Freshwater Biology

WILEY

Freshwater Biology

ORIGINAL ARTICLE OPEN ACCESS

Are Species Distribution Models That Use Climate Variables Suitable to Forecast the Changing Distribution of Freshwater Organisms?

David Carmelet-Rescan[1] | Steven A. Trewick[2] | Charlotte Bridger[2,3] | Stephen R. Pohe[4] | Mary Morgan-Richards[2] | Felix Zareie-Vaux[5]

[1]Centre d'Études Biologiques de Chizé, UMR 7372, University of La Rochelle-CNRS, Villiers-en-Bois, France | [2]Wildlife & Ecology, School of Food Technology and Natural Sciences, Massey University, Palmerston North, New Zealand | [3]Bioeconomy Science Institute, Havelock North, New Zealand | [4]Pohe Environmental, Whangārei, New Zealand | [5]Earth Sciences New Zealand, Wellington, New Zealand

Correspondence: Mary Morgan-Richards (m.morgan-richards@massey.ac.nz)

Received: 4 December 2025 | **Revised:** 23 July 2026 | **Accepted:** 14 August 2026

Keywords: ecological niche model | *Ichthybotus* | mayfly | mtDNA | New Zealand | wing shape

ABSTRACT

1. Species distribution models can be used to identify locations with environmental conditions suitable for populations in the past, present and future. However, climate variables commonly applied in species distribution models may be unsuitable for freshwater species because their explanatory power is difficult to assess. We examined the suitability of such models using a pair of extant mayfly species, *Ichthybotus hudsoni* and *I. bicolor*, that are allopatric in the main islands of Aotearoa–New Zealand.
2. These mayflies were confirmed as distinct species by concordance of hindwing colour, forewing shape and mtDNA. Using present-day species ranges and climate variables, we built models to predict the distribution of each during the climate of the last glacial maximum. Genetic data were used to validate predictions derived from hindcasted species distribution models.
3. Species distribution models suggested that climatically suitable habitat for the two mayfly species was widespread in lowland New Zealand during the last glacial maximum, from which we predicted a genetic signature of large, stable populations for both. High haplotype diversity and a signature of isolation by distance were detected as predicted for *I. hudsoni* in North Island, but models did not predict the observed association between latitude and nucleotide diversity. A small *I. bicolor* sample size in South Island limited our ability to distinguish patterns, but the low genetic diversity found suggests poor performance of species distribution models based on climate variables.
4. We conclude that species distribution models based solely on climate variables are unlikely to accurately predict changes in the distributions of freshwater organisms.

1 | Introduction

Species distribution models (SDMs), or environmental niche models, are widely used to estimate species' geographic ranges in the past, present and future. Forecasting has a particularly valuable role when considering the effects of anthropogenic climate change on species distributions (Malhi et al. 2020; Meza-Joya et al. 2025). Climate-based models can reflect local environmental conditions that are not directly measured, and therefore provide a means to assess the impact of large-scale shifts in climate. Hindcasting can, similarly, be used to estimate prehistoric biogeographic change, especially that associated with climate shifts due to Milankovitch cycles (e.g., Varela et al. 2011; Franklin et al. 2015; Maguire et al. 2015).

Using climate variables as the primary proxies to estimate suitable habitat space for species in the past can in turn help explain patterns of genetic diversity observed among living populations (e.g., Ye et al. 2014; Ramírez-Barahona et al. 2017; Rocha et al. 2020).

The availability of occurrence and environmental data can hamper the application of SDMs to freshwater species (Domínguez-Domínguez et al. 2006; Valencia-Rodríguez et al. 2021). When applied to invasive species that, by definition, have experienced recent, rapid range change, model transferability can be low for aquatic taxa (Liu et al. 2020). The inclusion of aquatic habitat variables (e.g., water depth, pH, flow rate, dissolved oxygen level) can improve the fit of distribution models for freshwater fish (e.g., see Valencia-Rodríguez et al. 2021), although topographic variables sometimes better predict species presence than hydrology (Parreira et al. 2019). The challenge is to assess the quality of models that predict decreasing suitability of habitat into the future (e.g., inferred for cold-adapted freshwater macroinvertebrates; Taubmann et al. 2011; Domisch et al. 2013).

Modelling the environmental envelope of species to forecast range shifts is now common (Araújo et al. 2019; Liu et al. 2020; Koot et al. 2022), however, the evaluation of such models and their predictions is difficult (Warren et al. 2020; Lee-Yaw et al. 2022). Major problems include determining whether the underlying assumptions of the models are credible (Wiens et al. 2010) and dealing with the illusion of certainty in high-resolution maps that mask the quality and quantity of the underlying data (Jansen et al. 2022). The assumption that all suitable habitat is occupied without dispersal limitations or evolutionary change may not be met, and overlooked species interactions will result in a disconnect between geographic distributions and environmental tolerances (Wisz et al. 2013; Dormann et al. 2018; Ortego and Knowles 2020). Competition among closely related species, particularly within confined landscapes such as islands, can restrict geographic ranges and conceal the maximum potential habitat of many species (Arntzen 2023). Reliance on geographic projections to evaluate and select optimal models may be insufficient (Smith et al. 2019), making genetic data particularly valuable in validating SDM predictions (He et al. 2013; Diniz-Filho et al. 2015; Lee-Yaw et al. 2022). The fundamental relationship between population size and neutral genetic diversity (Charlesworth 2009) and well-defined genetic consequences of range expansion (Excoffier et al. 2009) mean that genetic diversity provides an opportunity for model validation. Spatial and temporal dynamics of population demographics can also be simulated to generate expected patterns of genetic diversity (Currat et al. 2019). In simple terms, large stable populations retain genetic diversity, while reduced effective population size increases the rate at which alleles are lost, producing low diversity in recently colonised ranges (Figure 1).

Given the difficulty of inferring how aquatic habitat variables change over time, we investigate whether SDMs using climate variables are suitable proxies for predicting distribution shifts in freshwater invertebrates. Population genetic data from extant populations provides an independent assessment of the capacity to infer past distributions; a close fit to the model predictions will indicate that SDMs are suitable tools for inferring past and therefore potential future distribution shifts. We studied a pair of mayfly taxa (Insecta: Ephemeroptera; piriwai in te reo Māori) with separate geographic distributions, first determining whether they are genetically and morphologically distinct, then generating distribution models for each based on current climate data of Aotearoa–New Zealand (hereafter NZ). As these two putative species are allopatric we modelled their distributions independently using presence/absence data with standard climate variables. We test predictions arising from models that hindcast continual occupation at geographic locations during the last glacial maximum (LGM), using population samples from these locations in the current interglacial period (Figure 1a,d). For the mayfly *Ichthybotus hudsoni*, SDMs hindcasted continuous occupation at 78% of sampled locations during the last glacial cycle. Our SDMs for *I. bicolor* hindcasted high suitability for 67% of sampled locations during the last glacial cycle. Thus, for most samples, we predicted genetic signatures of large stable populations (Figure 1). We predicted the two mayfly species would have similar levels of genetic diversity (P1), and both species would have high genetic diversity over their range (P2), constant population size (P3), that spatially restricted gene flow would create genetic signatures of isolation by distance (P4) and haplotype networks would be spatially structured as expected of stable populations (P5).

2 | Methods

2.1 | Study Taxa

Mayflies have nymphs that live in freshwater, and short-lived, winged, terrestrial adults. NZ has 59 described species of endemic mayfly (Hitchings et al. 2022; Ridden et al. 2023) including two in the burrowing family Ichthybotidae Demoulin, 1957 in the endemic genus *Ichthybotus*. *Ichthybotus hudsoni* is restricted to North Island (Te Ika-a-Māui) while *I. bicolor* occurs in South Island (Te Waipounamu) (Figure 2; Hitchings 2008; Pohe 2019). These *Ichthybotus* sister taxa are relatively common in low elevation streams where native forest remains. They are distinguished by adult wing colouration (Tillyard 1923; Chisholm 1984), specifically the darker hindwings of *I. bicolor* (Figure 2). To confirm that the two putative species (hereafter species) of *Ichthybotus* are morphologically distinct we investigated variation in forewing shape and size using two-dimensional landmark-based geometric analysis. Full details of this morphometric analysis are presented in the Supporting Information (Appendix S1, Tables S3 and S6 and Figures S3–S10).

2.2 | Interspecific Genetic Analysis

To confirm the two *Ichthybotus* species are genetically differentiated we sequenced a fragment of mitochondrial DNA (mtDNA) from 225 specimens. Ethanol preserved specimens (nymphs, subimagos or imagos) were used for whole genomic DNA extraction using a high-salt method (Trewick et al. 2022). Mitochondrial DNA sequences were obtained using primers targeting a portion of the cytochrome oxidase subunit-I (COI) gene. Polymerase chain reactions (PCR) used the primers LCO1490 and HCO2198 (Folmer et al. 1994) with an annealing temperature of 48°C. PCR products were sequenced using big-dye chemistry (PerkinElmer; Waltham,

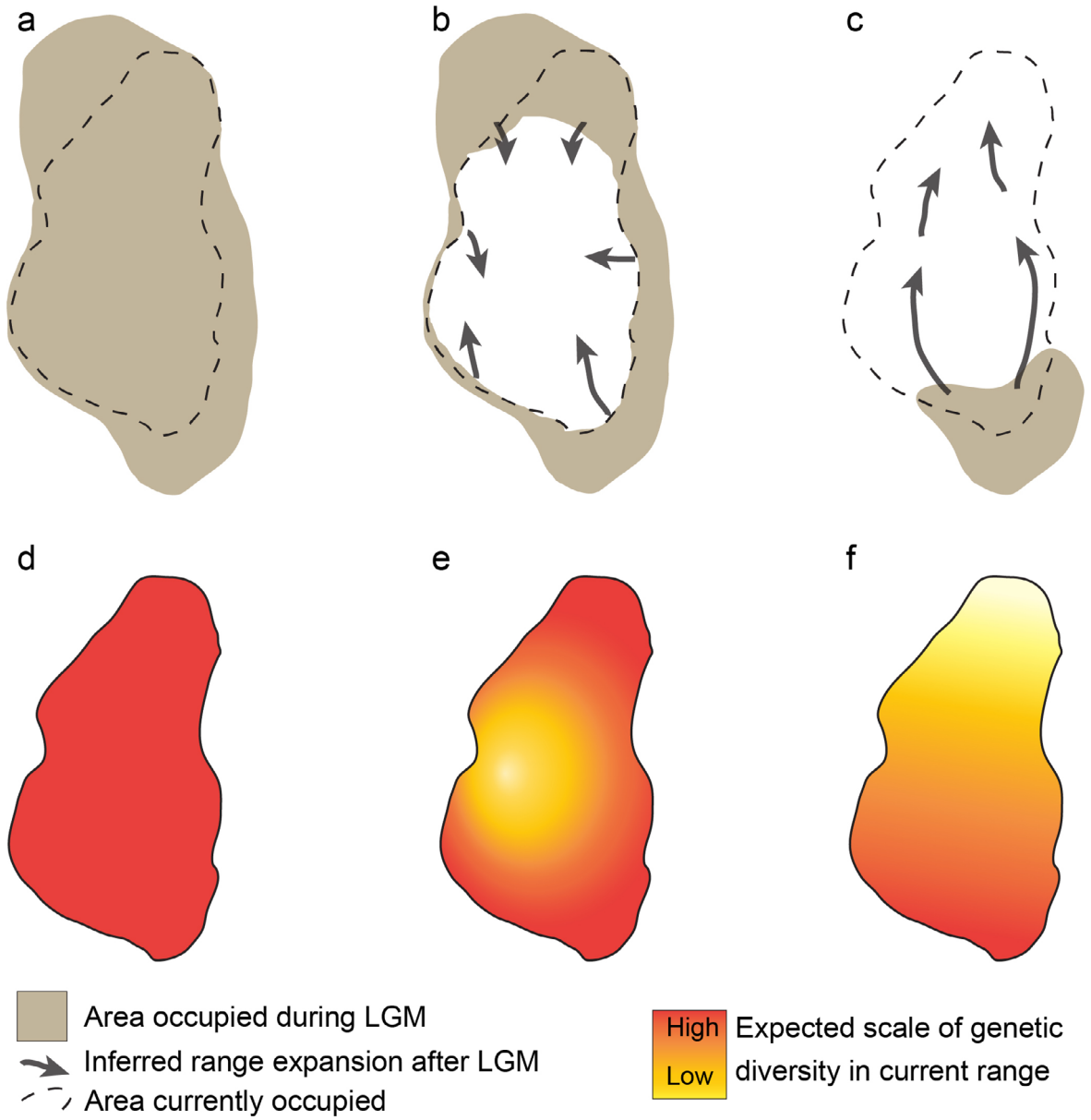

FIGURE 1 | Different species distributions (a–c) during the last glacial maximum (LGM) predict distinct patterns of intraspecific genetic diversity across the extant species range (e–g). Stable populations (a) predict uniform high genetic diversity (d), expanding populations (b, c) predict regions of relatively low genetic diversity (e, f).

MA, USA) following the manufacturer's protocol on an ABI3730 DNA analyser (Hitachi, Japan).

DNA sequences were checked for ambiguities, translated and aligned in Geneious Prime (Kearse et al. 2012) and haplotypes identified. Phylogenetic relationships among the *Ichthybotus* haplotypes were inferred using neighbour-joining implemented in Geneious and applying the HKY model of nucleotide substitution. Pairwise HKY genetic distances were used to estimate the time of the most recent common ancestor of the two lineages by applying an interspecific substitution rate of 0.0177 per site per million years derived for COI gene from invertebrate mitochondria (Papadopoulou et al. 2010), and assuming a strict molecular clock (Ho and Duchêne 2014). Representative haplotypes were submitted to GenBank (accession numbers: PP778198–PP778299) and are also available via BOLD.

2.3 | Species Distribution Models (SDMs)

2.3.1 | Occurrence Records

Data on the current distribution of the two *Ichthybotus* species, and other NZ mayflies, were gathered during a nationwide light trapping survey of 83 streams across the three main islands of NZ (Pohe 2019). Each stream had three light trapping sites to promote sufficient representation of the species present, which were pooled by location to minimise distributional bias (Zurell et al. 2020). Sampling of benthic macroinvertebrates from streams in 2020 and 2021 provided additional records of *Ichthybotus* nymphs. Verified presence data were also obtained from the New Zealand Ministry for the Environment's freshwater invertebrate monitoring database, specimens in the Otago Museum, Canterbury Museum, Museum of New Zealand–Te Papa Tongarewa, the New Zealand

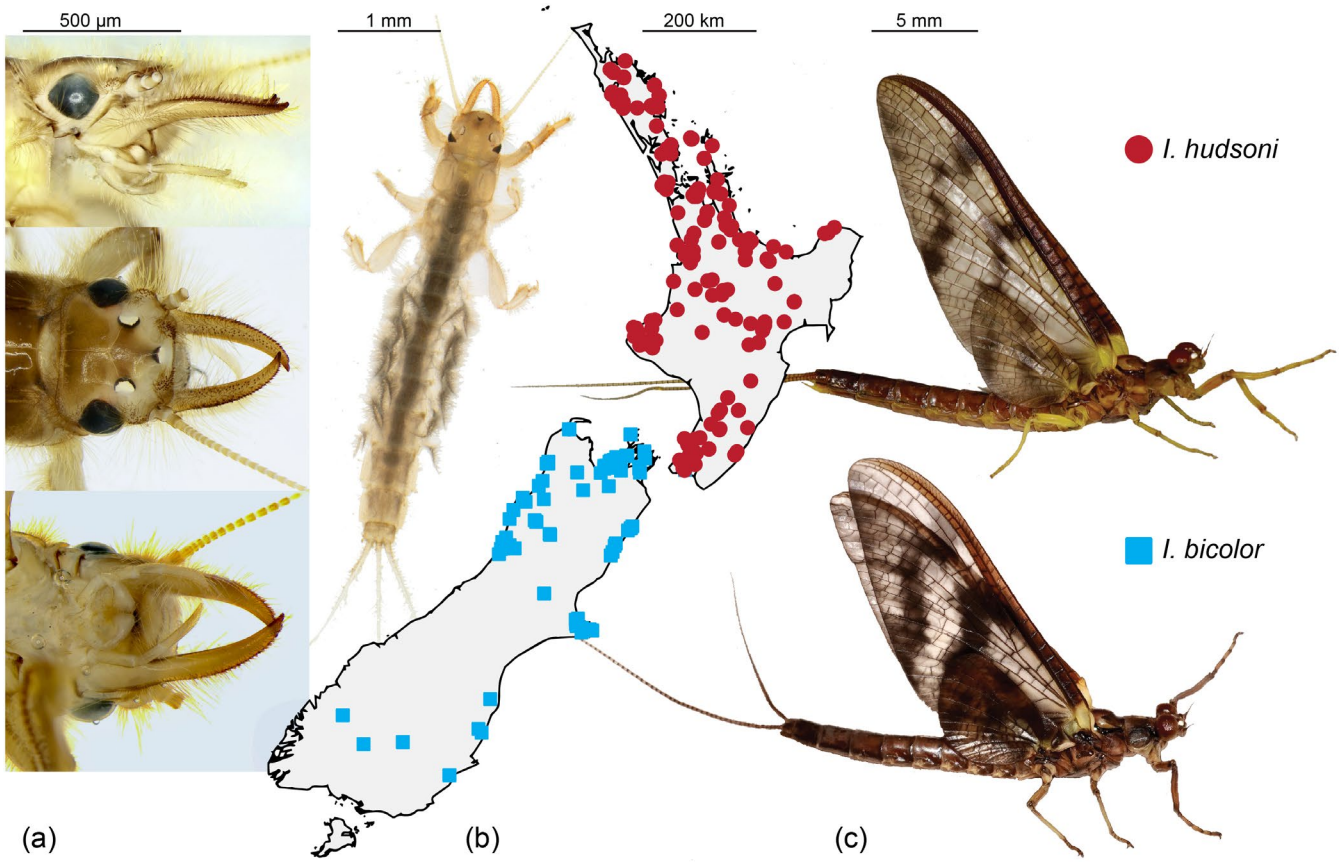

FIGURE 2 | *Ichthybotus* (Ichthybotidae) mayfly nymphs illustrating the large mandibular tusks (a); the distribution of *I. hudsoni* and *I. bicolor* in Aotearoa–New Zealand's North Island and South Island, respectively (b); and male subimagos of *I. hudsoni* and *I. bicolor* (c).

Arthropod Collection and the iNaturalist citizen science database (https://inaturalist.nz). Specimens or photographs were viewed to determine hindwing colour before incorporating records into the dataset as correct identification is crucial for modelling species distributions (Araújo et al. 2019; Zurell et al. 2020). Site coordinates were rounded to match the scale of the climate data, and replicate occurrence records were removed.

Binary presence/absence tables were constructed for each species (*I. hudsoni* present $n = 329$, *I. bicolor* present $n = 134$) including data from surveyed streams that held other mayflies but no *Ichthybotus* (absence $n = 45$; Pohe 2019). The two *Ichthybotus* species were modelled separately to relate their current distributions to climate variables. Our initial models treated stream sites as absences if the focal species had not been recorded (hereafter all-absence data) on the assumption that realised niche is not influenced by interactions between sister species (Zurell et al. 2020).

We repeated models for each species using an island-specific approach in which maps were cropped to consist of data for just South Island or just North Island. Presence data did not change, but absence data reduced and the difference between maximum and minimum climatic variables included in the models was narrower (reduced environmental amplitude) due to the limited land area that is currently accessible to each mayfly lineage. In addition, we considered the possibility of competitive exclusion between the species by modelling with a subset of 45 'absence' sites at which mayflies but neither *Ichthybotus* species were recorded. For example, if allopatry results from competition between the two *Ichthybotus* species, then the presence of *I. hudsoni* would not indicate unsuitable climate for *I. bicolor*. The subset SDMs used location data for which both species were absent in North Island ($n = 10$) and South Island ($n = 35$).

2.3.2 | Environmental Variables

We accessed 19 rainfall and temperature variables from WorldClim 2 (https://www.worldclim.org/; Fick and Hijmans 2017) averaged for 1960–1990, and data for the climate during the LGM at a resolution of 2.5 arc min (~5 km^2), cropped to the extent of NZ (or cropped to single island in alternative models). Variance inflation factor (VIF) analysis was implemented in the R package *usdm* (Naimi et al. 2014; R Core Team 2024) using the mutual information-VIF method (Cheng et al. 2022) to identify collinear variables and minimise overfitting during the modelling process. The 12 climate variables with strong collinearity (VIF > 5) were excluded leaving seven variables for further analysis (Table S1). Each variable included in downstream modelling was considered a proxy for all climatic variables with which it correlated.

2.3.3 | Distribution Modelling

Separate SDMs of *I. bicolor* and *I. hudsoni* were modelled with the R package *biomod2* 3.5.1 (Thuiller 2003) that implements

an ensemble approach by generating consensus distributions based on individual model projections. This approach reduces the effect of individual model discrepancies and can give more robust estimates through superior separation of signal and noise (Araújo and New 2007; Hao et al. 2019; Čengić et al. 2020). Four modelling algorithms were used to analyse the presence/absence data against the seven climate variables (Table S2): generalised boosting model/boosted regression tree, artificial neural network, maximum entropy and random forest (Diniz-Filho et al. 2015). We employed an algorithm sourced from the R package *SDMtune* 1.3.1 (Vignali et al. 2020) to enhance the hyperparameter configuration of the models for increased robustness. To calibrate models, we used 80% of the input data, with the remaining 20% used to test them. The 'prevalence' parameter was set at 0.5, which weighs absences and presences equally and VarImport was set to 3, which permits three permutations to estimate relative variable importance. Spatial-blocking and cross-validation were implemented with the R package *blockCV* 2.1.1 (Valavi et al. 2019) to generate four different blocks, each of which was run four times across the four models, leading to 16 runs and a total of 64 model runs. Well-performing runs were combined and weighted by each algorithm's performance to create the ensemble using a weighted mean approach (EMmw). Models were evaluated using the true skill statistic (TSS) and receiver operating characteristic (ROC) that assess model fit and predictive performance (Allouche et al. 2006). Accuracy of SDMs with TSS > 0.8 and ROC > 0.9 are generally considered excellent (Parreira et al. 2019; Liu et al. 2020).

Relative variable importance for the final ensemble SDM was calculated by applying the weightings used for the ensemble to the variable importance scores for individual models, summing values by variable and then dividing by the number of modelling methods applied from the qualifying subset of runs with ROC score > 0.80. Distribution models for each *Ichthybotus* species were projected onto climate models under current and LGM conditions when lower sea levels resulted in land connection between North Island and South Island NZ. Projections were generated using the ensemble SDM, and binary presence/absence maps were built using cut-off values determined by maximising sensitivity (the correctly predicted proportion of observed presence) and specificity (the correctly predicted proportion of observed pseudo-absence) of the model.

2.4 | Population Genetics

To assess model performance for either forecasting or hindcasting, we predicted how the inferred past distribution of each *Ichthybotus* species would affect current population genetic structure using population genetic principles (Charlesworth 2009). Stable populations retain genetic diversity whereas range expansion during colonisation is expected to leave a signature of reduced genetic variation at the leading edge (Excoffier et al. 2009), and these can be distinguished using population diversity metrics (Figure 1). We sampled 22 populations of *Ichthybotus* from locations inferred to have provided climatically suitable habitat continuously during the LGM, and nine populations from locations for which models suggested these mayflies were absent (Tables 1 and S3). The insect COI intraspecific divergence rate of 0.0792 substitutions per site per million years (Ney et al. 2018) was used to estimate the time since the most recent common ancestor of each lineage. We predicted similar levels of DNA sequence divergence in the two taxa (P1) due to widespread distribution in hindcasted SDMs of both.

We determined the total number of mtDNA COI haplotypes and calculated nucleotide diversity using DnaSP 6.0 (Rozas et al. 2017) for each population sample of > 5 individuals and looked for an association between these measures of diversity and latitude using a linear regression. Expansion of a range restricted species would lead to higher nucleotide diversity per population sample in source compared to sink locations (Bulgarella et al. 2014). Therefore, we predicted no association between nucleotide diversity and latitude where range has been relatively stable (P2). Tajima's D (Tajima 1989) implemented in DnaSP was used to detect departures from the mutation–drift equilibrium indicative of changes in historical demography when applied to selectively neutral markers. We assessed statistical significance using 1000 permutations. Values near zero for Tajima's D imply constant population size (P3), significant negative values suggest population expansion, while positive values indicate population subdivision. To estimate population differentiation we calculated population sample pairwise Φ_{ST} (Excoffier et al. 1992) with ARLEQUIN 3.5.2.2 (Excoffier et al. 2005). Pairwise geographic distances were triangulated from NZ map grid coordinates. To detect isolation by distance we compared pairwise Φ_{ST} with pairwise geographical distances among sampling sites using a Mantel test implemented in R in the package *vegan* 2.6–8 (Oksanen et al. 2024). Genetic structure resulting from spatially restricted gene flow would leave a signature of isolation by distance (P4) even in non-equilibrium populations (Slatkin 1993). Evolutionary relationships among haplotypes were inferred with median-joining network (Bandelt et al. 1999) implemented in PopArt (Leigh and Bryant 2015). In a growing population most coalescent events occur relatively early and the distribution of pairwise differences will be close to Poisson, therefore population expansion results in characteristic star-shaped haplotype networks (star phylogeny; Slatkin and Hudson 1991), that is not predicted if range has been stable (P5).

3 | Results

3.1 | Distinction Between the Two Taxa

North Island ($n = 145$) and South Island ($n = 80$) mayfly samples from 31 locations comprised 77 and 25 mtDNA COI haplotypes respectively. Despite sampling from the northern South Island and Taranaki in North Island, where the two islands were connected during the last glacial maximum (LGM; Trewick and Bland 2012), no haplotypes were shared between the two *Ichthybotus* mayfly species (Figure 3c). The two *Ichthybotus* species have distinct mtDNA haplotype clusters that are concordant with spatial distribution and colour variation of the hindwing and were distinguished by canonical variate analysis of forewing morphology (Figures 3 and S8, Table S6). The average and maximum pairwise interspecific genetic distances were 4.1% and 4.7% respectively. Using an interspecific insect substitution rate the most recent common ancestor of the two species was estimated to have existed in the Pleistocene (approximately 1.4 Mya ± 0.2).

TABLE 1 | Sampling of *Ichthybotus* mayflies for population genetics.

Site name	Species	LGM	*n*	h	Hd	π	Tajima D
Waiarakau Stream	*I. hudsoni*	Yes	7	6	0.952	0.0034	−0.734
Tapapa Stream	*I. hudsoni*	Yes	12	7	0.833	0.0062	0.166
Puketi Forest	*I. hudsoni*	Yes	1	1	N/A	N/A	N/A
Punaruku Stream	*I. hudsoni*	Yes	14	11	0.956	0.0069	−0.869
Waipoua River	*I. hudsoni*	Yes	10	7	0.911	0.0066	−1.520
Mangere Stream	*I. hudsoni*	Yes	8	8	1	0.0094	−0.291
Pohuehue Stream	*I. hudsoni*	Yes	12	7	0.894	0.0095	−0.093
Waitakere River	*I. hudsoni*	Yes	1	1	N/A	N/A	N/A
Karamatura Stream	*I. hudsoni*	Yes	8	7	0.964	0.0102	−1.359
Hunua Ranges	*I. hudsoni*	Yes	7	6	0.952	0.0050	−0.213
Fantail Bay	*I. hudsoni*	Yes	12	6	0.758	0.0036	−0.682
Te Puru Stream	*I. hudsoni*	Yes	7	6	0.952	0.0064	−0.233
Ratapihipihi Reserve	*I. hudsoni*	No	11	5	0.764	0.0028	−1.549
Kahutarawa Stream	*I. hudsoni*	No	12	5	0.667	0.0027	−1.074
Otari-Wiltons	*I. hudsoni*	Yes	8	5	0.786	0.0024	−1.640
Wainuiomata	*I. hudsoni*	Yes	5	4	0.9	0.0042	−0.668
Monckton Reserve	*I. hudsoni*	No	2	2	1	0.0049	0
Miller Reserve	*I. hudsoni*	No	7	4	0.714	0.0037	−0.339
Kaituna River	*I. bicolor*	Yes	9	6	0.889	0.003	−0.740
Lyell Creek	*I. bicolor*	No	1	1	N/A	N/A	N/A
Harvey Bay Stream	*I. bicolor*	Yes	12	3	0.758	0.0019	−1.633
Pukaka Stream	*I. bicolor*	Yes	12	3	0.318	0.0005	−1.451
Elvy Stream	*I. bicolor*	Yes	6	4	0.8	0.0022	−1.295
Lankey Creek	*I. bicolor*	No	1	1	N/A	N/A	N/A
Mororimu Stream	*I. bicolor*	Yes	1	1	N/A	N/A	N/A
Big Bush Gully	*I. bicolor*	Yes	3	3	1	0.0022	0
Okuti River	*I. bicolor*	No	1	1	N/A	N/A	N/A
Narbey Stream	*I. bicolor*	Yes	2	2	1	0.0008	0
Trotters Gorge	*I. bicolor*	Yes?	12	1	0	0	N/A
Taieri Beach Road	*I. bicolor*	No?	16	6	0.783	0.0023	−0.234
Boyd Creek	*I. bicolor*	No	4	2	0.5	0	−0.612

Note: LGM = hindcast projection inferred that the location was climatically suitable for continuous occupation by current *Ichthybotus* species (Yes) or not suitable (No). mtDNA summary statistics, n = sample size, h = number of haplotypes, Hd = haplotype diversity, π = nucleotide diversity.

3.2 | Species Distributions

All ensemble SDMs performed well according to receiver operating characteristics (ROC > 0.98) and true skill statistics (TSS > 0.86; Figure 4, Table S2). Models with all-absence data scored highest for specificity for both species (94.4%, 97.9%) and highest for sensitivity for *I. hudsoni* (99.4%). Evaluation scores, variable importance and projections of SDMs for the same species were not identical, due to the influence of absence data (Tables S2 and S4; Figures S1 and S2). In the SDM for *I. hudsoni* that used the maximum number of absence points, the relative importance of climate variables ranged from 0.39 to 0.03 (Table S4). In the *I. bicolor* all-absences SDM, climate variables ranged from 0.3 to 0.07 in relative importance (Table S4). The two most influential variables (and their proxies) for *I. hudsoni* and *I. bicolor* were temperatures: mean temperature of driest quarter and maximum temperature of warmest month (Table S4).

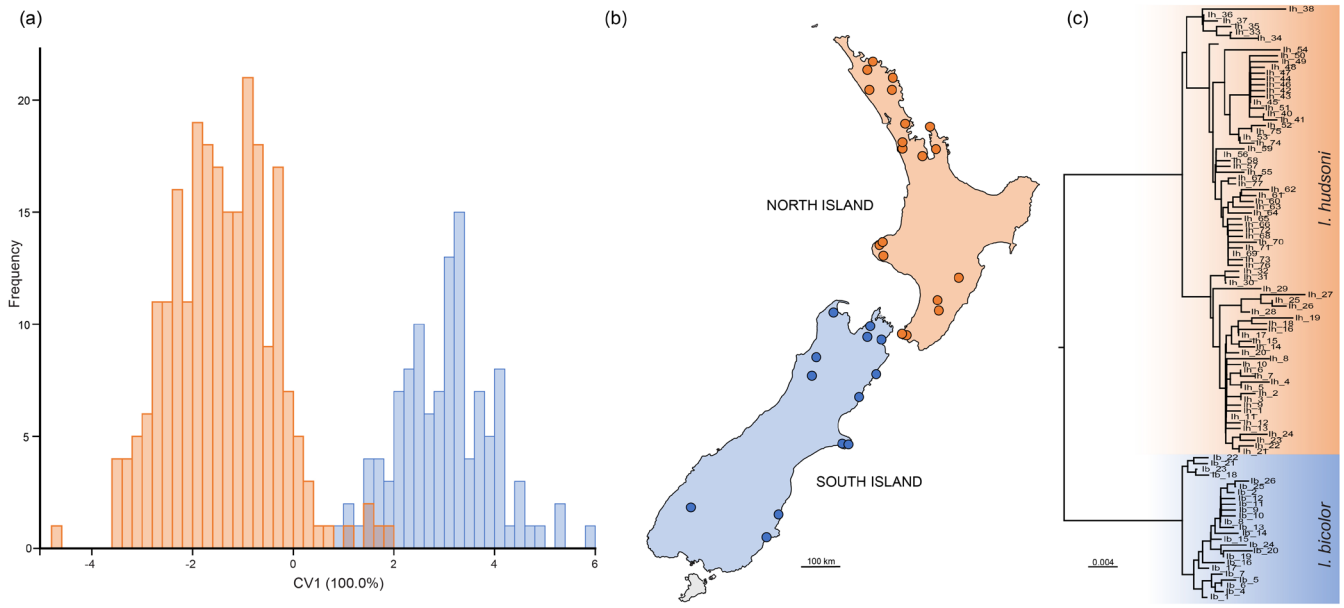

FIGURE 3 | Different species of *Ichthybotus* mayfly found on the two main islands of Aotearoa–New Zealand. Frequency distribution of canonical variate 1 (CV1) for forewing shape of *Ichthybotus* adults classified by island (a); collection locations of the mayflies *Ichthybotus hudsoni* (North Island) and *I. bicolor* (South Island) (b); and neighbour-joining tree of mtDNA COI haplotypes from the two mayfly species (c).

SDMs using fewer absence points in the island-specific and competitive exclusion (subset model) analyses also scored highly for ROC (0.99 to 0.98) and TSS (0.94 to 0.88) and inferred different relative influences of the seven climate variables. For *I. hudsoni*, the island-specific model had six variables ranging from 0.19 to 0.13 in relative importance, with only isothermality scoring low (0.05; Table S4). In contrast, the subset model suggested maximum temperature of warmest month was strongly influential (0.48; Table S4). For *I. bicolor*, the island-specific model had the same rankings for climate variables as the all-absences model except precipitation of the coldest quarter had the least influence (0.05; Table S4). In contrast, the subset model suggested maximum temperature of warmest month was most influential (0.46; Table S4). SDMs with all-absence data projected onto current climatic conditions matched current mayfly distributions more closely than SDMs with subsets of absence data (Figures 4 and S2). Island-specific SDMs could only be projected onto a single NZ island (Figure S1). The SDMs for *I. hudsoni* and *I. bicolor* with all-absence data identified relatively small areas of climatically suitable land where the mayfly species has not been recorded in northern South Island and central North Island, respectively (Figure 4).

Hindcasting distribution models onto estimates of climate during the LGM provided predictions of species ranges in the past. We projected the three SDMs for each species for comparison (Figures 4, S1 and S2). For *I. hudsoni*, a large area of suitable LGM habitat from the northernmost point southwards on the western side of NZ was identified. On eastern North Island, coastal areas as far as the Cook Strait, which currently divides North and South Islands, were identified as being climatically suitable for *I. hudsoni* during the LGM (Figure 4). The island-specific model for *I. hudsoni* inferred a greater area of climatically suitable space during LGM than either all-absences or a subset of absences. None of the models indicated less climatically suitable space compared to their current distribution. The potential distribution of *I. bicolor* during the LGM covered central North Island, extending south on both the western and eastern coasts of South Island (Figure 4). The island-specific approach for *I. bicolor* inferred climatically suitable space during LGM further north than the models with either all-absences or a subset of absences. All three models inferred climatically suitable habitat during the LGM on coastal South Island with no reduction in total area available for *I. bicolor*. Thus, our SDM projections imply neither species would have experienced a reduction in total area of climatically suitable space to occupy during the LGM, with minor changes in range limits. All of our SDMs for the same species derived similar population genetic predictions from hindcasting, despite incorporating a range of absence data and different environmental amplitudes.

3.3 | Population Genetics

For *I. hudsoni* we sampled 145 individuals in total from 12 locations inferred from hindcasted SDMs to have provided stable habitat (Table 1: LGM = yes). We sampled 30 *I. hudsoni* in total from three populations inferred to occupy colonised sites (Table 1: LGM = no). In South Island, we sampled 6–12 *I. bicolor* from five putatively stable population locations (51 individuals) and one population inferred to have colonised since the LGM ($n = 16$; Table 1). Sampled haplotypes differed by a maximum of 2.7% pairwise DNA sequence divergence within *I. hudsoni* (Figure 5c) and a maximum of 1.2% within *I. bicolor* (Figure 5c), contrary to our prediction of similar levels of DNA sequence divergence in the two species (P1; Figure 6). Given this level of differentiation, we estimate ~170,500–76,000 years since the most recent common mtDNA ancestor for each species (i.e., before LGM).

The highest haplotype diversity observed within a population sample of *I. hudsoni* was from Punaruku Stream in Northland

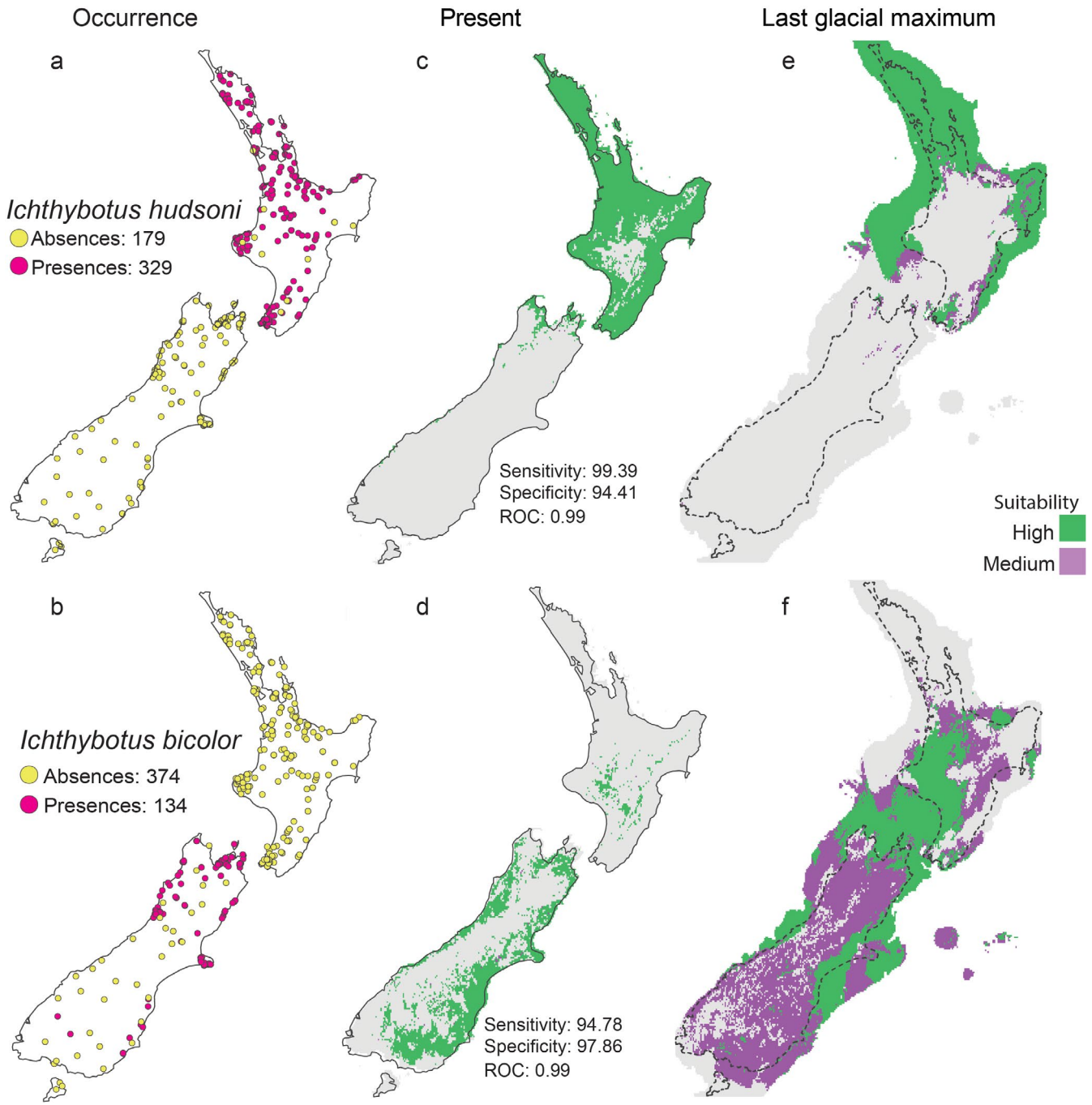

FIGURE 4 | Occurrence records for mayflies *Ichthybotus hudsoni* (a) and *I. bicolor* (b) in Aotearoa–New Zealand and species distribution model projections for present day (c, d) and the last glacial maximum (e, f) climate. Land available shown in grey. Projections indicate potential niche availability, with areas of high (≥ cutoff value maximising sensitivity and specificity) and moderate climate suitability (70% of cutoff value). Model fit determined by receiver operating characteristic (ROC).

(11 haplotypes from a sample of 14 individuals; Table 1). The most haplotypes within a population sample of *I. bicolor* were six, observed from the most southern site ($n = 16$) and the most northern site ($n = 9$; Table 1). Population sample nucleotide diversity (π) of *I. hudsoni* was higher in the north (Table 1, Figure 5a) with an association between latitude and π (linear regression, $R^2 = 0.3133$, $F = 6.387$, $p < 0.05$), contrary to our prediction that both species would have high genetic diversity over their entire range (P2). The cline in genetic diversity from north to south observed in *I. hudsoni* was not predicted by SDM hindcasting (Figure 6). Population sample π estimates for *I. bicolor* were not associated with latitude ($R^2 = 0.2301$, $F = 2.092$, $p = 0.191$; Figure 5a), and were lower than for *I. hudsoni*, contrary to P2 (Figure 6).

For 13 of 15 population samples of *I. hudsoni*, Tajima's D near zero suggested stable demographics, supporting our prediction of constant population size (P3), whereas Tajima's D was negative for two samples from locations within the inferred LGM species distribution (Waipoua River, −1.52, $p = 0.047$;

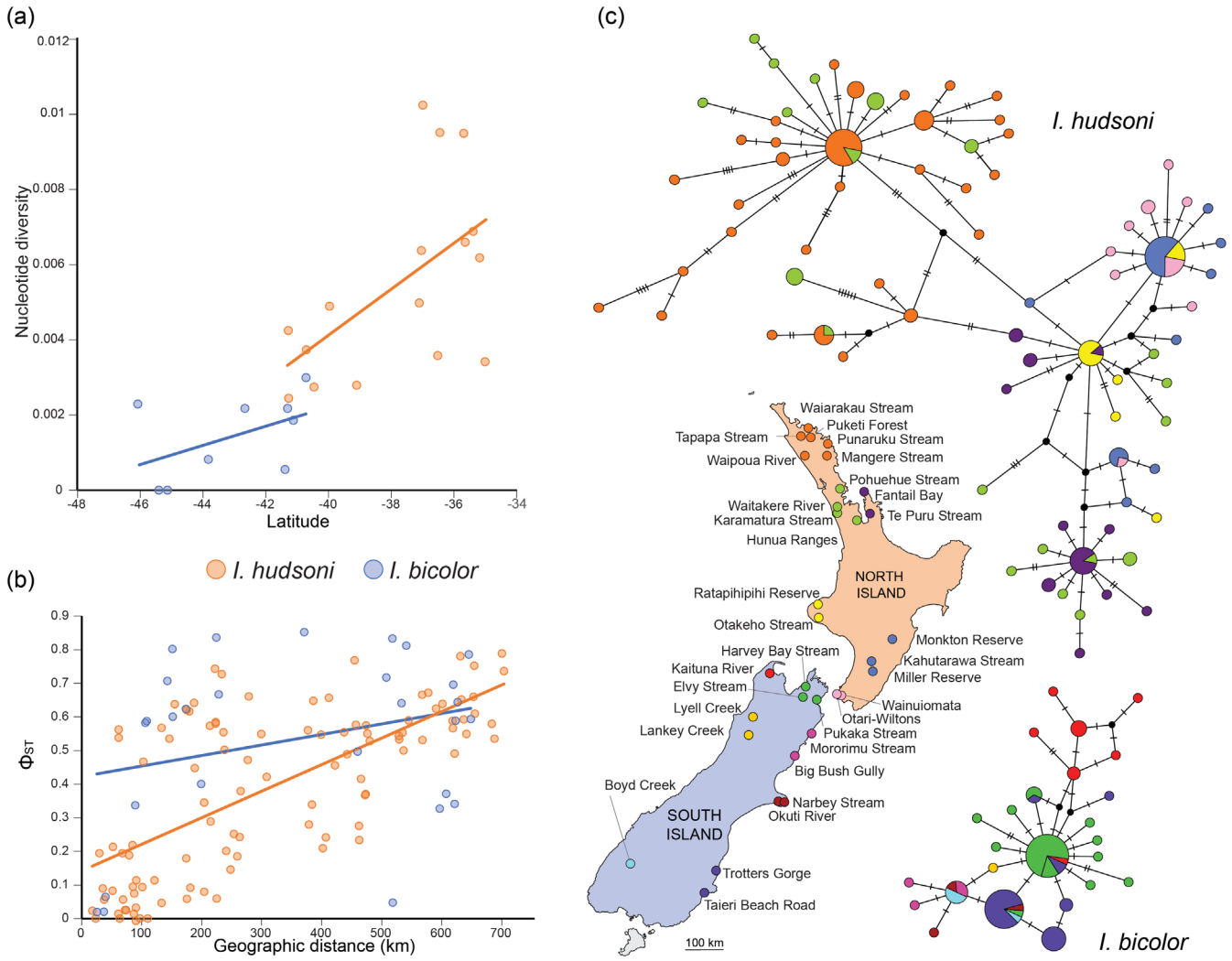

FIGURE 5 | Spatial genetic structure within two species of *Ichthybotus* mayflies inferred from COI mtDNA. Relationship between latitude and nucleotide diversity for 25 population samples (a); and between pairwise geographic distance (km) and genetic distance (Φ_{ST}) (b); and median-joining networks for the two species with colours indicating locations in the North Island (*I. hudsoni*) and South Island New Zealand (*I. bicolor*) (c).

Otari-Wilton, −1.64, $p = 0.034$). Tajima's D was near zero for five population samples of *I. bicolor*, supporting P3 (Figure 6) and negative in the sample from the putatively stable habitat of Harvey Bay Stream, suggesting population growth (−1.63, $p = 0.047$; Table 1).

We detected a positive association between pairwise geographic distance and pairwise genetic distance (Φ_{ST}) in *I. hudsoni* as predicted for isolation by distance (P4) but not in *I. bicolor* (Mantel test *I. hudsoni* $r = 0.68$, $p = 0.001$; *I. bicolor* $r = 0.28$, $p = 0.11$; Figure 5b). The evolutionary relationships inferred for *I. hudsoni* haplotypes (Figure 5c) formed a spatially structured network supporting P5 (Figure 6). Specimens from the northern half of the range (Northland, Auckland) had haplotypes from all clusters, whereas samples from the southern part of the *I. hudsoni* distribution had haplotypes from only one cluster (Figure 5c). The evolutionary relationships inferred for the haplotypes from the South Island species *I. bicolor* showed the star-shape characteristic of population expansion (Slatkin and Hudson 1991), contrary to our prediction of spatially structured haplotype networks (P5, Figure 6).

4 | Discussion

Studying closely related species can advance understanding of ecological differences and identify biotic interactions that limit geographic distributions (McCormack et al. 2010; Bulgarella et al. 2014; Sivyer et al. 2018). We found complete concordance between distinct hindwing pigmentation and mtDNA lineages of *I. hudsoni* and *I. bicolor*. However, minor wing shape and habitat differences suggest the two mayfly species are ecologically similar and their allopatry could indicate competitive exclusion. Competitive exclusion will impact SDMs if presence data is not a good estimate of the climatic envelope a species can tolerate, which is why our models incorporating data for locations where neither *Ichthybotus* was present are valuable. Hindcast distributions provided similar population genetic predictions from models with distinct absence data, suggesting only partial ecological overlap of the two species.

SDMs hindcast climatically suitable LGM habitat for *I. hudsoni* primarily in northern and western coastal North Island

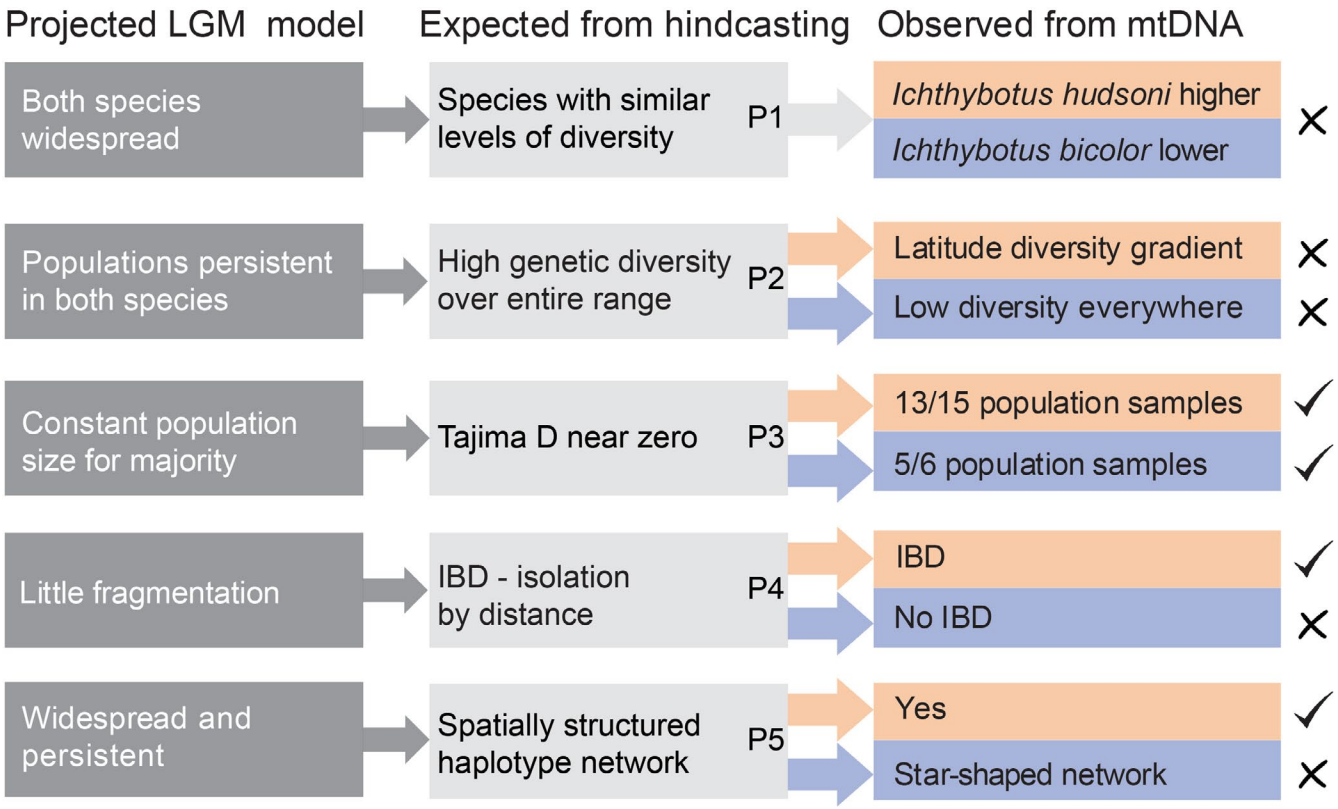

FIGURE 6 | Hindcasting species distributions during the last glacial maximum (LGM) based on models built on climate variables yield predictions (P1–P5) of intraspecific genetic diversity for New Zealand mayflies *Ichthybotus hudsoni* (peach) and *I. bicolor* (blue) in current conditions.

including a now submerged western area. As predicted, genetic diversity in most North Island population samples indicated large stable *Ichthybotus* populations during and after the LGM. Accordingly, our predictions for isolation by distance (P4) and haplotype network shape (P5) were met and we detected little evidence of population growth (P3). Nevertheless, the significant association between latitude and nucleotide diversity within *I. hudsoni* could have arisen from range expansion. SDMs for *I. hudsoni* projected onto LGM climate indicated a slightly reduced area of suitable habitat in southern North Island, which could be inferred as consistent with lower observed population genetic diversity here. However, inferred continuous suitable space in southern and coastal North Island during the last glacial period was sufficient to predict the maintenance of continuous populations of these insects and thus we predicted retention of genetic diversity. Thus, only three of the mayfly population genetic patterns matched the predictions of hindcasted SDMs for *I. hudsoni*. Volcanic activity in North Island NZ coinciding with the LGM (Muscheler et al. 2020) might explain the limited support for our population genetic predictions.

Ensemble SDMs for *I. bicolor* scored highly for accuracy and predicted climatically suitable areas around northern and coastal South Island, suggesting retention of large populations during the LGM. Contrary to predictions derived from SDMs of similar levels of DNA sequence divergence in the two taxa (P1), genetic diversity was lower within *I. bicolor* than *I. hudsoni*. Contrary to our prediction that both species would have high genetic diversity over their entire range (P2), we observed low genetic diversity in population samples of *I. bicolor*. Lack of association between latitude and nucleotide diversity (expected from expansion south from stable northern populations) might be explained by the limited number of population genetic samples analysed for *I. bicolor*. Our prediction of constant population size (P3) was supported by Tajima's D estimates, which were significantly negative in only one *I. bicolor* and two *I. hudsoni* population samples. Contrary to our prediction of spatial structuring (P5), the haplotype network for *I. bicolor* was star-shaped, consistent with population expansion (Slatkin and Hudson 1991). The pattern and level of genetic variation detected suggests *I. bicolor* is not at demographic equilibrium and may have only recently colonised the area we sampled (Slatkin 1993), despite projections of suitable continuous occupancy in large areas of central NZ and coastal South Island. Although the disparity between hindcast distribution from climate models and genetic diversity might result from sampling just 80 individuals of this species, analysed data represented its full distribution. As such, if haplotype divergence within *I. bicolor* was as high as within *I. hudsoni*, it would likely have been detected.

Although SDMs assume that realised distribution encompasses most potential habitat (Zurell et al. 2020), landscape dimensions and competition constrain the distribution of many species, limiting the ability to determine the full environmental tolerance (ecological amplitude) of a lineage (Arntzen 2023). Our comparison of SDMs for two congeneric mayflies suggests that climate data and geographic distributions can inform models that match their current ranges fairly well. However, despite the similarity between the observed and modelled range, we found evidence that these models may lack biological realism. Hindcasting

the potential LGM climate space for these mayflies did not accurately predict the limited genetic diversity detected in *I. bicolor* and cannot explain the association between latitude and diversity in *I. hudsoni*. Despite inference of climatically suitable space, it is plausible that interspecific competition between these ecologically similar species might have limited the distribution of one or both in southern North Island (Trewick et al. 2017). Our genetic data represent a single locus and therefore cannot provide a full picture of the population structure and demography of these two species. Thus, the mismatch between hindcasted potential distributions and current population genetic structure might reflect the resolution of the genetic marker used and/or sample sizes too small to detect significant associations. However, lack of resolution can explain neither the cline detected nor the difference in diversity levels observed in mtDNA.

SDMs are only biologically meaningful if they use relevant predictors of distribution (Mpakairi et al. 2017). Factors other than climate affecting freshwater biodiversity include habitat structure (i.e., substrate), water quality and land use (Daufresne and Boët 2007; Ormerod et al. 2010). Biotic interactions (e.g., predation, competition, food, disease and parasitism) that influence species distributions are usually not elucidated (Heikkinen et al. 2006; Lavergne et al. 2010; Ortego and Knowles 2020). Ecological competition can prevent species distributions reaching a stable state (Wisz et al. 2013; Arntzen 2023). For example, interpreting the distribution of two species of nocturnal forest Orthoptera relied on incorporating competitive exclusion (Bulgarella et al. 2014). More complex models can be built to capture occurrence dynamics including spatial dependence, therefore relaxing the species–environment equilibrium assumption that species fill their niche and do not occur elsewhere (Zurell et al. 2020). Inclusion of biotic factors would improve model fit and accuracy, although they are difficult to account for and even harder to model for climate change (Moilanen et al. 2008; Morin and Thuiller 2009; Boulangeat et al. 2012).

Species distribution modelling works on the assumption that the system of interest is in equilibrium. This is often not the case, particularly in freshwater systems that have a dynamic equilibrium where biodiversity constantly changes with phases of population reduction and expansion (Huston 1979). Floods, drought and other disturbances are important determinants of community structure in freshwater (Resh et al. 1988; Lepori and Hjerdt 2006), and these ecosystems are often recovering from disturbances that are not considered in SDMs (Haghkerdar et al. 2019). Climate variables may be too coarse to account for stochastic changes that alter community composition and genetic spatial structure (Daufresne and Boët 2007). The climate variables used in our models (temperature and rainfall) have limited capacity to represent such disturbances, and other phenomena such as volcanic eruptions were not included in our models. Incorporating hydrological, geological and biotic variables might improve forecasting, but model evaluation will be even more challenging.

Although outputs from SDMs give an impression of certainty, model assumptions and predictions are difficult to evaluate (Wiens et al. 2010; Warren et al. 2020). We used genetic data to evaluate the ability of SDMs to infer past population distributions. Metrics indicative of model accuracy suggest our SDMs fitted our data well. Nevertheless, deviations from predicted genetic patterns in both mayfly species suggest the models had limited capacity to predict past distributions. High-resolution genotypic data could give more sensitive validation of predictions from SDMs (Currat et al. 2019; Martin et al. 2023), but we recommend caution when hindcasting and forecasting potential environmental space for freshwater organisms as climate models may not provide realistic predictions of past or future distributions. The poor performance of our hindcast SDMs also reveals the limitations of studying species in isolation. As such, we suggest that SDMs using climate variables should incorporate multiple taxa from the same communities in order to identify the species with distributions most susceptible to climate change. The assumption that ecological amplitude can be inferred from presence data might be valid for only a minority of taxa. Biotic interactions will make freshwater communities vulnerable to the loss of temperature-sensitive species, but we are optimistic that incorporating additional biologically relevant variables including biotic interactions into models will improve predictions.

Author Contributions

Conceptualisation: S.A.T., S.R.P., M.M.-R. Developing methods: D.C.-R., S.A.T., M.M.-R., F.Z.-V. Data gathering and analysis: D.C.-R., S.A.T., C.B., S.R.P., F.Z.-V. Preparation of figures and tables, conducting the research, data interpretation, writing: D.C.-R., S.A.T., C.B., S.R.P., M.M.-R., F.Z.-V.

Acknowledgements

The Department of Conservation, Auckland Regional Council, Gisborne District Council and Queen Elizabeth II National Trust provided collecting permits or written approvals to undertake research on their respective administered lands. The following kaitiaki granted approval for site access and specimen collection: Te Rūnanga o Whaingaroa, Te Rūnanga o Te Rarawa, Puketi Forest Trust, Te iwi o Ngātiwai, Te iwi o Te Roroa, Pukenui Western Hills Forest Trust, Nga Ahi Kaa, Te Uri o Hau, Little Barrier Island Supporters Trust, Ngāti Manuhiri, Tūhoe Te Mana Motuhake mai Waikaremōana Tribal, Ngāi Tāmanuhiri Whānui, Eastern Taranaki Environmental Trust, Rangitāne o Manawatū, Te Rūnanga o Ngāi Tahu, Maurice White Forest Trust and five private landowners. Funding support was provided to C.B. by a Graduate Women New Zealand Manawatu Postgraduate Scholarship. Open access publishing was facilitated by Massey University, as part of the Wiley - Massey University agreement via the Council of Australasian University Librarians. Open access publishing facilitated by Massey University, as part of the Wiley - Massey University agreement via the Council of Australasian University Librarians.

Funding

This work was supported by Graduate Women New Zealand Manawatu.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are available from Figshare 10.6084/m9.figshare.30757607. DNA sequences are available on GenBank: accession numbers PP778198–PP778299 and via BOLD: https://portal.boldsystems.org/result?query=Ichthybotidae[tax].

References

Allouche, O., A. Tsoar, and R. Kadmon. 2006. "Assessing the Accuracy of Species Distribution Models: Prevalence, Kappa and the True Skill Statistic (TSS)." *Journal of Applied Ecology* 43, no. 6: 1223–1232. https://doi.org/10.1111/j.1365-2664.2006.01214.x.

Araújo, M. B., R. P. Anderson, A. Márcia Barbosa, et al. 2019. "Standards for Distribution Models in Biodiversity Assessments." *Science Advances* 5, no. 1: eaat4858. https://doi.org/10.1126/sciadv.aat4858.

Araújo, M. B., and M. New. 2007. "Ensemble Forecasting of Species Distributions." *Trends in Ecology & Evolution* 22, no. 1: 42–47. https://doi.org/10.1016/j.tree.2006.09.010.

Arntzen, J. W. 2023. "A Two-Species Distribution Model for Parapatric Newts, With Inferences on Their History of Spatial Replacement." *Biological Journal of the Linnean Society* 138, no. 1: 75–88. https://doi.org/10.1093/biolinnean/blac134.

Bandelt, H. J., P. Forster, and A. Rohl. 1999. "Median-Joining Networks for Inferring Intraspecific Phylogenies." *Molecular Biology and Evolution* 16, no. 1: 37–48. https://doi.org/10.1093/oxfordjournals.molbev.a026036.

Boulangeat, I., D. Gravel, and W. Thuiller. 2012. "Accounting for Dispersal and Biotic Interactions to Disentangle the Drivers of Species Distributions and Their Abundances." *Ecology Letters* 15, no. 6: 584–593. https://doi.org/10.1111/j.1461-0248.2012.01772.x.

Bulgarella, M., S. A. Trewick, N. A. Minards, M. J. Jacobson, and M. Morgan-Richards. 2014. "Shifting Ranges of Two Tree Weta Species (*Hemideina* spp.): Competitive Exclusion and Changing Climate." *Journal of Biogeography* 41, no. 3: 524–535. https://doi.org/10.1111/jbi.12224.

Čengić, M., J. Rost, D. Remenska, J. H. Janse, M. A. J. Huijbregts, and A. M. Schipper. 2020. "On the Importance of Predictor Choice, Modelling Technique, and Number of Pseudo-Absences for Bioclimate Envelope Model Performance." *Ecology and Evolution* 10, no. 21: 12307–12317. https://doi.org/10.1002/ece3.6859.

Charlesworth, B. 2009. "Effective Population Size and Patterns of Molecular Evolution and Variation." *Nature Reviews Genetics* 10, no. 3: 195–205. https://doi.org/10.1038/nrg2526.

Cheng, J., J. Sun, K. Yao, M. Xu, and Y. Cao. 2022. "A Variable Selection Method Based on Mutual Information and Variance Inflation Factor." *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* 268: 120652. https://doi.org/10.1016/j.saa.2021.120652.

Chisholm, W. P. 1984. *Aspects of the Biology of Ichthybotus hudsoni (McLachlan) (Ephemeroptera: Ephemeridae)*. https://ir.wgtn.ac.nz/handle/123456789/26846.

Currat, M., M. Arenas, C. S. Quilodràn, L. Excoffier, and N. Ray. 2019. "SPLATCHE3: Simulation of Serial Genetic Data Under Spatially Explicit Evolutionary Scenarios Including Long-Distance Dispersal." *Bioinformatics* 35, no. 21: 4480–4483.

Daufresne, M., and P. Boët. 2007. "Climate Change Impacts on Structure and Diversity of Fish Communities in Rivers." *Global Change Biology* 13, no. 12: 2467–2478. https://doi.org/10.1111/j.1365-2486.2007.01449.x.

Diniz-Filho, J. A., H. Rodrigues, M. P. Telles, et al. 2015. "Correlation Between Genetic Diversity and Environmental Suitability: Taking Uncertainty From Ecological Niche Models Into Account." *Molecular Ecology Resources* 15, no. 5: 1059–1066. https://doi.org/10.1111/1755-0998.12374.

Domínguez-Domínguez, O., E. Martínez-Meyer, L. Zambrano, and G. P.-P. De León. 2006. "Using Ecological-Niche Modeling as a Conservation Tool for Freshwater Species: Live-Bearing Fishes in Central Mexico." *Conservation Biology* 20, no. 6: 1730–1739. https://doi.org/10.1111/j.1523-1739.2006.00588.x.

Domisch, S., M. B. Araújo, N. Bonada, S. U. Pauls, S. C. Jähnig, and P. Haase. 2013. "Modelling Distribution in European Stream Macroinvertebrates Under Future Climates." *Global Change Biology* 19, no. 3: 752–762. https://doi.org/10.1111/gcb.12107.

Dormann, C. F., M. Bobrowski, D. M. Dehling, et al. 2018. "Biotic Interactions in Species Distribution Modelling: 10 Questions to Guide Interpretation and Avoid False Conclusions." *Global Ecology and Biogeography* 27, no. 9: 1004–1016. https://doi.org/10.1111/geb.12759.

Excoffier, L., M. Foll, and R. J. Petit. 2009. "Genetic Consequences of Range Expansions." *Annual Review of Ecology, Evolution, and Systematics* 40, no. 1: 481–501. https://doi.org/10.1146/annurev.ecolsys.39.110707.173414.

Excoffier, L., G. Laval, and S. Schneider. 2005. "Arlequin (Version 3.0): An Integrated Software Package for Population Genetics Data Analysis." *Evolutionary Bioinformatics* 1: 117693430500100003. https://doi.org/10.1177/117693430500100003.

Excoffier, L., P. E. Smouse, and J. M. Quattro. 1992. "Analysis of Molecular Variance Inferred From Metric Distances Among DNA Haplotypes: Application to Human Mitochondrial DNA Restriction Data." *Genetics* 131, no. 2: 479–491. https://doi.org/10.1093/genetics/131.2.479.

Fick, S. E., and R. J. Hijmans. 2017. "WorldClim 2: New 1-Km Spatial Resolution Climate Surfaces for Global Land Areas." *International Journal of Climatology* 37, no. 12: 4302–4315. https://doi.org/10.1002/joc.5086.

Folmer, O., M. Black, W. Hoeh, R. Lutz, and R. Vrijenhoek. 1994. "DNA Primers for Amplification of Mitochondrial Cytochrome c Oxidase Subunit I From Diverse Metazoan Invertebrates." *Molecular Marine Biology and Biotechnology* 3, no. 5: 294–299.

Franklin, J., A. J. Potts, E. C. Fisher, R. M. Cowling, and C. W. Marean. 2015. "Paleodistribution Modeling in Archaeology and Paleoanthropology." *Quaternary Science Reviews* 110: 1–14. https://doi.org/10.1016/j.quascirev.2014.12.015.

Haghkerdar, J. M., J. R. McLachlan, A. Ireland, and H. S. Greig. 2019. "Repeat Disturbances Have Cumulative Impacts on Stream Communities." *Ecology and Evolution* 9: 2898–2906. https://doi.org/10.1002/ece3.4968.

Hao, T., J. Elith, G. Guillera-Arroita, and J. J. Lahoz-Monfort. 2019. "A Review of Evidence About Use and Performance of Species Distribution Modelling Ensembles Like BIOMOD." *Diversity and Distributions* 25, no. 5: 839–852. https://doi.org/10.1111/ddi.12892.

He, Q., D. L. Edwards, and L. L. Knowles. 2013. "Integrative Testing of How Environments From the Past to the Present Shape Genetic Structure Across Landscapes." *Evolution* 67, no. 12: 3386–3402. https://doi.org/10.1111/evo.12159.

Heikkinen, R. K., M. Luoto, M. B. Araújo, R. Virkkala, W. Thuiller, and M. T. Sykes. 2006. "Methods and Uncertainties in Bioclimate Envelope Modelling Under Climate Change." *Progress in Physical Geography: Earth and Environment* 30, no. 6: 751–777. https://doi.org/10.1177/0309133306071957.

Hitchings, T., T. Hitchings, and J. D. Ridden. 2022. "*Nesameletus staniczeki*, a New Species of *Nesameletus* (Ephemeroptera: Nesameletidae) From New Zealand." *Records of the Canterbury Museum* 36: 169–177.

Hitchings, T. R. 2008. "The Post Glacial Distribution of New Zealand Mayflies." In *International Advances in the Ecology, Zoogeography, and Systematics of Mayflies and Stoneflies*, edited by F. R. Hauer and A. Welch, 89–101. California Scholarship Online. https://doi.org/10.1525/california/9780520098688.003.0007.

Ho, S. Y. W., and S. Duchêne. 2014. "Molecular-Clock Methods for Estimating Evolutionary Rates and Time Scales." *Molecular Ecology* 23, no. 24: 5947–5956. https://doi.org/10.1111/mec.12953.

Huston, M. 1979. "A General Hypothesis of Species Diversity." *American Naturalist* 113, no. 1: 81–101.

Jansen, J., S. N. C. Woolley, P. K. Dunstan, et al. 2022. "Stop Ignoring Map Uncertainty in Biodiversity Science and Conservation Policy." *Nature Ecology and Evolution* 6: 828–829. https://doi.org/10.1038/s41559-022-01778-z.

Kearse, M., R. Moir, A. Wilson, et al. 2012. "Geneious Basic: An Integrated and Extendable Desktop Software Platform for the Organization and Analysis of Sequence Data." *Bioinformatics* 28, no. 12: 1647–1649. https://doi.org/10.1093/bioinformatics/bts199.

Koot, E. M., M. Morgan-Richards, and S. A. Trewick. 2022. "Climate Change and Alpine-Adapted Insects: Modelling Environmental Envelopes of a Grasshopper Radiation." *Royal Society Open Science* 9: 211596. https://doi.org/10.1098/rsos.211596.

Lavergne, S., N. Mouquet, W. Thuiller, and O. Ronce. 2010. "Biodiversity and Climate Change: Integrating Evolutionary and Ecological Responses of Species and Communities." *Annual Review of Ecology, Evolution, and Systematics* 41: 321–350. https://doi.org/10.1146/annurev-ecolsys-102209-144628.

Lee-Yaw, J., J. L. McCune, S. Pironon, and S. N. Sheth. 2022. "Species Distribution Models Rarely Predict the Biology of Real Populations." *Ecography* 2022, no. 6: e05877. https://doi.org/10.1111/ecog.05877.

Leigh, J. W., and D. Bryant. 2015. "Popart: Full-Feature Software for Haplotype Network Construction." *Methods in Ecology and Evolution* 6, no. 9: 1110–1116. https://doi.org/10.1111/2041-210X.12410.

Lepori, F., and N. Hjerdt. 2006. "Disturbance and Aquatic Biodiversity: Reconciling Contrasting Views." *BioScience* 56, no. 10: 809–818. https://doi.org/10.1641/0006-3568(2006)56[809:DAABRC]2.0.CO;2.

Liu, C., C. Wolter, W. Xian, and J. M. Jesche. 2020. "Species Distribution Models Have Limited Spatial Transferability for Invasive Species." *Ecology Letters* 23: 1682–1692. https://doi.org/10.1111/ele.13577.

Maguire, K. C., D. Nieto-Lugilde, M. C. Fitzpatrick, J. W. Williams, and J. L. Blois. 2015. "Modeling Species and Community Responses to Past, Present, and Future Episodes of Climate and Ecological Change." *Annual Review of Ecology, Evolution, and Systematics* 46: 343–368. https://doi.org/10.1146/annurev-ecolsys-112414-054441.

Malhi, Y., J. Franklin, N. Seddon, et al. 2020. "Climate Change and Ecosystems: Threats, Opportunities and Solutions." *Philosophical Transactions of the Royal Society, B: Biological Sciences* 375, no. 1794: 20190104. https://doi.org/10.1098/rstb.2019.0104.

Martin, P., B. C. Haller, P. L. Ralph, and F. Racimo. 2023. "*Slendr*: A Framework for Spatio-Temporal Population Genomic Simulations on Geographic Landscapes." *Peer Community Journal* 3. https://doi.org/10.24072/pcjournal.354.

McCormack, J. E., A. J. Zellmer, and L. L. Knowles. 2010. "Does Niche Divergence Accompany Allopatric Divergence in *Aphelocoma* Jays as Predicted Under Ecological Speciation?: Insights From Test With Niche Models." *Evolution; International Journal of Organic Evolution* 64: 1231–1244. https://doi.org/10.1111/j.1558-5646.2009.00900.x.

Meza-Joya, F. L., M. Morgan-Richards, and S. A. Trewick. 2025. "Forecasting Range Shifts in Terrestrial Alpine Insects Under Global Warming." *Ecology and Evolution* 15: e70810. https://doi.org/10.1002/ece3.70810.

Moilanen, A., J. Leathwick, and J. Elith. 2008. "A Method for Spatial Freshwater Conservation Prioritization." *Freshwater Biology* 53, no. 3: 577–592. https://doi.org/10.1111/j.1365-2427.2007.01906.x.

Morin, X., and W. Thuiller. 2009. "Comparing Niche- and Process-Based Models to Reduce Prediction Uncertainty in Species Range Shifts Under Climate Change." *Ecology* 90, no. 5: 1301–1313. https://doi.org/10.1890/08-0134.1.

Mpakairi, K. S., H. Ndaimani, P. Tagwireyi, T. W. Gara, M. Zvidzai, and D. Madhlamoto. 2017. "Missing in Action: Species Competition Is a Neglected Predictor Variable in Species Distribution Modelling." *PLoS One* 12, no. 7: e0181088. https://doi.org/10.1371/journal.pone.0181088.

Muscheler, R., F. Adolphi, T. J. Heaton, et al. 2020. "Testing and Improving the IntCal20 Calibration Curve With Independent Records." *Radiocarbon* 62, no. 4: 1079–1094. https://doi.org/10.1017/RDC.2020.54.

Naimi, B., N. A. S. Hamm, T. A. Groen, A. K. Skidmore, and A. G. Toxopeus. 2014. "Where Is Positional Uncertainty a Problem for Species Distribution Modelling?" *Ecography* 37, no. 2: 191–203. https://doi.org/10.1111/j.1600-0587.2013.00205.x.

Ney, G., K. Frederick, and J. Schul. 2018. "A Post-Pleistocene Calibrated Mutation Rate From Insect Museum Specimens." *PLoS Currents* 10: ecurrents.tol.aba557de56be881793261f7e1565cf35. https://pmc.ncbi.nlm.nih.gov/articles/PMC6072552/.

Oksanen, J., G. Simpson, F. Blanchet, et al. 2024. *vegan: Community Ecology Package*. R Package Version 2.6-7. https://github.com/vegandevs/vegan, https://vegandevs.github.io/vegan/.

Ormerod, S. J., M. Dobson, A. G. Hildrew, and C. R. Townsend. 2010. "Multiple Stressors in Freshwater Ecosystems." *Freshwater Biology* 55, no. s1: 1–4. https://doi.org/10.1111/j.1365-2427.2009.02395.x.

Ortego, J., and L. L. Knowles. 2020. "Incorporating Interspecific Interactions Into Phylogeographic Models: A Case Study With Californian Oaks." *Molecular Ecology* 29: 4510–4524. https://doi.org/10.1111/mec.15548.

Papadopoulou, A., I. Anastasiou, and A. P. Vogler. 2010. "Revisiting the Insect Mitochondrial Molecular Clock: The Mid-Aegean Trench Calibration." *Molecular Biology and Evolution* 27, no. 7: 1659–1672. https://doi.org/10.1093/molbev/msq051.

Parreira, M. R., J. C. Nabout, G. Tessarolo, M. de Souza Lima-Ribeiro, and F. B. Teresa. 2019. "Disentangling Uncertainties From Niche Modeling in Freshwater Ecosystems." *Ecological Modelling* 391: 1–8. https://doi.org/10.1016/j.ecolmodel.2018.10.024.

Pohe, S. R. 2019. *Macroecology of New Zealand Ephemeroptera*. Unpublished PhD thesis. University of Canterbury. https://doi.org/10.26021/7728.

R Core Team. 2024. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. https://www.R-project.org/.

Ramírez-Barahona, S., C. González, A. González-Rodríguez, and J. F. Ornelas. 2017. "The Influence of Climate Niche Preferences on the Population Genetic Structure of a Mistletoe Species Complex." *New Phytologist* 214: 1751–1761. https://doi.org/10.1111/nph.14471.

Resh, V. H., A. V. Brown, A. P. Covich, et al. 1988. "The Role of Disturbance in Stream Ecology." *Journal of the North American Benthological Society* 7: 433–455. https://doi.org/10.2307/1467300.

Ridden, J. D., T. R. Hitchings, and T. R. Hitchings. 2023. "The Canterbury Museum Mayfly Collection Data Resource." *Journal of Limnology* 82: 2097. https://doi.org/10.4081/jlimnol.2023.2097.

Rocha, A. V., G. S. Cabanne, A. Aleixo, L. F. Silveira, P. Tubaro, and R. Caparroz. 2020. "Pleistocene Climate Oscillations Associated With Landscape Heterogeneity of the South American Dry Diagonal Explains the Phylogeographic Structure of the Narrow-Billed Woodcreeper (*Lepidocolaptes angustirostris*, Dendrocolaptidae)." *Journal of Avian Biology* 51, no. 9: e02537. https://doi.org/10.1111/jav.02537.

Rozas, J., A. Ferrer-Mata, J. C. Sánchez-DelBarrio, et al. 2017. "DnaSP 6: DNA Sequence Polymorphism Analysis of Large Data Sets." *Molecular Biology and Evolution* 34, no. 12: 3299–3302. https://doi.org/10.1093/molbev/msx248.

Sivyer, L., M. Morgan-Richards, E. Koot, and S. A. Trewick. 2018. "Anthropogenic Cause of Range Shifts and Gene Flow Between Two

Grasshopper Species Revealed by Environmental Modelling, Geometric Morphometrics and Population Genetics." *Insect Conservation and Diversity* 11, no. 5: 415–434. https://doi.org/10.1111/icad.12289.

Slatkin, M. 1993. "Isolation by Distance in Equilibrium and Non-Equilibrium Populations." *Evolution* 47, no. 1: 264–279. https://www.jstor.org/stable/2410134.

Slatkin, M., and R. R. Hudson. 1991. "Pairwise Comparisons of Mitochondrial DNA Sequences in Stable and Exponentially Growing Populations." *Genetics* 129: 555–562. https://pmc.ncbi.nlm.nih.gov/articles/PMC1204643/.

Smith, A. B., W. Godsoe, F. Rodríguez-Sánchez, H.-H. Wang, and D. Warren. 2019. "Niche Estimation Above and Below the Species Level." *Trends in Ecology & Evolution* 34, no. 3: 260–273. https://doi.org/10.1016/j.tree.2018.10.012.

Tajima, F. 1989. "The Effect of Change in Population Size on DNA Polymorphism." *Genetics* 105: 437–460. https://pubmed.ncbi.nlm.nih.gov/2599369/.

Taubmann, J., K. Theissinger, K. A. Feldheim, et al. 2011. "Modelling Range Shifts and Assessing Genetic Diversity Distribution of the Montane Aquatic Mayfly *Ameletus inopinatus* in Europe Under Climate Change Scenarios." *Conservation Genetics* 12, no. 2: 503–515. https://doi.org/10.1007/s10592-010-0157-x.

Thuiller, W. 2003. "BIOMOD—Optimizing Predictions of Species Distributions and Projecting Potential Future Shifts Under Global Change." *Global Change Biology* 9, no. 10: 1353–1362. https://doi.org/10.1046/j.1365-2486.2003.00666.x.

Tillyard, R. J. 1923. "Descriptions of Two New Species of May-Flies (Order Plectoptera) From New Zealand." *Transactions and Proceedings of the New Zealand Institute* 54: 226–230.

Trewick, S. A., and K. J. Bland. 2012. "Fire and Slice: Palaeogeography for Biogeography at New Zealand's North Island/South Island Juncture." *Journal of the Royal Society of New Zealand* 42: 153–183. https://doi.org/10.1080/03036758.2010.549493.

Trewick, S. A., I. M. Henderson, S. R. Pohe, and M. Morgan-Richards. 2022. "Spatial Variation of *Acanthophlebia cruentata* (Ephemeroptera), a Mayfly Endemic to Te Ika-a-Māui—North Island of Aotearoa, New Zealand." *Insects* 13, no. 7: 567. https://doi.org/10.3390/insects13070567.

Trewick, S. A., S. Pilkington, L. D. Shepherd, G. C. Gibb, and M. Morgan-Richards. 2017. "Closing the Gap: Avian Lineage Splits at a Young, Narrow Seaway Imply a Protracted History of Mixed Population Response." *Molecular Ecology* 36, no. 20: 5752–5772. https://doi.org/10.1111/mec.14323.

Valavi, R., J. Elith, J. J. Lahoz-Monfort, and G. Guillera-Arroita. 2019. "blockCV: An r Package for Generating Spatially or Environmentally Separated Folds for k-Fold Cross-Validation of Species Distribution Models." *Methods in Ecology and Evolution* 10, no. 2: 225–232. https://doi.org/10.1111/2041-210X.13107.

Valencia-Rodríguez, D., L. Jiménez-Segura, C. A. Rogéliz, and J. L. Parra. 2021. "Ecological Niche Modeling as an Effective Tool to Predict the Distribution of Freshwater Organisms: The Case of the Sabaleta *Brycon henni* (Eigenmann, 1913)." *PLoS One* 16, no. 3: e0247876. https://doi.org/10.1371/journal.pone.0247876.

Varela, S., J. M. Lobo, and J. Hortal. 2011. "Using Species Distribution Models in Paleobiogeography: A Matter of Data, Predictors and Concepts." *Palaeogeography, Palaeoclimatology, Palaeoecology* 310, no. 3–4: 451–463. https://doi.org/10.1016/j.palaeo.2011.07.021.

Vignali, S., A. G. Barras, R. Arlettaz, and V. Braunisch. 2020. "SDMtune: An R Package to Tune and Evaluate Species Distribution Models." *Wiley Online Library* 10, no. 20: 11488–11506. https://doi.org/10.1002/ece3.6786.

Warren, D. L., N. J. Matzke, and T. L. Iglesias. 2020. "Evaluating Presence-Only Species Distribution Models With Discrimination Accuracy Is Uninformative for Many Applications." *Journal of Biogeography* 47, no. 1: 167–180. https://doi.org/10.1111/jbi.13705.

Wiens, J. J., D. D. Ackerly, A. P. Allen, et al. 2010. "Niche Conservatism as an Emerging Principle in Ecology and Conservation Biology." *Ecology Letters* 13, no. 10: 1310–1324. https://doi.org/10.1111/j.1461-0248.2010.01515.x.

Wisz, M. S., J. Pottier, W. D. Kissling, et al. 2013. "The Role of Biotic Interactions in Shaping Distributions and Realised Assemblages of Species: Implications for Species Distribution Modelling." *Biological Reviews* 88, no. 1: 15–30. https://doi.org/10.1111/j.1469-185X.2012.00235.x.

Ye, Z., G. Zhu, P. Chen, D. Zhang, and W. Bu. 2014. "Molecular Data and Ecological Niche Modelling Reveal the Pleistocene History of a Semi-Aquatic Bug (*Microvelia douglasi douglasi*) in East Asia." *Molecular Ecology* 23, no. 12: 3080–3096. https://doi.org/10.1111/mec.12797.

Zurell, D., J. Franklin, C. König, et al. 2020. "A Standard Protocol for Reporting Species Distribution Models." *Ecography* 43, no. 9: 1261–1277. https://doi.org/10.1111/ecog.04960.

Supporting Information

Additional supporting information can be found online in the Supporting Information section. **Appendix S1:** Geometric morphometrics of forewing morphology methods, results and references. **Table S1:** The seven climatic variables retained for species distribution modelling of two Aotearoa–New Zealand mayflies. **Table S2:** Evaluation scores for ensemble species distribution models projected onto current climate for two New Zealand mayflies: Sensitivity is the percentage of presences correctly predicted, specificity is the percentage of absences correctly predicted. Relative operating characteristics (ROC) and true skill statistic (TSS) scores assess how well models fit the data, and how accurately the models predict a presence or absence correctly. **Table S3:** Sampling of *Ichthybotus* mayflies for species distribution models, forewing morphology (Wing) and genetics (mtDNA). LGM = hindcast projection of SDM inferred that the location was climatically suitable for continuous occupation by current *Ichthybotus* species. mtDNA summary statistics, n = sample size, h = number of haplotypes, Hd = haplotype diversity, π = nucleotide diversity. **Table S4:** The influence of each climate variable in the species distribution models for Aotearoa–New Zealand *Ichthybotus* mayflies recorded with predictor importance score. **Table S5:** New Zealand *Ichthybotus* mayfly samples used for geometric morphometric analysis of forewings. **Table S6:** Cross-validation results for the canonical variates analysis of forewing shape, testing three separate comparisons (island origin, life stage and sex). **Figure S1:** Sampling and species distribution model projections of *Ichthybotus hudsoni* and *I. bicolor* on current climate conditions in Aotearoa-New Zealand. Presence/absence (left), model projection (right). Projections show potential niche availability (≥ cutoff value maximising Sensitivity and Specificity) in green. Dotted line indicates ash deposit from Oruanui volcanic eruption (25.5 ka; Muscheler et al. 2020). **Figure S2:** Sampling and species distribution model projections of *Ichthybotus hudsoni* and *I. bicolor* on current climate conditions in Aotearoa-New Zealand using subset of absence data. Presence/absence showed on the left, model projection on right. Projections show potential niche availability (≥ cutoff value maximising Sensitivity and Specificity) in green. Lower maps show species distribution modelling projections of the mayflies *Ichthybotus hudsoni* (right) and *I. bicolor* (left) on Aotearoa-New Zealand last glacial maximum climatic conditions. Projections show potential species distributions (≥ cutoff value maximising Sensitivity and Specificity) in green colour, purple corresponds to areas with niche availability above 70% of the cutoff value. Dotted line indicates Oruanui volcanic eruption (25.5 ka) ash deposit > 1 m deep (Muscheler et al. 2020). **Figure S3:** Dorsal view of left *Ichthybotus* mayfly forewing, showing the placement of the virtual comb and landmarks used for the geometric morphometric analysis (fixed landmarks; 1–6; semi-landmarks: 7–30). **Figure S4:** The selection of retained principal components for the principal components analysis of forewing shape of New Zealand mayflies using Eigen values. **Figure S5:** Principal components analysis of *Ichthybotus* mayfly

forewing shape; (a) principal components 1 and 2; (b) principal components 3 and 4. Coloured ovals are 95% confidence ellipses of sample means. **Figure S6:** Results for the principal components analysis of forewing shape, showing principal components 1 and 2 with the classification of different groups: (a) the two clusters identified by *mclust* under the optimal VII2 model using principal components 1–6 and centroid size; (b) life stage (imago or subimago); and (c) sex (male or female). **Figure S7:** Thin plate spline wireframe figures visualising shape variation at each landmark associated with each principal component estimated by the principal components analysis of mayfly forewing shape. **Figure S8:** Canonical variates analyses plots for the morphometric analysis of mayfly forewing shape, with samples classified according to (a) geographic origin, (b) sex and (c) life stage (with overlap shown in intermediate colours). **Figure S9:** Plots of Bayesian information criterion scores for the fit of various Gaussian mixture clustering models in *mclust* to morphometric variation among mayfly forewing specimens; (a) scores using only shape variation (principal components 1–6); (b) scores using wing shape and size variation (principal components 1–6 and centroid size). **Figure S10:** Discrimination of the two mayfly wing-shape and size clusters identified under the VII2 model using *mclust* ordination plots, based on principal components 1–6 and centroid size scaled together (Cluster 1: purple; Cluster 2: yellow).