Alpine Botany (2018) 128:35–45
https://doi.org/10.1007/s00035-017-0196-8

ORIGINAL ARTICLE

The phylogeographic structure of *Arabis alpina* in the Alps shows consistent patterns across different types of molecular markers and geographic scales

Aude Rogivue[1] · René Graf[1] · Christian Parisod[2] · Rolf Holderegger[1,3] · Felix Gugerli[1]

Received: 17 July 2017 / Accepted: 6 November 2017 / Published online: 16 November 2017

Abstract
Glaciation during the Pleistocene confined alpine species to refugial areas. These range contractions had major impacts on the spatial genetic structure of alpine species. Consequently, one should take into account the often complex phylogeographic structure of species when performing genomic research, e.g. on signatures of local adaptation. Understanding the phylogeography of the widespread arctic and alpine *Arabis alpina* is particularly important, as this species is developing into a model species for ecological genetics. The first objective of this study was to assess the genetic variation of *A. alpina* across the Alps and to compare the spatial genetic patterns resulting from two different types of molecular markers, namely nuclear microsatellites and amplified fragment length polymorphisms (AFLPs). A second objective was to infer the distribution of genetic variation at the regional scale to understand the genetic structure of populations in the area of a previously suggested contact zone between genetic clusters that presumably recolonised their current range from different glacial refugia. We characterized the phylogeographic structure of 372 individuals from 127 populations across the entire Alpine range, complemented by 364 individuals from 22 populations in the western Swiss Alps. Nuclear microsatellite and AFLP markers described consistent population clustering, coherent with previous phylogeographic analyses. Furthermore, regional population structure in the western Alps of Switzerland highlighted a contact zone of genetic clusters associated with different presumed refugia. Again, this finding was in accordance with recolonisation routes formerly inferred for other plant taxa of the western Swiss Alps. Our results highlight the coincidence of large-scale patterns of genetic structure among alternative types of molecular markers and set a valuable basis for further studies on ecological genomics in *A. alpina*.

Keywords AFLPs · Alps · Brassicaceae · Microsatellites · Spatial genetic pattern

Introduction

The Alps were repeatedly covered by ice during the Pleistocene (Kelly et al. 2004), and it has been shown that the climate fluctuations during this period had an impact on the evolution of many species (Hewitt 2004) as a result of reducing their distribution ranges to refugia situated mostly in the South of the Alps (Taberlet et al. 1998; Hewitt 1999). However, it has also been shown that some tree species were present in northern refugia during the last glacial phase in Europe (Robin et al. 2016), substantiating that northerly refugia have contributed to the recolonisation of southern mountain areas such as the Alps. Such a scenario is even more realistic for herbaceous plant species that are adapted to tundra-like habitats, which were widely available in northern central Europe, between the arctic and alpine ice sheets. The survival of alpine plants to glaciation in or around the Alps left diverse and often species-specific patterns, as their glacial history is influenced by ecological demands, breeding system, preglacial distribution patterns and postglacial dispersal (Stehlik 2003). For example, ecological factors such

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00035-017-0196-8) contains supplementary material, which is available to authorized users.

✉ Aude Rogivue
aude.rogivue@gmail.com

[1] WSL Swiss Federal Research Institute, Zürcherstrasse 111, 8903 Birmensdorf, Switzerland

[2] Institute of Plant Sciences, University of Bern, Altenbergrain 21, 3013 Bern, Switzerland

[3] Institute of Integrative Biology, ETH Zürich, Universitätstrasse 16, 8092 Zurich, Switzerland

as the type and distribution of bedrock influenced the current spatial genetic structures (Alvarez et al. 2009), as did deep glacial valleys and high mountain ranges that have shaped the distribution of alleles and species in alpine plants (Thiel-Egenter et al. 2011). Following Holderegger and Thiel-Egenter (2009), three main types of glacial refugia supported the survival of mountain species during the (last) ice age. (1) Lowland glacial refugia are situated outside of mountain systems and beyond the limit of the ice sheet (Schönswetter et al. 2005). The absence of phylogeographic structure in *Oxytropis campestris* subsp. *tiroliensis* in the Alps may be explained by such a dynamics, whereby the species did not survive the Pleistocene glaciation in the interior of the Alps, but migrated to adjacent lowland refugia from where it recolonised its current Alpine range (Schönswetter et al. 2004). (2) Peripheral glacial refugia are situated on the edge of mountain systems at the forefront of the ice field. During the last glacial maximum, such glacial refugia were located along the southwestern, southern, eastern and northern border of the Alps, as described by Schönswetter et al. (2005) by comparing phylogeographic patterns with geological and palaeoenvironmental data of twelve alpine plant species. (3) Nunataks (Brochmann et al. 2003) protruded above the glaciers in the interior of a mountain range, harboring high-mountain taxa, with their respective phylogeographic patterns exemplified by *Eritrichium nanum* (Stehlik et al. 2002a) and *Senecio halleri* (Bettin et al. 2007).

In this study, we investigated the arctic–alpine Brassicaceae *Arabis alpina*, which has become increasingly used as a model species in different fields such as developmental biology (Wang et al. 2009; Wingler et al. 2014), evolutionary ecology (Manel et al. 2010; Poncet et al. 2010; Buehler et al. 2013; Zulliger et al. 2013), studies of mating system (Ansell et al. 2008; Tedder et al. 2011; Buehler et al. 2012) and ecological genomics, as the genome has recently been sequenced (Willing et al. 2015; Jiao et al. 2017). *Arabis alpina* is suggested to have originated in Asia Minor less than 2 million years ago (Ansell et al. 2011). One lineage migrated to the East African high mountains via the Arabian Peninsula, whereas a second lineage expanded across Europe (including arctic populations) and to northwestern Africa. Finally, a third lineage, which is still centred in Asia, migrated independently southward and got into secondary contact with the East African lineage (Koch et al. 2006; Assefa et al. 2007). Growing predominantly on calcareous, base-rich substrate types, the bedrock distribution had an important effect on its migration pathway in the European Alps (Alvarez et al. 2009).

In the Alps, recolonisation of *A. alpina* after glaciation has occurred at least twice, with an early postglacial recolonisation of the western Alps from a refugium presumably located in the Maritime Alps, followed by a second recolonisation of the central Alps from the Balkans (Koch et al. 2006; Ansell et al. 2008). Central areas of the Alps thus likely represent a melting pot of divergent lineages, but the coarse resolution of available studies prevents firm conclusions on where different recolonisation routes met or whether more local refugia also occurred. Furthermore, most of these studies either used uniparentally inherited cpDNA sequences or dominant amplified fragment length polymorphisms (AFLPs) as molecular markers. No study has yet been undertaken to investigate the spatial genetic structure of *A. alpina* across the Alps using highly polymorphic co-dominant markers such as microsatellites. These co-dominant markers with large numbers of alleles per locus should allow for a more detailed analysis of admixture zones among evolutionary lineages stemming from different refugia as well as inferring unusual genotypes potentially pointing to local glacial refugia.

The heterogeneous spatial genetic structure of *A. alpina* highlighted at the edge of several putative glacial refugia in the western Swiss Alps (Ehrich et al. 2007; Alvarez et al. 2009) illustrates how complex the glacial history of species can be at the regional scale within the Alpine arc. Such a situation calls for studies at higher spatial resolution to offer detailed inferences on the postglacial recolonisation process. As summarized by Parisod (2008), migration routes in the western Swiss Alps indeed are intricate, with three main pathways of recolonisation (Online Resource 1) postulated based on floristic evidence (Briquet 1906; Brockmann-Jerosch and Brockmann-Jerosch 1926; Merxmüller et al. 1952). A so-called Rhodanian pathway was likely used by thermophilous species recolonising from southwestern refugia at the front of the Rhone glacier (Christ 1907). Upon glacier melting, species migrated along the Rhone valley and recolonised the French Alps (Haute Savoie) and the Jura mountains (Briquet 1906). A transalpine eastern pathway was supported by the presence of plants that survived the last glaciation in eastern alpine refugia (Briquet 1906; Christ 1907; Brockmann-Jerosch and Brockmann-Jerosch 1926; Merxmüller et al. 1952). It is suggested that these species crossed high-elevation areas in the northern Alps to reach their western range. A third pathway is the transalpine southern route, from the south to the north through the Penninic Alps across high-elevation passes, connecting the Italian Alps to the Valais in Switzerland (Jaccard 1900; Christ 1907; Guyot 1934; Rytz 1951; Delarze 1987). A genetic study of *Biscutella laevigata* based on a dense sampling in western Switzerland illustrated these presumed recolonisation routes (Parisod and Besnard 2007). This species indeed survived the ice ages in refugia of the external Alps as well as in central nunataks and recolonised the western Alps of Switzerland with at least three independent lineages that immigrated through the Rhodanian pathway, the transalpine southern pathway, and via the Simplon Pass. However, there are no comparably detailed genetic studies of other alpine

species available for the western Swiss Alps, even though this area is suggested as an important region where different lineages of alpine plants came into contact during recolonisation (Schönswetter et al. 2005; Thiel-Egenter et al. 2011).

The first and main objective of the present study was to understand the distribution of genetic variation in *A. alpina* in its entire Alpine range (large scale) and to compare the patterns resulting from two alternative types of molecular markers, namely nuclear microsatellites and AFLPs. The second objective was to infer the regional distribution of genetic variation in the western Swiss Alps to understand the genetic structure of populations within this previously described contact zone of lineages from several glacial refugia (regional scale). Our results revealed that both sampling scales and types of marker identified spatial genetic structures that (1) match the previously described general genetic patterns found for calcicolous plant species across the entire Alps (Alvarez et al. 2009) and (2) support the assumption that various genetic clusters from separate refugial areas came into contact in the western Swiss Alps.

Materials and methods

Study species

Arabis alpina is a perennial arctic–alpine herb of the Brassicaceae, which has creeping shoots with a rosette of leaves below the upright flowering shoot. It grows along a wide elevational gradient and is adapted to a wide range of habitat types (Buehler et al. 2013). Most often, it occurs in moist, open pioneer vegetation on calcareous bedrock (Bovet et al. 2006). The species shows various reproductive strategies: it reproduces sexually mainly by selfing or asexually via stoloniferous growth. The small seeds are spread across the alpine landscape by wind, and insect-borne pollen is carried at least up to 1 km (Buehler et al. 2012).

Sampling design and genotyping

To investigate the phylogeography of *A. alpina* at the Alpine scale, we used AFLP data from the project IntraBioDiv (Gugerli et al. 2008; Taberlet et al. 2012a, b). This data set had previously been analysed by Ehrich et al. (2007) and Alvarez et al. (2009). In addition, microsatellite genotyping was performed on the DNAs of the same individuals, i.e. 372 samples from 127 populations (see below). At the regional scale, in the western Swiss Alps, 364 individuals from 22 populations (Online Resource 2) were sampled to investigate the population structure within a region in Western Switzerland, where a contact zone of genetic clusters, presumably originating from different glacial refugia, has previously been identified (Alvarez et al. 2009). For these samples, DNA was extracted from silicagel-dried leaf tissue with the DNeasy Plant Kit (96-well plates; Qiagen, Hilden, Germany) and then also genotyped with microsatellites.

Details on AFLP genotyping are given in Gugerli et al. (2008). All individuals at the Alpine and the regional scale were subjected to microsatellite analysis with 19 microsatellite markers and the protocol described in Buehler et al. (2011). The microsatellite loci were amplified in four multiplex PCRs in 10 µl volumes containing 4.5 µl of Master Mix (Type-it microsatellite PCR Kit, Qiagen), 3.5–4.0 µl H_2O and 0.045–0.250 µl of each primer (5 µM). PCR was performed on a Veriti Thermal Cycler (Applied Biosystems, Foster City, USA). Initial denaturation was set at 94 °C during 5 min and was followed by 27–29 cycles at 94 °C for 30 s, 57 °C for 90 s and 72 °C for 60 s. Final extension lasted for 30 min at 72 °C. PCR products were run with an internal size standard (Applied Biosystems) comprised of 9.5 µl of HiDi-Rox400 on a ABI plate (12 µl Rox400 in 1000 µl HiDi) on a 3730xl DNA Analyzer (Applied Biosystems), and electropherograms were analysed with GENMAPPER 5 (Applied Biosystems). Genotyping errors were minimized using the following strategies (Bonin et al. 2004; DeWoody et al. 2006): PCR conditions were optimized to reduce the error rate and minimize stuttering bands, 12 samples were replicated twice, and samples presenting suspect peaks were re-amplified to confirm their genotypes at particular loci. In the end, only 16 randomly distributed samples were eliminated due to bad amplification or more than 50% missing data, likely due to low DNA quality.

Identification of genetic structure

STRUCTURE (Pritchard et al. 2000; Hubisz et al. 2009) was used to assign individuals to genetic clusters using model-based Bayesian–Markov chain Monte Carlo (MCMC) clustering. Given predominant selfing in populations of *A. alpina* (Tedder et al. 2011; Buehler et al. 2012), we used a model without admixture, with prior information on location and uncorrelated allele frequencies. Ten runs of 1,000,000 MCMC simulations were performed for each value of K (from 1 to 15), after a burn-in period of 100,000 cycles. The appropriate *K* value was selected according to the highest likelihood [LnP(*D*)] and low variance among the 10 runs using the web-based tool STRUCTURE HARVESTER version 0.6.94 (Earl and vonHoldt 2012). We also compared the optimal value of *K* reported by LnP(*D*) to Evanno's summary statistic ΔK (Evanno et al. 2005), but we chose to disregard these values because this approach often gives an optimum of two clusters, which is not informative in our case (Meirmans 2015; Janes et al. 2017). In consequence, we defined the relevant clusters, i.e. those which remained constant and showed no substantial subdivision with increasing *K*. To account for label switching of clusters among

replicates and to combine STRUCTURE runs, CLUMPP 1.1.2 (Jakobsson and Rosenberg 2007) was used (Greedy for $K = 2$–5 and LargeKGreedy for $K > 6$). For the AFLP data, the same settings were used, except that the option RECESSIVEALLELES was set to 1. Finally, we visualized the population clusters in ARCMAP 10.2.1 (ESRI, Redlands, CA, USA). To compare the clustering of both markers, we used the online version of CLUMPAK 1.1 (Kopelman et al. 2015). We estimated the similarity of results from both marker types with the option "comparison" for each K value with the default settings.

The STRUCTURE analyses were done separately for the Alpine and the regional scales, but to feature the origin of samples from the western Swiss Alps (Fig. 2), four (out of 15) individuals per population were also included in the whole Alpine dataset (Online Resource 2) for a separate STRUCTURE analysis.

Because strong inbreeding may lead to incorrect estimation of the number of clusters K (Ehrenreich et al. 2009), we further applied INSTRUCT (Gao et al. 2007). This software is an extension of STRUCTURE that eliminates the assumption of Hardy–Weinberg equilibrium and instead uses inbreeding coefficients or selfing rates to estimate the cluster memberships of individuals. Mode 2 was used, which infers population structure and selfing rate at the population level. The program cannot deal with AFLP data and only uses a model with admixture. Hence, we performed the INSTRUCT analyses for the microsatellite data at the regional scale only (18 populations with 11–23 individuals per population). As for the STRUCTURE analyses, we used CLUMPP and ARCMAP for visualization.

Population genetic analysis

The number of individuals per population genotyped with microsatellites at the regional scale allowed for the assessment of the inbreeding coefficient (F_{IS}), expected and observed heterozygosity (H_e and H_o), calculated with GenAlEx 6.5 (Peakall and Smouse 2006). We calculated pairwise genetic differentiation among populations (pairwise F_{ST}) and the overall F_{ST} with GenAlEx as well. Finally, we checked for isolation by distance (IBD) with SPAGeDi 1.5 (Hardy and Vekemans 2002). We grouped each population pair into five distance classes of 46 pairs each. Permutation tests among locations were performed for tests on each distance class ($n = 10{,}000$) to assess the significance of the regression slope (b-log) between the pairwise F_{ST} and the logarithm of the pairwise distance. To estimate the range of genetic relatedness, we used a spatial autocorrelation approach in GenAlEx using matrices containing the pairwise geographic and genetic distances for the 18 populations. The autocorrelation coefficient (r) was calculated for a specific number of five distance classes chosen to comprise even sample size (0–17, 17–26, 26–34, 34–43 and 43–58 km). We used 1000 bootstraps to estimate 95% confidence intervals for the significance of r against the hypothesis of no spatial structuring.

Results

Population structure at the Alpine scale

Using the microsatellite data, the optimal grouping of the alpine populations was $K = 11$ (Fig. 1a, Online Resource 3a, b). To accurately describe the clustering, it is helpful to understand its development from $K = 2$ to $K = 11$ (Online Resource 4). At $K = 2$, a distinct, non-admixed cluster appeared (light blue) in the central eastern Alps. This group remained constant with increasing K until the optimal number of clusters at $K = 11$ was reached. At $K = 3$, a cluster in the southeastern Alps (red) was emerging and was well defined from $K = 3$ until $K = 11$. At $K = 4$, another split distinguished two major clusters in the North (yellow and turquoise), hence all the main clusters were formed. Once these two northern clusters have been delineated, increasing of K up to 11 mostly changed the clustering within the eastern and the French Alps. In these two regions, up to eight additional clusters with relatively high admixture within populations appeared at $K = 11$. Some individuals belonging to these clusters were sometimes also present in the North of the Alps, highlighting wide geographical disjunctions.

Using the AFLP dataset, the highest likelihood was reached at $K = 10$ (Fig. 1b, Online Resource 3c, d). Similar to the microsatellite analysis, main clusters were already distinctly defined at $K = 4$, and increasing K resulted in changes in the western Alps, but not in the eastern Alps (Online Resource 5). Notably, the overall pattern of the genetic structure for AFLP and microsatellite markers was similar, revealing a cluster arrangement that largely followed a North/South alignment. This observation is supported by the CLUMPAK results. For each K, we obtained a coefficient of similarity between 0.36 and 0.59 among the population membership coefficients of microsatellite and AFLP markers. The coefficient of similarity is slightly decreasing towards higher values of K (Online Resource 6), highlighting increasing discordance between genetic markers in the identification of nested clusters.

Population structure at the regional scale in the western Swiss Alps

The 18 populations from the western Swiss Alps showed an observed heterozygosity (H_o) between 0.02 and 0.21 and expected heterozygosity (H_e) between 0.17 and 0.37, resulting in F_{IS} values between 0.22 and 0.86

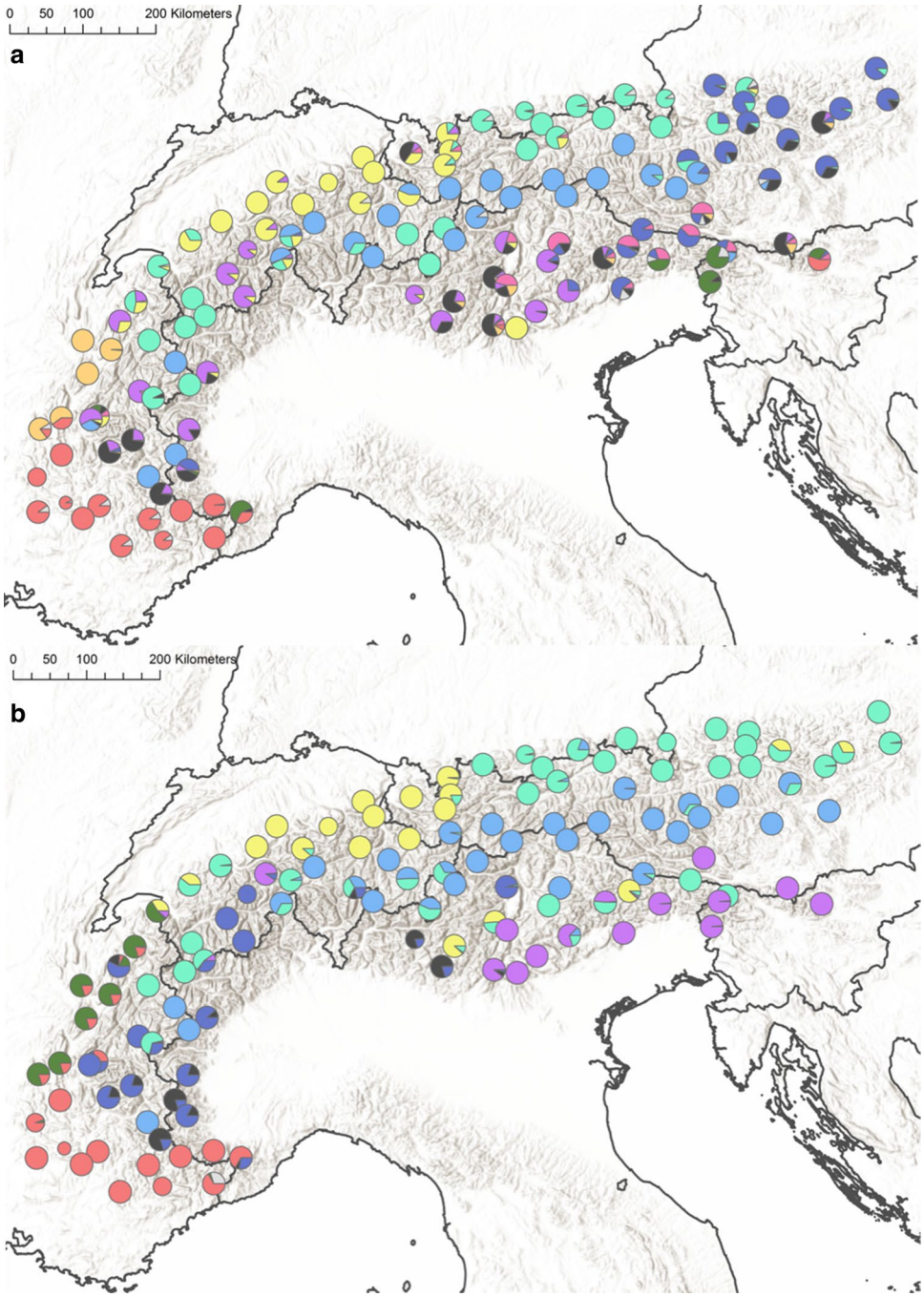

Fig. 1 Genetic clustering of *Arabis alpina*, determined by STRUCTURE (Pritchard et al. 2000; Hubisz et al. 2009); the size of the circles corresponds to sample size. **a** Clustering based on 19 nuclear microsatellite markers across the Alps. The clustering given is for $K = 11$ (for results for $K = 2$–11 see Online Resource 2). **b** Clustering based on 150 AFLP markers across the Alps. The clustering given is for $K = 10$ (for results for $K = 2$–10 see Online Resource 3)

(average = 0.68; more details for each population can be found in the Online Resource 7). The overall value of F_{ST} was 0.436 (± 0.017 SD; for pairwise F_{ST} values among populations see Online Resource 8). To highlight the origin of samples from the western Swiss Alps (Fig. 2), four (out of 15) individuals per population were included in the whole alpine dataset (Online Resource 2). When then considering the whole Alpine and the western Swiss Alpine microsatellite dataset together, $K = 15$ represented the best supported clustering (Online Resource 3e, f). In this context, the populations of the western Swiss Alps were split into three main clusters (yellow, violet and green; Fig. 2). The clustering analysis of the whole western Swiss Alps dataset only (i.e. with all 348 individuals from 18 populations) reached the highest likelihood at $K = 15$ (Online Resource 3g, h; Fig. 2b; Online Resource 9): we thus observed a very distinct and small-scale clustering at the western Swiss Alpine scale. In this region, peripheral populations showed less admixture than populations in the centre of the area. We observed gene pools that

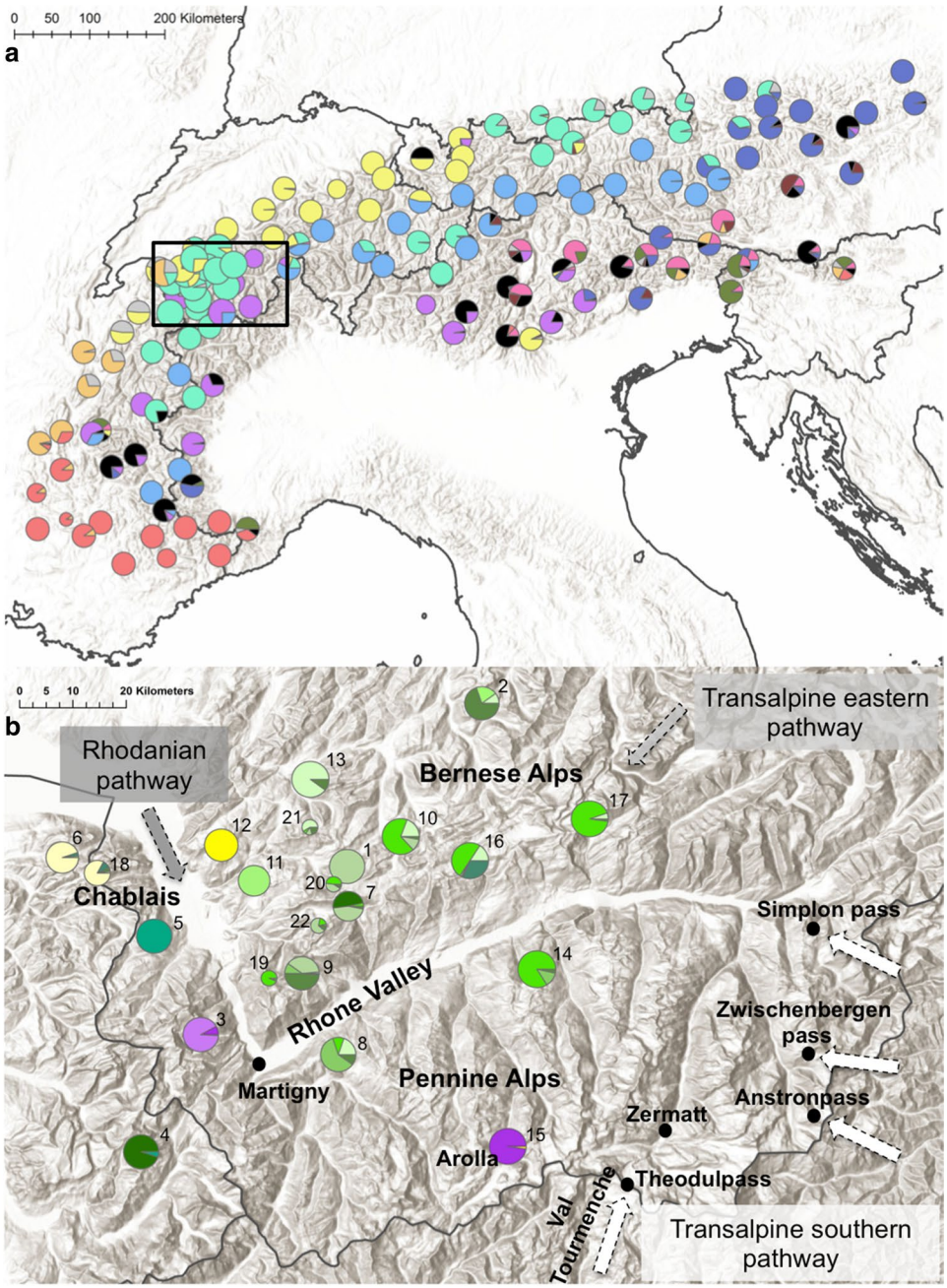

Fig. 2 Genetic clustering of *Arabis alpina*, determined by STRUCTURE (Pritchard et al. 2000; Hubisz et al. 2009), **a** for the Alps including a regionally dense sampling in the western Alps of Switzerland genotyped with 19 microsatellite markers and **b** enlarged for the regional scale; the size of the circles corresponds to sample size. Arrows in **b** show putative recolonisation pathways (Parisod 2008; Online Resource 1) of main clusters: Bernese Alps (green) and Penninic Alps (violet; transalpine southern pathway). The three main colours (yellow, green, violet) indicate to which clusters the regional samples belong at the Alpine scale. The derived colours indicate the sub-clustering when taking into account only the regional samples. Numbers refer to population codes (Online Resource 2)

were present on both sides of the Rhone valley, especially when looking at the clustering progression from $K = 2$ to $K = 15$ (Online Resource 9). We finally tested for IBD because the values of the mean likelihood LnP(D) consistently increased, considered to indicate increased genetic differentiation at larger distances. The results revealed a significant signal of IBD ($p = 0.0039$, slope = 0.08), also supported by the spatial autocorrelation results, plotted in a correlogram (Fig. 3). The latter showed the autocorrelation coefficient (r) as a function of distance classes. Only the first distance class up to 17 km showed significant (positive) spatial autocorrelation. Thereafter, r values decreased, but none significantly differed from the expected value for each given distance class.

The genetic clustering inferred from INSTRUCT was similar to that obtained with STRUCTURE (Online Resource 10), except that populations showed more admixture, although the selfing rate was estimated to be 82.5% for the western Swiss Alpine populations (at $K = 15$, Online Resource 3i, j).

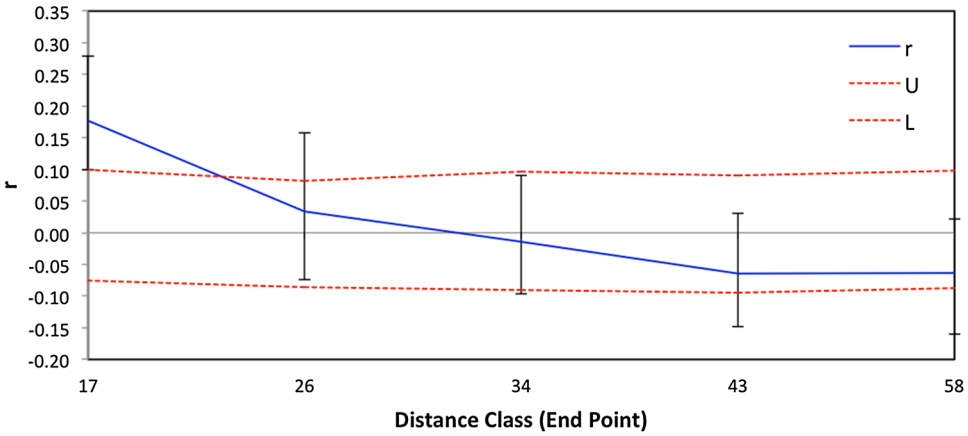

Fig. 3 Range of genetic relatedness estimated based on spatial autocorrelation. The correlogram shows the autocorrelation coefficient (*r*) as a function of five distance classes across the range of the regional study area in the western Swiss Alps, with the upper (U) and lower (L) limits for the 95% confidence interval

Discussion

When studying local adaptation or, more distinctly when identifying loci of adaptive relevance, it is pertinent to also take into account genetic patterns caused by demographic processes such as dispersal and gene flow (Holderegger et al. 2010). Demography and adaptation can indeed cause similar genetic patterns, as illustrated by postglacial recolonisation of Europe out of southern refugia that might be confounded with parallel adaptation to lower temperatures from the South to the North. It is thus important to discriminate demographic from adaptive genetic patterns (Sillanpää 2011). Knowing the neutral genetic structure of a study species a priori may thus be helpful when designing the sampling strategy, as it allows for selecting appropriate regions where genetic structure best matches the respective aims of a study.

Here, we provide such demographic background information for *A. alpina*, an emerging model species for ecological genomics. We present the large-scale genetic structure of this species across the entire Alpine range using two different types of molecular markers. The genetic clustering found was largely congruent with previous, but rather coarse phylogeographic analyses, irrespective of the molecular markers used. This finding highlights the fact that spatial genetic patterns based on AFLPs well reflect the phylogeographic structure in a given species. We further substantiate the North–South alignment of genetic clusters previously identified for calcicolous Alpine plant species, including *A. alpina* (Alvarez et al. 2009). At a finer spatial scale, the regional distribution of genetic variation in *A. alpina* was coherent with a recolonisation of the western Swiss Alps through the different pathways suggested in the floristic literature (Parisod 2008).

Large-scale phylogeography of *A. alpina*

The phylogeographic analyses of *A. alpina* across its entire Alpine range revealed a complex pattern of genetically differentiated groups (Fig. 1), which may be attributed to the convoluted topography and geology of the area in combination with demographic processes related primarily to the last glacial period. Noticeably, the analysis of the spatial genetic structure revealed that the eastern Alps and the French Alps comprised a high number of genetic clusters, which were further subdivided with increasing number of genetic clusters (*K*). Such a high genetic diversity and differentiation across areas described as glacial refugia for calcareous species (Schönswetter et al. 2005) may be related to long-term in situ persistence of populations (Hewitt 1996; Petit et al. 2003). The French Alps were not only surrounded by and comprised glacial refugia, but were further influenced by expanding populations of *A. alpina* from southern refugia like the Apuan Alps (Ansell et al. 2008). Similarly, the high number of clusters in the eastern Alps can be explained by the proximity of glacial refugia and gene exchange with southern and eastern populations such as the Carpathian Mountains, where similar genetic pools were observed (Ehrich et al. 2007). Similar general patterns were reported for several other species in Alvarez et al. (2009), e.g. for *Hornungia alpina*.

Glacial refugia of alpine species preferentially growing on calcareous bedrock were located almost all around the Alps (Schönswetter et al. 2005), with the largest ones occurring along both the southwestern and northeastern ends of the Alps as well as along the Southern Alps. Genetic patterns congruent with survival in small refugia located in the northern margin of the Alps have been recently reported as well (Schneeweiss and Schönswetter 2010). In the following, we attempt to associate these refugial areas with the

population clustering found for *A. alpina* (Fig. 1).The light blue cluster largely retained its central West/East position and range along the hierarchical development of the clustering from $K=2$ to $K=11$. The direction of recolonisation route could not be inferred from the distribution of genetic diversity, given that the number of individuals per population is rather low with an average of three. Even more so, paleoecological evidence would be required to substantiate inference on migration direction. Another well-defined cluster was the red one, found in the Maritime Alps. Such a western Alpine cluster was also observed in other species such as *Pritzelago alpina* (synonymous to the above mentioned *H. alpina*), in which western Alpine populations clustered with the Pyrenees, whereas central Alpine populations formed another cluster (Kropf et al. 2003). The positions of the different genetic clusters discussed coincide with regions of high genetic differentiation and mostly also harbour high genetic diversity, which can be interpreted as secondary contact zones of clusters from different glacial refugia during postglacial expansions (Duforet-Frebourg and Blum 2014). However, only inference based on allelic richness and the degree of heterozygosity may be informative regarding the distinction between a contact zone and glacial refugium; our data set with few samples per location remains inconclusive in this regard.

Based on the main glacial refugia observed across the Alps, Schönswetter et al. (2005) inferred biogeographic lines which correspond to biogeographic boundaries based on floristic evidence, supported by silicicolous and calcicolous species (Ozenda 1985). Genetic patterns of silicicolous plant taxa largely conformed to these floristic boundaries (Thiel-Egenter et al. 2011), but the distribution of genetic clusters in *A. alpina* and the respective contact zones did not follow these three biogeographic lines. The clustering indeed revealed primarily a North/South alignment of clusters that appear typical of calcicolous taxa (Alvarez et al. 2009). However, a comprehensive comparative study has not been done to identify areas of shared contact zones for calcicolous plant species at the inter- and intraspecific level.

Comparison between nuclear microsatellite and AFLP markers

Using alternative molecular markers may yield different results of genetic analyses (Nybom 2004; Woodhead et al. 2005). We assessed the coincidence of the spatial genetic structures as determined with (mainly) nuclear AFLPs and nuclear microsatellites. Both marker types showed largely consistent patterns of genetic variation at the Alpine scale, and the optimal values of K were almost identical for AFLPs and microsatellites. Generally, clusters based on AFLP markers were less admixed than those based on microsatellites, independent of the number of K, although a model without admixture was used. Apparently, individuals are more often assigned to the same cluster with AFLP markers than with microsatellites. This is possibly caused by the higher number of AFLP markers (150 loci) compared to 19 microsatellites, but could also be a result of the higher number of alleles in the latter as compared to the bi-allelic and dominant AFLP markers. Our comparison of both markers confirmed the conclusion of Estoup et al. (2002) that mutation rate does not represent a substantial problem in many population genetic analyses, although the higher mutation rate at microsatellite loci may lead to non-negligible homoplasy (Skrede et al. 2009). One of the major differences between both marker types was found in the eastern and the western part of the Alps. We described them as genetically diverse regions in the case of microsatellites, but these regions appeared less diverse with AFLPs. This observation does not unambiguously support these areas as located next to or even representing areas of glacial refugia. Instead, cluster diversity was increased in the northwestern part of the Alps (mostly from $K=8$ onwards, Online Resource 5), a region also surmised as a glacial refugium by Schönswetter et al. (2005). Our sampling design at the large scale, with largely three individuals genotyped per location, does not allow us to infer the underlying processes that have lead to the patterns observed in more detail.

To conclude, the two types of markers detected similar main patterns of genetic clustering, which was coherent with corresponding observations in other plant species (Teulat et al. 2000; Gaudeul et al. 2004). This finding is relevant when designing future studies, e.g. on signatures of local adaptation, in a taxon for which information on the spatial genetic structure is available only from AFLP data, often criticized for their anonymous nature.

Contact zone in western Swiss Alps

The second objective of this study was to infer the distribution of genetic variation in populations of *A. alpina* at the regional scale in the western Swiss Alps and compare the results to recolonisation routes outlined in Parisod (2008). At the regional scale, we argue that a higher genetic diversity could be explained by the junction of diverse genetic lineages (Fig. 2). Note, however, that this area was not described as a distinct contact zone of *A. alpina* by Duforet-Frebourg and Blum (2014), who used the same AFLP data set as in our study, but an alternative approach for identifying areas of high genetic differentiation.

When interpreting the clustering of the western Swiss Alps included in the analysis of the entire range of the Alps, we observed three genetic clusters (yellow, green and violet, Fig. 2b) in the western Swiss Alps. The three clusters could be associated with the Chablais, the Bernese and the Penninic Alps. This association contrasts with the relative

genetic homogeneity described over this area by Parisod and Besnard (2007) in support of the Rhodanian recolonisation pathway. In *A. alpina*, the genetic cluster of the Rhone Valley in the Chablais region (the "yellow" populations: 6, 12 and 18; Fig. 2b) indeed seemed to originate from a northern refugium, and not from a southern refugium as would be expected for the Rhodanian pathway. Furthermore, this genetic cluster is not expanding far into the Rhone Valley and came only into contact with the "green" cluster from the Bernese Alps, which seems to have a different source population. In turn, this "yellow" cluster came into contact with the "violet" populations (populations 3 and 15; Fig. 2b), which clustered with southern Alpine populations. Therefore, the transalpine southern pathway seems the most suitable explanation for the recolonisation history of the "violet" cluster. Pairwise F_{ST} values among populations confirmed the population assignment of STRUCTURE: samples from populations within clusters had lower pairwise F_{ST} values than samples from populations of different clusters. For example, the "violet" populations (populations 3 and 15; Fig. 2b) had a pairwise F_{ST} of 0.243 that is lower than values obtained when pairing these two populations with any of the surrounding populations that belong to the "green" cluster (Online Resource 8).

In this study, the assignment of the Bernese populations was consistent with the transalpine eastern pathway, which had been postulated based on floristic evidence (Welten 1982). However, it could also be that *A. alpina* survived in situ in the northern Prealps (for instance in the Chablais region) or the nearby Jura mountains as suggested for *Erinus alpinus* (Stehlik et al. 2002a, b; Stehlik 2003). This primarily northern-Alpine, rock-dwelling species is restricted to calcareous substrate as generally is *A. alpina*, and we observed an almost identical phylogeographic pattern in the western Swiss Alps as for *E. alpinus*. To confirm a transalpine recolonisation pathway for *A. alpina*, it would be necessary to sample more populations in the southern part of the Valais and in neighbouring Italian regions, but also plastid markers might be more informative in this regard. When only the regional data are considered, the optimal clustering at $K = 15$ in the western Swiss Alps revealed the complex clustering. Finally, the low but significant signal of IBD and the spatial autocorrelation results indicated that genetic differentiation gradually declined with increasing geographic distance, but with only a significantly positive correlation for population pairs located at up to 17 km (Fig. 3). It shows that gene flow between clusters is spatially limited. Accordingly, geography certainly plays an important role in the structure of the populations. Hence, our regionally dense sampling and microsatellite genotyping revealed a more complex genetic structure than hitherto assumed from large-scale studies. Our results suggest that *A. alpina* in the western Swiss Alps was made up of admixed populations because this area constitutes a contact zone between genetically distinct groups that immigrated from several glacial refugia.

No effect of high selfing rate on analyses of genetic structure

Finally, we took into account that the sexual reproduction of *A. alpina* occurs mostly by selfing (Ansell et al. 2008; Tedder et al. 2011; Buehler et al. 2012) and evaluated how this mating system might affect our analyses. When we looked at the population clustering not assuming Hardy–Weinberg equilibrium and performed in INSTRUCT, the results were not much different from those revealed by STRUCTURE. Hence, these results corroborated the general patterns found and discussed above. The inbreeding coefficients of the populations in the western Swiss Alps, on average 0.68 (SD = 0.18; Online Resource 7), are close to the values found in the literature for the western Alps based on allozyme data ($F_{IS} = 0.58$; Ansell et al. 2008). The inbreeding coefficient was also calculated in two studies on *A. alpina* from one population in the central Alps ($F_{IS} = 0.66$; Buehler et al. 2012) and across the Alpine range ($F_{IS} = 0.65$; Ansell et al. 2008). These highly positive values, indicative of a substantial excess of homozygote individuals as compared to expectations under random mating, support the observation that *A. alpina* is a predominant selfer. Differences in inbreeding coefficients between regions have been explained as changes in the breeding system of populations originating from different post-glacial recolonisation routes (Ansell et al. 2008). However, the estimation of the selfing rate is less accurate when there is no substructure or when subpopulations have similar selfing rates (Gao et al. 2007) and when selfing (or inbreeding) rates are calculated based on different molecular markers, e.g. microsatellites versus allozymes (Ansell et al. 2008). The clustering of populations was not much different between the STRUCTURE and the INSTRUCT analyses in our study (see above; Online Resource 10), probably because we accounted for selfing using a model without admixture in STRUCTURE. Other studies also did not find differences between these alternative approaches (McNally et al. 2009; Mosher et al. 2007).

In conclusion, our phylogeographic study of *A. alpina* across the entire Alpine range, combined with a regionally dense sampling, reveals that the spatial genetic structure of this species follows those previously presented for other calcicolous plant species. Knowledge about the spatial position and distribution of genetic clusters will be useful for planning further studies, for example when inferring local adaptation at the genome level. The results further indicate that the outcome of analyses on the genetic structure are not substantially influenced by the type of molecular marker used, nor the underlying breeding system, which lends support to the credibility of the results presented here.

Acknowledgements We thank Maurice Moor, Alexandra Foetisch and Sabine Brodbeck for the sampling in the western Alps of Switzerland and the IntraBioDiv Consortium for the DNA samples of the entire Alpine range. Christian Rellstab contributed to discussions, and two anonymous reviewers provided valuable comments. This study was financially supported by the Swiss National Science Foundation (GeneScale; CR32I3_149741/1).

Author contributions Aude Rogivue and Felix Gugerli designed the study, René Graf did the lab work and the genotyping, Aude Rogivue performed all the analyses and wrote the manuscript, with contributions from Christian Parisod, Rolf Holderegger and Felix Gugerli.

Compliance with ethical standards

Conflict of interest All authors declare that they have no conflicts of interests.

References

Alvarez N et al (2009) History or ecology? Substrate type as a major driver of patial genetic structure in Alpine plants. Ecol Lett 12:632–640

Ansell SW, Grundmann M, Russell SJ, Schneider H, Vogel JC (2008) Genetic discontinuity, breeding-system change and population history of *Arabis alpina* in the Italian Peninsula and adjacent Alps. Mol Ecol 17:2245–2257

Ansell SW, Stenøien HK, Grundmann M, Russell SJ, Koch MA, Schneider H, Vogel JC (2011) The importance of Anatolian mountains as the cradle of global diversity in *Arabis alpina*, a key arctic-alpine species. Ann Bot 108:241–252

Assefa A, Ehrich D, Taberlet P, Nemomissa S, Brochmann C (2007) Pleistocene colonization of afro-alpine 'sky islands' by the arctic-alpine *Arabis alpina*. Heredity 99:133–142

Bettin O, Cornejo C, Edwards P, 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

Bonin A, Bellemain E, Bronken Eidesen P, Pompanon F, Brochmann C, Taberlet P (2004) How to track and assess genotyping errors in population genetics studies. Mol Ecol 13:3261–3273

Bovet L, Kammer P, Meylan-Bettex M, Guadagnuolo R, Matera V (2006) Cadmium accumulation capacities of *Arabis alpina* under environmental conditions. Environ Exp Bot 57:80–88

Briquet J (1906) Le development des flores dans les Alpes occidentales avec aperçu sur les Alpes en général. In: von Wetterstein R, Wiesner J, Zahlbruckner A (eds) Wissenschaftliche Ergebnisse des Internationalen Botanischcn Kongresses Wien 1905. Gustav Fischer, Jena, pp 130–173

Brochmann C, Gabrielsen TM, Nordal I, Landvik JY, Elven R (2003) Glacial survival or tabula rasa? The history of North Atlantic biota revisited. Taxon 52:417–450

Brockmann-Jerosch H, Brockmann-Jerosch MC (1926) Die Geschichte der Schweizerischen Alpenflora. In: Schröter C (ed) Das Pflanzenleben der Alpen. Raustein, Zürich, pp 1110–1215

Buehler D, Graf R, Holderegger R, Gugerli F (2011) Using the 454 pyrosequencing-based technique in the development of nuclear microsatellite loci in the alpine plant *Arabis alpina* (Brassicaceae). Am J Bot 98:e103–e105

Buehler D, Graf R, Holderegger R, Gugerli F (2012) Contemporary gene flow and mating system of *Arabis alpina* in a Central European alpine landscape. Ann Bot 109:1359–1367

Buehler D, Poncet BN, Holderegger R, Manel S, Taberlet P, Gugerli F (2013) An outlier locus relevant in habitat-mediated selection in an alpine plant across independent regional replicates. Evol Ecol 27:285–300

Christ H (1907) La flore de la Suisse. Georg & Cie, Bâle-Genève-Lyon

Delarze R (1987) L'origine des pelouses steppiques valaisannes à la lumière de leurs liens de parenté avec les régions limitrophes. Bull Murith Soc Valais Sci Nat 105:41–70

DeWoody J, Nason JD, Hipkins VD (2006) Mitigating scoring errors in microsatellite data from wild populations. Mol Ecol Notes 6:951–957

Duforet-Frebourg N, Blum MG (2014) Non stationary patterns of isolation-by-distance: infering measures of local genetic differentiation with Bayesian kriging. Evol 68:1110–1123

Earl DA, von Holdt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv Genet Resour 4:359–361

Ehrenreich IM, Hanzawa Y, Chou L, Roe JL, Kover PX, Purugganan MD (2009) Candidate gene association mapping of *Arabidopsis* flowering time. Genetics 183:325–335

Ehrich D et al (2007) Genetic consequences of Pleistocene range shifts: contrast between the Arctic, the Alps and the East African mountains. Mol Ecol 16:2542–2559

Estoup A, Jarne P, Cornuet JM (2002) Homoplasy and mutation model at microsatellite loci and their consequences for population genetics analysis. Mol Ecol 11:1591–1604

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14:2611–2620

Gao H, Williamson S, Bustamante CD (2007) A Markov chain Monte Carlo approach for joint inference of population structure and inbreeding rates from multilocus genotype data. Genetics 176:1635–1651

Gaudeul M, Till-Bottraud I, Barjon F, Manel S (2004) Genetic diversity and differentiation in *Eryngium alpinum* L. (Apiaceae): comparison of AFLP and microsatellite markers. Heredity 92:508–518

Gugerli F et al (2008) Relationships among levels of biodiversity and the relevance of intraspecific diversity in conservation—a project synopsis. Perspect Plant Ecol Evol Syst 10:259–281

Guyot H (1934) Phytogéographie comparée du Valais et de la vallée d'Aoste. Bull Murith Soc Valais Sci Nat 52:16–35

Hardy OJ, Vekemans X (2002) SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. Mol Ecol Resour 2:618–620

Hewitt GM (1996) Some genetic consequences of ice ages, and their role in divergence and speciation. Biol J Linn Soc 58:247–276

Hewitt GM (1999) Post-glacial re-colonization of European biota. Biol J Linn Soc 68:87–112

Hewitt G (2004) Genetic consequences of climatic oscillations in the Quaternary. Phil Trans R Soc B 359:183–195

Holderegger R, Thiel-Egenter C (2009) A discussion of different types of glacial refugia used in mountain biogeography and phylogeography. J Biogeogr 36:476–480

Holderegger R, Buehler D, Gugerli F, Manel S (2010) Landscape genetics of plants. Trends Plant Sci 15:675–683

Hubisz MJ, Falush D, Stephens M, Pritchard JK (2009) Inferring weak population structure with the assistance of sample group information. Mol Ecol Resour 9:1322–1332

Jaccard P (1900) Contribution au problème de l'immigration postglacaire de la flore alpine. Bull Soc Vaud Sc Nat 36:87–130

Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics 23:1801–1806

Janes JK, Miller JM, Dupuis JR, Malenfant RM, Gorrell JC, Cullingham CI, Andrew RL (2017) The K = 2 conundrum. Mol Ecol 26:3594–3602
Jiao W-B et al (2017) Improving and correcting the contiguity of long-read genome assemblies of three plant species using optical mapping and chromosome conformation capture data. Genome Res 27.5:778–786
Kelly MA, Buoncristiani J-F, Schlüchter C (2004) A reconstruction of the last glacial maximum (LGM) ice-surface geometry in the western Swiss Alps and contiguous Alpine regions in Italy and France. Eclogae Geol Helv 97:57–75
Koch MA, Kiefer C, Ehrich D, Vogel J, Brochmann C, Mummenhoff K (2006) Three times out of Asia Minor: the phylogeography of *Arabis alpina* L. (Brassicaceae). Mol Ecol 15:825–839
Kopelman NM, Mayzel J, Jakobsson M, Rosenberg NA, Mayrose I (2015) Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. Mol Ecol Resour 15:1179–1191
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
Manel S, Poncet B, Legendre P, Gugerli F, Holderegger R (2010) Common factors drive adaptive genetic variation at different spatial scales in *Arabis alpina*. Mol Ecol 19:3824–3835
McNally KL et al. (2009) Genomewide SNP variation reveals relationships among landraces and modern varieties of rice. Proc Natl Acad Sci USA 106:12273–12278
Meirmans PG (2015) Seven common mistakes in population genetics and how to avoid them. Mol Ecol 24:3223–3231
Merxmüller H (1952) Untersuchungen zur Sippengliederung und Arealbildung in den Alpen. I. Jb. Ver. Schutze d. Alpenpflanzen u Tiere 17:96–133
Mosher DS, Quignon P, Bustamante CD, Sutter NB, Mellersh CS, Parker HG, Ostrander EA (2007) A mutation in the myostatin gene increases muscle mass and enhances racing performance in heterozygote dogs. PLoS Genet 3:779–786
Nybom H (2004) Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. Mol Ecol 13:1143–1155
Ozenda P (1985) La végétation de la chaine alpine dans l'espace montagnard européen. Masson, Paris
Parisod C (2008) Postglacial recolonisation of plants in the western Alps of Switzerland. Bot Helv 118:1–12
Parisod C, Besnard G (2007) Glacial in situ survival in the Western Alps and polytopic autopolyploidy in *Biscutella laevigata* L. (Brassicaceae). Mol Ecol 16:2755–2767
Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol Ecol Notes 6:288–295
Petit RJ et al (2003) Glacial refugia: hotspots but not melting pots of genetic diversity. Science 300:1563–1565
Poncet BN et al (2010) Tracking genes of ecological relevance using a genome scan in two independent regional population samples of *Arabis alpina*. Mol Ecol 19:2896–2907
Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959
Robin V, Nadeau MJ, Grootes PM, Bork HR, Nelle O (2016) Too early and too northerly: evidence of temperate trees in northern Central Europe during the Younger Dryas. New Phytol 212:259–268
Rytz W (1951) Environs de Zermatt et de Saas. Le rôle biogéographique des cols. In: Chouard P, Gauss H, Vischer W (eds) Coupe Botanique des Alpes du Tyrol à la France. Bulletin de la Société Botanique de France, Paris, pp 77–78
Schneeweiss G, 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
Schönswetter P, Tribsch A, Niklfeld H (2004) Amplified fragment length polymorphism (AFLP) reveals no genetic divergence of the Eastern Alpine endemic *Oxytropis campestris* subsp. *tiroliensis* (Fabaceae) from widespread subsp. *campestris*. Plant Syst Evol 244:245–255
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
Sillanpää M (2011) Overview of techniques to account for confounding due to population stratification and cryptic relatedness in genomic data association analyses. Heredity 106:511–519
Skrede I, Borgen L, Brochmann C (2009) Genetic structuring in three closely related circumpolar plant species: AFLP versus microsatellite markers and high-arctic versus arctic-alpine distributions. Heredity 102:293–302
Stehlik I (2003) Resistance or emigration? Response of alpine plants to the ice ages. Taxon 52:499–510
Stehlik I, Blattner F, Holderegger R, Bachmann K (2002a) Nunatak survival of the high Alpine plant *Eritrichium nanum* (L.) Gaudin in the central Alps during the ice ages. Mol Ecol 11:2027–2036
Stehlik I, Schneller JJ, Bachmann K (2002b) Immigration and in situ glacial survival of the low-alpine *Erinus alpinus* (Scrophulariaceae). Biol J Linn Soc 77:87–103
Taberlet P, Fumagalli L, Wust-Saucy AG, Cosson JF (1998) Comparative phylogeography and postglacial colonization routes in Europe. Mol Ecol 7:453–464
Taberlet P et al (2012a) Genetic diversity in widespread species is not congruent with species richness in alpine plant communities. Ecol Lett 15:1439–1448
Taberlet P et al (2012b) Data from: Genetic diversity in widespread species is not congruent with species richness in alpine plant communities. In: Dryad Data Repository, http://dx.doi.org/10.5061/dryad.s4q6s. Accessed 20 Oct 2016
Tedder A, Ansell S, Lao X, Vogel J, Mable B (2011) Sporophytic self-incompatibility genes and mating system variation in *Arabis alpina*. Ann Bot 108:699–713
Teulat B et al (2000) An analysis of genetic diversity in coconut (*Cocos nucifera*) populations from across the geographic range using sequence-tagged microsatellites (SSRs) and AFLPs. Theor Appl Genet 100:764–771
Thiel-Egenter C et al (2011) Break zones in the distributions of alleles and species in alpine plants. J Biogeogr 38:772–782
Wang R et al (2009) PEP1 regulates perennial flowering in *Arabis alpina*. Nature 459:423–427
Welten M (1982) Vegetationsgeschichtliche Untersuchungen in den westlichen Schweizer Alpen: Bern-Wallis. Mém Soc Hel Sci Nat 95:12–27
Willing E-M et al (2015) Genome expansion of *Arabis alpina* linked with retrotransposition and reduced symmetric DNA methylation. Nat Plants 1:14023
Wingler A, Juvany M, Cuthbert C, Munné-Bosch S (2014) Adaptation to altitude affects the senescence response to chilling in the perennial plant *Arabis alpina*. J Exp Bot 66:355–367
Woodhead M, Russell J, Squirrell J, Hollingsworth P, Mackenzie K, Gibby M, Powell W (2005) Comparative analysis of population genetic structure in *Athyrium distentifolium* (Pteridophyta) using AFLPs and SSRs from anonymous and transcribed gene regions. Mol Ecol 14:1681–1695
Zulliger D, Schnyder E, Gugerli F (2013) Are adaptive loci transferable across genomes of related species? Outlier and environmental association analyses in Alpine Brassicaceae species. Mol Ecol 22:1626–1639