Weak founder effects but significant spatial genetic imprint of recent contraction and expansion of European beech populations

Tonya A. Lander, Etienne E. Klein, Anne Roig, Sylvie Oddou-Muratorio

1 **Title: Weak founder effects but significant spatial genetic imprint of**
2 **recent contraction and expansion of European beech populations.**

3 **Authors:** Tonya A. Lander[1,2], Etienne K Klein[3], Anne Roig[1], Sylvie Oddou-Muratorio[1]
4
5 **Affiliations :**
6 [1] URFM, INRAE, 84 000 Avignon, France
7 [2] Department of Plant Sciences, University of Oxford, South Parks Road, Oxford OX1 3RB,
8 UK
9 [3] BioSp, INRAE, 84 000 Avignon, France
10 **Corresponding author:** Sylvie Oddou-Muratorio
11 sylvie.muratorio@inrae.fr
12

14

Abstract

15 **Abstract**

16 Understanding the ecological and evolutionary processes occurring during species range
17 shifts is important in the current context of global change. Here, we investigate the interplay
18 between recent expansion, gene flow and genetic drift, and their consequences for genetic
19 diversity and structure at landscape and local scales in European beech (*Fagus sylvatica* L.)

20 On Mont Ventoux, South-Eastern France, we located beech forest refugia at the time of
21 the most recent population minimum, approximately 150 years ago, and sampled 71
22 populations (2042 trees) in both refugia and expanding populations over an area of 15,000 ha.
23 We inferred patterns of gene flow and genetic structure using 12 microsatellite markers.

24 We identified six plots as originating from planting, rather than natural establishment,
25 mostly from local genetic material. Comparing genetic diversity and structure in refugia versus
26 recent populations did not support the existence of founder effects: heterozygosity (He =
27 0.667) and allelic richness (Ar = 4.298) were similar, and F_{ST} was low (0.031 overall). Still,
28 significant spatial evidence of colonization was detected, with He increasing along the
29 expansion front, while genetic differentiation from the entire pool (β_{WT}) decreased. Isolation
30 by distance was found in refugia but not in recently expanding populations.

31 Our study indicates that beech capacities for colonization and gene flow were sufficient to
32 preserve genetic diversity despite recent forest contraction and expansion. Because beech has
33 long distance pollen and seed dispersal, these results illustrate a 'best case scenario' for the
34 maintenance of high genetic diversity and adaptive potential under climate-change related
35 range change.

36

37

Introduction

Most species experience range expansion or contraction at some point in their history, which can have long-term consequences for population's genetic diversity (Excoffier *et al.* 2009). Range change studies are of increasing interest, as we attempt to predict how climate change may induce shifts in species distributions, for example migration to higher latitude or elevation (Lenoir *et al.* 2008), or expansion into previously inaccessible areas (Pluess 2011). Not only do we need to predict the likelihood and direction of range shifts, we also need to understand their genetic consequences, particularly on short time scales (a few generations), as this will facilitate the development of practical guidance for conservation and sustainable management of genetic resources.

When a population is increasing in number and is also expanding spatially, several factors determine genetic diversity and structure along the colonisation front (Excoffier *et al.* 2009). When only a few individuals contribute to the advance of the colonisation wave (i.e. pulled wave following Roques et al. 2012), we expect strong repeated founder effects, increasing frequencies of few neutral mutations ("gene surfing"), loss of genetic diversity and strong spatial genetic structure (SGS) along the expansion axis (Edmonds *et al.* 2004; Hallatschek and Nelson 2008). Alternatively, colonisation driven by many dispersing individuals (i.e. pushed wave) should result in higher genetic diversity at the colonisation front and a weaker SGS, especially if these many individuals originate from a variety of locations, demes or patches. The effective number of founders depends on long-distance dispersal and on a variety of demographic processes and life-history traits determining whether the colonisation wave is pushed or pulled (Hallatschek and Nelson 2008). The balance between dispersal distance relative to inter-patch distances, reproduction rate and carrying capacity is a first determinant of the pulled/pushed nature of a front (Klopfstein et al. 2006). The precise shape of the dispersal kernel has been shown to affect the effective number of founders in a complex way: roughly, fatter-tailed kernels (i.e. those that decrease more slowly at long distances) promote diversity in the front (Fayard *et al.* 2009) although this pattern is not completely monotonic (Paulose and Hallatschek, 2020) nor scale-free (Bialozyt *et al.*, 2006). Demographic processes, such as Allee effects (Roques *et al.* 2012), or life-history traits, such as a long juvenile stage (Austerlitz *et al.* 2000), also increase effective population size and limit the erosion of genetic diversity along the colonization front (i.e., pushed colonization waves). By extension, absence of Allee effect or short lifespan can favour the contribution of a few individuals to the colonization front (e.g., the further forward individuals, or first mature individuals), and thereby the rapid erosion of genetic diversity (i.e., pulled colonization waves).

Temperate forest trees are compelling study systems for investigating the relationship between range shifts and population genetic diversity, because experimental studies in tree species generally poorly support the expectations of classical models based on a drift-mutation model in the non-spatialized context of an isolated population or metapopulation. Classical models predict that population size reduction would be associated with decreased allelic richness and heterozygosity at neutral loci (Nei 1975). In contrast to these expectations, temperate forest trees retain high levels of within-population diversity despite their well-documented rapid post-glacial recolonization history during the last Quaternary (e.g. Petit *et al.* 2003; Hewitt 2004). Although decreasing trends of allelic richness along the postglacial expansion front have been reported in several tree species (e.g., Comps *et al.* 2001; Hoban *et al.* 2010), the founder events associated with postglacial range expansion have generally resulted in weak or no genetic drift. Studies of more recent and smaller scales natural

84   expansion also generally reveal only weak genetic drift associated with founder events, and
85   high levels of within-population diversity (Troupin *et al.* 2006; Born *et al.* 2008; Pluess 2011;
86   Shi and Chen 2012; Lesser *et al.* 2013; Elleouet and Aitken 2019). Similarly, recently
87   translocated tree populations generally combine a high level of differentiation for adaptive
88   traits, suggesting rapid genetic evolution, with a high level of within-population diversity,
89   indicating a limited impact of genetic drift and purifying selection (e.g. Lefevre *et al.*, 2004).
90   Hence, it is widely accepted that founder events can lead to genetic drift only in extremely
91   isolated tree populations, such as described by Ledig (2000), for *Pinus coulteri*, where highly
92   isolated populations are restricted to high elevations and separated by semiarid habitats
93   which severely limit gene flow.

94   The theoretical expectations and empirical results described above suggest that an optimal
95   strategy to detect the genetic signature of recent range shift in forest trees should combine
96   several indicators, including genetic diversity and differentiation, in a spatially explicit context,
97   including in particular spatial genetic structure (SGS). Although isolation by distance (IBD) and
98   SGS were first described by Wright (1942) in a stable metapopulation as the equilibrium
99   resulting from geographically restricted dispersal, ongoing processes that are not at
100 equilibrium can also be investigated by measuring the correlation between genetic divergence
101 and geographical distance (i.e. SGS). SGS can be investigated between individuals in a
102 continuous population (i.e., fine scale SGS) typically by individual (dis)similarities to estimate
103 genetic divergence, or between populations (i.e., inter-population SGS) typically by using F_{ST}
104 to estimate genetic divergence. The conceptual frameworks of SGS and IBD apply in similar
105 ways at these two scales of analysis (Rousset, 1997, 2000). Within a recently colonized
106 population, fine scale SGS among individuals is expected to start from no SGS, especially if the
107 different founders are distributed randomly at arrival, and then to increase with time,
108 especially if seed dispersal is spatially limited. It thereby provides a temporal proxy of the
109 establishment date (Slatkin 1993; Troupin *et al.* 2006). Successive founder events along a
110 colonization axis can also lead to significant SGS among populations, thus mimicking the
111 signature of IBD, particularly under stepwise expansion (de Lafontaine *et al.* 2013). In that
112 case, though, a decrease of genetic diversity occurs jointly with the establishment of the SGS,
113 unlike in the equilibrium IBD pattern. The comparison of inter-population SGS in refugia vs
114 expanding areas has seldom been investigated. One of the rare studies (de Lafontaine *et al.*
115 2013) found stronger genetic differentiation among populations in post-glacial refugia than in
116 recolonized areas, but regional SGS was lower within refugia than within recolonized areas.
117 By contrast, Pluess (2011) found significant fine-scale SGS in late successional sub-population
118 but no SGS in early successional sub-population.

119   Here, we investigate the genetic impact of range change in the tree species *Fagus*
120 *sylvatica* (European beech) on the slopes of Mont Ventoux, France. Across Europe in the 20th
121 century, large areas of agricultural land were abandoned and left to secondary succession
122 (Sluiter and De Jong 2007). In line with this pattern, the beech forests on Mont Ventoux
123 contracted until the 19th century due to human activities, but have now recolonized areas of
124 both the north and south slopes. In a previous study (Lander *et al.* 2011), we used historical
125 records to locate most of the probable remnant populations of the massif. These beech
126 populations hence provide a valuable model system for studying the genetic impacts of recent
127 local population contraction and expansion, which has occurred for many plant species across
128 Europe. We also demonstrated significant demographic fluctuations across the area using a
129 combination of historical information and Approximate Bayesian Computation (ABC) analyses

130 of modern genetic data. However, these ABC genetic analyses did not account for the spatial
131 component of genetic structure.

132   In this study, we improved our spatial sampling and more deeply analysed the
133 georeferenced genotypes to address two main questions:

134   (Q1) Can genetic diversity and structure provide evidence of expanding populations'
135 origin (i.e. natural recolonization versus establishment through planting)? Evidence of beech
136 plantations established using both local and non-local seeds was found in historical records
137 (Lander *et al.* 2011). A prerequisite to investigate the relationship between range shifts and
138 population genetic diversity in our study system is to identify the planted populations, to avoid
139 possibly confounding effects of plantation. Indeed, planted populations are expected to be
140 differentiated from the others, particularly if non-local material was used. Their diversity could
141 be higher than neighbouring populations (due to mixing of seed lots). They should also
142 decrease the overall pattern of inter-population SGS, and show no or weak fine-scale SGS.

143   (Q2) Did the contraction-recolonization history reduce genetic diversity? Along the
144 expansion front, we expect genetic diversity to decrease with increasing distance to refuges,
145 with a potentially strong impact of the modalities of expansion: under stepwise expansion, we
146 expect IBD patterns at landscape scale and a regular decrease in diversity with increasing
147 distance to refuges. Alternatively, under frequent events of long-distance colonisation, we
148 expect no IBD patterns at landscape scale, and more erratic patterns of diversity with
149 increasing distance to refuges. Within refuges, we expect higher levels of diversity, and IBD
150 patterns at landscape scale.

Material and methods

Study species

153   European beech (*Fagus sylvatica* L., Fagaceae) is a common European diploid ($2n = 24$),
154 monoecious tree species which typically begins to reproduce after 40 to 50 years. Pollen is
155 wind-dispersed , and mating occurs almost exclusively through outcrossing, though selfing is
156 possible (Gauzere *et al.* 2013). Seeds are produced in irregular mast years, and dispersed
157 primarily by gravity, and then by various animals. Previous studies of beech on Mont Ventoux
158 found that average dispersal distances were low for both seeds (18 m) and pollen (52 m), but
159 both seed and pollen dispersal kernels were fat-tailed. The proportion of seeds/seedlings
160 finding no compatible parents within plot (with typical size of 1.6 ha) was non-negligible: 46%
161 for male parent and 11.6% for female parent on average (Gauzere *et al.* 2013; Bontemps *et
162 al.* 2017; Oddou-Muratorio *et al.* 2018).

Study site and sampling design

164   Mont Ventoux is located at the warm and dry southern margin of the European beech
165 distribution (Figure S1), and the climate is typical of low altitude mountains with
166 Mediterranean influences (weather station of Mont Serein, 1 445 m a.s.l., 1993–2006; mean
167 annual temperature of 6.8∘C and mean annual rainfall of 1300 mm). Nevertheless, the strong
168 altitudinal variation over this large mountain, culminating at 1912 m, offers a wide array of
169 bioclimatic conditions. The forests on the mountain have changed species composition and
170 contracted and expanded many times due to climate cycles, however for the last five thousand
171 years the higher elevations have been dominated by European beech and European silver fir
172 (Barbero and Quezel 1987). Human activities caused extensive deforestation of the mountain
173 from the 13th to the 19th centuries, and in response a reforestation program was launched in
174 1861 (Jean 2008). In a previous study (Lander *et al.* 2011), we used historical data to identify

4

175   a modern population minimum in 1845, and we were able to distinguish areas of beech forest
176   which have been present for the last 200 years (four refugia) from two areas that appear to
177   be the result of recent forest expansion (Figure 1). That analysis found that the three regions
178   under study (North-West, North-East, South) were genetically distinct, with two remnant
179   areas (North-West, North-East) and one area of recent expansion (South). However, the areas
180   of recent expansion, as well as the remnant population on the South ridge were under-
181   represented in that previous study, which used 1932 trees in 51 plots. Moreover, the previous
182   study did not explicitly account for historical records showing that beech was planted on the
183   South edge of Mt Ventoux (Lander et al. 2011), although much less intensively than other
184   species (e.g., *Pinus nigra, Pinus sylvestris*). These beech plantations reportedly used either
185   local seeds (raised in local, non-permanent, "flying" nurseries) or non-local seed delivered by
186   commercial nurseries.

187
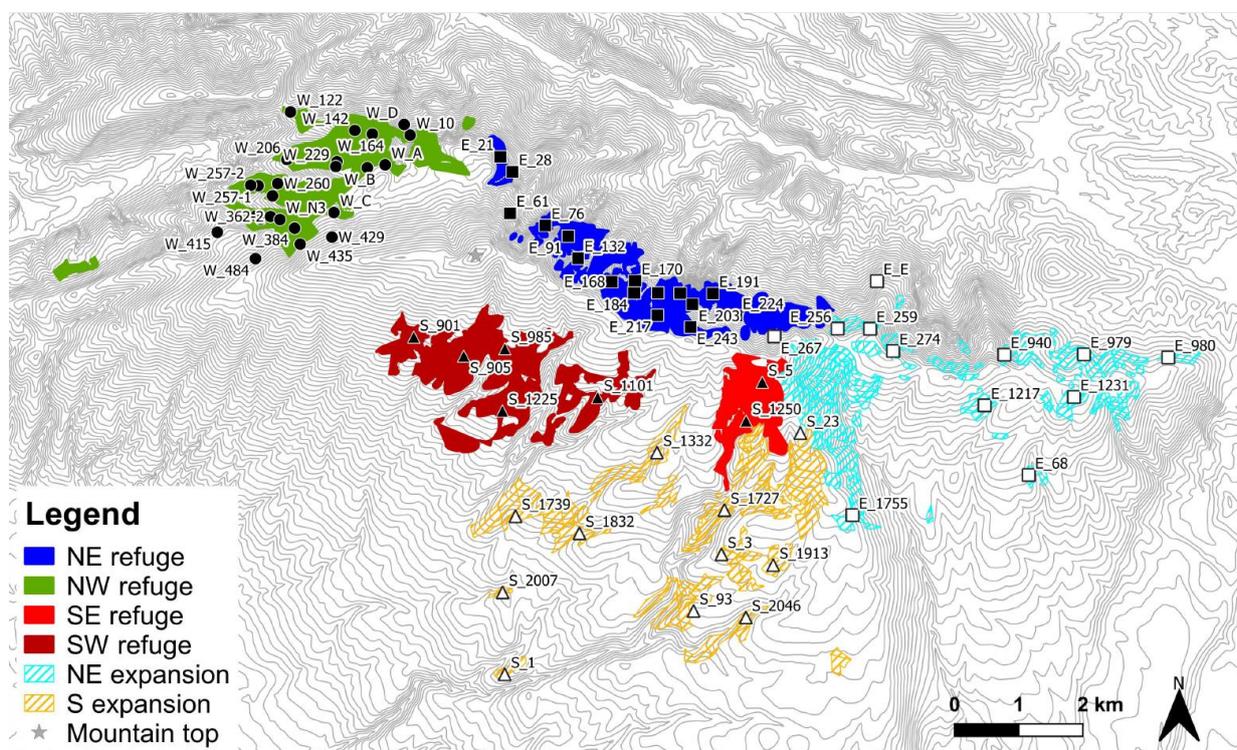
188
189

190   **Figure 1: Spatial distribution of the 71 studied plots overlaid on the topographical map of**
191   **Mont Ventoux.** Historical records allowed us to distinguish four refugia area (filled
192   polygons), where beech has been present for the last 200 years (NW_REF and NE_REF on the
193   Northern slope; SW-REF and SE-REF on the Southern slope) from two area of recent
194   expansion (hatched polygons; S-EXP and NE-EXP). The spatial delimitation between the
195   NW_REF and NE_REF was chosen to be a large terrace, while that between S-EXP and NE-
196   EXP was chosen to be the major crest line. SW-REF and SE-REF were aggregated in most
197   analyses as only few plots could be sampled in these areas. Plots are mapped with shape
198   indicating the region (filled dots= NW_REF; filled squares= NE_REF; empty squares= NE_EXP;
199   filled triangles= S_REF; empty triangles= S_EXP).

200
201

202   For this new study, we sampled 600 additional trees in 20 new plots, providing a total of
203   2532 adult trees in 71 plots covering five different regions of Mont Ventoux (Figure 1). We
204   retained 2042 trees for analyses (see Appendix A1 for selection), distributed as follows: (1)
205   748 trees in 25 plots in the north-western refuge (NW_REF), which is a tight mixture of
206   remnant and more recent communal forest under traditional management; (2) 464 trees in
207   16 plots in the north-eastern refuge (NE_REF) which is included in a Biosphere Reserve and is
208   unmanaged; (3) 316 trees in 12 plots in the far eastern region, recently recolonized by
209   European beech expanding out of the refuge areas (NE_EXP); (4) 208 beech in 7 plots at high
210   elevation on the south face of Mont Ventoux, a region identified as refuge forest (S_REF); and
211   (5) 306 beech in 11 plots at low elevation on the south face of Mont Ventoux, a region recently
212   recolonized by European beech (S_EXP).

213   Within each plot, 28.8 adult trees on average (up to a maximum of 40 individuals) were
214   sampled in an area of ~50 m radius so that all trees were separated by at least 3 meters. All
215   trees had a circumference at breast height > 160 mm. Trees were chosen so that half of them
216   had the largest circumference in the plot ("Old" trees, average mean/maximal circumference=
217   958/1495 mm) and the other half had the smallest circumference ("Young" trees, average
218   mean/minimal circumference= 444/309 mm). Geographical coordinates were recorded for all
219   sampled trees and a map of the study area was developed in ArcMap 10.4 (ESRI) using the
220   geographical coordinates of the trees, a map of current forest ownership and forest cover
221   (Direction Territoriale Méditerranée 2001), and a topographical map (IGN-PACA 2002). Plots'
222   altitudes were estimated in ArcMap. Finally, the maximal age of a tree within each plot was
223   estimated based on the tree ring profile of the largest possible tree (average maximal age =
224   155). Detailed information per region and plot is available in Table 1, Table S1, Figure S2 and
225   Appendix A1.

Genotyping and basic statistics

227   All individuals were genotyped using 13 microsatellite markers, one of which was excluded
228   due to high frequency of null alleles. Detailed information on genotyping and quality of the
229   marker set is available in Supplementary Appendix A1.

230   Statistical analyses were conducted using R 3.6.2 (R Core Team 2019) unless otherwise
231   indicated. We considered several statistics to describe population diversity at plot level. We
232   first used the package 'diveRsity' (Keenan *et al.* 2013) to compute the allelic richness (*Ar*) and
233   the expected heterozygosity (*He*). We also computed *He* and *Ar* values for each cohort within
234   plot, and derived the difference in *He* and *Ar* between old and young individuals (respectively
235   *difHe* and *difAr*). We used the package 'hierfstat' to compute Wright's inbreeding coefficient
236   (F_{IS}), pairwise F_{ST} among plots following Weir and Cockerham (1984), and β_{WT}, a plot-specific
237   index of genetic differentiation relative to the entire pool (Weir and Goudet 2017). Tests for
238   departures from Hardy–Weinberg equilibrium (HWE) and linkage equilibrium were conducted
239   using Fstat 2.9 (Goudet 2001). We used the package 'hierfstat' to estimate the components of
240   variance in allelic frequencies among regions, among plots within regions, and among
241   individuals within plots, and derived the associated F-statistics (F_{CT}, F_{SC}, F_{IS}, and F_{ST}).

Bayesian inference of population structure

243   The genetic structure was investigated using two different tools based on Bayesian
244   clustering algorithms. These methods have different prior distributions and assumptions, and
245   we used them simultaneously to evaluate the robustness of the genetic clusters. We
246   hypothesised that the number of possible clusters (K-values) was unlikely to be greater than
247   4 in our case, considering the continuous, rather than patchy, distribution of beech on Mont

248   Ventoux, and the presence of only four beech refugia on the mountain during the modern
249   population minimum. However, because of the possible planting using non-local seeds, we
250   investigated a wider range of K-values.

251   Bayesian clustering of the genetic data was first performed using STRUCTURE 2.3.3
252   (Pritchard *et al.* 2000), with K varying between 1 and 13, and 10 runs for each K value.
253   Parameters were 2,500 burn-in periods and 10,000 Markov Chain Monte Carlo repetitions
254   after burn-in, with allele frequencies correlated among populations and an admixture model
255   of population structure. To account for non-independence between two genotypes from the
256   same population, we used population identifiers as prior information to assist clustering. The
257   ΔK statistics allowed us to evaluate the change in likelihood and select the optimal K value
258   (Evanno *et al.* 2005). For the selected K-value, we averaged over 10 runs the proportion of
259   each cluster in each sampling plot and the individual probabilities of belonging to each cluster
260   using CLUMPAK (Kopelman *et al.* 2015).

261   TESS 2.3 (Chen *et al.* 2007) was also used to estimate the number of genetic clusters
262   present in the data by incorporating the geographical coordinates of individuals as prior
263   information to detect discontinuities in allele frequencies. We used an admixture model and
264   a burn-in of 10,000 iterations followed by 50,000 iterations from which estimates were
265   obtained. We performed 200 independent runs for each K value (K = 2 - 6), with spatial
266   interaction influence ψ at 0.6 (default value). The optimal K value was determined by the
267   lowest value of the deviance information criterion (DIC). The 200 runs for the best K were
268   averaged using CLUMPP (Jakobsson and Rosenberg 2007).

269   Spatial outputs of both STRUCTURE and TESS were visualised using the R script
270   'krigAdmixProportions' distributed with the TESS program. This script uses a kriging approach
271   to interpolate a surface model based on scattered, spatially explicit, data points. This consists
272   in using the proportions of the different clusters at each of the 71 sampled locations to
273   estimate the probabilities to belong to the different clusters at all locations of the landscape.

Spatial variation in diversity and connectivity

275   We visualized spatial patterns in genetic diversity and geneflow rates using the program
276   EEMS (Estimated Effective Migration Surfaces; Petkova *et al.* 2015). This method uses
277   sampling localities and pairwise dissimilarity matrices calculated from microsatellite data to
278   identify regions where genetic similarity decays more quickly than expected under IBD. A user-
279   selected number of demes determines the geographic grid size and resulting set of migration
280   routes, and the expected dissimilarity between two samples is approximated using resistance
281   distance. These estimates are calculated without the need to include environmental variables
282   or topographic information and are subsequently interpolated across the geographic space to
283   provide a visual summary of observed genetic dissimilarities, including regions with higher and
284   lower gene flow than expected. We tested three numbers of demes (400, 600, 800) using the
285   runeems_sats version of EEMS. For each deme size, we ran three independent analyses, with
286   a burn-in of 500,000 and MCMC length of 3,000,000. The results were combined across the
287   three independent analyses, and convergence of runs was assessed using the 'reemsplots' R
288   package. Using this package, we generated surfaces of effective diversity (q) and effective
289   migration rates (m) combining the nine independent runs for the three deme size.

Isolation by distance

291   We first estimated SGS among sampling plots across the whole study area and tested
292   whether geographic distances significantly shaped the patterns of genetic differentiation,
293   estimated by F_{ST}, among plots using the software SpaGeDi 1.4c (Hardy and Vekemans 2002).

294   To test IBD, the F_{ST} values were regressed on $\ln(d_{ij})$, where d_{ij} is 3D spatial distance accounting
295   for elevation between plots i and j, calculated using the 3D Analyst Tools in ArcMap 10.4. Then,
296   we tested the regression slope (b_{logFST}) using 5,000 permutations of genotypes among
297   population's positions. These analyses were run globally over the 71 plots, and within each
298   region. SGS estimates can be sensitive to outlier plots showing higher or lower differentiation
299   for the others plots (de Lafontaine *et al.* 2013). To account for possible biases due to planted
300   forest material, we ran conservative SGS analyses within each historical group (see Results).

301   We also estimated fine-scale SGS within each plot with SpaGeDi. Within each plot, genetic
302   relatedness between all pairs of individuals i and j was estimated using the kinship coefficient
303   F_{ij} (Loiselle *et al.* 1995). To estimate SGS, the F_{ij} values were regressed on $\ln(d_{ij})$, where d_{ij} is
304   the 2D spatial distance between individuals i and j. We tested the significance of SGS
305   (regression slope, b_{logFij}) using 5,000 permutations of genotypes among individual positions.
306   Following Vekemans and Hardy (2004) , the SGS intensity was quantified by $Sp= b_{logFij}/(F_1-1)$,
307   where F_1 is the average kinship coefficient between individuals of the first distance class ($<$
308   10m). Sp primarily depends upon the rate of decrease of pairwise kinship coefficients between
309   individuals with the logarithm of the distance, and is scaled by the average level of relatedness
310   between individuals, which allows inter-population comparison.

311   ***Impact of recolonization history on genetic diversity***

312   We tested the hypothesis that the distance and steepness of up-slope and down-slope
313   travel between each study population and the 'core area' of each of the refugia (NW_REF,
314   NE_REF, SW_REF and SE_REF) affects genetic diversity. As the refugia are irregularly shaped,
315   the 'core areas' were defined as the medial axes of the refugia polygons ('Thin', ArcMap 10.4).
316   Following Zafar (2011), the 2D line from the centroid of each sample plot to the nearest point
317   on the medial axis of each of the four refugia was drawn using Analysis Tools (ArcMap 10.4 ;
318   Supplementary Table S1). The 2D lines were then converted to 3D lines based on two aspect
319   rasters, one weighted for travel north to south, and one for travel south to north (3D Analyst,
320   ArcMap 10.4), providing data on the linear distance and travel up and down a seed or seed
321   vector would have had to travel on Mt Ventoux's surface between each refuge and each study
322   population.

323   Similar to Hoban et al. (2010), we used ANCOVA to investigate how the distance to refugia
324   shaped genetic variation at plot level, described by seven summary statistics (*He*, *Ar*, F_{IS}, *Sp*,
325   β_{WT}, *difHe*, and *difAr*), considered as response variables. For each summary statistic, we
326   considered the following models:

327   Response variable = wdistNE + wdistNW + wdistSE + wdistSW + HistGroup (model 1)

328   Response variable = (wdistNE + wdistNW + wdistSE + wdistSW) × HistGroup (model 2)

329   where all distances are quantitative variables, and HistGroup is a categorical variable
330   integrating the recent history of plots as supported by genetic clustering analyses (i.e., refuge,
331   expansion area, and planted plots; and see result section). The best linear model was selected
332   based on AIC with the stepwise algorithm implemented in the step procedure of the 'stats'
333   package.

334   **Results**

335   ***Genetic variation within and among sampled plots***

336   Genetic diversity estimates are summarised in Table 1 and detailed in Table S2. In total,
337   154 alleles were scored at the 12 loci, corresponding to an average of 12.8 alleles per locus
338   (range = 5–23). Mean allelic richness per population ranged between 3.1 and 5.11 (mean *Ar* =

339   4.3), while observed and expected heterozygosities per population ranged from 0.600 and
340   0.610 to 0.760 and 0.720 (mean Ho = 0.683 and mean He= 0.667). Ar, He, and Ho did not differ
341   significantly among the five studied regions.

342   Ten populations showed significant departure from HWE, four of which displayed
343   heterozygosity deficit and another six showed heterozygosity excess. F_{IS}-values ranged from
344   −0.096 to 0.108 (mean F_{IS} = 0), and did not differ between regions. Genetic differentiation of
345   each plot from the entire population ranged from -0.027 to 0.11 (mean $β_{WT}$=0.03). The region
346   NW_REF showed a significantly lower $β_{WT}$-value, likely due to the high contribution of the 25
347   plots of this region to the entire genepool.

348   Estimation of hierarchical variance components showed that most of the genetic variation
349   lies among individuals within plots (Table S3): genetic differentiation among-plots within
350   regions (F_{SC} = 0.029), and among-regions (F_{CT}=0.002) were weak, although significant. The
351   overall genetic differentiation among plots was F_{ST}=0.031.

352   Among-plots pairwise F_{ST}-values ranged between 0 and 0.097 with a mean value of 0.031
353   (median = 0.029) (Figure S3). The highest observed differentiation values involved plots
354   E_1231, S_1913 and S_2007. The lowest observed differentiation values occurred between
355   plots of the South region

356   ***Spatially distinct genetic clustering***

357   The genetic clustering analyses found weak but significant genetic structure. Using
358   STRUCTURE, the method of Evanno (2005) selected K=3 and K=6 as the most-likely values of
359   the number of clusters (Figure S4a). Retaining K=3 as the first major peak in ΔK statistics,
360   spatial kriging of the Q-matrix suggests that the three clusters are spatially distinct (Figure 2,
361   Figure S4b-e). Cluster C1 is predominant in North-West (22 of 25 plots) and North-East (13 of
362   16 plots) refuges. Cluster C2 is predominant in South refuge (4 of 7 plots), and present in all
363   other regions. Cluster C3 groups plots S_1727, S_1913, and S_3 (South Expansion), plots
364   E_1231 and E_1755 (East expansion) and plot S_23 (on the southern boundary of the SE
365   refuge). Finally, the average allelic divergence (F_{ST}) between clusters C1 and C2 is ~1.3%, while
366   F_{ST} between clusters C3 and C1 (C2 respectively) is 2.3% (2.6% respectively). When using
367   STRUCTURE without prior information, the power of plots' assignation to clusters C1 and C2
368   decreased, while cluster C3 remained distinct (Figure S5)

369   In the TESS analysis the lowest DIC value was for K=6 (Figure S6). For K=3 (Figure 2b), TESS
370   clustering is fully consistent with STRUCTURE, as illustrated by the strong correlations
371   between the membership coefficients of plots to clusters estimated with TESS and STRUCTURE
372   (ρ=0.96 for C1 and C2 ; ρ=0.99 for C3, p-values < 0.001). Cluster C3 is also the most supported:
373   it appears when the results from K=2 are graphed, and remains distinct up to K=6 (Figure S6).

374   In the following IBD and historical analyses, to test different expectations for expansion
375   areas vs refuges, we accounted for the detected genetic clusters and classified the 71 plots in
376   three historical groups. The "refugia" group (REF) includes 48 plots from the refuge regions
377   (i.e. 25 NW_REF plots, 16 NE_REF plots and 7 S_REF plots), all assigned either to clusters C1
378   and C2. The "likely-planted" group ("PLANTED") includes four S-EXP plots and two NE-EXP plots
379   predominantly assigned to cluster C3 (six plots in total). The "natural expansion" group (EXP)
380   includes all of the remaining 17 plots of the expansion area (i.e. 10 NE-EXP and 7 S-EXP plots),
381   assigned to clusters C1 and C2 (Table S2).

382
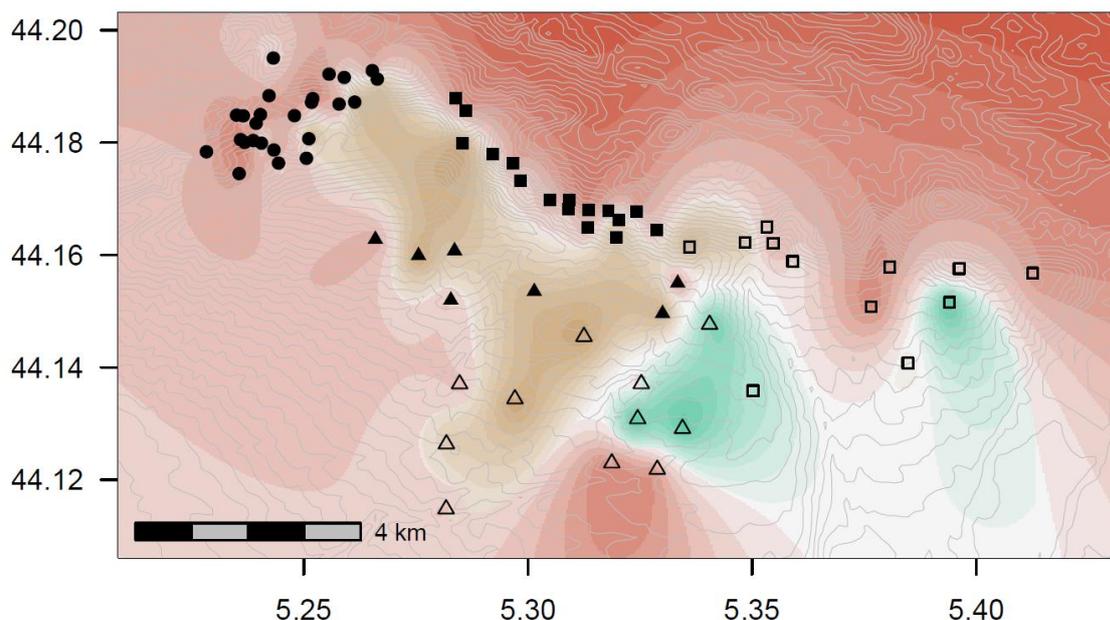
Figure 2: Spatial interpolates of the admixture coefficients estimated with STRUCTURE for K=3. The colours red coral, ochre and green correspond respectively to clusters C1, C2 and C3. The colour intensity indicates the probability to belong to the dominant cluster at a given position in space, based on spatial kriging of the individual q-matrix. Plots shape indicate the region (see legend of Fig 1). Grey lines represent topographic isoclines.

Spatial differences in geneflow and genetic diversity

EEMS spatial analyses highlight several barriers to migration resulting from either historical or contemporary patterns of gene flow (Figure 3, Figure S7). There is evidence for restricted migration around the mountaintop, and along the expansion paths towards the East and South. Spatial analyses of genetic diversity highlight four main regions of exceptionally high diversity, three of which are located along the expansion paths towards the East and South, and the last one in NW refuge. However, regions with lower-than-expected genetic diversity are also found along the expansion paths towards the East and South, resulting in a tight mosaic of diverse and homogenous areas in term of genetic composition.

Isolation by distance

SpaGeDi found a significant signal of IBD on genetic differentiation between the 71 plots (Table 2, Figure 4). Pairwise F_{ST} overall significantly increased with increasing 3D geographic distances accounting for elevation (b_{log3D}=0.003, p-value<0.001). However, this significant pattern of IBD is mainly driven by the 25 plots of the NW refuge (b_{log3D}=0.007, p-value<0.001), while no significant IBD patterns were detected in other regions. The signature of IBD remained significant between the 48 plots from the refugia group (REF), although weaker than that of the NW refuge ($b_{log3D-REF}$=0.0023 versus $b_{log3D-NW\_REF}$=0.007). No signature of IBD could be detected between the 17 plots from the expansion group (EXP), or between the six plots of the likely planted group (PLANTED), which may be due to weak testing power.

A significant signal of IBD on kinship coefficients among individuals within plot (i.e., fine-scale SGS) was detected in 37 of the 71 plots, corresponding to 17% of the NE_EXP, 27% of the S_EXP, 50% of the NE_REF, 72% of the NW_REF and 86% of the S_REF (Table S2). Although the prevalence of SGS was higher in refugia than in expansion areas (χ^2 test p-value= 0.003), the intensity of SGS, as depicted by Sp, did not significantly differ among regions.
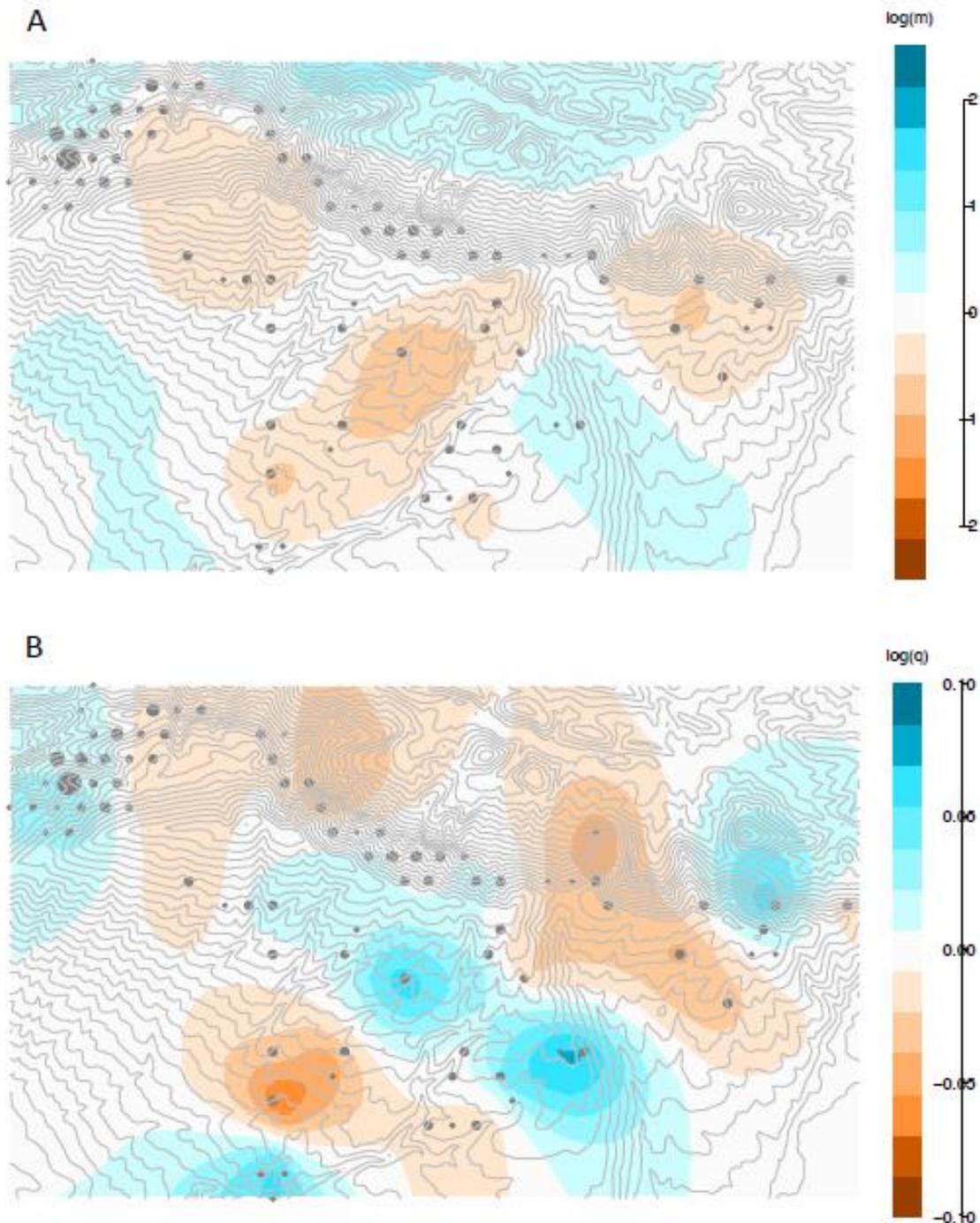
Figure 3: Contour maps representing the posterior mean of (a) effective migration surface and (b) effective diversity surface. In (a), blue colours represent areas of high migration, or dispersal corridors, whereas orange regions represent areas of low migration, or dispersal barriers. In (b), orange regions indicates areas of lower-than-expected genetic diversity, and blue colours represent higher levels of genetic diversity. The light grey dots illustrate the sampling design (bigger dots indicating a deme with more samples).

421
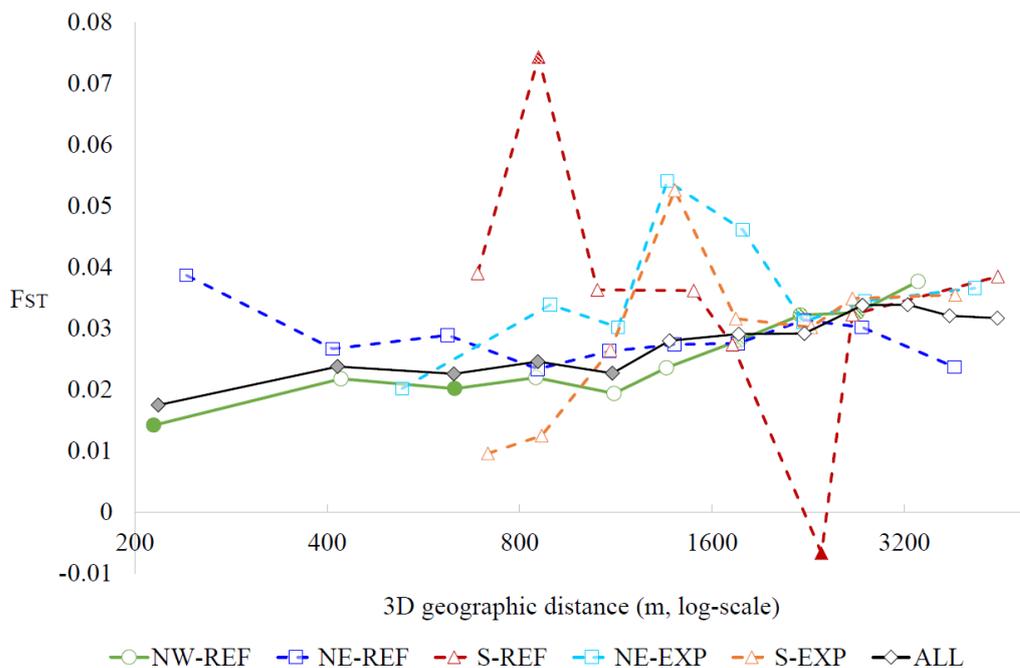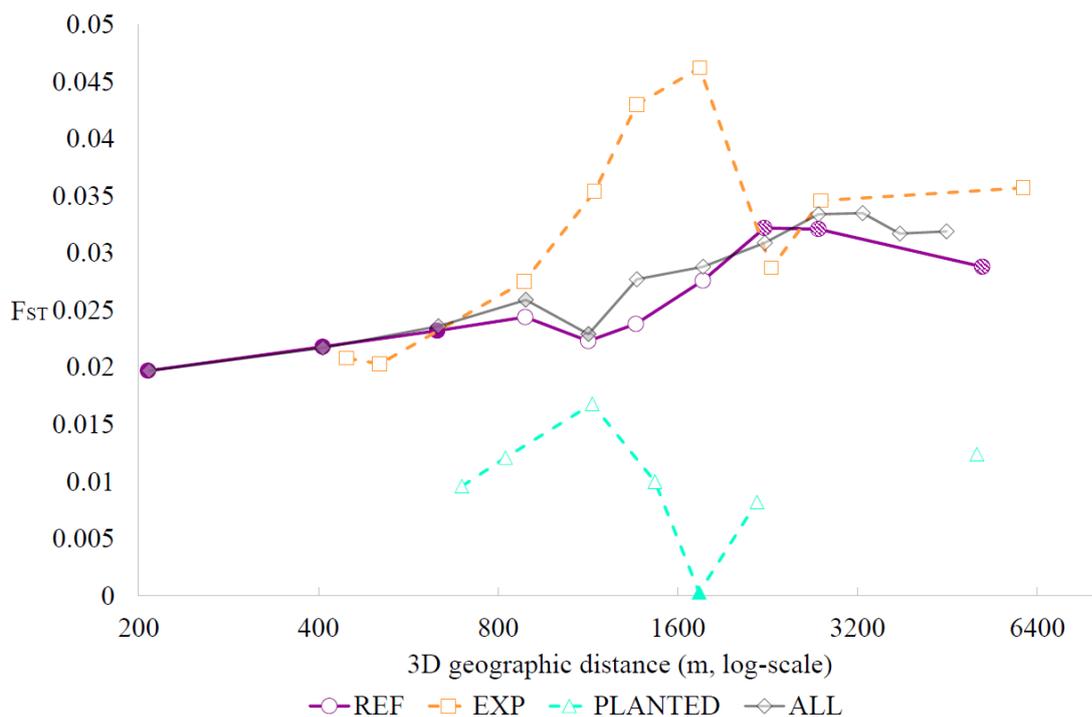
422
423 **Figure 4: Patterns of Spatial Genetic Structure (SGS) as depicted by average F_{ST}-values among**
424 **pairs of plots in function of the 3-dimensional geographic distance between them.** Solid
425 (respectively broken) lines indicates region where SGS is significant (respectively not
426 significant). Filled (respectively hatched) symbols represent average F_{ST} values lower
427 (respectively higher) than expected under complete spatial randomness. Shapes and colours
428 indicate (a) the region (green dots: NW_REF; blue squares: NE_REF; light blue squares:
429 NE_EXP; red triangles: S_REF; orange triangles: S_EXP); (b) the historical group (purple dots:
430 refugia; orange squares: natural expansion; aquamarine triangle: likely planted).

12

431 ***Impact of recolonization history on genetic diversity***
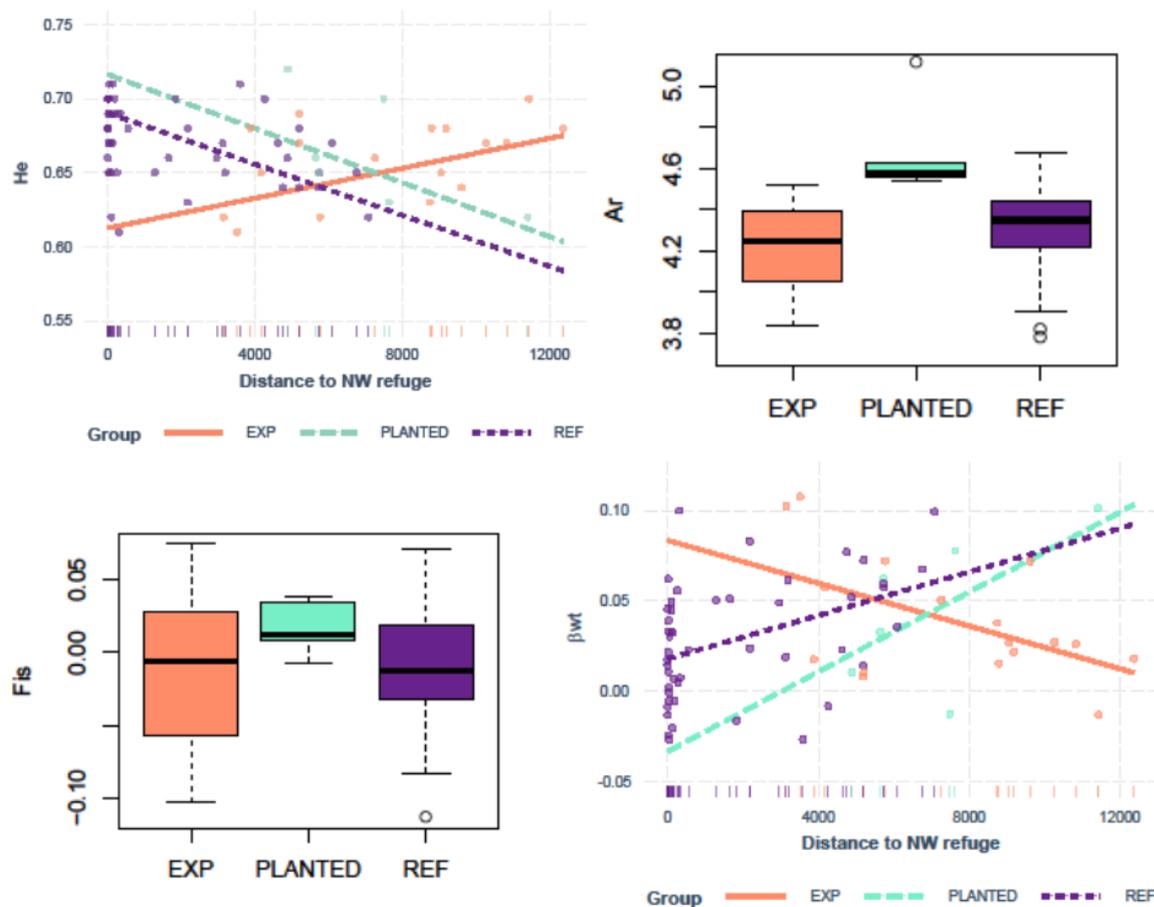
432
433 **Figure 5:** Impact of recolonization history on genetic diversity, as summarized by (a) expected
434 heterozygosity (*He*), (b) allelic richness (*Ar*), (c) differentiation from the entire gene pool (*βwt*)
435 and inbreeding coefficient (*F_{IS}*).

436   ANCOVA analyses showed that the impacts of distance to the refugia on genetic diversity
437 at plot level varied depending on the summary statistics considered (Table 3, Figure 5). We
438 found that Nei's genetic diversity (*He*) significantly decreased with increasing distance to the
439 NW refuge in the "REF" and "PLANTED" groups, while *He* significantly increased with
440 increasing distance to the NW refuge in the "EXP" group. We detected significantly higher
441 allelic richness (*Ar*) in the "PLANTED" group as compared to the "REF" and "EXP" groups, but
442 no significant effect of the distance to refuge on *Ar*.

443   Regarding Wright's inbreeding coefficient (*F_{IS}*), without accounting for distances to the
444 refuges, we found a significantly higher *F_{IS}*-level in the "PLANTED" group compared to the
445 "REF" and "EXP" groups (Figure 5d). Moreover, *F_{IS}* overall significantly increased with
446 increasing distance to the NE and NW refuge (which is partially confounded with the
447 "PLANTED" origin). The genetic differentiation relative to the entire pool ($β_{WT}$) increased
448 significantly with increasing distance to the NW refuge in the "REF" and "PLANTED" groups,
449 while $β_{WT}$ significantly decreased with increasing distance to the NW refuge in the "EXP" group
450 (Fig. 5c).

451   Fine-scale SGS was significant in 66% of the "REF" plots, in 29% of the "EXP" plots, and in
452 none of the "PLANTED" plots ($χ^2$ test p-value= 0.0008). However, the Sp statistics did not

453 reveal significant pattern variation in the intensity of SGS among groups, except a marginally
454 significant trend for lower SGS intensity in the "PLANTED" group as compared to the "REF"
455 and "EXP" groups (p-value=0.08). Finally, increasing distance to NE refuge and decreasing
456 distance to SW refuge were associated with decreasing difference in *Ar* between old and
457 young individuals (difAr), while no significant pattern was observed for difHe.

Discussion

459   This study aimed at investigating the genetic consequences of recent range shift, using a
460 spatially explicit theoretical framework and a valuable study system, that of the recent
461 expansion of beech on Mont Ventoux. We first discuss how the possible establishment of
462 some plots through planting may interfere with the signature of natural recolonization. We
463 then summarize how the observed patterns of genetic diversity and structure, including SGS,
464 support the theoretical expectations on the genetic consequences of spatial population
465 expansion. Finally, we discuss how these findings can be used to guide the management of
466 beech populations.

Genetic signatures of population origin

468   Beech has been a dominant species for 5000 years on Mont Ventoux, although its spatial
469 range has contracted and expanded several times, partly due to human activities in the last
470 1000 years (Lander *et al.* 2011).

471   Our results provide genetic evidence of tree planting events in six plots of the South and
472 North-East expansion areas. These plots all cluster together with Bayesian structure analyses,
473 and on the F_{ST}-based Neighbour-Joining tree (Figure S3). Moreover, they have a significantly
474 higher inbreeding coefficient (possibly due to Wahlund effect) and higher levels of allelic
475 richness, two features consistent with the mixing of seed lots from different origins. Only two
476 plots among the six identified are over-differentiated from all other plots (up to pairwise
477 F_{ST}=0.09, as compared to the mean F_{ST}=0.03), suggesting that the planted material was most
478 often of local origin. Interestingly, the EEMS analyses tend to associate these six plots with
479 areas combining low geneflow connectivity and high diversity, a paradox also consistent with
480 planting rather than natural establishment. Finally, none of these six plots shows significant
481 SGS at the individual level.

482   Thus six plots of the 23 sampled in the expansion area (26 %) appear to originate from
483 planting during the reforestation program launched in 1861. Hence, we cannot exclude the
484 possibility that planting by humans contributed to the spread rate previously estimated for
485 beech on Mont Ventoux (27–38 m/year, Lander et al. 2011). However, the remaining 17
486 expansion plots (73%) seem to have established naturally, confirming the high ability of beech
487 to spread and colonize new areas. To test theoretical hypothesis on the genetic consequences
488 of natural population expansion, we thus carefully distinguished the six probably planted plots
489 from the 17 plots probably originating from natural regeneration.

A weak but significant impact of contraction/expansion history on genetic diversity

491   Despite the short time elapsed to allow pollen to wipe out founder effects due to seed
492 dispersal, average levels of genetic diversity and structure did not show much evidence of the
493 impact of recent local range shifts. In particular, plots in expanding areas did not reveal the
494 classical signatures of strong genetic drift associated with founder events: their genetic
495 differentiation was overall low (although a few plots were over-differentiated from the
496 others), and their heterozygosity or allelic richness did not differ from the refuge areas. We
497 cannot exclude the possibility that the refugia themselves were subject to population size

498   contraction. However, consistent with previous studies in forest trees, our results confirm
499   that the long juvenile phase and the predominance of high pollen flow in wind-pollinated trees
500   strongly attenuate the genetic impacts of demographic changes (Shi and Chen 2012; Lesser *et*
501   *al.* 2013; Elleouet and Aitken 2019).

502   On the other hand, spatial patterns of genetic diversity and structure did reveal the
503   imprint of the expansion process, with a signature typical of genetic mixing between refuges.
504   Indeed, heterozygosity increased with increasing distance from the Northwest refuge, while
505   genetic differentiation from the entire gene pool (β_{WT}) decreased. These findings are
506   consistent with the increasing heterozygosity along post-glacial colonization axes at European
507   scale reported by Comps *et al.* (2001). Such an increase in *He* and decrease in β_{WT} is likely to
508   result from the combination of several processes and life history traits. First, a very fat-tailed
509   pollen dispersal curve as estimated for beech (Gauzere *et al.* 2013) can result in mixing of
510   propagules from different distant sources (Klein *et al.* 2006) and can thus promote high
511   genetic diversity at the colonization front (e.g. Fayard *et al.*, 2009; Paulose and Hallatschek,
512   2020). Second, the long juvenile phase strongly attenuates founder effects during colonisation
513   in a diffusive dispersal model (Austerlitz *et al.* 2000). Third, admixture at the intersection of
514   colonization fronts from different refugia can increase diversity (Comps *et al.* 2001), even
515   though this effect may be limited in our study case with weak genetic differentiation among
516   local refugia. Finally, selection pressures in the open environmental conditions at the
517   colonization front may support the maintenance of heterozygosity (Comps *et al.* 2001).

518   Another spatial signature of recent expansion was the absence of inter-population IBD,
519   whereas significant inter-population IBD was detected in some refuge areas. Moreover, fine-
520   scale SGS was much more prevalent in refuge as compared to expansion areas. This is
521   consistent with the theoretical work of Slatkin (1993) on IBD, which showed that a species
522   having restricted dispersal should exhibit SGS if enough time has elapsed after establishment,
523   assuming no initial structuring (Troupin *et al.* 2006). Testing this hypothesis in beech, De
524   Lafontaine et al. (2013) showed the reverse pattern, where recently colonized populations
525   displayed significant SGS whereas more ancient populations did not. However, they focused
526   on post-glacial recolonization, where several tens of generation probably allowed SGS to
527   develop, unlike our case of a recent colonization. Our study also highlights the idea that fine-
528   scale SGS does not systematically occur in refugia, for instance in the North East and South
529   refugia, which is also typical of a mix of founder origins under very fat-tailed kernels (Paulose
530   and Hallatschek, 2020).

531   Overall, this study confirms our previous results (Lander *et al.* 2011), and shows that the
532   genetic signatures of expansion predominate over those of genetic drift related to founder
533   events. The previous ABC approach selected the scenario were three populations (NorthWest,
534   NorthEast, South) expanded from a smaller ancestral population, rather than a scenario with
535   bottleneck supported by the historical data. This study additionally highlights specific spatial
536   signatures of the expansion process. Moreover, the South population was the most divergent
537   with the previous ABC approach. Here, we showed that two plots of this South population
538   (which included only five plots) originate from planting rather than natural establishment,
539   shedding light on this higher divergence.

540   ***Consequences for the management of beech populations.***

541   Knowledge of colonization and dispersal processes is crucial for management planning and
542   conservation efforts, particularly with a view to managing invasive species or genotypes
543   (Brandes *et al.* 2019), or to predicting species' response to climate change (Jump and Peñuelas

544   2005). However, empirical tests of how range expansion or contraction shape levels of
545   diversity within and among populations are limited by our ability to collect data on an
546   appropriate spatial and temporal scale, particularly in long-lived species. Many studies
547   investigate these issues at large spatial and temporal scales, which provide useful insights for
548   the conservation and management of genetic resources at species distribution scales.
549   However, local management also requires studies investigating recent and rapid events of
550   range change, similar to those expected under ongoing global and climate change.

551   Bioclimatic niche models predict a future reduction of beech at the rear edge of its range
552   over the next few decades (Cheaib *et al.*, 2012; Dyderski *et al.*, 2018). Moreover, forest areas
553   across Europe, and in the Mediterranean basin in particular, contracted and expanded many
554   times in the last centuries following socio-economic changes and their consequences on land-
555   use patterns (Sluiter and De Jong, 2007). Beech forests were no exception, and large areas of
556   beech forest have been cleared for agricultural production, and then recolonised following
557   field abandonment. This study found evidence of high adaptive potential of beech despite
558   such local range changes. First, our results demonstrate an overall increase in heterozygosity
559   and decrease in genetic differentiation along the expansion front, consistent with the genetic
560   mixing of founders from different origins. A major reservoir of genetic diversity was identified
561   within one refugia area (NorthWest). Moreover, the overall high levels of genetic diversity and
562   low genetic differentiation over the 15, 000 ha study area confirms the capacity for extensive
563   gene flow and the large effective population size previously estimated (Lander *et al.* 2011;
564   Oddou-Muratorio *et al.* 2018). These high levels of gene flow can be expected to foster rather
565   than hamper local adaptation, as shown by the microgeographic adaptation patterns along
566   the northern altitudinal gradient reported by Gauzere *et al.* (2020). Finally, this study also
567   demonstrates that the levels of allelic richness and heterozygosity are stable over time
568   between old and young cohorts (-0.6% and -0.1% respectively on average). Some reduction
569   up to 20% in *He* or *Ar* could be detected locally, but not associated with the expansion process,
570   or with differences in management practises (e.g., traditional management in NorthWest
571   refuge versus no management in the NorthEast refuge).

572   In conclusion, this study showed that range change on a local scale and over a small number
573   of generations did leave detectable genetic signatures, but overall did not increase genetic
574   differentiation, or reduce heterozygosity or allelic richness. These results paint a positive
575   picture of the potential for species to maintain genetic diversity and adaptive potential
576   through climate-change related range change. However, beech is both wind-pollinated and
577   biotically dispersed, and therefore expected to have long distance pollen and seed dispersal;
578   moreover, the seeds in this study would mainly have been dispersing downhill on the massif;
579   finally, the local persistence of several refugia and additional scattered beech trees even at
580   the population minimum is likely to have allowed continued geneflow between populations
581   to be maintained. For these different reasons, our results represent a 'best case scenario' for
582   the maintenance of high genetic diversity at the population perimeter during population
583   expansion, including during climate-change related range change. Other tree species with
584   more limited dispersal abilities, and/or more scattered distribution, such as *Pinus coulteri*
585   (Ledig, 2000), would represent an alternative 'worse case scenario'. The results therefore
586   should be conservatively interpreted as they relate to the development of management

587 recommendations for forest trees, and in particular for the high conservation priority forests
588 of the Mediterranean Basin (Madon and Médail 1997; Médail and Diadema 2009).

Competing Interests

597 The authors declare that they have no conflict of interest.

Data archiving

599 Data are available from the Dryad Digital Repository: doi:10.5061/dryad.nvx0k6dqt

References

601 Austerlitz F, Machon N, Gouyon P, Godelle B (2000). Effects of colonization processes on genetic
602   diversity: Differences between annual plants and tree species. *Genetics* **154**: 1309–1321.
603 Barbero M, Quezel P (1987). La végétation du Ventoux: diversité, stabilité et utilisation actuelles des
604   écosystèmes. *Etud Vauclus* **3**: 79–84.
605 Bialozyt R, Ziegenhagen B, Petit RJ (2006). Contrasting effects of long distance seed dispersal on genetic
606   diversity during range expansion. *J Evol Biol* **19**: 12–20.
607 Bontemps A, Davi H, Lefèvre F, Rozenberg P, Oddou-Muratorio S (2017). How do functional traits
608   syndromes covary with growth and reproductive performance in a water-stressed population of
609   Fagus sylvatica? *Oikos* **126**: 1472–1483.
610 Born C, Kjellberg F, Chevallier MH, Vignes H, Dikangadissi JT, Sanguié J, *et al.* (2008). Colonization
611   processes and the maintenance of genetic diversity: Insights from a pioneer rainforest tree,
612   Aucoumea klaineana. *Proc R Soc B Biol Sci* **275**: 2171–2179.
613 Brandes U, Furevik BB, Nielsen LR, Kjær ED, Rosef L, Fjellheim S (2019). Introduction history and
614   population genetics of intracontinental scotch broom (Cytisus scoparius) invasion (A Zhan, Ed.).
615   *Divers Distrib* **25**: 1773–1786.
616 Cheaib A, Badeau V, Boe J, Chuine I, Delire C, Dufrêne E, *et al.* (2012). Climate change impacts on tree
617   ranges: Model intercomparison facilitates understanding and quantification of uncertainty. *Ecol
618   Lett* **15**: 533–544.
619 Chen C, Durand E, Forbes F, François O (2007). Bayesian clustering algorithms ascertaining spatial
620   population structure: A new computer program and a comparison study. *Mol Ecol Notes* **7**: 747–
621   756.
622 Comps B, Gömöry D, Letouzey J, Thiébaut B, Petit RJ (2001). Diverging trends between heterozygosity
623   and allelic richness during postglacial colonization in the European beech. *Genetics* **157**: 389–397.
624 Dyderski MK, Paź S, Frelich LE, Jagodziński AM (2018). How much does climate change threaten
625   European forest tree species distributions? *Glob Chang Biol* **24**: 1150–1163.
626 Edmonds CA, Lillie AS, Cavalli-Sforza LL (2004). Mutations arising in the wave front of an expanding
627   population. *Proc Natl Acad Sci* **101**: 975–979.
628 Elleouet JS, Aitken SN (2019). Long-distance pollen dispersal during recent colonization favors a rapid
629   but partial recovery of genetic diversity in Picea sitchensis. *New Phytol* **222**: 1088–1100.
630 Evanno G, Regnaut S, Goudet J (2005). Detecting the number of clusters of individuals using the
631   software STRUCTURE: A simulation study. *Mol Ecol* **14**: 2611–2620.
632 Excoffier L, Foll M, Rémy J. Petit (2009). Genetic Consequences of Range Expansions. *Annu Rev Ecol*

633   *Evol Syst* **40**: 481–501.

634   Fayard J, Klein EK, Lefèvre F (2009). Long distance dispersal and the fate of a gene from the colonization
635   front. *J Evol Biol* **22**: 2171–2182.

636   Gauzere J, Klein EK, Brendel O, Davi H, Oddou-Muratorio. S (2020). Microgeographic adaptation and
637   the effect of pollen flow on the adaptive potential of a temperate tree species. *New Phytol*.

638   Gauzere J, Klein EK, Oddou-Muratorio S (2013). Ecological determinants of mating system within and
639   between three Fagus sylvatica populations along an elevational gradient. *Mol Ecol* **22**: 5001–
640   5015.

641   Goudet J (2001). FSTAT, a program to estimate and test gene diversities and fixation indices (version
642   2.9.3).

643   Hallatschek O, Nelson DR (2008). Gene surfing in expanding populations. *Theor Popul Biol* **73**: 158–170.

644   Hardy OJ, Vekemans X (2002). Spagedi : a Versatile Computer Program To Analyse Spatial. *Mol Ecol*
645   *Notes* **2**: 618–620.

646   Hewitt GM (2004). Genetic consequences of climatic oscillations in the Quaternary. *Philos Trans R Soc*
647   *B Biol Sci* **359**: 183–195.

648   Hoban SM, Borkowski DS, Brosi SL, McCleary TS, Thompson LM, McLachlan JS, *et al.* (2010). Range-
649   wide distribution of genetic diversity in the North American tree Juglans cinerea: A product of
650   range shifts, not ecological marginality or recent population decline. *Mol Ecol* **19**: 4876–4891.

651   Jakobsson M, Rosenberg NA (2007). CLUMPP: A cluster matching and permutation program for dealing
652   with label switching and multimodality in analysis of population structure. *Bioinformatics* **23**:
653   1801–1806.

654   Jean H (2008). Reboisement du Ventoux et restauration des terrains en montagne dans la vallée du
655   Toulourenc. *Les carnets du Vent* **61**: 72–75.

656   Jump AS, Peñuelas J (2005). Running to stand still: Adaptation and the response of plants to rapid
657   climate change. *Ecol Lett* **8**: 1010–1020.

658   Keenan K, Mcginnity P, Cross TF, Crozier WW, Prodöhl PA (2013). DiveRsity: An R package for the
659   estimation and exploration of population genetics parameters and their associated errors.
660   *Methods Ecol Evol* **4**: 782–788.

661   Klein EK, Lavigne C, Gouyon PH (2006). Mixing of propagules from discrete sources at long distance:
662   Comparing a dispersal tail to an exponential. *BMC Ecol* **6**.

663   Kopelman NM, Mayzel J, Jakobsson M, Rosenberg NA, Mayrose I (2015). Clumpak: A program for
664   identifying clustering modes and packaging population structure inferences across K. *Mol Ecol*
665   *Resour* **15**: 1179–1191.

666   de Lafontaine G, Ducousso A, Lefèvre S, Magnanou E, Petit RJ (2013). Stronger spatial genetic structure
667   in recolonized areas than in refugia in the European beech. *Mol Ecol* **22**: 4397–4412.

668   Lander TA, Oddou-Muratorio S, Prouillet-Leplat H, Klein EK (2011). Reconstruction of a beech
669   population bottleneck using archival demographic information and Bayesian analysis of genetic
670   data. *Mol Ecol* **20**: 5182–5196.

671   Ledig FT (2000). Founder effects and the genetic structure of Coulter pine. *J Hered* **91**: 307–315.

672   Lefevre F, Fady B, Ghosn D, Bariteau M (2004). Impact of founder population , drift and selection on
673   the genetic diversity of a recently translocated tree population. : 542–550.

674   Lenoir J, Gégout JC, Marquet PA, De Ruffray P, Brisse H (2008). A significant upward shift in plant
675   species optimum elevation during the 20th century. *Science (80-)* **320**: 1768–1771.

676   Lesser MR, Parchman TL, Jackson ST (2013). Development of genetic diversity, differentiation and
677   structure over 500 years in four ponderosa pine populations. *Mol Ecol* **22**: 2640–2652.

678   Loiselle BA, Sork VL, Nason J, Graham C (1995). Spatial genetic structure of a tropical understory shrub,
679   Psychotria officinalis (Rubiaceae). *Am J Bot* **82**: 1420–1425.

680   Madon O, Médail F (1997). The ecological significance of annuals on a Mediterranean grassland (Mt
681   Ventoux, France). *Plant Ecol* **129**: 189–199.

682   Médail F, Diadema K (2009). Glacial refugia influence plant diversity patterns in the Mediterranean
683   Basin. *J Biogeogr* **36**: 1333–1345.

684 Nei M (1975). The bottleneck effect and genetic variability in populations. *Evolution (N Y)* **29**: 1–10.

685 Oddou-Muratorio S, Gauzere J, Bontemps A, Rey JF, Klein EK (2018). Tree, sex and size: Ecological
686   determinants of male vs. female fecundity in three Fagus sylvatica stands. *Mol Ecol* **27**: 3131–
687   3145.

688 Paulose J, Hallatschek O (2019). From sectors to speckles : The impact of long range migration on gene
689   surfing. *bioRxiv*: 1–17.

690 Paulose J, Hallatschek O (2020). The impact of long-range dispersal on gene surfing. *Proc Natl Acad Sci
691   U S A* **117**: 7584–7593.

692 Petit RJ, Aguinagalde I, De Beaulieu JL, Bittkau C, Brewer S, Cheddadi R, *et al.* (2003). Glacial refugia:
693   Hotspots but not melting pots of genetic diversity. *Science (80-)* **300**: 1563–1565.

694 Petkova D, Novembre J, Stephens M (2015). Visualizing spatial population structure with estimated
695   effective migration surfaces. *Nat Genet* **48**: 94–100.

696 Pluess AR (2011). Pursuing glacier retreat: Genetic structure of a rapidly expanding Larix decidua
697   population. *Mol Ecol* **20**: 473–485.

698 Pritchard JK, Stephens M, Donnelly P (2000). Inference of Population Structure Using Multilocus
699   Genotype Data. *Genetics* **155**: 945–959.

700 Roques L, Garnier J, Hamel F, Klein EK (2012). Allee effect promotes diversity in traveling waves of
701   colonization. *Proc Natl Acad Sci U S A* **109**: 8828–8833.

702 Rousset F (1997). Genetic Differentiation and Estimation of Gene Flow from FStatistics Under Isolation
703   by Distance. *Genetics* **145**: 1219–1228.

704 Rousset (2000). Genetic differentiation between individuals. *J Evol Biol* **13**: 58–62.

705 Shi MM, Chen XY (2012). Leading-edge populations do not show low genetic diversity or high
706   differentiation in a wind-pollinated tree. *Popul Ecol* **54**: 591–600.

707 Slatkin M (1993). Isolation by distance in equilibrium and non-equilibrium populations. *Evolution (N Y)*
708   **47**: 264–279.

709 Sluiter R, De Jong SM (2007). Spatial patterns of Mediterranean land abandonment and related land
710   cover transitions. *Landsc Ecol* **22**: 559–576.

711 Team R Core (2013). R: A language and environment for statistical computing. R Foundation for
712   Statistical Computing, Vienna, Austria.

713 Troupin D, Nathan R, Vendramin GG (2006). Analysis of spatial genetic structure in an expanding Pinus
714   halepensis population reveals development of fine-scale genetic clustering over time. *Mol Ecol*
715   **15**: 3617–3630.

716 Vekemans X, Hardy OJ (2004). New insights from fine-scale spatial genetic structure analyses in plant
717   populations. *Mol Ecol* **13**: 921–935.

718 Weir BS, Cockerham CC (1984). Estimating F-statistics for the analysis of population structure.
719   *Evolution (N Y)* **38**: 1358–1370.

720 Weir BS, Goudet J (2017). A Unified Characterization of Population Structure and Relatedness. *Genetics*
721   **206**: 2085–2103.

722 Wright S (1942). Isolation by distance. *Genetics* **80**: 1514–1520.

723 Zafar SD (2011). Determination of an Optimal Bike Path Based on Slope & Distance Data. *Web site*: 17.

724

725