



**HAL**  
open science

## **The impact of Pleistocene glaciations and environmental gradients on the genetic structure of *Embothrium coccineum***

Francisco Sepúlveda-Espinoza, Ariana Bertin-Benavides, Rodrigo Hasbún, Óscar Toro-Núñez, Antonio Varas-Myrik, Diego Alarcón, Marie-laure Guillemin

### ► **To cite this version:**

Francisco Sepúlveda-Espinoza, Ariana Bertin-Benavides, Rodrigo Hasbún, Óscar Toro-Núñez, Antonio Varas-Myrik, et al.. The impact of Pleistocene glaciations and environmental gradients on the genetic structure of *Embothrium coccineum*. *Ecology and Evolution*, 2022, 12 (11), <10.1002/ece3.9474>. <hal-04061309>

**HAL Id: hal-04061309**

**<https://hal.science/hal-04061309v1>**

Submitted on 6 Apr 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

1 **Impact of Pleistocene glaciations and environmental gradients on *Embothrium coccineum* genetic**  
2 **structure**

3

4 Francisco Sepúlveda-Espinoza<sup>1</sup>, Ariana Bertin-Benavides<sup>1-2</sup>, Rodrigo Hasbún<sup>1</sup>, Oscar Toro-Núñez<sup>3</sup>,  
5 Antonio Varas-Myrik<sup>1</sup>, Diego Alarcón<sup>4</sup>, Marie-Laure Guillemin<sup>5-6</sup>

6 <sup>1</sup>Laboratorio de Epigenética Vegetal, Departamento de Silvicultura, Facultad de Ciencias Forestales,  
7 Universidad de Concepción, Victoria 631, Concepción, Chile.

8 <sup>2</sup>ONG Conciencia Sur, Freire 1375 casa C, Concepción, Chile.

9 <sup>3</sup>Departamento de Botánica, Facultad de Ciencias Naturales y Oceanográficas, Universidad de  
10 Concepción, Barrio Universitario s/n, Concepción, Chile.

11

12 <sup>4</sup>Departamento de Ciencias Ecológicas, Universidad de Chile. Instituto de Ecología y Biodiversidad.  
13 Las Palmeras 3425, Ñuñoa, Santiago, Chile.

14 <sup>5</sup>Facultad de Ciencias, Instituto de Ciencias Ambientales y Evolutivas, Universidad Austral de Chile,  
15 Casilla 567, Valdivia, Chile

16 <sup>6</sup>CNRS, Sorbonne Université, Pontificia Universidad Católica de Chile, Universidad Austral de Chile,  
17 IRL 3614, Evolutionary Biology and Ecology of Algae, Station Biologique de Roscoff, CS 90074, F-  
18 29688 Roscoff, France

19

20

21 \*Corresponding authors: Ariana Bertin-Benavides; email: [arianabertin@gmail.com](mailto:arianabertin@gmail.com); Marie-Laure  
22 Guillemin; email: [marielaure.guillemin@gmail.com](mailto:marielaure.guillemin@gmail.com).

23

24

25 **Abstract**

26 The South American temperate forests have been subjected to drastic past topographic and climatic  
27 changes during the Pliocene – Pleistocene linked to Andean orogeny and glacial cycles. Such changes  
28 are regarded standard drivers of genetic structure and adaptation for the local biota of this area.  
29 *Embothrium coccineum*, a member of the Proteaceae family and an emblematic tree of the South  
30 American temperate forest, is assumed to show patterns of genetic variation strongly derived by these  
31 historical drivers. Interestingly, recently documented genetic structure inferred from ecotype variants  
32 suggest that adaptive genetic responses could also reflect similar patterns of genetic differentiation. The  
33 goal of this study was to investigate the effects of both historical and contemporary landscape features  
34 affecting the genetic diversity and connectivity of *E. coccineum* throughout its natural distribution.  
35 Using more than 2,000 SNPs, two genetic groups (Northern and Central-Southern) that diverged some  
36 2.8 million years ago were observed. The level of genetic structure was higher between populations  
37 within the Northern genetic group than within the Central-Southern group. We propose that these  
38 contrasting patterns are related to differences in pollinator's assemblage and evolutionary histories  
39 between genetic groups. Moreover, we observed a strong pattern of isolation by the environment in *E.*  
40 *coccineum*, suggesting that selection could have led to adaptive divergence among localities. We  
41 propose that, within the Chilean temperate forest, the patterns of genetic variation in *E. coccineum*  
42 reflect both a Quaternary phylogenetic imprint and the impact of selection on the solid environmental  
43 gradient.

44

45 **Keywords:** Temperate forest - Population genomics - Proteaceae - Isolation by environment - Last  
46 Glacial Maxima - Divergence time

## 47 **Introduction**

48       Due to their isolation and a high degree of endemism, South American temperate forests are  
49 considered of great ecological and evolutionary importance (Smith-Ramirez et al., 2007). In Chile,  
50 temperate forests stretch from the 35°S down to the 55°S, forming a continuous area that can be  
51 considered a “biogeographic island” separated by impassable barriers (deserts, mountains, and oceans)  
52 from the rest of the ancestral sources of its biota (Villagrán, 1991). As a result, both climate and  
53 topography, which are distributed over a narrow strip (2,000 km long and 120 km wide) of Chilean  
54 temperate forests, define a highly heterogeneous landscape comprised of multiple forest and soil types,  
55 climatic conditions, and disturbance regimes (Loguercio et al., 2018). Under such conditions, it is  
56 expected that different, and even complementary, mechanisms equally affect genetic and phenotypic  
57 diversity. Adaptive processes linked to persistence in a mosaic of heterogeneous habitats may have  
58 contributed to plant species diversification in the temperate forests of southern South America (Johnson  
59 et al., 2014; Prunier et al., 2017). In the same way, the within-species genetic divergence of plants with  
60 extensive distribution under strong natural environmental gradients could also reflect the action of  
61 selection (Saldaña et al., 2007, Zeppel et al., 2014; Grossiord et al., 2004). Contrasting climatic and  
62 edaphic features on phenotypic and genetic diversity have been observed in temperate plants (Saldaña  
63 et al., 2007; Torres-Díaz et al., 2007).

64       For instance, distinct ecotypes have been linked to contrasting access to water and snow winter  
65 coverage in *Embothrium coccineum* J.R. Forts. & G. Forst (Fig. 1) (Dimitri, 1972), a widespread tree  
66 that inhabits the temperate forest of South America (Zegers, 1994). This species is an endemic tree of  
67 Gondwanan origin with a distribution range matching the extent of the temperate forest biome in Chile  
68 and Argentina, exhibiting different phenotypes throughout its distribution (Alberdi y Donoso, 2004;  
69 Chalcoff et al., 2008). These ecotypes are coincidentally distributed with observed pattern of genetic  
70 structure observed from isoenzymatic genetic markers (Souto & Premoli, 2007), suggesting that *E.*  
71 *coccineum* populations across the species range may be locally adapted to distinct environmental  
72 conditions (Souto and Premoli, 2007; Bertin-Benavides et al., 2020).



Fig. 1. *Embothrium coccineum* morphology tree. A) Mature tree, B) Leaf, C) Flower, D) Root, E) Follicle and F) Seed.

73 Besides the possible effect of current selective regimes, climate shifts during the Pleistocene have  
74 also strongly altered the spatial patterns of genetic variation of many southern South American taxa  
75 (Vidal-Russell et al., 2011; Turchetto-Zolet et al., 2013). During the Last Glacial Maximum (LGM, 25  
76 ,000 years Before Present, BP), given to the presence of ice caps, a large expanse of temperate forest  
77 located between the 42 °S and the 56 °S disappeared (Rabassa et al., 2005). Glacial maxima triggered a  
78 rapid range shift, with most temperate South American species surviving only in glacial refugia located  
79 north of the 42°S (Allnutt et al., 1999; Premoli et al., 2000, Sérsic et al., 2011). Some cold-tolerant  
80 species, such as *E. coccineum*, were also able to survive in refugia embedded within the main glaciated  
81 area in localities where microclimate, topography or geothermic activity hindered the formation of the  
82 ice cap (Allnutt et al., 1999; Premoli et al., 2000, Sérsic et al., 2011). These areas could represent  
83 sources of regional diversity recovery after deglaciation (Comps et al., 2001). The presence of 11,000  
84 years BP cold-tolerant plant pollen fossil in areas covered by ice during the LGM (Fesq-Martin et al.,  
85 2004, Steubing et al., 1983), plus results of genetic studies based on sequences of chloroplast genes or

86 genotyping of nuclear isoenzymatic markers (Vidal-Russel et al., 2011; Souto & Premoli, 2007)  
87 suggest the existence of a complex glacial-interglacial history in this species.

88         Recent developments in high throughput sequencing technologies and population genomics have  
89 allowed us to distinguish the influence of past environmental changes and selection to the current  
90 environmental conditions, particularly on spatial patterns of genetic variation in both South American  
91 plants and animals (Turchetto-Zolet et al., 2013; Hasbún et al., 2016; Varas-Myrik et al., 2021). If  
92 divergent natural selection has led to genetic differentiation between populations of *E. coccineum* along  
93 environmental gradients characterizing South American temperate forests, a pattern of isolation by the  
94 environment (IBE; Wang & Bradburd, 2014) should be expected. Indeed, when selection against  
95 migration limits gene flow among distinct environments, genetic differentiation increases with  
96 environmental differences (Wang & Summer, 2010; Wang & Bradburd, 2014). The correlation  
97 between genetic differentiation and environmental distances should be contrasted against the observed  
98 between genetic differentiation and geographic distances (isolation by distance; IBD; Slatkin, 1993) to  
99 test the IBE. IBD could be considered a null model for which genetic differentiation increases with  
100 geographical distance, given the local accumulation of genetic differences when the dispersion between  
101 populations is geographically restricted. In *E. coccineum*, gene flow could be limited by constraints on  
102 pollen transport that is linked to the distribution and behaviour of pollinators (i.e., more than 20  
103 species, including birds and insects, are reported as pollinators of *E. coccineum*; Chalcoff et al., 2008)  
104 and anemochorous seed dispersal (Rovere & Premoli, 2005).

105         The main goal of this study was to evaluate the effects of historical and contemporary landscape  
106 features affecting the genetic diversity and connectivity of *E. coccineum* throughout its natural  
107 distribution using SNPs obtained by genotyping-by-sequencing (GBS). Patterns of genetic structure  
108 were evaluated independently using SNPs found in genomic regions that may be subject to selection  
109 (i.e., outlier loci) and all loci. We hypothesized that patterns of genetic variation in *E. coccineum* would  
110 reflect both the impact of historical processes (i.e., glacial/interglacial cycles during the Pleistocene)  
111 and contemporary landscape features, including the selective effect of environmental heterogeneity.

## 112 **Materials and methods**

113

### 114 **Samples collection and DNA extraction**

115 The locations were selected to cover an extensive range of climatic conditions across *E.*  
116 *coccineum* natural distribution: i) three locations (Chillán, Nahuelbuta and Curacautín) representing the  
117 northern distribution, ii) four locations (Puerto Montt, Chiloé Norte, Chiloé Sur and Pumalín) the  
118 central distribution, and iii) three locations (Coyhaique, Chile Chico and Torres del Paine) the  
119 southern distribution (Fig. 2a and Appendix I). These three regions are subsequently named  
120 Northern, Central and Southern locations. Three to six individuals were sampled within each location  
121 and five mature leaves were collected per tree. A total of 42 trees were sampled. To calibrate  
122 divergence time among *E. coccineum* genetic groups, two individuals of *Lomatia hirsuta* Diels, a  
123 Proteaceae species from Embothrieae tribe (Sauquet et al., 2009), were included as an outgroup. The  
124 selecting mature leaves were deposited and transported to laboratory in a cooler. In the Laboratory, the  
125 samples were frozen with liquid nitrogen and stored at -80° C until the DNA extraction. Genomic DNA  
126 was obtained from tissue (50 mg) of mature leaves powdered in Precellys24 homogenizer (Precellys,  
127 USA) using the Qiagen DNeasy Plant kit (Qiagen Inc., USA) following the manufacturer's instructions.  
128 DNA samples were analyzed for their integrity by agarose gel electrophoresis, and the total amount of  
129 recovery was determined using a Qubit fluorometer (Invitrogen, USA).

130

### 131 **Climatic data collection**

132 To characterize patterns of spatial environmental variation among the locations, we used the historical  
133 values (1,970-2,000) of 19 environmental variables (bio1 - bio19) and elevation (Elev), which were  
134 obtained using the package raster v.2.7.15 and the Worldclim data of 2.5 arcmins (available at  
135 <http://worldclim.org/>). We tested for covariation among environmental variables using the vifcor()  
136 function from the “usdm” R package (Naime et al., 2014). Variables presenting no strong covariation  
137 ( $R < 0.8$ ; see Fitzpatrick and Keller, 2014) were conserved for subsequent analyses (Appendix I).

138

### 139 **Preparation of the library, high throughput sequencing and GBS genotyping**

140 Library preparation and high-throughput sequencing were performed at the University of  
141 Wisconsin Biotechnology Center (DNA Sequencing Facility). The preparation of the GBS genomic  
142 library was done following the protocol detailed by Elshire et al. (2011) using the ApeKI restriction  
143 enzyme and 44 specific barcodes (i.e., one per individual). High throughput sequencing was performed  
144 using an Illumina HiSeq 2,000 (Illumina, USA) and single-strand sequencing runs of 100 bp. The raw  
145 GBS dataset generated and analyzed during the current study is available in the SRA repository  
146 (<https://www.ncbi.nlm.nih.gov/sra/?term=PRJNA783610>). The quality of the sequenced raw data was  
147 assessed using FastQC (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>).

148 SNP calling was *de novo* assembled using Stacks v.2.2 pipeline (Catchen et al., 2013). Then,  
149 values of the two main parameters (-M: number of mismatches allowed between stacks within  
150 individuals; and -n: number of mismatches allowed between stacks between individuals) were chosen  
151 following the optimization procedure described in Paris et al. (2017).

152

### 153 **GBS data sets**

154 To test the possible effects of parameter assembly and missing data on patterns of population  
155 genetic structure, distinct sets of parameter of populations module on Stack v.2.2 (-p: minimum number  
156 of populations where a locus must be present for it to be included; and -r: percentage of individuals in a  
157 population that must possess a particular locus for it to be included), were examined before choosing  
158 optimal filtering options for further analysis. Eight combinations of parameters -p and -r were tested (-  
159 r= 0.6 and 0.7; -p= 6, 7, 8 and 9). For each combination the % of missing data, % of heterozygosity and  
160 number of SNPs were estimated (data not shown). The minimum minor allele frequency (min-maf) was  
161 set to 0.02, and the maximum observed heterozygosity (max-obs-het) was set to 0.5. The “Population  
162 Genetic data set” (i.e., PG\_dataset), was built using -p 9 and -r 0.7. These allowed us to obtain a final  
163 matrix with the largest SNPs and the lower number of missing data.

164 Furthermore, another data set was built to estimate the divergence time among the *E.*  
165 *coccineum* genetic groups, using the Bayesian multispecies coalescent model (Stange et al., 2018).  
166 First, we selected a subset of samples representing the Northern (Nah and Cu), the Central (ChlN and  
167 ChlS) and the Southern (Coy and TP) locations of *E. coccineum*. Then, we selected 12 *E. coccineum*  
168 individuals (i.e., two per location) to maximize the SNPs available and minimize the amount of missing  
169 data. These 12 *E. coccineum* individuals were processed jointly with the samples of *L. hirsuta*,  
170 included as an outgroup. A new SNP calling was conducted following the same parameters as those  
171 used for the PG\_dataset in Stacks v.2.2 pipeline (Catchen et al., 2013), eventually generating the  
172 “divergence time data set” (i.e., DT\_dataset). In this case, optimal values for -p and -r parameters were  
173 chosen to decrease the percentage of missing data: only the loci present in all locations, including the  
174 outgroup (-p 7), and present in the least 80 % of the individuals (-r 0.8) were kept for further  
175 analyses. Putative outliers loci were removed from the DT\_dataset. One SNP was randomly selected  
176 per locus for each data set to approximate unlinked loci variation.

### 177 **Populations diversity statistics**

178 To explore the distribution of genetic diversity, we calculated, for each location, the observed  
179 ( $H_o$ ) and expected ( $H_e$ ) levels of heterozygosity and the inbreeding coefficient FIS for each location  
180 with GenoDive v.3.02. The percentage of polymorphic loci and private alleles were estimated using the  
181 R package “hierfstat v. 0.04-22” (Goudet 2005).

182

### 183 **Genetic structure**

184 Four complimentary analyses were conducted to characterize the pattern of genetic structure  
185 within *E. coccineum* using the PG\_dataset. First, we calculated the pairwise FST values among all  
186 sampling localities using the function stampFst(), using 1,000 bootstrap replicates across loci and 95%  
187 confidence intervals with the R package StAMPP v.1.5.1 (Pembleton et al., 2013). Second, a principal  
188 component analysis (PCA) was performed using the glPCA() function with the R package adegenet  
189 v.2.1.1 (Jombart, 2008). Only the first three components were retained to estimate genetic groups.  
190 Third, a model-based STRUCTURE clustering analysis was performed using STRUCTURE v.2.3.4.  
191 (Falush et al., 2003). Ten independent simulations allowing admixture were run for each K (K=1-10),  
192 each with 200,000 Markov chain Monte Carlo replicates (MCMC) and 100,000 samples discarded as  
193 burn-in. STRUCTURE HARVESTER (<http://taylor0.biology.ucla.edu/structureHarvester/>) was used to  
194 determine the optimum K based on L(K) and  $\Delta K$  parameters (Zhang et al., 2018). Fourth, to describe  
195 the relationships between individuals and genetic groups among *E. coccineum*, the Nei's genetic  
196 distance between individuals was calculated within the R package StAMPP v.1.5.1 (Pembleton et al.,  
197 2013). Unrooted Neighbor-Joining (NJ) distances trees were constructed using the individual Nei's  
198 distance matrix, as implemented by the R package ape v.5.3 (Paradis & Schliep, 2018).

199

## 200 **Outlier detection and association of allele frequency with environmental variables**

201 Two complementary approaches were used to identify candidate loci influenced by selection  
202 within *E. coccineum* PG\_dataset. First, we used a Principal Component Analysis (PCA) to generate a  
203 "model-free" null distribution of individual genetic distances to detect outlier loci. This analysis was  
204 conducted using the R packages PCAdapt v.4.0.3 (Luu et al., 2017) and q-value v.2.16.0 (Storey et al.,  
205 2004) with default parameters and a search with K=10. The detection of outliers was confirmed by  
206 plotting the histograms of p-values, and the Mahalanobis (D2) test statistic, with a False Discovery  
207 Rate (FDR), set at 5%. Second, we implemented a Bayesian test to detect outlier loci using the  
208 BAYESCAN 2.1 software (Foll et al., 2008). All simulations were performed using the default  
209 parameters with 20 pilot runs of 5,000 iterations followed by 50,000 sampling iterations and an FDR of  
210 5%. Outlier loci detected by both PCAdapt and BAYESCAN were selected as candidate adaptive loci.

211 To search for candidate adaptive loci with allelic frequency potentially influenced by the  
212 environmental variation characterizing our study area (see Appendix I), a Gradient Forest (GF) analysis  
213 was implemented in the R package gradientForest (Ellis et al., 2012). GF provides a ranked list of the  
214 relative predictive power (R2) of all environmental variables allowing identifying those that best  
215 explain the observed genetic variation. Allele frequency of candidate adaptive loci (i.e., detected in  
216 both PCAdapt and BAYESCAN) was used as a response variable. GF was fitted using 2,000 regression  
217 trees per SNP and a variable correlation of 0.5 (Fitzpatrick & Keller, 2014).

218

## 219 **Isolation by distance and isolation by environment**

220 To better understand the drivers potentially explaining the observed patterns of spatial genetic  
221 variation in *E. coccineum*, we performed distance-based redundancy analysis (db-RDA) and partial db-  
222 RDA at the individual level on the complete PG\_dataset and loci potentially influenced by selection.  
223 RDA is a multiple linear regression method performed between a matrix of dependent variables and  
224 matrices of independent variables, which has been demonstrated as more appropriate than other  
225 methods (i.e., Mantel test) when multiple variables are analyzed to identify drivers of population  
226 genetic structure (Nadeu et al., 2016; Orsini et al., 2013).

227 A dependent matrix, calculated with a Bray-Curtis dissimilarity index (Bray & Curtis, 1957), was  
228 used as a response variable of each individual's co-dominant variant call format. Predictor matrices  
229 were geographic distance (IBD), environmental disparity (IBE) and degree of shared co-ancestry  
230 (IBA). As environmental data, we used seven environmental variables with a low level of covariation  
231 (see Appendix I). For the distance matrix, we used a Principal Coordinates of Neighbourhood Matrix  
232 (PCNM), using a truncation threshold of 0.05 for long distances with the function pcnm() within the R  
233 package vegan v.2.5.7 (Oksanen et al., 2007). PCNM is commonly used to transform spatial distances  
234 into rectangular data matrices suitable for constrained ordination or regression analyses (Borcard and  
235 Legendre 2002). Populations Q-values obtained with STRUCTURE K = 2, which separated  
236 populations into Northern and Central-Southern ancestral lineages (see results for more details), were  
237 used as a co-ancestry variable. Prior to the analysis, all three independent matrices were scaled to a  
238 mean of zero and variance of one with the scale() function in R.

239 Among populations, variation in *E. coccineum* was partitioned into components explained by  
240 geographic distance (IBD), ecological gradients (IBE) and co-ancestry (IBA), or their combination  
241 (i.e., constrained by the effects of the remaining two independent matrices), using the vartpart()  
242 function within R package vegan v.2.5.7 (Oksanen et al., 2007). The significance of each partition was  
243 tested with the anova.cca() function through 1,000 permutations on the R package vegan v.2.5.7  
244 (Oksanen et al., 2007).

245

#### 246 **Divergence time among *Embothrium coccineum* genetics groups**

247 We implemented a Bayesian multispecies coalescent model to estimate divergence times among  
248 *E. coccineum* genetic groups (Rannala & Yang, 2003) using the DT\_dataset. This analysis was  
249 conducted with the coalescent simulator SNAPP v.1.3.0 (Bryant et al., 2012) included in the BEAST  
250 v.2.4.5 package (Bouckaert et al., 2014). For this approach we followed a model accommodated to  
251 genome-wide SNP data, which implements a strict molecular clock and nodes calibrated with  
252 biogeographic constraints resulted from one or more divergence events (Stange et al. 2018). This  
253 approach assumes a symmetric substitution model and equal frequencies, plus the implementation of a  
254 single parameter for the rate of the strict molecular clock, as this is the only model currently supported  
255 by SNAPP for this type of data (Stange et al. 2018).

256 The set of parameters used exactly followed the “Divergence-time estimation with SNAPP”  
257 tutorial

258 ([https://github.com/mmatschiner/tutorials/tree/master/divergence\\_time\\_estimation\\_with\\_snp\\_data](https://github.com/mmatschiner/tutorials/tree/master/divergence_time_estimation_with_snp_data)),  
259 except with a few exceptions. These include the use of a lognormal distribution for the the crown age  
260 of the Embothriaceae tribe, which was estimated at 66 million years (Myr) (Mean = 66.6 Standard  
261 Deviation = 0.13, Mean In Real Space = TRUE) using a phylogeny of Proteaceae that included all the  
262 genera currently recognized in the family (Sauquet et al. 2009). This estimate was used to approximate  
263 the node age of the common ancestor of *Embothrium coccineum* in our dataset, which was derived  
264 from the sister relationship with *Lomatia* samples. Additionally, a generation time of 10 years per  
265 individual was set, following experimental greenhouse observations (Bertin-Benavides A., personal  
266 observation). Alpha and beta parameters were left with an uniform prior, with an arbitrary value of 1.0  
267 (Stange et al. 2018). For additional scrutiny, the XML file is available in supplementary material.

268       Species trees were inferred directly with no prior grouping assignment. We used SNAPP with  
269 four independent runs of 2,000,000 generations with a sampling every 200 generations each. The  
270 results (i.e., ESS > 300) were checked for mixing for all parameters with Tracer v1.7 (Rambaut et al.,  
271 2018) and convergence of split frequencies among runs (Appendix III) with the R package rwtly  
272 (Warren et al., 2019). The final trees from each run (corresponding to the 95% highest posterior density  
273 (HPD), after discarding a 10% burn-in of tree topologies) were combined and subsequently  
274 summarized using a maximum credibility tree with TreeSetAnalyser v.2.4.5. Finally, this species tree  
275 was visualized as a cloudgram with DensiTree v.2.2.7 (Bouckaert, 2010).

276

277 **Results**

278 In order to evaluate the population structure, the divergence time among genetics groups and the local  
 279 adaptation in *E. coccineum*, 38 trees were genotyped using GBS. A total of 116,606,122 reads were  
 280 generated and de novo assembled into 1,303,750 loci, from which 473,461 SNPs were called. The  
 281 application of filtering criteria resulted in a final PG\_dataset of 2,155 SNPs. For the DT\_dataset,  
 282 composed by 12 individuals of *E. coccineum* and two individuals of *L. hirsuta*, a total of 320 SNPs  
 283 were called.

284 **Genetic diversity**

285 The Expected and observed heterozygosity (He and Ho, respectively), inbreeding coefficient  
 286 (FIS), the percentages of polymorphic loci (PPL) and the number of private alleles (PA) are reported in  
 287 Table 1 for each sampling locality. Except for the low value observed in Chile Chico (ChCh;  
 288 PPL=12.668 %), PPL was reasonably homogeneous throughout the species range, varying from  
 289 18.794% in Chillán (Ch) to 42.367% in Coyhaique (Coy) (Table 1). Four sampling locations, Puerto  
 290 Montt (PM), Chiloé Norte (ChlN), Chiloé Sur (ChlS) and Chile Chico (ChCh), presented a lower PA  
 291 (PA≤30) than the rest of the sampled range (44<PA<83). Ch and Pumalín (Pu) were the only two  
 292 locations with a negative FIS value, while four locations (i.e., PM, ChlN, ChlS and ChCh) showed  
 293 positive values of FIS superior to 0.1.

294

295 **Table 1: Genetic diversity of *Embothrium coccineum*.** Locations ID are as in Appendix I. n, number  
 296 of genotyped individuals; He, expected heterozygosity; Ho, Observed heterozygosity; FIS, inbreeding  
 297 coefficient; PPL, percentage of polymorphic loci; PA, private alleles.

ID	n	He	Ho	Fis	PPL (%)	PA
Ch	3	0.1	0.105	-0.049	18.794	51
Nah	4	0.138	0.132	0.037	32.715	65
Cu	4	0.128	0.126	0.012	30.998	73
PM	3	0.152	0.135	0.114	29.791	30
ChlN	3	0.163	0.139	0.15	36.241	29
ChlS	3	0.152	0.128	0.155	33.132	29
Pu	3	0.134	0.154	-0.148	27.425	83
Coy	6	0.143	0.137	0.043	42.367	62
ChCh	3	0.124	0.106	0.146	12.668	10

TP	6	0.139	0.127	0.086	42.088	44
----	---	-------	-------	-------	--------	----

298

299

300

301 **Patterns of population genetic differentiation**

302 The highest pairwise values of FST (Table 2) were observed between Northern and Southern  
303 locations ( $0.319 < FST < 0.424$ ), followed by those calculated between Northern and Central locations  
304 ( $0.222 < FST < 0.429$ ) and then by the ones calculated between Central and Southern locations ( $0.070$   
305  $< FST < 0.193$ ). Lower levels of divergence were observed between locations sampled in the same  
306 macro region (Northern:  $0.089 < FST < 0.174$ ; Central  $0.010 < FST < 0.152$ ; Southern:  $0.043 < FST <$   
307  $0.061$ ) (Table 2). All FST values were significant.

308

309 **Table 2: Average FST values for pairwise comparisons among sampling locations of *Embothrium***  
310 ***coccineum* in Chile based on a set of 2,155 SNPs.** All *p*-values are significant ( $<0.05$ ). Sampling  
311 location codes are as in Appendix I. Northern: Ch, Cu and Nah; Central: PM, ChIN, ChIS and Pu;  
312 Southern: Coy, ChCh and TP.

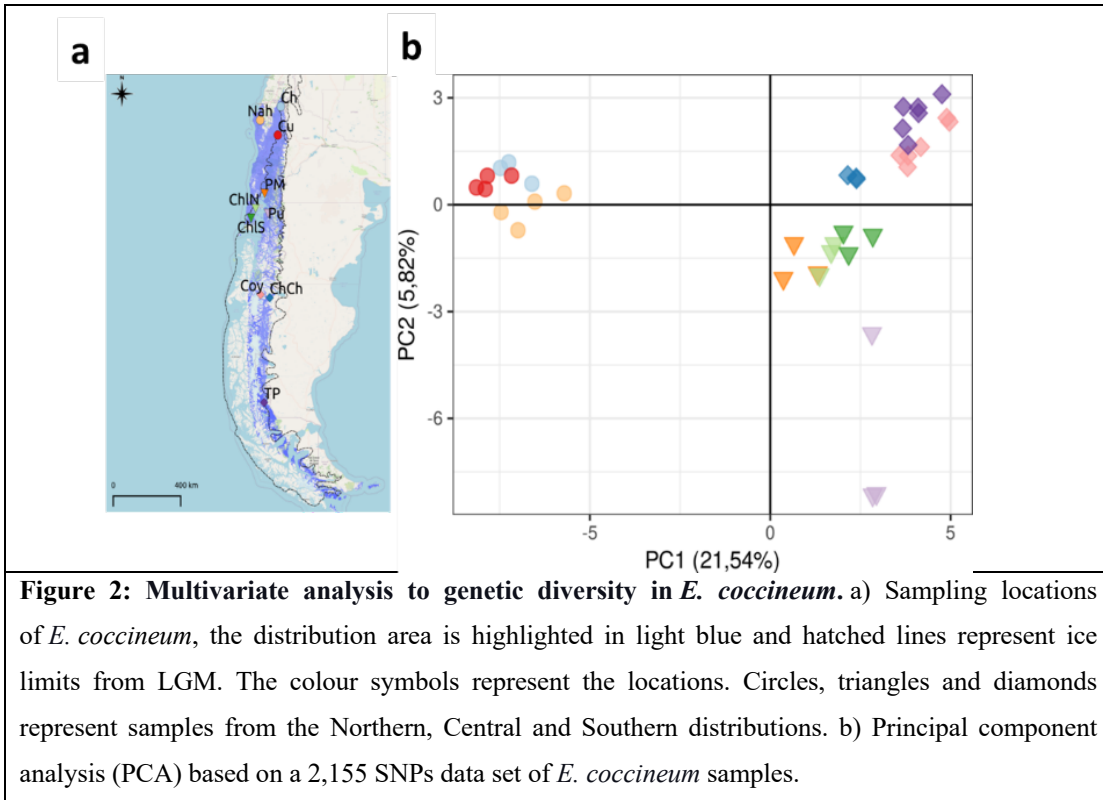
	Ch	Cu	Nah	PM	ChIN	ChIS	Pu	Coy	ChCh	TP
Ch	-									
Cu	0.163	-								
Nah	0.174	0.089	-							
PM	0.312	0.263	0.222	-						
ChIN	0.309	0.260	0.225	0.024	-					
ChIS	0.345	0.292	0.257	0.041	0.010	-				
Pu	0.429	0.369	0.338	0.152	0.126	0.135	-			
Coy	0.382	0.340	0.319	0.117	0.083	0.071	0.168	-		
ChCh	0.424	0.362	0.341	0.133	0.089	0.070	0.193	0.061	-	
TP	0.384	0.345	0.322	0.133	0.094	0.085	0.185	0.061	0.043	-

313

314 Regarding the distribution of *E. coccineum*, not all localities were under ice at LGM (Fig. 1a).  
315 The PCA showed that *E. coccineum* was divided into two major genetic groups: Northern and Central-  
316 Southern (Fig. 2b). The Northern locations were retrieved separated from those corresponding to the  
317 Central-Southern along the first axis of the PCA (PC1, representing 21.54% of the genetic variance). In  
318 contrast, samples from the Central and the Southern were separated along the second axis of the PCA  
319 (PC2, representing 5.82% of the genetic variance).

320

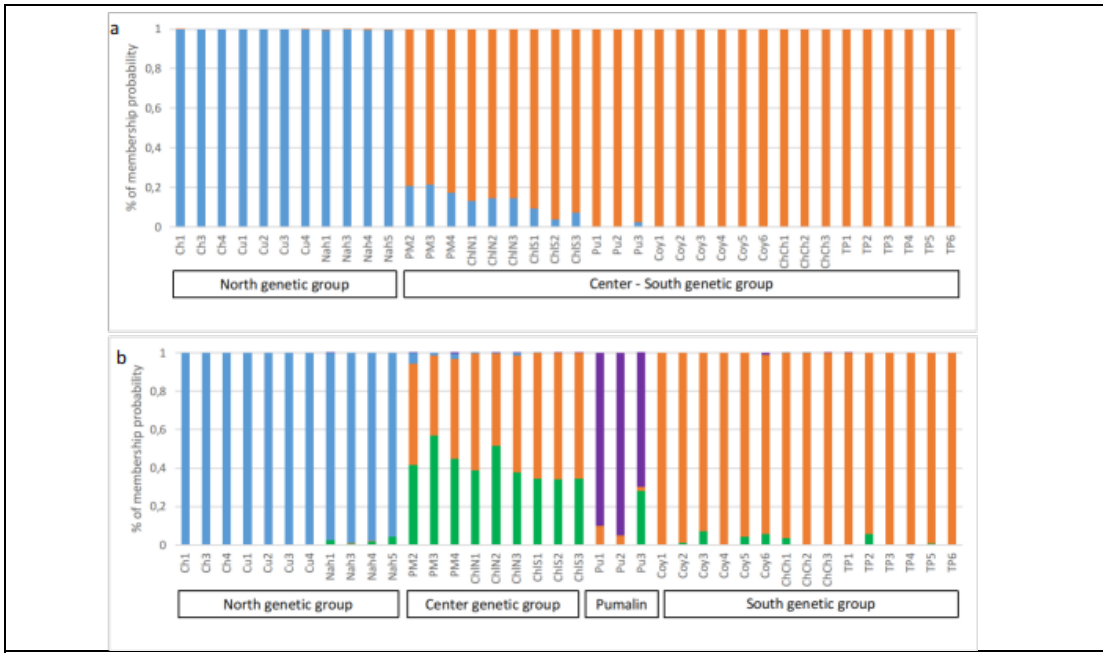
321  
322  
323  
324  
325



**Figure 2: Multivariate analysis to genetic diversity in *E. coccineum*.** a) Sampling locations of *E. coccineum*, the distribution area is highlighted in light blue and hatched lines represent ice limits from LGM. The colour symbols represent the locations. Circles, triangles and diamonds represent samples from the Northern, Central and Southern distributions. b) Principal component analysis (PCA) based on a 2,155 SNPs data set of *E. coccineum* samples.

326

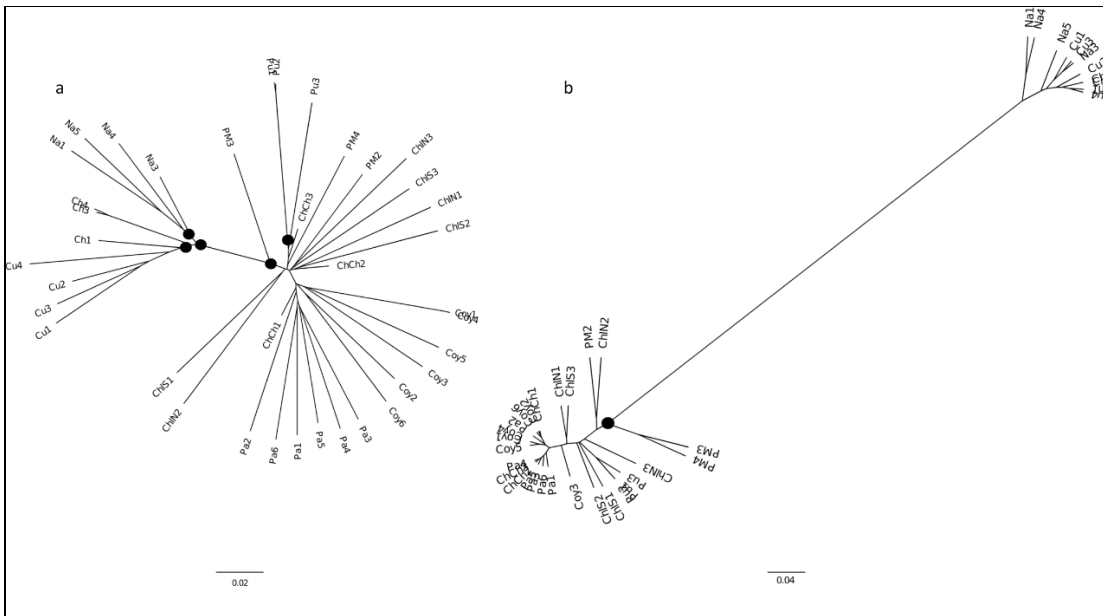
327 Results from the model-based Bayesian clustering approach showed a  $K = 2$  (highest  $\ln P(D)$  and  
328  $\Delta K$ ) as optimal and  $K = 4$  as a possible secondary level of substructure (Appendix I). The  
329 STRUCTURE clustering ( $K = 2$ ) revealed that all individuals sampled from the Northern were  
330 assigned to the blue genetic cluster. In contrast, all individuals sampled from both Central and Southern  
331 locations were assigned to the orange genetic cluster (Fig. 3a). In both Northern and Southern samples,  
332 almost no admixture was observed. Individuals from PM, ChlN and ChlS (Central) were majorly  
333 assigned to the orange genetic cluster but exhibited low levels of admixture from the blue genetic  
334 cluster. The level of assignment to the blue genetic cluster tends to recede across latitudes (from 20%  
335 of the genome assigned to the blue genetic cluster in PM to less than 10% in ChlS; Fig. 3a). The  
336 STRUCTURE clustering using  $K = 4$  showed two new genetic clusters within the Central-Southern  
337 part of the species range (Fig. 3b). Individuals from the Southern were assigned to the orange genetic  
338 cluster, individuals from Pumalín (PU, Central) mainly were assigned to the new purple genetic cluster,  
339 and the rest of the individuals from the Central (PM, ChlN and ChlS) were admixed between the  
340 orange genetic cluster and the newly detected green genetic cluster (Fig. 3b).



**Figure 3: *Embothrium coccineum* genetic structure.** Analyses for K = 2 (a) and K = 4 (b) of 38 *E. coccineum* individuals based on a 2,155 SNPs data set. Each individual is displayed from north to south, represented by a vertical bar. Sampling location codes are as in Appendix II. The Colour corresponds to the genetic group, and the proportion of the genome assigned to each genetic group is given along the Y-axis

341

342 The dendrogram was reconstructed with the PG\_dataset (2,155 SNPs) allowed to separate the  
 343 Northern and Central-Southern samples (Fig. 4a), being concurrent with the results obtained by the  
 344 PCA and the STRUCTURE clustering analyses. When compared with the results obtained with the  
 345 whole data set (i.e., PG\_dataset), the dendrogram retrieved with the 59 outlier loci putatively under  
 346 selection (see below for more details) revealed a much deeper separation between the Northern and  
 347 Central-Southern samples; yet, most sub clustering within major genetic groups was lost (Fig. 4b).



**Figure 4: NJ tree reconstructed based on Nei's distance between *Embothrium coccineum* individuals.** NJ reconstruction based on (a) the PG\_dataset (2,155 SNPs), and (b) the 59 outlier loci putatively under selection. Sampling location codes are as in Appendix I. Nodes with high support (> 0.95) are filled in black.

348

349

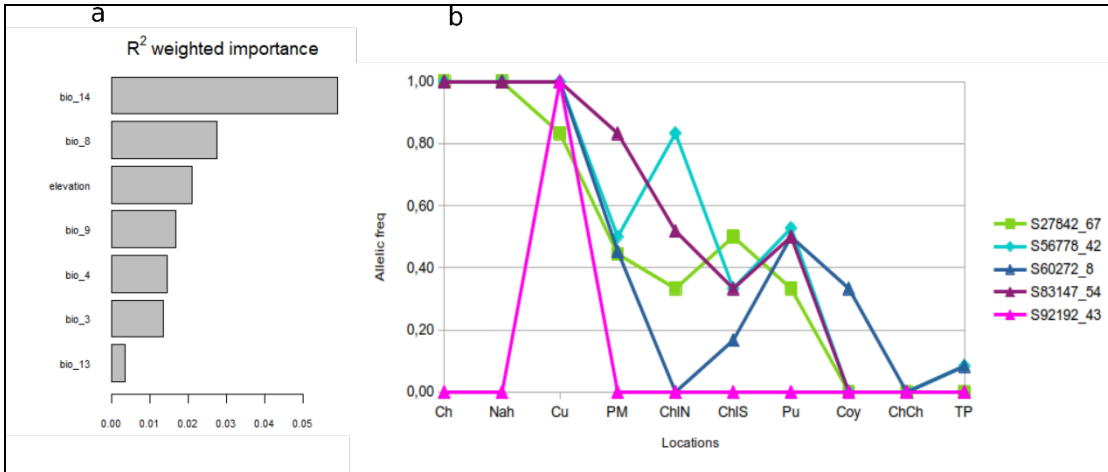
### 350 **Genome scan for outlier loci detection**

351 Within the 2,155 loci included in the PG\_dataset, PCAdapt detected 132 outlier loci and  
352 BAYESCAN 89 outlier loci. The two methods concur with the detection of 59 loci (2.73%) that were  
353 then identified as potentially adaptive loci in *E. coccineum*.

354 The gradient forest (GF) analysis revealed four environmental factors with the highest impact on  
355 *E. coccineum* populations (i.e., estimated as the variables presenting the highest values of R2 weighted  
356 importance), which correspond to Precipitation of the driest month (bio14), Mean temperature of the  
357 wettest quarter (bio8), Elevation, and the Mean temperature of the driest quarter (bio9) (Fig. 5a). The  
358 value of the R2 weighted importance for bio 14 (R2 weighted importance > 0.05) was at least two  
359 times higher than those observed for the other variables.

360 In addition, GF also produced individual turnover functions for each potentially adaptive loci  
361 having an  $R^2 > 0$  (Fitzpatrick & Keller, 2014). This value allows listing the potentially adaptive loci  
362 showing the highest correlation of allelic frequencies with the four environmental factors of the  
363 significant impact on *E. coccineum* populations. Results show that allelic frequencies of outliers  
364 S92192\_43 and S60272\_8 are strongly correlated with Precipitation of the driest month (bio14; Fig. 5b:  
365 pink and blue lines, respectively). S60272\_8 allelic frequency also exhibited a weak response to the  
366 Mean temperature of the driest quarter (bio9; Fig. 5b). Three outliers (S27842\_67, S56778\_42 and  
367 S83147\_54) exhibited a significant and concurrent response to the Mean Temperature of the Wettest  
368 Quarter (bio8; Fig. 5b: green, calypso and purple lines, respectively). The allelic frequency of two  
369 outliers (S272842\_67 and S56778\_42) also shows a weak response to elevation and Mean Temperature  
370 of the Wettest Quarter (bio9) (Fig. 5b; light green and light blue lines, respectively).

371 Four outliers showed allelic frequency gradually changing with latitude (Figure 5b; S60272\_8,  
372 S27842\_67, S56778\_42 and S83147\_54). Allelic frequencies among locations sampled in the northern  
373 distribution (i.e., Ch, Nah, Cu) were similar. The locus S92192\_43 showed the fixation of a private  
374 allele in the Curacautín (Cu) location.



**Figure 5: Environmental factors with the highest impact on *Embothrium coccineum* populations and allelic frequency.** a) The R<sup>2</sup> weighted importance of environmental variables with a low covariance explains the genetic variation across sampling locations obtained from GF analysis. Environmental variables used in the analyses are the same as in Appendix I; bio14, Precipitation of the driest month (mm); bio8, Mean temperature of the wettest quarter (°C); Elevation (masl); bio9, Mean temperature of the driest quarter (°C); bio4, Temperature seasonality (°C); bio3, Isothermality (%); bio13, Precipitation of the wettest month (mm). b) Change in allele frequency between sampling localities for the five potentially adaptive loci showing major response to environmental variables. (Loci: S92192\_43, S60272\_8, S27842\_67, S56778\_42 and S83147\_54). The sampling localities are ordered along a latitudinal gradient.

375

376 **Isolation by distance and isolation by environment**

377 The full RDA model using the PG\_datset explained 59.4 % of the total genetic variation and  
 378 supported an influence of environmental (Env), geographic (Geo) and coancestry (Anc) variables in  
 379 shaping allelic variation (Env: adjusted R<sup>2</sup> = 0.420, p < 0.001; Geo: adjusted R<sup>2</sup> = 0.226, p < 0.001;  
 380 Anc: adjusted R<sup>2</sup> = 0.105, p < 0.001; Table 3). With the partial db\_RDA, the contribution of the  
 381 environmental variables to genetic divergence was much higher than the contribution of the geographic  
 382 distance or the co-ancestry (Env | Geo+Anc: adjusted R<sup>2</sup> = 0.288; Geo | Env+Anc: adjusted R<sup>2</sup> =  
 383 0.118; Anc | Env+Geo: adjusted R<sup>2</sup> = 0.011; Table 3). In total, 0.178 explained variation was  
 384 confounded between the effects of IBE, IBD and IBA, and 0.405 of the variation remained unexplained  
 385 (Table 3). Similar results were obtained for the subset of 59 loci potentially influenced by selection  
 386 (i.e., outlier loci). The partial db-RDA highlights the important contribution of the environmental  
 387 variables to potential local adaptation in *E. coccineum* (Env | Geo+Anc: adjusted R<sup>2</sup> = 0.244; Table 3).

388

389 **Table 3: Redundancy analysis (RDA) partitioning among-population genetic variation in**  
 390 ***Embothrium coccineum* into three components: environmental (Env), geographic (Geo) and**  
 391 **Northern/Central-Southern coancestry (Anc). All p-values <0.001.**

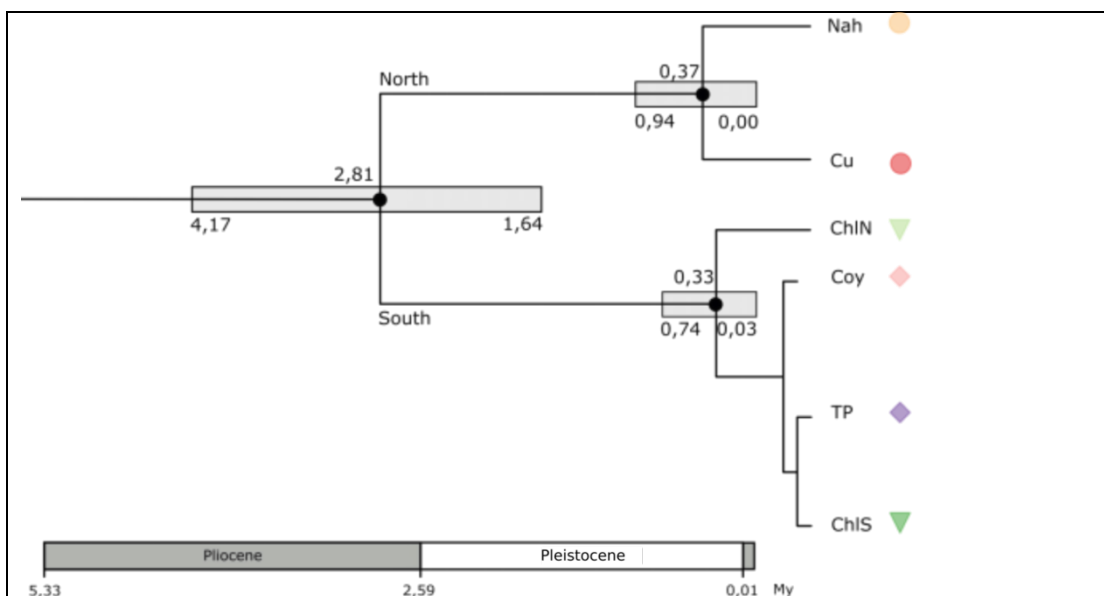
	All loci	Outlier loci		All loci	Outlier loci
Combined fractions	Adjusted R <sup>2</sup>	Adjusted R <sup>2</sup>	Individual fractions	Adjusted R <sup>2</sup>	Adjusted R <sup>2</sup>
Env	0.420	0.569	Env   Geo+Anc	0.288	0.244
Geo	0.226	0.378	Geo   Env+Anc	0.118	0.111
Anc	0.105	0.259	Anc   Env+Geo	0.011	0.002
Env+Geo	0.582	0.719	Total confounded	0.178	0.360
Env+Anc	0.476	0.606	Unexplained	0.405	0.283
Geo+Anc	0.305	0.473			
Env + Geo + Anc	0.594	0.717			

392

393

#### 394 Divergence time

395 The divergence time between the Northern and Central-Southern genetic groups was estimated at  
 396 2.81 My (1.64 – 4.17 My), near the end of the Pliocene and the beginning of the Pleistocene. However,  
 397 the locations within the Northern genetic group (Nah and Cu) or within the Central-Southern genetic  
 398 group (ChlN and Coy, TP, ChlS) seem to have diverged much more recently towards the end of the  
 399 Pleistocene (0.33 / 0.37 My ago; Fig. 6).



**Figure 6: Estimation of divergence time in *Embotrium coccineum* based on Bayesian coalescent analysis using SNAPP. Nodes with high support (posterior probability > 0.99) are filled in black. Median ages are provided above nodes with 95% highest posterior densities (HDP) below.**

The divergence time was inferred only for the nodes showing high support (posterior probability > 0.99). Sampling location codes are as in Appendix I; coloured symbols for each sampling location match Fig. 2a.

400

401

402

## 403 Discussion

404 In the present study, samples of *E. coccineum* from its distribution range along the Chilean  
405 temperate forest were genotyped through GBS to investigate patterns of genetic variation and their  
406 possible drivers using genome-wide data in this emblematic Chilean tree. The thousands of loci  
407 successfully genotyped provided a much better resolution and new knowledge of *E. coccineum*  
408 population genetic structure compared to previous studies. While these studies were based on  
409 allozymes genotypes or chloroplast sequences (Mathiasen et al., 2007; Souto and Premoli, 2007; Vidal-  
410 Russell et al., 2011; Souto and Smouse, 2014), their limited variability and genome coverage have not  
411 provided resolute evidence to delineate clear genetic limits and clarify the underlying processes  
412 shaping them.

413 Our data showed the existence of two main lineages, Northern and Central-Southern. These  
414 genetic groups were already reported in Souto and Premoli (2007), but the GBS data set also uncovered  
415 fine-scale genetic differentiation, especially in the Northern distribution. Inference of divergence time  
416 between these two genetic groups suggested that intraspecific divergence occurred some 2.8Myr ago, at  
417 the beginning of the Pleistocene, driven by climatic changes linked to glacial cycles. Moreover, the  
418 heterogeneous environment characterizing the South American temperate forests has been recognized  
419 as a significant driver explaining the actual genetic makeup of various plants and animals (Premoli et  
420 al., 2000). It could also have led to adaptive divergence among locations in *E. coccineum*. The strong  
421 IBE pattern detected in *E. coccineum* could reflect the impact of an actual gradient of temperature,  
422 precipitation and elevation, characteristic of the species distribution range, on its genetic diversity.  
423 These results suggest that selection linked to the strong environmental gradient could have led to  
424 adaptive divergence among localities.

425

### 426 Genetic differentiation within *Embothrium coccineum*: two genetic groups separated since the 427 early Pleistocene

428 Our result confirms previous studies in *E. coccineum* documenting the presence of two genetic  
429 groups, one in the Northern and one in the Southern distribution (Souto & Premoli, 2007; Vidal-Russell  
430 et al., 2011); yet these datasets exhibit different levels of spatial resolution. Alloenzymes revealed only  
431 a weak geographic structure, with some Northern and Southern locations clustered together (Vidal-  
432 Russell et al., 2011). The authors explained this pattern by proposing a possible effect of glacial  
433 survival in multiple refugia followed by recolonization and/or the existence of actual gene flow,  
434 blurring the trace of the historical geographic structure. Contrastingly, chloroplast markers revealed a  
435 distinct Southern genetic group characterized by a very low genetic variation (i.e., only a few  
436 haplotypes were observed in the Southern). In this last case, two hypotheses could explain this pattern:  
437 1) the existence of both northern and southern glacial refugia and recent colonization from the southern  
438 to the northern part of the country, or 2) a unique refugium in the north and a recent recolonization to  
439 the south (Souto and Premoli, 2007).

440

441        Given the higher resolution provided by genomic SNPs, the present study unambiguously  
442 supports the two proposed genetic groups and shows that individuals from the Central distribution are  
443 more genetically like those from the Southern. Furthermore, a divergence time of 2.8Myr was  
444 estimated between the Northern and Central-Southern genetic groups, corresponding to the Late  
445 Pliocene and Early Pleistocene. In contrast, the divergence between populations into Northern and  
446 Central-Southern groups was much more recent (i.e., some 0.35Myr ago). These results emphasize the  
447 possible role of the glacial cycles and associated environmental changes as a possible driver of  
448 divergence in *E. coccineum*. Our study provides a refined picture of the possible divergence time  
449 between the North and Central-Southern genetic groups. Most phylogeographic studies have shown  
450 that populations located at southern latitudes and impacted by glaciations have experienced a recent  
451 and rapid expansion after glacial retreat, whereas northern populations have remained more stable  
452 (Kozak et al., 2006; Fontanella et al., 2008). Indeed, various studies (e.g., Markgraf, 1983; Vidal et al.,  
453 2005) propose the existence of refugia Northern of the area heavily impacted by ice (i.e., ranging  
454 between 36°S - 40°S and 54°S).

455

456        In this respect, it has been reported that during the Pleistocene glaciation, and its consequent  
457 climate change, dramatically altered the landscape of southern South America (Ramos and Ghiglione,  
458 2008; Martínez and Kutscher, 2011; Ponce et al., 2011). The numerous glacial advances and retreats  
459 have impacted the abundance and distribution of the local biota (Rabassa et al., 2011). As a result,  
460 southern south American forests are composed of paleo-flora with variable origins and histories  
461 (Villagrán and Hinojosa, 1997; Segovia et al., 2013). Fossil records indicate that during the Paleocene,  
462 this region supported a widespread tropical flora (Hinojosa and Villagrán 1997). Then, lineages  
463 preadapted to cool temperate conditions, widely distributed in Gondwana (Poole et al., 2003), joined in  
464 this area during the mid-Eocene, associated with a northward displacement of vegetational belts  
465 (Hinojosa et al., 2006). Finally, during the late Eocene and early Miocene, a mixed paleo-flora with  
466 similar Neotropical, Australasian and Australantarctic elements was established in response to colder  
467 and drier conditions from Antarctic glaciation and the drop in global temperature (Zachos et al., 2001).  
468 At present, one-third of the woody genera in southern South America are related to ancient paleo flora  
469 that occupied southern South America in pre-Quaternary times (Villagrán and Hinojosa 1997; Wardle  
470 et al., 2001; Hinojosa et al., 2006). Comparative phylogeographic studies in southern South America  
471 (Bruno et al., 2016; Perez, 2017; Sérsic et al., 2011) propose the existence of at least six areas where  
472 both terrestrial plant and animal species survived during glacial maxima. These areas show higher  
473 genetic diversity than areas where post-glacial colonization occurred, as in some parts of the southern  
474 Patagonian Steppe (Hewitt, 2001; Pinceel et al., 2005; Huck et al., 2009).

475

476        Our study found relatively high genetic diversity and polymorphic loci in Nahuelbuta, Curacautín,  
477 Chiloe Norte, Coyhaique, and Torres del Paine, and the number of private alleles was also high in these  
478 exact locations. Vidal-Rusell et al. (2011), based on cpDNA haplotypes, suggested that southern  
479 populations of *E. coccineum* had been affected by glacial cycles, declining gradually in size since the  
480 Pleistocene (100ky - 12ky ago) and then rapidly increasing after the LGM (). Hence, evidence of areas

481 of local expansion in the north is supported, particularly from the existence of various microrefugia  
482 located in Nahuelbuta (37°S - 72°W), Manzanar (38°S - 71°W, nearby Curacautín), Lagüen Ñadi (41°S  
483 - 73°W, nearby Chiloé) and Lago Verde (42°S - 71°W, nearby Coyhaique), as suggested by these  
484 authors. Combined with the observed pattern of genetic structure, these results support the hypothesis  
485 of *E. coccineum* survival during glacial periods of the Pleistocene glacial cycles in various refugia,  
486 including at least two refugia located south of 42°S. The existence of southern local refugia in *E.*  
487 *coccineum* is supported by fossil records of cold-tolerant plant species characteristic of the austral  
488 forest reported south of 42°S (Villagran et al., 1986; Markgraf, 1995). High levels of population  
489 divergence and phylogeographic structure attributed to isolation in multiple refugia have been detected  
490 in other cold-tolerant native trees such as *Fitzroya cupressoides* (Allnutt et al., 1999; Premoli et al.,  
491 2000; Premoli et al., 2003), *Pilgerodendron uviferum* (Premoli et al., 2001; Premoli et al., 2002; Ausin  
492 et al., 2003), *Araucaria araucana* (Bekessy et al., 2002), *Podocarpus nubigena* (Quiroga & Premoli,  
493 2010), *Nothofagus alpina* (Marchelli et al., 1998; Marchelli & Gallo, 2006), and *Nothofagus pumilio*  
494 (Premoli, 2004; Mathiasen & Premoli, 2010; Premoli et al., 2010).

495

#### 496 **Possible regional differences in pollen and seeds dispersal: effect on *Embothrium coccineum* gene** 497 **flow**

498 Our analysis of the genetic differentiation drivers in *E. coccineum* identified a significant pattern  
499 of isolation by distance (IBD). The existence of IBD indicates that gene flow is limited as geographic  
500 distance increases (Wright, 1943). However, some level of connectivity was observed between nearby  
501 populations, mainly in the Central and Southern groups. In plants such as *Dactylis glomerata* (Sun et  
502 al., 2017) and *Protea rapens* (Prunier et al., 2017), IBD patterns were detected similarly geographic  
503 scale, pointing out opportunities for increased allelic exchange between neighbouring populations.  
504 Moreover, our STRUCTURE, PCA, and NJ distance tree analyses, and estimation of pairwise-FST  
505 values, showed that populations from the Northern genetic group of *E. coccineum* are highly structured.  
506 Therefore, we hypothesize that the population structure differences between localities within the  
507 Northern and the Central-Southern genetic groups could be, in part, linked to the existence of a more  
508 restricted gene flow in the northern distribution.

509 Plant gene flow may occur via pollen and/or seed dispersal and is directly affected by the species'  
510 reproductive system (Martins, 1987). Seed dispersal decreases competition in areas close to the parent  
511 plant while allowing colonization of new sites (Rovere & Premoli, 2005). It also influences the genetic  
512 structure of populations since seed transport favours gene flow between and within populations  
513 (Hamrick & Nason, 1996) while imposing restrictions on the potential for intraspecific differentiation  
514 into ecotypes and subpopulations (Willson et al., 1994). *E. coccineum* has an anemochory seed  
515 dispersal and the density of dispersed seeds declines with distance from the tree: 95% of the seeds are  
516 dispersed 5 m around the mother tree, but a distance of 20 m has been detected (Rovere & Premoli,  
517 2005). These short dispersal distances limit the distribution of the seeds to the surroundings of the  
518 parent trees, causing a restricted movement in the population. Thus, could have significant genetic  
519 consequences as the local establishment of propagules results in a clustering of genetically related

520 individuals in the vicinity of the seed tree, which would generate a marked genetic structure (Rovere &  
521 Premoli, 2005).

522 On the other hand, Mathiasen et al. (2007) consider *E. coccineum* a species with low sensitivity to  
523 forest fragmentation due to a predominantly outcrossing breeding system, favouring pollen exchange  
524 among nearby fragments (Rovere & Chalcof, 2010). Indeed, genetic erosion by drift in forest patches  
525 could be counteracted by pollinator's movements, reducing inbreeding, and allowing the populations to  
526 retain sufficient heterozygosity through a continuous influx of alleles from locally distinct gene pools  
527 (Mathiasen et al., 2007). *E. coccineum*, gene flow is highly dependent on pollinating agents  
528 (Mathiasen, 2004; Rovere & Chalcof, 2010), and its flowers are visited by more than 20 species,  
529 including birds of the Orders Passeriformes and Apodiformes, and insects of the orders Hymenoptera,  
530 Diptera, Lepidoptera and Coleoptera (Chalcoff, 2008).

531 The composition of the pollinator assemblage affects both pollen transport and seed production.  
532 In regions with fewer *E. coccineum* pollinators, pollen limitation can be high, leading to low  
533 reproductive efficiency (i.e., few fruits produced to the available flowers) (Chalcoff, 2008), which  
534 would translate to populations presenting reduced genetic diversity and high genetic structure due to  
535 the lack of genetic exchange (Rovere and Chalcof 2010). Pollinating insects, generally characterized by  
536 low migration capacities, are distributed throughout the *E. coccineum* distribution range, even if they  
537 are abundant, and visit the plants more frequently in populations located in drier and sunnier climates  
538 (i.e., Northern locations) (Rovere and Chalcof, 2010). Contrastingly, birds, such as the hummingbird  
539 *Sephanoides sephaniodes*, are abundant in the Southern distribution, characterized by the coldest and  
540 most humid conditions that act as principal pollinators (Chalcoff, 2008). We propose that differences  
541 between Northern and Central-Southern *E. coccineum* pollinators' distribution and behaviour could  
542 lead to the observed genetic structure, with a lower dispersal capacity in the Northern locations, leading  
543 to a higher level of structure. However, other differences between regions, such as the level of  
544 landscape fragmentation, could also generate differences in gene flow due to the different flowering  
545 times detected between the localities of *E. coccineum* (Rovere & Chalcof, 2010). Differences in  
546 climatic conditions through the distribution range affect *E. coccineum* flowering period, with a  
547 flowering period beginning later (i.e., November – December) in populations located at high latitude  
548 and altitude when compared to the rest of the distribution (i.e., September- October) (Hoffmann, 1997;  
549 Chalcoff, 2008; Rovere & Chalcof, 2010). Limited gene flow between populations presenting  
550 differences in flowering phenology could also lead to the observed pattern of divergence between and  
551 within genetic groups by limiting exchange between close populations located at contrasting altitudes.

552

### 553 **Signatures of local adaptation in *Embothrium coccineum***

554 On the other hand, we sought to elucidate the effects of local adaptation and isolation by  
555 environment (IBE) from neutral processes, such as isolation by distance (IBD) or co-ancestry linked to  
556 the species' glacial history (IBA), in shaping among-population genetic differentiation of *E.*  
557 *coccineum*. Our results show that the observed pattern of genetic variation is, at least in part, the result  
558 of adaptation to the solid environmental gradients characterizing the species distribution range (Daniield  
559 & Veblen, 2000; Souto & Premoli, 2007; Souto et al., 2009; Souto and Smouse, 2013). Indeed, results

560 from the redundancy analysis (RDA) show that IBE was a significant driver of population structure and  
561 explained the most considerable among-population variation in *E. coccineum*, even after controlling for  
562 IBD and IBA. These results were corroborated by the much higher divergence detected between the  
563 Northern and Central-Southern genetic groups in the NJ tree based on the genetic distance calculated  
564 using only the outlier loci putatively under selection than the one observed when using all the loci  
565 genotyped.

566 Our gradient forest analysis identifies the Precipitation of the driest month as the most critical  
567 explanatory environmental variable explaining the *E. coccineum* pattern of genetic variation. The level  
568 of precipitation of the driest month distinguishes the northern neighbouring Mediterranean biome and  
569 the Patagonian stepe to the east, both with deficient summer precipitation, from the temperate biome  
570 inhabited by *E. coccineum* (Escobar et al., 2006). The second explanatory environmental variable was  
571 the mean temperature of the wettest quarter. In addition, GF also produced individual turnover  
572 functions for each potentially adaptive loci having a  $R^2 > 0$  (Fitzpatrick & Keller, 2014). This allows  
573 listing the potentially adaptive loci showing the highest correlation of allelic frequencies with the four  
574 environmental factors of the significant impact on *E. coccineum* populations. For 5 of the 59 outlier  
575 loci putatively under selection, a correlation between allelic frequency and Precipitation of the driest  
576 month, Mean temperature of the driest quarter, Mean Temperature of the Wettest Quarter, and Mean  
577 Temperature of Wettest Quarter was revealed. The covariation between latitude and these  
578 environmental variables translates into four outliers (i.e., S60272\_8, S27842\_67, S56778\_42, and  
579 S83147\_54), showing allelic frequency-changing gradually with latitude. These results showed that  
580 genetic differentiation in *E. coccineum* mainly was associated with environmental differences in  
581 summer rainfall and winter temperature.

582 This result is consistent with studies focused on other Proteaceae; with summer rainfall associated  
583 with differences in functional traits in *P. rapens* (Carlson et al., 2016) and in Protea section Exsertae  
584 (i.e., a group of six Protea species; Carlson et al., 2011; Prunier et al., 2012, Prunier et al., 2017).  
585 Differences in morphology associated with environmental gradients have already been reported in *E.*  
586 *coccineum* (Chalcof, 2008; Souto & Premoli, 2007), supporting the idea of possible local adaptation in  
587 the species. Access to water seems vital in determining the ecotypes observed in each region, with  
588 individuals from the Northern distribution growing in the driest climate and developing as short plants,  
589 of 0.5-2.5 m in height, with small leaves. In contrast, the individuals in a wet temperate forest of the  
590 Central distribution can reach more than 10 m in height and develop longer leaves (Souto et al., 2009;  
591 Souto & Smouse, 2013). Plants that develop in areas characterized by steady and high rainfall may be  
592 less heat and water-stressed, favouring the production of more extensive and broader leaves and  
593 maximizing the photosynthetic area when risks of overheating and drought stress are low (Givnish,  
594 1987; Hallik et al., 2009). Indeed, a reduction in leaf size has been proposed as one of the critical traits  
595 allowing plants to withstand water deficit (Baldocchi and Xu, 2007, Peguero-Pina et al., 2014). For  
596 example, a reduction in *Quercus faginea* leaf size has been proposed as one of the critical traits  
597 allowing Mediterranean oaks to withstand the water deficit characteristic of the region (Baldocchi and  
598 Xu, 2007, Peguero-Pina et al., 2014).

599

## 600 **Conclusion and prospects**

601       Known patterns of morphological variations coupled with new genetic diversity and structure  
602 results obtained during the present study revealed a long history of geographic isolation and local  
603 adaptation in *E. coccineum* in South American temperate forests. Here, we observed a robust genetic  
604 structure in *E. coccineum* that could be linked to local adaptation, especially differential access to water  
605 during the driest months. These results support previous studies that have reported the existence of  
606 distinct ecotypes along with the natural distribution. Identifying and quantifying the environmental  
607 variables structuring population genetic variation could inform management decisions for conservation,  
608 restoration, or reforestation purposes. For example, knowledge about the association between genotype  
609 and environment is essential for selecting the proper seed when tree populations are locally adapted, as  
610 is the case of *E. coccineum*. However, as the environment changes, nonlocal seed sources may be  
611 considered for restoration or reforestation purposes based on their match to the novel environment.  
612 Therefore, seed transfer guidelines can benefit from the knowledge of factors structuring the landscape  
613 of genetic variation. Many conservation efforts rely on delineating distinct populations for management  
614 but often ignore the continuous nature of landscape variation and its potential relationship with local  
615 adaptation or other processes that lead to genotype-environment associations. The present results  
616 advocate for seed transfer guidelines that should be restricted within each ecological zone concerning  
617 land management decisions. The strong genetic structure also offers an opportunity to explore local  
618 adaptation among groups further and exploit these for agroforestry.

619

620

621

622 **References**

623

- 624 1. Abarzúa, A. M., Villagrán, C., and Moreno, P. I. (2004). Deglacial and postglacial climate history  
625 in east-center Isla Grande de Chiloé, southern Chile (43 S). *Quaternary Research*, 62(1), 49-59.
- 626 2. Alberdi, M. and C. Donoso. (2004). Variación en *Embothrium coccineum* J.R. et G.Forster (notro  
627 o ciruelillo), en C.Donosó, A. Premoli, L. Gallo and F. Ipinza (eds.). Variación intraespecífica en  
628 las especies arbóreas de los bosques templados de Chile y Argentina, *Editorial Universitaria*,  
629 *Santiago de Chile*, 345-355.
- 630 3. Allnutt, T. R., Newton, A. C., Lara, A., Premoli, A., Armesto, J. J., Vergara, R., and Gardner, M.  
631 (1999). Genetic variation in *Fitzroya cupressoides* (alerce), a threatened South American conifer.  
632 *Molecular Ecology*, 8(6), 975-987.
- 633 4. Allnutt, T. R., Newton, A. C., Premoli, A., and Lara, A. (2003). Genetic variation in the  
634 threatened South American conifer *Pilgerodendron uviferum* (Cupressaceae), detected using  
635 RAPD markers. *Biological Conservation*, 114(2), 245-253.
- 636 5. Ausin, I., Alonso-Blanco, C., and Martínez-Zapater, J. M. (2004). Environmental regulation of  
637 flowering. *International Journal of Developmental Biology*, 49(5-6), 689-705.
- 638 6. Baldocchi, D. D., and Xu, L. (2007). What limits evaporation from Mediterranean oak  
639 woodlands—The supply of moisture in the soil, physiological control by plants or the demand by  
640 the atmosphere?. *Advances in Water Resources*, 30(10), 2113-2122.
- 641 7. Bekessy, S. A., Allnutt, T. R., Premoli, A. C., Lara, A., Ennos, R. A., Burgman, M. A., Newton,  
642 A. C. (2002). Genetic variation in the vulnerable and endemic Monkey Puzzle tree, detected using  
643 RAPDs. *Heredity*, 88(4), 243-249.
- 644 8. Borcard, D., and Legendre, P. (2002). All-scale spatial analysis of ecological data by means of  
645 principal coordinates of neighbour matrices. *Ecological modelling*, 153(1-2), 51-68.
- 646 9. Bouckaert, R. R. (2010). DensiTree: making sense of sets of phylogenetic trees. *Bioinformatics*,  
647 26(10), 1372-1373.
- 648 10. Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C. H., Xie, D., Drummond, A. J. (2014).  
649 BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Comput Biol*, 10(4),  
650 e1003537.
- 651 11. Bray, J. R., and Curtis, J. T. (1957). An ordination of the upland forest communities of southern  
652 Wisconsin. *Ecological monographs*, 27(4), 326-349
- 653 12. Bruno, M. C., Mapelli, F. J., Casciotta, J. R., Almirón, A. E., and Lizarralde, M. S. (2016).  
654 Phylogeography of *Cnesterodon decemmaculatus* (Cyprinodontiformes: poeciilidae) in Southern

- 655 Pampas, Argentina: ancient versus recent patterns in freshwater fishes. *Environmental biology of*  
656 *fishes*, 99(2-3), 293-307.
- 657 13. Bryant, D., Bouckaert, R., Felsenstein, J., Rosenberg, N. A., and RoyChoudhury, A. (2012).  
658 Inferring species trees directly from biallelic genetic markers: bypassing gene trees in a full  
659 coalescent analysis. *Molecular biology and evolution*, 29(8), 1917-1932.
- 660 14. Carlson, J. E., Adams, C. A., and Holsinger, K. E. (2016). Intraspecific variation in stomatal  
661 traits, leaf traits and physiology reflects adaptation along aridity gradients in a South African  
662 shrub. *Annals of botany*, 117(1), 195-207.
- 663 15. Carlson, J. E., Holsinger, K. E., Prunier, R. (2011). Plant responses to climate in the Cape  
664 Floristic Region of South Africa: Evidence for adaptive differentiation in the Proteaceae.  
665 *Evolution* 65, 108-124.
- 666 16. Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., Cresko, W. A. (2013). Stacks: an  
667 analysis tool set for population genomics. *Molecular ecology*, 22(11), 3124-3140.
- 668 17. Chalcoff, V. (2008). Variación geográfica en los rasgos florales de *Embothrium coccineum*  
669 (Proteaceae) y su relación con los ensambles de polinizadores a ambos lados de los Andes  
670 australes. *Tesis doctoral. Universidad Nacional del Comahue, Bariloche*.
- 671 18. Chalcoff, V. R., Ezcurra, C., and Aizen, M. A. (2008). Uncoupled geographical variation between  
672 leaves and flowers in a South-Andean Proteaceae. *Annals of botany*, 102(1), 79-91.
- 673 19. Collins, B.G. and Rebelo, T. (2006). Pollination biology of the Proteaceae in Australia and  
674 southern Africa. *Australian Journal of Ecology*, 12(4), 387-421.
- 675 20. Comps, B., Gömöry, D., Letouzey, J., Thiébaud, B., Petit, R. J. (2001). Diverging trends between  
676 heterozygosity and allelic richness during postglacial colonization in the European beech.  
677 *Genetics*, 157(1), 389-397.
- 678 21. Dimitri, M. J. (1972). La región de los bosques andino-patagónicos: sinopsis general. *Colección*  
679 *científica del Instituto Nacional de Tecnología Agropecuaria (INTA), Buenos Aires, Argentina*.  
680 381 pp.
- 681 22. Ellis, N., Smith, S. J., and Pitcher, C. R. (2012). Gradient forests: calculating importance  
682 gradients on physical predictors. *Ecology*, 93(1), 156-168.
- 683 23. Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., Mitchell, S. E.  
684 (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species.  
685 *PloS one*, 6(5), e19379.
- 686 24. Escobar, B., C. Donoso, C., Souto A., Zúñiga, A. 2006. *Embothrium coccineum* J.R. et. G.  
687 Forster. In: Donoso, C. (ed.). *Autoecología de las especies arbóreas de los bosques templados de*  
688 *Chile y Argentina*. Marisa Cuneo Ediciones. pp. 345 - 356.

- 689 25. Falush, D., Stephens, M., Pritchard, J. K. (2003). Inference of population structure using  
690 multilocus genotype data: linked loci and correlated allele frequencies. *Genetics*, 164(4), 1567-  
691 1587.
- 692 26. Fesq-Martin, M., Friedmann, A., Peters, M., Behrmann, J., Kilian, R. (2004). Late-glacial and  
693 Holocene vegetation history of the Magellanic rain forest in southwestern Patagonia, Chile.  
694 *Vegetation History and Archaeobotany*, 13(4), 249-255.
- 695 27. Fitzpatrick, M., Keller, S. (2014). Ecological genomics meets community-level modelling of  
696 biodiversity: mapping the genomic landscape of current and future environmental adaptation.  
697 *Ecology Letters*, 18: 1-16.
- 698 28. Foll, M., Beaumont, M. A., and Gaggiotti, O. E. (2008). An approximate Bayesian computation  
699 approach to overcome biases that arise when using AFLP markers to study population structure.  
700 *Genetics*, 179(2), 927-39.
- 701 29. Fontanella, F. M., Feldman, C. R., Siddall, M. E., Burbrink, F. T. (2008). Phylogeography of  
702 *Diadophis punctatus*: extensive lineage diversity and repeated patterns of historical demography  
703 in a trans-continental snake. *Molecular Phylogenetics and Evolution*, 46(3), 1049-1070.
- 704 30. Goudet, J. (2005). Hierfstat, a package for R to compute and test hierarchical F-statistics.  
705 *Molecular Ecology Notes*, 5(1), 184-186.
- 706 31. Grossiord, C., Sevanto, S., Dawson, T. E., Adams, H. D., Collins, A. D., Hewitt, G. M. (2004).  
707 Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the*  
708 *Royal Society of London. Series B: Biological Sciences*, 359(1442), 183-195.
- 709 32. Hallik, L., Niinemets, Ü., Wright, I. J. (2009). Are species shade and drought tolerance reflected  
710 in leaf-level structural and functional differentiation in Northern Hemisphere temperate woody  
711 flora?. *New Phytologist*, 184(1), 257-274.
- 712 33. Hamrick, J. L., and Nason, J. D. (1996). Consequences of dispersal in plants. *Population*  
713 *dynamics in ecological space and time*, 203-236.
- 714 34. Hasbún, R., González, J., Iturra, C., Fuentes, G., Alarcón, D., Ruiz, E. (2016). Using genome-  
715 wide SNP discovery and genotyping to reveal the main source of population differentiation in  
716 *Nothofagus dombeyi* (Mirb.) Oerst. in Chile. *International journal of genomics*, 2016.
- 717 35. Heusser, C. J., and Heusser, L. E. (2006). Submillennial palynology and palaeoecology of the last  
718 glaciation at Taiquemo (~ 50,000 cal yr, MIS 2–4) in southern Chile. *Quaternary Science*  
719 *Reviews*, 25(5-6), 446-454.
- 720 36. Heusser, C. J., Heusser, L. E., Lowell, T. V. (1999). Paleocology of the southern Chilean lake  
721 district-isla grande de Chiloé during middle-late Illanquihue glaciation and deglaciation.  
722 *Geografiska Annaler: Series A, Physical Geography*, 81(2), 231-284.

- 723 37. Hewitt, G. M., and Ibrahim, K. M. (2001). Inferring glacial refugia and historical migrations with  
724 molecular phylogenies. In Silvertown, J. and Antonovics, J. (Eds.), *Integrating ecology and*  
725 *evolution in a spatial context*, 97– 123.
- 726 38. Hinojosa, L. F., Armesto, J. J., Villagrán, C. (2006). Are Chilean coastal forests pre-Pleistocene  
727 relicts? Evidence from foliar physiognomy, palaeoclimate, and phytogeography. *Journal of*  
728 *Biogeography*, 33(2), 331-341.
- 729 39. Hinojosa, L. F., and Villagrán, C. (1997). History of the forests of southern South America I:  
730 paleobotanical, geological and climatical background on Tertiary of southern South America.  
731 *Revista Chilena de Historia Natural*, 70, 225–39.
- 732 40. Hoffmann, J. (1997). Flora silvestre de Chile. Zona araucana. Arboles, arbustos y enredaderas  
733 leñosas. *Ediciones Fundación Claudio Gay*.
- 734 41. Huck, S., Büdel, B., Kadereit, J. W., Printzen, C. (2009). Range-wide phylogeography of the  
735 European temperate-montane herbaceous plant *Meum athamanticum* Jacq.: evidence for  
736 periglacial persistence. *Journal of Biogeography*, 36(8), 1588-1599.
- 737 42. Johnson, S. D., Linder, H. P., Allsopp, N., Colville, J. F., Verboom, G. A. (2014). Speciation and  
738 extinction in the Greater Cape Floristic Region. *Fynbos: Ecology, Evolution, and Conservation of*  
739 *a Megadiverse Region*, 119.
- 740 43. Jombart, T. (2008). adegenet: a R package for the multivariate analysis of genetic markers.  
741 *Bioinformatics*, 24(11), 1403-1405.
- 742 44. Loguercio, G. A., Donoso, P. J., Müller-Using, S., Dezzotti, A., Urretavizcaya, M. F., Navarro, C.  
743 O., Caselli, M. (2018). Silviculture of temperate mixed forests from South America. In *Dynamics,*  
744 *Silviculture and Management of Mixed Forests*, 271-317.
- 745 45. Luu, K., Bazin, E., Blum, M. G. (2017). pcadapt: an R package to perform genome scans for  
746 selection based on principal component analysis. *Molecular ecology resources*, 17(1), 67-77.
- 747 46. Markgraf, V. (1983). Late and postglacial vegetational and paleoclimatic changes in subantarctic,  
748 temperate, and arid environments in Argentina. *Palynology*, 7, 43-70.
- 749 47. Markgraf, V. (1995). Neogene paleoenvironmental and paleoclimatic changes in southern  
750 temperate ecosystems — A southern perspective. *Trends in Ecology and Evolution*, 10, 143-147.
- 751 48. Martínez, O. and Kutschker, A. (2011). The ‘Rodados Patagónicos’ (Patagonian shingle  
752 formation) of eastern Patagonia: environmental conditions of gravel sedimentation. *Biological*  
753 *Journal of the Linnean Society*, 103(2), 336-345.
- 754 49. Martins PS (1987) Estrutura populacional, fluxo gênico e conservação “*in situ*”. *Ipef*, 71-78.

- 755 50. Mathiasen, P. (2004). Efectos de la fragmentación del bosque templado sobre la demografía y  
756 estructura genética de *Embothrium coccineum* Forst. (Proteaceae) en el sur de Chile (*Doctoral*  
757 *dissertation, Tesis de Licenciatura, Universidad Nacional del Comahue, Bariloche, Argentina*).
- 758 51. Mathiasen, P. and Premoli, A. (2010). Out in the cold: Genetic variation of *Nothofagus pumilio*  
759 (*Nothofagaceae*) provides evidence for latitudinally distinct evolutionary histories in austral South  
760 America. *Molecular Ecology*, 19, 371- 385.
- 761 52. Mathiasen, P., Rovere, A. E., Premoli, A. C. (2007). Genetic structure and early effects of  
762 inbreeding in fragmented temperate forests of a self-incompatible tree, *Embothrium coccineum*.  
763 *Conservation biology*, 21(1), 232-240.
- 764 53. Moreno, P. I. (1997). Vegetation and climate near Lago Llanquihue in the Chilean Lake District  
765 between 20200 and 9500 14C yr BP. *Journal of Quaternary Science: Published for the*  
766 *Quaternary Research Association*, 12(6), 485-500.
- 767 54. Moreno, P. I., and León, A. L. (2003). Abrupt vegetation changes during the last glacial to  
768 Holocene transition in mid-latitude South America. *Journal of Quaternary Science*, 18(8), 787-  
769 800.
- 770 55. Moreno, P. I., Lowell, T. V., Jacobson Jr, G. L., Denton, G. H. (1999). Abrupt vegetation and  
771 climate changes during the last glacial maximum and last termination in the Chilean lake district: a  
772 case study from Canal de la Puntilla (41 S). *Geografiska Annaler: Series A, Physical Geography*,  
773 81(2), 285-311.
- 774 56. Oksanen, J., Kindt, R., Legendre, P., O'Hara, B., Stevens, M. H. H., Oksanen, M. J., Suggests, M.  
775 A. S. S. (2007). The vegan package. *Community ecology package*, 10(631-637), 719.
- 776 57. Orsini, L., Vanoverbeke, J., Swillen, I., Mergeay, J., De Meester, L. (2013). Drivers of population  
777 genetic differentiation in the wild: isolation by dispersal limitation, isolation by adaptation and  
778 isolation by colonization. *Molecular ecology*, 22(24), 5983-5999.
- 779 58. Paris, J. R., Stevens, J. R., Catchen, J. M. (2017). Lost in parameter space: a road map for stacks.  
780 *Methods in Ecology and Evolution*, 8(10), 1360-1373.
- 781 59. Pinceel, J., Jordaens, K., Van Houtte, N., Bernon, G., Backeljau, T. (2005). Population genetics  
782 and identity of an introduced terrestrial slug: *Arion subfuscus* sl in the north-east USA  
783 (*Gastropoda, Pulmonata, Arionidae*). *Genetica*, 125(2-3), 155-171.
- 784 60. Pembleton, L. W., Cogan, N. O., Forster, J. W. (2013). StAMPP: An R package for calculation of  
785 genetic differentiation and structure of mixed-ploidy level populations. *Molecular ecology*  
786 *resources*, 13(5), 946-952.
- 787 61. Perez, M. (2017). Comparative phylogeography of woody plant species from Mediterranean and  
788 temperate forest of southern South America: testing the influence of plant functional strategies on

- 789 the response of species to past climatic changes. *Fondecyt de iniciación*.  
790 <http://repositorio.conicyt.cl/handle/10533/236747>
- 791 62. Poole, I., Mennega, A. M., Cantrill, D. J. (2003). Valdivian ecosystems in the Late Cretaceous  
792 and Early Tertiary of Antarctica: further evidence from myrtaceous and eucryphiaceous fossil  
793 wood. *Review of Palaeobotany and Palynology*, 124(1-2), 9-27.
- 794 63. Ponce, J. F., Rabassa, J., Coronato, A., Borromei, A. M. (2011). Paleogeographic evolution of the  
795 Atlantic coast of Pampa and Patagonia since the Last Glacial Maximum to the Middle Holocene.  
796 *Biological Journal of the Linnean Society*, 103, 363-379.
- 797 64. Premoli, A. C., Kitzberger, T., Veblen, T. (2000). Isozyme variation and recent biogeographical  
798 history of the long-lived conifer *Fitzroya cupressoides*. *Journal of Biogeography*, 27(2), 251-260.
- 799 65. Premoli, A. C., Souto, C. P., Allnut, T. R., Newton, A. C. (2001). Effects of population  
800 disjunction on isozyme variation in the widespread *Pilgerodendron uviferum*. *Heredity*, 87(3),  
801 337-343.
- 802 66. Premoli, A. C., Souto, C. P., Rovere, A. E., Allnut, T. R., Newton, A. C. (2002). Patterns of  
803 isozyme variation as indicators of biogeographic history in *Pilgerodendron uviferum* (D. Don)  
804 Florin. *Diversity and Distributions*, 8(2), 57-66.
- 805 67. Premoli, A. C., Vergara, R., Souto, C. P., Lara, A., Newton, A. C. (2003). Lowland valleys shelter  
806 the ancient conifer *Fitzroya cupressoides* in the Center Depression of southern Chile. *Journal of*  
807 *the Royal Society of New Zealand*, 33(3), 623-631.
- 808 68. Premoli, A.C. (2004). Variación en *Nothofagus pumilio*. *Variación intraespecífica en las especies*  
809 *de los bosques templados de Chile y Argentina*, Editorial Universitaria, Santiago, Chile, 145-172.
- 810 69. Premoli, A. C., Mathiasen, P., Kitzberger, T. (2010). Southern-most *Nothofagus* trees enduring  
811 ice ages: genetic evidence and ecological niche retrodiction reveal high latitude (54° S) glacial  
812 refugia. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 298(3-4), 247-256.
- 813 70. Prunier, R., Akman, M., Kremer, C. T., Aitken, N., Chuah, A., Borevitz, J., Holsinger, K. E.  
814 (2017). Isolation by distance and isolation by environment contribute to population differentiation  
815 in *Protea repens* (Proteaceae L.), a widespread South African species. *American journal of*  
816 *botany*, 104(5), 674-684.
- 817 71. Prunier, R., Holsinger, K. E., Carlson, J. E. (2012). The effect of historical legacy on adaptation:  
818 do closely related species respond to the environment in the same way?. *Journal of evolutionary*  
819 *biology*, 25(8), 1636-1649.
- 820 72. Quiroga, M.P., and Premoli, A.C. (2010). Genetic structure of *Podocarpus nubigena*  
821 (Podocarpaceae) provides evidence of Quaternary and ancient historical events.  
822 *Palaeogeography, Palaeoclimatology, Palaeoecology*, 285(3-4), 186-193.

- 823 73. Rabassa, J., Coronato, A. M., Salemme, M. (2005). Chronology of the Late Cenozoic Patagonian  
824 glaciations and their correlation with biostratigraphic units of the Pampean region (Argentina).  
825 *Journal of South American Earth Sciences*, 20(1-2), 81-103.
- 826 74. Rabassa, J., Coronato, A., Martinez, O. (2011). Late Cenozoic glaciations in Patagonia and Tierra  
827 del Fuego: an updated review. *Biological Journal of the Linnean Society*, 103(2), 316-335.
- 828 75. Ramos, V. A., and Ghiglione, M. C. (2008). Tectonic evolution of the Patagonian Andes.  
829 *Developments in quaternary sciences*, 11, 57-71.
- 830 76. Rannala, B., and Yang, Z. (2003). Bayes estimation of species divergence times and ancestral  
831 population sizes using DNA sequences from multiple loci. *Genetics*, 164(4), 1645-1656.
- 832 77. Rovere, A. E., and Premoli, A. C. (2005). Dispersión asimétrica de semillas de *Embothrium*  
833 *coccineum* (Proteaceae) en el bosque templado de Chiloé, Chile. *Ecología austral*, 15(1), 1-7.
- 834 78. Rovere, A. E., and Chalcoff, V. R. (2010). *Embothrium coccineum* JR Forst. et G. Forst.  
835 *Kurtziana*, 35(2), 23-33.
- 836 79. Sauquet, H., Weston, P. H., Anderson, C. L., Barker, N. P., Cantrill, D. J., Mast, A. R.,  
837 Savolainen, V. (2009). Contrasted patterns of hyperdiversification in Mediterranean hotspots.  
838 *Proceedings of the National Academy of Sciences*, 106(1), 221-225.
- 839 80. Segovia, R. A., Pérez, M. F., Hinojosa, L. F. (2012). Genetic evidence for glacial refugia of the  
840 temperate tree *Eucryphia cordifolia* (Cunoniaceae) in southern South America. *American Journal*  
841 *of Botany*, 99(1), 121-129.
- 842 81. Sersic, A. N., Cosacov, A., Cocucci, A. A., Johnson, L. A., Pozner, R., Avila, L. J., Morando, M.  
843 (2011). Emerging phylogeographical patterns of plants and terrestrial vertebrates from Patagonia.  
844 *Biological Journal of the Linnean Society*, 103(2), 475-494.
- 845 82. Slatkin, M. (1993). Isolation by distance in equilibrium and non-equilibrium populations.  
846 *Evolution*, 47(1), 264-279.
- 847 83. Smith-Ramírez, C., Rovere, A. E., Núñez-Ávila, M. C., Armesto, J. J. (2007). Habitat  
848 fragmentation and reproductive ecology of *Embothrium coccineum*, *Eucryphia cordifolia* and  
849 *Aextoxicon punctatum* in southern temperate rainforests. *Biodiversity loss and conservation*, 102-  
850 119.
- 851 84. Souto, C. P., and Smouse, P. E. (2014). Correlated morphological and genetic patterns in  
852 *Embothrium coccineum* (Proteaceae) across climate and geography: can *Embothrium* survive  
853 patagonian climate change?. *Australian Journal of Botany*, 61(7), 516-527.
- 854 85. Souto, C. P., and Premoli, A. C. (2007). Genetic variation in the widespread *Embothrium*  
855 *coccineum* (Proteaceae) endemic to Patagonia: effects of phylogeny and historical events.  
856 *Australian Journal of Botany*, 55(8), 809-817.

- 857 86. Stange, M., Sánchez-Villagra, M., Salzburger, W., Matschiner, M. (2018). Bayesian Divergence-  
858 Time Estimation with Genome-Wide Single-Nucleotide Polymorphism Data of Sea Catfishes  
859 (Ariidae) Supports Miocene Closure of the Panamanian Isthmus. *Systematic Biology*, 67, 681-  
860 699.
- 861 87. Steubing, L., Alberdi, M., Wenzel, H. (1983). Seasonal changes of cold resistance of Proteaceae  
862 of the South Chilean laurel forest. *Vegetatio*, 52(1), 35-44.
- 863 88. Storey, J. D., Taylor, J. E., Siegmund, D. (2004). Strong control, conservative point estimation  
864 and simultaneous conservative consistency of false discovery rates: a unified approach. *Journal of*  
865 *the Royal Statistical Society: Series B (Statistical Methodology)*, 66(1), 187-205.
- 866 89. Torres-Díaz, C., Ruiz, E., González, F., Fuentes, G., Cavieres, L. A. (2007). Genetic diversity in  
867 *Nothofagus alessandrii* (Fagaceae), an endangered endemic tree species of the Coastal Maulino  
868 Forest of Center Chile. *Annals of Botany*, 100(1), 75-82.
- 869 90. Turchetto-Zolet, A. C., Pinheiro, F., Salgueiro, F., Palma-Silva, C. (2013). Phylogeographical  
870 patterns shed light on evolutionary process in South America. *Molecular Ecology*, 22(5), 1193-  
871 1213.
- 872 91. Varas-Myrik, A., Sepúlveda-Espinoza, F., Fajardo, A., Alarcón, D., Toro-Núñez, Ó., Castro-  
873 Nallar, E., & Hasbún, R. (2022). Predicting climate change-related genetic offset for the  
874 endangered southern South American conifer *Araucaria araucana*. *Forest Ecology and*  
875 *Management*, 504, 119856.
- 876 92. Vidal, M. A., Ortiz, J. C., Ramirez, C. C., Lamborot, M. (2005). Intraspecific variation in  
877 morphology and sexual dimorphism in *Liolaemus tenuis* (Tropiduridae). *Amphib-Repti*, 26, 343-  
878 351.
- 879 93. Vidal-Russell, R., Souto, C. P., Premoli, A. C. (2011). Multiple Pleistocene refugia in the  
880 widespread Patagonian tree *Embothrium coccineum* (Proteaceae). *Australian Journal of Botany*,  
881 59(4), 299-314.
- 882 94. Villagrán, C. (1988). Expansion of Magellanic Moorland during the late Pleistocene:  
883 Palynological evidence from northern Isla de Chiloé, Chile. *Quaternary Research*, 30(3), 304-  
884 314.
- 885 95. Villagrán, C. (1991). Historia de los bosques templados del sur de Chile durante el Tardiglacial y  
886 Postglacial. *Revista Chilena de Historia Natural*, 64(3), 447-460.
- 887 96. Villagrán, C., and Hinojosa, L. F. (1997). History of the forests of southern South America II:  
888 phytogeographical analysis. *Revista Chilena de Historia Natural*, 70, 241-67.
- 889 97. Wang, I. J., and Bradburd, G. S. (2014). Isolation by environment. *Molecular ecology*, 23(23),  
890 5649-5662.

- 891 98. Wang, I. J., and Summers, K. (2010). Genetic structure is correlated with phenotypic divergence  
892 rather than geographic isolation in the highly polymorphic strawberry poison-dart frog. *Molecular*  
893 *Ecology*, 19(3), 447-458.
- 894 99. Wardle, P., Ezcurra, C., Ramírez, C., Wagstaff, S. (2001). Comparison of the flora and vegetation  
895 of the southern Andes and New Zealand. *New Zealand Journal of Botany*, 39(1), 69-108.
- 896 100. Willson, M. F., De Santo, T. L., Sabag, C., Armesto, J. J. (1994). Avian communities of  
897 fragmented south-temperate rainforests in Chile. *Conservation biology*, 8(2), 508-520.
- 898 101. Wright, S. (1943). Isolation by distance. *Genetics*, 28, 114-138.
- 899 102. Zachos, J., Pagani, M., Sloan, L., Thomas, E., Billups, K. (2001). Trends, rhythms, and  
900 aberrations in global climate 65 Ma to present. *science*, 292(5517), 686-693.
- 901 103. Zegers, C. D. (1994). Bosques templados de Chile y Argentina. *Editorial Universitaria*.
- 902 104. Zeppel, M. J., Wilks, J. V., and Lewis, J. D. (2014). Impacts of extreme precipitation and seasonal  
903 changes in precipitation on plants. *Biogeosciences*, 11, 3083-3093.
- 904

905 **ACKNOWLEDGEMENTS**

906 A.B.B: Financial support was provided by both ANID/FONDECYT N° 3160551 and N° 11221214.  
907 D.A. was supported by FONDECYT 3200675 grant. O.T.N. was funded by CONICYT PAI  
908 Subvencion a la instalacion en la Academia convocatoria 2019 77190055. We thank Mr Cesar Tascón  
909 and Dr Luis Corcuera for allowing the access and the collection of samples within the Fundo Niagara  
910 and the Katalapi Park. Our thanks to Ximena Steinberg, Cristian Muñoz, Francisco Valdebenito and  
911 Alex Fajardo for their support and help during samples collection in Northern and Southern locations.  
912 The authors thank Catalina Mardones for her enthusiastic support in the samples processing.

913

914 **DATA AVAILABILITY STATEMENT**

915 The authors declare that all the work presented in this manuscript was done according to the standards  
916 of the Ecology and Evolution. Raw GBS data used in this study has been already published in ncbi-  
917 SRA database and is available at <http://www.ncbi.nlm.nih.gov/sra/?term=PRJNA783610>.

918

919 **CONFLICT OF INTEREST**

920 On behalf of all authors, the corresponding author states that there is no conflict of interest.

921

922 **AUTHOR CONTRIBUTIONS**

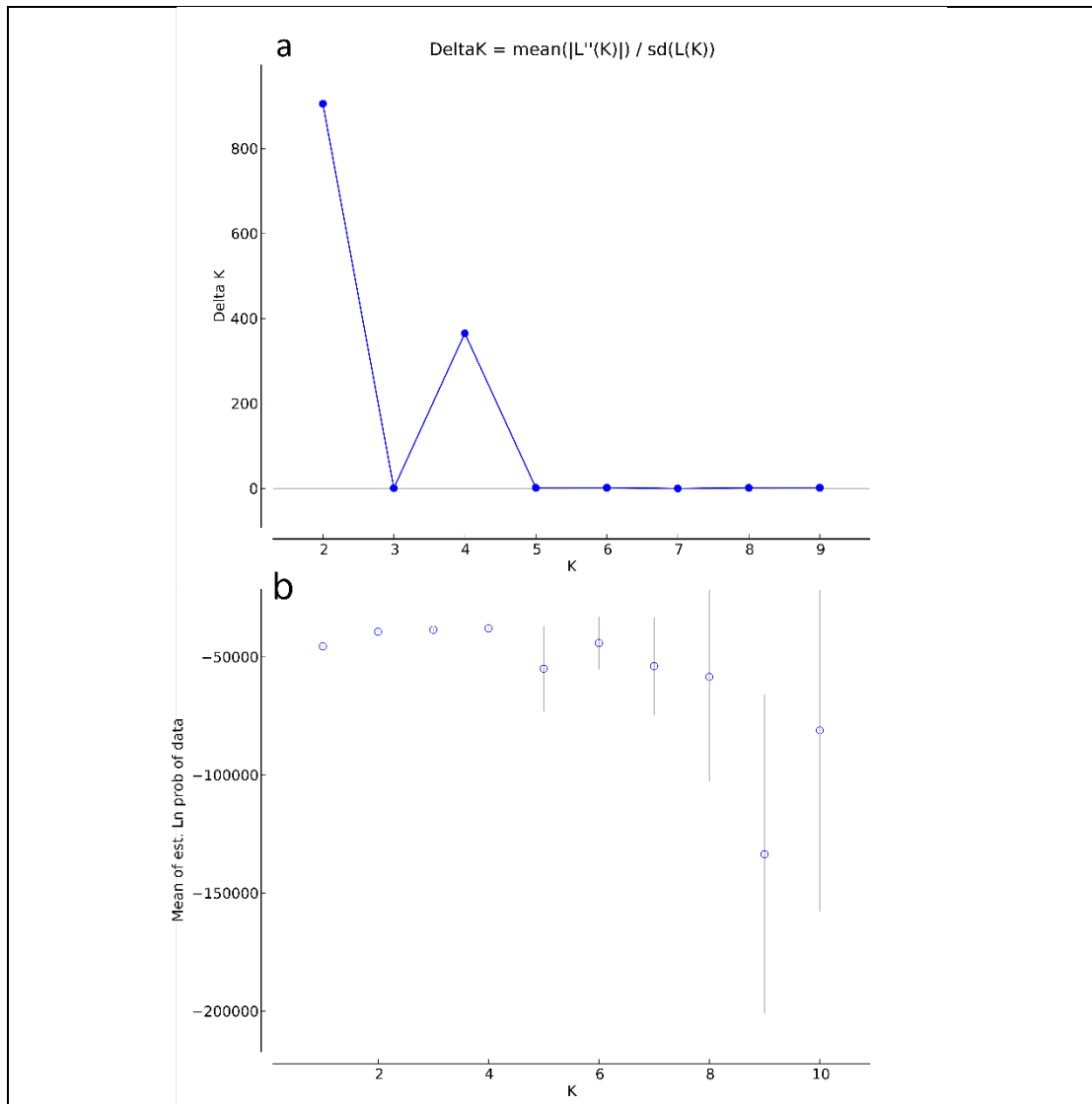
923 All authors contributed to the design of the study. Material preparation, data collection and analysis  
924 were carried out by Francisco Sepúlveda-Espinoza, Ariana Bertin-Benavides, Rodrigo Hasbún, Oscar  
925 Toro-Núñez, Antonio Varas-Myrik, Diego Alarcón and Marie-Laure Guillemin. The first draft of the  
926 manuscript was written by Francisco Sepúlveda-Espinoza, Ariana Bertin-Benavides and Marie-Laure  
927 Guillemin and all authors commented on earlier versions of the manuscript. All authors read and  
928 approved the final manuscript.

929 **Appendix I: Environmental characteristics of each sampling location of *Embothrium coccineum*.**

930 GPS coordinates are given for each sampling location. Environmental variables: bio 14, Precipitation of the driest month (mm); bio 8, Mean temperature of the wettest  
 931 quarter (°C); Elv, Elevation (mamsl); bio 9, Mean temperature of the driest quarter (°C); bio 4, Temperature seasonality (°C); bio 3, Isothermality (%); bio 13, Precipitation  
 932 of the wettest month (mm).

<b>Locations</b>	<b>ID</b>	<b>Latitude</b>	<b>Longitude</b>	<b>bio 14</b>	<b>bio 8</b>	<b>Elv</b>	<b>bio 9</b>	<b>bio 4</b>	<b>bio 3</b>	<b>bio 13</b>
Chillán	Ch	-36.908	-71.504	23	3.9	1216	14.0	4.1	55	257
Curacautín	Cu	-37.642	-73.099	28	4.4	989	11.2	2.8	57	297
Nahuelbuta	Nah	-38.455	-71.734	53	5.5	797	14.7	3.7	56	390
Puerto Montt	PM	-41.521	-72.753	95	7.9	81	14.9	2.9	49	267
Chiloe Norte	ChlN	-42.296	-73.459	90	7.7	129	13.5	2.4	47	333
Chiloe Sur	ChlS	-42.783	-73.806	85	8.0	97	13.4	2.3	45	303
Pumalín	Pu	-42.536	-72.496	96	7.4	22	15.4	3.3	49	309
Coyhaique	Coy	-46.505	-73.066	98	4.4	175	9.9	3.2	44	181
Chile Chico	ChCh	-46.649	-72.362	35	2.6	428	11.8	3.7	44	101
Torres del Paine	TP	-51.361	-72.799	37	9.1	33	9.7	3.4	48	67

933 Appendix II. Results of the model-based Bayesian clustering approach. Highest (a)  $\Delta K$  and (b)  
934  $\text{LnP(D)}$  show that the best  $K$  was  $K = 2$ .



935

936

