

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/20260> holds various files of this Leiden University dissertation.

Author: Becking, Leontine Elisabeth

Title: Marine lakes of Indonesia

Date: 2012-12-04

**Are marine lakes cradles or refuges of diversity?
A mussel's (*Brachidontes* sp.) perspective**

Leontine E. Becking, Abram J. Knegt, Nicole J. de Voogd, Katja T.C.A. Peijnenburg

manuscript in preparation

Abstract

The marine lake setting, landlocked seawater with clearly delineated contours, provides a unique model to study early stages of evolution in coastal marine taxa. The phylogeography of the mussel *Brachidontes* sp. from marine lakes in Indonesia was studied by collecting and analyzing sequence data of the mitochondrial Cytochrome Oxidase I (COI) gene. In addition, we examined shell shape using a geometric morphometric approach. We found strong genetic and morphological differentiation between *Brachidontes* sp. populations in Indonesian marine lakes. The Indonesian populations of *Brachidontes* sp. harbored deeply diverged lineages (14.8% sequence divergence) that correspond to 'lineage A & B' which was previously recorded from marine lakes in Palau. The mussel populations in the three Indonesian marine lakes differed significantly in shell outline shape, concordant with three genetic lineages and these lineages may constitute a species complex of at least two undescribed species. Analysis of two nuclear genes (28S and 18S) confirmed the designation of the marine lake samples to the genus *Brachidontes* but did not allow species distinction within the genus. Each lake contained within lineage diversity that was unique to each lake and we suggest that this resulted from *in situ* divergence. Combined effects of stochastic processes (e.g. founder effects), local adaptation and increased evolutionary rates could produce high levels of differentiation in small populations such as in marine lake environments. The lakes appear to be cradles of diversity resulting from recent divergence of evolving populations within the lakes. The lakes also may serve as refuges for ancient lineages from the sea or older anchialine populations. The role of marine lakes in supporting endemism in the Indo-Australian Archipelago region may reflect enhanced survival of endemics, with the possibility of population differentiation that in time may lead to speciation.

Keywords

Phylogeography • Indonesia • Indo-Australian Archipelago • Raja Ampat • Berau • anchialine

Introduction

Marine lake populations provide an opportunity to study the early stages of evolution coastal marine taxa. Marine lakes are anchialine systems: small bodies of landlocked seawater isolated to varying degrees from the surrounding marine environment (Holthuis 1973, Hamner & Hamner 1998). A large number of marine lakes are located in the countries Indonesia and Palau (Dawson et al. 2009, Becking et al. 2011). Indonesia and Palau are situated within the world's largest concentration of marine biodiversity, an area known as the Indo-Australian Archipelago (IAA) (Roberts et al. 2002, Hoeksema 2007). Numerous factors have been proposed that may account for the high biodiversity within the IAA, including geological history of the area (Renema et al. 2008), its position downstream of the Pacific (Connolly et al. 2003, Kool et al. 2011), the large area of shallow water habitat during the Pleistocene low sea level stands (Voris 2000, Hoeksema 2007), great habitat heterogeneity (Hoeksema 2007), and large reef area (Bellwood et al. 2005). The IAA is also remarkable for the high number of endemics. Given the areal definition of an endemic as spatially restricted species (Begon et al. 1996), centers of endemism could be areas where species arise and remain, and/or the last stand of previously widespread species (Briggs 2000, Roberts et al. 2002, Meyer et al. 2005, Bellwood & Meyer 2009). Interestingly, for many taxa in the IAA the majority of endemics lie in the peripheral areas which are characterized by a high degree of isolation (Bellwood & Meyer 2009, Malay & Paulay 2010).

Small peripatric populations such as those in marine lakes could provide an opportunity to study marine taxa in isolated environments (Dawson & Hamner 2005). The majority of marine lakes are shallower than 50 m., which means that during the Last Glacial Maximum when sea levels are estimated to have been approximately 110-140 m. lower than modern sea levels (Geyh et al. 1979, Voris 2000) the lakes would have been dry or contained fresh water (Dawson 2006). Based on the estimated sea level rise, the presumed dates of filling of the lakes with seawater are estimated at 5000-15000 years before present (Dawson 2006, Sathiamurthy & Voris 2006). Hence, marine lakes are young environments, yet their biodiversity is distinct from the adjacent sea (Dawson & Hamner 2005, CHAPTER 1 & 3). Several descriptive studies of marine lake fauna suggest high endemism or an abundance of species rare elsewhere (Tomascik & Mah 1994, Dawson 2005, Dawson & Hamner 2005, Azzini et al. 2007, Colin 2009, Becking et al. 2011, CHAPTER 1). The unique diversity within the lakes could have two origins: (1) it has resulted from recent divergence of rapidly evolving populations isolated from their ancestral population in the sea or (2) or it is composed of ancient lineages, which are relicts of the sea or of earlier anchialine populations. We refer to these hypotheses as the 'cradle' and 'refuge' hypothesis as adapted from Briggs (2000), Bellwood et al. (2005), and Bellwood & Meyer (2009) and do not consider them mutually exclusive.

Phylogeographic studies of the jellyfish *Mastigias papua* (Dawson 2005, Dawson & Hamner 2005), the fish *Sphaeramia orbicularis* (Gotoh et al. 2009), and the mussel *Brachidontes* sp. (Goto et al. 2011) from marine lakes of the islands of Palau show extreme genetic isolation, low genetic diversity, and in the cases of *Mastigias papua* and *Brachidontes* sp. rapid morphological evolution. In Indonesia the only phylogeographic study of marine lakes fauna conducted to date was on the sponge species *Suberites diversicolor* (CHAPTER 6). This study revealed two deeply diverged lineages and suggests that within one lineage there may have been local diversification in the largest and most isolated marine lake in Indonesia (Kakaban lake in East Kalimantan). Here we expand on these results by studying the phylogeography of a co-distributed but unrelated taxon, the mussel *Brachidontes* sp.

Species of the genus *Brachidontes* Swainson, 1840 (Mollusca; Bivalvia; Mytilidae) are marine mussels, which attach themselves to substrate in and below intertidal areas and can form large mytilid beds (Terranova et al. 2007). *Brachidontes* are broadcast spawners with external fertilization and only disperse during their planktonic larval stage for a duration of up to four weeks (Reunov et al. 1999, Monteiro-Ribas et al. 2006, Terranova et al. 2006). *Brachidontes* larvae can live in the plankton for up to four weeks (Monteiro-Ribas et al. 2006), yet it is unknown whether they are able to survive in the subterranean channels connecting marine lakes to the surrounding sea. Worldwide 31 species of *Brachidontes* have been described, but the phylogenetic position of the different species within this genus remains unclear with several reports of the occurrence of cryptic species (Lee & Foighil 2004, 2005, Aguirre et al. 2006). There is an undescribed species of *Brachidontes* that inhabits many marine lakes and when present is generally dominant in terms of space occupation and biomass in the lakes (Tomascik & Mah 1994, Colin et al. 2009, Becking et al. 2011, CHAPTER 1). In contrast, this species is extremely rare in coastal habitats (non-marine lake) (Colin 2009, Goto et al. 2011). A previous study of the *Brachidontes* sp. from marine lakes in Palau reported two genetically distinct and morphologically differentiated lineages that probably represent different species (Goto et al. 2011). Moreover, the spatial genetic structure of *Brachidontes* sp. from Palau indicated that the majority of the marine lake populations were highly differentiated from each other, each containing private haplotypes (Goto et al. 2011).

To study the phylogeography of *Brachidontes* sp. and to compare with data from Palau we collected sequence data of the mitochondrial Cytochrome Oxidase I (COI) gene from four populations of mussels. This is an informative marker at the within as well as between species level (Lee & Foighil 2004, 2005, Aguirre et al. 2006). We collected additional sequences of two nuclear ribosomal genes (28S and 18S) to determine species level relationships within Mytilidae (Lee & Foighil 2004, Aguirre et al. 2006). In addition we examined shell shape using a geometric morphometric approach to determine morphological differentiation. If marine lakes are isolated environments we would expect to find genetic and/or morphological differentiation between the lakes. With this data we addressed the following questions:

1. Are the same lineages present in *Brachidontes* sp. populations in the marine lakes of both Indonesia and Palau?
2. Are the Indonesian marine lake *Brachidontes* sp. populations isolated?
3. Can marine lakes be considered cradles and/or refuges of diversity?

Table 1 Sample localities of six *Brachidontes* sp. populations from Berau (East Kalimantan) and Raja Ampat (West Papua) in Indonesia. Per locality is provided: the salinity, size, relative degree of connection to the adjacent sea, number of individuals analyzed for morphometric analyses (morph.), number of individuals sequenced for 28S, 18S, and COI (number of unique sequences are given in brackets). For the Female-type (F-type) COI sequences the haplotype diversity (h) +/- standard deviations, nucleotide diversity (π) +/- standard deviations, and Tajima's D neutrality index was calculated per population, significant values in bold with asterisk.

														F-type	F-type	F-type	F-type
code	location	region	coordinates	Salinity (ppt)	size lake (x10 ³ m ³)	connection	morph	28S	18S	COI	h	π (10-3)	Tajima's D				
lake																	
KKB	Kakaban lake	Berau	N 02° 08' 23.5" E 118° 30' 31.9"	23-24	4000	most isolated	67	2 (2)	2 (1)	13 (4)	0.4231 +/- 0.1645	1.858 +/- 1.490	-1.29575				
TBB	Tanah Bamban lake	Berau	N 02° 13' 50.0" E 118° 34' 50.7"	26	120	least isolated	18			5 (1)	0	0	0				
RAJ	Sauwandarck lake	Raja Ampat	S 0° 35' 19.6" E 130° 35' 48.8"	28-30	84	isolated	33	4 (3)	3 (1)	32 (8)	0.3952 +/- 0.1100	0.898 +/- 0.877	-2.0796*				
coast																	
GAM	Gam mangrove	Raja Ampat	S 00° 30' 23.8" E 130° 38' 58.9"	32-33	-	open	22	4 (2)	3 (1)	15 (4)	0.3714 +/- 0.1532	1.208 +/- 1.099	-1.91084*				
MGR1	Maratua mangrove	Berau	N 02° 12' 16.3" E 118° 35' 18.9"	33-34	-	open	10	-	-	-	-	-	-				
MGR2	Samama mangrove	Berau	N 02° 09' 03.6" E 118° 19' 31.6"	33-34	-	open	3	-	-	-	-	-	-				

Material & Methods

Sample collection

All mussels collected in this study belong to the genus *Brachidontes* (Rob Moolenbeek, personal communication), and their morphology corresponds to morphological characters of *Brachidontes* provided by Goto et al. (2011). Mussels were collected from three marine lakes in Indonesia (Figure 1, Table 1): Kakaban lake (KKB, collected in 2003 & 2008), Tanah Bamban lake (TBB, in 2003, collected by R. Moolenbeek) in the Berau region of East Kalimantan province, and Sauwandarek lake (RAJ, in 2009) in Raja Ampat region of West Papua province. Mussels were abundant in the lakes displaying dense mussel beds along the shorelines or covering mangrove roots. *Brachidontes* sp. is rarely found outside of lakes (Colin 2009, Goto et al. 2011, Becking pers. obs.), but after extensive searching we were able to find small populations in Papua in a mangrove swamp by Gam island (GAM, in 2007) approximately 6km distance from RAJ on Mansuar island, and in Kalimantan in the large lagoon of Maratua island (MGR1, in 2003, collected by R. Moolenbeek) approximately 500m North of TBB, and in a tidal creek of Samama island (MGR2, in 2003, collected by R. Moolenbeek) approximately 20 km West from Kakaban island. We assume that these samples are well connected with populations in the surrounding sea. For a full description of the lakes see Becking et al. (2011). These landlocked pools of water have a tidal regime which is typically delayed (ranging from 20 minutes to 4 hours) and dampened (ranging from 20 cm to 1.5 m) compared to the adjacent sea (Hamner & Hamner 1998, Becking et al. 2011, CHAPTER 1). Based on the level of tidal dampening and delay, marine lakes can be ranked by their degree of connectivity to the surrounding sea. KKB was the most isolated, followed by RAJ, and TBB had the highest connection to the sea of all three lakes.

Additional sequences of *Brachidontes* sp. from Palau were obtained from GenBank (accession numbers given in Figure 2). Palau is located at approximately 1000 km north from Raja Ampat in Papua. The sample locations are indicated in Table 1. For a full description of the Palauan marine lakes see Hamner & Hamner (1998) and Colin (2009). All specimens were conserved intact in 96% ethanol, except for the samples collected in 2003, which were initially stored in 70% ethanol and transferred to 96% ethanol in March 2011.

DNA extraction, gene amplification and sequencing

Mantle and gill tissue samples were taken from *Brachidontes* sp. and incubated at room temperature for 24 hours in a lysis buffer consisting of 250mM EDTA, 5% SDS, 50mM Tris (pH = 8) (Holland 1993). Total DNA was purified using DNeasy Blood and Tissue Kit (Qiagen) following the manufacturer's protocol. Cytochrome Oxidase I (COI) is an informative marker at the within as well as between species level (e.g. Lee & Foighil 2004, Aguirre et al. 2006). Partial COI was amplified using the universal COI primers from Folmer (1994) LCO1490-F: 5'-GGTCAACAAATCATAAAGATATTGG-3' and HCO2198-R: 5'-TAAACTTCAGGGTGACCAAAAAATCA-3'. PCR reactions were performed in 25µL volumes containing 10.75µL ddH₂O, 2µL 25mM MgCl₂, 5µL dNTP's (1mM each), 2.5µL 10x PCR Buffer (SpheroQ), 1µL BSA (10mg/mL), 0.3µL of both primers (10µM), 0.15µL Taq polymerase (5units/µL, home made following Pluthero (1993)) and 3µL template DNA. The PCR program consisted of an initial denaturation step of 94°C for 3m followed by 35 PCR cycles of 94°C for 45s, 45°C for 60s and 72°C for 120s, with a final extension step of 72°C for 10m. Quality of PCR products was assessed using gel electrophoresis on 1% agarose gels. If PCR product quality was insufficient, a new PCR reaction was performed in a 25µL volume containing PCR Beads (Illustra, GE Healthcare) using 21.6µL ddH₂O, 0.2µL of both primers (10µM) and 3µL template DNA, and an identical PCR cycling program.

To test whether highly divergent mitochondrial lineages were also differentiated at nuclear loci, a subset

of 10 samples were amplified for two nuclear genes, 28S and 18S ribosomal DNA. The C1-D3 region of the nuclear ribosomal 28S was amplified using primers C1-F: 5'-ACCCGCTGAATTTAAGCAT-3' (Dayrat et al. 2001) and D3-R (5'-GACGATCGATTTCACGTCA-3' (Vonneman et al. 2005). The partial nuclear ribosomal 18S was amplified using the primers A1-F: 5'-CTGGTTGATCCTGCCAGTCATATGC-3' and 1800-R: 5'-GATCCTTCGCGAGGTTCACCTACG-3' (Vonneman et al. 2005), and sequenced with the internal primers KP-F: 5'-TGGAGGGCAAGTCTGGTG-3' and KP-R :5'-TTCCCGTGTGAGTCAAATTAAG-3' (Peijnenburg pers. comm.). PCR amplification of both 28S and 18S was performed in 20µL volumes containing 8.8µL ddH₂O, 4µL 5x Phire buffer (Phire), 2µL dNTP's (1mM each), 1.4µL 100% DMSO, 0.2µL BSA (10mg/ml), 0.5µL of each primer, 0.2µL Hot Start Taq (Phire) and 2.4µL template DNA. PCR cycling steps consisted of an initial denaturation step of 98°C for 30s, followed by 35 cycles of 98°C for 5s, 48°C for 5s and 72°C for 20s, a final extension step of 72°C for 60s and a cooling step of 4°C for 180s. PCR products were purified and sequenced by Macrogen Inc (Korea and The Netherlands).

Doubly uniparental inheritance (DUI)

Some species of the family Mytilidae display a special type of mitochondrial inheritance called 'doubly uniparental inheritance' (DUI) (Zouros et al. 1994), in which two different mitochondrial lineages, a male (M-type) and female (F-type) lineage, are present in male individuals (Terranova et al. 2007, Goto et al. 2011). Male and female mitochondrial lineages can be highly diverged (Saavedra et al. 1997, Kenchington et al. 2002). F-type mtDNA is commonly used as a genetic marker to study the phylogeny of mussels, because it exists in both sexes, and reference sequences for *Brachidontes* mussels are available in GenBank. The distinction between F and M-type sequences in our data was made by constructing a phylogeny and subsequent intraspecific analyses were focused on F-type sequences only.

Genetic data analysis

The bivalve origin of the obtained sequences was verified through BLAST searches (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>). Sequences were aligned and handled in Geneious Pro 4.8.5 (Drummond et al., 2010) and DAMBE 5.2.15 (Xia & Xie 2001). Outgroups of other species from the genus *Brachidontes* were collected from GenBank. Alignments were collapsed to contain only unique sequence types in DAMBE. The best-fit DNA substitution model was selected by the Akaike Information Criterion deployed in jMODELTEST v. 0.1.1 (Posada 2008) and this model (COI: GTR+G+I, 28S:GTR+G+I, 18S: K2+G) was used for subsequent Bayesian and maximum likelihood phylogeny inferences. Phylogenetic reconstructions were performed under Bayesian inference criteria implemented in MrBayes v. 3.1.2. (Huelsenbeck & Ronquist 2001). Each analysis comprised two independent runs of four Metropolis-coupled Markov chains, sampled at every 1000th generation at the default temperature (0.2). Analyses were terminated after the chains converged significantly as indicated by an average standard deviation of split frequencies of <0.001. Convergence of chains was subsequently checked in Tracer v. 1.5.0 (Rambaut & Drummond 2007). For comparison, maximum likelihood bootstrap analyses were conducted using MEGA v. 5.01 (Tamura et al. 2011) using a heuristic search with 1000 bootstrap replicates. The Bayesian and maximum likelihood phylograms were combined and visualized using Treegraph 2 (Stöver & Müller 2010). Intra- and interlineage uncorrected *p*-distances were calculated in MEGA v. 5.01. To assess genetic diversity within populations and to detect deviations of neutrality, haplotype diversity (*h*) and nucleotide diversity (π) were calculated and Tajima's D neutrality test (Tajima 1989) was performed using Arlequin v. 3.5.1.2 (Excoffier & Lischer 2010). Haplotype networks were computed to obtain graphical representations of evolutionary relationships between haplotypes using the minimum spanning network algorithm of Rohlf (1973) as and visualized using FigTree v1.3.1 and Adobe®

Illustrator© CS5. To test for spatial structuring of samples we performed an analysis of molecular variance (AMOVA) and calculated pairwise Φ_{st} values (based on uncorrected p -distances) between populations using Arlequin v. 3.5.1.2. Significance of pairwise Φ_{st} values was determined by 10000 permutations and exact tests of population differentiation in Arlequin.

Morphometric analysis

Brachidontes sp. mussels were photographed in a standardized orientation (Fig. 4) for geometric morphometric analyses. In total 153 digital images were stored as Nikon RAW format (.nef) and converted to 3008 x 2000 pixel JPEG images using Photoshop 5.0 (Adobe). JPEG images were sampled into TPS files using *tpsUtil* (Rohlf 2010a). Shell outlines were used to capture variation in shell shape of *Brachidontes* sp. We used a sliding semi-landmark analysis, in which semi-landmarks are allowed to slide along the outline of a shell in order to find the position that optimally matches the positions of corresponding semi-landmarks in a reference specimen, usually a consensus specimen (Bookstein 1997, Adams et al. 2004). Shell outlines were drawn as curves and digitized as 39 semi-landmarks at equal distance using *tpsDig* (Rohlf 2010b), using the beak of the mussel (umbo, see Fig. 4) as a standardized starting point for drawing an outline. A “sliders file” indicating sliding semi-landmarks was made using *tpsUtil* (Rohlf 2010a). To standardize for size and orientation we used *tpsRelw* (Rohlf 2010c) with Generalized Procrustes Superimposition (Rohlf 1999). Residuals from the superimposition were analysed with the thin-plate spline interpolating function, producing principal warps, followed by relative warp (RW) analysis. *TpsRelw* was used to obtain centroid size (Bookstein 1991) and RW scores for each individual. RW axes are analogous to the eigenvectors of principal component analysis, which combine the major patterns of shell shape variation in the data. Repeatability of RW axes was tested using regression analysis and a non-parametric analysis of similarity in PAST 2.11 (Hammer et al. 2001) of RW scores extracted from 16 specimens of *Brachidontes* sp., which were independently photographed by the same observer (A. Knecht). RW axes were considered repeatable when they showed a non-significant and close to zero R-value in the analysis of similarity and a strong ($r > 0.7$) and significant correlation. Only repeatable relative warp axes were included in further analyses of shell shape variation. Correlations of RW scores were tested with centroid size and if significant and strong ($r > 0.7$), we used residuals of the regression to correct for size in all further analyses of shell shape. Significant differentiation between populations was tested using a non-parametric analysis of similarity (ANOSIM, 10000 randomizations; Clarke 1993) based on Euclidian distance as implemented in PAST 2.11.

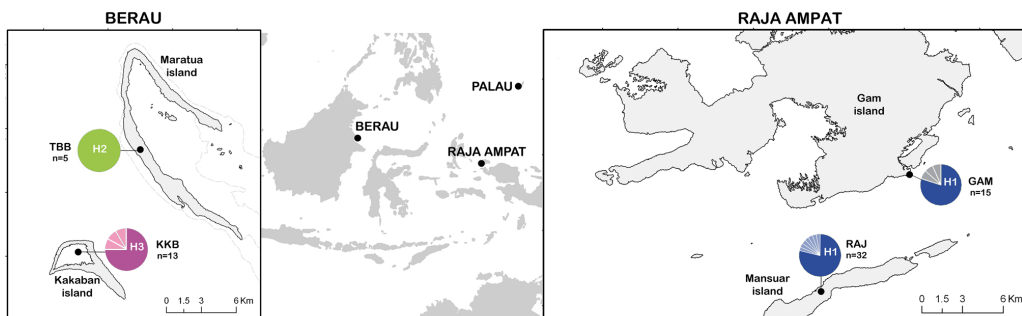


Figure 1 Sample localities of *Brachidontes* sp. populations. Left inset shows map of Berau, East Kalimantan province and right inset a map of Raja Ampat, Papua province (Indonesia). Haplotype frequencies of mitochondrial Cytochrome Oxidase I (COI) Female-type sequences given as pie charts per population of the mussel *Brachidontes* sp. in Kakaban lake (KKB), Tanah Bamban lake (TBB), Sauwandarek lake (RAJ) and Gam mangrove (GAM). Colors represent different haplotypes as shown in Figure 2. Haplotype H1, H2 & H3 refer to the three abundant haplotypes from the three lineage A, A2 and B respectively, as indicated in Figure 2.

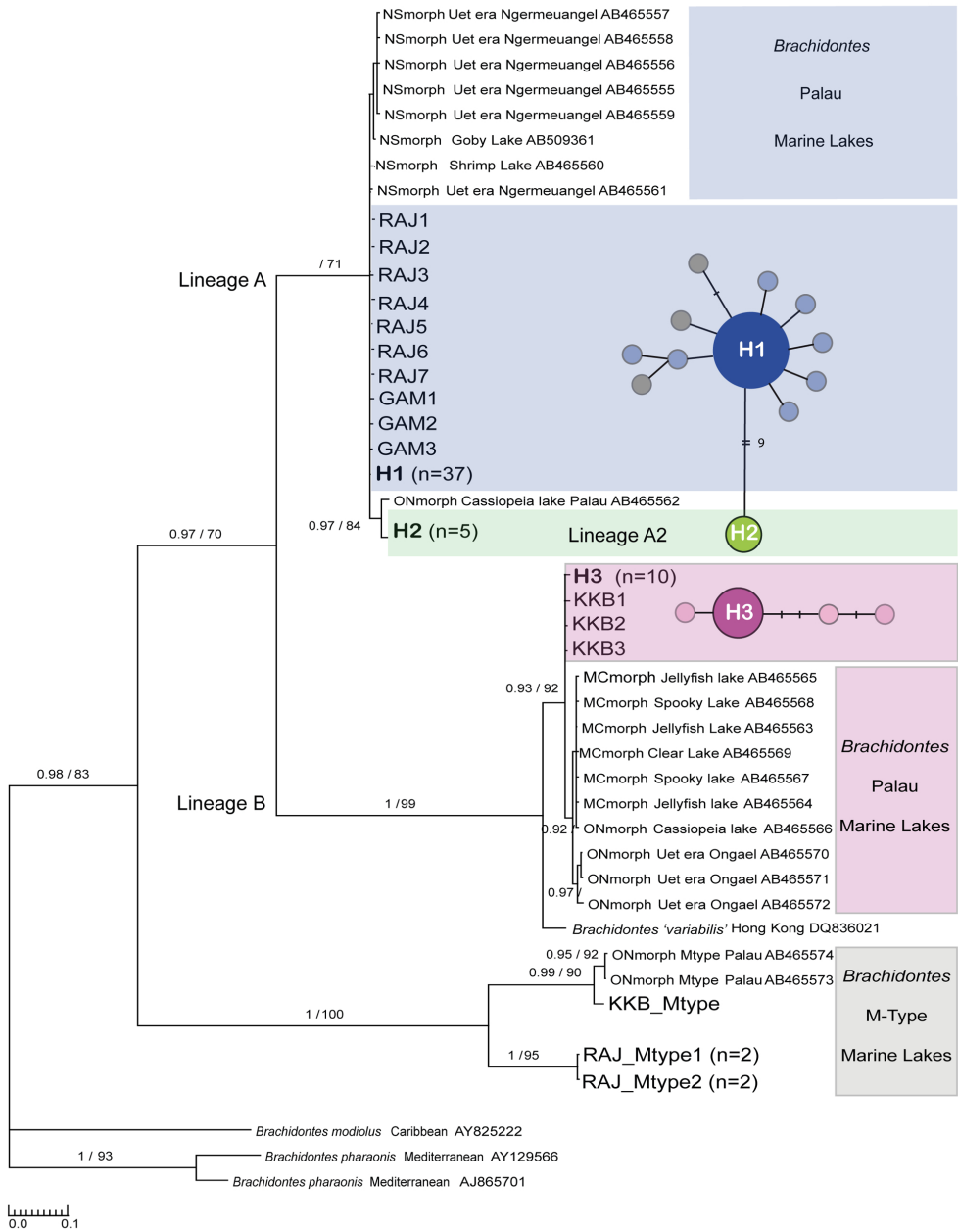


Figure 2 Bayesian/ maximum likelihood phylogram of 91 partial Cytochrome Oxidase I sequences of *Brachidontes* sp. sampled from Indonesia and Palau (see Fig. 1). The sequences from the present study from Indonesia are indicated by location codes (see Table 1), number of shared haplotypes indicated in brackets, no number in brackets indicates a unique haplotype. Sequences of *Brachidontes* from Palau marine lakes are from the study Goto et al. (2011) and are represented with the morphological type, lake name and Genbank accession number. The outgroup is presented with species name, sample location, and GenBank accession numbers. Only posterior probabilities of >90% and maximum likelihood support of >70% are indicated (see Material & Methods for more details). Scale bars indicate substitutions/site. To the right of the phylogram, minimum spanning networks are shown of haplotypes found in the Indonesian marine lakes of the present study. Each circle represents a single haplotype and its diameter is approximately proportional to the number of individuals carrying that haplotype, with the smallest circle representing a single individual. Lines connecting haplotypes represent one base substitution between two haplotypes with additional mutational steps indicated by hash marks and a number if more than three substitutions.

Results

Sequence diversity

We were able to amplify the genetic markers in the populations from Kakaban lake (KKB), Tanah Baman lake (TBB), Sauwandarek lake (RAJ) and Gam mangroves (GAM) to amplify the genetic markers, but in the other populations (MGR 1 & 2) the DNA-yield was too low to allow amplifications (Table 1). A total of 70 COI sequences of 552 bp were collected, resulting in 65 F-type sequences (16 haplotypes) and five M-type sequences (three haplotypes). Three abundant F-type haplotypes were collected: H1 in RAJ and GAM (n=37), H2 in TBB (n=5) and H3 in KKB (n=10). The remainder of the F-type haplotypes (n=13) were unique. No stop codons, indels or double peaks were observed in the F-type sequences. M-type sequences were designated based on a close similarity to the ON-morph M-type sequences from Palau (Figure 2). All M-type sequences revealed heteroplasmic “double peaks” at a small number of nucleotide positions in their sequence chromatograms. These double peaks are unlikely the result of amplification of both F- and M-type sequences in male individuals (resulting from DUI, explained in M&M) because both sequences were very distant from all sampled F-type sequences and were closely related to published M-type sequences. The heteroplasmic sequences did not contain any stopcodons, refuting the possibility of co-amplification of a pseudogene (Bensasson *et al.* 2001). All further phylogeographic analyses were focused on F type sequences only.

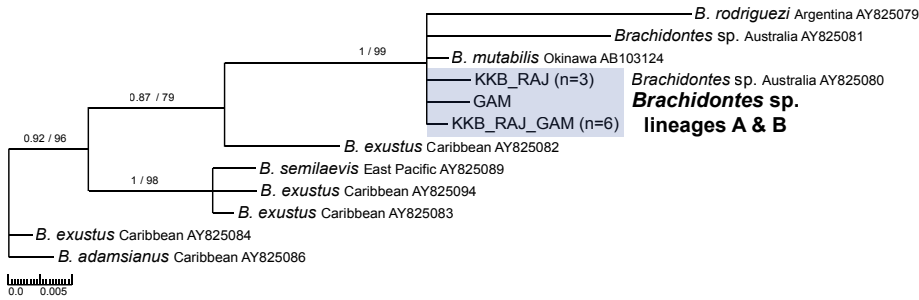
These double peaks were not a result of DUI (see Material and Methods) as the sequences were very distant from all sampled F-type sequences and M-type sequences could not be assembled to reference F-type sequences. The heteroplasmic sequences did not contain any stopcodons, refuting the possibility of a pseudogene (Bensasson *et al.* 2001). All further phylogeographic analyses was focused on F type sequences only. A subset of ten individuals from the four Indonesian populations were sequenced for 28S resulting in 636 bp length fragments and three gene variants (Figure 3A). For 18S the same subset of samples as 28S was amplified, but resulted in only eight successful sequences of 326 bp and two genetic variants (Figure 3B). Amplifications of 28S and 18S of samples from TBB samples were unsuccessful.

Phylogeography (COI)

The COI sequences represented two major lineages that were strongly supported by both Bayesian and maximum likelihood analyses: lineage A was represented by the two abundant haplotypes H1 and H2, and the single haplotypes GAM1-3 and RAJ1-7, while lineage B comprised the abundant haplotype H3 and the single haplotypes KKB1-3 (Figure 1 and 2). The two highly divergent lineages (p -distance is 14.8%, Table 3) correspond to ‘lineage A’ and ‘lineage B’ from the marine lakes in Palau (Goto *et al.* 2011) which naming we adopt. In the present study lineage A was present in TBB, RAJ and GAM, while lineage B was present in KKB only. No haplotypes were shared between Indonesia and Palau, but comparison of GAM and RAJ populations with lineage A from Palau shows that the two regions were closely related with a between group p -distance of 0.84% (Table 3). Similarly, COI haplotypes from the KKB population in Indonesia were closely related but not identical to lineage B haplotypes from Palau with a between group p -distance of 1.43% (Table 3). In addition, mussels from TBB constituted a highly supported lineage A2 within lineage A with a between-group p -distance of 1.8 % between TBB and RAJ & GAM (Figure 2 and Table 3). The samples from TBB were closely related to the Palauan ON-morph *Brachidontes* AB465562 from Northern Cassiopea Lake (Fig. 2; Goto *et al.* 2011).

Lineage A & B were not distinguished by the 28S and 18S sequences (Figure 3A&B). The phylogram of 28S

A.



B.

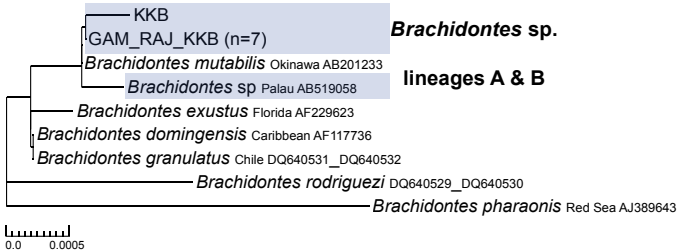


Figure 3 Bayesian/maximum likelihood phylograms of nuclear markers of *Brachidontes* sp. found in this study and related species from the same genus. **A.** partial 28S rDNA genes. **B.** partial 18S rDNA genes. Samples of the present study indicated by location codes (see Table 1) and separated by underscores followed by the number of individuals in brackets. For samples downloaded from GenBank species names are provided, followed by the sample location and GenBank accession number. *Brachidontes* sp. lineage A and B indicated by blue boxes. Only posterior probabilities of >90 and maximum likelihood values of >70 indicated. The 18S resulted in an unresolved tree. Scale bars indicate substitutions/site.

gene sequences (Figure 3A) shows a strongly supported polytomy including the Indonesian *Brachidontes* as well as the species *B. mutabilis* (collected in Okinawa), *Brachidontes* sp. (Darwin, Australia), and *B. rodriguezi* (Argentina). One 28S sequence representing both lineages A and B from the present study was identical with a *Brachidontes* sp. 28S sequence from Darwin, Australia (Genbank accession AY825080). The 18S sequences resulted in an unsupported phylogeny (both Bayesian and Maximum Likelihood), Fig. 3B. Our results of 28S and 18S confirmed the designation of the marine lake samples to the genus *Brachidontes* but did not allow species distinction within the genus (Lee & Foighil 2004, 2005, Aguirre et al. 2006, Santaclara et al. 2006). Genetic diversity indices based on F-type COI sequences are shown in Table 1. Haplotype diversity was highest in KKB (0.4231 +/- 0.1645) and lowest in TBB (0). Nucleotide diversity was highest in KKB (1.858 +/- 1.490) and lowest in TBB (0). Neutrality index Tajima's D was negative and significant for the populations of GAM and RAJ, suggesting deviations from neutrality in these samples (Table 1). All COI haplotypes found in KKB and TBB were unique for that location, whereas RAJ and GAM shared the COI haplotype H1 (Figure 1). The analysis of molecular variance (AMOVA) showed that there was a significant spatial genetic structure within the samples. All populations were strongly and significantly differentiated (Φ_{st} ranged from 0.95 to 0.99), except RAJ and GAM that were not significantly differentiated (Table 2).

Table 2 Pairwise Φ_{st} based on partial Cytochrome Oxidase I Female-type sequences between populations of *Brachidontes* sp. in Kakaban lake (KKB), Tanah Bamban lake (TBB), Sauwandarek lake (RAJ) and Gam mangrove (GAM). Significant Φ_{st} indicated in bold with asterisk ($p < 0.05$).

Φ_{st}	KKB	TBB	RAJ
TBB	0.99118*		
RAJ	0.99214*	0.95454*	
GAM	0.98985*	0.94805*	-0.00591

Table 3 Between- and within-group uncorrected p -distances based on partial Cytochrome Oxidase I sequences of *Brachidontes* sp. from Indonesia and Palau provided below the diagonal, standard deviations provided in italics above the diagonal. Black curvise along the diagonal indicates within-group uncorrected p -distance. For location codes refer to Table 1. Lineage A and B refer to the two major lineages obtained in Palau by Goto (2011), 'Outgroup' to *B.pharaonis* (GenBank accession numbers AY129566 and AJ865701) and *B.modiolus* (AY825222).

%	lineage A	lineage B	KKB	TBB	RAJ_GAM	OUTGROUP
lineage A	0.80	<i>1.43</i>	<i>1.42</i>	<i>0.45</i>	<i>0.25</i>	<i>1.29</i>
lineage B	<i>14.93</i>	0.59	<i>0.40</i>	<i>1.44</i>	<i>1.43</i>	<i>1.40</i>
KKB	<i>15.11</i>	<i>1.43</i>	0.51	<i>1.43</i>	<i>1.42</i>	<i>1.39</i>
TBB	<i>1.75</i>	<i>14.80</i>	<i>14.81</i>	0.00	<i>0.49</i>	<i>1.29</i>
RAJ_GAM	<i>0.84</i>	<i>14.85</i>	<i>14.86</i>	<i>1.83</i>	0.38	<i>1.29</i>
OUTGROUP	<i>17.29</i>	<i>18.79</i>	<i>18.66</i>	<i>17.10</i>	<i>17.10</i>	14.27

Morphometric analysis

The relative warp axes 1-4 were significantly repeatable, explaining 92.80% of total observed variation in shell shape. The population sampled at MGR2 was excluded from statistical analyses because its sample size was too small ($n < 5$). Variation in shell outline was found in shell length to width ratio and the position of the umbo relative to the longitudinal axis of the shell. A scatter plot was made for relative warp axes 1 and 2, showing how 85.11% of total observed variation in shell outline is distributed between sample locations and along these axes of shell outline variation (Fig. 5). RW1 represents variation in length to width ratio, whereas relative warp axis 2 incorporates the position of the umbo relative to the longitudinal axis of the shell. This suggests that individuals from RAJ and GAM have a larger length to width ratio than individuals from KKB, and individuals from MGR1 have an umbo that is positioned farther away under the longitudinal axis of the shell than individuals from other populations. Overall shell shape variation among populations was highly significant (ANOSIM, global $R = 0.3704$, $P = 0.0001$) and all pair wise comparisons were significant except between RAJ and GAM (Table 4).

Table 4 Analysis of similarity for *Brachidontes* sp. based on relative warp axes 1-4. R-values are shown below the diagonal, sequential Bonferroni corrected P -values are shown above the diagonal. Bold R-values with asterisk indicate significant values ($p < 0.05$). The population sampled at MGR2 was excluded from statistical analyses because its sample size was too small ($n < 5$).

	KKB	TBB	RAJ	GAM	MGR1
Sample size	67	18	33	22	10
KKB		0.0121	0.0001	0.0001	0.0001
TBB	0.1646*		0.0001	0.0001	0.0003
RAJ	0.392*	0.5065*		-0.03051	0.0001
GAM	0.2774*	0.5002*	0.7815		0.0001
MGR1	0.5144*	0.3608*	0.5865*	0.6728*	

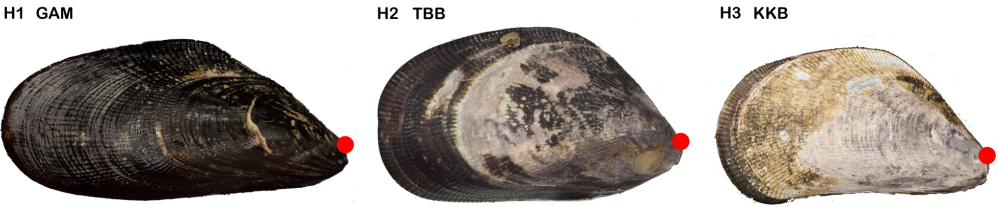


Figure 4 Standardized orientation the mussel *Brachidontes* sp. for the morphometric analyses. Morphology was characterized by a curve capturing shell outline, consisting of 39 semi-landmarks at equal distance starting from the umbo, indicated by red dot. Representatives are shown of the three dominant haplotypes (see Figure 2) from the localities Gam mangrove (GAM), Tanah Bamban lake (TBB), and Kakaban lake (KKB).

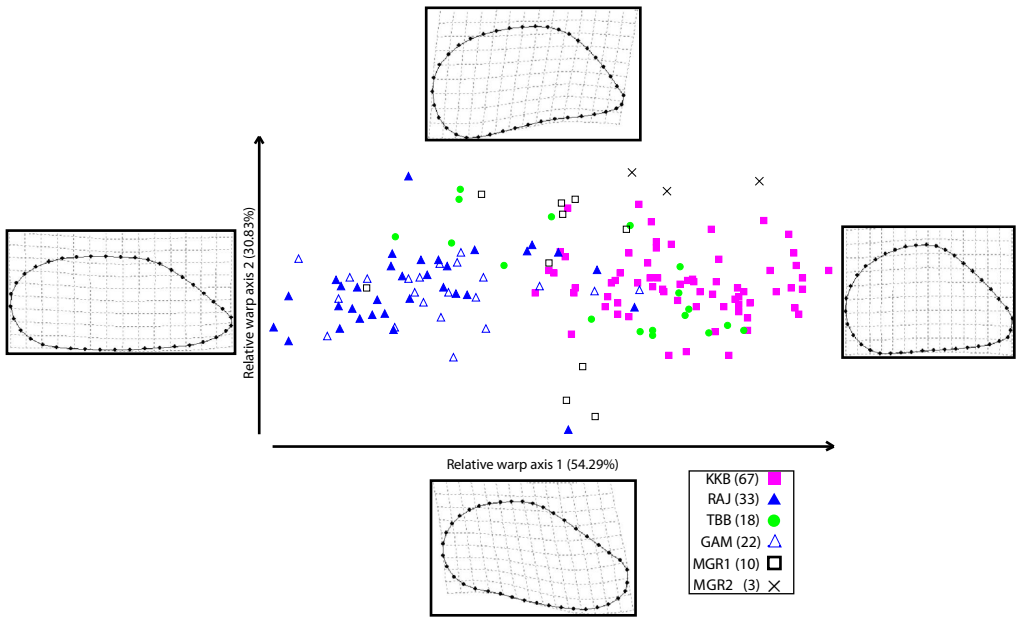


Figure 5 Shell outline shape variation in six populations of *Brachidontes* sp. mussels from Indonesia (sample codes are detailed in Table 1, sample size in brackets). Ordination of data in the plane identified by relative warp (RW) axis 1 and 2, together explaining 85.11% of the variance. Corresponding thin-plate splines of the most positive and negative deformations along the axes are shown. Both axes have been corrected for their correlation with centroid size.

Discussion

Divergent Brachidontes lineages in Indonesia and Palau

We detected three mitochondrial Female-type (F-type) lineages based on COI sequences, viz. lineage A, A2, and B, in *Brachidontes* sp. mussels sampled in marine lakes and adjacent sea habitats in Indonesia. The two major lineages are closely related to the lineages A and B reported by Goto et al. (2011) for *Brachidontes* sp. from marine lakes in Palau. Considering the degree of genetic divergence (almost 15%), lineages A and B of *Brachidontes* sp. in the marine lakes of Indonesia and Palau may constitute a species complex of at least two undescribed species. Furthermore, it is possible that the lineage A2 represents an incipient species that will continue to diverge if the populations remain isolated. The genetic distances between lineage A, A2, and B were comparable to those reported between (cryptic) species of other mussels (Lee and Foighil 2004, 2005, Terranova et al. 2007). Unfortunately, the two nuclear markers we used (18S and 28S) could not resolve species level relationships, probably because these markers were too conserved (Lee & Foighil 2004, 2005, Santaclara et al. 2006, Goto et al. 2011).

Patterns of COI genetic differentiation of the *Brachidontes* sp. populations were strikingly congruent with morphological differentiation. The mussel populations in the three Indonesian marine lakes differed significantly in shell outline shape, concordant with the three genetic lineages. Our results are supported by Wesselingh et al. (*in review*) who concluded that the *Brachidontes* sp. in Kakaban lake probably represents an as yet undescribed species as it could not be attributed to any known species from the region. Morphology can provide a proxy for underlying genetic variation and adaptive evolution (Dawson 2005, Dawson & Hamner 2005, Mariani et al. 2011). The morphometric differentiation we found is probably partly genetic but also may represent ecophenotypic plasticity related to different environments which is common in mussels (Aguirre et al. 2006, Groenenberg et al. 2011, Mariani et al. 2011). Each of the three lakes of the present study have a different environmental regime (Table 1; Becking et al. 2011, CHAPTER 1) which may also cause the differences in shell shape that we observe (Groenenberg et al. 2011).

In a study of the population structure of a co-distributed sponge species in marine lakes in Indonesia, Becking et al. (CHAPTER 6) found a highly divergent lineage in Kakaban lake that was not observed in any of the other populations. By combining our data with Goto et al. (2011), we find that unique *Brachidontes* sp. lineages in specific marine lakes in Indonesia are in fact, also found in other geographically very distant regions in the IAA. This finding stresses the importance of extensive sampling throughout the Indo-Pacific in order to establish true endemism.

Differentiation of populations in Indonesian marine lakes

We found strong genetic differentiation between *Brachidontes* sp. populations in Indonesian marine lakes that was congruent with differences in shell outline shape. Not one haplotype was shared between the different marine lakes populations of *Brachidontes* sp. from Indonesia and Palau, suggesting that dispersal between marine lakes is highly limited. If gene flow would be common between marine lakes and adjacent sea populations, we would have expected to see haplotype sharing between populations, particularly between geographically close lakes. Dawson & Hamner (2005) found that the genetic distance between marine lake and lagoon populations of the jellyfish *Mastigias papua* was highly correlated with the degree of physical isolation of the lake to the adjacent sea, and not the actual geographic proximity of populations to each other.

Our data shows that the largest and most isolated lake, Kakaban lake in Berau, harbors the highest nucleotide diversity of all sampled *Brachidontes* populations in Indonesia and Palau (Goto et al. 2011). Kakaban lake also harbors the highest genetic diversity of all populations of the co-distributed sponge *Suberites diversicolor* sampled from seven marine lakes and three coastal locations in Indonesia, Singapore and Australia (Becking et al. CHAPTER 6). Isolation acts to decrease the rate of immigration (MacArthur & Wilson 1967, Whittaker & Fernandez-Palacios 2007, Chen & He 2009, Chen et al. 2011, Rosindell et al. 2011), but isolation can also enhance diversification as diminished gene flow allows populations to diverge and ultimately form new species (Emerson & Gillespie 2008, Chen & He 2009, Rosindell et al. 2011). The degree of physical connection of a lake to the adjacent sea may, however, not fully dictate the genetic structure and diversity within a lake. Sauwandarek lake (RAJ) in Raja Ampat has a low physical connection with the surrounding sea, yet is not differentiated from the adjacent mangrove population (GAM). The lack of strong population differentiation between RAJ lake and the coastal mangrove may be caused by recurrent (recent and historic) gene flow. However, they do not share any of the unique haplotypes, which is suggestive that these populations are in fact isolated. The two populations may be too recently diverged to show strong differentiation in the molecular markers that we used.

Isolation of marine lake populations may be the result of strong isolating barriers and/or different selective regimes in the lakes. The subterranean channels that connect each lake with the surrounding sea may provide a formidable dispersal barrier for propagules. Alternatively, a propagule may be able to enter but may not be able to survive due to the environmental regime within the lake or competition with resident founder lineages/species. The pattern that is almost consistently seen in a variety of taxa (jellyfish, fish, and bivalves) is that each lake harbors a single lineage per taxon (Dawson 2005, Dawson & Hamner 2005, Gotoh et al. 2009, Goto et al. 2011). An explanation for this pattern may be that the first colonizers enter, proliferate and out-compete any subsequent migrating newcomers. This hypothesis is supported by the observation that lakes either contain mussels or oyster as the dominant bivalves, rarely both (Becking et al. 2011). The strong morphological differentiation in shell shape suggests that different lake habitats provide different selective regimes, as is also illustrated by differentiated species assemblages (Becking et al. 2011, Becking et al. *unpublished data*, CHAPTERS 1, 2, 3, 4 & 5).

Marine lakes: cradles or refuges of diversity?

The marine lakes of Indonesia and Palau are situated in the Indo-Australian Archipelago (IAA) of highest marine biodiversity, including a high number of endemics (Hoeksema 2007). The high biodiversity in this hotspot probably resulted from current ecological and environmental conditions as well as the biogeographical history of the area (e.g. Bellwood et al. 2005, Hoeksema 2007). The processes that have lead to this pattern of biodiversity remain much debated. In the 'cradle hypothesis' allopatric speciation in isolated marine environments during the Pleistocene low sea level stands is believed to have contributed to the current marine biodiversity in the IAA (e.g. Briggs 2000, 2005), while the 'refuge hypothesis' contends that isolated marine environments during Pleistocene may have provided refuge for marine shallow-water species that went extinct elsewhere (e.g. Bellwood et al. 2005, Renema et al. 2008, Bellwood & Meyer 2009). Recent phylogenetic studies strongly suggest that endemics in the IAA are not exceptionally young – most endemics having origins in the early Pliocene-Miocene age (Renema et al. 2008, Cowman & Bellwood 2011). Though there are still few examples in favor of the 'cradle hypothesis' (e.g. Barber et al. 2006), most studies have concluded that different and possibly combined processes will be relevant for different taxa (e.g. Barber & Bellwood 2005, Carpenter et al. 2011, Cowman & Bellwood 2011).

Both hypotheses are supported when taking the marine lakes as small natural laboratories for the scenario of the Pleistocene low sea level stands. We assume that the floodwaters that filled the lakes during the Holocene sea level rise (<15000 years before present) allowed for independent, chance colonization of lakes by propagules from the surrounding sea and that these propagules were the progenitors of the present day populations (Dawson & Hamner 2005, Dawson 2006). The deep divergences that are observed in the mussel populations are probably ancient lineages that have taken refuge there. It is unlikely that these have diverged within the timeframe of the lake because if this were the case we would have unrealistically high evolutionary rates of more than 100% COI sequence divergence per million years. Our results show that no marine lake shares any of the COI haplotypes with any other lake. With similar patterns of genetic variation in COI for the jellyfish *Mastigias papua* populations in marine lakes in Palau, Dawson & Hamner (2005) conclude that the pattern resulted from random redistribution of haplotypes that existed in the sea during the formation of each lake. It is, however, remarkable that of the 35 haplotypes of *Brachidontes* sp. that have been uncovered in Palau and Indonesia, almost 90% of the haplotypes are not shared among the 12 lakes whereas there is extensive haplotype sharing within each of the lakes. Therefore we suggest an alternative hypothesis, namely that the within lineage diversity that is unique to each lake resulted from *in situ* divergence. This scenario would also result in relatively rapid evolutionary rates, but such rates are not uncommon for recently diverged taxa (e.g. Genner et al. 2007, Ho et al. 2011). For example, if we take the within *p*-distances of the Kakaban lake population (0.51%) and the maximum age that the lake could be (15000 years), this would result in a mutation rate of 35% per million years for *Brachidontes* sp. This is higher than the evolutionary rate that has previously been estimated at 18.3-24.4% per million years for the F-type COI third codon positions in *Brachidontes* spp. (Lee & Ó Foighil 2004). Combined effects of stochastic processes (e.g. founder effects), local adaptation and increased evolutionary rates could produce high levels of differentiation in small populations such as in marine lake environments (Dawson & Hamner 2005, Ho et al. 2011). The patterns of genetic variation found so far in marine lake populations of *Mastigias papua* (Dawson & Hamner 2005), *Brachidontes* sp. (Gotoh et al. 2011, this study), *Suberites diversicolor* (Becking et al. CHAPTER 6), *Sphaeramia orbicularis* (Gotoh et al 2009) are generally consistent with taxa evolving in isolation in peripatric environments, such as islands or satellite lakes of ancient rift lakes in Africa (e.g. Genner et al. 2007, Emerson & Gillespie 2008, Chen & He 2009, Rosindell et al. 2011). The role of marine lakes in supporting endemism in the IAA region may reflect enhanced survival of endemics, with the possibility of population differentiation that in time may lead to speciation.

Acknowledgements

We would like to thank R. Moolenbeek for identification of the samples. The following people provided help with logistics for field- and/or labwork: B. Hoeksema, Suharsono, Y. Tuti, Bahrudin, M. Ammer, C. Huffard, M. Erdmann, S. Menken, H. Breeuwer, L. Dong, B. Voetdijk, P. Kuperus, and the staff of Papua Diving, of Nabucco Island Dive Resort, and of Derawan Dive Resort. The present study is part of a PhD-project of LEB funded by the Netherlands Organisation for Scientific Research (NWO), division Earth and Life Sciences (ALW IPJ-07002; # 817.01.008). KTCAP is supported by a VENI grant from NWO (863.08.024). Fieldwork in Indonesia was made possible through additional financial support of the Schure-Beijerinck-Popping Fund of the Royal Dutch Academy of Science (KNAW), 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 Biodiversity Center). We are grateful to the Indonesian Institute of Sciences (LIPI) and the Indonesian State Ministry of Research and Technology (RISTEK) for providing research permits in Indonesia. E. Gittenberger and W. Renema provided valuable comments on the original manuscript.

