quaternary

MDPI

Article

Postglacial Range Dynamics of the Common Roach (*Rutilus rutilus*) in the Baltic Region: Integrating Palaeodistribution Modelling and Mitochondrial Haplogroup Geography

Ieva Ignatavičienė [1,*], Dace Grauda [2], Dmitri Gudkov [3] and Dalius Butkauskas [1,2]

[1] State Scientific Research Institute Nature Research Centre, Akademijos 2, 08412 Vilnius, Lithuania; dalius.butkauskas@gamtc.lt
[2] Institute of Biology, Faculty of Medicine and Life Sciences, University of Latvia, Jelgavas 1, LV-1004 Riga, Latvia; dace.grauda@lu.lv
[3] Institute of Hydrobiology of the NAS of Ukraine, Geroyev Stalingrada Ave. 12, UA-04210 Kiev, Ukraine; digudkov@gmail.com
* Correspondence: ieva.ignataviciene@gamtc.lt; Tel.: +370-65926-290

Abstract

Quaternary climatic cycles reshaped freshwater fish distributions, with postglacial recolonisation influenced by climatic suitability and hydrological connectivity. We reconstructed the postglacial history of the common roach (*Rutilus rutilus*) in the Baltic region by integrating palaeodistribution modelling and mitochondrial haplogroup geography. A MaxEnt model calibrated with present-day occurrence records was projected onto eight periods from the Last Glacial Maximum (LGM) to the present. Mitochondrial haplogroups were inferred from concatenated ATP6 and D-loop sequences from three Lithuanian and Latvian drainages and the Ukrainian Dnieper River basin. Projected climatically suitable area changed non-monotonically, reaching its minimum during Heinrich Stadial 1 and increasing approximately 14-fold between the LGM and the present. Using the P10 suitability threshold (0.481) and MESS-masking criteria, no candidate cold-stage climatic refugia were predicted in the Baltic region. Haplogroup A predominated in the Lithuanian and Latvian drainages and comprised 96.7% of the Ukrainian Dnieper basin sample. It also occurred in Portuguese, British, Czech and Romanian references, whereas haplogroup D occurred only in Russian and Chinese reference sequences. Together, the results are consistent with postglacial recolonisation from populations outside the Baltic region, potentially facilitated during freshwater stages of the Baltic basin, but do not identify a unique refugial source or dispersal route.

Keywords: *Rutilus rutilus*; Quaternary; postglacial recolonisation; species distribution modelling; palaeodistribution; mitochondrial haplogroups; Baltic Sea Basin; freshwater connectivity

Check for updates

Academic Editor: Jane M. Bunting

Received: 22 September 2026
Revised: 1 November 2026
Accepted: 3 November 2026
Published: 5 November 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Repeated glacial–interglacial cycles during the Quaternary altered species distributions across Europe [1]. During glacial periods, many temperate species persisted in climatically suitable refugia, often in southern Europe, although persistence in extra-Mediterranean refugia has also been documented [1–3]. As ice sheets retreated, newly deglaciated freshwater habitats became available, while changes in proglacial lakes, meltwater channels and river networks either facilitated or restricted dispersal among freshwater systems [4]. During and following deglaciation, the Baltic Sea Basin underwent a well-documented

succession of lacustrine and marine-influenced stages, including the Baltic Ice Lake, Yoldia Sea, Ancylus Lake and Littorina Sea, accompanied by major changes in water level, salinity and connections with the North Sea and surrounding drainage systems [5–8]. These palaeohydrological changes repeatedly altered the availability of aquatic dispersal routes and the environmental conditions encountered by colonizing freshwater organisms.

The common roach (*Rutilus rutilus*) is a widespread, eurytopic freshwater fish inhabiting rivers and lakes, as well as low-salinity brackish coastal waters, across Europe and western Asia [9]. A previous phylogeographic analysis based on mitochondrial cytochrome *b* sequences identified two deeply divergent clades, designated the West European and Ponto-Caspian clades [10]. The few samples from Baltic Sea drainages included in that study fell within the West European clade, for which the Danube and Dniester basins were proposed as possible palaeorefugia [10]. Comparable studies of Eurasian perch (*Perca fluviatilis*) inferred colonisation of the Baltic region from refugia in western, south-eastern and north-eastern Europe [11]. Subsequent genomic work documented secondary contact among perch lineages in the Baltic region and provided evidence consistent with a possible late-glacial freshwater corridor between the Black Sea and northern Europe [12]. Although mitochondrial variation among Lithuanian drainage basins has recently been characterised [13], its broader geographic context and relationship with changes in climatically suitable area during the Late Glacial and Holocene remain incompletely resolved.

Species distribution models (SDMs) calibrated using present-day occurrence records and projected onto palaeoclimatic reconstructions provide spatially explicit estimates of climatic suitability through time, allowing candidate glacial refugia and potential directions of range expansion to be identified [14]. The geographic distribution of mitochondrial variation can then be used to assess whether genetic patterns correspond to the refugial areas and expansion scenarios suggested by palaeodistribution models [14,15]. To our knowledge, these approaches have not previously been combined to investigate the postglacial history of the common roach in the Baltic region. Here, we integrate SDM hindcasting with the geographic distribution of mitochondrial haplogroups to examine the postglacial history of *R. rutilus* in the Baltic region within a broader European context. Specifically, we (i) reconstruct climatically suitable habitat across eight climatic periods spanning the Last Glacial Maximum (LGM) to the present and quantify changes in suitable area and the geographic position of the suitable-area centroid, and (ii) characterize the geographic distribution of haplogroups inferred from concatenated mitochondrial ATP6 and D-loop sequences across three Lithuanian drainage basins and the Ukrainian Dnieper River basin, supplemented with reference sequences from other parts of the species' range.

2. Materials and Methods

2.1. Study Area and Occurrence Data

The study combined species distribution modelling (SDM) of the common roach across a study area covering Europe and western Asia with an analysis of mitochondrial haplogroups inferred from concatenated ATP6 and D-loop sequences. The SDM domain extended from 15° W to 60° E and from 35° to 72° N (WGS84). The genetic dataset comprised samples from the Nemunas, Daugava and Lielupe drainage basins in Lithuania and Latvia and the Dnieper basin in Ukraine (water bodies within the Chernobyl Exclusion Zone and the Ros' River sites; Figure 1a), supplemented with reference sequences obtained from GenBank (Table S1).

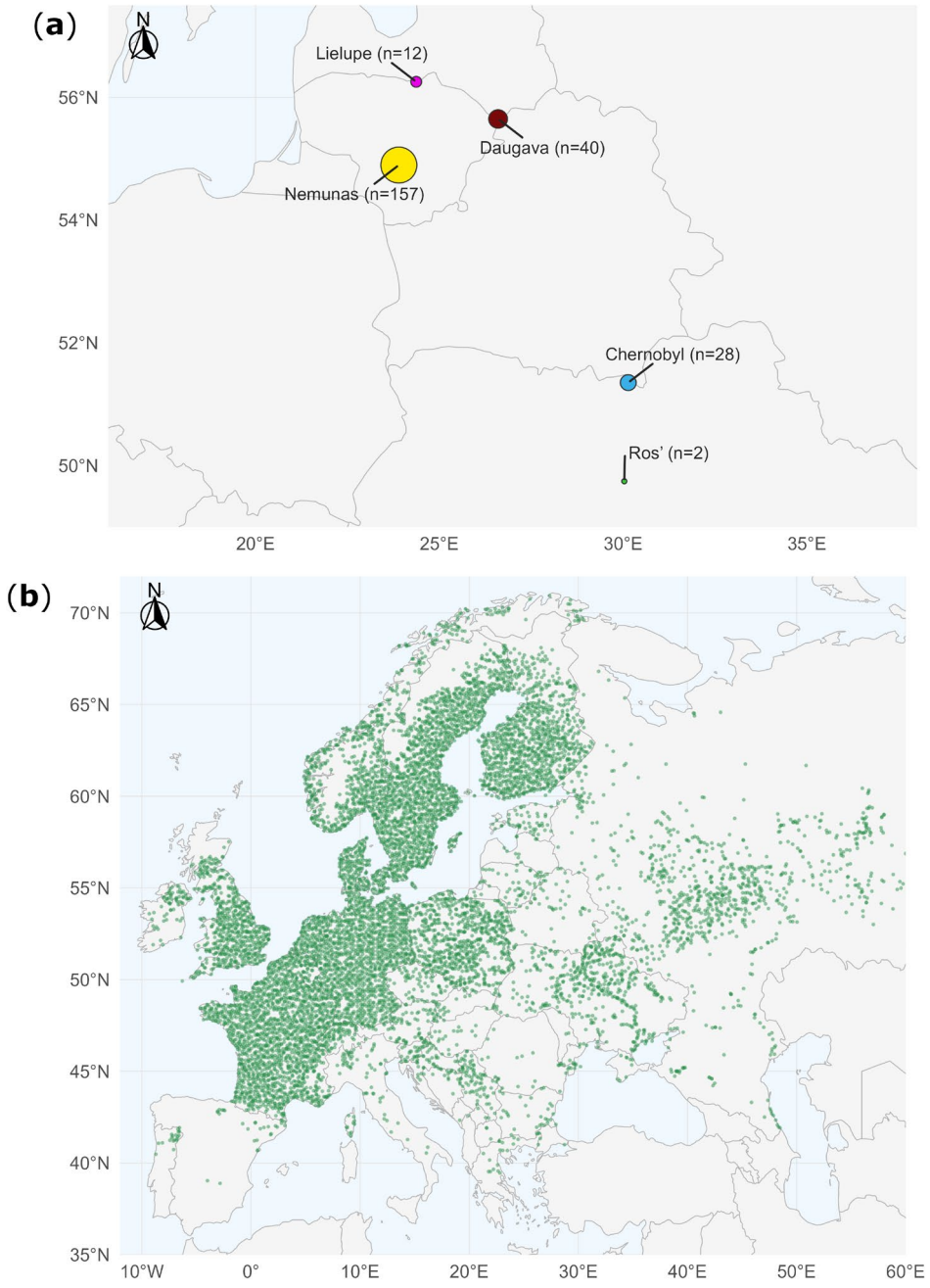


Figure 1. Study design. (**a**) Genetic sampling of *R. rutilus* in the Nemunas, Daugava and Lielupe drainage basins and at two Dnieper-basin sites (water bodies in the Chernobyl Exclusion Zone and the Ros' River). Symbol area is proportional to the aggregated sample size (*n*) within each basin or site. Colours distinguish the individual drainage basins and sampling sites. (**b**) Spatially filtered GBIF occurrence records retained within the European–western Asian study window before predictor-specific exclusion.

Georeferenced occurrence records of *R. rutilus* were obtained from the Global Biodiversity Information Facility (GBIF) [16]. Records were restricted to the study window and cleaned by removing duplicate coordinates, records with missing coordinates and records located at 0° latitude and 0° longitude. The cleaned dataset contained 118,685 records. To reduce spatial sampling bias, occurrence records were spatially filtered by retaining one record per occupied 0.2° × 0.2° grid cell [17], resulting in 9412 localities (Figure 1b). Of these, 9126 localities had complete values for all 19 bioclimatic variables and were used for predictor-correlation analysis.

2.2. Palaeoclimate Data and Predictor Selection

Bioclimatic variables (BIO1–BIO19) for collinearity screening were obtained at 2.5 arc-min resolution from WorldClim 2.1 [18]. The palaeoclimatic layers used for model fitting and projection were obtained at the same resolution from the PaleoClim database [19] using the rpaleoclim R package v1.1.0 [20]. Eight climatic datasets spanning the Last Glacial Maximum to the present were used. The CHELSA-derived datasets represented the Last Glacial Maximum (LGM; ~21 ka) and the present (1979–2013) and were reprocessed by PaleoClim to match the format of the other layers [21]. The remaining six datasets were based on PaleoView climate simulations: Heinrich Stadial 1 (HS1; 17.0–14.7 ka), the Bølling–Allerød interstadial (BA; 14.7–12.9 ka), the Younger Dryas stadial (YD; 12.9–11.7 ka), the Early Holocene (EH; 11.7–8.326 ka), the Mid Holocene (MH; 8.326–4.2 ka) and the Late Holocene (LH; 4.2–0.3 ka) [22]. The approximate ages used as labels in the text and figures (~21, ~16, ~14, ~12, ~10, ~6 and ~2 ka BP) denote the LGM time slice (~21 ka BP) and the rounded midpoints of the six PaleoView intervals listed above. The interval boundaries follow the period definitions used in PaleoClim and PaleoView [19,22].

To reduce multicollinearity, pairwise Pearson correlation coefficients were calculated among all 19 bioclimatic variables at the 9126 occurrence localities with complete environmental data [23]. Predictor variables were filtered using the find Correlation function in the caret R package v7.0-1 [24]. When the absolute pairwise correlation between two predictors exceeded 0.7 ($|r| > 0.7$), the predictor with the higher mean absolute correlation with the remaining predictors was removed. This procedure retained six variables: mean diurnal range (BIO2), isothermality (BIO3), mean temperature of the wettest quarter (BIO8), mean temperature of the warmest quarter (BIO10), precipitation of the wettest month (BIO13) and precipitation seasonality (BIO15). Among the four bioclimatic variables that combine temperature and precipitation information and may exhibit spatial discontinuities associated with changes in the quarter used for their calculation (BIO8, BIO9, BIO18 and BIO19) [25,26], only BIO8 remained after correlation filtering. BIO8 was subsequently excluded before model fitting to reduce the risk of artefacts when transferring the model across palaeoclimatic periods. The final predictor set comprised BIO2, BIO3, BIO10, BIO13 and BIO15. Although BIO3 is mathematically derived from BIO2 and annual temperature range (BIO3 = BIO2/BIO7 × 100) [27], the empirical correlation between BIO2 and BIO3 was below the predefined threshold, and both variables were retained.

2.3. Species Distribution Modelling and Evaluation

We selected MaxEnt because it is a presence–background method that can be applied to the opportunistic, presence-only occurrence records available for *R. rutilus* and does not require confirmed absence data. Although incomplete and spatially uneven records may affect model predictions, MaxEnt is widely used in species distribution modelling [28] and has also been applied in palaeodistribution studies [14], facilitating comparison with previous work. The climatic niche of *R. rutilus* was modelled using the maximum-entropy algorithm implemented in MaxEnt v3.4.3 [28] through the dismo R package v1.3-16 [29] in R v4.5.1 [30]. Of the 9412 spatially thinned occurrence records, 9322 had complete data for all five selected predictors and were included in the modelling workflow. The remaining 90 records occurred in cells with missing values for at least one predictor. The correlation analysis was based on the smaller subset of 9126 records with complete data for all 19 candidate predictors. Of the 9322 records available for modelling, 6992 (75%) were randomly assigned to the training dataset and 2330 (25%) were withheld for model evaluation. The model was fitted using linear, quadratic, product and hinge features, default regularisation settings and 10,000 background points sampled within the study area. Model output was expressed on the complementary log–log (cloglog) scale [31], bounded between

0 and 1 and interpreted as a relative index of climatic suitability rather than as absolute probability of occurrence.

Model performance was assessed using the area under the receiver operating characteristic curve (AUC) for the training and withheld test datasets. The Continuous Boyce Index (CBI) was calculated from the present-day climatic-suitability prediction using the ecospat.boyce function in the ecospat R package v4.1.4 [32]. The CBI evaluates whether the predicted-to-expected frequency of occurrence records increases monotonically along the suitability gradient. For the CBI, presence observations comprised all spatially thinned occurrence records with non-missing predicted suitability, including both training and test records. The available prediction distribution was defined by the suitability values of all non-missing raster cells within the study area. The CBI was calculated using a moving window (nclass = 0) and the function's default settings for window width, resolution and removal of successive duplicate predicted-to-expected ratios, with the predicted-to-expected ratio evaluated using Spearman correlation [33,34]. Because the CBI was computed from both training and test records, and because occurrence records were partitioned randomly rather than spatially, both AUC and CBI were interpreted as complementary measures of internal model fit rather than as spatially independent validation [35]. Predictor importance was assessed using MaxEnt percent contribution, permutation importance and the jackknife test of regularised training gain [31].

2.4. Hindcast Projections, Range Dynamics and Climatic Stability

The model calibrated under present-day climatic conditions was projected onto seven palaeoclimatic datasets, while the present-day prediction was retained as the baseline, yielding predictions for eight climatic periods in total. For each palaeoclimatic projection, a multivariate environmental similarity surface (MESS) [36] was calculated by comparing the values of the five predictors in each cell with their reference distributions in the present-day predictor layers. Cells with MESS < 0, indicating that at least one predictor fell outside its present-day reference range, were masked before calculating suitable area, area- and suitability-weighted centroids and candidate refugia.

To quantify changes in climatically suitable area through time, each MESS-masked continuous prediction was converted into a binary suitable–unsuitable map using the 10th-percentile training-presence threshold calculated from cloglog predictions for the training occurrences under present-day climatic conditions (P10 = 0.481) [37]. The same threshold was applied to all eight periods, with unmasked cells having cloglog suitability values ≥ 0.481 classified as suitable and those with lower values classified as unsuitable. Suitable area (km^2) was calculated for each period by summing the areas of all cells classified as suitable. Individual grid-cell areas were calculated in square kilometres using the cellSize function in the terra R package v1.9-27 [29], thereby accounting for variation in cell area with latitude.

The geographic position of projected climatic suitability in each period was summarised using an area- and suitability-weighted centroid, calculated from all MESS-unmasked cells as the weighted mean of their longitudes and latitudes, each cell weighted by the product of its continuous cloglog suitability value and area. Range-shift velocity was then calculated as the great-circle (haversine) distance between the centroids of successive periods, divided by the elapsed time between them, and expressed in metres per year.

The Last Glacial Maximum, Heinrich Stadial 1 and the Younger Dryas were selected to represent the three major cold intervals included in the analysed palaeoclimatic datasets [38]. Model-based candidate climatic refugia were defined as cells that remained MESS-unmasked and exceeded the P10 cloglog threshold in all three cold-stage projections. This criterion identified areas predicted to be suitable in all three selected time slices, but it

did not demonstrate continuous suitability between them or confirm the historical persistence of *R. rutilus* populations. Therefore, these areas were interpreted only as candidate climatic refugia requiring support from independent evidence [39].

2.5. Genetic Sampling and Haplogroup Assignment

R. rutilus were sampled from the Nemunas ($n = 157$), Daugava ($n = 40$) and Lielupe ($n = 12$) drainage basins and from the Dnieper basin in Ukraine ($n = 30$) between 2017 and 2026 (Figure 1; Table S1). The Ukrainian material comprised samples from the Chernobyl area ($n = 28$) and the Ros' River ($n = 2$). The 209 Lithuanian specimens and their ATP6–D-loop sequences were previously reported in the companion study [13], where they were collected as part of the national fish-monitoring programme using the standardised methods described therein. Ukrainian samples were provided by Dmitri Gudkov from the Institute of Hydrobiology of the NAS of Ukraine.

Genomic DNA was extracted from frozen muscle tissue following previously described protocols [13,40]. The mitochondrial ATP6 and D-loop regions were amplified separately and sequenced as described previously [13,40,41]. The same molecular procedures were applied to the Ukrainian samples. The final aligned fragment lengths were 454 bp for ATP6 and 368 bp for the D-loop.

Publicly available sequences from the United Kingdom, Portugal, the Czech Republic, Romania, Russia and China were retrieved from GenBank [42]. Reference sequences were included only when both mitochondrial regions originated from the same individual or could be extracted from a complete mitochondrial genome, and each was trimmed to the same ATP6 and D-loop regions used for the study sequences before concatenation. Accession numbers and geographic origins are provided in Table S1. Because these records represent individual reference sequences rather than population samples, they were used only to place the Baltic and Ukrainian samples within a broader Eurasian context.

The ATP6 and D-loop sequences from the Lithuanian and Ukrainian samples and the GenBank references were aligned separately using MUSCLE [43] in MEGA X [44] and then concatenated by individual to produce an 822-bp ATP6–D-loop alignment. Haplotype relationships were reconstructed from the combined dataset using a median-joining network [45] in NETWORK v10.2.0.0 (Fluxus Technology Ltd., Essex, UK). Haplogroup labels A–D were assigned by matching clusters in the expanded network to the descriptive groups defined by Ignatavičienė et al. [13], with the Ukrainian and GenBank reference sequences assigned according to their positions within that network. These haplogroups were treated as descriptive genetic clusters and do not represent phylogenetically validated mitochondrial lineages.

Haplogroup composition was mapped in R v4.5.1 using the sf v1.1.1 [46,47], rnaturalearth v1.2.0 [48], scatterpie v0.2.6 [49] and ggspatial v1.1.10 [50] packages. All spatial layers were transformed to ETRS89/LAEA Europe (EPSG:3035). For mapping, samples from multiple localities were pooled within the Nemunas and Daugava drainage basins. The Lielupe sample, which represented a single locality, was displayed separately, as were the Chernobyl-area and the Ros' River samples from Ukraine and each GenBank reference locality. Pie charts represented haplogroup proportions, and their radii were scaled in proportion to the square root of sample size, subject to a minimum radius to ensure the visibility of samples represented by one or two sequences.

3. Results

3.1. Model Performance, Present-Day Climatic Suitability and Predictor Importance

The MaxEnt model yielded closely matched training and test AUC values of 0.686 and 0.682 (SD = 0.005), against a maximum achievable test AUC of 0.677 for this dataset

(Table S2). The Continuous Boyce Index was 0.996. Under the present-climate baseline (1979–2013), predicted climatic suitability was highest across the British Isles and much of western and central Europe and was moderate to high in parts of southern Fennoscandia. Suitability generally declined eastwards across the East European Plain and western Asia and was also relatively low across much of the Mediterranean and south-eastern parts of the study area (Figure 2a).

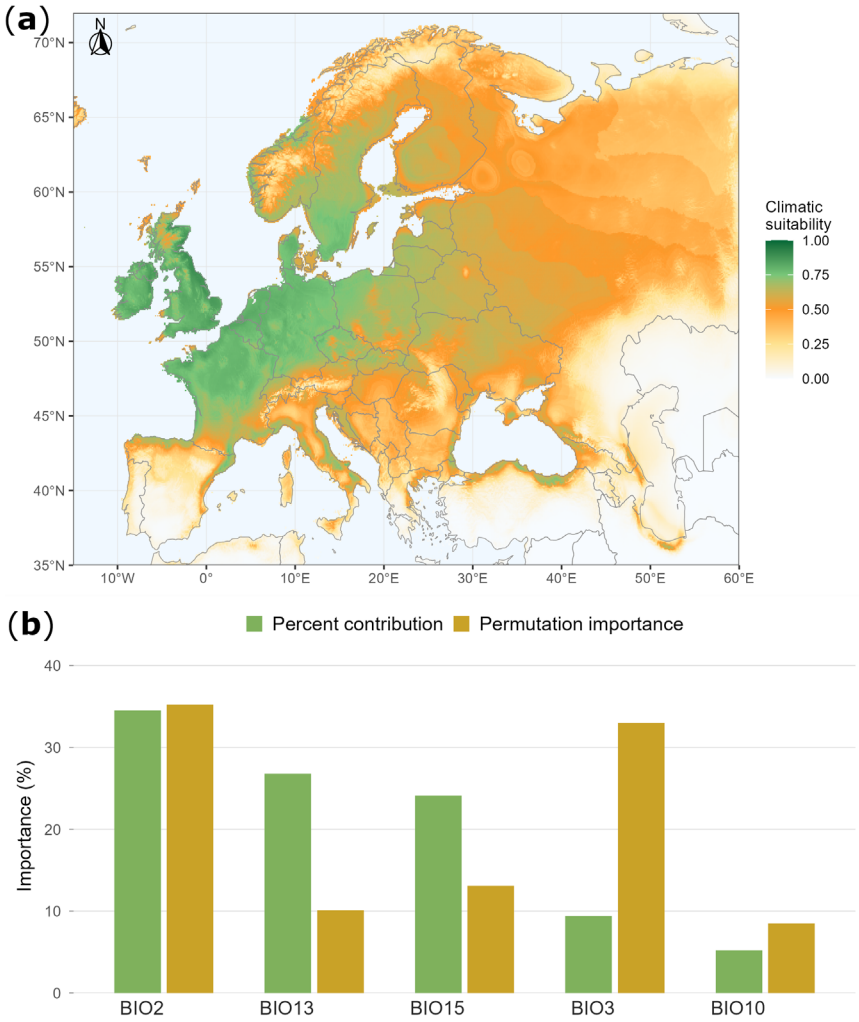


Figure 2. MaxEnt model of present-climate suitability for *R. rutilus* based on the 1979–2013 climatic baseline. (**a**) Predicted climatic suitability (0–1). (**b**) Predictor importance (percent contribution and permutation importance). BIO2, mean diurnal range; BIO13, precipitation of the wettest month; BIO15, precipitation seasonality; BIO3, isothermality; BIO10, mean temperature of the warmest quarter.

Mean diurnal range (BIO2) ranked first according to both MaxEnt importance measures, with the highest percent contribution (34.5%) and permutation importance (35.2%) (Figure 2b; Table S3). Precipitation of the wettest month (BIO13) and precipitation seasonality (BIO15) ranked second and third according to percent contribution (26.8% and 24.1%, respectively), although their permutation importance values were lower (10.1% and 13.1%, respectively). Isothermality (BIO3) and mean temperature of the warmest quarter (BIO10) had the lowest percent contributions (9.4% and 5.2%, respectively), whereas BIO3 ranked second according to permutation importance (33.0%). In the jackknife test of regularised training gain, BIO10 produced the highest gain when used alone, whereas omitting BIO2 caused the largest reduction in gain (Figure S1).

3.2. Palaeodistribution, Range Dynamics and Candidate Refugia Since the LGM

Projected climatic suitability varied in magnitude and spatial distribution among the seven palaeoclimatic time slices and the present-day baseline (Figure 3). During the Last Glacial Maximum (LGM; ~21 ka BP), relatively high predicted suitability was concentrated mainly in western and southern Europe, particularly along the Atlantic margins and in parts of the Italian Peninsula and northern Adriatic region. Climatically suitable areas contracted substantially during Heinrich Stadial 1 (HS1; ~16 ka BP) and expanded again during the Bølling–Allerød interstadial (~14 ka BP), particularly in western Europe and parts of Fennoscandia. This expansion was followed by renewed contraction during the Younger Dryas (~12 ka BP). From the Early Holocene onward, climatically suitable areas expanded progressively across central, northern and eastern Europe. During the Mid and Late Holocene, the spatial pattern of projected climatic suitability became increasingly similar to the present-day pattern. Cells with non-analog climatic conditions (MESS < 0) were generally most extensive in the earlier projections, particularly during the LGM and HS1, and became less extensive in the later periods.

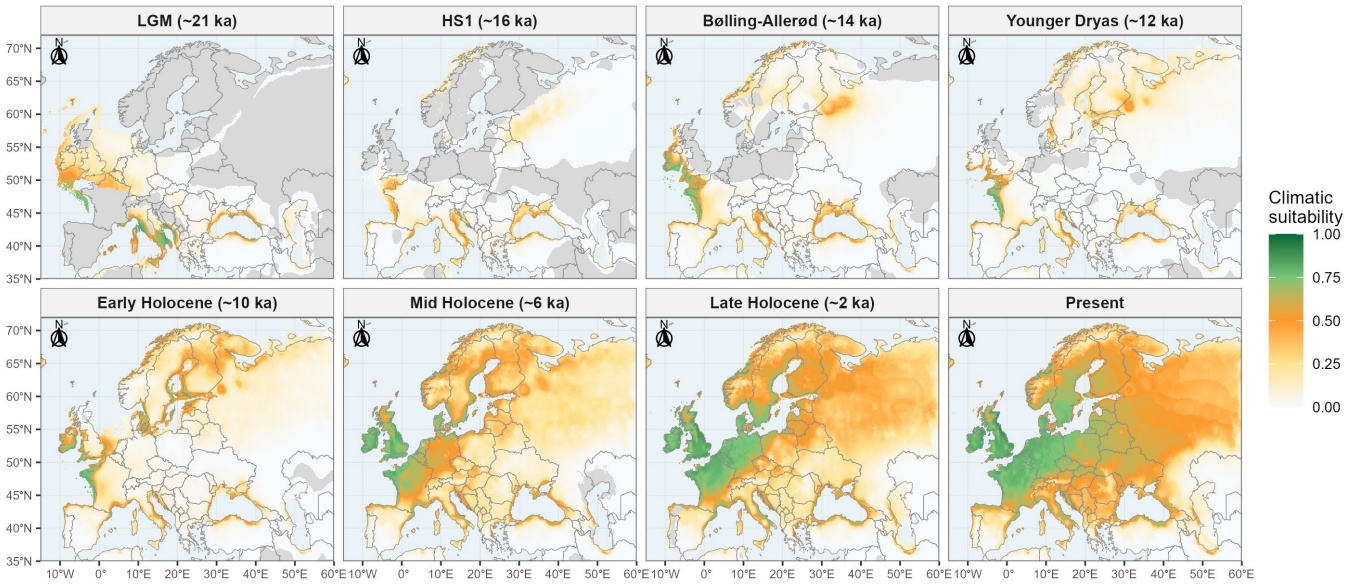


Figure 3. Hindcast climatic suitability for *R. rutilus* across eight periods from the LGM to the present. Climatically novel areas (MESS < 0) were masked (grey).

The projected climatically suitable area varied markedly among the eight time slices (Table S4; Figure 4a). It decreased from ~0.37 × 10^6 km^2 during the LGM to a minimum of ~0.05 × 10^6 km^2 during HS1. Suitable area subsequently increased to ~0.34 × 10^6 km^2 during the Bølling–Allerød and decreased again to ~0.19 × 10^6 km^2 during the Younger Dryas. From the Early Holocene onward, it increased consistently, reaching ~0.55, 1.62 and 2.80 × 10^6 km^2 during the Early, Mid and Late Holocene, respectively, and ~5.29 × 10^6 km^2 under present-day climatic conditions. The present-day estimate was therefore ~14.3 times greater than the LGM estimate.

The suitability-weighted centroid showed an overall shift towards the north and east (Figure 4b,c). Its latitude increased from 46.7° N during the LGM to a maximum of 54.6° N during the Early Holocene, before decreasing slightly to 53.1° N in the present-day projection. Its longitude increased from 11.1° E during the LGM to 24.3° E in the present-day projection.

Between the LGM and the present, the suitability-weighted centroid shifted by ~1180 km, corresponding to an average net shift of about 56 m yr^{-1} over 21 kyr. The

highest centroid-displacement rate was ~215 m yr^{-1} between the Bølling–Allerød and the Younger Dryas.

Candidate cold-stage refugial areas, defined as cells classified as suitable and environmentally analogous during all three selected cold stages—the LGM, HS1 and the Younger Dryas—covered 10,445 km^2 (Figure 5). These areas were concentrated in small parts of the Italian Peninsula and the northern Black Sea coastal region. Under the applied criterion, no candidate climatic refugia were identified in the Baltic region.

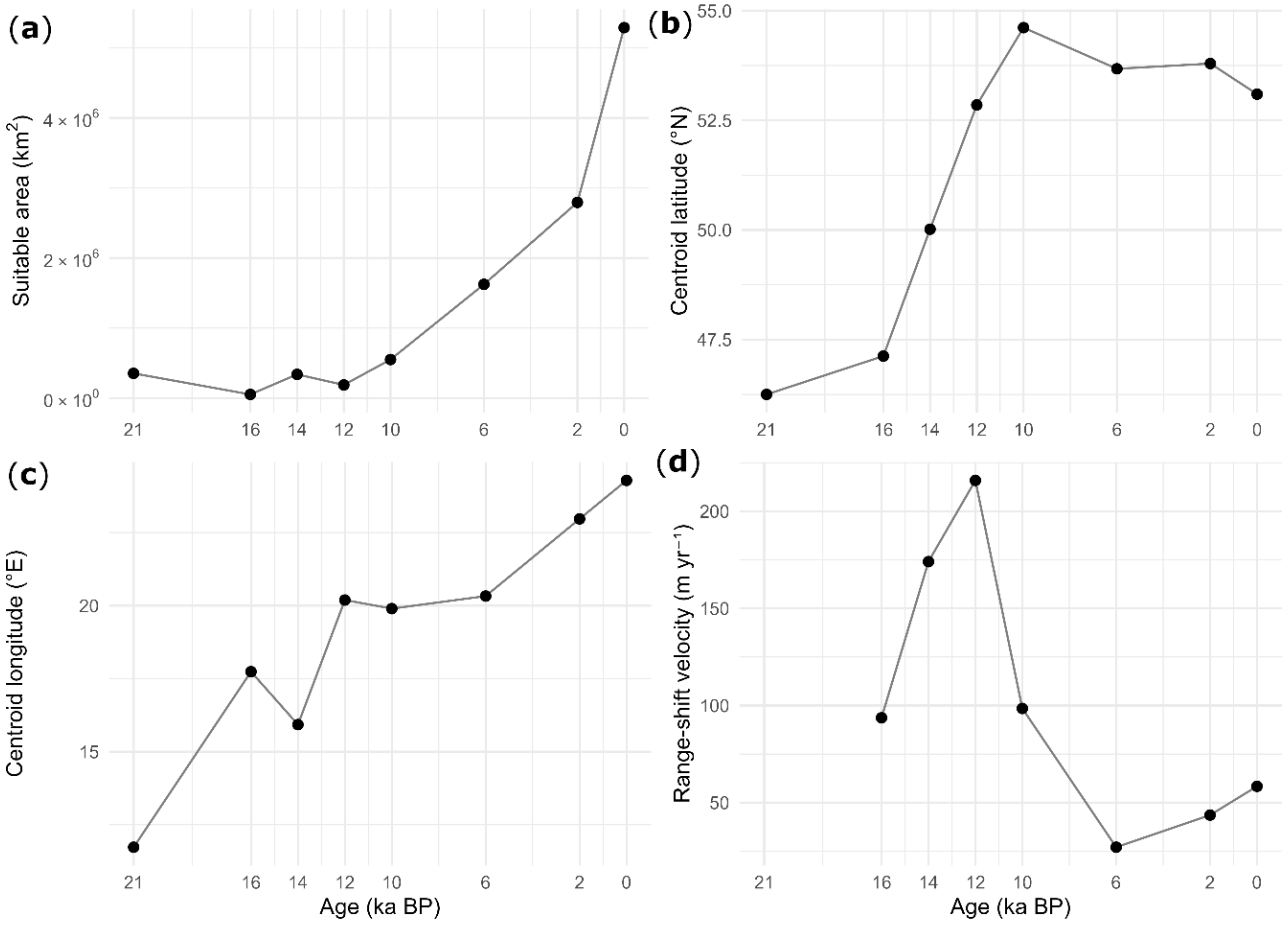


Figure 4. Range dynamics of climatically suitable habitat for *R. rutilus* across eight periods from the LGM to the present. (**a**) Suitable area; (**b**,**c**) area- and suitability-weighted centroid latitude and longitude; (**d**) range-shift velocity (great-circle displacement of the centroid between successive periods divided by the elapsed time).

3.3. Geographic Distribution of mtDNA Haplogroups

Across the combined study and reference dataset, mitochondrial haplogroup A was the most frequent and geographically widespread of the four haplogroups (Figure 6; Figure S2). It was the dominant haplogroup in all three drainage basins, accounting for 78.3% of individuals in the Nemunas basin (n = 157), 55.0% in the Daugava basin (n = 40) and 83.3% in the Lielupe basin (n = 12). Haplogroup A also represented 96.7% of the Ukrainian Dnieper basin samples (n = 30). All five reference sequences from Portugal, the United Kingdom, Czechia and Romania (the last represented by two sequences) were assigned to haplogroup A. Because sample sizes varied among basins, basin-sample haplogroup proportions are reported with 95% Wilson score confidence intervals (Table S1). The intervals are particularly wide for the smaller Lielupe sample (n = 12). The Ukrainian Dnieper sample also includes a very small Ros' subsample (n = 2).

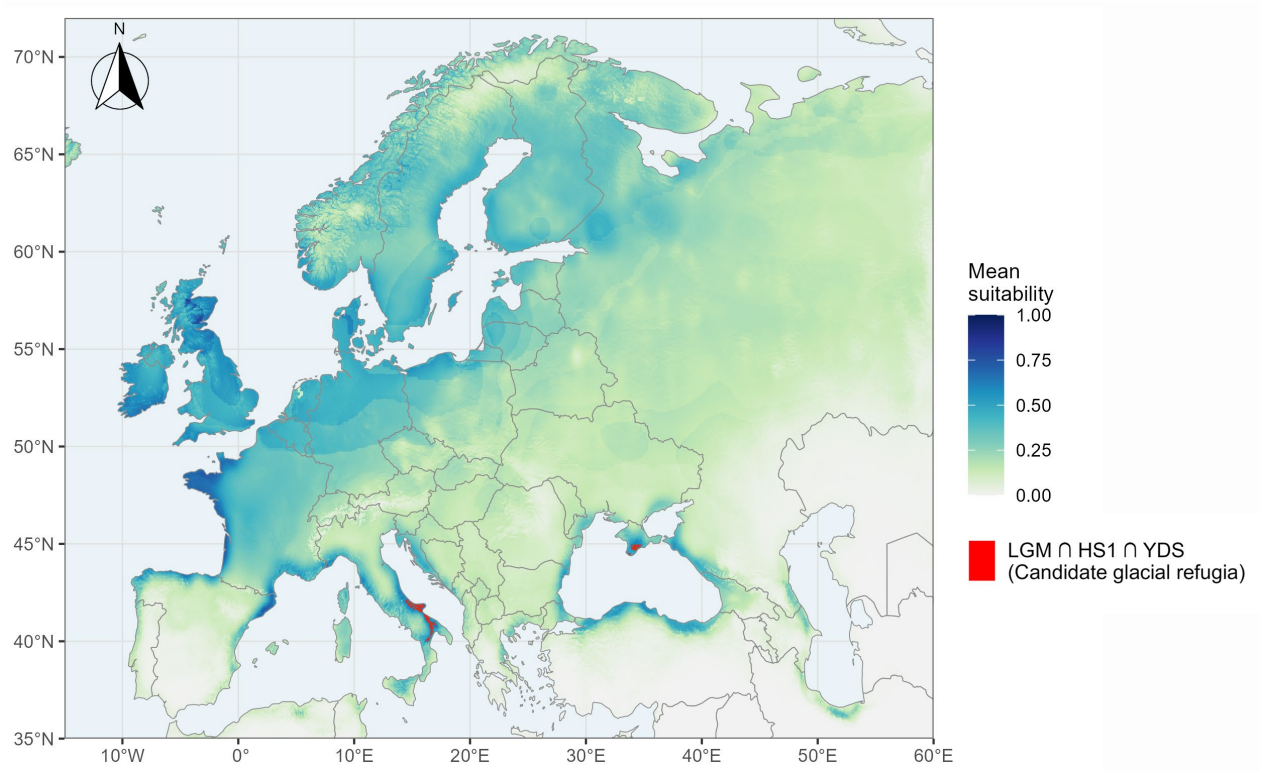


Figure 5. Candidate glacial refugia for *R. rutilus*. Candidate glacial refugia (red) are cells that remained suitable (cloglog suitability ≥ P10 (0.481)) and within the present-day predictor ranges (MESS ≥ 0) during all three cold stages (LGM, HS1 and the Younger Dryas). The background shows the mean climatic suitability across the eight periods, from low (pale) to high (dark).

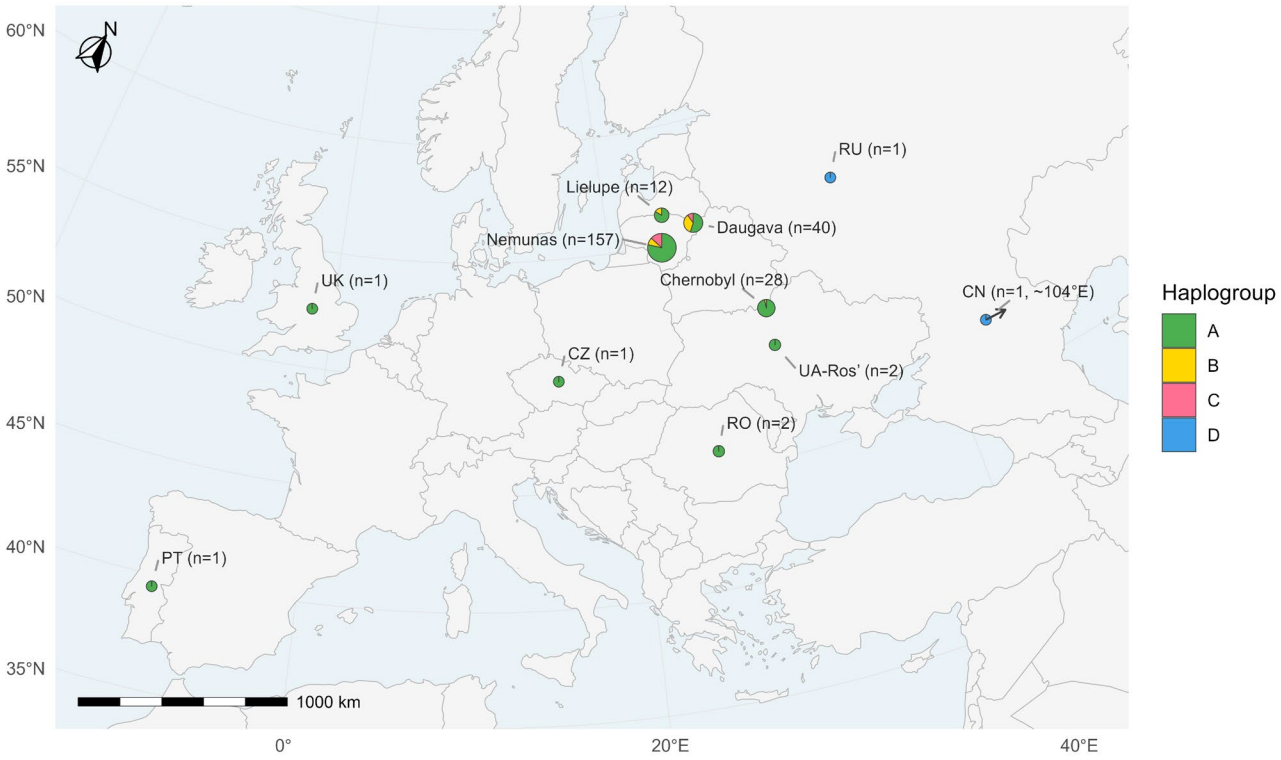


Figure 6. Geographic distribution of mitochondrial haplogroups (A–D) of *R. rutilus*. Sampling groups are shown as pie charts of haplogroup composition, with symbol radius scaled to the square root of sample size. Samples from multiple localities are pooled within the Nemunas and Daugava basins

and Chernobyl-area. The Lielupe and Ros' River samples are shown separately, as are the single- or double-sequence GenBank references (UK, PT, CZ, RO, RU, CN), which are plotted at approximate country-level positions because precise coordinates were unavailable. The Chinese haplogroup D reference (~104° E) lies outside the map extent and is indicated by an arrow at the margin. Country codes: LT, Lithuania; UA, Ukraine; UK, United Kingdom; PT, Portugal; CZ, Czechia; RO, Romania; RU, Russia; CN, China. Projection: ETRS89-LAEA Europe (EPSG:3035).

Haplogroup B was the second most frequent haplogroup among the Lithuanian, Latvian and Ukrainian samples. Its frequency was highest in the Daugava basin (35.0%), followed by the Lielupe (16.7%) and Nemunas (8.3%) basins, and it was detected in a single Ukrainian individual. Haplogroup C occurred only in the Nemunas (13.4%) and Daugava (10.0%) basins. Haplogroup D was not detected in any Lithuanian, Latvian or Ukrainian sample and was represented only by the two GenBank reference sequences from Russia and China.

4. Discussion

4.1. Model Performance and Variable Importance

The test AUC was 0.682, exceeding the random expectation of 0.5 and closely matching the maximum achievable test AUC estimated by MaxEnt (0.677), indicating that the model's discrimination was close to the maximum expected under the presence–background evaluation given the extensive overlap between occurrence and background climates. Because true absence data are unavailable, MaxEnt defines the false-positive rate using the fractional predicted area rather than the observed commission rate. Consequently, the maximum achievable AUC for presence–background data may be below 1 and generally decreases as the proportion of the study area occupied by the species increases. MaxEnt also reports a model-specific estimate of the maximum possible test AUC, assuming that test localities are drawn from the model's predicted distribution [28]. Such overlap is expected for a species with the wide geographic distribution and broad ecological tolerance of *R. rutilus* [9], and presence–background AUC is, in any case, not a purely species-specific measure of model quality [28,51,52]. The high CBI (0.996) showed a strong positive relationship between predicted climatic suitability and the predicted-to-expected frequency of occurrences. The CBI is suitable for evaluating presence-only models because it does not require absence data [33,34]. Together, the two metrics indicate that the model generally ranked locations from lower to higher predicted climatic suitability. However, they do not by themselves demonstrate that the predictions accurately represent the ecology of *R. rutilus*.

Percent contribution and permutation importance describe different aspects of predictor importance in MaxEnt and are therefore not directly interchangeable. Percent contribution is a path-dependent measure of a predictor's contribution during model training, whereas permutation importance measures how strongly the fitted model depends on that predictor [53,54]. Isothermality (BIO3) showed relatively low percent contribution but high permutation importance. Because BIO3 shares climatic information with mean diurnal range (BIO2), this contrast may reflect overlapping information or nonlinear relationships captured by the model. Even though the pairwise correlation between the two predictors did not exceed the filtering threshold. The importance measures alone, however, cannot establish its cause. The opposite pattern shown by precipitation of the wettest month (BIO13), which had a high percent contribution but lower permutation importance, further illustrates that the two measures capture different properties of the fitted model. BIO2 ranked highest on both measures, indicating a strong influence on model predictions. Nevertheless, these values describe the behaviour of the fitted model and should not be taken as evidence that individual climatic variables causally determine the distribution of *R. rutilus* [55].

4.2. Postglacial Dynamics of Climatically Suitable Area

The hindcast projections indicated a substantial net expansion of climatically suitable area for *R. rutilus* from the LGM to the present, accompanied by an overall northward and eastward displacement of the suitability-weighted centroid. These patterns represent changes in modelled climatic suitability rather than directly observed shifts in the distribution of the species. The broad climatic pattern consists of postglacial warming and the retreat of the Fennoscandian Ice Sheet. However, the extent to which *R. rutilus* was able to track this shifting climatic envelope would also have depended on the palaeohydrological reorganisation of northern European drainage systems and the availability of freshwater connections [5,8,56]. Molecular phylogeographic studies have similarly inferred postglacial recolonisation of northern Europe in several widespread freshwater fishes, including Eurasian perch (*Perca fluviatilis*), European grayling (*Thymallus thymallus*) and European bullhead (*Cottus gobio*) [11,57,58].

The non-monotonic changes in climatically suitable area from the LGM through the Late Glacial broadly correspond to the climatic oscillations documented during the last deglaciation [38], although this temporal correspondence does not by itself demonstrate a causal relationship. However, the magnitude of the difference between the LGM and HS1 projections should be interpreted cautiously because the two layers were derived from different underlying climate products. Notably, although the total area predicted to be climatically suitable decreased between the Bølling–Allerød and the Younger Dryas, the suitability-weighted centroid continued to shift northward. Thus, changes in the total area of climatic suitability did not always coincide with shifts in its geographic centre.

Areas along the northern Black Sea coast were predicted to have been climatically suitable during the LGM. These areas lie within the wider Pontic and Ponto–Caspian regions, where previous phylogeographic studies inferred glacial refugia for circum-Black Sea barbels (*Barbus* spp.) and spined loach (*Cobitis taenia*) [59,60]. However, this geographic correspondence does not demonstrate that the same region served as a refugium for *R. rutilus*. When climatic suitability was required in all three modelled cold-stage time slices—the LGM, HS1 and the Younger Dryas—candidate climatic refugia were restricted to small parts of the Italian Peninsula and the northern Black Sea coastal region. Because this criterion was based on three discrete projections, it cannot establish continuous climatic suitability during the intervening intervals. The absence of candidate refugial cells in the Baltic region is consistent with postglacial colonisation rather than local glacial persistence.

Finally, *R. rutilus* is primarily a freshwater fish. Climatically suitable conditions alone therefore do not demonstrate the presence of suitable aquatic habitats or hydrological connections among drainage basins. During deglaciation, changes in ice-sheet extent and sea level reshaped European coastlines [61,62], while ice-dammed lakes and river rerouting altered drainage configurations across northern Eurasia [63]. These palaeohydrological changes were not explicitly represented in the model, which was based only on climatic predictors. The identified areas should therefore be regarded as model-based candidates that were climatically suitable across the selected cold-stage projections, rather than as confirmed glacial refugia or reconstructed colonisation routes.

4.3. Climatic Suitability and Mitochondrial Haplogroup Geography

The SDM and the mitochondrial data reconstruct different aspects of postglacial history, and their spatial patterns are complementary rather than equivalent [39]. The SDM identified a broad area of climatic suitability across southern and western Europe, as well as the smaller candidate refugial areas. However, climatic suitability by itself does not indicate past occupancy. Correlative models identify areas with suitable environmental conditions, but actual occupancy also depends on whether these areas are accessible to the species

through dispersal [64,65]. For freshwater fishes, such accessibility is further constrained by hydrological connectivity among drainage basins [66]. The centroid-displacement rates reported here quantify shifts in the suitability-weighted geographic centre of projected climatic suitability. They do not measure the actual dispersal or range-shift rate of *R. rutilus*, nor are they equivalent to climate velocity as defined by Loarie et al. [67]. Because the species cannot immediately reach all newly suitable areas, its actual distribution may change more slowly than the geographic distribution of climatic suitability [68].

Our present-day and hindcast projections can be compared with the mitochondrial phylogeography of Larmuseau et al. [10]. The West European clade, to which Baltic-drainage samples in that study belonged, falls within the broad southern to south-eastern zone of glacial climatic suitability recovered here, so the two approaches are spatially consistent. However, because the two studies analysed different mitochondrial regions (cytochrome *b* versus concatenated ATP6–D-loop), our haplogroups A and D cannot be equated directly with the West European and Ponto-Caspian clades. The correspondence holds only at the level of broad geographic pattern, and the spatial overlap is therefore suggestive rather than conclusive.

Our own mitochondrial data add limited resolution. Haplogroup A predominated in all three Lithuanian and Latvian basins and in the Ukrainian Dnieper basin sample and was also present in the reference sequences from Portugal, the United Kingdom, Czechia and Romania. Haplogroup D, by contrast, was represented by only one Russian and one Chinese sequence and therefore provides no further phylogeographic resolution. The wide distribution of haplogroup A is consistent with broad-scale postglacial dispersal, but it could equally reflect the retention of a shared ancestral haplogroup across several populations, and our data do not distinguish between these explanations.

Taken together, the two lines of evidence are mutually informative but individually inconclusive. They are consistent with recolonisation from one or more refugia south of the Baltic, potentially including the Danube–Dniester region, but they cannot determine whether the Baltic populations originated there, from western refugia or from several sources.

4.4. A Provisional Postglacial Recolonisation Scenario

Together, the palaeodistribution and mitochondrial evidence narrow the possible origins of the Baltic populations but do not identify a single source population. Combined with the well-documented postglacial history of the Baltic basin, however, they allow a provisional recolonisation scenario to be outlined. During deglaciation and the Holocene, the Baltic basin passed through successive freshwater and marine-influenced stages—the Baltic Ice Lake, the Yoldia Sea, the Ancylus Lake, and the initial Littorina and Littorina Sea stages—accompanied by substantial changes in salinity, water level, outlets and connections with surrounding drainage systems [5,6,8]. These transitions would have repeatedly altered opportunities for freshwater-mediated dispersal among drainages. Because *R. rutilus*, although primarily a freshwater fish, also tolerates low-salinity brackish coastal waters [9,69], the marine-influenced stages are better interpreted as periods of reduced freshwater connectivity than as absolute barriers. An analogous eastern route from the Black Sea region towards the Baltic, via the Dnieper and Dniester, has recently been proposed for the freshwater mussel *Unio crassus* [70], indicating that such a palaeohydrological corridor may have been available to some freshwater organisms—though this does not establish that the same route was used by *R. rutilus*.

Figure 7 shows a broad temporal overlap between the palaeohydrological reorganisation of the Baltic basin and increasing modelled climatic suitability during the transition to the Holocene. Among the freshwater stages, the Ancylus Lake represents the interval of

greatest freshwater connectivity, whereas the Baltic Ice Lake offered more limited potential for dispersal. However, the figure does not demonstrate that climatically suitable conditions and a particular freshwater dispersal route were available simultaneously during any specific Baltic basin stage. The Ancylus Lake stage covered only a short part of the broader interval represented by the Early Holocene palaeoclimate layer. Moreover, in the ~10 cal ka BP projection, mean climatic suitability at the Baltic-drainage sampling localities remained below the P10 threshold. The Ancylus Lake, and to a lesser extent the Baltic Ice Lake, should therefore be regarded as candidate periods of hydrological opportunity rather than as dispersal windows demonstrated by the present analysis.

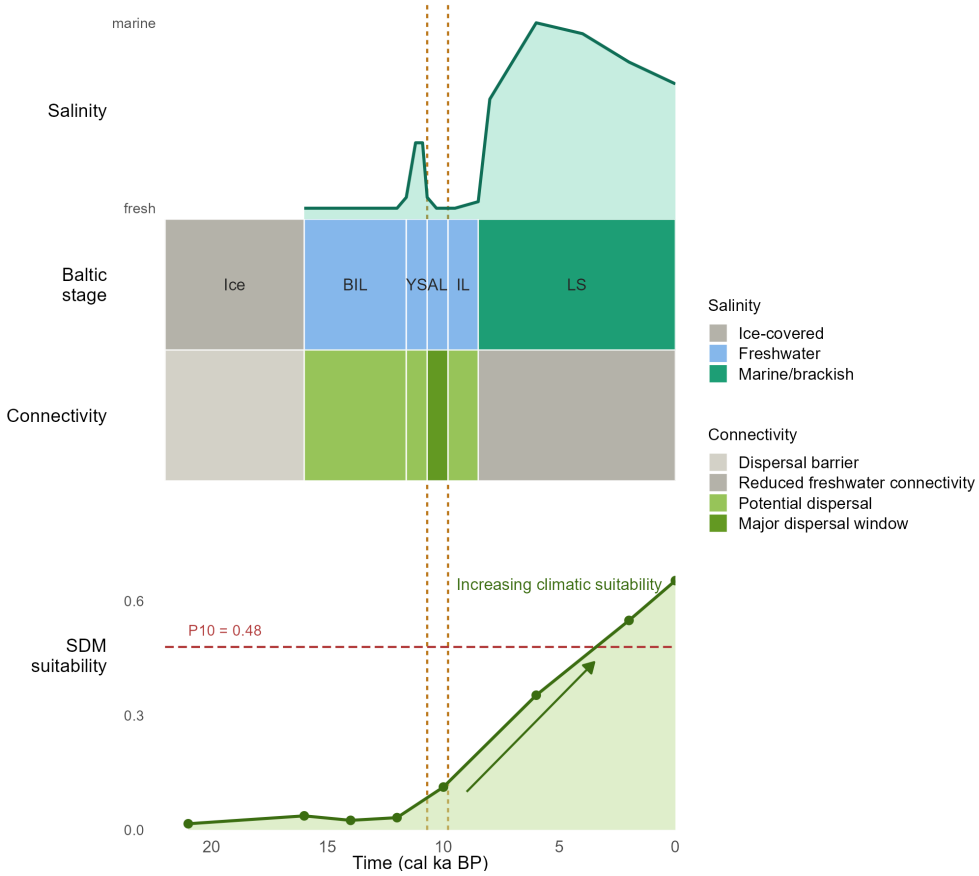


Figure 7. Late-glacial and postglacial history of the eastern Baltic drainage basins summarised on a common calendar-year timescale. From top to bottom: (i) a schematic relative-salinity curve for the Baltic basin (adapted from Björck [5]); (ii) Baltic basin stages and their dominant salinity conditions: Ice, ice-covered basin (LGM–Heinrich Stadial 1); BIL, Baltic Ice Lake; YS, Yoldia Sea; AL, Ancylus Lake; IL, initial Littorina Sea; and LS, Littorina Sea; (iii) inferred connectivity for freshwater fish; and (iv) mean MESS-masked SDM suitability across sampling localities within the Baltic Sea drainage basin, based on the PaleoClim projection closest to each plotted age. Dashed vertical lines delimit the Ancylus Lake stage (~10.7–9.8 cal ka BP). The Yoldia Sea was predominantly fresh, with its brief brackish phase shown as a minor peak in the salinity curve. The stage is therefore coloured as freshwater to represent its dominant condition. The initial Littorina Sea was a transitional stage characterised by very low salinity. Stage boundaries are approximate. The timing of the Baltic basin stages follows the calibrated synthesis of Andrén et al. [6]. The relative-salinity and basin-connection history are schematic adaptations of Björck [5] aligned with this chronology. Freshwater-fish connectivity was inferred from basin salinity and connection history and does not represent an independent measurement.

Given these limitations, we propose as a working hypothesis that the sampled Baltic drainage populations originated through postglacial colonisation from one or more pop-

ulations outside the Baltic region, as climatically suitable areas expanded and freshwater connections changed. The present-day mitochondrial composition of these populations is consistent with recolonisation during the Late Glacial and/or Holocene rather than with local survival through the main cold stages. However, the available data cannot determine when recolonisation occurred, which routes were followed, or how much the Baltic Ice Lake, the Ancylus Lake and individual river connections contributed to this process.

4.5. Limitations

The palaeoprojections are based on terrestrial bioclimatic variables, which represent only indirect proxies for the conditions experienced by a freshwater species and do not capture water temperature, discharge, salinity, aquatic habitat availability or connectivity within river networks [71]. MESS masking excluded areas where past climatic conditions fell outside the range of present-day conditions [72]. However, it did not account for uncertainty in the underlying palaeoclimate datasets, which may be particularly important for the older LGM and HS1 projections. The hindcasts are therefore more useful for identifying broad changes in climatic suitability through time than for precisely estimating the extent of suitable areas during any single palaeoclimatic period [36,73].

Estimates of suitable area and candidate refugia depend on the methods used to classify continuous model predictions and summarise them across discrete time slices. Because the choice of threshold can substantially affect estimates of range change, all reported areas are conditional on the selected P10 threshold of 0.481 [74]. Moreover, the eight palaeoclimatic layers are discrete time slices representing periods of unequal duration. A cell classified as suitable in two consecutive layers may not have remained suitable throughout the time between them. Therefore, candidate refugia represent areas predicted to be suitable in all three modelled cold stages, rather than areas demonstrated to have remained continuously suitable throughout the entire period. Finally, cells with MESS < 0 were assigned a value of zero in the candidate-refugium analysis. Therefore, the final classification did not distinguish cells with predicted suitability below the P10 threshold from cells masked because at least one predictor fell outside its present-day reference range. This treatment is conservative and may underestimate candidate refugial area, particularly for the colder periods, where non-analogue conditions were most extensive.

The genetic interpretation is likewise constrained by a concatenated mitochondrial dataset (ATP6–D-loop), which represents the history of a single maternally inherited, non-recombining genome rather than the nuclear genome [75]. Furthermore, the A–D haplogroups represent haplotype clusters identified in a median-joining network rather than phylogenetically validated mitochondrial lineages. Geographic coverage and sample sizes were uneven. The Lielupe basin was represented by a single sampling locality, while most GenBank reference localities were represented by only one or two sequences. Consequently, the available data cannot identify the exact source population of postglacial recolonisation or reliably estimate haplogroup frequencies across the species' wider range. They also cannot demonstrate that a particular haplogroup is truly absent from regions represented by few or no samples.

Finally, the proposed scenario integrates SDM-derived climatic suitability with published reconstructions of Baltic basin evolution but does not explicitly model palaeohydrological connectivity. Mitochondrial data alone also cannot distinguish unambiguously among alternative demographic processes, because similar spatial genetic patterns may arise through different combinations of refugial persistence, founder events, postglacial expansion and secondary contact [76]. Further progress will therefore require denser continental sampling, multilocus genomic data, threshold sensitivity analyses and explicit

palaeohydrological connectivity modelling to distinguish among alternative recolonisation pathways and identify the demographic sources of the Baltic populations more precisely.

5. Conclusions

By integrating climatic and mitochondrial evidence, we found that the area predicted to be climatically suitable for the common roach (*Rutilus rutilus*) expanded substantially from the LGM to the present, although changes during the Late Glacial were non-monotonic. Areas classified as suitable in all three cold-stage projections were restricted to small parts of the Italian Peninsula and the northern Black Sea coastal region, with no candidate climatic refugia identified in the Baltic region. These areas indicate persistent modelled climatic suitability rather than confirmed historical population refugia. Freshwater stages of the Baltic basin may have provided opportunities for dispersal as climatic suitability increased, but the timing and importance of these hydrological connections were not directly tested.

Haplogroup A dominated the sampled Baltic drainages and the Ukrainian Dnieper River basin and was also present in the limited western and central European reference sequences. In contrast, haplogroup D occurred only in the two easternmost reference sequences. This pattern is consistent with recolonisation of the Baltic region by populations sharing a geographically widespread mitochondrial haplogroup. However, the limited reference sampling and reliance on mitochondrial data cannot identify a specific refugial source or distinguish postglacial dispersal from the persistence of ancestral genetic variation.

The climatic modelling and mitochondrial analysis therefore provide complementary but incomplete evidence. Together, they are consistent with postglacial recolonisation from one or more regions outside the Baltic, potentially facilitated by freshwater stages of the Baltic basin, but they do not identify a unique source or dispersal route. Denser geographic sampling, nuclear genomic markers and explicit modelling of palaeohydrological connectivity will be required to test this scenario.

Supplementary Materials: The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/quat9060076/s1, Figure S1: Jackknife of regularized training gain for the present-day MaxEnt model of *R. rutilus*; Figure S2: Median-joining network of *R. rutilus* mitochondrial haplotypes based on the concatenated ATP6–D-loop dataset (822 bp). Table S1: Sampling populations, drainage basins and mitochondrial haplogroup composition of *R. rutilus* used in this study; Table S2: Performance and calibration metrics of the present-day MaxEnt model; Table S3: Contributions of the five bioclimatic predictors to the present-day MaxEnt model of *R. rutilus*; Table S4: Projected suitable area, suitability-weighted range centroid, and range-shift velocity of *R. rutilus* across eight time periods from the Last Glacial Maximum to the present.

Author Contributions: Conceptualization I.I. and D.B., D.G. (Dace Grauda) and D.G. (Dmitri Gudkov); methodology, I.I., D.B., D.G. (Dace Grauda) and D.G. (Dmitri Gudkov); software, I.I.; validation, D.B., D.G. (Dace Grauda) and D.G. (Dmitri Gudkov); formal analysis, I.I.; investigation, I.I., D.B., D.G. (Dace Grauda); resources, D.B., D.G. (Dace Grauda) and D.G. (Dmitri Gudkov); data curation, D.B.; writing—original draft preparation, I.I.; writing—review and editing, I.I., D.B., D.G. (Dmitri Gudkov); visualization, I.I.; supervision, D.B.; project administration, D.B., D.G. (Dace Grauda), D.G. (Dmitri Gudkov); funding acquisition, D.B., D.G. (Dmitri Gudkov). All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially funded from Latvian-Ukrainian Joint Programme of Scientific and Technological Cooperation Project Nr. LV_UA/2025/11.

Data Availability Statement: The species occurrence records used in this study are publicly available from GBIF under download DOI https://doi.org/10.15468/dl.3wbya3 (accessed on 18 June 2026). The climatic data are publicly available from PaleoClim (https://www.paleoclim.org/) (accessed on 18 June 2026) and were accessed using the rpaleoclim R package.

Acknowledgments: The study was supported by the Ministry of Education and Science of Ukraine (Project No. M/09-2025) and the National Academy of Sciences of Ukraine and was conducted in cooperation with the State Specialized Enterprise "Ecocentre" of the State Agency of Ukraine for Exclusion Zone Management.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

SDM	Species distribution model
LGM	Last Glacial Maximum
GBIF	Global Biodiversity Information Facility
HS1	Heinrich Stadial 1
BA	Bølling–Allerød
YD	Younger Dryas
EH	Early Holocene
MH	Mid Holocene
LH	Late Holocene
AUC	Area under the receiver operating characteristic curve
CBI	Continuous Boyce Index
MESS	Multivariate environmental similarity surface

References

1. Hewitt, G.M. The genetic legacy of the Quaternary ice ages. *Nature* **2000**, *405*, 907–913. [CrossRef] [PubMed]
2. Bhagwat, S.A.; Willis, K.J. Species persistence in northerly glacial refugia of Europe: A matter of chance or biogeographical traits? *J. Biogeogr.* **2008**, *35*, 464–482. [CrossRef]
3. Schmitt, T.; Varga, Z. Extra-Mediterranean refugia: The rule and not the exception? *Front. Zool.* **2012**, *9*, 22. [CrossRef]
4. Mandrak, N.E.; Crossman, E.J. Postglacial dispersal of freshwater fishes into Ontario. *Can. J. Zool.* **1992**, *70*, 2247–2259. [CrossRef]
5. Björck, S. A review of the history of the Baltic Sea, 13.0–8.0 ka BP. *Quat. Int.* **1995**, *27*, 19–40. [CrossRef]
6. Andrén, T.; Björck, S.; Andrén, E.; Conley, D.J.; Zillén, L.; Anjar, J. The development of the Baltic Sea Basin during the last 130 ka. In *The Baltic Sea Basin*; Harff, J., Björck, S., Hoth, P., Eds.; Springer: Berlin/Heidelberg, Germany, 2011; pp. 75–97. [CrossRef]
7. Jakobsson, M.; Björck, S.; Alm, G.; Andrén, T.; Lindeberg, G.; Svensson, N.-O. Reconstructing the Younger Dryas ice-dammed lake in the Baltic Basin: Bathymetry, area and volume. *Glob. Planet. Change* **2007**, *57*, 355–370. [CrossRef]
8. Rosentau, A.; Klemann, V.; Bennike, O.; Steffen, H.; Wehr, J.; Latinović, M.; Bagge, M.; Ojala, A.; Berglund, M.; Peterson Becher, G.; et al. A Holocene relative sea-level database for the Baltic Sea. *Quat. Sci. Rev.* **2021**, *266*, 107071. [CrossRef]
9. Kottelat, M.; Freyhof, J. *Handbook of European Freshwater Fishes*; Publications Kottelat: Cornol, Switzerland, 2007; 646p, ISBN 978-2-8399-0298-4.
10. Larmuseau, M.H.D.; Freyhof, J.; Volckaert, F.A.M.; Van Houdt, J.K.J. Matrilinear phylogeography and demographical patterns of *Rutilus rutilus*: Implications for taxonomy and conservation. *J. Fish Biol.* **2009**, *75*, 332–353. [CrossRef]
11. Nesbø, C.L.; Fossheim, T.; Vøllestad, L.A.; Jakobsen, K.S. Genetic divergence and phylogeographic relationships among European perch (*Perca fluviatilis*) populations reflect glacial refugia and postglacial colonization. *Mol. Ecol.* **1999**, *8*, 1387–1404. [CrossRef]
12. Lichman, V.; Ozerov, M.; López, M.-E.; Noreikiene, K.; Kahar, S.; Pukk, L.; Burimski, O.; Japoshvili, B.; Blazhekovikj-Dimovska, D.; Lajus, D.; et al. Pan-continental genomic analysis of Eurasian perch uncovers global diversity hotspots and postglacial recolonization patterns. *Ecol. Evol.* **2026**, *16*, e73502. [CrossRef]
13. Ignatavičienė, I.; Ragauskas, A.; Bencevičiūtė, A.; Rakauskas, V.; Butkauskas, D. Comparative Population Genetics of *Rutilus rutilus* and *Perca fluviatilis* in Lithuanian Freshwater Ecosystems. *Animals* **2026**, *16*, 2540. [CrossRef]
14. Wielstra, B.; Crnobrnja-Isailović, J.; Litvinchuk, S.N.; Reijnen, B.T.; Skidmore, A.K.; Sotiropoulos, K.; Toxopeus, A.G.; Tzankov, N.; Vukov, T.; Arntzen, J.W. Tracing glacial refugia of *Triturus* newts based on mitochondrial DNA phylogeography and species distribution modeling. *Front. Zool.* **2013**, *10*, 13. [CrossRef]
15. Hewitt, G.M. Post-glacial re-colonization of European biota. *Biol. J. Linn. Soc.* **1999**, *68*, 87–112. [CrossRef]
16. GBIF.org. GBIF Occurrence Download. 18 June 2026. Available online: https://doi.org/10.15468/dl.3wbya3 (accessed on 18 June 2026).
17. Boria, R.A.; Olson, L.E.; Goodman, S.M.; Anderson, R.P. Spatial filtering to reduce sampling bias can improve the performance of ecological niche models. *Ecol. Model.* **2014**, *275*, 73–77. [CrossRef]

18. Fick, S.E.; Hijmans, R.J. WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.* **2017**, *37*, 4302–4315. [CrossRef]
19. Brown, J.L.; Hill, D.J.; Dolan, A.M.; Carnaval, A.C.; Haywood, A.M. PaleoClim, high spatial resolution paleoclimate surfaces for global land areas. *Sci. Data* **2018**, *5*, 180254. [CrossRef]
20. Roe, J. *rpaleoclim: Download Paleoclimate Data from 'PaleoClim'*, R package version 1.1.0; CRAN: Vienna, Austria, 2025. [CrossRef]
21. Karger, D.N.; Conrad, O.; Böhner, J.; Kawohl, T.; Kreft, H.; Soria-Auza, R.W.; Zimmermann, N.E.; Linder, H.P.; Kessler, M. Climatologies at high resolution for the Earth's land surface areas. *Sci. Data* **2017**, *4*, 170122. [CrossRef] [PubMed]
22. Fordham, D.A.; Saltré, F.; Haythorne, S.; Wigley, T.M.L.; Otto-Bliesner, B.L.; Chan, K.C.; Brook, B.W. PaleoView: A tool for generating continuous climate projections spanning the last 21,000 years at regional and global scales. *Ecography* **2017**, *40*, 1348–1358. [CrossRef]
23. Dormann, C.F.; Elith, J.; Bacher, S.; Buchmann, C.; Carl, G.; Carré, G.; García Márquez, J.R.; Gruber, B.; Lafourcade, B.; Leitão, P.J.; et al. Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography* **2013**, *36*, 27–46. [CrossRef]
24. Kuhn, M. Building predictive models in R using the caret package. *J. Stat. Softw.* **2008**, *28*, 1–26. [CrossRef]
25. Booth, T.H. Checking bioclimatic variables that combine temperature and precipitation data before their use in species distribution models. *Austral Ecol.* **2022**, *47*, 1506–1514. [CrossRef]
26. Escobar, L.E.; Lira-Noriega, A.; Medina-Vogel, G.; Peterson, A.T. Potential for spread of the white-nose fungus (*Pseudogymnoascus destructans*) in the Americas: Use of Maxent and NicheA to assure strict model transference. *Geospat. Health* **2014**, *9*, 221–229. [CrossRef]
27. O'Donnell, M.S.; Ignizio, D.A. *Bioclimatic Predictors for Supporting Ecological Applications in the Conterminous United States*; Data Series 691; U.S. Geological Survey: Reston, VA, USA, 2012; 10p. [CrossRef]
28. Phillips, S.J.; Anderson, R.P.; Schapire, R.E. Maximum entropy modeling of species geographic distributions. *Ecol. Model.* **2006**, *190*, 231–259. [CrossRef]
29. Hijmans, R.J.; Phillips, S.; Leathwick, J.; Elith, J. *dismo: Species Distribution Modeling*, R package version 1.3-16; CRAN: Vienna, Austria, 2024. Available online: https://CRAN.R-project.org/package=dismo (accessed on 19 August 2026).
30. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2025. Available online: https://www.R-project.org/ (accessed on 19 August 2026).
31. Phillips, S.J.; Anderson, R.P.; Dudík, M.; Schapire, R.E.; Blair, M.E. Opening the black box: An open-source release of Maxent. *Ecography* **2017**, *40*, 887–893. [CrossRef]
32. Broennimann, O.; Di Cola, V.; Guisan, A. *ecospat: Spatial Ecology Miscellaneous Methods*, R package version 4.1.4; CRAN: Vienna, Austria, 2026. [CrossRef]
33. Hirzel, A.H.; Le Lay, G.; Helfer, V.; Randin, C.; Guisan, A. Evaluating the ability of habitat suitability models to predict species presences. *Ecol. Model.* **2006**, *199*, 142–152. [CrossRef]
34. Boyce, M.S.; Vernier, P.R.; Nielsen, S.E.; Schmiegelow, F.K.A. Evaluating resource selection functions. *Ecol. Model.* **2002**, *157*, 281–300. [CrossRef]
35. Roberts, D.R.; Bahn, V.; Ciuti, S.; Boyce, M.S.; Elith, J.; Guillera-Arroita, G.; Hauenstein, S.; Lahoz-Monfort, J.J.; Schröder, B.; Thuiller, W.; et al. Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography* **2017**, *40*, 913–929. [CrossRef]
36. Elith, J.; Kearney, M.; Phillips, S.J. The art of modelling range-shifting species. *Methods Ecol. Evol.* **2010**, *1*, 330–342. [CrossRef]
37. Pearson, R.G.; Raxworthy, C.J.; Nakamura, M.; Peterson, A.T. Predicting species distributions from small numbers of occurrence records: A test case using cryptic geckos in Madagascar. *J. Biogeogr.* **2007**, *34*, 102–117. [CrossRef]
38. Rasmussen, S.O.; Bigler, M.; Blockley, S.P.E.; Blunier, T.; Buchardt, S.L.; Clausen, H.B.; Cvijanovic, I.; Dahl-Jensen, D.; Johnsen, S.J.; Fischer, H.; et al. A stratigraphic framework for abrupt climatic changes during the Last Glacial period based on three synchronized Greenland ice-core records: Refining and extending the INTIMATE event stratigraphy. *Quat. Sci. Rev.* **2014**, *106*, 14–28. [CrossRef]
39. Gavin, D.G.; Fitzpatrick, M.C.; Gugger, P.F.; Heath, K.D.; Rodríguez-Sánchez, F.; Dobrowski, S.Z.; Hampe, A.; Hu, F.S.; Ashcroft, M.B.; Bartlein, P.J.; et al. Climate refugia: Joint inference from fossil records, species distribution models and phylogeography. *New Phytol.* **2014**, *204*, 37–54. [CrossRef]
40. Butkauskas, D.; Ignatavičienė, I.; Ragauskas, A.; Rakauskas, V. Temporal variation of genetic diversity in *Rutilus rutilus* populations from Lithuania using mtDNA markers in the context of anthropogenic activities. *Diversity* **2023**, *15*, 1113. [CrossRef]
41. Ignatavičienė, I.; Ragauskas, A.; Rakauskas, V.; Butkauskas, D. Quality of DNA extracted from freshwater fish scales and mucus and its application in genetic diversity studies of *Perca fluviatilis* and *Rutilus rutilus*. *Biol. Methods Protoc.* **2023**, *8*, bpad022. [CrossRef]
42. Sayers, E.W.; Cavanaugh, M.; Frisse, L.; Pruitt, K.D.; Schneider, V.A.; Underwood, B.A.; Yankie, L.; Karsch-Mizrachi, I. GenBank 2025 update. *Nucleic Acids Res.* **2025**, *53*, D56–D61. [CrossRef]

43. Edgar, R.C. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **2004**, *32*, 1792–1797. [CrossRef] [PubMed]
44. Kumar, S.; Stecher, G.; Li, M.; Knyaz, C.; Tamura, K. MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* **2018**, *35*, 1547–1549. [CrossRef] [PubMed]
45. Bandelt, H.J.; Forster, P.; Röhl, A. Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol.* **1999**, *16*, 37–48. [CrossRef]
46. Pebesma, E. Simple features for R: Standardized support for spatial vector data. *R J.* **2018**, *10*, 439–446. [CrossRef]
47. Pebesma, E.; Bivand, R. *Spatial Data Science: With Applications in R*; Chapman and Hall/CRC: Boca Raton, FL, USA, 2023. [CrossRef]
48. Massicotte, P.; South, A. *rnaturalearth: World Map Data from Natural Earth*, R package version 1.2.0; CRAN: Vienna, Austria, 2026. [CrossRef]
49. Yu, G. *scatterpie: Scatter Pie Plot*, R package version 0.2.6; CRAN: Vienna, Austria, 2025. [CrossRef]
50. Dunnington, D. *ggspatial: Spatial Data Framework for ggplot2*, R package version 1.1.10; CRAN: Vienna, Austria, 2025. [CrossRef]
51. Lobo, J.M.; Jiménez-Valverde, A.; Real, R. AUC: A misleading measure of the performance of predictive distribution models. *Glob. Ecol. Biogeogr.* **2008**, *17*, 145–151. [CrossRef]
52. Peterson, A.T.; Papeş, M.; Soberón, J. Rethinking receiver operating characteristic analysis applications in ecological niche modeling. *Ecol. Model.* **2008**, *213*, 63–72. [CrossRef]
53. Phillips, S.J.; Dudík, M. Modeling of species distributions with Maxent: New extensions and a comprehensive evaluation. *Ecography* **2008**, *31*, 161–175. [CrossRef]
54. Elith, J.; Phillips, S.J.; Hastie, T.; Dudík, M.; Chee, Y.E.; Yates, C.J. A statistical explanation of MaxEnt for ecologists. *Divers. Distrib.* **2011**, *17*, 43–57. [CrossRef]
55. Smith, A.B.; Santos, M.J. Testing the ability of species distribution models to infer variable importance. *Ecography* **2020**, *43*, 1801–1813. [CrossRef]
56. Makhrov, A.A.; Bolotov, I.N. Dispersal routes and species identification of freshwater animals in Northern Europe: A review of molecular evidence. *Russ. J. Genet.* **2006**, *42*, 1101–1115. [CrossRef]
57. Koskinen, M.T.; Ranta, E.; Piironen, J.; Veselov, A.; Titov, S.; Haugen, T.O.; Nilsson, J.; Carlstein, M.; Primmer, C.R. Genetic lineages and postglacial colonization of grayling (*Thymallus thymallus*, Salmonidae) in Europe, as revealed by mitochondrial DNA analyses. *Mol. Ecol.* **2000**, *9*, 1609–1624. [CrossRef]
58. Kontula, T.; Väinölä, R. Postglacial colonization of Northern Europe by distinct phylogeographic lineages of the bullhead, *Cottus gobio*. *Mol. Ecol.* **2001**, *10*, 1983–2002. [CrossRef]
59. Kotlík, P.; Bogutskaya, N.G.; Ekmekçi, F.G. Circum Black Sea phylogeography of Barbus freshwater fishes: Divergence in the Pontic glacial refugium. *Mol. Ecol.* **2004**, *13*, 87–95. [CrossRef]
60. Culling, M.A.; Janko, K.; Boron, A.; Vasil'ev, V.P.; Côté, I.M.; Hewitt, G.M. European colonization by the spined loach (*Cobitis taenia*) from Ponto-Caspian refugia based on mitochondrial DNA variation. *Mol. Ecol.* **2006**, *15*, 173–190. [CrossRef]
61. Clark, P.U.; Dyke, A.S.; Shakun, J.D.; Carlson, A.E.; Clark, J.; Wohlfarth, B.; Mitrovica, J.X.; Hostetler, S.W.; McCabe, A.M. The Last Glacial Maximum. *Science* **2009**, *325*, 710–714. [CrossRef]
62. Lambeck, K.; Rouby, H.; Purcell, A.; Sun, Y.; Sambridge, M. Sea level and global ice volumes from the Last Glacial Maximum to the Holocene. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 15296–15303. [CrossRef]
63. Mangerud, J.; Jakobsson, M.; Alexanderson, H.; Astakhov, V.; Clarke, G.K.C.; Henriksen, M.; Hjort, C.; Krinner, G.; Lunkka, J.-P.; Möller, P.; et al. Ice-dammed lakes and rerouting of the drainage of northern Eurasia during the Last Glaciation. *Quat. Sci. Rev.* **2004**, *23*, 1313–1332. [CrossRef]
64. Soberón, J.; Peterson, A.T. Interpretation of models of fundamental ecological niches and species' distributional areas. *Biodivers. Inform.* **2005**, *2*, 1–10. [CrossRef]
65. Barve, N.; Barve, V.; Jiménez-Valverde, A.; Lira-Noriega, A.; Maher, S.P.; Peterson, A.T.; Soberón, J.; Villalobos, F. The crucial role of the accessible area in ecological niche modeling and species distribution modeling. *Ecol. Model.* **2011**, *222*, 1810–1819. [CrossRef]
66. Carvajal-Quintero, J.; Villalobos, F.; Oberdorff, T.; Grenouillet, G.; Brosse, S.; Hugueny, B.; Jézéquel, C.; Tedesco, P.A. Drainage network position and historical connectivity explain global patterns in freshwater fishes' range size. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 13434–13439. [CrossRef]
67. Loarie, S.R.; Duffy, P.B.; Hamilton, H.; Asner, G.P.; Field, C.B.; Ackerly, D.D. The velocity of climate change. *Nature* **2009**, *462*, 1052–1055. [CrossRef]
68. Normand, S.; Ricklefs, R.E.; Skov, F.; Bladt, J.; Tackenberg, O.; Svenning, J.-C. Postglacial migration supplements climate in determining plant species ranges in Europe. *Proc. R. Soc. B Biol. Sci.* **2011**, *278*, 3644–3653. [CrossRef]
69. Vetemaa, M.; Eschbaum, R.; Verliin, A.; Albert, A.; Eero, M.; Lillemägi, R.; Pihlak, M.; Saat, T. Annual and seasonal dynamics of fish in the brackish-water Matsalu Bay, Estonia. *Ecol. Freshw. Fish* **2006**, *15*, 211–220. [CrossRef]

70. Egg, S.; Lopes-Lima, M.; Bayerl, H.; Froufe, E.; Stoeckle, B.C.; Kuehn, R.; Geist, J. The impact of glacial disturbance history upon the genetic diversity of *Unio crassus* and *Unio nanus* in Europe and implications for conservation. *Ecol. Evol.* **2025**, *15*, e72113. [CrossRef]
71. Domisch, S.; Jähnig, S.C.; Simaika, J.P.; Kuemmerlen, M.; Stoll, S. Application of species distribution models in stream ecosystems: The challenges of spatial and temporal scale, environmental predictors and species occurrence data. *Fundam. Appl. Limnol.* **2015**, *186*, 45–61. [CrossRef]
72. Owens, H.L.; Campbell, L.P.; Dornak, L.L.; Saupe, E.E.; Barve, N.; Soberón, J.; Ingenloff, K.; Lira-Noriega, A.; Hensz, C.M.; Myers, C.E.; et al. Constraints on interpretation of ecological niche models by limited environmental ranges on calibration areas. *Ecol. Model.* **2013**, *263*, 10–18. [CrossRef]
73. Elith, J.; Leathwick, J.R. Species distribution models: Ecological explanation and prediction across space and time. *Annu. Rev. Ecol. Evol. Syst.* **2009**, *40*, 677–697. [CrossRef]
74. Nenzén, H.K.; Araújo, M.B. Choice of threshold alters projections of species range shifts under climate change. *Ecol. Model.* **2011**, *222*, 3346–3354. [CrossRef]
75. Zink, R.M.; Barrowclough, G.F. Mitochondrial DNA under siege in avian phylogeography. *Mol. Ecol.* **2008**, *17*, 2107–2121. [CrossRef]
76. Knowles, L.L.; Maddison, W.P. Statistical phylogeography. *Mol. Ecol.* **2002**, *11*, 2623–2635. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.