



Population decline at distribution margins: Assessing extinction risk in the last glacial relictual but still functional metapopulation of a European butterfly

Stéphanie Sherpa¹  | Caroline Kebaïli^{1,2}  | Delphine Rioux¹ | Maya Guéguen¹ | Julien Renaud¹ | Laurence Després¹ 

¹Laboratoire d'Ecologie Alpine, UMR CNRS-UGA-USMB 5553, Université Grenoble Alpes, Grenoble Cedex 9, France

²Parc Naturel Régional du Haut Jura, Lajoux, France

Correspondence

Stéphanie Sherpa, Laboratoire d'Ecologie Alpine UMR CNRS-UGA-USMB 5553, Université Grenoble Alpes, CS 40700, 38058 Grenoble Cedex 9, France.
Email: stephanie.sherpa@hotmail.fr

Funding information

Conservatoire botanique national de Franche-Comté - Observatoire régional des Invertébrés (CBNFC-ORI)

Editor: Alan Andersen

Abstract

Aim: To determine the interplay between climate and land use changes in driving population dynamics in a butterfly species, *Coenonympha hero*, at the southern limit of its distribution.

Location: French Jura massif and Europe.

Methods: We analysed patterns of genetic diversity distribution at 817 loci in 136 butterflies from 31 sites using NGS to infer the genetic structure and population size changes over time, using two methods of demographic inference (SNP frequency spectrum analyses and coalescent ABC inferences). We then characterized the climate and land use descriptors of *C. hero* geographic distribution using species distribution modelling (SDM) and ordination method and compared demographic changes to changes in climatically suitable areas.

Results: *Coenonympha hero* persists in the Jura massif as three core populations that share a common history of decline in two steps: an old decline and a more recent decline that resulted in population fragmentation, the southernmost sites being the most threatened (lowest genetic diversity). Climate change during the Würm glaciation is presumably the main factor explaining the first demographic decline. The second decline started 2000 years ago possibly under increased human pressure as suggested by recent extinctions in several sites nowadays characterized by urban and agricultural surfaces. Both climate and land use variables are important descriptors of *C. hero* distribution, as SDM predictions were improved by adding aridity index, altitude and land use to bioclimatic predictors. Its habitats include forests in north-eastern Europe and grasslands in the Jura massif.

Main conclusions: Using SDM and genetic demographic inferences, we identified a persistent glacial refuge for the species in Europe. We show that although this relictual population has declined and fragmented under the combined effects of climate warming and anthropization, the metapopulation is still functional but requires

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *Diversity and Distributions* published by John Wiley & Sons Ltd.

particular conservation attention to maintain its connectivity, and to favour the local persistence of this highly endangered butterfly species.

KEYWORDS

Anthropization, climate change, *Coenonympha hero*, demographic history, genetic diversity, habitat fragmentation, Jura massif, metapopulation, Next-generation-sequencing, relict populations, species distribution models

1 | INTRODUCTION

Many species are declining in Europe as the result of climate warming and land use changes. The conversion of natural habitats by human activities is the threatening factor influencing the largest number of species, while climate change is expected to have a growing role in the next decades. Climate change affects the distribution and abundance of species, with increases in abundance in the cooler parts of species' ranges and declines in warmer parts (Fox et al., 2014; Lehtikoinen et al., 2013; Parmesan et al., 1999). Changing temperatures can have a particularly marked effect on butterflies, with species at the edge of their distribution showing the most dramatic range shifts (Menéndez et al., 2006; Parmesan et al., 1999; Warren et al., 2001).

The conjunction of climate warming and land use changes is likely to particularly affect cold-adapted boreal species at the southern margin of their distribution and habitat-specialist species that occupy semi-open habitats such as forest glades. Populations in these temporary habitat patches are often small and geographically isolated from each other and thus may strongly suffer under demographic and stochastic environmental fluctuations, making them particularly prone to local extinctions (Melbourne & Hastings, 2008). Furthermore, such isolated populations may express inbreeding depression causing reduced fitness and lower adaptive potential to environmental changes (Habel & Schmitt, 2018; Nonaka et al., 2019). These species persist over the landscape through a dynamic process of extinction/recolonization of local habitat patches with a metapopulation functioning (Hanski & Thomas, 1994). Dispersal is a central process in metapopulation dynamics, yet the effective connectivity between favourable patches is not easy to assess using direct methods such as mark–release–recapture that do not account for (1) rare long-distance dispersal events and (2) effective dispersal, that is the actual success in establishment after dispersion. Genetic methods can give better estimates of effective dispersal than direct methods with less effort. The large number of single nucleotide polymorphisms (SNPs) available from non-lethal sampling together with the development of computationally efficient inference methods such as approximate Bayesian computation (ABC) provides unprecedented amount of information about past and present population effective size, population fragmentation history and contemporary effective gene flow, even for endangered species persisting in small and fragmented populations at their range limits.

Although the recent decade has seen an explosion of species distribution modelling (SDM) based on climate variables to predict

future species ranges (reviewed in Zurell et al., 2020), there is increased concern that these global climate predictors might not capture the relevant habitat conditions encountered by individuals, because both topography and land cover can strongly influence local microclimate and population persistence, especially at range margins (Suggitt et al., 2011). For example, there is increasing evidence that climate warming is promoting altitudinal range shifts, which could drive many populations to the extinction in the future (Rödger et al., 2021), and that habitat fragmentation may present a major impediment to species poleward range shifts (Fourcade et al., 2021). Demographic genetic inferences can help improving the performance of SDM to capture past and current habitat suitability of species at their distribution margins, by choosing the environmental predictors best accounting for the population size fluctuations over time and space at fine-spatial scale. Here, we show the complementarity between demographic genetic inferences and species distribution models based on multiple predictors to understand past and current population functioning at species distribution margin, key information for conservation management.

The Scarce Heath *Coenonympha hero* (Linnaeus, 1761) (Lepidoptera, Nymphalidae) is a butterfly species occurring very locally from eastern France across central Europe and southern Scandinavia to temperate Asia. It is among the most threatened butterfly in Europe and has already become extinct in the Czech Republic, Denmark, Belgium, Luxembourg, and the Netherlands (Kudrna et al., 2011; Sielezniew & Nowicki, 2017; Van Swaay & Warren, 1999). France represents the south-western margin of the whole distribution of the species. The species was historically distributed locally throughout north-eastern France but has disappeared from Paris region in the 60s, from Massif Central in the 80s, and is nowadays only present in small, highly fragmented populations in the Jura massif in wet meadows, forest edges and glades up to 800m. The causes for its decline in Europe could be habitat lost through afforestation and/or urbanization and arable land expansion, but also climate change. Hybridization with the Pearly Heath *C. arcania* has also been mentioned as a possible threat to the species (Lafranchis, 2000).

Here, we investigated the genetic structure and diversity of *C. hero* populations in the Jura massif aiming at determining whether these remnant edge populations show evidence for genetic erosion and inbreeding depression, as found in Swedish isolated small populations of the species (Cassel & Tammaru, 2003; Cassel et al., 2001). Previous mark–release–recapture (MRR) studies in Sweden have shown a functioning in metapopulation, with migration between

suitable habitat patches (semi-open areas) maintaining a dynamical non-equilibrium population despite temporary local extinctions (Cassel-Lundhagen & Sjögren-Gulve, 2007), with large populations acting as sources for small satellite populations (Cassel-Lundhagen et al., 2008). Based on current pattern of genetic variation, we inferred the demographic history of the French population, using ABC to test for alternative demographic scenarios, involving ancient (e.g. several thousand years) or recent (historical) population decline and fragmentation. Based on European occurrences available for *C. hero*, we modelled the potential current suitable climatic conditions for the species and projected its putative distribution into the Mid-Holocene (MID, 6 kya), Last Glacial Maximum (LGM, 20 kya) and Last Interglacial (LIG, 120–140 kya) in order to link the demographic history of the species with broad-scale climatic changes. Because local conditions are likely to be important for this low mobile species, we tested whether adding more local environmental factors (global aridity index, altitude, land cover) improved model performance. We then focused on France to construct a habitat model taking into account occurrence status (still present, no-longer present = extinction) and their dates of report in order to evaluate which landscape features are the most important for the species persistence and to guide land use management decisions in a conservation perspective.

2 | METHODS

2.1 | Ecological analyses

2.1.1 | Species distribution models

In order to determine the best environmental predictors of *C. hero* current distribution in France and Europe and identify changes in climatically suitable areas over time in Europe, we constructed a series of SDMs at global (Europe) and regional (France) scales (Table 1), based on geo-referenced occurrences available for *C. hero* in the Global Biodiversity Information Facility database (GBIF, <https://www.gbif.org/>). We retained field observations recorded after 1900 and records from distribution atlas (unknown date; Kudrna et al., 2011) all with coordinate precision <5 km (Appendix S1) and added 31 new occurrences for the present study samples, resulting in 4615 occurrences in Europe (Table S1). Sampling biases due to unequal sampling effort (e.g. monitoring projects, sampling interests, accessibility) can affect the accuracy of SDM predictions by over-representing the ecological importance of the more sampled region (Fourcade et al., 2014). To reduce this possible bias, we randomly selected one occurrence when several points fell within the same raster cell of 1 km² (Appendix S1), resulting in 599 occurrences distributed in 12 countries (Figure 1; Table S1).

Depending on the model, environmental predictors included climatic, topographic and land use variables. We collected or calculated a total of 31 environmental variables (Table S2.1: Appendix S2) and retained non-collinear variables (Figures S2.1 and S2.2). We first constructed a climate-only model (SDM1) using six BIOCLIM

variables relevant with respect to the ecology of the species. Indeed, butterflies are highly sensitive to temperature fluctuations, and the main period of activity for *C. hero* (larval development and imago reproduction) occurs during the warmest months. We thus selected the following: mean annual temperature (BIO1), mean diurnal range (BIO2), isothermality (BIO3), mean temperature of warmest quarter (BIO8), annual precipitation (BIO12) and precipitation seasonality (BIO15) from WorldClim v2 (Fick & Hijmans, 2017). To determine current and past climatically suitable areas for *C. hero* in Europe and assess species distribution changes over time, we projected this model in the current (1970–2000), Mid-Holocene (MID), Last Glacial Maximum (LGM) and Last Inter-Glacial (LIG) periods. To determine whether a more detailed characterization of the physical environment improved model performance, we removed variables of low importance (BIO12) and added the Global Aridity Index (GAI) from the CGIAR-CSI v2 database (Trabucco & Zomer, 2019) and elevation (DEM) (European Environment Agency) to construct SDM2. Two other climate-related variables, the growing degree days (GDD) and soil moisture (SMC), were not included because they were strongly correlated to BIO1 and GAI, respectively (Appendix S2).

We then constructed a habitat-only model using the CORINE Land Cover (CLC) map of Europe level 3 (European Environment Agency) to characterize current land use (CLC2018). We reclassified the original 44 classes into 9 levels (Table S2.2), including artificial (urban) and natural (crops, bocage, grass, shrub, forest, water) surfaces, and calculated the percentage of each natural habitat type per km². To represent anthropization, we used the human footprint (HF) index (Venter et al., 2018) from the NASA Socioeconomic Data and Applications Center (SEDAC). We also included in this model the net primary productivity (NPP) as a proxy for global resource availability (Imhoff et al., 2004). Finally, we combined climatic, topographic and habitat predictors retained in SDM2 and SDM3 to construct SDM4 (Table 1; Appendix S2). For these four models, we used 599 presences and three datasets of 10,000 random pseudo-absences were selected using a surface range envelope model.

Among the 599 presences, some are historical records, where the species has not been reported since 1980 (Table S1). Since land use has dramatically changed in Europe during the last decades, considering extinct localities as presence data can potentially skew distribution models. While we have access to detailed information about the status (year of last record) for occurrences in France, distinguishing the areas where the species is extinct and the areas not surveyed since the 1980s for all European countries is difficult (Figure S2.3). We thus focused on France to build a fifth model (SDM5) based on 135 occurrences: 41 presences, 44 known extinctions (true absences) and 50 absences randomly sampled in areas where the species is absent without occurrences records (Figure 1; Appendix S2) to refine potentially suitable habitats for the species in France. For this model, we used the same climatic and topographic variables as SDM4 but the local habitat was characterized at the time of occurrence report using the last available CLC layer (CLC1992, CLC2000, CLC2006, CLC2012 and CLC2018) (Table S1; Appendix S2). We evaluated the possible impact of mismatch between climate

TABLE 1 Description of the species distribution models (SDM) performed

| Aim | Model | Scale | Occurrence dataset | Environmental data type | Environmental variables used | Projection |
|---|-------|--------|------------------------|------------------------------|--|-------------------------------|
| Determine past changes in geographical distribution | SDM1 | Europe | 599 P 10,000 PA | Climate | BIO1, BIO2, BIO3, BIO8, BIO12, BIO15 | Current, Past (MID, LGM, LIG) |
| Determine whether a more detailed description of the climate and topology improves model performance (compared to SDM1) | SDM2 | Europe | 599 P 10,000 PA | Climate, Topography | BIO1, BIO2, BIO3, BIO8, BIO15, GAI, DEM | Current |
| Characterize the best land use descriptors of geographic distribution | SDM3 | Europe | 599 P 10,000 PA | Habitat | HF, NPP, %crops, %bocage, %forest, %grass, %shrub, %water | Current |
| Determine whether using multiple environmental predictors improves model performance (compared to SDM1, SDM2, and SDM3) | SDM4 | Europe | 599 P 10,000 PA | Climate, Topography, Habitat | BIO1, BIO2, BIO3, BIO8, BIO15, GAI, DEM, HF, NPP, %crops, %bocage, %forest, %grass, %shrub, %water | Current |
| Better characterize suitable areas based on local habitat in France using known extinctions (compared to SDM4) | SDM5 | France | 41 P 44 TA 50 SA | Climate, Topography, Habitat | BIO1, BIO2, BIO3, BIO8, BIO15, GAI, DEM, HF, NPP, %crops, %bocage, %forest, %grass, %shrub, %water | Current |

Note: For occurrence and environmental data selection, see Appendices S1 and S2. Occurrence status: P, presences; PA, pseudo-absences; TA, true absences based on known local extinctions; SA, randomly sampled absences.

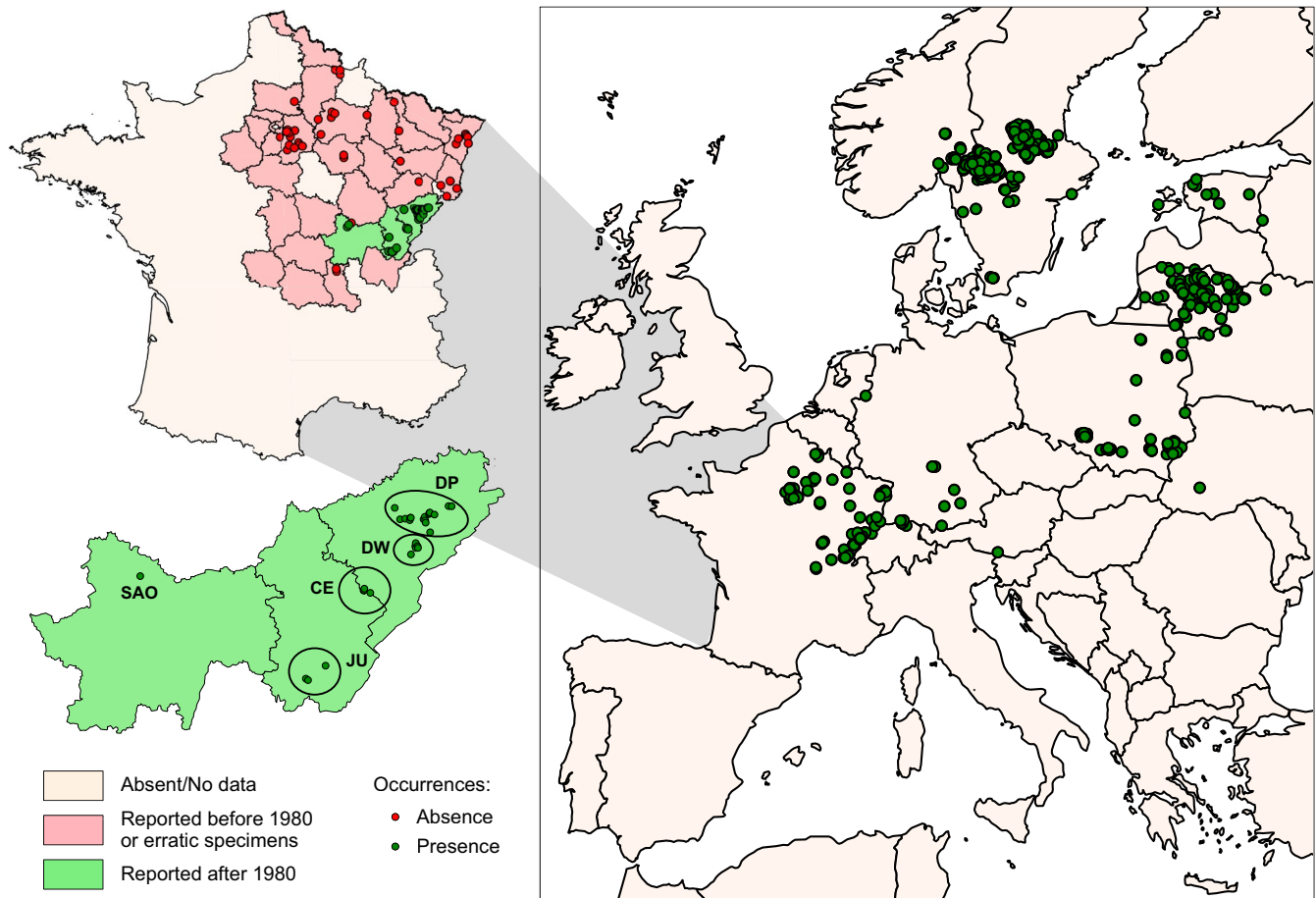


FIGURE 1 Distribution map of *Coenonympha hero* in Europe. Right: location of the 599 presences in Europe used for SDM1 to SDM4. Top left: location of the 85 occurrences in France used for SDM5. French departments are coloured according status of *C. hero* (present/extinct/absent). Bottom left: location of the 31 localities sampled in 2019 in the French departments of Doubs Jura, and Saône-et-Loire for genetic analyses

data (1970–2000) and species occurrence data (records between 1900 and 1970) on the mapped potential distributions (Appendix S2). SDM results including or not these 1900–1970 records are not significantly different so we included them in SDM1 to SDM4 models because excluding them would lead to underestimate the width of the species climatic niche and erroneous past (LIG, LGM, Mid-Holocene) distribution projections.

For each analysis, an ensemble of five statistical models was used, including GLM, GAM, GBM, MAXENT and RF (Araújo & New, 2007; Marmion et al., 2009; Appendix S2). Models were calibrated using 70% of observations randomly sampled from the initial data and evaluated against the remaining 30% data using the true skill statistic (TSS, Allouche et al., 2006) and the area under the curve (ROC, Swets, 1988). We performed a threefold internal cross-validation, resulting in a total of 45 models per analysis, that were merged using the weighted sum of probabilities (EMwmeanByTSS) or the committee averaging (EMcaByTSS) methods. For all analyses (SDM1 to SDM5), TSS and ROC statistics indicate that EMwmeanByTSS performs better on average and was thus used for projected distributions. The importance of each variable was assessed by calculating

Pearson's correlation between the fitted values and the predictions after randomly permuting the values of the variable (Thuiller et al., 2009). We normalized variable importance per model to 100 in order to compare values among models. Models and the ensemble forecasting procedure were performed using the biomod2 R package (Thuiller et al., 2016). Model parameters are provided in a Supplementary file.

2.1.2 | Ordination-based method

In addition to SDM, we performed principal component analyses (PCA) using the ade4 R package (Chessel et al., 2004). First, we compared current and past (MID, LGM and LIG) climate in Europe using BIOCLIM variables (data of SDM1). Second, we evaluated whether the still-occupied areas in France (departments of Doubs, Jura and Saône-et-Loire) are favourable or are situated at the margin of the species niche, as they represent the south-western distribution of *C. hero* (Figure 1). We used climatic and topographic data (data of SDM2) and land use variables (data of SDM3).

2.2 | Genetic analyses

2.2.1 | DNA samples

According to occurrence records, the southernmost margin of *C. hero* current distribution in Europe includes three French departments (Kudrna et al., 2011; Figure 1). *Coenonympha hero* is mainly found in the Doubs department, in the Plateau du Doubs (DP) and the nearby wetland (DW), and in Jura department (JU). Some populations have recently been reported at the limit between Doubs and Jura in a centre position (CE). The species was thought to be extinct in the Saône-et-Loire (SAO) department but was recently rediscovered. We collected 149 specimens from 31 localities in the five cited areas with three to six individuals per locality (Figure 1; Table S2). Individuals were captured and immediately released after one leg removal (non-lethal sampling). The sampling was performed in May–June 2018 by qualified persons from the Conservatoire Botanique National de Franche Comte—Observatoire Régional des Invertébrés (CBNFC-ORI) and the Conservatoire d'Espaces Naturels de Franche-Comté (CENFC) (Table S2). Legs were kept in ethanol 75° at –20°C till DNA extraction.

2.2.2 | Genomic data acquisition

DNA extraction and ddRADseq library preparation—DNA was extracted using the DNeasy Blood and Tissue Kit (Qiagen) according to manufacturer's instructions. For each individual, 50 ng of genomic DNA was digested in a 36- μ L reaction, using 30 units each of *Sbf*I-HF and *Msp*I restriction enzymes (New England Biolabs) and NEB CutSmart buffer. Digestions (2.5 h at 37°C) were ligated to Illumina P1 (individually indexed) and P2 adapters by 60 cycles of digestion at 37°C (2 min) and ligation at 16°C (4 min) with 10 units of T4 ligase (New England Biolabs), followed by a 10-min heat-inactivation step at 65°C. Digested-ligated products were pooled and purified with 1.5:1 ratio Ampure XP beads (Beckman Coulter), and 200–600 bp fragments were selected using a PIPPIN™-Prep 2% gel cassette (Sage Sciences). Each library was amplified in ten independent replicates of 15 PCR cycles (initial denaturation 30 s, 98°C; 15 cycles 98°C 10 s; 66°C 30 s and 72°C 1 min; final extension 10 min 72°C) in a volume of 20 μ l with 2 μ l of DNA template, 0.2 mM dNTPs, 0.25 μ M of each PCR primer, 3% DMSO and 0.4 U *Taq* Phusion-HF (New England Biolabs). The ten PCR products were pooled and purified with a QIAgen PCR Purification Kit (Qiagen), and the ddRAD library was sequenced on a single lane Hi-Seq 2500 using 125bp paired-end chemistry (Fasteris SA).

SNP genotyping—Among a total of 82 million reads obtained, we trimmed adapters using the BBduk tool of the BBmap package (Bushnell, 2014), removed reads with length <120 nucleotides and cut all reads to this value. We retained 91% high-quality (HQ) reads after removing reads of low quality or with uncalled bases (Table

S3). We used the *denovo_map.pl* pipeline in STACKS v2 (Rochette et al., 2019) to reconstruct loci and call genotypes, using a maximum of 7 mismatches to merge two stacks into a polymorphic locus (-M; *ustacks* function). Highly repetitive stacks were dropped using the "Removal" (-r) option. A catalog of the loci from all the individuals was built, with a maximum of 10 mismatches for merging two individual loci (-n; *cstacks* function). Loci within each individual were searched against the catalog (*sstacks* function). A SNP data set was produced with the genotype of each individual for every polymorphic position (*populations* function). We removed called genotypes from per-sample read depth <5. We required a locus to have <20% missing data, and individuals with >25% missing data were discarded ($N = 13$, Table S3), resulting in 136 individuals with 550,000 HQ reads per individual on average. The final sample sizes after data filtering for the five regions are as follows: $N_{DP} = 75$, $N_{DW} = 27$, $N_{JU} = 14$, $N_{CE} = 17$ and $N_{SAO} = 3$. Among 817 RAD loci (2373 SNPs), the median percentage of missing data per locus was 4% and the median coverage per locus was 85X (Table S3). Among the 2373 SNPs, we removed SNPs with minor allele frequency <1% (Linck & Battey, 2019) and retained only one random SNP per locus, resulting in 817 SNPs.

2.2.3 | Population structure and diversity

Population genetic structure and diversity was investigated using the 817 SNP data set. We first performed a PCA in the *pcadapt* R package (Luu et al., 2017). We then used ADMIXTURE (Alexander et al., 2009) for inferring admixture coefficients for all individuals. The optimal number of genetic clusters, K (range 1–10), was determined based on a 10-fold cross-validation. Genetic diversity was estimated using observed heterozygosity (H_o), population genetic diversity (=expected heterozygosity, H_e), allelic richness (A_R), inbreeding coefficient (F_{IS}) and population-specific F_{ST} using the *hierfstat* R package (Goudet, 2005). We also calculated pairwise relatedness coefficients among individuals (Wang, 2002) using the *related* R package (Pew et al., 2017) (Figures S3.6 and S3.7: Appendix S3).

We tested for isolation by distance (IBD) at two spatial scales: among all the populations, and within each region, by testing for correlations between normalized F_{ST} ($F_{ST}/[1-F_{ST}]$) and geographic distances. Pairwise F_{ST} were computed using the *hierfstat* R package, and IBD was tested using Mantel tests with 9999 permutation tests of significance in the *ecodist* R package (Goslee & Urban, 2007). Spatial patterns of genetic variation among populations were further investigated by computing genetic dissimilarities on a spatial grid of 500 demes among the three French departments using the EEMS programme (Petkova et al., 2016) (Figure S3.4). We ran three independent analyses with a burn-in of 2,000,000 and MCMC length of 10,000,000 with a thinning interval of 10,000. The results of the three analyses were combined to produce maps for within-demes (diversity rates) and between-demes (migration rates) estimates using the *rEEMSplots* R package (Petkova et al., 2016) (Figure S3.5).

2.2.4 | Demographic inferences

To reconstruct the demographic history of *C. hero*, we used two methods: the stairway plot method v2 based on the site frequency spectrum (SFS) (Liu & Fu, 2015) that reconstructs continuous changes in effective population size over time and the approximate Bayesian computation approach (ABC) implemented in DIYABC v2 (Cornuet et al., 2014) that allows to compare alternative scenarios and to estimate past and current population sizes and discrete times for size changes. Based on population genetic structure results, we defined three populations corresponding to three geographical regions (Appendix S4).

For the stairway plot method, the SFS of each population (DPW, JU, CE) was computed from imputed genotypes of full SNP data sets (2373 SNPs). Missing data were imputed via matrix completion (Chi et al., 2013) from allele frequencies and ancestry coefficients considering three genetic groups. A total sequence length of 254 kb was used, including monomorphic and polymorphic sites among polymorphic loci, as well as monomorphic loci (817 polymorphic loci among 2118 loci) with a mutation rate of 2×10^{-9} (Appendix S4). Effective population sizes through time for each population were estimated using the median and 95% confidence intervals of 100 bootstrap replicates.

For the ABC method, we used the 817 SNP data set. We first compared four alternative scenarios of population demographic history: constant population size (Scenario 1), recent decline (Scenario 2), old decline (Scenario 3) and two declines (Scenario 4) (Figure S4.3). After identifying the best scenario for each population separately, we determined the divergence time between the three populations by comparing three alternative scenarios: divergence before old decline (Scenario 1), divergence due to old decline (Scenario 2) or divergence after recent decline (Scenario 3) (Figure S4.5). All priors were set to uniform distribution. Prior intervals were defined based on the stairway plot results for ABC reconstruction of single-population demographic history, and using the prior distribution of parameters from ABC reconstruction of single-population demographic history for divergence scenarios (Appendix S4). For summary statistics, we used the within-population genetic diversity (H_d) for single-population demographic history and added the genetic differentiation (F_{ST}) between populations for the divergence scenarios.

We generated 1,000,000 data sets for each scenario and compared alternative scenarios by calculating their posterior probabilities using a logistic regression on the 1% simulated data closest to the observed data (Cornuet et al., 2010). The goodness of fit of the best scenario was assessed using a PCA, by checking the accordance between observed and simulated data sets from prior distributions. We evaluated confidence in scenario choice by generating 100 newly simulated data sets from priors to compute type I and type II errors (Cornuet et al., 2010). Once the most likely scenario was identified, the posterior distributions of parameters were estimated by computing a local linear regression on the 1% of the simulated data closest to observed data set, after applying a logit transformation to the parameters value (Cornuet et al., 2010).

3 | RESULTS

3.1 | Current and past distribution in Western Europe

The projected potential distribution in Europe of *C. hero* based on current BIOCLIM data (SDM1) predicts a high suitability around the Baltic Sea (between latitudes 52°N and 60°N) with continuous highly suitable areas (presence probability >0.7) in Sweden, Lithuania, Latvia and Estonia (Figure 2a). In contrast, the probability of the presence of *C. hero* between latitudes 45°N and 52°N is lower, with discontinued areas of intermediate suitability (presence probability between 0.4 and 0.6) especially in France, Belgium, Germany and the Carpathians. The climatic variables with the highest importance were BIO1 (mean annual temperature) and BIO3 (isothermality), explaining respectively 45% and 25% of *C. hero* distribution in Europe (Figure 2b).

The principal component analysis investigating climatic variation between the four periods compared (CUR, MID, LGM, LIG) revealed that climate at CUR, MID and LIG periods is similar with higher mean annual temperatures and precipitations than during the LGM (Figure 2c). The projected past distribution in Europe revealed several areas with high probability of presence (>0.8) at each of the period considered (MID, LGM, LIG) (Figure 2d). The potential distribution of *C. hero* during LIG is situated in Western Europe and includes Sweden, France, England, Belgium, the Netherlands and Spain. The projected distribution during LGM reveals two potential refugial areas: one in France (including the Paris basin and the Jura massif) and one in Northern Balkans. The projected distribution during MID is similar to the current geographic distribution of *C. hero* except that Southern edges (France, Balkans) are less favourable under current climate while Northern edges (Sweden, Estonia) are more favourable. The Jura massif constitutes one of the very few climatically suitable areas in Western Europe whatever the period.

In order to better characterize the current distribution of *C. hero* in Europe, we compared a series of models including only climatic (SDM1), climatic and topographic (SDM2), only habitat (SDM3) and all environmental variables (SDM4). Both TSS and ROC statistics reveal that a complex model combining different types of data (SDM4) performs better than simple models (SDM1 and SDM3) (Figure 3a). The three models that include climatic predictors are congruent to identify BIO1 as the most important descriptor of *C. hero* geographic distribution (between 32% and 45%) (Figure 3b). Comparing SDM1 and SDM2 show that adding GAI and DEM as predictors improves model performance and better predict the absence of the species in the Southernmost areas of Europe such as Italy and the Balkans (rainfall deficit) and in mountainous areas such as around the Alpine arc and Pyrenees (elevation) (Figure 3c). These two variables explain 28% of the geographic distribution of *C. hero* (Figure 3b). The habitat-only model (SDM3) is the less efficient model and identifies forest as an important descriptor of *C. hero* distribution (51%) (Figure 3b). The second descriptor is NPP (33%) but it is highly correlated with BIO1 (Figure S2.2), which explains why the importance of NPP in the last

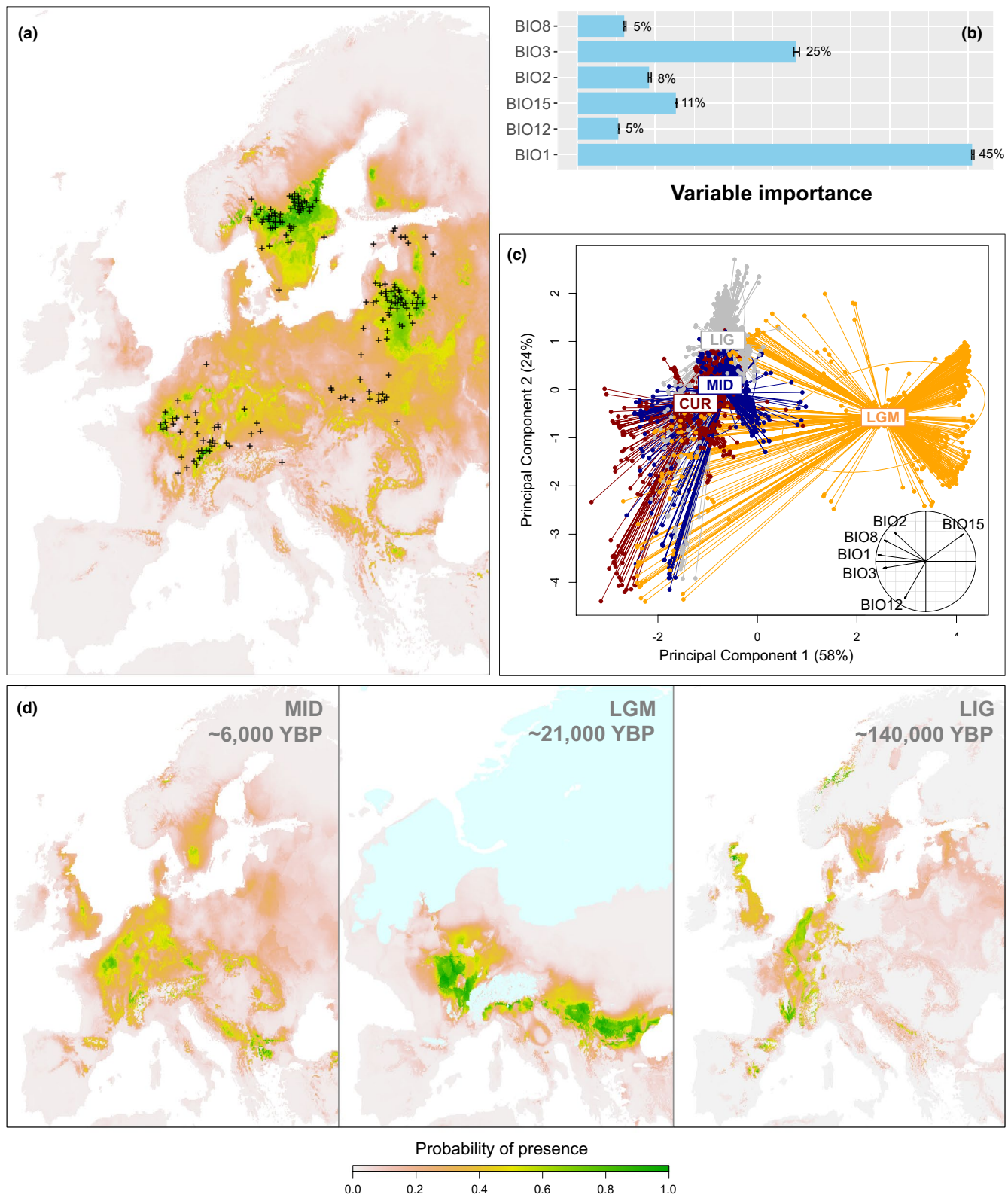


FIGURE 2 Past potential suitable areas for *Coenonympha hero* in Europe. (a) Probability of the presence of *C. hero* under current climate (SDM1) using 599 presences in Europe. (b) Variable importance of each of the six BIOCLIM variables. (c) Principal component analysis showing 79% of climatic variation between the four periods based on the six BIOCLIM variables. (d) Probability of the presence of *C. hero* under past climate (projection based on current model). CUR, current; LGM, Last Glacial Maximum including ice sheets in light blue (Becker et al., 2015); LIG, Last Inter-Glacial; MID, Mid-Holocene

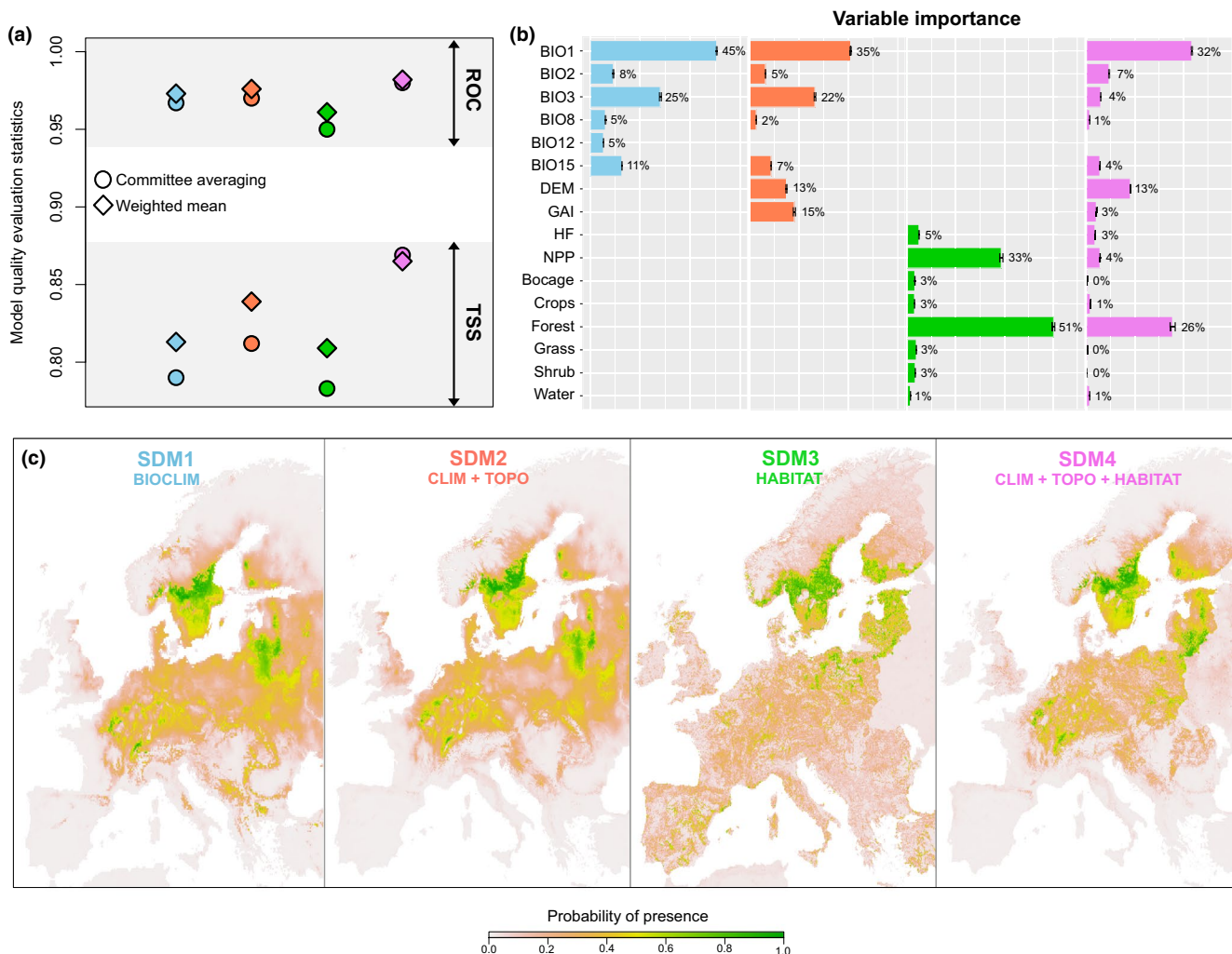


FIGURE 3 Current potential distribution of *Coenonympha hero* based on different sets of environmental variables. (a) Model quality performance comparing two methods (committee averaging and weighted mean) for four different models: SDM1 (BIOCLIM), SDM2 (BIOCLIM, GAI, DEM), SDM3 (HF, NPP, %CLC type) and SDM4 (combining all data). (b) Importance of each environmental variable set. (c) Probability of the presence of *C. hero* in European countries under current conditions for each SDM

model (SDM4) falls below 4%. This model combining all the types of environmental variables has the highest performance and shows that some large Northern areas (around the Baltic Sea) combine climatically favourable and suitable habitat characteristics, while the rest of Europe is mostly unfavourable except a few strongly fragmented areas across Western and central Europe (SDM4; Figure 3c).

3.2 | Current distribution in France

The comparison of environmental characteristics experienced by *C. hero* populations in Doubs, Jura and Saône-et-Loire, and the broader environmental niche occupied by *C. hero* in Europe using PCA revealed that both climatic and habitat characteristics are different (Figure 4). Concerning climatic variation (Figure 4a), PC1 is a precipitation gradient. Northern regions of *C. hero* distribution range are characterized by high seasonality in precipitation (higher

values for BIO15) whereas southern regions in France are characterized by more rainy (higher values for BIO12) and humid (higher values for GAI) conditions. Concerning temperatures (PC2), French populations are located in warmer climate than those from northern latitudes (lower values for BIO1). Concerning habitat variation, northern regions occupied by *C. hero* are composed of larger proportion of forest surfaces, whereas southern regions in France are mainly constituted of grassland surfaces (Figure 4b).

Based on SDM4 projections, we identified several areas in France showing a high probability of presence (Figure 4c). They include mainly the Parisian basin and the two departments of Jura and Doubs. All occurrences ($N = 599$) were considered as presences in European models. Nonetheless, some of the occurrences reported in France (FR in red), Switzerland and the Netherlands, recorded before 1990 (Table S1; Figure S2.3) all correspond to known areas of extinction. It is worth noting that these occurrences are located in the warmer climatic conditions (high values for BIO1; Figure 4a) where human pressure is high

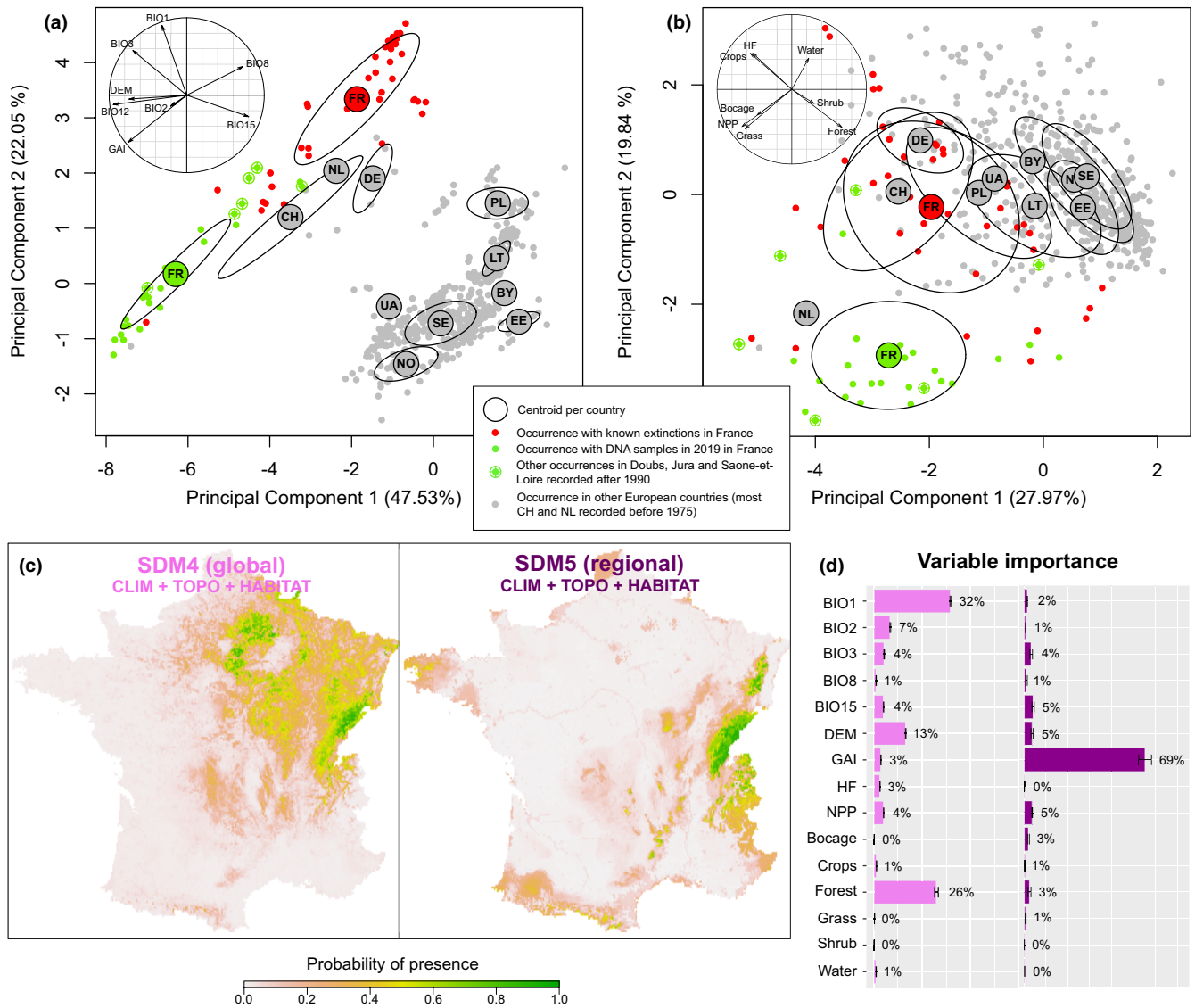


FIGURE 4 Environmental descriptors and probability of the presence of *Coenonympha hero* in France. Principal component analyses showing climatic (a) and habitat (b) variation between northern regions of *C. hero* distribution range and southern regions in France on the two first PCs. (c) Probability of the presence of *C. hero* in France under current environmental conditions using all occurrences as presences (left; SDM4) or considering true absences (right; SDM5). (d) Importance of each environmental variable for SDM4 and SDM5

(high values for HF and %crops; Figure 4b), suggesting that both climate and habitat are important factors for the persistence of *C. hero*. We thus performed a French model (SDM5) using this knowledge on the presence and absence (extinctions) of the species (Figure 1). This model revealed that the potential geographical distribution of *C. hero* in France is mainly reduced to the Jura massif. In accordance with the PCA on climatic variation, GAI is the best environmental predictor of *C. hero* distribution in France (69%; Figure 4d).

3.3 | Population genetic structure and diversity

Population genetics analysis revealed low genetic differentiation among *C. hero* populations (Figure 5). Indeed, the ADMIXTURE

analysis revealed that the 31 sampled localities constitute one large genetic population. Nonetheless, genetic differentiation between populations can be up to $F_{ST} = 0.21$ (Figure S3.3), and both PCA (Figure 4a) and ADMIXTURE (Figure 4b) show geographical structure. At $K = 2$, one genetic group is constituted of the southern populations in the regions JU and CE, and the second group of northern populations in the regions of Doubs (DP and DW) and Saône-et-Loire (SAO). Increasing K allows differentiating populations from JU and CE ($K = 3$), the two easternmost populations LOC and TIN in the Plateau du Doubs ($K = 4$), the population GIR and populations CER/VIR in JU ($K = 5$), the populations from DP and DW ($K = 6$), the north-eastern and south-western populations in Doubs ($K = 7$) and the northernmost population VUI in Doubs ($K = 8$) (Figure 5; Figures S3.1 and S3.2).

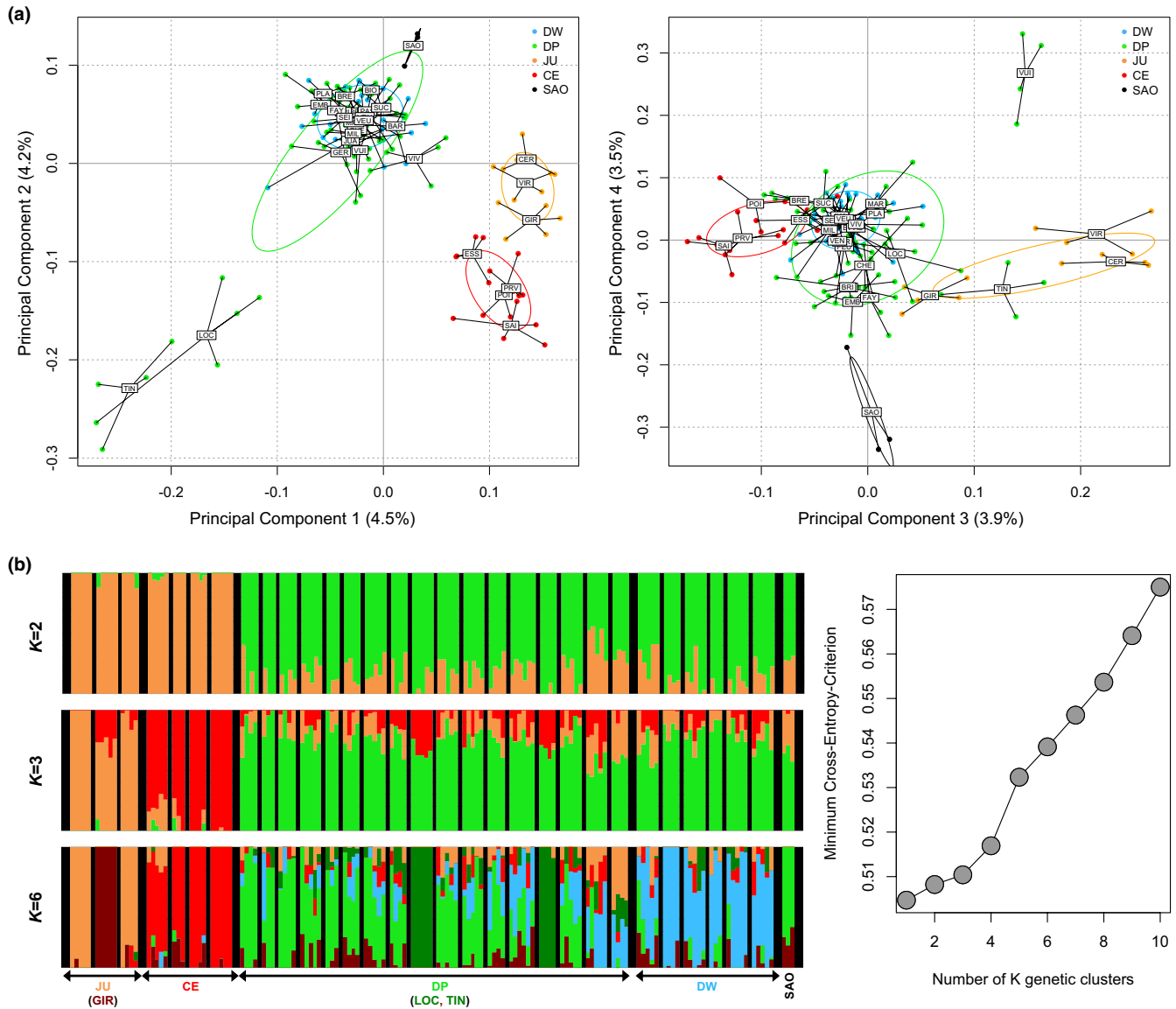


FIGURE 5 Population genetic structure of *Coenonympha hero* populations. (a) Principal component analysis showing genetic variation among populations on the three first PCs. (b) ADMIXTURE barplots (left) for $K = 2$, $K = 3$ and $K = 6$, and cross-validation errors (CVE) showing an optimal number of $K = 1$ (minimum value)

Investigating the determinants of spatial population structure (distance, barriers), we found significant IBD among all the populations (Mantel, $R = .586$, $p < .001$), and similar trends within each region although linear regressions were not significant for JU and CE (Figure S3.3). We further tested whether populations were more genetically distinct than expected under IBD (i.e. higher genetic drift in each region than gene flow between regions). Based on EEMS migration surfaces, we found genetic differences exceeding IBD suggesting barriers to gene flow between regions (Figure 6).

Genetic diversity was lower in the southern populations of JU and CE, with global $H_e = 0.249$ and $H_e = 0.250$, respectively, compared to genetic diversity in the northern populations of Doubs (DP and DW), with global $H_e = 0.276$ and $H_e = 0.268$, respectively

(Table 2). Average H_e was significantly different among regions (ANOVA, $p < .001$), and pairwise comparisons (Tukey) revealed that only JU showed a significantly lower genetic diversity than DP ($p = .002$) and DW ($p = .008$). EEMS diversity surfaces support this south–north gradient in genetic diversity (Figure 6) but we did not detect significantly lower or higher genetic diversity within demes (H_e and H_o ; Table 2). Furthermore, we found that most of the populations are at equilibrium based on F_{IS} test of significance (not significantly different from 0; Table 2). However, several populations in CE (PRV and SAI) and DP (LOC, TINC, VEN and VUI), and all the populations sampled in JU (CER, GIR and VIR) have high population-specific F_{ST} values and low allelic richness suggesting local genetic drift (Table 2).

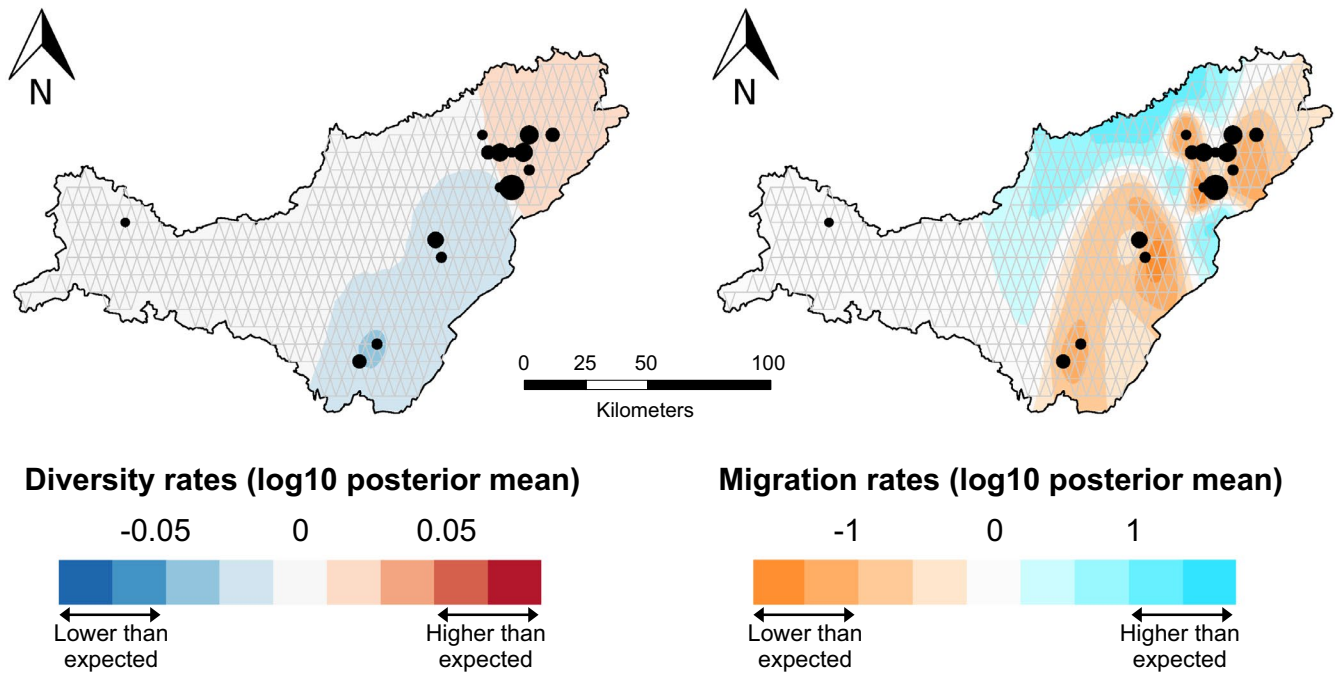


FIGURE 6 Spatial patterns of genetic variability among *Coenonympha hero* populations in France. Left: contour plot of average posterior distribution of diversity rates (log₁₀ scale). Right: contour plot of average posterior distribution of migration rates (log₁₀ scale). Darker font colours indicate significantly higher or lower genetic dissimilarities within- (diversity rates) or between- (migration rates) demes than expected under pure IBD (not significant for diversity rates)

3.4 | Demographic inferences

Based on population genetic results, we inferred the demographic history of three populations corresponding to three geographical regions sampled (DPW: DP-DW, CE and JU; Appendix S4). In the three regions, there is a deficit in rare alleles compared with neutral expectations, which is suggestive of population decline (Figure S4.1). In order to reconstruct the demographic history of *C. hero* populations in France, we first used the stairway plot method to reconstruct continuous fluctuations in effective population size over time. We then used the stairway plot results to design alternative scenario topologies and define prior intervals for ABC analysis to estimate the most likely effective population size and population decline timings (Appendix S4). The two methods are complementary, and each can be used to inform the other.

The three populations show similar demographic histories (Figure 7). For the stairway plot method, we found two events of population decline for DPW and CE while JU population show a strong old population decline followed by a progressive decline over time. For the ABC analysis, we found Scenario 4 as the best scenario: two demographic declines (Figure S4.3 and Table S4.1), which is congruent with the stairway plot results (Figure 7). Although the posterior probabilities of best scenarios range between 0.51 and 0.68, 95% confidence intervals of posterior probabilities do not overlap with intervals of other scenarios (Table S4.1). However, except for CE, type I errors were rather high (~0.30), suggesting that Scenarios 2 and 3 (one decline) are difficult to tell apart from Scenario 4 (two distinct declines) as also suggested by type II errors (Table S4.1), and

by PCA showing strong overlap between the summary statistics for the probability spaces of Scenarios 2, 3 and 4 (Figure S4.4). For JU, assuming discrete decline events potentially does not depict properly the population demographic history, as the stairway plot analyses showed a continuous population decline over time (Figure 7).

In accordance with genetic diversity indices, estimations of effective population size support a south–north gradient with current effective population size estimates twice higher in CE as compared to JU, and five times higher in DPW for both stairway plot and ABC methods (Figure 7; Table S4.2). Timings for old and recent population declines are similar for the three populations (Figure 7). *Coenonympha hero* started to decline between 1,03,000–1,16,000 (ABC) and 50,000–1,40,000 (stairway plot) years ago, with a reduction factor of N_e ranging from 3 (JU) to 5 (CE and DPW). The second more recent decline is dated between 2400–2700 (ABC) and 500–2500 (stairway plot) years ago, with a reduction factor of N_e of 5–7 for DPW, 9 for CE, and 25 for JU (Figure 7; Table S4.2). The first and second declines are of similar magnitude in DPW, while the second decline is much stronger in the two southern populations, especially in JU.

The best divergence scenario was Scenario 3 (divergence overlapping recent decline) with posterior probability ($pp = .997$ [0.997–0.998]) much higher than Scenario 1 (divergence older than first population decline, $pp = .000$ [0.000–0.000]) and Scenario 2 (divergence between first and second population decline, $pp = .003$ [0.003–0.003]). Type I and type II errors are small confirming our ability to discriminate among the three scenarios (type I: 0.002, type II: 0.00 for Scenario 1 and 0.09 for Scenario 2) (Figure 8). The

TABLE 2 Population genetic diversity statistics

| Region | Pop | N | H_o | H_e | F_{IS}^a | F_{ST}^a | A_R^a | N_{sing} | P_{sing} |
|--------|-----|-------|-------|--------|---------------|--------------|-------------|------------|------------|
| SAO | SAO | 3 | 0.214 | 0.238 | 0.022 | 0.154 | 1.23 | 12 | 4.00 |
| CE | All | 17 | 0.247 | 0.250 | 0.022 | 0.084 | 1.25 | 18 | 1.06 |
| | ESS | 5 | 0.258 | 0.250 | -0.043 | 0.088 | 1.25 | 3 | 0.60 |
| | POI | 3 | 0.249 | 0.249 | -0.054 | 0.096 | 1.25 | 5 | 1.67 |
| | PRV | 4 | 0.248 | 0.234 | -0.079 | 0.140 | 1.24 | 2 | 0.50 |
| | SAI | 5 | 0.231 | 0.235 | -0.006 | 0.145 | 1.24 | 8 | 1.60 |
| DP | All | 75 | 0.258 | 0.276 | 0.097 | -0.012 | 1.28 | 110 | 1.47 |
| | BEL | 4 | 0.268 | 0.270 | -0.038 | 0.032 | 1.27 | 9 | 2.25 |
| | BRE | 3 | 0.298 | 0.260 | -0.158 | 0.022 | 1.27 | 4 | 1.33 |
| | BRI | 4 | 0.268 | 0.266 | -0.029 | 0.031 | 1.27 | 12 | 3.00 |
| | CHE | 5 | 0.272 | 0.271 | -0.025 | 0.016 | 1.27 | 2 | 0.40 |
| | EMB | 3 | 0.244 | 0.267 | 0.013 | 0.049 | 1.26 | 5 | 1.67 |
| | FAY | 4 | 0.273 | 0.282 | -0.003 | -0.020 | 1.28 | 4 | 1.00 |
| | FEU | 5 | 0.266 | 0.260 | -0.033 | 0.051 | 1.26 | 5 | 1.00 |
| | GER | 4 | 0.266 | 0.277 | 0.002 | 0.001 | 1.27 | 3 | 0.75 |
| | LOC | 5 | 0.236 | 0.235 | -0.015 | 0.143 | 1.24 | 4 | 0.80 |
| | MAR | 5 | 0.267 | 0.258 | -0.045 | 0.057 | 1.26 | 10 | 2.00 |
| | MIL | 5 | 0.249 | 0.264 | 0.024 | 0.048 | 1.26 | 9 | 1.80 |
| | PAS | 4 | 0.266 | 0.258 | -0.062 | 0.059 | 1.26 | 9 | 2.25 |
| | SUC | 6 | 0.262 | 0.254 | -0.034 | 0.074 | 1.25 | 4 | 0.67 |
| | TIN | 4 | 0.228 | 0.236 | 0.000 | 0.146 | 1.23 | 6 | 1.50 |
| | VEN | 5 | 0.237 | 0.249 | 0.022 | 0.101 | 1.25 | 8 | 1.60 |
| VIV | 5 | 0.257 | 0.256 | -0.030 | 0.073 | 1.25 | 10 | 2.00 | |
| VUI | 4 | 0.241 | 0.249 | -0.011 | 0.101 | 1.25 | 6 | 1.50 | |
| DW | All | 27 | 0.260 | 0.268 | 0.057 | 0.017 | 1.27 | 41 | 1.52 |
| | BAR | 5 | 0.270 | 0.264 | -0.036 | 0.037 | 1.26 | 9 | 1.80 |
| | BIO | 4 | 0.254 | 0.247 | -0.055 | 0.094 | 1.25 | 6 | 1.50 |
| | JUA | 5 | 0.267 | 0.262 | -0.032 | 0.043 | 1.26 | 7 | 1.40 |
| | PLA | 3 | 0.244 | 0.265 | 0.010 | 0.050 | 1.26 | 2 | 0.67 |
| | SEI | 5 | 0.258 | 0.263 | -0.002 | 0.044 | 1.26 | 10 | 2.00 |
| | VEU | 5 | 0.255 | 0.257 | -0.011 | 0.070 | 1.26 | 7 | 1.40 |
| JU | All | 14 | 0.228 | 0.249 | 0.085 | 0.091 | 1.25 | 7 | 0.50 |
| | CER | 5 | 0.234 | 0.222 | -0.060 | 0.187 | 1.22 | 1 | 0.20 |
| | GIR | 5 | 0.219 | 0.226 | 0.007 | 0.184 | 1.22 | 0 | 0.00 |
| | VIR | 4 | 0.236 | 0.242 | -0.007 | 0.119 | 1.24 | 6 | 1.50 |

Note: N , sample size; H_o , observed heterozygosity; H_e , expected heterozygosity; F_{IS} , fixation index; F_{ST} , population-specific F_{ST} ; A_R , allelic richness; N_{sing} , total number of singletons; P_{sing} , proportion of singletons per individual.

^a F_{IS} significantly different from 0 (deviation); $F_{ST} \geq 0.10$ (genetic drift); and low allelic richness (≤ 1.24) indicated in bold.

goodness of fit of the best scenario was assessed using two PCAs to evaluate the position of the observed dataset against (i) the data sets simulated based on priors and (ii) the data sets simulated based on posterior median values of each parameter, showing that simulated data fit well the observed data (Figure S4.6). The divergence between populations is recent, estimated around 800 years ago, and overlaps with the recent population decline timing interval (Figure 8) suggesting a relation between population decline and fragmentation.

4 | DISCUSSION

4.1 | Identification of a Western European glacial refuge for *C. hero*

The Scarce Heath *Coenonympha hero* has a patchy Palaearctic distribution that extends from eastern France across Central Europe and southern Scandinavia to temperate Asia. Classical glacial refuges

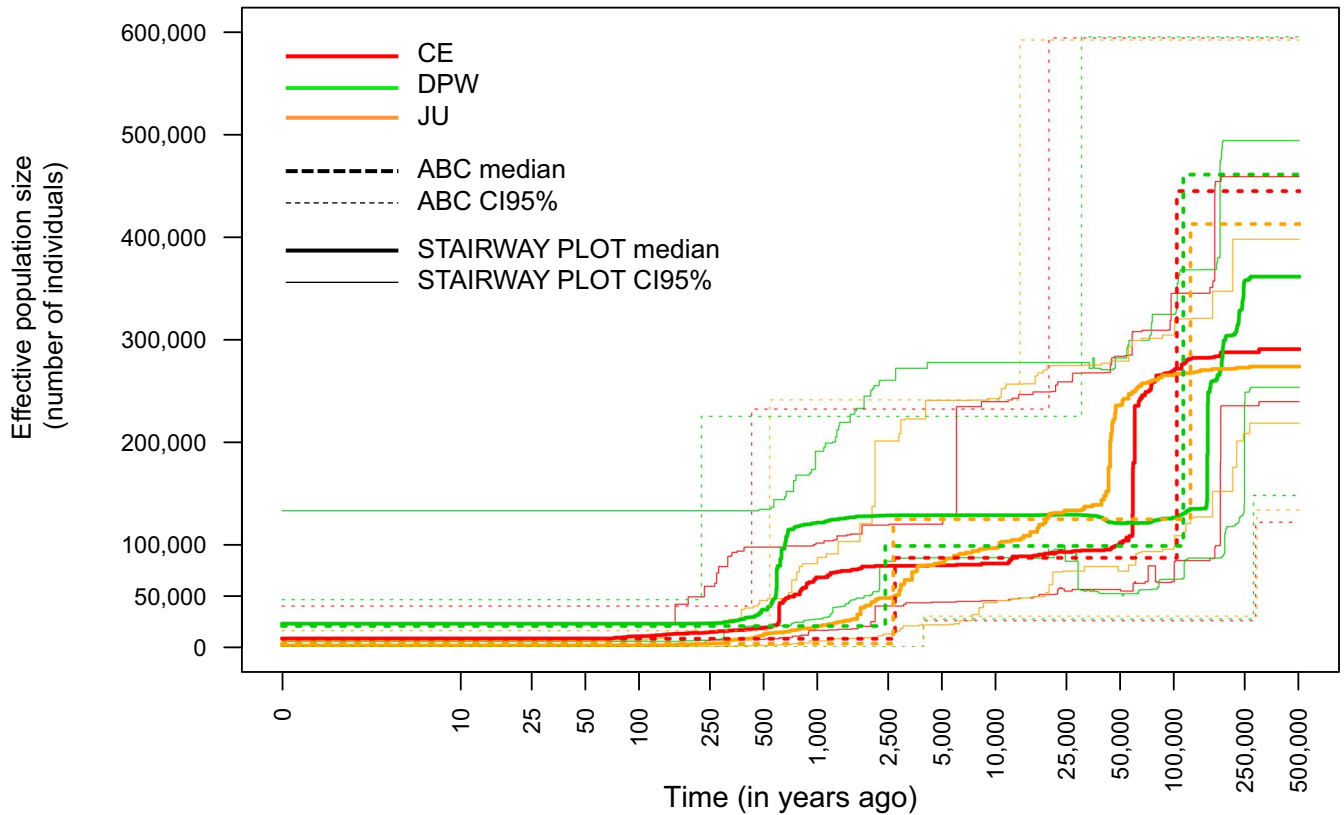


FIGURE 7 Comparison of changes in effective population sizes over time obtained by two different methods. The stairway plot method is based on the SFS and requires to fix the spontaneous mutation rate (here 2.0×10^{-9} ; Appendix S4) and reconstructs continuous fluctuations over time. The ABC method does not require information about the mutation rate but reconstructs discrete events of population decline and requires defining priors. Each population was analysed separately

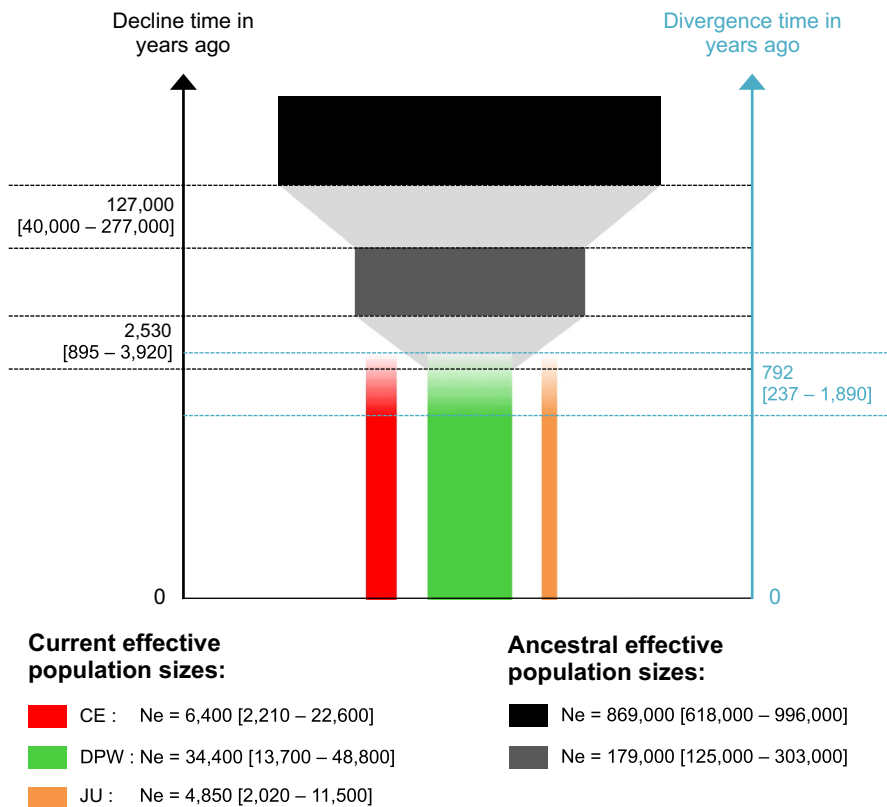


FIGURE 8 Demographic history of *Coenonympha hero* populations in Jura, France. The model includes estimations of current effective population size, population decline and divergence time

reported for European species include Mediterranean refuges (the Iberian, the Apennine and the Balkan peninsulas) and Central European refuges (Schmitt & Varga, 2012; Taberlet et al., 1998; Wendt et al., 2021). For boreal European species, extra-Europe Siberian refuges were also proposed (Schmitt & Varga, 2012). While most of the literature relies on mtDNA variation and haplotype geographical distribution to infer putative Pleistocene refuges (e.g. Dapporto et al., 2019; Ehl et al., 2021), with all the biases inherent to the particular transmission mode of mtDNA (Després, 2019), demographic inference based on multiple nuclear markers allows to go a step further in reconstructing in detail the demographic history of population expansion after the last glacial event. Combining demographic reconstructions and palaeoclimatic SDM further facilitates inferring the location of putative refuges, by identifying stable climatically suitable areas allowing populations to persist during glacial/interglacial oscillations (Ashcroft, 2010; Stewart et al., 2010).

SDM fitted with palaeoclimatic variables (Mid-Holocene, LGM, LIG) revealed that climatically suitable areas for *C. hero* during the LIG were mostly found in Western Europe from Spain to Norway across a large latitudinal range (Figure 2). During the LGM, the foothill of the Jura massif constituted one of the few identified climatically suitable areas together with the Balkans, but this later area was unsuitable during the LIG. This points the Jura region as the only climatically highly suitable area during the LGM, but also before, and after, during the mid-Holocene and current climates. SDM also provided insights into the factors shaping population demographic history. The two demographic inference approaches were congruent and identified a population decline in two steps: a first old decline before the LGM and a second more recent decline, which resulted in the fragmentation of the Jura population into three subpopulations (Figures 7 and 8). The first demographic decline corresponds to the transition from a warm and wet climate during the LIG to cold and dry climate with ice-sheet covering most northern Europe at LGM, supporting that *C. hero* persisted during the LGM in the Jura. Current populations in the Jura region being at the southernmost margin of *C. hero* distribution range, they constitute glacial relict populations.

During post-glacial warming, the ice-free area increased towards North where *C. hero* is currently found. Yet, despite evidence for a geographical northward expansion, we found no evidence for a population size expansion after the LGM (Figure 7). The climate during mid-Holocene was even less suitable than current climate, with only a few highly suitable patchy areas in north-western Europe (Figure 2). This mid-Holocene species-hostile picture despite evidence for successful species colonization northwards was possibly buffered by the presence of large deciduous forests that offer pockets of cooler microenvironment during hot and dry summers for climate-sensitive species (Bladon et al., 2020; Suggitt et al., 2011). Indeed, *C. hero* is currently mainly found in forested areas throughout Europe, especially in central and eastern Europe although it is an open-canopy butterfly (Figure 3). In accordance with a unique refuge with a limited population size, only a single mitochondrial haplotype (CO1 608 bp, BOLD) is found throughout Europe (Appendix S5), which is much lower than haplotype diversity generally observed in

other European butterflies (Cassel-Lundhagen et al., 2020; Després et al., 2019; Dincă et al., 2021).

Using ecological modelling and genetic demographic inferences, we found evidence for a unique Western glacial refuge for *C. hero* in Europe. Our finding is therefore a contribution to the ever-growing literature showing that many species, including butterflies, survived the LGM in Western Europe outside the Mediterranean peninsulas and that boreal species did not necessarily had a far-east Siberian refuge (Després et al., 2019; Minter et al., 2020). A second larger LGM refuge outside Europe for *C. hero* is pointed out by the higher mtDNA haplotype diversity found in eastern Asia, but this refuge did not contribute to Europe post-glacial recolonization (Appendix S5).

4.2 | Metapopulation functioning

Demographic inferences revealed that the three populations (JU, CE and DPW) constituted a single large population (same demographic history; Figures 7 and 8). Population genetic structure analyses (PCA and ADMIXTURE) further suggest a single large population of interconnected subpopulations with weak spatial structure. This overall weak genetic differentiation (global $F_{ST} = 0.076$) reflects the recent divergence of the three core populations (Figure 8) and is expected in a metapopulation functioning where extinction-recolonization takes place between well-connected sites. Furthermore, we found a significant correlation between genetic and geographical distance (IBD pattern), suggesting that gene flow is more likely to occur between geographically close sites. Despite this weak genetic structure, three main genetic clusters were identified (Figures 5 and 6) corresponding to the three main geographical regions sampled in Jura, along a south-north axis. The northern cluster gathered 23 sites from DP and DW and was more diversified than the two other clusters. Surprisingly, the only population sampled in Saône-et-Loire clustered with the northern cluster, although it is more than 100 km from all the other sampled sites, while two northern peripheric sites (TIN and LOC) were distinguished from the core population. These two sites only 1.6 km apart from each other were also the less genetically diverse among northern sites ($H_e = 0.236$ and 0.235, respectively, as compared to 0.276 on average in Doubs sites; Table 2) suggesting allelic loss through genetic drift. Genetic drift could result from a lack of connectivity between these sites and the core population due to their north-eastern isolation. Indeed, the core population is >15 km, far more than the dispersal capacity of this small butterfly which has been estimated by mark-release-recapture (MRR) to a few hundred metres (Cassel-Lundhagen & Sjögren-Gulve, 2007). It could also potentially be due to low local habitat quality and small local population, as suggested by SDM4 with a lower proportion of high-quality habitats in the southern regions (Figure 4). The lowest genetic diversity was found in the two southern clusters, CE and JU, and especially in one of the southernmost localities (CER, $H_e = 0.222$). This is unexpected based on main phylogeographic paradigms, which hypothesize lower genetic diversity to the north (e.g. Dincă et al., 2021). The unexpected northern

richness and southern lower genetic diversity could be due to rather recent (thousands of years) factors (increased aridity, habitat fragmentation) than to the northward expansion that occurred at the end of the Pleistocene.

Genetic diversity (H_e) within each sampled site ranged from 0.222 to 0.282 which is comparable to the diversity observed in other *Coenonympha* European butterfly populations using the same ddRAD procedure (Capblancq et al., 2015; Després et al., 2019), and nearly twice that of a currently expanding but young crop-feeding butterfly, *Pieris rapae* (Ryan et al., 2019). Most sites were at demographic equilibrium (F_{IS} non-significantly different from zero; Table 2), which suggests, together with the weak genetic structure observed at the regional scale, an overall good connectivity between sites and a dynamical metapopulation functioning. This was unexpected given the geographic distances between the three main core populations, which largely exceed the maximum flight distance observed by MRR (1200m, Cassel-Lundhagen & Sjögren-Gulve, 2007). Genetic estimates based on a large number of markers can capture rare long-distance events that are missed by direct dispersal observations. Furthermore, forecasting distribution in France suggests that the Jura massif is constituted of continuous suitable habitats (probability of presence >0.8; Figure 4). The genetic proximity between the Saône-et-Loire site and the northern core population (DPW) suggests that such long-distance events can take place in *C. hero*, either through passive transport of imagoes (wind) or larvae (hay trade), or through active flight during successive butterfly generations breeding in patchy suitable habitats between Doubs and Saône-et-Loire. Indeed, very small populations temporary occupying an habitat can be overlooked, as exemplified by the recently discovered patchy network of occupied sites in the CE region located between the long-known occupied sites in Jura and Doubs regions, and the even more recently discovered site in Saône-et-Loire where the species was thought to have gone extinct in the 80s. The three individuals sampled in this remote location of Saône-et-Loire were not related, suggesting that the founding event from the core population is not very recent (more than 3 generations ago) and that the local population is not restricted to the descendants of a single founding female. Furthermore, the overall genetic diversity in this site is not lower than in other more connected localities, yet the high population-specific F_{ST} is a signature of founding effect and further drift due to limited connectivity with the source population.

4.3 | Climate and land use constrain *C. hero* persistence

As all other Satyrinae, *C. hero* larvae feed on various types of widespread grass and are not specialized on a particular host plant (Tiitsaar et al., 2016). Neither do they depend on a very particular biotope such as two highly endangered wetland *Coenonympha* sp. only found in low altitude alkaline marshes (*C. oedippus*), or in cold raised bogs (*C. tullia*). These two habitat-specialist species are particularly affected by land drainage and are threatened throughout Europe

(van Swaay et al., 2006). Depending on the country in Europe, *C. hero* can be found in more or less forested steppes, meadows and grasslands, and in semi-open deciduous or mixed forests, fens or transition mires (Sielezniew & Nowicki, 2017; van Swaay et al., 2006). Despite this variety of biotopes used, the scarce heath *C. hero* is strongly declining in Europe.

This butterfly has a particularly narrow climatic tolerance. At the European scale, the mean annual temperature at which the species occurs is restricted between 3.5°C and 11.5°C. Recent probable extinctions (sites not reported after 1980) in Europe constitute almost all the occurrences found between 9°C and 11.5°C, suggesting that populations in warmest regions are more vulnerable. The southernmost populations of *C. hero* are located in France, where known extinctions occurred in warmest regions with high human footprint (Figure 4). Furthermore, records from Switzerland, Netherlands and Germany fall into this same ecological space, in accordance with the lack of recent observations of the species and suspected extinction in these countries (Kudrna et al., 2011; Lütolf et al., 2006; van Swaay & Warren, 1999). This local extinction pattern supports range contractions at the warm edges in response to climate and habitat changes (Cahill et al., 2014). The mean annual temperature in the Jura region ranging between 7°C and 10°C, these glacial relict populations are at major extinction risk given the unprecedented rate of contemporary climate change (Bennett et al., 2021). Interestingly, despite evidence for a dramatic recent decline especially in the southernmost populations, they retain high genetic diversity, meaning the whole metapopulation might have some evolutionary potential to respond to environmental change. This adaptive potential of *C. hero* French populations is suggested by their marginal position in the environmental space of the whole species in Europe (Figure 4) and could potentially contribute to evolutionary rescue northern European populations that appear to be already suffering from genetic erosion (Cassel & Tamaru, 2003) and inbreeding depression (Cassel et al., 2001).

In conjunction with warming, the reduction of suitable areas due to anthropization, as observed for *C. hero* in France, can have dramatic impacts on the most isolated populations, resulting in between-population genetic differentiation and within-population genetic erosion (Frankham et al., 2014; Hampe & Petit, 2005). Some 2000 years ago, genetic inferences identify a second population decline of similar (Doubs) or much stronger (Jura) amplitude as the first decline (Figure 7), although climatic projections suggest more suitable current conditions especially in northern Europe than during Holocene (Figure 2). This second decline is also supported by records of local extinctions during the last century in the most populated European countries (high human footprint index; Figure 4), suggesting that increasing anthropization and land use changes prevented *C. hero* population expansion and precipitated its decline. Indeed, human population in Europe has been multiplied by more than 1000 (Figure S4.7: Appendix S4) during the more recent *C. hero* population decline. Despite these threats on *C. hero*, we found no evidence for increased consanguinity nor genetic erosion within sampled sites, which appear to

maintain good connectivity with neighbouring sites and to function as a metapopulation, with high genetic diversity maintained at the regional scale. The current effective population sizes estimated, namely ~34,000, 6000 and 5000 individuals for the DPW, CE and JU populations, are compatible with the population size estimates by MRR in other studies (e.g. in Poland: ~200 individuals per ha, Sielezniew & Nowicki, 2017; in Sweden: ~50 individual per ha, Cassel et al., 2001) given the quantity of suitable habitats available in the Jura predicted by SDM.

The climatic niche of *C. hero* in France is best captured by the global aridity index (GAI) than by any other bioclimatic variable (Figure 4), a local index combining precipitation, potential evapotranspiration and temperature (Trabucco & Zomer, 2019). Furthermore, while most of the occupied areas in Europe are below 400 m, populations in Jura are found at an altitude between 400 m and 850 m. Therefore, both the habitat (altitude grasslands) and the climate (more hot and wet) currently occupied in its historical glacial refuge Jura massif are distinct and marginal as compared to the forested cooler habitats occupied in its current core distribution in northern and eastern Europe (Figure 4). Rather than a habitat-specialist, *C. hero* appears to be a climate tracker, and its local survival under warming conditions depends on the possibility to evade at higher altitude (in Jura) or latitude (in Doubs), or in semi-open forests where it occupies transitory open glades. Our results thus support that microclimate is a crucial aspect of habitat suitability at the margins of species' distribution range in butterflies (Eilers et al., 2013; Lawson et al., 2014; Örvössy et al., 2013; Turlure et al., 2010).

The evaluation of genetic diversity in glacial relict populations is critical for understanding and preventing extinctions in southern refuges (Jiménez-Alfaro et al., 2016; Stewart et al., 2010) but also throughout the species distribution range as these glacial relicts are expected to retain more genetic diversity and therefore more evolutionary potential than younger northern populations (Taberlet et al., 1998). In face of global warming and rapidly increasing temperature throughout Europe (Squintu et al., 2021), local adaptations selected in southernmost populations might prove to be useful to evolutionary rescue populations not yet exposed to extreme climatic conditions (Razgour et al., 2019). *Coenonympha hero* appears to maintain over time through small interconnected populations able to buffer climate change by shifting to cooler habitats (more forested, altitude). Nonetheless, the less diverse and most declining population is the southernmost population (JU) and particular attention is needed to maintain gene flow with the northern core population (DPW). Both climate and land use variables are needed to accurately predict the distribution of *C. hero* and ongoing changes (warming and increased human footprint) combine as factors of extinction risk, while we found no evidence for *C. arcania* introgression by mapping *C. hero* ddRAD data set to *C. arcania* genome (data not shown), indicating that hybridization with other *Coenonympha* species is not a driver of *C. hero* decline. Conservation actions should aim at maintaining and increasing patches of semi-open habitats and forest glades to help the species to maintain in

its relictual historical refuge in Jura, but also to favour the connectivity between these marginal southernmost populations and all the other populations in Europe, as it will favour the spreading of warm-adapted genes elsewhere in Europe.

ACKNOWLEDGEMENTS

We thank François Dehondt and Conservatoire Botanique National de Franche-Comté Observatoire Régional des Invertébrés for funding this research and all the collectors: Frédéric Mora, Perrine Jacquot, Catherine Duflo, Julien Ryelandt (Conservatoire Botanique National de Franche-Comté Observatoire Régional des Invertébrés), Anaëlle Bernard, Jocelyn Claude, Bruno Tissot (Association des Amis de la Réserve Nationale du Lac de Remoray), Luc Bettinelli, Magalie Mazuy and Magali Crouvezier (Conservatoire Espaces Naturels de Franche Comté). S. Sherpa and C. Kebaili were supported by fellowships from the University Grenoble Alpes. L. Despres is supported by the French Ministry for Higher Education and Research.

CONFLICT OF INTEREST

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

The raw sequences (fastq) analysed in the present study are available at the European Nucleotide Archive repository (<http://www.ebi.ac.uk/ena>) and accessible under study accession number PRJEB48550 (individual accession numbers in Supporting Information). The two SNP data sets (in VCF format) used for population genetics and demographic inferences, and the occurrence data sets used for each SDM analysis are provided in Supporting Information. In accordance with the French regulation regarding endangered species (article L.124-4 du code de l'environnement), the precision of coordinates of the localities sampled in the Jura massif was degraded to 10 × 10 km precision but are available from authors upon request.

ORCID

Stéphanie Sherpa  <https://orcid.org/0000-0001-9958-0073>

Caroline Kebaili  <https://orcid.org/0000-0002-8472-7981>

Laurence Després  <https://orcid.org/0000-0002-0660-6260>

REFERENCES

- Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, 19(9), 1655–1664. <https://doi.org/10.1101/gr.094052.109>
- Allouche, O., Tsoar, A., & Kadmon, R. (2006). Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology*, 43(6), 1223–1232. <https://doi.org/10.1111/j.1365-2664.2006.01214.x>
- Araújo, M. B., & New, M. (2007). Ensemble forecasting of species distributions. *Trends in Ecology & Evolution*, 22(1), 42–47. <https://doi.org/10.1016/j.tree.2006.09.010>
- Ashcroft, M. B. (2010). Identifying refugia from climate change. *Journal of Biogeography*, 37(8), 1407–1413. <https://doi.org/10.1111/j.1365-2699.2010.02300.x>

- Becker, D., Verheul, J., Zickel, M., & Willmes, C. (2015). LGM paleo-environment of Europe -Map. In *CRC806-database*. <https://doi.org/10.5880/SFB806.15>
- Bennett, J. M., Sunday, J., Calosi, P., Villalobos, F., Martínez, B., Molina-Venegas, R., Araújo, M. B., Algar, A. C., Clusella-Trullas, S., Hawkins, B. A., Keith, S. A., Kühn, I., Rahbek, C., Rodríguez, L., Singer, A., Morales-Castilla, I., & Olalla-Tárraga, M. Á. (2021). The evolution of critical thermal limits of life on Earth. *Nature Communications*, 12(1), 1–9. <https://doi.org/10.1038/s41467-021-21263-8>
- Bladon, A. J., Lewis, M., Bladon, E. K., Buckton, S. J., Corbett, S., Ewing, S. R., Hayes, M. P., Hitchcock, G. E., Knock, R., Lucas, C., McVeigh, A., Menéndez, R., Walker, J. M., Fayle, T. M., & Turner, E. C. (2020). How butterflies keep their cool: Physical and ecological traits influence thermoregulatory ability and population trends. *Journal of Animal Ecology*, 89(11), 2440–2450. <https://doi.org/10.1111/1365-2656.13319>
- Bushnell, B. (2014). *BBMap: A fast, accurate, splice-aware aligner* (No. LBNL-7065E). Lawrence Berkeley National Lab. (LBNL).
- Cahill, A. E., Aiello-Lammens, M. E., Caitlin Fisher-Reid, M., Hua, X., Karanewsky, C. J., Ryu, H. Y., Sbeglia, G. C., Spagnolo, F., Waldron, J. B., & Wiens, J. J. (2014). Causes of warm-edge range limits: systematic review, proximate factors and implications for climate change. *Journal of Biogeography*, 41(3), 429–442. <https://doi.org/10.1111/jbi.12231>
- Capblancq, T., Després, L., Rioux, D., & Mavárez, J. (2015). Hybridization promotes speciation in *Coenonympha* butterflies. *Molecular Ecology*, 24(24), 6209–6222.
- Cassel, A., & Tammaru, T. (2003). Allozyme variability in central, peripheral and isolated populations of the scarce heath (*Coenonympha hero*: Lepidoptera, Nymphalidae); implications for conservation. *Conservation Genetics*, 4(1), 83–93.
- Cassel, A., Windig, J., Nylin, S., & Wiklund, C. (2001). Effects of population size and food stress on fitness-related characters in the scarce heath, a rare butterfly in western Europe. *Conservation Biology*, 15(6), 1667–1673. <https://doi.org/10.1046/j.1523-1739.2001.99557.x>
- Cassel-Lundhagen, A., Schmitt, T., Wahlberg, N., Sarvašová, L., Konvička, M., Ryrholm, N., & Kaňuch, P. (2020). Wing morphology of the butterfly *Coenonympha arcania* in Europe: Traces of both historical isolation in glacial refugia and current adaptation. *Journal of Zoological Systematics and Evolutionary Research*, 58(4), 929–943.
- Cassel-Lundhagen, A., & Sjögren-Gulve, P. (2007). Limited dispersal by the rare scarce heath butterfly—potential consequences for population persistence. *Journal of Insect Conservation*, 11(2), 113–121. <https://doi.org/10.1007/s10841-006-9023-z>
- Cassel-Lundhagen, A., Sjögren-Gulve, P., & Sven-Ake, B. (2008). Effects of patch characteristics and isolation on relative abundance of the scarce heath butterfly *Coenonympha hero* (Nymphalidae). *Journal of Insect Conservation*, 12(5), 477–482. <https://doi.org/10.1007/s10841-007-9083-8>
- Chessel, D., Dufour, A. B., & Thioulouse, J. (2004). The ade4 package—One-table methods. *R News*, 4(1), 5–10.
- Chi, E. C., Zhou, H., Chen, G. K., Del Vecchio, D. O., & Lange, K. (2013). Genotype imputation via matrix completion. *Genome Research*, 23(3), 509–518. <https://doi.org/10.1101/gr.145821.112>
- Cornuet, J.-M., Pudlo, P., Veyssier, J., Dehne-Garcia, A., Gautier, M., Leblois, R., Marin, J.-M., & Estoup, A. (2014). DIYABC v2.0: A software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. *Bioinformatics*, 30, 1187–1189. <https://doi.org/10.1093/bioinformatics/btt763>
- Cornuet, J. M., Ravigné, V., & Estoup, A. (2010). Inference on population history and model checking using DNA sequence and microsatellite data with the software DIYABC (v1.0). *BMC Bioinformatics*, 11, 401.
- Dapporto, L., Cini, A., Vodá, R., Dincă, V., Wiemers, M., Menchetti, M., Magini, G., Talavera, G., Shreeve, T., Bonelli, S., Casacci, L. P., Balletto, E., Scalerio, S., & Vila, R. (2019). Integrating three comprehensive data sets shows that mitochondrial DNA variation is linked to species traits and paleogeographic events in European butterflies. *Molecular Ecology Resources*, 19(6), 1623–1636. <https://doi.org/10.1111/1755-0998.13059>
- Després, L. (2019). One, two or more species? Mitonuclear discordance and species delimitation. *Molecular Ecology*, 28(17), 3845–3847. <https://doi.org/10.1111/mec.15211>
- Després, L., Henniaux, C., Rioux, D., Capblancq, T., Zupan, S., ěelik, T., Sielezniew, M., Bonato, L., & Ficetola, G. F. (2019). Inferring the biogeography and demographic history of an endangered butterfly in Europe from multilocus markers. *Biological Journal of the Linnean Society*, 126(1), 95–113. <https://doi.org/10.1093/biolinnean/bly160>
- Dincă, V., Dapporto, L., Somervuo, P., Vodá, R., Cuvelier, S., Gascoigne-Pees, M., Huemer, P., Mutanen, M., Hebert, P. D. N., & Vila, R. (2021). High resolution DNA barcode library for European butterflies reveals continental patterns of mitochondrial genetic diversity. *Communications Biology*, 4(1), 1–11. <https://doi.org/10.1038/s42003-021-01834-7>
- Ehl, S., Ehl, S., Kramp, K., & Schmitt, T. (2021). Interglacials are driving speciation and intraspecific differentiation in the cold-adapted butterfly species group *Boloria pales/napaea* (Nymphalidae). *Journal of Biogeography*, 48(1), 134–146.
- Eilers, S., Pettersson, L. B., & Öckinger, E. (2013). Micro-climate determines oviposition site selection and abundance in the butterfly *Pyrgus armoricanus* at its northern range margin. *Ecological Entomology*, 38(2), 183–192.
- Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, 37(12), 4302–4315. <https://doi.org/10.1002/joc.5086>
- Fourcade, Y., Engler, J. O., Rödder, D., & Secondi, J. (2014). Mapping species distributions with MAXENT using a geographically biased sample of presence data: A performance assessment of methods for correcting sampling bias. *PLoS One*, 9(5), e97122. <https://doi.org/10.1371/journal.pone.0097122>
- Fourcade, Y., WallisDeVries, M. F., Kuussaari, M., van Swaay, C. A., Heliölä, J., & Öckinger, E. (2021). Habitat amount and distribution modify community dynamics under climate change. *Ecology Letters*, 24(5), 950–957. <https://doi.org/10.1111/ele.13691>
- Fox, R., Oliver, T. H., Harrower, C., Parsons, M. S., Thomas, C. D., & Roy, D. B. (2014). Long-term changes to the frequency of occurrence of British moths are consistent with opposing and synergistic effects of climate and land-use changes. *Journal of Applied Ecology*, 51(4), 949–957. <https://doi.org/10.1111/1365-2664.12256>
- Frankham, R., Bradshaw, C. J., & Brook, B. W. (2014). Genetics in conservation management: revised recommendations for the 50/500 rules, Red List criteria and population viability analyses. *Biological Conservation*, 170, 56–63. <https://doi.org/10.1016/j.biocon.2013.12.036>
- Goslee, S. C., & Urban, D. L. (2007). The ecodist package for dissimilarity-based analysis of ecological data. *Journal of Statistical Software*, 22(7), 1–19.
- Goudet, J. (2005). Hierfstat, a package for R to compute and test hierarchical F-statistics. *Molecular Ecology Notes*, 5(1), 184–186. <https://doi.org/10.1111/j.1471-8286.2004.00828.x>
- Habel, J. C., & Schmitt, T. (2018). Vanishing of the common species: Empty habitats and the role of genetic diversity. *Biological Conservation*, 218, 21–216. <https://doi.org/10.1016/j.biocon.2017.12.018>
- Hampe, A., & Petit, R. J. (2005). Conserving biodiversity under climate change: The rear edge matters. *Ecology Letters*, 8(5), 461–467. <https://doi.org/10.1111/j.1461-0248.2005.00739.x>
- Hanski, I., & Thomas, C. D. (1994). Metapopulation dynamics and conservation: A spatially explicit model applied to

- butterflies. *Biological Conservation*, 68(2), 167–180. [https://doi.org/10.1016/0006-3207\(94\)90348-4](https://doi.org/10.1016/0006-3207(94)90348-4)
- Imhoff, M. L., Bounoua, L., Ricketts, T., Loucks, C., Harriss, R., & Lawrence, W. T. (2004). Global patterns in human consumption of net primary production. *Nature*, 429(6994), 870–873.
- Jiménez-Alfaro, B., García-Calvo, L., García, P., & Acebes, J. L. (2016). Anticipating extinctions of glacial relict populations in mountain refugia. *Biological Conservation*, 201, 243–251. <https://doi.org/10.1016/j.biocon.2016.07.015>
- Kudrna, O., Harpke, A., Lux, K., Pennerstorfer, J., Schweiger, O., Settele, J., & Wiemers, M. (2011). *Distribution atlas of butterflies in Europe*. Gesellschaft für Schmetterlingsschutz eV.
- Lafranchis, T. (2000). *Les papillons de jour de France, Belgique et Luxembourg et leurs chenilles*. Biotope Editions.
- Lawson, C. R., Bennie, J., Hodgson, J. A., Thomas, C. D., & Wilson, R. J. (2014). Topographic microclimates drive microhabitat associations at the range margin of a butterfly. *Ecography*, 37(8), 732–740. <https://doi.org/10.1111/ecog.00535>
- Lehikoinen, A., Jaatinen, K., Vähätalo, A. V., Clausen, P., Crowe, O., Deceuninck, B., Hearn, R., Holt, C. A., Hornman, M., Keller, V., Nilsson, L., Langendoen, T., Tománková, I., Wahl, J., & Fox, A. D. (2013). Rapid climate driven shifts in wintering distributions of three common waterbird species. *Global Change Biology*, 19(7), 2071–2081. <https://doi.org/10.1111/gcb.12200>
- Linck, E., & Battey, C. J. (2019). Minor allele frequency thresholds strongly affect population structure inference with genomic data sets. *Molecular Ecology Resources*, 19, 639–647. <https://doi.org/10.1111/1755-0998.12995>
- Liu, X., & Fu, Y. X. (2015). Exploring population size changes using SNP frequency spectra. *Nature Genetics*, 47, 555–559. <https://doi.org/10.1038/ng.3254>
- Lütolf, M., Kienast, F., & Guisan, A. (2006). The ghost of past species occurrence: Improving species distribution models for presence-only data. *Journal of Applied Ecology*, 43(4), 802–815. <https://doi.org/10.1111/j.1365-2664.2006.01191.x>
- Luu, K., Bazin, E., & Blum, M. G. (2017). pcadapt: An R package to perform genome scans for selection based on principal component analysis. *Molecular Ecology Resources*, 17(1), 67–77.
- Marmion, M., Parviainen, M., Luoto, M., Heikkinen, R. K., & Thuiller, W. (2009). Evaluation of consensus methods in predictive species distribution modelling. *Diversity and Distributions*, 15(1), 59–69. <https://doi.org/10.1111/j.1472-4642.2008.00491.x>
- Melbourne, B. A., & Hastings, A. (2008). Extinction risk depends strongly on factors contributing to stochasticity. *Nature*, 454(7200), 100–103.
- Menéndez, R., Megías, A. G., Hill, J. K., Braschler, B., Willis, S. G., Collingham, Y., Fox, R., Roy, D. B., & Thomas, C. D. (2006). Species richness changes lag behind climate change. *Proceedings of the Royal Society B: Biological Sciences*, 273(1593), 1465–1470.
- Minter, M., Dasmahapatra, K. K., Thomas, C. D., Morecroft, M. D., Tonhasca, A., Schmitt, T., Siozios, S., & Hill, J. K. (2020). Past, current, and potential future distributions of unique genetic diversity in a cold-adapted mountain butterfly. *Ecology and Evolution*, 10(20), 11155–11168. <https://doi.org/10.1002/ece3.6755>
- Nonaka, E., Sirén, J., Somervuo, P., Ruokolainen, L., Ovaskainen, O., & Hanski, I. (2019). Scaling up the effects of inbreeding depression from individuals to metapopulations. *Journal of Animal Ecology*, 88(8), 1202–1214. <https://doi.org/10.1111/1365-2656.13011>
- Örvössy, N., Kőrösi, Á., Batáry, P., Vozár, A., & Peregovits, L. (2013). Potential metapopulation structure and the effects of habitat quality on population size of the endangered False Ringlet butterfly. *Journal of Insect Conservation*, 17(3), 537–547. <https://doi.org/10.1007/s10841-012-9538-4>
- Parmesan, C., Ryrholm, N., Stefanescu, C., Hill, J. K., Thomas, C. D., Descimon, H., Huntley, B., Kaila, L., Kullberg, J., Tammaru, T., Tennent, W. J., Thomas, J. A., & Warren, M. (1999). Poleward shifts in geographical ranges of butterfly species associated with regional warming. *Nature*, 399(6736), 579–583.
- Petkova, D., Novembre, J., & Stephens, M. (2016). Visualizing spatial population structure with estimated effective migration surfaces. *Nature Genetics*, 48(1), 94–100. <https://doi.org/10.1038/ng.3464>
- Pew, J., Wang, J., Muir, P., & Frasier, T. (2017). *related: an R package for analyzing pairwise relatedness data based on codominant molecular markers*. R package version 0.8/r2.
- Razgour, O., Forester, B., Taggart, J. B., Bekaert, M., Juste, J., Ibáñez, C., Puechmaile, S. J., Novella-Fernandez, R., Alberdi, A., & Manel, S. (2019). Considering adaptive genetic variation in climate change vulnerability assessment reduces species range loss projections. *Proceedings of the National Academy of Sciences of the United States of America*, 116(21), 10418–10423. <https://doi.org/10.1073/pnas.1820663116>
- Rochette, N. C., Rivera-Colón, A. G., & Catchen, J. M. (2019). Stacks 2: Analytical methods for paired-end sequencing improve RADseq-based population genomics. *Molecular Ecology*, 28(21), 4737–4754. <https://doi.org/10.1111/mec.15253>
- Rödder, D., Schmitt, T., Gros, P., Ulrich, W., & Habel, J. C. (2021). Climate change drives mountain butterflies towards the summits. *Scientific Reports*, 11(1), 1–12. <https://doi.org/10.1038/s41598-021-93826-0>
- Ryan, S. F., Lombaert, E., Espeset, A., Vila, R., Talavera, G., Dincă, V., Doellman, M. M., Renshaw, M. A., Eng, M. W., Hornett, E. A., Li, Y., Pfreder, M. E., & Shoemaker, D. (2019). Global invasion history of the agricultural pest butterfly *Pieris rapae* revealed with genomics and citizen science. *Proceedings of the National Academy of Sciences of the United States of America*, 116(40), 20015–20024.
- Schmitt, T., & Varga, Z. (2012). Extra-Mediterranean refugia: the rule and not the exception? *Frontiers in Zoology*, 9(1), 1–12. <https://doi.org/10.1186/1742-9994-9-22>
- Sielezniew, M., & Nowicki, P. (2017). Adult demography of an isolated population of the threatened butterfly Scarce Heath *Coenonympha hero* and its conservation implications. *Journal of Insect Conservation*, 21(4), 737–742. <https://doi.org/10.1007/s10841-017-0014-z>
- Squintu, A. A., van der Schrier, G., van den Besselaar, E., van der Linden, E., Putrasahan, D., Roberts, C., Roberts, M., Scoccimarro, E., Senan, R., & Klein Tank, A. (2021). Evaluation of trends in extreme temperatures simulated by HighResMIP models across Europe. *Climate Dynamics*, 56(7), 2389–2412. <https://doi.org/10.1007/s00382-020-05596-6>
- Stewart, J. R., Lister, A. M., Barnes, I., & Dalén, L. (2010). Refugia revisited: individualistic responses of species in space and time. *Proceedings of the Royal Society B: Biological Sciences*, 277(1682), 661–671.
- Suggitt, A. J., Gillingham, P. K., Hill, J. K., Huntley, B., Kunin, W. E., Roy, D. B., & Thomas, C. D. (2011). Habitat microclimates drive fine-scale variation in extreme temperatures. *Oikos*, 120(1), 1–8. <https://doi.org/10.1111/j.1600-0706.2010.18270.x>
- Swets, J. A. (1988). Measuring the accuracy of diagnostic systems. *Science*, 240(4857), 1285–1293.
- Taberlet, P., Fumagalli, L., Wust-Saucy, A. G., & Cosson, J. F. (1998). Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology*, 7(4), 453–464. <https://doi.org/10.1046/j.1365-294x.1998.00289.x>
- Thuiller, W., Georges, D., Engler, R., Breiner, F., Georges, M. D., & Thuiller, C. W. (2016). *Package 'biomod2'*. Species distribution modeling within an ensemble forecasting framework. R package. <http://R-Forge.R-project.org>
- Thuiller, W., Lafourcade, B., Engler, R., & Araújo, M. B. (2009). BIOMOD—a platform for ensemble forecasting of species distributions. *Ecography*, 32(3), 369–373. <https://doi.org/10.1111/j.1600-0587.2008.05742.x>
- Tiitsaar, A., Kaasik, A., Lindman, L., Stanevičs, T., & Tammaru, T. (2016). Host associations of *Coenonympha hero* (Lepidoptera: Nymphalidae) in northern Europe: Microhabitat rather than plant species. *Journal*

- of *Insect Conservation*, 20(2), 265–275. <https://doi.org/10.1007/s10841-016-9861-2>
- Trabucco, A. & Zomer, R. (2019). Global Aridity Index and Potential Evapotranspiration (ETO) climate database v2. figshare. Fileset. <https://doi.org/10.6084/m9.figshare.7504448.v3>
- Turlure, C., Choutt, J., Baguette, M., & Van Dyck, H. (2010). Microclimatic buffering and resource-based habitat in a glacial relict butterfly: significance for conservation under climate change. *Global Change Biology*, 16(6), 1883–1893. <https://doi.org/10.1111/j.1365-2486.2009.02133.x>
- van Swaay, C., & Warren, M. (1999). *Red data book of European butterflies (Rhopalocera)* (vol. 99). Council of Europe.
- van Swaay, C., Warren, M., & Lois, G. (2006). Biotope use and trends of European butterflies. *Journal of Insect Conservation*, 10(2), 189–209. <https://doi.org/10.1007/s10841-006-6293-4>
- Venter, O., Sanderson, E. W., Magrath, A., Allan, J. R., Beher, J., Jones, K. R., Possingham, H. P., Laurance, W. F., Wood, P., Fekete, B. M., Levy, M. A., & Watson, J. E. (2018). *Last of the Wild Project, Version 3 (LWP-3): 2009 Human Footprint, 2018 Release*. NASA Socioeconomic Data and Applications Center (SEDAC).
- Wang, J. (2002). An estimator for pairwise relatedness using molecular markers. *Genetics*, 160(3), 1203–1215. <https://doi.org/10.1093/genetics/160.3.1203>
- Warren, M. S., Hill, J. K., Thomas, J. A., Asher, J., Fox, R., Huntley, B., & Thomas, C. D. (2001). Rapid responses of British butterflies to opposing forces of climate and habitat change. *Nature*, 414(6859), 65–69.
- Wendt, M., Husemann, M., Kramp, K., & Schmitt, T. (2021). Reconstruction of forest dynamics in the Western Palearctic based on phylogeographic analysis of the ringlet butterfly *Erebia aethiops*. *Scientific Reports*, 11(1), 1–12. <https://doi.org/10.1038/s41598-020-79376-x>
- Zurell, D., Franklin, J., König, C., Bouchet, P. J., Dormann, C. F., Elith, J., Fandos, G., Feng, X., Guillerá-Arroita, G., Guisan, A., Lahoz-Monfort, J. J., Leitão, P. J., Park, D. S., Peterson, A. T., Rapacciuolo, G., Schmatz, D. R., Schröder, B., Serra-Diaz, J. M., Thuiller, W., ... Merow, C. (2020). A standard protocol for reporting species distribution models. *Ecography*, 43(9), 1261–1277. <https://doi.org/10.1111/ecog.04960>

BIOSKETCH

This research work is part of the programme “Des Ailes pour les tourbières” (https://www.researchgate.net/publication/342466014_Des_ailes_pour_les_tourbieres_un_programme_partiel_pour_mieux_connaître_Revue_scientifique_Bourgogne-Franche-Comte_Nature_n30) initiated by wildlife managers who performed all the field collection, and is part of the post-doc of S. Sherpa and PhD of C. Kebaïli supervised by L. Després. It combines two strong and internationally recognized expertise in LECA: ecology and genomics applied to the study of the adaptation of natural populations to their environment with major competences in NGS, and SDM applied to conservation biology.

Author contributions: LD and SS designed and coordinated the study, DR performed the molecular work, SS, CK and LD analyzed the molecular datasets, SS, MG and JR performed SDM analyses, SS and LD wrote the manuscript. All authors have read and approved the final manuscript.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Sherpa, S., Kebaïli, C., Rioux, D., Guéguen, M., Renaud, J., & Després, L. (2022). Population decline at distribution margins: Assessing extinction risk in the last glacial relictual but still functional metapopulation of a European butterfly. *Diversity and Distributions*, 28, 271–290. <https://doi.org/10.1111/ddi.13460>