Zoological Studies **62:**39 (2023)
doi:10.6620/ZS.2023.62-39

Zoological Studies

Open Access

Phylogeography of Long-spined Sea Urchin *Diadema setosum* Across the Indo-Malay Archipelago

Indra Bayu Vimono[1,2,3], Philippe Borsa[4,*], Régis Hocdé[5], and Laurent Pouyaud[3]

[1]*National Research and Innovation Agency Republic of Indonesia (BRIN), Research Center for Oceanography (RCO), Jakarta, Indonesia. E-mail: vimono@gmail.com (Vimono)*
[2]*Université de Montpellier, Ecole doctorale Gaia, Montpellier, France*
[3]*Institut de recherche pour le développement (IRD), UMR 226 ISEM, Montpellier, France. E-mail: laurent.pouyaud@ird.fr (Pouyaud)*
[4]*IRD, UMR 250 Entropi, Montpellier, France. *Correspondence: E-mail: philippe.borsa@ird.fr (Borsa)*
[5]*MARBEC, Univ Montpellier, CNRS, Ifremer, IRD, Montpellier, France. E-mail: regis.hocde@ird.fr (Hocdé)*

Received 26 November 2022 / Accepted 26 May 2023 / Published 26 July 2023
Communicated by Machida Ryuji

Widely distributed, broadcast-spawning *Diadema* sea urchins have been used as model invertebrate species for studying the zoogeography of the tropical Indo-Pacific. So far, the Indo-Malay archipelago, a wide and geographically complex maritime region extending from the eastern Indian Ocean to the western Pacific Ocean, has been under-sampled. This study aims to fill this sampling gap and uncover the phylogeographic structure of the long-spined sea-urchin *D. setosum* in the central Indo-West pacific region. *D. setosum* samples (total $N = 718$) were collected in 13 sites throughout the Indo-Malay archipelago. We sequenced over 1157 bp of *COI* gene. The Phylogeographic structure was derived from pairwise Φ_{ST} estimates using multidimensional scaling and hierarchical clustering analysis; biogeographic hypotheses were tested by AMOVA; genetic relationships between haplotypes were summarised in the form of a minimum-spanning network; and pairwise mismatch distributions were compared to the expectations from demographic and spatial expansion models. All samples from the Indo-West Pacific were of the previously uncovered *D. setosum*-a lineage. Phylogeographic structure was evident: the Andaman Sea population and the northern New Guinea population were genetically distinct. Subtler but significant haplotype-frequency differences distinguished two populations within the Indonesian seas, distributed in a parapatric-like fashion. The phylogeographic partition observed was insufficiently explained by previous biogeographic hypotheses. The haplotype network showed a series of closely related star-shaped haplogroups with a high proportion of singletons. Nucleotide-pairwise mismatch patterns in the two populations from the Indonesian seas were consistent with both demographic and spatial expansion models. While geographic barriers to gene flow were inferred at the western and eastern extremities of the Indo-Malay archipelago, the subtler parapatric pattern observed within the Indonesian seas indicated restriction in gene flow, in a fashion that can hardly be explained by geographic isolation given the dynamic current systems that cross this region. Our results thus raise the hypothesis of subtle reproductive isolation between ecologically incompatible populations. While the coalescence pattern of the Andaman-Sea population suggested demographic stability over evolutionary timescales, that of the two populations from the Indonesian seas indicated recent population expansion, possibly linked to the rapid changes in available *D. setosum* habitat caused by sea-level oscillations in the late Pleistocene. The phylogeographic patterns observed in this study point to likely allopatric differentiation in the central Indo-West Pacific region. Genetic differences between populations were likely reinforced during interglacials by some form of reproductive isolation.

Key words: Cytochrome oxidase subunit I (*COI*), Geographic barrier, Reproductive isolation, Demographic history, Indo-West Pacific

Citation: Vimono IB, Borsa P, Hocdé R, Pouyaud L. 2023. Phylogeography of long-spined sea urchin *Diadema setosum* across the Indo-Malay archipelago. Zool Stud **62:**39. doi:10.6620/ZS.2023.62-39.

BACKGROUND

Long-spined sea urchins of the genus *Diadema* play an important role in the ecology of coral reef ecosystems. They regulate marine algae thus helping to maintain or restore coral recruitment in coral reefs (Birkeland 1989; Carpenter and Edmunds 2006; Mumby et al. 2007; Muthiga and McClanahan 2007; Lessios 2016). Mass mortality of *Diadema* populations can contribute to phase shifts of tropical reefs from coral-dominated to algae-dominated communities (Lessios 2016). Conversely, high densities of *Diadema* sea urchins can result in increased bioerosion and, eventually, in the destruction of coral reefs (Dumont et al. 2013; Qiu et al. 2014). The demographic history of long-spined sea urchin populations may thus be linked to ecological phase shifts in modern coral reefs.

Diadema sea urchins have been used as model invertebrate species for studying marine speciation (Lessios et al. 2001). Of all extant *Diadema* lineages, the lineage ancestral to *D. setosum* separated early from the rest, presumably in the Miocene, then speciated 3–5 million years ago into two species, one ("*D. setosum*-b") that exclusively occurs in the northwestern Indian Ocean including the Red Sea and more recently the Levantine basin, the other one ("*D. setosum*-a") with wide Indo-West Pacific distribution (Lessios et al. 2001; Bronstein and Kroh 2018). Moderate mitochondrial differentiation was evident among *D. setosum*-a populations across the Indo-West Pacific (Lessios et al. 2001), consistent with its extended larval life (Dautov and Dautova 2016). Lessios et al. (2001) concluded that the phylogeographic structure uncovered in *D. setosum* may reflect either the effect of the Indo-Pacific barrier or isolation by distance.

So far, the Indo-Malay archipelago, a wide and geographically complex maritime region extending from the eastern Indian Ocean to the western Pacific Ocean, has been under-sampled. For instance, Lessios et al.'s (2001) extensive phylogeographic study only included one sample from the Indo-Malay archipelago. Apart from a DNA-barcoding study of *Diadema* spp. around Sulawesi Island, which included 14 individuals of *D. setosum* (Moore et al. 2019), we are unaware of any other phylogeographic study of *D. setosum* dedicated to studying patterns and processes of genetic differentiation of this species in the central Indo-West Pacific.

We expect that denser sampling in the Indo-Malay archipelago will lead to a refined understanding of the phylogeographic structure of *D. setosum* and a more precise ability to locate potential geographic barriers. The aim of the present work is to fill the sampling gap for *D. setosum* in the central Indo-West Pacific, provide a more accurate description of its phylogeographic structure (using nucleotide sequences of the mitochondrial marker *COI*), and infer related micro-evolutionary processes including insights into its demographic history.

MATERIALS AND METHODS

Sampling methodology and strategy

Long-spined sea urchins *Diadema setosum* (Leske, 1778) (total N = 718) were sampled from 13 locations throughout the Indo-Malay archipelago between July 2019 and November 2021 (Table 1, Fig. 1). Two distinct samples were collected from each Sabang, Andaman Sea, Mataram, Lombok Strait, and Kendari, Banda Sea; three distinct samples were collected from Pulau Pari, Java Sea and from around Sikka, Flores Sea (Table 1). Multiple samples from the same location were considered replicates. Long-spined sea-urchins were found on reef flats and crests at low tide, at depths between 0.5 m and 3 m. They were picked up using large metal tweezers or large wooden chopsticks and placed in a floating styrofoam box. All individuals sampled were identified as *D. setosum* by an orange ring on the anal cone, five interambulacral white spots, and dotted blue iridophore lines in the naked space of the interambulacral areas (Chow et al. 2014 2016). The gonad tissue of each individual was dissected, and an excised fragment of approximately 50 mm^3 was stored in 95% ethanol for subsequent molecular analysis.

DNA extraction, amplification and sequencing

Genomic DNA was extracted using the Nucleospin® 96 Tissue kit (Macherey-Nagel, Düren, Germany) following the manufacturer's instructions. A 1157-nucleotide long segment beginning at the 5' end of the mitochondrial cytochrome oxidase subunit I (*COI*) gene was amplified in two distinct polymerase-chain reaction (PCR) runs, hereafter PCR1 and PCR2. Specific primer pairs were designed from the alignment of *Diadema* spp. sequences available from GenBank (Benson et al. 2017) including *D. africanum* (GenBank no. KC622343), *D. antillarum* (MN683901, KC622344, KC626158, AY012729), *D. mexicanum* (MN683904, AY012734), *D. savignyi* (MH051887, MH051888, AY012743, AB909951), and *D. setosum* [NC_033522, KU496327, MN690257, AB909922, AY012733, AY012747). The PCR1 run used the DF1/DR1 oligonucleotide primer pair, whose sequences were, respectively, 5'-GCCATGAGAGTRATTATCCG-3' and 5'-CCAAGTATTCCRATAGCGATC-3'. The

PCR2 run used the primer pair DF2/DR2, respectively 5'-ACCTTTTTYGAYCCTGCMGG-3' and 5'-GAAATGYTGAGGRAAGAATG-3'. The PCR1 (resp. PCR2) amplicon was a fragment 717-base-pair (bp) long (resp. 599-bp) starting at homologous nucleotide position no. 5961 (resp. 6519) of NC_033522, the only complete mitochondrial DNA sequence of *D. setosum* published to date (Li et al. 2016). The 3' end of the PCR1 fragment overlapped with the 5' end of PCR2 between homologous nucleotide positions 6519 and 6677 of NC_033522, *i.e.*, over a length of 159 bp.

PCR amplification followed Ivanova and Grainger's (2007) protocol and was run in 96-well polycarbonate plates (ThermoFisher Scientific, Waltham MA, USA). Each well (0.2 ml) contained 20 µl reaction mixture comprising 10 µl 10% trehalose, 4.6 µl ddH_2O,

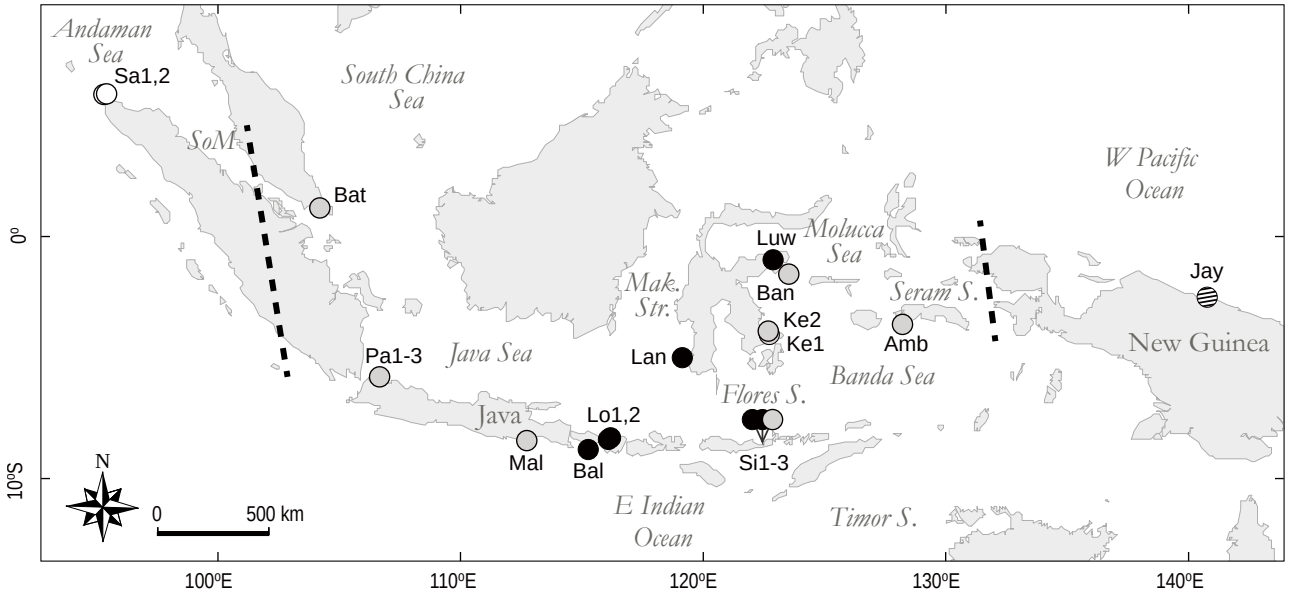

Fig. 1. *Diadema setosum* sampling locations. Colours refer to the results of MDS analysis (see RESULTS). Bold dotted lines indicate inferred geographic barriers. *SoM* Strait of Malacca; *Mak. Str.* Makassar Strait; other abbreviations as in table 1.

Table 1. *Diadema setosum* sampling details and GenBank accession numbers by sample. *N* sample size

Sample	N	Locality (see Fig. 1)	Coordinates	Sampling date	GenBank accession nos.
Sa1	11	Sabang, Andaman Sea	5.883N 95.254E	Nov. 2021	OP310072-OP310082
Sa2	22	Sabang, Andaman Sea	5.891N 95.310E	Jan. 2020	OP310083-OP310104
Bat	75	Batam, Malacca Strait	1.158N 104.145E	May 2021	OP310105-OP310179
Pa1	36	Pulau Pari, Java Sea	5.865S 106.614E	Feb. 2020	OP310180-OP310215
Pa2	31	Pulau Pari, Java Sea	5.865S 106.614E	Mar. 2020	OP310216-OP310246
Pa3	42	Pulau Pari, Java Sea	5.865S 106.614E	Aug. 2021	OP310247-OP310288
Mal	16	Malang, southern Java, E Indian O.	8.435S 112.682E	Aug. 2021	OP310289-OP310304
Bal	47	Southern Bali, E Indian O.	8.778S 115.228E	July 2021	OP310305-OP310351
Lo1	38	Mataram, Lombok Strait	8.404S 116.079E	Sep. 2019	OP310352-OP310389
Lo2	21	Mataram, Lombok Strait	8.360S 116.121E	June 2021	OP310390-OP310410
Lan	80	Langkai, Makassar Strait	5.030S 119.094E	Aug. 2021	OP310411-OP310490
Si1	31	Sikka, Flores Sea	8.607S 122.418E	Sep. 2021	OP310491-OP310521
Si2	30	Sikka, Flores Sea	8.526S 122.512E	Sep. 2021	OP310522-OP310551
Si3	23	Sikka, Flores Sea	8.493S 122.402E	Sep. 2021	OP310552-OP310574
Ke1	57	Kendari, Banda Sea	4.036S 122.673E	July 2019	OP310575-OP310631
Ke2	61	Kendari, Banda Sea	3.894S 122.615E	Aug. 2019	OP310632-OP310692
Luw	66	Luwuk, Molucca Sea	0.983S 122.791E	Aug. 2021	OP310693-OP310758
Ban	8	Banggai, Banda Sea	1.588S 123.472E	Nov. 2021	OP310759-OP310766
Amb	21	Ambon, Banda Sea	3.639S 128.199E	Sep. 2021	OP310767-OP310787
Jay	2	Jayapura, N New Guinea	2.533S 140.737E	Aug. 2021	OP310788-OP310789

2 µl 10X buffer, 1 µl 50 mM $MgCl_2$, 0.2 µl 10 µM each primer, 0.1 µl 10 mM dNTPs, 0.1 µl (5U/µl) Platinum *Taq* DNA Polymerase (Invitrogen, Carlsbad CA, U.S.A.), and 2 µl DNA template (1–15 ng). Conditions for PCR amplification included initial denaturation at 94°C for 2 min, followed by 35 amplification cycles (94°C for 30 s, 52°C for 40 s, and 72°C for 1 min), with final extension at 72°C for 10 min, followed by indefinite hold at 4°C. PCR products were checked on pre-cast SYBR-stained 2% agarose gels from the E-gel™-96 system of Invitrogen. PCR1 and PCR2 amplicons were sequenced according to the Sanger protocol by Microsynth SAS (Vaulx-en-Velin, France) using DR1 and DF2 as respective sequencing primers.

Characterization of nucleotide sequences

The sequence chromatograms were verified under Chromas v. 2.6.5 (Technelysium, Brisbane, Australia). The nucleotide sequences, once assembled (N = 718), were aligned using the ClustalW algorithm implemented in MegaX v. 10.0.4 (Kumar et al. 2018). All sequences were deposited into GenBank and allocated accession numbers OP310072 to OP310789.

The sequence dataset used in the phylogeny of the genus *Diadema* consisted of 739 sequences aligned over 586 bp, including all *Diadema* spp. *COI* gene sequences of Lessios et al. (2001), including *D. antillarum* (GenBank AY012728-29), *D. africanum* (AY012730-31), *D. setosum*-b (AY012132-33), *D. setosum*-a (AY012746-47 from the eastern Indian ocean and the Ryukyu islands), *D. mexicanum* (AY012734-35), *D. palmeri* (AY012736-37), *D. paucispinum* (AY012738-41), *D. savignyi* (AY012742-43), and *D. clarki* (AY012744-45), together with *D. setosum*-a from the Paracels Islands (NC_033522) and all *D. setosum* sequences of the present survey. The Diadematidae *Echinothrix calamaris* mitochondrial sequence (NC_050274) was added as an outgroup. A phylogenetic tree that was produced under MegaX using the maximum-likelihood (ML) algorithm based on Tamura's (1992) 3-parameter nucleotide substitution model with discrete gamma distribution (T92+G; G = 0.16), selected as the best model according to the Bayesian information criterion (under MegaX). Node robustness was tested by bootstrapping nucleotide sites (1000 random pseudo-replicates) in the alignment of sequences (MegaX). Mean genetic distances between and within groups were calculated with MegaX.

Phylogeographic analysis

For this analysis we used the complete *D. setosum COI* sequence dataset from the Indo-Malay archipelago, comprising 718 nucleotides sequences 1157 bp long (Vimono et al. 2022). Haplotype- and nucleotide-diversity indices were estimated for each sample with Arlequin 3.5 (Excoffier and Lischer 2010).

The pairwise genetic-differentiation index Φ_{ST} based on the T92+G model (see above) was estimated under Arlequin, together with significance levels (from 1000 random permutations) for all pairs of samples. The null hypothesis of no correlation between pairwise genetic distance $[\Phi_{ST}/(1-\Phi_{ST})]$ and geographic distance was tested with Mantel's test (10^4 random permutations under Genetix v. 4.05; Belkhir et al. 2004); we used as geographic distance the shortest sea route between two sampling locations, measured on GoogleEarth images using the *path* tool (https://earth.google.com/). A multidimensional scaling (MDS; Gower 1966) plot was derived from the matrix of pairwise Φ_{ST} to place samples in a two-dimensional genetic space. This was done using the *cmdscale* function in R (R Core Team 2020). Genetically consistent groups of samples were delineated by hierarchical clustering in R, using the *hclust* function with the *complete* method of Defays (1977).

Hierarchical analysis of molecular variance (AMOVA; Excoffier et al. 1992) was run to analyse population subdivision by partitioning genetic variance into its inter-individual (*i.e.*, within a local sub-population), between sub-populations, and inter-population covariance components. Covariance components were used to estimate the respective fixation index components Φ_{ST}, Φ_{SC} and Φ_{CT}, the significance of which was tested by random permutation. Three distinct phylogeographic hypotheses were tested by AMOVA: (i) a null hypothesis analogous to Wright's (1951) island model, where genetic differentiation was spatially uniform, *i.e.*, with no regional structure; (ii) Spalding et al.'s (2007: box 1) marine ecoregional structure where eight ecoregions were considered (Western Sumatra, Southern Java, Lesser Sundas, Malacca Strait, Sunda shelf, Sulawesi Sea and Makassar Strait, Banda Sea, and Papua); and (iii) regional structure according to the phylogeographic breaks proposed by Carpenter et al. (2011: fig. 4), delimiting five sub-regions (Andaman Sea, Sunda Shelf and Java Sea, Southern Java and Lesser Sunda islands, eastern Indonesia, Papua). The AMOVA parameter estimates in each case were compared with those derived from the phylogeographic structure inferred from MDS analysis.

Relationships between mitochondrial haplotypes were described by a minimum spanning network constructed with Popart 1.7 (Leigh and Bryant 2015). Branch length between haplotypes was adjusted under Illustrator (Smith 2010) to make it proportional to the number of nucleotide substitutions. The null hypothesis

of mutation-drift equilibrium was tested by comparing Tajima's (1989) *D* and Fu's (1997) F_S estimates to coalescent simulations under the infinite-site model (Hudson 1990) under Arlequin. Mismatch distributions computed from nucleotide sequences were compared to those expected under Rogers' (1995) model of sudden population expansion and Excoffier's (2004) model of spatial expansion under Arlequin.

Miscellaneous

Rosner's test for detecting outliers was run using the *rosnerTest* function of the EnvStats package v. 2.7.0 under R (Millard 2013; Millard and Kowarik 2022).

RESULTS

Phylogenetic relationships of *D. setosum* mitochondrial sequences

All *D. setosum* sequences from the Indo-Malay archipelago clustered within a single, highly supported clade on the ML tree (Fig. S1) that also included the three homologous sequences of *D. setosum*-a from the eastern Indian Ocean, the South China Sea and the Ryukyu archipelago. The average genetic distance between sequences of the *D. setosum*-a clade and the *D. setosum*-b clade was 0.093 ± 0.022; it reached 0.442 ± 0.122 between *D. setosum*-a and a third clade that included all homologous sequences available for all other *Diadema* species. The *D. setosum*-a clade displayed an average within-group genetic distance of 0.007 ± 0.002.

From the 718 mitochondrial *COI* gene sequences analysed, we observed a total of 259 haplotypes (details in Vimono et al. 2022). Sixty two haplotypes (23.9% of total) were shared between two or more sampling locations. The haplotype with the highest frequency in the total sample, H22, was shared by 83 individuals from 15 sampling locations. A large proportion of haplotypes were singletons, *i.e.*, haplotypes observed only once in the total sample (N = 191; 73.7%), or private, *i.e.*, shared by two or more individuals but exclusive to a particular sampling location (N = 6; 2.3%). The frequencies of the two private haplotypes sampled from the Andaman Sea reached 30% and 6%.

Phylogeographic structure

Sample-pairwise Φ_{ST} estimates ranged from -0.050 to 0.258; consistently high values among groups of samples (Table 2) indicated phylogeographic structure within the Indo-Malay archipelago. Three main genetically-differentiated groups of samples were observed: one group (G1) that included the two samples from the Andaman Sea, a second group (G2) that comprised the single sample from the western Pacific ocean, and a large group (G3) that comprised all samples from the Indonesian seas (Fig. 2A). The latter was subdivided into two subgroups: one subgroup (G3a) was geographically centred on Makassar strait and reached southern Bali towards the South, and the Molucca Sea and the Flores Sea towards the East, and the other subgroup (G3b) comprised all samples from the Strait of Malacca to the Java Sea and the southern coast of Java together with all samples of the Banda Sea and one sample from the Flores Sea (Fig. 1). Genetic grouping based on hierarchical clustering (Fig. 2A) was thus geographically consistent, in that samples that clustered within the same group were spatially contiguous (Fig. 1). Based on the total sample set, a correlation was apparent between genetic distance and geographic distance ($P < 0.001$) (Fig. 2B). However, the distribution of data points on the plot was not linear, as the pairwise comparisons involving the samples from the Andaman Sea and the western Pacific Ocean clustered well above the trend line (Fig. 2B). When these two geographically external groups of samples were omitted from the analysis, no significant correlation was observed ($P > 0.083$).

The partition of the total *D. setosum* sample into geographic groups defined according to either Spalding et al.'s (2007) or Carpenter et al.'s (2011) phylogeographic hypotheses explained respectively 2.7% and 3.3% of the total genetic variance, while the grouping derived from the present MDS analysis explained 15.8% (Φ_{CT} = 0.157; $P < 0.001$) (Table 3). The Φ_{ST} value between sub-groups G3a and G3b was Φ_{ST} = 0.025 ($P < 0.001$).

Departure from mutation-drift equilibrium

The nucleotide sequence dataset comprised a total of 259 haplotypes (Table S1). Variable nucleotide sites (n = 219; 18.9%) included 23 sites at first position, six at second position, and 190 at third position of a codon. The two Andaman-Sea samples displayed haplotype diversity values ($H \leq 0.891$) clearly lower than all the other samples ($H \geq 0.926$) (Rosner's test for detecting outliers; $P < 0.05$).

The network of haplotypes (Fig. 3) showed a series of closely related star-shaped haplogroups, each composed of a central haplotype at high frequency, surrounded by a high proportion of singletons. Tajima's *D* and Fu's F_S estimates were negative for all samples except those from the Andaman Sea (Table S2). After grouping samples according to the patterns of

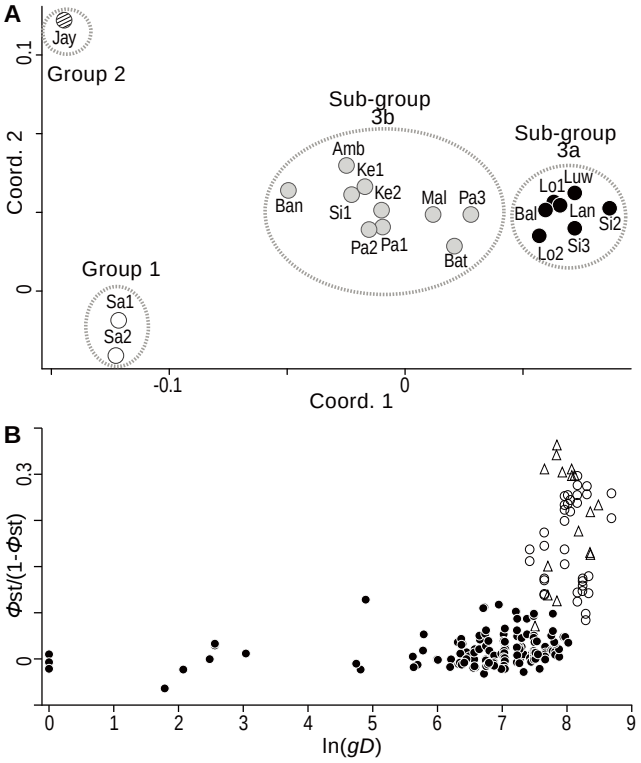

Fig. 2. Phylogeographic structure of *Diadema setosum* in the Indo-Malay archipelago. A: Multidimensional scaling plot based on the matrix of pairwise Φ_{ST} (T92+G model; Arlequin). Ellipses delineate groups of samples designated by hierarchical clustering. Abbreviations for samples as in table 1. B: Pairwise estimates of genetic differentiation plotted against the logarithm of geographic distance [ln(*gD*), *gD* in km]. Black circles (●) intra-group comparisons; open circles (○) (resp. triangles): inter-group comparisons involving Group 1 (resp. Group 2).

Table 2. *Diadema setosum* in the Indo-Malay archipelago. Population-pairwise Φ_{ST} estimates (T92+G model, Arlequin; below diagonal) and *p*-values (from 1023 random permutations of individuals in Arlequin; above diagonal)

Sample	Sample																			
	Sa1	Sa2	Bat	Pa1	Pa2	Pa3	Mal	Bal	Lo1	Lo2	Lan	Si1	Si2	Si3	Ke1	Ke2	Luw	Ban	Amb	Jay
Sa1	0	-	***	**	**	***	**	***	***	***	***	**	***	***	***	**	***	-	***	-
Sa2	-0.050	0	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	-	***	-
Bat	0.138	0.152	0	-	-	-	-	-	-	-	**	-	*	-	*	-	**	-	-	*
Pa1	0.095	0.114	-0.011	0	-	-	-	-	-	-	**	-	*	*	-	-	**	-	-	-
Pa2	0.096	0.116	0.006	-0.015	0	-	-	*	*	-	***	-	**	**	*	-	***	-	-	-
Pa3	0.156	0.171	-0.005	-0.005	0.008	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Mal	0.134	0.152	-0.016	-0.016	-0.009	-0.017	0	-	-	-	-	-	-	-	-	-	-	-	-	-
Bal	0.194	0.206	0.009	0.019	0.045	-0.005	-0.009	0	-	-	-	-	-	-	*	-	-	-	*	*
Lo1	0.201	0.210	0.010	0.019	0.048	-0.003	0.000	-0.007	0	-	-	-	-	-	-	-	-	*	*	*
Lo2	0.184	0.196	0.001	0.002	0.028	-0.016	-0.001	-0.017	-0.017	0	-	-	-	-	-	-	-	-	-	*
Lan	0.202	0.211	0.028	0.036	0.062	0.006	0.014	-0.006	0.000	-0.013	0	*	-	-	*	**	-	*	*	*
Si1	0.086	0.098	0.012	0.000	0.010	0.006	-0.010	0.021	0.016	0.011	0.027	0	-	-	-	-	+	-	-	-
Si2	0.221	0.230	0.025	0.037	0.069	0.003	0.010	-0.011	-0.014	-0.012	-0.008	0.025	0	-	*	*	-	*	*	-
Si3	0.203	0.211	0.027	0.034	0.066	0.013	0.029	0.000	0.000	-0.014	-0.006	0.022	0.000	0	-	*	-	*	*	**
Ke1	0.100	0.113	0.016	0.007	0.020	0.008	0.008	0.019	0.013	0.003	0.016	-0.009	0.023	0.009	0	-	*	-	-	-
Ke2	0.107	0.116	0.004	0.000	0.011	0.004	-0.021	0.017	0.016	0.015	0.031	-0.009	0.024	0.032	0.009	0	***	-	-	-
Luw	0.212	0.220	0.035	0.044	0.069	0.013	0.027	0.002	0.000	-0.006	-0.006	0.029	-0.006	-0.008	0.014	0.039	0	*	*	*
Ban	0.060	0.069	0.035	0.015	0.014	0.037	-0.004	0.062	0.072	0.072	0.082	-0.024	0.078	0.077	0.004	-0.013	0.089	0	-	-
Amb	0.097	0.119	0.026	0.007	0.008	0.012	0.008	0.039	0.033	0.020	0.039	-0.002	0.045	0.040	0.000	0.011	0.037	0.005	0	-
Jay	0.187	0.213	0.200	0.145	0.147	0.193	0.172	0.230	0.230	0.236	0.233	0.086	0.258	0.249	0.094	0.131	0.236	0.006	0.051	0

Abbreviations for samples as in table 1. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

geographic differentiation observed from pairwise Φ_{ST} values (Fig. 2A), the estimated *D* value were -0.75 ($P = 0.257$) for Group 1, -2.37 ($P < 0.001$) for Sub-group 3a and -2.43 ($P < 0.001$) for Sub-group 3b; the estimated *F*s values were, respectively -0.26 ($P = 0.472$), -24.73 ($P < 0.01$) and -25.11 ($P < 0.001$).

Both Rogers' sudden demographic expansion model and Excoffier-Lischer's spatial expansion models were rejected for the Andaman Sea population but not for the populations of the Indonesian seas (Table S2).

DISCUSSION

Phylogeographic patterns

Phylogeographic structure was evident among *D. setosum* populations across the Indo-Malay archipelago, where an Andaman Sea population and a New Guinean population were genetically distinct from a third population occupying the central Indonesian seas. Although genetic distance overall increased with geographic distance, the patterns observed at the scale of the entire Indo-Malay archipelago were not consistent with those expected under isolation-by-distance (Rousset 1997), allowing us to reject isolation-by-distance as the main driver of genetic differentiation. The estimated level of genetic differentiation among the three main population groups was markedly higher than the average value estimated by Lessios et al. (2001) between populations from Western Australia and the western Pacific Ocean ($\Phi_{ST} = 0.073 \pm 0.060$), *i.e.*, across the Indo-Malay archipelago but based on a sampling design partly distinct from ours and lacking samples from the Andaman Sea region. The Φ_{ST} values presented by Lessios et al. (2001) were not significant, which was interpreted by the latter and by Carpenter et al. (2011) as a "lack of phylogeographic break" across the Indo-Malay archipelago. In the present study, two phylogeographic breaks were detected, which had thus not been detected earlier because of a partly different sampling design and possibly because of weaker statistical power due to smaller sample sizes.

More-subtle but still highly significant haplotype-frequency differences distinguished two sub-populations within the Indo-Malay archipelago. One sub-population occurred across the archipelago West to East, from the Strait of Malacca to the Banda Sea, while the range of the other population extended from the Makassar Strait to Lombok Strait and to the Flores Sea. The latter seemed to deeply penetrate into the range of the former, precisely where major currents that constitute the Indonesian Throughflow carry Pacific waters from North to South across the Indo-Malay archipelago before reaching the eastern Indian Ocean (Gordon and Fine 1996). The geographic distribution of the two sub-populations may be either a stable pattern or a transitory one reflecting the ongoing intrusion of a genetically differentiated population.

Possible determinants of phylogeographic structure

In the late Pleistocene, the lowering and rising of sea levels greatly affected land mass configurations in the western and eastern parts of the Indo-Malay

Table 3. *Diadema setosum*. Summary of AMOVA results relative to nested spatial scales. Phylogeographic model refers to regional grouping of populations based on either marine ecoregions (Spalding: box 1 of Spalding et al. 2007), on phylogeographic breaks (Carpenter: figure 4 of Carpenter et al. 2011), or on genetic structure assessed from groupings derived from MDS analysis (MDS groups; Fig. 2A)

Phylogeographic model	Source of variation											
	Among groups				Among populations within groups				Within-population			
	d.f.	*var*	*%var*	Φ_{CT} [a]	*d.f.*	*var*	*%var*	Φ_{SC} [b]	*d.f.*	*var*	*%var*	Φ_{ST} [c]
No structure	–	–	–	–	19	0.080	3.06	–	698	2.554	96.94	0.031***
Spalding	7	0.074	2.79	0.028**	12	0.018	0.67	0.007	698	2.554	96.54	0.035***
Carpenter	4	0.086	3.25	0.032***	15	0.017	0.65	0.007*	698	2.554	96.10	0.039***
MDS (3 groups)	2	0.484	15.75	0.157***	17	0.032	1.05	0.012***	698	2.554	83.20	0.168***

d.f., degrees of freedom; *var*, variance components; *%var*, percentage of variation; Φ_{CT}, Φ_{SC}, and Φ_{ST}, resp. among-group, among-population and within-population components. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. [a]Probability of obtaining equal or lower Φ_{CT} value determined by 1,000 randomizations by permuting populations among groups. [b]Probability of obtaining equal or lower Φ_{SC} value determined by 1,000 randomizations by permuting haplotypes among populations within groups. [c]Probability of obtaining equal or lower Φ_{ST} value determined by 1,000 randomizations by permuting haplotypes among populations among groups.

archipelago. The Sunda shelf (which extends along the western half of the archipelago) and the Sahul shelf (at its eastern margin) were repeatedly above sea level, forming the Indo-Pacific barrier (Voris 2000; Gaither et al. 2011; Crandall et al. 2012; DeBoer et al. 2014; Bertrand et al. 2017; Lim et al. 2021). Large river systems draining the Sunda Shelf discharged massive amounts of freshwater into the oceanic waters of the

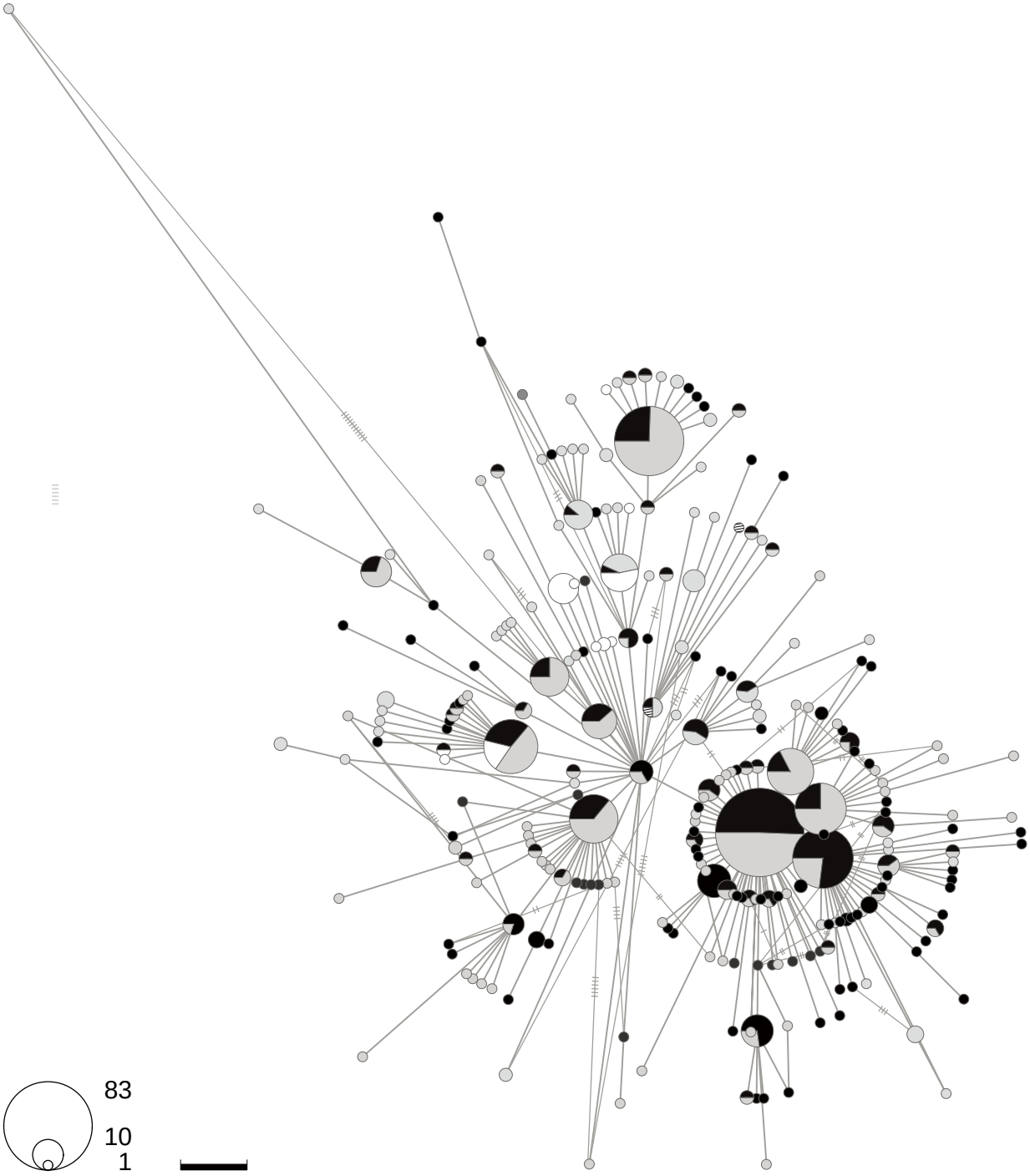

Fig. 3. *Diadema setosum*. Minimum-spanning network of *COI* gene haplotypes. Area of circle proportional to haplotype frequency in total sample; length of branch proportional to number of mutations (scale bar: one mutational step) except alternative links between haplotypes, where number of mutational steps is represented by a series of perpendicular dashes. A few such links between distant haplotypes were omitted for better readability. White: haplotype sampled from Group-1 *D*. setosum; hachured white: Group-2; black: Sub-group-3a; grey: Sub-group-3b.

Makassar Strait and similarly, from the Sahul Shelf into the Seram Sea and the Timor Sea (Voris 2000). This layer of turbid freshwater likely acted as an additional ecological barrier for oceanic larvae drifting at the surface. We interpret the phylogeographic structure reported here between the Andaman Sea and the rest of the Indo-Malay archipelago as a consequence of late Pleistocene vicariance on either side of the Indo-Pacific barrier. Even though the Indo-Pacific barrier was only intermittent and short-lasting, repeated geographic isolation of refuge populations may have accelerated their genetic isolation (He et al. 2019). The other phylogeographic discontinuity uncovered in the present study occurred between the eastern Indonesian seas and northern New Guinea. This pattern has been previously observed in giant clams (DeBoer et al. 2014). A possible geographic barrier could be the steep Pacific shores of northwestern New Guinea, where favourable shallow-water coastal habitats are mostly absent. The larva would have to cross a distance of hundreds of kilometres, which may exceed the range that a planktonic *D. setosum* larva is actually able to reach (see Shanks 2009). However, in the long term it is likely that even rare larval transport would be able to connect populations from either side of the barrier, suggesting that other mechanisms hampering gene flow might be involved.

Hypotheses for the persistence of phylogeographic patterns

In broadcast-spawning marine species with large effective population sizes, genetic differentiation between populations on either side of a geographic barrier is conceivably a long process; the fixation of genetic differences and traits that limit outcrossing may require many generations to occur (Norris and Hull 2012). The duration of the period of geographic isolation caused by the Indo-Pacific barrier in the late Pleistocene only lasted a few tens of thousands of years before the rise in sea level reconnected once-temporarily separated inshore habitats (Voris 2000). The pelagic larval duration of *Diadema setosum* estimated from laboratory experiments (40–45 d; Dautov and Dautova 2016) is potentially long enough to ensure gene flow at a broad geographic scale (Selkoe and Toonen 2011). Indeed, given their duration in the plankton, *D. setosum* larvae are potentially able to cross tens to possibly a few hundreds of kilometres, although propagule duration may only be loosely correlated with dispersal distance (Shanks 2009; Selkoe and Toonen 2011). Throughout the Indo-Malay archipelago, the monsoons seasonally reverse wind direction, ensuring surface-water connectivity in both the East-to-West and West-to-East directions, including through the Sunda strait between the eastern Indian ocean and the Java Sea, through the Karimata Strait between the Java Sea and the South China Sea, and through Makassar Strait on the opposite extremity of the Java Sea (He et al. 2015; Wang et al. 2019; Genda et al. 2022).

In this highly connective oceanographic context, it would be reasonable to expect genetic homogeneity among *D. setosum* populations of the Indo-Malay archipelago, even if that connectivity would be temporarily interrupted by episodes of sea-level lows. Instead, the sharp haplotype-frequency differences between adjacent populations near its western and eastern extremities demonstrate persistent genetic isolation. Therefore, we admit that repeated cycles of geographic isolation on either side of the Indo-Pacific barrier and secondary contact throughout the Pleistocene have led to reproductive isolation, like the kind recently documented in mangrove populations on either side of the Malacca Strait (He et al. 2019). The phylogeographic partition of populations within the Indonesian seas into two sub-groups, albeit more subtle, also deserves consideration. Again, this restriction in gene flow can hardly be explained by present geographic isolation given the highly dynamic current systems that cross the archipelago. Contiguous distributions ranges that overlap by a limited multiple of their dispersal range are defined as parapatric (Bull 1992). If the observed phylogeographic pattern observed within the Indonesian seas is stable, the maintenance of a parapatric-like boundary between populations implies competition, fitness differences, and reproductive isolation or semi-isolation (Bull 1992). Our observations thus raise the hypothesis of some degree of reproductive isolation between ecologically incompatible populations of *D. setosum* within the Indonesian seas. Reproductive isolation may have been initiated in the past when ancestral populations were allopatric, and reinforced in parapatry; it may not need to be complete to hamper gene flow in such a way that populations continue to diverge genetically (He et al. 2019; Westram et al. 2022).

Two possible mechanisms for reproductive isolation have been identified in long-spined sea urchins. The first one could be prezygotic reproductive isolation caused by distinct spawning periods, themselves linked to shifts in lunar spawning cycles (Lessios 1984). In *D. setosum*, different spawning periods have been observed between populations from northeastern Borneo, Rabaul, and Noumea (Pearse 1975), making this hypothesis a plausible one as it would explain the genetic differences observed by us at a similar geographic scale. The second mechanism hypothesized in sea urchins of the genera *Diadema*, *Echinometra*, *Paracentrotus* and

Strongylocentrotus is reproductive incompatibility caused by mismatch between the bindin receptor at the surface of the oocyte and the bindin form carried by the sperm that enables the membrane fusion of the two gametes (Glabe and Vacquier 1978; Schackmann and Shapiro 1981; Vacquier and Moy 1997; Geyer et al. 2020). However, the bindin gene in *D. setosum* shows no evidence of selection (Geyer et al. 2020). Although the latter result does not definitely eliminate the hypothesis of gamete recognition mechanisms between genetically distinct populations within *D. setosum*, it considerably weakens it.

Demographic change and speciation

The star-like radiation of haplogroups sampled within the Indo-Malay archipelago indicates rapid population expansion. The mutation-drift equilibrium hypothesis was not rejected for the Andaman Sea population, but it was for the other populations of the Indo-Malay archipelago. Considering that part of the habitat where *D. setosum* now occurs was unavailable during the last sea-level low (until ca. 10,000 years ago; Crandall et al. 2012), wide geographic colonisation likely happened within a relatively short evolutionary timeframe. The nucleotide-pairwise mismatch patterns in the populations of the Indonesian seas were consistent with both demographic and spatial expansion models. Thus, while the coalescence pattern of the Andaman-Sea population suggested demographic stability over evolutionary time scales, that of the populations of the Indonesian seas indicated recent expansion. This expansion likely coincided with the rapid increase in available *D. setosum* habitat caused by sea-level rise in the late Pleistocene (Voris 2000; Lambeck and Chappell 2001; Crandall et al. 2012). Negative *D* values, large negative F_S values, and the results of pairwise mismatch analysis were compatible with demographic and spatial expansions which occurred until the current geographic distributions were reached.

In conclusion, demographic expansion repeatedly occurred in *D. setosum* throughout the late Pleistocene, alternating each time with demographic contraction. We speculate that during sea-level lows, small populations of long-spined sea urchins survived in isolated refuges (Pellissier et al. 2014), a situation where genetic drift and possibly local adaptation entailed allopatric divergence. Once sea levels rose again and shallow-water habitat again expanded, surviving populations which had evolved towards reproductive isolation from one another re-expanded, with new mutations arising. The distribution of a population stopped when it met the expansion front of another. This peculiar pattern in the geographic distribution of populations implies at least partial reproductive isolation. Meanwhile, ecological compatibility of the different lineages has not yet been achieved. Regardless of which mechanisms are predominantly at play, the apparently rapid acquisition of reproductive isolation in *D. setosum* within the Indo-Malay archipelago would point towards yet unacknowledged incipient speciation.

Acknowledgments: We are grateful to Ucu Y. Arbi, Kadarusman, Elok Faiqoh, Riyana Subandi, La Ode Anshari, Olga Galih Rakha Siwi and Nadya Tirta for participating in the collection of samples. We thank an anonymous reviewer for insightful comments. This work was funded by IRD through recurrent funding to UMR 226 ISEM and UMR 250-S, and through the French-Indonesian SELAMAT International Joint Laboratory coordinated by H.Y. Sugeha (BRIN-P2O) and RH. IBV was supported by a PhD fellowship allocated by the Institut français d'Indonésie (French ministry of European and foreign affairs) and managed by Campus France (French ministry of higher education and research). Sampling in Langkai was authorised under research permit no. B-5935/III/KS.01.04/7/2021 and travel document no. 18624/S.01/PTSP/2021 issued by Lembaga Ilmu Pengetahuan Indonesia (LIPI). Transfer of biological material from Indonesia to France was authorised under material transfer agreement no. B-995/IPK.2/KS.01.04/XII/2020 between LIPI-RCO and Université de Montpellier.

Authors' contributions: PB, RH, LP and IBV designed the study; RH and IBV coordinated the collection of samples; LP and IBV did the laboratory analyses; PB, LP and IBV analysed the data; PB led the writing; all co-authors edited and approved the final version of the paper.

Competing interests: The authors declare no conflict of interests.

Availability of data and materials: Sequences generated in the study were deposited into the GenBank sequence database (https://www.ncbi.nlm.nih.gov/nucleotide). The alignment of sequences and haplotype distribution by sample were deposited into the DataSuds repository (Desconnets et al. 2019) and can be accessed using the following link: https://doi.org/10.23708/ZWQEFN (Vimono et al. 2022).

Ethics approval consent to participate: The present research was approved by the University of Montpellier ad hoc committee that evaluates Vimono's Ph.D. project on an annual basis; and by the scientific committee of IRD (IRD-CSS3), the research institute to

which P. Borsa, R. Hocdé and L. Pouyaud are affiliated.

REFERENCES

Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F. 2004. Genetix 4.05, logiciel sous Windows^{TM} pour la génétique des populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université Montpellier II, Montpellier.

Benson DA, Cavanaugh M, Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. 2017. GenBank. Nucl Acids Res **45:**D37–D42. doi:10.1093/nar/gkw1070.

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. Mol Ecol **26:**2527–2542. doi:10.1111/mec.14044.

Birkeland C. 1989. The influence of echinoderms on coral reef communities, pp. 1–79. *In*: Jangoux M. & Lawrence J.M. (eds.), Echinoderm studies, vol. 3. Balkema, Rotterdam, Netherland.

Bronstein O, Kroh A. 2018. Needle in a haystack – genetic evidence confirms the expansion of the alien echinoid *Diadema setosum* (Echinoidea: Diadematidae) to the Mediterranean coast of Israel. Zootaxa **4497:**593–599. doi:10.11646/zootaxa.4497.4.9.

Bull CM. 1992. Ecology of parapatric distributions. Annu Rev Ecol Syst **22:**19–36. doi:10.1146/ANNUREV.ES.22.110191.000315.

Carpenter KE, Barber PH, Crandall ED, Ablan-Lagman MCA, Ambariyanto G, Mahardika GN, Manjaji-Matsumoto BM, Juinio-Menez MA, Santos MD, Starger CJ, Toha AHA. 2011. Comparative phylogeography of the Coral Triangle and implications for marine management. J Mar Biol **2011:**396982. doi:10.1155/2011/396982.

Carpenter RC, Edmunds PJ. 2006. Local and regional scale recovery of *Diadema* promotes recruitment of scleractinian corals. Ecol Lett **9:**271–280. doi:10.1111/J.1461-0248.2005.00866.X.

Chow S, Kajigaya Y, Kurogi H, Niwa K, Shibuno T, Nanami A, Kiyomoto S. 2014. On the fourth *Diadema* species (*Diadema*-sp) from Japan. PLoS ONE **9:**e102376. doi:10.1371/journal.pone.0102376.

Chow S, Konishi K, Mekuchi M, Tamaki Y, Nohara K, Takagi M, Niwa K, Teramoto W, Manabe H, Kurogi H, Suzuki S, Ando D, Jinbo T, Kiyomoto M, Hirose M, Shimomura M, Kurashima A, Ishikawa T, Kiyomoto S. 2016. DNA barcoding and morphological analyses revealed validity of *Diadema clarki* Ikeda, 1939 (Echinodermata, Echinoidea, Diadematidae). Zookeys **585:**1–16. doi:10.3897/zookeys.585.8161.

Crandall ED, Sbrocco EJ, DeBoer TS, Barber PH, Carpenter KE. 2012. Expansion dating: Calibrating molecular clocks in marine species from expansions onto the Sunda Shelf following the last glacial maximum. Mol Biol Evol **29:**707–719. doi:10.1093/molbev/msr227.

Dautov SS, Dautova TN. 2016. The larvae of *Diadema setosum* (Leske, 1778) (Camarodonta: Diadematidae) from South China Sea. Invert Reprod Dev **60:**290–296. doi:10.1080/07924259.2016.1238411.

DeBoer TS, Abdon Naguit MR, Erdmann MV, Ablan-Lagman MCA, Carpenter KE, Toha AHA, Barber PH. 2014. Concordance between phylogeographic and biogeographic boundaries in the Coral Triangle: conservation implications based on comparative analyses of multiple giant clam species. Bull Mar Sci **90:**277–300. doi:10.5343/BMS.2013.1003.

Defays D. 1977. An efficient algorithm for a complete link method. Comput J **20:**364–366. doi:10.1093/comjnl/20.4.364.

Desconnets J-C, Aventurier P, Banon S, Doucouré C, Coupin T, Hensens H, Decker L. 2019. Entrepôt de données IRD: un service en ligne pour l'ouverture et le partage des données scientifiques au Sud. Poster, Colloque international "Science ouverte au Sud", Univ Cheikh Anta Diop, Dakar, 23–25 Oct. 2019.

Dumont CP, Lau DCC, Astudillo JC, Fong KF, Chak STC, Qiu J-W. 2013. Coral bioerosion by the sea urchin *Diadema setosum* in Hong Kong: Susceptibility of different coral species. J Exp Mar Biol Ecol **441:**71–79. doi:10.1016/J.JEMBE.2013.01.018.

Excoffier L. 2004. Patterns of DNA sequence diversity and genetic structure after a range expansion: lessons from the infinite-island model. Mol Ecol **13:**853–864. doi:10.1046/j.1365-294X.2003.02004.x.

Excoffier L, Lischer HEL. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resour **10:**564–567. doi:10.1111/j.1755-0998.2010.02847.x.

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. doi:10.1093/genetics/131.2.479.

Fu YX. 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics **147:**915–925. doi:10.1093/genetics/147.2.915.

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 Evol Biol **11:**189. doi:10.1186/1471-2148-11-189.

Genda A, Ikehara M, Suzuki A, Arman A, Inoue M. 2022. Sea surface temperature and salinity in Lombok Strait reconstructed from coral Sr/Ca and $\delta^{18}O$, 1962–2012. Front Clim **4:**918273. doi:10.3389/fclim.2022.918273.

Geyer LB, Zigler KS, Tiozzo S, Lessios HA. 2020. Slow evolution under purifying selection in the gamete recognition protein bindin of the sea urchin *Diadema*. Sci Rep **10:**9834. doi:10.1038/s41598-020-66390-2.

Glabe CG, Vacquier VD. 1978. Egg surface glycoprotein receptor for sea urchin sperm bindin. Proc Natl Acad Sci USA **75:**881–885. doi:10.1073/PNAS.75.2.881.

Gordon AL, Fine RA. 1996. Pathways of water between the Pacific and Indian oceans in the Indonesian seas. Nature **379:**146–149. doi:10.1038/379146A0.

Gower JC. 1966. Some distance properties of latent root and vector methods used in multivariate analysis. Biometrika **53:**325–328. doi:10.1093/BIOMET/53.3-4.325.

He Z, Feng M, Wang D, Slawinski D. 2015. Contribution of the Karimata Strait transport to the Indonesian Throughflow as seen from a data assimilation model. Continent Shelf Res **92:**16–22. doi:10.1016/J.CSR.2014.10.007.

He Z, Li X, Yang M, Wang X, Zhong C, Duke NC, Wu C-I, Shi S. 2019. Speciation with gene flow via cycles of isolation and migration: insights from multiple mangrove taxa. Natl Sci Rev **6:**275–288. doi:10.1093/nsr/nwy078.

Hudson RR. 1990. Gene genealogies and the coalescent process. *In*: Futuyma DJ, Antonovics JD (eds.) Oxford Surveys in Evolutionary Biology. Oxford University Press, New York, pp. 1–44.

Ivanova N, Grainger C. 2007. *COI* amplification: *Taq* polymerase choice. CCDB Protocols (Can Ctr DNA Barcoding, Guelph). Available at: www.dnabarcoding.ca.

Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X:

Molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol **35:**1547-1549. doi:10.1093/molbev/msy096.

Lambeck K, Chappell J. 2001. Sea level change through the last glacial cycle. Science **292:**679–686. doi:10.1126/science.1059549.

Leigh JW, Bryant D. 2015. POPART: full-feature software for haplotype network construction. Methods Ecol Evol **6:**1110–1116. doi:10.1111/2041-210X.12410.

Leske NG. 1778. Iacobi Theodori Klein naturalis dispositio echinodermatum accesserunt lucubratiuncula de aculeis echinorum marinorum et spicilegium de belemnitis, edita et descriptionibus novisque inventis et synonomis auctorem aucta. Officina Gleditschiana, Lipsiae, xx+278 p.

Lessios H. 1984. Possible prezygotic reproductive isolation in sea urchins separated by the isthmus of Panama. Evolution **38:**1144–1148. doi:10.1111/j.1558-5646.1984.tb00382.x.

Lessios HA. 2016. The great *Diadema antillarum* die-off: 30 years later. Annu Rev Mar Sci **8:**267–283. doi:10.1146/annurev-marine-122414-033857.

Lessios HA, Kessing BD, Pearse JS. 2001. Population structure and speciation in tropical seas: global phylogeography of the sea urchin *Diadema*. Evolution **55:**955–975. doi:10.1111/j.0014-3820.2001.tb00613.x.

Li C, Wu G, Fu W, Zeng X. 2016. The complete mitochondrial genome of *Diadema setosum* (Diadematacea: Diadematidae). Mitoch DNA Pt B Resour **1:**873–874. doi:10.1080/23802359.2016.1253039.

Lim HC, Habib A, Chen W-J. 2021. Comparative phylogeography and phylogeny of pennah croakers (Teleostei: Sciaenidae) in Southeast Asian waters. Genes **12:**1926. doi:10.3390/genes12121926.

Millard SP. 2013. EnvStats: an R package for environmental statistics. Springer, New York, xvi+291 p.

Millard SP, Kowarik A. 2022. EnvStats-package v. 2.7.0, package for environmental statistics, including US EPA guidance. Comprehensive R Archive Network, Vienna, 1384 p.

Moore AM, Tassakka ACM, Ambo-Rappe R, Yasir I, Smith DJ, Jompa J. 2019. Unexpected discovery of *Diadema clarki* in the Coral Triangle. Mar Biodiv **49:**2381–2399. doi:10.1007/s12526-019-00978-4.

Mumby PJ, Hastings A, Edwards HJ. 2007. Thresholds and the resilience of Caribbean coral reefs. Nature **450:**98–101. doi:10.1038/nature06252.

Muthiga NA, McClanahan TR. 2007. Ecology of *Diadema*. pp. 205-225. *In*: Lawrence JM (ed.) Edible sea urchins: biology and ecology, Elsevier, Amsterdam.

Norris RD, Hull PM. 2012. The temporal dimension of marine speciation. Evol Ecol **26:**393–415. doi:10.1007/s10682-011-9488-4.

Pearse JS. 1975. Lunar reproductive rhythms in sea urchins. A review. J Interdiscipl Cycle Res **6:**47–52.

Pellissier L, Leprieur F, Parravicini V, Cowman PF, Kulbicki M, Litsios G, Olsen SM, Wisz MS, Bellwood DR, Mouillot D. 2014. Quaternary coral reef refugia preserved fish diversity. Science **344:**1016–1019. doi:10.1126/science.1249853.

Qiu J-W, Lau DCC, Cheang C-c, Chow W-k. 2014. Community-level destruction of hard corals by the sea urchin *Diadema setosum*. Mar Pollut Bull **85:**783–788. doi:10.1016/j.marpolbul.2013.12.012.

R Core Team. 2020. R: a language and environment for statistical computing. R Found. Stat. Comput., Vienna.

Rogers A. 1995. Genetic evidence for a Pleistocene population explosion. Evolution **49:**608–615. doi:10.1111/j.1558-5646.1995.tb02297.x.

Rousset F. 1997. Genetic differentiation and estimation of gene flow from *F*-statistics under isolation by distance. Genetics **145:**1219–1228. doi:10.1093/genetics/145.4.1219.

Schackmann RW, Shapiro BM. 1981. A partial sequence of ionic changes associated with the acrosome reaction of *Strongylocentrotus purpuratus*. Dev Biol **81:**145–154. doi:10.1016/0012-1606(81)90357-2.

Selkoe KA, Toonen RJ. 2011. Marine connectivity: a new look at pelagic larval duration and genetic metrics of dispersal. Mar Ecol Prog Ser **436:**291–305. doi:10.3354/MEPS09238.

Shanks AL. 2009. Pelagic larval duration and dispersal distance revisited. Biol Bull **216:**373–385. doi:10.1086/BBLv216n3p373.

Smith J. 2010. Adobe Illustrator digital classroom. Wiley, Indianapolis IN, 349 p.

Spalding MD, Fox HE, Allen GR, Davidson N, Ferdaña ZA, Finlayson M, Halpern BS, Jorge MA, Lombana A, Lourie SA, Martin KD, McManus E, Molnar J, Recchia CA, Robertson J. 2007. Marine ecoregions of the world: a bioregionalization of coastal and shelf areas. Bioscience **57:**573–583. doi:10.1641/B570707.

Tajima F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics **123:**585–595. doi:10.1093/genetics/123.3.585.

Tamura K. 1992. Estimation of the number of nucleotide substitutions when there are strong transition-transversion and G+C-content biases. Mol Biol Evol **9:**678–687. doi:10.1093/OXFORDJOURNALS.MOLBEV.A040752.

Vacquier VD, Moy GW. 1997. The fucose sulfate polymer of egg jelly binds to sperm REJ and is the inducer of the sea urchin sperm acrosome reaction. Dev Biol **192:**125–135. doi:10.1006/DBIO.1997.8729.

Vimono IB, Borsa P, Pouyaud L. 2022. *CO1*-haplotype frequencies of long-spined sea urchin (*Diadema setosum*) from the Indo-Malay archipelago. DataSuds, 3 files. doi:10.23708/ZWQEFN.

Voris HK. 2000. Maps of Pleistocene sea levels in Southeast Asia: shorelines, river systems and time durations. J Biogeogr **27:**1153–1167. doi:10.1046/J.1365-2699.2000.00489.X.

Wang L, Zhou L, Xie L, Zheng Q, Li Q, Li M. 2019. Seasonal and interannual variability of water mass sources of Indonesian Throughflow in the Maluku Sea and the Halmahera Sea. Acta Oceanol Sin **38:**58–71. doi:10.1007/s13131-019-1413-7.

Westram AM, Stankowski S, Surendranadh P, Barton N. 2022. What is reproductive isolation? J Evol Biol **35:**1143–1164. doi:10.1111/jeb.14005.

Wright S. 1951. The genetical structure of populations. Ann Eugenics **15:**323–354. doi:10.1111/j.1469-1809.1949.tb02451.x.

Supplementary materials

Table S1. *Diadema setosum*. Summary statistics describing nucleotide-sequence variability by sample (from DNAsp and Arlequin outputs). *Eta*, total number of mutations; *H*, number of haplotypes; H_d, haplotype diversity; *K*, average number of nucleotide differences; *N*, sample size; *P*, number of parsimony-informative sites; π, nucleotide diversity per site; *S*, number of singletons (*i.e.*, haplotypes sampled only once in a given sample); S_i, number of nucleotide sites with a single substitution; S_s, number of variable sites. (download)

Table S2. *Diadema setosum*. Estimates of Tajima's *D* and Fu's F_S, and sums of squared deviations comparing observed and expected mismatch distributions of haplotype-pairwise differences under Rogers' (1995) pure demographic model and Excoffier's (2004) spatial expansion model, by sample. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$. (download)

Fig. S1. Maximum-likelihood tree of *Diadema* spp. partial nucleotide sequences of the *COI* gene, rooted by homologous sequence in *Echinotrix calamaris*. Placement of *D. setosum* sequences sampled from the Indo-Malay archipelago (present study). (download)