Phylogeography of the sponge Suberites diversicolor in Indonesia: insights into the evolution of marine lake populations.
Becking, L.E.; Erpenbeck, D.; Peijnenburg, K.T.C.A.; de Voogd, N.J.

Published in:
PLoS One

DOI:
10.1371/journal.pone.0075996

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (http://dare.uva.nl)
Phylogeography of the Sponge *Suberites diversicolor* in Indonesia: Insights into the Evolution of Marine Lake Populations

Leontine E. Becking, Dirk Erpenbeck, Katja T. C. A. Peijnenburg, Nicole J. de Voogd

Abstract

The existence of multiple independently derived populations in landlocked marine lakes provides an opportunity for fundamental research into the role of isolation in population divergence and speciation in marine taxa. Marine lakes are landlocked water bodies that maintain a marine character through narrow submarine connections to the sea and could be regarded as the marine equivalents of terrestrial islands. The sponge *Suberites diversicolor* (Porifera: Demospongiae: Suberitidae) is typical of marine lake habitats in the Indo-Australian Archipelago. Four molecular markers (two mitochondrial and two nuclear) were employed to study genetic structure of populations within and between marine lakes in Indonesia and three coastal locations in Indonesia, Singapore and Australia. Within populations of *S. diversicolor* two strongly divergent lineages (A & B) (COI: *p* = 0.4% and ITS: *p* = 7.3%) were found, that may constitute cryptic species. Lineage A only occurred in Kakaban lake (East Kalimantan), while lineage B was present in all sampled populations. Within lineage B, we found low levels of genetic diversity in lakes, though there was spatial genetic population structuring. The Australian population is genetically differentiated from the Indonesian populations. Within Indonesia we did not record an East-West barrier, which has frequently been reported for other marine invertebrates. Kakaban lake is the largest and most isolated marine lake in Indonesia and contains the highest genetic diversity with genetic variants not observed elsewhere. Kakaban lake may be an area where multiple putative refugia populations have come into secondary contact, resulting in high levels of genetic diversity and a high number of endemic species.

Introduction

It has long been hypothesized that marine species have large geographic ranges with large population sizes, and are faced with weaker barriers to dispersal than terrestrial organisms, thus resulting in relatively slow rates of speciation (e.g. [1]). The assumed presence of circum-tropical species has supported this view. However, recent phylogeographic and population genetic studies on marine taxa portray a situation of ecologically heterogeneous environments on small spatial scales with several morphologically cryptic species instead of cosmopolitan species (e.g. [2,3,4,5,6,7,8]). These results suggest that there may be many more barriers to dispersal at small spatial scales than we are able to observe [1,9,10]. The existence of multiple independently derived populations in landlocked marine lakes provides an opportunity for fundamental research into the role of isolation in population divergence and speciation in marine taxa [11]. Marine lakes are anchialine systems, which are landlocked water bodies that maintain a marine character through narrow submarine connections to the sea (Fig. 1; [12]). The marine lakes share many characteristics with island systems [13]: they are well-defined geographically [14,15,16], harbor unique biota with high endemism and/or an abundance of species rare that are elsewhere [16,17,18,19,20,21], and isolated populations [11,22,23]. The marine lakes in the Indo-Pacific were formed less than 12,000 years ago [13,24], yet their biodiversity is unique. Consistent with island biogeography theory [25,26,27] larger lakes harbor more species than smaller ones and the most isolated lakes contain the highest number of putative endemics, while the more connected lakes are dominated by reef species [15,16,19]. The degree of isolation thus appears to influence the species diversity within the lakes. In the present study our overall aim was to obtain insight into the role of isolation on the genetic diversity of marine lake populations.

Phylogeographic studies of anchialine systems across the world typically show high levels of genetic differentiation between marine...
lake populations, suggesting little to no gene flow at small spatial scales ranging from 10 to 100 km (Table 1). Furthermore, molecular markers revealed the presence of highly divergent, but morphologically cryptic species in a number of taxa such as cnidarians, crustaceans, fish and mollusks [11,22,23, 28,29,30,31,32]. There are, however, exceptions to this general pattern, which have been interpreted as resulting from life history strategies involving greater dispersal capabilities [33] or too limited sampling [34].

Here we have conducted the first phylogeographic study of Indonesian marine lake populations. The sponge species *Suberites diversicolor* [Porifera: Demospongiae: Suberitidae] is an ideal taxon to pursue this study as it allows comparison of multiple lakes at various scales and with varying degrees of connection to the sea (Fig. 1b). There are few other species that are prevalent in marine lakes [16]. *Suberites diversicolor* occurs in most moderately to highly isolated marine lakes in Indonesia [16], as well as in limited numbers of small populations in sheltered bays in Singapore, Indonesia and Australia [35]. This species shows great plasticity in adapting to harsh environments (low salinity and exposure to air) yet is absent in coral reefs [16,19,36,37]. Sponges are one of the most dominant taxa in marine lakes in terms of biomass and species diversity [16,17]. Recent comprehensive studies of sponge assemblages of marine lakes, coastal mangroves and coral reefs in various scales and with varying degrees of connection to the sea [21]. The specific aims of this study were: 1) to estimate levels of diversity and divergence of seven marine lake populations and three coastal populations using two mitochondrial and two nuclear markers, 2) to study the phylogeography of *S. diversicolor* populations in marine lakes across Indonesia, and 3) to investigate possible relationships between genetic diversity and the level of isolation of the lakes.

**Materials and Methods**

**Permits**


**Sampling**

Twenty four marine lakes and adjacent coastal habitats in Indonesia were thoroughly surveyed by snorkeling for the presence of the sponge *Suberites diversicolor*. Populations of *Suberites diversicolor* were located in seven marine lakes (29% of all surveyed lakes) in the region of Berau, East Kalimantan province (Kakaban lake, Haji Buang lake, Tanah Bamban lake) and the regions of Northern Raja Ampat (Cassiopeia lake, Utrani lake, Sawandarek lake) and Southern Raja Ampat (Misool Jellyfish lake) in West Papua province, and in mangroves along the coast of the island of Maratua in the region of Berau, East Kalimantan province (Fig. 2). Additional coastal populations were sampled from Johor Straight in Singapore (collected by S.C. Lim) and the man-made open Lake Alexander in Darwin, Australia (collected by B. Alvarez), resulting in a total of seven marine lake populations and three coastal populations sampled for this study (Fig. 2). The lakes Kakaban, Tanah Bamban, Haji Buang and Misool house immense perennial populations of the jellyfish *Mastigias papua* such as those that have been extensively documented in five marine lakes in Palau. For a full description of the sampled marine lakes,

**Table 1. Overview of published genetic variation in populations within anchialine systems.**

<table>
<thead>
<tr>
<th>Anchialine system</th>
<th>Location</th>
<th>Taxon</th>
<th>Marker(s)</th>
<th>Structure</th>
<th>Scale of differentiation</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lake Palau</td>
<td>Palau</td>
<td><em>Mastigias papua</em></td>
<td>mtDNA COI &amp; nDNA ITS</td>
<td>each lake private haplotypes</td>
<td>1–50 km</td>
<td>[11,59]</td>
</tr>
<tr>
<td>Lake Palau</td>
<td>Palau</td>
<td><em>Brachidontes</em> sp.</td>
<td>mtDNA COI</td>
<td>divergent species; each lake private haplotypes</td>
<td>1–50 km</td>
<td>[22]</td>
</tr>
<tr>
<td>Lake Palau</td>
<td>Palau</td>
<td><em>Sphaeramia orbicularis</em></td>
<td>mtDNA control region</td>
<td>lakes reduced diversity private haplotypes</td>
<td>1–50 km</td>
<td>[23]</td>
</tr>
<tr>
<td>Pool Hawaii island</td>
<td>Hawaii island</td>
<td><em>Holocaridina rubra</em></td>
<td>mtDNA COI</td>
<td>each pool private haplotypes</td>
<td>30–50 km</td>
<td>[28]</td>
</tr>
<tr>
<td>Pool Hawaii Archipelago</td>
<td>Hawaii Archipelago</td>
<td><em>Holocaridina rubra</em></td>
<td>mtDNA COI</td>
<td>each pool private haplotypes</td>
<td>10–50 km</td>
<td>[29]</td>
</tr>
<tr>
<td>Pool Maui &amp;Hawaii</td>
<td>Maui &amp;Hawaii</td>
<td><em>Holocaridina rubra</em></td>
<td>mtDNA COI</td>
<td>each pool private haplotypes</td>
<td>1–100 km</td>
<td>[94]</td>
</tr>
<tr>
<td>Pool Hawaii Archipelago</td>
<td>Hawaii Archipelago</td>
<td><em>Metabeneus lahen</em></td>
<td>mtDNA COI</td>
<td>panmixia</td>
<td>25–300 km</td>
<td>[33]</td>
</tr>
<tr>
<td>Cave Philippines</td>
<td>Philippines</td>
<td><em>Nentilia cavernicola</em></td>
<td>mtDNA COI</td>
<td>panmixia</td>
<td>200 km</td>
<td>[34]</td>
</tr>
<tr>
<td>Cave Australia</td>
<td>Australia</td>
<td><em>Stygioancus lancifera</em></td>
<td>mtDNA COI 16S</td>
<td>divergent species</td>
<td>10–100 km</td>
<td>[30]</td>
</tr>
<tr>
<td>Cave Spain</td>
<td>Spain</td>
<td><em>Metacranonym longipes</em></td>
<td>mtDNA COI 16S histone</td>
<td>divergent species</td>
<td>20–100 km</td>
<td>[32]</td>
</tr>
<tr>
<td>Cave Mexico</td>
<td>Mexico</td>
<td><em>Craseria marley</em></td>
<td>mtDNA COI 16S</td>
<td>divergent populations</td>
<td>10–100 km</td>
<td>[31]</td>
</tr>
</tbody>
</table>

doi:10.1371/journal.pone.0075996.t001

![Image](https://example.com/figure1.png)
Figure 2. A Sample locations of the sponge *Suberites diversicolor*: top three maps represent distribution and frequencies of haplotypes for partial Cytochrome Oxidase I (COI) and bottom three maps of genotypes of internal transcribed spacer region (ITS) in Indonesia Singapore and Australia with insets of Berau (East Kalimantan left) and Raja Ampat (West Papua right) in Indonesia; location codes are explained in Table 2; circles represent marine lakes and squares are coastal populations; haplo/genotypes are indicated by number code (COI: C1-4 and ITS: T1-9) and color codes as provided in B. Note that scale differs per map. B Bayesian/maximum likelihood phylogram of 105 COI sequences (right) and 104 ITS sequences (left); each haplo/genotype indicated by specific color followed by location code and total number of samples in squared brackets. Only posterior probabilities of > 90
and maximum likelihood values of >70 are indicated. Color blocks represent the same individuals for both molecular markers (i.e. lineage A (pink) and B (green) represented by the same individuals with both COI and ITS markers). Species of the family Halichondriidae were used for the outgroup followed by Genbank accession numbers. Scale bars indicate substitutions/site.

doi:10.1371/journal.pone.0075996.g002

see Becking et al. [16]. The number of marine lakes worldwide is estimated at approximately 200 with clusters of ten or more lakes occurring in areas with a karstic limestone landscape such as Croatia, Bermuda, Vietnam, Palau, and Indonesia [20]. The lakes are formed in natural inland depressions and are subjected to a tidal regime, which is typically delayed in phase (ranging from 20 minutes to 4 hours) and dampened in amplitude (tide ranging from 20 cm to 1.5 m) compared to the adjacent sea [16,17]. The level of obstruction of water exchange, i.e. the degree of isolation, differs per lake as does the salinity and environmental regimes within the lakes [14,16]. The relative degree of isolation of each marine lake is provided in Table 2.

Collections were made randomly along the entire coastline of each of the lakes and specimens were collected at least 25 m distance from each other to avoid collecting clone siblings. Our aim was to collect 20 individuals per location, but in most locations the resident population size was too small to attain this target (see Table 2). Hence, sample sizes are small for some locations. The color and substrate of each specimen was recorded, and a photograph was taken either in situ or within 2 hours after collection. After collection, portions of the choanosome were cut into approximately 125 mm³ cubes, avoiding the surface to minimize potential contamination with protists or other sponge associates, and preserved in 96% ethanol, which was refreshed after 24 hours. The remainder of the samples were preserved in 70% ethanol and deposited in the Porifera collection of the Naturalis Biodiversity Center, The Netherlands (RMNH POR.) as voucher specimens. The investigated specimens are listed in the Appendix A.

DNA Extraction, Amplification and Sequencing

Total DNA was extracted from 105 specimens using DNeasy tissue kit (Qiagen), following the instructions of the manufacturer. Portions of four markers were amplified: two mitochondrial genes, cytochrome oxidase subunit 1 (COI) and subunit 2 (COII), and two nuclear markers, the nuclear ribosomal operons consisting of partial 18S rDNA, full-length internal transcribed spacer 1 and 2, 5.8S, and partial 28S rDNA fragments (ITS) and the D3-D5 region of the nuclear ribosomal 28S gene (28S). The nuclear markers are independent from the mitochondrial markers and therefore provide extra support in case of congruent results.

The standard DNA-barcoding fragment of COI was amplified by using a specific forward primer designed for Suberites SUB-COI-F: GGAATGTACGGGACAGCTTTTAGCATG and a degenerated reverse primer from Folmer et al. [38] designed by Meyer et al. [39]: dhHCO2198: TAA ACT TCA GGG TGA CCA AAR AAY CA. COII was amplified with the primers from Rua et al. [39]: dgHCO2198: TAA ACT TCA GGG TGA CCA AAR AAY CA. COII was amplified with the primers from Rua et al. [39]. Amplifications were carried out in 25 µl Phire reaction volumes containing 5 µl Phire® Reaction Buffer, 3 µl dNTPs (1 mM), 0.625 µl of each primer (10 µM), 0.25 µl Phire® Hotstart-Tag polymerase DNA (Thermo Scientific, Finnzymes), and 1 µl of DNA (10–20 ng/µl). The temperature regime for amplification: 94°C for 30 s; followed by 35 cycles of 94°C for 5 s; 50°C for 5 s; 72°C for 12 s; followed by 72°C for 1 min. PCR products were purified and sequenced by Macrogen Inc (Korea and The Netherlands).

Data Analysis

The poriferan origin of the obtained sequences was verified through BLAST searches against Genbank [http://blast.ncbi.nlm.nih.gov/Blast.cgi]. Sequences were handled in SEQUENCHER 4.10.1 (Gene Codes Corporation) and aligned with CLUSTALW and MUSCLE as implemented in DAMBE [41] and SEAVIEW v 4.3.0 [42]. Alignment was conducted under default settings and optimized by eye. Alignments were collapsed to contain only unique sequence types in DAMBE. Haplo-/genotypes and nucleotide diversity as well as Tajima’s D neutrality test were calculated per population with Arlequin v. 3.11 [43].

Phylogeographic analyses were carried out for COI and ITS. We used ITS outgroup sequences obtained from Genbank from the family Halichondriidae (Figure 1), as the available sequences for ITS of other Suberitidae were more distant than those from Halichondriidae. Several studies have shown that the families Suberitidae and Halichondriidae are sister groups [44,45,46]. To be consistent we also used species of the family Halichondriidae for the outgroup of the COI phylogram. The relatively best-fit DNA substitution model was selected by the Akaike Information Criterion deployed in [MODELTEST v. 0.1.1 [47] and this model (COI: HKY and ITS: GTR+G+I) 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. [48]. Each analysis consisted of two independent runs of four Metropolis-coupled Markov-chains, sampled at every 1,000th 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 <0.001. Convergence was also checked in Tracer v. 1.5.0 [49]. For comparison, maximum likelihood bootstrap analyses were conducted using MEGA v. 5.01 [50] using a heuristic search with 1,000 bootstrap replicates. The Bayesian and maximum likelihood phylograms were combined and visualized using TreeGraph 2 [51]. Within group p-distance (uncorrected), as well as net nucleotide divergence between groups were calculated in MEGA. A Kruskall-Wallis test was performed to test whether color or substrate preference significantly differed between lineages. To test for spatial structuring of samples we performed an analysis of molecular variance (AMOVA) and calculated pairwise Φst values between separate populations using Arlequin 3.5.1.2 [43]. Significance of pairwise Φst values (based on p-distances) was determined by 10,000 permutations and exact tests of population differentiation in Arlequin.

Results

Sequence variation (COI, COII, ITS, 28S)

All sequences were submitted to GenBank with accession numbers KF568951-KF568965 (Table S1). We obtained final alignments (excluding primers) for the sponge S. diversicolor of 519 bp for COI with four haplotypes (C1-C4, 105 individuals, KF568960 - KF568963), 331 bp for COII with one haplotype (105 individuals, KF568964), 689 bp of ITS with nine genetic
Table 2. Sample locations of ten *Suberites diversicolor* populations from marine lakes and coastal locations in Berau (East Kalimantan) and Raja Ampat (West Papua) in Indonesia, Darwin in Australia, and Singapore.

<table>
<thead>
<tr>
<th>Code</th>
<th>Location</th>
<th>Region</th>
<th>Latitude</th>
<th>Longitude</th>
<th>Connection</th>
<th>density <em>S. diversicolor</em> (ind. 50m²)</th>
<th>color morphs <em>S. diversicolor</em></th>
<th>nCOI</th>
<th>nITS</th>
<th>nCOII</th>
<th>n28S</th>
<th>Size lake (1000 m³)</th>
</tr>
</thead>
<tbody>
<tr>
<td>KKB</td>
<td>Kakaban lake</td>
<td>Berau</td>
<td>N02° 08’ 23.5”</td>
<td>E118° 30’ 31.9”</td>
<td>most isolated</td>
<td>15-Jan</td>
<td>green, red</td>
<td>22</td>
<td>21</td>
<td>21</td>
<td>5</td>
<td>4000</td>
</tr>
<tr>
<td>HBL</td>
<td>Haji Buang lake</td>
<td>Berau</td>
<td>N02° 12’ 30.4”</td>
<td>E118° 35’ 40.8”</td>
<td>isolated</td>
<td>15–50</td>
<td>green, red, blue, purple, yellow</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>2</td>
<td>140</td>
</tr>
<tr>
<td>TBB</td>
<td>Tanah Bamban lake</td>
<td>Berau</td>
<td>N02° 13’ 50.0”</td>
<td>E118° 34’ 50.7”</td>
<td>least isolated</td>
<td>0–2</td>
<td>green</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>120</td>
</tr>
<tr>
<td>RAJ</td>
<td>Sauwandarek lake</td>
<td>Raja Ampat</td>
<td>S0° 35’ 19.6”</td>
<td>E130° 35’ 48.8”</td>
<td>very isolated</td>
<td>0–10</td>
<td>purple, blue green</td>
<td>21</td>
<td>21</td>
<td>21</td>
<td>2</td>
<td>84</td>
</tr>
<tr>
<td>CAS</td>
<td>Cassiopeia lake</td>
<td>Raja Ampat</td>
<td>N0° 08’ 36.6”</td>
<td>E130° 04’ 39.8”</td>
<td>least isolated</td>
<td>0–10</td>
<td>green</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>2</td>
<td>13</td>
</tr>
<tr>
<td>URA</td>
<td>Urani lake</td>
<td>Raja Ampat</td>
<td>N0° 06’ 05.1”</td>
<td>E130° 15’ 05.5”</td>
<td>isolated</td>
<td>0–2</td>
<td>green</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>1</td>
<td>68</td>
</tr>
<tr>
<td>MIS</td>
<td>Misool Jellfish Lake</td>
<td>Raja Ampat</td>
<td>S01° 55’</td>
<td>E130° 20’</td>
<td>isolated</td>
<td>0–2</td>
<td>green</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>BER</td>
<td>Maratua mangrove</td>
<td>Berau</td>
<td>N02° 12’ 52.3”</td>
<td>E118° 38’ 34.1”</td>
<td>open</td>
<td>0–1</td>
<td>green, yellow</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>DAR</td>
<td>Lake Alexander Darwin</td>
<td>Australia</td>
<td>S12° 25’</td>
<td>E130° 50’</td>
<td>open</td>
<td>0–1</td>
<td>green</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>SIN</td>
<td>Johor Strait</td>
<td>Singapore</td>
<td>N 01° 26’02.34”</td>
<td>E104° 02’54.31”</td>
<td>open</td>
<td>0–1</td>
<td>purple, blue, green</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>

Per locality relative connection to the adjacent sea is provided and for the marine lake size. In addition the density of the target sponge species *Suberites diversicolor* color morphs and number of samples per genetic marker (COI, COII, 28S, ITS) is provided per location. doi:10.1371/journal.pone.0075996.t002
variants (T1-9, 104 individuals, KF568951- KF568959) (Table 2, Table S1). For a subset of 20 specimens we obtained 574 bp for 28S resulting in one genetic variant (KF568965).

Divergent lineages in *S. diversicolor*

COI and ITS sequences were obtained from the same specimens and fall apart into two major lineages, termed A and B (Fig. 2B). These lineages represented reciprocally monophyletic groups for both markers and were strongly supported by both Bayesian and maximum likelihood inference methods (Fig. 2B). Lineage A was represented by haplotype C1 for COI and genotypes T1-5 for ITS. Lineage B is represented by COI haplotypes C2-4 and ITS genotypes T6-9 (Fig. 1B). Within lineage A there was no sequence variation in COI (n = 14), while the average p-distance within lineage B was 0.25% (n = 91). The net nucleotide divergence between lineages A & B for ITS was 0.38%. Haplotype C1 (lineage A) differed by two basepairs from C2 (the dominant haplotype from lineage B) of which one resulted in a non-synonymous substitution between two unpolar amino acids, from isoleucine to valine. For ITS the average p-distance within lineages A was 0.44% (n = 13), while the average p-distance within lineage B was 0.29% (n = 91). The net nucleotide divergence between lineages A & B for ITS was 7.26%. Several indels of 1–3 bp length were observed and were consistent within lineages and differed between lineages. There were insertions in lineage A with respect to lineage B from 102–103 bp (either CT or TT), 380–381 bp (CA), 470–473 bp (GGA or GAA). There were gaps in lineage A with respect to lineage B from 139, 178–180, 549–555 bp. No double peaks were observed, and it we therefore assumed that no intragenomic polymorphisms occur within this species. The level of intragenomic polymorphisms differs per species [52]. We consider the risk of analyzing paralogous rDNA sequences indicated in Table 2); gene diversity (h) nucleotide diversity (r) Tajima’s D neutrality test.

Diversity and spatial population structuring (COI & ITS)

Lineage A was only present in Kakaban lake while lineage B was present in all populations. The geographical distribution of COI haplotypes is shown in Fig. 2A. Of the four detected haplotypes in COI, haplotype C1 was restricted to Kakaban lake (East Kalimantan). Haplotype C3 only occurred in one individual in Urani lake (West Papua). The Darwin population was represented by haplotype C4, which was shared with no other population. Haplotype C2 was the most abundant haplotype, occurring in all populations except Darwin and was the dominant haplotype in the populations of Berau mangroves, Singapore, Sauwandarek lake, Cassopia lake, Urani lake, and Misool Jellyfish lake. Of the nine detected genotypes of ITS, five were restricted to Kakaban lake (genotypes T1-5), which are all representatives of lineage A. Genotype T7 (lineage B) was the most abundant and was shared by all sampled populations except Haji Buang lake and Tanah Bamban lake (Kalimantan) and Darwin (Australia). Darwin was represented by the private genotypes T8-9. Haji Buang lake and Tanah Bamban lake harbored a single genotype (T6) that was shared by Kakaban lake (Kalimantan) and Misool Jellyfish lake (Papua).

Within lineage A in Kakaban lake there was only a single haplotype of COI while the ITS gene diversity was 0.8242/-0.0567, and ITS nucleotide diversity was 0.005656+/-0.003392. Within lineage B all populations contained a single COI haplotype except Urani lake, which had two haplotypes with a haplotype diversity of 0.3333+/− 0.2152 and nucleotide diversity of 0.000624+/− 0.000822. For ITS, the majority of the populations contained only a single genotype, except for Kakaban lake, Darwin and Misool Jellyfish lake. The population in Kakaban lake had the highest gene and nucleotide diversity in lineage B, followed by Darwin and Misool Jellyfish lake (Table 3). Tajima’s D tests of neutrality were carried out per population. The majority of the populations had a zero value due to the presence of only one genetic variant. Values of Tajima’s D for ITS were negative, but not significant (p>0.1) in Misool lake and Darwin (Table 3).

Spatial analysis of genetic structure of lineage B COI and ITS sequences showed that the Darwin population was strongly (Φst between 0.53–1) and significantly differentiated from all marine lakes populations (Table 4 & Table S2). Besides Darwin there was no significant differentiation in COI between the different populations (Table S2). The ITS marker was more diverse and showed more structure among the populations than COI (Fig. 1, Table 3). The Berau lakes (East Kalimantan), Kakaban and Haji Buang lakes were all significantly differentiated. The Raja Ampat lakes (West Papua) were not genetically differentiated from each other.

Table 3. Genetic diversity indices based on ITS sequences per population of *Suberites diversicolor* of lineage A and B (location codes indicated in Table 2); gene diversity (h) nucleotide diversity (r) Tajima’s D neutrality test.

<table>
<thead>
<tr>
<th>Code</th>
<th>Lineage</th>
<th>n ITS</th>
<th>h ITS</th>
<th>r ITS</th>
<th>Tajima’s D</th>
</tr>
</thead>
<tbody>
<tr>
<td>KKB</td>
<td>A</td>
<td>13</td>
<td>0.8242+/− 0.0567</td>
<td>0.005656+/- 0.003392</td>
<td>1.3927</td>
</tr>
<tr>
<td>KKB</td>
<td>B</td>
<td>8</td>
<td>0.5357+/− 0.1232</td>
<td>0.001578+/- 0.001318</td>
<td>1.4488</td>
</tr>
<tr>
<td>HBL</td>
<td>B</td>
<td>20</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>TBB</td>
<td>B</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>RAJ</td>
<td>B</td>
<td>21</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CAS</td>
<td>B</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>URA</td>
<td>B</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>MIS</td>
<td>B</td>
<td>7</td>
<td>0.2857+/− 0.1964</td>
<td>0.000842+/- 0.000879</td>
<td>−1.23716</td>
</tr>
<tr>
<td>BER</td>
<td>B</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>SIN</td>
<td>B</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>DAR</td>
<td>B</td>
<td>6</td>
<td>0.3333+/− 0.2152</td>
<td>0.000980+/- 0.000997</td>
<td>−1.13197</td>
</tr>
</tbody>
</table>

The majority of populations had only one haplotype resulting in 0 values for all indices calculated. All Tajima D values are not significant.

doi:10.1371/journal.pone.0075996.t003
other except Misool, which was differentiated from all Raja Ampat lakes, yet not from the populations of the lakes Kakaban, Haji Buang and Tanah Bamban (East Kalimantan). The AMOVA analyses revealed that significant portions of the total variance within lineage B can be attributed to differences among the following three groups 1. Berau coast, Singapore coast, Northern Raja Ampat lakes (Sauwandarek, Cassopeia, Urani), 2. Berau lakes (Kakaban, Tanah Bamban, Haji Buang), Southern Raja Ampat (Misool), 3. Darwin. The among group variation was 84.6% (p<0.001) and the within population variation was 10% (p<0.001).

### Discussion

#### Divergent lineages in *S. diversicolor*

Two major lineages were uncovered in the populations of the sponge *S. diversicolor*. The congruent patterns of COI and ITS genetic markers and the degree of divergence between the two lineages (COI: 0.4% and ITS: 7.3%) are indicative of reproductive isolation, and thus we suggest that the two lineages (A and B) constitute different species. We searched for morphological and ecological characters to distinguish the two lineages, but did not find any. Genetic divergence can preclude morphological or ecological distinction. The skeletal structure and spicule lengths do not differ between lineages and fall within the natural variation of this species (see also [35]). The color and substrate preference are variable, but not consistent within a particular lineage. A related *Suberites* from Satonda lake (Sumbawa, Indonesia) displays different colors at different depths as a result of a symbiosis with the unicellular green algae *Chlorella* and symbiotic bacteria [53]. Phylogeographic studies in the Indo-Australian-Archipelago have uncovered numerous lineages in marine taxa that may represent undescribed cryptic species (e.g. [3,6,54]). Within sponges, molecular studies have revealed a high prevalence of morphologically cryptic sponge species (see review in [8,55]).

The divergence between lineage A and B points to a long isolation in spite of the fact that they are sympatric in Kakaban lake (East Kalimantan). Within sponges there are several reports of sympatric cryptic species: *Tedania* spp. in mangroves [56], *Scopalina lophypoda* [57], *Chiona* spp. [8], and *Hexadella* spp. [7]. Differential reproductive traits and output can promote the co-existence of sibling species (e.g. [57,58]). This observation of divergent lineages in one lake is, however, not common in the phylogeographic studies conducted thus far on populations in the marine lakes in Palau [11,22,23,59]. The Palauan studies on three distinct taxa (jellyfish, fish and bivalves) mostly show a pattern of one lineage occupying one lake [11,22,23,59]. One reason why Kakaban lake may contain two lineages is the sheer size of the lake – at almost 4 km² it is tenfold larger than any of the other marine lakes in Indonesia and the majority of Palau (Table 2; [15,16]). Alternatively, lineage B could be a recent introduction to the lake. Sponge fragments are known to be transported by waterfowl [60] and workers from the neighboring island Maratua who stay on Kakaban for short periods to attend small crops may also act as possible vectors of *Suberites diversicolor* from the Maratua lakes or the mangroves near their village.

### Phylogeography

Lineage A is only present in Kakaban lake, while lineage B is present in all populations. Within lineage B the spatial genetic structure shows three groups: 1. the three Berau lakes and southern Raja Ampat lake, 2. Berau coast, Singapore coast and the three northern Raja Ampat lakes, 3. Darwin, Australia. At present there is no comprehensive phylogeographic study of sponges spanning the Indonesian archipelago, yet pronounced genetic differences in populations of other marine invertebrates and vertebrates are present between the Java Sea, the Indonesian Through Flow, and the seas of East Sulawesi [3,5,9,61,62]. The marine phylogeographic patterns of these studies strongly support the existence of a barrier in the area between the Sunda and Sahul shelves, where populations from Kalimantan are genetically isolated from those in Papua. Our data do not show such a clear East to West phylogeographic break. The Darwin population of the present study, though small in sample size, is genetically differentiated from the other populations. This is consistent with phylogeographic studies of sponges and other invertebrates that show a barrier within the Torres Strait (e.g. [63,64]). Dispersal potential and habitat specialization may determine how lineages are distributed and how fauna of different geographic regions are connected (e.g. [9]). Many sponge population genetic and phylogeographic studies have revealed structured populations with in some cases evidence of (occasional) long distance dispersal events [65,66,67,68,69,70]. This pattern is congruent with philopatric, shortlived larvae that recruit at short distances from the parental locations [71,72], whilst at the same time the possibility of sponges to disperse as viable fragments in the currents or rafting on various floating material [73,74,75]. The reproductive cycle and larvae of *S. diversicolor* are unknown, but this species

---

**Table 4.** Pairwise *F*st values between all populations of lineage B based on ITS sequences of *Suberites diversicolor* (location codes indicated in Table 2).

<table>
<thead>
<tr>
<th></th>
<th>KKB</th>
<th>HBL</th>
<th>TBB</th>
<th>RAJ</th>
<th>CAS</th>
<th>URA</th>
<th>MIS</th>
<th>BER</th>
<th>SIN</th>
</tr>
</thead>
<tbody>
<tr>
<td>HBL</td>
<td></td>
<td>0.60591*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TBB</td>
<td>0.30435</td>
<td></td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RAJ</td>
<td>0.60591*</td>
<td>1*</td>
<td>1*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAS</td>
<td>0.4702*</td>
<td>1*</td>
<td>1*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>URA</td>
<td>0.42857</td>
<td>1*</td>
<td>1*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MIS</td>
<td>0.13514</td>
<td>0.16749</td>
<td>−0.09804</td>
<td>0.91393*</td>
<td>0.86315*</td>
<td>0.84646*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BER</td>
<td>0.25</td>
<td>1*</td>
<td>1*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SIN</td>
<td>0.30435</td>
<td>1*</td>
<td>1*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DAR</td>
<td>0.53451*</td>
<td>0.9512*</td>
<td>0.86348*</td>
<td>0.83584*</td>
<td>0.74359*</td>
<td>0.71049*</td>
<td>0.77327*</td>
<td>0.55882</td>
<td>0.60396</td>
</tr>
</tbody>
</table>

Values in bold and with asterisk indicate significant values (p<0.05).
doi:10.1371/journal.pone.0075996.t004
does produce asexual buds [Becking pers. obs.] which may survive a considerable amount of time in the plankton or by rafting before colonizing distant locations as proposed by Worheide et al. [67] for *Leucetta chagosensis*.

For lineage B we found no private haplotypes in any of the Indonesian marine lakes, and many lakes were identical in composition. The only other phylogeographic studies on marine lakes have been in the islands of Palau on the jellyfish *Mastigias papua* [11,59], the fish *Sphaeramia orbicularis* [22], and the mussel *Brachidontes* sp. [23] (see Table 1). These studies show extreme genetic isolation, low genetic diversity, and in the cases of *Mastigias papua* and *Brachidontes* sp. rapid morphological evolution in the marine lakes [11,22,23,59]. The lack of strong population structure between many of the Indonesian lakes of the present study may be caused by recurrent (recent and historic) gene flow among lakes. Alternatively, it is still possible that all these lakes are completely isolated, i.e. do not exchange any migrants, but that the lack of structure may be a result of the markers we used. These markers may not evolve fast enough for mutations to have accumulated to show the recent divergence. Of the four molecular markers used in the present study, ITS evolves the fastest and provided the highest resolution of spatial genetic structure. The difference in genetic diversity between COI and ITS is large. In contrast to most animals the mitochondrial DNA of sponges evolves slowly and generally slower than nuclear DNA [76,77].

The interspecific variation of COI in sponges can be as low as 0–0.5% (p-distances) (e.g. [55,78,79]). However, in some sponge taxa COI can provide low but sufficient genetic variation within species over relatively short geographic distances [63,69,80].

We found no sequence variation in 28S and COII markers between any of the populations or between the two lineages of *S. diversicolor*. The D3–D5 region of the 28S fragment has been used to distinguish genera and species of a wide range of demosponge taxa including halichondrinds [81,82], but was also reported too conserved to discriminate between closely related species in other sponge taxa [7]. COII was proposed as a polymorphic mitochondrial marker for sponge phylogeography by Rua et al. [40]. Rua et al. [40] indicated that the variation of this marker could be low in halichondrdid species *Hymeniacidon heliophila* but attributed their results to the selection of clone-mates. In the present study COII showed no variation between any of the samples spanning a wide geographic range. We conclude that COII is not a suitable marker for intraspecific variation or distinction between closely related species of the genus *Suberites* in particular, and probably more generally for the families Suberitidae and Halichondriidae.

**Isolation & genetic diversity**

Kakaban lake is the largest and most isolated lake in Indonesia (see Table 2), and it houses a high proportion of endemic sponge species [16,21]. In concordance, our study shows that the population of *Suberites diversicolor* displayed the highest genetic diversity with unique genetic variants that were not shared with two marine lakes at just 6 km distance (Fig. 1). These results indicate that Kakaban lake is very isolated both in physical and biological terms. Isolation acts to decrease the rate of immigration and thus to decrease the genetic diversity and the number of species expected at equilibrium in an island system [25,26,27,83]. Yet isolation can also enhance species formation, with the diminished gene flow allowing populations to diverge and ultimately form new species if they remain isolated [17,27,83]. In Palau the degree of genetic distance between marine lake and adjacent sea populations was strongly correlated with the degree of connection from the lake to the sea and not the actual geographic distance between the populations [11,23]. In the present study, the molecular markers used were not variable enough to detect a relationship between moderate levels of isolation and the genetic diversity of the lakes. For example, the populations in northern Raja Ampat lakes (West Papua) are not genetically differentiated, despite the limited physical connections to each other and to the adjacent sea.

**Biogeographic scenario**

Kakaban lake was probably filled with sea water less than 12,000 years ago [13,24]. Considering the deep divergence between lineages A & B in this lake, this divergence likely occurred well before the formation of Kakaban. Worheide et al. [52] estimated an evolutionary rate of 1% per million years for ITS in a suberitid sponge *Prosuberites ‘laughlini’* based on the formation of Isthmus of Panama. Implementing the 1% mutational rate would mean that the two lineages diverged approximately 7 million years ago. Though this is a rough estimation with great error bars and rates of evolution may be higher for recently diverged lineages [84], the age is consistent with recent phylogeographic studies that suggest that many endemics from the Indo-Australian-Archipelago have origins in the early Pliocene-Miocene (3-20 million years ago; e.g. [85,86,87]).

Kakaban lake houses a genetic and species diversity of sponges, that appears to be absent from the surrounding sea (see also [21,36]). Each lake is ephemeral, but the marine lakes ecosystem probably has occurred in various locations during the past glacial cycles [24]. The Sunda shelf, which includes Borneo (Kalimantan), was exposed during the Last Glacial Maximum (LGM) when sea levels are estimated to have been approximately 110–140 m lower than modern sea levels [88,89,90]. Multiple larger and smaller depressions in the shelf have been recorded which presumably represented palaeo-lakes during the LGM [24] that could have become brackish marine lakes with the increase in sea level. During the LGM the Sunda Land region was also dominated by mangroves [91], and the water around the Sunda area would have been brackish due to the multiple river outlets [90]. These are both environments amenable for *S. diversicolor*. Ancient lineages/endemics may have ‘hopped’ from lake to lake or from mangrove to lake, as the lakes formed and subsequently disappeared with the rise and fall in sea level during the Plio-Pleistocene glacial cycles. Genetic signatures of glacial refuge are expected to be characterized by high genetic diversity and a mixture of ancestral and private haplotypes [92,93]. While Kakaban matches this pattern, the lake could not have been a refuge during LGM (it was dry), however there may have been palaeo-lakes in the vicinity that served as such. Kakaban may be an area where multiple putative refuge populations have come into secondary contact, resulting in the high genetic diversity and the high number of endemics. Molecular studies on co-distributed taxa at larger scales including lakes from adjacent regions in Palau and Vietnam will enhance our understanding of the processes behind the unique marine lake diversity.

**Supporting Information**

**Table S1** List of *Suberites diversicolor* specimens studied. For each specimen the following information is provided: lineage to which it belongs (see Fig. 2) haplotype of Cytochrome Oxidase I (COI) GenBank Accession Number for the COI haplotype genotype of internal transcribed spacer region of nuclear ribosomal operons (ITS) GenBank Accession Number for ITS genotype location of collection (location codes indicated in Table 2) the color when
alive substrate it resided on collection number within the Porifera Collection of the Naturalis Biodiversity Center (RMNH POR).

(XLSX)

Table S2. Pairwise Φst values between all populations of the sponge *Suberites diversicolor* lineage B based COI of *Suberites diversicolor* (location codes indicated in Table 2). Values in bold indicate significant values (p<0.05).

(XLSX)

Acknowledgments

M. Ammer and B. Hoeksema were invaluable sources of information for the fieldwork. We would also like to thank the following people for their help with logistics of the fieldwork and samples: Suharsono, Y. Tuti, C. Huffard, M. Erdmann, S. Manugbai, Bahruddin, Estradivari, N. Santodomingo, E. Dourdop, W. Renema, B. Alvarez, S.-C. Lim, and the staff of TNC/WWF Berau Office, of Nabucco Island Dive Resort, of Derawan Dive Resort, of Conservation International Wayag station, and of The Nature Conservancy Misool station. We are grateful to the following people who facilitated the laboratory work: G. Worheide, S. Menken, H. Breeuwer, B. Voeldtij, P. Kuperus. 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, W. Renema, C. Vogler, and two anonymous reviewers provided valuable comments on the original manuscript.

Author Contributions

Conceived and designed the experiments: LEB DE KTCAP NJV. Performed the experiments: LEB DE. Analyzed the data: LEB DE KTCAP. Contributed reagents/materials/analysis tools: LEB DE NJV. Wrote the paper: LEB DE KTCAP NJV.

References


