



Addressing alpine plant phylogeography using integrative distributional, demographic and coalescent modeling

Dennis J. Larsson¹ · Da Pan² · Gerald M. Schneeweiss¹

Received: 7 March 2021 / Accepted: 5 July 2021 / Published online: 29 July 2021
© The Author(s) 2021

Abstract

Phylogeographic studies of alpine plants have evolved considerably in the last two decades from ad hoc interpretations of genetic data to statistical model-based approaches. In this review we outline the developments in alpine plant phylogeography focusing on the recent approach of integrative distributional, demographic and coalescent (iDDC) modeling. By integrating distributional data with spatially explicit demographic modeling and subsequent coalescent simulations, the history of alpine species can be inferred and long-standing hypotheses, such as species-specific responses to climate change or survival on nunataks during the last glacial maximum, can be efficiently tested as exemplified by available case studies. We also discuss future prospects and improvements of iDDC.

Keywords Alpine plants · iDDC · Demographic modeling · Distribution modeling · Phylogeography

Introduction

Phylogeography is the study of how historical processes have shaped the geographic distributions of genetic lineages within a species or among closely related species (Avice 2000). As such, it answers questions such as species response to past glaciations, gene flow (or lack thereof) between geographically disparate populations or the role of vicariance versus dispersal in shaping distributions. As a field phylogeography has grown considerably since its introduction by Avice (1987) both in application and in methods utilized (Stehlik et al. 2001; Ikeda et al. 2009; He et al. 2013; Theodoridis et al. 2017). For a general overview on phylogeography and its methodological development, the reader is

referred to the excellent reviews by Avice (2009), Knowles (2004, 2009), and Hickerson et al. (2010).

The progress of alpine phylogeography has seen our understanding of where alpine plants survived during the last glacial period (LGP, 115–11.7 ka) develop from tentative suggestions using qualitative data, usually derived from various nuclear and plastid markers under neutral evolution (Stehlik et al. 2001; Schönswetter et al. 2002; Kropf et al. 2003; Puşcaş et al. 2008; Ronikier et al. 2012; Gizaw et al. 2013; Wang et al. 2014), to explicit testing of hypotheses using sophisticated statistical models applied to large numbers of single nucleotide polymorphisms (SNPs) obtained from next generation sequencing (NGS) data (Theodoridis et al. 2017; Westergaard et al. 2019; Ikeda et al. 2020). This progress is also owed to the growing urgency to improve our insights into the phylogeographic histories of alpine plants: As many alpine species are under threat from the ongoing climate change (Freeman et al. 2018), comprehending how alpine plants survived past climate change may help us foresee if and how they will cope with future climate change. For instance, reduced snow depth and earlier timing of snowmelt induce changes in species-specific fitness of alpine plants growing in snowbeds, which may lead to changes in the species composition in those plant communities (Wipf et al. 2009). Other expected impacts due to a warming climate include changes in climatically suitable ranges (Hülber et al.

✉ Dennis J. Larsson
dennis.j.larsson@gmail.com

Da Pan
dapan@njnu.edu.cn

Gerald M. Schneeweiss
gerald.schneeweiss@univie.ac.at

¹ Department of Botany and Biodiversity Research, University of Vienna, Vienna, Austria

² Jiangsu Key Laboratory for Biodiversity and Biotechnology, College of Life Sciences, Nanjing Normal University, Nanjing, China

2020) and transformations of plant and pollinator compositions (Inouye 2020).

The phylogeography of alpine plants, however, is also interesting due to its model character. The rugged topography of mountains lead to complex interactions between temperature, precipitation, solar radiation, soil types, humidity, air pressure and the topography itself and gives rise to extreme environmental conditions that change dramatically over short distances, creating many different microhabitats (see Körner 2021 for an excellent overview on alpine plant ecology). Because alpine species are often restricted to their microhabitats, their distributions tend to be fragmentary and prone to considerable range shifts during changes in climate (Vargas 2003; Schönswetter et al. 2005; Schneeweiss and Schönswetter 2010). Range expansions and contractions and the associated changes in connectivity among populations result in complex phylogeographic patterns (Kadereit et al. 2004; Mráz et al. 2007; Dixon et al. 2009; Schneeweiss et al. 2017). By advancing our understanding of such patterns we can identify the general principles of how species distribute themselves over space and time.

In this review, we will outline the development of methodologies in alpine phylogeography, focusing on the relatively recent approach of integrative distributional, demographic and coalescent (iDDC) modeling. By integrating distributional data with spatially explicit demographic simulations and advanced statistical methods for model selection it becomes possible to test spatially explicit phylogeographic histories also in alpine plants, as already demonstrated by the few available studies. Finally, we will outline future prospects of the method and how it may be further improved.

Available methods

Interpretative phylogeography

Early phylogeographic studies on alpine plants used mostly qualitative methods, including ordinations such as principal coordinate analysis (PCoA; Schönswetter et al. 2002; Mráz et al. 2007), phylogenetic trees (Schönswetter et al. 2003; Puşcaş et al. 2008; Yang et al. 2012), hierarchical clustering using *STRUCTURE* (Pritchard et al. 2000) or related programs (Segarra-Moragues et al. 2007; Slovák et al. 2012), Mantel tests (Stehlik et al. 2001; Schönswetter et al. 2003), analysis of molecular variance (AMOVA; Stehlik et al. 2001; Kropf et al. 2003; DeChaine and Martin 2005; Ikeda and Setoguchi 2007; Chen et al. 2008) or the distribution of rare (i.e., being present in small proportion of individuals only) and private alleles (i.e., being restricted to usually a single population) of amplified fragment length polymorphisms (AFLPs) or of nuclear or plastid sequences (Mráz et al. 2007; Ronikier et al. 2012; Slovák et al. 2012; Gizaw

et al. 2013; Wang et al. 2014). These observed genetic patterns were then compared to those expected under specific hypotheses. For instance, the distribution of different gene pools in geographically distinct regions may be taken as evidence for the location of Pleistocene refugia (Schönswetter et al. 2002; Massatti and Knowles 2014; Schönswetter and Schneeweiss 2019; Westergaard et al. 2019). However, this approach relies upon ad hoc interpretations (hence, the term “interpretative phylogeography”) and lacks objective means to assess whether a preferred hypothesis actually does fit better than alternative hypotheses. Here, we also include nested clade phylogeographic analysis (NCPA) in interpretative phylogeography. NCPA attempts to find an association between haplotypes and geographic location (Templeton et al. 1995), but even if an association is found, no explanation of what created the association (e.g. historical process) is provided (Knowles and Maddison 2002). Irrespective of that, NCPA was only rarely used in alpine phylogeography (Stehlik 2002; Bettin et al. 2007) and the method has become obsolete because of prohibitively high levels of false positives (Panchal and Beaumont 2010).

Statistical phylogeography

From non-spatial to spatially explicit

To ascertain that a specific phylogeographic history explains observed genetic patterns it is necessary to show that the patterns can be replicated in a model that simulates the demographic conditions (population size changes, migration, etc.) of the specified phylogeographic scenario, but at the same time cannot be replicated in models of plausible alternative scenarios. The most popular way to do this is by simulating genealogies under the coalescent theory (Kingman 1982). This mathematical model describes how sampled alleles may have originated from the most recent common ancestor (MRCA) by merging (“coalescing”) them randomly backwards in time (Rosenberg and Nordborg 2002; for a detailed overview on coalescent theory see Rosenberg and Nordborg 2002 and Fu and Li 1999). Plausible genealogies can thus be simulated under models with conditions as expected under hypothesized phylogeographic histories and then be compared to observed genealogies within a statistical framework. Statistical frameworks are usually based on maximum likelihood or Bayesian Inference; common methods to evaluate models within these frameworks are Akaike information criterion (AIC; Akaike 1974) and Bayes factors (Jeffreys 1961), respectively (for more in depth information on the statistical frameworks see Marjoram and Tavaré 2006). Popular programs that apply these statistical frameworks include *FASTSIMCOAL2* (Excoffier et al. 2013), *DIYABC* (Cornuet et al. 2014) or *IMA3* (Hey et al. 2018). By testing models within a statistical framework it is possible to

objectively determine the best supported model, which is the one considered most likely to have occurred. This approach was coined by Knowles and Maddison (2002) as “statistical phylogeography” and will be referred to as such henceforth. Examples of its application in alpine phylogeography include Theodoridis et al. (2017), Fu et al. (2018), Wang et al. (2019) and Ikeda et al. (2020). For further reviews on statistical phylogeography and how it is applied see Knowles (2004) and Knowles (2009).

For over a decade, it has been repeatedly argued that integrating spatial data in statistical phylogeography could greatly improve phylogeographic inferences (Knowles 2009; Hickerson et al. 2010; Chan et al. 2011; Alvarado-Serrano and Knowles 2014). One of the most popular types of spatially explicit data to be used are ecological niche models (ENMs), nearly exclusively based on climate data; for an overview on how ENMs can be integrated see Richards et al. (2007) and Alvarado-Serrano and Knowles (2014). Because ENMs, both for the present and the past, can provide a priori information on present and historical distributions, they can be used to inform the design of demographic models (e.g., number and relationships of demes), including the choice of priors for parameters (e.g., population size, divergence time), without making ad hoc assumptions (Carstens and Richards 2007; Knowles et al. 2007; Dépraz et al. 2008; Forster et al. 2013; Theodoridis et al. 2017). As this approach is not directly integrating the spatial information into the models, it is here referred to as “spatially-informed”.

As many phylogeographic patterns are driven by complex interactions that vary across the landscape, such as interactions with other species (Ortego and Knowles 2020), distributional shifts due to climate change (Knowles and Massatti 2017) or anthropogenic activity (González-Serna et al. 2019), simulating phylogeographic histories explicitly on a two dimensional landscape adds an additional layer of realism. Hereinafter, we will refer to those models as “spatially explicit”. Such spatially explicit simulations are often run on a “flat” landscape (Espíndola et al. 2012; Dellicour et al. 2014a, 2017); whereas the extent of the landscape is informed by environmental data such as ENMs, local differences in carrying capacity or factors affecting movement are not considered. There are two main types of spatially explicit demographic simulators: forward-time simulators and reverse-time simulators (Hoban et al. 2012). Specifically, forward-time simulators such as *SIMADAPT* (Rebaudo et al. 2013) or *NEMO* (Guillaume and Rougemont 2006) simulate the life history of individuals within populations, including birth, selection, mating system, mutation, migration and death, as time progresses forward. The high level of detail in these types of simulators enables accurate predictions of impact of ecological changes (e.g. climate change and human impact) on present populations, making them important tools in fields such as conservation biology and

ecology (Yang et al. 2007; Bruford et al. 2010). The downside of forward-time simulators is that simulation on an individual level is relatively slow, especially when simulating over long time periods and with large population sizes. Furthermore, simulating from the past to the present requires hard to obtain knowledge about the past genetic composition of the initial individuals. In contrast, reverse-time simulators such as *IBMSIM* (Leblois et al. 2009) or *PHYLOGEOSIM 1.0* (Dellicour et al. 2014b) simulate from present time backwards using the previously mentioned coalescent theory, thus only requiring knowledge of present initial conditions. Moreover, because they simulate genealogies rather than individuals, they can simulate large population sizes over long periods of time faster than forward-time simulators. The downside of reverse-time simulators is that they usually cannot simulate complex life histories, making them less suitable for simulations where significant deviations from the Wright–Fisher model are possible (for an example, see Thalmann et al. 2007). However, considering the speed of simulating over long time periods and the simplicity of only requiring knowledge of present initial conditions, reverse-time simulators appear to be more suitable for phylogeographic inferences. Most reverse-time simulators are only able to integrate spatial data to the extent of defining which cells are inhabitable. To our knowledge and at the time of writing, *SPLATCHE* (Currat et al. 2004) and its sequels, *SPLATCHE 2* (Ray et al. 2010) and *SPLATCHE 3* (Currat et al. 2019), are the only spatially explicit coalescent simulators that are able to integrate spatial data into models as a heterogeneous landscape.

The iDDC approach

Although spatially explicit demographic simulations have been applied to phylogeographic questions for some time (Currat and Excoffier 2004, 2005; Ray et al. 2005; François et al. 2008), a full integration of distributional data into these demographic models was first introduced by Knowles and Alvarado-Serrano (2010) and Brown and Knowles (2012). Their approach was later termed the integrative Distributional, Demographic and Coalescent (iDDC) method (He et al. 2013), whose major steps will be shortly outlined in the following.

In the first step (corresponding to the “Distributional” part of iDDC), distribution data (step 1 in Fig. 1) and environmental layers (step 2 in Fig. 1) are used to construct ENMs (step 3 in Figs. 1 and 2A, B) for the relevant time periods (e.g., the last glacial maximum (LGM, 22 ka) and the present) using ecological niche modeling tools, such as *MAXENT* (Phillips et al. 2006) or *BIOMOD* (Thuiller et al. 2009; concerning best practices see Warren and Seifert 2011 and Elith et al. 2011 for *MAXENT* and Hao et al. 2019 for *BIOMOD*). The construction of ENMs assumes no changes

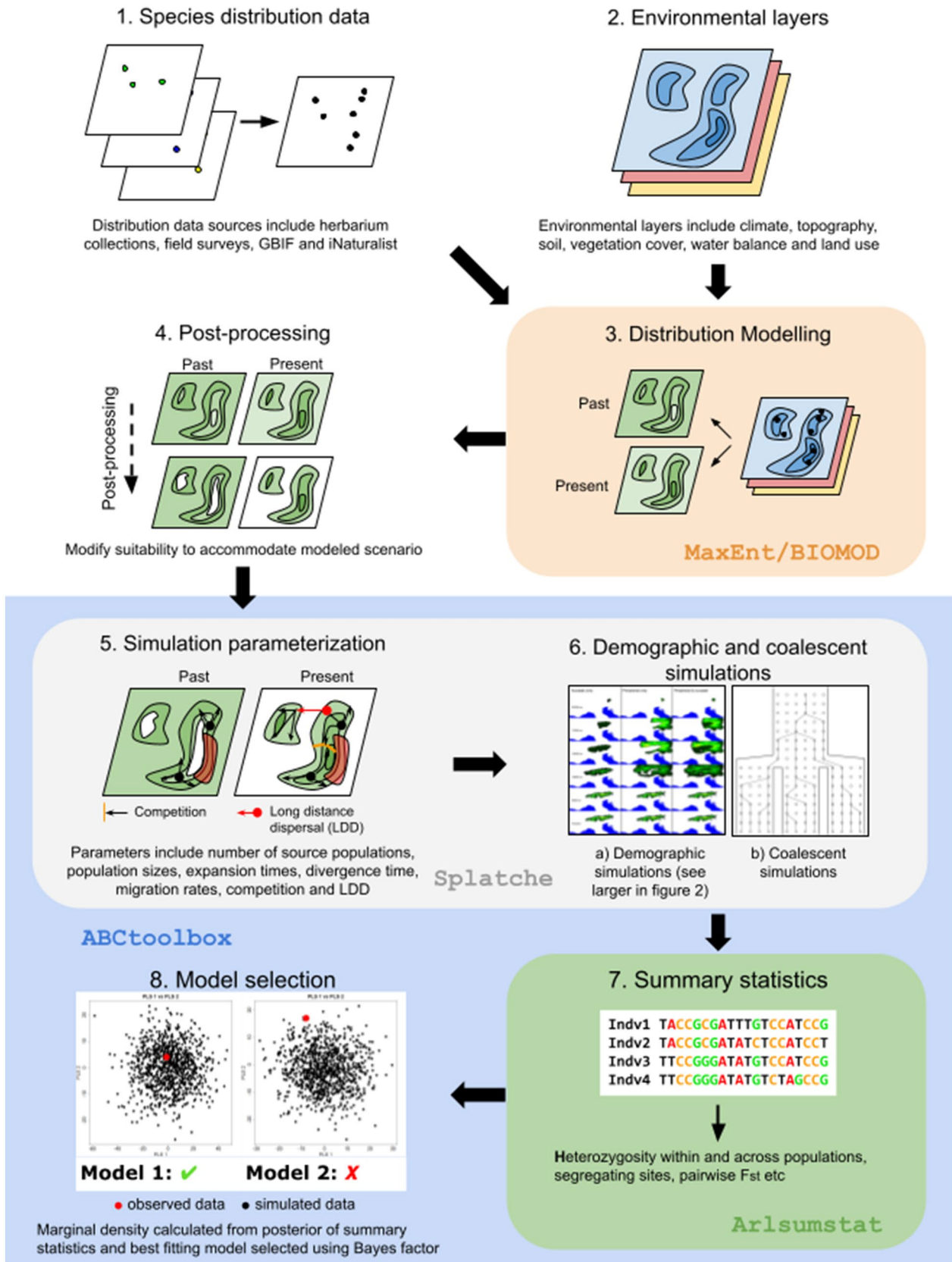


Fig. 1 The iDDC workflow. Colored areas highlight central steps in the iDDC workflow and what programs are frequently used in that step. Below several steps are short descriptions of resources, parameters, etc. that can be used or modified. (1) Species distribution data and (2) relevant environmental data are collected and used to (3) conduct ecological niche modeling using programs such as *MAXENT* or *BIOMOD*; (4) the resulting models are modified to accommodate hypothesized scenarios (see Fig. 2). After defining friction layers (these are optional) and (5) setting up simulation parameters in *SPLATCHE*, (6) the demographic and coalescent simulations are run multiple times for each model and (7) summary statistics are estimated for each of these simulations; finally, (8) the best fitting model is selected

in ecological niche since the past time period. This is a safe assumption when operating within tens of thousands of years, but may no longer hold when modeling longer time periods (Peterson 2011). The environmental layers usually include bioclimatic data from sources such as WorldClim (Hijmans et al. 2005; Fick and Hijmans 2017), but other factors that might restrict the range of the species, such as topography (Bemmels et al. 2016), can be included. Potential challenges for constructing accurate ENMs include local adaptation counteracting niche conservatism (Smith et al. 2019) and correlation among environmental variables (Braunisch et al. 2013). For more information on ENMs in phylogeographic studies the reader is referred to the excellent review of Alvarado-Serrano and Knowles (2014).

Once the ENMs for the different time periods have been produced, they can be modified to accommodate specific hypotheses (step 4 in Fig. 1). This is, for instance, done by making areas uninhabitable (setting the habitat suitability in that area to zero) as seen in Fig. 2C–F, thus acting as barriers to dispersal during that time period.

The simulated space is made up of grid cells corresponding to or derived from the cells in the habitat suitability layers. Each of these grid cells is treated as a deme (a panmictic group of individuals). Habitat suitabilities derived from the ENM define the carrying capacities of inhabitable cells, i.e., the higher the habitat suitability the higher the carrying capacity; thus, the impact of environmental factors can be taken directly into account. Upon this landscape of cells the demographic simulations are run to simulate demographic expansions and contractions in two dimensions (corresponding to the “Demographic” part of iDDC). The habitat suitability layer can be changed at a defined time (step 5 in Fig. 1; Fig. 3), thus changing the carrying capacities of demes, to accommodate changes in climate or in other factors (e.g., barriers to migration). The tool (a script originally devised by He et al. 2013) used to convert ENMs from different time periods into the input files required by *SPLATCHE*, by translating the changes in habitat suitability between different time periods into classes of change, can only handle two time periods. The demographic simulation (step 6a in Fig. 1)

is run forward in time, from a set time point in the past until present as seen in Fig. 3, to obtain a demographic history over the studied time period. In the third step (corresponding to the “Coalescent” part of iDDC), *SPLATCHE* uses this demographic history (population sizes at sampled demes, etc.) to run a coalescent simulation (step 6b in Fig. 1) backwards in time, resulting in a genealogy and thereupon simulated genetic data.

Summary statistics (step 7 in Fig. 1) over this genetic data are calculated using *ARLSTAT* (Excoffier and Lischer 2010). The statistical framework used in iDDC is Approximate Bayesian Computation (ABC; Beaumont et al. 2002), which approximates the posterior distribution of parameters using rejection-sampling (e.g. only retain a specified number of best runs or only runs that are above a tolerance threshold) of summary statistics with regression adjustment and weighting (see Marjoram and Tavaré 2006, Csilléry et al. 2010 and Beaumont 2010 for overviews on ABC). ABC is thus well suited for cases where the evaluation of the likelihood function is too costly or analytically not possible for the full data, as is the case for complex problems, such as spatially explicit phylogeographic studies. For iDDC, ABC is done in *ABCTOOLBOX* (Wegmann et al. 2010), which is used jointly with *SPLATCHE* and *ARLSTAT* (steps 5–8 in Fig. 1). Specifically, *ABCTOOLBOX* draws the parameter values for the demographic simulations in *SPLATCHE* (e.g., migration rates, ancestral population sizes) from prior distributions (instead of using fixed values as done in studies using ad hoc methods: Currat and Excoffier 2005; Ray et al. 2005) and sets the demographic simulation (and in consequence the genetic simulation) to be run multiple times (at least 10^5 times, preferably longer) to achieve sufficiently good sampling of the parameters. As the number of summary statistics can be very high, possibly causing a “curse of dimensionality” effect (i.e., a high dimensional input causes large approximation errors: Bellman 1957), partial linear squares (PLS) components of the summary statistics are used instead (Wegmann et al. 2009). The type of genetic data available determines what summary statistics can be used, however there is no specific limitation on type of genetic data. For example, out of the seven iDDC studies cited in this review that use ABC, four used SNPs from restriction-site associated (RAD) sequencing (Massatti and Knowles 2016; González-Serna et al. 2019; Ortego and Knowles 2020; Pan et al. 2020), two used full sequence data from multiple nuclear loci (He et al. 2013; Knowles and Massatti 2017) and one used 13 genotyped nuclear microsatellite markers (Bemmels et al. 2016). The data and summary statistics used is highly dependent on the study. Summary statistics frequently used in the above studies include number of segregating sites for each population and across populations, mean heterozygosity across loci for each population and across populations, and pairwise population F_{ST} .

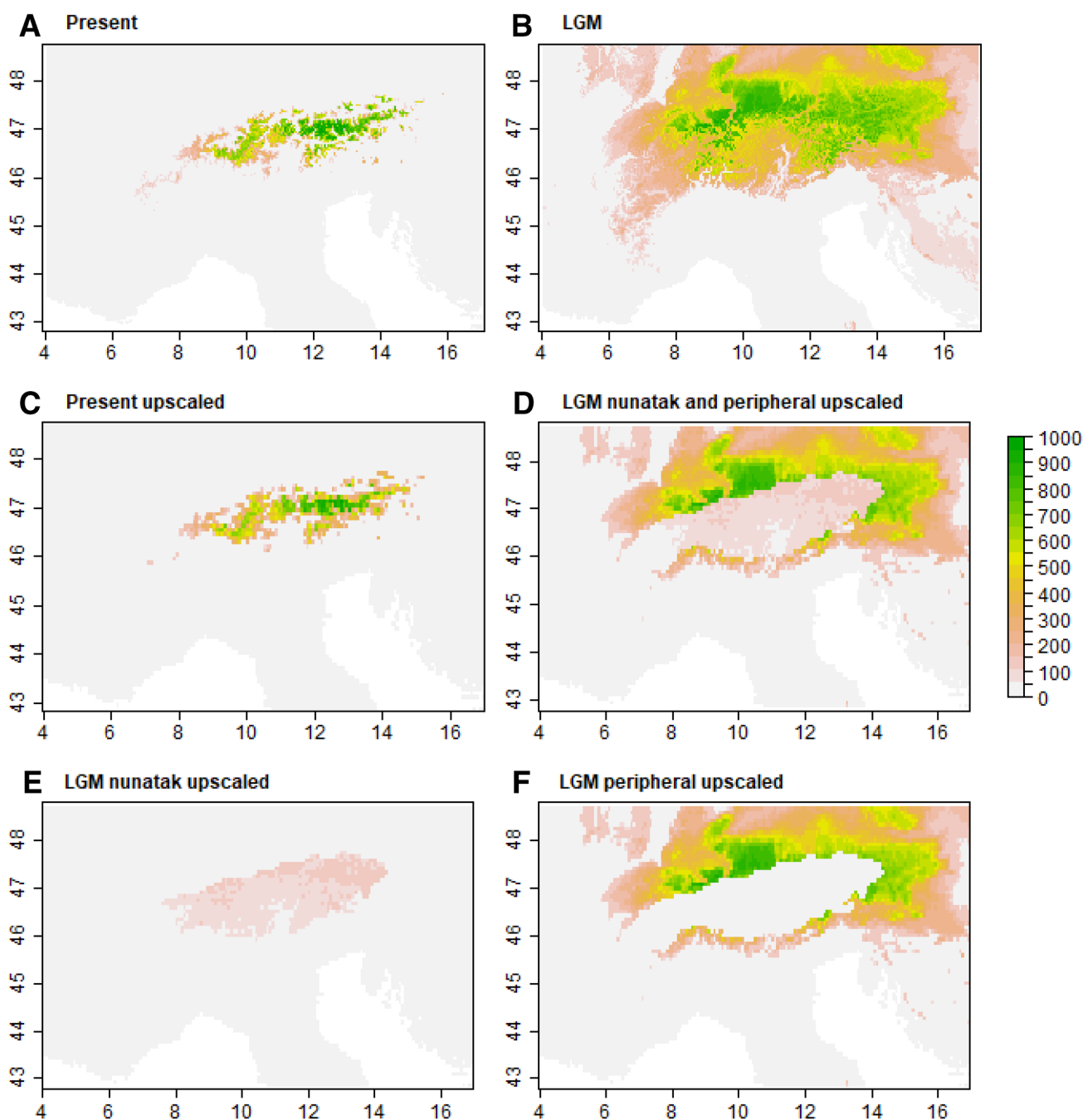


Fig. 2 Raw and modified ecological niche models (ENMs) (data taken from Pan et al 2020). Unmodified ENMs for **A** the present and **B** the Last Glacial Maximum (LGM), respectively, are generated in *BIOMOD*. These are then modified by upscaling (reducing the resolution) and by setting habitat suitabilities (ranging from 0 to 1000) below a certain threshold (10% of maximum habitat suitability) to zero. In contrast to **C** the thus modified ENM for the present, the ENMs for the LGM were further modified: **D** all grid cells are

available (i.e., both peripheral and interior refugia are permitted) with habitat suitabilities of grid cells inside the Alpine ice sheet reduced to 15% of the original habitat suitability; **E** only grid cells inside the Alpine ice sheet are available (i.e., only interior refugia are permitted) with the same reduction in habitat suitability as described before; **F** only grid cells outside the Alpine ice sheet are available (i.e., only peripheral refugia are permitted)

Finally, the marginal density is calculated from the best fitting runs (commonly 0.5% of all simulations). Model selection (step 8 in Fig. 1) is done via Bayes factors

calculated from the marginal densities of different models and is used to quantify how much better the best model is compared to the other models. Generally, a Bayes factor

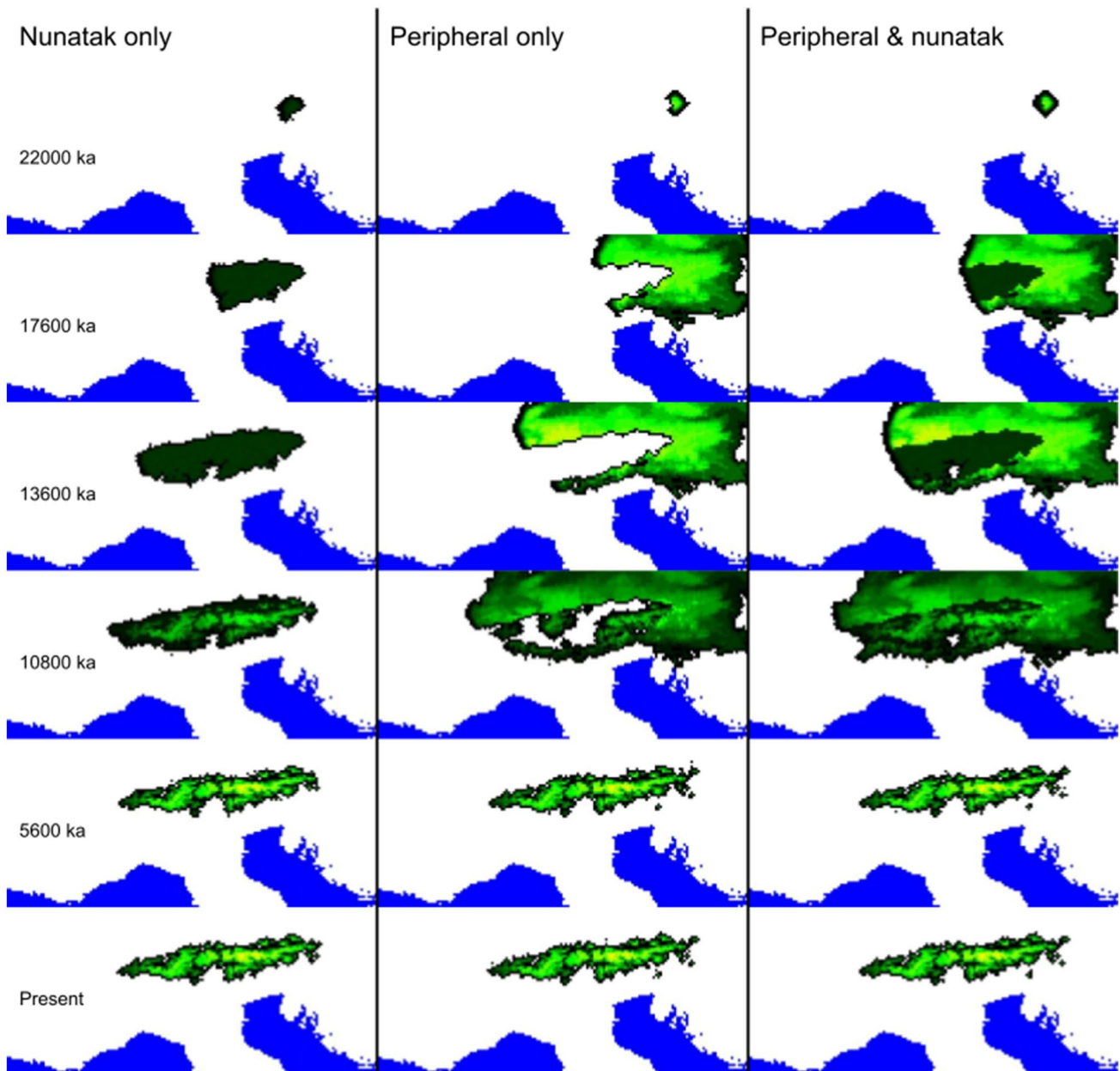


Fig. 3 The extent of expansion (with eastern refugium as starting point) at selected time intervals during the spatially explicit demographic simulations in *SPLATCHE 3* for each of the three models tested by Pan et al. (2020)

greater than 3.2 is considered significant, between 10 and 100 is considered strong and above 100 is decisive (Kass and Raftery 1995). In addition to the marginal density, the goodness-of-fit is assessed and expressed as a p value, which represents the proportion of the best fitting runs with a likelihood higher or equal to the observed data. A higher p value indicates a better fit between the simulated and the observed data. If the best model (highest marginal density) has a low p value, this model is just the best among a set of poor models, none of which can reliably explain the observed data; hence, the model set up has to be modified.

iDDC case studies

The first time iDDC was applied to alpine plants was by Massatti and Knowles (2016). The authors tested whether the microhabitat preferences of *Carex nova* and *C. chalciolepis* (Cyperaceae), two sedges from the Southern Rocky Mountains, influenced their ability to survive and migrate through glaciated mountains during the LGM. Specifically, *C. nova*, which prefers wet microhabitats (moist sub-alpine and alpine swards, streams and lake margins; Ball and Reznicek 2021), was not expected to have been able to

grow at higher altitudes due to the large accumulation of snow and ice inside drainages during the LGM. In contrast, *C. chalciolepis*, which prefers dry microhabitats (subalpine to alpine swards; Ball and Reznicek 2021), may have been able to inhabit ice-free ridges at higher altitudes, enabling this species to have been more interconnected during the glacial periods. To test these hypotheses the authors created two models, one where high altitude areas that were glaciated during the LGM were permeable (permeable model), albeit with low suitability, and one where these areas acted as barriers (barrier model). This was done by first reconstructing, for each species separately, ENMs for the LGM and for the present using *MAXENT* and then averaging the ENMs across the two species for each time period, resulting in one averaged ENM for the LGM and one averaged ENM for the present. Thus, differential outcomes of the models due to differences in the ENMs of the two species were avoided. Then the authors modified the ENM for those parts that were known to have been glaciated during the LGM to have habitat suitability reduced to 15% of the original suitability in the permeable model and to zero suitability in the barrier model. For each of the two models one million simulations were generated for each species. For *C. nova* the barrier model fitted significantly better to the observed data (Bayes factor: 22.69) than the permeable model. The reverse was true for *C. chalciolepis*, where the permeable model fitted better than the barrier model, yet only with weak support (Bayes factor: 2.84). The goodness-of-fit of the best models for both species was high (0.970 for *C. chalciolepis* and 0.844 for *C. nova*), indicating that the models performed well at reconstructing the observed data.

The second iDDC study on alpine plants was conducted on two species in the European Alps by Pan et al. (2020). The purpose of the study was to test the hypothesis of nunatak survival (i.e. survival on mountain peaks protruding above the continuous ice sheets), which has been suggested in multiple studies (Stehlik et al. 2002; Escobar García et al. 2012; Schönswetter and Schneeweiss 2019), but has rarely been tested explicitly (Westergaard et al. 2019). The two study species, *Pedicularis asplenifolia* (Orobanchaceae) and *Carex fuliginosa* (Cyperaceae), both grow in alpine swards and, especially the former, also in subnival communities. The authors hypothesized that these cold-adapted species could have survived the last LGM within the Alpine ice sheet on nunataks (alone or in addition to refugia at the periphery of the Alps). For each species, ENMs were created for the LGM and for the present using the R package *BIO-MOD* (Thuiller et al. 2009, 2020). The ENMs from the LGM were then modified to accommodate three models: survival in the periphery of the Alpine ice sheet only (peripheral survival = Peri), survival both in the periphery of and on nunataks within the ice sheet (peripheral and nunatak survival = Peri + Nun), and survival only on nunataks within the

ice sheet (nunatak survival = Nun). For Peri the suitability in ice-covered regions was set to zero, for Peri + Nun and Nun the suitability of ice-covered regions was set to 15% of the original, and for the Nun model additionally the suitability outside the ice-covered regions was set to zero. Due to computational limitations, Pan et al. (2020) only used one starting point (one refugium) in the demographic simulations. Thus, effectively six models were simulated for each species, where the Peri and Peri + Nun models were simulated two and three times, respectively, each time with a different location for the ancestral refugium (southern periphery, eastern periphery and central nunatak). Similar to Massatti and Knowles (2016), for each model one million simulations were generated. For *P. asplenifolia* the Peri_{East} + Nun model (eastern periphery plus central nunataks as refugia) fitted the observed data best, overwhelmingly better compared to the alternative models (Bayes factor > 100), and had a decent goodness-of-fit (*p* value of 0.736). However, for *C. fuliginosa* the results were less decisive. The Peri_{East} model (eastern periphery as refugium) was decisively better (Bayes factor > 100) than all other models except the Peri_{East} + Nun model, where the Bayes factor was only 5.58. For both models, the *p*-values were very high (0.993 and 0.992 for the Peri_{East} and the Peri_{East} + Nun model, respectively), indicating a good reconstruction.

The results of the two studies show that the iDDC method can be successfully used to contrast biologically informed hypotheses and provide insight into the phylogeographic history of alpine plants. Both studies suggest that alpine plants did not experience a universal response to Pleistocene climate change, but rather show species- or possibly “cohort”—(a group of species with similar ecological traits) specific responses, which are determined by adaptations to specific ecological conditions. The signal for this is particularly strong in the two North American sedges (Massatti and Knowles 2016), whose primary ecological difference concerns their microhabitat preferences. The good fit of the simulated data from the permeable model to the observed data in the dry adapted species and, likewise, for the barrier model in the wet adapted species is in line with the hypothesis that dry adapted species could have survived (also) on dry ice-free ridges at higher altitudes, whereas wet adapted species had to survive outside the ice sheet at lower altitudes. Although Pan et al. (2020) primarily investigated nunatak survival, the unambiguous support for nunatak survival in the species more tolerant to harsher conditions (as it can grow also in the subnival belt) and the unclear model choice for the species not tolerant to subnival conditions are indicative of differing response to climate change for the two species. However, as Pan et al. (2020) point out, the equivocal result in *C. fuliginosa* may also be the result of postglacial genetic swamping and the resulting genetic homogenization, which appears likely in the wind-pollinated sedge species.

Either way, the study provides strong support in favor of the nunatak survival hypothesis, especially for species that grow in the subnival belt. Although it is undoubtedly true that the best model is not necessarily the correct one, the results of the two studies clearly point to the direction that there are species-specific responses to (Pleistocene) climate change.

Outlook

The iDDC method has proven itself useful in progressing our understanding of the phylogeography of alpine plants, but there are a number of drawbacks where improvements are desirable. One of those issues concerns the rather coarse resolution used for the demographic simulations. As the number of grid cells in the demographic simulation directly impacts the run-time, it is necessary to upscale (i.e., decrease resolution) the geographic layers to reduce the number of cells, leading to a coarser resolution. This is expected to be a particular problem in the alpine and subnival zones, where habitat conditions often differ strongly even over short distances (Körner 2021), and such differences will be more easily lost with decreasing resolution. Randin et al. (2009) found that species persistence (survival under future climate conditions) in alpine plants was predicted to be higher in a high spatial resolution model (25×25 m) versus a low resolution model (16×16 km). The authors suggest that this is likely due to the higher resolution model better representing the rugged topography of mountains, which allows micro-refugia that are overlooked in the low resolution model to be captured. Similarly, Trivedi et al. (2008) showed that habitat loss in montane plants is underestimated in a low resolution model compared to a high resolution model. A cross-scale comparison revealed that the low resolution model overestimated the thermal tolerance of mountain plants because cold high altitude areas were averaged out in the coarser model.

While higher resolution layers (30 arc seconds) do exist for the present (Karger et al. 2017) and the LGM (Karger et al. 2021) from CHELSA, their usefulness may be limited for several reasons. Firstly, to our knowledge only González-Serna et al. (2019) have been able to achieve a computationally tractable resolution for the ENM smaller than 5 arc minutes (González-Serna et al. 2019 used 2.5 arc minutes) in a study applying iDDC. Secondly, although the evaluation of LGM and mid-Holocene (6 ka) climate models against paleodata (ice-core, marine and terrestrial archives) shows that paleoclimate models are able to reproduce large-scale changes, these climate models tend to underestimate regional variability (Braconnot et al. 2012). For this reason, estimates of climate at a local and regional level are likely to be inaccurate. This is further corroborated in the benchmark of paleoclimate models by Harrison et al. (2014), who found that models tend to agree on the direction of change, but

differ in the amplitude of change. Thirdly, the microhabitats that many alpine plants are dependent on are usually shaped by factors that vary on a very small scale (meter-by-meter, Körner 2021), a level of resolution impossible to be modeled for past climates. Downscaling (i.e., increasing resolution) of paleoclimate models is a notoriously complex process (Lima-Ribeiro et al. 2015; Beyer et al. 2020; review by Harris et al. 2014), where minor biases on a global scale can create major biases on a regional scale (see the jet stream example of Hall 2014). While Hall (2014) mention a number of cases where downscaling can result in more accurate estimates, there is no guarantee that models on a scale relevant for alpine plant microhabitats will be accurate. Although the difficulties in obtaining paleoclimate data at a resolution relevant for alpine species may intuitively appear detrimental, the situation is not that bleak. For many alpine plants the most important phylogeographic patterns tend to occur over relatively large geographic distances, for example across mountain ranges (Schönswetter et al. 2005; Alvarez et al. 2009). Capturing microhabitats at a local scale certainly is needed to accurately estimate local population sizes during past climates, but for most species, except the geographically most strongly restricted ones, accurately recreating the overall habitat connectivity will be far more important. Therefore, for alpine plants distributed across mountain ranges and mountain systems the currently available relatively coarse resolution is sufficiently informative.

Connected to the issue of spatial resolution is temporal resolution (but note the current restrictions concerning the number of different time periods described in section *The iDDC approach*). Being able to cover time periods other than the LGM would likely improve the simulation of the phylogeographic history of alpine species, especially since time periods such as “the last glaciation” were climatically rather heterogeneous (Andersen et al. 2004; Jouzel et al. 2007). Both Paleoclim (Brown et al. 2018) and Worldclim offer bioclimatic variables for the last interglacial (LIG, (130–115 ka), and data from Paleoclim extend as far back as the Pliocene (ca. 3.3 Ma). Whereas Worldclim does have bioclimatic variables for the mid-Holocene, Paleoclim and CHELSA cover more time points between the LGM and the present; CHELSA even has model data for every 100 years from 21 to 1 ka. Beyer et al. (2020) similarly offer bioclimatic variables for more time periods, however with greater focus on the LIG, the LGM and the Holocene, by providing bioclimatic variables for every 2000 years between the LIG (120 ka) and the LGM and every 1000 years between the LGM and present. However, this dataset is only available at 30 arc minutes resolution and would certainly need to be downscaled for use with alpine plants (i.e., in topographically complex regions like mountains). Even if paleoclimate data would be available in optimal temporal resolution going far beyond the LGM, it has to be kept in mind that when

modeling ecological niches hundreds of thousands of years into the past the assumption of niche constancy is more likely to be violated.

Another issue concerns the type of data used for ecological niche modeling. In both Pan et al. (2020) and Massatti and Knowles (2016) only climatic data was used, but other types of environmental data are expected to be relevant as well and can be implemented. For example, snowpack was included as an environmental variable in Forester et al. (2013) for the alpine plant *Rhodiola integrifolia* (Crassulaceae) and subsoil pH was included in de Melo et al. (2016) for the dry tropical forest tree *Tabebuia rosealba* (Bignoniaceae). Piedallu et al. (2013) showed that using soil water balance (calculated using soil and precipitation data) instead of precipitation or climate related proxies for water balance significantly improved the distribution model of the majority of 37 tree species, although somewhat unevenly; specifically, 71–100% of species, ranging from hygrophilous to xerophilous ones, had improved distribution models (the strongest in hygrophilous species), whereas for mesophilous species only 25% of species had improved distribution models. Data for environmental parameters beyond climate data are available, for example soil data on a global scale in the Harmonized World Soil Database v 1.2 (Fischer et al. 2008) or on a European scale in the European Soil Database v 2 (ESDB 2004; Panagos 2006). Their relevance for alpine habitats may, however, be limited due to insufficient spatial resolution. Despite the intuitive relevance of including more environmental data, improving ecological niche models beyond bioclimatic variables may not improve iDDC models. In an iDDC study by Bemmels et al. (2016) on canyon live oak (*Quercus chrysolepis*, Fagaceae), two simple models including either only climate variables or drought/growth related variables had significantly higher marginal densities than all other models, including those that contained additional variables (e.g. trade offs between growth rate and cold tolerance, topography). Whereas this suggests that including more variables does not necessarily make the model better, more studies are necessary to assess to what extent the performance of iDDC models are affected by the quality of ENMs based on different environmental variables.

Although the most common approach to construct competing models in iDDC is to modify the ENMs (step 4 in Fig. 1), there are other functionalities in *SPLATCHE* that may be used for hypothesis testing. For example, a friction layer can be used to modify the rate of migration into demes, allowing testing of hypotheses that only concern rate of migration, but not habitat suitability. For example, in areas with directional winds, migrating upwind might be difficult for wind dispersed species, even though the habitats are suitable. In this case, the hypothesis could be tested by defining a friction layer to make demes upwind of the source population more difficult to migrate into. Friction layers can also

be used to facilitate migrations, allowing spatially distributed factors to be tested that may act as dispersal corridors, such as roads and rivers (Johansson et al. 1996; Tikka et al. 2001). Another functionality available since *SPLATCHE 2* that could be used for hypothesis testing is its two-populations mode, which can be used to test interactions, including competition, between lineages (e.g., species or subspecies). An example where this might be relevant is *Phyteuma globulariifolium* (Campanulaceae). In this species, the boundary between its two major lineages (essentially corresponding to previously identified subspecies) in the European Alps is largely defined by wide inhospitable valleys, except in the northern part where there is no clear ecological or geomorphological barrier (Schönswetter et al. 2002). By testing models with or without competition, a better understanding of what maintains the genetic barrier between the two subspecies could be gained. In contrast to previous versions of the program, *SPLATCHE 3* can also simulate long distance dispersal (LDD), which could be used to test whether currently isolated populations of alpine plants are the result of LDD or of vicariance following upward range shifts. It remains to be seen, though, whether currently available data, both environmental and genetic, are of sufficient resolution to distinguish such scenarios.

A certain disadvantage of *SPLATCHE* is that functions for demographic and genetic simulations are hard-coded, potentially restricting the range of questions possible to address with iDDC. Using a set of modular libraries, such as are becoming available via *QUETZAL* developed by Becheler et al. (2019), may enable improved integration and model refinement, allowing more hypotheses with greater detail and accuracy to be tested.

Conclusion

Despite some potential drawbacks, which are mostly technical (e.g., low resolution of environmental layers), the iDDC approach has opened up new avenues of phylogeographic research by the integration of spatially explicit demographic simulations with distributional data. This is well exemplified by the few available case studies, as both Massatti and Knowles (2016) and Pan et al. (2020) have furthered our understanding of the phylogeographic history of alpine species by testing both long-standing hypotheses in alpine phylogeography (nunatak versus peripheral survival) and new hypotheses on how past pre-adaptation to different ecologies may facilitate differential responses to recent climate change. We expect that with iDDC's ability to address new phylogeographic questions using spatially explicit demographic modeling and the increasing ease of obtaining large-scale genetic data from alpine plants, approaches like the iDDC

will play a central role in alpine phylogeography in the future.

Acknowledgements We would like to thank the Editor Christian Parisdod for inviting us to write this manuscript and the three reviewers for their excellent feedback that decidedly improved its quality. The writing of this review was made possible through the funding by project P30098 (Austrian Science Fund to GMS).

Authors' contributions DJL is the corresponding author and was responsible for conceptualization, visualization, writing of original draft and review and editing; DP is a co-author and was responsible for review and editing and contributed resources from his previously published study for figure creation; GMS is a co-author and was responsible for funding acquisition, supervision, project administration and review and editing.

Funding Open access funding provided by University of Vienna.

Declarations

Conflict of interest The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Akaike H (1974) A new look at the statistical model identification. *IEEE Trans Automat Contr* 19:716–723. <https://doi.org/10.1109/TAC.1974.1100705>
- Alvarado-Serrano DF, Knowles LL (2014) Ecological niche models in phylogeographic studies: applications, advances and precautions. *Mol Ecol Resour* 14:233–248. <https://doi.org/10.1111/1755-0998.12184>
- Alvarez N, Thiel-Egenter C, Tribsch A et al (2009) History or ecology? Substrate type as a major driver of patial genetic structure in Alpine plants. *Ecol Lett* 12:632–640. <https://doi.org/10.1111/j.1461-0248.2009.01312.x>
- Andersen KK, Azuma N, Barnola JM et al (2004) High-resolution record of Northern Hemisphere climate extending into the last interglacial period. *Nature* 431:147–151. <https://doi.org/10.1038/nature02805>
- Avise JC (1987) Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annu Rev Ecol Syst* 18:489–522. <https://doi.org/10.1146/annurev.ecolsys.18.1.489>
- Avise JC (2000) *Phylogeography: the history and formation of species*. Harvard University Press, Cambridge
- Avise JC (2009) *Phylogeography: retrospect and prospect*. *J Biogeogr* 36:3–15. <https://doi.org/10.1111/j.1365-2699.2008.02032.x>
- Ball PW, Reznicek AA (2021) *Carex*. In: *Flora of North America Editorial Committee* (eds. 1993+) *Flora of North America North of Mexico* [Online]. 21+ vols. New York and Oxford. Vol. 23. <http://beta.floranorthamerica.org/Carex>. Accessed 3 Mar 2021.
- Beaumont MA (2010) Approximate Bayesian computation in evolution and ecology. *Annu Rev Ecol Syst* 41:379–406. <https://doi.org/10.1146/annurev-ecolsys-102209-144621>
- Beaumont MA, Zhang W, Balding DJ (2002) Approximate Bayesian computation in population genetics. *Genetics* 162:2025–2035. <https://doi.org/10.1093/genetics/162.4.2025>
- Becheler A, Coron C, Dupas S (2019) The Quetzal coalescence template library: a C++ programmers resource for integrating distributional, demographic and coalescent models. *Mol Ecol Resour* 19:788–793. <https://doi.org/10.1111/1755-0998.12992>
- Bellman RE (1957) *Dynamic programming*. Princeton University Press, Princeton
- Bemmels JB, Title PO, Ortego J, Knowles LL (2016) Tests of species-specific models reveal the importance of drought in postglacial range shifts of a Mediterranean-climate tree: insights from integrative distributional, demographic and coalescent modelling and ABC model selection. *Mol Ecol* 25:4889–4906. <https://doi.org/10.1111/mec.13804>
- Bettin O, Cornejo C, Edwards PJ, Holderegger R (2007) Phylogeography of the high alpine plant *Senecio halleri* (Asteraceae) in the European Alps: in situ glacial survival with postglacial stepwise dispersal into peripheral areas. *Mol Ecol* 16:2517–2524. <https://doi.org/10.1111/j.1365-294X.2007.03273.x>
- Beyer RM, Krapp M, Manica A (2020) High-resolution terrestrial climate, bioclimate and vegetation for the last 120,000 years. *Sci Data* 7:1–9. <https://doi.org/10.1038/s41597-020-0552-1>
- Braconnot P, Harrison SP, Kageyama M et al (2012) Evaluation of climate models using palaeoclimatic data. *Nat Clim Chang* 2:417–424. <https://doi.org/10.1038/NCLIMATE1456>
- Braunisch V, Coppes J, Arlettaz R et al (2013) Selecting from correlated climate variables: a major source of uncertainty for predicting species distributions under climate change. *Ecography* (cop) 36:971–983. <https://doi.org/10.1111/j.1600-0587.2013.00138.x>
- Brown JL, Knowles LL (2012) Spatially explicit models of dynamic histories: Examination of the genetic consequences of Pleistocene glaciation and recent climate change on the American Pika. *Mol Ecol* 21:3757–3775. <https://doi.org/10.1111/j.1365-294X.2012.05640.x>
- Brown JL, Hill DJ, Dolan AM et al (2018) Paleoclim, high spatial resolution paleoclimate surfaces for global land areas. *Sci Data* 5:1–9. <https://doi.org/10.1038/sdata.2018.254>
- Bruford MW, Ancrenaz M, Chikhi L et al (2010) Projecting genetic diversity and population viability for the fragmented orang-utan population in the Kinabatangan floodplain, Sabah, Malaysia. *Endanger Species Res* 12:249–261. <https://doi.org/10.3354/esr00295>
- Carstens BC, Richards CL (2007) Integrating coalescent and ecological niche modeling in comparative phylogeography. *Evolution* 61:1439–1454. <https://doi.org/10.1111/j.1558-5646.2007.00117.x>
- Chan LM, Brown JL, Yoder AD (2011) Integrating statistical genetic and geospatial methods brings new power to phylogeography. *Mol Phylogenet Evol* 59:523–537. <https://doi.org/10.1016/j.ympev.2011.01.020>
- Chen S, Wu G, Zhang D et al (2008) Potential refugium on the Qinghai-Tibet Plateau revealed by the chloroplast DNA phylogeography of the alpine species *Metagentiana striata* (Gentianaceae). *Bot J Linn Soc* 157:125–140. <https://doi.org/10.1111/j.1095-8339.2008.00785.x>

- Cornuet JM, Pudlo P, Veyssier J et al (2014) DIYABC v2.0: a software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. *Bioinformatics* 30:1187–1189. <https://doi.org/10.1093/bioinformatics/bt763>
- Csilléry K, Blum MGB, Gaggiotti OE, François O (2010) Approximate Bayesian Computation (ABC) in practice. *Trends Ecol Evol* 25:410–418. <https://doi.org/10.1016/j.tree.2010.04.001>
- Currat M, Excoffier L (2004) Modern humans did not admix with Neanderthals during their range expansion into Europe. *PLoS Biol* 2:e421. <https://doi.org/10.1371/journal.pbio.0020421>
- Currat M, Excoffier L (2005) The effect of the Neolithic expansion on European molecular diversity. *Proc R Soc B Biol Sci* 272:679–688. <https://doi.org/10.1098/rspb.2004.2999>
- Currat M, Ray N, Excoffier L (2004) SPLATCHE: a program to simulate genetic diversity taking into account environmental heterogeneity. *Mol Ecol Notes* 4:139–142. <https://doi.org/10.1046/j.1471-8286.2003.00582.x>
- Currat M, Arenas M, Quilodrán CS et al (2019) SPLATCHE3: simulation of serial genetic data under spatially explicit evolutionary scenarios including long-distance dispersal. *Bioinformatics* 35:4480–4483. <https://doi.org/10.1093/bioinformatics/bt311>
- DeChaine EG, Martin AP (2005) Marked genetic divergence among sky island populations of *Sedum lanceolatum* (Crassulaceae) in the Rocky Mountains. *Am J Bot* 92:477–486. <https://doi.org/10.3732/ajb.92.3.477>
- de Melo WA, Lima-Ribeiro MS, Terribile LC, Collevatti RG (2016) Coalescent simulation and paleodistribution modeling for *Tabebuia rosealba* do not support south American dry forest refugia hypothesis. *PLoS ONE* 11:e0159314. <https://doi.org/10.1371/journal.pone.0159314>
- Dellicour S, Fearnley S, Lombal A et al (2014a) Inferring the past and present connectivity across the range of a North American leaf beetle: combining ecological niche modeling and a geographically explicit model of coalescence. *Evolution* 68:2371–2385. <https://doi.org/10.1111/evo.12426>
- Dellicour S, Kastally C, Hardy OJ, Mardulyn P (2014b) Comparing phylogeographic hypotheses by simulating DNA sequences under a spatially explicit model of coalescence. *Mol Biol Evol* 31:3359–3372. <https://doi.org/10.1093/molbev/msu27>
- Dellicour S, Kastally C, Varela S et al (2017) Ecological niche modelling and coalescent simulations to explore the recent geographical range history of five widespread bumblebee species in Europe. *J Biogeogr* 44:39–50. <https://doi.org/10.1111/jbi.12748>
- Dépraz A, Cordellier M, Hausser J, Pfenninger M (2008) Postglacial recolonization at a snail's pace (*Trochulus villosus*): confronting competing refugia hypotheses using model selection. *Mol Ecol* 17:2449–2462. <https://doi.org/10.1111/j.1365-294X.2008.03760.x>
- Dixon CJ, Schönswetter P, Vargas P et al (2009) Bayesian hypothesis testing supports long-distance Pleistocene migrations in a European high mountain plant (*Androsace vitaliana*, Primulaceae). *Mol Phylogenet Evol* 53:580–591. <https://doi.org/10.1016/j.ympev.2009.07.016>
- Elith J, Phillips SJ, Hastie T et al (2011) A statistical explanation of MaxEnt for ecologists. *Divers Distrib* 17:43–57. <https://doi.org/10.1111/j.1472-4642.2010.00725.x>
- Escobar García P, Winkler M, Flatscher R et al (2012) Extensive range persistence in peripheral and interior refugia characterizes Pleistocene range dynamics in a widespread Alpine plant species (*Senecio carniolicus*, Asteraceae). *Mol Ecol* 21:1255–1270. <https://doi.org/10.1111/j.1365-294X.2012.05456.x>
- ESDB (2004) The European soil database distribution version 2.0. In: European Commission and the European Soil Bureau Network. <https://esdac.jrc.ec.europa.eu/content/european-soil-database-v20-vector-and-attribute-data>. Accessed 29 Jan 2021
- Espindola A, Pellissier L, Maiorano L et al (2012) Predicting present and future intra-specific genetic structure through niche hindcasting across 24 millennia. *Ecol Lett* 15:649–657. <https://doi.org/10.1111/j.1461-0248.2012.01779.x>
- Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Resour* 10:564–567. <https://doi.org/10.1111/j.1755-0998.2010.02847.x>
- Excoffier L, Dupanloup I, Huerta-Sánchez E et al (2013) Robust demographic inference from genomic and SNP data. *PLoS Genet* 9:e1003905. <https://doi.org/10.1371/journal.pgen.1003905>
- Fick SE, Hijmans RJ (2017) WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *Int J Climatol* 37:4302–4315. <https://doi.org/10.1002/joc.5086>
- Fischer G, Nachtergaele F, Prieler S, van Velthuizen HT, Verelst L, Wiberg D (2008) Global Agro-ecological Zones Assessment for Agriculture (GAEZ 2008). IIASA, Laxenburg
- Forester BR, Dechaine EG, Bunn AG (2013) Integrating ensemble species distribution modelling and statistical phylogeography to inform projections of climate change impacts on species distributions. *Divers Distrib* 19:1480–1495. <https://doi.org/10.1111/ddi.12098>
- François O, Blum MGB, Jakobsson M, Rosenberg NA (2008) Demographic history of European populations of *Arabidopsis thaliana*. *PLoS Genet* 4:e1000075. <https://doi.org/10.1371/journal.pgen.1000075>
- Freeman BG, Lee-Yaw JA, Sunday JM, Hargreaves AL (2018) Expanding, shifting and shrinking: The impact of global warming on species' elevational distributions. *Glob Ecol Biogeogr* 27:1268–1276. <https://doi.org/10.1111/geb.12774>
- Fu YX, Li WH (1999) Coalescing into the 21st century: an overview and prospects of coalescent theory. *Theor Popul Biol* 56:1–10. <https://doi.org/10.1006/tpbi.1999.1421>
- Fu P-C, Ya H-Y, Liu Q-W et al (2018) Out of Refugia: Population Genetic Structure and Evolutionary History of the Alpine Medicinal Plant *Gentiana lawrencei* var. *farreri* (Gentianaceae). *Front Genet* 9:564. <https://doi.org/10.3389/fgene.2018.00564>
- Gizaw A, Kebede M, Nemomissa S et al (2013) Phylogeography of the heathers *Erica arborea* and *E. trimeria* in the afro-alpine “sky islands” inferred from AFLPs and plastid DNA sequences. *Flora Morphol Distrib Funct Ecol Plants* 208:453–463. <https://doi.org/10.1016/j.flora.2013.07.007>
- González-Serna MJ, Cordero PJ, Ortego J (2019) Spatiotemporally explicit demographic modelling supports a joint effect of historical barriers to dispersal and contemporary landscape composition on structuring genomic variation in a red-listed grasshopper. *Mol Ecol* 28:2155–2172. <https://doi.org/10.1111/mec.15086>
- Guillaume F, Rougemont J (2006) Nemo: an evolutionary and population genetics programming framework. *Bioinformatics* 22:2556–2557. <https://doi.org/10.1093/bioinformatics/btl415>
- Hall A (2014) Projecting regional change. *Science* 346:1461–1462. <https://doi.org/10.1126/science.aaa0629>
- Hao T, Elith J, Guillera-Aroita G, Lahoz-Monfort JJ (2019) A review of evidence about use and performance of species distribution modelling ensembles like BIOMOD. *Divers Distrib* 25:839–852. <https://doi.org/10.1111/ddi.12892>
- Harris RMB, Grose MR, Lee G et al (2014) Climate projections for ecologists. *Wiley Interdiscip Rev Clim Chang* 5:621–637. <https://doi.org/10.1002/wcc.291>
- Harrison SP, Bartlein PJ, Brewer S et al (2014) Climate model benchmarking with glacial and mid-Holocene climates. *Clim Dyn* 43:671–688. <https://doi.org/10.1007/s00382-013-1922-6>
- He Q, Edwards DL, Knowles LL (2013) Integrative testing of how environments from the past to the present shape genetic structure across landscapes. *Evolution* 67:3386–3402. <https://doi.org/10.1111/evo.12159>

- Hey J, Chung Y, Sethuraman A et al (2018) Phylogeny estimation by integration over isolation with migration models. *Mol Biol Evol* 35:2805–2818. <https://doi.org/10.1093/molbev/msy162>
- Hickerson MJ, Carstens BC, Cavender-Bares J et al (2010) Phylogeography's past, present, and future: 10 years after *Avisé*, 2000. *Mol Phylogenet Evol* 54:291–301. <https://doi.org/10.1016/j.ympev.2009.09.016>
- Hijmans RJ, Cameron SE, Parra JL et al (2005) Very high resolution interpolated climate surfaces for global land areas. *Int J Climatol* 25:1965–1978. <https://doi.org/10.1002/joc.1276>
- Hoban S, Bertorelle G, Gaggiotti OE (2012) Computer simulations: Tools for population and evolutionary genetics. *Nat Rev Genet* 13:110–122. <https://doi.org/10.1038/nrg3130>
- Hülber K, Kuttner M, Moser D et al (2020) Habitat availability disproportionately amplifies climate change risks for lowland compared to alpine species. *Glob Ecol Conserv* 23:e01113. <https://doi.org/10.1016/j.gecco.2020.e01113>
- Ikeda H, Setoguchi H (2007) Phylogeography and refugia of the Japanese endemic alpine plant, *Phyllodoce nipponica* Makino (Ericaceae). *J Biogeogr* 34:169–176. <https://doi.org/10.1111/j.1365-2699.2006.01577.x>
- Ikeda H, Fujii N, Setoguchi H (2009) Application of the isolation with migration model demonstrates the Pleistocene origin of geographic differentiation in *Cardamine nipponica* (brassicaceae), an endemic Japanese alpine plant. *Mol Biol Evol* 26:2207–2216. <https://doi.org/10.1093/molbev/msp128>
- Ikeda H, Yakubov V, Barkalov V et al (2020) East Asian origin of the widespread alpine snow-bed herb, *Primula cuneifolia* (Primulaceae), in the northern Pacific region. *J Biogeogr* 47:2181–2193. <https://doi.org/10.1111/jbi.13918>
- Inouye DW (2020) Effects of climate change on alpine plants and their pollinators. *Ann New York Acad Sci* 1469:26–37
- Jeffreys H (1961) *Theory of probability*. Clarendon Press, Oxford
- Johansson ME, Nilsson C, Nilsson E (1996) Do rivers function as corridors for plant dispersal? *J Veg Sci* 7:593–598. <https://doi.org/10.2307/3236309>
- Jouzel J, Masson-Delmotte V, Cattani O et al (2007) Orbital and millennial antarctic climate variability over the past 800,000 years. *Science* (80-) 317:793–796. <https://doi.org/10.1126/science.1141038>
- Kadereit JW, Griebeler EM, Comes HP (2004) Quaternary diversification in European alpine plants: pattern and process. *Philos Trans R Soc B Biol Sci* 359:265–274. <https://doi.org/10.1098/rstb.2003.1389>
- Karger DN, Conrad O, Böhrner J et al (2017) Climatologies at high resolution for the earth's land surface areas. *Sci Data* 4:1–20. <https://doi.org/10.1038/sdata.2017.122>
- Karger DN, Nobis MP, Normand S, Graham CH, Zimmermann NE (2021) CHELSA-TraCE21k v1.0. Downscaled transient temperature and precipitation data since the last glacial maximum. *Clim past Discuss*. <https://doi.org/10.5194/cp-2021-30> (preprint)
- Kass RE, Raftery AE (1995) Bayes factors. *J Am Stat Assoc* 90:773–795. <https://doi.org/10.1080/01621459.1995.10476572>
- Kingman JFC (1982) The Coalescent. *Stoch Process Their Appl* 13:235–248. [https://doi.org/10.1016/0304-4149\(82\)90011-4](https://doi.org/10.1016/0304-4149(82)90011-4)
- Knowles LL (2004) The burgeoning field of statistical phylogeography. *J Evol Biol* 17:1–10. <https://doi.org/10.1046/j.1420-9101.2003.00644.x>
- Knowles LL (2009) Statistical phylogeography. *Annu Rev Ecol Evol Syst* 40:593–612. <https://doi.org/10.1146/annurev.ecolsys.38.091206.095702>
- Knowles LL, Alvarado-Serrano DF (2010) Exploring the population genetic consequences of the colonization process with spatio-temporally explicit models: Insights from coupled ecological, demographic and genetic models in montane grasshoppers. *Mol Ecol* 19:3727–3745. <https://doi.org/10.1111/j.1365-294X.2010.04702.x>
- Knowles LL, Maddison W (2002) Statistical phylogeography. *Mol Ecol* 11:2623–2635. <https://doi.org/10.1046/j.1365-294X.2002.01410.x>
- Knowles LL, Massatti R (2017) Distributional shifts—not geographic isolation—as a probable driver of montane species divergence. *Ecography* 40:1475–1485. <https://doi.org/10.1111/ecog.02893>
- Knowles LL, Carstens BC, Keat MLL (2007) Coupling genetic and ecological-niche models to examine how past population distributions contribute to divergence. *Curr Biol* 17:940–946. <https://doi.org/10.1016/j.cub.2007.04.033>
- Körner C (2021) *Alpine Plant Life*. Springer, Cham. <https://doi.org/10.1007/978-3-030-59538-8>
- Kropf M, Kadereit JW, Comes HP (2003) Differential cycles of range contraction and expansion in European high mountain plants during the Late Quaternary: insights from *Pritzelago alpina* (L.) O. Kuntze (Brassicaceae). *Mol Ecol* 12:931–949. <https://doi.org/10.1046/j.1365-294X.2003.01781.x>
- Leblois R, Estoup A, Rousset F (2009) IBDSim: a computer program to simulate genotypic data under isolation by distance. *Mol Ecol Resour* 9:107–109. <https://doi.org/10.1111/j.1755-0998.2008.02417.x>
- Lima-Ribeiro MS, Varela S, González-Hernández J et al (2015) EcoClimate: a database of climate data from multiple models for past, present, and future for macroecologists and biogeographers. *Biodivers Inform* 10:1–21. <https://doi.org/10.17161/bi.v10i0.4955>
- Marjoram P, Tavaré S (2006) Modern computational approaches for analysing molecular genetic variation data. *Nat Rev Genet* 7:759–770. <https://doi.org/10.1038/nrg1961>
- Massatti R, Knowles LL (2014) Microhabitat differences impact phylogeographic concordance of codistributed species: genomic evidence in montane sedges (*Carex* L.) from the Rocky Mountains. *Evolution* 68:2833–2846. <https://doi.org/10.1111/evo.12491>
- Massatti R, Knowles LL (2016) Contrasting support for alternative models of genomic variation based on microhabitat preference: species-specific effects of climate change in alpine sedges. *Mol Ecol* 25:3974–3986. <https://doi.org/10.1111/mec.13735>
- Mráz P, Gaudeul M, Rioux D et al (2007) Genetic structure of *Hypochaeris uniflora* (Asteraceae) suggests vicariance in the Carpathians and rapid post-glacial colonization of the Alps from an eastern Alpine refugium. *J Biogeogr* 34:2100–2114. <https://doi.org/10.1111/j.1365-2699.2007.01765.x>
- Ortego J, Knowles LL (2020) Incorporating interspecific interactions into phylogeographic models: a case study with Californian oaks. *Mol Ecol* 29:4510–4524. <https://doi.org/10.1111/mec.15548>
- Pan D, Hülber K, Willner W, Schneeweiss GM (2020) An explicit test of Pleistocene survival in peripheral versus nunatak refugia in two high mountain plant species. *Mol Ecol* 29:172–183. <https://doi.org/10.1111/mec.15316>
- Panagos P (2006) The European Soil Database. *GEO Connex* 5:32–33
- Panchal M, Beaumont MA (2010) Evaluating nested clade phylogeographic analysis under models of restricted gene flow. *Syst Biol* 59:415–432. <https://doi.org/10.1093/sysbio/syq022>
- Peterson AT (2011) Ecological niche conservatism: a time-structured review of evidence. *J Biogeogr* 38:817–827. <https://doi.org/10.1111/j.1365-2699.2010.02456.x>
- Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. *Ecol Modell* 190:231–259. <https://doi.org/10.1016/j.ecolmodel.2005.03.026>
- Piedallu C, Gégout JC, Perez V, Lebourgeois F (2013) Soil water balance performs better than climatic water variables in tree species distribution modelling. *Glob Ecol Biogeogr* 22:470–482. <https://doi.org/10.1111/geb.12012>

- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155:945–959. <https://doi.org/10.1093/genetics/155.2.945>
- Puşçaş M, Choler P, Tribsch A et al (2008) Post-glacial history of the dominant alpine sedge *Carex curvula* in the European Alpine System inferred from nuclear and chloroplast markers. *Mol Ecol* 17:2417–2429. <https://doi.org/10.1111/j.1365-294X.2008.03751.x>
- Randin CF, Engler R, Normand S et al (2009) Climate change and plant distribution: local models predict high-elevation persistence. *Glob Change Biol* 15:1557–1569. <https://doi.org/10.1111/j.1365-2486.2008.01766.x>
- Ray N, Currat M, Berthier P, Excoffier L (2005) Recovering the geographic origin of early modern humans by realistic and spatially explicit simulations. *Genome Res* 15:1161–1167. <https://doi.org/10.1101/gr.370850>
- Ray N, Currat M, Foll M, Excoffier L (2010) SPLATCHE2: a spatially explicit simulation framework for complex demography, genetic admixture and recombination. *Bioinformatics* 26:2993–2994. <https://doi.org/10.1093/bioinformatics/btq579>
- Rebaudo F, Le Rouzic A, Dupas S et al (2013) SimAdapt: an individual-based genetic model for simulating landscape management impacts on populations. *Methods Ecol Evol* 4:595–600. <https://doi.org/10.1111/2041-210X.12041>
- Richards CL, Carstens BC, Lacey Knowles L (2007) Distribution modelling and statistical phylogeography: an integrative framework for generating and testing alternative biogeographical hypotheses. *J Biogeogr* 34:1833–1845. <https://doi.org/10.1111/j.1365-2699.2007.01814.x>
- Ronikier M, Schneeweiss GM, Schönswetter P (2012) The extreme disjunction between Beringia and Europe in *Ranunculus glacialis* s. l. (Ranunculaceae) does not coincide with the deepest genetic split—a story of the importance of temperate mountain ranges in arctic-alpine phylogeography. *Mol Ecol* 21:5561–5578. <https://doi.org/10.1111/mec.12030>
- Rosenberg NA, Nordborg M (2002) Genealogical trees, coalescent theory and the analysis of genetic polymorphisms. *Nat Rev Genet* 3:380–390. <https://doi.org/10.1038/nrg795>
- Schneeweiss GM, Schönswetter P (2010) The wide but disjunct range of the European mountain plant *Androsace lactea* L. (Primulaceae) reflects Late Pleistocene range fragmentation and post-glacial distributional stasis. *J Biogeogr* 37:2016–2025. <https://doi.org/10.1111/j.1365-2699.2010.02350.x>
- Schneeweiss GM, Winkler M, Schönswetter P (2017) Secondary contact after divergence in allopatry explains current lack of ecological isolation in two hybridizing alpine plant species. *J Biogeogr* 44:2575–2584. <https://doi.org/10.1111/jbi.13071>
- Schönswetter P, Schneeweiss GM (2019) Is the incidence of survival in interior Pleistocene refugia (nunataks) underestimated? Phylogeography of the high mountain plant *Androsace alpina* (Primulaceae) in the European Alps revisited. *Ecol Evol* 9:4078–4086. <https://doi.org/10.1002/ece3.5037>
- Schönswetter P, Tribsch A, Barfuss M, Niklfeld H (2002) Several Pleistocene refugia detected in the high alpine plant *Phyteuma globulariifolium* Sternb. & Hoppe (Campanulaceae) in the European Alps. *Mol Ecol* 11:2637–2647. <https://doi.org/10.1046/j.1365-294X.2002.01651.x>
- Schönswetter P, Tribsch A, Niklfeld H (2003) Phylogeography of the high alpine cushion plant *Androsace alpina* (Primulaceae) in the European Alps. *Plant Biol* 5:623–630. <https://doi.org/10.1055/s-2003-44686>
- Schönswetter P, Stehlik I, Holderegger R, Tribsch A (2005) Molecular evidence for glacial refugia of mountain plants in the European Alps. *Mol Ecol* 14:3547–3555. <https://doi.org/10.1111/j.1365-294X.2005.02683.x>
- Segarra-Moragues JG, Palop-Esteban M, González-Candelas F, Catalán P (2007) Nunatak survival vs. tabula rasa in the Central Pyrenees: a study on the endemic plant species *Borderea pyrenaica* (Dioscoreaceae). *J Biogeogr* 34:1893–1906. <https://doi.org/10.1111/j.1365-2699.2007.01740.x>
- Slovák M, Kučera J, Turis P, Zozomová-Lihová J (2012) Multiple glacial refugia and postglacial colonization routes inferred for a woodland geophyte, *Cyclamen purpurascens*: patterns concordant with the Pleistocene history of broadleaved and coniferous tree species. *Biol J Linn Soc* 105:741–760. <https://doi.org/10.1111/j.1095-8312.2011.01826.x>
- Smith AB, Godsoe W, Rodríguez-Sánchez F et al (2019) Niche Estimation above and below the species level. *Trends Ecol Evol* 34:260–273. <https://doi.org/10.1016/j.tree.2018.10.012>
- Stehlik I (2002) Glacial history of the alpine herb *Rumex nivalis* (Polygonaceae): a comparison of common phylogeographic methods with nested clade analysis. *Am J Bot* 89:2007–2016. <https://doi.org/10.3732/ajb.89.12.2007>
- Stehlik I, Schneller JJ, Bachmann K (2001) Resistance or emigration: Response of the high-alpine plant *Eritrichium nanum* (L.) Gaudin to the ice age within the Central Alps. *Mol Ecol* 10:357–370. <https://doi.org/10.1046/j.1365-294X.2001.01179.x>
- Stehlik I, Blattner FR, Holderegger R, Bachmann K (2002) Nunatak survival of the high Alpine plant *Eritrichium nanum* (L.) Gaudin in the central Alps during the ice ages. *Mol Ecol* 11:2027–2036. <https://doi.org/10.1046/j.1365-294X.2002.01595.x>
- Templeton AR, Routman E, Phillips CA (1995) Separating population structure from population history: a cladistic analysis of the geographical distribution of mitochondrial DNA haplotypes in the tiger salamander, *Ambystoma tigrinum*. *Genetics* 140:767–782. <https://doi.org/10.1093/genetics/140.2.767>
- Thalmann OH, Fischer AH, Lankester FH et al (2007) The complex evolutionary history of gorillas: Insights from genomic data. *Mol Biol Evol* 24:146–158. <https://doi.org/10.1093/molbev/msl160>
- Theodoridis S, Randin C, Szovenyi P et al (2017) How do cold-adapted plants respond to climatic cycles? Interglacial expansion explains current distribution and genomic diversity in *Primula farinosa* L. *Syst Biol* 66:715–736. <https://doi.org/10.1093/sysbio/syw114>
- Thuiller W, Lafourcade B, Engler R, Araújo MB (2009) BIOMOD—a platform for ensemble forecasting of species distributions. *Ecography* 32:369–373. <https://doi.org/10.1111/j.1600-0587.2008.05742.x>
- Thuiller W, Georges D, Engler R, Breiner F (2020) biomod2: Ensemble platform for species distribution modeling. R package version 3.4.6. <https://cran.r-project.org/web/packages/biomod2/index.html>. Accessed 4 Mar 2021.
- Tikka PM, Högmander H, Koski PS (2001) Road and railway verges serve as dispersal corridors for grassland plants. *Landsc Ecol* 16:659–666. <https://doi.org/10.1023/A:1013120529382>
- Trivedi MR, Berry PM, Morecroft MD, Dawson TP (2008) Spatial scale affects bioclimate model projections of climate change impacts on mountain plants. *Glob Change Biol* 14:1089–1103. <https://doi.org/10.1111/j.1365-2486.2008.01553.x>
- Vargas P (2003) Molecular evidence for multiple diversification patterns of alpine plants in Mediterranean Europe. *Taxon* 52:463–476. <https://doi.org/10.2307/3647383>
- Wang GN, He XY, Miede G, Mao KS (2014) Phylogeography of the Qinghai-Tibet Plateau endemic alpine herb *Pomatosace filicula* (Primulaceae). *J Syst Evol* 52:289–302. <https://doi.org/10.1111/jse.12089>
- Wang ZM, Meng SY, Rao GY (2019) Quaternary climate change and habitat preference shaped the genetic differentiation and phylogeography of *Rhodiola* sect. *Prainia* in the southern Qinghai-Tibetan Plateau. *Ecol Evol* 9:8305–8319. <https://doi.org/10.1002/ece3.5406>

- Warren DL, Seifert SN (2011) Ecological niche modeling in Maxent: the importance of model complexity and the performance of model selection criteria. *Ecol Appl* 21:335–342. <https://doi.org/10.1890/10-1171.1>
- Wegmann D, Leuenberger C, Excoffier L (2009) Efficient approximate Bayesian computation coupled with Markov Chain Monte Carlo without likelihood. *Genetics* 182:1207–1218. <https://doi.org/10.1534/genetics.109.102509>
- Wegmann D, Leuenberger C, Neuenschwander S, Excoffier L (2010) ABCtoolbox: a versatile toolkit for approximate Bayesian computations. *BMC Bioinform* 11:116. <https://doi.org/10.1186/1471-2105-11-116>
- Westergaard KB, Zemp N, Bruederle LP et al (2019) Population genomic evidence for plant glacial survival in Scandinavia. *Mol Ecol* 28:818–832. <https://doi.org/10.1111/mec.14994>
- Wipf S, Stoeckli V, Bebi P (2009) Winter climate change in alpine tundra: Plant responses to changes in snow depth and snowmelt timing. *Clim Change* 94:105–121. <https://doi.org/10.1007/s10584-009-9546-x>
- Yang Z, Hu J, Liu N (2007) The influence of dispersal on the meta-population viability of giant panda (*Ailuropoda melanoleuca*) in the Minshan Mountains. *Acta Zool Acad Sci Hungaricae* 53:169–184
- Yang Z-Y, Yi T-S, Pan Y-Z, Gong X (2012) Phylogeography of an alpine plant *Ligularia vellerea* (Asteraceae) in the Hengduan Mountains. *J Syst Evol* 50:316–324. <https://doi.org/10.1111/j.1759-6831.2012.00199.x>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.