Delineating freshwater fish biogeographic regions in the Philippine archipelago

Brian Wade Jamandre[1*], Nico Jose Leander[2], Rodulf Anthony Balisco[3], Faith Santiago-Tadeo[4]

[1]National Fish and Wildlife Foundation, Washington, D.C., USA; [2]Bureau of Fisheries and Aquatic Resources III, Philippines; [3]College of Fisheries and Natural Sciences, Western Philippines University, Philippines; [4]College of Fisheries, Don Mariano Marcos Memorial State University, Philippines

*Corresponding Author: bwjamandre@gmail.com

Abstract

Aim

To delineate native freshwater fish biogeographic regions in the Philippine archipelago and identify the contemporary and historical factors that structure freshwater fish diversity, endemism, and species turnover.

Location

The Philippine archipelago

Taxon

Native Freshwater Fishes

Methods

We compiled an archipelago-wide dataset of native freshwater fish occurrences and integrated these with environmental variables and historical biogeographic groupings, including Pleistocene Aggregate Island Complex (PAIC) configurations. Species distributions, diversity, and endemism were analyzed using cluster analysis of a Simpson's beta (βsim) dissimilarity matrix to delineate provinces, with provincial assignments interpreted in light of the PAIC framework and prior biogeographic evidence. Inventory completeness per province was quantified using coverage-based rarefaction (iNEXT). Generalized Dissimilarity Modelling (GDM) was used to quantify the relative contributions of geographic, environmental, and historical predictors to compositional turnover, with predictor importance and statistical significance assessed by matrix permutation. Indicator species analysis (IndVal) was applied to both the full dataset and the

obligate primary-freshwater (PF) subset, and dendrogram congruence between the two analyses was tested with a Mantel test and cophenetic correlation.

Results

Cluster analysis resolved five freshwater fish biogeographic provinces: Northern Luzon, Central/Southern Luzon, Greater Visayas, Mindanao, and Palawan–Mindoro. Geographic distance was the strongest correlate of compositional turnover, while PAIC grouping and contemporary climate variables contributed comparatively little once geographic distance was accounted for. Two patterns stood out. First, the Visayan islands of Leyte and Samar clustered with the rest of the Greater Visayas rather than with Greater Mindanao, despite their conventional PAIC assignment to the latter, a pattern consistent with contemporary marine straits exerting biogeographic influence beyond what Pleistocene land-bridge predicts. Second, Palawan–Mindoro stood out as a Sundaic outlier within the Philippine archipelago, anchored by obligate primary-freshwater Cyprinidae of Bornean affinity. By contrast, Greater Visayas lacked any significant obligate-freshwater indicator and was characterized instead by an estuarine–freshwater interface assemblage, supporting the interpretation of this province as a hydrologically open, marine-influenced biogeographic system.

Main Conclusions

This study provides the first comprehensive biogeographic framework for Philippine freshwater fishes, clarifying how distance, palaeogeographic history, and habitat context are associated with assemblage boundaries in a complex oceanic archipelago. The contrasting biogeographic character of Palawan–Mindoro and Greater Visayas illustrates how a single archipelago can host fundamentally distinct freshwater biogeographic modes shaped by geological origin and

contemporary hydrological characteristics. The framework supplies a spatial template for conservation planning in Philippine inland waters and more broadly illustrates the value of integrating historical and contemporary drivers to explain freshwater biogeographic patterns in island systems.

Keywords

assemblage; islands; map; Southeast Asia; tropical

Introduction

Freshwater ecosystems cover less than 1% of the Earth's surface but support a disproportionately high percentage of global biodiversity (Dudgeon et al., 2006; Strayer and Dudgeon, 2010). The Philippines, an archipelago of over 7,600 islands, is no exception to this pattern. With its complex geography, tropical climate, and unique geological history, the country boasts a rich array of freshwater habitats, from mountain streams to lowland rivers and lakes, each supporting distinct assemblages of aquatic life shaped by both contemporary ecological processes and historical biogeographic events.

The concept of biogeographic regions, areas characterized by distinctive evolutionary histories, species assemblages, and environmental features, has proven invaluable in understanding large-scale patterns of biodiversity distribution and evolution (Holt et al., 2013). In the context of freshwater systems, biogeographic regions can be particularly useful for elucidating the interplay between current ecological conditions and historical processes in shaping the distribution patterns of aquatic organisms, including fish (Abell et al., 2008; Stelbrink et al., 2012; Tedesco et al., 2017).

While extensive research on freshwater biogeography has been conducted in continental systems, the delineation of such regions in complex archipelagos like the Philippines remains largely unexplored. This knowledge gap is particularly striking given the Philippines' status as a biodiversity hotspot. The archipelago's location at the intersection of several biogeographic realms, coupled with its geological history of island formation, isolation, and connectivity, has led to high levels of endemism and unique distribution patterns among its fauna (Heaney, 1986; Vallejo, 2011; Brown et al., 2013).

The Philippines presents a unique challenge and opportunity for biogeographic analysis. Its fragmented landmass, varied topography, and diverse climatic conditions have created a mosaic of freshwater habitats. From the extensive river systems of Luzon and Mindanao, to the numerous crater lakes scattered across the archipelago, each water body potentially hosts distinct fish communities shaped by local environmental conditions, dispersal events, and long-term evolutionary processes. Moreover, the archipelago's geological history is marked by repeated cycles of connection and isolation driven by sea-level fluctuations during the Pleistocene. During glacial maxima, sea levels dropped by approximately at least 120 m below present (Voris, 2000; MGB, 2010), exposing the shallow continental shelves and uniting adjacent islands into larger landmasses known as Pleistocene Aggregate Island Complexes (PAICs) (Heaney, 1985; Robles, 2013). These PAICs (Figure 1) included Greater Luzon, Greater Palawan, Mindoro, Greater Negros-Panay, Greater Mindanao, and the Sulu Archipelago. During low-stands, expanded drainage systems would have enabled the colonization and exchange of freshwater fishes among islands within each PAIC; during interglacials, rising seas re-established marine barriers and imposed renewed isolation, promoting lineage divergence and endemism (Brown and Diesmos, 2002). The fossil record of Philippine freshwater fishes, though sparse, supports this dynamic history (Mediodia et al., 2025; Přikryl et al., 2025). These PAIC dynamics add a fundamental layer of historical complexity to the biogeographic patterns of the freshwater fauna across the archipelago.

Despite the ecological and evolutionary importance of Philippine freshwater ecosystems, comprehensive studies on the distribution patterns and biogeographic classification of freshwater fishes across the archipelago are lacking. While there have been localized surveys and taxonomic studies (e.g., Herre, 1924; Conlu, 1986), a broader, integrative approach to understanding

freshwater fish distributions at the national scale, incorporating both ecological and historical factors, remains to be undertaken. Thus, this study aims to address this knowledge gap by: (1) compiling and analyzing existing data on freshwater fish distributions across the Philippines, (2) identifying distinct freshwater fish biogeographic regions based on similarities in species composition, environmental factors, and historical connectivity, (3) investigating the relationships between fish distribution patterns and key environmental variables, and (4) examining the influence of historical biogeography, particularly PAICs, on current distribution patterns. Lastly, the implications of these biogeographic regions for understanding evolutionary processes, conservation planning, and management in the Philippines are discussed.

Materials and Methods

Study Area

The Philippines is an archipelago located in Southeast Asia and comprises 7,641 islands (Baldia et al., 2017; Romeroso et al., 2021), with a total land area of approximately 300,000 km² (Figure 1) (Amoroso, 2012). The archipelago is characterized by complex topography, ranging from coastal lowlands to mountain ranges. The geological history of the Philippine archipelago is complex and diverse, playing a crucial role in shaping its current biodiversity patterns. The islands have a heterogeneous origin, with some formed through volcanic activity, others through tectonic uplift, and some with continental origins (Voris, 2000; Hall, 2002). The archipelago can be broadly divided into three geological components: the eastern portion, including parts of Luzon and eastern Mindanao, originated from volcanic island arcs formed during the Cretaceous to early Paleogene periods (Voris, 2000; Pubellier et al., 2004). The western portion, including Palawan and parts of Mindoro, is composed of continental fragments that rifted from the southeastern margin of continental Asia during the opening of the South China Sea in the

Oligocene (Yumul et al., 2003). The central Philippines, including the Visayas, represent a complex amalgamation of island arc systems and ophiolite complexes (Aurelio et al., 2013).

Catchment Boundaries and Environmental Data

Catchment boundaries or units were delineated using multiple sources. The primary catchment unit used in the analysis was level-7 HydroBASINS (Linke et al., 2019), supplemented by data from Boothroyd et al. (2023). Boundaries were further refined based on traditional river delineations (e.g., Cagayan River, Agno, and Pampanga Rivers) and the Philippine ecogeographic maps (BMB-DENR, 2015) (Table S2). Integration of smaller tributaries and poorly researched rivers considered fish distribution quality, geomorphological features, and coastal river size. QGIS ver. 3.44 (QGIS, 2023) was utilized for catchment unit delineation.

Three categories of environmental data were collected to explore potential drivers of freshwater fish distribution patterns. Geographic and topographic variables included geographic distance between catchments, catchment area, mean stream length, drainage density, mean catchment elevation and slope (Table S3). Topographic data were obtained from Boothroyd et al. (2023). Catchment area and stream length were included because they scale with habitat availability and species richness in island freshwater systems, with larger catchments supporting greater habitat heterogeneity (He and Legendre, 2002; Tisseuil et al., 2013). Drainage density was included as a proxy for within-catchment connectivity and the degree of habitat subdivision, which influences dispersal opportunities and population isolation in fragmented island drainages (Leprieur et al., 2011). Mean elevation and slope were included as proxies for habitat type (lowland floodplain versus upland torrent), thermal regime, and flow regime — all known to structure fish assemblages along elevational gradients in Southeast Asian Island systems (Ng and

Kottelat, 2007). Geographic distance (Euclidean centroid distance between catchments) was included to capture isolation-by-distance effects, which are expected to be strong in archipelago settings where dispersal requires crossing marine barriers (Setiawan et al., 2020). Climate data were sourced from the Philippine Atmospheric, Geophysical and Astronomical Services Administration (PAGASA, 2024), including annual mean and maximum temperature, annual precipitation, and precipitation seasonality (Table S4). Annual mean and maximum temperature were selected because they set physiological limits on the distribution of ectothermic freshwater fishes and are among the strongest climatic predictors of fish assemblage composition globally (Heino et al., 2009; Tisseuil et al., 2013). Annual precipitation and precipitation seasonality were included because they determine river discharge, flood-pulse dynamics, and dry-season habitat availability, all of which structure fish assemblage composition and dispersal windows in tropical river systems (Junk et al., 1989; Tedesco et al., 2017). To test the PAIC hypothesis, the classification of Pleistocene Aggregate Island Complexes as defined by Brown and Diesmos (2002) was used, allowing investigation of whether freshwater fish diversity patterns align with the archipelago's paleo-geographic history. Catchment units were categorized based on which PAIC group they belong to (Table S2, Table S3).

Fish Data

A comprehensive dataset of freshwater fish species occurrences across the Philippines was compiled. Data collection encompassed historical surveys, including the seminal works by Herre (1924, 1953), which provided a foundational understanding of Philippine freshwater fish distributions and taxonomy. Recent taxonomic studies were incorporated, with the primary species list derived from Jamandre (2023) and Balisco and Liao (2025), which themselves document voucher material, localities, and collection dates for the field-collected records

underlying the present dataset. Occurrence data were extracted from global biodiversity databases including the Global Biodiversity Information Facility (GBIF, 2024) and FishNet2 (2024), supplemented by peer-reviewed literature, regional ichthyological journals, ecological surveys, and conservation reports.

The dataset includes all native freshwater fishes, comprising primary freshwater (PF), diadromous (DI), and euryhaline (EU) species that spend at least part of their life cycle in freshwater environments. Each species was assigned to one of these three ecological categories based first on direct evidence from species-specific life-history studies where available, and second on inference from closely related congeners following the criteria of Berra (2007), Kottelat (2013), the FishBase migration category, and species-specific revisionary literature where recent updates were available (e.g., Keith and Mennesson, 2023, for Indo-Pacific *Hypseleotris*). These categories are necessarily simplifications: many tropical species exhibit life-history polymorphism, the boundary between DI and EU is often ambiguous, and detailed migration studies of Philippine populations are largely absent for most taxa. The categories used here should therefore be regarded as best estimated current inferences rather than empirically verified life-history types. Non-native and introduced species were excluded from the analysis (Leprieur et al., 2008; Vitule et al., 2009; Olden et al., 2011). Prior to analysis, the species list was cross-referenced against Eschmeyer's Catalog of Fishes (Fricke et al., 2024), FishBase (Froese and Pauly, 2024), and recent revisionary literature to identify and correct taxonomic synonymies, misidentifications, and outdated family or ecological-category assignments. The final dataset comprises 299 species (Table S1), of which 79 are obligate primary freshwater (PF), 49 are diadromous (DI), and 171 are euryhaline (EU).

We note that sampling effort across the archipelago is uneven, being historically more intensive in Luzon than in Visayan and Mindanao headwaters. Simpson's beta dissimilarity (βsim) was chosen as the principal compositional metric for its robustness to unequal sampling effort (Tuomisto, 2010), and inventory completeness per province was quantified explicitly using coverage-based rarefaction (see Sample Coverage Assessment below). The dataset reflects the current state of knowledge of Philippine freshwater fish taxonomy and distribution, which remains incomplete; entries in Table S1 likely include some unresolved synonymies, misidentifications carried over from older records, and species *inquirendae* whose validity awaits re-examination. The most uncertain entries are individually flagged in Table S1.

Cluster Analysis

A pairwise matrix of Simpson's beta (βsim) species turnover was generated for all catchment pairs. βsim was chosen to minimize the effect of unequal sampling effort across the study area (Tuomisto, 2010; Shelley et al., 2019). Using the βsim distance matrix, an agglomerative cluster analysis was performed to generate a WPGMA (weighted pair-group method using arithmetic averages) hierarchical dendrogram in the software Biodiverse ver. 5.0 (Laffan et al., 2010). The PF-only WPGMA dendrogram and the tanglegram comparing the full-dataset and PF-only analyses were computed using the equivalent WPGMA algorithm implemented in Biodiverse.

Biogeographic provinces were identified by examining the WPGMA dendrogram in conjunction with the Pleistocene Aggregate Island Complex framework (Table S2) and prior knowledge of Philippine Island biogeography (Brown et al., 2013; Brown and Diesmos, 2002), rather than by a single uniform-height cut. This approach is standard practice in biogeographic

regionalization (e.g., Holt et al., 2013; Vilhena and Antonelli, 2015) and allowed finer-scale structure to be recognized where supported by independent geological and ecological evidence.

Sample Coverage Assessment

To quantify inventory completeness across the five biogeographic provinces and assess whether differential sampling effort could confound subsequent biogeographic interpretation, we computed sample coverage ($\hat{C}$) using the Chao-Jost coverage-based framework (Chao and Jost, 2012; Chao et al., 2014). For each province we treated the constituent catchments as incidence-based sampling units and computed (i) observed species richness; (ii) sample coverage $\hat{C}$ at the observed sample size; and (iii) asymptotic species richness (Chao 2 estimator) with 95% bootstrap confidence intervals (200 bootstrap replicates). Analyses were conducted in R using the iNEXT package version 3.0 (Hsieh, Ma, and Chao, 2016).

Species Richness and Corrected Weighted Endemism

Species richness (SR) and corrected weighted endemism (CWE) were calculated for each catchment using Biodiverse ver. 5.0. SR represents the number of species present in each catchment, providing a measure of local diversity. CWE was calculated in relation to the entire Philippine archipelago, allowing identification of areas of high endemism within the country. CWE weights species by the inverse of their range size, giving higher weight to species with restricted distributions. This metric is designed to identify catchments that not only have a higher number of endemic species but also harbor species found nowhere else in the Philippines.

Generalized Dissimilarity Modelling

Generalized Dissimilarity Modelling (GDM) was used to investigate the environmental and historical correlates of fish compositional turnover (Ferrier et al., 2007; Mokany et al.,

2022). GDM models pairwise compositional dissimilarity between sites as a non-linear function of pairwise environmental, geographic, and categorical predictors, using monotonic I-spline basis functions to capture rates of compositional change along each predictor gradient. GDM was analysed using the 'gdm' package (Manion et al., 2018) implemented in R (R Core Team, 2023).

Candidate predictors were drawn from three blocks: (i) geographic/topographic (geographic distance, catchment area, mean elevation, mean stream length, drainage density, mean stream slope), (ii) climatic (annual rainfall, maximum temperature, minimum temperature, mean temperature, precipitation seasonality), and (iii) historical biogeography (PAIC group) (Table S3 and S4). Continuous predictors were screened for collinearity using pairwise Spearman rank correlations; for any pair with $|\rho| > 0.7$, the more derived or less ecologically interpretable variable was iteratively removed until no pair exceeded the threshold. Specifically, mean stream length ($\rho = 0.928$ with catchment area), mean temperature ($\rho = 0.847$ with minimum temperature), and mean stream slope ($\rho = -0.725$ with drainage density) were removed. The retained continuous predictor set comprised catchment area, mean elevation, drainage density, annual rainfall, maximum temperature, and minimum temperature.

The categorical PAIC predictor was encoded as a pairwise binary distance variable (0 if two catchments share a PAIC group, 1 if not), following the standard gdm-package convention for categorical site-level factors. Continuous predictors were not pre-standardized in the user workspace because the gdm package internally rescales each predictor to the [0,1] range over the observed data prior to fitting I-splines (Manion et al., 2018). Three I-spline basis functions per predictor were used (the package default). Simpson's beta (βsim) was used as the response, computed via the betapart package (Baselga and Orme, 2012).

GDM was implemented using the gdm() function with geo = TRUE so that geographic distance is included as a predictor. Variable importance and per-predictor significance were assessed using gdm.varImp() with 500 matrix permutations. Variable importance is reported as the percent contribution to deviance explained, with permutation-based p-values (Table 2). To verify that the primary GDM results are robust to the listwise deletion of 12 catchments with incomplete predictor data, we re-ran the model on the full 32-catchment dataset with mean-imputed missing values. Overall, the sensitivity result is consistent with the primary analysis.

To verify that residual spatial structure does not bias inference beyond what is captured by geographic distance as an explicit predictor, we computed the absolute model residuals between predicted and observed βsim for each site pair and tested correlation against pairwise geographic distance with a Mantel test (Spearman, 9999 permutations) using the vegan package (Oksanen et al., 2022).

I-spline partial response curves were plotted using plot.gdm(). The maximum height of each I-spline indicates the total contribution of that predictor to compositional turnover; the slope at any point indicates the local rate of turnover along that predictor gradient.

Indicator Species Analysis and PF-only Robustness Testing

To identify species characteristic of each delineated biogeographic province, an indicator species analysis was performed (Dufrêne and Legendre, 1997) using the indicspecies package (De Cáceres and Legendre, 2009) implemented in R. The IndVal index combines specificity (the probability that a given catchment belongs to a particular biogeographic province, given the presence of the species) and fidelity (the probability of finding the species in catchments of that biogeographic province). IndVal ranges from 0 to 1, with higher values indicating stronger

species-province associations. Statistical significance was assessed via permutation tests with 999 iterations; associations with $p < 0.05$ were considered significant.

To evaluate the robustness of the indicator-species assemblages and the cluster-derived provincial scheme to the inclusion of diadromous and euryhaline taxa, we conducted a parallel analysis using only the 79 obligate primary-freshwater (PF) species. The PF-only βsim matrix was computed identically to the full-dataset matrix and used to generate a parallel WPGMA dendrogram. To formally assess congruence between the full-dataset and PF-only biogeographic signals, we computed (i) a Mantel test between the two βsim distance matrices using Spearman rank correlation with 9999 permutations (vegan package in R; Oksanen et al., 2022), and (ii) the Pearson cophenetic correlation between the two dendrograms (Sneath and Sokal, 1973). We treat the PF-only analysis as a co-equal robustness analysis rather than a sensitivity check: convergence between the two analyses strengthens confidence in the provincial scheme, while discordance highlights provinces whose distinctiveness depends on the inclusion of broader life-history groups.

Results

Delineation of Freshwater Fish Biogeographic Provinces

The agglomerative cluster analysis based on Simpson's beta (βsim) species turnover matrix resolved five biogeographic provinces for freshwater fish in the Philippines (Table 1; Figure 2): Northern Luzon, Central/Southern Luzon, Greater Visayas, Mindanao, and Palawan–Mindoro.

Palawan–Mindoro, comprising four catchments, shows the highest merge height in the WPGMA dendrogram (0.544 in βsim; Table 1, Figure 2), indicating that its constituent

catchments span the largest compositional range among all provinces. Mindanao, encompassing eight catchments in this study, shows the second-highest merge height (0.509), reflecting substantial internal heterogeneity across its diverse drainage systems. Greater Visayas (7 catchments; 0.451) and Central/Southern Luzon (9 catchments; 0.450) exhibited nearly identical merge heights, indicating comparable levels of internal compositional variation. Northern Luzon, comprising four catchments, shows the lowest merge height (0.409), reflecting the greatest internal compositional cohesion among its catchments.

Inventory Completeness

Per-province sample-coverage analysis (Table S5) showed variation in inventory completeness across the five provinces. Greater Visayas had the highest sample coverage ($\hat{C}$ = 0.940), with observed richness (152 species) reaching 92% of the Chao 2 asymptotic estimate (165 species). Central/Southern Luzon was also well-sampled ($\hat{C}$ = 0.926; 86% complete). Palawan–Mindoro and Northern Luzon showed intermediate coverage ($\hat{C}$ = 0.874 and 0.854, respectively; 86% and 80% complete). Mindanao had the lowest sample coverage ($\hat{C}$ = 0.838; 71% complete relative to the Chao 2 asymptotic estimate of 216 species), with the largest gap between observed and asymptotic richness across all provinces. These patterns inform the interpretation of indicator-species results below and are revisited in the Discussion and Limitations sections.

Species Richness and Endemism

Species richness (SR) and corrected weighted endemism (CWE) varied distinctly across the five biogeographic provinces (Figure 3, Table 1). Northern Luzon exhibited moderate diversity and endemism, with SR ranging from 32 to 55 species and CWE values between 0.127

and 0.230. Central/Southern Luzon showed the widest range of species richness among all provinces (24 to 105 species across its nine catchments) but relatively lower CWE values (0.095 to 0.191), indicating high diversity with comparatively few range-restricted species. Greater Visayas showed intermediate SR and CWE values (34–98 species; CWE 0.120–0.181 across its seven catchments). Mindanao showed moderate SR (25 to 75 species) but the widest CWE range of any province (0.088–0.388), reflecting the coexistence of areas of low endemism with hotspots of irreplaceable endemism. The Agus River/Lanao Lake catchment within Mindanao stands out with the highest CWE score in the dataset (0.388). Palawan–Mindoro exhibited the highest lower bounds for both SR (55–72 species) and CWE (0.150–0.279), indicating consistently moderate to high diversity and endemism across its four catchments.

Robustness of the Provincial Scheme to Life-History Composition

Cluster analysis restricted to the 79 obligate primary-freshwater species (shown as the lower dendrogram in the Figure S2) was moderately congruent with the full-dataset analysis. The Mantel test between the full-dataset and PF-only βsim distance matrices was significant (Mantel $r = 0.46$, $p = 0.0001$, Spearman, 9999 permutations), and the Pearson cophenetic correlation between the two dendrograms was 0.49 ($p < 0.001$). Four of the five provinces retained 75% or more of their catchments in the same cluster across the two analyses: Mindanao (88%, 7 of 8 catchments), Greater Visayas (86%, 6 of 7 catchments), Northern Luzon (75%, 3 of 4 catchments), and Palawan–Mindoro (75%, 3 of 4 catchments). Central/Southern Luzon showed the lowest retention (56%, 5 of 9 catchments), with the four remaining CSL catchments distributed across other clusters in the PF-only analysis. The moderate cross-dataset congruence indicates that diadromous and euryhaline species contribute biogeographic information that is

not fully recoverable from the obligate-freshwater fauna alone, while the provincial structure remains broadly recoverable from obligate freshwater taxa for the majority of provinces.

Generalized Dissimilarity Modelling

The Generalized Dissimilarity Model indicated that geographic distance was the dominant correlate of compositional turnover among Philippine freshwater fish catchments (Table 2, Figures S1). The model explained 36.39% of the deviance in pairwise compositional dissimilarity. Of the seven retained predictors, only geographic distance was statistically significant under matrix permutation, contributing 13.5% to deviance explained ($p = 0.028$). The I-spline for geographic distance showed a steep rise in compositional turnover over the first ~400 km, followed by a plateau, consistent with the limited dispersal of freshwater taxa across marine barriers in the archipelago. Catchment area contributed 11.0% ($p = 0.17$), maximum temperature 9.3% ($p = 0.23$), and mean elevation 8.1% ($p = 0.18$), but none were significant under permutation. PAIC grouping contributed only 2.8% ($p = 0.23$), indicating that once geographic distance was included in the model, PAIC-based historical structure added little additional explanatory power. A Mantel test of residuals against pairwise geographic distance returned $r = -0.040$ ($p = 0.706$), indicating no significant residual spatial autocorrelation and supporting the adequacy of geographic distance as an explicit model predictor. A sensitivity analysis using the imputed 32-catchment dataset returned a qualitatively consistent pattern, in which geographic distance as the dominant and only significant predictor (32.2%, $p < 0.001$), with PAIC again non-significant (3.1%, $p = 0.25$).

Per-predictor importance (Table 2) indicates that geographic and topographic predictors collectively accounted for the majority of deviance attributable to any single block (geographic distance 13.5% + catchment area 11.0% + mean elevation 8.1% + drainage density 0.6% =

33.2%), climatic predictors contributed modestly (maximum temperature 9.3% + minimum temperature and rainfall combined ~3.3% = ~12.6%), and PAIC contributed least (2.8%). A Mantel test of residuals against pairwise geographic distance returned $r = -0.040$ ($p = 0.706$), indicating no significant residual spatial autocorrelation and supporting the adequacy of geographic distance as an explicit model predictor.

Indicator Species Analysis

Indicator species analysis identified 60 species as significantly associated with a single province ($p \leq 0.05$, 999 permutations). Of these, 18 remained significant as indicators in the parallel PF-only analysis restricted to obligate primary-freshwater species (Table 3). Palawan–Mindoro returned the largest number of indicator species (24 significant indicators, of which 12 also significant in the PF-only analysis), reflecting the high level of endemism in the province and its biogeographic distinctiveness. In the sections that follow, we describe the indicator assemblage of each province in turn.

The full-dataset analysis identified 11 indicator species for Northern Luzon (Table 3). The strongest were the Anguillid eels *Anguilla japonica* (IndVal = 0.943, $p = 0.001$), *Anguilla luzonensis* (IndVal = 0.899, $p = 0.001$), and the mullets *Crenimugil crenilabis* (IndVal = 0.846, $p = 0.003$) and *C. heterocheilos* (IndVal = 0.760, $p = 0.004$). Additional indicators included *Mugil cephalus, Plicofollis magatensis, Lamnostoma taylori*, and other diadromous and euryhaline taxa. The dominance of diadromous and euryhaline species is consistent with the open-system ecology of Northern Luzon's short coastal drainages. The PF-only analysis returned only a single significant indicator for this province: *Rhinogobius philippinus* (IndVal = 0.737, $p = 0.019$). Although the cross-dataset cluster retention for Northern Luzon is moderately high (3 of 4 catchments; 75%), the sharp reduction in indicator counts from eleven to one and the province's

moderate sample coverage ($\hat{C}$ = 0.854) together suggest that upland endemic PF taxa are likely underrepresented in the current dataset for this province.

Central/Southern Luzon returned eight significant indicators. Two species tied for the highest IndVal: the endemic silver perch *Leiopotherapon plumbeus* (IndVal = 0.816, p = 0.001) and the estuarine catfish *Arius manillensis* (IndVal = 0.816, p = 0.002). Additional significant indicators included the *Zenarchopterus philippinus* (IndVal = 0.767, p = 0.006), *Cynoglossus puncticeps* (IndVal = 0.670), *Gobiopterus brachypterus* (IndVal = 0.667), *Mistichthys luzonensis* (IndVal = 0.667), *Psammogobius biocellatus* (IndVal = 0.595), and *Hippichthys heptagonus* (IndVal = 0.580). Of these eight, four are obligate primary-freshwater species (*Leiopotherapon plumbeus, Zenarchopterus philippinus, Gobiopterus brachypterus,* and *Mistichthys luzonensis*) and all four were also recovered as significant indicators in the PF-only analysis. Near-complete sample coverage ($\hat{C}$ = 0.926) supports the interpretation that this endemic-anchored PF signal reflects a genuine biogeographic feature of the province.

Greater Visayas returned 11 significant indicator species, all of which were euryhaline or diadromous. The strongest indicators were the snake eel *Ophichthus polyophthalmus* (IndVal = 0.894, p = 0.001), the estuarine goby *Glossogobius bicirrhosus* (IndVal = 0.845, p = 0.002), *Moringua raitaborua* (IndVal = 0.756), and *Glossogobius aureus* (IndVal = 0.693). Additional indicators included several other estuarine gobies, mullets, and snappers. Notably, no obligate primary-freshwater species emerged as significant indicators for this province: the PF-only analysis returned zero significant indicators for Greater Visayas. Combined with the province's high sample coverage ($\hat{C}$ = 0.940), this pattern indicates that the biogeographic signal of Greater Visayas is genuinely carried by estuarine and diadromous taxa rather than by obligate freshwater endemics.

The Mindanao indicator assemblage in the full-dataset analysis comprised six significant species: *Clarias macrocephalus* (IndVal = 0.739, p = 0.001), *Anguilla borneensis* (IndVal = 0.707, p = 0.005), *Anguilla bicolor* (IndVal = 0.640), *Butis gymnopomus* (IndVal = 0.624), *Belobranchus belobranchus* (IndVal = 0.584), and *Sillago sihama* (IndVal = 0.524). Only *Clarias macrocephalus* is an obligate primary-freshwater species, and it was the sole significant PF-only indicator for this province. The relatively small number of indicators (six species, one of which is PF) and the broad geographic ranges of many Mindanao species across multiple provinces are consistent with Mindanao's position as a biogeographic mixing zone.

Palawan–Mindoro returned the largest indicator assemblage of any province, with 24 significant species. The strongest indicator was the cyprinid *Rasbora argyrotaenia* (IndVal = 1.000, p = 0.002), followed by *Neostethus bicornis* (IndVal = 0.866, p = 0.003), *Nematabramis alestes* (IndVal = 0.802, p = 0.004), and *Mugilogobius mertoni* (IndVal = 0.794, p = 0.002). A large group of species — including multiple *Rasbora, Barbodes, Dermogenys*, and *Rhinogobius* endemics — returned identical IndVal values of 0.707 reflecting perfect fidelity to a small subset of catchments within the province. The PF-only analysis confirmed 12 of these 24 indicators as obligate primary-freshwater species: *Rasbora argyrotaenia, Nematabramis alestes, Barbodes hemictenus, Barbodes palavanensis, Rasbora everetti, Rasbora lateristriata, Rhinogobius estrellae, Rhinogobius tandikan, Dermogenys bispina, Dermogenys palawanensis, Dermogenys robertsi,* and *Pterocryptis taytayensis*. This large PF indicator assemblage strongly supports Palawan–Mindoro's distinctiveness as a Sundaic biogeographic province, dominated by obligate freshwater endemics of primary-freshwater families (Cyprinidae, Adrianichthyidae, Zenarchopteridae).

Discussion

Philippine Freshwater Fish Biogeographic Delineation

This study presents the first biogeographic analysis specifically focused on freshwater fishes in the Philippines, addressing a significant gap in our understanding of aquatic biodiversity patterns in the archipelago. Cluster analysis resolved five biogeographic provinces (Northern Luzon, Central/Southern Luzon, Greater Visayas, Mindanao, and Palawan–Mindoro) each with distinctive faunal assemblages and endemism patterns. These provinces refine previous biogeographic hypotheses for the Philippines (Brown et al., 2013), and the present results reflect not only historical PAIC influences but also contemporary environmental and ecological factors that shape freshwater fish distributions.

Northern Luzon Province

The Northern Luzon province occupies the oldest emergent Philippine landmass, with the Cordillera and Northern Sierra Madre ranges generating a mosaic of high-gradient streams, lowland river systems, and the extensive Cagayan River corridor (27,684 km^2), the largest in the Philippines (Heaney et al., 2016; Hall, 2002, Boothroyd et al., 2023). The long geological tenure of Luzon, dating from Oligocene volcanic arc accretion, has provided extended opportunity for *in situ* diversification and the persistence of range-restricted endemics such as the Philippine medaka or ricefish *Oryzias luzonensis*. The province's constituent catchments form a cohesive Pleistocene PAIC unit (Greater Luzon) that remained above sea level during glacial low-stands (Brown and Diesmos, 2002).

The Northern Luzon indicator assemblage in the full-dataset analysis is dominated by catadromous and euryhaline taxa, particularly Anguillidae and Mugilidae species. This composition reflects the open-system ecology of the region's short coastal drainages rather than

an absence of freshwater endemism. The PF-only analysis recovered only a single significant obligate freshwater indicator, *Rhinogobius philippinus*, a sharp reduction from the eleven indicators recovered by the full-dataset analysis. Although Northern Luzon retains a moderately high fraction of its catchments in the PF-only clustering (3 of 4; 75%), the diminished indicator assemblage under the PF-only analysis underscores that the province's biogeographic identity is carried primarily by taxa with broader marine-freshwater life histories rather than by an obligate primary-freshwater fauna. The sample-coverage analysis (Table S5) identifies Northern Luzon as among the less-completely sampled provinces, indicating that some range-restricted PF species likely remain undetected in the province's headwater systems.

Central/Southern Luzon Province

The Central/Southern Luzon province encompasses the most hydrologically diverse region of the Philippines, including the lowland floodplains of the Central Luzon Basin, the volcanic crater lakes of the Batangas and Bicol regions, and the Pampanga and Pasig-Marikina river systems. Laguna de Bay (~900 km²), the largest lake in the Philippines, supports a rich ichthyofauna reflecting its connectivity with both the Pasig River estuary and upland tributaries (Aquino et al., 2011; Vallejo, 1985). Lake Buhi in the Bicol region is the type locality of *Mistichthys luzonensis*, the world's smallest commercially harvested fish and one of the most ecologically significant single-catchment endemics in the archipelago.

The Central/Southern Luzon indicator assemblage includes both estuarine and obligate freshwater components. Four of the eight significant full-dataset indicators are obligate primary-freshwater species and all four were recovered as significant PF-only indicators as well: *Leiopotherapon plumbeus*, *Zenarchopterus philippinus, Gobiopterus brachypterus*, and *Mistichthys luzonensis*. The concordance between the full-dataset and PF-only IndVal results

supports a genuine freshwater-endemism signal for the province carried by these species. However, the cross-dataset cluster retention for Central/Southern Luzon is the lowest among the five provinces (56%), indicating that several Central/Southern Luzon catchments cluster with adjacent Greater Visayas catchments in the PF-only analysis. Two interpretations of this pattern are possible and are not mutually exclusive. First, several Central/Southern Luzon catchments share obligate-freshwater species with adjacent Greater Visayas catchments, so the boundary between the two provinces is less sharp for the PF fauna than for the full assemblage. Second, the strong PF-indicator signal for Central/Southern Luzon is carried by a small number of highly diagnostic endemic species rather than by broad, province-wide PF assemblage similarity. A good sample coverage ($\hat{C} = 0.926$) argues against undersampling as an explanation. The province's biogeographic distinctiveness therefore rests on the fidelity of a small set of characteristic endemic species rather than on the coherence of the broader PF assemblage.

Greater Visayas Province

The Greater Visayas province encompasses a highly fragmented archipelago of islands separated by shallow seas and narrow straits (Cebu, Negros, Panay, Leyte, Samar, and smaller islands). This configuration produces predominantly short, steeply graded coastal rivers; few exceed 50 km in length, and total freshwater habitat per unit land area is substantially less than in the larger island provinces. The Visayan islands originated as oceanic volcanic arcs and never formed part of the Sunda continental shelf (Hall, 2002), although Pleistocene low-stands united several adjacent islands into PAIC units (Greater Negros-Panay, Greater Leyte-Samar) that may have permitted limited within-province faunal exchange (Brown and Diesmos, 2002).

The Greater Visayas indicator assemblage in the full-dataset analysis comprised 11 indicator species, all euryhaline or diadromous taxa; the PF-only analysis returned zero

significant indicators. Crucially, the per-province sample-coverage analysis shows that Greater Visayas has very good sample coverage and highest of any province in the dataset, and 86% of its catchments retain their cluster assignment in the PF-only analysis. The absence of significant obligate-freshwater indicators is therefore not an artefact of undersampling; it reflects a real biological feature of the province. Greater Visayas is best understood as a hydrologically open, estuarine-dominated system whose faunal identity is shaped by marine connectivity rather than prolonged freshwater isolation. This is consistent with island biogeography theory (MacArthur and Wilson, 1967): the smaller Visayan islands lack the catchment size and drainage continuity required to sustain large assemblages of obligate freshwater species. The province remains a meaningful biogeographic unit within the clustering framework, and integrated coastal-freshwater management is appropriate for protecting its distinctive estuarine communities.

A particularly noteworthy result of the cluster analysis is the placement of Leyte and Samar within Greater Visayas, despite both islands being formally assigned to the Greater Mindanao PAIC under the Brown and Diesmos (2002) classification. This represents a meaningful discordance between the historical PAIC framework and the faunal clustering recovered by the analysis. Under the PAIC framework, Leyte and Samar are grouped with Greater Mindanao because Pleistocene sea-level low-stands exposed shelves between these islands and the Mindanao landmass; however, the Surigao Strait reaches depths exceeding 300 m at its narrowest points and would have remained an open marine passage even at glacial maximum, constituting a persistent dispersal barrier for obligate freshwater taxa (Voris, 2000; Hall, 2002). By contrast, the much shallower Camotes Sea (~80–100 m maximum depth) between Leyte–Samar and the central Visayan islands would have been more extensively bridged during glacial low-stands. The clustering pattern is therefore consistent with the central Visayan

islands, rather than Mindanao, having served as the dominant source of freshwater colonists for Leyte and Samar.

This discordance has two implications. First, PAIC membership is a necessary but not sufficient condition for faunal exchange: two islands can be assigned to the same PAIC while remaining effectively isolated if the connecting marine passage was never shallow enough to permit land-bridge formation. Second, freshwater fish beta dissimilarity is a more sensitive recorder of realized faunal exchange than the binary PAIC classification. The Leyte–Samar pattern has precedent in terrestrial biogeographic analyses showing similar Eastern Visayas affinities (Heaney et al., 1998; Brown et al., 2013; Siler et al., 2012), and reinforces calls to treat PAIC groupings as working hypotheses to be tested against distributional data.

Mindanao Province

The Mindanao province encompasses the second-largest island in the Philippines, with the Rio Grande de Mindanao draining the Cotabato Basin, and major lake basins including Lake Mainit in the northeast and Lake Lanao in Lanao del Sur. Mindanao's position near the Celebes Sea and the Wallacea transition zone places it at the interface of multiple biogeographic realms; its freshwater fauna accordingly blends Sundaic, Wallacean, and oceanic Philippine elements more than any other Philippine province (Kottelat et al., 1993; Hall, 2002). The province may have received Sundaic faunal elements via the Sulu Archipelago, a chain of islands extending between Zamboanga and Sabah that is included with Mindanao in the same freshwater ecoregion and may have functioned as a Pleistocene land bridge (Voris, 2000).

Lake Lanao is one of the world's ancient lakes and a global hotspot of lacustrine endemism (Myers, 1960; Cristescu et al., 2010). Herre (1924) described 18 endemic cyprinid

species from the lake, constituting a textbook adaptive radiation; 15 of these are now considered extinct or extinct in the wild, representing one of the most severe documented freshwater extinction events in Southeast Asia (Ismail et al., 2014; IUCN, 2020). The Agus River/Lake Lanao catchment carries the highest corrected weighted endemism of any catchment in the dataset, reflecting the evolutionary significance of this system even after the recent extinctions.

The Mindanao indicator assemblage in the full-dataset analysis comprised six significant species: the walking catfish *Clarias macrocephalus* and the eels *Anguilla borneensis* and *A. bicolor* were the strongest indicators, alongside *Butis gymnopomus, Belobranchus belobranchus*, and *Sillago sihama*. Only *Clarias macrocephalus* is an obligate primary-freshwater species, and it was the sole PF-only indicator retained for the province. Although Mindanao shows high cross-dataset cluster retention in the PF-only analysis (88%), the small number of obligate-PF indicators and the broad geographic ranges of many Mindanao species across multiple provinces are consistent with Mindanao's position as a mixing zone for biogeographic affinities. The presence of two Anguilla species and *Belobranchus belobranchus* among the indicators reflects Mindanao's position as a receiving area for diadromous taxa from both Philippine and Bornean sources. The sample-coverage analysis identifies Mindanao as the least-completely sampled province in the dataset, with the largest gap between observed and asymptotic richness. This indicates that some range-restricted Mindanao species likely remain undetected, particularly in the interior catchments, and we identify targeted surveys of these areas as a priority for future work.

Palawan–Mindoro Province

The Palawan–Mindoro province is geologically the most distinctive of the five biogeographic provinces, having originated as a fragment of the Asian continental shelf that

rifted southward from the southern China margin during the mid-Oligocene (Hall, 2002; Carpenter and Springer, 2005). This continental origin fundamentally distinguishes Palawan and Mindoro from the volcanic oceanic-arc origins of the other Philippine islands. During Pleistocene glacial maxima, the Balabac Strait between Palawan and Borneo narrowed substantially and possibly formed a land bridge, facilitating faunal exchange with the Sunda Shelf (Heaney, 1985; Rohling et al., 1998). Palawan is accordingly recognized as the north-eastern boundary of the Sundaic biogeographic region (Esselstyn et al., 2004), a distinction supported by mammals, reptiles, amphibians (Blackburn et al., 2010; Siler et al., 2012; Brown et al., 2016), and the present freshwater fish data.

The obligate-freshwater indicator assemblage for Palawan–Mindoro is now anchored by Cyprinidae (*Barbodes* spp., *Rasbora* spp., *Nematabramis*) and *Dermogenys palawanensis*. These PF indicators are characteristic of Sunda Shelf drainages (Kottelat et al., 1993; Kottelat and Widjanarti, 2005), supporting the interpretation of Palawan–Mindoro as a Sundaic outlier within the Philippine archipelago. The relatively high number of endemic indicators indicates that this province might also serve as a site of *in situ* speciation, likely due to its isolation and unique environmental conditions.

Patterns of Species Richness and Endemism

The variation in species richness across biogeographic provinces, with Central/Southern Luzon hosting the highest diversity, challenges the traditional expectation of higher species richness on larger islands like Mindanao. This pattern may reflect several factors: the complex topography and diverse river systems of Central/Southern Luzon provide a wide array of habitat niches (Rahel, 2007); the region may have served as a refugium during past climatic fluctuations (Woodruff, 2010); and differential sampling effort across regions is a contributing factor

(Lomolino, 2004), with the per-province sample-coverage analysis showing that Mindanao's lower observed richness partly reflects its lower inventory completeness.

The high endemism observed in Palawan–Mindoro, as indicated by CWE scores, underscores the evolutionary uniqueness of this region and supports its recognition as a distinct biogeographic unit, often referred to as the Palawan Micro-continent (Blackburn, 2015). The exceptional endemism in the Agus River/Lanao Lake catchment within Mindanao highlights the importance of ancient lakes as cradles of biodiversity, paralleling patterns in other ancient lake systems worldwide (Cristescu et al., 2010).

The IndVal analysis and CWE metric capture complementary, non-overlapping aspects of endemism. IndVal detects species that are both specific to a province and consistently present across multiple catchments within it — rewarding broad fidelity within a group — whereas highly range-restricted point endemics may score perfectly on specificity but fail to reach significance because their fidelity is inherently low when they occupy only one or two catchments. For example, *Oryzias luzonensis*, endemic to a single Northern Luzon catchment, has perfect specificity (A = 1.0) but a fidelity of only 0.25, yielding a non-significant IndVal value. The endemic cyprinids of Lake Lanao, confined to a single catchment within a ten-catchment province, have a maximum possible IndVal value regardless of their degree of endemism. CWE, designed to weight species by the inverse of their range size, is precisely the metric that reveals Lanao Lake (site 26) as the highest endemism hotspot in the dataset. The two analyses serve distinct biogeographic purposes: IndVal characterizes the broad faunal identity of each province; CWE identifies the catchments most critical for conserving irreplaceable, range-restricted diversity.

Comparative Biogeographic Context in Southeast Asia

The five Philippine freshwater fish provinces sit east of Wallace's Line within the Wallacean biogeographic transition zone, but the heterogeneous origins of the islands produce marked intra-archipelago variation in biogeographic affinity. Palawan–Mindoro is the clearest Sundaic outlier: its obligate primary-freshwater indicators are characteristic of Sunda Shelf drainages (Kottelat et al., 1993; Kottelat, 2013), and its assembly is best explained by Pleistocene land-bridge connectivity with Borneo across a narrowed Balabac Strait rather than by dispersal from the rest of the Philippines. This mechanism differs from internal Sundaland connectivity events involving drainage capture across the lowland Sunda Shelf (Voris, 2000; Woodruff, 2010), and helps explain why Palawan–Mindoro harbors Sundaic primary freshwater genera but lacks the full richness of Bornean assemblages (Tedesco et al., 2017).

The other three provinces — Northern Luzon, Central/Southern Luzon, and Greater Visayas — occupy the oceanic Philippines and lack continental Sundaic primary freshwater genera. Their faunas are composed of diadromous, catadromous, and euryhaline taxa capable of crossing marine barriers, together with in-situ island endemics derived from such colonists. This is the standard assembly mode for oceanic island freshwater faunas (McDowall, 2004; Lohman et al., 2011) and contrasts with continental Southeast Asian River basins (Mekong, Chao Phraya) where climate and drainage network structure dominate freshwater fish beta diversity (Tedesco et al., 2017). The Philippine GDM result reflects the overriding importance of marine dispersal barriers and glacial land-bridge dynamics in structuring island freshwater faunas (Lohman et al., 2011; Setiawan et al., 2020).

Mindanao occupies an intermediate position. Its proximity to Sulawesi and the Moluccas means it receives biogeographic influence from both the Wallacean and oceanic Philippine realms; several of its eel indicator species (*Anguilla bicolor*, *A. borneensis*, *A. celebesensis*) have

Indo-Pacific rather than Philippine endemic distributions, consistent with a dispersal-assembly model (Aoyama, 2009). The blending of Sundaic, Wallacean, and oceanic-Philippine elements in Mindanao's fauna parallels mixed affinities reported for Sulawesi freshwater fishes (Moss and Wilson, 1998; Kottelat, 2013), and suggests that Mindanao functions as an eastern mixing zone in Philippine freshwater biogeography, analogous to Sulawesi's role in the broader Wallacean transition.

Environmental Drivers of Fish Distribution

The GDM results indicate that the principal correlate of compositional turnover among Philippine catchments is geographic distance, with PAIC grouping and contemporary climate variables contributing comparatively little once geographic distance is in the model. In both the primary and sensitivity catchment analyses, geographic distance was the only statistically significant predictor. The dominance of distance-decay is consistent with the strong dispersal limitation expected for freshwater taxa in an archipelagic setting, where marine straits are persistent barriers (Setiawan et al., 2020; Lohman et al., 2011).

The limited contribution of climate variables in the GDM is consistent with the relative climatic homogeneity of the Philippine archipelago, where most catchments lie within the same humid tropical zone. This contrasts with continental tropical systems, where temperature, precipitation, and seasonality are often important drivers of freshwater fish beta diversity. Crucially, the weak contribution of PAIC grouping in the GDM does not mean PAIC history is irrelevant to Philippine freshwater biogeography. Geographic distance and PAIC are strongly confounded by construction (islands within the same PAIC are by definition close together), and the GDM cannot easily separate them. The cluster analysis still resolves the five biogeographic provinces in a pattern broadly congruent with PAIC structure (with the notable exception of

Leyte–Samar discussed above), and the indicator-species results identify obligate primary-freshwater taxa of Bornean affinity that anchor the Palawan–Mindoro province. The PAIC framework remains useful as a conceptual organizer for the Philippine freshwater fauna; what the GDM shows is that its explanatory power is largely captured by geographic distance once that variable is in the model.

Conservation Implications

These findings have significant implications for freshwater fish conservation in the Philippines. The delineation of five biogeographic provinces provides a biologically grounded framework for identifying conservation priorities and aligning management effort with the spatial structure of freshwater biodiversity. We recommend that each province be considered as a distinct Freshwater Conservation Unit within the national biodiversity planning framework, consistent with the approach used for terrestrial taxa under the Philippine Biodiversity Strategy and Action Plan (BMB-DENR, 2015).

Palawan–Mindoro is the highest conservation priority in terms of obligate freshwater endemism. Its indicator assemblage is dominated by obligate primary-freshwater Cyprinidae with no occurrence outside the Philippines, and its CWE scores are the highest of any Philippine province outside the Lanao catchment. Priority actions: (i) enforcement of the Strategic Environmental Plan for Palawan (Republic Act 7611); (ii) expansion of Key Biodiversity Area coverage to underprotected lowland river systems in southern Palawan and northwestern Mindoro; and (iii) regulation of non-native fish introductions, the primary documented threat to obligate primary-freshwater taxa on the island.

The Agus River/Lake Lanao catchment within Mindanao is the single most critical site for irreplaceable endemism in the entire dataset. Priority actions: (i) urgent eradication or suppression of invasive species through coordinated removal programs; (ii) rapid population surveys to determine whether endemic cyprinid populations persist in remnant habitats or tributary headwaters; (iii) captive assurance colonies for any surviving endemic taxa; and (iv) engagement with local government units to incorporate freshwater biodiversity objectives into watershed management.

Greater Visayas represents a significant conservation gap in the current protected-area network. Despite harboring a distinct estuarine-freshwater interface fauna, its river systems are among the least represented in the BMB-DENR framework. Given that the province's indicator assemblage reflects the open-system ecology of short coastal drainages — sensitive to land-use change, sedimentation, and mangrove loss — priority actions include freshwater-inclusive assessments for major Visayan river systems and integration of riverine habitat protection into coastal resource management plans.

Across all provinces, the dominance of geographic distance as the correlate of compositional turnover, together with the role of within-PAIC connectivity in maintaining the broader biogeographic provinces, underscores the importance of within-province river connectivity. Fragmentation of river systems by dams, road development, and irrigation infrastructure reduces the effective catchment area available to migratory, catadromous, and diadromous species, and can accelerate the differentiation and local extinction of isolated populations. Connectivity conservation should be integrated into environmental impact assessments for infrastructure projects across all five biogeographic provinces.

Limitations and Future Directions

The biogeographic patterns reported here are based entirely on taxonomic occurrence data, and three limitations warrant explicit acknowledgement. First, compositional similarity is not equivalent to phylogenetic relatedness. Simpson's beta dissimilarity reflects shared species lists rather than shared evolutionary history; two catchments can share many species and yet contain populations of independent colonization origin, while two catchments with few shared species can nevertheless harbor closely related sister-species pairs. Interpretations involving evolutionary history, dispersal, Sundaic affinity, and colonization in this manuscript should be read as hypotheses generated by the compositional patterns rather than as conclusions about phylogenetic relatedness or historical population connectivity. Direct tests require comparative phylogenetic and phylogeographic data, which are not yet available for most Philippine freshwater fish lineages.

Second, inventory completeness varies across provinces. The per-province sample-coverage analysis identified Mindanao and Northern Luzon as the least-completely sampled provinces, with the largest gaps between observed and asymptotic species richness. Some range-restricted species in remote Mindanao catchments and Northern Luzon headwater systems likely remain undetected. Greater Visayas, Central/Southern Luzon, and Palawan–Mindoro showed higher coverage in the analysis. Targeted surveys of Mindanao interior catchments and Northern Luzon headwaters are therefore identified as priorities for future work.

Third, taxonomic and life-history assignments in Philippine freshwater fishes remain incomplete. The data entries here were all cross-referenced against recent revisionary literature and addressed numerous specific cases but does not constitute a comprehensive systematic revision. Several species are flagged as *inquirendae*, and ecological-category assignments for many taxa rest on inference from congeners rather than direct life-history observation in

Philippine populations. Three research priorities follow from these limitations. First, focused taxonomic revisions of Philippine *Eleotridae, Gobiidae*, and *Atherinomorpha* are needed, where unresolved synonymies and likely cryptic diversity continue to obscure assemblage-level patterns. Second, detailed life-history studies of Philippine populations are needed to test migration-type assignments. Third, expanded survey effort is required in undersampled provinces. Together these three lines of work would substantially refine, and where appropriate revise, the provincial framework presented here.

Beyond these immediate priorities, integrating functional trait data could illuminate the ecological roles and ecosystem services provided by different fish assemblages; integrating phylogenetic and population genetic data is essential for elucidating evolutionary history and connectivity; and incorporating data on anthropogenic impacts (land-use change, pollution, overexploitation) is needed to evaluate vulnerabilities and inform management. Comparative analyses combining freshwater fish data with other taxonomic groups would further enrich our understanding of Philippine biogeography by revealing shared patterns or unique aspects of freshwater ecosystems.

Conclusion

This study provides the first comprehensive biogeographic analysis of freshwater fishes in the Philippine archipelago, resolving five biogeographic provinces whose structure reflects both Pleistocene Island connectivity and contemporary spatial and ecological processes. By integrating Simpson's beta-based clustering with sample-coverage assessment, primary-freshwater-only robustness testing, and Generalized Dissimilarity Modelling, the analysis distinguishes provinces whose distinctiveness rests on obligate freshwater endemism (Palawan–Mindoro, Northern Luzon and Central/Southern Luzon) from those characterized by estuarine–

freshwater interface assemblages (Greater Visayas) and from those positioned as biogeographic mixing zones (Mindanao). The unique perspective offered by freshwater fishes underscores the importance of taxon-specific analyses in understanding archipelagic biodiversity, and provides a foundation for biogeographically informed conservation of Philippine freshwater ecosystems under increasing anthropogenic pressure.

Author Contributions

Brian Wade Jamandre: conceptualization, data curation and validation, formal analysis, writing – original draft, writing – review and editing. **Nico Jose Leander**: data validation, writing – review and editing. **Rodulf Anthony Balisco**: data validation, writing - review and editing. **Faith Santiago-Tadeo**: writing – review and editing.

Data availability

The data supporting the findings of this study are provided in the supplementary information.

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., Higgins, J. V., Heibel, T. J., Wikramanayake, E., Olson, D., Lopez, H. L., Reis, R. E., Lundberg, J. G., Sabaj Pérez, M. H., & Petry, P. (2008). Freshwater ecoregions of the world: A new map of biogeographic units for freshwater biodiversity conservation. *BioScience, 58*(5), 403–414. https://doi.org/10.1641/B580507

Amoroso, V. B. (2012). Plant Diversity in Two Philippine Long-Term Ecological Research Sites. In *The Biodiversity Observation Network in the Asia-Pacific Region: Toward Further Development of Monitoring* (pp. 39-51). Tokyo: Springer Japan.

Aoyama, J. (2009). Life history and evolution of migration in catadromous eels (Anguilla spp.). *Aqua-BioScience Monographs, 2*(1), 1–42.

Aquino, L. M. G., Tango, J. M., Canoy, R. J. C., Fontanilla, I. K. C., Basiao, Z. U., Ong, P. S., & Quilang, J. P. (2011). DNA barcoding of fishes of Laguna de Bay, Philippines. *Mitochondrial DNA*, *22*(4), 143-153.

Aurelio, M. A., Peña, R. E., & Taguibao, K. J. L. (2013). Sculpting the Philippine archipelago since the Cretaceous through rifting, oceanic spreading, subduction, obduction, collision and strike-slip faulting: Contribution to IGMA5000. *Journal of Asian Earth Sciences, 72*, 102–107.

Baldia, S. F., Kabiling, M. T. C., Gabriel, C. A. L., Abeleda, A. N. D., Aguinaldo, R. C. A., Clemente, K. J. E., & Cordero, P. A., Jr. (2017). Marine macroalgal diversity assessment of Biri Island and Dalupirit Island, Northern Samar, Philippines. *Acta Manilana, 65*, 29–38.

Balisco, R. A. T., & Liao, T.-Y. (2025). An annotated checklist of fish species described from the Philippines. *ZooKeys, 1246*, 1–287. https://doi.org/10.3897/zookeys.1246.145752

Baselga, A. (2010). Partitioning the turnover and nestedness components of beta diversity. *Global Ecology and Biogeography*, 19(1), 134–143.

Baselga, A., & Orme, C. D. L. (2012). betapart: An R package for the study of beta diversity. *Methods in Ecology and Evolution*, 3(5), 808–812.

Berra, T. M. (2007). Freshwater Fish Distribution. University of Chicago Press.

Blackburn, D. C., Bickford, D. P., Diesmos, A. C., Iskandar, D. T., & Brown, R. M. (2010). An ancient origin for the enigmatic flat-headed frogs (Bombinatoridae: Barbourula) from the islands of Southeast Asia. *PLoS ONE, 5*(8), e12090. https://doi.org/10.1371/journal.pone.0012090

Blackburn, D. C., Siler, C. D., Diesmos, A. C., McGuire, J. A., Cannatella, D. C., & Brown, R. M. (2015). An adaptive radiation of frogs in a Southeast Asian island archipelago. *Evolution, 69*(6), 139–1397. https://doi.org/10.1111/evo.12660

BMB-DENR (Biodiversity Management Bureau, Department of Environment and Natural Resources). (2015). Philippine Biodiversity Strategy and Action Plan 2015–2028. Biodiversity Management Bureau, DENR, Quezon City, Philippines.

Boothroyd, R. J., Williams, R. D., Hoey, T. B., MacDonell, C., Tolentino, P. L. M., Quick, L., Guardian, E. L., Reyes, J. C. M. O., Sabillo, C. J., Perez, J. E. G., & David, C. P. C. (2023). National-scale geodatabase of catchment characteristics in the Philippines for river management applications. *PLOS ONE, 18*(3), e0281933. https://doi.org/10.1371/journal.pone.0281933

Brown, R. M., & Diesmos, A. C. (2002). Application of lineage-based species concepts to oceanic island frog populations: The effects of differing taxonomic philosophies on the estimation of Philippine biodiversity. *Silliman Journal, 42*(1), 133–162.

Brown, R. M., Siler, C. D., Oliveros, C. H., Esselstyn, J. A., Diesmos, A. C., Hosner, P. A., Linkem, C. W., Barley, A. J., Oaks, J. R., Sanguila, M. B., Welton, L. J., Blackburn, D. C., Moyle, R. G., Peterson, A. T., & Alcala, A. C. (2013). Evolutionary processes of diversification in a model island archipelago. *Annual Review of Ecology, Evolution, and Systematics, 44*, 411–435. https://doi.org/10.1146/annurev-ecolsys-110411-160323

Brown, R. M., Su, Y.-C., Barger, B., Siler, C. D., Sanguila, M. B., Diesmos, A. C., & Blackburn, D. C. (2016). Phylogeny of the island archipelago frog genus Sanguirana: Another endemic Philippine radiation that diversified "Out-of-Palawan." *Molecular Phylogenetics and Evolution, 94*, 531–536. https://doi.org/10.1016/j.ympev.2015.10.010

Carpenter, K. E., & Springer, V. G. (2005). The center of the center of marine shore fish biodiversity: The Philippine Islands. *Environmental Biology of Fishes, 72*(4), 467–480. https://doi.org/10.1007/s10641-004-3154-4

Chao, A., & Jost, L. (2012). Coverage-based rarefaction and extrapolation: Standardizing samples by completeness rather than size. *Ecology*, 93(12), 2533–2547.

Chao, A., Gotelli, N. J., Hsieh, T. C., Sander, E. L., Ma, K. H., Colwell, R. K., & Ellison, A. M. (2014). Rarefaction and extrapolation with Hill numbers: A framework for sampling and estimation in species diversity studies. *Ecological Monographs*, 84(1), 45–67.

Conlu, P. V. (1986). Guide to Philippine Flora and Fauna, Volume IX: Fishes. Natural Resources Management Center, Ministry of Natural Resources and University of the Philippines.

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*(22), 4837–4851. https://doi.org/10.1111/j.1365-294X.2010.04832.x

De Cáceres, M., & Legendre, P. (2009). Associations between species and groups of sites: Indices and statistical inference. *Ecology, 90*(12), 3566–3574. https://doi.org/10.1890/08-1823.1

Dudgeon, D., Arthington, A. H., Gessner, M. O., Kawabata, Z.-I., Knowler, D. J., Lévêque, C., Naiman, R. J., Prieur-Richard, A.-H., Soto, D., Stiassny, M. L. J., & Sullivan, C. A. (2006). Freshwater biodiversity: Importance, threats, status and conservation challenges. *Biological Reviews, 81*(2), 163–182. https://doi.org/10.1017/S1464793105006950

Dufrêne, M., & Legendre, P. (1997). Species assemblages and indicator species: The need for a flexible asymmetrical approach. *Ecological Monographs, 67*(3), 345–366. https://doi.org/10.1890/0012-9615(1997)067[0345:SAAIST]2.0.CO;2

Esselstyn, J. A., & Brown, R. M. (2009). The role of repeated sea-level fluctuations in the generation of shrew (Soricidae: Crocidura) diversity in the Philippine Archipelago. *Molecular Phylogenetics and Evolution, 53*(1), 171–181. https://doi.org/10.1016/j.ympev.2009.05.034

Esselstyn, J. A., Widmann, P., & Heaney, L. R. (2004). The mammals of Palawan Island, Philippines. *Proceedings of the Biological Society of Washington, 117*(3), 271–302.

Ferrier, S., Manion, G., Elith, J., & Richardson, K. (2007). Using generalized dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment. *Diversity and Distributions, 13*(3), 252–264. https://doi.org/10.1111/j.1472-4642.2007.00341.x

FishNet 2. (2024). *FishNet 2 portal*. Retrieved November 25, 2024, from http://www.fishnet2.net

Fricke, R., Eschmeyer, W. N., & Van der Laan, R. (Eds.). (2024). Eschmeyer's Catalog of Fishes: Genera, Species, References. California Academy of Sciences. Retrieved June 06, 2026, from http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp

Froese, R., & Pauly, D. (Eds.). (2024). FishBase. World Wide Web electronic publication. Retrieved June 06, 2026, from http://www.fishbase.org

GBIF.org. (2024). *GBIF home page*. Global Biodiversity Information Facility. Retrieved November 25, 2024, from https://www.gbif.org

Hall, R. (2002). Cenozoic geological and plate tectonic evolution of SE Asia and the SW Pacific: Computer-based reconstructions, model and animations. *Journal of Asian Earth Sciences, 20*(4), 353–431. https://doi.org/10.1016/S1367-9120(01)00069-4

He, F., & Legendre, P. (2002). Species diversity patterns derived from species-area models. *Ecology, 83*(5), 1185–1198. https://doi.org/10.1890/0012-9658(2002)083[1185:SDPDFS]2.0.CO;2

Heaney, L. R. (1985). Zoogeographic evidence for middle and upper Pleistocene land bridges to the Philippine islands. In *Indo-Pacific Prehistory Association Philippines. Congress. 12* (pp. 127-143).

Heaney, L. R. (1986). Biogeography of mammals in Southeast Asia: Estimates of rates of colonization, extinction and speciation. *Biological Journal of the Linnean Society, 28*(1–2), 127–165. https://doi.org/10.1111/j.1095-8312.1986.tb01752.x

Heaney, L. R., Balete, D. S., Dolar, M. L., Alcala, A. C., Dans, A. T. L., Gonzales, P. C., Ingle, N. R., Lepiten, M. V., Oliver, W. L. R., Ong, P. S., Rickart, E. A., Tabaranza, B. R., Jr., & Utzurrum, R. C. B. (1998). A synopsis of the mammalian fauna of the Philippine Islands. *Fieldiana Zoology. New Series, 88*, 1–61.

Heaney, L. R., Balete, D. S., & Rickart, E. A. (2016). The mammals of luzon island: biogeography and natural history of a philippine fauna. Johns Hopkins University Press.

Heino, J., Virkkala, R., & Toivonen, H. (2009). Climate change and freshwater biodiversity: Detected patterns, future trends and adaptations in northern regions. *Biological Reviews, 84*(1), 39–54. https://doi.org/10.1111/j.1469-185X.2008.00060.x

Herre, A. W. C. T. (1924). Distribution of the true freshwater fishes in the Philippines. I. The Philippine Cyprinidae. *Philippine Journal of Science*, 24, 249–307.

Herre, A. W. C. T. (1953). Check List of Philippine Fishes (Research Report 20). U.S. Fish and Wildlife Service.

Holt, B. G., Lessard, J.-P., Borregaard, M. K., Fritz, S. A., Araújo, M. B., Dimitrov, D., Fabre, P.-H., Graham, C. H., Graves, G. R., Jønsson, K. A., Nogués-Bravo, D., Wang, Z., Whittaker, R. J., Fjeldså, J., & Rahbek, C. (2013). An update of Wallace's zoogeographic regions of the world. *Science, 339*(6115), 74–78. https://doi.org/10.1126/science.1228282

Hsieh, T. C., Ma, K. H., & Chao, A. (2016). iNEXT: An R package for rarefaction and extrapolation of species diversity (Hill numbers). *Methods in Ecology and Evolution*, 7(12), 1451–1456.

IUCN (International Union for Conservation of Nature). (2020). *The IUCN Red List of Threatened Species* (Version 2020-3). https://www.iucnredlist.org

Ismail, G. B., Sampson, D. B., & Noakes, D. L. G. (2014). The status of Lake Lanao endemic cyprinids (Puntius species) and their conservation. *Environmental Biology of Fishes, 97*(4), 425–434. https://doi.org/10.1007/s10641-013-0155-3

Jamandre, B. W. (2023). Freshwater fishes of the Philippines: A provisional checklist. *Zootaxa, 5301*(2), 151–181. https://doi.org/10.11646/zootaxa.5301.2.1

Junk, W. J., Bayley, P. B., & Sparks, R. E. (1989). The flood pulse concept in river–floodplain systems. In D. P. Dodge (Ed.), *Proceedings of the International Large River Symposium* (pp. 110–127). *Canadian Special Publication of Fisheries and Aquatic Sciences*, 106.

Keith, P., & Mennesson, M. I. (2023). Revision of Hypseleotris (Teleostei: Eleotridae) from Indo-Pacific Islands using molecular and morphometric approaches, with description of one new species. *Zoological Journal of the Linnean Society*, *198*(4), 1035-1069.

Kottelat, M. (2013). The fishes of the inland waters of Southeast Asia: A catalogue and core bibliography of the fishes known to occur in freshwaters, mangroves and estuaries. *Raffles Bulletin of Zoology, Supplement 27*, 1–663.

Kottelat, M., Whitten, A. J., Kartikasari, S. N., & Wirjoatmodjo, S. (1993). *Freshwater fishes of Western Indonesia and Sulawesi*. Periplus Editions.

Kottelat, M., & Widjanarti, E. (2005). The fishes of Danau Sentarum National Park and the Kapuas Lakes area, Kalimantan Barat, Indonesia. *Raffles Bulletin of Zoology, Supplement 13*, 139–173.

Laffan, S. W., Lubarsky, E., & Rosauer, D. F. (2010). Biodiverse, a tool for the spatial analysis of biological and related diversity. *Ecography, 33*(4), 643–647. https://doi.org/10.1111/j.1600-0587.2010.06237.x

Leprieur, F., Beauchard, O., Blanchet, S., Oberdorff, T., & Brosse, S. (2008). Fish invasions in the world's river systems: When natural processes are blurred by human activities. *PLoS Biology, 6*(2), e28. https://doi.org/10.1371/journal.pbio.0060028

Leprieur, F., Beauchard, O., Hugueny, B., Grenouillet, G., & Brosse, S. (2011). Partitioning global patterns of freshwater fish beta diversity reveals contrasting signatures of past climate changes. *PLOS ONE, 6*(3), e19570. https://doi.org/10.1371/journal.pone.0019570

Linke, S., Lehner, B., Ouellet Dallaire, C., Ariwi, J., Grill, G., Anand, M., Beames, P., Burchard-Levine, V., Maxwell, S., Moidu, H., Tan, F., & Thieme, M. (2019). Global hydro-environmental sub-basin and river reach characteristics at high spatial resolution. *Scientific Data, 6*, 283. https://doi.org/10.1038/s41597-019-0300-6

Lohman, D. J., 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 of Ecology, Evolution, and Systematics, 42*, 205–226. https://doi.org/10.1146/annurev-ecolsys-102710-145001

Lomolino, M. V. (2004). Conservation biogeography. In M. V. Lomolino & L. R. Heaney (Eds.), *Frontiers of biogeography: New directions in the geography of nature* (pp. 293–296). Sinauer Associates.

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

Manion, G., Lisk, M., Ferrier, S., Nieto-Lugilde, D., Mokany, K., & Fitzpatrick, M. C. (2018). gdm: Generalized Dissimilarity Modeling. Version 1.6.0-7 [R package]. https://CRAN.R-project.org/package=gdm

McDowall, R. M. (2004). Ancestry and amphidromy in island freshwater fish faunas. *Fish and Fisheries, 5*(1), 75–85. https://doi.org/10.1111/j.1467-2979.2004.00138.x

Mediodia, D., Castro, A., Tablizo, M., Policarpio, D., Calvelo, J., Baguio, J. P., Borja II, A., Lin, C.-H., & Fernando, A. G. (2025). Paleoichthyology in the Philippines: A review of Cenozoic fish fossils with insights on its current status and future opportunities. *Geobios, 88*, 163–174.

MGB (Mines and Geosciences Bureau). (2010). *Geology of the Philippines* (2nd ed.; M. A. Aurelio & R. E. Peña, Eds.). Mines and Geosciences Bureau, Department of Environment and Natural Resources.

Mokany, K., Ware, C., Woolley, S. N. C., Ferrier, S., & Fitzpatrick, M. C. (2022). A working guide to harnessing generalized dissimilarity modelling for biodiversity analysis and conservation assessment. *Global Ecology and Biogeography, 31*(4), 802–821. https://doi.org/10.1111/geb.13459

Moss, S. J., & Wilson, M. E. J. (1998). Biogeographic implications of the Tertiary palaeogeographic evolution of Sulawesi and Borneo. In R. Hall & J. D. Holloway (Eds.), *Biogeography and geological evolution of SE Asia* (pp. 133–163). Backhuys Publishers.

Myers, G. S. (1960). The endemic fish fauna of Lake Lanao, and the evolution of higher taxonomic categories. *Evolution, 14*(3), 323–333. https://doi.org/10.1111/j.1558-5646.1960.tb03094.x

Ng, H. H., & Kottelat, M. (2007). Descriptions of three new species of torrent catfishes (Teleostei: Sisoridae: Glyptothorax) from Indochina and the Philippines. *Ichthyological Exploration of Freshwaters, 18*(3), 193–214.

Oksanen, J., Simpson, G., Blanchet, F., Kindt, R., Legendre, P., Minchin, P. R., O'Hara, R. B., Solymos, P., Stevens, M. H. H., Szoecs, E., Wagner, H., Barbour, M., Bedward, M., Bolker, B., Borcard, D., Carvalho, G., Chirico, M., De Cáceres, M., Durand, S., Evangelista, H. B. A., FitzJohn, R., Friendly, M., Furneaux, G., Hill, M. O., Lahti, L., McGlinn, D., Ouellette, M.-H., Cunha, E. R., Smith, T., Stier, A., ter Braak, C. J. F., Weedon, J., & Borman, T. (2022). *vegan: Community ecology package* (Version 2.6-4) [R package]. CRAN. https://cran.r-project.org/package=vegan

Olden, J. D., Lockwood, J. L., & Parr, C. L. (2011). Biological invasions and the homogenization of faunas and floras. In R. J. Whittaker & R. J. Ladle (Eds.), *Conservation biogeography* (pp. 224–243). Wiley-Blackwell.

PAGASA (Philippine Atmospheric, Geophysical and Astronomical Services Administration). (2024). Climate Data. Department of Science and Technology, Philippines. https://www.pagasa.dost.gov.ph

Přikryl, T., Brignon, A., & Carnevale, G. (2025). Fossil fish assemblage of the Laguna Formation, Philippines: Unveiling the uniqueness of Pleistocene freshwater ecosystems in Southeast Asia. *Swiss Journal of Palaeontology, 144*(1), 2. https://doi.org/10.1186/s13358-024-00347-0

Pubellier, M., Monnier, C., Maury, R. C., & Tamayo, R. A. (2004). Plate kinematics, origin and tectonic emplacement of supra-subduction ophiolites in SE Asia. *Tectonophysics, 392*(1–4), 9–36. https://doi.org/10.1016/j.tecto.2004.04.028

QGIS Development Team. (2023). *QGIS geographic information system [Computer software]*. Open Source Geospatial Foundation Project. http://www.qgis.org

R Core Team. (2023). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Rahel, F. J. (2007). Biogeographic barriers, connectivity and homogenization of freshwater faunas: It's a small world after all. *Freshwater Biology, 52*(4), 696–710. https://doi.org/10.1111/j.1365-2427.2006.01708.x

Robles, E. C. (2013). Estimates of Quaternary Philippine coastlines, land bridges, submerged river systems and migration routes: A GRASS GIS approach. *Hukay, 18*(1), 31–53.

Rohling, E. J., Fenton, M., Jorissen, F. J., Bertrand, P., Ganssen, G., & Caulet, J. P. (1998). Magnitudes of sea-level lowstands of the past 500,000 years. *Nature*, 394(6689), 162–165.

Romeroso, R. B., Tandang, D. N., & Navarrete, I. A. (2021). New distributional record of Phyllanthus securinegoides Merr. (Phyllanthaceae) and Rinorea niccolifera Fernando (Violaceae) of Homonhon Island, Philippines. *Biodiversitas, 22*(1), 501–506. https://doi.org/10.13057/biodiv/d220148

Setiawan, F., Laffan, S. W., & Doherty, T. S. (2020). Isolation by marine barriers structures freshwater fish diversity in an oceanic island group. *Diversity and Distributions, 26*(12), 1753–1766. https://doi.org/10.1111/ddi.13161

Shelley, J. J., Dempster, T., Le Feuvre, M. C., Unmack, P. J., Laffan, S. W., & Swearer, S. E. (2019). A revision of the bioregionalisation of freshwater fish communities in the Australian Monsoonal Tropics. *Ecology and Evolution, 9*(8), 4568–4588. https://doi.org/10.1002/ece3.5059

Siler, C. D., Oaks, J. R., Esselstyn, J. A., Diesmos, A. C., & Brown, R. M. (2010). Phylogeny and biogeography of Philippine bent-toed geckos (Gekkonidae: Cyrtodactylus) contradict a prevailing model of Pleistocene diversification. *Molecular Phylogenetics and Evolution, 55*(2), 699–710. https://doi.org/10.1016/j.ympev.2010.01.027

Siler, C. D., Oaks, J. R., Welton, L. J., Linkem, C. W., Swab, J. C., Diesmos, A. C., & Brown, R. M. (2012). Did geckos ride the Palawan raft to the Philippines? *Journal of Biogeography, 39*(7), 1217–1234. https://doi.org/10.1111/j.1365-2699.2011.02680.x

Sneath, P. H. A., & Sokal, R. R. (1973). Numerical taxonomy: the principles and practice of numerical classification. W. H. Freeman.

Stelbrink, B., Albrecht, C., Hall, R., & von Rintelen, T. (2012). The biogeography of Sulawesi revisited: Is there evidence for a vicariant origin of taxa on Wallace's "anomalous island"? *Evolution, 66*(7), 2252–2271. https://doi.org/10.1111/j.1558-5646.2012.01588.x

Strayer, D. L., & Dudgeon, D. (2010). Freshwater biodiversity conservation: Recent progress and future challenges. *Journal of the North American Benthological Society, 29*(1), 344–358. https://doi.org/10.1899/08-171.1

Tedesco, P. A., Beauchard, O., Bigorne, R., Blanchet, S., Buisson, L., Conti, L., Cornu, J.-F., Dias, M. S., Grenouillet, G., Hugueny, B., Jézéquel, C., Leprieur, F., Brosse, S., & Oberdorff, T. (2017). A global database on freshwater fish species occurrence in drainage basins. *Scientific Data, 4*, 170141. https://doi.org/10.1038/sdata.2017.141

Tisseuil, C., Cornu, J.-F., Beauchard, O., Brosse, S., Darwall, W., Holland, R., Hugueny, B., Tedesco, P. A., & Oberdorff, T. (2013). Global diversity patterns and cross-taxa convergence in freshwater systems. *Journal of Animal Ecology, 82*(2), 365–376. https://doi.org/10.1111/1365-2656.12018

Tuomisto, H. (2010). A diversity of beta diversities: Straightening up a concept gone awry. Part 2. Quantifying beta diversity and related phenomena. *Ecography, 33*(1), 23–45. https://doi.org/10.1111/j.1600-0587.2009.06148.x

Vallejo, A. N. (1985). Fishes of Laguna de Bay. *Natural and Applied Science Bulletin, 37*(4), 285–346.

Vallejo, B. J. (2011). The Philippines in Wallacea. In D. Telnov (Ed.), *Biodiversity, biogeography and nature conservation in Wallacea and New Guinea* (Vol. 1, pp. 27–42). Entomological Society of Latvia.

Vilhena, D. A., & Antonelli, A. (2015). A network approach for identifying and delimiting biogeographical regions. *Nature Communications*, 6, 6848.

Vitule, J. R. S., Freire, C. A., & Simberloff, D. (2009). Introduction of non-native freshwater fish can certainly be bad. *Fish and Fisheries, 10*(1), 98–108. https://doi.org/10.1111/j.1467-2979.2008.00312.x

Voris, H. K. (2000). Maps of Pleistocene sea levels in Southeast Asia: Shorelines, river systems and time durations. *Journal of Biogeography, 27*(5), 1153–1167. https://doi.org/10.1046/j.1365-2699.2000.00489.x

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 refugial-phase biodiversity. *Biodiversity and Conservation, 19*(4), 919–941. https://doi.org/10.1007/s10531-010-9783-3

Yumul, G. P., Jr., Dimalanta, C. B., Tamayo, R. A., Jr., & Maury, R. C. (2003). Collision, subduction and accretion events in the Philippines: A synthesis. *Island Arc, 12*(2), 77–91. https://doi.org/10.1046/j.1440-1738.2003.00382.x

Tables

Table 1. List of freshwater fish biogeographic provinces in the Philippines with their merge height values from the WPGMA dendrogram (the βsim value at which each province's constituent catchments cluster together, indicating internal compositional variation within the province), number of catchments, corrected weighted endemism (CWE) ranges, and observed species richness (SR) ranges.

Biogeographic province	Merge height	N catchments	CWE range	SR range
Palawan–Mindoro	0.544	4	0.150–0.279	55–72
Mindanao	0.509	8	0.088–0.388	25–75
Greater Visayas	0.451	7	0.120–0.181	34–98
Central/Southern Luzon	0.450	9	0.095–0.191	24–105
Northern Luzon	0.409	4	0.127–0.230	32–55

Table 2. Generalized dissimilarity model (GDM) variable importance for explaining the variation in freshwater fish compositional turnover across Philippine catchments. Values are percent contribution to deviance explained, computed via the matrix-permutation procedure. p-values are from the permutation test. Predictor names: "Geographic" = pairwise Euclidean distance between catchment centroids; "PAIC" = binary categorical distance between Pleistocene Aggregate Island Complex groupings.

Predictor	**Importance (% deviance)**	**p-value**
Geographic distance	13.5	0.028
area_km2 (catchment area)	11.0	0.172
temp_max_c (maximum temperature)	9.3	0.226
elev_m (elevation)	8.1	0.178
matrix_1 (PAIC)	2.8	0.228
rain_mm (annual precipitation)	1.6	0.506
drain_density (drainage density)	0.6	0.598
Total deviance explained	**36.39%**	

Table 3. Indicator Value (IndVal) analysis results for each biogeographic province. Species are listed in descending IndVal within each province. Ecological category (Eco.): PF — obligate primary freshwater; DI — diadromous; EU — euryhaline.

Species	Eco.	IndVal	p-value
Northern Luzon (11 indicators)			
Anguilla japonica	DI	0.943	0.001
Anguilla luzonensis	DI	0.899	0.001
Crenimugil crenilabis	DI	0.846	0.003
Crenimugil heterocheilos	DI	0.760	0.004
Rhinogobius philippinus †	PF	0.737	0.019
Plicofollis magatensis	EU	0.707	0.029
Mugil cephalus	DI	0.707	0.029
Crenimugil buchanani	EU	0.688	0.015
Oligolepis acutipennis	EU	0.673	0.019
Lamnostoma taylori	EU	0.588	0.021
Sicyopterus macrostetholepis	DI	0.586	0.049
Central/Southern Luzon (8 indicators)			
Arius manillensis	EU	0.816	0.002
Leiopotherapon plumbeus †	PF	0.816	0.001
Zenarchopterus philippinus †	PF	0.767	0.006
Cynoglossus puncticeps	EU	0.670	0.025
Gobiopterus brachypterus †	PF	0.667	0.007
Mistichthys luzonensis †	PF	0.667	0.008
Psammogobius biocellatus	EU	0.595	0.039
Hippichthys heptagonus	EU	0.580	0.044
Greater Visayas (11 indicators)			

Species	Eco.	IndVal	p-value
Ophichthus polyophthalmus	EU	0.894	0.001
Glossogobius bicirrhosus	EU	0.845	0.002
Moringua raitaborua	EU	0.756	0.013
Glossogobius aureus	EU	0.693	0.007
Osteomugil cunnesius	EU	0.662	0.031
Hippichthys penicillus	EU	0.630	0.019
Lutjanus malabaricus	EU	0.627	0.013
Ambassis gymnocephalus	EU	0.624	0.021
Pomadasys argenteus	EU	0.599	0.02
Favonigobius reichei	EU	0.596	0.033
Dichotomyctere nigroviridis	EU	0.579	0.041
Mindanao (6 indicators)			
Clarias macrocephalus †	PF	0.739	0.001
Anguilla borneensis	DI	0.707	0.005
Anguilla bicolor	DI	0.640	0.002
Butis gymnopomus	EU	0.624	0.026
Belobranchus belobranchus	DI	0.584	0.032
Sillago sihama	EU	0.524	0.002
Palawan–Mindoro (24 indicators)			
Rasbora argyrotaenia †	PF	1.000	0.002
Neostethus bicornis	EU	0.866	0.003
Nematabramis alestes †	PF	0.802	0.004
Mugilogobius mertoni	EU	0.794	0.002
Hemigobius hoevenii	EU	0.737	0.011
Barbodes hemictenus †	PF	0.707	0.029

Species	Eco.	IndVal	p-value
Barbodes palavanensis †	PF	0.707	0.025
Rasbora everetti †	PF	0.707	0.026
Rasbora lateristriata †	PF	0.707	0.025
Gobiosoma pallida	EU	0.707	0.029
Mangarinus waterousi	EU	0.707	0.029
Pandaka trimaculata	EU	0.707	0.029
Periophthalmus malaccensis	EU	0.707	0.027
Rhinogobius estrellae †	PF	0.707	0.026
Rhinogobius tandikan †	PF	0.707	0.025
Stiphodon palawanensis	DI	0.707	0.025
Dermogenys bispina †	PF	0.707	0.029
Dermogenys palawanensis †	PF	0.707	0.025
Dermogenys robertsi †	PF	0.707	0.022
Pterocryptis taytayensis †	PF	0.707	0.026
Eleutherochir opercularis	EU	0.682	0.009
Zenarchopterus gilli	EU	0.678	0.022
Planiliza macrolepis	EU	0.613	0.046
Zenarchopterus dispar	EU	0.608	0.041

† Species also significant ($p \leq 0.05$) as an indicator in the PF-only IndVal analysis restricted to the obligate primary-freshwater species.

Figure Captions

Figure 1. [Single Column] Map of the Philippines with elevation shading, islands and catchment/river basin boundaries included in the analysis. Blue shading indicates sea floor depth of 120 m (estimated sea level during the last glacial maximum, Pleistocene period). Pleistocene Aggregate Island Complexes (PAIC) are labelled.

Figure 2. [Double Column] WPGMA hierarchical dendrogram based on Simpson's beta (βsim) pairwise dissimilarity matrix among all catchments, overlaid on a map of the Philippines to show the spatial correspondence of the five proposed freshwater fish biogeographic provinces. Catchment numbers correspond to those listed in Table S2. Gray shaded catchments were excluded from analysis due to insufficient fish data.

Figure 3. [Single Column] Corrected weighted endemism (A), and species richness (B) of freshwater fishes per catchment unit across the Philippines. Gray shaded catchments were excluded from analysis due to insufficient fish data.

1 **Figure 1.**

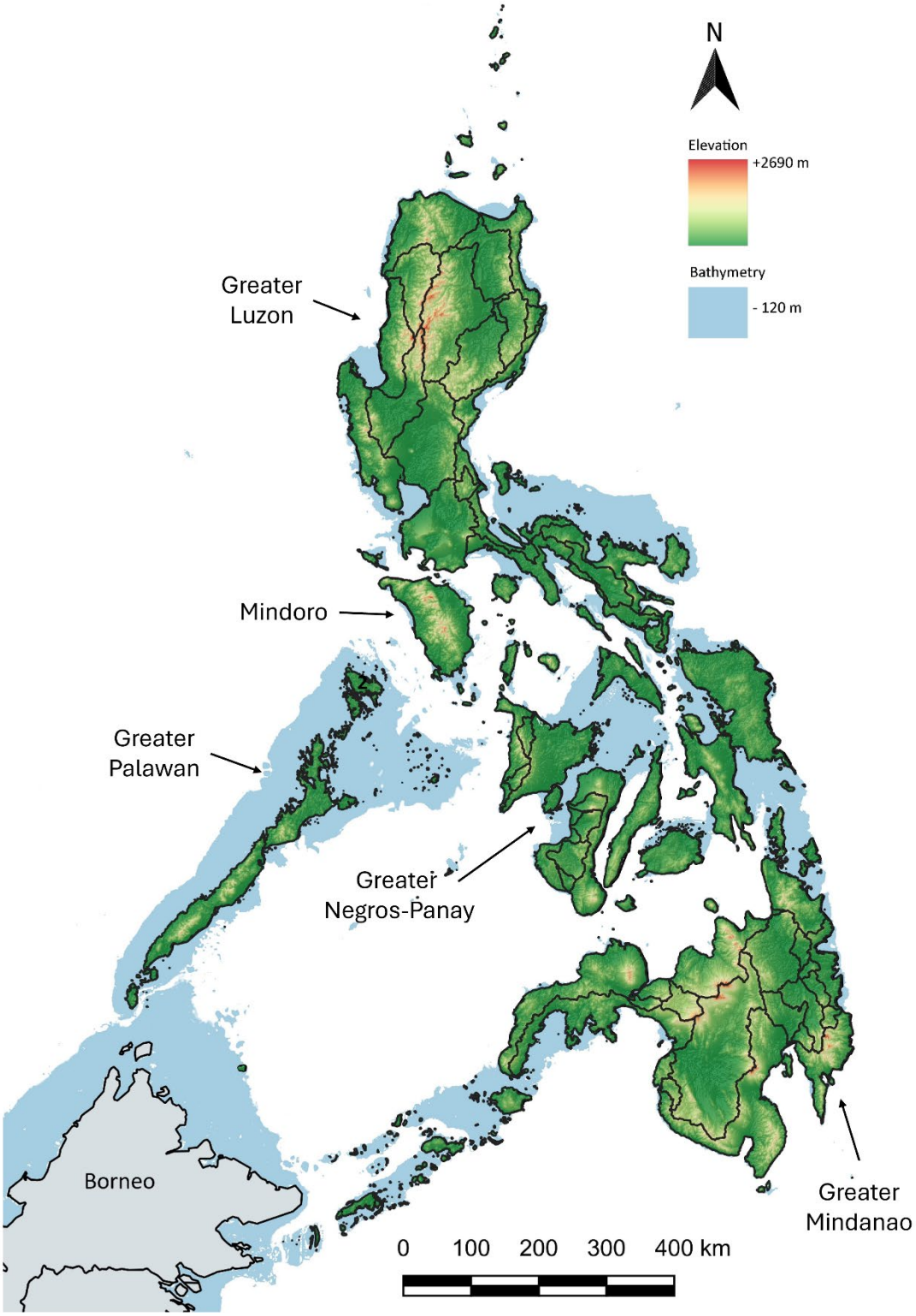

N
Elevation
+2690 m
Bathymetry
- 120 m
Greater Luzon
Mindoro
Greater Palawan
Greater Negros-Panay
Borneo
Greater Mindanao
0 100 200 300 400 km

Figure 2.

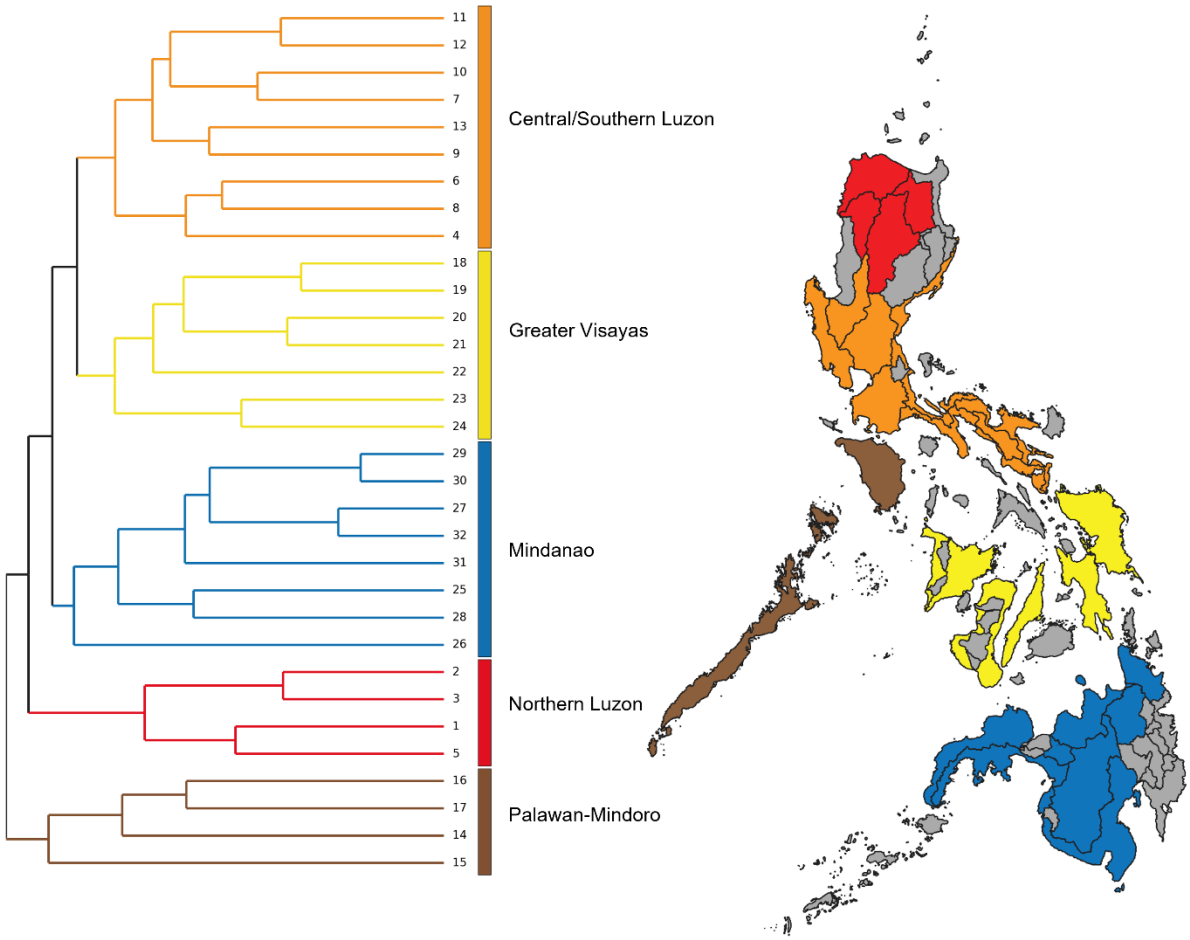
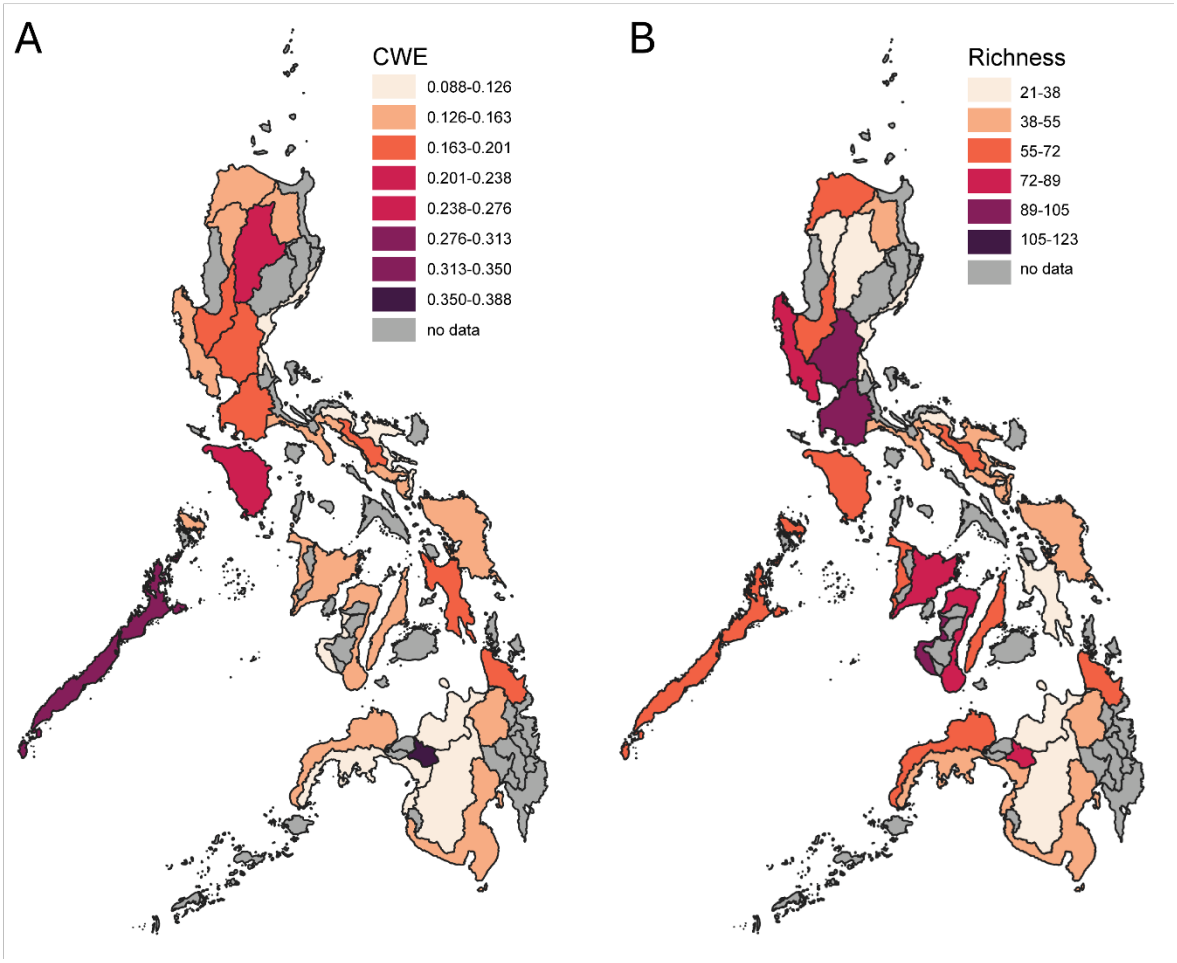

11
12
10
7
13
9
6
8
4
18
19
20
21
22
23
24
29
30
27
32
31
25
28
26
2
3
1
5
16
17
14
15
Central/Southern Luzon
Greater Visayas
Mindanao
Northern Luzon
Palawan-Mindoro

Figure 3.


A
CWE
0.088-0.126
0.126-0.163
0.163-0.201
0.201-0.238
0.238-0.276
0.276-0.313
0.313-0.350
0.350-0.388
no data
B
Richness
21-38
38-55
55-72
72-89
89-105
105-123
no data