Shin *et al. BMC Ecology and Evolution* (2025) 25:105
https://doi.org/10.1186/s12862-025-02442-3

BMC Ecology and Evolution

RESEARCH **Open Access**

Climatic data sources and limitations of ecological niche models impact the estimations of historical ranges and niche overlaps in distantly related Korean salamanders

Yucheol Shin[1,2,3], Amaël Borzée[4,5] and Daesik Park[6*]

Abstract

Background Ecological niche models (ENMs) and analyses of niche overlap/divergence have become popular methods in ecology and evolutionary biology. These analyses rely on environmental data available from several databases. However, the influence of data sources on these analyses is rarely tested. Here, we test the impact of climatic data choice on the prediction of current and Plio-Pleistocene suitable habitats for two distantly related, but broadly sympatric, salamanders endemic to the Korean Peninsula. We ran MaxEnt separately on WorldClim and CHELSA climate data. We then hindcasted ENMs to five time periods of the Plio-Pleistocene, bracketing the estimated intraspecific divergence times for these species. We then quantified the differences in predictions between WorldClim- and CHELSA-based models. Also, given the sympatry and similar habitat requirements of the two species, we tested for niche overlaps using niche identity and background tests and tested the sensitivity of the results to climatic data choice.

Results The ENMs successfully predicted contemporary suitable habitats for the two species. However, the predictions were highly sensitive to climatic data choice as well as variable combinations. The hindcasted ENMs produced contrasting predictions depending on the choice of climatic dataset and failed to predict suitable habitats for some Pleistocene time periods regardless of the climatic data choice. The niche analyses were also sensitive to climatic data choice, with results suggesting either niche overlaps or divergence depending on the climatic dataset used for the analyses.

Conclusions Our study highlights the influence of climatic data choice on the outcomes of ENMs and niche analyses. Our results also underscore the limitations of macroclimate-based ENMs, especially when the species is likely buffered from macroclimatic changes by microhabitat. We argue for the need for additional ecological,

*Correspondence:
Daesik Park
parkda@kangwon.ac.kr

Full list of author information is available at the end of the article

ecophysiological, and population genomic studies to better understand the range formation of these enigmatic species.

Keywords Ecological niche modeling, Korean peninsula, Macroclimate, MaxEnt, Niche analyses, Salamander

Background

Correlative ecological niche models (ENMs) and analyses of niche overlap/divergence have become hugely popular methods in ecology and evolutionary biology. With increasing data availability and computational advancements, these methods have been widely applied to answer questions in ecology, biogeography, evolution, and conservation biology [1–5]. Both of these methods, as implemented in popular software packages, generally require the following data types: geographic coordinates of species occurrence records, coordinates for absence or background data, and environmental variables. Most of these data can be obtained from public data sources or sampled as part of the analyses. For example, environmental data, and especially climatic data, for these analyses are generally obtained from one of several popular public databases, including WorldClim [6], CHELSA [7], and ENVIREM [8].

Numerous studies investigated the effects of occurrence and absence/background point sampling and model complexity on the model prediction outcomes [9–14]. Relatively less attention has been given, however, to the prediction uncertainties associated with the choice of climatic data sources [15–17]. Different climatic databases are generated using different methodological approaches. The choice of climatic data, therefore, can introduce significant uncertainties into model outcomes and downstream interpretations [17]. The implications of these uncertainties become greater when the models are extrapolated to different environmental conditions across time and space. Most studies select climatic data from one of many available databases, but justifications for selection are rarely provided. Given that one of the primary applications of ecological niche modeling and niche overlap/divergence analyses is spatiotemporal projections and the study of niche shifts, the influence of climatic data choice on the outcomes of these analyses needs to be better understood.

On the other hand, the resolution of available data and species-specific ecology can result in genuine limitations of ecological niche models to predict species distributions. For example, macroclimatic variables available from databases such as WorldClim or CHELSA may not be suitable to model suitable habitats of small, forest-dwelling species that are buffered from macroclimatic fluctuations and are likely more sensitive to microclimatic changes [18]. While mechanistic niche modeling [19, 20] can produce more realistic predictions for these species by accounting for ecophysiological factors and microclimate, this method is usually more data-intensive and, therefore, not applicable to many species that are in a juxtaposition of poorly known ecology and lack of data.

In this study, we investigate the influence of climatic data choice and limitations of ENMs in predicting the current and past distributions of two distantly related lungless salamander species endemic to the Korean Peninsula: the Korean Clawed Salamander (*Onychodactylus koreanus*) and the Korean Crevice Salamander (*Karsenia koreana*). *Onychodactylus koreanus* is a hynobiid salamander with a biphasic life cycle and a prolonged aquatic larval stage [21]. Both larvae and adults are strict habitat specialists of forests and mountain streams [22–24]. *Karsenia koreana* is the only known Asian representative of the family Plethodontidae [25], and unlike *O. koreanus*, this species is fully terrestrial and has direct-developing eggs [26]. It is also found strictly in forested areas adjacent to mountain streams [27].

While the geographic distributions of both species are not fully characterized within the Korean Peninsula [28], the two species occur in broad sympatry from the central to southern part of the peninsula along a major mountain range (Fig. 1). Based on previous mtDNA-based phylogeographic studies, the current range of *O. koreanus* was most likely formed from a historical southward dispersal of an ancestral lineage to the Korean Peninsula [23]. This was likely followed by rapid diversification and isolation along the mountains of the Korean Peninsula between 1.6 and 2.7 Ma [23]. For *K. koreana*, the dispersal of ancestral lineages from North America through the Bering land bridge around 65 Ma has been inferred based on phylogenetic analyses [29, 30]. Following the Pleistocene glacial period, the populations that persisted in the southern part of the Korean Peninsula expanded their ranges northward, followed by the isolation of genetically distinct regional populations along the mountains of the Korean Peninsula. Most of the intraspecific divergence in this species occurred within 1 Ma [31].

Given the physiological similarity, distant evolutionary relationships, broad sympatry, and apparent overlaps in habitat use, understanding the processes of range formation in these species carries substantial biogeographical and ecological significance. In this context, it is crucial to understand the impacts of past climatic shifts on the range formation processes of these species as both species are lungless and likely sensitive to climate change [24, 32]. This is especially relevant considering the influence of Quaternary climatic oscillations on current biodiversity [33]. Furthermore, geographic distributions

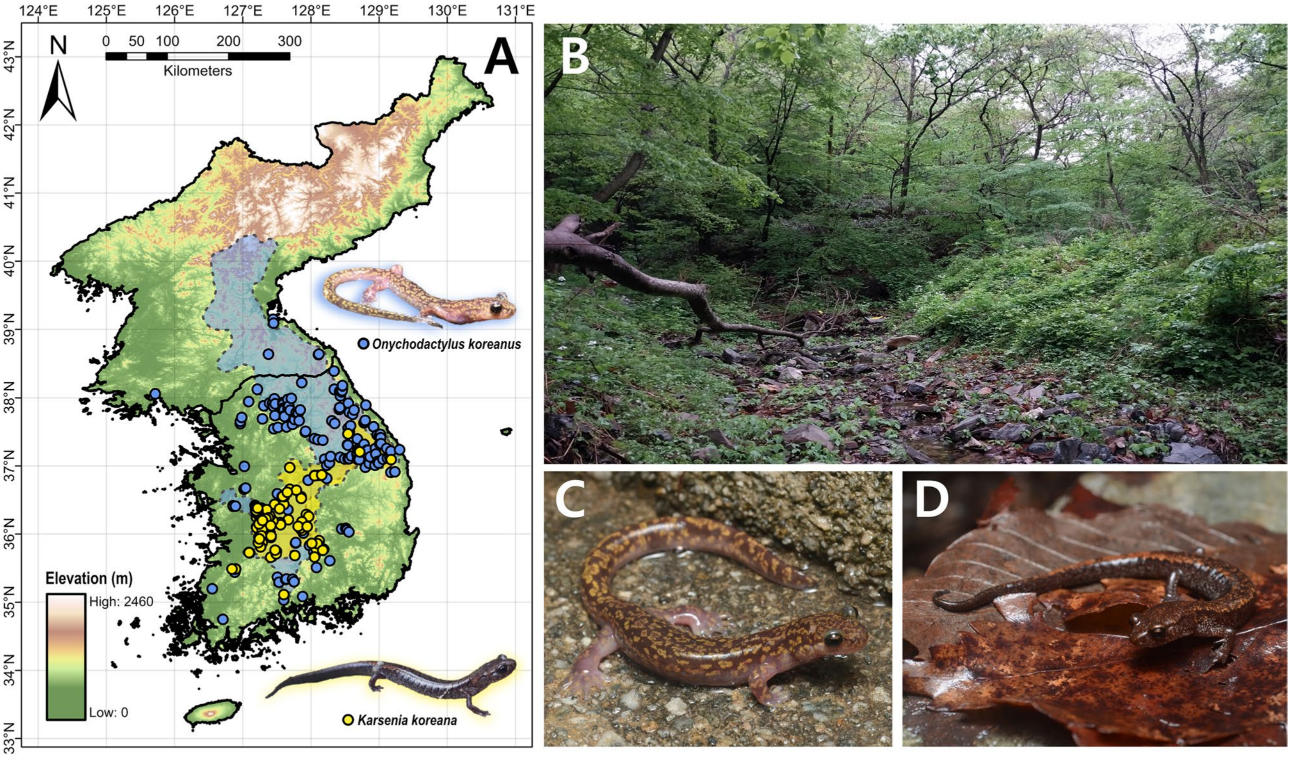

Fig. 1 Geographic distribution and habitat of the Korean Clawed Salamander (*Onychodactylus koreanus*) and Korean crevice salamander (*Karsenia koreana*). **A** The known geographic distributions of *O. koreanus* (blue dots) and *K. koreana* (yellow dots) across the Korean Peninsula. Note the overlapping ranges of the two species in the central and southern parts of the peninsula. The geographic range of each species based on Borzée et al. (2024) [42] is shown with polygons corresponding to the color of occurrence points. Note that the range polygons used here do not provide a complete representation of species ranges. **B** A representative forest habitat in the Republic of Korea, where both *O. koreanus* (**C**) and *K. koreana* (**D**) occur. Both species can be found along the streams and adjacent forest floors with no apparent segregation in habitat types. Photographs in (**B**) – (**D**) and the map inset image of *O. koreanus* were taken by Yucheol Shin, and the map inset image of *K. koreana* was taken by Amaël Borzée

and ecological niches are usually compared between sister species or closely related species within the same genus. These species, therefore, provide rare and interesting opportunities to investigate responses of ecologically and physiologically similar but distantly related species to common climatic fluctuations.

While ENMs can complement phylogeographic studies by visualizing the shifts of potentially suitable habitats over time, and niche analyses can provide insights into contemporary niche overlaps given broad sympatry, the influence of climatic data choice and associated uncertainties should also be evaluated to avoid spurious conclusions. Therefore, our study aims to address the following questions: (1) Can current environmental datasets sufficiently predict the contemporary distributions of *O. koreanus* and *K. koreana* given known distributions? (2) Can contemporary ENMs predict suitable habitats during the Quaternary, consistent with the population demography estimated from genetic data? (3) Do these species show patterns of niche overlaps given broad sympatry and similar habitat requirements? (4) What are the influences of climatic data choice on the results of these analyses?

Methods

Data Preparation

To generate ENMs for the two study species under current environmental conditions, we obtained spatially rarefied occurrence dataset of *O. koreanus* from a previous study ($n = 187$; [24]), which was initially compiled from survey records, Global Biodiversity Information Facility (GBIF), VertNet, and natural history museum data. We compiled occurrence records of *K. koreana* from the results of National Ecosystem Surveys (NES; accessed via the Ecobank platform; [34]), a previous study by Jeon et al. (2021) [31], and the GBIF (via the *megaSDM* package; [35]), resulting in a total of 362 occurrence points. We spatially rarefied this dataset to 137 occurrence points using the “thinData” function of the *SDMtune* package [36] in R v4.2.2 [37]. We note that our occurrence datasets compiled data from different biodiversity databases and surveys, each with different data collection methods, biases, and temporal and spatial coverages. For example, the NES are standardized surveys conducted in specified surveys grids, whereas the GBIF data incorporates various data types including non-standardized citizen science observations and museum records. However, we justify the use of these datasets as this provided the most comprehensive spatial coverage of occurrence points within the known distributions of the two species. We further compensated for the spatial sampling bias of occurrence points using a modified background sampling method (see below).

Although not fully characterized, the geographic distributions of *O. koreanus* and *K. koreana* are likely to extend into the Democratic People's Republic of Korea (DPR Korea) based on the continuous distribution of mountain ranges and forested habitats [28]. However, the general paucity of reliable occurrence records for these species from DPR Korea and the relative abundance of occurrence points from the Republic of Korea (R Korea) introduce a significant sampling bias towards R Korea that cannot be compensated by the spatially thinning of occurrence points alone. Therefore, we modified the sampling of background points considering two levels of spatial sampling bias in the occurrence datasets. The first level (background set 1; BG1 hereafter) was sampled from a kernel density surface generated from pooled occurrence points of amphibians recorded from across the Korean Peninsula, representing the overall sampling effort for amphibians across the study area (see [24] for specific methods). The amphibian occurrences were obtained from GBIF, NES, and Borzée et al. (2021) [28]. The second level (BG2) was sampled from a kernel density surface generated from pooled occurrences of *O. koreanus* and *K. koreana*, representing spatial sampling effort constrained specifically to the two study species. The generation of kernel density surfaces was conducted using the R package *MASS* [38]. For each level of spatial bias, we sampled three background sets with different sample sizes ($n = 5{,}000$, $n = 10{,}000$, $n = 15{,}000$). Thus, we used a total of six different background datasets (2 levels of spatial bias × 3 different sample sizes) for model testing.

For environmental data, we initially considered 22 variables: 19 bioclimatic variables from WorldClim (https://www.worldclim.org/; [6]), elevation, slope, and global consensus raster layers for needleleaf forest, deciduous broadleaf forest, and mixed/other forest types obtained from the EarthEnv database (https://www.earthenv.org/; [39]). We note that slope is a highly scale-dependent variable and that salamanders may have different preferences for finer-scale slopes. Therefore, our application of the slope layer here is to capture the general, larger-scale habitat characteristics of these salamanders rather than capturing finer scale habitat conditions. Also, because these salamanders are generally associated with heavily forested habitats, we merged the three forest cover layers with the "Mosaic to New Raster" tool in ArcGIS Pro v2.6.0 (ESRI, Redlands, CA), using the "Max" mosaic operator to assign a maximum percent forest cover value among overlapping forest cover layers. All raster layers were in 1 km spatial resolution (= 0.008333 dd) and were masked to the geographic extent of the Korean Peninsula using the *raster* R package [40]. We used the boundary of the Korean Peninsula to define the model calibration range because the distributions of these two species fall within this area (e.g., regions in northeastern China adjacent to DPR Korea are inhabited by other species of *Onychodactylus* and no other Asian plethodontid has been reported outside of Korean Peninsula; [41, 42]). This area thus represents the range of environmental conditions suitable for these species. In addition, using alternative approaches such as alpha or convex hulls will result in smaller calibration areas that cannot fully capture the range of environmental conditions available to these species.

From this set of 22 variables, we first removed five bioclimatic variables not represented in the paleoclimatic dataset used for hindcasting. These variables were mean diurnal range (bio2), isothermality (bio3), maximum temperature of the warmest month (bio5), minimum temperature of the coldest month (bio6), and temperature annual range (bio7). We then conducted a Pearson's correlation test on the remaining 17 variables and removed highly correlated variables with $| r | > 0.8$. This resulted in the following set of eight environmental variables: annual mean temperature (bio1), temperature seasonality (bio4), annual precipitation (bio12), precipitation of the wettest month (bio13), precipitation of the driest month (bio14), precipitation seasonality (bio15), forest cover, and slope. Considering the known ecology of the species, we considered the selection of these variables to be appropriate. The variable selection step was conducted with the R package *ntbox* [43].

Model development

To estimate the current habitat suitability of *O. koreanus* and *K. koreana*, we used the maximum entropy (MaxEnt; [44]) algorithm and conducted extensive model testing using the *ENMeval* R package (version 2.0; [45]). One of the primary goals of this study was to transfer the current model to historical climate conditions. However, no matching topographic and vegetation raster data were available for the time periods of model transfer and our study area, and keeping the non-climatic variables constant across model transfers was deemed unrealistic. Therefore, we first generated current ENMs for both species using only the bioclimatic variables ("climate-only model" hereafter).

To do so, we considered six bioclimatic variables with low multicollinearity and six background datasets (BG1, BG2, each with three different sample sizes). For each combination of climatic and background sets, we tested combinations of 13 MaxEnt feature classes (L, Q, H, P,

LQ, LP, QH, QP, HP, LQH, LQP, LQHP, LQHPT; where L = Linear, Q = Quadratic, H = Hinge, P = Product) and 10 regularization multipliers ranging from 0.5 to 5 at a 0.5 increment. We evaluated the candidate models using a 4-fold hierarchical spatial checkerboard cross-validation, implemented in the R package *ENMeval* with the "checkerboard2" method [45]. We selected this data partitioning scheme as spatial cross-validation methods are generally recommended for model transfer [46, 47] and because it produced a generally similar number of data points and environmental representations across the four cross-validation folds. All other data partitioning strategies, such as the 4-fold spatial blocks, random *k*-fold partitioning, and user-specified spatial blocks, produced irregular allocation of data points and/or irregular environmental representation across the cross-validation folds. Furthermore, clamping was used in all model runs to limit extrapolation beyond the range of calibration dataset.

Thus, we tested a total of 3120 models for each species (1 modeling algorithm x 2 levels of spatial bias correction x 3 sets of background points x 4 cross-validation runs x 13 MaxEnt feature combinations x 10 regularization values). While ΔAICc is commonly used as a model selection criterion [48, 49], preliminary model selection using this criterion resulted in the selection of models with excessively high omission rates. Therefore, we selected the optimal model for each species in a sequential fashion by applying the lowest 10% omission rate (OR_{10}) as a primary filtering criterion, the lowest AUC_{DIFF} (difference between the area under the curve [AUC] values calculated from training and testing data [AUC_{TRAIN} - AUC_{TEST}]; [50]) as a secondary filtering criterion, and the highest AUC_{TEST} as a tertiary filtering criterion, thereby minimizing omission and overfitting while maximizing predictive performance [51, 52]. Although the use of AUC as an absolute measure of model performance has been criticized for presence-background ENMs [53, 54], we used AUC as a relative measure of model performance to compare a suite of candidate ENMs for each species. In addition, we used the Continuous Boyce Index (CBI; [55]) as another measure of model performance. We also visually inspected the output prediction maps to assess the geographical consistency of predictions with the known distributions of the two species.

Based on these criteria, the optimal model for *O. koreanus* was made with BG1 (with $n = 10{,}000$) and LQ features combined with a regularization multiplier of 1.0. For *K. koreana*, the optimal model was made with BG1 (with $n = 10{,}000$) and an LP feature with a regularization multiplier of 5.0. Finally, we tested the selected optimal model for each species against the null ENMs [56] and assessed whether the empirical ENMs significantly outperform the null ENMs based on CBI and AUC_{TEST}. The null ENMs were computed in 1,000 iterations using the "ENMnulls" function of the *ENMeval* package. For the optimal model for each species, we assessed variable importance through permutation importance and percent contribution.

We also generated ENMs for both species by adding two non-climatic variables (slope and forest cover) that are relevant to species ecology to the six bioclimatic variables ("full model" hereafter) to test if the predictions from climate-only models deviated considerably from the full models. All models generated in this study were in complementary log-log (cloglog; [44]) format, and we used the WGS 84 coordinate reference system for all geographic projections. The full details of ENMs can be found in the ODMAP reporting protocol associated with this study ([57, 58]; Supplementary Material 1).

Hindcasting ENMs

We transferred the optimal climate-only ENMs for *O. koreanus* and *K. koreana* to five paleoclimatic models of the Pliocene and Pleistocene, corresponding to the major climatic events bracketing the intraspecific divergence times of the two species estimated from genetic data [23, 31]. These time periods are: mid-Pliocene Warm Period (mPWP; ca. 3.2 Ma), Marine Isotope Stage 19 (MIS19; ca. 787 Ka), Last Interglacial (LIG; ca. 130 Ka), Last Glacial Maximum (LGM; ca. 21 Ka), and Mid-Holocene (MH; ca. 6 Ka). The paleoclimatic models were downloaded from PaleoClim (http://www.paleoclim.org/; [59–61]) at the 5 km spatial resolution (= 0.041666 dd). To match the spatial resolution of current and paleoclimatic data, we statistically downscaled the paleoclimatic layers to 1 km spatial resolution using bilinear interpolation, implemented with the "disaggregate" function of the *raster* package. For each model transfer, we assessed the extrapolation risk through Multivariate Environmental Similarity Surface (MESS; [62]) computed in the R package *ntbox* [43].

Influence of climatic data sources on ENMs

To assess the effect of climatic data choice on the output model predictions [17, 63], we repeated the modeling steps outlined above by using the same set of six bioclimatic variables downloaded from CHELSA (https://chelsa-climate.org/; [7]). We note that conducting a separate Pearson's correlation test on CHELSA bioclimatic variables resulted in the selection of the same six variables as the WorldClim-based selection. First, to directly compare climate values between WorldClim and CHELSA datasets within the geographic ranges of our target species, we extracted raster pixel values of the two datasets separately from the occurrence points of *O. koreanus* and *K. koreana*. For each species, and separately for each bioclimatic variable, we conducted the Mann-Whitney *U*-test to test if climatic values differed between datasets. To spatially visualize the difference between the WorldClim and CHELSA bioclimatic variables, we subtracted the CHELSA rasters from WorldClim rasters. Thus, in these "difference rasters," pixels with positive values indicate higher values in WorldClim layers and pixels with negative values indicate regions where the CHELSA layers had higher values.

Next, we fitted the same optimal model for both species (LQ 1.0 for *O. koreanus* and LP 5.0 for *K. koreana*) but using the CHELSA bioclimatic variables. We generated both climate-only and full models and then spatially visualized the areas in which the WorldClim-based and CHELSA-based models differed in current habitat suitability predictions. We then repeated model transfer with the CHELSA-based climate-only models. To quantify the difference between spatial predictions based on WorldClim and CHELSA data, we followed the method provided by Dubos et al. (2023) [17], using the following equation slightly modified from that study:

$$\frac{|\sum P_{WorldClimj} - \sum P_{CHELSAj}|}{\sum P_{WorldClimj}} \times 100$$

Here, $P_{WorldClimj}$ and $P_{CHELSAj}$ denote the estimated suitability score of a given raster pixel j in predictions based on WorldClim and CHELSA data, respectively. Therefore, this method quantifies the overall percentage difference between a pair of model predictions relative to the baseline prediction (WorldClim-based predictions in this case) across the entire area of projection for a given time period, with a higher level of difference indicating a stronger effect of climate data source in causing the difference in predictions [17].

In addition, following Dubos et al. (2023) [17], we computed the Schoener's D overlap between WorldClim-based and CHELSA-based model predictions to account for spatial information. The Schoener's D index ranges from 0 to 1, where a value closer to 1 indicates higher spatial similarity between a pair of predictions and thus a lower effect of climatic data source. We used the following equation to calculate Schoener's D [64, 65]:

$$D\left(p_x, p_y\right) = 1 - \frac{1}{2}\sum_i |p_{x_i} - p_{y_i}|$$

As originally defined for the application of Schoener's D to ENM predictions, p_{xi} (or p_{yi}) denotes the estimated suitability value of a grid cell i for the ENM prediction of species X (or Y) [65]. In the context of our analysis, X and Y correspond to the prediction based on WorldClim and CHELSA data, respectively. Furthermore, as Schoener's D is typically applied as an index of niche overlap, we converted the calculated D values using $(1 - D) \times 100$ to the index of percent difference between predictions [17]. The spatial overlap between model predictions was calculated using the "raster.overlap" function of the *ENMTools* package [66].

While the use of same model parameters optimized for WorldClim data for CHELSA data is convenient for the purpose of model comparisons, it may lead to suboptimal parameterization for the CHELSA dataset. Therefore, we conducted separate model tuning runs for the CHELSA dataset, using the same occurrence and background datasets, data partitioning scheme, and testing the same number of feature classes and regularization values for each species. The optimal parameters for the CHELSA data were HP feature classes and regularization multiplier of 4 for *O. koreanus* and LQ feature classes with a regularization multiplier of 0.5 for *K. koreana*. We then hindcasted these CHELSA-optimized models to test the influence of CHELSA-specific parameter optimization on model transfer results.

Contemporary niche overlaps

We conducted niche identity test and asymmetric background test to quantify and compare the ecological niches of *O. koreanus* and *K. koreana*. The niche identity test uses randomized pseudoreplicates to generate a null distribution of niche overlap values and compare these values to the observed overlap value [65]. Here, the randomization process is conducted by pooling the occurrence points of a species pair and randomizing the species identity. Next, new sets of occurrence points are sampled from this randomized pool, retaining the original number of occurrence points for each species [65]. The null hypothesis of niche identity is rejected when the test result is statistically significant. On the other hand, in the asymmetric background test, the null distribution is generated by randomly shifting the occurrence points of one species and calculating the niche overlap value with the other species [65]. This null distribution is then compared to the observed value of niche overlap. Therefore, this test compares the niche of one species to the broader environmental background of another species [65]. A significantly higher niche overlap value indicates higher niche similarity than expected. On the other hand, a significantly lower niche overlap value may indicate niche divergence between the species pair. We conducted the background test in two directions, first comparing the niche of *O. koreanus* to the environmental background of *K. koreana* and then comparing the niche of *K. koreana* to the environmental background of *O. koreanus*.

We implemented these analyses in environmental space (E-space; [67, 68]) and within the framework of the Niche Overlap Test (NOT) and Niche Divergence Test (NDT) of Brown & Carnaval (2019) [68] to explicitly consider the spatial distribution and availability of environmental conditions. The NOT represents the identity and background statistics calculated over the total accessible E-space within the geographic ranges of two species, whereas NDT represents the identity and background statistics calculated only within the accessible E-space shared between two species [68]. We conducted these analyses using the R package *humboldt* [68]. As input environmental data, we used eight environmental variables (six climatic and two non-climatic variables) initially selected from Pearson's correlation test (see above), following the recommendations in the *humboldt* package documentation [69]. To test for the sensitivity of results to climatic data choice, we separately conducted NOT and NDT using the bioclimatic variables from WorldClim and CHELSA and used Schoener's *D* [64] as the metric of niche overlap.

Results

Current ENMs of O. koreanus

The climate-only ENMs for *O. koreanus* based on WorldClim data had low model overfitting and adequate predictive performance (CBI = 0.909; AUC_{TEST} = 0.787; AUC_{DIFF} = 0.014; OR_{10} = 0.089; Table 1). In addition, the comparison of climate-only empirical models against null ENMs suggested that the empirical models have significantly higher predictive abilities ($p < 0.001$ for both CBI and AUC_{TEST}; Table 1). The climate-only ENMs based on WorldClim data predicted suitable habitats along the major mountains of the Korean Peninsula. Overall, predictions from the full model did not deviate considerably from the climate-only prediction (Fig. 2A; See Supplementary Fig. 1 A for binary presence/absence maps and Supplementary Table 1 for conversion thresholds). However, the full model predicted suitable habitats for *O. koreanus* in areas with the highest elevation within DPR Korea, while the climate-only model predicted this area to be unsuitable.

The climate-only ENMs for *O. koreanus* based on CHELSA data also had low model overfitting and good predictive performance (CBI = 0.899; AUC_{TEST} = 0.784; AUC_{DIFF} = 0.025; OR_{10} = 0.104; Table 1). The climate-only empirical model significantly outperformed the null ENMs ($p < 0.001$ for both CBI and AUC_{TEST}; Table 1). The climate-only ENMs based on CHELSA data predicted suitable habitats for *O. koreanus* across the major mountains of the Korean Peninsula. However, unlike the predictions from WorldClim-based models, CHELSA-based models predicted the entire geographic extent of DPR Korea to be suitable for *O. koreanus* (Fig. 2A; See Supplementary Fig. 1 A for binary presence/absence maps and Supplementary Table 1 for conversion thresholds). Except for the smaller predicted area of suitable habitats in DPR Korea, the prediction from the CHELSA-based full model did not deviate significantly from the climate-only prediction (Supplementary Fig. 1 A).

For both climate-only models based on WorldClim and CHELSA data, mean annual temperature (bio1) and precipitation of the wettest month (bio13) were the two most important variables. However, slope was the most important variable for the full model based on WorldClim data, and precipitation of the wettest month (bio14) was the second most important variable. For the full model based on CHELSA data, annual mean temperature (bio1) was the most important variable, and slope was the second most important variable based on permutation importance. See Table 2 for the importance of each variable based on permutation importance, and Supplementary Table 2 for variable importance values based on percent contribution.

Current ENMs of K. koreana

The climate-only ENMs for *K. koreana* based on WorldClim data had good predictive performance and a low degree of model overfitting (CBI = 0.771; AUC_{TEST} = 0.837; AUC_{DIFF} = 0.011; OR_{10} = 0.092; Table 1). The climate-only empirical model also performed significantly better than the null ENMs ($p < 0.001$ for both CBI and AUC_{TEST}; Table 1). The climate-only ENMs based on WorldClim data predicted suitable habitats of *K. koreana* to be primarily in the central and southern parts of R Korea, along the major mountains (Fig. 2B; Supplementary Fig. 1B). The suitable habitats predicted from the full model did not deviate considerably from predictions

Table 1 Predictive performance of climate-only empirical and null maxent models for *Onychodactylus koreanus* and *karsenia koreana* based on four evaluation metrics

Species	Data	Model	CBI (Mean ± SD)	AUC_{TEST} (Mean ± SD)	AUC_{DIFF} (Mean ± SD)	OR_{10} (Mean ± SD)
O. koreanus	WorldClim	Empirical	0.909 ± 0.029	0.787 ± 0.011	0.014 ± 0.002	0.089 ± 0.031
O. koreanus	WorldClim	Null	0.056 ± 0.198	0.516 ± 0.038	0.117 ± 0.027	0.160 ± 0.046
O. koreanus	CHELSA	Empirical	0.899 ± 0.041	0.784 ± 0.021	0.025 ± 0.014	0.104 ± 0.030
O. koreanus	CHELSA	Null	0.608 ± 0.241	0.517 ± 0.038	0.113 ± 0.027	0.170 ± 0.044
K. koreana	WorldClim	Empirical	0.771 ± 0.036	0.837 ± 0.009	0.011 ± 0.007	0.092 ± 0.088
K. koreana	WorldClim	Null	0.526 ± 0.248	0.520 ± 0.051	0.137 ± 0.027	0.035 ± 0.044
K. koreana	CHELSA	Empirical	0.839 ± 0.099	0.849 ± 0.023	0.020 ± 0.017	0.109 ± 0.024
K. koreana	CHELSA	Null	0.111 ± 0.218	0.538 ± 0.068	0.182 ± 0.035	0.131 ± 0.091

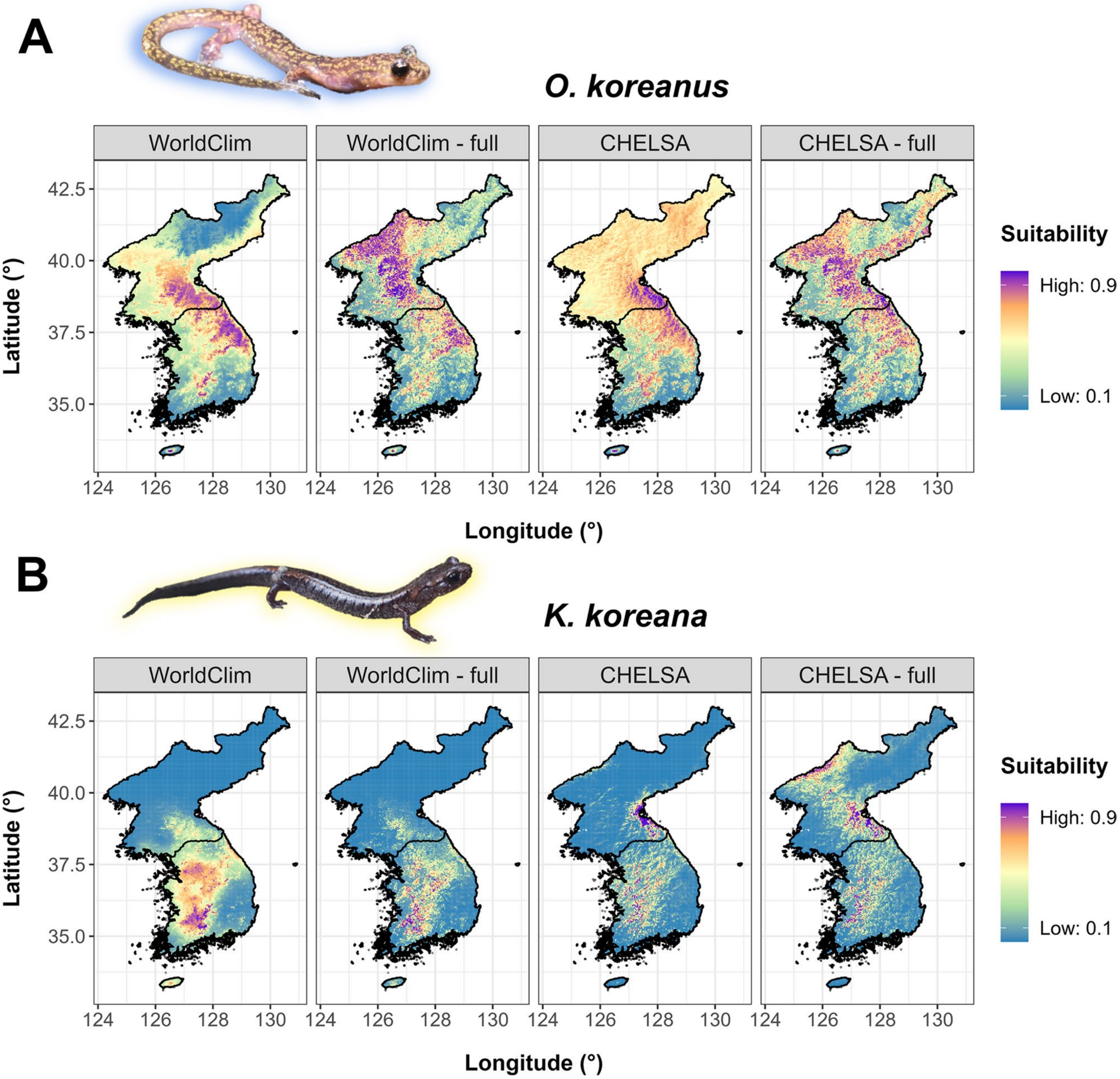

Fig. 2 Spatial predictions of current habitat suitability for *Onychodactylus koreanus* (**A**) and *Karsenia koreana* (**B**) based on different climatic data sources (WorldClim vs. CHELSA) and predictor variable sets. The models for each species were calibrated separately using WorldClim and CHELSA bioclimatic variables. For the "full" models, slope and forest cover variables were added to the bioclimatic variables

Table 2 The variable importance values for maxent models for *Onychodactylus koreanus* and k*arsenia koreana* based on permutation importance. The first and second most important variables for each model are highlighted in bold

Species	Data	Model	Bio1	Bio4	Bio12	Bio13	Bio14	Bio15	Forest cover	Slope
O. koreanus	WorldClim	Climate-only	70.40	1.73	2.60	14.36	10.80	0.10	N/A	N/A
O. koreanus	WorldClim	Full	5.99	3.55	0.48	1.31	25.72	12.46	4.19	46.30
O. koreanus	CHELSA	Climate-only	62.57	0.54	7.34	18.55	1.68	9.31	N/A	N/A
O. koreanus	CHELSA	Full	28.50	0	12.59	21.86	10.23	0.14	0.70	25.97
K. koreana	WorldClim	Climate-only	2.56	8.83	3.55	53.97	24.23	6.86	N/A	N/A
K. koreana	WorldClim	Full	15.55	0.01	0.81	13.11	24.36	18.23	7.26	20.65
K. koreana	CHELSA	Climate-only	1.04	32.73	25.19	5.55	32.53	2.96	N/A	N/A
K. koreana	CHELSA	Full	10.31	29.65	2.21	4.68	28.76	0.63	2.37	21.37

based on the climate-only model, although the predicted extent of suitable habitats in DPR Korea was broader in the full model.

The climate-only ENMs for *K. koreana* based on CHELSA data also had adequate predictive performance and a low degree of model overfitting (CBI = 0.839; AUC_{TEST} = 0.849; AUC_{DIFF} = 0.020; OR_{10} = 0.109; Table 1). When compared to the null ENMs, the climate-only empirical model performed significantly better ($p < 0.001$ for both CBI and AUC_{TEST}; Table 1). The climate-only model based on CHELSA data predicted suitable habitats along the central and southwestern parts of R Korea, as well as along the eastern coast of the Korean Peninsula (Fig. 2B; Supplementary Fig. 1B). The predicted area of suitable habitats based on the full model differed considerably from the prediction based on the climate-only model. For example, the full model predicted a much broader area of suitable habitats within DPR Korea, while the suitable habitats based on the climate-only model were limited to the eastern coast (Fig. 2B; Supplementary Fig. 1B).

For the climate-only model based on WorldClim data, the two most important variables were precipitation of the wettest month (bio13) and precipitation of the driest month (bio14). However, temperature seasonality (bio4) and precipitation of the driest month (bio14) were the two most important variables for the climate-only model based on CHELSA data. For the full model based on WorldClim data, precipitation of the driest month (bio14) and slope were the two most important variables, whereas temperature seasonality (bio4) and precipitation of the driest month (bio14) were the two most important variables for the full model based on CHELSA data. See Table 2 for the importance of each variable based on permutation importance, and Supplementary Table 2 for variable importance values based on percent contribution.

Hindcasting ENMs of O. koreanus

According to the hindcasted ENMs of *O. koreanus* based on WorldClim data (Fig. 3A), low to intermediate levels of habitat suitability were predicted at mPWP in northeastern China, the southern part of the Russian Far East, and the northeastern and eastern Korean Peninsula. At MIS19, areas of high habitat suitability were located along the eastern coast and northwestern part of the Korean Peninsula and Japan. The overall extent of highly suitable areas decreased somewhat during the LIG, but then expanded significantly during the LGM, with highly suitable areas covering the western and southern coasts of the Korean Peninsula, Japan, and the Yellow Sea basin. The habitat suitability along the eastern mountain ranges in the Korean Peninsula, however, was low during LGM. During the MH, the areas of high habitat suitability were located along the major mountain ranges of the Korean Peninsula, similar to the prediction under current climatic conditions. The MESS results suggest higher extrapolation risk in regions corresponding to the northern part of the Korean Peninsula and northeastern China compared to the southern part of the Korean Peninsula, which generally showed low to intermediate extrapolation risk. The extrapolation risk in this region is greatest at LIG (Supplementary Fig. 2).

The hindcasted ENMs of *O. koreanus* based on CHELSA data showed similar trends of changes in habitat suitability through time (Fig. 3A). Nevertheless, there were some notable differences between predictions based on the two different climatic datasets. For example, habitat suitability in the northern part of the Korean Peninsula, northeastern China, and the Russian Far East was higher during the mPWP in CHELSA-based prediction compared to the WorldClim-based prediction. During the LIG, predicted habitat suitability across the Korean Peninsula was higher overall for the CHELSA-based prediction compared to the WorldClim-based prediction. On the other hand, habitat suitability across the Korean Peninsula at LGM was lower in the CHELSA-based prediction compared to the WorldClim-based prediction. The MESS results based on CHELSA data showed considerably lower extrapolation risk across the landscape compared to the WorldClim data (Supplementary Fig. 2).

Hindcasting ENMs of K. koreana

According to the hindcasted ENMs of *K. koreana* based on the WorldClim data (Fig. 3B), habitat suitability was high across most of the Korean Peninsula and Japan at mPWP. However, habitat suitability within the Korean Peninsula decreased significantly at MIS19, with highly suitable habitats remaining on the eastern and southern coasts of the region. At LIG, the predicted habitat suitability was low across most of the projection area, with the exception of a small area in Japan. At LGM, areas with high suitability were found between the Korean Peninsula and Japan, along the contemporary Tsushima Strait and the western coast of Japan, but habitat suitability was again low across most of the projection area. At MH, highly suitable habitats were predicted on the eastern and southwestern mountain ranges of the Korean Peninsula. The MESS results showed higher extrapolation risk in the northern part of the Korean Peninsula and northeastern China compared to the southern part of the Korean Peninsula, which generally showed low to intermediate extrapolation risk (Supplementary Fig. 2).

The hindcasted ENMs of *K. koreana* based on CHELSA data (Fig. 3B) generally showed contrasting trends from those based on WorldClim data. For example, at mPWP, CHELSA-based ENMs predicted high habitat suitability across broad areas of the Korean Peninsula and Japan,

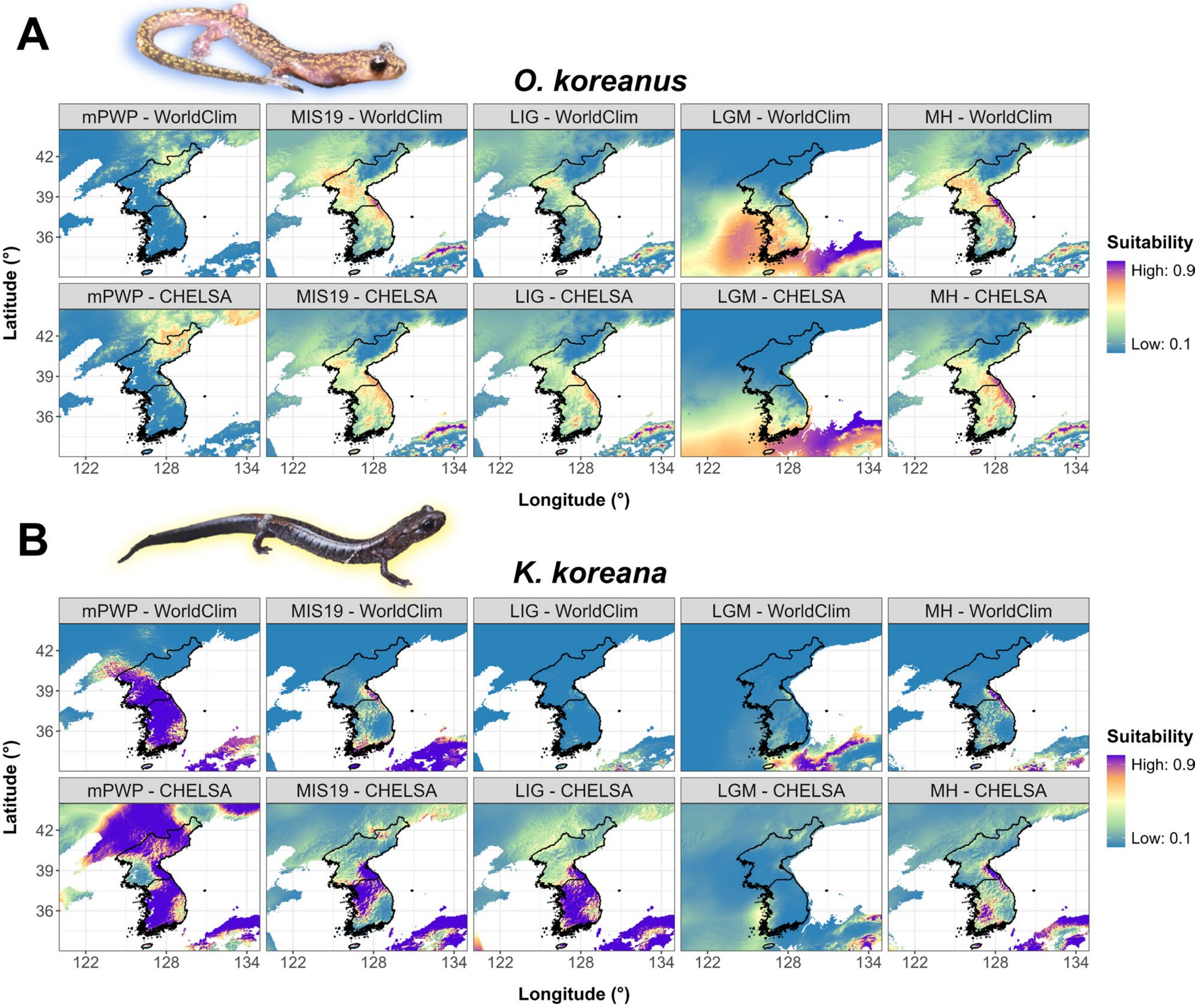

Fig. 3 Ecological niche models of *Onychodactylus koreanus* (**A**) and *Karsenia koreana* (**B**) hindcasted to the climatic conditions of the Plio-Pleistocene. The abbreviated names of time periods are as follows: mid-Pliocene Warm Period (mPWP; 3.205 Ma), Marine Isotope Stage 19 (MIS19; 787 Ka), Last Interglacial (LIG; 130 Ka), Last Glacial Maximum (LGM; 21 Ka), and mid-Holocene (MH; ca. 6.3 Ka). For each species, the upper panel represents model transfers from WorldClim-based models, and the lower panel represents model transfers from CHELSA-based models. Note the contrasting predictions depending on the initial climatic data used for model calibration

but also in northeastern China and the Russian Far East. At MIS19, the overall habitat suitability was higher across the projection area compared to the WorldClim-based predictions, with CHELSA-based ENMs predicting broad areas of highly suitable habitats across the central part of the Korean Peninsula. The CHELSA-based model predicted an expansion of highly suitable habitats at LIG, contrary to the WorldClim-based prediction. At LGM, however, the CHELSA-based model predicted a drastic decrease in suitable habitats, with an almost complete disappearance of highly suitable habitats across the Korean Peninsula, similar to the WorldClim-based prediction. At MH, areas of high habitat suitability were predicted along the eastern and southwestern mountain ranges of the Korean Peninsula, similar to the WorldClim-based prediction. However, the overall extent of highly suitable areas was broader for the CHELSA-based prediction. The MESS results based on CHELSA data showed considerably lower extrapolation risk across the landscape compared to the WorldClim data (Supplementary Fig. 2).

Influence of climatic data sources on ENMs

For *O. koreanus*, the differences in climatic values between WorldClim and CHELSA data were significant for the annual mean temperature (bio1; $W = 11{,}392.5$; $p < 0.001$; Fig. 4A), temperature seasonality (bio4; $W = 29{,}109$; $p < 0.001$; Fig. 4A), annual precipitation

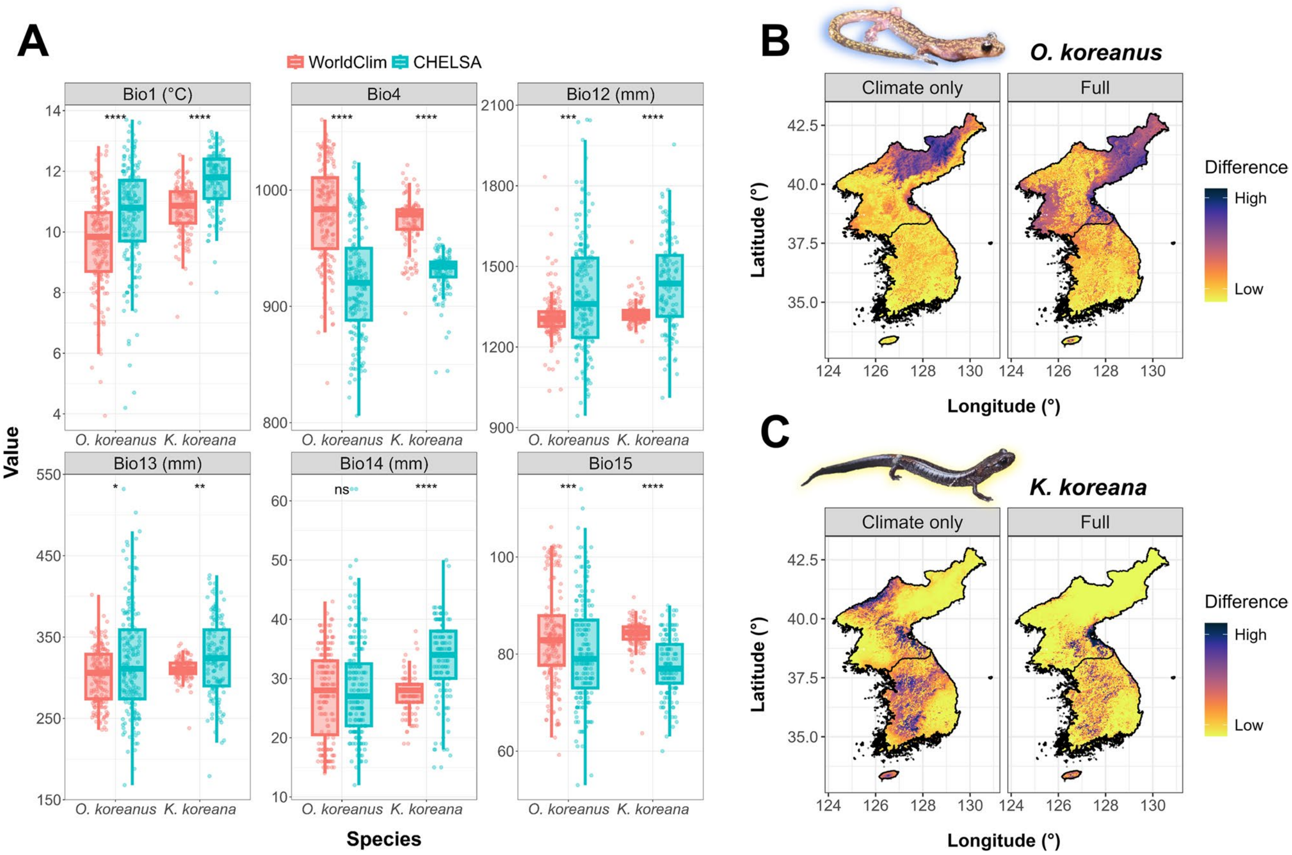

Fig. 4 The differences in model predictions and climate values across occurrence points of *Onychodactylus koreanus* and *Karsenia koreana*. **A** Comparisons of the range of climatic values between WorldClim and CHELSA datasets. The values were extracted from the occurrence points of each species ($n = 187$ for *O. koreanus* and $n = 137$ for *K. koreana*). The level of statistical significance based on Mann-Whitney *U*-tests are denoted with the following symbols above the boxplots – ns: $p > 0.05$, *: $p \le 0.05$, **: $p \le 0.01$, ***: $p \le 0.001$, ****: $p \le 0.0001$. **B** Differences in spatial predictions between WorldClim-based and CHELSA-based models for *O. koreanus*. **C** Differences in spatial predictions between WorldClim-based and CHELSA-based models for *K. koreana*. In (**B**) and (**C**), the CHELSA-based models used the same parameters as WorldClim-based models

(bio12; $W = 13{,}567.5$; $p < 0.001$; Fig. 4A), precipitation of the wettest month (bio13; $W = 15{,}312$; $p = 0.038$; Fig. 4A), and precipitation seasonality (bio15; $W = 21{,}091$; $p < 0.001$; Fig. 4A). On the other hand, the difference between the values of precipitation of the driest month (bio14; $W = 17{,}105$; $p = 0.717$; Fig. 4A) was not significant. For *K. koreana*, the differences in climatic values between WorldClim and CHELSA data were significant for all six variables. See Supplementary Table 3 for the full results of Mann-Whitney *U* tests.

When the difference between WorldClim and CHELSA bioclimatic layers were visualized spatially (Supplementary Fig. 3), CHELSA had higher values of temperature seasonality (bio4) along the northeastern coast of the Korean Peninsula, whereas WorldClim had higher values along the mountains of the peninsula. The values for annual precipitation (bio12) and precipitation of wettest month (bio13) were higher in CHELSA in central and northwestern regions of the Korean Peninsula than WorldClim. The values for precipitation of driest month (bio14) were higher in CHELSA in the mountains along the eastern coast of the Korean Peninsula, whereas WorldClim had higher values in southeastern part of the peninsula. The values for precipitation seasonality (bio15) were higher in CHELSA across broad regions of the Korean Peninsula, including the northwestern mountains and northeastern coasts.

When the differences in current habitat suitability predictions between WorldClim- and CHELSA-based models were visualized spatially for *O. koreanus*, the regions of greatest difference were in the mountains of northern Korean Peninsula for both climate-only and full models (Fig. 4B). For *K. koreana*, the climate-only predictions had a high level of difference in the northwestern mountains, eastern mountains, western lowlands, and southwestern mountains of the Korean Peninsula (Fig. 4C). For the full model, the regions of high difference were identified in the eastern mountains (Fig. 4C).

Based on the overall comparisons between WorldClim-based and CHELSA-based predictions, the highest level of difference between the two sets of predictions was at LIG for *O. koreanus*, followed by LGM, MH, mPWP, and

MIS19 (Table 3). The difference was the lowest for prediction under the current climate. For *K. koreana*, the prediction difference was also highest at LIG, followed by LGM, prediction under the current climate, MH, MIS19, and mPWP (Table 3). When the spatial information is incorporated using Schoener's *D*, the level of disagreement between WorldClim-based and CHELSA-based predictions was highest at LGM for both species. For *O. koreanus*, this was followed by MH, LIG and MIS19 (both having the same level of difference), mPWP, and prediction under the current climate. For *K. koreana*, mPWP had the highest level of prediction difference after LGM, followed by LIG, MH, MIS19, and prediction under the current climate. See Table 3 for the prediction difference values for each species, method, and time period.

For both species, the CHELSA-optimized models had comparable evaluation metrics compared to the models calibrated with parameters optimized for WorldClim data (Supplementary Table 4). The CHELSA-optimized predictions of current habitat suitability were also highly similar to CHELSA-based predictions with parameters optimized for the WorldClim data (Supplementary Fig. 4). When CHELSA-optimized models were hindcasted to Plio-Pleistocene climate conditions, however, the predictions for *O. koreanus* had a broader area of intermediate habitat suitability across all time periods than predictions from the CHELSA-based model with parameters optimized for the WorldClim data (Supplementary Fig. 5). For *K. koreana*, the predicted area of high habitat suitability was generally smaller in CHELSA-optimized hindcast predictions than predictions from the CHELSA-based model with parameters optimized for the WorldClim data (Supplementary Fig. 5).

Contemporary niche overlaps

The *p*-values for the background tests are shown as p_{12} for the tests comparing the E-space of *O. koreanus* (species 1) to the randomly shifted E-space of *K. koreana* (species 2), and p_{21} for the tests comparing the E-space of *K. koreana* to the randomly shifted E-space of *O. koreanus*. The niche overlap test (NOT) based on WorldClim data resulted in a nonsignificant identity test ($p > 0.05$; Fig. 5A) and background tests ($p_{12} > 0.05$; $p_{21} > 0.05$; Fig. 5B), and Schoener's *D* value calculated from the total available E-space was 0.19. When calculated only on the shared E-space, the NDT resulted in a significant identity test ($p < 0.05$; Fig. 5A) and nonsignificant background tests ($p_{12} > 0.05$; $p_{21} > 0.05$; Fig. 5B), with Schoener's *D* value of 0.10. The nonsignificant NOT background statistic indicates that the E-space available for both species is similar and the significant NDT niche identity statistic and nonsignificant NDT background metric indicates that there is evidence of niche divergence between *O. koreanus* and *K. koreana* ([68]; Fig. 5).

The NOT based on CHELSA data had Schoener's *D* value of 0.32, and resulted in a nonsignificant identity test ($p > 0.05$; Fig. 5A). For background tests (Fig. 5B), the comparison of observed niche similarity to the overlap between the E-space of *O. koreanus* and the randomly shifted E-space of *K. koreana* was nonsignificant ($p_{12} > 0.05$). However, the test was significant ($p_{21} < 0.05$) when the observed niche similarity was compared to the overlap between the E-space of *K. koreana* and the randomly shifted E-space of *O. koreanus*. The NDT conducted on CHELSA data had a Schoener's *D* value of 0.36 and a nonsignificant identity test result ($p > 0.05$; Fig. 5A). The background test was nonsignificant ($p_{12} > 0.05$) when observed niche similarity was compared to the overlap between the E-space of *O. koreanus* and the randomly shifted E-space of *K. koreana*. On the other hand, the test was significant ($p_{21} < 0.05$) when the observed niche similarity was compared to the E-space of *K. koreana* and the randomly shifted E-space of *O. koreanus* (Fig. 5B). As the niche identity tests were nonsignificant and the background tests were significant in only one direction for both NOT and NDT, the results indicate that the E-space available to the two species are similar and that the two species have equivalent niches ([68]; Fig. 5), contrary to the results based on the WorldClim data.

Table 3 Quantification of prediction differences between WorldClim- and CHELSA-based ENMs for *Onychodactylus koreanus* and *karsenia koreana*, based on the methods of Dubos et al. (2023) [17]. These are pairwise comparisons between WorldClim- and CHELSA-based predictions per time period. The "overall" method is based on the absolute difference between the sum of raster pixel values of WorldClim- and CHELSA-based predictions divided by the sum of WorldClim-based prediction pixel values and then multiplied by 100. On the other hand, the "spatial" method incorporates Spatial information based on schoener's *D* metric. The values for the "spatial" method are represented as (1 - *D*) x 100. For both methods, higher values indicate higher disagreement between WorldClim- and CHELSA-based predictions for each time period

Species	Method	Current (%)	MH (%)	LGM (%)	LIG (%)	MIS19 (%)	mPWP (%)
O. koreanus	Overall	32	142	163	207	122	126
O. koreanus	Spatial	20	35	47	32	32	22
K. koreana	Overall	49	35	86	960	34	23
K. koreana	Spatial	47	56	87	78	55	86

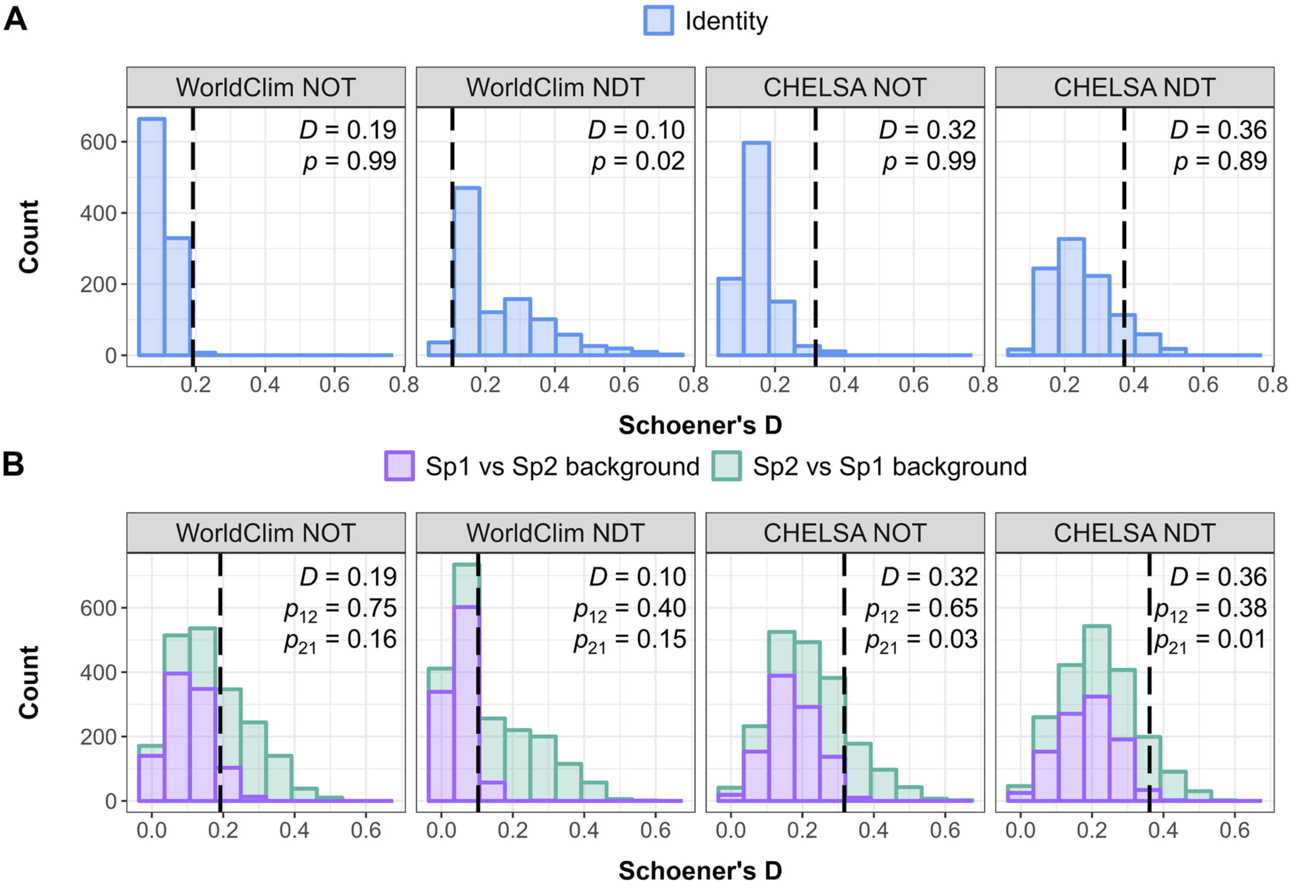

Fig. 5 The results of niche analyses based on (**A**) niche identity and (**B**) background tests. The analyses were conducted separately for WorldClim and CHELSA data, within the framework of the niche overlap test (NOT) and niche divergence test (NDT) of Brown & Carnaval (2019) [68]. The degree of niche overlap is based on Schoener's *D*. For the background test, p_{12} designates the test *p*-value when the environmental space (E-space) of species 1 (*Onychodactylus koreanus*) is compared to the randomly shifted E-space of species 2 (*Karsenia koreana*). Likewise, p_{21} designates the test *p*-value when the E-space of *K. koreana* is compared to the randomly shifted E-space of *O. koreanus*

Discussion

Ecological niche models and transferability

Our results demonstrate that the results of ENMs can sufficiently predict contemporary suitable habitats for *O. koreanus* and *K. koreana*. The predicted area of suitable habitats for *O. koreanus* was generally consistent with a previous ENM study, even with different model parameterizations and input variables [24]. On the other hand, the current range estimation for *K. koreana* predicted a broader area of suitable habitat compared to a previous study [32]. This is most likely the effect of correcting for spatial sampling bias through modified background point selection and adding occurrence points from the northern edge of the species' known distribution. Nevertheless, the prediction outputs were highly sensitive to input variable combinations (climate-only vs. full models) and climatic data sources (WorldClim vs. CHELSA), which were not tested in previous ENM studies on these species.

The uncertainties associated with climatic data choice are visualized when the distributions of climatic values at the species occurrence points are compared between WorldClim and CHELSA, where the range of values is considerably different between the two climate datasets for some variables (Fig. 4A). In addition, spatial visualization of difference between WorldClim and CHELSA climate variables showed that several climate layers differed consistently in similar regions within the Korean Peninsula. For example, one notable pattern is that the CHELSA layers had higher values of temperature seasonality (bio4), annual precipitation (bio12), precipitation of wettest month (bio13), and precipitation of driest month (bio14) in the eastern part of the Korean Peninsula, and higher annual precipitation (bio12), precipitation of wettest month (bio13), and higher precipitation seasonality (bio15) in the northwestern part of the Korean Peninsula. These two regions were where CHELSA-based models consistently predicted suitable habitats for *K. koreana* (Supplementary Fig. 1) and where the current habitat suitability predictions differed the most between WorldClim- and CHELSA-based models (Fig. 4B - C). This suggests that the difference between the two climate datasets is likely driving the differences in outcome predictions and identifies the choice of climate dataset as a source of uncertainty in niche modeling.

While we show that the choice of climate dataset has an important effect on model predictions, the climate-only ENMs failed to predict suitable habitats for these species during certain time periods of the Pleistocene (especially during LIG and LGM) regardless of the data source. This is in direct conflict with range formation histories inferred from genetic data. For example, the results of Jeon et al. (2021) [31] imply the persistence of relict *K. koreana* populations in the southern part of the Korean Peninsula, followed by a relatively recent, unidirectional northward dispersal forming the current geographic distribution. However, this pattern is not supported by hindcasted ENMs for some time periods (e.g., LIG and LGM). While regions of non-zero suitability are present for these time periods, these areas are small and patchily distributed. If we consider these predictions to be similar to the actual distributions during these time periods, we would expect to see a demographic signatures of rapid population expansion following the expansion of suitable habitats towards the present. Such a pattern is not supported from genetic data [31]. The same conflict holds for *O. koreanus*, which likely persisted within the mountains of the Korean Peninsula during the Pleistocene based on genetic data [23]. Furthermore, the hindcasted predictions appear to be sensitive to climatic data sources initially used to fit the current ENMs (WorldClim vs. CHELSA), with predictions for the same time period showing contrasting results (e.g., LIG predictions for *K. koreana*).

The failure of hindcasted ENMs to predict suitable Pleistocene habitats may stem from methodological approaches used in our ENMs. Compiling occurrence datasets from different sources, each with different coordinate accuracy and bias, as well as temporal and spatial resolutions of occurrence and environmental datasets, can all influence model outcomes [12–14, 70], and failure to account for these factors may lead to spurious model predictions. For our purposes, it was necessary to compile occurrence data from multiple sources, as no single source could provide a representative sampling of species ranges for either species. We nevertheless ascertained the accuracy of occurrence points whenever possible, by visualizing the distribution of occurrence points across the landscape, matching them with known distributions, and filtering out potentially inaccurate records. Considering the data filtering methods and prediction outcomes, the compilation of occurrence points from different sources is unlikely to be the cause of poor transferability.

Regarding the temporal resolution, it was not possible to fully match our occurrence datasets exactly to the timeframes of the climatic data. However, this mismatch in temporal resolution between occurrence and climatic data does not seem to have adverse effects on model performance. For example, we were not able to match the temporal resolution between the WorldClim dataset (1970–2000 climatic averages) and occurrence points of *K. koreana* because the species was formally described only in 2005 [25]; except for three records based on historical specimens [71, 72], all other occurrence points were recorded after 2005. The CHELSA climatology, spanning climate data from 1979 to 2013, better matches the occurrence timeframe of *K. koreana*. Despite this, WorldClim-based models outperformed CHELSA-based models in terms of consistency with the known distribution of the species, with all the other statistical metrics having comparable values between the two model sets (Table 1). The same holds for *O. koreanus*: the occurrence points for this species encompass even broader types of records, including literature-based data, georeferenced historical specimens, herpetological survey records, and citizen science data [24], spanning time periods that are much broader than the coverage of either climatic dataset. Nevertheless, the models could sufficiently predict the known current distribution of this species.

The spatial resolution of our input raster data also does not seem to be the root cause of poor transferability either. We used statistical downscaling of the raw 5 km PaleoClim dataset to match its spatial resolution with the contemporary climate datasets used for calibration. However, additional experiments with ENMs calibrated on 5 km resolution WorldClim/CHELSA datasets and hindcasting them to raw 5 km PaleoClim dataset still resulted in poor model transferability (not shown). Similarly, hindcasting ENMs calibrated with a 1 km WorldClim/CHELSA dataset to 1 km paleoclimate at LIG [59] also resulted in failure to predict suitable habitats during this time period (not shown).

In addition, potential suboptimal parameterization of CHELSA-based models due to the use of parameters optimized for WorldClim data could be another reason for poor transferability in CHELSA-based models. We tested this by running a separate model parameter tuning for CHESLA data and generated current and hindcasted predictions. For both species, the predictions of current habitat suitability were similar between CHELSA-based models that used parameters optimized for WorldClim data and CHELSA-based models with parameters specifically optimized for CHELSA dataset. On the other hand, the hindcasted predictions from CHELSA-based models with parameters optimized for CHELSA data were different those from CHELSA-based models that used parameters optimized for WorldClim data. For example, CHELSA-optimized hindcast predictions for *O. koreanus* predicted broader areas of intermediate suitability than hindcast predictions based on WorldClim-optimized paramters. Nevertheless, specifically optimizing model parameters for CHELSA data did not resolve poor model transferability. This is clearly demonstrated in model predictions for *K. koreana*, where the model still failed to predict suitable habitats in the Korean Peninsula during the LGM despite specific parameter optimization for CHELSA data. These results suggest that, while the parameter specification do have an influence in output predictions, this is not necessarily a root cause of poor model transferability.

Thus, barring all these possibilities, we seek explanations from ecological factors and fundamental limitations of macroclimate-based correlative ENMs. Both *O. koreanus* and *K. koreana* are strict habitat specialists that are only found in heavily forested mountains [27]. Within these habitats, they utilize specific microhabitats including rock piles, leaf litter, or cold streams (for *O. koreanus*). In such environments, the microclimates directly experienced by these animals can differ significantly from the regional macroclimate [18]. Taking fine-scale microclimate and physiologically relevant variables into account, therefore, is likely to generate more realistic estimations of habitat suitability [18, 73]. However, the ecology of both *O. koreanus* and *K. koreana* remains poorly studied, and physiologically important parameters for salamanders, such as rates of water loss and respiration, are unknown. Therefore, further research on the ecophysiology of these species and mechanistic niche modeling approaches are the obvious next steps.

The need to incorporate microclimate-scale modeling capacity is exemplified in the behavior of our macroclimate-based ENMs. Based on the importance of each environmental variable for ENMs, it appears that the hindcasted predictions for both species are largely driven by the most important variable in the initial model calibration. For example, annual mean temperature (bio1) is disproportionately important in climate-only models of *O. koreanus* (Table 2). However, when the slope and forest cover variables were added, the importance of bio1 decreased considerably. Therefore, a disproportionate impact of bio1 on the climate-only ENMs of *O. koreanus* seems to explain the apparent failure to predict suitable habitats for this species during LGM, considering the significantly colder climate of the Korean Peninsula during this time period [74, 75]. For *K. koreana*, temperature seasonality (bio4) and/or precipitation extremes (bio13 and bio14) had high importance in current ENMs (Table 2). Considering that the climate of the Korean Peninsula during LGM was much colder and drier [74, 75], the high importance of temperature and precipitation variables is likely to be the reason for poor model transfer.

The goal of our study is not to denounce the values of correlative ENMs altogether. Indeed, correlative ENMs are powerful and robust tools for understanding many aspects of biogeography, evolution, paleontology, and invasion ecology [2, 76–78]. Rather, our aim is to caution the users of correlative ENMs against their blind applications without critical assessments of modeling aims, model assumptions, available data, and ecological information on target species. For example, directly associating population trends estimated from genetic data with the increase/decrease of predicted habitat suitability under the Pleistocene climate could lead to spurious conclusions if climatic data sources strongly influence the predictions. Therefore, studies seeking to complement the results from phylogeographic studies (e.g., Bayesian skyline plots) with hindcasted ENMs should carefully test the impacts of climatic data sources and select environmental variables relevant to the target species as much as possible. Furthermore, the modeling protocols should follow the available best practices [57, 79].

Such cautions equally apply to studies that aim to estimate future suitable habitats under climate change. For example, previous studies have demonstrated that accounting for plasticity and other physiological parameters can produce more conservative predictions of species response to future climate change compared to the conventional correlative ENMs [80–82]. Our previous studies on the potential response of *O. koreanus* and *K. koreana* to future climate changes predicted considerable range shifts with concomitant conservation implications [24, 32]. Considering the uncertainties and limitations associated with correlative ENMs, these conclusions should be re-evaluated and refined as better modeling capabilities, as well as additional ecological and ecophysiological data, become available. On the other hand, further testing of the influences of model parameterization, spatial and temporal resolutions, and alternative modeling approaches, such as ensemble modeling [83, 84] and minimum volume ellipsoids [85], and comparing these results to our MaxEnt-based outputs presented here could be useful avenues of methodological investigations.

Niche overlaps and range formation

We initially considered niche overlaps between *O. koreanus* and *K. koreana* as more likely than niche divergence at the macroclimatic scale based on generally similar habitat requirements, overlapping geographic distributions, and broad sympatry between the two species. However, the results of niche analyses suggested that these tests are also sensitive to climatic data sources. In our case, this is likely due to the different ranges of climatic values between WorldClim and CHELSA datasets (Fig. 4), as well as data resolution. Also, while not directly related to the framework of the tests implemented here, niche analyses are scale-dependent [86]. For example, a previous study based on diet and stable isotope data suggests trophic niche differentiation between *O. koreanus* and *K. koreana* [87], while niche overlap/divergence with respect to other variables (e.g., morphology, microhabitat use) and scales remain poorly understood for these species. As these two species vary in key ecological characteristics (e.g., morphology and life histories), detailed ecological studies investigating finer-scale habitat preferences of these two species in regions of sympatry may yield further insights into patterns of niche overlaps and divergence. On the other hand, as numerous studies have employed niche analyses to test hypotheses on the range formations and niche evolution between species pairs [4, 88–90], the influence of climatic data sources on the results of these tests, scale dependence of niche estimations, and subsequent impacts on interpretations warrant further investigation.

As it currently stands, it is difficult to gain detailed insights into the range formation processes of *O. koreanus* and *K. koreana* from ENMs. The apparent conflicts between model-based predictions and phylogeographic patterns based on genetic data are difficult to reconcile. To resolve this issue, further ecological studies, additional datasets, and improved modeling capacities are required to incorporate microclimatic, physiological, and other non-climatic variables (e.g., topography) into the ENM framework. In addition, phylogeographic studies based on phylogenomic data could be fruitful avenues for future investigations. While whole-genome sequencing for salamanders remains challenging due to large genome size (but see [91]), reduced-representation genome sequencing methods can be readily applied to salamanders [92–95]. The results from such studies may also provide further resolutions to some of the puzzling patterns revealed by mtDNA-based phylogeography, such as geographically proximate populations possessing divergent haplotypes and populations separated by complex mountain ranges sharing the same haplotypes [23, 31]. For *O. koreanus*, studies using phylogenomic data may also provide insights into the processes of speciation between this species and a microendemic sister species, *O. sillanus* [96]. For *K. koreana*, additional fieldwork may uncover additional populations, especially in the poorly sampled and heavily forested northern mountain ranges. These efforts can benefit from predictions and ground validation based on fine-scale ENMs and environmental DNA [97, 98]. Overall, such integrative approaches will provide further insights into the evolutionary histories and biogeography of these unique and poorly understood lineages of salamanders.

Conclusion

Using two distantly related lungless salamanders found in broad sympatry and two popular climatic databases, we investigated the influences of climatic data choice on the outcomes of correlative niche model predictions and niche analyses. Our results demonstrate that, while the models can sufficiently predict contemporary distributions, the choice of input climate data source has a strong influence on model predictions. Despite the ability of these models to predict current distributions, Plio-Pleistocene range predictions failed to predict suitable habitats regardless of the input climate data, even though previous genetic studies suggested the persistence of populations during this time period. The analyses of niche overlaps and divergence were also sensitive to climatic data choice, leading to different conclusions depending on the input data source. Taken together, we conclude that the influence of input climatic data and limitations of macroclimate-based analyses can be major sources of uncertainties in model transfers and analyses of niche overlaps and divergence, especially when the species is likely buffered from large-scale environmental changes by microhabitat use. The impacts of climatic data choice on niche modeling and associated analyses thus warrant further investigations.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12862-025-02442-3.

Supplementary Material 1.

Supplementary Material 2.

Acknowledgements
Y.S. was supported by the Richard Gilder Graduate School Doctoral Fellowship from the American Museum of Natural History during the writing of this manuscript. We thank Jong Yoon Jeon for fruitful discussions on the population genetic structure and historical demography of K. koreana. We also thank the editor and three anonymous reviewers for their comments that significantly improved the manuscript.

Authors' contributions

Y.S. and D.P. conceived the project, Y.S. collected the data, carried out the analyses, generated figures, and wrote the main manuscript text. A.B. and D.P. provided feedback on the initial manuscript draft, and D.P. provided funding for the study. All authors reviewed the manuscript.

Funding
This work was supported by the National Research Foundation of Korea (NRF) grant IRIS RS-2024-00436579, funded by the Ministry of Science and ICT (MSIT) of the Korean government.

Data availability

The datasets, package dependencies, and custom R scripts used in this study can be found in a dedicated GitHub repository for this project (https://github.com/yucheols/TwoSalDist). In addition, the environmental variables used for the analyses are available from public databases and can be retrieved through the information provided in the main text. The detailed niche modeling protocol is available as an ODMAP protocol in Supplementary Material 1.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Department of Biological Sciences, Kangwon National University, Chuncheon 24341, Republic of Korea
[2]Richard Gilder Graduate School, American Museum of Natural History, Central Park West at 79th Street, New York, NY 10024, USA
[3]Department of Herpetology, Division of Vertebrate Zoology, American Museum of Natural History, Central Park West at 79th Street, New York, NY 10024, USA
[4]Laboratory of Animal Behaviour and Conservation, College of Life Sciences, Nanjing Forestry University, Nanjing, Jiangsu 210037, People's Republic of China
[5]Department of Zoology and General Biology, Faculty of Life Sciences, Fergana State University, Fergana 150100, Uzbekistan
[6]Division of Science Education, Kangwon National University, Chuncheon 24341, Republic of Korea

Received: 10 March 2025 / Accepted: 26 August 2025

References
1. Pena JC, de Kamino C, Rodrigues LHY, Mariano-Neto M, de Siqueira E. MF. Assessing the conservation status of species with limited available data and disjunct distribution. Biol Conserv. 2014;170:130–6. https://doi.org/10.1016/j.biocon.2013.12.015.
2. Meseguer AS, Lobo JM, Ree R, Beerling DJ, Sanmartín I. Integrating fossils, phylogenies, and niche models into biogeography to reveal ancient evolutionary history: the case of *Hypericum* (Hypericaceae). Syst Biol. 2015. https://doi.org/10.1093/sysbio/syu088
3. Murphy SJ, Smith AB. What can community ecologists learn from species distribution models? Ecosphere. 2021. https://doi.org/10.1002/ecs2.3864.
4. Jones LN, Leaché AD, Burbrink FT. Biogeographic barriers and historic climate shape the phylogeography and demography of the common gartersnake. J Biogeogr. 2023. https://doi.org/10.1111/jbi.14709.
5. Jeon JY, Shin Y, Mularo AJ, Feng X, DeWoody JA. The integration of whole-genome resequencing and ecological niche modelling to conserve profiles of local adaptation. Divers Distrib. 2024;30:e13847.
6. Fick SE, Hijmans RJ. Worldclim 2: new 1-km spatial resolution climate surfaces for global land areas. Int J Climatol. 2017. https://doi.org/10.1002/joc.5086.
7. Karger DN, Conrad O, Böhner J, Kawohl T, Kreft H, Soria-Auza RW, et al. Climatologies at high resolution for the earth's land surface areas. Sci Data. 2017. https://doi.org/10.1038/sdata.2017.122.
8. Title PO, Bemmels JB. Envirem: an expanded set of bioclimatic and topographic variables increases flexibility and improves performance of ecological niche modeling. Ecography. 2018. https://doi.org/10.1111/ecog.02880.
9. Boria RA, Olson LE, Goodman SM, Anderson RP. Spatial filtering to reduce sampling bias can improve the performance of ecological niche models. Ecol Modell. 2014. https://doi.org/10.1016/j.ecolmodel.2013.12.012.
10. Varela S, Anderson RP, García-Valdés R, Fernández-González F. Environmental filters reduce the effects of sampling bias and improve predictions of ecological niche models. Ecography. 2014. https://doi.org/10.1111/j.1600-0587.2013.00441.x.
11. Warren DL, Wright AN, Seifert SN, Shaffer HB. Incorporating model complexity and spatial sampling bias into ecological niche models of climate change risks faced by 90 California vertebrate species of concern. Divers Distrib. 2014. https://doi.org/10.1111/ddi.12160.
12. Anderson RP. Harnessing the world's biodiversity data: promise and peril in ecological niche modeling of species distributions. Ann N Y Acad Sci. 2012. https://doi.org/10.1111/j.1749-6632.2011.06440.x.
13. Botella C, Joly A, Monestiez P, Bonnet P, Munoz F. Bias in presence-only niche models related to sampling effort and species niches: lessons for background point selection. PLoS ONE. 2020;15:e0232078. https://doi.org/10.1371/journal.pone.0232078.
14. Moudrý V, Bazzichetto M, Remelgado R, Devillers R, Lenoir J, Mateo RG, et al. Optimising occurrence data in species distribution models: sample size, positional uncertainty, and sampling bias matter. Ecography. 2024. https://doi.org/10.1111/ecog.07294.
15. Baker DJ, Hartley AJ, Butchart SHM, Willis SG. Choice of baseline climate data impacts projected species' responses to climate change. Glob Chang Biol. 2016;22:2392–404. https://doi.org/10.1111/gcb.13273.
16. Bobrowski M, Udo S. Why input matters: selection of climate data sets for modelling the potential distribution of a treeline species in the Himalayan region. Ecol Modell. 2017;359:92–102. https://doi.org/10.1016/j.ecolmodel.2017.05.021.
17. Dubos N, Fieldsend TW, Roesch MA, Augros S, Besnard A, Choeur A et al. Choice of climate data influences predictions for current and future global invasion risks for two *Phelsuma* geckos. Biol Invasions. 2023;25:2929–48. https://doi.org/10.1007/s10530-023-03082-8.
18. Haesen S, Lenoir J, Gril E, De Frenne P, Lembrechts JJ, Kopecký M, et al. Microclimate reveals the true thermal niche of forest plant species. Ecol Lett. 2023. https://doi.org/10.1111/ele.14312.
19. Kearney M, Porter W. Mechanistic niche modelling: combining physiological and spatial data to predict species' ranges. Ecol Lett. 2009. https://doi.org/10.1111/j.1461-0248.2008.01277.x.
20. Briscoe NJ, Morris SD, Mathewson PD, Buckley LB, Jusup M, Levy O, et al. Mechanistic forecasts of species responses to climate change: the promise of biophysical ecology. Glob Change Biol. 2023. https://doi.org/10.1111/gcb.16557.
21. Lee JH, Ra NY, Eom J, Park D. Population dynamics of the long-tailed clawed salamander larva, *Onychodactylus fischeri*, and its age structure in Korea. J Ecol Field Biology. 2008;31:31–6.
22. Poyarkov NA, Che J, Min MS, Kuro-O M, Yan F, Li C et al. Review of the systematics, morphology and distribution of Asian clawed salamanders, genus *Onychodactylus* (Amphibia, caudata: Hynobiidae), with the description of four new species. Zootaxa. 2012:3465:1–106. https://doi.org/10.11646/zootaxa.3465.1.1.
23. Suk HY, Lee MY, Bae HG, Lee SJ, Poyarkov N, Lee H, et al. Phylogenetic structure and ancestry of Korean clawed salamander, *Onychodactylus koreanus* (Caudata: Hynobiidae). Mitochondr DNA A DNA Mapp Seq Anal. 2018. https://doi.org/10.1080/24701394.2017.1339187
24. Shin Y, Min MS, Borzée A. Driven to the edge: species distribution modeling of a clawed salamander (Hynobiidae: *Onychodactylus koreanus*) predicts range shifts and drastic decrease of suitable habitats in response to climate change. Ecol Evol. 2021;11:14669–88.
25. Min MS, Yang SY, Bonett RM, Vieites DR, Brandon RA, Wake DB. Discovery of the first Asian plethodontid salamander. Nature. 2005. https://doi.org/10.1038/nature03474.
26. Moon KY, Park D. Report of *Karsenia koreana* eggs oviposited within a semi-natural terrarium constructed at natural habitat. Korean J Herpetol. 2016;7:1–5.
27. Lee J-H, Park D. The encyclopedia of Korean amphibians. Seoul, South Korea: Nature and Ecology; 2016.
28. Borzée A, Litvinchuk SN, Ri K, Andersen D, Nam TY, Jon GH, et al. Update on distribution and conservation status of amphibians in the Democratic People's Republic of Korea: conclusions based on field surveys, environmental modelling, molecular analyses and call properties. Animals. 2021. https://doi.org/10.3390/ani11072057.
29. Vieites DR, Min MS, Wake DB. Rapid diversification and dispersal during periods of global warming by plethodontid salamanders. Proc Natl Acad Sci U S A. 2007. https://doi.org/10.1073/pnas.0705056104.
30. Shen XX, Liang D, Chen MY, Mao RL, Wake DB, Zhang P. Enlarged multilocus data set provides surprisingly younger time of origin for the plethodontidae, the largest family of salamanders. Syst Biol. 2016;65:66–81. https://doi.org/10.1093/sysbio/syv061.
31. Jeon JY, Jung Jhwa, Suk HY, Lee H, Min MS. The Asian plethodontid salamander preserves historical genetic imprints of recent Northern expansion. Sci Rep. 2021. https://doi.org/10.1038/s41598-021-88238-z.
32. Borzée A, Andersen D, Groffen J, Kim HT, Bae Y, Jang Y. Climate change-based models predict range shifts in the distribution of the only Asian plethodontid salamander: *Karsenia koreana*. Sci Rep. 2019. https://doi.org/10.1038/s41598-019-48310-1.
33. Hewitt G. The genetic legacy of the quaternary ice ages. Nature. 2000. https://doi.org/10.1038/35016000.
34. Kim HW, Yoon S, Kim M, Shin M, Yoon H, Kim K. EcoBank: A flexible database platform for sharing ecological data. Biodivers Data J. 2021;9:e61866. https://doi.org/10.3897/BDJ.9.e61866.

35. Shipley BR, Bach R, Do Y, Strathearn H, McGuire JL, Dilkina B. megaSDM: integrating dispersal and time-step analyses into species distribution models. Ecography. 2022;2022:e05450. https://doi.org/10.1111/ecog.05450.
36. Vignali S, Barras AG, Arlettaz R, Braunisch V. SDMtune: an R package to tune and evaluate species distribution models. Ecol Evol. 2020;10:11488–506. https://doi.org/10.1002/ece3.6786.
37. R Core Team. R: A language and environment for statistical computing, Version 4.2.2. 2022. https://www.Rproject.org//. Accessed 19 Sep 2024.
38. Venables WN, Ripley BD. Modern Applied Statistics with S. 4th Edition. New York: Springer; 2002.
39. Tuanmu MN, Jetz W. A global 1-km consensus land-cover product for biodiversity and ecosystem modelling. Glob Ecol Biogeogr. 2014. https://doi.org/10.1111/geb.12182.
40. Hijmans RJ. raster: Geographic Data Analysis and Modeling. 2023.
41. Peng Y, Li Y, Cao G, Li H, Shin Y, Piao Z, et al. Estimation of habitat suitability and landscape connectivity for Liaoning and Jilin clawed salamanders (Hynobiidae: *Onychodactylus*) in the transboundary region between the People's Republic of China and the Democratic People's Republic of Korea. Glob Ecol Conserv. 2023. https://doi.org/10.1016/j.gecco.2023.e02694.
42. Borzée A. Continental Northeast Asian Amphibians Origins, Behavioural Ecology, and Conservation. 1st edition. Amsterdam: Elsevier; 2024.
43. Osorio-Olvera L, Lira-Noriega A, Soberón J, Peterson AT, Falconi M, Contreras-Díaz RG, et al. Ntbox: an R package with graphical user interface for modelling and evaluating multidimensional ecological niches. Methods Ecol Evol. 2020. https://doi.org/10.1111/2041-210X.13452.
44. Phillips SJ, Anderson RP, Dudík M, Schapire RE, Blair ME. Opening the black box: an open-source release of maxent. Ecography. 2017. https://doi.org/10.1111/ecog.03049.
45. Kass JM, Muscarella R, Galante PJ, Bohl CL, Pinilla-Buitrago GE, Boria RA, et al. ENMeval 2.0: redesigned for customizable and reproducible modeling of species' niches and distributions. Methods Ecol Evol. 2021. https://doi.org/10.1111/2041-210X.13628.
46. Roberts DR, Bahn V, Ciuti S, Boyce MS, Elith J, Guillera-Arroita G, et al. Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. Ecography. 2017. https://doi.org/10.1111/ecog.02881.
47. Valavi R, Elith J, Lahoz-Monfort JJ, Guillera-Arroita G, blockCV. An r package for generating spatially or environmentally separated folds for k-fold cross-validation of species distribution models. Methods Ecol Evol. 2019;10:225–32. https://doi.org/10.1111/2041-210X.13107.
48. Muscarella R, Galante PJ, Soley-Guardia M, Boria RA, Kass JM, Uriarte M et al. ENMeval: an R package for conducting spatially independent evaluations and estimating optimal model complexity for maxent ecological niche models. Methods Ecol Evol. 2014;5:1198–205. https://doi.org/10.1111/2041-210X.12261.
49. Melton AE, Clinton MH, Wasoff DN, Lu L, Hu H, Chen Z, et al. Climatic niche comparisons of Eastern North American and Eastern Asian disjunct plant genera. Glob Ecol Biogeogr. 2022. https://doi.org/10.1111/geb.13504.
50. Warren DL, Seifert SN. Ecological niche modeling in maxent: the importance of model complexity and the performance of model selection criteria. Ecol Appl. 2011. https://doi.org/10.1890/10-1171.1.
51. Radosavljevic A, Anderson RP. Making better maxent models of species distributions: complexity, overfitting and evaluation. J Biogeogr. 2014. https://doi.org/10.1111/jbi.12227.
52. Low BW, Zeng Y, Tan HH, Yeo DCJ. Predictor complexity and feature selection affect maxent model transferability: evidence from global freshwater invasive species. Divers Distrib. 2021. https://doi.org/10.1111/ddi.13211.
53. Lobo JM, Jiménez-valverde A, Real R. AUC: a misleading measure of the performance of predictive distribution models. Glob Ecol Biogeogr. 2008. https://doi.org/10.1111/j.1466-8238.2007.00358.x.
54. Jiménez-Valverde A. Insights into the area under the receiver operating characteristic curve (AUC) as a discrimination measure in species distribution modelling. Glob Ecol Biogeogr. 2012. https://doi.org/10.1111/j.1466-8238.2011.00683.x.
55. Hirzel AH, Le Lay G, Helfer V, Randin C, Guisan A. Evaluating the ability of habitat suitability models to predict species presences. Ecol Modell. 2006. https://doi.org/10.1016/j.ecolmodel.2006.05.017.
56. Bohl CL, Kass JM, Anderson RP. A new null model approach to quantify performance and significance for ecological niche models of species distributions. J Biogeogr. 2019. https://doi.org/10.1111/jbi.13573.
57. Zurell D, Franklin J, König C, Bouchet PJ, Dormann CF, Elith J et al. A standard protocol for reporting species distribution models. Ecography. 2020;43:1261–77. https://doi.org/10.1111/ecog.04960.
58. Fitzpatrick MC, Lachmuth S, Haydt NT. The ODMAP protocol: a new tool for standardized reporting that could revolutionize species distribution modeling. Ecography. 2021. https://doi.org/10.1111/ecog.05700.
59. Otto-Bliesner BL, Marshall SJ, Overpeck JT, Miller GH, Hu A. Simulating arctic climate warmth and icefield retreat in the last interglaciation. Science. 2006;311:1751–53. https://doi.org/10.1126/science.1120808.
60. Fordham DA, Saltré F, Haythorne S, Wigley TML, Otto-Bliesner BL, Chan KC, et al. Paleoview: a tool for generating continuous climate projections spanning the last 21 000 years at regional and global scales. Ecography. 2017. https://doi.org/10.1111/ecog.03031.
61. Brown JL, Hill DJ, Dolan AM, Carnaval AC, Haywood AM. Paleoclim, high Spatial resolution paleoclimate surfaces for global land areas. Sci Data. 2018;5:180254.https://doi.org/10.1038/sdata.2018.254.
62. Elith J, Kearney M, Phillips S. The Art of modelling range-shifting species. Methods Ecol Evol. 2010;1:330–42. https://doi.org/10.1111/j.2041-210X.2010.00036.x.
63. Dubos N, Augros S, Deso G, Probst JM, Notter JC, Roesch MA. Here be dragons: important Spatial uncertainty driven by climate data in forecasted distribution of an endangered insular reptile. Anim Conserv. 2022;25:704–17. https://doi.org/10.1111/acv.12775.
64. Schoener TW. The *Anolis* lizards of bimini: resource partitioning in a complex fauna. Ecology. 1968. https://doi.org/10.2307/1935534.
65. Warren DL, Glor RE, Turelli M. Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. Evolution. 2008. https://doi.org/10.1111/j.1558-5646.2008.00482.x.
66. Warren DL, Matzke NJ, Cardillo M, Baumgartner JB, Beaumont LJ, Turelli M, et al. ENMTools 1.0: an R package for comparative ecological biogeography. Ecography. 2021. https://doi.org/10.1111/ecog.05485.
67. Broennimann O, Fitzpatrick MC, Pearman PB, Petitpierre B, Pellissier L, Yoccoz NG, et al. Measuring ecological niche overlap from occurrence and spatial environmental data. Glob Ecol Biogeogr. 2012. https://doi.org/10.1111/j.1466-8238.2011.00698.x.
68. Brown JL, Carnaval AC. A Tale of two niches: methods, concepts, and evolution. Front Biogeogr. 2019;11:e44158. https://doi.org/10.21425/F5FBG44158.
69. Brown JL. humboldt R package documentation. 2019. https://jasonleebrown.github.io/humboldt/. Accessed 8 Mar 2025.
70. Roubicek AJ, VanDerWal J, Beaumont LJ, Pitman AJ, Wilson P, Hughes L. Does the choice of climate baseline matter in ecological niche modelling? Ecol Modell. 2010. https://doi.org/10.1016/j.ecolmodel.2010.06.021.
71. Nishikawa K. The first specimen of *Karsenia Koreana* (Caudata: Plethodontidae) collected 34 years before its description. Curr Herpetol. 2009;28:27–8. https://doi.org/10.3105/018.028.0104.
72. Shin Y, Jang Y, Kim T, Borzée A. A specimen of *Karsenia koreana* (caudata: plethodontidae) misidentified as *Hynobius leechii* 27 years before the species' description and additional historical record. Curr Herpetol. 2020;39:75–9. https://doi.org/10.5358/hsj.39.75.
73. Guillon M, Martínez-Freiría F, Lucchini N, Ursenbacher S, Surget-Groba Y, Kageyama M, et al. Inferring current and last glacial maximum distributions are improved by physiology-relevant climatic variables in cold-adapted ectotherms. J Biogeogr. 2024. https://doi.org/10.1111/jbi.14828.
74. Kim S-J, Park Y-M, Lee B-Y, Choi T-J, Yoon Y-J, Suk B-C. Study of East Asia climate change for the last glacial maximum using numerical model. Korean J Quaternary Res. 2006;20:51–66.
75. Kim SJ, Kim JW, Kim BM. Last glacial maximum climate over Korean Peninsula in PMIP3 simulations. Quat Int. 2015. https://doi.org/10.1016/j.quaint.2015.02.062.
76. Escoriza D, Hassine JB. Niche partitioning at local and regional scale in the North African Salamandridae. J Herpetol. 2015;49:276–83. https://doi.org/10.1670/13-151.
77. Barbet-Massin M, Rome Q, Villemant C, Courchamp F. Can species distribution models really predict the expansion of invasive species? PLoS One. 2018. https://doi.org/10.1371/journal.pone.0193085.
78. Chiarenza AA, Mannion PD, Lunt DJ, Farnsworth A, Jones LA, Kelland SJ et al. Ecological niche modelling does not support climatically-driven dinosaur diversity decline before the cretaceous/paleogene mass extinction. Nat Commun. 2019;10:1091. https://doi.org/10.1038/s41467-019-08997-2.
79. Feng X, Park DS, Walker C, Peterson AT, Merow C, Papeş M. A checklist for maximizing reproducibility of ecological niche models. Nat Ecol Evol. 2019. https://doi.org/10.1038/s41559-019-0972-5.
80. Riddell EA, Odom JP, Damm JD, Sears MW. Plasticity reveals hidden resistance to extinction under climate change in the global hotspot of salamander diversity. Sci Adv. 2018. https://doi.org/10.1126/sciadv.aar5471.

81. Lyons MP, Kozak KH. Vanishing islands in the sky? A comparison of correlation- and mechanism-based forecasts of range dynamics for montane salamanders under climate change. Ecography. 2020. https://doi.org/10.1111/ecog.04282.
82. Riddell E, Sears MW. Terrestrial salamanders maintain habitat suitability under climate change despite trade-offs between water loss and gas exchange. Physiol Biochem Zool. 2020. https://doi.org/10.1086/709558.
83. Velazco SJE, Rose MB, de Andrade AFA, Minoli I, Franklin J. Flexsdm: an r package for supporting a comprehensive and flexible species distribution modelling workflow. Methods Ecol Evol. 2022;13:1661–9. https://doi.org/10.1038/s41467-019-08997-2.
84. Thuiller W, Lafourcade B, Engler R, Araújo MB. Biomod - a platform for ensemble forecasting of species distributions. Ecography. 2009. https://doi.org/10.1111/j.1600-0587.2008.05742.x.
85. Van Aelst S, Rousseeuw P. Minimum volume ellipsoid. Wiley Interdisciplinary Reviews: Comput Stat. 2009;1:71–82. https://doi.org/10.1002/wics.19.
86. Ficetola GF, Lunghi E, Canedoli C, Padoa-Schioppa E, Pennati R, Manenti R. Differences between microhabitat and broad-scale patterns of niche evolution in terrestrial salamanders. Sci Rep. 2018. https://doi.org/10.1038/s41598-018-28796-x.
87. Jung J-H. Habitat characteristics of three salamanders (Caudata: Amphibia) in forests and genetic diversity of *Karsenia koreana*. PhD Dissertation. Seoul National University; 2020.
88. Oboudi R, Malekian M, Khosravi R, Fadakar D, Adibi MA. Genetic structure and ecological niche segregation of Indian Gray mongoose (*Urva edwardsii*) in Iran. Ecol Evol. 2021. https://doi.org/10.1002/ece3.8168.
89. Niwa K, Tran D, Van, Nishikawa K. Differentiated historical demography and ecological niche forming present distribution and genetic structure in coexisting two salamanders (Amphibia, urodela, Hynobiidae) in a small island, Japan. PeerJ. 2022;10:e13202.
90. Tran D, Van, Tominaga A, Pham LT, Nishikawa K. Ecological niche modeling shed light on new insights of the speciation processes and historical distribution of Japanese fire-bellied Newt *Cynops pyrrhogaster* (Amphibia: Urodela). Ecol Inf. 2024;79:102443. https://doi.org/10.1016/j.ecoinf.2023.102443.
91. Pyron RA, Pirro S, Hains T, Colston TJ, Myers EA, O'Connell KA et al. The draft genome sequences of 50 salamander species (Caudata, Amphibia). Biodiversity Genomes. 2024. https://doi.org/10.56179/001c.116891.
92. Pyron RA, O'Connell KA, Lemmon EM, Lemmon AR, Beamer DA. Phylogenomic data reveal reticulation and incongruence among mitochondrial candidate species in Dusky salamanders (Desmognathus). Mol Phylogenet Evol. 2020;146:106751. https://doi.org/10.1016/j.ympev.2020.106751.
93. Burgon JD, Vences M, Steinfartz S, Bogaerts S, Bonato L, Donaire-Barroso D et al. Phylogenomic inference of species and subspecies diversity in the Palearctic salamander genus *Salamandra*. Mol Phylogenet Evol. 2021;157:107063. https://doi.org/10.1016/j.ympev.2020.107063.
94. Pyron RA, O'connell KA, Duncan SC, Burbrink FT, Beamer DA. Speciation hypotheses from phylogeographic delimitation yield an integrative taxonomy for seal salamanders (*Desmognathus monticola*). Syst Biol. 2023. https://doi.org/10.1093/sysbio/syac065.
95. Talavera A, Palmada-Flores M, Burriel-Carranza B, Valbuena-Ureña E, Mochales-Riaño G, Adams DC, et al. Genomic insights into the Montseny brook newt (*Calotriton arnoldi*), a critically endangered glacial relict. iScience. 2024. https://doi.org/10.1016/j.isci.2023.108665.
96. Borzée A, Shin Y, Poyarkov NA, Jeon JY, Baek HJ, Lee CH, et al. Dwindling in the mountains: description of a critically endangered and microendemic *Onychodactylus* species (Amphibia, Hynobiidae) from the Korean Peninsula. Zool Res. 2022;43:750–5. https://doi.org/10.24272/j.issn.2095-8137.2022.048.
97. Rhoden CM, Peterman WE, Taylor CA. Maxent-directed field surveys identify new populations of narrowly endemic habitat specialists. PeerJ. 2017;5:e3632. https://doi.org/10.7717/peerj.3632.
98. Sakai Y, Kusakabe A, Tsuchida K, Tsuzuku Y, Okada S, Kitamura T, et al. Discovery of an unrecorded population of Yamato salamander (*Hynobius vandenburghi*) by GIS and edna analysis. Environ DNA. 2019. https://doi.org/10.1002/edn3.31.

Publisher's Note