

Supplementary materials

Extended Material and Methods

2.3. Species Occurrence Data

Helleborus odorus subsp. *cyclophyllus* occurrence data (167 occurrences) were obtained from the Global Biodiversity Information Facility database through functions from the 'rgbif' 1.4.0 [36] package. To avoid pseudoreplication and associated spatial sampling biases, we selected occurrences with a minimum distance from each other (1 km). We removed the fewest records necessary to substantially reduce the effects of sampling bias, while simultaneously retaining the greatest amount of useful information. This subsampling reduces the spatial aggregation of records to prevent that SDMs reflect a possible over-representation of environmental conditions associated with regions of higher sampling, which hinder interpretation and application of models [73,74]. This data cleaning and organizing procedure followed the protocols as set out in Robertson et al. [37] and we used the 'biogeo' 1.0 [37] and 'spThin' 0.1.0 [74] packages. To tackle sample bias uncertainty, we also employed an environmental filtering procedure that can improve model performance [75], based on the representative and uncorrelated environmental variables occurring in the study area (see Environmental data below). Finally, we evaluated whether any geographical sampling bias existed in our species occurrence data by comparing the statistical distance distribution observed in our dataset to a simulated distribution expected under random sampling via the 'sampbias' 0.1.1 [76] package.

2.4. Environmental data

Current and future climatic data were obtained from the WorldClim database [38] and the CHELSA database [39,40] at a 30 sec resolution. We used two climate databases to assess the bioclimatic consistency and congruence of model predictions [41], a crucial source of uncertainty in SDMs [42].

We constructed 16 more climatic variables at the same resolution via the 'envirem' 1.1 [77] package based on the 19 bioclimatic variables from WorldClim and CHELSA for current and future climate conditions. We selected three Global Circulation Models (GCMs) that are rendered more suitable and realistic for the study area's future climate based on McSweeney et al. [78] and four different IPCC scenarios from the Representative Concentration Pathways (RCP) family. Soil variables were obtained from the SoilGrids 250 m database [79,80]. Elevation data were derived from the CGIAR-CSI data-portal [43,81] and then aggregated and resampled using 'raster' 2.6.7 package [44] to match the resolution of the other environmental variables. Additional topographical variables (slope, aspect, heat load index, topographic position index and terrain ruggedness index) were computed based on elevation data using functions from the 'raster' 2.6.7 package [44] and 'spatialEco' 1.2-0 package [45]. Finally, we rasterized the Geological Map of Greece [46] and created a layer containing the occurrence of calcareous substrate, using functions from the 'sf' 0.8.0 package [82] and the 'fasterize' 1.0.0 package [83].

From this initial set of 50 predictors, only fifteen and seventeen variables (depending on the extent of the distributional area – see below) were not highly correlated (Spearman rank correlation < 0.7 and VIF < 5) [47]. Multicollinearity assessment was performed with the 'usdm' 1.1.18 package [84].

2.5. Species distribution models

2.5.1. Model parameterization and evaluation

We modelled the realized climatic niche of *Helleborus odorus* subsp. *cyclophyllus* by combining the available occurrence data with current environmental predictors with the 'biomod2' 3.3.7 package [49] and 'ecospat' 2.2.0 package [85]. We used three different modelling algorithms for our

study species: Random Forest (RF), Classification Tree Analysis (CTA) and Multiple Adaptive Regression Splines (MARS) in an ensemble modelling scheme, as ensemble forecasting integrates the results of multiple SDM algorithms into a single geographical projection for each time period, reducing the uncertainties associated with the use of a single model algorithm [42,48]. Since these algorithms require presence/absence (PA) data, we generated PAs following the recommendations of Barbet-Massin et al. [86]. Since data regarding the exact distribution area of *Helleborus odorus* subsp. *cyclophyllus* in Greece are lacking, we estimated the species' background area using the 'EOO.computing' function from the 'ConR' 1.1.1 package [87] for both the convex and alpha hull methods. Thus, pseudo-absences were generated at a minimum distance of 39.4 and 35.7 km for the alpha and convex hull method, respectively, from presence locations to reduce the probability of false absences. We chose that minimum distance due to the median autocorrelation of 39.3 and 35.6 km for each method respectively, among the non-collinear environmental variables, which we computed with 'blockCV' 1.0.0 [88] package. We followed the ensemble of small models (ESM) framework [89], since the occurrence/predictors ratio was lower than 20 [73]. We calibrated ESMs by fitting numerous bivariate models, which were then averaged into an ensemble model using weights based on model performances. For all models, the weighted sum of presences was equal to that of the PAs. The models' predictive performance was evaluated via the True Skill Statistic (TSS) [90] based on a repeated (10 times) split-sampling approach in which models were calibrated with 80% of the data and evaluated over the remaining 20%. We used null model significance testing [91] to evaluate the performance of all models and estimated the probability that each model performed better than 100 null models. All models were found to outperform the null expectation at $P < 0.001$. Difference of predictive ability (TSS) between the different climate databases, distribution areas and thinning procedures was investigated via a Kruskal-Wallis non-parametric test (KWA).

2.5.2. Model projections

Calibrated models were used to project the suitable area for our species in the study area under current and future conditions through an ensemble forecast approach [48]. The contribution of each model to the ensemble forecast was weighted according to its TSS score. Models with a TSS score < 0.8 were excluded from building projections, to avoid working with poorly calibrated models. Note that while model evaluation was carried out using the above-mentioned data-splitting procedure, the final models used for spatial projections were calibrated using 100% of the data, thus allowing taking advantage of all available data. Binary transformations were carried out using the value maximizing the TSS score as the threshold for distinguishing presence and absence predictions.

As a conservative approach, the suitability of all cells showing variable values not experienced during the model training (values greater than zero in the clamping mask) was set to zero [73]. We subsequently applied a mask representing urban and suburban areas to avoid projections at locations that are unsuitable regardless of the prevailing environmental conditions.

2.5.3. Area range change

To assess whether *Helleborus odorus* subsp. *cyclophyllus* will experience range contraction or expansion under future conditions, we used the 'biomod' 3.3.7 package [49].

2.6. Bioclimatic congruence and consistency

We followed the framework of Morales-Barbero & Álvarez [41] in order to construct the bioclimatic congruence and consistency maps for *Helleborus odorus* subsp. *cyclophyllus* for every time-period that was available in both climate databases.

2.7. Generalized Dissimilarity Modelling

We used Generalized Dissimilarity Modelling (GDM) [50] in the framework laid out by Fitzpatrick and Keller [51] to investigate the spatial and environmental drivers of genetic beta

diversity, as well as to explore the potential variation in future genetic diversity patterns due to climate change.

2.7.1. Statistical modelling

We used the same environmental variables as in the SDM analyses. The significance of all variables was assessed through a Monte Carlo permutation test (999 repetitions) [52] and thus, we identified the most significant predictor variables. For each of these variables, we extracted the fitted I-spline (a curvilinear line expressing the relationship between species turnover and each predictor – each I-spline has three coefficients). We quantified the magnitude of turnover along each gradient by using the sum of the I-spline's coefficients [it defines the proportion of compositional turnover explained by that variable and is determined by the maximum height of its I-spline [50,92]. We assessed model fit via percent deviance explained by the model [52]. The relative importance of each gradient in driving species turnover was explained as the percent change in deviance explained by the full model and the deviance explained by a model fit with that variable permuted (999 permutations [52].

2.7.2. Visualisation of genetic variation patterns

We followed Fitzpatrick and Keller [51] to visualise the spatial patterns of genetic variation. We reduced into three factors the important and uncorrelated environmental variables via a Principal Component Analysis (PCA) and assigned them to RGB colour palette. Thus, similar colours indicate similarity of the expected genetic composition patterns [51].

2.7.3. Population level variability to climate change

Finally, we projected our GDM based on current environmental conditions to every GCM/RCP combination we included in our study, so as to predict the areas where the relationship between genetic composition and future environmental conditions will experience the greatest change (‘genetic offset’)[51]. We then mapped the mean genetic offset from the twelve GCMs/RCPs to indicate the spatial distribution of population-level vulnerability to climate change [51]. By doing so, we can infer the intensity of the change of the genetic composition across the landscape needed to preserve the gene-environment relationships observed under current environmental conditions [51]. Based on the GDM results and following the framework of Fitzpatrick et al. [93], we estimated the current and future spatial clusters of genetic composition via an unsupervised classification procedure, using two clustering algorithms: k-means and CLARA. We assessed the optimal number of clusters via the Silhouette index [94] for each time-period and GCM/RCP combination. Finally, we quantitatively assessed the similarity of the different bioregionalizations via the V-measure index of spatial association [95,96]. All analyses were performed using functions from the ‘raster’ 3.0.7 [44], ‘cluster’ 2.0.7-1 [97], ‘clusterCrit’ 1.2.8 [98] and ‘sabre’ 0.3.1 [96] R packages.

All GDM analyses were performed with the ‘gdm’ 1.3.7 [52] R package in the R 3.5.3 [53].

Supplementary Figures

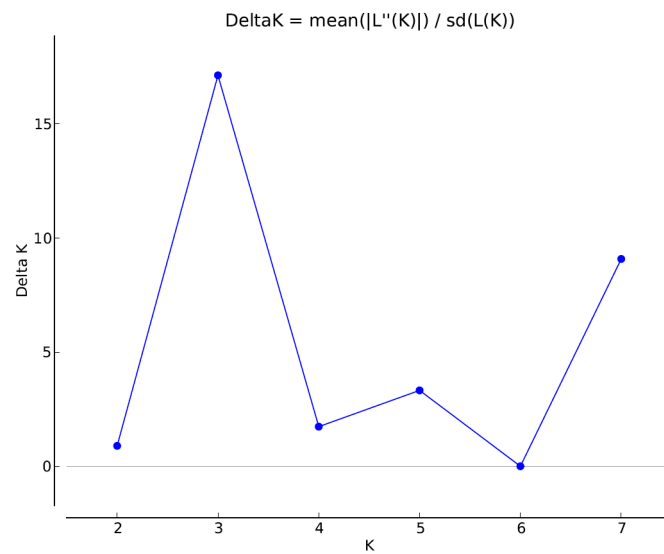


Figure S1. Delta K values for 1 to 8 genetic clusters for *Helleborus odorus* subsp. *cyclophyllus* populations studied. Delta K was calculated according to Evanno et al. (2005).

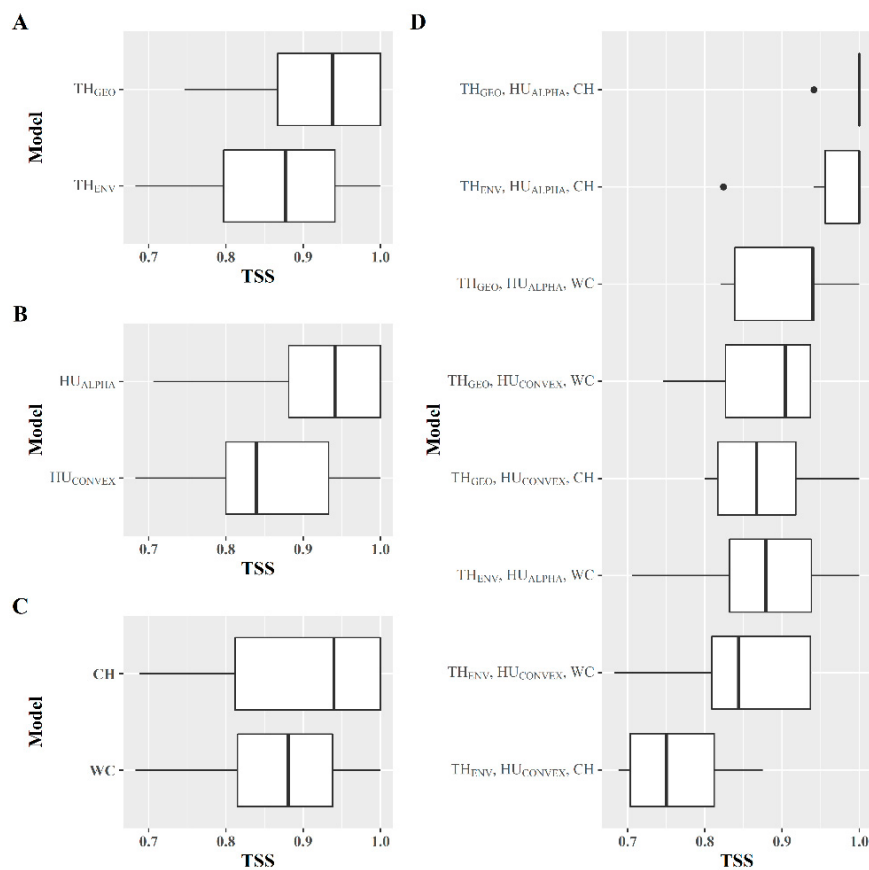


Figure S2. True Skill Statistic (TSS) values for every source of uncertainty included in the present study. TH_{GEO}: geographical thinning. TH_{ENV}: environmental thinning. HU_{ALPHA}: distribution area estimated with the alpha-hull method. HU_{CONVEX}: distribution area estimated with the convex-hull method. CH: CHELSA climate database. WC: WorldClim climate database. Panels A, B, C and D correspond to the thinning procedure, distribution area method, climate database and all the aforementioned uncertainty sources combined, respectively.

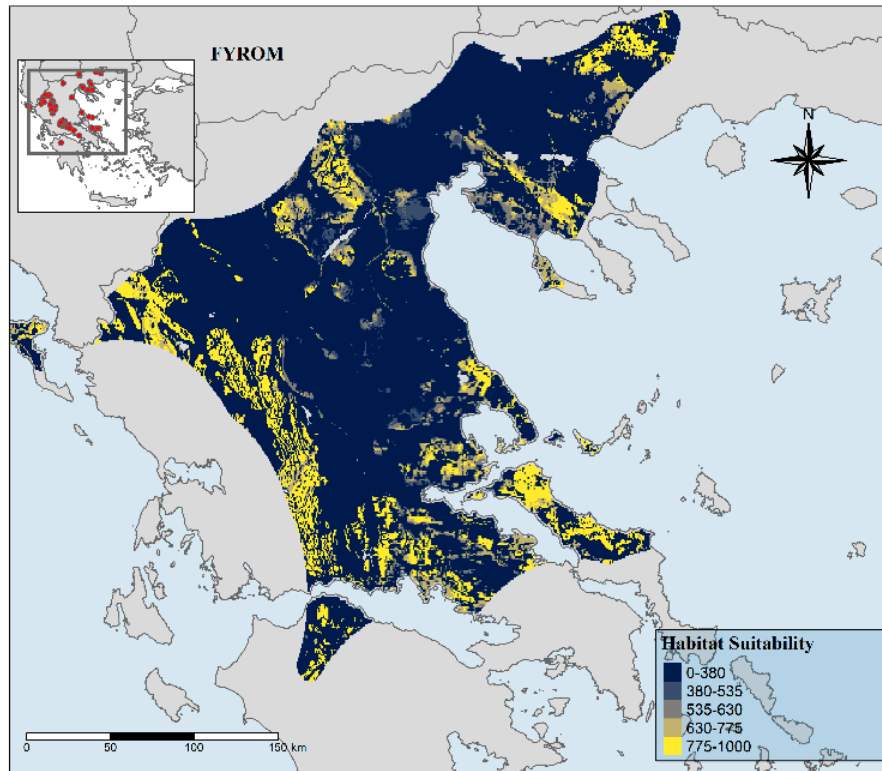


Figure S3. Habitat suitability map based on the CHELSA climate database, geographically thinned coordinates and distribution area estimated with the alpha-hull method for the present-time period.

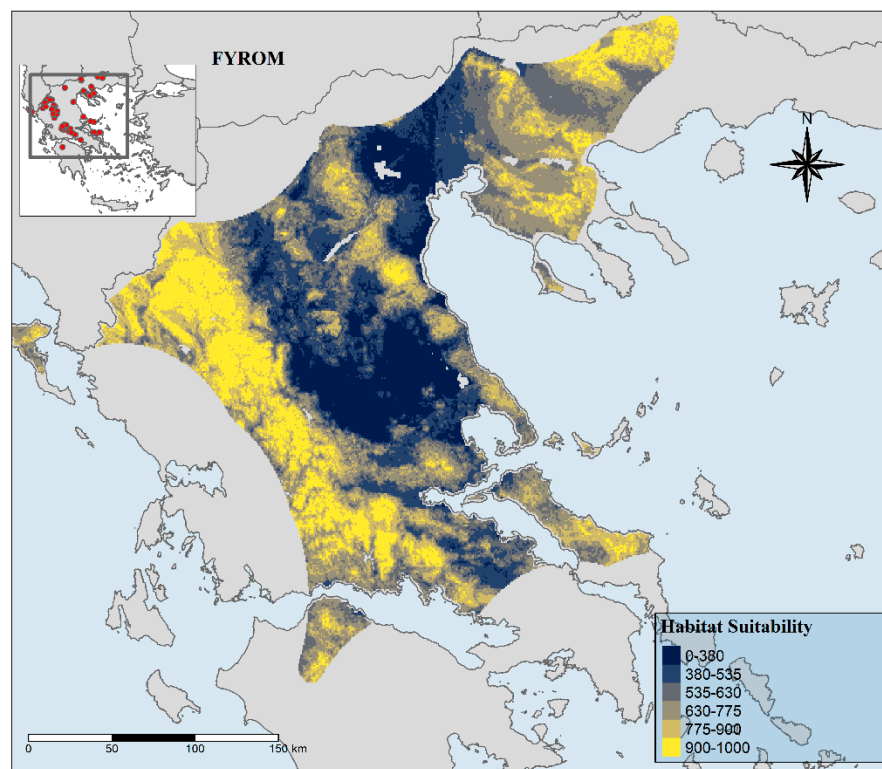


Figure S4. Habitat suitability map based on the WorldClim climate database, geographically thinned coordinates and distribution area estimated with the alpha-hull method for the present-time period.

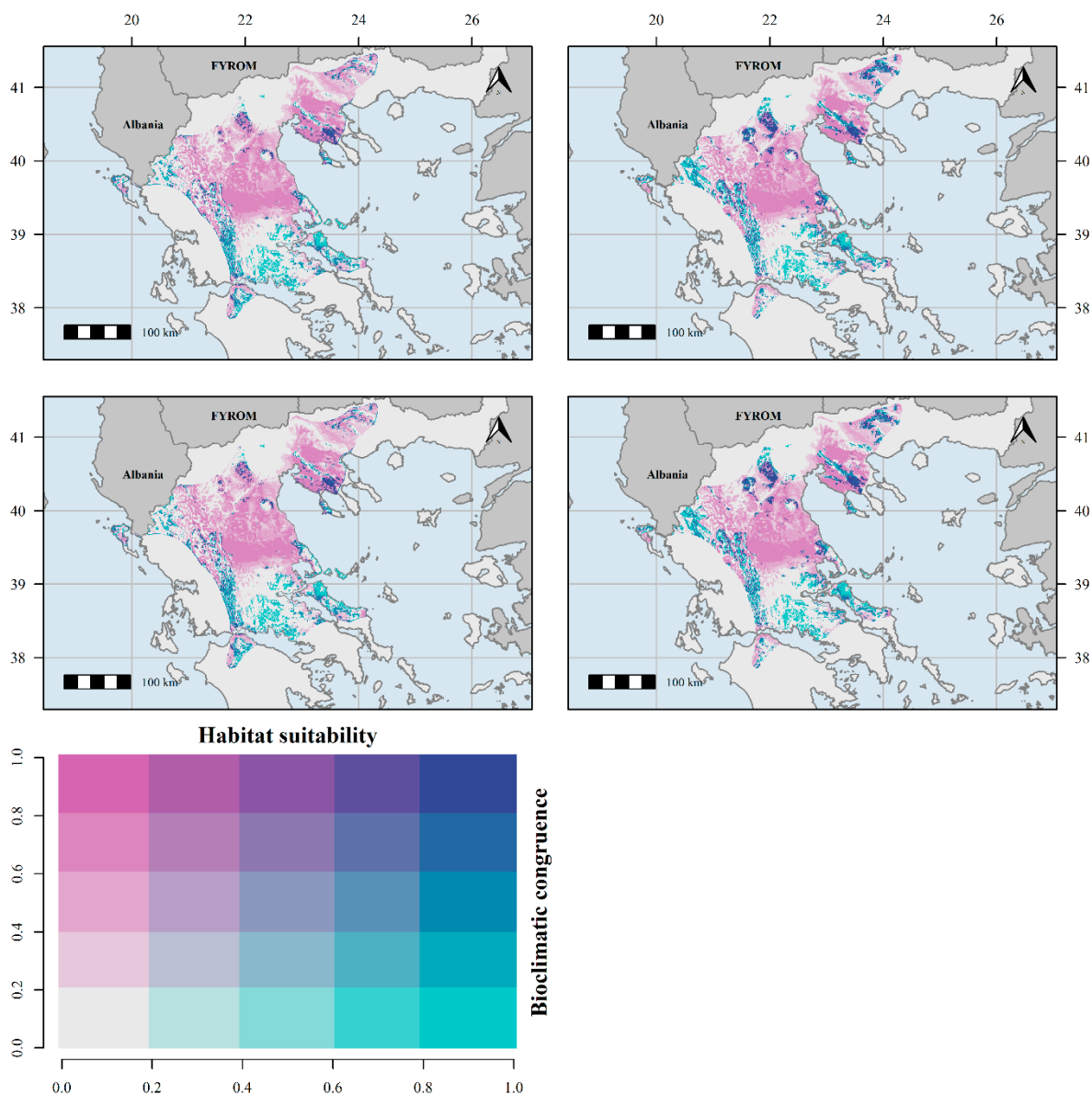


Figure S5. Bioclimatic consistency maps for both thinning procedures (geographical and environmental thinning) based on the alpha-hull method regarding the distribution area of *Helleborus odorus* subsp. *cyclophyllus*. From top-left to bottom-right: Geographically-thinned, based on CHELSA; Geographically-thinned, based on WorldClim; Environmentally-thinned, based on CHELSA; Environmentally-thinned, based on WorldClim.

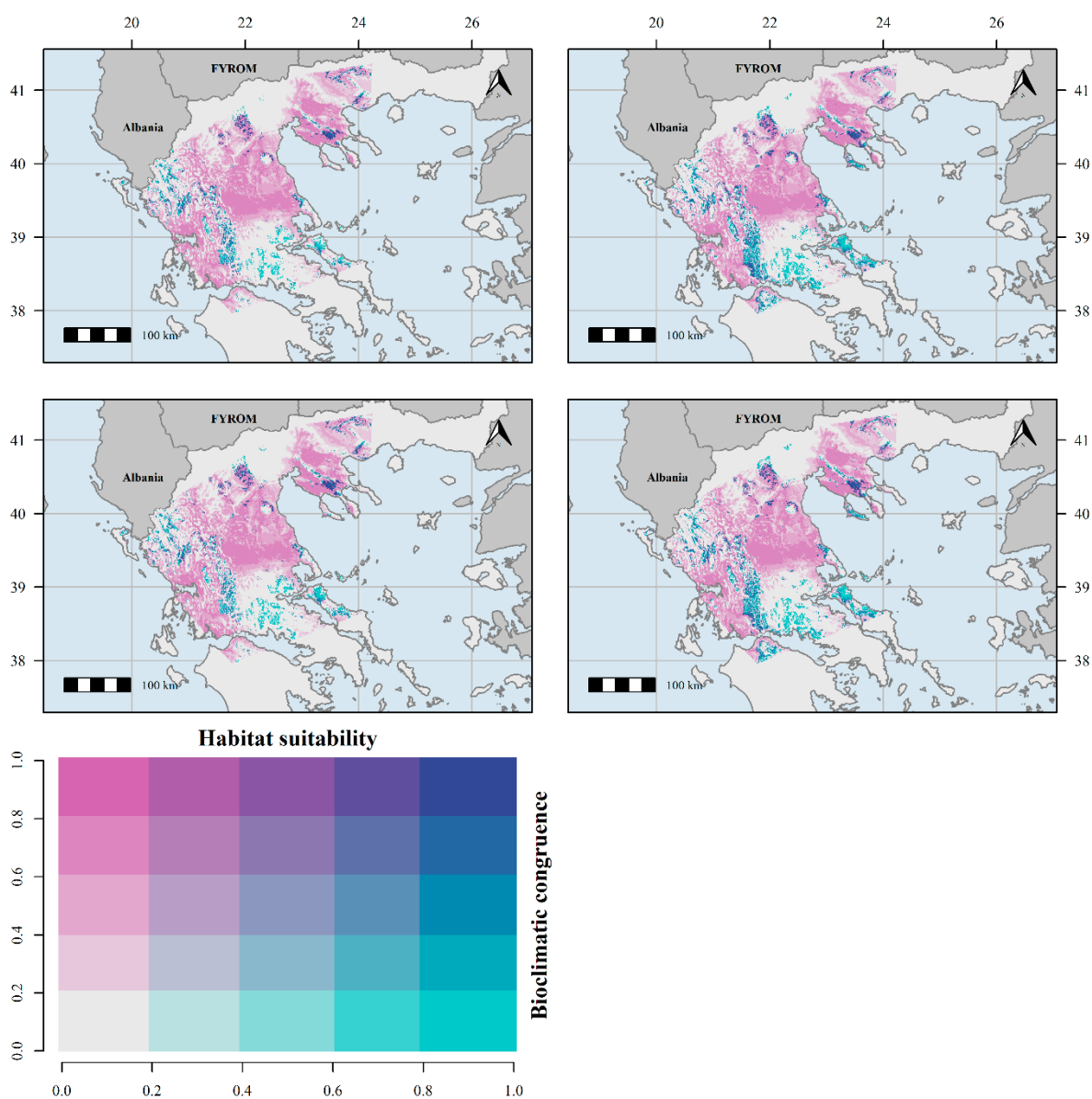


Figure S6: Bioclimatic consistency maps for both thinning procedures (geographical and environmental thinning) based on the convex-hull method regarding the distribution area of *Helleborus odorus* subsp. *cyclophyllus*. From top-left to bottom-right: Geographically-thinned, based on CHELSA; Geographically-thinned, based on WorldClim; Environmentally-thinned, based on CHELSA; Environmentally-thinned, based on WorldClim.

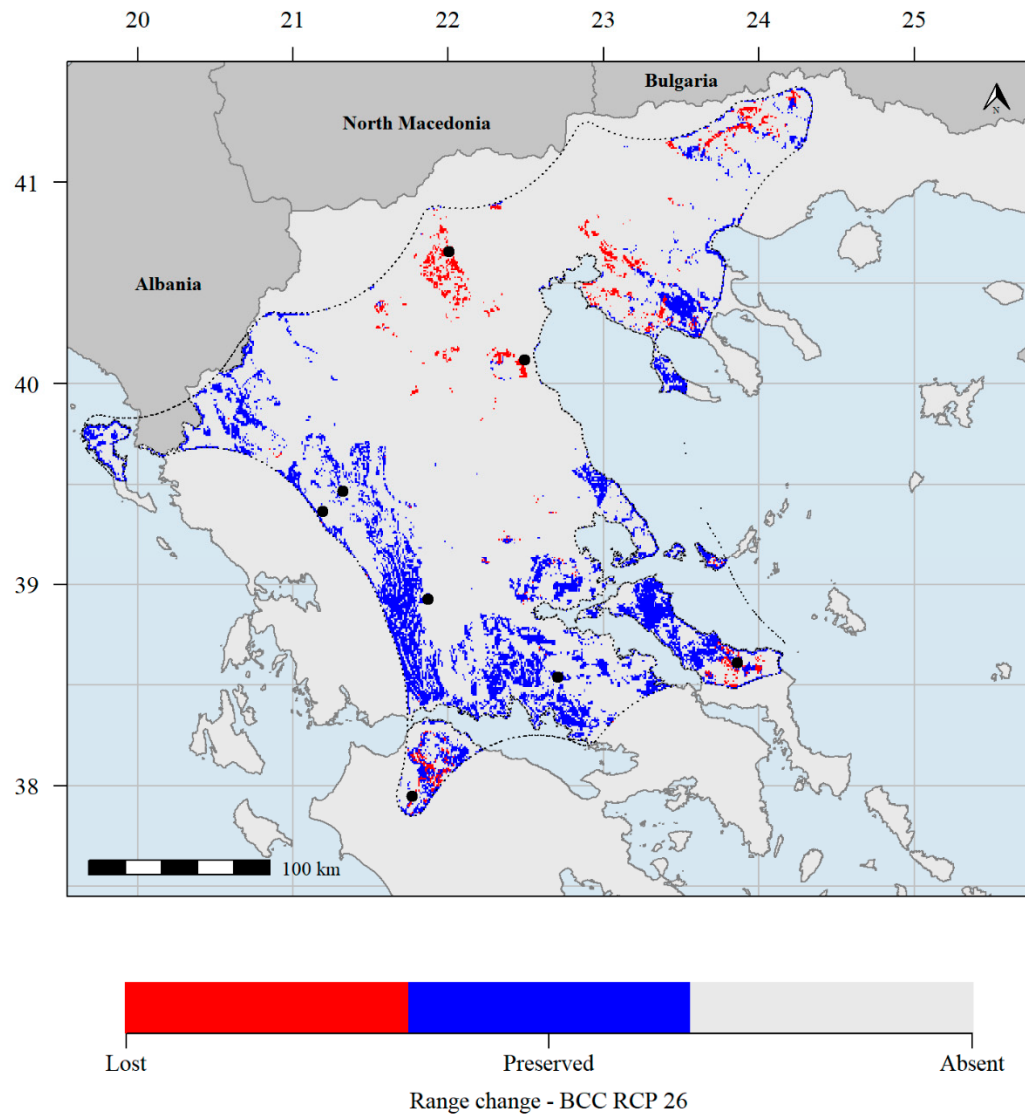


Figure S7: Predicted potential distribution map for 2070 and the BCC GCM and the RCP 2.6 scenario. Red grid cells: the species currently occupies those areas but will not in the future. Blue grid cells: the species currently occupies those areas and will continue to occupy them in the future. Grey grid cells: the species does not currently occupy those areas and it will not occupy them in the future. The dotted line indicates the distribution area of *Helleborus odorus* subsp. *cyclophyllus* based on the alpha-hull method. Climate data refer to the WorldClim database.

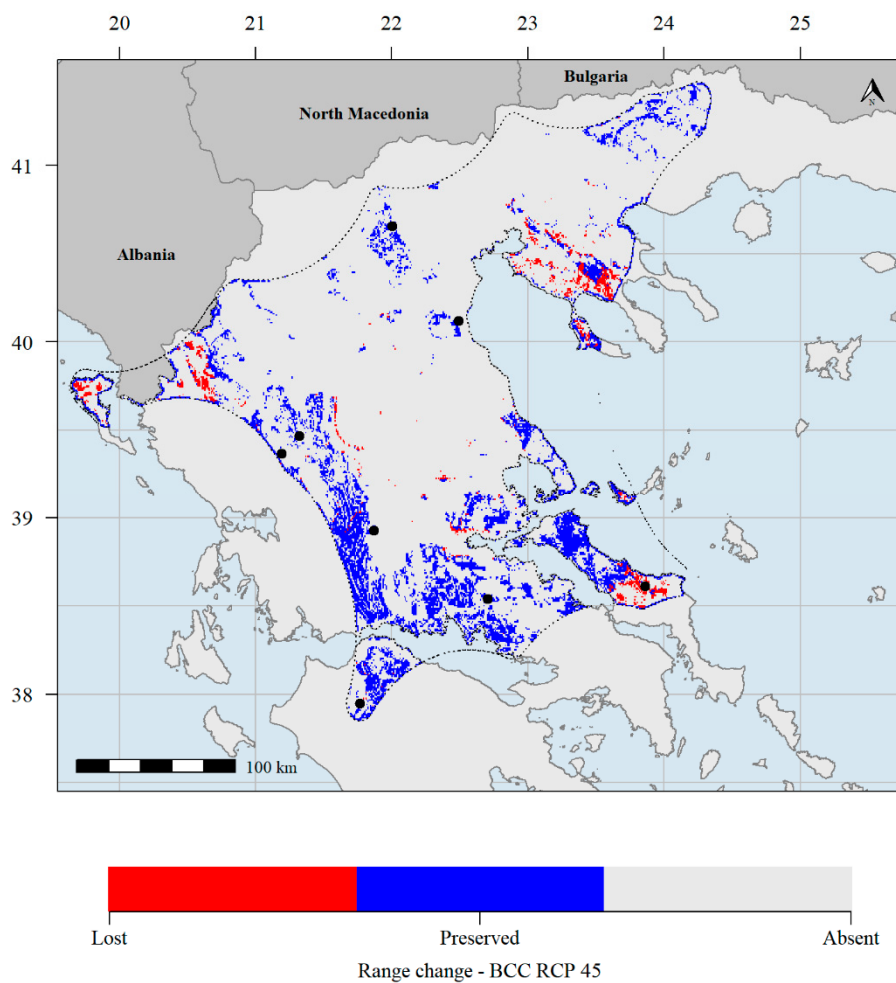


Figure S8: Predicted potential distribution map for 2070 and the BCC GCM and the RCP 4.5 scenario. Red grid cells: the species currently occupies those areas but will not in the future. Blue grid cells: the species currently occupies those areas and will continue to occupy them in the future. Grey grid cells: the species does not currently occupy those areas and it will not occupy them in the future. The dotted line indicates the distribution area of *Helleborus odoratus* subsp. *cyclophyllus* based on the alpha-hull method. Climate data refer to the WorldClim database.

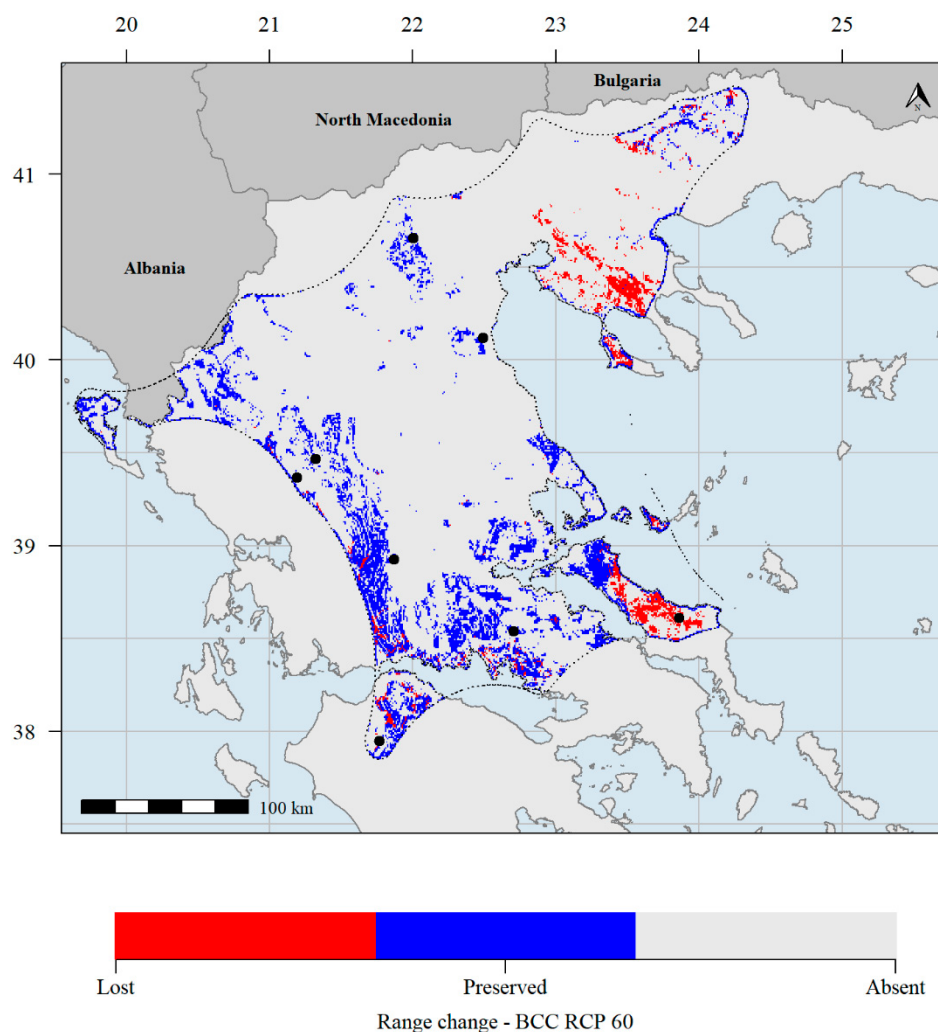


Figure S9: Predicted potential distribution map for 2070 and the BCC GCM and the RCP 6.0 scenario. Red grid cells: the species currently occupies those areas but will not in the future. Blue grid cells: the species currently occupies those areas and will continue to occupy them in the future. Grey grid cells: the species does not currently occupy those areas and it will not occupy them in the future. The dotted line indicates the distribution area of *Helleborus odorus* subsp. *cyclophyllus* based on the alpha-hull method. Climate data refer to the WorldClim database.

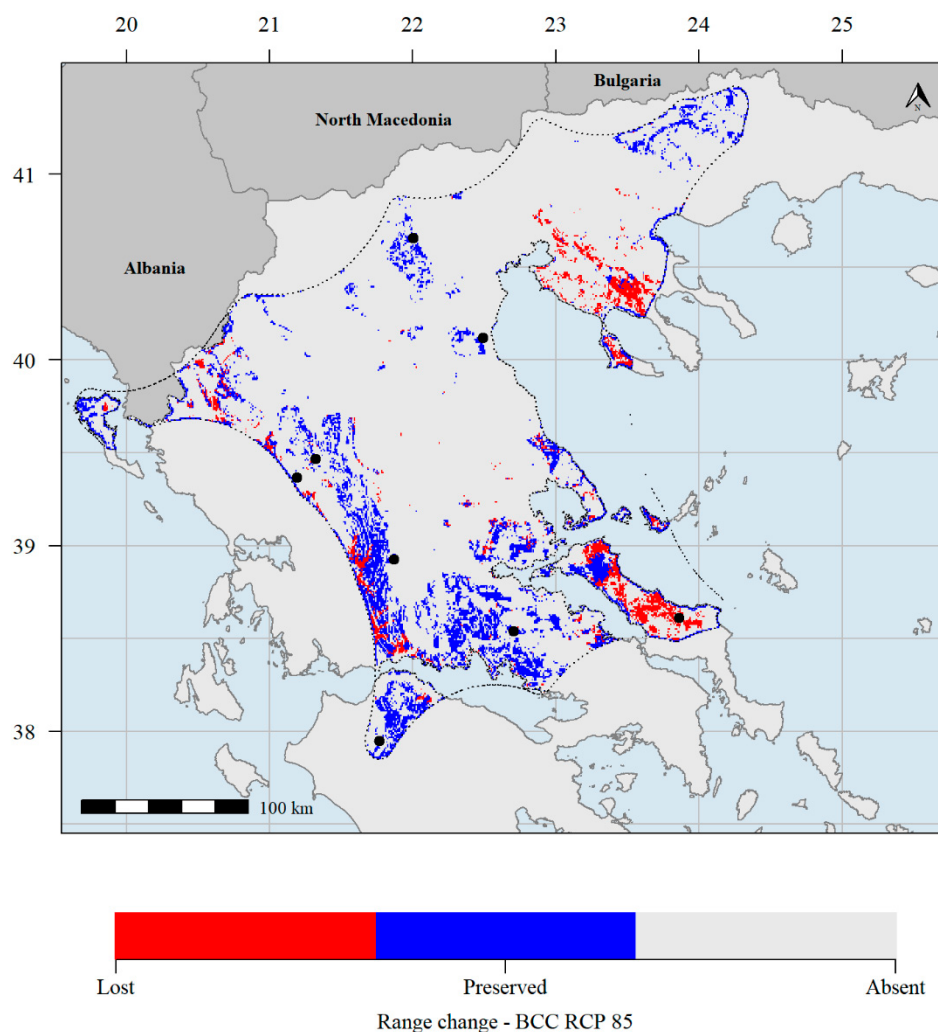


Figure S10: Predicted potential distribution map for 2070 and the BCC GCM and the RCP 8.5 scenario. Red grid cells: the species currently occupies those areas but will not in the future. Blue grid cells: the species currently occupies those areas and will continue to occupy them in the future. Grey grid cells: the species does not currently occupy those areas and it will not occupy them in the future. The dotted line indicates the distribution area of *Helleborus odorus* subsp. *cyclophyllus* based on the alpha-hull method. Climate data refer to the WorldClim database.

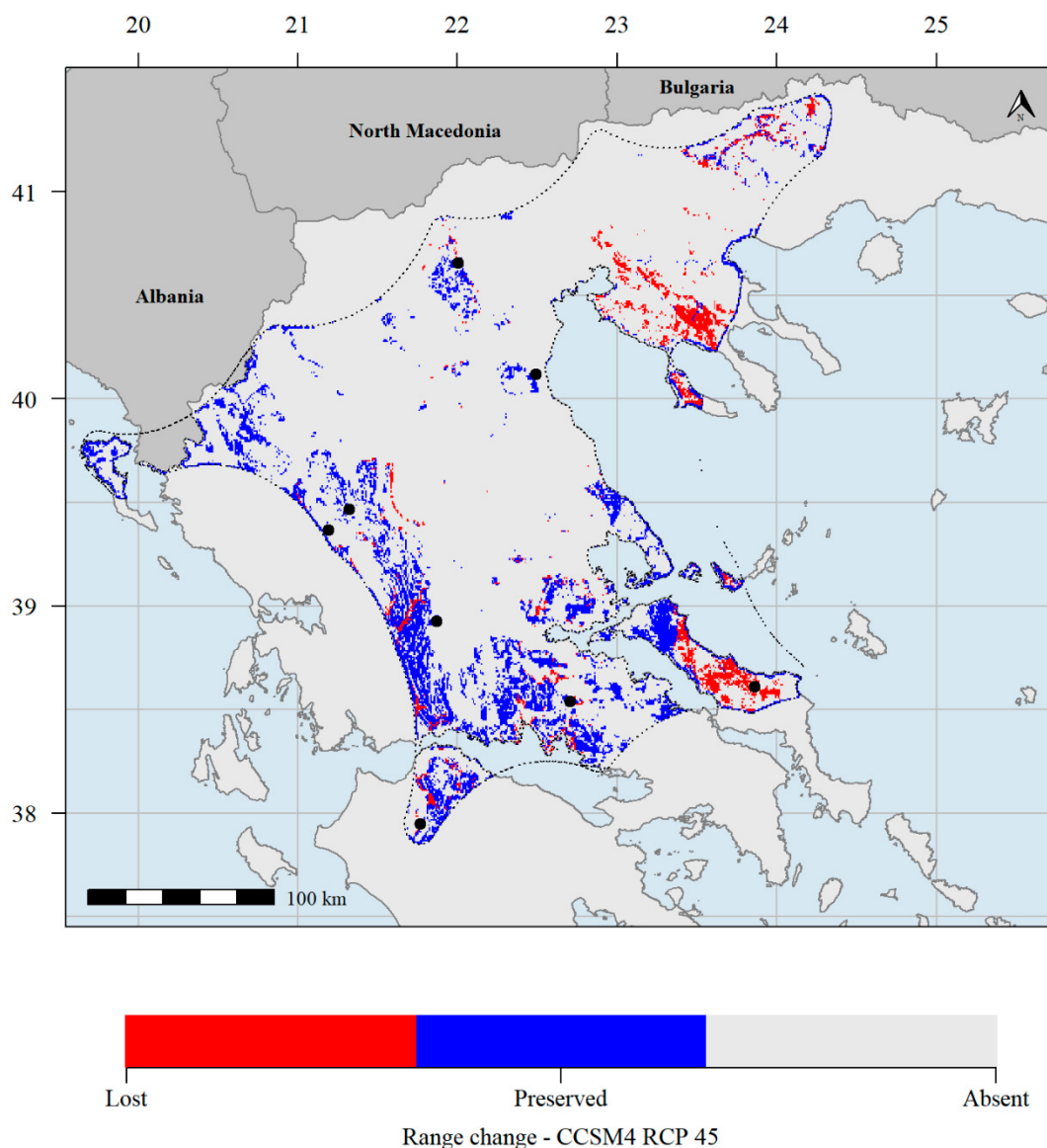


Figure S11: Predicted potential distribution map for 2070 and the CCSM4 GCM and the RCP 4.5 scenario. Red grid cells: the species currently occupies those areas but will not in the future. Blue grid cells: the species currently occupies those areas and will continue to occupy them in the future. Grey grid cells: the species does not currently occupy those areas and it will not occupy them in the future. The dotted line indicates the distribution area of *Helleborus odorus* subsp. *cyclophyllus* based on the alpha-hull method. Climate data refer to the WorldClim database.

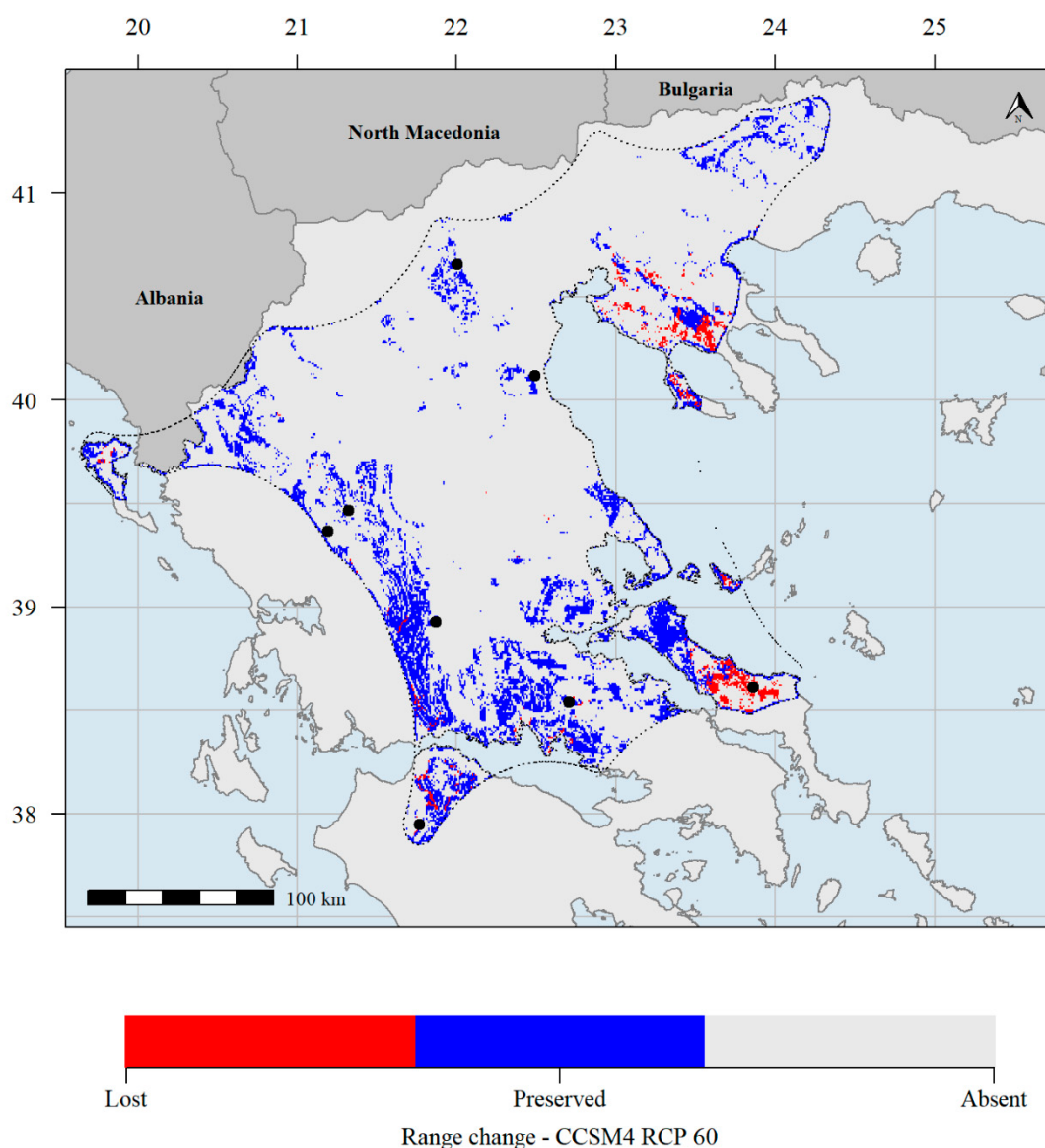


Figure S12: Predicted potential distribution map for 2070 and the CCSM4 GCM and the RCP 6.0 scenario. Red grid cells: the species currently occupies those areas but will not in the future. Blue grid cells: the species currently occupies those areas and will continue to occupy them in the future. Grey grid cells: the species does not currently occupy those areas and it will not occupy them in the future. The dotted line indicates the distribution area of *Helleborus odorus* subsp. *cyclophyllus* based on the alpha-hull method. Climate data refer to the WorldClim database.

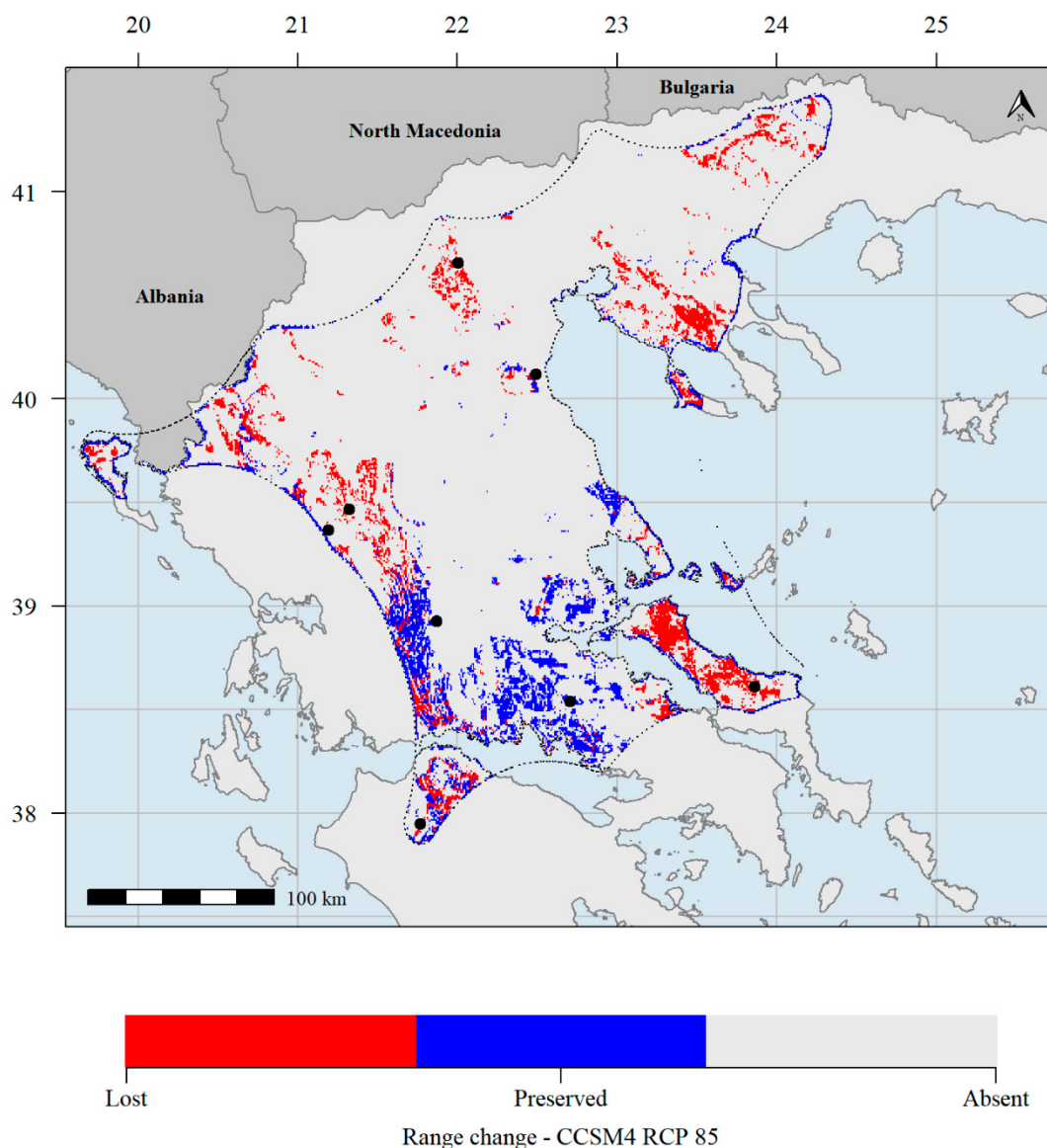


Figure S13: Predicted potential distribution map for 2070 and the CCSM4 GCM and the RCP 8.5 scenario. Red grid cells: the species currently occupies those areas but will not in the future. Blue grid cells: the species currently occupies those areas and will continue to occupy them in the future. Grey grid cells: the species does not currently occupy those areas and it will not occupy them in the future. The dotted line indicates the distribution area of *Helleborus odorus* subsp. *cyclophyllus* based on the alpha-hull method. Climate data refer to the WorldClim database.

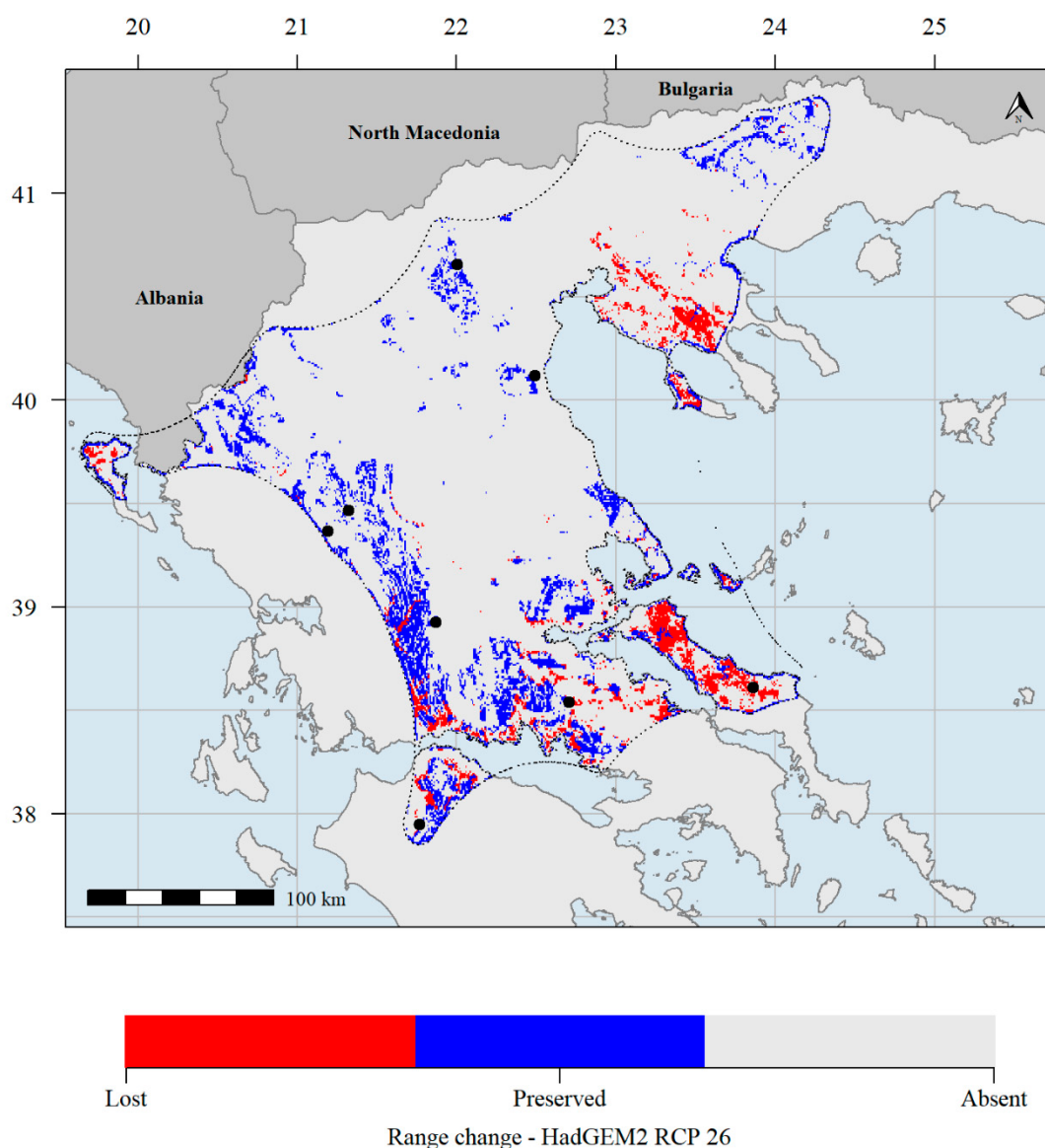


Figure S14: Predicted potential distribution map for 2070 and the HadGEM2 GCM and the RCP 2.6 scenario. Red grid cells: the species currently occupies those areas but will not in the future. Blue grid cells: the species currently occupies those areas and will continue to occupy them in the future. Grey grid cells: the species does not currently occupy those areas and it will not occupy them in the future. The dotted line indicates the distribution area of *Helleborus odorus* subsp. *cyclophyllus* based on the alpha-hull method. Climate data refer to the WorldClim database.

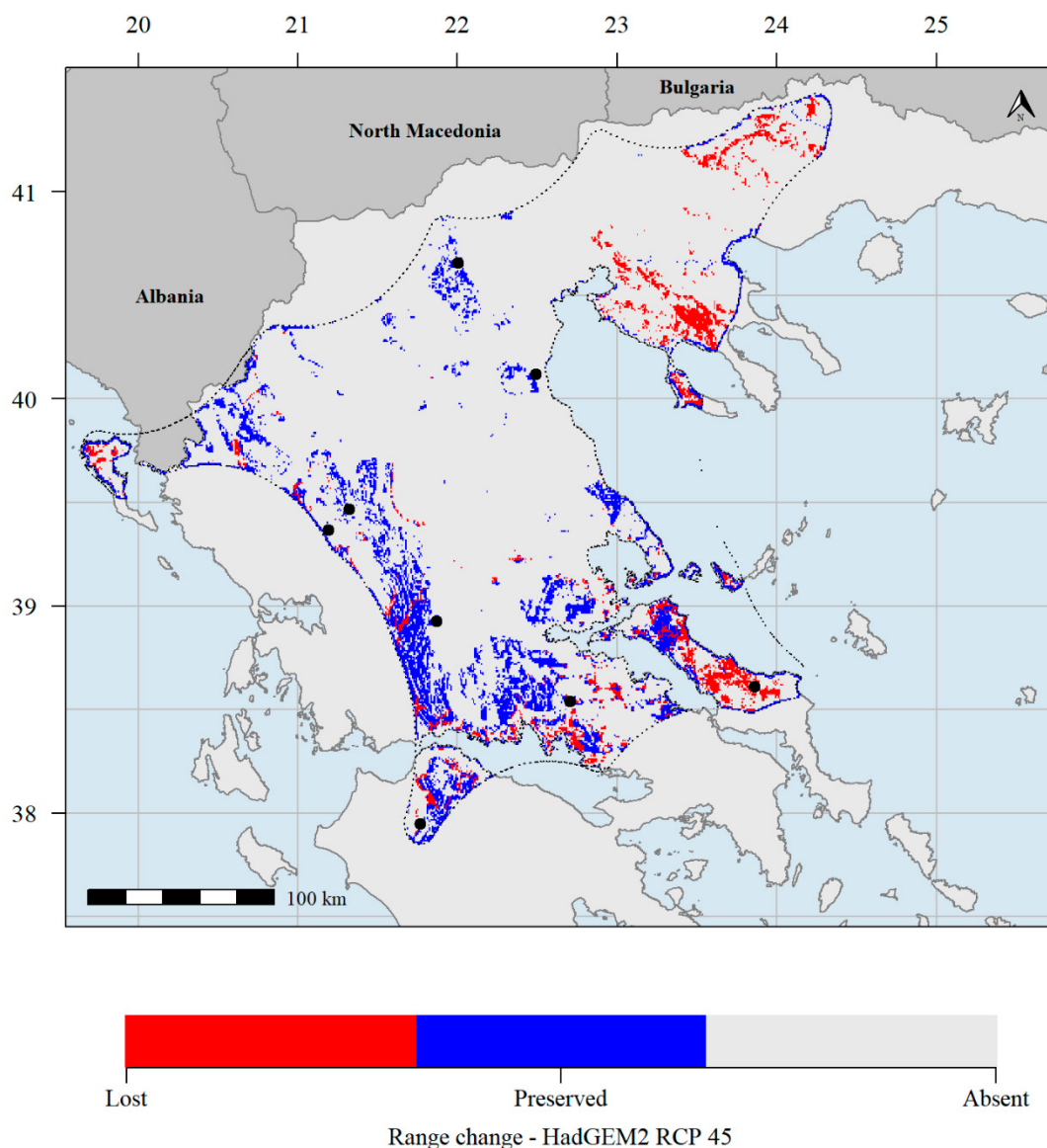


Figure S15. Predicted potential distribution map for 2070 and the HadGEM2 GCM and the RCP 4.5 scenario. Red grid cells: the species currently occupies those areas but will not in the future. Blue grid cells: the species currently occupies those areas and will continue to occupy them in the future. Grey grid cells: the species does not currently occupy those areas and it will not occupy them in the future. The dotted line indicates the distribution area of *Helleborus odorus* subsp. *cyclophyllus* based on the alpha-hull method. Climate data refer to the WorldClim database.

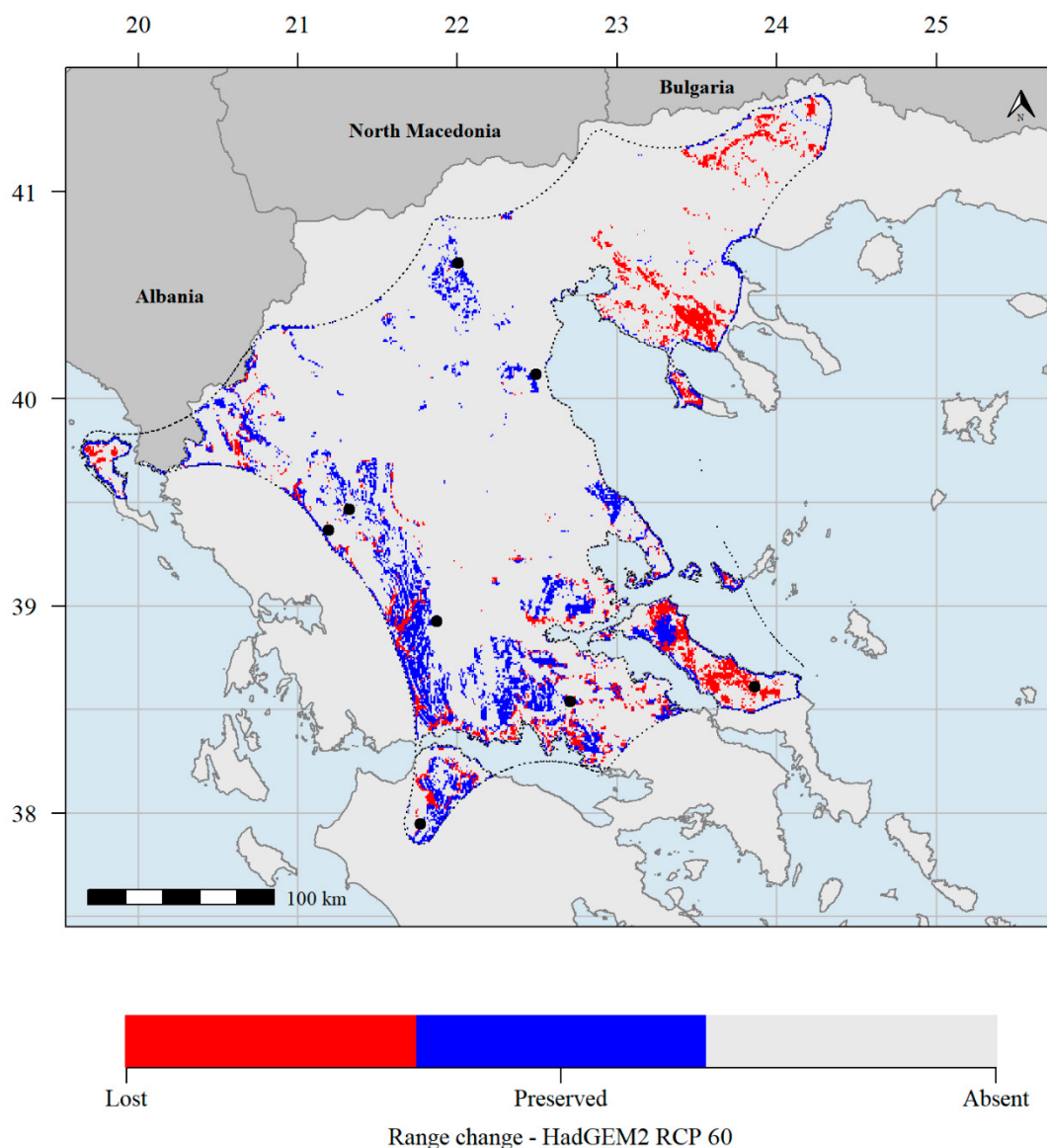


Figure S16: Predicted potential distribution map for 2070 and the HadGEM2 GCM and the RCP 6.0 scenario. Red grid cells: the species currently occupies those areas but will not in the future. Blue grid cells: the species currently occupies those areas and will continue to occupy them in the future. Grey grid cells: the species does not currently occupy those areas and it will not occupy them in the future. The dotted line indicates the distribution area of *Helleborus odorus* subsp. *cyclophyllus* based on the alpha-hull method. Climate data refer to the WorldClim database.

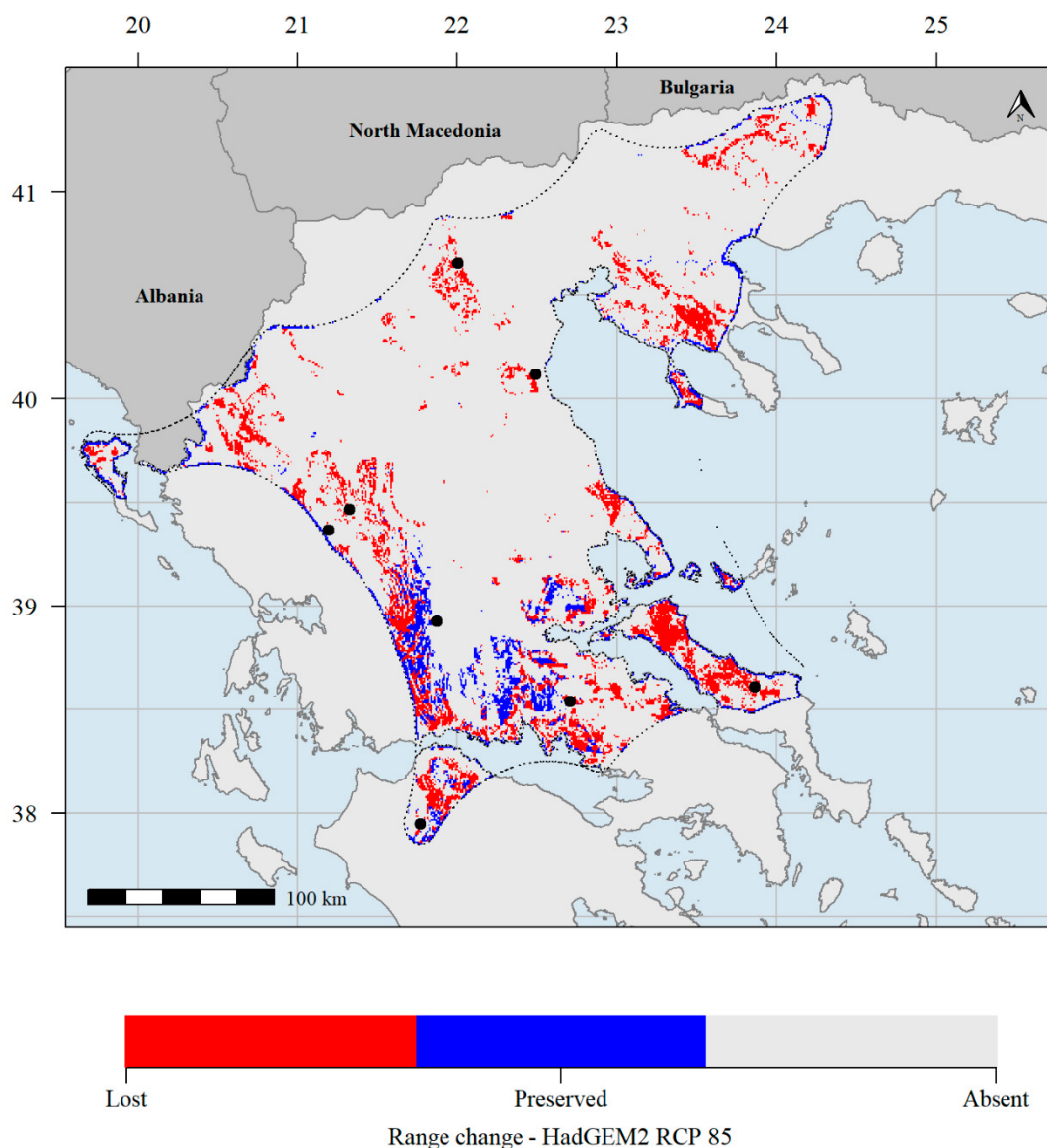


Figure S17: Predicted potential distribution map for 2070 and the HadGEM2 GCM and the RCP 8.5 scenario. Red grid cells: the species currently occupies those areas but will not in the future. Blue grid cells: the species currently occupies those areas and will continue to occupy them in the future. Grey grid cells: the species does not currently occupy those areas and it will not occupy them in the future. The dotted line indicates the distribution area of *Helleborus odorus* subsp. *cyclophyllus* based on the alpha-hull method. Climate data refer to the WorldClim database.

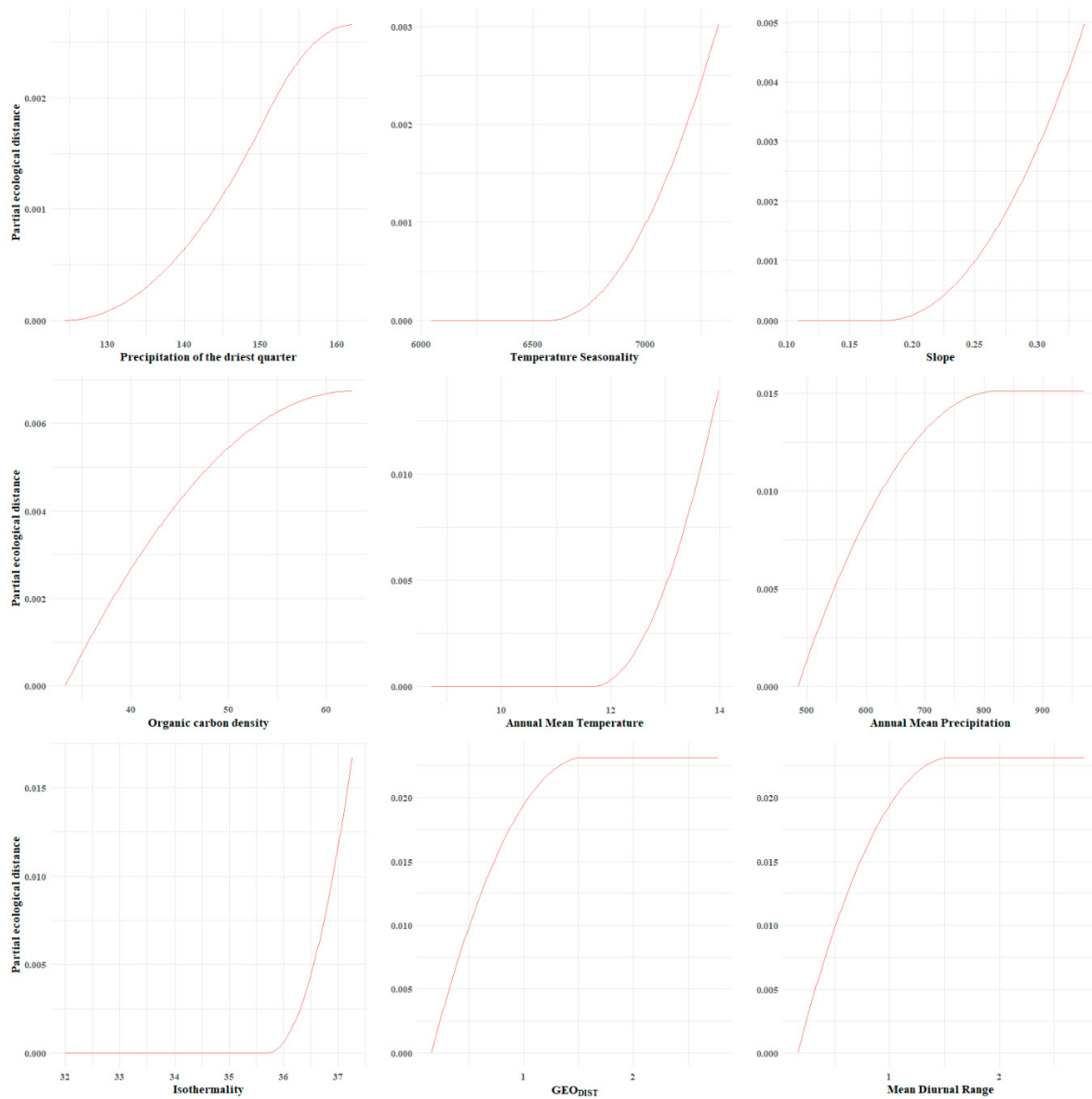


Figure S18: GDM-fitted I-splines (partial regression fits) for variables significantly associated with the genetic composition of the populations under study of *Helleborus odorus* subsp. *cyclophyllus*. The maximum height reached by each curve indicates the total amount of turnover associated with that variable, holding all other variables constant. The shape of each function indicates the variation of turnover rate along the gradient. GEO_{DIST}: The geographical distance (in km) between the populations under study.

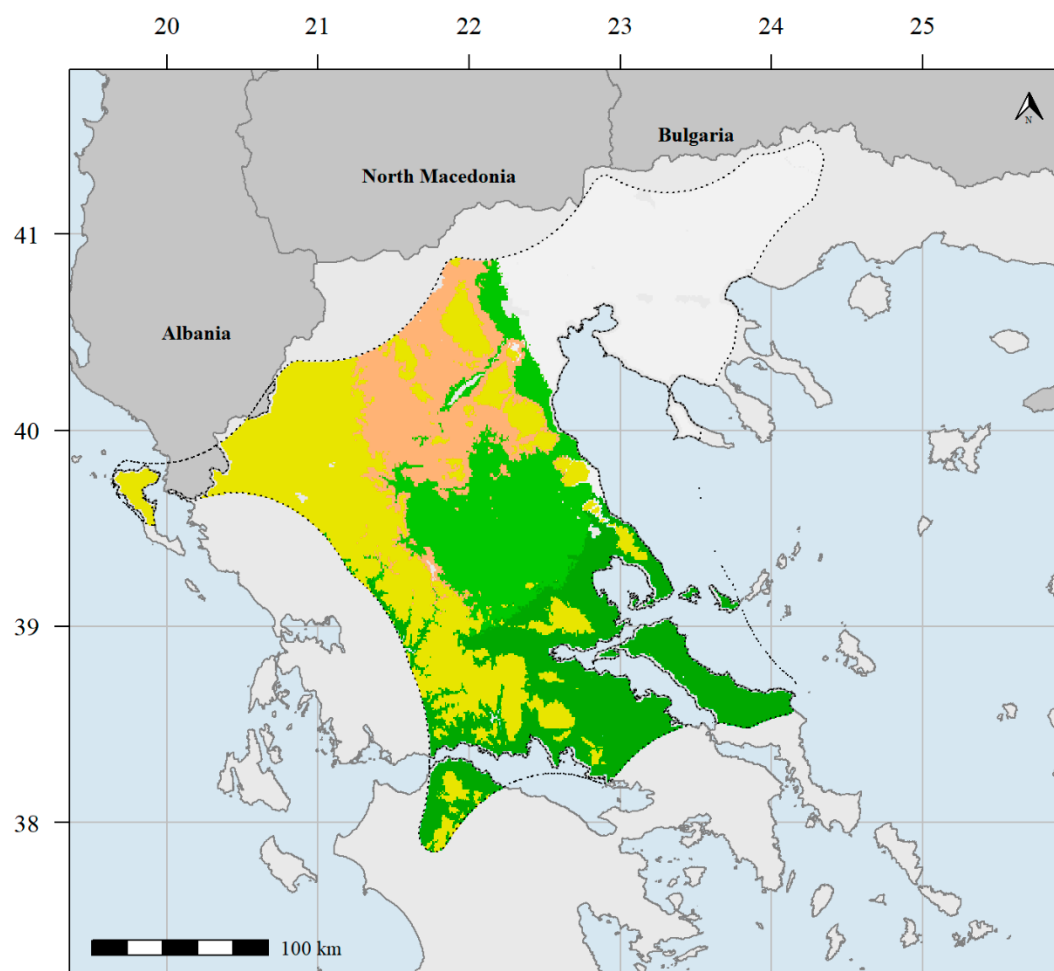


Figure S19: Predicted spatial variation for the current time-period in population-level genetic composition from GDM. Different colours represent gradients in genetic turnover. Locations with similar colours are expected to harbour populations with similar genetic composition.

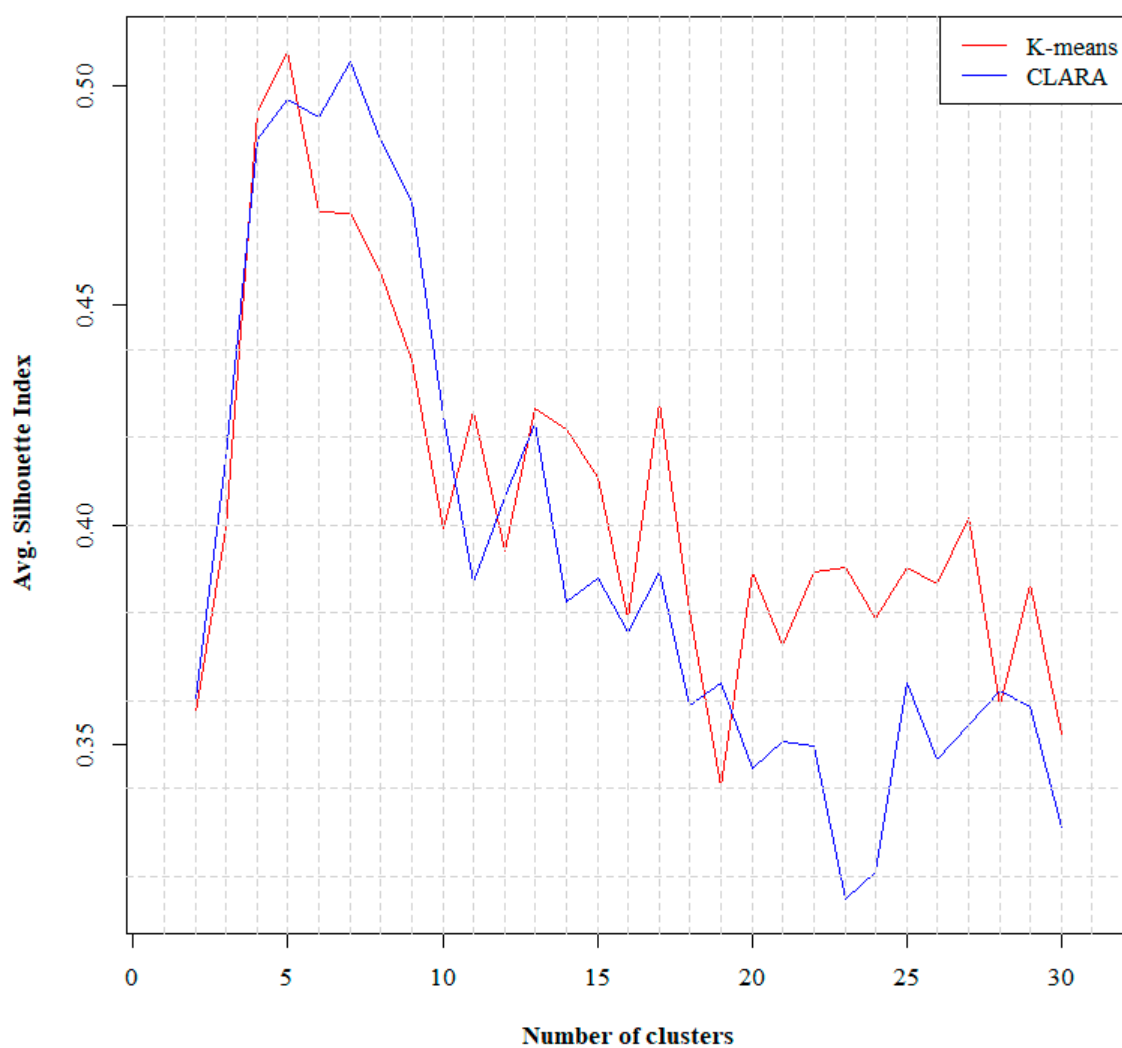


Figure S20: The values of the Silhouette index for the k-means and the CLARA unsupervised clustering algorithms regarding the optimal number of spatial clusters of genetic composition for *Helleborus odorus* subsp. *cyclophyllus* for the current time-period.

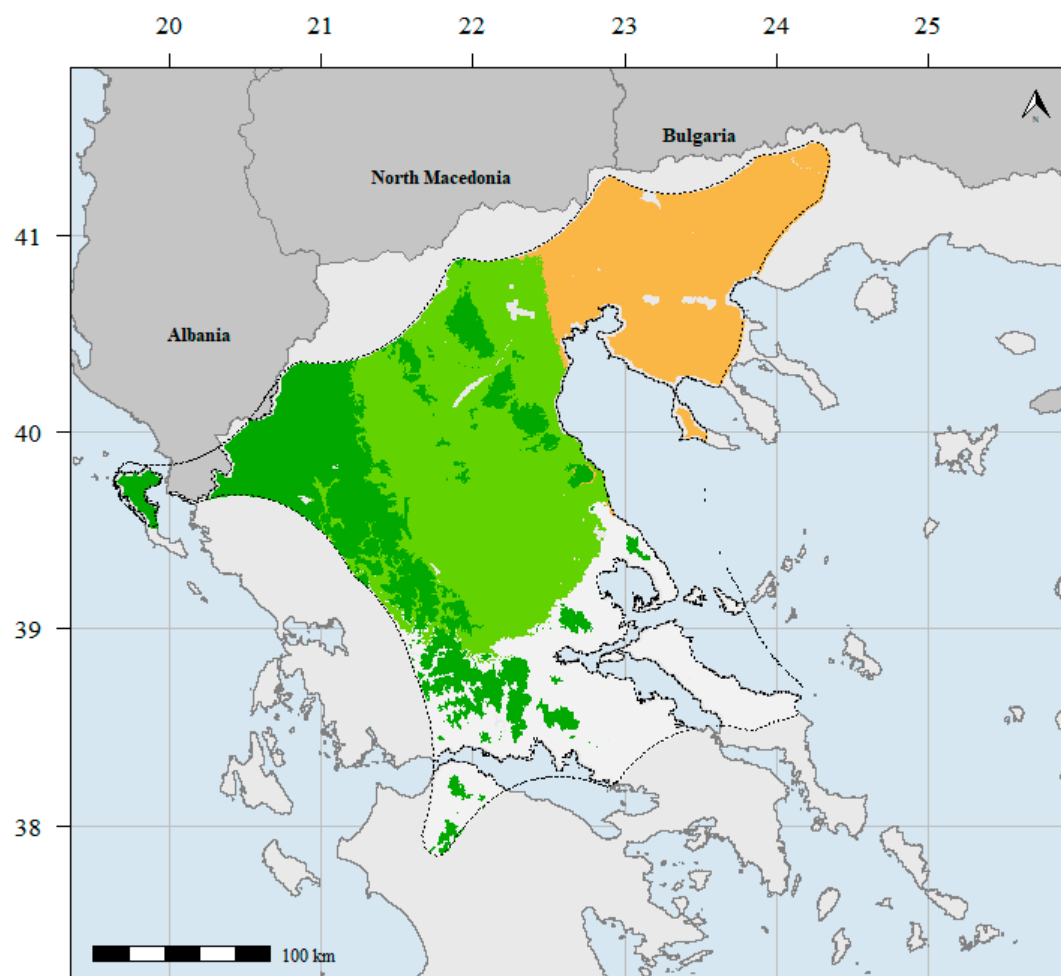


Figure S21: Predicted spatial variation for the BCC RCP 2.6 time-period in population-level genetic composition from GDM. Different colours represent gradients in genetic turnover. Locations with similar colours are expected to harbour populations with similar genetic composition.

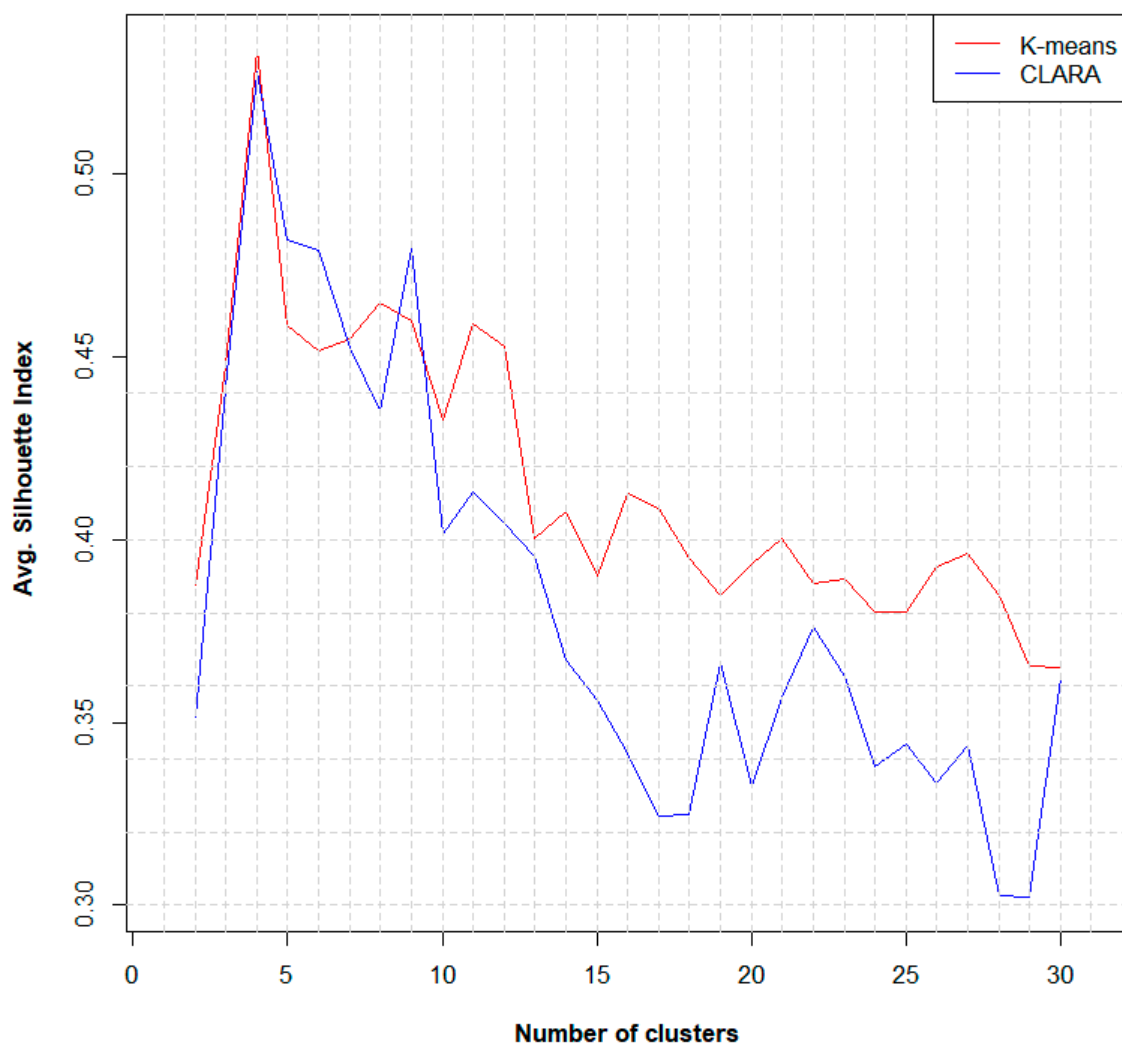


Figure S22: The values of the Silhouette index for the k-means and the CLARA unsupervised clustering algorithms regarding the optimal number of spatial clusters of genetic composition for *Helleborus odorus* subsp. *cyclophyllus* for the BCC RCP 2.6 time-period.

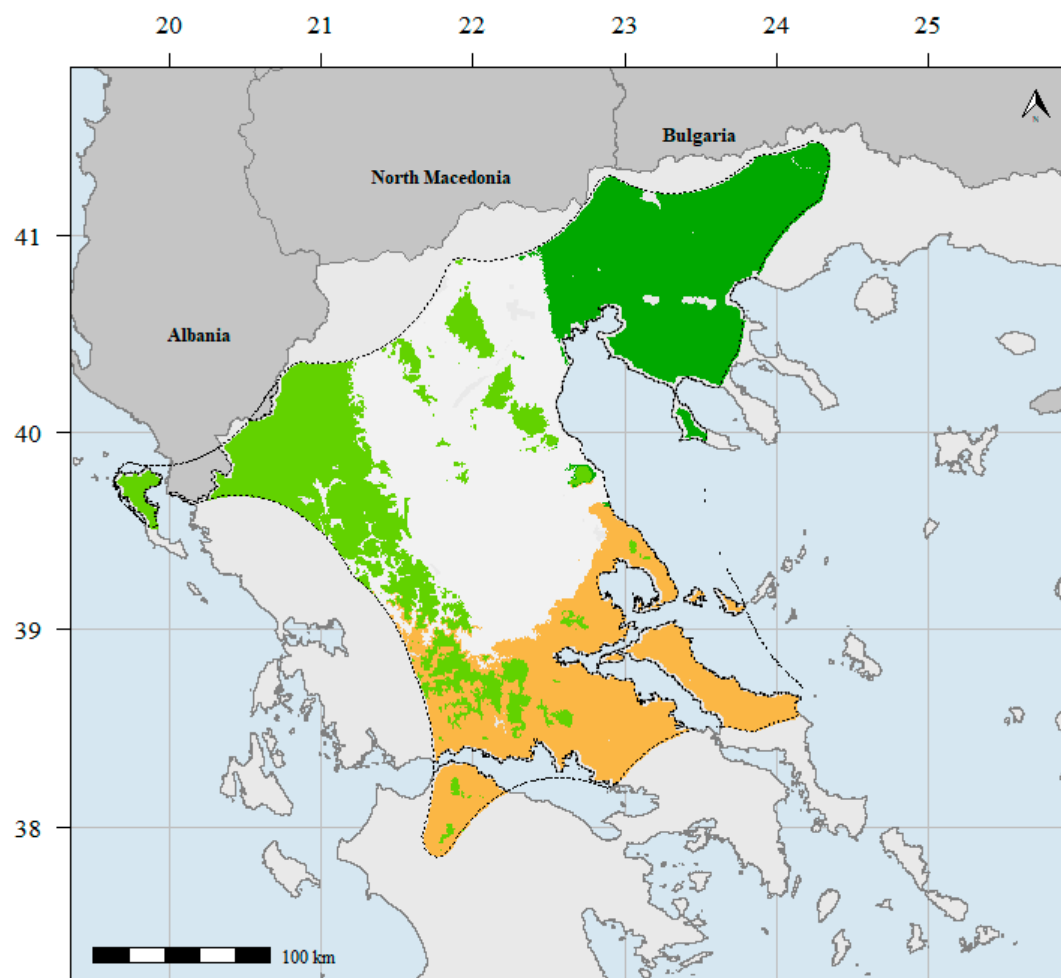


Figure S23: Predicted spatial variation for the BCC RCP 4.5 time-period in population-level genetic composition from GDM. Different colours represent gradients in genetic turnover. Locations with similar colours are expected to harbour populations with similar genetic composition.

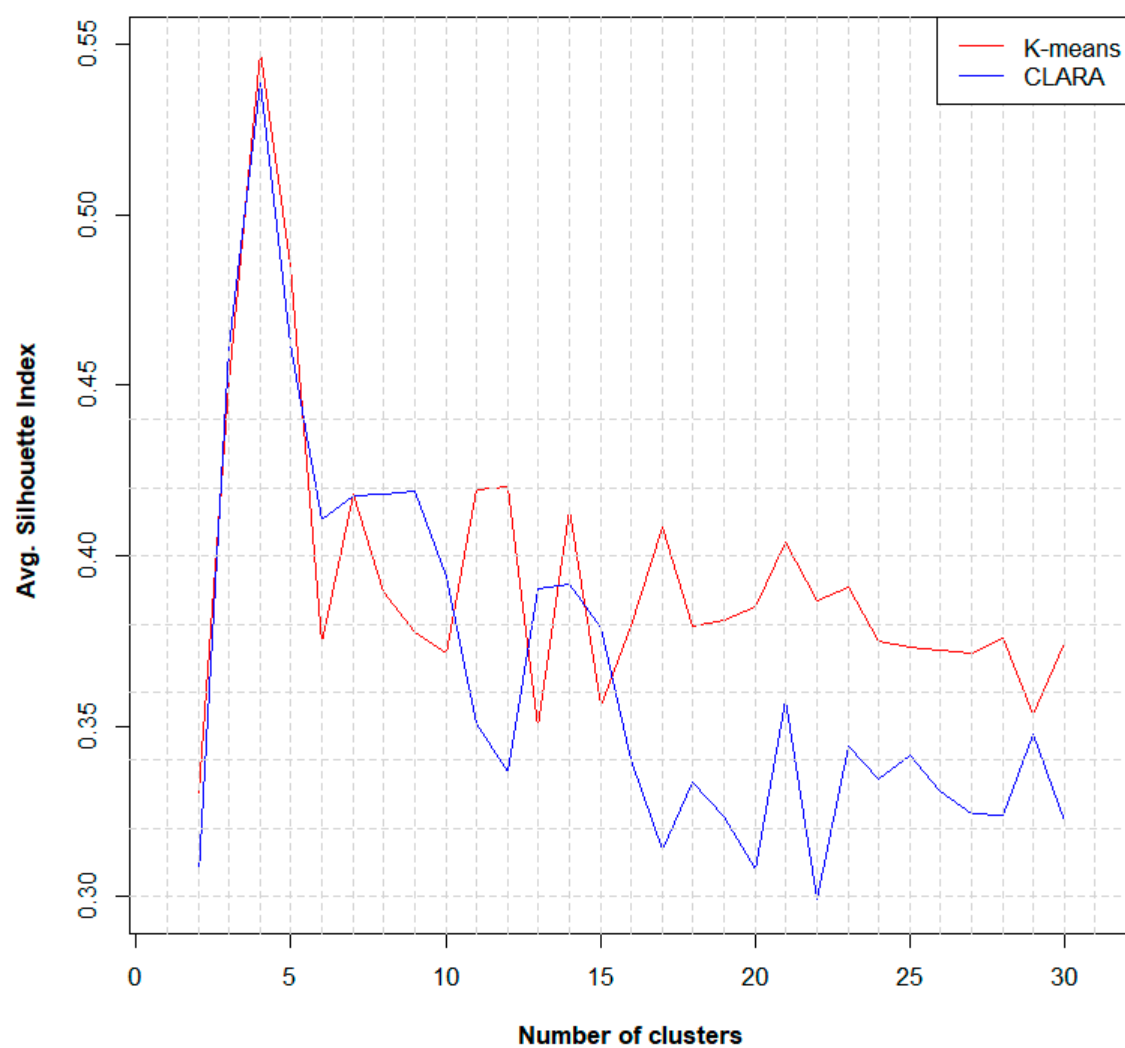


Figure S24. The values of the Silhouette index for the k-means and the CLARA unsupervised clustering algorithms regarding the optimal number of spatial clusters of genetic composition for *Helleborus odorus* subsp. *cyclophyllus* for the BCC RCP 4.5 time-period.

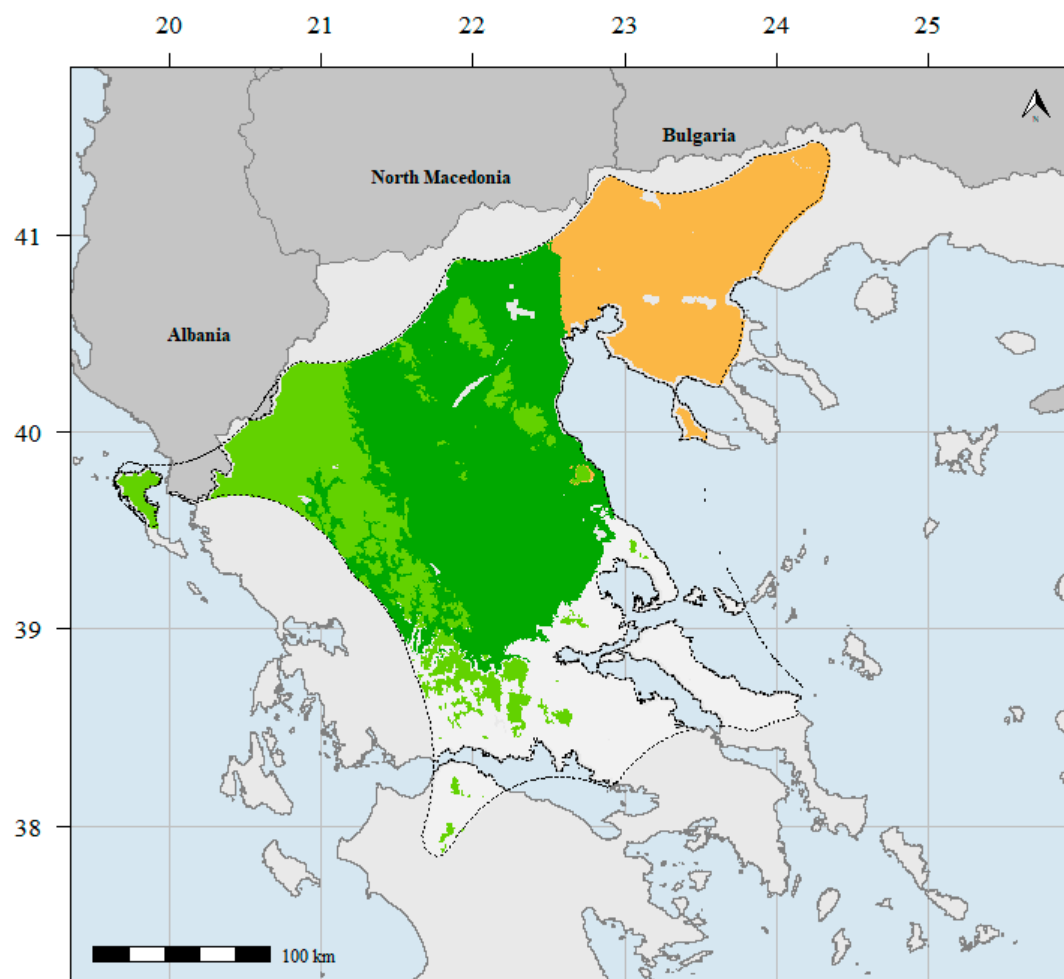


Figure S25: Predicted spatial variation for BCC RCP 6.0 time-period in population-level genetic composition from GDM. Different colours represent gradients in genetic turnover. Locations with similar colours are expected to harbour populations with similar genetic composition.

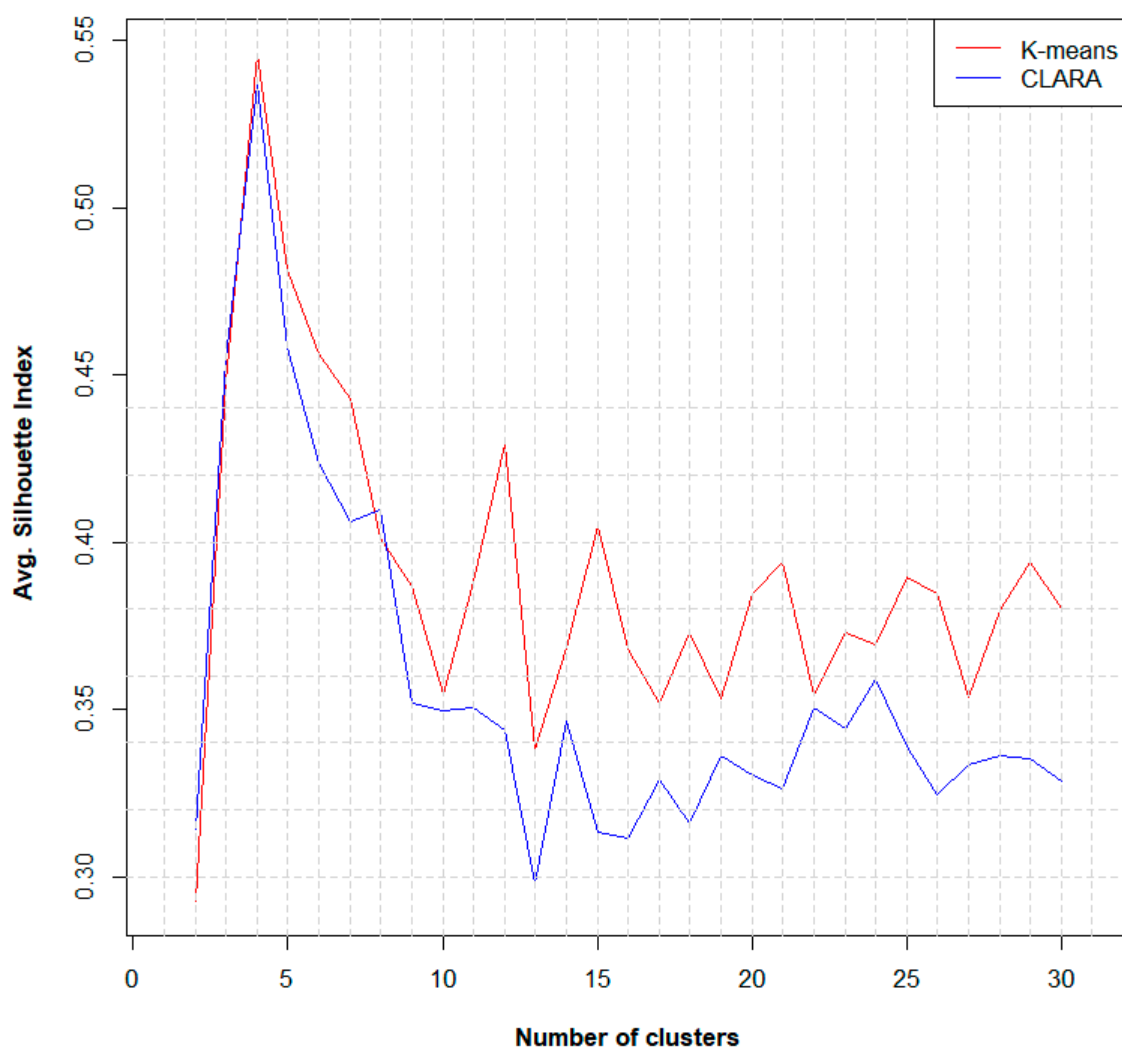


Figure S26: The values of the Silhouette index for the k-means and the CLARA unsupervised clustering algorithms regarding the optimal number of spatial clusters of genetic composition for *Helleborus odorus* subsp. *cyclophyllus* for the BCC RCP 6.0 time-period.

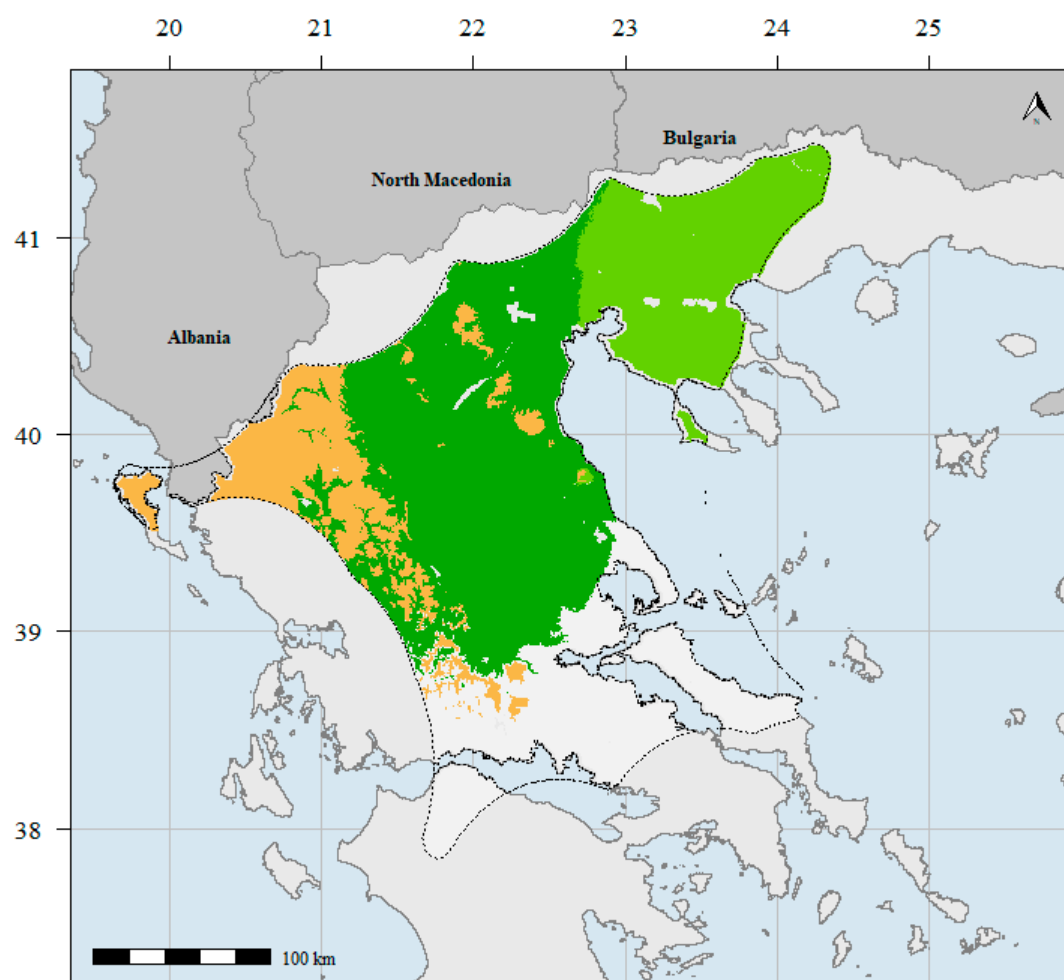


Figure S27: Predicted spatial variation for the BCC RCP 8.5 time-period in population-level genetic composition from GDM. Different colours represent gradients in genetic turnover. Locations with similar colours are expected to harbour populations with similar genetic composition.

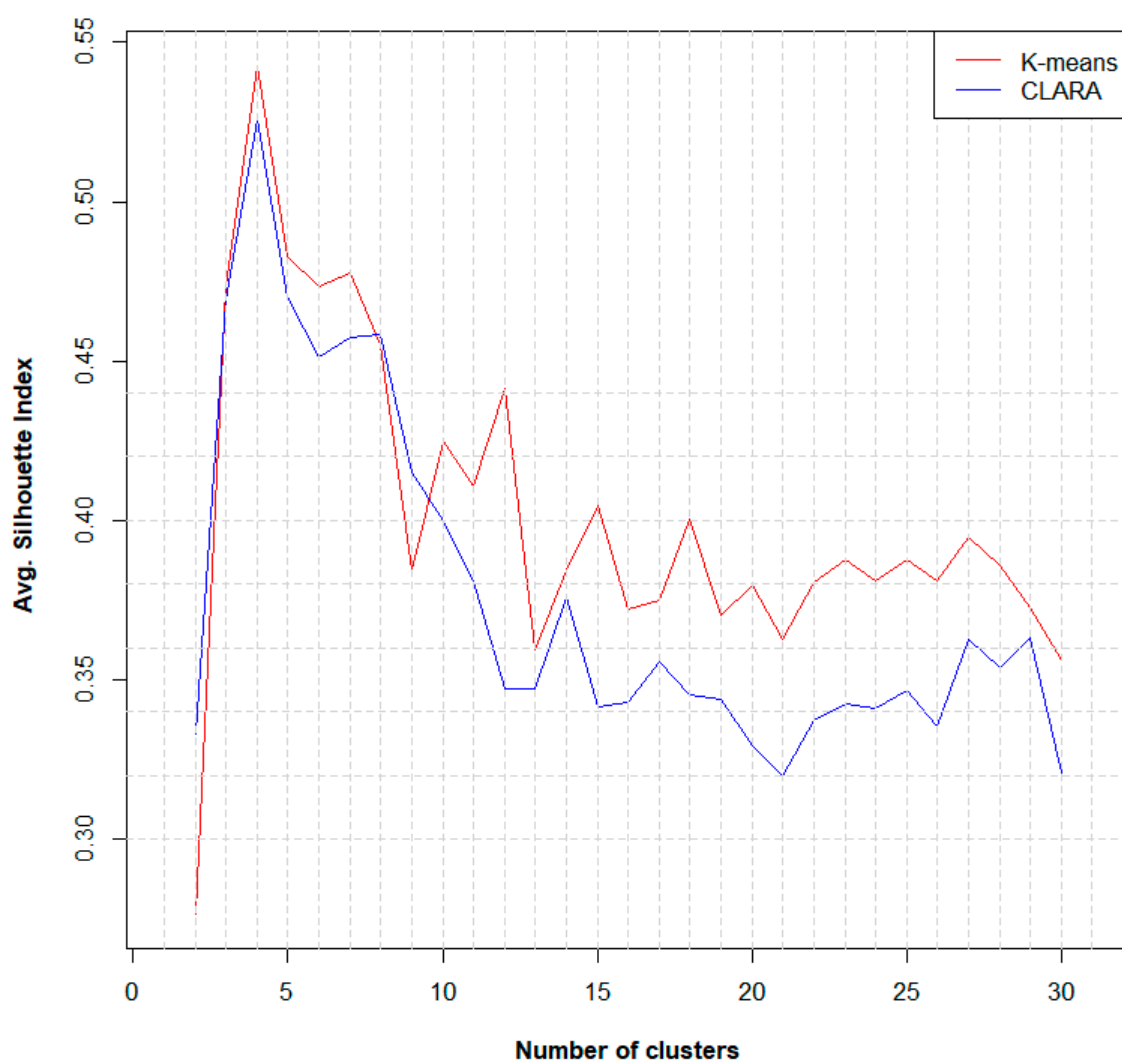


Figure S28: The values of the Silhouette index for the k-means and the CLARA unsupervised clustering algorithms regarding the optimal number of spatial clusters of genetic composition for *Helleborus odorus* subsp. *cyclophyllus* for the BCC RCP 8.5 time-period.

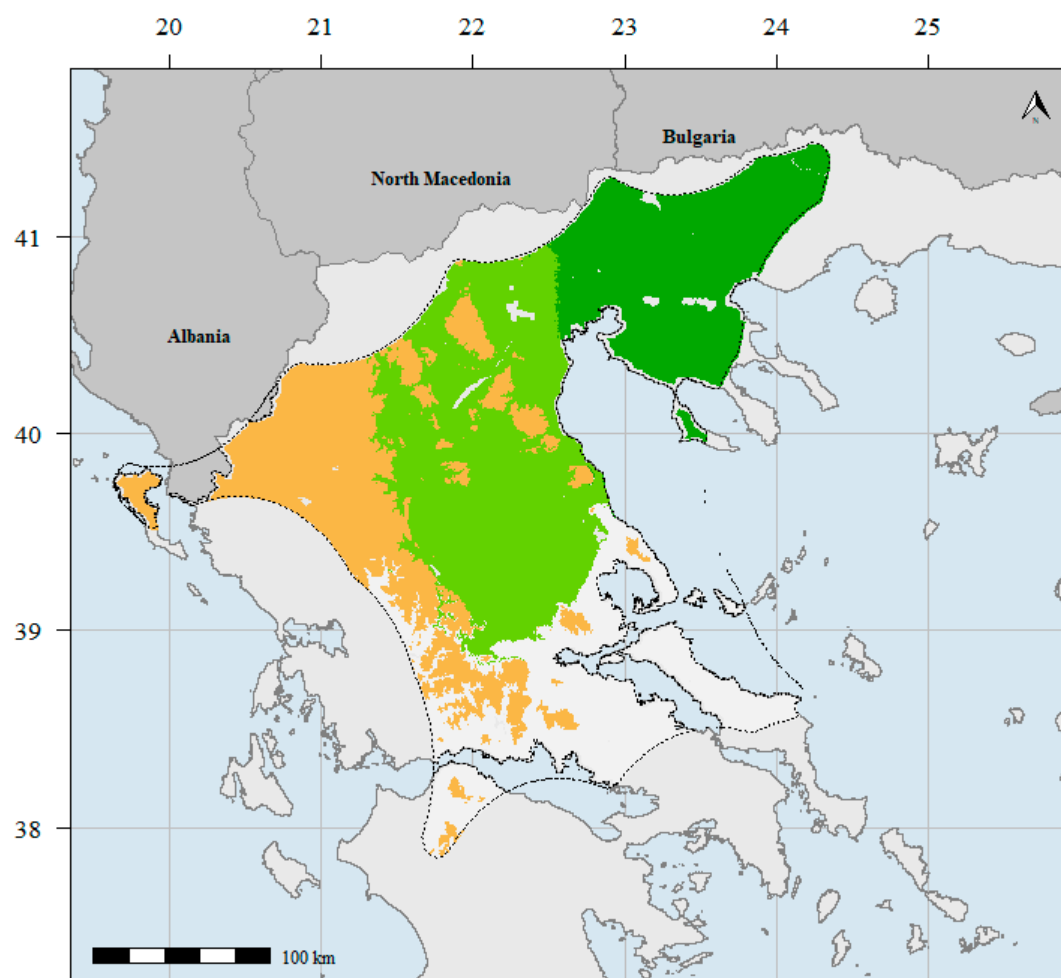


Figure S29: Predicted spatial variation for the CCSM4 RCP 2.6 time-period in population-level genetic composition from GDM. Different colours represent gradients in genetic turnover. Locations with similar colours are expected to harbour populations with similar genetic composition.

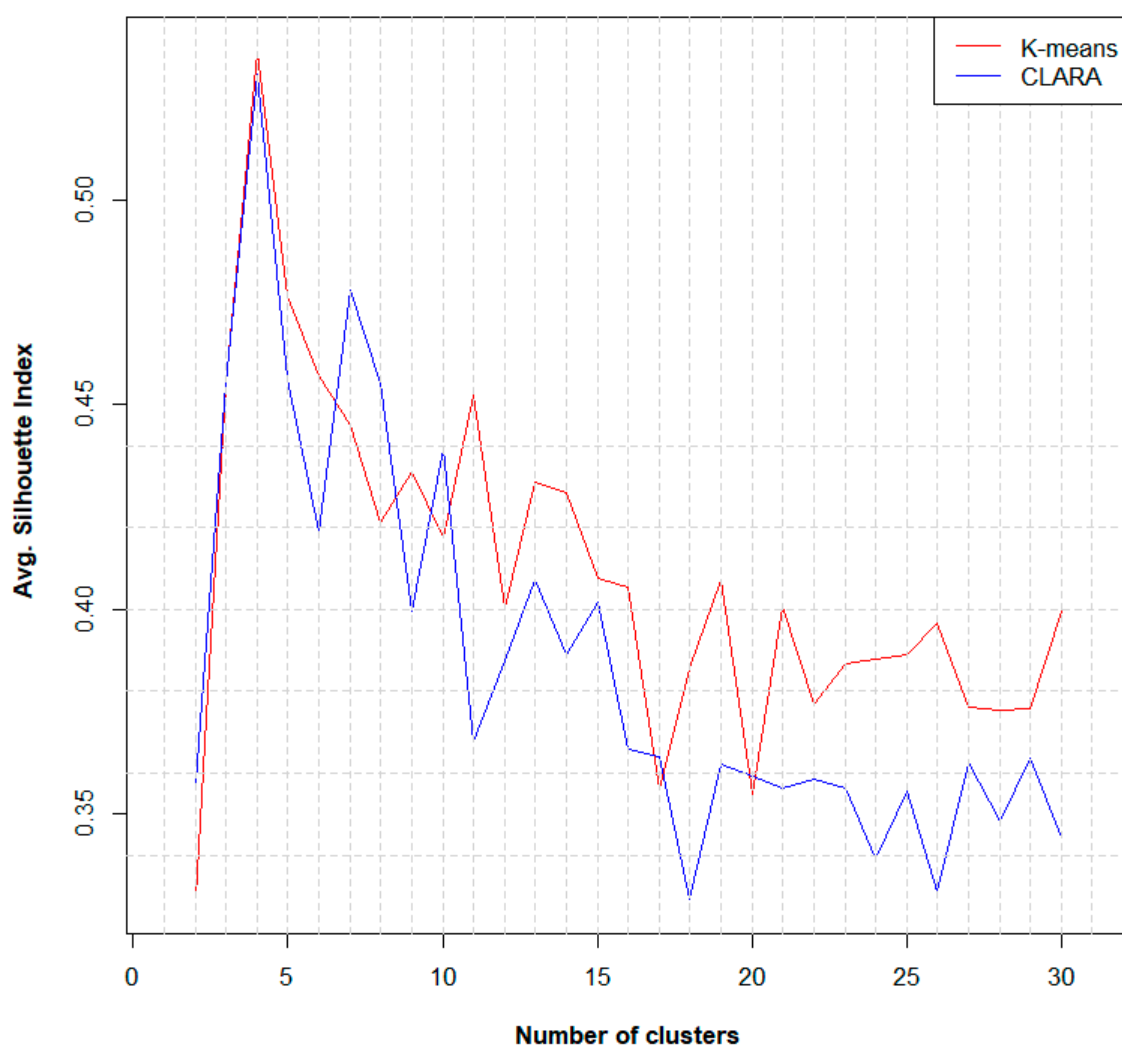


Figure S30: The values of the Silhouette index for the k-means and the CLARA unsupervised clustering algorithms regarding the optimal number of spatial clusters of genetic composition for *Helleborus odorus* subsp. *cyclophyllus* for the CCSM4 RCP 2.6 time-period.

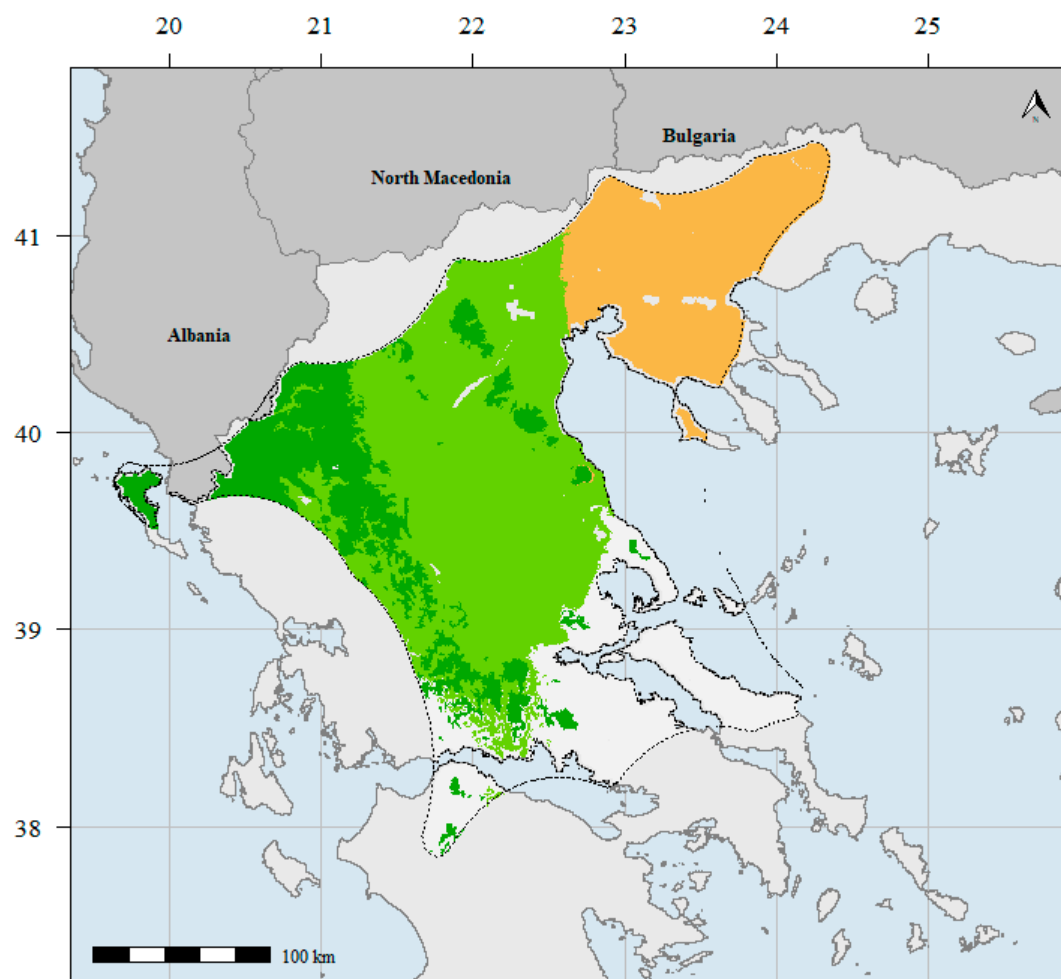


Figure S31: Predicted spatial variation for CCSM4 RCP 4.5 time-period in population-level genetic composition from GDM. Different colours represent gradients in genetic turnover. Locations with similar colours are expected to harbour populations with similar genetic composition.

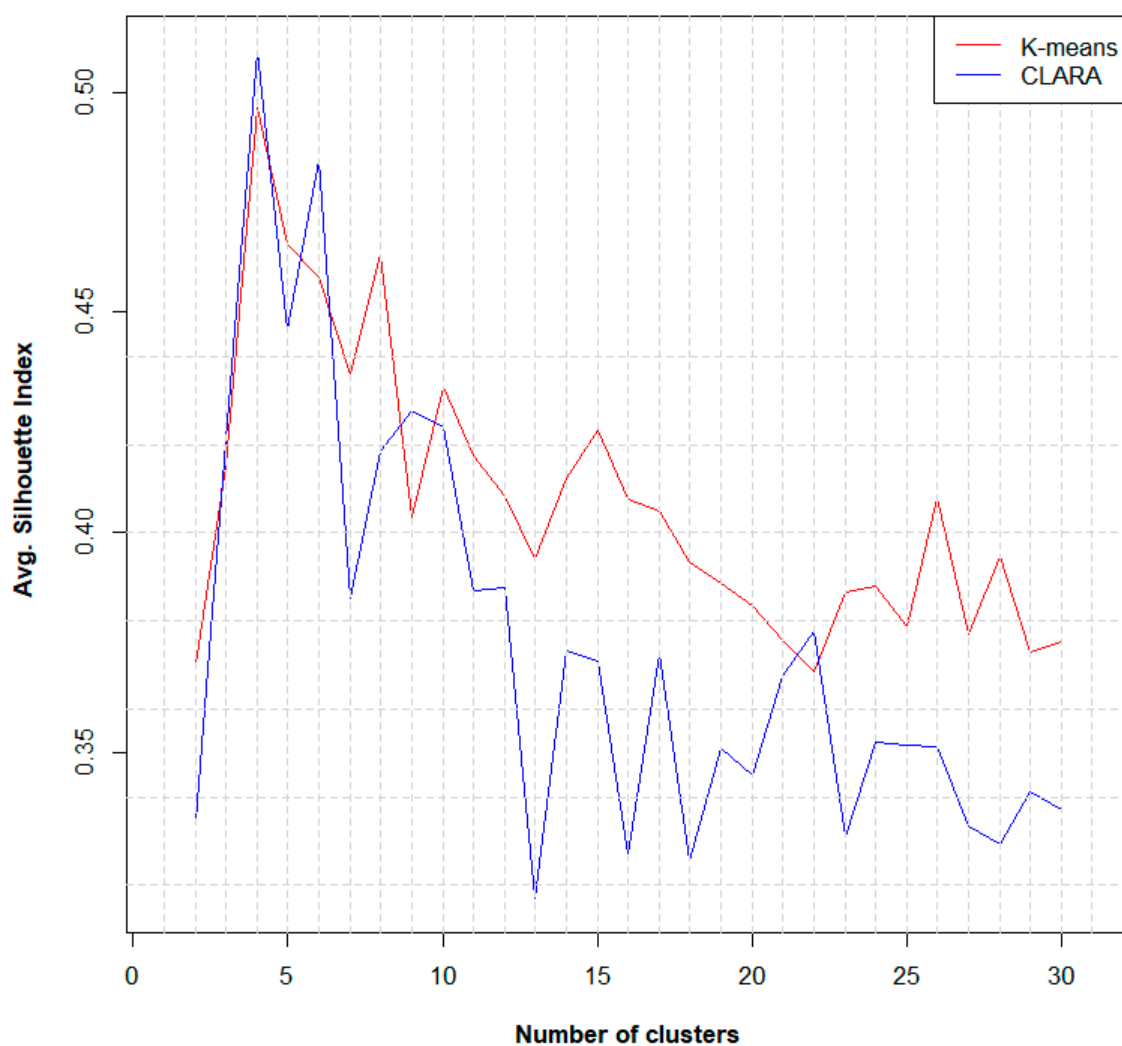


Figure S32: The values of the Silhouette index for the k-means and the CLARA unsupervised clustering algorithms regarding the optimal number of spatial clusters of genetic composition for *Helleborus odorus* subsp. *cyclophyllus* for the CCSM4 RCP 4.5 time-period.

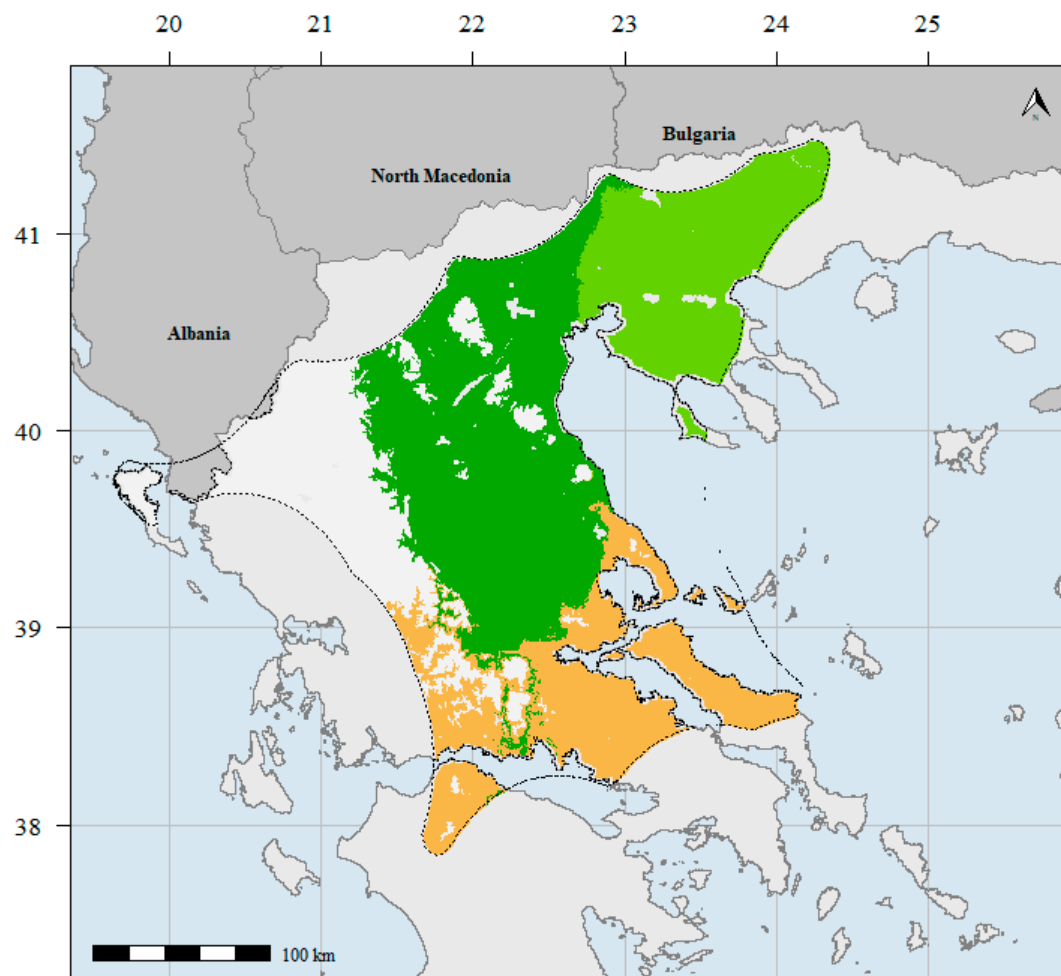


Figure S33: Predicted spatial variation for the CCSM4 RCP 6.0 time-period in population-level genetic composition from GDM. Different colours represent gradients in genetic turnover. Locations with similar colours are expected to harbour populations with similar genetic composition.

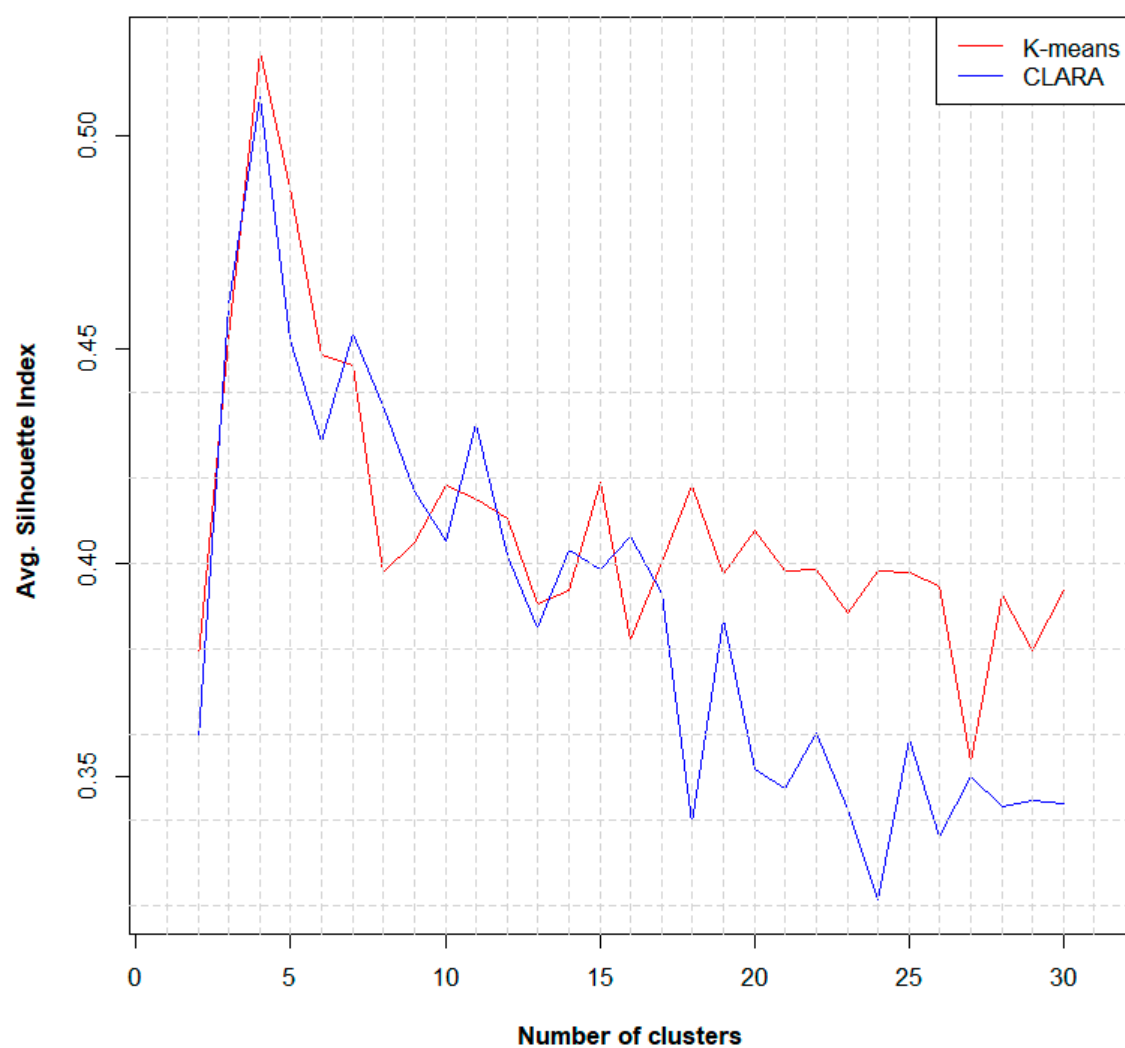


Figure S34: The values of the Silhouette index for the k-means and the CLARA unsupervised clustering algorithms regarding the optimal number of spatial clusters of genetic composition for *Helleborus odorus* subsp. *cyclophyllus* for the CCSM4 RCP 6.0 time-period.

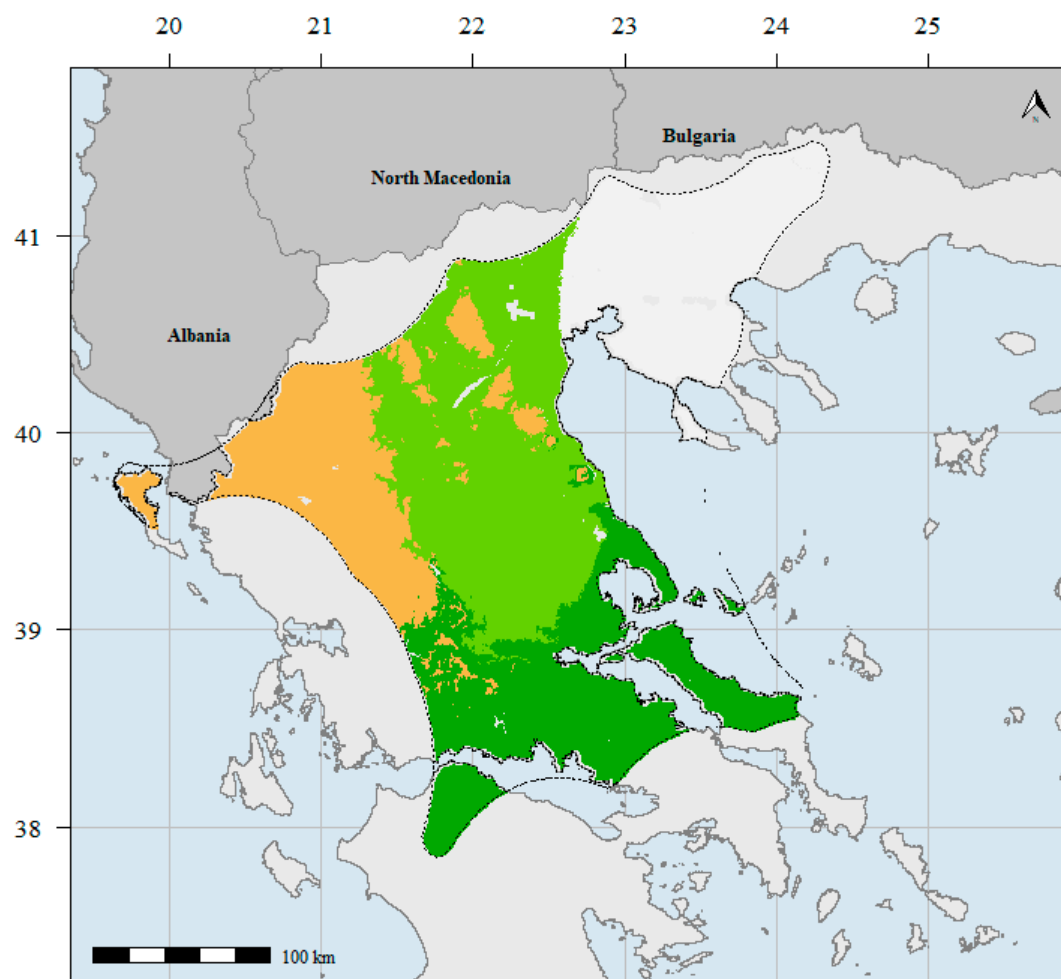


Figure S35: Predicted spatial variation for the CCSM4 RCP 8.5 time-period in population-level genetic composition from GDM. Different colours represent gradients in genetic turnover. Locations with similar colours are expected to harbour populations with similar genetic composition.

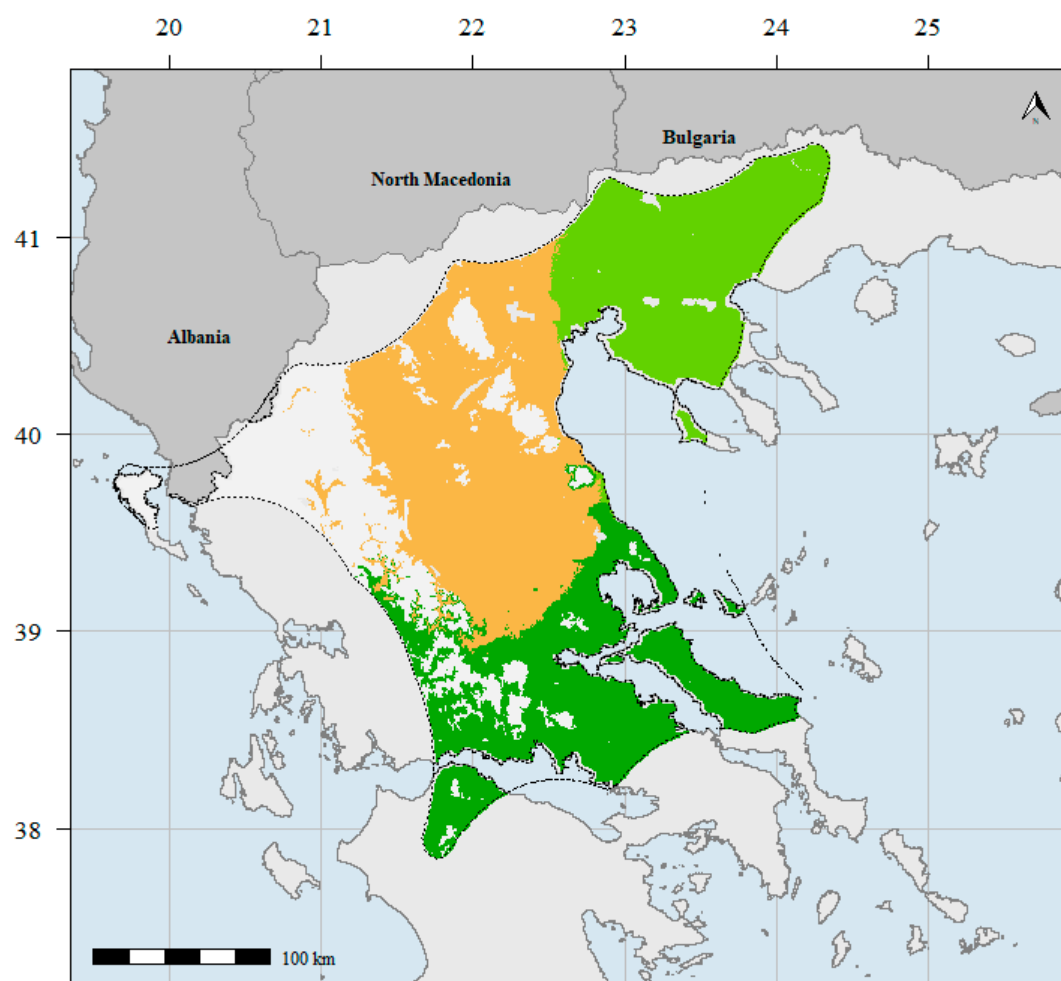


Figure S36: The values of the Silhouette index for the k-means and the CLARA unsupervised clustering algorithms regarding the optimal number of spatial clusters of genetic composition for *Helleborus odorus* subsp. *cyclophyllus* for the CCSM4 RCP 8.5 time-period.

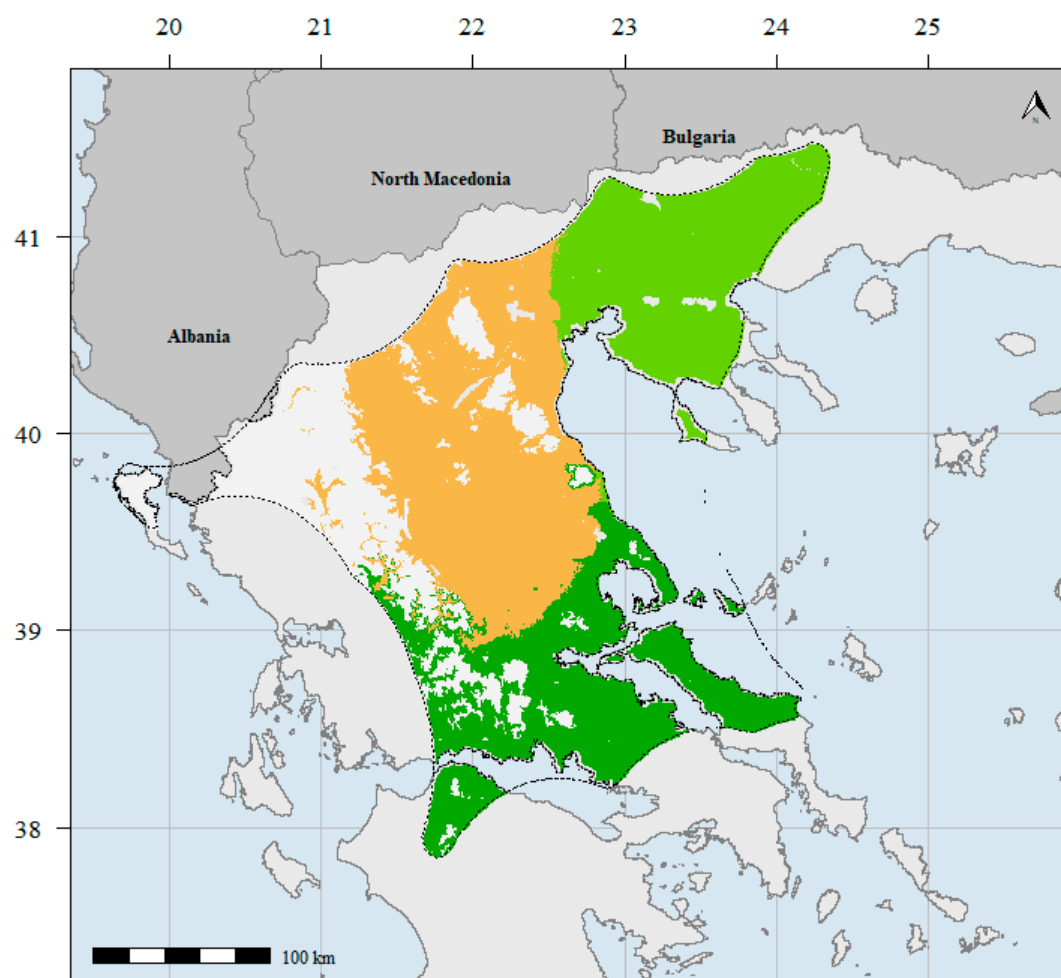


Figure S37. Predicted spatial variation for the HadGEM2 RCP 2.6 time-period in population-level genetic composition from GDM. Different colours represent gradients in genetic turnover. Locations with similar colours are expected to harbour populations with similar genetic composition.

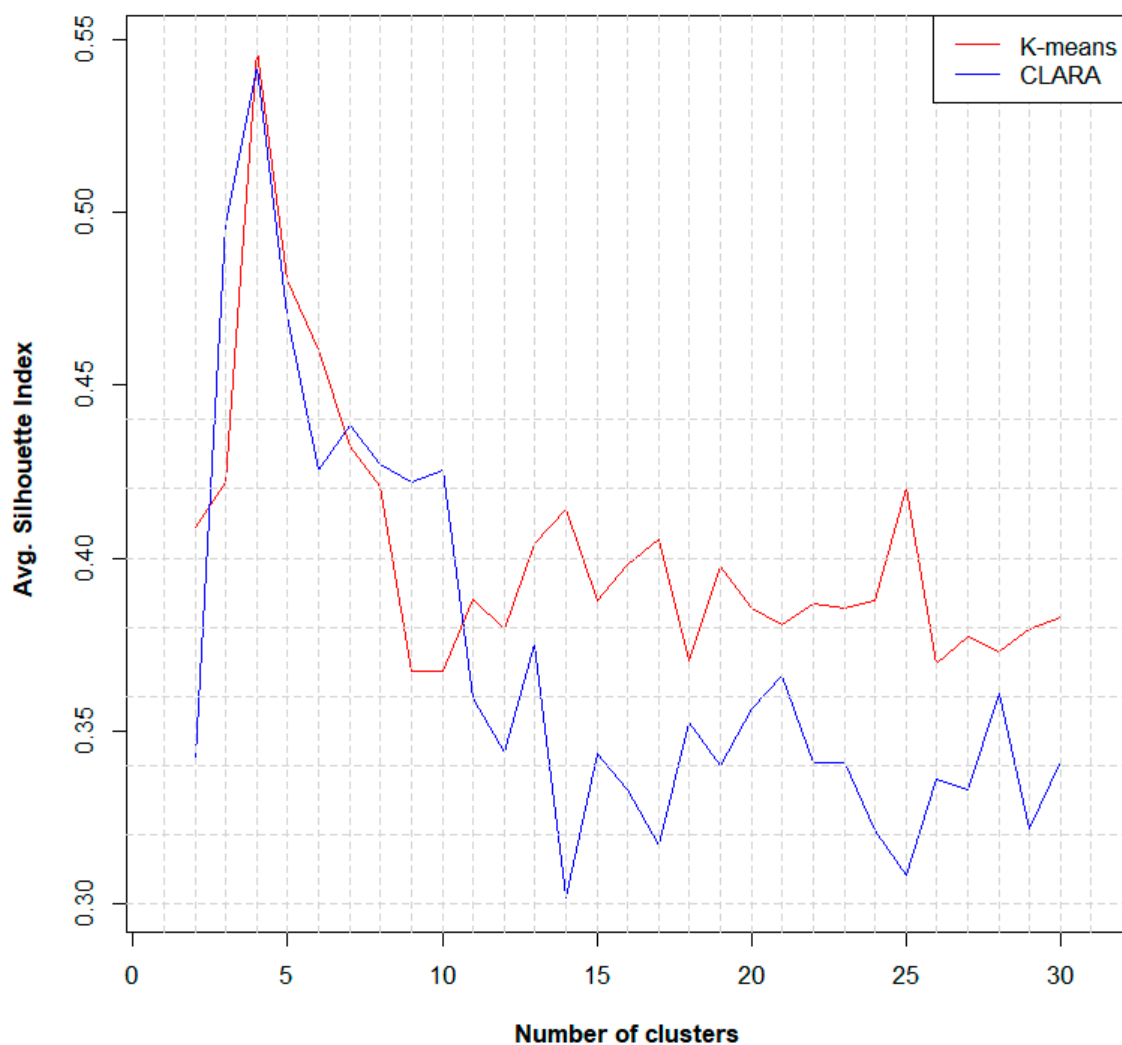


Figure S38: The values of the Silhouette index for the k-means and the CLARA unsupervised clustering algorithms regarding the optimal number of spatial clusters of genetic composition for *Helleborus odorus* subsp. *cyclophyllus* for the HadGEM2 RCP 2.6 time-period.

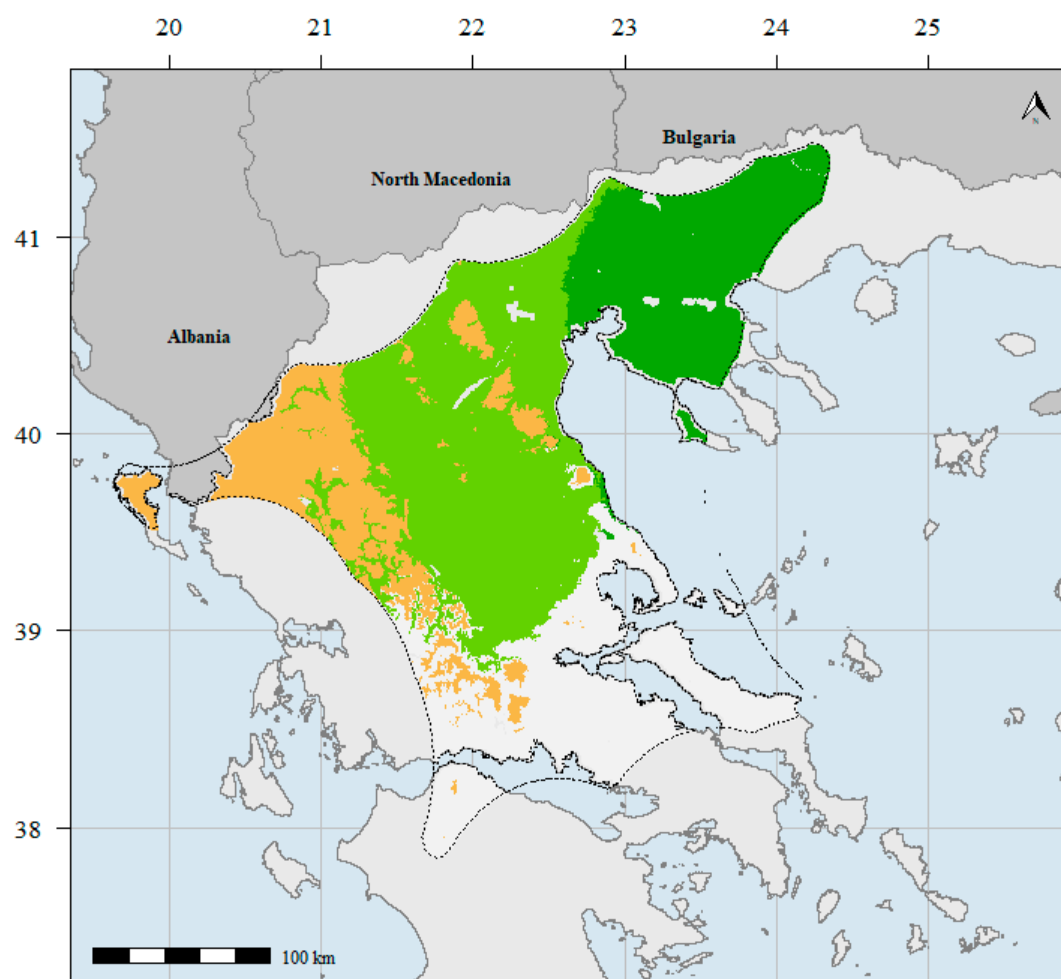


Figure S39: Predicted spatial variation for the HadGEM2 RCP 4.5 time-period in population-level genetic composition from GDM. Different colours represent gradients in genetic turnover. Locations with similar colours are expected to harbour populations with similar genetic composition.

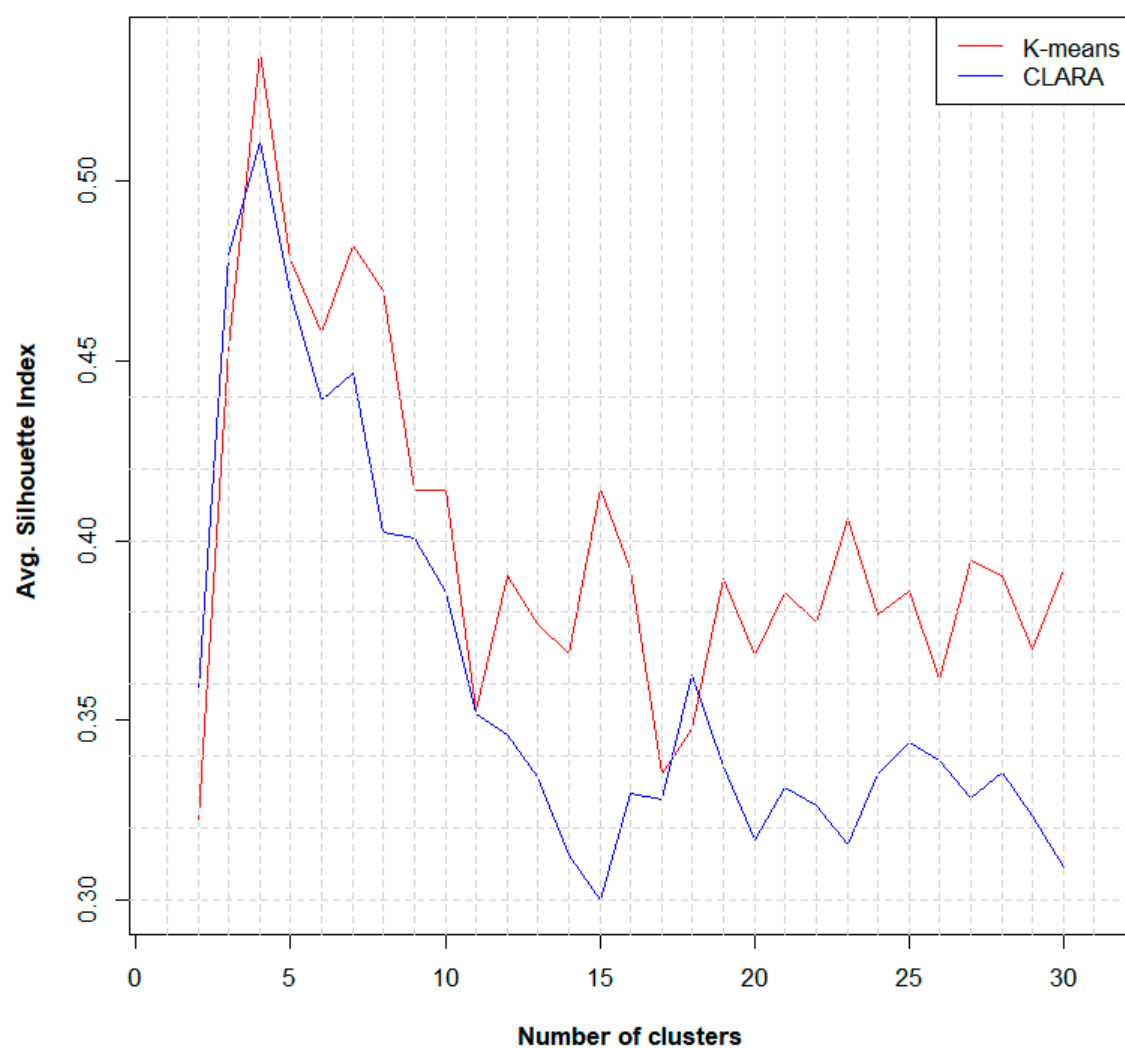


Figure S40: The values of the Silhouette index for the k-means and the CLARA unsupervised clustering algorithms regarding the optimal number of spatial clusters of genetic composition for *Helleborus odorus* subsp. *cyclophyllus* for the HadGEM2 RCP 4.5 time-period.

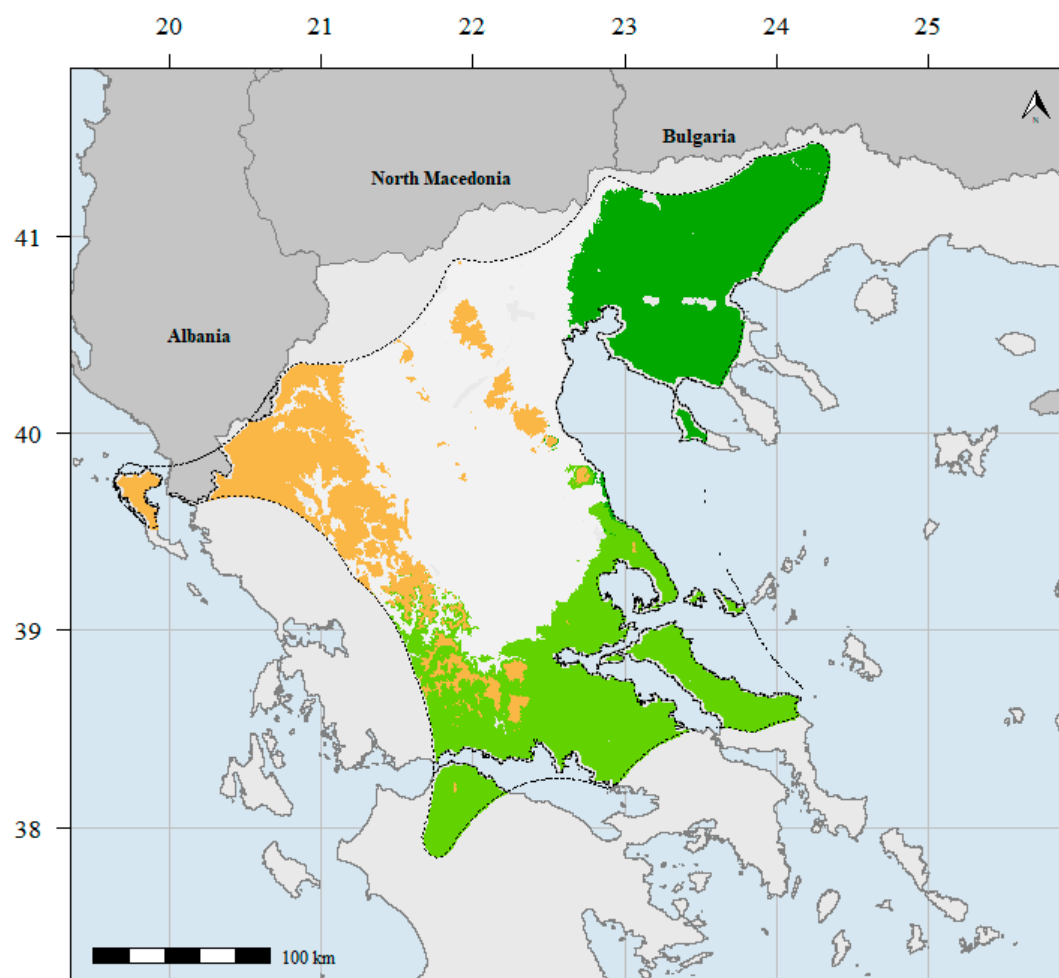


Figure S41: Predicted spatial variation for the HadGEM2 RCP 6.0 time-period in population-level genetic composition from GDM. Different colours represent gradients in genetic turnover. Locations with similar colours are expected to harbour populations with similar genetic composition.

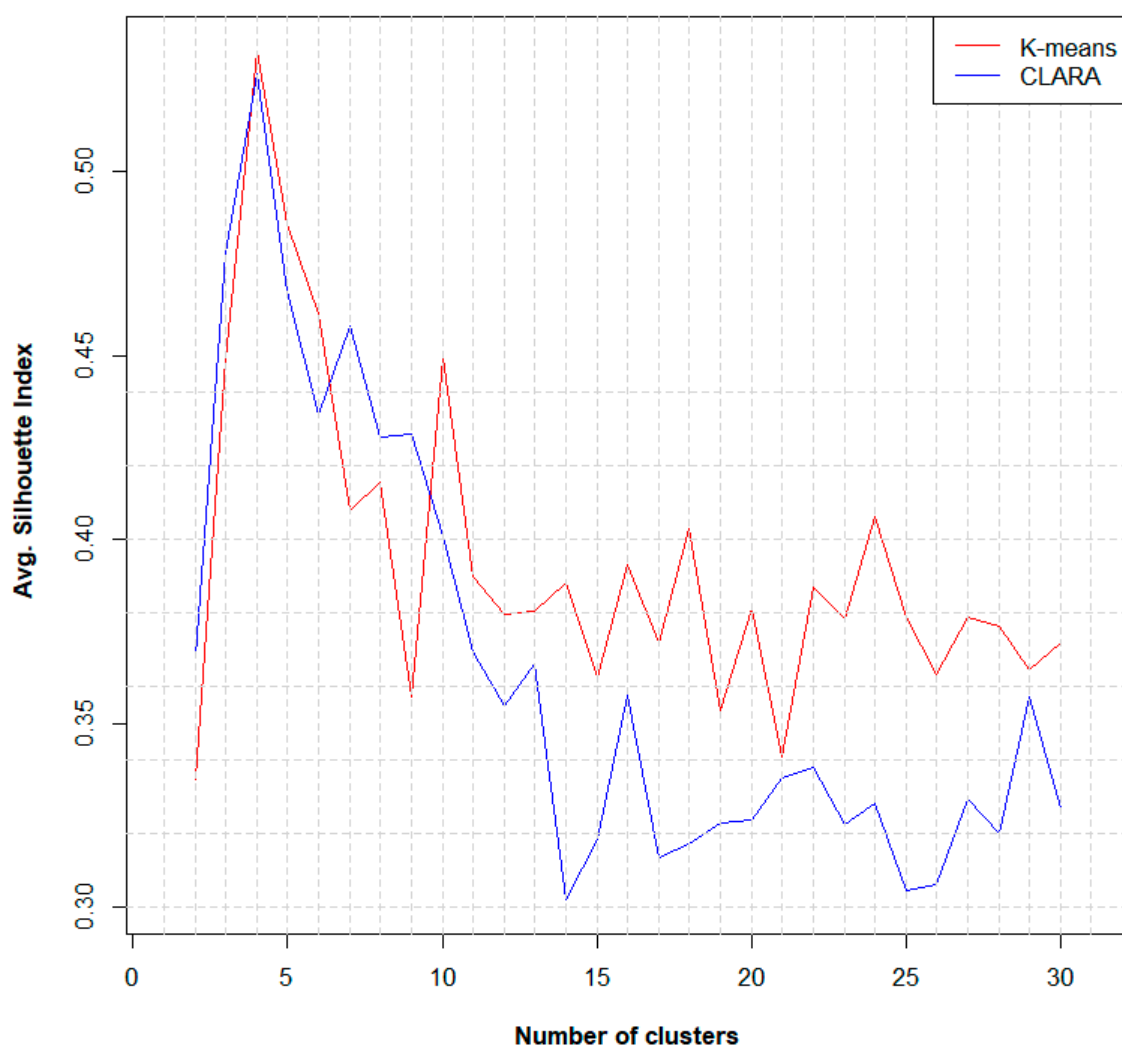


Figure S42: The values of the Silhouette index for the k-means and the CLARA unsupervised clustering algorithms regarding the optimal number of spatial clusters of genetic composition for *Helleborus odorus* subsp. *cyclophyllus* for the HadGEM2 RCP 6.0 time-period.

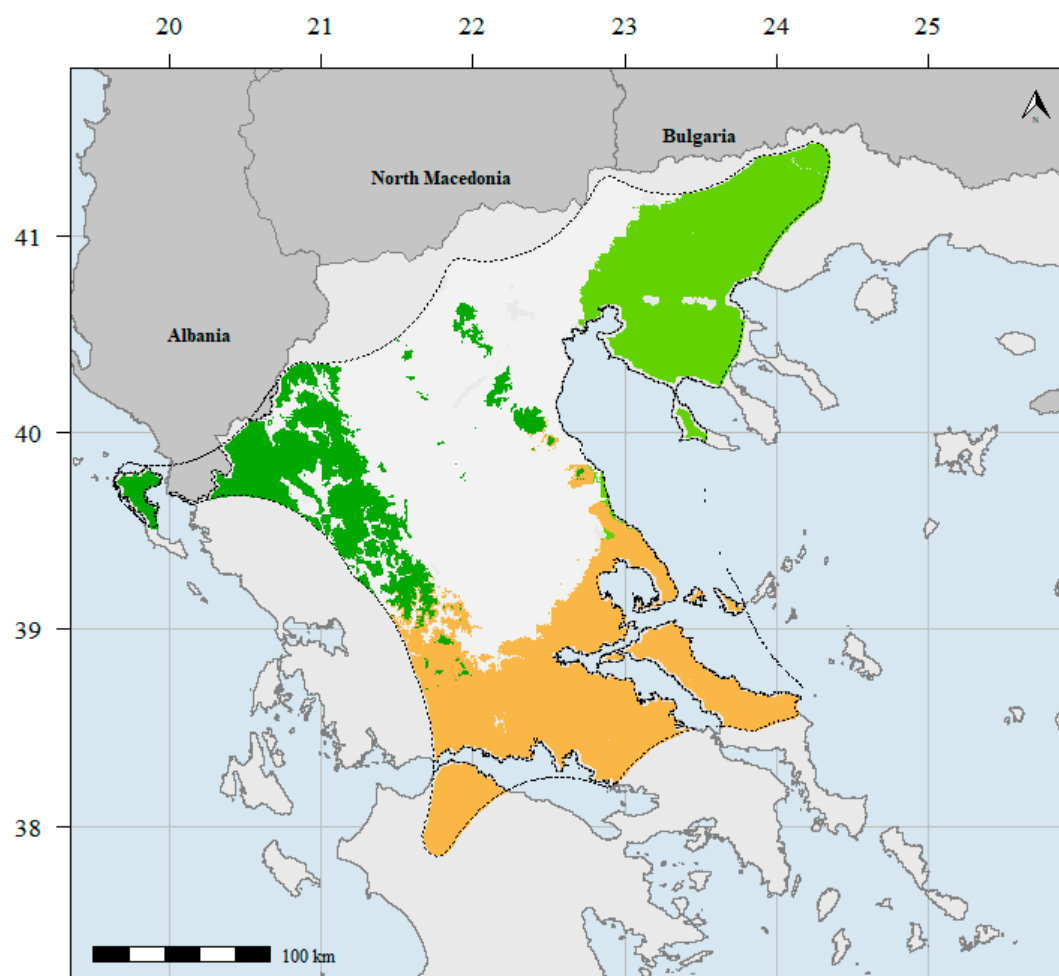


Figure S43: Predicted spatial variation for the HadGEM2 RCP 8.5 time-period in population-level genetic composition from GDM. Different colours represent gradients in genetic turnover. Locations with similar colours are expected to harbour populations with similar genetic composition.

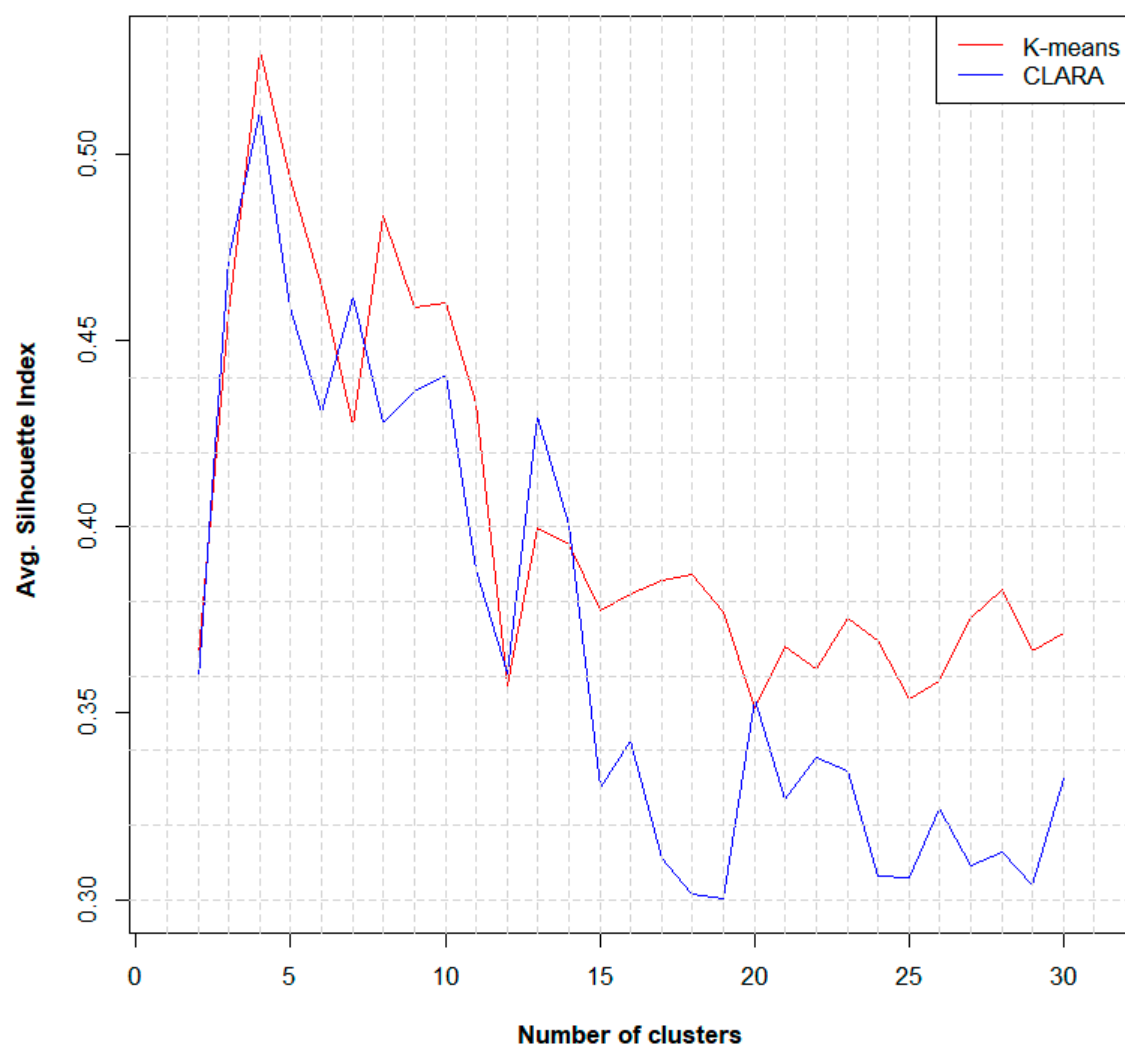


Figure S44: The values of the Silhouette index for the k-means and the CLARA unsupervised clustering algorithms regarding the optimal number of spatial clusters of genetic composition for *Helleborus odorus* subsp. *cyclophyllus* for the HadGEM2 RCP 8.5 time-period.

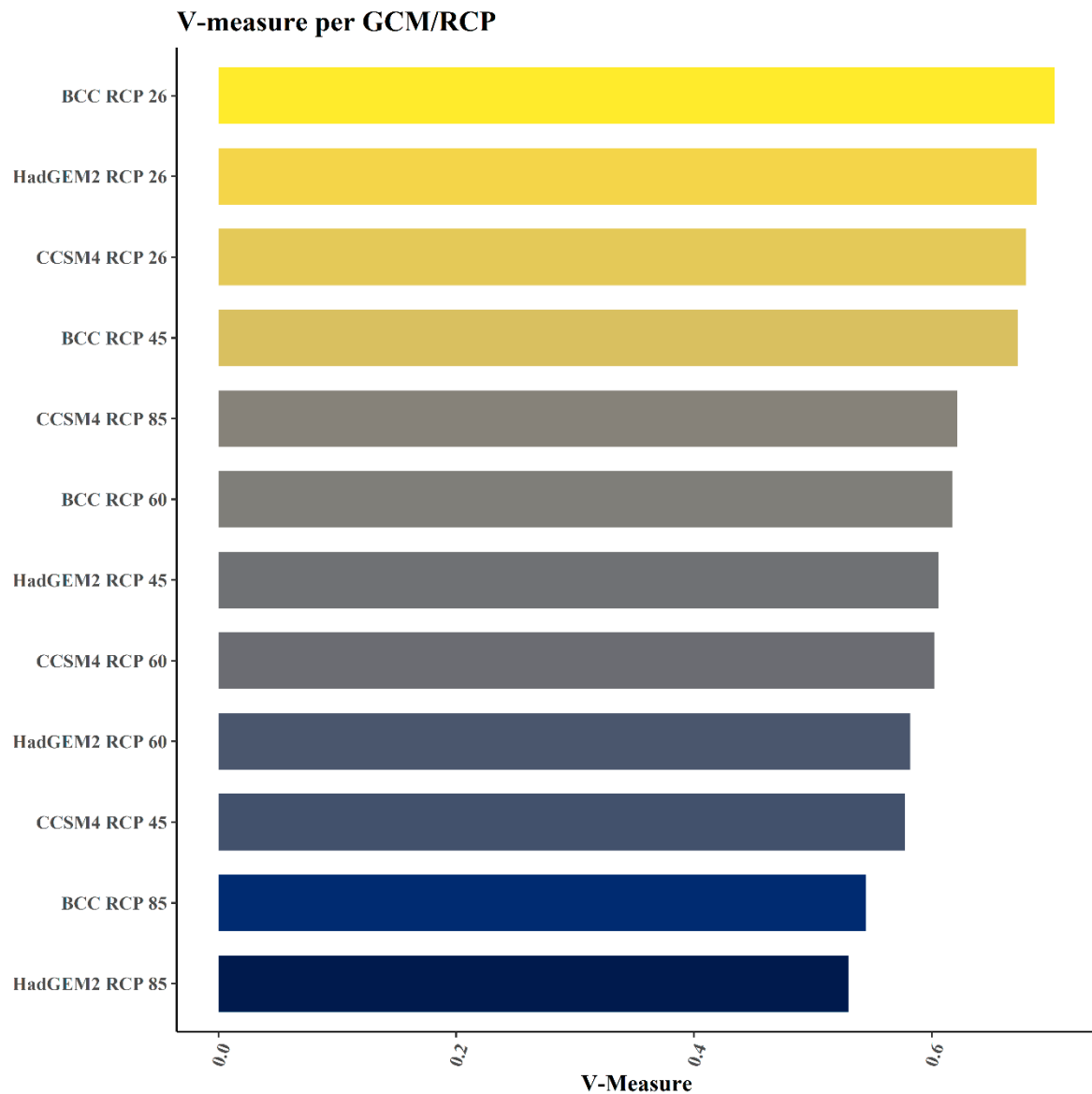


Figure S45: Similarity regarding the spatial clusters of genetic composition for *Helleborus odorus* subsp. *cyclophyllus* between the present and each Global Circulation Model (GCM) and Representative Concentration Pathway (RCP), based on the V-measure index.

Table 1. Pairwise Fst values between populations.

ERY	*						
VEL	0.1018	*					
PAR	0.1757	0.1101	*				
VOU	0.1015	0.0572	0.0862	*			
FRAG	0.1234	0.0597	0.1272	0.0317	*		
NAOU	0.1326	0.0743	0.1254	0.0787	0.1079	*	
OLY	0.2140	0.1500	0.2906	0.1922	0.2134	0.2212	*
DIR	0.2171	0.1740	0.1756	0.1591	0.2010	0.1530	0.3316 *

References

72. Sagarin, R.; Gaines, S.D. The 'abundant centre' distribution: To what extent is it a biogeographical rule? *Ecol. Lett.* **2002**, *5*, 137–147, doi:10.1046/j.1461-0248.2002.00297.x.
73. Elith, J.; Kearney, M.R.; Phillips, S. The art of modelling range-shifting species. *Methods Ecol. Evol.* **2010**, *1*, 330–342, doi:10.1111/j.2041-210x.2010.00036.x.

74. Aiello-Lammens, M.E.; Boria, R.A.; Radosavljevic, A.; Vilela, B.; Anderson, R.P.; Silva, D.P. spThin: An R package for spatial thinning of species occurrence records for use in ecological niche models. *Ecography* **2015**, *38*, 541–545, doi:10.1111/ecog.01132.
75. Varela, S.; Anderson, R.P.; García-Valdés, R.; Fernández-González, F. Environmental filters reduce the effects of sampling bias and improve predictions of ecological niche models. *Ecography* **2014**, *37*, 1084–1091, doi:10.1111/j.1600-0587.2013.00441.x.
76. Zizka, A.; Antonelli, A.; Silvestro, D. Sampbias, a Method for Quantifying Geographic Sampling Biases in Species Distribution Data. *bioRxiv* **2020** doi:10.1101/2020.01.13.903757.
77. Title, P.; Bemmels, J.B. ENVIREM: An expanded set of bioclimatic and topographic variables increases flexibility and improves performance of ecological niche modeling. *Ecography* **2017**, *41*, 291–307, doi:10.1111/ecog.02880.
78. McSweeney, C.; Jones, R.G.; Lee, R.W.; Rowell, D.P. Selecting CMIP5 GCMs for downscaling over multiple regions. *Clim. Dyn.* **2014**, *44*, 3237–3260, doi:10.1007/s00382-014-2418-8.
79. Hengl, T.; de Jesus, J.M.; Heuvelink, G.B.M.; Gonzalez, M.R.; Kilibarda, M.; Blagotić, A.; Shangguan, W.; Wright, M.N.; Geng, X.; Bauer-Marschallinger, B.; et al. SoilGrids250m: Global gridded soil information based on machine learning. *PLoS ONE* **2017**, *12*, e0169748, doi:10.1371/journal.pone.0169748.
80. SoilGrids250m Database Available online: <https://www.soilgrids.org> (accessed on 31 April 2020).
81. CGIAR-CSI Data-Portal Available online: <http://srtm.csi.cgiar.org> (accessed on 31 April 2020).
82. Pebesma, E. Simple Features for R: Standardized Support for Spatial Vector Data. *R J.* **2018**, *10*, 439–446, doi:10.32614/rj-2018-009.
83. Ross, N. *fasterize: Fast Polygon to Raster Conversion*, version 1.0.2; CRAN: 2018.
84. Naimi, B.; Hamm, N.; Groen, T.A.; Skidmore, A.K.; Toxopeus, A.G. Where is positional uncertainty a problem for species distribution modelling? *Ecography* **2013**, *37*, 191–203, doi:10.1111/j.1600-0587.2013.00205.x.
85. Broennimann, O.; Cola, V. Di; Petitpierre, B.; Breiner, F.; D'Amen, M.; Randin, C.; Engler, R.; Hordijk, W.; Pottier, J.; Febbraro, M. Di; et al. *ecospat: Spatial Ecology Miscellaneous Methods*, version 3.1; CRAN: 2018.
86. Barbet-Massin, M.; Jiguet, F.; Albert, C.H.; Thuiller, W. Selecting pseudo-absences for species distribution models: How, where and how many? *Methods Ecol. Evol.* **2012**, *3*, 327–338, doi:10.1111/j.2041-210x.2011.00172.x.
87. Dauby, G. *ConR: Computation of Parameters Used in Preliminary Assessment of Conservation Status*, version 1.2.4; CRAN: 2019.
88. Valavi, R.; Elith, J.; Lahoz-Monfort, J.J.; Guillera-Arroita, G. block CV: An r package for generating spatially or environmentally separated folds for k-fold cross-validation of species distribution models. *Methods Ecol. Evol.* **2018**, *10*, 225–232, doi:10.1111/2041-210x.13107.
89. Breiner, F.T.; Guisan, A.; Bergamini, A.; Nobis, M.P. Overcoming limitations of modelling rare species by using ensembles of small models. *Methods Ecol. Evol.* **2015**, *6*, 1210–1218, doi:10.1111/2041-210x.12403.
90. Allouche, O.; Tsoar, A.; Kadmon, R. Assessing the accuracy of species distribution models: Prevalence, kappa and the true skill statistic (TSS). *J. Appl. Ecol.* **2006**, *43*, 1223–1232, doi:10.1111/j.1365-2664.2006.01214.x.
91. Raes, N.; Ter Steege, H. A null-model for significance testing of presence-only species distribution models. *Ecography* **2007**, *30*, 727–736, doi:10.1111/j.2007.0906-7590.05041.x.
92. Fitzpatrick, M.C.; Sanders, N.J.; Normand, S.; Svenning, J.-C.; Ferrier, S.; Gove, A.D.; Dunn, R.R. Environmental and historical imprints on beta diversity: Insights from variation in rates of species turnover along gradients. *Proc. R. Soc. B Biol. Sci.* **2013**, *280*, 20131201, doi:10.1098/rspb.2013.1201.
93. Fitzpatrick, M.C.; Sanders, N.J.; Ferrier, S.; Longino, J.T.; Weiser, M.D.; Dunn, R.R. Forecasting the future of biodiversity: A test of single- and multi-species models for ants in North America. *Ecography* **2011**, *34*, 836–847, doi:10.1111/j.1600-0587.2011.06653.x.
94. Rousseeuw, P.J. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.* **1987**, *20*, 53–65, doi:10.1016/0377-0427(87)90125-7.
95. Rosenberg, A.; Hirschberg, J. V-Measure: A conditional entropy-based external cluster evaluation measure. In Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL), Prague, Czech Republic, June 2007; pp. 410–420.

96. Nowosad, J.; Stepinski, T.F. Spatial association between regionalizations using the information-theoretical V-measure. *Int. J. Geogr. Inf. Sci.* **2018**, *32*, 1–16, doi:10.1080/13658816.2018.1511794.
97. Maechler, M.; Rousseeuw, P.; Struyf, A.; Hubert, M.; Hornik, K. *cluster: Cluster Analysis Basics and Extensions*, version 2.1.0; CRAN: 2019.
98. Desgraupes, B. *clusterCrit: Clustering Indices*, version 1.2.8; CRAN: 2018.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).