



# UNIVERSITÀ DI PARMA

## ARCHIVIO DELLA RICERCA

University of Parma Research Repository

Within-Population Genetic Structure in Beech (*Fagus sylvatica* L.) Stands Characterized by Different Disturbance Histories: Does Forest Management Simplify

This is the peer reviewed version of the following article:

*Original*

Within-Population Genetic Structure in Beech (*Fagus sylvatica* L.) Stands Characterized by Different Disturbance Histories: Does Forest Management Simplify Population Substructure? / Piotti, Andrea; Leonardi, Stefano; Myriam, Heuertz; Joukje, Buiteveld; Thomas, Geburek; Sophie, Gerber; Koen, Kramer; Cristina, Vettori; Giovanni Giuseppe, Vendramin. - In: PLOS ONE. - ISSN 1932-6203. - 8:9(2013), pp. e73391-e73391. [10.1371/journal.pone.0073391]

*Availability:*

This version is available at: 11381/2652127 since: 2018-10-11T15:02:05Z

*Publisher:*

*Published*

DOI:10.1371/journal.pone.0073391

*Terms of use:*

Anyone can freely access the full text of works made available as "Open Access". Works made available

*Publisher copyright*

note finali coverpage

(Article begins on next page)

# Within-Population Genetic Structure in Beech (*Fagus sylvatica* L.) Stands Characterized by Different Disturbance Histories: Does Forest Management Simplify Population Substructure?

Andrea Piotti<sup>1,2</sup>, Stefano Leonardi<sup>2</sup>, Myriam Heuertz<sup>3</sup>, Joukje Buiteveld<sup>4</sup>, Thomas Geburek<sup>5</sup>, Sophie Gerber<sup>6,7</sup>, Koen Kramer<sup>4</sup>, Cristina Vettori<sup>1</sup>, Giovanni Giuseppe Vendramin<sup>1\*</sup>

**1** Plant Genetics Institute, National Research Council, Sesto Fiorentino, Italy, **2** Dipartimento di Bioscienze, Università di Parma, Parma, Italy, **3** Department of Forest Ecology and Genetics, CIFOR-INIA, Madrid Spain, **4** Alterra, Wageningen, The Netherlands, **5** Federal Research Centre for Forests, Vienna, Austria, **6** UMR 1202 BIOGECO, INRA, Cestas, France, **7** UMR 1202 BIOGECO, University of Bordeaux, Talence, France

## Abstract

The fine-scale assessment of both spatially and non-spatially distributed genetic variation is crucial to preserve forest genetic resources through appropriate forest management. Cryptic within-population genetic structure may be more common than previously thought in forest tree populations, which has strong implications for the potential of forests to adapt to environmental change. The present study was aimed at comparing within-population genetic structure in European beech (*Fagus sylvatica* L.) plots experiencing different disturbance levels. Five plot pairs made up by disturbed and undisturbed plots having the same biogeographic history were sampled throughout Europe. Overall, 1298 individuals were analyzed using four highly polymorphic nuclear microsatellite markers (SSRs). Bayesian clustering within plots identified 3 to 11 genetic clusters (within-plot  $\theta_{ST}$  ranged from 0.025 to 0.124). The proportion of within-population genetic variation due to genetic substructuring ( $F_{C_{Plot}} = 0.067$ ) was higher than the differentiation among the 10 plots ( $F_{PlotTot} = 0.045$ ). Focusing on the comparison between managed and unmanaged plots, disturbance mostly explains differences in the complexity of within-population genetic structure, determining a reduction of the number of genetic clusters present in a standardized area. Our results show that: *i*) genetic substructuring needs to be investigated when studying the within-population genetic structure in forest tree populations, and *ii*) indices describing subtle characteristics of the within-population genetic structure are good candidates for providing early signals of the consequences of forest management, and of disturbance events in general.

**Citation:** Piotti A, Leonardi S, Heuertz M, Buiteveld J, Geburek T, et al. (2013) Within-Population Genetic Structure in Beech (*Fagus sylvatica* L.) Stands Characterized by Different Disturbance Histories: Does Forest Management Simplify Population Substructure? PLoS ONE 8(9): e73391. doi:10.1371/journal.pone.0073391

**Editor:** Pär K. Ingvarsson, University of Umeå, Sweden

**Received:** October 26, 2012; **Accepted:** July 17, 2013; **Published:** September 5, 2013

**Copyright:** © 2013 Piotti et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This study was financially supported by the European Commission through the Dynabech project (5th Framework Programme, QLRT-1999-01210) and finalised based on the financial support by the European Commission through the FP7-project FORGER: Towards the Sustainable Management of Forest Genetic Resources in Europe (KBBE - 289119). AP was supported by the Italian MIUR project "Approccio multitaxa allo studio delle risposte della biodiversità italiana al cambiamento climatico" (RBAP10A2T4). KK was in addition supported with co-funding by the Dutch Ministry of Economics, Agriculture and Innovation in the project (KB-14-002-033) of the Knowledge-Base Green Space and Biodiversity. This support is gratefully acknowledged. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: giovanni.vendramin@igv.cnr.it

## Introduction

Within-population genetic structure is shaped by the complex interplay of genetic and demographic factors. In forest tree populations subject to anthropogenic influence, management can alter both processes influencing the amount and distribution of genetic variation [1–4]. Within-population genetic structure follows spatial as well as temporal dynamics whose comprehension is one of the bases for understanding how populations evolve [5]. In addition, it has recently been pointed out that the study of within-population genetic structure is a key prerequisite to correctly interpret the results of association genetic studies, since cryptic genetic structure can yield spurious statistical associations between genotypic and phenotypic traits [6,7]. The probability of such misinterpretations is particularly high when studying forest

tree populations because they have long been seen as large, random-mating units with minimal structure [8].

The partitioning of within-population genetic diversity into well-distinguished genetic clusters indicates the existence of subpopulations (i.e. 'population stratification', [9,10]). However, genetic clusters are not necessarily spatially clumped, since individuals belonging to the same genetic cluster may be spatially randomly distributed, and clusters may be highly intermingled. In forest trees, the focus of most studies has mainly been directed towards the spatial component of within-population genetic structure, while its non-spatial component has only rarely been assessed [5]. Slavov et al. [7] surprisingly found a strong non-spatial genetic substructure in a small continuous *Populus trichocarpa* stand characterized by an extensive gene flow via pollen. The stand

**Table 1.** Characteristics of investigated beech plots.

Country	Plot	Code <sup>a</sup>	Disturbance history	N	Plot size (ha)	Density (trees/ha)	Latitude/ Longitude	Altitude (m)	Stand age
Germany	Flecken-Zechlin 1	Gl	Semi-natural	120	0.86	140	53°11′/12°43′	85	75–140
	Flecken-Zechlin 2	Gh	Shelterwood	120	0.91	132	53°11′/12°44′	85	46–155
The Netherlands	Pijpebrandje	NLI	Semi-natural	120	0.68	178	52°15′/5°43′	50	130–200
	Solse Bosje	Nlh	Plantation	120	0.91	132	52°14′/5°39′	50	130
Austria <sup>b</sup>	Dobra1	Al	Natural	110	1.87	58	48°35′/15°23′	390–550	250–300
	Dobra2	Ah	Shelterwood	110	0.58	191	48°35′/15°23′	550–580	–
France	St. Baume	Fl	Natural	286	1.91	150	43°19′/5°45′	750	–
	Mt. Ventoux	Fh	Colonisation	90	1.32	68	44°10′/5°16′	1450	–
Italy <sup>c</sup>	Abruzzo A	Il	Natural	112	0.56	196	42°30′/13°29′	1270	–
	Abruzzo C	Ih	Coppice before 1850, then shelterwood	110	0.21	537	42°30′/13°29′	1155	70

<sup>a</sup>Plot codes were formed by the indication of country (G = Germany, NL = The Netherlands, A = Austria, F = France, I = Italy) and intensity of disturbance (l = low, h = high).

<sup>b</sup>Austrian plots are subplots of Piotti et al. [26] plots.

<sup>c</sup>Italian plots studied by Paffetti et al. [25] are subplots of the ones analyzed here.

doi:10.1371/journal.pone.0073391.t001

was composed of individuals that were unambiguously assigned to two distinct subpopulations that probably originated from different seedling cohorts. Jansson & Ingvarsson [8] underlined how such cryptic structures may be more common than previously thought in forest tree populations. Besides temporally separated founder events during colonization, other temporal factors, such as phenological and recruitment dynamics, can determine non-spatial clustering of genetic variation. In addition, re-forestation using mixed seed lots from distinct provenances and other silvicultural practices might also result in non-spatial clustering of genotypes.

Natural and anthropogenic disturbances can alter the distribution of genetic variation within populations. This variation is essential for the potential of forests to adapt to environmental change, such as climate change [11–13]. The comprehension of the consequences of disturbances, such as fires and silvicultural practices, on within-population genetic structure can therefore be crucial to preserve forest genetic resources through appropriate management [4,14]. A few studies investigated the effects of common forest management practices on the within-population genetic structure. In some cases, a negative impact of management practices on genetic diversity and spatial genetic structure (SGS hereafter) was shown [3,15–18], although no effect or a weak effect on gene flow, genetic diversity and SGS was found elsewhere (e.g. [17,19,20]). The effects of some regeneration practices on genetic diversity was demonstrated to be species-specific [17] and, interestingly, even population-specific [15,19].

The lack of experiments specifically designed to compare the within-population genetic structure between disturbed and undisturbed forest tree stands has retarded our understanding of the fine-scale consequences of forest management [21,22]. The present study was aimed at comparing stands experiencing different disturbance levels using five plot pairs of *Fagus sylvatica* (L.). In each pair, plots were chosen as geographically close as possible to avoid the effect of potential confounding factors. An exhaustive spatial sampling allowed us to perform an in-depth analysis of within-population genetic structure using different approaches (spatial autocorrelation analysis and Bayesian clustering) that have proved effective for revealing cryptic genetic structuring. Therefore, the specific aims of the study were: *i*) to investigate the existence and, in case, the characteristics of within-

population genetic structure in the ten plots analyzed, and *ii*) to check whether disturbance had an impact on within-population genetic structure in each plot pair.

## Materials and Methods

### Study Sites, Sample Collection and Microsatellite Analysis

Study sites and sampling strategy have been described in details in previous works [23–26]. Briefly, five study sites were chosen in five European countries: Austria, France, Germany, Italy and The Netherlands. Each study site consists of 2 plots located in close proximity of each other, with the exception of the French ones which are also characterized by a marked difference in altitude (Table 1, Table S1). In the other 4 sites, environmental conditions, climate and genetic background are assumed to be similar between plots. Within each site, the 2 plots differ in the intensity of natural or anthropogenic disturbance, mainly due to forest management practices. Differences in stand structure were generally marked except between the Dutch plots (Table S1). The undisturbed Dutch plot has been subject to low-intensity human management during the last three centuries (cattle grazing, fire wood and construction wood for local use) with a gradual cessation of human activities. However, the managed Dutch plot is a plantation, probably established in 1870 with local forest reproductive material. Therefore, the recent management histories of the two plots are markedly different (Table 1).

All 1298 sampled individuals were genotyped at four highly polymorphic unlinked SSRs (FS1–15, FS4–46, FS1–25 and mfc5, [24]). The four SSRs displayed a total of 25, 41, 26, 31 alleles, respectively, and provided exclusion probabilities as high as ~0.985 when used for paternity analysis in French plots [26]. Non negligible frequencies of null alleles have been reported previously for some marker-site combinations [26,27].

### Data Analysis

General estimates of genetic diversity as well as genetic differentiation and phylogeographic analyses are reported in Buiteveld et al. [24]. We recalculated inbreeding coefficients ( $F_{IS}$ ) taking into account the possible presence of null alleles by using the program INEst, running the individual inbreeding model with a Gibbs sampler of  $10^5$  iterations [28].

Fine-scale within-population genetic structure was investigated following two approaches: classical spatial autocorrelation analysis and Bayesian clustering.

Spatial autocorrelation analysis was performed using the multivariate method by Smouse & Peakall [29] implemented in the program GenAlEx 6.5 [30]. It provides a multi-locus estimate of pair-wise relatedness between individuals ( $r$ ) which minimizes the stochasticity found in single locus or single allele estimates of relatedness. Tests for statistical significance were conducted by  $i$ ) random shuffling (1000 times) of individual geographic locations to define the upper and lower bounds of the 95% confidence interval for each distance class, and  $ii$ ) estimating 95% confidence intervals around mean  $r$  values by bootstrapping pair-wise comparisons within each distance class (1000 repeats). Analyses were performed using both the even distance classes option (using 20 m wide distance classes) and the even sample size option (distributing all possible pairs in seven distance classes with similar numbers of pairs per class). The nonparametric heterogeneity test proposed by Smouse et al. [21] was used to compare correlograms between plots within each pair, setting the number of bootstrap resamplings to 9999.

The intensity of SGS was also measured by the  $S_p$  statistic [31].  $S_p$  is computed as  $S_p = b_F / (F_1 - 1)$ , where  $b_F$  is the regression slope of the kinship estimator  $F_{ij}$  computed among all pairs of individuals against their geographical distances, and  $F_1$  is the average kinship coefficient between individuals of the first distance class (0–20 m).  $S_p$  has the desirable characteristic of being comparable among stands in a single study and among studies. The statistical significance of  $F_1$  and  $b_F$  was tested based on 1000 permutations of individual locations among individuals. All analyses were performed using SPAGeDi 1.3 [32].

To assess the power of the marker set to detect SGS in the studied stands we performed spatially explicit simulations of a population life cycle (pollen and seed dispersal, reproduction, death and replacement) for 64 generations following the simulation approach described in Heuertz et al. [33] and De-Lucas et al. [34]. At generation null, individuals were given random genotypes at microsatellite loci according to the observed allele frequencies in stand Fh (the plot with the strongest SGS). Spatial genetic structure was then allowed to build up over generations according to four scenarios characterized by different combinations of dispersal parameters (the axial standard deviation of gene dispersal,  $\sigma_g$ ) and density, spanning from an unrealistically high SGS for beech ( $F_1 > 0.09$ ) to very low SGS ( $F_1 < 0.01$ ). Intermediate scenarios were chosen to mimic SGS in the 2 French plots, according to pollen flow parameters and effective number of pollen donors estimated in the same plots [26] and seed dispersal parameters from Oddou-Muratorio et al. [35] and Millerón et al. [36]. For the low-SGS scenario, a parameter combination that produced an almost absent SGS was selected. On the other hand, for the high-SGS scenario, dispersal and density values significantly higher than the highest found among the study plots were chosen. In the simulation procedure the generation overlap was 50%. Sixty-four generations were modelled, taking measures of SGS at generations 1, 2, 4, 8, 16, 32, 48, and 64 and using a sampling scheme similar to the Fh stand. Simulations were repeated 100 times with 4 and 20 loci (5 independent replicates of the original 4 loci).

The simulation results were used to assess possible errors introduced by the low number of loci by:  $i$ ) Comparing the no-structure confidence intervals observed in Fh and Fl with the distribution of simulated  $F_1$  values from intermediate scenarios. The rationale is that the four microsatellites have sufficient power to detect SGS if SGS estimates from realistic simulations fall

consistently outside the no-SGS 95% confidence interval.  $ii$ ) Comparing  $F_1$  values observed in Fh and Fl with those obtained from the simulated extreme scenarios. The overlap of  $F_1$  values from the observed *vs.* simulated scenarios will inform on the power of the four loci to discriminate between realistic and unrealistic SGS scenarios.  $iii$ ) Comparing simulations with four and 20 loci. This comparison will help us assess the gain in accuracy and precision of SGS estimates when increasing the number of SSR loci.

The spatially explicit Bayesian clustering algorithm implemented in the R package GENELAND v. 4.0.2 was used to assign individuals to putative sub-populations within each plot [37]. We used the spatial model with correlated allele frequencies, setting the maximum number of sub-populations to 20, and running the analyses for  $10^6$  iterations, with a thinning value of 1000. The possible presence of null alleles was explicitly taken into account by using the `filter.null.alleles` option. Each analysis was repeated 10 times. The highest median number of clusters of the 10 repetitions was chosen as the most representative one for each plot (see also Figure S1). To test whether clusters identified by Bayesian clustering were differentiated, Weir & Cockerham [38]  $\theta_{ST}$  among clusters within each plot was calculated with Fstat [39]. Statistical significance of  $\theta_{ST}$  values was tested after 1000 randomization of genotypes among clusters. To account for the possible presence of null alleles, differentiation among clusters was also calculated with FreeNA applying the ENA correction method to correct efficiently for the positive bias induced by the presence of null alleles on  $\theta_{ST}$  estimation [40].

Once reliable Bayesian clustering results were obtained, a hierarchical analysis of the genetic structure of all plots was performed to assess the importance of population substructure relative to structure at higher hierarchical scales. The hierarchical estimates of  $F$ -statistics were obtained by the R package *HIERFSTAT* [41] using the method by Weir & Cockerham [38] based on the estimation of variance components of gene frequencies. Two intermediate hierarchical levels were chosen: plot (Plot) and clusters (Clu). Therefore, in the following  $F_{PlotTotal}$  refers to the correlation of genes within plots relative to the total,  $F_{CluPlot}$  to the correlation of genes within clusters relative to the plots, and  $F_{IndClu}$  to the correlation of genes within individuals relative to the clusters.  $F$ -statistics are related to each other by the following expression  $(1 - F_{IndTotal}) = (1 - F_{IndClu}) \times (1 - F_{CluPlot}) \times (1 - F_{PlotTotal})$ . Statistical significance of  $F$ -statistics was tested using 1000 randomization of the units defined by the level just below that of interest in the hierarchy.

Plots with larger areas are expected to contain a higher number of clusters just by chance. To compare the number of clusters produced by GENELAND among sites with different areas, we calculated  $\overline{Nc}$ , that is the mean number of clusters ( $Nc$ ) in a standardized area of 0.21 ha (equal to the area of Ih, the smallest of our plots) in each plot. Since there is very large number of ways to subset an area, we performed the standardization procedure as follows: in each plot, the  $Nc$  was recorded within a window of 0.21 ha drawn around each individual and the number of clusters was then averaged for all windows around all individuals. To avoid border effects,  $Nc$  was calculated only for individuals more than 23 m from the border (*<i.e.* half-side of a 0.21 ha window).  $\overline{Nc}$  and its confidence interval was then calculated for each plot as a measure of complexity of the genetic structure.

## Results

Spatial autocorrelation analyses based on different methods to estimate pair-wise genetic relatedness were concordant in detect-

ing a significant spatial genetic structure in the first distance class (0–20 m) in all plots with the exception of Al and Ih, where no evidence of SGS was observed (Figure 1). Correlograms were similar using the ‘even sample size’ or the ‘even distance class’ option (results not shown). Fh is the only plot where SGS extends up to 40 m. Simulations confirmed the robustness of results from spatial correlation analysis with respect to the power of the marker set used (Figure 2) and the presence of null alleles (Text S1). In particular, Figure 2 shows that: *i*) the  $F_1$  values computed from the spatially explicit simulations for scenarios with intermediate strength of SGS fell consistently outside the no-SGS 95% confidence intervals obtained by random shuffling of individual geographic locations, *ii*) the confidence intervals of observed  $F_1$  values in Fh and Fl showed a limited overlap with  $F_1$  values from simulated extreme scenarios, with an error rate not higher than 19%, and *iii*) increasing the number of loci from 4 to 20 moderately increased the precision but had no effect on accuracy: mean values did not significantly differ between 4 and 20 loci scenarios as assessed by t-tests with Welch modification for unequal variances between groups (HIGH-SGS scenario:  $t=0.38$ ,  $df=168.7$ ,  $P=0.70$ ; Fh-like SGS scenario:  $t=0.25$ ,  $df=167.0$ ,  $P=0.80$ ; Fl-like SGS scenario:  $t=0.75$ ,  $df=163.9$ ,  $P=0.46$ ; LOW-SGS scenario:  $t=-0.20$ ,  $df=171.3$ ,  $P=0.84$ ). Therefore, we inferred that the 4 microsatellite loci have high power to detect SGS in a dispersal context that is realistic for the beech stands studied in this work.

The highest  $F_1$  values were recorded for Fh, Gl and Fl, the last two being the only plots with a statistically significant excess of homozygotes (Table 2). The  $S_p$  statistic ranged from 0.0040 in Ih to 0.0293 in Fh and, in general, it reflected results from spatial autocorrelation analyses obtained from GenAlEx:  $S_p$  was high in plots where genetic variability is spatially structured (Gl, Fh, and Il) whereas the lowest  $S_p$  values were recorded in plots characterized by an absence of SGS (Al and Ih). In pair-wise comparisons, correlograms of the Fh-Fl and Ih-Il pairs were statistically different. This was mainly due to a large difference in the first distance class, with  $r_{1,Fh} > r_{1,Fl}$  and  $r_{1,Il} > r_{1,Ih}$  (Table S2, Figure 1).

The highest median number of clusters into which individuals are grouped by Bayesian clustering (averaged over 10 independent GENELAND runs) ranged from 3 (Ih and NLh) to 11 (Fl) (Table 2, Figure S1). Within-plot  $\theta_{ST}$  among GENELAND clusters ranged from 0.025 (Gh) to 0.124 in the recently colonized plot (Fh), with several plots showing high within-plot differentiation (Tab. 2). The  $\theta_{ST}$  values estimated applying the ENA correction method using FreeNA were not different from the ones without correction (Paired  $t$ -test,  $t=0.44$ ,  $df=9$ ,  $P=0.67$ ), indicating a negligible influence of the possible presence of null alleles on differentiation estimates. Differentiation among clusters within plots ( $F_{CltPlot} = 0.067$ ,  $P<0.001$ ) was actually larger than the differentiation among plots, which is determined by both post-glacial recolonization history at the biogeographical scale and disturbance history at the local scale ( $F_{PlotTot} = 0.045$ ,  $P<0.001$ ).  $F_{IndClt}$  was 0.186, including the effect of both inbreeding ( $F_{IS}>0$ ) within clusters and null alleles.

Standardizing the number of clusters over an area of 0.21 ha (the area of the smallest plot, Ih), we found that undisturbed plots are generally characterized by a more complex structure (Figure 3). The mean  $\overline{Nc}$  was 9.20 and 6.16 in Dutch and Italian undisturbed plots (NLi and Il), respectively, about three and two times the mean  $\overline{Nc}$  observed in nearby disturbed plots (3.00). Mean  $\overline{Nc}$  was higher in the disturbed compared to the undisturbed plot only in Austria (6.50 and 4.66, respectively).

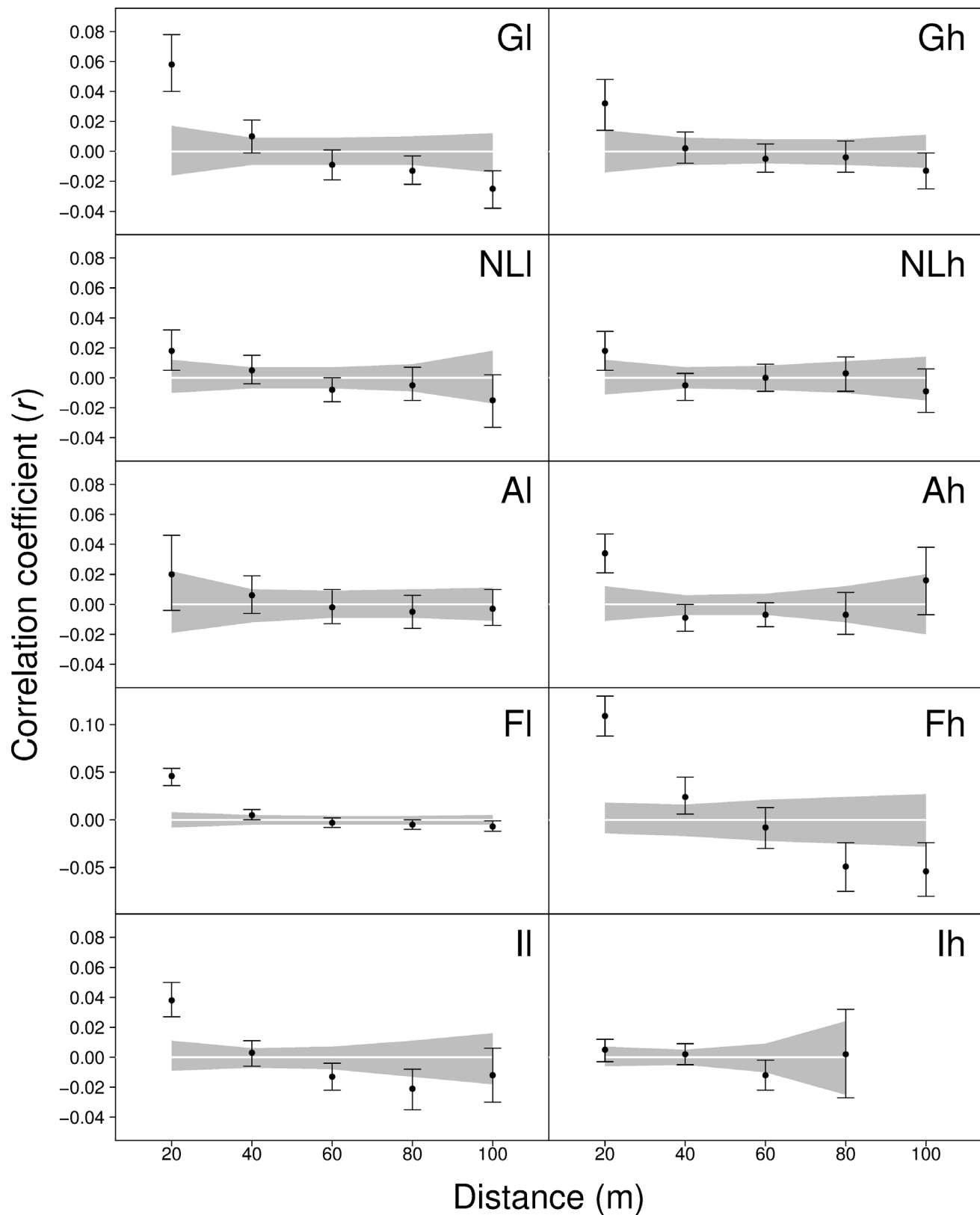
## Discussion

Our experimental setup allowed us to obtain a reliable set of parameters describing the within-population genetic structure of European beech in a wide variety of ecological and management conditions, spanning from undisturbed to highly managed stands. Two relevant results for the comprehension of within-population dynamics in *F. sylvatica* were achieved: *i*) despite the variety of conditions explored, we found that a large proportion of within-population genetic variation is due to genetic substructuring in all plots, and *ii*) in plot pairs sharing the same bio-geographic history, disturbance explains most of the difference in the complexity of within-population genetic structure.

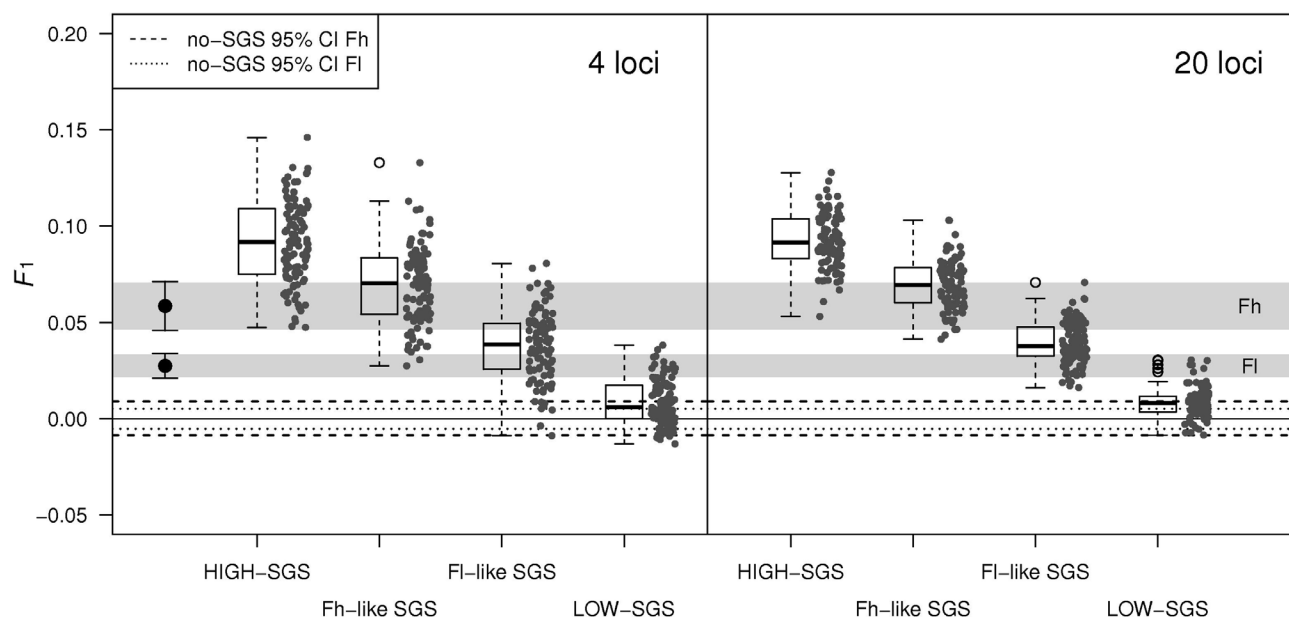
### Within-population Genetic Structure in Beech

Our work confirms what emerged in previous studies, that SGS in beech extends up to 40 m and not further [35,42–44]. An SGS up to larger distances was found only with AFLPs [45]. This is likely to be related to the peculiar characteristics of these markers, namely the much higher number of loci that can be scored. In fact, Jump et al. [44] concluded that SGS estimates should only be compared within marker types and for similar marker numbers. According to the theory of isolation by distance, the intensity of SGS is negatively related to effective density ( $d_e$ ) and dispersal distances [31]. European beech has a high potential for pollen dispersal, with large (~75%) pollen immigration within small to medium size (1–8 ha) plots and mean within-population pollen dispersal distances ranging from 40 to 180 m in natural stands [5,26,35]. Seed dispersal is much more limited. In general, seed immigration from outside the plot is <30% and mean seed dispersal distance is approximately 10 m [5,35]. Chybicki et al. [43] estimated a ratio of seed to pollen dispersal distances ranging from 1:10 to 1:100 for beech. Our results showed how the interplay between limited seed dispersal and large pollen flow is likely to have determined the spatially restricted SGS (up to 20 m) detected in all plots at regular tree density ( $100 < d < 200$  trees  $ha^{-1}$ ). The lack of a spatial signal in the genetic structure of Ih could have been a result of high local density, which should result in a considerable overlap of individual gene shadows. Conversely, the SGS up to 40 m in Fh, an area characterized by high pollen immigration (~80%, [26]), can be explained by the low density in this colonization area. This is in agreement with what was previously found in beech colonization areas [35]. High SGS in Fh could also be an effect of local pollen dispersal distances being extremely low [26], further reducing the overlap between individual gene shadows. In Al, a low density plot from a pristine part of the Dobra forest, we curiously found an absence of SGS. This result exemplifies that an observed SGS pattern can result from a series of determinants that affect spatial patterns [1,2], and how their interaction can determine unexpected results. For instance, Vekemans & Hardy [31] noted in several studies a positive, and thus counterintuitive, effect of density on the intensity of SGS. On the other hand, dispersal can be strongly enhanced when density is low and, as a consequence, its effect can prevail over the density effect *per se* in shaping SGS.

Our results together with previous studies [35,43–44] allowed us to obtain a reliable distribution of  $S_p$  values ( $n=19$ ), ranging from 0.0054 in a low intensity managed stand [44] to 0.0354 in an expanding population [35]. About 60% of studied stands are characterized by  $S_p$  values between 0.005 and 0.015, confirming the general conclusion by Chybicki et al. [43] that  $S_p$  in *F. sylvatica* is around 0.01, which is typical for outcrossing species whose pollen is wind-dispersed. Still, according to Vekemans & Hardy [31], values of  $S_p$  between ~0.01 and ~0.035 characterize species



**Figure 1. Correlograms from spatial autocorrelation analysis using the correlation coefficient  $r$  by Smouse & Peakall [29] and even distance classes.** Shaded areas represent the 95% confidence interval obtained through random shuffling (1000 times) of individual geographic locations, black lines around mean  $r$  values represent 95% confidence intervals around mean  $r$  values generated by bootstrapping (1000 times) pairwise comparisons within each distance class.  
doi:10.1371/journal.pone.0073391.g001



**Figure 2. Assessment of the power of the marker set to detect SGS by spatially explicit simulations.** For illustration of the results, the distribution of the kinship coefficient  $F_1$  between neighbours at generation 64 was used as the focal statistic (grey dots and boxplots) and compared to *i)* the no-structure 95% confidence intervals of  $F_1$  from the Fh and FI populations (dotted lines, see legend in the left panel) obtained by random shuffling of individual geographic locations, and *ii)* real  $F_1$  values from Fh and FI (black dots in the left panel) and their confidence intervals (grey areas). Results from simulations with 4 and 20 loci (right and left panels, respectively) are reported. Parameter settings for the 4 simulated scenarios were  $\sigma_g = 12$  m and  $D = 20$  trees/ha (HIGH-SGS),  $\sigma_g = 12$  m and  $D = 35$  trees/ha (FI-like SGS),  $\sigma_g = 29$  m and  $D = 50$  trees/ha (Fh-like SGS),  $\sigma_g = 72$  m and  $D = 145$  trees/ha (LOW-SGS). doi:10.1371/journal.pone.0073391.g002

whose seeds are dispersed by wind and gravity, respectively. The distribution of  $S_p$  values suggests that seed dispersal might not be consistently limited in beech. In fact, although within-population seed dispersal distances are usually limited to a few dozens of meters, long distance dispersal can be strongly enhanced by secondary dispersal by birds, reaching up to several kilometres in beech [46,47]. In an expanding oak (*Quercus robur*) population, for instance, seedlings growing under their mothers were responsible for high kinship at short distance whereas dispersed seeds accounted for most of the SGS pattern at distances beyond tens of meters [48]. In addition, although  $F_1$  is likely to be mainly determined by seed dispersal, low  $S_p$  values can also, in part, be determined by the capability of long distance pollen dispersal in beech.

A relevant result emerging from our survey is that local substructure explained a larger proportion of genetic variation ( $F_{\text{CltPlot}} = 0.067$  and  $0.025 < \theta_{ST} < 0.124$ , Table 2) than differentiation among all ten plots ( $F_{\text{PlotTot}} = 0.045$ ). Notably, a relevant local substructure is present when the spatial signal is weak or even absent, making spatial autocorrelation analysis insufficient to characterize all aspects of population stratification at the local scale. Contrarily to what was previously thought, the presence of a complex within-population structure is emerging as a common characteristic in forest trees, even though until now few studies have addressed this topic and contrasting results have been found [7,25,49]. For instance, a single genetic cluster was found when analyzing a large low-density *P. trichocarpa* population whereas, in the same study, a small population surprisingly showed a clear substructure [7].

The presence of neutral genetic structuring at any geographical level can have profound implications for experimental design in studies aimed at searching phenotype-genotype associations and molecular signatures of selection [6,50]. In fact, even a subtle

neutral genetic differentiation observed at the population spatial scale needs to be statistically accounted for in association studies to avoid spurious associations [51]. Once stratification has been demonstrated, refined methods such as mixed models are now available to take into account even the hierarchically lowest stratification (i.e. family structure) in association studies [6].

### Effect of Disturbance on the within-population Genetic Structure

Our results indicate that disturbance histories had a site-specific influence on within-population genetic structure at a finer level than what is generally detected by methods commonly employed for investigating genetic consequences of forest management. In the literature, several studies aimed at comparing genetic diversity, SGS and gene flow between managed and unmanaged stands, or among stands managed by different techniques, are present (e.g. [3,16,20,24]). Some contrasting results have been found but, in general, effects of silvicultural practices on the adult cohort are weak or absent (e.g. [24]). In the managed *vs.* unmanaged comparisons Gh-Gl, NLh-NLl, and Ih-Il, we detected a substantial reduction in the complexity of the genetic structure as measured by the mean number of clusters in a standardized area ( $\overline{Nc}$ ) and by within-population  $\theta_{ST}$  in disturbed plots. The reduction was low between the German plots, reflecting the recent divergence in their management history (Gl is a strict reserve only since 1961, [23]). The difference in the complexity of genetic structure was large in the comparison of the even-aged Dutch plantation (NLh) and the Italian formerly coppiced stand (Ih) with their respective unmanaged stands, indicating a possible high impact of such management regimes on the heterogeneity of within-population genetic structure in beech. In plantations, such a reduction can be generated by the use of a limited number of seed sources or

**Table 2.** Parameters describing within-population genetic structure in the studied beech plots.

Site	SGS parameter			$N_c$	$\theta_{ST}$	$F_{IS}$
	$F_1$	$b_F (\pm SE)$	$Sp (\pm SE)$			
GI	0.0371***	-0.0258 $\pm$ 0.0049***	0.0268 $\pm$ 0.0043	9	0.087***	0.170 $\pm$ 0.055***
Gh	0.0187***	-0.0115 $\pm$ 0.0048***	0.0117 $\pm$ 0.0076	7	0.025***	0.025 $\pm$ 0.025
NLI	0.0115***	-0.0101 $\pm$ 0.0021***	0.0102 $\pm$ 0.0041	10	0.105***	0.068 $\pm$ 0.055
NLh	0.0104**	-0.0095 $\pm$ 0.0031***	0.0096 $\pm$ 0.0052	3	0.062***	0.063 $\pm$ 0.042
AI	0.0111	-0.0068 $\pm$ 0.0015**	0.0069 $\pm$ 0.0059	8	0.043***	0.051 $\pm$ 0.040
Ah	0.0183***	-0.0131 $\pm$ 0.0032***	0.0133 $\pm$ 0.0034	8	0.042***	0.023 $\pm$ 0.022
FI	0.0274***	-0.0096 $\pm$ 0.0017***	0.0099 $\pm$ 0.0033	11	0.050***	0.083 $\pm$ 0.038*
Fh	0.0585***	-0.0276 $\pm$ 0.0025***	0.0293 $\pm$ 0.0065	9	0.124***	0.033 $\pm$ 0.028
II	0.0224***	-0.0186 $\pm$ 0.0039***	0.0190 $\pm$ 0.0049	8	0.049***	0.042 $\pm$ 0.036
Ih	0.0015	-0.0040 $\pm$ 0.0035*	0.0040 $\pm$ 0.0024	3	0.043***	0.023 $\pm$ 0.022

$F_1$ , average kinship coefficient between individuals of the first distance class (0–20 m);  $b_F$ , regression slope of the kinship estimator  $F_{ij}$  computed among all pairs of individuals against geographical distances;  $Sp$ , intensity of SGS;  $N_c$ , mean number of clusters from GENELAND analyses;  $\theta_{ST}$ , differentiation among clusters within each plot;  $F_{IS}$ , inbreeding coefficient estimated by INEst.

\* $P < 0.05$ ,

\*\* $P < 0.01$ ,

\*\*\* $P < 0.001$ .

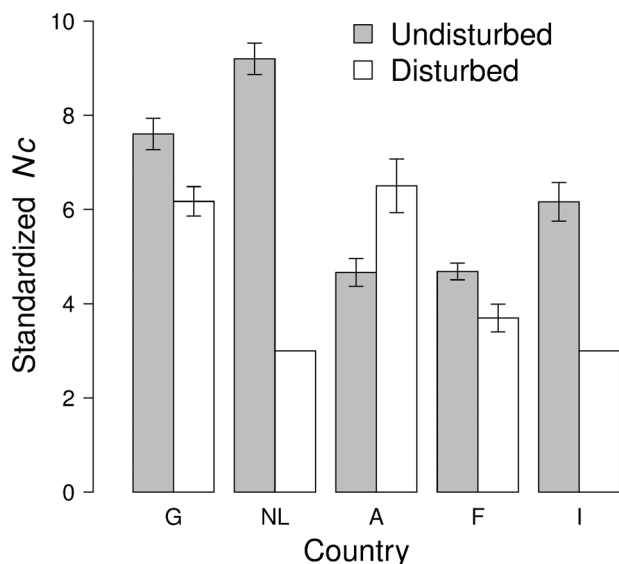
doi:10.1371/journal.pone.0073391.t002

genetically homogeneous material after a clear-cut [52,53], which can also explain the weak SGS found by spatial autocorrelation analysis in NLh. In the Italian site, Ih has been subject to a complex recent management history: it was coppiced until 1850, then a conversion to high forest began through thinning from below and regeneration felling following the uniform shelterwood system [23,25]. Such high anthropogenic pressure has produced the loss of the spatial component and a simplification of the non-spatial component of genetic structure in this stand, as also reported analyzing a subset of our Italian datasets with different methods and markers [25]. It is noteworthy, however, that this difference in the genetic structures does not correspond to a

comparable reduction of within-population genetic diversity [24]. Paffetti et al. [25] reached similar results and the only difference in genetic diversity detected between managed and unmanaged stands was a slight reduction in the frequency of rare alleles in the managed stand.

The shelterwood system produced weak to negligible genetic effects in the stand pairs analyzed (AI-Ah, GI-Gh), with no difference in genetic diversity [24] and only a slight reduction in the intensity of genetic structure in the German plot pair. Given the wide application of shelterwood-based techniques, the genetic consequences of this management system has been previously investigated and contrasting results have been found. Gene flow patterns were not affected by this silvicultural technique in *Pinus sylvestris* populations [20], whereas it decreased the SGS comparing three managed *P. strobus* stands with an old growth stand [3]. El-Kassaby et al. [17] found different genetic consequences of shelterwood in *Abies amabilis* and *Tsuga heterophylla*, two species with contrasting mating system dynamics. In two separate studies on *Pseudotsuga menziesii*, only one showed a reduction in genetic diversity in managed stands [15,19]. This range of results shows that the response of forest trees to management techniques can be highly species-specific, and that several replicates of properly sampled plot pairs (managed *vs.* unmanaged) are needed to draw conclusions about the effects of natural and anthropogenic disturbances on the characteristics of within-population genetic structure.

Indices of genetic diversity as well as spatial structure and gene flow parameters are well suited for monitoring in long-term modelling studies, but detectable changes have rarely been found when comparing managed and unmanaged sites in beech [24,26,42]. Although not appropriate to describe changes in the evolutionary potential of populations, our results show that indices describing subtle changes in the within-population genetic structure (such as the mean number of clusters in a standardized area presented here) can be good candidates for providing early signals of the consequences of forest management on neutral genetic variation. A special effort should be put into properly characterizing within-population genetic structure in follow-up



**Figure 3. Complexity of within-population genetic structure as measured by the standardized number of clusters in an area of 0.21 ha.** 0.21 ha is the area of Ih, the smallest of sampling plots.  
doi:10.1371/journal.pone.0073391.g003



experimental and modelling studies for a deeper comprehension of short- and long-term responses of long-lived forest tree species to climatic as well as anthropogenic pressures.

## Supporting Information

**Figure S1 Bayesian clustering results.** Distribution of results from the 10 repetitions of GENELAND analyses for each plot. The highest median number of clusters of the 10 repetitions was chosen as the most representative one for each plot. (TIF)

**Table S1 Stand characteristics of the 10 study plots.** For each social class the mean value of tree diameter at breast height and height, the stem number, the basal area and total volume are reported. (PDF)

**Table S2 Heterogeneity tests of SGS between disturbed and undisturbed plots using GenAIEx (Peakall & Smouse [30]).** The reported values represent the degree of differentiation

of SGS between two plots from the same site in each distance class ( $r^2$ ) and total  $\omega$ . (PDF)

**Text S1 Assessment of the potential impact of null alleles on spatial correlation analyses in the studied *Fagus sylvatica* stands.** (PDF)

## Acknowledgments

We thank two anonymous reviewers and the Editor for useful comments on an earlier version of this manuscript.

## Author Contributions

Conceived and designed the experiments: KK TG JB SL CV GGV. Performed the experiments: JB TG CV GGV. Analyzed the data: AP SL MH GGV. Contributed reagents/materials/analysis tools: AP SL TG JB CV GGV. Wrote the paper: AP SL GGV. Contributed to analysis and interpretation of data: KK SG JB MH. Critically reviewed the manuscript: KK SG JB TG MH.

## References

- Loveless M, Hamrick JL (1984) Ecological determinants of genetic structure in plant populations. *Annu Rev Ecol Syst* 15: 65–95.
- Troupin D, Nathan R, Vendramin GG (2006) Analysis of spatial genetic structure in an expanding *Pinus halepensis* population reveals development of fine-scale genetic clustering over time. *Mol Ecol* 15: 3617–3630.
- Marquardt PE, Echt CS, Epperson BK, Pubanz DM (2007) Genetic structure, diversity, and inbreeding of eastern white pine under different management conditions. *Can J For Res* 37: 2652–2662.
- Sagnard F, Oddou-Muratorio S, Pichot C, Vendramin GG, Fady B (2011) Effects of seed dispersal, adult tree and seedling density on the spatial genetic structure of regeneration at fine temporal and spatial scales. *Tree Genet Genomes* 7: 37–48.
- Oddou-Muratorio S, Klein EK, Vendramin GG, Fady B (2011) Spatial vs. temporal effects on demographic and genetic structures: the roles of dispersal, mating and differential mortality on patterns of recruitment in *Fagus sylvatica*. *Mol Ecol* 20: 1997–2010.
- Price AL, Zaitlen NA, Reich D, Patterson N (2010) New approaches to population stratification in genome-wide association studies. *Nat Rev Genet* 11: 459–463.
- Slavov GT, Leonardi S, Adams WT, Strauss SH, DiFazio SP (2010) Population substructure in continuous and fragmented stands of *Populus trichocarpa*. *Heredity* 105: 348–57.
- Jansson S, Ingvarsson PK (2010) Cohort-structured tree populations. *Heredity* 105: 331–332.
- Waples RS (2010) Spatial-temporal stratifications in natural populations and how they affect understanding and estimation of effective population size. *Mol Ecol Res* 10: 785–796.
- Sebro R, Risch NJ (2012) A brief note on the resemblance between relatives in the presence of population stratification. *Heredity* 108: 563–568.
- Kramer K, Buiteveld J, Forstreuter M, Geburek T, Leonardi S, et al. (2008) Bridging the gap between ecophysiological and genetic knowledge to assess the adaptive potential of European beech. *Ecol Model* 216: 333–353.
- Jump AS, Marchant R, Peñuelas J (2009) Environmental change and the option value of genetic diversity. *Trends Plant Sci* 14: 51–58.
- Kramer K, Degen B, Buschbom J, Hickler T, Thuiller W, et al. (2010) Modelling exploration of the future of European beech (*Fagus sylvatica* L.) under climate change—Range, abundance, genetic diversity and adaptive response. *For Ecol Manage* 259: 2213–2222.
- Bradshaw RHW (2004) Past anthropogenic influence on European forests and some possible consequences. *For Ecol Manage* 197: 203–212.
- Adams WT, Zuo J, Shimizu JY, Tappeiner JC (1998) Impact of alternative regeneration methods on genetic diversity in coastal Douglas-fir. *For Sci* 44: 390–396.
- Epperson BK, Chung MG (2001) Spatial genetic structure of allozyme polymorphisms within populations of *Pinus strobus* (Pinaceae). *Am J Bot* 88: 1006–1010.
- El-Kassaby YA, Dunsworth BG, Krakowski J (2003) Genetic evaluation of alternative silvicultural systems in coastal montane forests: western hemlock and amabilis fir. *Theor Appl Genet* 107: 598–610.
- Marquardt PE, Epperson BK (2004) Spatial and population genetic structure of microsatellites in white pine. *Mol Ecol* 13: 3305–3315.
- Neale DB (1985) Genetic implications of shelterwood regeneration of Douglas-fir in Southwest Oregon. *For Sci* 15: 995–1005.
- Robledo-Arnuncio JJ, Smouse PE, Gil L, Alia R (2004). Pollen movement under alternative silvicultural practices in native populations of Scots pine (*Pinus sylvestris* L.) in central Spain. *For Ecol Manage* 197: 245–255.
- Smouse PE, Peakall R, Gonzales E (2008) A heterogeneity test for fine-scale genetic structure. *Mol Ecol* 17: 3389–3400.
- García-Gil MR, Olivier F, Kamruzzahan S, Waldmann P (2009) Joint analysis of spatial genetic structure and inbreeding in a managed population of Scots pine. *Heredity* 103: 90–96.
- Forstreuter M (2004) Characterization of the study plots. In: Kramer K, editor. Effects of silvicultural regimes on dynamics of genetic and ecological diversity of European beech forests. Impact assessment and recommendations for sustainable forestry. Final Report of 5th framework project DynaBeech, QLK5-CT-1999–1210. Wageningen: Alterra. 9–33.
- Buiteveld J, Vendramin GG, Leonardi S, Kramer K, Geburek T (2007) Genetic diversity and differentiation in European beech (*Fagus sylvatica* L.) stands varying in management history. *For Ecol Manage* 247: 98–106.
- Paffetti D, Travaglini D, Buonamici A, Nocentini S, Vendramin GG, et al. (2012) The influence of forest management on beech (*Fagus sylvatica* L.) stand structure and genetic diversity. *For Ecol Manage* 284: 34–44.
- Piotti A, Leonardi S, Buiteveld J, Geburek T, Gerber S, et al. (2012) Comparison of pollen gene flow among four European beech (*Fagus sylvatica* L.) populations characterized by different management regimes. *Heredity* 108: 322–331.
- Oddou-Muratorio S, Vendramin GG, Buiteveld J, Fady B (2009) Population estimators or progeny tests: what is the best method to assess null allele frequencies at SSR loci? *Conserv Genet* 10: 1343–1347.
- Chybicki JJ, Burczyk J (2009) Simultaneous estimation of null alleles and inbreeding coefficients. *J Hered* 100: 106–133.
- Smouse PE, Peakall R (1999) Spatial autocorrelation analysis of individual multiallele and multilocus genetic structure. *Heredity* 82: 561–573.
- Peakall R, Smouse PE (2012) GenAIEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics* 28: 2537–2539.
- Vekemans X, Hardy OJ (2004) New insights from fine-scale spatial genetic structure analyses in plant populations. *Mol Ecol* 13: 921–935.
- Hardy OJ, Vekemans X (2002) SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Mol Ecol Notes* 2: 618–620.
- Heuertz M, Vekemans X, Hausman JF, Palada M, Hardy OJ (2003) Estimating seed vs. pollen dispersal from spatial genetic structure in the common ash. *Mol Ecol* 12: 2483–2495.
- De-Lucas AI, Gonzalez-Martinez SC, Vendramin GG, Hidalgo E, Heuertz M (2009) Spatial genetic structure in continuous and fragmented populations of *Pinus pinaster* Aiton. *Mol Ecol* 18: 4564–4576.
- Oddou-Muratorio S, Bontemps A, Klein EK, Chybicki I, Vendramin GG, et al. (2010) Comparison of direct and indirect genetic methods for estimating seed and pollen dispersal in *Fagus sylvatica* and *Fagus cremata*. *For Ecol Manage* 259: 2151–2159.
- Millerón M, Lopez de Heredia U, Lorenzo Z, Alonso J, Dounavi A, et al. (2013) Assessment of spatial discordance of primary and effective seed dispersal of European beech (*Fagus sylvatica* L.) by ecological and genetic methods. *Mol Ecol* 22: 1531–1545.
- Guillot G, Mortier F, Estoup A (2005) GENELAND: a computer package for landscape genetics. *Mol Ecol Notes* 5: 712–715.
- Weir BS, Cockerham CC (1984) Estimating *F*-statistics for the analysis of population structure. *Evolution* 38: 1358–1370.

39. Goudet J (1995) FSTAT (Version 1.2): a computer program to calculate F-statistics. *J Hered* 86: 485–486.
40. Chapuis M-P, Estoup A (2007) Microsatellite null alleles and estimation of population differentiation. *Mol Biol Evol* 24: 621–631.
41. Goudet J (2005) HIERFSTAT, a package for R to compute and test hierarchical F-statistics. *Mol Ecol Res* 5: 184–186.
42. Leonardi S, Menozzi P (1996) Spatial structure of genetic variability in natural stands of *Fagus sylvatica* L. (beech) in Italy. *Heredity* 77: 359–368.
43. Chybicki JJ, 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.
44. Jump AS, Rico L, Coll M, Peñuelas J (2012) Wide variation in spatial genetic structure between natural populations of the European beech (*Fagus sylvatica*) and its implications for SGS comparability. *Heredity* 108: 633–639.
45. Jump AS, Penuelas 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.
46. Nilsson SG (1985) Ecological and evolutionary interactions between reproduction of beech *Fagus sylvatica* and seed eating animals. *Oikos* 44: 157–164.
47. Kunstler G, Thuiller W, Curt T, Bouchaud M, Jouvie R, et al. (2007) *Fagus sylvatica* L. recruitment across a fragmented Mediterranean landscape, importance of long distance effective dispersal, abiotic conditions and biotic interactions. *Divers Distrib* 13: 799–807.
48. Hampe A, El Masri L, Petit RJ (2010) Origin of spatial genetic structure in an expanding oak population. *Mol Ecol* 19: 459–471.
49. Chybicki JJ, Oleksa A, Burczyk J (2011) Increased inbreeding and strong kinship structure in *Taxus baccata* estimated from both AFLP and SSR data. *Heredity* 107: 589–600.
50. Neale DB, Ingvarsson PK (2008) Population, quantitative and comparative genomics of adaptation in forest trees. *Curr Opin Plant Biol* 11: 149–155.
51. Slavov GT, DiFazio SP, Martin J, Schackwitz W, Muchero W, et al. (2012) Genome resequencing reveals multiscale geographic structure and extensive linkage disequilibrium in the forest tree *Populus trichocarpa*. *New Phyt* 196: 713–725.
52. Finkeldey R, Ziehe M (2004) Genetic implications of silvicultural regimes. *Forest Ecol Manage* 197: 231–244.
53. Carnus J-M, Parrotta JA, Brockerhoff E, Arbez M, Jactel H, et al. (2006) Planted forests and biodiversity. *J For* 104: 65–77.