

Article

Historical connectivity and environmental filtering jointly determine the freshwater fish assemblages on Taiwan and Hainan Islands of China

Haoxian LIN, Chao DAI, Hongyin YU, Jiahao TU, Jiehua YU, Jiekun HE*, and Haisheng JIANG*

Spatial Ecology Laboratory, School of Life Sciences, South China Normal University, Guangzhou 510631, China

*Address correspondence to Jiekun He, E-mail: jiekunhe@m.scnu.edu.cn and Haisheng Jiang, E-mail: jhs@scnu.edu.cn.

Handling editor: Zhi-Yun Jia

Received on 9 December 2021; accepted on 28 February 2022

Abstract

The biotas of Taiwan and Hainan Islands are of continental origin, but the manner with which historical and ecological factors shaped these insular species is still unclear. Here, we used freshwater fish as a model to fill this gap by quantifying the phylogenetic structure of the insular faunas and disentangling the relative contribution of potential drivers. Firstly, we used clustering and ordination analyses to identify regional species pools. To test whether the insular freshwater fish faunas were phylogenetically clustered or overdispersed, we calculated the net relatedness index (NRI) and the nearest taxon index (NTI). Finally, we implemented logistic regressions to disentangle the relative importance of species attributes (i.e., maximum body length, climatic niche dissimilarity [ND], and diversification) and historical connectivity (HC) in explaining the insular faunas. Our results showed that the most possible species pools of Taiwan are Zhejiang and Fujian provinces, and those of Hainan are Guangdong and Guangxi provinces. These insular faunas showed random phylogenetic structures in terms of NRI values. According to the NTI values, however, the Taiwanese fauna displayed more phylogenetic clustering, while the Hainanese one was more overdispersed. Both the standard and phylogenetic logistic regressions identified HC and climatic ND as the 2 top explanatory variables for species assemblages on these islands. Our reconstruction of the paleo-connected drainage basins provides insight into how historical processes and ecological factors interact to shape the freshwater fish fauna of the East Asian islands.

Key words: biogeography, community assemblage, paleo-drainage, sea level change.

Island systems harbor a substantial fraction of global biodiversity and provide excellent opportunities to understand the evolutionary processes underlying community assemblages (MacArthur and Wilson 1967; Wallace 1869). In general, the composition of an insular species assemblage is proposed to emerge from geographical barriers (Wallace 1869), biogeographical processes (MacArthur and Wilson 1967), and island characteristics (Gillespie and Roderick 2002). Apart from abiotic factors, the

extent to which species inhabit in island systems is likely related to their intrinsic ecological and evolutionary features (Carvajal-Endara et al. 2017). Nevertheless, how these factors interact to shape the present-day insular biotas is still poorly understood, as underlying drivers vary among island systems and taxonomic groups (Cardillo et al. 2008; Weigelt et al. 2015); this hinders our understanding of the effects of global changes on insular biodiversity (Bellard et al. 2014; Gallardo et al. 2018; Leclerc et al. 2020).

Naturalists have observed for centuries that an insular biota is a non-random combination of species (Darwin 1859; Wallace 1869); it represents either phylogenetic clustering or overdispersion that results from the interplay between evolutionary and ecological processes (Cardillo et al. 2008; Weigelt et al. 2015). From this perspective, the most popular hypothesis proposes that biotas on island systems are primarily shaped by dispersal limitation (e.g. Burns 2005; Kubota et al. 2014; Guimarães et al. 2021). This process has been invoked to explain the higher biotic similarity of mainland islands than that of oceanic islands with the mainland (Kubota et al. 2014), which is also known as the distance decay of taxonomic similarity (Nekola and White 1999). However, considering the geographical barrier as the sole driver in insular biotas is likely an oversimplification, as dispersal is only the first step in the colonization process. The occurrence of species in a community also depends on their ability to sustain beyond these environments, that is, the environmental filtering hypothesis (Graham et al. 2014). Several studies have highlighted the importance of environmental filtering in determining island community assembly (Carvajal-Endara et al. 2017; Si et al. 2017). Nevertheless, due to substantial differences in island systems (Gillespie and Roderick 2002), historical structure (Ali 2017), and bioclimatic environments (Weigelt et al. 2013), the patterns and underlying determinants of freshwater fish assembly may be distinct among island systems.

Apart from extrinsic factors, several ecological and evolutionary features of insular species might enable them to persist on island systems. For example, insular species should have a faster diversification rate. This is because long-term isolation can often lead to rapid radiation possibly by establishing geographical barriers and creating novel habitats where species evolve and diversify (Kier et al. 2009; Losos and Ricklefs 2009). In addition, the successful colonization and survival of insular species have usually been linked to their morphological traits (Lomolino 2005; Meiri et al. 2008; Benitez-Lopez et al. 2021). For example, we would expect insular freshwater fish to have small body sizes because the rivers on islands are generally characterized by short length, small catchment areas, steep channel gradients, and fast flow velocity. Overall, species co-occurring on islands are expected to have small body sizes, share similar climatic niche properties to their mainland counterparts, have a fast diversification rate, and originate from the same paleo-connected landmasses.

Freshwater fish are important organisms in studies that address questions regarding the eco-evolutionary processes underlying the current insular biodiversity (Dias et al. 2014). Unlike terrestrial organisms (mostly terrestrial vertebrates and plants), strictly freshwater species (i.e., organisms that live and disperse exclusively in freshwater) scarcely disperse across marine waters, mountain ranges, and expansive lands (Griffiths 2017; Carvajal-Quintero et al. 2019; Leroy et al. 2019). Therefore, analyses of freshwater fishes can identify some biogeographical signals that might not be captured by terrestrial groups (Campbell Grant et al. 2007; Rinaldo et al. 2014; Leroy et al. 2019). To our knowledge, few studies have explored the imprint of historical riverine connectivity in present-day insular freshwater fish faunas (but see Dias et al. 2014). Here, we fill this gap by exploring the patterns and factors underlying the community structure of freshwater fishes on East Asian islands, where the largest land extension has occurred during low sea-level periods caused by Quaternary glacial cycles (Qiu et al. 2011).

Taiwan and Hainan Islands are isolated islands from the east of Asia continent and are known for their high levels of biodiversity and endemism (Tzeng 1986; Kang et al. 2014; Chen et al. 2016; Feng et al. 2016). Previous studies have suggested that the periodic

emergence of a land bridge during Quaternary glacial cycles facilitated biotic interchanges between these islands and the neighboring mainland Asia (Figure 1; Chiang et al. 2010; He et al. 2018; Wang et al. 2021), followed by community assembly via environmental filtering (Lin et al. 2021). In this study, we aimed to identify the species pools with phylogenetic beta dissimilarity among different drainage basins and disentangle the relative importance of morphology, environmental filtering, historical connectivity (HC), and speciation in structuring the freshwater fish faunas of Taiwan and Hainan Islands, China. We addressed these objectives in a 3-step workflow as follows: 1) to identify the species pools of Taiwan and Hainan Islands using clustering and ordination analyses based on phylogenetic beta dissimilarity; 2) to analyze the phylogenetic structure of the insular fish faunas relative to potential species pools; and 3) to disentangle the effect of eco-evolutionary factors on fish communities using standard and phylogenetic logistic regression models.

Materials and Methods

Study area

Our study area comprises the drainage basins of Taiwan, Hainan, and other neighboring mainland Asia regions, including the Guangxi, Guangdong, Fujian, and Zhejiang Provinces in China, Vietnam, and Laos (Figure 1). Particularly, the freshwater fish assemblages of Taiwan and Hainan Islands are our focus. Taiwan (~35,800 km²) and Hainan (~33,900 km²) are the 2 largest islands of China, and their drainage basins were connected to the mainland when sea level decreased (Figure 1). It has been estimated that Taiwan and Hainan have, respectively, appropriately 90 species (Chen et al. 2008) and 138 species (Xiong et al. 2018) of native freshwater fish, and among those, 42 and 14 are endemic species (Chen et al. 2008; Xiong et al. 2018), accounting for 46.7% and 10.1% of total native species. The differences in total number and endemic levels of freshwater fish may be structured by the different geographical contexts between Taiwan and Hainan Islands, such as distances and period of isolation from the mainland, source species pools, and inner environmental heterogeneity (Ali 2018; He et al. 2018; Lin et al. 2021).

Species data

We assembled species lists of freshwater fishes from our field surveys, published fauna data, and regional checklists (Supplementary Table S1). To increase the comparability among these basins, we used watersheds of primary tributaries as basic units. We used only native and strictly freshwater species (e.g. obligate freshwater dispersal) and assumed that drainage basins were separated from one another by land and marine water (Leprieur et al. 2011). We combined all infraspecific taxa with their species and standardized the species names according to FishBase (<http://www.fishbase.in>) and Global Biodiversity Information Facility (<https://www.gbif.org>).

We combined the species checklist with a time-calibrated phylogenetic tree of freshwater fishes (31,526 species) from Rabosky et al. (2018). In this phylogeny, the topology of 11,638 species with genetic data was fixed and the remaining 19,888 species that were unsampled for genetic data were assigned using stochastic polytomy resolution based on taxonomy, resulting in a distribution of 100 trees. As a result, our dataset included 81 basins with sizes ranging from 276 to 86,586 km² (Figure 1 and Supplementary Table S2), comprising a total of 540 primary freshwater fishes belonging to 186 genera and 27 families (Supplementary Tables S3 and S4).

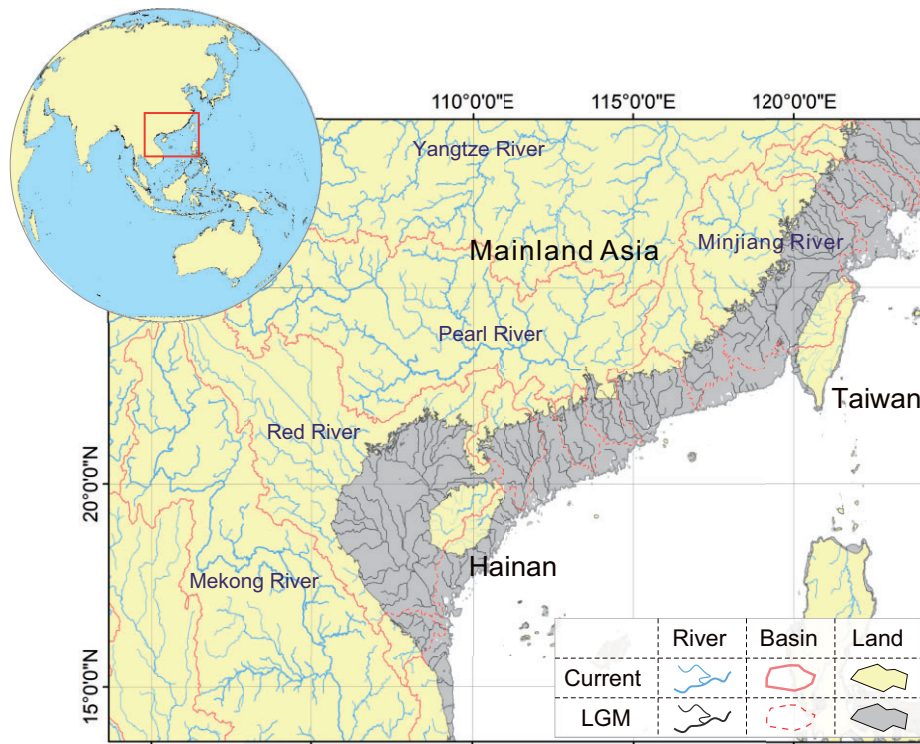


Figure 1. The paleo- and current landmasses, drainage basins, and rivers of Taiwan, Hainan Island, and Mainland Asia. The map of current landmasses was downloaded from the Resource and Environment Science and Data Center of China (<http://www.resdc.cn/>). The map of paleo-landmass was reconstructed from ETOPO1 1 arc-minute global relief models, with paleo-landmass indicating the exposed land bridge during the Last Glaciation Maximum when the sea level was -120 m below the current sea level.

Identification of potential species pools

We used distance-based clustering analysis to identify the potential species pools of the insular freshwater fish assemblages. We recorded the presence or absence of the species in each basin (Supplementary Table S5) and calculated the phylogenetic Simpson dissimilarity index ($p\beta_{sim}$) between each pair of freshwater fish assemblages using the R package “betapart” (Baselga and Orme 2012). $p\beta_{sim}$ is a widely used index for community dissimilarity because it accounts for the evolutionary histories among species (Graham and Fine 2008) and is not affected by differences in species richness (Koleff et al. 2003). This index calculates the pairwise phylogenetic dissimilarity between 2 communities by replacing shared and unique species with shared and unique branch lengths, respectively (Leprieur et al. 2012), as follows:

$$p\beta_{sim} = \frac{\min(PD_{Total} - PD_k, PD_{Total} - PD_j)}{PD_k + PD_j - PD_{Total} + \min(PD_{Total} - PD_k, PD_{Total} - PD_j)},$$

where PD_j and PD_k are the total branch lengths of communities j and k , respectively. PD_{Total} is the total branch length of a phylogenetic tree containing all species present in both j and k communities.

Then, we used the unweighted pair group method with arithmetic mean (UPGMA) clustering to investigate the faunistic relationships between Taiwan and Hainan and neighboring landmasses. We assessed the performance of the UPGMA in transferring the matrices into dendrograms using cophenetic correlation coefficients (Kreft and Jetz 2010). The optimal number of clusters was determined using the “optimal_phyloregion” function in the R package “phyloregion” (Daru et al. 2020), which divides the hierarchical dendrogram into meaningful clusters based on the “elbow” or

“knee” of an evaluation graph that corresponds to the point of optimal curvature (Daru et al. 2020). We performed non-metric multidimensional scaling (NMDS) ordination to illustrate the faunistic dissimilarity using the “monoMDS” function in the R package “vegan” (Oksanen et al. 2019). We calculated stress values to assess the fit between the NMDS and the original dissimilarity matrix. This index ranges from 0 to 1, with smaller values indicating better NMDS results.

Phylogenetic structure analyses

To test whether the insular freshwater fish communities in Taiwan and Hainan were phylogenetically clustered or overdispersed, we used the net relatedness index (NRI) and the nearest taxon index (NTI) of Webb et al. (2002). These 2 indices emphasize the different depths of evolutionary history (Qian et al. 2020). NRI reflects the overall structure throughout the phylogeny, whereas NTI is more sensitive to structures near the tips. NRI and NTI test the significance of the observed pattern and calculate the standardized effect size of the mean pairwise phylogenetic distance (MPD) and the mean nearest taxon distance (MNTD) between all species with respect to null expectations generated by selecting species randomly across the tips of the phylogeny (999 runs)

$$NRI \text{ or } NTI = -1 \times \frac{X_{sample} - X_{rnsample}}{sd(X_{rnsample})}.$$

For NRI, X_{sample} is the observed MPD and $X_{rnsample}$ is the expected MPD of the randomized assemblages. For NTI, X_{sample} is the observed MNTD and $X_{rnsample}$ is the expected MNTD of the randomized assemblages. Positive NRI or NTI values indicate

phylogenetic clustering, whereas negative values indicate phylogenetic overdispersion. As NRI and NTI approximate the Z distribution, a value greater than 1.96 (or less than -1.96) is equivalent to significance at $P < 0.05$. We calculated the NRI and NTI using 100 phylogenetic trees and reported the results as medians. To ensure the robustness of our results and evaluate sensitivity to different sizes of species pools, we calculated the NRI and NTI metrics using 3 alternative source pools (small, medium, and large) based on different depths of cut-off points in the clustering dendrogram.

Potential drivers of community assembly

Based on the ecological theory and hypotheses proposed in previous studies, we developed a set of predictions regarding the potential drivers of freshwater fish communities in Taiwan and Hainan.

Maximum body length

We described fish body size using the log-transformed maximum body length (BL) that was obtained from FishBase (<http://www.fishbase.se>) and published literature (see [Online Appendix S1](#) for details). In cases in which maximum BL data of specific species were not found ($n = 67$ out of 540 species; 12.4%), we used either genus ($n = 62$) or family ($n = 5$) mean maximum BL values instead, as suggested by several previous critical studies (e.g., [Blanchet et al., 2010](#) [Blanchet et al. 2010](#); [Griffiths 2012](#)).

Climatic niche dissimilarities

Climate data were obtained from the WorldClim v.2 database with a resolution of 30 arc-seconds ([Fick and Hijmans 2017](#); <http://www.worldclim.org>). We calculated climate niche dissimilarity (ND) based on the mean annual temperature (BIO1), minimum temperature of the coldest month (BIO6), and mean annual precipitation (BIO12) using the “hypervolume_distance” function in the R package “hypervolume” ([Blonder and Harris 2019](#)). This metric represents the Euclidean distance between each species’ hypervolume centroid and the centroid of the climatic hypervolume represented by Taiwan and Hainan.

Diversification

We defined the rate of diversification for each species using the species-per-genus ratio, that is, the total number of species within each genus. This metric indicates that species belonging to one genus have the same diversification rate and have been used in several previous studies ([Verdú 2002](#); [Carvajal-Quintero et al. 2019](#)). We obtained the total number of species per genus from the FishBase. We also used the tip diversification rate of each species provided by [Rabosky et al. \(2018\)](#). In cases in which the tip diversification rate of a species was not found ($n = 248$ out of 540 species; 45.9%), we used the mean values of all species within a specific genus ($n = 221$) or family ($n = 27$).

Historical connectivity

To assess the effect of HC between basins on the insular freshwater fish community assemblages, we recorded the presence or absence of the species in the paleodrainages connecting to either Taiwan or Hainan Island. The paleodrainages during the Last Glacial Maximum (LGM) were reconstructed with a spatial resolution of 15 arc-seconds (approximately 0.5 km) from a global terrain model (GEBCO_2021; <http://www.gebco.net/>). We considered the land area exposed during the LGM when the sea level was -120 m and we then identified the paleo-connected drainage basins in ArcGIS

10.2 (ESRI, Redlands, CA, USA) using hydrological tools ([Dias et al. 2014](#)).

Statistical analyses

We implemented logistic regressions to disentangle the relative importance of species attributes (i.e., maximum BL, climatic ND, and diversification) and HC in explaining insular freshwater fish communities. We recorded species presence/absence on the islands as the binary response variable and used species’ maximum BL, climatic ND, diversification (DR), and HC as predictor variables

$$0/1 \sim \text{BL} + \text{ND} + \text{DR} + \text{HC}.$$

We performed all models separately for Taiwan and Hainan Islands for each source pool definition. All predictors were standardized (mean = 0, SD = 1) to allow comparisons between model coefficients. We assessed the collinearity among variables by calculating variance inflation factors and found that collinearity was low (all VIFs < 2). We used a multimodel approach based on the sample-size-corrected AIC (AICc) to evaluate model-averaged parameter estimates in a likelihood-based framework. We ran 15 models including all possible combinations of the 4 predictors using the “dredge” function in R package “MuMIn” ([Bartoń 2019](#)); we then calculated the relative importance of each variable by summing the AICc weights (Σw_i) over all models that included the variable ([Burnham and Anderson 2002](#)).

In addition, we used phylogenetic logistic regressions ([Ives and Garland 2010](#)) in the R package “phylolm” ([Ho and Ané 2014](#)) to assess the phylogenetic relatedness among species. Compared with standard (non-phylogenetic) logistic regression, this method adds a measure of the phylogenetic signal (i.e., alpha; α) to the model. α is a parameter derived from the Ornstein–Uhlenbeck model of evolution, in which a large α value indicates low phylogenetic correlations among species ([Ives and Garland 2010](#)). To account for uncertain topography in phylogenetic trees, we ran phylogenetic logistic regressions over a set of 100 phylogenetic trees ([Rabosky et al. 2018](#)) and reported the median value and an inference interval as the 2.5–97.5th percentiles for model coefficients. The variance (R^2) explained by each model was estimated using the R package “rr2” ([Ives and Li 2018](#)). All analyses were performed using the R software v. 4.0.4 (R Foundation for Statistical Computing, <http://www.r-project.org>).

Results

Identification of species pools on islands

The UPGMA clustering of basin-based freshwater fish assemblages yielded 9 groups ([Figure 2A,B](#)). Hainan Island was first merged with China’s Guangdong and Guangxi provinces at a phylogenetic β_{sim} value of approximately 0.26 ([Figure 2A](#); node 1). Taiwan was first grouped with China’s Zhejiang and Fujian provinces at approximately 0.31 ([Figure 2A](#); node 2). Finally, these 2 groups merged ([Figure 2A](#); node 3) and then clustered with Vietnam at a phylogenetic β_{sim} value of approximately 0.40 ([Figure 2A](#); node 4). The cophenetic correlation coefficient was significant ($r > 0.67$; $P < 0.001$), indicating good agreement between the cluster assignments and the original phylogenetic β_{sim} matrix. Additionally, the major groups identified by clustering analysis were clearly separated from each other in the NMDS ordinations ([Figure 2C](#)). The stress value of the NMDS was 0.25, indicating that the NMDS ordination had an acceptable projection of the dissimilarity matrix into the 2-

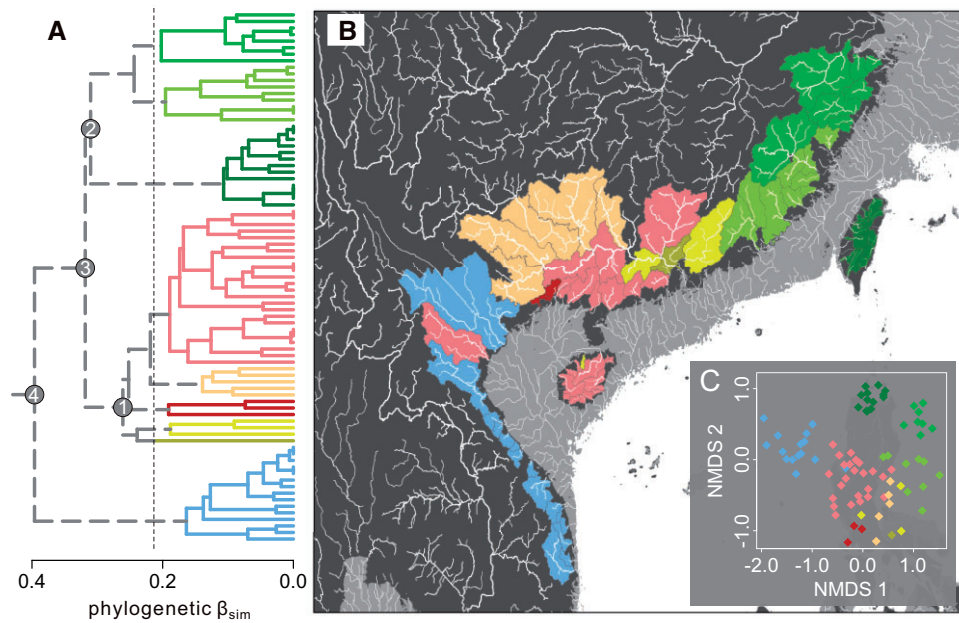


Figure 2. Classification maps of freshwater fish communities of Taiwan, Hainan Island, and Mainland Asia. (A) Dendrograms are yielded using the UPGMA hierarchical clustering based on the phylogenetic Simpson's dissimilarity index. The phenon line indicates the optimal cut-off location determined by the “optimal_phyloregion” function in the R package “phyloregion”. The colors of the UPGMA clusters are identical to those on (B) maps and in (C) NMDS ordinations.

dimensional ordination space. As a result, we used nodes 3 and 4 in the clustering dendrogram as the small species source pools for Taiwan and Hainan (Figure 2A), respectively. Nodes 1 and 2 were the large and medium species pools for both Taiwan and Hainan Islands.

Phylogenetic structure of the faunas

The NRI values of community structure for Taiwan and Hainan Islands were between -1.96 and 1.96 (Figure 3A) and none of the 100 randomized phylogenies were significant ($P > 0.05$). This indicated that the insular faunas represented phylogenetic subsets of species randomly drawn from their regional species pools. This result was generally consistent across all 3 regional species pools (small, medium, and large), except that species assemblages showed stronger phylogenetically clustering toward larger species source pools (Figure 3A). However, we found some differences between the NRI and NTI values (compare Figure 3A and B) and between Taiwan and Hainan Islands. For example, the freshwater fish fauna of Hainan was phylogenetically structured at random in terms of its NRI value, whereas it tended to be overdispersed based on its NTI value. In contrast, the Taiwanese fauna was much more strongly phylogenetically clustered as indicated by the NTI value than by the NRI metric, implying that phylogenetic clustering emerged recently in the phylogeny.

Predictors explaining the faunas

The results of the standard (non-phylogenetic) logistic regressions showed that species co-occurrence on the Taiwan and Hainan Islands are controlled by different drivers (Figure 4; see also Supplementary Figures S1 and S2 and Supplementary Tables S6 and S7). Among all the possible combinations of standard logistic regressions for Taiwan (Supplementary Table S6), the highest ranked model based on the AICc included ND and HC as explanatory variables (Figure 2A). The priority of these 2 predictors was also

identified by the sum of the Akaike weights (both $\Sigma wi = 1$, Figure 2A), while maximum BL ($\Sigma wi = 0.30$) and diversification ($\Sigma wi = 0.30$) had little explanatory power (Figure 2A). For Hainan Island, the best-fitting model included all variables except for diversification (Supplementary Table S7). The model averaging results showed that HC was the most important factor ($\Sigma wi = 1.00$; Figure 2B), followed by ND ($\Sigma wi = 0.92$), and maximum BL ($\Sigma wi = 0.87$). Our results consistently identified the relative low importance of diversification either using the measure of species-per-genus ratio or tip diversification rate (compare Supplementary Figure S1 with Supplementary Figure S3). The relative importance of variables identified by phylogenetic logistic regressions was highly consistent with that of standard logistic regression (compare Figure 4C,D with A,B) and among different sizes of species pools (compare Figure 4 with Supplementary Figures S1 and S2). Notably, phylogenetic signals existed in the community assembly of species on Hainan Island as the phylogenetic alpha coefficient (α) was almost zero ($\alpha = 0.13 \pm 0.10$). In contrast, the phylogenetic signals for Taiwan were relatively weak ($\alpha = 0.28 \pm 0.08$).

Discussion

This study delineated the regional species pools for Taiwan and Hainan Islands under a quantitative framework and supported their close relationships with the nearby regions of mainland Asia. Overall, both Taiwan and Hainan Islands represented phylogenetic subsets of species randomly drawn from their regional species pools. However, the Taiwanese fauna exhibited more phylogenetic clustering and the Hainanese one was more overdispersed toward the shallower structure in the phylogeny, compared with each other; this implies the important role of biogeographical processes (e.g., *in situ* speciation and multiple colonization pathways) in shaping the community structure of insular freshwater fish. To our knowledge, our reconstruction of the paleo-connected drainage basins is the first to

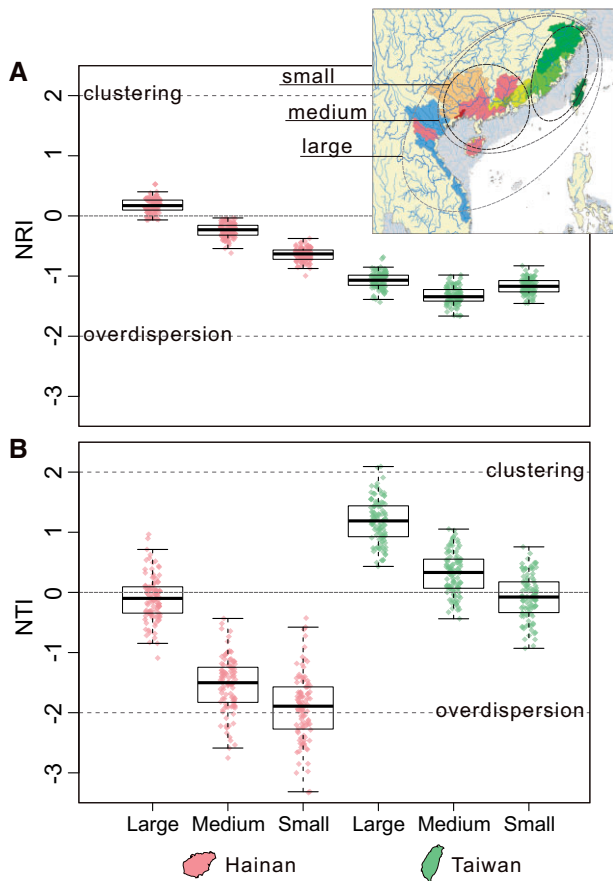


Figure 3. Phylogenetic community assemblage structure of freshwater fish on Taiwan and Hainan Islands. The phylogenetic structures were estimated using (A) the NRI and (B) the NTI within 3 (large, medium, and small) potential source pools.

provide empirical evidence on East Asian areas that support the joint effects of historical riverine connectivity and environmental filtering on the present-day patterns of freshwater fish biodiversity.

The UPGMA clustering and NMDS ordinations results indicate that the Taiwanese fauna resembled most closely those of China's Zhejiang and Fujian provinces. This finding is in line with that of a meta-analysis of phylogeographical studies by He et al. (2018), who revealed that South China shares a majority of terrestrial vertebrates with Taiwan. Indeed, the biotas of Taiwan Islands have long been recognized to originate from South China (Li et al. 2006; Jing et al. 2007; Qu et al. 2015). However, recent phylogeographical studies have proposed 4 other possible source regions, including North China (Päckert et al. 2009), the Eastern Himalayas (Mays et al. 2015), Japan (Li et al. 2011), and Indochina (Chen et al. 2015). The differences among these findings arise probably because previous analyses were largely based on the interpretation of single-taxon studies. Our study, based on freshwater fish, lends support to the traditional view that South China is the most likely source pool for Taiwanese fauna.

Our results showed that the freshwater fish fauna of Hainan is more affiliated with China's Guangdong and Guangxi provinces, followed by Vietnam (Figure 2). This finding differs from those of the recent floristic studies by Lin et al. (2021), who found the flora of Hainan to resemble most closely that of Vietnam. One possible reason for this discrepancy is that freshwater fish lack terrestrial

modes of dispersal (Shurin et al. 2009) and thus are generally more constrained by geographic barriers than plants. Given that freshwater fishes must disperse via freshwater channels, the affinities between Hainan Island and neighboring landmasses not only depend on the land bridges formed during glaciations, but are also determined by the extent to which drainage basins are connected with each other (Carvajal-Quintero et al. 2019). As the paleo-drainage basins of Hainan first merged with those of Guangdong (Figure 1), it is reasonable to assume that freshwater fish in Hainan share higher species similarity with those of China's Guangdong than with those of Vietnam.

Our results revealed that both Taiwan and Hainan Islands represented phylogenetic subsets of species randomly drawn from their regional species pools in terms of NRI values (Figure 3). This indicates that the freshwater fish faunas of the 2 islands generally resulted from random dispersal within the source-pool phylogeny. However, the freshwater fish fauna of Taiwan showed much stronger phylogenetic clustering based on the NTI metric than based on the NRI metric (Figure 3). This pattern implies that some specific clade of freshwater fish on Taiwan Island diversified recently, and thus the processes shaping its fauna are not independent of phylogeny. For example, a phylogeographical study showed that the Taiwanese endemic hillstream loaches (i.e., *Hemimyzon*) have diversified locally (Wang et al. 2007). Four species of *Pseudobagrus* have been recorded and are endemic to Taiwan (Chen 2008). This is similar to the cases of other freshwater fishes endemic to Taiwan, such as *Aphyocypris kikuchii*, *Candidia barbata*, *Gobiobotia cheni*, *Onychostoma alticorpus*, and *Opsariichthys pachycephalus* (Chen 2008). The high degree of endemism is related not only to the emergence of the Taiwan Strait, but also to the uplift of the Central Mountain Range, which formed barriers for dispersal (Wang et al. 2007). Thus, the discrepancy in the phylogenetic structure revealed by the NRI and NTI metrics highlights the importance of the most recent macroevolutionary processes, such as speciation, in shaping the freshwater fish fauna in Taiwan.

The freshwater fish fauna of Hainan showed more phylogenetic overdispersion based on the NTI value than on the NRI value (Figure 3A). This finding is in accordance with those of a global-scale study on insular floras by Weigelt et al. (2015), who showed that angiosperms on Hainan Island represent extreme phylogenetic overdispersion. One possible explanation is that freshwater fish native to Hainan are sourced from a variety of drainage basins, such as the Yangtze River, the Pearl River, and the Mekong River (Pearl River Fisheries Research Institute 1986). Multiple colonization pathways might increase the probability of Hainan's clade being overdispersed in the tip-level structure of phylogeny. Furthermore, compared with Taiwan, Hainan is less isolated and its environmental heterogeneity is too small to provide sufficient opportunities for species to rapidly diversify via spatial or ecological segregation. For example, although the land area of Hainan (33,900 km²) is nearly equal to that of Taiwan (35,800 km²), the number of species of freshwater fish endemic to Taiwan ($n=42$; Chen 2008) is 3-fold higher than that of Hainan ($n=14$; Xiong et al. 2018). Accordingly, the joint effects of low isolation and multiple colonization pathways would have resulted in high phylogenetic overdispersion in the freshwater fish fauna of Hainan Island.

Our study revealed that HC and environmental filtering jointly govern the co-occurrence of freshwater fish on Taiwan and Hainan Islands (Figure 4). It is not surprising that HC emerged as the most important factor. Regional-scale evidence from seed plants (Lin et al. 2021), birds (Chen et al. 2015), reptiles (Huang et al. 2013),

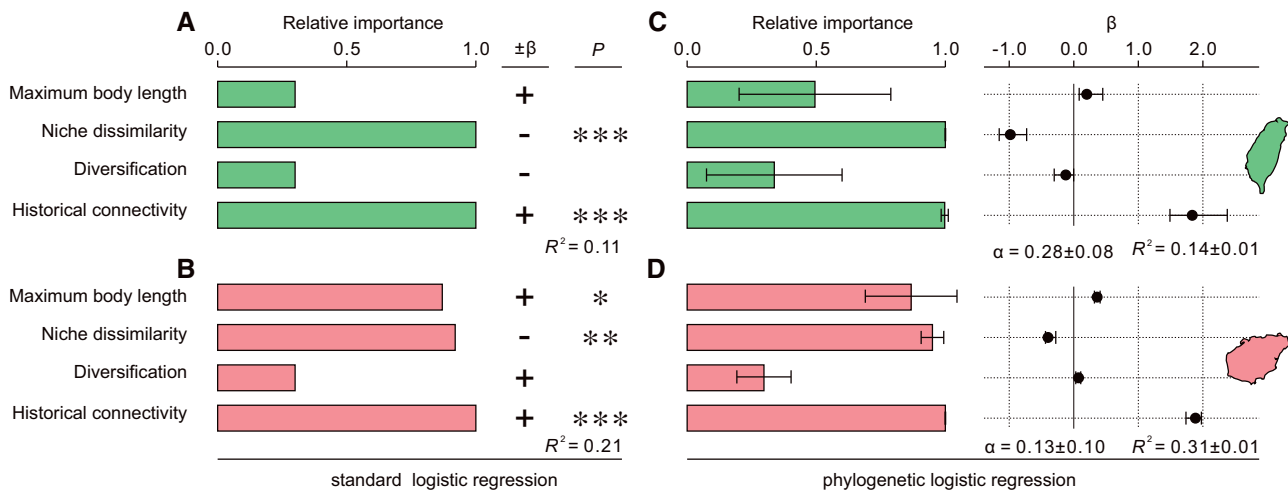


Figure 4. Regression coefficients of predictors and model diagnostics for logistic regression models explaining the occurrence of freshwater fish species on Taiwan and Hainan Islands. (A and B) Results of standard (non-phylogenetic) logistic regression for species assemblages on Taiwan (A) and Hainan Island (B). Bar lengths show that the relative importance value of each variable is calculated by the sum of the Akaike weight of models including them. Plus and minus symbols indicate the sign of model-averaged logistic regression coefficient estimates (β). Asterisks indicate the level of significance of the variables in logistic regression analyses (* $P < 0.05$, ** $P < 0.01$, and *** $P < 0.001$). R^2 indicates the variance explained by the best model. (C and D) Results of phylogenetic logistic regression for species assemblages on Taiwan (C) and Hainan Island (D). Bar lengths show the mean relative importance values with a vertical bar representing standard deviations based on the calculations of 100 phylogenetic trees. Filled dots represent the median values of regression coefficient (β) estimates for phylogenetic models with an inference interval as the 2.5–97.5th percentiles for 100 phylogenetic trees. The phylogenetic correlation (α) and the variance (R^2) explained by a model are shown as the mean values and standard deviation of the best models among 100 repetitions.

and other vertebrates (He et al. 2018) has supported the importance of a land bridge in community assembly on Taiwan and Hainan Islands. Our reconstruction of the paleo-connected drainage basins revealed that the rivers in western Taiwan were parts of the Minjiang River and rivers in Hainan belonged to the Red River during sea-level lowering (Figure 1). Thus, strictly freshwater fish can disperse from mainland Asia to these islands via the paleo-connected rivers (Carvajal-Quintero et al. 2019; Leroy et al. 2019). Besides, our result showed that the insular freshwater fish faunas were structured by climate ND (Figure 4). This finding is consistent with a regional-scale floristic study on the taxonomic and phylogenetic turnover of species between Hainan Island and neighboring landmasses (Lin et al. 2021). Studies on the Thousand Island lake system (Si et al. 2017) and the Galápagos Islands (Carvajal-Endara et al. 2017) also identified importance of environmental filtering in structuring species assemblages, which corroborates our findings.

Notably, our results did not show a strong effect of diversification on the insular species assemblages (Figure 4), regardless of the size of species pools, or between standard and phylogenetic logistic regressions (Supplementary Figures S1 and S2). This is unexpected because many studies have documented rapid clade radiations on island systems owing to the increased probability of allopatric divergence (Kier et al. 2009; Losos and Ricklefs 2009). One potential explanation is that when the paleo-connected drainage basins are submerged by rising sea levels, both watersheds of mainland and islands undergo fragmentation and experience similar geographical isolation. As a result, the species on Taiwan and Hainan Islands do not necessarily have a higher diversification rate than that of their mainland congeners. An alternative explanation is that it is simply the effects of ecological and biogeographical processes, such as HC and environmental filtering, that are strong enough to obscure the influence of diversification.

Surprisingly, fish BL had a positive effect on species occurrence in Hainan, despite being marginally significant (Figure 4), thereby indicating that freshwater fish species on Hainan, on average, have a

larger BL than that of their mainland congeners. This contradicts our prediction that insular species might dwarf owing to islands' shorter river lengths and steeper channel gradients. Indeed, our initial prediction ignores that paleo-rivers often have large flow volumes and facilitate the dispersal of species with large body sizes. In contrast, species with small body sizes favor the headwaters within the mountainous regions of mainland Asia and are isolated by larger mainstem channels (Brown and Swan 2010), which decreases their opportunities to colonize Hainan Island. This explanation is supported by a faunistic analysis, which suggests that the majority of freshwater fish on Hainan (approximately 90%) are sourced from plain rivers of mainland Asia and only 10% of species originated from montane streams (Pearl River Fisheries Research Institute 1986). A previous study in mountainous areas of Hainan also suggested that freshwater fishes are concentrated in low- to mid-elevation streams and are rarely encountered above 600 m asl (Chan and Chen 2008), which supports this pattern.

Our study had several limitations that deserve explicit mention. First, the different sizes of species pools might have influenced our findings. Although this study used distance-based clustering analysis to identify regional source pools, some discrepancies emerged among different pool sizes (Figure 3 and Supplementary Figures S1 and S2). Nevertheless, identifying the species source pool accurately for a specific biota is often challenging; thus, it needs to be determined by different methods and criteria in future studies (Carstensen et al. 2013). Second, this study did not fully cover all aspects of drivers underlying the occurrence of insular species. For example, a stronger dispersal capacity of freshwater fish has been proposed to increase their travel distances, thereby increasing their geographical ranges (Lester et al. 2007; Carvajal-Quintero et al. 2019). However, our analysis failed to include the dispersal capacity of freshwater fish as a covariable because this information is still lacking for most fishes (Brosse et al. 2021). Finally, we did not use comprehensive data on species morphological traits to test the community assembly processes, but we instead assessed the community

structure under a phylogenetic framework. Although this method has been validated by numerous studies at large spatial scales (Cardillo et al. 2008; Weigelt et al. 2015; Qian et al. 2020), the community assembly processes should ideally be assessed by integrating the functional and phylogenetic structures (Sobral and Cianciaruso 2016; Ramm et al. 2018). Therefore, future studies should incorporate phylogenetic information and comprehensive morphological trait data (Brosse et al. 2021) to better understand the forces underlying insular community assembly.

Author Contributions

J.H. and H.J. designed the research. H.L., C.D., H.Y., J.T., and J.Y. collected the data. H.L. and J.H. performed the analyses with the help of H.J. H.L., J.H., and H.J. analyzed the results and wrote the manuscript. All authors contributed to the final version of the manuscript. J.H. was responsible for the communication during the submission process.

Acknowledgments

The idea presented in this paper was developed during inspiring discussions with H. Zhu. We thank him and our colleagues X. Liu, S. L. Lin, and Y. Xu for valuable discussions and comments that substantially improved this manuscript. We are grateful to the Handling Editor Z. Jia and 2 anonymous reviewers for their constructive comments on the manuscript. J.H. was supported by grants from the National Natural Science Foundation of China (grant no. 31900324), Guangdong Basic and Applied Basic Research Foundation (grant no. 2020A1515011472), and Star-up Fund from South China Normal University.

Data Accessibility

The freshwater fish data used in this study are available from the [Supplementary Material](#).

Supplementary Material

[Supplementary material](#) can be found at <https://academic.oup.com/cz>.

Conflict of Interest Statement

The authors declare that they have no competing interests.

References

- Ali JR, 2017. Islands as biological substrates: Classification of the biological assemblage components and the physical island types. *J Biogeogr* 44: 984–994.
- Ali JR, 2018. New explanation for elements of Hainan Island's biological assemblage may stretch things a little too far. *Ecography* 41:457–460.
- Bartoń K, 2019. MUMIn: multi-model inference. R package version 1.43.15. Available from: <https://CRAN.R-project.org/package=MUMIn>
- Baselga A, Orme CDL, 2012. Betapart: An R package for the study of beta diversity. *Methods Ecol Evol* 3:808–812.
- Bellard C, Leclerc C, Courchamp F, 2014. Impact of sea level rise on the 10 insular biodiversity hotspots. *Glob Ecol Biogeogr* 23:203–212.
- Benitez-Lopez A, Santini L, Gallego-Zamorano J, Mila B, Walkden P et al., 2021. The island rule explains consistent patterns of body size evolution in terrestrial vertebrates. *Nat Ecol Evol* 5:768–786.
- Blanchet S, Grenouillet G, Beauchard O, Tedesco PA, Leprieur F et al., 2010. Non-native species disrupt the worldwide patterns of freshwater fish body size: implications for Bergmann's rule. *Ecol Lett* 13:421–431.
- Blonder B, Harris DJ, 2019. hypervolume: High dimensional geometry and set operations using kernel density estimation, support vector machines, and convex hulls. R package version 2.0.0. Available from: <https://CRAN.R-project.org/package=hypervolume>
- Brosse S, Charpin N, Su G, Toussaint A, Herrera-R GA et al., 2021. FISHMORPH: A global database on morphological traits of freshwater fishes. *Glob Ecol Biogeogr* 30:2330–2336.
- Brown BL, Swan CM, 2010. Dendritic network structure constrains metacommunity properties in riverine ecosystems. *J Anim Ecol* 79:571–580.
- Burnham KP, Anderson DR, 2002. *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. New York: Springer.
- Burns KC, 2005. A multi-scale test for dispersal filters in an island plant community. *Ecography* 28:552–560.
- Campbell Grant EH, Lowe WH, Fagan WF, 2007. Living in the branches: Population dynamics and ecological processes in dendritic networks. *Ecol Lett* 10:165–175.
- Cardillo M, Gittleman JL, Purvis A, 2008. Global patterns in the phylogenetic structure of island mammal assemblages. *Proc R Soc B* 275:1549–1556.
- Carstensen DW, Lessard J-P, Holt BG, Krabbe Borregaard M, Rahbek C, 2013. Introducing the biogeographic species pool. *Ecography* 36: 1310–1318.
- Carvajal-Endara S, Hendry AP, Emery NC, Davies TJ, 2017. Habitat filtering not dispersal limitation shapes oceanic island floras: Species assembly of the Galápagos archipelago. *Ecol Lett* 20:495–504.
- Carvajal-Quintero J, Villalobos F, Oberdorff T, Grenouillet G, Brosse S et al., 2019. Drainage network position and historical connectivity explain global patterns in freshwater fishes' range size. *Proc Natl Acad Sci USA* 116: 13434–13439.
- Chan BPL, Chen X, 2008. Species diversity and distribution of freshwater fishes at Mt. Yinggeling, Hainan Island, China. *Biodiversity Science* 16:44–52.
- Chen D, Chang J, Li SH, Liu Y, Liang W et al., 2015. Was the exposed continental shelf a long-distance colonization route in the ice age? The southeast Asia origin of Hainan and Taiwan partridges. *Mol Phylogenet Evol* 83: 167–173.
- Chen IS, 2008. The taxonomic status of inland-water fishes of Taiwan. In: Shao KT, Peng CI, Wu WJ, editors. *2008 Taiwan Species Diversity-I Recent Study*. Taipei: Forestry Bureau, Council of Agriculture, Executive Yuan. 235–248.
- Chen Y, Yang X, Li D, Long W, 2016. Status of vascular plant species on Hainan Island. *Biodiv Sci* 24:948–956.
- Chiang T-Y, Lin H-D, Shao K-T, Hsu K-C, 2010. Multiple factors have shaped the phylogeography of Chinese spiny loach *Cobitis sinensis* in Taiwan as inferred from mitochondrial DNA variation. *J Fish Biol* 76: 1173–1189.
- Daru BH, Karunaratne P, Schliep K, Silvestro D, 2020. Phyloregion: R package for biogeographical regionalization and macroecology. *Methods Ecol Evol* 11:1483–1491.
- Darwin C, 1859. *On the Origin of Species*. London: John Murray.
- Dias MS, Oberdorff T, Huguency B, Leprieur F, Jezequel C et al., 2014. Global imprint of historical connectivity on freshwater fish biodiversity. *Ecol Lett* 17:1130–1140.
- Feng G, Mao L, Sandel B, Swenson NG, Svenning J-C, 2016. High plant endemism in China is partially linked to reduced glacial-interglacial climate change. *J Biogeogr* 43:145–154.
- Fick SE, Hijmans RJ, 2017. Worldclim 2: New 1-km spatial resolution climate surfaces for global land areas. *Int J Climatol* 37:4302–4315.
- Gallardo B, Bogan AE, Harun S, Jainih L, Lopes-Lima M et al., 2018. Current and future effects of global change on a hotspot's freshwater diversity. *Sci Total Environ* 635:750–760.
- Gillespie RG, Roderick GK, 2002. Arthropods on islands: Colonization, speciation, and conservation. *Annu Rev Entomol* 47:595–632.
- Graham CH, Carnaval AC, Cadena CD, Zamudio KR, Roberts TE et al., 2014. The origin and maintenance of montane diversity: Integrating evolutionary and ecological processes. *Ecography* 37:711–719.

- Graham CH, Fine PV, 2008. Phylogenetic beta diversity: Linking ecological and evolutionary processes across space in time. *Ecol Lett* 11:1265–1277.
- Griffiths D, 2012. Body size distributions in North American freshwater fish: Large-scale factors. *Glob Ecol Biogeogr* 21:383–392.
- Griffiths D, 2017. Connectivity and vagility determine beta diversity and nestedness in North American and European freshwater fish. *J Biogeogr* 44:1723–1733.
- Guimarães TFR, Petry AC, Hartz SM, Becker FG, 2021. Influence of past and current factors on the beta diversity of coastal lagoon fish communities in South America. *J Biogeogr* 48:639–649.
- He J, Gao Z, Su Y, Lin S, Jiang H, 2018. Geographical and temporal origins of terrestrial vertebrates endemic to Taiwan. *J Biogeogr* 45:2458–2470.
- Ho L, Ané C, 2014. A linear-time algorithm for Gaussian and non-Gaussian trait evolution models. *Syst Biol* 63:397–408.
- Huang Y, Guo X, Ho SYW, Shi H, Li J et al., 2013. Diversification and demography of the oriental garden lizard *Calotes versicolor* on Hainan Island and the adjacent mainland. *PLoS ONE* 8:e64754.
- Ives AR, Garland T Jr, 2010. Phylogenetic logistic regression for binary dependent variables. *Syst Biol* 59:9–26.
- Ives AR, Li D, 2018. rr2: An R package to calculate r2s for regression models. *J Open Source Softw* 3:1028.
- Jing M, Yu HT, Wu SH, Wang W, Zheng X, 2007. Phylogenetic relationships in genus *niviventer* (Rodentia: Muridae) in China inferred from complete mitochondrial cytochrome *b* gene. *Mol Phylogenet Evol* 44:521–529.
- Kang B, Deng J, Wu Y, Chen L, Zhang J et al., 2014. Mapping China's freshwater fishes: Diversity and biogeography. *Fish Fish* 15:209–230.
- Kier G, Kreft H, Lee TM, Jetz W, Ibsch PL et al., 2009. A global assessment of endemism and species richness across island and mainland regions. *Proc Natl Acad Sci USA* 106:9322–9327.
- Koleff P, Gaston KJ, Lennon JJ, 2003. Measuring beta diversity for presence-absence data. *J Anim Ecol* 72:367–382.
- Kreft H, Jetz W, 2010. A framework for delineating biogeographical regions based on species distributions. *J Biogeogr* 37:2029–2053.
- Kubota Y, Hirao T, Fujii S-J, Shiono T, Kusumoto B, 2014. Beta diversity of woody plants in the Japanese archipelago: The roles of geohistorical and ecological processes. *J Biogeogr* 41:1267–1276.
- Leclerc C, Villéger S, Marino C, Bellard C, Di ME, 2020. Global changes threaten functional and taxonomic diversity of insular species worldwide. *Divers Distrib* 26:402–414.
- Leprieur F, Albouy C, De Bortoli J, Cowman PF, Bellwood DR et al., 2012. Quantifying phylogenetic beta diversity: Distinguishing between “true” turnover of lineages and phylogenetic diversity gradients. *PLoS ONE* 7:e42760.
- Leprieur F, Tedesco PA, Hugué B, Beauchard O, Durr HH et al., 2011. Partitioning global patterns of freshwater fish beta diversity reveals contrasting signatures of past climate changes. *Ecol Lett* 14:325–334.
- Leroy B, Dias MS, Giraud E, Hugué B, Jézéquel C et al., 2019. Global biogeographical regions of freshwater fish species. *J Biogeogr* 46:2407–2419.
- Lester SE, Ruttenberg BI, Gaines SD, Kinlan BP, 2007. The relationship between dispersal ability and geographic range size. *Ecol Lett* 10:745–758.
- Li J, Fu C, Lei G, 2011. Biogeographical consequences of Cenozoic tectonic events within East Asian margins: A case study of *Hynobius* biogeography. *PLoS ONE* 6:e21506.
- Li SH, Li JW, Han LX, Yao CT, Shi H et al., 2006. Species delimitation in the Hwamei *Garrulax canorus*. *Ibis* 148:698–706.
- Lin S, Chen L, Peng W, Yu J, He J et al., 2021. Temperature and historical land connectivity jointly shape the floristic relationship between Hainan Island and the neighbouring landmasses. *Sci Total Environ* 769:144629.
- Lomolino MV, 2005. Body size evolution in insular vertebrates: generality of the island rule. *J Biogeogr* 32:1683–1699.
- Losos JB, Ricklefs RE, 2009. Adaptation and diversification on islands. *Nature* 457:830–836.
- MacArthur RH, Wilson EO, 1967. *The Theory of Island Biogeography*. Princeton: Princeton University Press.
- Mays HL Jr, McKay BD, Tietze DT, Yao C-T, Miller LN et al., 2015. A multi-locus molecular phylogeny for the avian genus *Liocichla* (Passeriformes: Leiothrichidae: Liocichla). *Avian Res* 6:17.
- Meiri S, Cooper N, Purvis A, 2008. The island rule: made to be broken?. *Proceedings. Biological Sciences* 275: 141–148. 10.1098/rspb.2007.1056 17986433
- Nekola JC, White PS, 1999. The distance decay of similarity in biogeography and ecology. *J Biogeogr* 26:867–878.
- Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P et al., 2019. *vegan*: Community ecology package. R package version 2.5-6. Available from: <https://CRAN.R-project.org/package=vegan>.
- Päckert M, Martens J, Severinghaus LL, 2009. The Taiwan firecrest *Regulus goodfellowi* belongs to the Goldcrest assemblage (*Regulus regulus* s. l.): Evidence from mitochondrial DNA and the territorial song of the Regulidae. *J Ornithol* 150:205–220.
- Qian H, Cao Y, Li D, Chu C, Sandel B et al., 2020. Geographic patterns and environmental correlates of phylogenetic relatedness and diversity for freshwater fish assemblages in North America. *Ecography* 43:1814–1824.
- Qiu Y-X, Fu C-X, Comes HP, 2011. Plant molecular phylogeography in China and adjacent regions: Tracing the genetic imprints of Quaternary climate and environmental change in the world's most diverse temperate flora. *Mol Phylogenet Evol* 59:225–244.
- Qu Y, Song G, Gao B, Quan Q, Ericson PGP et al., 2015. The influence of geological events on the endemism of East Asian birds studied through comparative phylogeography. *J Biogeogr* 42:179–192.
- Rabosky DL, Chang J, Title PO, Cowman PF, Sallan L et al., 2018. An inverse latitudinal gradient in speciation rate for marine fishes. *Nature* 559: 392–395.
- Ramm T, Cantalapiedra JL, Wagner P, Penner J, Rodel MO et al., 2018. Divergent trends in functional and phylogenetic structure in reptile communities across Africa. *Nat Commun* 9:4697.
- Rinaldo A, Rigon R, Banavar JR, Maritan A, Rodriguez-Iturbe I, 2014. Evolution and selection of river networks: Statics, dynamics, and complexity. *Proc Natl Acad Sci USA* 111:2417–2424.
- Shurin JB, Cottenie K, Hillebrand H, 2009. Spatial autocorrelation and dispersal limitation in freshwater organisms. *Oecologia* 159:151–159.
- Si X, Cadotte MW, Zeng D, Baselga A, Zhao Y et al., 2017. Functional and phylogenetic structure of island bird communities. *J Anim Ecol* 86: 532–542.
- Sobral FL, Cianciaruso MV, 2016. Functional and phylogenetic structure of forest and savanna bird assemblages across spatial scales. *Ecography* 39: 533–541.
- Tzeng C-S, 1986. Distribution of the freshwater fishes of Taiwan. *J Taiwan Mus* 39:127–146.
- Verdú M, 2002. Age at maturity and diversification in woody angiosperms. *Evolution* 56:1352–1361.
- Wallace AR, 1869. *The Malay Archipelago: The Land of the Orang-Utan and the Bird of Paradise; a Narrative of Travel, with Studies of Man and Nature*. New York: Harper.
- Wang J, Li C, Chen J, Wang J, Jin J et al., 2021. Phylogeographic structure of the dwarf snakehead *Channa gachua* around Gulf of Tonkin: Historical biogeography and pronounced effects of sea-level changes. *Ecol Evol* 11: 12583–12595.
- Wang T-Y, Liao T-Y, Tzen C-S, 2007. Phylogeography of the Taiwanese endemic hillstream loaches, *Hemimyzon formosanus* and *H. taitungensis* (Cypriniformes: Balitoridae). *Zool Stud* 46:547–560.
- Webb CO, Ackerly DD, McPeck MA, Donoghue MJ, 2002. Phylogenies and community ecology. *Annu Rev Ecol Syst* 33:475–505.
- Weigelt P, Jetz W, Kreft H, 2013. Bioclimatic and physical characterization of the world's islands. *Proc Natl Acad Sci USA* 110:15307–15312.
- Weigelt P, Kissling WD, Kisel Y, Fritz SA, Karger DN et al., 2015. Global patterns and drivers of phylogenetic structure in island floras. *Sci Rep* 5:12213.
- Xiong W, Wang Q, Xie D, Fletcher DH, He D, 2018. Factors influencing tropical island freshwater fishes: Species, status, threats and conservation in Hainan Island. *Knowl Manag Aquat Ecosyst* 6.