Received: 7 December 2017 | Revised: 25 February 2018 | Accepted: 26 February 2018

DOI: 10.1111/mec.14556

ORIGINAL ARTICLE

WILEY MOLECULAR ECOLOGY

Rapid divergence of mussel populations despite incomplete barriers to dispersal

Diede L. Maas[1] | Stefan Prost[2,3] | Ke Bi[4,5] | Lydia L. Smith[4] | Ellie E. Armstrong[2] | Ludi P. Aji[6] | Abdul Hamid A. Toha[7] | Rosemary G. Gillespie[8] | Leontine E. Becking[1,8,9]

[1]Department of Marine Animal Ecology, Wageningen University & Research, Wageningen, The Netherlands

[2]Department of Biology, Stanford University, Stanford, CA, USA

[3]Department of Integrative Biology, Center for Theoretical Evolutionary Genomics, University of California, Berkeley, CA, USA

[4]Museum of Vertebrate Zoology, University of California, Berkeley, CA, USA

[5]Computational Genomics Resource Laboratory, California Institute for Quantitative Biosciences, University of California, Berkley, CA, USA

[6]Research Centre for Oceanography, Indonesian Institute of Sciences, Jakarta, Indonesia

[7]Department of Fisheries, University of Papua, Manokwari, Indonesia

[8]Department of Environmental Science, Policy, and Management, University of California, Berkeley, CA, USA

[9]Wageningen Marine Research, Den Helder, The Netherlands

Correspondence
Diede L. Maas and Leontine E. Becking, Department of Marine Animal Ecology, Wageningen University & Research, Wageningen, The Netherlands.
Emails: diede.maas@wur.nl and lisa.becking@wur.nl

Funding information
This work used the Vincent J. Coates Genomics Sequencing Laboratory at UC Berkeley, supported by NIH S10 Instrumentation grants S10RR029668 and S10RR027303. The Netherlands Organisation for Scientific Research provided funding to Leontine E. Becking through the grants RUBICON #825.12.007 and VENI #863.14.020. Fieldwork in Indonesia was made possible through additional financial support of National Geographic Society Waitt Grant, De Beukelaar-van der Hucht Stichting, World Wildlife Fund, Netherlands-INNO Fund, the Schure-Beijerinck-Popping Fund of the Royal Dutch Academy of Science (KNAW), Conservation International Ecosystem Based Management program (funded by the David and Lucile Packard Foundation), the Treub-Maatschappij Fund, the Leiden University Fund (LUF)/Slingelands, Singapore Airlines, the A.M. Buitendijk Fund and the J.J. ter Pelkwijk Fund (Naturalis Biodiveristy Center). The funders had no role in the study design, data collection and analysis,

Abstract

Striking genetic structure among marine populations at small spatial scales is becoming evident with extensive molecular studies. Such observations suggest isolation at small scales may play an important role in forming patterns of genetic diversity within species. Isolation-by-distance, isolation-by-environment and historical priority effects are umbrella terms for a suite of processes that underlie genetic structure, but their relative importance at different spatial and temporal scales remains elusive. Here, we use marine lakes in Indonesia to assess genetic structure and assess the relative roles of the processes in shaping genetic differentiation in populations of a bivalve mussel (*Brachidontes* sp.). Marine lakes are landlocked waterbodies of similar age (6,000–10,000 years), but with heterogeneous environments and varying degrees of connection to the sea. Using a population genomic approach (double-digest restriction-site-associated DNA sequencing), we show strong genetic structuring across populations (range F_{ST}: 0.07–0.24) and find limited gene flow through admixture plots. At large spatial scales (>1,400 km), a clear isolation-by-distance pattern was detected. At smaller spatial scales (<200 km), this pattern is maintained, but accompanied by an association of genetic divergence with degree of connection. We hypothesize that (incomplete) dispersal barriers can cause initial isolation, allowing priority effects to give the numerical advantage necessary to initiate strong genetic structure. Priority effects may be strengthened by local adaptation, which the data may corroborate by showing a high correlation between mussel genotypes

decision to publish, or preparation of the manuscript. Permits were obtained from RISTEK (0094/FRP/SM/V/2009; 098/SIP/FRP/SM/V/2011).

and temperature. Our study indicates an often-neglected role of (evolution-mediated) priority effects in shaping population divergence.

KEYWORDS
ddRAD-seq, isolation by distance, isolation by environment, marine biodiversity, population genomics, priority effects

1 | INTRODUCTION

In the marine realm, it has been assumed that barriers to dispersal are few and speciation rates are slow when contrasted to terrestrial ecosystems (Carr et al., 2003; Cowen, Lwiza, Sponaugle, Paris, & Olson, 2000; Palumbi, 1994). Recently, these long-held assumptions have been overturned by studies in marine taxa showing high population genetic structuring at small spatial scales (Barber, Erdmann, & Palumbi, 2006; Carpenter et al., 2011; Gonzalez et al., 2017; Marshall, Monro, Bode, Keough, & Swearer, 2010; Neves et al., 2016). Isolating processes at small spatial scales may therefore play an important role in forming genetic diversity within species. The relative importance of processes shaping genetic structure on different spatial and temporal scales, however, remains elusive (Bowen, Rocha, Toonen, & Karl, 2013; Vellend, 2005). Meanwhile, anthropogenic impacts on ecosystems via climate change and fragmentation of habitats are evident (Haddad et al., 2015; Hoegh-Guldberg & Bruno, 2010). Therefore, insights into drivers of biodiversity and how populations may respond and disperse to colonize new habitats will be valuable knowledge for adequate management of, for example, marine-protected areas.

Isolation acts to decrease rates of immigration and thus decreases gene flow, which subsequently can enhance speciation by allowing populations to locally adapt, diverge and ultimately form new species (Hendry, Bolnick, Berner, & Peichel, 2009). Within the context of population genetics, generally two modes of isolation are considered: (i) isolation-by-distance due to geographic distance or physical barriers (Wright, 1943), and (ii) isolation-by-environment due to environmental dissimilarity (Wang & Bradburd, 2014; Wang & Summers, 2010). Recently, a third mode has come into more consideration in the context of population genetics: (iii) historical priority effects, due to the arrival order of species or genotypes (De Meester, Vanoverbeke, Kilsdonk, & Urban, 2016; Fukami, 2015; Orsini, Vanoverbeke, Swillen, Mergeay, & de Meester, 2013). The three modes are not mutually exclusive and may have different relative importance at different spatial and temporal scales in the formation of population genetic structure (Orsini et al., 2013). They provide insights into the underlying types of processes that shape diversity, such as larval mobility, selection against migrants and competition.

The three modes differ in the predictions they make about the relationship between genetic distance and geographic and environmental distance. Isolation-by-distance patterns arise when there is a spatial reduction in movement of individuals through the process of dispersal limitation, which is contrasted to a panmictic population with freely dispersing propagules (Wright, 1943). It predicts that genetic differentiation increases with geographic distance between populations. Additionally, in the oceans propagules will become more diluted with increasing geographic distances, which lowers the chance of gene flow or colonization (Johannesson, 1988). With restricted gene flow, populations can locally accumulate genetic differences via stochastic genetic drift. Secondly, isolation-by-environment predicts gene flow between environmentally divergent habitats is limited due to biased dispersal, selection against migrants via reduced fitness of immigrants in the new environment or reduced hybrid fitness (Rundle & Nosil, 2005; Schluter, 2009; Wang & Bradburd, 2014). Environmentally similar habitats are in contrast expected to have ongoing gene flow, which may enhance local adaptation through incorporation of preadapted alleles (Wang & Bradburd, 2014). Finally, priority effects emphasize the importance of first colonizers in shaping subsequent population genetic structure, also termed historical contingency (Fukami, 2015; Orsini et al., 2013). Priority effects may have an ecological and an evolutionary component. The ecological advantage of being the first colonizer through competitor or predator release may result in a numerical dominance of certain genotypes that can be difficult to overcome by subsequent immigrating genotypes (De Meester et al., 2016). The advantage may be enhanced by an evolutionary component via rapid local adaptation of the first genotypes. This gives a head start to early colonizers towards locally adapting and therefore becoming stronger competitors to late arrivers, effectively fixing the initial stochastic numerical advantage in the long term (de Meester, Gómez, Okamura, & Schwenk, 2002; De Meester et al., 2016). Priority effects predict that even on small spatial scales, and with incomplete barriers to dispersal, gene flow can still be limited between populations due to density-dependent and evolution-mediated dominance of early genotypes, irrespective of geographic and environmental distance. Priority effects are expected to be stronger in more isolated habitat patches, as the time lag between initial colonization and subsequent immigration will be sufficient to attain a numerical advantage for early colonists (Fukami, 2015).

A key issue remains that modes of isolation can be difficult to distinguish, as they are frequently confounded (Lee & Mitchell-Olds, 2011; Legendre, 1993). Environmental data are often spatially structured at multiple scales and may change over time, which clouds the ultimate cause of divergence. Island systems have classically been used as model systems as they alleviate confounding factors and provide a clearly defined spatial, temporal and environmental context (MacArthur & Wilson, 1967; Warren et al., 2015). Therefore, island-

like systems are ideal to test relative importance of modes of isolation underlying genetic structure. Here, we focus on marine lakes which, like terrestrial islands, are "natural laboratories" as they harbour discrete marine populations, are replicated over space and time, and represent different environmental regimes (Becking et al., 2011). Marine lakes are land-locked bodies of sea water with varying degrees of connection to the surrounding ocean via subterranean fissures and pores (Dawson, Martin, Bell, & Patris, 2009; Hamner & Hamner, 1998; Holthuis, 1973; Tomascik & Mah, 1994). The lakes originated after the Last Glacial Maximum when natural depressions in karstic landscapes filled with rising seawater level (Sathiamurthy & Voris, 2006). Hence, they are relatively young systems estimated to be about 6,000–12,000 years old (Dawson et al., 2009). Large numbers of marine lakes are found in the Coral Triangle in Indonesia (Becking, de Leeuw, & Vogler, 2015; Becking et al., 2011). This region is characterized by extremely high marine biodiversity (Hoeksema, 2007; Mangubhai et al., 2012). All marine lakes maintain a connection to the surrounding sea, which ensures a continued vector for propagules to move in and out of the lakes with tidal fluctuations.

Most marine lakes in Indonesia harbour the diploid bivalve mussel *Brachidontes* spp. (Swainson 1840) (Mollusca; Bivalvia; Mytilidae). *Brachidontes* spp. shells have been found in the deepest layers of sediment cores from marine lakes and have likely been some of the first colonists of the marine lakes. Species of the genus *Brachidontes* can form large mussel beds by attaching themselves to substrates in and below intertidal areas (Terranova, Lo Brutto, Arculeo, & Mitton, 2007). They are broadcast spawners and have a dispersive planktonic larval stage for a duration of up to four weeks (Monteiro-Ribas, Rocha-Miranda, Romano, & Quintanilha, 2006). Previous work on these mussels has shown that *Brachidontes* spp. from marine lakes and mangroves in the Indo-Pacific fall into six genetically distinct lineages that likely represent six separate species (Becking et al., 2016; Goto, Tamate, & Hanzawa, 2011). Recently, studies using a mitochondrial DNA marker (COI) found that for 22 marine lakes studied each lake contained only one lineage, which may be diverging in situ (Becking et al., 2016; C. L. De Leeuw, K. T. C. A. Peijnenburg, R. G. Gillespie, D. L. Maas, N. Hanzawa, L. P. Aji, Abdunnur, A. H. A. Toha, L. E. B. Becking, unpublished data, January 03, 2018). We currently expand on these single marker studies, by using thousands of markers generated from high-throughput double-digest restriction-site-associated DNA (ddRAD) sequencing to assess genome-wide signals.

By comparing marine lakes in Indonesia on multiple spatial scales, with similar ages and sizes, but with varying degrees of connection to the sea and differing environmental regimes, we tested the relative contribution of isolation-by-distance, isolation-by-environment and historical priority effects in the formation of population genetic structure of a bivalve mussel (*Brachidontes* sp.). On large spatial scales (>1,400 km), we expect to find a pattern of isolation-by-distance, as gene flow is unlikely to be maintained across this distance due to currents and landmasses. On scales where propagule dispersal is still expected (<200 km), we expect to find low levels of genetic differentiation, reflecting panmictic populations. If isolation-by-environment plays a role, we expect environmentally similar marine lakes will also be genetically similar. Finally, if priority effects are at play, we would expect to see high levels of differentiation even on small spatial scales (<40 km), where lakes with higher connection to the sea would be more similar to each other as they would have less time lag between colonists and later immigrating genotypes.

2 | MATERIALS AND METHODS

2.1 | Sample locations

The study encompasses two regions, Berau in East Kalimantan (Figure 1a) and Raja Ampat (including Gam and Misool islands) in West Papua (Figure 1b,c), and three spatial scales (>1,400 km: including Berau and Raja Ampat areas, <200 km: including only Raja Ampat areas, and <40 km: including only the Misool area). The islands of Berau are part of the Berau marine-protected area (Becking, Cleary, & De Voogd, 2013). Berau has a tropical rainforest climate with no clear difference in rainy and dry seasons, except for an increase in winds between December and March (Becking et al., 2013; Tomascik & Mah, 1994). We included a lake from the Berau region as a distant population to compare its patterns to those found in West Papua. Raja Ampat is part of the Coral Triangle, which is famous for its high marine biodiversity (Hoeksema, 2007). Raja Ampat lies on the equator and has a tropical climate with yearly precipitation ranging from 2,500 to 4,500 mm, with monsoons being the main characteristics of seasonal change. Northwestern monsoons from November to March and southeastern monsoons from May to October are characterized by persistent winds. Within the Raja Ampat Regency, there is a multitude of islands of karstic rock, which causes the coastline to be complex (Becking et al., 2011). Currents among complex coastlines create local turbulence, which is expected to benefit larval connectivity among reefs (Starger, Van Nydeck Erdmann, Toha, Baker, & Barber, 2015). Reefs of Raja Ampat are mostly shallow fringing, lagoon and atoll reefs (Mangubhai et al., 2012). Average sea surface temperature is 29.0°. At least 45 marine lakes have been identified in Raja Ampat, with the highest density found in the Misool area (Becking et al., 2011; Figure 1c). Most marine lakes in this area are in a pristine state with no apparent anthropogenic influence.

2.2 | Sample collection and environmental profiling

We collected specimens of the bivalve mussel *Brachidontes* sp. from seven marine lakes. From sediment cores, we see that these mussels were present at the early stages of formation of the lake (*Klei, 2016*). All *Brachidontes* sp. samples were confirmed to be of the same genetic lineage (Lineage A, as defined by Goto et al. (2011) and Becking et al. (2016)). One lake was sampled in East Kalimantan (Figure 1a) and six in Raja Ampat (Figure 1b,c). Codes, locations, number of samples per lake and lake characteristics are recorded in Table 1. In total, 125 *Brachidontes* sp. samples were collected.

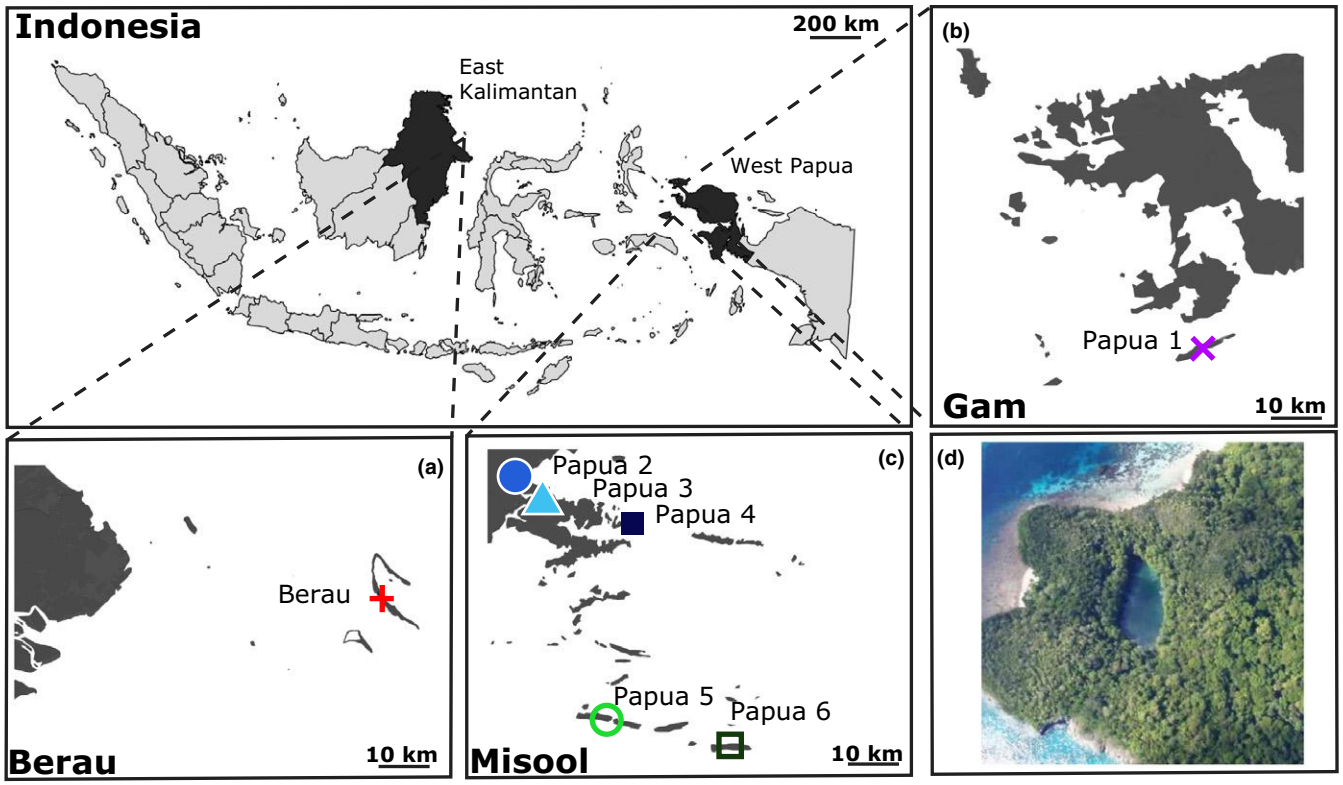

FIGURE 1 Overview of *Brachidontes* sp. sampling sites in seven marine lakes in Indonesia. Land is indicated in dark grey, and sea is indicated in white. Marine lakes are located inland on different landmasses, indicated with a symbol. (a) Berau, East Kalimantan. (b) Papua 1, Gam, Raja Ampat. (c) Papua 2–6, Misool, Raja Ampat. (d) Marine lake, photograph by L.E. Becking. There is approximately 1,400 km distance between East Kalimantan and West Papua [Colour figure can be viewed at wileyonlinelibrary.com]

Adductor muscles from *Brachidontes* sp. samples were excised and preserved in 99% ethanol at 0–4°C while in the field (4–8 weeks). In the laboratory, tissue samples were stored in a −80°C freezer until further use.

We logged lake coordinates using a Global Positioning System (Garmin Gpsmap 64S). Lake coordinates were used to calculate minimum pairwise geographic distances between lakes via the R function *distm* implemented in the geosphere package. Environmental characteristics defined as temperature (°C), salinity (ppt) and pH were measured using a YSI Professional Plus Multimeter from an inflatable boat (Table 1). Measurements were taken at 1-m intervals from 1 m to 5 m depth in each lake, as the bulk of *Brachidontes* sp. bivalves are found within this range. In each lake, we measured at least 10 different sites. Pairwise environmental distance between lakes was defined via calculating Euclidean distances between lakes via a Principal Component Analysis ran with three environmental variables (Temperature, salinity and pH). The tidal amplitude was measured simultaneously inside of the lakes as well as in the adjacent sea using Hobo water-level loggers. Physical connection of the lake to the sea was defined as the ratio of maximum tidal amplitude in metres of the lakes as compared to the sea (this was termed "Connection," Equation 1).

$$\text{Connection} = \frac{\text{Max. tidal amplitude lake (m)}}{\text{Max. tidal amplitude sea (m)}} \quad (1)$$

To get a measure of "Isolation" instead of "Connection," Connection was subtracted from 1 (Equation 2). Now, lakes with a high connection ratio (e.g., 0.8) have a low value for Isolation (e.g., 0.2). Conversely, low connected lakes (e.g., connection ratio of 0.1) had high values for Isolation (e.g., 0.9).

$$\text{Isolation} = 1 - \text{Connecton} \quad (2)$$

Finally, pairwise distances between lakes were obtained by multiplying the isolation values of each pairwise combination of lakes (*i* and *j*, Equation 3), making sure the diagonal (lake *i* with itself) was put at zero. Low connection distance values indicate two lakes both with high connection (e.g., Isolation value 0.2 × Isolation value 0.2 = 0.04), and high connection distance values indicate two lakes with low connection (e.g., Isolation value 0.9 × Isolation value 0.9 = 0.81). Intermediate connection distance values indicate lakes with varying degrees of connection (e.g., 0.9 × 0.2 = 0.18).

$$\text{Connection distance}_{i,j} = \text{Isolation}_i \times \text{Isolation}_j \quad (3)$$

2.3 | Library preparation and sequencing

DNA from the excised adductor muscles was extracted using Qiagen DNeasy kit (Qiagen, Germantown, MD, USA). We built ddRAD-seq libraries for the 125 individuals following an adapted protocol of Peterson, Weber, Kay, Fisher, and Hoekstra (2012) (See Supporting Information for protocol). Briefly, we started the ddRAD-seq

TABLE 1 Overview of measured variables in sampled locations. Codes and sizes, physiographic, environmental and genetic parameters are displayed. Temperature, salinity and pH were measured at 1-m intervals from 1 m to 5 m depth in at least 10 locations per lake. Genetic measures nucleotide diversity (θ_π) and Tajima's *D* neutrality values (*D*) are displayed. Means are displayed with standard deviations. Codes correspond with Figure 1

Code	Sample size (*n*)	Surface area (m^2)	Max. depth (m)	Fraction tidal amplitude Lake/Sea	Temperature (°C)	Salinity (ppt)	pH	θ_π (%)	Tajima's *D*
Berau	15	12×10^4	–	0.38	29.5 (0.71)	26.0	–	0.90 (0.82)	−0.24 (0.85)
Papua 1	18	8.6×10^4	19	0.07	32.3 (0.13)	24.0 (0.71)	7.6 (0.05)	0.81 (0.76)	−0.30 (0.85)
Papua 2	19	1.5×10^4	7	0.31	33.6 (0.58)	25.6 (0.52)	7.8 (0.08)	0.87 (0.78)	−0.33 (0.90)
Papua 3	15	2.1×10^4	7.5	0.51	32.6 (0.62)	30.7 (0.84)	8.0 (0.06)	0.93 (0.76)	−0.62 (0.77)
Papua 4	16	1.3×10^4	20	0.80	31.7 (0.35)	25.9 (0.84)	8.1 (0.06)	0.94 (0.78)	−0.52 (0.84)
Papua 5	19	0.4×10^4	5	0.26	31.5 (0.28)	28.9 (0.26)	8.1 (0.05)	0.88 (0.76)	−0.52 (0.84)
Papua 6	23	0.3×10^4	12	0.78	31.9 (0.31)	28.3 (0.26)	7.9 (0.03)	0.89 (0.75)	−0.36 (0.88)

protocol with 600 nanograms of DNA. The DNA was digested using SphI-HF (rare-cutting) and MlucI (frequent-cutting) restriction enzymes. The fragmentation was investigated with Bioanalyzer High Sensitivity chip (Agilent). Following the spreadsheet of Peterson et al. (2012), "Locus count from Bioanalyzer % in region," the combination of enzymes was predicted to yield 64,770 sequenceable fragments when excising a region of 400–500 bp, assuming a genome size of 1 Gb. We pooled 17–18 individuals per ddRAD library and came to a total of 7 libraries for 125 specimens. We used a Sage Science Pippen Prep to size select 476–576 bp (including internal adapters) fragments. We confirmed the sizes by Bioanalyzer High Sensitivity chip (Agilent). Then, ten indexing polymerase chain reaction cycles (PCRs) were run on each library to enrich for double-digested fragments and to incorporate a unique external index for each library pool. The seven libraries were sequenced as 100-bp single read (SR100) sequences on two Illumina HiSeq 2500 lanes at the Vincent J. Coates Genomic Sequencing Facility at UC Berkeley.

2.4 | Assembly reference and filtering

We used pipelines implemented in a custom perl script invoking a variety of external programs to process ddRAD-seq data (RADTOOLKIT v0.13.10; https://github.com/CGRL-QB3-UCBerkeley/RAD). Raw fastq reads were first demultiplexed based on the sequences of internal barcodes with a maximum tolerance of one mismatch. Demultiplexed reads were removed if expected cutting sites were not found at the beginning of the 5′-end of the sequences. The resulting reads were then filtered using cutadapt (Martin, 2011) and Trimmomatic (Bolger, Lohse, & Usadel, 2014) to trim off Illumina adapter contaminations and low-quality reads. We also removed low complexity and potential bacterial, viral and human contamination sequences, using the genomes from GenBank represented in Table S1. Exact duplicates, either derived from PCR or from sequencing, could not be determined due to the nature of ddRAD data, so these were not eliminated.

The resulting cleaned reads of each individual were clustered using CD-HIT (Fu, Niu, Zhu, Wu, & Li, 2012; Li & Godzik, 2006), and only clusters with at least three reads supported were kept. For each cluster, the representative sequence determined by CD-HIT was retained. The markers were then masked for putative repetitive elements, low complexities and short repeats with Ns using RepeatMasker (Smit, Hubley, & Green, 2014), with "Mytilidae" as a database. After masking, we eliminated markers if more than 60% of the nucleotides were Ns. The resulting RAD markers from each individual were then combined and clustered for all individuals to search for those shared by at least 70% of all the individuals, which then served as the reference.

Cleaned sequence reads from all individuals were aligned to the reference using Novoalign (http://www.novocraft.com), and only reads that mapped uniquely to the reference were kept. We used Picard (http://www.picard.sourceforge.net) to add read groups and GATK (McKenna et al., 2010) to perform realignment on alignment files in BAM format, generated by SAMtools (Li et al., 2009). We then used SAMtools/BCFtools to generate data quality control information in VCF format. We filtered out any markers where more than two alleles were called on any site. We masked SNPs and sites within 10 bp upstream and downstream around an indel. Individual sites were eliminated if their depth fell outside 1st and 99th percentile of the overall coverage (among all samples). These data were further filtered using a custom sites filtering program, SNPcleaner (Bi et al., 2013), which was modified and implemented in our pipelines.

2.5 | Genotype likelihoods and allele frequency estimation

To assess potential artefacts of coverage of SNPs and percentage included individuals, we tested four filtering options in total (Table S2), varying these two parameters. We tested coverage of 3 ×, 5 × and 10 ×, and percentage included individuals of 40% and 70%. We found that downstream patterns remained highly similar (see Figures S1, S2 and Table S3). This indicated that there was no large influence of minimum coverage or percentage of included individuals on the observed biological patterns. All further results are

based on SNPs having at least 3 × coverage and a minimum of 70% included individuals.

SNP and genotype calls based on allele counts might show high uncertainty and could cause potential bias or introduce noise in downstream analyses (Johnson & Slatkin, 2008; Lynch, 2008). To account for uncertainty in the data, we used genotype likelihoods instead of genotype calls whenever possible. Genotype likelihoods were calculated in an empirical Bayesian framework, implemented in analysis of next-generation sequencing data (ANGSD; http://www.popgen.dk/angsd/index.php/ANGSD; Korneliussen, Albrechtsen, & Nielsen, 2014). This software is specialized in analysing low to medium coverage next-generation sequencing data. Most of the downstream analyses implemented in ANGSD were performed based on likelihood of site allele frequencies, genotype likelihood or genotype posterior probabilities. For some analyses performed by external programs that rely on called genotypes, we used genotype posterior probability of .95 as a cut-off to output a list of high confidence variants.

2.6 | Within-population diversity and demography

Overall genetic variation within marine lakes was estimated via two methods of calculating nucleotide diversity. First, we calculated average number of pairwise differences between sequences, theta pi (θ_π; Nei, 1987), and second we calculated Watterson's theta as the total number of segregating sites (θ_W; Watterson, 1975). We ran Pearson's correlations of nucleotide diversity (using θ_π) versus environmental variables and connection fractions to see how these variables affect within-population diversity. From both thetas, we computed Tajima's *D* as a neutrality test to examine genomic evidence for population expansion or decline (Tajima, 1989). We computed per-individual inbreeding coefficients (*F*), calculated from degree of deviation from Hardy–Weinberg equilibrium via ngsF (Vieira, Fumagalli, & Albrechtsen, 2013). Furthermore, we computed stairway plots (Liu & Fu, 2015) to estimate changes in effective population size (N_e) over time, using a generation time of 1 year (Morton, 1988) and a range of mutation rates from 1.0×10^{-8} to 3.5×10^{-8} per site per generation (Salojärvi *et al*., 2017).

2.7 | Population genetic structure

We summarized genotypic differentiation among marine lake populations using different strategies. First, we performed a Principal Components Analysis (PCA) of the covariance matrix of posterior genotype probabilities as implemented in ngsTools (http://github.com/mfumagalli/ngsTools; Fumagalli, Vieira, & Linderoth, 2014). PCAs are commonly used in analysing SNP data, as it is an unsupervised clustering method, which may discern population structure in an unbiased manner. The first four principal components were included based on their eigenvalues. We further explored the data by performing a neighbour-joining network (NeighborNet) analysis based on uncorrelated *p*-distances in Splitstree (Huson, 1998; Huson & Bryant, 2006). This shows how well the data would fit a phylogenetic tree, without forcing a tree-like structure onto the data. Furthermore, a genetic distance matrix was computed from genotype probabilities via the program NGSDIST (Vieira, Lassalle, Korneliussen, & Fumagalli, 2016). A bootstrapped neighbour-joining tree matrix was computed from 1,000 possible trees via RAXML (Stamatakis, 2014), converted to a phylogenetic tree via FASTME (Lefort, Desper, & Gascuel, 2015) and visualized in FigTree v.1.4.2 (Rambaut, 2009).

Genetic differentiation among populations as summarized by the fixation index (F_{ST}) was calculated in ANGSD using the shared site frequency spectrum of each pairwise combination of lakes (2dSFS; Korneliussen et al., 2014). Next, we explored genetic structure across populations via admixture analysis implemented in ngsAdmix (Skotte, Korneliussen, & Albrechtsen, 2013). By calculating admixture proportions per individual, the ancestry of populations could be defined. Finally, a connectivity network was computed using the Nei's G_{ST} calculation implemented in the DIVERSITY package of R (Sundqvist, Keenan, Zackrisson, Prodohl, & Kleinhans, 2016).

2.8 | Inferences on modes of isolation

We used Mantel's tests (Mantel, 1967; Slatkin, 1993) to test significance among genetic, geographic, environmental distance matrices to elucidate the importance the different processes causing isolation. As genetic distance, we used normalized pairwise genetic differentiation ($F_{ST}/(1-F_{ST})$; Slatkin, 1995). Mantel's tests were run via the function *mantel* from the R package VEGAN (Oksanen et al., 2016), with 10,000 permutations. A correlation of $r > .6$ was considered strong, and significance was assigned when the *p*-value was smaller than .05. We confirmed the absence of a correlation between geographic and environmental distance (Mantel's test, $r = -.25$, *p*-value = .78), which allowed us to distinguish between scenarios of dispersal limitation and adaptation to the environment. We ran similar tests for geographic and connection distance, and environmental and connection distance, and similarly, no correlations were found (Mantel's test, $r = .34$, *p*-value = .21, and $r = .25$, *p*-value = 0.24, respectively).

Finally, we further explored the effect of environmental variables influencing genetic structure. A Spearman correlation test was run between the first principal component from Principal Component Analysis of the lakes in Raja Ampat (explaining 10.73% of total genetic variation, Figure S13) and the environmental variables.

3 | RESULTS

3.1 | Filtering

We obtained 103 million reads from the Illumina sequencer after demultiplexing. From these reads, we obtained 116,416 anonymous SNPs usable in downstream analyses. Principal Component Analyses, Splitstree and F_{ST} estimates remained highly similar with different filtering options (coverage: 3×, 5× or 10×, included individuals 40% or 70%; Figures S1 and S2; Table S3). Therefore, all subsequent results are based on SNPs having at least 3× coverage and 70% included individuals. On average, samples had a coverage of 12× (range: 3–42×) across all loci.

3.2 | Within-population diversity and demography

First, we investigated genetic diversity within *Brachidontes* sp. populations from seven marine lakes. Per-population nucleotide diversity (θ_π) was similar among all lakes (Table 1, θ_π). The highest nucleotide diversity was observed for Papua 4 (0.0094) and lowest for Papua 1 (0.008). Although the standard deviation of both θ_π and θ_W was high, they showed consistent patterns per lake (Table 1, Table S4 and Figure S4). We observed a trend towards higher nucleotide diversity with increasing connection to the sea (Pearson's correlation, $r = .78$, p-value = .04; Figure S5). Inbreeding coefficients (F) were generally found to be low: range (0.003–0.041; Table S4).

All lakes were found to have slightly negative Tajima's D values (average D-value: −0.41, SD: 0.14; Table 1, Figure S6). The lowest value of Tajima's D was observed for Papua 3 (−0.62) and the highest for Berau (−0.24). Negative Tajima's D values are generally associated with populations showing recent expansion after a bottleneck. The stairway plots consistently showed a bottleneck at 4,000 to 6,000 years before present for each population, indicating the early colonization of the lakes by *Brachidontes* sp. (Figure S7).

3.3 | Population differentiation and structure

Next, we investigated population structure. Clear structuring was observed in the first four principal components of the Principal Component analysis, together explaining 28.81% of total genetic variation (Figure 2a). The first Principal Component, explaining 10.87% of the total variation, separated East Kalimantan (Berau) from the others. The second Principal Component (explaining 8.50%) separated Papua 1 and 2 from the remaining Misool lakes. Principal Components 3 and 4 (together explaining 9.44% of the variation) clearly separated the lakes in southern Misool (Papua 5 and 6) from the others. None of the Principal Components separated Papua 3 and 4.

The pattern observed in the Principal Component Analysis was supported by the neighbour-joining network constructed using Splitstree (Figure 2b). The network was based on 390 splits and had a fit of 99.3. The degree of reticulation was assessed with a delta score of 0.224 and a Q-residual score of 9.28×10^{-5}. These relatively low scores show that the data followed a tree-like pattern (Holland, Huber, Dress, & Moulton, 2002). All lakes showed distinct separation, with Berau and Papua 1 being most distant from the others. Papua 6 showed additional splits with Papua 2 and 3, indicating that some individuals might share genetic variation (Figure S8). The neighbour-joining tree showed congruent patterns (Figure S9). The obtained connectivity network confirmed patterns from F_{ST} comparisons and the Splitstree network (Figure S10). Lowest bidirectional connectivity was found between Berau to all other lakes, whereas highest bidirectional connectivity was found between Papua 3 and Papua 4.

Pairwise population fixation indices (F_{ST}) showed moderate-to-high genetic structuring between all pairwise lake comparisons (Table 2, Figure S11). Pairwise F_{ST} values varied between 0.069 (Papua 3 and 4) and 0.235 (Berau and Papua 1). On average, marine lakes had a global fixation index of 0.145 (SD = 0.05). Admixture analysis showed clear distinction between lakes (Figure 3). Conversion of likelihood values of 5 replicate runs showed the lowest variance for seven ancestral populations (Figure S12). When putative ancestral populations were set to $K = 7$, all seven marine lakes were separated. Increasing the number of ancestral populations from $K = 3$ to 7 showed a high similarity of Papua 3 and 4, as these lakes only became separate at $K = 7$. Some individual admixture between Papua 2, 3 and 6 could be seen, consistent with the additional splits in the neighbour-joining network (Figure 2b).

3.4 | Relative importance of modes of isolation

Finally, we found clear evidence for isolation by distance on large spatial scales (>1,400 km) with a strong and significant correlation between geographic distance and genetic distance (Mantel's test, $r = .92$, p-value < .001), whereas environmental and connection distance both were not significant (Mantel's test, $r = .44$, $p = .10$, and $r = .46$, $p = .11$, respectively). On the spatial scale of 200 km, we found associations with geographic distance (Mantel's test, $r = .82$, $P = .01$) and connection distance (Mantel's test, $r = .63$, $p = .03$), but not for environmental distance (Mantel's test, $r = .21$, $p = .29$; Figure 4). Remarkable is the wide range of genetic differentiation among relatively well-connected lakes (left part of Figure 4c). On the smallest spatial scale (40 km), all associations with geographic, environmental and connection distance become less pronounced and insignificant (Mantel's tests, geography: $r = .50$, $p = .05$, environment: $r = .16$, $P = .35$, connection: $r = .40$, $p = .19$). Furthermore, we observed a significant correlation of water temperature to the first Principal Component (Figure 1) on the two smaller spatial scales (200 km and 40 km; Spearman's correlation, $\rho = 0.49$ and 0.78, respectively, p-value < .001; Figure 5).

4 | DISCUSSION

Understanding processes underlying population genetic structure in the marine realm is critical for predicting ecosystem responses to natural and anthropogenic change. By comparing marine lakes in Indonesia with similar ages and sizes, but varying degrees of connection to the sea and differing environmental regimes, we tested the relative contribution of isolation-by-distance, isolation-by-environment and historical priority effects in the formation of population genetic patterns of a bivalve (*Brachidontes* sp.). The results indicate strong genetic structure despite incomplete dispersal barriers and provide important insights into the role that (evolution-mediated) priority effects may play in influencing rapid population divergence in peripheral environments.

As marine lakes are landlocked water bodies with subterranean connection to the sea, they can be regarded as the marine equivalent of terrestrial islands. After the origin of marine lakes, they provided novel habitat to be colonized, and we assume that the floodwaters that filled the lakes brought in the propagules of populations from the source pool and that these were the progenitors of current marine lake populations (Dawson & Hamner, 2005).

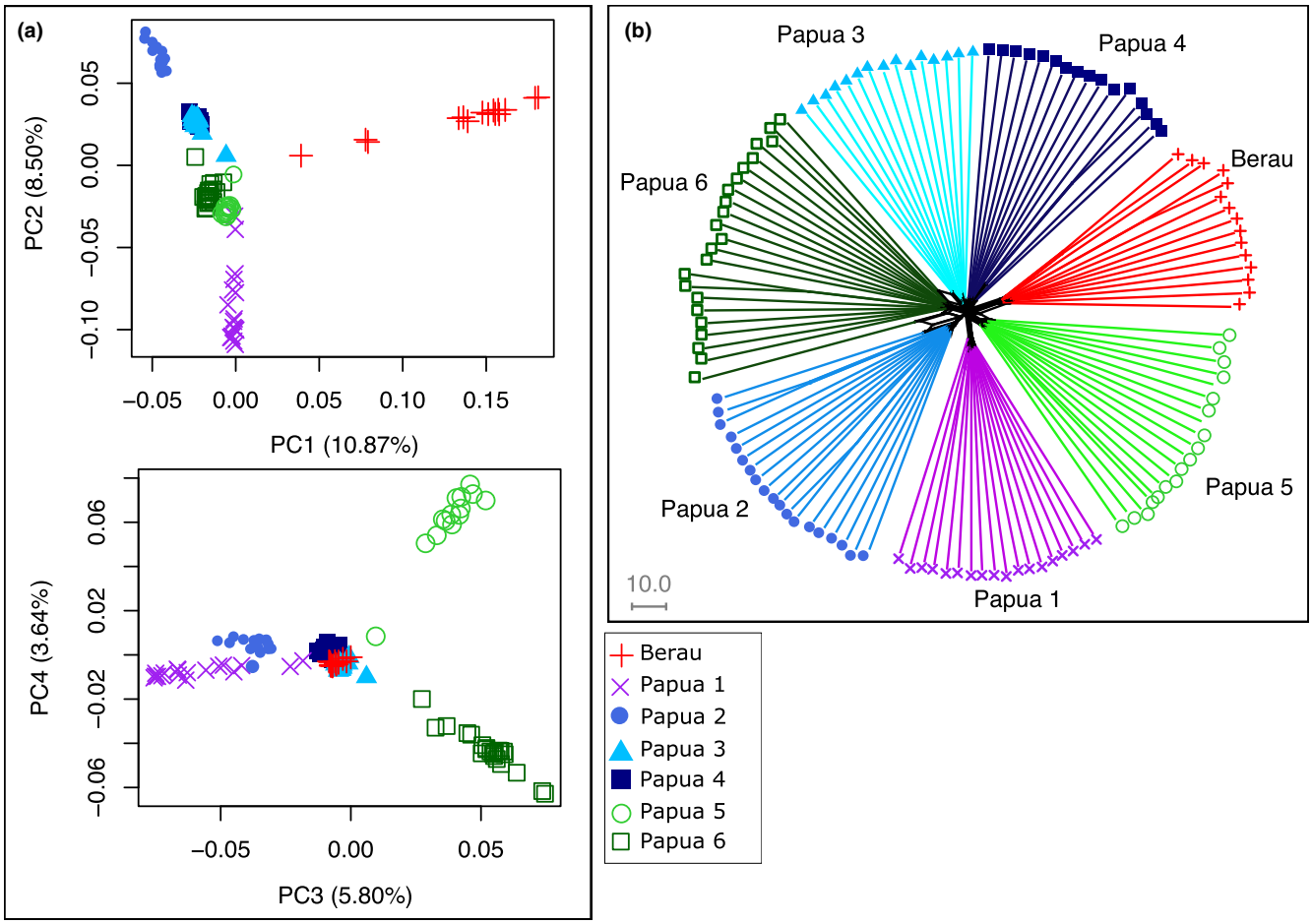

FIGURE 2 Genetic structure of seven populations of *Brachidontes* sp. mussels in marine lakes in Indonesia. (a) Principal component analysis (PCA) based on pairwise genetic covariance among 125 individuals with 116,415 SNPs, from all seven locations. First four axes represent 28.8% of total genetic variation. Each dot is one individual. (b) Neighbor-joining network with equal angles computed in Splitstree from pairwise genetic distances. Splitstree based on 390 splits and a delta score of 0.224. Scale bar indicates number of substitutions per site. Colours and codes correspond to Figure 1 and Table 1 [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 2 Genetic differentiation among seven *Brachidontes* sp. populations from marine lakes in Indonesia. Genetic fixation indices (F_{ST}) are displayed, showing strong genetic differentiation. Codes correspond with Figure 1 and Table 1

	Berau	Papua 1	Papua 2	Papua 3	Papua 4	Papua 5
Papua 1	0.235	–				
Papua 2	0.213	0.179	–			
Papua 3	0.184	0.148	0.082	–		
Papua 4	0.191	0.155	0.104	0.069	–	
Papua 5	0.198	0.153	0.138	0.106	0.116	–
Papua 6	0.198	0.151	0.126	0.089	0.108	0.102

Preliminary studies of sediment cores from the marine lakes indicate that *Brachidontes* specimens were present at the onset of the lakes in Raja Ampat (Klei, 2016). We assume that the degree of connection between lakes and the sea has remained similar since the formation of the lake, and the potential for novel migrants to come into the lakes remained.

We show that short-term (<6,000 years) reduction in connection between populations, potentially followed by (evolution-mediated) priority effects, can lead to strong population divergence. The role of priority effects is often neglected in population genetic and phylogeography studies, but may be relevant in the context of ongoing environmental change and habitat fragmentation, all influencing landscape connectivity. We first discuss within-lake structure and demography, then elaborate on the observed population structure and finally disentangle the importance of the three modes of isolation: isolation-by-distance, isolation-by-environment and (evolution-mediated) priority effects.

4.1 | Marine lake colonization and population structure

We investigated early colonization of *Brachidontes* sp. mussels of the marine lakes based on analyses for bottlenecks (Table 1, Figure S7). Our findings are consistent with mitochondrial DNA data (COI) based on the same mussel lineage analysed in this study (Becking

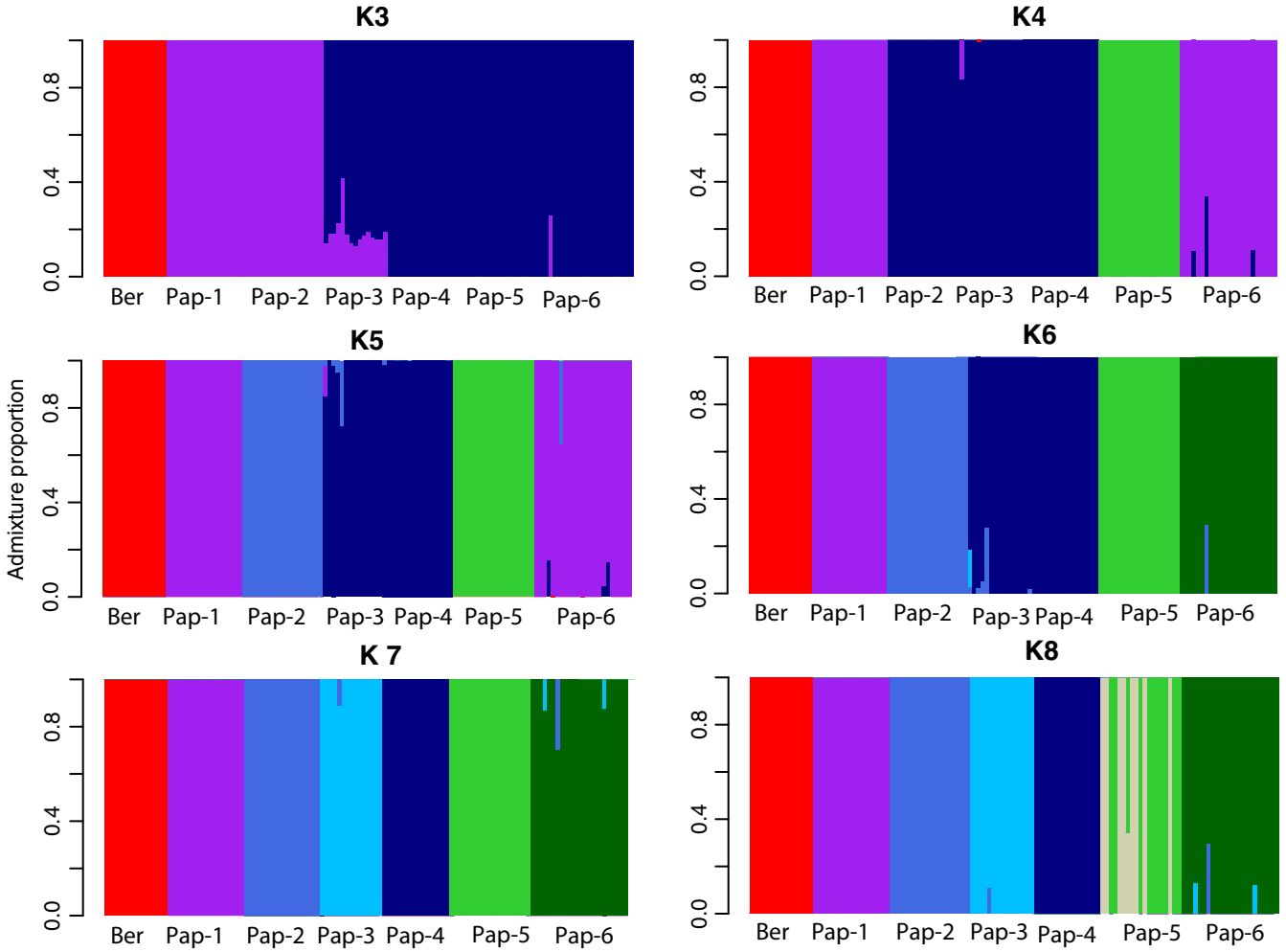

FIGURE 3 Individual admixture plots of *Brachidontes* sp. populations from seven marine lakes in Indonesia. Bayesian admixture analysis for a range of putative ancestral populations (*K*) based on genotype likelihoods via ngsAdmix. Highest likelihood was obtained for $K = 7$ (Figure S12). Each bar represents one individual. Colours within bars represent admixture proportions and correspond to Figure 1 and Table 1 [Colour figure can be viewed at wileyonlinelibrary.com]

et al., 2016; de Leeuw et al., *under review*). Their mismatch distributions followed a unimodal distribution which is typical for bottlenecked populations and subsequent expansions (Becking et al., 2016). In the current study, the estimated timing of bottlenecks events found for the majority of lakes is approximately 4,000 to 6,000 years before present, which corresponds to the presumed dates of filling of the marine lakes (Sathiamurthy & Voris, 2006). We do note that the estimations may be crude, as they have large confidence intervals, and are dependent on a variety of assumptions (mutation rate and generation time). Bottlenecks with subsequent expansions have also been found in studies of other peripheral environments (Dawson & Hamner, 2005; Goto et al., 2011; Gotoh, Chiba, Goto, Tamate, & Hanzawa, 2011; Hanzawa et al., 2012), which show rapid accumulation of mutations in populations after colonization of the habitat in a typical pattern of founder effects.

The nucleotide diversities of marine lake populations were relatively high (ranging between 0.008 and 0.010, Table 1; Pazmiño, Maes, Simpfendorfer, León, & Van Herwerden, 2017). High nucleotide diversity is generally not expected in bottleneck scenarios due to a loss of alleles and increased inbreeding (Fauvelot, Bernardi, & Planes, 2003; Gotoh et al., 2011; Hohenlohe et al., 2010). Indeed, inbreeding coefficients for marine lake populations were found to be low (*F* mean 0.018 ± 0.024), where severe inbreeding is indicated by *F* values larger than 0.1 (Hazzouri et al., 2015; Vieira et al., 2013; Wang et al., 2016). This indicates that our species may not suffer from inbreeding depression after bottlenecks, for instance due to its high reproductive potential. We did observe a trend towards higher nucleotide diversity with increasing connection of the lake to the surrounding sea. With higher connection to the sea, the influx of genetically different propagules in highly connected lakes will be more frequent. Higher nucleotide diversity in more connected lakes may therefore be the result of a higher rate of (successful) immigration.

4.2 | Rapid population differentiation

We observed striking genetic structure among the seven marine lake populations, indicating limitations to gene flow, even on small spatial

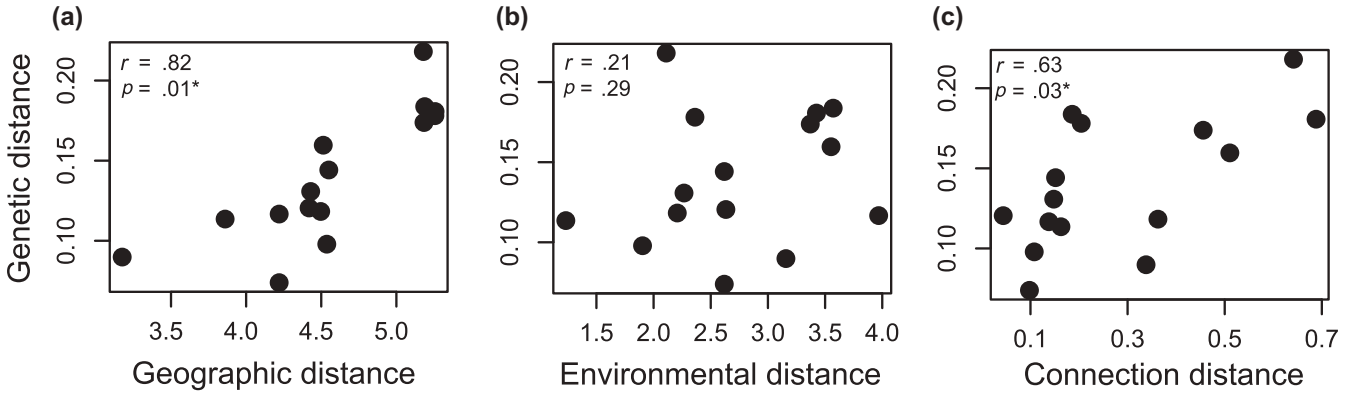

FIGURE 4 Relative importance of three modes of isolation tested for *Brachidontes* sp. populations from six marine lakes in West Papua, Indonesia (scale < 200 km). Mantel test results of genetic distance (normalized F_{ST}) versus geographic distance (log-transformed m), (b) versus environmental distance (Euclidean distances based on principal component analysis or temperature, salinity, pH and dissolved oxygen), (c) versus connection distance. Mantel's observation values (*r*) and *p*-values are displayed. Each dot represents a pairwise comparison between two populations. Asterisks represent significance at the alpha = 0.05 level.

(<200 km) and temporal (~6,000 years) scales. Compared to other studies using RADseq in marine organisms, mussels from marine lakes show moderate-to-high pairwise F_{ST} values, and clear structure in admixure analyses (Lal, Southgate, Jerry, & Zenger, 2016; Tariel, Longo, & Bernardi, 2016; Van Wyngaarden et al., 2016). There is not a complete barrier to the sea, as the vector of water that can carry propagules in and out of the lakes is ongoing. However, there may be other barriers to dispersal or successful establishment of immigrating genotypes. Such barriers would effectively block individuals from outside the lakes to interbreed with the standing population. In the subheading "Isolation-by-distance, isolation-by-environment and priority effects," we discuss which of these modes of isolation may underlie the genetic structure.

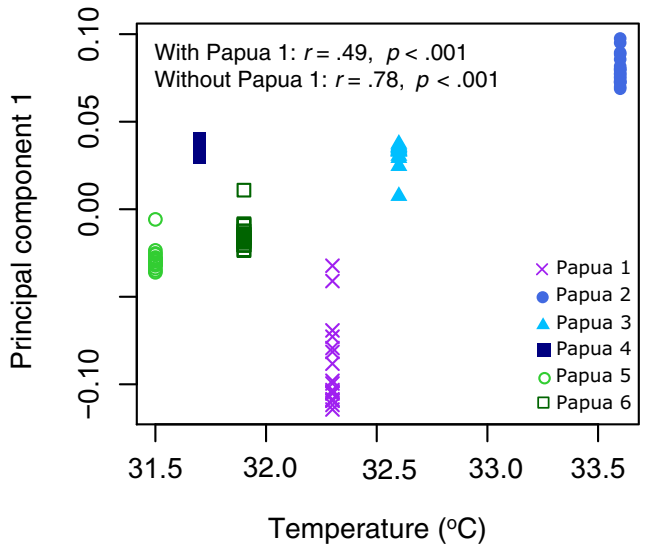

FIGURE 5 Correlation of genetic variation of *Brachidontes* sp. populations with water temperature in marine lakes from West Papua, Indonesia (scale <200 km). Correlations between first principal component of Figure S13 (explaining 10.73% of genetic variation among populations Papua 1–6) and temperature. Spearman's correlation values (ρ) and *p*-values are displayed for correlations including Papua 1 (spatial scale 200 km) and excluding Papua 1 (spatial scale 40 km). Colours and codes correspond to Figure 1 and Table 1 [Colour figure can be viewed at wileyonlinelibrary.com]

As colonization of the lakes occurred a relatively short time ago (~6,000 years), any in situ differentiation can be considered rapid. Assuming that all marine lakes were colonized from the same ancestral source pool in the sea, at least on the scale of <200 km, genetic fixation indices of >0.07 are relatively high. Still, multiple studies of three-spine stickleback, African rift lake cichlids, and organisms in anchialine ponds have found genetic and morphological diversification at even shorter time scales, as little as 150 years (Genner et al., 2007; Gonzalez et al., 2017; Lucek, Sivasundar, Roy, & Seehausen, 2013; Marques et al., 2016; Weber, Bradburd, Stuart, Stutz, & Bolnick, 2016; Weese, Fujita, Hidaka, & Santos, 2012). Even though post-glacial-derived mutations would imply rapid evolutionary rates, such rates are not uncommon (e.g., Genner et al., 2007; Ho et al., 2011). Furthermore, postglacial divergence may also come from divergent selection on standing genetic mutation (Wang & Bradburd, 2014).

Emerging patterns for multiple marine taxa in population genetic studies suggest spatial genetic structure and limited gene flow at small spatial scales, despite a lack of clear physical barriers and high dispersal potential (Barber, 2009; Carpenter et al., 2011; Hoeksema, 2007; Starger et al., 2015; Waters, Fraser, & Hewitt, 2013). For example, genetic structure was found for stomatopods and caridean shrimp (range 200–300 km; Barber et al., 2006; Haig, Connolly, & Hughes, 2010), damselfish (range 25–150 km; Timm & Kochzius, 2008), giant boring clams (range 25–50; DeBoer, Subia, Erdmann, Kovitvongsa, & Barber, 2008) and starfish (range 10–15 km; Crandall et al., 2008). A number of processes have been suggested underlying the high marine biodiversity in the Coral Triangle, such as glacial cycles, heterogeneous environments, complex bathymetries and ocean current systems forming eddies to effectively trap larvae (Hoeksema, 2007). Glacial cycles and complex bathymetries may

enhance patterns of isolation-by-distance or priority effects. As sea levels were approximately 120 m lower than modern sea levels at the Last Glacial Maximum (20,000 years BP), the reef setting of the Sahul shelf, including West Papua, would have consisted of many semi-isolated basins with reduced connection among each other (Hoeksema, 2007; Sathiamurthy & Voris, 2006). Hence, peripheral environments, similar to modern-day marine lakes, have likely been present throughout history in the Coral Triangle area. Additionally, heterogeneous environments may facilitate isolation-by-environment as populations may become locally adapted.

Elucidating the mechanisms of divergence is challenging as means to speciation are often obscure (Coyne & Orr, 2004; Palumbi, 1994), and typically a reliable historical component is lacking (Ho et al., 2011; Sanderson, 1997). The existence of multiple independently derived populations in landlocked marine lakes with varied environments provides an opportunity for fundamental research into the role of short-term and incomplete isolation in population divergence.

4.3 | Isolation-by-distance, isolation-by-environment and priority effects

We aimed to disentangle relative importance of three modes of isolation (Figure 6). We assume that small populations colonized the marine lakes after their formation and have likely experienced founder effects (Barton & Charlesworth, 1984; Mayr, 1954). Genetic drift may be particularly strong in small, founding populations, as alleles rare in the source population may stochastically be driven to fixation due to a subset of individuals colonizing new habitats. Therefore, founding populations can be expected to be genetically distinct from the ancestral populations and from conspecifics colonizing different marine lakes within several thousand generations. There has, however, been a continued connection between lakes and the sea and a potential for new migrants. Genetic signatures resulting from founder effects would then be expected to be overwhelmed by ongoing dispersal from the source population or from other marine lakes or by forces of natural selection (Mayr, 1963; Waters et al., 2013).

On large spatial scales (>1,400 km), we find support for isolation-by-distance, formed likely by dispersal limitation. Due to the geographic distance, currents and obstructing landmasses between the lake in Berau and lakes in Papua, larvae are unlikely to be able to maintain genetic connectivity (Barber, Cheng, & Erdmann, 2011). With decreased dispersal between populations, there is subsequently decreased gene flow, and genetic drift may cause populations to become genetically distinct (Slatkin, 1985; Wright, 1943). Theoretical models have showed that larvae of marine organisms are able to traverse large distances, via passive dispersal via oceanic currents (Armsworth, 2002). This may be aided by long dispersive larval stages, which species of the genus Brachidontes are known to have (up to four weeks). However, those predictions are often overestimations of actual dispersal (Marshall et al., 2010), as shown

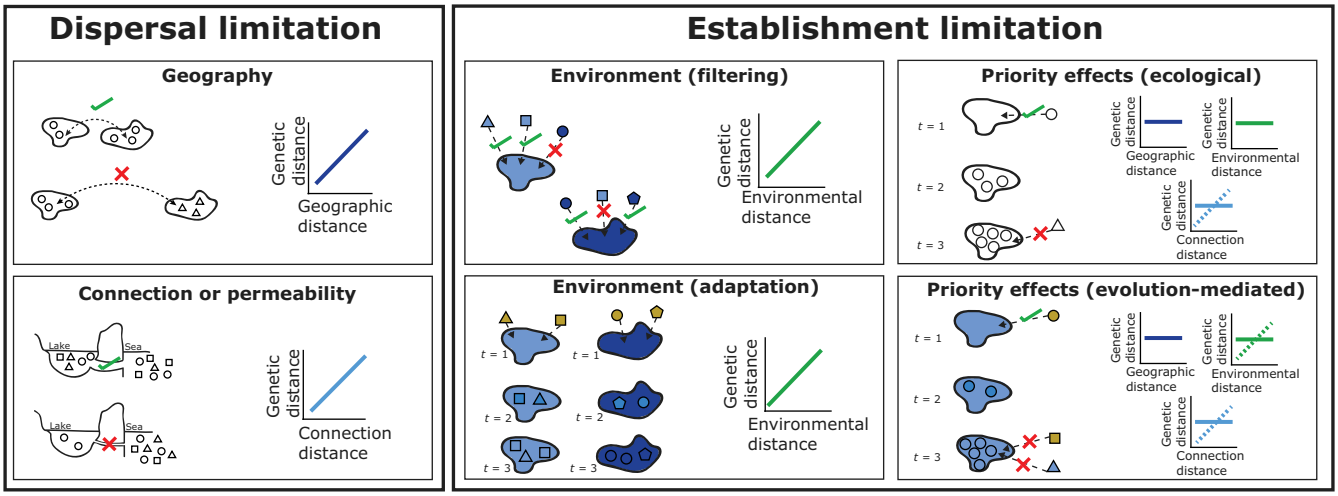

FIGURE 6 Modes of isolation and predictions for population genetic differentiation (F_{ST}). Dispersal limitation is caused by a reduction in homogenizing gene flow due to reduced dispersal success. Geographic distance can play a role when genetic distance between populations will increase with increasing geographic distance in a pattern of isolation-by-distance. Alternatively, permeability of habitats or landscape barriers can play a role. Connection distance represents the permeability of habitats to entrance of propagules, where higher connection distance means less permeable habitats. This results in a pattern of high genetic structure among populations inhabiting less permeable habitats, compared to populations from more permeable habitats. Alternatively, establishment limitation is caused by a reduction in homogenizing gene flow due to reduced establishment success of propagules within a habitat. It arises when there are differential environments, which either act as filters or which cause local adaptation of populations. Both result in an isolation-by-environment pattern, which entails populations in environmentally similar habitats will have ongoing gene flow, while populations in environmentally dissimilar habitats will not. Establishment limitation can also be caused by historical priority effects, which results in colonizing populations being able to outcompete any new immigrants via a numerical advantage, potentially aided by evolutionary local adaptation. Dotted lines in predictions for connection distance indicate the time lag necessary for priority effects to occur and be maintained is more likely in habitats that have low connection. Dotted lines in environmental distance indicate local adaptation may result in a pattern of isolation-by-environment [Colour figure can be viewed at wileyonlinelibrary.com]

particularly for the Coral Triangle (Treml, Roberts, Halpin, Possingham, & Riginos, 2015). Furthermore, there is an issue of dilution, where a high density of larvae is needed to maintain gene flow between habitats at large distances (Johannesson, 1988).

On smaller spatial scales (<200 km), we find that geographic distance also plays a significant role, although it becomes less pronounced and not significant at the scale of 40 km (Figure 4). No signatures of isolation-by-environment were found on any spatial scale; thus we do not find evidence for isolation-by-environment based on the environmental data we used: temperature, salinity and pH (Orsini et al., 2013; Wang & Bradburd, 2014). It is possible we have not accounted for certain environmental variables that may be important in population adaptation. Including parameters such as nutrient level or biotic interactions may increase resolution in finding signatures of local adaptation. On the scale of <200 km, we do find a positive relationship between genetic variation and degree of connection to the sea, indicating that the extent of connection to a source of propagules is important in influencing genetic structure among populations, which can also be thought of as an isolation-by-resistance pattern (McRae, 2006). We expect more isolated lakes will have a more pronounced delay in the arrival of new genotypes, which is corroborated by the finding of higher nucleotide diversity within populations inhabiting marine lakes with higher connections to the sea. The observation of more isolated lakes being more genetically distinct may be consistent with the hypothesis that priority effects are stronger in more isolated environments (Fukami, 2015; Orsini et al., 2013). Density-dependent ecological priority effects may further be mediated by evolution via adaptation of first colonizers to local conditions (De Meester et al., 2016). By having a head start in locally adapting to the new environment, early colonizers may become even more strong competitors to any future immigrants. If this were the case, we would expect that genotypes show a correlation to important environmental parameters. The data corroborate this by showing a strong correlation to water temperature, particularly on the scale of <40 km, which potentially indicates local adaptation may have occurred. Furthermore, priority effects predict extensive small-scale genetic differentiation (De Meester et al., 2016), which we observe even at the scale of <40 km. Incomplete dispersal barriers facilitated by physical barriers may provide the ideal setting for priority effects to arise and subsequently be fixed in the long-term via (evolution-mediated) adaptation.

The influence of priority effects on genetic structure we find is supported by the broadcast spawning life strategy of mussels, which allows them to reproduce rapidly, which is beneficial to the colonization of novel habitats. Mussels show a range of tolerance towards temperature and salinity (Dowd & Somero, 2012; Sarà, Romano, Widdows, & Staff, 2008; Spidle, May, & Mills, 1995). Particularly species of the genus *Brachidontes*, such as *Brachidontes pharaonis*, appear to be highly tolerant to low salinities and high temperatures (Sarà et al., 2008). The high plasticity would benefit colonization of diverse environments. As we find potential influences of temperature in determining genetic variation, we hypothesize that environment has a role in short-term population differentiation. To further explore the extent of evolution-mediated priority effects in marine lake systems, neutral and adaptive loci need to be distinguished and signatures of local adaptation further identified.

High differentiation in peripheral populations which are assumed to have ongoing gene flow is found in multiple marine taxa and could support a role for priority effects (Dawson & Hamner, 2005; Goto et al., 2011; Gotoh et al., 2011; Pinheiro et al., 2017; Swift, Gómez Daglio, & Dawson, 2016). These studies showed evidence for genetic isolation on similar time scales as this current study for jellyfish ($\phi_{ST} > 0.74$; Dawson & Hamner, 2005), fish (ϕ_{ST} ranging from 0.040 to 0.728; Gotoh et al., 2011) and within the *Brachidontes* genus (*p*-distance = .146; Goto et al., 2011). Isolation-by-distance, isolation-by-environment and priority effects are not mutually exclusive and may have different relative importance on different spatial and temporal scales. They form umbrella patterns for specific types of processes driving genetic structure, such as larval mobility (isolation-by-distance), selection against immigrants (isolation-by-environment) and competition (priority effects). We hypothesize that density-dependent priority effects, potentially strengthened by local adaptation, may be more prevalent than previously assumed. In coastal reef systems, ongoing habitat fragmentation and climate change may accelerate processes of priority effects (Legrand et al., 2017), which may have important implications for conservational efforts. Our study supports a role of eco-evolutionary dynamics in early stages of habitat colonization and subsequent population divergence.

4.4 | Low-cost method for nonmodel organisms

On a final note, we would like to stress the importance of developing low-cost next-generation sequencing methods for nonmodel organisms. We used a low-cost method of population genomic library preparation, and an extensive step-by-step protocol is provided in the Supporting Information to enhance the use in low-budget projects. We found major structures and conclusions to remain the same when trying different filtering options, showing that low coverage data can still provide accurate information. Hence, high numbers of individuals can be sequenced on one lane, reducing costs. This is especially relevant for laboratories with limited funding and working on nonmodel organisms, for example in the Coral Triangle. We hope that low-cost protocols such as the one used in the current study promotes and facilitates more extensive studies into tropical marine biodiversity.

ACKNOWLEDGEMENTS

We would like to thank two anonymous reviewers and the subject editor for valuable comments on the original manuscript. We would like to thank The Nature Conservancy and Conservation International for their logistical support during our sampling period in Indonesia. Furthermore, we would like to thank the following people for assistance with the library preparation in the laboratory and

consultation about analyses: Cerise Chen, Emily Farrer, Maria Meijer, Milan Malinksy, George Roderick and Norah Saarman. The following people provided help with logistics for field and/or laboratory work: RCO-LIPI, RISTEK; Bahruddin; Bert Hoeksema; Purwanto; A.S. Sidik; Suharsono; Ricardo Tapilatu; Yosephine Tuti; Nicole de Voogd and the staff of Nabucco Island Dive Resort, Derawan Dive Resort and Misool Eco Resort. For that, they have our sincere thanks.

DATA ACCESSIBILITY

Data is accessible via the Dyrad Digital Repository (https://doi.org/10.5061/dryad.1r89hq8). Available data consist of "Brachidontes_Rawreads" including raw reads in fastq format for all 125 individuals sequenced during this study, "Populations.txt" allocating individual raw reads to one of the seven marine lakes and "Brachidontes_Reference" including the reference in fasta format generated during this study to which individual reads were aligned.

AUTHOR CONTRIBUTIONS

D.L.M. conceived and designed the experiments, performed the experiments, analysed the data, wrote the manuscript, prepared figures and/or tables and reviewed drafts of the manuscript. S.P. contributed analysis tools, analysed the data, wrote the manuscript and reviewed drafts of the manuscript. K.B. analysed the data, contributed reagents/materials/analysis tools, wrote the manuscript and reviewed drafts of the manuscript. L.S. contributed analysis tools, wrote the manuscript and reviewed drafts of the manuscript. E.E.A. contributed reagents/materials/analysis tools and reviewed drafts of the manuscript. L.P.A and A.H.A.T. contributed reagents/materials/analysis tools, reviewed drafts of the manuscript and permit application and material transfer. R.G. wrote the manuscript and reviewed drafts of the manuscript. L.E.B. conceived and designed the experiments, performed the experiments, contributed reagents/materials/analysis tools, wrote the manuscript and reviewed drafts of the manuscript.

REFERENCES

Armsworth, P. R. (2002). Recruitment limitation, population regulation, and larval connectivity in reef fish metapopulations. *Ecology*, *83*, 1092–1104. https://doi.org/10.1890/0012-9658(2002) 083[1092:RLPRAL]2.0.CO;2

Barber, P. H. (2009). The challenge of understanding the Coral Triangle biodiversity hotspot. *Journal of Biogeography*, *36*, 1845–1846. https://doi.org/10.1111/j.1365-2699.2009.02198.x

Barber, P. H., Cheng, S. H., & Erdmann, M. V. (2011). Evolution and conservation of marine biodiversity in the Coral Triangle: Insights from stomatopod Crustacea. *Crustacean Issues*, *19*, 129–156. https://doi.org/10.1201/CRCCRUSTAISS

Barber, P. H., Erdmann, M. V., & Palumbi, S. R. (2006). Comparative phylogeography of three codistributed stomatopods: Origins and timing of regional lineage diversification in the coral triangle. *Evolution*, *60*, 1825–1839. https://doi.org/10.1111/j.0014-3820.2006.tb00526.x

Barton, N. H., & Charlesworth, B. (1984). Genetic revolutions, founder effects, and speciation. *Annual Reviews of Ecology and Systematics*, *15*, 133–164. https://doi.org/10.1146/annurev.es.15.110184.001025

Becking, L. E., Cleary, D. F. R., & De Voogd, N. J. (2013). Sponge species composition, abundance, and cover in marine lakes and coastal mangroves in Berau, Indonesia. *Marine Ecology Progress Series*, *481*, 105–120. https://doi.org/10.3354/meps10155

Becking, L. E., de Leeuw, C. A., Knegt, B., Maas, D. L., de Voogd, N. J., Suyatna, I., & Peijnenburg, K. T. (2016). Highly divergent mussel lineages in isolated Indonesian marine lakes. *PeerJ*, *4*, e2496. https://doi.org/10.7717/peerj.2496

Becking, L. E., de Leeuw, C., & Vogler, C. (2015). Newly discovered "jellyfish lakes" in Misool, Raja Ampat, Papua, Indonesia. *Marine Biodiversity*, *45*, 597–598. https://doi.org/10.1007/s12526-014-0268-6

Becking, L. E., Renema, W., Santodomingo, N. K., Hoeksema, B. W., Tuti, Y., & de Voogd, N. J. (2011). Recently discovered landlocked basins in Indonesia reveal high habitat diversity in anchialine systems. *Hydrobiologia*, *677*, 89–105. https://doi.org/10.1007/s10750-011-0742-0

Bi, K., Linderoth, T., Vanderpool, D., Good, J. M., Nielsen, R., & Moritz, C. (2013). Unlocking the vault: Next-generation museum population genomics. *Molecular Ecology*, *22*, 6018–6032. https://doi.org/10.1111/mec.12516

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Bowen, B. W., Rocha, L. A., Toonen, R. J., & Karl, S. A. (2013). The origins of tropical marine biodiversity. *Trends in Ecology and Evolution*, *28*, 359–366. https://doi.org/10.1016/j.tree.2013.01.018

Carpenter, K. E., Barber, P. H., Crandall, E. D., Ablan-Lagman, M. C. A., Mahardika, G. N., Manjaji-Matsumoto, B. M., ... Toha, A. H. A. (2011). Comparative phylogeography of the coral triangle and implications for marine management. *Journal of Marine Biology*, *2011*, 1–14. https://doi.org/10.1155/2011/396982

Carr, M. H., Neigel, J. E., Estes, J. A., Andelman, S., Warner, R. R., & Largier, J. L. (2003). Comparing marine and terrestrial ecosystems: Implications for the design of coastal marine reserves. *Ecological Applications*, *13*, S90–S107. https://doi.org/10.1890/1051-0761(2003)013[0090:CMATEI]2.0.CO;2

Cowen, R. K., Lwiza, K. M. M., Sponaugle, S., Paris, C. B., & Olson, D. B. (2000). Connectivity populations: Open or closed? *Science*, *287*, 857–859. https://doi.org/10.1126/science.287.5454.857

Coyne, J. A., & Orr, H. A. (2004). *Speciation*. Sunderland, MA: Sinauer Associates Inc.

Crandall, E. D., Jones, M. E., Munoz, M. M., Akinronbi, B., Erdmann, M. V., & Barber, P. H. (2008). Comparative phylogeography of two seastars and their ectosymbionts within the Coral Triangle. *Molecular Ecology*, *17*, 5276–5290. https://doi.org/10.1111/j.1365-294X.2008.03995.x

Dawson, M. N., & Hamner, W. M. (2005). Rapid evolutionary radiation of marine zooplankton in peripheral environments. *Proceedings of the National Academy of Sciences of the United States of America*, *102*, 9235–9240. https://doi.org/10.1073/pnas.0503635102

Dawson, M. N., Martin, L. E., Bell, L. J., & Patris, S. (2009). Marine lakes. In R. Gillespie & D. A. Clague (Eds.), *Encyclopedia of islands* (pp. 603–607). Berkeley, CA: University of California Press.

de Meester, L., Gómez, A., Okamura, B., & Schwenk, K. (2002). The Monopolization Hypothesis and the dispersal – gene flow paradox in aquatic organisms. *Acta Oecologica*, *23*, 121–135. https://doi.org/10.1016/S1146-609X(02)01145-1

De Meester, L., Vanoverbeke, J., Kilsdonk, L. J., & Urban, M. C. (2016). Evolving perspectives on monopolization and priority effects. *Trends in Ecology and Evolution*, *31*, 136–146. https://doi.org/10.1016/j.tree.2015.12.009

DeBoer, T. S., Subia, M. D., Erdmann, M. V., Kovitvongsa, K., & Barber, P. H. (2008). Phylogeography and limited genetic connectivity in the endangered boring giant clam across the coral triangle. *Conservation Biology*, *22*, 1255–1266. https://doi.org/10.1111/j.1523-1739.2008.00983.x

Dowd, W. W., & Somero, G. N. (2012). Behavior and survival of Mytilus congeners following episodes of elevated body temperature in air and seawater. *Journal of Experimental Biology*, *216*, 502–514. https://doi.org/10.1242/jeb.076620

Fauvelot, C., Bernardi, G., & Planes, S. (2003). Reductions in the mitochondrial DNA diversity of coral reef fish provide evidence of population bottlenecks resulting from Holocene sea-level change. *Evolution*, *57*, 1571–1583. https://doi.org/10.1111/j.0014-3820.2003.tb00365.x

Fu, L., Niu, B., Zhu, Z., Wu, S., & Li, W. (2012). CD-HIT: Accelerated for clustering the next-generation sequencing data. *Bioinformatics*, *28*, 3150–3152. https://doi.org/10.1093/bioinformatics/bts565

Fukami, T. (2015). Historical contingency in community assembly: Integrating niches, species pools, and priority effects. *Annual Review of Ecology, Evolution, and Systematics*, *46*, 1–23. https://doi.org/10.1146/annurev-ecolsys-110411-160340

Fumagalli, M., Vieira, F. G., & Linderoth, T. (2014). ngsTools : Methods for population genetics analyses from Next-Generation Sequencing data © The Authors 2014. Published by Oxford University Press on behalf of the Japanese. *Bioinformatics*, *30*, 1486–1487. https://doi.org/10.1093/bioinformatics/btu041

Genner, M. J., Nichols, P., Carvalho, G. R., Robinson, R. L., Shaw, P. W., Smith, A., & Turner, G. F. (2007). Evolution of a cichlid fish in a Lake Malawi satellite lake. *Proceedings of the Royal Society of London B: Biological Sciences*, *274*, 2249–2257. https://doi.org/10.1098/rspb.2007.0619

Gonzalez, B. C., Martínez, A., Borda, E., Iliffe, T. M., Fontaneto, D., & Worsaae, K. (2017). Genetic spatial structure of an anchialine cave annelid indicates connectivity within - but not between - islands of the Great Bahama Bank. *Molecular Phylogenetics and Evolution*, *109*, 259–270. https://doi.org/10.1016/j.ympev.2017.01.003

Goto, T. V., Tamate, H. B., & Hanzawa, N. (2011). Phylogenetic characterization of three morphs of mussels (Bivalvia, Mytilidae) inhabiting isolated marine environments in Palau Islands. *Zoological Science*, *28*, 568–579. https://doi.org/10.2108/zsj.28.568

Gotoh, R. O., Chiba, S. N., Goto, T. V., Tamate, H. B., & Hanzawa, N. (2011). Population genetic structure of the striped silverside, *Atherinomorus endrachtensis* (Atherinidae, Atheriniformes, Teleostei), inhabiting marine lakes and adjacent lagoons in Palau: Marine lakes are "Islands" for marine species. *Genes & Genetic Systems*, *86*, 325–337. https://doi.org/10.1266/ggs.86.325

Haddad, N. M., Brudvig, L. A., Clobert, J., Davies, K. F., Gonzalez, A., Holt, R. D., … Townshend, J. R. (2015). Habitat fragmentation and its lasting impact on Earth's ecosystems. *Science Advances*, *1*, e1500052–e1500052. https://doi.org/10.1126/sciadv.1500052

Haig, J. A., Connolly, R. M., & Hughes, J. M. (2010). Little shrimp left on the shelf: The roles that sea-level change, ocean currents and continental shelf width play in the genetic connectivity of a seagrass-associated species. *Journal of Biogeography*, *37*, 1570–1583. https://doi.org/10.1111/j.1365-2699.2010.02301.x

Hamner, W., & Hamner, P. (1998). Stratified marine lakes of Palau (Western Caroline Islands). *Physical Geography*, *19*, 175–220. https://doi.org/10.1080/02723646.1998.10642647

Hanzawa, N., Gotoh, R. O., Sekimoto, H., Goto, T. V., Chiba, S. N. & Kuriiwa, K. (2012). Genetic diversity and evolution of marine animals isolated in Marine Lakes. In M. Caliskan (Ed.), *Analysis of genetic variation in animals* (pp. 121–150). Croatia: InTech Open Access Publisher.

Hazzouri, K. M., Flowers, J. M., Visser, H. J., Khierallah, H. S. M., Rosas, U., Pham, G. M., … Purugganan, M. D. (2015). Whole genome resequencing of date palms yields insights into diversification of a fruit tree crop. *Nature Communications*, *6*, 8824. https://doi.org/10.1038/ncomms9824

Hendry, A. P., Bolnick, D. I., Berner, D., & Peichel, C. L. (2009). Along the speciation continuum in sticklebacks. *Journal of Fish Biology*, *75*, 2000–2036. https://doi.org/10.1111/j.1095-8649.2009.02419.x

Ho, S. Y. W., Lanfear, R., Bromham, L., Phillips, M. J., Soubrier, J., Rodrigo, A. G., & Cooper, A. (2011). Time-dependent rates of molecular evolution. *Molecular Ecology*, *20*, 3087–3101. https://doi.org/10.1111/j.1365-294X.2011.05178.x

Hoegh-Guldberg, O., & Bruno, J. F. (2010). The impact of climate change on the world's marine ecosystems. *Science*, *328*, 1523–1528. https://doi.org/10.1126/science.1189930

Hoeksema, B. W. (2007). Delineation of the Indo-Malayan centre of maximum marine biodiversity: The Coral Triangle. In W. Renema (Ed.), *Biogeography, time, and place: Distributions, barriers, and islands* (pp. 118–178). Dordrecht, the Netherlands: Springer. https://doi.org/10.1007/978-1-4020-6374-9

Hohenlohe, P. A., Bassham, S., Etter, P. D., Stiffler, N., Johnson, E. A., & Cresko, W. A. (2010). Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLoS Genetics*, *6*, 1–23. https://doi.org/10.1371/journal.pgen.1000862

Holland, B. R., Huber, K. T., Dress, A., & Moulton, V. (2002). Delta plots: A tool for analyzing phylogenetic distance data. *Molecular Biology and Evolution*, *19*, 2051–2059. https://doi.org/10.1093/oxfordjournals.molbev.a004030

Holthuis, L. B. (1973). Caridean shrimps found in land-locked salt-water pools at four Indo-West Pacific localities (Sinai Peninsula, Funafuti Atoll, Maui and Hawaii Islands), with the description of one new genus and four new species. *Zoologische Verhandelingen*, *128*, 1–48.

Huson, D. H. (1998). SplitsTree: Analyzing and visualizing evolutionary data. *Bioinformatics (Oxford, England)*, *14*, 68–73. https://doi.org/10.1093/bioinformatics/14.1.68

Huson, D. H., & Bryant, D. (2006). Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution*, *23*, 254–267. https://doi.org/10.1093/molbev/msj030

Johannesson, K. (1988). The paradox of Rockall: Why is a brooding gastropod (*Littorina saxatilis*) more widespread than one having a planktonic larval dispersal stage (*L. littorea*)? *Marine Biology*, *99*, 507–513. https://doi.org/10.1007/BF00392558

Johnson, P. L. F., & Slatkin, M. (2008). Accounting for bias from sequencing error in population genetic estimates. *Molecular Biology and Evolution*, *25*, 199–206. https://doi.org/10.1093/molbev/msm239

Klei, K. A. B. (2016). *Holocene mollusk faunas in three marine lakes in Misool, Raja Ampat, Indonesia* MSc thesis, University of Amsterdam, 75pp.

Korneliussen, T. S., Albrechtsen, A., & Nielsen, R. (2014). ANGSD: Analysis of next generation sequencing data. *BMC Bioinformatics*, *15*, 356. https://doi.org/10.1186/s12859-014-0356-4

Lal, M. M., Southgate, P. C., Jerry, D. R., & Zenger, K. R. (2016). Fishing for divergence in a sea of connectivity : The utility of ddRADseq genotyping in a marine invertebrate, the black-lip pearl oyster. *Marine Genomics*, *25*, 57–68. https://doi.org/10.1016/j.margen.2015.10.010

Lee, C.-R., & Mitchell-Olds, T. (2011). Quantifying effects of environmental and geographical factors on patterns of genetic differentiation. *Molecular Ecology*, *20*, 4631–4642. https://doi.org/10.1111/j.1365-294X.2011.05310.x

Lefort, V., Desper, R., & Gascuel, O. (2015). FASTME 2.0: A comprehensive, accurate, and fast distance-based phylogeny inference program. *Molecular Biology and Evolution*, *32*, 2798–2800. https://doi.org/10.1093/molbev/msv150

Legendre, P. (1993). Spatial autocorrelation: Trouble or New Paradigm ? *Ecology*, *74*, 1659–1673. https://doi.org/10.2307/1939924

Legrand, D., Cote, J., Fronhofer, E. A., Holt, R. D., Ronce, O., Schtickzelle, N., … Clobert, J. (2017). Eco-evolutionary dynamics in fragmented landscapes. *Ecography*, *40*, 9–25. https://doi.org/10.1111/ecog.02537

Li, W., & Godzik, A. (2006). CD-HIT: A fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics*, *22*, 1658–1659. https://doi.org/10.1093/bioinformatics/btl158

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., … Durbin, R. (2009). The sequence alignment/map format and

SAMtools. *Bioinformatics*, *25*, 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Liu, X., & Fu, Y.-X. (2015). Exploring population size changes using SNP frequency spectra. *Nature Genetics*, *47*, 555–559. https://doi.org/10.1038/ng.3254

Lucek, K., Sivasundar, A., Roy, D., & Seehausen, O. (2013). Repeated and predictable patterns of ecotypic differentiation during a biological invasion: Lake-stream divergence in parapatric Swiss stickleback. *Journal of Evolutionary Biology*, *26*, 2691–2709. https://doi.org/10.1111/jeb.12267

Lynch, M. (2008). Estimation of nucleotide diversity, disequilibrium coefficients, and mutation rates from high-coverage genome-sequencing projects. *Molecular Biology and Evolution*, *25*, 2409–2419. https://doi.org/10.1093/molbev/msn185

MacArthur, R. H., & Wilson, E. O. (1967). *The theory of island biogeography*. Princeton, NJ: Princeton University Press.

Mangubhai, S., Erdmann, M. V., Wilson, J. R., Huffard, C. L., Ballamu, F., Hidayat, N. I., … Wen, W. (2012). Papuan Bird's Head Seascape: Emerging threats and challenges in the global center of marine biodiversity. *Marine Pollution Bulletin*, *64*, 2279–2295. https://doi.org/10.1016/j.marpolbul.2012.07.024

Mantel, N. (1967). The detection of disease clustering and a generalized regression approach. *Cancer Research*, *27*, 209–220.

Marques, D. A., Lucek, K., Meier, J. I., Mwaiko, S., Wagner, C. E., Excoffier, L., & Seehausen, O. (2016). Genomics of rapid incipient speciation in sympatric threespine stickleback. *PLoS Genetics*, *12*, 1–34. https://doi.org/10.1371/journal.pgen.1005887

Marshall, D. J., Monro, K., Bode, M., Keough, M. J., & Swearer, S. (2010). Phenotype-environment mismatches reduce connectivity in the sea. *Ecology Letters*, *13*, 128–140. https://doi.org/10.1111/j.1461-0248.2009.01408.x

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. Journal*, *17*, 10–12. https://doi.org/10.14806/ej.17.1.200

Mayr, E. (1954). *Change of genetic environment and evolution*. London, UK: Allen & Unwin.

Mayr, E. (1963). *Animal species and evolution*. Cambridge, MA: Harvard University Press. https://doi.org/10.4159/harvard.9780674865327

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., … DePristo, M. A. (2010). The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, *20*, 1297–1303. https://doi.org/10.1101/gr.107524.110

McRae, B. H. (2006). Isolation by resistance. *Evolution*, *60*, 1551–1561. https://doi.org/10.1111/j.0014-3820.2006.tb00500.x

Monteiro-Ribas, W., Rocha-Miranda, F., Romano, R. C., & Quintanilha, J. (2006). Larval development of *Brachidontes solisianus* (Bivalvia, Mytilidae), with notes on differences between its hinge system and that of the mollusk Perna perna. *Brazilian Journal of Biology*, *66*, 109–116. https://doi.org/10.1590/s1519-69842006000100014

Morton, B. (1988). The population dynamics and reproductive cycle of *Brachidontes variabilis* (Bivalvia: Mytilidae) in a Hong Kong mangrove. *Malacological Review*, *21*, 109–117.

Nei, M. (1987). *Molecular evolutionary genetics*. New York, NY: Columbia University Press.

Neves, J. M. M., Lima, S. M. Q., Mendes, L. F., Torres, R. A., Pereira, R. J., & Mott, T. (2016). Population Structure of the rockpool blenny entomacrodus vomerinus shows source-sink dynamics among ecoregions in the tropical Southwestern Atlantic. *PLoS ONE*, *11*, 1–15. https://doi.org/10.1371/journal.pone.0157472

Oksanen, J., Blanchet, F. G., Firendly, M., Kindt, R., Legendre, P., Minchin, P. R., … Wagner, H. (2016). VEGAN: Community ecology package.

Orsini, L., Vanoverbeke, J., Swillen, I., Mergeay, J., & de Meester, L. (2013). Drivers of population genetic differentiation in the wild: Isolation by dispersal limitation, isolation by adaptation and isolation by colonization. *Molecular Ecology*, *22*, 983–5999. https://doi.org/10.1111/mec.12561

Palumbi, S. R. (1994). Genetic divergence, reproductive isolation, and marine. *Annual Review of Ecology and Systematics*, *25*, 547–572. https://doi.org/10.1146/annurev.es.25.110194.002555

Pazmiño, D. A., Maes, G. E., Simpfendorfer, C. A., León, P. S. D., & Van Herwerden, L. (2017). Genome-wide SNPs reveal low effective population size within confined management units of the highly vagile Galapagos shark (*Carcharhinus galapagensis*). *Conservation Genetics*, *18*, 1151–1163. https://doi.org/10.1007/s10592-017-0967-1

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*, e37135. https://doi.org/10.1371/journal.pone.0037135

Pinheiro, H. T., Bernardi, G., Simon, T., Joyeux, J.-C., Macieira, R. M., Gasparini, J. L., … Rocha, L. A. (2017). Island biogeography of marine organisms. *Nature*, *549*, 82. https://doi.org/10.1038/nature23680

Rambaut, A. (2009). FigTree.

Rundle, H. D., & Nosil, P. (2005). Ecological speciation. *Ecology Letters*, *8*, 336–352. https://doi.org/10.1111/j.1461-0248.2004.00715.x

Salojärvi, J., Smolander, O. P., Nieminen, K., Rajaraman, S., Safronov, O., Safdari, P., … Rastas, P. (2017). Genome sequencing and population genomic analyses provide insights into the adaptive landscape of silver birch. *Nature Genetics*, *49*(6), 904.

Sanderson, M. J. (1997). A nonparametric approach to estimating divergence times in the absence of rate constancy. *Molecular Biology and Evolution*, *14*, 1218–1231. https://doi.org/10.1093/oxfordjournals.molbev.a025731

Sarà, G., Romano, C., Widdows, J., & Staff, F. J. (2008). Effect of salinity and temperature on feeding physiology and scope for growth of an invasive species (*Brachidontes pharaonis* - MOLLUSCA: BIVALVIA) within the Mediterranean sea. *Journal of Experimental Marine Biology and Ecology*, *363*, 130–136. https://doi.org/10.1016/j.jembe.2008.06.030

Sathiamurthy, E., & Voris, K. H. (2006). Maps of holocene sea level transgression and submerged lakes on the Sunda shelf. *Tropical Natural History*, *2*, 1–44.

Schluter, D. (2009). Evidence for ecological speciation and its alternative. *Science*, *323*, 737–741. https://doi.org/10.1126/science.1160006

Skotte, L., Korneliussen, T. S., & Albrechtsen, A. (2013). Estimating individual admixture proportions from next generation sequencing data. *Genetics*, *195*, 693–702. https://doi.org/10.1534/genetics.113.154138

Slatkin, M. (1985). Gene flow in natural populations. *Annual Reviews of Ecology and Systematics*, *16*, 393–430. https://doi.org/10.1146/annurev.es.16.110185.002141

Slatkin, M. (1993). Isolation by distance in equilibrium and non-equilibrium populations. *Evolution*, *47*, 264–279. https://doi.org/10.1111/j.1558-5646.1993.tb01215.x

Slatkin, M. (1995). A measure of population subdivision based on microsatellite allele frequencies. *Genetics*, *139*, 457–462.

Smit, A. F. A., Hubley, R., & Green, P. (2014). RepeatMasker Open-4.0. 2013–2015.

Spidle, A. P., May, B., & Mills, E. L. (1995). Limits to tolerance of temperature and salinity in the quagga mussel and the zebra mussel. *Canadian Journal of Fish and Aquatic Science*, *52*, 2108–2119. https://doi.org/10.1139/f95-804

Stamatakis, A. (2014). RAXML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, *30*, 1312–1313. https://doi.org/10.1093/bioinformatics/btu033

Starger, C. J., Van Nydeck Erdmann, M., Toha, A. H. A., Baker, A. C., & Barber, P. H. (2015). Strong genetic structure among coral populations within a conservation priority region, the Bird's Head Seascape (Papua and West Papua, Indonesia). *Frontiers of Biogeography*, *7*, 91–106. https://escholarship.org/uc/item/6ds1g7bt

Sundqvist, L., Keenan, K., Zackrisson, M., Prodohl, P., & Kleinhans, D. (2016). Directional genetic differentiation and relative migration. *Ecology and Evolution*, *6*, 3461–3475. https://doi.org/10.1002/ece3.2096

Swift, H. F., Gómez Daglio, L., & Dawson, M. N. (2016). Three routes to crypsis: Stasis, convergence, and parallelism in the Mastigias species complex (Scyphozoa, Rhizostomeae). *Molecular Phylogenetics and Evolution*, *99*, 103–115. https://doi.org/10.1016/j.ympev.2016.02.013

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, *123*, 585–595. https://doi.org/PMC1203831

Tariel, J., Longo, G. C., & Bernardi, G. (2016). Tempo and mode of speciation in Holocanthus angelfishes based on RADseq markers. *Molecular Phylogenetics and Evolution*, *98*, 84–88. https://doi.org/10.1016/j.ympev.2016.01.010

Terranova, M. S., Lo Brutto, S., Arculeo, M., & Mitton, J. B. (2007). A mitochondrial phylogeography of *Brachidontes variabilis* (Bivalvia: Mytilidae) reveals three cryptic species. *Journal of Zoological Systematics and Evolutionary Research*, *45*, 289–298. https://doi.org/10.1111/j.1439-0469.2007.00421.x

Timm, J., & Kochzius, M. (2008). Geological history and oceanography of the Indo-Malay Archipelago shape the genetic population structure in the false clown anemonefish (*Amphiprion ocellaris*). *Molecular Ecology*, *17*, 3999–4014. https://doi.org/10.1111/j.1365-294X.2008.03881.x

Tomascik, T., & Mah, A. J. (1994). The ecology of Halimeda Lagoon: An anchialine Lagoon of a raised atoll, Kakaban Island, East Kalimantan, Indonesia. *Tropical Biodiversity*, *2*, 385–399.

Treml, E. A., Roberts, J., Halpin, P. N., Possingham, H. P., & Riginos, C. (2015). The emergent geography of biophysical dispersal barriers across the Indo-West Pacific. *Diversity and Distributions*, *21*, 465–476. https://doi.org/10.1111/ddi.12307

Van Wyngaarden, M., Snelgrove, P. V. R., DiBacco, C., Hamilton, L. C., Rodríguez-Ezpeleta, N., Jeffery, N. W., … Bradbury, I. R. (2016). Identifying patterns of dispersal, connectivity, and selection in the sea scallop, *Placopecten magellanicus*, using RAD-seq derived SNPs. *Evolutionary Applications*, *10*, 1–16. https://doi.org/10.1111/eva.12432

Vellend, M. (2005). Species diversity and genetic diversity: Parallel processes and correlated patterns. *The American Naturalist*, *166*, 199–215. https://doi.org/10.1086/431318

Vieira, F. G., Fumagalli, M., & Albrechtsen, A. (2013). Estimating inbreeding coefficients from NGS data: Impact on genotype calling and allele frequency estimation. *Genome Research*, *23*, 1852–1861. https://doi.org/10.1101/gr.157388.113

Vieira, F. G., Lassalle, F., Korneliussen, T. S., & Fumagalli, M. (2016). Improving the estimation of genetic distances from Next-Generation Sequencing data. *Biological Journal of the Linnean Society*, *117*, 139–149. https://doi.org/10.1111/bij.12511

Wang, I. J., & Bradburd, G. S. (2014). Isolation by environment. *Molecular Ecology*, *23*, 5649–5662. https://doi.org/10.1111/mec.12938

Wang, I. J., & Summers, K. (2010). Genetic structure is correlated with phenotypic divergence rather than geographic isolation in the highly polymorphic strawberry poison-dart frog. *Molecular Ecology*, *19*, 447–458. https://doi.org/10.1111/j.1365-294X.2009.04465.x

Wang, H., Xu, X., Vieira, F. G., Xiao, Y., Li, Z., Wang, J., … Chu, C. (2016). The power of inbreeding: NGS-Based GWAS of rice reveals convergent evolution during rice domestication. *Molecular Plant*, *9*, 975–985. https://doi.org/10.1016/j.molp.2016.04.018

Warren, B. H., Simberloff, D., Ricklefs, R. E., Aguilée, R., Condamine, F. L., Gravel, D., … Thébaud, C. (2015). Islands as model systems in ecology and evolution: Prospects fifty years after MacArthur-Wilson. *Ecology Letters*, *18*, 200–217. https://doi.org/10.1111/ele.12398

Waters, J. M., Fraser, C. I., & Hewitt, G. M. (2013). Founder takes all: Density-dependent processes structure biodiversity. *Trends in Ecology and Evolution*, *28*, 78–85. https://doi.org/10.1016/j.tree.2012.08.024

Watterson, G. A. (1975). On the number of segregating sites in genetical models without recombination. *Theoretical Population Biology*, *7*, 256–276. https://doi.org/10.1016/0040-5809(75)90020-9

Weber, J. N., Bradburd, G. S., Stuart, Y. E., Stutz, W. E., & Bolnick, D. I. (2016). Partitioning the effects of isolation by distance, environment, and physical barriers on genomic divergence between parapatric threespine stickleback. *Evolution*, *71*, 1–15. https://doi.org/10.1111/evo.13110

Weese, D. A., Fujita, Y., Hidaka, M., & Santos, S. R. (2012). The long and short of it: Genetic variation and population structure of the Anchialine Atyid Shrimp Caridina rubella on Miyako-Jima, Japan. *Journal of Crustacean Biology*, *32*, 109–117. https://doi.org/10.1163/193724011x615389

Wright, S. (1943). Isolation by distance. *Genetics*, *28*, 114–138.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Maas DL, Prost S, Bi K, et al. Rapid divergence of mussel populations despite incomplete barriers to dispersal. *Mol Ecol*. 2018;27:1556–1571. https://doi.org/10.1111/mec.14556