

Scale matters: genome-wide signatures of local adaptation to high-resolution environmental variation in an alpine plant

Running title Scale effects in local adaptation

Aude Rogivue ^{1,§*}, Kevin Leempoel ^{2,3,§}, Annie S. Guillaume ², Rimjhim Roy Choudhury ^{4,5}, François Felber ⁶, Michel Kasser ⁷, Stéphane Joost ², Christian Parisod ^{4,5*}, Felix Gugerli ^{1*}

¹ Swiss Federal Research Institute for Forest, Snow and Landscape Research, Zürcherstrasse 111, CH–8903 Birmensdorf, Switzerland

² Ecole Polytechnique Fédérale de Lausanne (EPFL), Geospatial Molecular Epidemiology group (GEOME), Laboratory for Biological Geochemistry (LGB), Civil and Environmental Engineering (ENAC), Station 2, 1015 Lausanne, Switzerland

³ current address: Biolizard, Ottergemsesteenweg Zuid 808 bus B511, 9000 Ghent, Belgium

⁴ University of Berne, Institute of Plant Sciences, Altenbergrain 21, 3013 Bern, Switzerland

⁵ current address: University of Fribourg, Department of Biology, Chemin du Musée 10, 1700 Fribourg, Switzerland

⁶ Musée et Jardins botaniques cantonaux, av. de Cour 14B, 1007 Lausanne, Switzerland

⁷ Haute-Ecole d'Ingénierie et de Gestion (HEIG), INSIT Laboratory, Route de Cheseaux 1, 1401 Yverdon-Les-Bains, Switzerland

[§] Aude Rogivue and Kevin Leempoel are considered joint first authors

^{*} corresponding authors: aude.rogivue@gmail.com, christian.parisod@unifr.ch, felix.gugerli@wsl.ch

Abstract

Microevolutionary processes shape adaptive responses to heterogeneous environments, where these effects vary both among and within species. However, it remains largely unknown to which degree signatures of adaptation to environmental drivers can be detected based on the choice of spatial scale and genomic marker. We studied signatures of local adaptation across different levels of spatial extents, investigating complementary types of genomic variants—single nucleotide polymorphisms (SNPs) and polymorphic transposable elements (TEs)—in populations of the alpine model plant species *Arabis alpina*. We coupled environmental factors, derived from remote sensing digital elevation models at very high resolution (0.5m), with whole-genome sequencing data of 304 individuals across four populations. We demonstrate that responses of *A. alpina* to similar amounts of abiotic variation are largely governed by local evolutionary processes, and we find minimally overlapping signatures of local adaptation between SNPs and polymorphic TEs. Notably, functional annotations of high-impact genomic variants revealed several defence-related genes associated with the abiotic factors studied, which could indicate indirect selective pressure of biotic agents. Our results highlight the importance of considering different spatial extents and types of genomic polymorphisms when searching for signatures of adaptation to environmental variation. Such insights provide key information on microevolutionary processes and could guide management decisions to mitigate negative impacts of climate change on alpine plant populations.

Key words: *Arabis alpina*, genomic variation, genotype–environment associations, high-resolution environmental variation, local adaptation, remote sensing

Introduction

Local adaptation enables populations to evolve traits that confer a fitness advantage in their respective habitats (Kawecki & Ebert, 2004). Genomic variation may reveal signatures of such adaptive processes, but identifying the underlying environmental drivers and relevant spatial scale remains challenging (Savolainen et al., 2013). To improve our understanding of traits under selection and their underlying functional processes, one may decipher the imprints of local adaptation based on whole-genome patterns of variation. However, to our knowledge, fine-grained sampling of individuals and environmental data, at a resolution matching genomic variation, has never been used to assess environmental drivers of adaptation.

Studies investigating local adaptation often cover broad spatial extents (commonly referred to as scales), involving continent-wide sampling for example (Gougherty et al., 2021; Pais et al., 2017; Yeaman et al., 2016), but see (de Villemereuil et al., 2018; Eckert et al., 2015; Fischer et al., 2013). However, adaptation can be truly local (Rellstab et al., 2017), particularly in highly heterogeneous alpine landscapes where environmental conditions vary over short distances (Rellstab et al., 2020).

In their local study on the alpine Brassicaceae *Biscutella laevigata*, Leempoel et al. (2018) highlight the importance of considering the ratio between grain (spatial resolution) and extent of a study area to assess signatures of adaptation, which should match the spatial context at which selection operates (Dauphin et al., 2023). In rugged alpine terrain, an organism's habitat can be adequately described using ecological data derived from digital elevation models (DEM; Kozak et al., 2008), which correlate significantly with climatic factors measured using ground-based sensors (Leempoel et al., 2015)(Pradervand et al., 2014). Such topography-derived environmental descriptors provide continuous information across habitats, reflecting long-term patterns of local site conditions, as compared to point-based measurements acquired on-site or interpolated from local weather stations, which provide short-term and often patchy environmental information.

Whole-genome studies of adaptation in natural plant populations mostly remain restricted to model organisms like *Arabidopsis thaliana* (Exposito-Alonso et

al., 2018; Fournier-Level et al., 2011; Hancock et al., 2011) or crop wild relatives (e.g., Yoder et al., 2014). Additionally, they are often based on pooled sequencing (Fischer et al., 2013; Rellstab et al., 2020; Turner et al., 2010), or only consider a few individuals that do not necessarily represent the local populations (Laenen et al., 2018). Moreover, structural variants, often involved in adaptation beyond what can be captured by commonly analysed single-nucleotide polymorphisms (SNPs), have been largely ignored in these studies. Transposable elements (TEs), for example, represent a major fraction of many plant genomes (Bennetzen & Wang, 2014; Quesneville, 2020) and are known to induce adaptive responses to abiotic stress (Baduel et al., 2021; Kalendar et al., 2000), yet have remained largely overlooked in studies of local adaptation. Indeed, biallelic SNPs and TEs are expected to have complementary effects on the adaptive responses of plants to local conditions. For example, polymorphic TEs are known to affect recombination along chromosomes, such that they may promote entire blocks of linked loci to reveal signatures of selection along chromosomes (Choudhury et al., 2019), while genomic variation in non-synonymous SNPs may induce rather subtle changes in the respective proteins. Therefore, it is expected that these marker types identify few common, but rather complementary signals of local adaptation.

It is commonly assumed that similar environmental contrasts across sites leave consistent signatures of selection in plants with a common genomic background (Booker et al., 2021; Lotterhos & Whitlock, 2015) (Rellstab et al., 2020). However, for populations with limited gene flow, particularly in species with mixed mating systems or selfing, one may expect population-specific adaptive responses with limited convergence (Rellstab et al., 2017; 2020).

Arabis alpina (Brassicaceae) has recently emerged as a model species for ecological genomics (Wötzel et al., 2022). Studies have provided empirical evidence of how local adaptation acts on the species' genome (Laenen et al., 2018; Lobréaux & Miquel, 2020; e.g., Poncet et al., 2010), with notable effects on phenotypes (Chopra et al., 2019; de Villemereuil et al., 2018; Vayssières et al., 2020; Wang et al., 2009). As detailed in (Wötzel et al., 2022), several field studies, including reciprocal transplant experiments, substantiate the potential of *A. alpina* to adapt on small spatial scales, e.g., in response to elevational differences (Buckley et al., 2019; de Villemereuil et al.,

2018; Wingler et al., 2015). Due to this species' responsiveness to environmental conditions, and the availability of a high-quality reference genome, it is an intriguing organism to further investigate whether signatures of selection to fine-scale environmental drivers are found across the whole genome.

Here, we re-sequenced 304 individuals of *A. alpina*, sampled in local populations with a common history, using a spatially explicit, hierarchical design (individuals in plots nested within four local populations, Figures 1 and 2). We computed the environmental factors at a very high resolution (0.5 m) to describe the local habitat for each individual. Through associations with genomic variation, we inferred drivers of local adaptation (i) at the regional level, i.e., among the four local populations, and (ii) at the local level, i.e., among plots within each local population. In parallel, we performed outlier detection analyses to identify possible genomic signatures of local adaptation irrespective of underlying environmental drivers. We focused on whether signals of adaptation between regional and local spatial extents, as well as between types of genomic markers (SNPs vs. TEs), show common or different patterns and functions, and to what degree patterns of local adaptation are driven by particular environmental factors at different spatial extents.

Materials and Methods

Study species and sampling

Arabis alpina L. has a broad ecological niche covering a large altitudinal and latitudinal range, making it suitable for investigating adaptation to variation across heterogeneous environments (Wötzel et al., 2022). This perennial Brassicaceae is found throughout the Northern Hemisphere, predominantly on calcareous bedrock. In the European Alps, *A. alpina* is found across an elevation range of 400 m to 3200 m a.s.l. This species reproduces sexually, predominantly via selfing (Ansell et al., 2008; Buehler et al., 2012; Laenen et al., 2018; Tedder et al., 2011), though outcrossing populations can be found in parts of its range (e.g., Ansell et al., 2008; Tedder et al., 2011). SNP-based estimates of individual inbreeding coefficients (F), as well as positive F_{IS} values for both SNPs and polymorphic TEs, suggest substantial inbreeding

in the studied populations, likely due to extensive selfing (Choudhury et al., 2019; Rogivue et al., 2019b).

We sampled four local populations of *A. alpina* from the western Swiss Alps in summer 2014 (Figure 1 and 2), from temperate alpine climates with generally abundant rain and snowfall (Drake et al., 2006). Individuals were sampled predominantly from calcareous bedrock in the alpine vegetation zone above 2000 m a.s.l. in the regions Para (N = 69), Pierredar (N = 69), Essets (N = 70) and Martinets (N = 96) (Table S1; see also Rogivue et al., 2019b). The four studied populations are considered descendants from the same refugial population and share a similar genomic background (Rogivue et al., 2018). Each population is situated within a different catchment valley (Figure 1a), presumably preventing regular gene flow between these populations. However, mating within each population can occur over large distances, as insects may transport pollen up to 1 km (Buehler et al., 2012) and their small, light seeds can be disseminated by wind. Within each population, we selected ten plots from contrasting habitat types to encompass large ecological gradients, within which we sampled the leaves of six to ten plants at least 1 m apart to avoid sampling families (Figure 1 and Table S1). We stored the samples in silica gel and extracted DNA as described elsewhere (Rogivue et al., 2019b). The precise geographic coordinates of each individual (± 2 cm) were obtained with a differential global positioning system (DGPS) receiver, using information provided by Real Time Kinematik (RTK) Global Navigation Satellite System (GNSS) (Table S1).

Environmental factors

Using SAGA GIS (Conrad et al., 2015), we computed 13 topography-derived environmental factors (Table S2) from a precise and high-resolution DEM at 0.5 m pixel resolution, based on light detection and ranging (LiDAR) data obtained from the regional authorities (Canton Vaud, Switzerland). These topographic factors are known proxies for ecologically relevant descriptors, including elevation (and therefore temperature), climate, hydrology, soil conditions, light availability and exposure (Guillaume et al., 2021; Lecours et al., 2017; Leempoel et al., 2015; Wilson & Gallant, 2000).

Primary terrain attributes including slope, aspect as northness (AspC – cosine of aspect) and eastness (AspS – sine of aspect), and curvature are often used in species distribution models at local and regional scales, as they underlie multiple biophysical processes (such as water flow, snow movement, erosion and solar radiation (Guisan & Zimmermann, 2000; Lyon et al., 2008; Moore et al., 1991). Slope (Slo) describes the steepness of the terrain and aspect describes its orientation. Both are good proxies for solar radiation and temperature, especially when the surrounding relief is of minor importance (Leempoel et al., 2015). Horizontal curvature (Hcu) affects the convergence and the divergence of the water flow, while vertical curvature (Vcu) describes the acceleration and the deceleration of the water flow (Wilson & Gallant, 2000).

Over the last two decades, more elaborate secondary terrain attributes have been developed as proxies to model specific biophysical features. Vector ruggedness measure (VRM) describes the degree of terrain unevenness and distinguishes between rocky and smooth terrain (Sappington et al., 2007). This descriptor correlates with soil moisture and is a surrogate for stony soil in the western Alps (Leempoel et al., 2015). Positive topographic openness (TOP) and negative topographic openness (TON) express the protection of a focal point from the surrounding relief (Yokoyama et al., 2002). These factors are based on the maximum angle found at the zenith (TOP) or nadir (TON) from the point over a defined radius. By describing the terrain convexities and concavities, these two factors reflect soil drainage as well as protection from wind (Doneus, 2013).

Other factors have been designed to specifically model hydrological processes. The downslope distance gradient (DDG) quantifies the influence of slope on local water drainage (Hjerdt et al., 2004). The topographic wetness index (TWI) quantifies the topographic control of hydrological processes and is formulated as the logarithm of the ratio between the catchment area and the tangent of slope. It has been used to predict soil pH, snow cover and ambient humidity (Beven & Kirkby, 1979; Böhner & Selige, 2006), where negative values were correlated with mean and minimum air temperature, while positive values were correlated with the daily range of air temperature (Leempoel et al., 2015).

Light availability and temperature can be described by secondary terrain attributes that account for surrounding terrain. The sky view factor (SVF) expresses the ratio of the radiation received by a planar surface to the radiation emitted by the entire hemispheric environment (Böhner & AntoniĆ, 2009), such that SVF proxies for water and light availability. At a microclimatic level, SVF also describes wind protection. Finally, total insolation (Ti6, direct and diffuse radiation) and direct solar insolation in the growing month of June (Di6, direct radiation) account for sun angle, surrounding relief and the SVF. DEM-derived solar radiation is a direct proxy for light availability affecting photosynthesis rates and indirectly accounts for air and soil temperature, influencing evapotranspiration, snow melt and soil moisture (Tovar-Pescador et al., 2006).

As environmental factors often covary, we applied Spearman correlation tests to pairs of factors and retained only one factor of the pair if Spearman's $\rho > |0.8|$. Finally, we performed a principal component analysis (PCA) on the retained 12 standardised environmental factors at the locations of the 304 sampled individuals, using the 'prcomp' function in the *stats* package in R 4.1.0 (R Development Core Team, 2021).

Genomic data

We used genomic data described elsewhere (Rogivue et al., 2019a, b), representing SNPs and polymorphic TEs (presence/absence). In short, we sequenced each individual with Illumina HiSeq2500 (ATLAS Biolabs GmbH, Berlin, Germany; 125-bp paired-end reads). After filtering and mapping (BWA 0.7.12; Li & Durbin, 2010) to the high-quality, chromosome-level reference genome (Jiao et al., 2017), the remaining reads corresponded to an average coverage of 11.7x. We performed SNP calling with FreeBayes 1.0.2 (Garrison & Marth, 2012), with 291,396 biallelic SNPs remaining at the regional extent (across all four sites) after stringent filtering (excluding SNPs within TE sequences; Rogivue et al., 2019b). At the local extent, the number of SNPs retained were 220,214 in Essets, 113,900 in Martinets, 287,261 in Para and 160,322 in Pierredar (minor allele frequency < 0.025). To describe the spatial genetic structure of the sampled individuals, we carried out an analysis of model-based ancestry,

implemented in the program Admixture (Alexander et al., 2009), on the retained SNPs at the regional level ($N = 291,396$) for $K = 1-20$.

For the TEs, we genotyped presence/absence of polymorphic copies of long terminal repeat retrotransposons (LTR-RTs) using TEPIID (Stuart et al., 2016) and inferred their zygosity from read coverage at breakpoints to treat them similarly to SNPs. Of the 20,548 identified polymorphic TEs (Rogivue et al., 2019b), 3,874 loci with no missing data were retained for subsequent environmental association analyses and outlier detection.

Genotype–environment association analyses (GEA)

We performed GEA analyses (Lasky et al., 2022; Rellstab et al., 2015) at two hierarchical levels (Figure 1): (i) the regional level, i.e., among the four studied populations; and (ii) the local catchment level, i.e., among plots within each population. We applied latent factor mixed models (LFMM; Frichot et al., 2013) implemented in the R package *LFMM2* (Caye et al., 2019), a univariate GEA method in which neutral genetic structure is intrinsically considered.

We used the ‘snmf’ function in the R package *LEA* to replace the missing observations (10%) in the SNP dataset with imputed data based on the population structure from the sNMF analysis ($K = 1-10$, entropy = T, and 10 repetitions). We then imputed the data with the ‘impute’ function in *LEA*, where latent factors (K) were obtained using the lowest genomic inflation factor (gif) value (method = ‘mode’). Once the SNP datasets were complete, we ran the *LEA* function ‘lfmm_ridge’ for latent factors $K = 1-8$ for each environmental factor.

We defined the best number of latent factors (K) as the value with the genomic inflation factor (λ , averaged over the 12 environmental factors) closest to 1 (Table S3), as suggested by Caye et al. (2019). We did not choose a specific K for each environmental factor, as K has no specific biological meaning. Therefore, for the SNP dataset at the regional level (i) we used $K = 6$, and at the local level (ii) we used $K = 2$ for Essets, $K = 3$ for Martinets, $K = 6$ for Para, and $K = 2$ for Pierredar. For the TE dataset, we used $K = 1$ for all analyses. Following Rogivue et al. (2019b), we treated the two marker types (SNPs and polymorphic TEs) with different K values, reflecting

the differential patterns of variation observed among populations (e.g., 6% of SNPs were shared among populations, compared with 90% of polymorphic TEs; SNPs had higher pairwise F_{ST} values (0.09–0.18) than those for polymorphic TEs (0.03–0.09)). We ran the 'lfmm_test' function in *LFMM2* (Caye et al., 2019) to calculate the z scores and p values based on the λ and χ^2 distributions. To control for false positives, we applied the Benjamini-Hochberg algorithm (Benjamini & Hochberg, 1995) with a false discovery rate (FDR) of 1%.

Outlier analysis

To complement the GEA analyses, we used BayeScan 2.1 (Foll & Gaggiotti, 2008) to identify outlier SNPs and polymorphic TEs putatively under divergent selection, i.e., showing genetic differentiation above that expected for neutral markers. BayeScan is an F_{ST} -based approach that considers differences in allele frequencies between the common pool and each population. Outlier loci are those resulting from divergent selection beyond the level of what demographic processes, such as random genetic drift, may impose in terms of genetic differentiation.

Like with LFMM, we ran BayeScan (i) at the regional level among the four populations, and (ii) at the local level among plots within each local population. We converted the SNP and TE datasets from LFMM to BayeScan format using a custom R script. We defined the prior odd at 100, with 5,000 outputted iterations, a thinning interval size of 10, 20 pilot runs of length 5,000, and a burn-in of 50,000. We visualised BayeScan results using Manhattan plots created with the R package *qqman* (Turner, 2018).

GO term enrichment analyses among levels of spatial extent, statistical methods, and marker types

We characterised the putative functions of candidate loci identified in the above analyses using the same annotation as in Rogivue et al. (2019b). For simplicity, we considered only the SNPs with a high impact on their respective protein based on SnpEff (Cingolani et al., 2012), where high-impact SNPs were those identified as missense, splice acceptor and splice donor variants, as well as start/stop codon loss

or gain. For similar reasons, we retained only polymorphic TEs located within 2 kb upstream or downstream of annotated genes as possibly high-impact due to either gene disruption or spreading of epigenetic marks lowering gene expression (Hollister & Gaut, 2009).

We performed a gene ontology (GO) term enrichment analysis of the significantly associated SNPs and polymorphic TEs with topGO 2.44.0 (Alexa et al., 2006). We established the significance of the terms according to Fisher's exact tests (p value cut off at 0.01 and not adjusted for multiple testing, as recommended by the topGO authors).

We compared the annotated genes and enriched GO terms between hierarchical levels (regional vs. local extents), statistical methods (GEA vs. outlier analyses), and marker types (SNPs vs. polymorphic TEs) to identify analogous signals of adaptation. We searched for common high-impact SNPs and clusters (see below) among local populations and between levels of spatial extent, which we visualised using upset plots created with the R package *UpSetR* (Conway et al., 2017).

Clusters of loci identified as adaptive

We further compared common clusters of high-impact SNPs across the genome by defining 6228 blocks of 100 kbp, with a step of 50 kbp. We counted the number of high-impact SNPs, identified using either LFMM or BayeScan, per block and per region, where a cluster was assigned if a block or region contained at least two significant SNPs. We note that we summed all the LFMM-derived high-impact SNPs, irrespective of their associated environmental factors.

Results

Spatial genetic structure

The 291,396 SNPs retained at the regional level were used to determine the genetic structure across the studied populations. Using the program *Admixture* with $K = 1-20$, we found substantial genetic differentiation among populations despite the common genomic background reported in a phylogeographic study across the Alps (Rogivue et al., 2018). *Admixture* converged to a minimum cross-validation for 15 clusters, to

which 60% of individuals could be unambiguously assigned (membership coefficient > 0.8), whereas admixed individuals were typically found in a single population. All clusters were well nested within the four populations (3 in Essets, 4 in Martinets, 4 in Para, 4 in Pierredar) and contained $< 4\%$ of the variants originating from other populations (Figure 1).

Environmental factors and genotype–environment association (GEA) analysis

We evaluated 13 environmental factors derived from LiDAR-acquired DEMs at 0.5 m resolution. This spatial resolution matches our sampling grain, i.e., individuals' sites within plots (minimum distance between sampled plants > 1 m). Among these environmental factors, direct insolation in June was highly positively correlated with total insolation in June ($\rho = 0.99$; Table S4), so we discarded it to reduce collinearity between explanatory variables in subsequent analyses. A principal component analysis (PCA) on the retained 12 environmental factors indicated that the environmental conditions across the four sampled populations were similar overall, with the habitat conditions of sampled individuals spanning similar gradients along the environmental factors considered (Figure S1). This finding substantiates the reasoning that our sampling scheme provides optimal premises for investigating convergent signatures of adaptation across populations (Rellstab et al., 2020) based on genomic and environmental variation.

In GEA analyses (Figure S2), horizontal curvature was most prominently associated with genomic variation across all analyses, while eastness (with only a few significant associations) and slope (with none) seemed to be uninformative regarding signatures of selection (Table 1, Table S5). Notably, the environmental factors with the largest number of significant associations differed substantially among the local populations (Table S5).

The specificity of the environmental descriptors in driving local adaptation was assessed through patterns of interaction between the markers and the environmental factors. The majority of loci were associated with only one environmental factor, while up to 32.8% showed an association with multiple environmental factors in either the regional or local analyses (Figure 3a–e). The

spatial extent considered in the GEA analyses had a considerable effect on the number and relative frequency of significant associations per environmental factor; notably, each local population showed a distinct frequency distribution (Table 1, Tables S5 and S6). Accordingly, only 260 high-impact SNPs and polymorphic TEs were found to be significant between the regional and the four local analyses (Figure 2f). These were located in 209 candidate adaptive genes shared among the two levels of spatial extent (Figure S3a). Only one gene was identified in all five GEA analyses. However, no ortholog of this gene is known from *A. thaliana* and, hence, its function remains unknown.

For the two marker types, SNPs and polymorphic TEs, we found similar percentages of loci that were significantly associated with at least one environmental factor at the regional level: 2.46% of the SNPs and 2.48% of the polymorphic TEs. Despite such broad variation, only one candidate gene (Aa_G559430), homologous to the terpenoid cyclase AT3G14490 in *A. thaliana*, was highlighted by both SNPs and polymorphic TEs.

Outlier analysis

At the regional level, BayeScan identified only two outlier SNPs, of which only one was identified as high-impact. Conversley, between 15 (Essets) and 134 (Martinets) high-impact SNPs significantly exceeded the neutral expectation at the local level (q values < 0.1 ; Figure 4, Figure S4, Table 2, Table S7), though only one SNP was detected across all local populations. Most of the outlier SNPs identified were tightly flanked by other SNPs in linkage disequilibrium, which is known to extend across a considerable physical distance (Rogivue et al., 2019b). Conversley, outlier analyses for polymorphic TEs showed an inverted trend: few significant outliers were identified at the local level (between zero in Essets and three in Martinets), while 108 significant outliers were detected across the regional level (Table 2 and Figure S5). Comparing the two statistical methods, GEA and outlier analyses, found neither common SNPs nor genes at the regional level, with a few SNPs detected with both analyses at the level of local populations (Table 2).

GO term enrichment

To substantiate the functional relevance of the loci putatively involved in divergent selection, we performed tests for gene ontology (GO) term enrichment. We followed the advice of the authors of topGO and did not apply any correction for multiple testing to avoid excessive false negatives, at the cost of obtaining false positives. Accordingly, the results of the enrichment analyses point towards, but do not necessarily substantiate, genes of interest for being involved in adaptive processes. Of the 7,180 associated SNPs identified by GEA at the regional level, we found 653 high-impact SNPs across all 12 environmental factors occurring within 519 annotated genes (Table 1, Table S8), highlighting 42 significantly enriched GO terms (Table S9). In turn, the 15 polymorphic TEs occurring within 2 kb upstream or downstream of annotated genes highlighted 11 significantly enriched GO terms (Table 1, Tables S9 and S10). At the local level, high-impact SNPs and genes highlighted 22 significantly enriched GO terms that were shared among three or four local populations (Figure S3b). No gene flanking an associated polymorphic TE was identified in Essets or Pierredar, whereas one was detected in Martinets and 62 in Para (Table S5). GO term enrichment analyses were not considered for polymorphic TEs, as only a few linked genes were identified.

At the regional-level, GEA analyses identified only two enriched GO terms based on more than five genes comprising of at least one high-impact SNP, and these were therefore considered top candidates (Table 1, highlighted in Table S9): ‘defence response’ (GO:0006952) and ‘innate immune response’ (GO:0045087). The remaining 40 SNP-based significant GO terms and all 11 TE-based ones indicated a weaker signal of selection (Table S9). Of the two outliers detected at the regional level, the SNP at position 145,431 on chromosome 7 is a missense variant located in the gene Aa_G219310 (orthologous to AT4G38420 in *A. thaliana*), encoding a ‘copper ion binding’ protein involved in the oxidation–reduction process (GO:0055114, <http://atgenie.org/transcript?id=AT4G38420.1>). The outlier SNP located on chromosome 8 is a synonymous variant situated in the gene Aa_G499380 (orthologous to AT5G16970 in *A. thaliana*; Table S11) involved in responses to oxidative stress; however, as a synonymous SNP it is likely not the locus directly

under selection, but may be linked to the causative variation (i.e., not captured possibly for statistical reasons). No GO term analysis was possible because only one high-impact SNP was annotated (Table 2).

At the local level, few GO terms were classified as top candidates: two in Essets ('defence response' (GO:0006952) and 'innate immune response' (GO:0045087)), one in Martinets ('protein phosphorylation' (GO:0006468)), three in Para (three times 'protein phosphorylation' (GO:0006468)), and none in Pierredar (Tables S5a and S9). Among the 134 identified high-impact outlier SNPs in Martinets (Table S11), we found two top candidate GO terms: 'response to oomycetes' (GO:0002239) and 'defence response to fungus, incompatible interaction' (GO:0009817); Table S12). The three other regions presented no top candidate GO terms (Table S12), despite the occurrence of several high-impact outliers (Table S11).

Common clusters of adaptive loci

We identified 12 common blocks (genomic regions spanning 100 kbp, with a step of 50 kbp) of high-impact SNPs detected by GEA analyses, most of which were shared between at least two local populations (Figure 5). Only one genomic block was common among three populations, and two (consecutive) genomic blocks were shared between all four populations. Based on the outlier loci identified with BayeScan, we found three common blocks, two of which were consecutive on chromosome 3, which were common to the same two populations (Essets and Pierredar) and in common with one population based on LFMM (Pierredar). The remaining block was shared between Para and Pierredar. These 15 common blocks comprised 211 significant associations in GEA analyses for 94 SNPs.

Discussion

Investigating the genes and functions that are under selection and identifying the factors driving selection are essential for understanding local adaptation, particularly in view of global change. The high-quality and high-resolution genomic and environmental data now available provide increasing insights into the evolution of species and how environmental conditions can shape evolution. By assessing genomic

variation at the whole-genome level at both single nucleotides (SNPs) and TE markers for 304 individuals, and by characterising each individual's micro-habitat, we found that environmental factors describing soil water availability and solar radiation are important drivers of local adaptation in *A. alpina*. While we detected loci showing signals of local adaptation in the genic (high-impact SNPs and polymorphic TEs) and non-genic (polymorphic TEs near to genes) genome regions, the detected signals differed substantially among the four local populations, with inconsistencies across the regional and local spatial extents. These findings highlight the specificity of local adaptation for populations, despite these populations sharing a common genomic background and experiencing similar environmental conditions (Fig. S1; Rogivue et al., 2018). Such truly local responses to abiotic environmental cues may reflect in part the high frequency of selfing observed in Alpine populations of *A. alpina* (Buehler et al., 2012; Tedder et al., 2011), which could reinforce genomic imprints of local adaptation once they have become established (Trickovic & Glémin, 2022). This finding provides evidence that convergent signatures of selection, even within closely situated populations of a given species, may only be expected under particularly strong selective pressure and, thus, are rarely realised in populations of mostly selfing individuals, as investigated here.

Singular signals of genomic adaptation across levels of spatial extent

In our hierarchical study design, we identified few gene candidates for local adaptation common between the regional and local analyses (Figure 3f). While the literature about the spatial extent of adaptation is still limited, two recent studies on forest trees also showed that adaptation at different levels of spatial extent appears complementary (Brousseau et al., 2021; von Takach et al., 2021). These authors concluded that such scale issues are important to consider for understanding genomic signatures of local adaptation, which is in line with our findings. The very local adaptive response of populations reported here was already highlighted by Rellstab et al. (2017) in a study on the strictly outcrossing *Arabidopsis halleri*. There, only 31% of previously identified candidate adaptive SNPs were confirmed in an

independent set of populations, underlining the truly local nature of adaptation in response to a highly heterogeneous environment.

The low degree of overlapping results between the two marker types, SNPs and polymorphic TEs, implies that their respective variation reflects adaptive responses involving complementary candidate genes and gene functions. We cannot exclude potential effects of using different bioinformatic tools to characterize each marker type, stemming from fundamental differences between the types of polymorphisms. For example, SNP-derived estimates of population-specific inbreeding coefficients F_{IS} were lower than respective values calculated from polymorphic TEs (Rogivue et al., 2019b). By ignoring sequence variation, we underestimate heterozygosity in TEs, resulting in inflated F_{IS} values. Nevertheless, our findings highlight the relevance of incorporating structural polymorphisms, such as TEs and copy number variants, when studying complex evolutionary responses to changing environments. Similar findings have been reported previously for Brassicaceae (Niu et al., 2019; Quadrana et al., 2016), emphasising the need to consider TE variation when investigating adaptation to global warming (Rey et al., 2016). The present study is, to our knowledge, the first to show that SNPs within genes and polymorphic TE copies indicate complementary candidates of local adaptation.

Unsurprisingly, the two statistical methods used (GEA and outlier detection) revealed complementary outcomes and did not point to the same candidate genes. To some degree, this inconsistency may be attributed to the smaller number of candidates found in the outlier analysis (Table 2). Such a result may be associated with sampling individuals in plots of contrasting environments, rather than along abiotic gradients (Richardson et al., 2014). Additionally, the fundamental differences between the two methodological approaches (statistical, environment-related GEA, population genomic outlier detection) suggest that one should expect complementary, rather than overlapping, signals of adaptation. Our results indeed integrate effects of selection across various abiotic and (unaccounted for) biotic factors. Accordingly, conditional neutrality (Mee & Yeaman, 2019) might at least partly explain the small number of outliers detected that contrasts with the numerous loci associated with environmental heterogeneity. Relatively high genetic

load was indeed detected within those expending alpine populations (Zeitler et al., 2023), and corresponding loci may be conditionally deleterious to a large extent, hence showing limited allelic differentiation despite significant genotype-by-environment interactions. As simulations show that GEA analyses have a high power of detection under scenarios of weak selection and low dispersal (Forester et al., 2016), such approaches may be sensitive enough to identify conditionally neutral loci (Lasky et al., 2022; Yoder & Tiffin, 2018).

Selective factors are distinct among SNPs and polymorphic TEs, and are predominantly consistent between levels of spatial extent

In our study, the type of genomic marker investigated considerably affected the relative importance of the associated environmental factors (Table 1, Figure 3). Horizontal curvature, a proxy for drainage and hydrological processes, detected the most candidate SNPs. Conversely, GEAs using polymorphic TEs highlighted that the factor with the largest number of significant candidates was total insolation in June (or the highly correlated direct insolation in June), a factor representing solar radiation. Although mechanistic underpinnings remain out of scope, this finding again highlights the complementarity of SNPs and polymorphic TEs in revealing responses to different environmental cues.

Across the two hierarchical levels of spatial extent, the order of importance of the environmental factors differed slightly, while the top environmental factors largely remained consistent (Table 1 and Table S5). This result suggests that the studied populations inhabit similar environments and that our study design was appropriate for detecting potentially convergent signatures of adaptation (Figure S1). Nevertheless, environmental contrasts such as (indirectly accounted for) differences in biotic interactions as well as restricted gene flow, and hence large effects of genetic drift, likely contributed to overruling selective effects to impose a common signal in the genomes of *A. alpina* individuals.

Defence-related functions are key in heterogeneous alpine environments

In alpine environments, plants must deal with harsh, highly variable conditions to survive. In our study, high-impact polymorphisms in genes related to defence responses are overrepresented among candidates of local adaptation (Table S9). Likewise, loci identified by the outlier analyses also highlighted defence-related functions under selection, though outlier loci did not directly link to any environmental drivers. Biotic interactions are usually considered to play a minor role when compared to abiotic environmental factors for plants at high elevations (Zvereva & Kozlov, 2022). Nevertheless, despite our focus on potential abiotic drivers of local adaptation, our study indicated that genes with biotic-related biological functions are key in adaptive responses in alpine plants. Therefore, we advocate also including biotic drivers in future environmental association analyses to better represent potential environmental drivers of adaptation.

The two biological functions under selection at the regional level, i.e., ‘defence response’ and ‘innate immune response’, were also found at the local level in Essets. No other common biological function was found among the two levels of spatial extent, again stressing the notion that local adaptation is truly local. Moreover, adaptive functions previously described in studies on *A. alpina*, e.g., related to survival, flowering and fecundity (de Villemereuil et al., 2018; Toräng et al., 2015; Wötzel et al., 2022), were not found in our analyses. This discrepancy might be due to differences in the spatial extent and resolution of the studies, the complementary methodological approaches, and the different types of environmental factors considered. Beyond this, high levels of selfing encountered in Alpine populations of *A. alpina* may amplify population-specific adaptive responses, though such an effect depends on various conditions such as patterns of dominance, levels of selective pressure, and migration rate (Trickovic & Glémin, 2022).

In conclusion, by combining whole-genome re-sequencing data with high-resolution descriptions of environmental factors, our study pointed to several genes and biological functions potentially under selection. These findings pave the ground for a better understanding of adaptive processes and could guide further investigations of *A. alpina* and related (alpine) species. Further, our findings suggest

that signatures of local adaptation are dependent on the spatial extent of the study, the type of genomic markers employed, and the statistical methods used. The common genomic blocks indicative of imprints of adaptation (Figure 5) refer to regions in the genome that are under sufficiently strong selection to overcome the presumed high level of genetic drift induced by demography and life-history traits of *A. alpina* in our study regions. Taken together, our results suggest that microevolutionary processes may likely result in non-overlapping signatures of local adaptation across spatial scales.

Acknowledgements

We thank Jessica Joaquim and Lolita Ammann for help with the sampling, Nicolas Delley and Patrick Thee for introducing to the GPS and GNSS, and Bertrand Cannelle, Stéphane Cretegny, Ivo Widmer and Christophe Randin for their technical advice. Stefan Zoller (Genetic Diversity Centre ETH Zurich) provided bioinformatic support, and Christian Rellstab, Benjamin Dauphin, Katalin Csilléry, Rolf Holderegger and Christian Sailer contributed valuable discussions and critical reviews, and Melissa Dawes helped with language editing of an earlier version of the manuscript. Financial support for AR, KL, RRC, SC and CP was provided through a Swiss National Science Foundation grant (GeneScale, CR32I3_149741) awarded to FG, FF, MK and SJ, and for RRC through the Bundesbehörden der Schweizerischen Eidgenossenschaft, (grant number: 2014.0821).

References

- Alexa, A., Rahnenführer, J., & Lengauer, T. (2006). Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. *Bioinformatics*, 22(13), 1600–1607. doi:10.1093/bioinformatics/btl140
- Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, 19(9), 1655–1664. doi:10.1101/gr.094052.109
- Ansell, S. W., Grundmann, M., Russell, S. J., Schneider, H., & Vogel, J. C. (2008). Genetic discontinuity, breeding-system change and population history of *Arabis alpina* in the Italian Peninsula and adjacent Alps. *Molecular Ecology*, 17(9), 2245–2257. doi:10.1111/j.1365-294X.2008.03739.x
- Baduel, P., Leduque, B., Ignace, A., Gy, I., Gil, J., Loudet, O., Colot, V., & Quadrana, L. (2021). Genetic and environmental modulation of transposition shapes the evolutionary potential of *Arabidopsis thaliana*. *Genome Biology*, 22, 26. doi:10.1186/s13059-021-02348-5
- Benjamini, Y., & Hochberg, Y. (1995). Controlling for the false discovery rate — a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B – Statistical Methodology*, 57(1), 289–300. doi:10.1111/j.2517-6161.1995.tb02031.x
- Bennetzen, J. L., & Wang, H. (2014). The contributions of transposable elements to the structure, function, and evolution of plant genomes. *Annual Review of Plant Biology*, 65, 505–530. doi:10.1146/annurev-arplant-050213-035811
- Beven, K., & Kirkby, M. J. (1979). A physically based, variable contributing area model of basin hydrology. *Hydrological Sciences*, 24(1), 43–69. doi:10.1080/02626667909491834
- Böhner, J., & AntoniĆ, O. (2009). Land-surface parameters specific to topo-climatology. In T. Hengl & H. I. Reuter (Eds.), *Geomorphometry: Concepts, Software, Applications* (pp. 195–126): Elsevier.
- Böhner, J., & Selige, T. (2006). Spatial prediction of soil attributes using terrain analysis and climate regionalisation. *Göttinger Geographische Abhandlungen*, 115, 13–28.
- Booker, T. R., Yeaman, S., & Whitlock, M. C. (2021). Global adaptation complicates the interpretation of genome scans for local adaptation. *Evolution Letters*, 5(1), 4–15. doi:10.1002/evl3.208
- Brousseau, L., Fine, P. V. A., Dreyer, E., Vendramin, G. G., & Scotti, I. (2021). Genomic and phenotypic divergence unveil microgeographic adaptation in the Amazonian hyperdominant tree *Eperua falcata* Aubl. (Fabaceae). *Molecular Ecology*, 30, 1136–1154. doi:10.1111/mec.15595
- Buckley, J., Widmer, A., Mescher, M. C., & De Moraes, C. M. (2019). Variation in growth and defence traits among plant populations at different elevations: Implications for adaptation to climate change. *Journal of Ecology*, 107(5), 2478–2492. doi:10.1111/1365-2745.13171
- Buehler, D., Graf, R., Holderegger, R., & Gugerli, F. (2012). Contemporary gene flow and mating system of *Arabis alpina* in a Central European alpine landscape. *Annals of Botany*, 109(7), 1359–1367. doi:10.1093/aob/mcs066
- Caye, K., Jumentier, B., Lepeule, J., & François, O. (2019). LFMM 2.0: Latent factor models for confounder adjustment in genome and epigenome-wide association studies. *Molecular Biology and Evolution*, 36(4), 852–860. doi:10.1093/molbev/msz008

- Chopra, D., Mapar, M., Stephan, L., Albani, M. C., Deneer, A., Coupland, G., Willing, E.-M., Schellmann, S., Schneeberger, K., Fleck, C., Schrader, A., & Hülkamp, M. (2019). Genetic and molecular analysis of trichome development in *Arabis alpina*. *Proceedings of the National Academy of Sciences of the United States of America*, 116(Genetic and molecular analysis of trichome development in *Arabis alpina*), 12078–12083. doi:10.1073/pnas.1819440116
- Choudhury, R. A., Rogivue, A., Gugerli, F., & Parisod, C. (2019). Impact of polymorphic transposable elements on linkage disequilibrium along chromosomes. *Molecular Ecology*, 28(6), 1550–1562. doi:10.1111/mec.15014
- Cingolani, P., Platts, A., Wang, L. L., Coon, M., Nguyen, T., Wang, L., Land, S. J., Lu, X., & Ruden, D. M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w(1118); iso-2; iso-3. *Fly*, 6(2), 80–92. doi:10.4161/fly.19695
- Conrad, O., Bechtel, B., Bock, M., Dietrich, H., Fischer, E., Gerlitz, L., Wehberg, J., Wichmann, V., & Böhner, J. (2015). System for automated geoscientific analyses (SAGA) v. 2.1.4. *Geoscientific Model Development*, 8(7), 1991–2007. doi:10.5194/gmd-8-1991-2015
- Conway, J. R., Lex, A., & Gehlenborg, N. (2017). UpSetR: an R package for the visualization of intersecting sets and their properties. *Bioinformatics*, 33(18), 2938–2940. doi:10.1093/bioinformatics/btx364
- Dauphin, B., Rellstab, C., Wüest, R. O., Karger, D. N., Holderegger, R., Gugerli, F., & Manel, S. (2023). Re-thinking the environment in landscape genomics. *Trends in Ecology & Evolution*, 38(3), 261–274. doi:10.1016/j.tree.2022.10.010
- de Villemereuil, P., Mouterde, M., Gaggiotti, O. E., & Till-Bottraud, I. (2018). Patterns of phenotypic plasticity and local adaptation in the wide elevation range of the alpine plant *Arabis alpina*. *Journal of Ecology*, 106(5), 1952–1971. doi:10.1111/1365-2745.12955
- Doneus, M. (2013). Openness as visualization technique for interpretative mapping of airborne lidar derived digital terrain models. *Remote Sensing*, 5(12), 6427–6442. doi:10.3390/rs5126427
- Drake, J. M., Randin, C. F., & Guisan, A. (2006). Modelling ecological niches with support vector machines. *Journal of Applied Ecology*, 43(3), 424–432. doi:10.1111/j.1365-2664.2006.01141.x
- Eckert, A. J., Maloney, P. E., Vogler, D. R., Jensen, C. E., Delfino Mix, A., & Neale, D. B. (2015). Local adaptation at fine spatial scales: an example from sugar pine (*Pinus lambertiana*, Pinaceae) *Tree Genetics & Genomes*, 11(3), 42. doi:10.1007/s11295-015-0863-0
- Exposito-Alonso, M., Vasseur, F., Ding, W., Burbano, H. A. A., & Weigel, D. (2018). Genomic basis and evolutionary potential for extreme drought adaptation in *Arabidopsis thaliana*. *Nature Ecology & Evolution*, 2, 352–358. doi:10.1038/s41559-017-0423-0
- Fischer, M. C., Rellstab, C., Tedder, A., Zoller, S., Gugerli, F., Shimizu, K. K., Holderegger, R., & Widmer, A. (2013). Population genomic footprints of selection and associations with climate in natural populations of *Arabidopsis halleri* from the Alps. *Molecular Ecology*, 22(22), 5594–5607. doi:10.1111/mec.12521
- Foll, M., & Gaggiotti, O. (2008). A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. *Genetics*, 180, 977–993.

- Forester, B. R., Jones, M. R., Joost, S., Landguth, E. L., & Lasky, J. R. (2016). Detecting spatial genetic signatures of local adaptation in heterogeneous landscapes. *Molecular Ecology*, 25(1), 104–120. doi:10.1111/mec.13476
- Fournier-Level, A., Korte, A., Cooper, M. D., Nordborg, M., Schmitt, J., & Wilczek, A. M. (2011). A map of local adaptation in *Arabidopsis thaliana*. *Science*, 334(6052), 86–89. doi:10.1126/science.1209271
- Frichot, E., Schoville, S. D., Bouchard, G., & François, O. (2013). Testing for associations between loci and environmental gradients using latent factor mixed models. *Molecular Biology and Evolution*, 30(7), 1687–1699.
- Garrison, E., & Marth, G. (2012). Haplotype-based variant detection from short-read sequencing. *arXiv*.
- Gougherty, A. V., Keller, S. R., & Fitzpatrick, M. C. (2021). Maladaptation, migration and extirpation fuel climate change risk in a forest tree species. *Nature Climate Change*, 11, 166–171. doi:10.1038/s41558-020-00968-6
- Guillaume, A. S., Leempoel, K., Rogivue, A., Rochat, E., Kasser, M., Gugerli, F., Parisod, C., & Joost, S. (2021). Multiscale very-high resolution topographic models in Alpine ecology: pros and cons of airborne LiDAR and drone-based stereo-photogrammetry technologies. *Remote Sensing*, 13. doi:10.3390/rs13081588
- Guisan, A., & Zimmermann, N. (2000). Predictive habitat distribution models in ecology. *Ecological Modelling*, 135(2–3), 147–186.
- Hancock, A. M., Brachi, B., Faure, N., Horton, M. W., Jarymowycz, L. B., Sperone, F. G., Toomajian, C., Roux, F., & Bergelson, J. (2011). Adaptation to climate across the *Arabidopsis thaliana* genome. *Science*, 334(6052), 83–86.
- Hjerdt, K., McDonnell, J., Seibert, J., & Rodhe, A. (2004). A new topographic index to quantify downslope controls on local drainage. *Water Resources Research*, 40(5), 199–207. doi:10.1029/2004WR003130
- Hollister, J. D., & Gaut, B. S. (2009). Epigenetic silencing of transposable elements: A trade-off between reduced transposition and deleterious effects on neighboring gene expression. *Genome Research*, 19(1419–1428). doi:10.1101/gr.091678.109
- Jiao, W.-B., Garcia Accinelli, G., Hartwig, B., Kiefer, C., Baker, D., Severing, E., Willing, E.-M., Piednoël, M., Woetzel, S., Madrid-Herrero, E., Huettel, B., Hürmann, U., Reinhard, R., Koch, M. A., Swan, D., Clavijo, B., Coupland, G., & Schneeberger, K. (2017). Improving and correcting the contiguity of long-read genome assemblies of three plant species using optical mapping and chromosome conformation capture data. *Genome Research*, 27, 778–786. doi:10.1101/gr.213652.116
- Kalendar, R., Tanskanen, J., Immonen, S., Nevo, E., & Schulman, A. H. (2000). Genome evolution of wild barley (*Hordeum spontaneum*) by BARE-1 retrotransposon dynamics in response to sharp microclimatic divergence. *Proceedings of the National Academy of Sciences of the United States of America*, 97(12), 6603–6607. doi:10.1073/pnas.110587497
- Kawecki, T. J., & Ebert, D. (2004). Conceptual issues in local adaptation. *Ecology Letters*, 7(12), 1225–1241. doi:10.1111/j.1461-0248.2004.00684.x
- Kozak, K. H., Graham, C. H., & Wiens, J. J. (2008). Integrating GIS-based environmental data into evolutionary biology. *Trends in Ecology & Evolution*, 23(3), 141–148.
- Laenen, B., Tedder, A., Nowak, M. D., Toräng, P., Wunder, J., Wötzel, S., Steige, K. A., Kourmpetisd, Y., Odong, T., Drouzas, A. D., Bink, M. C. A. M., Ågren, J., Coupland, G., & Slotte, T. (2018). Demography and mating system shape the genomewide impact of purifying selection in *Arabis alpina*. *Proceedings of the*

- National Academy of Sciences of the United States of America, 115(4), 816–821.
doi:10.1073/pnas.1707492115
- Lasky, J. R., Josephs, E. B., & Morris, G. P. (2022). Genotype–environment associations to reveal the molecular basis of environmental adaptation. *The Plant Cell*, 2023(35), 1. doi:10.1093/plcell/koac267
- Lecours, V., Devillers, R., Simms, A. E., Lucieer, V. L., & Brown, C. J. (2017). Towards a framework for terrain attribute selection in environmental studies. *Environmental Modeling & Software*, 89, 19–30. doi:10.1016/j.envsoft.2016.11.027
- Leempoel, K., Parisod, C., Geiser, C., Daprà, L., Vittoz, P., & Joost, S. (2015). Very high resolution digital elevation models: are multi-scale derived variables ecologically relevant? *Methods in Ecology and Evolution*, 6(12), 1373–1383. doi:10.1111/2041-210X.12427
- Leempoel, K., Parisod, C., Geiser, C., & Joost, S. (2018). Multiscale landscape genomic models to detect signatures of selection in the alpine plant *Biscutella laevigata*. *Ecology and Evolution*, 8, 1794–1806. doi:10.1002/ece3.3778
- Li, H., & Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics*, 26(5), 589–595. doi:10.1093/bioinformatics/btp698
- Lobréaux, S., & Miquel, C. (2020). Identification of *Arabis alpina* genomic regions associated with climatic variables along an elevation gradient through whole genome scan. *Genomics*, 112(1), 729–735. doi:10.1016/j.ygeno.2019.05.008
- Lotterhos, K. E., & Whitlock, M. C. (2015). The relative power of genome scans to detect local adaptation depends on sampling design and statistical method. *Molecular Ecology*, 24(5), 1031–1046. doi:10.1111/mec.13100
- Lyon, S. W., Troch, P. A., Broxton, P. D., Molotch, N. P., & Brooks, P. D. (2008). Monitoring the timing of snowmelt and the initiation of streamflow using a distributed network of temperature/light sensors. *Ecohydrology*, 1(3), 215–224. doi:10.1002/eco.18
- Mee, J. A., & Yeaman, S. (2019). Unpacking conditional neutrality: genomic signatures of selection on conditionally beneficial and conditionally deleterious mutations. *The American Naturalist*, 194(4), 529–540. doi:10.1086/702314
- Moore, I. D., Grayson, R., & Ladson, A. (1991). Digital terrain modelling: a review of hydrological, geomorphological, and biological applications. *Hydrological Processes*, 5(1), 3–30. doi:10.1002/hyp.3360050103
- Niu, X.-M., Xu, Y.-C., Li, Z.-W., Bian, Y.-T., Hou, X.-H., Chen, J.-F., Zou, Y.-P., Jiang, J., Wu, Q., Ge, S., Balasubramanian, S., & Guo, Y.-L. (2019). Transposable elements drive rapid phenotypic variation in *Capsella rubella*. *Proceedings of the National Academy of Sciences of the United States of America*, 116(14), 6908–6913. doi:10.1073/pnas.1811498116
- Pais, A. L., Whetten, R. W., & Xiang, Q.-Y. (2017). Ecological genomics of local adaptation in *Cornus florida* L. by genotyping by sequencing. *Ecology and Evolution*, 7(1), 451–465. doi:10.1002/ece3.2623
- Poncet, B., Herrmann, D., Gugerli, F., Taberlet, P., Holderegger, R., Gielly, L., Rioux, D., Thuiller, W., Aubert, S., & Manel, S. (2010). Tracking genes of ecological relevance using a genome scan: application to *Arabis alpina*. *Molecular Ecology*, 19(14), 2896–2907. doi:10.1111/j.1365-294X.2010.04696.x
- Pradervand, J.-N., Dubuis, A., Pellissier, L., Guisan, A., & Randin, C. F. (2014). Very high resolution environmental predictors in species distribution models: moving beyond topography? *Progress in Physical Geography*, 38, 79–96.

- Quadrana, L., Bortolini Silveira, A., Mayhew, G. F., LeBlanc, C., Martienssen, R. A., Jeddeloh, J. A., & Colot, V. (2016). The *Arabidopsis thaliana* mobilome and its impact at the species level. *eLIFE*, 5, 25. doi:10.7554/eLife.15716
- Quesneville, H. (2020). Twenty years of transposable element analysis in the *Arabidopsis thaliana* genome. *Mobile DNA*, 11, 13. doi:10.1186/s13100-020-00223-x
- R Development Core Team. (2021). R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. Retrieved from <http://www.R-project.org>
- Rellstab, C., Fischer, M. C., Zoller, S., Graf, R., Tedder, A., Shimizu, K. K., Widmer, A., Holderegger, R., & Gugerli, F. (2017). Local adaptation (mostly) remains local: reassessing environmental associations of climate-related candidate SNPs in *Arabidopsis halleri*. *Heredity*, 118(1), 193–201. doi:10.1038/hdy.2016.82
- Rellstab, C., Gugerli, F., Eckert, A. J., Hancock, A. M., & Holderegger, R. (2015). A practical guide to environmental association analysis in landscape genomics. *Molecular Ecology*, 24(17), 4348–4370. doi:10.1111/mec.13322
- Rellstab, C., Zoller, S., Sailer, C., Tedder, A., Gugerli, F., Shimizu, K. K., Holderegger, R., Widmer, A., & Fischer, M. C. (2020). Genomic signatures of convergent adaptation to Alpine environments in three Brassicaceae species. *Molecular Ecology*, 29(22), 4350–4365. doi:10.1111/MEC.15648
- Rey, O., Danchin, E., Mirouze, M., Loot, C., & Blanchet, S. (2016). Adaptation to global change: a transposable element–epigenetics perspective. *Trends in Ecology & Evolution*, 31(7), 514–526. doi:10.1016/j.tree.2016.03.013
- Richardson, J. L., Urban, M. C., Bolnick, D. I., & Skelly, D. K. (2014). Microgeographic adaptation and the spatial scale of evolution. *Trends in Ecology & Evolution*, 29(3), 165–176.
- Rogivue, A., Choudhury, R. R., Zoller, S., Joost, S., Felber, F., Kasser, M., Parisod, C., & Gugerli, F. (2019a). Data from: Genome-wide variation in nucleotides and retrotransposons in alpine populations of *Arabis alpina* (Brassicaceae). Retrieved from: <https://datadryad.org/resource/doi:10.5061/dryad.58g217k>
- Rogivue, A., Choudhury, R. R., Zoller, S., Joost, S., Felber, F., Kasser, M., Parisod, C., & Gugerli, F. (2019b). Genome-wide variation in nucleotides and retrotransposons in alpine populations of *Arabis alpina* (Brassicaceae). *Molecular Ecology Resources*, 19(3), 773–787. doi:10.1111/1755-0998.12991
- Rogivue, A., Graf, R., Parisod, C., Holderegger, R., & Gugerli, F. (2018). The phylogeographic structure of *Arabis alpina* in the Alps shows consistent patterns across different types of molecular markers and of geographic scales. *Alpine Botany*, 128(1), 35–45. doi:10.1007/s00035-017-0196-8
- Sappington, J. M., Longshore, K. M., & Thompson, D. B. (2007). Quantifying landscape ruggedness for animal habitat analysis: a case study using bighorn sheep in the Mojave Desert. *Journal of Wildlife Management*, 71, 1419–1426.
- Savolainen, O., Lascoux, M., & Merilä, J. (2013). Ecological genomics of local adaptation. *Nature Reviews Genetics*, 14, 807–820.
- Stuart, T., Eichten, S. R., Cahn, J., Karpievitch, Y. V., Borevitz, J. O., & Lister, R. (2016). Population scale mapping of transposable element diversity reveals links to gene regulation and epigenomic variation. *eLIFE*, 5, 27. doi:10.7554/eLife.2077
- Tedder, A., Ansell, S. W., Lao, X., Vogel, J. C., & Mable, B. K. (2011). Sporophytic self-incompatibility genes and mating system variation in *Arabis*. *Annals of Botany*, 108(4), 699–713. doi:10.1093/aob/mcr157

- Toräng, P., Wunder, J., Obeso, R. J., Herzog, M., Coupland, G., & Ågren, J. (2015). Large-scale adaptive differentiation in the alpine perennial herb *Arabis alpina*. *New Phytologist*, 206, 459–470. doi:10.1111/nph.13176
- Tovar-Pescador, J., Pozo-Vázquez, D., Ruiz-Arias, J., Batlles, J., López, G., & Bosch, J. (2006). On the use of the digital elevation model to estimate the solar radiation in areas of complex topography. *Meteorological Applications*, 13(3), 279–287. doi:10.1017/S1350482706002258
- Trickovic, B., & Glémin, S. (2022). Establishment of local adaptation in partly self-fertilizing population. *Genetics*, 220. doi:10.1093/genetics/iyab201
- Turner, S. D. (2018). qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. *Journal of Open Source Software*, 3, 2. doi:10.21105/joss.00731
- Turner, T. L., Bourne, E. C., Von Wettberg, E. J., Hu, T. T., & Nuzhdin, S. V. (2010). Population resequencing reveals local adaptation of *Arabidopsis lyrata* to serpentine soils. *Nature Genetics*, 42(3), 260–263.
- Vayssières, A., Mishra, P., Roggen, A., Neumann, U., Ljung, K., & Albani, M. C. (2020). Vernalization shapes shoot architecture and ensures the maintenance of dormant buds in the perennial *Arabis alpina*. *New Phytologist*, 227(1), 99–115. doi:10.1111/nph.16470
- von Takach, B., Ahrens, C. W., Lindenmayer, D. B., & Banks, S. C. (2021). Scale-dependent signatures of local adaptation in a foundation tree species. *Molecular Ecology*, 30(10), 2248–2261. doi:10.1111/mec.15894
- Wang, R., Farrona, S., Vincent, C., Joecker, A., Schoof, H., Turck, F., Alonso-Blanco, C., Coupland, G., & Albani, M. C. (2009). *PEP1* regulates perennial flowering in *Arabis alpina*. *Nature*, 459(7245), 423–427. doi:10.1038/nature07988
- Wilson, J. P., & Gallant, J. C. (Eds.). (2000). *Terrain analysis: principles and applications*. New York: Wiley.
- Wingler, A., Juvany, M., Cuthbert, C., & Munné-Bosch, S. (2015). Adaptation to altitude affects the senescence response to chilling in the perennial plant *Arabis alpina*. *Journal of Experimental Botany*, 66(1), 355–367. doi:10.1093/jxb/eru426
- Wötzel, S., Andrello, M., Albani, M. C., Koch, M. A., Coupland, G., & Gugerli, F. (2022). *Arabis alpina*: a perennial model plant for ecological genomics and life-history evolution. *Molecular Ecology Resources*, 22(2), 468–486. doi:10.1111/1755-0998.13490
- Yeaman, S., Hodgins, K. A., Lotterhos, K. E., Suren, H., Nadeau, S., Degner, J. C., Nurkowski, K. A., Smets, P., Wang, T., Gray, L. K., Liepe, K. J., Hamann, A., Holliday, J. A., Whitlock, M. C., Rieseberg, L. H., & Aitken, S. N. (2016). Convergent local adaptation to climate in distantly related conifers. *Science*, 353(6306), 1431–1433. doi:10.1126/science.aaf7812
- Yoder, J. B., Stanton-Geddes, J., Zhou, P., Briskine, R., Young, N. D., & Tiffin, P. (2014). Genomic signature of adaptation to climate in *Medicago truncatula*. *Genetics*, 196(4), 1263–1275. doi:10.1534/genetics.113.159319
- Yoder, J. B., & Tiffin, P. (2018). Effects of gene action, marker density, and timing of selection on the performance of landscape genomic scans of local adaptation. *Journal of Heredity*, 109(1), 16–28. doi:10.1093/jhered/esx042
- Yokoyama, R., Shirasawa, M., & Pike, R. J. (2002). Visualizing topography by openness: a new application of image processing to digital elevation models. *Photogrammetric Engineering and Remote Sensing*, 68(3), 257–266.

889 Zeitler, L., Parisod, C., & Gilbert, K. J. (2023). Purging due to self-fertilization does
890 not prevent accumulation of expansion load. *PLoS Genetics*, 19, e1010883.
891 doi:10.1371/journal.pgen.1010883
892 Zvereva, E. L., & Kozlov, M. V. (2022). Meta-analysis of elevational changes in the
893 intensity of trophic interactions: Similarities and dissimilarities with latitudinal
894 patterns. *Ecology Letters*, 25(9), 2076–2087. doi:10.1111/ele.14090
895

Data Accessibility and Benefit-Sharing

All data and custom-made codes are available in the DRYAD repository at <http://doi.org/....>, reference number

A Benefit-Sharing statement does not apply, given that all samples originated from the country where the research was conducted.

Author contributions

A.R., K.L., R.R.C., F.F., M.K., S.J., C.P. and F.G. initiated and conceived the project. AR performed the experiments. A.R., K.L., A.S.G. and R.R.C. analysed the data. A.R., K.L., R.R.C., A.S.G., M.K., S.J. and C.P. contributed analysis tools. A.R., K.L. and F.G. wrote the initial manuscript, with contributions from all co-authors.

Conflict of interest

The authors declare that they do not have any conflict of interest.

Tables and Figures

TABLE 1 Results of environmental association analyses in *Arabidopsis* at the regional level (latent factor mixed model, LFMM). The number (#) of significant single-nucleotide polymorphisms (SNPs), high-impact SNPs (i.e., missense, splice acceptor and splice donor variants, as well as start/stop codon loss or gain) and polymorphic transposable elements (TEs) within 2 kbp of a gene, as well as the number of associated genes and significantly enriched gene ontology (GO) terms associated with each environmental factor, with less (†) or more (‡) than five genes significant. The three environmental factors most frequently associated with SNPs and polymorphic TEs are highlighted in bold and italics, respectively.

Environmental factor	# associated loci			# associated genes		# significantly enriched GO terms †	
	SNPs	High-impact SNPs	TEs	From SNPs	From TEs	From SNPs	From TEs
<i>Total insolation in June</i>	0	0	42	0	9	0	5
Sky view factor	243	31	0	17	0	5	0
Topographic wetness index	52	1	0	1	0	0	0
Neg. topographic openness	151	6	0	6	0	0	0
Pos. topographic openness	82	21	13	9	4	7	6
Vector ruggedness measure	1767	179	3	148	0	5	0
Downslope distance gradient	224	16	0	14	0	5	0
<i>Horizontal curvature</i>	3355	286	36	236	0	8	0
Vertical curvature	924	78	0	62	0	6 + 2‡	0
Northness	382	35	0	26	0	4	0
Eastness	0	0	2	0	2	0	0
Slope	0	0	0	0	0	0	0
Total	7180	653	96	519	15	42	11

† $p < 0.01$, < 5 genes significant

‡ $p < 0.01$, > 5 genes significant

923 TABLE 2 Number of outlier loci detected by BayeScan in *Arabis alpina*. The number of significant single-nucleotide polymorphisms (SNPs) and
 924 polymorphic transposable elements (TEs) at the regional level and for each of the four local populations, the number of high-impact SNPs and
 925 TEs, the number of genes containing SNPs or TEs, and the number of significantly enriched gene ontology (GO) terms for SNPs and TEs, with
 926 less (†) or more (‡) than five genes significant.
 927

Spatial extent	Local population	# outliers (<i>q</i> value < 0.1)		# high-impact		# genes		# enriched GO terms †		# common with LFMM		# common gene with LFMM	
		SNPs	TEs	SNPs	TEs	SNPs	TEs	SNPs	TEs	SNPs	TEs	SNPs	TEs
Regional		2	108	1	2	1	0	0	0	0	6	0	1
Local	Essets	176	0	15	0	14	0	4	0	0	0	0	0
	Martinets	2032	3	134	0	134	0	5 + 2‡	0	26	0	24	0
	Para	1193	2	100	0	100	0	6	0	1	0	2	0
	Pierredar	599	1	27	0	27	0	5	0	2	0	3	0

928 † $p < 0.01$, < 5 genes significant

929 ‡ $p < 0.01$, > 5 genes significant

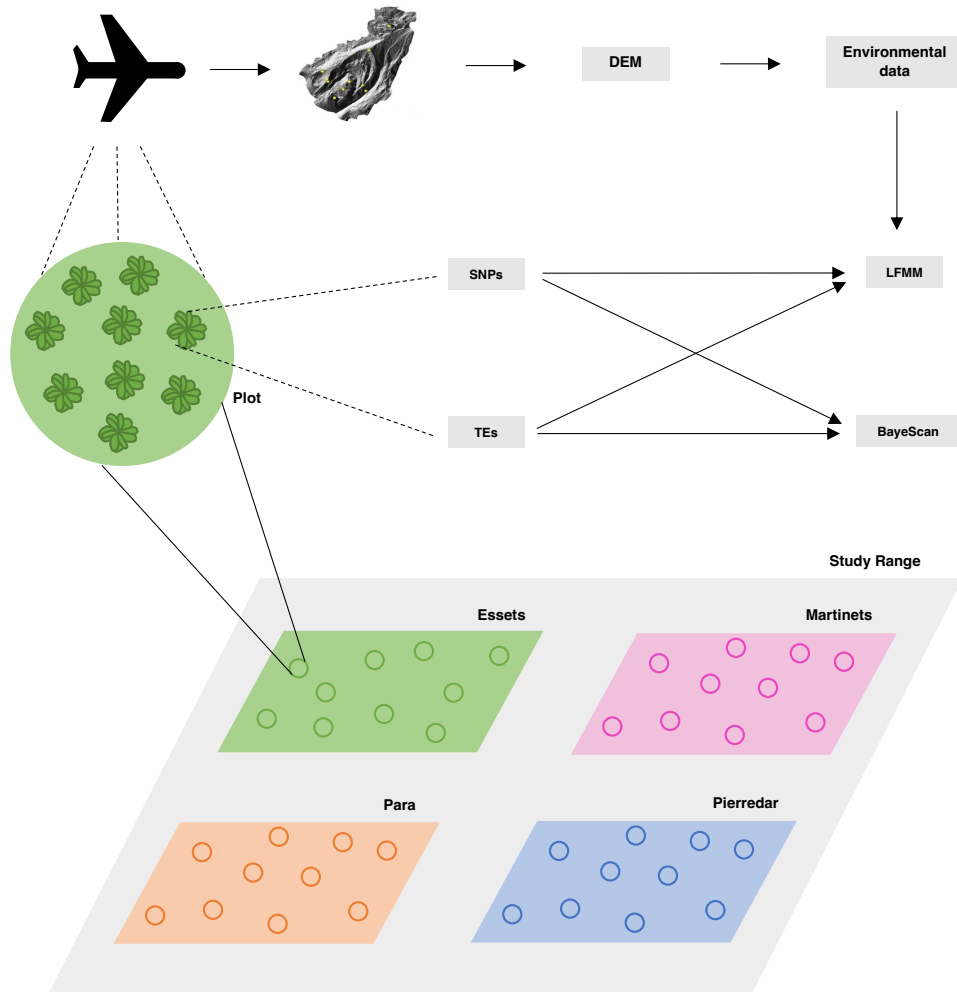
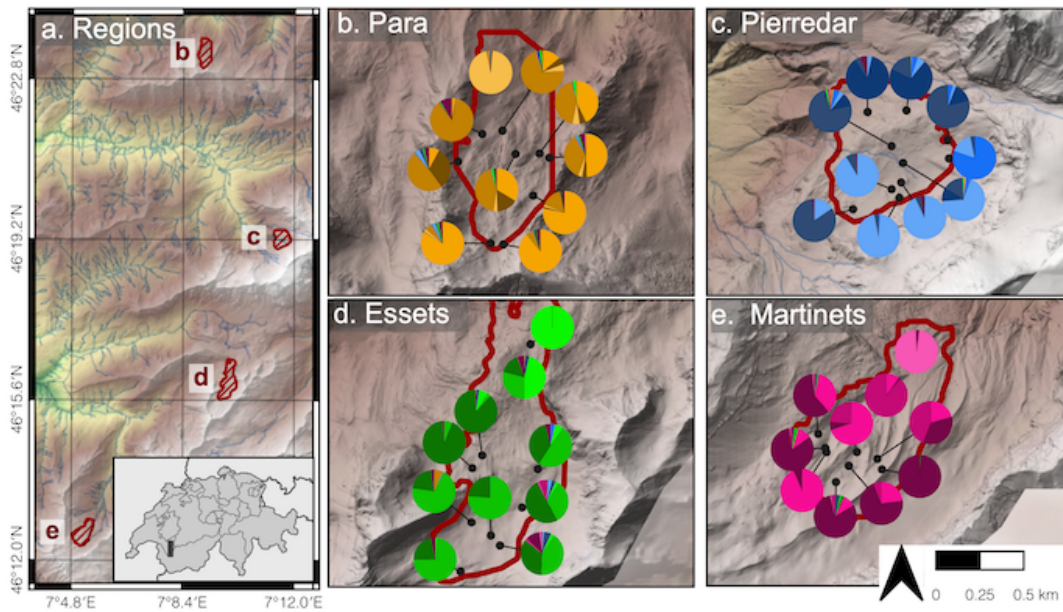


FIGURE 1 Schematic illustration of the study design and analyses. Four local populations were selected (illustrated by rectangles), each comprising ten plots (circles) with six to ten plants each. Individual plants were sampled and sequenced to identify single-nucleotide polymorphisms (SNPs) and polymorphic transposable elements (TEs). Environmental data used to characterise the habitat of each individual were computed from a regional high-resolution digital elevation model (DEM) at 0.5 m pixel resolution, based on light detection and ranging (LiDAR). Two types of analyses were performed to search for genomic signatures of selection: (i) genotype–environment association (GEA) analyses using a latent factor mixed model (LFMM; Frichot et al., 2013) and (ii) outlier analyses using BayeScan (Foll & Gaggiotti, 2008). Both types of analyses were performed at two hierarchical levels differing in spatial extent: (i) at the regional level among the four local populations and (ii) at the local level among plots within each population.



945 FIGURE 2 Locations of the four local populations of *Arabis alpina* and genetic
 946 assignment. (a) The study region situated in the western Swiss Alps, with four local
 947 populations (b) Para, (c) Pierredar, (d) Essets and (e) Martinets. Pie charts in (b)–(e)
 948 represent the proportions of assignment probabilities for each plot (black dot) within
 949 each studied region (delimited by a red line).
 950

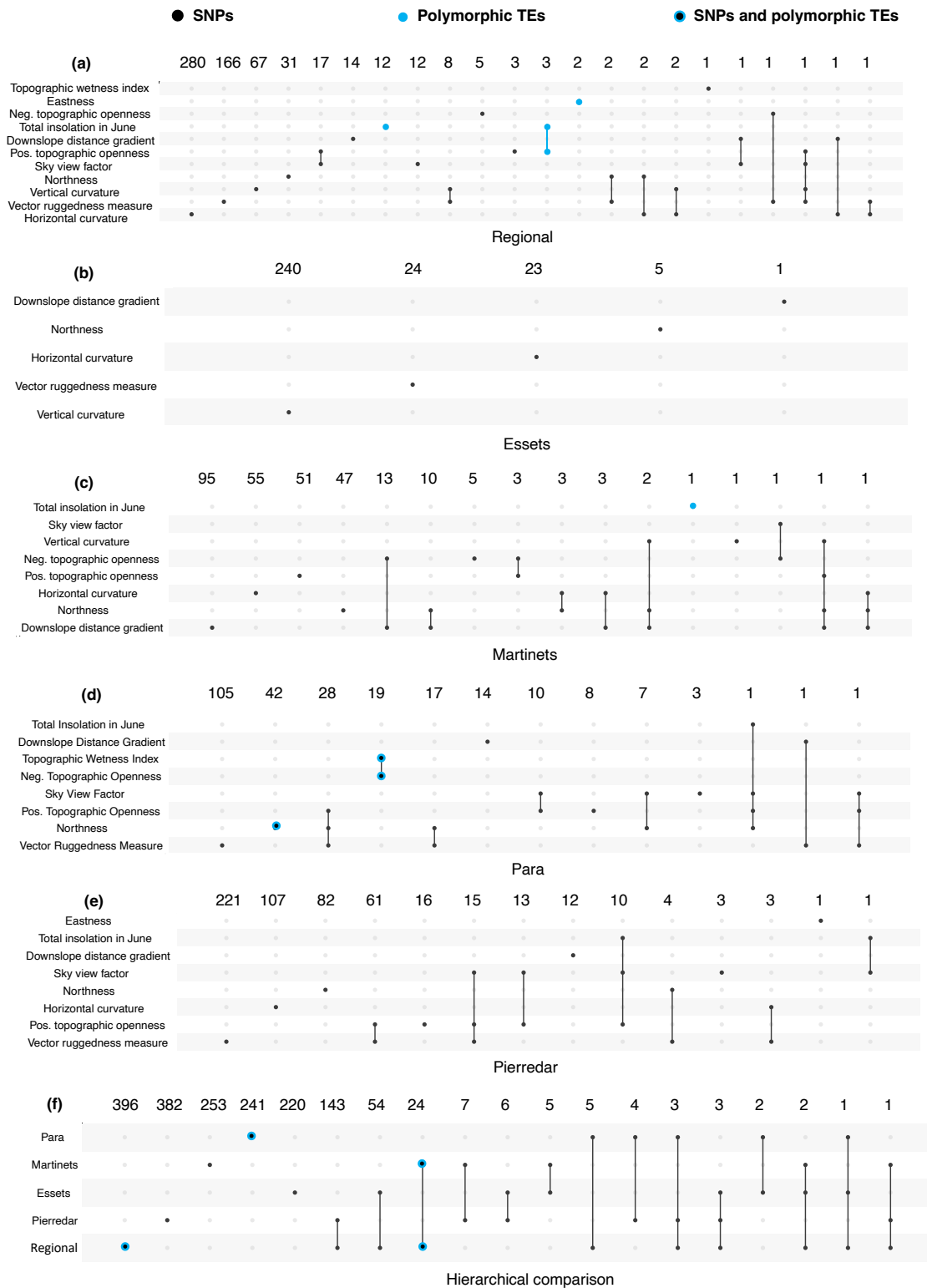


FIGURE 3 Common associations between environmental factors and hierarchical levels. Combination and magnitude of intersections of the significantly associated single-nucleotide polymorphisms (SNPs, high-impact) and polymorphic transposable elements (TEs) from the latent factor mixed models (LFMM) comparing environmental factors within each analysis. (a) Regional level, (b)–(e) local level, and (f) between hierarchical levels. The number of loci involved in each category is given above the corresponding plot.

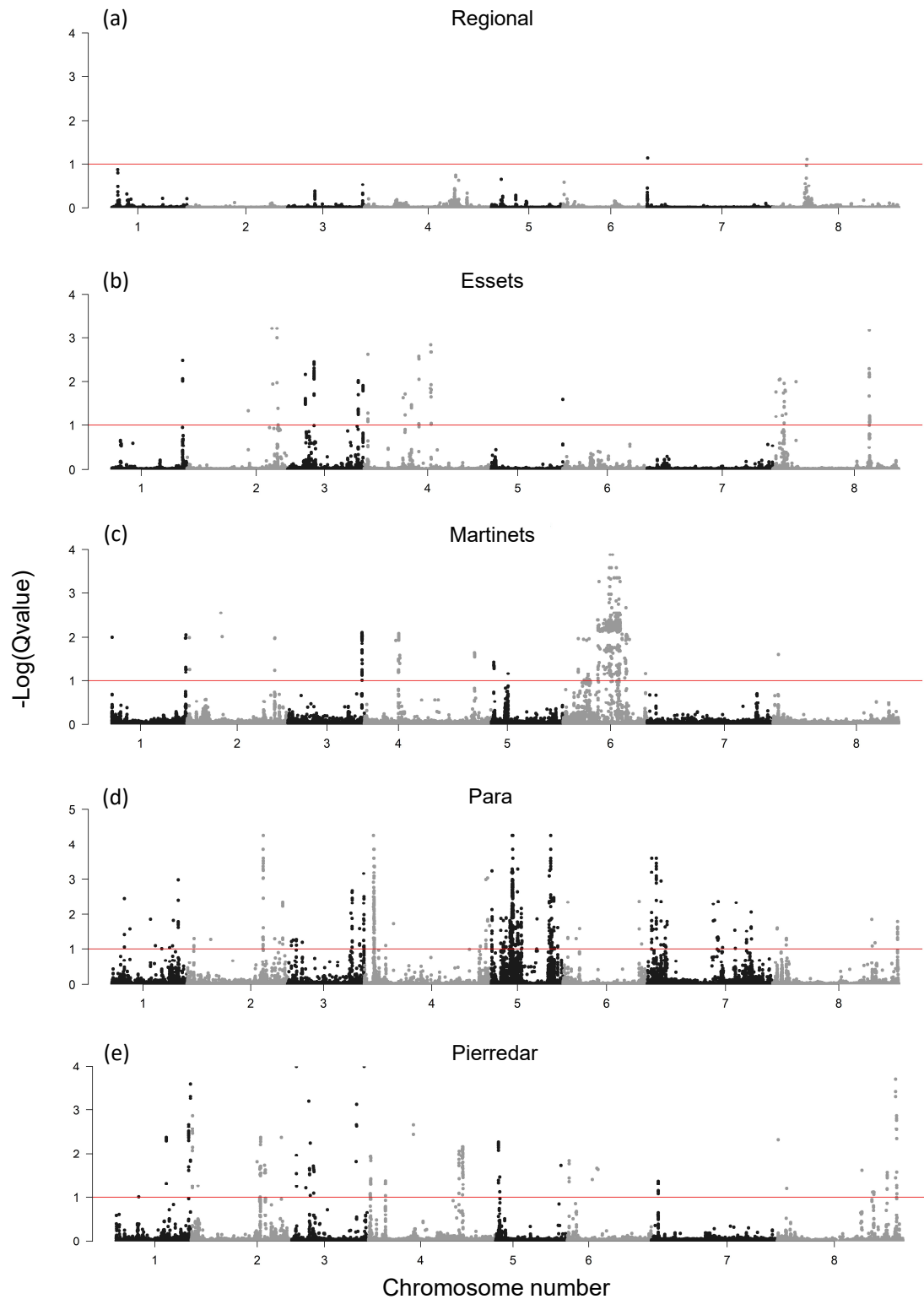
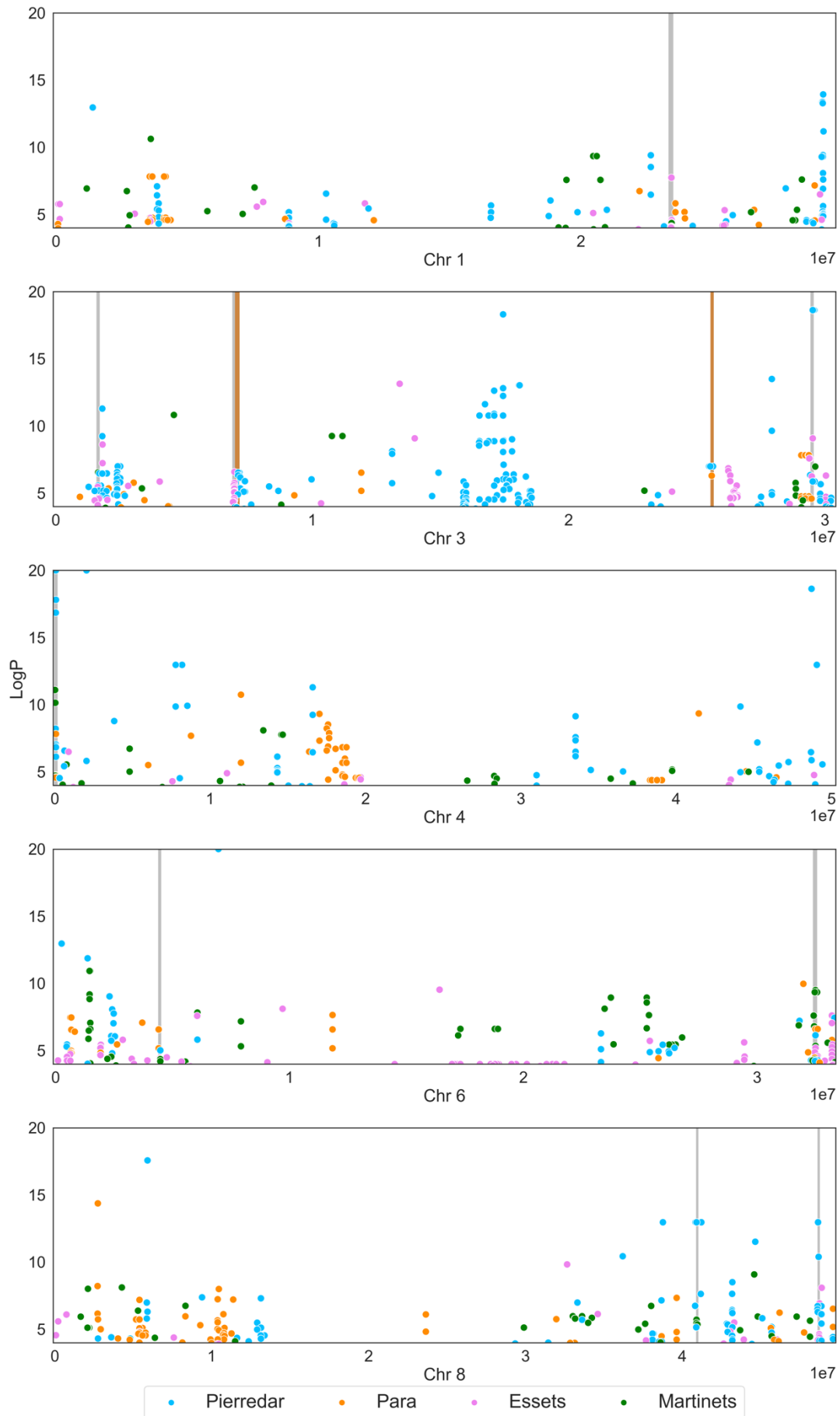


FIGURE 4 Manhattan plots showing BayeScan outliers. The analysis is based on the genome scan of single-nucleotide polymorphisms (SNPs) in *Arabis alpina* for the two levels of spatial extent: (a)–(d) local and (e) regional. Alternating black/grey dots mark detected SNPs within each of eight chromosomes (numbers); red lines represent the false discovery rate (FDR, 10%). Values of genetic differentiation F_{ST} as a function of the $\log(q)$ value can be found in Table S7.



967 FIGURE 5 Genomic regions indicating signatures of adaptation. Significant high-
968 impact single-nucleotide polymorphisms (SNPs) in *Arabis alpina*, identified in
969 genotype–environment association (GEA) analyses, with common chromosomal
970 blocks. Panels show the log p -value of all high-impact SNPs identified per local
971 population (denoted with different colours) for each of the five (out of eight)
972 chromosomes (CHR) of *A. alpina* for which common blocks could be found. Blocks
973 are denoted if detected in at least two populations, with vertical lines in grey for GEA
974 analyses and in brown for outlier analyses. Numbers in x-axis denote length of
975 chromosome in Mbp.

Supporting Information

Supportin information is available for this paper currently as a zip archive, except for Figure S2 shared via a link to Google Doc (see below).

Table S1 Geographical coordinates and values of the 13 environmental factors for each sampled individual of *Arabis alpina* in the four local populations.

Table 2 Definitions of the 13 environmental factors.

Table S3 The genomic inflation factor λ , calculated using latent factor mixed models (LFMMs), for single-nucleotide polymorphisms (SNPs) and polymorphic transposable elements (TEs) in *Arabis alpina*.

Table S4 Spearman's ρ between pairs of 13 environmental factors, measured in the four local populations of *Arabis alpina*.

Table S5 Summary results of the latent factor mixed models (LFMMs) for the single-nucleotide polymorphisms (SNPs) and the polymorphic transposable elements (TEs) in *Arabis alpina*.

Table S6 Associated single-nucleotide polymorphisms (SNPs) and polymorphic transposable elements (TEs) in *Arabis alpina*, based on the latent factor mixed models (LFMMs). (a) SNPs regional, (b) TEs regional, (c) SNPs local and (d) TEs local.

Table S7 Results of the outlier detection analyses, using BayeScan, based on single-nucleotide polymorphisms (SNPs) in *Arabis alpina* at the regional and local levels.

Table S8 Annotated single-nucleotide polymorphisms (SNPs) in *Arabis alpina* at the regional and local levels, based on the results of latent factor mixed models (LFMMs).

1007 Table S9 Results of the gene ontology (GO) enrichment analysis for latent factor
 1008 mixed models (LFMMs) with single-nucleotide polymorphisms (SNPs) and
 1009 polymorphic transposable elements (TEs) in *Arabis alpina* at the regional and local
 1010 levels.
 1011

1012 Table S10 Annotated associated polymorphic transposable elements (TEs) in *Arabis*
 1013 *alpina* at the regional and local levels, based on the latent factor mixed models
 1014 (LFMMs).
 1015

1016 Table S11 Annotated single-nucleotide polymorphisms (SNPs) in *Arabis alpina* at the
 1017 regional and local levels, based on outlier detection analyses using BayeScan.
 1018

1019 Table S12 Gene ontology (GO) terms in *Arabis alpina* at the regional and local levels,
 1020 based on the results of outlier detection analyses of single-nucleotide
 1021 polymorphisms (SNPs) using BayeScan.
 1022
 1023

1024 Figure S1 Principal component analysis (PCA) including the 304 individuals of *Arabis*
 1025 *alpina* from the four local populations, based on the twelve retained environmental
 1026 factors.
 1027

1028 Figure S2 Manhattan plots of the latent factor mixed model (LFMM) analyses for
 1029 *Arabis alpina*.
 1030 [https://drive.google.com/file/d/1_VbJxP9xHOdYLqtuMdzN3rcrosPwjqeg/view?usp=s](https://drive.google.com/file/d/1_VbJxP9xHOdYLqtuMdzN3rcrosPwjqeg/view?usp=share_link)
 1031 [hare link](https://drive.google.com/file/d/1_VbJxP9xHOdYLqtuMdzN3rcrosPwjqeg/view?usp=share_link)
 1032

1033 Figure S3 Shared genes and gene ontology terms between levels of spatial extent in
 1034 *Arabis alpina*.
 1035

1036 Figure S4 Plots of the results of the outlier detection analyses (using BayeScan),
 1037 based on the genome scan of single-nucleotide polymorphisms (SNPs) in *Arabis*

1038 *alpina*: (a) regional analysis and (b)–(e) local analyses. (b) Essets, (c) Martinets, (d)
1039 Para and (e) Pierredar.
1040
1041 Figure S5 Manhattan plots of the results of the outlier detection analysis (using
1042 BayeScan), based on the genome scan of polymorphic transposable elements (TEs) in
1043 *Arabis alpina*.