ORIGINAL ARTICLE

MOLECULAR ECOLOGY WILEY

Linking large-scale genetic structure of three Argynnini butterfly species to geography and environment

Daniela Polic[1] | Yeşerin Yıldırım[1,2] | Kyung Min Lee[3,4] | Markus Franzén[3] |
Marko Mutanen[3] | Roger Vila[5] | Anders Forsman[1]

[1]Department of Biology and Environmental Science, Linnaeus University, Kalmar, Sweden

[2]Department of Ecology and Genetics, Uppsala University, Uppsala, Sweden

[3]Ecology and Genetics Research Unit, University of Oulu, Oulu, Finland

[4]Zoology Unit, Finnish Museum of Natural History, University of Helsinki, Helsinki, Finland

[5]Institut de Biologia Evolutiva (CSIC-Universitat Pompeu Fabra), Barcelona, Spain

Correspondence
Daniela Polic, Department of Biology and Environmental Science, Linnaeus University, Kalmar, Sweden.
Email: daniela.polic@lnu.se

Abstract

Understanding which factors and processes are associated with genetic differentiation within and among species remains a major goal in evolutionary biology. To explore differences and similarities in genetic structure and its association with geographical and climatic factors in sympatric sister species, we conducted a large-scale (>32° latitude and >36° longitude) comparative phylogeographical study on three Argynnini butterfly species (*Speyeria aglaja*, *Fabriciana adippe* and *F. niobe*) that have similar life histories, but differ in ecological generalism and dispersal abilities. Analyses of nuclear (ddRAD-sequencing derived SNP markers) and mitochondrial (*COI* sequences) data revealed differences between species in genetic structure and how genetic differentiation was associated with climatic factors (temperature, solar radiation, precipitation, wind speed). Geographical proximity accounted for much of the variation in nuclear and mitochondrial structure and evolutionary relationships in *F. adippe* and *F. niobe*, but only explained the pattern observed in the nuclear data in *S. aglaja*, for which mitonuclear discordance was documented. In all species, Iberian and Balkan individuals formed genetic clusters, suggesting isolation in glacial refugia and limited postglacial expansion. Solar radiation and precipitation were associated with the genetic structure on a regional scale in all species, but the specific combinations of environmental and geographical factors linked to variation within species were unique, pointing to species-specific responses to common environments. Our findings show that the species share similar colonization histories, and that the same ecological factors, such as niche breadth and dispersal capacity, covary with genetic differentiation within these species to some extent, thereby highlighting the importance of comparative phylogeographical studies in sympatric sister species.

KEYWORDS
COI-sequencing, ddRAD-sequencing, evolution, mitonuclear discordance, phylogeography, Pleistocene

---

Daniela Polic and Yeşerin Yıldırım share first authorship and contributed equally to the work.

----

1 | INTRODUCTION

Evolutionary biologists and ecologists have long been studying the biogeographical history (phylogeography) of species to understand how different factors act together in forming genetic patterns of populations (Avise, 2000; Avise et al., 2016; Bermingham & Moritz, 1998; Shikano et al., 2010). In particular, large-scale comparative phylogeography among closely related species with similar yet diverging habitat requirements and overlapping distribution areas can help understand if and how common historical factors, and biogeographical and environmental parameters shape biological diversity (Michaux et al., 2005; Papadopoulou & Knowles, 2016; Victoriano et al., 2008). Similar responses to the same factors would be expected to result in convergent phylogeographical patterns across species, whereas discrepant patterns might reflect species-specific responses (Avise, 2000, Avise et al., 2016, Bermingham & Moritz, 1998, Hickerson et al., 2010, Zink et al., 2001). Incongruence among species could stem from different historical and contemporary reactions to drivers of spatial genetic divergence: geographical boundaries such as landscape barriers (vicariance), spatial isolation (isolation by distance), selective regimes and differences in effective population size (Avise et al., 2016, Crispo et al., 2006, Riddle, 2016, Sobel et al., 2010, Wright, 1943). Ultimately, differential responses to spatial and temporal variation in environmental conditions may be determined by species traits such as dispersal capacity, thermal regulation, body size, physiological tolerance, life history traits, adaptive capacity and plasticity (Bernardo & Spotila, 2006; Kawecki & Ebert, 2004; Paz et al., 2015; Rundle & Nosil, 2005).

The Quaternary climatic oscillations are considered crucial for today's species' distribution and genetic structure (Davis & Shaw, 2001; Hewitt, 2004). During the Pleistocene (2.58 million to 12,000 years ago), many populations have been isolated in different glacial refugia which presumably resulted in allopatric differentiation, intraspecific variation and local adaptation (Ren et al., 2017; Stewart et al., 2010). Depending on how long populations and species have been confined to such refugia, and how fast populations have expanded their ranges after the last glacial maximum (LGM, 26,500–20,000 years ago), gene flow between previously separated populations would have occurred at different rates, thereby differently influencing the contemporary genetic structure of species (Roed et al., 2016; Stewart et al., 2010; Xun et al., 2016).

For ectothermic organisms like butterflies, abiotic factors such as temperature, solar radiation, precipitation and wind speed influence the need for thermoregulation, define selective regimes, and set boundaries on activity and dispersal, thereby potentially affecting the genetic and phenotypic variability of natural populations (Kingsolver & Watt, 1983; Kingsolver & Wiernasz, 1991; Polgar et al., 2013; Zaman et al., 2019). Further, geographical factors such as altitude may differently affect distinct butterfly species. In high-altitude habitats, the combination of relatively short vegetation periods (which translates to short foraging times during the larval stage), low temperatures, high solar radiation, and frequent and sudden changes in weather might lead to adaptive divergence in butterfly populations at different altitudinal clines (Karl et al., 2008). Depending on the species' mobility, high mountain ranges may hinder dispersal and restrict gene flow. In addition to the direct effects that climatic and geographical factors may have on the evolution of species, there is potential for differential and divergent selection owing to spatial variation in community species composition and biotic interactions with host plants, competitors, predators and enemies to influence local adaptations and genetic structure (Magalhaes et al., 2011; Mopper & Strauss, 2013; Pilot et al., 2006; Runquist et al., 2020; Xie et al., 2019). Again, large-scale comparative biogeographical approaches can inform about the relative roles of different abiotic and biotic factors.

Argynnini butterfly species are well suited for analysing and comparing the genetic structure in closely related species with partially overlapping ecological niches and distribution ranges, as they can occur sympatrically across large parts of their shared distribution areas. Here, we use three Argynnini species, *Speyeria aglaja*, *Fabriciana adippe* and *Fabriciana niobe*, as model species. These species belong to the same tribe within the fritillaries (i.e., Nymphalidae butterflies with mostly checkered wing patterns) and share overall morphological similarities. Despite having similar distribution ranges, they differ in their northern range limits, presumably owing to differences in their temperature tolerance (Ellis et al., 2010). *Speyeria aglaja* is distributed from Northern Africa over Europe and Russia to Southeast Asia, reaching as far north as Southern Lapland, whereas *F. adippe*'s northern distribution only extends to central Sweden (Fox et al., 2011; Franzén & Johannesson, 2007; Tolman & Lewington, 2012; Warren, 1995), and *F. niobe* only reaches southern Sweden (Eliasson et al., 2005; Reinhardt & Bolz, 2011; van Swaay & Warren, 1999). Further, these species differ in their degree of ecological generalism, dispersal abilities, habitat preferences and host plant use (Eliasson et al., 2005; Forster & Wohlfahrt, 1955; Fric et al., 2005; Higgins & Riley, 1970; Komonen et al., 2004; Öckinger et al., 2006; Polic et al., 2020; Pollard & Yates, 1994; Zima et al., 2013). *Speyeria aglaja* is a widespread and rather common species that occupies a wide range of habitats and can breed at lower temperatures than the other two fritillaries (Ellis et al., 2010; Zimmermann et al., 2009). By contrast, both *F. adippe* and *F. niobe* are habitat specialists confined to warm microhabitats; *F. adippe* is restricted to woodland-dominated areas, while *F. niobe* is associated with dry habitats and has strongly declined in Europe in recent decades (Ellis et al., 2019; Robertson et al., 1995; Salz & Fartmann, 2009; Spitzer et al., 2009). All three species are univoltine and overwinter as first-instar larvae or as formed larvae within the eggs, which are usually placed directly on or nearby their larval food plants (Eliasson et al., 2005). While the larvae of *F. adippe* and *F. niobe* only feed on *Viola* species, *S. aglaja* larvae have also been found foraging on *Bistorta major* and *Persicaria* spp., depending on their geographical location (Forster & Wohlfahrt, 1955; Fric et al., 2005; Higgins & Riley, 1970; Tolman & Lewington, 2012). According to their dispersal ability and to rankings based on "mobility indices" as assessed by butterfly experts, all three species are considered as intermediately mobile (Komonen et al., 2004; Öckinger et al., 2006; Pollard & Yates, 1994; Zima

et al., 2013), although *F. adippe* has been reported to be more likely to engage in long-distance dispersal of several kilometres than *S. aglaja* (Polic et al., 2020). These interspecific divergences could lead to different large-scale genetic patterns among the species, despite them being phylogenetically close and exposed to similar environmental conditions and selection regimes.

To quantify and compare population genetic structures among closely related species and gain insights into the underlying evolutionary drivers and historical changes, different molecular markers can be used. Traditionally, population genetic studies of non-model organisms relied on relatively limited numbers of neutral nuclear DNA (nuDNA) markers, mostly microsatellites, or on mitochondrial DNA (mtDNA) sequences, such as the mitochondrial cytochrome *c* oxidase I (*COI*) gene (Guichoux et al., 2011; Hajibabaei et al., 2007; Harrison, 1989). Thanks to the rapid development of high-throughput technologies, approaches such as double-digest restriction-site associated DNA sequencing (ddRAD-sequencing) generate thousands of mainly nuclear SNP (single nucleotide polymorphism) loci randomly distributed across the genome, allowing for genotyping of even non-model organisms (Andrews et al., 2016; Peterson et al., 2012). While nuclear loci are inherited biparentally, mtDNA is haploid, nonrecombinant and only inherited maternally (Lansman et al., 1983), resulting in a four-fold (assuming a balanced sex ratio) smaller effective population size of the mitochondrial genome. Further, mutation rates are higher for mtDNA than nuDNA in most organisms (Allio et al., 2017; Ballard & Whitlock, 2004). These intrinsic characteristics could lead to mitonuclear discordance

following incomplete lineage sorting, sex-biased dispersal, introgression, horizontal gene transfer, selective sweeps and demographic expansion (Despres, 2019; Gompert et al., 2008; Hinojosa et al., 2019; Soucy et al., 2015; Yıldırım et al., 2018). Some of these effects can be mediated by *Wolbachia* bacteria as infection by them has been documented to be one of the most frequent causes of mitonuclear discordance in insects (Toews & Brelsford, 2012). These maternally inherited bacteria may induce cytoplasmic incompatibility and male-killing, potentially resulting in selective sweeps favouring the mitochondrial genome of the *Wolbachia*-infected lineage (Hurst & Jiggins, 2005; Ritter et al., 2013; Werren et al., 2008). As different markers might tell different stories about the genetic structuring of species, incorporating more than one marker type into phylogeographical studies can provide a more comprehensive picture of the history and variation of contemporary genetic patterns.

Here, we studied 148 individuals belonging to three Argynnini species (i.e., *S. aglaja*, *F. adippe* and *F. niobe*) from various geographical locations covering large parts of Europe (and Morocco for *S. aglaja*, Figure 1) to analyse their genetic structure using nuclear SNP markers derived from ddRAD-sequencing and mtDNA *COI* sequences. To our knowledge, this is the first genome-wide study that addresses large-scale comparative phylogeographical patterns in butterflies using a genome-wide nuDNA approach in addition to mtDNA. Specifically, we aim to answer the following questions: (i) Are the large-scale population genetic structures of the studied species similar? (ii) Are the genetic structures of the studied species associated with the same or different environmental and geographical
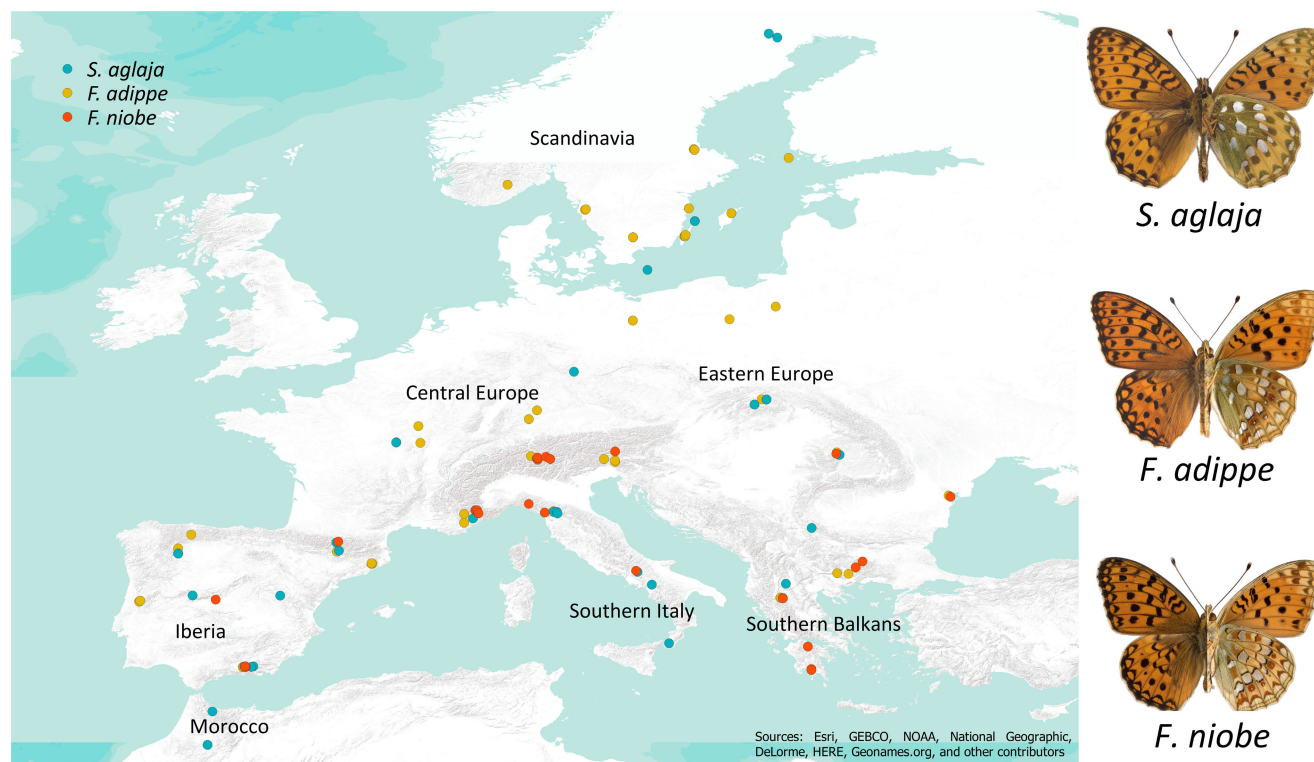
FIGURE 1 Map of sampling locations of *Speyeria aglaja* (turquoise), *Fabriciana adippe* (yellow) and *Fabriciana niobe* (red). The right side of the butterflies shows the dorsal wings and the left side the ventral wings

variables, and can any differences be accounted for by species characteristics? (iii) Do the respective marker types (ddRADseq, *COI*) generate different large-scale genetic patterns? In this study, we highlight the importance of studying spatial genetic variation in multiple species simultaneously to understand how even sister species may respond differently to common biotic and abiotic factors, possibly due to their unique history and ecology.

2 | MATERIAL AND METHODS

2.1 | Study area and data collection

We studied 148 individuals belonging to three Argynnini species, 43 *Speyeria aglaja*, 74 *Fabriciana adippe* and 31 *F. niobe* individuals, selected from butterfly collections at the Institut de Biologia Evolutiva, Universitat Pompeu Fabra in Barcelona and the Department of Biology and Environmental Science, Linnaeus University in Kalmar. Samples were chosen from various locations (one to three individuals per sampling location for 22 sampling locations in *S. aglaja*, one to nine individuals per sampling location for 26 sampling locations in *F. adippe*, and one to seven individuals per sampling location for 13 sampling locations in *F. niobe*), so that large parts of their European (and Moroccan for *S. aglaja*) distribution areas were covered (33.542–65.856 degrees latitude, and −8.227 to 28.689 degrees longitude) and that individuals in all species stemmed from a wide range of locations (Figure 1; Table S1). All Moroccan samples belonged to *S. aglaja*, as the other two species do not occur in the Maghreb. The butterflies were collected between 2008 and 2017. After capture, the specimens were either directly transferred to 95–99% ethanol or dried, and stored at −20°C upon arrival to the laboratory until DNA extraction. The coordinates for each sampling location were recorded directly in the field.

2.2 | Environmental data

Climate data for each sample location, namely information on temperature, solar radiation, precipitation and wind speed, were attained from WorldClim – Global Climate Data (Fick & Hijmans, 2017). We used the R package *raster* 3.5 (Hijmans et al., 2021) to import the WorldClim data as raster layers, and the function *extract* to add the environmental values to each butterfly sampling point. We extracted information on altitude for each sampling location from raster data obtained from the U.S. Geological Survey Earth explorer using QGIS 2.18.25 (QGIS.org, 2022). Information about the environmental and geographical data for each sample location can be found in Table S1.

2.3 | Library preparation and ddRAD-sequencing

For DNA extraction, thorax tissue from each sample was ground with stainless steel beads (Next Advance) in a bullet blender, and

an E.Z.N.A. Genomic DNA Isolation Kit (OMEGA Bio-Tek) was used for further processing. The quality of genomic DNA (gDNA) was estimated by performing agarose gel electrophoresis to identify signs of DNA degradation. DNA samples of relatively low quality were subjected to whole genome amplification with a REPLI-g Mini Kit (Qiagen) prior to the ddRADseq library preparation to obtain sufficient gDNA quantity and quality. We prepared ddRADseq libraries at the molecular laboratory at the Ecology and Genetics Research Unit, University of Oulu, Finland, following the protocol in Lee et al. (2018) and Peterson et al. (2012) with a few exceptions. A total of 500 ng of DNA for each sample and *Pst*I and *Mse*I restriction enzymes (New England Biolabs) were used, a purification step with AMPure XP magnetic beads (Agencourt) was carried out in a Biomek automated liquid handler (Beckman Coulter), and the size distribution and concentration of the pools were measured with a Bioanalyzer (Agilent Technologies). The individuals used in this study were a subset of 285 Argynnini samples, which were pooled into three libraries (95 samples per pool), and sequenced on an Illumina HiSeq 2500 PE 101 v4 machine at the Institute for Molecular Medicine Finland (FIMM, Helsinki). The demultiplexed fastq data are archived in the NCBI SRA under the BioProject ID PRJNA70744.

2.4 | Quality control and data trimming

We demultiplexed the raw data and performed quality control using the *process_radtags* unit in STACKS version 2.2 (Catchen et al., 2011; Catchen et al., 2013) with default settings and the following additional specifications: *-c* (reads with an uncalled base were discarded), *-q* (reads were discarded when the raw Phred quality score dropped below 10 within a sliding window spanning 15% of the read length), and *-r* (reads were rescued in case of one nucleotide mismatch at the enzyme cut-site or individual barcodes). The filtering process retained 90% of the raw data. We used TRIMMOMATIC 0.36 (Bolger et al., 2014) to remove adapters and to trim all the reads to 92 bp.

2.5 | Building the catalogue of loci

STACKS' de novo pipeline (Catchen et al., 2011; Catchen et al., 2013) was used for assembly of ddRADtag loci and SNP calling by running each step of the pipeline (*ustacks, cstacks, sstacks, tsv2bam* and *gstacks*) separately. The parameters used for further processing in the STACKS pipeline were determined after parameter optimization on a randomly selected subset of individuals covering a wide geographical range (16 *S. aglaja*, 16 *F. adippe*, 12 *F. niobe* and 16 individuals across all three species), as described in Paris et al. (2017), and Rochette and Catchen (2017). Parameter *m* is the minimum number of raw reads to form a stack or putative allele within an individual, *M* is the number of mismatches tolerated between stacks to form a putative locus and *n* is the number of mismatches allowed between individual loci across samples to

build a catalogue of loci across individuals. We decided to use the same parameter setup for all four data sets, that is $m = 4$, $M = 5$ and $n = 6$ (Figures S1 and S2). The coverage at $m4$ was around 80× for all data sets (Figure S1). We ran the de novo pipeline on 47 *S. aglaja*, 84 *F. adippe*, 36 *F. niobe* and 167 samples across all three species using these parameters and otherwise default settings, which resulted in catalogues containing 395,548, 221,181, 420,316 and 931,511 loci, respectively. The catalogues were further processed with the *populations* unit in STACKS. For all data sets, we chose the *r80* setup, where a locus had to be present in at least 80% of individuals in order to be retained. This rather conservative approach was used, as the number of retained loci was still relatively high, while we can be confident to avoid spurious associations between the presence of a locus and genetic distance. For all data sets, we chose one random SNP per locus, a minimum minor allele frequency of 0.05 and a maximum observed heterozygosity of 0.7. After *populations*, samples with more than 25% of missing data were excluded, which resulted in 43 *S. aglaja*, 74 *F. adippe*, 31 *F. niobe* and 148 individuals across all three species. Then, *populations* was repeated with the reduced data sets. The final data sets consisted of 8627, 7503, 9492 and 2022 SNPs for *S. aglaja*, *F. adippe*, *F. niobe* and all three species pooled, respectively.

Since our analyses showed a strong genetic differentiation of Iberian and/or Moroccan populations in *S. aglaja* and *F. adippe* (the latter is only valid for *S. aglaja*, as the other species do not occur in the Maghreb, see details in the Results), potential finer structuring between populations in the rest of Europe might have been concealed. We therefore repeated *populations* in STACKS with *r80* excluding individuals from Iberia (and Morocco), and ran the downstream analyses for the reduced data sets, which consisted of 30 (*S. aglaja*) and 59 (*F. adippe*) individuals.

To detect and eliminate potential mitochondrial sequences from ddRADseq data sets, the sequences obtained after *populations* in STACKS were blasted against three mitochondrial Lepidoptera genomes: *Melitaea cinxia* (GenBank ID: CM002851, Ahola et al., 2014), *Zerene cesonia* (GenBank: CM022631, Rodriguez-Caro et al., 2020) and *Heliconius melpomene rosina* (GenBank: KP100653, Meng et al., 2016). Three reads from *F. adippe* and one read from *F. niobe* sequences mapped to the mitochondrial genomes, and all these loci were removed from the data sets.

2.6 | Assessing potential *Wolbachia* infections

The resulting sequences obtained after *populations* in STACKS were mapped to the *Wolbachia pipientis* genome (GenBank: NZ_JQAM01000001) using Geneious mapper in GENEIOUS 11.1.05 (Biomatters) to evaluate the risk of contamination by the bacterial parasite *Wolbachia*, which could influence the genetic structure of species and thus the outcome of the population genetic analyses (Gompert et al., 2008; Werren et al., 2008). Only one sequence for *F. niobe* mapped to the *Wolbachia* genome, and it was excluded from the data set for the analyses. This sequence was also blasted against *Wolbachia* genomes in NCBI's nucleotide database, and mapped to the genome.

2.7 | Sequencing of the mitochondrial *COI* gene

The barcode region of mitochondrial cytochrome *c* oxidase 1 (*COI*) was amplified using the primers LepF1 and LepR1 (Cooper et al., 2007), and the PCR (polymerase chain reaction) amplification conditions as stated in Dincă et al. (2019). Sanger sequencing of the amplicons was carried out by Eurofins Genomics (Eurofins Genomics Europe Shared Services). Six samples with sequencing quality issues were not included in the *COI* analyses (Table 1). After aligning and trimming the *COI* reads in GENEIOUS, sequences with a length of 601 bp of the barcode region were used for further analyses. The sequences were submitted to the NCBI nucleotide database (Accession codes: MZ026157–MZ026168, MZ026170, MZ026171, MZ026174–MZ026200, MZ026202, MZ026242, MZ026245–MZ026261, MZ026263–MZ026290). Sequencing data for 20 of the samples were retrieved from the NCBI nucleotide database (see Table S1 for accession codes).

2.8 | Data analysis

2.8.1 | Genetic structure

We examined patterns of genetic differentiation in the study organisms by using multiple approaches with different algorithms for the purpose of comparing results from differing analyses and thereby obtaining a more comprehensive picture of the genetic structure.

TABLE 1 Diversity indices for each species according to nuclear data obtained through ddRADseq and mitochondrial data obtained through *COI* sequencing

		ddRADseq	COI					
	Number of samples (ddRADseq/mtDNA)	Number of SNP sites	Size (bp)	Number of haplotypes	Nucleotide diversity	Haplotype diversity	*k*	Segregating sites
S. aglaja	43/40	8627	601	10	0.005	0.835	3.29	16
F. adippe	74/73	7503	601	17	0.009	0.881	5.38	23
F. niobe	31/30	9492	601	7	0.005	0.823	3.03	9

Note: *k*: average number of nucleotide differences.

The program FASTSTRUCTURE 1.0 (Raj et al., 2014) was used as a Bayesian approach to determine the most likely number of genetic clusters (*K*) by running the analysis for each number of *K* (1–22, 1–26 and 1–13 for *S. aglaja*, *F. adippe* and *F. niobe*, respectively; based on the number of sampling locations) with 10 iterations using a simple prior and five-fold cross-validation. The most likely number of *K* was determined using the FASTSTRUCTURE script *chooseK.py* based on two criteria: "model complexity that maximizes marginal likelihood" (henceforth marginal likelihood), and "model components used to explain structure in data" (henceforth model complexity). The plots were created using the FASTSTRUCTURE script *distruct.py* for the chosen values of *K* (Raj et al., 2014). Next, FINERADSTRUCTURE (Malinsky et al., 2018) was used as a formal clustering analysis to evaluate shared ancestry among individuals in each species. This program uses Bayesian haplotype-based statistics to group individuals with high levels of shared co-ancestry. Using the R scripts fineRADstructurePlot.R and FinestructureLibrary.R that come with the software, we visualized the shared co-ancestry between individuals in a heatmap. We also assessed the genetic structure in each species visually via principal component analysis (PCA) of the ddRADseq data, using a distance matrix. Pairwise distances between the individuals were based on the proportion of shared alleles. PCA was performed using the *prcomp* function in RSTUDIO 3.6.1 (R Core Team, 2020). To statistically test for an association of genetic distance with geographical distance, we performed linear regressions for each species separately in RSTUDIO. First, we calculated the geographical distance between all points using the R package *Imap* 1.32 (Wallace, 2012), converted the distance matrix as used for PCA into a Manhattan dissimilarity matrix, and then regressed the individual geographical distances against the genetic distances using the *lm* function.

A colonization pattern towards northern regions following the LGM is expected to lead to a lower genetic diversity in northern regions compared to southern regions (Adams & Hadly, 2013; Hughes & Hughes, 2007; Pereira, 2016). We performed linear regressions for each species separately in RSTUDIO to test for associations between individual genetic diversity (expected heterozygosity [H_E]) and latitude using the SNP data derived from ddRADseq. Expected heterozygosity H_E was calculated in ARLEQUIN 3.5 (Excoffier et al., 2005).

2.8.2 | Phylogenetic analyses

To assess phylogenetic relationships between individuals from different sampling locations for each species separately and all three species pooled based on concatenated ddRADseq data, we used the maximum-likelihood (ML) tree inference in IQ-TREE 2.1.2 (Nguyen et al., 2015) and branch support bootstrap values were calculated with ultrafast bootstrap (UFBoot, 1000 bootstraps), as implemented in IQ-TREE (Hoang et al., 2017). Prior to analyses, the best fitting mutational model was obtained through IQ-TREE's MODELFINDER (Kalyaanamoorthy et al., 2017) based on Bayesian information criterion (BIC; see Table S2 for all models and corresponding BIC values). The chosen best fitting model was TVMe+G4 (Transversion model,

AG = CT, equal base frequency, discrete Gamma model with default 4 rate categories) for *S. aglaja* (BIC = 223,146.3), PMB+F+G4 (Probability Matrix from Blocks, empirical base frequencies, discrete Gamma model with default 4 rate categories) for *F. adippe* (BIC = 1,506,379.7) and for *F. niobe* (BIC = 197,869.5), and for all three species together (BIC = 78,144.5), the best fitting model was TVM+F+G4 (Transversion model, AG = CT, unequal base frequency, empirical base frequencies, discrete Gamma model with default 4 rate categories). We visualized the phylogenetic trees with Interactive Tree Of Life (ITOL) 6 (Letunic & Bork, 2019).

Genetic diversity indices for *COI* sequences were calculated in DNASP 6 (Rozas et al., 2017). The evolutionary relationships between the haplotypes were assessed by haplotype networks using the Median Joining Network method (Bandelt et al., 1999), as implemented in POPART 1.7 (Leigh & Bryant, 2015). The phylogenetic analyses for the three fritillary species based on *COI* sequences were also performed in IQ-TREE. We only included the unique haplotypes in the analysis, and used *Boloria pales* (GenBank ID: HM901782.1, Dapporto et al., 2019) as an outgroup for the *COI* tree. We included *COI* sequences from *Speyeria zerene* (GenBank ID: HQ561199.1) and *Argynnis paphia* (GenBank ID: MN145295.1) to show the relationship between the study species at a higher resolution. MODELFINDER, which was used to estimate the best-fitting model prior to construction of the tree, suggested TIM2+F+R2 (AC = AT, CG = GT, unequal base frequency, empirical base frequencies, FreeRate model with 2 categories; Soubrier et al., 2012; Yang, 1995) as the best-fitting model based on BIC (=4040.6); see Table S2 for all models and corresponding BIC values. The consensus phylogenetic trees were visualized with ITOL. We calculated divergence percentages between *COI* haplotypes in GENEIOUS based on nucleotide substitutions.

2.8.3 | Associations of the genetic structure with geography and environmental factors

Potential associations of the genetic structure and environmental and spatial factors—mean annual temperature, mean annual solar radiation, mean summer wind speed, mean annual precipitation, latitude, longitude and altitude—were investigated using distance-based redundancy analysis (db-RDA), a constrained version of a PCA (Legendre & Anderson, 1999). We chose mean annual temperature, mean annual solar radiation and mean annual precipitation, as these variables influence the vegetation periods in temperate climates (Piao et al., 2006), thereby potentially affecting the studied butterfly species' host plants growth and distribution, as well as the availability of nectar plants. This in turn may affect body size in butterflies, as longer vegetation periods might prolong the larval food stage, thereby increasing the body size. Body size itself is often used as a proxy for dispersal capacity in insects, which can have implications for gene flow (Kuussaari et al., 2014; Lenormand, 2002). Similarly, wind speed during the butterflies' flight period (i.e., summer) may influence wing architecture and size, thus impacting mobility. We included altitude because climatic variables change with variations in

this factor (e.g., higher solar radiation at higher altitudes, but same latitudes). The distance matrix used for PCA (see above) was converted into a Manhattan dissimilarity matrix and used as input to a principal coordinates analysis (PCoA). We used the *capscale* function in the R package *vegan* 2.5 (Oksanen et al., 2019) to build db-RDA models in order to assess which variables contribute the most to the PCoA ordination. When building the models, we first checked for multicollinearity between variables by calculating the variance inflation factors (VIFs), and excluded variables with a VIF > 10 from the model, as such high values are indicative of a problematic level of collinearity (James et al., 2013). The overall statistical significance of the model and the predictor variables was determined with the function *anova.cca* using 999 permutations. Nonsignificant variables were removed stepwise, and adjusted R^2 was calculated for each model, which was then used as the basis for model selection. The *p*-values were adjusted for multiple testing by using Holm's correction (Holm, 1979). To account for spatial autocorrelation between samples, we used longitude and latitude as a condition in the model, thereby partialling out these terms from the test before analysing the other variables. As the environmental and geographical variables used in this study were heterogeneous in dimension, we standardized them via z-scoring, that is, for each variable, the mean value was subtracted from the observed value and then divided by the standard deviation of all observed values.

3 | RESULTS

3.1 | Genetic structure

For *Speyeria aglaja*, the most likely number of genetic clusters determined with FASTSTRUCTURE was either three (marginal likelihood) or four (model complexity); that is, Morocco and Southern Spain constituted clear distinct clusters, whereas the rest of the individuals seemed to present one cluster (*K* = 3, Figure 2), or a vague north–south structure (*K* = 4, Figure S3). For *Fabriciana adippe*, FASTSTRUCTURE identified two distinct genetic clusters (*K* = 2 with both marginal likelihood and model complexity) with individuals from Iberia grouping together and the rest of the European samples forming a second cluster. Individuals from Catalonia were intermediate and about half of their analysed sequences were attributed to each cluster, constituting a transition from the Iberian Peninsula to the rest of Europe (Figure 2). For *Fabriciana niobe*, FASTSTRUCTURE identified either one (model complexity) or three (marginal likelihood) clusters; that is, samples from Eastern Europe and the Southern Balkans formed one cluster, and the rest of the samples composed two clusters with a somewhat east–west structuring (*K* = 3, Figure 2).

The FINERADSTRUCTURE results were largely in line with FASTSTRUCTURE (Figure 2): in *S. aglaja*, the highest levels of shared co-ancestry were revealed between individuals from Morocco, followed by a group of Iberian individuals and some of the Southern Balkan individuals, while in *F. adippe*, individuals from Iberia constituted the most distinct group. In *F. niobe*, most of the Iberian samples formed the

furthest diverged cluster, followed by the Southern Balkan individuals. Lastly, central European *F. niobe* individuals showed comparably high levels of shared co-ancestry.

Prinicipal component analyses visually supported the FASTSTRUCTURE and FINERADSTRUCTURE results. However, PCA suggested a finer structure than FASTSTRUCTURE (Figure 3). In *S. aglaja*, Southern Iberian and Moroccan samples formed two clusters, which were well separated from each other and from the rest of the samples. When excluding these individuals from the analyses, a finer structure in the remaining samples was revealed, with Scandinavian, Eastern European and central European individuals, respectively, loosely grouping together (Figure 3). In *F. adippe*, PCA illustrated the main axis of variation between samples from Iberia and the rest of the locations. Individuals from Catalonia were recovered at an intermediate position from the other two. Non-Iberian individuals visually clustered into three vague groups, namely Scandinavia, Central Europe, and Southern Italy together with Southern Balkans; some Eastern European samples showed overlaps between the clusters. These loose clusters became more obvious when individuals from Iberia were excluded from the analysis (Figure 3). The strongest differentiation in *F. niobe* was visualized between individuals from the Southern Balkans and the rest. Central Europe and Iberian samples also formed two separate clusters, but the divergence of the latter from the remaining individuals was not as strong as in *F. adippe* (Figure 3).

There was a significant and positive association of geographical and genetic distance in *S. aglaja* (regression, $R^2 = 0.13$, $F_{1,901} = 131.9$; effect of geographical distance, $b = 1.008e^{-08}$, $p < .001$), *F. adippe* (regression, $R^2 = 0.44$, $F_{1,2699} = 2119$; effect of geographical distance, $b = 3.591e^{-08}$, $p < .001$) and *F. niobe* (regression, $R^2 = 0.48$, $F_{1,463} = 420.3$; effect of geographical distance, $b = 2.129e^{-08}$, $p < .001$, Figure 3). Heterozygosity decreased significantly towards the North in *S. aglaja* (regression, $R^2 = 0.09$, $F_{1,41} = 4.16$; effect of latitude, $b = -5.723e^{-05}$, $p = .05$, Figure 3). The trend of a decrease in diversity towards higher latitudes was also observed in *F. adippe* (regression, $R^2 = 0.02$, $F_{1,72} = 1.64$; effect of latitude, $b = -3.542e^{-05}$, $p = .21$, Figure 3) and *F. niobe* (regression, $R^2 = 0.04$, $F_{1,29} = 1.26$; effect of latitude, $b = -7.395e^{-05}$, $p = .3$, Figure 3), although it was not significant.

3.2 | Phylogenetic analyses

In all data sets obtained by ddRADseq, the phylogenetic (ML) trees revealed similarities in the genetic patterns with FASTSTRUCTURE, FINERADSTRUCTURE and PCA. Phylogenetic analyses of nuclear data retrieved through ddRADseq could separate the three species from each other with high bootstrap support (100%), with *S. aglaja* forming a sister lineage to *F. adippe* and *F. niobe*. The trees of each individual species were in line with the PCA plots, forming clades associated with geographical proximity (Figure 4).

Corroborating the results from the phylogenetic analyses of nuclear data obtained by ddRADseq, both the haplotype network
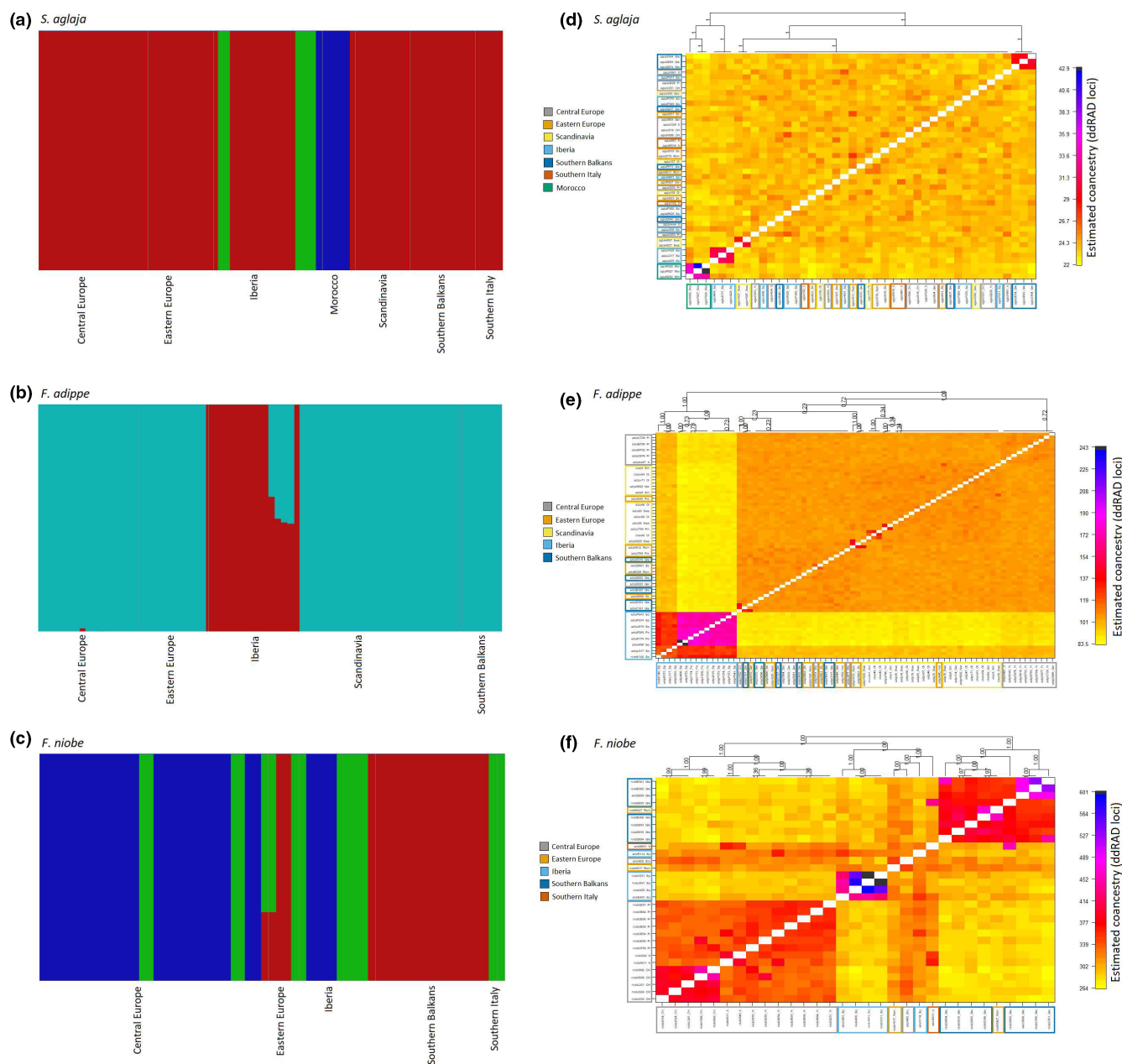
FIGURE 2 (a–c) Genetic clusters according to FASTSTRUCTURE analyses of nuclear data obtained through ddRADseq with the chosen number of *K* for each species (*Speyeria aglaja*; *K* = 3, *Fabriciana adippe*: *K* = 2, *Fabriciana niobe*: *K* = 3), where each bar represents one individual and each colour a genetic cluster. In (a), the green bars represent individuals from Southern Iberia. (d–f) Co-ancestry matrices for each species according to FINERADSTRUCTURE analyses based on genetic similarity. Higher levels of co-ancestry (darker colours) imply less divergence between two individuals and frame colours of individuals show the sampling region

(Figure 4) and the phylogenetic tree (Figure S4) constructed from the *COI* sequences showed that *S. aglaja* is a sister to the two *Fabriciana* species. The divergence between the *Fabriciana* species was 2.39–3.49%, whereas it was 5.16–6.82% between *S. aglaja* and the two *Fabriciana* species. Many individuals shared common haplotypes for the *COI* gene in all three species. The number of unique haplotypes was 10 for *S. aglaja*, 18 for *F. adippe* and seven for *F. niobe* (Figure 4). Structuring based on geographical regions could be observed, but these structures were not parallel among the three species. Haplotype diversity was richer in *F. adippe* than in the other two

species (Table 1). Further, there was some incongruence between data obtained via mtDNA *COI* sequencing and nuclear data obtained via ddRADseq, especially in *S. aglaja*.

According to the network analysis in *S. aglaja*, the *COI* data lacked a clear geographical structure; almost every geographical region had more than one haplotype with a divergence of 1–5 bp between them, and the distribution of the haplotypes among the regions was patchy (Figure 4). The most common haplotype (Ag5) was shared among several regions, namely Southern Balkans, Central Europe and Eastern Europe. According to both the network
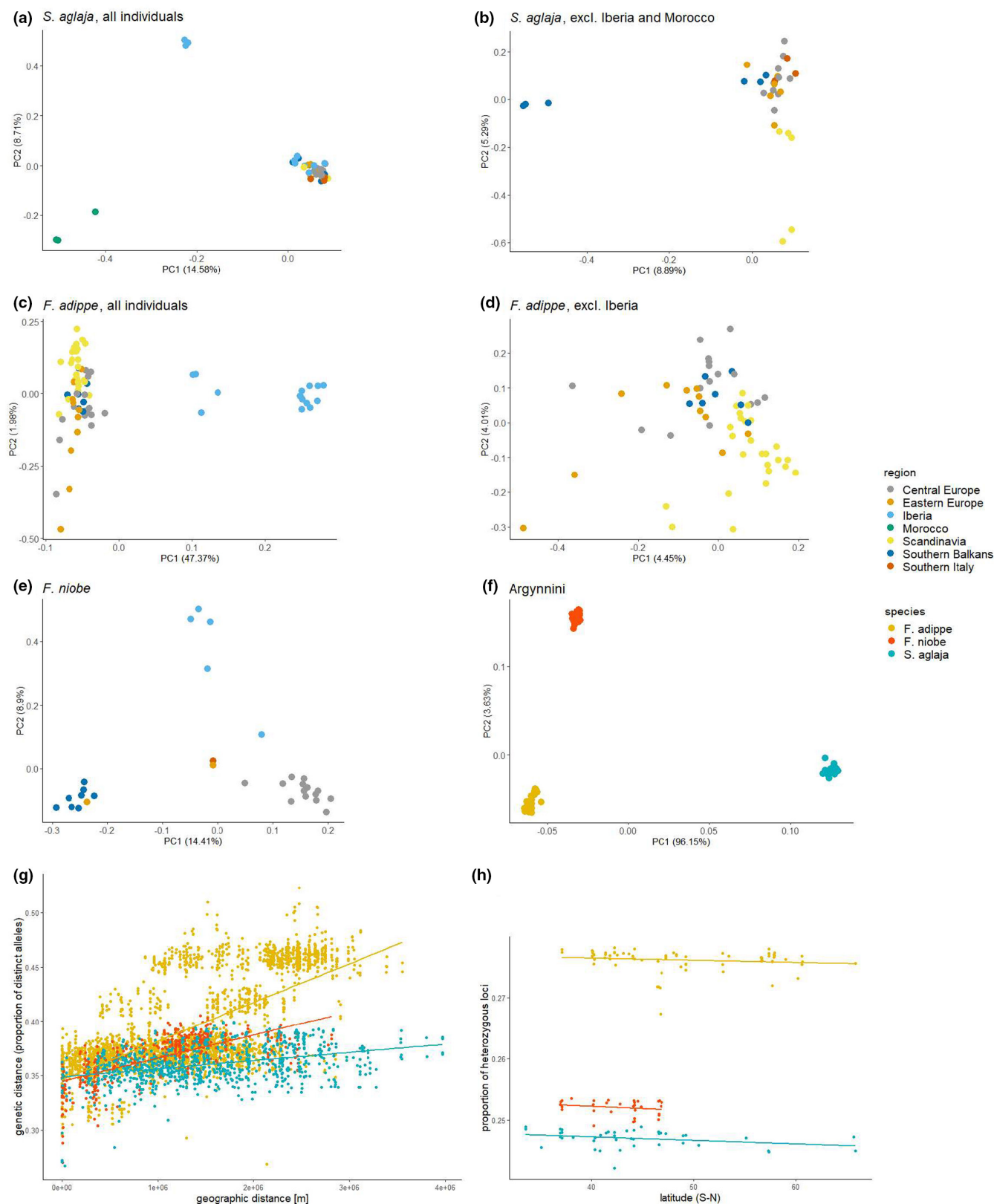
FIGURE 3 (a–f) Biplots showing the genetic clustering according to PCA of nuclear data obtained by ddRADseq for the three Argynnini species. The colours represent geographical regions, or species identity, respectively (f). For *Speyeria aglaja* (a, b) and *Fabriciana adippe* (c, d), PCAs were performed for the data sets including all individuals (a, c), and the data sets excluding individuals from Iberia (and Morocco for *S. aglaja*) (b, d). (g) Scatterplot showing the significant relationship of pairwise geographical distance and genetic distance based on the proportion of distinct alleles. (h) Scatterplot showing a decrease in heterozygosity H_E from southern towards northern sampling regions (only significant in *S. aglaja*)
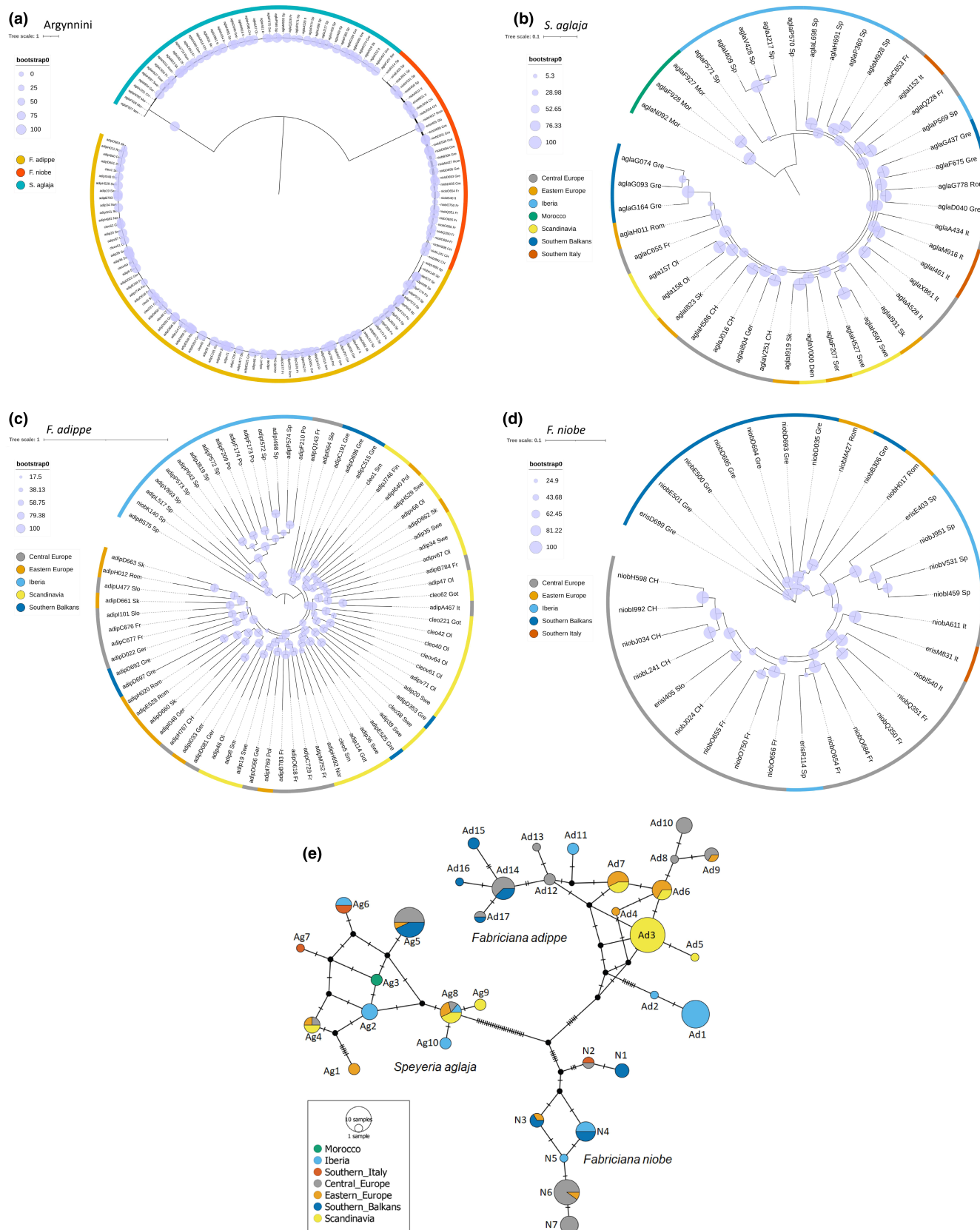
FIGURE 4 (a–d) Maximum likelihood trees according to nuclear data obtained by ddRADseq for (a) all three Argynnini species together (rooted at the base of *Speyeria aglaja*), (b) *S. aglaja*, (c) *Fabriciana adippe* and (d) *Fabriciana niobe*. Trees (b–d) were rooted at the furthest diverged geographical group in each respective species according to PCA. (e) Median Joining Haplotype network for three Argynnini species according to mtDNA *COI* data. The coloration is based on the sampling regions. Black nodes represent the unsampled haplotypes connecting the present haplotypes, and the hashes indicate the mutational steps between the haplotypes. Each distinct haplotype was given a unique ID

TABLE 2 Geographical and environmental variables associated with the genetic structure in the three studied fritillary species according to db-RDA of nuclear data obtained through ddRADseq

	S. aglaja		F. adippe		
	Whole data set	Excl. Iberia and Morocco	Whole data set	Excl. Iberia	F. niobe
Longitude	—	—	—	—	—
Latitude	***	—	—	—	***
Altitude	—	—	***	—	—
Solar radiation	—	***	—	***	*
Temperature	*	—	*	—	—
Precipitation	**	—	*	—	*
Wind speed	—	—	**	**	—

*** $p < .001$, ** $p < .01$, * $p < .05$.

and the ML tree (Figure S4), one common haplotype shared by several regions (Ag8; Iberia, Eastern Europe, Central Europe and Scandinavia) gave rise to the other haplotypes. The majority of the haplotypes were closely related to each other (1–4 bp), and similar to the *F. adippe* network, the relationship between the haplotypes was complex with ambiguities in the evolutionary relationships among them. We found high incongruence with the nuclear data retrieved from ddRADseq insofar as the Moroccan (Ag3) and some Iberian samples (Ag2) were not distant from the rest. Instead, one Eastern European haplotype shared between some Romanian and Slovakian samples (Ag1) formed the most distant lineage (6 bp from the closest haplotypes).

In *F. adippe*, the outcome of *COI* analyses was similar to that of the nuclear data from ddRADseq analyses. Iberian individuals (except one Catalan individual; haplotypes Ad1 and Ad2, Figure 4) were differentiated from the rest of the regions by 8 bp according to the *COI* haplotype network. The divergence between these haplotypes and the rest varied between 1.64% and 2.49%. Apart from the Iberian cluster, two main groups were observed: Scandinavian, Eastern European and some Central European samples formed one clade (Ad3–Ad10), while Southern Balkan, Southern Italian and the rest of the Central European samples formed another cluster (Ad14–Ad17). The haplotypes connecting the clusters were almost fully connected to each other with high ambiguity in the order of mutational changes. Overall, the geographical proximity determined the population structure. The ML tree supported the network by separating the Iberian samples with high bootstrap values (95%) (Figure S4). However, the resolution of the evolutionary relationship among the other haplotypes was mostly backed by only ~50% bootstrap support. This was in parallel with the haplotype network, pointing towards ambiguity in the evolutionary relationship among the present haplotypes. The relationship between the geographical regions was mostly congruent with the nuclear data (Figures 3 and 4).

Despite some similarities, a clear geographical separation as nuclear data obtained from ddRADseq suggested was not observed for *F. niobe* haplotypes, yet haplotypes clustered based on geographical proximity (Figure 4). The haplotypes were closely related and showed a vague latitudinal gradient. Some Southern Balkan and Central European haplotypes (Peloponnesian, and the majority of

Italian samples, haplotypes N1 and N2) formed a distinct cluster, which was 4–5 bp distant from the rest of the haplotypes. The second distinct cluster included the haplotypes from Eastern European, Central European and Iberian samples (N3–N7). Strong ambiguity was not observed between the *F. niobe* haplotypes data. For *F. niobe*, the ML tree supported the network, and the separation of the two clusters had relatively high bootstrap support (84%). Further separation between the haplotypes had intermediate to high support (62–99%).

3.3 | Associations of the genetic structure with geographical and environmental factors

A different set of environmental variables was associated with the genetic structure in the three studied species, following db-RDA of nuclear data obtained by ddRADseq (Table 2). In *S. aglaja* (whole data set), we found a significant association of latitude, mean annual temperature and mean annual precipitation with the genetic structure. The genetic structure in *F. adippe* (whole data set) was associated with altitude, mean annual temperature, mean annual precipitation and mean summer wind speed. *F. niobe* individuals clustered together according to latitude, mean annual solar radiation and mean annual precipitation.

In *S. aglaja*, the final selected model for the whole data set (all *S. aglaja* individuals) included mean annual temperature, mean annual precipitation, mean summer wind speed, latitude and longitude as explanatory variables (adjusted $R^2 = 0.04$). Latitude ($F_{1,37} = 1.86$, $p < .001$), mean annual temperature ($F_{1,37} = 1.31$, $p = .02$) and mean annual precipitation ($F_{1,37} = 1.38$, $p < .01$) were significantly associated with the genetic structure in *S. aglaja*, when analysing all individuals together. Examination of the biplot reveals that the separation of the Moroccan samples from the rest of the individuals is mainly linked to temperature and precipitation (Figure 5). When we excluded individuals from Iberia and Morocco, the selected model included the variables mean annual solar radiation, mean annual temperature, mean annual precipitation, mean summer wind speed and longitude (adjusted $R^2 = 0.03$). There was a significant association of mean annual solar radiation ($F_{1,24} = 1.55$, $p < .001$) with the

(a) *S. aglaja*, all individuals
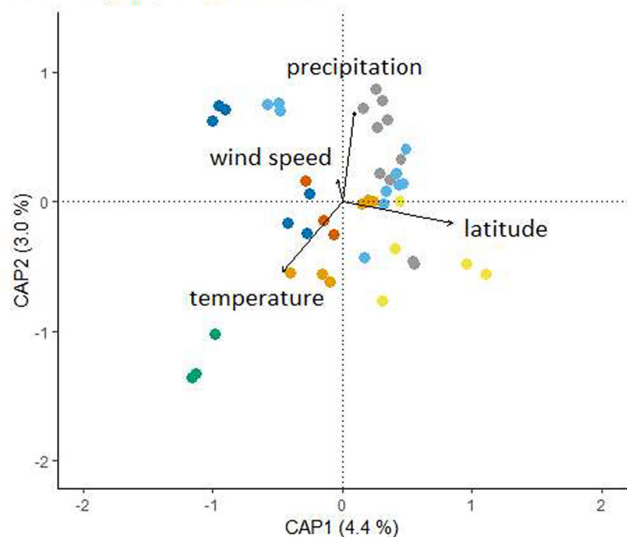
(b) *S. aglaja*, excl. Iberia and Morocco
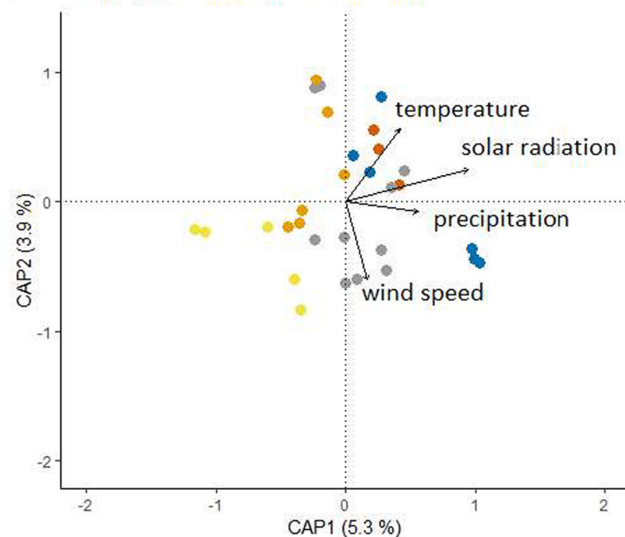
(c) *F. adippe*, all individuals
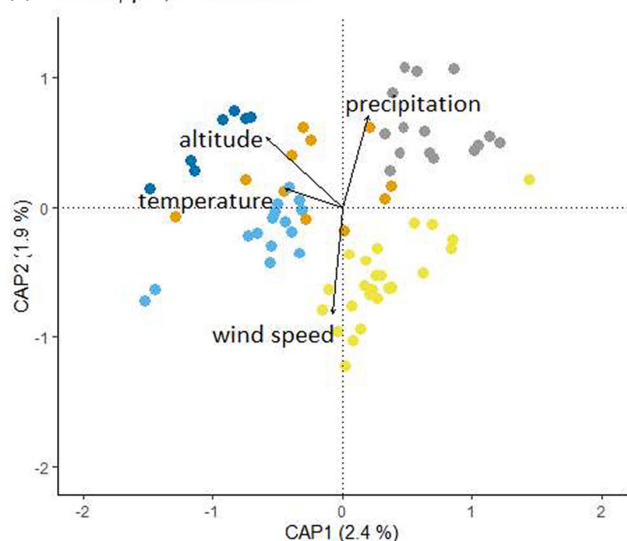
(d) *F. adippe*, excl. Iberia
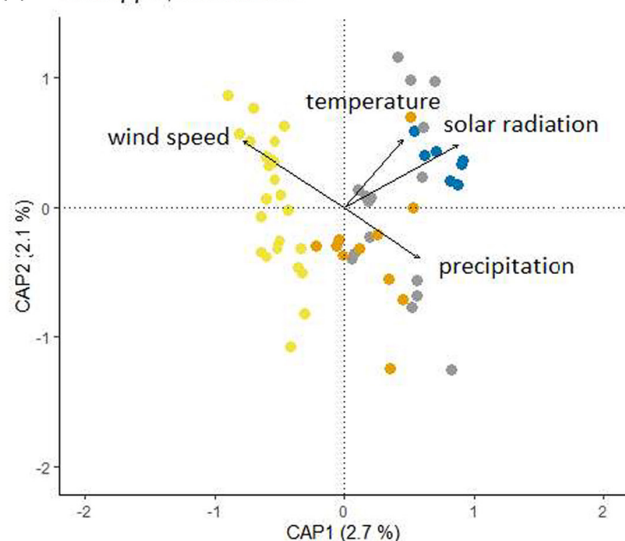
(e) *F. niobe*
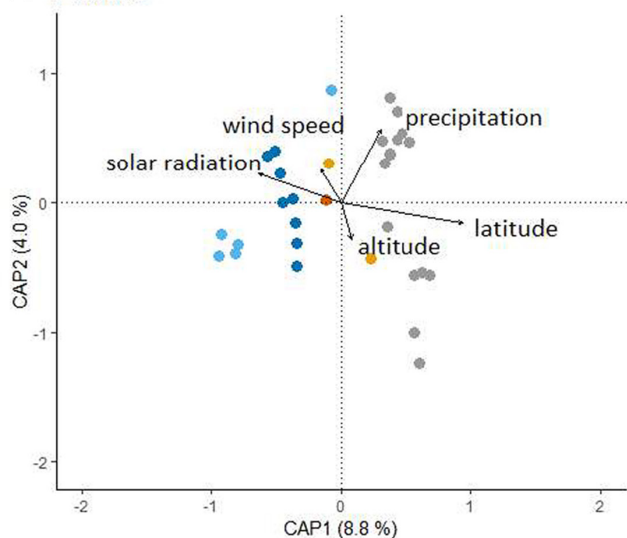
region
- Central Europe
- Eastern Europe
- Iberia
- Morocco
- Scandinavia
- Southern Balkans
- Southern Italy

genetic structure. The Scandinavian samples seem to be separated from most of the other samples following the direction associated with solar radiation and temperature (Figure 5, Table 2).

The final selected model for *F. adippe* when using the whole data set (all *F. adippe* individuals) included mean annual temperature, mean annual precipitation, mean summer wind speed, longitude and altitude as explanatory variables (adjusted $R^2 = 0.02$). Altitude ($F_{1,68} = 1.65$, $p < .001$), mean summer wind speed ($F_{1,68} = 1.4$, $p < .01$), mean annual precipitation ($F_{1,68} = 1.29$, $p = .01$) and mean annual temperature ($F_{1,68} = 1.19$, $p = .03$) were significantly associated with the genetic structure in *F. adippe*. The biplot showed that the direction of separation associated with altitude and temperature was linked to the differentiation between Iberian and Southern Balkan samples from Scandinavian samples in *F. adippe*. Precipitation and wind speed were associated with the separation of butterflies from central Europe from Iberia and some Scandinavian samples (Figure 5, Table 2). When we excluded the individuals from Iberia, the final selected model included mean annual solar radiation, mean annual temperature, mean summer wind speed, mean annual precipitation and longitude (adjusted $R^2 = 0.01$). The genetic structure in *F. adippe* when excluding Iberian individuals was significantly associated with mean annual solar radiation ($F_{1,53} = 1.49$, $p < .001$), and mean summer wind speed ($F_{1,53} = 1.23$, $p < .01$). The separation of genetic clusters was associated with similar environmental and geographical factors as in the whole data set (Figure 5, Table 2).

In *F. niobe*, the final selected model included mean annual solar radiation, mean annual precipitation, mean summer wind speed, latitude, longitude and altitude as predictor variables (adjusted $R^2 = 0.08$). The genetic structure in *F. niobe* was significantly associated with latitude ($F_{1,24} = 2.84$, $p < .001$), mean annual solar radiation ($F_{1,24} = 1.27$, $p = .03$) and mean annual precipitation ($F_{1,24} = 1.34$, $p = .02$). According to the biplot, the separation of the Iberian, Southern Balkan and Southern Italian samples from the rest was associated with mean annual solar radiation and latitude (Figure 5, Table 2).

4 | DISCUSSION

In this study, we compared phylogeographical patterns of three Argynnini species throughout Europe (and Morocco for *Speyeria aglaja*) to shed light on their postglacial colonization histories, and explored similarities and differences in genetic variation and structure among closely related sympatric species. Our data suggest that the contemporary structure is mainly linked to geographical proximity, but also reflects the species' unique refugial and postglacial

histories. Apparent differences in the species' present genetic structure might be at least partly explained by species-specific ecological requirements and differences in their dispersal abilities, showing that even sister species with partially overlapping distribution ranges may respond differently to common ecological factors. Note that the sampling localities only partially overlap between the three studied species, which is why comparisons have to be made with care. However, since the results generated by the different analytical approaches used were largely in line with each other, we are confident that our conclusions are valid.

4.1 | Distinct genetic clusters suggest differences in the biogeographical history of the study species

Studying and comparing the phylogeography of three fritillary species using nuclear data obtained through ddRADseq revealed three main genetic clusters in all studied species in Europe, namely Iberia, Southern Balkans (and Italy) and the rest of Europe. In *S. aglaja*, samples from Morocco formed a separate cluster (the other two species do not occur in the Maghreb).

According to nuclear data attained via ddRADseq, Iberian samples were clearly separated from the rest, forming the most distinct cluster for both *S. aglaja* and *Fabriciana adippe*. An Iberian cluster was also observed for *Fabriciana niobe*, although it did not constitute the furthest diverged group. We propose that this points towards an important glacial refugium in Iberia for these species, as Iberian populations were probably separated from the rest of Europe at a relative early stage as a result of population retraction to warmer areas during the ice ages. This peninsula has previously been identified as an important refugium during the LGM for many plant and animal species (e.g., Cheddadi et al., 2006; Ferrero et al., 2011).

Unlike in *F. adippe*, the northern Iberian samples were distinct from the southern Iberian ones for *S. aglaja*, possibly because the Baetic and Sierra Morena Mountain Ranges might limit dispersal for butterflies across these landscape elements. Such topographical features can shape major axes of genetic differentiation, which is also evident in multiple other terrestrial taxonomic groups, such as grasshoppers (Gonzalez-Serna et al., 2019), hummingbirds (Benham & Witt, 2016) and mammals (Castillo et al., 2014). In *F. adippe*, the whole of Iberia formed the most distinct cluster separated from the rest of Europe, with the Pyrenees presumably acting as a geographical barrier for eastward movement that might have reduced gene flow after the LGM. Catalan *F. adippe* individuals seemed to constitute a mixture of the two main clusters, which may represent a very balanced outcome of a recent

contact between the two lineages. The very high divergence of the Iberian lineage coincides with the distribution area of the subspecies *F. adippe* subsp. *chlorodippe*, whose status needs to be revised according to some authors (Tolman & Lewington, 2012; Tshikolovets, 2011). The highest divergence recovered between the Iberian haplotypes (obtained via mitochondrial *COI* sequencing) and the rest was within the range of the differentiation between *F. adippe* and *F. niobe* haplotypes, which may suggest the Iberian clade of *F. adippe* as a sister species. However, since all Catalan individuals in this study were genetically exactly intermediate according to nuclear data obtained via ddRADseq, occasional F_1 hybridization can be dismissed, and thus Catalonia could represent a recent contact zone between the two main European lineages. Further testing of this scenario with a more balanced and larger samples size would be worthwhile. In light of our present results, *F. adippe* can be considered as a single species with high probability.

The Southern Balkans emerged as another important refugium for all studied species, with Southern Balkan samples forming the most distinct cluster in *F. niobe*. During the LGM, the Balkans and Asia Minor were probably connected by a land bridge, rendering the Balkans an important European refugial region for species confined to Asia Minor during the last glaciation (Hewitt, 1999; Perissoratis & Conispoliatis, 2003). After removing Iberian and Moroccan samples in *S. aglaja*, a clear separation of Southern Balkan samples was also evident in this species, suggesting this area as another glacial refugium. However, the lack of geographical elements that hinder gene flow between the Balkans and central Europe has presumably allowed some admixture between these regions during interglacial periods (Hewitt, 2000). The stronger divergence of Balkan samples relative to the Iberian ones in *F. niobe* suggests a longer isolation of the Balkan populations in this species.

According to nuclear data obtained via ddRADseq, there was a gradual differentiation among the samples in *S. aglaja* and *F. adippe* that could be explained by geographical distance. In both species, samples from Scandinavia formed one clade along with northern European samples, following a latitudinal gradient. The evolutionary relationships between the mitochondrial *COI* haplotypes along with the analyses of nuclear ddRADseq data for *F. adippe* suggest a northwards population expansion from the glacial refugia in Iberia and Southern Balkans following the LGM, which was supported by a decreased genetic diversity in the northern regions.

Overall, the detected large-scale genetic structuring in the studied species is in line with the main European biogeographical paradigm, proposing that most temperate species were restricted to isolated refugia in Southern Europe (Iberia, Italy, Balkans and Mediterranean islands) during the Pleistocene, leading to genetically diverged lineages along these regions in many European species (Dapporto et al., 2019; Hewitt, 2000). For example, a broad-scaled genetic structuring with three major groups for Iberia, Italy and the Balkans was also found in the butterfly species *Polyommatus coridon* and *Melanargia galathea* (Habel et al., 2005; Schmitt & Seitz, 2001). Further, a European-wide study on *Maculinea alcon* butterflies

revealed a genetic pattern mainly explained by geographical proximity, which corresponded with the above-mentioned three refugial regions (Koubínová et al., 2017).

4.2 | The unique ecology of species might explain differences in the genetic structure

Besides geographical space and geological time, one factor that could contribute to the structural differences between the species is variation in dispersal capacity. In *S. aglaja*, northern and southern Iberian samples differed substantially from each other, possibly enforced by the Baetic and Sierra Morena Mountain Ranges, while such a structuring was not observed in *F. adippe*. As *F. adippe* seems to more likely perform long-distance dispersal than *S. aglaja* (Polic et al., 2020), there might have been more exchange between northern and southern Iberian populations in *F. adippe* than in *S. aglaja*, resulting in the slight differences in the contemporary genetic structure in these species. Differences in dispersal capacity have been shown to affect spatial genetic patterns in other taxa as well, with an increased genetic structure in organisms of limited mobility as opposed to highly mobile species, such as in grasshoppers (Ortego et al., 2021) and songbirds (Cros et al., 2020).

Another factor potentially contributing to differences in genetic architecture between sister species is ecological niche breadth. In Iberia, *S. aglaja* is found at higher altitudes, whereas *F. adippe* also occupies habitats at lower elevations (Vila et al., 2018). This relatively higher degree of generalism in Iberian *F. adippe* populations might have led to a less structured genetic pattern on the peninsula.

On a large scale, the stronger geographical structuring in the mitochondrial *COI* data as observed in the two specialist species *F. adippe* and *F. niobe* might be due to the less widespread warm microhabitats to which these species are confined. By contrast, the detected pattern of the relative generalist *S. aglaja* hints at higher levels of gene flow between local populations, probably because of a higher temperature tolerance, a less patchy distribution of suitable habitats and a slightly wider larval food plant range. It should be mentioned that this generalist–specialist differentiation is based on large-scale distribution patterns and habitat preferences of these species. On a local scale, this distinction may vary; for example, on the Iberian Peninsula, *S. aglaja* is considered a specialist species relative to the other two, as mentioned above (Vila et al., 2018). Similar broad-scale phylogeographical patterns to our findings have been observed in a comparative genetic study of two closely related, but cold-adapted fritillary species, namely the specialist *Boloria eunomia*, which showed a more distinct geographical structuring than the generalist *Boloria selene* (Marešová et al., 2019). These findings show that ecological specialization may affect population expansion, thereby potentially leading to more pronounced regional population differentiation in specialist than in generalist species, which has been shown for many organisms (Cros et al., 2020; Engelhardt et al., 2011; Matern et al., 2009). Our data also suggest an association of ecological niche breadth and range expansion in *F. niobe*. Although the

Balkans are not separated from central Europe by mountain ranges or other landscape elements (as is Iberia), the furthest diverged genetic cluster in *F. niobe* consisted of individuals from this region. One explanation could be that the ecological restrictions of the highly specialized *F. niobe*, such as the need for warm microhabitats (Salz & Fartmann, 2009; Spitzer et al., 2009), might have constrained its northward movement at a similar speed as in *F. adippe* and *S. aglaja*. Our findings thereby suggest that species-specific traits, such as dispersal ability and the degree of ecological generalism, might influence genetic population divergences over time.

4.3 | Mitonuclear discordance

While nuclear data obtained via ddRADseq and mitochondrial *COI* data were mostly congruent in *F. adippe* and *F. niobe*, there was discrepancy between the two data sets in *S. aglaja*: The Moroccan and the Southern Iberian samples from the Sierra Nevada formed distinct and well-separated lineages based on ddRADseq data, but they were poorly differentiated from the rest based on *COI*. Common haplotypes were shared among several regions without forming clear geographical lineages. Although our sampling does not allow for elucidating the species' biogeography in complete detail, we suggest that gene flow between the isolated lineages during interglacial cycles or selective sweeps affecting the mtDNA could have resulted in decreased haplotype diversity and thus might have erased the signs of early differentiation among the populations in the *COI* gene. Because mtDNA is more vulnerable to selective sweeps, such phenomena can be observed more strongly when sequencing the mtDNA. In contrast to nuclear data attained via ddRADseq, the most distinct lineage formed by Eastern European samples (Ag1, 6 bp from the rest) suggests that some areas in the Carpathians might have constituted an extra-Mediterranean microrefugium (Schmitt & Varga, 2012) for *S. aglaja*, as observed for some other Lepidoptera species, such as *Polyommatus coridon* (Kühne et al., 2017), *Erebia euryale* (Paučulová et al., 2017) and *Coenonympha arcania* (Cassel-Lundhagen et al., 2020). An early split might have happened between this region and the rest of Europe (Schmitt & Varga, 2012), with the Carpathians staying separated for a relatively long time and thus being less affected by gene flow or selective sweeps.

Incongruence between nuclear data as obtained through ddRADseq and mtDNA data as obtained through *COI* sequencing (mitonuclear discordance) is not uncommon (e.g., Campbell et al., 2020; Hinojosa et al., 2019; Ivanov et al., 2018), and can have several reasons, one of which is sex-biased dispersal (e.g., Nietlisbach et al., 2012). In a study on demographic patterns in *S. aglaja* in the Czech Republic, Zimmermann et al. (2009) report on higher dispersal distances in females and a lifespan twice as long as in males, which would support the hypothesis of sex-biased dispersal as the reason for the observed mitonuclear discordance. Most other explanations for mitonuclear discordance such as a faster mutation rate in mtDNA, the nonrecombinant nature of mtDNA and introgression events often lead to a higher structure in mtDNA as opposed to nuDNA

(Toews & Brelsford, 2012), which is not the case in our study. In fact, the pattern of nuDNA being far more structured than mtDNA across a large range is rather unusual, and frequently the opposite applies: potential cryptic species highlighted by *COI* that are not when studied with nuDNA (e.g., Hinojosa et al., 2019; Thielsch et al., 2017). Besides sex-biased dispersal, *Wolbachia*-mediated genetic sweeps could explain homogeneity of mtDNA, but *Wolbachia* infection was not detected in the *S. aglaja* data set. Although we cannot rule out the other explanations for mitonuclear discordance with certainty, we argue that sex-biased dispersal, with females dispersing further than males, is one of the likely reasons for the detected pattern. Note that the furthest diverged groups according to genome-wide SNP data obtained through ddRADseq (in discordance with the *COI* data) are two isolated sampling locations at the extreme of the distribution range: Sierra Nevada in Southern Iberia and Morocco. Thus, low population numbers and stochasticity have to be considered as important factors. The detected mitonuclear discrepancy suggests that some females dispersed from the European lineage to these two isolated enclaves. However, the impact of these immigrants on the nuclear genome seems to have been minimal, suggesting that the dispersal was done by extremely few individuals. This might have happened during the LGM, considering that *S. aglaja* inhabited lower altitudes then (as higher mountains were covered by ice) in these regions, and that *COI* haplotypes between Southern Iberia and Morocco are slightly differentiated (Figure 4). It is worth mentioning that the probably higher gene flow among the *S. aglaja* populations due to its relative generalist ecology and tolerance of lower temperatures might also have led to the mitonuclear discordance as shown in another Lepidoptera species (Hinojosa et al., 2019). More rigorous sampling would allow to formally test various hypotheses regarding mitonuclear discordance in *S. aglaja*.

4.4 | Associations of the genetic structure with geographical and environmental factors

Overall, it seems that geographical distance is the main driver of genetic divergence between populations in the studied species and that differences in the prevailing environmental factors may further refine the differentiation. Our analyses showed that the genetic structure is associated with a unique combination of environmental variables in each of the studied Argynnini species, although some factors were seemingly important for more than one species. Solar radiation and precipitation were associated with the genetic structure in all three species, and temperature was linked with the genetic structure in *S. aglaja* and *F. adippe*, while wind speed and altitude were only linked with the genetic patterns in *F. adippe*. For ectothermic organisms such as butterflies, solar radiation, temperature, precipitation and wind speed are important factors connected to thermoregulation, and these parameters might therefore contribute to the current genetic structure in butterfly species. However, this interpretation is potentially complicated by the fact that the extent and location of the sampling regions

of the three studied species only partially overlap. Further, while these environmental variables could contribute to the contemporary structure by exerting selective pressures or creating heterogeneous habitats which results in resistance to migration (Manel et al., 2003), the interaction of processes shaping the genetic structure of species is complex. Besides geographical features, species' intrinsic characteristics, such as ecological and life history traits, play an important role in determining phylogeographical patterns (Paz et al., 2015). The genetic structure we see today probably results from a combination of several interacting phenomena: events in the phylogeographical history of the species, stochastic processes, adaptation to prevailing environmental factors and gene flow (Cassel-Lundhagen et al., 2020; Joshi, 1997). Because environmental conditions also change over time, a joint consideration of landscape genetics and phylogeography could provide a more comprehensive explanation of the factors shaping the current distribution of genetic variation.

5 | CONCLUSIONS AND FUTURE DIRECTIONS

We utilized both genome-wide and mitochondrial sequence data in a comparative phylogeographical analysis framework to investigate how the large-scale genetic structure of three closely related fritillary butterfly species is associated with environmental and geographical factors. Taken together, the results suggest that, following isolation in several refugia during the LGM and expansion after the Pleistocene, the diverging ecology of the species together with geographical distances and barriers to dispersal has resulted in differentiation among regions, with responses to environmental factors being partly unique to each species. The mitonuclear discordance that was observed in *Speyeria aglaja* may be the result of sex-biased dispersal, selective sweeps or admixture events between lineages after geographical isolation during the Pleistocene. However, because of low sample sizes at each location it remains uncertain what caused the discrepancy, something that may be investigated in the future using a strategy with larger sample sizes.

Our phylogeographical comparisons suggest that ecological specialization and dispersal capacity might have modulated the responses to geographical and environmental factors, and thereby contributed to the variation in genetic structure between species. Future studies of small-scale structure and biogeographical histories in distinct regions across the species' distribution ranges are required to further tease apart the contributions of different species characteristics and identify the underlying mechanisms responsible for the species-specific evolutionary responses to key drivers of population genetic structure indicated by the present results. As more studies become available, systematic evaluations and meta-analysis approaches may be used to assess the generality of findings and determine whether and how genetic structure is modified by species characteristics, geography and environmental conditions.

This would deepen our understanding of how contemporary distribution ranges are shaped, improve the ability to predict a species' future development in response to environmental change, and hence inform conservation management in the respective countries of occurrence about appropriate measures.

AUTHOR CONTRIBUTIONS

D.P., Y.Y., A.F., M.F. and R.V. designed the study. Funding was secured by A.F., M.F., D.P. and Y.Y. D.P., Y.Y. and K.M.L. performed the laboratory work. D.P., M.F. and R.V. contributed to field sampling. M.M. provided laboratory infrastructure to prepare the ddRADseq libraries. D.P. and Y.Y. analysed the data, and wrote the first draft of the manuscript. All authors contributed to the final version and approved the submitted manuscript.

CONFLICT OF INTEREST

The authors have no conflicts of interest to disclose.

DATA AVAILABILITY STATEMENT

COI sequences were deposited in the NCBI nucleotide database (Accession codes: MZ026157–MZ026168, MZ026170, MZ026171, MZ026174–MZ026200, MZ026202–MZ026242, MZ026245–MZ026261, MZ026263–MZ026290). Raw ddRAD sequences can be found in the NCBI SRA under the BioProject ID PRJNA70744.

ORCID

Daniela Polic https://orcid.org/0000-0001-6168-6489
Yeşerin Yıldırım https://orcid.org/0000-0002-0431-8417
Kyung Min Lee https://orcid.org/0000-0001-9055-1073
Markus Franzén https://orcid.org/0000-0001-8022-5004
Marko Mutanen https://orcid.org/0000-0003-4464-6308

Roger Vila https://orcid.org/0000-0002-2447-4388
Anders Forsman https://orcid.org/0000-0001-9598-7618

REFERENCES

Adams, R. I., & Hadly, E. A. (2013). Genetic diversity within vertebrate species is greater at lower latitudes. *Evolutionary Ecology*, *27*(1), 133–143. https://doi.org/10.1007/s10682-012-9587-x

Ahola, V., Lehtonen, R., Somervuo, P., Salmela, L., Koskinen, P., Rastas, P., Välimäki, N., Paulin, L., Kvist, J., Wahlberg, N., Tanskanen, J., Hornett, E. A., Ferguson, L. C., Luo, S., Cao, Z., de Jong, M. A., Duplouy, A., Smolander, O.-P., Vogel, H., ... Hanski, I. (2014). The Glanville fritillary genome retains an ancient karyotype and reveals selective chromosomal fusions in Lepidoptera. *Nature Communications*, *5*, 4737.

Allio, R., Donega, S., Galtier, N., & Nabholz, B. (2017). Large variation in the ratio of mitochondrial to nuclear mutation rate across animals: Implications for genetic diversity and the use of mitochondrial DNA as a molecular marker. *Molecular Biology and Evolution*, *34*(11), 2762–2772. https://doi.org/10.1093/molbev/msx197

Andrews, K. R., Good, J. M., Miller, M. R., Luikart, G., & Hohenlohe, P. A. (2016). Harnessing the power of RADseq for ecological and evolutionary genomics. *Nature Reviews Genetics*, *17*(2), 81–92. https://doi.org/10.1038/nrg.2015.28

Avise, J. C. (2000). *Phylogeography: the history and formation of species.* Harvard university press.

Avise, J. C., Bowen, B. W., & Ayala, F. J. (2016). In the light of evolution X: Comparative phylogeography. *Proceedings of the National Academy of Sciences*, *113*(29), 7957–7961. https://doi.org/10.1073/pnas.1604338113

Ballard, J. W. O., & Whitlock, M. C. (2004). The incomplete natural history of mitochondria. *Molecular Ecology*, *13*(4), 729–744.

Bandelt, H.-J., Forster, P., & Röhl, A. (1999). Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution*, *16*(1), 37–48.

Benham, P. M., & Witt, C. C. (2016). The dual role of Andean topography in primary divergence: functional and neutral variation among populations of the hummingbird, Metallura tyrianthina. *BMC Evolutionary Biology*, *16*(1), 22. https://doi.org/10.1186/s12862-016-0595-2

Bermingham, E., & Moritz, C. (1998). Comparative phylogeography: concepts and applications. *Molecular Ecology*, *7*(4), 367–369. https://doi.org/10.1046/j.1365-294x.1998.00424.x

Bernardo, J., & Spotila, J. R. (2006). Physiological constraints on organismal response to global warming: Mechanistic insights from clinally varying populations and implications for assessing endangerment. *Biology Letters*, *2*(1), 135–139. https://doi.org/10.1098/rsbl.2005.0417

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120.

Campbell, E. O., Gage, E. V., Gage, R. V., & Sperling, F. A. H. (2020). Single nucleotide polymorphism-based species phylogeny of greater fritillary butterflies (Lepidoptera: Nymphalidae: Speyeria) demonstrates widespread mitonuclear discordance. *Systematic Entomology*, *45*(2), 269–280. https://doi.org/10.1111/syen.12393

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. https://doi.org/10.1111/jzs.12360

Castillo, J. A., Epps, C. W., Davis, A. R., & Cushman, S. A. (2014). Landscape effects on gene flow for a climate-sensitive montane species, the American pika. *Molecular Ecology*, *23*(4), 843–856. https://doi.org/10.1111/mec.12650

Catchen, J. M., Amores, A., Hohenlohe, P., Cresko, W., & Postlethwait, J. H. (2011). Stacks: Building and genotyping Loci de novo from short-read sequences. *G3 (Bethesda)*, *1*(3), 171–182. https://doi.org/10.1534/g3.111.000240

Catchen, J. M., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). Stacks: An analysis tool set for population genomics. *Molecular Ecology*, *22*(11), 3124–3140. https://doi.org/10.1111/mec.12354

Cheddadi, R., Vendramin, G. G., Litt, T., François, L., Kageyama, M., Lorentz, S., Laurent, J.-M., de Beaulieu, J.-L., Sadori, L., Jost, A., & Lunt, D. (2006). Imprints of glacial refugia in the modern genetic diversity of *Pinus sylvestris*. *Global Ecology and Biogeography*, *15*(3), 271–282. https://doi.org/10.1111/j.1466-8238.2006.00226.x

Cooper, J. K., Sykes, G., King, S., Cottrill, K., Ivanova, N. V., Hanner, R., & Ikonomi, P. (2007). Species identification in cell culture: A two-pronged molecular approach. *In Vitro Cellular & Developmental Biology-Animal*, *43*(10), 344–351.

Crispo, E., Bentzen, P., Reznick, D. N., Kinnison, M. T., & Hendry, A. P. (2006). The relative influence of natural selection and geography on gene flow in guppies. *Molecular Ecology*, *15*, 49–62.

Cros, E., Chattopadhyay, B., Garg, K. M., Ng, N. S. R., Tomassi, S., Benedick, S., Edwards, D. P., & Rheindt, F. E. (2020). Quaternary land bridges have not been universal conduits of gene flow. *Molecular Ecology*, *29*(14), 2692–2706. https://doi.org/10.1111/mec.15509

Dapporto, L., Cini, A., Vodă, R., Dincă, V., Wiemers, M., Menchetti, M., Magini, G., Talavera, G., Shreeve, T., Bonelli, S., Casacci, L. P., Balletto, E., Scalercio, 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.

Davis, M. B., & Shaw, R. G. (2001). Range shifts and adaptive responses to Quaternary climate change. *Science*, *292*(5517), 673–679.

Despres, 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

Dincă, V., Lee, K. M., Vila, R., & Mutanen, M. (2019). The conundrum of species delimitation: A genomic perspective on a mitogenetically super-variable butterfly. *Proceedings of the Royal Society B: Biological Sciences*, *286*(1911), 20191311. https://doi.org/10.1098/rspb.2019.1311

Eliasson, C., Ryrholm, N., & Gärdenfors, U. (2005). *Nationalnyckeln till Sveriges flora och fauna. Fjärilar. Dagfjärilar, Hesperiidae – Nymphalidae*. Artdatabanken SLU.

Ellis, S., Wainwright, D., Dennis, E. B., Bourn, N. A. D., Bulman, C. R., Hobson, R., Jones, R., Middlebrook, I., Plackett, J., Smith, R. G., Wain, M., & Warren, M. S. (2019). Are habitat changes driving the decline of the UK's most threatened butterfly: The High Brown Fritillary Argynnis adippe (Lepidoptera: Nymphalidae)? *Journal of Insect Conservation*, *23*, 351–367. https://doi.org/10.1007/s10841-019-00134-0

Ellis, S., Warren, M., Brereton, T., Toynton, P., & Chandler, D. (2010). *Fact Sheet - Dark Green Fritillary*. Butterfly Conservation, Dorset. https://butterfly-conservation.org/sites/default/files/dark-green-fritillary-regional

Engelhardt, C. H. M., Haase, P., & Pauls, S. U. (2011). From the Western Alps across Central Europe: Postglacial recolonisation of the tufa stream specialist Rhyacophila pubescens (Insecta, Trichoptera). *Frontiers in Zoology*, *8*(1), 10. https://doi.org/10.1186/1742-9994-8-10

Excoffier, L., Laval, G., & Schneider, S. (2005). Arlequin (version 3.0): An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics*, *1*, 117693430500100003.

Ferrero, M. E., Blanco-Aguiar, J. A., Lougheed, S. C., Sánchez-Barbudo, I., De Nova, P. J. G., Villafuerte, R., & Dávila, J. A. (2011). Phylogeography and genetic structure of the red-legged partridge (Alectoris rufa): more evidence for refugia within the Iberian glacial refugium. *Molecular Ecology*, 20(12), 2628–2642. https://doi.org/10.1111/j.1365-294X.2011.05111.x

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, 4302–4315.

Forster, W., & Wohlfahrt, T. A. (1955). *Die Schmetterlinge Mitteleuropas. 2. Tagfalter: Diurna (Rhopalocera und Hesperiidae)*. Franckh'sche Verlagshandlung.

Fox, R., Warren, M. S., Brereton, T. M., Roy, D. B., & Robinson, A. (2011). A new Red List of British butterflies. *Insect Conservation and Diversity*, 4(3), 159–172. https://doi.org/10.1111/j.1752-4598.2010.00117.x

Franzén, M., & Johannesson, M. (2007). Predicting extinction risk of butterflies and moths (Macrolepidoptera) from distribution patterns and species characteristics. *Journal of Insect Conservation*, 11(4), 367–390. https://doi.org/10.1007/s10841-006-9053-6

Fric, Z., Klimova, M., Hula, V., & Konvicka, M. (2005). Caterpillars of Argynnis aglaja (LINNAEUS, 1758) feeding on Bistorta major. *Atalanta*, 36, 119–121.

Gompert, Z., Forister, M. L., Fordyce, J. A., & Nice, C. C. (2008). Widespread mito-nuclear discordance with evidence for introgressive hybridization and selective sweeps in Lycaeides. *Molecular Ecology*, 17(24), 5231–5244. https://doi.org/10.1111/j.1365-294X.2008.03988.x

Gonzalez-Serna, M. J., Cordero, P. J., & Ortego, J. (2019). Spatiotemporally explicit demographic modelling supports a joint effect of historical barriers to dispersal and contemporary landscape composition on structuring genomic variation in a red-listed grasshopper. *Molecular Ecology*, 28(9), 2155–2172. https://doi.org/10.1111/mec.15086

Guichoux, E., Lagache, L., Wagner, S., Chaumeil, P., Léger, P., Lepais, O., Lepoittevin, C., Malausa, T., Revardel, E., Salin, F., & Petit, R. J. (2011). Current trends in microsatellite genotyping. *Molecular Ecology Resources*, 11(4), 591–611. https://doi.org/10.1111/j.1755-0998.2011.03014.x

Habel, J. C., Schmitt, T., & Müller, P. (2005). The fourth paradigm pattern of post-glacial range expansion of European terrestrial species: the phylogeography of the Marbled White butterfly (Satyrinae, Lepidoptera). *Journal of Biogeography*, 32(8), 1489–1497.

Hajibabaei, M., Singer, G. A., Hebert, P. D., & Hickey, D. A. (2007). DNA barcoding: how it complements taxonomy, molecular phylogenetics and population genetics. *Trends in Genetics*, 23(4), 167–172. https://doi.org/10.1016/j.tig.2007.02.001

Harrison, R. G. (1989). Animal mitochondrial DNA as a genetic marker in population and evolutionary biology. *Trends in Ecology & Evolution*, 4(1), 6–11. https://doi.org/10.1016/0169-5347(89)90006-2

Hewitt, G. M. (1999). Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society*, 68(1–2), 87–112.

Hewitt, G. M. (2000). The genetic legacy of the Quaternary ice ages. *Nature*, 405, 907–913.

Hewitt, G. M. (2004). Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 359(1442), 183–195.

Hickerson, M. J., Carstens, B. C., Cavender-Bares, J., Crandall, K. A., Graham, C. H., Johnson, J. B., Rissler, L., Victoriano, P. F., & Yoder, A. D. (2010). Phylogeography's past, present, and future: 10 years after. *Molecular Phylogenetics and Evolution*, 54, 291–301.

Higgins, L. G., & Riley, N. D. (1970). *A field guide to the butterflies of Britain and Europe*. Collins.

Hijmans, R. J., van Etten, J., Sumner, M., Cheng, J., Baston, D., Bevan, A., Bivand, R., Busetto, L., Canty, M., Fasoli, B., Forrest, D., Ghosh, A., Golicher, D., Gray, J., Greenberg, J. A., Hiemstra, P., Hingee, K., Ilich, A., Institute for Mathematics Applied Geosciences ... Wueest, R. (2021). raster: Geographic Data Analysis and Modeling. R package version 3.5-11. Retrieved from http://CRAN.R-project.org/package=raster.

Hinojosa, J. C., Koubínová, D., Szenteczki, M. A., Pitteloud, C., Dincă, V., Alvarez, N., & Vila, R. (2019). A mirage of cryptic species: Genomics uncover striking mitonuclear discordance in the butterfly Thymelicus sylvestris. *Molecular Ecology*, 28(17), 3857–3868. https://doi.org/10.1111/mec.15153

Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q., & Vinh, L. S. (2017). UFBoot2: Improving the Ultrafast Bootstrap Approximation. *Molecular Biology and Evolution*, 35(2), 518–522. https://doi.org/10.1093/molbev/msx281

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6, 65–70.

Hughes, A. L., & Hughes, M. A. K. (2007). Coding sequence polymorphism in avian mitochondrial genomes reflects population histories. *Molecular Ecology*, 16(7), 1369–1376. https://doi.org/10.1111/j.1365-294X.2007.03242.x

Hurst, G. D., & Jiggins, F. M. (2005). Problems with mitochondrial DNA as a marker in population, phylogeographic and phylogenetic studies: The effects of inherited symbionts. *Proceedings of the Royal Society B: Biological Sciences*, 272(1572), 1525–1534.

Ivanov, V., Lee, K. M., & Mutanen, M. (2018). Mitonuclear discordance in wolf spiders: Genomic evidence for species integrity and introgression. *Molecular Ecology*, 27(7), 1681–1695. https://doi.org/10.1111/mec.14564

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning (Vol. 112)*. Springer.

Joshi, A. (1997). Adaptive evolution and the footprints of history. *Current Science*, 72(12), 944–949.

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A., & Jermiin, L. S. (2017). ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nature Methods*, 14(6), 587–589. https://doi.org/10.1038/nmeth.4285

Karl, I., Janowitz, S. A., & Fischer, K. (2008). Altitudinal life-history variation and thermal adaptation in the copper butterfly Lycaena tityrus. *Oikos*, 117, 778–788.

Kawecki, T. J., & Ebert, D. (2004). Conceptual issues in local adaptation. *Ecology Letters*, 7(12), 1225–1241. https://doi.org/10.1111/j.1461-0248.2004.00684.x

Kingsolver, J. G., & Watt, W. B. (1983). Thermoregulatory strategies in colias butterflies: Thermal stress and the limits to adaptation in temporally varying environments. *The American Naturalist*, 121, 32–55.

Kingsolver, J. G., & Wiernasz, D. C. (1991). Seasonal polyphenism in wing-melanin pattern and thermoregulatory adaptation in pieris butterflies. *The American Naturalist*, 137, 816–830.

Komonen, A., Grapputo, A., Kaitala, V., Kotiaho, J. S., & Päivinen, J. (2004). The role of niche breadth, resource availability and range position on the life history of butterflies. *Oikos*, 105, 41–54.

Koubínová, D., Dincă, V., Dapporto, L., Vodă, R., Suchan, T., Vila, R., & Alvarez, N. (2017). Genomics of extreme ecological specialists: Multiple convergent evolution but no genetic divergence between ecotypes of Maculinea alcon butterflies. *Scientific Reports*, 7(1), 13752. https://doi.org/10.1038/s41598-017-12938-8

Kühne, G., Kosuch, J., Hochkirch, A., & Schmitt, T. (2017). Extra-Mediterranean glacial refugia in a Mediterranean faunal element: The phylogeography of the chalk-hill blue Polyommatus coridon (Lepidoptera, Lycaenidae). *Scientific Reports*, 7(1), 1–11.

Kuussaari, M., Saarinen, M., Korpela, E. L., Poyry, J., & Hyvonen, T. (2014). Higher mobility of butterflies than moths connected to habitat suitability and body size in a release experiment. *Ecology and Evolution*, 4(19), 3800–3811. https://doi.org/10.1002/ece3.1187

Lansman, R. A., Avise, J. C., & Huettel, M. D. (1983). Critical experimental test of the possibility of "paternal leakage" of mitochondrial DNA. *Proceedings of the National Academy of Sciences*, 80(7), 1969–1971.

Lee, K. M., Kivelä, S. M., Ivanov, V., Hausmann, A., Kaila, L., Wahlberg, N., & Mutanen, M. (2018). Information dropout patterns in restriction site associated DNA phylogenomics and a comparison with multi-locus Sanger data in a species-rich moth genus. *Systematic Biology*, *67*(6), 925–939.

Legendre, P., & Anderson, M. J. (1999). Distance-based redundancy analysis: Testing multispecies responses in multifactorial ecological experiments. *Ecological Monographs*, *69*(1), 1–24.

Leigh, J. W., & Bryant, D. (2015). POPART: Full-feature software for haplotype network construction. *Methods in Ecology and Evolution*, *6*(9), 1110–1116.

Lenormand, T. (2002). Gene flow and the limits to natural selection. *Trends in Ecology & Evolution*, *17*(4), 183–189. https://doi.org/10.1016/S0169-5347(02)02497-7

Letunic, I., & Bork, P. (2019). Interactive Tree Of Life (iTOL) v4: Recent updates and new developments. *Nucleic Acids Research*, *47*(W1), W256–W259. https://doi.org/10.1093/nar/gkz239

Magalhaes, I. S., Gleiser, G., Labouche, A.-M., & Bernasconi, G. (2011). Comparative population genetic structure in a plant–pollinator/seed predator system. *Molecular Ecology*, *20*(22), 4618–4630. https://doi.org/10.1111/j.1365-294X.2011.05296.x

Malinsky, M., Trucchi, E., Lawson, D. J., & Falush, D. (2018). RADpainter and fineRADstructure: Population Inference from RADseq Data. *Molecular Biology and Evolution*, *35*(5), 1284–1290. https://doi.org/10.1093/molbev/msy023

Manel, S., Schwartz, M. K., Luikart, G., & Taberlet, P. (2003). Landscape genetics: Combining landscape ecology and population genetics. *Trends in Ecology & Evolution*, *18*(4), 189–197.

Marešová, J., Habel, J. C., Neve, G., Sielezniew, M., Bartonova, A., Kostro-Ambroziak, A., & Fric, Z. F. (2019). Cross-continental phylogeography of two Holarctic Nymphalid butterflies, Boloria eunomia and Boloria selene. *PLoS One*, *14*(3), e0214483. https://doi.org/10.1371/journal.pone.0214483

Matern, A., Desender, K., Drees, C., Gaublomme, E., Paill, W., & Assmann, T. (2009). Genetic diversity and population structure of the endangered insect species Carabus variolosus in its western distribution range: Implications for conservation. *Conservation Genetics*, *10*(2), 391–405.

Meng, Z., Lei, C., Chen, X., & Jiang, S. (2016). Complete mitochondrial genome sequence of Heliconius melpomene rosina (Insecta: Lepidoptera: Nymphalidae). *Mitochondrial DNA Part A*, *27*(6), 3911–3912.

Michaux, J. R., Libois, R., & Filippucci, M. G. (2005). So close and so different: comparative phylogeography of two small mammal species, the Yellow-necked fieldmouse (*Apodemus flavicollis*) and the Woodmouse (*Apodemus sylvaticus*) in the Western Palearctic region. *Heredity*, *94*(1), 52–63. https://doi.org/10.1038/sj.hdy.6800561

Mopper, S., & Strauss, S. Y. (2013). *Genetic structure and local adaptation in natural insect populations: Effects of ecology, life history, and behavior.* Springer Science & Business Media.

Nguyen, L.-T., Schmidt, H. A., von Haeseler, A., & Minh, B. Q. (2015). IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution*, *32*(1), 268–274. https://doi.org/10.1093/molbev/msu300

Nietlisbach, P., Arora, N., Nater, A., Goossens, B., Van Schaik, C. P., & Kruetzen, M. (2012). Heavily male-biased long-distance dispersal of orang-utans (genus: Pongo), as revealed by Y-chromosomal and mitochondrial genetic markers. *Molecular Ecology*, *21*(13), 3173–3186.

Öckinger, E., Hammarstedt, O., Nilsson, S. G., & Smith, H. G. (2006). The relationship between local extinctions of grassland butterflies and increased soil nitrogen levels. *Biological Conservation*, *128*(4), 564–573.

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., O'Hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H., & Wagner, H. (2019). *vegan: Community Ecology Package. R package version 2.5-5.* Retrieved from: https://CRAN.R-project.org/package=vegan.

Ortego, J., Gutiérrez-Rodríguez, J., & Noguerales, V. (2021). Demographic consequences of dispersal-related trait shift in two recently diverged taxa of montane grasshoppers. *Evolution*, *75*, 1998–2013.

Papadopoulou, A., & Knowles, L. L. (2016). Toward a paradigm shift in comparative phylogeography driven by trait-based hypotheses. *Proceedings of the National Academy of Sciences*, *113*(29), 8018–8024. https://doi.org/10.1073/pnas.1601069113

Paris, J. R., Stevens, J. R., Catchen, J. M., & Johnston, S. (2017). Lost in parameter space: A road map for stacks. *Methods in Ecology and Evolution*, *8*(10), 1360–1373. https://doi.org/10.1111/2041-210x.12775

Paučulová, L., Šemeláková, M., Mutanen, M., Pristaš, P., & Panigaj, Ľ. (2017). Searching for the glacial refugia of Erebia euryale (Lepidoptera, Nymphalidae)–insights from mt DNA-and nDNA-based phylogeography in the Western Carpathians. *Journal of Zoological Systematics and Evolutionary Research*, *55*(2), 118–128.

Paz, A., Ibanez, R., Lips, K. R., & Crawford, A. J. (2015). Testing the role of ecology and life history in structuring genetic variation across a landscape: A trait-based phylogeographic approach. *Molecular Ecology*, *24*(14), 3723–3737. https://doi.org/10.1111/mec.13275

Pereira, H. M. (2016). A latitudinal gradient for genetic diversity. *Science*, *353*(6307), 1494–1495. https://doi.org/10.1126/science.aah6730

Perissoratis, C., & Conispoliatis, N. (2003). The impacts of sea-level changes during latest Pleistocene and Holocene times on the morphology of the Ionian and Aegean seas (SE Alpine Europe). *Marine Geology*, *196*(3), 145–156. https://doi.org/10.1016/S0025-3227(03)00047-1

Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., & Hoekstra, H. E. (2012). Double digest RADseq: An inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS One*, *7*(5), e37135. https://doi.org/10.1371/journal.pone.0037135

Piao, S., Fang, J., Zhou, L., Ciais, P., & Zhu, B. (2006). Variations in satellite-derived phenology in China's temperate vegetation. *Global Change Biology*, *12*(4), 672–685. https://doi.org/10.1111/j.1365-2486.2006.01123.x

Pilot, M., Jędrzejewski, W., Branicki, W., Sidorovich, V. E., Jędrzejewska, B., Stachura, K., & Funk, S. M. (2006). Ecological factors influence population genetic structure of European grey wolves. *Molecular Ecology*, *15*(14), 4533–4553. https://doi.org/10.1111/j.1365-294X.2006.03110.x

Polgar, C. A., Primack, R. B., Williams, E. H., Stichter, S., & Hitchcock, C. (2013). Climate effects on the flight period of Lycaenid butterflies in Massachusetts. *Biological Conservation*, *160*, 25–31. https://doi.org/10.1016/j.biocon.2012.12.024

Polic, D., Tamario, C., Franzén, M., Betzholtz, P. E., Yıldırım, Y., & Forsman, A. (2020). Movements and occurrence in two closely related fritillary species. *Ecological Entomology*, *46*(2), 428–439. https://doi.org/10.1111/een.12987

Pollard, E., & Yates, T. J. (1994). *Monitoring butterflies for ecology and conservation: The British butterfly monitoring scheme.* Springer Science & Business Media.

QGIS.org. (2022). *QGIS geographic information system.* QGIS Association. http://www.qgis.org

R Core Team. (2020). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing. https://www.R-project.org/

Raj, A., Stephens, M., & Pritchard, J. K. (2014). fastSTRUCTURE: Variational inference of population structure in large SNP data sets. *Genetics*, *197*(2), 573–589. https://doi.org/10.1534/genetics.114.164350

Reinhardt, R., & Bolz, R. (2011). Rote Liste und Gesamtartenliste der Tagfalter (Rhopalocera)(Lepidoptera: Papilionoidea et Hesperioidea) Deutschlands. *Naturschutz und Biologische Vielfalt*, *70*(3), 165–194.

Ren, G., Mateo, R. G., Liu, J., Suchan, T., Alvarez, N., Guisan, A., Conti, E., & Salamin, N. (2017). Genetic consequences of Quaternary climatic oscillations in the Himalayas: Primula tibetica as a case study based on restriction site-associated DNA sequencing. *New Phytologist*, *213*(3), 1500–1512. https://doi.org/10.1111/nph.14221

Riddle, B. R. (2016). Comparative phylogeography clarifies the complexity and problems of continental distribution that drove AR Wallace to favor islands. *Proceedings of the National Academy of Sciences*, *113*, 7970–7977.

Ritter, S., Michalski, S. G., Settele, J., Wiemers, M., Fric, Z. F., Sielezniew, M., Šašić, M., Rozier, Y., & Durka, W. (2013). Wolbachia infections mimic cryptic speciation in two parasitic butterfly species, Phengaris teleius and P. nausithous (Lepidoptera: Lycaenidae). *PLoS One*, *8*(11), e78107.

Robertson, P., Clarke, S., & Warren, M. (1995). Woodland management and butterfly diversity. In A. S. Pullin (Ed.), *Ecology and conservation of butterflies* (pp. 113–122). Springer.

Rochette, N. C., & Catchen, J. M. (2017). Deriving genotypes from RAD-seq short-read data using Stacks. *Nature Protocols*, *12*, 2640–2659. https://doi.org/10.1038/nprot.2017.123

Rodriguez-Caro, L., Fenner, J., Benson, C., Van Belleghem, S. M., & Counterman, B. A. (2020). Genome assembly of the dogface butterfly Zerene cesonia. *Genome Biology and Evolution*, *12*(1), 3580–3585.

Roed, K. H., Kvie, K. S., Hasle, G., Gilbert, L., & Leinaas, H. P. (2016). Phylogenetic lineages and postglacial dispersal dynamics characterize the genetic structure of the tick, Ixodes ricinus, in Northwest Europe. *PLoS One*, *11*(12), e0167450. https://doi.org/10.1371/journal.pone.0167450

Rozas, J., Ferrer-Mata, A., Sánchez-DelBarrio, J. C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S. E., & Sánchez-Gracia, A. (2017). DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Molecular Biology and Evolution*, *34*(12), 3299–3302.

Rundle, H. D., & Nosil, P. (2005). Ecological speciation. *Ecology Letters*, *8*(3), 336–352. https://doi.org/10.1111/j.1461-0248.2004.00715.x

Briscoe Runquist, R. D., Gorton, A. J., Yoder, J. B., Deacon, N. J., Grossman, J. J., Kothari, S., Lyons, M. P., Sheth, S. N., Tiffin, P., & Moeller, D. A. (2020). Context dependence of local adaptation to abiotic and biotic environments: A quantitative and qualitative synthesis. *The American Naturalist*, *195*(3), 412–431. https://doi.org/10.1086/707322

Salz, A., & Fartmann, T. (2009). Coastal dunes as important strongholds for the survival of the rare Niobe fritillary (*Argynnis niobe*). *Journal of Insect Conservation*, *13*, 643–654. https://doi.org/10.1007/s10841-009-9214-5

Schmitt, T., & Seitz, A. (2001). Allozyme variation in Polyommatus coridon (Lepidoptera: Lycaenidae): Identification of ice-age refugia and reconstruction of post-glacial expansion. *Journal of Biogeography*, *28*(9), 1129–1136.

Schmitt, T., & Varga, Z. (2012). Extra-Mediterranean refugia: The rule and not the exception? *Frontiers in Zoology*, *9*(1), 22. https://doi.org/10.1186/1742-9994-9-22

Shikano, T., Shimada, Y., Herczeg, G., & Merila, J. (2010). History vs. habitat type: explaining the genetic structure of European nine-spined stickleback (*Pungitius pungitius*) populations. *Molecular Ecology*, *19*(6), 1147–1161. https://doi.org/10.1111/j.1365-294X.2010.04553.x

Sobel, J. M., Chen, G. F., Watt, L. R., & Schemske, D. W. (2010). The biology of speciation. *Evolution*, *64*, 295–315.

Soubrier, J., Steel, M., Lee, M. S., Der Sarkissian, C., Guindon, S., Ho, S. Y., & Cooper, A. (2012). The influence of rate heterogeneity among sites on the time dependence of molecular rates. *Molecular Biology and Evolution*, *29*(11), 3345–3358.

Soucy, S. M., Huang, J., & Gogarten, J. P. (2015). Horizontal gene transfer: Building the web of life. *Nature Reviews Genetics*, *16*(8), 472–482.

Spitzer, L., Beneš, J., & Konvička, M. (2009). Oviposition of the Niobe fritillary (Argynnis niobe [Linnaeus, 1758]) at sub-mountain conditions in the Czech Carpathians (Lepidoptera, Nymphalidae). *Nachrichten des entomologischen Vereins Apollo*, *30*, 165–168.

Stewart, J. R., Lister, A. M., Barnes, I., & Dalen, L. (2010). Refugia revisited: Individualistic responses of species in space and time. *Proceedings of the Royal Society B: Biological Sciences*, *277*(1682), 661–671. https://doi.org/10.1098/rspb.2009.1272

Thielsch, A., Knell, A., Mohammadyari, A., Petrusek, A., & Schwenk, K. (2017). Divergent clades or cryptic species? Mito-nuclear discordance in a Daphnia species complex. *BMC Evolutionary Biology*, *17*(1), 227. https://doi.org/10.1186/s12862-017-1070-4

Toews, D. P. L., & Brelsford, A. (2012). The biogeography of mitochondrial and nuclear discordance in animals. *Molecular Ecology*, *21*(16), 3907–3930.

Tolman, T., & Lewington, R. (2012). *Schmetterlinge Europas und Nordwestafrikas* (2nd ed.). Franckh-Kosmos Verlag.

Tshikolovets, V. V. (2011). *Butterflies of Europe & the Mediterranean area*. Tshikolovets.

van Swaay, C., & Warren, M. (1999). *Red Data book of European butterflies (Rhopalocera). Nature and Environment, No. 99*. Council of Europe Publishing.

Victoriano, P. F., Ortiz, J. C., Benavides, E., Adams, B. J., & Sites, J. W., Jr. (2008). Comparative phylogeography of codistributed species of Chilean Liolaemus (Squamata: Tropiduridae) from the central-southern Andean range. *Molecular Ecology*, *17*(10), 2397–2416. https://doi.org/10.1111/j.1365-294X.2008.03741.x

Vila, R., Stefanescu, C., & Sesma, J. M. (2018). *Guia de les papallones diürnes de Catalunya*. Lynx edicions.

Wallace, J. R. (2012). Imap – Interactive Mapping. R package version 1.32. Retrieved from http://CRAN.R-project.org/package=Imap.

Warren, M. S. (1995). Managing local microclimates for the high brown fritillary, Argynnis adippe. In A. S. Pullin (Ed.), *Ecology and Conservation of Butterflies* (pp. 198–210). Springer Netherlands.

Werren, J. H., Baldo, L., & Clark, M. E. (2008). Wolbachia: Master manipulators of invertebrate biology. *Nature Reviews Microbiology*, *6*(10), 741–751. https://doi.org/10.1038/nrmicro1969

Wright, S. (1943). Isolation by distance. *Genetics*, *28*, 114.

Xie, P., Zhao, G., Niu, J.-G., Wang, J., Zhou, Q., Guo, Y., & Ma, X.-F. (2019). Comprehensive analysis of population genetics of Phoxinus phoxinus ujmonensis in the Irtysh River: Abiotic and biotic factors. *Ecology and Evolution*, *9*(14), 7997–8012. https://doi.org/10.1002/ece3.5320

Xun, H., Li, H., Li, S., Wei, S., Zhang, L., Song, F., Jiang, P., Yang, H., Han, F., & Cai, W. (2016). Population genetic structure and post-LGM expansion of the plant bug Nesidiocoris tenuis (Hemiptera: Miridae) in China. *Scientific Reports*, *6*, 26755. https://doi.org/10.1038/srep26755

Yang, Z. (1995). A space-time process model for the evolution of DNA sequences. *Genetics*, *139*(2), 993–1005.

Yıldırım, Y., Anderson, M. J., Hansson, B., Patel, S., Millar, C. D., & Rainey, P. B. (2018). Genetic structure of the grey side-gilled sea slug (Pleurobranchaea maculata) in coastal waters of New Zealand. *PLoS One*, *13*(8), e0202197.

Zaman, K., Hubert, M. K., & Schoville, S. D. (2019). Testing the role of ecological selection on colour pattern variation in the butterfly Parnassius clodius. *Molecular Ecology*, *28*(23), 5086–5102. https://doi.org/10.1111/mec.15279

Zima, J., Leština, D., & Konvička, M. (2013). Characterization of ten polymorphic microsatellite markers for an endangered butterfly Argynnis niobe and their cross-species utility in the closely related species A. adippe (Lepidoptera: Nymphalidae). *European Journal of Entomology*, *110*(2), 383–387. https://doi.org/10.14411/eje.2013.098

Zimmermann, K., Konvička, M., Fric, Z., & Čiháková, V. (2009). Demography of a common butterfly on humid grasslands: *Argynnis aglaja* (Lepidoptera: Nymphalidae) studied by mark-recapture. *Polish Journal of Ecology*, *57*(4), 715–727.

Zink, R. M., Kessen, A. E., Line, T. V., & Blackwell-Rago, R. C. (2001). Comparative phylogeography of some aridland bird species. *The Condor*, *103*, 1–10.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.