography (IBO) (García-De León et al., 2018; Gilg & Hilbish, 2003; Young et al., 2015).

The oceanographic, environmental and geographical features of the Arabian Seas and the Western Indian Ocean provide an excellent in situ laboratory for testing connectivity under IBB, IBD, IBE and IBO (Figure 1; e.g. Giles et al., 2015; Ngugi et al., 2012; Roberts et al., 2016; Robitzch et al., 2015) and their influence on biogeography. The Red Sea harbours some of the longest continuous coral reefs systems in the world (i.e. over 2000 km) and is fairly young. It started forming after the northeast African continental crust began drifting eastwards around 31 MYA (Bosworth et al., 2005), establishing a connection to the West Indian Ocean approximately 5 MYA through the Strait of Bab El Mandeb in the south (Bosworth et al., 2005; Siddall et al., 2003). The Gulf of Aqaba, located in the far northeast of the Red Sea, is home to the northernmost, coolest, most landlocked and most oligotrophic coral reefs in the world. The only connection the Gulf of Aqaba has to the Red Sea is the Strait of Tiran, 13 km wide and 290 m deep. Due to this geography, the Gulf of Aqaba may experience IBB and potentially IBE (due to its unique climate dominated by desert surroundings). Similarly, at the southern end of the Red Sea, the Strait of Bab el Mandeb is the only connection of the Red Sea with other tropical reef ecosystems. This very shallow and narrow strait (only 29 km wide and 137 m deep) is another well-founded example for IBB that has played an important role in the divergence of populations and the origin of high endemism (e.g. DiBattista et al., 2013, 2016; Klausewitz, 1989). During sea-level drops throughout the Pleistocene (2.58 MYA–0.01 MYA) and glacial maxima (including recent sea level fluctuation until about 11,700 years ago), these two shallow and narrow sills may have reduced connectivity, gene flow and dispersal of many marine species, potentially leading to the separation of populations on either side of the sill. Due to continuous coral reefs along the entire Red Sea coastline, and the lack of estuaries or strong current regimes, IBD can be studied without the influence of IBB or IBO. With latitude, the marine environment also gradually changes (in salinity, temperature and productivity) in the Red Sea, as does the bathymetry of its basin. More abruptly, a southward widening of the eastern and western continental shelves of Saudi Arabia and Eritrea leads to a transformation from steep and oligotrophic reef walls to shallower, warmer, nutrient-richer and more turbid reefs
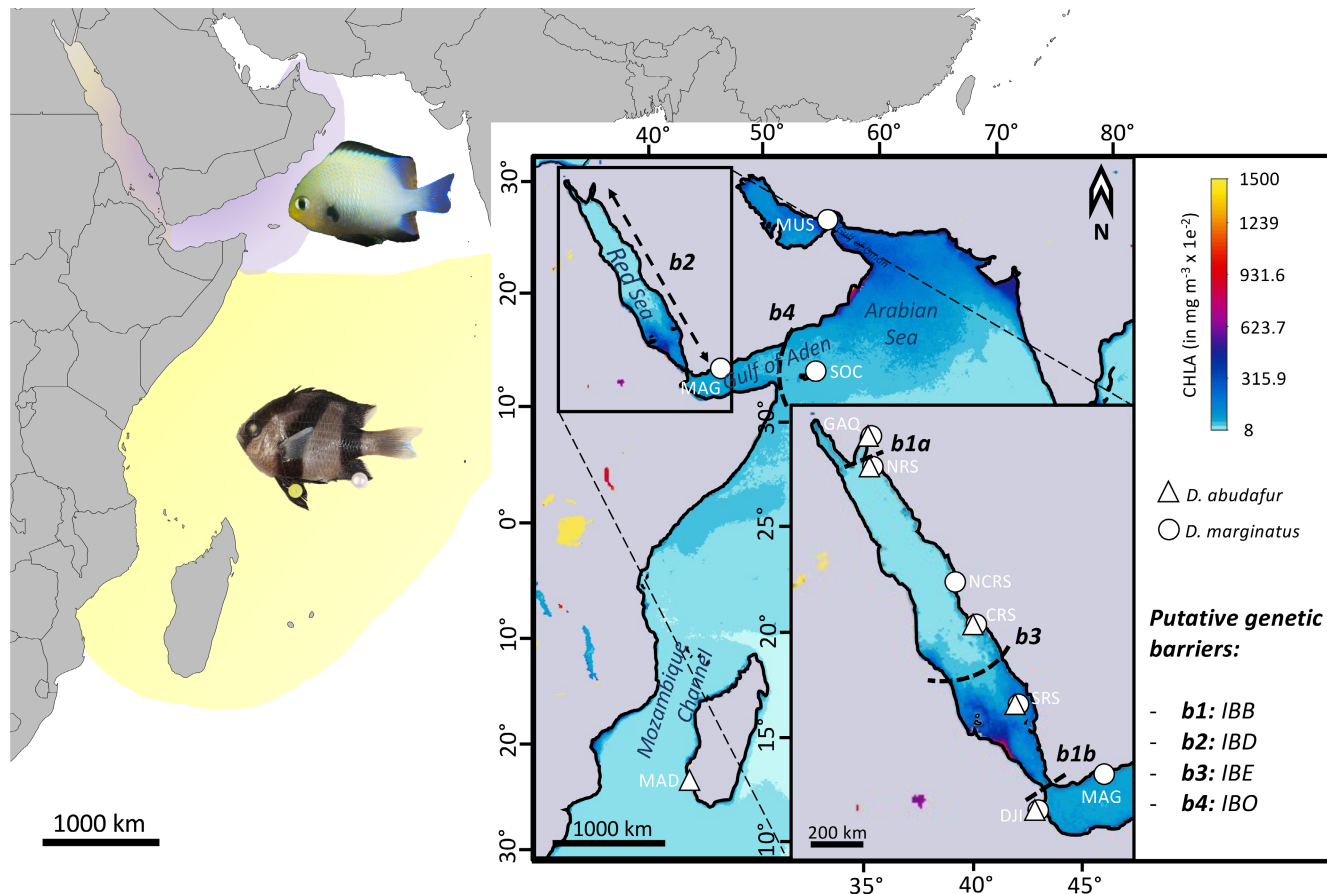
FIGURE 1 Sampling sites inside and outside the Red Sea (RS) for the assessment of single nucleotide polymorphisms (SNPs) data in *Dascyllus abudafur* (triangles) and *D. marginatus* (circles). The distribution range in the Western Indo-Pacific of the endemic *D. marginatus* is indicated in purple and that for the widespread *D. abudafur* in yellow, with an overlap of both distributions within the Red Sea. More information on the sites is found in Table 1. Average chlorophyll *a* concentrations (CHLA, in mg m^{-3} × 1 e^{-2}) of the Indian Ocean and the Red Sea are displayed from the NASA Giovanni website (http://oceancolor.gsfc.nasa.gov; for the season: December, January, and February from 2002 to 2019) to visualize approximate environmental differences among locations. Black dashed lines are putative barriers to gene flow among coral reef species: *b1*: IBB = 'Isolation by barrier' at the Strait of Tiran (*b1a*, northern Red Sea) and/or the Strait of Bab El Mandeb (*b1b*, southern Red Sea); *b2*: IBD = 'Isolation by distance' along the Saudi Arabian Red Sea coast with continuous coral reefs; *b3*: IBE = 'Isolation by environment' at the environmental transition between 16°N and 20°N within the Red Sea; and *b4*: IBO = 'Isolation by oceanography' at the oceanographic conditions present in the Gulf of Aden within the Oman upwelling region (*b1b*, *b3*, and *b4*, located senso Saenz-Agudelo et al., 2015).

in the south. These differences are a potential barrier to gene flow by IBE, previously identified between 16°N and 20°N (in Giles et al., 2015 and Nanninga et al., 2014; also reflected in fish communities in Roberts et al., 1992). Outside the Red Sea, entering the Indian Ocean, the marine environment changes substantially. The Gulf of Aden is highly impacted by monsoonal conditions, resulting in upwelling and IBO at the Yemeni-Omani coast. This region is also characterized by cool waters (Racault et al., 2015; Raitsos et al., 2013) and high abundances of macroalgae, seasonally induced by changes in the environment (i.e. additional IBE). Hence, cemented coral reef structures are less prominent (i.e. additional IBD; Kemp, 2000; Kemp & Benzoni, 2000). Due to this upwelling and IBO, connectivity from the Red Sea to nearby regions around the Arabian Peninsula may be restricted for coral reef-dependent organisms (e.g. DiBattista et al., 2015, 2020; Saenz-Agudelo, DiBattista, et al., 2015). Resolving more precisely the location of

this barrier within the Gulf of Aden requires broader sampling efforts along the continental Yemeni coast to increase resolution.

Dascyllus is a genus of coral reef damselfishes comprising 11 species distributed in the Indian and Pacific Oceans, with two widespread species complexes (i.e. *D. aruanus/abudafur* (Borsa et al., 2014) and *D. trimaculatus* (Leray et al., 2010)) that share habitats with other local/endemic *Dascyllus* species. Our two studied species, *D. marginatus* and *D. abudafur*, commonly occur in branching corals throughout the Red Sea. Outside the Red Sea, the regional endemic *D. marginatus* is ubiquitous in coral habitats along the southern end of the Arabian Peninsula, from Djibouti to Musandam (the innermost Gulf of Oman). At Socotra, *D. marginatus* is present, but *D. abudafur* is absent. The widespread *D. abudafur*, however, can be found throughout most of the rest of the Western Indian Ocean. The two species are biologically and ecologically very similar but there are notable differences in pelagic

TABLE 1 Sampling sites for the preparation of genetic libraries of single nucleotide polymorphisms (SNPs) for the species *Dascyllus marginatus* (Dm) and *D. abudafur* (Da) from locations across the Western Indo-Pacific realm, including number of samples, site names, location acronyms (Abb.) and coordinates

Province	Ecoregion	Site name	Abb.	Lib. prep. N (Dm)	Data set N (Dm)	Lib. prep. N (Da)	Data set N (Da)	Latitude	Longitude
RS	Gulf of Aqaba	Haql	GAQ	20	9	20	13	29°15′11.45′′N	34°56′20.11′′E
	North/Central RS	Burqan	NRS	20	20	20	18	27°54′35.8″N	35°03′55.4″E
		Thuwal	NCRS	4	4	-	-	22°17′46.80′′N	39°04′26.64′′E
	South RS	Al Lith	CRS	16	9	20	15	20°08′05.10′′N	40°06′04.39′′E
		Farasan	SRS	20	15	20	17	16°37′07.9″N	41°55′58.9″E
		Djibouti	DJI	20	16	20	18	11°43′16.03″N	43°09′53.94″E
		Total		100	73	100	81		
Arabian	Gulf of Aden	Magateen	MAG	20	20	-	-	13°24′38.3″N	46°22′36.3″E
	West Arabian Sea	Socotra	SOC	20	20	-	-	12°36′17.3″N	54°21′03.1″E
	Gulf of Oman	Musandam	MUS	20	17	-	-	26°13′08.9″N	56°17′02.7″E
WIO	W/N Madagascar	Madagascar	MAD	-	-	20	19	23°22′59.68″S	43°38′25.70″E
		Total		60	57	20	19		
ALL				160	**130**	130	**100**		

Note: Locations are split into the provinces Red Sea (RS), Arabian and West Indian Ocean (WIO), following Spalding et al. (2007). Six sites are inside the RS-Djibouti ecoregion and four additional ones in the wider West Indo-Pacific. The sites have a letter-coded abbreviation (Abb.). The total number of samples used for the generation of SNP libraries is given as 'Lib. Prep. N' and the number of samples which were retained for the analysis of genetic is given as 'Data set N' for Dm and Da, respectively.

larval duration (PLD) plasticity as a response to environmental conditions (Robitzch et al., 2016). With biogeographical overlap, contrasting distribution ranges and potentially differing adaptive strategies and responses to environmental changes (Chen et al., 2019; Robitzch et al., 2016), these two *Dascyllus* species are good models to assess the relationship between genetic structure and distribution range in coral reef fishes.

We assessed genetic signatures of isolation processes at potential barriers to dispersal using single nucleotide polymorphism (SNP) data for these *Dascyllus* species across their distributions (similar to Casas et al., 2021). We hypothesized that in the scenarios of evidence for IBD, IBB and IBO, we will find neutral genetic differentiation between the populations on each side of the barrier due to limited gene flow and divergence via restricted dispersal and genetic drift (Riginos & Liggins, 2013; Slatkin, 1987; Wright, 1943). In the case of IBE, these populations will instead show signatures of selective differentiation, which leads to divergence via adaptation and restricted/selective gene flow even if there is high initial dispersal among the populations (Nosil et al., 2009; Via, 2002).

2 | MATERIALS AND METHODS

2.1 | Study region and sampling

Dascyllus marginatus was collected between 2009 and 2015 from nine sites around the Arabian Peninsula (from the Gulf of Aqaba (28°N, 34°E) to Musandam (12°N, 43°E)) and across potential dispersal barriers (*b1 to b4*, Figure 1). Simultaneously, *D. abudafur* was

collected from six sites: the Gulf of Aqaba, three sites along the eastern coast of the Red Sea, Djibouti and the Mozambique Channel (Toliara, Madagascar). Although it has a wider distribution, *D. abudafur* was absent from the continental Yemeni coast ('Magateen'), Socotra Island and the inner Gulf of Oman ('Musandam') (Table 1). Its sampling comprised the provinces of the Red Sea and Gulf of Aden, and the Western Indian Ocean (following Spalding et al., 2007). In total, 160 *D. marginatus* and 130 *D. abudafur* were used to create double digest restriction-associated DNA (ddRAD) libraries and the generation of SNP data (Table 1). This study did not involve live specimens or endangered or protected species.

2.2 | DNA extraction, SNP library preparation and sequencing

Genomic DNA was extracted from fin or gill tissue preserved in 96% ethanol using a Nucleospin-96 Tissue kit (Macherey-Nagel). Four genomic libraries were prepared, using 500 ng of high-quality DNA per sample, following the protocol described by Peterson et al. (2012) with some modifications (see Appendix S1). Sequencing took place on four lanes of a HiSeq 2000 Illumina sequencer (single end reads, 1 × 101 bp; v3 reagents).

2.3 | De-novo assembly

Sequences were demultiplexed and filtered for quality using the 'process_radtags.pl' pipeline in STACKS v.2.5 (Catchen et al., 2011). Individual reads with uncalled bases (-c), low-quality score (- q) and

phred-scores below an average of 20 (on a sliding window, -s = 20) were discarded. RAD-tags and barcodes were rescued (-r). After demultiplexing, individuals with less than 500,000 reads recovered were removed. Loci for the remaining samples were assembled *de-novo* using the 'denovo_map.pl' pipeline (STACKS v.2.5). Different parameter combinations were evaluated, which resulted in different numbers of loci but gave similar results in genetic comparisons (i.e. genetic clustering and pairwise *F*ST among sites). For the final data-set, the maximum number of mismatches allowed between stacks within and between individuals (for ustacks and cstacks, respectively) was set to three (-M and -n), similar to recommendations by Mastretta-Yanes et al. (2015).

2.4 | SNP filtering

The '*populations*' component of STACKs v.2.5 was used to create and export a vcf file containing SNPs (only the first per locus to be called, --write-single-snp) that were present in at least 95% of the samples (-R 0.95), with a minimum allele frequency of 0.02 (--min-maf) and a maximum observed heterozygosity of 0.6 (--max-obs-het); and to calculate population statistics (--fstats).

2.5 | Population genetics and statistics

Several analyses were conducted to characterize the genetic structure and compare the presence of barriers to gene flow in the endemic vs. the widespread damselfish. A principal component analysis (PCA) summarized graphically the distribution of genetic variation among individual samples; and PCA plots were built in R v.3.6.2 (in R Studio v.1.2.5033) using the packages *poppr* (Kamvar et al., 2014) and *ggplot2* (Kamvar et al., 2015). Genetic clustering analysis using *K* values from 1 to the total number of sampling locations of the species were performed for 10 replicates in the R-package *LEA* (Frichot & François, 2015) using the snmf function (Frichot et al., 2014). Both analyses only used polymorphic loci and individuals missing less than 10% data, and were performed to elucidate the spatial patterns of genetic admixture along the distribution range of *D. abudafur* and *D. marginatus*. Input files were generated with the genomic_converter function from the R-package *radiator* (Gosselin et al., 2020) transforming the vcf from 'populations' (STACKS v.2.5) into a STRUCTURE format. Plots were built for each species separately using (1) all samples or (2) solely Red Sea samples. Expected and observed heterozygosity and fixation indices (F_{IS}) among variant sites; and π values for all sites were calculated by 'populations' and were used as proxies for genetic diversity/richness within a location. To quantify the genetic differentiation among locations, the distribution of pair-wise F'STs (Meirmans, 2006) between locations was investigated. Pairwise F'STs were calculated using 'populations'. Pairwise ge-netic distance matrices were tested for significant relations to pairwise geographical and environmental distance matrices using

Mantel tests and 10,000 permutations (GenAIEx v.6.501; Peakall & Smouse, 2006, 2012). The matrices for pairwise geographic dis-tances were generated taking the shortest possible path along the marine trajectory between each pair of sites, measured in Google Earth Pro v.7.3.2.5776. Environmental distances between sites were calculated using satellite gridded environmental data from BioOracle (Tyberghein et al., 2012) provided by E. F. Cagua follow-ing the methodology described in DiBattista et al., (2020; scripts in https://github.com/efcaguab/fish_data_matrix), to reduce the dimensionality of the data and perform PCAs. The environmental variables were log-transformed and normalized to have a mean of zero and unit variance, a priori. From these computations, we took the PC1 values to generate the pairwise environmental distances.

To investigate the nature of putative genetic differentiation within the Red Sea in both species, we searched for loci putatively under natural selection. The Reversible Jump Markov Chain Monte Carlo from BayeScan v.2.1 (Fischer et al., 2011; Foll et al., 2010; Foll & Gaggiotti, 2008) detects outlier loci based on the decomposition of the overall genetic variation into a population-specific and a locus-specific component. The analyses were done with pre-set chain pa-rameters and only changing the likelihood of the neutral model (i.e. to 10, 100 or 1000-fold) in three independent runs to verify that differ-ences between datasets were independent from these priors at a false discovery rate of 0.05. The same BayeScan tests were run between the populations at which we found evidence of a genetic break with LEA to assess whether or not the break may be a result of selective pressure.

3 | RESULTS

3.1 | SNP data

From a total of 274,622,464 reads, 248,877,760 were retained (91%) in *D. marginatus*. In *D. abudafur,* from 259,982,400 reads, 236,921,182 were retained (91%). In *D. marginatus*, 24 individuals had under 500,000 recovered reads and were removed; 17 were re-moved in *D. abudafur* using the same cut-off. The STACKS catalogue comprised 1,095,692 loci with an average coverage of 13.7x for 136 *D. marginatus*; and 630,763 loci with an average coverage of 11.7× for 107 *D. abudafur*. After filtering, the final dataset contained 5648 SNPs for the former and 10,667 SNPs for the latter species.

3.2 | Population genetics and statistics

Basic population statistics suggested highest expected heterozygo-sity, π and FIS at the entrance to the Red Sea, in the Gulf of Tadjoura (DJI, Table 2), for *D. marginatus*; while the highest number of private al-leles was found at the easternmost location of the Arabian Peninsula, in the Strait of Hormuz, inner Gulf of Oman (MUS Table 2). Preliminary analysis on genetic structure also suggested the site of Djibouti to be genetically highly connected to the Red Sea sites, for which we here-after refer to these locations jointly as the Red Sea-Djibouti region.

TABLE 2 Summary statistics of population genetics of *Dascyllus marginatus* and *Dascyllus abudafur* SNPs data from locations across its distribution in the Western Indo-Pacific realm

Location	He$_{obs}$	He$_{exp}$	π	F_{IS}	Priv. alleles
Dascyllus marginatus					
GAQ	0.1163	0.1133	0.00094	0.0135	0
NRS	0.1174	0.1170	0.00094	0.0162	0
NCRS	0.1153	0.1035	0.00093	0.0073	0
CRS	0.1193	0.1158	0.00096	0.0135	0
SRS	0.1232	0.1209	0.00098	0.0110	0
DJI	0.1208	0.1312	0.00106	0.0696	0
MAG	0.1017	0.0999	0.00080	0.0113	2
MUS	0.1025	0.1010	0.00082	0.0129	5
SOC	0.1002	0.0997	0.00080	0.0126	1
Dascyllus abudafur					
GAQ	0.1519	0.1485	0.00119	0.0113	1
NRS	0.1554	0.1507	0.00120	0.0046	3
CRS	0.1531	0.1497	0.00120	0.0102	0
SRS	0.1557	0.1528	0.00122	0.0112	1
DJI	0.1540	0.1532	0.00122	0.0167	0
MAD	0.1604	0.1622	0.00129	0.0270	106

Note: Observed (He$_{obs}$) and expected (He$_{exp}$), genetic diversity (π), fixation indices (F_{IS}) and private alleles were calculated with the 'populations' component of STACKS v.2.5. over all sites. More information on the sampling location is found in Table 1. For π values among all sites, please see Figure 2.

For the widespread *D. abudafur*, basic population statistics indicated highest heterozygosity, π FIS, and number of private alleles to be outside the Red Sea-Djibouti region and much further away, in Madagascar (MAD, Table 2). The differences were not significant and much smaller than in *D. marginatus* for similar sampling locations. Genetic diversity was also generally higher among *D. abudafur* compared to *D. marginatus* and increased from north to south within the Red Sea-Djibouti region in both species. Outside the Red Sea-Djibouti region, the genetic diversity of the endemic species decreased, while it increased for *D. abudafur*. Additionally, private alleles were found at almost every location inside the Red Sea for *D. abudafur*, except for the central Red Sea (CRS) and the entrance to the Red Sea (DJI) (Table 2). The second highest indicators of genetic diversity/richness were also found just outside the Red Sea (DJI, Table 2).

3.3 | Distribution of genetic variation and clustering

The PCA and genetic clustering results were based on two datasets for each species. In *D. marginatus* it comprised 5648 polymorphic loci and 130 individuals for all nine sampling sites; and 5155 polymorphic loci and 73 individuals for six Red Sea-Djibouti sites. For *D. abudafur*, analyses among all six sampling sites comprised 10,667 polymorphic loci for 100 individuals, and 10,604

polymorphic loci for 81 individuals among the five Red Sea-Djibouti locations. Among all sites, the first axis of the PCAs of both species segregated individuals from inside versus outside the Red Sea, where those from Djibouti clustered with Red Sea sites, forming 'the Red Sea-Djibouti' cluster (Figure 2a). Among *D. marginatus*, the second axis of the PCA differentiated two clusters within the Red Sea-Djibouti region, separating sites above and below 22°N/20°N latitude, where Djibouti clustered with southern Red Sea sites, except for one specimen, which clustered with the Yemeni-Omani cluster (see Figure 2a, top plot). Contrastingly, among *D. abudafur*, Madagascar was the only location displaying variability along the second axis, while all Red Sea-Djibouti sites densely clustered (Figure 2a, bottom plot). In *D. marginatus*, solely discriminating among sites within the Red Sea-Djibouti region (Figure 2b), the Djibouti 'outlier' belonging to the Yemeni-Omani cluster remained separated (Figure 2b, top plot). For *D. abudafur*, no further discrimination among Red Sea-Djibouti locations was present, but a subtle arrangement of individuals by latitude became evident along the first axis (Figure 2b, bottom plot), where the furthest differing individuals belonged to the most distant locations within the Red Sea-Djibouti cluster (in orange, the Gulf of Aqaba/GAQ; and in grey, Djibouti/DJI). In the PCA of Yemeni-Omani *D. marginatus*, all three sites (MAG, MUS and SOC) clustered together except for two outliers (in red and pink), which interestingly belonged to the sites at the western and eastern extremes of the Arabian Peninsula (MAG and MUS, respectively, Figure 2c), and not to the isolated island of Socotra (SOC). Overall, the relatively low explanatory power of the PCAs was lowest for the widespread species and more so among Red Sea-Djibouti samples due to decreased genetic differentiation within that region.

The genetic clustering barplots (from *LEA*) supported the PCAs and displayed a clear differentiation in *D. marginatus* between the Red Sea-Djibouti region and sites outside of it (see Figure 3a, top plot). The minimal cross-entropy analysis also suggested $K = 2$ as the total number of populations (K). Here, the Djibouti outlier was represented by an exclusive, red bar, confirming that its genetic background belonged to the Yemeni-Omani cluster. Among *D. abudafur*, a similar split between the Red Sea-Djibouti and the Madagascar samples was present (Figure 3a, bottom plot), with the minimal cross-entropy also suggesting $K = 2$.

Inside the Red Sea-Djibouti cluster (Figure 3b), the barplots and minimal cross-entropy plots differed between the two species: there was evidence for a genetic break and two distinct populations for *D. marginatus* within the Red Sea-Djibouti region (Figure 3b, top plots), while only one population was suggested among *D. abudafur* (i.e. the smallest minimal cross-entropy is for $K = 1$, Figure 2b, bottom plots). Nonetheless, the Red Sea population seemed to have disparities in ancestry among locations and samples; and the genetic composition did not seem homogeneous across the Red Sea. A more detailed picture of these genetic disparities was attempted by generating barplots for the highest yet likely number of K for the two species ($K = 4$, for *D. marginatus* and $K = 5$, for *D. abudafur*, Figure 3b). Among these higher K values, a north–south genetic
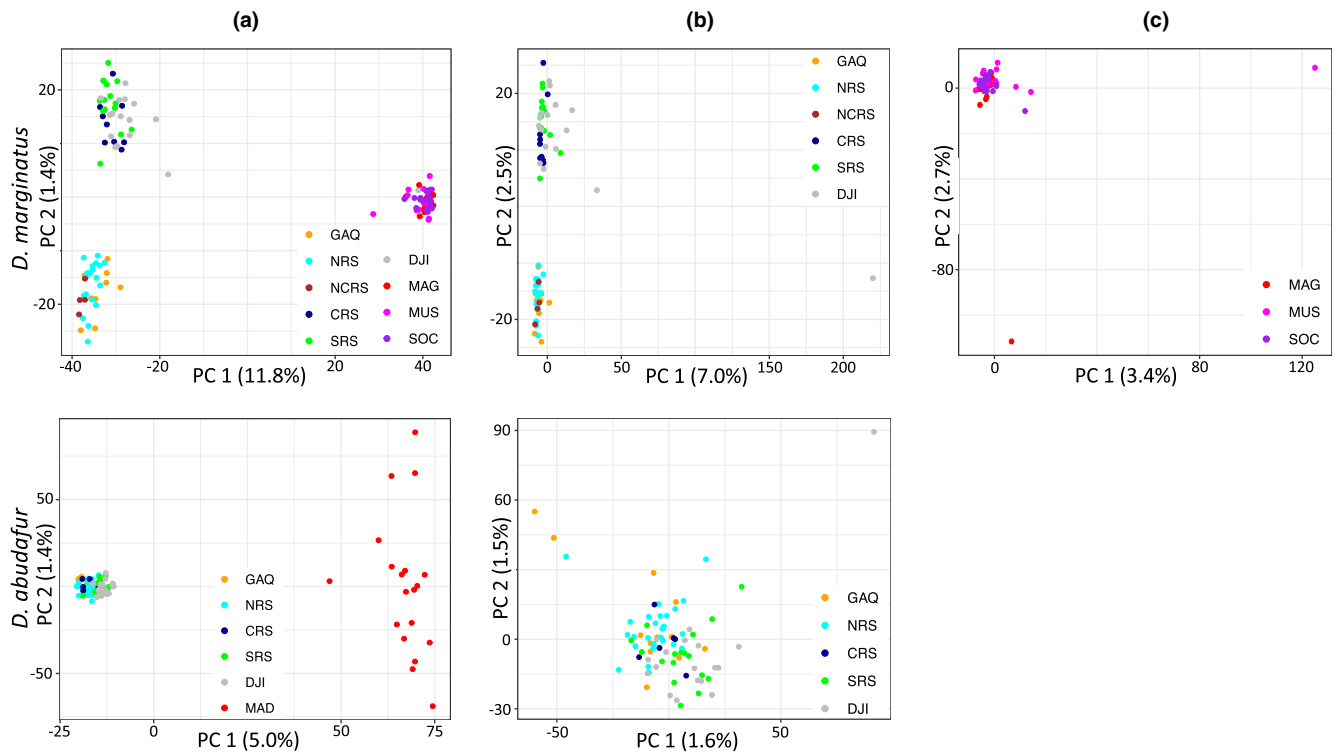
FIGURE 2 Principal component analysis (PCA) using the first and second principal components (PC1 and PC2, respectively) comparing sites inside and outside the Red Sea-Djibouti region for the endemic damselfish *Dascyllus marginatus* (plots in the top row) and the widespread *D. abudafur* (plots in the bottom row). Panel (a) displays clustering among all the sampling sites, panel (b) only for sites within the Red Sea-Djibouti region and panel (c) only for sites outside the Red Sea-Djibouti region.

gradient seemed to structure the hypothetical populations of *D. abudafur*, resulting in three main clusters: one north, one central/south and one south/far-south, in red, orange and green, respectively (Figure 3b, bottom right). Contrastingly, the genetic differentiation remained abrupt in *D. marginatus*: the barplots in Figure 3b (top) display a division of the Red Sea-Djibouti region into two populations, one north of 22°N (comprising the NCRS, the NRS and the GAQ sites, Figure 3b, top right) and the other one south of 20°N (including the CRS, the SRS and the DJI sites, Figure 3b, top right); and the genetically Yemeni-Omani sample from Djibouti forms its own population (a unique red line) in the *K* = 4 barplot.

3.4 | Genetic distances and connectivity

Pairwise *F*'ST comparisons were in line with the former two analyses. The highest differentiation in the endemic *D. marginatus* was between Red Sea-Djibouti sites and those from the Yemeni-Omani region, followed by sites below and above the 22°N/20°N latitudinal break inside the Red Sea-Djibouti region. *F*'ST comparisons comprising the sites NCRS and/or CRS (located at the latitudinal break) could be affected by their low sample sizes. Nonetheless, the consistency in the overall genetic differentiation of the species among all aforementioned analyses supports our main results. Further interesting was that Socotra (SOC) did not seem more isolated than any other Yemeni-Omani site along the Gulf of Aden and

around the southern Arabian Peninsula (SOC vs. MAG and SOC vs. MUS, Appendix S2). Among *D. abudafur*, Madagascar was most isolated (with highest *F*'ST values; MAD, Appendix S3), while all other sites were genetically similarly isolated from another. Genetic distance (pairwise *F*'STs) in relation to geographical distances within the Red Sea-Djibouti region revealed a stronger correlation among *D. abudafur* than *D. marginatus* (Figure 4). However, the genetic differentiation of *D. marginatus* was of an order of magnitude higher than that of *D. abudafur*. Oppositely, the correlation between genetic and environmental distances was stronger in *D. marginatus* than in *D. abudafur*. Although all correlations were marginally significant, the genetic distance was better explained by the geographic distance in *D. abudafur*, while the genetic distance in *D. marginatus* was relatively better explained by the environmental distance (Figure 4). The natural correlation between environmental and geographical distances, present in the latitudinal environmental gradient of the Red Sea (Mantel's *r* = 0.979; *p* = 0.008), complicates the independent interpretation of these correlations.

3.5 | Loci under selection

BayeScan was used to analyse the possibility of isolation due to natural selective barriers to gene flow (such as IBE and IBO). Here, the analyses on outlier loci were done excluding the Yemeni-Omani outlier of Djibouti and gave consistent results independent from
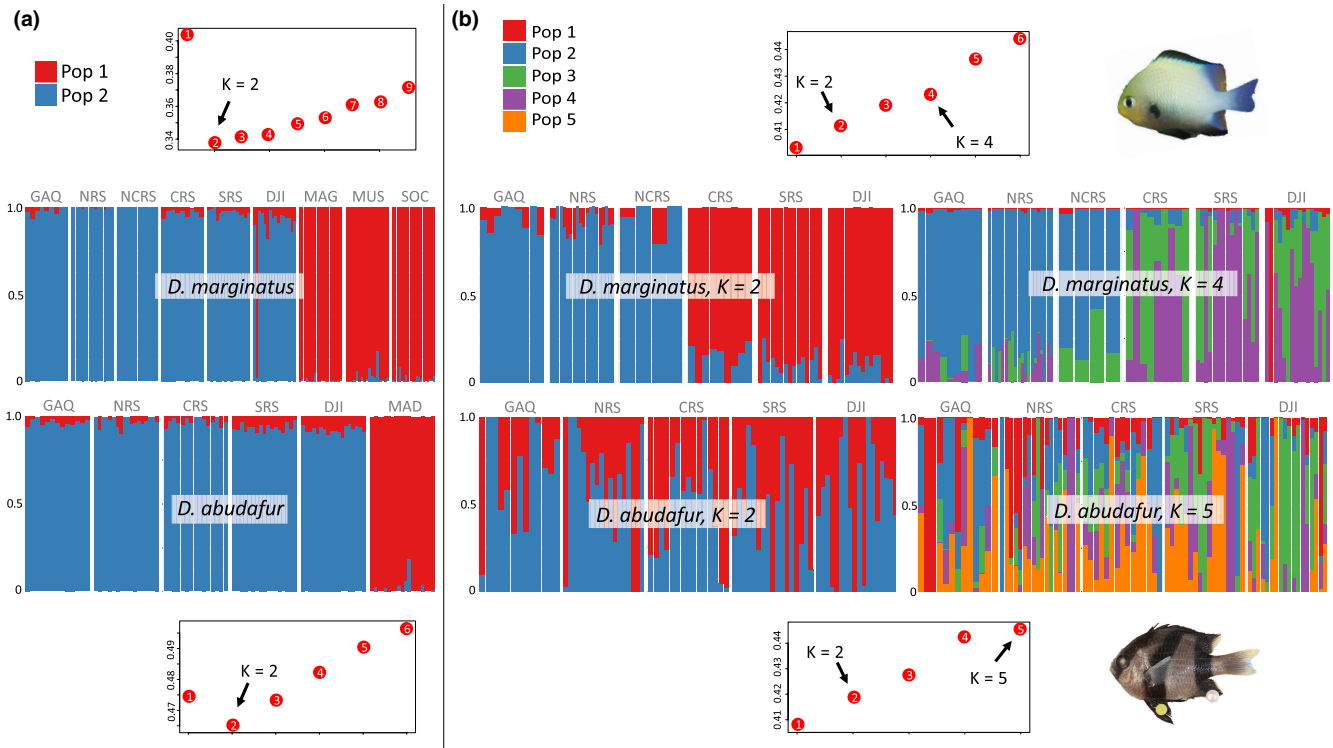
FIGURE 3 Genetic clustering analysis (R package *LEA*) for single nucleotide polymorphisms (SNPs) data of *Dascyllus marginatus* (top row) and *D. abudafur* (bottom row), from the Western Indo-Pacific. The results are divided into those for the entire data (panel a) and those only inside the Red Sea-Djibouti region (panel b). The graphs at the periphery of each panel give the most likely number of populations (K, *x*-axis) using the minimum cross-entropy (*y*-axis): Top plots for *D. marginatus* and bottom plots for *D. abudafur*. The barplots in the centre of each panel display the proportional probability of assignment of an individual into the different inferred populations of origin (K's), each of which has a unique colour. Each individual is represented by one individual bar and individuals are separated into groups by a thicker white line according to their sampling location, given at the top of each group of individuals (with a letter code as in Table 1).

the chosen priors of the neutral model (i.e. 10, 100 or 1000). No loci under natural selection were detected for the widespread *D. abudafur*, while several were detected in the endemic *D. marginatus*. The number of these loci under selection in *D. marginatus* were in 69, 32 and 18 loci for the priors of 10, 100 and 1000, respectively. However, the pairwise analyses of outliers between the populations of *D. marginatus* on either side of the genetic breaks detected with LEA did not identify any outlier loci. Within the Red Sea, this could be due to the low sampling size of NCRS and CRS. The outlier loci found within the Red Sea discriminated solely three individuals: one from SRS and another two from GAQ (see PCA in Appendix S4).

4 | DISCUSSION

We found that two biologically and ecologically similar damselfish species displayed different patterns of genetic structure in the Red Sea and Western Indian Ocean. Genetic breaks were not found at the Strait of Tiran nor at the Strait of Bab el Mandeb, suggesting that IBB is no longer an active barrier to dispersal in the region. Genetic diversity diminished drastically for populations east of the Red Sea-Djibouti region for the endemic species but was similar across all sampled sites for the widespread species. The genetic differentiation

among populations of the endemic species was an order of magnitude higher compared to those of the widespread species. Strong differentiation between the Red Sea-Djibouti specimens and the Yemeni-Omani specimens suggests a deep gap in the genetic connectivity of *D. marginatus* or even the presence of a putative cryptic species. Altogether, our results indicated that various evolutionary processes were shaping differences in genetic structure and that the previously recognized genetic IBO/IBE barrier in the Gulf of Aden might be more specifically located at the entrance to the Red Sea, between Yemen and Djibouti. This IBO/IBE barrier is not directly related to upwelling conditions but potentially may be the combined outcome of historical IBB at the Strait of Bab El Mandeb and species-specific IBE, demographics and colonization events.

4.1 | Genetic structuring, connectivity and adaptation relative to isolating processes in the endemic *D. marginatus*

Much higher genetic differentiation between populations inside vs. outside the Red Sea-Djibouti region were found in the endemic *D. marginatus* than in the widespread *D. abudafur* or *D. trimaculatus* (see Salas et al., 2019 for the latter). We identified two main clusters in
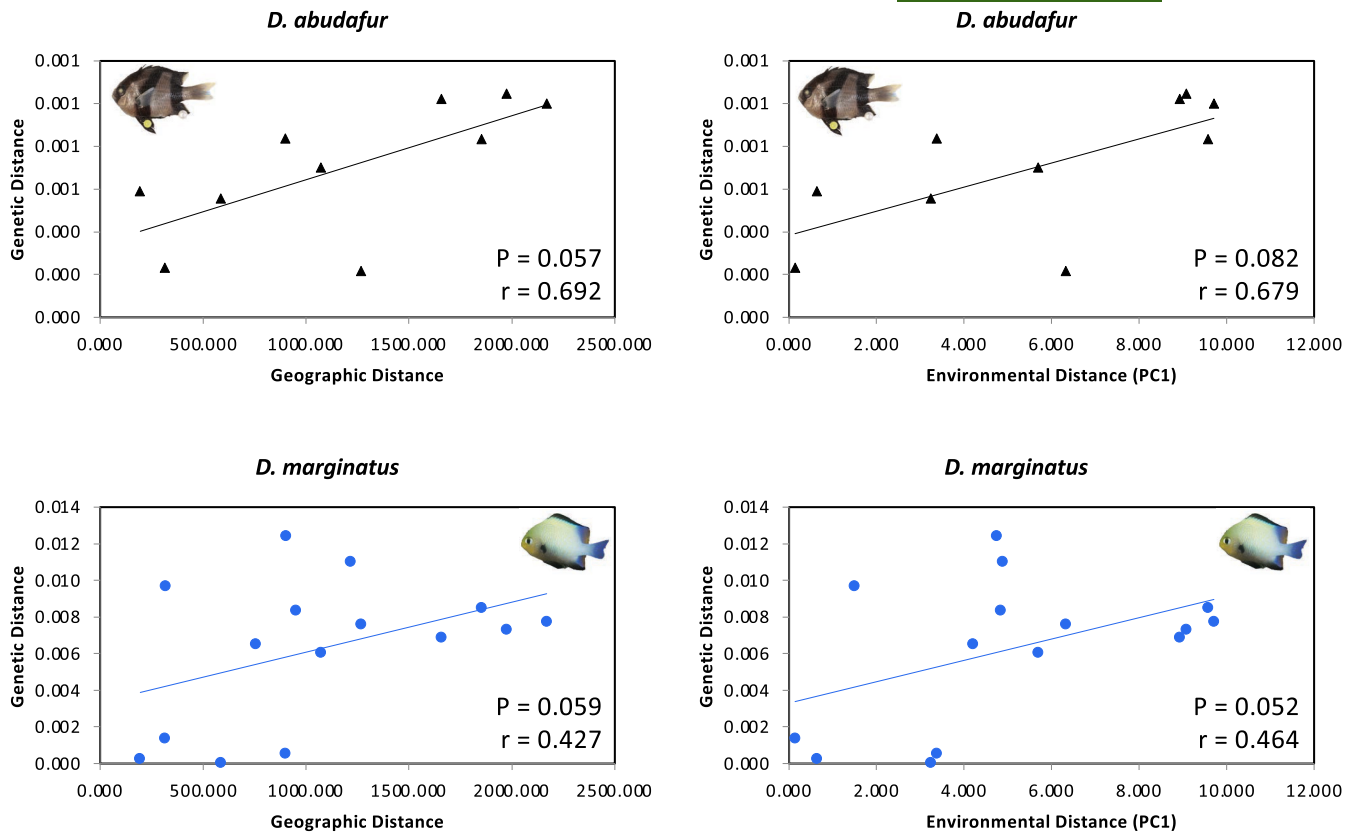
FIGURE 4 Correlation plots of pairwise genetic distances ($F'ST$, *y*-axis) with geographical (left panels, in km, *x*-axis) or environmental distance (right panels, PC1, *x*-axis) for single nucleotide polymorphism (SNPs) data of *Dascyllus abudafur* (top row, values in black) and *D. marginatus* (bottom row, values in blue) for all locations in the Western Indo-Pacific versus only those inside the Red Sea/Djibouti region. Coefficients of determination (R^2) for linear regression are given in each plot.

D. marginatus: one comprising the Red Sea-Djibouti region and another comprising the Yemeni-Omani region (including Socotra). The genetic break was located along the western continental Yemeni coast, far from and consequently not driven by the Omani upwelling. We also lacked evidence for natural selection acting at the upwelling barrier, leading us to eliminate IBE as a possible barrier. The highest genetic diversity in *D. marginatus* was found immediately adjacent to the entrance to the Red Sea, in the Gulf of Tadjoura, Djibouti, which may suggest a larger population size of the Red Sea-Djibouti endemic compared to the Yemeni-Omani endemic (Kimura, 1983). Alternatively, this possibly reflects the junction of two historically vicariant populations in the Gulf of Tadjoura. In Djibouti, all samples genetically belonged to the Red Sea population, except for one individual with Yemeni-Omani genetic signature. This implies sporadic migration at least in one direction: from the Yemeni-Omani population to the Red Sea-Djibouti population. However, because we find no evidence of introgression or hybridization, further phylogenetic analysis of these two highly divergent clusters is needed to investigate potential cryptic speciation and reproductive isolation. The origin for this speciation and isolation is most likely historical IBB at the Strait of Bab el Mandeb during sea-level drops in the Pleistocene 5 MYA, which may have been strengthened by differing environmental conditions (IBE). Demographic analyses may help reconstruct extinction and colonization events at the entrance of the Red Sea.

Among Yemeni-Omani *D. marginatus*, we found the highest number of private alleles in the easternmost location (the Strait of Hormuz) in the Gulf of Oman. The peripheral location of that site, in combination with harsher environmental changes at reefs of higher latitudes, may be an additional selective force (D'Agostino et al., 2021).

Within the Red Sea-Djibouti *D. marginatus*, we found further population structure dividing two populations: one north and one south of 22°N/20°N. The location of this genetic break is likely linked to abruptly differing environmental conditions around these latitudes, but an examination of selection was restricted by low sample sizes at these latitudes. Similar genetic splits within the Red Sea have been reported in a sea anemone (25°N; Emms et al., 2020), a clownfish (19°N; Nanninga et al., 2014) and a sponge (16°N; Giles et al., 2015). We suggest that although IBE within the Red Sea is likely present throughout various taxa, the exact environmental threshold active as a selective barrier may be species specific, related to differences in life histories.

4.2 | Genetic structuring, connectivity and adaptation relative to isolating processes in the widespread *D. abudafur*

The widespread *D. abudafur* also showed genetic differentiation inside vs. outside the Red Sea-Djibouti region, maintaining the

same clustering pattern of Red Sea specimens with Djibouti specimens and no sign of restricted connectivity at the Strait of Bab el Mandeb, eliminating IBB as a mechanism. The collection site of *D. abudafur* outside the Red Sea-Djibouti region was Madagascar and thus an order of magnitude more distant than the Yemeni-Omani sites sampled for *D. marginatus*. Despite the much larger geographical distance, the genetic distance between the furthermost populations of *D. abudafur* was a magnitude lower than in *D. marginatus*. When comparing geographically equidistant sites (i.e. the north Red Sea vs Djibouti), the widespread species showed an eighth of the genetic differentiation found in the endemic species. Furthermore, there was no sign of selective pressure among populations of the widespread species, even from the Gulf of Aqaba to Madagascar. The genetic differences found between distant geographical ecoregions were solely linked to neutral genetic drift via IBD. Although none of the locations within the Red Sea-Djibouti region were significantly structured in *D. abudafur*, a relationship between genetic and geographical distance was present in this widespread species, but not in the endemic *D. marginatus*. We concluded that the widespread species is more capable of wide-ranging dispersal and much higher gene flow/connectivity between distant locations than the endemic, even extending beyond ecoregions and putative environmental barriers. Additionally, a larger population size of *D. abudafur* compared to *D. marginatus* may lead to less genetic structure in the former. A recent study on *D. abudafur* larvae reported high numbers of genetically diverse siblings within a single recruitment cohort, suggesting panmixia and a large population size within the Red Sea (Robitzch et al., 2020). Recent findings on phenotypic variation in *D. abudafur* (Chen et al., 2019) also warrant further testing of the relationship between dispersal capabilities, adaptation and plasticity.

Lastly, we interpret the notably high number of private alleles in Madagascar (i.e. nearly twice as diverse as the Red Sea-Djibouti) as an indication that the Red Sea-Djibouti population is possibly marginal. It may have been founded by a less diverse population that, although panmictic and well connected, remained genetically poorer but without indication of bottlenecks or inbreeding. Hence, the Red Sea-Djibouti specimens could represent a fairly young population (much younger than the population of the endemic *D. marginatus*) that still lacks signatures of selection/adaptation. Sampling more populations outside the Red Sea-Djibouti region along the eastern African coastline as well as demographic studies on historical population growth and migration rates may help address some of these remaining questions.

4.3 | What is defining contrasting genetic structures and biogeographies?

Two main results were striking: (1) *D. abudafur* is widespread and does not show signs of selection (driven by IBE), but it is not present further east and beyond the Red Sea-Djibouti region into the Yemeni-Omani region (where IBO/IBE or IBB may be occurring along the southern Arabian Peninsula). Contrastingly, (2) the genetic structure and connectivity of the endemic *D. marginatus* seem affected by differing environmental conditions but not by oceanographic phenomena such as upwelling at the Gulf of Aden (IBO/IBE) and inhabits areas east of the Gulf of Aden. While the biogeographical range of the widespread species seems delimited by the challenging environmental conditions generated by oceanographic phenomena (IBO/IBE), the biogeographical range of the endemic species may rather be delimited by population sizes, dispersal potential (IBD), and adaptation to and selection by local environmental conditions (IBE). The widespread species may maintain connectivity over large distances but may not adapt quickly enough to contrasting environments and is thus specialized to dispersal and plasticity rather than adaptation. Contrarily, the endemic species may adapt so quickly and successfully that once it finds a niche, it differentiates into a new population, maintaining smaller population sizes and increasing genetic drift through a feedback loop. Consequently, it is specialized to local adaptation at the cost of dispersal, a potential hallmark of endemism.

On first consideration, our results may seem contradictory to recent findings on population genetics of endemic versus widespread butterflyfishes and surgeonfishes. DiBattista et al. (2020) showed that widespread species (those inhabiting the Red Sea and Arabian Seas) were genetically structured but not the endemic species (those restricted either to the Red Sea or the Arabian Seas). They concluded that the range limits for Red Sea endemics were largely due to upwelling in the Gulf of Aden but they did not have samples from continental Yemen (DiBattista et al., 2020). Our findings are concordant with their conclusions, despite the much stronger genetic divergence we found in the endemic *D. marginatus* compared to the widespread *D. abudafur*. First, the stronger genetic structure in *D. marginatus* is putatively indicative of two different species and therefore two endemics: one endemic to the Red Sea-Djibouti region and the other endemic to the Yemeni-Omani region. In this scenario, the Yemeni-Omani endemic would not show signs of genetic structuring and seems 'adapted to a diversity of local conditions' such as strong seasonal changes related to upwelling in the Gulf of Aden (as suggested among endemics by DiBattista et al. (2020)). Second, DiBattista et al. (2020) suggested that widespread species may be subject to 'ecological filtering, where different combinations of genotypes persist under divergent ecological regimes'. Our data support a similar conclusion. First, the widespread *D. abudafur* is absent in the upwelling region. Second, the absence of genetic structure in neutral loci in this species could be because neutral markers in its ancestral population were 'ecologically filtered' and not tightly linked to adaptation. Accordingly, genetic structuring among the differently adapted endemic species is to be expected. DiBattista et al. (2020) may have missed these signs of genetic differentiation among endemics because they primarily focused on *Chaetodon* butterflyfishes that have extraordinarily fast speciation rates (Bellwood et al., 2010) (Hemingson et al., 2019). Additionally, *Chaetodon* are broadcast spawners with long PLDs. Therefore, subtle signs of genetic structure may not be captured at the species level as new endemics could

emerge quickly. Furthermore, previous comparative phylogeographical studies in other regions have suggested that population genetic differences in sympatric species of other marine taxa, such as gobies (Dawson et al., 2002) or limpets (Bird et al., 2007), are related to differences in dispersal potential and microhabitat use. However, biogeographical studies conducted in the in situ laboratory of the Arabian region highlight the complexity of understanding barriers to dispersal among coral reef fishes and the challenge of discerning common patterns across taxa (if common patterns exist at all, see DiBattista et al., 2013).

In conclusion, our results emphasize that the deep genetic break found along the Yemeni coast is the result of historic processes or the combination of historic and environmental processes. Within seemingly similar fishes, there may be overlooked or subtle differences in life histories that are important determinants in structuring their populations. Moreover, the coast of continental Yemen seems to hold crucial information to understand the evolutionary history of the fish fauna of this region, but remains undersampled. Increased sampling, past demographic reconstructions and the implementation of full genomic approaches for species ranging throughout the Gulf of Aden would help elucidate common drivers of genetic differentiation within this unique collision point of biogeographical provinces.

CONFLICT OF INTEREST

The authors declare no competing interests.

DATA AVAILABILITY STATEMENT

The genomic SNPs data as well as barcodes and population information used for the preparation of libraries and downstream processing in STACKs v2.5 are deposited in Dryad under doi: https://doi.org/10.5061/dryad.59zw3r29b. The raw data from Illumina next generation sequencing of the four ddRAD libraries can be found in the NCBI SRA repository under SUB11121710 (BioProject: PRJNA809575) and SUB11132900 (BioProject: PRJNA810838).

ORCID

Vanessa Robitzch https://orcid.org/0000-0002-6965-8498
Pablo Saenz-Agudelo https://orcid.org/0000-0001-8197-2861
Michael L. Berumen https://orcid.org/0000-0003-2463-2742

REFERENCES

Bellwood, D. R., Klanten, S., Cowman, P. F., Pratchett, M. S., Konow, N., & Van Herwerden, L. (2010). Evolutionary history of the butterflyfishes (f: Chaetodontidae) and the rise of coral feeding fishes. *Journal of Evolutionary Biology*, 23(2), 335–349. https://doi.org/10.1111/j.1420-9101.2009.01904.x

Bird, C. E., Holland, B. S., Bowen, B. W., & Toonen, R. J. (2007). Contrasting phylogeography in three endemic Hawaiian limpets (*Cellana* spp.) with similar life histories. *Molecular Ecology*, 16(15), 3173–3186.

Borsa, P., Sembiring, A., Fauvelot, C., & Chen, W.-J. (2014). Resurrection of Indian Ocean humbug damselfish, *Dascyllus abudafur* (Forsskål) from synonymy with its Pacific Ocean sibling, *Dascyllus aruanus* (L.). *Comptes Rendus Biologies*, 337(12), 709–716. https://doi.org/10.1016/j.crvi.2014.09.001

Bosworth, W., Huchon, P., & McClay, K. (2005). The Red Sea and Gulf of Aden basins. *Journal of African Earth Sciences*, 43(1–3), 334–378. https://doi.org/10.1016/j.jafrearsci.2005.07.020

Casas, L., Saenz-Agudelo, P., Villegas-Ríos, D., Irigoien, X., & Saborido-Rey, F. (2021). Genomic landscape of geographically structured colour polymorphism in a temperate marine fish. *Molecular Ecology*, 30(5), 1281–1296. https://doi.org/10.1111/mec.15805

Catchen, J. M., Amores, A., Hohenlohe, P., Cresko, W., & Postlethwait, J. H. (2011). Stacks: Building and genotyping loci *de novo* from short-read sequences. *Genes, Genomes, Genetics*, 1, 171–182. https://doi.org/10.1534/g3.111.000240

Chen, C. T., Robitzch, V., Sturaro, N., Lepoint, G., Berumen, M. L., & Frédérich, B. (2019). "Homemade": The phenotypic diversity of coral reef damselfish populations is driven by the local environment. *Biological Journal of the Linnean Society*, 127(2), 361–376. https://doi.org/10.1093/biolinnean/blz049

Cooke, G. M., Chao, N. L., & Beheregaray, L. B. (2012). Divergent natural selection with gene flow along major environmental gradients in Amazonia: Insights from genome scans, population genetics and phylogeography of the characin fish *Triportheus albus*. *Molecular Ecology*, 21(10), 2410–2427.

D'Agostino, D., Burt, J. A., Santinelli, V., Vaughan, G. O., Fowler, A. M., Reader, T., Taylor, B. M., Hoey, A. S., Cavalcante, G. H., Bauman, A. G., & Feary, D. A. (2021). Growth impacts in a changing ocean: Insights from two coral reef fishes in an extreme environment. *Coral Reefs*, 40(2), 433–446.

Dawson, M. N., Louie, K. D., Barlow, M., Jacobs, D. K., & Swift, C. C. (2002). Comparative phylogeography of sympatric sister species, *Clevelandia ios* and *Eucyclogobius newberryi* (Teleostei, Gobiidae), across the California transition zone. *Molecular Ecology*, 11(6), 1065–1075. https://doi.org/10.1046/j.1365-294X.2002.01503.x

DiBattista, J. D., Berumen, M. L., Gaither, M. R., Rocha, L. A., Eble, J. A., Choat, J. H., Craig, M. T., Skillings, D. J., & Bowen, B. W. (2013).

After continents divide: Comparative phylogeography of reef fishes from the Red Sea and Indian Ocean. *Journal of Biogeography*, *40*(6), 1170–1181. https://doi.org/10.1111/jbi.12068

DiBattista, J. D., Roberts, M. B., Bouwmeester, J., Bowen, B. W., Coker, D. J., Lozano-Cortés, D. F., Howard Choat, J., Gaither, M. R., Hobbs, J. P. A., Khalil, M. T., Kochzius, M., Myers, R. F., Paulay, G., Robitzch, V., Saenz-Agudelo, P., Salas, E., Sinclair-Taylor, T. H., Toonen, R. J., Westneat, M. W., … Berumen, M. L. (2016). A review of contemporary patterns of endemism for shallow water reef fauna in the Red Sea. *Journal of Biogeography*, *43*(3), 423–439. https://doi.org/10.1111/jbi.12649

DiBattista, J. D., Rocha, L. A., Hobbs, J. A., He, S., Priest, M. A., Sinclair-taylor, T. H., Bowen, B. W., & Berumen, M. L. (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. https://doi.org/10.1111/jbi.12526

DiBattista, J. D., Saenz-Agudelo, P., Piatek, M. J., Cagua, E. F., Bowen, B. W., Choat, J. H., Rocha, L. A., Gaither, M. R., Hobbs, J. P. A., Sinclair-Taylor, T. H., McIlwain, J. H., Priest, M. A., Braun, C. D., Hussey, N. E., Kessel, S. T., & Berumen, M. L. (2020). Population genomic response to geographic gradients by widespread and endemic fishes of the Arabian peninsula. *Ecology and Evolution*, *10*(10), 4314–4330. https://doi.org/10.1002/ece3.6199

Duputié, A., & Massol, F. (2013). An empiricist's guide to theoretical predictions on the evolution of dispersal. *Interface Focus*, *3*, 20130028.

Emms, M. A., Saenz-Agudelo, P., Giles, E. C., Gatins, R., Nanninga, G. B., Scott, A., Hobbs, J. A., Frisch, A. J., Mills, S. C., Beldade, R., & Berumen, M. L. (2020). Comparative phylogeography of three host sea anemones in the Indo-Pacific. *Journal of Biogeography*, *47*(2), 487–500. https://doi.org/10.1111/jbi.13775

Fischer, M. C., Foll, M., Excoffier, L., & Heckel, G. (2011). Enhanced AFLP genome scans detect local adaptation in high-altitude populations of a small rodent (*Microtus arvalis*). *Molecular Ecology*, *20*(7), 1450–1462.

Foll, M., Fischer, M. C., Heckel, G., & Excoffier, L. (2010). Estimating population structure from AFLP amplification intensity. *Molecular Ecology*, *19*(21), 4638–4647.

Foll, M., & Gaggiotti, O. (2008). A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. *Genetics*, *180*(2), 977–993.

Frankham, R. (1997). Do Island populations have less genetic variation than mainland populations? *Heredity*, *78*(3), 311–327.

Frichot, E., & François, O. (2015). LEA: An R package for landscape and ecological association studies. *Methods in Ecology and Evolution*, *6*(8), 925–929.

Frichot, E., Mathieu, F., Trouillon, T., Bouchard, G., & François, O. (2014). Fast and efficient estimation of individual ancestry coefficients. *Genetics*, *196*(4), 973–983.

García-De León, F. J., Galvan-Tirado, C., Sanchez Velasco, L., Silva-Segundo, C. A., Hernandez-Guzman, R., Barriga-Sosa, I. d. l. A., Díaz Jaimes, P., Canino, M., & Cruz-Hernandez, P. (2018). Role of oceanography in shaping the genetic structure in the North Pacific hake *Merluccius productus*. *PLoS ONE*, *13*(3), e0194646.

Giles, E. C., Saenz-Agudelo, P., Hussey, N. E., Ravasi, T., & Berumen, M. L. (2015). Exploring seascape genetics and kinship in the reef sponge *Stylissa carteri* in the Red Sea. *Ecology and Evolution*, *5*(13), 2487–2502. https://doi.org/10.1002/ece3.1511

Gilg, M. R., & Hilbish, T. J. (2003). The geography of marine larval dispersal: Coupling genetics with fine-scale physical oceanography. *Ecology*, *84*(11), 2989–2998.

Gosselin, T., Lamothe, M., Devloo-Delva, F., & Grewe, P. (2020). Radiator: RADseq data exploration, manipulation and visualization using R. R package version 1.1.5.

Heinz, S. K., Mazzucco, R., & Dieckmann, U. (2009). Speciation and the evolution of dispersal along environmental gradients. *Evolutionary Ecology*, *23*(1), 53–70. https://doi.org/10.1007/s10682-008-9251-7

Hellberg, M. E. (2007). Footprints on water: The genetic wake of dispersal among reefs. *Coral Reefs*, *26*(3), 463–473. https://doi.org/10.1007/s00338-007-0205-2

Hemingson, C. R., Cowman, P. F., Hodge, J. R., & Bellwood, D. R. (2019). Colour pattern divergence in reef fish species is rapid and driven by both range overlap and symmetry. *Ecology Letters*, *22*(1), 190–199. https://doi.org/10.1111/ele.13180

Hughes, T. P., Bellwood, D. R., & Connolly, S. R. (2002). Biodiversity hotspots, centres of endemicity, and the conservation of coral reefs. *Ecology Letters*, *5*(6), 775–784. https://doi.org/10.1046/j.1461-0248.2002.00383.x

Jones, G. P., Almany, G. R., Russ, G. R., Sale, P. F., Steneck, R. S., Oppen, M. J. H., & Willis, B. L. (2009). Larval retention and connectivity among populations of corals and reef fishes: History, advances and challenges. *Coral Reefs*, *28*(2), 307–325. https://doi.org/10.1007/s00338-009-0469-9

Kamvar, Z. N., Brooks, J. C., & Grünwald, N. J. (2015). Novel R tools for analysis of genome-wide population genetic data with emphasis on clonality. *Frontiers in Genetics*, *6*, 208.

Kamvar, Z. N., Tabima, J. F., & Grünwald, N. J. (2014). Poppr: An R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ*, *2*, e281.

Kemp, J. M. (2000). Zoogeography of the coral reef fishes of the North-Eastern Gulf of Aden, with eight new records of coral reef fishes from Arabia. *Fauna of Arabia*, *18*(October), 293–321.

Kemp, J. M., & Benzoni, F. (2000). A preliminary study of coral communities in the northern Gulf of Aden. *Fauna of Arabia*, *18*(November 2000), 67–86.

Kimura, M. (1983). *The neutral theory of molecular evolution*. Cambridge University Press.

Klausewitz, W. (1989). Evolutionary history and zoogeography of the Red Sea ichthyofauna. *Fauna of Saudi Arabia*, *10*, 310–337.

Leray, M., Beldade, R., Holbrook, S. J., Schmitt, R. J., Planes, S., & Bernardi, G. (2010). Allopatric divergence and speciation in coral reef fish: The three-spot dascyllus, *Dascyllus trimaculatus*, species complex. *Evolution; International Journal of Organic Evolution*, *64*(5), 1218–1230. https://doi.org/10.1111/j.1558-5646.2009.00917.x

Mastretta-Yanes, A., Arrigo, N., Alvarez, N., Jorgensen, T. H., Piñero, D., & Emerson, B. C. (2015). Restriction site-associated DNA sequencing, genotyping error estimation and de novo assembly optimization for population genetic inference. *Molecular Ecology Resources*, *15*(1), 28–41. https://doi.org/10.1111/1755-0998.12291

Meirmans, P. G. (2006). Using the AMOVA framework to estimate a standardized genetic differentiation measure. *Evolution*, *60*(11), 2399–2402.

Mittermeier, R. A., Gil, P. R., Hoffman, M., Pilgrim, J., Brooks, T., Goettsch-Mittermeier, C., Lamoreux, J., & DaFonseca, G. A. B. (2005). *Hotspots revisited: Earth's biologically richest and most endangered terrestrial ecoregions* (1st ed.). Conservation International.

Myers, N., Mittermeier, R. A., Mittermeier, C. G., Da Fonseca, G. A. B., & Kent, J. (2000). Biodiversity hotspots for conservation priorities. *Nature*, *403*(6772), 853–858.

Nanninga, G. B., Saenz-Agudelo, P., Manica, A., & Berumen, M. L. (2014). Environmental gradients predict the genetic population structure of a coral reef fish in the Red Sea. *Molecular Ecology*, *23*(3), 591–602. https://doi.org/10.1111/mec.12623

Ngugi, D. K., Antunes, A., Brune, A., & Stingl, U. (2012). Biogeography of pelagic bacterioplankton across an antagonistic temperature-salinity gradient in the Red Sea. *Molecular Ecology*, *21*(2), 388–405. https://doi.org/10.1111/j.1365-294X.2011.05378.x

Nosil, P., Funk, D. J., & Ortiz-Barrientos, D. (2009). Divergent selection and heterogeneous genomic divergence. *Molecular Ecology*, *18*(3), 375–402. https://doi.org/10.1111/j.1365-294X.2008.03946.x

Peakall, R., & Smouse, P. E. (2006). GenAlEx 6: Genetic analysis in excel. Population genetic software for teaching and

research. *Molecular Ecology Notes*, *6*(1), 288–295. https://doi.org/10.1111/j.1471-8286.2005.01155.x

Peakall, R., & Smouse, P. E. (2012). GenAlEx 6.5: Genetic analysis in excel. Population genetic software for teaching and research - an update. *Bioinformatics (Oxford, England)*, *28*(19), 2537–2539. https://doi.org/10.1093/bioinformatics/bts460

Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., & Hoekstra, H. E. (2012). Double digest RADseq: An inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS ONE*, *7*(5), e37135. https://doi.org/10.1371/journal.pone.0037135

Racault, M.-F., Raitsos, D. E., Berumen, M. L., Brewin, R. J. W., Platt, T., Sathyendranath, S., & Hoteit, I. (2015). Phytoplankton phenology indices in coral reef ecosystems: Application to ocean-color observations in the Red Sea. *Remote Sensing of Environment*, *160*, 222–234. https://doi.org/10.1016/j.rse.2015.01.019

Rahel, F. J. (2007). Biogeographic barriers, connectivity and homogenization of freshwater faunas: It's a small world after all. *Freshwater Biology*, *52*(4), 696–710.

Raitsos, D. E., Pradhan, Y., Brewin, R. J. W., Stenchikov, G., & Hoteit, I. (2013). Remote sensing the phytoplankton seasonal succession of the Red Sea. *PLoS ONE*, *8*(6), e64909. https://doi.org/10.1371/journal.pone.0064909

Riginos, C., & Liggins, L. (2013). Seascape genetics: Populations, individuals, and genes marooned and adrift. *Geography Compass*, *7*(3), 197–216.

Roberts, C., McClean, C., & Veron, J. (2002). Marine biodiversity hotspots and conservation priorities for tropical reefs. *Science*, *295*(5558), 1280–1284. http://www.sciencemag.org/content/295/5558/1280.short

Roberts, C. M., Shepherd, A. R. D., & Ormond, R. F. G. (1992). Large-scale variation in assemblage structure of Red Sea butterflyfishes and angelfishes. *Journal of Biogeography*, *19*(6), 239–250. http://www.jstor.org/stable/2845449

Roberts, M. B., Jones, G. P., McCormick, M. I., Munday, P. L., Neale, S., Thorrold, S., Robitzch, V., & Berumen, M. L. (2016). Homogeneity of coral reef communities across 8 degrees of latitude in the Saudi Arabian Red Sea. *Marine Pollution Bulletin*, *105*(2), 558–565. https://doi.org/10.1016/j.marpolbul.2015.11.024

Robitzch, V., Banguera-Hinestroza, E., Sawall, Y., Al-Sofyani, A., & Voolstra, C. R. (2015). Absence of genetic differentiation in the coral *Pocillopora verrucosa* along environmental gradients of the Saudi Arabian Red Sea. *Frontiers in Marine Science*, *2*(February), 1–10. https://doi.org/10.3389/fmars.2015.00005

Robitzch, V., Lozano-Cortés, D., Kandler, N. M., Salas, E., & Berumen, M. L. (2016). Productivity and sea surface temperature are correlated with the pelagic larval duration of damselfishes in the Red Sea. *Marine Pollution Bulletin*, *105*(2), 566–574. https://doi.org/10.1016/j.marpolbul.2015.11.045

Robitzch, V., Saenz-Agudelo, P., & Berumen, M. L. (2020). Travel with your kin ship! Insights from genetic sibship among settlers of a coral damselfish. *Ecology and Evolution*, *10*(15), 8265–8278. https://doi.org/10.1002/ece3.6533

Saenz-Agudelo, P., DiBattista, J. D., Piatek, M. J., Gaither, M. R., Harrison, H. B., Nanninga, G. B., & Berumen, M. L. (2015). Seascape genetics along environmental gradients in the Arabian peninsula: Insights from ddRAD sequencing of anemonefishes. *Molecular Ecology*, *24*(24), 6241–6255. https://doi.org/10.1111/mec.13471

Salas, E. M., Bernardi, G., Berumen, M. L., Gaither, M. R., & Rocha, L. A. (2019). RADseq analyses reveal concordant Indian Ocean biogeographic and phylogeographic boundaries in the reef fish *Dascyllus trimaculatus*. *Royal Society Open Science*, *6*(5), 1–9. https://doi.org/10.1098/rsos.172413

Selkoe, K., & Toonen, R. (2011). Marine connectivity: A new look at pelagic larval duration and genetic metrics of dispersal. *Marine Ecology Progress Series*, *436*, 291–305. https://doi.org/10.3354/meps09238

Siddall, M., Rohling, E. J., Almogi-Labin, A., Hemleben, C., Meischner, D., Schmeltzer, I., & Smeed, D. A. (2003). Sea-level fluctuations during the last glacial cycle. *Nature*, *423*, 853–858. https://doi.org/10.1038/nature01687.1

Slatkin, M. (1987). Gene flow and the geographic structure of natural populations. *Science*, *236*(4803), 787–792.

Spalding, M. D., Fox, H. E., Allen, G. R., Davidson, N., Ferdaña, Z. A., Finlayson, M. A. X., Halpern, B. S., Jorge, M. A., Lombana, A. L., Lourie, S. A., Martin, K. D., Mcmanus, E., Molnar, J., Recchia, C. A., & Robertson, J. (2007). Marine ecoregions of the world: A bioregionalization of coastal and shelf areas. *Bioscience*, *57*(7), 573–583.

Stebbins, G. L. (1942). The genetic approach to problems of rare and endemic species. *Madrono*, *6*(8), 241–258.

Stebbins, G. L. (1980). Rarity of plant species: A synthetic viewpoint. *Rhodora*, *82*(829), 77–86.

Tyberghein, L., Verbruggen, H., Pauly, K., Troupin, C., Mineur, F., & De Clerck, O. (2012). Bio-ORACLE: A global environmental dataset for marine species distribution modelling. *Global Ecology and Biogeography*, *21*(2), 272–281.

Via, S. (2002). The ecological genetics of speciation. *The American Naturalist*, *159*(S3), S1–S7.

Wright, S. (1943). Isolation by distance. *Genetics*, *28*(2), 114.

Young, E. F., Belchier, M., Hauser, L., Horsburgh, G. J., Meredith, M. P., Murphy, E. J., Pascoal, S., Rock, J., Tysklind, N., & Carvalho, G. R. (2015). Oceanography and life history predict contrasting genetic population structure in two Antarctic fish species. *Evolutionary Applications*, *8*(5), 486–509.

Zellmer, A. J., Hanes, M. M., Hird, S. M., & Carstens, B. C. (2012). Deep phylogeographic structure and environmental differentiation in the carnivorous plant *Sarracenia alata*. *Systematic Biology*, *61*(5), 763–777.

BIOSKETCH

The authors' interests are focused on assessing gene flow, its barriers, and thus, the evolutionary processes structuring populations and leading to speciation of marine species. The authors have mainly targeted population and evolutionary genetic approaches in combination with morphological, biological and ecological features to elucidate drivers of marine biodiversity.

Author contributions: V.R. conceived the idea for this study, collected the tissue samples, produced SNP libraries, analysed the data and led the writing. In addition to writing and the joint interpretation of results, P.S-A., T.A., B.F. and M.L.B further analysed the data, collected the tissue samples and/or provided computing facilities.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.