Genetic variation and signatures of natural selection in populations of European beech (*Fagus sylvatica* L.) along precipitation gradients

Laura Cuervo-Alarcon[1], Matthias Arend[2,3], Markus Müller[1], Christoph Sperisen[2], Reiner Finkeldey[4], Konstantin V. Krutovsky[1,5,6,7*]

[1]Department of Forest Genetics and Forest Tree Breeding, Georg-August University of Göttingen, Buesgenweg 2, 37077 Göttingen, Germany

[2]Swiss Federal Institute for Forest, Snow and Landscape Research WSL, Zürcherstrasse 111, 8903 Birmensdorf, Switzerland

[3]Department of Environmental Sciences, University of Basel, Schönbeinstrasse 6, 4056 Basel, Switzerland

[4]University of Kassel, Mönchebergstrasse 19, 34109 Kassel, Germany

[5]Vavilov Institute of General Genetics, Russian Academy of Sciences, 3 Gubkina Str., Moscow 119333, Russia

[6]Laboratory of Forest Genomics, Genome Research and Education Center, Siberian Federal, University, 50a/2 Akademgorodok, Krasnoyarsk 660036, Russia

[7]Department of Ecosystem Science and Management, Texas A&M University, 2138 TAMU, College Station, TX 77843-2138, USA

*****Corresponding author. Tel: +49 551 393 35 37; Fax: +49 551 39 83 67; e-mail: konstantin.krutovsky@forst.uni-goettingen.de; ORCID: 0000-0002-8819-7084

Abstract

European beech (*Fagus sylvatica* L.) is one of the most important forest tree species in Europe, and its genetic adaptation potential to climate change is of great interest. Saplings and adults from 12 European beech populations were sampled along two steep precipitation gradients in Switzerland. All individuals were genotyped at 13 microsatellite markers and 70 SNPs in 24 stress response and phenology related candidate genes. Both SSR and SNP markers had high genetic diversity in the studied populations and low but statistically significant population differentiation. Two approaches were used to discover SNPs with signatures of selection: search for F_{ST} SNP outliers and analyses of SNP associations with environmental variables such as temperatures, precipitation and humidity. Three (4.3%) SNPs were consistently identified as outliers in the adults by more than one method, and they were very likely under positive selection. Twenty (28.6%) SNPs in the saplings and 10 (14.3%) SNPs in the adults were associated with environmental variables found by more than one method. In general, there were 22 (31.4%) SNPs in 17 (70.8%) candidate genes in the saplings, and 16 (22.9%) SNPs in 10 (41.6%) candidate genes in the adults, consistently identified by at least two of the five methods used, indicating that they are very likely under selection. Genes with SNPs showing signatures of selection are involved in a wide range of molecular functions, such as oxidoreductases (*IDH*), hydrolases (*CysPro*), transferases (*XTH*), transporters (*KT2*), chaperones (*CP10*) and transcription factors (*DAG*, *NAC* transcription factor). The obtained data will help us better understand the genetic variation underlying adaptation to environmentally changing conditions in European beech, which is of great importance for the development of scientific guidelines for the sustainable management and conservation of this important species.

Keywords: Adaptation • Climate change • Environmental association analysis • Microsatellite • Outlier analysis • SNP

Introduction

Climate change scenarios predict not only higher annual temperatures, but also changes in precipitation patterns, increasing the risk of extreme events, such as floods and droughts (Trenberth 2011). In Central Europe, an increment in the temperature of 1.3°C has been observed during the first decade of the 21st century compared to the last half of the 19th century. Similarly, the frequency of hot days, tropical nights and heat waves has increased since the last half of the 20th century, whereas cold periods and frost days have been reduced (Kovats et al. 2014). Additionally, an increase in the duration and intensity of summer droughts has also been observed, and this trend is expected to continue through the 21st century (Beniston and Goyette 2007; Kovats et al. 2014).

Changes in climate will very likely affect the survival of forest trees, altering the composition and distribution of forests (Allen et al. 2010; Crookston et al. 2010; Chmura et al. 2011). European beech (*Fagus sylvatica* L.) is one of the most important and widely distributed forest tree species in Europe (Ellenberg 1988). In Switzerland, *F. sylvatica* is the second most important tree species, being predominant in the sub-montane and lower montane range (Weber et al. 2011). Similar to other beech species, its distribution depends mainly on temperature, followed by moisture availability (Fang and Lechowicz 2006). Under climate change, the distribution of beech is expected to be affected, with a reduction in the south and expansion in the north, and a shift in distribution towards higher elevations (Kramer et al. 2010; Bugmann et al. 2014).

Genetic variation is needed for a species to cope with environmental changes. Genetic studies on beech using isozyme, RAPD, AFLP and microsatellite (SSR) markers have found high genetic variation, high gene flow and low population structure in European beech (Sander et al. 2000; Emiliani et al. 2004; Jump and Peñuelas 2007; Kraj and Sztorc 2009; Pluess and Weber 2012). However, those markers have limited potential to study adaptation. In particular, SSR markers are mainly located in non-coding regions (random genomic SSRs) and thus, likely represent selectively neutral genetic variation, i.e., not being under natural selection (Holderegger et al. 2006). Instead, single nucleotide polymorphisms (SNPs) in coding sequences are the most common polymorphisms in genes that can be under selection. They are considered to be more suitable markers to study adaptive genetic variation (Morin et al. 2004). Recently, multiple SNP markers have been developed in climate adaptation related candidate genes in *F. sylvatica* (Seifert et al. 2012; Lalagüe et al. 2014; Müller et al. 2015a), but so far, only a few studies have used them to detect genetic variation showing signatures of selection (Csilléry et al. 2014; Müller et al. 2015b; Pluess et al. 2016; Krajmerová et al. 2017).

Different approaches can be used to identify genetic variation under selection. F_{ST} outlier tests are among the most broadly used methods. They rely on the assumption that non-selective processes have the same effect on all loci in genome, while selection would affect only certain loci (Lewontin and Krakauer 1973). Thus, loci with genetic differentiation (measured by the F_{ST} parameter) higher or lower than expected under neutrality are considered to be under positive or balancing selection, respectively (Vitti et al. 2013).

Environmental association analyses (EAA) are among the most efficient approaches to detect signatures of selection, since they include directly the environmental variables that could drive adaptation (Schoville et al. 2012). They aim at identifying associations between allele frequencies and environmental variables (Rellstab et al. 2015; Stephan 2016), relying on the assumption that alleles in a locus under selection and affected by a particular environmental factor might demonstrate a change in allele frequency following environmental change (Holderegger et al. 2010). Using this approach in plants, associations between genetic variation with temperature and precipitation have been detected in different species, such as *Quercus lobata* (Sork et al. 2010), *Arabis alpina* (Poncet et al. 2010; Manel et al. 2010), *Pseudotsuga menziesii* (Eckert et al. 2009), *Pinus taeda* (Eckert et al. 2010a,b), *P. pinaster* and *P. halepensis* (Grivet et al. 2011). Likewise, in *F. sylvatica*, genetic variation at AFLP markers has been associated with temperature (Jump et al. 2006) and water availability (Pluess and Weber 2012). More recently, SNPs in candidate genes that might be under climate induced selection have been found (Csilléry et al. 2014; Lalagüe et al. 2014), and their associations with environmental variables such as temperature, precipitation and drought have been determined (Pluess et al. 2016). However, the genetic variation underlying adaptation to different environmental conditions in *F. sylvatica* remains insufficiently studied.

Precipitation gradients may cause differences in water availability for plants, and thus, reflect differences in selection pressure acting on forest populations. In this study, populations of *F. sylvatica* occurring along two steep precipitation gradients in Switzerland were selected, and the patterns of selectively neutral genetic variation and population genetic structure were studied by using 13 SSR markers. Additionally, SNPs in candidate genes potentially involved in important traits such as phenology and stress response were used for the detection of genetic variation showing signatures of selection. Firstly, outlier SNPs showing genetic differentiation higher or lower than expected under neutrality were identified by using three different methods implemented in LOSITAN, Arlequin, and BayeScan software. Secondly, SNPs showing association with important environmental variables such as precipitation, temperature and humidity were tested using two different methods: LEA (an R package for Landscape and Ecological Associations studies) and Samβada. SNPs identified by at least two of the five methods were considered very likely to be under selection.

Materials and methods

Plant material

Twelve populations of *F. sylvatica* located in the dry inner-alpine Rhone and Rhine valleys in Switzerland were used in this study (six populations per valley). The populations were located at similar elevations (550-850 m above sea level), with a mean annual temperature between 9.8 and 10.1 °C. The mean annual precipitation ranged between 849 and 1334 mm in the Rhine valley, and between 603 and 1012 mm in the Rhone valley (Table 1). Leaves from 2-4 saplings underneath the same adult tree were collected, for a total of 60-64 saplings sampled per population. Additionally, leaves from 25 adult trees per population were collected. In total, 755 saplings and 300 adult trees were sampled. The leaves were dehydrated with silica gel and stored at room temperature.

DNA isolation

DNA was isolated from dry leaves using the DNeasy™ 96 Plant Kit (Qiagen, Hilden, Germany). The amount and quality of the DNA were examined using electrophoresis in agarose gel at 1% with 1X TAE as running buffer. DNA was stained with Roti®-Safe GelStain (Roth, Karlsruhe, Germany), visualized by UV illumination, and compared with a Lambda DNA size ladder (Roche, Mannheim, Germany).

SSR amplification and genotyping

Individuals were genotyped at 13 SSR loci. Ten SSR loci were random genomic SSRs representing noncoding regions: Six of them were originally developed for *F. sylvatica: FS3-04* (Pastorelli et al. 2003), *msf11* (Vornam et al. 2004), *csolfagus_06*, *csolfagus_19* (Lefèvre et al. 2012), *Fagsyl_002929* and *Fagsyl_003994* (Pluess and Määttänen 2013). Four markers - *sfc0018*, *sfc0161*, *sfc1063* and *sfc1143* - were originally developed for *F. crenata* (Asuka et al. 2004). The other three SSR loci - *GOT066*, *FIR065* and *FIR004* - were EST-linked (EST-SSRs). They were originally developed for *Quercus robur* (Durand et al. 2010), and successfully used for *F. sylvatica* in this study.

The PCR amplifications were performed using fluorescent dye labeled primers as follows: 6-carboxyfluorescein (FAM) dye for *mfs11*, *sfc0161*, *sfc1063*, *csolfagus_06*, *csolfagus_19*, *Fagsyl_003994* and *FIR004*; and 6-hexachlorofluorescein (HEX) dye for *sfc0018*, *sfc1143*, *Fagsyl_002929*, *GOT066*, *FIR065* and *FS3-04*. This allowed us to assemble four different PCR amplification multiplexes. The 1st multiplex was composed of the *FS3-04* and *msf11* markers, the 2nd multiplex - all four *sfc* markers, the 3rd - the *csolfagus* and *Fagsyl* markers, and the 4th - all three EST markers. The PCR amplifications were performed in a total volume of 15 µL containing 2 µL of genomic DNA (about 10 ng), 1X reaction buffer (0.8 M Tris-HCl pH 9.0, 0.2 M $(NH_4)_2SO_4$, 0.2% *w/v* Tween-20; Solis BioDyne, Tartu, Estonia), 2.5 mM $MgCl_2$, 0.2 mM of each dNTP, 0.3 µM of each forward and reverse primer and 1 unit of *Taq* DNA polymerase (HOT FIREPol® DNA Polymerase, Solis BioDyne, Tartu, Estonia). The amplification conditions were as follows: an initial denaturation step at 95 °C for 15 min, followed by 30 cycles consisting of a denaturing step at 94 °C for 1 min, an annealing step at 55 °C (first, second and third multiplexes) or at 47 °C (EST multiplex) for 30 s and an extension step at 72 °C for 1 min. After 30 cycles, a final extension step at 72 °C for 20 min was executed. The PCR fragments were separated and sized on an ABI PRISM® 3100 Genetic Analyzer (Applied Biosystems, Foster City, USA). The GS 500 ROX™ (Applied Biosystems, Foster City, USA) was used as an internal size standard. The genotyping was done using the GeneMapper 4.1® software (Applied Biosystems, Foster City, USA).

Candidate genes and SNPs

SNPs in candidate genes involved in phenology and drought stress tolerance from previously published studies for *F. sylvatica* were selected (Seifert et al. 2012; Lalagüe et al. 2014; Müller et al. 2015a). For the candidate genes that contained several SNPs, linkage disequilibrium (LD) blocks were identified using the htSNPer 1.0 software (Ding et al. 2005), and a subset of SNPs representing the majority of haplotypes (haplotype tag SNPs) was selected for further genotyping. In addition, SNPs showing signatures of natural selection in previous studies (Csilléry et al. 2014; Müller et al. 2015b) were also selected. Finally, 24 genes and 76 SNPs (21 non-synonymous, 27 synonymous and 28 non-coding SNPs) were selected for genotyping (Supplementary material 1 Table S1). Nucleotide sequences neighboring selected SNPs were sent to LGC Genomics Ltd. for primer design and SNP genotyping using the PCR-based KASP™ genotyping assay (Hoddesdon, UK).

Environmental data

Data on climatic variables collected by meteorological stations located near the populations were downloaded from the website of the Federal Office of Meteorology and Climatology MeteoSwiss (http://www.meteoswiss.admin.ch). Climate normals for the reference period 1961-1990 were used as a proxy for the climate that imposed selection pressure on the early life stages of adult trees, whereas climate normals for the reference period 1981-2010 were used for the saplings. The environmental variables included data on annual and growing season (May-September) temperature and precipitation, heat days (HD) and summer days (SD), as well as latitude and longitude (Table 2). Three derived climatic variables were additionally calculated: potential annual direct incident solar radiation (ASR), the Thornthwaite's moisture index (I_m) (Thornthwaite 1948) and the Ellenberg's climatic quotient (EQ) (Jahn 1991) (Table 2). ASR was calculated using data on latitude, slope and aspect according to McCune and Keon (2002). To calculate I_m, first, monthly potential evapotranspiration (PET) was calculated according to (Thornthwaite 1948) using the R package SPEI 1.6 (R Core Team 2016). Then, I_m was calculated according to the formula $I_m = \frac{100s-60d}{n}$, where s is the sum of surplus water for the months when precipitation exceeds PET, d is the sum of water deficiency for the months when PET exceeds precipitation, and n is water need (annual PET) (Thornthwaite 1948; Maliva and Missimer 2012). According to (Thornthwaite 1948), moist climates have positive values of I_m, and dry climates have negative values. The Ellenberg's climatic quotient (EQ), which is widely used to describe habitats suitable for the genus *Fagus*, was calculated as $\mathrm{EQ} = \frac{\text{Temperature of July (°C)}}{\text{Annual precipitation (mm)}} \times 1000$, (Jahn 1991; Fang and Lechowicz 2006). According to Jahn (1991), regions with values of EQ below 20 represent a pure beech climate, while the beech competitiveness slowly decreases in regions with EQ values between 20 and 30 and disappears in regions with EQ > 30. Information about the environmental variables per population and for the reference periods 1961-1990 and 1981-2010 are presented in Supplementary material 1—Table S2.

Spearman's rank correlation coefficients between all pairs of environmental variables were calculated. Principal component analysis (PCA) was used to reduce dimensionality of the environmental variables; variables were standardized to a mean of 0 and standard deviation of 1 before PCA analysis. Principal components (PCs) with eigenvalues greater than 1 were kept for the environmental association analysis; these PCs will be referred further as environmental PCs. All analyses were conducted using the software Statistica 12 (Dell Inc 2015). Environmental PCs as well as individual environmental variables were used further to find their association with SNPs.

Data analysis

Tentative neutral genetic variation (SSRs)

Allelic richness was calculated taking into account differences in sample size with the HP-Rare program (Kalinowski 2005) using a sample size of 50 individuals. Additionally, the diversity parameters, such as observed (H_o) and expected (H_e) heterozygosities and the fixation index (F_{IS}), as well as deviation from Hardy-Weinberg equilibrium, were calculated using the GenAlEx 6.5 software (Peakall and Smouse 2006, 2012). Furthermore, the MICRO-CHECKER software (Van Oosterhout et al. 2004) was used to identify and correct genotyping errors, such as null alleles. Differences in genetic diversity parameters between saplings and adults were tested for significance using the FSTAT 2.9.3.2 software (Goudet 1995). The GENEPOP 4.2 program (Raymond and Rousset 1995; Rousset 2008) was used to test for linkage disequilibrium (LD) between pairs of the SSR loci using 10000 dememorizations, 1000 batches and 10000 iterations per batch for Markov chain parameters.

To assess genetic differentiation, F_{ST} and Hedrick's standardized G''_{ST} (Meirmans and Hedrick 2011) were calculated with the GenAlEx 6.5 software (Peakall and Smouse 2006, 2012) using 999 permutations. Population structure was inferred using the Bayesian approach implemented in the STRUCTURE 2.3.4 software (Pritchard et al. 2000); the analysis was done for genomic SSRs and EST-SSR separately, and for all SSR together. The admixture model with correlated allele frequencies was used. We used 100000 iterations for both the MCMC (Markov chain Monte Carlo) burn-in period and the following MCMC. We tested from 1 to 20 possible populations or clusters (K), using 20 iterations for each of them. The most likely number of clusters K was determined considering mean posterior probability of the data (LnP(D)) and also according to the ΔK method (Evanno et al. 2005), which is implemented in the STRUCTURE HARVESTER 0.6.94 software (Earl and vonHoldt 2012). The CLUMPAK software (Kopelman et al. 2015) was used for summation and graphical representation of the results obtained by STRUCTURE.

Tentative adaptive genetic variation (SNPs)

The genetic diversity parameters H_o and H_e, the index F_{IS} and deviations from Hardy-Weinberg equilibrium, LD between pairs of SNP loci, F_{ST} and Hedrick's standardized G''_{ST} and population structure were analyzed the same way as it is described above for the SSR markers.

Signatures of natural selection

Two different approaches were used to detect SNPs showing signatures of selection: outlier detection and environmental association analyses, respectively. For the detection of outlier SNPs three different methods with different demographic assumptions were used. The first method was developed by Beaumont and Nichols (1996) and is implemented in the LOSITAN software (Antao et al. 2008). This method determines the expected thresholds for distribution of F_{ST} along H_E for loci with selectively neutral variation under an island model of migration. The analysis was done using the infinite allele model with 200000 simulations, a confidence interval of 95% and a false discovery rate (FDR) of 0.1. To run LOSITAN we used a procedure typically used in similar studies (Krutovsky et al. 2009). LOSITAN was run first using all loci to estimate the mean neutral F_{ST}. After the first run, all loci outside the 95% confidence interval were removed, and using only putatively neutral loci that were not removed, LOSITAN was run again to estimate a second mean neutral F_{ST}. Finally, a third run was done using all loci and the second mean neutral F_{ST}. This procedure lowers the bias when estimating the mean neutral F_{ST} by removing, at the end of the first run, the most extreme loci from the estimation (Antao et al. 2008). LOSITAN analysis was done taking into account the entire set of populations, and also for each region (Rhine or Rhone) separately.

The second method is implemented in the Arlequin 3.5 software (Excoffier and Lischer 2010) and is similar to the one implemented in LOSITAN, but considers a hierarchical island model of migration, in which populations exchange more migrants within groups than between groups (Excoffier et al. 2009). Populations of saplings and adults were grouped hierarchically according to the region; furthermore, populations of saplings were also grouped according to the groups suggested by the STRUCTURE analysis based on the all SSR markers. Then, 50000 simulations were carried out, using 10 groups of 100 demes as running conditions as recommended by Excoffier et al. (2009). A FDR of 0.1 was applied using the Benjamini & Hochberg (1995) method implemented in the R script “p.adjust” (R Core Team 2016).

The third outlier detection method is implemented in the BayeScan 2.1 software (Foll and Gaggiotti 2008). It assumes that populations diverged from an ancestral gene pool, and their allele frequencies show different degrees of differentiation from it. Running conditions used in BayeScan were as follows: a burn-in period with 50000 iterations, a thinning interval of 10, a sample size of 5000 and 20 pilot runs with 5000 iterations each, for a total of 100000 iterations. A locus was considered outlier if its q value was less than FDR < 0.05 or 0.1. The BayeScan analysis was done taking into account the entire set of populations, and also for each region separately.

Environmental PCs as well as individual environmental variables were used for the detection of associations with SNPs. Two different methods were used – one implemented in an R package for Landscape and Ecological Associations (LEA) studies (Frichot and François 2015; http://membres-timc.imag.fr/Olivier.Francois/LEA/software.htm) and another implemented in the software Samβada (Stucki et al. 2016; https://lasig.epfl.ch/sambada).

The LEA method tests for associations between allele frequencies and environmental variables based on latent factor mixed models (LFMM), in which associations are tested while estimating the effects of hidden confounding factors, such as population structure and spatial autocorrelation (Frichot et al. 2013). A burning period of 5000 and a total number of 10000 cycles were used. Based on the results of the STRUCTURE analysis using all SSR markers (see Results), the number of clusters (K) was set to 2 in the saplings and 1 in the adults. Five runs were performed; the z-scores obtained from the different runs were combined using a robust variant of the Stouffer method (Whitlock 2005), and the genomic inflation factor λ (Devlin and Roeder 1999) was computed. P-values from the combined z-scores were calibrated by the computed λ as described in the manual of LEA, and if necessary, further calibrated manually by using different values of λ until the histograms showed that the P-values were uniformly distributed (François et al. 2016). The Benjamini-Hochberg procedure (Benjamini and Hochberg 1995) with an expected FDR equalled to 10% was used to correct the P-values for multiple testing.

The Samβada method tests for associations between genotypes and environmental variables using logistic regressions, and allows for the inclusion of population structure (Stucki et al. 2016). SNPs were coded as presence/absence of a given genotype in each individual. Given the results obtained with STRUCTURE 2.3.4 software (Pritchard et al. 2000) for all SSR markers, a multivariate analysis was run in the saplings including population structure as the coefficients of membership (*Q*) for each individual; the *G* scores to assess significance were calculated according to Samβada manual. For the adults, a univariate analysis (without including population structure, see Results) was run. The *G* scores obtained in both multivariate and univariate analyses were used to compute the corresponding *P*-values using a χ^2 distribution with one degree of freedom. Correction for multiple testing was done by adjustment of *P*-values for a FDR equal to 0.1 using the Benjamini & Hochberg (1995) method implemented in the R function "p.adjust" (R Core Team 2016). A SNP was considered to be candidate under selection if at least one of its tree genotypes showed significant association with an environmental PC or environmental variable (Stucki et al. 2016). Graphical representation of logistic regression fits was done with the software JMP®, Version 13.1.0 SAS Institute Inc., Cary, NC, 1989-2007.

Results of the five different methods (LOSITAN, Arlequin, BayeScan, LEA and Samβada) were compared, and loci detected by two or more of them were considered as likely true candidates under selection.

Results

Relationships between environmental variables

Latitude was strongly positively correlated with minimum temperatures, precipitation variables and the moisture index I_m, and moderately negatively correlated with maximum temperatures, SD, HD and EQ based on Spearman's rank correlation coefficients

(Supplementary material 2—Fig. S1). Longitude had either no correlation or weak positive correlations with most of the variables, most of which were not significant. Maximum temperatures were strongly and positively correlated with *SD* and *HD*, while negatively correlated with minimum temperatures and precipitation variables. The Thornthwaite's moisture index I_m was strongly negatively correlated with maximum temperatures and *SD* and *HD*, and strongly positively correlated with precipitation. In contrast, the *EQ* index was positively correlated with maximum temperatures and *SD* and *HD*, and negatively correlated with minimum temperatures and precipitation. *ASR* had either weak or no correlation with all the environmental variables (Supplementary material 2—Fig. S1).

The PCA showed that the top three PCs had eigenvalues higher than 1 and captured the most of the overall variance of the environmental variables for both reference periods: 95.54% for 1961-1990, and 95.99% for 1981-2010 (Table 3). To interpret each environmental PC, environmental variables showing strong correlation coefficients with values more than |0.8| with a given environmental PC were considered (Supplementary material 1—Table S3). Thus, for both reference periods, the environmental PC1 was strongly and positively correlated with latitude, minimum temperatures, precipitation variables and the moisture index I_m, whereas negatively correlated to maximum temperatures, *SD*, *HD* and the *EQ* index (Table 3; Supplementary material 1—Table S3). This indicates that positive values of PC1 represent more humid/colder environments, while negative values indicate drier/warmer environments. The environmental PC2 was strongly correlated only with mean annual temperature, and the environmental PC3 was strongly and positively correlated only with solar radiation (Table 3; Supplementary material 1—Table S3).

Tentative neutral genetic variation (SSRs)

For 13 SSR markers, 4-19 alleles were detected in the saplings and 3-17 alleles were detected in the adults. The F_{IS} indices per locus were close to zero and overall, no significant deviations from Hardy-Weinberg equilibrium were found (Supplementary material 1—Table S4). No loci showed evidence of null alleles. In general, EST-SSRs demonstrated lower genetic diversity than genomic SSRs (Supplementary material 1—Table S4). Analysis of genetic diversity revealed no significant differences between saplings and adults: $A = 6.36$ vs. 6.37 ($P = 0.9$), $H_e = 0.649$ vs. 0.645 ($P = 0.6$) (Table 4). Likewise, there were no significant differences between the two regions neither in the saplings: $A = 6.49$ vs. 6.23 ($P = 0.3$), $H_e = 0.656$ vs. 0651 ($P = 0.1$) nor in the adults: $A = 6.59$ vs. 6.14 ($P = 0.1$), $H_e = 0.651$ vs. 0.650 ($P = 0.8$) (Table 4). The F_{IS} indices were close to zero, and no significant deviations from Hardy-Weinberg equilibrium were found, except for the adult trees in the Saxon population. Significant LD was observed for 15 pairs of all 78 possible pairs (19.2%) of the 13 SSR loci in the populations of saplings (Supplementary material 2—Fig. S2), but only for the *Sfc0018-FIR065* pair (1.3%) in the populations of adults. This pair was in LD also in the saplings.

Genetic differentiation among populations was low but significant for saplings ($F_{ST} = 0.017$, $P < 0.001$; $G''_{ST} = 0.029$, $P < 0.001$) and adults ($F_{ST} = 0.027$, $P < 0.001$; $G''_{ST} = 0.027$, $P < 0.001$). Analysis of population structure based on all SSR together, as well as based on genomic SSRs and EST-SSR separately, revealed that there is no strong clustering neither among saplings nor adults or possibly two clusters (*K*) in the saplings due to Chamoson as a population likely the most genetically different from others (Fig. 1a,b and Supplementary material 2—Fig. S3 - S7), which is supported also by its high pairwise F_{ST} and G''_{ST} values in the adults (Supplementary material 2—Fig. S8).

Tentative adaptive genetic variation (SNPs)

Among the 76 SNPs genotyped, 6 were monomorphic (*APX1_2, PhyB, 50_320, 52_1_249, 92_166, 110_1_111*). Based on the remaining 70 SNPs, both observed and expected heterozygosities were not much different between each other and between saplings and adults: H_o = 0.301 *vs*. 0.311, H_e = 0.309 *vs*. 0.310 for saplings *vs*. adults, respectively (Table 5 and Supplementary material 1—Table S5). Overall, F_{IS} was close to zero, and no significant deviations from Hardy-Weinberg equilibrium were found, except for the Mastrils, Sargans and Ollon populations in the saplings, and the population Mastrils in the adults (Table 5).

In both saplings and adults, LD was mainly found between SNPs in the same gene. In the saplings, significant LD was observed for 134 pairs of all 2415 possible pair combinations of SNPs (5.5%), and 68 of them were found between SNPs in the same gene (Supplementary material 2—Fig. S9). Similarly, for populations of adults, 107 pairs (4.4%) of all the possible pairs showed significant LD, and 59 of them were found between SNPs in the same gene (Supplementary material 2—Fig. S9).

Genetic differentiation was low but significant for populations of both saplings (F_{ST} = 0.020, $P < 0.001$; G''_{ST} = 0.020, $P < 0.001$) and adults (F_{ST} = 0.028, $P < 0.001$; G''_{ST} = 0.016, $P < 0.001$). Likewise, analysis of population structure using SNP markers revealed that there is a weak population structure in both saplings and adults (Fig. 1c and d and Supplementary material 2—Fig. S10).

Signatures of natural selection

In the saplings, no outlier SNPs were identified by LOSITAN when doing the analysis with all populations together and with populations from the Rhine valley. However, the analysis including populations from the Rhone valley detected the SNP *ALDH_4* as outlier possibly being under balancing selection in (Table 6 and Supplementary material 2—Fig. S11).

Arlequin identified the SNPs *ERD*, *CysPro_202* and *NAC_962* as outliers that are likely under positive selection (Table 6 and Supplementary material 2—Fig. S11). No significant outlier SNPs were identified by BayeScan.

More outlier SNPs were identified in the adults than in the saplings. In the LOSITAN analysis for adults, 15 SNPs fell outside the 95% confidence interval when analyzing all populations and populations from each valley separately (Table 6 and Supplementary material 2—Fig. S12). In the Arlequin analysis, 5 SNPs fell outside the 95% interval (Table 6 and Supplementary material 2—Fig. S12), but no significant outliers were detected by BayeScan in the adults. Thus, among the detected outliers, 3 (4.3%) SNPs (CysPro_202, NAC_962 and 92_352 SNPs) are very likely true outliers under selection, because they were detected by both LOSITAN and Arlequin methods in the adults (Table 6). Interestingly, the SNPs CysPro_202 and NAC_962 were also detected by Arlequin in the saplings.

EAA carried out with LEA and Samβada identified additional SNPs showing significant association with the environmental variables and PCs, indicating that they are potentially subject to selection. LEA detected 25 (35.7%) and 27 (38.6%) SNPs in the saplings and adults, respectively (Table 7), while Samβada identified 44 (62.9%) and 16 (22.9%) SNPs in the saplings and adults, respectively (Table 7). Details of the genotypes per SNP showing significant associations with the environmental variables as detected by Samβada can be found in Supplementary material 1—Table S6 and S7. We considered SNPs identified by both LEA and Samβada as very likely under selection: in the saplings, 20 (28.6%) SNPs in 16 (66.6%) genes were detected by both methods, while 10 (14.3%) SNPs in 7 (29.2%) genes were identified by both methods in the adults (Table 7). SNPs detected by both methods showed differences in allele and genotype frequencies along the environmental gradient. For instance, the frequency of the allele G in the *CP10_442* SNP declined with increasing moisture in the saplings (Fig. 2a); similarly, the allele C in the *52_1_235* SNP decreases in frequency with increasing *AP* in the adults (Fig. 2b). On the other hand, in the *50_232* SNP, the frequency of allele A increases with *MaxAT* in the saplings (Fig. 2b), while in the adults the frequency of the allele T in the *IDH_1 SNP* increases with positive values of PC1, i.e., in populations with humid/colder environments (Fig. 2d). Such differences in allele frequencies were also reflected in differences in genotype frequencies (Fig. 3).

Comparing the results from the five different methods used to detect candidate SNPs under selection (LOSITAN, Arlequin, BayeScan, LEA and Samβada), it was found that 22 (31.4%) SNPs in the saplings and 16 (22.9%) SNPs in the adults were detected by at least two methods, and thus, they were considered as very likely true candidates under selection. These SNPs are located in 17 (70.8%) and 10 (41.6%) genes in saplings and adults, respectively.

Discussion

Putative neutral genetic variation (SSRs)

A high genetic variation was found in all the studied populations of *F. sylvatica* (Table 4). No significant differences in genetic variation between saplings and adults were found, suggesting that the saplings represent the genetic variation of the adult populations. Similar levels of genetic variation have been found in other studies based on similar sets of SSR loci (Seifert 2012; Müller 2013; Bontemps et al. 2013; Rajendra et al. 2014), and slightly lower when compared to the studies based on other SSR loci (Buiteveld et al. 2007; Kraj and Sztorc 2009; Chybicki et al. 2009; Bilela et al. 2012). It is known that a high genetic variation, characteristic of woody plants, is due to their large geographic ranges, long lives, outcrossing mating systems and wide pollen and seed dispersal (Hamrick et al. 1992). Among SSRs, EST-SSRs presented lower variation than genomic SSRs (Supplementary material 1—Table S4). Similar results have been reported in other studies (Seifert 2012; Müller 2013), and can be attributed to the location of EST-SSRs in coding regions, making them more conserved and thus, less polymorphic (Varshney et al. 2005; Ellis and Burke 2007).

Null alleles are alleles that fail to amplify due to mutations in the primer annealing site, causing misgenotyping heterozygotes as homozygotes and resulting in a biased estimation of allele frequencies and a reduced observed heterozygosity (Ellis and Burke 2007). They are more likely to occur when SSR loci are transferred from other species. Although seven SSR loci used in this study were transferred from *F. crenata* and *Q. robur*, no loci showed evidence of null alleles, which is supported by the fixation indices (F_{IS}) close to zero (Supplementary material 1—Table S4). These results confirmed the observations from other studies indicating that the transferability of SSR loci among species of the genus *Fagus* is relatively high (Pastorelli et al. 2003; Lefèvre et al. 2012) and that transferability of EST-SSR can be successful even in species from different genus but the same family (Ellis and Burke 2007), as was the case for the EST-SSR transferred from *Q. robur*.

LD between SSR loci was found for 19.2% of all the possible pairs in the saplings. In contrast, 1.3% of all the possible pair combinations were in LD in the adults, which is comparable to the low percentage found in a similar study (Lefèvre et al. 2012). The higher percentage of SSR loci in LD in the saplings could be an effect of relatedness, since groups of 2-4 saplings were collected underneath the same adult tree. In fact, those saplings had higher pairwise relatedness coefficient than saplings collected under different trees (data not shown). Furthermore, since there are no genetic linkage data for the studied loci, it is impossible to see if observed LD is due to close linkage.

The low F_{ST} and G''_{ST} values and the STRUCTURE analysis demonstrated that population differentiation was very weak in the studied populations of *F. sylvatica* (Fig. 1). These findings are in consensus with other studies in beech that also reported low genetic differentiation in Germany (Sander et al. 2000; Rajendra et al. 2014; Müller et al. 2015b), Italy (Paffetti et al. 2012), France (Csilléry et al. 2014) and other parts of Europe (Buiteveld et al. 2007). High gene flow may explain the low differentiation even in populations from different valleys, since *F. sylvatica* is an outcrossing wind-pollinated tree species with long distance pollen flow (Oddou-Muratorio et al. 2011; Piotti et al. 2012). In fact, beech pollen can travel for thousands of kilometers, from Germany and North Italy to Catalonia in Spain (Belmonte et al. 2008). This high pollen dispersal capability can explain the low genetic differentiation, even between populations from the two different valleys. However, despite the low genetic differentiation in general, STRUCTURE analysis with SSRs identified Chamoson as a genetically distinct population (Fig. 1a); additionally, Chamoson also had the highest pairwise population differentiation in the adults. Some past forest management cannot be ruled out as a reason for this pattern.

Tentative adaptive genetic variation (SNPs)

Similar to the SSR markers, SNPs also revealed high genetic variation in the studied populations of European beech (Table 5), comparable to the genetic variation found in other studies using SNP markers (Seifert et al. 2012; Müller et al. 2015a). LD analysis revealed that 5.5% and 4.4% of all the possible SNP pairs were found to be in LD in the saplings and adults, respectively. These values are comparable to the percentage (5.01%) reported by Pluess et al. (2016), but considerably lower than 18.45% reported by Müller et al. (2015b). Furthermore, the low F_{ST} and G''_{ST} values and the inferred population structure also demonstrated that there is a weak population differentiation (Fig. 1c,d). In general, low LD and weak population differentiation should be expected for a highly outcrossing, wind-pollinated tree species, such as European beech (Jump et al. 2006; Aitken et al. 2008).

Signatures of natural selection

The F_{ST} outlier tests are among the most commonly used methods to detect adaptive genetic variation. They assume that loci with genetic differentiation higher or lower than expected under neutrality could be under positive or balancing selection, respectively (Vitti et al. 2013). However, one of the disadvantages of outlier detection tests is that they can produce false outliers due to hidden population structure and other confounding effects such as migrations, recent demographic expansions and bottlenecks (Schoville et al. 2012; Vitti et al. 2013). To address this problem, outlier methods with different demographic assumptions can be used (Li et al. 2012). Thus, loci appearing as outliers when considering different methods will be more likely to be real loci under selection. In this study, three different outlier detection methods were used, and they detected only partly overlapping sets of outlier SNPs (Table 6). Discrepancies between different outlier detection methods are common and have been reported also in other studies (Russello et al. 2012; Tsumura et al. 2014; Konijnendijk et al. 2015). This can be attributed, on the one hand, to the different demographic assumptions underlying each method, and, on the other hand, to the different rates of type I (false positives) and type II (false negatives) errors (Narum and Hess 2011). Interestingly, no SNPs were identified as outliers by BayeScan. Indeed, BayeScan is considered more conservative in identifying outlier SNPs than other methods (Narum and Hess 2011). In total, only three SNPs (4.3%) were detected as outliers under positive selection by at least two methods in the adults - *CysPro_202*, *NAC_962* and *92_352* (Table 6). We consider them as likely true outlier SNPs under selection. The first two of them were also detected as outliers in the saplings. The small proportion of outlier loci detected is in line with other studies carried out in forest trees, such as boreal black spruce (Prunier et al. 2011), *Cryptomeria japonica* (Tsumura et al. 2014) and *Quercus petraea* (Alberto et al. 2013). This may be due to the limited sensitivity of outlier methods to identify markers under weak selection (Narum and Hess 2011). Indeed, detection of outliers can be difficult, if there are subtle changes in allele frequencies, such as in the case of polygenic traits, in which adaptation involves subtle changes in allele frequencies at the loci controlling the polygenic trait, or when there is a high gene flow counteracting selection (Rellstab et al. 2015; Stephan 2016).

Unlike outlier detection tests, EAA are more sensitive to detect subtle changes in allele frequencies caused by weak selection (De Mita et al. 2013; Stephan 2016); this would explain the higher number of SNPs potentially under selection detected by EAA when compared to outlier methods (Table 6 and 7). However, EAA approaches could be prone to false positives, especially if a hidden population structure is unaccounted (Rellstab et al. 2015). In this study, weak population structure was found both in saplings and adults, although there are possibly two clusters in the saplings (Fig. 1a). Thus, the potential confounding effect of neutral genetic structure in the saplings was accounted for in the analysis with LEA and Samβada. In general, the two methods detected different sets of SNPs as potential candidates under selection (Table 7); similar findings have been reported in other studies (Christmas et al. 2016; Stucki et al. 2016), and are expected given the different statistical frameworks of the methods (Frichot et al. 2013; Lotterhos and Whitlock 2015; Frichot and François 2015; Stucki et al. 2016). Consequently, when a marker is detected by several methods, it could be considered a very likely true positive (de Villemereuil et al. 2014). Thus, we considered SNPs detected simultaneously by LEA and Samβada as the most likely true candidates to be under selection. In total, 28.6% and 14.3% of the 70 SNPs were consistently identified by both EAA methods in saplings and adults, respectively (Table 7), and they showed differences in allele and genotype frequencies in contrasting environments, as demonstrated for some example SNPs in Figs 2 and 3.

In total, 31.4% and 22.9% SNPs were detected by at least two of the five methods (LOSITAN, Arlequin, BayeScan, LEA and Samβada) in saplings and adults, respectively, and were considered as the most likely true candidates under selection in the studied populations. Some of these SNPs have also shown evidence of selection in other studies of European beech; for example, the *CP10_442*, *CysPro_728*, *DAG_289*, *NAC_854*, *NAC_1300* and *PPC2C_1200* SNPs have been associated with the important trait - bud burst (Müller et al. 2015b), and the *50_232*, *52_1_235*, *52_1_368*, *68_277*, *91_2_57*, *91_2_141* and *91_2_479* SNPs have shown evidence of epistatic selection (Csilléry et al. 2014). Although the rest of the SNPs found to be very likely under selection in this study have not been reported as such by other studies on European beech where they were genotyped, those studies showed that other SNPs from the same genes could be under selection (Csilléry et al. 2014; Müller et al. 2015b; Pluess et al. 2016; Krajmerová et al. 2017), stressing the importance of the studied candidate genes in the adaptation of European beech to different environmental conditions. Besides, SNPs in these genes have also shown signatures of selection in other plant species. For example, SNPs in the *Dhn* gene have been associated with temperature in *Pinus pinaster* (Grivet et al. 2011), SNPs in the *NAC* gene have been detected as potentially under selection by outlier analyses in white and black spruce (Namroud et al. 2008; Prunier et al. 2011), SNPs in the *CAT* gene have been identified as outliers in *Quercus petraea* (Alberto et al. 2013), and SNPs in the *DAG* and *PP2C* genes have been associated with environmental variables such as temperature and water availability in *Dodonaea viscosa* (Christmas et al. 2016).

Interestingly, some different SNPs showing signatures of selection were detected in saplings and adults (Table 6 and Table 7). Not only the environment can exert different selection pressures at different life stages (Petit and Hampe 2006), but also different sets of genes could be involved in the same trait at different stages (Prunier et al. 2013). Therefore, different SNPs could be under selection at the different ages. Moreover, due to high competition and mortality, only a small fraction of seeds survive until the adult stage (Petit and Hampe 2006), which means that adult trees have passed through different selection pressures during their life, and this could be reflected in the different set of SNPs showing signatures of selection in saplings and adults.

Not only non-synonymous SNPs showed signatures of selection, but also synonymous and non-coding SNPs. Since non-synonymous SNPs represent amino acid replacements and thus, a change in protein sequence, they have been traditionally thought to be the main target of natural selection. However, some studies indicated that synonymous substitutions may affect mRNA splicing, stability and translation kinetics (Chamary et al. 2006; Komar 2007), and thus, also affect the production of the final protein (Pagani et al. 2005). Similarly, SNPs in non-coding regions may also be involved in regulation of gene expression (Barrett et al. 2012). Therefore, synonymous and non-coding SNPs can also be subjected to natural selection directly, and not only due to a tight linkage with selective loci.

SNPs showing signatures of selection were located in 70.8% and 41.6% of the studied candidate genes in saplings and adults, respectively. They are involved in a wide range of cellular functions and represent oxidoreductases, hydrolases, oxidases, transferases, transporters, chaperones and transcription factors. This is expected since many traits in plants are polygenic, involving complex interactions among several genes (Ingvarsson and Street 2011). In addition, several SNPs at the same gene showed signatures of selection in this study, and even though some of them were identified only by one method and could be considered false positives, they should not be disregarded for further investigation, especially since some of them have been found to be associated with important climate-related traits and environmental variables in other studies (Müller et al. 2015b; Pluess et al. 2016). Thus, to determine their participation in the adaptation to different environmental conditions of populations of European beech, other approaches could be used. For example, haplotypes can have a substantial advantage over single SNP analysis for the detection of adaptive genetic variation (Balding 2006; Rajora et al. 2016), as well models incorporating polygenic and epistatic selection (Pritchard and Di Rienzo 2010; Fu and Akey 2013; Csilléry et al. 2014).

Additionally, it is possible that other environmental factors that were not accounted for could also exert selection pressure on the studied populations. In this study, climate data were taken from stations less than 10 km away from the actual populations. However, the Alps have high variation in topography, and climatic factors such as temperature and precipitation can vary over short distances (Baruck et al. 2016). Therefore, small-scale heterogeneity and microclimatic conditions specific to a respective population that were not accounted for, could explain some of the differences in allele frequencies. Furthermore, although precipitation and temperature are the main climatic factors influencing plants' distribution, which is supported by several studies that showed their association with potential adaptive genetic variation in the Alps (Poncet et al. 2010; Manel et al. 2012; Pluess et al. 2016), soil properties might also affect plants' distribution because water availability depends on the interaction between climatic variables and soil characteristics (Piedallu et al. 2013). For example, (Gärtner et al. 2008) found that lower humidity can be compensated for by greater available soil water storage capacity (ASWSC) that allows the growth of beech. Low soil water availability affects survival and competitive interactions between beech and other species (Fotelli et al. 2002; Fotelli et al. 2004) and determines the transition from beech to *Quercus pubescens*, a more drought tolerant tree species (Gärtner et al. 2008). In the Alps, soil properties affect not only the present distribution of plants, but also determined the migration pathways during the post-glacial recolonization (Alvarez et al. 2009). Thus, the identification of adaptive genetic variation might be improved by including not only climatic variables but also soil characteristics and microclimatic conditions. However, characteristics of alpine soils vary considerably over short spatial ranges, and soil information is still limited (Baruck et al. 2016).

In this study, a candidate gene approach was used to investigate adaptive genetic variation in beech. By combining genetic variation in SNPs in candidate genes, outlier detection tests and environmental association analysis, it was possible to identify loci showing signatures of selection. This opens new perspectives for understanding the genetic basis of adaptation of *F. sylvatica* to different environmental conditions.

Acknowledgements We thank Florian Schreyer and Jhon Rivera-Monroy for their help in the fieldwork, and Alexandra Dolynska for her assistance in the laboratory. We also thank Hadrien Lalagüe for providing the SNP positions and haplotype sequences from his study to select the haplotype tag SNPs. We also thank COLFUTURO and the Administrative Department of Science, Technology and Innovation COLCIENCIAS for supporting Laura Cuervo-Alarcon during this study. This study was financially supported by the Swiss Federal Office for the Environment FOEN and the Swiss Federal Institute for Forest, Snow and Landscape Research WSL.

Author's contributions LCA collected the samples, generated and analyzed data, and wrote the manuscript. MA helped with the sample collection; MA, CS, KVK and RF conceived and designed the study, developed experimental plan, coordinated the research, and participated in the drafting of the manuscript. MM helped with data analysis, interpretation, and manuscript editing. All authors read and approved the final manuscript.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflicts of interest.

Data Archiving Statement (in progress)

References

Aitken SN, Yeaman S, Holliday JA, Wang T, Curtis-McLane S (2008) Adaptation, migration or extirpation: climate change outcomes for tree populations. Evol Appl 1:95–111. https://doi.org/10.1111/j.1752-4571.2007.00013.x

Alberto FJ, Derory J, Boury C, Frigerio J-M, Zimmermann NE, Kremer A (2013) Imprints of Natural Selection Along Environmental Gradients in Phenology-Related Genes of *Quercus petraea*. Genetics 195:495–512. https://doi.org/10.1534/genetics.113.153783

Allen CD, Macalady AK, Chenchouni H, Bachelet D, McDowell N, Vennetier M, Kitzberger T, Rigling A, Breshears DD, Hogg EH (Ted), Gonzalez P, Fensham R, Zhang Z, Castro J, Demidova N, Lim J-H, Allard G, Running SW, Semerci A, Cobb N (2010) A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests. For Ecol Manag 259:660–684. https://doi.org/10.1016/j.foreco.2009.09.001

Alvarez N, Thiel-Egenter C, Tribsch A, Holderegger R, Manel S, Schönswetter P, Taberlet P, Brodbeck S, Gaudeul M, Gielly L, Küpfer P, Mansion G, Negrini R, Paun O, Pellecchia M, Rioux D, Schüpfer F, Van Loo M, Winkler M, Gugerli F, IntraBioDiv Consortium (2009) History or ecology? Substrate type as a major driver of spatial genetic structure in Alpine plants. Ecol Lett 12:632–640. https://doi.org/10.1111/j.1461-0248.2009.01312.x

Antao T, Lopes A, Lopes RJ, Beja-Pereira A, Luikart G (2008) LOSITAN: A workbench to detect molecular adaptation based on a Fst-outlier method. BMC Bioinformatics 9:323. https://doi.org/10.1186/1471-2105-9-323

Arend M, Gessler A, Schaub M (2016) The influence of the soil on spring and autumn phenology in European beech. Tree Physiol 36:78–85. https://doi.org/10.1093/treephys/tpv087

Asuka Y, Tani N, Tsumura Y, Tomaru N (2004) Development and characterization of microsatellite markers for *Fagus crenata* Blume. Mol Ecol Notes 4:101–103. https://doi.org/10.1046/j.1471-8286.2003.00583.x

Balding DJ (2006) A tutorial on statistical methods for population association studies. Nat Rev Genet 7:781–791. https://doi.org/10.1038/nrg1916

Barrett LW, Fletcher S, Wilton SD (2012) Regulation of eukaryotic gene expression by the untranslated gene regions and other non-coding elements. Cell Mol Life Sci 69:3613–3634. https://doi.org/10.1007/s00018-012-0990-9

Baruck J, Nestroy O, Sartori G, Baize D, Traidl R, Vrščaj B, Bräm E, Gruber FE, Heinrich K,

Geitner C (2016) Soil classification and mapping in the Alps: The current state and future challenges. Geoderma 264, Part B:312–331. https://doi.org/10.1016/j.geoderma.2015.08.005
Beaumont MA, Nichols RA (1996) Evaluating Loci for Use in the Genetic Analysis of Population Structure. Proc Biol Sci 263:1619–1626.
Belmonte J, Alarcón M, Avila A, Scialabba E, Pino D (2008) Long-range transport of beech (*Fagus sylvatica* L.) pollen to Catalonia (north-eastern Spain). Int J Biometeorol 52:675–687. https://doi.org/10.1007/s00484-008-0160-9
Beniston M, Goyette S (2007) Changes in variability and persistence of climate in Switzerland: Exploring 20th century observations and 21st century simulations. Glob Planet Change 57:1–15. https://doi.org/10.1016/j.gloplacha.2006.11.004
Benjamini Y, Hochberg Y (1995) Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J R Stat Soc Ser B Methodol 57:289–300.
Bilela S, Dounavi A, Fussi B, Konnert M, Holst J, Mayer H, Rennenberg H, Simon J (2012) Natural regeneration of *Fagus sylvatica* L. adapts with maturation to warmer and drier microclimatic conditions. For Ecol Manag 275:60–67. https://doi.org/10.1016/j.foreco.2012.03.009
Bontemps A, Klein EK, Oddou-Muratorio S (2013) Shift of spatial patterns during early recruitment in *Fagus sylvatica*: Evidence from seed dispersal estimates based on genotypic data. For Ecol Manag 305:67–76. https://doi.org/10.1016/j.foreco.2013.05.033
Buiteveld J, Vendramin GG, Leonardi S, Kamer K, Geburek T (2007) Genetic diversity and differentiation in European beech (*Fagus sylvatica* L.) stands varying in management history. For Ecol Manag 247:98–106. https://doi.org/10.1016/j.foreco.2007.04.018
Bugmann H, Brang P (Lead authors), Elkin C, Henne P, Jakoby O, Lévesque M, Lischke H, Psomas A, Rigling A, Wermelinger B, Zimmermann NE (Contributing authors) (2014) Climate change impacts on tree species, forest properties, and ecosystem services. Chapter 8 in Appenzeller C, Fischer EM, Fuhrer J, Grosjean M, Hohmann R, Joos F, Raible C, Ritz C (Coordinating Group). CH2014-Impacts. Toward Quantitative Scenarios of Climate Change Impacts in Switzerland. OCCR, FOEN, MeteoSwiss, C2SM, Agroscope, and ProClim, Bern, Switzerland, pp. 79–88
Chamary JV, Parmley JL, Hurst LD (2006) Hearing silence: non-neutral evolution at synonymous sites in mammals. Nat Rev Genet 7:98–108. https://doi.org/10.1038/nrg1770
Chmura DJ, Anderson PD, Howe GT, Harrington CA, Halofsky JE, Peterson DL, Shaw DC,

Brad St.Clair J (2011) Forest responses to climate change in the northwestern United States: Ecophysiological foundations for adaptive management. For Ecol Manag 261:1121–1142. https://doi.org/10.1016/j.foreco.2010.12.040

Christmas MJ, Biffin E, Breed MF, Lowe AJ (2016) Finding needles in a genomic haystack: targeted capture identifies clear signatures of selection in a non-model plant species. Mol Ecol 25:4216–4233. https://doi.org/10.1111/mec.13750

Chybicki IJ, Trojankiewicz M, Oleksa A, Dzialuk A, Burczyk J (2009) Isolation-by-distance within naturally established populations of European beech (*Fagus sylvatica*). Botany 87:791–798. https://doi.org/10.1139/B09-049

Crookston NL, Rehfeldt GE, Dixon GE, Weiskittel AR (2010) Addressing climate change in the forest vegetation simulator to assess impacts on landscape forest dynamics. For Ecol Manag 260:1198–1211. https://doi.org/10.1016/j.foreco.2010.07.013

Csilléry K, Lalagüe H, Vendramin GG, González-Martínez SC, Fady B, Oddou-Muratorio S (2014) Detecting short spatial scale local adaptation and epistatic selection in climate-related candidate genes in European beech (*Fagus sylvatica*) populations. Mol Ecol 23:4696–4708. https://doi.org/10.1111/mec.12902

De Mita S, Thuillet A-C, Gay L, Ahmadi N, Manel S, Ronfort J, Vigouroux Y (2013) Detecting selection along environmental gradients: analysis of eight methods and their effectiveness for outbreeding and selfing populations. Mol Ecol 22:1383–1399. https://doi.org/10.1111/mec.12182

De Villemereuil P, Frichot É, Bazin É, François O, Gaggiotti OE (2014) Genome scan methods against more complex models: when and how much should we trust them? Mol Ecol 23:2006–2019. https://doi.org/10.1111/mec.12705

Dell Inc (2015) Dell Statistica (data analysis software system), version 12. software.dell.com

Devlin B, Roeder K (1999) Genomic Control for Association Studies. Biometrics 55:997–1004. https://doi.org/10.1111/j.0006-341X.1999.00997.x

Ding K, Zhang J, Zhou K, Shen Y, Zhang X (2005) htSNPer1.0: software for haplotype block partition and htSNPs selection. BMC Bioinformatics 6:38. https://doi.org/10.1186/1471-2105-6-38

Durand J, Bodénès C, Chancerel E, Frigerio J-M, Vendramin G, Sebastiani F, Buonamici A, Gailing O, Koelewijn H-P, Villani F, Mattioni C, Cherubini M, Goicoechea PG, Herrán A, Ikaran Z, Cabané C, Ueno S, Alberto F, Dumoulin P-Y, Guichoux E, Daruvar A de, Kremer A, Plomion C (2010) A fast and cost-effective approach to develop and map EST-SSR markers: oak as a case study. BMC Genomics 11:570.

https://doi.org/10.1186/1471-2164-11-570
Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv Genet Resour 4:359–361. https://doi.org/10.1007/s12686-011-9548-7
Eckert AJ, Bower AD, González-Martínez SC, Wegrzyn JL, Coop G, Neale DB (2010a) Back to nature: ecological genomics of loblolly pine (*Pinus taeda*, Pinaceae). Mol Ecol 19:3789–3805. https://doi.org/10.1111/j.1365-294X.2010.04698.x
Eckert AJ, Bower AD, Wegrzyn JL, Pande B, Jermstad KD, Krutovsky KV, Clair JBS, Neale DB (2009) Association Genetics of Coastal Douglas Fir (*Pseudotsuga menziesii* var. *menziesii*, Pinaceae). I. Cold-Hardiness Related Traits. Genetics 182:1289–1302. https://doi.org/10.1534/genetics.109.102350
Eckert AJ, Heerwaarden J van, Wegrzyn JL, Nelson CD, Ross-Ibarra J, González-Martínez SC, Neale DB (2010b) Patterns of Population Structure and Environmental Associations to Aridity Across the Range of Loblolly Pine (*Pinus taeda* L., Pinaceae). Genetics 185:969–982. https://doi.org/10.1534/genetics.110.115543
Ellenberg H (1988) Vegetation Ecology of Central Europe. Cambridge University Press, Cambridge
Ellis JR, Burke JM (2007) EST-SSRs as a resource for population genetic analyses. Heredity 99:125–132. https://doi.org/10.1038/sj.hdy.6801001
Emiliani G, Paffetti D, Vettori C, Giannini R (2004) Geographic distribution of genetic variability of *Fagus sylvatica* L. in southern Italian populations. For Genet Slovak Repub 11(3-4):231–237.
Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software structure: a simulation study. Mol Ecol 14:2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x
Excoffier L, Hofer T, Foll M (2009) Detecting loci under selection in a hierarchically structured population. Heredity 103:285–298. https://doi.org/10.1038/hdy.2009.74
Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resour 10:564–567. https://doi.org/10.1111/j.1755-0998.2010.02847.x
Fang J, Lechowicz MJ (2006) Climatic limits for the present distribution of beech (*Fagus* L.) species in the world. J Biogeogr 33:1804–1819. https://doi.org/10.1111/j.1365-2699.2006.01533.x
Foll M, Gaggiotti O (2008) A Genome-Scan Method to Identify Selected Loci Appropriate

for Both Dominant and Codominant Markers: A Bayesian Perspective. Genetics 180:977–993. https://doi.org/10.1534/genetics.108.092221

Fotelli MN, Rennenberg H, Geßler A (2002) Effects of Drought on the Competitive Interference of an Early Successional Species (*Rubus fruticosus*) on *Fagus sylvatica* L. Seedlings: 15N Uptake and Partitioning, Responses of Amino Acids and other N Compounds. Plant Biol 4:311–320. https://doi.org/10.1055/s-2002-32334

Fotelli MN, Rienks M, Rennenberg H, Geßler A (2004) Climate and forest management affect 15N-uptake, N balance and biomass of European beech seedlings. Trees 18:157–166. https://doi.org/10.1007/s00468-003-0289-4

François O, Martins H, Caye K, Schoville SD (2016) Controlling false discoveries in genome scans for selection. Mol Ecol 25:454–469. https://doi.org/10.1111/mec.13513

Frichot E, François O (2015) LEA: An R package for landscape and ecological association studies. Methods Ecol Evol 6:925–929. https://doi.org/10.1111/2041-210X.12382

Frichot E, Schoville SD, Bouchard G, François O (2013) Testing for Associations between Loci and Environmental Gradients Using Latent Factor Mixed Models. Mol Biol Evol 30:1687–1699. https://doi.org/10.1093/molbev/mst063

Fu W, Akey JM (2013) Selection and adaptation in the human genome. Annu Rev Genomics Hum Genet 14:467–489. https://doi.org/10.1146/annurev-genom-091212-153509

Gärtner S, Reif A, Xystrakis F, Sayer U, Bendagha N, Matzarakis A (2008) The drought tolerance limit of *Fagus sylvatica* forest on limestone in southwestern Germany. J Veg Sci 19:757–768. https://doi.org/10.3170/2008-8-18442

Goudet J (1995) FSTAT (Version 1.2): A Computer Program to Calculate F-Statistics. J Hered 86:485–486.

Grivet D, Sebastiani F, Alía R, Bataillon T, Torre S, Zabal-Aguirre M, Vendramin GG, González-Martínez SC (2011) Molecular Footprints of Local Adaptation in Two Mediterranean Conifers. Mol Biol Evol 28:101–116. https://doi.org/10.1093/molbev/msq190

Hamrick JL, Godt MJW, Sherman-Broyles SL (1992) Factors influencing levels of genetic diversity in woody plant species. New For 6:95–124. https://doi.org/10.1007/BF00120641

Holderegger R, Buehler D, Gugerli F, Manel S (2010) Landscape genetics of plants. Trends Plant Sci 15:675–683. https://doi.org/10.1016/j.tplants.2010.09.002

Holderegger R, Kamm U, Gugerli F (2006) Adaptive vs. neutral genetic diversity: implications for landscape genetics. Landsc Ecol 21:797–807.

https://doi.org/10.1007/s10980-005-5245-9

Ingvarsson PK, Street NR (2011) Association genetics of complex traits in plants. New Phytol 189:909–922. https://doi.org/10.1111/j.1469-8137.2010.03593.x

Jahn G (1991) Temperate deciduous forests. In: Röhrig E, Ulrich B (eds) Ecosystems of the world. Elsevier, Amsterdam, The Netherlands, pp 377–502

Jump AS, Hunt JM, Martínez-Izquierdo JA, Peñuelas J (2006) Natural selection and climate change: temperature-linked spatial and temporal trends in gene frequency in *Fagus sylvatica*. Mol Ecol 15:3469–3480. https://doi.org/10.1111/j.1365-294X.2006.03027.x

Jump AS, Peñuelas J (2007) Extensive spatial genetic structure revealed by AFLP but not SSR molecular markers in the wind-pollinated tree, *Fagus sylvatica*. Mol Ecol 16:925–936. https://doi.org/10.1111/j.1365-294X.2006.03203.x

Kalinowski ST (2005) hp-rare 1.0: a computer program for performing rarefaction on measures of allelic richness. Mol Ecol Notes 5:187–189. https://doi.org/10.1111/j.1471-8286.2004.00845.x

Komar AA (2007) Genetics. SNPs, silent but not invisible. Science 315:466–467. https://doi.org/10.1126/science.1138239

Konijnendijk N, Shikano T, Daneels D, Volckaert FAM, Raeymaekers JAM (2015) Signatures of selection in the three-spined stickleback along a small-scale brackish water – freshwater transition zone. Ecol Evol 5:4174–4186. https://doi.org/10.1002/ece3.1671

Kopelman NM, Mayzel J, Jakobsson M, Rosenberg NA, Mayrose I (2015) CLUMPAK: a program for identifying clustering modes and packaging population structure inferences across K. Mol Ecol Resour 15:1179–1191. https://doi.org/10.1111/1755-0998.12387

Kovats RS, Valentini R, Bouwer LM, Georgopoulou E, Jacob D, Martin E, Rounsevell M, Soussana JF (2014) Europe. In: Barros V., Field CB, Dokken DJ, Mastrandrea MD, Mach KJ, Bilir TE, Chatterjee M, Ebi KL, Estrada YO, Genova RC, Girma B, Kissel ES, Levy AN, MacCracken S, Mastrandrea PR, White LL (eds) Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part B: Regional Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp 1267–1326

Kraj W, Sztorc A (2009) Genetic structure and variability of phenological forms in the European beech (*Fagus sylvatica* L.). Ann For Sci 66:203–203. https://doi.org/10.1051/forest/2008085

Krajmerová D, Hrivnák M, Ditmarová Ľ, Jamnická G, Kmeť J, Kurjak D, Gömöry D (2017)

Nucleotide polymorphisms associated with climate, phenology and physiological traits in European beech (*Fagus sylvatica* L.). New For 1–15. https://doi.org/10.1007/s11056-017-9573-9

Kramer K, Degen B, Buschbom J, Hickler T, Thuiller W, Sykes MT, de Winter W (2010) Modelling exploration of the future of European beech (*Fagus sylvatica* L.) under climate change—Range, abundance, genetic diversity and adaptive response. For Ecol Manag 259:2213–2222. https://doi.org/10.1016/j.foreco.2009.12.023

Krutovsky KV, Clair JBS, Saich R, Hipkins VD, Neale DB (2009) Estimation of population structure in coastal Douglas-fir [*Pseudotsuga menziesii* (Mirb.) Franco var. *menziesii*] using allozyme and microsatellite markers. Tree Genet Genomes 5:641–658. https://doi.org/10.1007/s11295-009-0216-y

Lalagüe H, Csilléry K, Oddou-Muratorio S, Safrana J, Quattro C de, Fady B, González-Martínez SC, Vendramin GG (2014) Nucleotide diversity and linkage disequilibrium at 58 stress response and phenology candidate genes in a European beech (*Fagus sylvatica* L.) population from southeastern France. Tree Genet Genomes 10:15–26. https://doi.org/10.1007/s11295-013-0658-0

Lefèvre S, Wagner S, Petit RJ, De Lafontaine G (2012) Multiplexed microsatellite markers for genetic studies of beech. Mol Ecol Resour 12:484–491. https://doi.org/10.1111/j.1755-0998.2011.03094.x

Lewontin RC, Krakauer J (1973) Distribution of Gene Frequency as a Test of the Theory of the Selective Neutrality of Polymorphisms. Genetics 74:175–195.

Li J, Li H, Jakobsson M, Li S, Sjödin P, Lascoux M (2012) Joint analysis of demography and selection in population genetics: where do we stand and where could we go? Mol Ecol 21:28–44. https://doi.org/10.1111/j.1365-294X.2011.05308.x

Lotterhos KE, Whitlock MC (2015) The relative power of genome scans to detect local adaptation depends on sampling design and statistical method. Mol Ecol 24:1031–1046. https://doi.org/10.1111/mec.13100

Maliva R, Missimer T (2012) Aridity and Drought. In: Arid Lands Water Evaluation and Management, 1st edn. Springer Berlin Heidelberg, pp 21–39

Manel S, Gugerli F, Thuiller W, Alvarez N, Legendre P, Holderegger R, Gielly L, Taberlet P, IntraBioDiv Consortium (2012) Broad-scale adaptive genetic variation in alpine plants is driven by temperature and precipitation. Mol Ecol 21:3729–3738. https://doi.org/10.1111/j.1365-294X.2012.05656.x

Manel S, Poncet BN, Legendre P, Gugerli F, Holderegger R (2010) Common factors drive adaptive genetic variation at different spatial scales in *Arabis alpina*. Mol Ecol 19:3824–3835. https://doi.org/10.1111/j.1365-294X.2010.04716.x

McCune B, Keon D (2002) Equations for potential annual direct incident radiation and heat load. J Veg Sci 13:603–606. https://doi.org/10.1111/j.1654-1103.2002.tb02087.x

Meirmans PG, Hedrick PW (2011) Assessing population structure: FST and related measures. Mol Ecol Resour 11:5–18. https://doi.org/10.1111/j.1755-0998.2010.02927.x

Morin PA, Luikart G, Wayne RK, group the S workshop (2004) SNPs in ecology, evolution and conservation. Trends Ecol Evol 19:208–216. https://doi.org/10.1016/j.tree.2004.01.009

Müller M (2013) A candidate gene-based association study to investigate potentially adaptive genetic variation in European beech (*Fagus sylvatica* L.). Dissertation, Georg-August-University Göttingen

Müller M, Seifert S, Finkeldey R (2015a) Identification of SNPs in candidate genes potentially involved in bud burst in European beech (*Fagus sylvatica* L.). Silvae Genet 64:1–20.

Müller M, Seifert S, Finkeldey R (2015b) A candidate gene-based association study reveals SNPs significantly associated with bud burst in European beech (*Fagus sylvatica* L.). Tree Genet Genomes 11:1–13. https://doi.org/10.1007/s11295-015-0943-1

Namroud M-C, Beaulieu J, Juge N, Laroche J, Bousquet J (2008) Scanning the genome for gene single nucleotide polymorphisms involved in adaptive population differentiation in white spruce. Mol Ecol 17:3599–3613. https://doi.org/10.1111/j.1365-294X.2008.03840.x

Narum SR, Hess JE (2011) Comparison of FST outlier tests for SNP loci under selection. Mol Ecol Resour 11:184–194. https://doi.org/10.1111/j.1755-0998.2011.02987.x

Oddou-Muratorio S, Klein EK, Vendramin GG, Fady B (2011) Spatial vs. temporal effects on demographic and genetic structures: the roles of dispersal, masting and differential mortality on patterns of recruitment in *Fagus sylvatica*. Mol Ecol 20:1997–2010. https://doi.org/10.1111/j.1365-294X.2011.05039.x

Paffetti D, Travaglini D, Buonamici A, Nocentini S, Vendramin GG, Giannini R, Vettori C (2012) The influence of forest management on beech (*Fagus sylvatica* L.) stand structure and genetic diversity. For Ecol Manag 284:34–44. https://doi.org/10.1016/j.foreco.2012.07.026

Pagani F, Raponi M, Baralle FE (2005) Synonymous mutations in CFTR exon 12 affect splicing and are not neutral in evolution. Proc Natl Acad Sci USA 102:6368–6372.

https://doi.org/10.1073/pnas.0502288102
Pastorelli R, Smulders MJM, Van'T Westende WPC, Vosman B, Giannini R, Vettori C, Vendramin GG (2003) Characterization of microsatellite markers in *Fagus sylvatica* L. and *Fagus orientalis* Lipsky. Mol Ecol Notes 3:76–78. https://doi.org/10.1046/j.1471-8286.2003.00355.x
Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol Ecol Notes 6:288–295. https://doi.org/10.1111/j.1471-8286.2005.01155.x
Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. Bioinformatics 28:2537–2539. https://doi.org/10.1093/bioinformatics/bts460
Petit RJ, Hampe A (2006) Some Evolutionary Consequences of Being a Tree. Annu Rev Ecol Evol Syst 37:187–214. https://doi.org/10.1146/annurev.ecolsys.37.091305.110215
Piedallu C, Gégout J-C, Perez V, Lebourgeois F (2013) Soil water balance performs better than climatic water variables in tree species distribution modelling. Glob Ecol Biogeogr 22:470–482. https://doi.org/10.1111/geb.12012
Piotti A, Leonardi S, Buiteveld J, Geburek T, Gerber S, Kramer K, Vettori C, Vendramin GG (2012) Comparison of pollen gene flow among four European beech (*Fagus sylvatica* L.) populations characterized by different management regimes. Heredity 108:322–331. https://doi.org/10.1038/hdy.2011.77
Pluess AR, Frank A, Heiri C, Lalagüe H, Vendramin GG, Oddou-Muratorio S (2016) Genome–environment association study suggests local adaptation to climate at the regional scale in *Fagus sylvatica*. New Phytol 210:589–601. https://doi.org/10.1111/nph.13809
Pluess AR, Määttänen K (2013) Characterization of eighteen novel microsatellite markers and multiplex PCR protocol for *Fagus sylvatica*. Conserv Genet Resour 5:311–314. https://doi.org/10.1007/s12686-012-9791-6
Pluess AR, Weber P (2012) Drought-Adaptation Potential in *Fagus sylvatica*: Linking Moisture Availability with Genetic Diversity and Dendrochronology. PLoS ONE 7:e33636. https://doi.org/10.1371/journal.pone.0033636
Poncet BN, Herrmann D, Gugerli F, Taberlet P, Holderegger R, Gielly L, Rioux D, Thuiller W, Aubert S, Manel S (2010) Tracking genes of ecological relevance using a genome scan in two independent regional population samples of *Arabis alpina*. Mol Ecol 19:2896–2907. https://doi.org/10.1111/j.1365-294X.2010.04696.x

Pritchard JK, Di Rienzo A (2010) Adaptation – not by sweeps alone. Nat Rev Genet 11:665–667. https://doi.org/10.1038/nrg2880

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

Prunier J, Laroche J, Beaulieu J, Bousquet J (2011) Scanning the genome for gene SNPs related to climate adaptation and estimating selection at the molecular level in boreal black spruce. Mol Ecol 20:1702–1716. https://doi.org/10.1111/j.1365-294X.2011.05045.x

Prunier J, Pelgas B, Gagnon F, Desponts M, Isabel N, Beaulieu J, Bousquet J (2013) The genomic architecture and association genetics of adaptive characters using a candidate SNP approach in boreal black spruce. BMC Genomics 14:368. https://doi.org/10.1186/1471-2164-14-368

R Core Team (2016) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria

Rajendra KC, Seifert S, Prinz K, Gailing O, Finkeldey R (2014) Subtle human impacts on neutral genetic diversity and spatial patterns of genetic variation in European beech (*Fagus sylvatica*). For Ecol Manag 319:138–149. https://doi.org/10.1016/j.foreco.2014.02.003

Rajora OP, Eckert AJ, Zinck JWR (2016) Single-locus versus multilocus patterns of local adaptation to climate in eastern white pine (*Pinus strobus,* Pinaceae). PLoS ONE 11(7): e0158691. https://doi.org/10.1371/journal.pone.0158691

Raymond M, Rousset F (1995) GENEPOP (Version 1.2): Population Genetics Software for Exact Tests and Ecumenicism. J Hered 86:248–249.

Rellstab C, Gugerli F, Eckert AJ, Hancock AM, Holderegger R (2015) A practical guide to environmental association analysis in landscape genomics. Mol Ecol 24:4348–4370. https://doi.org/10.1111/mec.13322

Rousset F (2008) GENEPOP’007: a complete re-implementation of the GENEPOP software for Windows and Linux. Mol Ecol Resour 8:103–106. https://doi.org/10.1111/j.1471-8286.2007.01931.x

Russello MA, Kirk SL, Frazer KK, Askey PJ (2012) Detection of outlier loci and their utility for fisheries management. Evol Appl 5:39–52. https://doi.org/10.1111/j.1752-4571.2011.00206.x

Sander T, König S, Rothe GM, Janssen A, Weisgerber H (2000) Genetic variation of

European beech (*Fagus sylvatica* L.) along an altitudinal transect at Mount Vogelsberg in Hesse, Germany. Mol Ecol 9:1349–1361.

Schoville SD, Bonin A, François O, Lobreaux S, Melodelima C, Manel S (2012) Adaptive Genetic Variation on the Landscape: Methods and Cases. Annu Rev Ecol Evol Syst 43:23–43. https://doi.org/10.1146/annurev-ecolsys-110411-160248

Seifert S (2012) Variation of candidate genes related to climate change in European beech (*Fagus sylvatica* L.). Dissertation, Georg-August-University Göttingen

Seifert S, Vornam B, Finkeldey R (2012) A set of 17 single nucleotide polymorphism (SNP) markers for European beech (*Fagus sylvatica* L.). Conserv Genet Resour 4:1045–1047. https://doi.org/10.1007/s12686-012-9703-9

Sork VL, Davis FW, Westfall R, Flint A, Ikegami M, Wang H, Grivet D (2010) Gene movement and genetic association with regional climate gradients in California valley oak (*Quercus lobata* Née) in the face of climate change. Mol Ecol 19:3806–3823. https://doi.org/10.1111/j.1365-294X.2010.04726.x

Stephan W (2016) Signatures of positive selection: from selective sweeps at individual loci to subtle allele frequency changes in polygenic adaptation. Mol Ecol 25:79–88. https://doi.org/10.1111/mec.13288

Stucki S, Orozco-terWengel P, Forester BR, Duruz S, Colli L, Masembe C, Negrini R, Landguth E, Jones MR, The NEXTGEN Consortium, Bruford MW, Taberlet P, Joost S (2016) High performance computation of landscape genomic models including local indicators of spatial association. Mol Ecol Resour 17(5):1072–1089. https://doi.org/10.1111/1755-0998.12629

Thornthwaite CW (1948) An Approach toward a Rational Classification of Climate. Geogr Rev 38:55–94. https://doi.org/10.2307/210739

Trenberth K (2011) Changes in precipitation with climate change. Clim Res 47:123–138.

Tsumura Y, Uchiyama K, Moriguchi Y, Kimura MK, Ueno S, Ujino-Ihara T (2014) Genetic Differentiation and Evolutionary Adaptation in *Cryptomeria japonica*. G3 GenesGenomesGenetics 4:2389–2402. https://doi.org/10.1534/g3.114.013896

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) micro-checker: software for identifying and correcting genotyping errors in microsatellite data. Mol Ecol Notes 4:535–538. https://doi.org/10.1111/j.1471-8286.2004.00684.x

Varshney RK, Graner A, Sorrells ME (2005) Genic microsatellite markers in plants: features and applications. Trends Biotechnol 23:48–55. https://doi.org/10.1016/j.tibtech.2004.11.005

Vitti JJ, Grossman SR, Sabeti PC (2013) Detecting Natural Selection in Genomic Data. Annu Rev Genet 47:97–120. https://doi.org/10.1146/annurev-genet-111212-133526

Vornam B, Decarli N, Gailing O (2004) Spatial Distribution of Genetic Variation in a Natural Beech Stand (F*agus sylvatica* L.) Based on Microsatellite Markers. Conserv Genet 5:561–570. https://doi.org/10.1023/B:COGE.0000041025.82917.ac

Weber P, Pluess, A, Mühlethaler, U (2011) Genetic resources of beech in Switzerland. In: Frydl J, Novotny P, Fennessy J, von Wühlisch G (eds) COST Action E 52 Genetic resources of beech in Europe - current state. Johann Heinrich von Thünen-Institut, Germany, pp 248–255

Whitlock MC (2005) Combining probability from independent tests: the weighted Z-method is superior to Fisher's approach. J Evol Biol 18:1368–1373. https://doi.org/10.1111/j.1420-9101.2005.00917.x

Figure legend

Fig. 1 Structure analysis based on the 13 SSR markers (**a** and **b**) and the 70 SNPs (**c** and **d**) for $K = 2$. Bar plot indicates the assignment probability of each individual to two different clusters (K) in saplings (**a** and **c**) and adults (**b** and **d**). Population name abbreviations: Fel - Felsberg; Chu - Chur; Mal - Malans; Mas - Mastrils; Sar - Sargans; Mel - Mels; Ard - Ardon; Cha - Chamoson; Sax - Saxon; Mar - Martigny; Col - Collombey; Oll – Ollon

Fig. 2 Examples of some SNP allele frequencies calculated for each population and plotted against of environmental variables *AP*, I_m and *MaxAt* and environmental PC1 that were identified as being very likely under selection by EAA**.** Black and open circles denote Rhine and Rhone populations, respectively

Fig. 3 Examples of logistic regression fit of SNP allele frequencies along environmental variables *AP*, I_m and *MaxAt* and environmental PC1 for four SNPs identified as being very likely under selection by EAA ($P < 0.1$)

Supplementary material

Supplementary material 1 - Tables S1-S7

Supplementary material 2 - Figs S1-S12