






Origin of endemic species in a moderately isolated ancient lake: The case of a snakehead in Inle Lake, Myanmar

Yusuke Fuke¹  | Prachya Musikasinthorn² | Yuichi Kano^{3,4}  | Ryoichi Tabata^{1,5}  |
Shoko Matsui^{1,6}  | Sein Tun^{7,8} | L. K. C. Yun^{7,9} | Bunthang Touch¹⁰ |
Phanara Thach¹⁰ | Katsutoshi Watanabe¹ 

¹Laboratory of Animal Ecology, Graduate School of Science, Kyoto University, Kyoto, Japan

²Department of Fishery Biology, Faculty of Fisheries, Kasetsart University, Bangkok, Thailand

³Institute of Decision Science for Sustainable Society, Kyushu University, Fukuoka, Japan

⁴Kyushu Open University, Fukuoka, Japan

⁵Lake Biwa Museum, Shiga, Kusatsu, Japan

⁶Osaka Museum of Natural History, Osaka, Japan

⁷Inlay Lake Wildlife Sanctuary, Nature and Wildlife Conservation Division, Forest Department, Ministry of Natural Resources and Environmental Conservation, the Republic of the Union of Myanmar, Shan State, Nyaung Shwe, Myanmar

⁸Natma Taung National Park, Chin State, Kanpalet Township, Myanmar

⁹Hkakaborazi National Park, Kachin State, Putao, Myanmar

¹⁰Inland Fisheries Research and Development Institute of Fisheries Administration, Phnom Penh, Cambodia

Correspondence

Yusuke Fuke and Katsutoshi Watanabe,
Laboratory of Animal Ecology,
Graduate School of Science, Kyoto
University, Kitashirakawa-oiwakecho,
Sakyo, Kyoto 606-8502, Japan.
Email: yfa73986@gmail.com and
watanak@terra.zool.kyoto-u.ac.jp

Funding information

Japan Society for the Promotion of
Science; Kurita Water and Environment
Foundation; Sumitomo Foundation

Abstract

Inle Lake is an ancient lake in Myanmar, which is an important area with unique and diverse fauna. Its ichthyofauna is believed to have formed non-radiatively, but the historical processes are poorly understood. To elucidate the mechanisms that shape species diversity in this moderately isolated biogeographical ‘island’, this study focused on a typical endemic fish of Inle Lake, *Channa harcourtbutleri* (Channidae, Anabantiformes), with its widely distributed sister species, *C. limbata*, and estimated the historical distribution and diversification processes of the endemic fish based on genome-wide polymorphism (MIG-seq) and mitochondrial DNA data. *Channa harcourtbutleri* contained two genetically and morphologically distinct groups inhabiting Inle Lake and the surrounding rivers respectively. These two groups were genetically the closest to each other; however, the riverine group showed some similarity to the closely related species, *C. limbata* from Southeast Asia. The mtDNA haplotypes of the endemic species were not monophyletic; most of the riverine group had haplotypes identical or close to those of *C. limbata* from the upper Irrawaddy and Salween rivers. The time tree suggested that *C. harcourtbutleri* diverged from *C. limbata* in the early Pleistocene and then experienced secondary contact with *C. limbata* in the late

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2023 The Authors. *Zoologica Scripta* published by John Wiley & Sons Ltd on behalf of Royal Swedish Academy of Sciences.

Pleistocene. Genetic and morphological differentiation within *C. harcourtbutleri* suggests that local adaptation to different environments has played an important role for the coexistence of its two forms with some reproductive isolation. Further, the results highlight the importance of multiple colonization and allopatric speciation in shaping biodiversity in the long-term, moderately isolated environments.

KEYWORDS

biogeography, *Channa*, freshwater fish, population structure, secondary contact

1 | INTRODUCTION

Understanding how local faunas form and transition is one of the major topics in biogeography. In the study of community formation, island-like habitats with distinct boundaries, such as islands, lakes and mountain tops, have long been a focus of attention (Brooks, 1950; Losos & Ricklefs, 2009; MacArthur & Wilson, 1963, 1967; Rahbek et al., 2019). Island biogeographical theory has attempted to explain species composition in an 'island' using the dynamics of colonization from the species pool 'continent' and speciation and extinction within the island (Whittaker et al., 2008, 2017). When the distance between an island and the continent (source area) increases, the contribution to species composition shifts from colonization to speciation (Emerson & Gillespie, 2008). In strongly isolated oceanic islands and ancient lakes, in situ diversification often accounts for the species diversity in various taxa (Gillespie et al., 2020). However, in more general communities with moderate isolation, the continent (source) and timing of colonization should have a more important role in the community assembly (Hauffe et al., 2020; Mittelbach & Schemske, 2015; Rosindell & Phillimore, 2011). Colonization can occur multiple times and from multiple source regions (Johnson et al., 2019; Kimura et al., 2022; Šlechtová et al., 2021); therefore, detailed phylogeographical studies using molecular markers based on regionally exhaustive sampling are needed for identifying the processes of community formation with colonization (Cox et al., 2020).

Ancient lakes, often defined as those that have existed since 130,000 BP (Hampton et al., 2018), have unique biota because they maintain a relatively closed, stable aquatic environment for long periods. They have attracted the attention of many biologists as 'natural experiments' for understanding the processes and mechanisms that lead to adaptation and speciation (Cristescu et al., 2010). Adaptive radiation is the prominent process underlying the biodiversity in ancient lakes; it is a factor for the explosive in situ speciation (Henning & Meyer, 2014; Schluter, 2000).

Non-radiative processes, such as colonization and long-term isolation of multiple ancestors, also greatly contribute to the formation of unique fauna (Hauffe et al., 2020; Tabata et al., 2016). However, this has received little attention compared to radiation (Seehausen, 2006). Ancient lakes where explosive radiation did not occur are good systems to examine the combined effects of isolation and colonization from surrounding water systems on the formation of the lake fauna.

Inle Lake is the only ancient lake in the continental part of Southeast Asia; it is located in Shan Plateau of Myanmar. The age of Inle Lake is estimated at about 1.5 million years (Hampton et al., 2018), although the geological evidence remains unclear. Inle Lake is a small lake with an area of 116 km² and a depth of 2 m (Toke et al., 2013); however, it has rich biodiversity and several endemic species (Abell et al., 2008; Annandale, 1918; Fuke et al., 2021). Fifteen endemic fish species have been reported from the lake, and there might still be several undescribed species (Kano et al., 2016, 2022). The native fish fauna of Inle Lake is composed of a wide range of taxa (28 genera from 13 families; Kano et al., 2016, 2022), which are believed to be formed through non-radiative processes (Fuke et al., 2022). Annandale (1918) suggested that the unique fish fauna of Inle Lake includes widely distributed species that colonized the lake from surrounding areas and some endemic species derived through geographical isolation. In addition, several species treated as endemic species of Inle Lake also occur in the surrounding areas, and some of them are known to show genetic differentiation among regions (Fuke et al., 2022; Kano et al., 2022). The fishes of the Inle region, defined here as the area including Inle Lake and surrounding regions in Shan Plateau, will provide good materials for exploring the processes by which the interactions of in situ diversification and colonization shape local biodiversity. However, detailed historical biogeographical studies are lacking in this region.

In this study, we aimed (1) to estimate the historical patterns of colonization from the surrounding water systems to Inle Lake and (2) to characterize the in situ diversification of endemic fish species within the region. We

focused on *Channa harcourtbutleri* (Annandale, 1918) (Channidae, Anabantiformes), an endemic snakehead species occurring in and around Inle Lake (Rüber et al., 2020). *Channa harcourtbutleri* is one of the local populations derived from *Channa limbata* (Cuvier in Cuvier & Valenciennes, 1831), which is distributed widely from Southeast Asia to southern China except the Inle region (Courtenay & Williams, 2004; Rüber et al., 2020). *Channa limbata* is genetically differentiated by region (Conte-Grand et al., 2017; Rüber et al., 2020); therefore, it would be possible to estimate the origin of *C. harcourtbutleri*, to understand from which region and when the ancestral population of this endemic species colonized the Inle region. We first estimated the phylogeographical patterns of *C. harcourtbutleri* and *C. limbata* based on genome-wide SNPs and mitochondrial DNA (mtDNA) to infer the origin of the populations in the Inle region. Second, we examined the genetic population structure and morphological characters of *C. harcourtbutleri* to determine whether population diversification occurred within the region. These results assess the formation history of the non-radiative biodiversity in this region and provide insights into the dynamic relationships between regional biodiversity and connectivity with and isolation from the surrounding areas.

2 | MATERIALS AND METHODS

2.1 | Taxonomic background and identification

In this study, we focused on two species of snakeheads, *Channa harcourtbutleri*, an endemic species in Inle Lake and its surroundings, and *C. limbata*, widely distributed in Southeast Asia. These species belong to Channidae, which includes two genera with 56 species distributed in Africa and South to East Asia (Adamson et al., 2010; Fricke et al., 2022). The species of this family have a suprabranchial organ for air-breathing and are adapted to a variety of environments, including streams, lakes and wetlands (Musikasinthorn, 2003). Among them, *C. harcourtbutleri* and *C. limbata* are small snakehead species with a maximum body size of about 20 cm, inhabiting hill streams, lakes and ponds. *Channa limbata* had been regarded as a junior synonym of *Channa gachua* (Hamilton, 1822) (Roberts, 1993); however, considering the distinct phylogenetic divergence, the western populations are referred to as the *C. gachua* group and the eastern populations, as the *C. limbata* group, bound by the Indo-Burman ranges (Conte-Grand et al., 2017). *Channa harcourtbutleri* is a regional clade nested within the *C. limbata* group (Rüber et al., 2020).

Ng et al. (1999) documented that *C. harcourtbutleri* can be morphologically distinguished from *C. gachua* (including *C. limbata*) by body colour and head shape. However, they did not provide detailed quantitative comparisons that account for allometry and geographical variation. *Channa harcourtbutleri* occurs not only in Inle Lake but also in stream habitats surrounding Inle Lake (Kano et al., 2016, 2022). Although intraspecific differentiation in morphological and genetic features is often observed among freshwater fishes inhabiting different environments (e.g. lakes and streams; Ravinet et al., 2013; Theis et al., 2014), no detailed comparisons have been made for the *Channa* species. Therefore, we tentatively treated the specimens of the *C. limbata* group based on their locality as follows: specimens from the Inle region as *C. harcourtbutleri* following Annandale (1918) and those from Indochina, China and Myanmar except the Inle region as *C. limbata* following Conte-Grand et al. (2017). Recently, Laskar et al. (2023) treated specimens from the Andaman Islands and mainland India as *C. harcourtbutleri* based on mtDNA barcoding only. However, this cannot be validated considering the mtDNA diversity of the *C. limbata* group (Conte-Grand et al., 2017; Rüber et al., 2020; see also the results of this study), as well as the nomenclatural priority.

2.2 | Fish sampling

We used a total of 152 specimens of *C. harcourtbutleri* collected from various habitats or purchased from local markets in and around Inle Lake by Kano et al. (2016) and Kano et al. (2022) from 2014 to 2016, and 2020 (Tables S1 and S2). Among these, 137 individuals were used for genetic analysis, of which 53 were used for morphological analysis. Specimens obtained from local markets could potentially have been collected far from the market areas, but based on local interviews and the style of selling, we judged that they were not transported long distances and used these samples (Table S1). We grouped the sampling sites into the following 10 areas (Figure 1): Inle Lake, collected at four sites on the lake and purchased at two local markets; streams north of the lake (North, seven sites); streams west of the lake (West, two sites); Inn Dein Stream (Inndein, three sites and one local market), a stream flowing into the lake from the southwest; Min Ywar Village (South, one site), about 15 km downstream from the lake; Moe Byel Dam Lake (Dam, one local market), 60 km downstream from Inle Lake; Loikaw (two sites and one local market), 20 km downstream from the dam; Heho (three sites), about 20 km northwest of Inle Lake separated by highlands; Aungpan (one site), about 10 km further west of Heho, separated from Inle Lake by highlands; and Hopong (two sites), about 30 km northeast

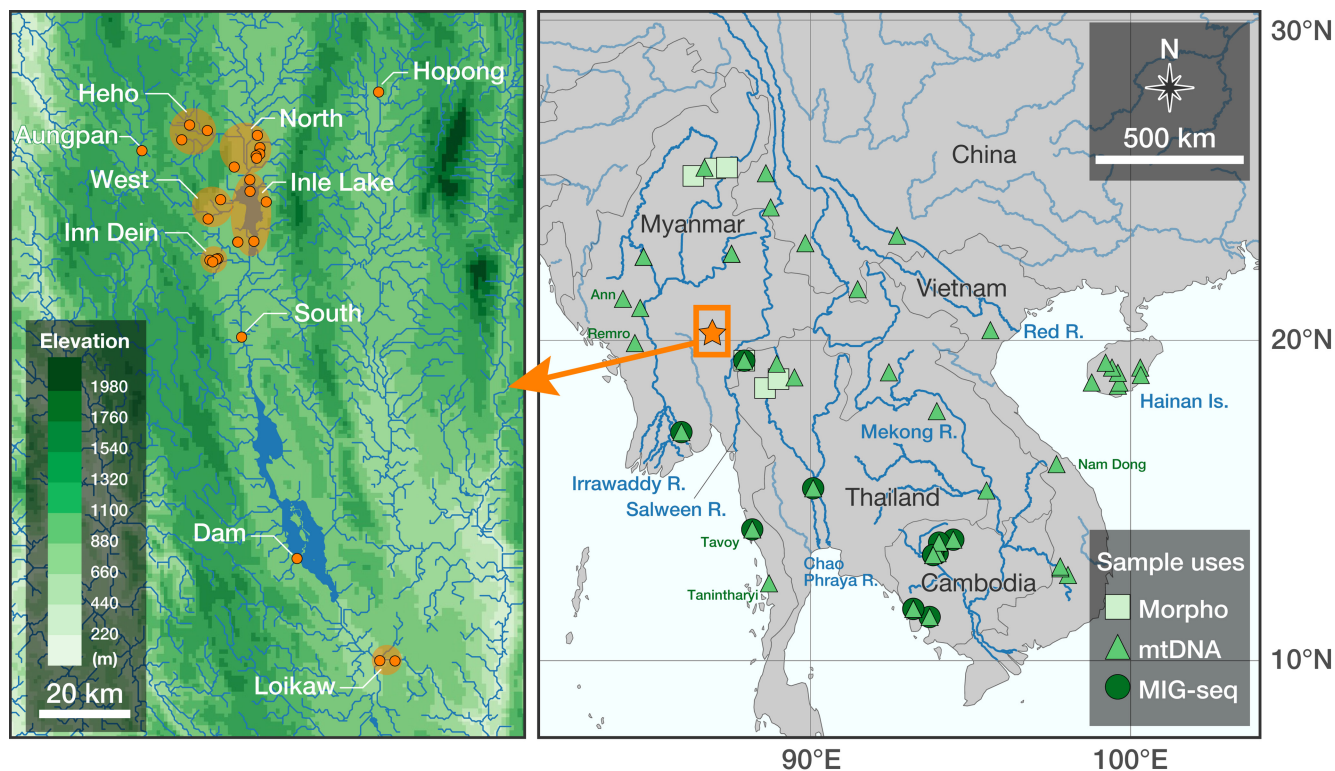


FIGURE 1 Collection sites for *Channa harcourtbutleri* (left) and *C. limbata* (right) used in this study. In the left map, the orange circle plots represent sampling sites of *C. harcourtbutleri* and the translucent circles represent the range of pooled sampling sites. See [Table S2](#) for details of sampling sites. In the right map, the plots represent the locality of *C. limbata* samples, and their shapes represent the uses of the samples. Data for samples represented only by triangles (mtDNA) were obtained from the International Nucleotide Sequence Database. See [Table S1](#) for specimen details.

of Inle Lake, separated from the lake through the river by about 400 km (Kano et al., 2022).

A total 33 specimens of *C. limbata* were collected from Myanmar, Thailand and Cambodia for genetic analysis. In addition, we used 22 *C. limbata* specimens from Myanmar and Thailand for morphological analysis, which were not used for genetic analysis (Figure 1).

We used the DNA samples shared with Kano et al. (2016, 2022). DNA from some samples was newly extracted using the Wizard Genomic DNA Purification Kit (Promega).

The mtDNA cytochrome *b* (*cytb*) sequences of 67 individuals of *C. limbata* for which locality information was available were obtained from the International Nucleotide Sequence Database (Figure 1; [Table S1](#)). We additionally included eight *Channa* species and 30 other anabantiform species as outgroups from the database (see [Table S1](#)).

2.3 | DNA sequencing

To assess the genome-wide polymorphism, we used the multiplexed ISSR genotyping by sequencing method (MIG-seq; Suyama & Matsuki, 2015; Suyama et al., 2022).

The MIG-seq library was prepared according to Watanabe et al. (2020) and Onuki and Fuke (2022) for 136 individuals of *C. harcourtbutleri* and 31 of *C. limbata*. The library was sequenced on MiSeq (Illumina) with MiSeq Reagent Kit v3 (150 cycle) or outsourced to Novogene and sequenced on NovaSeq 6000 (Illumina).

The mtDNA *cytb* region was amplified using PCR with the following primers: L-Glu (5'-CTA ACC AGG ACT AAT GGC TTG AA-3') and H-Thr (5'-CGG CTT ACA AGA CCG GCG CTC TGA-3') (Takahashi et al., 2016). The PCR conditions were as follows: initial denaturation at 94°C for 120s; 30 cycles of denaturation at 94°C for 15s, annealing at 52°C for 15s, extension at 72°C for 30s and final extension at 72°C for 420s. Sequencing was performed on an ABI 3130xl Genetic Analyzer (Applied Biosystems) following the method of Fuke et al. (2022) or outsourced to Macrogen Japan and sequenced on ABI 3730xl DNA analyser. The sequences were checked for quality and edited using the mapping function of Unipro UGENE 41.0 (Okonechnikov et al., 2012); multiple alignments were performed using MAFFT (Kato & Standley, 2013). Finally, the *cytb* sequences were obtained for a total of 137 and 98 specimens for *C. harcourtbutleri* and *C. limbata* respectively.

2.4 | MIG-seq data treatment and SNP detection

The raw data sequenced by NovaSeq 6000 were demultiplexed using the `process_shortread` program in Stacks 2.49 (Rochette et al., 2019). Quality control and trimming of the obtained raw reads were performed using `fastp` program (Chen et al., 2018). All reads were trimmed to 80 bases: the first 14 bases of Read 2 were trimmed for Miseq data; the first 35 and last 29 bases of Read 1 and 17 and 52 bases of Read 2 were trimmed for NovaSeq data. In addition, the row quality bases (<Q30) and primer sequences were removed.

Detection of single nucleotide polymorphisms (SNPs) was performed using the `denovo_map.pl` pipeline of Stacks with paired-end mode with the setting: maximum distance allowed between stacks (M)=5; a minimum depth of coverage required to create a stack (m)=5. Outfiles for the following analyses were generated using `populations` program in Stacks with the following setting: minimum percentage of variant sites shared by all individuals (R)=0.8; only the first SNP per locus (`--write-single-snp`). We excluded sites with heterozygosity greater than 75% (`--max-obs-het=0.75`) and less than two minor alleles (`--min-mac=2`). All other parameters were set as the default setting.

2.5 | Population structure and phylogenetic analyses based on SNPs

To estimate the population structure of *C. harcourtbutleri* and *C. limbata*, we conducted principal component analysis (PCA) using PLINK (Purcell et al., 2007). To estimate individual admixture proportions, we performed unsupervised clustering using the likelihood model-based program ADMIXTURE 1.3.0 (Alexander et al., 2009) at each number of genetic populations (K) from 1 to 10 with a convergence criterion set to 10^{-4} ($C=0.0001$). These analyses were repeated 100 times with different seed values, and the optimal K value was estimated based on the lowest mean cross-validation error (CV-error) value. The outputs of the runs with the lowest CV-error for each K -value were plotted. In the following analysis, individuals with a major ancestry proportion below 90% in the best K were treated as putative hybrids. Summarization of CV-error values and visualization of the output data were performed in R 4.1.1 (R Core Team, 2021).

To estimate the relationship among clades, the network tree inference was performed using SplitsTree 4.14.7 (Huson & Bryant, 2006) based on the Neighbour-net method with the uncorrected p-distance.

2.6 | Haplotype network and phylogenetic analyses based on mtDNA

To estimate the population structure of *C. harcourtbutleri* and *C. limbata* based on mtDNA *cytb* sequences, we detected the haplotypes and conducted a haplotype network analysis using POPART 1.7 (Leigh & Bryant, 2015) based on the TCS algorithm (Clement et al., 2000). The networks were generated based on two datasets, one with *C. harcourtbutleri* only and the other with both species. *Channa limbata* used in the network analysis were limited to samples from the Salween and Irrawaddy River systems, which formed a monophyletic group including *C. harcourtbutleri*.

To estimate phylogenetic relationships among mtDNA *cytb* haplotypes detected in *C. harcourtbutleri* and *C. limbata*, a maximum likelihood (ML) tree inference was performed using IQ-TREE 2.1.3 (Minh et al., 2020) with 38 outgroup sequences. The *cytb* sequences were partitioned by codon position, and a substitution model was selected for each position using ModelFinder (Kalyaanamoorthy et al., 2017). The estimated trees were visualized using FigTree 1.4.4 (<http://tree.bio.ed.ac.uk/software/figtree/>).

2.7 | Divergence time estimation

To estimate the divergence time of *C. harcourtbutleri* from *C. limbata*, a time-calibrated tree was inferred based on mtDNA using BEAST 2.6.7 (Bouckaert et al., 2019). All haplotypes of *C. harcourtbutleri* and *C. limbata* were used in the analysis. The *cytb* sequences were partitioned by codon position, and the following setting was applied for all codon positions; the substitution and site models were set to the same as in the IQ-TREE2 analysis, the clock model to the Relaxed Clock Log Normal model, and the tree prior to the Birth-death model. The analysis was carried out with three calibration points according to Rüber et al. (2020) and Wu et al. (2019): concestor of *Parachanna* stem (log-normal, offset = 33.0, log mean = 0.1, log stdev = 0.8), concestor of *Osphronemus* stem (log-normal, offset = 28.5, log mean = 0.23, log stdev = 1.0), concestor of Anabantidae stem (uniform, lower = 24.3 Ma, upper = 163.5 Ma). All other parameters were set at the default setting. MCMC was performed for 300,000,000 generations with sampling every 10,000 generations. We confirmed that all parameters had ESS values above 200 using Tracer 1.7.1 (Rambaut et al., 2018). TreeAnnotator excluded the first 10% of trees as burn-in and combined all trees into a single tree.

2.8 | Morphological analyses

Three phenotypic datasets were obtained to identify morphological characteristics of *C. harcourtbutleri*: body shape based on geometric morphometrics, and six counting and 11 metric traits. Specimens were fixed in 10% formalin and then placed in 70% ethanol. For *C. harcourtbutleri*, only specimens grouped by genetic analysis were used for morphological analysis. For *C. limbata*, all specimens were pooled and treated as one group, regardless of whether they were genetically analysed or not. Specimens smaller than 100 mm in total length were not included in the morphological analysis, with referring to the mature body size of *C. gachua*, the closely related species to the two species (Silva, 1991).

The number of fin-rays and lateral line scales (LLS) were counted, as counting traits. The length of dorsal fin base (LDFB), anal fin base (LAFB), pectoral fin (PFL), ventral fin (VFL), caudal fin (CFL), 1st ray of dorsal fin (DFL1), and nasal tube (NTL), mouth width (MW), head width (HW), body width (BW) and inter orbital width (IOW) were measured as metric traits and calculated as a percentage of standard length (SL) (Hubbs & Lagler, 1958; Musikasinthorn, 1998). The final dataset included 53 *C. harcourtbutleri* and 22 *C. limbata* specimens.

Sample images for geometric morphological analysis were taken using a digital camera under the same condition. These images are deposited in ffish.asia (<https://ffish.asia>), a biodiversity database of freshwater fish and freshwater organisms in Asia (Kano et al., 2013; Watanabe et al., 2010). Landmarks were plotted on tpsDIG2 2.3 (<https://www.sbmorphometrics.org/soft-dataacq.html>) for the ventral and left images of the specimens. The ventral side and left side of the body were plotted with eight and nine landmarks respectively. In order to alleviate the potential effects of body axis curvature, the position of the landmark of the ventral side was limited to the front half of the body, which was hardly affected by the curvature. For the lateral side analysis, specimens with obviously distorted body axes were not included. The final datasets consisted of $n=72$ (51 *C. harcourtbutleri* and 21 *C. limbata*) for the ventral side and $n=51$ (37 *C. harcourtbutleri* and 14 *C. limbata*) for the lateral side. The differences in position, direction and size of the samples were corrected using Procrustes analysis; the principal component analysis (PCA) was carried out using MorphoJ 1.06 d (Klingenberg, 2011).

The results were compared among the genetic groups based on the SNP analysis. The Kruskal–Wallis test was performed for each PC score, counting and metric trait to identify the differences among groups. We performed multiple comparisons using the Steel–Dwass test to examine which groups differed. The significance level was set at

1% in all statistical tests. These calculations and graphical visualization were performed using R 4.1.1.

3 | RESULTS

3.1 | Population structure and phylogenetic relationships based on SNPs data

PCA results based on 231 SNPs for 167 individuals of *Channa harcourtbutleri* ($n=136$) and *C. limbata* ($n=31$) showed that *C. harcourtbutleri* formed a single cluster, while *C. limbata* clustered by location, with the lower Irrawaddy River, middle Salween River and Indochina clusters in PC1 in this order (Figure 2a). For the *C. harcourtbutleri* only data, PCA based on 236 SNPs for 136 individuals showed two dense clusters and some intermediate individuals in PC1. Some of the intermediate individuals from Hopong separated from the others in PC2. One cluster consisted mainly of individuals from Inle Lake-to-Loikaw (outflow river); another consisted of individuals from the surrounding rivers. Hereafter, these two clusters of the endemic species, which were also supported by ADMIXTURE analysis (see below), are referred to as SNP Cluster 1 (Inle Lake and its outflow) and 2 (surrounding rivers) respectively. Some sites, such as North, Innsein and Hopong, had individuals from both clusters. The intermediate individuals in PC1 consisted of some specimens from North and Hopong.

ADMIXTURE analysis, including the two species, showed the optimal number of clusters was $K=4$ or 5 where the mean or minimum CV-error value was minimized respectively (Figure 2b). The two clusters of *C. harcourtbutleri* shown in $K=3-5$ corresponded to the two clusters in the PCA results (i.e. SNP Clusters 1 and 2); individuals consisting of both elements corresponded to the intermediate individuals of the two clusters in the PCA results. Individuals from the two clusters were observed sympatrically in Inle Lake, Innsein, North and Hopong. Among them, individuals having both cluster elements, as putative hybrids between the two clusters, were found in Inle Lake, North, Heho and Hopong. North and Hopong were particularly notable for the sympatry and admixture of the two clusters. These results were also supported in the dataset for endemic species only (Figure S3). In the case of $K=5$, *C. limbata* was composed of three clusters that were divided by region: Myanmar, Thailand and Cambodia. *Channa limbata* from the Irrawaddy and Salween rivers partially contained a genetic component of SNP Cluster 2 of *C. harcourtbutleri* in $K=2-4$.

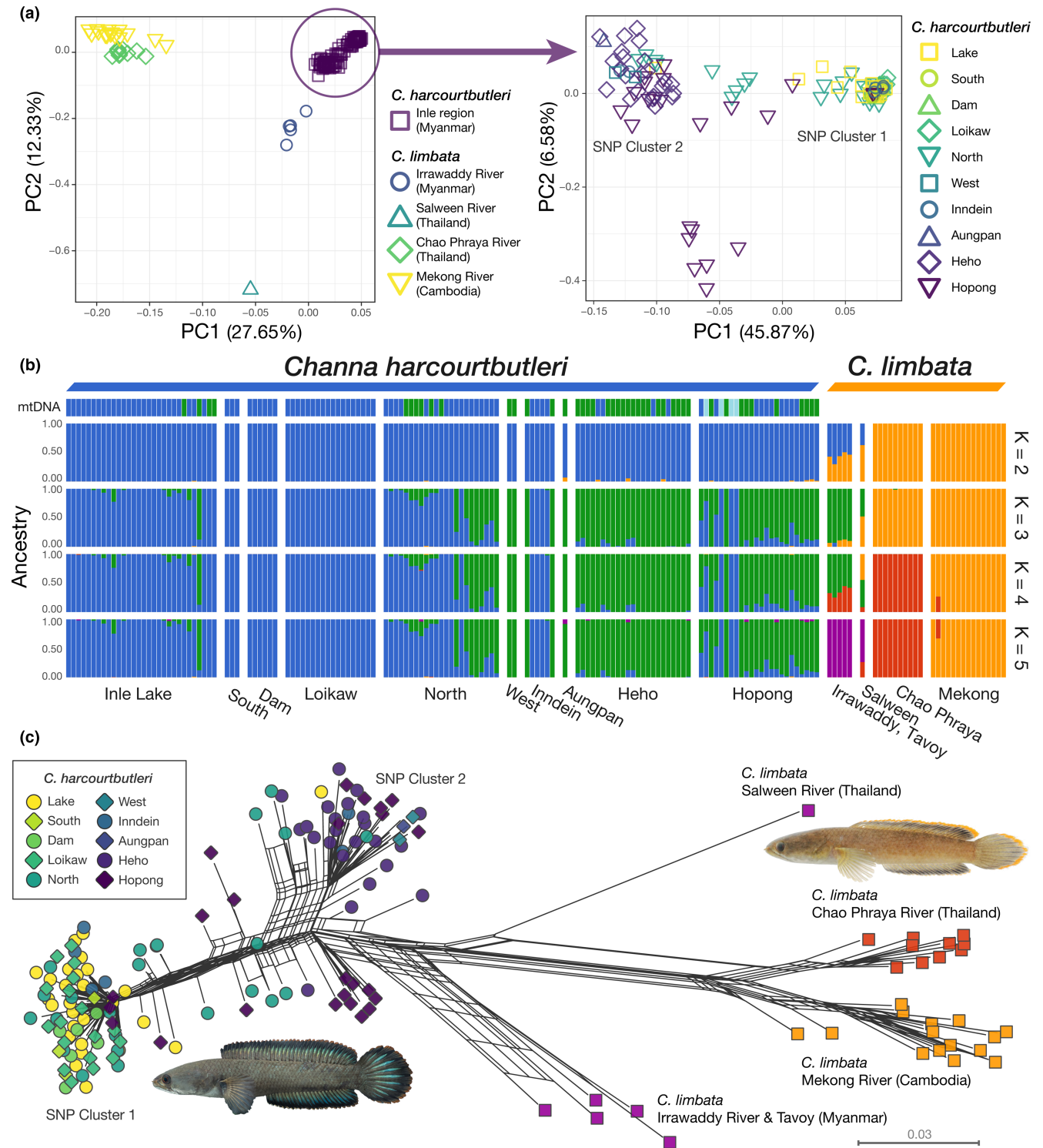


FIGURE 2 Results of single nucleotide polymorphism (SNP)-based analyses. (a) Plots of principal component analyses of *Channa harcourtbutleri* and *C. limbata* based on 231 SNPs (left) and *C. harcourtbutleri* only based on 236 SNPs (right). (b) Ancestry barplots of ADMIXTURE analysis for *C. harcourtbutleri* and *C. limbata* based on 231 SNPs showing $K=2-5$. The optimal number of clusters estimated via cross-validation error comparison was four or five. The colour of the plot above the bar plot for *C. harcourtbutleri* represents the mtDNA clades; blue: Clade 1, sky blue: Clade 2, green: Clade 3. (c) Neighbour-net network for *C. harcourtbutleri* and *C. limbata* based on 231 SNPs.

The Neighbour-net network showed two genetically differentiated groups in *C. harcourtbutleri* corresponding to SNP Clusters 1 and 2 with some putative hybrids

between them (Figure 2c). SNP Cluster 2 and *C. limbata* were closer in the network than SNP Cluster 1 and *C. limbata*. Four populations of *C. limbata* were detected,

in which the population most closely related to *C. harcourtbutleri* was from the Irrawaddy River, followed by that from the Salween River.

3.2 | Phylogenetic relationships based on mtDNA

Haplotype analysis based on the *cytb* sequences (1134 bp) for 137 individuals of *C. harcourtbutleri* and 98 of *C. limbata* detected a total of 27 and 84 haplotypes respectively. Haplotypes of *C. harcourtbutleri* were included in three clades (Clades 1, 2 and 3); 21 haplotypes in Clade 1, 3 in Clade 2 and 3 in Clade 3 (Figure S2b,c). All haplotypes in Clade 2 were found in Hopong specimens. Haplotypes in Clades 1 and 3 were found at several sites, the former mainly from Inle Lake-to-Loikaw and the latter mainly from the surrounding rivers, such as Heho and Hopong. In Clade 3, the most major haplotype ch3-1 was identical to the haplotype widely found in *C. limbata* in the upper Salween, upper Mekong and upper Red rivers (cl16; Hap 16 in Wang et al., 2021). The haplotype ch3-2 had one substitution with the haplotype of *C. limbata* from the upper Irrawaddy River (cl70). The uncorrected *p*-distances between Clade 1 and 2 haplotypes were 1.9%–2.5%, and those between Clades 1+2 and 3 were 3.1%–3.6%.

The ML tree based on the 27 mtDNA *cytb* haplotypes of *C. harcourtbutleri* and 84 haplotypes of *C. limbata* revealed polyphyly of mtDNA in *C. harcourtbutleri*. The tree supported that Clades 1 and 2 were monophyletic, and Clade 3 was involved in a clade of *C. limbata* from the upper regions of the Salween, Irrawaddy, Mekong and Red river systems. This mtDNA polyphyly of *C. harcourtbutleri* contrasted with the results of the SNP-based analysis, which showed a clear segregation of *C. harcourtbutleri* and *C. limbata*. Specimens having Clades 1 or 2 haplotypes were generally included in SNP Cluster 1 (62/68, 3/3, respectively, calculated excluding putative hybrids), and those having Clade 3 mtDNA were included in SNP Cluster 2 (30/34). The correspondence between mtDNA clades and SNP clusters was unclear in the North and Hopong populations (Figure 2b).

Channa limbata included several deeply divergent clades that showed geographical cohesion. The basal clades were found from the middle Salween River and the western coast of Myanmar (i.e. cl59, 77, 78; Figure S2a). There were several haplotypes (e.g. cl14, 16 and 17) found from different water systems; they were all distributed in the upper reaches of these rivers (Figure 3; Figure S2a).

3.3 | Divergence time estimation

Using all 27 haplotypes of *C. harcourtbutleri* and 84 haplotypes of *C. limbata*, the divergence time was estimated based on the partitioned *cytb* region (Figure 3). The sister species to *C. limbata* group that included *C. harcourtbutleri* was estimated as *C. barca*; their divergence time was estimated at 9.69 Ma [95% highest posterior density (HPD), 5.87–12.87]. The most recent common ancestor (tMRCA) of the *C. limbata* group was estimated at 4.47 Ma (95% HPD, 2.84–6.31). The divergence time of Clade 1 and Clade 2 of *C. harcourtbutleri* was estimated at 1.11 Ma (95% HPD, 0.58–1.77), and Clade 1+2 and *C. limbata* including Clade 3 was estimated at 2.01 Ma (95% HPD, 1.22–2.88). The tMRCA of Clade 3 was estimated at 0.13 Ma (95% HPD, 0.03–0.18).

3.4 | Morphological analyses

The ranges of standard length (SL) of each genetically classified group (*C. harcourtbutleri* SNP Clusters 1 and 2, putative hybrid and *C. limbata*) largely overlapped (Figure S4; Table S3) and showed no significant differences except for one group pair (Steel–Dwass test, $p = .007$ for SNP Cluster 2 vs. hybrid, $p > .07$ for other pairs). Among the six counting and 11 metric traits, two counting traits (dorsal fin rays and LLS) were more abundant in the endemic species than in *C. limbata* (Figure 4c; Figure S5; Table S3). SNP Cluster 2 showed intermediate values between SNP Cluster 1 and *C. limbata* in three traits (anal fin rays, LDFB/SL and MW/SL), and more similar values to *C. limbata* than to SNP Cluster 1 in five traits (LAFB/SL, CFL/SL, IOW, HW/SL and BW/SL). Ventral fin rays showed differences between SNP Cluster 2 and *C. limbata* (Steel–Dwass test, $p = .004$); however, the ranges were consistent among all groups (5–6). PFL/SL showed differences among the groups (Kruskal–Wallis rank sum test, $p = .006$) but did not reach significance levels (1%) in any multiple comparison (Steel–Dwass test, the smallest $p = .02$ between SNP Clusters 1 and 2). No significant difference among the groups was found in the other five traits (Kruskal–Wallis rank sum test, $p > .013$).

Morphometric analysis based on the eight landmarks on the ventral side of *Channa* species showed significant differences among the genetically classified groups in PC1 (36.0% contribution) and PC2 (16.2%) (Kruskal–Wallis rank sum test, $p < .001$). Negative correlations were found between body size and PC1 in SNP Cluster 1 and *C. limbata* (Spearman's rank correlation test, $\rho = -.58, -.67, p = .002, .001$ respectively); however, no correlations were detected in other groups. In PC1, SNP Cluster 1 had a longer and narrower head, while SNP

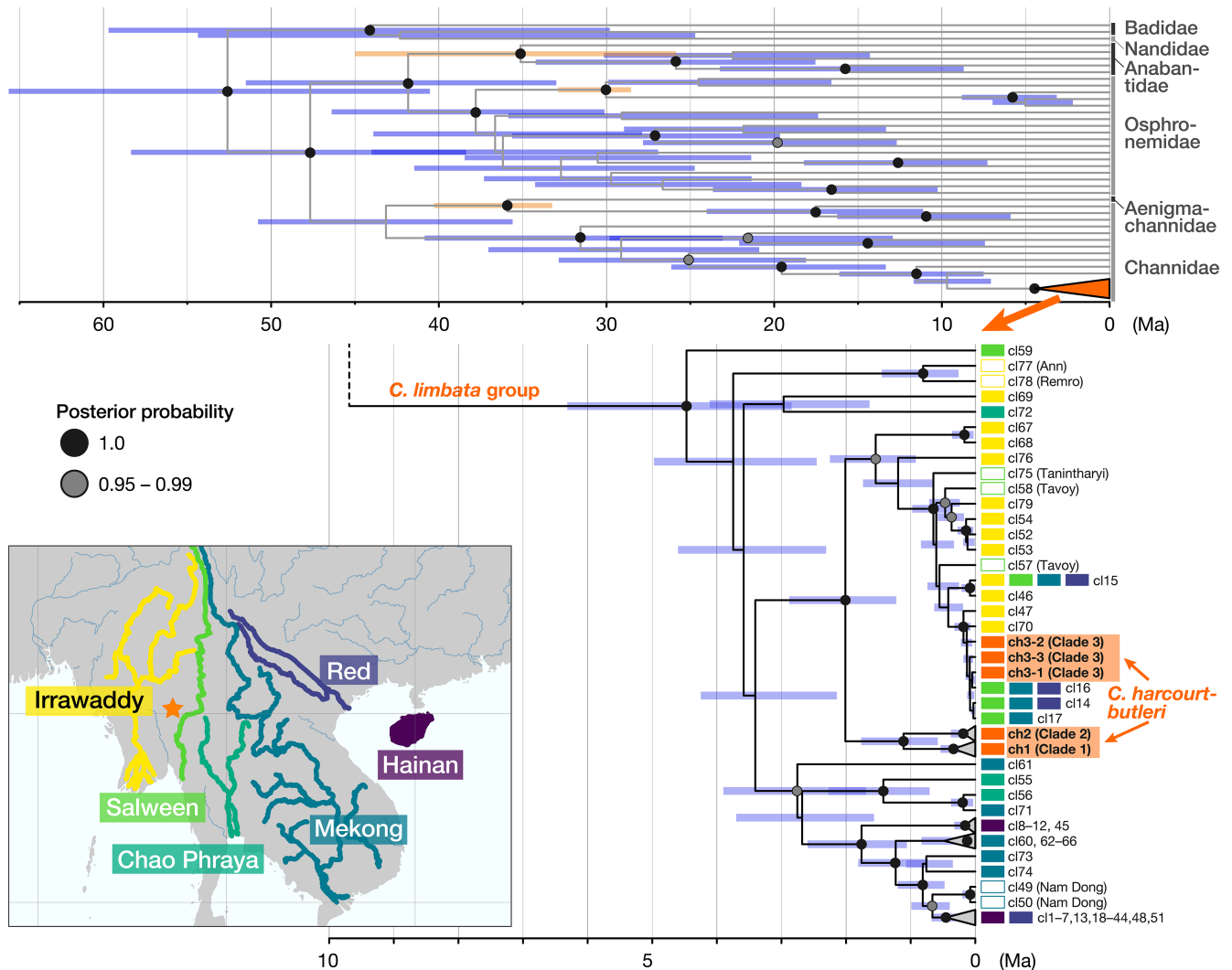


FIGURE 3 A time tree with the divergence time of *Channa limbata* group including *C. harcourtbutleri* based on mtDNA *cytb* (1134 bp). The node circles represent the nodes with ≥ 0.95 posterior probability, and node bars represent 95% highest posterior density of divergence time. Nodes with orange bars represent time calibration points based on fossil records.

Cluster 2, hybrid and *C. limbata* had a shorter and wider head (Steel–Dwass test, $p < .001$; Figure 4). In PC2, SNP Cluster 2 had a narrower mouth width and more anterior isthmus position, whereas *C. limbata* had a wider mouth width and more posterior isthmus position; the SNP Cluster 1 was found to be intermediate between them (Steel–Dwass test, $p < .01$)

The analysis for the nine landmarks on the left lateral side of the *Channa* species showed significant differences among groups in PC2 (11.9% contribution; Kruskal–Wallis rank sum test, $p < .001$). No significant correlations between body size and PC scores were found in any of the four groups (Spearman's rank correlation test, $p > 0.03$). The variation in PC1 (50.9% contribution) probably reflected the curvature of the specimens during fixation (Figure 4a). The PC2 results indicated that the distance between the snout and the dorsal fin origin was shorter in SNP Cluster 1 and longer in SNP Cluster 2 (Steel–Dwass test, $p = 0.002$).

4 | DISCUSSION

Species diversity on an 'island' is hypothesized to be formed through the dynamics of species colonization through dispersal from 'continent(s)' and extinction and speciation within the island (Heaney, 2000; Whittaker et al., 2008). Especially on moderately isolated 'islands', where immigration could occur multiple times through multiple routes (Hirao et al., 2015; Spironello & Brooks, 2003), detangling the complex patterns of isolation, colonization and in situ speciation over the islands' history is necessary to understand the mechanism underlying the formation of the regional fauna (Cavender-Bares et al., 2009; Ricklefs, 1987). In this study, to understand the mechanisms underlying the formation of the unique fauna in and around Inle Lake, a moderately isolated biogeographical island, we examined the phylogeography, population structure and morphological divergence of the

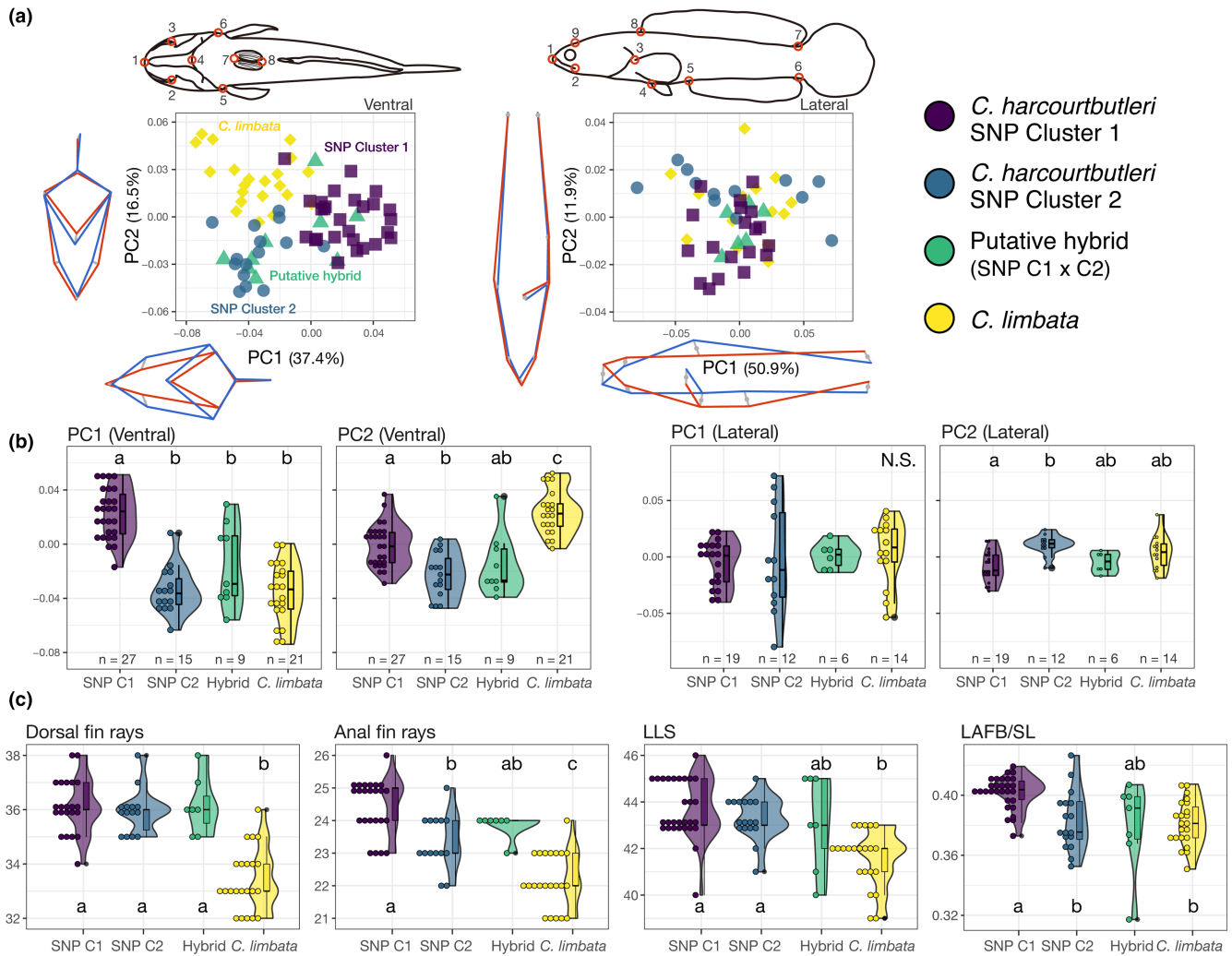


FIGURE 4 Results of morphological analysis of *Channa harcourtbutleri* and *C. limbata*. (a) Results of principal component analysis for PC1 and PC2. The position of the landmarks is shown above the plots. The wireframes represent the maximum (red lines) and minimum (blue lines) values of the calculated principal components of PC1 and PC2 respectively. The colour and shape of the plots represent the SNP clusters, putative hybrid and *C. limbata*. (b) Violin plots show the PC scores by the group for PC1 and PC2. The letters above the violins represent the results of multiple comparisons using the Steel–Dwass test, indicating significant differences among the groups. (c) Violin plots show the representative counting and metric traits. See [Figure S5](#) for other traits. The letters above the violins represent the results of multiple comparisons using the Steel–Dwass test, indicating significant differences among the groups.

endemic fish *Channa harcourtbutleri*. The results provide important insights into the evolutionary history including colonization, divergence and secondary contact events experienced by the species in this region ([Figure 5](#)).

4.1 | The origin and secondary contact with closely related species

Genome-wide SNP and mtDNA sequenced data successfully reconstructed the complex history of *C. harcourtbutleri*, which was derived from and experienced secondary contact with the widely distributed species *C. limbata* in a geological time framework. *Channa limbata* exhibited clear geographical differentiation, with its

population in the Irrawaddy or Salween River system, which were geographically near the Inle region, being most closely related to *C. harcourtbutleri* ([Figures 2 and 5a](#)). SNP-based results suggest that the colonization originated from the Irrawaddy River system, but it is not conclusive due to insufficient samples from the Salween River. The mtDNA monophyletic group (Clade 1+2) endemic to the Inle region was estimated to have diverged from their sister lineage at 2.0 Ma (Early Pleistocene), suggesting that the ancestral population of *C. harcourtbutleri* colonized the Inle region from the Irrawaddy or Salween River system around this time. This timing was slightly older than that estimated by [Rüber et al. \(2020\)](#) (1.4 Ma), which was obtained based on the phylogeny of mtDNA *cytb* and ribosomal RNA

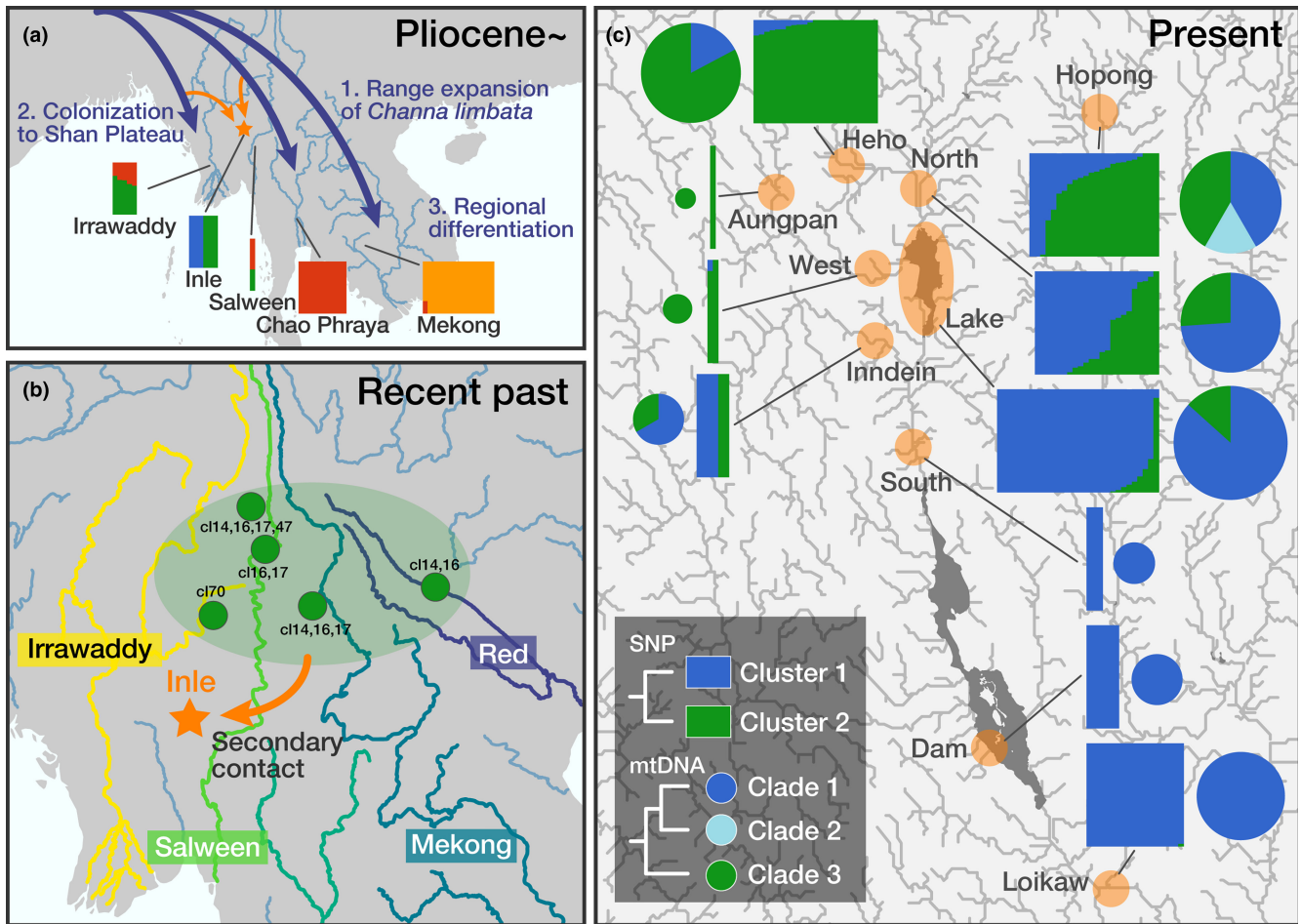


FIGURE 5 Graphical summary of the genetic analysis in this study. (a) The origin of *Channa harcourtbutleri* and its present genetic structure including *C. limbata* based on SNPs data. (b) Secondary contact from *C. limbata* to *C. harcourtbutleri* in the recent past (the late Pleistocene). (c) Genetic structure of *C. harcourtbutleri*. Bar plots show ancestral proportion in $K = 2$ based on SNPs data, and pie charts show the proportion of three mtDNA clades in the populations.

regions' and RAG1 sequences calibrated using the fossil records common to this study. However, both overlapped at credible intervals and supported their differentiation in the Early Pleistocene. The tMRCA for the *C. limbata* group including *C. harcourtbutleri* was estimated at ca. 4.5 Ma (Early Pliocene), also similar to (or slightly older than) the estimation (3.7 Ma) by Rüber et al. (2020), even though our samples included the basal populations of *C. limbata* from the Salween River system not used in the previously analysis. The *C. limbata* group is inferred to have originated in the Eastern Himalayan region (see Figure 3a) in Rüber et al., 2020. Following the range expansion of *C. limbata* towards Sundaland and East Asia probably from the Early Pliocene (tMRCA of the species group), the ancestor of *C. harcourtbutleri* could have been isolated on the Shan Plateau since at 2.0 Ma (Pleistocene). After the Indo–Eurasian collision in ca. 35 Ma (Aitchison et al., 2007), the Shan Plateau continued regional uplifting during the Oligocene and the Pliocene (Morley, 2009). Such geological formation processes of

the plateau might have impacted the isolation of the ancestral population of *C. harcourtbutleri*.

In addition to Clades 1 and 2 haplotypes, a part of *C. harcourtbutleri* specimens had highly differentiated haplotypes that composed the third clade (Clade 3) together with several haplotypes of *C. limbata* (Figure 5b). The occurrence of mtDNA Clade 3 in *C. harcourtbutleri* suggests that *C. limbata* re-immigrated to the Inle region after speciation of the former (i.e. secondary contact). This event would have occurred in the recent past, probably in the late Pleistocene, because the Clade 3 haplotypes from *C. harcourtbutleri* were identical or closely related (1–2 nucleotide differences, <0.2% in uncorrected p-distance) to those of *C. limbata* from the upper Irrawaddy, Salween, Mekong and Red rivers (Figure 3). This similarity is in contrast with their greater differentiation from the haplotypes of Clades 1 and 2 (3.1%–3.6% in uncorrected p-distance, roughly corresponding to 2.0 Ma). Two clusters revealed by genome-wide SNP data would reflect these multiple invasions of the *C. limbata* group; SNP Cluster 2, which mainly occurs

outside Inle Lake and the outflow river, is inferred to have originated from the second colonizers. The morphological similarity between the riverine populations and *C. limbata* supported this scenario. However, the genetic cohesion in the nuclear DNA of *C. harcourtbutleri* (Figure 2) suggested the partial genetic admixture in the Inle region after the second colonization. In this scenario, mtDNA Clade 2, the rare group that is a sister group of Clade 1, occurring only in Hopong together with Clades 1 and 3 (Figure 5), could be a trace of past geographical isolation.

The gene flow observed in the *C. limbata* group across river systems in the Indo-Burma region could be explained by the historical water system connections via river captures. The common haplotype found in the Inle region, upper Salween, upper Mekong and upper Red rivers (ch3-1 = cl16; this study; Wang et al., 2021) provides evidence of river captures in the upper regions of these three adjacent rivers (Clark et al., 2004) and dispersal of *C. limbata* among the water systems. There is no direct geological evidence of a river capture between the Irrawaddy and Salween rivers (Clark et al., 2004; Robinson et al., 2014); however, the shared haplotype in *C. limbata* in those areas (cl15; Wang et al., 2021) suggests that these water systems were connected in the past. The complex history of connections between the major rivers surrounding the Shan Plateau could have contributed to the exchange and differentiation of species in the Indo-Burma region (Glaubrecht & Köhler, 2004; He & Chen, 2006; Ratmuangkhwang et al., 2014; Rüber et al., 2004).

4.2 | Genetic and morphological differentiation within the Inle region

This study revealed the intraspecific differentiation of *C. harcourtbutleri*, that is, the genetic and morphological differences between the lake–outflow populations and the surrounding river populations. The two groups were inferred to have been formed via secondary invasion of the closely related species rather than through in situ differentiation. However, it should be noted that the second colonizers have rarely intruded into Inle Lake and its outflowing rivers. The association among genetic cohesion, morphological similarity and habitats in the two groups suggests that the intraspecific differentiation has been maintained by genetic isolation under divergent selection for adaptation to different environments (Heaney, 2000; Nosil, 2012; Schluter, 2000).

The specimens composing SNP Cluster 1, mainly distributed in the lake–outflow, have more elongated head and anal fin base and larger number of anal fin rays than those in specimens composing SNP Cluster 2, mainly occurring in the surrounding rivers, as well as *C. limbata*.

The latter two resembled each other in multiple morphological traits representing body and head shape; the results will be helpful in future taxonomic studies. The differences in these traits could reflect in some of the adaptations to flow velocity, space for swimming and food resources in different habitats (Friedman et al., 2020; Langerhans, 2008; Walker, 1997). Adaptation to contrasting environments often leads to reproductive isolation via restriction of dispersal and maladaptation in another environment (Nosil et al., 2005; Waters et al., 2020). Indeed, genetic admixture between the two SNP clusters was only partially observed at the sites where individuals of both clusters co-occurred (e.g. Hopong, North); overall, these SNP clusters were distributed exclusively to each other in the Inle region. This occurrence pattern suggests the presence of some degree of reproductive isolation between them. The pre-establishment of partial reproductive isolation between *C. harcourtbutleri* and *C. limbata* before secondary contact cannot be ruled out; however, the contrasting habitats of the two groups (lake vs. stream) associated with morphological differences imply that local adaptation has contributed to their isolation. On the other hand, SNP Clusters 1 and 2 showed similarities in some traits (dorsal fin rays and LLS), or the latter showed intermediate values between the former and *C. limbata* (anal fin rays and MW/SL). These could reflect local adaptation to unique environments common to the lake and rivers in this region and/or partial gene flow; however, the relationships between their morphology and adaptation could not be determined in this study.

The presence of mtDNA Clade 2 in Hopong provided insight into the formation of regional faunas in the Inle region and the diversity of distribution history among members of the faunas. Clade 2 showed differentiation corresponding to 1.11 (0.58–1.77) Ma from Clade 1, widely distributed in the Inle region; this range includes the formation of Inle Lake (1.5 Ma; Hampton et al., 2018). Similar to that in *C. harcourtbutleri*, regional genetic differentiation was reported in several endemic fish species in the Inle region (Fuke et al., 2022; Kano et al., 2022). Three of the five species examined showed genetic differentiation between Inle Lake and Hopong populations (Kano et al., 2022). Among these, the red dwarf rasbora, *Microrasbora rubescens* (Cyprinidae), endemic to the Inle region, showed a contrasting pattern to *C. harcourtbutleri*. In the latter, two genetic groups co-occur with morphological divergence, whereas the rasbora exhibited a large genetic differentiation (roughly corresponding to 2.7 Ma) among regions, without any morphological differentiation (Fuke et al., 2022). This morphological invariability could be explained through niche conservatism. The rasbora was adapted to a stagnant water environment; therefore, dispersal through rivers has been strongly limited. On the other hand, snakeheads, including *C. harcourtbutleri*, have

airbreathing organs and most of them have some degree of terrestrial migration ability (Musikasinthorn, 2003). These examples suggest the presence of several different mechanisms that could have generated the genetic, morphological and species diversity in the Inle region. Further studies on more species are necessary to better understand the biodiversity formation in this region. In addition, the adaptive significance of morphological and ecological characteristics and their contribution to reproductive isolation should be studied to determine the maintenance mechanisms of differentiated populations.

5 | CONCLUSION






The phylogeographical study of *Channa harcourtbutleri*, endemic to the Inle region, with closely related species specifically characterized Annandale's (1918) insights that the ichthyofauna of this region was formed via colonization and isolation from adjacent areas. The colonization events could have occurred at least twice; first as an immigration of the ancestral population of *C. harcourtbutleri* from the Irrawaddy or Salween River system to the Inle region (ca. 2 Ma), followed by the secondary contact of *C. limbata* in the relatively recent past. The coexistence of the old and new lineages could have been maintained through divergent adaptation to different environments. These multiple colonization patterns suggest that the Inle region is not strictly isolated from the adjacent large river systems and that historical connectivity of the upper Salween or Irrawaddy River system has influenced the formation of its ichthyofauna. However, the mountainous topography and the presence of several endemic species and genera in the Inle region imply that the degree of isolation was high enough to exclude frequent colonization. The long-lived, possibly stable environment and moderate isolation of the Inle region might have contributed to its unique ichthyofauna, where endemic and widely distributed non-endemic species coexist. Future studies on the origins of other endemic species will allow for an interspecific comparative approach and contribute to further understanding of the community assembly and evolutionary history of fish diversity in this region.

ACKNOWLEDGEMENTS

We deeply appreciate the kind assistance from Nyi Nyi Kyaw (Forest Department, Ministry of Natural Resources and Environmental Conservation, Myanmar [FD-MOECF]), Win Naing Thaw (Nature and Wildlife Conservation Division, FD-MOECF), the late Than Htay (Inle Lake Wildlife Sanctuary, FD-MOECF), the late Ohn and Than Nwai (Forest Resource Environment Development and Conservation Association, Myanmar [FREDA]), Nyi Nyi Lwin (Green Leaf, Myanmar) and Akihisa Iwata

(Kyoto University) in the fieldwork in and around Inle Lake. We are grateful to members of the Inland Fisheries Research and Development Institute (IFReDI, Cambodia), Yoshinori Kumazawa (Nagoya City University, Japan) and members of Research Laboratory of Ichthyology, Faculty of Fisheries, Kasetsart University (RLIKU, Thailand) for supporting the fieldwork in Southeast Asia. We also thank members of the Animal Ecology Laboratory, Kyoto University for supporting the molecular work and data analyses. This study was supported by the Research Foundation of Kurita Water and Environment Foundation (17B028, 18K013, 19K008), JSPS KAKENHI Grant (26304007, 19J23130) and the Sumitomo Foundation (the Grant for Environmental Research Project, 193271).

ORCID

Yusuke Fuke  <https://orcid.org/0000-0001-6404-6241>
 Yuichi Kano  <https://orcid.org/0000-0001-7158-753X>
 Ryoichi Tabata  <https://orcid.org/0000-0003-0550-0904>
 Shoko Matsui  <https://orcid.org/0000-0002-4065-4960>
 Katsutoshi Watanabe  <https://orcid.org/0000-0003-2244-2902>

REFERENCES

- Abell, R., Thieme, M. L., Revenga, C., Bryer, M., Kottelat, M., Bogutskaya, N., Coad, B., Mandrak, N., Balderas, S. C., Bussing, W., Stiassny, M. L. J., Skelton, P., Allen, G. R., Unmack, P., Naseka, A., Ng, R., Sindorf, N., Robertson, J., Armijo, E., ... Petry, P. (2008). Freshwater ecoregions of the world: A new map of biogeographic units for freshwater biodiversity conservation. *Bioscience*, 58, 403–414. <https://doi.org/10.1641/B580507>
- Adamson, E. A. S., Hurwood, D. A., & Mather, P. B. (2010). A reappraisal of the evolution of Asian snakehead fishes (Pisces, Channidae) using molecular data from multiple genes and fossil calibration. *Molecular Phylogenetics and Evolution*, 56, 707–717. <https://doi.org/10.1016/j.ympev.2010.03.027>
- Aitchison, J. C., Ali, J. R., & Davis, A. M. (2007). When and where did India and Asia collide? *Journal of Geophysical Research*, 112, B05423. <https://doi.org/10.1029/2006JB004706>
- Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, 19, 1655–1664. <https://doi.org/10.1101/gr.094052.109>
- Annandale, N. (1918). Fish and fisheries of the Inle Lake. *Records of the Indian Museum*, 14, 33–64, pls 1–7.
- Bouckaert, R., Vaughan, T. G., Barido-Sottani, J., Duchêne, S., Fourment, M., Gavryushkina, A., Heled, J., Jones, G., Kühnert, D., De Maio, N., Matschiner, M., Mendes, F. K., Müller, N. F., Ogilvie, H. A., du Plessis, L., Poppinga, A., Rambaut, A., Rasmussen, D., Siveroni, I., ... Drummond, A. J. (2019). BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. *PLoS Computational Biology*, 15, e1006650. <https://doi.org/10.1371/journal.pcbi.1006650>
- Brooks, J. (1950). Speciation in ancient lakes. *The Quarterly Review of Biology*, 25, 131–176.
- Cavender-Bares, J., Kozak, K. H., Fine, P. V. A., & Kembel, S. W. (2009). The merging of community ecology and phylogenetic

- biology. *Ecology Letters*, 12, 693–715. <https://doi.org/10.1111/j.1461-0248.2009.01314.x>
- Chen, S., Zhou, Y., Chen, Y., & Gu, J. (2018). fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*, 34, i884–i890. <https://doi.org/10.1093/bioinformatics/bty560>
- Clark, M. K., Schoenbohm, L. M., Royden, L. H., Whipple, K. X., Burchfiel, B. C., Zhang, X., Tang, W., Wang, E., & Chen, L. (2004). Surface uplift, tectonics, and erosion of eastern Tibet from large-scale drainage patterns. *Tectonics*, 23, TC1006. <https://doi.org/10.1029/2002TC001402>
- Clement, M., Posada, D., & Crandall, K. A. (2000). TCS: A computer program to estimate gene genealogies. *Molecular Ecology*, 9, 1657–1659. <https://doi.org/10.1046/j.1365-294x.2000.01020.x>
- Conte-Grand, C., Britz, R., Dahanukar, N., Raghavan, R., Pethiyagoda, R., Tan, H. H., Hadiaty, R. K., Yaakob, N. S., & Rüber, L. (2017). Barcoding snakeheads (Teleostei, Channidae) revisited: Discovering greater species diversity and resolving perpetuated taxonomic confusions. *PLoS One*, 12, e0184017. <https://doi.org/10.1371/journal.pone.0184017>
- Courtenay, W. R., & Williams, J. D. (2004). *Snakeheads (Pisces, Channidae): A biological synopsis and risk assessment*. U.S.
- Cox, C. B., Moore, P. D., & Ladle, R. J. (2020). *Biogeography: An ecological and evolutionary approach* (Tenth ed.). Wiley.
- Cristescu, M. E., Adamowicz, S. J., Vaillant, J. J., & Haffner, D. G. (2010). Ancient lakes revisited: From the ecology to the genetics of speciation. *Molecular Ecology*, 19, 4837–4851. <https://doi.org/10.1111/j.1365-294X.2010.04832.x>
- Cuvier, G., & Valenciennes. (1831). *Histoire naturelle des poissons*. Chez F. G. Levrault. <https://doi.org/10.5962/bhl.title.7339>
- Emerson, B. C., & Gillespie, R. G. (2008). Phylogenetic analysis of community assembly and structure over space and time. *Trends in Ecology & Evolution*, 23, 619–630. <https://doi.org/10.1016/j.tree.2008.07.005>
- Fricke, R., Eschmeyer, W. N., & van der Laan, R. (2022). *Eschmeyer's catalog of fishes: Genera, species, references*. Retrieved October 14, 2022, from <http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp>
- Friedman, S. T., Price, S. A., Corn, K. A., Larouche, O., Martinez, C. M., & Wainwright, P. C. (2020). Body shape diversification along the benthic–pelagic axis in marine fishes. *Proceedings of the Royal Society B: Biological Sciences*, 287, 20201053. <https://doi.org/10.1098/rspb.2020.1053>
- Fuke, Y., Kano, Y., Tun, S., Yun, L. K. C., Win, S. S., & Watanabe, K. (2022). Cryptic genetic divergence of the red dwarf rasbora, *Microrasbora rubescens*, in and around Inle Lake: Implications for the origin of endemism in the ancient lake in Myanmar. *Journal of Fish Biology*, 101, 1235–1247. <https://doi.org/10.1111/jfb.15195>
- Fuke, Y., Satoh, T. P., Kano, Y., & Watanabe, K. (2021). Annandale's collection of freshwater fishes from Inle Lake, Myanmar, housed in the Kyoto University Museum. *Ichthyological Research*, 68, 556–560. <https://doi.org/10.1007/s10228-021-00806-5>
- Gillespie, R. G., Bennett, G. M., De Meester, L., Feder, J. L., Fleischer, R. C., Harmon, L. J., Hendry, A. P., Knope, M. L., Mallet, J., Martin, C., Parent, C. E., Patton, A. H., Pfennig, K. S., Rubinoff, D., Schluter, D., Seehausen, O., Shaw, K. L., Stacy, E., Stervander, M., ... Wogan, G. O. U. (2020). Comparing adaptive radiations across space, time, and taxa. *Journal of Heredity*, 111, 1–20. <https://doi.org/10.1093/jhered/esz064>
- Glaubrecht, M., & Köhler, F. (2004). Radiating in a river: Systematics, molecular genetics and morphological differentiation of viviparous freshwater gastropods endemic to the Kaek River, central Thailand (Cerithioidea, Pachychilidae). *Biological Journal of the Linnean Society*, 82, 275–311. <https://doi.org/10.1111/j.1095-8312.2004.00361.x>
- Hamilton, F. (1822). *An account of the fishes found in the river Ganges and its branches*. Archibald Constable and company. <https://doi.org/10.5962/bhl.title.6897>
- Hampton, S. E., McGowan, S., Ozersky, T., Virdis, S. G. P., Vu, T. T., Spanbauer, T. L., Kraemer, B. M., Swann, G., Mackay, A. W., Powers, S. M., Meyer, M. F., Labou, S. G., O'Reilly, C. M., DiCarlo, M., Galloway, A. W. E., & Fritz, S. C. (2018). Recent ecological change in ancient lakes. *Limnology and Oceanography*, 63, 2277–2304. <https://doi.org/10.1002/lno.10938>
- Hauffe, T., Delicado, D., Etienne, R. S., & Valente, L. (2020). Lake expansion elevates equilibrium diversity via increasing colonization. *Journal of Biogeography*, 47, 1849–1860. <https://doi.org/10.1111/jbi.13914>
- He, D., & Chen, Y. (2006). Biogeography and molecular phylogeny of the genus *Schizothorax* (Teleostei: Cyprinidae) in China inferred from cytochrome *b* sequences. *Journal of Biogeography*, 33, 1448–1460. <https://doi.org/10.1111/j.1365-2699.2006.01510.x>
- Heaney, L. R. (2000). Dynamic disequilibrium: A long-term, large-scale perspective on the equilibrium model of island biogeography. *Global Ecology and Biogeography*, 9, 59–74. <https://doi.org/10.1046/j.1365-2699.2000.00163.x>
- Henning, F., & Meyer, A. (2014). The evolutionary genomics of cichlid fishes: Explosive speciation and adaptation in the postgenomic era. *Annual Review of Genomics and Human Genetics*, 15, 417–441. <https://doi.org/10.1146/annurev-genom-090413-025412>
- Hirao, T., Kubota, Y., & Murakami, M. (2015). Geographical patterns of butterfly species diversity in the subtropical Ryukyu Islands: The importance of a unidirectional filter between two source islands. *Journal of Biogeography*, 42, 1418–1430. <https://doi.org/10.1111/jbi.12501>
- Hubbs, C. L., & Lagler, K. F. (1958). *Fishes of the Great Lakes region*. University of Michigan Press.
- Huson, D. H., & Bryant, D. (2006). Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution*, 23, 254–267. <https://doi.org/10.1093/molbev/msj030>
- Johnson, M. A., Pillon, Y., Sakishima, T., Price, D. K., & Stacy, E. A. (2019). Multiple colonizations, hybridization and uneven diversification in *Cyrtandra* (Gesneriaceae) lineages on Hawai'i Island. *Journal of Biogeography*, 46, 1178–1196. <https://doi.org/10.1111/jbi.13567>
- Kalyanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A., & Jermini, L. S. (2017). ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nature Methods*, 14, 587–589. <https://doi.org/10.1038/nmeth.4285>
- Kano, Y., Adnan, M. S., Grudpan, C., Grudpan, J., Magtoon, W., Musikasinthorn, P., Natori, Y., Ottomanski, S., Praxaysonbath, B., Phongsa, K., Rangsiruji, A., Shibukawa, K., Shimatani, Y., So, N., Suvarnaksha, A., Thach, P., Thanh, P. N., Tran, D. D., Utsugi, K., & Yamashita, T. (2013). An online database on freshwater fish diversity and distribution in Mainland Southeast Asia. *Ichthyological Research*, 60, 293–295. <https://doi.org/10.1007/s10228-013-0349-8>
- Kano, Y., Fuke, Y., Musikasinthorn, P., Iwata, A., Soe, T. M., Tun, S., ... Watanabe, K. (2022). Fish diversity of a spring field in Hopong Town, Taunggyi District, Shan State, Myanmar (the Salween River Basin), with genetic comparisons to some

- “species endemic to Inle Lake”. *Biodiversity Data Journal*, 10, e80101. <https://doi.org/10.3897/BDJ.10.e80101>
- Kano, Y., Musikasinthorn, P., Iwata, A., Tun, S., Yun, L. K. C., Win, S. S., ... Watanabe, K. (2016). A dataset of fishes in and around Inle Lake, an ancient lake of Myanmar, with DNA barcoding, photo images and CT/3D models. *Biodiversity Data Journal*, 4, e10539. <https://doi.org/10.3897/BDJ.4.e10539>
- Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution*, 30, 772–780. <https://doi.org/10.1093/molbev/mst010>
- Kimura, T., Yamada, T., Sakaguchi, S., Ito, M., & Maki, M. (2022). Multiple colonizations and genetic differentiation in goldenrod populations on recently formed nearshore islands. *Journal of Biogeography*, 49, 836–852. <https://doi.org/10.1111/jbi.14342>
- Klingenberg, C. P. (2011). MorphoJ: An integrated software package for geometric morphometrics. *Molecular Ecology Resources*, 11, 353–357. <https://doi.org/10.1111/j.1755-0998.2010.02924.x>
- Langerhans, R. B. (2008). Predictability of phenotypic differentiation across flow regimes in fishes. *Integrative and Comparative Biology*, 48, 750–768. <https://doi.org/10.1093/icb/icn092>
- Laskar, B. A., Adimalla, H., Kundu, S., Jaiswal, D., & Chandra, K. (2023). Genetic evidence on the occurrence of *Channa harcourtbutleri* (Annandale, 1918) in Eastern Ghats, India: First report from mainland India. *Journal of Threatened Taxa*, 15, 22834–22840. <https://doi.org/10.11609/jott.6894.15.3.22834-22840>
- Leigh, J. W., & Bryant, D. (2015). POPART: Full-feature software for haplotype network construction. *Methods in Ecology and Evolution*, 6, 1110–1116. <https://doi.org/10.1111/2041-210X.12410>
- Losos, J. B., & Ricklefs, R. E. (2009). Adaptation and diversification on islands. *Nature*, 457, 830–836. <https://doi.org/10.1038/nature07893>
- MacArthur, R. H., & Wilson, E. O. (1963). An equilibrium theory of insular zoogeography. *Evolution*, 17, 373–387. <https://doi.org/10.1111/j.1558-5646.1963.tb03295.x>
- MacArthur, R. H., & Wilson, E. O. (1967). *The theory of island biogeography*. Princeton University Press.
- Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., von Haeseler, A., & Lanfear, R. (2020). IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *Molecular Biology and Evolution*, 37, 1530–1534. <https://doi.org/10.1093/molbev/msaa015>
- Mittelbach, G. G., & Schemske, D. W. (2015). Ecological and evolutionary perspectives on community assembly. *Trends in Ecology & Evolution*, 30, 241–247. <https://doi.org/10.1016/j.tree.2015.02.008>
- Morley, C. K. (2009). Geometry and evolution of low-angle normal faults (LANF) within a Cenozoic high-angle rift system, Thailand: Implications for sedimentology and the mechanisms of LANF development. *Tectonics*, 28, TC5001. <https://doi.org/10.1029/2007TC002202>
- Musikasinthorn, P. (1998). *Channa panaw*, a new channid fish from the Irrawaddy and Sittang River basins, Myanmar. *Ichthyological Research*, 45, 355–362. <https://doi.org/10.1007/BF02725188>
- Musikasinthorn, P. (2003). Channoidei (snakeheads). In M. Hutchins, A. Thoney, P. V. Loiselle, & N. Schlager (Eds.), *Grzimek's animal life encyclopedia, fishes I-II* (Vol. 4–5, 2nd ed.). Gale.
- Ng, H. H., Ng, P. K. L., & Britz, R. (1999). *Channa harcourtbutleri* (Annandale, 1918): A valid species of snakehead (Perciformes: Channidae) from Myanmar. *Journal of South Asian Natural History*, 4, 57–63.
- Nosil, P. (2012). *Ecological speciation*. Oxford University Press.
- Nosil, P., Vines, T. H., & Funk, D. J. (2005). Reproductive isolation caused by natural selection against immigrants from divergent habitats. *Evolution*, 59, 705–719. <https://doi.org/10.1111/j.0014-3820.2005.tb01747.x>
- Okonechnikov, K., Golosova, O., & Fursov, M. (2012). Unipro UGENE: A unified bioinformatics toolkit. *Bioinformatics*, 28, 1166–1167. <https://doi.org/10.1093/bioinformatics/bts091>
- Onuki, K., & Fuke, Y. (2022). Rediscovery of a native freshwater shrimp, *Neocaridina denticulata*, and expansion of an invasive species in and around Lake Biwa, Japan: Genetic and morphological approach. *Conservation Genetics*, 23, 967–980. <https://doi.org/10.1007/s10592-022-01467-1>
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J., Sklar, P., de Bakker, P. I. W., Daly, M. J., & Sham, P. C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics*, 81, 559–575. <https://doi.org/10.1086/519795>
- R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Rahbek, C., Borregaard, M. K., Antonelli, A., Colwell, R. K., Holt, B. G., Noguez-Bravo, D., CMØ, R., Richardson, K., Rosing, M. T., Whittaker, R. J., & Fjeldsø, J. (2019). Building mountain biodiversity: Geological and evolutionary processes. *Science*, 365, 1114–1119. <https://doi.org/10.1126/science.aax0151>
- Rambaut, A., Drummond, A. J., Xie, D., Baele, G., & Suchard, M. A. (2018). Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology*, 67, 901–904. <https://doi.org/10.1093/sysbio/syy032>
- Ratmuangkhwang, S., Musikasinthorn, P., & Kumazawa, Y. (2014). Molecular phylogeny and biogeography of air sac catfishes of the *Heteropneustes fossilis* species complex (Siluriformes: Heteropneustidae). *Molecular Phylogenetics and Evolution*, 79, 82–91. <https://doi.org/10.1016/j.ympev.2014.05.009>
- Ravinet, M., Prodöhl, P. A., & Harrod, C. (2013). Parallel and non-parallel ecological, morphological and genetic divergence in lake - stream stickleback from a single catchment. *Journal of Evolutionary Biology*, 26, 186–204. <https://doi.org/10.1111/jeb.12049>
- Ricklefs, R. E. (1987). Community diversity: Relative roles of local and regional processes. *Science*, 235, 167–171. <https://doi.org/10.1126/science.235.4785.167>
- Roberts, T. R. (1993). The freshwater fishes of Java, as observed by Kuhl and van Hasselt in 1820–23. *Zoologische Verhandlungen*, 285, 3–94.
- Robinson, R. A. J., Brezina, C. A., Parrish, R. R., Horstwood, M. S. A., Oo, N. W., Bird, M. I., Thein, M., Walters, A. S., Oliver, G. J. H., & Zaw, K. (2014). Large rivers and orogens: The evolution of the Yarlung Tsangpo–Irrawaddy system and the eastern Himalayan syntaxis. *Gondwana Research*, 26, 112–121. <https://doi.org/10.1016/j.gr.2013.07.002>
- Rochette, N. C., Rivera-Colón, A. G., & Catchen, J. M. (2019). Stacks 2: Analytical methods for paired-end sequencing improve

- RADseq-based population genomics. *Molecular Ecology*, 28, 4737–4754. <https://doi.org/10.1111/mec.15253>
- Rosindell, J., & Phillimore, A. B. (2011). A unified model of island biogeography sheds light on the zone of radiation: A unified model of island biogeography. *Ecology Letters*, 14, 552–560. <https://doi.org/10.1111/j.1461-0248.2011.01617.x>
- Rüber, L., Britz, R., Kullander, S. O., & Zardoya, R. (2004). Evolutionary and biogeographic patterns of the Badidae (Teleostei: Perciformes) inferred from mitochondrial and nuclear DNA sequence data. *Molecular Phylogenetics and Evolution*, 32, 1010–1022. <https://doi.org/10.1016/j.ympev.2004.04.020>
- Rüber, L., Tan, H. H., & Britz, R. (2020). Snakehead (Teleostei: Channidae) diversity and the Eastern Himalaya biodiversity hotspot. *Journal of Zoological Systematics and Evolutionary Research*, 58, 356–386. <https://doi.org/10.1111/jzs.12324>
- Schluter, D. (2000). *The ecology of adaptive radiation*. Oxford University Press.
- Seehausen, O. (2006). African cichlid fish: A model system in adaptive radiation research. *Proceedings of the Royal Society B: Biological Sciences*, 273, 1987–1998. <https://doi.org/10.1098/rspb.2006.3539>
- Silva, K. H. G. M. (1991). Population ecology of the paddy field-dwelling fish *Channa gachua* (Gunther) (Perciformes, Channidae) in Sri Lanka. *Journal of Fish Biology*, 38, 497–508. <https://doi.org/10.1111/j.1095-8649.1991.tb03137.x>
- Šlechtová, V., Musilova, Z., Tan, H. H., Kottelat, M., & Bohlen, J. (2021). One northward, one southward: Contrasting biogeographical history in two benthic freshwater fish genera across Southeast Asia (Teleostei: Cobitoidea: *Nemacheilus*, *Pangio*). *Molecular Phylogenetics and Evolution*, 161, 107139. <https://doi.org/10.1016/j.ympev.2021.107139>
- Spirorello, M., & Brooks, D. R. (2003). Dispersal and diversification: Macroevolutionary implications of the MacArthur-Wilson model, illustrated by *Simulium* (*Inseliellum*) Rubstov (Diptera: Simuliidae): Radiation of Pacific black flies. *Journal of Biogeography*, 30, 1563–1573. <https://doi.org/10.1046/j.1365-2699.2003.00945.x>
- Suyama, Y., Hirota, S. K., Matsuo, A., Tsunamoto, Y., Mitsuyuki, C., Shimura, A., & Okano, K. (2022). Complementary combination of multiplex high-throughput DNA sequencing for molecular phylogeny. *Ecological Research*, 37, 171–181. <https://doi.org/10.1111/1440-1703.12270>
- Suyama, Y., & Matsuki, Y. (2015). MIG-seq: An effective PCR-based method for genome-wide single-nucleotide polymorphism genotyping using the next-generation sequencing platform. *Scientific Reports*, 5, 16963. <https://doi.org/10.1038/srep16963>
- Tabata, R., Kakioka, R., Tominaga, K., Komiya, T., & Watanabe, K. (2016). Phylogeny and historical demography of endemic fishes in Lake Biwa: The ancient lake as a promoter of evolution and diversification of freshwater fishes in western Japan. *Ecology and Evolution*, 6, 2601–2623. <https://doi.org/10.1002/ece3.2070>
- Takahashi, H., Møller, P. R., Shedko, S. V., Ramatulla, T., Joen, S.-R., Zhang, C.-G., Sideleva, V. G., Takata, K., Sakai, H., Goto, A., & Nishida, M. (2016). Species phylogeny and diversification process of Northeast Asian *Pungitius* revealed by AFLP and mtDNA markers. *Molecular Phylogenetics and Evolution*, 99, 44–52. <https://doi.org/10.1016/j.ympev.2016.03.022>
- Theis, A., Ronco, F., Indermaur, A., Salzburger, W., & Egger, B. (2014). Adaptive divergence between lake and stream populations of an East African cichlid fish. *Molecular Ecology*, 23, 5304–5322. <https://doi.org/10.1111/mec.12939>
- Toke, Y. Y., Aung, N. N., Oo, K. S., Tun, K. M., & Gyi, M. M. (2013). Fish species diversity according to spatial habitat degradation of Inle Lake. In G. T. Pramono, D. Ramdani, B. Barus, & R. M. Ariansyah (Eds.), *34th Asian Conference on Remote Sensing* (pp. 1766–1771). Asian Association on Remote Sensing.
- Walker, J. A. (1997). Ecological morphology of lacustrine threespine stickleback *Gasterosteus aculeatus* L. (Gasterosteidae) body shape. *Biological Journal of the Linnean Society*, 61, 3–50. <https://doi.org/10.1111/j.1095-8312.1997.tb01777.x>
- Wang, J., Li, C., Chen, J., Wang, J., Jin, J., Jiang, S., Yan, L., Lin, H.-D., & Zhao, J. (2021). Phylogeographic structure of the dwarf snakehead (*Channa gachua*) around Gulf of Tonkin: Historical biogeography and pronounced effects of sea-level changes. *Ecology and Evolution*, 11, 12583–12595. <https://doi.org/10.1002/ece3.8003>
- Watanabe, K., Kano, Y., Takahashi, H., Mukai, T., Kakioka, R., & Tominaga, K. (2010). GEDIMAP: A database of genetic diversity for Japanese freshwater fishes. *Ichthyological Research*, 57, 107–109. <https://doi.org/10.1007/s10228-009-0129-7>
- Watanabe, K., Tabata, R., Nakajima, J., Kobayakawa, M., Matsuda, M., Takaku, K., Hosoya, K., Ohara, K., Takagi, M., & Jang-Liaw, N.-H. (2020). Large-scale hybridization of Japanese populations of Hinamoroko, *Aphyocypris chinensis*, with *A. kikuchii* introduced from Taiwan. *Ichthyological Research*, 67, 361–374. <https://doi.org/10.1007/s10228-019-00730-9>
- Waters, J. M., Emerson, B. C., Arribas, P., & McCulloch, G. A. (2020). Dispersal reduction: Causes, genomic mechanisms, and evolutionary consequences. *Trends in Ecology & Evolution*, 35, 512–522. <https://doi.org/10.1016/j.tree.2020.01.012>
- Whittaker, R. J., Fernández-Palacios, J. M., Matthews, T. J., Borregaard, M. K., & Triantis, K. A. (2017). Island biogeography: Taking the long view of nature's laboratories. *Science*, 357, eaam8326. <https://doi.org/10.1126/science.aam8326>
- Whittaker, R. J., Triantis, K. A., & Ladle, R. J. (2008). A general dynamic theory of oceanic island biogeography: A general dynamic theory of oceanic island biogeography. *Journal of Biogeography*, 35, 977–994. <https://doi.org/10.1111/j.1365-2699.2008.01892.x>
- Wu, F., He, D., Fang, G., & Deng, T. (2019). Into Africa via docked India: A fossil climbing perch from the Oligocene of Tibet helps solve the anabantid biogeographical puzzle. *Science Bulletin*, 64, 455–463. <https://doi.org/10.1016/j.scib.2019.03.029>

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Fuke, Y., Musikasinthorn, P., Kano, Y., Tabata, R., Matsui, S., Tun, S., Yun, L. K. C., Touch, B., Thach, P., & Watanabe, K. (2024). Origin of endemic species in a moderately isolated ancient lake: The case of a snakehead in Inle Lake, Myanmar. *Zoologica Scripta*, 53, 16–31. <https://doi.org/10.1111/zsc.12633>