

1 Temperature and day length drive local adaptation  
2 in the Patagonian foundation tree species  
3 *Nothofagus pumilio*  
4

5 **Authors:** Jill Sekely <sup>^, #, @</sup>, Paula Marchelli <sup>%</sup>, Verónica Arana <sup>%</sup>, Benjamin Dauphin <sup>f</sup>, María  
6 Gabriela Mattera <sup>%</sup>, Mario Pastorino <sup>%</sup>, Ivan Scotti <sup>§</sup>, Carolina Soliani <sup>%</sup>, Katrin Heer <sup>^</sup>, Lars  
7 Opgenoorth <sup>#, f</sup>  
8

9 **Affiliations:**

10 <sup>^</sup> Forest Genetics, Albert-Ludwigs Universität Freiburg, Bertoldstraße 17, 79098 Freiburg,  
11 Germany

12 <sup>#</sup> Plant Ecology and Geobotany, Philipps-Universität Marburg, Karl-von-Frisch-Straße 8, 35032  
13 Marburg, Germany

14 <sup>%</sup> INTA Bariloche, Instituto de Investigaciones Forestales y Agropecuarias Bariloche IFAB (INTA-  
15 CONICET). Modesta Victoria 4450 (8400) S.C. Bariloche, ARGENTINA

16 <sup>f</sup> Swiss Federal Research Institute WSL, Zürcherstrasse 111, 8903 Birmensdorf, Switzerland

17 <sup>§</sup> Institut national de recherche pour l'agriculture, l'alimentation et l'environnement (INRAE),  
18 URFM, 228 Route de l'Aérodrome, 84914 Avignon, FRANCE

19

20 **@Author for correspondence:** [jtsekely@gmail.com](mailto:jtsekely@gmail.com)

21

22 **Key words:** climate, genotype-environment association, genome scan, outlier loci, *Nothofagus*

23

24

25 **Summary**

26 Climate change alters relationships among environmental conditions and thus has the potential to  
27 change the selection pressures acting on adaptive gene variants. Using a landscape genomic  
28 approach, we show that the southern beech species *Nothofagus pumilio* has notable genetic  
29 adaptations to climate along its 2000-kilometer-long range in the Andes. We screened 47,336 SNP  
30 loci in 1,632 contigs and found that high-latitude sampling sites have lower genetic diversity, likely  
31 due to greater impact of glacial oscillations at high latitudes. Using four genome scan methods, we  
32 identified 457 outlier SNPs that are either strongly differentiated among subpopulations or  
33 associated with environmental covariates related to temperature, day length, and precipitation.

34 Temperature and day length parameters were associated with notably more outliers than  
35 precipitation (n = 133, 113, and 61 outliers, respectively), and almost half of all annotated outliers  
36 were related to stress response (n=38, 21%) or catabolism-metabolism (n=43, 24%). Our findings  
37 suggest that *Nothofagus pumilio* is an ideal Andean model of genetic adaptation to climate change  
38 because it is locally adapted to extant climate conditions, and shifting patterns among  
39 environmental parameters may be detrimental to its future survival and adaptation potential.

## 40 **Introduction**

41 Contemporary climate change is expected to have acute impacts on forests, including shifts in  
42 species ranges, tree growth rates, and phenology (Brondizio, Settele, Díaz, & Ngo, 2019). Tree  
43 populations typically have high levels of standing genetic diversity due to their widespread  
44 distributions across diverse habitats and large effective population sizes, and therefore they may  
45 have high local adaptation potential even when faced with rapidly changing conditions (Kremer et  
46 al., 2012; Savolainen, Pyhäjärvi, & Knürr, 2007). However, the extraordinary challenge of  
47 contemporary climate change is that it will likely create novel combinations of precipitation,  
48 temperature, and photoperiod that neither occur within the current range nor have occurred for  
49 millions of years (Burke et al., 2018; Williams & Jackson, 2007). By decoupling current  
50 relationships among environmental conditions, no-analog conditions could impose unique  
51 selection pressures that will challenge tree populations' ability to survive.

52 While photoperiod is unaffected by climate change, temperature and precipitation patterns will  
53 shift across regions (Barros et al., 2015; Williams & Jackson, 2007). The consequences are myriad.  
54 In extratropical species, phenology is mediated by a combination of photoperiod and temperature  
55 cues (Howe, Hackett, Furnier, & Klevorn, 1995; Singh, Svystun, AlDahmash, Jönsson, &  
56 Bhalerao, 2017). Climate shifts also have implications for drought. Drought is a direct consequence  
57 of water availability, but its severity is influenced by temperature, since high temperatures can  
58 increase evapotranspiration rates and drought stress during the growing season (Vicente-Serrano,  
59 Beguería, & López-Moreno, 2010). Furthermore, pests and pathogen species including insects,  
60 bacteria, and fungi may likewise experience range or phenology shifts due to climate change.  
61 Therefore, no-analog climate combinations will likely affect many genes and traits related to  
62 phenology (Hänninen & Tanino, 2011), extreme temperature and drought response (e.g. Niinemets,  
63 2010), and immune response (Haynes, Liebhold, Lefcheck, Morin, & Wang, 2022). Selection upon

64 these genes leaves signatures of adaptation along the genome, and searching for signatures among  
65 putative adaptive loci can provide critical information about how tree populations might respond  
66 to climate change. An initial step is to establish whether adaptation to environmental clines is  
67 currently observed.

68 The Patagonia region of the southern Andes mountain range presents an ideal study location due  
69 to its north-south orientation and two geographically orthogonal environmental gradients. The first  
70 is a north-south gradient of day-length and temperature that is driven by latitude, and the second is  
71 a west-east precipitation gradient driven by prevailing winds and a montane rain shadow. Patagonia  
72 is strongly affected both by glacial oscillations and the El Niño-Southern Oscillation (Morales et  
73 al., 2020). Climate change has already intensified the acute drought stress following La Niña events  
74 (Cai et al., 2015). The most widespread native tree species in this region is the southern beech  
75 “lenga” (*Nothofagus pumilio* ([Poepp. & Endl.] Krasser)), a cold-tolerant deciduous tree that  
76 inhabits a nearly continuous range more than 2,000 kilometers long. Its range encompasses a  
77 diverse climate space, from 5,000 mm of annual precipitation in the west to just 200 mm in the east  
78 (Veblen, Donoso, Kitzberger, & Rebertus, 1996 ; Fig 1). Little is known about adaptation patterns  
79 at the fine geographical scale in this non-model montane species. Previous studies have examined  
80 the neutral genetic diversity and phenotypic plasticity using neutral markers (Arana et al., 2016;  
81 Mathiasen & Premoli, 2013, 2016; A C Premoli, 2003). However, adaptive variation along  
82 environmental clines using high-throughput SNPs has not yet been assessed.

83 The objective of this study was to characterize extant local adaptation in *N. pumilio*. We searched  
84 for signatures of local adaptation within candidate genes in situ using a landscape genomics  
85 approach to assess how evolutionary processes and environmental variation have shaped genetic  
86 variation (Capblancq & Forester, 2021; Rellstab, Gugerli, Eckert, Hancock, & Holderegger, 2015).  
87 Local adaptation depends on a fine balance among many factors whose individual effects can be  
88 difficult to differentiate. Thus, choosing an appropriate sampling design and analysis methods is  
89 crucial for improving study power to detect signatures (Lotterhos & Whitlock, 2015; Meirmans,  
90 2015). We used a paired-site sampling design, which aims to disentangle environmental effects  
91 from neutral population structure by maximizing the climatic distance between pairs of sampling  
92 sites while minimizing the neutral genetic divergence (Lotterhos & Whitlock, 2015; Scotti et al.,  
93 2023). We distributed sampling site pairs along the two orthogonal gradients on the eastern side of  
94 the Andes, and assessed variation using univariate and multivariate genome scan methods that

95 incorporate population structure. We hypothesize that the two gradients have exerted strong  
96 selection pressure on *N. pumilio* and have resulted in signatures of local adaptation in candidate  
97 genes. We address these questions by quantifying the strength of correlations between  
98 environmental covariate predictors and genetic SNP responses. We predicted that (i) allele  
99 frequencies in candidate genes that are linked to growth and stress response will correlate with  
100 temperature and photoperiod clines and (ii) allele frequencies in candidate genes linked to drought  
101 response will correlate with precipitation, albeit to a weaker degree given the narrower precipitation  
102 gradient covered.

## 103 **Materials and Methods**

### 104 **Study species**

105 *Nothofagus pumilio*, common name “lenga,” is a deciduous tree species native to the southernmost  
106 temperate forest of the Andes mountains. It is a wind-pollinated and strictly outcrossing species  
107 that grows between latitudes 35° to 56 °S (Veblen et al., 1996). A member of the Fagaceae family,  
108 its closest relatives in the Northern Hemisphere are *Fagus* species (Vento & Agraín, 2018). It is  
109 cold-tolerant and often forms monospecific stands up to the montane tree line. North of 41°S, lenga  
110 grows in the subalpine zone, but it also grows at sea level in the southernmost (i.e. poleward) parts  
111 of its range. *Nothofagus pumilio* is an important local forestry species, although much of the local  
112 timber industry has historically focused on introduced genera from the Northern Hemisphere such  
113 as *Pinus* and *Eucalyptus* (Gea-Izquierdo, Pastur, Cellini, & Lencinas, 2004). It is a non-model  
114 species without a reference genome and a de novo transcriptome was recently assembled (Estravis-  
115 Barcala et al., 2021).

### 116 **Sampling design**

117 To disentangle neutral and adaptive genetic variation, we used a paired-site study design after  
118 Lotterhos and Whitlock (2015). Sites within a pair are geographically close enough to share a  
119 demographic history but are distant enough that they experience different environmental selection  
120 pressures. According to Lotterhos and Whitlock (2015), this sampling design has greater power to  
121 detect signatures of local adaptation compared to transect or random sampling designs, particularly  
122 when combined with genome scan methods based on latent factor mixed models (LFMM) and  
123 Bayesian methods (see genome scan methods below). We selected eight localities that were  
124 distributed along the species’ full latitudinal range on the eastern slope of the Andes (Fig 1), and

125 each locality contained two (or, in one locality, three) sampling sites. Linear distance among paired  
126 sites within a locality was always less than three kilometers, and elevation difference among the  
127 sites' centroids was between 150 - 320 meters to capture an approximate 1-2 °C difference in mean  
128 annual temperature due to lapse rate (Whiteman, 2000). The high-elevation sites were located  
129 below the alpine treeline to avoid sampling trees with shrub-like krummholz formation (Table 1).  
130 In addition, we sampled three singleton localities in marginal habitats (i.e. located at the edge of  
131 the species distribution). Each singleton locality has one sampling site (Epulauquen (site 1), La  
132 Hoya (6), and Jose de San Martin (12)). The latter two sites are located at approximately the same  
133 latitude as one of the paired localities, in an attempt to capture a wider portion of the east-west  
134 precipitation gradient.

135  
136 We sampled between 21 and 25 adult trees per site for a total of 496 individuals. Selected trees  
137 were dominant or co-dominant and at least 50 years old, as confirmed by annual tree rings (Sekely  
138 et al, manuscript in progress). Intertree distances were at least 30 meters to reduce the chance of  
139 sampling directly related individuals. Geographic coordinates for each tree were recorded with a  
140 handheld GPS device (Garmin model GPSMAP 64st). We collected fresh leaf buds for DNA  
141 extraction and stored them at -80°C. Immediately before extraction, buds were manually descaled,  
142 flash-frozen with liquid nitrogen, and ground with mortar and pestle. Samples were randomly  
143 assigned to extraction batches. Total genomic DNA was extracted from 0.1 g of plant material  
144 using the CTAB protocol by Doyle (1990) with minor modifications, since *N. pumilio* leaf buds  
145 have high levels of polysaccharides and polyphenols that can impact the quality and quantity of  
146 extracted DNA. Therefore we added 1% soluble Polyvinylpyrrolidone (PVP) and Dithiothreitol  
147 (DTT) to the lysis buffer (Porebski, Bailey, & Baum, 1997). Extracted DNA quantity was measured  
148 with a QUBIT 1.0 Fluorometer (Invitrogen, Carlsbad, CA) and the quality was spot-checked with  
149 Nanodrop™ 2000 (ThermoFisher Scientific, catalog ND-2000). Extracted DNA samples were  
150 randomized among plates for downstream sequencing.

151  
152 **Table 1.** Characteristics of the *Nothofagus pumilio* sampling sites, ordered from North (site number 1) to  
153 South (20).

Site number	Locality	Elevation class	Elevation (m a.s.l.)	Latitude (°)	Longitude (°)	# samples
1	Epulauquen	singleton	1511	-36.8321	-71.1134	25

2	San Martín dl Andes	high	1478	-40.1263	-71.4886	25
3	San Martín dl Andes	low	1253	-40.1281	-71.4799	25
4	Cerro Otto	high	1382	-41.1482	-71.3783	25
5	Cerro Otto	low	1146	-41.1512	-71.3658	24
6	La Hoya	singleton	1442	-42.8341	-71.2592	25
7	Trevelin	high	1360	-43.0565	-71.5877	21
8	Trevelin	middle	1312	-43.0548	-71.5847	25
9	Trevelin	low	1085	-43.0663	-71.574	25
10	Lago Guacho	high	1314	-43.8121	-71.4513	25
11	Lago Guacho	low	1162	-43.823	-71.4629	25
12	José de San Martín	singleton	1317	-43.8281	-70.757	24
13	El Triana	high	915	-45.6044	-71.7387	25
14	El Triana	low	729	-45.6119	-71.7181	25
15	El Chaltén	high	670	-49.0749	-72.9001	25
16	El Chaltén	low	505	-49.0986	-72.9007	25
17	El Calafate	high	614	-50.4683	-72.9687	24
18	El Calafate	low	295	-50.4729	-72.979	25
19	Ushuaia	high	326	-54.8191	-68.5575	25
20	Ushuaia	low	20	-54.8223	-68.5675	25

154

## 155 **Environmental data and covariate choice**

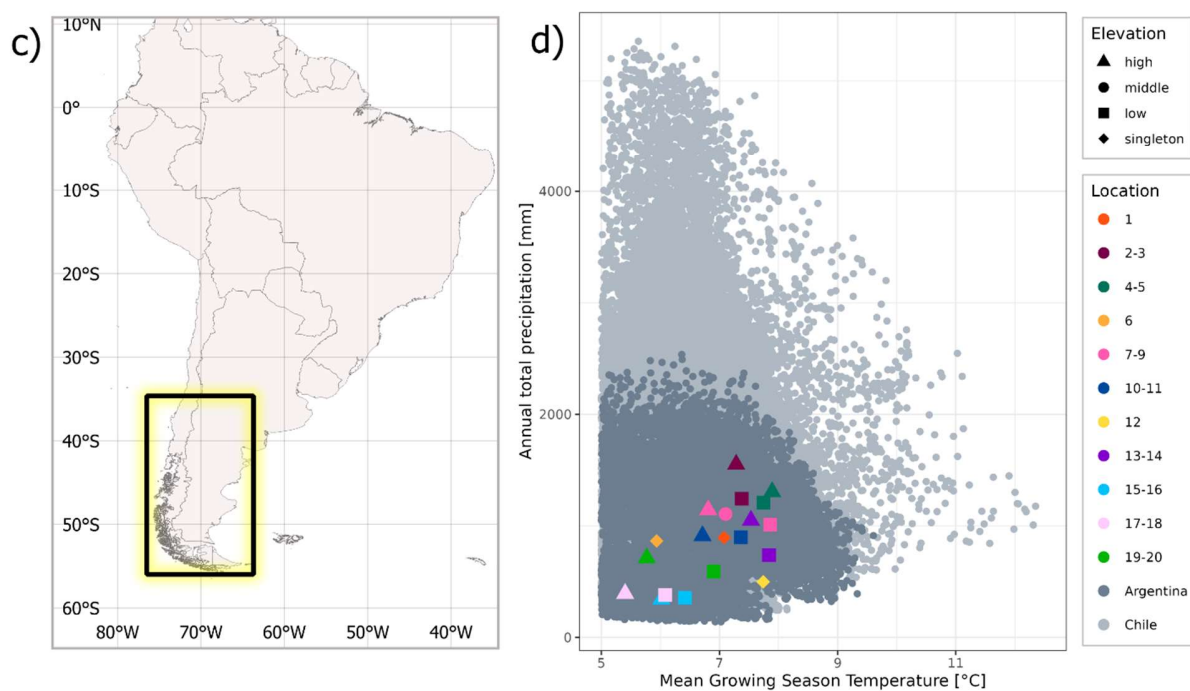
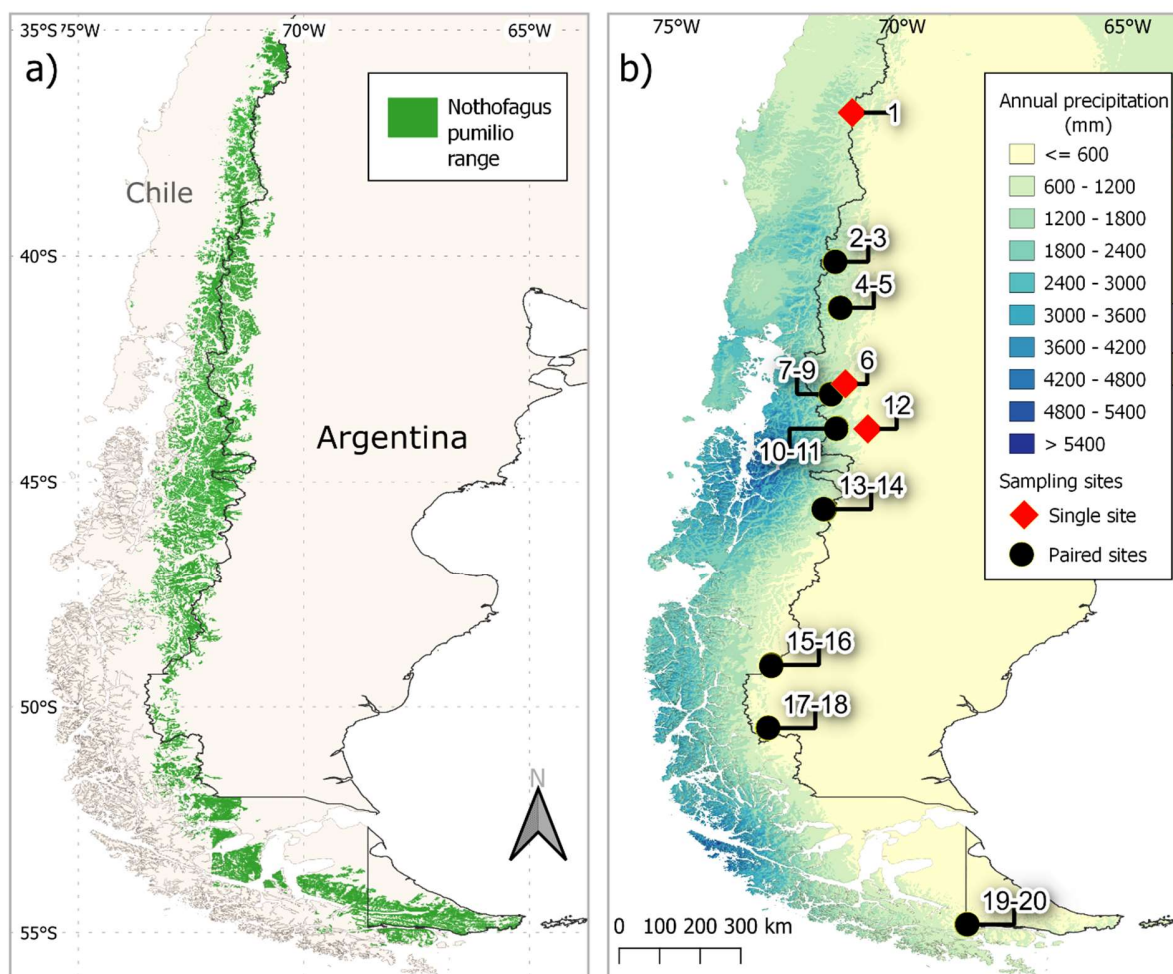
156 Empirical climate data is limited for the Andes region, so environmental covariates were extracted  
157 from the public repository climate dataset CHELSA v.1.2 (Karger et al., 2017). CHELSA  
158 incorporates empirical climate data from 1979-2013, from which further bioclim variables were  
159 derived and extrapolated across the globe at a resolution of 30 arc sec (~1 km<sup>2</sup>). We chose this  
160 dataset since it has been shown to represent more accurate orographic conditions than WorldClim  
161 (e.g. Bobrowski, Weidinger, & Schickhoff, 2021). We extracted tree-level data from climate layers  
162 with R::raster package (v. 3.6.14) using the extract() command for individual tree GPS locations  
163 and the “bilinear” option, which interpolates values from the four nearest raster cells to approximate  
164 finer-scale climate parameters (Hijmans, 2023).

165

166 Genome scans are sensitive to collinearity, so covariate pruning prior to analysis is a critical step  
167 (Dormann et al., 2013; Rellstab et al., 2015). From the CHELSA dataset we first selected a short  
168 list of variables related to temperature and drought stress (Fig S1). We calculated pairwise Pearson  
169 correlation values among these covariates, using the R package psych (v 2.2.9, Revelle, 2022), and  
170 retained only the most relevant variables that had a value less than  $|0.8|$  (Fig S2). Ultimately, we  
171 selected two CHELSA temperature parameters (number of frost days and isothermality (bioclim  
172 3)), one precipitation parameter (annual precipitation amount (bioclim 12)) and one temperature-  
173 limited precipitation parameter (precipitation during growing season (gsp\_9)). Finally, since we  
174 also investigated circadian clock candidate genes, we calculated average day length in the  
175 midsummer month of January (dl.Jan) using the R package geosphere (v 1.5.18, Hijmans, 2022).  
176 All environmental variables were scaled prior to association analyses.

177







179 **Figure 1 : Characterisation of the *Nothofagus pumilio* distribution and study area.** a) species distribution map, b)  
180 example of a bioclim layer from CHELSA (total annual precipitation) with sampling sites, c) overview map, d) climate  
181 space inhabited by *N. pumilio* in terms of mean temperature of all growing season days (°C) based on TREELIM and  
182 mean annual total precipitation (mm). Both values are the mean across years 1979-2013. Values are shown for the full  
183 distribution range in Argentina and Chile and sampling sites are grouped by color, with local elevation classes within  
184 sites differentiated by shape. Mean growing season temperature was truncated at 5°C to compensate for low spatial  
185 resolution of CHELSA data, which showed erroneous low-temperature artifacts at high elevations due to sharp  
186 mountain slopes.  
187

## 188 **Probe design**

189 Trees were genotyped with targeted sequencing (i.e. exome capture), for which we assembled a  
190 starting set of candidate genes. A recent study investigated 811 candidate gene orthogroups in seven  
191 European tree species that are closely- or distantly-related to *N. pumilio*, including Fagaceae  
192 members (Milesi et al., 2023 in review; Opgenoorth et al., 2021). Candidate genes were pertinent  
193 to environmental variables and were therefore selected from Gene Ontology (GO) and KEGG gene  
194 regulation networks related to cold, heat, drought, and immune response (Ashburner et al., 2000;  
195 Carbon et al., 2009). These orthogroups are represented by 1,789 candidate genes in *Arabidopsis*  
196 *thaliana*. We additionally investigated 415 species-specific *N. pumilio* candidate genes, including  
197 those that were differentially expressed in a recent heat stress transcriptomic study (Estravis-  
198 Barcala et al., 2021) or are affiliated with wood growth or circadian clock rhythms (Estravis-  
199 Barcala et al., 2020).

200  
201 We used BLASTn (Altschul, Gish, Miller, Myers, & Lipman, 1990) to align the de novo *N. pumilio*  
202 transcriptome with the 2,204 candidate genes, retained the best transcriptome hit per gene by  
203 applying an e-value threshold of  $10^{-5}$ , and finally selected the best sequence hit per gene. This  
204 resulted in 1,467 contig hits that covered 2.58 Mb. To reach the probe design target size of 3 Mb,  
205 we complemented the list with the longest available contigs from the *N. pumilio* transcriptome. The  
206 final probe design encompassed 1,913 contigs, each containing between 200 - 11,873 nucleotides.  
207

## 208 **Library preparation and target sequencing**

209 Libraries were prepared with SeqCap EZ-HyperPlus (Roche Sequencing Solutions). Library size  
210 was analyzed with Bioanalyzer High Sensitivity DNA assay (Agilent technologies), library  
211 quantity was analyzed with Qubit 2.0 Fluorometer, and libraries were sequenced on NovaSeq 6000  
212 (Illumina) in pair-end mode with 150 cycles per read. A total of 3.1 trillion reads were produced.  
213 Base-calling and demultiplexing were completed with Illumina bcl2fastq v.2.20. Reads were

214 trimmed using ERNE v1.4.6 (Del Fabbro, Scalabrin, Morgante, & Giorgi, 2013) and cutadapt  
215 (Martin, 2011) then mapped onto the transcriptome with BWA-MEM v0.7.17 (Li & Durbin, 2009).  
216 Variants were called using gatk-4.0 (Poplin et al., 2017), first with HaplotypeCaller and then joint  
217 genotyping was performed using GenotypeGVCFs (DePristo et al., 2011). Variants were selected  
218 with GATK SelectVariants and coarsely quality-filtered in VariantFiltration. Default parameters  
219 were used for each step.

220

### 221 **Variant quality filtering**

222 Variants were quality-filtered using general best practice thresholds in vcftools v 0.1.16 (Danecek  
223 et al 2011). These thresholds were minimum read depth per locus  $> 8$  (command: `--minDP 8`),  
224 minimum quality  $> 20$  (`--minGQ 20`), and maximum missing data per locus 20% (`--max-missing`  
225 `0.8`) (Carson et al., 2014). We calculated genotype missingness per individual (`--imiss`). GATK in  
226 docker mode was used to remove all newly-created monomorphic loci and all individuals with  $>$   
227 50% missing data ( $n = 3$  individuals). Multiallelic loci were removed, leaving only biallelic loci (`-`  
228 `min-alleles 2` and `--max-alleles 2`). After these initial quality filtering steps, our dataset contained  
229 116,136 SNPs in 1,783 contigs.

230

231 Paralogous loci were identified with the HDplot method and its accompanying R script (McKinney,  
232 Waples, Seeb, & Seeb, 2017), then pruned based on author recommendations ( $H > 0.6$  and/or  $D >$   
233  $|20|$ ). We used R version 4.2.2 for all analyses (R Core Team, 2022). We pruned the dataset of loci  
234 in linkage disequilibrium using plink v1.9 (Chang et al., 2015) to retain the allele with the greater  
235 minor allele frequency. We used the following settings: window size 50, stepwise progression 10,  
236 and  $r^2$  threshold 0.5 (`plink v1.9 --indep-pairwise 50 10 0.5`). This pipeline created our “main  
237 dataset,” which contained 47,336 SNPs (in 1,632 contigs). As a subsequent step, we applied a minor  
238 allele frequency filter of 5% to create a “maf-filtered dataset” that contained 9,601 SNPs (1,437  
239 contigs).

240

### 241 **Descriptive genetic diversity statistics and population structure**

242 We ran all genetic diversity and population structure analyses using the main dataset, with the  
243 exception of nucleotide diversity. Pairwise  $F_{ST}$  statistics were calculated in vcftools for every  
244 possible pair of sampling sites ( $n=190$ ) with the weighted  $\theta$  correction (Weir & Cockerham, 1984).  
245 The rarefied count of private alleles was calculated with the R::poppr package (v 2.9.3, Kamvar,

246 Tabima, & Grünwald, 2014). We calculated heterozygosity and  $F_{IS}$  using hierfstat (v 0.5.11,  
247 Goudet & Jombart, 2022). Nucleotide diversity was calculated with pixy (v 1.2.7; Korunes &  
248 Samuk, 2021), which includes invariant sites to calculate unbiased values. Therefore our input  
249 dataset contained the main dataset plus every called invariant site, which were quality-filtered using  
250 the same thresholds. Per pixy user guidelines, we aggregated values within a sampling site by  
251 summing raw count differences and dividing by summed comparisons. Pearson correlation  
252 coefficients between genetic diversity statistics and latitude were calculated using ggpubr package  
253 and stat\_cor command (v 0.5.0, Kassambara, 2022). Population structure was analyzed using the  
254 ADMIXTURE software (Alexander, Novembre, & Lange, 2009). We assessed every K value from  
255 1 to 20, to represent the 20 sampling sites. Singletons can confound model-based inference of  
256 population structure such as ADMIXTURE (Linck & Battey, 2019), so they were removed prior  
257 to analysis.

258

## 259 **Genome scan methods**

260 To determine whether SNPs are under selection, we applied genome scan methods to the maf-  
261 filtered dataset. Genome scans compare genetic variation of SNP loci across the targeted genome  
262 areas and identify over-differentiated loci, hereafter called outlier SNPs. There is an ever-growing  
263 list of genome scan tools and algorithms (see Bourgeois & Warren, 2021), each of which has its  
264 own benefits and pitfalls (e.g. Rellstab et al., 2015; Waldvogel, Schreiber, Pfenninger, &  
265 Feldmeyer, 2020). Common practice is to analyze a SNP dataset with multiple methods and inspect  
266 overlap among their results, since this provides stronger evidence that a SNP is a true-positive  
267 outlier (de Villemereuil, Frichot, Bazin, François, & Gaggiotti, 2014; Waldvogel et al., 2020).

268

269 Genome scans search for loci that are strongly differentiated among genetic clusters (e.g.  
270 subpopulations) and/or strongly associated with environmental gradients (Savolainen, Lascoux, &  
271 Merilä, 2013). We use both methods and classify them respectively as “population differentiation”  
272 (sensu Beaumont & Nichols, 1996) and “genotype-environment association” (sensu Hedrick,  
273 Ginevan, & Ewing, 1976). Population differentiation (PD) tests are advantageous because they  
274 require no prior knowledge about environmental selection pressures and therefore are less  
275 susceptible to errors related to missing environmental data or suboptimal choice of climatic  
276 variables. We used pcadapt, which is a multiple linear regression method that identifies outlier loci  
277 via correlation to genetic structure ordination axes (Duforet-Frebourg, Bazin, & Blum, 2014) and

278 is implemented in the R package `pcadapt` (v 4.3.3, Privé, Luu, Vilhjálmsson, & Blum, 2020). The  
279 number of principal components ( $K = 3$ ) was chosen based on the lowest genomic inflation factor  
280 ( $\lambda = 1.35$ ) and the deflation of explained variance of the first three principal components. On  
281 the other hand, genotype-environment associations (GEA) can provide evidence about which  
282 environmental variables are associated with adaptive differentiation. The null hypothesis in a GEA  
283 is that there is no correlation between allele frequencies and environmental covariates (Manel et  
284 al., 2010). GEA methods have greater power than PD to detect weakly selected loci, which may  
285 only show small allele frequency shifts but are crucial for adaptation (e.g. De La Torre, Wilhite, &  
286 Neale, 2019). We used three GEA methods that assess SNP frequency variations and environmental  
287 covariates in different combinations of univariate and multivariate approaches.

288  
289 The Bayesian hierarchical model BayPass (v. 2.31, Gautier, 2015) is a univariate method for both  
290 genetic and environmental components. It computes  $X_T X$  values, which are analogous to the SNP-  
291 specific  $F_{ST}$  that is calculated by PD methods, and Bayes Factor (BF), which measures the strength  
292 of correlation between an individual SNP and an individual environmental covariate. BayPass is a  
293 stochastic algorithm, so we ran three iterations with the core model (i.e. without environmental  
294 covariates) to calculate the population covariance matrix and  $X_T X$  values, then calculated median  
295 values of each statistic. Next we ran three iterations of the auxiliary covariate model using the  
296 median covariate matrix to calculate Bayes Factor values and again retained median values. An  
297 important distinction is that BayPass requires allele frequencies to be pooled by sampling site, thus  
298 treating each of the 20 sites as its own population, unlike the other GEA methods.

299  
300 The two additional GEA methods assess multivariate environmental parameters to account for  
301 interaction among environmental factors across the Patagonian landscape. The first is latent factor  
302 mixed models (LFMMs), which search for correlations between an individual SNP and multivariate  
303 environmental predictors while simultaneously correcting for hidden (i.e. latent) factors (Frichot,  
304 Schoville, Bouchard, & François, 2013). Latent factors can include unobserved demographic  
305 patterns or unmeasured environmental variables. This method is demonstrably effective in  
306 continuous ranges and has low false positive rates even if isolation-by-distance patterns are present  
307 (Frichot et al., 2013). We ran LFMM 2 using the R package LEA (v 3.9.5, Frichot & Francois,  
308 2015) and the `lfmm2()` command. The number of latent factors ( $K=3$ ) was chosen based on the  
309 deflation of explained variance along the first three principal components. LFMM cannot handle

310 missing locus data, so we imputed data (LEA::impute() command) using K=3 prior to analysis.  
311 The second GEA method, redundancy analysis (RDA), is a multivariate approach in terms of both  
312 environmental predictor and the genetic response variables (Capblancq & Forester, 2021).  
313 Multivariate genetic analysis may account for polygenic architecture in adaptive traits. We used  
314 the rda() command in the vegan package (v. 2.6.4, Oksanen et al., 2022). The imputed genotype file  
315 that was created for the LFMM analysis was also used in this analysis for consistency across  
316 methods. SNPs were scaled prior to analysis (scale=T). We used a full RDA model with climatic  
317 variables only, rather than a partial model, because there were strong correlations among genetic  
318 principal components, geographic variables, and climatic variables that strongly reduced the signal  
319 (Fig S2) (Capblancq & Forester, 2021). The custom command “rdadapt()” from the accompanying  
320 R script was then used to calculate q-values using K=3 (Figs S3, S4).

321

## 322 **Identifying signatures of selection across methods**

323 Combining multiple genome scan analyses improves study power to reject neutrality, but false  
324 discovery rates within and among tests must be controlled to account for multiple testing and  
325 confounding effects (François, Martins, Caye, & Schoville, 2016). The null hypothesis underlying  
326 all genome scan tests is that a locus is not under selection, and all four aforementioned methods  
327 use chi-squared distributions of putatively neutral alleles to reject this null hypothesis on a per-  
328 locus scale. This shared methodological basis makes it possible to compare results among tests, as  
329 long as the significance values are properly calibrated. The first calibration step occurs within  
330 methods and is based on inherent test statistics used to reject the null hypothesis. In LFMM,  
331 pcadapt, and RDA, test statistics were calibrated via the genomic inflation factor. Inflation  
332 calibration is automatically implemented in pcadapt and RDA, but it is manually implemented in  
333 LFMM using the lfmm2.test() command. BayPass is calibrated with the population covariance  
334 matrix created during core model analysis.

335

336 Next we curated a study-wide list of candidate outliers by first converting the calibrated p-values  
337 to q-values using the p.adjust() command in the base R stats package with the Benjamini-Hochberg  
338 equation (Benjamini & Hochberg, 1995), then applying the same false discovery rate control  
339 threshold across tests (François et al., 2016). An ideal false discovery rate threshold balances Type  
340 I errors (false positives) and Type II errors (false negatives). We applied a fairly lenient study-wide

341 false discovery rate threshold of 0.01 (i.e. < 1% false positives). Finally, we compared overlap  
342 among all four tests to determine evidence strength for true positive outliers.

### 343 **Annotating outliers**

344 To determine whether outlier SNPs would change the amino acid produced by its containing codon  
345 (i.e. synonymous or nonsynonymous substitution), we again used nucleotide BLAST with the  
346 contigs containing the candidate SNPs and then translated amino acids within the top gene hit. A  
347 recent study found that both synonymous and nonsynonymous mutations can affect the level of  
348 mRNA expression of a mutated gene, and that the effect's magnitude partially predicted the  
349 phenotypic fitness outcome (Shen, Song, Li, & Zhang, 2022). In our study, many BLASTn top  
350 gene hits included either frame shifts or premature stop codons before the target locus, calling into  
351 question the impact of substitution. Therefore, we reported the substitution type in the full outlier  
352 list (Table S1), but we did not differentiate among them in our gene function analysis or discussion.  
353 Gene functions were obtained from the UniProt database (The UniProt Consortium, 2023). We  
354 used PANTHER to summarize GO results and run a GO-term statistical overrepresentation test,  
355 using our starting candidate gene list as the background list (Thomas et al., 2022).

## 356 **Results**

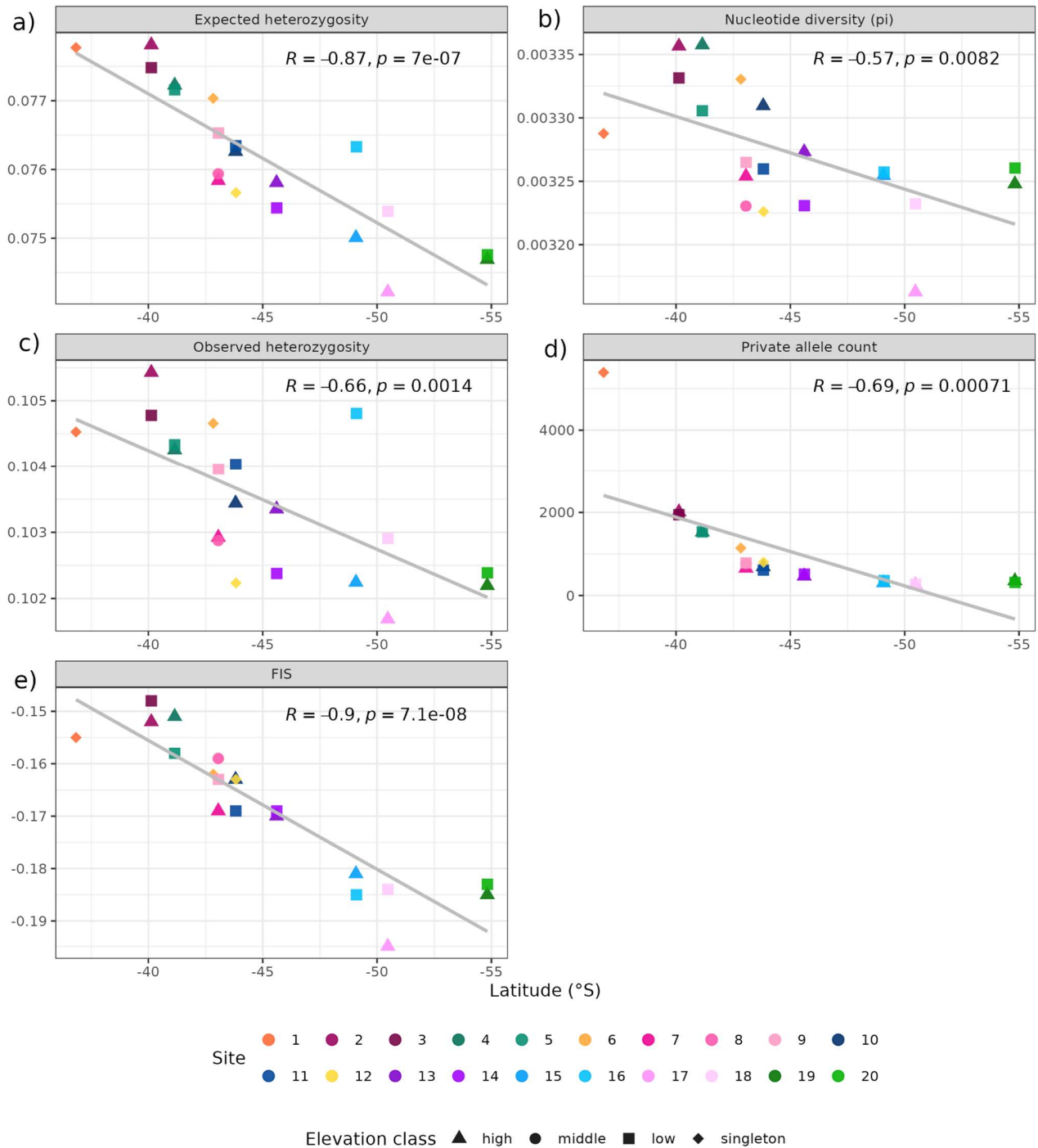
### 357 **Population genetic structure and diversity**

358 We found significant negative correlations between each genetic diversity parameter and latitude  
359 (Fig 2), meaning diversity values are highest in low-latitude sampling sites and decrease poleward.  
360 For example, the correlation between latitude and expected heterozygosity has a high R-value of -  
361 0.87 and significant p-value of  $7e-07$ . On the contrary, there are no consistent significant local  
362 elevation trends within paired sites, meaning the locally higher sites do not always have lower  
363 diversity. In the north, diversity values tend to be greater in high-elevation sites (e.g. sites 2 & 4),  
364 but in the south they tend to be lesser in high-elevation sites (sites 15, 17, 19). The same patterns  
365 hold true for nucleotide diversity, which had overall the weakest correlation with latitude ( $R = -$   
366  $0.57$ ,  $p = 0.0082$ ).

367  
368 Observed heterozygosity per sampling site was always greater than expected and therefore there is  
369 heterozygote excess (negative  $F_{IS}$ ), as can be expected from a self-incompatible outcrossing  
370 species. Pairwise subpopulation differentiation ( $F_{ST}$ ) values were all less than 0.036 (Fig 3) and



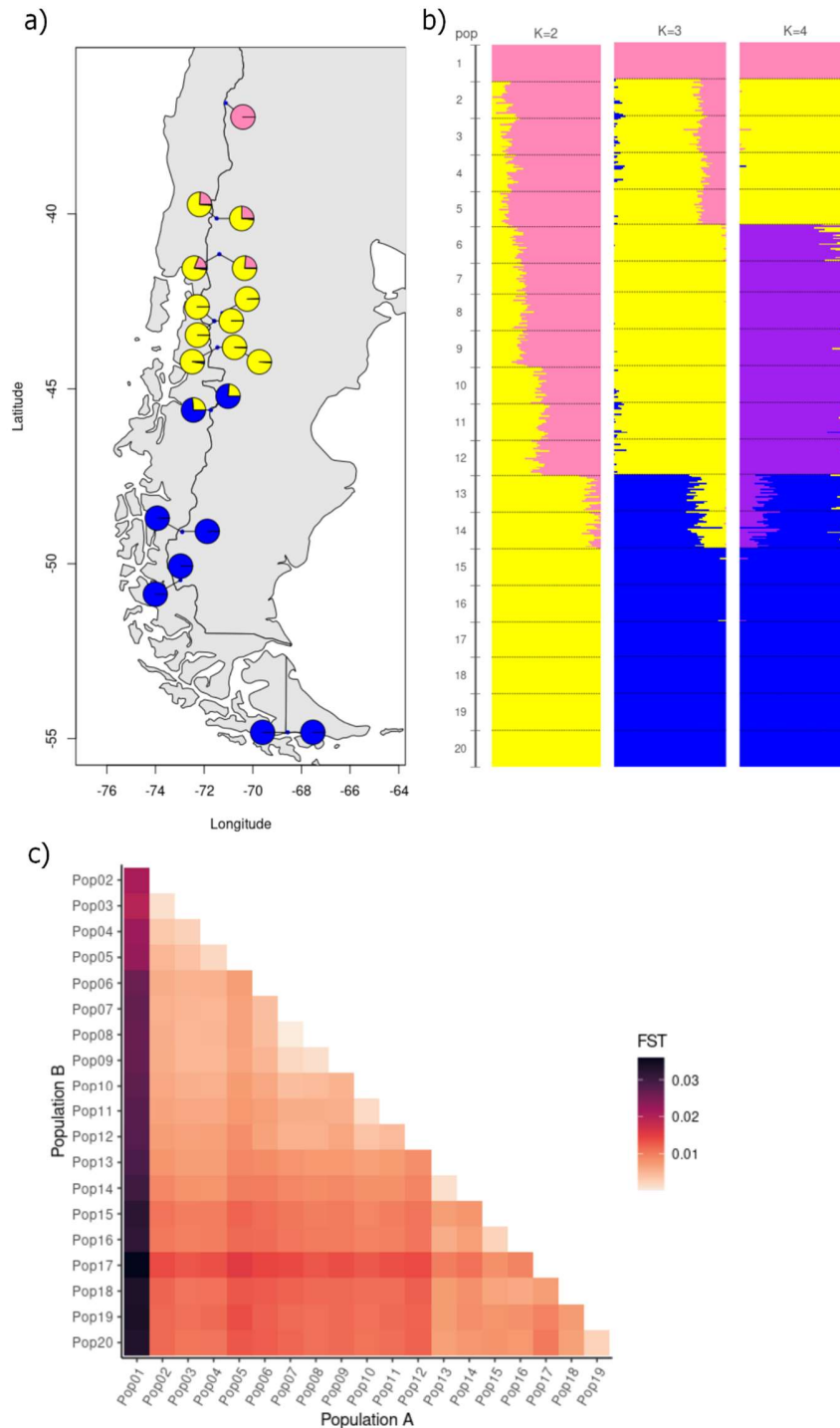
371 had a mean of 0.0102. Paired sites generally had the lowest pairwise  $F_{ST}$  values, with most ranging  
372 from 0.00024 (sites 7 vs. 8) to 0.0025 (19 vs. 20). The notable exception is sites 17 and 18, which  
373 have a value of 0.0072. Singleton locality values in relation to all sampling sites ranged from 0.0025  
374 to 0.0359. Site 1 (Epulaufquen) consistently had the highest pairwise values with all other  
375 populations (range 0.0196 - 0.0359). Epulaufquen also had the greatest endemic diversity (number  
376 of private alleles), and then the values sharply decreased poleward.



377  
378  
379  
380  
381  
382  
383  
384

**Figure 2. Genetic diversity statistics and their correlations with latitude:** a) expected heterozygosity, b) nucleotide diversity including all invariant called sites, c) observed heterozygosity, d) rarefied private allele count, and e) fixation index  $F_{IS}$ . Colors indicate site number and shapes indicate relative elevation class (high, middle, and low within paired localities, or singleton). R-squared and p-values for linear regression models are included in each graph.

385 Population structure is also oriented along the latitudinal gradient, although the exact number of  
386 genetic clusters is ambiguous. According to cross-validation values in ADMIXTURE, the optimal  
387 number of genetic clusters (K) is 2. However, principal component and snmf (LEA package)  
388 analyses suggested that the optimal number is 3 (Fig 4a). We present K values from 2-4 (Fig 4b),  
389 since all are informative about the hierarchical population structure (Meirmans, 2015). Across K-  
390 values, a break consistently occurs between sites 12 and 13 (i.e. between 43.8 - 45.6 °S), with  
391 admixture appearing in sites 13 and 14. At K=3, sites 2-5 show admixture, and at K=4 this region  
392 becomes its own cluster, with a break between sites 5 and 6 (41.2 - 42.8 °S). The singleton  
393 northernmost site 1 (Epulaufquen) also isolates into its own cluster at K=4.



394  
395 **Figure 3. Population genetic structure of *Nothofagus pumilio*.** a) Average cluster per sampling site using  $K = 3$   
396 results from ADMIXTURE. b) Individual ADMIXTURE plots for  $K$  values of 2, 3, and 4, ordered top to bottom from  
397 north (sampling site 1) at the top to south (20) at the bottom. Each line indicates one individual and colors indicate  
398 population clusters. (c) Pairwise genetic difference between and among sampling sites as assessed with Weir and  
399 Cockerham pairwise  $F_{ST}$  values. Sampling sites are ordered from North (site 1) at top and left to South (20) at bottom  
400 and right. Color indicates  $F_{ST}$  value, from small (light orange) to large (dark purple).

## 401 **Climate conditions per site**

402 Climate covariates also show strong gradients along the latitude axis, and many have strong within-  
403 locality gradients (Fig S1). Higher-elevation sites experience more frost days and greater total  
404 annual precipitation, although the majority is received as snow in winter (Veblen et al., 1996, Fig  
405 S1b & 1d). Northern localities generally have higher temperatures and annual precipitation  
406 amounts than central and southern localities, although the low-latitude Epulauquen locality (site  
407 1) is among the drier localities. The El Chaltén (sites 15-16) and El Calafate (17-18) localities are  
408 directly adjacent to the Southern Patagonian icefield, where most precipitation falls as snow even  
409 in summer, and therefore they are among the coldest and driest sites overall. The southernmost  
410 locality Ushuaia (19-20) has a maritime climate, with relatively warmer temperatures and higher  
411 precipitation than other poleward localities. The two geographically marginal singleton localities  
412 in the central portion of the range (6-12) have lower precipitation and more frost days than their  
413 counterpart paired localities at similar latitudes (sites 7-9 and 10-11, respectively), demonstrating  
414 that they are also environmentally marginal. A PCA indicated 81% of the variance can be described  
415 by two axes, the first mainly comprising annual precipitation and day length in January, and the  
416 second comprising number of frost days and growing season precipitation (Fig S1f).

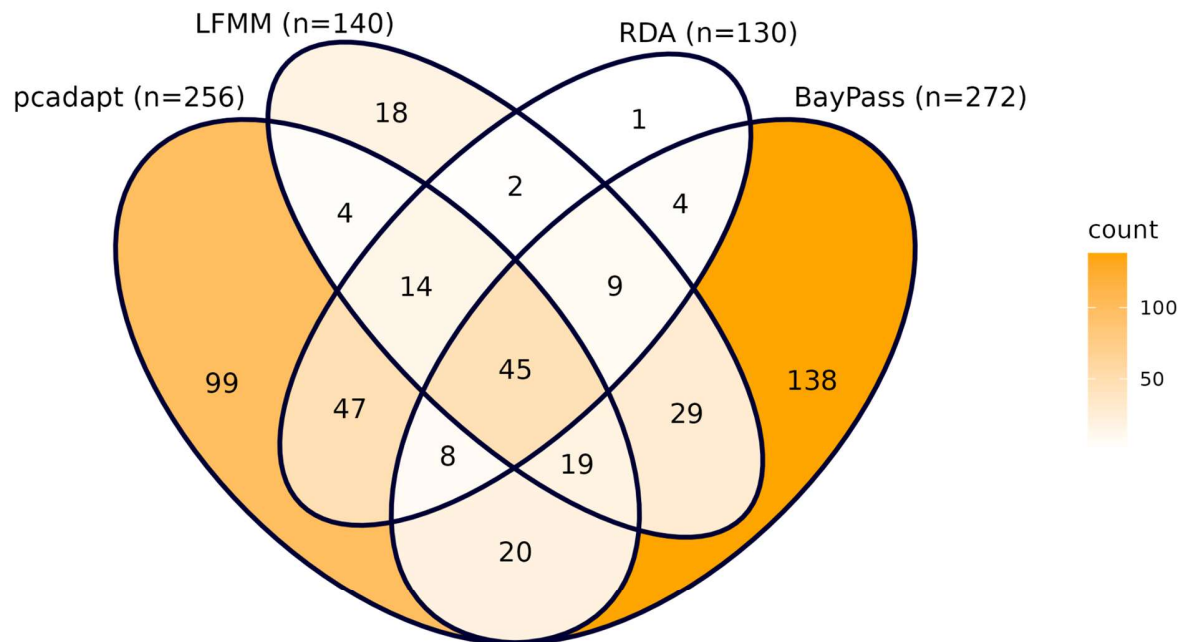
417

## 418 **Genes containing SNP outliers**

419 A total of 457 SNPs in 329 contigs (5.2% of analyzed SNPs) were identified as outliers by at least  
420 one genome scan method under a false discovery rate threshold of 1% (Table S1). BayPass  
421 identified the greatest number of outliers (n=272) and also had the greatest number of unique  
422 outliers (Fig 4). RDA identified the fewest outliers (n=130), and only one outlier was unique to this  
423 analysis. Of the RDA outliers, 114 were also found by the other multivariate-genetic method,  
424 pcadapt. Meanwhile, pcadapt found 99 unique outliers that were not identified by any other method.

425

426 Of all outliers, 201 SNPs were identified by at least two methods, and 45 of these were identified  
427 by all four algorithms (Fig 4). These two outlier lists will be called moderate-evidence and strong-  
428 evidence outliers (i.e. for being true positives), respectively. The majority of moderate- and strong-  
429 evidence outliers are located within coding regions of annotated target genes (125 annotated of 201  
430 total outliers, Table S1). The GO enrichment analysis with PANTHER indicated that no gene  
431 functions were significantly overrepresented in relation to the background starting gene list.

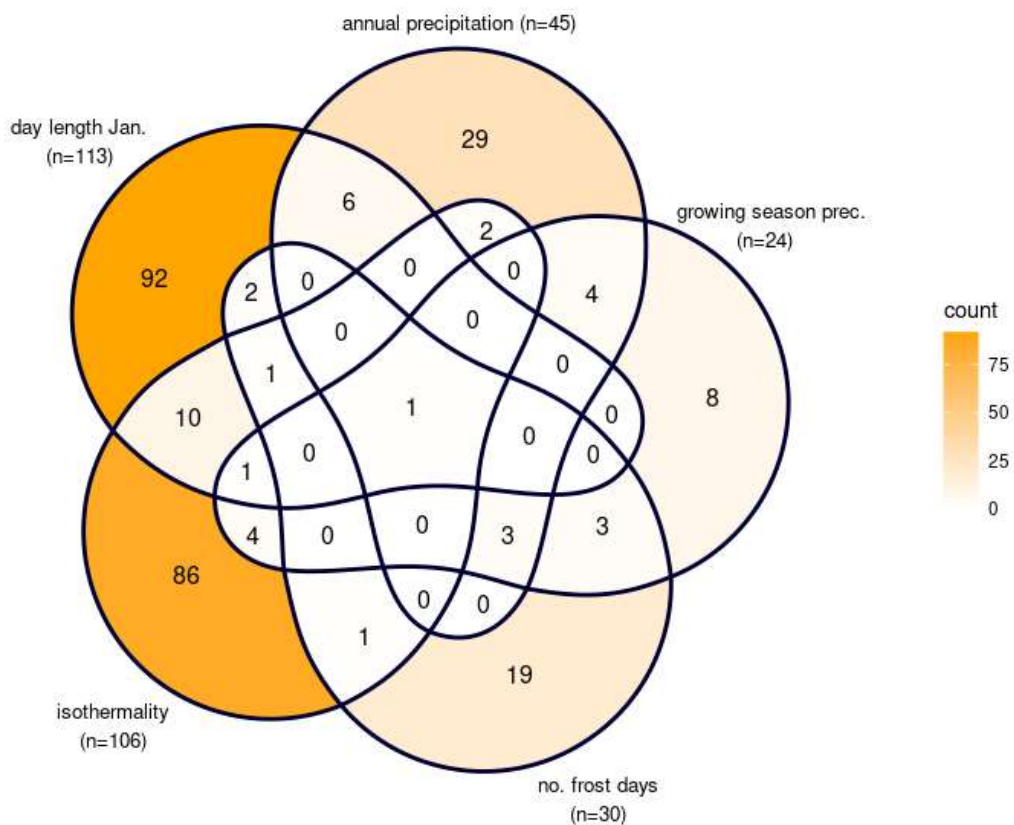


432  
433 **Figure 4. Overlap among significant outlier SNPs found by each genome scan method.** Numbers inside each Venn  
434 section indicate the number of SNP(s) found by the method(s), and color also indicates count, from low (white) to high  
435 (dark orange). Total number of SNPs found by a method is shown in parentheses after the method name.  
436

437 BayPass is a univariate method for both predictor and response variables, so it is possible to identify  
438 individual predictors per SNP (Fig 5). As an aside, LFMM could have been used in univariate-  
439 environment mode, but we chose to use its multivariate configuration to reflect the real-world  
440 dimensionality of environmental covariates. Day length in January and isothermality were  
441 significantly associated with the most outliers (n=113 and 106, respectively). Among the 45 strong-  
442 evidence outliers, all but 3 were associated with one or both of these covariates. Growing season  
443 precipitation was significantly associated with the fewest outliers (n=24). Thirty-eight of the  
444 BayPass outliers were associated with more than one covariate, and one SNP was significantly  
445 associated with all five covariates. An example of a strong-evidence SNP that associated with more  
446 than one environmental covariate is sequence “chain\_2392, locus 1061” (Table 2). The allele cline



447 is plotted against the three environment clines with which it associates in Figure 6. The allele  
448 frequency cline also demonstrates within-locality allele differences, most noticeably in El Calafate  
449 (sites 17-18). This SNP is located within the MYC2 gene, a transcription factor that may be  
450 involved in light signaling pathways and abscisic acid signaling pathways, which is related to  
451 drought stress response (Yadav, Mallappa, Gangappa, Bhatia, & Chattopadhyay, 2005).

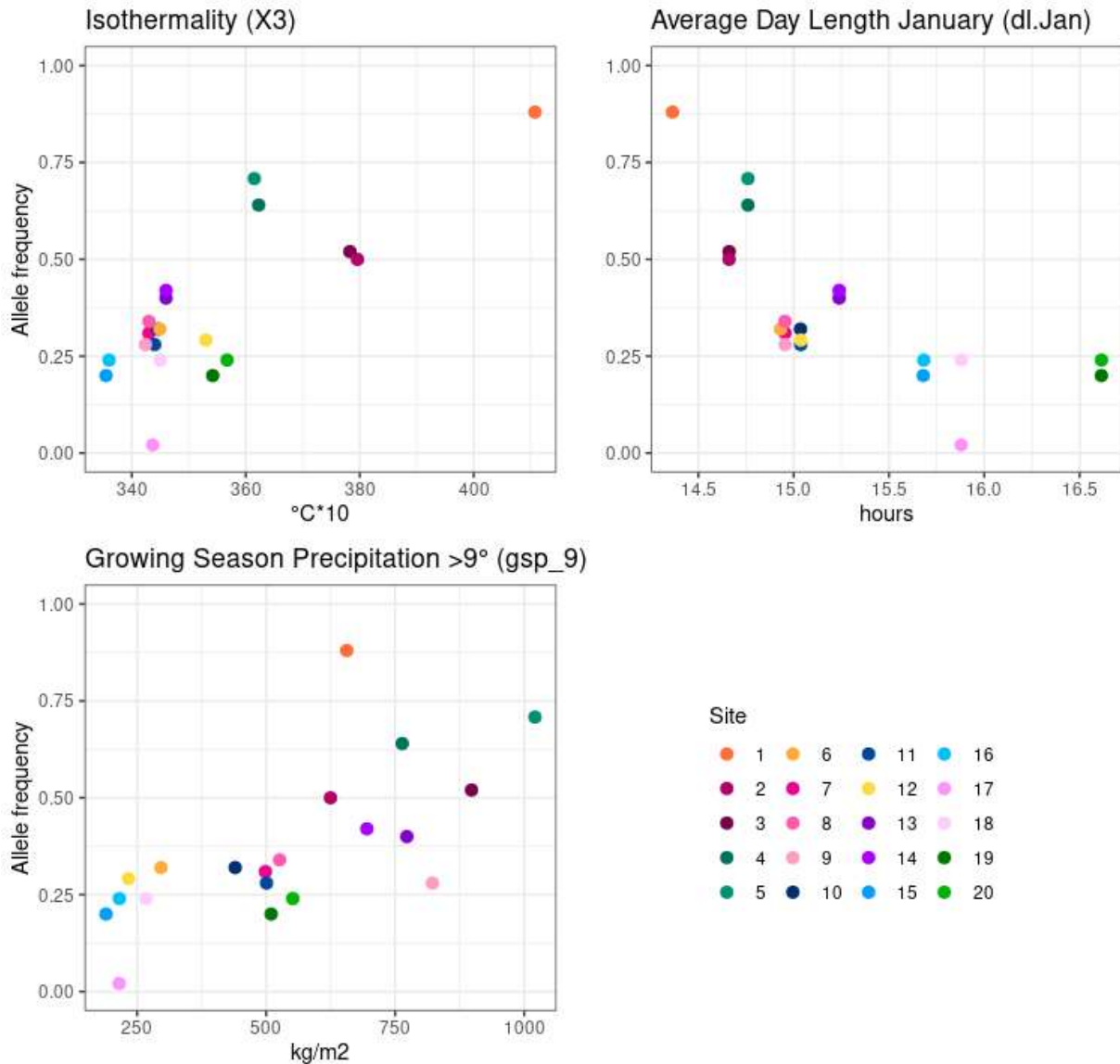


452  
453  
454 **Figure 5. Number of outlier SNPs significantly associated with individual covariates in the univariate method**  
455 **BayPass.** Threshold for significance is Bayes Factor > 10. Color indicates count, from low (white) to high (dark  
456 orange). Total number of associated SNPs is shown in parentheses after the covariate name. Numbers inside each Venn  
457 section indicate the number of SNP(s) associated with the covariate(s).

Table 2. Gene function summary for a subset of the strongest evidence outliers within the synthesis-metabolism and stress response protein function groups. UniProt accession code, protein name, and function were sourced from the UniProt database. Q-values for each GEA method are shown, as are all significant Bayes Factor values for environmental covariates that had significant q-values in BayPass. For further details on these and other outliers, see Supplemental Table 1.

Function group	UniProt	Protein name	Type - Function	GEA method q-values					Environmental covariate Bayes Factors				
				pcadapt	LFMM2	RDA	BayPass	NFD	Ann Rec	Isotherm	GSP 9	DLJ	
Synthesis	B3RFJ6	CYP86A22	Cytochrome - lipid metabolism	2.12E-20	6.97E-11	8.51E-07	0.00E+00			18.412	12.558		
	Q9ZNZ7	GLU1	glutamate synthase – ammonium metabolism	1.19E-08	2.89E-04	3.03E-05	2.40E-09			20.043			
	Q9LV03	GLT1	glutamate synthase	1.19E-08	2.89E-04	3.03E-05	3.32E-09			25.204			
	Q9LIB2	PHS1	alpha-1,4 <del>glucan</del> phosphorylase L-2 isozyme	2.32E-07	1.61E-03	4.16E-03	0.00E+00					15.481	
	Q8LFK6	DUF1195	sugar transporter, putative	5.85E-07	5.04E-08	2.18E-04	3.86E-09					26.339	
Stress Response	Q9C8E7	GLR3.3	glutamate receptor – Biotic stress	1.69E-18	1.66E-05	1.23E-06	8.18E-13			12.491			
	Q9XJ74	PUMP3	Protection against oxidative stress damage	9.50E-15	3.26E-06	2.26E-06	0.00E+00			17.959			
	O22264	MYB12	Flavonoid biosynthesis, stress (UV-B)	4.27E-10	2.97E-10	2.34E-05	0.00E+00			52.964			
	Q3YL57	NHX8	Salt tolerance	1.29E-09	5.30E-05	2.43E-05	2.63E-05			23.111			
	Q39204	MYC2	transcription factor MYC2 – Biotic stress	3.44E-08	1.07E-07	7.33E-04	2.86E-12			19.852	18.358	12.109	

459



460  
461 **Figure 6. Example allele frequency clines for a single SNP locus that is significantly associated with multiple**  
462 **covariates.** The sequence is “contig “chain\_2392,” nucleotide position 1061. The covariates are isothermality, average  
463 day length in January, and growing season precipitation. Colors indicate sampling site. This SNP is located within gene  
464 MYC2. Allele frequencies are not corrected for population genetic structure.

## 465 Discussion

466 Climate change will likely disrupt relationships among environmental covariates, thereby exerting  
467 unique selection pressures on trees such as the Andean foundation species *Nothofagus pumilio*. We  
468 used a landscape genomics approach to determine which genes show signatures of adaptation and

469 which environmental factors might be influencing selection. In particular, we investigated  
470 environmental covariates related to temperature, day length, and precipitation, both as univariate  
471 and multivariate factors, since interplay among covariates can affect biological processes such as  
472 phenology and drought response. We found that population structure and genetic diversity in *N.*  
473 *pumilio* are mainly structured along the major north-south spine of the Andes, likely due to past  
474 glacial cycles. Temperature and day length variables were significantly associated with the greatest  
475 numbers of outlier SNPs, while precipitation variables were associated with fewer. However, many  
476 outliers were identified only by multivariate analyses or were associated with more than one  
477 variable, suggesting genes are responding to a combination of environmental covariates. Climate  
478 change may have a dire impact on *N. pumilio* survival and adaptation, particularly if it decouples  
479 relationships among environmental selection pressures to which these genes are currently adapted.

#### 480 **Latitude-oriented population structure and genetic diversity patterns reflect** 481 **glacial oscillations**

482 The latitudinal orientation of the hierarchical population structure (Fig 2) is likely due to glacial  
483 oscillations. Our population structure analyses ( $K = 2-4$ , Fig 4) indicate the first major cluster  
484 division occurs at mid-latitudes between 43–45 °S (sites 12 and 13). Previous neutral marker  
485 studies also found evidence for two geographically segregated chloroplast and microsatellite  
486 lineages, namely a northern and a southern clade (Mathiasen & Premoli, 2010; Mattera, Pastorino,  
487 Lantschner, Marchelli, & Soliani, 2020; Soliani, Gallo, & Marchelli, 2012; Soliani et al., 2015).  
488 Those analyses identified the major division slightly further north, near 42 °S (Mathiasen &  
489 Premoli, 2010; Mattera et al., 2020), or between 42 - 44 °S (Soliani et al., 2015). At  $K = 4$ , we also  
490 observed a division there, specifically between 41.1 - 42.8 °S (sites 5 and 6). This phylogenetic  
491 divide has also been observed in other regional taxa (e.g. Sersic et al., 2011) and has been attributed  
492 to divergent glacial patterns. Glaciers north of 41 °S were alpine-style and restricted to valleys,  
493 while glaciers south of 45 °S formed a more continuous ice sheet (Glasser, Jansson, Harrison, &  
494 Kleman, 2008). Palynological and genetic data indicate *Nothofagus* species responded to these  
495 glaciations by migrating towards more favorable northern latitudes (Villagran, 1990) and by  
496 retreating to refugia in various parts of the range (Markgraf, 1993).

497  
498 The distribution of refugia and subsequent postglacial expansion have helped shape the genetic  
499 landscape of *N. pumilio*. In general, high heterozygosity is expected in regions near refugia (Petit

500 et al., 2003; Roberts & Hamann, 2015) and also in admixture zones along recolonization routes  
501 where secondary contact occurred (Soliani et al., 2015). We found higher heterozygosity values in  
502 the north, providing evidence for more northern refugia than southern. However, we found no  
503 elevated heterozygosity near the mid-latitude population cluster divisions, even though admixture  
504 in sites 13 and 14 (El Triana), which suggests this region was a secondary contact zone. It is  
505 possible that this lack of elevated heterozygosity is due to unobserved evolutionary forces such as  
506 genetic drift or bottlenecks. Assuming a similar mutation rate across populations, higher nucleotide  
507 diversity in the north may also reflect larger historical effective population sizes there (Nei &  
508 Takahata, 1993). This would be a likely consequence of northward migration patterns and denser  
509 refugia distribution. Greater endemic diversity (private alleles) in the north also support the idea of  
510 enduring populations that may have been temporarily isolated from other populations. Notably,  
511 southern sampling sites also had at least 250 private alleles each, supporting the claim that there  
512 were multiple southern refugia in previous Glacial Maxima (Paula & Leonardo, 2006; Andrea C  
513 Premoli, Mathiasen, & Kitzberger, 2010). Further evidence regarding postglacial expansion  
514 patterns comes from gene flow and migration speed estimates. Species-specific gene flow  
515 information is limited, although it has been postulated that wind-dispersed pollen can travel 10-100  
516 times further than the anemochorous seeds (Mathiasen & Premoli, 2010). High male-gametic gene  
517 flow could explain our low overall population differentiation values, which have a mean of 0.0102  
518 (Fig 3). On the other hand, a prior estimate of migration speed suggests that the maximum  
519 postglacial expansion distance following the Last Glacial Maximum is no more than 800 km  
520 (Mathiasen & Premoli, 2010).

521 There were no significant range-wide elevation patterns among diversity statistics, although local  
522 gradients and regions showed trends. High-elevation sites in poleward localities generally had  
523 lower diversity, which aligns with previous studies that found high-elevation populations of *N.*  
524 *pumilio* have reduced polymorphism due to recent postglacial expansion, genetic drift, and/or  
525 inbreeding (A C Premoli, 2003). Suboptimal climate conditions at higher elevations may also be at  
526 play (Mathiasen & Premoli, 2013), for example the increased number of frost days (Fig S1e). In  
527 contrast, high-elevation sites in low-latitude localities generally had greater diversity. A possible  
528 explanation is that our “high-elevation” sites were located below local treelines, meaning they  
529 could be more accurately classified as intermediate elevation. Meta-analyses have shown

530 intermediate elevation areas have high genetic diversity due to locally optimal environmental  
531 conditions and larger effective population sizes (Ohsawa & Ide, 2008).

532 A peculiar case is observed in the northernmost site, Epulaufquen (site 1), which had the highest  
533 number of private alleles and ubiquitously high pairwise  $F_{ST}$  values, but also had relatively low  
534 observed heterozygosity and nucleotide diversity (Fig 2). Despite strong genetic differentiation  
535 from other sites (Fig 3c), this site only segregates into its own population cluster at  $K=4$  (Fig 3b).  
536 This may be due to its relatively small sampling size (e.g. Rosenberg et al., 2002) in comparison to  
537 the paired localities, or there is relatively stronger differentiation between north and south clusters  
538 at the range-wide level. Epulaufquen is an ecologically and topographically unique site, and related  
539 species with populations there such as *Nothofagus obliqua* also display distinct genetic and  
540 morphological characteristics (Azpilicueta et al., 2014). The forest is located on the valley floor  
541 and is surrounded by mountain peak barriers that likely impede migration and gene flow, which  
542 may explain the high  $F_{ST}$  values and private allele count. Epulaufquen is near the species' current  
543 low-latitude range margin, which are usually areas that experience higher temperatures and less  
544 precipitation in comparison to the rest of the range and are thus exposed to increased drought risk  
545 (Hampe & Petit, 2005). This locality may harbor alleles that are adapted to extreme conditions, but  
546 it is also possible that this locality will not remain a suitable habitat for *N. pumilio* under warmer  
547 conditions.

## 548 **Evidence sources**

549 Univariate and multivariate genome scan methods provide different but equally valuable  
550 information about signatures of local adaptation. Univariate genetic methods can identify large-  
551 effect alleles that may be under strong selection, and univariate environment methods can identify  
552 the drivers exerting the strongest selection pressures. Meanwhile, multivariate genetic methods can  
553 identify small-effect alleles, and multivariate environment methods can characterize interplay  
554 among environmental factors. While multivariate environment methods may be more realistic, an  
555 important caveat is that they combine covariates into principal components, which are inherently  
556 difficult to interpret (Rellstab et al., 2015). In this case, it can be more informative to examine the  
557 biological functions of the outlier-containing genes. Therefore, we used a combination of univariate  
558 and multivariate results and the biological processes of outlier-containing genes to characterize  
559 selection pressures.



560 PCAdapt identified 99 outliers that were not indicated by any of the GEA analyses. Population  
561 differentiation methods may have more power than GEA analyses when demographic history has  
562 caused collinearity between neutral allele frequencies and environmental clines (Lotterhos &  
563 Whitlock, 2015), for example in a post-glacial orographic habitat like Patagonia. This explanation  
564 is supported by our results, which indicate that genetic diversity statistics (Fig 2), population  
565 structure (Fig 3), and environmental covariates (Fig S1) all had strong relationships with the  
566 latitude cline. However, allele frequencies do differ among sites within localities (Fig 8),  
567 demonstrating that latitude is not the only axis along which adaptation occurs. Therefore, it is  
568 possible that these 99 putatively adaptive regions are associated with unobserved climatic factors  
569 or they may be influenced by factors beyond climate (e.g. Meirmans, 2015). The pcadapt-unique  
570 outliers may be worth investigating in further studies.

### 571 **Temperature and day length most important environmental covariates**

572 In the univariate GEA test BayPass, we found notably more SNP outliers that significantly  
573 associated with a temperature (n=133) and/or photoperiod (n=113) covariate than with a  
574 precipitation covariate (n = 61) (Fig 6, Table S1). Temperature and photoperiod parameters may  
575 have a stronger effect than precipitation does, an interpretation that aligns with some previous  
576 studies of *N. pumilio* adults and seedlings. In adults, significant correlations were found between  
577 radial growth and temperature but not precipitation (Castellano, Srur, & Bianchi, 2019). In  
578 seedlings, temperature was also a stronger factor than air humidity for mortality rate along elevation  
579 clines (Cagnacci et al., 2020). Our study is the first to explicitly assess photoperiod in relation to  
580 *N. pumilio* genetic adaptation and the 113 day length-associated outliers suggest that photoperiod  
581 has a strong effect. Since empirical information is limited for *N. pumilio*, we turn to supporting  
582 evidence from other tree species. Delayed spring bud-burst in response to short photoperiod has  
583 been observed in related late-successional species *Fagus sylvatica* and *Quercus petraea* under  
584 common garden conditions (Basler & Körner, 2012; Vitasse & Basler, 2013). Similarly, autumnal  
585 dormancy in perennial plants is largely initiated by the environmental cues of shortened  
586 photoperiod and low temperature (Howe et al., 1995; Singh et al., 2017); for example in *Populus*  
587 species, dormancy-related genes (phytochromes) have shown stark latitude clines (Ingvarsson,  
588 Garcia, Hall, Luquez, & Jansson, 2006). These results highlight the importance of temperature,  
589 photoperiod, and their interaction in regards to phenology and adaptation.

590

591 Fifteen outliers are associated with both day length and at least one temperature covariate, and the  
592 containing genes are predominantly related to stress response (Table S1). This temperature-  
593 photoperiod overlap value is somewhat lower than expected, considering that both temperature and  
594 photoperiod clines are mainly oriented along the latitude axis (Fig S1), but covariate pruning is an  
595 important consideration when examining these results. We chose isothermality and number of frost  
596 days for the temperature variables due to their biological relevance, but also in part because their  
597 correlations with January day length were moderate ( $R = -0.44$  and  $R = -0.39$ , respectively; Fig  
598 S1). However, some pruned temperature parameters had much stronger correlations with day length  
599 ( $R > 0.9$ , data not shown). Environmental associations must always be carefully interpreted (e.g.  
600 Rellstab et al., 2015), and outliers that we found to be associated with photoperiod might actually  
601 be associated with unobserved but highly correlated temperature variable(s) instead of, or in  
602 addition to, photoperiod. In any case, unobserved relationships may well be disrupted by global  
603 climate change, meaning the outlier-containing genes will still experience different selection  
604 pressures than those under which they evolved.

605  
606 A smaller number of precipitation-associated outliers were also identified ( $n=61$ ), and the  
607 containing annotated genes have diverse functions including metabolism, growth, and stress  
608 response (Table S1). However, only one of these annotated genes is directly related to drought  
609 response (probable protein phosphatase 2C 24). A notable point is that 11 annotated outlier-  
610 containing genes were assigned Gene Ontology terms related to water deprivation response, but  
611 the SNPs were not significantly associated with either precipitation covariate (e.g. genes *NCED1*,  
612 *BFRUCT3*). Possible reasons for the lower count of precipitation-associated outliers may include  
613 sampling locations or *N. pumilio* biology. Regarding the study design, our sampling area  
614 encompassed the drier portion of the species' range, namely in Argentina. Our sampling sites  
615 received between 380-1600 mm mean total annual precipitation between 1979-2013, but the  
616 species also inhabits Chilean locations that received up to 5500 mm of mean precipitation during  
617 those years (Figs 1, S1). It is possible that our sampled range was insufficient to capture more  
618 precipitation-related signatures of adaptation. From a biology standpoint, common garden studies  
619 performed on young *N. pumilio* that were collected along local precipitation or elevation gradients  
620 consistently show trait differentiation in water use and morphology when those plants are grown  
621 under drought conditions, but there is little consensus about whether genetics or phenotypic  
622 plasticity is responsible. Some studies suggest a genetic basis (Ignazi, Bucci, & Premoli, 2020;

623 Soliani & Aparicio, 2020), others suggested that responses are plastic (Ivancich et al., 2012), and  
624 still others found supporting evidence for both mechanisms (Mathiasen & Premoli, 2016; Andrea  
625 C Premoli & Brewer, 2007; Soliani et al., 2012). Phenotypic plasticity is advantageous when  
626 physical conditions are highly variable, for instance in northern Patagonian locations with a  
627 Mediterranean climate (Villalba et al., 2003). Plasticity can allow plants to evade temporary  
628 suboptimal conditions, but sidestepping the selection pressures required for adaptation means fewer  
629 loci may show adaptive signatures. Joining the common garden evidence with our genome scan  
630 results, it seems likely that water use and drought response may be complicated processes that  
631 implicates many biological processes and could also incorporate plasticity.

632  
633 The critical question of how interplay among environmental covariates affects selection has  
634 complicated answers. Interactions between precipitation and temperature have been noted as  
635 important for *N. pumilio* survival and radial growth (Lara et al., 2001) as well as seedling  
636 establishment patterns at alpine treelines (Daniels & Veblen, 2004). Interactions among day length  
637 and temperature are also known to be important in the annual growth cycle and phenology (Singh  
638 et al., 2017). Of the 272 BayPass outliers, 38 were associated with more than one univariate  
639 environmental factor. We explicitly assessed environmental interplay through two multivariate  
640 environmental analysis methods, which identified 40 annotated SNPs that were not identified by  
641 the univariate method. While synthetic variables used in multivariate analyses can represent  
642 environment space more realistically, they are inherently difficult to interpret biologically.  
643 Therefore, it is informative to examine the biological functions of the outlier-containing genes.

644

## 645 **Biological functions of outliers mainly related to stress and metabolism**

646 The tradeoff between growth and survival is a fundamental challenge for plants, particularly when  
647 they undergo stress events such as high temperature or drought. Nearly half of all annotated outliers  
648 were located in genes related to stress response (n=38, 21%) or metabolism-catabolism synthesis  
649 (n=43, 24%) (Table S1). A previous study regarding *N. pumilio* transcriptome expression under  
650 heat stress found that genes related to photosynthesis and carbon metabolism were down-regulated  
651 under heat treatment, while stress response genes were up-regulated (Estravis-Barcala et al., 2021).  
652 Many of our outlier-containing genes were similar in form or function to those identified in the  
653 heat stress study. For example, we found 6 related outlier-containing genes related to the ABA

654 signaling pathway, which was up-regulated by heat stress. ABA is produced under water deficit  
655 and confers tolerance to water and salt stress (Abe et al., 2003). Among our ABA-related outliers  
656 were CBL9 and CCD1, which had no univariate associations, and MYC2, which is also a regulator  
657 of light and jasmonic acid (Yadav et al., 2005) and was associated with all three major variables  
658 (Fig 8). Taken together, this provides evidence that the ABA pathway experiences diverse selection  
659 pressures. Another family of genes, the WRKY transcription factors, contained multiple outliers  
660 and was also up-regulated under heat stress (Estravis-Barcala et al., 2021). While the exact genes  
661 differed between studies, the genes' functions were comparable, namely defense response to  
662 drought or pathogens (WRKY 4, associated with day length and temperature, and 33, associated  
663 with day length).

664 Nested within the growth-survival tradeoff is phenology, which also has implications for  
665 reproduction and future adaptation. Growth and development comprised 16 genes (9%) and  
666 included phenological functions such as flowering time. PHYA is a critical component for  
667 flowering time expression through light perception (Yanovsky & Kay, 2002) and a PHYA homolog  
668 in *Populus* trees has also been linked to autumnal growth cessation and bud set (Böhlenius et al.,  
669 2006). Our results indicate this gene is under selection in *N. pumilio* and is associated with day  
670 length. Other flowering time genes, like PFT1, are not associated with any covariates. In the  
671 closely-related species *Fagus sylvatica*, spring bud burst phenology is also photoperiodically  
672 controlled (Vitasse & Basler, 2013). A disconnect between day length and temperature could mean  
673 a longer growing season that trees could not take advantage of, particularly those at higher  
674 elevations whose growing-season window is already condensed (Barrera, Frangi, Richter,  
675 Perdomo, & Pinedo, 2000).

## 676 **Synthesis and outlook**

677 Climate change could decouple photoperiod from contemporary temperature and precipitation  
678 patterns, thus placing novel selection pressures upon genes related to phenology and stress  
679 response. Climate change has already exacerbated ENSO effects, including acute drought stress,  
680 and the Patagonian precipitation gradient is expected to sharpen further (Barros et al., 2015).  
681 Populations of *N. pumilio* are adapted to current specific patterns and combinations of  
682 environmental variables, and as these patterns and combinations shift, the adaptations could  
683 become maladaptations. A promising avenue for predicting climate change impact is to quantify

684 genomic offset (reviewed in Capblancq, Fitzpatrick, Bay, Exposito-Alonso, & Keller, 2020), which  
685 characterizes mismatch between extant allele compositions and those that would be required under  
686 future conditions. In order to evaluate genomic offset under no-analog conditions, it must first be  
687 established that local adaptation clines are associated with environmental clines, as we have shown  
688 here.

689 We found evidence that patterns of local adaptation in *N. pumilio* are associated with temperature,  
690 day length, and, to a lesser univariate extent, precipitation. Individual genes related to stress  
691 response and development were often associated with temperature and day length. This supports  
692 our prediction (i) regarding growth and stress response genes being correlated with temperature  
693 and photoperiod. However, against our prediction (ii), genes related to drought response were only  
694 rarely associated with precipitation covariates, and were more likely to be associated with  
695 temperature and/or day length. This suggests a more complex response to drought than precipitation  
696 and genetic factors alone. These results have many implications under future climate change.  
697 Cataloging extant genetic diversity may help us identify ideal candidate genes and populations for  
698 assisted gene flow or migration, and identifying the most influential local selection drivers may  
699 help predict response to climate change.

## 700 **Conclusion**

701 Revealing genetic adaptations to climate is a critical exercise for forests in the face of climate  
702 change, especially if relationships among environmental variables become decoupled. Here, we  
703 investigated relationships among environmental factors and allele frequencies of outlier SNPs in  
704 candidate genes in the Andean foundation species *Nothofagus pumilio*. We found that population  
705 structure and overall genetic diversity have strong relationships with latitude. Temperature and  
706 photoperiod covariates were associated with the greatest number of outlier-containing genes, while  
707 precipitation was associated with fewer, even among genes related to drought response. Our results  
708 suggest that stress response and catabolism-metabolism may be under tighter genetic control than  
709 drought response traits, which could be more plastic. This has great relevance for this species'  
710 ability to survive and thrive into the future.

711  
712  
713

## 714 Supplemental materials

715 Supplemental Document 1

## 716 Sources

- 717 Abe, H., Urao, T., Ito, T., Seki, M., Shinozaki, K., & Yamaguchi-Shinozaki, K. (2003). Arabidopsis  
718 AtMYC2 (bHLH) and AtMYB2 (MYB) function as transcriptional activators in abscisic acid  
719 signaling. *The Plant Cell*, *15*(1), 63–78.
- 720 Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in  
721 unrelated individuals. *Genome Research*, *19*(9), 1655–1664.
- 722 Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment  
723 search tool. *Journal of Molecular Biology*, *215*(3), 403–410.
- 724 Arana, M. V., Gonzalez-Polo, M., Martinez-Meier, A., Gallo, L. A., Benech-Arnold, R. L.,  
725 Sánchez, R. A., & Batlla, D. (2016). Seed dormancy responses to temperature relate to  
726 Nothofagus species distribution and determine temporal patterns of germination across  
727 altitudes in Patagonia. *New Phytologist*, *209*(2), 507–520. <https://doi.org/10.1111/nph.13606>
- 728 Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., ... others. (2000).  
729 Gene ontology: tool for the unification of biology. *Nature Genetics*, *25*(1), 25–29.
- 730 Azpilicueta, M. M., Pastorino, M. J., Puntieri, J., Barbero, F., Martinez-Meier, A., Marchelli, P., &  
731 Gallo, L. A. (2014). Robles in Lagunas de Epulauquen, Argentina: previous and recent  
732 evidence of their distinctive character. *Revista Chilena de Historia Natural*, *87*(1), 1–12.
- 733 Barrera, M. D., Frangi, J. L., Richter, L. L., Perdomo, M. H., & Pinedo, L. B. (2000). Structural  
734 and functional changes in Nothofagus pumilio forests along an altitudinal gradient in Tierra  
735 del Fuego, Argentina. *Journal of Vegetation Science*, *11*(2), 179–188.
- 736 Barros, V. R., Boninsegna, J. A., Camilloni, I. A., Chidiak, M., Magrì, G. O., & Rusticucci, M.  
737 (2015). Climate change in Argentina: trends, projections, impacts and adaptation. *Wiley*  
738 *Interdisciplinary Reviews: Climate Change*, *6*(2), 151–169.
- 739 Basler, D., & Körner, C. (2012). Photoperiod sensitivity of bud burst in 14 temperate forest tree  
740 species. *Agricultural and Forest Meteorology*, *165*, 73–81.
- 741 Beaumont, M. A., & Nichols, R. A. (1996). Evaluating loci for use in the genetic analysis of  
742 population structure. *Proceedings of the Royal Society of London. Series B: Biological*  
743 *Sciences*, *263*(1377), 1619–1626.
- 744 Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and  
745 powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B*  
746 *(Methodological)*, *57*(1), 289–300.
- 747 Bobrowski, M., Weidinger, J., & Schickhoff, U. (2021). Is new always better? Frontiers in global  
748 climate datasets for modeling treeline species in the Himalayas. *Atmosphere*, *12*(5), 543.
- 749 Böhlenius, H., Huang, T., Charbonnel-Campaa, L., Brunner, A. M., Jansson, S., Strauss, S. H., &  
750 Nilsson, O. (2006). CO/FT regulatory module controls timing of flowering and seasonal  
751 growth cessation in trees. *Science*, *312*(5776), 1040–1043.
- 752 Bourgeois, Y. X. C., & Warren, B. H. (2021). An overview of current population genomics methods  
753 for the analysis of whole-genome resequencing data in eukaryotes. *Molecular Ecology*,  
754 *30*(23), 6036–6071.
- 755 Brondizio, E. S., Settele, J., Díaz, S., & Ngo, H. T. (2019). *Global assessment report on biodiversity*  
756 *and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and*  
757 *Ecosystem Services*.



- 758 Burke, K. D., Williams, J. W., Chandler, M. A., Haywood, A. M., Lunt, D. J., & Otto-Bliesner, B.  
759 L. (2018). Pliocene and Eocene provide best analogs for near-future climates. *Proceedings of*  
760 *the National Academy of Sciences*, *115*(52), 13288–13293.
- 761 Cagnacci, J., Estravis-Barcala, M., L'via, M. V., Mart'inez-Meier, A., Polo, M. G., & Arana, M.  
762 V. (2020). The impact of different natural environments on the regeneration dynamics of two  
763 *Nothofagus* species across elevation in the southern Andes. *Forest Ecology and Management*,  
764 *464*, 118034.
- 765 Cai, W., Wang, G., Santoso, A., McPhaden, M. J., Wu, L., Jin, F.-F., ... others. (2015). Increased  
766 frequency of extreme La Niña events under greenhouse warming. *Nature Climate Change*,  
767 *5*(2), 132–137.
- 768 Capblancq, T., Fitzpatrick, M. C., Bay, R. A., Exposito-Alonso, M., & Keller, S. R. (2020).  
769 Genomic prediction of (mal) adaptation across current and future climatic landscapes. *Annual*  
770 *Review of Ecology, Evolution, and Systematics*, *51*, 245–269.
- 771 Capblancq, T., & Forester, B. R. (2021). Redundancy analysis: A Swiss Army Knife for landscape  
772 genomics. *Methods in Ecology and Evolution*, *12*(12), 2298–2309.
- 773 Carbon, S., Ireland, A., Mungall, C. J., Shu, S., Marshall, B., Lewis, S., ... Group, W. P. W. (2009).  
774 AmiGO: online access to ontology and annotation data. *Bioinformatics*, *25*(2), 288–289.
- 775 Carson, A. R., Smith, E. N., Matsui, H., Brækkan, S. K., Jepsen, K., Hansen, J.-B., & Frazer, K. A.  
776 (2014). Effective filtering strategies to improve data quality from population-based whole  
777 exome sequencing studies. *BMC Bioinformatics*, *15*(1), 1–15.
- 778 Castellano, P. L., Srur, A. M., & Bianchi, L. O. (2019). Climate-growth relationships of deciduous  
779 and evergreen *Nothofagus* species in Southern Patagonia, Argentina. *Dendrochronologia*, *58*,  
780 125646.
- 781 Chang, C. C., Chow, C. C., Tellier, L. C. A. M., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015).  
782 Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience*,  
783 *4*(1), s13742–015.
- 784 Daniels, L. D., & Veblen, T. T. (2004). Spatiotemporal influences of climate on altitudinal treeline  
785 in northern Patagonia. *Ecology*, *85*(5), 1284–1296.
- 786 De La Torre, A. R., Wilhite, B., & Neale, D. B. (2019). Environmental genome-wide association  
787 reveals climate adaptation is shaped by subtle to moderate allele frequency shifts in loblolly  
788 pine. *Genome Biology and Evolution*, *11*(10), 2976–2989.
- 789 de Villemereuil, P., Frichot, É., Bazin, É., François, O., & Gaggiotti, O. E. (2014). Genome scan  
790 methods against more complex models: when and how much should we trust them? *Molecular*  
791 *Ecology*, *23*(8), 2006–2019.
- 792 Del Fabbro, C., Scalabrin, S., Morgante, M., & Giorgi, F. M. (2013). An extensive evaluation of  
793 read trimming effects on Illumina NGS data analysis. *PLoS One*, *8*(12), e85024.
- 794 DePristo, M. A., Banks, E., Poplin, R., Garimella, K. V, Maguire, J. R., Hartl, C., ... others. (2011).  
795 A framework for variation discovery and genotyping using next-generation DNA sequencing  
796 data. *Nature Genetics*, *43*(5), 491–498.
- 797 Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., ... others. (2013).  
798 Collinearity: a review of methods to deal with it and a simulation study evaluating their  
799 performance. *Ecography*, *36*(1), 27–46.
- 800 Doyle, J. J. (1990). A rapid total DNA preparation procedure for fresh plant tissue. *Focus*, *12*, 13–  
801 15.
- 802 Duforet-Frebourg, N., Bazin, E., & Blum, M. G. B. (2014). Genome scans for detecting footprints  
803 of local adaptation using a Bayesian factor model. *Molecular Biology and Evolution*, *31*(9),  
804 2483–2495.
- 805 Estravis-Barcala, M., Heer, K., Marchelli, P., Ziegenhagen, B., Arana, M. V., & Bellora, N. (2021).



- 806 Deciphering the transcriptomic regulation of heat stress responses in *Nothofagus pumilio*.  
807 *Plos One*, 16(3), e0246615.
- 808 Estravis-Barcala, M., Mattera, M. G., Soliani, C., Bellora, N., Opgenoorth, L., Heer, K., & Arana,  
809 M. V. (2020). Molecular bases of responses to abiotic stress in trees. *Journal of Experimental*  
810 *Botany*, 71(13), 3765–3779.
- 811 François, O., Martins, H., Caye, K., & Schoville, S. D. (2016). Controlling false discoveries in  
812 genome scans for selection. *Molecular Ecology*, 25(2), 454–469.
- 813 Frichot, E., & Francois, O. (2015). {LEA}: an {R} package for {L}andscape and {E}cological  
814 {A}ssociation studies. *Methods in Ecology and Evolution*. Retrieved from [http://membres-](http://membres-timc.imag.fr/Olivier.Francois/lea.html)  
815 [timc.imag.fr/Olivier.Francois/lea.html](http://membres-timc.imag.fr/Olivier.Francois/lea.html)
- 816 Frichot, E., Schoville, S. D., Bouchard, G., & François, O. (2013). Testing for associations between  
817 loci and environmental gradients using latent factor mixed models. *Molecular Biology and*  
818 *Evolution*, 30(7), 1687–1699.
- 819 Gautier, M. (2015). Genome-wide scan for adaptive divergence and association with population-  
820 specific covariates. *Genetics*, 201(4), 1555–1579.
- 821 Gea-Izquierdo, G., Pastur, G. M., Cellini, J. M., & Lencinas, M. V. (2004). Forty years of  
822 silvicultural management in southern *Nothofagus pumilio* primary forests. *Forest Ecology and*  
823 *Management*, 201(2–3), 335–347.
- 824 Glasser, N. F., Jansson, K. N., Harrison, S., & Kleman, J. (2008). The glacial geomorphology and  
825 Pleistocene history of South America between 38 S and 56 S. *Quaternary Science Reviews*,  
826 27(3–4), 365–390.
- 827 Goudet, J., & Jombart, T. (2022). *hierfstat: Estimation and Tests of Hierarchical F-Statistics*.  
828 Retrieved from <https://cran.r-project.org/package=hierfstat>
- 829 Hampe, A., & Petit, R. J. (2005). Conserving biodiversity under climate change: the rear edge  
830 matters. *Ecology Letters*, 8(5), 461–467.
- 831 Hänninen, H., & Tanino, K. (2011). Tree seasonality in a warming climate. *Trends in Plant Science*,  
832 16(8), 412–416.
- 833 Haynes, K. J., Liebhold, A. M., Lefcheck, J. S., Morin, R. S., & Wang, G. (2022). Climate affects  
834 the outbreaks of a forest defoliator indirectly through its tree hosts. *Oecologia*, 198(2), 407–  
835 418.
- 836 Hedrick, P. W., Ginevan, M. E., & Ewing, E. P. (1976). Genetic polymorphism in heterogeneous  
837 environments. *Annual Review of Ecology and Systematics*, 7(1), 1–32.
- 838 Hijmans, R. J. (2022). *geosphere: Spherical Trigonometry*. Retrieved from [https://cran.r-](https://cran.r-project.org/package=geosphere)  
839 [project.org/package=geosphere](https://cran.r-project.org/package=geosphere)
- 840 Hijmans, R. J. (2023). *raster: Geographic Data Analysis and Modeling*. Retrieved from  
841 <https://cran.r-project.org/package=raster>
- 842 Howe, G. T., Hackett, W. P., Furnier, G. R., & Klevorn, R. E. (1995). Photoperiodic responses of  
843 a northern and southern ecotype of black cottonwood. *Physiologia Plantarum*, 93(4), 695–  
844 708.
- 845 Ignazi, G., Bucci, S. J., & Premoli, A. C. (2020). Stories from common gardens: Water shortage  
846 differentially affects *Nothofagus pumilio* from contrasting precipitation regimes. *Forest*  
847 *Ecology and Management*, 458, 117796.
- 848 Ingvarsson, P. K., García, M. V., Hall, D., Luquez, V., & Jansson, S. (2006). Clinal variation in  
849 phyB2, a candidate gene for day-length-induced growth cessation and bud set, across a  
850 latitudinal gradient in European aspen (*Populus tremula*). *Genetics*, 172(3), 1845–1853.
- 851 Ivancich, H. S., Lencinas, M. V., Pastur, G. J. M., Esteban, R. M. S., Hernández, L., & Lindstrom,  
852 I. (2012). Foliar anatomical and morphological variation in *Nothofagus pumilio* seedlings  
853 under controlled irradiance and soil moisture levels. *Tree Physiology*, 32(5), 554–564.

- 854 Kamvar, Z. N., Tabima, J. F., & Grünwald, N. J. (2014). `poppr`: an R package for genetic  
855 analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ*, 2,  
856 e281. <https://doi.org/10.7717/peerj.281>
- 857 Karger, D. N., Conrad, O., Böhrer, J., Kawohl, T., Kreft, H., Soria-Auza, R. W., ... Kessler, M.  
858 (2017). Climatologies at high resolution for the earth's land surface areas. *Scientific Data*,  
859 4(1), 1–20.
- 860 Kassambara, A. (2022). *ggpubr: "ggplot2" Based Publication Ready Plots*. Retrieved from  
861 <https://cran.r-project.org/package=ggpubr>
- 862 Korunes, K. L., & Samuk, K. (2021). pixy: Unbiased estimation of nucleotide diversity and  
863 divergence in the presence of missing data. *Molecular Ecology Resources*, 21(4), 1359–1368.
- 864 Kremer, A., Ronce, O., Robledo-Arnuncio, J. J., Guillaume, F., Bohrer, G., Nathan, R., ... others.  
865 (2012). Long-distance gene flow and adaptation of forest trees to rapid climate change.  
866 *Ecology Letters*, 15(4), 378–392.
- 867 Lara, A., Aravena, J. C., Villalba, R., Wolodarsky-Franke, A., Luckman, B., & Wilson, R. (2001).  
868 Dendroclimatology of high-elevation *Nothofagus pumilio* forests at their northern distribution  
869 limit in the central Andes of Chile. *Canadian Journal of Forest Research*, 31(6), 925–936.
- 870 Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler  
871 transform. *Bioinformatics*, 25(14), 1754–1760.
- 872 Linck, E., & Battey, C. J. (2019). Minor allele frequency thresholds strongly affect population  
873 structure inference with genomic data sets. *Molecular Ecology Resources*, 19(3), 639–647.
- 874 Lotterhos, K. E., & Whitlock, M. C. (2015). The relative power of genome scans to detect local  
875 adaptation depends on sampling design and statistical method. *Molecular Ecology*, 24(5),  
876 1031–1046.
- 877 Manel, S., Joost, S., Epperson, B. K., Holderegger, R., Storfer, A., Rosenberg, M. S., ... FORTIN,  
878 M.-J. (2010). Perspectives on the use of landscape genetics to detect genetic adaptive variation  
879 in the field. *Molecular Ecology*, 19(17), 3760–3772.
- 880 Markgraf, V. (1993). Paleoenvironments and paleoclimates in Tierra del Fuego and southernmost  
881 Patagonia, South America. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 102(1–2),  
882 53–68.
- 883 Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads.  
884 *EMBnet. Journal*, 17(1), 10–12.
- 885 Mathiasen, P., & Premoli, A. C. (2010). Out in the cold: genetic variation of *Nothofagus pumilio*  
886 (*Nothofagaceae*) provides evidence for latitudinally distinct evolutionary histories in austral  
887 South America. *Molecular Ecology*, 19(2), 371–385.
- 888 Mathiasen, P., & Premoli, A. C. (2013). Fine-scale genetic structure of *Nothofagus pumilio* (*lenga*)  
889 at contrasting elevations of the altitudinal gradient. *Genetica*, 141, 95–105.
- 890 Mathiasen, P., & Premoli, A. C. (2016). Living on the edge: adaptive and plastic responses of the  
891 tree *Nothofagus pumilio* to a long-term transplant experiment predict rear-edge upward  
892 expansion. *Oecologia*, 181(2), 607–619.
- 893 Mattera, M. G., Pastorino, M. J., Lantschner, M., Marchelli, P., & Soliani, C. (2020). Genetic  
894 diversity and population structure in *Nothofagus pumilio*, a foundation species of Patagonian  
895 forests: defining priority conservation areas and management. *Scientific Reports*, 10(1), 1–13.
- 896 McKinney, G. J., Waples, R. K., Seeb, L. W., & Seeb, J. E. (2017). Paralogs are revealed by  
897 proportion of heterozygotes and deviations in read ratios in genotyping-by-sequencing data  
898 from natural populations. *Molecular Ecology Resources*, 17(4), 656–669.  
899 <https://doi.org/10.1111/1755-0998.12613>
- 900 Meirmans, P. G. (2015). *Seven common mistakes in population genetics and how to avoid them*.  
901 Wiley Online Library.

- 902 Milesi, P., Kastally, C., Dauphin, B., Cervantes, S., Bagnoli, F., Budde, K. B., ... others. (2023).  
903 Synchronous effective population size changes and genetic stability of forest trees through  
904 glacial cycles. *BioRxiv*, 2001–2023.
- 905 Morales, M. S., Cook, E. R., Barichivich, J., Christie, D. A., Villalba, R., LeQuesne, C., ... others.  
906 (2020). Six hundred years of South American tree rings reveal an increase in severe  
907 hydroclimatic events since mid-20th century. *Proceedings of the National Academy of*  
908 *Sciences*, 117(29), 16816–16823.
- 909 Nei, M., & Takahata, N. (1993). Effective population size, genetic diversity, and coalescence time  
910 in subdivided populations. *Journal of Molecular Evolution*, 37, 240–244.
- 911 Niinemets, Ü. (2010). Responses of forest trees to single and multiple environmental stresses from  
912 seedlings to mature plants: past stress history, stress interactions, tolerance and acclimation.  
913 *Forest Ecology and Management*, 260(10), 1623–1639.
- 914 Ohsawa, T., & Ide, Y. (2008). Global patterns of genetic variation in plant species along vertical  
915 and horizontal gradients on mountains. *Global Ecology and Biogeography*, 17(2), 152–163.
- 916 Oksanen, J., Simpson, G. L., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., ... Weedon,  
917 J. (2022). *vegan: Community Ecology Package*. Retrieved from [https://cran.r-](https://cran.r-project.org/package=vegan)  
918 [project.org/package=vegan](https://cran.r-project.org/package=vegan)
- 919 Opgenoorth, L., Dauphin, B., Benavides, R., Heer, K., Alizoti, P., Martínez-Sancho, E., ...  
920 others. (2021). The GenTree Platform: growth traits and tree-level environmental data in 12  
921 European forest tree species. *GigaScience*, 10(3), giab010.
- 922 Paula, M., & Leonardo, G. (2006). Multiple ice-age refugia in a southern beech of South America  
923 as evidenced by chloroplast DNA markers. *Conservation Genetics*, 7(4), 591.
- 924 Petit, R. J., Aguinagalde, I., de Beaulieu, J.-L., Bittkau, C., Brewer, S., Cheddadi, R., ... others.  
925 (2003). Glacial refugia: hotspots but not melting pots of genetic diversity. *Science*, 300(5625),  
926 1563–1565.
- 927 Poplin, R., Ruano-Rubio, V., DePristo, M. A., Fennell, T. J., Carneiro, M. O., der Auwera, G. A.,  
928 ... others. (2017). Scaling accurate genetic variant discovery to tens of thousands of samples.  
929 *BioRxiv*, 201178.
- 930 Porebski, S., Bailey, L. G., & Baum, B. R. (1997). Modification of a CTAB DNA extraction  
931 protocol for plants containing high polysaccharide and polyphenol components. *Plant*  
932 *Molecular Biology Reporter*, 15, 8–15.
- 933 Premoli, A. C. (2003). Isozyme polymorphisms provide evidence of clinal variation with elevation  
934 in *Nothofagus pumilio*. *Journal of Heredity*, 94(3), 218–226.
- 935 Premoli, Andrea C, & Brewer, C. A. (2007). Environmental v. genetically driven variation in  
936 ecophysiological traits of *Nothofagus pumilio* from contrasting elevations. *Australian Journal*  
937 *of Botany*, 55(6), 585–591.
- 938 Premoli, Andrea C, Mathiasen, P., & Kitzberger, T. (2010). Southern-most *Nothofagus* trees  
939 enduring ice ages: genetic evidence and ecological niche retrodiction reveal high latitude (54  
940 S) glacial refugia. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 298(3–4), 247–256.
- 941 Privé, F., Luu, K., Vilhjálmsson, B. J., & Blum, M. G. B. (2020). Performing highly efficient  
942 genome scans for local adaptation with {R} package pcadapt version 4. *Molecular Biology*  
943 *and Evolution*. <https://doi.org/10.1093/molbev/msaa053>
- 944 R Core Team. (2022). *R: A Language and Environment for Statistical Computing*. Retrieved from  
945 <https://www.r-project.org/>
- 946 Rellstab, C., Gugerli, F., Eckert, A. J., Hancock, A. M., & Holderegger, R. (2015). A practical  
947 guide to environmental association analysis in landscape genomics. *Molecular Ecology*,  
948 24(17), 4348–4370.
- 949 Revelle, W. (2022). *psych: Procedures for Psychological, Psychometric, and Personality*

- 950 *Research*. Retrieved from <https://cran.r-project.org/package=psych>
- 951 Roberts, D. R., & Hamann, A. (2015). Glacial refugia and modern genetic diversity of 22 western  
952 North American tree species. *Proceedings of the Royal Society B: Biological Sciences*,  
953 282(1804), 20142903.
- 954 Rosenberg, N. A., Pritchard, J. K., Weber, J. L., Cann, H. M., Kidd, K. K., Zhivotovsky, L. A., &  
955 Feldman, M. W. (2002). Genetic structure of human populations. *Science*, 298(5602), 2381–  
956 2385.
- 957 Savolainen, O., Lascoux, M., & Merilä, J. (2013). Ecological genomics of local adaptation. *Nature*  
958 *Reviews Genetics*, 14(11), 807–820.
- 959 Savolainen, O., Pyhäjärvi, T., & Knürr, T. (2007). Gene flow and local adaptation in trees. *Annu.*  
960 *Rev. Ecol. Evol. Syst.*, 38, 595–619.
- 961 Scotti, I., Lalagüe, H., Oddou-Muratorio, S., Scotti-Saintagne, C., Ruiz Daniels, R., Grivet, D., ...  
962 others. (2023). Common microgeographical selection patterns revealed in four European  
963 conifers. *Molecular Ecology*, 32(2), 393–411.
- 964 Sersic, A. N., Cosacov, A., Cocucci, A. A., Johnson, L. A., Pozner, R., Avila, L. J., ... Morando,  
965 M. (2011). Emerging phylogeographical patterns of plants and terrestrial vertebrates from  
966 Patagonia. *Biological Journal of the Linnean Society*, 103(2), 475–494.
- 967 Shen, X., Song, S., Li, C., & Zhang, J. (2022). Synonymous mutations in representative yeast genes  
968 are mostly strongly non-neutral. *Nature*, 606(7915), 725–731.
- 969 Singh, R. K., Svystun, T., AlDahmash, B., Jönsson, A. M., & Bhalerao, R. P. (2017). Photoperiod-  
970 and temperature-mediated control of phenology in trees--a molecular perspective. *New*  
971 *Phytologist*, 213(2), 511–524.
- 972 Soliani, C., & Aparicio, A. G. (2020). Evidence of genetic determination in the growth habit of  
973 *Nothofagus pumilio* (Poepp. & Endl.) Krasser at the extremes of an elevation gradient.  
974 *Scandinavian Journal of Forest Research*, 0(0), 1–10.  
975 <https://doi.org/10.1080/02827581.2020.1789208>
- 976 Soliani, C., Gallo, L., & Marchelli, P. (2012). Phylogeography of two hybridizing southern beeches  
977 (*Nothofagus* spp.) with different adaptive abilities. *Tree Genetics & Genomes*, 8, 659–673.
- 978 Soliani, C., Tsuda, Y., Bagnoli, F., Gallo, L. A., Vendramin, G. G., & Marchelli, P. (2015). Halfway  
979 encounters: Meeting points of colonization routes among the southern beeches *Nothofagus*  
980 *pumilio* and *N. antarctica*. *Molecular Phylogenetics and Evolution*, 85, 197–207.
- 981 The UniProt Consortium, T. U. (2023). UniProt: the Universal Protein knowledgebase in 2023.  
982 *Nucleic Acids Research*, 51(D1), D523–D531.
- 983 Thomas, P. D., Ebert, D., Muruganujan, A., Mushayahama, T., Albou, L.-P., & Mi, H. (2022).  
984 PANTHER: Making genome-scale phylogenetics accessible to all. *Protein Science*, 31(1), 8–  
985 22.
- 986 Veblen, T. T., Donoso, C., Kitzberger, T., & Rebertus, A. J. (1996). Ecology of southern Chilean  
987 and Argentinean *Nothofagus* forests. *The Ecology and Biogeography of Nothofagus Forests*,  
988 10, 93–353.
- 989 Vento, B., & Agraín, F. A. (2018). Phylogenetic relationships and time-calibration of the South  
990 American fossil and extant species of southern beeches (*Nothofagus*). *Acta Palaeontologica*  
991 *Polonica*, 63(4).
- 992 Vicente-Serrano, S. M., Beguería, S., & López-Moreno, J. I. (2010). A multiscalar drought index  
993 sensitive to global warming: the standardized precipitation evapotranspiration index. *Journal*  
994 *of Climate*, 23(7), 1696–1718.
- 995 Villagran, C. (1990). Glacial climates and their effects on the history of the vegetation of Chile: a  
996 synthesis based on palynological evidence from Isla de Chiloé. *Review of Palaeobotany and*  
997 *Palynology*, 65(1–4), 17–24.



- 998 Villalba, R., Lara, A., Boninsegna, J. A., Masiokas, M., Delgado, S., Aravena, J. C., ... Ripalta, A.  
999 (2003). Large-scale temperature changes across the southern Andes: 20th-century variations  
1000 in the context of the past 400 years. *Climate Variability and Change in High Elevation*  
1001 *Regions: Past, Present & Future*, 177–232.
- 1002 Vitasse, Y., & Basler, D. (2013). What role for photoperiod in the bud burst phenology of European  
1003 beech. *European Journal of Forest Research*, 132, 1–8.
- 1004 Waldvogel, A.-M., Schreiber, D., Pfenninger, M., & Feldmeyer, B. (2020). Climate change  
1005 genomics calls for standardized data reporting. *Frontiers in Ecology and Evolution*, 8, 242.
- 1006 Weir, B. S., & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population  
1007 structure. *Evolution*, 1358–1370.
- 1008 Whiteman, C. D. (2000). *Mountain meteorology: fundamentals and applications*. Oxford  
1009 University Press.
- 1010 Williams, J. W., & Jackson, S. T. (2007). Novel climates, no-analog communities, and ecological  
1011 surprises. *Frontiers in Ecology and the Environment*, 5(9), 475–482.
- 1012 Yadav, V., Mallappa, C., Gangappa, S. N., Bhatia, S., & Chattopadhyay, S. (2005). A basic helix-  
1013 loop-helix transcription factor in Arabidopsis, MYC2, acts as a repressor of blue light-  
1014 mediated photomorphogenic growth. *The Plant Cell*, 17(7), 1953–1966.
- 1015 Yanovsky, M. J., & Kay, S. A. (2002). *Molecular basis of seasonal time measurement in*  
1016 *Arabidopsis*. 419(September). <https://doi.org/10.1038/nature01061.1>.  
1017  
1018  
1019