Global Ecology and Conservation 12 (2017) 170–187

Contents lists available at ScienceDirect

Global Ecology and Conservation

journal homepage: http://www.elsevier.com/locate/gecco

Original Research Article

Identifying spatially concordant evolutionary significant units across multiple species through DNA barcodes: Application to the conservation genetics of the freshwater fishes of Java and Bali

Aditya Hutama [a], Hadi Dahruddin [b], Frédéric Busson [c], Sopian Sauri [b], Philippe Keith [c], Renny Kurnia Hadiaty [b], Robert Hanner [d], Bambang Suryobroto [a], Nicolas Hubert [e,*]

[a] Bogor Agricultural University, Faculty of Mathematics and Natural Science, Animal Bioscience, Jl. Raya Darmaga, Bogor 16680, Indonesia
[b] LIPI, Research Center for Biology, Zoology Division, MZB, Gedung Widyasatwaloka, Jl. Raya Jakarta Bogor Km. 46, Cibinong- Bogor 16911, Indonesia
[c] Muséum national d'Histoire naturelle, UMR 7208 (MNHN-CNRS-UPMC-IRD-UCBN), CP 026, 43, rue Cuvier, F-75231 Paris Cedex 05, France
[d] Biodiversity Institute of Ontario and Department of Integrative Biology, University of Guelph, Guelph, ON, Canada
[e] Institut de Recherche pour le Développement, UMR 226 ISEM (UM2-CNRS-IRD), Place Eugène Bataillon, CC 065, F-34095 Montpellier cedex 05, France

ARTICLE INFO

Article history:
Received 3 July 2017
Received in revised form 16 November 2017
Accepted 16 November 2017
Available online 1 December 2017

Keywords:
Phylogeography
Biodiversity hotspot
Pleistocene
Southeast Asia
Coalescent theory
Cryptic diversity

ABSTRACT

Delineating Evolutionary Significant Units for conservation purposes is a crucial step in conservation. Across a distribution range, species frequently display population structure that drives the distribution of genetic diversity. These patterns of genetic structure and diversity result from intricate interactions between biogeographic history and demographic dynamics. Prior biogeographic knowledge, however, is scarcely available, a trend particularly pronounced in the tropics where the taxonomic impediment is hampering biogeographic studies and conservation efforts. DNA barcoding has been initially proposed to foster taxonomic studies through the development of an automated molecular system of species identification. While its utility for species identification is increasingly acknowledged, its usefulness for fast and large-scale delineation of ESU remains to be explored. If proved to be useful for that purpose, DNA barcoding may also open new perspectives in conservation by quickly providing preliminary information about population conservation status. The present study aims at assessing the utility of DNA barcoding for the delineation of ESUs among the most common freshwater fish species of Java and Bali through the comparison of population genetic structures and diversification patterns across multiple species. Substantial levels of cryptic diversity are discovered among the three widely distributed freshwater fish species analyzed with a total of 21 evolutionary independent mitochondrial lineages (BINs) observed in *Barbodes binotatus*, *Channa gachua* and *Glyptothorax platypogon*. The maximum genetic distance for each coalescent tree ranges from 6.78 to 7.76 K2P genetic distances for *C. gachua* and *G. platypogon*, respectively. Diversification and population genetic analyses support a scenario of allopatric differentiation. The analysis of the BINs spatial distribution indicates concordant

* Corresponding author.
E-mail address: nicolas.hubert@ird.fr (N. Hubert).

https://doi.org/10.1016/j.gecco.2017.11.005

distribution patterns among the three species that allow identifying 18 ESUs. Implications for the conservation genetics of these species are discussed at the light of the history of the region.

1. Introduction

Conservation aims at preserving species evolutionary potential to sustain their adaptive abilities in fluctuating environments and fuel evolution on a long-term perspective. The nature of the biological units to be targeted for achieving this goal, however, has been a contentious issue since the 1990s (Crandall et al., 2000). With the objective to get around taxonomic confusion, Ryder (1986) proposed the concept of Evolutionary Significant Units (ESU) that he defined as 'subset of the more inclusive entity species, which possess genetic attributes significant for the present and future generations of the species'. He proposed to delineate ESUs based on concordant evidences from ecological, physiological and genetic perspectives (Ryder, 1986). Following this initial formulation, several recognition criteria were proposed with varying emphasis on reproductive isolation (Waples, 1991, 1995), historical trends in population structure (Avise, 1989; Moritz, 1994), shared character states (Vogler and DeSalle, 1994) and genetic or ecological exchangeability (Crandall et al., 2000). From a genetic perspective, most of the criteria focus on detecting the imprint of disrupted gene flow in mitochondrial and nuclear genomes, either from a short-term perspective when ESUs are delineated based on differences in allele frequencies (Waples, 1991, 1995; Dizon et al., 1992; Waples, 1995) or long-term perspective when ESUs are based on reciprocal monophyly of their genes genealogies (Avise, 1989; Moritz, 1994).

Fraser and Bernatchez (2001) highlighted that the procedures for delineating ESUs are based on criteria, not mandatory properties, which applicability depends on biogeographical and ecological context. Since then, major improvements happened in terms of genome analysis and statistical tools to link population genetic data with landscape ecology (Holderegger and Wagner, 2006; Anderson et al., 2010; Sork and Waits, 2010) or landscape history (Beheregaray, 2008). These methodological developments opened new perspectives in conservation, in particular, landscape genetics and niche modeling shed a new light at the way we look at dispersal and gene flow (Endo et al., 2014; Gutiérrez-Tapia and Palma, 2016). Usually implemented at small spatial scales, this process-based approach requires a prior knowledge of population structure at the regional scale and identifying ESUs is still a preliminary step in conservation (Fraser and Bernatchez, 2001; Pearse and Crandall, 2004).

Delineating ESU is context dependent as Earth biotas originated from diverse biogeographical histories resulting in species of varying age, distribution range and population structure (Wiens and Donoghue, 2004; Mittelbach et al., 2007; Weir and Schluter, 2007; Hua and Wiens, 2010). Thus, prior biogeographic knowledge should provide a useful framework to guide the delineation of ESUs (Avise, 2000; Fraser and Bernatchez, 2001). This knowledge, however, is not sufficiently detailed for many regions of the world, a situation frequently observed in the tropics where high species richness has largely amplified the problem (Beheregaray, 2008; Hubert and Hanner, 2015). DNA barcoding, the use of the cytochrome oxidase I gene as an internal species tag for molecular identifications, has opened new perspectives on collecting genomic resources across multiple species (Hebert et al., 2004; Hebert and Gregory, 2005; Janzen et al., 2005; Smith et al., 2005; Steinke et al., 2009; Hubert et al., 2012). A quick scan of population structure across multiple species through DNA barcode may provide a preliminary and insightful approach in conservation, prior to a more comprehensive assessment using both mitochondrial and nuclear markers, by (i) delineating mitochondrial lineages that depart from population dynamics and display independent mutation/drift dynamics (Hajibabaei et al., 2007; Vernooy et al., 2010; Ratnasingham and Hebert, 2013; Kekkonen and Hebert, 2014), (ii) providing genomic resources for quick species and ESUs molecular identification (Hajibabaei et al., 2011; Gibson et al., 2014), (iii) easing the characterization of ESUs range distribution and their spatial match across multiple species to identify conservation priorities.

Such strategy may be particularly relevant in areas facing massive anthropogenic threats and where conservation strategies are impeded by the lack of appropriate knowledge on species evolutionary dynamics and taxonomic confusion, a trend further amplified by the recent rarefaction of taxonomists worldwide (*i.e.* taxonomic impediment). This situation is currently observed in South-East Asia where the four biodiversity hotspots identified are among the most threatened to date (Myers et al., 2000; Lamoreux et al., 2006; Hoffman et al., 2010). This is particularly evident for the insular hotspots of the Indonesian archipelago (*i.e.* Sundaland and Wallacea), where the impact of anthropogenic activities is amplified by their refugial state, particularly so in the Sundaland hotspot (Kottelat, 1989; Woodruff, 2010; Lohman et al., 2011). Considering that Indonesia increased its Gross Domestic Product (GDP) and carbon emissions by 1000% and 400% during the last two decades (World Bank), respectively, with a population reaching 260 Millions people, it becomes evident that anthropogenic threats have increased severely. The present study focuses on delineating ESUs and their spatial concordance through the development of a DNA barcode reference library for the widespread freshwater fishes of the islands of Java and Bali, two of the less explored islands of the Sundaland hotspot (Hubert et al., 2015b; Dahruddin et al., 2017). Our objective is to provide a 3-step general framework for the delineation of ESUs through the analysis of the spatial and temporal population structures among

multiple species based on mitochondrial coalescent trees at the COI gene (Fig. 1). Finally, this framework is used to provide recommendations for evidence-based conservation strategies in the area.

2. Materials and methods

2.1. Sampling strategy and collection management

A large scale DNA barcoding campaign was conducted by the authors between November 2012 and May 2015 across 95 sites in Java and Bali islands (Dahruddin et al., 2017). A total of 3310 specimens, including 162 species, 110 genera and 53 families were collected, providing a comprehensive assessment of the Javanese and Balinese ichthyofauna. Among those 162 species, three species (*Barbodes binotatus*, *Glyptothorax platypogon* and *Channa gachua*) were sampled in more than 50% of the sites visited, displayed unusually high maximum within-species genetic distances based on a previous assessment using a restricted set of individuals (Dahruddin et al., 2017), and as such represented suitable candidates considering the objective of the present studies. Each of the three species belong to different orders displaying varied ecological preferences (Froese and Pauly, 2011), and as such, they ensure that the concordance of their ESUs spatial patterns is unlikely to results from the history of a restricted set of aquatic habitats but originated through common evolutionary dynamics of broad impact on aquatic ecosystems (Avise, 2000; Avise et al., 2016).

The three species were collected across 51 sites distributed across the islands of Java and Bali (Fig. 2, Supplementary Table S1). Specimens were captured using various gears including electrofishing, seine nets, cast nets and gill nets. Specimens were photographed, individually labeled and voucher specimens were preserved in a 5% formalin solution. A fin clip or a muscle biopsy was taken for each specimen and fixed in a 96% ethanol solution for further genetic analyses. Both tissues and voucher specimens were deposited at the national collections at the Museum Zoologicum Bogoriense (MZB) in the Research Centre for Biology (RCB) from the Indonesian Institute of Sciences (LIPI).

2.2. Sequencing and international repositories

Genomic DNA was extracted using a Qiagen DNeasy 96 tissue extraction kit following the manufacturer's specifications. A 651-bp segment from the 5′ region of the cytochrome oxidase I gene (COI) was amplified using primers cocktails C_FishF1t1/C_FishR1t1 including a M13 tails (Ivanova et al., 2007). PCR amplifications were done on a Veriti 96-well Fast (ABI-Applied Biosystems) thermocycler with a final volume of 10.0 μl containing 5.0 μl Buffer 2X, 3.3 μl ultrapure water, 1.0 μl each primer (10 μM), 0.2 μl enzyme Phire® Hot Start II DNA polymerase (5U) and 0.5 μl of DNA template (~50 ng). Amplifications were conducted as follow: initial denaturation at 98 °C for 5 min followed by 30 cycles denaturation at 98 °C for 5s, annealing at 56 °C for 20s and extension at 72 °C for 30s, followed by a final extension step at 72 °C for 5 min. The PCR products were purified with ExoSap-IT® (USB Corporation, Cleveland, OH, USA) and sequenced in both directions. Sequencing reactions were performed using the "BigDye® Terminator v3.1 Cycle Sequencing Ready Reaction" and sequencing was performed on the automatic sequencer ABI 3130 DNA Analyzer (Applied Biosystems). The sequences and collateral information have been deposited in BOLD (Ratnasingham and Hebert, 2007) in the projects BIFGA and BIFG in the container 'Barcoding Indonesian Fishes' of the 'Barcoding Fish (FishBOL)' campaign and DNA sequences were submitted to GenBank (accession numbers are accessible directly at the individual records in BOLD).

2.3. Population genetic structure (Fig. 1, step 1a)

We examined the distribution of molecular variance through an additive partitioning across increasing spatial scales as implemented in AMOVA (Excoffier et al., 1992; Excoffier and Smouse, 1994). We opted for the SAMOVA version that both defines groups of populations without *a priori* partitioning scheme and estimates molecular variance within populations, among populations within groups and among groups, based on a simulated annealing approach (Dupanloup et al., 2002). SAMOVA is not a decision-based method pointing to the optimal number of groups, we hence applied an empirical threshold and stopped increasing the number of groups once the partitioning scheme produced at least one group consisting of a single sampling site. The SAMOVA were performed for each species using the software SAMOVA 2.0 (Dupanloup et al., 2002). SAMOVA partitioning schemes were further compared to the results of a hierarchical clustering based on genetic distances. Mean genetic K2P distances were computed among sampling sites using MEGA 6 (Tamura et al., 2013) and used to produce hierarchical clusters derived from the complete linkage algorithm as implemented in hclust function of the R Stats ver. 3.1.2 package (R_Core_Team, 2014). Finally, traditional parameters of population genetic diversity including the number of haplotypes, nucleotidic diversity and mean pairwise K2P distance for each groups were computed using ARLEQUIN 3.5 (Excoffier et al., 2005).

2.4. Delineating mitochondrial lineages and inferring their diversification (Fig. 1, step 1b)

Several alternative methods have been proposed for delineating molecular lineages (Schloss and Handelsman, 2005; Pons et al., 2006; Puillandre et al., 2012). They have all in common to detect transition zones in branching patterns resulting from different segments of the gene genealogies that originated from phylogenetic diversification (speciation and extinction) or

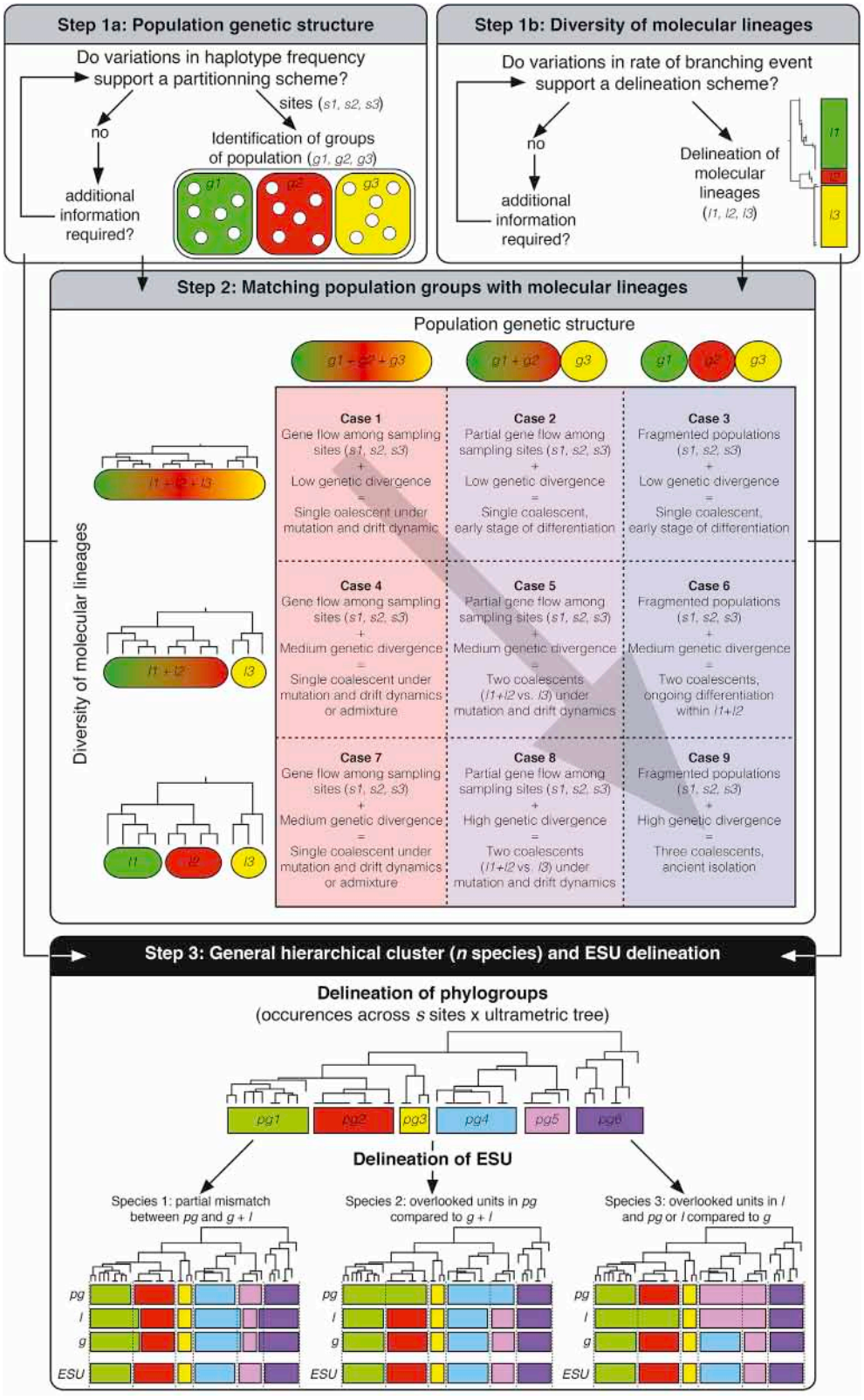

Fig. 1. Conceptual framework developed in the present study for the delineation of ESUs. **Step 1a**, detection of groups of populations differentiated by their haplotypes frequencies. **Step 1b**, detection of molecular lineages with independent evolutionary dynamics (*e.g.* lower connectance). **Step 2**, comparing population groups with molecular lineages. If population genetic structure results from ancient fragmentation of the populations, a correlation between genetic groups and BINs is expected. Conversely, the lack of correlation would indicate that population groups originated recently and either share ancient polymorphism or have been connected by gene flow in a recent past (Fu, 1999; Nielsen and Wakeley, 2001; Wakeley, 2001, 2003). Along the same line, several BINs may be delineated within a genetic group as a consequence of the stochastic nature of the coalescent but not disrupted gene flow (Hudson, 1982; Kingman, 1982; Tajima, 1983). **Step 3**, Delineation of ESUs for individual species. ESUs are defined based on either variation of haplotype frequencies or independent mitochondrial lineages after comparisons with the phylogroups defined as groups of population sharing similar sets of mitochondrial lineages.

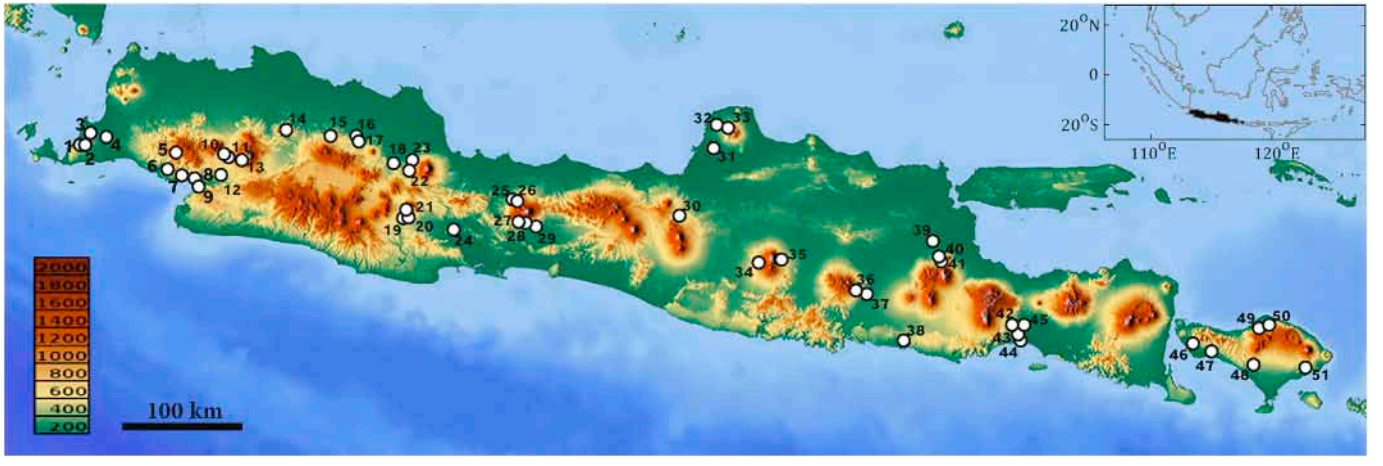

Fig. 2. Location of the 51 collection sites for the samples analyzed in this study.

Table 1
Partitioning of the molecular variance at various spatial scales, fixation indexes and number of population groups as inferred from SAMOVA. The percentage column indicates the amount of total variance explained by each of the hierarchical levels according to the number of groups of population. Φ-statistics estimate the correlation among haplotypes at each of the hierarchical levels examined and their significant departure from a random distribution of the haplotype was tested through randomization across 1000 permutations.

Number of Group	*C. gachua*				*G. platypogon*				*B. binotatus*			
	5				7				6			
Variance component	Variance	% of total	*P*	Φ-statistics	Variance	% of total	*P*	Φ-statistics	Variance	% of total	*P*	Φ-statistics
Among groups	8.788	70.56	<0.001	*ΦCT*=0.706	8.042	78.75	<0.001	*ΦCT*=0.788	6.958	69.90	<0.001	*ΦCT*=0.702
Among populations within groups	3.071	24.69	<0.001	*ΦSC*=0.839	1.516	11.34	<0.001	*ΦSC*=0.534	2.094	10.75	<0.001	*ΦSC*=0.361
Within populations	0.591	4.75	<0.001	*ΦST*=0.953	1.015	9.91	<0.001	*ΦST*=0.901	0.903	19.05	<0.001	*ΦST*=0.890

coalescent dynamics (mutation and drift). We opted for the Refined Single Linkage (RESL) algorithm that considers the number of connections of each sequences in a network estimated through the silhouette index (Rousseuw, 1987) as implemented in BOLD. Sequence connectivity is explored through random walks and optimal partitioning schemes are identified through Markov clustering. At the end, each cluster of sequence is assigned to a Barcode Index Number (BIN) in BOLD (Ratnasingham and Hebert, 2013).

Once BINs were delineated (Table S2), their timing of diversification was explored through the Bayesian approach implemented in BEAST 1.8.1 (Drummond et al., 2012). In order to establish robust prior, the best-fit substitution model was selected through the Bayesian Information Criterion (BIC) as implemented in JMODELTEST 2.1.7 (Darriba et al., 2012) and further used as a prior for the joint reconstruction of tree topology and divergence times. The initial tree topology was obtained with an UPGMA starting tree and a Coalescent model was used as a tree prior (Kingman, 1982). We used the canonical fish substitution rate of 1.2% of genetic divergence per Millions years for mitochondrial protein coding gene (Bermingham et al., 1997) and applied it to the maximum K2P distance within each species to estimate the age of the Most Recent Common Ancestor (MRCA) to be used as a prior for Bayesian analyses. The prior upper and lower bounds for the MRCAs age intervals were derived from the known highest and smallest substitution rates for fish mitochondrial genomes (Hardman and Lundberg, 2006; Read et al., 2006). We ran one MCMC of 10×10^6 step long, sampled every 1000 states with a burn-in period of 10 000. The maximum credibility tree was obtained with TreeAnnotator 1.8.1 after an additional burn-in period of 10 000. Median node ages and 95% highest posterior density (HPD) intervals were plotted in the chronogram. Duplicated sequences were removed for these analyses. We further explored the properties of BINs diversification through Lineage Through Time (LTT) plots (Harvey et al., 1994) and the Generalized Skyline Plots (GSP) (Strimmer and Pybus, 2001) methods. The Bayesian LTT was conducted through similar MCMC parameters as of the tree analyses and duplicated sequences were removed. The Bayesian GSP analyses were conducted on the entire data set, including duplicated sequences, a HKY substitution model and MCMC chains of 50×10^6 steps long sampled as described for LTT and tree reconstruction analyses.

2.5. Delineating ESU (Fig. 1, steps 2 & 3)

We examined the relationship between the genetic groups delineated by SAMOVA and the BINs (Fig. 1, Step 2) by testing the independence of each partitioning scheme using the Pearson chi-square test of independence as implemented in the R Stats ver. 3.1.2 package (R_Core_Team, 2014). The objective was to examine to what extent the SAMOVA groups of populations were determined by the BINs delineated by the RESL algorithm (Fig. 1, Step 2). We further examined the concordance of the BINs spatial distribution among species by performing a general hierarchical cluster analysis based on the average phylogenetic distance of the BINs among sites as implemented in hclust function of the R Stats ver. 3.1.2 package (R_Core_Team, 2014). The matrix of phylogenetic distance among sites was computed based on the composite chronogram of the three grafted maximum credibility trees of *Barbodes binotatus*, *Channa gachua* and *Glyptothorax platypogon* and BINs occurrence data using the R package PICANTE (Kembel et al., 2010). The composite chronogram was constructed by adding internal branches set to 0.001 Millions years between (i) the MRCA of *B. binotatus* and the MRCA of *G. platypogon*, (ii) the MRCA of both *B. binotatus* and *G. platypogon* and the MRCA of the three species. The internal branch between the MRCA of *C. gachua* and the MRCA of the three species was further adjusted to produce an ultrametric tree. BINs were considered as absent from sites out of their distribution range but BINs absence was coded as missing data at sites within their distribution range in order to account for sampling uncertainty *i.e.* a BIN present but not sampled (Table S3). We finally produced distribution maps of the population groups (Fig. 1, Step 1a), the BINs (Fig. 1, Step 1b) and the general hierarchical cluster (Fig. 1, Step 3). The general hierarchical cluster was furthered used to define phylogroups *i.e.* groups of sites hosting phylogenetically related BINs. The distribution of the phylogroups was used as a template to spot phylogeographic breaks, as exemplified by the phylogroup geographic boundaries, and compared to the distribution of SAMOVA groups and BINS to propose ESUs (Fig. 1, Step 3).

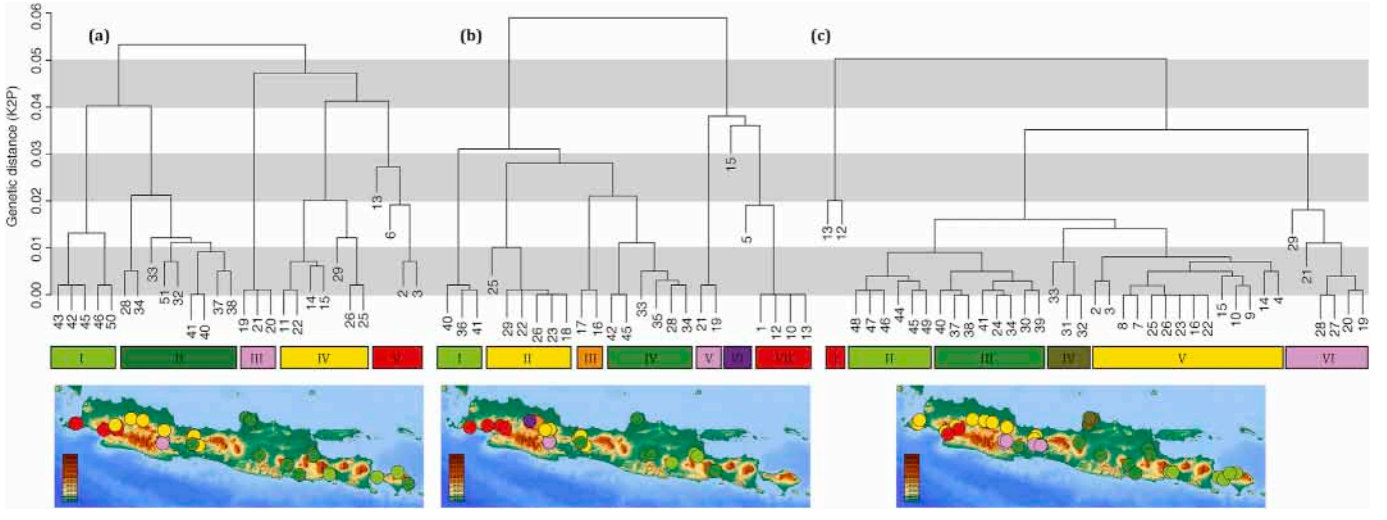

Fig. 3. Population genetic structure of *Barbodes binotatus* (**a**), *Channa gachua* (**b**) and *Glyptothorax platypogon* (**c**) as inferred from the hierarchical cluster analysis and SAMOVA.

3. Results

A total of 317 new sequences were successfully generated for the three species including 108 sequences of *Barbodes binotatus*, 99 sequences of *Channa gachua* and 110 sequences of *Glyptothorax platypogon*. Together with the 37 sequences previously published for these species (Dahruddin et al., 2017), a total of 354 sequences were analyzed including 122 sequences of *Barbodes binotatus*, 109 sequences of *Channa gachua* and 123 sequences of *Glyptothorax platypogon* (Table S1). All the sequences were above 500 bp of length and no codon stops were detected, suggesting that the sequences collected represent functional coding regions. The 354 sequences were collected from 39 sites for *B. binotatus*, 28 sites for *C. gachua* and 25 sites for *G. platypogon*.

3.1. Population genetic structure and molecular variance

A total of 5, 7 and 6 groups of populations were delineated for *Channa gachua, Glyptothorax platypogon* and *Barbodes binotatus* respectively (Table 1, Table S1). For the three species, most of the molecular variance is explained by differences among groups of populations with a percentage of the total variance ranging from 69.9 percent in *B. binotatus* to 78.75 percent in *G. platypogon*. 'Among populations within groups' is the second level in terms of the total variance explained with 24.69 percent for *C. gachua* and 11.34 percent for *G. platypogon*. *B. binotatus* differs from the two other species in having a slightly higher proportion of the total variance explained by differences within populations, with 19.05 percent, than among populations within groups, with 10.75 percent. The fixation indexes are significant at all spatial scales supporting spatially structured populations in the three species. This result was further confirmed by the hierarchical cluster analysis of the populations that indicates a substantial genetic differentiation of each group with K2P genetic distances among groups above 0.04, 0.02 and 0.01 for *C. gachua*, *G. platypogon* and *B. binotatus*, respectively (Fig. 3). Within groups of populations, haplotype diversity is high on average with *h* above 0.5 excepting for population groups II, V and VII of *G. platypogon* with *h* below 0.2 (Table 2). This high haplotype diversity is opposed to the low nucleotide diversity and the low mean-pairwise differences within groups for the three species.

3.2. Delineation and diversification of BINs

A total of 9, 7 and 3 molecular lineages (*i.e.* BINs) were delineated by the RESL algorithm on BOLD for *Channa gachua*, *Glyptothorax platypogon* and *Barbodes binotatus*, respectively (Table S1). Both maximum and average K2P distances are low on average within BINs and contrast with the high maximum K2P distances observed for each of the three coalescent trees (Table 3). Among the 19 BINs delineated, four were represented by singletons in *C. gachua* (ACQ3951, ACQ6941, ACQ6939, ACQ6940). The Bayesian analyses yielded three maximum credibility trees (Fig. 4a, b, 4C) that confirmed the contrast between the deep divergence among BINs, ranging from 0.47 to 2.75 Million years ago (Ma), and the shallow coalescent depth of the BINs ranging from 0.08 to 0.89 Ma (Table 3). The age of the MRCA was very similar among the three species ranging from 2.71 Ma for *B. binotatus* to 3.14 Ma *for G. platypogon* (Table 3). Plotting the BINs coalescent depth on the LTT curves further confirmed that the molecular diversity within BINs accumulated very recently within the three species (Fig. 4d, e, 4f). The Bayesian GSP

Table 2
Summary statistics of the genetic diversity including the sampling size (N), haplotype diversity (*h*), nucleotide diversity (π) and mean number of pairwise differences among haplotypes (mean-pairwise differences) for each of the population groups.

	N	*h*	π	mean-pairwise differences
Barbodes binotatus				
I	4	1	0.019	12.17
II	22	0.87	0.002	1.41
III	33	0.83	0.002	1.37
IV	12	0.55	0.005	2.73
V	36	0.82	0.004	2.35
VI	15	0.89	0.011	6.88
Channa gachua				
I	14	0.88	0.006	3.91
II	34	0.93	0.010	6.07
III	11	0.51	0.008	0.51
IV	36	0.81	0.009	5.85
V	14	0.58	0.011	6.97
Glyptothorax platypogon				
I	16	0.58	0.001	0.75
II	34	0.28	0.004	2.29
III	6	0.60	0.001	0.60
IV	39	0.75	0.006	3.62
V	10	0.20	0.002	1.20
VI	7	0.91	0.032	20.80
VII	11	0.18	0.001	0.18

Table 3
Summary statistics of the genetic K2P distances and age estimates. The maximum and average K2P distances are provided for each BINs and the maximum K2P distances are provided for the entire coalescent trees. The age estimate of the MRCA is provided based on three alternative hypotheses of molecular clock including 0.005 genetic divergence per million years (Hardman and Lundberg, 2006; Read et al., 2006), 0.012 genetic divergence per million years (Bermingham et al., 1997) and 0.02 genetic divergence per million years (Read et al., 2006). Divergence estimates are derived from the Bayesian analyses based on the H2 calibration with upper and lower bounds of the prior age intervals defined by the molecular clock hypotheses H1 and H3.

	K2P distance (%)		MRCA H1[a] (Myr)	MRCA H2[b] (Myr)	MRCA H3[c] (Myr)	Divergence estimates (Myr)	Coalescent depth (Myr)
	Maximum	Average					
Channa gachua							
Coalescent	6.78		6.78	2.83	1.70	2.75 (6.02-1.70)	
ACQ0292	2.1	0.87				2.04 (4.56-0.82)	0.66 (1.60-0.26)
ACQ0290	1.4	0.61				2.04 (4.56-0.82)	0.61 (1.49-0.16)
ACQ0291	0.15	0.08				2.75 (6.02-1.70)	0.08 (0.31-0.01)
ACQ3951	-	-				1.23 (2.87-0.41)	-
ACQ6941	-	-				0.73 (1.75-0.24)	-
ACQ6940	-	-				0.47 (1.17-0.24)	-
ACQ6939	0.00	0.00				0.47 (1.17-0.24)	-
ACQ3952	1.4	0.31				0.87 (2.08-0.31)	0.14 (1.40-0.14)
ACQ3950	0.93	0.32				0.87 (2.08-0.31)	0.33 (0.84-0.08)
Glyptothorax platypogon							
Coalescent	7.76		7.76	3.24	1.94	3.14 (6.74-1.94)	
ACQ5850	-	-				1.05 (2.68-0.21)	0.13 (0.45-0.01)
AAY1028	0.00	0.00				1.05 (2.68-0.21)	0.08 (0.33-0.01)
ACQ5898	0.93	0.23				1.75 (4.22-0.51)	0.39 (1.13-0.05)
ACP6225	0.31	0.12				1.98 (4.71-0.73)	0.23 (0.72-0.03)
ACQ6223	1.24	0.55				1.09 (2.63-0.32)	0.50 (1.35-0.14)
ACP6117	0.15	0.09				0.69 (1.75-0.18)	0.08 (0.34-0.01)
ACQ6224	0.46	0.05				0.69 (1.75-0.18)	0.21 (0.56-0.05)
Barbodes binotatus							
Coalescent	6.80		6.8	2.84	1.70	2.71 (5.96-1.70)	
ACP6290	0.01	0.00				2.71 (5.96-1.70)	0.12 (0.41-0.01)
ACP6025	0.57	1.09				1.74 (4.26-0.48)	0.38 (1.12-0.08)
ACP5712	1.94	0.73				1.74 (4.26-0.48)	0.89 (2.33-0.48)

[a] 0.005.
[b] 0.012.
[c] 0.02.

yielded very similar demographic trajectories for the three species with a global trend of constant population size over the last 2 millions years and a steep decline of population size during the last 100.000 years (Fig. 4g, h, 4i).

3.3. Identifying phylogroups and delineating ESUs

The null hypothesis of independence between population groups delineated by SAMOVA and BINs was rejected by the Pearson chi-square test ($X^2 = 4486.7$, p-value<0.001) suggesting that the delineation of population groups is largely supported by alternative distribution of BINs among populations. The general hierarchical cluster constructed using the phylogenetic distance among sites resulted in the delineation of five major phylogroups, diverging by more than 5 Million years on average, and ordered into two main clusters (Fig. 5). The geographic range of phylogroups I and II showed no overlap and allowed identifying three phylogeographic breaks (Fig. 5b). Phylogroup I is restricted to the western part of Java while phylogroup II shows an alternative distribution in eastern Java and Bali. The first phylogeographic break identified is located in the Java-Bali straight based on the segregation of phylogroups I and II on each side of the straight (Fig. 5, break 1). The second phylogeographic break identified is located in central Java, in between the western and central volcanic arches (Fig. 5, break 2). The third phylogeographic break is located on the western most slope of the western volcanic arch based on the western limit of phylogroup II (Fig. 5, break 3). The distribution ranges of phylogroups III, IV and V are more intricated but show some similarities with range distribution of phylogroups I and II. The Java-Bali straight is also identified as a phylogeographic break segregating phylogroups IV and V, each being alternatively distributed in Java or Bali (Fig. 5, break 1). The central Java phylogeographic break is also associated with the segregation of phylogroups III and V (Fig. 5, break 2) but two additional phylogeographic breaks are identified in central java, in between the western and central volcanic arches, that are associated with the segregation of phylogroup III from the phylogroup V (Fig. 5, break 4) and the segregation of an isolated site of phylogroup V from the eastern sites (Fig. 5, break 5). The western most phylogeographic break is also associated with the segregation of phylogroups IV and V in their western range but with some overlaps (Fig. 5, break 3).

The boundaries of population groups and BINs range distribution are generally associated with the five phylogeographic breaks for the three species with some exceptions (Fig. 6). Three cases of trans distribution across phylogeographic breaks are detected in *C. gachua* (Fig. 6a and b). The first case is located at the phylogeographic break 3 with the population group V (Fig. 6a, red). The population group V, however, is associated with four independent BINs. The second case is associated with the phylogeographic break 2 with the population group IV (Fig. 6a, yellow). Finally, the population groups I and II are

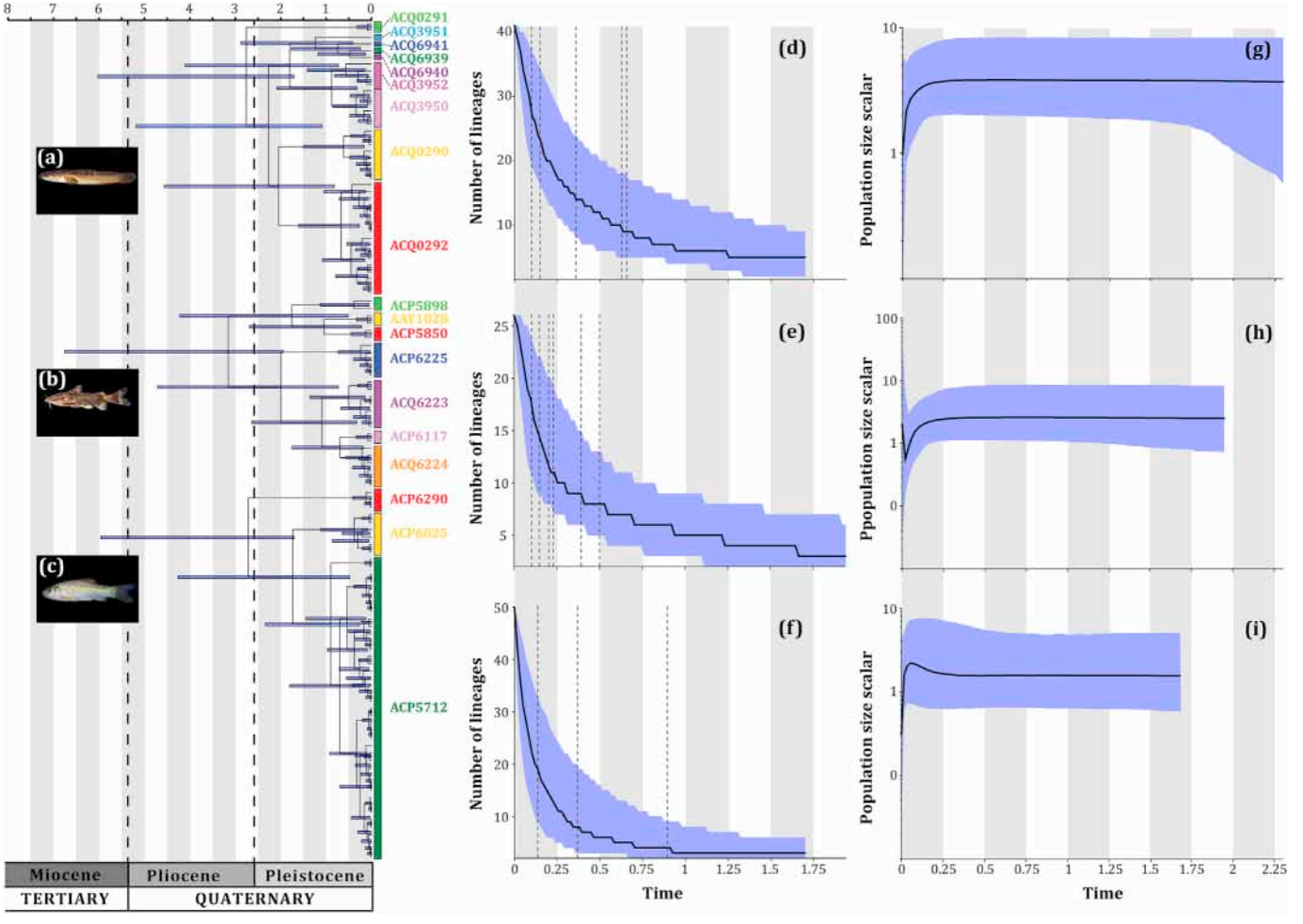

Fig. 4. BINs diversification patterns of *Channa gachua* (**a**, **d** and **g**), *Glyptothorax platypogon* (**b**, **e** and **h**) and *Barbodes binotatus* (**c**, **f** and **i**) including Maximum Credibility Trees (**a**, **b**, **c**); Lineage Through Time plot (**d**, **e**, **f**) and Bayesian Generalized Skyline Plots (**g**, **h**, **i**). Solid black line in **d**, **e** and **f** is the median diversity accumulation curve, blue shaded area is the 95% highest posterior density intervals and dotted lines represent the BIN coalescent depth in Million years. The solid black line in **g**, **h** and **i** represent the median effective population size, blue shaded area is the 95% highest posterior density intervals and Population size scalar (in millions) = effective population size x generation time. Calibrations are derived from previously published molecular clock hypotheses applied to the maximum K2P distance for each species - *i.e.* age of the MRCA (see Table 3). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

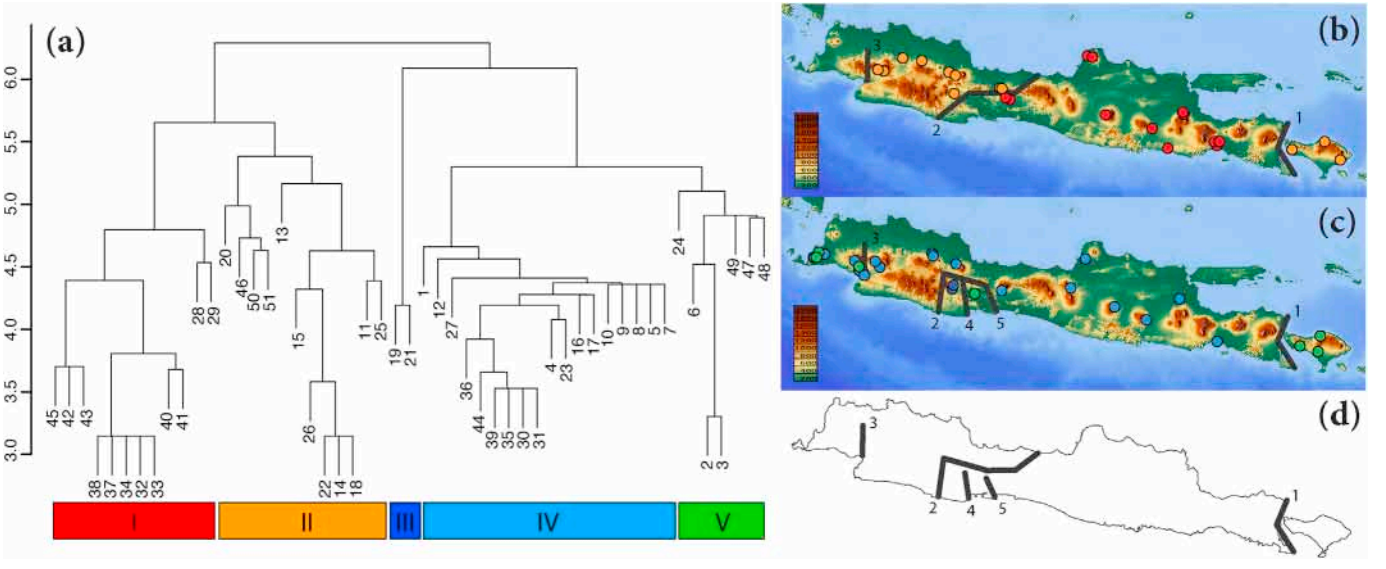

Fig. 5. Delineation of phylogroups and detection of phylogeographic breaks. **a** general hierarchical cluster based on the phylogenetic distances among sites established from the grafted chronograms of *C. gachua*, *G. platypogon* and *B. binotatus*, and occurrence data (see Table S2). **b** distribution range of phylogroups I and II, and associated phylogeographic breaks (solid black lines). **c** distribution range of phylogroups III, IV and V, and associated phylogeographic breaks (solid black lines). **d** phylogeographic breaks identified from the phylogroup geographic boundaries and further used for delineating ESUs.

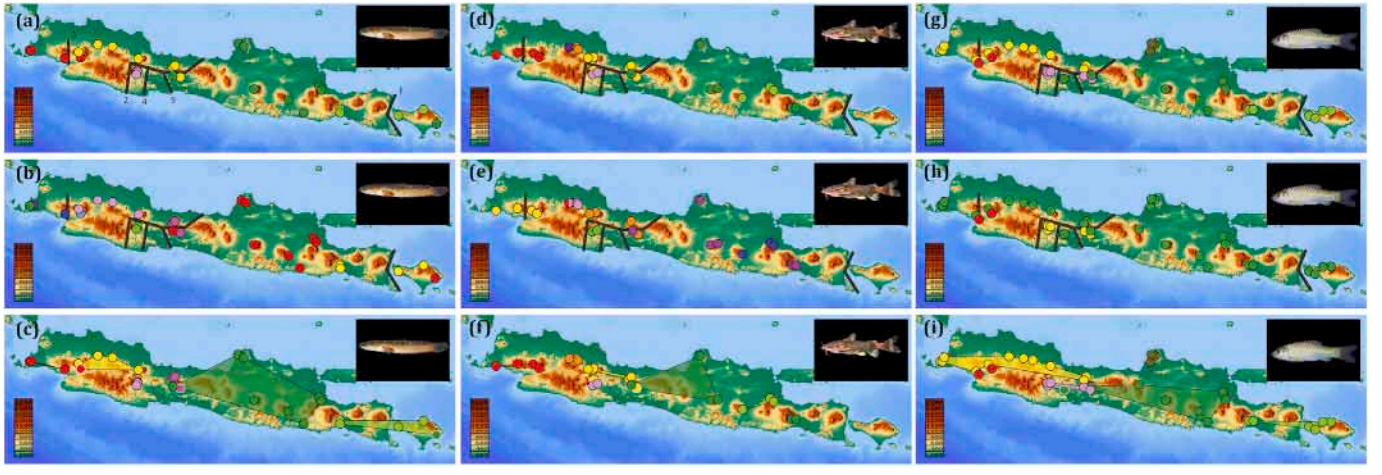

Fig. 6. Mapping of the population groups identified by SAMOVA (**a**, **d** and **g**), the BINs delineated by the RESL algorithm (**b**, **e** and **h**) and ESUs candidates (**c**, **f** and **i**) for *C. gachua* (**a**, **b** and **c**), *G. platypogon* (**d**, **e** and **f**) and *B. binotatus* (**g**, **h** and **i**). Ambiguous assignment of populations to ESUs due to trans distribution across phylogeographic breaks are highlighted by solid white circle (**c**, **f** and **i**).

distributed on each side of the phylogeographic break 1 (Fig. 6a, green groups). The two cases of trans distribution across phylogeographic breaks 1 and 2 are also observed for *G. platypogon* with population group VII overlapping break 1 (Fig. 6d, red) and population group II overlapping break 2 (Fig. 6d, yellow). In addition, two groups of population corresponding to two distinct BINS in Eastern Java are detected that are not associated with a phylogeographic break (Fig. 6d, green groups; Fig. 6e, purple and blue). Trans distribution across the phylogeographic breaks 1 and 3 are also observed in *B. binotatus* for population groups II and III (Fig. 6g, green groups) and population group V (Fig. 6g, yellow), respectively. Other trans distributions are also observed for the population group VI (Fig. 6g, pink) across the phylogeographic breaks 4 and 5, and the population group III (Fig. 6g, dark green) across the phylogeographic breaks 2 and 5. Similarly with *G. platypogon*, additional population groups not associated with previously identified phylogeographic breaks are detected for *B. binotatus* (Fig. 6g, green groups). These populations groups, however, do not match BINs boundaries as population groups II, III and IV belong to the same BIN.

Following the identified match between phylogeographic breaks, population groups and BINs, 6 ESUs are proposed for *C. gachua* (Fig. 6c), *G. platypogon* (Fig. 6f) and *B. binotatus* (Fig. 6i) that show similar range distribution. In the three species, most population groups and BINs show great overlap in their distribution. Assigning sites to ESUs was conflicting in only a few cases associated with the cases of trans distribution above mentioned (Fig. 6c, f and 6i; solid white circle). Following the theoretical framework of the present study (Fig. 1), populations associated with trans distribution were assigned to ESUs according to their relative position to phylogeographic breaks. Along the same line, ESUs were delineated based on either population groups or BINs whenever each grouping scheme provided conflicting sorting *e.g.* two population groups within the same BIN or several BINs within a population group. A single exception was done to this general principle for the BINs in *C. gachua* represented by singletons (Fig. 6b; purple, dark blue, light blue, dark green) and associated to the same population group (Fig. 6a; red) that were grouped into a single ESU.

4. Discussion

4.1. Regional patterns and evolutionary dynamics

Our study emphasizes how the joint use of population clustering and molecular lineage delineation methods enhances our understanding of population evolutionary dynamics and resulting patterns of genetic structure. Population dynamics are ruled by stochastic and deterministic mechanisms acting at varying scales. While deterministic mechanisms, such as fragmentation through geological dynamics, are usually predominant in shaping population structure by limiting gene flow (Hanski, 1991; Hastings and Harrison, 1994; Moilanen and Hanski, 1998), stochastic dynamics predominate at local scale within populations or tightly connected populations where they determine genetic diversity through genetic drift (Vellend, 2005; Vellend and Geber, 2005; Hubert et al., 2015a). Landscapes are dynamic, particularly so in insular systems (Warren et al., 2014), and deterministic mechanisms may predominate during major landscape changes while stochastic mechanisms may take over during equilibrium dynamics (Alonso et al., 2006; Hubert et al., 2015a).

The present study illustrates the benefits of comparing patterns of population structure across multiple species to disentangle the relative contribution of deterministic and stochastic mechanisms on population genetic structure, particularly so in the absence of *a priori* biogeographic information. The global concordance in the population genetic structure of *B. binotatus*, *C. gachua* and *G. platypogon* argue for a predominant influence of deterministic processes in shaping population structure. This trend could be expected considering the volcanic origin of Java and Bali islands and the stringent constraint of geological history on species range evolution (Lohman et al., 2011). For instance, Java originated from the merging of two volcanic arches, one is the West and a second in the east including East Java and Bali, that emerged between 10 and 5 Ma and further aggregated during the last 5 Million years. This scenario was reflected in the closer genetic affinities between the central-eastern Java and Bali population groups than between the central-eastern and the western population groups in Java of *C. gachua* and *B. binotatus*. This pattern supported the identification of the phylogeographic break 2 located in between the central and western volcanic arches of Java.

The orogeny of Java started at 10 million years and landscapes likely achieved their modern configuration during the last 5 million years (Lohman et al., 2011). The ancient history of Java landscapes is reflected in the lack of independence between the groups of population and the delineated molecular lineages, stressing that the population structures detected have been established for a long time. This trend was further reflected in the large difference between genetic distances and age estimates within and among population groups. The spatial structures depicted here and inferred timing of population settlement further highlight the limited dispersal opportunities for aquatic biotas during the geological history of Java and Bali. This trend suggest that the colonization of Java and Bali rivers by primary freshwater fishes happened during the earliest geological stage, as exemplified by the deep divergence among BINs, with a subsequent fragmentation of the populations rather than multiple colonization through time. This assumption is further supported by the similar age estimates of the BINs in between 750.000 and 150.000 years ago, suggesting a synchronous fragmentation of the populations. Along the same line, multiple colonization events through range expansion are frequently associated with a signature of past population growth (Slatkin and Hudson, 1991; Rogers and Harpending, 1992; Harpending et al., 1998). The inferred demographic trajectories, and their remarkable similarity for the three species, do not support a scenario of colonization through range expansion as constant population sizes are inferred during the last 2 Million years.

4.2. Implications for the conservation of aquatic ecosystems in South Sundaland

The settlement of Sundaland results from intricate interactions between plate tectonics and eustatic changes (Kottelat, 1989; Kottelat et al., 1993; Woodruff, 2010; Lohman et al., 2011). The emergence of Sundaland resulted from the subduction of the Australian plate beneath Sundaland during the last 10 Ma (Lohman et al., 2011; Hall, 2013), a process initiated with the isolation of Borneo and continued with the emergence of Sumatra and Java from 5 Ma onward. Once the maximum depth of the Java sea reached nearly 120 m, however, sea levels fluctuations associated with Milankovitch cycle (Hays et al., 1976) started to interfere with the orogeny of Sundaland islands (Kottelat et al., 1993; Woodruff, 2010). During sea level low-stands, Sundaland islands were connected to each others, prompting the settlement of four large palaeodrainages straddling across islands (Kottelat, 1989; Kottelat et al., 1993). Such extended watershed surface happened repeatedly during the late Pleistocene and as a consequence, Sundaland ecosystems are currently in a refugial state, occupying only 50–75% of their maximal Pleistocene extent (Woodruff, 2010).

The demographic inferences produced here, as well as population genetic diversity estimates, are consistent with a scenario of biome contraction during the Pleistocene. Steep declines in population size are detected in *C. gachua*, *G. platypogon* and *B. binotatus* and their inferred time-frame are consistent with a late Pleistocene fragmentation of aquatic biotas in Sundaland. The shallow coalescent trees of the population groups and low nucleotide diversity further support that populations experienced bottleneck in the past. The high haplotype diversity and observed reciprocal monophyly among most population groups, however, argue that population size reduction is historical and populations are recovering. This observation is suggested by the demographic inference of *G. platypogon* where a recent population growth is detected. This trend, however, is not observed in *B. binotatus* and *C. gachua* and considering the notorious difficulties to disentangle multiple demographic events in structured populations (Markovtsova et al., 2000; Heller et al., 2013), the recovery state of the populations remain to be explicitly addressed.

The present study highlight several properties of the ESUs defined here that provides some guidelines for future conservation plans of the Javanese ichthyofauna. First, each ESU display substantial levels of genetic divergence, with divergence age estimates beyond 1 Million years. The population genetic structure inferred argue for a differentiation in allopatry and considering that these BINs have been consistently assigned to the same nominal species by taxonomists (Kottelat, 2013; Ng and Kottelat, 2016), they show conserved morphological attributes. This trend could be expected when differentiation happens in allopatry across homogeneous environmental and biotic conditions (Hubert et al., 2015a). These large levels of genetic divergence, however, question the reproductive status of those ESUs as genomic incompatibilities may accumulate through time, despite conserved eco-morphological characteristics, and produce post-zygotic isolation (Orr and Turelli, 2001; Brideau et al., 2006). Second, these ESUs show remarkably restricted range distribution, particularly so in Western Java, and low genetic diversity at the nucleotide level. Considering the demographic trend of population contraction observed in the three species, these ESUs should be treated as small populations, particularly sensitive to further size reduction that may result from anthropogenic perturbations. Nevertheless, Java is currently the most densely populated island in Sundaland and the reduction of aquatic habitats has been severe during the last two decades. This situation is clearly of concern as it is now widely acknowledged that important shifts in life history traits (LHT) and reduction of genetic diversity are precursor of extinction vortex - reduced population size and increased demographic variance induce either a spatial fragmentation or a decrease of population adaptive potential (Gilpin and Soulé, 1986; Fagan and Holmes, 2006). These two characteristics of the 18 ESUs call for a specific assessment of their genetic exchangeability if breeding programs were to be considered. In particular, the success of restoration program through genetic rescue, aiming at increasing population fitness through the repletion of genetic diversity by immigration, is tightly dependent on the exchangeability of the immigrants (Whiteley et al., 2014). From an ecological perspective, the 18 ESUs show no pattern of altitudinal zonation (Fig. S1) that usually results from adaptive changes to heterogeneous thermal regimes and biotic interactions (Angilletta et al., 2006). This was expected considering the commonness of these three species during the field sampling. From a genetic perspective, however, the high level of genetic divergence among BINs calls for a conservative approach in breeding program that focus on maintaining the genetic integrity of the ESUs identified here. A few tenth of generation might be sufficient to accumulate adaptive combinations of alleles for selected genes and disrupting those allelic combinations may result in significant decreases of fitness (Whiteley et al., 2014). Before further evidences are collected from ecological and population genetic perspectives, the present study argue against any translocation program at large spatial scale for those species that, if not accounting for the deep population structure observed, may have dramatic consequences on populations fitness.

4.3. Validation of the approach and generalization

Defining biological units for conservation purposes is challenging due to the multiple interactions that drive population dynamics in nature (Vellend and Geber, 2005; Urban et al., 2008; Vellend, 2010; Hubert et al., 2015a). As a consequence, the concept of ESUs has been subject of much debate since its earliest development (Crandall et al., 2000). Preserving populations genetic diversity is a mandatory step for successful conservation or restoration programs as genetic diversity has direct consequences on ecological characteristics (Hughes et al., 2008) and preserving genetic diversity usually implies accounting for population genetic structure in delineating ESUs that comprehensively cover species genetic diversity (Waples, 1991, 1995; Dizon et al., 1992; Moritz, 1994; Waples, 1995; Fraser and Bernatchez, 2001). The present study proposes a heuristic

framework that enables the delineation of ESUs in the absence of *a priori* phylogeographic knowledge through the joint use of population clustering and molecular delineation methods across multiple species.

The three species present marked population genetic structure, a trend that may constitute a particularity of the Java and Bali aquatic systems. The concordance between the groups of population defined by SAMOVA and the BINs delineated by the RESL algorithm argue that most population groups achieved reciprocal monophyly. In this context, population groups may be treated as independent cryptic lineages instead of groups of population, a trend that has been previously observed for several lineages of freshwater fishes in Sundaland (Nguyen et al. 2008; Pouyaud et al. 2009; De Bruyn et al., 2013). Such system, however, exemplify the potential of a community-level assessment of population genetic structure through a fast method of molecular screening such as DNA barcoding. First, the concordance between the gene genealogies of *C. gachua*, *B. binotatus* and *G. platypogon* helped identify common trends in population grouping and associated phylogeographic breaks in the absence of *a priori* phylogeographic knowledge. This general trend further enabled the delineation of ESUs based on a newly established biogeographic background that helped refine population assignments in case of conflicting grouping among species. Second, this approach allowed to identify populations that may have a specific status, such as populations located in secondary contact zone and having experienced hybridization and introgression. Several cases were identified here with populations that have been assigned to a different ESU based on their trans distribution across phylogeographic breaks. Further addressing the status of these populations may gather important information about their reproductive status. For instance, none of the populations analyzed in the present study harbor more than one BIN, excepting one population of *G. platypogon*, despite the mosaic distribution across several phylogeographic breaks suggesting a dynamic of secondary contact. Third, the present study is based on a fast method of molecular screening based on mitochondrial DNA. The ease of amplification of COI sequences for fishes (Ward et al., 2009; April et al., 2011; Hubert et al., 2012; Pereira et al. 2013) offers a fast and effective heuristic approach to the exploration of population genetic structure and ESUs delineation (Hajibabaei et al., 2007). This approach is thus particularly welcomed when phylogeographic and taxonomic knowledge is scarce and the extent of anthropogenic threat on biodiversity urge for an objective assessment of conservation priorities.

The effectiveness of our approach for delineating ESUs in the present system is due in part to the deep population genetic structure of the species analyzed. As such, the usefulness of the method remains to be explored for less fragmented and open systems. Insular biotas are frequently built upon *in situ* diversification instead of immigration, as observed in continental systems, and evolutionary diversification happen at smaller spatial scales (Emerson and Gillespie, 2008; Warren et al., 2014; Hubert et al., 2015a). The usefulness of the present method depends on the biogeographic context that dictates the spatial scale of diversification and population genetic structure (Holt, 1993; Ricklefs and Schluter, 1993). In opened systems, shared patterns of population genetic structure may not be a common trend and ESU should be delineated based on an individual basis as each species may have experienced distinct evolutionary history (Bowen et al., 2016). Along the same line, shallow population genetic structure may be determined by subtle variations in haplotype frequencies and substitution rates at COI might not be sufficient to reliably detect groups of population. In that case, alternative approaches should be favored based on the assessment of molecular markers displaying higher levels of polymorphism, such as those resulting from length polymorphism (Selkoe and Toonen, 2006). In addition, the maternal inheritance of mitochondrial DNA is a further limit for assessing gene flow, particularly so when subtle variations in haplotypes frequencies are observed (Nielsen and Wakeley, 2001). Here, both groups of population and molecular lineages were determined based on the same data set. The present framework, however, is flexible and both population grouping and molecular lineages may be delineated based on independent data sets. In addition, the general hierarchical cluster was established based on phylogenetic relationships of the BINs as the groups of population were largely delineated based on BINs boundaries. The general cluster may be established based on population groups instead of BINs providing that both delineating schemes are independent and shallow population genetic structure are observed. In this case, variations in allelic frequencies might be more successful at identifying common phylogeographic breaks, particularly so if population genetic structures established recently.

5. Conclusion

While we are facing a major biodiversity crisis, fast, cheap and universal methods for delineating ESUs are needed. The inventory of earth diversity and our understanding of its origin through space and time are still fragmentary, particularly in the tropics where the taxonomic impediment dramatically hampers conservation (Garnett and Christidis, 2017). DNA barcoding has already established as a new reference for species identification and its utility for the inventory of tropical biotas is increasingly acknowledged (Hebert et al., 2004; Smith et al., 2005, 2007; Monaghan et al., 2009; Hubert et al., 2012; Tänzler et al., 2012; Riedel et al., 2013). Its application for the delineation of ESUs, however, has been much less explored. The present study suggests that DNA barcoding may be successfully used for this purpose, providing that the limits of the mitochondrial genome, due to its maternal inherence, are balanced by the comparison of population genetic structure across multiple species. In the context of Java and Bali aquatic biotas, our approach helped identify ESUs in the absence of prior phylogeographic knowledge and produce recommendations for the conservation of the Java and Bali freshwater fishes.

Acknowledgements

The study was designed by R.H., B.S., P.K., R.K.H. and N.H. These results were obtained in the framework of AH master degree under the supervision of N.H and B.S. Fieldwork was carried out by A.H., H.D., S.S., F.B., R.K.H., P.K. and N.H. S.S and H.D.

prepared and deposited the specimens in museum collection, N.H. and R.K.H. identified the specimens and R.H. supervised the sequencing. A.H. and N.H. analyzed the data and wrote the manuscript, and all authors contributed to revisions.

The authors wish to thank Siti Nuramaliati Prijono, Witjaksono, Mohammad Irham, Marlina Adriyani, Ruliyana Susanti, Hari Sutrisno and Sri Sulandari, at Research Centre for Biology (RCB-LIPI); Jean-Paul Toutain, Robert Arfi and Jean-François Agnèse from the 'Institut de Recherche pour le Développement'; Joel Le Bail and Nicolas Gascoin at the French embassy in Jakarta for their continuous support. We are thankful to Daisy Wowor and Ujang Nurhaman at RCB-LIPI, Sumanta and Bambang Dwisusilo at IRD Jakarta for their help during the field sampling in Java and Bali. Part of the present study was funded by the Institut de Recherche pour le Développement (UMR226 ISE-M and IRD through incentive funds), the MNHN (UMR BOREA), the RCB-LIPI, the French Ichthyological Society (SFI), the Foundation de France and the French embassy in Jakarta. The Indonesian Ministry of Research approved this study and field sampling was conducted according to the research permits 097/SIP/FRP/SM/IV/2014 for Philippe Keith, 60/EXT/SIP/FRP/SM/XI/2014 for Frédéric Busson and 41/EXT/SIP/FRP/SM/VIII/2014 for Nicolas Hubert. Sequence analysis was aided by funding from the government of Canada through Genome Canada and the Ontario Genomics Institute in support of the International Barcode of Life project. We thank Paul Hébert and Evgeny Zakharov as well as BOLD and CCDB staffs at the University of Guelph for their valuable support. This publication has ISEM number 2017-278 SUD.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.gecco.2017.11.005.

References

Alonso, D., Etienne, R.S., McKane, A.J., 2006. The merits of neutral theory. Trends Ecol. Evol. 21, 451–457.
Anderson, C.D., Epperson, B.K., Fortin, M.-J., Holderegger, R., James, P., Rosenberg, M.S., Scribner, K.T., Spear, S., 2010. Considering spatial and temporal scale in landscape-genetic studies of gene flow. Mol. Ecol. 19, 3565–3575.
Angilletta, M.J.J., Bennett, A.F., Guderley, H., Navas, C.A., Seebacher, F., Wilson, R.S., 2006. Coadaptation: a unifying principle in evolutionary thermal biology. Physiological Biochem. Zool. 79, 282–294.
April, J., Mayden, R., Hanner, L.,R.H., Bernatchez, L., 2011. Genetic calibration of species diversity among North America's freshwater fishes. Proc. Natl. Acad. Sci. U. S. A. 108, 10602–10607.
Avise, J.C., 1989. Molecular Markers, Natural History and Evolution. New York.
Avise, J.C., 2000. Phylogeography: the History and Formation of Species. Harvard University Press.
Avise, J.C., Bowen, B.W., Ayala, F., 2016. In the light of evolution X: comparative phylogeography. Proc. Natl. Acad. Sci. U. S. A. 113, 7957–7961.
Beheregaray, L.B., 2008. Twenty years of phylogeography: the state of the field and the challenges for the Southern hemisphere. Mol. Ecol. 17, 3754–3774.
Bermingham, E., McCafferty, S., Martin, A.P., 1997. Fish biogeography and molecular clocks: perspectives from the Panamanian isthmus. In: Kocher, T.D., Stepien, C.A. (Eds.), Molecular Systematics of Fishes. CA Academic Press, San Diego, pp. 113–128.
Bowen, B.W., Gaither, M.R., DiBattista, J.D., Iacchei, M., Andrews, K.R., Grant, W.S., Toonen, R.J., Briggs, J.C., 2016. Comparative phylogeography of the ocean planet. Proc. Natl. Acad. Sci. U. S. A. 113, 7962–7969.
Brideau, N.J., Flores, H.A., Wang, J., Maheshwari, S., Wang, X., Barbash, D.A., 2006. Two dobzhansky-Muller genes interact to cause hybrid lethality in Drosophila. Science 314, 1292–1295.
Crandall, K.A., Bininda-Emonds, O.R.P., MAce, G.M., Wayne, R.K., 2000. Considering evolutionary processes in conservation biology. Trends Ecol. Evol. 15, 290–295.
Dahruddin, H., Hutama, A., Busson, F., Sauri, S., Hanner, R., Keith, P., Hadiaty, R.K., Hubert, N., 2017. Revisiting the ichthyodiversity of Java and Bali through DNA barcodes: taxonomic coverage, identification accuracy, cryptic diversity and identification of exotic species. Mol. Ecol. Resour. 17, 288–299.
Darriba, D., Taboada, G.L., Doallo, R., Posada, D., 2012. jModelTest 2: more models, new heuristics and parallel computing. Nat. Methods 9, 772.
De Bruyn, M., Rüber, L., Nylinder, S., Stelbrink, B., Lovejoy, N.R., Lavoué, S., Heok Hui, T., Nugroho, E., Wowor, D., Ng, P.K.L., Siti Azizah, M.N., Von Rintelen, T., Hall, R., Carvalho, G.R., 2013. Paleo-drainage basin connectivity predicts evolutionary relationships across three Southeast asian biodiversity hotspots. Syst. Biol. 62, 398–410.
Dizon, A.E., Lickyer, C., Perrin, W.F., et al., 1992. Rethinking th stock concept: a phylogeographical approach. Conserv. Biol. 6, 24–36.
Drummond, A., Suchard, M., Xie, D., Rambault, A., 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol. Biol. Evol. 29.
Dupanloup, I., Schneider, S., Excoffier, L., 2002. A simulated annealing approach to define the genetic structure of populations. Mol. Ecol. 11, 2571–2581.
Emerson, B.C., Gillespie, R.G., 2008. Phylogenetic analysis of community assembly and structure over space and time. Trends Ecol. Evol. 23, 519–530.
Endo, Y., Nash, M., Hoffmann, A.A., Slatyer, R., Miller, A.D., 2014. Comparative phylogeography of alpine invertebrates indicates deep lineage diversification and hisgorical refugie in the Australian Alps. J. Biogeogr. 42, 89–102.
Excoffier, L., Smouse, P., 1994. Using allele frequencies and geographic subdivision to reconstruct gene genealogies within a species. Molecular variance parsimony. genetics 136, 343–359.
Excoffier, L., Laval, G., Schneider, S., 2005. Arlequin ver. 3.0: an integrated software package for population genetics data analysis. Evol. Bioinforma. online 1, 47–50.
Excoffier, L., Smouse, P., Quattro, J.M., 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. genetics 131, 479–491.
Fagan, W.F., Holmes, E.E., 2006. Quantifying the extinction vortex. Ecol. Lett. 9, 51–60.
Fraser, D.J., Bernatchez, L., 2001. Adaptative evolutionary conservation: toward a unified concept for defining conservation units. Mol. Ecol. 10, 2741–2752.
Froese, R., Pauly, D., 2011. Fishbase. Worldwide web electronic publication. http://www.fishbase.org. version (06/2011).
Fu, Y.-X., 1999. Coalescing into the 21st Century: an overview and prospects of coalescent theory. Theor. Popul. Biol. 56, 1–10.
Garnett, S.T., Christidis, L., 2017. Taxonomy anarchy hampers conservation. Nature 546, 25–27.
Gibson, J., Shokralla, S., Porter, T.M., King, I., van Konynenburg, S., Janzen, D.H., Hallwachs, W., Hajibabaei, M., 2014. Simultaneous assessment of the macrobiome and microbiome in a bulk sample of tropical arthropods through DNA metasystematics. Proc. Natl. Acad. Sci. 111, 8007–8012.
Gilpin, E., Soulé, M., 1986. Minimum viable populations: processes of species extinction. In: Soulé, M.E. (Ed.), Conservation Biology: the Science of Scarcity and Diversity. Sinauer, Sunderland, Mass, pp. 19–34.
Gutiérrez-Tapia, P., Palma, R.E., 2016. Integrating phylogeography and species distribution models: cryptic distributional reponses to past climate change in an endemic rodent from central Chile hotspot. Divers. Distrib. 22, 638–650.
Hajibabaei, M., Shokralla, S., Zhou, X., Singer, G.A.C., Baird, D.J., 2011. Environmental barcoding: a next-generation sequencing approach for biomonitoring applications using river benthos. PLoS One 6, e17497.

Hajibabaei, M., Singer, G.A.C., Hebert, P.D.N., Hickey, D.A., 2007. DNA barcoding: how it complements taxonomy, molecular phylogenetics and population genetics. Trends Ecol. Evol. 23 (4), 167–172.
Hall, R., 2013. The palaeogeography of Sundaland and Wallacea since the late jurassic. J. Limnol. 72, 1–17.
Hanski, I., 1991. Single-species metapopulation dynamics: concepts, models and observations. Biol. J. Linn. Soc. 42, 17–38.
Hardman, M., Lundberg, J.G., 2006. Molecular phylogeny and chronology of diversification for "phractocephaline" catfishes (Siluriformes: Pimelodidae) based on mitochondrial DNA and DNA recombination activating gene 2 sequences. Mol. Phylogenet. Evol. 40, 410–418.
Harpending, H., Batzer, M., Geurven, M., Jorde, L., Rogers, A., Sherry, S., 1998. Genetic traces of ancient demography. Proc. Natl. Acad. Sci. 95.
Harvey, P.H., May, R.M., Nee, S., 1994. Phylogenies without fossils. Evolution 48, 523–529.
Hastings, A., Harrison, S., 1994. Metapopulation dynamics and genetics. Annu. Rev. Ecol. Syst. 25, 167–188.
Hays, J.D., Imbrie, J., Shackleton, N.J., 1976. Variations in the earth's orbit: pacemaker of the ice ages. Science 194, 1121–1132.
Hebert, P.D.N., Gregory, T.R., 2005. The promise of DNA barcoding for taxonomy. Syst. Biol. 54, 852–859.
Hebert, P.D.N., Penton, E.H., Burns, J.M., Janzen, D.H., Hallwachs, W., 2004. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly Astraptes fulgerator. Proc. Natl. Acad. Sci. U. S. A. 101, 14812–14817.
Heller, R., Chikhi, L., Siegismund, H.R., 2013. The confounding effect of population structure on Bayesian skyline plot inferences of demographic history. Plos One 8, e62992.
Hoffman, M., Hilton-Taylor, C., Angulo, A., Böhm, M., Brooks, T.M., Butchart, S.H.M., Carpenter, K.E., Chanson, J., Collen, B., Cox, N.A., et al., 2010. The impact of Conservation on the status of the world's vertebrates. Science 330, 1503–1509.
Holderegger, R., Wagner, H.H., 2006. A brief guide to landscape genetics. Landsc. Ecol. 21, 793–796.
Holt, R., 1993. Ecology at the mesoscale: the influence of regional processes on local communities. In: Ricklefs, R., Schluter, D. (Eds.), Species Diversity in Ecological Communities: Historical and Geographical Perspectives. University of Chicago Press, Chicago.
Hua, X., Wiens, J.J., 2010. Latitudinal variation in speciation mechanisms in frogs. Evolution 64, 429–443.
Hubert, N., Calcagno, V., Etienne, R.S., Mouquet, N., 2015a. Metacommunity speciation models and their implication for diversification theory. Ecol. Lett. 18, 864–881.
Hubert, N., Hanner, R., 2015. DNA barcoding, species delineation and taxonomy: a historical perspective. DNA Barcodes 3, 44–58.
Hubert, N., Kadarusman, A., Wibowo, F., Busson, Caruso, D., Sulandari, S., Nafiqoh, N., Rüber, L., Pouyaud, L., Avarre, J.C., Herder, F., Hanner, R., Keith, P., Hadiaty, R.K., 2015b. DNA barcoding Indonesian freshwater fishes: challenges and prospects. DNA Barcodes 3, 144–169.
Hubert, N., Meyer, C., Bruggemann, J.H., Guérin, F., Komeno, R.J.L., Espiau, B., Causse, R., Williams, J.T., Planes, S., 2012. Cryptic diversity in Indo-Pacific coral reef fishes revealed by DNA-barcoding provides new support to the centre-of-overlap hypothesis. Plos One 7, e28987.
Hudson, R., 1982. Testing the constant rate neutral allele model with protein sequence data. Evolution 37, 203–217.
Hughes, A.R., Inouye, B.D., Johnson, M.T.J., Underwood, N., Vellend, M., 2008. Ecological consequences of genetic diversity. Ecol. Lett. 11, 609–623.
Ivanova, N.V., Zemlak, T.S., Hanner, R.H., Hébert, P.D.N., 2007. Universal primers cocktails for fish DNA barcoding. Mol. Ecol. notes 7, 544–548.
Janzen, D.H., Hajibabaei, M., Burns, J.M., Hallwachs, W., Remigio, E., Hebert, P.D.N., 2005. Wedding biodiversity inventory of a large and complex lepidoptera fauna with DNA barcoding. Phil. Trans. R. Soc. Serie B 360, 1835–1845.
Kekkonen, M., Hebert, P.D.N., 2014. DNA barcode-based delineation of putative species: efficient start for taxonomic workflows. Mol. Ecol. Resour. 14, 706–715.
Kembel, S.W., Cowan, P.D., Helmus, M.R., Cornwell, W.K., Morlon, H., Ackerly, D.D., Blomberg, S.P., Webb, C.O., 2010. Picante: R tools for integrating phylogenies and ecology. Bioinformatics 26, 1463–1464.
Kingman, J.F.C., 1982. The coalescent. Stoch. Process Appl. 13, 245–248.
Kottelat, M., 1989. Zoogeography of the fishes from Indochinese inland water with annotated checklist. Bull. Zool. Mus. Univ. Van Amst. 12, 1–55.
Kottelat, M., 2013. The fishes of the inland waters of Southeast Asia: a catalog and core bibliography of the fishes known to occur in freshwaters, mangroves and estuaries. Raffles Bull. Zool. Suppl. 27, 1–663.
Kottelat, M., Whitten, A.J., Kartikasari, S.R., Wirjoatmodjo, S., 1993. Freshwater Fishes of Western Indonesia and Sulawesi. Periplus editions, Singapore.
Lamoreux, J.F., Morrison, J.C., Ricketts, T.H., Olson, D.M., Dinerstein, E., McKnight, M., Shugart, H.H., 2006. Global tests of biodiversity concordance and the importance of endemism. Nature 440, 212–214.
Lohman, K., De Bruyn, M., Page, T., Von Rintelen, K., Hall, R., Ng, P.K.L., Shih, H.-T., Carvalho, G.R., Von Rintelen, T., 2011. Biogeography of the Indo-Australian archipelago. Annual Review Ecolo. Evol. Syst. 42, 205–226.
Markovtsova, L., Marjoram, P., Tavaré, S., 2000. The effects of rate variation on ancestral inference in the coalescent. genetics 156, 1427–1436.
Mittelbach, G.G., Schemske, D.W., Cornell, H.V., Allen, A.P., Brown, J.M., Bush, M.B., Harrison, S.P., Hurlbert, A.H., Knowlton, N., Lessios, H.A., McCain, C.M., McCune, A.R., McDade, L.A., McPeek, M.A., Near, T.J., Price, T.D., Ricklefs, R.E., Roy, K., Sax, D.F., Schluter, D., Sobel, J.M., Turelli, M., 2007. Evolution and the latitudinal diversity gradient: speciation, extinction and biogeography. Ecol. Lett. 10, 315–331.
Moilanen, A., Hanski, I., 1998. Metapopulation dynamics: effects of habitat quality and landscape structure. Ecology 79, 2503–2515.
Monaghan, M.T., Wild, R., Elliot, M., Fujisawa, T., Balke, M., Inward, D., Lees, D.C., Ranaivosolo, R., Eggleton, P., Barraclough, T.G., Vogler, A.P., 2009. Accelerated secies inventories on Madagascar using coalescent-based models of species delineation. Syst. Biol. 58, 298–311.
Moritz, C., 1994. Defining 'evolutionary significant units' for conservation. Trends Ecol. Evol. 9, 373–375.
Myers, N., Mittermeier, R.A., Mittermeier, C.G., da Fonseca, G.A.B., Kent, J., 2000. Biodiversity hotspots for conservation priorities. Nature 403, 853–858.
Ng, H.H., Kottelat, M., 2016. The Glyptothorax of Sundaland: a revisionary study (Teleostei: Sisoridae). Zootaxa 4188, 001–092.
Nguyen, T.T.T., Na-Nakorn, U., Sukmanomon, S., ZiMing, C., 2008. A study on phylogeny and biogeography of mahseer species (Pisces: Cyprinidae) using sequences of three mitochondrial DNA gene regions. Mol. Phylogenet. Evol. 48, 1223–1231.
Nielsen, R., Wakeley, J., 2001. Distinguishing migration from isolation: a markov chain approach. genetics 158, 885–896.
Orr, H.A., Turelli, M., 2001. The evolution of post-zygotic isolation: accumulating Dobzhansky-Muller incompatibilities. Evolution 55, 1085–1094.
Pearse, D.E., Crandall, K.A., 2004. Beyond FST: analysis of population genetic data for conservation. Conserv. Genet. 5, 585–602.
Pereira, L.H.G., Hanner, R., Foresti, F., Oliveira, C., 2013. Can DNA barcoding accurately discriminate megadiverse Neotropical freshwater fish fauna. BMC Genet. 14, 20.
Pons, J., Barraclough, T.G., Gomez-Zurita, J., Cardoso, A., Duran, D.P., Hazell, S., Kamoun, S., Sumlin, W.D., Vogler, A.P., 2006. Sequence-based species delimitation for the DNA taxonomy of undescribed insects. Syst. Biol. 55, 595–606.
Pouyaud, L., Sudarto, Paradis, E., 2009. The phylogenetic structure of habitat shift and morphological convergence in Asian Clarias (Teleostei, Siluriformes: Clariidae). J. Zool. Syst. Evol. Res. 47, 344–356.
Puillandre, N., Lambert, A., Brouillet, S., Achaz, G., 2012. ABGD, automatic barcode gap discovery for primary species delimitation. Mol. Ecol. 21, 1864–1877.
R_Core_Team, 2014. R: a Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna.
Ratnasingham, S., Hebert, P.D.N., 2007. Bold: the barcode of life data system. Mol. Ecol. notes 7, 355–364. http://www.barcodinglife.org.
Ratnasingham, S., Hebert, P.D.N., 2013. A DNA-based registry for all animal species: the barcode index number (BIN) system. PLoS One 8, e66213.
Read, C.I., Bellwood, D.R., Herwerden van, L., 2006. Ancient origins of Indo-Pacific coral reef fish biodiversity: a case study of the Leopard wrasses (Labridae: Macropharyngodon). Mol. Phylogenet. Evol. 38, 808–819.
Ricklefs, R.E., Schluter, D., 1993. Species Diversity in Ecological Communities: Historical and Geographical Perspectives. University of Chicago Press, Chicago.
Riedel, A., Sagata, K., Suhardjono, Y.R., Tänzler, R., Balke, M., 2013. Integrative taxonomy on the fast track - towards more sustainability in biodiversity research. Front. Zool. 10, 15.
Rogers, A.R., Harpending, H., 1992. Population growth makes waves in the distribution of pairwise genetic differences. Mol. Biol. Evol. 9, 552–569.
Rousseuw, P.J., 1987. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. J. Comput. Appl. Math. 20, 53–65.
Ryder, O.A., 1986. Species conservation and systematics: the dilemma of subspecies. Trends Ecol. Evol. 1, 9–10.

Schloss, P.D., Handelsman, J., 2005. Introducing DOTUR, a computer program for defining operational taxonomic units and estimating species richness. Appl. Environ. Microbiol. 71, 1501–1506.
Selkoe, K.A., Toonen, R.J., 2006. Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. Ecol. Lett. 9, 615–629.
Slatkin, M., Hudson, R.R., 1991. Pairwise comparisons of mitochondrial DNA sequences in stable and exponantially growing populations. genetics 129, 555–562.
Smith, A.M., Fisher, B.L., Hebert, P.D.N., 2005. DNA barcoding for effective biodiversity assessment of a hyperdiverse arthropod group: the ants of Madagascar. Phil. Trans. R. Soc. Serie B 360, 1825–1834.
Smith, M.A., Wood, D.M., Janzen, D.H., Hallwachs, W., Hebert, P.D.N., 2007. DNA barcodes affirm that 16 species of apparently generalist tropical parasitoid flies (Diptera, Tachinidae) are not all generalists. Proc. Natl. Acad. Sci. U. S. A. 104, 4967–4972.
Sork, V.L., Waits, L., 2010. Contributions of landscape genetics–approaches, insights, and future potential. Mol. Ecol. 19, 3489–3495.
Steinke, D., Zemlak, T.S., Hebert, P.D.N., 2009. Barcoding Nemo: DNA-based identifications for the ornamental fish trade. Plos One 4, e6300.
Strimmer, K., Pybus, O.G., 2001. Exploring the demographic history of DNA sequences using the generalised skyline plot. Mol. Biol. Evol. 18, 2298–2305.
Tajima, F., 1983. Evolutionary relationships of DNA sequences in finite populations. genetics 105, 437–460.
Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. Molecular evolutionary genetics analysis version 6.0. Mol. Biol. Evol. 30, 2725–2729.
Tänzler, R., Sagata, K., Surbakti, S., Balke, M., Riedel, A., 2012. DNA barcoding for community ecology - how to tackle a hyperdiverse, mostly undescribed melanesian fauna. PLoS One 7, e28832.
Urban, M.C., LEibold, M.A., Amarasekare, P., De Meester, L., Gomulkiewicz, R., Hochberg, M.E., Klausmeier, C.A., Loeuille, N., De Mazancourt, C., Norberg, J., Pantel, J.H., Strauss, S.Y., Vellend, M., Wade, M.J., 2008. The evolutionary ecology of metacommunities. Trends Ecol. Evol. 23, 311–317.
Vellend, M., 2005. Species diversity and genetic diversity: parallel processes and correlated patterns. Am. Nat. 166, 199–215.
Vellend, M., 2010. Conceptual synthesis in community ecology. Q. Rev. Biol. 85, 183–206.
Vellend, M., Geber, M.A., 2005. Connections between species diversity and genetic diversity. Ecol. Lett. 8, 767–781.
Vernooy, R., Haribabu, E., Muller, M.R., Vogel, J.H., Hebert, P.D.N., Schindel, D.E., Shimura, J., Singer, G.A.C., 2010. Barcoding life to conserve biological diversity: beyond the taxonomic imperative. PLoS Biol. 8, e1000417.
Vogler, A.P., DeSalle, R., 1994. Diagnosing units of conservation management. Conserv. Biol. 6, 170–178.
Wakeley, J., 2001. The coalescent in an island model of population subdivision with variation among demes. Theor. Popul. Biol. 59, 133–144.
Wakeley, J., 2003. Polymorphism and divergence for island-model species. genetics 163, 411–420.
Waples, R.S., 1991. Pacific salmon, Oncorhynchus spp. & the definition of 'species' under the endagered species act. Mar. Fish. Rev. 53, 11–22.
Waples, R.S., 1995. Evolutionary significant units and the conservation of biological diversity under the endangered species act. In: Nielsen, J.L., Power, G.A. (Eds.), Evolution and the Aquatic Ecosystem: Defining Unique Units in Population Conservation. American Fisheries Society, Bethesda, Maryland, pp. 8–27.
Ward, R.D., Hanner, R.H., Hebert, P.D.N., 2009. The campaign to DNA barcode all fishes, FISH-BOL. J. Fish Biol. 74, 329–356.
Warren, B.H., Simberloff, D., Ricklefs, R.E., Aguilée, R., Condamine, F.L., Gravel, D., Morlon, H., Mouquet, N., Rosindell, J., Casquet, J., Conti, E., Cornuault, J., Fernandez-Palacios, J.M., Hengl, T., Norder, S.J., Rijsdijk, F., Sanmartin, I., Strasberg, D., Triantis, K.A., Valente, L.M., Whittaker, R.J., Gillespie, R., Emerson, B.C., Thébaud, C., 2015. Islands as model systems in ecology and evolution: prospects fifty years after MacArthur-Wilson. Ecol. Lett. 18 (2), 200–217.
Weir, J.T., Schluter, D., 2007. The latitudinal gradient in recent speciation and extinction rates of birds and mammals. Science 315, 1574–1576.
Whiteley, A.R., Fitzpatrick, S.W., Funk, W.C., Tallmon, D.A., 2014. Genetic rescue to the rescue. Trends Ecol. Evol. 30, 42–49.
Wiens, J.J., Donoghue, M.J., 2004. Historical biogeography, ecology and species richness. Trends Ecol. Evol. 19, 639–644.
Woodruff, D.S., 2010. Biogeography and conservation in Southeast asia: how 2.7 million years of repeated environmental fluctuations affect today's patterns and the future of the remaining refugium-phase biodiversity. Biodivers. Conserv. 19, 919–941.