Vulnerability of the iconic mountain butterfly Parnassius apollo to climate change - a genomic approach

© C. Barriuso

Thomas FRANCISCO
Rapport de stage
Master Biodiversité, Ecologie et Evolution
Parcours recherche
Mention Biodiversité et fonctionnement des écosystèmes terrestres
Université de Bordeaux
Mémoire défendu le **20/06/2023**

Encadrement : Laurence DESPRES
Laboratoire d'Ecologie Alpine
laurence.despres@univ-grenoble-alpes.fr

Laboratoire d'Ecologie Alpine
Bâtiment D, 2233 Rue de la Piscine,
38610 Gières

Abstract

Major changes are expected under ongoing climate change, especially in mountain areas. Models that predict the vulnerability of the biodiversity by incorporating the intraspecific variation are rapidly emerging. They enable a better comprehension of the vulnerability within the species level. Here, we investigated the vulnerability of the Apollo - a mountain butterfly - under climate change. Based on the genetic data of 293 Apollo spread in French mountains obtained by ddRAD sequencing method, a genetic environment association approach was performed, using RDA, to incorporate the local adaptation as intraspecific variation. Then, the disturbance of this genetic environmental association was calculated under future climatic conditions as a metric of maladaptation called genomic offset. Our results showed patterns of local adaptation with the identification of 32 loci that seem to be close to genomic regions involved in adaptation to cold and seasonal temperature and high precipitations and involved in ongoing adaptation to warm temperature and low and seasonal precipitations. Besides, we also identified areas where the vulnerability of Apollo to climate change will be important. Indeed, we found that the Pyrenees, Northern Alps, and South of the Central Massif will be affected by an important disruption of the genetic environment association and the Cévennes, Southern Alps and Western of the Pyrenees will face climatic conditions outside the current adaptive range. This study has provided some information about the patterns of local adaptation occurring in the French population of Apollo and his vulnerability under changing climate. This could be the starting point for further studies investigating these patterns through, for instance, common garden experiments.

INTRODUCTION

Climate change is one of the main growing threats for biodiversity. It affects ecosystems through changes in environmental conditions inducing modifications of biotic interactions, selective pressure, or spatial distribution of favourable habitats (Díaz *et al.* 2019). According to climate modelling, the global mean temperature has already increased by 1.1 degrees since the pre-industrial period (1880-1900) (IPCC, 2022) and will continue to increase in the years to come as well as the frequency of extreme climatic events, affecting ecosystem functioning as a whole (Christidis *et al.* 2015). There are already some evidence of the impact of climate change at community, species and population levels like decline, extinction, shift in the phenology or modification of the geographic range (Chen *et al.* 2011; Hill *et al.* 2021). Indeed, in response to changing environments, populations can move or adapt (Chevin *et al.* 2010).

Investigating and predicting the effects of climate change on biodiversity through the development of diverse models was, and still is, one major objective of many researches. At or above the species level, species distribution modelling (SDM) has been widely used to predict the spatial distribution of species or communities to changing environmental conditions by quantifying the correlation between current distributions of species and environmental drivers (Miller 2010). These models identify environmental conditions which are required for species and then project the shift in the spatial distribution of species, using prediction of future climatic conditions (Miller 2010). However, considering for example the intraspecific variation in the form of standing adaptive genetic variation (SAGV) - the pre-existing genetic variation at loci involved in adaptation (Chhatre *et al.* 2019) - seems necessary since natural forces such as selective pressure often varies across space inducing genomic divergences between populations due to specific genetic - environmental associations (Kawecki & Ebert 2004). Indeed, the presence of local adaptation has been demonstrated in many species. It is characterized by genetic variations between populations resulting from evolutions to locally optimal phenotypes in response to local environment (Hereford 2009). SAGV can therefore be non-randomly distributed along the geographic area of the species, and as a result, the impacts of climate change may vary among populations (Razgour et al. 2019).

Recent studies have tackled this limitation by using genomic approaches which incorporate adaptive genetic variation to identify the most at-risk populations facing climate change (Fitzpatrick & Keller 2015). Indeed, maladaptation of populations - when phenotypes of individuals are distant from the optimal phenotype in a given environment - depends on both the selective pressure induced by changing environment and the genetic composition of populations (Brady *et al.* 2019). Studies working on maladaptation generally first search for gene - environment associations using genetic-environment association (GEAs) methods (Forester *et al.* 2018; Razgour *et al.* 2019) such as Redundancy Analysis (RDA) (Legendre & Legendre 2012). However, discriminating genetic variations associated with environmental factors from genetic variations associated with neutral mechanisms like historical range movements of species or isolation by distance (IBD) - the increase of genetic differentiation between populations as the geographical distance increase (Wright 1943) - can be difficult (Rellstab *et al.* 2015). Indeed, neutral processes such as genetic structuration or IBD might, in some cases, produce similar patterns to those resulting from gene - environment associations (De Mita *et al.* 2013). To improve the identification of actual adaptive genetic variation linked with environmental variations, these models can be corrected for neutral processes (Rellstab *et al.* 2015). The identified genetic variation correlated with environmental drivers with GEA's method is then used to predict the mismatch of populations between current and required adaptive genomic compositions to maintain their current fitness under changing climatic conditions (Capblancq *et al.* 2020a). Various metrics of climate maladaptation, such as the genomic offset (Fitzpatrick & Keller 2015), can be used to predict the response of organisms to changing environments.

Mountain biodiversity is especially vulnerable to climate change (Parmesan 2006). Indeed, the modifications of environmental conditions have been and will continue to be particularly important in mountain habitats with rising air temperatures and radiative forces. They induce more frequent long-lasting droughts, the reduction of the ratio snow to rainfall and the ice sheet (Nogués-Bravo *et al.* 2007; Huss *et al.* 2017). These changes can lead to negative effects on cold-adaptive populations and species through shifts in the phenology, habitat loss or direct biological effects (Pauli *et al.* 1996; Descimon *et al.* 2006). It is then expected that populations of mountain habitats will face a more severe maladaptation due to climatic changes.

To investigate this assumption, we conducted a study on the demes of an iconic and endangered grassland butterfly of Eurasian mountains, the Apollo, *Parnassius apollo.* This species and more widely ectotherms species are expected to be particularly vulnerable to climatic changes especially to temperature fluctuations since they cannot regulate their body temperature by using metabolic heat (Sheridan & Bickford 2011; Kingsolver *et al.* 2013). Besides, his long-standing decline since the end of The Last Glacial Maximum (LGM) – exacerbated since the last 70 years - could highlight his vulnerability to climate change (Nakonieczny *et al.* 2007). Considering this, we suspect that this species could be seriously subject to future maladaptation due to climatic changes. Given the large distribution of Apollo in Eurasia and more locally in the studied area, we expect to find signals of local adaptation through genetic variation correlated with climatic variables.

This study aims to: (I) establish if there is adaptive genetic variation linked to climatic drivers in the studied area and identify the climatic factors that explain the most the adaptive genetic variation, then (II) to predict the potential maladaptation of *Parnassius apollo* to future climatic conditions with metrics considering the adaptive genetic variation.

MATERIALS AND METHODS

Study system and sampling

The Apollo, *Parnassius apollo*, is an iconic mountain butterfly species. It is present in small populations widely distributed across Eurasia in open, rocky Palearctic habitats (Brommer & Fred 1999). In France, it is mostly in mountains between 1000 and 2400 m.a.s.l (meters above sea level). This species is univoltine, its eggs winter on stems under the snow protecting them from frost events, its caterpillars feed mostly on *Sedum sp* and imago flight period is from May to September (Deschamps-Cottin *et al.* 1997). Its population size dramatically declined during the past 70 years to the point that this species is classified as “vulnerable” by the IUCN and protected in Europe by the annexe IV of the appendix II of the Habitat Directive of the European Union. This decline can be attributed to a combination of factors such as habitat loss, genetic erosion or long-term climatic changes (Nakonieczny et al. 2007). Examples of local extinction due to climatic events have been identified in France like in Causse of Larzac in the 1980’s (Descimon *et al.* 2006).

The sampling was conducted in 2019, 2021 and 2022. 317 butterflies were sampled accross 96 localities spread into French mountains (Ardèche (2), Ariège (4), Auvergne (6), Bauges (7), Belledonne (4), Cévennes (2), Chartreuse (9), Devoluy (3), Ecrins (9), Hautes-Pyrénées (3), Haute-Savoie (3), Jura (9), Mercantour (4), Mont Ventoux (5), Pyrénées-Atlantiques (6), Pyrénées-Orientales (6), Queyras (9) and Vercors (5)). For each locality, 1 to 10 individuals were sampled. The distance between the localities of the same massif greatly varies as well as the distance between the massifs. To reduce the impact of this study on the population of Apollo, only one-leg was collected for each butterfly to conduct the DNA extraction. For each individual, the coordinates of the capture point were also collected.

Genetic data and SNP calling

The genetic dataset was obtained by double-digest Restriction-site Associated DNA sequencing method (ddRADseq) (Peterson *et al.* 2012) using two enzymes *sbf I* (R1) and *msp I* (R2). Further filters were performed using *Stacks* software (Catchen *et al.* 2013). Genetic fragments (R1 and R2) were filtered based on their quality score (> 90%) and length, keeping only fragments of 110bp and cutting the longer ones. Remaining fragments were demultiplexed. Then, for every locus within an individual, R1 and R2 fragments were merged and mapped onto the reference genome of Apollo (Podsiadlowski *et al.* 2021) and SNPs calling was performed.

Finally, we did - using *vcftools* software (Danecek *et al.* 2011) - a series of filters to retain only the most accurate variants for further analysis. We retained SNPs present in at least 70% of all individuals and minor allele count of 3 (avoiding rare alleles and keeping only loci with variation), with reads coverage between 5 and 150 (to reduce sequencing error and repeated sequences/pseudogenes) and only 1 SNP every 220 pb (reducing linkage disequilibrium between SNPs). From this SNPs dataset, we removed individuals with more than 40% missing data and we were left with 1999 SNPs for 293 individuals. For genetic structure, we performed a more stringent filtering to reduce the bias of position in the PCA space due to missing data (Reeves *et al.* 2016). We kept SNPs only present in 95% of individuals (558 SNPs). Missing data were imputed with the median genotype of all the individuals.

Climatic data

Environmental data at high resolution were used to explore the influence of climatic factors on the genetic variation and to identify the main drivers of this variation. Present and future environmental data were extracted from WorldClim (Fick & Hijmans 2017) - at 30 arc second (~1km²) resolution - for every individual using the coordinates (longitude and latitude) of the capture point. Indeed, a previous study of Brommer & Fred (1999) on Apollo estimated - using capture- mark- recapture - that adults were not so mobile with a mean distance between the capture and recapture points of less than 400 meters. 8 variables were pre-selected including 6 bioclimatic (bio1, 2, 3, 4, 12, 15) variables (see table 1) and 2 other variables (snow cover days and mean humidity). These variables are known to be involved in local adaptation in some species of butterflies or to be relevant according to the ecology of Apollo (Roy *et al.* 2015; McDermott Long *et al.* 2017) and they are predicted in the future by climatic models. To avoid multicollinearity, we only retained bio 1, 2, 4, 12 and 15 (see table 1) because they were not too correlated (< 0.7).

Table 1: Signification of the climatic variables used.

Variable	Signification	Abbreviation	Analysis
Bio1	Mean annual air temperature	an_tc	PCA, pRDA, RDA, SDM
Bio2	Mean diurnal air temperature range	diur_range_tc	PCA, pRDA, RDA, SDM
Bio3	Isothermality		SDM
Bio4	Temperature seasonality	tc_se	pRDA, RDA, PCA
Bio7	Annual range of air temperature		SDM
Bio8	Mean daily mean air temperatures of the wettest quarter		SDM

Bio9	Mean daily mean air temperatures of the driest quarter		SDM
Bio12	Annual precipitation amount	an_precip	PCA, pRDA, RDA, SDM
Bio15	Precipitation seasonality	precip_se	PCA, pRDA, RDA, SDM
Hurs-mean	Mean monthly near-surface relative humidity		SDM
SfcWind_mean	Mean monthly near-surface wind speed		SDM

Present climatic data are the means for the period 1970-2000. Future climatic data are the mean values predicted by the climatic model *IPSL-CM6A-LR* SSP3-7.0 (O'Neill *et al.* 2016) for the years 2060 to 2080. Present and future climatic data were standardized together (means and standard deviations) allowing the comparison of indices based on these climatic factors (Legendre & Legendre 2012).

We also performed a principal component analysis (PCA), using core R-package (R core Team 2022), with the 5 climatic factors retained for the 293 individuals to characterize the climatic environment of the sampling (figure 1) and identify the mains climatic gradients.

Apollo 's geographic range studied

To interpolate in space the indexes calculated in this study, we needed a map of Apollo's presence in the French and also in the Pyreneans Spain and a part of Swiss and Italian Alps (extent: longitude [-2;8] and latitude [42;47]). This map was estimated with SDM analysis (using crossvalidation) from Maxent software (Phillips *et al.* 2004). We performed this analysis by using filtered Gbif occurrences of *Parnassius apollo* in this area (4704 presence occurrences), 10 environmental factors from Chelsa (Karger *et al.* 2017) that were not too correlated (<0.7) including: bio1, 2, 3, 7, 8, 9, 12, 15, hurs-mean and sfcWind_mean (see table 1) and a map of the probability of presence in the same area of *Sedum album*, the most common host plant for Apollo (Deschamps-Cottin *et al.* 1997) that were also obtain by SDM analysis using filtered Gbif occurrences (12000 presence occurrences) and the same factors used above. The geographic range of Apollo was obtained by selecting a minimum presence probability equals to the lower presence probability of a sampled point (all points with probability > 0.2).

Neutral processes

Considering nonadaptive processes when looking for relations between adaptive genetic variation and environmental factors can be useful. For instance, the spatial structuration between individuals due to IBD, or historical demography can, in some cases, generate similar patterns to those observed due to local adaptation (De Mita *et al.* 2013). To account for the IBD, we used the coordinates of each individual (Orsini *et al.* 2013). To account for the historical demography and other neutral processes, we conducted an analysis of the genetic structure. We performed a PCA analysis with the data set containing 558 SNPs using the *pcadapt* R-package

(R core Team 2022). We used for each individual their scores into the two first PCA axes as proxies of the genetic structuration according to Capblancq & Forester (2021)

Variance partitioning

One of the hypothesis in GEA's analysis is that the environment explains some of the genetic variations (Rellstab *et al.* 2015). In order to confirm this, we needed to unravel patterns of genetic variation correlated to the environment from similar patterns arising from genetic structuration or IBD. To do this, we first performed a redundancy analysis (RDA), using *vegan* R-package, with the 1999 SNPs as response variable and the climatic factors, the coordinates and the two first axes of PCA as explanatory variables (full model). Then, we did partial redundancy analysis (pRDA) to visualize the contribution of each pool of variables in the total explained variation. This type of analysis enables to perform a RDA after removing the variance of the response variables (ex: genetic variation) explained by the covariables and also after removing the variance of the explanatory variables (ex: climate) explained by the covariables. To do this, we performed the RDA on the residuals of the two previous RDA: one between the response variables and covariables and one between the explanatory variables and the covariables. Therefore, working on residuals enables to visualize the importance of each pool of factors (climate, geography, and genetic structuration) after removing the effects of others. We did pRDA for each of the three pools of variables. The full model provides the total explained variance and informs on the contribution of each group of factors on the variance by comparing the pRDA results to the full model (table 2). The effect of the geography was marginal, thus, for further analysis we would not consider this variable.

Gene - Environment association

After looking for the importance of climatic variables on the genetic variation, we searched to identify genetic variants associated with some of these factors, according to the procedure in Capblancq and al. 2021. We performed a pRDA with the set of 1999 SNPs as response variable and the climatic factors as explanatory variable after removing the effect of the genetic structure. Then, we identified SNPs with an atypical projection onto the 2 first axes of the pRDA compared to the projection of the majority of SNPs (later called outliers). To realize this identification, we followed the genome scan procedure of Luu et al., (2017) to transform the Mahalanobis distance into a P-value for each variant. To minimize the false discovery rate, we applied a Bonferroni correction and retained only variants with a p-value lower than *5.00 e-6*.

Adaptive space

From the outliers, we performed an RDA to create an adaptively enriched RDA space to identify the main climatic drivers of the adaptive variation and visualize the adaptive genetic variation across the Apollo's geographic range studied. Then, for each RDA axis retained, we calculated an index that uses the scores of the climatic variables along the RDA axis, coming from the relations between the climatic variables and the variants, to interpolate/extrapolate the values of RDA in areas where the climatic data are available according to the formula in Steane *et al.* 2014: $Adaptive\ composition = \sum_{i=1}^{n} a_i\, b_i$

Where $\boldsymbol{a}$ is the value of the standardised climatic factor, $\boldsymbol{b}$ is the score of the climatic factor on the RDA axis considered and $\boldsymbol{i}$ corresponds to the different climatic factors used to construct the RDA model. These RDA values were calculated for the first axis of RDA along the Apollo's geographic area studied. We also estimated these RDA's values into future climatic conditions by using the future climatic dataset of variables.

The potential impact of climate change

The predictions of the adaptive composition were used to represent the range and density of RDA current and future values along the geographic range. The range of values that regroups 99% of the current data was considered to be the current climatic range of the adaptive gradient. Then, we identified as "outsiders" all the future predicted values which were outside this range.

These RDA values were also used to estimate the change in the adaptive composition needed to face climatic changes along the studied area. We calculated a metric proposed by Fitzpatrick and Keller (2015) called "genomic offset". To do so, we calculated the Euclidian distance between the current and future adaptive compositions for each pixel. The present and future RDA values projected on the map (figure 5) were scaled together between -1 and 1 to enable the comparison between them. The values of genomic offset were also scaled between -1 and 1 (figure 6).

Population level metrics: PAI and SAGV

We also calculated two other metrics at the massif level to further investigate the standing adaptive genetic variation across the studied area. We based these analysis at the massif level to have a sufficient number of individuals (at least 5) to calculate metrics of genetic diversity (Sherpa *et al.* 2022). We removed the Hautes-Pyrénées massif because only 2 individuals remained after the filtration. The 17 other massifs were retained with 7 to 28 individuals. We calculated the SAVG per massif as the mean frequency (p*q) of the SNPs outliers. We also wanted to estimate the extremeness of the massifs along the adaptive gradient using the PAI index which is the mean of the absolute differences for each SNPs outliers between the allele frequency of the massif compared to the mean allele frequency of all the massifs for the considered SNP (Capblancq *et al.* 2020b). However, we found that this metric can be biased by a sampling effect. Indeed, it estimates the extremeness of a massif compared to other massifs along the adaptive gradient. However, if the majority of the massifs are extreme then the most important values of PAI will be the actual less extreme massifs. We proposed another metric not biased by the sampling, the Specificity index (SI). We calculated for each variant of a massif the absolute difference between his allele frequency and 0.5 and then, we mean these results for all the variants per massif: $Specificity\ index = \frac{\left(\sum_{k=i}^{n} |AF_i - 0.5\,|\right)}{\sum i}$

Where AF is the allele frequency and *i* is the SNP considered. This index ranges from 0 (the mean allele frequency is 0.5 in the massif) to 0.5 (alleles are fixed in the massif regardless of the frequency: 0 or 1). We also calculated these 2 metrics for the neutral loci to compare the patterns observed between adaptive and neutral loci. Neutral loci were obtained by excluding the outliers from the dataset of 1999 SNPs. SGV and SI metrics were then calculated on the 1967 neutral loci.

RESULTS

Climatic environment

Figure 1 represents the results of the PCA realised on the sampling points and their climatic environment. The position of the points on this space depends on their values for the 5 climatic factors considered. We retained only the two first axes because they explain the majority of the variance (72.4%). Two main climatic gradients can be identified. PC1 axis is driven by the annual temperature, diurnal range of temperature and seasonality of the precipitation

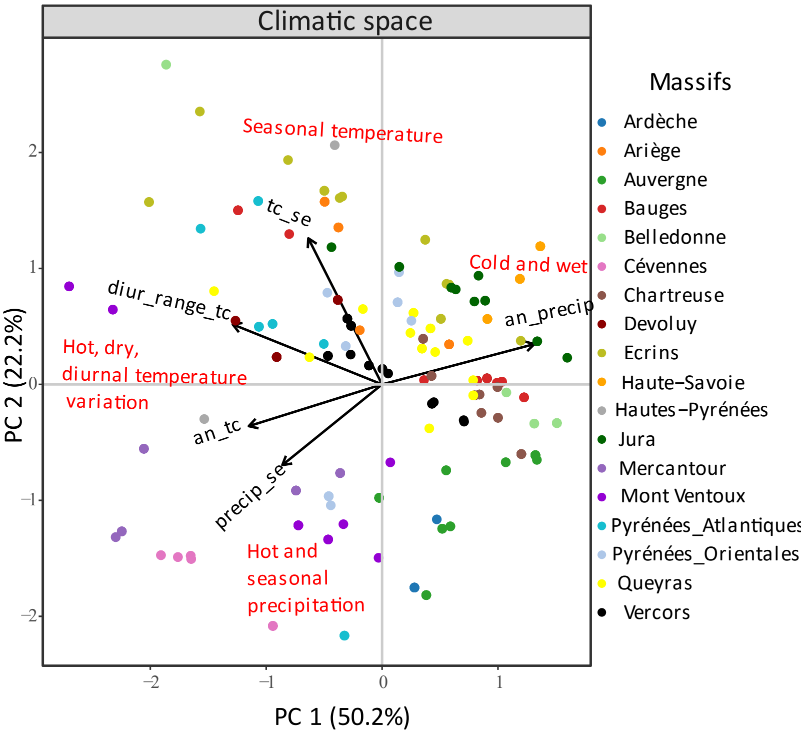

Figure 1: Principal component analysis where the black arrows represent the climatic factors and each points represent a sampled locality. The colours of the points correspond to the different massifs.

opposed to the annual precipitation. Along this axis, we can identify groups of massifs with, on the negative values, the Cévennes, Mont Ventoux, Pyrénées-Orientales or Mercantour and along the positive values the Jura, Bauges, Chartreuse, Belledonne or Queyras. On the second PCA axis, we can identify a gradient mostly driven by the seasonality of the temperature and the diurnal range of temperature on one side and by the seasonality of the precipitation on the other. Groups of massifs can also be identified with Ariège, Ecrins and Pyrénées-Atlantiques for example along the positive values and with Auvergne, Mont Ventoux or Cévennes along the negative values. Besides, within a massif, the position of the points can be very different such as the Vercors, Pyrénées-Atlantiques or the Mont Ventoux pointing the differences of the climatic conditions intra-massif in addition to the differences inter-massif.

Genetic structuration

Figure 2 represents the results of the PCA realised on the genetic markers. The position of the individuals in this space depends on their genetic composition for the 558 SNPs considered. We only kept the 2 first axes because they explain the majority of the variations (65.2%). We can see that 4 genetic groups can be identified: Auvergne, Pyrénées, Ardèche/ Cévennes and the Alps.

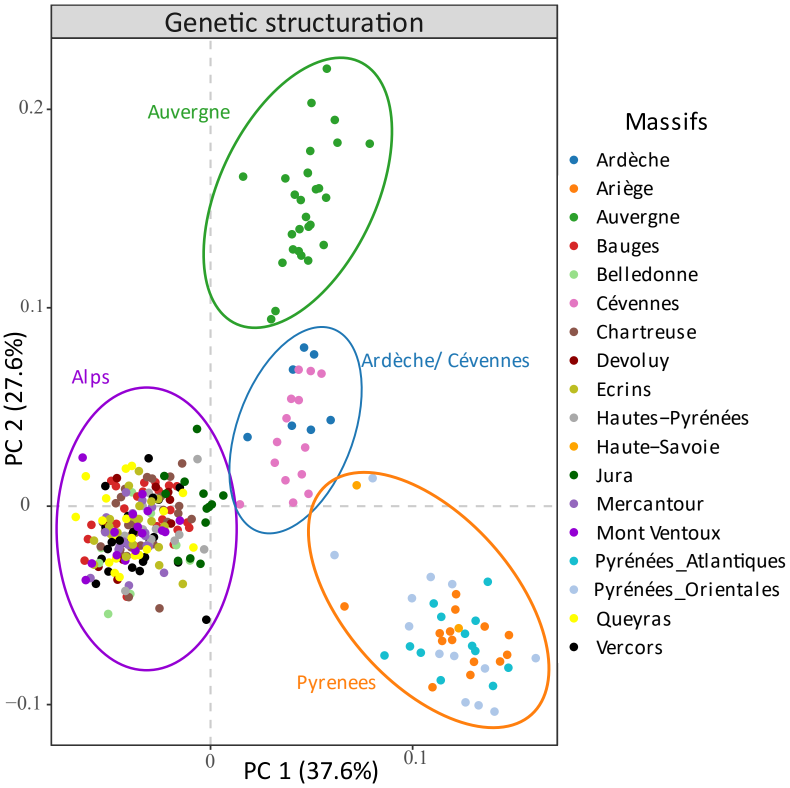

Figure 2: Principal component analysis on the genetic markers where each point represents a sampled locality. The colours of the points correspond to the different massifs.

Variance partitioning

The results of the RDA full model and pRDA for the climatic, genetic structuration and geographic predictors are shown in table 3.

Table 3: Variance partitioning between the climate, structure, and geography. The full model is used as a reference to calculate the percentage of explainable variance for each pool of variables.

RDA and pRDA models	**Inertia**	**R²**	**Explainable variance (percentage)**	**Total variance (percentage)**
Full model: F ~ *clim.* + *struc.* + *geog.*	199.1	0.10	100	10.0
Pure climate: F ~ *clim.* | (*struc.* + *geog.*)	52.3	0.026	26	2.6
Pure structure: F ~ *struc.* | (*clim.* + *geog.*)	30.9	0.016	16	1.6
Pure Geography: F ~ *geog.* | (*clim.* + *struc.*)	18.8	0.009	9	0.9
Confounded *climate* / *structure* / *geography*	97.1		49	4.9
Total unexplained	1800			90.0
Total inertia	1999			100

The full model explains 10% of the genetic variation. The climate represents 26% of the explainable genetic variation (2.6% of the total variation), the structure represents 16% (1.6% total variation) and the geography only 9% (0.9% total variation). Almost half of the explainable variance cannot be separated from those 3 pools of variables suggesting a strong covariation between the neutral mechanisms and climatic drivers as expected (De Mita *et al.* 2013).

Gene - environment association and adaptive space

The pRDA conditioned only by the structure was performed and only the two first axes were retained (RDA1 35% and RD2 25%) to select the outliers. After performing the genome scan procedure, 32 SNPs were identified to be the most correlated to climatic drivers and considered as outliers. The adaptive enriched space was realised on these outliers (figure 3). Figure 3 shows the position of the outliers and the position of the individuals, based on their allelic frequencies for the considered outliers, along the first two RDA axes. The vast majority of the variance is explained by the first RDA axis (86.5% versus 5.1% for

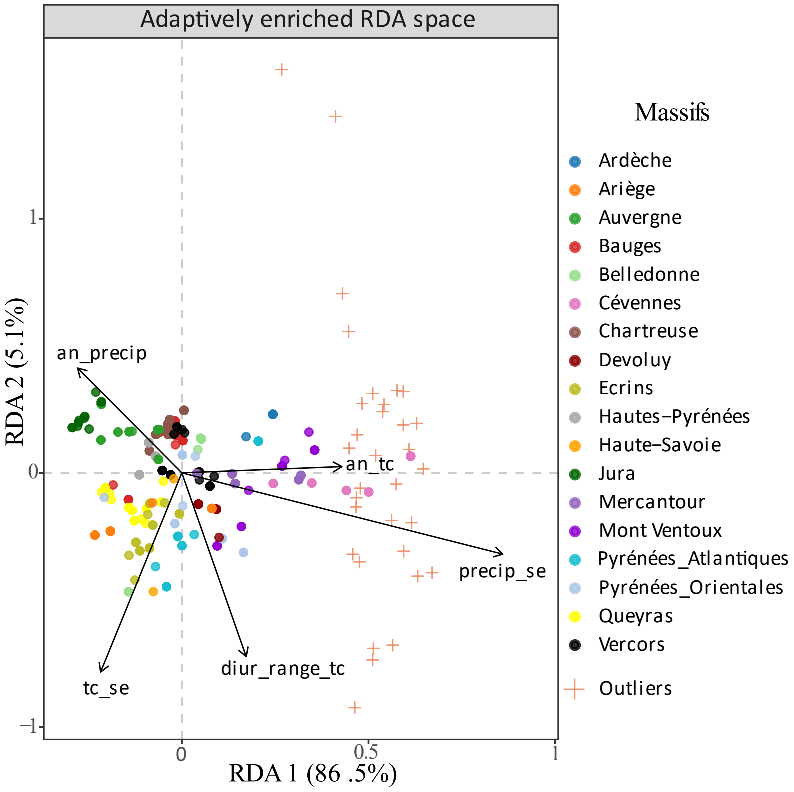

Figure 3: RDA analysis with the 32 outliers as response variable and the 5 climatic factors as explainable variables. The black arrows represent the climatic factors, the orange crosses represent the position of the outliers on this RDA space and the points represent the position of the sampled localities. Colours correspond to the different massifs.

RDA2). Therefore, for further analysis we only kept this axis. We can observe that the first RDA axis is mostly correlated with the seasonality of the precipitation and the annual temperature on one side and with the annual precipitation and seasonality of the temperature on the other. Some massifs are grouped together such as the Cévennes, Mercantour, Ardèche and Mont Ventoux along the positive value of RDA1 and on the other side, the massifs of Auvergne, Jura, Ariège and Queyras are also grouped. Furthermore, some massifs are in the middle of the gradient such as the Northern Alps with Chartreuse, Bauges, Belledonne, Vercors and the Pyrénées-Atlantiques and Orientales. We can see that these results are quite similar to those of figure 1. Indeed, the groups of massifs and their position along the climatic variables are comparable between these two analyses.

The potential impact of climate change

The interpolation/extrapolation of the values of RDA1 (adaptive composition) in current and future climatic conditions along the studied area and characterization of the current climatic range of the adaptive gradient are presented in figure 4. First, we can see that the densities of current and future RDA values are quite different. Indeed, a shift of values seems to occur with a more important density of higher values in the future. Moreover, a significant proportion of RDA values in the future - 7.9 % - is outside the current climatic range of the adaptive gradient.

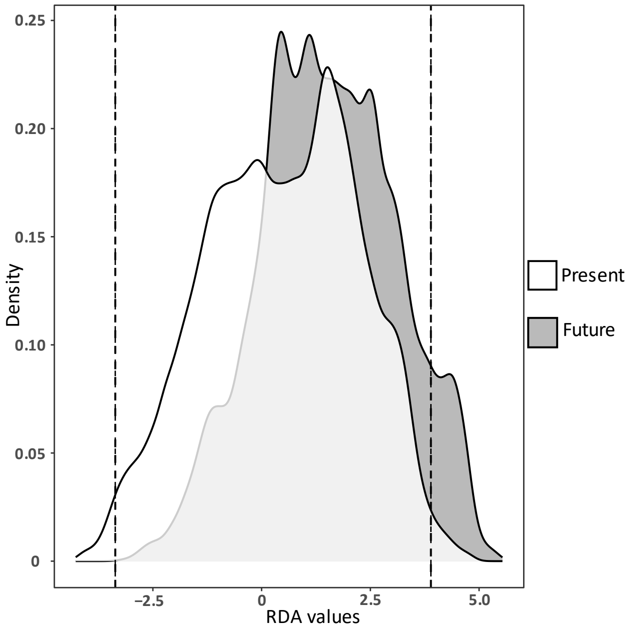

Figure 4: Representation of the RDA values along RDA1. Colours correspond to the present (white) and future (grey). Dotted lines represent the interval where 99% of the values are present.

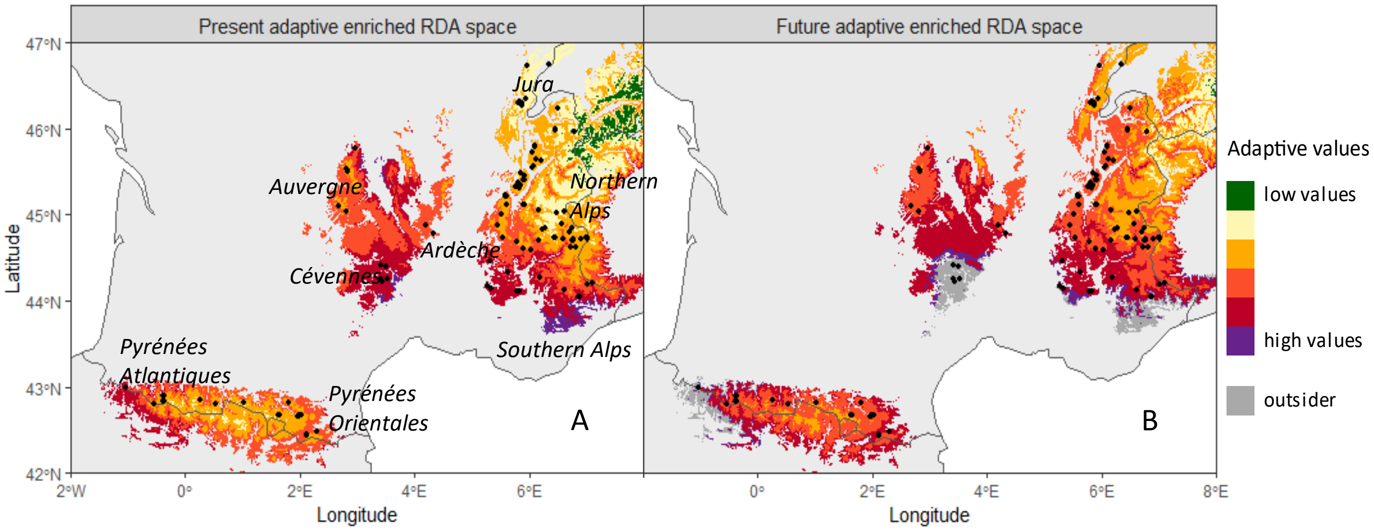

Figure 5: Maps representing the present (A) and future (B) adaptive enriched RDA spaces along the studied area. Colours represent the range of the RDA values. Black points represent the localities sampled.

Then, we projected the predicted current and future adaptive compositions along the studied area to visualize the adaptive gradient and identify the outsider's area (figure 5). We can see on figure 5.A that areas in warm/high seasonal precipitation, low altitude such as the Cévennes, Ardèche and Southern Alps have higher RDA values than the cold/high seasonal temperature, high precipitation, and high-altitude areas like the Northern Alps, including the Jura, and the central Pyrenees. The future adaptive gradient, figure 5.B shows overall that every area will face an increase in the RDA values. This observation is particularly verified in areas with currently low RDA values such as the Jura. Moreover, in this figure, we can see that the outsider's areas are exclusively in areas with current high RDA values like the Cévennes.

We also calculated and projected the genomic offset in the studied area. These results are presented in figure 6. We can see that the change in adaptive allelic frequency required will be much more important in the Northern Alps than in the South, quite high and homogeneous in the Pyrenees with higher change required in the Pyrénées-Atlantiques and more important in the South of the Central Massif (Cévennes, Auvergne, Ardèche) than in the North.

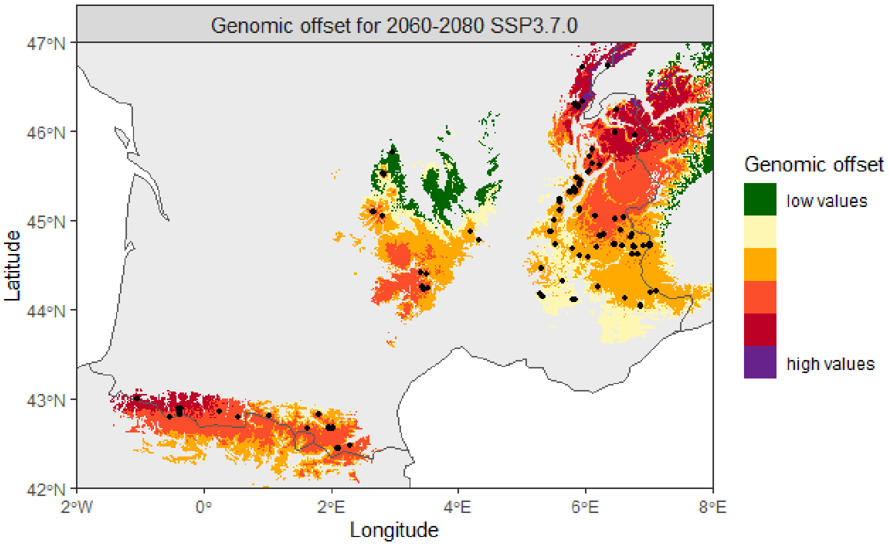

Figure 6: Map representing the values of genomic offset along the studied area. Colours represent the range of the genomic offset values. Black points represent the localities sampled.

Table 3: Massif-level indices of the variation and fixation of the standing adaptive (SAGV/ SI adaptive) and neutral (SGV/SI neutral) genetic variation.

Massifs	**SAGV**	**SI adaptive**	**SGV**	**SI neutral**
Cévennes	0.202	**0.189**	0.038	**0.446**
Ardèche	0.050	**0.428**	0.031	**0.454**
Mont Ventoux	0.032	**0.463**	0.052	**0.430**
Pyrénées-Orientales	0.030	**0.462**	0.046	**0.440**
Chartreuse	0.027	**0.468**	0.053	**0.428**
Devoluy	0.027	**0.469**	0.051	**0.432**
Mercantour	0.026	**0.471**	0.051	**0.431**
Ecrins	0.026	**0.472**	0.052	**0.431**
Bauges	0.025	**0.472**	0.053	**0.429**
Ariège	0.025	**0.468**	0.045	**0.436**

Then, we calculated the SAGV and the specificity index (SI) for the outliers and the SGV and SI for the neutral SNPs. The results of these four indexes are presented in table 3. We can see that the values of SGV and SI for the neutral SNPs vary much less along the massifs than the SAGV and adaptive SI. We overall observed high values of SI

Belledonne	0.024	**0.472**	0.050	**0.432**
Pyrénées-Atlantiques	0.023	**0.474**	0.045	**0.436**
Queyras	0.022	**0.476**	0.052	**0.430**
Haute-Savoie	0.021	**0.477**	0.052	**0.430**
Vercors	0.021	**0.477**	0.054	**0.428**
Auvergne	0.020	**0.470**	0.030	**0.457**
Jura	0.012	**0.485**	0.045	**0.437**

and low values of SGV. However, it seems that the SGV is less important in the Central Massif than in other massifs. Concerning the SAGV, the variations are much more important ranging from the Cévennes with high values (0.202) to the Jura with low values (0.012). The same observation can be made with the adaptive SI with values ranging from the Cévennes with lower values (0.189) to the Jura with higher values (0.485). Besides, some massifs can also be identified with overall higher values of SAGV and lower of SI: the Cévennes, Ardèche, Mont Ventoux and Pyrénées-Orientales. On the contrary, we can notice that the Jura has the lowest values of SAGV and the highest of SI. The rest of the massifs is intermediate for these metrics.

DISCUSSION

Overview of the genetic - environmental association

The first goal of this study was to investigate the presence of local adaptation to climate within the studied area of Apollo and disentangle the drivers of genetic variations. First, it seems that there are some important climatic variations in the studied area with the identification of two climatic gradients in figure 1. These climatic variations indicate that the selective pressure could be different within the studied area which may lead to local adaptation. Besides, a significant proportion of genetic variation could be attributed to the 5 climatic factors retained (table 2), suggesting that a part of the genetic variation can be explained by the climatic variables. These two elements suggest that along the studied area, there might be genetic environment association and then potentially local adaptation like in Hecht and al 2015 where 5.8 % of the total genetic variance were explained by the climate and this was sufficient to identify 568 loci close to genomic regions potentially involved in adaptation (outliers).

The results of the GEA analysis seem to confirm this assumption with the identification of 32 SNPs that are greatly correlated to climatic factors. These outliers are mostly correlated to the first axis of RDA (85%) which is driven by a similar climatic gradient (PC1) identified in figure 1 which is the seasonality of precipitation and annual temperature on one side and annual precipitation and seasonality of temperature on the other. Moreover, the results of figures 1 and 3 showed that the positions of the individuals per massifs in the climatic space and adaptive space are quite similar. This result could confirm that the loci outliers greatly characterise the first axis of the climatic space. Besides, this gradient seems to be the result of the confluence of several gradients such as the continentality with colder temperature in continentality area (Jura), the elevation with colder and more seasonality of temperature and high precipitation regions in high altitude (Centre of Pyrenees and Northern Alps) and the latitude with warmer temperature and higher seasonality of precipitation in the South (Southern Alps). Overall, this climatic gradient is a common gradient found in many studies especially the ones focusing on mountain species (Karl *et al.* 2010; Muir *et al.* 2014; Capblancq *et al.* 2020b). Moreover, the results are in accordance with previous studies working on *Parnassius apollo* (Ashton 2002; Sbaraglia *et al.* 2023) showing that the temperature and the precipitation are the main drivers for the favourability of the habitat.

Patterns of adaptation

The previous results suggest that along the most important climatic gradient identified in the studied area, there seems to be some signals of adaptation. The results of table 3 enable to further investigate the signals identified. Indeed, under a pattern of directional selection, we expected to find a low SAGV and a high adaptive SI in both extremes of the climatic gradient (Holsinger 2001). In fact, the selective pressure is expected to be greater, inducing a strong selection resulting in an important allelic fixation for the adaptive loci. This would lead to a reduction of the SAGV and an increase of the adaptive SI as found in other studies (Chhatre *et al.* 2019; Capblancq *et al.* 2020b). However, we did not find this exact pattern along the studied area. Indeed, we found this pattern for the extreme side of cold/ seasonal temperature and high precipitation of the gradient with the Jura for example but in the other extreme part of the gradient with Cévennes, Mont Ventoux or Ardèche for example, we did not find a reduction of the SAGV but on the contrary an increase. These results are quite surprising, especially because the neutral genetic variation (SGV) is low in the majority of these extreme massifs. Indeed, if the amount of neutral genetic variation were high, this could have been an evidence that something is happening in these massifs and impacts the genetic variation like a high migration rate (Berthier *et al.* 2006; Alcala *et al.* 2013). This migration would have maintained an important genetic variation on the neutral markers and on the adaptive ones if the gene flow was maladaptive like in Aubree *et al.* 2023. However, this was not the case in the studied area according to the SGV.

One hypothesis to explain these results can be made by looking at the ancestral movements of Apollo in this area. Indeed, Kebaïli and al. 2022 found that, between the last interglacial (LIG) and the last maximum glacial (LGM), climatic conditions in low altitude were favourable for Apollo because the air temperature was colder than currently. At the end of the LGM, Apollo have started to climb in altitude to avoid the warming. This ascension in altitude continued until, in some massifs, the Apollo reached the summit like in the Cévennes.

Then, the demes faced an important selective pressure because they are exposed to new environmental conditions. Under this selective pressure, some adaptive variants to warmer temperatures and more important seasonality of precipitation might have been selected in the demes. This hypothesis could explain the pattern observed. Indeed, the demes are potentially under ongoing selection that could explain why the SAGV are more important and the adaptive SI less important in the extreme part of the gradient associated with warmer temperatures and higher seasonality of precipitation. This hypothesis seems supported by the pattern observed in the other part of the gradient with the Jura for instance. Indeed, the individuals in this massif have the lowest amount of SAGV and the highest of adaptive SI and it is also the more extreme massifs for the cold and seasonal temperature and amount of precipitation (figure 1). This suggests that the outliers are mainly adaptive to this gradient and some variations present in demes under warmer temperature and more seasonality of precipitation are due to ongoing selection to new climatic conditions.

To conclude, it seems that we found some outliers close to genomic regions involved in local adaptation by detecting a pattern that can be associated to a signal of adaptation to colder and seasonal temperature and important precipitation and also a pattern that can be referring to ongoing selection to warmer temperature and more seasonality of precipitation. However, we kept in mind that the dataset of genetic markers is quite low and that they only represent a very small proportion of the genome (less than 0.015%) due to the use of ddRADseq method. However, we still could have identified some signal of adaptation. Indeed, we made the hypothesis - according to the infinitesimal model suggested by Fisher in 1919 - that most phenotypes are the result of a huge number of genes with additive effect and not only one gene for one phenotype. Therefore, we surely did not find all the markers close to genomic regions potentially involved in the adaptation for this gradient. We also might not have identified all the gradients of local adaptation. Indeed, the other climatic gradient (PC2) in figure 1 is not identified by our GEA analysis. This could be due to an absence of adaptation to this gradient in the studied area or due to a weaker signal that was not detected by our analysis. Besides, we cannot exclude the hypothesis that a part of the signal of adaptation identified is biased by the discovery of false positives. Indeed, the important SAGV in the Cévennes could be explained by the fact that some outliers identified are neutral rare loci only present in this massif and because the Cévennes is the extreme part of one side of the gradient, this could lead to false positive if these alleles are only present in this massif. Nevertheless, we are quite confident in the general pattern discovered because we corrected the identification of outliers by the genetic structure (using pRDA) and by the Bonferroni correction to minimise the false positive.

Moreover, all the massifs with warmer temperature and important seasonality of precipitation (figure 1) have the same pattern of respective higher values of SAGV and lower values of SI and are in the same side of the adaptive enriched RDA space despite the fact that they are not closer genetically than other massifs according to the genetic structure analysis (figure 2).

The vulnerability of Apollo to climate change

The second goal of this study was to investigate the potential maladaptation of Apollo to climate change. Indeed, major climate changes are expected to occur in the future and especially in altitude area (Nogués-Bravo *et al.* 2007; Huss *et al.* 2017). Thus, we suspected that the Apollo could be greatly impacted by climate change. First, we identified that in the future, 2060-2080 under the SSP 3-7.0, the adaptive space will be significantly different from the current adaptive space. Indeed, the results of figure 3 showed an important shift of RDA values to higher values in the future, indicating that the demes will face important climatic changes with warmer temperature and more seasonality of precipitation and will require a change in the adaptive genetic composition to face it. This pattern is overall observable for all the areas but mostly in current areas with low RDA values (figure 5). This result indicates that current demes adapted to cold and seasonal temperature and high precipitation will need an important change in the adaptive composition due to more severe changes in the climatic conditions. Indeed, the genomic offset of these areas such as the Pyrenees and Northern Alps is important. This result is in accordance with other studies indicating that species in altitude areas are predicted to be more impacted by climate change (Parmesan 2006). Massifs with high genomic offset and isolated like the Jura are also expected to be particularly vulnerable to climate change because they will need an important change in adaptive genomic composition but they are isolated from other massifs so the adaptive gene flow could be limited or impossible. Thus, this could have been interesting to look at other metrics which take into account the distance between massifs like the geographic offset proposed by Capblancq and al 2020b to identify these isolated areas.

We also found that some of the demes will, in the future, face climatic conditions that are not currently encountered by Apollo in the studied range (called outsiders). A part or the totality of these massifs - Cévennes, Southern Alps, Western of the Pyrenees - will be outside the current adaptive climatic range identified. This indicates that these climatic conditions have never been encountered by Apollo’s population in France and thus, we suspect that the fitness of these demes will be impacted and potentially drive the demes to extinction due to an incapacity to migrate in altitude anymore. However, we also found that the demes in warmer and more seasonal precipitation regimes may be under ongoing selection, we might expect that they will select new adaptive genetic variation to these extreme future climatic conditions.

Overall, we identified two types of areas that are mostly vulnerable to climate change: the areas with higher values of genomic offset: the Pyrenees, Northern Alps and South of the Central Massif, and the outsider's areas: the Cévennes, Southern Alps and Western Pyrenees. However, we kept in mind that the information given by the values of genomic offset, without any link with the fitness, are limited (Capblancq *et al.* 2020a). Indeed, we have no clue if a genomic offset of 0.1 will be associated to an important reduction of fitness or on the contrary if a genomic offset of 5 will have a small impact on the fitness. Thus, the difference of values between areas are more informative than the values themselves. Furthermore, other adaptive gradients could occur in this species so values of maladaptation could be different with the addition of other gradients. Finally, the prediction of maladaptation is based on the hypothesis that the relation gene - environment identified is not the result of false positives. It could be interesting, for further studies, to use common garden experiments to test the relation gene - environment identified and to link the genetic offset with the fitness.

Ideas of management such as assisted gene flow (AGF), "the managed movement of individuals or gametes between populations within species ranges to mitigate local maladaptation in the short and long term"(Aitken & Whitlock 2013), could emerge from these results to mitigate the maladaptation due to climate change. Indeed, some examples of successful AGF are present in the literature (Dixon *et al.* 2015; Hagedorn *et al.* 2021). However, it seems that there is no management that could directly assist the outsider's area because they will encounter climatic conditions outside the adaptive range identified. Besides, we also need to keep in mind that many other factors could influence the future of Apollo in this area. Indeed, for example, the treeline will also change due to climate change with an expected migration of trees in altitude (Smith *et al.* 2009). Thus, the favourable habitats for Apollo are expected to be reduced. Moreover, these climatic changes will impact other species and we might expect that the close interaction butterflies/host plant can be disturbed, such as for the marsh fritillary species (Brunbjerg *et al.* 2017), impacting as well the repartition of favourable habitats.

CONCLUSION

This study is a first step to a better comprehension of the adaptive patterns occurring in Apollo along the Pyrenees, Western Alps, and Central massif. Indeed, our results seem to show a pattern of adaptation to colder and seasonal temperature and important precipitation plus an ongoing pattern of adaptation to warmer and seasonal precipitation. Moreover, our study successfully incorporates genomic data to investigate the vulnerability of Apollo to climate change. Indeed, some areas were identified to be at risk in the future and needing an important change in the adaptative genetic composition or predicted to be outside the current climatic range. This study might be the starting point for further studies confirming the adaptive patterns with common garden and exploring the link between genomic offset and fitness.

BIBLIOGRAPHY

Aitken, S. N., & Whitlock, M. C. (2013). Assisted Gene Flow to Facilitate Local Adaptation to Climate Change. *Annual Review of Ecology, Evolution, and Systematics*, *44*(1), 367–388. https://doi.org/10.1146/annurev-ecolsys-110512-135747

Alcala, N., Streit, D., Goudet, J., & Vuilleumier, S. (2013). Peak and Persistent Excess of Genetic Diversity Following an Abrupt Migration Increase. *Genetics*, https://doi.org/10.1534/genetics.112.147785

Ashton, S., Gutiérrez, D., & Wilson, R. J. (2009). Effects of temperature and elevation on habitat use by a rare mountain butterfly: Implications for species responses to climate change. *Ecological Entomology*, *34*(4), 437–446. https://doi.org/10.1111/j.1365-2311.2008.01068.x

Aubree, F., Lac, B., Mailleret, L., & Calcagno, V. (2023). Migration pulsedness alters patterns of allele fixation and local adaptation in a mainland-island model. *Evolution*, https://doi.org/10.1093/evolut/qpac067

Berthier, K., Charbonnel, N., Galan, M., Chaval, Y., & Cosson, J.-F. (2006). Migration and recovery of the genetic diversity during the increasing density phase in cyclic vole populations. *Molecular Ecology*, *15*(9), 2665–2676. https://doi.org/10.1111/j.1365-294X.2006.02959.x

Brady, S. P., Bolnick, D. I., Angert, A. L., Gonzalez, A., Barrett, R. D. H., Crispo, E., Derry, A. M., Eckert, C. G., Fraser, D. J., Fussmann, G. F., Guichard, F., Lamy, T., McAdam, A. G., Newman, A. E. M., Paccard, A., Rolshausen, G., Simons, A. M., & Hendry, A. P. (2019). Causes of maladaptation. *Evolutionary Applications*, *12*(7), 1229–1242. https://doi.org/10.1111/eva.12844

Brommer, Jon. E., & Fred, M. S. (1999). Movement of the Apollo butterfly Parnassius apollo related to host plant and nectar plant patches. *Ecological Entomology*, https://doi.org/10.1046/j.1365-2311.1999.00190.x

Brunbjerg, A. K., Høye, T. T., Eskildsen, A., Nygaard, B., Damgaard, C. F., & Ejrnæs, R. (2017). The collapse of marsh fritillary (Euphydryas aurinia) populations associated with declining host plant abundance. *Biological Conservation*, *211*, 117–124. https://doi.org/10.1016/j.biocon.2017.05.015

Capblancq, T., Fitzpatrick, M. C., Bay, R. A., Exposito-Alonso, M., & Keller, S. R. (2020). Genomic Prediction of (Mal)Adaptation Across Current and Future Climatic Landscapes. *Annual Review of Ecology, Evolution, and Systematics*, *51*(1), 245–269. https://doi.org/10.1146/annurev-ecolsys-020720-042553

Capblancq, T., & Forester, B. R. (2021). Redundancy analysis: A Swiss Army Knife for landscape genomics. *Methods in Ecology and Evolution*, *12*(12), 2298–2309. https://doi.org/10.1111/2041-210X.13722

Capblancq, T., Morin, X., Gueguen, M., Renaud, J., Lobreaux, S., & Bazin, E. (2020). Climate-associated genetic variation in Fagus sylvatica and potential responses to climate change in the French Alps. *Journal of Evolutionary Biology*, *33*(6), 783–796. https://doi.org/10.1111/jeb.13610

Catchen, J., 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

Chen, I.-C., Hill, J. K., Ohlemüller, R., Roy, D. B., & Thomas, C. D. (2011). Rapid Range Shifts of Species Associated with High Levels of Climate Warming. *Science*, https://doi.org/10.1126/science.1206432

Chevin, L.-M., Lande, R., & Mace, G. M. (2010). Adaptation, Plasticity, and Extinction in a Changing Environment: Towards a Predictive Theory. *PLOS Biology*, https://doi.org/10.1371/journal.pbio.1000357

Chhatre, V. E., Fetter, K. C., Gougherty, A. V., Fitzpatrick, M. C., Soolanayakanahally, R. Y., Zalesny, R. S., & Keller, S. R. (2019). *Climatic niche predicts the landscape structure of locally adaptive standing genetic variation* (p. 817411). bioRxiv. https://doi.org/10.1101/817411

Christidis, N., Jones, G. S., & Stott, P. A. (2015). Dramatically increasing chance of extremely hot summers since the 2003 European heatwave. *Nature Climate Change*, https://doi.org/10.1038/nclimate2468

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., Durbin, R., & 1000 Genomes Project Analysis Group. (2011). The variant call format and VCFtools. *Bioinformatics*, https://doi.org/10.1093/bioinformatics/btr330

De Mita, S., Thuillet, A.-C., Gay, L., Ahmadi, N., Manel, S., Ronfort, J., & Vigouroux, Y. (2013). Detecting selection along environmental gradients: Analysis of eight methods and their effectiveness for outbreeding and selfing populations. *Molecular Ecology*, *22*(5), 1383–1399. https://doi.org/10.1111/mec.12182

Deschamps-Cottin, M., Roux, M., & Descimon, H. (1997). Valeur trophique des plantes nourricières et préférence de ponte chez Parnassius apollo L. (Lepidoptera, Papilionidae). *Comptes Rendus de l'Académie des Sciences - Series III - Sciences de la Vie*. https://doi.org/10.1016/S0764-4469(97)85028-7

Descimon, H., Bachelard, P., Boitier, E., & Pierrat, V. (2006). Decline and extinction of Parnassius apollo populations in France—Continued. *Studies on the Ecology and Conservation of Butterflies in Europe*.

Díaz, S. M., Settele, J., Brondízio, E., Ngo, H., Guèze, M., Agard, J., Arneth, A., Balvanera, P., Brauman, K., Butchart, S., Chan, K. M. A., Garibaldi, L. A., Ichii, K., Liu, J., Subramanian, S., Midgley, G., Miloslavich, P., Molnár, Z., Obura, D., … Zayas, C. (2019). *The global assessment report on biodiversity and ecosystem services: Summary for policy makers*. Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. https://ri.conicet.gov.ar/handle/11336/116171

Dixon, G. B., Davies, S. W., Aglyamova, G. V., Meyer, E., Bay, L. K., & Matz, M. V. (2015). Genomic determinants of coral heat tolerance across latitudes. *Science*, https://doi.org/10.1126/science.1261224

Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, *37*(12), 4302–4315. https://doi.org/10.1002/joc.5086

Fisher, R. A. (1919).—The Correlation between Relatives on the Supposition of Mendelian Inheritance. *Earth and Environmental Science of The Royal Society of Edinburgh*, https://doi.org/10.1017/S0080456800012163

Fitzpatrick, M. C., & Keller, S. R. (2015). Ecological genomics meets community-level modelling of biodiversity: Mapping the genomic landscape of current and future environmental adaptation. *Ecology Letters*, *18*(1), 1–16. https://doi.org/10.1111/ele.12376

Forester, B. R., Lasky, J. R., Wagner, H. H., & Urban, D. L. (2018). Comparing methods for detecting multilocus adaptation with multivariate genotype-environment associations. *Molecular Ecology*, https://doi.org/10.1111/mec.14584

Frichot, E., Schoville, S. D., Bouchard, G., & François, O. (2013). Testing for Associations between Loci and Environmental Gradients Using Latent Factor Mixed Models. *Molecular Biology and Evolution*, *30*(7), 1687–1699. https://doi.org/10.1093/molbev/mst063

Hagedorn, M., Page, C. A., O'Neil, K. L., Flores, D. M., Tichy, L., Conn, T., Chamberland, V. F., Lager, C., Zuchowicz, N., Lohr, K., Blackburn, H., Vardi, T., Moore, J., Moore, T., Baums, I. B., Vermeij, M. J. A., & Marhaver, K. L. (2021). Assisted gene flow using cryopreserved sperm in critically endangered coral. *Proceedings of the National Academy of Sciences*, https://doi.org/10.1073/pnas.2110559118

Hecht, B. C., Matala, A. P., Hess, J. E., & Narum, S. R. (2015). Environmental adaptation in Chinook salmon (Oncorhynchus tshawytscha) throughout their North American range. *Molecular Ecology*, *24*(22), 5573–5595. https://doi.org/10.1111/mec.13409

Hereford, J. (2009). A quantitative survey of local adaptation and fitness trade-offs. *The American Naturalist*, *173*(5), 579–588. https://doi.org/10.1086/597611

Hill, G. M., Kawahara, A. Y., Daniels, J. C., Bateman, C. C., & Scheffers, B. R. (2021). Climate change effects on animal ecology: Butterflies and moths as a case study. *Biological Reviews*, *96*(5), 2113–2126. https://doi.org/10.1111/brv.12746

Holsinger, K. E. (2001). Natural Selection*. In S. Brenner & J. H. Miller (Eds.), *Encyclopedia of Genetics* (pp. 1291–1297). Academic Press. https://doi.org/10.1006/rwgn.2001.1161

Huss, M., Bookhagen, B., Huggel, C., Jacobsen, D., Bradley, R. s., Clague, J. j., Vuille, M., Buytaert, W., Cayan, D. r., Greenwood, G., Mark, B. g., Milner, A. m., Weingartner, R., & Winder, M. (2017). Toward mountains without permanent snow and ice. *Earth's Future*, https://doi.org/10.1002/2016EF000514

IPCC. (2022). *Climate Change 2022: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. https://doi.org/10.1017/9781009325844

Karger, D. N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R. W., Zimmermann, N. E., Linder, H. P., & Kessler, M. (2017). Climatologies at high resolution for the earth's land surface areas. *Scientific Data*, *4*(1), Article 1. https://doi.org/10.1038/sdata.2017.122

Karl, I., Hoffmann, K. H., & Fischer, K. (2010). Cuticular melanisation and immune response in a butterfly: Local adaptation and lack of correlation. *Ecological Entomology*, *35*(4), 523–528.

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

Kebaïli, C., Sherpa, S., Rioux, D., & Després, L. (2022). Demographic inferences and climatic niche modelling shed light on the evolutionary history of the emblematic cold-adapted Apollo butterfly at regional scale. *Molecular Ecology*, *31*(2), 448–466. https://doi.org/10.1111/mec.16244

Kingsolver, J. G., Diamond, S. E., & Buckley, L. B. (2013). Heat stress and the fitness consequences of climate change for terrestrial ectotherms. *Functional Ecology*, *27*(6), https://doi.org/10.1111/1365-2435.12145

Legendre, P., & Legendre, L. (2012). *Numerical Ecology*. Elsevier.

Luu, K., Bazin, E., & Blum, M. G. B. (2017). pcadapt: An R package to perform genome scans for selection based on principal component analysis. *Molecular Ecology Resources*. https://doi.org/10.1111/1755-0998.12592

McDermott Long, O., Warren, R., Price, J., Brereton, T. M., Botham, M. S., & Franco, A. M. A. (2017). Sensitivity of UK butterflies to local climatic extremes: Which life stages are most at risk? *Journal of Animal Ecology*, *86*(1), 108–116. https://doi.org/10.1111/1365-2656.12594

Miller, J. (2010). Species Distribution Modeling. https://doi.org/10.1111/j.1749-8198.2010.00351.x

Muir, A. P., Biek, R., Thomas, R., & Mable, B. K. (2014). Local adaptation with high gene flow: Temperature parameters drive adaptation to altitude in the common frog (Rana temporaria). *Molecular Ecology*, *23*(3), 561–574. https://doi.org/10.1111/mec.12624

Nakonieczny, M., Kedziorski, A., & Michalczyk, K. (2007). Apollo butterfly (Parnassius apollo L.) in Europe: Its history, decline and perspectives of conservation. *Funct Ecosyst Commun*, *1*, 56–79.

Nogués-Bravo, D., Araújo, M. B., Errea, M. P., & Martínez-Rica, J. P. (2007). Exposure of global mountain systems to climate warming during the 21st Century. *Global Environmental Change*, *17*(3), 420–428. https://doi.org/10.1016/j.gloenvcha.2006.11.007

Ohmura, A. (2012). Enhanced temperature variability in high-altitude climate change. *Theoretical and Applied Climatology*, *110*(4), 499–508. https://doi.org/10.1007/s00704-012-0687-x

O'Neill, B. C., Tebaldi, C., van Vuuren, D. P., Eyring, V., Friedlingstein, P., Hurtt, G., Knutti, R., Kriegler, E., Lamarque, J.-F., Lowe, J., Meehl, G. A., Moss, R., Riahi, K., & Sanderson, B. M. (2016). The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6. *Geoscientific Model Development*, *9*(9), 3461–3482. https://doi.org/10.5194/gmd-9-3461-2016

Orsini, L., Vanoverbeke, J., Swillen, I., Mergeay, J., & De Meester, L. (2013). Drivers of population genetic differentiation in the wild: Isolation by dispersal limitation, isolation by adaptation and isolation by colonization. *Molecular Ecology*, *22*(24), 5983–5999. https://doi.org/10.1111/mec.12561

Parmesan, C. (2006). Ecological and Evolutionary Responses to Recent Climate Change. *Annual Review of Ecology, Evolution, and Systematics*, https://doi.org/10.1146/annurev.ecolsys.37.091305.110100

Pauli, H., Gottfried, M., & Grabherr, G. (1996). Effects of climate change on mountain ecosystems—Upward shifting of alpine plants. *World Resource Review*, *8*(3). https://www.osti.gov/biblio/379668

Phillips, S. J., Dudík, M., & Schapire, R. E. (2004). A maximum entropy approach to species distribution modeling. *Proceedings of the Twenty-First International Conference on Machine Learning*, 83. https://doi.org/10.1145/1015330.1015412

Podsiadlowski, L., Tunström, K., Espeland, M., & Wheat, C. W. (2021). The Genome Assembly and Annotation of the Apollo Butterfly Parnassius apollo, a Flagship Species for Conservation Biology. *Genome Biology and Evolution*, *13*(8), evab122. https://doi.org/10.1093/gbe/evab122

R core Team. (2022). R: A language and environment for statistical computing. *R for Statistical Computing*.

Razgour, O., Forester, B., Taggart, J. B., Bekaert, M., Juste, J., Ibáñez, C., Puechmaille, S. J., Novella-Fernandez, R., Alberdi, A., & Manel, S. (2019). Considering adaptive genetic variation in climate change vulnerability assessment reduces species range loss projections. *Proceedings of the National Academy of Sciences*, *116*(21), 10418–10423. https://doi.org/10.1073/pnas.1820663116

Reeves, P. A., Bowker, C. L., Fettig, C. E., Tembrock, L. R., & Richards, C. M. (2016). *Effect of error and missing data on population structure inference using microsatellite data* (p. 080630). bioRxiv. https://doi.org/10.1101/080630

Rellstab, C., Gugerli, F., Eckert, A. J., Hancock, A. M., & Holderegger, R. (2015). A practical guide to environmental association analysis in landscape genomics. *Molecular Ecology*, *24*(17), 4348–4370. https://doi.org/10.1111/mec.13322

Roy, D. B., Oliver, T. H., Botham, M. S., Beckmann, B., Brereton, T., Dennis, R. L. H., Harrower, C., Phillimore, A. B., & Thomas, J. A. (2015). Similarities in butterfly emergence dates among populations suggest local adaptation to climate. *Global Change Biology*, *21*(9), 3313–3322. https://doi.org/10.1111/gcb.12920

Sbaraglia, C., Samraoui, K. R., Massolo, A., Bartoňová, A. S., Konvička, M., & Fric, Z. F. (2023). Back to the future: Climate change effects on habitat suitability of Parnassius apollo throughout the Quaternary glacial cycles. *Insect Conservation and Diversity*, *16*(2), 231–242. https://doi.org/10.1111/icad.12615

Sheridan, J. A., & Bickford, D. (2011). Shrinking body size as an ecological response to climate change. *Nature Climate Change*, *1*(8), Article 8. https://doi.org/10.1038/nclimate1259

Sherpa, S., Tutagata, J., Gaude, T., Laporte, F., Kasai, S., Ishak, I. H., Guo, X., Shin, J., Boyer, S., Marcombe, S., Chareonviriyaphap, T., David, J.-P., Chen, X.-G., Zhou, X., & Després, L. (2022). Genomic Shifts, Phenotypic Clines, and Fitness Costs Associated With Cold Tolerance in the Asian Tiger Mosquito. *Molecular Biology and Evolution*, *39*(5), msac104. https://doi.org/10.1093/molbev/msac104

Smith, W. K., Germino, M. J., Johnson, D. M., & Reinhardt, K. (2009). The Altitude of Alpine Treeline: A Bellwether of Climate Change Effects. *The Botanical Review*, https://doi.org/10.1007/s12229-009-9030-3

Steane, D. A., Potts, B. M., McLean, E., Prober, S. M., Stock, W. D., Vaillancourt, R. E., & Byrne, M. (2014). Genome-wide scans detect adaptation to aridity in a widespread forest tree species. *Molecular Ecology*, *23*(10), 2500–2513. https://doi.org/10.1111/mec.12751

Wright, S. (1943). Isolation by Distance. *Genetics*, *28*(2), 114–138.

CONTRIBUTION

Mon stage s'inscrit dans le projet *Apollon (2019-2024)* porté par Laurence Després au sein du Laboratoire d'Ecologie Alpine (LECA). Ce projet est financé par le LECA ainsi que par la DREAL Auvergne-Rhône-Alpes dans le cadre du plan National d'Actions en faveur des papillons de jour. La mise en place du projet et son suivi, la définition de la question de recherche ainsi que du protocole d'échantillonnage ont été réalisés par Laurence Després. Les échantillons ont été obtenus grâce à la collaboration de nombreux organismes : les Conservatoires des Espaces Naturels (Ariège, Haute-Savoie, PACA), le Parc National des Cévennes, de nombreux Parcs Naturels Régionaux (Chartreuse, Massif des Bauges, Haut-Jura, Volcans Auvergne), ainsi que de nombreuses Réserves Naturelles (Ardèche, Hauts de Chartreuse, Les Partias, Le Mantet) et Réserves Naturelles Nationales (Chastreix-Sarcy, Hauts-Plateaux du Vercors, Vallée de Chaudefour , Gorges de Daluis) et également d'autres organismes (laboratoire, association …) comme la Société d'Histoire naturelle Alcide-d'Orbigny, le Conservatoire Botanique National de Franche Comté, Cistude Nature, FLAVIA APE, la fédération des réserves Catalanes, le Muséum Nationale d'Histoire Naturelle et le Centre d'Ecologie Fonctionnelle et Evolutive. L'extraction de l'ADN ainsi que la préparation des librairies ddRAD ont été réalisées par Delphine Rioux.

La recherche bibliographique a été menée par Thomas Francisco. L'obtention des données climatiques (Chelsa et Worldclim) ainsi que des occurrences Gbif et leur filtration ont été réalisées par Thomas Francisco avec les conseils de Laurence Després et Thibaut Capblancq. Les analyses bio-informatiques sous Rstudio ainsi que la réalisation des résultats ont été réalisées par Thomas Francisco avec les précieux conseils de Laurence Després et Thibaut Capblancq. L'interprétation des résultats a été réalisée par Laurence Després, Thibaut Capblancq et Thomas Francisco. L'écriture du manuscrit a été réalisée par Thomas Francisco. Le manuscrit a été relu et commenté par Laurence Després. Ce stage a majoritairement été supervisé par Laurence Després ainsi que par Thibaut Capblancq qui ont su se rendre disponible pour répondre aux interrogations de tout type et apporter des suggestions / réponses face aux problèmes rencontrés.

Remerciements

Je souhaite tout d'abord remercier Laurence Després pour m'avoir donné l'opportunité de réaliser ce stage et d'avoir su se rendre disponible lors de ces 6 mois pour me permettre de mener à bien ce projet. Je souhaite également remercier Thibaut Capblancq qui m'a beaucoup aidé à la compréhension des concepts clés abordés. Je remercie également toutes les personnes du LECA qui m'ont aidé de près comme de loin. Enfin, je remercie ma famille qui a toujours cru en moi ainsi que ma copine pour son soutien et la relecture de ce rapport (et des nombreux précédents depuis déjà plusieurs années).